A Deep Learning Model for Automatic Plastic Waste Monitoring Using Unmanned Aerial Vehicle (UAV) Data

Wenlong Han[1], Wei Luo[1], Yongtao Jin[1], and Mengxu Zhu[1]

[1]North China Institute of Aerospace Engineering;Hebei Province

November 23, 2022

Abstract

Plastic waste is one of the main factors causing environmental pollution and affecting biodiversity, and identification and detecting plastic waste is the premise of removal and treatment. Unmanned aerial vehicle (UAV) is gradually applied to identify and classify plastic waste because of its advantages of simplicity, convenience, and safe operation, but the current visual interpretation method is inefficient and cumbersome. To support the detection of plastic waste, researchers have developed various automatic and semi-automatic algorithms. Among these algorithms, deep learning technology has outstanding performance in river garbage detection, but there are also practical problems such as small floating garbage volume, sparse samples, complex garbage environment, In this paper, a classification plus target detection (C+D) model is proposed, and a lightweight floating plastic waste detection model based on deep learning is constructed. The EfficientNet classification algorithm and Yolov5 target detection algorithm are combined and improved for experimental verification, and various floating plastic wastes are automatically identified and located. In this paper, the UAV image data set obtained from the flight in Longhe River Basin, Langfang City, Hebei Province, China, is used to investigate the plastic floating garbage. The algorithm verification experiment shows that the detection accuracy of the three kinds of plastic garbage is higher than 85% (AP: plastic bag: 0.95; Plastic foam: 0.90; Plastic bottle: 0.87), which shows its excellent floating plastic recognition ability. The FPS of UAV equipment can reach 40.23 on edge, which shows that its recognition speed is fast and meets the real-time demand.

Hosted file

`essoar.10507932.1.docx` available at https://authorea.com/users/544841/articles/601842-a-deep-learning-model-for-automatic-plastic-waste-monitoring-using-unmanned-aerial-vehicle-uav-data

A Deep Learning Model for Automatic Plastic Waste Monitoring Using Unmanned Aerial Vehicle (UAV) Data

Wenlong Han [1,2], Wei Luo[1,2*], Yongtao Jin [1,2] and Mengxun Zhu[1,2]

[1] North China Institute of Aerospace Engineering,China

[2] Collaborative Innovation Center of Aerospace Remote Sensing Information Processing and Application of Hebei Province,China.

Key Points: (1) This model can accurately count the three types of plastic waste in UAV images. (2) As the data set increases, the classification algorithm can be multi-classified to solve practical problems. (3) The model can be reasonably extended to other fields.

Abstract

Plastic waste is one of the main factors causing environmental pollution and affecting biodiversity, and identification and detecting plastic waste is the premise of removal and treatment. Unmanned aerial vehicle (UAV) is gradually applied to identify and classify plastic waste because of its advantages of simplicity, convenience, and safe operation, but the current visual interpretation method is inefficient and cumbersome. To support the detection of plastic waste, researchers have developed various automatic and semi-automatic algorithms. Among these algorithms, deep learning technology has outstanding performance in river garbage detection, but there are also practical problems such as small floating garbage volume, sparse samples, complex garbage environment, In this paper, a classification plus target detection (C+D) model is proposed, and a lightweight floating plastic waste detection model based on deep learning is constructed. The EfficientNet classification algorithm and Yolov5 target detection algorithm are combined and improved for experimental verification, and various floating plastic wastes are automatically identified and located. In this paper, the UAV image data set obtained from the flight in Longhe River Basin, Langfang City, Hebei Province, China, is used to investigate the plastic floating garbage. The algorithm verification experiment shows that the detection accuracy of the three kinds of plastic garbage is higher than 85% (AP: plastic bag: 0.95; Plastic foam: 0.90; Plastic bottle: 0.87), which shows its excellent floating plastic recognition ability. The FPS of UAV equipment can reach 40.23 on edge, which shows that its recognition speed is fast and meets the real-time demand. This study shows that UAV technology combined with this deep learning algorithm can efficiently, accurately, and cheaply realize floating plastic waste detection.

1 Introduction

Plastic pollution has become one of the most critical global ecological and environmental problems. The global plastic output increases year by year, but only 9% of the nearly 10 billion tons of plastic once produced has been recycled((Fadeeva & Van Berkel.,2021) , and the unrecycled plastic waste eventually flows into the ocean (Lebreton et al.,2017). Rivers are the main conduit for land-based transport of waste to the oceans, with just 10 river systems transporting more than 90% of the global waste input, according to The Guardian . According to statistics, marine plastic waste accounts for about 60% to 80% of all waste, and in some areas, the proportion of plastic can reach 90% to 95% (Thompson.,2017;Thushari & Senevirathna., 2020). Plastic waste pollution causes ecological problems such as the entanglement of marine organisms by swallowing, habitat degradation, and chemical contamination (Haward et al.,2018 ;Monteiro et al.,2018;Haward, 2018)), which kills more than 1 million seabirds and 100,000 animals, including whales, dolphins, and seals every year from just one type of plastic waste, plastic bags((Thushari & Senevirathna., 2020). Meanwhile, plastics are non-biodegradable, and over time, large plastics decompose into microplastics that rise through the food chain and enter the human body, endangering health (Mukherjee et al., 2020). Current plastic waste in the ocean reaches 150 million tons, about one-fifth of the total weight of marine fish, and it is expected that by 2050 the weight of plastic in the ocean will exceed that of fish(da Costa et al., 2020). Preventing the negative impacts of marine plastics first requires understanding the sources and locations of plastics and their trends.

At present, plastic waste detection methods are mainly manual investigation, aerial survey, and satellite monitoring. The manual investigation, Generally, plastic waste is visually inspected along the cross-section, time-consuming, labor-intensive, and unsafe for operators. For example, to understand the source and impact of marine debris, the Australian Scientific and Industrial Research Organization organized thousands of volunteers to conduct a national garbage survey in 175 locations around Australia for 18 months and only counted 575 cross-sections. Aerial survey is also used for the marine plastic waste survey(Hardesty et al.,2017). Moy et al. (2018)carried out the aerial survey on each island of Hawaii to collect high-resolution photos, visually interpret orthophoto mosaic with a sampling distance of 2 cm on the ground, and draw a hot spot map of debris on the beach of Hawaii Island. The survey time and survey area for manned aircraft operation are flexible, but it is relatively expensive. Professional surveyors are needed, and the route is affected by human factors. Satellite monitoring usually uses satellite images with high spatial, temporal, and spectral resolutions. Davaasuren et al. (2018)used Sentinel-1A and COSMO-SkyMed SAR to identify microplastics in the ocean. Topouzelis et al. (2019) used WorldView-2 images to study the optical properties of wet plastics and dry plastics and evaluated the possibility of detecting floating plastics in water by multispectral images. Themistocleous et al. (2020)used Sentinel II satellite images to construct plastic index (PI) and reverse normalized difference vegetation index (RNDVI) to identify plastic wastes in the sea. Plastic index (PI) was able to identify plastic objects floating on the water surface accurately. Although current experts have validated the effectiveness of satellite monitoring of plastic waste, the accuracy of identifying plastic waste is still limited by temporal and spatial resolution.

Unmanned aerial vehicle (UAV) survey is gradually applied to plastic waste monitoring with its advantages of cheap, convenient, and safe operation. Compared with the manual investigation, UAV investigation not only ensures the safety of investigators in field activities but also reduces the influence of humans on some animals. Unlike aerial exploration, which requires high-quality professionals, a UAV survey is cost-effective and straightforward. Unmanned aerial vehicle (UAV) is equipped with advanced sensors, providing higher resolution images and a more flexible revisit cycle than satellite remote sensing. Deidun et al. (2018) used unmanned aerial vehicles to find stranded and floating marine garbage and generated the density map of beach garbage to help identify the same garbage. Geraeds et al. (2019)used images obtained from drones at different flight altitudes to manually mark river banks and floating plastic. The UAV survey ensured that although manual identification of classified river plastic litter in UAV images is relatively accurate and reliable, manually viewing UAV images is both time-consuming and laborious. Researchers have developed various machine learning algorithms to detect plastic waste in UAV images. C Martin et al. (2018) uses SVM and random forest algorithm to identify plastic debris detection and three types of classification in UAV images. Random forest algorithm is superior to support vector machine in multi-type plastic classification tasks and is several times higher than visual interpretation and recognition. Gonçalves et al. (2020) used multiple machine learning to compare detection and mapping of river trash objects, and the results showed that random forest had an F1 score of 70, which was slightly better than other methods. Although non-deep neural network-based machine learning methods produce better detection results in simple cases, this method cannot exploit complex trash features and has a limited accuracy rate.

