A Few-Shot Learning Framework for Rotor Unbalance and Shaft Crack Fault Diagnostic Based on Physics-Informed Neural Network

WEIKUN DENG, KHANH T. P. NGUYEN, CHRISTIAN GOGU, JEROME MORIO and KAMAL MEDJAHER

ABSTRACT

This study aims to detect and localize rotor unbalance and shaft crack damage in a few-shot data scenario. It proposes a reinforcement learning-based approach with physics preferences (called RLP2). RLP2 is used to guide the physics consistency in the rotor finite element mimetic neural network (RFEMNN). The RFEMNN is first trained in an unsupervised manner using mixed simulation and experimental datasets in the task of reconstructing rotor's vibration signals. Then, the RFEMNN is fine-tuned in the RLP2 framework using a physics preference reward as policy loss to ensure similarity between hidden layer output and rotor system parameters. The output of the RFEMNN is fed into a downstream multi-output convolutional neural network (CNN) for fault diagnostic and localization. The proposed method's effectiveness is demonstrated through experiments on a PT500 platform under zero-shot, one-shot, and few-shot learning scenarios. The obtained results indicate the potential of this method for the predictive maintenance of rotor systems in real-world applications with limited training data.

INTRODUCTION

Nowadays, deep learning (DL) techniques have demonstrated promising capabilities for diagnosing structural faults in rotating machinery. These techniques typically rely on extracting fault indicators from vast amounts of data containing rich fault information. However, the use of uninterpretable, data-centralized black-box models poses challenges to the development and trustworthiness of Prognostics and Health Management (PHM) technology. To address the issue of limited data availability and to achieve improved generalization performance, the application of few-shot learning (FSL) in the PHM domain has garnered increasing interest [1]. In literature, FSL can be categorized into three groups based on different algorithmic principles, as described in [2]: meta-learning, metric learning, memory learning, and data augmentation. For instance, Tao et al. [3] employed an unknown matching network model, which integrates parameter optimization-based meta-learning and metric-based learning. This approach addresses the challenges of sparse fault samples and cross-domain datasets in real industrial set-

Weikun Deng, PhD Student, Email: weikun.deng@enit.fr. Laboratoire genie de production, Ecole Nationale d'Ingenieurs de Tarbes, Universite de Toulouse, France

tings. Zhang et al. [4] introduced a few-shot learning framework for bearing fault diagnostic based on model-agnostic meta-learning. By designing auxiliary classifiers as a regularizer corresponding to hidden inception layer features, the framework enables the training of an effective fault classifier using limited data. Regarding model architecture, Cui et al. [5] and Zhang et al. [6] proposed a Siamese network model with residual blocks for few-shot learning in bearing fault diagnostic. Wu et al. [7] presented a residual prototype network (RPN) for few-shot fault diagnostic of bearings by employing meta-learning concepts and updating the model parameters using a query set.

Although the promising results demonstrated in the aforementioned articles are notable, the literature also mentions some remaining major challenges. Specifically, there is a concern that purely data-centric learning models may not be sufficient to obtain feature representations capable of interpreting and extrapolating damage information in a manner similar to physics formulations. Therefore, this paper aims to fill this gap by proposing a few-shot learning solution from the perspective of physics-informed machine learning (PIML) . It introduces, in Section 2, a physics preference feedback in the reinforcement learning of the rotor finite element mimetic neural network (RFEMNN) architecture. To validate the few-shot learning capability of theproposed approach, experiments are conducted in Section 3, using a rotor unbalance and shaft crack test bench. Finally, the conclusion and perspective are presented in Section 4.

