Real-Time Ship Monitoring Algorithm and Measurement for Single-Shot Multi-Box Detection and Navigation Based on Deep Learning

Xiang Wang, Jingxian Liu, Zhao Liu and Zhi Yuan

March 31, 2020

Real-time ship monitoring algorithm and measurement for single-shot multi-box detection and navigation based on deep learning

Xiang Wang[a,b], Jingxian Liu [a], Zhao Liu [a,1], Zhi Yuan[a]

[a]*Hubei Key Laboratory of Inland Shipping Technology, School of Navigation, Wuhan University of Technology, 1178 Heping Avenue, Wuchang District, Wuhan, Hubei Province, China*

[b]*Donghai Navigation Safety Administration (DNSA) of the Ministry of Transport (MOT), 190 Siping Avenue, Shanghai, China*

Abstract. Man-made ship collisions have greatly affected the marine environment. Automatic identification system (AIS) is widely used. Currently, researchers use AIS-based data analysis methods to predict navigation risks, but they have not been able to solve the problem of real-time ship detection. This paper studies the real-time detection method of a ship's single-shot multi-box detector (SSD) framework with typical application scenarios. SSD is a single-order deep convolutional neural network (CNN) learning algorithm that uses a feedforward CNN to generate a set of fixed-size bounding boxes for each object from a different feature map. We evaluated a number of different feature extractors, including Faster RCNN (VGG16), YOLO, YOLOv2, YOLOv3, SSD300, SSD512, RefineDet320, RefineDet512. In 2019, we collected a detection dataset of a ship sailing video atlas. To verify the method, we identified it using various ship detection methods and compared the SSD with YOLO and RefineDet. Our results show that the method has good test results and surpasses all ship detection methods. Specifically, in terms of detection speed, our proposed method is superior to all methods and can meet the actual needs of ships when detecting ships in the surrounding waters in real time. In short, the SSD-based real-time ship detection method performs well and has the potential to improve accuracy and efficiency.

Keywords: automatic identification system; single shot multi-box detector; yolo

[1] Corresponding author. Zhao Liu, Department of Hubei Key Laboratory of Inland Shipping Technology, School of Navigation, Wuhan University of Technology, Wuhan 430063, P.R. China. E-mail: zhaoliu@whut.edu.cn.

1. Introduction

In 2018, the maritime accident of the "Sanji" in the East China Sea caused considerable marine pollution and economic losses. After a joint investigation by the International Maritime Organization (IMO) and a number of stakeholder countries, the results showed that the main factor was human error, as shown in Figure 1. Based on this, this paper proposes an SSD navigation vision real-time ship detection method based on a deep convolutional neural network.

Fig. 1. "Sanji" accident scene

Artificial intelligence technology has developed rapidly, and experts and scholars from various countries worldwide have conducted much research on deep learning methods. Real-time ship detection methods based on deep learning are mainly divided into detection methods based on ship characteristic data, such as automatic identification system (AIS) and detection methods based on machine learning, such as convolutional neural network (CNN).

The traditional ship collision avoidance method based on AIS big data has the characteristics of positioning, navigation, timing error, and being deceived, which are greatly affected by human factors. Therefore, the real-time ship detection method based on SSD has practical application requirements.

The detection method based on ship characteristic data mainly detects the ship position based on edge data [1-2], texture [3], corner points [5-6], colour [7], and morphological feature data in the image. Reference [8] improved the mixed Gaussian background model, combined with [9] morphological and other feature hypothesis verification, and good results were achieved [10]. The Sobel operator was used to detect all vertical edge image data in the image, and the vertical integration projection was used. Features determine the area of interest. Literature [11] analysed the ship's edge data, used Gabor filters to extract corresponding features, and achieved good results using SVM classifiers. The literature [12] combined shallow-level features that are prone to error classification but are high-resolution and high-level features that are low-resolution but abstract, resulting in better candidate regions and further improving the accuracy of detection. Literature [13] used the Faster RCNN network to extract deep convolution features of ships, which not only addressed the problem of traditional dependence on manual features but also greatly improved the efficiency of target detection. However, these two deep learning methods do not meet real-time requirements.