With the enhancement of the computing power of graphics processors and the increase of open training data sets, deep learning(Lecun, Y et al 2015)is widely used in remote sensing image recognition and classification tasks ((Ma et al., 2019; Zhu et al., 2019; Liang et al., 2017), such as automatic classification, target detection, and semantic segmentation. The deep learning model has the advantages of automatically selecting image features, etc. The models such as VGGNet((Karen Simonyan et al.,2018), FCN(Zhuang et al., 2019), Faster R-CNN (Hanna & Cardillo., 2013), Yolo(Redmon et al., 2016), U-Net((Hammernik et al., 2017) have reached the most advanced accuracy in the marine plastic garbage detection task of UAV images. Kyriaki et al. (2019)used the VGG16 model to train on three types of plastic marine garbage (i.e., bottles, barrels, and straws), and the classifier can successfully identify the floating objects in front with a success rate of 86%. Li et al.(2020)used the modified YOLOv3 model to detect underwater marine life and debris floating on the sea surface, and the average accuracy was 69.6% and 77.2%, respectively. Jakovljevic et al. (2020) used the U-net model to identify and distinguish floating plastics in the ocean. The classification accuracy improved with the improvement of spatial resolution, and the F1 score was up to 92%, showing the ability to identify plastic types. Grays et al(2019)trained five deep learning networks and compared them. The results showed that the VGG19 model performed best with an accuracy rate of 77.6% and an F value of 77.42%. Although the deep learning algorithm has outstanding performance in UAV images, the current algorithm parameters are complex, and the detection rate is slow. Deep learning cannot be directly applied to marine plastic detection scenes when faced with the problems of small and sparse marine garbage samples and complex and changeable garbage environments. Therefore, the feasibility of deep learning plastic garbage detection for UAV images deserves further study in a wide range of hydrological environments.

The main objective of this study is to use UAV and deep learning to solve practical problems in floating plastic waste monitoring and propose a classification plus target detection (C+D) model, and fuse the EfficientNet (Tan et al.,2019) image classification algorithm and YoloV5 (ultralytics et al.,2020)target detection algorithm, and improve the YoloV5 target detection algorithm for plastic floating material features to build A lightweight floating plastic garbage detection model based on deep learning can automatically identify and locate a variety of floating plastic garbage. The experimental results show that the algorithm has excellent recognition ability for three kinds of experimental floating garbage: plastic bags, plastic foam, and plastic bottles, with mAP up to 0.91 and FPS up to 40.23, which can monitor plastic garbage in real-time.

2 Materials and Methods

2.1 Study Area and UAV Data

The drone image was taken in the Longhe River Basin, Langfang City, Hebei Province, China (as shown in Figure. 1). The Longhe River is typical in northern China, and the results of the environmental survey showed that some sections of the Long River are polluted by plastic waste. A total of 5682 RGB UAV images and 4,234 valid orthophotos were obtained from multiple shooting campaigns, covering an effective aerial survey area of about 20km2.

The image was taken by a fixed-wing drone equipped with an Intel d435i binocular depth camera. There are four round holes on the front of the D435 camera, including infrared and visible light sensors. The maximum distance captured by the camera can reach 10 meters, and the video transmission rate can reach 90fps. The flying height is set to 7 meters above the ground, and the image resolution is 5472×3678. These images captured the floating plastic garbage, and three kinds of garbage among plastic garbage, plastic bottles, plastic foam, and plastic bags, were selected as garbage for deep learning detection. Other garbage was not used as the object of this experimental study due to the lack of samples for training deep learning models spatial resolution.

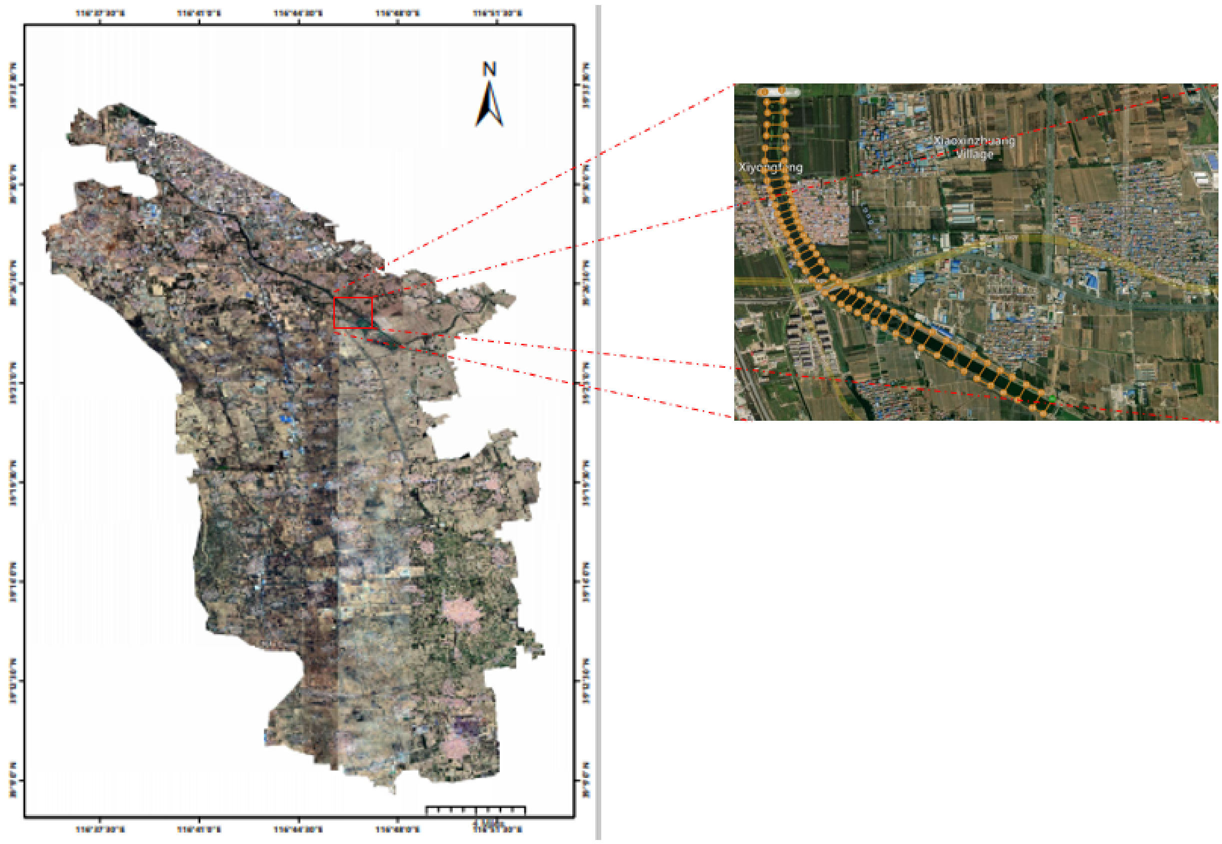


Figure 1. UAV survey area and flight trajectory

2.2. Data Preparation

Image classification and target detection are different task requirements and require separate preparation of the dataset. The image classification algorithm determines the presence or absence of plastic litter and classifies the data in two categories: water without litter or with other floating objects (plant branches, leaves, etc.) and water with floating plastic litter. The task was relatively simple, with one expert recognizing the differentiated images in the classification process, obtaining 1208 images containing plastic trash and 3026 images without plastic trash. Typical sample images are shown in Figure 2 in (a) and (b).

