

Article

# A Deep Learning-Based Fault Warning Model for Exhaust Temperature Prediction and Fault Warning of Marine Diesel Engine

Zhenguo Ji, Huibing Gan \*  and Ben Liu

Marine Engineering College, Dalian Maritime University, Dalian 116026, China;  
jzg1120211158@dmlu.edu.cn (Z.J.); lb1120200087@dmlu.edu.cn (B.L.)

\* Correspondence: ghbzq@dmlu.edu.cn

**Abstract:** Marine diesel engines are essential for safe navigation. By predicting the operating conditions of diesel engines, the performance of marine diesel engines can be improved, failures can be prevented to reduce maintenance costs, and emissions can be controlled to protect the environment. To this end, this paper proposes a hybrid neural network (HNN) prediction model (CNN-BiLSTM-Attention) based on deep learning (DL) for predicting the exhaust gas temperature (EGT) of marine diesel engines. CNN is used to extract features from time-series data, BiLSTM is used to predict the time series through modeling, and Attention is used to improve the accuracy and robustness of fault prediction. Moreover, through comparison experiments with other neural network prediction models, it has been proven that the CNN-BiLSTM-Attention method is more accurate. This article also presents an approach to fault prediction by integrating the Mahalanobia distance and the mathematical model. Based on the Mahalanobia distance between the prediction result and the actual value, the function mapping method combined with the criterion is used to set the alarm value and threshold of the monitoring indicators, and the failure data set is used for experimental verification. The results indicate that the approach presented in this article can accurately realize the operating condition monitoring and fault early warning of marine diesel engines, which provides a new way of thinking for the research of fault early warning and health management of marine diesel engines.

**Keywords:** marine diesel engines; CNN; BiLSTM; attention; failure warning; Mahalanobia distance



**Citation:** 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. <https://doi.org/10.3390/jmse11081509>

Academic Editor: Leszek Chybowski

Received: 28 June 2023  
Revised: 27 July 2023  
Accepted: 27 July 2023  
Published: 29 July 2023



**Copyright:** © 2023 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

A ship's diesel engine is the primary propulsion unit of a ship. Its reliability and stability are critical to the safe and economical operation of the vessel [1]. Due to the complex structure of marine diesel engines and their long-term operation under high temperatures and pressure, abnormal conditions are likely to occur [2]. The traditional maintenance of marine diesel engines mainly adopts the way of regular maintenance and after-sales maintenance [3]. Regular maintenance requires downtime and a lot of money, which affects productivity and may not be able to predict all malfunctions and lead to timely repair, while after-sales maintenance cannot prevent accidents from happening and may threaten the safety of ship personnel and property in serious cases [4]. Therefore, it is very important to study the early warning technology of the ship's diesel engine.

With the rapid development of modern industrial technology, marine equipment systems are becoming more automated and intelligent [5]. Conventional diesel engine detection and warning techniques are mostly concerned with the external parameters of the diesel engine. It is only when the fault has reached a certain level of deterioration that the parameters show a noticeable abnormality [6]. This means that conventional methods of monitoring and alerting ships are unable to provide a timely warning. Failure warning technology is a system monitoring and prediction technology that can collect and analyze data from equipment or systems to predict possible failures and take timely action to avoid

them [7]. Fault warnings can improve the reliability and stability of the equipment and reduce the damage and maintenance costs caused by failures. Traditional fault warning methods are mainly based on rules and statistical models, and their accuracy and reliability are limited [8]. DL techniques have achieved good results in the fields of image and speech and have therefore been introduced to fault warning for marine diesel engines [9–11].

EGT is one of the most important heat parameters for diesel engines' combustion and dynamical properties in current operating conditions [12]. Because the exhaust gas temperature changes slowly and is less subject to disturbances, it is a strong indicator of failure. By predicting the vessel's EGT, it is possible to effectively demonstrate the healthy performance of the diesel engine [13]. This prediction method can help to identify potential faults early and take appropriate repair and maintenance measures to ensure the stability of the diesel engine.

There is less research on marine diesel engine fault warnings. Nguyen et al. introduced a new model for forecasting maintenance, which uses sensor measurements to estimate the likelihood of system failure in various time scales [14]. Jiang et al. presented a new algorithm to diagnose the failure of a diesel engine by using a double-assistant two-aided diagnosis algorithm based on the high angle relation and the transient nonsmooth state of the diesel engine [15]. Liu et al. introduced a kind of ship diesel engine fault alarm system based on CNN's characteristic extracting ability and BiGRU's time forecast ability [16]. Liu et al. combined long- and short-term neural network (LSTM) characteristic extracting ability and time sequence storage ability to build a forecast model for marine diesel engine EGT [17]. Patil et al. proposed a swarm-based Cauchy Particle Swarm Optimization (CPSO) method to search for the hyperparameters of LSTM to predict the optimal LSTM hyperparameters for temperature [18]. Xie et al. presented a multilayer perceptron (CGMP) based on the convolutional gated recurrent unit (GRU) to predict the South China Sea surface temperature. Convolution CGMP can efficiently catch adjacent effects in space, while GRU and MLP can deal with historical information efficiently [19]. Raptodimos et al. presented a nonlinear autoregression neural network (NARX) for the prediction of the exit temperature of a vessel's primary engine and showed that NARX was well-performing and stable under different assumptions [20].

Tan et al. developed a multistage prediction model for reheating steam in coal-fired power boilers based on an LSTM, which was able to accurately predict the reheat steam temperature within 2.5 min, providing an important reference for reheat steam temperature control [21]. Yan et al. used a data-driven hybrid approach to locomotive axle temperature prediction using particle swarm optimization and gravitational search algorithm (PSoGSA) to optimize and integrate bidirectional long- and short-term memory (LSTM) network units to achieve locomotive axle temperature prediction [22]. Cheliotis and his colleagues used the Expectation Behavior (EB) model and index-weighted moving mean to forecast the operation status of the primary engine [23]. Karatug and his team used an artificial nerve network (ANN) to determine the real-time running state of diesel engines [24]. Lazarus and his team used a fault tree and fault model to predict and locate faults in diesel engines [25].

In the current research, researchers usually classify diesel engine fault prediction and early warning methods into two types: condition prediction and condition classification [26]. Condition classification requires a lot of early failure information, but this method is not applicable, because early fault information on marine diesel engines is difficult to collect. Regular operation of the ship's diesel engine is an important guarantee of the safety of the ship and its crew. Thus, we present a CNN-BiLSTM-Attention prediction model, which predicts the operating status of marine diesel engines based on real-time operating data. Then, the normal EGT and the predicted EGT of the ship's diesel engine are used to calculate the Mahalanobias distance, and the function mapping method is used to construct the transformation function, to construct the ship's EGT health index model. The fault warning values and thresholds are set in conjunction with  $3\sigma$  criteria, where the fault warning values are used to indicate that a fault may occur, and the fault thresholds are

used to monitor whether the EGT monitoring index has crossed the limit and to warn of a fault. Finally, the validity of this approach has been experimentally proven.

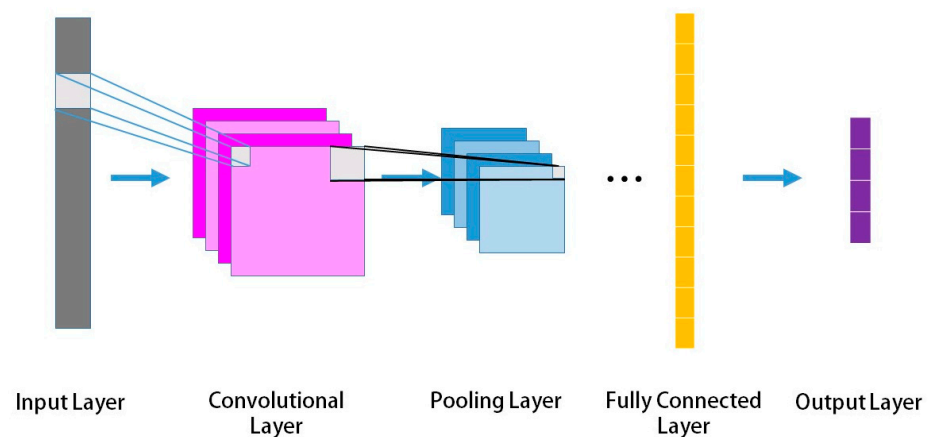
The rest of the thesis is as follows: This article introduces the theoretical part of DL used, including CNN, BiLSTM, and the attention mechanism (Attention). The data processing Section 3 introduces the sources of the prediction sample data and the data processing procedures such as data normalization and dimensionality reduction. The predictive model construction section presents how to adjust the model hyperparameters and the results of optimal hyperparameter selection. A fault warning model for marine diesel engine EGT is established, and the effectiveness of the model is verified through comparative experiments. In the fault warning Section 4, the methods for determining the monitoring index alarm values and thresholds are presented and experimentally verified. In the conclusion Section 5, the whole paper is summarized and prospects are presented.