PROPOSED METHODOLOGY

This study presents a model that integrates the PIML architecture with RL. The prior physics knowledge and its application, which is viewed as the preferences in RL reward, will be presented in the following subsections. This paper assumes negligible motion and deformation of the rotor along the axial direction and represents the dynamic behavior of rotor systems by using the rotor finite element method (FEM). Specifically, the rotor is modeled as a continuum structure consisting of discrete elementary nodes, which are represented by a symmetry matrix ($m \times n, m \times n$), where m denotes the degree of freedom data for each node. In this study, m is set to 4. The main idea of FEM in fault type recognition and defects location is shown in Eq.(1). It compares the differences between the actual monitoring vibration signals and the predicted ones with the identification of the structure matrix changes caused by faults [8]. For example, when the rotor has an unbalance, it causes abnormal vibration because of the mass increases or decreases, which can be modeled by the theoretical value changes in the mass matrix $\mathbf{M}$. Similarly, the material stiffness at the crack location decreases and varies over operating time, leading the time-varying stiffness elements in stiffness matrix $\mathbf{K}$. The equivalent matrix $\mathbf{A}$ comes from the New-mark β numerical integration that contains mass $\mathbf{M}$, damping $\mathbf{D}$ and stiffness $\mathbf{S}$ information:

$$\begin{gathered} dt = 1/fs, \alpha_0 = 1/(\gamma \times dt^2), \alpha_1 = \beta/(\gamma \times dt) \\ \mathbf{A} = \mathbf{S} + \alpha_0 \mathbf{M} + \alpha_1 \mathbf{D} \end{gathered} \tag{1}$$

Establishing the relationship between structural faults, changes in the value of parameter $\mathbf{A}$, and vibration response can be challenging in finite element (FE) models due to their high dimensionality and a large number of unknown parameters. Various factors,

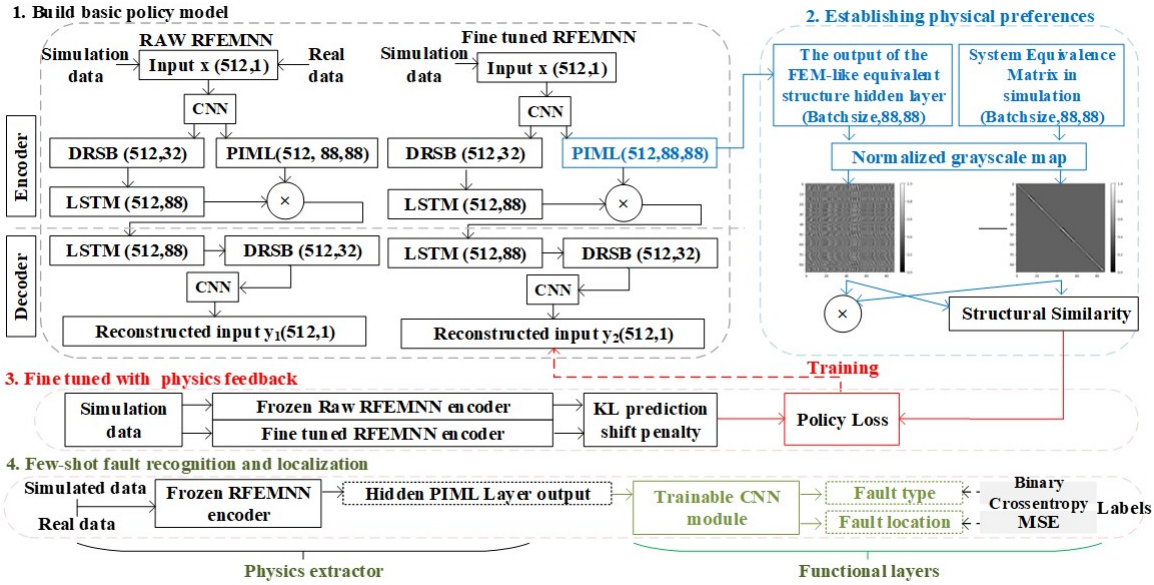


Figure 1. Overview of the proposed methodology

including rotor rotation speed, geometry, and material, can affect this relationship, and quantifying their effects can be difficult. Moreover, the vibration variations induced by small changes in $\mathbf{A}$ may be masked by noise in most fault conditions. Additionally, the types and locations of faults in the rotor shaft structure can be diverse and complexly coupled, leading to incomplete physics knowledge and making the modeling difficulty.