In 2016, Wei Liu proposed the SSD detection method [14]. SSD is a new end-to-end detection method. SSD is also a type of CNN, but SSD not only retains the previous CNN algorithm but can ensure accuracy and meet real-time requirements.

In this paper, we propose a real-time ship detection method based on SSD. In addition, we carried out multiple tests in 2019 to test SSD-based real-time ship detection methods in Zhoushan City, Zhejiang

Province, Shanghai City, and Fuzhou City, Fuzhou Province. We validate our method using different data sets, the Kaggle Airbus Ship Detection Challenge, and Kaggle digital.

Compared to traditional methods and CNN-based ship detection methods, our system achieves an excellent balance between efficiency and accuracy. With extremely high detection speeds, our system obtains the highest number of ship identifications. Specifically, our proposed SSD algorithm is faster than Faster RCNN [15,16]. Similarly, our system is much more accurate than the YOLO of the same stage method.

Finally, we conducted a large number of experiments using SSD-based ship identification methods. The results show that we verified that with other image datasets and FasterRCNN, YOLO, YOLOv2, YOLOv3, SSD300, SSD512, RefineDet320, RefineDet512, compared with the research of SSD (DP-SSD300 and DP-SSD512), our detection model shows higher detection performance.

2. Methodology

Deep learning methods have developed rapidly, and visual deep learning methods and models in the computer field have also emerged rapidly. For example, Faster RCNN, YOLO [17], and SSD [18] have made breakthrough progress in natural target recognition and detection, such as China's Megavision Technology Co., Ltd. Unlike Faster RCNN [19], YOLO and SSD do not require regional recommendations because they can completely eliminate and suggest the generation and subsequent function of multiple sampling stages and encapsulate all the calculated data in the network. Therefore, SSD and YOLO as one-stage methods are different from Faster RCNN [20] as a two-stage method. The tests show that the SSD detection method can obtain higher accuracy and increase the detection speed and has a good use effect. However, in all detection models, the SSD method can achieve higher accuracy and excellent data. Therefore, it has application potential in real-time detection.

2.1. SSD algorithm

The CNN network algorithm is used to detect ships in the following steps:

Step 1: Generate a title box in the image;

Step 2: Feature extraction for the title box;

Step 3: Perform feature verification on the trained classifier.

Therefore, it is possible to ensure that the detected SSD algorithm emerges as necessary, and that it has both real-time capability and accuracy. Such a target detection algorithm is a considerable breakthrough. To improve the accuracy of SSD-based real-time ship detection algorithms, we adjusted the basic network structure and adopted a multilayer network structure, which is a major breakthrough in the history of target detection. To further improve the accuracy, we adopt a multilayer structure for the feature extraction method of the basic network.

2.2. SSD framework

An SSD network based on a feedforward CNN generates a fixed-size bounding box based on size or expansion. Additionally, a scoring system is used to obtain the score of the placed pair category. Finally, a nonmaximum suppression algorithm is used to eliminate duplicate frames to obtain the final detection frame. In the experiment, we use the VGG-16 network and the basic CNN network and use this feature extraction network to perform feature extraction on the input image. Different from the ordinary CNN network, we did not use a fully connected layer but designed a new type of auxiliary structure, which resulted in the following two special detectors.

Multiscale feature map detector: Several convolutional feature layers are added at the end of the basic network. The sizes of these layers gradually change to form a pyramid structure, which enhances the network's robustness to vehicle size. Moreover, the convolution models detected on each feature layer are different.

Detected convolution predictor: We designed a convolution filter for the newly added feature layer to obtain a fixed set of predictions. For a feature layer with c channels and a size of m × n, we use a convolution kernel of size 3×3×c to obtain the score of the category or the coordinate transformation relative to the alternating box. In the feature map, an output result is obtained at the corresponding position through the convolution kernel function of size m×n, as shown in Figure 2.

