




Article

Deep Hybrid Model for Fault Diagnosis of Ship's Main Engine

Se-Ha Kim ^{1,†}, Tae-Gyeong Kim ^{1,†}, Junseok Lee ² , Hyoung-Kyu Song ^{3,4} , Hyeonjoon Moon ⁵ 
and Chang-Jae Chun ^{6,*}

¹ Department of Artificial Intelligence, Sejong University, Seoul 05006, Republic of Korea; 23114417@sju.ac.kr (S.-H.K.); ktk23114418@sju.ac.kr (T.-G.K.)

² Artificial Intelligence Laboratory, Okestro Co., Ltd., Seoul 07335, Republic of Korea; js.lee6@okestro.com

³ Department of Information and Communication Engineering, Sejong University, Seoul 05006, Republic of Korea; songhk@sejong.ac.kr

⁴ Department of Convergence Engineering for Intelligent Drone, Sejong University, Seoul 05006, Republic of Korea

⁵ Department of Computer Science and Engineering, Sejong University, Seoul 05006, Republic of Korea; hmoon@sejong.ac.kr

⁶ Department of Data Science and Artificial Intelligence, Sejong University, Seoul 05006, Republic of Korea

* Correspondence: cchun@sejong.ac.kr

† These authors contributed equally to this work.

Abstract

Ships play a crucial role in modern society, serving purposes such as marine transportation, tourism, and exploration. Malfunctions or defects in the main engine, which is a core component of ship operations, can disrupt normal functionality and result in substantial financial losses. Consequently, early fault diagnosis of abnormal engine conditions is critical for effective maintenance. In this paper, we propose a deep hybrid model for fault diagnosis of ship main engines, utilizing exhaust gas temperature data. The proposed model utilizes both time-domain features (TDFs) and time-series raw data. In order to effectively extract features from each type of data, two distinct feature extraction networks and an attention module-based classifier are designed. The model performance is evaluated using real-world cylinder exhaust gas temperature data collected from the large ship low-speed two-stroke main engine. The experimental results demonstrate that the proposed method outperforms conventional methods in fault diagnosis accuracy. The experimental results demonstrate that the proposed method improves fault diagnosis accuracy by 6.146% compared to the best conventional method. Furthermore, the proposed method maintains superior performance even in noisy environments under realistic industrial conditions. This study demonstrates the potential of using exhaust gas temperature using a single sensor signal for data-driven fault detection and provides a scalable foundation for future multi-sensor diagnostic systems.

Keywords: attention mechanism; deep learning; degradation; exhaust gas temperature; fault diagnosis; feature fusion; hybrid model; marine main engine; time-domain feature



Received: 30 May 2025

Revised: 18 July 2025

Accepted: 20 July 2025

Published: 23 July 2025

Citation: Kim, S.-H.; Kim, T.-G.; Lee, J.; Song, H.-K.; Moon, H.; Chun, C.-J. Deep Hybrid Model for Fault Diagnosis of Ship's Main Engine. *J. Mar. Sci. Eng.* **2025**, *13*, 1398. <https://doi.org/10.3390/jmse13081398>

Copyright: © 2025 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/>).

1. Introduction

1.1. Background

In modern society, ships play a pivotal role in global trade, with approximately 90% of international trade facilitated through maritime transport [1]. Among the critical components of a ship, the main engine is essential for its operation [2]. Consequently, a malfunction or failure of the main engine can significantly disrupt operations, resulting in substantial

losses. To mitigate such risks, accurate fault diagnosis of the main engine is crucial. Effective fault detection and diagnosis of the ship's main engine enable early problem identification, reducing maintenance costs, enhancing reliability, and proactively preventing accidents [3].

Recent advancements in artificial intelligence have catalyzed significant progress in fault detection and diagnosis across various industries. In the maritime domain, research on machine-learning-based fault diagnosis for main engines of ships has gained traction. Recently, with the emergence of unmanned ships, highly accurate fault diagnosis technologies have become a critical requirement. To meet these demands, deep learning algorithms have recently been adopted for fault diagnosis in core ship systems. Despite the significant advances in deep-learning-based fault diagnosis, a key challenge lies in the scarcity of fault data compared to normal data [4]. Many studies relied on simulator-generated data or datasets with insufficient fault instances, limiting the ability to account for diverse real-world abnormal conditions. To address this issue, researchers have explored techniques for synthesizing fault data.

In order to improve the fault diagnosis performance, it is crucial to address the issue of insufficient fault data, as mentioned earlier, and to enhance the feature extraction performance of the model for effectively capturing fault information in the input data.

1.2. Literature Review

In this subsection, we provide a comprehensive review of the existing literature pertaining to fault diagnosis and degradation. Our focus centers on three key themes: machine learning, deep learning, data synthesis.

1.2.1. Machine-Learning-Based Fault Diagnosis Methods

With the advancement of artificial intelligence technologies, research on fault detection and diagnosis based on machine learning techniques has been actively conducted. For instance, in study [5], diverse machine learning algorithms were applied, such as linear regression, support vector machine, and ensemble methods, using simulator-generated data. In [6], classification algorithms such as naïve Bayes, support vector machine, and simple neural networks were utilized for fault diagnosis. In a follow-up study [7], six primary causes of engine failure were identified, focusing the study on defect detection and cause classification. Cheliotis et al. [8] proposed a multi-polynomial ridge regression model to predict cylinder exhaust gas temperature using the output, speed, and exhaust gas pressure data of a ship's main engine. In [9,10] study, support-vector-machine-based fault diagnosis methods were proposed for each of the four ship systems and the ship's main engine, respectively. In [11], a fault diagnosis framework based on a support vector machine was proposed for condition-based maintenance for the oil purifier of ships. In [12], a method was proposed that employs K-Means clustering to detect faults and defects in a ship's main engine and further aims to prevent their occurrence in advance by implementing a real-time monitoring framework based on the clustering results.

1.2.2. Deep-Learning-Based Fault Diagnosis Methods

Recently, deep-learning-based approaches have been proposed to enable more precise fault diagnosis in complex ship systems. For instance, in [13], fault detection for six thrusters was conducted using one-dimensional CNN with data obtained from a simulator. In [14], a real-time anomaly detection intelligent system was proposed by estimating the normalized root mean square error matrix using LSTM-based variational autoencoder and by generating images using multi-step Otsu thresholding. In [15], a fault detection method employing auto-associative kernel regression, a nonlinear regression algorithm based on k-means clustering, was proposed. In [16], an early fault detection method based on CNN and bi-directional GRU was proposed using exhaust gas temperature data of

ships. In [17], an early fault detection algorithm was proposed by combining an attention mechanism with CNN and bi-directional LSTM architecture to predict the exhaust gas temperature of ships. An artificial-neural-network-based prediction method for the exhaust gas temperature of the main engine was proposed in [18]. In [19], a fault detection algorithm was proposed based on CNN in the ship operation domain. In [20], a fault diagnosis methodology based on CNN was proposed for the electrical power systems of ships. Ai et al. [21] proposed a multi-head graph attention neural network that combines probabilistic and rank-order similarity-based graph structures to enhance fault diagnosis performance. Senemmar et al. [22] proposed a wavelet-based graph neural network (WGNN) for non-intrusive fault detection in naval shipboard power systems, achieving high accuracy under noisy and transient conditions and demonstrating real-time applicability. Recently, transformer-based approaches have shown promise in time-series fault diagnosis by capturing long-range dependencies without recurrence. Li et al. [23] proposed synchro squeezing S-transform (SSST) and vision transformer (ViT), demonstrating that converting vibration signals into time-frequency images using SSST and classifying them with ViT enables accurate diesel engine fault status identification, achieving high accuracy. Similarly, Liu et al. [24] introduced iTransformer, a fault diagnosis and prediction model applied to marine battery management systems, enabling early detection and accurate forecasting of voltage inconsistencies using battery cluster data. While GNNs are effective for modeling structural relationships, they require manual graph construction and may face scalability issues. Transformers offer greater flexibility and scalability for long-range sequence modeling, though they typically require large datasets and computational resources.