The proposed hybrid approach combines physics supervision with a trial-and-error method, as illustrated in Fig. 1. It enables few-shot learning by leveraging information about the structure matrix inferred from unlabeled simulation data and integrating it into the diagnostic process. It consists of 4 major steps: 1) Building the rotor finite element mimetic neural network (RFEMNN); 2) Design of the physics preferences; 3) Reinforcement learning for fine-tuning with physics consistency preferences; 4) Few-shot learning for fault recognition and localization. These steps are detailed as follows.

To efficiently model the matrix $\mathbf{A}$ in the physics-informed hidden layer structure, the geometric architecture of $\mathbf{A}$ is considered in the design of the neural network (NN) structure. Since $\mathbf{A}$ is a sparse symmetric matrix with a large number of zero elements, a masking method is proposed to simulate its mathematical form by processing the hidden layer outputs. This allows for the rapid construction of the network of finite element method (FEM) topology-based physics-informed machine learning (PIML) architecture. The purpose of the masking method is to restrict the positions of the weight parameters in the neuron layers that can be updated during training, in order to shape the latent space. The mask matrix is binary and has the same dimensions and non-zero distribution as the real symmetric square matrix shown in Figure 2.

The “Mask” process can implicitly incorporate physics knowledge using the same mathematical expression form by computing the Mask matrix deterministically. For unknown parameters of the physics knowledge, an NN layer serves to learn them, while for known parameters, a direct embedding is performed. After “Mask” process, the numerical distribution of the structure in the PIML layer becomes mathematically consistent

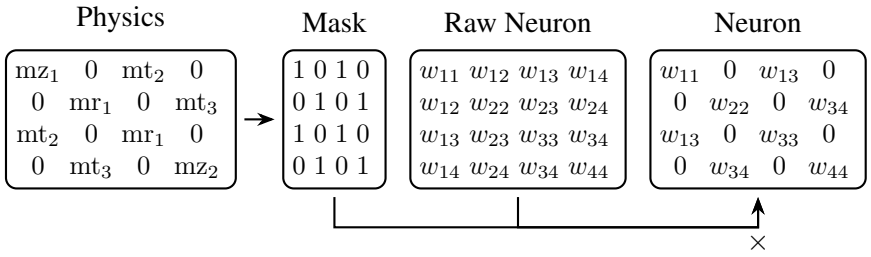


Figure 2. Illustration of the topology process applied to NN weight matrix.

with the rotor dynamics.

In the proposed PIML paradigm, we design a RL reward function that is based on the physics consistency between the "PIML output" and the equivalence matrix of the actual FEM system $\mathbf{A}$. During the simulation of the actual FEM system, $\mathbf{A}$ is automatically generated according to a manually given structure, which can be obtained inexpensively. The "Structural Similarity Index (SSIM)" is used to measure the difference between the two matrices as shown in Eq. (2).

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{2}$$

The SSIM metric takes into account three aspects of the matrix: values, contrast, and structure. In Eq. (2), x and y are two matrices being compared. The numerator of the equation represents the similarity of the values and contrast of the two matrices, while the denominator represents the similarity of the structural information. Here, μ_x and μ_y represent the average elements intensities of the two matrices, σ_x^2 and σ_y^2 represent the variances of the element's intensities, and σ_{xy} represents the covariance between the elements intensities of the two matrices. The constants c_1 and c_2 are small positive values added to the equation to avoid instability when the denominator is close to zero. The SSIM metric ranges from 0 to 1, with 1 indicating perfect similarity between the two matrices, i.e, the best physics consistency. The RFEMNN model is first trained on the

