Ship feature recognition methods for deep learning in complex marine environments

Xiang Wang\textsuperscript{1,2} · Jingxian Liu\textsuperscript{1} · Xiangang Liu\textsuperscript{1,3} · Zhao Liu\textsuperscript{1} · Osamah Ibrahim Khalaf\textsuperscript{4} · Jing Ji\textsuperscript{5,6} · Quan Ouyang\textsuperscript{6}

Received: 20 January 2021 / Accepted: 22 January 2022 / Published online: 24 March 2022
© The Author(s) 2022

Abstract
With the advancement of edge computing, the computing power that was originally located in the center is deployed closer to the terminal, which directly accelerates the iteration speed of the "sensing-communication-decision-feedback" chain in the complex marine environments, including ship avoidance. The increase in sensor equipment, such as cameras, have also accelerated the speed of ship identification technology based on feature detection in the maritime field. Based on the SSD framework, this article proposes a deep learning model called DP-SSD. By adjusting the size of the detection frame, different feature parameters can be detected. Through actual data learning and testing, it is compatible with Faster RCNN, SSD and other classic algorithms. It was found that the proposed method provided high-quality results in terms of the calculation time, the processed frame rate, and the recognition accuracy. As an important part of future smart ships, this method has theoretical value and an influence on engineering.

Keywords Marine environments · SSD · CNN · Deep learning · Edge computing

Abbreviations

| Abbreviation | Acronym | Description |
|--------------|---------|-------------|
| AI | Artificial intelligence | |
| AIS | Automatic identification system | |
| AP | Average precision | |
| BN | Batch normalization | |
| CBFF-SSD | Context-based feature fusion single-shot multi-box detector | |
| CNN | Convolutional neural network | |
| CSSD | Context-aware single-shot multi-frame object detector | |
| DBN | Deep belief network | |
| DL | Deep learning | |
| DSSD | Deconvolution single shot detector | |
| ECS | East China sea | |
| Faster R-CNN | Faster regional convolutional neural networks | |
| FFESSD | Feature fusion and enhancement for single shot detection | |
| FN | False negative | |
| FP | False positive | |
| FPS | Frames per second | |
| FSSD | Function single-shot multi-box detector | |
| IALA | International association of marine aids to navigation and lighthouse authorities | |
| IMO | International maritime organization | |
| IOU | Intersection over union | |
| mAP | Mean average precision | |
| NavCom | Integration of navigation and communication | |
| NPU | Neural processing units | |
| PE | Processing elements | |
| Prec | Precision | |

\textsuperscript{1} Hubei Key Laboratory of Inland Shipping Technology, School of Navigation, Wuhan University of Technology, Wuhan 430063, China
\textsuperscript{2} Donghai Navigation Safety Administration (DNSA) of the Ministry of Transport (MOT), Shanghai 20086, China
\textsuperscript{3} Air Force Early Warning Academy, Wuhan 430019, China
\textsuperscript{4} AlNahrain Non-Renewable Energy Research Centre, Al-Nahrain University, Baghdad 10072, Iraq
\textsuperscript{5} School of Information Engineering, Wuhan University of Technology, Wuhan 430070, China
\textsuperscript{6} Wuhan Xingtu Electronics Co., Ltd, Wuhan 430070, China
Introduction

A large cruise or tanker collision is a major maritime disaster. A collision causes a series of consequences, such as personal injury, property damage, and pollution of the wrecked sea area. As shown in Fig. 1 the tanker, Sanchi, which sank in the East China Sea, passed through the International Maritime Organization (IMO). The expert group investigated and determined that its crash was mainly caused by human factors, specifically, the eyes’ misrecognition of targets in complex marine environments, which cause serious consequences, such as that collision. In terms of DP-SSD, a real-time identification and navigation method is proposed based on deep learning (DL) for ships.

Artificial intelligence (AI) recognition technology has been developing rapidly, and target recognition and tracking methods based on complex environments have been researched and experimented [9, 53]. At the same time, the AI and DL research on maritime engineering has mainly focused on the real-time monitoring data fusion method for integrating navigation and communication (NavCom) in the complex maritime environment and mainly process the data of the automatic ship identification system (AIS) [2, 11]. As previous researchers did not use experimental platforms for feature recognition of marine ships, no real-time monitoring method has been based on NavCom in the complex maritime environment for DL [43].

AIS time intervals do not reflect ship data in real-time. In specific ship navigation scenarios, such as in narrow waterways, specific waters with high ship exchange flow density, or extremely low navigation, AIS may be invalid.

The ship’s single-shot multi-box detector (SSD) in the real-time identification method for NavCom in a complex maritime environment is a brand-new method that can help ships avoid collisions and identify obstacles in autonomous navigation. Compared with the traditional ship AIS, the SSD ship real-time identification method is based on human vision. In addition, ships sailing in polar regions, narrow waterways, and ship anchorages have strong practical needs. In recent years, there have been numerous marine accidents and marine pollution incidents caused by marine vessel navigation and marine engineering construction. Only new methods can reduce these types of incidents.

In the past two decades, the large-scale use of automatic ship identification systems has not substantially reduced maritime accidents. Therefore, the multisource heterogeneous data fusion method for NavCom in a complicated maritime environment is important to research.

This paper focuses on three points: first, there is an actual need for a ship feature identification method based on deep learning for the integration of communication and navigation in a complex maritime environment; second, the paper explains the shortcomings and limitations of the current AIS, and the limitations of the real-time dynamic ship monitoring and identification method based on AIS data are proposed; finally, the paper explains the role of the ship feature identification method based on deep learning for the integration of communication and navigation in a complex maritime environment. The results show that the use of the deep learning-based ship feature recognition method with integrated communication and navigation in a complex marine environment on the surface can mimic biological visual recognition and identify ships.

The rest of this paper is organized as follows: in Sect. 2, we present a literature review on the research of multisource heterogeneous data fusion methods for the integration of communication and navigation in a complex maritime environment. In Sect. 3, the experimental framework is used to reveal a ship’s feature recognition method based on DL for NavCom in complex environments. Section 4 introduces the research area and dataset, before presenting the results and discussion. Finally, a conclusion is made.

