Article

Multiple Feature Extraction Long Short-Term Memory Using Skip Connections for Ship Electricity Forecasting

Ji-Yoon Kim and Jin-Seok Oh

1 Contents AI Research Center, Romantique, 27 Daeyeong-ro, Busan 49227, Republic of Korea; rlawldbs2918@nate.com
2 Division of Marine System Engineering, Korea Maritime and Ocean University, 727 Taejong-ro, Yeong-do-gu, Busan 49112, Republic of Korea
* Correspondence: ojs@kmou.ac.kr; Tel.: +82-410-4283

Abstract: The power load data of electric-powered ships vary with the ships’ operational status and external environmental factors such as sea conditions. Therefore, a model is required to accurately predict a ship’s power load, which depends on changes in the marine environment, weather environment, and the ship’s situation. This study used the power data of an actual ship to predict the power load of the ship. The research on forecasting a ship’s power load fluctuations has been quite limited, and the existing models have inherent limitations in predicting these fluctuations accurately. In this paper, a multiple feature extraction (MFE)-long short-term memory (LSTM) model with skip connections is introduced to address the limitations of existing deep learning models. This novel approach enables the analysis and forecasting of the intricate load variations in ships, thereby facilitating the prediction of complex load fluctuations. The performance of the model was compared with that of a previous convolutional neural network-LSTM network with a squeeze and excitation (SE) model and deep feed-forward (DFF) model. The metrics used for comparison were the mean absolute error, root mean squared error, mean absolute percentage error, and R-squared, wherein the best, average, and worst performances were evaluated for both models. The proposed model exhibited a superior predictive performance for the ship’s power load compared to that of existing models, as evidenced by the performance metrics: mean absolute error (MAE) of 55.52, root mean squared error of (RMSE) 125.62, mean absolute percentage error (MAPE) of 3.56, and R-squared ($R^2$) of 0.86. Therefore, the proposed model is expected to be used for power load prediction during electric-powered ship operations.

Keywords: ship power load; multiple feature extraction; LSTM prediction

1. Introduction

Many diesel-powered propulsion vessels currently used for maritime transportation have the disadvantage that they emit high levels of sulfur oxides (SOx), nitrogen oxides (NOx), and carbon monoxide [1,2]. To protect the marine environment, the Marine Environment Protection Committee of the International Maritime Organization regulates air pollution through the International Convention for the Prevention of Pollution from Ships [3]. In eco-friendly electric-powered vessels, the entire power load of the vessel is supplied by fuel cells or batteries, and the vessel operations aim to meet zero emission goals [4,5]. Furthermore, research on green ship technologies has been conducted to reduce fuel consumption and mitigate CO2 emissions. These include aspects such as hull design [6,7], engine models [8,9], propulsion propeller selection [10,11], steering gear design [12], alternative fuels [13,14], post-treatment technologies to reduce CO2 emissions [15], heat recovery systems [16], power distribution systems [17], and ship operation systems [18,19]. However, there exists a research gap regarding the forecasting of the variations in a ship’s power loads. Thus, for a comprehensive study on the control of electric-powered ships, it
is essential to study the power load prediction model of the ship, which changes according to sea conditions. Previous research on power prediction investigated various power prediction models, centering on urban power load prediction [20], power load prediction for a specific country [21], solar power generation prediction [22], and building power demand prediction [23]. However, unlike the power load data from the previous research, the power load data of current vessels are characterized by rapid variations in response to changes in the vessel’s operational status and external environmental factors. Furthermore, it is difficult to identify the trends, seasonality, and periodicity that can indicate changes in power load. Lastly, several proposed ship power load prediction models have yielded limited results in interpreting and expressing the characteristics of a ship’s power load [24–26]. Therefore, it is crucial to develop a model that can accurately interpret and predict a ship’s power load according to changes in the marine environment, weather conditions, and the ship’s situation.

