Journal of *Marine Science and Engineering*

MDPI

Article

Ship Abnormal Behavior Detection Method Based on Optimized GRU Network

Hongdan Liu, Yan Liu, Bing Li * and Zhigang Qi

College of Intelligent Science and Engineering, Harbin Engineering University, Harbin 150001, China; liuhongdan@hrbeu.edu.cn (H.L.); liuyan1011@hrbeu.edu.cn (Y.L.); qizhigang@hrbeu.edu.cn (Z.Q.)
* Correspondence: libing265@hrbeu.edu.cn

Abstract: Ship abnormal behavior detection is an essential part of maritime supervision. It can assist maritime departments to conduct real-time supervision on a certain sea area, avoid ship risks, and improve the efficiency of sea area supervision. Given the problems of complex detection methods, poor detection effectiveness, and low detection accuracy, a Gated Recurrent Unit (GRU) was proposed for ship abnormal behavior detection. Under the premise of introducing the attention mechanism into a GRU, the optimal GRU structure parameters were obtained through the intelligent algorithm to perform deeper feature extraction and train the ship abnormal behavior based on the optimized GRU neural network, so as to realize the detection and recognition of the trajectory data to be measured. Finally, based on the public data set and the trajectory data of the inward and outward ports of ships issued by Nanjing Section, Jiangsu Maritime Bureau, the TensorFlow frame was used to establish an abnormal behavior detection model. The simulation results demonstrated that the abnormal behavior detection model shortened the abnormal detection time. The abnormal behavior detection model used in the detection of ship abnormal behavior enhanced the accuracy and stability of the abnormal behavior identification and verified the validity and superiority of this method.

Keywords: ship abnormal behavior detection; anomaly classification; the optimal GRU; intelligent algorithm

check for updates

Citation: Liu, H.; Liu, Y.; Li, B.; Qi, Z. Ship Abnormal Behavior Detection Method Based on Optimized GRU Network. *J. Mar. Sci. Eng.* **2022**, *10*, 249. https://doi.org/10.3390/jmse10020249

Academic Editor: Md Jahir Rizvi

Received: 9 December 2021
Accepted: 9 February 2022
Published: 12 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the long-term development of the marine and shipbuilding industry, people's maritime activities have penetrated into all aspects of marine production. To ensure the effective conduct of marine production activities, maritime supervision is needed to ensure the safety of maritime activities. In the relevant maritime supervision activities, the main work is to identify the abnormal behavior of the ship. Thus, timely and accurate identification of the abnormal behavior of the ship is a powerful guarantee for the normal navigation of the ship and an effective method to reduce the incidence of maritime and port traffic accidents.

To realize the effective monitoring of the ship's navigation state, the relevant government departments have established a ship traffic service monitoring system in marine environments. However, this monitoring system does not have the ability to identify abnormal behaviors, and the processing of ship trajectory information is manually completed. Therefore, the monitoring and management of ships are prone to incomplete and unmeticulous phenomena [1]. Considering the defects of the traffic service monitoring system, a large number of researchers and engineers have tried to model ship behavior, so as to realize the detection of abnormal behavior. The existing ship anomaly behavior detection methods mainly include clustering analysis, statistical modeling, neural networks, and other anomaly detection methods [2–4]. Ma [5] classified abnormal behavior detection methods according to whether they were data-driven, model-driven, or hybrid-driven [6,7].

On clustering-based anomaly behavior detection methods, Liu et al. [8] proposed an anomaly behavior detection method based on an auxiliary feature vector and a density-based spatial clustering of applications with noise (DBSCAN). Compared with the existing anomaly behavior detection techniques, this method has the ability to accurately identify and distinguish normal data patterns, effective anomalies, and invalid anomalies. Reference [9] proposed a data flow clustering algorithm based on dimension maximum entropy, and realized the anomaly detection of the data flow by comparing the information entropy size and distribution characteristics of micro-clusters. In view of the multi-dimensional characteristics of ship track features, Zhao [10] proposed a ship anomaly behavior detection method and recurrent neural network based on DBSCAN, which can detect the anomalies of the multi-dimensional features of ships. However, there are some problems in density clustering, such as the setting of initial parameters, which greatly affects the accuracy of anomaly detection. Wang et al. [11] improved the DBSCAN algorithm, combined it with the isolated forest algorithm, and selected the internal distribution law of the data set, changing the clustering effect in a large number of unrelated ship distribution data to find abnormalities. Based on the ship anomaly behavior detection method of statistical modeling, Tang [12] proposed a ship anomaly detection method based on the probabilistic digraph model, which has the ability to detect most ship abnormal behaviors. Wang et al. [13] used the electronic chart system to collect data for the repeated abnormal behavior of ships in navigation, and used the time-space analysis method and state determination method to process the collected data, so as to detect the trajectory characteristics of the ship. This method has a large amount of calculation in the data collection and analysis process, and the time complexity is high.

In the method based on neural networks, the recurrent neural network (RNN) takes ship trajectory data as the research object, and proposes an application method of ship anomaly behavior detection using a RNN [14–16]. However, most of the problems considering the selection of training data are studied. Yang et al. [17] proposed a ship abnormal behavior detection method based on bi-long short-term memory (LSTM) for ship abnormal behavior recognition and detection. The results showed that the recognition accuracy of the Bi-LSTM model was higher, and the ship abnormal behavior could be found in time. However, there were problems in the selection of model parameters, and the recognition accuracy needed to be improved. Hu et al. [18] proposed a parallel LSTM-FCN (parallel LSTM-FCN, PLSTM-FCN) model, whose accuracy for multi-dimensional feature ship track prediction is significantly better than that of the traditional recurrent neural network. Liu et al. [19] proposed a new deep learning model called a bidirectional Gated Recurrent Unit (GRU) and attention mechanism (BGRUA), which can accurately identify web services running on HTTPS connections. BGRUA uses a bidirectional gating loop unit (GRU) and attention mechanism to improve the accuracy of HTTPS traffic classification. Wang et al. [20] proposed a Bi-GRU-Caps net model based on the attention mechanism to detect the epistasis between composite entities. The attention mechanism aims to focus on the difference between the two composite entities. The experimental results show that the model has high accuracy in the detection of symptoms and corresponding diseases in English and Chinese corpus, which is superior to the traditional method. References [21–23] use the characteristics of deep learning and time series to propose a ship trajectory prediction method based on recurrent neural networks—the long–short-term memory (RNN-LSTM) model. With the prediction of the correct ship trajectory, the ship abnormal behavior is judged by the threshold, and the LSTM model is superior to the ability to process sequence data, while the selection of the threshold has a great influence on the results.

There are few studies using neural networks for ship anomaly behavior detection in existing methods. Based on the existing methods, this paper proposes a GRU network model with an attention mechanism. This model uses an intelligent algorithm to optimize the structural parameters and convergence speed of the GRU network, and then constructs a ship abnormal behavior detection model based on the optimized GRU network. The model improves the real-time accuracy of ship abnormal behavior detection and solves

the problems of low detection accuracy and long detection time caused by the complex detection methods in the existing methods.

2. Ship Abnormal Behavior Detection

As a crucial part of the Vessel Traffic Service (VTS), ship abnormal behavior detection measures the performance of ship motion data, which is different from the methods and laws of similar organizations. When describing the abnormal behavior of a single ship with AIS data, the most basic abnormality is the abnormal speed, position, and course of the ship, as illustrated in Figure 1. The course of the normal ship is indicated by a yellow arrow, assuming the normal range of ship speed is 8 to 15 kn, where ① is the normal ship position. In ②, the red arrow is outside the boundary, indicating the abnormal position of the ship. In ③, the course of the red-arrow ship differs greatly from that of the normal ship in the model. In ④, ship speed is not within normal range and is regarded as an abnormal speed. This paper mainly studies the abnormal behavior of a single ship.