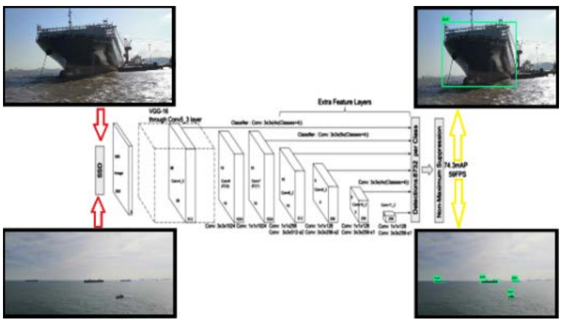


Fig. 2. Proposed SSD polyp detection framework architecture. A feature extractor is used to generate features from different spatial resolution layers.

2.3. SSD network training

When training the SSD network model, we assign the ship tag information to a specific output of the detector output set. Once the specified correspondence is determined, we can apply the loss function and backpropagation in an end-to-end manner.

The training phase is divided into the following three steps:

Step 1: Establish the correspondence between the ship's actual label and the default box in the training set. For each label box, we choose from the determined default box. At the beginning of training, we match each real label box with the best Jaccard of the default box to overlap, which can greatly simplify the need for learning computing power. Therefore, we obtain higher confidence from multiple overlapping default boxes.

We make $x_{ij}^{p} = \{0,1\}$, where $x_{ij}^{p} = 1$ indicates that the i default box matches the j real label box of category P. According to this matching strategy, we obtain $\sum_i x_{ij}^{p} \geq 1$, that is, at least one can match the j real label box. The overall target loss function is the weighted sum of confidence loss and position loss.

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

where N is the total number of matching default boxes. When $N = 0$, we consider the loss to be 0.

$$L_{conf}(x,c) = -\sum_{i\in Pos}^{N} x_{ij}^{p} \log\left(\widehat{c_i^p}\right) - \sum_{i\in Neg} \log\left(\widehat{c_i^0}\right) \quad (2)$$

where

$$\widehat{c_i^p} = \frac{exp(c_i^p)}{\sum_p exp(c_i^p)} \quad (3)$$

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

$$L_{loc}(x,l,g) = \sum_{i\in Pos}^{N} \sum_{m\in\{cx,xy,w,h\}} x_{ij}^{k}\, smooth_{L_1}\left(l_i^m - \widehat{g_j^m}\right) \quad (4)$$

where,

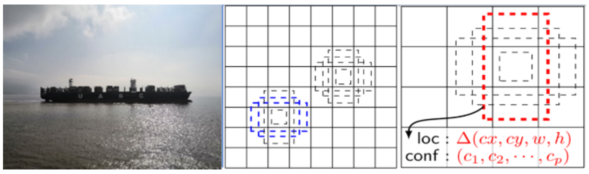
$$\widehat{g_j^{cx}} = \left(g_j^{cx} - d_i^{cx}\right)/d_i^{w} \quad (5)$$

$$\widehat{g_j^{cy}} = \left(g_j^{cy} - d_i^{cy}\right)/d_i^{h} \quad (6)$$

$$\widehat{g_j^{w}} = \log\left(g_j^{w}/d_i^{w}\right) \quad (7)$$

$$\widehat{g_j^{h}} = \log\left(g_j^{h}/d_i^{h}\right) \quad (8)$$

Step 2: the number of network layers increases, and the pooling process reduces the obtained feature maps. Therefore, the need for computing power can be greatly reduced. Saving memory can ensure the translation of the features and the same scale. To make the algorithm robust to scales, Sermanet et al. used images of different sizes as training sets and obtained better results. We can combine the prediction results of feature maps extracted at different levels in the base layer to achieve similar results. The experimental results of Hariharan et al. can guide the feature maps of lower layers and improve the semantic segmentation results. Similarly, global text sampled in a high-level feature map can improve the result of smooth segmentation. Within the SSD framework, the default frame settings do not need to correspond to the actual field of view of each layer. We can obtain the prediction set containing the different sizes and shapes of the input object through the combination of different sizes of feature maps and the prediction results of the default boxes with different aspect ratios. For example, in Figure 3, the default box of the ship 4×4 feature map matches but does not match the default box of the 8×8 feature map. This is because the default boxes all have different sizes but do not match the ship's box, so they are considered negative samples during training.