## 2. Principles of Deep Learning Models

### 2.1. The Convolutional Neural Networks (CNN)

A convolutional neural network (CNN) is a set of layers consisting of an input layer, a convolutional layer, a pooling layer, and a full connection layer [27]. Compared to conventional neural networks, this model is characterized by a convolution and a pooling layer. In the convolution layer, every neural network is associated with only a few adjacent nodes, and every convolution is composed of several cells that form a rectangle [28]. After initialization of the convolution kernel, the convolution kernel obtains the proper weights. The convolution kernel has the advantage of minimizing the number of links between different layers while minimizing the risk of overfitting [29]. Pooling layers, also known as subsampling layers, are usually available in the form of mean pooling and maximum pooling. Pooling can be thought of as a special type of convolution [30]. The fully connected layer combines and integrates the features extracted by the convolution and pooling layers for classification, recognition, or prediction [31]. The CNN architecture is illustrated in Figure 1. Its convolution and pooling layers are calculated as shown in Formula (1).

$$\begin{cases} C_t = (N_{i1} - f_1 + 2p_1) / s_1 + 1 \\ P_t = (N_{i2} - f_2 + 2p_2) / s_2 + 1 \end{cases} \quad (1)$$



**Figure 1.** Structure of convolutional neural networks.

In the formula,  $C_t$  and  $P_t$  are the output matrix sizes of the convolution and pooling layer operations, respectively.  $N_{i1}$  and  $N_{i2}$  are the input matrix sizes.  $f_1$  and  $f_2$  are the weights of the convolution and pooling layers, respectively.  $p_1$  and  $p_2$  are the number of padding fills in the convolution and pooling layers, respectively.  $s_1$  and  $s_2$  are the step sizes of the convolution and pooling layers, respectively.

CNN has shown that it is possible to extract features from data by convolution, pooling, weight distribution, etc. CNN is widely used in image processing, including

two-dimensional convolutional neural networks (Convolution2d, 2DCNN), which can be applied in many fields such as image classification, object detection, image segmentation, etc. However, its dimension is different from that of time series, so it is not suitable for temporal sequence forecasts. To solve this problem, one-dimensional convolutional neural networks (Convolution1d, 1DCNN) are used to perform data mining on time-series data to extract local features of time-series data and improve the accuracy of prediction models [32].

### 2.2. Long- and Short-Term Memory Neural Network Unit (LSTM)

LSTM is a variant of Recurrent Neural Network (RNN) and is a solution proposed to overcome the short-term memory problem for processing sequential data [33]. Compared to traditional RNNs, LSTM has stronger memory and long-term dependency modeling capabilities. The key concept of LSTM is to introduce a structure called a “memory unit”, which stores and accesses the information at different time steps in the sequence [34]. The memory unit consists of an oblivion gate, an input gate, and an output gate, which control the flow of information through a series of mathematical operations [35]. The forgetting gate determines how much information in the memory state of the previous time step is retained during the current time step. The input gate determines which parts of the current input are to be stored at the current time step. The output of the input gate is multiplied by the input of the current time step to update the memory state. This gating mechanism allows the LSTM to selectively remember and forget information, allowing it to better handle long-term dependencies. The LSTM has performed well in many natural language processing tasks [36].

The overall architecture of the LSTM model is composed of an input word  $x_t$ , a cell state  $C_t$ , a temporary cell state  $D_t$ , a hidden layer state  $h_t$ , a forgetting gate  $f_t$ , a remembering gate  $i_t$ , and an output gate  $O_t$  at time  $t$ . The calculation works as follows: by forgetting the information and memorizing the new information in the cell state, the information that is useful for calculation at a later time is transferred, whereas the useless information is discarded, and the hidden level status  $h_t$  is the output at every time step. The input  $x_t$  of the current time step is first merged with the hidden state  $h_{t-1}$  of the previous time step to obtain  $[x_t, h_{t-1}]$ , which is activated by a sigmoid function to obtain the oblivious gate value  $f_t$ , where the sigmoid function is used to regulate the values flowing through the network. The input gate determines what information is to be added to the storage status from the input data in the present time step. The formula for generating the value  $i_t$  for the memory gate is almost identical to the formula for the forgetting gate, the only difference being the target on which it will subsequently act. The second formula for the input gate is the current cell state  $D_t$ . The input gate consists of a sigmoid activation function and a tanh function. Finally, the input gate multiplies the sigmoid output with the tanh output and adds the result to the current memory state. The structure and formula for the cell update are that the newly acquired forgetting gate value  $f_t$  is multiplied by the one obtained in the preceding time step, and then a result of the multiplication of the input gate value by the un-updated  $D_t$  acquired in the present time stage, and the final result is acquired as a portion of the input in the subsequent time step. The output gate value  $O_t$  is calculated in the same way as the forgetting gate and the input gate. The entire procedure of the output gating is the generation of the hidden state  $h_t$ . Figure 2 illustrates the configuration of long-term and short-term storage networks, and the calculation equation is illustrated in (2).

$$\begin{cases} f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \\ i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \\ D_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \\ C_t = f_t \times C_{t-1} + i_t \times D_t \\ O_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \\ h_t = o_t \times \tanh(C_t) \end{cases} \quad (2)$$

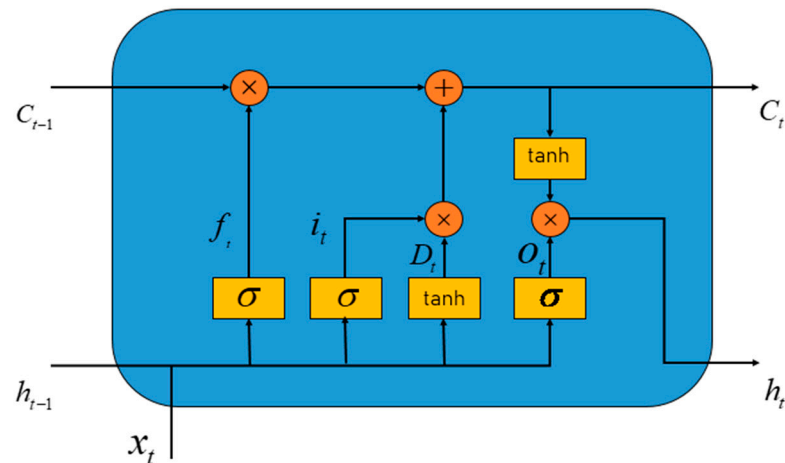


Figure 2. Structure of the long- and short-term memory network.

In the formula,  $\sigma$  is the sigmoid function,  $D_t$  is the candidate neuron,  $b_f, b_i, b_c, b_o$  are the offset matrices of the forgetting gate, memory gate, temporary cell, and output gate parts, respectively,  $W_f, W_i, W_c, W_o$  are the weight coefficients of the forgetting gate, memory gate, temporary cell, and output gate parts, respectively,  $C_{t-1}$  is the cell state input at time  $t-1$ ,  $h_{t-1}$  is the output at time  $t-1$ ,  $x_t$  is the input at the current time,  $h_t$  is the output and  $C_t$  is the cell state.

The LSTM is a powerful recurrent neural network structure that can effectively handle long-term dependencies in sequential data by introducing memory units and gating mechanisms [37].

### 2.3. Bidirectional Long- and Short-Term Memory Neural Network Unit (BiLSTM)

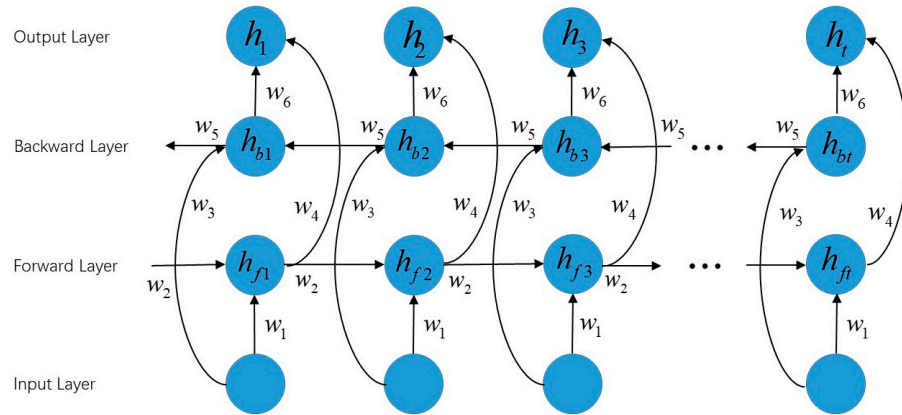
In the problem of predicting the EGT of a ship’s diesel engine, the operating state of the diesel engine is influenced not only by information from the historical moment but also from the future moment. However, the conventional LSTM can only use the information from the historical moment to predict the state output at the future moment, and it cannot encode the information from backward to front. To solve this problem, a bidirectional long- and short-term memory network (BiLSTM) is proposed [38]. BiLSTM consists of a combination of the forward LSTM and the reverse LSTM and can use information from both the past and the future to predict the operating state of the diesel engine at the current moment [39]. Through forward and backward propagation, BiLSTM can capture bidirectional semantic dependencies more comprehensively, resulting in a better understanding of the temporal data [40]. As a result, BiLSTM has better feature representation, stronger modeling capability, and better prediction representation in the marine diesel engine EGT prediction problem. BiLSTM is an effective model for improving the accuracy of a ship’s diesel engine EGT prediction.