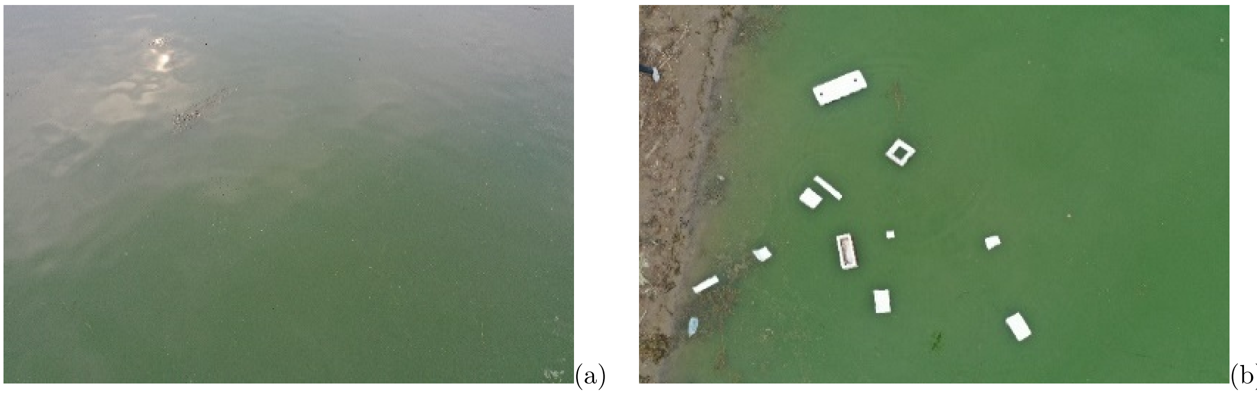
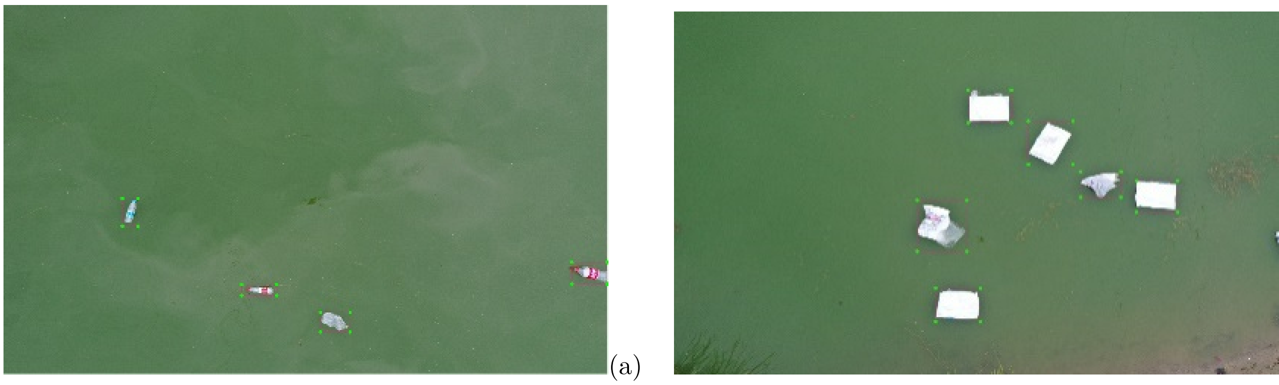


Figure 2. data for classification algorithm (a) Clean waterbody surface ,(b) Surface of waterbody with plastic waste.

The target detection algorithm needs to label the river garbage in the image, and the "labeling" (2019) graphic image annotation tool is used to draw the bounding box around each identified plastic garbage in the image. Two experts in two steps identified the plastic waste in the drone image, and the examples annotated by the two scientists were adopted as basic facts. Second, instances that one scientist only annotated are rechecked to verify whether they should be adopted as basic facts or discarded. The image marking process obtains images with three types of plastic waste. A typical sample diagram is shown in Figure 3 (a) (b)and (c). In order to enhance the reliability and robustness of the model, the data is augmented, including random cropping of the data, rotation, scaling, and horizontal flipping to generate multiple similar images.



(a)

(c)

Figure 3. data for target detection algorithm(a) plastic bottle,(b) polyfoam, (c) plastic bag.

2.3. Model

The target detection of floating plastic waste is very challenging, and there are problems such as low training data, highly imbalanced data sets, and frequent target position and scene changes caused by constant plastic movement. This paper proposes a classification plus target detection (C+D) modeling mode to solve the above problems. First, train the EfficientNet classification network on two categories (with garbage and no garbage) to simplify the garbage classification and identification problem, and ignore Details of various types of garbage to determine whether there is garbage. Then train three sub-categories on floating plastic waste to simplify the classification problem. Focus on the details of each type of floating waste on the Yolov5 target detection algorithm, and optimize the target detection algorithm by modifying the anchor frame, adding the attention mechanism and weighted frame fusion. Finally, the two models are integrated by setting the high and low thresholds. The overall process is shown in Figure 4. The specific details are introduced below.

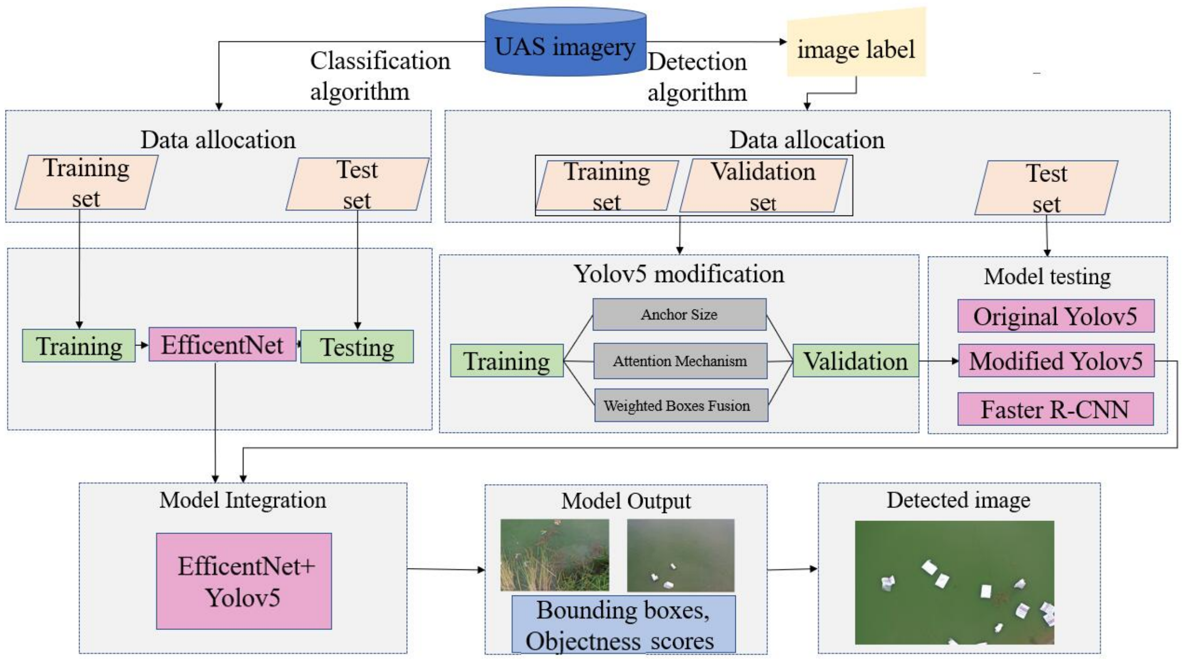
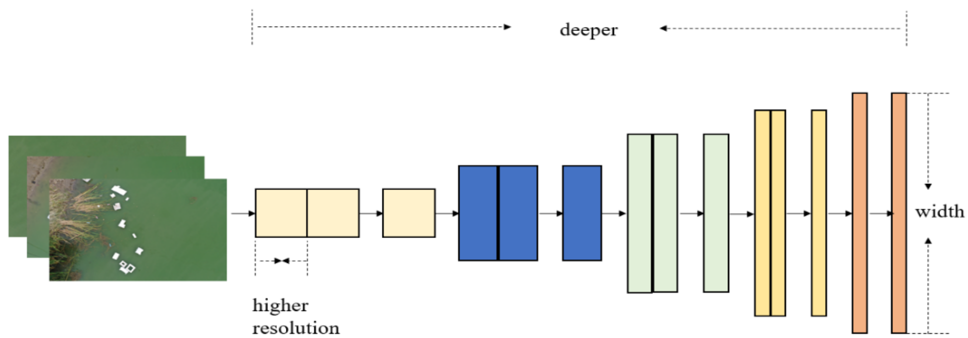


