Design of a Smart Navigation Mark Light System Based on Deep Learning

YihengWu[a], XinLi[b], ZhangjieYin[c], JianLi[d] and ChunLiu[e, *]

Institute of Computer, Hubei University of Technology, Hubei 100049, China

[a]514773500@qq.com, [b]335800283@qq.com, [c]973414732@qq.com, [d]775322174@qq.com, [e]liuchun.hust@gmail.com

*Corresponding author

Keywords: Deep learning, Smart navigation mark light, MobileNet-SSD, Waterway monitoring

Abstract: Based on deep learning, the design of smart navigation mark light system proposes to extend traditional navigation mark light as an unattended platform for waterway. While actively acquiring the working status of navigation mark light and timely feeding back the damage information to the back-end server, the system can also monitor the passing vessels on waterway at all times. In view of the main cause of damage to navigation mark light for ship impact, this system can retain evidence and warn the management so the economic loss caused by hit-and-run vessels. After real-scene experiments on the Yangtze River, the navigation mark light equipped with this system can be used as an unattended platform and implement all functions.

1. Introduction

The navigation mark light is a type of traffic light installed on some navigation aids. It emits the specified light color and flash frequency at night to reach the specified illumination angle and visibility distance, thus guides the ships at night. The malfunction of the navigation mark light to will cause great safety hazards to the waterway, and the shipwreck caused by the damage of the navigation mark light is not rare. The impact of the passing vessels is a major cause of damage to the navigation mark light. However, it is difficult to obtain evidence of the crime after the vessel has escaped. The damage of the navigation mark light has caused the waterway management department large economic losses every year.

At present, the technology to obtain the status of the navigation mark light and timely feedback of the damage information to the administrator has matured, and there are also many studies on how to extend the life of the navigation mark light. Paper [1] proposed a hybrid power supply system consisting of a super capacitor and a battery in parallel, which can improve the endurance of the navigation mark light. Song et al. [2] proposed that combining the navigation light with remote sensing technology to monitor the water level of the channel in real time while ensuring the normal operation of the navigation light, and to transmit the information in real time through GPRS. In the aspect of ship detection, Zeng [3] proposed a vessel detection method combined mean-shift algorithm with horizontal and vertical gradient information, ZHOU et al. [4] combined image morphology model with wavelet transform, proposed sequence image fusion method to achieve the detection of surface vessels, but there is still no good algorithm at present for both near-target and far-off targets.

With the rapid development of Internet technology, deep learning is widely used in the field of image [5, 6] and has achieved good development with its powerful self-learning ability [7]. Convolutional neural networks (CNN) show advantages in target detection and classification tasks, and the development of computer hardware also provides a possibility for the promotion of convolutional networks on the mobile terminal. The data processing speed of the embedded chip combined with artificial intelligence is greatly improved compared with the traditional chip, and the power consumption of the CPU and the sensor is reduced which solved the problem that the convolutional neural network is difficult to apply to unattended scenarios because of its complex structure and large computational complexity.

The system can obtain the status information while ensuring navigation mark light's normal operation, retain the video evidence of destruction behaviors against navigation mark light and monitor waterway at the same time. Once an abnormal event is found, the system immediately sends a warning to the back-end server to notify the administrator to take measures while shooting videos. The video preserved provides traceability for the investigation. Real-time monitoring of the passing vessel improved the automation of navigation monitoring, and reduced the delay in discovering the abnormality of the waterway and the labor cost. The system is realized by a standard navigation mark light equipped with an embedded control board which is also the core hardware. Compared with the traditional embedded device, this system uses an artificial intelligence chip which greatly reduces the power consumption of deep learning and the transmitted data of the sensor, therefore purposes of relying on solar energy to work and unattended equipment running on the waterway have been achieved.

2. Overall system design

This system consists of 7 modules. After the image received by the system is transmitted to vessel identification and classification module in order to detect whether the target exists, and according to the result, choose whether to go to classification distance discrimination module or not. If the warning distance is reached, video module starts recording video and stores it in SD card at the same time. If the target vessel leaves the warning distance, the system cancels the alarm and stops recording to save storage space and power. In the case of no abnormal events, the oldest video file is overwritten when there is no space in the local storage.