Literature survey

Collision avoidance among objects covering multidisciplinary areas has been studied by many scholars and engineers. The solutions introduced can be divided based on positioning and navigation information, radio frequency identification (RFID) information, and on image information detection and multi-sensor data fusion [41]. Among them, the method based on positioning information emphasizes the use of positioning and navigation information as the identification of collision detection. Its advantage is that it can use position, speed and time information to solve collision prediction and real-time detection issues. The disadvantage is that it strongly relies on PVT data. The refresh rate is high,
and it is difficult to have real-time alarms for the positioning deviation caused by multiple factors [29, 35, 57]. The method based on RFID information converts the multi-object detection problem into the anti-collision problem between multiple readers and tags. The advantage is that the use of the time division multiple access (TDMA) ALOHA algorithm can effectively increase the number of object detections, but the disadvantage is that a time reference is required to synchronize all tags in the reader’s reading area [1, 3, 51, 52]. The method based on image information uses image information as a sign to detect collisions. Its advantage is that it is more intuitive and focuses on the principle of bionics, such as human eyes. The disadvantage is that if accurate and fast recognition is required, it needs to rely on training sets and algorithms [12, 19, 59].

Certain specific fields have already possessed unique advantages over manual operations. For example, the detection of the human colon, gallbladder, appendix, and stomach, as well as the identification of diseased bodies can provide doctors with robust and reliable pathological data [14, 39, 40, 58] based on the application foundation of artificial intelligence technology. AI target recognition methods based on visual recognition have been applied in research on multisource heterogeneous data fusion for the integration of communication and navigation in complex maritime environments [34, 50].

Machine learning has been widely used in various fields. The literature [20] is focuses on recognition tasks of natural language and studies the sentence clustering recognition method of the network. The concept of pre-reliability is used to measure the credibility of the sentence recognition results. The network simulation program analyzes the influence of the neural network on sentence recognition and performs a post-reliability evaluation index for the credibility of the model’s construction. The literature [8, 33, 36, 37, 46] examines research on multiple application scenarios, including B2B communication services, wireless sensor network keys and personal wireless networks. Among them, [46] proposes a deep Boltzmann machine (DBM) for learning a data generation model that is composed of multiple input modes. The experimental results of bimodal data composed of images and text show that multimodal DBM can learn a good generation model. The joint space for image and text input is useful for retrieving information from unimodal and multimodal queries.

There are also many reviews on AI target recognition. The literature [44] systematically introduces the deep learning technology of marine target recognition. The method chapter is divided into two parts: supervised learning and unsupervised learning. The part of supervised learning focuses on introducing the respective principles and progress of AlexNet, ZFNet, VGG-16, GoogLeNet, ResNet, and SENet, while focusing on the deep belief network (DBN) and automatic coding model in the unsupervised learning part; it briefly introduces the target dataset on the surface and underwater, lists the existing datasets of Fish4K, Kyutech-10 K, QuickBird, SPOT-5, HRSC2016, R2VV-p, VAIS, E2S2-Vessel, FleetMon, MARVEL, CCF-BDCI, and performs comparative research. The method is divided into three parts: data preprocessing, feature extraction and recognition, and model optimization for analysis. Among them, image preprocessing is a key step in the recognition of ocean targets in deep learning models because images or videos are the premises and application foundations of deep learning methods. The literature [22, 28, 31, 55] compares target detection algorithms (R-CNN, Fast-RCNN, Faster-RCNN), end-to-end detection algorithms (YOLO series, SSD algorithm) and novel algorithms proposed recently (Cascaded RCNN algorithm, RefineDet algorithm, SNIP algorithm, R-FCN-3000 algorithm, DES algorithm, STDN algorithm). Comparing accuracy and speed as measurement indicators, it analyzes
and summarizes their respective advantages and disadvantages using public datasets.

The majority task of target detection is composed of object recognition and object location. The task of the former is to classify objects, and the task of the latter is to obtain the position of the object in the picture [32, 42].

The mainstream target detection model can be divided into a two-stage model and a one-stage model. The first step is to generate a series of candidate frames. The candidate frame generation methods include selective search, edgeboxes, deepmask, region proposal network (RPN), etc., and then the second step is to perform classification regression using Region convolutional neural networks (R-CNN), Fast R-CNN, faster-RCNN and other algorithms as typical representatives. The latter performs classification and regression while generating candidate boxes in each cell of the feature map, such as You only look once (YOLO), Single shot multi-Box detector (SSD), etc. The algorithm is a typical representative.

A two-stage technology evolution model is developed from R-CNN to fast-RCNN and then to faster-RCNN. R-CNN target detection can be divided into four steps: setting the extraction frame, extracting features by frame, training the classifier for image classification, and using regression to fine-tune the position of the selection frame. During the evolution of R-CNN, the idea of spatial pyramid pooling (SPP Net) contributed substantially to it [23, 30]. Fast R-CNN was proposed by Microsoft in 2015. Compared with the R-CNN algorithm, there are two differences. When training the classifier, a neural network is used to replace the support vector machine (SVM) in terms of efficiency, the training phase and the test. In this stage, Fast R-CNN is 9 times and 213 times that of R-CNN [15]. The Faster-R-CNN algorithm consists of two modules: the PRN candidate frame extraction module and the Fast R-CNN detection module. Among them, RPN is a fully convolutional neural network used to extract candidate frames. Fast R-CNN detects and recognizes the target in the proposal based on the proposal extracted by RPN [6, 27, 47].

In the one-stage technology evolution model, the YOLO algorithm represented by YOLO uses a simple convolutional neural network to predict the border and the prediction category at the same time, while using the characteristics of the entire image to predict and reduce errors. It does not require preprocessing and postprocessing. The problem with processing is that the test scale must be consistent with the training scale and each grid can only predict one object. YOLOv2 uses YOLO, which is faster and more accurate, with more than 9,000 recognition categories. However, the disadvantage is that when using the $13 \times 13$ output feature map, the smaller object may not have obvious features or even be ignored, though YOLOv2 uses a $1 \times 1$ convolution to reduce the number of channels from 512 to 64. In addition, after the convolutional layer, batch normalization (BN) is added to achieve a 2% improvement, which helps the regularization model remove the dropout. YOLOv3 implements logistic regression to regress the box confidence, and the prior and actual box Intersection over Union (IOU) is greater than 0.5 as a positive example. Unlike SSD, if there are multiple priors that satisfy the goal, only one prior with the largest IOU is taken [18, 45, 49, 56].

In summary, for the target detection network, two stages have higher accuracy, and one-stage speed is faster.

Proposed methodology

At present, in the industry, target feature learning and recognition technology has made great progress. For example, leading technological companies such as China SenseTime Technology and Questyle Technology have incorporated classic algorithms into engineering applications. However, there is no fundamental breakthrough in innovative research based on classic algorithms. Analyzing the principle of classic algorithms, YOLO and SSD are both very concise. They do not perform deep fitting on the sample dataset, though they do only on the linear connection of multidimensional data sample points. Therefore, the meaning of the test is only in the test set and the data sample. The numerical value of the points is simple and relative; therefore, only a single stage allows them to have a fast detection effect relative to Faster RCNN.

