

Article

Study on Few-Shot Fault Diagnosis Method for Marine Fuel Systems Based on DT-SViT-KNN

Shankai Li ^{1,2}, Liang Qi ^{1,2,*} , Jiayu Shi ¹, Han Xiao ¹, Bin Da ^{1,2}, Runkang Tang ¹ and Danfeng Zuo ¹

¹ School of Automation, Jiangsu University of Science and Technology, Zhenjiang 212100, China; 221210301111@stu.just.edu.cn (S.L.)

² Jiangsu Shipbuilding and Ocean Engineering Design and Research Institute, Zhenjiang 212100, China

* Correspondence: alfred_02030210@just.edu.cn

Abstract: The fuel system serves as the core component of marine diesel engines, and timely and effective fault diagnosis is the prerequisite for the safe navigation of ships. To address the challenge of current data-driven fault-diagnosis-based methods, which have difficulty in feature extraction and low accuracy under small samples, this paper proposes a fault diagnosis method based on digital twin (DT), Siamese Vision Transformer (SViT), and K-Nearest Neighbor (KNN). Firstly, a diesel engine DT model is constructed by integrating the mathematical, mechanism, and three-dimensional physical models of the Medium-speed diesel engines of 6L21/31 Marine, completing the mapping from physical entity to virtual entity. Fault simulation calculations are performed using the DT model to obtain different types of fault data. Then, a feature extraction network combining Siamese networks with Vision Transformer (ViT) is proposed for the simulated samples. An improved KNN classifier based on the attention mechanism is added to the network to enhance the classification efficiency of the model. Meanwhile, a Weighted-Similarity loss function is designed using similarity labels and penalty coefficients, enhancing the model's ability to discriminate between similar sample pairs. Finally, the proposed method is validated using a simulation dataset. Experimental results indicate that the proposed method achieves average accuracies of 97.22%, 98.21%, and 99.13% for training sets with 10, 20, and 30 samples per class, respectively, which can accurately classify the fault of marine fuel systems under small samples and has promising potential for applications.



Academic Editors: Yongbo Li, Bing Li, Khandaker Noman and Teng Wang

Received: 30 November 2024

Revised: 19 December 2024

Accepted: 21 December 2024

Published: 24 December 2024

Citation: Li, S.; Qi, L.; Shi, J.; Xiao, H.; Da, B.; Tang, R.; Zuo, D. Study on Few-Shot Fault Diagnosis Method for Marine Fuel Systems Based on DT-SViT-KNN. *Sensors* **2025**, *25*, 6. <https://doi.org/10.3390/s25010006>

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: marine fuel system; few-shot fault diagnosis; Siamese network; transformer; KNN

1. Introduction

Currently, the majority of vessels globally are propelled by diesel engines [1]. In contrast to other shipboard equipment, diesel engines consist of various subsystems, including the fuel system, lubricating oil system, and cooling water system [2]. These subsystems exhibit varying probabilities of fault due to their distinct functions and operating conditions. An analysis of shutdown faults provided by the Association of British Diesel Engine Engineers and Users indicates that fuel system faults account for 60% of the total diesel engine faults, which represents the highest proportion [3]. Consequently, research on fault diagnosis for the fuel system is pivotal for enhancing the reliability of diesel engines and ensuring the safe navigation of vessels.

Digital twin, as an emerging virtualization method, enables more accurate fault diagnosis by building virtual models of physical entities for simulated computation and deduction [4]. The scholarly community has initiated a series of research on the application

of DT for fault diagnosis. Liu et al. [5] leveraged structural digital twin technology to generate simulated monitoring signals and as an input to the designed network, significantly enhancing its performance of fatigue damage detection across different structures. Xu et al. [6] developed a DT model of a centrifugal pump, which was used to simulate a spectrum of health conditions and augmented the dataset, finally constructing a graphical convolutional neural network, and facilitating cross-domain fault diagnosis. Yi et al. [7] made a significant stride by integrating a bearing dynamics model within a neural network architecture. Their innovative method involved using actual vibration data to construct a DT model of the bearing, thereby enriching the fault repository and enhancing the diagnostic accuracy. Liu et al. [8] analyzed faults under constant and variable speed conditions by establishing a fault diagnosis test platform. Then, the obtained data are integrated into the Unity3D platform to realize online diagnosis and the digital twin model is conducted for subsequent fault diagnosis of reducer. Despite the notable achievements of DT technology in these areas, its application in diesel engine fault diagnosis still needs to be explored.

Alternatively, deep learning has achieved excellent results in the field of fault diagnosis due to its ability to extract data features and use them fully [9]. Nonetheless, deep learning-based diagnostic methodologies typically necessitate an extensive array of fault samples, which is difficult to realize in practice due to many constraints. Few-shot learning provides a solution to this problem [10,11], enabling models to learn new tasks or categories with few labeled samples quickly. At present, most few-shot learning methods are based on data augmentation [12], transfer learning [13], and meta-learning [14], with researchers worldwide actively engaged in exploring these approaches. Li et al. [15] introduced a generalized few-shot classification framework founded on Convolutional Neural Networks (CNN). This framework innovatively incorporates an orthogonal Softmax layer as the network's classifier, thereby maximizing the distinction between training and test samples, optimizing the network's architecture, and achieving structural simplification. Li et al. [16] proposed a fault diagnosis method which converts raw vibration signals into two-dimensional time-frequency images and extracts deep features using KANs-CNN. Then, the FAN module aggregates features from multiple levels, and data generation through diffusion networks addresses the small sample issue. Fu et al. [17] proposed a few-shot fault diagnosis method based on Deep Auto-Encoder (DAE) and transfer learning. This approach pre-trains a DAE on a normal sample corpus, then applies SVM classification to fault samples via transfer learning, as proven effective on a civil aviation dataset. Despite these advancements, which have enhanced diagnostic accuracy in scenarios with small samples, there remain certain challenges. For instance, data augmentation may introduce noise and other negative effects during the generation process. Transfer learning not only relies on a large amount of source domain data but also requires that the data distribution difference between the source and target domains should not be too large. Meanwhile, meta-learning requires too many computational resources.

Compared with the above methods, the Siamese network, as a comparison network, known for its simple architecture and ease of training, is extensively applied in fields such as image processing [18], fingerprint recognition [19], target tracking [20]. During training, two samples are randomly selected and input into the network for training, when the number of training samples is n , a total of C_n^2 times of effective training for the network, so in the case of small samples, the Siamese network can expand the number of training times of the network, which effectively solves the problem of network underfitting due to the lack of training samples [12]. However, the utilization of Siamese networks also entails a suite of considerations that require resolution:

(1) Common Siamese networks mostly use CNNs as their subnetworks, which are limited by the influence of the receptive field; CNNs can only extract local features and cannot focus on the global information of the input samples.

(2) The Siamese network is a comparison network; its output is the degree of similarity of the two input samples. Consequently, when confronted with multiclass classification tasks, the samples to be tested need to be compared with all types of samples one by one; this approach inherently diminishes classification efficiency.

In response to the challenges previously mentioned, this paper introduces a novel DT-SViT-KNN few-shot fault diagnosis method. Initially, a diesel engine DT model is constructed by integrating the mathematical model, mechanism model, and three-dimensional physical model of the Medium-speed diesel engines of 6L21/31 Marine. Then, the DT model is used to calculate fault simulation to obtain different types of fault samples. For the acquired fault samples, this paper proposes a feature extraction network that merges Siamese network architecture with ViT. Finally, an improved KNN classifier based on the attention mechanism is added on the basis of this network to improve the model classification efficiency. The specific contributions of this paper are delineated as follows:

(1) This study constructs a comprehensive DT model that integrates diesel engine mathematical, mechanical, and three-dimensional physical models, which facilitates the mapping from the physical entity to its virtual counterpart. In conjunction, a novel feature extraction network is devised by combining the Siamese network with the ViT, enabling the extraction of global multiscale features from the input data.

(2) For the issue of the model's limited discriminative capacity between similar sample categories, we designed a Weighted-Similarity loss function. This innovation introduces similarity labels and penalty coefficients, thereby enhancing the model's efficiency in distinguishing between similar and dissimilar samples. This advancement elevates the differentiation accuracy of similar and distinct sample pairs to a novel level of precision.

(3) The incorporation of an attention mechanism during the KNN search phase allows the model to account for the relevance and significance of each neighbor rather than relying solely on distance-based classification. The attention-scoring mechanism enables the model to assign greater weight to more pertinent samples, enhancing overall classification accuracy.

The rest of this paper is organized as follows: Section 2 presents the detailed theoretical information of the proposed method. Experimental validation is carried out in Section 3. Section 4 gives the conclusion.

2. Proposed Method

Figure 1 shows the overview of the fault diagnosis method proposed in this paper, which consists of diesel engine entity, mathematical model, twin model, and state identification. On the basis of the diesel engine entity, the diesel engine DT model is established. Through the diesel engine DT model for fault simulation to obtain different types of fault data, the system uses the fault diagnosis algorithm for fault diagnosis and outputs the diagnosis results.

2.1. Diesel Engine Digital Twin Model Creation

2.1.1. Mathematical Model Construction of Diesel Engine

The mathematical model of a diesel engine can be divided into the following parts by components [21]:

(1) Diesel engine body

The diesel engine dynamics equations are as follows:

$$\frac{\pi I}{30} \cdot \frac{dN_D}{dt} = M_d - M_l \quad (1)$$

$$M_d = C_0 \cdot g_{cyl} \cdot \eta_{cy} - M_f \quad (2)$$

where M_f represents the friction loss of the diesel engine, η_{cy} represents the combustion efficiency of the diesel engine, and N_D is the rotational speed of the diesel engine.

