

Model-Agnostic Meta-Learning for Fault Diagnosis of Industrial Robots

Yuxin Liu^a, Chong Chen^{a*}, Tao Wang^a, Lianglun Cheng^a and Jian Qin^b

^aGuangdong Provincial Key Laboratory of Cyber-Physical System, Guangdong University of Technology, Guangzhou, China

^bWelding Engineering and Laser Processing Centre, School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, UK
chenc2021@gdut.edu.cn

Abstract—The success of deep learning in the field of fault diagnosis depends on a large number of training data, but it is a challenge to achieve fault diagnosis of multi-axis industrial robots in the case of few-shot. To address this issue, this paper proposes a method called Model-Agnostic Meta-Learning (MAML) for fault diagnosis of industrial robots. Its goal is to train an effective industrial robot fault classifier using minimal training data. Additionally, it can learn to recognize faults in new scenarios with high accuracy based on the training data. Experimental results based on a six-axis industrial robot dataset show that the proposed method is superior to traditional convolutional neural network (CNN) and transfer learning, and that the diagnostic results with the same amount of data in few-shot cases are better than existing intelligent fault diagnosis methods.

Keywords—Fault Diagnosis; Deep Learning; Industrial Robots; Meta Learning

I. INTRODUCTION

The rapidly developing manufacturing industry urgently needs high-end automated production equipment to meet the production requirements of large-scale, high-precision and high-complexity products. As the most representative equipment in the field of intelligent manufacturing, industrial robots have expanded from the traditional field of electronics and electrical engineering to various intelligent manufacturing systems and can complete various tasks through programmable mechanical arms [1]. However, as the operating time of industrial robots increases, the probability of failures also increases. Unexpected shutdowns caused by failures for different reasons may interrupt the entire production line, and the time to operate safely or economically has become an important basis for measuring enterprise production efficiency [2-4]. Therefore, researching on industrial robot fault diagnosis and health prediction technology has become a hot topic in recent years.

Early mechanism analysis and manual feature extraction are the main basis of traditional methods, that is, after extracting enough features, machine learning is used for classification. In the field of fault diagnosis, machine learning-based methods have been shown to be effective [5]. Later, fault diagnosis methods are roughly divided into: mathematical model-based methods, signal processing-based methods and knowledge-based methods. The mathematical model-based fault diagnosis method is to establish an accurate model of the object, analyze the residual between the model output and the measurement to

achieve fault diagnosis, which is suitable for systems with relatively accurate mathematical models. However, if the model is inaccurate or the measurement information contains noise, the diagnostic accuracy will be affected [6]. The signal processing-based method utilizes certain feature quantities of the system output signal or state. When phenomena such as singular data and frequency changes occur in the measurement signal, it is considered that the system has a fault. Although it avoids mathematical modeling, when the signal fault features are not obvious, this method will have missed diagnosis and misdiagnosis issues [7-9]. Knowledge-based fault diagnosis is a new intelligent fault diagnosis method based on knowledge or learning, which has self-learning, self-organization and self-reasoning capabilities, and improves fault diagnosis accuracy by using domain expert knowledge. Currently, deep learning has been widely applied in the field of fault diagnosis due to its powerful learning and feature extraction capabilities [10], but the convergence of deep learning models requires a large amount of prior fault data for training. In actual industrial applications, due to the low frequency and long cycle of fault occurrences, the collected fault data is often insufficient. In addition, heavy background noise and the variability of working conditions also reduce the performance of existing fault diagnosis methods. In recent years, researchers have made significant achievements in accurate fault diagnosis in few-shot situations [11-13]. This type of research can be divided into two categories: the first category is to generate enough samples through data augmentation, and the second category uses data from other working conditions for transfer learning.

