

Ship Recognition and Tracking System for Intelligent Ship Based on Deep Learning Framework

B. Liu, S.Z. Wang, Z.X. Xie, J.S. Zhao & M.F. Li
Merchant Marine College, Shanghai Maritime University, Shanghai, China

ABSTRACT: Automatically recognizing and tracking dynamic targets on the sea is an important task for intelligent navigation, which is the prerequisite and foundation of the realization of autonomous ships. Nowadays, the radar is a typical perception system which is used to detect targets, but the radar echo cannot depict the target's shape and appearance, which affects the decision-making ability of the ship collision avoidance. Therefore, visual perception system based on camera video is very useful for further supporting the autonomous ship navigational system. However, ship's recognition and tracking has been a challenge task in the navigational application field due to the long distance detection and the ship itself motion. An effective and stable approach is required to resolve this problem. In this paper, a novel ship recognition and tracking system is proposed by using the deep learning framework. In this framework, the deep residual network and cross-layer jump connection policy are employed to extract the advanced ship features which help enhance the classification accuracy, thus improves the performance of the object recognition. Experimentally, the superiority of the proposed ship recognition and tracking system was confirmed by comparing it with state of-the-art algorithms on a large number of ship video datasets.

1 INTRODUCTION

With the rapid development of information technology such as big data, artificial intelligence and deep learning, the shipbuilding industry is moving towards informatization and intelligence (Alexander 2011, Russell 2010, Lecun 2015). Intelligent ship visual perception is the premise and foundation for unmanned navigation (Zhang 2010). It can obviously reduce the marine traffic accidents caused by human factors, optimize the ship route, reduce the fuel consumption, reduce the cost of ship operation, and improve the safety of ship navigation. Ship recognition and tracking is an indispensable part of intelligent ship visual perception. It can identify dangerous ship types, monitor the surrounding environment and make reasonable ship collision

avoidance decisions, which is of great significance to the promotion and development of intelligent ships.

In recent years, in order to deal with the visual perception challenges of intelligent ships, research institutions and scholars have processed the visual perception data under the background of intelligent ships (Johansson 1973). Ship detection, ship tracking and ship type recognition have been extensively studied. In ship detection, the emergence of visual mechanism provides a good research idea for detecting surface targets based on visible video sequences. Kim et al. proposed an adaptive focusing region of interest detection algorithm, and achieved ideal detection results (Kim et al. 2015). Li accelerated class detection and recognition by sharing convolution neural network, which provided a new

idea for water surface target detection (Li et al. 2016). Some scholars excavated the video containing water targets according to the learning mechanism of large perspective, and generated the most possible set of water surface targets according to the maximum likelihood probability method (Tjampens et al. & Zou et al. 2016). Some scholars use multi-view method to extract multiple features of water targets (such as texture features, structural features, color features, etc.). Sparse learning and multi-task learning are used to fuse features, eliminate false targets and retain the detected water targets (Albrecht 2011, Hong 2015, Bergamasco 2016). In ship tracking, the traditional method is to abstract the tracking ship as a particle through the automatic identification system (AIS) and radar (Xiao et al. 2015). Domel et al. applied the correlation filtering algorithm to the tracking field, and use a single gray feature to represent the target for tracking (Bolme et al. 2010). In order to overcome the shortcomings of traditional ship tracking algorithms, Chen et al. proposed a ship moving position tracking algorithm based on support vector machine regression and game theory. Support vector machine was used to estimate the position of the ship to be tracked in order to improve the accuracy of ship tracking position (Chen et al. 2017). Chen et al. fused instruction filter and back stepping method to construct a robust adaptive neural network tracking controller for ship course (Chen et al. 2016). In the aspect of ship type recognition, the above goal is achieved by fusing sensor data information such as self-identification system (AIS) and radar (Robards 2016, Shu 2017, Sang 2015). Jiang et al. also proposed a ship type recognition method based on structural feature analysis, which can effectively extract high-resolution COSMO-SkyMed image features of bulk carriers, container ships and tankers (JIANG et al. 2014). Chen et al. took into account the computational complexity, recognition accuracy and the difference of features extracted by various algorithms in ship type recognition, and used support vector machine algorithm to fuse the ship features extracted by the above operators (Chen et al. 2016). The methods mentioned above have achieved certain results for ship visual perception under certain specific conditions. However, in the era of intelligent ships, higher requirements have been put forward in ship detection, tracking and ship type recognition. It is necessary to detect and recognize small targets of ships by using relevant information collection and sensing technology to determine potential collision risk and help the decision-making system of intelligent ships to determine interested ship targets.

