

Research Article

Risk Prediction for Ship Encounter Situation Awareness Using Long Short-Term Memory Based Deep Learning on Intership Behaviors

Jie Ma,^{1,2,3} Wenkai Li,¹ Chengfeng Jia ¹, Chunwei Zhang,¹ and Yu Zhang⁴

¹School of Navigation, Wuhan University of Technology, Wuhan 430063, China

²Hubei Inland Shipping Technology Key Laboratory, Wuhan 430063, China

³National Engineering Research Center for Water Transportation Safety, Wuhan 430063, China

⁴School of Logistics Engineering, Wuhan University of Technology, Wuhan 430063, China

Correspondence should be addressed to Chengfeng Jia; jcf@whut.edu.cn

Received 3 September 2020; Revised 20 October 2020; Accepted 9 November 2020; Published 29 November 2020

Academic Editor: Petr Dolezel

Copyright © 2020 Jie Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Encounter risk prediction is critical for safe ship navigation, especially in congested waters, where ships sail very near to each other during various encounter situations. Prior studies on the risk of ship collisions were unable to address the uncertainty of the encounter process when ignoring the complex motions constituting the dynamic ship encounter behavior, which may seriously affect the risk prediction performance. To fill this gap, a novel AIS data-driven approach is proposed for ship encounter risk prediction by modeling intership behavior patterns. In particular, multidimensional features of intership behaviors are extracted from the AIS trace data to capture spatial dependencies between encountering ships. Then, the challenging task of risk prediction is to discover the complex and uncertain relationship between intership behaviors and future collision risk. To address this issue, we propose a deep learning framework. To represent the temporal dynamics of the encounter process, we use the sliding window technique to generate the sequences of behavioral features. The collision risk level at a future time is taken as the class label of the sequence. Then, the long short-term memory network, which has a strong ability to model temporal dependency and complex patterns, is extended to establish the relationship. The benefit of our approach is that it transforms the complex problem for risk prediction into a time series classification task, which makes collision risk prediction reliable and easier to implement. Experiments were conducted on a set of naturalistic data from various encounter scenarios in the South Channel of the Yangtze River Estuary. The results show that the proposed data-driven approach can predict future collision risk with high accuracy and efficiency. The approach is expected to be applied for the early prediction of encountering ships and as decision support to improve navigation safety.

1. Introduction

Water traffic has become increasingly busy with the rapid development of the shipping industry in recent years, which has led to an increased risk to individuals and society in terms of various aspects, especially ship-ship collision accidents. Owing to the frequent occurrences and serious consequences of collisions, research on reducing collision accidents from both theoretical and practical points of view has always been a major topic of concern for navigational experts and scholars. Perceiving risk and predicting encounter situations between ships are crucial for the

prevention of collision accidents, especially in busy traffic areas, where congested ships sail relatively close to each other [1].

To understand the risk level and take actions to decrease the possibility of collisions occurring in the waters, numerous efforts have been devoted to the risk analysis and assessment of ship collisions. Some focus on risk surveys among maritime experts and the conduct of qualitative collision risk analyses, primarily through empirical studies. Cohen et al. [2] used highly stressful training scenarios generated by a ship simulator to measure the heart rates of participants to estimate the collision risk. Chin and Debnath

[3] examined the risks of different ship types by developing a survey conducted by Singapore port pilots under both day and night conditions. However, the above qualitative methods do not take into account the ship navigation data, so it is difficult for them to reflect highly dynamic and continuous vessel movement as well as the evolutionary collision risk trends.

In recent years, with the wide application of automatic identification systems (AIS) in water traffic control and surveillance, AIS data have been proven to be a valuable source of ship behavior monitoring and analysis [4–6]. The AIS can transmit motion information (e.g., speed, course, etc.) between ships, from ships to shore, or vice versa. This makes it possible to quantitatively analyze collision risk by means of massive AIS data. Silveira et al. [7] used AIS data to model traffic patterns off the coast of Portugal, based on which the probability of a ship collision occurring was calculated. A related method was adopted by Christian and Kang [8] to develop a probabilistic risk assessment. In addition, the motion data obtained from AIS can be used to calculate the distance to the closest point of approach (DCPA) and time to closest point of approach (TCPA), which can quantify the collision risk from spatial and temporal aspects, respectively. Ahn et al. [9] defined the membership functions of DCPA and TCPA based on the simulation results. Collision-avoidance maneuvers were then obtained using multilayer perceptron neural networks. Similarly, Hwang et al. [10] designed a fuzzy collision-avoidance expert system, where the DCPA and TCPA were considered simultaneously. With the help of their system, ships can be advised to make proper maneuvers to avoid collisions at the right time.

However, as navigational experts and some studies have found, the DCPA and TCPA do not fully reflect the actual collision risk level, and using only these two parameters may lead to misjudgments regarding the collision risk [11, 12]. Therefore, modeling the collision risk using multiple parameters has gradually been adopted by most researchers. Ren et al. [13] presented a linear model for evaluating ship collisions, which considered several factors, such as the ship type, velocity, and route. Silveira et al. [14] estimated the distances between ships by using sampled positions, courses, and speeds, based on which the number of varying collision candidates was evaluated by comparing it with a predefined collision diameter. Zhang et al. [15] developed a vessel conflict ranking operator (VCRO) model, which considered the relative ship speeds, the course difference, and distance between two ships. Then, the Northern Baltic Sea AIS data were used to assess the risk of a near-miss collision, and the results indicate that the model is adequate for ranking the encounters. Based on the VCRO model, Zhang et al. [16] combined the density complexity of open waters with the multivessel VCRO model to assess the regional near-miss collision risk.

The ship domain is supposed to be a feasible metric to make collision risk predictions based on the assumption that the risk of collision is high when the ship's domain is invaded by the target ship. Szlapczynski [17] proposed a novel method to measure collision risk by adopting the concept of

an ellipse-shaped ship domain. Further, Szlapczynski and Szlapczynska [18] addressed the domain violation problem by combining two parameters, that is, the degree and time of domain violation, to offer an intuitive assessment of the collision risk. Wu et al. [19] also employed the ship domain violation rule to study the frequency of ship conflicts, which considered elliptical and circular domains individually, and a series of hot spots with high collision risk in the Sabine-Neches Waterway were identified by using the two domain types. Wang [20] proposed a novel ship domain model termed the fuzzy quaternion ship domain (FQSD). The domain sizes are determined by the quaternion, including the forward, aft, starboard side, and port side radius. The FQSD model uses fuzzy boundaries (e.g., the ship boundary could be linear or nonlinear as well as thin or fat) to estimate the collision risk, aiming at providing a reasonable and dependable evaluation method. By taking advantage of the FQSD model, Qu et al. [21] estimated the number of ship domain overlaps to evaluate the collision risk in the Singapore Strait, assuming that increased ship domain overlap indicates a higher ship collision probability. However, the ship domain uncertainty will seriously affect prediction performance. Several measures, such as the length and width of the encounter ship, which should be known when calculating the ship domain are not always available. In addition, most collision risk prediction approaches based on ship domains assume that the speed and course of the ship are constant at the moment of sampling [22], which does not sufficiently take into account the evolutionary factors of an encounter process affecting the risk.

The primary limitation of prior studies on the risk of ship collisions is that they cannot address the uncertainty of the encounter process when neglecting the complex motions constituting the dynamic behavior of encountering ships. Thus, it is necessary to incorporate the spatiotemporal behaviors of ships encountering each other (intership behavior) to make the risk prediction more reasonable because the intership behavior will determine the subsequent risk state to a certain extent with the evolutionary process of the ship encounter but was rarely considered and implemented in previous research. To bridge this gap, we propose a novel AIS data-driven approach for ship encounter risk prediction by modeling intership behavior patterns. The primary contributions of this study are summarized as follows:

- (i) Intership behavior is essentially a stochastic process consisting of the motion behaviors of any encountered ships. Following this rationale, this paper proposes modeling intership behavior by transforming the AIS traces into a sequence of behavioral features by combining a fixed set of parameters, including the relative velocity, course difference, and relative distance as well as three azimuthal types. With this time series structure, the process of ships encountering each other and the corresponding spatiotemporal dynamics can be effectively characterized.
- (ii) As previously discussed, ship collisions are often closely related to the navigator's behavior. Thus, a

novel method to model the relationship between intership behavior and collision risk involvement is necessary to accurately predict the risk. To address this challenge, we relate the sequence of behavioral features involving a specified time window to the future risk level. Then, the mapping between them can be established through a supervised learning approach, and the problem of risk prediction is formed as a time series classification task [23], which makes the prediction process easier to implement by taking full advantage of the benefits of data-driven modeling with AIS. Inspired by the recent achievements of long short-term memory (LSTM) networks for various time series learning tasks such as text categorization [24, 25] and trajectory prediction [26, 27], we extend them to our mapping modeling between the intership behavior and the collision risk. To the best of our knowledge, we are the first to address this issue through the utilization of LSTM networks.

- (iii) With our proposed approach, the potential collision risk associated with the uncertain encounter process could be recognized and identified at an early stage. Thus, early warnings can be provided so that ship officers have sufficient time to react to emergencies and take evasive actions in advance. Additionally, the outcome of this research can provide useful support to human operators in charge of large and crowded water areas and encourage safe navigation under specific scenarios to reduce the incidence of ship collisions.

The remainder of this paper is organized as follows: First, we provide a brief description of the issue of risk prediction in Section 2. Next, Section 3 develops a thorough discussion regarding the extraction of AIS data as well as constructing the sequence of behavioral features. The ship encounter risk prediction frameworks are proposed in Section 4. Finally, Section 5 is dedicated to a summary of our numerical results and a discussion of the model's performance.