To address the issue of insufficient samples for data augmentation, many scholars have adopted Generative Adversarial Network (GAN) to solve this problem. Su et al. [14] proposed an optimization-based GAN method for few-shot fault diagnosis. The method firstly uses gradient boosting to optimize the iterative strategy of the generator, by building weak learners to optimize deep learning efficiency. Then, the k -Nearest Neighbor algorithm is used to establish the decision boundary for faulty samples, and the Mahalanobis distance is used to measure the continuity of probability space distribution. Finally, fault classification and diagnosis are achieved by scoring and fusing the fault data with a dual-stream convolutional network. Ding et al. [15] believed that a single GAN model cannot achieve good diagnostic results. To achieve adaptive feature extraction and high diagnostic accuracy

under few-shot conditions, they proposed an intelligent fault diagnosis method based on GANs for rotating machinery. Each GAN is trained through adversarial training. In this process, the discriminator is pre-trained in an unsupervised manner to determine whether the input sample comes from the learned distribution. Then, a supervised training set selected from the original few-shot training set is used to semi-supervised enhance the discriminator to improve its decision-making ability. Finally, the enhanced discriminators obtained from all samples are integrated and fed into the integrated model, and the discriminator with the highest output corresponds to the predicted health condition. This method differs from the idea of using data augmentation or resampling methods to increase the number of training samples. It only requires a small amount of samples to map the relationship between corresponding measurement signals and different health conditions of equipment, and eliminates the interaction between inherent distributions of different fault modes during the training process. Due to the problem of gradient vanishing that GAN networks are prone to encounter in practical working conditions, many researchers had considered transferring knowledge from other working condition data. Shao et al. [16] proposed a dual threshold attention GAN for rotation machinery fault diagnosis. Zhang et al. [17] developed an iterative matching network with an enhanced selective sample reuse strategy. Pseudo-labeling was performed on unlabeled signals under invisible conditions, and the matching network was trained using these labeled samples to reduce the difference between the source and target domains. In addition, a similarity measurement based on cosine distance was designed to filter out samples with erroneous pseudo-labels. Xie et al. [18] proposed a multi-task attention-guided network that can train a multi-task model with limited data. The model designed a task-sharing network to capture global features. Subsequently, the extracted features were sent to two sub-networks to perform fault type recognition and fault level recognition tasks. By leveraging the advantages of multi-labeling, the requirement for labeled data was reduced. However, all of the aforementioned transfer learning methods require a large number of data from the source domain to achieve good results.

In the field of fault diagnosis, data augmentation and transfer learning have achieved certain results. However, GAN networks are prone to gradient vanishing problems under practical working conditions, and transfer learning requires a large amount of data from the source domain to achieve good results. Moreover, current transfer learning methods mostly rely on deep structures [19], which may not work well when the target domain is complex. Meta-learning demonstrates unique advantages in solving the few-shot problem [20], aiming to train a model that can classify new classes with limited labeled examples. Some scholars have implemented meta-learning based fault diagnosis in other fields [21-23], but few scholars have studied the advantages of meta-learning in the industrial robot field. To address the aforementioned issues, this paper proposes using MAML to solve the problem of difficult convergence of network models under few-shot scenarios. The aim is to achieve effective fault diagnosis during the meta-training phase with limited data, and to

perform well in unknown fault scenarios. This model trains the initial parameters of the model by learning to the learning process itself. After updating the parameters through one or more gradient steps, the model has strong generalization ability for new tasks. The remainder of this paper is organised as follows: Section 2 presents the methodology, Section 3 reports the experimental study, and Section 4 concludes.

II. METHODOLOGY

A. Meta-Learning model

Meta-learning is different from traditional machine learning methods, aiming to improve the learning algorithm itself based on experiences from multiple learning events. It helps to address the fundamental issues of data, computation bottlenecks, and generalization in deep learning. The typical MAML algorithm is model-agnostic, where the neural network architecture and loss function are unknown. Its core is to optimize the parameters for fast-adaptation gradient descent in two loops (inner and outer loops). All these features provide MAML with much flexibility, making it suitable for both reinforcement learning problems that maximize the expected cumulative reward function, and supervised learning problems that minimize a loss function (such as cross-entropy and mean squared error). An example of fault diagnosis of a six-axis industrial robot is shown in Figure 1 below. The specific steps for fault diagnosis are as follows:

