Ocean Engineering 282 (2023) 115020

Contents lists available at ScienceDirect

Ocean Engineering

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

Toward Multimodal Vessel Trajectory Prediction by modeling the distribution of modes

Siyang Guo [a], Hui Zhang [a,*], Yaming Guo [b]

[a] Inner Mongolia University, School of Electronic Information Engineering, Hohhot, 010021, China
[b] Jilin University, School of Artificial Intelligence, Changchun, 130012, China

ARTICLE INFO

Keywords:
Vessel trajectory prediction
Multimodal prediction
Trajectory modes
Adversarial learning
Invertible mapping
Smooth mapping

ABSTRACT

Vessel trajectory prediction using AIS data plays an important role in maritime navigation warning and safety. A key aspect of trajectory prediction is multimodal because of the uncertainty of vessel behavior. However, complex trajectory modes are difficult to be learned from low-dimensional AIS data with noise. In this paper, we propose a new method for multimodal vessel trajectory prediction, called Multimodal Vessel Trajectory Prediction via Modes Distribution Modeling (VT-MDM). This approach addresses the above challenges by introducing additional hiding regimes to characterize complex trajectory modes independently. Specifically, we introduce an additional latent vector as the encoding of the trajectory modes, which is randomly sampled from a multivariate Gaussian distribution to generate multiple predicted trajectories. To enable this Gaussian distribution for capturing the vessel trajectory modes, we use adversarial learning to enforce all its realizations to generate realistic predicted trajectories. Furthermore, we also encourage the mapping between the latent vectors of the modes and the predicted trajectories to be invertible and smooth, which prompts VT-MDM to produce truly and gradually multimodal predicted trajectories. Experiments on the real AIS dataset show that our method is capable of multimodal trajectory prediction with high accuracy.

1. Introduction

The maritime transportation environment has become increasingly complex in the past few decades, with the rapid development of the vessel industry. This complexity has led to an increased risk of collisions, which have caused casualties, environmental damage, and substantial economic losses. In light of this, it is important for the maritime industry, risk assessment, and traffic management to fully utilize maritime surveillance data for vessel trajectory prediction (Ozbas, 2013; Rajabi et al., 2018). Automatic Identification System (AIS) is now widely used in maritime traffic due to its unprecedented high resolution and real-time accuracy in analyzing vessel behavior (Mao et al., 2018). The International Maritime Organization (IMO) adopted the AIS navigation system in 2000 for the purposes of traffic control and coastal surveillance. AIS can provide static data (e.g., IMO number, name, type, and length, etc.), dynamic navigation data (e.g., longitude, latitude, speed, and course, etc.), voyage-related data, as well as short safety messages (Robards et al., 2016). Additionally, it enables automatic and continuous data exchange within the range of vessels and between vessels and coastal authorities (Goudossis and Katsikas, 2019). For maritime traffic monitoring, AIS has distinguished itself among the existing self-reporting positioning systems as a reliable source of information regarding spatial coverage, vessel coverage, and frequency of information transmission. As a result, numerous studies have been conducted to predict the future trajectories of vessels using historical AIS data, thereby improving the management of marine traffic and enhancing the safety of marine navigation (Hexeberg et al., 2017; Suo et al., 2020).

Multimodality A key aspect of vessel trajectory prediction is the multimodality of behavior (Huang et al., 2022). Due to the natural uncertainty in the behavior of vessels, there are multiple possibilities for their future trajectories. Specifically, vessel behavior is highly dependent on the current environmental conditions, such as sea breezes, waves, and currents, as well as the decisions made by the navigator, which can be influenced by their physical and psychological state (Chroni et al., 2015). In summary, the vessel may behave differently in future actions, such as accelerating, decelerating, changing direction, etc. In such scenarios, the traditional vessel trajectory prediction techniques that produce a single deterministic trajectory have a low practical application value. As an illustration, Fig. 1 depicts a simulation scenario in which the vessel exhibits different behaviors. Firstly, Figs. 1(a) and 1(b) show the vessels traveling in different directions in an encounter situation. The green vessel does not collide

* Corresponding author.
E-mail address: hui.zhang@imu.edu.cn (H. Zhang).

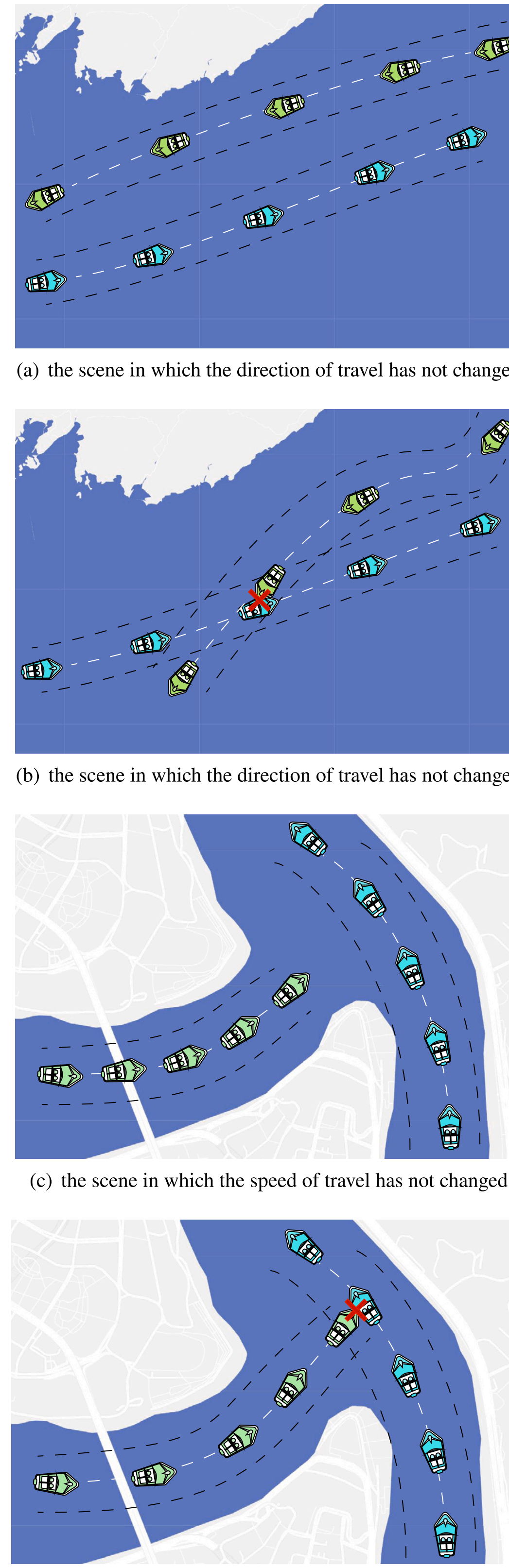
https://doi.org/10.1016/j.oceaneng.2023.115020
Received 14 February 2023; Received in revised form 23 April 2023; Accepted 31 May 2023
Available online 13 June 2023

with the blue vessel under the original direction of travel, but collides when it changes its steering angle. Secondly, Figs. 1(c) and 1(d) show the vessels traveling through an intersection at different speeds. At their initial speeds, the two vessels can avoid one another perfectly, but they collide when the green vessel speeds up. This example highlights that the deterministic method cannot provide information about the potential future trajectory of the green vessel, which might result in a collision. Therefore, it is necessary to consider the multimodality of vessel trajectory, which can reduce collision risk. Some preliminary efforts (Murray and Perera, 2020; Sørensen et al., 2022) have been made on this issue, most of which attempt to mine trajectory modes from AIS data and use them as a priori knowledge for subsequent multimodal vessel trajectory prediction.

Motivation However, complex trajectory modes are difficult to be learned from low-dimensional AIS data with noise. On the one hand, AIS messages are usually represented using a low-dimensional vector (e.g., the 5-D real-valued vector consisting of timestamps, latitude, longitude, speed, and course), which is generated by multiple explanatory factors jointly (Nguyen et al., 2018, 2021). Since the mode factor is tangled with other explanatory factors, disentangling the underlying trajectory modes from low-dimensional vectors is difficult for the model. On the other hand, AIS data is frequently inundated with noise due to inevitable errors, such as environmental disturbances, system failures, and human mistakes (Harati-Mokhtari et al., 2007; Zhang et al., 2018). These noises make it even more difficult for the model to learn the trajectory modes from the AIS data. To address the above challenge, we revisit the multimodal trajectory prediction problem from a new perspective, i.e., by introducing additional hiding regimes to characterize the complex trajectory modes independently. In this way, we eliminate the necessity to equip the model with the ability to distinguish mode factors from AIS data.

