A Semi-Supervised Deep Learning Model for Ship Encounter Situation Classification

Xiang Chen[a], Yuanchang Liu[b,*], Kamalasudhan Achuthan[a,*], Xinyu Zhang[c] and Jinhai Chen[d]

[a] *Department of Civil, Environmental and Geomatic Engineering, University College London, Chadwick Building, London WC1E 6BT, UK*
[b] *Department of Mechanical Engineering, University College London, Torrington Place, London WC1E 7JE, UK*
[c] *Key Laboratory of Maritime Dynamic Simulation and Control of Ministry of Transportation, Dalian Maritime University, Dalian 116026, China*
[d] *Navigation College of Jimei University, National-local Joint Engineering Research Centre for Marine Navigation Aids Services, Xiamen 361021, China*

ARTICLE INFO

Keywords:
Automatic Identification System (AIS)
Semi-Supervised Learning
Deep Learning
Convolutional Neural Network
Encoder-Decoder
Trajectory Data
Encounter Situation Classification

ABSTRACT

Maritime safety is an important issue for global shipping industries. Currently, most of collision accidents at sea are caused by the misjudgement of the ship's operators. The deployment of maritime autonomous surface ships (MASS) can greatly reduce ships' reliance on human operators by using an automated intelligent collision avoidance system to replace human decision-making. To successfully develop such a system, the capability of autonomously identifying other ships and evaluating their associated encountering situation is of paramount importance. In this paper, we aim to identify ships' encounter situation modes using deep learning methods based upon the Automatic Identification System (AIS) data. First, a segmentation process is developed to divide each ship's AIS data into different segments that contain only one encounter situation mode. This is different to the majority of studies that have proposed encounter situation mode classification using hand-crafted features, which may not reflect the actual ship's movement states. Furthermore, a number of present classification tasks are conducted using substantial labelled AIS data followed by a supervised training paradigm, which is not applicable to our dataset as it contains a large number of unlabelled AIS data. Therefore, a method called Semi-Supervised Convolutional Encoder-Decoder Network (SCEDN) for ship encounter situation classification based on AIS data is proposed. The structure of the network is not only able to automatically extract features from AIS segments but also share training parameters for the unlabelled data. The SCEDN uses an encoder-decoder convolutional structure with four channels for each segment (distance, speed, Time to the Closed Point of Approach (TCPA) and Distance to the Closed Point of Approach (DCPA)) been developed. The performance of the SCEDN model are evaluated by comparing to several baselines with the experimental results demonstrating a higher accuracy can be achieved by our proposed model.

1. Introduction

Safety is always a significant concern in maritime industry. Based on the survey from the Lloyd's report (Hassel et al., 2011), 67% of waterborne transportation accidents are caused by collisions between ships. Studies on accidents show that the ship crew is the significant factor when it comes to maritime safety since approximately 80-90% of maritime accidents are caused by human errors (Berg et al., 2013). Therefore, it is important to carry out research on ship intelligence development to replace the human decision (Parsons et al., 2008). In recent years, academics and industries have largely aimed at developing and building autonomous ships by following the initiatives advocated in land-based transportation (Ramos et al., 2019). The concept of Maritime Autonomous Surface Ships (MASS) has been proposed by the International Maritime Organisation in 2017 and one of the motivations for the MASS is to use a collection of sensors to replace the human decision-making and to large extend avoid collision accidents. In order to realise a MASS collision avoidance system, it is important to develop an artificial intelligence system to help precept the relative encounter situations between ships (Zhang et al., 2015).

Recent developments of MASS bring a lot of requirements for the intelligence, especially each ship's collision avoidance. Traditional collision avoidance is based on the operator's decision with the help of sensor such as Radar (Bovcon and Kristan, 2020). For the MASS, one of main challenges is the timely detection and avoidance of close-range obstacles. With the development of sensor technologies, it becomes promising to use the sensor fusion information to replace the human operator's subjective judgements. Currently, there are many sensors on board to provide information and one of the widely used devices for collision avoidance is the Automatic Identification System (AIS). The AIS is an automatic tracking system that uses transceivers on ships to broad each ship's unique identification, position, course, and speed. For human operator, these information can be used to calculate the Close Point of Approach (CPA) (shown in Figure 1) to help them make a collision avoidance action.

Figure 1 illustrates one of representative scenarios for the collision avoidance. Two ships denoted as V_1 and V_2 have a distance on the left side of Figure 1 at time $t+1$. If two ships keep the directions and speed, they will eventually lead to a collision after time $t + 3$. The red star in Figure 1 represents the closest point of approach, which is the latest point to take action. If actions (changing course or reducing speed) are taken later than CPA, the two ships will inevitably col-

*Corresponding author

xiang.chen.17@ucl.ac.uk (X. Chen); yuanchang.liu@ucl.ac.uk (Y. Liu); k.achuthan@ucl.ac.uk (K. Achuthan)

ORCID(s): 0000-0002-1084-9302 (X. Chen)

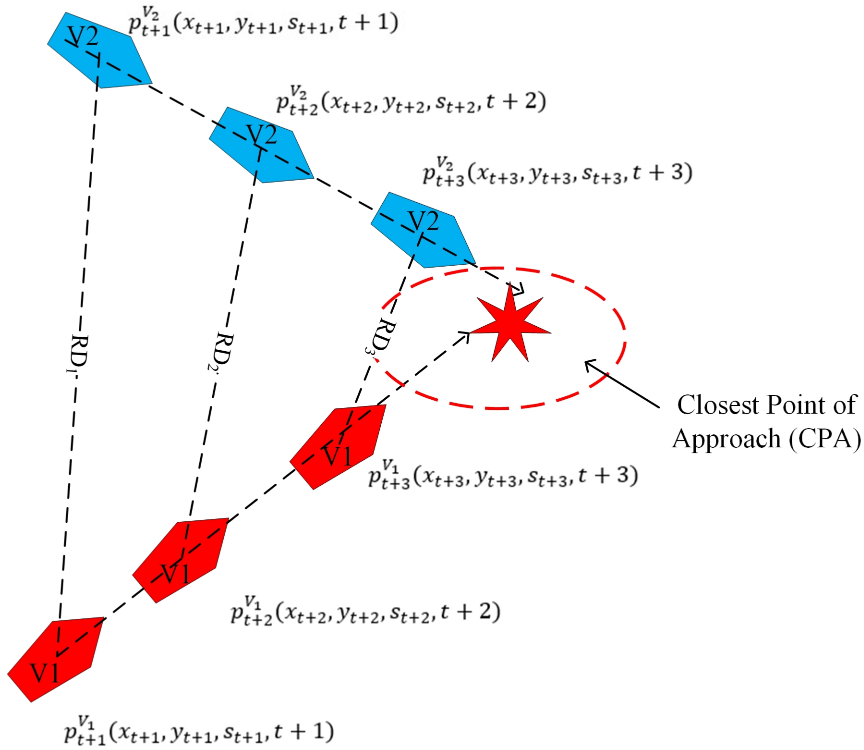

Figure 1: CPA Illustration: The two ships V_1 and V_2 keep the current directions in time from $t+1$ to $t+3$. If the two ships are continuing the direction after $t+3$, the two ships will collide. The latest decision-making point represented by the star is called the Closest Point of Approach (CPA).

lide. The CPA can be calculated by the time or the distance separately known as $DCPA$ and $TCPA$.

As shown in Figure 2, there are four basic different types of ship encounter situations prescribed by the International Regulations for Preventing Collisions at sea 1972 (COLREGs) (Mankabady, 1986). For crossing situation in Figure 2, if ship V_2 is on the right hand side of ship V_1, ship V_1 should give way for ship V_2 and vice versa for the crossing 2 scenario. If ship V_1 and ship V_2 has the opposite direction, each ship has the same responsibility to make a way for safety reasons. The last scenario is the overtaking situation, ship V_1 is behind ship V_2 and plans to surpass ship V_2 that is playing a give-way role in this scenario. Based on the COLREGs rule, it needs to classify the ship encounter situation into different types for the development of intelligent collision avoidance system. When it comes to the development of the learning-based collision avoidance model that is compliant with COLREGs, it is evident that a sufficient amount of data containing four encounter situations should be provided for training purposes (Naeem et al., 2012).

However, a bulk of currently available AIS data are only used for the normal navigation without explicitly revealing more complex navigation details including the encountering situations. Currently, there are lots of publications on the collision avoidance based on various methods such as velocity obstacle (Yuan et al., 2021), decision-making (Zheng et al.; Mizythras et al., 2021; Gao and Shi, 2020). These methods are mainly on the global situational awareness and they do not contain any two ship's encounter situation strategy. In our previous publication in Chen et al. (2020), a supervised learning technique has been proposed to classify AIS data into three basic categories, i.e. normal navigation, static state and manoeuvring state. The manoeuvring state (equal to encounter situation) is the data that we require for collision avoidance, but it does not indicate properly which types of encounter situation a ship is dealing with. Moreover, the AIS data only records the ship movement state with most of them incapable of specifying which category it belongs to. Therefore, it becomes necessary to design a new model to classify the ship encounter situations based on the AIS data. Under the foundation of the work in Schwenker and Trentin (2014), Kostopoulos et al. (2018), Liu et al. (2018), etc., in this paper we propose a new Semi-Supervised Convolutional Encoder Decoder Network for ship encounter situation classification. The main contributions of this work can be summarised as follows:

- *Designing an efficient representation for AIS data.* In order to make a good performance for the encounter situation classification, a new method capable of augmenting conventional AIS segments to an efficient representation that incorporates encounter situation has been proposed. The new representation can enable the use of a convolutional neural network (CNN) architectures to extract high-level features from AIS.

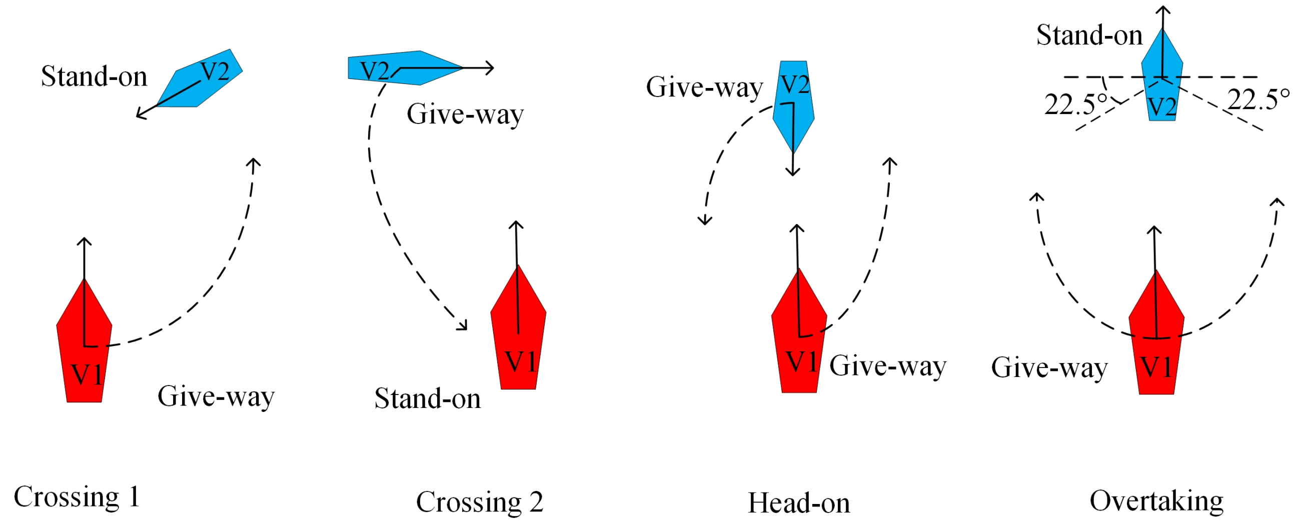

Figure 2: Encounter situations: Based on the COLREGS rule from IMO, the ship-ship encounters are categorised into three types (crossing, head-on and overtaking). Using the different relative position for the ship V_1 and ship V_2, the crossing situation can be simplified into two groups ('give-way' or 'stand-on').

