Renewable Energy 231 (2024) 121022

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Fault diagnosis of a wave energy converter gearbox based on an Adam optimized CNN-LSTM algorithm

Jichuan Kang [a], Xu Zhu [a], Li Shen [a], Mingxin Li [b,*]

[a] College of Shipbuilding Engineering, Harbin Engineering University, Harbin, China
[b] Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, United Kingdom

ARTICLE INFO

Keywords:
Wave energy converter
Gearbox
Fault diagnosis
Deep learning
CNN-LSTM
Renewable energy

ABSTRACT

The complex structure and harsh operating environment of wave energy converters can result in various faults in transmission components. Environmental noise in practical operating situations may obscure the effective information in collected vibration signals, significantly increasing the difficulty of fault diagnosis. This paper presents a fault diagnosis model for the gearbox of the point absorber wave energy converter. The model integrates a convolutional neural network with long short-term memory to realize efficient extraction of local features from signals and enhance the performance in time-series analysis. Moreover, the model incorporates the Adaptive Moment Estimation algorithm to address the situations where gradients within tensors exhibit unstable changes in the model. A rigid-flexible coupled dynamics simulation model is developed to simulate vibration signals used to train and verify the fault diagnosis model. Experimental tests of the proposed model on a vibration dataset acquired from real vibration experiments demonstrate its efficacy in diagnosing various types of faults under interference of operating conditions. Comparative studies with other models demonstrate the superiority of the proposed model in terms of fault feature extraction, learning convergence efficiency, and diagnostic accuracy, indicating that the proposed model can achieve faster and more accurate fault diagnosis of wave energy converter gearboxes.

1. Introduction

1.1. Background

The world is moving toward the usage of clean and renewable energy sources to fulfill energy demand and reduce carbon emissions [1,2]. Renewable energy is expected to become the main source of global energy consumption by 2050, with its share in power generation expected to experience a significant increase [3,4]. Mainstream renewable energy options include wind power, solar energy, hydropower, bioenergy and geothermal power [5,6]. Among the various renewable energy sources, wave energy is attracting increasing attention due to its vast potential [7]. The global reserves of wave energy are estimated to be as high as 10 TW [8]. The exploitable wave energy resource is estimated to be 29,500 TWh per year [9], which exceeds the global electricity consumption in 2022. The advantages of wave energy over other renewable resources, such as solar and wind, include higher power density, longer operable time, and less impact on marine environment [10].

Despite the promising potential, wave energy converter systems remain in the demonstration phase. Challenges still exist in transforming the concept into a consolidated industrial technology [11]. For example, wave-to-wire models play an important role in the development of wave energy converters, which can be classified into frequency-domain models, time-domain models and Computational Fluid Dynamics (CFD) models. The computational efficiency of these models decreases in the sequence, while the fidelity increases [12]. Balancing efficiency and accuracy in the development of models presents a significant challenge in the design of wave energy converters. For another example, the optimal design of wave energy converters is determined under a specified set of conditions, which is a challenging process due to the interaction among stochastic wave climates, nonlinear hydrodynamics, optimal control strategies, non-ideal Power Take-Off (PTO) dynamics [13].

Wave energy converters can be generally categorized into oscillating water column devices, oscillating body devices and overtopping devices [11]. The functioning of the oscillating water columns is based on the principle of wave-induced air pressurization. As the free water surface rises and falls, the air trapped within the air chamber is compressed and

* Corresponding author. University of Strathclyde - Royal College Building, 204 George Street, G11RX, Glasgow, United Kingdom.
E-mail address: mx.li@strath.ac.uk (M. Li).

https://doi.org/10.1016/j.renene.2024.121022
Received 22 November 2023; Received in revised form 9 July 2024; Accepted 16 July 2024
Available online 17 July 2024

expanded, inducing a pressure difference from that outside the chamber and driving the turbine at the airflow passage to rotate and generate electricity [14]. Overtopping devices produce electricity by utilizing the difference between the high-water level in the reservoir and the low-water level at the exit, forcing fluid through a low-head turbine coupled to a generator [15]. Oscillating body devices are floats oscillating under the excitement of incident waves and capture their energy, which is then converted into electrical energy using PTO [16]. Oscillating body devices are further categorized into point absorbers, attenuators, and terminators. Among the classifications, the point absorber wave energy converter has advantageous characteristics including small dimensions, ease of fabrication and installation, and cost-effective maintenance, which is gaining increasing attention in recent years [7,17]. For clarity purposes, subsequent references to wave energy converters in this paper specifically refer to point absorber wave energy converters.

Wave energy converters can also be categorized based on the technologies suitable for shoreline, nearshore, and offshore deployment. The water depth influences the design architecture and operating conditions of wave energy converters [18]. The potential fault types are possibly related to the different operating environment. Shoreline devices are situated in the shoreline. The gears in the PTO system primarily suffer from risk of wear due to random mechanical motion. Nearshore converters are installed a few hundred meters from the shore in moderate water depths. The structure withstands the stress caused by passing waves. In this case, gears may develop pitting faults due to immersion in seawater.

As a complex offshore structure, the wave energy converter consists of buoy systems, PTO systems, pre-tensioning systems and anchoring systems, operating in harsh and variable offshore environments [19,20]. In offshore waters, the average wave height is usually higher than in nearshore waters. The wave period is between 10 s and several tens of seconds. Wave energy converters are placed in deep waters (more than 40 m), far from the shore, where their structure endures very high loads. In this environment, the wave energy converter suffers from negative environmental influences, such as salt spray erosion, huge waves, and typhoons [21]. These environmental impact causes various types of external and internal damage to the system and seriously restrict the operation of wave energy converters. For instance, gears may experience broken tooth fault due to intense wave impacts [22]. Once a fault occurs, performing maintenance in the marine environment is costly and time-consuming [23,24,62].

1.2. Literature review

The fault diagnosis of wave energy converters is a challenging task because the operational data of faulty components collected are often accompanied by noise interference and waveform distortion, which makes it difficult to extract information about the fault features of the critical component directly from the spectrogram of the collected signals.

In recent years, there have been studies addressing the fault analysis of wave energy converters. In past research [25,26], simulation models of the test component were built based on real-world cases, and the sensitivity of fault detection on the wave energy converter was analyzed using structural analysis. Authors of [27] used the Operation and Maintenance (O&M) simulation tool to derive an estimate of the failure rate of different wave energy converter components. Authors of [17] applied an approach that combines parametric finite element analysis modeling, response surface modeling, and reliability analysis to build a reliability assessment framework for the floating system of the wave energy converter, which identifies the key components of the floating system and its fault features. In Ref. [28], the authors performed real-time and smoothing estimates for fault diagnosis by using predicted and measured wave height data for damping subsystems. The aforementioned studies focus on the fault analysis of the overall wave energy converter, but they do not consider the critical components whose failures could trigger a chain reaction throughout the entire system.

When studying the fault diagnosis in specific components of wave energy converters, past research mainly focused on control systems. In Ref. [29], a fault diagnosis approach was proposed to effectively detect sensor and actuator faults in real-time, aiming to maximize energy production. A fault diagnosis method was proposed where an unknown input observer was designed to estimate the fault in real time, which is robust against model uncertainties [30]. Although the gearbox in the PTO system is the key component with the highest failure rate and risk priority according to the experience of wind turbines [31,32], the research focusing on fault diagnosis of wave energy converter gearboxes still lacks.

There has been extensive research on fault diagnosis across various fields, such as renewable energy [64] and transportation engineering [63]. These approaches can be applied to wave energy converters as well [33]. One of the most common methods is the machine learning method [34,35]. Numerous studies have been conducted to address fault diagnosis of gearboxes in wind turbines using machine learning methods [36]. An artificial neural network (ANN) model with a back propagation network was first used for fault diagnosis in wind turbine gearboxes in Ref. [37].

With the continuous iteration and improvement of machine learning methods, the ANN algorithm is found to have a single network structure with poor expressive and fitting capabilities. Authors of [38] proposed a fault diagnosis model for gearboxes based on Convolutional Neural Network (CNN), where the CNN increases the depth of the network by adding convolutional layers to enhance the learning ability for fault features. The input vibration signals are decomposed to improve the robustness of the CNN when dealing with uncertain fault output scales [39]. In order to deal with faulty health indices with time-series relationships, a Recurrent Neural Network (RNN) was used for fault verification of a simulated drive train of a gearbox, which can prove the effectiveness of this approach [40]. In Ref. [41], in order to overcome the problems of gradient vanishing and gradient explosion of RNN, a faster converging Long Short-Term Memory (LSTM) was used to accurately categorize the types of faults in the gearbox. By adding cosine loss to the conventional structure of LSTM, elimination of the effect of signal strength to improve the accuracy of fault diagnosis can be realized [42]. A CNN-RNN fault diagnosis model was proposed by unifying the capacity of CNN in extracting high-level local features and the excellent performance of RNN in learning the contextual dependencies of vibration signals [43,44]. In Refs. [45,46], CNN and residual learning was utilized for local feature extraction and dimension reduction, and then LSTM is used to deal with the long-range temporal dependencies.

1.3. Contributions of this paper

The above research has proven that the deep learning method has a powerful ability to identify the fault features and shows great potential in terms of accuracy and efficiency of fault diagnostics. To the best of the authors' knowledge, although the deep learning method has been applied in fault diagnosis for various systems, no research before introduced it to deal with fault diagnosis problems in wave energy converter gearboxes. The utilization of deep learning in fault diagnosis is expected to enhance the performance and reliability of wave energy converters.

Compared to other rotating machinery, there are challenges existing in the current fault diagnosis of wave energy converter gearboxes: (1) The wave energy converter gearbox is subject to the corrosive effect of seawater and the abrasive effect of wave impact. The gearbox suffers from multiple interrelated fault modes, and the ambient noise complicates the extraction of fault features from vibration signals. (2) Previous research lacks vibration experiments and datasets for the wave energy converter gearbox. The practical effects of this deficiency cannot be adequately addressed by relying solely on simulation data and publicly

available datasets. This gap hinders testing the performance of fault diagnosis models in real-world scenarios.

Considering the above challenges, in this paper, a fault diagnosis model based on the CNN-LSTM method is proposed. In addressing the issue where the frequency component of environmental noise overlaps with the frequency range of faulty gear vibration signal, the proposed model utilizes the convolutional and pooling layers to encode the spatial information of the data, so as to realize the efficient extraction of local features from vibration signals. Considering the deficiency of CNN method in learning long-term dependence in vibration signals, the forget gate structure of LSTM is introduced to enhance the performance of the model in time-series analysis. Moreover, the model based on end-to-end automatic learning feature extraction is improved by adding the Adaptive Moment Estimation (Adam) optimization algorithm, which calculates an adaptive learning rate for each parameter. The issue of unstable gradient changes in tensors can be effectively addressed, enhancing the stability and efficiency of the diagnostic model. The fault diagnosis model is trained using the vibration signal data derived from the simulation of gearbox multibody dynamics modeling. The performance of the model is subsequently verified by the dataset generated by the numerical simulation and validated by the dataset collected from the vibration experiment on the gearbox of a point absorber wave energy converter. Additionally, the performance of the trained model is compared and analyzed with the other neural networks.

In summary, the contributions of the paper are.

(1) A fault diagnosis model based on CNN-LSTM is presented for the point absorber wave energy converter gearbox, addressing the challenge of distinguishing vibration signals under various fault modes. To the best of the author's knowledge, this is the first paper using the deep learning method for the fault diagnosis of the wave energy converter gearbox.
(2) The network structure is improved by using Adam algorithm to calculate the adaptive learning rate for each parameter, elevating the performance of the fault diagnosis model.
(3) A rigid-flexible coupling dynamics simulation model is established considering four gears in different conditions (i.e., normal, broken tooth, pitting, cracked). The simulation model generates simulated vibration signals corresponding to different fault modes, effectively serving as a dataset for training and verifying the fault diagnosis model.
(4) A vibration experiment is conducted to acquire vibration signals of gear under different fault modes. This real dataset is used to validate the effectiveness of the fault diagnosis model in handling environmental noise interference. To the best of the author's knowledge, this is the first paper where a real experiment is performed to collect vibration signals from the wave energy converter gearbox and validate the fault diagnosis model.

1.4. Outline

The remainder of this paper is organized as follows. Section 2 provides the relevant preliminary theoretical knowledge used in the model. Section 3 shows the network architecture of the model and the process of fault diagnosis. In Section 4, the simulation model of the gearbox is established, and the CNN-LSTM fault diagnosis model is trained based on the simulation dataset. Section 5 shows the experimental validation of the model, and comparative studies with other neural network models are also presented. Finally, the conclusions and further work are shown in Section 6.

2. Theoretical background

In this section, the theoretical background of CNN and LSTM is elaborated respectively. This study combines the advantages of CNN in data feature extraction and the excellent performance of LSTM in processing the vibration signal of the wave energy converter gearboxes with temporal features [47]. The original information and temporal features in the vibration signal are retained to achieve excellent fault diagnosis effect and generalization ability to the maximum extent.

2.1. Convolutional neural network

A one-dimensional convolutional layer is used for feature extraction of the time-series vibration signals in this model. The original vibration signal is pre-processed by normalization and sample segmentation and then used as input to the fault diagnosis model. In other words, the internal features of the original vibration signals are extracted directly by the neural network.

(1) Convolutional layer

The convolutional layer implements the function of information mining and feature extraction from the input data, which is essentially a mapping of local features from the upper layer to multiple filters in the lower layer, also known as convolutional kernels [48]. The process of feature extraction operation inside the convolution kernel can be represented as [49]:

$$C_j^n = \sigma\left(w_j^n * C_i^{n-1} + b_j^n\right) \tag{1}$$

where C_j^n is the jth feature mapping of the nth convolutional layer; $\sigma(\cdot)$ is the non-linear activation function; $*$ denotes the convolution operation; w_j^n is the jth weight matrix of the nth convolutional layer; C_i^{n-1} is the ith feature output of the (n-1)th convolutional layer; b_j^n is the jth bias vector of the nth convolutional layer.

The architecture of a CNN is a stack of one-dimensional convolutional layers and max pooling layers, which extracts sequence segments from the input sequence to be dotted with a convolutional kernel and output [50], as shown in Fig. 1. When combined with the LSTM, it is placed at the front end of the LSTM layer as the input.