With the increasing scale of maritime traffic and the increasing complexity and diversification of the surrounding environment of ships in voyage, at present, it is mainly through the crew to judge the type of ships around and the state of navigation artificially. There are certain subjective errors, which cannot meet the basic requirements of intelligent ships. In this paper, a visual perception system for intelligent ship is constructed based on the in-depth learning framework. The ship-borne camera can monitor the information around the ship and identify other ship types in real time. This paper improves the shallow network structure and multi-scale prediction method of targets in the traditional deep learning, introduces the idea of residual network (He et al.

2015) effectively overcomes the problems of gradient dispersion and gradient explosion, and improves the ability of data feature learning. The network depth is increased by cross-layer connection, and advanced features of ships are extracted for combination learning. On this basis, target region prediction and classification prediction are integrated into a single neural network model to realize the global information of the image for target recognition. In the case of high accuracy, fast target detection and ship type recognition are realized.

2 INTELLIGENT VISUAL PERCEPTION SYSTEM FOR SHIPS

With the rapid development of computer vision theory and technology (Moeslund et al. 2001), it provides a favorable technical support for data visualization of intelligent ships. Intelligent ships perceive the surrounding environment and their own state through various sensors, and make decisions on the perceived environment so as to realize the auxiliary sailing and active safety of ships, even autonomous sailing. Figure 1, shows the framework of the proposed intelligent ship visual perception system based on computer vision.

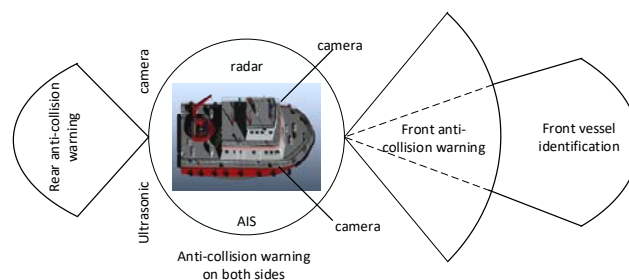


Figure 1. Schematic diagram of ship visual perception system.

Intelligent ship visual perception system collects video image information during navigation by visual sensors installed around ships, and processes it with automatic identification system (AIS) and radar (Merchant et al. 2012). Vessel tracking and recognition in the navigation area is the premise and foundation of the intelligent ship visual perception. By installing cameras on both sides of the ship, the dynamic information of surrounding ships is monitored and tracked. The possible dangerous can be obtained, and the collision risk between the target ship and the own-ship can be further determined. Provide the crew with different degrees of danger signals to help the crew make correct judgments. On this basis, the types of other ships in the situation of intersection and encounter are identified, and the heading speed of the ship is adjusted dynamically and timely to avoid the risk of collision and ensure the safety of navigation. As shown in Figure 2, the visual perception flow chart of a ship is constructed by a ship-borne camera which is used to sense the surrounding environment to obtain the traffic situation near the ship, detect and identify the ship type of another ship.

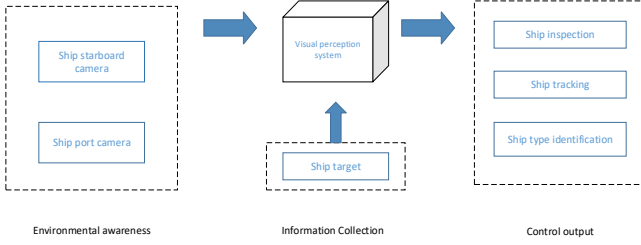


Figure 2. Visual perception flow chart of shipborne camera.