- *Developing a novel semi-supervised convolutional encoder decoder architecture.* A deep semi-supervised model has been proposed to balance unlabelled and labelled AIS data for classifying encounter situations. A CNN classifier dealing with supervised training using labelled is embedded with a encoder-decoder structure which is responsible for unlabelled data training.

- *Building an efficient training scheme for finding optimal hyperparameters.* At present, most of AIS data are unlabelled making the labelled AIS data for encounter situation only account for a small proportion. We therefore propose a new small-sample training strategy, where a new training scheme first uses a small amount of labelled AIS data as a warm-up and then shares the training weights for unlabelled AIS data for large-scale training. A joint training is then employed for boosting the accuracy and reducing the whole training loss.

- *Conducting a series of extensive experiments for performance evaluation.* To compare the results of the proposed model, classic methods including the supervised decision tree, support vector machine and k-nearest neighbour classification methods are used and implemented.

The rest of this paper is organised as follows. Section 2 provides the related work on trajectories classification and deep semi-supervised learning. After that, Section 3 provides the preliminaries for the encounter situation classification task and Section 4 shows the details of our model framework. The experiment and results are shown in Section 5. Finally, the conclusion is presented in Section 6.

2. Related Work

Until now, a number of studies have employed various data sources (e.g., AIS, GPS or sonar data) to classify ships' movements (Pallotta et al., 2013; Wijaya and Nakamura, 2013; Yuan et al., 2019; El Mekkaoui et al., 2020; Gao et al., 2021). In this section, we specifically review the studies that have used the AIS data for collision avoidance model development. In order to expand the vision, we also review the research on the movement trajectories classification in other transportation sectors such as road transport. After reviewing the movement classification, we will also discuss various semi-supervised deep learning architectures that have been used in different applications related to the transportation studies.

2.1. Ship Movement Classification based on AIS data

Feature extraction and classification are two main tasks for the studies on ships AIS data (Sheng et al., 2018). A lot of classification tasks are often conducted by using traditional supervised learning algorithms such as support vector machines (Lang et al., 2018), decision tree (Krüger, 2018; Chen et al., 2018), k-nearest neighbour algorithm (Damastuti et al., 2019), random forest (Zhang et al., 2020a) etc. and the feature extraction is a significant research question due to various application aims.

There are two mainstream studies for the feature extraction using the AIS data with one of them being the fishing movement activity detection (Jiang et al., 2016; Yang et al., 2019). De Souza et al. (2016). These works use the satellite AIS data to research the fishing fleets activities by employing a hidden Markov model to train a model in a large Satellite AIS data. Jiang et al. (2016) uses the autoencoder network to detect the fishing activities and in their model, a fine-tune feed forward network is introduced. The results show that the algorithm is able to detect 46427 AIS data points with 77.96% of them classified for fishing movement AIS squeeze. Compared to SVM and decision tree, the deep neural network has a better performance than the traditional classification methods. Sánchez Pedroche et al. (2020) studies the AIS data features and it extracts 5 general features (speed, distance, course variation, speed variation and time gaps) to build a binary classification (fishing and non-fishing) model. The experiment shows that the accuracy reaches to 82% after 10-folds cross validation. Based on above studies, the feature extraction methods could be summarised into two categories:

- Using the hand-crafted features for different fishing movements detection, such as speed, duration, course variation and distance;
- Using a neural network such as auto-encoder architecture or the hidden Markov model for automatically detecting fishing activities.

The results from the studies show that the hand-crafted features may lose a large number of hidden high-level features. Also, with the increasing of hand-crafted features, the computing complexity also grows exponentially. In addition, one drawback for the hand-crafted feature building is the difficulty in labelling (Svanberg et al., 2019). Therefore, a completely hand-crafted classification method is difficult for a large scale AIS dataset.

Another study for the feature extraction using AIS data is the intelligent maritime application developments (Tu et al., 2017), especially the ship automatic collision avoidance decision making system. The AIS brings an enriched kinematic information and the collected AIS data contain a wealth of useful information for maritime safety (Tu et al., 2017). Liu et al. (2020) proposes a data-driven method to estimate the navigable capacity of busy waterways for ship's encounter. They use the K-means clustering algorithm to classify the structural characteristics of the traffic flow and then based on the filed study to verify the results. Sheng et al. (2018) supposes that the historical AIS trajectory data could help to boost the maritime efficiency and help classify the unknown types of ships. It first classifies the ship movement into three basic types and continues to classify the ship into fishing or cargo ship types and the work mainly aims at behaviour pattern mining and outlier behaviour detection. Also, Chen et al. (2020) uses the convolutional neural network to classify the ship movement into three basic types (static, normal navigation and manoeuvring) and they use the CNN capability for image classification to transfer the AIS movement data into different types of movement images for different types of ships. Gao et al. (2018) develops an online realtime ship behaviour prediction model based on bidirectional long short-term memory recurrent neural network for realtime AIS data. It helps ships to achieve an intelligent collision avoidance and ship route planning. Similar studies based on the long short-term memory recurrent neural network are anomaly detection (Ginoulhac et al., 2019) and collision avoidance (Ginoulhac et al., 2019; Shi and Liu, 2020). The above researches for ship collision avoidance are on the single ship response. However, there are a number of applications such as multi-ASVs sailing at sea, where the collision avoidance is different from the single ship scenario. For example, Ma et al. (2021) resolved the collision avoidance issues for multiple ships according to ships' manoeuvrability and encounter situation division. An ad-hoc network was developed to construct a negotiation communication between vessels, which ensures an effective and reliable avoidance coordination for swarm vessels.

Based on above reviews for ship movement detection using AIS data, we can summarise main research gaps as:

- A number of AIS data classification algorithms are relying on the hand-crafted features, which require a lot of features to be labelled by human;
- Using the neural network automatically to study the AIS data distribution and features is becoming a research trend due to the capacity imposed by the end-to-end study learning. However, it also requires a large amount of accurate labelled data to train neural networks. If the training dataset is relatively small or the labelled data is not sufficient, the classification ability is impaired and the performance is degraded compared to the hand-crafted methods.

2.2. Semi-Supervised Deep Learning Architectures

The definition of semi-supervised learning is a learning paradigm concerned with the study of both labelled and unlabelled data (Zhu and Goldberg, 2009). Semi-supervised learning architectures based on deep learning algorithms have been used in many tasks, mainly in computer vision and natural language processing (Dabiri et al., 2019). Currently, the semi-supervised learning methods are widely used in the road transportation for the origin-destination prediction (Bachir et al., 2019), pedestrian behaviour prediction (Radu et al., 2014) and transportation mode classification based on GPS data (Leodolter et al., 2017; Dabiri et al., 2019). Due to increased human activities and an extensive use of mobile phone today, a lot of GPS and sensor data can be generated. However, most of the data are not well labelled or detailed, as well as cannot be directly used to train a supervised deep learning model. Therefore, one of the common characteristics for these studies is to use the semi-supervised learning architecture. The training paradigm on the semi-supervised architecture falls into two research directions:

- Two-step training process, in which the semi-supervised network is first trained in the unsupervised learning method as a warm-up step, after that the supervised learning process is a fine tuning process using the labelled data.
- Joint training process, in which both the supervised and unsupervised learning process are simultaneously trained.

However, the semi-supervised learning architecture has one drawback which is the unsupervised data are commonly in a larger proportion as opposed to the supervised data. If the weight trained from the unlabelled data cannot be fitted with the supervised data, the fine tuning may not work, and the final classification accuracy can be worse than the handcrafted features and the loss value cannot converge. Hence, there are many effective methods introduced to solve the problem such as using balancing hyperparameters, hybrid loss function to achieve a balance between the supervised and unsupervised objective functions (Socher et al., 2011).

Based upon above discussions, a new semi-supervised deep learning model (or the SCEDN), has been proposed to specifically address the ship encounter situation classification. The proposed model adopts the semi-supervised learning architecture and is trained using a novel schedule for tuning a set of balanced hyperparameters to solve the imbalance label problem, which is also common to AIS data. Unlike the methods in the literature, the proposed training scheme addresses the issue by introducing two new balancing hyperparameters integrated in training process.

Note that although the semi-supervised learning architecture is widely used in the image semantic segmentation and image classification, a common feature for these tasks is that the input data is uniformly distributed with salient features. However, the AIS segments are all in one dimension rather than two, and contain no explicit information for ships encounter situation classification. The existing methods may not be well suited to the special requirement imposed by AIS segments. Therefore, this paper has proposed a new data augmentation method to make AIS data suitable for deep learning based encounter situation classification. The details of this method will be introduced in next section.

3. AIS Data Augmentation and Preprocessing

The raw AIS data is made up by a series of chronological ordered points but without an explicit indication on the ship encounter situations. When ships are in operation in open sea areas, three main encounter situations (head-on, overtaking and crossing) may occur, and a reliable prediction of encounter situation is critical to ensure the subsequent evasive actions taken by ships. In this paper, a dedicated AIS based data augmentation method has been proposed to incorporate temporal ship encounter situation information into AIS. More specifically, in order to facilitate an AIS data based training processes and avoid any training ambiguity, the encounter situation should be integrated into AIS in a way that each sub-divided segment AIS only contains one encounter situation. At the same time, the ship encounter situation is highly relevant to the ship movement characteristics. Therefore, the ship movement characteristics are needed to be considered. In this section, the definition and problem statements are firstly presented.

