Research on Machine Vision Detection Method of Ship Sulfur Emission Based on Convolutional Neural Network

Yang Han¹, Guangzheng Li¹*, Qi Qin¹, Shipeng Wang¹, Yuanhao Li¹

¹Shipping college, Shandong Jiaotong University, Weihai, Shandong Province, China
*email: ligz@sdjtu.edu.cn

Abstract. The traditional method of ship exhaust gas detection has the shortcomings of poor timeliness, blindness and hysteresis, which cannot meet the needs of marine supervision departments for comprehensive, efficient and real-time supervision of ship sulfur emissions. To solve this problem, a convolution neural network (CNN) based machine vision detection method for ship sulfur emission is introduced, and the plume characteristics of heavy sulfur black smoke ship exhaust are adaptively extracted by convolution operation. An improved LeNet-5 network model structure is proposed. By changing the activation function and increasing the depth of the model, the batch normalization operation is introduced to improve the accuracy and generalization ability of the model. The experimental results show that the detection accuracy of the model for heavy sulfur black smoke ships can reach 94.52%, which is better than the ordinary CNN network model and the VGG-16 network model. This method can quickly screen suspicious ships with excessive sulfur content, broaden the way for maritime supervision departments to detect ships under the background of "sulfur limit order", and effectively improve the supervision efficiency of maritime personnel.

1. Introduction

The International Maritime Organization (IMO) has made it clear that a global sulfur emission limit of 0.5% will be implemented on January 1, 2020, requiring all ships sailing outside the sulfur emission control zone to use fuel with a sulfur content of no more than 0.5%.[1] At present, the maritime departments of various countries in the ship sulfur emission control zone have gradually strengthened their supervision.

Currently, detection methods of sulfur emission from ships can be divided into two types: contact and non-contact, the snifing technology is contact detection, differential Optical Absorption Spectroscop (DOAS) and infrared remote sensing technology belong to non-contact detection. Many scholars have done related research on ship emission gas detection. In the existing research of short-term instrument monitoring, the smoke plume sniffing method based on the carbon balance method proposed by the IMO is taken as the representative. Mellqvist[2] uses this method to reverse calculate the sulfur content of the ship's fuel oil in a certain water area of Denmark, and the calculation accuracy is up to 0.1%. In China, Zhong R[3] proposed a rapid detection method of sulfur content in ship fuel based on gas sensor, which uses the tail gas detection pod carried by UAV to detect the gas concentration of SO₂ and CO₂ in the tail gas of a single ship. Zhuo Hongming et al. [4] designed a rapid and intelligent detection of ship exhaust by using remote sensing detection technology of ship exhaust and gas detection equipment mounted on unmanned air vehicle (UAV). This kind of method has good timeliness and short detection cycle, but it is greatly affected by the weather and environment, the scope of supervision area is limited, and testing personnel need to be trained in UAV...
operation. In the remote non-contact monitoring and control, Wu J[5] proposed a ship exhaust remote monitoring method based on WebGIS, Internet of things, MyBatis framework and other technologies, which can realize real-time visual monitoring of ship emissions. Liu Y et al.[6] using the combination of the real-time emission calculation model of ship air pollutants and the DOAS shore-based on-line observation method, the accuracy of sulfur content in smoke plume is obtained by inverse calculation of ship plume identification and ship fuel sulfur content.

In recent years, with the theory of deep learning put forward by Hinton et al.[7] in 2006, this has opened a wave of deep learning research in the academic world. As a representative neural network in deep learning technology, CNN is widely used in the field of image classification, and CNN has achieved good results in the recognition of smoke. CNN can use multi-layer neural network to extract features such as the chromaticity, texture, and shape of the ship's sulfur exhaust plume.[8-10]. This paper uses machine vision to put forward a method to directly detect whether the exhaust plume of ships is emitted illegally. Machine vision has a wide spectrum corresponding range, which can broaden the scope of supervision. Compared with contact detection such as UAV, this method can be used for long-term supervision and identification tasks. In addition, machine vision belongs to non-contact detection, and the system has good stability. Compared with traditional non-contact detection of machines, there is no need to build a large gas detection station, which greatly reduces the detection time and labor cost. Based on this idea, an improved LeNet-5 structure is proposed, and the improved network model structure is named Model_A, which realizes the extraction of the exhaust plume characteristics of heavy sulfur black smoke ships. The detected ship exhaust image categories are divided into polluted ships and non-polluted ships. Experimental results show. The experimental results show that Model_A achieves 94.52% accuracy on the test set, which exceeds the performance of ordinary CNN and VGG16 network models.