(2) Fuel Injection Pump

The characteristics of the injection pump are represented by the injection volume g_{cyl} , which is a function of the displacement F_r of the throttle rack and the speed N_D of the diesel engine:

$$g_{cyl} = f(N_D, F_r) \quad (3)$$

For a multicylinder diesel engine, its calculation can be calculated separately for each cylinder injection, and then its average value can be taken. However, the calculation is more complicated, so this paper gives an approximate formula according to the literature [22]:

$$\tau = 15/N_D \quad (4)$$

(3) Intake air flow

The intake air flow \dot{m} consists of 2 parts, the cylinder closed air flow \dot{m}_{tr} and the swept air flow \dot{m}_{SC} .

$$\dot{m} = \dot{m}_{tr} + \dot{m}_{SC} \quad (5)$$

$$\dot{m}_{tr} = V \cdot N_{cy} \cdot \frac{N_D}{120} \cdot \eta_v \cdot \frac{P_I}{RT_I} \quad (6)$$

$$\dot{m}_{SC} = \frac{A_v \cdot P_I}{\sqrt{T_I}} \left\{ \frac{2K_1}{K_1 - 1} \frac{1}{R} \left[\left(\frac{P_E}{P_I} \right)^{\frac{2}{K_1}} - \left(\frac{P_E}{P_I} \right)^{\frac{K_1+1}{K_1}} \right] \right\}^{\frac{1}{2}} \quad (7)$$

where η_v is the inflation coefficient, A_v is the swept volume coefficient, N_{cy} is the number of diesel cylinders, P_I , T_I denotes the intake pressure and temperature, and P_E denotes the pre-turbine pressure.

(4) Intercooler

Diesel intake air temperature:

$$T_I = T_C(1 - \varepsilon) + \varepsilon T_W \quad (8)$$

where ε is the efficiency of the intercooler, and T_W is the cooling water temperature.

Diesel intake pressure:

$$P_i = P_c - \Delta P_{cool} \quad (9)$$

where ΔP_{cool} is the intercooler pressure loss.

(5) Pre-turbine temperature

The pre-turbine temperature T_E (diesel exhaust temperature) determines the amount of exhaust energy, which is a function of the overall air-fuel ratio A_{F2} and the rotational speed:

$$T_E = \Delta T_E + T_i \quad (10)$$

$$\Delta T_E = \frac{x}{1 + A_{F2}} \cdot \frac{\omega \eta_{cy}}{C_{pex}} + f(N_D) \quad (11)$$

where ω is the low calorific value of the fuel and x is the proportion of the energy carried into the exhaust pipe by the exhaust gas to the amount of fuel injected into the cylinder.

(6) Turbocharger

$$\frac{\pi I_{tc}}{30} \cdot \frac{dN_{tc}}{dt} = M_T \eta_{MT} - \frac{M_C}{\eta_{MC}} \quad (12)$$

where η_{MT} and η_{MC} denote the mechanical efficiencies of the turbine and compressor, respectively.

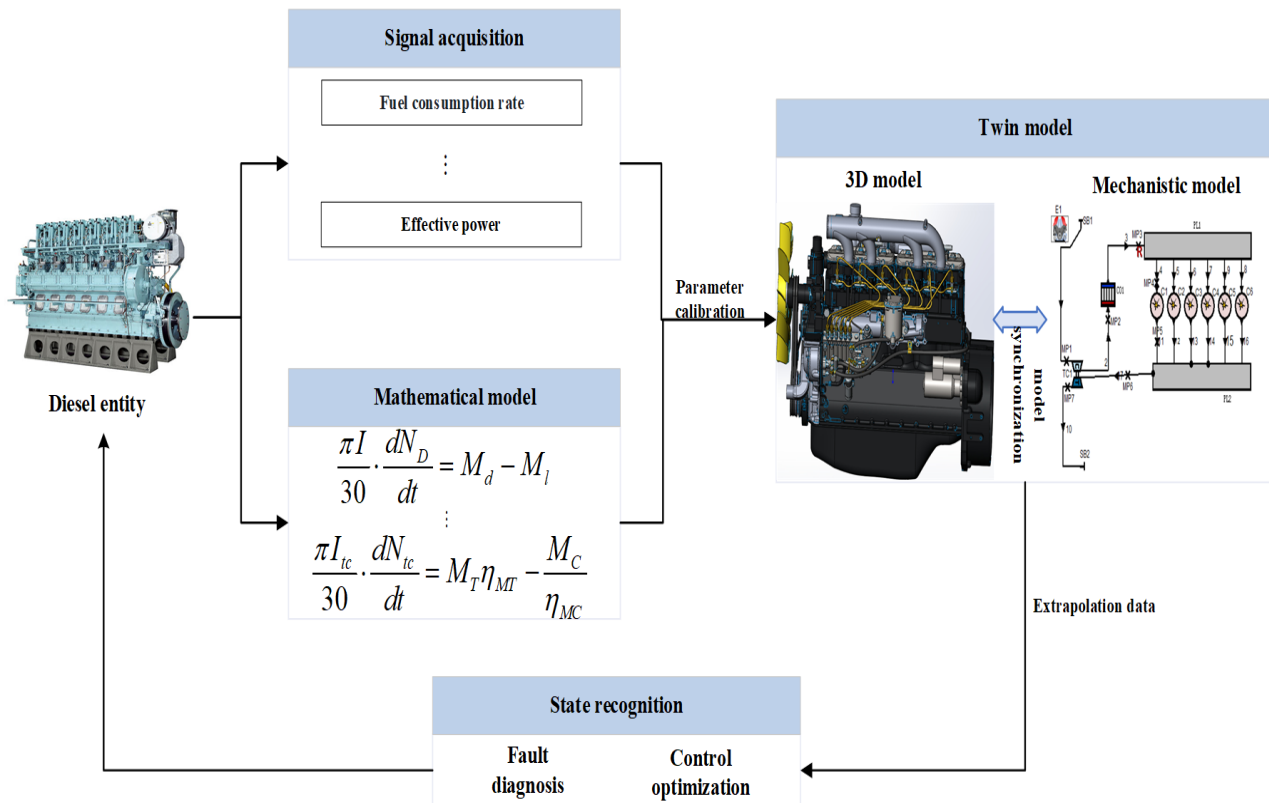


Figure 1. Composition of diesel engine digital twin system.

2.1.2. Diesel Engine Simulation Model Construction

In this study, the diesel engines of 6L21/31 Marine are modeled by utilizing the specialized diesel engine modeling software AVL-Boost R2020. According to the structural composition and parameters of the diesel engine shown in Table 1. The corresponding modules within the AVL-Boost software are used to represent each subsystem [23–25], and the simulation model of the whole engine is obtained as shown in Figure 2.

Table 1. Main parameters of 6L21/31 marine diesel engine.

Items	Parameters
Number of cylinders	6
Stroke count	4
Bore (mm)	210
Piston stroke (mm)	310
Compression ratio	15.5
Rotation speed (r/min)	900
Cylinder displacement (dm ³)	10.73
Mean effective pressure (bar)	24.6
Fuel consumption rate (g/(kW · h))	195
Maximum combustion pressure (bar)	200
Rated power (kW)	1290

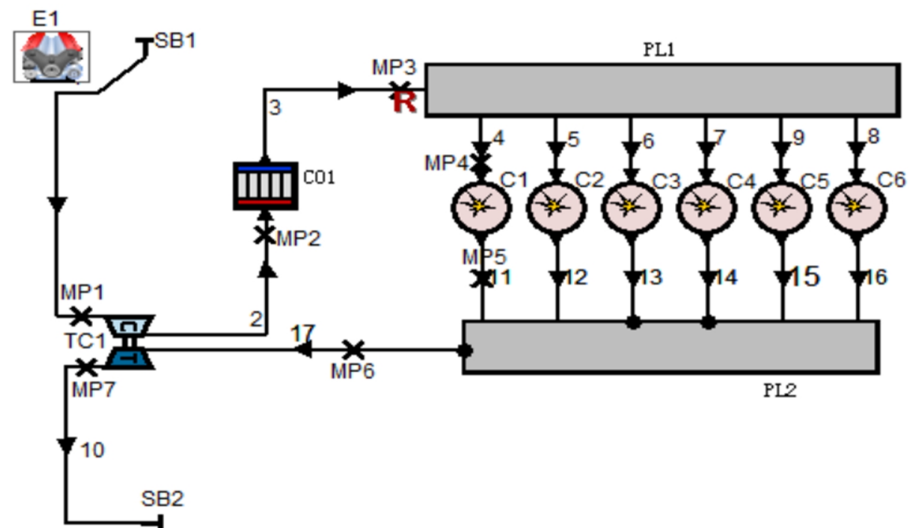


Figure 2. Simulation schematic of the 6L21/31 marine diesel engine.

2.1.3. Diesel Engine 3D Model Construction

In this section, we use SolidWorks 2017 modeling software to complete the visual modeling of the diesel engine with reference to the relevant data of the diesel engines of 6L21/31 Marine, and on this basis, we use the Animator plug-in to complete the simulation of the working process of the diesel engine [26]. The specific modeling steps are as follows:

- (1) Obtain the information of each part of the diesel engine and establish the three-dimensional model of each part of the diesel engine;
- (2) Correct the dimensions of each part model to meet the fit requirements;
- (3) Determine the assembly order and assembly relationship of each component, and complete the assembly design;
- (4) Complete the assembly work, and carry out the diesel engine operation simulation.