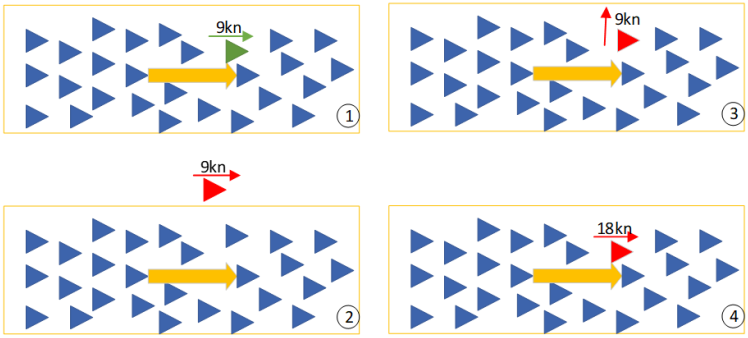


Figure 1. Abnormal trajectory of single ship.

3. Optimize GRU Neural Network

3.1. GRU Network Model

Cyclic neural networks are good at processing time series signals and can accurately obtain the logic between input information. Therefore, it is suitable for the detection of the abnormal behavior of ships. Among them, GRU is an improved variety based on RNN [24], and its internal structure is shown in Figure 2.

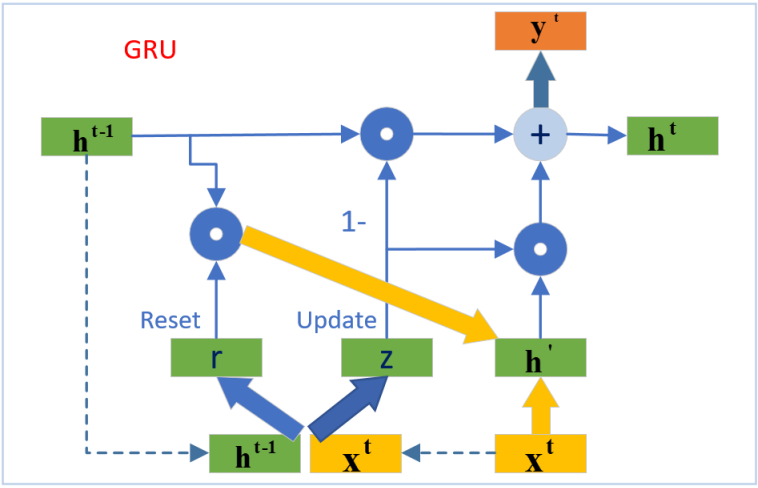


Figure 2. GRU neural network structure.

In Equation (1), [ ] is to concatenate the matrices in square brackets: $[2, 3]$ and $[11]$ are concatenated into $[2, 3, 11]$. $*$ represents the Hadamard product of the matrix, that is, the element-to-bit multiplication. σ is the *sigmoid* function, which scales the data to $(0, 1)$, *tanh*

is the hyperbolic tangent function, which scales the data to $(-1, 1)$. Reset gate: determines the degree of discarding the state information of the previous moment—the smaller the r, the more discarded (the range of r is $(0, 1)$). Update gate: here, $(1 - z)$ constitutes a forgetting gate (refer to LSTM) to selectively forget the previous state information h_{t-1}, selectively remember the current hidden state information $h^{'}$, and finally get h_t. The structure diagram of GRU is shown in Figure 2.

$$\begin{cases} r_t = \sigma([h_{t-1}, x_t] \cdot W_r) \\ z_t = \sigma([h_{t-1}, x_t] \cdot W_z) \\ h'_t = tanh([r_t * h_{t-1}, x_t] \cdot W_{h'}) \\ h_t = (1 - z_t) * h_{t-1} + z_t * h'_t \\ y_t = \sigma(h_t \cdot W_o) \end{cases} \tag{1}$$

where *sigmoid* and *tanh* functions are the following equations:

$$\begin{cases} sigmoid = \frac{1}{1+e^{-x}} \\ tanh = \frac{e^x - e^{-x}}{e^x + e^{-x}} \end{cases} \tag{2}$$

where x is the input variable.

Compared with the basic RNN, the GRU has deeper feature extraction and self-learning ability, and has memory mechanism for long-term state information. However, there are problems in the GRU network, such as unstable training and learning results, the complexity of structural parameters, and a poor optimization of initial weights and feature extraction. Therefore, this paper designs a parameter optimization method for the GRU model. The structure parameters of the GRU neural network are optimized based on an intelligent algorithm to improve the convergence speed and accuracy of the GRU model.

3.2. GRU Neural Network with Attention Mechanism

The inspiration for the attention mechanism comes from the human itself. When processing information, the human brain adjusts the distribution of brain attention according to knowledge and experience to improve efficiency. Therefore, the attention mechanism is an algorithm related to biological attention-allocation mechanisms. Its core idea is to optimize a series of weight parameters, namely to adjust the impact size of each element in the sequence and merge the elements according to the impact size. The weight parameter is similar to the 'attention' allocation coefficient, which determines the 'attention' assigned to each element. Using mathematical principles, the probability distribution of each element is calculated. A series of weight parameters are obtained to reflect the effect of important elements on the output, so as to overcome the relatively poor ability of extracting temporal relations in RNNs [25].

To improve the detection stability of the GRU recurrent neural network, this paper introduces the attention mechanism to calculate the attention probability, so as to highlight the influence of a certain trajectory point on the whole trajectory. For the ship anomaly detection model, adding an attention mechanism increases the association between the front and rear trajectory points. The input of the anomaly detection model is $(x_1, x_2, x_3, \ldots, x_t)$. The hidden layer output of the GRU neural network is $(o_1, o_2, o_3, \ldots, o_t)$. It can be regarded as the feature capture of each track point. The output of $\boldsymbol{x_i}$ is set to be o_i. The attention probability of the attention mechanism is a_i. The formulation:

$$f_i = W_1 tanh(W_2 o_i + b) \tag{3}$$

$$a_i = \frac{\exp(f_i)}{\sum_j^t \exp(f_j)} \tag{4}$$

$$O_i = a_i o_i \tag{5}$$

where W_1 and W_2 are the trainable weight matrix, b is the $\boldsymbol{b_i}$ as vector, f_i is the additive algorithm, and the output after the attention mechanism is set as $O_i = (O_1, O_2, O_3, \ldots, O_t)$.

The process is shown in Figure 3:

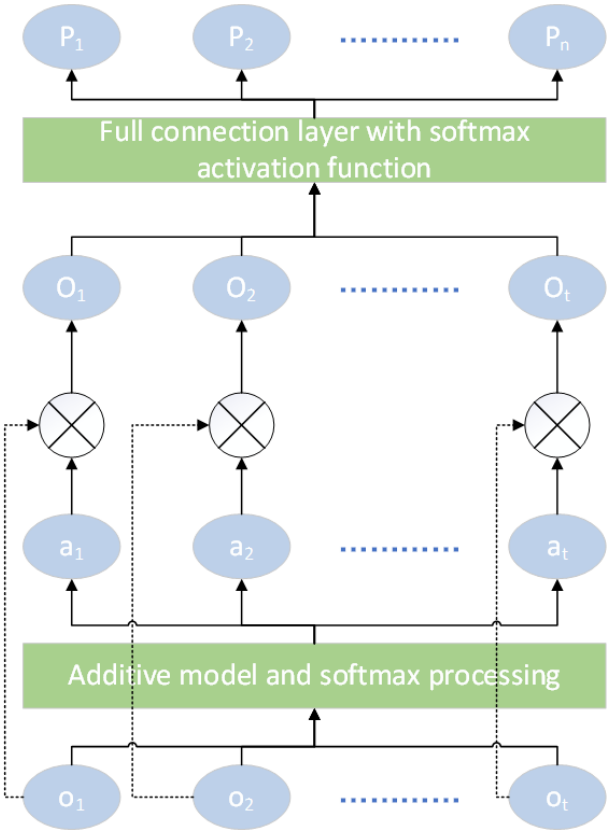


Figure 3. Attention mechanism.