a. Real map b. 8×8 Feature map c.4×4 Feature map

Fig. 3. Schematic diagram of SSD

Step 3: Most of the default boxes centred on real labels are negative samples, which causes a serious imbalance in the proportion of positive and negative samples during training, thereby reducing the accuracy of detection. Therefore, we sort the highest confidence of each default box. By selecting the box that is ranked first, we guarantee that the ratio of positive and negative samples is at most 3:1, instead of using all negative samples.

During the training process, to make the model more robust to the size and shape of the input image, the data are augmented by random sampling of the training image. If the box centre of the real label is within the sampled segment, the overlapping part is retained. At the same time, to maintain the uniformity of the network, we set the size of each sample slice to a fixed value and horizontally flip it with a probability of 0.5.

3. Experiment

3.1. Data selection

We evaluated the SSD model using the Kaggle Airbus Ship Detection Challenge and the Kaggle digital datasets. The datasets contain 2,000 images obtained from 50 videos of 10 voyages from Shanghai, Zhoushan City, Zhejiang Province, and Fuzhou City, Fujian Province. We also evaluated SSD in the dataset. Details of these datasets are listed below.

The Kaggle Airbus Ship Detection Challenge contains 1,000 images with ship detection results obtained from 25 video sequences obtained from 5 voyages.

The Kaggle digital dataset contains 1,000 images with relevant ship pictures, backgrounds (including multiple ships, reefs, and lighthouses), and is segmented from 25 video sequences obtained within 5 voyages. We have a database of frames extracted from the ship video detection video. These frameworks contain multiple examples of ships.

All annotations provided in the dataset are per-pixel masks, while SSD requires ground-truth bounding boxes. Therefore, to feed data using SSD, we convert the mask to a rectangular frame and ensure that all ships are inside the frame, as shown in Figure 4.

Fig. 4. Bounding box with a green rectangle is the notation used during SSD training.

3.2. Evaluation indicators

We use a new evaluation method. Specifically, when testing ships in the area, each ship considers only one true negative, which can be considered (TN). Any test value other than the ship is regarded as a false positive and can be considered (FP). The situation in which the ship's image is not identified is considered a false negative and can be considered (FN). If the ship is not found in the picture, and no picture is provided for the picture output, we define this value as true negative (TN). Among them, true positives, we define as (TP); false positives, we define as (FP); false negatives, we define as (FN); true negatives, we define as (TN). After testing, we give different details. The performance indicators (precision, recall, specificity, F1 measurement, F2 measurement) are listed.

Table.1.

Performance metrics for polyp detection.

Metric	Abbreviation	Calculation
Precision	Prec	$Prec = \frac{TP}{TP+FP}$
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	$F2 = \frac{5 \times Prec \times Rec}{4 \times Prec + Rec}$

3.3. Implementation

We use the programming language PyTorch to carry out the experiments. The graphics card used is NVIDIA GeForce RTX 2080Ti. The processor is an Intel (R) Core (TM) i9 9900KF CPU @ 3.6 GHz, with 32.0 GB of RAM. Due to the lack of sufficient data, we use the dataset for training to extract feature models. In the training process, we use the mean method to train the model. To improve robustness and reduce overfitting, the following data theory should be adopted:

Random cropping, scene changed to colour;

Adjust all pictures to 300×300, 320×320, 416×416, 448×448, 512×512, 544×544;

Random sampling

Finally, we conduct multiple groups of experiments for comparison.

