Object Detection for Small Water Floater

Contributor: Fuxun Chen, Lanxin Zhang, Siyu Kang, Lutong Chen, Honghong Dong, Dan Li, Xiaozhu Wu

Object detection is one of the most widely used applications in UAV missions. Detection of small objects in unmanned aerial vehicle (UAV) images remains a persistent challenge due to the limited pixel values and interference from background noise.

1. Introduction

Water is the source of life, and oceans and rivers cover about 71% of the Earth's area [1]. Rivers constitute a vital component of the global water cycle and serve a crucial function in facilitating the transfer [2]. However, recent global economic expansion and urbanization have caused significant damage to the natural environment [3] and severe water pollution [4]. Therefore, real-time monitoring of water environments has become crucial due to the ongoing deterioration of water quality. At present, the water's surface is susceptible to problems, such as exposure or dim light due to the intensity of sunlight, which results in less information about the target features and difficult recognition results.

Object detection is one of the most widely used applications in UAV missions. Due to the angle and flight height of UAV photography, small objects account for a large proportion of UAV images compared to general scenes [5]. Small object detection has been one of the key difficulties in the field of object detection [6] with problems, such as invalid image feature information and blurred object features. To solve this problem, many scholars have proposed solutions. Xia et al. proposed an automated driving system (ADS) data acquisition and analytics platform for vehicle trajectory extraction, reconstruction, and evaluation. In addition to collecting various sensor data, the platform can also use deep learning to detect small targets, such as vehicle trajectories [7]. Liu et al. proposed a model for tassel detection in maize, named yolov5-tassel. The authors trained the yolov5-tassel model based on UAV remote sensing images and achieved an mAP value of 44.7%, which is an improvement compared to FCOS, RetinaNet, and YOLOv5 in terms of detecting small tassels in maize [8]. Liu et al. proposed multibranch parallel feature pyramid networks (MPFPN) to detect small objects under UAV images. The MPFPN model applies the SSAM attention module [9] to attenuate the effect of background noise and uses cascade architecture in the Fast R-CNN stage to achieve a more powerful localization [10]. Chen et al. proposed small object detection networks based on a classification-oriented super-resolution generative adversarial network (CSRGAN), which is a model that adds classification branches and introduces classification losses into a typical SRGAN [11]. The experimental results demonstrate that CSRGAN outperforms VGG16 in classification [12]. In this research, UAV technology and deep learning are applied to small object detection at the same time. However, the water's surface is susceptible to problems, such as exposure or dim light due to the intensity of sunlight, which results in less information about the water surface floater's features and difficult recognition results. Deep learning models also face some problems when detecting a small water floater, such as small water surface floaters losing some of their features in the process of down-sampling, which leads to a lack of ability to extract features from global and low detection accuracy and poor feature recognition of a small water surface floater. Therefore, the direct application of the original YOLOv5 to UAV image object detection is not very effective. At the same time, UAV is also known as a powerful complement to the conventional water environment and assessment and has been gradually applied to water environment detection. So, this research proposes an efficient and accurate method based on YOLOV5 for the detection of small water surface floaters in UAV-captured images. The model can effectively locate the water surface floater and thus assist in the water surface floater salvage work.

2. Object Detection

Computer vision research is mainly concerned with image classification, object detection, object tracking, semantic segmentation, and instance segmentation. Object detection is one of the most fundamental and challenging tasks in the field of computer vision [13]. Exploring efficient real-time object detection models has been a hot research topic in recent

years [14]. Traditional object detection methods usually include region proposal, feature extraction, feature fusion, and classifier training, all of which require the laborious manual production of object features and must be completed step by step [15]. Therefore, traditional object detection methods have the drawbacks of massive redundancy, time-consuming computation, and low accuracy. With the rapid development of deep learning theory, deep learning has made great progress in the field of computer vision, such as image classification and object detection [16]. Since then, object detection has entered a new phase. In the 2012 ImageNet competition, compared with the traditional method, A. Krizhevsky et al. used a convolutional neural network (CNN) [17] to make image classification results much improved. CNN gives region suggestions through a region suggestion network (RPN) [18] at a low cost, which can significantly improve the efficiency of object detection. In 2014, R. Girshick et al. [19] used a region convolutional neural network for the first time in the field of object detection. The detection results were much improved. Compared to traditional methods, this end-to-end network is more popular because it reduces the complex steps, such as data preprocessing and manual design of object features [20]. Since then, deep learning has started to develop rapidly in object detection and has been widely used in practice [21].