3.1. Definitions and Problem Statements

Definition 1 (Ship encounter situation data). *A ship's AIS data X is defined as a sequence of time-stamped points $p \in X$ and $X = \{p_1, p_2, ..., p_n\}$ and the encounter situation data can be labelled as X', $X' \in X$. Each ship encounter situation point p is a tuple of latitude, longitude, course, speed and time, $p = \{lat, lng, c, s, t\}$.*

Definition 2 (Ship encounter situation segment). *A ship's segment is a sub-division of a ship's encounter situation which is only contains one encounter situation mode $y \in Y$, where Y is a set of encounter situation modes (crossing, overtaking and head-on) and y is one encounter situation mode in each se segment. A ship's encounter situation segment represents as $se = \{p_1, p_2, ..., p_m\}$, where m is the total number of ship's encounter situation AIS points.*

Based on above definitions, the ship encounter situation classification problem is defined as follows:

Problem Description for Ship Encounter Situation data Classification *Given the training data $\{(X'_i, y_i)\}_{i=1}^n$, for n training data of se_i, the encounter situation classification problem is defined as training an optimal classifier for dividing the ship encounter situation into three parts (crossing, overtaking and head-on) based on its features X'.*

Since the ship encounter situation data has the homogeneous feature and each data contains an irrelevant context information characteristics, the trained model weight can be used to classify the unlabelled ship encounter situation data (Gao et al., 2018).

3.2. AIS Augmentation by Incorporating Ship Movement Features

AIS messages contain a ship's dynamic navigational data and is reasonably accurate as it transmits absolute navigational information of a ship from its on-board sensors such as the GPS and electronic compass (Robson and Qj, 2006). The message commonly contain static, dynamic and voyage information along with ship safety alert. Static information, such as the ship's call sign, name and its Maritime Mobile Service Identity (MMSI) is permanently stored in the on-board AIS transponder. Dynamic information contains the ship's absolute position, speed and course, along with the target ship's. Voyage related information includes ship's destination, hazardous cargo type, etc. is set up at the beginning of the voyage (Harati-Mokhtari et al., 2007).

The raw AIS data sequence is composed by a series of ship movement points. In order to incorporate ship encounter situation, processes including encounter situation calculation and AIS augmentation need to be carried out. First, in order to incorporate the ship encounter situation states into each AIS data sequence, it is necessary to divide each AIS movement data into different segments according to the change of ship movement states (relative distance, relative speed, TCPA or DCPA). These features can be computed based on the AIS data geographic coordinates and time stamps. The relative distance between two consecutive AIS points in a sequence is computed using Equation 1, in which the lat and lng is the abbreviation of latitude and longitude, respectively. The time interval can be simply computed using Equation 2. Based upon the relative distance and time interval, the kinematic parameter, i.e. the relative speed, can be

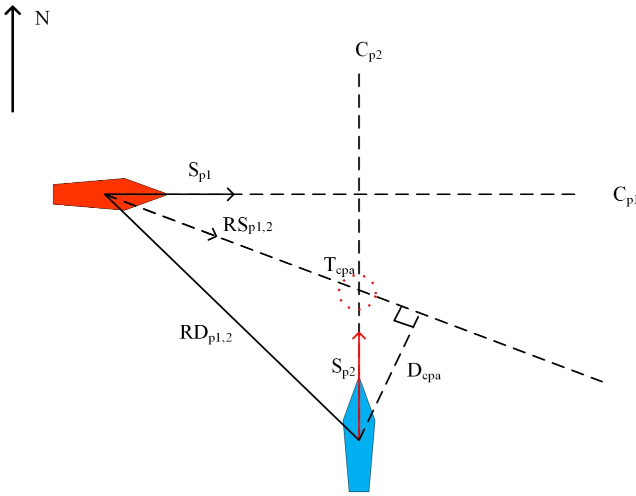

Figure 3: D_{cpa} and T_{cpa} calculation illustration

further obtained from Equation 3.

$$RD_{p_{1,2}} = \sqrt{(lat_{p_2} - lat_{p_1})^2 + (lng_{p_2} - lng_{p_1})^2} \tag{1}$$

$$\Delta t_{p_{1,2}} = p_2\{t\} - p_1\{t\} \tag{2}$$

$$RS_{p_{1,2}} = \frac{RD_{p_{1,2}}}{\Delta t_{p_{1,2}}} \tag{3}$$

In order to make a clear collision avoidance warning, the T_{cpa} and D_{cpa} are introduced. Figure 3 is an illustration to help understand the concepts of T_{cpa} andD_{cpa}, where S_{p1} and S_{p2} denote the speeds from p, C_{p1} and C_{p2} are the course of the ships. When we need to calculate the T_{cpa}, it firstly needs the relative speed from the x, y coordinate ($RS^x_{p_{1,2}}$, $RS^y_{p_{1,2}}$ respectively in Equation 4 and Equation 5) and the relative speed from Equation 3. The D_{cpa} could be easily obtained based on T_{cpa} result, shown in Equation 7.

$$RS^x_{p_{1,2}} = S_{p1} * sinC_{p1} - S_{p2} * sinC_{p2} \tag{4}$$

$$RS^y_{p_{1,2}} = S_{p1} * cosC_{p1} - S_{p2} * cosC_{p2} \tag{5}$$

$$T_{cpa} = \frac{(lng_{p2} - lng_{p1}) * RS^x_{p_{1,2}} + (lat_{p2} - lat_{p1}) * RS^y_{p_{1,2}}}{RS^2_{p_{1,2}}} \tag{6}$$

$$D_{cpa} = \sqrt{RD^2_{p1,2} - (RS^2_{p_{1,2}} * |T_{cpa}|^2)} \tag{7}$$

3.3. AIS Data Preprocessing

Apart from incorporating relevant encounter information into AIS, in order to train a ship encounter situation model, further data preprocessing is also required. Unlike the conventional image classifications, where the input information normally has uniform format, the augmented AIS sample processed by a neural network can potentially be different in sizes, i.e. a longer voyage will inevitably contains more AIS segments points. Therefore, before feeding the encounter situation segments into the training network, a further step needs to be undertaken to divide or fill each segment into the same size.

In this paper, a new AIS preprocessing method including a data filling and a division procedures has been proposed. The method and its overall process is shown in Figure 4. First, all AIS data based within one hour should be collected and counted together with the split time t and the total number of AIS points p_m for each ship in one hour. Based on the split time t, the number of split time interval can be calculated. Then, all the AIS points in each time interval are gathered by the MMSI. For each time interval AIS data, based on the equations described in Section 3.2, all ship's RD, RS, TCPA and DCPA are calculated and sorted according to the corresponding MMSI. In terms of data preprocessing, the method for dividing is based on the dataset size and specific navigational information, whereas for filling, the zero-padding method has been used. It should be noted that different AIS dataset can follow the same paradigm but may require different configurations to process the data.

With regards to the detailed data filling and division procedures, as shown in Figure 4, after calculating the four movement characteristics for each ship, the AIS points for every ship in the time interval are counted and if the current ship's number of AIS points is larger than a threshold p_m, a data division is needed to reduce the AIS points dimension into $(p_1, ..., p_m)$, as shown in the right side bottom of Figure 4 highlighted in green. On the contrary, if the current ship's number of AIS points is less than the threshold p_m, a data filling is required to increase the AIS points dimension into $(p_1, ..., p_m)$. The zero-padding method is shown in the right side upper of Figure 4 in red. Finally, the training data are collected and prepared for the next training.

After obtaining cleaned and cleared ship encounter situation segments, four features can be stacked into a tensor for each segment to enable a learning process. In the next section, a semi-supervised CNN based ship encounter learning process will be introduced.

4. Semi-Supervised Convolutional Encoder-Decoder Network

In this section, an efficient representation for each ship encounter situation from AIS segments is introduced with the whole architecture of our semi-supervised deep learning framework explained. Within the learning framework, a small sized supervised Convolutional Neural Network (CNN) is firstly introduced as a warm-up to learn the weights of neu-

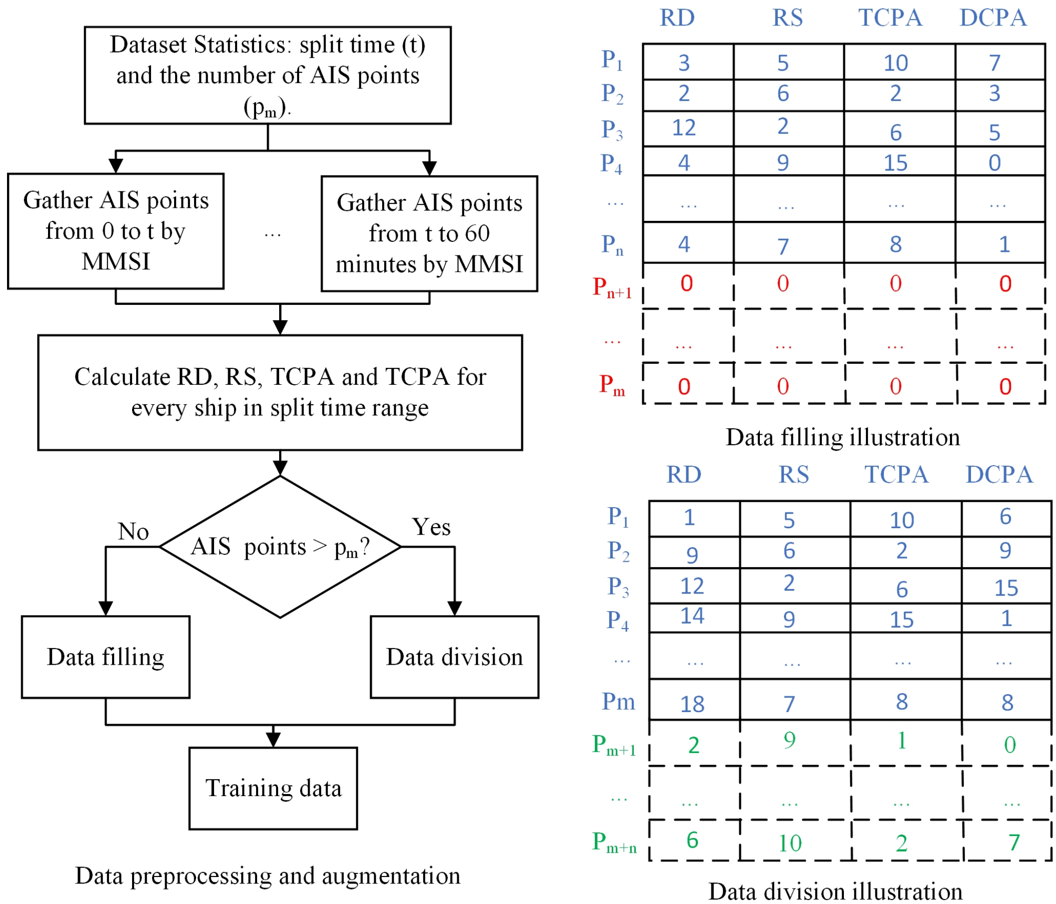

Figure 4: Data preprocessing and augmentation. The left side of the figure shows the overall work flow for the data preprocessing. The data augmentation is split into data filling and data division as shown in the right side of the figure.

ral networks. And then, a large amount of unsupervised data will be trained based on the shared weights. Finally, the joint training combining the labelled data and unlabelled data will play a significant fine-tuning role to improve the precision of the proposed classification algorithm.

