Original Article

SHM Structural Health Monitoring

Structural Health Monitoring
1–14

DOI: 10.1177/14759217231188002
journals.sagepub.com/home/shm
Sage

Dual-input anomaly detection method based on deep reinforcement learning

Yuxiang Kang[1], Guo Chen[2], Hao Wang[3], Wenping Pan[1] and Xunkai Wei[3]

Abstract

Aiming at the problem of low accuracy of unsupervised learning anomaly detection algorithm, a dual-input anomaly detection method based on deep reinforcement learning was proposed. The proposed model mainly consists of a feature extractor and anomaly detector. Based on the deep reinforcement learning framework, the feature extractor uses a dual-input deep neural network to form the current value network and the target value network, which are used to extract the low-dimensional feature vectors. Based on the 3σ principle, the reward function of reinforcement learning is designed to reward and punish the output results of the model during training. The model was trained only with the normal data, and the extracted feature vector of the normal class was used as the input of the anomaly detector to complete the learning of the detector. During the test, the input anomaly detection was realized based on the dual-input convolutional neural network, and the anomaly detector was completed by learning. To illustrate the generality and generalization performance of the proposed method, four sets of image data and two sets of rolling bearing fault data in different fields were verified respectively. At the same time, the proposed method is applied to the fault detection of a real aero-engine rolling bearing.The results show that the proposed model has high anomaly detection accuracy, which is superior to the current optimal method.

Keywords

Deep reinforcement learning, anomaly detection, dual-input deep neural network, unsupervised learning, rolling bearings

As one of the universal and key components of aero-engine, rolling bearings will increase the cost of maintenance and even cause unexpected accidents once a failure occurs. Therefore, it is of great significance to explore more accurate, more efficient, and more intelligent early fault detection technology, so that the monitoring of aero-engine rolling bearings can be realized in the early stage of bearing faults.[1,2]

In recent years, signal analysis methods represented by empirical wavelet transform (EWT), minimum entropy deconvolution (MED), singular spectrum analysis (SSA), and so on have achieved good detection results in rolling bearing fault detection. Sawalhi et al.[3] applied MED to fault diagnosis of rolling bearings and showed that impact characteristics in fault signals of rolling bearings could be significantly improved by using MED. Gilles[4] proposed the Empirical EWT, which has been applied to the resonance frequency band extraction of rolling bearing fault signals.[5] SSA[6] is another powerful signal-processing method, which has been well applied in rolling bearing fault detection in recent years. To improve the detection efficiency of SSA, Bhowmik et al.[7] combining with eigen-perturbation (EP) theory, proposed an error-corrected First-order eigen perturbation (FOEP)-based formulation of Real-time singular spectrum analysis (RSSA) for adaptive filtering and verified the proposed model. At the same time, Bhowmik et al.[8] studied the application of first-order EP techniques FOEP in structural health monitoring and verified it through experiments. Although the above methods have achieved good detection results in rolling bearing fault detection, however, the above methods are mainly verified in the laboratory environment. For the aero-engine, the working environment is more complex, and the signal contains more noise components, so it

[1]College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, China
[2]College of General Aviation and Flight, Nanjing University of Aeronautics and Astronautics, Nanjing, China
[3]Beijing Aeronautical Engineering Technical Research Center, Beijing, China

Corresponding author:
Guo Chen, Nanjing University of Aeronautics and Astronautics, No.29 Jiangjun Avenue, Jiangning District, Nanjing City, Jiangsu Province, China.
Email: cgnuaacca@163.com

still needs further verification to directly apply the above method to the fault detection of the aero-engine rolling bearing.

At present, with the development of artificial intelligence technology in the fields of image and voice,[9,10] the technology has also shone in the field of rolling bearing fault diagnosis.[11] In recent years, Deep Learning (DL) methods such as Convolutional Neural Network (CNN),[12] Transfer Learning,[13] Deep Belief Network,[14] and Autoencoder[15] had been widely used in the field of rolling bearing fault diagnosis and obtained good diagnostic results. To be effective, such methods typically require that the data has a similar distribution, a comparable amount of data, and contains human labels. However, in actual engineering, it is difficult to obtain the operation data of the mechanical system in the "sick" state. As a result, the obtained samples contain a large number of normal data and the abnormal data amount is very small or even zero in various typical fault states. In this reality, the unsupervised anomaly detection (AD)[16] method driven only by normal data has become an effective way to realize the early AD of rolling bearing.

In recent years, researchers have proposed a variety of AD methods, including traditional methods based on probability and statistical theory and methods based on DL. Traditional AD methods such as the Support Vector Data Description (SVDD),[2] Hyperspherical Distance Discrimination (HDD), or Principal Component Analysis (PCA), were usually used for rolling bearing AD. Wang[17] used Sparse Non-negative Matrix Factorization results as the input of SVDD, for establishing a composite fault AD method for rolling bearings, and realized the accurate AD of composite faults of rolling bearings. Lin et al.[18] proposed a novel HDD method to assess the performance of aero-engine bearings, which can solve problems with many parameters and high computational complexity. Liu et al.[19] used PCA and the decision tree algorithm to realize the early fault warning of civil engine rolling bearings, and the warning rate is up to 99.8%. However, Classical AD methods often required human experience to provide features and fail in high-dimensional samples.[20]

In view of some inevitable defects of traditional AD methods, AD based on DL has become a research hotspot in image, video, finance, and other fields in recent years. Huang et al.[21] proposed a novel method, namely, memory residual regression autoencoder to realize the AD of rolling bearings, which can achieve 97.97% and 93.51% accuracy on the IMS (Developed by NSFI/UCR Intelligent Maintenance System Center, University of Cincinnati, OH, USA) and XJTU-SY (Made by Xi'an Jiaotong University, Xi'an, Shaanxi, China) datasets. Zhao et al.[22] combined sparse autoencoder and transfer learning to propose a network model for rolling bearing AD. To reduce the dependence on manual feature extraction, Ruff et al.[20] proposed the Deep Support Vector Data Description (DSVDD) method by combining DL and SVDD. DSVDD uses a CNN for feature extraction. Then, a hypersphere describing normal samples is constructed. When the sample falls into the sphere, it is expressed as normal, otherwise, it is abnormal. Akcay S et al.[23] proposed an AD method based on Generative Adversarial Networks (GANs) and verified the effectiveness of the method on several common datasets. Perera P et al.[24] proposed a classification model based on GAN. In addition, there are Deep AD methods based on autoencoders, Boltzmann machine-based methods, etc.[25] Mao et al.[26] used DSVDD for early AD of rolling bearings and achieved good results. At the same time, Deep one-class classification neural networks (OC-NN)[27] had also been applied in the early AD of rolling bearings. Although the current unsupervised Deep AD methods have achieved good results in many fields, however, Deep AD based on unsupervised learning methods still has room for further improvement in detection accuracy, and few researchers have introduced Deep Reinforcement Learning (DRL) theory into AD.

Given this, based on the comprehensive study of the existing DSVDD methods, this paper proposes a dual-input AD method based on DRL (DADDRL). DADDRL uses a dual-input DL network to extract data features and detects the extracted features based on traditional AD methods. The method is verified on multiple data sets.

The novelty of this paper mainly includes the following points:

(1) The Deep AD method and reinforcement learning are effectively integrated, and a high-precision AD method is proposed.
(2) The method proposed in this paper is tested on multiple real aero-engines of the same type, which proves the effectiveness of the method proposed in this paper, which is one of the novelties of the method in this paper.

Related work

Before introducing the work of DADDRL, we first briefly introduce the related Deep SVDD and DRL methods.