4. Results and discussion

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 light, the test results will also affect the speed and real-time detection results of the ship. This article tests the SSD framework with high accuracy, low computational complexity, and fast speed. In order to explore the power of the SSD method, we also performed multiple sets of experiments, including Faster RCNN (VGG16), YOLO, YOLOv2, YOLOv3, SSD300, SSD512, RefineDet320, RefineDet512, SSD (DP-SSD300 and DP-SSD512). Finally, we tested the SSD method using image and video data from Shanghai, Zhoushan City, Zhejiang Province, and Fuzhou City, Fujian Province. Compare and evaluate real-time ship monitoring data of Faster RCNN (VGG16), YOLO, YOLOv2, YOLOv3, SSD300, SSD512, RefineDet320, RefineDet512, SSD (DP-SSD300 and DP-SSD512).

Table.2.

Multi-group test results for real-time monitoring of ships

Method	Input	mAP(%)	FPS
Faster RCNN(VGG16)	-	72.69	11.27
YOLO	448×448	62.50	42.33
YOLO v2	416×416	73.84	64.66
YOLOv2 544×544	544×544	75.97	39.15
YOLOv3	416×416	88.11	51.27
SSD 300	300×300	74.21	58.79
SSD 512	512× 512	76.85	27.77
RefineDet320	320×320	76.81	46.81
RefineDet512	512×512	77.71	29.46
Ours (DP-SSD300)	300×300	76.45	54.49
Ours (DP-SSD512)	512× 512	77.97	25.13

4.1. Precision comparison

As shown in Figure 2, we use SSD and other monitoring methods for comparison, using different feature extractors respectively. The experiments show that in the tests of Faster RCNN (VGG16), YOLO, YOLOv2, YOLOv3, SSD300, SSD512, RefineDet320, RefineDet512, SSD (DP-SSD300 and DP-SSD512), we found that SSD has a lower false alarm rate. Therefore, the ship recognition method using SSD has the strongest recognition accuracy of ship picture information.

We use three methods of end-to-end, hybrid method and manual plotting to test the target picture. We found that using the SSD method can generate feature extractors from layers with different spatial resolutions, so that the largest proportion of ship features can be captured, while having the least false positive results. Through experiments with multiple sets of CNN feature extraction methods, when the picture quality and hardware are in the same operating environment, the real-time ship detection method using the SSD method has the highest accuracy.

4.2. Detection efficiency

We have tested Faster RCNN (VGG16), YOLO, YOLOv2, YOLOv3, SSD300, SSD512, RefineDet320, RefineDet512, SSD (DP-SSD300 and DP-SSD512). The test results show that the real-time monitoring method for ships in SSD has the characteristics of simultaneous classification and regression, and has the image recognition type. It has the strongest detection efficiency in the same environment as the hardware operating environment. When the application scene is at different angles, light intensity and environment rendering, the main factor for SSD detection is the elimination of the region replacement step. The two-stage detection method (for example, Faster RCNN) between this step. Therefore, this method is better than other methods. More effective and accurate.

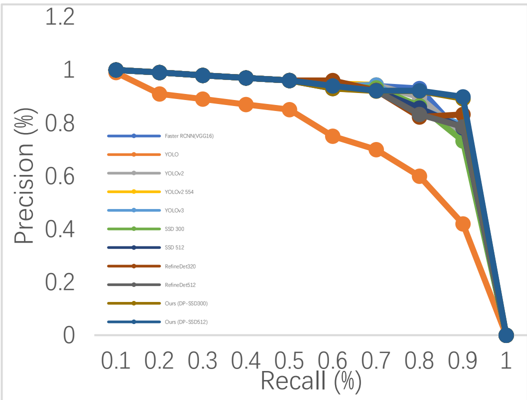


Fig. 5. Precision-recall curves for all the methods

The performance of SSD with three backbones is much better than that of the teams that attended the other methods but slightly lower than that of the two-stage method.