There are two major categories of mainstream deep learning object detection algorithms: One is the one-stage object detector, such as the YOLO series, SSD [22], RetinaNet [23], FCOS [24], etc. The second is two-stage object detectors, such as R-CNN [25], Fast R-CNN [26], Faster R-CNN [27], Mask R-CNN [28], etc. For example, the two-stage object detector Faster R-CNN integrates RPN (region proposal network) on top of Fast R-CNN. The Faster R-CNN mainly includes the shared convolutional layer module, RPN, and Fast R-CNN detector. However, due to the large number of prediction frames of the two-stage object detector and the large computational effort and slow detection speed, it is not suitable for real-time detection tasks. In 2015, J. Redmon et al. proposed a one-stage object detector, You Only Look Once (YOLO) [29], which directly obtains the class and bounding box information of the object and greatly improves the object detection speed. Ma et al. proposed an improved YOLOv3 [30], which has stronger noise immunity and better generalization ability. Deep learning also has good applications in the detection of the water surface floater. Lieshoutet et al. constructed a plastic floating debris monitoring dataset using videos collected from five different waterway bridges in Jakarta, Indonesia, and used Faster R-CNN to first detect regions that may contain plastic floating materials and then Inception V2 [31], pre-trained based on COCO [32], to determine whether these regions are packed with plastic [33]. Li et al. acquired ocean images of the South China Sea using a camera mounted on an unmanned ship and used a fusion model based on YOLOv3 with DenseNet [34] to detect sea surface objects [35]. Stofa et al. selected the DenseNet model to detect ships in remotely sensed images and determined by fine-tuning the hyperparameters a batch size of 16 and a learning rate of 0.0001. The model uses the Adam optimizer [36] and optimized the parameter settings [37]. In contrast, YOLOv5 [38] discarded the candidate frame generation phase and directly performed classification and regression operations on the objects, which improved the real-time detection speed of the object detection algorithm, and the model complexity of YOLOv5 was reduced by about 10% compared to YOLOv4 [39]. YOLOv5 has a stronger generalization capability and a lighter network structure compared to other object detection models. Moreover, generally speaking, the computational resources of UAV platforms are limited, and the impact of the model on the object detection speed must be considered. One-stage segment object detector is more suitable for object detection in UAV platforms compared to the two-stage object detector. Therefore, the YOLOv5 model is finally selected as the original model for small object detection of a water surface floater in this reseach.

Small objects usually have lower resolution and less distinctive features. Therefore, achieving precise detection of small objects is a hot issue in the field of object detection. Many scholars have conducted a lot of research work on small object detection. Kim et al. inserted a high-resolution processing module (HRPM) and a sigmoid fusion module (SFM), which not only reduced computational complexity but also improved the detection accuracy of small targets. They obtained good detection results in drone reconnaissance images and small vehicles [40]. Wang et al. proposed a bidirectional attention network called BANet, which solved the problems of inaccurate and inefficient detection of small and multiple targets. The model achieved an AP improvement of 0.55–2.93% compared to YOLOX on the VOC2012 dataset and a, AP improvement of 0.3–1.01% on the MSCOCO2017 dataset [41]. Yang et al. proposed QueryDet, which uses a new query mechanism called cascade sparse query (CSQ) to speed up inference and calculate detection results using sparse queries. The model obtains high-resolution feature maps by avoiding redundant calculations in the background area. QueryDet is applied in FCOS and Faster-R CNN and is tested on the COCO dataset and the visDrone dataset for small objects, achieving better results than the original algorithms in terms of accuracy and inference speed [42]. Liu et al. addressed the problem of low accuracy in small target recognition by changing the ROI alignment method, which reduced the quantization error of Faster R-CNN and improved its accuracy by 7% compared to the original model [43].