3 TARGET RECOGNITION AND TRACKING FRAMEWORK BASED ON DEEP LEARNING

In recent years, with the introduction of computer vision and deep learning algorithms into the field of target tracking and recognition, great breakthroughs have been made in performance gradually, which provides a new idea for the research of visual perception of intelligent ships. In the visual perception task for maritime traffic, it is necessary to detect ships in the video sequence quickly, efficiently and accurately through the ship-borne camera, and identify the types of ships in real-time to help intelligent ships more accurately judge the collision risk and ensure the safe navigation of intelligent ships.

Traditional deep learning network framework mainly includes input layer, hidden layer and output layer (Xu et al. 2016). The number of network layers is relatively shallow, which cannot meet the basic requirements of intelligent ships. Considering the changes of ship imaging size, illumination, angle of view, overlap of ship imaging and artificial participation in the situation of intersection and encounter, etc. Figure 3 is a network framework based on deep learning structure for intelligent ship tracking and recognition.

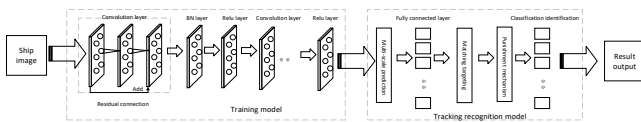


Figure 3. Target recognition and tracking framework based on deep learning.

3.1 Training Model

In order to ensure the recognition accuracy and the stability of the tracking of the visual perception system, Residual structure is employed into the deep neural network model to ensure that the network structure is deep and convergent. The input samples are convoluted to extract the corresponding features, and then combined learning is carried out to get the feature map model of the object, which initializes the follow-up tracking and recognition model.

Convolution layer is the core component of the neural network structure. The number of training parameters of the neural network is reduced by the sharing of receptive fields and weights. In convolutional networks, the latter layer of neurons

extracts the local features of different locations of the former layer of feature map to get the next layer of feature map. In order to overcome effectively the shortcomings of the deep neural network training and accelerate the convergence speed of network training, Batch Normalization (BN) operations are added after each convolution layer to normalize the distribution of input data into a mean value of 0 and a variance of 1.

On this basis, a cross-layer jump connection method is added. By using the residual function $F(u) = H(u) - u$, the layer-by-layer training of the deep neural network structure is changed into stage-by-stage training. The network structure is divided into several sub-segments, each sub-segment contains a relatively shallow number of network layers, and each sub-segment contains a part of the total learning deficit (total loss), which ultimately achieves a relatively small overall loss. According to the training network of ships, the mean square and error are used as loss functions, which are composed of coordinate error, IOU error and classification error. The expressions are as follows:

$$\begin{aligned}
 & \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B \perp_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \\
 & + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B \perp_{ij}^{obj} [(\sqrt{\omega_i} - \sqrt{\hat{\omega}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] \\
 & + \sum_{i=0}^{S^2} \sum_{j=0}^B \perp_{ij}^{obj} (\sqrt{C_i} - \sqrt{\hat{C}_i})^2 \\
 & + \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B \perp_{ij}^{noobj} (\sqrt{C_i} - \sqrt{\hat{C}_i})^2 \\
 & + \sum_{i=0}^{S^2} \perp_{ij}^{obj} \sum_{c \in classes} (p_i(c) - \hat{p}_i(c))^2
 \end{aligned} \tag{1}$$

Among them, the first two lines represent the coordinate error, the first line is the prediction of the center coordinate of the bounding box, the second line is the prediction of the width and height of the bounding box, the third and fourth lines represent the loss of confidence of the bounding box, and the fifth line is the error of the prediction category. If there is no target in a cell, the classification error will not be propagated backward. When the object in the bounding box and the one with the highest IOU in the real frame propagate backward. The rest will not proceed.