Deep Support Vector Data Description

The basic principle of DSVDD is shown in Figure 1. It uses large-scale normal training data for learning, and obtains a DL network transformation with the weight

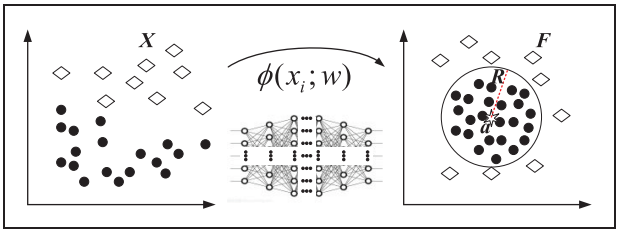

Figure 1. Basic principle of Deep Support Vector Data Description (DSVDD).

of w. $\phi(x_i; w)$ maps the data representation of most input spaces to a minimum hypersphere in the output space. Normal and abnormal samples fall into the interior and exterior of the hypersphere after $\phi(x_i; w)$ mapping. DSVDD does not need to manually label the samples during the initial sample input and training process. Therefore, DSVDD is an unsupervised AD method.

$\phi(x_i; w)$ in DSVDD uses a CNN model. By establishing the objective function shown in Equation (1), the stochastic gradient descent algorithm is used to complete the training of the network.

$$L = \min_{R,w} R^2 + \frac{1}{vn}\sum_{i=1}^{n} \max\left\{0, \|\phi(x_i; w) - a\|^2 - R^2\right\} + \frac{\lambda}{2}\sum_{l=1}^{L} \|w_l\|_F^2 \tag{1}$$

Where: R and a denote the radius and center of the hypersphere; x_i denotes the i-th training sample; $\|\cdot\|_F$ denotes the Frobenius norm; and v and λ are hyperparameters. The first term in the objective function minimizes the radius of the hypersphere to make the volume of the hypersphere as small as possible. The second term punishes the data points falling outside the hypersphere. The third term regularizes the network weights.

Deep reinforcement learning method based on classification Markov decision process

DRL, which combines the advantages of RL and DL, has been favored by many scholars since it was proposed. The purpose of DRL is to maximize the benefits of agents in the process of trial and error with the environment. Classification tasks can be regarded as sequential decision problems of agents, which are classification Markov decision process (CMDP).[28] The process is usually composed of state set S, action set A, decision reward R, state transition matrix P, and

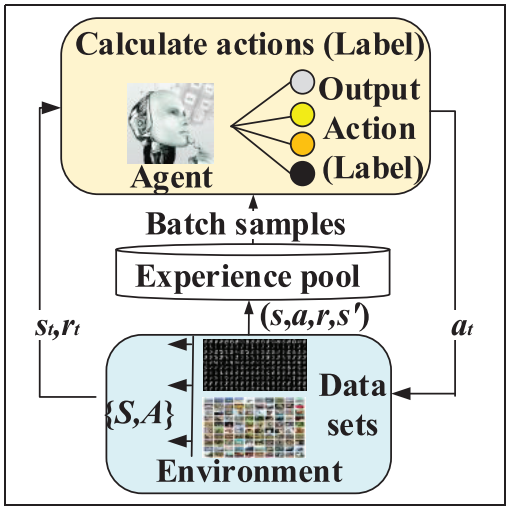

Figure 2. Classification Markov decision process.

discount factor γ. Firstly, it is assumed that the training data set is $D = \{(x_1, l_1), (x_2, l_2), \ldots, (x_N, l_N)\}$, where (x_i, l_i) represents the ith sample x_i and label l_i, N represents the total number of samples. If the agent correctly judges the label of the sample x_i during interaction with the environment, the environment gives a positive reward and vice versa. Figure 2 shows the specific process of CMDP.

The specific definitions of the relevant variables in CMDP and Figure 2 are as follows :

State set S: The state of the environment is determined by the training sample, and the state s_t at time t corresponds to the t-th sample x_t in the training set. When the T-step training is completed, new training is started after randomly disrupting the sample set D.

Action set A: The action of the agent is determined by the label of the sample, $A = \{0, 1, 2, \ldots, K-1\}$, where K is the number of classes of the sample. Action a_t is the label of the output after the agent accepts the current state s_t.

Decision reward R: The reward r_t represents the feedback given by the environment after the agent performs the action a_t under the state s_t, that is $S \times A \rightarrow R$, it is used to reflect whether the action performed by the agent is correct.

State transition matrix P: The state transition probability p represents the probability that the state s_t transferred to the next state s_{t+1} according to the existing training sample D after the agent executes the action a_t in the state s_t, expressed as $p(s_{t+1}|s_t, a_t)$. Since the training sample set D is determined in CMDP, the state transition matrix p is determined.

Discount factor γ: $\gamma \in [0, 1]$ is used to balance future and current rewards. The larger the γ, the more the

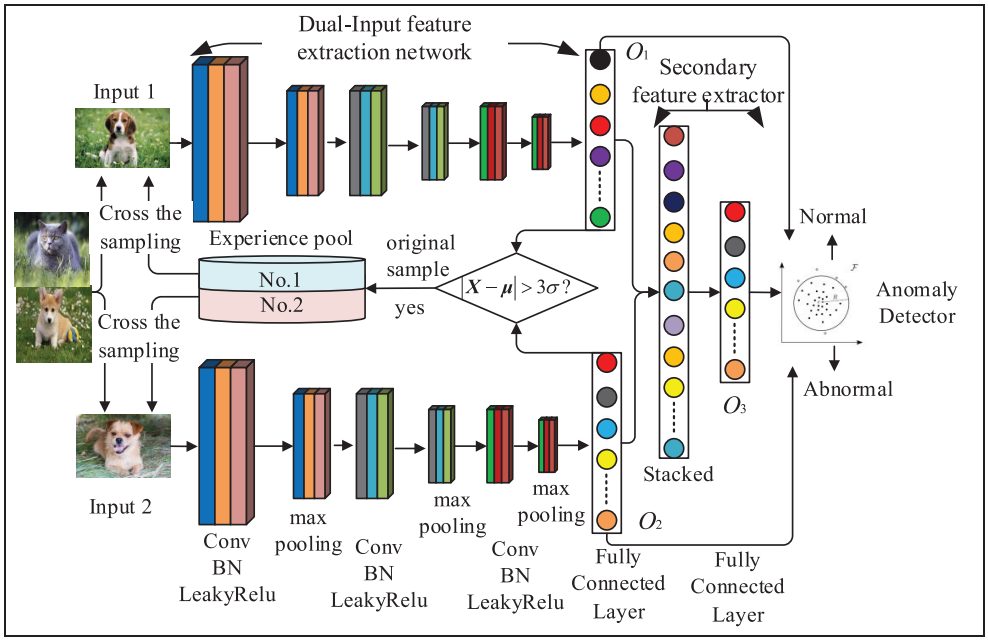

Figure 3. Dual-input deep neural network.

agent pays attention to long-term return; The smaller the γ, the more the agent cares about immediate interests.

Episode: Represents the transition trajectory from the initial state s_1 to the termination state s_T. Episode $= \{s_1, a_1, r_1, s_2, a_2, r_2, \ldots, s_T, a_T, r_T\}$.

Policy $\pi(\theta)$: Indicates the mapping between the state s_t and the corresponding action a_t, $S \rightarrow A$. In this article, agents with parameters θ are represented $\pi(\theta)$.

To achieve more accurate classification, it is necessary to find the optimal state-to-action mapping strategy $\pi(\theta)$, and the process of exploring the optimal strategy $\pi(\theta)$ can be realized through deep reinforcement learning.

Dual-input AD method based on DRL

First, a dual-input deep neural network with an experience pool structure is introduced. Then the overall structure of the DADDRL model and the calculation methods of the loss function are introduced.