Figure 4. Overview of C+D model

2.3.1. EfficientNet Algorithm

EfficientNet [33] is currently one of the most outstanding image classification algorithms. Compared with other classification networks, it has higher accuracy and efficiency and reduces the parameter size and FLOPS order of magnitude. Its network structure is shown in Figure 5. Unlike the ResNet structure, EfficientNet only optimizes the depth of the network. It uses a simple and efficient composite coefficient to enlarge the network structure in a more structured manner. First, the composite coefficient is fixed, and then the depth, width, and resolution coefficient pairs are used for optimization. The network search, fix three coefficients, expand the baseline network through the compound adjustment formula, and achieve the purpose of weighing the three characteristics of the network model's depth, width, and resolution, and uniform scaling. The model obtained by the compound scaling tends to focus on and more The area related to the target details can distinguish different types of images well. Since the image classification task is mature and EfficientNet can automatically scale the network structure, the EfficientNet classification algorithm network structure has not been adjusted too much. The EfficientNet-B4 version is selected from EfficientNet, and the model accuracy rate is trained to more than 95% so that Junk



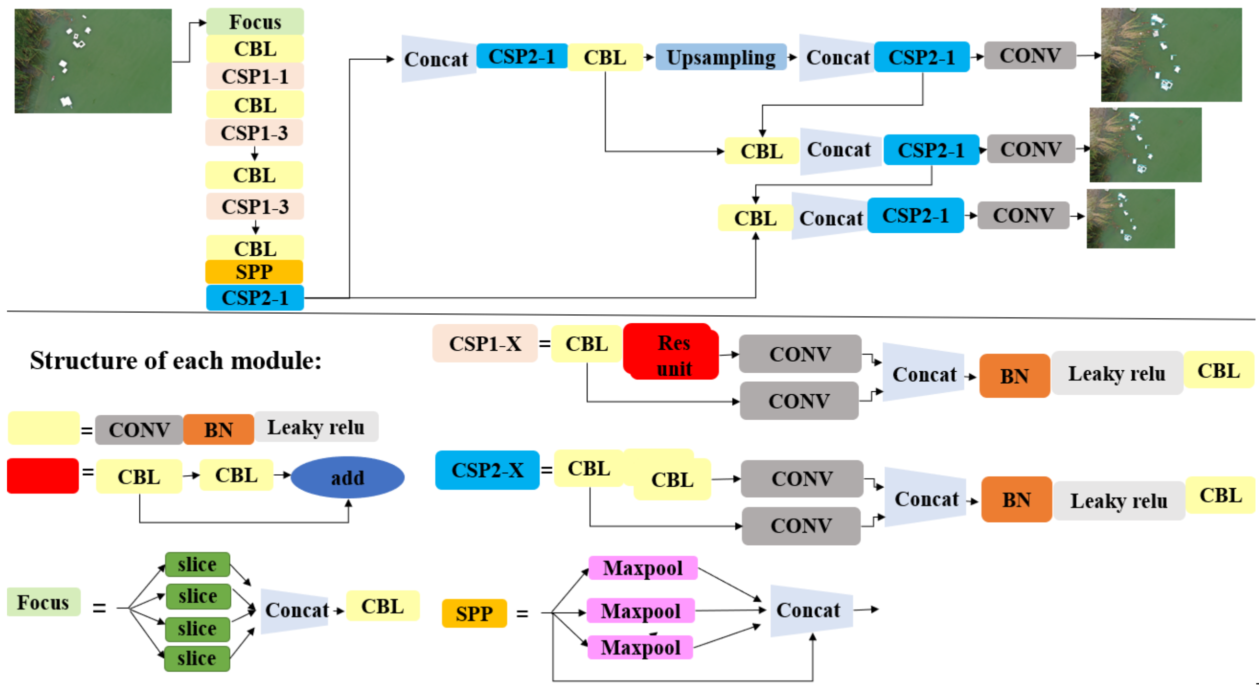
images are quickly judged.

Figure 5. EfficientNet algorithm structure

2.3.2. Modified Yolov5 Algorithm

As one of the most typical target detection algorithms, Yolo is a single-stage algorithm that integrates the object proposal stage and the classification stage, and the detection rate is better than the two-stage RNN algorithm. Yolov5 is the latest version of the YOLO architecture. Yolov5 inherits Yolov4's CSPDarknet as the backbone feature extraction network. It integrates the advantages of the typical Darknet and CSP structure, improves the learning efficiency of the network structure, and dramatically improves the feature extraction rate. The Neck part uses FPN plus PAN to generate a dense grid of reference frames, called "anchors," with a specified scale and aspect ratio on the feature map. FPN plus PAN pre-designates a score for each anchor point, which indicates whether the anchor point contains the top-ranked anchor point of the object of interest, which is retained as a target suggestion and input to the second stage of the network. Finally, the low-scoring bounding box is filtered by setting a threshold, non-maximum suppression (NMS) processing is performed on the retained bounding box, and the final detection result is obtained after the overlapping bounding box is removed and the target detection is completed. The YOLOv5 architecture includes four architectures [34], specifically named YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. In order to prevent the model from being too large and over-fitting, this paper chose YOLOv5s with a relatively simple structure as the baseline model. The specific structure is shown in Figure 6. In order to solve the actual problem of UAV detection of plastic waste, this research proposes a variety of optimization strategies based on Yolov5 target detection, including

anchor size revision, attention mechanism, and weighted frame fusion. The three strategies are specifically introduced below.



Fi

6. Yolov5s algorithm structure

The anchor structure, an important part of the Yolo series of target detection algorithms, produces suggestions for predicting potential objects. The original anchor structure performs well in detecting various objects in data sets such as COCO. However, the size of these anchors is not suitable for small objects, such as plastic waste in drone images, whose average size is less than 30 cm, and examples occupy about 1% of the total image area. In small target detection, setting a smaller anchor scale is a feasible solution to this problem. However, it is arbitrary to evaluate the performance of the model by comparing the anchor size and the sample size, and the model also can find a more suitable size through bounding box regression. In order to select the appropriate anchor size, we optimize the anchor size selection setting through the K-means clustering algorithm (Jain., 2010), which is the most commonly used iterative clustering algorithm among many clustering methods. K objects are randomly selected as the initial clustering centers, and then the distance between each object and each seed clustering center is calculated, and finally, each object is assigned to the nearest clustering center. The value of each clustering center is continuously updated during the iterative process until the best clustering result is obtained. The target detection is performed by experimentally setting three groups of anchor structures [45,62;25,20;16,28], [13,9;31,44;10,26], [24,54;15,21;23,30], as shown in Figure 7.

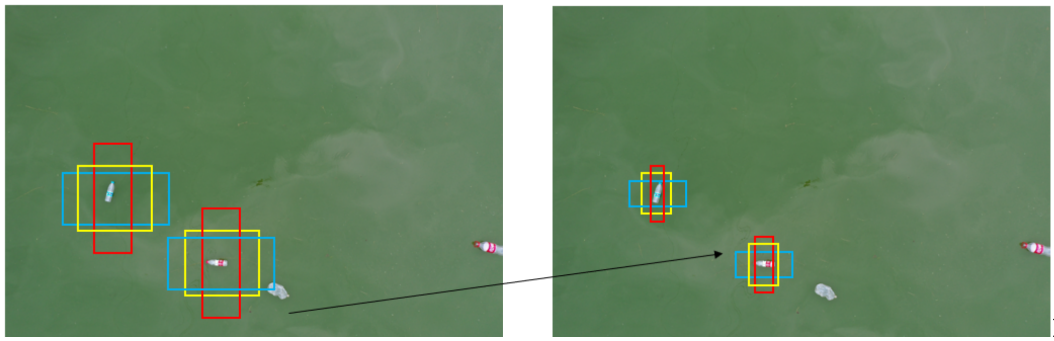
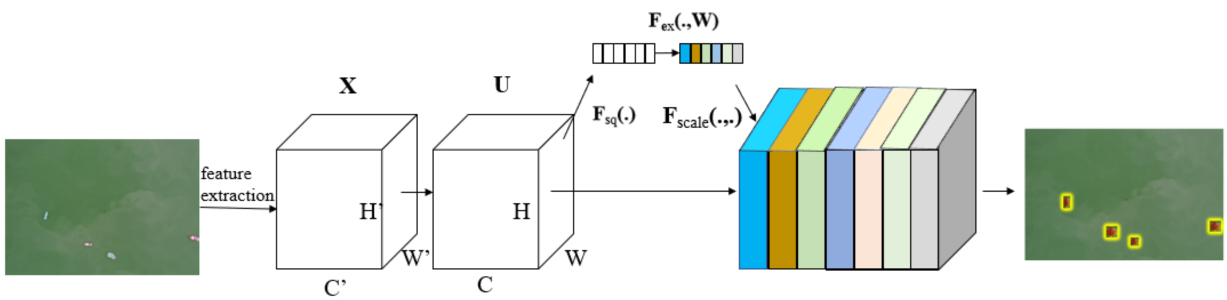