2. Problem Formulation

The goal of designing the methodological framework is to investigate the key issues (e.g., ship encounter risk prediction) affecting the intership behavior. The collision risk level of the encountering ships at time t is represented by R_t . R_t is divided into five categories according to the risk level from low to high, and class labels of 1, 2, 3, 4, and 5 represent the following:

$$\text{CRI} = \begin{cases} 1, & \text{low risk level,} \\ 2, & \text{low - middle risk level,} \\ 3, & \text{middle risk level,} \\ 4, & \text{middle - high risk level,} \\ 5, & \text{high risk level.} \end{cases} \quad (1)$$

We denote these risk levels as follows:

- (i) Low risk level: A situation where risk begins to be present and two ships are free to maneuver.
- (ii) Low-middle risk level: A situation in which the ships approaching each other have a collision risk and the given-way ship should maneuver in advance.
- (iii) Middle risk level: A situation in which a safe passing distance cannot be ensured if only the given-way ship fully maneuvers.
- (iv) Middle-high risk level: A situation in which collision cannot be avoided if only the given-way ship fully maneuvers.
- (v) High risk level: A situation in which two ships should fully maneuver to avoid the collision.

In this study, the collision risk can be defined as a continuum spectrum of colors, as shown in Figure 1. This spectrum ranges from the safest situation (a near-zero chance of collision) to the riskiest situation during encounters (both ships need to take evasive actions to avoid collision). A collision risk index (CRI) [28] was employed to calculate the risk spectrum. In terms of collision avoidance, the CRI is essential for a ship officer to evaluate the risk of a ship encounter as well as for performing an evasion strategy [29].

As previously mentioned, the collision risk could be affected by the uncertain and complex behavior of encountering ships. A ship encounter is essentially a dynamic evolutionary process commonly utilized to perceive the situation of encountering ships. The evolution of the encounter process is subjected to the specific motions of each ship as well as pairwise exchanges of influences between the ships, thus indicating that the spatiotemporal kinematics of the ships involved in the encounter have dependency and correlation. To associate the collision risk prediction with the evolution of the encounter process, we aim to model the relationship between the sequence of behavioral features and the future risk level. For one encounter pair, we denote the behavioral features as follows:

$$U = [u^1, \dots, u^i, \dots, u^N], \quad (2)$$

$$u^i = [V_R^i, A^i, D_{ot}^i, \alpha^i, \alpha_o^i, \alpha_t^i]^T,$$

where U is an $N \times 6$ -dimensional variable composed of u^i , where N represents N sampling points. V_R^i, A^i, D_{ot}^i are relative velocity, course difference, and relative distance between two ships, respectively. $\alpha^i, \alpha_o^i, \alpha_t^i$ are three types of azimuths (the details are introduced in Section 3). If the time window is 2δ and the sliding step is δ , then the entire track U can be divided into L time windows. Therefore, an encounter process consisting of a sequence of behavioral features can be reformulated as follows:

$$w^t = [u^{t-2\delta+1}, u^{t-2\delta+2}, \dots, u^t], \quad (3)$$

where w^t represents the observation window with length of 2δ before time t . The goal of this study is to predict the risk level of a ship at a future time ΔT , so we need to match w^t with $R^{t+\Delta T}$ and generate the pairs of sample datasets

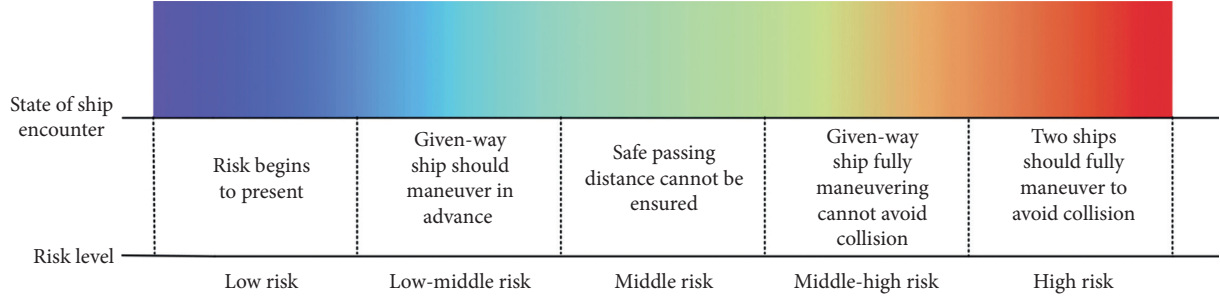


FIGURE 1: Spectrum of the collision risk level.

$(w^t, R^{t+\Delta T})$. We want to find a function f that can best model the relationship between w^t and $R^{t+\Delta T}$:

$$f: w^t \longrightarrow R^{t+\Delta T}. \quad (4)$$

By means of equation (5), the issue of risk prediction can be transformed into a time series classification task. To evaluate the model prediction, a confusion matrix was designed to assess the predictability. The size of the square matrix represents the categories of various risk levels. Table 1 shows the confusion matrix with five risk levels. As presented in Table 1, each diagonal element of the confusion matrix represents the correct category; for example, T_L is the proportion of low risk level that is correctly predicted. FM_L represents False Middle give Low, which is the proportion of low risk level that is wrongly predicted as middle risk level. In addition, the misclassification error rate (MER) is employed to estimate the overall performance of the model. The MER can be obtained by comparing the predicted risk level with the actual risk level as follows:

$$MER = \frac{1}{N_R} \sum_1^{N_R} (\hat{R}_s \neq R_s), \quad (5)$$

where N_R is the number of the windows. Furthermore, a tenfold cross-validation method is utilized to obtain the best model, which has been empirically shown to yield estimates that suffer neither from overly high bias nor from excessively high variance [30].

3. Data Preparation and Feature Extraction

In this section, we describe the process of extracting the behavior features from the original AIS trace data, which can effectively characterize the navigation activities and corresponding spatiotemporal dynamics. The process comprises two components. First, we clean and integrate the enormous volume of original AIS data, by which the AIS data will be purified and selected into a time series structure. Then, through the space-time registration, the synchronous pairwise trajectory of encountering ships can be obtained. Next, we transform the pairwise trajectory data into a sequence of behavioral features by combining a fixed set of parameters.

3.1. Data Preparation. The AIS is an automatic tracking system to improve navigation safety and avoid collision accidents by providing the navigation information of various

TABLE 1: Confusion matrix for prediction result of risk level.

Risk level	L	LM	M	MH	H
L	T_L	FL_{LM}	FL_M	FL_{MH}	FL_H
LM	FLM_L	T_{LM}	FLM_M	FLM_{MH}	FLM_H
M	FM_L	FM_{LM}	T_M	FM_{MH}	FM_H
MH	FMH_L	FMH_{LM}	FMH_M	T_{MH}	FMH_H
H	FH_L	FH_{LM}	FH_M	FH_{MH}	T_H

ships. In general, this navigation information in AIS messages is broadly classified as either dynamic information or static information. The dynamic information includes the ship location (longitude and latitude), speed over ground (SOG), course over ground (COG), destination, and estimated arrival time. The static information contains the ship name, ship maritime mobile service identity (MMSI), ship type, ship size, current time, and other information. As the AIS data contains the above information, it can serve as the data source for understanding the traffic situations [31]. In particular, SOG and COG have substantial impacts on dangerous encounter situations. Many existing studies take SOG and COG into consideration in ship collision risk assessment [32–34]. However, there are some errors in the AIS data, such as messy codes and data irrationalities, which may contribute to misjudgments of collision accidents. Therefore, certain preprocessing methods are essential to ensure the reliability and applicability of the AIS data to gain a better investigation of the collision risk.

3.1.1. Data Cleaning and Trajectory Interpolation. This part aims to eliminate the above-mentioned errors in the AIS data. A mathematical data cleaning method is used. We denote a trajectory Traj as follows:

$$\text{Traj} = [\text{traj}^1, \dots, \text{traj}^m, \dots, \text{traj}^M], \quad (6)$$

$$\text{traj}^m = [\text{lng}^m, \text{lat}^m, \text{sog}^m, \text{cog}^m]^T,$$

where M is the number of trajectory sampling points. traj^m denotes the four-dimensional vector of the m -th sampling points, which contains the location information and kinematics information of the ship. With this background, the method filters out the outliers by taking the statistics of data distribution statistics into account. Assuming that these parameters are normally distributed, the distribution can be identified by the mean and the variance calculated from the

samples. According to the 3σ rule, the outlier points in the data can be eliminated. Taking the longitude lng as an example, formulas (7)–(9) show how to eliminate outlier points. If equation (10) is satisfied, lng^m of the sampling point traj^m is considered as an abnormal value. lng^m needs to be removed and replaced with blank placeholders.

$$\overline{\text{lng}} = \frac{1}{M} \sum_1^M \text{lng}^m, \quad (7)$$

$$\varepsilon_{\text{lng}}^m = \text{lng}^m - \overline{\text{lng}}, \quad (8)$$

$$\sigma_{\text{lng}} = \sqrt{\frac{1}{M-1} \sum_1^M \varepsilon_{\text{lng}}^m}, \quad (9)$$

$$|\varepsilon_{\text{traj}}^m| > 3\sigma_{\text{traj}}. \quad (10)$$

Because of the AIS system broadcasting frequency and the above outlier elimination process, there will be some missing data at different time points. That is, the time intervals between the sampling points in a track Traj may not be equal. For example, the time interval between point traj^i and point traj^{i+1} may not be the same as that between point traj^{i+1} and point traj^{i+2} . The purpose of this portion is to form a continuous time series with equal frequencies using the interpolation method. In particular, different interpolation methods are used to fill in the blanks according to the variable sparsity of a track Traj . Through the initial window length of 240 s, the whole track Traj can be divided into L windows to identify its sparsity. The smaller the size of L is, the sparser the data is (i.e., the smaller the sampling frequency is).