As shown in Figure 3, the BiLSTM model plugs the same input sequence into two LSTMs, the forward LSTM hidden layer state  $h_{ft}$  and the reverse LSTM hidden layer  $h_{bt}$ , respectively, then connects the two hidden layers and plugs them together into the output layer for prediction to obtain the BiLSTM model output  $h_t$ . The formula is given in (3).

$$\begin{cases} h_{ft} = f(w_1x_t + w_2h_{ft-1} + b_{ft}) \\ h_{bt} = f(w_3x_t + w_5h_{bt-1} + b_{bt}) \\ h_t = g(w_4h_{ft} + w_6h_{bt} + b_{ot}) \end{cases} \quad (3)$$

In the formula,  $w_1$  is the weighting coefficient from the input level to the forward LSTM,  $w_2$  is the weighting coefficient between the forward LSTM cell layers,  $w_3$  is the weighting coefficient from the input level to the reverse LSTM,  $w_5$  is the weighting coefficient between the backward LSTM cell layers,  $w_4$  is the weighting coefficient from the forward LSTM to

the output layer,  $w_6$  is the weighting coefficient from the reverse LSTM to the output layer, and  $b_{ft}$ ,  $b_{bt}$ , and  $b_{ot}$  are the bias matrices of the respective parts.



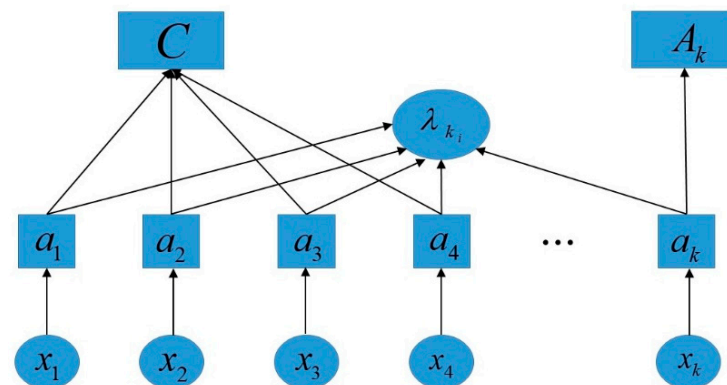
**Figure 3.** Structure of a two-way long- and short-term memory neural network.

2.4. Mechanisms of Attention (Attention)

Attention is a mechanism used to increase a model’s attention to input in different parts. It is used in many DL domains [41]. Its operation is the calculation of a relevance score for each element in the input sequence: for the current moment and then weighting the input sequence according to the score to produce a representation that focuses more on the important elements [42]. Thus, the model can selectively focus on the parts that are relevant to the current task, improving its expressiveness and ability to generalize. The introduction of an attention mechanism based on the BiLSTM network model can solve the long dependency problem, and so on [43]. In summary, the Attention allows the model to better focus on the key information by dynamically weighting the input sequence, thereby improving the predictive power of the model.

As shown in Figure 4, the input sequence is  $x_k$ , the hidden layer state of the input sequence is  $a_k$ , the attention weight of the hidden layer state of the historical input to the current input is  $\lambda_{k_i}$ , and the hidden layer state value of the final node of the final output is  $A_k$ . See (4) for the formula.

$$\begin{cases} \lambda_{k_i} = \frac{\exp(S_{k_i})}{\sum_{i=j}^{T_x} \exp(S_{k_i})} \\ S_{k_i} = v \tanh(W h_k + U h_i + b) \\ C = \sum_{i=j}^{T_x} \lambda_{k_i} a_i \end{cases} \quad (4)$$



**Figure 4.** Structure of the attention mechanism.

The last characteristic vector  $A_k$  is the hidden vector of the last node. See (5) for the formula.

$$A_k = A(C, a_k, x_k) \tag{5}$$

In the formula,  $S_{k_i}$  is a matrix measuring the degree of influence between data,  $\lambda_{k_i}$  is the weight data, and  $C$  is the weight after weighting.

### 2.5. CNN-BiLSTM-Attention Prediction Model

Since the traditional single neural network can no longer meet the contemporary needs of complexity and diversity of marine diesel engine fault warning, in this article, an HNN is proposed to predict the failure of a ship’s diesel engine. The hybrid model combines CNN, BiLSTM, and Attention. Among them, CNN can be used to extract the local features of the data through the convolving operation. BiLSTM can capture the long-range dependencies in the sequences, and by using the CNN-BiLSTM model can make full use of the local and global information to make predictions. Attention can automatically learn the association scores of different positions in the input sequence, which enables the prediction model to further improve the model’s expressiveness and generalization ability. The CNN-BiLSTM-Attention model can reduce the dependence on noise and irrelevant information, and it can improve the robustness and stability of the model. The bidirectional LSTM can solve the dependency problem in long sequences that cannot be handled by the unidirectional LSTM, and it can increase the precision and validity of the model. Figure 5 illustrates a CNN-BiLSTM-Attention prediction model.

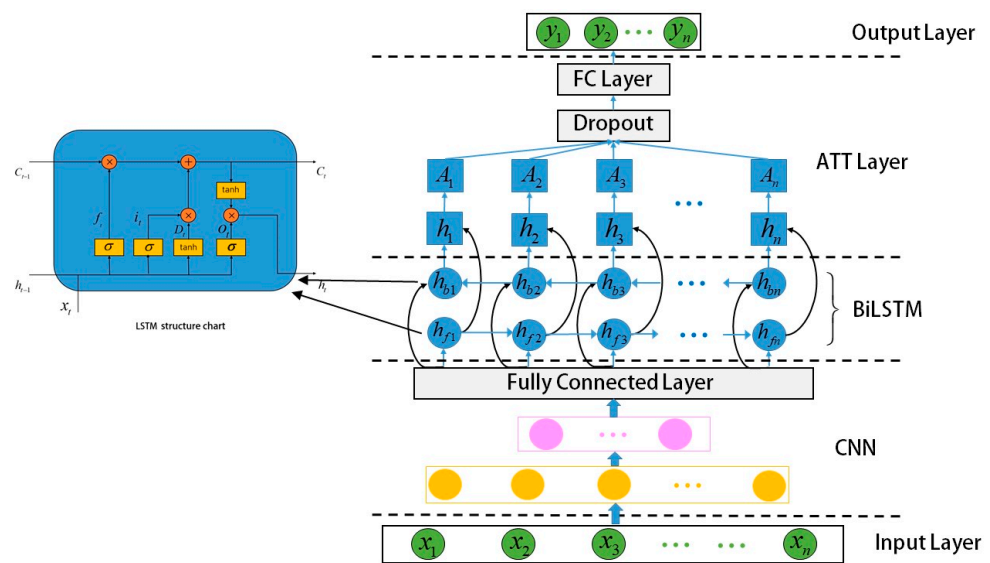


Figure 5. Structure of the CNN-BiLSTM-Attention prediction model.

## 3. Forecasting Process

### 3.1. Data Processing

This paper uses the actual operating data collected from a 6L34DF dual-fuel power generation diesel engine installed on an LNG carrier. The actual operating data were recorded using the principle of recording once per second. To decrease computation times, increase process speed, and decrease cabin noise impact, the 6 s interval was used to select a monitoring point for interval recording, and a total of 7200 surveillance sites were chosen, i.e., twelve hours of operating parameters. A total of eight operating parameters were selected for this diesel engine model, as can be seen in Table 1.

The data collected were taken with the ship’s diesel engine running under stable operating conditions, with only small fluctuations. The selected thermal parameters contain a lot of information about the condition of the diesel engine. The EGT reflects combustion efficiency and engine performance. If the EGT is too high, it may mean incomplete combus-

tion or problems with the diesel engine. The air cooler outlet temperature reflects whether the engine cooling water circulation is normal and whether the engine load is too high. The compressor speed reflects the load of the main engine. The gas inlet pressure reflects the working condition of the gas supply system and the stability of the gas quality. The EGT at the outlet of the compressor reflects the combustion efficiency of the ship’s main engine and the working condition of the compressor. The inlet pressure of the high-temperature water of the cylinder liner reflects the working condition of the ship’s cooling system, and the outlet temperature of the high temperature of the cylinder reflects the thermal load of the engine.