Figure 7. modification of anchor size



Riv

plastic waste spatial information is variable; the target is not easy to detect. In order to quickly and accurately identify river waste in the model to introduce an attention mechanism, the attention mechanism can ignore irrelevant information and focus on effective local information. Common attention mechanism modules include SE module (Woo et al., 2018), CBAM module(Hu et al., 2020), etc. In this paper, we choose the SE module to be introduced into the yolov5 model. The SE module is lightweight, and its structural features are shown in Figure 8. By processing the convolved feature map, compressing the spatial dimension through the squeeze operation, and learning the direct correlation of the channels by weights, a one-dimensional vector equal to the number of channels is finally obtained as the evaluation score of each channel, And then apply the modified scores to the corresponding channels. The results obtained can improve the sensitivity of the model to channel features. Only a small amount of calculation is needed to increase performance. Embed it in the backbone network of the YOLOv5s architecture. Improve the detection accuracy of the model.

Figure 8. SE attention module

Non-maximum suppression (NMS) is a common prediction frame generation algorithm used in target detection. Searching for local maxima and suppressing non-maxima elements has good results in the case that only a single object in the picture is detected. However, the classical NMS algorithm has some problems in drifting plastic detection, which is prone to the proximity and overlap of multiple floating objects in plastic floating trash due to the point of river flow. The classical NMS algorithm filters out one of the multiple prediction frames with low confidence, resulting in inaccurate counts, not conducive to later removal efforts. Weighted box fusion(Solovyev et al., 2021)is the new bounding box fusion method to solve the above problems of the target detection model. The weighted box fusion workflow is as follows: First, it sorts all bounding boxes in descending order of confidence scores. It then generates another list of possible box "fusions" and tries to check whether they match the original box this is achieved by checking whether IoU is more significant than a specified threshold. Then, it uses a formula to adjust the coordinates and the confidence scores of all boxes in the box list. The new confidence is the average confidence of all the boxes that are fused. The new confidence is the average confidence of all the boxes being fused, and the new coordinates are fused and weighted in a similar way to finally generate the most appropriate prediction box, generating the correct coordinates with the quantity information, and the NMS is compared with the WBF as shown in Figure 9.

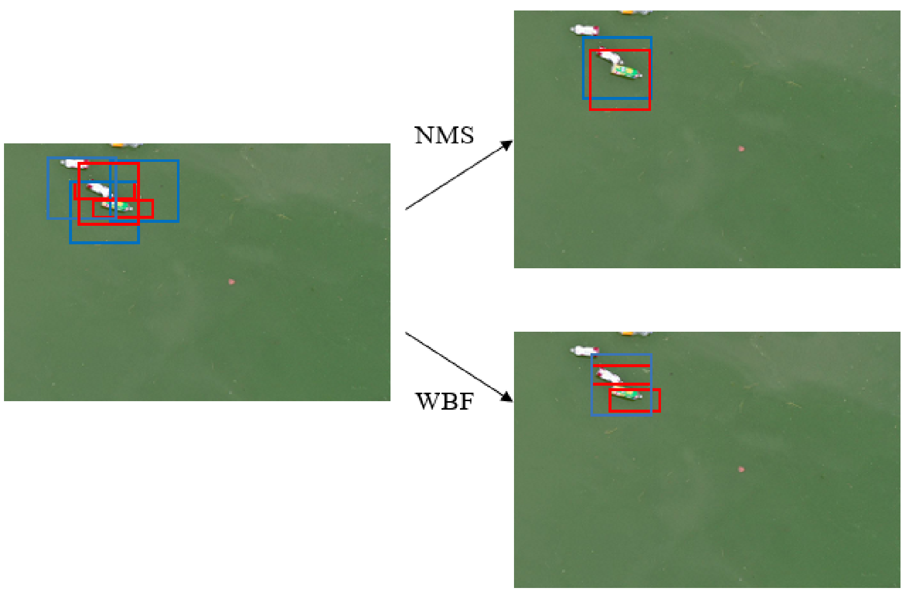


Figure 9. Comparison of NMS and WBF generated prediction frames

2.3.3. Model Integration

This study selects a simple and effective high and low threshold method to integrate the model and experimentally set the low threshold to 0.1 and the high threshold to 0.9. By checking each classification prediction. If the EfficientNet prediction probability is less than the low threshold, we set the prediction to ”no garbage.” If the classification prediction is between the low and high thresholds, a ”garbage” prediction is obtained, which has the confidence level of EfficientNet (not Yolov5). Finally, if the classification prediction is higher than the high threshold, it means that the yolov5 target detection network is highly confident that the water surface contains a high probability of floating plastic garbage. The floating plastic garbage is classified and identified, modeled as a classification combined with a detection model (see Fig10.)

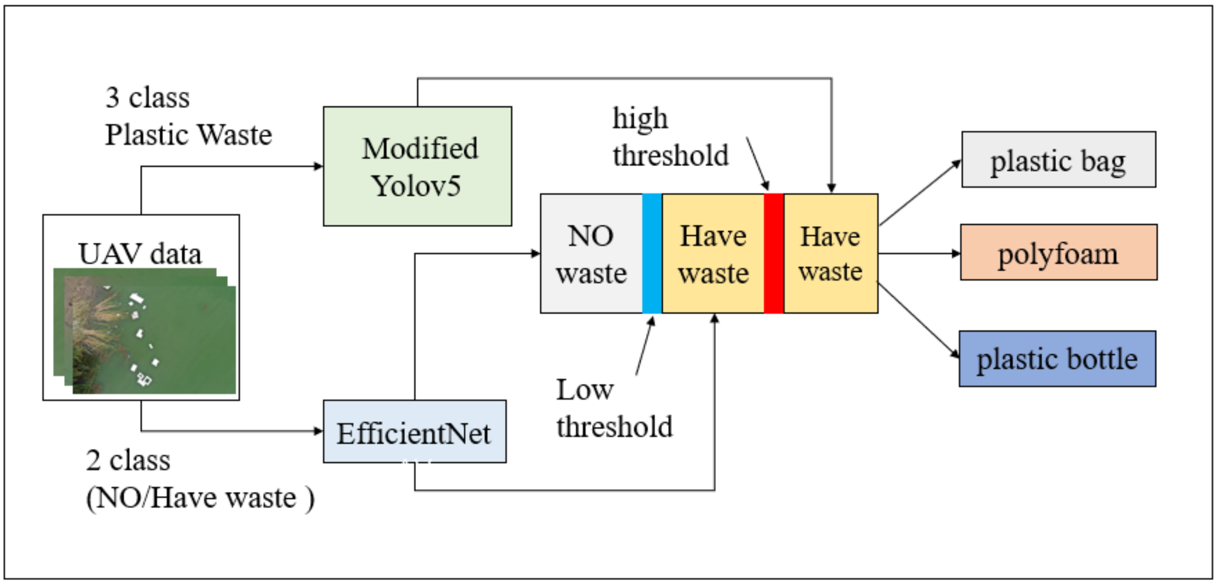


Fig **10.** C+D model

2.4. Model Evaluation

This study used FP, true positive (TP), and false negative (FN) assessment schemes. When a predicted bounding box corresponds to a unique garbage target, and the IOU threshold reaches 0.5, it is calculated as TP. Otherwise, the predicted bounding box will be regarded as FP. The prediction of plastic waste in the study is based on the recall rate and accuracy, which is defined as follows:

$$Precision = TP/(TP + FP) \quad (1)$$

$$Recall = TP/(TP + FN) \quad (2)$$

Since recall and precision only reflect part of the performance of the target detection model, they cannot respond well to the overall floating garbage classification and detection model. Therefore, FPS, average precision rate, and average accuracy rate of all categories are used to evaluate the results comprehensively. FPS is the number of pictures that can be processed per second. The higher the FPS, the faster the processing speed. FPS of more than 30 can meet the real-time requirement.