The diesel engine visualization results are shown in the following Figure 3:

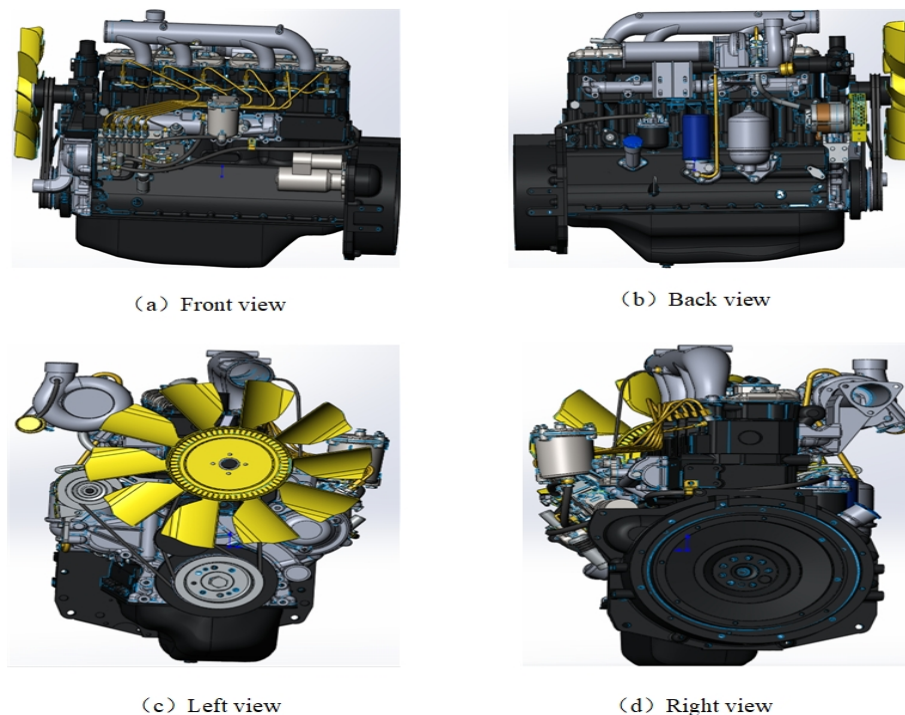


Figure 3. Twin model of 6L21/31 marine diesel engine.

2.1.4. Validation of Model Validity

To substantiate the precision of the developed model further, an assessment of the DT model's effectiveness is conducted. This involves simulating the diesel engine under three distinct operational conditions: full load (100%), three-quarter load (75%), and half load (50%). The DT model's performance is stabilized at each load level, and the corresponding outcomes are documented. Subsequently, the simulated results from the DT model are compared against the actual data. The comparative analysis is presented in Table 2, illustrating the model's accuracy and reliability.

Table 2. Comparative Analysis of Deduced Data and Actual Data.

Parameters	Operating Conditions								
	100%			75%			50%		
	Real Value	Simulation Value	Errors	Real Value	Simulation Value	Errors	Real Value	Simulation Value	Errors
Fuel consumption rate g/kW · h	190.5	187.73	1.45%	186.5	184.97	0.82%	191.9	188.52	1.76%
Effective power kW	1200	1190.91	0.76%	900	888.25	1.31%	600	597.19	0.47%
Combustion pressure bar	197.4	202.35	2.51%	161.4	160.52	0.55%	120.8	117.37	2.84%
Mean effective pressure bar	23.93	24.65	3%	18.02	18.45	2.39%	12.03	12.32	2.41%

According to Table 2, the model's simulated data under the three operational states closely match the actual data, with a maximum divergence of 3%. This result confirms that the developed DT model adheres to the stringent criteria for high fidelity.

2.1.5. Fault Deduction

To address the problem of scarcity of fuel system fault data, this paper analyzes the causes of fuel system faults according to the literature [27], and then modifies the corresponding parameters in the DT model to carry out fault deduction. In this paper, based on the digital twin model, the fuel system normal state, supercharger failure, injection advance, injection lag, cooler failure, injector wear, and fuel supply pipe blockage are deduced to obtain the relevant fault data in seven health states. The specific derivation is shown in Table 3:

Table 3. Types of faults and the method of deduction.

Fault Type	Method of Deduction	Fault Label
Normal state	Normal system operation	F1
Turbocharger failure	Adjust turbocharger pressure efficiency	F2
Fuel injection advance	Adjust the injection angle	F3
Fuel injection lag	Adjust the injection angle	F4
Cooler failure	Adjust cooler efficiency	F5
Injector wear	Adjust the cylinder oil supply volume	F6
Clogged oil supply lines	Adjust the cylinder oil supply volume	F7

2.2. Key Theories

Owing to the limited information available on fault characteristics within small samples, coupled with the propensity of traditional neural networks to underfitting, a Siamese network architecture is used to expand the number of training. Subsequently, a ViT encoder is utilized as a subnetwork to extract global multiscale features from the input data, facilitating deep learning processes. The fault features thus extracted are then integrated into an improved KNN algorithm for fault classification. This section delves into the theoretical underpinnings of Siamese networks, the ViT encoder, and the KNN classifier.

2.2.1. Siamese Networks

The idea of siamese network is to compare 2 groups of input samples, its structure is shown in Figure 4, the model firstly uses 2 subnetworks with shared weights to extract the features of the input sample (X_1, X_2) , then calculates the distance between the 2 groups of features in the feature space, and measures the similarity of the 2 groups of samples by the distance between the sample features and outputs them [28]. When the 2 samples belong to the same category, the more the similarity label $Y_{i,j}$ of the model output tends to 1, and $Y_{i,j}$ tends to 0 if the 2 samples belong to different categories.

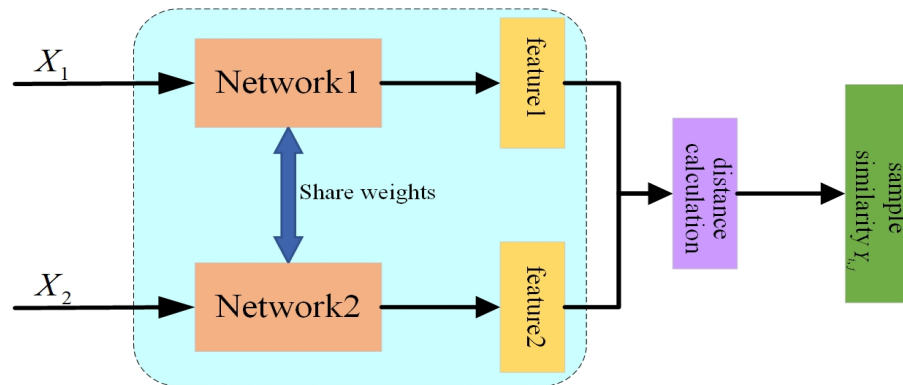


Figure 4. Structure of the siamese network.

2.2.2. Vision Transformer Encoder

The encoder is the main component of the Vision Transformer, which consists of a stack of N encoders with similar structures. Each encoder is in turn mainly composed of Multi-Head Self-Attention (MSA) and MultiLayer perceptron (MLP) layers [29,30]. Among them, MSA, as the core component of the encoder, can compute the attention weights of the input vectors through multiple dimensions to extract multifaceted features, which has excellent global feature extraction capability. MSA consists of multiple single-head self-attention and the single-head self-attention mechanism is computed as shown below:

$$\begin{cases} Q = W_q \cdot T \\ K = W_k \cdot T \\ V = W_v \cdot T \end{cases} \quad (13)$$

$$Attention(Q, K, V) = \text{soft max}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (14)$$

where Q, K, V are three different weight matrices obtained by three different linear transformations of the input vector T , $\sqrt{d_k}$ denotes the dimensionality of Q, K, V , and $Attention(\cdot)$ is the self-attention computation.

MSA as an improved version of the single-head attention mechanism consists of self-attention, which allows the model to focus on several different locations from different subspaces to capture different levels of features and information, and finally the outputs of the h single-head attention mechanism are connected and then fused by the linear mapping weight matrix W^0 to obtain the final feature information, which is shown as follows:

$$head_i = Attention(Q_i, K_i, V_i) \quad i = 1, 2 \dots h \quad (15)$$

$$Multihead(Q, K, V) = Concat(head_1, \dots, head_h)W^0 \quad (16)$$

where $head_i$ is the output of h single-headed self-attentive mechanisms and $Concat(\cdot)$ is the splicing operation.

2.2.3. K-Nearest Neighbor Classification Algorithm

The KNN algorithm is one of the most widely used classification algorithms proposed by Cover and Hart [31], with the advantages of fast computation speed and no need for training. Its principle is to find the K training samples in the feature space closest to the sample to be tested. If the vast majority of the K samples belong to a certain category, it is assumed that the sample to be tested also belongs to this category. The steps of KNN classification are as follows:

- (1) Select the parameter K .
- (2) Calculate the distance between the sample to be tested and all known samples.
- (3) Select the K samples closest to the sample to be tested.
- (4) Output the sample category to be tested as the most numerous category among the K nearest samples according to the voting law of minority to the majority.

2.3. Fault Diagnosis Method Based on SViT-KNN

2.3.1. Data Preprocessing

When a fuel system operates, it inevitably generates various noises, such as mechanical noise and combustion noise. The noise will make the already limited data contain more misleading information with small samples, thus affecting the generalization ability of the model [32]. Especially in ViT encoder, the attention mechanism may be more sensitive to the noise in the samples, and the noise may be extracted as an important feature in the self-attentive computation, affecting the classification of the subsequent model. So, it is meaningful to denoise the samples. In this paper, the traditional wavelet threshold denoising method is used, which has sound localization and multiresolution characteristics and can effectively distinguish between data and noise; the steps are as follows:

- (1) Wavelet decomposition: the noise-containing signal $x(t)$ is decomposed to different frequencies by wavelet transform with the following equation:

$$W_k = DWT(x(t), \phi_k) \quad (17)$$

$$W_k = A_j + \sum_{k=1}^j D_k \quad (18)$$

where W_k denotes the decomposed wavelet signal, DWT denotes the discrete wavelet transform, ϕ_k is the wavelet basis function, A_j is the low-frequency part of layer j , and D_k is the high-frequency part of layer k . In this paper, ϕ_k chooses haar wavelet, and the number of decomposition layers k is taken as 3.