(2) Pooling layer

The concept of pooling layer is similar to the convolutional layer in that its purpose is to reduce the number of elements to be processed in the feature map. Additionally, it expand the observation window of successive convolutional layers, which refers to the proportion of the original layer covered by the window. This expansion allows downsampling the feature map, effectively decreasing redundant information while retaining distinct features. The max pooling process is [49]:

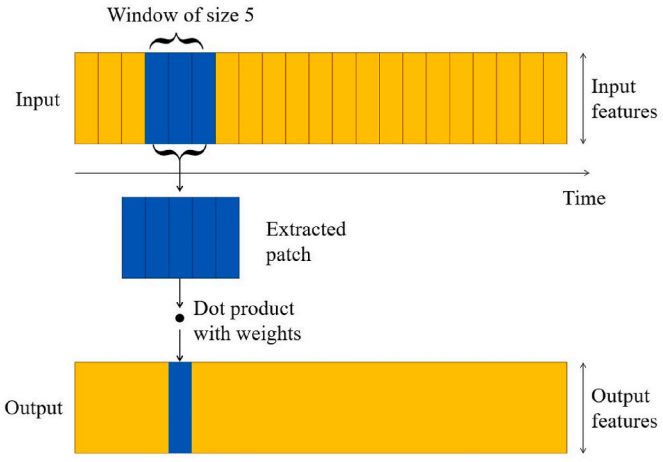


Fig. 1. Schematic of convolution.

$$P_j^n = \max\left(C_J^{n-1}, s\right) \tag{2}$$

where P_j^n is the feature output of the pooling layer; $\max(\cdot)$ denotes the downsampling function of the maximum value; C_J^{n-1} is the feature vector of the convolutional layer output; s is the pooling size.

2.2. Long short-term memory

LSTM is an evolution of the traditional RNN. It is designed to capture and retain a richer and more extended history of information from the input data, which makes it highly effective for tasks involving sequential data [51]. It addresses the drawbacks of gradient disappearance and gradient explosion during backpropagation in traditional RNNs by adding input, output and forget gate [52].

As shown in Fig. 2, the LSTM adds long-term memory storage units C_t to the short-term memory units h_t and updates both the hidden units h_t and the memory units C_t during the iterative update process. The iterative update process is implemented through forget gate, input gate and output gate.

(1) Forget gate

The mathematical procedure of the forget gate, is taking the input x_t at the current moment and the output h_{t-1} at the previous moment as inputs and using the Sigmoid activation function to select whether to forget or not. Finally, the output f_t at the moment t, is obtained as follows [53]:

$$f_t = \sigma\left(w_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{3}$$

where w_f is the weight matrix of the forgotten gate; b_f is the bias vector of the input gate; $\sigma(\cdot)$ is the Sigmoid activation function.

(2) Input gate

The process of input gate typically consists of several stages. Firstly, h_{t-1} and x_t are used as inputs to determine which information to update by outputting i_t through the Sigmoid activation function. Secondly, h_{t-1} and x_t are used as inputs to create a new vector of marquee values $\widehat{C}$ through the Tanh activation function. Finally, the output of the Sigmoid activation function is multiplied by the vector of marquee values, and the output of the Sigmoid activation function determines which information in the marquee vector is important and needs to be retained.

The mathematical procedure for the output i_t at time t in the input gate is as follows [53]:

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \tag{4}$$

where w_i is the weight matrix of the input gate; b_i is the bias vector of the input gate.

The mathematical procedure for a vector $\widehat{C}$ of marquee values is as follows [53]:

$$\widehat{C}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \tag{5}$$

where w_c is the weight matrix of the candidate values; b_c represents the bias vector in this operation; $\tanh(\cdot)$ represents the Tanh activation function.

(3) Memory cell update

After completing the operations of the forget gate and the input gate respectively, the updated value C_t of the memory cell at time t is determined by multiplying the previous layer memory cell C_{t-1} by the output f_t of the forget gate and then adding the output value of the input gate. The output value of the input gate is the result of multiplying i_t and the candidate value $\widehat{C}$. The process of updating a memory cell is as follows [45]:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widehat{C}_t \tag{6}$$

(4) Output gate

The output o_t is obtained by taking x_t and h_{t-1} as inputs using the Sigmoid activation function and then multiplying it with the result of the memory cell C_t at time t after activation by the Tanh activation function to obtain the output h_t at time t of this layer.

At moment t, the mathematical form of the output gate o_t operation is [45]:

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \tag{7}$$

where w_o is the weight matrix of the output gate; b_o is the bias vector of the output gate.

The final output of the LSTM at time t is C_t and h_t, and the operation is as follows [45]:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widehat{C}_t \tag{8}$$

$$h_t = o_t \cdot \tanh(C_t) \tag{9}$$

In the data update process, the LSTM determines which relevant information from the previous step needs to be retained through the forget gate. The input gate determines which information in the current input is important and needs to be added. The output gate determines what the next hidden state should be.

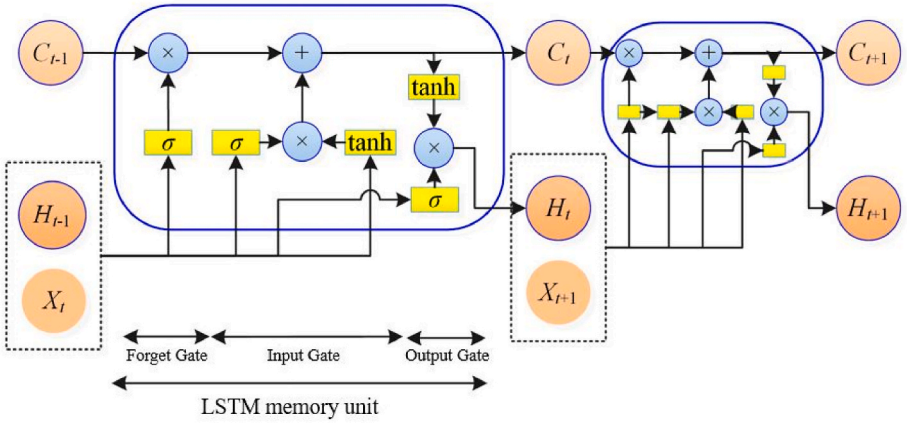


Fig. 2. Structure of the LSTM network.

2.3. Comparison indicator

Comparison indicator play a crucial role in evaluating the performance of models, and various tasks require different metrics for evaluation. In this paper, accuracy and cross-entropy loss are chosen as comparison indicators for the fault classification task.

The accuracy is an important indicator to evaluate classification models, which refers to the proportion of correctly predicted samples out of the total samples. Accuracy provides a direct reflection of the prediction performance of the model and is used to assess the fitting effect and the generalization ability. The accuracy is calculated as:

$$\text{Accuracy} = \frac{N_P}{N_t} \tag{10}$$

where N_P means the quantity of correctly classified samples; N_t signifies the total quantity of samples.

The loss function is an evaluation indicator used to measure the distance between the predicted probability distribution and the true distribution. It provides guidance on the direction for model optimization, indicating how the model should adjust its parameters to improve its predictions. For the classification problem where the true labels are one-hot encoding, the most commonly used loss function is the cross-entropy loss. The cross-entropy loss is differentiable, which facilities training using gradient descent algorithms, meanwhile, it has better numerical stability and is less likely to cause outliers.

The cross-entropy loss is calculated as:

$$\text{Cross} - \text{Entropy Loss} = -\frac{1}{m}\sum_{i=1}^{m}(y_i \log(f(x_i; w)) + (1 - y_i)\log(1 - f(x_i; w))) \tag{11}$$

where m is the number of samples; $f(x_i; w)$ is the actual output of the ith sample; y_i is the expected output of the ith sample; x_i is the input of the ith sample; w is the input weight.

3. Proposed methodology

3.1. Model structure

The complexity and diversity of vibration signals of wave energy converter gearboxes can lead to long training times for neural networks and influence the final classification results. In this context, CNN is well-suited for processing large amounts of high-dimensional and nonlinear data. By employing a sequence of dimensionality reduction and feature

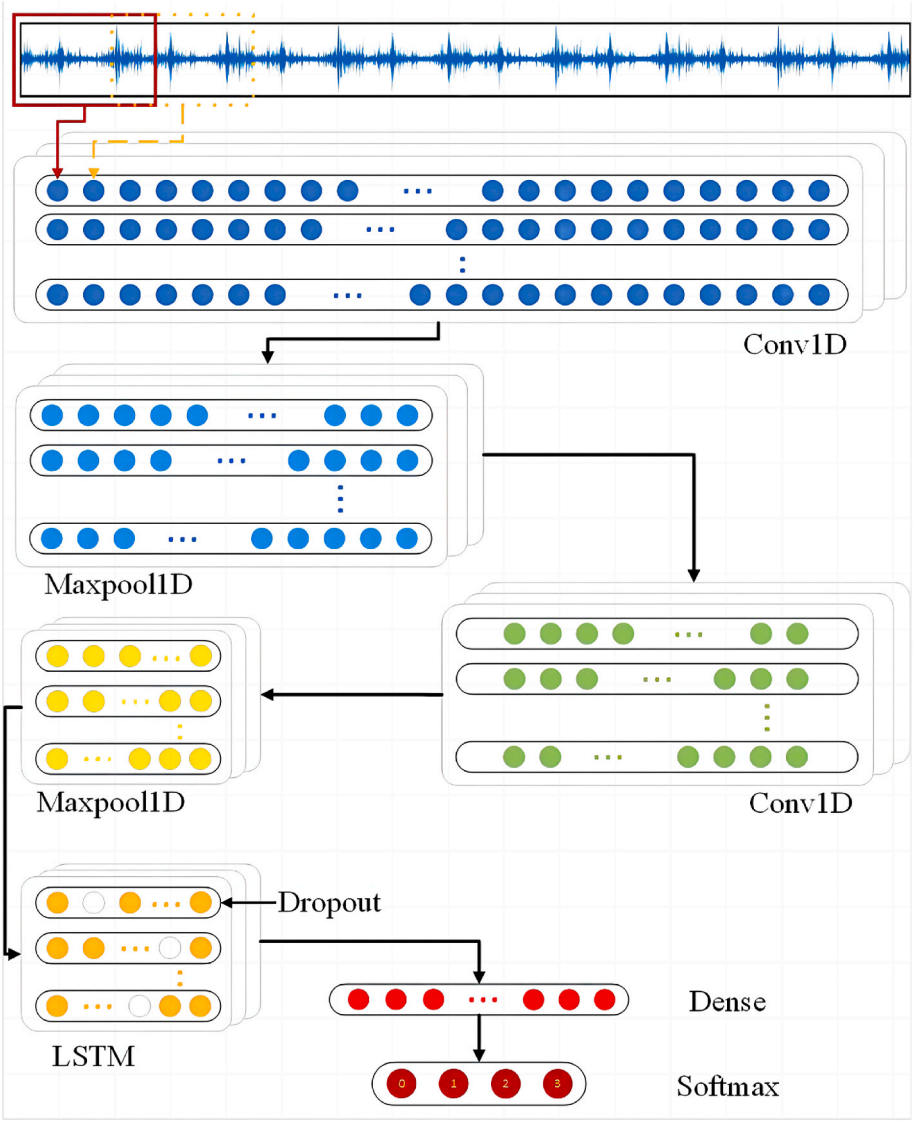


Fig. 3. Brief schematic of the overall structure of the proposed CNN-LSTM.

extraction stages, including CNN layers, batch normalization layers, activation function layers, and pooling layers, the vibration signals undergo significant parameter reduction while preserving essential features from the original data. Furthermore, LSTM is able to addresses issues such as gradient explosion and vanishing gradients, making LSTM ideal for processing time series data and performing classification tasks. Utilizing LSTM enhances the learning performance when dealing with time series data from gearbox vibration signals. Consequently, this paper proposes the integration of a CNN before the LSTM to reduce data dimensionality, forming a CNN-LSTM model. The data undergoes a sequential process involving convolutional layers, batch normalization, activation functions, pooling, LSTM layers, fully connected layers, and a softmax layer, ultimately producing the desired results.

Fig. 3 briefly reveals the structure of the proposed CNN-LSTM model, which consists of three parts: feature extraction, weight optimization and fault classification.

The proposed method consists of three parts, i.e., data preprocessing, model training, fault diagnosis. The overall process is shown in Fig. 4.

In the first part, the simulated signals are imported and the signals with large fluctuation periods are eliminated, followed by normalization of the data and one-hot encoding of the labels. The normalized data is sliced and sampled using overlapping techniques. The training set, validation set and test set are divided proportionally. Then, CNN-LSTM model construction and compilation are carried out. Hyperparameters are set and the number of iterations is defined, followed by the training of the network to optimize weights and biases through backpropagation. In the final step, the trained network preserves weight data information and completes the diagnosis of the target data by utilizing the test set (the data to be diagnosed). Model evaluation involves assessing the fault diagnosis accuracy of the test samples and considering the loss value to judge the strength of the model. The final result of the test set is presented in the form of a confusion matrix, showcasing the effect of the classification of various types of faults.

3.2. Adaptive Moment Estimation

The traditional stochastic gradient descent method employs a single learning rate for all weight updates. While it can yield good results in handling shallow neural networks, it tends to suffer from slow convergence and a strong reliance on learning rate parameters when dealing with large and complex data network models. Moreover, it is susceptible to getting stuck in local optimal solutions.

To address these issues, the Adam algorithm is introduced to enhance model optimization and fine-tune the learning rate. The Adam algorithm combines the strengths of adaptive gradient methods and root mean square propagation. It efficiently explores the parameter space while also correcting biases, all while conserving computational resources by providing distinct adaptive learning rates for various parameters.

The first moment estimation m_t and second moment estimation n_t of the gradient is as follows [54,55]:

$$\begin{aligned} m_t &= \mu * m_{t-1} + (1-\mu) * g_t \\ n_t &= \nu * n_{t-1} + (1-\mu) * g_t^2 \end{aligned} \tag{12}$$

where μ and ν are the decay rates of the moment estimates; g_t represents the gradient information.