Vessels use various types of equipment that utilize electrical loads to ensure the reliable transportation of cargo. The power loads of vessels are characterized by three factors. The first factor is the marine environment. Changes in the marine environment, such as wind speed, waves, and currents, have a direct impact on the hull resistance of ships sailing in the ocean [27]. The change in hull resistance affects the travel direction and speed of the ship; thus, continuous control is required to maintain the speed and direction of travel. Consequently, changes occur in the amount of power supplied to power-consuming equipment [28], such as steering equipment [29,30] and auxiliary blowers of the main engine [31] installed on the ship. The second factor is weather conditions. Wind direction and wind speed have a direct impact on the drag of a ship [32]. Moreover, the drag of the hull can increase considerably in foggy or rainy weather [33]. As the ship requires continuous control because of changes in its resistance, its power consumption also changes continuously. The third factor is the ship’s situation. The ship’s power load is characterized by considerable changes based on the ship’s situation [34]. For example, ships use a hydraulic crane installed on the ship when unloading or receiving cargo at a port. Here, the power consumption of the electric motor installed in the hydraulic crane changes according to the weight of the cargo [35]. Furthermore, when the ship enters or leaves the port, it uses a bow thruster to control its lateral movement [36]. Because the bow thruster consumes a lot of power, the power load of the ship also changes considerably. When the ship anchors, the heavy anchor is controlled by the winch motor [37], which also changes the power load of the ship. Hence, the ship’s power load is also dependent on changes in the ship’s situation.

This study used the power data measured on an actual ship to predict the power load of the ship. A multiple feature extraction (MFE)-long short-term memory (LSTM) model based on skip connections was developed. The performance of the model was compared with that of a previous convolutional neural network (CNN)-LSTM network with an SE model and DFF model. The comparison test results showed that the proposed model outperformed other models in predicting the ship’s power loads. Thus, the proposed model can be useful for power load prediction during ship operations.

Table 1 lists the nomenclature used in this study.

<table>
<thead>
<tr>
<th>Nomenclature</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>MFE</td>
<td>Multiple Feature Extraction</td>
</tr>
<tr>
<td>LSTM</td>
<td>Long Short-Term Memory</td>
</tr>
<tr>
<td>SE</td>
<td>Squeeze and Excitation</td>
</tr>
<tr>
<td>DFF</td>
<td>Deep Feed-Forward</td>
</tr>
<tr>
<td>CNN</td>
<td>Convolutional Neural Network</td>
</tr>
<tr>
<td>RELU</td>
<td>Rectified Linear Unit</td>
</tr>
<tr>
<td>TEU</td>
<td>Twenty-foot Equivalent Unit</td>
</tr>
</tbody>
</table>
2. Theoretical Background

2.1. One-Dimensional (1D) CNN

A convolutional neural network (CNN) [38] is a deep learning model that mimics the structure of the human optic nerve and generates a feature map of the data. Examples of data to which CNN models have been applied include videos in 3D format [39], images in 2D format [40], and signal data in 1D format [41]. For many years, the application of CNN models has been focused on image classification, face recognition, and object recognition, which require computer vision using 2D CNN. The 1D CNN [42], which features a 1D format, has been widely used for characterizing time series data such as particulate matter [43], individual residential loads [44], loads of commercial buildings [45], and ATM cash demand [46], yielding good performance results. As the ship’s power load data are also time series data, this study used 1D CNN to extract and analyze the features of the ship’s power load data.

The 1D CNN model consists of a convolutional layer and a pooling layer. The filters in the convolutional layer are used to extract features from the input data. The following equation describes the behavior of the convolutional layer:

$$y = \sigma(\sum_{i=1}^{n} conv1D(Wx_i) + b_i),$$

where $n$ is the number of feature maps in the layer; $conv1D$ is the behavior of a one-dimensional convolutional neural network within the same padding, $W$ is a trainable 1D convolutional kernel, $x_i$ is the $i$-th feature map, $b_i$ is the bias of the $i$-th feature map, and $\sigma()$ denotes the activation function. The 1D CNN model proposed in this study uses a rectified linear unit (ReLU) [47] as the activation function. The ReLU can be expressed by the following equation:

$$ReLU(x) = \max(0, x)$$

The ReLU activation function outputs a straight line with a slope of 1 if the value of the input $x$ is greater than 0, and it outputs 0 otherwise. The feature maps extracted from the convolution layer are input to the max pooling layer [48,49]. Max pooling has the advantage that it can suppress the overfitting problems and excessive computation that may occur during the training of deep learning models by down-sampling the input feature data.