Dual-input deep neural network

The established dual-input deep neural network is mainly composed of a dual-input feature extraction main framework, a secondary feature extractor, an anomaly detector, and an experience pool. The specific structure is shown in Figure 3.

(1) Dual-input deep neural networks can use models such as Resnet[29] and Long Short-Term Memory (LSTM).[30] This paper uses a CNN with a three-layer structure. The output results $O_1 \in R^{1\times m}$, $O_2 \in R^{1\times m}$ of the two networks are stacked for secondary feature extraction.

(2) Experience pool structure is introduced into the dual-input deep neural network. For the batch samples of this training, samples whose output deviates greatly from the overall mean are stored in the designed experience pool cache.

(3) *Anomaly detector*. The extracted features O_1, O_2, O_3 are stacked as input to the anomaly detector (it is worth noting that the anomaly detector does not participate in the network training process and is only used for testing). Different anomaly detectors have different detection effects. The anomaly detector used in this paper is ① SVDD detector[2]; ② Gaussian Mixture Model (GMM) detector[31]; ③ PCA detector[32]; ④ Empirical Cumulative Outlier Detection (ECOD) detector.[33] For the specific algorithm please refer to the relevant literatand (https://github.com/Minqi824/ADBench).

Dual-input AD method based on DRL

The dual-input deep neural network that removes the anomaly detector is used as the current value network and target value network structure of DRL, and the specific model structure of DADDRL is constructed as shown in Figure 4.

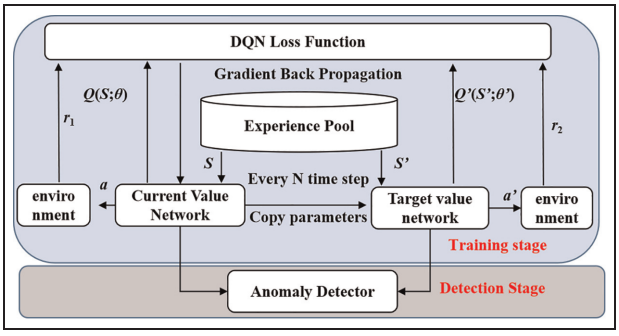

Figure 4. Structure of dual-input anomaly detection method based on deep reinforcement learning (DADDRL).

The environment in DADDRL is mainly used to calculate the reward value r of the result of the current value network and the target value network; the experience pool and network structure are shown in Section "Dual-input deep neural network"; the lossfunction and reward value r are calculated as shown in Sections "New loss function" and "Experience pool structure."

New loss function

The loss function proposed in this paper mainly includes two parts. One is the reconstruction error loss of the output results of the current value network and the target value network. The second is the reward loss of two networks.

Assuming that n training samples have p features after CNN mapping $\mathbf{X} = (\mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}, \ldots, \mathbf{X}_{ip})$, $i = 1, 2, \ldots, n$. $\mathbf{X}$ obeys the distribution $\mathbf{X} \sim \mathbf{F}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with expectation $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \ldots, \mu_p)^T$ and covariance matrix $\boldsymbol{\Sigma}$.

We can use Equations (2) and (3) to calculate the expectation $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$,denoted as $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$

$$\hat{\boldsymbol{\mu}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i \tag{2}$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n} (\mathbf{X}_i - \hat{\mathbf{u}})(\mathbf{X}_i - \hat{\mathbf{u}})^T \tag{3}$$

Dual-input deep neural network with three outputs O_1, O_2, O_3. The global $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ of all n samples can be calculated from Equations (2) and (3), denoted as global distribution $\hat{\boldsymbol{\kappa}} = [\hat{\boldsymbol{\mu}}_1, \hat{\boldsymbol{\mu}}_2, \hat{\boldsymbol{\mu}}_3]$, $\hat{\boldsymbol{\chi}} = [\hat{\boldsymbol{\Sigma}}_1, \hat{\boldsymbol{\Sigma}}_2, \hat{\boldsymbol{\Sigma}}_3]$. Simultaneously, we calculate the $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ of the sample for the current batch size, denoted as local distribution $\bar{\boldsymbol{\kappa}} = [\bar{\boldsymbol{\mu}}_1, \bar{\boldsymbol{\mu}}_2, \bar{\boldsymbol{\mu}}_3]$, $\bar{\boldsymbol{\chi}} = [\bar{\boldsymbol{\Sigma}}_1, \bar{\boldsymbol{\Sigma}}_2, \bar{\boldsymbol{\Sigma}}_3]$.

To make the local and the global distribution similar, the covariance matrix and the expectation should be equal. So, we can get the sub-loss function as shown in Equation (4)

$$\begin{aligned} \hat{L}_t &= \sum_{i=1}^{p} \left(\hat{\boldsymbol{\kappa}}_i^t - \hat{\boldsymbol{\kappa}}_i^t\right)^2 \\ \bar{L}_t &= \sum_{j=1}^{p}\sum_{i=1}^{p} \left(\hat{\boldsymbol{\chi}}^t{}_{ij} - \hat{\boldsymbol{\chi}}^t{}_{ij}\right)^2 \end{aligned} \tag{4}$$

Where: $\bar{L}_t$ is the t-th ($t = 1, 2, 3$) covariance matrix error and $\hat{L}_t$ is expectation error.

For the dual-input deep neural network driven only by normal samples, we consider that its outputs O_1 and O_2 should have the same distribution, so we can get the second sub-loss function as shown in Equation (5).

$$\begin{aligned} \hat{L}_{12} &= \sum_{i=1}^{p} \left(\bar{\mu}^1{}_i - \bar{\mu}^2{}_i\right)^2 \\ \bar{L}_{12} &= \sum_{j=1}^{p}\sum_{i=1}^{p} (\bar{\Sigma}_{ij}^1 - \bar{\Sigma}_{ij}^2)^2 \end{aligned} \tag{5}$$

According to the established sub-loss function, we use the maximum-minimum algorithm to calculate the reconstruction error of the dual-input deep neural network, as shown in Equation (6):

$$L_{Re} = \min\max\left([\hat{L}_t, \bar{L}_t, \hat{L}_{12}, \bar{L}_{12}]\right) \tag{6}$$

Therefore, according to Equation (6), the loss function of the current value network and the target value network in deep reinforcement learning is shown in Equation (7).

$$L_{CT} = \left(L_{\mathrm{Re}}^C - \gamma L_{\mathrm{Re}}^T\right)^2 \tag{7}$$

Where: L_{Re}^C and L_{Re}^T is the reconstruction error of the current value network and the target value network. γ is the discount factor.

For reward loss, the 3σ method is used to calculate the reward value for the current output. After using Equation (2) to calculate the global expectation $\hat{\boldsymbol{\mu}}$ (in this case, it is a p-dimensional vector), we calculate its mean β and variance σ and the mean $\hat{\mu}_r$ of the p-dimensional characteristics of the three outputs of a single sample O_1, O_2, O_3, using Equation (8) as the reward function of DRL

$$r = \begin{cases} 0 & |\hat{\mu}_r - \beta| - 3\sigma < 0 \\ -1 & |\hat{\mu}_r - \beta| - 3\sigma > 0 \end{cases} \tag{8}$$

After calculating the reward values of the three outputs of a single sample through Equation (8), the actual reward value of the network is the mean of the three. For the current value network, it is recorded as r_1, and for the target value network, it is recorded as r_2. The reward loss is shown in Equation (9):

$$L_r = (r_1 - \gamma r_2)^2 \tag{9}$$

The loss function used in this paper is shown in Equation (10):

$$L = L_r + L_{CT} \tag{10}$$

Experience pool structure

The main function of the experience pool structure is to store heterogeneous samples in normal samples during training. These samples can obtain more training opportunities by balanced cross-sampling. We use 3σ method to judge heterogeneous samples during training. First, the samples with negative reward values calculated by Equation (8) are regarded as heterogeneous samples in the normal class and stored in the experience pool. Then, during training, the composition of the batch sample is 30% from the experience pool and 70% from the original dataset. In the training process, it is guaranteed that the two input samples selected each time are different samples. In the detection stage, two input samples are the same sample.