4.3. Comparison of experimental results

We compare the real-time ship detection method of SSD with the literature [12]. As shown in Table 2, the accuracy of detection speed F1 and F2 has faster real-time ship detection results. We conducted tests on the best data set of 198 ships, and the test results can detect 180 ships. As shown in Table 3, during the training process, the post-processing period and other data sets are tested, and the training is repeated many times. The results show that the real-time ship monitoring method using SSD has the strongest real-time ship detection capability and the shortest detection time when using the same test data set and hardware environment.

a. Compared with YOLO

We compare YOLO and SSD in two groups. The test results show that when using YOLO for real-time ship inspection, YOLO is lower than SSD in both the accuracy of the test results and the detection speed. This is mainly because when YOLO is used for image recognition, the feature map generates candidate frames, which is affected by the spatial resolution of the picture, resulting in lower accuracy and detection time. On the other hand, SSD can use multiple spatial resolutions and different feature layers to generate real-time ship detection results. Therefore, the test results using the SSD method have stronger detection accuracy and higher detection efficiency.

b. Comparison with RefineDet

We compared the RefineDet and SSD trials. The test results show that the RefineDet method of the multi-step target detection network is used to perform real-time video detection on ship targets.

Step 1, generates a series of multiple candidate frames;

Step 2, performs classification regression on the target pictures;

Step 3, confirms the ship recognition result.

SSD method, which uses multiple candidate frames generated in each unit to perform classification and regression at the same time. Therefore, the test results using the SSD method have higher detection accuracy and detection efficiency.

4.4. Detection and Identification

We use different methods to evaluate the impact of the test feature extractor. The test results are shown in Table 3. The values in the table show the confidence level of the ship's detection area. Specifically, as shown in Table 3, training evaluation indexes of feature extractors using different methods.

Table.3.

Detection identification evaluation

Method	Map (%) Moderate	Easy	Hard	Runtime
Faster RCNN (VGG16)				
YOLO	91.98	91.99	84.55	0.37s
YOLO v2	85.35	89.21	74.81	0.1s
YOLOv2 544×544	84.74	88.97	72.66	0.1s
YOLOv3	83.86	87.38	70.72	0.08s
SSD300	82.39	88.15	71.71	0.07s
SSD512	81.55	76.22	66.84	0.7 s
RefineDet320	81.37	90.37	70.31	0.7 s
RefineDet512	79.78	89.79	78.59	0.34s
Ours (DP-SSD300)	79.23	90.13	65.69	0.3 s
Ours (DP-SSD512)	79.23	90.13	65.69	0.3 s

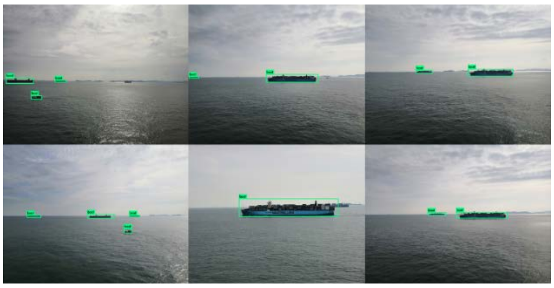
Different ship appearances, different photo angles, and different light intensities make it difficult to detect the ship. To solve this problem, SSD uses multi-layered functional maps with different spatial resolutions. At the same time, it can use filters or accept different combinations of different fields for image feature extraction.

The SSD structure enables the network to obtain ship features covering different sizes and shapes through different convolution sizes, so it can reduce and large changes in shape. For SSD, the fast connection network used to reconstruct the convolutional layer still retains more ship characteristics when deep. This structure helps to improve the detection area of ships based on the SSD detection method.

It is shown that different filter sizes and shortcut connections are crucial for the extraction of ship features. I compare the SSD with YOLO and RefineDet. We use the SSD method to detect ships in real time. In summary, low-level features have typical geometric data, and high-level features have more semantic data, which can effectively improve the detection accuracy.

We use the SSD method for real-time video detection of marine vessels, and the successful detection result is green. We extract single target detection results, multiple target detection results, and different angle target detection results, as shown in Figure 6.