The detection model needs to classify the types of data. Hence, the model output should be a vector with dimension n. Each dimension represents the possibility of each type, and n represents the classified type. $\boldsymbol{Pi} = (P_{i0}, P_{i1}, P_{i2}, \ldots, P_{ic})$, i is the ith trajectory of the ship data, c is the label type. Therefore, this paper needs to add a full-connected layer with a softmax activation function after the attention mechanism. It will be used as the input of the full-connected layer. After the full-connected layer processing, a vector with dimension n (n is the number of pre-set types) is obtained, and each dimension represents the probability of the type represented by this dimension. The type represented by the dimension, where the maximum probability is located, is taken as the final identified label.

3.3. GA Optimizes GRU Network

For the parameter optimization problem of the neural network, the key to the evaluation is whether the optimized design can improve the learning ability of the model. The GA (genetic algorithm) is a mature optimization algorithm [26], which does not rely on the gradient information in the optimization. Therefore, this algorithm is applied to optimize the super parameters of the GRU neural network, including the number of layers and the number of nodes in each layer. The model includes the GRU layer and the full-connection layer to improve the convergence speed and accuracy of the neural network model.

Each chromosome is set to the same length (in this article, we set 2 GRU layers and 2 fully connected layers, plus the first two genes representing the number of layers; thus, each chromosome has $2 + 2 + 2 = 6$ genes), and zeros are added after the length requirement is not met.

In Figure 4, the GA algorithm structure is divided into chromosome structure, GA operation, adaptation function. In optimizing neural networks, traditional algorithms are not easy to implement. The following methods are used to manage the complications:

1. First, set the first two genes. Thus, their ranges are between one and three. Then, determine the following genes related to the number of neurons in each layer according to these two genes;
2. For the modification of the crossover function, first determine the positions that need to be exchanged on the two extracted chromosomes (set as 'a' chromosome and 'b' chromosome), and then traverse the genes of the two chromosomes at these positions. If the gene is 0 or the gene to be swapped is about the number of layers, the swap at this position is canceled;
3. For the modification of the variogram, only the genes related to the number of neurons mutate, and the genes related to the number of layers do not mutate.

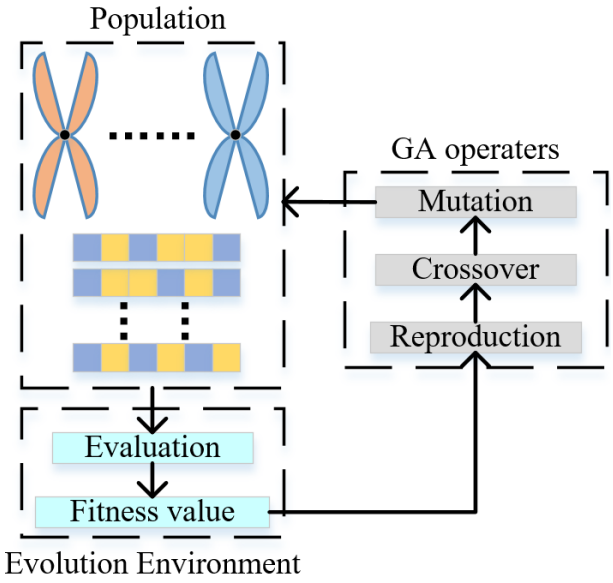


Figure 4. GA algorithm structure.

Parameter settings: population C, selection probability P_s, crossover probability P_c, mutation probability P_m, and iteration number T. We randomly selected and repeatedly randomly sampled C populations from C populations to perform population initialization operations. Here, the roulette method is used for selection. The selected population is formed into the initial population. Then, the crossover operation is performed, and the crossover occurs when the random number is less than the crossover probability. We selected a population from C populations for crossover. Next, the mutation operation is performed, and the mutation operation is only for the last 6-bit parameters. Finally, within the number of iterations, the fitness function of the population is calculated according to Equation (6):

$$\begin{cases} F' = \alpha^* Fit(X) + \beta \\ \alpha = \frac{(C_{\text{mult}} - 1) \cdot Fit_{\text{avg}}}{Fit_{\max} - Fit_{\text{avg}}} \\ \beta = \frac{(Fit_{\max} - C_{\text{mult}} \cdot Fit_{\text{avg}}) \cdot Fit_{\text{avg}}}{Fit_{\max} - Fit_{\text{avg}}} \end{cases} \tag{6}$$

where F' is the fitness function, $Fit(X)$ is the fitness of each chromosome, Fit_{max} is the largest fitness in the chromosome, Fit_{avg} is the average fitness of each chromosome, and C_{mult} is $(1.0, 2.0)$.

In the experiment, when the neural network layer number of the model is four, the output result is the best. According to the characteristics of ship abnormal behavior detection, the number of layers of the neural network is four layers, and the number of

nodes in each layer is between [32, 128]. The genetic algorithm coding in this paper is not a binary coding, but a one-dimensional array composed of various parameters. Each chromosome is a variable of a 10-bit integer. Positions 0 and 1 represent the number of layers, and positions 2–9 represent the number of specific nodes in each layer. The specific implementation steps are described as follows :

1. According to the network layer number limit, the neuron number limit, and the chromosome coding method, the chromosomes with a certain population number are randomly generated. The specific genotype of the chromosome was decoded, and the model was built according to the layer information and node information obtained by decoding. The data were input to train for a certain number of times to obtain the accuracy of the model after training, and the accuracy was used as the individual fitness index;
2. According to the calculated fitness, the selection operator is used to select the individuals that are inherited to the next generation. The crossover operator operation is carried out according to the crossover probability to exchange genes for individuals. The mutation operator is calculated according to the mutation rate, and new individuals are mutated;
3. The fitness of the new population obtained by Step2 is calculated. The various calculation iterations of Step2 are repeated to the optimal individual genotype in the whole process, namely the optimal layer node information. The specific flow chart is shown in Figure 5:

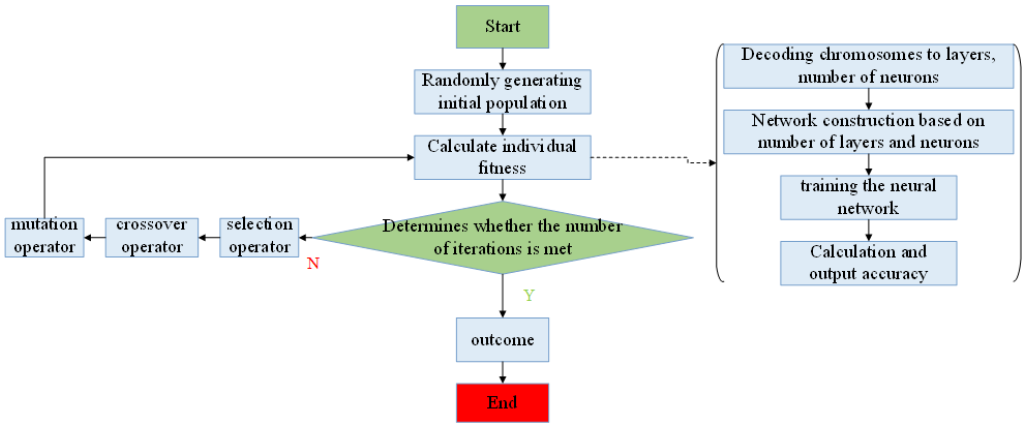


Figure 5. Structure parameter optimization process based on GA.