3.2 Ship Recognition and Tracking Model

By training the visual perception model obtained by the network, feature maps of a certain size can be obtained from the input image. Drawing on the idea of Yolo algorithm (Redmon J et al. 2015), the input image is divided into corresponding size grids. Each grid prediction priori box (clustered values) on the feature graph contains four predictive values t_x, t_y, w, h of which the first four are coordinates. The process of obtaining b_x, b_y, b_w, b_h from the actual prediction t_x, t_y, w, h is expressed as:

$$b_x = \sigma(t_x) + c_x \tag{2}$$

$$b_y = \sigma(t_y) + c_y \quad (3)$$

$$b_w = P_w \cdot e^w \quad (4)$$

$$b_h = P_h \cdot e^h \quad (5)$$

Among them, c_x and c_y are the number of the first grid in the upper left corner where the central coordinates of the border are located. t_x and t_y are the center coordinates of the predicted border. The σ represents the logistic function, which normalizes coordinates to 0-1. The final b_x and b_y are normalized values relative to the grid position. The width and height of the predicted border are w and h . P_w , P_h are the width and height of the candidate box. The final b_w and b_h are normalized values relative to candidate box positions.

In order to prevent the drift of ship tracking frame and make the moving of target more robust, a penalty mechanism is constructed to process the model features and to represent and learn these features in order to achieve the purpose of ship tracking in video sequence. In order to obtain better tracking effect, we introduce the coordinate prediction value as the cost function. The penalty mechanism is as follows:

$$\text{Min} \sum_K^N f_L(D^K U^K - V^K) + \lambda_1 [P]_1 + \lambda_2 [Q^T]_2 \quad (6)$$

$$U^K = P^K + Q^K \quad (7)$$

For each feature K in the model, D^K is used to represent the ship image tracking sequence, V^K is the feature matrix of the ship feature graph, and U^K is the matrix representation of the sequence D^K . f_L is a cost function, which is used to evaluate the degree of difference between ship target and ship feature graph matrix in the first feature. P^K is the matrix representation of the k -th feature. The global representation matrix P is obtained by filling P^K horizontally, which is similar to the global coefficient matrix Q . The parameter $[P]_1$ represents the independence of each feature of the model. $[Q^T]_2$ represents the abnormal result of model tracking. The parameter λ_1 represents the penalty degree of the global coefficient matrix P , and the parameter λ_2 is the penalty coefficient corresponding to the global coefficient matrix P .

For the recognition of ship types at sea, the spatial distribution of ships overlaps, and the same frame detection corresponds to two different ships. Thus, only one ship type can be identified, resulting in a decline in recognition rate. In this paper, multi-label classification is used to predict the target category, and the logical regression layer of multi-label and multi-classification are added to the network structure. The sigmoid function is used as the logistic regression unit to classify each category. At the same time, the cross-entropy cost function is used to measure the difference between the predicted value

and the actual value of the neural network. The expression is as follows:

$$y = \frac{1}{e^{-x} + 1} \quad (8)$$

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m [T^i \log(f(x^i)) + (1-T^i) \log(1-f(x^i))] \quad (9)$$

Among them, m is the total number of samples, T is the label, with a value of 0 or 1. i represents the ship sample and $f(x)$ represents the predicted output.

4 EXPERIMENTAL ANALYSIS

4.1 Ship Data Image Set

At present, there are fewer ship-related data sets in object detection related data sets, and fewer image sets for merchant ship recognition tasks. Therefore, this study collects five kinds of common merchant ship images as sample sets by means of network search. It mainly includes container ships, bulk carriers, oil tankers, LNG vessels and fishing vessels. In this study, 7402 images were collected, including 2320 images of container ships, 1050 images of tankers, 1140 images of liquefied natural gas vessels, 1860 images of grocery vessels and 1032 images of fishing vessels. 80% images of each ship type are selected as training set, and the remaining 20% images are selected as test set. Figure 4 shows pictures of different ship types.