3. Object Detection in UAV-Captured Images

UAV has been widely applied in industries, such as agriculture, forestry, electric power, atmospheric detection, and mapping, among others. Compared with traditional methods, UAVs [44] have the advantages of flexibility and mobility, high efficiency and energy saving, diversified results, and low operating costs. Lan et al. combined UAV technology and deep learning to complete the detection of diseased plants and abnormal areas in orchards using the Swin-T YOLOX lightweight model. Additionally, the improved model compared with the original YOLOVX had a detection accuracy of 1.9% [45]. Bajić et al. used the improved YOLOV5 for object detection of unexploded remnants of war based on UAV thermal imaging, and the accuracy was above 90% for all 11 detection objects [46]. Liang et al. proposed a UAV-based low-altitude remote-sensing-based helmet detection system, which improved the AP value of small object detection to 88.7% and significantly improved the network's detection performance for small objects [47]. Wang et al. built enhanced CSPDarknet (ECSP) and weighted aggregate feature re-extraction pyramid modules (WAFR) based on the YOLOX-nano network, which solved the problem of low recognition accuracy for grazing livestock due to the small number of occupied pixels [48]. However, due to the limited number of pixels and features occupied by objects in UAV images, the detection accuracy is low. Thus, small object detection based on UAV imagery remains a significant challenge.

4. Object Detection for Small Water Floater

As deep learning is increasingly applied to object detection, some researchers have also applied it to water surface floater detection. The application of deep learning in floating objects on the water's surface has contributed greatly to timely detection and processing of floating objects and to improving the supervision level of rivers and lakes. Li et al. collected marine images from the South China Sea using a camera mounted on an unmanned boat and obtained a dataset containing 4000 images after data augmentation. A fusion model based on YOLOv3 and DenseNet was used to detect sea surface vessel targets [35]. Zhang et al. used a dataset extracted from three days of video monitoring of a certain location on the Beijing North Canal to detect plastic bags and bottles floating on the water's surface. The detection model used was Faster R-CNN with VGG16 as the feature extractor, in which conv4_3 and conv5_3 were selected from VGG16 for feature fusion, to improve the detection accuracy of small objects [49]. In 2020, van et al. constructed a plastic floating garbage monitoring dataset using videos collected from five different waterway bridges in Jakarta, Indonesia, and identified plastic floating objects through two rounds of detection. Faster R-CNN was used to detect regions that may contain plastic floating objects, and then Inception V2 based on pre-trained COCO was used to determine whether these regions contain plastic [33].

References

1. Jambeck, J.R.; Geyer, R.; Wilcox, C.; Siegler, T.R.; Perryman, M.; Andrady, A.; Narayan, R.; Law, K.L. Plastic waste inputs from land into the ocean. Science 2015, 347, 768–771.

2. Suwal, N.; Kuriqi, A.; Huang, X.; Delgado, J.; Młyński, D.; Walega, A. Environmental flows assessment in Nepal: The case of Kaligandaki River. Sustainability 2020, 12, 8766.

3. Zhang, L.; Xie, Z.; Xu, M.; Zhang, Y.; Wang, G. EYOLOv3: An Efficient Real-Time Detection Model for Floating Object on River. Appl. Sci. 2023, 13, 2303.

4. Qiu, H.; Cao, S.; Xu, R. Cancer incidence, mortality, and burden in China: A time-trend analysis and comparison with the United States and United Kingdom based on the global epidemiological data released in 2020. Cancer Commun. 2021, 41, 1037–1048.

5. Zhu, J.; Yang, G.; Feng, X.; Li, X.; Fang, H.; Zhang, J.; Bai, X.; Tao, M.; He, Y. Detecting wheat heads from UAV low-altitude remote sensing images using deep learning based on transformer. Remote Sens. 2022, 14, 5141.

6. Liu, C.; Yang, D.; Tang, L.; Zhou, X.; Deng, Y. A Lightweight Object Detector Based on Spatial-Coordinate Self-Attention for UAV Aerial Images. Remote Sens. 2022, 15, 83.