4. Ship Abnormal Behavior Detection Method Based on GRU Neural Network

4.1. Ship Anomaly Detection Framework Based on GRU

The abnormal behavior of a single ship can be divided into 16 specific behaviors [27]. According to the classification method of ship motion anomalies and ship position anomalies, specific labels should be set for each type. The established anomaly behavior detection model is to train the trajectory data that have been labeled to achieve the detection and identification of the trajectory data that have not been labeled. In this paper, the abnormal behavior of the ship's speed and trajectory point positions are detected. Firstly, the ship anomaly classification model based on the GRU is built. The anomaly classification framework that is built mainly includes three parts: input, hidden, and output. The internal framework of the GRU-based detection model is also three layers, as exhibited in Figure 6:

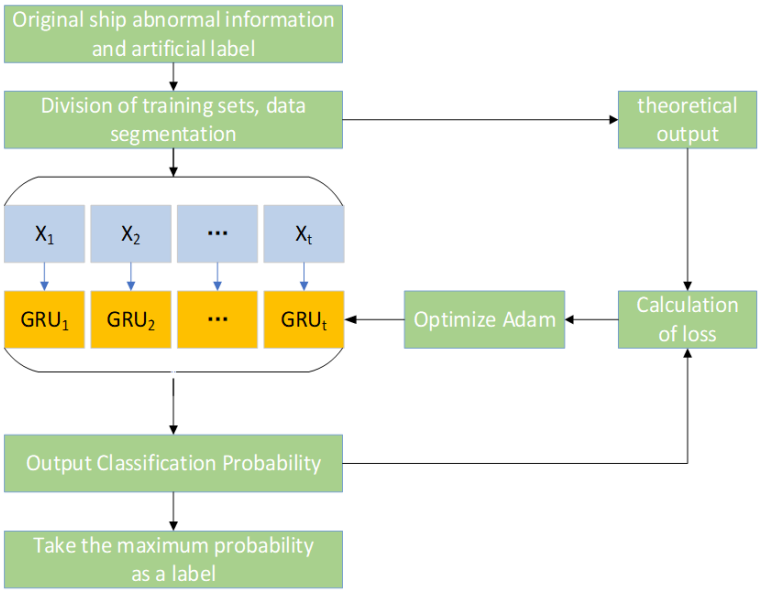


Figure 6. Ship anomaly detection framework based on GRU.

As suggested in Figure 6, the GRU network anomaly recognition framework is divided into input-processing information, the GRU neural network's hidden layer capture characteristics, and output layer label extraction, according to the calculation process. Firstly, the labels of the initial ship anomaly information and artificial tags are input by the input section, and the train set and test set are divided by the input section. The number of training batches is delimited by data segmentation, and the artificial tags are stripped as the theoretical output. The input part outputs a series of multi-dimensional time series signals of length t, with the input information $X = (X_1, X_2, X_3, \ldots, X_t)$. After GRU network processing, a length of the same multi-dimensional time series is output, $o = (o_1, o_2, o_3, \ldots, o_t)$. Through the attention mechanism layer, the multi-dimensional time series with length t is weighted to obtain the output of the same dimension as the number of categories. After the softmax function (such as Equation (7)) of the output layer is processed, the probability of each label of a ship is obtained. $\boldsymbol{P_i} = (P_{i1}, P_{i2}, P_{i3}, \ldots, P_{ic})$, where c is the label type. The label output probability is compared with the theoretical output to calculate the loss. In this paper, we choose cross entropy to calculate the loss [28]. The cross entropy formula is Equation (8):

$$\text{softmax} = \frac{\exp(f_y)}{\sum_{y=1}^{N} \exp(f_y)} \tag{7}$$

where $f_{y_i} = \sum_{i=1}^{N} W_{y_i} \cdot x_i$, W_y is the weight of output. In this paper, the softmax function converts the output class into the form of the probability output, and finally labels the abnormal categories.

$$L = \frac{1}{N} \sum_{i=1}^{N} L_i = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{M} y_{ic} log(P_{ic}) \tag{8}$$

where M is the number of categories, N is the total number of samples, and i is the ith trajectory of the ship data. If the ship has an abnormal behavior, $y_{ic} = 1$, otherwise, $y_{ic} = 0$. P_{ic} means the detection probability of the observed sample i belonging to category c.

4.2. Optimization Model of Ship Anomaly Detection Framework

The loss value is obtained by calculation, and the weight value in the detection framework is optimized by Adam, and the result is obtained by iterating multiple times. Adam is an adaptive learning rate method [29]. It uses the first-order moment estimation and second-order moment estimation of the gradient to dynamically adjust the learning

rate of each parameter. The advantage of Adam is that after bias correction, each iterative learning rate has a certain range, making the parameters more stable.

After completing the GRU-based ship anomaly detection framework, we start to build a GRU parameter optimization framework outside the ship anomaly detection framework. Due to the manual testing of the optimal value of each parameter, not only is the workload is huge but the lack of directionality and reliability are also not strong. Therefore, in order to improve the GRU learning ability, based on the characteristics of optimizing each parameter, two commonly used intelligent algorithms are selected to optimize the parameters of the detection framework. First, the genetic algorithm is used to optimize the number of hidden layers and the number of nodes in each layer. After multiple iterations, the optimal number of layers and nodes are selected to build a GRU-based ship anomaly detection framework.

Among them, the fitness can select the accuracy or loss value output by the test set. The design of the optimization part in Figure 7 can make the ship anomaly detection model automatically complete the selection of the optimal parameters, and improve the detection performance of the detection model.

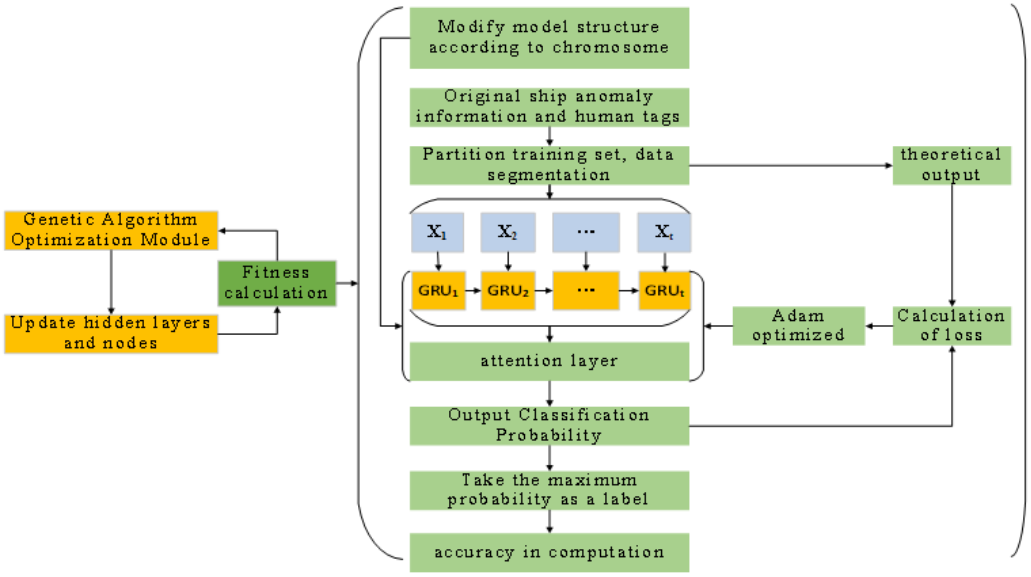


Figure 7. Anomaly detection framework optimization model.