When multi-sensor abnormal event detection module detects an impact event, then the file management module quickly locks the local related video evidence, uploads it to the back-end server, and issues a warning to remind administrator to handle the lock file which can only be released by the administrator.

The heartbeat process continuously transmits the working status of the navigation mark light to the back-end server, including the working information of all hardware and module in the system. Therefore, the heartbeat process is also used as an important basis for judging whether the navigation mark light is damaged or not.

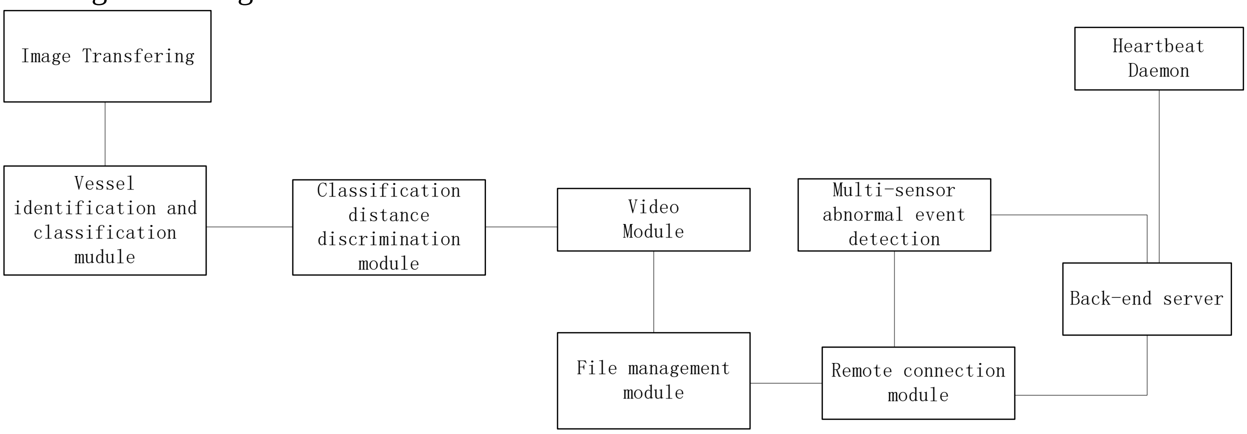

Figure 1. Overall system design

3. Vessel IDENTIFICATION and classification

Vessel identification and classification is the core module of the system. It is not only related to waterway monitoring, but also premise for obtaining video evidence. This module can also be described as target detection and classification. In the field of target detection, the Convolutional Neural Network (CNN) has excellent characteristics due to its weight-sharing network structure which reduce the complexity and weight of the network model [8, 9]. However, a large number of convolution operations make it less efficient and less real-time. In recent years, studies of neural network models' compression and advances in computer hardware have given this challenge a breakthrough.

Among the commonly used CNN models, the YOLO [10] network model detection accuracy is at a low level, and its network performance is not good. The traditional SSD [11] network model is not good for small target detection, and the Faster-RCNN [12] model's network structure leads to memory consumption and computational complexity [13]. With the needs for target recognition in mobile terminal increases, Andrew G. Howard et al proposed MobileNet, an efficient model for embedded device. The main feature of MobileNet is replacing the standard convolution of traditional network structures with depthwise separable convolution to solve the computational inefficiencies and huge parameters of convolutional networks. Andrew G. Howard et al. compared the results of the VCO model with the MobileNet model using COCO dataset to train and test based on the SSD framework [14]. The results show that the SSD framework combined with the MobileNet network structure to achieve target detection superior despite the detection accuracy is slightly reduced, but the amount of calculation and parameters are greatly decreased. For embedded platform applications, hardware resources are limited, so lightweight and low-latency network models such as MobileNet-SSD can effectively improve target detection performance.