2.2. Long Short-Term Memory

LSTM [50] adopts an extended structure of the memory cell in RNN (recurrent neural network) to store and retrieve data. Furthermore, it possesses the advantage that it can train temporal relationships based on long timescales. The LSTM models are mainly used in time series data processing and natural language processing. It overcomes the key problem of the traditional recurrent neural network, which is the failure to remember information that is far from the output. Figure 1 depicts the structure of the LSTM cell.
LSTMs utilize the concept of gating, which involves component-wise multiplication of the input. The LSTM cell state is updated based on the activation of these gates. The input provided to an LSTM is processed through various gates, such as the input gate, output gate, or forget gate, each controlling specific operations on the cell memory. Here, $C_{t-1}$ is the previous cell state, $h_{t-1}$ is the previous hidden state, $x_t$ is the data input to the LSTM, $C_t$ is a new cell state, $h_t$ is a new hidden state, $\oplus$ and $\otimes$ are vector operations, and $\sigma$ represents the sigmoid function, which can be expressed as follows:

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$$

where $x$ is the input data, and $e$ is a natural constant. Lastly, $\tanh$ stands for the hyperbolic tangent \cite{36}, which can be expressed by the following equation:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

where $x$ is the input data, and $e$ is a natural constant.

The forget gate of LSTM uses the computation of the current information and the value of the past hidden layer to determine how much past information to forget. The operation of the forget gate at time $t$ can be expressed by the following equation:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f),$$

where $x_t$ is the input vector, $h_{t-1}$ is the vector of a past hidden layer, $W_f$ is the weight, $b_f$ is the bias value, $f_t$ is the output of the forget gate, and $\sigma$ is the sigmoid operation. The output from the forget gate is input to the input gate, which decides the importance of the data at hand and writes them to a cell. The layers of the input gate can be represented by the following equation:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i).$$

In addition, the operations of the input gate layer and the tanh layer can be expressed as follows:
\[ \hat{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c). \] (7)

The output gate obtains a new cell state \( C_t \) using the following expression of the input gate:
\[ C_t = f_t \times C_{t-1} + i_t \times \hat{C}_t. \] (8)

Lastly, the output gate uses the following equation to determine the new hidden state \( h_t \):
\[ o_t = \sigma_t(W_o[h_{t-1}, x_t] + b_o), \] (9)
\[ h_t = \sigma_t \times \tanh(C_t). \] (10)

2.3. Skip Connection

Deep learning models with deep architectures generally have better learning outcomes. However, the deeper the architecture of a deep learning model, the worse the performance. Skip connections, which emerged to solve this problem, proved effective in preventing the performance degradation of deep learning models with deep layers.

The skip connection method skips input data in a deep learning model and connects them directly to the output. Examples include addition skip connection [51] and concatenation [52] skip connection; their structures are depicted in Figure 2.

![Figure 2. Structures of addition and concatenation skip connections.](image)

The addition skip connection method skips the convolutional layer and adds the input data directly to the output, which allows information from the input data to flow to the output even in deep learning models, preventing the performance of the model from declining. In Figure 2a, \( x \) is the input data, \( F(x) \) is the result output from Layer 2, and \( F(x) + x \) is the result of the addition skip connection method, which adds the input data to the result output from Layer 2. Unlike the addition skip connection in Figure 2b, the concatenation skip connection performs the concatenation of the vector of the input data with the vector value of the output from Layer 1. Thus, the maximum amount of information is stored in each layer of the deep learning model, thereby improving the model accuracy. This study adopted the concatenation skip connection method.

2.4. Dataset

The data used in this study were gathered from a 6800 twenty-foot equivalent unit (TEU) container ship named Hyundai Bangkok that was in actual operation between 15 November 2014 and 9 April 2015. The vessel consisted of one MAN B&W diesel engine to propel the vessel and four 3800 kW generators to feed the ship’s power load. The detailed specifications of the vessel are presented in Table 2.
Table 2. Detailed specifications of the vessel.