5. Simulation and Analysis of Ship Abnormal Behavior Detection

5.1. Simulation Data Processing

The GRU neural network simulation in this paper is based on the existing advanced neural network framework TensorFlow2.0. The raw ship data is downloaded from www.hifleet.com (accessed on 31 May 2021), mainly for general cargo ships and container ships. In our study, a ship enters (or leaves) once as a trajectory. We selected ship data for approximately 70 vessels during the periods 1 March 2021 and 30 March 2021. We first use the ADTK module in Python to process the ship data, so as to obtain the initial abnormal data of the ship. The ADTK module in Python has the function of labeling data. We use the 'ThresholdAD' algorithm in the ADTK module. This algorithm can detect the anomaly of input data according to the set threshold. The trajectory points above the threshold are marked as anomalies, which can provide convenience for subsequent research. In this study, according to the 'COLREGS—international collision avoidance rule', and the threshold research method adopted by the scholars in reference [30], we analyzed the threshold of this study area. Finally, based on the empirical threshold, the speed threshold is set to 3.589 kn (when the speed of the ship entering and leaving the port exceeds this threshold, it is a speed anomaly), and the course threshold is set to 8.867° (when the angle deviating from the route exceeds this threshold, it is a course anomaly). In the data preprocessing, the ship's longitude, dimension, speed, and heading exceeding the threshold in the detection are

abnormal. Our research area is the inland navigation section of Nanjing Port (32°10.405 N, 118°51.942 E), as shown in Figure 8.

In the data processing stage of the early simulation stage, the data is screened for abnormal classification, and the ship data is summarized and classified. The trajectory data segment with abnormal behavior is intercepted, and labels are preset to obtain the data set of the abnormal behavior trajectory of the ship. The abnormal behavior data accounts for 10% and the overall data is divided into a training set (including 20% in a validation set) and a test set in a ratio of 9:1. The training set is used to train the model, the validation set is used to adjust the parameters, and the test set is used to verify the effects of the model. The input format of the GRU neural network is a multi-dimensional matrix. The construction of the ship trajectory input vector is to convert the ship's AIS trajectory information into a vector form, and each to-be-detected ship trajectory segment is converted into a vector form. Since the ship's AIS trajectory information itself is the time series, the input vector can be designed directly according to the physical meaning. The characteristics of each ship at a certain moment can be embodied by the latitude, longitude, speed, course, and other information in the AIS information. Therefore, the input information x_i can be designed as a four-dimensional direction, including longitude, latitude, speed, and course information, which is $x_i = (Longitude, Latitude, Course, Speed)$.

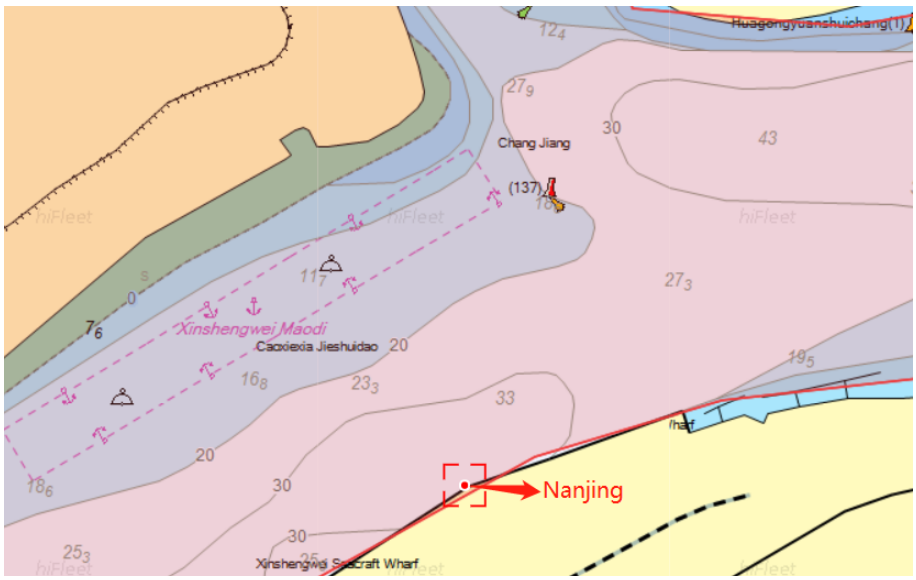


Figure 8. Research on area channel map.

The recurrent neural network has the concept of the number of steps, and the processing information is generally expanded from a vector to a tensor, and the tensor can be regarded as a time series, so this paper combines multiple track points as an input tensor. Based on the duration of abnormal ship behavior, this paper selects the feature states of 20 trajectory points (20 time steps after the ship in the research port and 20 time steps before departure) as the training tensor. In the simulation, the batch size of network training is set to 40, which takes into account the computer operation speed while calibrating the gradient optimization direction of the model. The shape of the multi-dimensional input matrix for this model are in studies [40, 20, 4]. The output is four categories, and the categories are labeled. The normal mark is 0, the position abnormality is marked as 1, the track abnormality is marked as 2, and the speed abnormality is marked as 3.

5.2. Evaluation Index of Anomaly Detection Model

This paper chooses to evaluate the classification effect of ship abnormal behavior detection frameworks by cross entropy. Cross entropy is an important concept in information theory that is mainly used to measure the difference in information between two

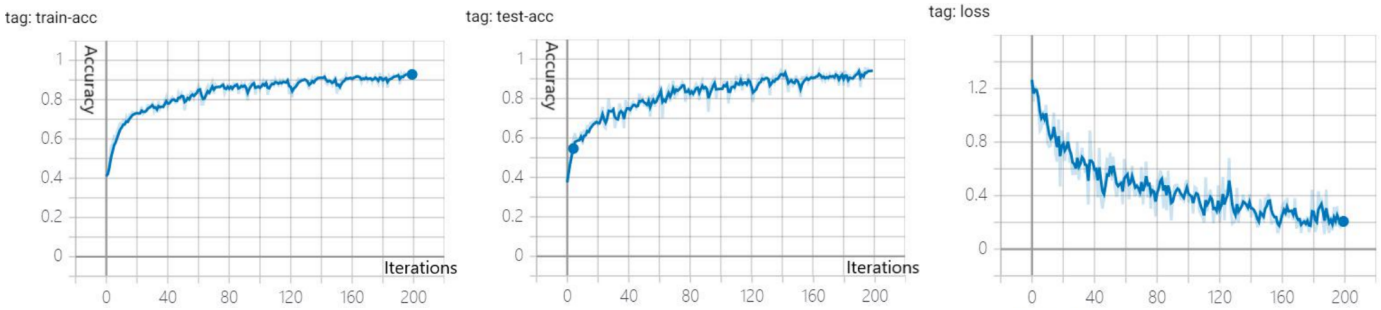
probability distributions. It depicts the distance between the two probability distributions. Cross entropy can measure the similarity between classification markers and prediction markers. The calculation formula is Equation (8). The smaller the cross-entropy error index, the more accurate the classification of the anomaly detection.

Besides using a loss function to represent the model performance, this paper also adds the output accuracy of test-set labels as one of the evaluation criteria of the model. Compared with the accuracy of cross entropy, the physical meaning is clearer and more intuitive. If the correct number of predictions is y_r, the total number of test sets is y_s, and the accuracy is a_c, then the greater the detection accuracy of $a_c = \frac{y_r}{y_s}$, and the better the final effect of the detection model.

5.3. GRU Neural Network Comparison Simulation

The input part of the GRU model is a fully connected layer. The hidden layer is mainly composed of multiple GRUs superimposed. The output part is a fully connected layer whose output is a vector. The output of the model represents the inferred probability of each label. The basic GRU model selects the hidden layer as a two-layer GRU, and sets the random initial weight seed.