2. Methods

2.1. The network structure

The classic CNN model LeNet-5 structure was first proposed by LeCun[11] in 1988, It was first used to recognize handwritten digits (MNIST) image sets. The LeNet-5 structure has a total of 7 layers including 3 convolutional layers, 2 pooling layers, 1 fully connected layer and 1 output layer. The structure is the pioneering work of the CNN model, which determines the basic architecture of the CNN model structure.

Our network has 11 layers, as shown in Figure 1. The network includes 5 convolutional layers, 5 pooling layers, 2 fully connected layers and 1 output layer. Our changes to LeNet-5 are as follows: (1) Increase network depth, increase the recognition performance of the model; (2) Change the activation function to increase the non-linear expression ability of the network, the ReLu function effectively reduces the problem of gradient disappearance, has high computational efficiency and speeds up the convergence speed of the model; (3) Add a Dropout layer, add a Dropout layer after the first fully connected layer, the Dropout parameter is set to 0.5, the model randomly turns off 50% of the neurons during training, which can effectively prevent overfitting for small sample data sets. (4) Introduce batch normalization operation. Add batch normalization after convolution operation, improve the learning rate of training, reduce the dependence of the gradient on the parameter or its initial value scale, accelerate the convergence of the network, and improve the accuracy of the model on the test set. The batch normalization operation calculation process is

$$y_i = \text{BN}_{\gamma, \beta}(x_i) = \gamma \frac{x_i - \frac{1}{n} \sum_{i=1}^{n} x_i}{\sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \frac{1}{n} \sum_{i=1}^{n} x_i)^2} + \varepsilon} + \beta$$

Where $y_i$ is the I output value after batch normalization, $x_i$ is the I input value, $\varepsilon$ is a small constant.
to prevent the denominator equal to 0, γ is the scale parameter, and β is the shift learning parameter.

**Fig. 1** Model_A network structure

Input image size 150 * 150, C1-C5 are convolutional layers, the size of the convolution kernel is unified to 3×3, and their strides are 1, P1-P5 is the max pooling layer, the pooling size is 2×2, and their strides are 2. The number of output feature maps for each pooling layer is 16, 32, 64, 128, 256, respectively. After these layers, FC1 and FC2 are fully connected layers. FC1 has a total of 1024 neurons. After passing through the Dropout layer, it enters the FC2 layer, and the FC2 layer outputs 1 neuron. The last is the output layer, because this article is a two-classification problem and only needs to distinguish between polluted ships and unpolluted ships.

### 3. Experiment part

The experimental data in this article comes from: ①Collecting pictures from relevant domestic and foreign shipping websites ②SeaShips data set ③On-site shooting of ship videos. A total of 1,000 pictures of ships illegally emitting heavy sulphur and black smoke and 2,400 pictures of ships with standard sulphur emissions have been sorted out. In this paper, the data set is expanded, and the number of positive samples is expanded to 2,400 after performing image enhancement methods such as horizontal mirroring, flipping, and adding gaussian noise to 1000 pictures of illegal ship emissions. The data set compiled in this paper is randomly divided into training set and test set according to the ratio of 5:1. Each sub-data set contains two categories of polluting ships and non-polluting ships. The polluting ship data set and the non-polluting ship data set are shown in Figure 2. and Figure 3.

**Fig. 2** Unpolluted_ship  **Fig. 3** Polluted_ship

#### 3.1. The experimental setup

The experimental environment was Tensorflow, a deep learning framework developed by Google, and programming with python language. Computer configuration for the experiment: Intel Core i7-8700 CPU, 32 GB running memory, GPU NVIDIA GeForce GTX 1080, memory size 1 TB.

As the experimental control group, two groups of comparative network models were set up, namely the ordinary convolution neural network model and the migration network. The convolution size and pooling size of the two network models are the same, which are 3×3 and 2×2 respectively. In this paper, the full connection layer of VGG16 network trained in Imagenet is used to retain the parameters for migration. The parameter setting of the last full connection layer of the VGG16 network model is 2,
Finally, the images are divided into two different categories: polluted ships or non-polluted ships. In order to ensure the credibility of the performance, the hyper parameters selected after multiple rounds of adjustment are shown in Table 1.

Tab. 1 The hyper parameters

| Image size | Initial learning rate | Batch size | Epoch | Optimizer |
|------------|-----------------------|------------|-------|-----------|
| 224×224    | 0.001                 | 32         | 100   | Adam      |

3.2 Training results and analysis

In the experiment, we randomly load and train the dataset three times in three different network models. To compare with other methods, we conducted a comparative experiment, and the results are shown in Table 2. the test accuracies are evolving along the training epochs, as shown in Figure 4.