Motivated by representation learning (Bengio et al., 2013), we propose *Multimodal **V**essel **T**rajectory Prediction via **M**odes **D**istribution **M**odeling* (VT-MDM) in this paper, which directly models the distribution of trajectory modes for multimodal trajectory prediction. Specifically, we introduce an additional latent vector to encode the trajectory modes, which is randomly sampled from a Gaussian distribution to generate multiple predicted trajectories. To enable this Gaussian distribution to capture the vessel trajectory modes, we use adversarial learning to enforce all of its realizations to generate realistic predicted trajectories. Furthermore, we encourage the mapping between the latent vectors of the modes and the predicted trajectories to be invertible, allowing the Gaussian distribution to characterize the truly multimodal distribution instead of the single mode with high probability. Considering that the vessel behavior will not change abruptly, we also encourage the above mapping to be smooth so that the transition between trajectory modes involving the Gaussian distribution is gradual. We summarize our main contributions as follows:

1. We propose a method to directly model the distribution of trajectory modes, called VT-MDM, aiming to address the multimodal trajectory prediction problem using historical AIS data.
2. Our method uses adversarial learning to enable an additional Gaussian distribution to capture the vessel's trajectory modes. We emphasize that VT-MDM does not need to disentangle the mode factor from the low-dimensional AIS data, and that the mode space is continuous.
3. The proposed method attempts to create an invertible and smooth mapping between the latent vectors of the modes and the predicted trajectories, which aids in truly and gradually multimodal trajectory prediction. As far as we know, this is the first time that the truly and gradually multimodal distribution has been concerned.
4. We validate the effectiveness of the proposed method using the real AIS dataset. The results present that VT-MDM achieves truly and gradually multimodal trajectory prediction while maintaining high prediction accuracy.



(a) the scene in which the direction of travel has not changed.

(b) the scene in which the direction of travel has not changed.

(c) the scene in which the speed of travel has not changed.

(d) the scene in which the speed of travel has changed.

Fig. 1. The simulation scenario.

Organization The paper's remainder is structured as follows. Firstly, Section 2 provides an overview of related work and analyses the limitations of existing models for AIS trajectory prediction. Secondly, Section 3 presents background information about the least squares GANs and seq2seq model. The details of the proposed approach are presented in Section 4. In Section 5, we perform a thorough experimental exploration of the proposed method's performance on the real AIS dataset. Finally, conclusions and future work are discussed in Section 6.

2. Related work

2.1. AIS based vessel trajectory prediction

AIS data provide insights into the historical behavior of vessels in given regions, which can be used for maritime traffic data mining and prediction techniques. Predicting the vessel's trajectory is a process of using historical AIS data around the expected position of the vessel to predict its location at one or more future time points. Due to the limitations of AIS technology, such as highly irregular temporal sampling, poor data quality and completeness, and various behaviors exhibited by vessels (depending on, for example, their size, navigation status, traffic rules, etc.), vessel trajectory prediction is still a challenging problem. Currently, several results have been achieved in the research of vessel trajectory prediction based on AIS data. Based on the basic implementation mechanisms of various approaches, they can be broadly classified into two categories: Kinematic model-based approach and Deep learning based approach.

2.1.1. Kinematic model-based approach

Early approaches to vessel trajectory prediction rely on physical models of vessel movement. Sutulo et al. (2002) proposed a dynamic mathematical model for vessel trajectory prediction based on the vessel's current speed and acceleration. The model also considers trajectory prediction in maneuvers involving changing course. Perera et al. (2012) proposed an Extended Kalman Filter (EKF)-based method for predicting vessel trajectories by adding estimated noise to the kinematic model. For underdriven surface vessels, Assaf et al. (2020) uses the Unscented Kalman Filter (UKF) to predict their trajectory position. Perera et al. (2010) proposed an EKF to estimate the vessel's state, which was then used to predict its trajectory. However, these methods have the following drawbacks. First, developing a precise mathematical model for vessel trajectories, which accounts for dynamic factors like wind and current, can be challenging. Second, kinematic models can be unreliable for longer prediction ranges, as vessel trajectories are often highly non-linear due to navigator decisions. Most kinematic models have difficulty in modeling continuous motion behavior, which can lead to unsatisfactory multi-step continuous prediction results.

2.1.2. Deep learning based approach

In the majority of cases, one might easily surpass the boundaries of manually created kinematic models. This has sparked research into more adaptable, data-driven statistical methods. The widespread use of neural networks has brought vessel trajectory prediction to a new stage. Gao et al. (2021) enable multi-step prediction of vessel trajectories by combining TPNet and long and short-term memory networks (LSTM). However, since the hyperparameters of the neural networks are difficult to obtain the optimal solution manually, Qian et al. (2022) suggested using the genetic algorithm (GA) to optimize the key hyperparameters of the LSTM network, which effectively improves the accuracy and speed of trajectory prediction. Murray and Perera (2021) pointed out that decomposing the historical vessel behavior in a given geographical area into clusters with similar behavioral characteristics in advance can effectively improve the prediction accuracy. Park et al. (2021) applied a spectral clustering approach to the similarity measured by the longest common subsequence (LCSS) distance. Based on the clustering results, a vessel trajectory prediction model was developed using bidirectional long and short-term memory (Bi-LSTM). Recently, attention networks have also been used extensively for vessel trajectory prediction, yielding promising results. Capobianco et al. (2021b) proposed an attention-based mechanism for aggregation layers to learn the relation between the observed and the predicted kinematics states while preserving the spatio-temporal structure of the input. Furthermore, to model the prediction uncertainty of future estimates, they extend the deep learning framework for trajectory prediction tasks by generating the corresponding prediction uncertainty via Bayesian modeling of epistemic and aleatoric uncertainties (Capobianco et al., 2021a, 2022). However, traditional deep learning approaches use the Euclidean distance between the ground truth and the predicted trajectory to train the model, causing the model to learn the "average behavior" instead of the actual trajectory. Gupta et al. (2018) propose using generative adversarial networks for pedestrian trajectory prediction and find that it can better generate more realistic trajectories.

2.2. Multimodal vessel trajectory prediction

Some scholars have also done corresponding work on multimodal trajectory prediction of vessels. Rong et al. (2019) proposed a probabilistic trajectory prediction model based on the Gaussian process to describe the uncertainty of the future vessel position through a continuous probability distribution. However, their model only considered the uncertainty in the lateral and longitudinal motion of the vessel, which may limit its ability to capture the diversity of vessel trajectory modes. Sørensen et al. (2022) modeled vessel trajectories using an eleven-dimensional Gaussian distribution and introduced a hybrid density network that predicts probabilistic future positions of the vessel and generates multiple predicted trajectories through multiple sampling. Murray and Perera (2020) estimated the distribution of potential possible future vessel trajectories by training a dual linear autoencoder, and by sampling from this distribution, multiple trajectories can be predicted. All the above methods attempt to mine vessel trajectory modes directly from AIS data and incorporate them as prior knowledge into later predictions. Nguyen and Fablet (2021) extended the original AIS data into a high-dimensional space and transformed the regression problem into a classification problem using a transformer to extract relevant information from the historical AIS data of the target vessel to predict the future position distribution. However, higher dimensional representations may result in additional noise. As discussed earlier, complex modes are difficult to learn from low-dimensional AIS data with noise. Thus, existing methods invariably make a strong assumption that the modes are discrete. We believe this leads to a loss of information, as the modes should be described using a continuous space. Lastly, none of the above works considered the gradualness problem of multimodal trajectory prediction.

3. Preliminaries

3.1. Least square GANs

Generative adversarial networks (GANs) (Goodfellow et al., 2020) have been extensively studied in the past few years. A GANs is a generative model that consists of two deep neural networks: one for generating fake samples and another for discriminating between real and fake samples, which compete with each other until convergence.

The GANs are implemented as follows. The user provides some training samples as real samples $\{x_i | i = 1, 2, \dots, m\}$, called the sample distribution $p_r(x)$. Two deep neural networks, called discriminator D and generator G, are constructed. We randomly sample from distribution $z \sim p_z(z)$ and feed it to generator G to generate samples $G(z)$, which essentially maps the distribution $z \sim p_z(z)$ to the distribution $p_g(x)$ (called the generating distribution). The input to $D(x)$ is the sample x, which may come from either the sample distribution or the generating distribution. And the output of $D(x)$ is a scalar representing

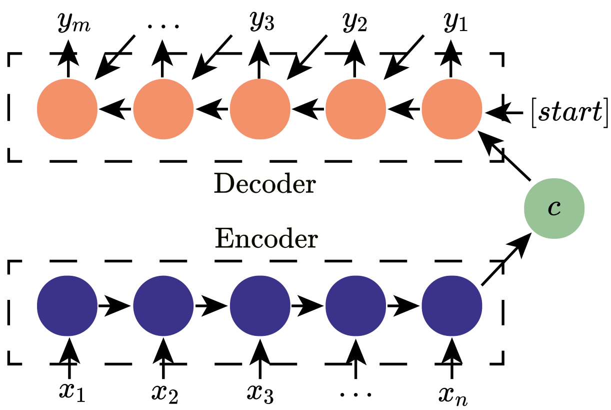


Fig. 2. The seq2seq model.

the probability that x is obtained from the sample distribution. We train $D(x)$ to maximize the probability of assigning the correct labels to the real and generated samples and train $G(z)$ to minimize this probability. $D(x)$ and $G(z)$ are trained alternately until a Nash equilibrium is reached. In other words, the optimization objective of the GANs is as follows:

$$\max_D \mathcal{L}_{\mathrm{GAN}}(D) = \mathbb{E}_{x\sim p_r(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))] \tag{1}$$

$$\min_G \mathcal{L}_{\mathrm{GAN}}(G) = \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))], \tag{2}$$

where real sample is labeled 1 and fake sample is labeled 0. The discriminator of regular GANs performs 0/1 classification of real and fake samples, and uses binary cross-entropy function for training. This makes GANs extremely unstable during the training process (Arjovsky and Bottou, 2017). To solve this problem, Mao et al. (2017) proposed the Least Squares Generative Adversarial Networks (LSGANs), in which the discriminator is trained using a least-squares loss function. The optimization objective of LSGANs is defined as:

$$\min_D \mathcal{L}_{\mathrm{LSGAN}}(D) = \mathbb{E}_{x\sim p_r(x)}\left[(D(x)-1)^2\right] + \mathbb{E}_{z\sim p_z(z)}\left[D^2(G(z))\right] \tag{3}$$

$$\min_G \mathcal{L}_{\mathrm{LSGAN}}(G) = \mathbb{E}_{z\sim p_z(z)}\left[(D(G(z))-\gamma)^2\right], \tag{4}$$

where γ is the probability that the generator wants to fool the discriminator. In the trajectory prediction task, γ is generally taken to be 1.