- (1) If $L < 15$, then the trajectory data is too sparse, and it is difficult to restore the missing information even through the interpolation method. For such cases, we discarded these trajectories.
- (2) If $15 \leq L < 20$, then the trajectory data are sparse for a portion of the time windows. Thus, we reduce the window length to 120 s to guarantee the density of the data in shorter windows. Then the linear interpolation is selected for the sequences in each shorter window.
- (3) If $20 \leq L$, it means that the sampling frequency of the data is relatively consistent. For such dense data, the Hermitian cubic interpolation achieves better results than linear interpolation.

As the proposed method is employed to interpolate various sparsity situations of the trajectory data, a continuous time series with equal frequency can be obtained, in which the frequency is 1 Hz.

3.1.2. Pairwise Trajectory Selection. Through data cleaning and trajectory interpolation, a dataset of a fine single trajectory was obtained. To predict the collision risk in an encounter situation, it is necessary to match the pairwise

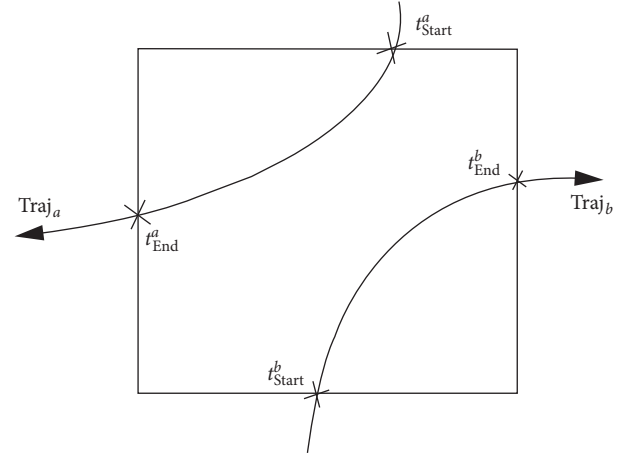


FIGURE 2: Diagram of space-time constraints during the encounter process.

trajectories of these ship pairs. The matching rule takes both time and space constraints into account. Specifically, as shown in Figure 2, these selected pairwise trajectories ($\text{Traj}_a, \text{Traj}_b$) should have intersections in the time dimension and be close to each other in the space dimension.

- (1) If $[t_{\text{Start}}^a, t_{\text{End}}^a]$ and $[t_{\text{Start}}^b, t_{\text{End}}^b]$ denote the time intervals of the ship a and ship b respectively, then $[t_{\text{Start}}^a, t_{\text{End}}^a] \cap [t_{\text{Start}}^b, t_{\text{End}}^b] \neq \emptyset$.
- (2) If $D_{ab} = [d^1, d^2, \dots, d^t]$ represents the relative distance between two ships, d_{max} is the distance threshold for assessing the encounter between ships. Then, $\forall d^t \in D_{ab}, d^t < d_{\text{max}}$.

Those trajectories that are subject to the above two constraints can be selected as pairwise samples. In addition, according to the experience of experts and the definition of an encounter, $d_{\text{max}} = 6$.

In crossing situation, two ships are crossing to involve a collision risk. One ship is coming from either the left or right direction of the other ship's bow, and the relative azimuth between the two ships is 5.7° to 112.5° .

- (i) Head-on situation: Two ships are meeting on reciprocal or nearly reciprocal courses to involve a collision risk. One ship sees the other ahead or nearly ahead, and the relative azimuth between two ships is -5.7° to 5.7° .
- (ii) Overtaking situation: Two ships are sailing on identical or nearly identical course to involve a collision risk. One ship comes up to another ship from 112.5° to 247.5° .

3.2. Feature Extraction of Intership Behaviors. This part aims to obtain insights into the dynamic encounter process through utilizing a sequence of behavioral features. These features have been established by merging the six parameters in a fixed time window, including the relative velocity, course difference, and relative distance as well as three types of azimuths. The coordinate system presented in Figure 3 is

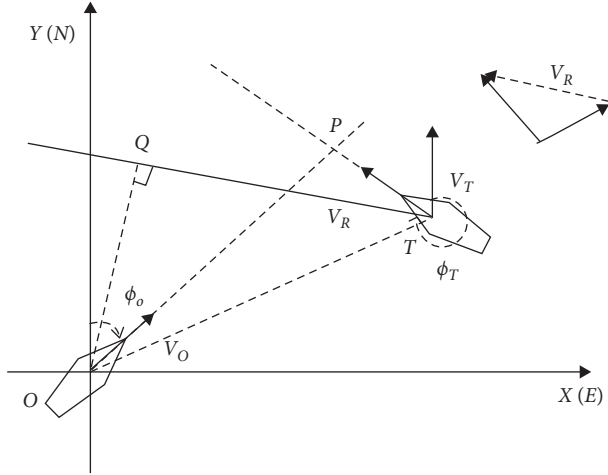


FIGURE 3: Rectangular coordinate system of own ship and target ships.

established to offer insights into calculating a set of parameters by modeling the spatial relationship between ships encountering each other. As shown in Figure 3, the point O indicates the position of the own ship, and lng_o , lat_o , V_O , and ϕ_O are the longitude, latitude, SOG, and COG of own ship. Moreover, the point T represents the location of the target ship, and lng_t , lat_t , V_T , and ϕ_T are longitude, latitude, SOG, and COG of the target ship. The relative velocity, which is denoted as V_R , is as follows:

$$V_R = |V_O - V_T|. \quad (11)$$

A represents the course difference between the own ship and target ship through incorporating ϕ_O and ϕ_T as follows:

$$A = \begin{cases} \phi_O - \phi_T, & |\phi_O - \phi_T| \leq 180^\circ, \\ 360^\circ - (\phi_O - \phi_T), & |\phi_O - \phi_T| > 180^\circ. \end{cases} \quad (12)$$

D_{ot} denotes the relative distance between the two ships, which is estimated by merging with a set of parameters including lng_o , lat_o , lng_t , and lat_t and the Earth radius denoted as R , as follows:

$$D_{ot} = R \times \arccos(\sin(\text{lat}_o)\sin(\text{lat}_t) + \cos(\text{lat}_o)\cos(\text{lat}_t)\cos(\text{lng}_t - \text{lng}_o)), \quad (13)$$

and α is the true azimuth between the two ships, which can be computed as follows:

$$\alpha = \begin{cases} \arccos\left(\frac{\text{lat}_o - \text{lat}_t}{D_{ot}} \times 60\right), & \text{lat}_o > \text{lat}_t, \\ 360^\circ - \arccos\left(\frac{\text{lat}_o - \text{lat}_t}{D_{ot}} \times 60\right), & \text{lat}_o \leq \text{lat}_t. \end{cases} \quad (14)$$

α_o and α_t denote relative azimuths of two ships, respectively, which are defined as follows:

$$\begin{aligned} \alpha_o &= \alpha - \phi_O, \\ \alpha_t &= \alpha - \phi_T. \end{aligned} \quad (15)$$

Thus, the parameters $w = (V_R, A, D_{ot}, \alpha, \alpha_o, \alpha_t)$ are regarded as the behavioral features of encountering ship pairs.

4. Collision Risk Prediction Model

In this section, we propose a novel collision risk prediction algorithm, which can perceive the potential risk at an early stage by mapping current behavior to future collision risk. To this end, first, the risk level of the current encounter situation is calibrated through the widely used CRI method. Then a deep recurrent neural network structure is used to establish the mapping between the ship's current behavior and future collision risk; then the problem of risk prediction is formed as a time series classification task.

4.1. Collision Risk Calibration. Collision risk calibration is a process used to calculate the risk level for encountering ship pairs. It should be noted that the risk level obtained from the calibration is only an assessment based on the current situation. However, the purpose of this study is to predict future collision risk. Therefore, it is necessary to establish the mapping relationship between the current behavior w^t and the risk level $R^{t+\Delta T}$ after a period of time. ΔT is the prediction horizon, which represents the time interval between observed behavior and predicted risk.

During the training process, a large set of w^t and $R^{t+\Delta T}$ will be prepared to train the model. In this section, the calibration process of the risk level R will be described. As a widely used way of risk calibration, the CRI is used to warn of the collision risk by setting off a collision alarm based on diverse factors influencing the collision risk. In particular, various parameters are taken into account in our calibration process, including the DCPA, TCPA, relative distance, and course difference between two ships [35].

$$\text{DCPA} = D_{ot} \times \sin(\angle \text{OTQ}),$$

$$\text{TCPA} = D_{ot} \times \frac{\cos(\angle \text{OTQ})}{V_R},$$

$$\text{CRI}_{\text{basic}} = \left[a_{\text{dcpa}} \left(\frac{\text{DCPA}}{D_s} \right)^2 + a_{\text{tcpa}} \left(\frac{\text{TCPA}}{T_s} \right)^2 + a_d \left(\frac{D_{ot}}{D_s} \right)^2 \right]^{-(1/2)},$$

$$\text{CRI} = \text{CRI}_{\text{basic}} F_{\text{DCPA}} F_{\text{TCPA}} F_{\text{cd}}, \quad (16)$$

where D_{ot} denotes the relative distance between the ships encountering each other. D_s and T_s are the minimum safe distance and time necessary to perform evasive maneuvers; we set them as 0.5 miles and 10 minutes, respectively. Moreover, a_{dcpa} , a_{tcpa} , and a_d are the weights coefficients depending on the state of visibility at sea, the length and beam of the ship, and the type of water area. According to [11], F_{cd} is a multiplier reflecting the encounter danger degree in different encounter situations. Specifically, regarding the course difference between the ships involved in the encounter, the encounter situations can be divided into three categories and the corresponding value of each

multiplier F_{cd} is obtained in Table 2. Moreover, F_{DCPA} and F_{TCPA} are the amplification coefficients of DCPA and TCPA, which are somewhat inversely proportional to the values of DCPA and TCPA. The formulas for calculating the amplification coefficients are given as follows:

$$\begin{aligned} F_{TCPA} &= \exp^{-(TCPA/10)}, \\ F_{DCPA} &= \exp^{-DCPA}. \end{aligned} \quad (17)$$