4.1. Ship Encounter Situation Representation for each Segment

Since the main part of the proposed network is a convolutional network, it first needs to convert a ship encounter situation AIS data into a format that is not only compatible with CNN but also easy to represent a ship's movement characteristics. As discussed in Section 3, a ship encounter situation could be represented by the relative distance (RD), relative speed (RS), Time to Closest Point of Approach (T_{cpa}) and Distance to Closest Point of Approach (D_{cpa}). For each AIS segment, a sequence can be expressed by placing the corresponding value in a chronological order, in which the feature values are computed by the Equation 1, 2, 6 and 7. The sequence is regarded as one-dimensional channel. Stacking the four calculated features leads to a 4-channel representation and it is compatible to a CNN architecture. All AIS segments are represented into a 4-channel tensor with the shape of $(1xMx4)$, where M is the total size of all AIS segments. Figure 5 demonstrates the 4 channel components for each AIS segment. Finally, all channel values are processed by a *min-max* normalisation method to have a *0-1* distribution for model training purposes.

4.2. Semi-Supervised Convolutional Encoder-Decoder Network (SCEDN)

The AIS dataset records lots of ship movement information, but unlike the image classification dataset ImageNet (Deng et al., 2009), the AIS dataset does not have a large amount of accurate manual labelling data for training. At the same time, compared to the image classification, the ship encounter situations contain only three categories. If we use a small portion of labelled ship encounter situation data for training a high accuracy network, it becomes easier to transfer the weight into the large ship encounter situation dataset than the image classification. The model generalisation ability can also be guaranteed because the AIS data is non-heterogeneous.

As shown in Figure 6, the SCEDN architecture includes two parts: (1) a supervised CNN classifier on the left side of Figure 6, which is used to train the labelled AIS data, denoted as X_l, and (2) a Encoder-Decoder structure combining the left side except the last part of Softmax in Figure

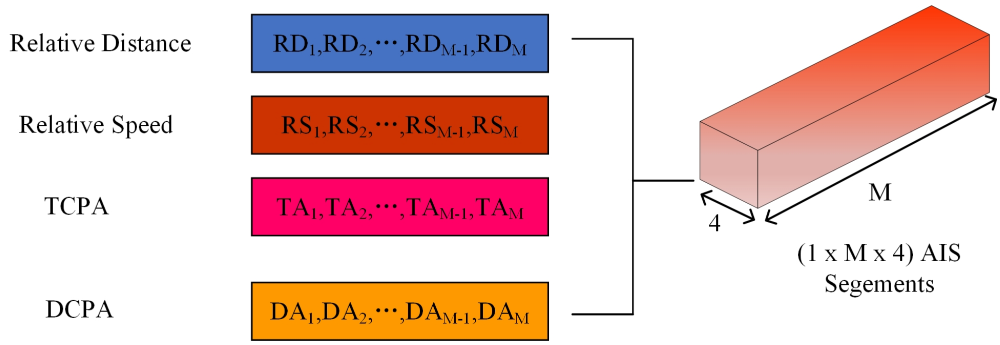

Figure 5: A 4-Channel representation for AIS segments: The relative distance and relative speed are directly derived from the AIS data. The T_{CPA} and D_{CPA} are calculated based on Equation 6 and Equation 7.

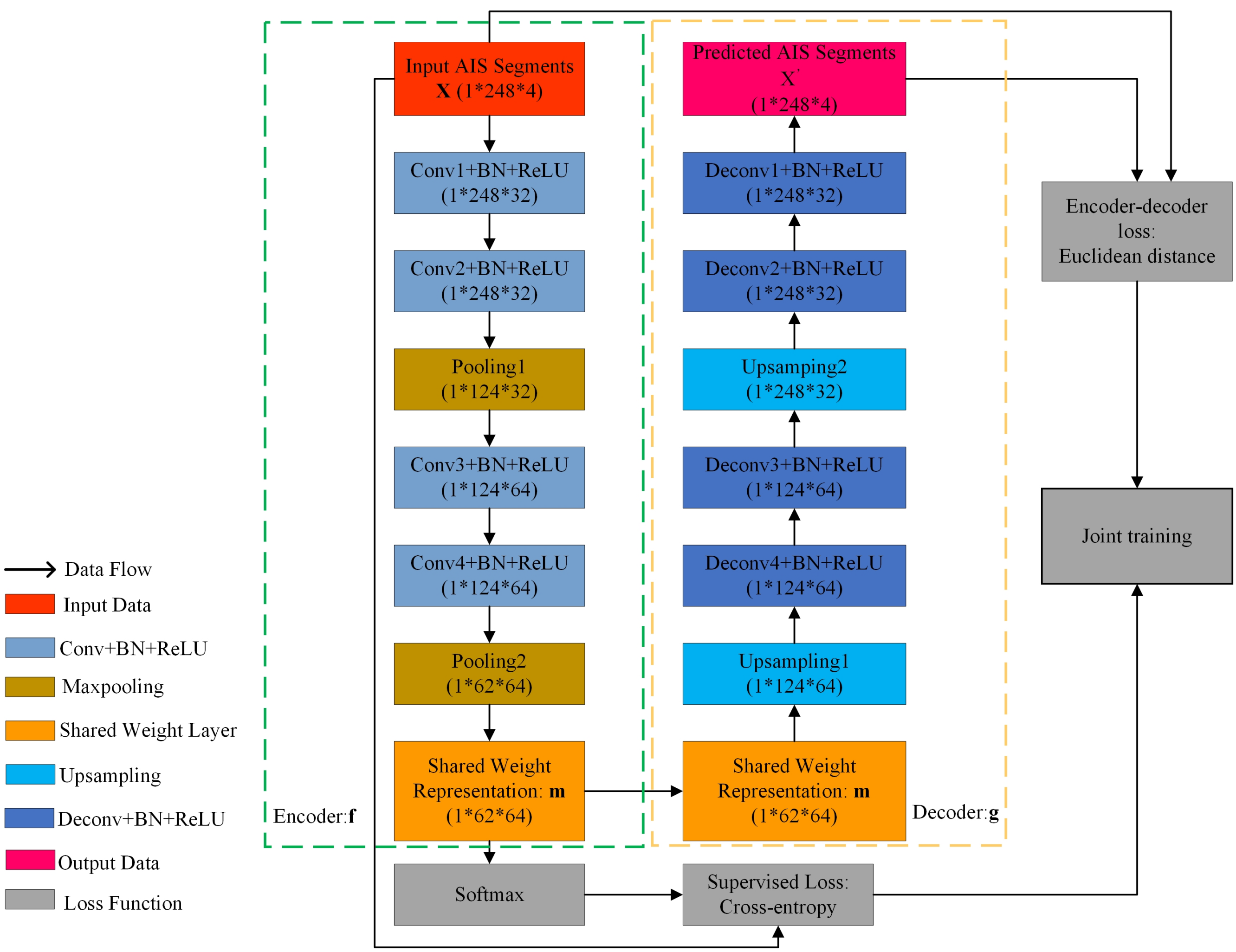

Figure 6: The overall architecture of SCEDN: It includes three basic networks (supervised CNN, encoder part and decoder part). In the left side of this figure, the supervised CNN combines the encoder part and Softmax layer. The different colour indicates various functions, shown in the left-bottom of this figure.

6 trains both labelled and unlabelled AIS data, denoted as $X_{comb} = X_l + X_r$. The X_l and X_r represent the labelled and unlabelled AIS data for 4-channel tensors, which is described in previous subsection.

4.2.1. *Supervised CNN classifier*

In the left part of Figure 6, the CNN classifier contains a stack of convolutional layers, Maxpooling layers and Softmax layer. The convolutional layers are the same as the encoder function derived from Dabiri et al. (2019) architecture but with two main differences:

- There is no pre-training step and no decoder component in the network proposed in Dabiri et al. (2019). This is because the AIS dataset in this paper contains a number of unlabelled data with only a small portion of manual labelled data. In order to train a highly accurate and robust classifier, it needs to use the manual labelled data as a warm-up role. Different to GPS which is made of heterogeneous features, the AIS data has a homogeneous feature, which facilitates to include a small pre-trained model in the whole training process to improve the semi-supervised classification ability. Hence, one of innovations in this paper is that we designed a shared weight representation layer to store the pre-trained weights to help the decoder sub-network to train the unlabelled AIS data.

- The second difference between this paper and the work in Dabiri et al. (2019) is the differences in the representation of the input data. In this paper, in order to properly classify the ship encounter situations, an unique 4-channel representation for AIS has been designed, as shown in Figure 5. The relative distance, relative speed, TCPA and DCPA have been embedded in the encoder-decoder network. On the contrary, in Dabiri et al. (2019), the movement features are consisting of the relative distance, time interval, speed, acceleration and jerk, with the aim to identify the people's commute ways and the movement features, which is quite different from the work in this paper.

The shared weight representation m means that it saves all weights and bias values to provide the decoder part to use. The Softmax layer is used to classify the labelled AIS segments to generate a probability distribution. The cross entropy loss function is used for the supervised CNN. The loss function for the labelled AIS segments can be expressed as:

$$Loss1_{S-CNN} = -\sum_{i=1}^{k} y_{l,i} log(p_{l,i}) \tag{8}$$

where $Loss1_{S-CNN}$ denotes the supervised CNN loss, $y_{l,i}$ is a binary value. If value is 1, it means the true ship encounter situation for the sample x_l else 0.

4.2.2. Encoder-Decoder classifier

The encoder-decoder classifier is used widely in the image segmentation to help extract the deep features. Based on these deep features, the decoder parts use a upsampling strategy to recover the image to help identification process (Socher et al., 2011). The encoder-decoder architecture commonly contains two parts: (1) an encoder part to map the input data into a latent representation, denoted as $m = f(X)$, and (2) a decoder part used to reconstruct the original data from a latent representation $X' = g(m)$. Functions f and g are downsampling and upsampling operations, respectively. The latent representation h called deep feature representation contains more high-level extraction information than the input data (Dabiri et al., 2019). As shown in Figure 6, the encoder part f includes four convolutional layers. In each convolutional layer, the batch normalisation and activation function ReLU are contained to help accelerate the convergence speed. After thoroughgoing each convolutional layer, the max-pooling layer are used to help finish the downsampling to reduce the computation complexity.