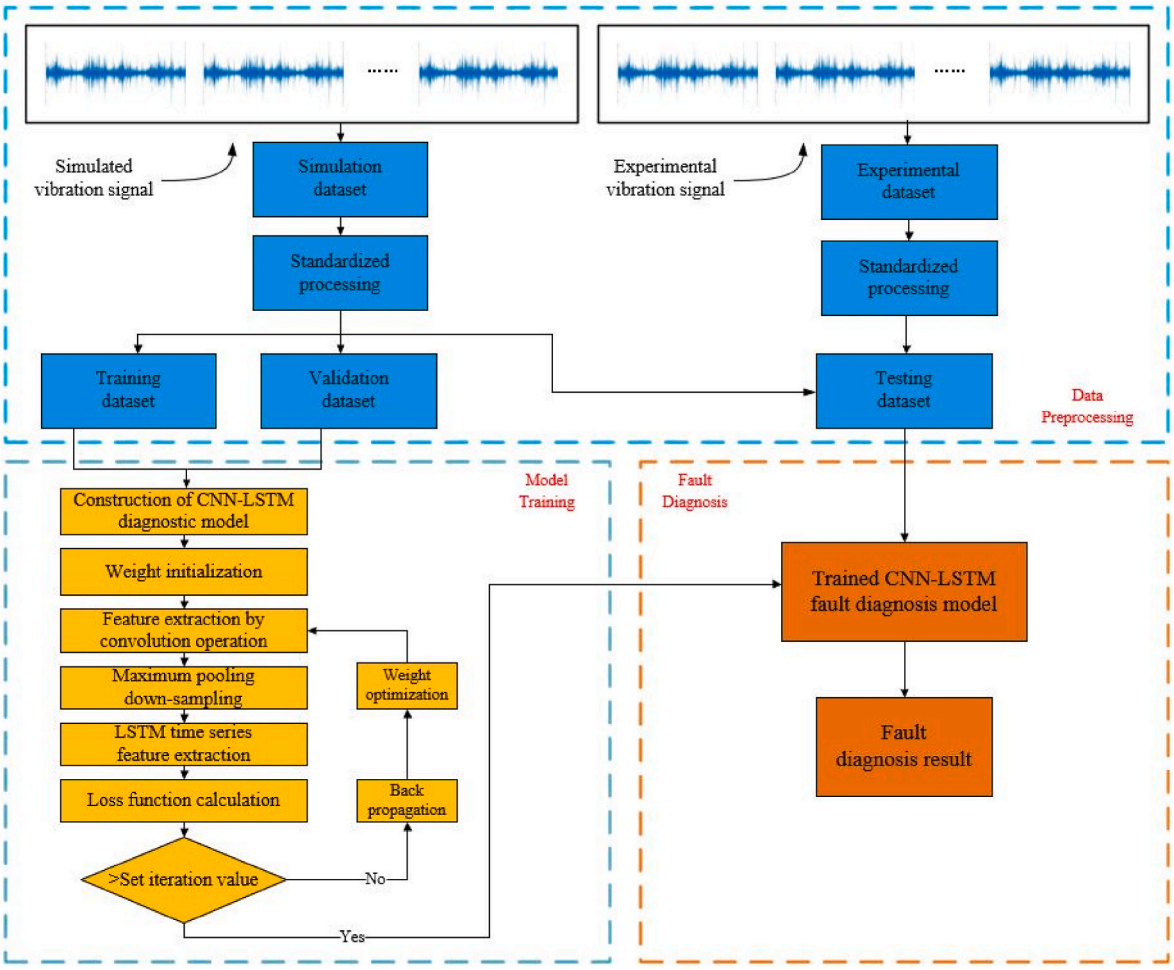


Fig. 4. Flowchart of fault diagnosis process based on CNN-LSTM model.

The correction process of moment estimation is as follows:

$$m_t^{'} = \frac{m_t}{1-\mu^t} \tag{13}$$

$$n_t^{'} = \frac{n_t}{1-\nu^t} \tag{14}$$

The parameter update of learning rate is as follows:

$$\theta_t = \theta_{t-1} - \eta \frac{m_t^{'}}{\sqrt{n_t^{'}} + \varepsilon} \tag{15}$$

where θ is the learning rate; ε represents the numerical stability constant; η represents the initial learning rate.

The integration of the Adam optimization algorithm aims to yield substantial improvements in the diagnosis outcomes of the gearbox fault diagnosis model established on the CNN-LSTM neural network. Specifically, Adam optimization demonstrates a pronounced ability to dynamically fine-tune the learning rate and efficiently promote the model convergence speed while strengthening its generalization capacity. Moreover, this algorithm manifests resilience when dealing with noise-corrupted or data-sparse datasets. These advantages of the introduced Adam optimization algorithm are anticipated to improve the capacity of the fault diagnosis model.

4. Numerical simulation

4.1. Numerical model of gearboxes

Drawing from an extensive collection of fault information related to point absorber wave energy converters, components are categorized based on their functional modules and environmental loads, and a Failure Mode and Effects Analysis (FMEA) method [56] is used to identify that the gearbox is the critical components with the highest risk priority.

The gear exhibits the highest failure rate and a wide range of fault modes, making it the most critical factor affecting gearbox reliability. Therefore, this paper analyzes the common fault modes of gears in gearboxes. Typical gear fault modes include broken gear tooth, gear pitting and cracked gears. Table 1 concludes the percentage of various types of gear faults, with a notably high occurrence of broken tooth and pitting faults.

In this study, the rack and pinion transmission gearbox of a point absorber wave energy converter is examined, with a consideration of the effects of flexible deformation and changes in the gear transmission process [57]. A rigid-flexible coupling dynamics simulation model is constructed through three-dimensional modeling, with a focus on addressing critical component flexibility, defining constraints and contact forces, and applying driving and loading conditions [58]. The simulation model is based on pertinent structural data parameters [7] and is constructed at a 1:4 scaling ratio. When modeling the numerical model of the gearbox, the Wave Test Station, referred to as the prototype in Ref. [59], operates in water depth conditions of 30m and has a wave power density of 4 kW/m and a maximum test capacity of 3 × 100 kW. Its PTO system has a conversion efficiency of 90 %.

Fig. 5 displays a schematic representation of the internal drive of the point absorber wave energy converter gearbox.

The key parameters for the gears and the double-sided rack are shown in Table 2 and Table 3, respectively.

In order to explore and analyze the vibration signal characteristics of the gearbox under different tooth conditions, a model of the target gear with various fault tooth conditions is developed. The modeling of the fault condition is shown in Fig. 6. The gear fault modes are described in detail below.

Table 1
Distribution of gear fault types.

Gear fault	Broken tooth	Pitting	Cracked	Scratch	Other
Proportion (%)	41	31	10	10	8

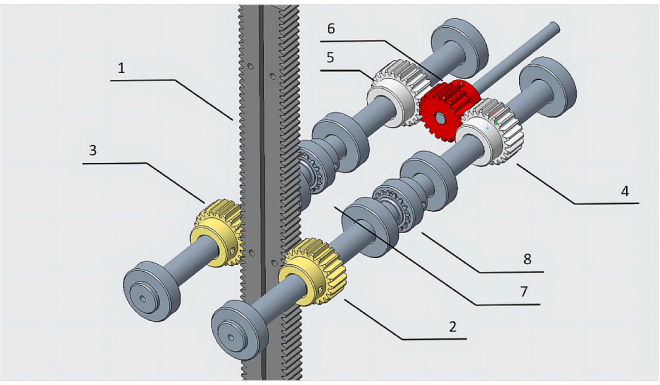


Fig. 5. Schematic diagram of the main subcomponents of the gearbox. 1 - rack; 2 - gear I; 3 - gear I; 4 - gear II; 5 - gear II; 6 - gear III; 7, 8 - ratchet.

- Broken tooth refers to the phenomenon where gear teeth wear out or develop cracks due to various factors, eventually leading to fracture. This is a severe gear failure that significantly impacts the normal operation of the gear mechanism. Moreover, broken gear components can pose additional risks to other machinery, potentially causing damage to the gearbox mechanism. Simulation of broken tooth involves modeling the gear by stretching and softening it in the modeling process.
- Pitting is caused by fatigue wear on the gear surface, typically manifested as irregular depressions on the gear surface. These depressions gradually enlarge and deepen, eventually leading to surface fatigue failure of the gear. Simulation of pitting involves stretching and softening irregular arcs.
- Cracked gear usually occur on the gear root surface or inside the gear, mainly due to gear overload causing stress at the gear root to exceed its fatigue limit. Gear cracks can cause increased wear on the gear surface, further exacerbating crack propagation. They can also cause a reduction in gear strength and negatively affect transmission performance. Simulation of cracks involves stretching the gear component at the root in the modeling process, cutting it in the tooth width direction, and performing softening treatment.

The gearbox structure of the point absorber wave energy converter is completely symmetrical on both sides of the rack. The unidirectional motion only drives one side of the structure to produce axial rotation. To improve the efficiency of the simulation calculation, only one side of the gearbox structure is flexibly modelled. The completed rigid-flexible coupling dynamics model of the gearbox is shown in Fig. 7.

A vertical simple harmonic motion is applied to the rack, with the same amplitude as the wave height and aligned with the wave phase. The relative motion of the PTO system at wave height $H = 0.6$ m and wave periods $T = 4$ s and $T = 6$ s in the sea is taken as the displacement change state of the rack input according to reference [59] and related sea wave information [60]. To facilitate the normalization of vibration signals from numerical simulation results and to align with the marine conditions simulated by the experimental platform described in Section 5, the selected wave conditions for this study only consider average conditions and do not account for the impact of seasonal variations on wave energy converter performance. The step size is set to 0.001 s, and the default solver method is used to resolve the simulation of the rigid-flexible coupled gearbox dynamics model. The vibration acceleration signal states of the gearbox are investigated for four tooth conditions with periods $T = 4$ s and $T = 6$ s, including normal, broken tooth, pitting and cracked.

Table 2
Key parameters of gears.

	Module number (mm)	Number of teeth	Diameter of indexing circle (mm)	Diameter of tooth top circle (mm)	Degree of tooth (mm)	Tooth thickness (mm)
Gear I (subcomponent 2 in Fig. 5)	2	23	46	50	20	3.14
Gear II (subcomponent 4 in Fig. 5)	2	23	46	50	20	3.14
Gear III (subcomponent 6 in Fig. 5)	2	20	40	44	20	3.14

Table 3
Parameters of the double-sided rack.

Modulus (mm)	Number of teeth	Rack length (mm)	Rack width (mm)
2	78	500	20

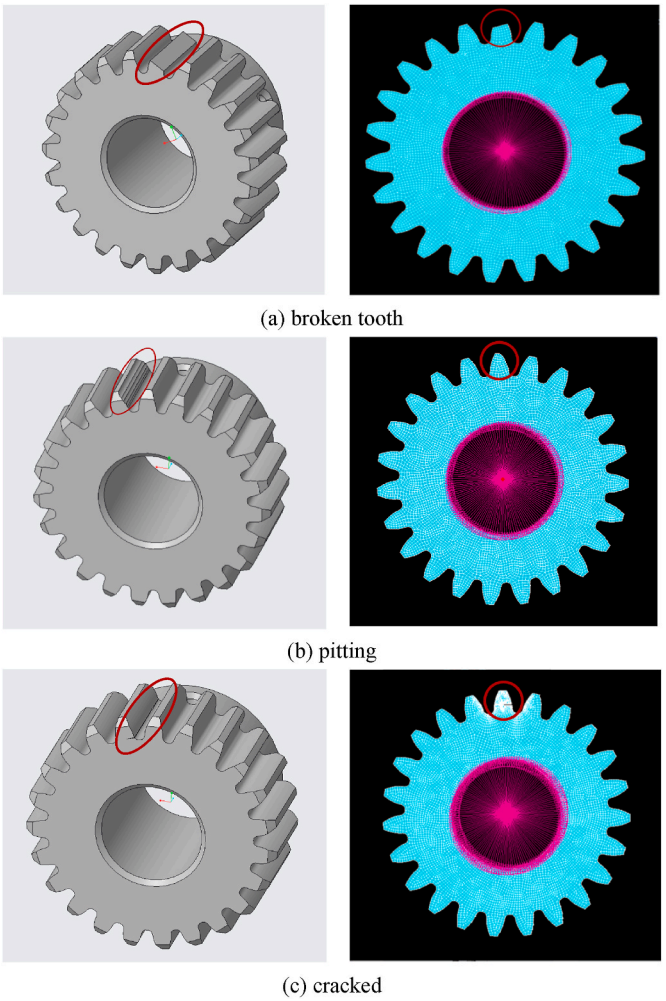


Fig. 6. Modeling of the fault tooth condition.

A time-domain plot is generated to display the acceleration signals for different tooth conditions, as depicted in Fig. 8. The acceleration signals for normal tooth conditions are represented in red, while those for broken, pitting, and cracked teeth are depicted in blue. The figure shows that the acceleration signal of the gearbox in the broken tooth condition is more variable, and the vibration distribution is significantly different from the normal tooth condition. The acceleration signals in the pitting and cracked conditions show less disparity from the normal condition. Although there are still abrupt changes in the vibration signal

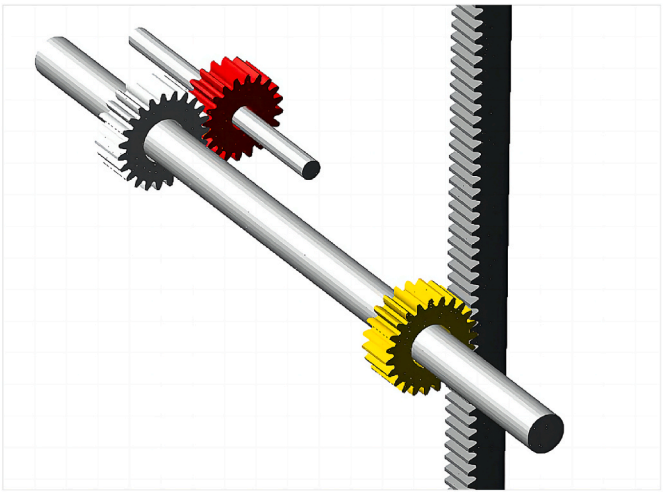


Fig. 7. Rigid-flexible coupling dynamics model of the gearbox.

in specific local areas, they contain fewer distinctive characteristics of a fault signal compared to the broken tooth condition.