**Table 1.** Type of feature parameter.

Data Type	Symbol	Unit
EGT	$T_p$	°C
High-temperature water air cooler outlet temperature	$T_1$	°C
Speed of supercharger	$N$	r/min
Gas inlet pressure for diesel engines	$P_1$	MPa
A load of diesel engine	$L$	KW
Exhaust temperature at turbocharger outlet	$T_2$	°C
High-temperature water cylinder liner inlet pressure	$P_2$	MPa
High-temperature cylinder liner outlet temperature	$T_3$	°C

The obtained data were used to build a marine diesel engine EGT prediction data set, with 70% trained and 30% tested, to predict the EGT of the ship’s diesel engine and validate its effectiveness. To eliminate the influence of the magnitude and scale between different predictor parameters on the predictive effect, the collected sample data are normalized. The formula is shown in (6).

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{6}$$

In this equation,  $X_{norm}$  is the standard data,  $X$  is the raw sample data, and  $X_{max}$  and  $X_{min}$  are the maximal and minimal values of the sample data set, respectively.

### 3.2. Principal Component Analysis (PCA)

PCA is one of the most popular methods for reducing high-dimensional data to low-dimensional data. With PCA dimensionality reduction, the dimensionality of the data can be reduced while retaining the main information of the data. The reduced data set can be better visualized, processed, and analyzed while reducing computational complexity. However, some data may be lost due to dimension reduction, so it is necessary to consider the relationship between dimension reduction and information preservation.

As this paper acquires the data set of a ship diesel engine under normal operation as a multidimensional feature data set, the method of principal component analysis can simplify the model structure and improve the model convergence rate as well as the computational efficiency. The sample data collected were subjected to PCA using Python language code to perform dimensionality reduction on eight preselected thermal parameters to better understand the nature of the data. The formula is shown in (7).

$$\begin{cases} F_i = \sum_{j=1}^j W_{ij}X_j \\ W_{ij} = \frac{\theta_i}{\sqrt{\lambda_i}} \end{cases} \tag{7}$$

In the formula,  $F_i$  is the score of the  $i$ -th principal component,  $X_j$  is the sample parameter,  $W_{ij}$  is the weight of each variable of the principal component, and  $\theta_i$  is the coefficient corresponding to each variable in the component matrix.

According to the results of the PCA of the eight parameters in the original sample, the KMO and Bartlett’s test showed a KMO of 0.786 and a Sig of 0.000, indicating that



the degree to which the characteristic factors correlate with the parameters of the sample was high enough to meet the requirements of the PCA. Plots of principal component contributions and cumulative contributions are shown in Figure 6.

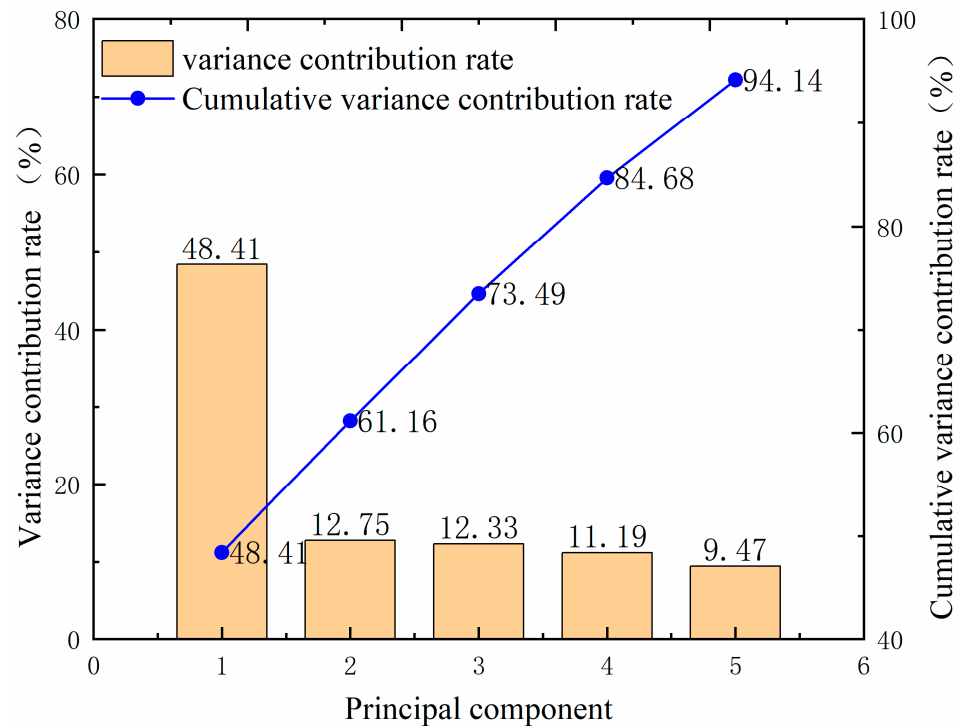


Figure 6. The plot of principal component contribution and cumulative contribution.

The principal component analysis was completed to extract the principal components with eigenvalues greater than 0.5 and a variance contribution rate of approximately 95%. As illustrated in Figure 6, the cumulative variance contribution rate of the top five main constituents was 94.145%, indicating that these five principal components can represent 94.145% of the information of the original eight operational parameters and can represent the total data of the selected samples. Therefore, this paper uses EGT  $T_p$ , high-temperature water air cooler outlet temperature  $T_1$ , supercharger speed  $N$ , engine load  $L$ , and high-temperature water cylinder liner outlet temperature  $T_3$  for marine diesel engine EGT prediction study, which reduces data redundancy and improves the diagnostic efficiency of the model.

### 3.3. Martens Distance Screening for Outliers

The Mahalanobias distance is a method used to measure the distance between multi-variate data, taking into account the correlation between dimensions. In this experiment, the Mahalanobias distance was used to filter outliers, i.e., those outliers that are significantly different from other data points. For the data after the PCA dimensionality reduction process, the Mahalanobias distance between the EGT and the other four characteristic parameters was determined by calculating the Mahalanobias distance. The formula is shown in (8).

$$D_{MH} = \sqrt{(x_i - y_j)S^{-1}(x_i - y_j)} \tag{8}$$

In the formula,  $D_{MH}$  is the computed Mahalanobias distance,  $x_i$  and  $y_j$  are column vectors, and  $S^{-1}$  is the inverse matrix of the covariance.

Then, the Chi-square test was used to filter outliers for the EGT, with a total degree of freedom of 4 for the sample data set, and a critical value significance level of 0.005 was determined for the martingale distance. The formula is shown in (9).

$$x^2 = \sum_{i=1}^k \frac{(A_i - np_i)^2}{np_i} \tag{9}$$

In the formula,  $x^2$  is the cardinality of freedom,  $A_i$  is the cell observation, and  $p_i$  is the cell premise probability.

According to the Chi-square test, the Chi-square degree of freedom was calculated to be 14.86026, so points with a martingale distance greater than 14.86026 were considered outliers. A total of 80 outliers were removed from the sample data by screening the outliers using the martingale distance, thus eliminating inaccurate and unrealistic sample data due to the particular working environment of the ship and improving the authenticity of the failure prediction.

To address the problem of discontinuity in the time series after outlier removal, we will compensate for the abnormal point using the multi-order Lagrangian value method. The formula is shown in (10).

$$x_t = \frac{\sum_{i=1}^{n_1} x_{t-i} + \sum_{k=1}^{n_2} x_{t+k}}{n_1 + n_2} \tag{10}$$

In the formula,  $x_t$  is the lost value at time  $t$ ,  $n_1$  is the advance cycle, and  $n_2$  is the reverse cycle.

### 3.4. Indicators for Model Evaluation

Evaluating the model can help determine the accuracy and effectiveness of the model. First of all, it is necessary to study the generalization of a model to compare the various models and determine which one is superior or not. Next, we can progressively improve the performance of our models with these metrics. In this paper, four prediction evaluation metrics commonly used in machine learning are selected, namely, mean-squared error (*MSE*), root-mean-squared error (*RMSE*), mean absolute error (*MAE*), and mean absolute percentage error (*MAPE*). The formula is shown in (11).

$$\left\{ \begin{array}{l} MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \\ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \\ MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \\ MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \end{array} \right. \tag{11}$$

In the formula,  $\hat{y}_i$  is the predicted value and  $y_i$  is the actual value.

The *MSE* is a measure of the extent to which the expected amount differs from the actual amount. The *RMSE* is the square root of the ratio of square deviation of predicted value from true value to ratio of observed value,  $n$ . The *MAE* is the absolute difference between the individual measurements and the mean. The *MAE* eliminates the problem of mutually exclusive errors, so it is possible to accurately measure the size of the true prediction error. The reason that *MAPE* can describe accuracy is that it is often used as a statistical indicator of forecast accuracy itself, for example, in time-series forecasting.