- (2) Threshold processing: After the signal has undergone wavelet transformation, the distribution of the signal and noise on the spectrum is often different. The signal portion usually appears at low frequencies, while the noise primarily appears at high frequencies. An appropriate threshold is selected to filter out the abnormal frequency parts of the signal by utilizing this characteristic. The formula is as follows:

$$\hat{W}_k = \begin{cases} W_k & \text{if } |W_k| > \lambda \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

where \hat{W}_k is the denoised wavelet signal, λ is the threshold, the Sqtwolog threshold is selected as the wavelet threshold.

- (3) Wavelet reconstruction: the thresholded wavelet signal is de-noised by inverse transformation reconstruction, and the reconstructed signal $\hat{x}(t)$ is obtained by the following equation:

$$\hat{x}(t) = IDWT(\hat{A}_j, \{\hat{D}_k\}_{k=1}^j) \quad (20)$$

where $IDWT$ denotes the discrete wavelet inverse transform and $\hat{x}(t)$ is the denoised data.

To further improve the data quality under small sample sizes, the denoised data are finally subjected to maximum and minimum normalization to eliminate the influence of data dimensions. The resulting data are used as input for subsequent models.

2.3.2. SViT Feature Extraction Network

In contemporary Siamese network designs, CNNs are predominantly used as subnetworks for feature extraction, with these networks adept at capturing local features of the input data via convolutional kernels of varying dimensions. Nevertheless, the constraints imposed by the receptive field size and the depth of the CNN model impede its ability to attend to the global features within the input data effectively. In contrast, the ViT encoder, leveraging its intrinsic attention mechanism, demonstrates a superior capacity to concentrate on the global characteristics of the input samples [33–35]. To concurrently extract both local details and global information from the data, a more sophisticated architecture termed the Siamese Vision Transformer is introduced. This architecture is integrated with an improved KNN classifier to facilitate the automatic classification of input samples. The structure of the proposed model is delineated in Figure 5, with the detailed parameters enumerated in Table 4.

After preprocessing, samples $X = [x_1, x_2, \dots, x_{11}]^T$ and $Y = [y_1, y_2, \dots, y_{11}]^T$ are transmitted to the SViT feature extraction network for the extraction of pertinent features. Within this network, the encoder embedding module translates the input samples into vectors of dimension 11×1 , each encapsulating a set of thermal parameters relevant to the diesel fuel system. Subsequently, the model deploys a multihead self-attention mechanism to compute the self-attention of the input vectors concurrently. This mechanism enables each self-attention unit to concentrate on distinct aspects of the input vectors, facilitating the capture of a hierarchy of features and information. The output of the multihead self-attention is then directed to the multilayer perceptron layer, where it undergoes nonlinear transformations and mappings to yield the final encoder output. This output comprises a rich representation of the input data across multiple dimensions. Finally, the extracted features from the encoder are transmitted to the KNN classifier for automatic classification.

Table 4. Main parameters of the model.

Network	Model Parameters	Value
Feature extraction network	Input data dimensions	11×1
	Batch size	64
	Learning rate	0.001
	Transformer code block	2
	Number of multihead attention	8
	Embedding dimensions	512
Classifier	Hidden dimensions	512
	Nearest neighbor K	1

The Transformer encoder relies on its multiattention mechanism to extract features from the input data, which focuses on both global and local features of the data and makes the model better understand the correlation and importance of different features by calculating the attention weights of the features in each dimension, which effectively avoids the shortcomings of the local feature extraction and provides a convenient method of subsequent data classification.

2.3.3. Weighted-Similarity Loss Function

Due to the complexity and diversity of fuel system faults, the values of thermal parameters of different types of faults may be closer. Traditional loss functions, during the training of models, primarily concentrate on the relationship between model outputs and labels,

neglecting the connections between sample categories. Leading to suboptimal performance when classifying samples of similar categories. This study introduces a novel Weighted-Similarity loss function to bolster further the model's capability to discern between identical category samples. By incorporating similarity labels and penalty coefficients, the model can effectively distinguish between samples belonging to similar categories. This enhanced loss function significantly elevates the precision in distinguishing between similar and dissimilar sample pairs, with its composition comprising two distinct components: similarity loss and penalty loss.

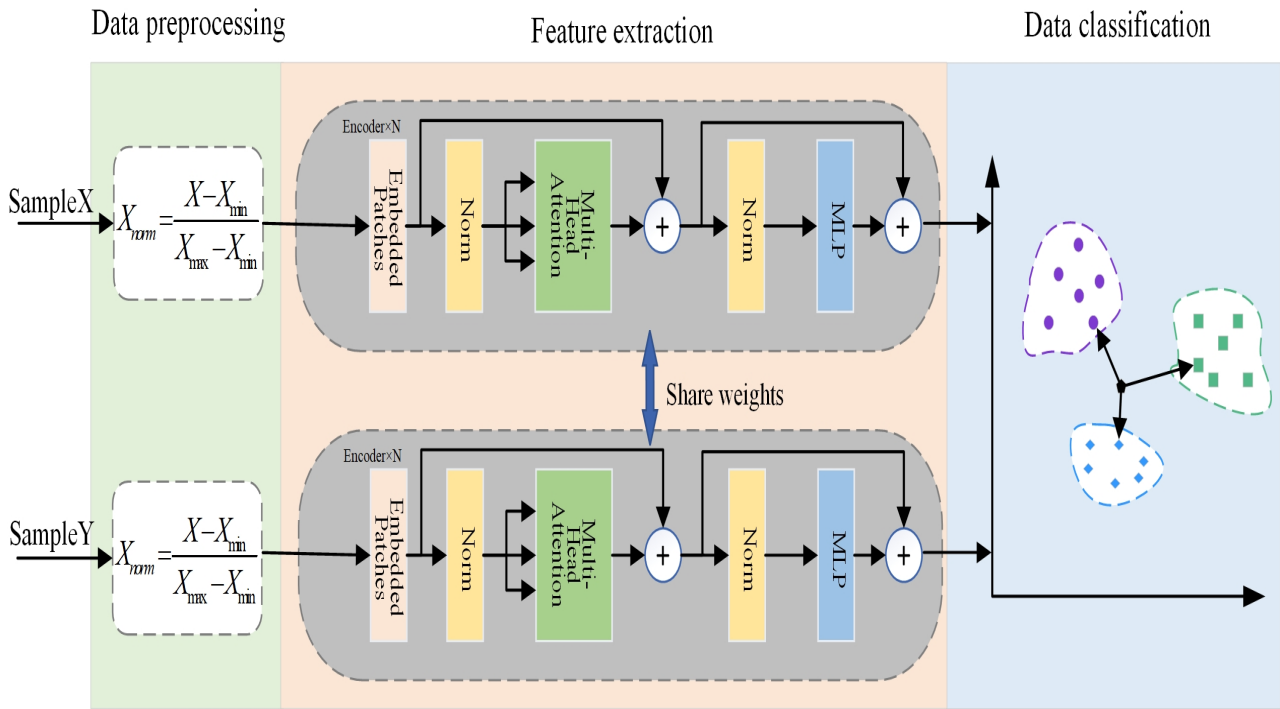


Figure 5. Structure of SViT-KNN model.

(1) Loss of similarity: for pairs of similar samples (the similarity label $Y_{i,j}$ is 1), the loss function calculates the square of the Euclidean distance, which encourages the model to reduce the distance between the features of the similar samples by the following formula:

$$L_S = \sum_{i,j=1, i \neq j}^N [Y_{i,j}(1 - R_{i,j})^2] \quad (21)$$

(2) Penalty Loss: For dissimilar sample pairs (the similarity label $Y_{i,j}$ is 0), the loss function calculates a penalty loss. The penalty coefficient is determined based on the difficulty of distinguishing between different categories of sample pairs. The penalty loss incentivizes the model to augment the feature distance between samples of distinct categories beyond a predefined threshold, as per the following formula:

$$L_S = \sum_{i,j=1, i \neq j}^N [\alpha_{i,j}(1 - Y_{i,j})(0 - R_{i,j})^2] \quad (22)$$

Integrating the aforementioned two components, the final formula for the Weighted-Similarity loss function is as follows:

$$L_S = \sum_{i,j=1, i \neq j}^N [Y_{i,j}(1 - R_{i,j})^2 + \alpha_{i,j}(1 - Y_{i,j})(0 - R_{i,j})^2] \quad (23)$$

where $Y_{i,j}$ denotes the similarity label between the two samples and $\alpha_{i,j}$ denotes the penalty coefficient when sample i and sample j belong to different fault types. When the two samples belong to different categories, if the $R_{i,j}$ value of the network output is not close to zero, a larger $\alpha_{i,j}$ value is assigned to increase the loss. During the training process the loss value can motivate the model to learn more accurately the subtle differences between pairs of samples with similar categories. After determining the loss function the model is trained using the gradient descent method, the goal of training is to make the samples of the same category as close to the feature distance as possible, while the samples of different categories are as far away from the feature distance as possible, so as to significantly improve the accuracy of the model to classify samples of similar categories.

2.3.4. Improved KNN Classifier

Traditionally, the KNN algorithm classifies samples solely based on their distances to all training samples, which are computed across all features. Among these features, some are strongly correlated with the classification task, others are weakly correlated, and others are entirely unrelated [36]. In light of this, this paper introduces an attention mechanism within the KNN search process. This mechanism allows the model to extend beyond mere distance-based classification, leveraging the attention mechanism to weigh the neighbor voting in the KNN framework impartially. By doing so, the model moves away from a simplistic majority voting approach, thereby enhancing the accuracy of the classification process. The operational steps are outlined as follows:

(1) Find the k nearest neighbors $\{x_1, x_2, \dots, x_k\}$ of the sample x to be tested based on the Euclidean distance with the following equation:

$$d_f^2(x_1, x_2) = \sqrt{\sum_{j=1}^n (f_j(x) - f_j(x_i))^2} \quad (24)$$

where f_j is the number of features for j samples and n is the total number of features.