Furthermore, target recognition has been listed as a trending research topic in various areas. Shi et al. [38] proposed an accurate and effective target detection method in the literature, which is called the feature fusion and enhancement for single shot detection (FFESSD) model. It is verified through experiments, proving that when the input rate is 54.3 FPS and the input image size reaches $300 \times 300$, the average accuracy rate of the FFESSD can reach 79.1%. When the input rate is 30.2 FPS and the input image size is $512 \times 512$, the average accuracy rate can reach 81.8%, which is significantly better than the conventional SSD model, the deconvolution single shot detector (DSSD), feature fusion SSD (FSSD) and other advanced detectors. Scholars, such as Liu et al. [25] proposed a new type of SSD model in the literature. Through verification in the MICCAI challenge, it was found to have the fastest and best accuracy and recall rate. In general, it has excellent performance in polyp detection. The target detection method in SAR ship detection has also become a very prevalent study point. Scholars, such as Zhang et al. [53] proposed an improved Faster-RCNN method in the literature. The method could improve the detection results of small-scale distribution ships, while simultaneously improving the recall rate, which could provide a high-resolution remote sensing image-based detection.
method for offshore and inland ships. Chang et al. [5] proposed a YOLOv2-reduced method in the literature. Through testing and verification on the SAR (SSDD) and DSSD, it was found that compared with Faster R-CNN, the YOLOv2-reduced method improved the accuracy of ship detection and substantially reduced the calculation time. Zhang et al. [54] proposed a high-speed ship detection method for SAR images based on a convolutional neural network (ABYI, G-CNN for short) in the literature. The experimental results show that the detection speed of G-CNN is faster than that of other existing methods, such as faster regional convolutional neural networks (Faster R-CNN), SSD, and the YOLO method. Li et al. [21] proposed a SAR-based context-based feature fusion single-shot multi-box detector (CBFF-SSD) framework in the literature and adopted parallel deep learning processing with multiple neural processing units. The high-efficiency hardware architecture (NPU) of the processor is composed of two-dimensional processing elements (PE), which can calculate multiple output element maps at the same time. Experiments show that it can be further optimized in the detection of small targets, reducing the computational complexity and parameter size. Scholars, such as Ivan Lorencin used AI algorithms in the literature to identify marine targets from the air [16] and classified marine targets into ships and other targets through classification.

In 2016, Liu et al. [26] proposed a deep neural network method called SSD, which uses a single object to detect objects in images, according to the literature. It discretizes the output space of the bounding box into a set of default boxes at different aspect ratios and proportions for each function graph position. For a 300 × 300 input, SSD reaches a 74.3% mAP. When VOC2007 is tested at 59 FPS, the SSD reaches a 76.9% mAP for 512 × 512 inputs, which is better than the latest Faster R-CNN model. Fu et al. [10] proposed a method to introduce other contexts into the latest general object detection in the literature in 2017. In the PASCAL VOC and COCO tests, the DSSD of a 513 × 513 input reaches an 81.5% mAP in the VOC2007 test, an 80.0% mAP in the VOC2012 test, and a 33.2% mAP in the COCO test, which is better than the latest R-FCN method in each dataset. Based on the traditional SSD, literature [17] obtains better generalizability by changing the structure to make it close to the classifier network instead of adding layers. For the Pascal VOC 2007 test set, when using the VOC 2007 and VOC 2012 training sets, the mAP is 78.5% at a 300 × 300 input and a rate of 35.0 FPS. With an input of 512 × 512 and 16.6 FPS, the mAP can reach 80.8%. The accuracy is better than those of the traditional SSD, YOLO, Faster-RCNN and RFCN, and it is faster than the RCNN and RFCN accuracies. Reference [48] also proposed a context-aware single-shot multi-frame object detector (CSSD) based on a traditional SSD. In addition, the reference combined a high detection accuracy and real-time speed and experimentally demonstrated that CSSD can achieve better detection efficiency, especially when detecting small objects. The literature [4] proposed a multilevel feature fusion method for small object detection. The experimental results show that the mAP obtained by the multilevel feature fusion method on PASCAL VOC2007 is 2–3% higher than the baseline SSD, especially for detecting small objects. Their speeds are 43 FPS and 40 FPS, which are better than the 29.4 and 26.4 FPS of the latest DSSD. Reference [24] proposed a novel method based on pyramids and an SSD, called the function single shot detector. In the Pascal VOC 2007 test, the network can reach an 82.7 mAP with an input size of 300 × 300 and a speed of 65.8 FP. The performance on other test sets is also better than those of the other datasets. With multiscale context fusion technology, reference [7] proposes the WeaveNet method, which greatly simplifies the structure. The PASCAL VOC 2007 and PASCAL VOC 2012 tests show that WeaveNet brings vast performance improvements.

**DP-SSD framework**

To suppress the detection frame that eliminates the number of repeated times throughout the process of optimizing the computing power of a large amount of data, we propose an optimization algorithm using the SSD network topology. The algorithm is a deep learning algorithm based on a feedforward CNN optimized computing power, which arbitrarily identifies the picture boundary according to the size of the target picture. We used a feature extractor to extract features from network images and compared the VGG-16 network with a convolutional neural network (CNN) in the experiment. We designed a new feature recognition framework to obtain two completely different feature detectors. As shown in Fig. 2 compared with the traditional convolutional neural network (CNN), it has a stronger feature recognizability.

In Fig. 2 we use image features generated from different spatial and different resolution layers to construct a new feature extractor. Due to the irregular pyramid structure’s characteristics of multilayer images with different sizes and resolutions, it has extremely high robustness requirements for network feature recognizabilities. Based on this, we construct several convolutional feature layers at the network endpoint of each feature map detector of different scales. Due to the different characteristics of the characteristic layer of each convolutional network, the design refers to the output result of the optimized calculation of the matched convolutional network filter on the prediction set. There are c channels in the feature recognition, and the resolution of the image feature is designed as an m × n feature layer. Therefore, the calculation kernel of the 3 × 3 × c convolutional network is used for classification calculation, otherwise the coordinate system of the network feature recognition of the alternating frame is replaced.
Accordingly, the output function of the calculation result of the $m \times n$ convolutional neural network prediction model is in the corresponding data result.