To obtain the frequency domain acceleration curves for the different tooth conditions, the Fourier transform is applied to the acceleration signals, as shown in Fig. 9. At each frequency order, the distribution between normal and faulty tooth conditions appears similar. The spectrum of the faulty gear is distinguished by the presence of specific sidebands, and the width of the frequency band varies slightly depending on the tooth condition. However, the overall differences are not highly significant.

Based on the simulation results, a database of gear acceleration and speed signals is created. This database includes data from normal operating conditions and three distinct fault conditions, i.e., broken gear, pitting, and cracked gear. This database serves as a basis for developing fault diagnosis method models. It also acts as a reference and data comparison source for subsequent experimental verification of the gearbox vibration signal acquisition scheme.

4.2. Data preprocessing

The training samples used are the simulated speed and acceleration signals, each having four dimensions. These dimensions include the acceleration signals along the *X*, *Y*, and *Z* axes of the gears as well as the speed signals along these axes. All signals are generated with a simulation step size of 0.001 s.

To ensure sufficient training and testing datasets, the sample size is increased by performing overlapping slicing operations on the data. The dataset is further split into training and validation sets, while the test set data, which does not enter the training network, is reserved and handled in the same way.

The simulation data exhibits irregular and large fluctuations in the signal during the transition from the initial stage to the stable stage. Thus, the first segment (0 s–0.2 s) of the simulation data is removed

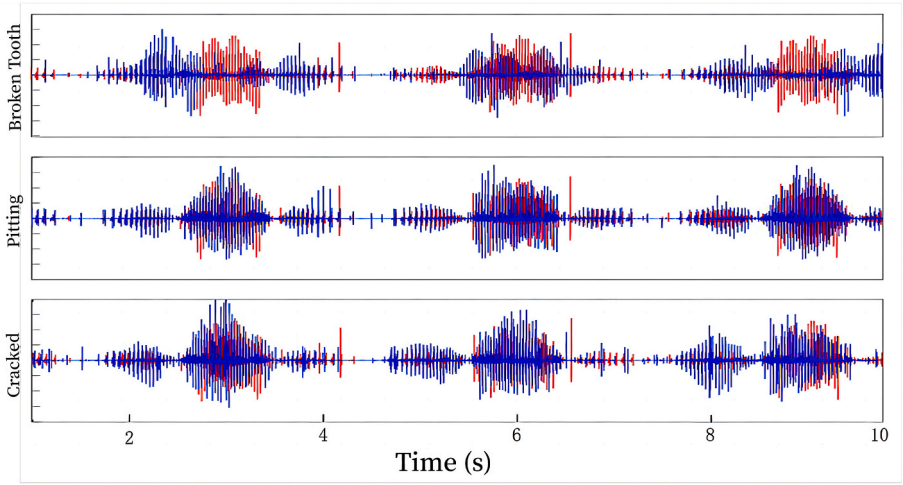


Fig. 8. Comparison of acceleration vibration signals of gear I under different tooth conditions.

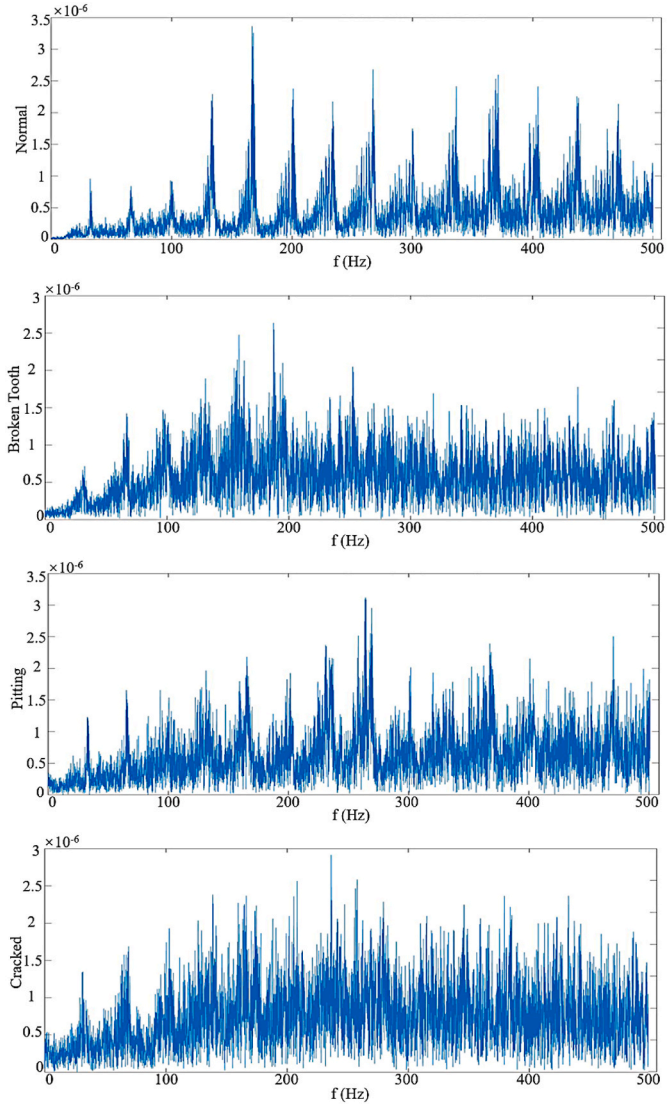


Fig. 9. Spectrum of acceleration of gear I in the *Y*-axis direction under different conditions.

before vectorization [61], and only the subsequent stable data signal is retained. For smoother learning convergence, the original data features need to be normalized. The mean value of all data is set to 0, and the standard deviation is set to 1. To address the problem of label distances in classification problems, the input labels are presented as probability distributions. The one-hot encoding of data labels in this paper is shown in Table 4.

The simulation data is sampled at a rate of 1 kHz. Total 2000 sets of data points are set up, each covering a duration of 2 s in a single sample. The data is sampled using a sliding window approach with a fixed step.

4.3. Results

Fig. 10 shows the changes in both the accuracy and loss function of the model trained on the simulated dataset. It can be seen that the model converges very quickly during the initial iterations. The verification accuracy stabilizes after around 20 iterations and finally maintains at 100.00 %. The loss values show similar pattern, dropping rapidly and then settling at a level close to zero.

For the diagnostic results of the test set, the confusion matrix obtained from the test set when using the trained model is shown in Fig. 11. The results show that the proposed method shows good learning and fitting capabilities on the simulation dataset. The fault diagnosis accuracy reaches 94.40 %, in which the accuracy for the broken tooth condition reaches as high as 100.00 %, but there are still the cases of misjudging the normal tooth condition.

Moreover, to address the problem of model overfitting on the validation dataset and to consider the number of rounds required to achieve a stable validation accuracy, the Dropout algorithm is introduced to the LSTM layer for network regularization. By comparing the confusion matrices of the fault diagnosis test with different Dropout ratios, it is found that the most favorable results when training the model with a Dropout value of 0.3, as shown in Fig. 12. The fault diagnosis accuracy reaches 99.47 %.

Table 4
One-hot encoding of labels.

Category Tags	Digital labels	One-hot encoding
Normal	0	[1,0,0,0]
Broken tooth	1	[0,1,0,0]
Pitting	2	[0,0,1,0]
Cracked	3	[0,0,0,1]

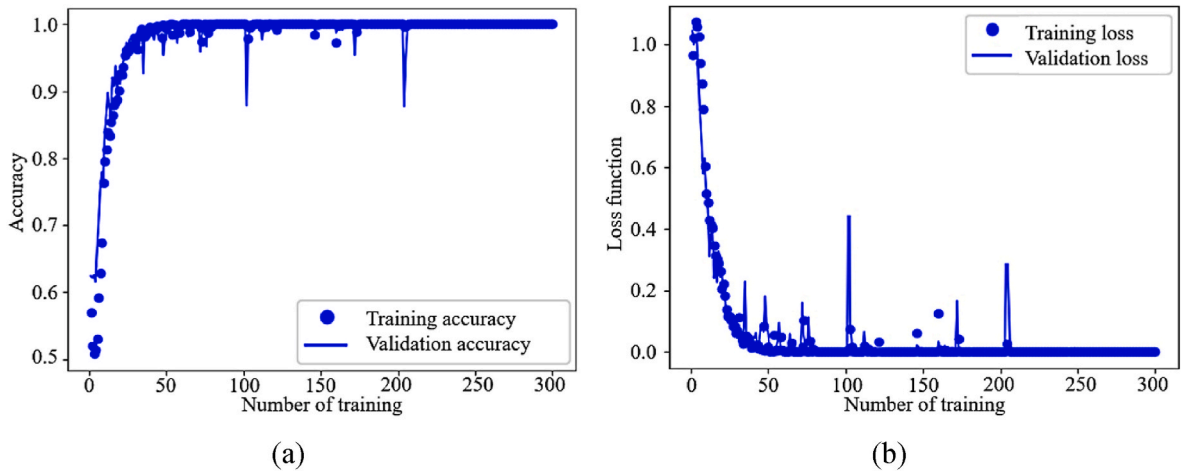


Fig. 10. (a) Change in accuracy with training rounds; (b) Change in loss with training rounds.

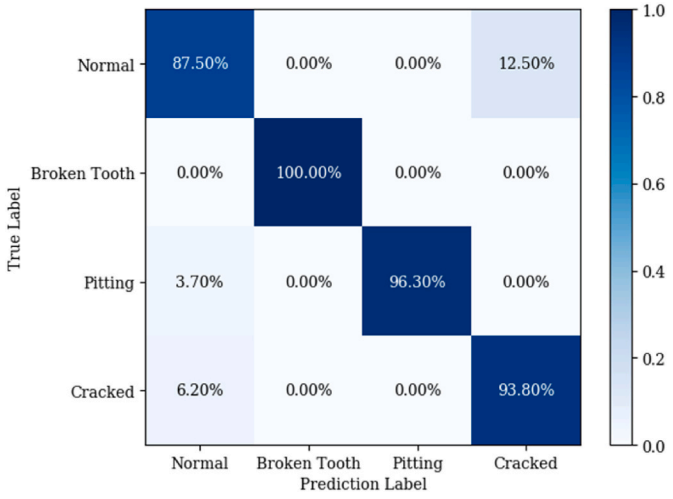


Fig. 11. Confusion matrix of CNN-LSTM on the simulation dataset.

4.4. Comparative study and discussion

In order to assess the fault diagnosis capability and demonstrate the superiority of the proposed method, several commonly used methods, including CNN, LSTM, RNN, CNN-RNN, CNN-Gated Recurrent Unit (GRU), are selected for comparison with the CNN-LSTM method.

Among the various methods mentioned, CNNs are capable of automatically learning fault-related features within the data and extracting these features through convolutional kernels in the convolutional layers. However, the pooling layer may overlook some subtle signal changes, leading to a deterioration in the diagnosis of certain types of faults. RNN, as a traditional and simple neural network for time series processing, can capture dynamic characteristics within time series, which is crucial for fault diagnosis. In RNN, due to the multiplicative effect of weights, gradients during backpropagation may diminish or increase rapidly, making it difficult for the network to learn long-term dependencies. This can significantly impact the performance of fault diagnosis models trained by RNN. GRU and LSTM networks have evolved from the foundation of RNN. The core advantage of LSTM is its ability to learn long-term dependencies involved in time series data, while also introducing gating mechanisms to control the flow of fault feature information, effectively addressing the issue of gradient vanishing. However, compared to other networks, LSTMs have higher computational complexity. GRUs simplify the network structure of LSTM by merging the forget gate and input gate of LSTM into an update gate. GRU also introduces a reset gate to determine how much of the previous information the network should forget and which parts of the state to update. These configurations make GRU more effective when dealing with shorter sequences or when the dataset is not very large.

CNN-RNN, CNN-GRU, and CNN-LSTM are the models that combine convolutional and recurrent networks, integrating the powerful feature extraction capabilities of CNN with the advantages of recurrent networks in handling time series data. The analysis of fault signals in gearboxes presents challenges such as strong time-variant characteristics, difficulties in fault feature extraction, and significant time series dependencies. Compared to single neural network models, these hybrid methods leverage the strengths of different methods, bringing about significant potential to enhance fault diagnosis performance.

To ensure the objectivity in the comparison results, all methods are subjected to the same optimization function, computational capacity, learning rate and other hyperparameters. Moreover, the identical dataset is utilized for both training and testing. The outcomes are shown in Fig. 13 and Table 5.

The comparison results reveal that the CNN model converges faster and achieves high diagnostic accuracy in broken tooth and pitting conditions. However, it fails to differentiate between normal and cracked conditions effectively. RNN and LSTM models exhibit slow convergence during training. RNN fails to achieve stable convergence, and the accuracy is very low. LSTM, on the other hand, provides better training results, with a fault diagnosis accuracy rate of 74.33 %. However, it encounters challenges in accurately identifying local features of the cracked signal.

The three convolutional-recurrent neural networks, including the methods proposed in this paper, perform well across various aspects, including validation accuracy, test accuracy, and the number of iterations needed to reach stability. Combining CNN and RNN, two networks that individually struggle to achieve good diagnostic results, results in fault accuracy no less than LSTM. It ranks third in terms of convergence speed, following the CNN-GRU and CNN-LSTM. The CNN-GRU and CNN-LSTM have remarkably close performance. GRU, as a simplified version of LSTM, performs well when combined with convolutional networks but falls slightly short of the CNN-LSTM, which achieves a diagnostic accuracy of 99.47 %.