(2) Calculate the weight α_i of each neighbor x_i on x through the attention mechanism, where the Key, Value, Query relationship is as follows. First, use the dot product of data sample K and query Q to calculate the similarity between the samples with the following formula:

$$\text{Similarity}(x_i, x) = Q^T K_i \quad (25)$$

Then, the obtained similarity is normalized to obtain the attention coefficient with the following formula:

$$s_i = \text{soft max}(\text{Similarity}(x_i, x)) \quad (26)$$

Finally, the weight α_i of each nearest neighbor x_i to x is obtained by the dot product operation of the attention coefficients and the attention value Value , which is calculated as follows:

$$\alpha_i = \sum_{i=1}^k s_i \cdot x_i \quad (27)$$

(3) These weights are then used to calculate the weighted voting result V , which is calculated as follows:

$$V = \sum_{i=1}^k \alpha_i \cdot l_i \quad (28)$$

where l_i is the category of the i neighbor.

Finally, the category of x is determined based on the weighted vote V . The highest weighted vote is chosen as the category with the following formula:

$$\hat{y} = \arg \max_l V_l \quad (29)$$

where \hat{y} is the final category label, which is the maximum value chosen based on the weighted vote V .

2.3.5. Overall Diagnostic Process

Figure 6 illustrates the comprehensive diagnostic workflow for the ship's fuel system under conditions of few available samples. The detailed steps are as follows:

(1) Data Acquisition: fault simulation is conducted to generate fault data utilizing the established diesel engine DT model.

(2) Data Preprocessing: raw data are standardized by normalizing the maximum and minimum values to eliminate the influence of the scale.

(3) Model Training: Preprocessed data are segmented into training and testing subsets. From the training set. Two samples are randomly selected from the training set to form pairs, which are then input into the Siamese network. The loss function is computed by Equation (19), and the model's parameters are updated using the gradient descent method. Training concludes when the model's accuracy meets the predefined threshold or the maximum number of training iterations is reached, with the model being saved at this juncture.

(4) Fault Diagnosis: Input sample pairs are randomly generated from the test set and fed into the trained model. Following feature extraction, the classifier autonomously generates diagnostic outcomes.

(5) Output the diagnostic results and troubleshoot the diesel engine based on the results.

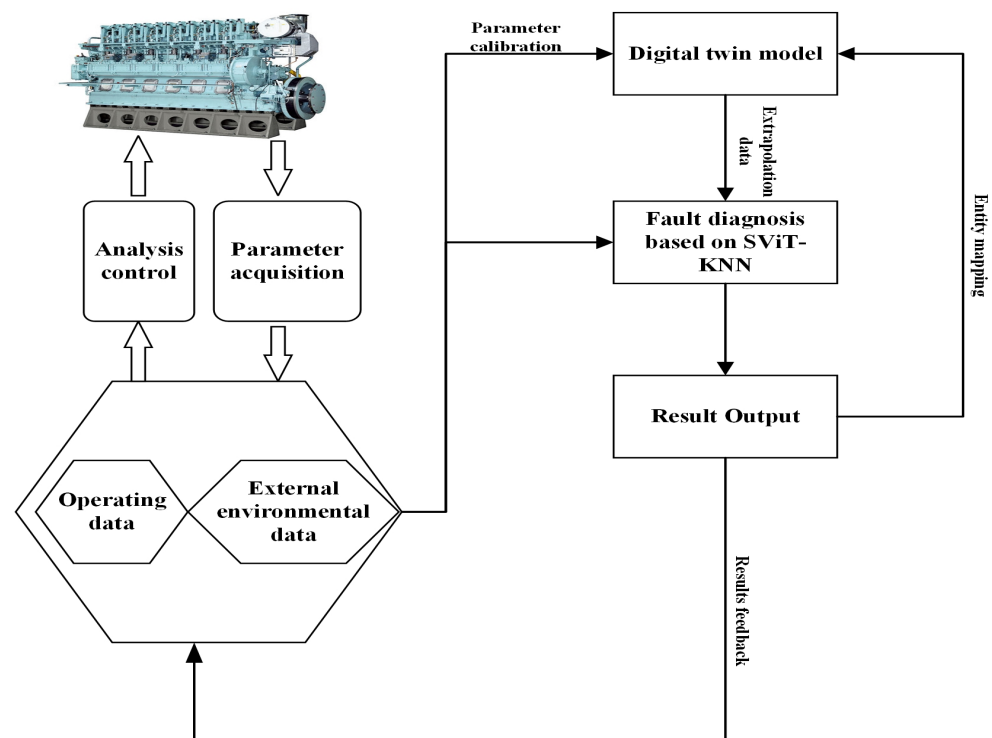


Figure 6. Overall Diagnostic Flowchart.

3. Experimental Validation

In this section, ablation and comparison experiments are set up, and the results are analyzed to demonstrate the effectiveness of the proposed method. All the experiments are run in the same environment, the deep learning framework is built by Pytorch 2.0.1, and the computers are configured as follows: the GPU is NVIDIA GeForce RTX 3060, and the CPU is Intel Core i7 12700H. Three sets of experiments are conducted, with the number of samples per class participating in the network training, n , set at 10, 20, and 30 for each set of experiments, respectively. The model's accuracy, precision, recall, and F1 score are selected as the evaluation indexes. Each group of experiments is carried out 10 times, respectively, to ensure the reliability of the experimental results, and the final results are averaged for comparison.

3.1. Ablation Experiments

This section validates the effectiveness of the method proposed in this study by adding specific steps to the Siamese Convolutional Neural Network (SCNN) model from the literature [37] and comparing the experimental results.

(1) SCNN, which uses Siamese Convolutional Neural Networks for feature extraction and FC layer-based classification, and loss function using traditional cross-entropy loss function.

(2) SViT-KNN, using ViT encoder instead of CNN for feature extraction, improved KNN for classification and loss function using traditional cross-entropy loss function.

(3) The diagnostic model of the method proposed in this paper uses SViT-KNN, and the loss function uses the Weighted-Similarity loss function. The SCNN model structure and parameters are shown in Figure 7:

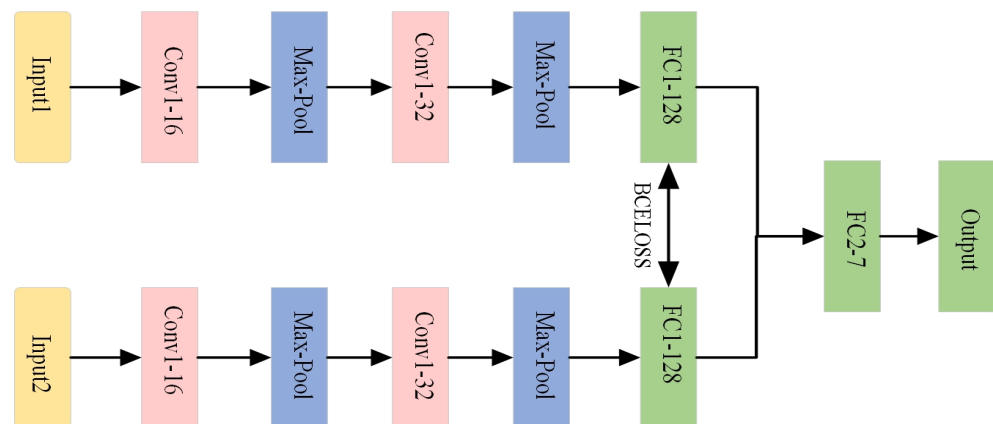


Figure 7. Structure of SCNN model.

The experimental outcomes are presented in Table 5, alongside Figures 8 and 9, which collectively assess the performance of the three methods across varying sample sizes using metrics such as accuracy, precision, F1 score, and recall.

Observations indicate that the SCNN model achieves an accuracy of 94.6% when the sample count is substantial. However, as the number of samples engaged in training diminishes, the model's accuracy experiences a steep decline. Particularly noteworthy is the accuracy of 83.63% when the training sample count per class is reduced to 10, suggesting that with a lack of samples, the CNN-based model struggles to extract a sufficient quantity of informative features, subsequently compromising its performance.

The SViT-KNN approach adopts the SViT network to extract global multiscale data features, which are then inputted into the KNN classifier for classification. In contrast to the SCNN model across all three experiments, SViT-KNN demonstrates a marked en-

hancement in accuracy, attributed to the superior feature extraction ability of SViT over SCNN. Nevertheless, both models exhibit many misclassifications when differentiating samples within similar categories, as evidenced in Figure 9. This is a consequence of the traditional loss function's exclusive emphasis on the discrepancy between model outputs and labels, disregarding the categorical relationships between samples. Consequently, errors are prevalent when classifying samples with overlapping characteristics, thus hindering the potential for further augmentation in model performance.

Table 5. Comparison of the three methods at different sample sizes.