Experimental validation of image data sets

To verify the effectiveness and versatility of the DADDRL method in AD, it was verified on six sets of data sets. There are four sets of image data sets and two sets of rolling bearing fault data sets. The information of the four sets of image data sets is:

Mnist[34] is a well-known digital recognition data set composed of 70,000 handwritten grayscale digital images with a size of 28×28. The Mnist training set contains 60,000 samples and the test set contains 10,000 samples.

Cifar10[35] is a natural image dataset. The objects in the image come from objects in daily life. It consists of 60,000 32×32 size 10 color images, each class of 6000 images.

Fashion-mnist;[36] the training set of Fashion-mnist contains 60,000 samples and the test set contains 10,000 samples. Compared with Mnist, Fashion-mnist is a more challenging data set. The dataset contains 10 categories, and the size of each sample is 28×28 grayscale images.

Street View House Number (SVHN)[37] is a real-world digital image data set obtained from the house number of google street view images, which consists of more than 600,000 digital images. SVHN's style is similar to that of Mnist, but it comes from a very difficult, unsolved real-world problem (identifying numbers and numbers in natural scene images).

In the test, one of the classes is the normal class and samples from the remaining classes are used to represent anomalies. We evaluate the results quantitatively via Area Under Curve (AUC) measure by using ground truth labels in testing.

The Graphics Processing Unit (GPU) used in this paper is NVIDIA GTX1660 6G; i5-9600K processor; the running system is Windows 10; 8G memory; the programming language is python 3.7; the framework of all DL models is Pytorch 1.11; the batch sample size is 256; the number of iterations is 200; using Adam optimization algorithm, the learning rate is 0.001.

Effects of different data preprocessing and detectors on the results

To verify the different data preprocessing methods, the influence of abnormal detectors on the detection results and the generalization performance of the model, the DADDRL model is verified by using the multiple anomaly detector proposed in Section "Dual-input deep neural network" (the calculation method of abnormal scores in the detection process refers to the corresponding references) and five data preprocessing methods. The data preprocessing methods are: (1) L1-norm; (2) L2-norm; (3) Zero-Centered (ZC); (4) Maximum Normalization (MN); and (5) Raw Data (RD). In the verification process $\gamma = 0.9$ and the experience pool size is 2000. Each class is calculated 10 times, and then the average value is calculated as the final result of the data set. The experimental results are shown in Table 1.

The results show that for different anomaly detectors, the AD results are not the same after using different data preprocessing methods. Specifically, on the four datasets, the detection accuracy of the anomaly detector ④ (ECOD) is significantly better than that of the other three detectors. When ECOD is selected, for the Mnist data set, the detection accuracy reaches the maximum 97.8% after using the data preprocessing method MN. For the Cifar10 data set, the optimal value is 75.5% when the data preprocessing method is L1. The best data preprocessing methods for Fashion-mnist and SVHN datasets are ZC and MN, respectively. The results show that different data preprocessing methods are needed to improve the detection accuracy for different data sets.

To show that the proposed method still has higher detection accuracy than other methods under different data preprocessing methods, DSVDD,[20] SVDD,[2] Deep convolutional autoencoder (DCAE),[25] Kernel Density Estimation (KDE),[13] Anomaly Detection with Generative Adversarial Networks (ANOGAN),[24] OC-NN[27] are used for comparative verification. The specific parameters of several methods used are consistent with the original references. DADDRL uses an ECOD anomaly detector. The verification results are shown in Table 2.

Table 1. DADDRL detection results under different AD and data preprocessing methods.

Data preprocessing methods	Mnist (%)				Cifar10 (%)				Fashion-mnist (%)				SVHN (%)			
	①	②	③	④	①	②	③	④	①	②	③	④	①	②	③	④
L1	**95.7**	94.5	92.6	95.1	74.9	74.3	72.8	**75.5**	**93.6**	91.8	90.6	92.4	57.5	57.2	56.4	**57.8**
L2	96.3	95.4	93.2	**96.5**	62.3	60.6	57.8	**64.1**	91.9	91.5	91.2	**92.1**	57.9	55.6	52.1	**58.9**
ZC	97.4	96.6	95.3	**97.7**	70.9	68.9	67.8	**71.4**	92.8	92.2	91.3	**93.3**	61.8	61.4	58.2	**62.7**
MN	97.5	94.7	90.6	**97.8**	68.9	65.6	63.4	**69.4**	93.4	91.3	89.2	**93.8**	64.1	62.6	57.3	**64.4**
RD	97.3	97.2	94.6	**97.6**	65.8	66.5	63.8	**67.7**	91.5	92.6	91.4	**93.3**	57.9	56.3	54.7	**58.2**

AD: anomaly detection; ZC: zero-centered; MN: maximum normalization; RD: raw data; DADDRL: dual-input anomaly detection method based on deep reinforcement learning. The bold represents the optimal detection result obtained by the current data preprocessing method.

Table 2. Detection results of various AD methods under different data preprocessing methods.

Data set	Detector	DSVDD	SVDD	DCAE	KDE	ANOGAN	OC-NN	DADDRL
Mnist (%)	L1	92.5	90.8	89.4	86.8	90.9	93.7	95.1
	L2	94.2	91.3	94.9	89.6	93.4	94.1	96.5
	ZC	92.7	90.9	92.8	90.5	90.6	93.8	97.7
	MN	95.6	91.7	91.7	89.7	92.4	94.6	97.8
	RD	91.5	86.3	84.6	88.8	94.3	89.7	97.6
Average		93.3	90.2	90.7	89.1	92.3	93.2	96.9
Cifar10 (%)	L1	64.9	64.5	59.8	65.2	63.0	62.6	75.5
	L2	59.0	55.6	51.3	62.6	57.8	58.9	64.1
	ZC	59.0	57.9	59.6	65.9	66.8	67.2	71.4
	MN	62.2	55.8	61.4	63.5	66.7	63.8	69.4
	RD	58.5	49.7	53.8	56.9	60.2	60.5	67.7
Average		60.7	56.7	57.2	62.8	62.9	62.6	69.4
Fashion-mnist (%)	L1	90.7	79.6	89.7	78.4	88.9	90.2	92.4
	L2	75.6	78.4	88.6	81.6	85.4	90.5	92.1
	ZC	92.1	86.3	90.9	86.3	86.7	91.5	93.3
	MN	91.5	81.9	92.3	80.6	88.9	88.8	93.8
	RD	92.1	87.8	91.2	85.4	92.2	89.7	93.3
Average		88.6	82.8	90.5	82.5	88.4	90.1	93.0
SVHN (%)	L1	54.3	50.3	55.2	50.6	52.4	52.6	57.8
	L2	50.6	48.6	54.8	48.7	53.9	55.9	58.9
	ZC	52.1	49.8	55.9	47.9	54.2	53.4	62.7
	MN	49.6	54.3	53.7	50.8	53.8	52.6	64.4
	RD	53.9	49.2	50.6	52.3	51.7	50.7	58.2
Average		52.1	50.4	54.0	50.1	53.2	53.0	60.4

AD: anomaly detection; DSVDD: Deep Support Vector Data Description; SVDD: Support Vector Data Description; ZC: zero-centered; MN: maximum normalization; RD: raw data; DADDRL: dual-input anomaly detection method based on deep reinforcement learning.