3.2. Seq2seq model

Standard recurrent neural networks (RNNs) can be effectively used in tasks where the alignment between inputs and outputs is known in advance. Yet, there is little research on how to apply it to tasks where the input–output alignment is unclear and complex. In response Cho et al. (2014) first proposed a novel neural network called RNN encoder–decoder, which allows variable length input and output sequences, addressing the shortcomings of standard RNNs. And then Sutskever et al. (2014) proposed a sequence-to-sequence (seq2seq) model to improve it. The seq2seq model is an end-to-end network model that can be trained to learn arbitrary temporal structures in input sequences and has achieved great success in machine translation, speech recognition, text summarization, question-and-answer systems, and other fields (Zhang et al., 2019; Ma et al., 2018; Ghazvininejad et al., 2018).

The seq2seq model contains two main components, an encoder, and a decoder. The encoder learns the input and encodes it into a low-dimensional latent vector, which is subsequently passed to the decoder, which outputs it by learning the latent vector. The decoder has two learning methods. The first is that the potential vector only participates in the initial operation, and the second is that the potential vector participates in each subsequent decoding operation. Fig. 2 illustrates case one.

The seq2seq model uses the idea of maximizing the likelihood function for the joint training of the encoder and decoder, which is represented by the following problem.

$$\max_\theta \frac{1}{N}\sum_{i=1}^{N} \log p_\theta\left(y_i \mid x_i\right), \tag{5}$$

where θ denotes the model parameters, N is the number of samples in the training set, x_i and y_i are the input and output sequences, respectively.

4. Methodology

In this section, we present a detailed description of the proposed method's implementation. The first part provides the problem definition (Section 4.1). Then, the framework of the proposed method (Section 4.2) as well as the architecture of the model used (Section 4.3) are presented. Finally, the loss functions used in the proposed method are summarized (Section 4.4)

4.1. Problem definition

We assume that the N-selected historical vessel trajectories are defined as:

$$\Delta P_{f,i} = \{(\Delta lon_i^t, \Delta lat_i^t) \in \mathbb{R}^2 | t = 1, \dots, t_{obs}\}, \tag{6}$$

where i is the index of the vessel for $\forall i \in \{1, \dots, N\}$, and Δlon_i^t and Δlat_i^t denote the relative longitude and relative latitude, respectively. To help VT-MDM learn the distribution of the actual vessel trajectory modes more accurately, the model uses asymmetric input and output. Therefore, the model uses the initial state of the vessel as input, which is defined as:

$$s_i = \{(\Delta P_{f,i}, v_i^t, c_i^t, \Delta T_{s,i}^t) \in \mathbb{R}^4 | t = 1, \dots, t_{obs}\}, \tag{7}$$

where v_i^t is the speed over ground, c_i^t is the course over ground, and $\Delta T_{s,i}^t$ represents the time difference between two adjacent feature points of the vessel. And the predicted trajectory can be defined as:

$$\Delta \hat{P}_{b,i} = \{(\Delta \hat{l}on_i^t, \Delta \hat{l}at_i^t) \in \mathbb{R}^2 | t = t_{obs+1}, \dots, t_{pred}\}. \tag{8}$$

And the future trajectory (ground truth) can be defined similarly as:

$$\Delta P_{b,i} = \{(\Delta lon_i^t, \Delta lat_i^t) \in \mathbb{R}^2 | t = t_{obs+1}, \dots, t_{pred}\}. \tag{9}$$

4.2. Overall framework

In this subsection, we outline the framework of the proposed method, VT-MDM. Our goal is to achieve multimodal trajectory prediction by modeling the distribution of trajectory modes. We first use adversarial learning to enable an additional Gaussian distribution that captures the vessel trajectory modes. Then, we use a reconstruction task to construct an invertible mapping between the latent vectors of the modes and the predicted trajectories. Finally, we use global regularization and local interpolation to encourage the above mapping to be smooth as well. The overall framework of VT-MDM is shown in Fig. 3.

Modeling distribution To avoid learning complex trajectory modes from low-dimensional AIS data with noise, we introduce an additional hidden regime to model the trajectory modes' distribution directly. The distribution of trajectory modes can be regarded as the root of the uncertainty in the vessel behavior, which governs the vessel's movement. For this purpose, we introduce an additional latent vector $z \in \mathbb{R}^n$ as the encoding of the trajectory modes, which randomly sampled from a multivariate Gaussian distribution $\mathcal{N}(\mathbf{0}, \mathbf{1})$. Firstly, we input the latent vector z and the historical vessel trajectories obs_traj into

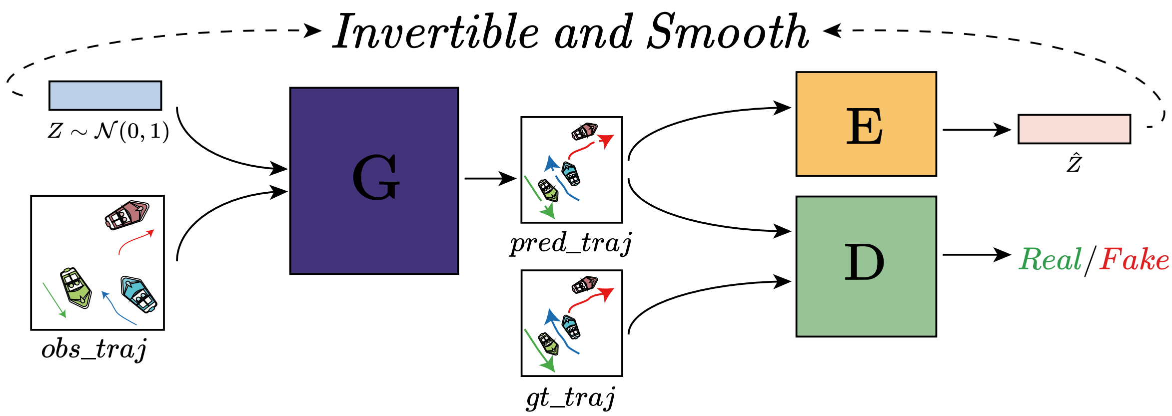


Fig. 3. The framework of VT-MDM.

the generator G to produce the predicted trajectories $pred\_traj$. Next, we feed the generated predicted trajectories $pred\_traj$ and the ground truth gt_traj to the discriminator D to classify the real from the fake, where $pred\_traj$ is fake and gt_traj is real. We then train the generator and the discriminator alternately, and they compete against each other until the discriminator cannot differentiate between the predicted and the ground truth trajectories. As a result, the generator can generate multiple realistic predicted trajectories by picking different z. As all realizations of the Gaussian distribution can generate realistic predicted trajectories, we can consider that it successfully captures the vessel trajectory modes. It is worth mentioning that such a modal space is continuous, which does not lead to information loss and is more practical.

Invertible mapping Although the Gaussian distribution captures the vessel trajectory modes, there may be a many-to-one mapping from the latent vector of modes to the predicted trajectory. In this case, the Gaussian distribution characterizes a single mode with high probability instead of a truly multimodal mode. Therefore, we attempt to construct invertible mappings between the latent vector z and the predicted trajectory $pred\_traj$. Inspired by the reconstruction task (Zhu et al., 2017), and we perform the transformation starting from the latent vector, i.e., $z \rightarrow pred\_traj \rightarrow \hat{z}$. In more detail, we generate the predicted trajectory $pred\_traj$ from a latent vector z and try to recover z from $pred\_traj$ using the latent encoder E. By training the generator G and the latent encoder E simultaneously, we can achieve the one-to-one mapping between the latent vectors and predicted trajectory, thus allowing the Gaussian distribution to characterize a larger distribution of trajectory modes.

Smooth mapping Unlike small objects, such as vehicles or pedestrians, vessels exhibit smoother movement patterns, which means that the transition between vessel trajectory modes should be gradual. For instance, a vessel traveling in a low-speed mode will not abruptly switch to a fast mode. To ensure that the transition between trajectory modes is gradual, we aim to construct a smooth mapping between the latent vector z and the predicted trajectory $pred_traj$. Liu et al. (2021, 2022) point out that the term "smooth" implies a compact latent space. Therefore, to compress the latent space, we employ both global regularization and local interpolation techniques. At the global level, we encourage the latent distribution of the ground truth encoded by E to be close to the Gaussian distribution, which forces the latent vector of modes to shrink near the origin. Locally, we expect that linear interpolation between different latent vectors can also generate realistic predicted trajectories to fool the discriminator, which corresponds to removing the unrealistic part between two different latent vectors. With the two techniques mentioned above, we prompt the latent space of the modes to become more compact, thus allowing for a gradual transition between trajectory modes.