### 3.5. Hybrid Neural Network Prediction Models

Since the input level of the HNN depends on the dimension of the input parameters, there are five input parameters, so the input level has five neurons. The model ultimately predicts that the output result will be the EGT only, so the number of neurons in the output layer is one. The HNN is made up of a CNN frame, a BiLSTM frame, and an Attention frame: the main hyperparameters of CNN and convolution and layer pooling, etc.; the main hyperparameters of BiLSTM are num layers, input size, seq len, batch size, epochs, learning rate, loss, optimizer, etc.; the main hyperparameters of the Attention layer are dropout, etc.

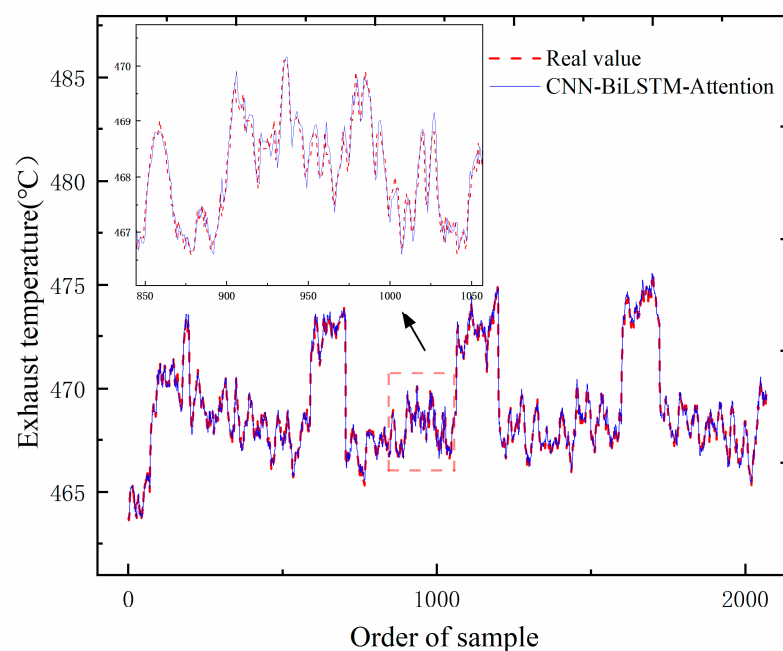
The above parameters were debugged several times, and the parameters were selected using the control variable method. MAE, RMSE, MSE, and MAPE were used as evaluation indicators for the combined prediction model, and the optimal parameters were selected according to Table 2, and the optimal CNN-BiLSTM-Attention model was trained.

**Table 2.** CNN-BiLSTM-Attention network structure parameters.

Network Parameter Name	Optimal Parameter Values
CNN layers	1
Pooling layer	Max Pool
Num layers	1
Input size	5
Hidden size	64
Seq len	5
Batch size	100
Epochs	1000
Learning rate	0.0005
Dropout	0.4
Loss	MSE loss
Optimizer	Adam

### 3.6. Analysing the Predicted Results

In this thesis, we input the sample data from the finished data into the training prediction model. The prediction result graph is shown in Figure 7.



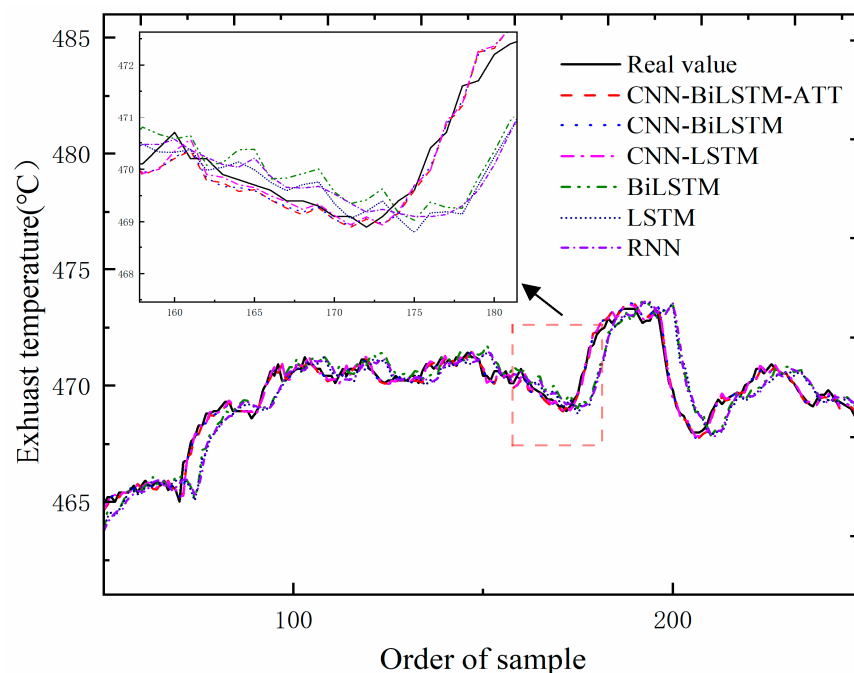
**Figure 7.** CNN-BiLSTM-Attention prediction result graph.

As shown in Figure 7, the operation of a diesel engine is fairly stable over this period, there is little difference between expected and actual EGT, the range of data waves is within the normal operating range of the diesel engine, the prediction and the actual data were in good agreement, and there was a lot of overlap. The assessment indicators for forecast results are presented in Table 3. The results show that the CNN-BiLSTM-Attention prediction model proposed in this thesis has high precision and accuracy.

**Table 3.** Forecast indicator assessment summary.

Evaluation Indicators	Optimal Model Evaluation Results
MAE	0.1988029
MSE	0.0958985
MAPE	0.0004238
RMSE	0.3096749

To prove that the HNN combined prediction model proposed in this article is real and effective, and can have high reliability and accuracy in practical parameter prediction work, this paper uses experimental comparison to verify the model’s accuracy. The CNN-BiLSTM-Attention prediction model is compared with five prediction models, namely, RNN, LSTM, BiLSTM, CNN-LSTM, and CNN-BiLSTM. Before performing the experimental comparison, ensure that the input data set of each model is the same standard data set after data processing is completed, and ensure that the hyperparameters of the other models are the same as those of the CNN-BiLSTM-Attention prediction model. Ensure that the input parameters of the models and the conditions of the experimental models are as consistent as possible during the comparison experiments to better compare the prediction accuracies, model training speeds, etc. of the different models. By comparing the performance of the models under different evaluation metrics, it can accurately reflect the strengths and weaknesses of the models in terms of actual prediction work. A comparison of the predicted results from different models is shown in Figure 8.



**Figure 8.** Comparison of prediction results of different models.

Indicators used to evaluate the different forecasting models are presented in Table 4.

**Table 4.** Summary of evaluation indicators for different model prediction results.

Predictive Models	MSE	RMSE	MAE	MAPE
RNN	0.1216556	0.3487916	0.2203313	0.0004698
LSTM	0.1131162	0.3363276	0.2211981	0.0004717
BiLSTM	0.1069281	0.3217019	0.2138589	0.0004559
CNN-LSTM	0.1057659	0.3252167	0.2135331	0.0004552
CNN-BiLSTM	0.1034921	0.326999	0.2092867	0.0004461
CNN-BiLSTM-Attention	0.0958982	0.3096749	0.1988029	0.0004238

As shown in Figure 8, the CNN-BiLSTM-Attention HNN prediction model has the best fit and the highest similarity to the actual sample data than other prediction models. As can be seen from Table 4, the evaluation index of the CNN-BiLSTM-Attention prediction model has the lowest error compared to others, indicating a small discrepancy between the predicted result of a model and the actual value. In summary, the CNN-BiLSTM-Attention has a high prediction accuracy and can meet the basic application in predicting the EGT time series of marine diesel engines.

#### 4. Fault Early Warning Research

##### 4.1. Setting of the Monitoring Index

The EGT of a ship’s diesel engine depends on several factors. In this paper, through feature selection, five feature parameters with the highest correlation with the EGT are selected as the input parameters of the HNN, to obtain the predicted value of the EGT. A real-time monitoring method of the ship’s diesel engine EGT is proposed as a rapid assessment of the integrity, stability, and accuracy of the ship’s diesel engine, and to ensure that the failure warning of the ship’s diesel engine is achieved. The proposed ship’s diesel engine EGT fault warning method is based on the function mapping metric of the Mahalanobia distance, which first calibrates the sample similarity between the predicted and actual values by calculating the Mahalanobia distance between the predicted and actual values. The formula for its calculation is given in Equation (12).

$$D_{MH} = \sqrt{(x_i - y_j)S^{-1}(x_i - y_j)} \tag{12}$$

In the formula,  $D_{MH}$  is the computed Mahalanobia distance,  $x_i$ , and  $y_j$  are column vectors, and  $S^{-1}$  is the inverse matrix of the covariance.