Sample Size n	Indexs	Methods		
		SCNN	SViT-KNN	Proposed
10	Accuracy	0.8363	0.9345	0.9722
	Recall	0.8409	0.8571	0.9836
	Precision	0.8842	0.9375	0.9843
	F1	0.7852	0.8955	0.9836
20	Accuracy	0.9098	0.9632	0.9821
	Recall	0.9208	0.9231	0.9816
	Precision	0.9328	0.9767	0.9837
	F1	0.9172	0.9492	0.9815
30	Accuracy	0.9460	0.9768	0.9913
	Recall	0.9595	0.9320	0.9895
	Precision	0.9632	0.9716	0.9897
	F1	0.9585	0.9514	0.9895

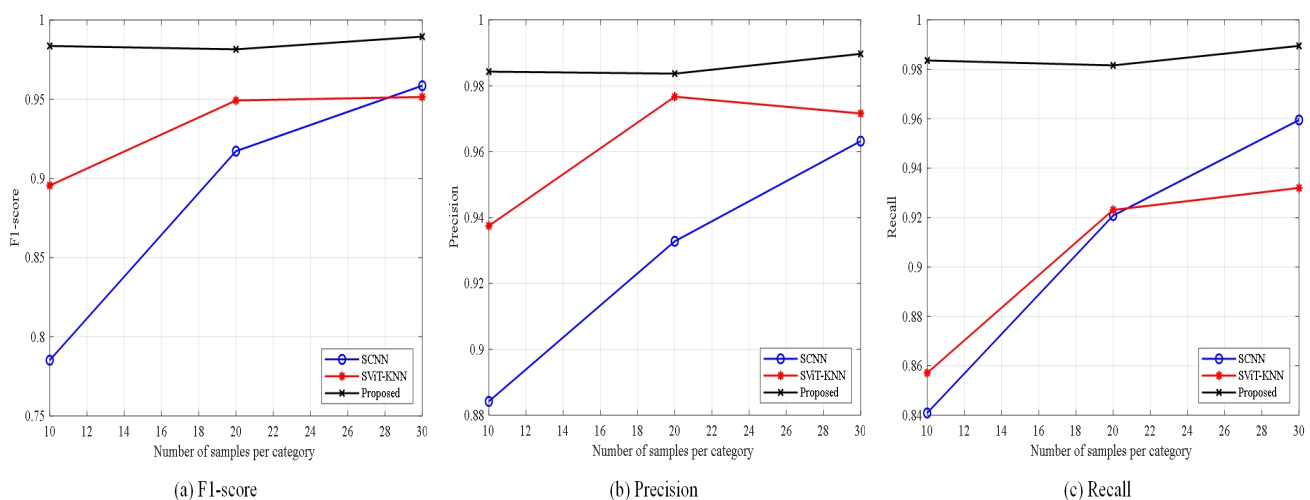


Figure 8. Experimental results of the three methods with different sample sizes.

The methodology introduced in this paper uses the SViT-KNN fault diagnosis model, augmented by the Weighted-Similarity loss function outlined in Section 2.3.3 for training. From Figures 8 and 9, the proposed method in this paper not only performs stably but also has the highest accuracy, precision, F1 score, and recall in the three sets of experiments. Remarkably, even when the training sample count is reduced to 10, the method maintains an accuracy of 97.97%. Furthermore, the examination of Figure 9c,f,i illustrates a marked enhancement in the model's capability to discern between samples of similar categories, underscoring the efficiency of the loss function devised in this paper in bolstering diagnostic accuracy.

In summary, the ablation experiments (1) and (2) substantiate that the integration of the Siamese network and ViT as a feature extractor outperforms SCNN when training samples are scarce. The cumulative evidence from (1), (2), and (3) corroborates that the loss function designed in this paper surpasses the traditional cross-entropy loss function in differenti-

ating similar categories of samples, thereby enhancing the accuracy rate. Consequently, the methodologies and steps outlined herein are indispensable and efficacious.

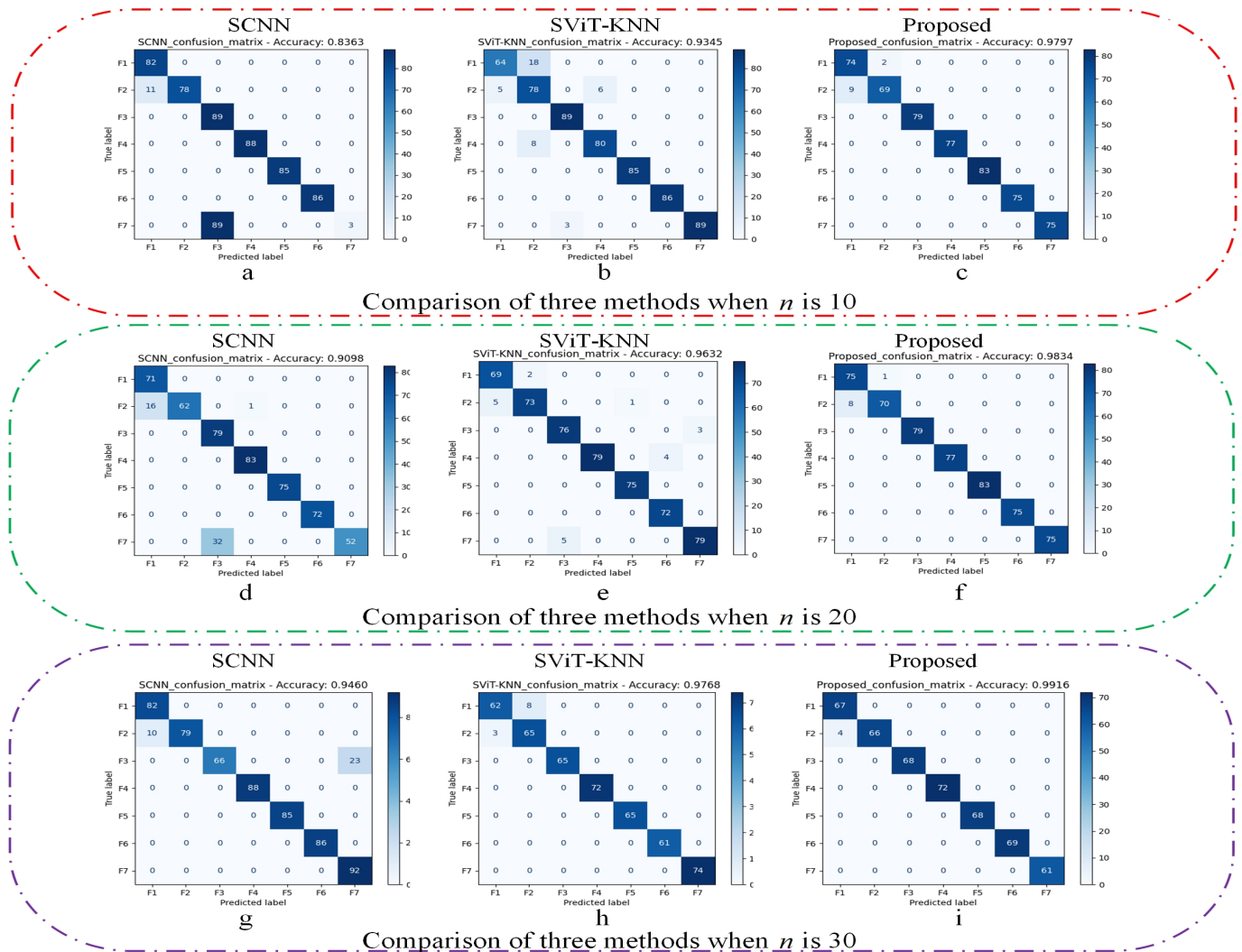


Figure 9. Confusion matrix of the three methods for different sample sizes.

3.2. Comparative Experiments

This section presents a comparative analysis to substantiate the proposed method's superiority in the context of few-shot fault diagnosis, aligning the method with 1D-CNN classical methods, as well as a method outlined in [15,38].

(1) 1D-CNN is a classical deep learning network; the specific structure and parameters of the 1D-CNN model selected in this paper are shown in Figure 10.

(2) DCSN-DRN is another kind of Siamese network proposed in the [38] for marine engine fault diagnosis with small samples. The Siamese network was built by two Deep Residual Networks (DRNs) with shared weights, then it was used to learn discriminative information from limited fault samples. What's more, the Siamese network applies strong forces to hard samples towards their corresponding correct distribution areas. Finally, the distribution area of intra-class samples was shrunk and further improved the accuracy of the classification.

(3) OSLNet is a generalized few-shot fault classification network proposed by [15], which designs an orthogonal Softmax layer on top of CNN as the classifier of the network, so that the model maximizes the difference between samples during training and testing, and the structure of the network is simplified.

The 1D-CNN and OSLNet model structure and parameters are shown in Figure 10:

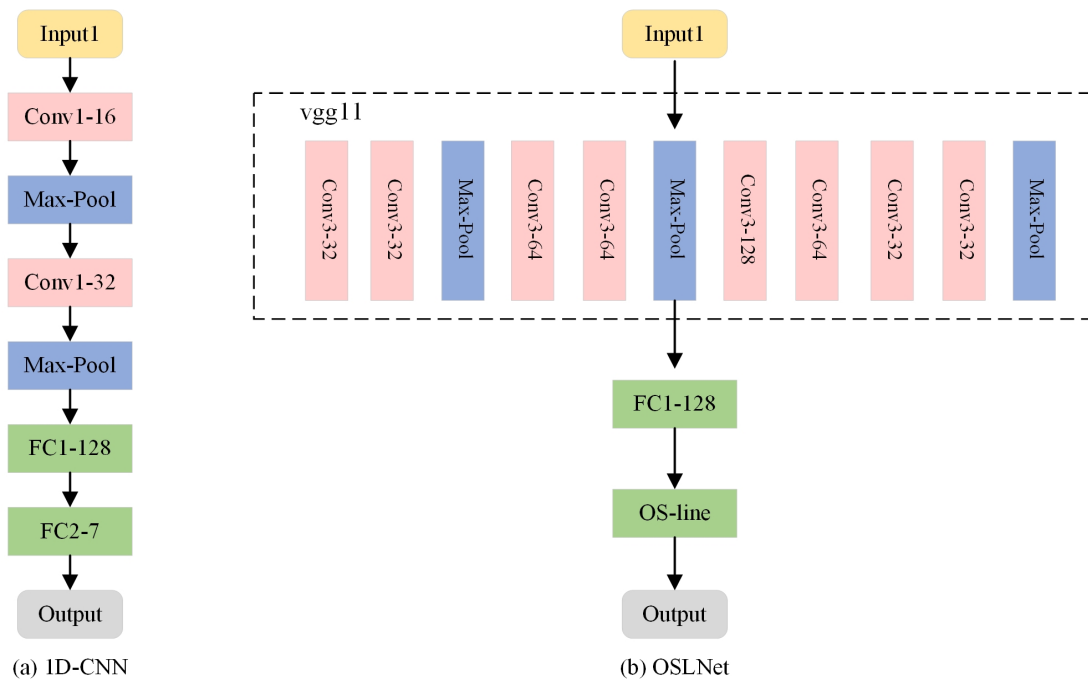


Figure 10. Structure of 1D-CNN and OSLNet models.