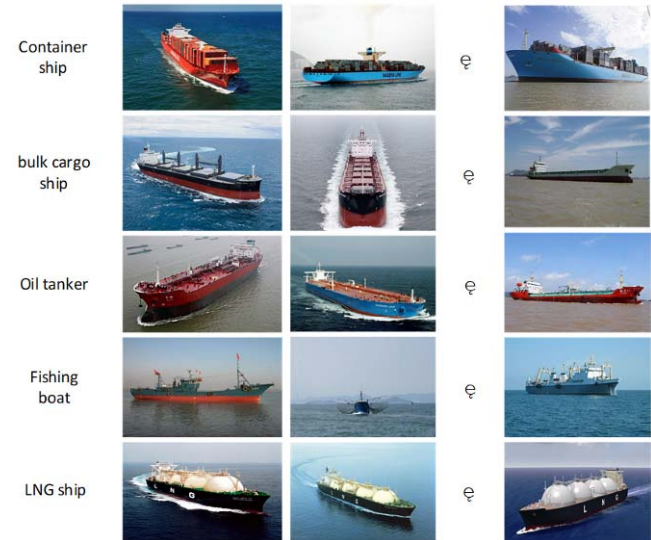


Figure 4. Training sample set of different ship type pictures

4.2 Experimental Platform and Parameter Settings

The experimental platform of this study is Windows 10 operating system, 16G RAM, CPU processor's main frequency is 3.2GHz, GPU is NVIDIA GTX 1050Ti, display memory is 4G, test platform is PyCharm (2018 version). The specifications and parameters of the ship-borne camera are listed in Table 1.

Table 1. Camera Specifications

Pixel	5 million
Lens size	8mm
Monitoring angle	41.5°
Monitoring distance	20-25m

The parameter setting of deep network structure is the main task of network training. According to the idea of transfer learning, the pre-training network framework can be fine-tuned with its own training data on the existing basic network, which can achieve better training effect. This study is based on the pre-training Darknet model. Some parameters are initialized as shown in Table 2.

Table 2. Initialization tuning settings for network structure parameters

Parameter	Initial value
momentum	0.9
decay	0.0005
angle	0
saturation	1.5
exposure	1.5
hue	0.1
learning_rate	0.001
burn_in	1000
max_batches	500200
policy	steps
steps	40000, 45000

4.3 Video Data Source for Ship Monitoring

The experimental data are based on the video data collected by the cameras on both sides of the container ship YUFENG of Shanghai Maritime University. Figure 5 shows the installation position of the camera of the container ship. The collected surveillance video is divided into two groups to evaluate the performance of ship detection algorithm. The first group of ship surveillance video is used to evaluate the detection performance of ship detection model under different traffic conditions based on good navigation environment. The second group of ship surveillance video is based on the foggy navigation environment, which is used to test the robustness and accuracy of ship detection model under very low visibility.



Figure 5. Installation location of ship-borne camera.

4.4 Analysis of experimental results

In order to verify the validity and reliability of the detection, the training pictures contain pictures of various meteorological and environmental scenarios. According to the characteristics of convolution neural network, illumination, sea surface environment and other important factors will be automatically learned by the model. In addition, since batch normalization operation is included in our training process, the generalization ability of the model can be greatly improved, and the effects of different light intensity can be effectively overcome. According to the average loss curve of the number of iterations in the training process shown in Figure 6, it is found that the loss of the type is basically stable around 0.3 when the number of iterations is 12,000. With the increase of the number of iterations, the value of the average loss function remains basically unchanged and tends to be stable. It shows that the algorithm has fast convergence in the training process.