**Training model**

First, automatically identifying and labeling the ship tag information is carried out on the target ship picture recognized by the system. Second, the SSD network model is used to input the image with the ship information label for the feature recognizer. Finally, we discuss the output of the detection results after the recognition of the computing power. When the corresponding relationship is clearly specified, the loss function and reverse compensation can be performed on the data point-to-point calculation method Fig. 3.

Network training is divided into three steps:

\[ X_{ij}^p = \{0, 1\} \]

where $X_{ij}^p = 1$ indicates that the $i$-th default box matches the $j$-th real label box of category $P$. According to this matching strategy, $\sum_i X_{ij}^p \geq 1$ is obtained, at least one can match the $j$-th real label box. The overall target loss function is the weighted sum of confidence loss and position loss and can be expressed by:

\[ L(x, c, b, g) = \frac{1}{N} \left( L_{conf}(x, c) + \alpha L_{loc}(x, b, g) \right) \]  

(1)

where $x$ denotes whether the matched box belongs to class $P$ and $c$ is the confidence level. If $x$ is 1, the default box is positive; otherwise, $x$ is 0, and the default box is negative. $b$ indicates the predicted box, and $g$ denotes the truth box. $N$ is the total number of matching default boxes. When $N = 0$, the loss is considered to be 0.

\[ L_{conf}(x, c) = -\sum_{i \in Pos} x_i^p \log(c_i^p) - \sum_{i \in Neg} \log(\hat{c}_i^0) \]  

(2)

where $c_i^p$ represents the probability that the $i$-th default box belongs to $P$, which can be defined as:

\[ \hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)} \]  

(3)

The position loss is the smoothing loss between the prediction box and the true label value box parameters.

\[ L_{loc}(x, l, g) = \sum_{i \in Pos} \sum_{m \in \{cx, cy, w, h\}} x_{ij}^k \text{smooth}_L_1(l_i^m - \hat{g}_j^m) \]  

(4)

where $l_i^m$ denotes the offsets of the predicted box relative to the default box, $\hat{g}_j^m$ represents the truth box matching the
The smoothing loss is given as:

$$\text{smooth}_{L_1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases}$$

where $c_x, c_y$ are the coordinates of the $j$-th sea truth box, and $w, h$ are the width and height of the $i$-th default box, respectively. $d_i^w, d_i^h$ is the width and height of the $i$-th default box, respectively.

Step 2: to reduce the high demand for computing power of the ship identification system in the complex marine environments, the calculation process is greatly reduced, besides the computer’s feature recognition method for the GPU. This method makes the robust optimization and the calculation results of the training set of pictures of different scales to obtain the best results. At the same time, the experimental results of Hariharan et al. [13] show that in the special image of the highly guided lower layer, the segmentation ratio of the image is increased, and the result is output. Similarly, the global text smoothing feature recognition method can be used to optimize the results. Using the settings of each scene of the default frame in the SSD framework, the prediction results of feature maps of different sizes and default frames of different aspect ratios, a prediction set containing different sizes and shapes of input objects can be obtained. For example, (for example, in Fig. 4a), the green box of the ship matches the 4 × 4 feature map. The red box matches the 8 × 8 feature map, as shown in Fig. 4b and c. However, sometimes the default box cannot match the corresponding object. This is because the default box has different sizes, but it does not match the ship’s box, and thus, it is considered a negative sample during training.

Step 3: we use the ship target identification label of real experimental data to take part of the default frame in the center as sampling samples for training, which leads to a serious imbalance of the sample test results during training. This also leads to a decrease in the detection rate of marine ship target monitoring results. We select the output result of 3:1 with the first test result in the sample set and the ratio of positive and negative samples for comparison, rather than all of the samples in the overall dataset.

Our experiments show that in the process of ship target recognition training in a complex maritime navigation environment, different models are used to output robust image sizes and shapes. The graphics training method is used for randomly sampling to solve different image data enhancements. If the label information of the real ship target picture is used as the sampling block diagram of the center frame, the sampling overlap model is retained. At the same time, to ensure the robustness and reliability of the feature recognizability, we adopt a uniform network and a fix sample resolution value, and then flip and compare the results with a probability level of 0.5 to enhance the final detection result.

**Proposed algorithm**

The DP-SSD network algorithm is used to detect ships in the following steps:

---

Fig. 4 Schematic diagram of SSD. a Real image with boxes, b Feature map with $4 \times 4$, c Feature map with $8 \times 8$
Step 1: similar to SSD, the training method classifies different default boxes into positive samples (objects) and negative samples (background).
Step 2: by calculating the IOU values of the default box and true box (truth value box), the label with IOU > 0.5 is a positive sample. Then, the label of the default box is the label of the truth box with the largest IOU of this default box, and IOU < 0.5 are marked as negative samples.
Step 3: this is caused by the classification vector setting a background, making negative samples able to participate in the loss calculation of the category, though they cannot participate in the loss calculation of the coordinate regression.

In summary, we first select the timeliness and accuracy of the appearance of the SSD that can be detected by the characteristics of the SSD algorithm. Accordingly, the DP-SSD is a new feature recognition method that uses a basic network structure, network feature extraction method and optimized multilayer network computing architecture. This is a brand-new feature recognition method for automatic identification of marine target features.

**Experiment**

**Data selection**

We use the DP-SSD model to measure and evaluate Kaggle’s ship dataset. We adopt experimental scenes at different times in the same year and under different complicated maritime navigation environments. For example, Shanghai, Fuzhou City, Fujian Province, and Zhoushan City, Zhejiang Province have a total of more than 10 voyages, collecting more than 50 video sets and more than 2000 experimental pictures. The test results of the detailed datasets are as follows:

First, the Kaggle ship dataset is used to test the feature recognition of more than 1000 ships, including special scenes with different backgrounds and complex environments, such as multitarget ships, coral reefs, lighthouses, etc. Our experiment adaptively selects more than 1000 multitarget and special scene pictures from a total of 25 video detection results in 5 voyages.

We test the source code pixel by pixel based on the results of ship feature recognition in complex maritime scenes constructed by the SSD model. Our experimental results show that the maritime ship identification method constructed by the SSD method has the relative limitations of the real frame boundary. To avoid this situation, we convert the mask to a frame structure.

According to this, all ship targets can be placed in the framework of the model identified by SSD, as shown in Fig. 5.