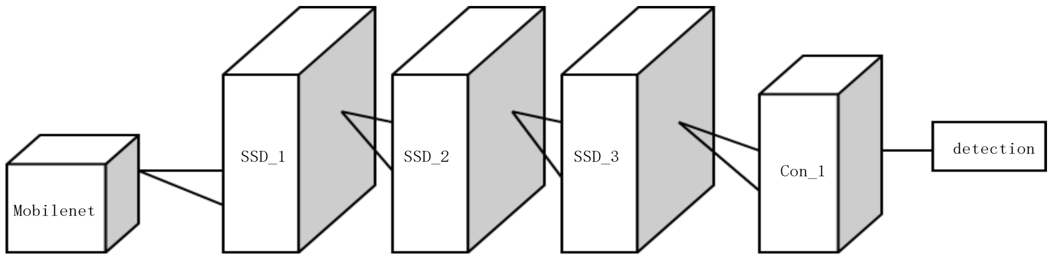

Figure 2. MobileNet-SSD structure

MobileNet-SSD adopts SSD model as the basic model, and combines MobileNet network to improve structural parameter redundancy or other issues, which reduce parameter size and computational overhead. After multi-layer convolutional network, the output is calculated separately for each channel by deep separation and convolution operations. Finally, the feature is fused by the 1×1 convolution kernel and input to the next layer network structure.

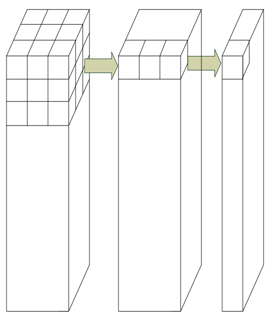

Figure 3. Depth separable convolution

The size of input feature map F is (DF, DF, M), the standard convolution K used is (DK, DK, M, N), and the size of output features are mapped to G (DG, DG, N). The number of channels input and output is M and N, the corresponding calculation amount is

$$C_1 = D_k \cdot D_k \cdot M \cdot N \cdot D_f \cdot D_f \quad (1)$$

When using depth separable convolution, the standard convolution (DK, DK, M, N) is split into deep convolution and point-by-point convolution, deep convolution is responsible for filtering whose size is (DK, DK, 1, M). The output characteristics are (DG, DG, M). The point-by-point convolution is responsible for the conversion channel whose size is (1, 1, M, N), and the final output is (DG, DG, N), the amount of calculation is

$$C_2 = D_K \cdot D_K \cdot M \cdot D_F \cdot D_F + M \cdot N \cdot D_F \cdot D_F \tag{2}$$

The amount of calculation is reduced to the original

$$\triangle C = \frac{DK \cdot DK \cdot M \cdot DF \cdot DF + M \cdot N \cdot DF \cdot DF}{DK \cdot DK \cdot M \cdot N \cdot DF \cdot DF} = \frac{1}{N} + \frac{1}{DK \cdot DK} \tag{3}$$

To make the model lighter, the MobileNet network architecture also introduces two hyper parameters, width multiplier and resolution multiplier. The first parameter width multiplier is mainly to reduce the number of channels proportionally. This parameter is denoted as α whose value range is (0, 1], then the number of input and output channels will become αM and αN for depth separable convolution

The amount of calculation becomes

$$C_3 = D_K \cdot D_K \cdot \alpha M \cdot D_F \cdot D_F + \alpha M \cdot \alpha N \cdot D_F \cdot D_F \tag{4}$$

The amount of calculation is reduced to

$$\triangle \acute{C} = \frac{DK \cdot DK \cdot \alpha M \cdot DF \cdot DF + \alpha M \cdot \alpha N \cdot DF \cdot DF}{DK \cdot DK \cdot M \cdot N \cdot DF \cdot DF} = \frac{\alpha}{N} + \frac{\alpha \cdot \alpha}{DK \cdot DK} \tag{5}$$

The second parameter resolution multiplier is mainly to scale down the size of the feature map, which is recorded as the meat resolution factor to control the resolution of the input. This parameter is denoted as ρ whose value range is (0, 1], then the calculation amount of the depth separable convolution is

$$C_4 = D_K \cdot D_K \cdot \alpha M \cdot \rho D_F \cdot \rho D_F + \alpha M \cdot \alpha N \cdot \rho D_F \cdot \rho D_F \tag{6}$$