(1) During the inner loop stage, any network model can be adapted to extract task features. In this stage, the updated parameter vector θ^* is first calculated for each class i using the training data D_i . Then, the updated model parameters θ^* are used to calculate the loss term for the validation data D_i sampled from the same class. The purpose of this stage is to optimize the parameter set θ of the neural network along the descending direction of the aggregation gradient of the fault scenario data of different industrial robots. The initialization parameter θ is updated after passing through the inner loop structure:

$$\theta^* = \theta - \alpha \nabla_{\theta} L_i(f_{\theta}, D_i) \quad (1)$$

Where θ^* represents the optimal parameters obtained after updating on the task; D_i represents the i th task in the task set; α is the learning rate of the inner loop network; f_{θ} is the inner loop network structure with parameters θ ; L_i is the loss on the task; and $\nabla_{\theta} L_i(f_{\theta}, D_i)$ is the loss gradient.

(2) During the outer loop stage, the initial parameters of the model are used as input, and the fast training process of multiple tasks is used to update the initial parameters of the model:

$$\min_{\theta} \sum_i^n L_i(f_{\theta^*}, D_i) = \sum_i^n L_i \cdot f(\theta - \alpha \nabla_{\theta} L_i(f_{\theta}, D_i)) \quad (2)$$

Where n represents the total number of meta-learning tasks.

(3) Finally, after aggregating each task, MAML updates the loss and performs meta-gradient updates on the original model parameters θ :

$$\theta \leftarrow \theta - \beta \nabla_{\theta} \sum_i^n L_i(f_{\theta^*}, D_i) \quad (3)$$

Where β represents meta-learning rate.

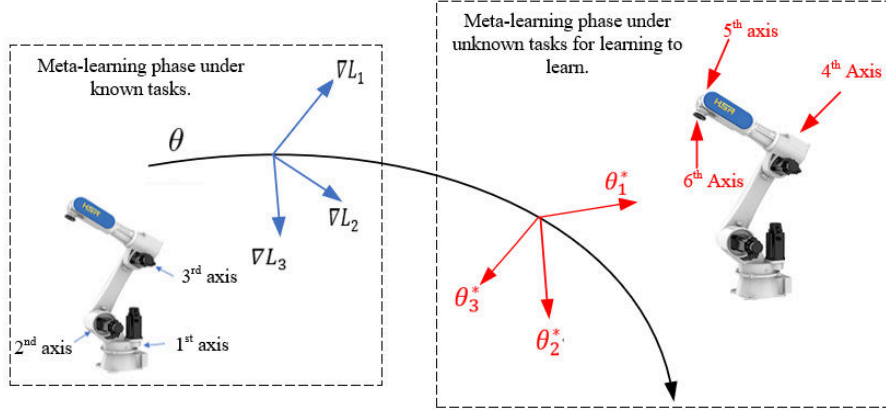


Figure 1 The application of MAML in fault diagnosis of a six-axis industrial robot.

B. Few-shot industrial robot fault diagnostic based on MAML

Few-shot classification is a field of supervised learning, which can be described as an N-way K-shot problem, where N is the unknown number of classes in the meta-testing process, and K refers to the number of training samples available for each class. The MAML algorithm is agnostic to the model, including its network architecture and loss function. The flowchart is shown in Figure 2.

The core of MAML is to optimize gradient parameters in two loops: the inner and outer loops. In the inner loop,

the parameters θ are updated based on each task dataset D_i from the meta-training phase, and the updated model parameters are used to sample and validate the same class dataset, calculating the loss. In the outer loop, MAML aggregates the updated losses of each task to calculate the model parameters, and performs meta-gradient updates.

Overall, MAML attempts to find neural network parameters that may not necessarily have the best performance on different classes of data provided in the meta-training phase, but can quickly adapt to new (unseen) tasks.