Analysis of the different approaches employed in the study shows that CNN is efficient in capturing local patterns and spatial relationships through the convolutional and pooling layers. Although it may be inadequate in capturing long-term temporal dependencies that are necessary to differentiate between normal and cracked conditions, the results still show that this approach has a good accuracy and outstanding convergence speed. RNN can capture temporal dependencies, but suffers from the vanishing gradient problem, which can make it difficult to

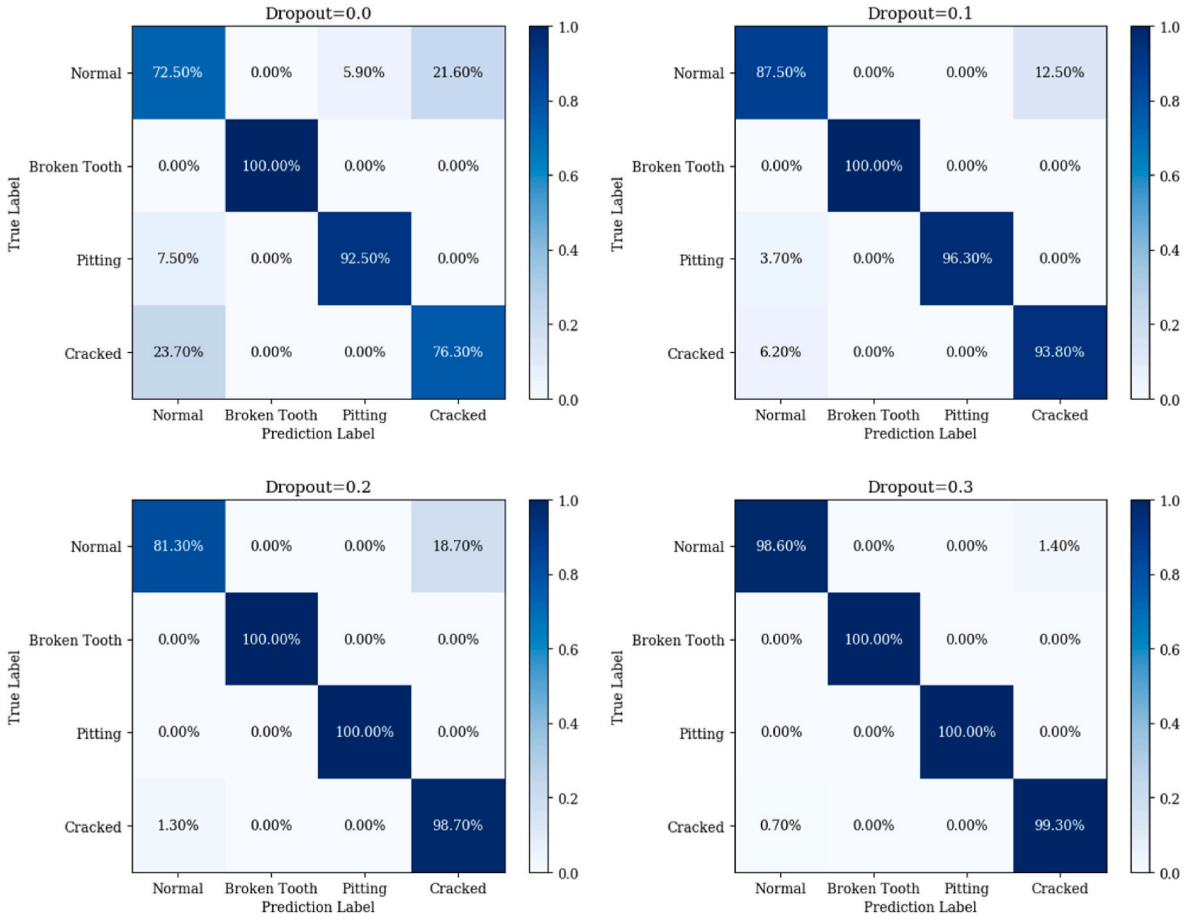


Fig. 12. Confusion matrices of CNN-LSTM on the simulation dataset with different Dropout ratios.

identify complex patterns in vibration signals correctly. This drawback leads to the result that the performance of RNN is the worst among different methods. LSTM networks have a specific design to model sequential data and can leverage memory cells and gates to learn long-term dependencies, enhancing the capacity to diagnosis fault modes. While LSTM networks can capture temporal information, they may not efficiently extract less spatial features that are crucial in analyzing cracked signals, resulting in the second longest stable round. CNN-RNN, CNN-GRU, and CNN-LSTM have achieved high test accuracy. However, CNN-RNN may not fully exploit spatial features in vibration signals, resulting in suboptimal performance in test accuracy compared to CNN-GRU and CNN-LSTM. Additionally, CNN-GRU has gating mechanisms that are simplified compared to LSTM and may limit their effectiveness in capturing complex dependencies in vibration signal data. Its performance in test accuracy is slightly worse than CNN-LSTM.

In summary, the CNN-LSTM model theoretically offers the best fault diagnosis performance among the compared methods because it possesses the advantages of CNN in local feature information extraction and the proficiency of LSTM in processing time-series signals. Comparative results demonstrate these unique advantages of CNN-LSTM, leading to superior learning convergence efficiency and diagnostic accuracy. It achieves outstanding fault diagnosis results with fewer iterations. The method outperforms others in terms of the number of iterations required to achieve the same accuracy, as well as its performance on the test set.

5. Experimental studies

The vibration signals generated by the gearbox of the wave energy converter in the experiment usually contain significant noise, environmental information and periodic clutter generated with the driving force. Consequently, these signals carry complex feature information, making it essential for the fault diagnosis model to possess strong noise reduction and feature extraction capabilities to identify fault types accurately. Moreover, different faults may exhibit similar characteristics, further complicating the task of fault classification and identification. To validate the effectiveness of the proposed CNN-LSTM-based fault diagnosis method in processing real signals, a gearbox vibration signal acquisition platform is established.

5.1. Experimental setup

The gearbox of the point absorber wave energy converter, mainly consists of a variable speed drive motor, a control circuit board, a crank linkage mechanism and the internal transmission components of the gearbox. The experiment uses a variable speed drive motor combined with a crank linkage mechanism to simulate the vertical motion input of a buoy to the rack and a speed control circuit board to achieve variable speed regulation (rack motion cycle range of 5 s–8 s). The motion of the rack drives the subsequent gear, and this motion is then transferred to the output gear through the use of a ratchet overrunning clutch and a transmission gear. To simulate the load of a generator, a fixed load is applied at the end of the output shaft. Fig. 14 shows the experimental platform and internal structure. To study the vibration signals of the gearbox under different faulty gear conditions, the experiments are equipped with faulty gear kits with different fault modes for the target gears. The specific parameters of the gears and gear racks inside the gearbox used in this study are consistent with the structural parameters analyzed in the dynamics simulation.

The experimental signal acquisition equipment includes two sets of ICP/IEPE tri-axis piezoelectric acceleration sensors of CT1010SLFP, one

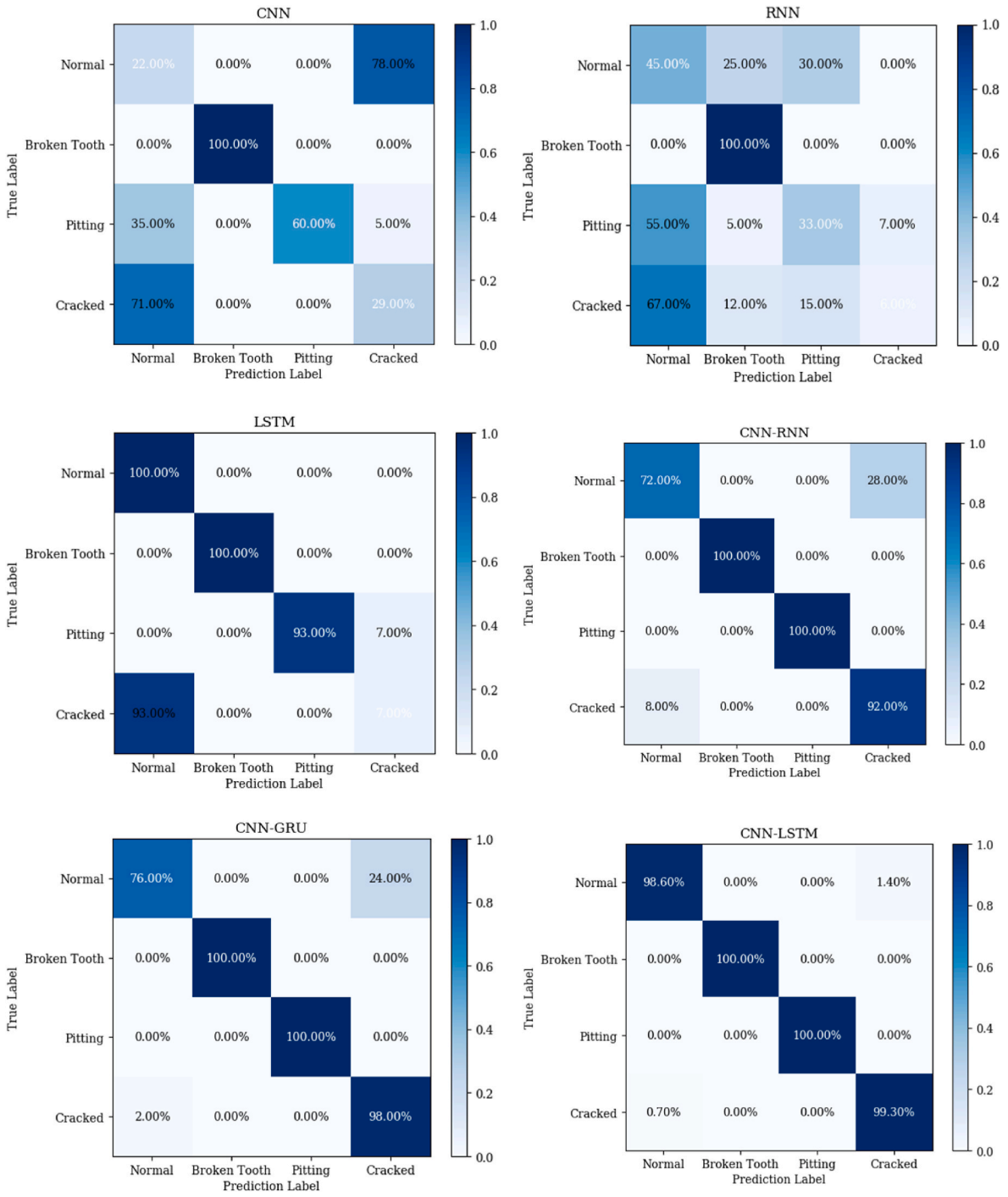


Fig. 13. Confusion matrices of different models on the simulation dataset.

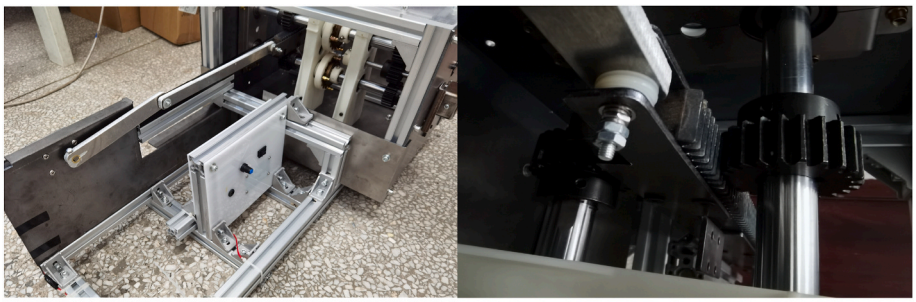
Table 5
Performance comparison of the different models for fault diagnosis.

Models	Verification accuracy (%)	Test accuracy (%)	Stable rounds
CNN	96.82	53.60	6
RNN	55.36	46.12	83
LSTM	99.79	74.33	52
CNN-RNN	100.00	89.00	15
CNN-GRU	100.00	95.50	12
CNN-LSTM	100.00	99.47	12

set of CZ400 Hall-type tachometer sensors, one set of CT5210 ten-channel constant current adapters, and one MCC USB-1608G data acquisition card. The specific experimental equipment and arrangement are shown in Fig. 15.

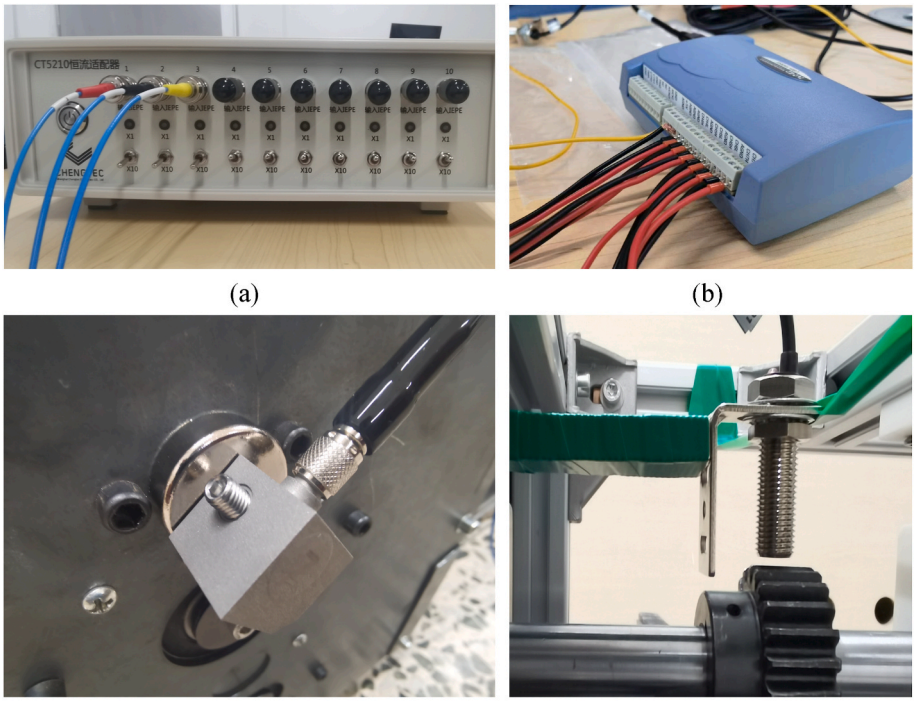
5.2. Results

The experimental sampling frequency is set as 1 kHz, and data is collected from seven sampling channels, including six sets of acceleration vibration signals from two tri-axial acceleration sensors and a square wave pulse signal from a speed sensor. Data collection for each operating condition is carried out for 20 min. The sampling conditions



(a) (b)

Fig. 14. (a) Experimental platform; (b) Internal structure.



(a) (b)

(c) (d)

Fig. 15. Experiment equipment: (a) CT 5200 series constant current adapter; (b) USB-1608G data acquisition card; (c) ICP/IEPE piezoelectric accelerometer; (d) CZ400 Hall speed sensor.

include normal, broken, pitting and cracked gears at three different periods of $T = 5.6$ s, $T = 6.0$ s and $T = 6.7$ s.