1.2.3. Fault Data Synthesis Approach

A persistent challenge in the field of fault diagnosis is the scarcity of fault data compared to normal data [4]. To overcome this issue, numerous studies have relied on simulator-generated data or datasets with limited fault cases [13–20]. However, such approaches often fall short in capturing the complexity of diverse abnormal scenarios encountered in real-world environments. To address these limitations, recent research has focused on techniques for synthesizing fault data. For instance, in [4], a study on fault diagnosis was conducted using the collective anomalies technique [25,26] that injects noise into data during the abnormal data generation process and the method of generating abnormal data with a degradation pattern. In [27], a fault diagnosis model was designed by generating data that considers point anomalies where abnormal situations occur at multiple points and transitional occurrences where the operating state of the system changes depending on environmental factors or operating conditions. Additionally, self-supervised learning has emerged as a promising direction to mitigate the shortage of labeled fault data. Wang et al. [28] proposed a self-supervised contrastive learning framework with nearest-neighbor matching (SCLNNM) that effectively extracts discriminative features from unlabeled marine engine data, improving fault diagnosis performance under limited annotation conditions. However, self-supervised learning also presents challenges such as high computational cost, reliance on carefully designed pretext tasks, and difficulty in interpretation. Moreover, noisy or biased input data can lead to suboptimal representations, limiting generalization to unseen fault patterns.

1.3. Motivation

In recent years, considerable research has focused on improving fault diagnosis performance by designing advanced feature extraction models or applying structural transformations to input data to effectively capture fault-relevant information. In conventional studies, fault detection and time-series prediction were performed by using time-series data as the

input [13–19] and by extracting patterns for various abnormal situations utilizing image data transformed from time-series data as input [4,20,27]. In the previously mentioned studies [4,13–20,27], a model that employs time-series or image data with a single feature extractor was proposed. However, information about anomalies may be contained in the time-series raw data itself, as well as in both the image-based data and the time-domain feature (TDF) transformed from the time-series data. Specifically, time-series raw data contains micro-level information, such as temporal patterns, and image-based data, such as spectrograms, which have temporal-frequency information; TDF captures information about macro-level changes in the data, such as statistical characteristics of time-series data. Therefore, considering these diverse data types and feature extraction processes when constructing the model can help improve performance. Recently, conventional hybrid fault diagnosis models have been proposed for fault diagnosis. They utilized two types of input data such as time-frequency image data and numerical data [29] and a scalogram obtained by applying a continuous wavelet transform to the time-series data and image data transformed from the time-series data [30]. However, in the marine environment, various and complex factors such as waves, wind speed, corrosion, and high humidity affect the sensor signals of ship systems and machinery, so it is very difficult to diagnose faults of a ship's main engine. Therefore, a high-performance and robust fault diagnosis model of a ship's main engine is very important for the stable operation of ships considering this marine environment. To the best of our knowledge, there has been no work in the literature to design a hybrid model involving multiple feature extractors for fault diagnosis of a ship's main engine.

1.4. Contribution

In this paper, we propose a hybrid model with multiple feature extraction processes for ship main engine fault diagnosis. The model combines time-series data and time-domain features (TDFs), enabling it to capture both fine-grained temporal patterns and macro-level statistical characteristics. To evaluate its performance, we utilize cylinder exhaust gas temperature data collected from real-world environments. We address the issue of fault data scarcity by synthesizing fault data from normal data. Furthermore, the final training dataset is constructed by incorporating equipment degradation levels, considering the marine operating environment. The experimental results demonstrate that the proposed fault diagnosis model outperforms conventional methods. The main contributions of this paper are as follows:

- We propose a hybrid model for fault diagnosis of a ship's main engine. Since the proposed hybrid model consists of two separate feature extractors for time-series raw data and TDF, it can effectively extract features that lead to achieving high fault diagnosis accuracy.
- We analyzed the performance of the proposed model by additionally considering the environment with noise signals. We demonstrated through simulation that the performance of the proposed model is better than the existing methods even in noisy environments.
- In order to evaluate the performance of the proposed hybrid model, we created training data by simulating six main engine abnormal classes according to the degree of equipment degradation based on the actual data collected from a two-stroke ship diesel engine. We trained and verified our proposed model using the data created based on the actual collected data.

The rest of this paper is organized as follows. In Section 2, a brief theoretical background of the attention mechanism used in this paper is introduced. In Section 3, the proposed framework for fault diagnosis of a ship's main engine and the detailed structure of

the proposed model are presented. We also describe the performance of the proposed model in the presence of noise. In Section 4, the experimental setup to evaluate the performance of the proposed model is described, and an analysis of the results is provided. Finally, Section 5 concludes the paper.

2. Theoretical Background

Attention Mechanism

Humans naturally focus on specific parts of information when processing large volumes of data. Inspired by this biological system, the attention mechanism has been introduced and widely applied in artificial intelligence [31]. Originally proposed by Bahdanau et al. [32], the attention mechanism was later expanded with the introduction of the self-attention mechanism by Vaswani et al. [33], achieving remarkable success, particularly in natural language processing. Recent advancements have extended the use of attention mechanisms to enhance feature representation by incorporating channel attention and spatial attention. Channel attention assigns weights to individual channels, while spatial attention assigns weights to specific spatial regions or pixels [34]. As illustrated in Figure 1a, channel attention computes weight values for each channel from a multi-channel feature map, assigning weights accordingly. Similarly, spatial attention, depicted in Figure 1b, calculates weights for each spatial location from a multi-space feature map. This approach enables more accurate predictions by focusing on channels or spatial regions that require greater emphasis.

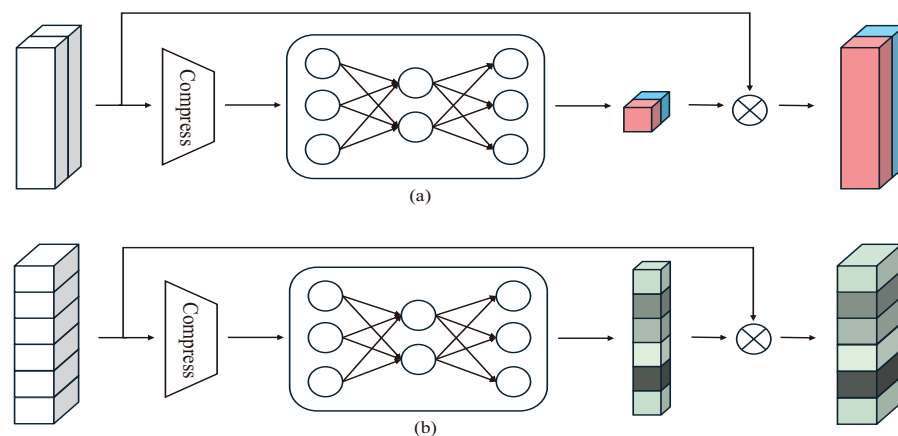


Figure 1. Attention mechanisms: (a) channel attention; (b) spatial attention.

3. Proposed Method

In this paper, we propose a deep-learning-based hybrid model for fault diagnosis of a ship's main engine. The objective of this study is to develop a diagnostic model that can detect abnormal conditions in large marine main engines using cylinder exhaust gas temperature. To achieve this, a deep learning model was trained on a large volume of time-series data collected under both normal and abnormal conditions. The model is designed to automatically identify thermal deviations from typical patterns, thereby enabling real-time anomaly detection when applied to operational data during ship navigation. By utilizing such real-world operational data, the proposed model is designed to reflect a wide range of engine conditions encountered in actual maritime environments, laying a practical foundation for reliable fault diagnosis research.

Section 3.1 introduces the overall architecture of the proposed hybrid model, along with a description of the input data, which includes time-series raw data and TDF. Section 3.2 elaborates on the three key modules of the model: (i) the TDF extractor, responsible for extracting features from TDF; (ii) the raw signal feature extractor, which

processes and extracts features from the raw data; and (iii) the fault diagnosis classifier, designed to identify and classify fault degrees. Finally, Section 3.3 provides an analysis of the fault diagnosis performance under various noise conditions.

3.1. Overall Model Structure

In this paper, time-series cylinder exhaust gas temperature data are utilized for fault diagnosis of the ship's main engine. The overall architecture of the proposed deep hybrid model is illustrated in Figure 2. The model takes time-series raw data and TDFs as input. Specifically, the raw signal feature extractor processes the raw data, while the TDF extractor handles the TDFs to extract meaningful features. The output from each extractor is a 1×6 vector, which serves as input to the fault diagnosis classifier. This classifier performs multi-class classification, categorizing the data into six fault classes based on the severity of the fault. Our proposed fault diagnosis model is detailed by the following two steps. (i) Feature Extraction Step: Feature extraction is performed through a parallel architecture composed of a raw signal feature extractor and a TDF extractor, which process the raw signal and TDF features, respectively. (ii) Classification Step: The feature information extracted from the two separate feature extractors are fused using spatial and channel attention mechanism to emphasize informative representations. Then, based on the fused feature information, the final classification is performed to determine the status of the ship main engine. A detailed explanation of the raw data and TDFs, used as model inputs, is provided below.