Obviously, from the above equation, CRI is a continuous value. However, continuous CRI values do not directly indicate the urgency of a ship collision risk. In other words, even if we know the value of the CRI, we cannot be certain about the danger level it represents. Here, we apply the different risk stages of ship encounters to divide the CRI into five different risk levels: low (L), low-middle (LM), middle (ML), middle-high (MH), and high (H):

$$\left\{ \begin{array}{l} 1, \quad \text{if } 0 \leq \text{CRI} < \tau_1 \text{ [low risk level],} \\ 2, \quad \text{if } \tau_1 \leq \text{CRI} < \tau_2 \text{ [low - middle risk level],} \\ 3, \quad \text{if } \tau_2 \leq \text{CRI} < \tau_3 \text{ [middle risk level],} \\ 4, \quad \text{if } \tau_3 \leq \text{CRI} < \tau_4 \text{ [middle - high risk level],} \\ 5, \quad \text{if } \tau_4 \leq \text{CRI} \text{ [high risk level],} \end{array} \right. \quad (18)$$

where τ_1 , τ_2 , τ_3 , and τ_4 are threshold values that need to be determined to separate different risk levels. We can determine them through analyzing the distribution of CRI, which is computed from AIS data of the encounter ships involved in the encounter. In addition, it has been put forward that the statistical probability of the CRI is equal in each encounter stage [36]. In following this reasoning, we calculate the corresponding CRI values for all the samples by using equation (18). All the CRI values are sorted and divided into five equal intervals according to the frequency. The endpoint of the i -th interval is the threshold τ_i . Figure 4 shows the thresholds selected for each risk level, the left side of Figure 4 illustrates the cumulative probability of the CRI in these time windows, and the right side of Figure 4 counts the number of each encounter stage, which is closely related to the corresponding risk levels. Thereby, the threshold values of the five risk levels are provided as follows:

- (1) The CRI values between 0.00 and 0.13 are ranked as the low risk level.
- (2) The CRI values between 0.13 and 0.20 are ranked as the low-middle risk level.
- (3) The CRI values between 0.20 and 0.28 are ranked as the middle risk level.
- (4) The CRI values between 0.28 and 0.45 are ranked as the middle-high risk level.
- (5) The CRI values larger than 0.45 are ranked as the high risk level.

Following the above process of risk discretization, the risk level R^t at a different time t is obtained. As mentioned earlier, we will match the behavior sequence w^t with the risk $R^{t+\Delta T}$ to obtain the training set. From the perspective of machine learning classification, w^t is the temporal feature,

TABLE 2: The relationship between F_{cd} and course difference.

Course difference	$0^\circ - 60^\circ$	$60^\circ - 150^\circ$	$150^\circ - 180^\circ$
F_{cd}	1	8.5	2.34

and $R^{t+\Delta T}$ is the label. Thus, the problem of risk prediction is transformed into a problem of sequence classification. The following section will introduce the sequence classification method used in this paper.

4.2. Risk Prediction Model. With the fast development of deep learning, recurrent neural networks (RNNs) have gained great success in recent years [37] in terms of sequence classification. While an RNN has the ability to make full use of the information of the historical input, it is difficult to manage the long-term dependence caused by the fast failure of nodes. As one of the advanced RNNs, LSTM networks address this issue by modifying the internal RNN cell structure. In particular, LSTM contains a set of memory blocks consisting of one or more autocorrelative memory cells and three gates, that is, input, output, and forget gates. In following this structure, a memory block can retain the relevant historical information [38]. Besides, the sequence of behavioral features is considered to be a typical time series; thus, it follows that the issue of risk prediction can be treated as a time series classification task. In view of the above, it is reasonable to think that LSTM networks can provide valuable insight for predicting the collision risk between ships encountering each other. With this modeling framework, an understanding of the sequence of behavioral features and their relationships with collision risk can be achieved.

In our implementation, we assume that w^t is the sequence of behavioral features in the t -th time window; in addition, $R^{t+\Delta T}$ represents the risk level profiles in the $t + \Delta T$ -th time window computed in terms of equation (18). With the evolution of the encounter process, the LSTM networks are employed to learn the mapping between w^t and $R^{t+\Delta T}$. Figure 5 shows the modeling framework of this mapping; it can be clearly observed from Figure 5 that the sequence of behavioral features involved in the time window is effectively related to the risk level. Thus, the ship encounter risk prediction is achieved by utilizing the encounter dataset under three encounter situations.

5. Experimental Results and Discussion

5.1. Study Areas. The South Channel intersection waterway, an important and busy shipping channel located on the Yangtze Estuary, was selected as the study area. Figure 6 shows an electronic chart of the South Channel intersection waterway. Figure 6 indicates that a large number of ships in this waterway lead to complex encounter situations. In such a water area with dense traffic flow, the early identification of risk is very important for navigation safety. In this study, we use an AIS dataset for 1729 ships in the South Channel intersection waterway from 07/01/2019 to 08/31/2019. Subsequently, the sequence of behavioral features is

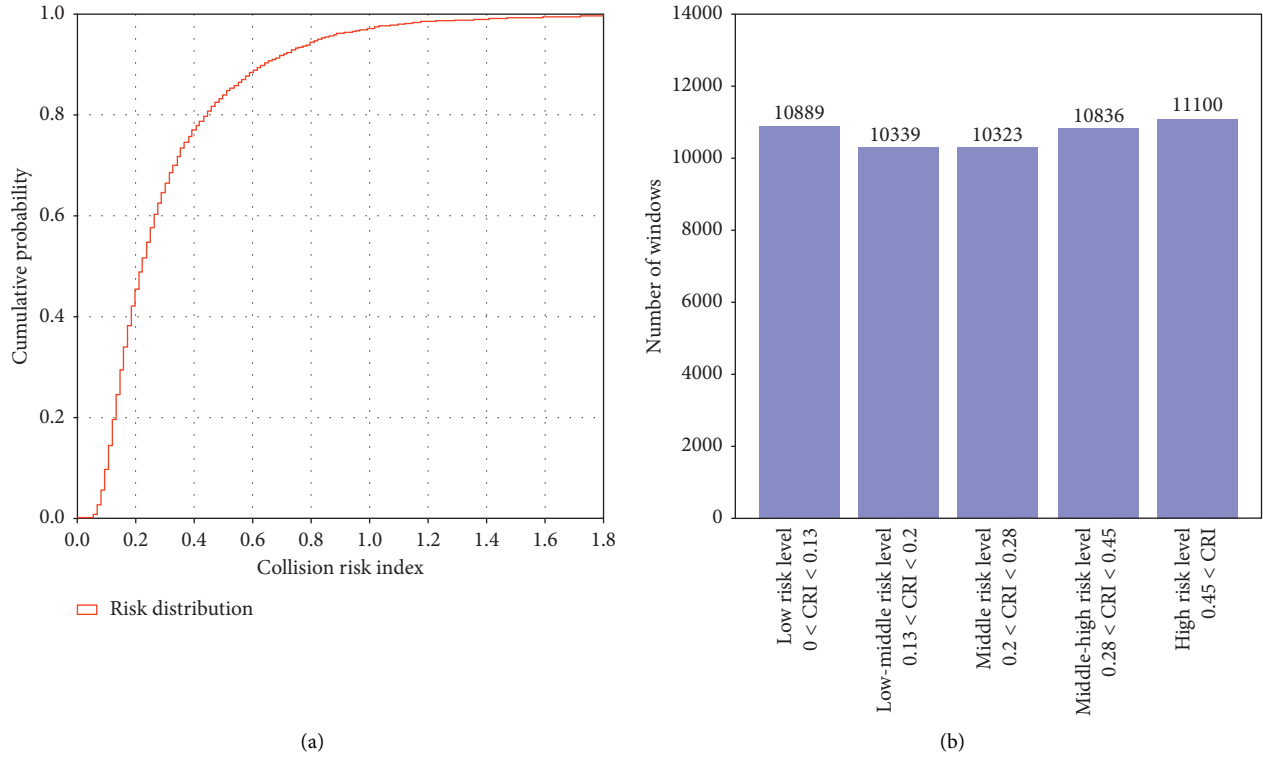


FIGURE 4: Cumulative probability of CRI and number of CRI in each stage.