<table>
<thead>
<tr>
<th>Vessel Type</th>
<th>Container Ship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>303 m</td>
</tr>
<tr>
<td>Breadth</td>
<td>40 m</td>
</tr>
<tr>
<td>Draught</td>
<td>12.1 m</td>
</tr>
<tr>
<td>Engine power</td>
<td>91,886 BHP</td>
</tr>
<tr>
<td>Generator power</td>
<td>3800 kW × 4</td>
</tr>
<tr>
<td>TEU</td>
<td>6800 TEU</td>
</tr>
</tbody>
</table>

Vessels use rudders for direction control. The power consumption of the rudder fluctuates according to the hull resistance, where rudder angle, water speed, wind speed, and wind angle can be classified as hull resistance variables. The ship utilizes an electric hydraulic system for rudder angle control. It is composed of a hydraulic pump with an electric pump, which exploits the rudder angle by controlling the hydraulic cylinder. The multivariate auto regressive exogenous model is implemented for rudder control, which can be expressed as follows.

\[
X(n) = \sum_{m=1}^{M} A(m)X(n-m) + \sum_{m=1}^{M} B(m)Y(n-m) + U(n) \tag{11}
\]

where \(X(n)\) is a two-dimensional vector which has yaw and roll. \(Y(n)\) is a one-dimensional rudder’s vector that can be controlled. \(U(n)\) refers to Gaussian white noise; \(X(n-m)\) and \(Y(n-m)\) represent the difference between the instrumented input and output vector values. \(M\) is the value of the \(M\)th command. The established heading set by the multivariate auto regressive exogenous model remains constant. However, the ship is influenced by consistent waves and wind conditions, and under severe wave and wind conditions, the heading angle can change. Four external environmental data were selected in this study, and all the data were measured and acquired every 10 min, with a total acquisition of 20,935 data. The types of data used in this study are listed in Table 3.

Table 3. Data used in experiments.

<table>
<thead>
<tr>
<th>Type of Data</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electric Load</td>
<td>kW</td>
</tr>
<tr>
<td>No. 1 Diesel Generator Power</td>
<td>kW</td>
</tr>
<tr>
<td>No. 2 Diesel Generator Power</td>
<td>kW</td>
</tr>
<tr>
<td>No. 3 Diesel Generator Power</td>
<td>kW</td>
</tr>
<tr>
<td>No. 4 Diesel Generator Power</td>
<td>kW</td>
</tr>
<tr>
<td>No. 1 Reefer Container Group Load</td>
<td>kW</td>
</tr>
<tr>
<td>No. 2 Reefer Container Group Load</td>
<td>kW</td>
</tr>
<tr>
<td>No. 3 Reefer Container Group Load</td>
<td>kW</td>
</tr>
<tr>
<td>No. 4 Reefer Container Group Load</td>
<td>kW</td>
</tr>
<tr>
<td>No. 5 Reefer Container Group Load</td>
<td>kW</td>
</tr>
<tr>
<td>No. 6 Reefer Container Group Load</td>
<td>kW</td>
</tr>
<tr>
<td>Rudder Angle</td>
<td>Degree</td>
</tr>
<tr>
<td>Water Speed</td>
<td>(\text{m.s}^{-1})</td>
</tr>
<tr>
<td>Wind Speed</td>
<td>(\text{m.s}^{-1})</td>
</tr>
<tr>
<td>Wind Angle</td>
<td>Degree</td>
</tr>
</tbody>
</table>

The operational states of a ship can be classified into underway, standby, and inport. First, the underway state indicates that the ship is in motion, utilizing all equipment for navigation. It is characterized by minimal fluctuations in power load. Second, the standby state represents that the ship is in the process of entering or leaving the port. It is marked
by significant variations in the total power consumption of the auxiliary devices, depending on the ship’s speed changes. Third, the inport state denotes that the ship is either loading or unloading cargo. During this time, both the power consumption and load variations are minimal due to the stationary situation of the ship. Figure 3 depicts the total power load of the ship measured in 10 min increments. Rapid changes in load were observed as the ship operated. Particularly during the ship’s entry and departure, significant changes were observed in the power consumption of the electrically operated steering gear and bow thrusters, leading to substantial fluctuations in the overall power load of the vessel.

Figure 3. Total power load of vessel measured in 10 min intervals.