After each convolutional layer, the max-pooling layers are used to help finish the downsampling to reduce the computation complexity. We test different layer number combinations in the warm-up training phase and find that setting the filter size to 32 is the best option for our work in the first convolutional layer filters setting. It also should be noted that the layer design of neural networks is rather empirical with no clear theory for the design of deep neural networks been established. Many models including Inception (Szegedy et al., 2017), VGG (Simonyan and Zisserman, 2014), ResNet (He et al., 2016) are built under a larger number of experiments. Based on these models, there are some deep neural network designing principles for specific applications. For example, Krizhevsky et al. (2012) suggests that the filter numbers should increased with the deeper layer increasing. In addition, Szegedy et al. (2016) argues that balancing the width and depth of the deep network can get a higher quality and the filter numbers should be increased exponentially. In our model designing, the aim is to classify the ship encounter situations. The input AIS data dimension is (1x 284 x4), with the second dimension being the length of AIS and the third dimension being the ship movement characteristics which are consisting of the four parameters (relative distance, relative speed, TCPA and DCPA). Because the four ship movement parameters are the most important features to distinguish different types of ship encounters, we should increase the spatial features of the dimensions of the four parameters, while reducing the computing complexity of the second dimension to accelerate the convergence speed. Therefore, we set the filter size of the second convolutional layers as 64. The stride for each layer is configured to be 1 because the AIS segments have a continuous property. The max-pooling layer is $(1 * 2)$ with the stride of 2 to decrease the dimensional complexity.

The decoder part g has the same number of layers as the encoder and make the inverse operations. The tensor is firstly passed from the shared weight representation m to the upsampling generating a tensor with size $(1 * 124 * 64)$. After that, the output features from the upsampling layer is passed to a deconvolutional layer, in which the tensor shape is the same with the upsampling layer. After finishing a series of deconvolutional and upsampling operations, the output shape X' is the same with the input shape X. For the encoder-decoder architecture, the loss function here uses the Enclidean distance to measure. The equation is:

$$Loss2_{U-EDN} = \sum_{i} (X_i' - X_i)^2 \tag{9}$$

where, the $Loss2_{U-EDN}$ means the loss function for the encoder-decoder network. $X_i^{'}$ and X_i are output and input elements.

4.3. Training Scheme for SCEDN

The training scheme for SCEDN is to first train a Supervised CNN followed by training a Encoder-Decoder network. In order to realise a joint training, we firstly need to extract useful information from the labelled ship encounter situation data, and the converted AIS segments in Section 4.1 belong to homologous data and do not contain any context information, which makes the training weights easy to be reused. Therefore, in this paper, we can first use a small portion of labelled AIS data to train a high accurate supervised CNN classifier. The trained learning weights and bias values from CNN can be saved as initial weights for further encoder-decoder network training when unlabelled data is used.

More specifically, as shown in Figure 7, the input X, y represents the labelled data and is used to train a supervised CNN classifier. The training weights are then saved in a shared weight representation m which can be reused by the Encoder-Decoder network. Then, when the unlabelled data ($Unlabelled\ data\ :\ x^{'}$) is inputted, it is firstly fed through the CNN layer, which now acts as the encoder, using the shared weight to learn the distribution and then the decoder network is used to recover the original data representation. The difference between the raw data and the recovered data through the decoder network is calculated by the metric of Euclidean distance, as a loss value for the unlabelled data.

However, the proportion of unlabelled data in the AIS dataset is usually large and the training process through the supervised CNN and the encoder-decoder network can easily result in a label imbalance problem which reduces the model generalisation ability. Many studies have discussed such a problem (Cao et al., 2019; Cheng et al., 2020; Zhang et al., 2020b) with the common methods to deal with the label imbalance being: (1) increasing the sample data; (2) introducing hyperparameters. In this research, because the proportion of the labelled ship encounter situation data is relatively small and can not be increased in an easy way, the second paradigm that is to introduce the hyperparameters has been adopted. More specifically, two hyperparameters α, β are introduced in this paper in Equation 10 as:

$$Loss_{joint} = \alpha * Loss2_{U-END} + \beta * Loss1_{S-CNN} \quad (10)$$

where, α, β are the joint training loss function hyperparameters to solve the label imbalance problem combining Equation 8 and Equation 9. The range of the α, β is from 0 to 1. When $\alpha = 0$, the network is purely on the S-CNN; whereas, when $\alpha = 1$, the training process relies on the U-END. In general, in order to get a high generalisation performance, we use the S-CNN in the initial training stage to get a higher proportion of shared weights by configuring β larger than α. In the later training process, α value will be configured to be larger than β to enable a better performance for the unlabelled data using the U-END. Note that varying hyperparameters will undoubtedly affect the final training results, and a detailed study revealing how this is resolved together with training details will be explained in Section 5.

5. Experiments and Results

In this section, the performance of SCEDN algorithm will be evaluated and the relevant comparisons will be presented. The cleaning and preprocessing of dataset will be firstly discussed. Then, various supervised and semi-supervised models are described to compare with the proposed SCEDN model. The analysis and discussion will be presented in the final part of this section.

5.1. Experimental Setup

5.1.1. Dataset Description and Data Preprocessing

The training data in this paper uses a historical AIS data obtained from the Tianjin seaport and the open AIS dataset from Danish Maritime Authority. Note that at the moment, using the designed network, two conditions including ships' draft and water depth have not been considered due to following reasons: 1) our training data is collected from two locations, i.e. the Tianjin sea port (outside the seaport operational area) and Danish open sea area. Both of them belong to open sea areas, of which the depth satisfies the ship's safety restriction; 2) the proposed semi-supervised classification model has been designed to use AIS as the primary data input, which may not necessarily contain water depth or draft information for certain ships. Also, by using the most common navigational information within AIS, such as speed, distance, position and course, accurate identification results can be well obtained. In addition, this research is mainly focused on analysing ships' movement within middle or long range (5 or 10 nautical miles), where the draft is not a main factor affecting the proposed classification model. Before each AIS segments are fed to the SCEDN, a preposessing for AIS segments is conducted to remove errors and anomaly data. Each AIS segment is filtered by following principles:

- Each AIS segment time stamp over the 20 minutes thresholds is splitted to two different segments.
- For labelled AIS segments, the speed or course do not exceed a certain and realistic range of movement characteristics.
- For unlabelled AIS segments, any speed over 30 knots or course over 20 degree should be identified and removed.

After removing and cleaning the impractical AIS segments, the cleaned data features are shown in Table 1. The size of total labelled data is 15,653 segments in which crossing situation is 8,238 segments, overtaking is 3,802 and head-on situation is 3,613. The size of unlabelled data is 53,260 for training and the number AIS segments for testing is 6,745. Our ship encounter situation mode list is:

$$y = \{crossing, overtaking, head - on\}$$

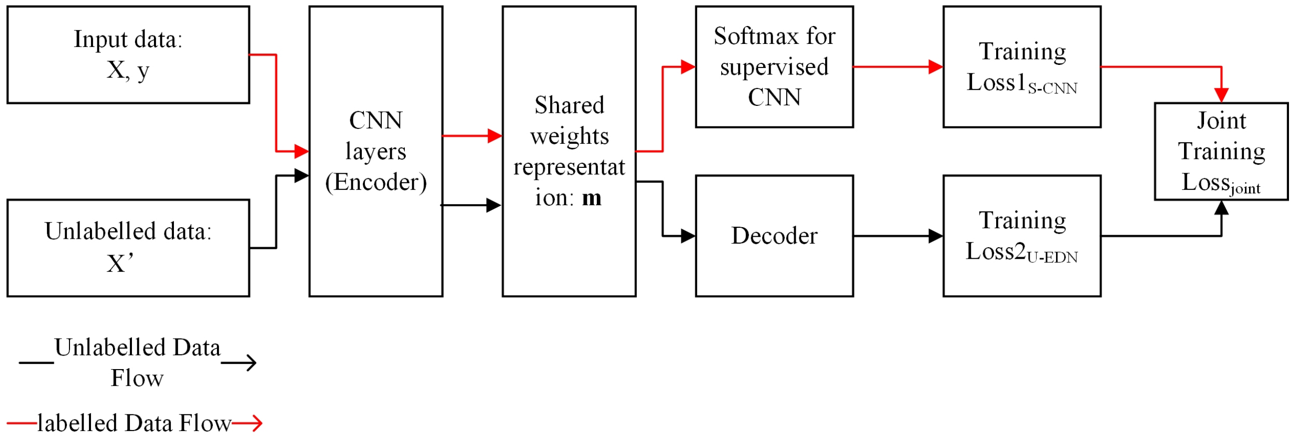

Figure 7: Training scheme for the SCEDN: The training scheme contains two different parts. The labelled input data through the supervised training scheme as a warm-up role, which the data flow in figure is shown as red colour. After getting the shared weighted representation from the labelled data, the unlabelled data training is carried out and the data flow is denoted as black colour in figure. The shared training layers can save the weights for the unlabelled data using. After that, the joint training is deployed for fine tuning.

Table 1
The details for training and testing data: ship encounter types are labelled manually and total number for training and testing are 15,653 and 8,456 respectively.

Ship Encounter type	Number of segments	
	Training Dataset	Test Dataset
Crossing	8238	5230
Overtaking	3802	1356
Head-on	3613	1870
Total labelled	15653	8456
Total unlabelled	53260	6745

and time-interval threshold for each AIS segment is set to be 20 minutes and each AIS segment size is 248.

5.1.2. Baseline Models

In order to get a realistic performance of our model, two baseline models are set: (1) supervised algorithms and (2) semi-supervised algorithms, for comparison. In the comparison experiments, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree (DT) and Supervised Convolutional Neural Network (S-CNN) are selected as the role of supervised algorithms to evaluate the algorithm performance compared to proposed SCEDN model.

For semi-supervised algorithms, the *semi-two-steps* method is used to compare with the proposed SCEND model where the joint training schemes are used. The *semi-two-steps* method is firstly trained on both labelled and unlabelled AIS segments. The labelled AIS segments are transferred to the shared weights layer using the encoder architecture f. Then, the transferred data are trained using the supervised CNN algorithm which is a logistic regression to evaluate the performance. The difference between the *semi-two-steps* with proposed SCEDN is that SCEDN is firstly trained by the labelled data using the supervised CNN as a warm-up. And then, the shared training weights are used for the unlabelled data using.

5.1.3. Performance Evaluation Metrics

In order to evaluate the performance of SCEDN, a confusion matrix, as shown in Table 2, is introduced to help with the assessment. The precision indicating the ratio of correctly predicted positive classes to the total predicted positive classes, is firt calculated using Eq.11. The recall, also called sensitivity, denoting the ratio of correctly predicted positive classes to the all actual labelled classes, is then calculated using Eq.12. Based upon precision and recall values, the accuracy, which is the most intuitive performance measurement and simply a ratio of correctly predicted classes to the total classes can be evaluated using Eq.13). The F_1 score, which can be calculated using Eq.14, is also evaluated to assess the weighted average of precision and recall to evaluate the comprehensive classification ability of the model.

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

Table 2
Confusion Matrix for Classification

Encounter Type	Predicted as class i	Predicted as other class
Class i	True Positive (TP)	False Negative (FN)
other class	False Positive (FP)	True Negative (TN)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

$$F_1 = \frac{2 * Precision * Recall}{Precision + Recall} \tag{14}$$

5.2. Performance Evaluations

In all conducted experiments, models are trained and evaluated using 5-fold validation and getting average value for comparison, where 5-fold validation is only used on the labelled AIS segments. The data processing and models are implemented under Python programming environment with Tensorflow used for deep learning models building and the package of *scikit-learn* package in Python is integrated in the environment for proposed supervised baseline models implementation. These experiments are trained on a High Performance Computing (HPC) facility at UCL with one GPU node.