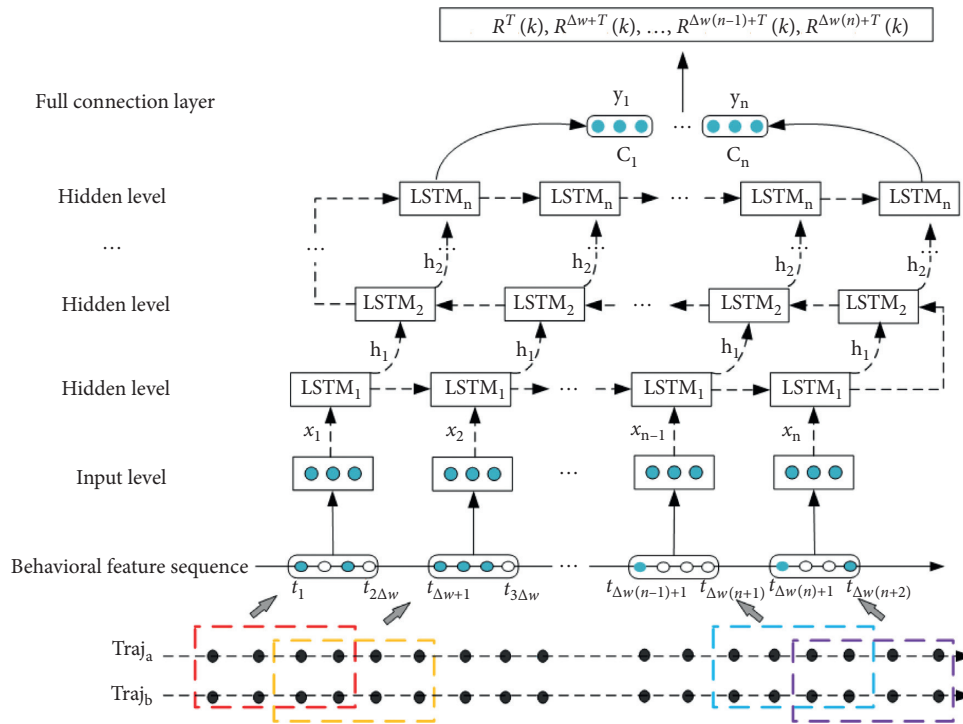


FIGURE 5: The model framework of ship encounter risk prediction.

constructed by determining the length of time window and sliding step. In our implementation, we set the different window lengths to 20 s and 10 s; thus, there are 185,208

records in the dataset. The records of the individual encounter situations are shown in Table 3. Among them, 159,798 records were employed to train the risk prediction



FIGURE 6: Electronic chart of the South Channel intersection waterway of the Yangtze Estuary.

TABLE 3: Records of time windows.

Category of situation	Total	Training set	Validation set	Test set
		Individual	Individual	Individual
Crossing situation		45774	5100	2600
Head-on situation	185208	74805	8310	2250
Overtaking situation		39219	4350	2800

model, 17,760 records were employed to determine the optimal parameters of the model, and the remaining records were used for testing. Figure 7 shows the risk level distribution in the test data, thus providing an opportunity to advance our test data knowledge.

5.2. Parameters in the Experiment. This section discusses the various experimental parameters to find an optimal parameter combination to accurately predict the risk of ship encounters. First, we compare different prediction horizons, that is, 30 seconds and 40 seconds. It is reasonable that ship officers have sufficient time to react to emergencies with these prediction horizons. For improving the accuracy of the model, the grid search method is adopted to determine the optimal number of hidden layers and the learning rate of LSTM in the cross-validation set. Figure 8 shows the cross-validation results under the two prediction horizons. In the case of 30 seconds in advance, the peak value of the prediction accuracy is obtained when the hidden layer is 2 and the learning rate is 0.1, and the accuracy is 0.8712. When it is 40 seconds in advance, the optimal number of hidden layers and the learning rate should be 3 and 0.00010, respectively, and the corresponding accuracy is 0.8676. Finally, the number of LSTM units in the individual hidden layer is 18, which is closely related to the six types of parameters in the sequence of behavioral features.

5.3. Experimental Result. In this section, we evaluate the prediction accuracy and robustness in a typical scenario of three encounter situations (crossing, head-on, and

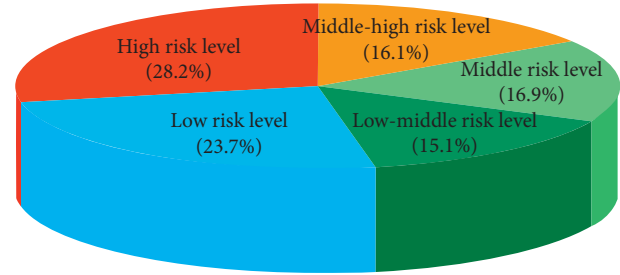


FIGURE 7: Proportion of risk levels in the test set.

overtaking). In particular, Figures 9–11 display a series of comparisons between the ground truth of the risk level and the predicted results in an individual scenario, and each of them contains eight subgraphs. As in the above process, the ground truth and predicted risk level here both refer to the future risk level corresponding to the current window. Figures 9(a)–11(b) and 9(b)–11(b) show the spatial distribution of the predicted values under various prediction horizons, including 30 seconds and 40 seconds, respectively. Moreover, Figures 9(c)–11(c) present the spatial distribution of the real risk level to evaluate the accuracy of the risk prediction. Furthermore, Figure 9(d)–11(d) and 9(e)–11(e) demonstrate the risk level of each window under the two prediction horizons. It can be observed that the parts above and below the horizontal line are the real risk level and the predicted risk level, respectively. Finally, to study the dynamic change of the risk in the encountering process, we divide the whole encounter process into five stages according to time. Figures 9(f)–11(f) and 9(g)–11(g) are the histograms displaying the predicted risk level of each stage under the three encounter scenarios, and the ratios of individual risk level are intuitively presented in Figure 9(h)–11(h), which are closely related to the prediction accuracy of each stage. The following three typical encounter scenarios are analyzed.

Figure 9 shows the predicted results in the crossing situation and compares them with the actual values. As shown in Figure 9(c), the collision risk is initially at the low risk level, and it is continuing at that level for a while until the ships involved in the encounter sail into the warning zone (area indicated by the red dotted line). Subsequently, the collision risk gradually rises to the high risk level, while one of the ships is in the center of the warning zone, and it commences evasive maneuvers to achieve a safe encounter. Moreover, as shown in Figures 9(a) and 9(b), the risk level predicted by both models is almost inconsistent with the real values at the beginning. With the evolution of the encounter process, there are certain deviations. Later, all the models correctly predict the risk, especially while the collision risk is at a high risk level. It follows that effective predictions can be made by taking advantage of the sequence of behavioral features. Figures 9(d)–9(g) show that the model has the ability to yield more superior predictions, while the prediction horizon is shorter. However, the variation tendencies of true values and predicted values are fairly consistent. As shown in Figure 9(h), the model has achieved high prediction accuracy in general. The proposed approach is

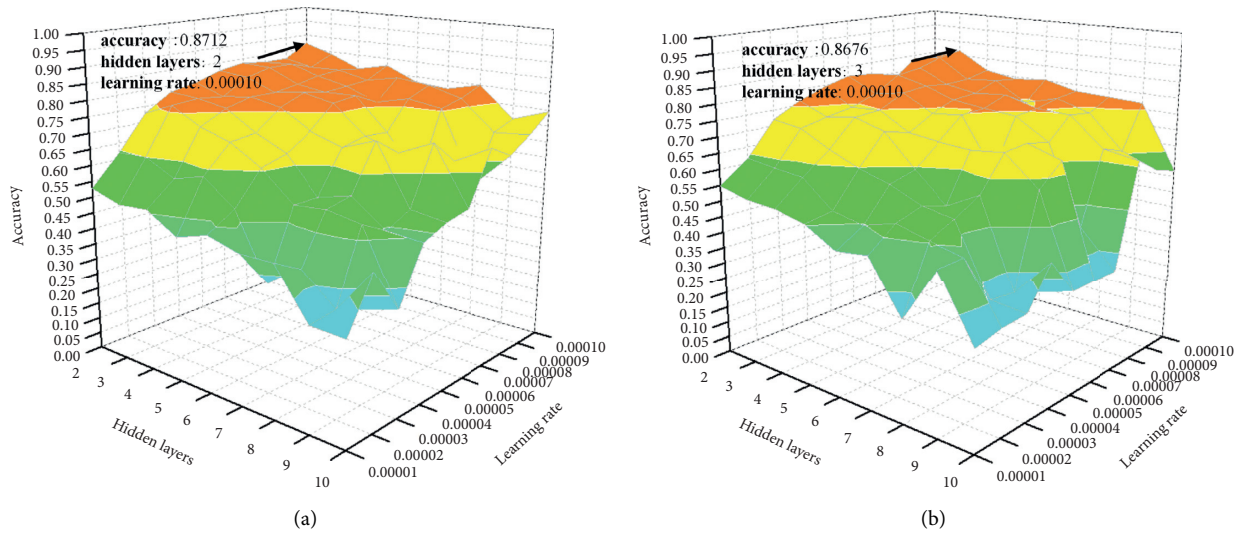


FIGURE 8: Determining the optimal hidden layers and learning rate of LSTM via a grid search in the validation set.