4.3. Model architecture

The overall architecture of the proposed VT-MDM is presented in Fig. 4. The model consists of three main components. The first component is the prediction trajectory generator, which is used to generate future trajectories of vessels. The second component is the trajectory discriminator, which is used to determine whether the input trajectory belongs to the ground truth or the trajectory generated by the generator. The third component is the latent encoder, which is used to construct an invertible and smooth mapping between the predicted trajectories and the latent vectors.

4.3.1. Generator

Similar to the seq2seq model (see Section 3.2), the generator has two main components in our model: a feature encoder and a trajectory decoder. The feature encoder encodes the historical vessel trajectories into low-dimensional features and then feeds them to the decoder. The trajectory decoder decodes the features and outputs the multi-step future trajectories.

1. Feature encoder
   The feature encoder consists of a two-layers network. Firstly, a linear layer embeds the historical vessel trajectory into a higher dimension, facilitating subsequent network layers to learn more features. Then, the high-dimensional features are fed into an LSTM to learn the correlation between different vessel features across time steps and embed them into a single output. The final feature encoding $h_{f,i}$ is calculated as follows:

$$\begin{aligned} e_{f,i} &= Linear_{emb}(s_i; W_{emb}) \\ h_{f,i} &= LSTM_{encoder}(e_{f,i}, h_{encoder}(i); W_{encoder}), \end{aligned} \tag{10}$$

   where W_{emb} and $W_{encoder}$ denote the network weights that can be trained.
2. Trajectory decoder
   Similar to the feature encoder, the trajectory decoder also contains linear and LSTM networks. First, the latent vector z sampled from the normal distribution is connected to the feature encoding $h_{f,i}$ to obtain $L_{f,i}$ which contains the historical trajectory features. $L_{f,i}$ is then fed into the LSTM network and decoded into the context information. Finally, the context information is reshaped by two linear layers to output the future trajectories $\Delta\hat{P}_{b,i}$.

$$\begin{aligned} h_{b,i} &= LSTM_{decoder}((L_{f,i}, \Delta P_{f,i}), h_{decoder}(i); W_{decoder}) \\ e_{b,i} &= Linear_{embb}(h_{b,i}; W_{embb}) \\ \Delta\hat{P}_{b,i} &= Linear_{output}(e_{b,i}; W_{output}), \end{aligned} \tag{11}$$

   where $W_{decoder}$, W_{embb} and W_{output} denote the weights of the network.

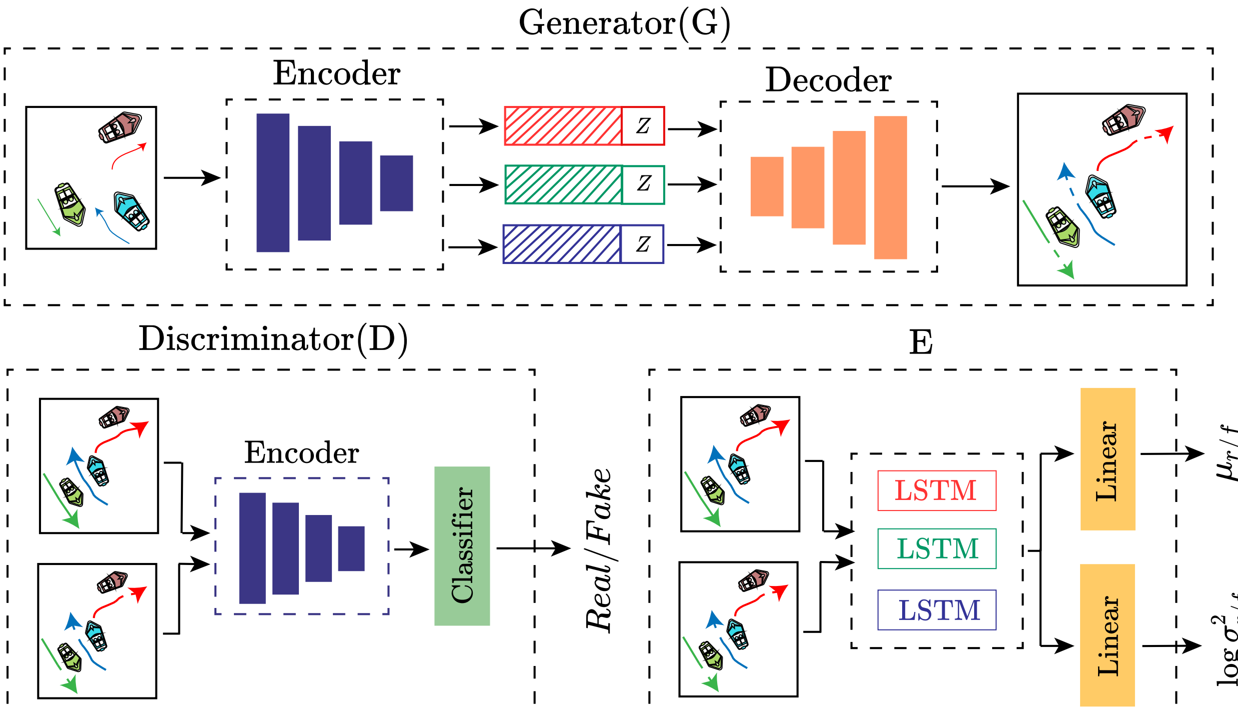


Fig. 4. The model architecture of VT-MDM.

4.3.2. Discriminator

As shown in Fig. 4, the architecture of the discriminator continues the architecture of the feature encoder in the generator. Specifically, it embeds the input trajectory into a low-dimensional encoding and feeds it into a classification network, i.e., a fully connected layer, and finally outputs the scores. Similar to conditional generation adversarial networks (Mirza and Osindero, 2014), the discriminator takes the complete vessel trajectory as input, which contains both historical and future trajectories. The real sample $P_i = (\Delta P_{f,i}, \Delta P_{b,i})$ and the fake sample $\hat{P}_i = (\Delta P_{f,i}, \Delta \hat{P}_{b,i})$ are calculated according to Eq. (10), and then the final score of the trajectory is given by Eq. (12).

$$Sco = Linear_{score}(\tilde{P}_i, W_{score}), \tag{12}$$

where $\tilde{P}_i \sim p(P_i, \hat{P}_i)$ is randomly sampled from the either real or fake sample, and W_{score} denotes the weights of the network.

4.3.3. Latent encoder

We also train a latent encoder that learns latent vectors from different vessel trajectories and generates the mean and variance of the latent vector that best represent the selected vessel's trajectory mode. Specifically, the mean and variance of the future trajectories are calculated using the following equation:

$$\begin{aligned} q_i &= LSTM_q(\tilde{P}_i, h_q(i); W_q) \\ \mu_i &= Linear_\mu(q_i; W_\mu) \\ \log\sigma_i^2 &= Linear_\sigma(q_i; W_\sigma), \end{aligned} \tag{13}$$

where W_q, W_μ and W_σ denote the weights of the network.

4.4. Loss function

Modeling distribution loss. In the adversarial training process, we use real complete vessel trajectories $P_i = (\Delta P_{f,i}, \Delta P_{b,i})$ as real samples. The prediction trajectories $\hat{P}_i = (\Delta P_{f,i}, \Delta \hat{P}_{b,i})$ generated by the generator are used as fake samples, where $\hat{P}_{b,i} = G(z, \Delta P_{f,i})$. The discriminator is trained to distinguish between real and fake samples, and the generator is trained to fool the discriminator. Finally, realistic vessel trajectory modes can be captured by playing the generator and discriminator against each other. The least squares objective (see Section 3.1) is used as the adversarial loss in the proposed method and is written as:

$$\mathcal{L}_{GAN-G}(G) = \mathbb{E}[(D(\hat{P}_i) - 1)^2] \tag{14}$$

$$\mathcal{L}_{GAN-D}(D) = \mathbb{E}[(D(P_i) - 1)^2] + \mathbb{E}[D^2(\hat{P}_i)]. \tag{15}$$

Moreover, to force the generator to generate predicted trajectories that match the ground truth, we use L2 loss between the output trajectories and the ground truth.

$$\mathcal{L}_{pred} = \|\Delta P_{b,i} - \Delta \hat{P}_{b,i}\|_2. \tag{16}$$

The modeling loss is obtained by adding the three losses mentioned above:

$$\mathcal{L}_{modeling}(G, D) = \mathcal{L}_{GAN-G}(G) + \mathcal{L}_{GAN-D}(D) + \lambda_{pred}\mathcal{L}_{pred}, \tag{17}$$

where λ_{pred} controls the importance of $\mathcal{L}_{pred}$.

Invertible mapping loss. To create an invertible mapping between the latent vector and the predicted trajectories, the reconstruction loss is employed. Specifically, we use a randomly sampled latent vector z to generate a predicted trajectory $\hat{P}_i$ and wish to recover that latent vector through the latent encoder. The invertible mapping loss is defined as:

$$\mathcal{L}_{invertible}(E) = \lambda_{invertible}\|E(\hat{P}_i) - z\|_1, \tag{18}$$

where z is a randomly sampled latent vector from the Gaussian distribution, and $\lambda_{invertible}$ controls the importance of $\mathcal{L}_{invertible}$. Following Zhu et al. (2017), we use the re-parameterization trick to obtain z, allowing direct back-propagation.