5.2.1. SCEDN Evaluation

Figure 8 and 9 are the performance results of SCEDN model and baseline models under the metrics of accuracy and F_1 score. Each model is trained using different proportions of labelled AIS segments (10%, 25%, 50%, 75%, 100%) to reveal the model effectiveness.

In Figure 8, it shows the advantage of proposed SCEDN model and its training strategy in comparison with other baseline models. The result can be concluded that DT has a better performance in the 10% labelled AIS segments scenario. Under the other proportion groups, it shows that SCEDN model has a better classification ability than other baseline models. From the perspective of supervised algorithms, the S-CNN and DT are more competitive because the accuracy for the other models are relatively low.

With regards to the comparison between the supervised and semi-supervised algorithms, the results of *semi-two-steps* and our proposed model (SCEDN) show that the semi-supervised classification ability is better than the supervised algorithms in most cases. More specifically, by comparing with the supervised algorithms, it shows that using the semi-supervised strategy under different proportions of labelled AIS segments can help improve the algorithm classification ability for the unlabelled AIS segments. This is due to the warm-up strategy provided by the labelled data during the training process.

Furthermore, if we evaluate the SCEDN accuracy in Figure 8 with the *semi-two-steps*, the adopted joint training strategy in SCEDN can significantly improve the accuracy performance. Such results show that with the help of initialisation, training a semi-supervised model combined with labelled data can get a better performance than the *semi-two-steps* method. Also, The SCEDN could improve around 6.0 percent accuracy than the other baseline models and even more for the S-CNN model, which has the worst performance.

The accuracy is only one of important metrics for evaluating the training performance for a complex classification task. Another effective and explicit method for evaluating the performance of SCEDN is to use the F_1 score based on Equations 11, 12 and 14. The F_1 score result is shown in Figure 9, which also well demonstrates that the proposed SCEDN model provides a better result compared to all the other baseline models, especially in the scenarios where labelled AIS data accounts for 25%, 50%, 75% and 100% of the whole dataset, respectively.

The proposed SCEDN in this paper has been well trained off-line and the inference time for the model is validated in the test dataset. Although the training phase may take a relatively long time depending upon specific computing resources, the inference time for our model is efficiently fast. In Figure 10, the results for the SCEDN under different proportions of labelled AIS segments are provided. It has been well demonstrated that highest inference speed could be achieved when a 100% proportion of labelled AIS segments is used because all the data is labelled and the model (S-CNN in this case) does not contain any decoder component. When data is not 100% labelled, our SCEDN model will be used to enable a semi-supervised learning process, where the decoder is adopted to deal with the unlabelled AIS data. It is evident that a highly efficient inference speed can also be achieved by the proposed SCEDN with the longest time less than 30 ms when 75% of data is labelled.

5.2.2. Overall Performance Evaluation

Table 3 shows the accuracy and F_1 score of SCEDN model under three different segments:

- True Labelled AIS Segments: The different kinds of encounter situations are segmented by the true labelled data.
- Normalised AIS Segments: The ship encounter situations are partitioned into different segments and 20 minutes for each segment is set in our dataset. This can make sure that each AIS segment just only have one encounter situation in every 20 minutes.
- SCEDN: The ship encounter situations are divided based on the proposed SCEDN method.

From Table 3, the overall accuracy and F1 score of SCEDN model is only 4% lower than the true labelled group, which

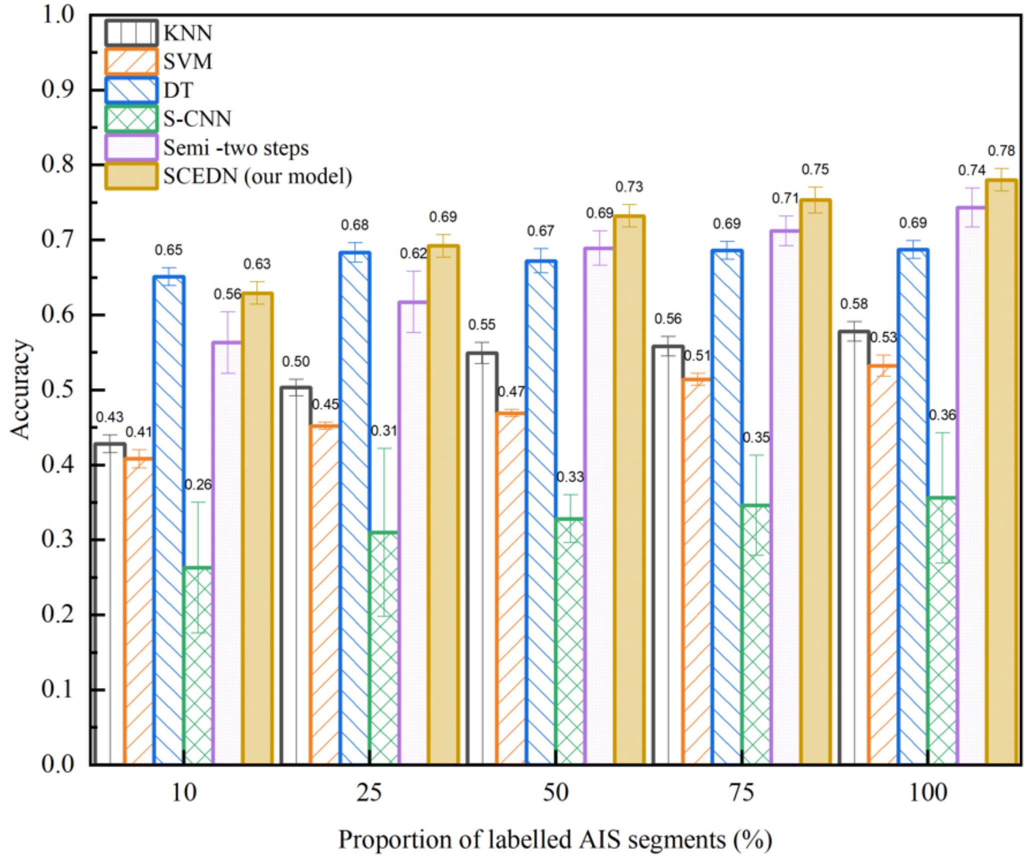

Figure 8: Accuracy for different models with various proportion of labelled data

Table 3
Comparison of accuracy and F1 score for our SCEDN model with different proportion of labelled AIS segments

Segmentation methods	Proportion of labelled AIS segments in the training phrase									
	10%		25%		50%		75%		100%	
	Accuracy	F1	Accuracy	F1	Accuracy	F1	Accuracy	F1	Accuracy	F1
True labelled AIS	0.635	0.623	0.697	0.674	0.74	0.732	0.756	0.753	0.769	0.763
Normalisaed	0.54	0.512	0.587	0.58	0.625	0.621	0.648	0.64	0.674	0.668
Two-step	0.603	0.594	0.652	0.645	0.683	0.675	0.71	0.694	0.724	0.714

indicates that our SCEDN model classification ability is acceptable. However, if we use the normalised AIS segments method, the accuracy and F1 score will be compromised compared with the other two groups. The performance of the second method is 10% lower than the first actual labelled AIS segments method and 6% less than our proposed two-step segments.

5.3. Further Analysis and Discussion

In this subsection, we evaluate our SCEDN model based on the label imbalance problem, the SCEDN model architecture analysis and the prediction ability for the ship encounter situations.

5.3.1. Label Imbalance Problem

In Section 4.3, a series of training scheme methods to alleviate the label imbalance problem are set. The hyperparameters α and β are introduced in the proposed SCEDN model. Figure 11 shows test accuracy based on several training schemes for the fine tuning hyperparameters in the loss function.

In order to find optimal parameters for α and β, 6 training schemes are designed in SCEDN model. The hyperparameter fine-tuning schemes are divided into two group: Group1: #1 - #3; Group2: #4 - #6, as shown in Figure 11. In scheme #1, the parameter α decreases gradually from 1 to 0.1 and the parameter β keeps the value of 1 to compare to the scheme #2. The training scheme #1 gradually trans-

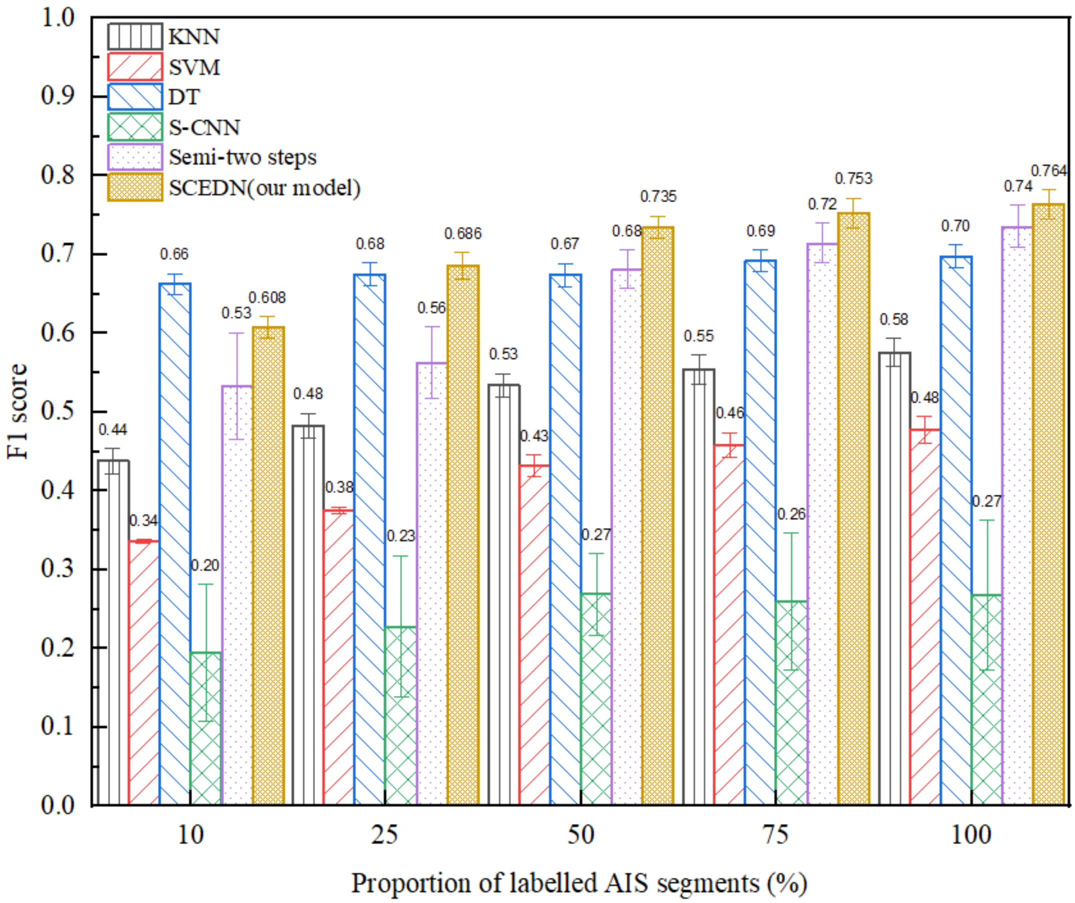

Figure 9: F1 scores for different models with various proportion of labelled data

forms the training attention from the unsupervised learning into the supervised CNN. Scheme #2 is similar to scheme #1 by keeping the training attention on the unsupervised learning compared with scheme #1. Scheme #3 keeps α and β to 1 during the whole training to compare with scheme #1 and #2. The scheme from #1 to #3 only include one training stage and the training process will stop when no improvements can be made after 3 epochs. From Figure 11, the hyperparameters α and β has not obvious improvements for the accuracy in different proportion of labelled AIS segments.