3. Proposed Model

The MFE-LSTM model based on the skip connection model was developed to improve the performance of the existing ship power consumption prediction model. The proposed model is largely composed of a data input layer, multiple feature extraction, concatenation layer, LSTM layer, skip connection layer, dense layer, and forecasts of the ship’s power. Figure 4 illustrates the structure of the proposed MFE-LSTM model based on the skip connection model.
Figure 4. Structure of the MFE-LSTM model based on skip connections.
The model process is described as follows.

Step 1. Input layer

The collected ship data are subjected to data scaling and are input to the input layer to train the model. The MinMaxScaler [53] was used, which can be expressed as follows:

$$\text{MinMaxScaler}: x' = \frac{x - \text{min}(x)}{\text{max}(x) - \text{min}(x)}$$  \quad (12)

where $x'$ is the new value obtained by the MinMaxScaler; $x$ is the original input data; $\text{min}(x)$ is the minimum value of the original data column; and $\text{max}(x)$ is the maximum value of the original data column.

Step 2. MFE

MFE plays a key role in the proposed model. The ship’s power load data are characterized by large variations. Hence, extracting various features of the data improves the performance of the model, and three CNNs with different kernel sizes and filters are connected in parallel to extract various features of the ship’s power load data. Each CNN has the following structure: 1D convolution layer—1D convolution layer—1D convolution layer—1D convolution layer—max pooling layer—batch normalization. Each layer uses the ReLU as an activation function. The structure of each 1D convolutional layer of the CNN is presented in Table 4.

Table 4. Structures of 1D convolution layers of CNN models.

<table>
<thead>
<tr>
<th>Number of CNN</th>
<th>Structure of Each 1D Convolution Layer (Filters, Kernel Size)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN1</td>
<td>(256, 128)—(128, 64)—(64, 32)</td>
</tr>
<tr>
<td>CNN2</td>
<td>(512, 256)—(256, 128)—(128, 64)</td>
</tr>
<tr>
<td>CNN3</td>
<td>(1024, 512)—(512, 256)—(256, 128)</td>
</tr>
</tbody>
</table>

Next, the feature values output from the 1D convolution layers are down-sampled by max pooling, and batch normalization [54] is performed to prevent internal covariate shifts.

Step 3. Concatenation layer

For the concatenation layer, the concatenation skip connection method is adopted. The advantage of this approach is that the feature values output from the MFE can all be utilized in the subsequent step.

Step 4. LSTM layer

The LSTM layer predicts the power load of the ship according to the input values from the concatenation layer and is composed of two LSTMs with 1024 neurons. The activation function is ReLU. The values output from the LSTM layer are input to the skip connection layer in the next step.

Step 5. Skip connection layer

The skip connection layer combines the ship’s power load feature values collected from the concatenation layer and the ship’s power load predicted using the LSTM layer into one vector using the concatenation skip connection method. The combined vector value is input to the dense layer. This method enables the next layer to utilize both the analyzed feature values of the ship’s power load and the power load values predicted by the LSTM layer.

Step 6. Dense layer

The dense layer analyzes the values input through the skip connection layer and predicts the power load value of the vessel. The layer consists of six perceptrons in total, each with 1024, 512, 256, 128, 32, and 1 neurons. Here, the dense layer can be expressed as follows:

$$\theta_1 = \text{Relu}(W_1\text{SkipConnectionLayer} + b_1),$$

$$\theta_n = \text{Relu}(W_n\theta_{n-1} + b_n),$$  \quad (13)
where $\theta_1$ is the first perceptron; $ReLU$ is the activation function; $SkipConnectionLayer$ is the vector input through the skip connection layer; and $W$ and $b$ are weight and bias, respectively. Lastly, the predicted value of the ship’s power load is input with inverse scaling.

Step 7. Ship power forecasts and inverse scaling

Because the values input through the dense layer are output according to the data-scaled values, their predicted values are small. Consequently, inverse scaling must be performed to obtain the ship’s power value predicted by the model. Here, inverse scaling can be expressed as follows:

$InverseScale: x = y'(\text{max}(x) - \text{min}(x)) + \text{min}(x)$, (14)

where $y'$ is the ship’s power forecast value with MinMaxScaler applied, and $x$ is the actual value obtained after inverse scaling.