Table 6 and Figures 11 and 12 present the diagnostic results of each method in the comparative experimentation. When the quantity of training samples is adequate, the accuracy of all fault diagnosis approaches surpasses 90%. However, as the sample count in the training set diminishes, the diagnostic accuracy of the 1D-CNN and OSLNet networks experiences a drastic decline. Particularly, the 1D-CNN yields the poorest results, which can be attributed to its classification within the domain of traditional deep learning networks. This classification renders the 1D-CNN highly susceptible to the number of training samples, posing challenges in extracting sufficient features for learning under limited sample conditions.

Table 6. Comparison of the four methods at different sample sizes.

Sample Size n	Indexs	Methods			
		1D-CNN	DCSN-DRN	OSLNet	Proposed
10	Accuracy	0.5303	0.9023	0.6372	0.9722
	Recall	0.5303	0.8940	0.6339	0.9836
	Precision	0.3978	0.9062	0.7438	0.9894
	F1	0.4232	0.8847	0.5966	0.9836
20	Accuracy	0.8250	0.9246	0.8692	0.9821
	Recall	0.8250	0.9282	0.7238	0.9816
	Precision	0.7528	0.9335	0.6285	0.9893
	F1	0.7745	0.9261	0.9815	0.9815
30	Accuracy	0.9011	0.9676	0.9642	0.9913
	Recall	0.9937	0.9663	0.8653	0.9895
	Precision	0.9939	0.9699	0.9306	0.9917
	F1	0.9937	0.9661	0.9895	0.9895

Despite improved overall accuracy, DCSN-DRN and OSLNet show higher misclassification within similar categories, indicating a struggle to differentiate under limited samples. In contrast, the method presented in this paper demonstrates a remarkable resilience to variations in the number of training samples, achieving the highest performance across all four evaluation indexes in the three experimental groups. This comparative analysis

underscores the substantial advantage of the proposed method in the diagnosis of diesel engine fuel system faults under conditions of small sample availability.

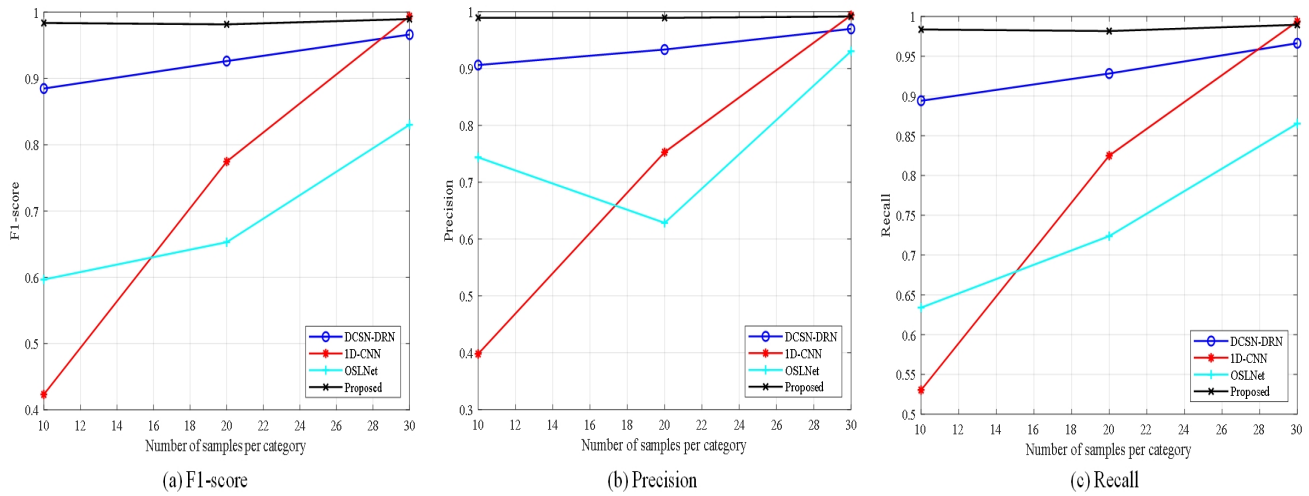


Figure 11. Experimental results of the four methods with different sample sizes.

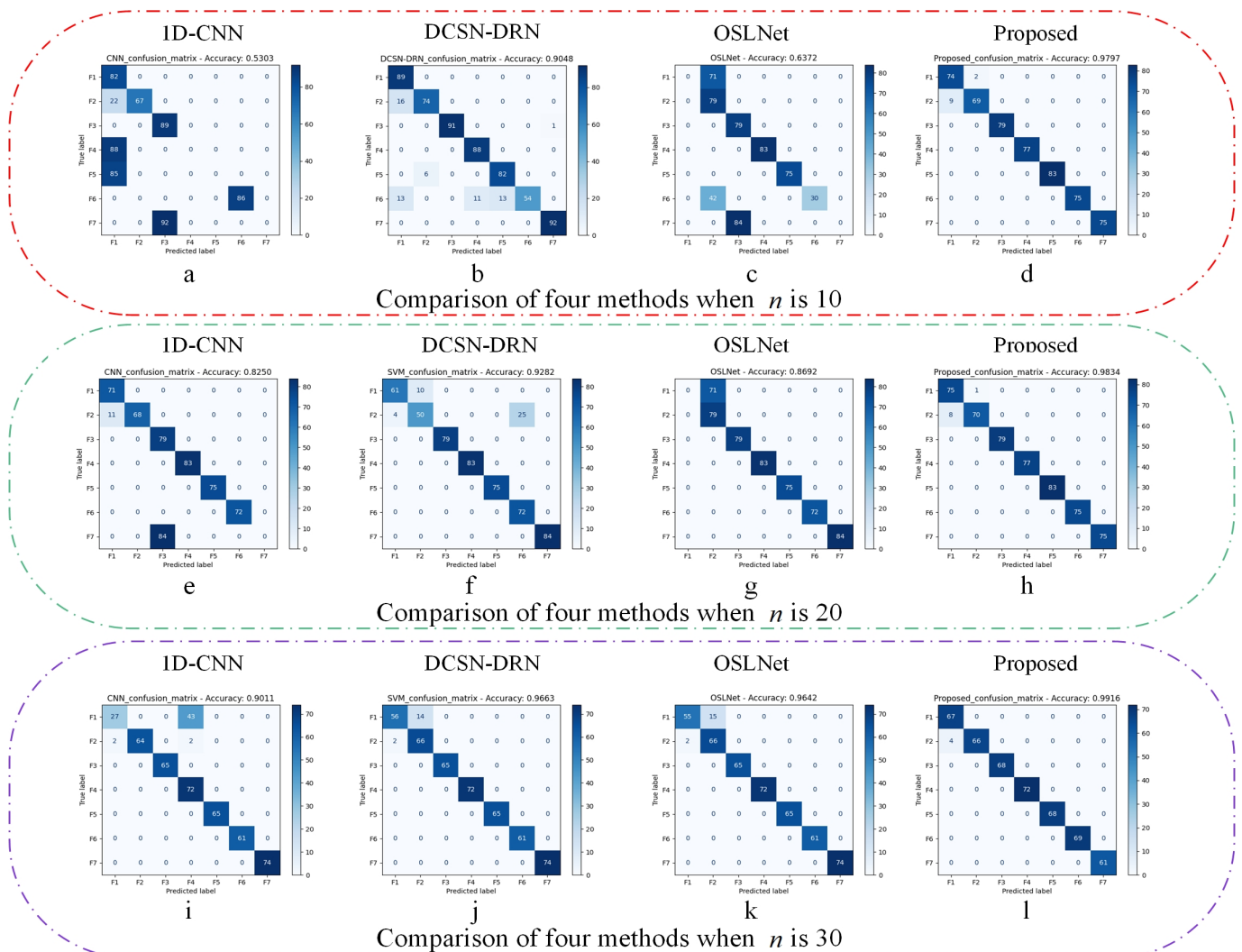


Figure 12. Confusion matrix of the four methods for different sample sizes.

4. Conclusions

This paper proposes a DT-SViT-KNN few-shot fault diagnosis method to realize accurate and efficient fault diagnosis. Firstly, a DT model of the diesel engine is established, and the fault data are obtained through model deduction. Then, a new Siamese network was designed by combining the Siamese network with ViT to extract the global multiscale features of the input data. The Weighted-Similarity loss function is also designed to improve the model's differentiation accuracy for similar samples to a new level. Finally, an improved KNN classifier based on the attention mechanism is added to the traditional Siamese network, enabling the model to classify and improve the diagnostic efficiency automatically. Ablation experiments and comparison experiments were conducted successively to verify that DT-SViT-KNN is an accurate and reliable few-shot fault diagnosis method.

However, there are still some challenges in practical application scenarios. Fault signals may be interfered with by a variety of noises, and the levels of these noises are usually unknown, which makes it difficult to select the number of wavelet decomposition layers and thresholds in wavelet denoising and affects the denoising effect. On the other hand, as the operating conditions and time of the system change, the data will become diversified, and the difference in the distribution of the same type of data will become larger, so how to carry out fault diagnosis under variable operating conditions with small samples is also a difficult problem. In future research work, we will investigate the performance of the proposed method on real data and further solve the problems encountered in the practical application of this method.