The recall-accuracy curve is a performance index of a classifier, which is used to reflect the accuracy and accuracy of ship type recognition. In this experiment, four common types of ships were selected, namely container ship, bulk carrier, oil tanker and fishing vessel. As shown in Figure 7, the relationship curve between the recall rate and the accuracy of the improved ship type is compared on the basis of the original method. From the experimental data, it can be seen that the area enclosed by the accuracy and recall rate of ship detection in this method is higher than that of the original method, reflecting that the value of AP in the data is obviously larger than that of the original method. The recall rate can reach 85% without loss of precision. When the recall rate reaches 80%, the accuracy can still reach 80%, which fully illustrates the accuracy of this method.

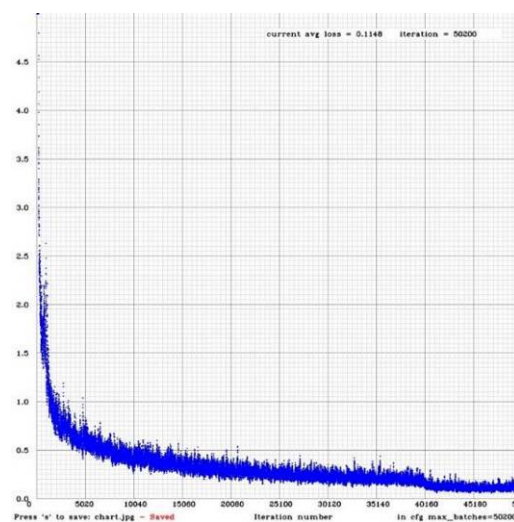
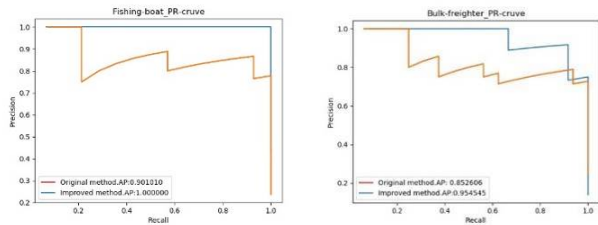
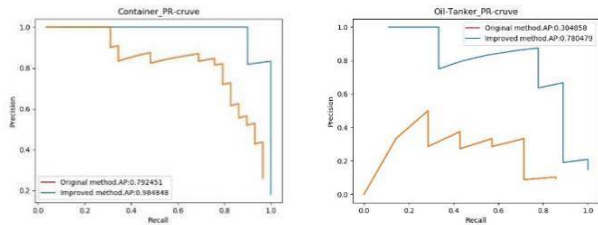


Figure 6. Average loss function curve.



(a) P-R Curve of Fishing Ship. (b) P-R Curve of Bulk-freighter Ship.



(c) P-R Curve of Container Ship. (d) P-R Curve of Oil Tanker.

Figure 7. Contrast of P-R Curves of enhanced visual perception method.

4.5 Contrastive experiments of different methods

In order to illustrate the effectiveness of this method, the recognition performance of this model is further validated by comparing the commonly used deep learning target recognition methods. K-Nearest Neighbor (KNN), Artificial Neural Network (ANN), Traditional Neural Network (CNN) and Deep Convolution Neural Network (DCNN) are used to compare different ship types. Under the conditions of this experiment, the unified agreement that IOU value is greater than 0.75 is the correct target detection. The results are shown in Table 3.

Table 3. Comparison of recognition accuracy of different ship type recognition algorithms.

	KNN	ANN	CNN	DCNN	Proposed method
General cargo ship	34.20%	30.10%	80.00%	86.00%	89.50%
Bulk cargo ship	31.20%	33.10%	71.20%	72.50%	88.20%
Container ship	53.20%	61.20%	85.20%	90.70%	96.80%
LNG ship	42.10%	37.00%	63.20%	66.70%	82.70%
Oil tanker	46.10%	45.30%	79.50%	84.60%	90.50%
average value	40.90%	42.60%	76.70%	81.40%	89.50%

Among the above five types of ship recognition tasks, KNN and ANN algorithms have the lowest recognition accuracy for grocery and bulk carriers, which are 31.2% and 30.1% respectively, while CNN method has the lowest recognition accuracy for LNG ships, which is only 63.2%. The recognition accuracy of DCNN algorithm for bulk carrier and LNG ship is 72.5% and 66.7% respectively, while the recognition accuracy of the two ship types mentioned above is 88.2% and 82.7% respectively, and 96.8% for container ship. The accuracy of ship type recognition based on KNN (ANN) and DCNN (CNN, DCNN) shows that the ship type recognition method based on this model

cannot extract the characteristics of different ship types very well, and the ship type recognition method based on this model can find the depth features of different ship types better, and can obtain better ship recognition effect.