4. Methods

4.1. Model Training Process

The proposed MFE-LSTM model based on skip connections was trained using the data collected from the vessel. The data in the dataset were measured every 10 min, and the model utilized the previous 50 min of data to predict the next 10 min of the ship’s power load. Hence, the time step was 5. To prevent the model from overfitting, the dataset was classified into training, validation, and test sets for training and evaluation. The training and evaluation sets accounted for 80% and 20% of the total dataset, respectively. Furthermore, the training dataset was divided into 70% and 30% for the testing and validation sets, respectively.

4.2. Evaluation Metrics

Four evaluation metrics were selected to evaluate the performance of the model. The selected metrics have characteristics that are commonly used in regression and forecasting models. They consist of mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE), and R-squared ($R^2$). The evaluation metrics can be expressed by the following equations:

$MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \bar{y}_i|$, (15)

$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{y}_i)^2}$, (16)

$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{y_i - \bar{y}_i}{\bar{y}_i} \right| \times 100$, (17)

$R^2 = 1 - \frac{\sum(y_i - \bar{y}_i)^2}{\sum(y_i - \bar{y})^2}$ (18)

where $N$ is the number of data points used to evaluate the model performance, $y_i$ is the $i$-th correct answer, $\bar{y}_i$ is the $i$-th predicted value, and $\bar{y}$ is the average value of the correct answers. MAE is the average of the differences between the predicted and actual values. RMSE is the square root of the averages of the squared differences between the predicted and actual values. MAPE indicates the relative difference between the predicted and actual values by dividing the difference between each predicted value and actual value by the actual value, then taking the absolute values and averaging them. $R^2$ indicates how well the model describes the data and can be used to determine the correlation between the predicted and actual values.

4.3. Model Performance Evaluation Process
The objective performance of the model was evaluated using the MAE, RMSE, MAPE, and $R^2$ metrics, under the following conditions:

- The CNN-LSTM neural network with the SE model and DFF model, which are reputed to be the best predictors of the ship’s power load among the existing models, was selected for the performance evaluation comparison of the proposed MFE-LSTM model based on skip connections.
- The same data were used for training and evaluation of the model, and the training was repeated five times.
- The experimental results of the proposed model and the comparison model were compared in terms of the selected evaluation metrics, and a detailed comparative evaluation was performed using the maximum, average, and minimum performance results.

5. Results and Discussion

In this chapter, the performance of the proposed model is assessed by comparing it with the conventional models. The comparison experiments were conducted using Python and TensorFlow libraries. First, the experimental results of each model were examined. Table 5 presents the results of five experiments conducted with the proposed MFE-LSTM model based on skip connections. The highest performance was an MAE of 47.19, an RMSE of 122.26, a MAPE of 3.01, and an $R^2$ of 0.87. The lowest performance was an MAE of 69.40, an RMSE of 130.98, a MAPE of 4.41, and an $R^2$ of 0.85. Lastly, the average MAE was 55.52, the average RMSE was 125.62, the average MAPE was 3.56, and the average $R^2$ was 0.86. These results indicated that the overall performance of the ship’s power load prediction was excellent.

Table 5. Experimental results of MFE-LSTM model based on skip connections.

<table>
<thead>
<tr>
<th>Evaluation Metric</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>47.19</td>
<td>51.17</td>
<td>69.40</td>
<td>55.92</td>
<td>53.93</td>
</tr>
<tr>
<td>RMSE</td>
<td>122.26</td>
<td>125.07</td>
<td>130.98</td>
<td>123.64</td>
<td>126.19</td>
</tr>
<tr>
<td>MAPE</td>
<td>3.01</td>
<td>3.48</td>
<td>4.41</td>
<td>3.49</td>
<td>3.43</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.87</td>
<td>0.87</td>
<td>0.85</td>
<td>0.87</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Table 6 summarizes the results of five experiments for predicting the ship’s power load with the CNN-LSTM neural network with the SE model selected for performance evaluation comparison. The highest performance was an MAE of 78.28, an RMSE of 146.64, a MAPE of 5.18, and an $R^2$ of 0.82. The lowest performance was an MAE of 90.6, an RMSE of 148.96, a MAPE of 6.21, and an $R^2$ of 0.81. The average MAE was 85.64, the average RMSE was 149.02, the average MAPE was 5.75, and the average $R^2$ was 0.81.