Single target ship detection



Multi-target ship detection

Multi-angle target ship detection, the first is the original picture

Fig.6. Ship target detection results

4.5. Limitations

Our research has limitations. The experimental data and real-time scene samples are limited, and larger experimental samples are needed in the future. Among them, the lack of more tests on factors such as light intensity, photo angle, complex environment, and ship sway. The detector is specifically trained to distinguish ship shapes in different scenarios. However, as shown in Figure 7, they still have difficulty identifying ships (small ships in distant scenes). Experiments show that the method of obtaining more training data can improve the accuracy of SSD-based ship identification methods. The test dataset contains more ship data, so it cannot be directly compared with the actual scene of manipulating the ship. Considering that maritime accidents triggered by ship navigation will cause inestimable losses, we need to conduct comprehensive training on the ship dataset. Further prospective research is needed to verify the performance of its algorithm.

Fig.7. Sample failure detection. There are failure results in the first ship detection target, marked in red.

5. Conclusions

In this article, we have studied a method based on the SSD (DP-SSD300 and DP-SSD512) framework. We use different feature extractors to perform multi-dataset and multi-learning on ship pictures. Subsequently, we embedded video and pictures into the real-time ship detection system of SSD, and compared with the methods of YOLO and RefineDet, we verified that the ship detection method using SSD architecture has high accuracy and high efficiency. Our research shows the feasibility of SSD-based methods to provide ship pilots with additional information. The development of a real-time ship automatic identification platform will have a positive impact on future ship navigation based on artificial intelligence technology. At the same time, we will further develop hardware that improves detection accuracy and efficiency, develops cheaper and more stable. At the same time, we will make recommendations to the International Maritime Organization (IMO) and the International Association of Marine Aids to Navigation and Lighthouse Authorities (IALA) for the establishment of demonstration projects.

Acknowledgments

This work was partially supported by National Natural Science Foundation of China under Grant No. 51809207 and the Fundamental Research Funds for the Central Universities No. 2019IVB040 and the National key research and development plan under Grant No. 2018YFC1407404.

References

[1] C. Hassan, E. Quintero, J.-M. Dumonceau, J. Regula, C. Brandão, S. Chaussade, E. Dekker, M. Dinis-Ribeiro, M. Ferlitsch, A. Gimeno-García, Y. Hazewinkel, R. Jover, M. Kalager, M. Loberg, C. Pox, B. Rembacken, and D. Lieberman, Post-polypectomy colonoscopy surveillance: European society of gastrointestinal endoscopy (ESGE) guideline, Endoscopy 45(10) (2013), 842–851.

[2] M. Løberg, M. Kalager, O. Holme, G. Hoff, H.-O. Adami, and M. Bretthauer, Long-term colorectal-cancer mortality after adenoma removal, England J. Med 371(9) (2014), 799–807.

[3] Y. Fang, Z. Wang, W. Lin, and Z. Fang, Video saliency incorporating spatiotemporal cues and uncertainty weighting, IEEE Transactions on Image Processing, vol. 23, no. 9, pp. 3910–3921, 2014.

[4] Y. Fang, Z. Wang, W. Lin, and Z. Fang, Video saliency incorporating spatiotemporal cues and uncertainty weighting, IEEE Transactions on Image Processing 23(9) (2014) 3910–3921.

[5] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein et al., Imagenet large scale visual recognition challenge, International Journal of Computer Vision 115(3) (2015) 211–252.

[6] P. Mital, T. J. Smith, S. Luke, and J. Henderson, Do low-level visual features have a causal influence on gaze during dynamic scene viewing? Journal of Vision 13(9) (2013) 144–144.

[7] W. Zhu, S. Liang, Y. Wei, and J. Sun, Saliency optimization from robust background detection, in IEEE Conference on Computer Vision and Pattern Recognition (2014) 2814–2821.

[8] I. Sluimer, A. Schilham, M. Prokop, and B. van Ginneken, Computer analysis of computed tomography scans of the lung: A survey, IEEE Trans. Med. Imag., 25(4) (2006) 385–405.

[9] S. Park, S. M. Lee, N. Kim, J. B. Seo, and H. Shin, Automatic reconstruction of the arterial and venous trees on volumetric chest CT, Med. Phys., 40(7) (2013) 071906.