The average accuracy AP can be simply considered as the region under the accuracy-recall curve for the application, or mathematically expressed as:

$$\mathrm{AP} = \Sigma_{i=1}^{n} \mathrm{Precision}_i \left(\mathrm{Recall}_i - \mathrm{Recall}_{i-1}\right), \mathrm{withRecall}_{i=0} = 0 \quad (3)$$

The average accuracy of the whole class is the average value of the whole class AP, mathematically expressed as:

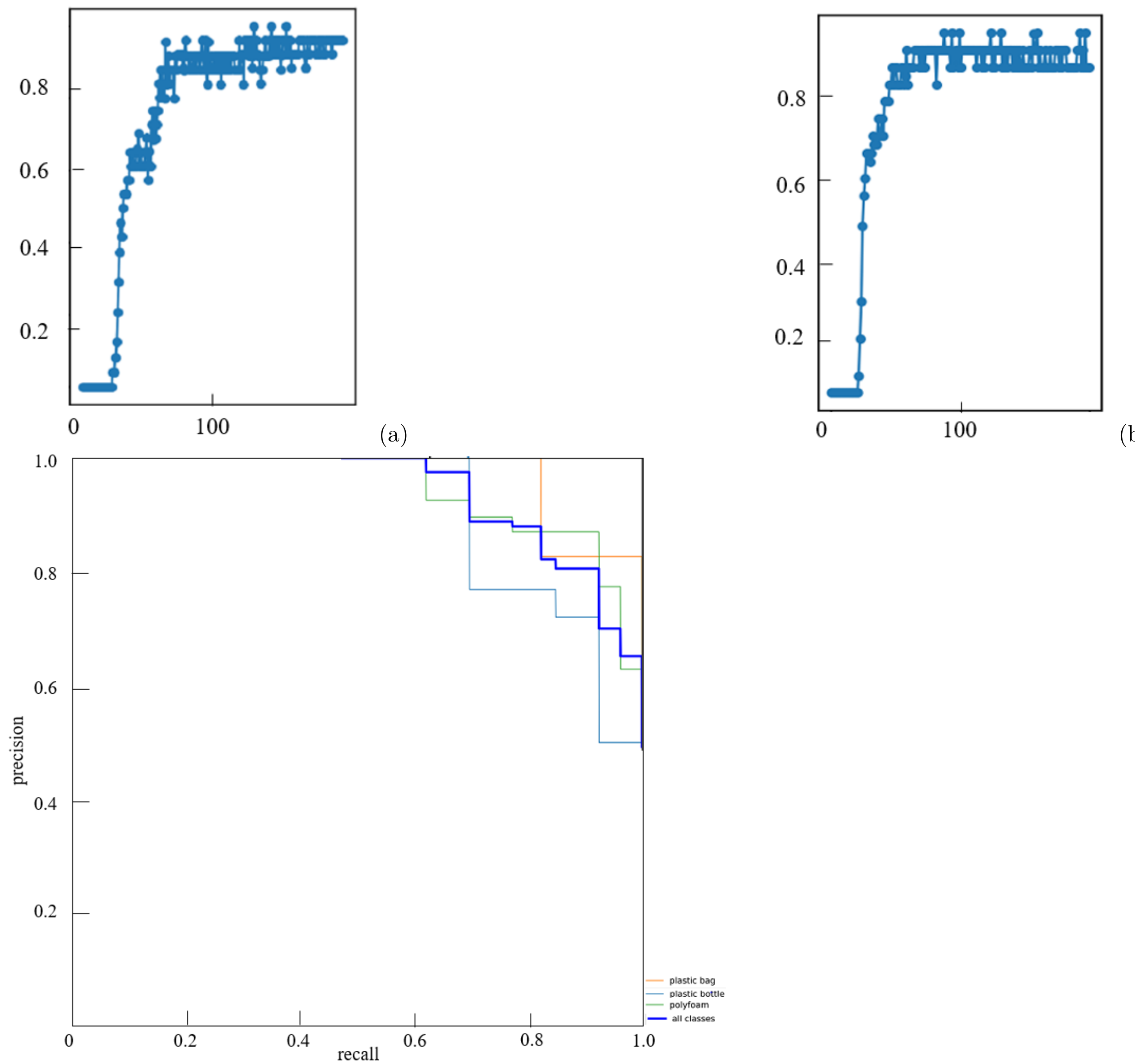
$$\mathrm{mAP} = \Sigma_{i=1}^{n} \mathrm{AP}/n \quad (4)$$

The score threshold of the algorithm is set to 0.8 to suppress low score prediction. High score predictions were compared with surface facts to obtain a set of TP, FP, FN, precision, recall, and AP, mAP.

3. Results

3.1. Training results of Modified Yolov5 algorithm

This section presents the results of floating plastic trash detection using deep learning. The training experiments in this study were conducted on a graphics workstation with an Intel Core i9-7900X CPU, an NVIDIA GeForce GTX 1080 Ti graphics card, and eight 8 GB memory cards. During the model training, stochastic gradient descent was used for network optimization. No overfitting problem was found in the experiments. The deep learning model was implemented on pytorch ("pytorch", 2019) and trained for 200 iterations. After about 6.5 hours of training, all target detection models converged, and the modified, modified Yolov5 P, R, AP, and mAP curves are shown in Figure 11.



(a) (b)

(c)

Figure 11. model training results. (a)precision curve (b)recall curve (c)AP, mAP curve

3.2. Comparison with the Recognition Results Using Different Object

Detection Algorithms

The model was tested with 300 UAV images in the test set, including 225 no-junk images, 75 junk images, and 323 junk examples. Compare the improved YOLOv5s network with the original YOLOv5s and Faster-RCNN on the 300 test set images under study. The target detection evaluation results are shown in Table 1. According to Table 1, Modified Yolo v5's P, R, and mAP were 0.86, 0.89, 0.86, respectively. The AP of plastic bottles plastic bags, and polyfoams were 0.80, 0.89 0.87, respectively. In terms of recognition accuracy, the algorithm proposed in the mAP study of the improved YOLOv5s recognition model is up to 0.86, which is 7% higher than the original YOLOv5 network, and the accuracy of plastic bottles, plastic tubes, and plastic foams are increased by 2%, 6%, and 11 respectively. %, the result is slightly better than the two-level network Faster-CNN. In terms of the recognition speed of the model, the average detection speed of the improved YOLOv5s model is 45.63 FPS, a decrease of 2.74, slightly lower than the original structure, and about three times faster than Faster-CNN. The results show that the modified yolov5 model balances floating plastic's real-time performance and accuracy, and the algorithm optimization strategy is effective.

Table 1. Performance comparison of three object detection networks.

Network	FPS	P	R	AP_{bottle}	AP_{bag}	$AP_{polyfoam}$	mAP
Faster-CNN							
Yolo v5							
Modified Yolo v5							

3.3. Results of Plastic Waste Targets Detection

EfficientNet and Modified Yolo v5 were integrated according to the high and low threshold methods and recognized on the test set of 300 images using the classification plus target detection algorithm. P, R, and mAP of EfficientNet+Modified Yolov5 were 0.93, 0.92, and 0.91, respectively. The AP of plastic bottles, plastic bags and polyfoam were 0.87, 0.95 0.90, respectively. Plastic waste targets detection results are shown in Figure. 12, plastic bottles in the red border, plastic foam in the blue border, and plastic bags in the purple border, from the results it can be seen that most of the floating plastic garbage is identified, and the IoU threshold is kept chiefly above 0.9.