Due to the high uncertainty of the magnitude of Mahalanobia’s distance, a function mapping is considered to fix it within a certain range. The invention constructs a monitoring indicator function  $MF$ , as in Equation (13), to map the Mahalanobia distance to the interval  $[0, 1]$ , and uses this method to accurately and intuitively monitor the diesel engine operating conditions. Based on the monitoring indicator  $MF$ , the operating condition of the diesel engine can be monitored more intuitively and effectively by setting the alarm value and the threshold value of the EGT, wherein the alarm value can provide the reference value of the possible abnormal operation of the ship’s diesel engine. The threshold value can provide the reference value of the failure of the ship’s diesel engine.

$$MF = \frac{2}{1 + e^{\alpha D_{MH}}} \tag{13}$$

In the formula,  $\alpha$  is the adjustment factor calculated according to Equation (14).

$$\alpha = -\frac{\ln \frac{MF_Q}{2 - MF_Q}}{\bar{D}_{MH}} \tag{14}$$

In the formula,  $\bar{D}_{MH}$  is the mean Mahalanobia distance,  $MF_Q$  is the confidence factor for the normal EGT, and in this paper, we make  $MF_Q = 0.95$ .

According to Equation (13), the calculated martingale distance between the predicted and actual values of the EGT is used as an input with a definition range of  $(0, +\infty)$ , but the value range of  $MF$  is  $(0, 1)$ . The output value  $MF$  is positively correlated with the health status of the monitored EGT indicator so that the state change of the EGT can be consistently and accurately monitored using this mathematical conversion method.

After obtaining the EGT of  $MF$ , this paper better characterizes the health of the EGT of a marine diesel engine by setting a monitoring index warning value and a threshold value to alert the reference value for a possible abnormal EGT, and a threshold value for an abnormal EGT criterion was used to set the EGT warning values and thresholds. The  $3\sigma$  criterion was used to set the EGT warning values and thresholds. First, the calculated  $D_{MH}$  under normal operation of the marine diesel engine was tested for normality of distribution. As shown in Figure 9, the Mahalanobia distance between predicted and actual values lies between 0 and 8, and most of them lies between 0 and 4.

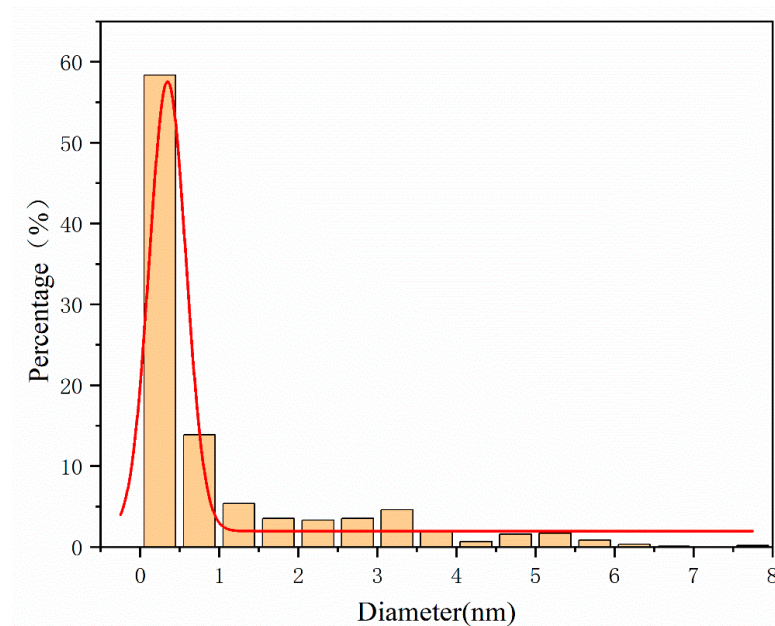


Figure 9. Map of the normal distribution of the Mahalanobia distance test.

Then, the alarm value and the threshold value of the Mahalanobia distance are determined. The formula is shown in (15).

$$\begin{cases} K_{2\sigma} = \mu_{D_{MH}} + 2\lambda_{D_{MH}} \\ K_{3\sigma} = \mu_{D_{MH}} + 3\lambda_{D_{MH}} \end{cases} \quad (15)$$

In the formula,  $\mu_{D_{MH}}$  is the average of Mahalanobia’s distance, and  $\lambda_{D_{MH}}$  is Mahalanobia’s standard deviation.

Finally, the alarm value  $MF_w$  and the threshold  $MF_f$  for the monitoring index are calculated according to and using Equation (16).

$$\begin{cases} MF_w = \frac{2}{1+e^{\alpha \cdot K_{2\sigma}}} \\ MF_f = \frac{2}{1+e^{\alpha \cdot K_{3\sigma}}} \end{cases} \quad (16)$$

The warning value and threshold value of the monitoring indices calculated by the function construction method can satisfy us to make a clear distinction between the normal or abnormal operating state of the EGT of the ship’s diesel engine.

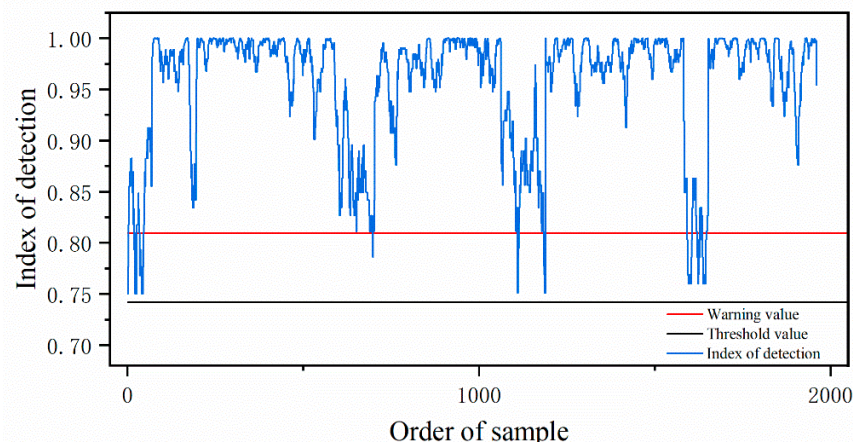
Based on the method proposed above, the predicted EGT values obtained from the CNN-BiLSTM-Attention HNN predictions in this article were evaluated against the actual values of the sample for the monitoring index status. The monitoring indices for the sample

data, as well as the healthy operating conditions alarm values and thresholds for the EGT, were calculated. The healthy operating alarm values and thresholds are shown in Table 5.

**Table 5.** Alarms and thresholds for monitoring indices.

Index Name	Numerical Values	Note Description
Value of the alarm	0.809632	Monitoring concern
Value of the threshold	0.741913	Monitoring of the bottom line

A graph of the EGT monitoring index for a diesel engine in normal operation is shown in Figure 10.



**Figure 10.** Normal operating conditions diesel engine exhaust temperature control indicator chart.

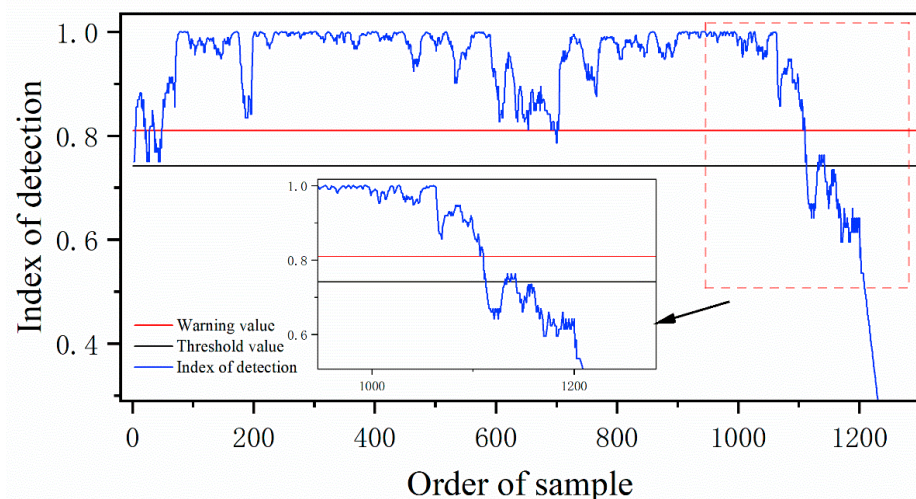
Figure 10 shows that the ship diesel engine EGT prediction value under normal working conditions of the monitoring index has been higher than the monitoring index threshold. Although the monitoring index exceeded the monitoring index warning value at some points in time, the monitoring threshold was not exceeded, and therefore, no marine diesel engine failure alarms occurred. During the monitoring intervals shown in the figure, the monitoring index values fluctuated smoothly within the normal range most of the time, indicating that the vessel was in normal sailing condition. It shows that the ship’s diesel engine failure prediction and early warning method proposed can meet the monitoring of the ship’s diesel engine EGT in the operation of state characteristics.