Smooth mapping loss. To create a smooth mapping between the latent vectors and the predicted trajectories, we use an additional smooth mapping loss for training. First, to compact the space on a global scale, we perform a forced prior, i.e., encourage the encoding of real trajectories to obey a Gaussian distribution:

$$\mathcal{L}_{kl}(E) = \mathbb{E}[\mathcal{D}_{\mathrm{KL}}(E(P_i) \parallel \mathcal{N}(\mathbf{0}, \mathbf{1}))], \tag{19}$$

where $\mathbf{1}$ is the identity matrix, $\mathcal{D}_{\mathrm{KL}}(\cdot \parallel \cdot)$ is the Kullback–Leibler (KL) divergence. In this way, our method is able to perform random sampling during the prediction time.

Second, we use local interpolation to encourage a more compact space. Specifically, we interpolate two mode codes z_1 and z_2 of the same vessel to obtain the mixture vector $Mix(\alpha, z_1, z_2) = \alpha z_1 + (1-\alpha) z_2$, as suggested in Verma et al. (2019). Here, $\alpha \in [0,1]$ is sampled from a $Beta(2,2)$ distribution, and it denotes the weight of the z_1 features included in the mixed variable $Mix(\alpha, z_1, z_2)$. Then, $Mix(\alpha, z_1, z_2)$ is fed to the generator to generate a predicted trajectory $\Delta\hat{P}_{b,mix}$. Finally, $\hat{P}_{mix} = (\Delta P_{f,mix}, \Delta\hat{P}_{b,mix})$ is fed to the discriminator, and we expect the discriminator to classify it as real. Specifically, the local interpolation is achieved by the following loss terms:

$$\mathcal{L}_{Mix}(G) = \mathbb{E}[(D(\hat{P}_{mix}) - 1)^2]. \tag{20}$$

Summing the above losses to obtain $\mathcal{L}_{smooth}$:

$$\mathcal{L}_{smooth}(G, E) = \lambda_{kl}\mathcal{L}_{kl}(E) + \mathcal{L}_{Mix}(G), \tag{21}$$

where λ_{kl} controls the importance of $\mathcal{L}_{kl}$.

Full loss We combine the above losses to obtain the full loss:

$$\min_{G,E}\min_{D}\mathcal{L}_{full}(G, D, E) = \mathcal{L}_{modeling} + \mathcal{L}_{invertible} + \mathcal{L}_{smooth}. \tag{22}$$

Similarly to Kosaraju et al. (2019), the generator G and latent encoder E are trained simultaneously.

5. Experiments

In this section, we evaluate the predictive performance of VT-MDM. We use the real AIS dataset of the Yellow Sea for analysis to validate the efficacy of the proposed approach. The dataset corresponds to that collected from December 13 to December 15, 2021, and consists of the historical trajectories of 1,063 cargo vessels. First, we start by pre-processing the AIS data (Section 5.1). Next, we evaluate its prediction accuracy (Section 5.2), its multimodal prediction performance (Section 5.3), its smoothing analysis (Section 5.4), as well as the uncertainty analysis (Section 5.5) and finally scalability analysis (Section 5.6).

5.1. Data preprocessing

The AIS data is used to develop models for predicting vessel trajectories. Before being used as input data, it requires appropriate pre-processing to ensure its reliability, accuracy, and availability. First, the dataset is filtered according to the following:

- Feature extraction. Select the following information from the raw AIS data: time, MMSI, latitude, longitude, speed over ground, and course over ground.
- Separate the vessel's trajectory. After dividing the dataset according to MMSI, it is sorted by time.
- Time interval splitting. When the interval between two points is greater than 3 h, the two points are used as splitting points to separate two consecutive vessel trajectories.
- Elimination of invalid data. Delete the vessel data whose latitude and longitude at the before and after moments do not meet the given threshold to remove moored or anchored vessels.

After the above processing, the final matrix representation of the AIS data for the ith vessel is as follows:

$$X_i = [lon_i, lat_i, v_i, c_i, T_{s,i}, \mathrm{MMSI}]^T, \tag{23}$$

where $T_{s,i}$ denotes the timestamps.

IMO forces vessels to comply with AIS performance standards, which require static information to be updated every 6 min or as required, and dynamic information to be reported between 2 s and 3 min (Yang et al., 2019). However, AIS data is affected by various factors, such as varying broadcast frequency and packet loss, resulting in inconsistent time intervals. Cubic spline interpolation was introduced to recover vessel trajectories, but different resampling times must be considered for different application scenarios. We argue that the uncertainty of short-term vessel trajectory may have more practical implications, such as avoiding close collisions in busy waters. Therefore, we set the resampling time to 60 s in our experiments. We emphasize that the proposed method can be adapted to predict longer vessel trajectories by adjusting the resampling time. The latitude, longitude, and time data are then differentially processed. At this point, the temporal and spatial dimensions of the vessel trajectory are qualitatively represented by the time interval and relative position, respectively. The speed and heading information are normalized using the maximum–minimum normalization method, which is defined by the following equation:

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \tag{24}$$

where $x_{\min}$ and $x_{\max}$ denote the minimum and maximum values in the sample, x and x^* denote the original data and the normalized data, respectively.

Finally, sliding window extraction decomposes the trajectory data into a list of input/target sequences of a specific length.

5.2. Trajectory prediction accuracy

In this subsection, we validate the prediction accuracy of the proposed method. The vessel's trajectory during the last 24 min is used to predict the trajectory of the next 6 and 12 min. We compare the prediction results of our model with several other deterministic baselines, including the linear regression (Linear), and the LSTM, as well as the seq2seq model.

Evaluation Metrics Similar to prior work (Sørensen et al., 2022; Chen et al., 2022), we use three error metrics to describe the distance error of the predicted value and the ground truth :

1. *Mean Absolute Error* (MAE): The absolute value of the difference between the model prediction and the ground truth over all prediction steps.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{t_{pred}} |(lon_i^t, lat_i^t) - (\hat{lon}_i^t, \hat{lat}_i^t)|, \tag{25}$$

where $(\hat{lon}_i, \hat{lat}_i)$ is the predicted latitude and longitude, and (lon_i, lat_i) is the ground truth.

2. *Root Mean Square Error* (RMSE): The arithmetic square root of the expected value of the square of the difference between the model prediction and the ground truth.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{t_{pred}} [(lon_i^t, lat_i^t) - (\hat{lon}_i^t, \hat{lat}_i^t)]^2}. \tag{26}$$

3. *Haversine formula*: The Haversine formula determines the great circle distance between two points on a sphere based on longitude and latitude. This paper calculates the mean prediction error (in kilometers) within the selected prediction time step using the Haversine distance formula, which is given by:

$$hav(\frac{\mathcal{D}_{hav}}{R}) = hav(lat_i - \hat{l}at_i) + \cos(lat_i)\cos(\hat{lat}_i)hav(lon_i - \hat{l}on_i), \tag{27}$$

where $hav(\theta) = \sin^2(\frac{\theta}{2})$, R is the radius of the earth, and $\mathcal{D}_{hav}$ is the actual distance between the predicted point and the ground truth.

The prediction error results for two prediction lengths of different methods are presented in Tables 1 and 2, where Table 1 corresponds to the error result for a prediction length of 6 min, and Table 2 corresponds to the result for 12 min. The results presented here are the average of 30 times of training for each model. It is worth noting that the LSTM and Linear can only predict the trajectory at one future moment at a time, so the multi-step prediction in recursive form is used.

The results demonstrate that Linear can make accurate predictions for short-time trajectories, but its performance deteriorates rapidly for

Table 1
Comparison of prediction results for different methods at t_{pred} = 6 min.

	Method	Linear	LSTM	seq2seq	VT-MDM(Ours)
MAE (e-3°)	Avg	8.449	5.728	2.011	**1.963**
	Min	8.377	5.299	1.891	1.874
	Max	8.537	6.124	2.137	2.061
RMSE (e-3°)	Avg	11.15	7.641	2.676	**2.521**
	Min	11.06	7.042	2.512	2.405
	Max	11.26	8.141	2.844	2.640
$\mathcal{D}_{hav}$ (km)	Avg	1.399	0.911	0.344	**0.310**
	Min	1.389	0.841	0.323	0.296
	Max	1.411	0.973	0.365	0.325

Table 2
Comparison of prediction results for different methods at t_{pred} = 12 min.