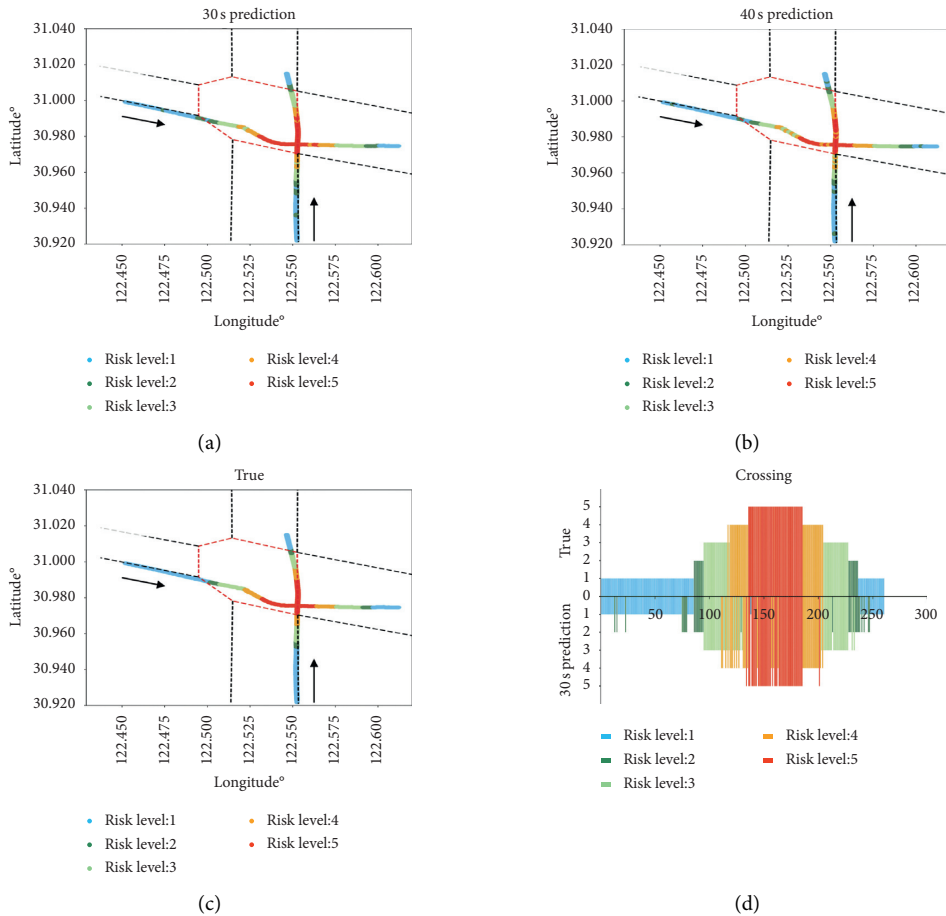


FIGURE 9: Continued.

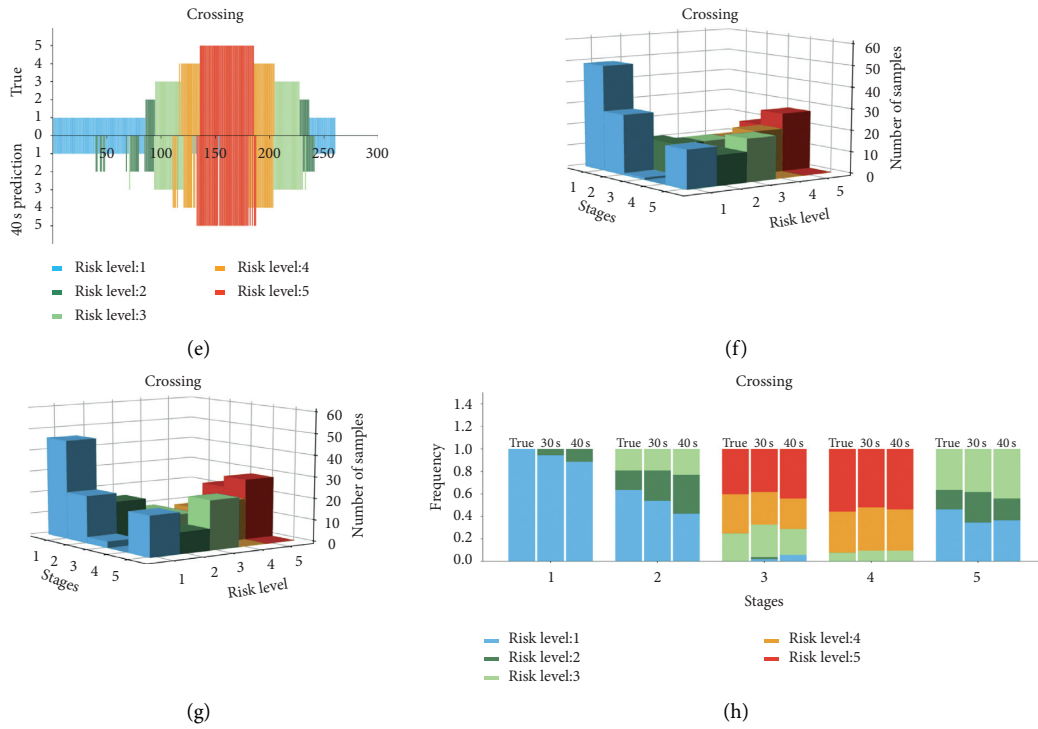


FIGURE 9: Crossing situation: (a) and (b) are the spatial distributions of risk level under various prediction horizons; (c) is the corresponding spatial distribution of realistic risk level. (d) and (e) show the risk level statistics of each window under two prediction horizons. (f) and (g) are the number of samples of each risk level in the five stages. (h) is the frequency of each risk level at various stages.

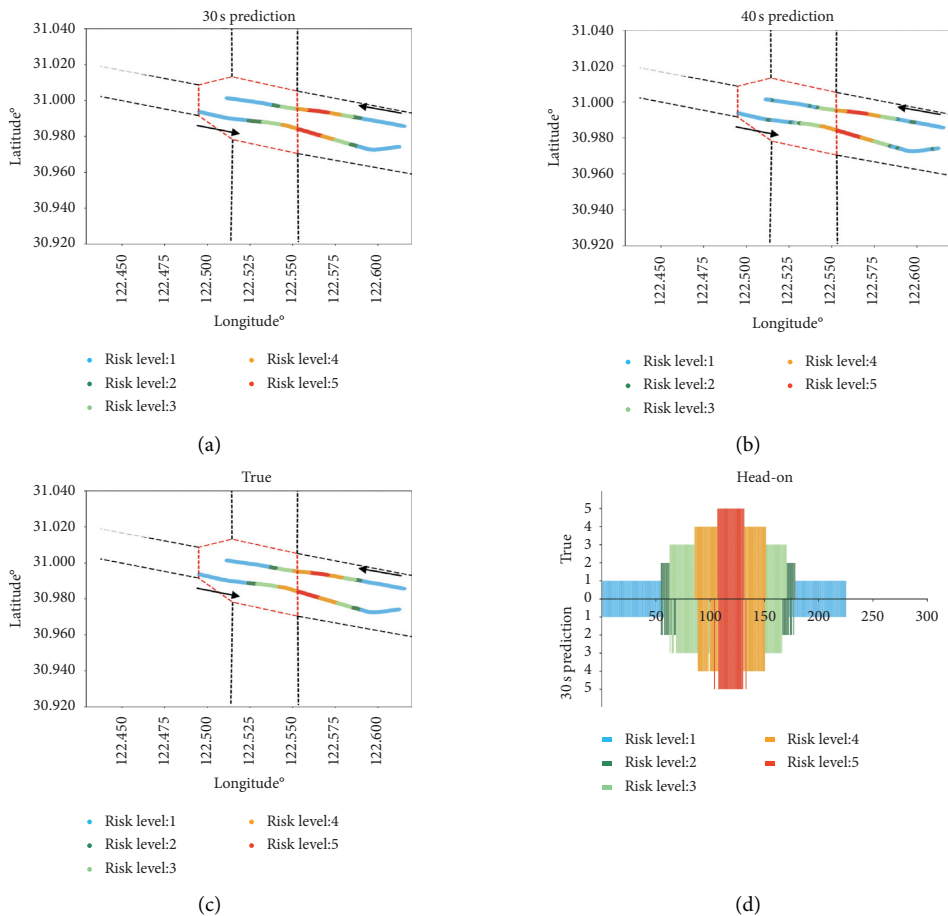


FIGURE 10: Continued.

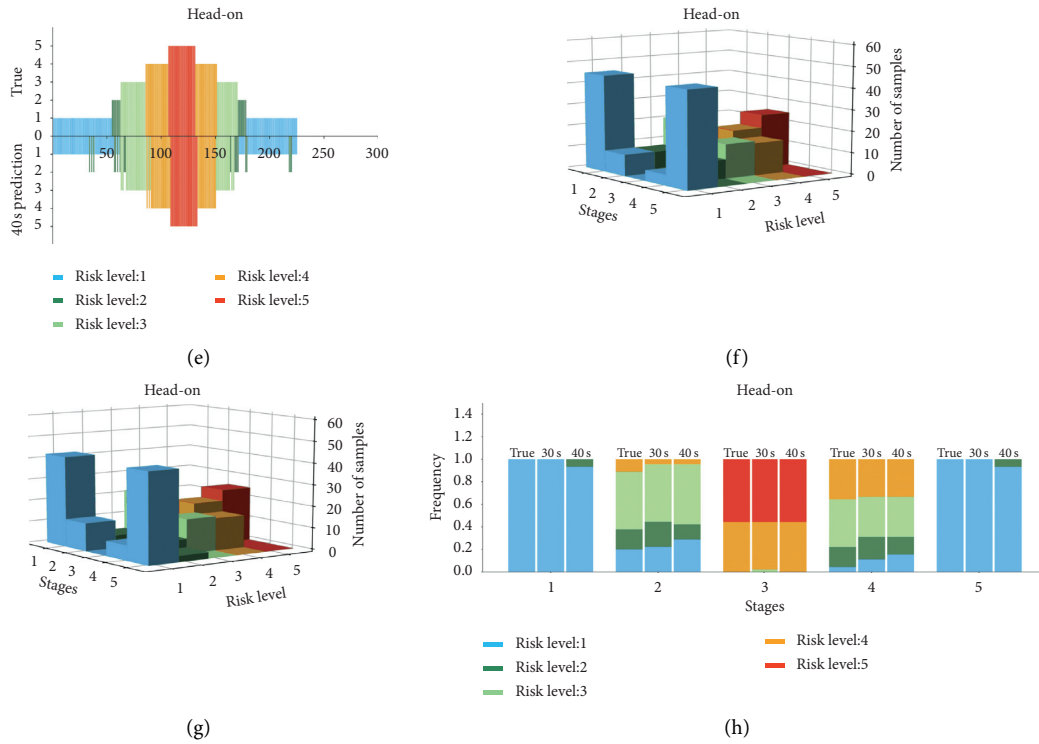


FIGURE 10: Head-on situation: (a) and (b) are the spatial distribution of risk level under various prediction horizons; (c) is the corresponding spatial distribution of realistic risk level. (d) and (e) show the risk level statistics of each window under two prediction horizons. (f) and (g) are the number of samples of each risk level in the five stages. (h) is the frequency of each risk level at various stages.