**Evaluation indicators**

Our article reveals a new method of ship target recognition and evaluation in complex maritime navigation scenarios. First, we identify whether the sailing scenes of sea ships are true or false in different situations. If any target other than the ship is detected, it is considered a false alarm. If the ship target is not found, the ship picture is not output, and the value range is defined as a false negative. If the ship monitoring system does not find or output a picture with ship characteristics, it is defined as a true negative value. Among them, true positives are defined as TP; false-positives are defined as FP; false negatives are denoted as FN; true negatives are denoted as TN. The measurement results show that there are huge measurement differences in different complex environments at sea in different scenarios. We use different indicators, such as accuracy, recall, specificity, F1 measurement, and F2 measurement, to measure the experiment. The measurement results are as follows: (see Table 1).

Our measurement results propose key evaluation indicators for the fourth part, which can accurately measure the performance of DP-SSD target feature recognition and the adaptability of different scenarios in complex environments. Finally, we use methods, such as accurate recall histograms, mAP, FPS and running time to measure and evaluate the evaluation algorithm of the ship target recognition method in a complex navigation environment.

In summary, the fourth part proposes the key performance indicators to evaluate all of the algorithm models, such as the accurate recall histogram, mAP, FPS and running time.

**Implementation**

We use the Python programming language for measurement and evaluation, NVIDIA GeForce RTX 2080Ti graphics card to measure and evaluate ship target recognition data in complex sailing environments, and Intel (R) Core (TM) i9 9900 KF CPU @ 3.6 GHz processor for experimental measurement. All data are shown in parameter Table 2.

Our experiment is based on the training dataset, which has the characteristics of the limitation of the number of sampling targets. Based on this, we use the process of repeated training and sampling, apply the mean method to train the model to improve the reliability of ship feature recognition and reduce the number of overfits, and use different data processing theories to conduct experiments and evaluations, as follows:

- Random sampling, different scenarios;
- Adjust all pictures to 300 × 300, 320 × 320, 416 × 416, 448 × 448, 512 × 512, 544 × 544;
- Random sampling;
- Finally, the results of multiple sets of experiments are evaluated.
Detection efficiency

We use Faster RCNN (VGG16), YOLO, YOLOv2, YOLOv3, SSD300, SSD512, RefineDet320, RefineDet512, and SSD (DP-SSD300 and DP-SSD512) for measurement and evaluation, as shown in Fig. 6.

The experimental results show that the method of ship target recognition in a complex sailing environment using SSD has the characteristics of classifying and regressing target pictures, while also having the ability to recognize picture pixels. The method of ship target recognition in a complex navigation environment has the strongest detection efficiency for the hardware operating environment and has the ability to identify ship targets in different scenarios, as well as in complex navigation environments. Accordingly, this method is more accurate and effective than two-stage detection methods, such as Faster RCNN.
Table 1 Performance indicators for detection

| Indicators             | Abbreviation | Definition                                      |
|-----------------------|--------------|-------------------------------------------------|
| Precision             | Prec         | \( Prec = \frac{TP}{TP+FP} \)                   |
| Average precision     | AP           | \( AP = \frac{\sum Prec}{TP+FN} \)              |
| Mean average precision| mAP          | \( mAP = \frac{\sum AP}{Classes} \)             |
| Recall                | Rec          | \( Rec = \frac{TP}{TP+FN} \)                    |
| Specificity           | Spec         | \( Spec = \frac{TN}{FP+TN} \)                   |
| F1-measure            | F1           | \( F1 = \frac{2 \times Prec \times Rec}{Prec + Rec} \) |
| F2-measure            | F2           | \( F2 = \frac{5 \times Prec \times Rec}{4 \times Prec + Rec} \) |

Table 2 Parameters of implementation platforms

| Indicators          | Parameters       |
|---------------------|------------------|
| GPU                 | GeForce RTX 2080Ti |
| CPU                 | Intel (R) Core (TM) i9 9900KF |
| Clock speed         | 3.6 G Hz         |
| RAM                 | 32 GB            |
| Operation system    | 64bits Windows 10 |
| ML platform         | TensorFlow       |
| Versions            | 2.0              |

As shown in the Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 on the right side of the above figure indicate that the recall rate is 10–100%, and the step size is 10%. Legend 1 indicates that the recall rate is 10%, 2 indicates that the recall rate is 20%, and 10 indicates that the recall rate is 100%. Therefore, the target detection characteristics of the SSD, which has three backbones, are more accurate than those of the other methods in the ship target recognition method in a complex navigation environment. However, they are slightly lower than the performance characteristics of the two-stage target recognition method.

Results and analysis

Real-time detection of ships is a very difficult task because ships have different shapes and video monitoring angles. During the test, due to the effects of ship roll, pitch, and strong and weak lighting, the test results also affect the speed and real-time detection results of the ship. In this article, the tested SSD framework has a high accuracy, a low computational complexity, and a fast speed. To explore the power of the SSD method, multiple sets of experiments were also performed, including Faster RCNN (VGG16), YOLO, YOLOv2, YOLOv3, SSD300, SSD512, RefineDet320, RefineDet512, and SSD (DP-SSD300 and DP-SSD512). Finally, the SSD method was tested using image and video data from Shanghai, Zoushan City, Zhejiang Province, and Fuzhou City, Fujian Province, to compare and evaluate real-time ship monitoring data of Faster RCNN (VGG16), YOLO, YOLOv2, YOLOv3, SSD300, SSD512, RefineDet320, RefineDet512, and SSD (DP-SSD300 and DP-SSD512).

As shown in Fig. 7a, the average recognition precision indicates that the blue legend of DP-SSD is slightly better than Faster RCNN (VGG16), but the difference is only approximately 5%. The red legend of DP-SSD300 indicates the recognition frame rate is up to 54.49, and it is significantly better than those of DP-SSD512 and Faster RCNN (VGG16). In Fig. 7b, the average recognition precision displayed with the blue legend of YOLO is up to 90.5%, which is significantly better than DP-SSD300’s 76.45% and DP-SSD512’s 77.97%. The recognition frame rate of YOLO is between DP-SSD300 and DP-SSD512, at up to 42.33. YOLO v2 is roughly the same as DP-SSD in average recognition precision, but the recognition frame rate is better than that.
Fig. 7 The mAP and FPS histogram for all algorithm models