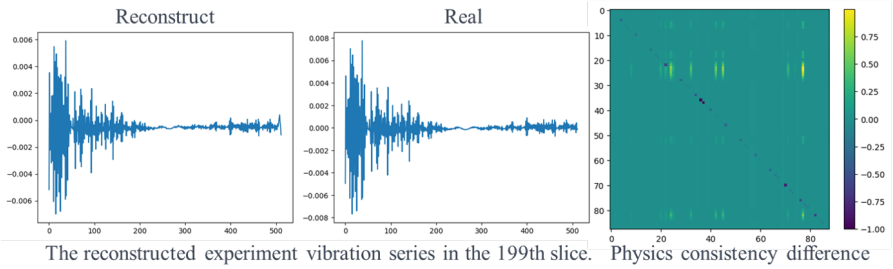


Figure 3. Investigation of the vibration reconstruction ability (left) and physics consistency(right) before RL fine-tuning

dataset that includes the real data and simulation data. This task is based on a reconstruction of the vibration sequence. The results are shown in Fig.3. Initially, despite the reconstruction error being relatively small (with an overall MSE of $3.2e^{-5}$), there is a noticeable divergence between the embedded physical knowledge, denoted as "PIML

out", and the simulated $\mathbf{A}$, as illustrated in Fig.3. This suggests that the model optimization trajectory does not align with the goal of achieving physical consistency. To address this, a fine-tuning procedure is introduced to maintain the feature encoding capability while improving the physical consistency, which is quantified by the difference between "PIML out" and simulated $\mathbf{A}$. This procedure corresponds to the third step depicted in Fig.1. In this process, we clone the pre-trained RFEMM as $Policy_{RFEMNN}$, while keeping the original pre-trained RFEMM intact and frozen as Raw_{RFEMNN}. Both these models are fed with a (512, 1) vector as input, derived from data slicing results of the raw vibration series. To refine the RL, we employ the Deep Deterministic Policy Gradient (DDPG) algorithm [9], using the terms in Eq.(3) as reward metrics. In this setup, the $Policy_{RFEMNN}$ serves as the actor, with its hidden layer output labeled as "PIML out" and Reconstructed vibration" forming the policy output.

$$Loss_{\text{Policy}} = -R_{\text{Policy RFEMNN}} = -(\alpha SSIM - \beta \text{KL}\left[||P_{\text{policy}}(s_t) - P_{\text{raw}}(s_t)||\right]) \quad (3)$$

where α and β are the hyperparameters that control the weight of the loss function. $P_{\text{RFEMNN}}(s_t)$ is the prediction of the Policy RFEMNN on the state s_t. $P_{\text{policy}}(s_t)$ and $P_{\text{raw}}(s_t)$ are the reconstructed vibration of the Raw_{RFEMNN} and $Policy_{RFEMNN}$ on the state s_t respectively. Its goal is to keep the time series recovery ability of the actor by considering the KL divergence between the output of two PIML modules, $Policy_{RFEMNN}$ and Raw_{RFEMNN}, and simultaneously enhance their physics consistency with the equivalence matrix $\mathbf{A}$. So that the "PIML out" can be viewed as a beneficial additional generalized knowledge representation for the few-shot fault diagnostic. A critic NN is designed to estimate the value function (Q-value) for a given state-action pair. It consists of three dense layers with 64 neurons each, using the ReLU activation function for the first two layers and no activation for the last layer, which outputs a single Q-value. In the policy update step, the difference between the reward and the policy output is used as the loss function, which is calculated using TensorFlow's Gradient-Tape and updated using the Adam optimizer for the "$Policy_{RFEMNN}$". During training in the few-shot test, a self-encoder based on DRSN-LSTM architecture serves as the policy autoencoder in the downstream model, see Fig.4, which is connected to the encoder part of the RFEMNN with frozen model parameters via a trainable CNN. The self-encoder establishes a compressed representation of relevant physics in the hidden "PIML" layer, which provides additional rich information for PHM tasks.