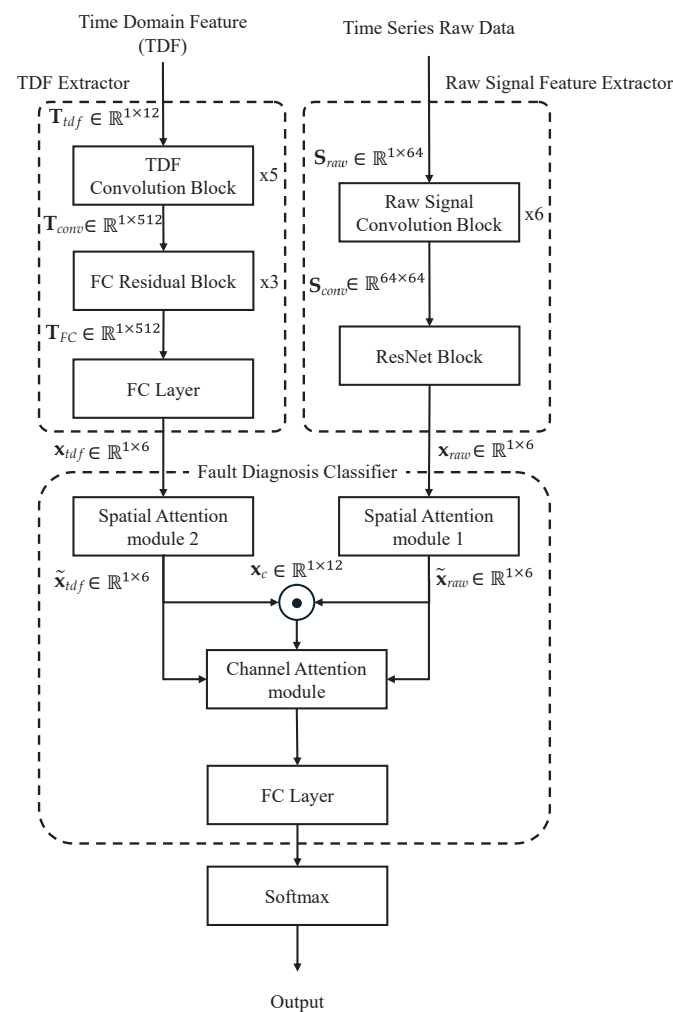


Figure 2. Proposed fault diagnosis model structure for marine diesel engine.

3.1.1. Time-Series Raw Data

One of the inputs to the proposed hybrid model is time-series raw data from cylinder exhaust gas temperature measurements. However, this data suffers from a significant lack of fault-related samples, making it difficult to use effectively for model training. To overcome this limitation, normal data are used to generate a dataset comprising six classes based on the fault conditions of a ship's main engine. The final dataset is then constructed, taking into account multiple levels of equipment degradation under marine environmental conditions. In the raw signal feature extractor of Figure 2, micro-level information, such as temporal patterns of defects, is captured from the raw data. This ensures that the model effectively identifies and learns complicated fault-related patterns.

3.1.2. TDF

TDF is used to capture statistical characteristics of time-series raw data for model training. Specifically, various statistical features called TDFs are extracted from the raw data and used as one of the inputs to the proposed hybrid model. While raw data input allows the model to learn fine-grained patterns, TDFs enable the model to capture macro-level changes via statistical features. In this paper, 12 TDFs are extracted from the raw data and used as one of the inputs to the hybrid model. These 12 TDFs are listed in Table 1.

Table 1. TDF calculation.

| No. | Feature Name | Feature Formula |
|-----|--------------------|------------------------------------------------------------------|
| 1 | Mean | $Mean = \frac{1}{n} \sum_{i=1}^n x_i$ |
| 2 | Max | $Max = \max(x_i)$ |
| 3 | Min | $Min = \min(x_i)$ |
| 4 | Standard Deviation | $Std = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - Mean)^2}$ |
| 5 | Skewness | $Skewness = \frac{\sum_{i=1}^n (x_i - x)^3}{(n-1) \times Std^3}$ |
| 6 | Kurtosis | $Kurtosis = \frac{\sum_{i=1}^n (x_i - x)^4}{(n-1) \times Std^4}$ |
| 7 | Peak-To-Peak | $P2P = Max - Min$ |
| 8 | Root Mean Square | $RMS = \sqrt{\frac{\sum_{i=1}^n (x_i)^2}{n}}$ |
| 9 | Variataion | $Var = \frac{1}{n} \sum_{i=1}^n (x_i - Mean)^2$ |
| 10 | Median | $Median = \{(n+1)/2\}^{th} term$ |
| 11 | Q1 | $Q1 = \{(n+1)/4\}^{th} term$ |
| 12 | Q3 | $Q3 = \{3(n+1)/4\}^{th} term$ |

3.2. Proposed Architecture

This subsection provides a detailed explanation of the structure and operating principles of the three key modules in the proposed hybrid model: the TDF extractor, the raw signal feature extractor, and the fault diagnosis classifier.

3.2.1. TDF Extractor

The TDF extractor takes 12 statistical features, extracted from the time-series raw data, as the input. As illustrated in Figure 2, the TDF extractor consists of five TDF convolution blocks, three fully connected (FC) residual blocks, and an FC layer. The output, $\mathbf{x}_{tdf} \in \mathbb{R}^{1 \times 6}$, is a vector that serves as one of the input feature maps to the fault diagnosis classifier.

The detailed architectures of the TDF convolution block and the FC residual block are presented in Figures 3 and 4, respectively.

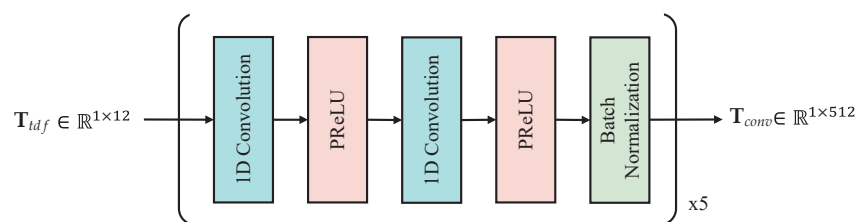


Figure 3. Structure of the TDF convolution block.

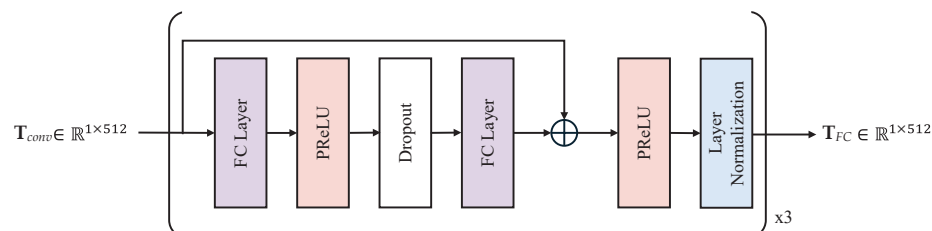


Figure 4. Structure of the FC residual block.

The TDF convolution block performs feature extraction by capturing the correlations within the TDF data. As shown in Figure 3, the TDFs are processed through the TDF convolution block, where 1D convolution is applied to capture nonlinear features and extract patterns reflecting the correlations among the data. To prevent the model from converging to a local optimum, PReLU activation and batch normalization are applied. This process is iteratively repeated five times. The detailed configuration of the 1D convolution layers within the TDF convolution block is presented in Table 2. Both layers are configured with a kernel size of 1×1 , a stride of 1, and zero padding.

Table 2. Structural parameters of TDF convolution block.

| Block | Convolution Layer | Input | Kernel Size | Output |
|--------|-------------------|-------|--------------|--------|
| First | First | 12 | 1×1 | 32 |
| | Second | 32 | 1×1 | 32 |
| Second | First | 32 | 1×1 | 64 |
| | Second | 64 | 1×1 | 64 |
| Third | First | 64 | 1×1 | 128 |
| | Second | 128 | 1×1 | 128 |
| Fourth | First | 128 | 1×1 | 256 |
| | Second | 256 | 1×1 | 256 |
| Fifth | First | 256 | 1×1 | 512 |
| | Second | 512 | 1×1 | 512 |

In the FC residual block, skip connections and dropout are used to address the vanishing gradient problem and reduce the risk of performance degradation. The FC layer is employed to extract high-dimensional features. As shown in Figure 4, each FC residual block consists of two fully connected layers for feature extraction, with PReLU and

dropout applied to avoid local optima. Specifically, the first FC layer reduces the input dimensionality from 512 to 256, and the second FC layer increases the dimensionality back to 512 to enhance nonlinear expressiveness. The dropout rate is set to 0.5. Skip connections help mitigate the vanishing gradient problem and improve the generalization ability of the model [35]. By applying skip connections followed by layer normalization, the model ensures that values within each layer follow the same distribution, thereby reducing instability during training and enhancing performance. This FC residual block is repeated three times.