The amount of calculation is reduced to

$$\triangle \hat{C} = \frac{DK \cdot DK \cdot \alpha M \cdot \rho DF \cdot \rho DF + \alpha M \cdot \alpha N \cdot \rho DF \cdot \rho DF}{DK \cdot DK \cdot M \cdot N \cdot DF \cdot DF} = \frac{\alpha\rho}{N} + \frac{\alpha \cdot \alpha \cdot \rho \cdot \rho}{DK \cdot DK} \tag{7}$$

It can be seen that the MobileNet-SSD model structure replaces the original redundant parameter structure with a small-scale parameter network, which greatly reduces the amount of computation, reduces the resource requirements for hardware, and shortens the training time while improving the performance of the model.

The ship image data used in the experiment was collected from the Yangtze River. After marking, a total of 2670 images were obtained for training 800 images for testing. In order to increase the generalization of the model, deliberately added adverse weather conditions in the data set. For example, a backlight or a cloudy ship picture, and in order to increase the data set and enhance the ability of the model to extract features, the model is pre-trained based on the vessel class data set of the COCO data set.

Figure 4. Dataset for training

3.1 Waterway Monitoring

The system includes four cameras for polling work, each camera pointing in one direction, so the purpose of monitoring the waterway without dead angle has been achieved. The waterway detection of this system is mainly for abnormal ships, for example a fishing boat in territorial waters where fishing is prohibited, or more extremely, found the shipwreck. When a camera starts working, the image is transmitted to the vessel classification and detection module. If it is not an abnormal ship, follow the normal steps to transmit the picture to the classification distance discrimination module, otherwise immediately issue a warning to the back-end server and send the location information.

4. Abnormal event monitoring of Navigation mark light

The anomalous events of navigation mark lights mainly refer to the impact of ships. The abnormal event can be detected by the sensors carried by the navigation mark light, and since the navigation mark light continuously sends its own working status information to the back-end server, the abnormal event can be known by the administrator in time, so the navigation mark light can be maintained at the fastest speed, thus traffic risk of the waterway is greatly reduced. In addition to monitor abnormal event, the biggest advantage of this system is that it can retain the video evidence of abnormal events. When receiving the information of the vessel detection and classification module, a fuzzy determination of the distance starts to function. If the distance is too close, the photography is turned on. When the sensor does detect an abnormal event, the video will become the retained evidence for the administrator to trace. For anomalous vessels on the channel, such as the discovery of a gradually sinking vessel or vessel wreck, the video evidence is also retained and the warning immediately sends.

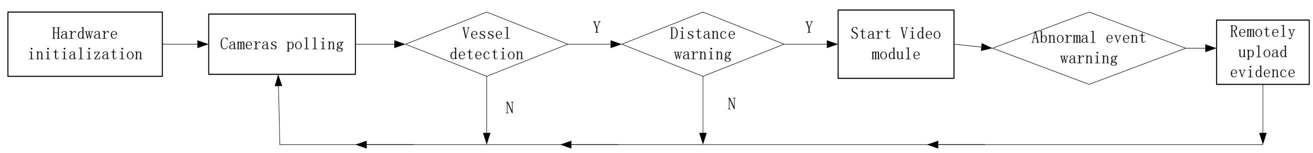

Figure 5. Process of abnormal event monitoring

4.1 Distance fuzzy calculation

Traditional ranging solutions rely on large, energy-intensive devices such as lasers, ultrasonic, and radars that are completely unsuitable to use in waterway environments. The system proposes fuzzy distance estimation based on ship classification which is a soft measurement scheme. According to the classification of the image after the target is detected, the vessel is given a rough height, and an approximate distance can be obtained according to the similar triangle principle. This method is not only greatly reduced compared to the measurement of traditional sensors, but also greatly reduced in cost.