Author Contributions: Conceptualization, S.L. and L.Q.; methodology, S.L.; software, S.L.; validation, S.L., B.D. and J.S.; formal analysis, H.X.; investigation, L.Q.; resources, R.T.; data curation, D.Z.; writing—original draft preparation, S.L.; writing—review and editing, L.Q.; visualization, R.T.; supervision, L.Q.; project administration, L.Q.; funding acquisition, L.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the National Natural Science Foundation of China (grant number: 51875270) and the Zhenjiang Science and Technology Program (grant number: JC2024021).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Xu, X.; Zhao, Z.; Xu, X.; Yang, J.; Wang, G. Machine learning-based wear fault diagnosis for marine diesel engine by fusing multiple data-driven models. *Knowl.-Based Syst.* **2019**, *190*, 105324. [[CrossRef](#)]
2. Yaqiong Lv, Xueting Yang, Yifan Li, Jialun Liu, Shijie Li, Fault detection and diagnosis of marine diesel engines: A systematic review. *Ocean. Eng.* **2024**, *294*, 116798. [[CrossRef](#)]
3. Zhang, Z.; Wang, D.; Zhang, Y. Research on Fault Monitoring and Diagnosis Method Based on PSO Improved KPCA-SVM. *Mod. Manuf. Eng.* **2020**, *9*, 7.
4. Yao, H.; Zhang, X.; Guo, Q.; Miao, Y.; Guan, S. Fault diagnosis method for oil-immersed transformers integrated digital twin model. *Sci. Rep.* **2024**, *14*, 20355. [[CrossRef](#)] [[PubMed](#)]
5. Liu, C.; Chen, Y.; Xu, X.; Che, W. Domain generalization-based damage detection of composite structures powered by structural digital twin. *Compos. Sci. Technol.* **2024**, *258*, 110908. [[CrossRef](#)]
6. Xu, Z.; Wang, Z.; Gao, C.; Zhang, K.; Lv, J.; Wang, J.; Liu, L. A digital twin system for centrifugal pump fault diagnosis driven by transfer learning based on graph convolutional neural networks. *Comput. Ind.* **2024**, *163*, 104155. [[CrossRef](#)]
7. Qin, Y.; Liu, H.; Wang, Y.; Mao, Y. Inverse physics-informed neural networks for digital twin-based bearing fault diagnosis under imbalanced samples. *Knowl.-Based Syst.* **2024**, *292*, 1.1–1.13. [[CrossRef](#)]
8. Liu, W.; Han, B.; Zheng, A.; Zheng, Z. Fault Diagnosis for Reducers Based on a Digital Twin. *Sensors* **2024**, *24*, 2575. [[CrossRef](#)]

9. Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R.X. Deep Learning and Its Applications to Machine Health Monitoring: A Survey. *arXiv* **2016**, arXiv:1612.07640. [[CrossRef](#)]
10. Cao, C.; Zhang, Y. Learning to Compare Relation: Semantic Alignment for Few-Shot Learning. *arXiv* **2020**, arXiv:2003.00210. [[CrossRef](#)] [[PubMed](#)]
11. Dumoulin, V.; Houlsby, N.; Evci, U.; Zhai, X.; Larochelle, H. Comparing Transfer and Meta Learning Approaches on a Unified Few-Shot Classification Benchmark. *arXiv* **2021**, arXiv:2104.02638.
12. Wang, N.; Jin, W.; Bi, H.; Xu, C.; Gao, J. A Survey on Deep Learning for Few-Shot PolSAR Image Classification. *Remote Sens.* **2024**, *16*, 4632. [[CrossRef](#)]
13. Sunal, C.E.; Velisavljevic, V.; Dyo, V.; Newton, B.; Newton, J. Centrifugal Pump Fault Detection with Convolutional Neural Network Transfer Learning. *Sensors* **2024**, *24*, 2442. [[CrossRef](#)]
14. Li, L.; Su, T.; Ma, F. Fault Diagnosis Method for Small Sample in High-Pressure Common Rail System. *J. Intern. Combust. Engines* **2023**, *41*, 255–262. [[CrossRef](#)]
15. Li, X.; Chang, D.; Ma, Z.; Tan, Z.H.; Xue, J.H.; Cao, J.; Yu, J.; Guo, J. OSLNet: Deep Small-Sample Classification with an Orthogonal Softmax Layer. *IEEE Trans. Image Process.* **2020**, *29*, 6482–6495. [[CrossRef](#)] [[PubMed](#)]
16. Li, Y.; Gu, X.; Wei, Y. A Deep Learning-Based Method for Bearing Fault Diagnosis with Few-Shot Learning. *Sensors* **2024**, *24*, 7516. [[CrossRef](#)] [[PubMed](#)]
17. Fu, S.; Zhong, S.; Lin, L.; Zhang, Y. Fault Diagnosis of Civil Aviation Engines with Few Samples Based on Transfer Learning. *J. Comput. Integr. Manuf. Syst.* **2021**, *27*, 12.
18. Perera, S.; Allali, M.; Linstead, E.; El-Askary, H. Landuse Landcover Change Detection in the Mediterranean Region Using a Siamese Neural Network and Image Processing. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 4368–4371.
19. Zhu, L.; Xu, P.; Zhong, C. Siamese Network Based on CNN for Fingerprint Recognition. In Proceedings of the 2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI), Fuzhou, China, 24–26 September 2021; pp. 303–306.
20. Fan, L. and Kim, P. Anchor free based Siamese network tracker with transformer for RGB-T tracking. *Sci. Rep.* **2023**, *13*, 13294. [[CrossRef](#)] [[PubMed](#)]
21. Chen, H.; Ao, C. Mathematical Model of Diesel Engine in Ship Propulsion System Simulation. *Ship Eng.* **2000**, *5*, 33–37+4.
22. Zeng, F.; Chen, Y.; Wu, J.; Chen, G. Dynamic Modeling and Simulation of Marine Diesel Engine Using Elman Networks. In Proceedings of the 2003 International Conference on Neural Networks and Signal Processing, Nanjing, China, 14–17 December 2003.
23. Zhao, Z.; Zhang, J.; Luo, C.; Jia, B.; An, L. AVL BOOST-based typical fault simulation and data analysis of marine diesel engine. *Chin. J. Ship Res.* **2021**, *16*, 176–181.
24. Neshat, E.; Honnery, D.; Saray, R.K. Multi-zone model for diesel engine simulation based on chemical kinetics mechanism. *Appl. Therm. Eng.* **2017**, *121*, 351–360. [[CrossRef](#)]
25. Min, X.; Han, X.-P. Research on Co-simulation of Thermodynamics and Dynamic Analysis of Marine Diesel Engine. *J. Syst. Simul.* **2014**. [[CrossRef](#)]
26. Matulis, M.; Harvey, C. A robot arm digital twin utilising reinforcement learning. *Comput. Graph.* **2021**, *95*, 106–114. [[CrossRef](#)]
27. Liu, X.; Zeng, H.; Dong, J.; He, T.; Liu, L. Ship Diesel Engine Fault Diagnosis Based on Optimized Stacked Autoencoders. *Navig. China* **2022**, *45*, 45–51+57.
28. Wen, C.; Xue, Y.; Liu, W.; Chen, G.; Liu, X. Bearing fault diagnosis via fusing small samples and training multi-state siamese neural networks. *Neurocomputing* **2024**, *576*, 127355. [[CrossRef](#)]
29. Zhu, Q.X.; Qian, Y.S.; Zhang, N.; He, Y.L.; Xu, Y. Multi-scale Transformer-CNN domain adaptation network for complex processes fault diagnosis. *J. Process Control* **2023**, *130*, 103069. [[CrossRef](#)]
30. Hajj-Ali, Z.; Dosso, Y.S.; Greenwood, K.; Harrold, J.; Green, J.R. Depth-Based Intervention Detection in the Neonatal Intensive Care Unit Using Vision Transformers. *Sensors* **2024**, *24*, 7753. [[CrossRef](#)] [[PubMed](#)]
31. Cover, T.M.; Hart, P.E. Nearest neighbour pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [[CrossRef](#)]
32. Sun, T.; Gao, J. New Fault Diagnosis Method for Rolling Bearings Based on Improved Residual Shrinkage Network Combined with Transfer Learning. *Sensors* **2024**, *24*, 5700. [[CrossRef](#)] [[PubMed](#)]
33. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S. An Image is Worth 16 × 16 Words: Transformers for Image Recognition at Scale. In Proceedings of the International Conference on Learning Representations, ICLR 2021, Vienna, Austria, 4 May 2021.
34. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *arXiv* **2017**, arXiv:1706.03762.
35. Wang, S.; Liu, Z.; Jia, Z.; Zhao, W.; Li, Z.; Wang, L. Intermittent fault diagnosis of analog circuit based on enhanced one-dimensional vision transformer and transfer learning strategy. *Eng. Appl. Artif. Intell.* **2024**, *127*, 107281. [[CrossRef](#)]

36. Zhou, H.; Mo, L.; Liu, Y.; Wang, Y.; Qing, X. Fault Diagnosis of Aero-Engines Based on Feature Optimization and Improved KNN. *Aeronaut. Comput. Technol.* **2023**, *53*, 45–49.
37. Gao, H.; Chao, Q.; Xu, Z.; Tao, J.; Liu, M.; Liu, C. Piston pump fault diagnosis based on siamese neural network with small samples. *Beijing Aeronaut. Astronaut. Univ. J.* **2021**, *49*, 155–164.
38. Zhong, B.; Zhao, M.; Wang, L.; Fu, S.; Zhong, S. DCSN: Focusing on hard samples mining in small-sample fault diagnosis of marine engine. *Measurement* **2024**, *235*, 114929. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.