3.2.2. Raw Signal Feature Extractor

The raw signal feature extractor consists of two components. First is the raw signal convolution block, which extracts features from the time-series raw data. Second is the ResNet [35] block, which extracts features from the feature map generated by the raw signal convolution block. In the raw signal convolution block, 1D convolution is applied to capture the correlations within the time-domain raw data. Additionally, batch normalization and PReLU are used to ensure stable learning. This process is repeated six times. The detailed structure of the raw signal convolution block is shown in Figure 5. The raw signal convolution block takes time-domain raw data of length 64 as the input, doubling the channel size during each step and transforming the data into a 64×64 feature map in image form. Specifically, a kernel size of 1×3 is used throughout the block, with output channels progressively set to 2, 4, 8, 16, 32, and 64 across the six layers. The stride is fixed at 1, and padding is set to 1 to maintain the resolution. This feature map is then fed into the ResNet block, known for its strong performance in multi-class image classification, enabling more effective feature extraction. In this way, the proposed method combines two blocks with distinct characteristics: the raw signal convolution block, which captures one-dimensional correlations, and the ResNet block, which captures two-dimensional correlations. The extracted feature map $\mathbf{x}_{raw} \in \mathbb{R}^{1 \times 64}$ from the raw signal feature extractor is used as one of the inputs to the classifier.

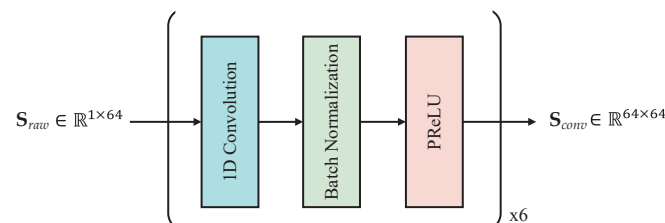


Figure 5. Structure of the raw signal convolution block.

3.2.3. Fault Diagnosis Classifier

The structure of the fault diagnosis classifier, which determines the fault status and degree in the proposed model is shown in Figure 6. The output vectors \mathbf{x}_{raw} from the raw signal feature extractor and \mathbf{x}_{tdf} from the TDF are input into spatial attention modules 1 and 2, respectively. In each spatial attention module, an FC block followed by a sigmoid activation function is used to compute weights for the elements of \mathbf{x}_{raw} and \mathbf{x}_{tdf} (i.e., each spatial element). These weights are then used to multiply the corresponding spaces, producing the outputs $\tilde{\mathbf{x}}_{raw}$ and $\tilde{\mathbf{x}}_{tdf}$, respectively.

In the channel attention module, the concatenated feature \mathbf{x}_c is defined as $\mathbf{x}_c = \text{Concat}(\tilde{\mathbf{x}}_{raw}, \tilde{\mathbf{x}}_{tdf})$, where *Concat* denotes the concatenation operation, and $\tilde{\mathbf{x}}_{raw}$ and $\tilde{\mathbf{x}}_{tdf}$ represent the refined features extracted from the spatial attention module 1 and 2, respectively. This concatenated vector \mathbf{x}_c is processed through an FC block and a sigmoid activation function to compute channel-wise weight elements ω_1 and ω_2 . The output of the channel attention module \mathbf{x}_a is computed as $\mathbf{x}_a = \text{Add}(\omega_1 \cdot \tilde{\mathbf{x}}_{raw}, \omega_2 \cdot \tilde{\mathbf{x}}_{tdf})$, where *Add*

denotes the element-wise addition operation, allowing the model to assign appropriate importance to each input branch. Finally, the feature map \mathbf{x}_a is processed through the FC layer and softmax activation function to predict the fault class. The structure of the FC block within the fault diagnosis classifier is shown in Figure 7. Specifically, the detailed parameter configurations of the FC block used in both spatial attention modules and the channel attention module are summarized in Table 3.

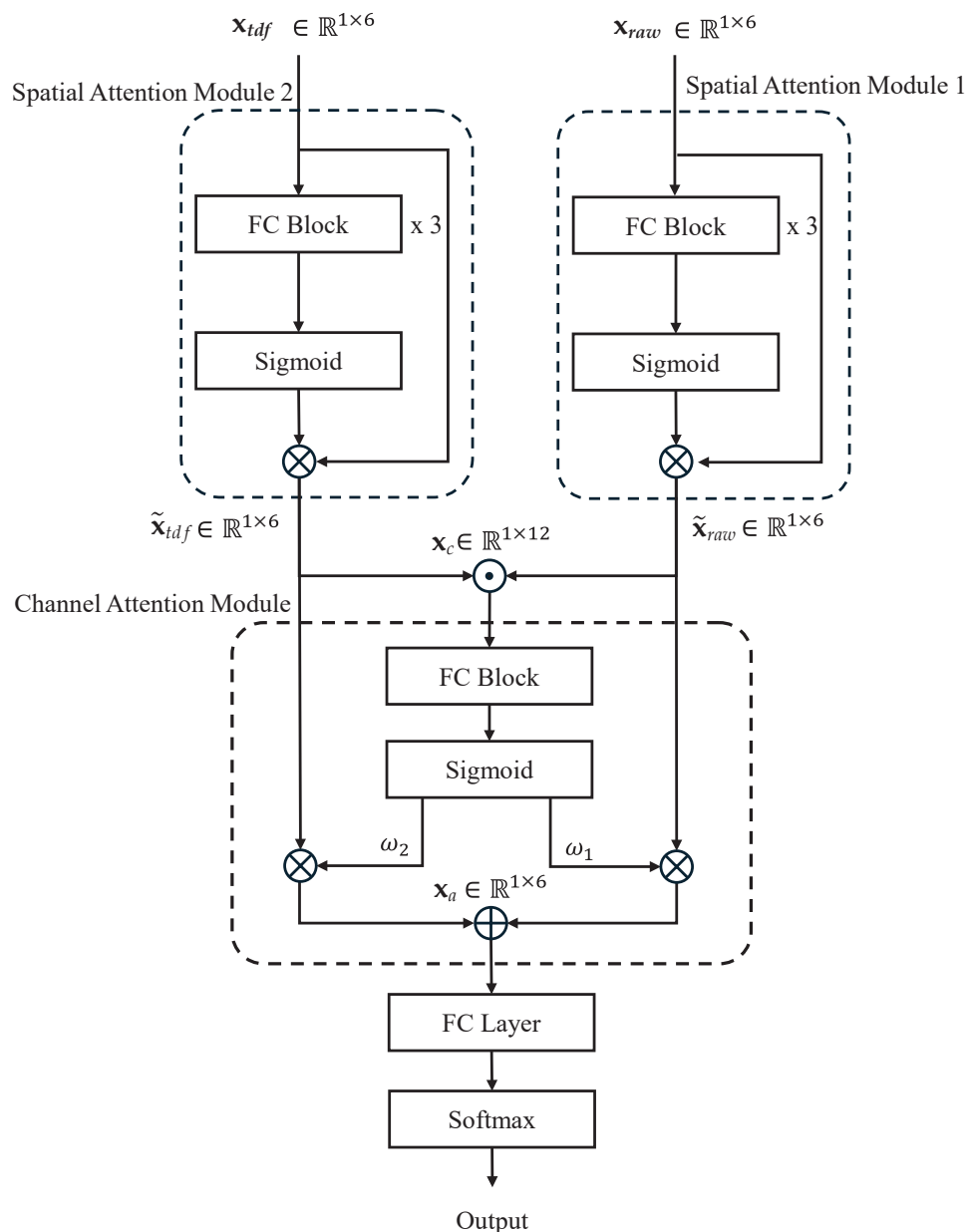


Figure 6. Structure of the fault diagnosis classifier.

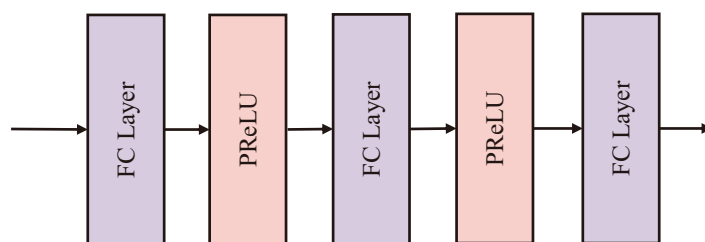


Figure 7. Structure of the FC block in attention module.