4.2. Experimental Verification of the Fault Warning Function

Since the model data are the actual operating data of the ship’s diesel engine, it is difficult to obtain fault data in the actual operation of marine diesel engines. Therefore, to validate the failure alert approach presented in this article, a manual linear adjustment of the data set was used to simulate the state data of a ship’s diesel engine at fault. The EGT of a ship’s diesel engine is mainly affected by combustion efficiency, load size, intake air temperature, and cooling system effect. This paper simulates the abnormally high EGT fault condition of a ship’s diesel engine caused by the clogging of the high-temperature water air cooler water pipe, which makes the air cooler unable to dissipate heat effectively and reduces the cooling effect of the air cooler. The monitoring index diagram of the abnormal increase in EGT is shown in Figure 11.

As shown in Figure 11, the high-temperature water air cooler blockage simulated by the manual linear adjustment data set causes the EGT to rise abnormally until it leads to the high temperature of the ship’s diesel engine and the shutdown fault. As shown in Figure 11, the ship’s diesel engine EGT is in a relatively stable operating condition before the fault is introduced (sample sequence before 1110). The 0–50 sequence points fluctuate, but the fluctuation range is within the monitoring index threshold, and after

a short fluctuation, it returns to the monitoring index warning value, so no fault alarm occurs during this period of operation. The monitoring index fluctuates steadily from 50 to 1100 and rarely exceeds the warning value, indicating that the ship's diesel engine is in a stable operating condition. After the artificial introduction of faults (sample sequence after 1100), the monitoring index suddenly displayed large fluctuations and was extremely unstable; after crossing the monitoring index warning value, it did not fluctuate to the normal working range, but directly crossed the monitoring index threshold, fluctuated up and down for some time below the monitoring index threshold, and then fell off the cliff, resulting in the shutdown of the diesel engine if the temperature is too high. If the monitoring index exceeds the monitoring index warning value line, but does not exceed the monitoring index threshold line, a diesel engine early failure warning is generated. If the monitoring index continues to fluctuate below the threshold line, a diesel engine exhaust temperature abnormality fault alarm is generated. It can be seen that the marine diesel engine fault warning method proposed in this paper can satisfy both the real-time monitoring and evaluation of the exhaust temperature of the marine diesel engine, and it can also issue a fault warning alert for the exhaust temperature abnormality in a future period of time. The method has high precision and sensitivity and can basically meet the needs of modern marine diesel engine exhaust temperature fault warning.



**Figure 11.** Abnormally high exhaust temperature monitoring index table.

## 5. Conclusions

This article presents an approach to the prediction of EGT in marine diesel engines using a CNN-BiLSTM-Attention prediction model. The main conclusions are as follows:

The normalization method is used in the data processing to eliminate the influence of the magnitude and order of the different prediction parameters. PCA is used to extract the characteristic parameters that have a greater influence on diesel EGT, simplify the model structure, and increase the model convergence rate and computational efficiency. The Mahalanobis distance outlier screening method can consider multidimensional data, adaptively identify outliers, and increase the precision of outlier monitoring to increase the precision of sample data.

The CNN-BiLSTM-Attention prediction model proposed in this article combines CNN, BiLSTM, and Attention. Among them, CNN can extract features from time-series data; BiLSTM can automatically learn the extracted time-series features; and Attention assigns weights to time-series features, which can better capture the sequence of fault-related features. It overcomes the limitations of the previous single neural network and can extract the temporal and spatial features of the EGT of the ship's diesel engine more comprehensively and improve the prediction accuracy.



By conducting model comparison experiments, the prediction model proposed in this article was compared with RNN, LSTM, BiLSTM, CNN-LSTM, and CNN-BiLSTM prediction models to analyze the prediction results. According to the prediction model evaluation indices, the CNN-BiLSTM-Attention prediction model proposed in this article has higher prediction accuracy, indicating that the HNN prediction model proposed in this paper has certain advantages in predicting time series.

Addressing the issue of early warning of faults, this paper adopts the method of converting a mathematical function model by combining the Mahalanobia distance to construct the EGT monitoring index value of a ship's diesel engine and uses the principle to set the alarm value and threshold value of the EGT monitoring index, with the alarm value used for fault warning indication and the threshold value used for fault warning. The proposed fault warning method can be verified through fault experiments, in that it can continuously monitor the operating status of diesel engines and provide a timely warning when an abnormal operation occurs, which can meet the health management needs of modern marine diesel engines.

To summarize, the CNN-BiLSTM-Attention-based marine diesel engine fault warning model has high prediction accuracy and early warning lead time, which can provide interpretable fault prediction results. The model can capture the characteristic patterns and provide early warning before the fault occurs, which is conducive to taking measures such as repair or spare parts in advance, reducing downtime and maintenance costs. It has important application value for fault prevention and the health management of marine diesel engines.

In our future work, we will not limit our research to one monitoring parameter, but also study other monitoring parameters, such as turbocharger EGT, to ensure the complete prediction results of diesel engines. The selection of prediction methods and the length of the prediction period are also worthy of further study, with a focus on improving the prediction capability for long time series. We will improve the model function and implement the fault classification function based on fault prediction, to provide the fault category together with the fault warning and facilitate subsequent repair and maintenance.

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

**Funding:** This research was funded by the National Key R&D Program of China (grant number 2022YFB4301400), and the High-technology Ship Research Program (grant number CBG3N21-3-3).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation of this paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Zhang, H.; Fei, T.M.; Guan, W.; Zhang, S.W.; Jin, Z.P.; Tang, W.K. Research on visual 3D assembly process design and simulation for marine diesel engine. *Clust. Comput. J. Netw. Softw. Tools Appl.* **2019**, *22*, S5505–S5519. [[CrossRef](#)]
2. Zhang, H.B.; Cui, Y.; Liang, G.; Li, L.T.; Zhang, G.Y.; Qiao, X.Q. Fatigue life prediction analysis of high-intensity marine diesel engine cylinder head based on fast thermal fluid solid coupling method. *J. Braz. Soc. Mech. Sci. Eng.* **2021**, *43*, 1–15. [[CrossRef](#)]
3. Lan, F.; Jiang, Y.; Wang, H.Y.; Publishing, I.O.P. Performance Prediction Method of Prognostics and Health Management of Marine Diesel Engine. In Proceedings of the 2020 3rd International Conference on Applied Mathematics, Modeling and Simulation, Beijing, China, 27–29 March 2020. [[CrossRef](#)]