Tab. 2 The contrasting accuracies between the proposed method and the others

| Network mode | Accuracy rating(%) | Batch size | Training epochs |
|--------------|--------------------|------------|-----------------|
| CNN          | 75.8               | 32         | 42              |
| VGG-16       | 90.98              | 32         | 80              |
| Model_A      | 94.52              | 32         | 73              |

The experimental results show that in the comparison of the test accuracy of the ordinary CNN and the Model_A model, the accuracy curve of the Model_A model is obviously above the CNN network model from the beginning of the training epochs, and the accuracy is significantly higher than the CNN model. When the ordinary CNN network model is trained for 42 epochs, the CNN model curve oscillates up and down 75.6%, when the number of training epochs is 73 times, the Model_A network model shows a stable trend, and its accuracy rate reaches 94.52%. In the comparison between VGG-16 and Model_A network models. When the number of the training epochs is 3, the accuracy of the test set of Model_A in this paper exceeds that of the VGG-16 network model, at this time, the Model_A model rapidly rises to 78%, and the accuracy has always been better than VGG-16 network model. When the number of the training epochs is 80, the VGG-16 network model shows a stable trend, and its accuracy curve reaches 90.98%. From the comparison results, it can be seen that the convergence speed and the test accuracy of the Model_A model are obviously superior.

4. Conclusions

In this paper, in the face of the regulatory requirements of ship sulfur emissions and the shortcomings of existing regulatory methods, We propose an improved LeNet-5 network model structure and build a suitable for heavy sulfur black smoke detection Deep Learning Network Model_A, this makes the...
detection of ships that emit heavy sulfur and black smoke in violation of regulations reached 94.52%, which is better than other algorithms. Compared with the traditional sniffing technology, which is contact detection, ultraviolet differential absorption spectroscopy technology, infrared remote sensing technology and other non-contact detection, it has the advantages of strong timeliness and can quickly locate suspected polluting ships, widening the supervision way of maritime law enforcement department. Our research still needs to be further strengthened, there are fewer image data sets collected in this article, Continuing expansion of contaminated ship data sets is also required, and if the image segmentation technology is used before the image classification task, the ship exhaust plume features in the image can be better extracted.

Acknowledgments
This study was supported financially by Shandong Big Data Development Innovation Laboratory for Shipping Safety and Management, ‘Dynamic Evolution Mechanism and Monitoring of Water Traffic Situation for Complex Navigation Systems’(ZR202102181062), and Shandong Jiaotong University Graduate Science and Technology Innovation Project (2021YK02).

References
[1] Ministry of Transport 2019 Low sulfur oil will be fully used by ships traveling in China’s jurisdiction in 2020 https://www.msa.gov.cn/html/xxgk/hsyw/20191028
[2] Berg N, Mellqvist J, Jalkanen J P, et al 2012 Ship emissions of SO2 and NOx: DOAS measurements from airborne platforms Atmospheric Measurement Techniques 5(5) pp 1085-1098
[3] Zhong R, An B, Chen W, Ni X and Pan S 2020 Detection system of sulfur content in marine fuel based on gas sensor Journal of Sensing Technology China 33 (05) pp 757-762
[4] Zhuo H, Miao Bi and Xu P 2021 Design of ship exhaust intelligent detection system based on UAV Wireless Internet Technology China 18 (06) pp 53-55
[5] Wu J, Yuan X and Liu L 2020 The telemetry of ship tail gas based on WebGIS is set up according to the analytical system The current generation of electric technology China 43 (12) pp 111-114
[6] Liu Y, Zhang Y, Yuan Y, Wang S, Guo J,Wang L and Zhou B 2021 Identification method of sulfur content in fuel oil based on real-time emission from ships and on-line monitoring of smoke plume Journal of environmental science China 41(07) pp 2624-2632
[7] Hinton E and Salakhutdinov R 2006 Reducing the Dimensionality of Data with Neural Networks Science 313 (5786) pp 504-507
[8] Chen T, Yin Y, Huang S, et al The smoke detection for early fire-alarming system based on video processing. Proc. Int. Conf. on intelligent information hiding and multimedia signal processing (Pasadena) pp 427-430
[9] Yu C, Zhang Y and Fang J 2009 Proc. Int. Conf. on computer science and engineering (Qingdao, Shandong) vol 2 pp 511-515
[10] Lee E C, Jung H W and Kim D 2010 Infrared Imaging based Finger Recognition Method International Conference on Convergence and Hybrid Information Technology pp 228-230
[11] Lecun Y, Matan O, Boser B E, Denker J S, Henderson D, Howard R E, Hubbard W E, Jacket L D and Baird H S 1990 Handwritten zip code recognition with multilayer networks Proc. Int. Conf. on Pattern Recognition (Atlantic City, New Jersey) vol 2 pp 35-40