Table 3. Structural parameters of FC block in attention module.

| Module | Block | Layer | Input | Output |
|-------------------|--------|--------|-------|--------|
| Spatial Attention | First | First | 6 | 6 |
| | | Second | 6 | 6 |
| | | Third | 6 | 3 |
| | Second | First | 3 | 3 |
| | | Second | 3 | 3 |
| | | Third | 3 | 6 |
| | Third | First | 6 | 6 |
| | | Second | 6 | 6 |
| | | Third | 6 | 6 |
| Channel Attention | First | First | 12 | 12 |
| | | Second | 12 | 12 |
| | | Third | 12 | 12 |

3.3. Noise Environment

The main engine system of a ship is a complex structure with multiple interconnected components. During the data acquisition process using various sensors, noise can affect these sensors due to interference from surrounding equipment. To evaluate the performance of the proposed hybrid model, Gaussian noise is added to the time-series raw data, resulting in a noisy signal. Specifically, to assess the model performance under different noise conditions, five signal-to-noise ratio (SNR) cases ranging from -4 dB to 4 dB with 2 dB intervals are considered. The SNR is calculated as follows:

$$SNR = 10 \times \log_{10} \frac{P_{signal}}{P_{noise}} \quad (1)$$

where P_{signal} and P_{noise} denote the signal and noise power, respectively.

As seen in conventional studies, if only a single feature extraction process is employed, the model may be more susceptible to strong noise. However, as proposed in this paper, utilizing a hybrid model with two feature extraction processes allows the model to extract features through multiple pathways and achieve more robust performance against noise. In particular, in the proposed hybrid model, the use of TDF and the skip-connection structure emphasizes only the essential features of the data. This enables the model to learn effectively while progressively reducing irrelevant noise as it propagates through the layers. Moreover, channel and spatial attention processes allow the model to focus on important features by assigning higher weights to relevant data, thus mitigating the influence of noise during the training process.

4. Experiments

In Section 4, we provide a detailed description of the dataset and simulation environment used in this paper, followed by the performance evaluation results of the proposed hybrid model. In Section 4.1, the dataset utilized in the experiments is introduced. Section 4.2 outlines the experimental scenario. Section 4.3 describes the data preparation process, considering the corresponding scenario. Finally, Sections 4.4 and 4.5 present the experimental setup and results, respectively.

4.1. Dataset

The experiments are conducted using cylinder exhaust gas temperature data collected over approximately 12 months at 30-s intervals (i.e., at a sampling rate of $1/30$ Hz) from a

MAN B&W low-speed two-stroke marine diesel engine (Augsburg, Germany). The engine has five cylinders, a maximum continuous rating of 3200 kW, an operating speed range of 148–195 rpm, and a mean effective pressure of 21.0 bar. Each cylinder is equipped with a sensor to measure the exhaust gas temperature. Typically, the normal exhaust gas temperature of the cylinders ranges from 250 °C to 395 °C, while a temperature exceeding 600 °C indicates a fault. For the main engine fault diagnosis experiments, the average value of data collected from the five sensors is utilized.

4.2. Fault Scenario

The main components of a ship main engine include the fuel system, cooling system, lubrication system, and air intake and exhaust system. These components work together to ensure the efficient and stable operation of the diesel engine [36]. The exhaust gas temperature of a ship main engine is one of the key indicators of its overall condition, including combustion efficiency, cooling, and lubrication [14,17]. If these components degrade due to long-term operation or external environmental influences, abnormalities in the combustion process may occur, leading to an increase in the exhaust gas temperature [37]. Unlike automobiles, ships are not free to anchor and are heavily influenced by the external environment. As a result, the engine operating conditions are not constant, and excessive operation can occur depending on the situation. Due to external factors such as the aging of internal components, heat dissipation may be ineffective, leading to heat accumulation [17]. This heat accumulation can cause the exhaust gas temperature to be higher than that of the current state of the engine.

In this paper, when preparing data for model training, we classify the data into six categories based on the engine condition and ensure that data from adjacent classes overlapped by a certain ratio account for heat accumulation. The training data are configured to reflect different levels of heat accumulation, ranging from no overlap (representing no heat accumulation) to significant overlap (representing severe heat accumulation). Model performance is then evaluated through experiments considering these varying scenarios.

4.3. Data Preparation

The exhaust gas temperature data collected in the actual operating environment represents normal operating conditions. However, to create a fault diagnosis model for the ship's main engine, engine status data categorized by stage according to the degree of failure are required. Therefore, we divide the exhaust gas temperature data into six classes based on the degree of failure. Among these six classes, the data from the 4th, 5th, and 6th classes are defined as failure states. Specifically, the 4th class represents the initial stage of failure, the 5th class represents the stage in which the failure is in progress, and the 6th class corresponds to a serious failure state. Since the exhaust gas temperature of a two-stroke marine diesel engine exceeds 600 °C during failure, the average temperature of the 4th class (the initial failure stage) is set to 600 °C. The 1st class represents normal data collected from actual ship operations without modification. The average temperature of the 2nd and 3rd class data is determined by linear interpolation between the average temperatures of the 1st and 4th classes, and data for these classes are created using the interpolated temperature values. The average temperatures of the 5th and 6th classes are set to maintain the increasing rate of temperature values from the 1st to the 4th classes, and training data are created accordingly.

Additionally, heat accumulation in the engine is considered according to the scenario described earlier. In other words, the data are generated under the assumption that heat may accumulate in the engine, causing the exhaust gas temperature to exceed the expected value for the current engine condition. Specifically, the data of adjacent classes are scaled to overlap at certain ratios, and fault diagnosis performance is evaluated through experiments

with overlap ratios of 0%, 10%, 20%, 30%, 40%, and 50%. The overall data generation procedure based on the proposed fault progression scenario is detailed in Algorithm 1. The original temperature dataset D is evenly divided into six subsets, D_1 through D_6 , each corresponding to a specific fault class. The average temperature of the fourth class is fixed at 600 °C, and to ensure a uniform difference in mean temperature between adjacent classes, we calculate the temperature gap as $\Delta c = (600 - \text{mean}_1)/3$, where mean_1 denotes the average of the first class. This value Δc represents the inter-class mean difference and is used to derive the per-timestep temperature increment Δx . Data are then generated to reflect both a linearly increasing temperature pattern and a predefined thermal baseline. In this process, two hyperparameters, α and β , are introduced. Specifically, α controls the slope of the temperature increase within each time series, while β adjusts the thermal baseline level. These hyperparameters are employed to regulate the degree of overlap between adjacent classes. To achieve the desired class overlap ratios, α and β are determined experimentally. For overlap levels ranging from 10% to 40%, the parameters were finely tuned to closely approximate the target ratios. The final parameter settings for each case are summarized in Table 4. This data generation approach effectively captures the fault progression and thermal accumulation scenarios. When data from adjacent classes overlap, it becomes challenging to determine the current engine condition accurately based solely on the temperature at a specific point in time. Therefore, a strategy that utilizes the temporal patterns of time-series data is necessary to more precisely assess the engine's condition. In the experiments, the performance of the proposed hybrid model is compared with various other models to evaluate its effectiveness.

Algorithm 1 Data Generation Process

Input: Exhaust gas temperature data D
Hyperparameters: α (trend slope), β (intercept)
 1: Split D into 6 equal parts: $D_1, D_2, D_3, D_4, D_5, D_6$
 2: Compute $\text{mean}_1 \leftarrow \text{mean}(D_1)$
 3: Compute $\Delta c \leftarrow (600 - \text{mean}_1) / 3$
 4: Compute $\Delta x \leftarrow \Delta c / \text{len}(D_1)$
 5: Compute $\text{slope} \leftarrow \alpha \cdot \Delta x$
 6: **for** $i = 2$ to 6 **do**
 7: $\text{mean}_i \leftarrow \text{mean}(D_i)$
 8: $\text{offset} \leftarrow \text{mean}_1 - \text{mean}_i$
 9: $\text{linear} \leftarrow \{1, 2, \dots, \text{len}(D_i)\} \cdot \text{slope}$
 10: $D'_i \leftarrow D_i + \text{offset} + (i - 2 + \beta) \cdot \Delta c + \text{linear}$
 11: **end for**
 12: **Output:** Synthetic datasets $D_1, D'_2, D'_3, D'_4, D'_5, D'_6$

Table 4. Hyperparameter settings for class overlapping percentage.

| α | β | Overlapping Percentage (%) |
|----------|---------|----------------------------|
| 1.00 | 0.50 | 0 |
| 1.10 | 0.45 | 10 |
| 1.26 | 0.37 | 20 |
| 1.44 | 0.28 | 30 |
| 1.68 | 0.16 | 40 |
| 2.00 | 0.00 | 50 |

4.4. Experiment Setting

In this experiment, the Adam optimizer is used with the hyperparameters set to $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The learning rate is initially set to 0.002, and the cosine annealing warm-up restart algorithm [38] is applied to adjust the learning rate at each epoch during training. Cross entropy is employed as the loss function to optimize the model with a batch size of 256 and 500 epochs. The experiments are conducted in a Linux 20.08 environment, using Python version 3.9.13 and PyTorch version 2.0.1, and performed on an Intel Core i9-13900K CPU and an NVIDIA A6000 GPU Santa Clara, CA, USA).