Scheme #4 and #5 have two training stages to compare with the one-stage training strategy from Group 1 training schemes. For the two training stages, parameter α and β are used with different values. In the first training stage, α and β are set separately ($\alpha = 1, \beta = 0$) and ($\alpha = 0, \beta = 1$). In the second training stage, the parameters are set into ($\alpha = 1, \beta = 1$) and ($\alpha = 1, \beta = 1$). The reason for such a training scheme setting is that it needs to find which part of the training hyperparameters could be fitted with the data distribution under different proportions of AIS segments. The stopping condition in the second stage is that no further improvement of the loss value can be generated. From the training results shown in Figure 11, it can be concluded that the training scheme #6 provides a better model performance than the other training schemes under different proportions of labelled AIS segments. Therefore, during the joint training, the hyperparameters α and β should be set to ($\alpha = 1, \beta = 1$) in first the training stage. When the loss values are converged by using the supervised CNN, the hyperparameters α and β should be readjust to ($\alpha = 1, \beta = 0.1$). It can therefore be summarised that the label imbalance issue for the ship encounter situation data can be solved through finding the different combinations for the hyperparameters.

5.3.2. SCEDN Model Architecture Analysis

With the increase of convolutional layers, the network feature extraction ability can be potentially improved as well (Simonyan and Zisserman, 2014). However, there is no relevant proof for how deep a neural network should be making the depth of the CNN a key hyperparameter to be investigated. Therefore, a series of experiments to find the optimal depth for our SCEDN model are designed, and the results are shown in Table 4. The test are conducted in a way that the number of layers in a neural network is initially set to be 2 and gradually increased towards 8 with an incremental step of 2. Training results are assessed by mainly considering the accuracy under different proportions of labelled AIS data with the average accuracy value for different number of layers summarised in Table 4. It is clear that increasing

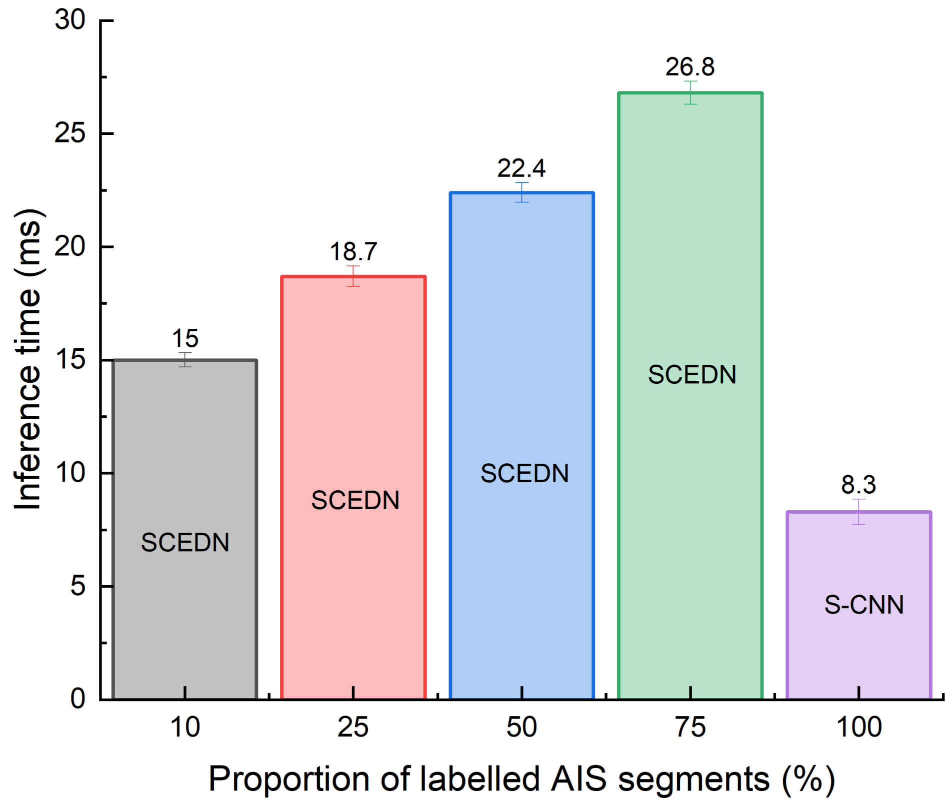

Figure 10: Inference time comparison for the SCEDN and Supervised Convolutional Neural Network (S-CNN). The unit for inference time is millisecond. In the x-axis, the meaning is the proportion of labelled AIS segments for testing, e.g. 10% means the labelled AIS segments are in proportion of 10 percent and the rest of unlabelled AIS segments are 90%. If the proportion of labelled AIS segments is 100%, it means all the AIS data is labelled and it is a supervised-CNN (S-CNN). The inference speed for the S-CNN is the fastest, because the S-CNN does not necessarily do the decoder part.

Table 4
Average Accuracy evaluation of the model configuration by different number of convolutional layers based on different labelled data

Number of Layers	Proportion of labelled AIS segments					Avgerage Accuracy
	10%	25%	50%	75%	100%	
2	0.625	0.681	0.723	0.741	0.751	0.704
4	0.628	0.694	0.734	0.752	0.769	**0.715**
6	0.628	0.694	0.734	0.752	0.759	0.713
8	0.628	0.697	0.724	0.748	0.764	0.712

layer numbers to 4 can help to improve SCEDN accuracy by 1 percent.

5.3.3. *Prediction Ability for Ship Encounter Situations*

In order to evaluate the prediction ability for our SCEDN model, the confusion matrix for the labelled AIS segments is introduced. The evaluation is conducted on the test dataset. Table 5 shows the confusion matrix, precision and recall for each label. The crossing situation has a large proportion in the test dataset and can get a higher recall around 95%. We can see that the overtaking is the smallest proportion in our test dataset, but the recall and precision may not be lower than the crossing and head-on situations. This is due to the training scheme of our proposed in the Section 4.3 can specifically solve the label imbalance problem. Another interesting finding is that the head-on situation's recall and precision are not so high with the reason being that the angle for head-on situations for two ships is difficult to identify com-

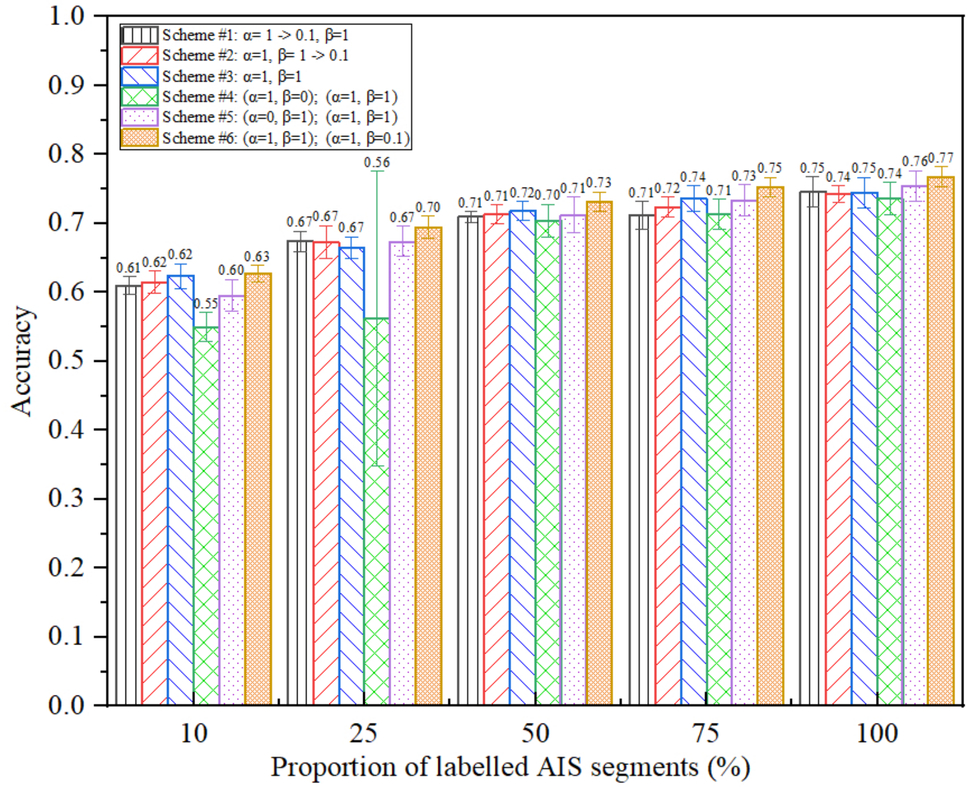

Figure 11: Comparison of accuracy for different hyperparameter schedules with different proportion of labelled data

Table 5
Confusion Matrix for our SCEDN in Test Dataset

Confusion Matrix		Predicted ship encounter			
		crossing	overtaking	head-on	Recall
True label	crossing	4980	30	220	**0.95**
	overtaking	88	1208	60	0.89
	head-on	182	38	1650	0.88
	Precision	0.94	0.94	0.85	

pared to that in crossing situations. One of potential ways to address such a compromised classification is via providing more head-on situation data.

6. Conclusions

In this paper, we have proposed a Semi-Supervised Convolutional Encoder-Decoder Network (SCEDN) for ship encounter situation classification purely based on ship AIS data. First, a segmentation process is developed based on the ship movement characteristics such as relative distance, relative speed, time to closest point of approach and distance to closest point of approach. Then based on these four features, a 4-channel tensor representation for each ship encounter situation is built for the SCEDN network. Due to the shortage of labelled ship encounter situation data and a large proportion of unlabelled ship encounter situation, a semi-supervised architecture coupled with a designated training scheme have been designed. The labelled data plays a warm-up role in the whole training and shares the training weights to the unlabelled data to speed up the model convergence. Our extensive experiments show that our model has a better performance than the other baseline classification models.