To facilitate the identification of faulty parts, they were marked sequentially. Taking the gearbox vibration signal acquisition under normal gear conditions as an example, the motor speed is adjusted to the preset value using the speed control board. Signal acquisition commences once the rotor had achieved stable operation. Data acquisition is performed for a single working condition, and the collected data is saved. The speed is then adjusted to the next speed condition for further data collection. This process continues until data has been collected for each cycle condition.

After collecting data for various cycle and torque conditions, torque adjustments were made to collect data for each cycle condition under different torques. Once data collection for different cycles and torque conditions is completed, the gear to be measured is replaced, allowing for signal data acquisition for different cycles and torque conditions under different gear conditions.

The simulated and experimental acceleration vibration time-series signals are extracted. Figs. 16 and 17 are the change curves of the simulated and experimental acceleration vibration signals of gear I in the same motion state in the time domain respectively. According to the position of the peaks and valleys, it can be seen that the two signals have the same cyclic movement pattern. The acceleration movement period, peak values, and change trends are consistent under identical rack movement inputs. However, it is noted that the experimental data is influenced by noise, environmental factors, and mechanical linkages during the acquisition process. As a result, the average amplitude of the experimental data is higher at the same sampling rate when compared to the simulated data. The smaller amplitude vibrations are still present during periods of low energy, even in the experimental data.

The fault diagnosis model and parameters established based on the simulated data are trained and verified using data collected from the experiments under different working conditions. Fig. 18 shows the changes in both the accuracy and loss function of the model trained on the experimental dataset.

Fig. 19 shows the confusion matrix of CNN-LSTM on the experimental dataset. The analysis reveals that the overall fault diagnosis accuracy of the experimental data is 83.15 %. Examining the results from the confusion matrix for the test set, it can be concluded that the proposed method demonstrates strong fault diagnosis accuracy in normal

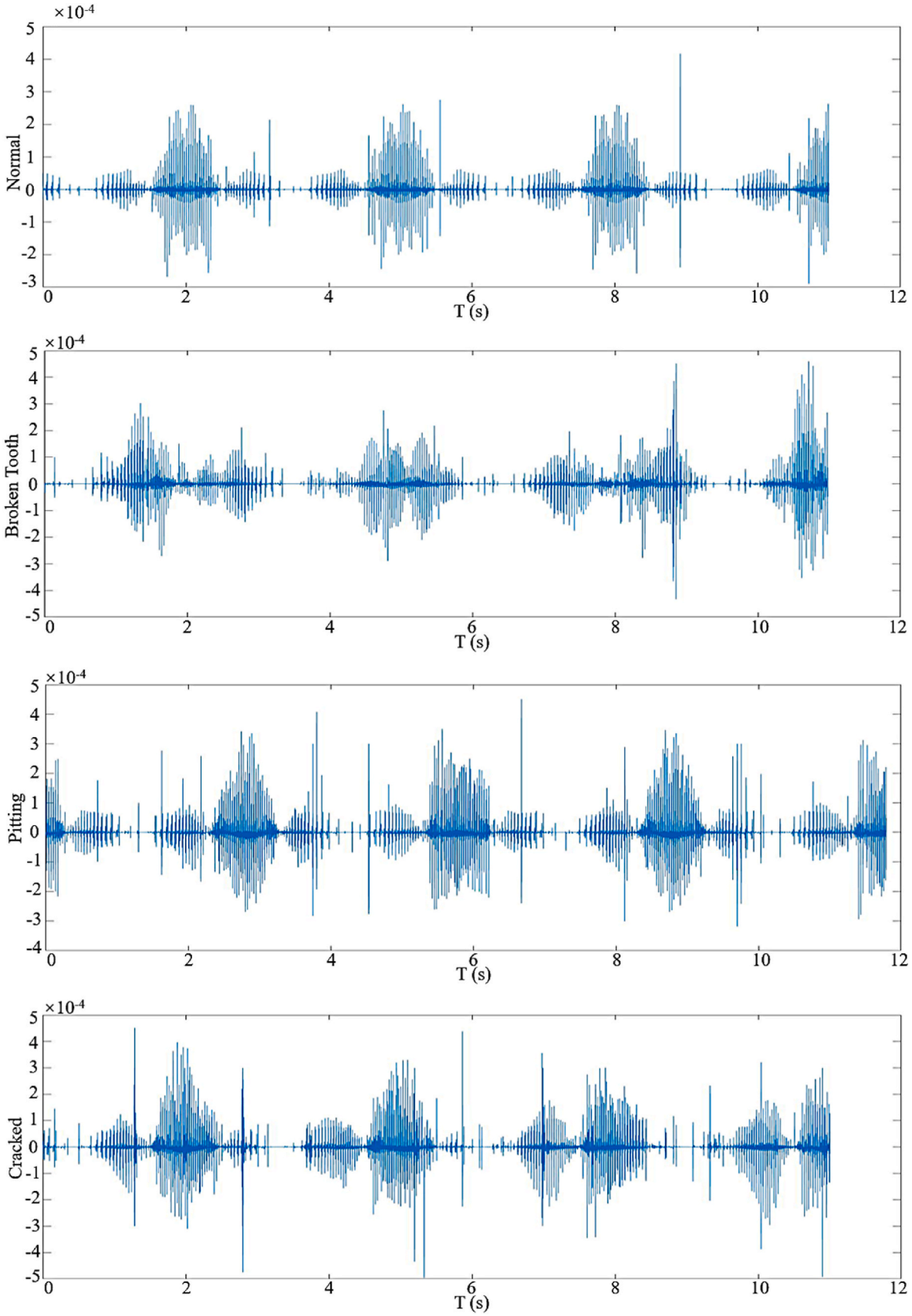


Fig. 16. Time domain diagram of simulated signals.

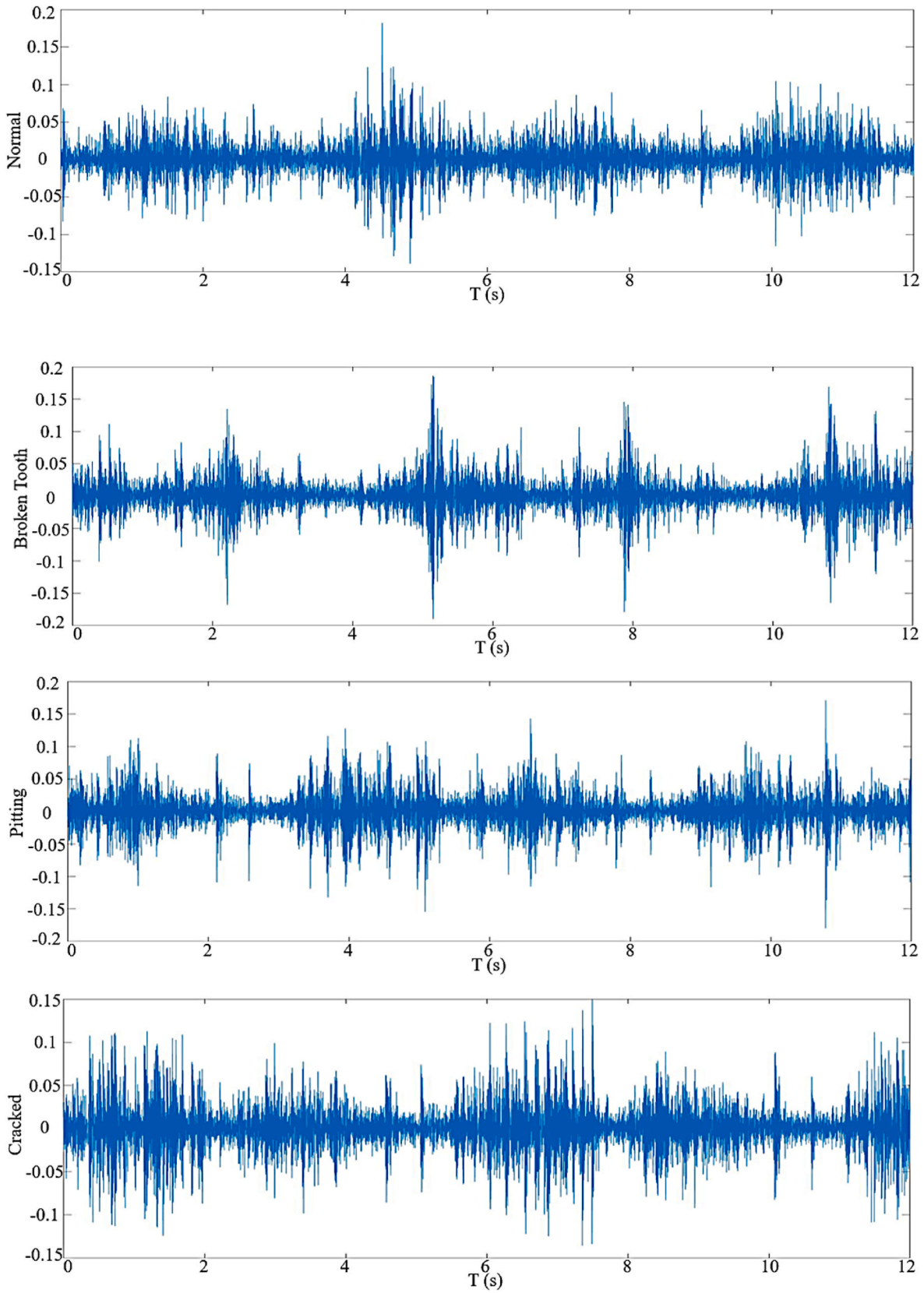


Fig. 17. Time domain diagram of experiment signals.

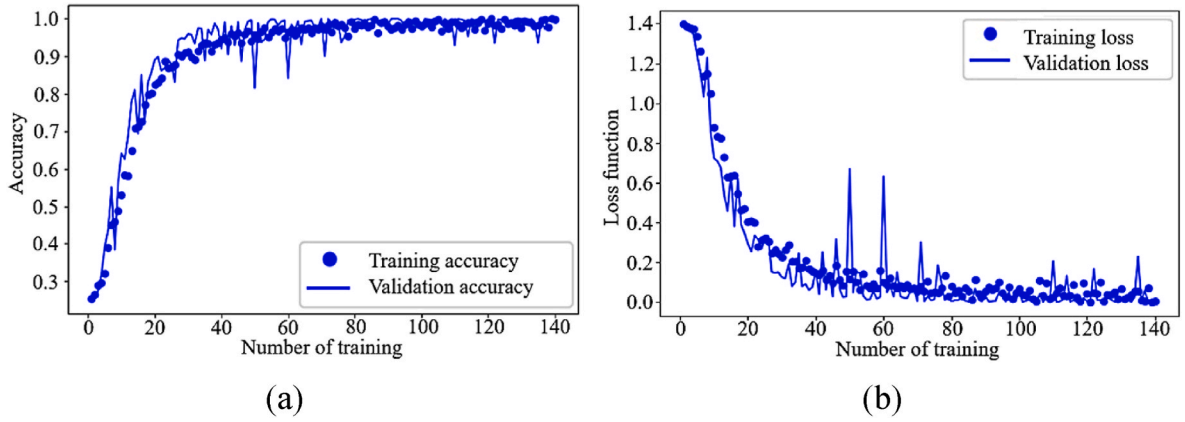


Fig. 18. (a) Change in accuracy with training rounds; (b) Change in loss with training rounds.

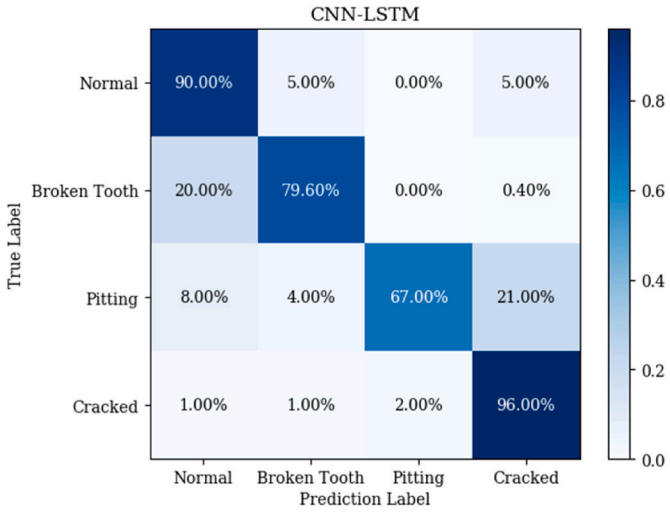


Fig. 19. Confusion matrix of CNN-LSTM on the experimental dataset.

and cracked tooth conditions. However, its performance is less consistent when diagnosing pitting and broken tooth conditions, as it may easily confuse these conditions with other vibration characteristics.

5.3. Comparative study and discussion

The fault diagnosis results of the different models introduced in Section 4.4 are compared using the experimental vibration signals. A comparison of the performance of different models for fault diagnosis is shown in Fig. 20 and Table 6. It can be seen that the diagnostic accuracies of all the models on the four fault conditions have decreased. The CNN method is not capable enough to distinguish between the normal condition and the cracked condition, achieving 90.00 % accuracy in the diagnosis of the fault of the broken tooth. The RNN remains the worst performing model, and the number of convergence rounds further increases. The LSTM model has an overall fault diagnosis accuracy of 57.45 %, and can hardly identify the cracked condition of the faults accurately. The overall fault diagnosis accuracies of CNN-RNN, CNN-GRU, and CNN-LSTM are 73.67 %, 76.67 %, and 83.15 %, respectively. CNN-RNN and CNN-GRU both have the problem of identifying a portion of normal condition and pitting condition cases as cracked condition. Compared to other models, CNN-LSTM has the highest accuracy in identifying experimental faulty vibration signals affected by ambient noise, while the number of convergence rounds is lower. The varying results of different configurations can be attributed to the characteristics of the model, as explained in Section 4.4.

5.4. Improvement of the model

In this section, to improve its training efficiency on real vibration signals and increase the fault diagnosis accuracy, the CNN-LSTM model is improved by adding Adam algorithm and improving network structure to handle complex experimental conditions. The generalization ability and convergence speed are enhanced to improve the diagnostic performance. The improvements of network structure focus on two main aspects, i.e., optimization and generalization of the deep learning network.