7. Xia, X.; Meng, Z.; Han, X.; Li, H.; Tsukiji, T.; Xu, R.; Zheng, Z.; Ma, J. An automated driving systems data acquisition and analytics platform. Transport. Res. Part C-Emerg. Technol. 2023, 151, 104120.

8. Liu, W.; Quijano, K.; Crawford, M.M. YOLOv5-Tassel: Detecting tassels in RGB UAV imagery with improved YOLOv5 based on transfer learning. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2022, 15, 8085–8094.

9. Xue, M.; Chen, M.; Peng, D.; Guo, Y.; Chen, H. One spatio-temporal sharpening attention mechanism for light-weight YOLO models based on sharpening spatial attention. Sensors 2021, 21, 7949.

10. Liu, Y.; Yang, F.; Hu, P. Small-Object Detection in UAV-Captured Images via Multi-branch Parallel Feature Pyramid Networks. IEEE Access 2020, 8, 145740–145750.

11. Demiray, B.Z.; Sit, M.; Demir, I. D-SRGAN: DEM super-resolution with generative adversarial networks. SN Comput. Sci. 2021, 2, 48.

12. Chen, Y.; Li, J.; Niu, Y.; He, J. Small Object Detection Networks Based on Classification-Oriented Super-Resolution GAN for UAV Aerial Imagery. In Proceedings of the 2019 Chinese Control And Decision Conference (CCDC), Nanchang, China, 3–5 June 2019.

13. Jiang, W.; Ren, Y.; Liu, Y.; Leng, J. Artificial neural networks and deep learning techniques applied to radar target detection: A review. Electronics 2022, 11, 156.

14. Ju, M.; Luo, J.; Zhang, P.; He, M.; Luo, H. A simple and efficient network for small target detection. IEEE Access 2019, 7, 85771–85781.

15. Lang, K.; Yang, M.; Wang, H.; Wang, H.; Wang, Z.; Zhang, J.; Shen, H. Improved One-Stage Detectors with Neck Attention Block for Object Detection in Remote Sensing. Remote Sens. 2022, 14, 5805.

16. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436–444.

17. Benghanem, M.; Mellit, A.; Moussaoui, C. Embedded Hybrid Model (CNN–ML) for Fault Diagnosis of Photovoltaic Modules Using Thermographic Images. Sustainability 2023, 15, 7811.

18. Wang, R.; Jiao, L.; Xie, C.; Chen, P.; Du, J.; Li, R. S-RPN: Sampling-balanced region proposal network for small crop pest detection. Comput. Electron. Agric. 2021, 187, 106290.

19. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.

20. Ding, J.; Zhang, J.; Zhan, Z.; Tang, X.; Wang, X. A Precision Efficient Method for Collapsed Building Detection in Post-Earthquake UAV Images Based on the Improved NMS Algorithm and Faster R-CNN. Remote Sens. 2022, 14, 663.

21. Pan, Y.; Zhu, N.; Ding, L.; Li, X.; Goh, H.-H.; Han, C.; Zhang, M. Identification and Counting of Sugarcane Seedlings in the Field Using Improved Faster R-CNN. Remote Sens. 2022, 14, 5846.

22. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. Ssd: Single shot multibox detector. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; pp. 21–37.

23. Joochim, O.; Satharanond, K.; Kumkun, W. Development of Intelligent Drone for Cassava Farming. In Recent Advances in Manufacturing Engineering and Processes: Proceedings of ICMEP 2022; Springer: Singapore, 2023; pp. 37–45.

24. Tian, Z.; Shen, C.; Chen, H.; He, T. Fcos: Fully convolutional one-stage object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, South Korea, 27 October–2 November 2019; pp. 9627–9636.

25. Xie, X.; Cheng, G.; Wang, J.; Yao, X.; Han, J. Oriented R-CNN for object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 3520–3529.

26. Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.

27. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Proceedings of the Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, Montreal, QC, Canada, 7–12 December 2015.

28. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.

29. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.