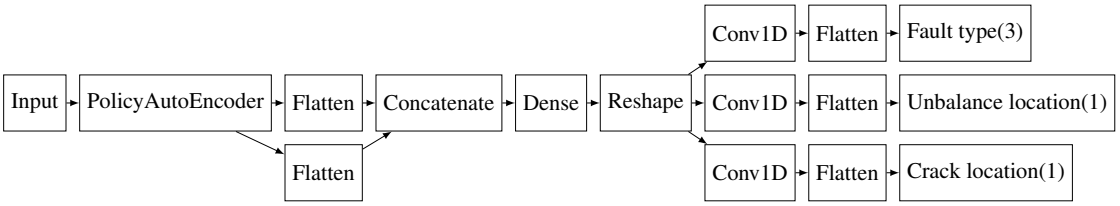
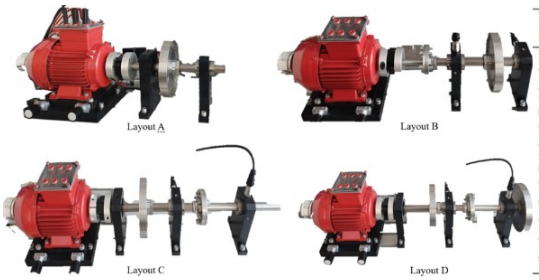


Figure 4. Downstream diagnostic module structure

EXPERIMENTAL VALIDATION



Health state	Shaft length (m)	Rotor structure	Rotating speed (rpm)	Fault position (m)	Samples
H_1	0.355	Layout A	1200, 1500, 1800, 2100, 2400, 2700, 3000	-	300
U_1	0.355	Layout A	1200, 1500, 1800, 2100, 2400, 2700, 3000	0.175	300
H_2	0.409	Layout B	1500, 1800	-	40
C_1	0.409	Layout B	1500, 1800	0.120	40
U_2	0.409	Layout B	1500, 1800, 2400	0.290	60
H_3	0.605	Layout C	1200, 1500, 1800	-	60
C_2	0.605	Layout D	1500, 1800, 2400	0.355	60
U_3	0.605	Layout D	1500, 1800, 2400	0.207	60
U_4	0.605	Layout D	1500, 1800, 2400	0.110	60
U_5	0.605	Layout D	1500, 1800, 2400	0.155	60
U&C_1	0.605	Layout C	1200, 1500, 1800, 2100, 2400	C:0.155, U:0.586	60
U&C_2	0.605	Layout C	1200, 1500, 1800, 2100, 2400	C:0.355, U:0.175	60

Figure 5. Fault experiment arrangement

This study is based on the Python-based rotor dynamics simulation library, ROSS [10], to model a 44-node rotor system and simulate various faults, including unbalanced defects, shaft cracks, and a combination of both. We generated 8940 vibration samples in the rotor fault conditions, each consisting of 10240 data points, with randomized fault locations and speeds. Additionally, we conducted experiments on a multi-structured test rig, shown in Fig. 5. The fine-tuning results of RLP2 are detailed in Fig.6. It shows the results of the vibration reconstruction after fine tune and the difference between the PIML output and the matrix "$\mathbf{A}$".

As can be seen in Fig.6, the fine-tuning process is a dynamic balance of satisfying physics knowledge and data-driven-based signal reconstruction accuracy. Referring to Eq. 3, the discrepancy observed in the comparison between the "Rewards" and "SSIM" curves denotes the KL divergence of the reconstructed time series results between the $Policy_{RFEMNN}$ and Raw_{RFEMNN} models, named time reconstruction MSE. At the end of the fine-tuning process, the time reconstruction MSE is found to be small enough (0.04), to ensure the similarity of the trend of the real and reconstructed vibration signals. In addition, significant improvements are made to the physics consistency of the RFEMNN structure. The color map representing channel differences in the square matrix shows that before training in Fig.3, the hidden layer output exhibited distinct channel traits and differences, with a widespread different values and the largest difference concentrated in a few diagonal neurons and channels. This indicates that the distribution did not align with the band symmetry of the physics matrix and the diagonal distribution in $\mathbf{A}$. Post fine-tuning, the PIML output showed marked similarity with matrix $\mathbf{A}$, as validated by a maximum SSIM value of 0.984 and a reward of 0.941. The square matrix color map in Fig.6 underscores a decreased difference between the PIML output and $\mathbf{A}$, particularly along the diagonal. This signifies RFEMNN ability to consistently represent structural and vibrational responses.