As shown in Fig. 21, the impact of varying the number of channels in the first and second convolution layers on the fault diagnosis accuracy is explored, while keeping all other parameters constant. The color gradient represents the accuracy of the model for different combinations of channel numbers in the two layers. It is found that accuracy above 90.00 % is mainly achieved when using 64–92 channels in the first layer and 92–128 channels in the second layer.

For the experimental dataset, this section adjusts the network structure of the original CNN-LSTM model to reduce the fit of the learning model to the training data by reducing the number of model layers and the number of units per layer. The adjusted parameters include the number of layers, the convolution kernel sizes of the convolutional layer, the convolutional step sizes, the size of the max pooling layer, LSTM layer nodes, and the size of the fully connected layer size.

The improved CNN-LSTM fault diagnosis model is trained and validated on the experimentally collected data under different working conditions. The training and validation accuracy and loss function variation curves with the number of iterations are shown in Fig. 22.

The fault diagnosis results of the improved model on the test set are shown in Fig. 23. The model achieves a fault diagnosis accuracy of over 86.00 % for each tooth condition, 96.70 % for cracks and an overall fault diagnosis rate of over 92.00 %.

It can be seen that, after the introduction of the Adam optimization algorithm, different learning rates are calculated adaptively for the network structure parameters of CNN-LSTM models. Compared with the 83.15 % fault diagnosis accuracy before optimization, the average fault diagnosis accuracy of the improved model reaches 92.00 %. The recognition rate for pitting faults increased from 67.00 % to 92.00 %. In summary, the improved model enhances the convergence speed and validation accuracy while maintaining the same training data. It substantially reduces the risk of overfitting, thus improving the generalization capability and fault diagnosis accuracy.

6. Conclusions and future research

The gearbox is the critical component of wave energy converters. Efficient and accurate fault diagnosis is significant for the development of the wave energy converter and its future large-scale application. In

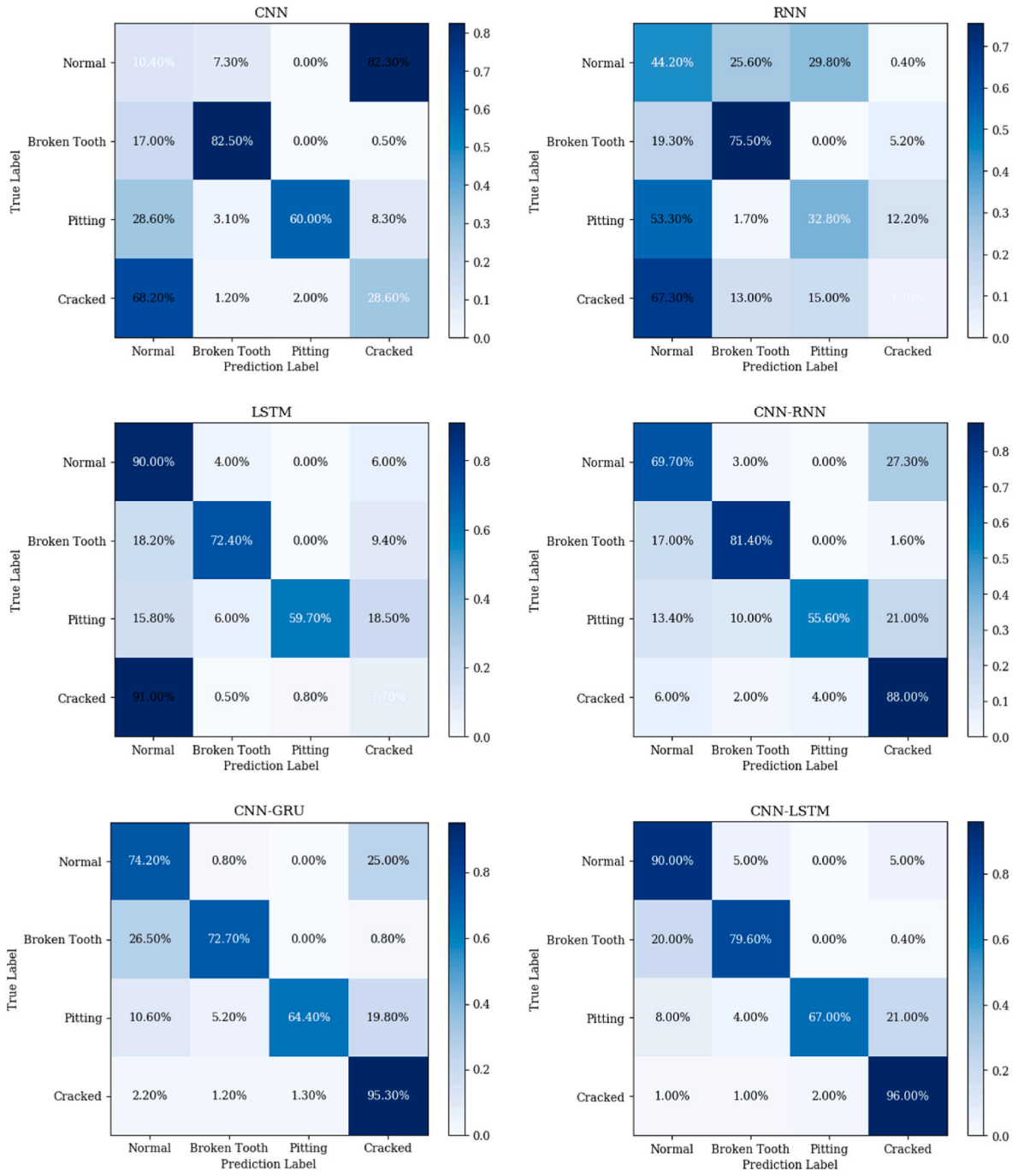


Fig. 20. Confusion matrices of different models on the experimental dataset.

Table 6
The comparison of the performance of different models for fault diagnosis.

Models	Time (s)	Test accuracy (%)	Stable rounds
CNN	70	45.38	9
RNN	1043	39.30	118
LSTM	526	57.45	73
CNN-RNN	155	73.68	22
CNN-GRU	148	76.65	18
CNN-LSTM	142	83.15	17

this paper, a fault diagnosis model for the gearbox of the point absorber wave energy converter is presented based on CNN-LSTM networks, with the strengthen of the efficient extraction of local features from signals and the effective analysis of time-series data. The Adam algorithm is introduced to the model to address situations where gradients within tensors exhibit unstable changes. Vibration signals from the gearbox of a point absorber wave energy converter under different fault modes are simulated using a rigid-flexible coupled dynamics simulation model. These simulated vibration signals are used as training datasets and test datasets. The performance of the fault diagnosis method is verified by using the dataset generated. To assess the effectiveness of the CNN-LSTM-based fault diagnosis method in processing real vibration signals, an experiment is carried out to acquire gearbox vibration signals. The fault diagnosis performance achieved with the experimental data across various working conditions is compared. Improvements are made to the structure and parameters of the fault diagnosis model in response to the experimental data. The main conclusions are drawn as follows.

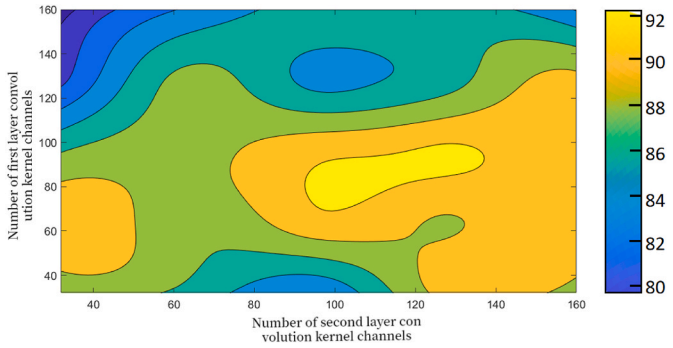


Fig. 21. Effect of the number of convolutional kernel channels on fault diagnosis rate.

(1) In the time-domain plot, it is observed that the acceleration signal under the broken tooth condition shows a significant deviation from normal conditions, with a distinct difference in vibration distribution. In comparison, under pitting and cracked conditions, the acceleration signal shows less difference compared to the normal condition. Compared with broken tooth condition, the fault signal characteristics under these two conditions are less. In the frequency-domain plot, it is observed that the distribution between normal and fault tooth conditions is similar, and the width of the frequency band of fault gears varies slightly depending on the tooth condition. However, the overall difference is not significant.

(2) The CNN-LSTM model integrates two deep learning methods to realize efficient extraction of local features from fault signals and enhance the performance in time-series analysis. Through comparisons with other network structures, i.e., CNN, LSTM, RNN, CNN-RNN, CNN-GRU, it is observed that the proposed model has the highest average accuracy of 99.47 % and the lowest stable round of 12 on the simulation dataset. The CNN-GRU model and CNN-RNN model achieves accuracies of 95.50 % and 89.00 %, respectively. Although the LSTM model has high accuracies in normal, broken tooth, and pitting conditions, it can hardly recognize the cracked condition and requires a high stable round. The CNN model has the lowest stable round of 6, but its average fault diagnosis accuracy only reached 53.60 %. The RNN model performs the worst among all the compared models, with the accuracy of only 46.12 % and the highest stable round of 83.

(3) When assessing the performance of different methods by using the experimental vibration signals, it is observed that the diagnostic accuracies of all the models have decreased. Among them, the CNN-LSTM model has the highest accuracy of 83.15 % and a stable round of 17. The CNN-GRU model and CNN-RNN model present similar performances in fault diagnosis for different fault modes of gears, with accuracies of 76.65 % and 73.68 %, and required stable rounds of 18 and 22, respectively. The LSTM model has a fault diagnosis accuracy of 57.45 % and can hardly identify the cracked condition of faults. The CNN model cannot distinguish between the normal condition and the cracked condition, achieving an average fault diagnosis accuracy of only 45.38 %. The RNN model remains the worst-performing model with an accuracy of 39.30 % and a stable round of 118. To address complex experimental conditions, improvements are further made to the CNN-LSTM model by incorporating the Adam algorithm and refining the network structure. The enhanced model achieves a fault diagnosis accuracy of over 86.00 % for each tooth condition and an average fault diagnosis rate exceeding 92.00 %. As a result, the CNN-LSTM model fault diagnosis rate reaches a higher level, demonstrating its generalizability and training effectiveness in processing complex real-world signals.

The main implications of this study are that potential gear faults can be diagnosed in real-time by monitoring vibration signals by using the wave energy converter gearbox fault diagnosis model based on CNN-LSTM. When the condition monitoring system detects fault signals, operators and maintenance service providers can perform maintenance activities within a specific weather window by analyzing and interpreting fault patterns, thereby avoiding further damage and long system

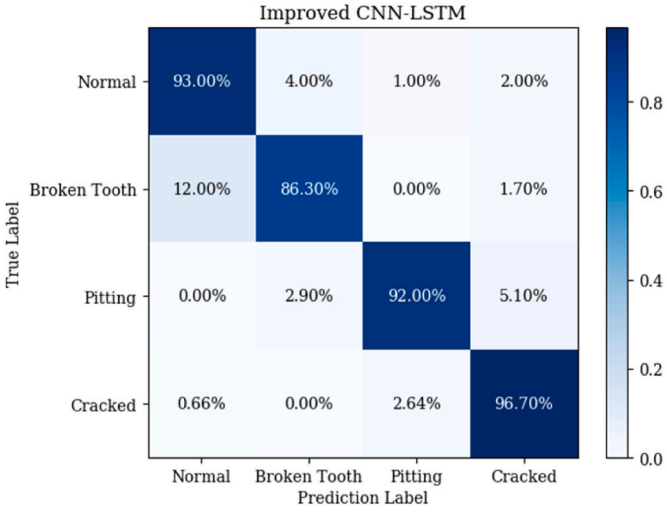


Fig. 23. Confusion matrix of the improved model on the experimental dataset.

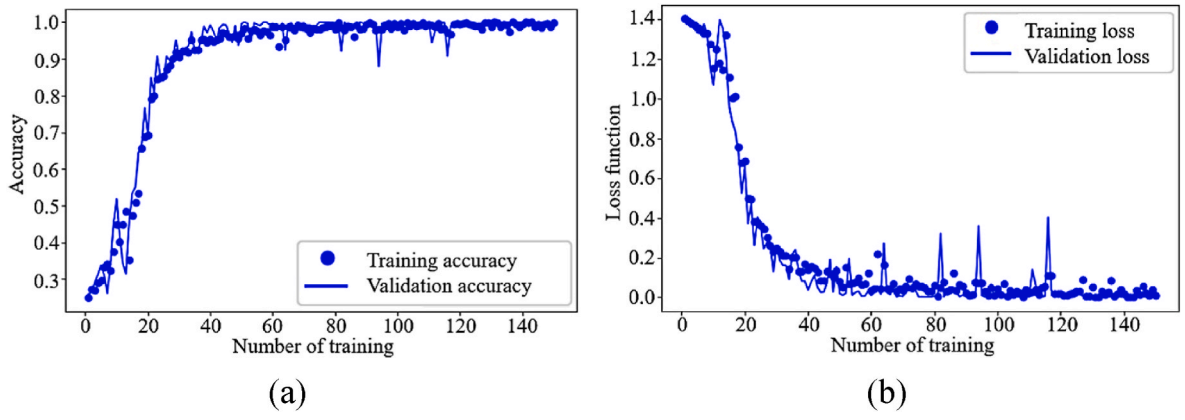


Fig. 22. (a) Change in accuracy with training rounds; (b) Change in loss with training rounds.

downtime. The proposed method mitigates the losses and risks caused by gearbox faults and reduces the time and cost associated with maintenance. This will help improve equipment reliability and maintainability, enhance production efficiency, save important maintenance resources over long-term operations, reduce cost of wave energy, and finally benefit its development.

However, there are still limitations of the research. The vibration signal of the gearbox of the point absorber wave energy converter has been somewhat simplified in this paper, and the influence of the complex marine environment has not been fully considered. Furthermore, the proposed fault diagnosis method does not consider variations in fault severity or compound faults, and the examination of fault conditions is not comprehensive enough. These limitations should be addressed in the future research (Table 6).