30. Ma, H.; Liu, Y.; Ren, Y.; Yu, J. Detection of collapsed buildings in post-earthquake remote sensing images based on the improved YOLOv3. Remote Sens. 2019, 12, 44.

31. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 448–456.

32. Kim, D.-H. Evaluation of COCO validation 2017 dataset with YOLOv3. Evaluation 2019, 6, 10356–10360.

33. van Lieshout, C.; van Oeveren, K.; van Emmerik, T.; Postma, E. Automated river plastic monitoring using deep learning and cameras. Earth Space Sci. 2020, 7, e2019EA000960.

34. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.

35. Li, Y.; Guo, J.; Guo, X.; Liu, K.; Zhao, W.; Luo, Y.; Wang, Z. A novel target detection method of the unmanned surface vehicle under all-weather conditions with an improved YOLOV3. Sensors 2020, 20, 4885.

36. Zhang, Z. Improved adam optimizer for deep neural networks. In Proceedings of the 2018 IEEE/ACM 26th International Symposium on Quality of SERVICE (IWQoS), Banff, AB, Canada, 4–6 June 2018; pp. 1–2.

37. Stofa, M.M.; Zulkifley, M.A.; Zaki, S.Z.M. A deep learning approach to ship detection using satellite imagery. Proc. IOP Conf. Ser. Earth Environ. Sci. 2020, 540, 012049.

38. Li, R.; Ji, Z.; Hu, S.; Huang, X.; Yang, J.; Li, W. Tomato Maturity Recognition Model Based on Improved YOLOv5 in Greenhouse. Agronomy 2023, 13, 603.

39. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y.M. Scaled-yolov4: Scaling cross stage partial network. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 13029–13038.

40. Kim, M.; Kim, H.; Sung, J.; Park, C.; Paik, J. High-resolution processing and sigmoid fusion modules for efficient detection of small objects in an embedded system. Sci. Rep. 2023, 13, 244.

41. Wang, S.-Y.; Qu, Z.; Li, C.-J.; Gao, L.-Y. BANet: Small and multi-object detection with a bidirectional attention network for traffic scenes. Eng. Appl. Artif. Intel. 2023, 117, 105504.

42. Yang, C.; Huang, Z.; Wang, N. Querydet: Cascaded sparse query for accelerating high-resolution small object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 13668–13677.

43. Liu, Q.-p.; Wang, Q.-j.; Hanajima, N.; Su, B. An improved method for small target recognition based on faster RCNN. In Proceedings of the 2021 Chinese Intelligent Systems Conference: Volume II, Fuzhou, China, 16–17 October 2021; pp. 305–313.

44. Imai, N.; Otokawa, H.; Okamoto, A.; Yamazaki, K.; Tamura, T.; Sakagami, T.; Ishizaka, S.; Shimojima, H. Abandonment of Cropland and Seminatural Grassland in a Mountainous Traditional Agricultural Landscape in Japan. Sustainability 2023, 15, 7742.

45. Lan, Y.; Lin, S.; Du, H.; Guo, Y.; Deng, X. Real-Time UAV Patrol Technology in Orchard Based on the Swin-T YOLOX Lightweight Model. Remote Sens. 2022, 14, 5806.

46. Bajić, M., Jr.; Potočnik, B. UAV Thermal Imaging for Unexploded Ordnance Detection by Using Deep Learning. Remote Sens. 2023, 15, 967.

47. Liang, H.; Seo, S. UAV Low-Altitude Remote Sensing Inspection System Using a Small Target Detection Network for Helmet Wear Detection. Remote Sens. 2023, 15, 196.

48. Wang, Y.; Ma, L.; Wang, Q.; Wang, N.; Wang, D.; Wang, X.; Zheng, Q.; Hou, X.; Ouyang, G. A Lightweight and High-Accuracy Deep Learning Method for Grassland Grazing Livestock Detection Using UAV Imagery. Remote Sens. 2023, 15, 1593.

49. Zhang, L.; Zhang, Y.; Zhang, Z.; Shen, J.; Wang, H. Real-time water surface object detection based on improved faster R-CNN. Sensors 2019, 19, 3523.