(a) DP-SSD vs. Faster RCNN (VGG16)
(b) DP-SSD vs. YOLO
(c) DP-SSD vs. YOLOv2
(d) DP-SSD vs. YOLOv2 544×544
(e) DP-SSD vs. YOLOv3
(f) DP-SSD vs. SSD300
(g) DP-SSD vs. SSD512
(h) DP-SSD vs. RefineDet320
(i) DP-SSD vs. RefineDet512

of DP-SSD, as shown in Fig. 7c. In Fig. 7d, YOLO v2 544 × 544 is roughly the same as DP-SSD in average recognition precision, but the recognition frame rate is between DP-SSD300 and DP-SSD512. The average recognition precision of YOLO v3 reaches 85.11%, which is better than those of DP-SSD300 and DP-SSD512. However, the mAP of YOLO v2 reaches 74.21%, which is slightly weaker than that of DP-SSD300 and much stronger than that of DP-SSD512. As shown in Fig. 7f, SSD300 is slightly smaller than DP-SSD300 and DP-SSD512 in terms of average recognition precision, while FPS is up to 58.79, proving to be better than DP-SSD. In Fig. 7g, the mAP of SSD-512 is close to that of DP-SSD, and the FPS of SSD-512 is slightly higher than that of DP-SSD512 but obviously below that of DP-SSD300. In Fig. 7h, the mAP of RefineDet320 can reach 76.81%, and it is comparable to DP-SSD. However, the FPS of RefineDet reaches 46.81, which is significantly superior to DP-SSD512 but slightly lower than DP-SSD300.
In Fig. 7i, the average recognition precision of RefineDet512 can be comparable to DP-SSD, reaching 77.71%, while the FPS of RefineDet512 is 29.46, and it is much smaller than that of DP-SSD300 and slightly larger than that of DP-SSD512.

**Precision comparison**

It can be clearly seen from the results that DP-SSD is considerably better than the other method in terms of target detection performance, whether it is the average accuracy or the detection speed. The root cause analysis is due to the multiscale network proposed in this paper. At the same time, it also has the ability to detect multilevel features, which makes the final detection result more accurate and more efficient.

**Comparison with RefineDet**

The target detection of the RefineDet method requires multiple stages for achievement. It first generates several detection frames at the same time, then performs statistical regression classification on the detected target pictures, and finally determines the target recognition result.

Because the RefineDet method is very cumbersome in the process of ship detection, the proposed scheme has comparative advantages over the former in terms of detection accuracy and time.

**Detection and identification**

The confidence test of the ship area is shown in Fig. 8. The different methods and even the different scales in the same method can have a considerable impact on the average accuracy of detection. When the confidence level of the ship recognition area is easy, the average recognition precision of DP-SSD300 and DP-SSD512 proposed in this paper are both up to 90.13%, the performance is superior to 89.21% of YOLO v2, 88.97% of YOLO v2 544 × 544, 87.38% of YOLO v3, 88.15% of SSD300, 76.22% of SSD512 and 89.79% of RefineDet512, and slightly lower than YOLO’s 91.99% and 90.37% of RefineDet320. However, when the confidence level is more moderate and harder, the performance of our proposed models cannot achieve the desired results.

Regarding the complexity of all of the algorithms, as illustrated in Fig. 9 the running time of DP-SSD is approximately 0.3 s, which is less than 0.37 s of YOLO, 0.7 s of SSD512 and RefineDet320, though it is slightly better than the 0.34 s of RefineDet512.

The difference between the object and the external environment gives the ship the opportunity to miss the detection situation. In response to this problem, the DP-SSD network we propose adopts a variety of feature combinations in the learning stage to try to make the learning network converge to a target. The environmental differentiation has a robust combination of features to achieve the generalization of testing.

In addition, the multiscale, multi-shape convolution allows the method proposed in this article to obtain more differentiated features, from contours to textures.

The choice of filter is of great importance to the ship’s target. It is manifested at least at the level of connection speed and feature extraction. Specifically, lower-level data features can reflect direct features, such as geometric features, and the higher the level High is, the more hidden the characteristics reflected by the data characteristics are.

The DP-SSD method for real-time ship detection is shown in Fig. 10. When the proposed scheme can correctly detect the ship wave object, the target frame is green. It can be seen from the figure that our method is capable of operating at different distances and scales. We identify single or even multiple targets, which is more feasible for the scheme proposed in this article.

**Conclusions and future work**

In this research, the authors propose a neural network named DP-SSD. In the feature extraction training stage, it uses different sizes of extraction boxes to learn from the sample training set. In the sample testing stage, it uses video. The sample combined with the picture sequence is used as the test set. From the perspective of calculation time, processed frame rate, and recognition accuracy, a comprehensive evaluation of the three dimensions shows that the former is better than the classic algorithm for comparison. As an important part of future intelligent ship automatic identification edge platforms, this research is particularly of theoretical value and practical engineering.
What needs to be emphasized here is that due to the limitation of the number of data samples collected in this study, sample differences and the limitations of the equipment itself, there is still a certain gap between the feasibility of the project and the actual application of the scene. At the same time, the external environment of the terminal sensor is also impossible. Completely consistent, this will cause the performance results to have short-term stability and long-term time-varying characteristics. However, the detection of ship
Fig. 10 Ship target detection results

(a) Single target ship detection

(b) Multitarget ship detection

(c) Multiangle target ship detection. The first is the original figure
contour features is time-invariant and robust at the same time. This is the guarantee of success and the weight of the feature. For factors, the future will be further adjusted from the perspective of weights, giving the algorithm stronger time-invariant characteristics and robustness.

In the future, hardware that improves detection accuracy and efficiency will be further developed, and it will be cheaper and more stable. In addition, recommendations will be made to the IMO and the IALA for the establishment of demonstration projects Fig. 11.

Acknowledgements The authors thank Maritime Safety Administration of the People’s Republic of China for providing the raw dataset and sharing the source code. The authors would also like to thank Wuhan Xingtu Electronic Co., Ltd for providing powerful edge-computational machine to verify our ideas.

Author contributions XW and JL designed the study and performed the experiments; XW, XL and ZL performed the experiments, OIK and JJ analyzed the data, XW and JJ wrote the manuscript, and JJ, QO revised the manuscript.

Funding This work was partially supported by the Wuhan Science and Technology Plan Enterprise Technology Innovation Project under Grant No. 2019010702011291, the National Key Research and Development Plan under Grant No. 2018YFC1407404, the National Natural Science Foundation of China under Grant No. 51809207 and No. 52171351 and the Fundamental Research Funds for the Central Universities No. 2019IVB040.

Availability of data and materials The datasets generated and/or analyzed during the current study are available from the first author on specific requests.

Code availability The source codes generated and/or analyzed during the current study are available from the first author on specific requests.