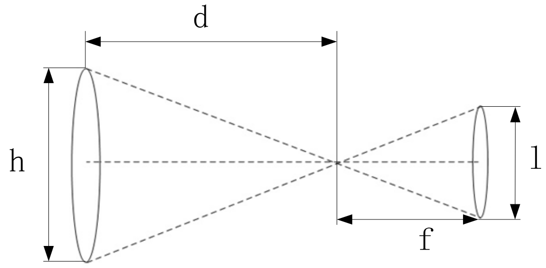

Figure 6. Principle of ranging

The distance between the object and the camera head satisfies the following relationship,

$$d = \frac{f \times h}{l} \tag{8}$$

In this formula, d is the actual distance between the object and the camera, f is the focal length of the camera. The camera used in this article is a fixed focus camera, h is the actual size of the object being detected, and l is the size of the object after it has been imaged by the camera.

4.2 File Management

File management uses a circular overlay file system. The camera records video frame by frame. When the video size reaches a certain size, it is saved as an mp4 file named after the current time and stored in the local SD card. When the system detects that the SD card has insufficient storage space, the oldest recorded video file is automatically deleted. Important video evidence retained as evidence will be set and the file lock will not be automatically deleted. For locked files, they will be deleted according to the first in first out (FIFO) principle when the capacity reaches the limit or get unlock by administrator.

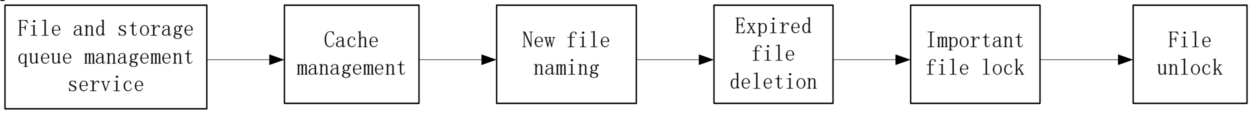

Figure 7. File management processing

5. Test and result analysis

The pre-training model used in this experiment is ssd_mobilenet_v1_coco, and the incoming data set is further trained based on the tensorflow deep learning framework. The training platform is Fedora 28 and accelerated with NVIDIA 1080Ti. The training was carried out by a random gradient descent method with an impulse of 0.85, the initial learning rate was 0.002 which decreased to its 0.1 times for each 25,000 training sessions, and the training steps were set to 100,000 times, the regularization parameter was 0.0004. Tensorboard was used to see the convergence after training.

The test platform is a common navigation mark light. The model is compressed based on Tengine. The main board of the system is arm RK3399 equipped with 4 monocular cameras with a resolution of 1600×1200, ultrasonic and acceleration sensors.

The navigation mark light was placed on the ferry of the Yangtze River in Wuhan for testing. The test time totaled three days. The test content was the recognition rate under different weather conditions and the accuracy of the classification of the vessel.

Table.1. Various types of vessel identification results

Types	**Number of vessels**	**Vessel identification accuracy**	**Vessel classification accuracy**	**Average detection time**
Small yacht	199	182	91.46%	0.0182s
Large ferry	183	169	92.35%	0.0196s
Cargo ship	203	187	92.12%	0.0123s
Warship	18	17	94.94%	0.0188s

Table.2. Vessel identification results under four lighting conditions

Weather	**Number of vessels**	**Number of identified vessels**	**Vessel identification accuracy**	**Average detection time**
Cloudy	208	188	90.38%	0.0180s
glare	216	193	89.35%	0.0177s
Evening	106	88	83.02%	0.0163s
Backlighting	216	200	92.60%	0.0186s