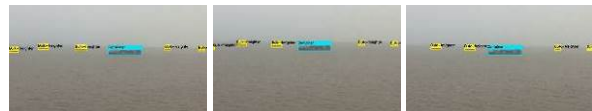
4.6 Experimental results

Key frames are extracted from ship-borne camera surveillance video in different environments and traffic flows to evaluate the detection performance of the algorithm. This system detects the ship in video sequence. From Figure 8, it is shown that in the video scene with good navigation environment, the ship tracking process shows a good tracking effect, and accurately real-time display of the ship type.

In order to verify the robustness and accuracy of the proposed algorithm, ship detection experiments were carried out on ship-borne camera surveillance video during fog navigation. Figure 9 shows that the proposed algorithm can effectively overcome the effects of haze weather and illumination changes, and still track ships and identify ship types in the case of low visibility. On the basis of identifying ship types, ship visual perception tasks are further processed. As shown in Figure 10, various ship types and important parts of ships are accurately identified.



(a) 166 Frame (b) 427 Frame (c) 877 Frame
Figure 8. Tracking and Recognition Results of the System in Video Scene with Good Navigation Environment.



(a) 211Frame (b) 524 Frame (c) 660 Frame
Figure 9. Tracking and Recognition Results of the System Diagram in a Foggy Navigation Environment Video Scene.



Figure 10. Recognition results of ship type and position.

5 CONCLUSION

In this paper, an intelligent ship vision enhancement system based on deep learning framework is proposed to solve the problem of ship tracking and recognition for intelligent navigation visual perception tasks. It effectively overcomes the shortcomings of different illumination, different weather, wind and wave conditions and artificial participation. Future research work will integrate radar, infrared and AIS data to obtain more long-distance marine vessel monitoring and real-time display of the ship's geographical location under poor visual conditions.

ACKNOWLEDGEMENTS

The research presented in this paper has been supported by Shanghai Shuguang Plan Project (15SG44), National Natural Science Foundation of China (51709167), National Science Foundation of Shanghai (18ZR1417100), Shanghai Pujiang Program (18PJ017), and Shanghai Science and Technology Innovation Action Plan (18DZ1206101), and the Young Teacher Training Program of Shanghai Municipal Education Commission (ZZHS18053).