After testing, when the number of nodes is greater than 64, the accuracy drops significantly; as such, this study chooses 64 as the number of hidden layer nodes of the neural network and inputs the prepared AIS data set to train the model. When the number of hidden layer nodes is 64, the recorded result of 200 iterations is shown in Figure 9. As the number of iterations increases, the accuracy of the test results gradually increases, and the error loss gradually decreases. The final curve tends to be stable, which is in line with the model's expectations. The final indicator of the detection result of the ship trajectory data of the Jiangsu Nanjing Waterway is connection stable. In this simulation, the average value of the four iterations of the later training curve is selected to represent the final effect. After calculating, the post-average value of the cross entropy of the training set is 0.23, the post-average value of the training set accuracy is 0.905, and the post-average value of the test set accuracy is 0.902, thereby meeting the anomaly detection requirements.



(**a**) (**b**) (**c**)

Figure 9. Initial GRU model simulation graph: (**a**) initial GRU training set accuracy, (**b**) initial GRU test set accuracy, (**c**) initial GRU loss error value.

5.3.1. GRU Simulation Analysis with Attention Mechanism

As shown in Figure 10, the simulation results suggest that the effect of the attention mechanism is not significant when the number of cells is small; when the number of nodes is more than 45, the accuracy of the detection algorithm is significantly higher than that of the algorithm without attention. The above analysis implies that the learning ability of the GRU neural network with the attention mechanism is significantly better than that of the original GRU neural network, and adding the attention layer has little effect on the training time. Therefore, the attention layer can be added to the GRU to improve its learning ability based on the ship trajectory data of the Nanjing Waterway in Jiangsu Province. By introducing the GRU model with the attention mechanism, the training results

of 100 iterations after the anomaly detection simulation are exhibited in Figure 11a–c. The horizontal axis in the figure is the number of iterations.

As observed in Figure 11a–c, with the increase of iteration times, the accuracy of the test results improves, and the error loss degree decreases. The final curve tends to be stable, which is consistent with the expectations of the model. In this simulation, the average of the four iterative results of the training curve in the later period is selected to represent the final index. After calculation, the average value of the cross entropy of the training set is 0.20, the average value of accuracy of the training set is 0.942, and the average value of accuracy of the test set is 0.941. Compared with the original GRU, the accuracy and convergence speed are improved to some extent.

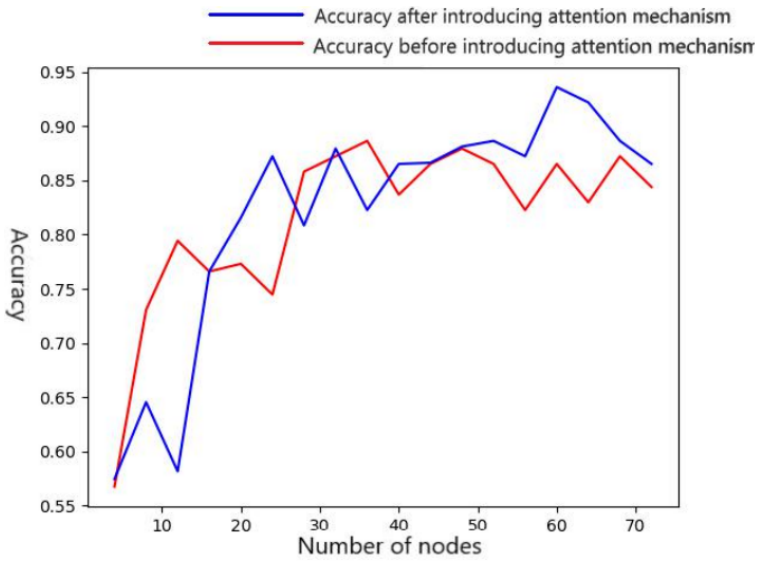
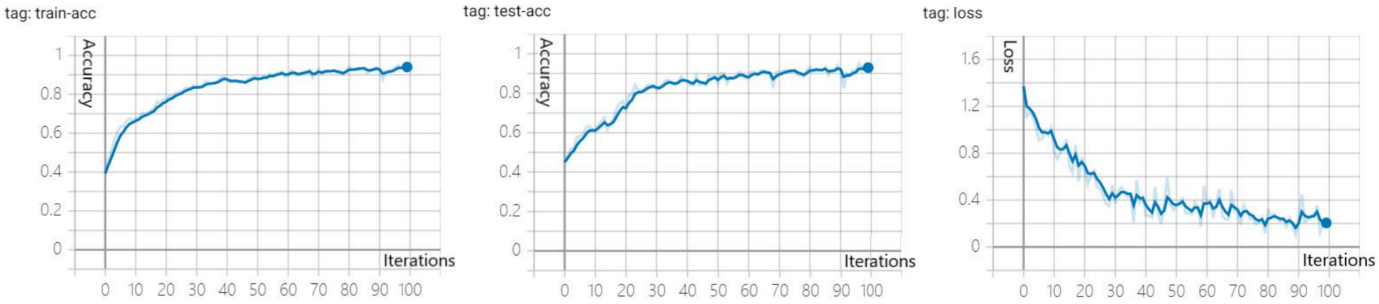


Figure 10. GRU simulation diagram of citing attention.



(**a**) (**b**) (**c**)

Figure 11. Simulation diagram of GRU network with attention mechanism: (**a**) training set accuracy, (**b**) test set accuracy, (**c**) loss error curve.

5.3.2. GA Algorithm Optimizes GRU Simulation

After introducing the attention mechanism, the accuracy of the GRU model is improved, while the convergence speed is too slow. To accelerate the convergence speed, the GA algorithm is used to improve the convergence speed of the detection based on the attention mechanism, and the mechanism parameters of the GRU network are optimized at the same time. In the simulation, the specific parameters of the genetic algorithm were set as follows: the length of each chromosome was set to be 10, the population size to be $C = 20$, the selected probability $P_s = F_i^{'} / \sum_{i=1}^{20} F_i^{'}$ ($F_i^{'}$ from Equation (6)), the crossover rate was set to $P_c = 0.5$, and the mutation rate was set to $P_m = 0.02$. The number of iterations was 40 times. After 40 iterations, the optimal individual is a four-layer GRU, a two-layer fully connected layer, and the number of nodes is 64, 72, 112, 56, 78, and 47, respectively.

The specific training results of the optimized individual are shown in Figure 12a–c, and the horizontal axis in the figure represents the number of iterations. It can be seen from Figure 12a–c that with the increase in the number of iterations, the clarity of the test results gradually increases, and the error loss gradually decreases. The final curve tends to be stable, which is in line with the model's expectations. In this simulation, the average value of the four iterations of the later training curve is selected to represent the final index. After calculation, the post-average cross entropy of the training set is 0.0914, the post-average accuracy of the training set is 0.991, and the post-average accuracy of the test set is 0.989. This comparison reflects that the anomaly detection model after optimizing the structural parameters has the same convergence speed and a highly improved accuracy of abnormal behavior detection.

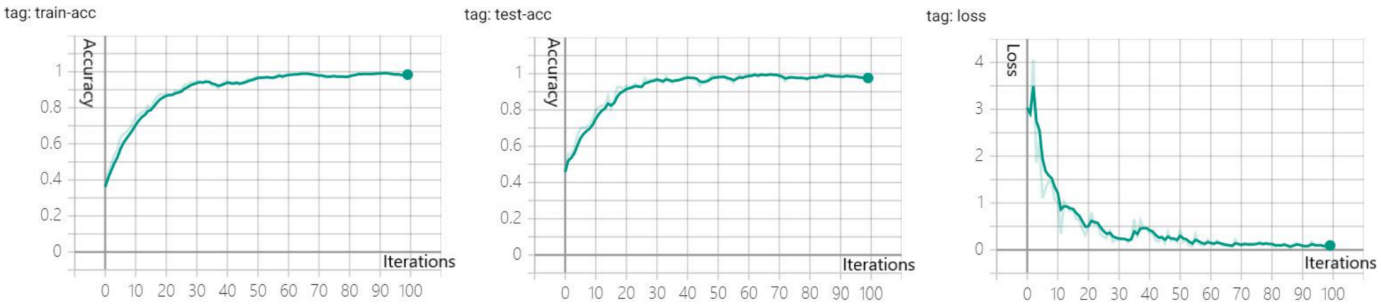


Figure 12. GA algorithm optimizes GRU network.: (**a**) training set accuracy, (**b**) test set accuracy, (**c**) loss error curve.