Declarations

Conflict of interest The authors have no conflicts of interest to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Ahmed MS (2019) Application of RFID systems to collision avoidance. Electron Eng 9:1–8. https://doi.org/10.5923/j.eee.20190901.01
2. Aloig G, Loscri V, Borgia A, Natalizio E, Costanzo S, Pace P, Di Massa G, Spadafora F (2011) Software defined radar: synchronization issues and practical implementation. In: Proceedings of the 4th international conference on cognitive radio and advanced spectrum management - CogART ’11. ACM Press, New York, NY, USA, pp 1–5
3. Biswas P, Chakraborty M, Bera R, Shome S (2021) Ensuring reliability in vehicular collision avoidance using joint RFID and radar-based vehicle detection. In: Chakraborty M, Jha RK, Balas VE, Sur SN, Kandar D (eds) Lecture notes in electrical engineering. Springer, Singapore, pp 99–105
4. Cao G, Xie X, Yang W, Liao Q, Shi G, Wu J (2018) Feature-fused SSD: fast detection for small objects. In: SPIE 10615, ninth international conference on cognitive and image processing (ICGIP 2017). SPIE, Qingdao, China, p 106151E
5. Chang YL, Anagaw A, Chang L, Wang Y, Hsiao CY, Lee WH (2019) Ship detection based on YOLOv2 for SAR imagery. Remote Sens 11:786. https://doi.org/10.3390/rs11070786
6. Chen D, Manning C (2014) A fast and accurate dependency parser using neural networks. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). Association for Computational Linguistics, Doha, Qatar, pp 740–750
7. Chen Y, Li J, Zhou B, Feng J, Yan S (2017) Weaving multi-scale context for single shot detector. arXiv:171203149
8. Dalal S, Khalaf OI (2021) Prediction of occupation stress by implementing convolutional neural network techniques. J Cases Inf Technol 23:27–42. https://doi.org/10.4018/jcit.2021070103
9. Debatty T (2010) Software defined RADAR—a state of the art. In: 2010 2nd international workshop on cognitive information processing. IEEE, Elba, Italy, pp 253–257
10. Fu CY, Liu W, Ranga A, Tyagi A, Berg AC (2017) DSSD: deconvolutional single shot detector. arXiv:1701.06659
11. Garmatyuk D, Schuering J, Kauffmann K (2011) Multifunctional software-defined radar sensor and data communication system. IEEE Sens J 11:99–106. https://doi.org/10.1109/jsen.2010.2052100
12. Gaucci J, Zammit-Mangion D, Sabatini R (2012) Correspondence and clustering methods for image-based wing-tip collision avoidance techniques. In: 28th international congress of the aeronautical sciences (ICAS 2012). International Council of the Aeronautical Science, Brisbane, Australia, pp 1–13
13. Lu L, Pillai TS, Gopalarmishnan H, Arpaci-Dusseau AC, Arpaci-Dusseau RH (2017) Wiskey: Separating keys from values in SSD-conscious storage. ACM Transactions on Storage (TOS) 13(1):1–28
14. Hatipoglu N, Bilgin G (2014) Classification of histopathological images using convolutional neural network. In: 2014 4th international conference on image processing, tools and applications (IPTA). IEEE, Paris, France, pp 1–6
15. Hinton GE, Srivastava N, Krizhevsky A (2012) Improving neural networks by preventing co-adaptation of feature detectors. arxiv:abs/12070580
16. Hong SJ, Han Y, Kim SY, Lee AJ, Kim G (2019) Application of deep-learning methods to bird detection using unmanned aerial vehicle imagery. Sensors (Basel, Switzerland) 19:1651. https://doi.org/10.3390/s19101651
17. Jeong J, Park H, Kwak N (2017) Enhancement of SSD by concatenating feature maps for object detection. arXiv:170509587
18. Kim YH, Kim DG, Han JW, Song KH, Kim HN (2017) Analysis of sensor-emitter geometry for emitter localisation using TDOA and FDOA measurements. IET Radar Sonar Navig 11:341–349. https://doi.org/10.1049/iet-rsn.2016.0314
19. Lee SJ, Roh MI, Oh MJ (2020) Image-based ship detection using deep learning. Ocean Syst Eng 10:415–434. https://doi.org/10.12989/oase.2020.10.4.415
20. Li G, Liu F, Sharma A, Khalaf OI, Aloataibi Y, Alsufyani A, Alghamdi S (2021) Research on the natural language recognition method based on cluster analysis using neural network. Math Probl Eng 2021:1–13. https://doi.org/10.1155/2021/9982305
21. Li L, Zhang S, Wu J (2019) Efficient object detection framework and hardware architecture for remote sensing images. Remote Sens 11:2376. https://doi.org/10.3390/rs1102376
22. Mingbo Li (2019) Overview of object detection algorithms based on machine learning. Technol Inf 006:154–155
23. Li Q, Ji H (2014) Incremental joint extraction of entity mentions and relations. In: Proceedings of the 52nd annual meeting of the association for computational linguistics (volume 1: long papers). Association for Computational Linguistics, Baltimore, Maryland, pp 402–412
24. Li Z, Zhou F (2017) FSSD: feature fusion single shot multibox detector. arXiv:171200960
25. Liu M, Jiang J, Wang Z (2019) Colonic polyp detection in endoscopic videos with single shot detection based deep convolutional neural network. IEEE Access 7:75058–75066. https://doi.org/10.1109/access.2019.2921027
26. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, Berg AC (2016) SSD: single shot multibox detector. In: Leibe B, Matas J, Sebe N, Welling M (eds) Computer vision – ECCV 2016. Springer International Publishing, Cham, pp 21–37
27. Liu Y, Guo F, Yang L, Jiang W (2016) Source localization using a moving receiver and noisy TOA measurements. Signal Process 119:185–189. https://doi.org/10.1016/j.sigpro.2015.07.029
28. Luo HL, Chen HK (2020) Survey of object detection based on deep learning. Acta Electronica Sinica 48(6):1230–1239
29. Masiero A, Fissore F, Guarnieri A, Pirotti F, Vettore A (2015) UAV positioning and collision avoidance based on RSS measurements. Int Arch Photogramm Remote Sens Spatial Inf Sci XL-1/W4 219–225. https://doi.org/10.5194/isprsarchives-xl-1-w4-219-2015
30. Pennington J, Socher R, Manning C (2014) Glove: global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). Association for Computational Linguistics, Doha, Qatar, pp 1532–1543
31. Ping SL, Qiang D (2020) A survey of research on image target recognition based on deep learning. Command Control Simul 41:1–5
32. Prasoon A, Petersen K, Igel C, Lauze F, Dam E, Nielsen M (2013) Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. Medical image computing and computer-assisted intervention—MICCAI 2013—16th international conference. Springer, Berlin, Heidelberg, pp 246–253
33. Rajasoundaran S, Prabu AV, Subrahmanyam JBV, Rajendran R, Sateesh Kumar G, Kiran S, Khalaf OI (2021) Secure watchdog selection using intelligent key management in wireless sensor networks. Mater Today Proc. https://doi.org/10.1016/j.mtpr.2020.12.027
34. Ribeiro E, Uhl A, Hafner M (2016) Colonic polyp classification with convolutional neural networks. In: 2016 IEEE 29th international symposium on computer-based medical systems (CBMS). IEEE, Belfast and Dublin, Ireland, pp 253–258
35. Sato Y, Shimonaka Y, Maruoka T, Wada T, Okada H (2007) Vehicular collision avoidance support system v2 (VCASSv2) by GPS+INS hybrid vehicular positioning method. In: 2007 Australasian telecommunication networks and applications conference. IEEE, Christchurch, New Zealand, pp 29–34
36. Sengan S, Rao GRK, Khalaf OI, Babu MR (2021) Markov mathematical analysis for comprehensive real-time data-driven in healthcare. Math Eng Sci Aerosp (MESA) 12:77–94
37. Sengan S, Sagar PV, Ramesh R, Khalaf OI, Dhanapal R (2021) The optimization of reconfigured real-time datasets for improving classification performance of machine learning algorithms. Math Eng Sci Aerosp (MESA) 12(1):43–54
38. Shi W, Bao S, Tan D (2019) FFESSD: an accurate and efficient single-shot detector for target detection. Appl Sci 9:4276. https://doi.org/10.3390/app9204276
39. Tajbaksh N, Gurudu SR, Liang J (2015) A comprehensive computer-aided polyp detection system for colonoscopy videos. In: Proceedings of the 24th international conference on information processing in medical imaging (IPMI’15). Sabahar Mor Ostag, Isle of Skye, UK, pp 327–338
40. Tajbaksh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J (2016) Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Trans Med Imaging 35:1299–1312. https://doi.org/10.1109/tmi.2016.2535302
41. Tam CK, Bucknall R, Greig A (2009) Review of collision avoidance and path planning methods for ships in close range encounters. J Navig 62:455–476. https://doi.org/10.1016/j.conengprac.2009.11.001
42. Tieleman T, Hinton G (2012) Lecture 6.5-rmsprop: divide the gradient by a running average of its recent magnitude. Coursera Neural Netw Mach Learn 4:26–31
43. Venneri F, Costanzo S, Di Massa G (2013) Design of a reconfigurable reflect array unit cell for wide angle beam-steering radar applications. Adv Intell Syst Comput 206:1007–1013. https://doi.org/10.1007/978-3-642-36981-0_95
44. Wang N, Wang Y, Er MJ (2020) Review on deep learning techniques for marine object recognition: architectures and algorithms. Control Eng Pract. https://doi.org/10.1016/j.conengprac.2020.104458
45. Wang Y, Wu Y (2017) An efficient semidefinite relaxation algorithm for moving source localization using TDOA and FDOA measurements. IEEE Commun Lett 21:80–83. https://doi.org/10.1109/Lcomm.2016.2614936