	Method	Linear	LSTM	seq2seq	VT-MDM(Ours)
MAE (e-3°)	Avg	12.02	8.413	3.852	**2.376**
	Min	11.88	7.927	3.622	2.171
	Max	12.19	8.928	4.091	2.580
RMSE (e-3°)	Avg	17.05	11.17	5.258	**3.206**
	Min	16.87	10.52	4.947	2.927
	Max	17.27	11.81	5.612	3.506
$\mathcal{D}_{hav}$ (km)	Avg	2.223	0.911	0.681	**0.392**
	Min	2.120	0.841	0.641	0.358
	Max	2.250	0.973	0.724	0.423

longer predictions. Similarly, the prediction errors of LSTM increase as the sliding window moves backward, causing a substantial decline in its performance for longer predictions. Regarding the prediction of the future 6 min vessel trajectory, both the seq2seq model and VT-MDM demonstrate comparable performance. However, for longer predictions of 12 min, VT-MDM significantly outperforms the seq2seq model in terms of accuracy, revealing the decrease in accuracy of the seq2seq model as the prediction time increases. In contrast, the prediction accuracy of VT-MDM remains more stable because it captures the actual distribution of vessel trajectory modes.

5.3. Multimodal trajectory prediction

In this subsection, we present the multimodal trajectory prediction results of the proposed method to illustrate the impact of modifying the latent vectors on the prediction.

As explained in Section 4.3.1, we input the vessel's historical trajectory and a randomly chosen latent vector from a Gaussian distribution into the generator to generate the anticipated trajectory. The outcomes of the final predictions are shown in Fig. 5. Figs. 5(a), 5(c) and 5(e) corresponds to the case where the prediction duration is 6 min, whereas Figs. 5(b), 5(d) and 5(f) corresponds to the case with a prediction duration of 12 min. The red lines in the figure represent the historical trajectories of the vessels, while the other colors indicate the predicted trajectories. Figs. 5(a) and 5(b) demonstrate the prediction results for the vessel traveling in different directions at future moments. In the case of Fig. 5(b), the yellow line depicts the predicted outcome of the vessel continuing to travel in the original direction, whereas the orange line indicates the predicted outcome of the vessel changing its direction of travel. Although the two predictions maintain parallel travel directions after some time, the straight-line distance between them differs by 531.91 m. Figs. 5(c) and 5(d) present the predicted results of the vessel traveling at different speeds at future moments. The predicted directions of the same vessel almost overlap; however, there is a deviation of several hundred meters in the final arrival position of the two different predictions, corresponding to different travel speeds. Figs. 5(e) and 5(f) demonstrate the prediction results obtained by inputting five different latent vectors. Different latent vectors correspond to different predicted trajectories, as presented in the figure. The results show that by constructing an invertible mapping between the latent vectors and the predicted trajectories, the model's multimodal prediction capability can be effectively enhanced.

Table 3
Percentage of the true vessel position inside the corresponding σ-contour.

Prediction steps	6 min	12 min
1σ	85%	50%
2σ	89%	59%
3σ	95%	90%

5.4. Prediction smoothness analysis

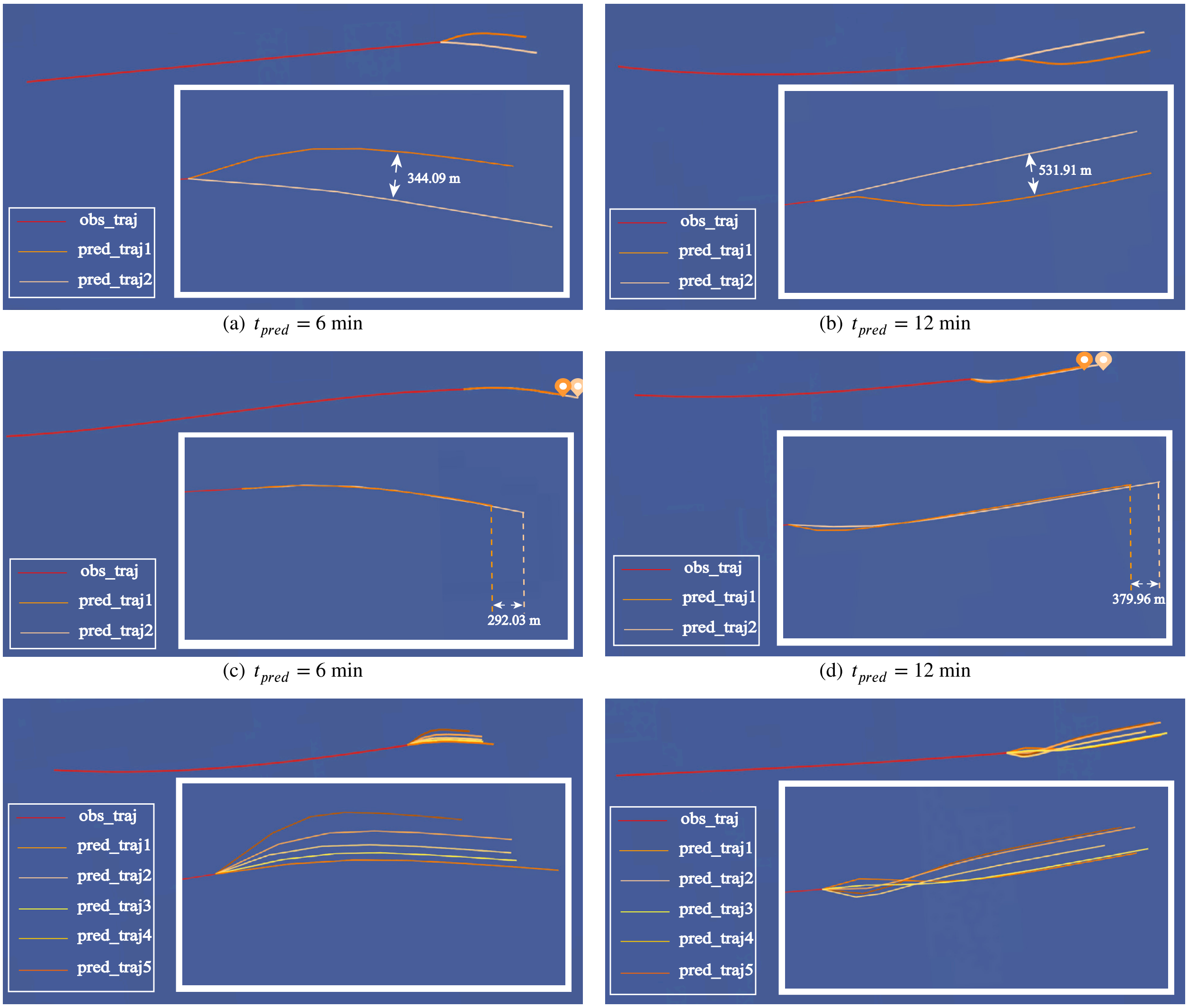
In this subsection, we illustrate the smoothing effect of the proposed method by visualizing the predicted trajectories obtained after interpolation using various modes.

To ensure a gradual transition between the different prediction modes of the vessel, we compact the latent space using global regularization and local interpolation, as detailed in Section 4.2. We perform linear interpolation for two randomly sampled latent vectors, $z1$ and $z2$, where we choose α from the interval $[0, 1]$ with a step size of 0.2. The resulting mixture vector $Mix(\alpha, z_1, z_2)$ is then fed into the generator to generate the predicted trajectory. Fig. 6 presents the different predicted trajectories generated by uniformly varying α. For illustrative purposes, we only display the 12 min predicted duration results. The yellow and orange lines represent the predicted trajectories generated by $z1$ and $z2$, respectively, while the green line represents the predicted trajectory generated by the mixed vector $Mix(\alpha, z_1, z_2)$. Specifically, Fig. 6(a) depicts the scenario with a gradual change in the vessel direction of travel, and Fig. 6(b) depicts the scenario with a gradual change in travel speed. As shown in the figure, the predicted trajectory generated by the mixed vector $Mix(\alpha, z_1, z_2)$ contains both $z1$ and $z2$ features with a given scale, and the predicted trajectory varies uniformly with α. These results demonstrate that constructing a smooth mapping between the latent vector and the predicted trajectory can achieve a gradual transition from one prediction mode to the next.

5.5. Uncertainty analysis

In order to further evaluate the prediction risk of the proposed method, we perform an analysis of prediction uncertainty.

The primary cause of uncertainty in the predicted locations for each time interval is the uncertainty that is associated with the latent vectors. To quantify the predicted risk, it is suggested to run a Monte Carlo simulation (Raychaudhuri, 2008; Capobianco et al., 2021a). As presented in Fig. 7, the prediction results are given for two prediction durations of 6 and 12 min. The red dashed line represents the historical trajectory of the vessel, the green dashed line represents the actual future trajectory, and the blue dashed line corresponds to the mean of the trajectories obtained from 500 predictions generated by our model. Additionally, we provide the 1σ, 2σ, and 3σ contours for the final predicted position. While the model can accurately capture the actual future trajectory of the vessel in the majority of cases, deviations in the predictions do arise. Consequently, statistical analysis was performed on the overall predictions to evaluate the model's capacity to accurately forecast the future vessel position. For each trajectory, 500 predictions were performed, and the percentage of vessels whose ground truth values were within the regions bounded by the corresponding σ-contour was investigated. Table 3 presents the final results. It was discovered that the ability of the σ-contour to capture the true trajectory decreased as the prediction duration increased, which is consistent with the findings presented in Tables 1 and 2. In the worst-case scenario, the 3σ contour still captures 90% of the true future vessel trajectories.