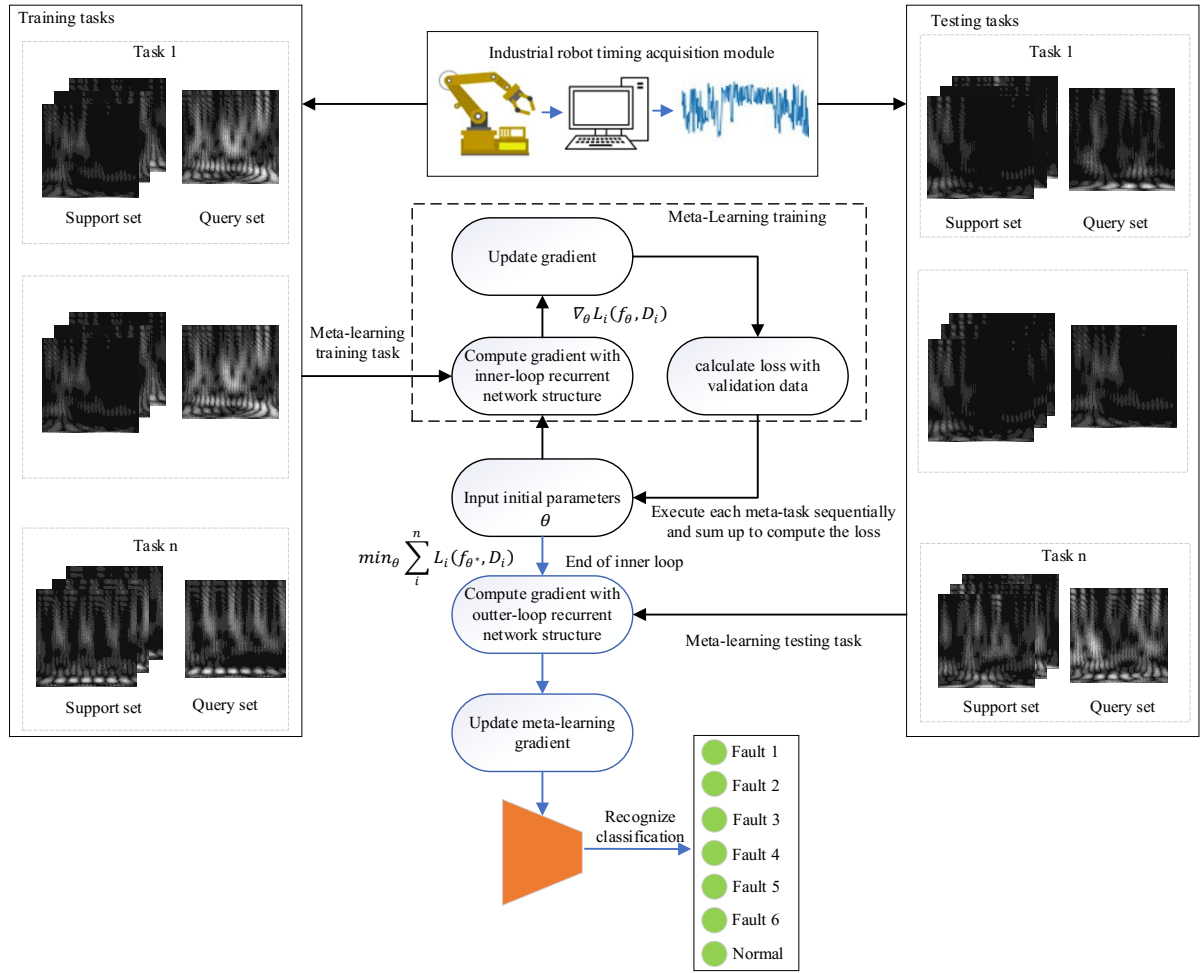


Figure 2 Fault diagnosis flow chart of industrial robot based on MAML

III. EXPERIMENTAL STUDY

A. Experimental setup

This experiment uses Google's Pytorch deep learning framework, a Linux-based operating system, and a PC with an Nvidia Geforce RTX 3060 GPU and 12GB RAM. The trained models were evaluated using a test set.