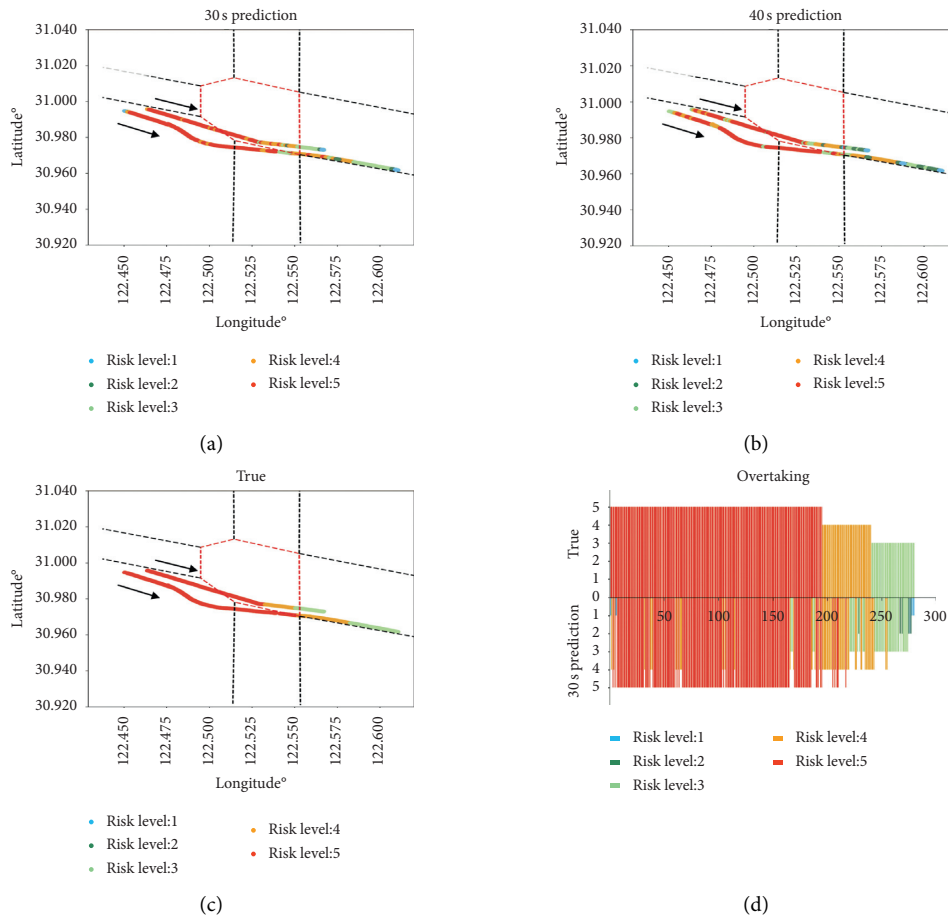


FIGURE 11: Continued.

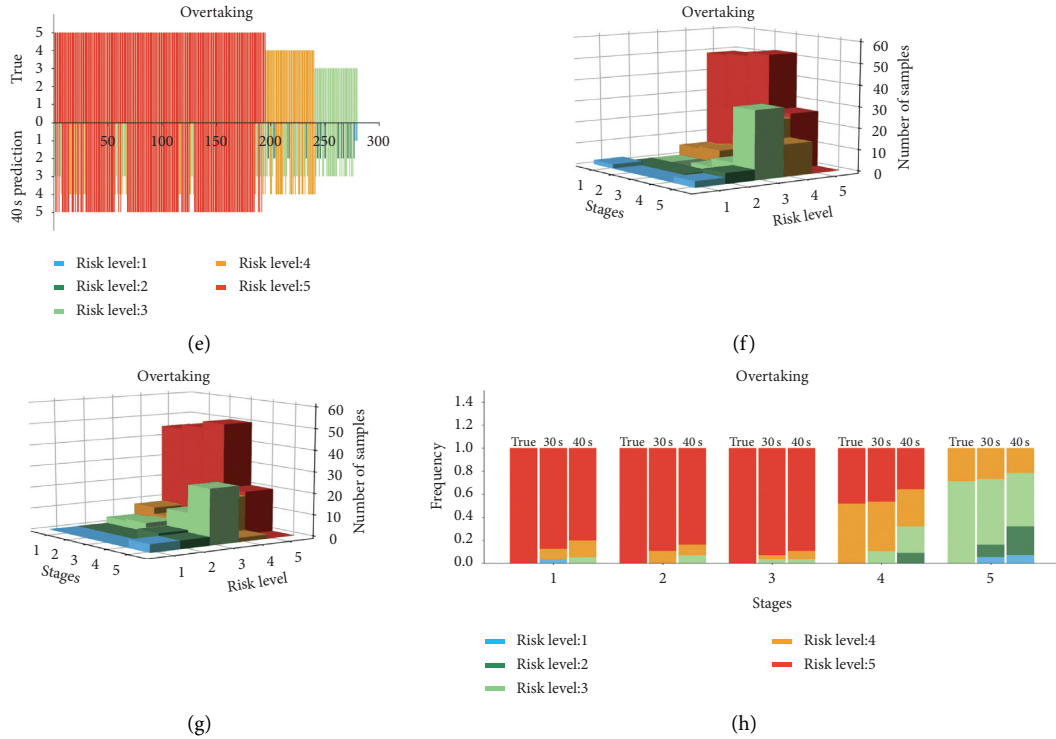


FIGURE 11: Overtaking situation: (a) and (b) show the spatial distribution of the risk level under various prediction horizons; (c) is the corresponding spatial distribution of the realistic risk level; (d) and (e) show the risk level statistics of each window under the two prediction horizons. (f) and (g) are the number of samples of each risk level in the five stages. (h) is the frequency of each risk level at various stages.

capable of predicting the collision risk under a crossing situation by making full use of the spatiotemporal behaviors of the ships involved in the encounter.

As shown in Figure 10(c), collision risk is initially at a low risk level since the two ships are far apart. As the head-on process evolves, the risk of collision increases gradually. However, since both ships sail in their respective channels, neither of them takes anticollision maneuvers under such a circumstance, although the high risk level has been maintained for a certain period. Eventually, the risk of collision gradually disappears, because the two ships have passed each other (Past and Clear). Moreover, the spatial distribution of risk level in Figures 10(a) and 10(b) is consistent with that in Figure 10(c). In this example, it seems that there is no difference between the predicted results in the 30-second horizon and those in the 40-second horizon. This may be because the motion state of the ship does not change during the course in terms of speed and direction keeping, and the prediction accuracy of the collision risk has only a small relationship with the advance of time. As shown in Figures 10(d)–10(g), it can be observed that the predicted results differ from the actual collision risk, thereby confirming the suitability of our approach in a head-on situation. In particular, accurate prediction of a high risk level is of great significance to avoid potential conflicts in congested waterways because dangerous encounters occur occasionally in these high-risk zones. Therefore, it necessitates additional attention and caution in these areas to ensure safe encounters between ships.

TABLE 4: Confusion matrix under a prediction horizon of 30 seconds.

Risk level	L	LM	M	MH	H
L	903	7	8	15	18
LM	131	117	63	26	31
M	32	20	468	43	43
MH	18	15	62	449	78
H	16	13	32	82	1135
SUM	1100	172	633	615	1305
ACCURACY (%)	82	68	74	73	87

For the overtaking situation, Figure 11(c) shows that the collision risk has continually been at a high risk level. This is primarily because the relative distance and course between the two ships are small, which makes it easier for a potential collision accident to occur. Soon afterward, the collision risk tends to decline gradually, while one ship leaves the warning zone, marking a safe encounter between two ships as well. The risk distributions with the high risk level in Figure 11(c) are fairly consistent with those in Figures 11(a) and 11(b), which shows that the risk prediction model has a high recall rate for high risk cases. However, for other risk levels, there are certain deviations between the predicted results and the real value. Eventually, we can observe from Figure 11(h) that the model can perceive a high risk from the beginning, which suggests that once the overtaking situation is formed, there is a high risk in the initial stage. In this case, the ship officer can pay additional attention to the possible collision risk according to this early warning model.

TABLE 5: Confusion matrix under a prediction horizon of 40 seconds.

Risk level	L	LM	M	MH	H
L	893	6	13	21	15
LM	135	118	58	24	27
M	49	19	456	37	59
MH	13	16	72	431	83
H	10	13	34	102	1121
SUM	1100	172	633	615	1305
ACCURACY (%)	81	69	72	70	86

To evaluate the model performance in all the test sets, the comparisons of the overall prediction precision results of the collision risk for the two prediction horizons are shown in Tables 4 and 5. From the overall sample, this model can predict the risk accurately in different horizons. The ability to identify the risk situations can effectively warn the ship officers of potential collisions, which could provide the basis for a navigation decision. Moreover, in terms of the different horizons, the model is more accurate in predicting collision risks that may occur in the near future than in predicting those further away. In practical applications, this model needs to balance the tradeoff between prediction accuracy and horizon length.