Table 6. Experimental results of CNN-LSTM neural network with SE model.

<table>
<thead>
<tr>
<th>Evaluation Metric</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>86.64</td>
<td>90.60</td>
<td>90.60</td>
<td>78.28</td>
<td>81.16</td>
</tr>
<tr>
<td>RMSE</td>
<td>151.53</td>
<td>148.96</td>
<td>148.96</td>
<td>146.64</td>
<td>149.01</td>
</tr>
<tr>
<td>MAPE</td>
<td>5.68</td>
<td>6.21</td>
<td>6.21</td>
<td>5.18</td>
<td>5.47</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.81</td>
<td>0.81</td>
<td>0.81</td>
<td>0.82</td>
<td>0.81</td>
</tr>
</tbody>
</table>

Table 7 shows the results of five experiments for forecasting the ship’s power load with the DFF model for a comparison test. The highest performance was an MAE of 80.71, an RMSE of 152.65, a MAPE 5.63 and an $R^2$ of 0.81. The lowest performance was an MAE of 91.75, an RMSE of 160.81, a MAPE 6.47 and an $R^2$ of 0.78. The average MAE was 85.97, the average RMSE was 156, the average MAPE was 6.04, and the average $R^2$ was 0.79.
Table 7. Experimental results of DFF model.

<table>
<thead>
<tr>
<th>Evaluation Metric</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>86.13</td>
<td>86.56</td>
<td>80.71</td>
<td>84.72</td>
<td>91.75</td>
</tr>
<tr>
<td>RMSE</td>
<td>155.33</td>
<td>156.30</td>
<td>152.65</td>
<td>154.92</td>
<td>160.81</td>
</tr>
<tr>
<td>MAPE</td>
<td>6.12</td>
<td>6.14</td>
<td>5.63</td>
<td>5.83</td>
<td>6.47</td>
</tr>
<tr>
<td>$\textbf{R}^2$</td>
<td>0.80</td>
<td>0.79</td>
<td>0.81</td>
<td>0.80</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Figures 5–7 illustrate the results of the ship’s power consumption prediction experiments using the MFE-LSTM model based on skip connections, the CNN-LSTM neural network with the SE model, and the DFF model. Overall, the latter was a better predictor of the ship’s power load; however, it achieved an inferior performance in predicting the ship’s power load.
Figure 5. MFE-LSTM using skip connections.
Figure 6. CNN-LSTM neural network with SE.
Table 8 summarizes a comparison of the best performances of the models. When the MFE-LSTM model based on skip connections was used to predict the ship’s power load, the model improved the ship’s power load prediction performance by an MAE of 31.09, an RMSE of 24.38, a MAPE of 2.14, and an $R^2$ of 0.05 compared to the CNN-LSTM neural network with SE, and the proposed model reinforced the prediction performance by an MAE of 33.52, an RMSE of 30.39, a MAPE of 2.62, and an $R^2$ of 0.06 compared to the DFF model.
Table 8. Best performance comparison of both models.

<table>
<thead>
<tr>
<th>Model</th>
<th>MAE</th>
<th>RMSE</th>
<th>MAPE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>MFE-LSTM using skip connection</td>
<td>47.19</td>
<td>122.26</td>
<td>3.01</td>
<td>0.87</td>
</tr>
<tr>
<td>CNN-LSTM neural network with SE</td>
<td>78.28</td>
<td>146.64</td>
<td>5.15</td>
<td>0.82</td>
</tr>
<tr>
<td>DFF</td>
<td>80.71</td>
<td>152.65</td>
<td>5.63</td>
<td>0.81</td>
</tr>
</tbody>
</table>

Table 9 summarizes a comparison of the lowest performances of the models. Compared to the performance of the CNN-LSTM neural network with the SE model, the performance of the MFE-LSTM model based on skip connections was improved by an MAE of 21.2, an RMSE of 17.98, a MAPE of 1.8, and an $R^2$ of 0.04 in predicting the ship’s power load, and the proposed model improved the performance by an MAE of 22.35, an RMSE of 29.83, a MAPE of 2.06, and an $R^2$ of 0.07 compared to the DFF model.