REFERENCE

- [1] Albrecht T, West G A, Tan T, et al. 2011. Visual maritime attention using multiple low-level features and naive bayes classification: Digital Image Computing Techniques and Applications (DICTA), International Conference on [C]. IEEE.
- [2] Bergamasco F, Benetazzo A, Barbariol F, et al. 2016. Multi-view horizon-driven sea plane estimation for stereo wave imaging on moving vessels[J]. *Computers & Geosciences* 95:105-117.
- [3] Bolme D S , Beveridge J R , Draper B A , et al. 2010. Visual object tracking using adaptive correlation filters[J].
- [4] Chen Weiqiang, Chen Jun, Zhang Wei, et al. 2016. Robust tracking control for ship heading adaptive neural network[J]. *Journal of Ship Engineering* (09): 15-20.
- [5] Chen Wenting, Liu Nantong, Ji Kefeng, et al. 2014. Ship Recognition for SAR Image Based on Multi-Classifer Fusion[J]. *Remote Sensing Information* (5): 90-95.
- [6] Chen Xiaojun, Yang Zhangqiong. 2017. Application of Support Vector Regression and Game Theory in Ship Moving Position Tracking [J].*Ship science and technology* (08):19-21.
- [7] He K , Zhang X , Ren S , et al. 2015. Deep Residual Learning for Image Recognition[J].
- [8] Hong Z, Chen Z, Wang C, et al. 2015. Multi-store tracker (muster): A cognitive psychology inspired approach to object tracking: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, [C].
- [9] Hong Z, Mei X, Prokhorov D, et al. 2013. Tracking via Robust Multi-task Multi-view Joint Sparse Representation[J]:649-656.
- [10] JIANG Shaofeng, WANG Chao, WU Fan, et al. 2014. COSMO-SkyMed Image Commercial Ship Classification Algorithm Based on Structural Feature Analysis[J]. *Remote Sensing Technology and Application*.29(4):607-615.
- [11] Johansson G. 1973. Visual perception of biological motion and a model for its analysis[J]. *Perception & Psychophysics* 14(2):201-211.
- [12] Kim D, Kim H, Jung S, et al. 2015. A vision-based detection algorithm for moving jellyfish in underwater environment: Ubiquitous Robots and Ambient Intelligence (URAI), 2015 12th International Conference on [C]. IEEE.
- [13] Lecun Y, Bengio Y, Hinton G. 2015. Deep learning[J]. *Nature* 521(7553):436.
- [14] Li X, Shang M, Hao J, et al. 2016. Accelerating fish detection and recognition by sharing CNNs with objectness learning: OCEANS 2016-Shanghai [C]. IEEE.
- [15] Mayerschönberger V, Cukier K. 2014. Big data: A revolution that will transform how we live, work, and think.[J]. *Mathematics & Computer Education* 47(17):181-183.
- [16] Merchant N D, Witt M J, Blondel P, et al. 2012. Assessing sound exposure from shipping in coastal waters using a single hydrophone and Automatic Identification System (AIS) data[J]. *Marine Pollution Bulletin* 64(7):1320-1329.
- [17] Moeslund T B, Granum E. 2001. A Survey of Computer Vision-Based Human Motion Capture[J]. *Computer Vision & Image Understanding* 81(3):231-268.
- [18] Robards M, Silber G, Adams J, et al. 2016. Conservation science and policy applications of the marine vessel Automatic Identification System (AIS)—a review[J] 92(1):75-103.
- [19] Russell S J, Norvig P. 2010. Artificial intelligence: a modern approach[J]. *Applied Mechanics & Materials* 263(5):2829-2833.
- [20] Sang L, Wall A, Mao Z, et al. 2015. A novel method for restoring the trajectory of the inland waterway ship by using AIS data[J] 110:183-194.
- [21] Shu Y, Daamen W, Ligteringen H, et al. 2017. Influence of external conditions and vessel encounters on vessel behavior in ports and waterways using Automatic Identification System data[J] 131:1-14.
- [22] T'Jampens R, Hernandez F, Vandecasteele F, et al. 2016. Automatic detection, tracking and counting of birds in marine video content: Image Processing Theory Tools and Applications (IPTA), 2016 6th International Conference on [C]. IEEE.
- [23] Xiao F, Han L, Gulijk C V, et al. 2015. Comparison study on AIS data of ship traffic behavior[J]. *Ocean Engineering* 95(3):84-93.
- [24] Xu C, Lu C, Liang X, et al. 2016. Multi-loss Regularized Deep Neural Network[J]. *IEEE Transactions on Circuits & Systems for Video Technology* 26(12):2273-2283.
- [25] Zhang Z, Zhang X W, Liang R Y, et al. 2010. Research and Implementation of Ship-lock Monitoring System Based on SVM and Visual Perception[J]. *Modern Electronics Technique*.
- [26] Zou Z, Shi Z. 2016. Ship detection in spaceborne optical image with SVD networks[J]. *IEEE Transactions on Geoscience and Remote Sensing* 54(10):5832-5845.