4.5. Experiment Results

In this subsection, the fault diagnosis performance of the proposed methods is evaluated and analyzed. First, the performance of the proposed method is compared with that of existing methods. Second, the performance of both the proposed and existing methods is evaluated across various noise levels. Third, the effectiveness of different feature fusion techniques used in the fault diagnosis classifier is compared. Finally, the impact of applying channel attention and spatial attention modules is evaluated by comparing performance with and without these modules.

4.5.1. Fault Diagnosis Accuracy

We evaluate and compare the performance of the proposed fault diagnosis model with that of conventional models. The conventional models include (i) random forest classifier [39], (ii) CNN and bi-directional GRU (CNN+BiGRU) [16], (iii) a model that combines CNN and bi-directional LSTM with Attention (CNN+BiLSTM+Attention) [17], and (iv) a model that incorporates the raw signal convolution block with vision transformer (RSCB+ViT) [40]. To evaluate the performance considering the degree of engine heat accumulation, the overlap between engine state classes is varied from 0% to 50%. The accuracy results for the proposed hybrid model and the conventional methods are presented in Table 5. The random forest classifier is trained using TDF, while the CNN+BiGRU, CNN+BiLSTM+Attention, and RSCB+ViT models are trained using time-series raw data as input.

Table 5. Fault diagnosis results for four conventional models and six overlap percentages.

| Model | Fault Diagnosis Accuracy (%) for 10 % to 50 % Overlapping Percentage | | | | | |
|-------------------------------|----------------------------------------------------------------------|--------|--------|--------|--------|--------|
| | 0% | 10% | 20% | 30% | 40% | 50% |
| Random Forest Classifier [39] | 98.426 | 96.775 | 95.037 | 92.219 | 89.902 | 88.234 |
| CNN+BiGRU [16] | 97.842 | 95.215 | 92.379 | 87.154 | 85.151 | 79.682 |
| CNN+BiLSTM+Attention [17] | 98.345 | 95.140 | 92.218 | 91.378 | 88.584 | 83.713 |
| RSCB+ViT [40] | 97.113 | 94.264 | 91.671 | 84.737 | 81.430 | 73.781 |
| Proposed | 99.148 | 98.976 | 98.293 | 97.574 | 96.857 | 94.380 |

The experimental results demonstrate that the proposed method achieved the highest accuracy across all levels of degradation, i.e., for all overlap percentages. As shown in Table 5, the fault diagnosis accuracy generally decreases as the overlap percentage increases. While the performance of most conventional methods significantly deteriorates, the proposed method shows a relatively smaller decrease. Specifically, for a 50% overlap, the proposed method achieves an accuracy improvement of up to 20.599%, with a minimum improvement of 6.146% over the existing methods. This demonstrates that the proposed

method maintains strong fault diagnosis capability, even under conditions of significant heat accumulation caused by high levels of equipment degradation.

To evaluate the real-time applicability of the proposed hybrid model, we measured its inference time and computational complexity in terms of floating-point operations (FLOPs). Table 6 presents the results of the inference time measured on both CPU and GPU platforms, along with the computational complexity based on FLOPs. The experimental results show that the proposed model achieves an inference time of 10.864 ms on GPU and 11.293 ms on CPU, with a total computational load of approximately 0.3 GFLOPs. These results suggest that the proposed model features a lightweight architecture, making it well suited for real-time applications.

Table 6. Inference time and computational complexity of the proposed model.

| Inference Time on GPU | Inference Time on CPU | FLOPs |
|-----------------------|-----------------------|-------|
| 10.864 ms | 11.293 ms | 0.3 G |

4.5.2. Noisy Environment

The performance of the proposed hybrid model is compared with that of conventional models under noisy conditions. Specifically, the fault diagnosis accuracy of the proposed model and the conventional models is presented in Table 7 for different SNR values. In this experiment, to evaluate the robustness of the fault diagnosis model against noisy signals, the overlap percentage is set to 10%, and the SNR is tested at five values: -4 , -2 , 0 , 2 , and 4 . Compared to the results in Table 5, which do not account for noise, all conventional methods in Table 7 show a significant decrease in accuracy across all SNR ranges. In particular, when the SNR is 0 , the accuracy decreases by up to 13.240% and at least 9.121% compared to the case without noise (i.e., when the overlap percentage in Table 5 is 10%). In contrast, the proposed hybrid method showed a relatively small accuracy decrease of 1.392%. This result demonstrates the robustness of the proposed hybrid method in maintaining fault diagnosis performance even under noisy conditions.

Table 7. Fault diagnosis results with different SNRs in 10% overlapping percentage.

| Model | Accuracy (%) with Different SNRs (dB) | | | | |
|-------------------------------|---------------------------------------|--------|--------|--------|--------|
| | -4 | -2 | 0 | 2 | 4 |
| Random Forest Classifier [39] | 80.727 | 81.791 | 83.535 | 85.125 | 87.334 |
| CNN+BiGRU [16] | 80.110 | 81.505 | 83.471 | 85.083 | 86.800 |
| CNN+BiLSTM+Attention [17] | 81.693 | 83.646 | 86.019 | 87.846 | 89.586 |
| RSCB+ViT [40] | 79.339 | 85.955 | 82.150 | 85.739 | 86.551 |
| Proposed | 97.292 | 97.323 | 97.584 | 97.880 | 98.337 |

4.5.3. Feature Fusion

In the fault diagnosis classifier of the proposed model, feature fusion is performed twice: through concatenation (\odot) and addition (\oplus), as shown in Figure 6. Specifically, the first feature fusion F_1 occurs when the outputs of the spatial attention modules, \tilde{x}_{raw} and \tilde{x}_{idf} , are combined. The second instance of feature fusion F_2 occurs when the two feature maps are merged before the channel attention output. In the proposed hybrid model, F_1 and F_2 are designed as concatenation (\odot) and addition (\oplus), respectively. The choice of F_1 as concatenation minimizes information dilution by preserving distinct features, and F_2 as an addition facilitates efficient training by putting information on the presence

or absence of faults in the positional elements of the added feature vector. As shown in Table 8, the proposed model demonstrates superior performance compared to other feature fusion combinations, with accuracy improvements ranging from a minimum of 1.930% to a maximum of 2.678%.

Table 8. Fault diagnosis accuracy for various feature fusion cases.

| Case | F_1 | F_2 | Accuracy (%) |
|----------|---------------|---------------|--------------|
| Case 1 | Addition | Addition | 94.927 |
| Case 2 | Addition | Concatenation | 94.179 |
| Case 3 | Concatenation | Addition | 94.231 |
| Proposed | Concatenation | Addition | 96.857 |

4.5.4. Attention Mechanism

Both spatial and channel attention modules are used simultaneously in the classifier of the proposed hybrid model, as shown in Figure 6. Specifically, spatial attention is applied to \mathbf{x}_{raw} and \mathbf{x}_{tdf} , which are features extracted by the raw signal feature extractor and the TDF extractor, respectively, as depicted in Figure 2. The spatial attention module assigns element-wise weights to \mathbf{x}_{raw} and \mathbf{x}_{tdf} , producing the feature maps $\tilde{\mathbf{x}}_{raw}$ and $\tilde{\mathbf{x}}_{tdf}$. Subsequently, channel attention assigns weights to the relationship between $\tilde{\mathbf{x}}_{raw}$ and $\tilde{\mathbf{x}}_{tdf}$, generated through spatial attention. These two attention mechanisms complement each other, enabling the model to focus on critical information across specific dimensions of the feature maps. This synergy plays a pivotal role in enhancing the model's overall performance. To validate this, the fault diagnosis accuracy is compared across three variations of the classifier: one using only channel attention (channel-attention-based classifier), one using only spatial attention (spatial-attention-based classifier), and the proposed classifier that incorporates both mechanisms as shown in Figure 8. The proposed model significantly outperforms the other two classifiers, demonstrating the effectiveness of integrating both attention mechanisms as shown in Table 9.