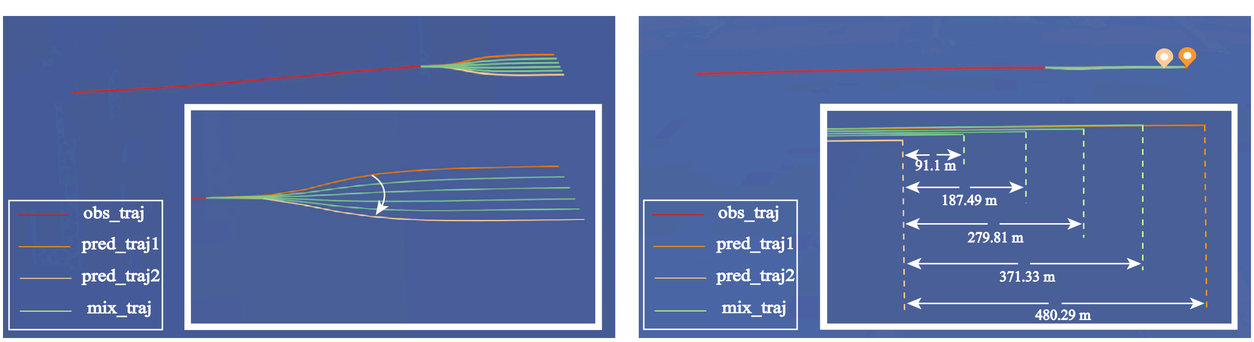


(a) $t_{pred} = 6$ min (b) $t_{pred} = 12$ min

(c) $t_{pred} = 6$ min (d) $t_{pred} = 12$ min

(e) $t_{pred} = 6$ min (f) $t_{pred} = 12$ min

Fig. 5. Visualization of prediction results.

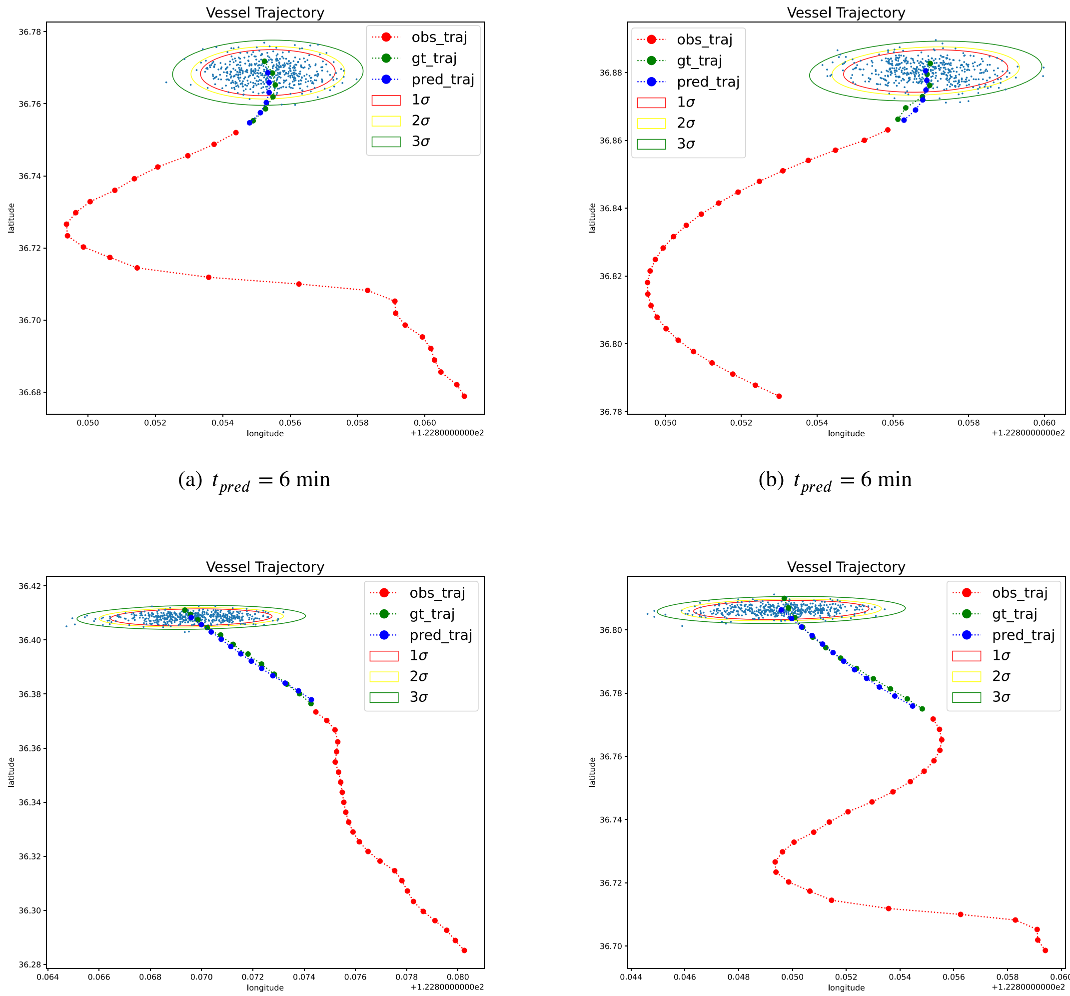


(a) $t_{pred} = 12$ min (b) $t_{pred} = 12$ min

Fig. 6. Visualization of prediction results.

5.6. Extensibility analysis

In our setting, the generator uses the most basic encoder–decoder structure. This structure, although simple, has been shown to handle variable-length input and output sequence prediction tasks efficiently. Nevertheless, recent research has introduced new model architectures, including the encoder–decoder structure with the attention mechanism (Capobianco et al., 2021b). The attention mechanism can establish a relationship between the hidden states of the encoder and decoder, thus improving the expressiveness of the model. To further explore the extensibility of the proposed network architecture, we conduct an illustrative experiment. Specifically, we replace the original



(a) $t_{pred} = 6$ min (b) $t_{pred} = 6$ min

(c) $t_{pred} = 12$ min (d) $t_{pred} = 12$ min

Fig. 7. Predicted trajectory results with uncertainty ellipse.

Table 4
Comparison of prediction results for different network architectures at $t_{pred} = 6$ min.

	Method	VT-MDM	VT-MDM(attention)
MAE (e-3°)	Avg	1.963	**1.252**
	Min	1.874	1.189
	Max	2.061	1.324
RMSE (e-3°)	Avg	2.521	**1.692**
	Min	2.405	1.603
	Max	2.640	1.805
$\mathcal{D}_{hav}$ (km)	Avg	0.310	**0.200**
	Min	0.296	0.190
	Max	0.325	0.212

Table 5
Comparison of prediction results for different network architectures at $t_{pred} = 12$ min.

	Method	VT-MDM	VT-MDM(attention)
MAE(e-3°)	Avg	2.376	**2.263**
	Min	2.171	2.144
	Max	2.580	2.384
RMSE(e-3°)	Avg	3.206	**3.077**
	Min	2.927	2.912
	Max	3.506	3.243
$\mathcal{D}_{hav}$(km)	Avg	0.392	**0.357**
	Min	0.358	0.338
	Max	0.423	0.377

network architecture of the generator using an encoder–decoder structure with an attention mechanism. And the final results are reported in Tables 4 and 5.

The results demonstrate that the prediction performance of the proposed method can be effectively improved by using a more advanced network architecture. Indeed, our proposed method is not limited to any specific network architecture and can be implemented with most network architectures designed for sequence prediction. This implies that our proposed model can be integrated with other model structures to improve the generative power further. This provides good insight for future research to explore various new model structures and techniques and integrate them into our method for better performance.

6. Conclusions

In this paper, we study the multimodal vessel trajectory prediction problem from a novel perspective, i.e., utilizing additional hiding regimes to characterize the complex trajectory modes independently. To this end, we propose VT-MDM, a new method for vessel trajectory prediction based on historical AIS data that performs better in multimodal prediction by modeling the distribution of trajectory modes. In more detail, we use adversarial learning to enable a Gaussian distribution to capture the vessel trajectory modes and create an invertible and smooth mapping between the latent vectors of modes and the predicted trajectories. To the best of our knowledge, this is the first method to achieve the truly and gradually multimodal trajectory prediction, providing richer information for improving maritime safety. Extensive experiments demonstrate that VT-MDM can achieve promising multimodal vessel trajectory predictions with high accuracy, providing a general framework for multimodal trajectory prediction of vessels.

In this study, we only consider the uncertainty of short-term vessel behavior. A major reason is that short-term multimodal trajectory prediction is more realistic for preventing close-range encounter situations from arising. A future direction of our work is to systematically evaluate multimodal vessel trajectory prediction methods, including our method, in long-term settings. Furthermore, while we employ a simple seq2seq model as our base prediction architecture, investigating the potential of more complex network architectures to improve prediction performance is an important area of further research.

CRediT authorship contribution statement

Siyang Guo: Conceptualization, Methodology, Software, Writing – original draft, Visualization. **Hui Zhang:** Supervision, Writing – review & editing. **Yaming Guo:** Conceptualization, Writing – review & editing.

Declaration of competing interest

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

Data availability

The data that has been used is confidential.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 61701263). The authors also appreciate the valuable comments and suggestions from Reviewers and Editor.