6. Conclusions

An AIS data-driven approach has been derived for collision risk prediction in a vessel encounter situation by learning the intership behavior. The approach considers the relationship between intership behavior and future collision risk, which helps to predict the potential collision risk in various encounter situations in advance. To illustrate the approach, the intership behavior is transformed from AIS traces to a sequence of behavioral features by combining a fixed set of parameters. Then, we related the sequence of behavioral features involved in a specified time window to the risk level at a future time; then, the mapping between them was established through an RNN. Furthermore, we tested the approach over encounter cases in the South Channel intersection waterway with various prediction horizons. The prediction results demonstrated that the approach has reasonable and effective ability and that the risk predicted in advance is consistent with the ship encounter situations. In particular, the model has an outstanding ability to identify risk through intership behavior when the potential collision risk is at a high level. This research offers a valuable insight into collision risk prediction by intership behaviors, and the approach is expected to be applied to the implementation of a new collision warning system.

Data Availability

This Data Statement has been confirmed, and there is no further revision by the authors.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant nos. 51679182 and 71874132, the Green Intelligent Inland Ship Innovation Programme, and the Fundamental Research Funds for the Central Universities under Grant 2020-YB-035.

References

- [1] Y. He, Y. Jin, L. Huang, Y. Xiong, P. Chen, and J. Mou, "Quantitative analysis of COLREG rules and seamanship for autonomous collision avoidance at open sea," *Ocean Engineering*, vol. 140, pp. 281–291, 2017.
- [2] I. Cohen, W.-P. Brinkman, and M. A. Neerinx, "Modelling environmental and cognitive factors to predict performance in a stressful training scenario on a naval ship simulator," *Cognition, Technology & Work*, vol. 17, no. 4, pp. 503–519, 2015.
- [3] H. C. Chin and A. K. Debnath, "Modeling perceived collision risk in port water navigation," *Safety Science*, vol. 47, no. 10, pp. 1410–1416, 2009.
- [4] G. Pallotta, M. Vespe, and K. Bryan, "Vessel pattern knowledge discovery from AIS data: a framework for anomaly detection and route prediction," *Entropy*, vol. 15, no. 12, pp. 2218–2245, 2013.
- [5] A. Dobrkovic, M. E. Iacob, and J. van Hillegersberg, "Using machine learning for unsupervised maritime waypoint discovery from streaming AIS data," in *Proceedings of the 15th International Conference on Knowledge Technologies and Data-Driven Business*, ACM, Graz, Austria, October 2015.
- [6] Z. Xiao, L. Ponnambalam, X. Fu, and W. Zhang, "Maritime traffic probabilistic forecasting based on vessels' waterway patterns and motion behaviors," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 11, pp. 3122–3134, 2017.
- [7] P. Silveira, A. Teixeira, C. Guedes Soares, and T. Santos, *Assessment of Ship Collision Estimation Methods Using AIS Data*, Taylor & Francis Group, London, UK, 2015.
- [8] R. Christian and H. G. Kang, "Probabilistic risk assessment on maritime spent nuclear fuel transportation (part II: ship collision probability)," *Reliability Engineering & System Safety*, vol. 164, pp. 136–149, 2017.
- [9] J. H. Ahn, K. P. Rhee, and Y. J. A. You, "Study on the collision avoidance of a ship using neural networks and fuzzy logic," *Applied Ocean Research*, vol. 37, pp. 162–173, 2012.
- [10] C. N. Hwang, J. M. Yang, and C. Y. Chiang, "The design of fuzzy collision-avoidance expert system implemented by H-autopilot," *Journal of Marine Science and Technology*, vol. 9, pp. 25–37, 2001.
- [11] J. M. Mou, C. v. d. Tak, and H. Ligteringen, "Study on collision avoidance in busy waterways by using AIS data," *Ocean Engineering*, vol. 37, no. 5–6, pp. 483–490, 2010.
- [12] W. Zhang, F. Goerlandt, J. Montewka, and P. Kujala, "A method for detecting possible near miss ship collisions from AIS data," *Ocean Engineering*, vol. 107, pp. 60–69, 2015.
- [13] Y. Ren, J. Mou, Q. Yan, and F. Zhang, "Study on assessing dynamic risk of ship collision," in *Proceedings of the First International Conference on Transportation Information and Safety (ICTIS)*, pp. 2751–2757, Wuhan, China, April 2012.
- [14] P. A. M. Silveira, A. P. Teixeira, and C. G. Soares, "Use of AIS data to characterise marine traffic patterns and ship collision risk off the coast of Portugal," *Journal of Navigation*, vol. 66, no. 6, pp. 879–898, 2013.

- [15] W. Zhang, F. Goerlandt, P. Kujala, and Y. Wang, "An advanced method for detecting possible near miss ship collisions from AIS data," *Ocean Engineering*, vol. 124, pp. 141–156, 2016.
- [16] W. Zhang, X. Feng, Y. Qi, F. Shu, Y. Zhang, and Y. Wang, "Towards a model of regional vessel near-miss collision risk assessment for open waters based on AIS data," *Journal of Navigation*, vol. 72, no. 6, pp. 1449–1468, 2019.
- [17] R. Szlapczynski, "A unified measure of collision risk derived from the concept of a ship domain," *Journal of Navigation*, vol. 59, no. 3, pp. 477–490, 2006.
- [18] R. Szlapczynski and J. Szlapczynska, "An analysis of domain-based ship collision risk parameters," *Ocean Engineering*, vol. 126, pp. 47–56, 2016.
- [19] X. Wu, A. L. Mehta, V. A. Zaloom, and B. N. Craig, "Analysis of waterway transportation in Southeast Texas waterway based on AIS data," *Ocean Engineering*, vol. 121, pp. 196–209, 2016.
- [20] N. Wang, "An intelligent spatial collision risk based on the quaternion ship domain," *Journal of Navigation*, vol. 63, no. 4, pp. 733–749, 2010.
- [21] X. Qu, Q. Meng, and L. Suyi, "Ship collision risk assessment for the Singapore Strait," *Accident Analysis & Prevention*, vol. 43, no. 6, pp. 2030–2036, 2011.
- [22] E. Tu, G. Zhang, L. Rachmawati, E. Rajabally, and G. B. Huang, "Exploiting AIS data for intelligent maritime navigation: a comprehensive survey from data to methodology," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, pp. 1559–1582, 2017.
- [23] G. A. Susto, A. Cenedese, and M. Terzi, "Time-series classification methods: review and applications to power systems data," in *Big Data Application in Power Systems*, pp. 179–220, Elsevier, Amsterdam, Netherlands, 2018.
- [24] A. Graves, "Generating sequences with recurrent neural networks," <https://arxiv.org/abs/1308.0850>.
- [25] A. Graves and N. Jaitly, "Towards end-to-end speech recognition with recurrent neural networks," in *Proceedings of the International Conference on Machine Learning*, pp. 1764–1772, Beijing, China, June 2014.
- [26] Z. Shi, M. Xu, Q. Pan, B. Yan, and H. Zhang, "LSTM-based flight trajectory prediction," in *Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN)*, July 2018.
- [27] L. Sun, Z. Yan, S. M. Mellado, M. Hanheide, and T. Duckett, "3DOF pedestrian trajectory prediction learned from long-term autonomous mobile robot deployment data," in *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA)*, May 2018.
- [28] L. Gang, Y. Wang, Y. Sun, L. Zhou, and M. Zhang, "Estimation of vessel collision risk index based on support vector machine," *Advances in Mechanical Engineering*, vol. 8, 2016.
- [29] J.-B. Yim, D.-S. Kim, and D.-J. Park, "Modeling perceived collision risk in vessel encounter situations," *Ocean Engineering*, vol. 166, pp. 64–75, 2018.
- [30] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer Science & Business Media, Berlin, Germany, 2009.
- [31] Z. He, F. Yang, Z. Li, K. Liu, and N. Xiong, "Mining channel water depth information from IoT-based big automated identification system data for safe waterway navigation," *IEEE Access*, vol. 6, pp. 75598–75608, 2018.
- [32] T. Macduff, "The probability of vessel collisions," *Ocean Industry*, vol. 9, 1974.
- [33] D. Chen, C. Dai, X. Wan, and J. Mou, "A research on AIS-based embedded system for ship collision avoidance," in *Proceedings of the 2015 International Conference on Transportation Information and Safety (ICTIS)*, pp. 512–517, IEEE, Wuhan, China, June 2015.
- [34] F. Kaneko, "Methods for probabilistic safety assessments of ships," *Journal of Marine Science and Technology*, vol. 7, no. 1, pp. 1–16, 2002.
- [35] Y. Koldenhof, C. Van der Tak, and C. Glansdorp, "Risk Awareness; a model to calculate the risk of a ship dynamically," in *Proceedings of the XIII International Scientific and Technical Conference on Marine Traffic Engineering*, pp. 112–119, Malmö, Sweden, October 2009.
- [36] H. Yixiong, Y. Xiong, H. Liwen et al., "Studies of last steering point/CRI basis on MMG and ship domain," *Journal of Wuhan University of Technology: Transportation Science and Engineering*, vol. 38, pp. 1088–1091, 2014.
- [37] A. Zyner, S. Worrall, and E. Nebot, "A recurrent neural network solution for predicting driver intention at unsignaled intersections," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1759–1764, 2018.
- [38] J. Zhao, Y. Gao, Z. Bai, H. Wang, and S. Lu, "Traffic speed prediction under non-recurrent congestion: based on LSTM method and BeiDou navigation satellite system data," *IEEE Intelligent Transportation Systems Magazine*, vol. 11, no. 2, pp. 70–81, 2019.