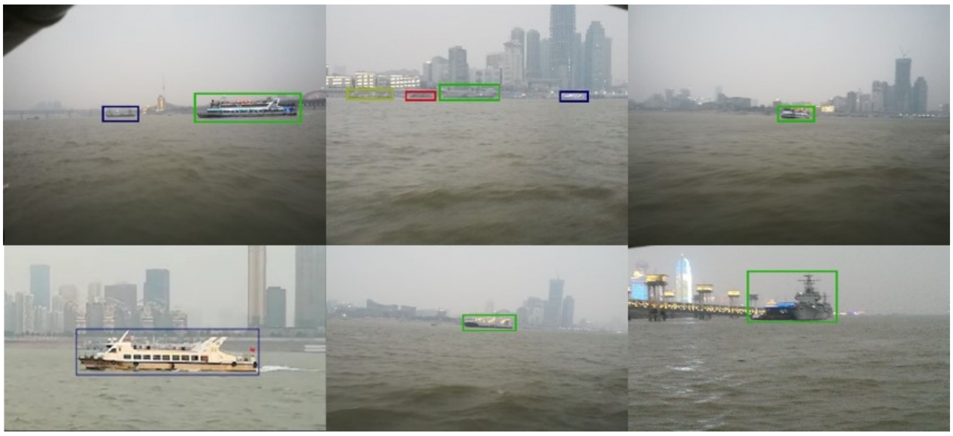

Figure 8. Real-scene recognition

6. Conclusion

Introducing deep learning into the design of the navigation mark light can not only make thousands of navigation lights on the entire waterway powerful tools for waterway monitoring, but also solve the problem of how to trace after the damage of the navigation mark light which bothers waterway manager for long. Finally the application of the waterway unmanned platform is realized. After real-scene trials, the system can fully realize all functions, shows high practical value and economic benefits. In addition, the system has strong robustness and can be widely applied to other unattended scenarios, such as forest fire warning and wild endangered animal tracking. It can be seen from the test results that the recognition accuracy of the system is mainly related to light, and will be affected a lot in a darker weather environment, so corresponding image enhancement processing will be added in the subsequent system upgrade.

Acknowledgments

This work is financially supported by National Key R&D Program of China (2017YFC1405403).

References

[1] FENG Hu-liang, “Design of LED pharos and avigation obstacle light for super capacitor”, International Electronic Elements, 2007, (11): pp. 31-33.

[2] SONG Guangjun, GU Hewei, LIU Kaiwen, CHEN Ruixing, ZHANG Wenshuai, “Design of beacon intelligence telemetry and remote control systems based on Beidou satellite communication”, Shipping Management, 2018, 40(7): pp.32-34.

[3] Zeng Wenjing, “Marine Object Detection and Tracking of USV Based on Visual System” Ph.D. thesis, Harbin Engineering University, 2013.

[4] Weina, Zhou, X. Kaiyu and L. Shuangyi. “A Method for Ship Target Detection Based on Image Fusion.” International Conference on Electronic Measurement & Instruments IEEE, 2007.

[5] Lee, Honglak, et al. “Unsupervised learning of hierarchical representations with convolutional deep belief networks.” Communications of the ACM 54.10(2011):95.

[6] Xu, Yan, et al. "Deep learning of feature representation with multiple instance learning for medical image analysis." 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) IEEE, 2014.

[7] Rayar F, “ImageNet MPEG-7 Visual Descriptors - Technical Report”, 2017.

[8] ZOU J, SONG R. “Microarray camera image segmentation with Faster-RCNN”, 2018 IEEE International Conference on Applied System Invention (ICASI), 2018: pp.86-89.

[9] Lecun, Yann, K.Kavukcuoglu and Clement Farabet, “Convolutional Networks and Applications in Vision.” Proceedings of 2010 IEEE International Symposium on Circuits and Systems IEEE, 2010.

[10] Redmon, Joseph, et al. “You Only Look Once: Unified, Real-Time Object Detection.” Proceedings of 2016 IEEE Computer Vision and Pattern Recognition. Las Vegas, 2016, pp. 779–788.

[11] Liu, Wei, et al. “SSD: Single Shot MultiBox Detector.” (2015).

[12] Ren S, He K, Girshick R, et al. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks” [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6), pp. 1137-1149.

[13] YANG Jie1, CHEN Mengzi, WU Zhi-qin1, CHEN Ling-na1*, LIN Ying, “Research on Video Target Detection Based on SSD Convolution Network”, Journal of University of South China (Science and Technology) Vol. 32 No. 1 Feb. 2018.

[14] Howard A G, Zhu M, Chen B, et al. “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications”. 2017.