In terms of the future work, although a good training performance can be achieved using the proposed algorithm, further improvements such as replacing the training scheme by an end-to-end architecture can potentially help with the hyperparameters fine tuning process.Also, to better understand complex encountering situations in practical maritime environments, the ship encounter situation representation can also be improved by adding with more features apart from the four channel features already used in this paper to help extract more details.

Acknowledgement

The authors acknowledge the use of the UCL Myriad High Performance Computing Facility (Myriad@UCL), and associated support services, in the completion of this work. This research is supported by the National Nature Science Foundation of China (No.51879119). And also, this research is funded by a grant (No.JMCBZD202008) from Navigation College of Jimei University, National-local Joint Engineering Research Centre for Marine Navigation Aids Services.

References

Authority, D.M., . Danish maritime authority ais dataset. [EB/OL]. https://www.dma.dk/SikkerhedTilSoes/Sejladsinformation/AIS/Sider/default.aspx Accessed 16, Nov, 2020.

Bachir, D., Khodabandelou, G., Gauthier, V., El Yacoubi, M., Puchinger, J., 2019. Inferring dynamic origin-destination flows by transport mode using mobile phone data. Transportation Research Part C: Emerging Technologies 101, 254–275.

Berg, N., Storgård, J., Lappalainen, J., 2013. The impact of ship crews on maritime safety. Publications of the Centre for Maritime Studies, University of Turku A 64.

Bovcon, B., Kristan, M., 2020. A water-obstacle separation and refinement network for unmanned surface vehicles. arXiv preprint arXiv:2001.01921 .

Cao, K., Wei, C., Gaidon, A., Arechiga, N., Ma, T., 2019. Learning imbalanced datasets with label-distribution-aware margin loss, in: Advances in Neural Information Processing Systems, pp. 1567–1578.

Chen, P., Shi, G., Liu, S., Zhangi, Y., 2018. Decision support based on artificial fish swarm for ship collision avoidance from ais data, in: 2018 International Conference on Machine Learning and Cybernetics (ICMLC), IEEE. pp. 31–36.

Chen, X., Liu, Y., Achuthan, K., Zhang, X., 2020. A ship movement classification based on automatic identification system (ais) data using convolutional neural network. Ocean Engineering 218, 108182.

Cheng, K., Gao, S., Dong, W., Yang, X., Wang, Q., Yu, H., 2020. Boosting label weighted extreme learning machine for classifying multi-label imbalanced data. Neurocomputing 403, 360–370.

Dabiri, S., Lu, C.T., Heaslip, K., Reddy, C.K., 2019. Semi-supervised deep learning approach for transportation mode identification using gps trajectory data. IEEE Transactions on Knowledge and Data Engineering 32, 1010–1023.

Damastuti, N., Aisjah, A.S., Masroeri, A.A., 2019. Classification of ship-based automatic identification systems using k-nearest neighbors, in: 2019 International Seminar on Application for Technology of Information and Communication (iSemantic), IEEE. pp. 331–335.

De Souza, E.N., Boerder, K., Matwin, S., Worm, B., 2016. Improving fishing pattern detection from satellite ais using data mining and machine learning. PloS one 11, e0158248.

Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database, in: 2009 IEEE conference on computer vision and pattern recognition, Ieee. pp. 248–255.

El Mekkaoui, S., Benabbou, L., Berrado, A., 2020. Predicting ships estimated time of arrival based on ais data, in: Proceedings of the 13th International Conference on Intelligent Systems: Theories and Applications, pp. 1–6.

Gao, M., Shi, G., Li, S., 2018. Online prediction of ship behavior with automatic identification system sensor data using bidirectional long short-term memory recurrent neural network. Sensors 18, 4211.

Gao, M., Shi, G.Y., 2020. Ship collision avoidance anthropomorphic decision-making for structured learning based on ais with seq-cgan. Ocean Engineering 217, 107922.

Gao, S., Liu, F., Jiang, C., 2021. Improvement study of modal analysis for offshore structures based on reconstructed displacements. Applied Ocean Research 110, 102596.

Ginoulhac, R., Barbaresco, F., Schneider, J.Y., Pannier, J.M., Savary, S., 2019. Coastal radar target recognition based on kinematic data (ais) with machine learning, in: 2019 International Radar Conference (RADAR), IEEE. pp. 1–5.

Harati-Mokhtari, A., Wall, A., Brooks, P., Wang, J., 2007. Automatic identification system (ais): data reliability and human error implications. The Journal of Navigation 60, 373.

Hassel, M., Asbjørnslett, B.E., Hole, L.P., 2011. Underreporting of maritime accidents to vessel accident databases. Accident Analysis & Prevention 43, 2053–2063.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Jiang, X., Silver, D.L., Hu, B., de Souza, E.N., Matwin, S., 2016. Fishing activity detection from ais data using autoencoders, in: Canadian Conference on Artificial Intelligence, Springer. pp. 33–39.

Kostopoulos, G., Karlos, S., Kotsiantis, S., Ragos, O., 2018. Semi-supervised regression: A recent review. Journal of Intelligent & Fuzzy Systems 35, 1483–1500.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems 25, 1097–1105.

Krüger, M., 2018. Experimental comparison of ad hoc methods for classification of maritime vessels based on real-life ais data, in: 2018 21st International Conference on Information Fusion (FUSION), IEEE. pp. 1–7.

Lang, H., Wu, S., Xu, Y., 2018. Ship classification in sar images improved by ais knowledge transfer. IEEE Geoscience and Remote Sensing Letters 15, 439–443.

Leodolter, M., Widhalm, P., Plant, C., Brandle, N., 2017. Semi-supervised segmentation of accelerometer time series for transport mode classification, in: 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), IEEE. pp. 663–668.

Liu, C., Liu, J., Zhou, X., Zhao, Z., Wan, C., Liu, Z., 2020. Ais data-driven approach to estimate navigable capacity of busy waterways focusing on ships entering and leaving port. Ocean Engineering 218, 108215.

Liu, J., Timsina, P., El-Gayar, O., 2018. A comparative analysis of semi-supervised learning: the case of article selection for medical systematic reviews. Information Systems Frontiers 20, 195–207.

Ma, Y., Zhao, Y., Incecik, A., Yan, X., Wang, Y., Li, Z., 2021. A collision avoidance approach via negotiation protocol for a swarm of usvs. Ocean Engineering 224, 108713.

Mankabady, S., 1986. The International Maritime Organization, Volume 1: International Shipping Rules.

Mizythras, P., Pollalis, C., Boulougouris, E., Theotokatos, G., 2021. A novel decision support methodology for oceangoing vessel collision avoidance. Ocean Engineering 230, 109004.

Naeem, W., Irwin, G.W., Yang, A., 2012. Colregs-based collision avoidance strategies for unmanned surface vehicles. Mechatronics 22, 669–678.

Pallotta, G., Vespe, M., Bryan, K., 2013. Vessel pattern knowledge discovery from ais data: A framework for anomaly detection and route prediction. Entropy 15, 2218–2245.

Parsons, M.G., Chung, H., Nick, E., Daniels, A., Liu, S., Patel, J., 2008. Intelligent ship arrangements: a new approach to general arrangement. Naval Engineers Journal 120, 51–65.

Radu, V., Katsikouli, P., Sarkar, R., Marina, M.K., 2014. A semi-supervised learning approach for robust indoor-outdoor detection with smartphones, in: Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems, pp. 280–294.

Ramos, M.A., Utne, I.B., Mosleh, A., 2019. Collision avoidance on maritime autonomous surface ships: Operators' tasks and human failure events. Safety science 116, 33–44.

Robson, J., Qj, O.O., 2006. Overview of collision detection in the ukcs .

Sánchez Pedroche, D., Amigo, D., García, J., Molina, J.M., 2020. Architecture for trajectory-based fishing ship classification with ais data. Sensors 20, 3782.

Schwenker, F., Trentin, E., 2014. Pattern classification and clustering: A review of partially supervised learning approaches. Pattern Recognition

Letters 37, 4–14.

Sheng, K., Liu, Z., Zhou, D., He, A., Feng, C., 2018. Research on ship classification based on trajectory features. The Journal of Navigation 71, 100–116.

Shi, J.h., Liu, Z.j., 2020. Deep learning in unmanned surface vehicles collision-avoidance pattern based on ais big data with double gru-rnn. Journal of Marine Science and Engineering 8, 682.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 .

Socher, R., Pennington, J., Huang, E.H., Ng, A.Y., Manning, C.D., 2011. Semi-supervised recursive autoencoders for predicting sentiment distributions, in: Proceedings of the 2011 conference on empirical methods in natural language processing, pp. 151–161.

Svanberg, M., Santen, V., Hörteborn, A., Holm, H., Finnsgård, C., 2019. Ais in maritime research. Marine Policy 106, 103520.

Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A., 2017. Inception-v4, inception-resnet and the impact of residual connections on learning, in: Proceedings of the AAAI Conference on Artificial Intelligence.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818–2826.

Tu, E., Zhang, G., Rachmawati, L., Rajabally, E., Huang, G.B., 2017. Exploiting ais data for intelligent maritime navigation: A comprehensive survey from data to methodology. IEEE Transactions on Intelligent Transportation Systems 19, 1559–1582.

Wijaya, W.M., Nakamura, Y., 2013. Predicting ship behavior navigating through heavily trafficked fairways by analyzing ais data on apache hbase, in: 2013 First International Symposium on Computing and Networking, IEEE. pp. 220–226.

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. Transport Reviews 39, 755–773.

Yuan, X., Zhang, D., Zhang, J., Zhang, M., Soares, C.G., 2021. A novel real-time collision risk awareness method based on velocity obstacle considering uncertainties in ship dynamics. Ocean Engineering 220, 108436.

Yuan, Z., Liu, J., Liu, Y., Li, Z., et al., 2019. A novel approach for vessel trajectory reconstruction using ais data, in: The 29th International Ocean and Polar Engineering Conference, International Society of Offshore and Polar Engineers.

Zhang, C., Bin, J., Wang, W., Peng, X., Wang, R., Halldearn, R., Liu, Z., 2020a. Ais data driven general vessel destination prediction: A random forest based approach. Transportation Research Part C: Emerging Technologies 118, 102729.

Zhang, J., Zhang, D., Yan, X., Haugen, S., Soares, C.G., 2015. A distributed anti-collision decision support formulation in multi-ship encounter situations under colregs. Ocean Engineering 105, 336–348.

Zhang, M.L., Li, Y.K., Yang, H., Liu, X.Y., 2020b. Towards class-imbalance aware multi-label learning. IEEE Transactions on Cybernetics .

Zheng, Y., Zhang, X., Shang, Z., Guo, S., Du, Y., . A decision-making method for ship collision avoidance based on improved cultural particle swarm. Journal of Advanced Transportation 2021.

Zhu, X., Goldberg, A.B., 2009. Introduction to semi-supervised learning. Synthesis lectures on artificial intelligence and machine learning 3, 1–130.