CRediT authorship contribution statement

Jichuan Kang: Writing – original draft, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Xu Zhu:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation. **Li Shen:** Writing – original draft, Visualization, Validation, Resources, Investigation, Formal analysis, Data curation, Conceptualization. **Mingxin Li:** Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research was funded by the National Natural Science Foundation of China (Grant No. 52101305).

References

[1] M.X. Li, X.L. Jiang, J. Carroll, R.R. Negenborn, A closed-loop maintenance strategy for offshore wind farms: incorporating dynamic wind farm states and uncertainty-awareness in decision-making, Renew. Sustain. Energy Rev. 184 (2023) 113535.
[2] S. Li, A. Coraddu, F. Brennan, A framework for optimal sensor placement to support structural health monitoring, J. Mar. Sci. Eng. 10 (12) (2022) 1819.
[3] M.X. Li, X.L. Jiang, R.R. Negenborn, Opportunistic maintenance for offshore wind farms with multiple-component age-based preventive dispatch, Ocean Eng. 231 (2021) 109062.
[4] S. Li, D.K. Kim, Ultimate strength characteristics of unstiffened cylindrical shell in axial compression, Ocean Eng. 243 (2022) 110253.
[5] M. Centeno-Telleria, J.I. Aizpurua, M. Penalba, Impact of accessibility on O&M of floating offshore wind turbines: sensitivity of the deployment site, Trends in Renewable Energies Offshore (2022) 847–855.
[6] M.X. Li, M. Wang, J.C. Kang, L.P. Sun, P. Jin, An opportunistic maintenance strategy for offshore wind turbine system considering optimal maintenance intervals of subsystems, Ocean Eng. 216 (2020) 108067.
[7] B.Y. Guo, T.Y. Wang, S.Y. Jin, S.L. Duan, K.D. Yang, Y.M. Zhao, A review of point absorber wave energy converters, J. Mar. Sci. Eng. 10 (10) (2022) 1534.
[8] S. Ahn, V.S. Neary, K.A. Haas, Global wave energy resource classification system for regional energy planning and project development, Renew. Sustain. Energy Rev. 162 (2022) 112438.
[9] T. Aderinto, H. Li, Ocean wave energy converters: status and challenges, Energies 11 (5) (2018) 1250.
[10] F. Mwasilu, J.W. Jung, Potential for power generation from ocean wave renewable energy source: a comprehensive review on state-of-the-art technology and future prospects, IET Renew. Power Gener. 13 (3) (2019) 363–375.
[11] L. Ciappi, I. Simonetti, A. Bianchini, L. Cappietti, G. Manfrida, Application of integrated wave-to-wire modelling for the preliminary design of oscillating water column systems for installations in moderate wave climates, Renew. Energy 194 (2022) 232–248.
[12] J. Tan, W. Tao, A.J. Laguna, H. Polinder, Y. Xing, S. Miedema, A spectral-domain wave-to-wire model of wave energy converters, Appl. Ocean Res. 138 (2023) 103650.
[13] B.Y. Guo, J.V. Ringwood, Geometric optimisation of wave energy conversion devices: a survey, Appl. Energy 297 (2021) 117100.
[14] J. Tovar-Facio, B. Cansino-Loeza, J.M. Ponce-Ortega, Management of renewable energy sources, in: Sustainable Design for Renewable Processes, 2022, pp. 3–31.
[15] D.V.E. Barbosa, A.L.G. Santos, E.D. dos Santos, J.A. Souza, Overtopping device numerical study: openfoam solution verification and evaluation of curved ramps performances, Int. J. Heat Mass Tran. 131 (2019) 411–423.
[16] W. Sheng, Wave energy conversion and hydrodynamics modelling technologies: a review, Renew. Sustain. Energy Rev. 109 (2019) 482–498.
[17] A. Kolios, M.L.F. Di, L. Wang, L. Cui, Q.H. Sheng, Reliability assessment of point-absorber wave energy converters, Ocean Eng. 163 (2018) 40–50.
[18] A. Clément, P. McCullen, A. Falcão, A. Fiorentino, F. Gardner, K. Hammarlund, G. Lemonis, T. Lewis, K. Nielsen, S. Petroncini, M.T. Pontes, P. Schild, B. O. Sjöström, H.C. Sørensen, T. Thorpe, Wave energy in Europe: current status and perspectives, Renew. Sustain. Energy Rev. 6 (5) (2002) 405–431.
[19] S. Li, D.K. Kim, Q.Q. Liang, Fibre-Based modelling for predicting the progressive collapse of cylindrical shells under combined axial compression and bending moment, Eng. Struct. 272 (2022) 114988.
[20] M. Centeno-Telleria, J.I. Aizpurua, M. Penalba, Computationally efficient analytical O&M model for strategic decision-making in offshore renewable energy systems, Energy 285 (2023) 129374.
[21] S. Li, D.K. Kim, S. Benson, A probabilistic approach to assess the computational uncertainty of ultimate strength of hull girders, Reliab. Eng. Syst. Saf. 213 (2021) 107688.
[22] L. Iraide, A. Jon, C. Salvador, I.M. Alegría, I. Kortabarria, Review of wave energy technologies and the necessary power-equipment, Renew. Sustain. Energy Rev. 27 (2013) 413–434.
[23] B. Drew, A.R. Plummer, M.N. Sahinkaya, A review of wave energy converter technology, Proc. Inst. Mech. Eng. A J. Power Energy 223 (8) (2009) 887–902.
[24] M.X. Li, X.L. Jiang, J. Carroll, R.R. Negenborn, A multi-objective maintenance strategy optimization framework for offshore wind farms considering uncertainty, Appl. Energy 321 (2022) 119284.
[25] A. González-Esculpi, C. Verde, P. Maya-Ortiz, Fault-tolerant control for a wave energy converter by damping injection, in: Proceedings of the 2021 IEEE Conference on Control Technology and Applications (CCTA), 2021, pp. 673–678.
[26] A. González-Esculpi, C. Verde, P. Maya-Ortiz, FDI study for a Wave Energy Converter by structural analysis, IFAC-PapersOnLine 53 (2) (2020) 13721–13726.
[27] G. Anthony, D. Beth, B. Tom, A. Ian, J. Lars, Reliability and O&M sensitivity analysis as a consequence of site specific characteristics for wave energy converters, Ocean Eng. 141 (2017) 493–511.
[28] A. González-Esculpi, C. Verde, P. Maya-Ortiz, Comparison of estimates of the excitation force for fault diagnosis in a Wave Energy Converter, IFAC-PapersOnLine 55 (6) (2022) 396–401.
[29] Y. Zhang, T. Zeng, Z. Gao, Fault diagnosis and fault-tolerant control of energy maximization for wave energy converters, IEEE Trans. Sustain. Energy 13 (3) (2022) 1771–1778.
[30] Y. Zhang, T. Zeng, Z. Gao, S. Turnock, D. Hudson, Fault diagnosis for wave energy converters with model uncertainties, IFAC-PapersOnLine 56 (2) (2023) 10880–10885.
[31] M. Li, J.C. Kang, L.P. Sun, M. Wang, Development of optimal maintenance policies for offshore wind turbine gearboxes based on the non-homogeneous continuous-time Markov process, J. Mar. Sci. Appl. 18 (2019) 93–98.
[32] M.X. Li, Towards Closed-Loop Maintenance Logistics for Offshore Wind Farms: Approaches for Strategic and Tactical Decision-Making, Delft University of Technology, 2023. PhD thesis.
[33] S.A. Mortazavizadeh, R. Yazdanpanah, D.C. Gaona, O. Anaya-Lara, Fault Diagnosis and condition monitoring in wave energy converters: a Review, Energies 16 (19) (2023) 6777.
[34] P. Bangalore, S. Letzgus, D. Karlsson, M. Patriksson, An artificial neural network-based condition monitoring method for wind turbines, with application to the monitoring of the gearbox, Wind Energy 20 (8) (2017) 1421–1438.
[35] Z.Y. Zhang, K.S. Wang, Wind turbine fault detection based on SCADA data analysis using ANN, Advances in Manufacturing 2 (2014) 70–78.
[36] A.P. Marugán, F.P.G. Márquez, J.M.P. Perez, R.H. Diego, A survey of artificial neural network in wind energy systems, Appl. Energy 228 (2018) 1822–1836.
[37] S.L. Yang, W.H. Li, C.L. Wang, The intelligent fault diagnosis of wind turbine gearbox based on artificial neural network, in: Proceedings of the International Conference on Condition Monitoring and Diagnosis, 2008, pp. 1327–1330.
[38] L.W. Zhang, Q. Fan, J. Lin, Z.C. Zhang, X.H. Yan, C. Li, A nearly end-to-end deep learning approach to fault diagnosis of wind turbine gearboxes under nonstationary conditions, Eng. Appl. Artif. Intell. 119 (2023) 105735.
[39] D. Huang, W.A. Zhang, F.H. Guo, W.J. Liu, X.M. Shi, Wavelet packet decomposition-based multiscale CNN for fault diagnosis of wind turbine gearbox, IEEE Trans. Cybern. 53 (1) (2023) 443–453.
[40] Y.Y. Peng, F.Z. Cheng, W. Qiao, L.Y. Qu, Fault prognosis of drivetrain gearbox based on a recurrent neural network, in: Proceedings of the IEEE International Conference on Electro Information Technology (EIT), 2017, pp. 593–599.
[41] L.X. Cao, J.Y. Zhang, J.Y. Wang, Z. Qian, Intelligent fault diagnosis of wind turbine gearbox based on Long short-term memory networks, in: Proceedings of the IEEE 28th International Symposium on Industrial Electronics (ISIE), 2019, pp. 890–895.
[42] A.J. Yin, Y.H. Yan, Z.Y. Zhang, C. Li, R.V. Sánchez, Fault diagnosis of wind turbine gearbox based on the optimized LSTM neural network with cosine loss, Sensors 20 (8) (2020) 2339.
[43] K.W. Liang, N. Qin, D.Q. Huang, Y.Z. Fu, Convolutional recurrent neural network for fault diagnosis of high-speed train bogie, Complexity 2018 (2018) 4501952.
[44] J. Xiao, C.Y. Li, B. Liu, J. Huang, L. Xie, Prediction of wind turbine blade icing fault based on selective deep ensemble model, Knowl. Base Syst. 242 (2022) 108290.

[45] K.N. Ravikumar, A. Yadav, H. Kumar, K.V. Gangadharan, A.V. Narasimhadhan, Gearbox fault diagnosis based on multi-scale deep residual learning and stacked LSTM model, Measurement 186 (2021) 110099.
[46] L. Wei, Z. Cheng, J.S. Cheng, N.Q. Hu, Y. Yang, A fault detection method based on an oil temperature forecasting model using an improved deep deterministic policy gradient algorithm in the helicopter gearbox, Entropy 24 (10) (2022) 1394.
[47] D.T. Hoang, H.J. Kang, A survey on deep learning based bearing fault diagnosis, Neurocomputing 335 (2019) 327–335.
[48] X.J. Guo, L. Chen, C.Q. Shen, Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis, Measurement 93 (2016) 490–502.
[49] J.Y. Jiao, M. Zhao, J. Lin, K.X. Liang, A comprehensive review on convolutional neural network in machine fault diagnosis, Neurocomputing 417 (2020) 36–63.
[50] L. Eren, T. Ince, S. kiranyaz, A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier, Journal of Signal Processing Systems 91 (2019) 179–189.
[51] W. Abed, S. Sharma, R. Sutton, A robust bearing fault detection and diagnosis technique for brushless DC motors under non-stationary operating conditions, Journal of Control, Automation and Electrical Systems 26 (2015) 241–254.
[52] T. Huang, Q. Zhang, X.A. Tang, A novel fault diagnosis method based on CNN and LSTM and its application in fault diagnosis for complex systems, Artif. Intell. Rev. 55 (2022) 1289–1315.
[53] Y.C. Zhu, C.C. Zhu, J.J. Tan, Y.L. Wang, J.Q. Tao, Operational state assessment of wind turbine gearbox based on long short-term memory networks and fuzzy synthesis, Renew. Energy 181 (2022) 1167–1176.
[54] L. Xiang, P. Wang, X. Yang, A.J. Hu, H. Su, Fault detection of wind turbine based on SCADA data analysis using CNN and LSTM with attention mechanism, Measurement 175 (2021) 109094.
[55] R.P. Varshney, D.K. Sharma, Optimizing Time-Series forecasting using stacked deep learning framework with enhanced adaptive moment estimation and error correction, Expert Syst. Appl. 249 (2024) 123487.
[56] S. Ambühl, F. Ferri, J.P. Kofoed, Fatigue reliability and calibration of fatigue design factors of wave energy converters, International Journal of Marine Energy 10 (2015) 17–38.
[57] W.Z. Yang, J.F. Cao, W. Zhang, Study on dynamic simulation of wind turbine gearbox based on Pro/E and ADAMS, Appl. Mech. Mater. 644–650 (2014) 442–445.
[58] S.M. Wang, X.Z. Ai, C. Lv, L.N. Ma, Study on fault diagnosis of gear transmission which is based on ADAMS, Appl. Mech. Mater. 215–216 (2012) 812–816.
[59] T. Aderinto, H. Li, Review on power performance and efficiency of wave energy converters, Energies 12 (22) (2019) 4329.
[60] C.H. Ni, Development of ocean energy test field in China, J. Shipp. Ocean Eng. 6 (2015) 44–49.
[61] Y. Xie, J.S. Zhao, B.H. Qiang, Attention mechanism-based CNN-LSTM model for wind turbine fault prediction using SSN ontology annotation, Wireless Commun. Mobile Comput. 1 (2021) 6627588.
[62] M. Li, B. Bijvoet, K. Wu, X. Jiang, R.R. Negenborn, Optimal chartering decisions for vessel fleet to support offshore wind farm maintenance operations, Ocean Eng. 298 (2024) 117202.
[63] Z. Liu, D. Peng, M.J. Zuo, J. Xia, Y. Qin, Improved Hilbert–Huang transform with soft sifting stopping criterion and its application to fault diagnosis of wheelset bearings, ISA Trans. 125 (2022) 426–444.
[64] Y. Zhang, K. Yu, Z. Lei, J. Ge, Y. Xu, Z. Li, K. Feng, Integrated intelligent fault diagnosis approach of offshore wind turbine bearing based on information stream fusion and semi-supervised learning, Expert Syst. Appl. 232 (2023) 120854.