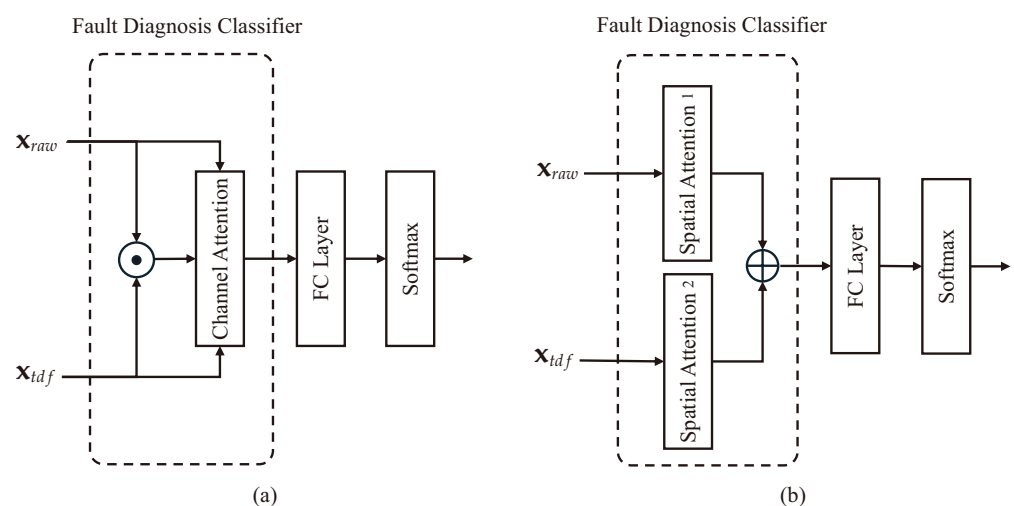


Figure 8. A structure of the fault diagnosis classifier with single attention: (a) channel-attention-based classifier; (b) spatial-attention-based classifier.

Table 9. Accuracy comparison between various attention mechanism methods.

| Classifier | Accuracy |
|------------------------------------|----------|
| Channel-attention-based classifier | 92.544 |
| Spatial-attention-based classifier | 94.872 |
| Proposed classifier | 96.857 |

4.5.5. Number of Convolution Blocks

As illustrated in Figure 2, the proposed hybrid model simultaneously utilizes time-series raw data and TDF as inputs. These inputs are processed through the raw signal feature extractor and the TDF extractor, respectively, to produce the feature vectors \mathbf{x}_{raw} and \mathbf{x}_{tdf} . In the implemented architecture, the raw signal feature extractor comprises six raw signal convolution blocks, whereas the TDF extractor is composed of five TDF convolution blocks. This subsection presents experimental results aimed at determining the optimal number of convolution blocks for each pathway. Table 10 compares fault diagnosis accuracy across three configurations with different numbers of raw signal and TDF convolution blocks. The configuration using five raw signal blocks and four TDF blocks exhibited slightly lower accuracy than the proposed setting, primarily due to insufficient feature extraction capacity in both branches. Conversely, employing seven raw signal blocks and six TDF blocks resulted in significantly degraded performance, which is attributed to overfitting caused by excessive model complexity. These results demonstrate that the proposed configuration comprising six raw signal convolution blocks and five TDF convolution blocks achieves the best balance between model complexity and feature extraction capability, thereby yielding the highest fault diagnosis accuracy.

Table 10. Accuracy comparison with varying numbers of raw signal and TDF convolution blocks.

| The Number of Raw Signal Blocks | The Number of TDF Blocks | Accuracy (%) |
|---------------------------------|--------------------------|-------------------|
| 5 | 4 | 95.831 |
| 6 | 5 | 96.857 (Proposed) |
| 7 | 6 | 91.989 |

5. Discussion

5.1. Limitations

While the proposed model demonstrates promising diagnostic performance using cylinder exhaust-gas temperature data, its current formulation has two key limitations. First, it relies on a single sensor modality, which restricts its ability to capture the full complexity of the engine's operational behavior. Exhaust gas temperature alone may not reflect certain types of failures and thus the proposed method may have limitations in localizing faults or identifying their specific causes. Second, the model's robustness was evaluated only under additive Gaussian noise, which does not fully represent the range of disturbances encountered in real-world marine environments. These limitations suggest the need for both input diversity and more realistic noise modeling to enhance generalization and reliability.

5.2. Model Applicability

The proposed model was developed and validated using data collected from a large ship low-speed two-stroke main engine installed on an ocean-going vessel. This type of engine has distinct thermal and operational characteristics, such as low rotational speed, large bore size, and electronically controlled fuel injection. Therefore, the applicability of the model is explicitly limited to similar low-speed main propulsion engines equipped with comparable sensors and operating under similar load profiles. To extend the model to other engine types such as four-stroke engines or auxiliary generators, further adaptation is required. This may include transfer learning using a small number of labeled samples from the target system or domain adaptation techniques to align the feature distributions between different engine domains.

5.3. Future Works

In the present study, we relied solely on cylinder exhaust-gas temperature and evaluated robustness only under additive Gaussian noise, yet real-world marine data are exposed to more complex disturbances, non-Gaussian noise, long-term sensor drift, and intermittent data loss, which can undermine diagnostic accuracy. Future work will, therefore, address these two limitations simultaneously by (1) introducing richer noise models such as impulsive bursts, drift trends, and missing segments while employing data-imputation and noise-adaptive learning to alleviate their impact, (2) extending the input space beyond exhaust-gas temperature to include cylinder pressure, vibration, and fuel-injecting timing so that the full work cycle of a low-speed main engine is captured. Furthermore, since the current model focuses on early warning and anomaly detection rather than root cause identification, future research will aim to incorporate causal reasoning by fusing additional signals such as injection timing, intake pressure, and cylinder dynamics. This integration will help enable reliable cause-specific diagnosis beyond general anomaly detection. We will also investigate unsupervised pre-training so that the expanded multi-sensor framework can generalize under noisy or incomplete conditions, ultimately paving the way for a robust, system-level diagnostic platform. Furthermore, the current dataset does not include labeled information for identifying the specific causes of engine faults. Constructing a cause-annotated dataset and developing supervised learning models to classify fault causes such as mechanical wear, timing misadjustments, or fuel quality issues will be an important extension of this work. This will contribute to the development of a more interpretable and practically useful diagnostic system capable of pinpointing the root causes of observed anomalies.

6. Conclusions

In this paper, we propose a hybrid model that diagnoses faults in ship main engines using exhaust gas temperature data, specifically targeting faults caused by equipment degradation. The proposed hybrid model was designed to extract features by simultaneously leveraging time-series raw data and time-domain features (TDFs). We used two attention mechanisms (spatial and channel) and found that better classification performance can be achieved by properly combining these two mechanisms given the structure of the hybrid model. To evaluate the performance of the proposed model, fault scenarios involving heat accumulation and overheating due to equipment degradation were considered based on exhaust-gas temperature data extracted from the main engines of large ships. Additionally, extended experiments were conducted on various overlap percentages to analyze the impact of degradation severity on model performance. The experimental results showed that the proposed model achieved the highest accuracy compared to the existing methods. Additionally, the model demonstrated significantly greater robustness

under noisy conditions. These findings indicated that the proposed fault diagnosis model was less affected by both equipment degradation and noise.

Author Contributions: Conceptualization, S.-H.K. and T.-G.K.; Methodology, S.-H.K. and T.-G.K.; Software, S.-H.K. and T.-G.K.; Validation, J.L. and C.-J.C.; Formal Analysis, S.-H.K. and C.-J.C.; Investigation, T.-G.K. and J.L.; Resources, H.-K.S.; Data Curation, J.L. and H.M.; Writing—Original Draft Preparation, S.-H.K. and T.-G.K.; Writing—Review and Editing, C.-J.C. and H.-K.S.; Visualization, H.M.; Supervision, C.-J.C.; Project Administration, C.-J.C.; Funding Acquisition, C.-J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) under the metaverse support program to nurture the best talents (IITP-2025-RS-2023-00254529) grant funded by the Korea government (MSIT) and supported by Culture, Sports and Tourism R&D Program through the Korea Creative Content Agency (KOCCA) grant funded by the Ministry of Culture, Sports and Tourism (MCST) in 2024 (Project Name: Cultivating masters and doctoral experts to lead digital-tech tourism, Project Number: RS-2024-00442006, and Contribution Rate: 50%).

Data Availability Statement: The datasets presented in this article are not readily available because they are proprietary data provided by an industrial partner and are subject to confidentiality agreements. Therefore, the data cannot be shared publicly.

