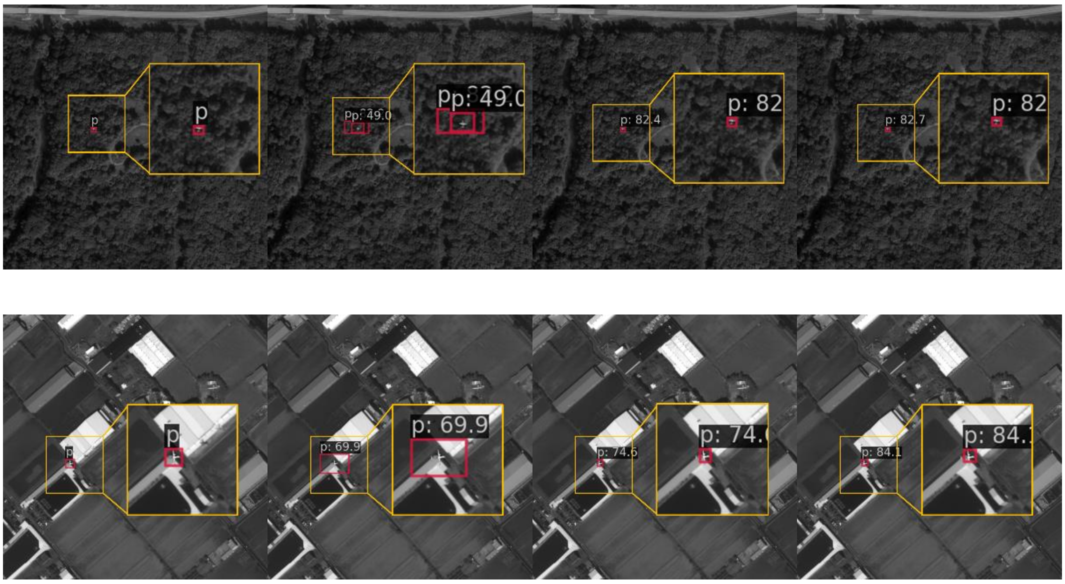
More Experiments

Aircraft detection is an important and basic research in remote sensing applications. However, the existing dataset only include the aircraft parked at the airport, so we need to create a new dataset to ensure that the new dataset can more truly reflect the actual situation. Based on AI-TOD, we added some aircraft target data for flight status. Fig. 1 shows the results of the object inspection in flight. We did not include these results in the manuscript due to space constraints.



(a) (b) (c) (d)

Fig. 1. Example results on the test set. (a) Ground truth. (b) NWD. (c) YOLOv8. (d) HFF-YOLO.

To further test the performance of the proposed method, we use AI-TOD dataset for further testing. The method proposed in this paper is compared with several mainstream methods on AI-TOD, and the results are shown in Table I. The method proposed in this paper is compared with several mainstream methods on AI-TOD. When training the model, Adam [30] was used as the optimizer, with a batch size of 2 and a total of 100 epochs. Train using an E5-2630 v3 CPU and a single RTX-4080 GPU. It should be noted that due to page limitations, we were unable to include the results and related analyses in the manuscript.