This experiment was conducted to validate the application using a certain brand's six-axis robot fault dataset. The dataset includes multiple sets of normal/abnormal robot driver feedback current data. The data samples were taken at 1-second intervals and contain a total of 7 different operating state data. Specifically, Monitoring data for a six-axis industrial robot was collected through fault injection experiments. In the experiment, a faulty part was replaced with a normal one and installed on an industrial robot, and then the robot was run to collect the feedback current signals. In addition, relevant compound-fault data was collected through

composite fault injection experiments, as shown in Table I. Figure 3 shows the amplitude of the first 100 data points.

TABLE I. EXPERIMENTAL DATASET

Data size	Faults description	Label
180000	Compound fault of the 1-axis reducer and 2-axis motor	Fault 1
180000	Compound fault of the 1-axis reducer and 3-axis reducer	Fault 2
180000	Compound fault of the 3-axis reducer and 4-axis reducer	Fault 3
180000	Fault of the 3-axis reducer	Fault 4
180000	Fault of the 2-axis motor	Fault 5
300000	Fault of the 4-axis reducer	Fault 6
330000	Normal	Normal

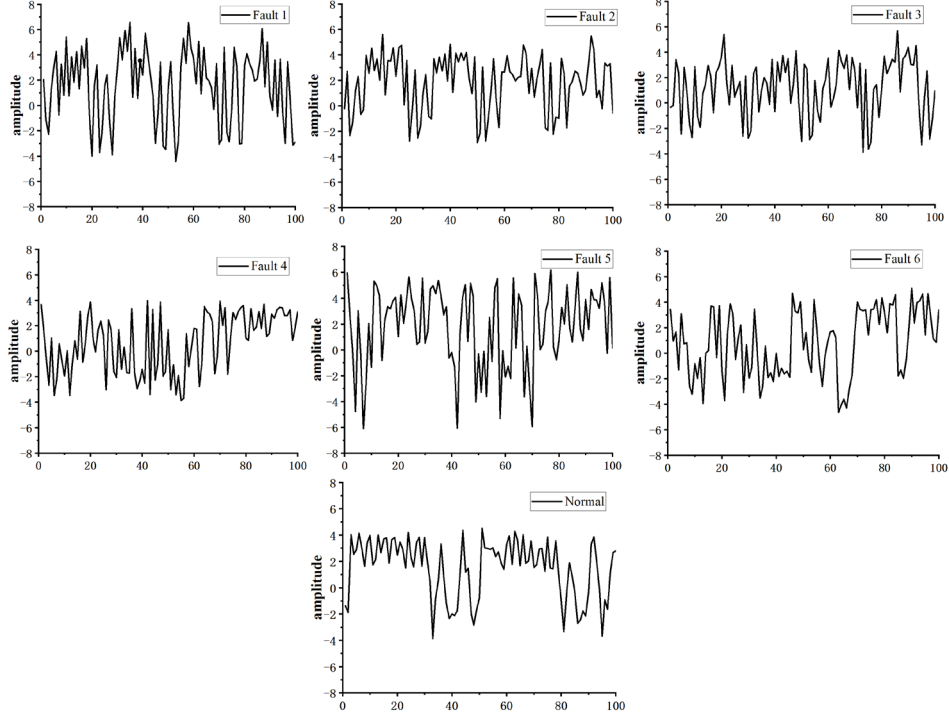


Figure 3 Amplitude plot of the first 100 data points for a single fault on the 2nd axis

During the experiment, the time-series data collected at different time points were processed by removing abnormal values and empty or erroneous values caused by certain human factors. The data from different dates under the same fault were merged, normalized, and labeled with corresponding fault tags. Then, the sample order was randomly shuffled separately to form the final dataset. Finally, for the N-way K-shot classification problem under meta-learning, the number of categories (N) and the number of samples (K) were dynamically specified for each meta-training and meta-testing.