VALIDATION OF THE PHYSICS PROPERTIES' CONSISTENCY

In this study, 2000 labeled vibration signal samples are generated using the ROSS library, involving random simulation of a 0.5g unbalance, cracks with a depth of 0.01 shaft diameter, and axis diameter variation within a 44-node FE model. Slices of length 512 with a step size of 256 are extracted using a sliding window approach from each sample in the auxiliary dataset D_A. The proposed few-shot learning method is validated under different settings: zero-shot, one-shot, and few-shot, using various sample

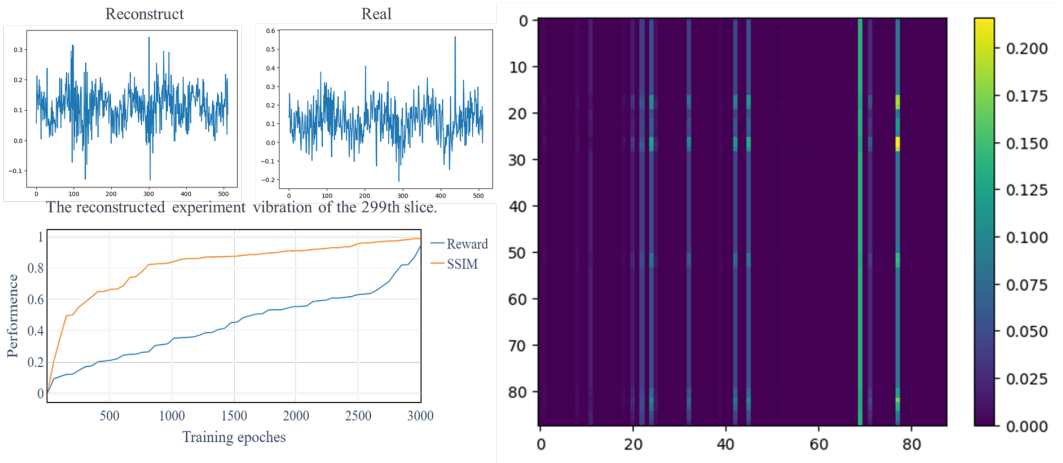


Figure 6. Investigation of the vibration reconstruction ability (left) and physics consistency(right) after RL fine-tuning

sources in the training data. Location accuracy is evaluated using MAE on one of the 10 folds of the original experimental data, and the results are presented in Table I. The zero-shot learning achieves a diagnostic accuracy rate of 44.87%, correctly classifying 2866 samples. However, the RFEMNN model tends to classify most samples as either unbalanced or combined faults due to the dataset's configuration and distribution bias. One-shot learning improves the diagnostic performance, especially in identifying crack faults, by adding only one real fault data to the training. Few-shot learning yields the best results, with an accuracy of 90.7%, enabling clearer identification and localization of faults with few-shot samples. Nonetheless, false alarm situations persist, particularly misclassifications of unbalanced and healthy samples. To enhance accuracy and reliability, further refinements such as increasing data slice length and selecting representative simulated samples for D_A are necessary. Additionally, the model exhibits partial diagnostic capability for fault location even under zero-shot and one-shot scenarios, owing to the physical consistency of "PIML out" with the simulation data from 2000 sets.