The comparison results show that under the premise of different data preprocessing, the detection results of various methods are not the same. The detection effect of DADDRL is significantly better than that of the other methods. For example, for the Mnist dataset DADDRL, the detection result is 96.9%, while the result of the second-ranked DSVDD method is 93.3%, which is 3.6% higher than DSVDD. This value reached 6.5% for the Cifar10 dataset; the result on the Fashion-mnist dataset is 2.5%; it is 6.4% on the SVHN dataset. It can be seen from the above statistical results that DADDRL has stronger AD ability than the other methods.

To further verify the generalization performance and AD advantages of DADDRL, the L1 preprocessing method is used to compare the detection results of various methods on the specific categories of Mnist and Cafar10 datasets. As shown in Table 3.

The results show that DADDRL has the highest AUC value on both datasets. In terms of subdivision, in the detection of class 0 on the Mnist dataset DADDRL, compared with the second-ranked DCAE, the AUC value is increased by about 2.1%; the detection of class 3 increased by about 7.1%; in the detection of class 4, the AUC value is 3.9% higher than the second-ranked SVDD result; in the detection of class

Table 3. Detection results of various AD methods on Mnist and Cifar10 datasets.

Normal class	DSVDD	SVDD	DCAE	KDE	ANOGAN	OC-NN	DADDRL
0	96.9	95.5	97.6	97.3	96.1	96.2	99.7
1	99.5	99.1	98.6	98.6	99.2	99.4	99.9
2	86.3	82.3	85.4	79.1	85.8	98.6	98.8
3	86.9	88.2	86.2	86.6	89.6	88.1	96.7
4	93.5	94.3	84.8	87.3	90.2	93.3	98.2
5	84.5	77.1	77.6	73.5	83.6	86.8	96.0
6	98.1	96.6	93.8	87.9	92.7	98.5	99.5
7	94.9	93.4	92.6	90.9	94.1	94.7	97.6
8	88.9	88.5	86.8	78.8	84.3	87.2	95.2
9	95.9	93.2	90.8	87.6	93.8	94.1	97.5
Airplane	65.7	61.2	59.6	61.4	68.7	61.2	80.1
Automobile	67.3	63.4	58.5	64.8	55.8	63.6	77.6
Bird	51.5	49.6	49.3	50.7	53.6	64.7	67.5
Cat	60.6	55.3	59.5	56.9	56.5	54.5	64.8
Deer	57.2	65.8	54.2	66.8	64.3	68.5	75.6
Dog	67.7	61.9	62.8	61.7	63.8	59.2	70.1
Frog	65.6	74.8	51.6	75.2	58.9	62.7	78.3
Horse	63.7	62.9	58.9	63.4	62.6	63.9	75.1
Ship	78.9	74.5	75.6	74.8	77.8	65.8	83.3
Truck	71.1	75.6	68.2	75.9	67.8	61.9	82.4

AD: anomaly detection; DSVDD: Deep Support Vector Data Description; SVDD: Support Vector Data Description; DADDRL: dual-input anomaly detection method based on deep reinforcement learning.

5, the AUC value is 9.2% higher than the second-ranked OC-NN result; compared with the second-ranked DSVDD results, it is improved by 6.3% in the detection of class 8. On the Cifar10 dataset, the best detection results of DADDRL are in the detection of Airplan and Automobile. The results of DADDRL reached more than 77.0%, while the results of ANOGAN and DSVDD models were 68.7% and 67.5%, which were 11.4% and 10.3% higher than those of Airplane and Automobile. In the detection of normal class Horse, the result of DADDRL is 75.1%, while the result of the second-ranked DSVDD model is 63.7%, which is 11.4% higher than that of the second-ranked DSVDD model. In the Ship and Truck, the AUC values of DADDRL reached 83.3% and 82.4%, while the second-ranked results were 78.9% and 75.9%, an increase of 4.4% and 6.5%. In addition, the accuracy of DADDRL is also improved to varying degrees in the detection of the remaining class. The above comparison results show that the AD accuracy of DADDRL is significantly better than that of the other models. It is also proved that DADDRL has excellent AD ability.

Influence of experience pool structure and γ on detection results

To illustrate the influence of experience pool structure and γ value on the detection results. Select the optimal preprocessing method and ECOD detector corresponding to each data set in Section "Effects of different data preprocessing and detectors on the results." The model is verified by setting (Note EP) or without experience pool (Note WEP) and different γ values. When $\gamma = 0.0$, only the loss caused by the current value network is retained in the loss function, and the influence of each part of the loss function on the result can be further verified by this value. The verification results are shown in Table 4.

The results show that different γ values have different effects on the model, but as the γ value increases, the detection accuracy of the model on the three data sets increases. It reaches its maximum when the γ value is 0.8–0.9. When $\gamma = 0.0$, the detection results on the four datasets are the minimum, which further illustrates that the components of the loss function proposed in this paper are indispensable. In addition, the experience pool structure proposed in this paper also has a certain impact on detection accuracy. The detection accuracy of the experience pool on the four data sets is generally higher than that of the non-experience pool structure, which also shows that the experience pool structure in the model has a certain impact on the accuracy of the model.

Rolling bearing fault data set validation

To verify the effectiveness of DADDRL in rolling bearing fault detection, the fault test data sets of the aero-

Table 4. The influence of experience pool structure and γ on test results.

γ	Mnist (MN)		Cifar10 (L1)		Fashion-mnist (MN)		SVHN (MN)	
	EP	WEP	EP	WEP	EP	WEP	EP	WEP
0.0	96.1	95.1	70.4	68.6	92.1	91.5	62.5	61.7
0.1	96.2	96.0	71.6	71.4	92.5	91.8	63.0	62.1
0.2	96.6	95.8	72.8	70.2	92.4	91.7	63.4	62.5
0.3	96.4	95.9	73.3	70.3	92.8	92.0	63.4	62.8
0.4	96.8	96.4	73.2	71.6	92.9	92.2	63.8	63.2
0.5	97.1	96.5	73.8	72.2	93.1	92.1	63.6	63.6
0.6	97.2	96.3	74.6	73.4	93.0	92.5	64.1	62.9
0.7	97.4	97.4	74.1	72.9	93.4	92.7	64.0	63.7
0.8	97.9	97.8	74.3	73.2	93.3	93.0	64.3	64.0
0.9	**97.8**	97.6	**75.5**	73.6	**93.8**	92.9	**64.4**	63.9
0.95	97.6	97.7	74.9	73.1	**93.7**	93.3	64.5	64.2

EP: eigen-perturbation. The bold represents the optimal results of the structure with or without experience pool under different *r*, where EP represents the structure with experience pool and WEP represents the structure without experience pool.

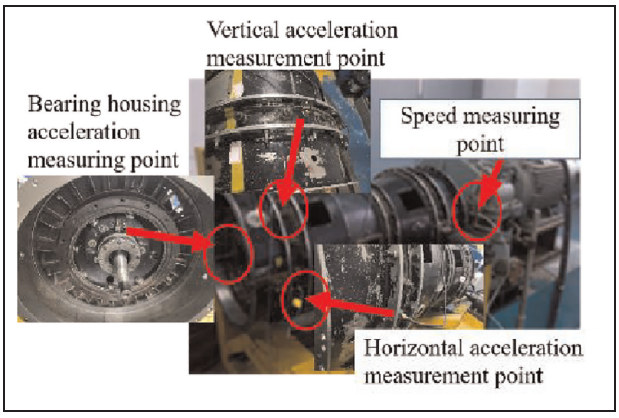

Figure 5. Acro-engine rotor tester.