The interior of MAML consists of four ConvBlocks, each of which consists of MaxPool, BatchNorm, ReLU activation function, and Conv. The cross-entropy function is used as the loss function, Adam algorithm is used as the optimization algorithm, and the classification results are output through a Linear layer. In addition, in the inner and outer loops of MAML, the inner loop learning rate is 0.005, the outer loop learning rate is 0.0005, the epoch iteration is set to 100, and the model is compiled.

B. Results and analysis

To compare the accuracy of the dataset on traditional algorithms, it was compared to multiple models on the same dataset, such as CNN, Dilation CNN, FCNN and LSTM. It is worth noting that the dataset used for model training consisted of the 180,000 data points mentioned earlier. Prior to model training, data processing was necessary: first, the faulty and normal data were labeled separately, merged, and shuffled. Second, data sampling was performed using a sliding window mechanism, with the sampling features being mean, variance, and kurtosis. Finally, the processed data was used as the final dataset for model training. The experimental results of different models are shown in Figure 4. From Figure 4, it can be seen that after data processing and input into the model, the

accuracy can reach over 80%. Among them, the accuracy of the improved dilated convolutional neural network based on CNN is higher, reaching 93%.

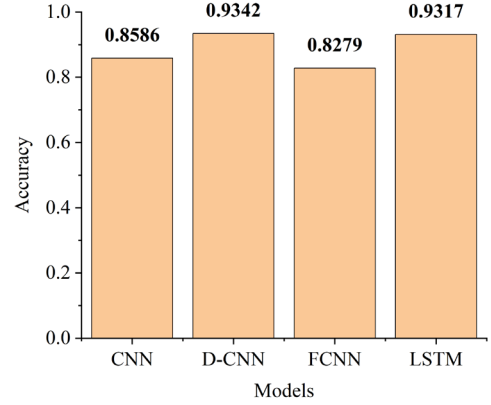


Figure 4 Comparison of accuracy of different models and different

TABLE II. EXPERIMENTAL RESULTS

N-way	K-shot	Meta Train Accuracy	Meta Valid Accuracy	Meta Test Accuracy
2-way	1-shot	1.0000	0.8750	0.6500
	5-shot	1.0000	0.9062	0.7050
	10-shot	1.0000	0.9688	0.7250
3-way	1-shot	0.5000	0.4167	0.4533
	5-shot	0.9958	0.6875	0.5733
	10-shot	1.0000	0.7667	0.6633
4-way	1-shot	0.5156	0.3542	0.4667
	5-shot	0.7719	0.5833	0.5600
	10-shot	0.9906	0.7563	0.5900
5-way	1-shot	0.6125	0.6875	0.6750
	5-shot	0.6500	0.8750	0.7500
	10-shot	0.9825	0.9969	0.7600

The experimental results based on MAML are shown in Table II. It should be noted that there are 7 categories in total, so there are different combinations of categories for different n . Here, the average value of the 5 combinations selected is taken as the final conclusion. Please refer to Table II for details (class N is for testing and class $7 - N$ is for training). As can be seen from Table II: (1) selecting appropriate training and testing sample numbers can optimize the model, such as 5-way 10-shot, which can achieve an accuracy rate of 98.25%, and the classification accuracy rate for unknown fault categories reaches 76%. It is already very good to perform so well with so few samples. (2) Under the same number of categories, the accuracy rate of the meta-training and meta-testing stages increases with the increase of each category; (3) The more fault categories there are in the meta-training stage, the stronger the model's generalization ability and the better the performance on the test set.

Through the above comparison, it can be found that: with a large amount of data, a fault diagnosis accuracy of 93% can be achieved through data processing and model parameter optimization. However, compared to MAML which can achieve 98% accuracy with very few samples, MAML demonstrates a significant advantage. Further experiments have shown that when K continues to increase to 20, the accuracy of MAML in industrial robot fault diagnosis can reach 100%, demonstrating the effectiveness of MAML in this field of fault diagnosis.