Figure 12. Plastic waste targets detection results

3.4. Comparison with the Recognition Results Using C+D Model

To verify the gain effect of the classification algorithm on the target detection algorithm, the trained classification network EfficientNet was introduced in front of the three target detection models, respectively, and then tested and compared the application effect on the dataset. The evaluation results are shown in Table

2. According to Table 2, the EfficientNet classification algorithm reduced the Faster-CNN, Yolo v5, and Modified Yolo v5 target detection algorithms FPS by 4.92, 5.13, and 5.4, and mAP increased by 5%, 6%, and 5%, respectively. The results show that EfficientNet reduces FPS by a small amount and has a sound improvement inaccuracy. Among them, the accuracy of the EfficientNet+Modified Yolo v5 model reached the highest mAP of 0.91, and the accuracy of plastic bottles, plastic tubes, and plastic foams were 0.87, 0.95, 0.90, and the FPS could also reach 40.23, meeting real-time and accuracy requirements.

Table 2. Performance comparison of three C+D Model

Network	**FPS**	**P**	**R**	AP_{bottle}	AP_{bag}	$AP_{polyfoam}$	**mAP**
EfficientNet+ Faster-CNN							
EfficientNet+ Yolo v5							
EfficientNet+ Modified Yolo v5							

4. Discussion

Experimental results show that Modified yolov5 is an excellent floating object detection algorithm. The Faster-RCNN algorithm is equivalent to the algorithm in this paper in terms of accuracy. However, it cannot achieve real-time performance due to the computational burden of the two-level network and is not suitable for river litter monitoring in complex environments. By modifying the original yolov5, the three optimization strategies have different degrees and effects of plastic identification and classification optimization. The K-mean clustering algorithm is used to adjust the anchor frame to focus on solving the disadvantages of small target river garbage that is difficult to identify. Adding an attention mechanism makes the algorithm more quickly focus on the plastic garbage target in the UAV image and integrate the plastic garbage information. The weighted frame fusion is to fuse and filter the prediction frame to locate the plastic waste more accurately. The plastic detection accuracy can reach 0.86 through strategy optimization, and the FPS can reach 45.63 the fastest. Although the model is improved, the network parameter structure is increased to reduce the FPS rate by about 3, but A 7% improvement is achieved. The overall performance of the algorithm structure is superior to other algorithms. In a complex and changeable river environment, the algorithm can process images in real-time and provide timely information on plastic floating objects for UAVs and other equipment, but the river environment is complicated. In addition, frequent target position and scene changes caused by continuous plastic movement can easily lead to misjudgment by the target detection algorithm.

Through C+D algorithm mode, Modified Yolov5 combines with robust EfficientNet image classification algorithm to achieve more excellent results in water surface floating objects. The current C plus D algorithm is slightly slower than the target detection algorithm in terms of speed, mainly because the algorithm needs to judge the accuracy of the algorithm according to the threshold value after processing. In contrast, the current classification algorithm is limited by the collective amount of data; with the expansion and completion of the data set, the classification algorithm will be further effective in terms of speed. 40.23 and mAP can reach 0.91, which meets the practical requirements. The algorithm can effectively overcome the main practical problems and investigate the plastic river waste in a more extensive range while satisfying the accuracy and real-time requirements.

The AP accuracy rate of the three types of plastic waste is above 85%. The main reason for limiting the accuracy is that the occlusion of riverside reeds and other plants causes abnormal light, angle, and time. Some floating objects sink and cause feature changes to cause inaccurate model recognition, as shown in Fig13(a). Among the three types of plastic garbage, the floating plastic bag has a single characteristic, and the highest accuracy is 0.95, while the plastic foam shape and color are slightly different, and the accuracy is slightly lower than that of the plastic bag, which is 0.90, as shown in 3 in Fig.13(b). The lowest accuracy of plastic bottles is 0.87, mainly due to the large difference in the volume and size of plastic bottles and the failure to identify some flat and compressed plastic bottle samples, as shown in Fig.13, 1and 2.

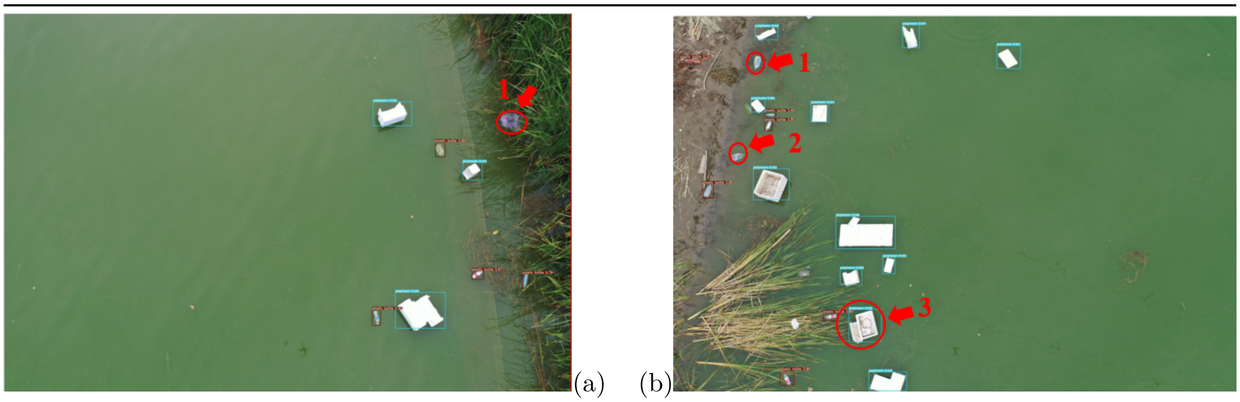


Figure 13. Error monitoring example

Based on the above experimental results, the algorithm proposed in this paper is superior to previous studies and solves the various drawbacks of floating garbage that are difficult to identify and monitor, and is suitable for real-time dynamic monitoring of river floating garbage, and can have better robustness in complex water environment with the expansion of data set.

5. Conclusions

A fusion classification detection deep learning algorithm is proposed in this paper to solve various problems arising from remote sensing monitoring of river garbage. The classification algorithm regulates the imbalance between river background images and plastic garbage samples and solves the problem of confusing other floating objects with plastic garbage to a certain extent. The target detection algorithm and three strategies solve the problem of sparse samples in actual river garbage monitoring, the environment is complex and changeable, and some garbage is not easy to detect. The classification and monitoring algorithm can reach 91mAP on GTX 1080 (AP: plastic bag: 0.95; plastic foam: 0.90; plastic bottle: 0.87), FPS can reach 40.63, ensuring the real-time and accuracy of floating garbage detection, making the river float Garbage monitoring has become intelligent and automated. It is cheap, effective, and accurate to carry the development algorithm on the UAV platform of edge computing to identify the plastic waste in the river channel, reduce the influence of subjective factors of manual identification and classification, obtain more accurate floating plastic waste information, and reduce the investment in human resources. Ensure the safety of cleaners, improve the efficiency of the waste cleaning industry, and make outstanding contributions to the protection of the water surface environment. Combining knowledge of other disciplines such as water environment ecology can further research and explore the source, spatial distribution, and dynamic waste changes.

Current algorithms are currently only accurate real-time extraction of floating plastic trash in rivers. However, this research model approach provides a new way of thinking about classification plus detection deep learning algorithms in remote sensing target detection is worth exploring. With the development of deep learning, various algorithms can be selected to combine according to different remote sensing task requirements to enhance the experimental results. Meanwhile, the classification categories and levels can be increased or decreased according to the complexity of the features to solve the practical problems arising from remote sensing target detection. The main result of this research is to propose a remote sensing feature detection algorithm model of classification plus detection, apply relevant deep learning algorithms for plastic floating object detection on a more complex river environment, and eventually, verify the effectiveness of the algorithm model.

Acknowledgments:

Funded by Key Laboratory of Surveying and Mapping Science and Geospatial Information Technology of Ministry of Natural Resources Open Research Fund Project Contract(2020-2-5)and Foundation of Hebei Educational Department(QN2019213)

Data Availability Statement :

Data for this research are available in ScienceDatabank (doi/10.11922/sciencedb.01121). Please contact us directly for more information.