46. Wisesa O, Andriansyah A, Khalaf OI (2020) Prediction analysis for business to business (B2B) sales of telecommunication services using machine learning techniques. Majlesi J Electr Eng 14:145–153. https://doi.org/10.29252/mjee.14.4.145

47. Wu N, Yuan W, Wang H, Kuang J (2016) TOA-based passive localization of multiple targets with inaccurate receivers based on belief propagation on factor graph. Digit Signal Process 49:14–23. https://doi.org/10.1016/j.dsp.2015.10.013

48. Xiang W, Zhang DQ, Yu H, Athitsos V (2018) Context-aware single-shot detector. In: 2018 IEEE winter conference on applications of computer vision (WACV). IEEE, Lake Tahoe, NV, USA, pp 1784–1793

49. Yu H, Wilamowski BM (2012) Neural network training with second order algorithms. Adv Intell Soft Comput 99:463–476. https://doi.org/10.1007/978-3-642-23172-8_30

50. Yu JS, Chen J, Xiang ZQ, Zou YX (2015) A hybrid convolutional neural networks with extreme learning machine for WCE image classification. In: 2015 IEEE international conference on robotics and biomimetics (ROBIO). IEEE, Zhuhai, China, pp 1822–1827

51. Yu X, Zhao Z, Zhang X (2021) Physical theory of RFID system physical anti-collision. Springer, Singapore, pp 59–108

52. Zhang H, Li L, Wu K (2007) 24GHz software-defined radar system for automotive applications. In: 2007 European conference on wireless technologies. IEEE, Munich, Germany, pp 138–141

53. Zhang S, Wu R, Xu K, Wang J, Sun W (2019) R-CNN-based ship detection from high resolution remote sensing imagery. Remote Sens 11:631. https://doi.org/10.3390/rs11060631

54. Zhang T, Zhang X (2019) High-speed ship detection in SAR images based on a grid convolutional neural network. Remote Sens 11:1206. https://doi.org/10.3390/rs11101206

55. Zhao JH, Zhang XG, Yang L (2020) Ship detection in remote sensing based on deep learning. Sci Surv Mapp 45:110–116

56. Zhu GH, Feng DZ, Xie H, Zhou Y (2016) An approximately efficient bi-iterative method for source position and velocity estimation using TDOA and FDOA measurements. Signal Process 125:110–121. https://doi.org/10.1016/j.sigpro.2015.12.013

57. Zhuang M, Tan L, Li K, Song S (2021) Fixed-time position coordinated tracking control for spacecraft formation flying with collision avoidance. Chin J Aeronaut 34:182–199. https://doi.org/10.1016/j.cja.2020.12.024

58. Zou Y, Li L, Wang Y, Yu J, Li Y, Deng WJ (2015) Classifying digestive organs in wireless capsule endoscopy images based on deep convolutional neural network. In: 2015 IEEE international conference on digital signal processing (DSP). IEEE, Singapore, pp 1274–1278

59. Zsedrovits T, Zarandy A, Vanek B, Peni T, Bokor J, Roska T (2011) Collision avoidance for UAV using visual detection. In: 2011 IEEE international symposium of circuits and systems (ISCAS). IEEE, Rio de Janeiro, Brazil, pp 2173–2176

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.