C. Discussions

Applying MAML to industrial robot fault diagnosis can effectively identify fault categories, and adjusting the n to K -shot quantity appropriately can significantly improve accuracy, approaching 100%. Therefore, MAML performs well in the application of few-shot industrial robots and has stronger generalization ability than traditional CNN. Although the Transformers algorithm has been applied in fault diagnosis, its large parameter size hinders its combination with other learners. Therefore, in future work, it is possible to further explore the combination of other excellent models for feature extraction during meta-training and meta-testing processes, making it easier to fine-tune and quickly adjust model parameters to adapt to new tasks.

IV. CONCLUSIONS

Traditional machine learning algorithms have strict requirements on the number of fault samples, which limits the accuracy of fault diagnosis. In this study, MAML was used for small-sample fault diagnosis and further used to identify new fault samples that did not appear in meta-training. The effectiveness of this method was verified using a six-axis industrial robot dataset. The results show that the model can achieve near 100% accuracy with fewer samples, which is higher than traditional CNN even with more data. In addition, the model has stronger generalization and scalability and can be further combined with other excellent models for quick parameter tuning in inner and outer loops, reducing the operating costs of industrial robots.

ACKNOWLEDGEMENT

Our work is supported by multiple funds in China, including the Key Program of NSFC-Guangdong Joint Funds (U2001201), Industrial core and key technology plan of ZhuHai City (ZH22044702190034HJL), the Natural Science Foundation of Guangdong Province (2020B1515120010), the Dedicated Fund for Promoting High-Quality Economic Development in Guangdong Province (Marine Economic Development Project) GDNRC [2021]44. Our work is also supported by Guangdong Provincial Key Laboratory of Cyber-Physical System (2020B1212060069).

REFERENCES

- [1] X. Zhou, H. Zeng, C. Chen, H. Xiao, and Z. Xiang, "An attention-enhanced multi-modal deep learning algorithm for robotic compound fault diagnosis," *Measurement Science and Technology*, vol. 34, no. 1, p. 014007, 2022.
- [2] R. S. Beebe and R. S. Beebe, *Predictive maintenance of pumps using condition monitoring*. Elsevier, 2004.
- [3] Y. Liu, C. Chen, T. Wang, and L. Cheng, "An attention enhanced dilated CNN approach for cross-axis industrial robotics fault diagnosis," *Autonomous Intelligent Systems*, vol. 2, no. 1, pp. 1-11, 2022.
- [4] K. Lu, C. Chen, T. Wang, L. Cheng, and J. Qin, "Fault diagnosis of industrial robot based on dual-module attention convolutional neural network," *Autonomous Intelligent Systems*, vol. 2, no. 1, pp. 1-12, 2022.
- [5] Z. Zhang, J. Wu, J. Ma, X. Wang, and C. Zhou, "Fault diagnosis for rolling bearing based on lifting wavelet and morphological fractal dimension," in *The 27th Chinese Control and Decision Conference (2015 CCDC)*, 2015: IEEE, pp. 6351-6354.
- [6] D. Stavrou, D. G. Eliades, C. G. Panayiotou, and M. M. Polycarpou, "Fault detection for service mobile robots using model-based method," *Autonomous Robots*, vol. 40, no. 2, pp. 383-394, 2016.
- [7] G. Haisong, X. Fang, J. Kai, and B. Hongye, "Design and research of a robot joint torque sensor," *Modern Manufacturing Engineering*, vol. 440, no. 5, p. 33, 2017.
- [8] H. YAN, R. BAI, and F. JI, "An Improved Dynamics Parameter Identification Method for SCARA Robots," *China Mechanical Engineering*, vol. 28, no. 22, p. 2707, 2017.
- [9] A. A. Jaber and R. Bicker, "Fault diagnosis of industrial robot gears based on discrete wavelet transform and artificial neural network," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 58, no. 4, pp. 179-186, 2016.
- [10] D. Zhu, X. Song, J. Yang, Y. Cong, and L. Wang, "A bearing fault diagnosis method based on L1 regularization transfer learning and LSTM deep learning," in *2021 IEEE International Conference on Information Communication and Software Engineering (ICICSE)*, 2021: IEEE, pp. 308-312.
- [11] L. Zou, Y. Li, and F. Xu, "An adversarial denoising convolutional neural network for fault diagnosis of rotating machinery under noisy environment and limited sample size case," *Neurocomputing*, vol. 407, pp. 105-120, 2020.
- [12] C. Wang and Z. Xu, "An intelligent fault diagnosis model based on deep neural network for few-shot fault diagnosis," *Neurocomputing*, vol. 456, pp. 550-562, 2021.
- [13] S. R. Saufi, Z. A. B. Ahmad, M. S. Leong, and M. H. Lim, "Gearbox fault diagnosis using a deep learning model with limited data sample," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6263-6271, 2020.
- [14] Y. Su, L. Meng, X. Kong, T. Xu, X. Lan, and Y. Li, "Small sample fault diagnosis method for wind turbine gearbox based on optimized generative adversarial networks," *Engineering Failure Analysis*, vol. 140, p. 106573, 2022.
- [15] Y. Ding, L. Ma, J. Ma, C. Wang, and C. Lu, "A generative adversarial network-based intelligent fault diagnosis method for rotating machinery under small sample size conditions," *IEEE Access*, vol. 7, pp. 149736-149749, 2019.
- [16] H. Shao, W. Li, B. Cai, J. Wan, Y. Xiao, and S. Yan, "Dual-Threshold Attention-Guided Gan and Limited Infrared Thermal