Table 9. Worst performance comparison of both models.

<table>
<thead>
<tr>
<th>Model</th>
<th>MAE</th>
<th>RMSE</th>
<th>MAPE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>MFE-LSTM using skip connection</td>
<td>69.40</td>
<td>130.98</td>
<td>4.41</td>
<td>0.85</td>
</tr>
<tr>
<td>CNN-LSTM neural network with SE</td>
<td>90.60</td>
<td>148.96</td>
<td>6.21</td>
<td>0.81</td>
</tr>
<tr>
<td>DFF</td>
<td>91.75</td>
<td>160.81</td>
<td>6.47</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Table 10 summarizes a comparison of the average performances of the models. The MFE-LSTM model based on skip connections improved the performance of the prediction of the ship’s power load by an MAE of 30.13, an RMSE of 23.4, a MAPE of 2.19, and an $R^2$ of 0.05 compared to the CNN-LSTM neural network with SE, and the proposed model improved the performance by an MAE of 30.45, an RMSE of 30.38, a MAPE 2.48, and an $R^2$ of 0.07 compared to the DFF model.

Table 10. Average performance comparison of both models.

<table>
<thead>
<tr>
<th>Model</th>
<th>MAE</th>
<th>RMSE</th>
<th>MAPE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>MFE-LSTM using skip connection</td>
<td>55.52</td>
<td>125.62</td>
<td>3.56</td>
<td>0.86</td>
</tr>
<tr>
<td>CNN-LSTM neural network with SE</td>
<td>85.65</td>
<td>149.02</td>
<td>5.75</td>
<td>0.81</td>
</tr>
<tr>
<td>DFF</td>
<td>85.97</td>
<td>156</td>
<td>6.04</td>
<td>0.79</td>
</tr>
</tbody>
</table>

The results of the experimental analysis reveal that the proposed model, MFE-LSTM using skip connections, obtained the lowest RMSE, MAE, MAPE, and $R^2$ values. In particular, the proposed model demonstrated its capability to extract and analyze significantly more intricate load data from ships compared to the existing models. Therefore, it can be concluded that the proposed model is the most effective in forecasting the complex variations in the power loads of ships.

6. Conclusions

This paper describes an MFE-LSTM model based on skip connections that is capable of comprehensively extracting various features from a ship’s power load. The proposed model leverages the advantages of LSTM and CNN structures to address the large variations in the power loads of ships. LSTM excels at managing time series data, and CNN
models are better suited for extracting features from data. Furthermore, the skip connection layer preserves the information in the input data to prevent the performance degradation of the model. The performance of the model was compared with previous models, i.e., the CNN-LSTM neural network with the SE and DFF models. The results of the comparative test indicate that the proposed model outperformed the others, with the best values of MAE = 55.52, RMSE = 125.62, MAPE = 3.56, and $R^2 = 0.86$. The main conclusions drawn from this study are summarized as follows:

1. The ship’s power load prediction performance was improved by extracting various features of the ship’s power load using MFE.
2. The skip connection layer combines the results of the MFE from the concatenate layer into a singular vector representing the ship’s power load predicted by the LSTM. Consequently, the MFE-LSTM model based on skip connections excels in predicting the intricate dynamics of the ship’s power load compared to the conventional models.
3. A dedicated feature extraction model for predicting heavy loads intermittently is required to improve the performance of intermittent heavy load prediction.

This study demonstrated that multiple feature extraction models can extract and analyze various features from the data to provide an improved forecasting of the power load performance of vessels. However, the power load prediction performance of the heavy load generated by ships was insufficient. Therefore, in forthcoming research, alternative deep learning models will be explored to improve the power load prediction performance for heavy loads in ships.

Author Contributions: Conceptualization, methodology, and software, J.-Y.K.; project administration, funding acquisition, J.-S.O. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Technology Innovation Program (RS-2023-00252794, Development of Efficient Power Energy Management System and Verification for Hybrid Propulsion Ship) funded by the Ministry of Trade, Industry, and Energy of Korea.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References


**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.