Conflicts of Interest: Author Junseok Lee was employed by the company Okestro. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Grbic, T. The Importance of Shipping in the Transportation Industry. Technical Report. 2016. Available online: <https://blog.greencarrier.com/the-importance-of-shipping-in-the-transportation-industry/> (accessed on 12 February 2016).
2. Li, Z.; Yan, X.; Yuan, C.; Peng, Z. Intelligent fault diagnosis method for marine diesel engines using instantaneous angular speed. *J. Mech. Sci. Technol.* **2012**, *26*, 2413–2423. [\[CrossRef\]](#)
3. Yang, W.; Zimroz, R.; Papaelias, M. Advances in machine condition monitoring and fault diagnosis. *Electronics* **2022**, *11*, 1563. [\[CrossRef\]](#)
4. Velasco, C.; Lazakis, I. Analysis of time series imaging approaches for the application of fault classification of marine systems. In Proceedings of the European Safety and Reliability Conference, Dublin, Ireland, 28 August–1 September 2022.
5. Kocak, G.; Gokcek, V.; Genc, Y. Condition monitoring and fault diagnosis of a marine diesel engine with machine learning techniques. *Pomorstvo* **2023**, *37*, 32–46. [\[CrossRef\]](#)
6. Tsaganos, G.; Papachristos, D.; Nikitakos, N.; Dalaklis, D.; Ölcner, A. Fault detection and diagnosis of two-stroke low-speed marine engine with machine learning algorithms. In Proceedings of the 3rd International Symposium on Naval Architecture and Maritime, Istanbul, Turkey, 24–25 April 2018.
7. Tsaganos, G.; Nikitakos, N.; Dalaklis, D.; Ölcner, A.; Papachristos, D. Machine learning algorithms in shipping: Improving engine fault detection and diagnosis via ensemble methods. *WMU J. Marit. Aff.* **2020**, *19*, 51–72. [\[CrossRef\]](#)
8. Cheliotis, M.; Lazakis, I.; Theotokatos, G. Machine learning and data-driven fault detection for ship systems operations. *Ocean Eng.* **2020**, *216*, 107968. [\[CrossRef\]](#)
9. Cai, C.; Weng, X.; Zhang, C. A novel approach for marine diesel engine fault diagnosis. *Clust. Comput.* **2017**, *20*, 1691–1702. [\[CrossRef\]](#)
10. Tong, S.; Yanqiao, C.; Yuan, Z. Fault prediction of marine diesel engine based on time series and support vector machine. In Proceedings of the 2020 International Conference on Intelligent Design (ICID), Xi'an, China, 11–13 December 2020; pp. 75–81.
11. Lee, S.; Lee, T.; Kim, J.; Lee, J.; Ryu, K.; Kim, Y.; Park, J.W. A Study on the Application of Discrete Wavelet Decomposition for Fault Diagnosis on a Ship Oil Purifier. *Processes* **2022**, *10*, 1468. [\[CrossRef\]](#)
12. Diez-Olivan, A.; Pagan, J.A.; Sanz, R.; Sierra, B. Data-driven prognostics using a combination of constrained K-means clustering, fuzzy modeling and LOF-based score. *Neurocomputing* **2017**, *241*, 97–107. [\[CrossRef\]](#)
13. Han, P.; Li, G.; Skulstad, R.; Skjong, S.; Zhang, H. A deep learning approach to detect and isolate thruster failures for dynamically positioned vessels using motion data. *IEEE Trans. Instrum. Meas.* **2020**, *70*, 1–11. [\[CrossRef\]](#)

14. Velasco-Gallego, C.; Lazakis, I. RADIS: A real-time anomaly detection intelligent system for fault diagnosis of marine machinery. *Expert Syst. Appl.* **2022**, *204*, 117634. [\[CrossRef\]](#)
15. Brandsæter, A.; Vanem, E.; Glad, I.K. Cluster based anomaly detection with applications in the maritime industry. In Proceedings of the 2017 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC), Shanghai, China, 16–18 August 2017; pp. 328–333.
16. Liu, B.; Gan, H.; Chen, D.; Shu, Z. Research on fault early warning of marine diesel engine based on CNN-BiGRU. *J. Mar. Sci. Eng.* **2022**, *11*, 56. [\[CrossRef\]](#)
17. Ji, Z.; Gan, H.; Liu, B. A deep learning-based fault warning model for exhaust temperature prediction and fault warning of marine diesel engine. *J. Mar. Sci. Eng.* **2023**, *11*, 1509. [\[CrossRef\]](#)
18. Lazakis, I.; Raptodimos, Y.; Varelas, T. Predicting ship machinery system condition through analytical reliability tools and artificial neural networks. *Ocean Eng.* **2018**, *152*, 404–415. [\[CrossRef\]](#)
19. Theodoropoulos, P.; Spandonidis, C.C.; Giannopoulos, F.; Fassois, S. A deep learning-based fault detection model for optimization of shipping operations and enhancement of maritime safety. *Sensors* **2021**, *21*, 5658. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Yu, C.; Qi, L.; Sun, J.; Jiang, C.; Su, J.; Shu, W. Fault diagnosis technology for ship electrical power system. *Energies* **2022**, *15*, 1287. [\[CrossRef\]](#)
21. Ai, Z.; Cao, H.; Wang, J.; Cui, Z.; Wang, L.; Jiang, K. Research method for ship engine fault diagnosis based on multi-head graph attention feature fusion. *Appl. Sci.* **2023**, *13*, 12421. [\[CrossRef\]](#)
22. Senemmar, S.; Jacob, R.A.; Zhang, J. Non-intrusive fault detection in shipboard power systems using wavelet graph neural networks. *Meas. Energy* **2024**, *3*, 100009. [\[CrossRef\]](#)
23. Li, S.; Liu, Z.; Yan, Y.; Wang, R.; Dong, E.; Cheng, Z. Research on diesel engine fault status identification method based on synchro squeezing S-transform and vision transformer. *Sensors* **2023**, *23*, 6447. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Liu, Y.; Jin, H.; Yang, X.; Tang, T.; Song, Q.; Chen, Y.; Liu, L.; Jiang, S. Early Fault Diagnosis and Prediction of Marine Large-Capacity Batteries Based on Real Data. *J. Mar. Sci. Eng.* **2024**, *12*, 2253. [\[CrossRef\]](#)
25. Zhao, Z.; Cerf, S.; Birke, R.; Robu, B.; Bouchenak, S.; Mokhtar, S.B.; Chen, L.Y. Robust anomaly detection on unreliable data. In Proceedings of the 2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), Portland, OR, USA, 24–27 June 2019; pp. 630–637.
26. Li, Y.; Huang, X.; Ding, P.; Zhao, C. Wiener-based remaining useful life prediction of rolling bearings using improved Kalman filtering and adaptive modification. *Measurement* **2021**, *182*, 109706. [\[CrossRef\]](#)
27. Velasco-Gallego, C.; Lazakis, I. Development of a time series imaging approach for fault classification of marine systems. *Ocean Eng.* **2022**, *263*, 112297. [\[CrossRef\]](#)
28. Wang, R.; Chen, H.; Guan, C. A self-supervised contrastive learning framework with the nearest neighbors matching for the fault diagnosis of marine machinery. *Ocean Eng.* **2023**, *270*, 113437. [\[CrossRef\]](#)
29. Sinitsin, V.; Ibryaeva, O.; Sakovskaya, V.; Ereemeeva, V. Intelligent bearing fault diagnosis method combining mixed input and hybrid CNN-MLP model. *Mech. Syst. Signal Process.* **2022**, *180*, 109454. [\[CrossRef\]](#)
30. Liang, M.; Cao, P.; Tang, J. Rolling bearing fault diagnosis based on feature fusion with parallel convolutional neural network. *Int. J. Adv. Manuf. Technol.* **2021**, *112*, 819–831. [\[CrossRef\]](#)
31. Niu, Z.; Zhong, G.; Yu, H. A review on the attention mechanism of deep learning. *Neurocomputing* **2021**, *452*, 48–62. [\[CrossRef\]](#)
32. Bahdanau, D. Neural machine translation by jointly learning to align and translate. *arXiv* **2014**, arXiv:1409.0473.
33. Ashish, V. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1.
34. Wang, Z.; Oates, T. Encoding time series as images for visual inspection and classification using tiled convolutional neural networks. In Proceedings of the Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015.
35. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
36. Youssef, A.; Noura, H.; Amrani, A.E.; Adel, E.M.E.; Ouladsine, M. A Survey on Data-Driven Fault Diagnostic Techniques for Marine Diesel Engines. *arXiv* **2024**, arXiv:2404.10363. [\[CrossRef\]](#)
37. Lv, Y.; Yang, X.; Li, Y.; Liu, J.; Li, S. Fault detection and diagnosis of marine diesel engines: A systematic review. *Ocean Eng.* **2024**, *294*, 116798. [\[CrossRef\]](#)
38. Loshchilov, I.; Hutter, F. Sgdr: Stochastic gradient descent with warm restarts. *arXiv* **2016**, arXiv:1608.03983.
39. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
40. Dosovitskiy, A. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv* **2020**, arXiv:2010.11929.

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.