- Images for Rotating Machinery Fault Diagnosis Under Speed Fluctuation," *IEEE Transactions on Industrial Informatics*, pp. 1-10, 2023.
- [17] K. Zhang, J. Chen, T. Zhang, S. He, T. Pan, and Z. Zhou, "Intelligent fault diagnosis of mechanical equipment under varying working condition via iterative matching network augmented with selective Signal reuse strategy," *Journal of Manufacturing Systems*, vol. 57, pp. 400-415, 2020.
- [18] Z. Xie, J. Chen, Y. Feng, K. Zhang, and Z. Zhou, "End to end multi-task learning with attention for multi-objective fault diagnosis under small sample," *Journal of Manufacturing Systems*, vol. 62, pp. 301-316, 2022.
- [19] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345-1359, 2009.
- [20] Y. Feng, J. Chen, J. Xie, T. Zhang, H. Lv, and T. Pan, "Meta-learning as a promising approach for few-shot cross-domain fault diagnosis: Algorithms, applications, and prospects," *Knowledge-Based Systems*, vol. 235, p. 107646, 2022.
- [21] C. Li, S. Li, A. Zhang, Q. He, Z. Liao, and J. Hu, "Meta-learning for few-shot bearing fault diagnosis under complex working conditions," *Neurocomputing*, vol. 439, pp. 197-211, 2021.
- [22] S. Zhang, F. Ye, B. Wang, and T. G. Habetler, "Few-shot bearing fault diagnosis based on model-agnostic meta-learning," *IEEE Transactions on Industry Applications*, vol. 57, no. 5, pp. 4754-4764, 2021.
- [23] C. Yu, Y. Ning, Y. Qin, W. Su, and X. Zhao, "Multi-label fault diagnosis of rolling bearing based on meta-learning," *Neural Computing and Applications*, vol. 33, pp. 5393-5407, 2021.



DSPACE

<https://dspace.org/>

Model-agnostic meta-learning for fault diagnosis of industrial robots

Liu, Yuxin; Chen, Chong; Wang, Tao; Cheng, Lianglun; Qin, Jian

2023-10-16

Institute of Electrical and Electronics Engineers (IEEE)

<https://doi.org/10.1109/ICAC57885.2023.10275255>

Liu Y, Chen C, Wang T, et al., (2023) Model-agnostic meta-learning for fault diagnosis of industrial robots. In: 2023 28th International Conference on Automation and Computing (ICAC), 30 August - 1 September 2023, Birmingham, UK