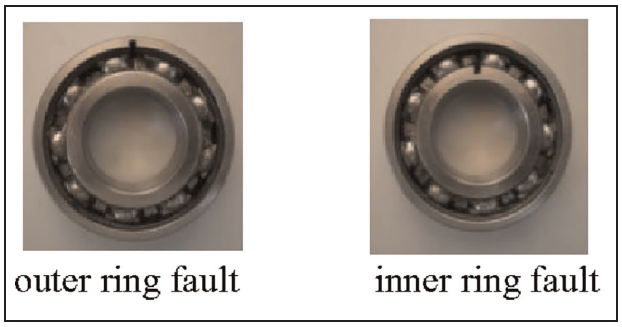

Figure 6. Rolling bearing fault diagram.

engine bearing outer ring and inner ring based on casing signal are verified in the Intelligent Diagnosis and Expert System of Nanjing University of Aeronautics and Astronautics, and compared with the DSVDD, OC-NN, and ANOGAN methods with better detection results in Section "Experimental validation of image data sets."

The aero-engine bearing failure test based on casing signals was performed on the platform shown in Figure 5, which was a 1:3 scale imitation of a real engine. The test platform can effectively reflect the attenuation characteristics of the aero-engine vibration signal in the transmission process. The bearing model used in the test is a 6206 single row deep groove ball bearing. During the test, the following fault defects were processed by electrical discharge machining (EDM) cutting: cracks with a width of 6mm in the outer ring and inner ring respectively. The specific defect diagram is shown in Figure 6.

The vibration acceleration sensor (B&K4805) and NI USB9234 data collector are used in the test. The sampling frequency is 10,240 Hz and the sample data point is 8192. The test speed is 2400 r/min. The sensor installation position is shown in Figure 5. The obtained vibration acceleration signal is directly converted into two-dimensional matrix sample data. When the algorithm proposed in this paper is applied to the aero-engine bearing fault detection based on casing signal, only the samples in normal state are used to train the model, and the threshold is set by the 3σ method. The samples in normal and abnormal states are tested at the same time, and the samples larger or smaller than the threshold range are regarded as abnormal (fault). Sample information is shown in Table 5.

The detection results on two sets of rolling bearing fault data sets are shown in Table 6. The results show that the proposed method can better realize the AD of rolling bearings, and the AD accuracy can reach 100%, followed by DSVDD detection accuracy of 99.6%, followed by OC-NN of 98.4%, and the effect is poor. The ANOGAN method has an accuracy of 97.8%.

Table 5. Bearing fault detection sample information.

Bearing state	Speed (rpm)	Sample size
Normal	2400	150
Outer ring fault	2400	116
Inner ring fault	2400	110

Table 6. Detection result.

Fault location	DADDRL	DSVDD	OC-NN	ANOGAN
Outer ring	100	99.7	98.4	98.6
Inner ring	100	99.6	98.8	97.8

DSVDD: Deep Support Vector Data Description; DADDRL: dual-input anomaly detection method based on deep reinforcement learning.

To illustrate the effectiveness of the proposed method, the results of DADDRL are visualized, as shown in Figure 7. The results show that the output results of DADDRL in three states are in three different intervals and the performance is stable. This difference also provides a basis for the fault detection of rolling bearings by formulating relevant thresholds.

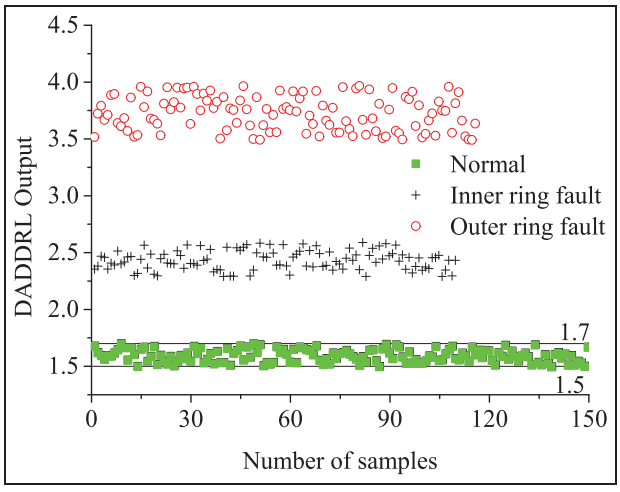

Figure 7. Test results of rolling bearing data set.

A real aero-engine rolling bearing fault AD validation

Validation on a full-life dataset

To demonstrate the practical efficacy of the DADDRL method, it has been validated on a real aero-engine rolling bearing life-cycle fault data set. The aero-engine testing was conducted in Beijing from September to November 2021, culminating in successful completion. The total duration of the test was about 150 h, the sampling frequency was 200,000 Hz, the data length of a single sample was 1 s, and the storage interval of the sample was 3 s. During the test, the Endevco vibration acceleration sensor was installed on the intermediate casing of the aero-engine. The final test was stopped due to the excessive vibration parameters of the entire machine. After disassembly, it was found that the outer ring of the three-point rolling bearing had serious spalling. Expert analysis revealed that this fault began to occur after approximately 120 h of use. Therefore, this paper only uses normal class data from the first 40 h to train the model. To reduce computation, approximately 50 groups of samples with rotation speeds greater than 14,000 r/min were screened every hour during verification, resulting in a total of 7010 data samples, of which 2000 were used for training. During the data preprocessing process, FFT is first applied to a single sample of data. The resulting spectrum data is then converted into a grayscale image with dimensions of $316 \times 316 \times 1$ (approximately 100,000 spectrum points), which serves as the input for the model. Part of the original vibration acceleration data and the corresponding spectrum are shown in Figure 8 (to improve the resolution, only the spectrum data in the range of 0–10,000 Hz are shown). The results show that the spectrum of vibration acceleration also shows corresponding changes as the test proceeds, mainly showing an increase in the amplitude of the low frequency band and a decrease in the amplitude of the high frequency band. Figure 9 and Table 7 show the final comparison results using various methods.

The comparative results demonstrate that DADDRL can achieve a detection accuracy of over 98%, which is 10.36% higher than the ANOGAN method with the highest detection accuracy among the other three methods. These findings indicate that the DADDRL approach is highly practical for actual fault detection in aero-engine rolling bearings. The output values of DADDRL and ANOGAN in Figure 9 further verify the above conclusions. The output value of the DADDRL method tends to be 0 in the normal stage but is much larger than 0 in the abnormal stage, which confirms our previous findings. However, for most samples of ANOGAN in the fault stage, there is no significant difference between the normal and abnormal output values, making it difficult to determine the current operating state of the bearing. The output difference between the two models during the abnormal stage also indicates that DADDRL performs better than ANOGAN.