5.3.3. Comparative Analysis of GRU Anomaly Detection Models before and after Optimization

The comparison simulation curve of the anomaly detection model before and after the GRU optimization is illustrated in Figure 13a–c. The comparison between the training curve that introduces the attention layer and the basic GRU training curve suggests that the GA improves the convergence speed of the abnormal behavior detection framework in the early stage of the iterative optimization, and the curve converges more smoothly in the iterative process. When the curve is finally stable, the final accuracy and loss of the GA-optimized curve are slightly better than the original model.

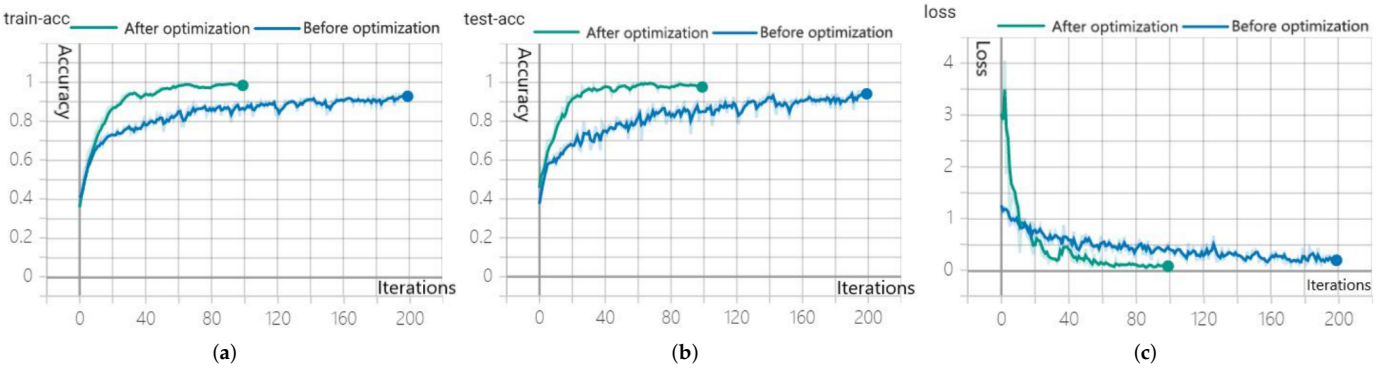


Figure 13. Comparison of simulation diagrams before and after optimization: (**a**) training set accuracy, (**b**) test set accuracy, (**c**) loss error curve.

Comparing the model training curves before and after the GA optimization of the GRU structure parameters, it can be observed that the curve before optimization requires 200 iterations to converge to the optimal value. After using the genetic algorithm to optimize the model hyperparameters, the model only needs about 60 iterations to reach a stable state.

The convergence ability of the anomaly detection model is further improved. When the curve becomes stable, the accuracy and loss are better than the original model. Besides this, the learning ability of the detection framework has been significantly optimized.

5.4. Optimized GRU-Based Simulation and Analysis of Ship Abnormal Behavior Detection

This paper uses the optimized GRU model to detect and identify the anomalies of ships entering and leaving the port. The data source is the Nanjing Regional Channel of Jiangsu Maritime between 1 March 2021 and 30 March 2021. Test data were 270 tracks, of which 10% were abnormal, and identifications of specific detection results are shown in Table 1:

Table 1. Ship anomaly detection results.

Number	Ship's Name	True Anomaly	Detect Abnormal Tags
1	AN PING8268	Abnormal position	1
2	JUN YAN169	Abnormal track	2
3	LU TONG0999	Abnormal speed	3
…	…	…	…
270	YONGAN996	Normal track	0

In the end, the number of correct detections was 266, and the correct rate was 96.58%. However, the detection results in the table can only reflect the maximum possible label output by the model, and cannot intuitively reflect the likelihood that the abnormal ship belongs to this label or the degree of differentiation from other labels. Therefore, this paper also draws the label probability output by the model on the three-dimensional image to intuitively reflect the classification of the three abnormal behaviors. In this paper, cross entropy is selected as the loss degree of the weight optimization of the detection model, that is, the final output of the model can be regarded as a vector representing the detection possibilities of various behaviors, and the maximum probability type is taken as the final label.

In this paper, normal data and three types of anomalies are selected for identification. The output is a four-dimensional space vector, so as to make the output of the model more intuitive, and we chose three types of anomalies to draw. Then, the vectors representing the differences among the three anomaly types are three-dimensional, and each dimension represents the possibility that a certain trajectory is identified as a certain anomaly. The exception outputs of the original anomaly detection framework and the final anomaly detection framework are presented in Figure 14a,b:

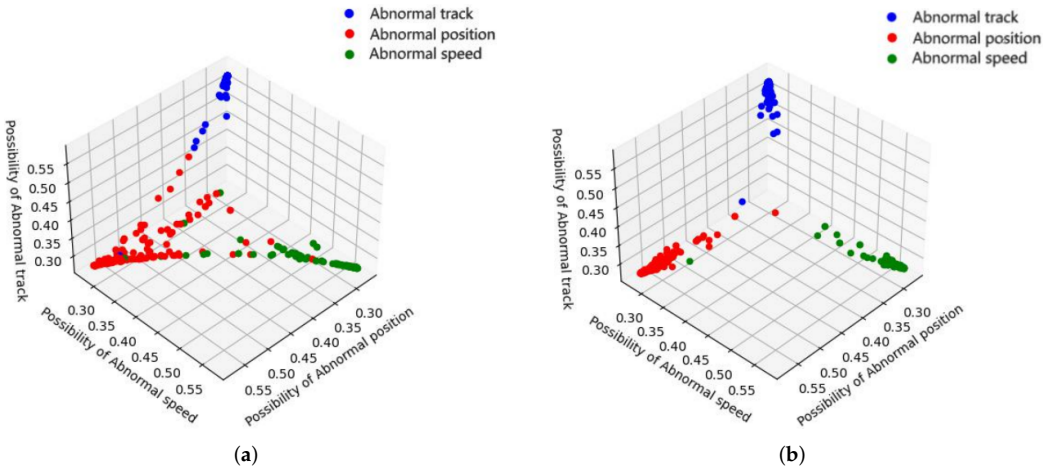


Figure 14. Output results of detection framework before and after optimization: (**a**) output result before optimization, (**b**) output result after optimization.

From the output, three kinds of anomalous trajectories should be clearly divided in the possibility space. Meanwhile, the closer each kind of anomalous trajectory is to the standard axis, the higher the value on the nearest coordinate axis, and the better the classification effect. In Figure 14a, the division of abnormal trajectories is not clear; there are a large number of trajectories close to other abnormal-possibility axes and origins (normal trajectories). Therefore, it is easy to increase the false detection rate. The detection results of the anomaly detection framework after parameter optimization are shown in Figure 14b, and the spatial distribution of abnormal trajectories is more clear. Only a few trajectories are identified as other anomalies because they are close to other classification axes, or are identified as normal trajectories because they are close to the origin.