[10] E. A. Regan et al., Genetic epidemiology of COPD (COPDGene) study design, J. Chronic Obstructive Pulmonary Disease 7(1) (2010) 32–43.

[11] W. Huang, R. T. Yen, M. McLaurine, and G. Bledsoe, Morphometry of the human pulmonary vasculature, J. Appl. Physiol., 81(5) (1996) 2123–2133.

[12] A. M. Leufkens, M. G. H. Van Oijen, F. P. Vleggaar, and P. D. Siersema, Factors influencing the miss rate of polyps in a back-to-back colonoscopy study, Endoscopy 44(5) (2012), 470–475.

[13] J. C. Van Rijn, J. B. Reitsma, J. Stoker, P. M. Bossuyt, S. J. Van Deventer, and E. Dekker, Polyp miss rate determined by tandem colonoscopy: A systematic review, Amer. J. Gastroenterol 101(2) (2006),343.

[14] C. Van Wijk, V. F. Van Ravesteijn, F. M. Vos, and L. J. Van Vliet, Detection and segmentation of colonic polyps on implicit isosurfaces by second principal curvature flow, IEEE Trans. Med. Imag 29(3) (2010),688–698.

[15] B. V. Dhandra, R. Hegadi, M. Hangarge, and V. S. Malemath, Analysis of abnormality in endoscopic images using combined hsi color space and watershed segmentation, in Proc. 18th Int. Conf. Pattern Recognit. (ICPR) 4, Aug. (2006), 695–698.

[16] Y. Wang, W. Tavanapong, J. Wong, J. Oh, and P. C. de Groen, Partbased multiderivative edge cross-sectional profiles for polyp detection in colonoscopy, IEEE J. Biomed. Health Informat, 18(4) (2014) 1379–1389.

[17] Y. Lecun, B. Boser, J.S. Denker, D. Henderson, and R.E. Howard, W. Hubbard, Backpropagation Applied to Handwritten Zip Code Recognition, Neural Comput 1(1989) 541–551.

[18] P. Dollar, C. Wojek, B. Schiele, and P. Perona, Pedestrian detection: An evaluation of the state of the art, IEEE Transactions on Pattern Analysis and Machine Intelligence 34(4) (2012) 743–761.

[19] S. J. Raudys and A. K. Jain, Small Sample Size Effects in Statistical Pattern Recognition: Recommendations for Practitioners, IEEE Transactions on Pattern Analysis and Machine Intelligence 13(3) (1991) 252–264.

[20] C. Cortes and V. Vapnik, Support-Vector Networks, Machine Learning 20(3) (1995) 273–297.

XIANG WANG received the M.Sc degree in Transportation Engineering Department of Jimei University's School of Navigation, Xiamen, China, in 2015. He is currently a senior engineer of the Donghai Navigation Safety Administration (DNSA) of the Ministry of Transport (MOT) of the People's Republic of China. His research interests include intelligent transportation systems, vessel traffic flow and artificial intelligence, and ocean engineering.

JINGXIAN LIU received the PhD degree in Vehicle Operation Engineering from the School of Energy and Power Engineering, Wuhan University of Technology, Wuhan, China, in 2008. He is currently a Full Professor at the School of Navigation, Wuhan University of Technology, Wuhan, China. He was a Visiting Scholar at the Illinois Institute of Technology, Illinois, USA, and Massachusetts Maritime Academy, Massachusetts, USA. He has authored or co-authored over 100 international journal and conference papers. His research interests include intelligent transportation system, vessel traffic flow, and artificial intelligence.

ZHAO LIU received the Ph.D. degree in Traffic Information Engineering and Control from the School of Navigation, Wuhan

University of Technology, Wuhan, China, in 2017. He is currently a lecture at the School of Navigation, Wuhan University of Technology, Wuhan, China. His research interests include intelligent transportation system, vessel traffic flow, and artificial intelligence.

ZHI YUAN is currently working toward the PhD degree at the School of Navigation, Wuhan University of Technology (WUT), China. His research interests include intelligent transportation system, vessel traffic flow, and artificial intelligence.