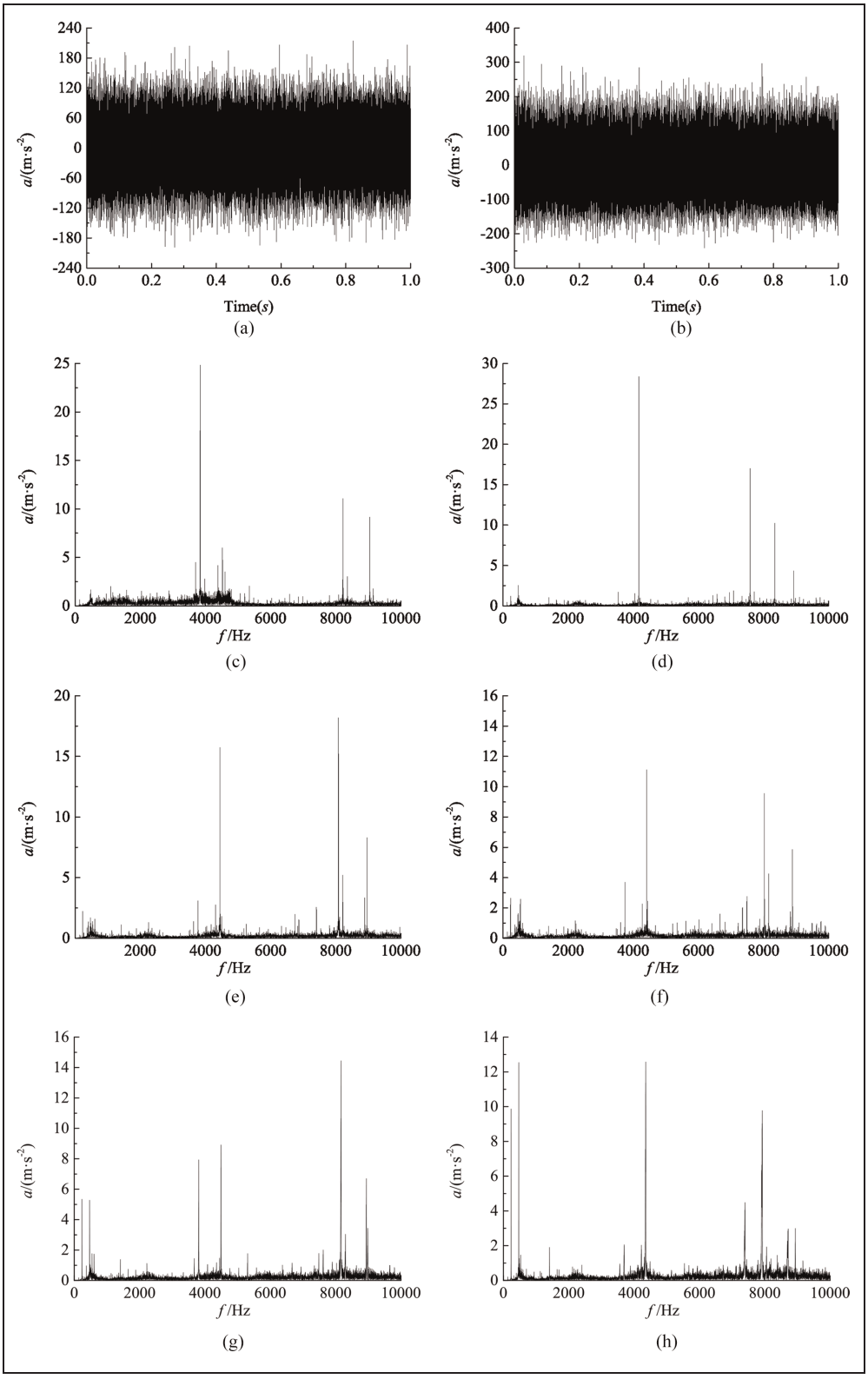

Figure 8. Partial raw vibration acceleration data and corresponding spectrogram: (a) vibration acceleration data at the 100th hour, (b) vibration acceleration data at the 150th hour, (c) spectrogram of vibration data at the 100th hour, (d) spectrogram of vibration data at the 110th hour, (e) spectrogram of vibration data at the 120th hour, (f) spectrogram of vibration data at the 130th hour, (g) spectrogram of vibration data at the 140th hour, and (h) spectrogram of vibration data at the 150th hour.

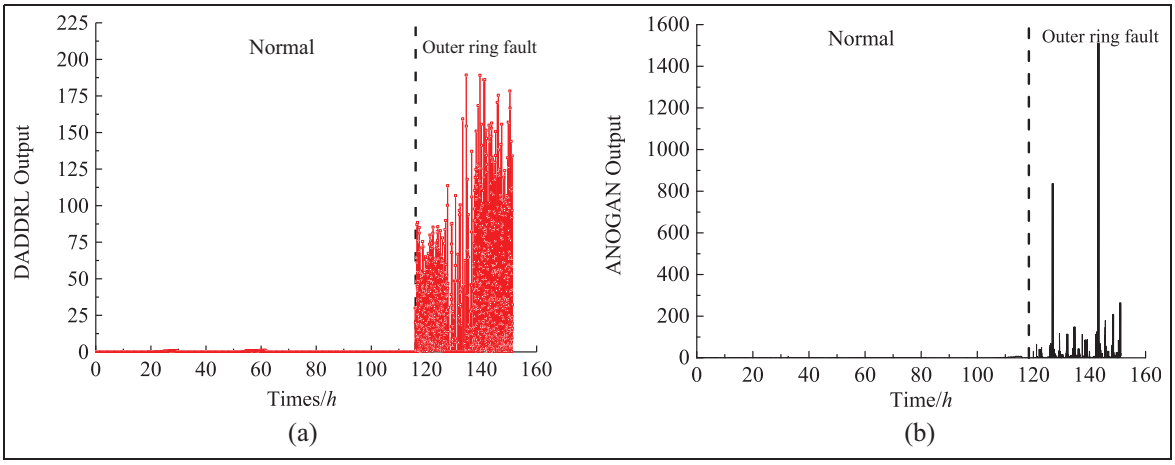

Figure 9. Outputs of both models on real aero-engine data: (a) output of DADDRL on real aero-engine data and (b) output of ANOGAN on real aero-engine data.
DADDRL: dual-input anomaly detection method based on deep reinforcement learning.

Table 7. Test results of rolling bearing data of an aero-engine(%).

Fault location	DADDRL	DSVDD	OC-NN	ANOGAN
Outer ring	98.82	82.67	85.32	88.46

DSVDD: Deep Support Vector Data Description; DADDRL: dual-input anomaly detection method based on deep reinforcement learning.

Table 8. Validation results on different aero-engines.

Aero-engine	Aero-engine 1 (Training)	Aero-engine 2	Aero-engine 3
AUC(%)	98.82	100.0	100.0

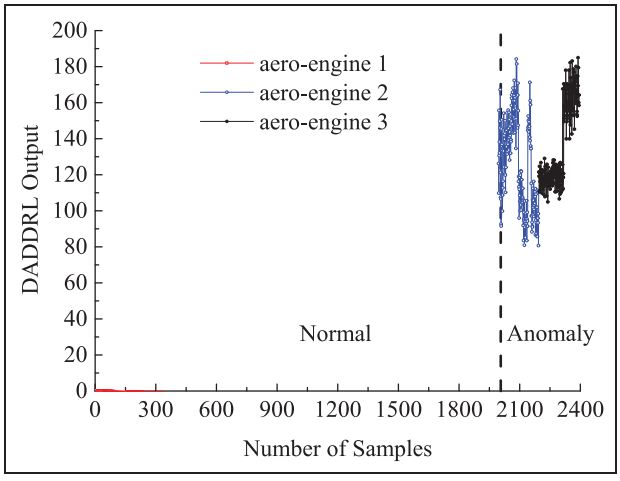

Figure 10. Output results of DADDRL on different aero-engines.
DADDRL: dual-input anomaly detection method based on deep reinforcement learning.

Verification of model generalization performance

To further verify the effectiveness and generalization performance of DADDRL in the fault AD of aero-engine rolling bearings, the model trained in Section "Validation on a full-life dataset" is used to verify two other aero-engines of the same type. The two aero-engines used are three-point rolling bearing outer ring spalling faults, and 200 sets of samples are selected for each engine for testing. The verification results are shown in Table 8 and Figure 10.

The results show that the trained DADDRL model achieves 100% detection accuracy on the other two engines. The results in Figure 10 show that the output value of DADDRL is distributed around 0 in the normal stage, while the output value on the other two engines is far greater than 0, and the output value on the other two engines has the same trend. Comparing the output value can intuitively distinguish normal and abnormal. The results further show that the model has strong generalization performance and good practicability.