References

da Costa, J. P., Mouneyrac, C., Costa, M., Duarte, A. C., & Rocha-Santos, T. (2020). The Role of Legislation, Regulatory Initiatives and Guidelines on the Control of Plastic Pollution. Frontiers in Environmental Science, 8(July), 1–14. https://doi.org/10.3389/fenvs.2020.00104Davaasuren, N., Marino, A., Boardman, C., Alparone, M., Nunziata, F., Ackermann, N., & Hajnsek, I. (2018). Detecting microplastics pollution in world oceans using SAR remote sensing. International Geoscience and Remote Sensing Symposium (IGARSS), 2018-July, 938–941. https://doi.org/10.1109/IGARSS.2018.8517281Deidun, A., Gauci, A., Lagorio, S., & Galgani, F. (2018). Optimising beached litter monitoring protocols through aerial imagery. Marine Pollution Bulletin, 131(February), 212–217. https://doi.org/10.1016/j.marpolbul.2018.04.033Fadeeva, Z., & Van Berkel, R. (2021). 'Unlocking circular economy for prevention of marine plastic pollution: An exploration of G20 policy and initiatives.' Journal of Environmental Management, 277(October 2020). https://doi.org/10.1016/j.jenvman.2020.111457Geraeds, M., van Emmerik, T., de Vries, R., & bin Ab Razak, M. S. (2019). Riverine plastic litter monitoring using Unmanned Aerial Vehicles (UAVs). Remote Sensing, 11(17), 6–8. https://doi.org/10.3390/rs11172045Gonçalves, G., Andriolo, U., Pinto, L., & Duarte, D. (2020). Mapping marine litter with Unmanned Aerial Systems: A showcase comparison among manual image screening and machine learning techniques. Marine Pollution Bulletin, 155(February), 111158. https://doi.org/10.1016/j.marpolbul.2020.111158Gray, P. C., Fleishman, A. B., Klein, D. J., McKown, M. W., Bézy, V. S., Lohmann, K. J., & Johnston, D. W. (2019). A convolutional neural network for detecting sea turtles in drone imagery. Methods in Ecology and Evolution, 10(3), 345–355. https://doi.org/10.1111/2041-210X.13132Hammernik, K., Tobias, W., Pock, T., & Maier, A. (2017). Invited Talk : U-Net Convolutional Networks for Biomedical Image, 54345. https://doi.org/10.1007/978-3-662-54345-0Hanna, E., & Cardillo, M. (2013). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. Biological Conservation, 158, 196–204.Hardesty, B. D., Lawson, T. J., van der Velde, T., Lansdell, M., & Wilcox, C. (2017). Estimating quantities and sources of marine debris at a continental scale. Frontiers in Ecology and the Environment, 15(1), 18–25. https://doi.org/10.1002/fee.1447Haward, M. (2018). Plastic pollution of the world's seas and oceans as a contemporary challenge in ocean governance. Nature Communications, 9(1), 9–11. https://doi.org/10.1038/s41467-018-03104-3Hu, J., Shen, L., Albanie, S., Sun, G., & Wu, E. (2020). Squeeze-and-Excitation Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(8), 2011–2023. https://doi.org/10.1109/TPAMI.2019.2913372Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 31(8), 651–666. https://doi.org/10.1016/j.patrec.2009.09.011Jakovljevic, G., Govedarica, M.,

& Alvarez-Taboada, F. (2020). A deep learning model for automatic plastic mapping using unmanned aerial vehicle (UAV) data. Remote Sensing, 12(9). https://doi.org/10.3390/RS12091515Karen Simonyan∗ & Andrew Zisserman+. (2018). VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION Karen. American Journal of Health-System Pharmacy, 75(6), 398–406.Kylili, K., Kyriakides, I., Artusi, A., & Hadjistassou, C. (2019). Identifying floating plastic marine debris using a deep learning approach. Environmental Science and Pollution Research, 26(17), 17091–17099. https://doi.org/10.1007/s11356-019-05148-4Lebreton, L. C. M., Van Der Zwet, J., Damsteeg, J. W., Slat, B., Andrady, A., & Reisser, J. (2017). River plastic emissions to the world's oceans. Nature Communications, 8, 1–10. https://doi.org/10.1038/ncomms15611Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539Li, X., Tian, M., Kong, S., Wu, L., & Yu, J. (2020). A modified YOLOv3 detection method for vision-based water surface garbage capture robot. International Journal of Advanced Robotic Systems, 17(3), 1–11. https://doi.org/10.1177/1729881420932715Liang, H., Sun, X., Sun, Y., & Gao, Y. (2017). Text feature extraction based on deep learning: a review. Eurasip Journal on Wireless Communications and Networking, 2017(1), 1–12. https://doi.org/10.1186/s13638-017-0993-1Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing, 152(March), 166–177. https://doi.org/10.1016/j.isprsjprs.2019.04.015Martin, C., Parkes, S., Zhang, Q., Zhang, X., McCabe, M. F., & Duarte, C. M. (2018). Use of unmanned aerial vehicles for efficient beach litter monitoring. Marine Pollution Bulletin, 131(January), 662–673. https://doi.org/10.1016/j.marpolbul.2018.04.045Monteiro, R. C. P., Ivar do Sul, J. A., & Costa, M. F. (2018). Plastic pollution in islands of the Atlantic Ocean. Environmental Pollution, 238, 103–110. https://doi.org/10.1016/j.envpol.2018.01.096Moy, K., Neilson, B., Chung, A., Meadows, A., Castrence, M., Ambagis, S., & Davidson, K. (2018). Mapping coastal marine debris using aerial imagery and spatial analysis. Marine Pollution Bulletin, 132(June), 52–59. https://doi.org/10.1016/j.marpolbul.2017.11.045Mukherjee, A., Misra, S., Sukrutha, A., & Raghuwanshi, N. S. (2020). Distributed aerial processing for IoT-based edge UAV swarms in smart farming. Computer Networks, 167. https://doi.org/10.1016/j.comnet.2019.107038Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem, 779–788. https://doi.org/10.1109/CVPR.2016.91Solovyev, R., Wang, W., & Gabruseva, T. (2021). Weighted boxes fusion: Ensembling boxes from different object detection models. Image and Vision Computing, 107. https://doi.org/10.1016/j.imavis.2021.104117Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.Themistocleous, K., Papoutsa, C., Michaelides, S., & Hadjimitsis, D. (2020). Investigating detection of floating plastic litter from space using sentinel-2 imagery. Remote Sensing, 12(16). https://doi.org/10.3390/RS12162648Thompson, R. (2017). Environment: A journey on plastic seas. Nature, 547(7663), 278–279. https://doi.org/10.1038/547278aThushari, G. G. N., & Senevirathna, J. D. M. (2020). Plastic pollution in the marine environment. Heliyon, 6(8), e04709. https://doi.org/10.1016/j.heliyon.2020.e04709Topouzelis, K., Papakonstantinou, A., & Garaba, S. P. (2019). Detection of floating plastics from satellite and unmanned aerial systems (Plastic Litter Project 2018). International Journal of Applied Earth Observation and Geoinformation, 79(March), 175–183. https://doi.org/10.1016/j.jag.2019.03.011ultralytics. yolov5. Available online: https://github.com/ultralytics/yolov5 (accessed on 18 May 2020)Wabnitz, C., & Nichols, W. J. (2010). Editorial: Plastic Pollution: An Ocean Emergency. Marine Turtle Newsletter, (129), 1–4. Retrieved from http://search.proquest.com/docview/924334169?accountid=27795Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11211 LNCS, 3–19. https://doi.org/10.1007/978-3-030-01234-2_1Zhu, M., He, Y., & He, Q. (2019). A Review of Researches on Deep Learning in Remote Sensing Application. International Journal of Geosciences, 10(01), 1–11. https://doi.org/10.4236/ijg.2019.101001Zhuang, J., Yang, J., Gu, L., & Dvornek, N. (2019). Fully Convolutional Networks for Semantic Segmentation. Proceedings - 2019 International Conference on Computer Vision Workshop, ICCVW 2019, 847–856. https://doi.org/10.1109/ICCVW.2019.00113