4. Sana, S.S. Optimum buffer stock during preventive maintenance in an imperfect production system. *Math. Methods Appl. Sci.* **2022**, *45*, 8928–8939. [[CrossRef](#)]
5. Duan, Y.P.; Li, Z.; Tao, X.M.; Li, Q.; Hu, S.Z.; Lu, J.H. EEG-Based Maritime Object Detection for IoT-Driven Surveillance Systems in Smart Ocean. *IEEE Internet Things J.* **2020**, *7*, 9678–9687. [[CrossRef](#)]
6. Gao, Z.L.; Jiang, Z.N.; Zhang, J.J. Identification of power output of diesel engine by analysis of the vibration signal. *Meas. Control.* **2019**, *52*, 1371–1381. [[CrossRef](#)]
7. Zhu, Y.Q.; Wu, P.H.J.; Liu, F.; Kanagavelu, R.; Society, I.C. Disk Failure Prediction for Software-Defined Data Centre (SDDC). In Proceedings of the 2021 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech), Virtual, 25–28 October 2021; pp. 264–268. [[CrossRef](#)]
8. Liu, J.R.; Yang, G.T.; Li, X.L.; Hao, S.M.; Guan, Y.M.; Li, Y.Q. A deep generative model based on CNN-CVAE for wind turbine condition monitoring. *Meas. Sci. Technol.* **2023**, *34*, 035902. [[CrossRef](#)]
9. Jan, B.; Farman, H.; Khan, M.; Imran, M.; Ul Islam, I.; Ahmad, A.; Ali, S.; Jeon, G. Deep learning in big data Analytics: A comparative study. *Comput. Electr. Eng.* **2019**, *75*, 275–287. [[CrossRef](#)]
10. Han, S.H.; Rahim, T.; Shin, S.Y. Detection of Faults in Solar Panels Using Deep Learning. In Proceedings of the 20th International Conference on Electronics, Information, and Communication (ICEIC), Jeju, South Korea, 31 January–3 February 2021; IEEE: Piscataway, NJ, USA, 2021.
11. Shao, S.; McAleer, S.; Yan, R.; Baldi, P. Highly Accurate Machine Fault Diagnosis Using Deep Transfer Learning. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2446–2455. [[CrossRef](#)]
12. Kumar, A.; Srivastava, A.; Goel, N.; McMaster, J. Exhaust gas temperature data prediction by autoregressive models. In Proceedings of the 2015 IEEE 28th Canadian Conference on Electrical and Computer Engineering (CCECE), Halifax, NS, Canada, 3–6 May 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 976–981.
13. Zhang, Y.F.; Liu, P.P.; He, X.; Jiang, Y.P. A prediction method for exhaust gas temperature of marine diesel engine based on LSTM. In Proceedings of the 2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT), Weihai, China, 14–16 October 2020; pp. 49–52.
14. Nguyen, K.T.P.; Medjaher, K. A new dynamic predictive maintenance framework using deep learning for failure prognostics. *Reliab. Eng. Syst. Saf.* **2019**, *188*, 251–262. [[CrossRef](#)]
15. Jiang, J.J.; Li, H.; Mao, Z.W.; Liu, F.C.; Zhang, J.J.; Jiang, Z.N.; Li, H. A digital twin auxiliary approach based on adaptive sparse attention network for diesel engine fault diagnosis. *Sci. Rep.* **2022**, *12*, 251–262. [[CrossRef](#)]
16. Liu, B.; Gan, H.B.; Chen, D.; Shu, Z.P. Research on Fault Early Warning of Marine Diesel Engine Based on CNN-BiGRU. *J. Mar. Sci. Eng.* **2023**, *11*, 56. [[CrossRef](#)]
17. Liu, Y.; Gan, H.B.; Cong, Y.J.; Hu, G.T. Research on fault prediction of marine diesel engine based on attention-LSTM. *Proceed. Inst. Mech. Eng. Part M J. Eng. Mar. Environ.* **2023**, *237*, 508–519. [[CrossRef](#)]
18. Patil, M.M.; Rekha, P.M.; Solanki, A.; Nayyar, A.; Qureshi, B. Big Data Analytics Using Swarm-Based Long Short-Term Memory for Temperature Forecasting. *Cmc-Comput. Mater. Contin.* **2022**, *71*, 2347–2361. [[CrossRef](#)]
19. Xie, J.; Ouyang, J.M.; Zhang, J.Y.; Jin, B.G.; Shi, S.X.; Xu, L.Y. An Evolving Sea Surface Temperature Predicting Method Based on Multidimensional Spatiotemporal Influences. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 2347–2361. [[CrossRef](#)]
20. Raptodimos, Y.; Lazakis, I. Application of NARX neural network for predicting marine engine performance parameters. *Ships Offshore Struct.* **2020**, *15*, 443–452. [[CrossRef](#)]
21. Tan, P.; Zhu, H.Y.; He, Z.Q.; Jin, Z.Y.; Zhang, C.; Fang, Q.Y.; Chen, G. Multi-Step Ahead Prediction of Reheat Steam Temperature of a 660 MW Coal-Fired Utility Boiler Using Long Short-Term Memory. *Front. Energy Res.* **2022**, *10*, 443–452. [[CrossRef](#)]
22. Yan, G.X.; Yu, C.Q.; Bai, Y. A New Hybrid Ensemble Deep Learning Model for Train Axle Temperature Short Term Forecasting. *Machines* **2021**, *9*, 312. [[CrossRef](#)]
23. Cheliotis, M.; Lazakis, I.; Theotokatos, G. Machine learning and data-driven fault detection for ship systems operations. *Ocean. Eng.* **2020**, *216*, 312. [[CrossRef](#)]
24. Karatu, C.; Arslanoglu, Y. Development of condition-based maintenance strategy for fault diagnosis for ship engine systems. *Ocean. Eng.* **2022**, *256*, 111515. [[CrossRef](#)]
25. 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](#)]
26. Dashti, R.; Daisy, M.; Mirshekali, H.; Shaker, H.R.; Aliabadi, M.H. A survey of fault prediction and location methods in electrical energy distribution networks. *Measurement* **2021**, *184*, 404–415. [[CrossRef](#)]
27. Lee, L.; Xie, L.J.; Zhang, D.; Yu, B.; Ge, Y.F.; Lin, F.C. Condition Assessment of Power Transformers Using a Synthetic Analysis Method Based on Association Rule and Variable Weight Coefficients. *IEEE Trans. Dielectr. Electr. Insul.* **2013**, *20*, 2052–2060.
28. Tek, F.B.; Cam, I.; Karli, D. Adaptive convolution kernel for artificial neural networks? *J. Vis. Commun. Image Represent.* **2021**, *75*, 2052–2060. [[CrossRef](#)]
29. He, J.C.; Li, L.; Xu, J.C. Approximation properties of deep ReLU CNNs. *Res. Math. Sci.* **2022**, *9*, 103015. [[CrossRef](#)]
30. Vigneron, V.; Maaref, H.; Syed, T.Q. A New Pooling Approach Based on Zeckendorf’s Theorem for Texture Transfer Information. *Entropy* **2021**, *23*, 279. [[CrossRef](#)]

31. Zheng, T.Y.; Wang, Q.; Shen, Y.; Lin, X.T. Gradient rectified parameter unit of the fully connected layer in convolutional neural networks. *Knowl. Based Syst.* **2022**, *248*, 279. [[CrossRef](#)]
32. Chen, Z.S.; Zhang, L.K.; Hua, J.M.; Kim, B.; Li, K.; Xue, X.Y. A framework of data-driven wind pressure predictions on bluff bodies using a hybrid deep learning approach. *Meas. Control.* **2023**, *56*, 237–256. [[CrossRef](#)]
33. Mishra, A.; Tripathi, K.; Gupta, L.; Singh, K.P. Long Short-Term Memory Recurrent Neural Network Architectures for Melody Generation. In Proceedings of the Soft Computing for Problem Solving, Nagpur, India, 26–27 November 2021; pp. 41–55.
34. Li, Q.Y.; Wang, B.; Jin, J.; Wang, X.Y. Comparison of CNN-Uni-LSTM and CNN-Bi-LSTM based on single-channel EEG for sleep staging. In Proceedings of the 2020 5th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS 2020), Okinawa, Japan, 18–20 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 76–80.
35. Shu, X.B.; Tang, J.H.; Qi, G.J.; Liu, W.; Yang, J. Hierarchical Long Short-Term Concurrent Memory for Human Interaction Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 1110–1118. [[CrossRef](#)]
36. Kent, D.; Salem, F. Performance of Three Slim Variants of The Long Short-Term Memory (LSTM) Layer. In Proceedings of the 2019 IEEE 62ND International Midwest Symposium on Circuits and Systems (Mwscas), Dallas, TX, USA, 4–7 August 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 307–310.
37. Yang, H.F.; Hu, J.J.; Cai, J.H.; Wang, Y.P.; Chen, X.; Zhao, X.J.; Wang, L.L. A New MC-LSTM Network Structure Designed for Regression Prediction of Time Series. *Neural Process. Lett.* **2023**. [[CrossRef](#)]
38. Shi, H.F.; Miao, K.; Ren, X.C. Short-term load forecasting based on CNN-BiLSTM with Bayesian optimization and attention mechanism. *Concurr. Comput. Pract. Exp.* **2021**, *35*, e6676. [[CrossRef](#)]
39. De Lhoneux, M.; Ballesteros, M.; Nivre, J.; Assoc Computat, L. Recursive Subtree Composition in LSTM-Based Dependency Parsing. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Naacl Hlt 2019), Minneapolis, MN, USA, 4 June 2019; Volume 1, pp. 1566–1576.
40. Li, X.T.; Wang, H.L.; Xiu, P.F.; Zhou, X.Y.; Meng, F.H.; Soc, I.C. Resource Usage Prediction Based on BiLSTM-GRU Combination Model. In Proceedings of the 2022 IEEE 13th International Conference on Joint Cloud Computing (JCC 2022), Fremont, CA, USA, 15–18 August 2022; pp. 9–16.
41. Li, X.L.; Yuan, A.H.; Lu, X.Q. Vision-to-Language Tasks Based on Attributes and Attention Mechanism. *IEEE Trans. Cybern.* **2021**, *51*, 913–926. [[CrossRef](#)] [[PubMed](#)]
42. Wu, J.J.; Yu, Z.T.; Gao, S.X.; Guo, J.J.; Song, R. Chinese-Vietnamese News Documents Summarization Based on Feature-related Attention Mechanism. In Proceedings of the Computer Supported Cooperative Work and Social Computing, CHINESECSCW 2019, Kunming, China, 16–18 August 2019; pp. 526–539.
43. Du, J.; Cheng, Y.Y.; Zhou, Q.A.; Zhang, J.M.; Zhang, X.Y.; Li, G. Power Load Forecasting Using BiLSTM-Attention. In Proceedings of the 2019 5th International Conference on Environmental Science and Material Application, Singapore, 21–23 June 2019.

**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.