Conclusion

This paper proposes a DADDRL. Four image data sets and two rolling bearing fault data sets in different fields were verified respectively. The following conclusions can be drawn:

(1) Compared with the existing unsupervised AD methods, DADDRL has higher detection accuracy.
(2) The detection results of various detection models are different under different data preprocessing methods, but the DADDRL method has the highest detection accuracy.
(3) DADDRL uses two-stage detection, namely the feature extraction stage and the AD stage. Different anomaly detectors have different detection accuracy. The effect of the ECOD detector in this paper is better than that of the other detectors.
(4) The detection results on the rolling bearing dataset show that bearing fault detection can be realized only by relying on the normal class samples.
(5) The validation results on real aero-engine rolling bearing fault data sets show that the proposed method has strong practicability.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work has been supported by the National Science and Technology Major Project (J2019-IV-004-0071); National Natural Science Foundation of China (52272436).

ORCID iD

Yuxiang Kang https://orcid.org/0000-0002-5973-7766

Data availability

The data that has been used is confidential.

References

1. Guo C. Feature extraction and intelligent diagnosis for ball bearing early faults. *Acta Aeronaut Astronaut Sin* 2009; 30(2): 362–367.
2. Pacheco-Chérrez J, Fortoul-Díaz JA, Cortés-Santacruz F, et al. Bearing fault detection with vibration and acoustic signals: Comparison among different machine learning classification methods. *Eng Fail Anal* 2022; 139: 106515.
3. Sawalhi N, Randall RB and Endo H. The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis. *Mech Syst Signal Process* 2007; 21(6): 2616–2633.
4. Gilles J. Empirical wavelet transform. *IEEE Trans Signal Process* 2013; 61(16): 3999–4010.
5. Cao H, Fan F, Zhou K, et al. Wheel-bearing fault diagnosis of trains using empirical wavelet transform. *Measurement* 2016; 82: 439–449.
6. Bhowmik B, Krishnan M, Hazra B, et al. Real-time unified single-and multi-channel structural damage detection using recursive singular spectrum analysis. *Struct Health Monit* 2019; 18: 563–589.
7. Bhowmik B, Panda S, Hazra B, et al. Feedback-driven error-corrected single-sensor analytics for real-time condition monitoring. *Int J Mech Sci* 2022; 214: 106898.
8. Bhowmik B, Tripura T, Hazra B, et al. First-order eigenperturbation techniques for real-time damage detection of Vibrating systems: theory and applications. *Appl Mech Rev* 2019; 71(6): 060801.
9. Hinton GE, Osindero S and Teh Y. A fast learning algorithm for deep belief nets. *Neural Comput* 2006; 18: 1527–1554.
10. He K, Zhang X, Ren S, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: *Proceedings of the IEEE international conference on computer vision*, 2015, Santiago, Chile, pp. 1026–1034.
11. Lei Y, Yang B, Jiang X, et al. Applications of machine learning to machine fault diagnosis: a review and roadmap. *Mechanical Systems and Signal Process* 2020; 138: 106587.
12. Xu G, Liu M, Jiang Z, et al. Online fault diagnosis method based on transfer convolutional neural networks. *IEEE Trans Instrum Meas* 2019; 69: 509–520.
13. Xia M, Shao H, Williams D, et al. Intelligent fault diagnosis of machinery using digital twin-assisted deep transfer learning. *Reliab Eng Syst Saf* 2021; 215: 107938.
14. Li W-H, Shan W-P and Zeng X-Q. Bearing fault identification based on deep belief network. *J Vibr Eng* 2016; 29(2): 340–347.
15. Wen J, Yan C, Sun J, et al. Bearing fault diagnosis method based on compressed acquisition and deep learning. *Chin J Sci Instrum* 2018; 39(1): 171–179.
16. Seliya N, Abdollah Zadeh A and Khoshgoftaar T. A literature review on one-class classification and its potential applications in big data. *J Big Data* 2021; 8(1): 122.
17. Wang H. Fault diagnosis of rolling element bearing compound faults based on sparse no-negative matrix factorization-support vector data description. *J Vibr Control* 2018; 24(2): 272–282.
18. Lin T, Chen G, Ouyang W, et al. Hyper-spherical distance discrimination: a novel data description method for aero-engine rolling bearing fault detection. *Mech Syst Signal Process* 2018; 109: 330–351
19. Liu Y, Wang C and Zhou P. An early warning method for rolling bearing fault of civil aero-engine. *J Propul Technol* 2022; 43(2): 295–304.
20. Ruff L, Vandermeulen RA, Görnitz N, et al. Deep one-class classification. In: *International conference on machine learning*, Stockholm, Sweden, 2018.
21. Huang X, Wen G, Dong S, et al. Memory residual regression autoencoder for bearing fault detection. *IEEE Trans Instrum Meas* 2021; 70: 351551.

22. Zhao X, Jia M and Liu Z. Fault diagnosis framework of rolling bearing using adaptive sparse contrative auto-encoder with optimized unsupervised extreme learning machine. *IEEE Access* 2020; 8: 99154–99170.
23. Akcay S, Atapour-Abarghouei A and Breckon TP. GANomaly: semi-supervised anomaly detection via adversarial training. In: *Computer Vision – ACCV, Perth, Australia, 2018*, pp. 622–637.
24. Perera P, Nallapati R and Xiang B. OCGAN: one-class novelty detection using GANS with constrained latent representations. In: *IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, Long Beach, USA, 2019, pp. 2893–901.
25. Ruff L, Kauffmann JR, Vandermeulen RA, et al. A unifying review of deep and shallow anomaly detection. *Proc IEEE* 2021; 109(5): 756–795.
26. Mao W, Chen J, Liang X, et al. A new online detection approach for rolling bearing incipient fault via self-adaptive deep feature matching. *IEEE Trans Instrum Meas* 2020; 69(2): 443–456.
27. Chalapathy R, Menon AK and Chawla S. Anomaly detection using one-class neural networks.*arXiv preprint arXiv:1802.06360*, 2018.
28. Wiering MA, Hasselt HV, Pietersma AD, et al. Reinforcement learning algorithms for solving classification problems. In: *Adaptive dynamic programming and reinforcement learning (ADPRL)*, Paris, France, 2011.
29. He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition. In: *IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, USA, 2016, pp. 770–778.
30. Hochreiter S and Schmidhuber J. Long short-term memory. *Neural Comput* 1997; 9(8): 1735–1780.
31. Zong B, Song Q, Min MR, et al. Deep autoencoding gaussian mixture model for unsupervised anomaly detection. In: *ICLR 2018 conference blind submission*, vancouver, canada, 2018.
32. Shyu M-L, Chen S-C, Sarinnapakorn K, et al. A novel anomaly detection scheme based on principal component classifier. Technical report, Miami University Coral Gables Florida Department of Electrical and Computer Engineering, 2003.
33. Li Z, Zhao Y, Hu X, et al. ECOD: unsupervised outlier detection using empirical cumulative distribution functions. IEEE Trans-actions on Knowledge and Data Engineering. arXiv preprint arXiv:2201.00382, 2022.
34. LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *Proc IEEE* 1998; 86(11): 2278–2324.
35. Krizhevsky A. Learning multiple layers of features from tiny images. Master's Thesis, Department of Computer Science, University of Toronto, Canada, 2009.
36. Xiao H, Rasul K and Vollgraf R. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747*, 2017.
37. Netzer Y, Wang T, Coates A, et al. Reading digits in natural images with unsupervised feature learning. In: *NIPS workshop on deep learning and unsupervised feature learning*, Sierra Nevada, Spain, 2011.