6. Conclusions

Ship abnormal behavior detection technology has an extremely high application value in the field of maritime situation awareness. In this regard, this paper proposes a method for detecting abnormal ship behavior based on the improved GRU neural network. This method introduces intelligent algorithms to optimize the structural parameters of the GRU neural network, contributing to the improvement of the detection ability of the GRU model. Besides this, the optimized GRU model is employed for the detection of abnormal ship behavior. The training set and test set of the ship trajectory are established. With the TensorFlow framework, the detection of ship abnormal behaviors such as abnormal ship position and speed is completed. The comparative analysis of the simulation results based on the evaluation index of cross entropy verifies that the method effectively shortens the detection time and improves the detection performance. Simulation experiment results demonstrate that the method has a certain degree of practicability and innovation.

Author Contributions: Conceptualization, Y.L. and H.L.; methodology, Y.L. and H.L.; software, Y.L. and B.L.; validation, Y.L. and H.L.; formal analysis, Z.Q.; writing—original draft preparation, Y.L. and H.L.; writing—review and editing, Y.L., H.L. and B.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Natural Science Foundation of Heilongjiang Province (grant number LH2019E035) and the Natural Science Foundation of Heilongjiang Province (grant number LH2021F017).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study.

References

1. Sun, J. On the supervision of maritime law enforcement in China. *Adm. Law* **2019**, *12*, 92–98.
2. Rong, H.; Teixeira, A.P.; Soares, C.G. Data mining approach to shipping route characterization and anomaly detection based on AIS data. *Ocean. Eng.* **2020**, *198*, 106936. [CrossRef]
3. Filipiak, D.; Strózyna, M.; Wecel, K.; Abramowicz, W. Anomaly Detection in the Maritime Domain: Comparison of Traditional and Big Data Approach. In Proceedings of the NATO IST-160- RSM Specialists' Meeting on Big Data & Artificial Intelligence for Military Decision Making, Bordeaux, France, 30 May–1 June 2018; pp. 143–159.
4. Zhang, B.; Ren, H.; Wang, P.; Wang, D. Research Progress on Ship Anomaly Detection Based on Big Data. In Proceedings of the 2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 16–18 October 2020.
5. Ma, W.Y.; Wu, Z.L.; Li, W.F. Consistency detection algorithm for abnormal ship behavior. *J. Traffic Transp. Eng.* **2017**, *17*, 149–158.
6. Xia, L.; Li, Z. A new method of abnormal behavior detection using LSTM network with temporal attention mechanism. *J. Super Comput.* **2020**, *77*, 3223–3241. [CrossRef]
7. Tang, H.; Yin, Y.; Shen, H. Summary of the detection of abnormal behavior of sea ships. *J. Chongqing Jiaotong Univ. Sci. Ed.* **2019**, *38*, 109–115.

8. Liu, H.; Wang, Y.; Chen, W.G. Anomaly detection for condition monitoring data using auxiliary feature vector and density-based clustering. *IET Gener. Transm. Distrib.* **2020**, *14*, 108–118. [CrossRef]
9. Geng, Z.Q.; Ji, W.; Han, Y.M.; Cao, J. Anomaly detection method based on dimensional maximum entropy data stream clustering. *Control. Decis.* **2016**, *31*, 343–348.
10. Zhao, L.; Shi, G. Maritime anomaly detection using density-based clustering and recurrent neural network. *J. Navig.* **2019**, *72*, 894–916. [CrossRef]
11. Wang, Y.T. *Ship Trajectory Prediction Based on CNN and LSTM*; Dalian Maritime University: Dalian, China, 2020; pp. 36–42.
12. Tang, H. *Intelligent Analysis of Behavior Characteristics of Ships Entering and Leaving the Port*; Dalian Maritime University: Dalian, China, 2020; pp. 89–93.
13. Wang, Z.R.; Zhao, K.Y.; Cai, C.; Ding, M.Z.; Wang, P. Analysis of abnormal ship behavior based on DBSCAN and iForest algorithm. *Ship Electron. Eng.* **2021**, *41*, 89–94.
14. Daranda, A. Neural Network Approach to Predict Marine Traffic. *Trans. Balt. J. Mod. Comput.* **2016**, *4*, 483–495.
15. Zhang, L.; Suzuki, H.; Koyama, A. Recognition of Meal Information Using Recurrent Neural Network and Gated Recurrent Unit. *Internet Things* **2021**, *13*, 100–106. [CrossRef]
16. Burnaev, E.; Koptelov, I.; Novikov, G.; Khanipov, T. Automatic construction of a recurrent neural network based classifier for vehicle passage detection. In Proceedings of the SPIE 10341, Ninth International Conference on Machine Vision 2016, Nice, France, 18–20 November 2016.
17. Yang, F.; He, Z.; He, F. Detection method of abnormal ship behavior based on LSTM neural network. *J. Wuhan Univ. Technol. (Transp. Sci. Eng. Ed.)* **2019**, *43*, 886–892.
18. Hu, D.; Meng, X.; Lu, S.; Xing, L.N. Application of a Parallel LSTM-FCN Model in Ship Track Prediction. *Control Decis.* **2021**, *21*, 1–7.
19. Liu, X.; You, J.; Wu, Y.; Li, T.; Li, L.; Zhang, Z.; Ge, J. Attention-Based Bidirectional GRU Networks for Efficient HTTPS Traffic Classification. *Inf. Sci.* **2020**, *541*, 297–315. [CrossRef]
20. Wang, Q.; Xu, C.; Zhou, Y.; Ruan, T.; Gao, D.; He, P. An Attention-based BI-GRU-CapsNet Model for Hypernymy Detection between Compound Entities. In Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Madrid, Spain, 3–6 December 2018; pp. 1031–1035.
21. Ren, Y.X.; Zhao, J.S.; Liu, W.; Wang, S.Z.; Wei, Y.H. Dynamic prediction of ship navigation based on AIS data and LSTM network. *J. Shanghai Marit. Univ.* **2019**, *40*, 32–37.
22. Liang, Y.; Zhang, H. Ship Track Prediction Based on AIS Data and PSO Optimized LSTM Network. *Int. Core J. Eng.* **2020**, *6*, 23–33.
23. Tang, H.; Yin, Y.; Shen, H. A model for vessel trajectory prediction based on long short-term memory neural network. *J. Mar. Eng. Technol.* **2019**, *3*, 1–10. [CrossRef]
24. Xie, X.; Wang, B.; Wan, T.; Tang, W. Multivariate Abnormal Detection for Industrial Control Systems Using 1D CNN and GRU. *IEEE Access* **2020**, *8*, 88348–88359. [CrossRef]
25. Xiong, L.; Zhang, L.; Huang, X.; Yang, X.; Huang, W.; Zeng, H.; Tang, H. DCAST: A Spatiotemporal Model with DenseNet and GRU Based on Attention Mechanism. *Math. Probl. Eng.* **2021**, *2021*, 8867776. [CrossRef]
26. Town, J.S.; Gao, Y.; Hancox, E.; Liarou, E.; Shegiwal, A.; Atkins, C.J.; Haddleton, D. Automatic peak assignment and visualisation of copolymer mass spectrometry data using the 'genetic algorithm'. *Rapid Commun. Mass Spectrom.* **2020**, *34*, e8654. [CrossRef]
27. Jiang, B.C.; Guang, J.; Zhou, W.; He, Y. Mining and Recognition Analysis of Abnormal Ship Behavior in Maritime Traffic. *Comput. Simul.* **2017**, *34*, 329–334.
28. Zhou, F.; Li, Y.; Fan, X.Y. Improved feedback loss calculation method of image classification convolutional neural network. *Small Microcomput. Syst.* **2019**, *40*, 1532–1537.
29. Zhang, Z. Improved Adam Optimizer for Deep Neural Networks. In Proceedings of the 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS), Banff, AB, Canada, 4–6 June 2018.
30. Peng, X.W. *Research on Spark-Based Real-Time Anomaly Detection Method of Inland Ship Trajectory*; Wuhan University of Technology: Wuhan, China, 2017; pp. 67–69.