CONCLUSION

In this paper, a novel few-shot fault diagnostic method based on Physics-Informed Machine Learning (PIML) is presented. The approach achieved a generalized physics representation of system behavior by enforcing physics constraints during RL fine-tuning. The validation of the method's physics consistency is done by comparing similarity changes between "PIML out" and the equivalent matrix $\mathbf{A}$ of simulated data pre- and post-fine-tuning via reinforcement learning. Experimental results demonstrated the few-shot fault diagnostic capability using real data from varying label percentages. The proposed approach effectively captured the structural information of the system matrix for prognostics and health management. Notably, reasonable diagnostic accuracy is achieved in one-shot and few-shot scenarios. Future work will focus on efficient embedding techniques for heuristic physics knowledge and faster methods for designing physics-informed structures with incomplete knowledge, further enhancing the method's applicability and effectiveness.

TABLE I. Diagnostic results in different few-shot sample scenarios (Z: Zero-shot, O: one-shot, F: Few-shot, H: Healthy, U: Unbalance, C: Crack, U&C: Unbalance and Crack, LA: Location accuracy.)

		Diagnostics results														
		H			U			C			U&C			LA(%)		
		Z	O	F	Z	O	F	Z	O	F	Z	O	F	Z	O	F
Real results	H	0	680	1255	1102	1134	560	0	1	0	713	0	0	-	-	-
	U	0	245	8	1097	2074	2285	0	1	63	1259	36	0	65.79	92.0	99.2
	C	0	2	0	55	61	0	0	262	411	256	86	0	51.79	84.8	97.1
	U&C	0	0	0	36	158	0	0	2	0	1769	1645	1805	-	-	-

ACKNOWLEDGMENT

This work is carried at LGP, INP-ENIT, with the funding from French-American Cultural Exchange foundation and Chinese Scholarship in Education and Science.

REFERENCES

1. Ding, P., M. Jia, and X. Zhao. 2021. "Meta deep learning based rotating machinery health prognostics toward few-shot prognostics," *Applied Soft Computing*, 104:107211.
2. Li, X., Z. Sun, J.-H. Xue, and Z. Ma. 2021. "A concise review of recent few-shot meta-learning methods," *Neurocomputing*, 456:463–468.
3. Tao, H., L. Cheng, J. Qiu, and V. Stojanovic. 2022. "Few shot cross equipment fault diagnosis method based on parameter optimization and feature mertic," *Measurement Science and Technology*, 33(11):115005.
4. Zhang, S., F. Ye, B. Wang, and T. Habetler. 2021. "Few-Shot Bearing Fault Diagnosis Based on Model-Agnostic Meta-Learning," *IEEE Transactions on Industry Applications*, 57(5):4754–4764.
5. Cui, Z., X. Kong, and P. Hao. 2021. "Few-shot Learning for Rolling Bearing Fault Diagnosis Based on Residual Convolutional Neural Network," in *2021 4th International Conference on Artificial Intelligence and Big Data (ICAIBD)*, IEEE.
6. Zhang, A., S. Li, Y. Cui, W. Yang, R. Dong, and J. Hu. 2019. "Limited Data Rolling Bearing Fault Diagnosis With Few-Shot Learning," *IEEE Access*, 7:110895–110904.
7. Wu, K., C. Chen, Z. Song, K. Yu, J. Zhou, and J. Wu. 2022. "Residual Prototype Network-based Method for Few-Shot Fault Diagnosis of Bearing," in *2022 Global Reliability and Prognostics and Health Management (PHM-Yantai)*, IEEE.
8. Kushwaha, N. and V. Patel. 2020. "Modelling and analysis of a cracked rotor: a review of the literature and its implications," *Archive of Applied Mechanics*, 90(6):1215–1245.
9. Lillicrap, T. P., J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra. 2015. "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*.
10. Timbó, R., R. Martins, G. Bachmann, F. Rangel, J. Mota, J. Valério, and T. G. Ritto. 2020. "ROSS - Rotordynamic Open Source Software," *Journal of Open Source Software*, 5(48):2120.