References

Arjovsky, M., Bottou, L., 2017. Towards principled methods for training generative adversarial networks. arXiv preprint arXiv:1701.04862.
Assaf, M.H., Groza, V., Petriu, E.M., 2020. The use of Kalman filter techniques for ship track estimation. WSEAS Trans. Syst. 19, 7–13.
Bengio, Y., Courville, A., Vincent, P., 2013. Representation learning: A review and new perspectives. IEEE Trans. Pattern Anal. Mach. Intell. 35 (8), 1798–1828.
Capobianco, S., Forti, N., Millefiori, L.M., Braca, P., Willett, P., 2021a. Uncertainty-aware recurrent encoder-decoder networks for vessel trajectory prediction. In: 2021 IEEE 24th International Conference on Information Fusion. FUSION, IEEE, pp. 1–5.
Capobianco, S., Forti, N., Millefiori, L.M., Braca, P., Willett, P., 2022. Recurrent encoder-decoder networks for vessel trajectory prediction with uncertainty estimation. IEEE Trans. Aerosp. Electron. Syst..
Capobianco, S., Millefiori, L.M., Forti, N., Braca, P., Willett, P., 2021b. Deep learning methods for vessel trajectory prediction based on recurrent neural networks. IEEE Trans. Aerosp. Electron. Syst. 57 (6), 4329–4346.
Chen, J., Chen, H., Zhao, Y., Li, X., 2022. FB-BiGRU: A deep learning model for AIS-based vessel trajectory curve fitting and analysis. Ocean Eng. 266, 112898.
Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y., 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.
Chroni, D., Liu, S., Plessas, T., Papanikolaou, A., 2015. Simulation of the maneuvering behavior of ships under the influence of environmental forces. Towards Green Marine Technology and Transport 111–119.
Gao, D.-w., Zhu, Y.-s., Zhang, J.-f., He, Y.-k., Yan, K., Yan, B.-r., 2021. A novel MP-LSTM method for ship trajectory prediction based on AIS data. Ocean Eng. 228, 108956.
Ghazvininejad, M., Brockett, C., Chang, M.-W., Dolan, B., Gao, J., Yih, W.-t., Galley, M., 2018. A knowledge-grounded neural conversation model. In: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 32. (1).
Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2020. Generative adversarial networks. Commun. ACM 63 (11), 139–144.
Goudossis, A., Katsikas, S.K., 2019. Towards a secure automatic identification system (AIS). J. Mar. Sci. Technol. 24, 410–423.
Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S., Alahi, A., 2018. Social gan: Socially acceptable trajectories with generative adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 2255–2264.
Harati-Mokhtari, A., Wall, A., Brooks, P., Wang, J., 2007. Automatic identification system (AIS): data reliability and human error implications. J. Navig. 60 (3), 373–389.
Hexeberg, S., Flåten, A.L., Brekke, E.F., et al., 2017. AIS-based vessel trajectory prediction. In: 2017 20th International Conference on Information Fusion (Fusion). IEEE, pp. 1–8.
Huang, Z., Mo, X., Lv, C., 2022. Multi-modal motion prediction with transformer-based neural network for autonomous driving. In: 2022 International Conference on Robotics and Automation. ICRA, IEEE, pp. 2605–2611.
Kosaraju, V., Sadeghian, A., Martín-Martín, R., Reid, I., Rezatofighi, H., Savarese, S., 2019. Social-bigat: Multimodal trajectory forecasting using bicycle-gan and graph attention networks. Adv. Neural Inf. Process. Syst. 32.
Liu, Y., Sangineto, E., Chen, Y., Bao, L., Zhang, H., Sebe, N., Lepri, B., De Nadai, M., 2022. Smooth image-to-image translations with latent space interpolations. arXiv preprint arXiv:2210.00841.
Liu, Y., Sangineto, E., Chen, Y., Bao, L., Zhang, H., Sebe, N., Lepri, B., Wang, W., De Nadai, M., 2021. Smoothing the disentangled latent style space for unsupervised image-to-image translation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10785–10794.
Ma, S., Sun, X., Lin, J., Wang, H., 2018. Autoencoder as assistant supervisor: Improving text representation for Chinese social media text summarization. arXiv preprint arXiv:1805.04869.
Mao, X., Li, Q., Xie, H., Lau, R., Wang, Z., Smolley, S., 2017. Least squares generative adversarial networks. In: Proceedings of the IEEE international conference on computer vision, pages=2794–2802,.
Mao, S., Tu, E., Zhang, G., Rachmawati, L., Rajabally, E., Huang, G.-B., 2018. An automatic identification system (AIS) database for maritime trajectory prediction and data mining. In: Proceedings of ELM-2016. Springer, pp. 241–257.
Mirza, M., Osindero, S., 2014. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784.
Murray, B., Perera, L.P., 2020. A dual linear autoencoder approach for vessel trajectory prediction using historical AIS data. Ocean Eng. 209, 107478.
Murray, B., Perera, L.P., 2021. An AIS-based deep learning framework for regional ship behavior prediction. Reliab. Eng. Syst. Saf. 215, 107819.
Nguyen, D., Fablet, R., 2021. TrAISformer-A generative transformer for AIS trajectory prediction. arXiv preprint arXiv:2109.03958.
Nguyen, D., Vadaine, R., Hajduch, G., Garello, R., Fablet, R., 2018. A multi-task deep learning architecture for maritime surveillance using AIS data streams. In: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics. DSAA, IEEE, pp. 331–340.
Nguyen, D., Vadaine, R., Hajduch, G., Garello, R., Fablet, R., 2021. GeoTrackNet—A maritime anomaly detector using probabilistic neural network representation of AIS tracks and a contrario detection. IEEE Trans. Intell. Transp. Syst. 23 (6), 5655–5667.
Ozbas, B., 2013. Safety risk analysis of maritime transportation: review of the literature. Transp. Res. Rec. 2326 (1), 32–38.
Park, J., Jeong, J., Park, Y., 2021. Ship trajectory prediction based on bi-LSTM using spectral-clustered AIS data. J. Mar. Sci. Eng. 9 (9), 1037.
Perera, L.P., Oliveira, P., Soares, C.G., 2012. Maritime traffic monitoring based on vessel detection, tracking, state estimation, and trajectory prediction. IEEE Trans. Intell. Transp. Syst. 13 (3), 1188–1200.
Perera, L.P., Soares, C.G., et al., 2010. Ocean vessel trajectory estimation and prediction based on extended Kalman filter. In: The Second International Conference on Adaptive and Self-Adaptive Systems and Applications. Citeseer, pp. 14–20.
Qian, L., Zheng, Y., Li, L., Ma, Y., Zhou, C., Zhang, D., 2022. A new method of inland water ship trajectory prediction based on long short-term memory network optimized by genetic algorithm. Appl. Sci. 12 (8), 4073.
Rajabi, A., Saryazdi, A.K., Belfkih, A., Duvallet, C., 2018. Towards smart port: an application of AIS data. In: 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS). IEEE, pp. 1414–1421.
Raychaudhuri, S., 2008. Introduction to monte carlo simulation. In: 2008 Winter Simulation Conference. IEEE, pp. 91–100.
Robards, M., Silber, G., Adams, J., Arroyo, J., Lorenzini, D., Schwehr, K., Amos, J., 2016. Conservation science and policy applications of the marine vessel automatic identification system (AIS)—A review. Bull. Mar. Sci. 92 (1), 75–103.
Rong, H., Teixeira, A., Soares, C.G., 2019. Ship trajectory uncertainty prediction based on a Gaussian process model. Ocean Eng. 182, 499–511.
Sørensen, K.A., Heiselberg, P., Heiselberg, H., 2022. Probabilistic maritime trajectory prediction in complex scenarios using deep learning. Sensors 22 (5), 2058.
Suo, Y., Chen, W., Claramunt, C., Yang, S., 2020. A ship trajectory prediction framework based on a recurrent neural network. Sensors 20 (18), 5133.

Sutskever, I., Vinyals, O., Le, Q.V., 2014. Sequence to sequence learning with neural networks. Adv. Neural Inf. Process. Syst. 27.

Sutulo, S., Moreira, L., Soares, C.G., 2002. Mathematical models for ship path prediction in manoeuvring simulation systems. Ocean Eng. 29 (1), 1–19.

Verma, V., Lamb, A., Beckham, C., Najafi, A., Mitliagkas, I., Lopez-Paz, D., Bengio, Y., 2019. Manifold mixup: Better representations by interpolating hidden states. In: International Conference on Machine Learning. PMLR, pp. 6438–6447.

Yang, D., Wu, L., Wang, S., Jia, H., Li, K.X., 2019. How big data enriches maritime research–A critical review of automatic identification system (AIS) data applications. Transp. Rev. 39 (6), 755–773.

Zhang, J.-X., Ling, Z.-H., Jiang, Y., Liu, L.-J., Liang, C., Dai, L.-R., 2019. Improving sequence-to-sequence voice conversion by adding text-supervision. In: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing. ICASSP, IEEE, pp. 6785–6789.

Zhang, L., Meng, Q., Xiao, Z., Fu, X., 2018. A novel ship trajectory reconstruction approach using AIS data. Ocean Eng. 159, 165–174.

Zhu, J.-Y., Zhang, R., Pathak, D., Darrell, T., Efros, A.A., Wang, O., Shechtman, E., 2017. Toward multimodal image-to-image translation. Adv. Neural Inf. Process. Syst. 30.