TABLE I

COMPARISON OF DETECTION RESULTS ON AI-TOD

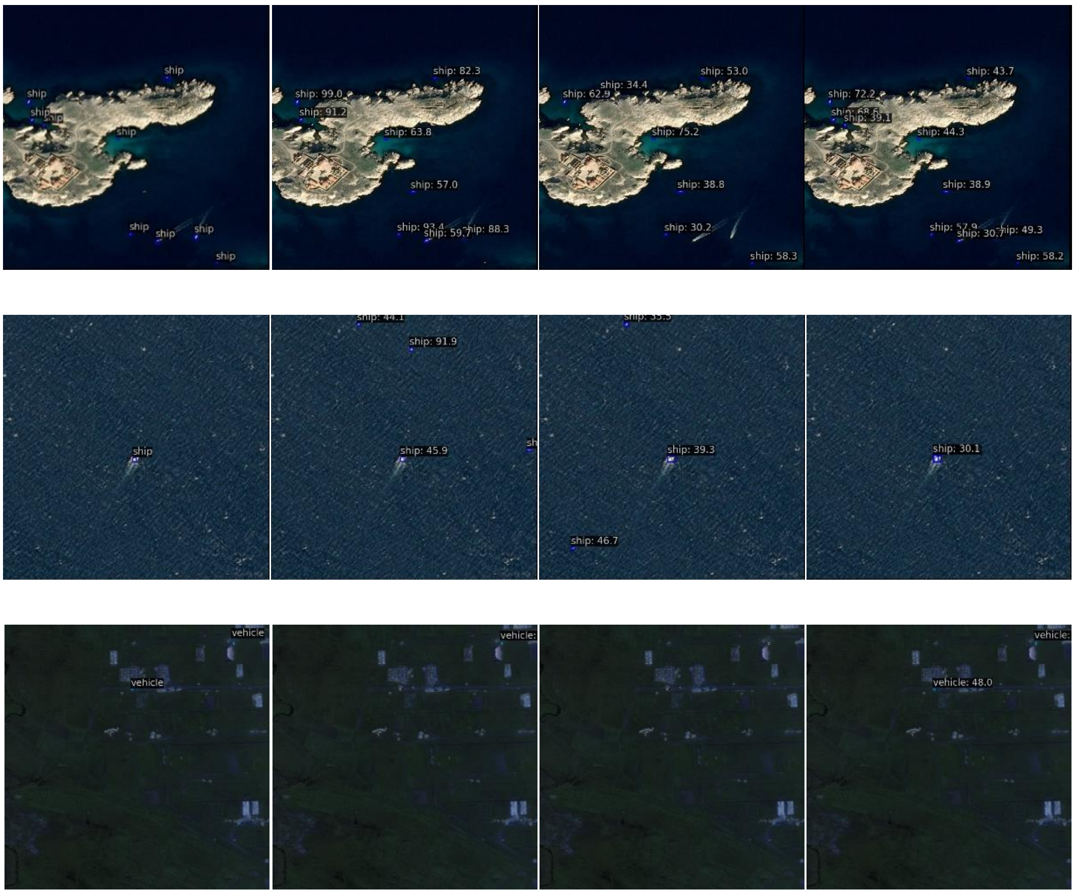
Network	Fusion method	AR@0.5:0.95	AP@0.5	AP@0.5:0.95	AP_{VS}@0.5:0.95	AP_S@0.5:0.95	AP_M@0.5:0.95
Faster_rcnn [1]	FPN	11.7%	17.2%	7.0%	5.9%	10.8%	14.8%
Cascade R-CNN [2]	FPN	11.4%	20.4%	6.5%	4.8%	20.4%	30.2%
Co-DETR [3]	FPN	4.4%	2.5%	0.6%	0.9%	0.4%	4.2%
Retinanet [4]	FPN	9.9%	13.2%	5.4%	4.4%	9.4%	12.6%
CFINet [5]	FPN	11.2%	20.2%	8.9%	4.3%	20.3%	30.2%
DetectoRS [6]	RFP	12.1%	20.8%	9%	4.6%	20.4%	31.3%

NWD [7]	RFP	20.6%	30.6%	13.0%	11.4%	19.0%	24.1%
PANet [8]	PAN	12.9%	19.0%	7.7%	6.8%	11.5%	16.0%
YOLOv5 [9]	PAN	18.5%	27.5%	10.8%	9.8%	16.7%	24.6%
YOLOv8 [9]	PAN	20.6%	27.4%	12.1%	10.6%	16.8%	24.1%
HFF-YOLO	PAN+HFFS	**24.7%**	**35.7%**	**15.4%**	**13.5%**	**21.3%**	**31.9%**

The results of the comparative experiment on AI-TOD dataset are shown in Table I. The experimental results show that compared with the baseline model, the model detection performance improved after adding the HFFS module, and it also has competitive performance compared to current mainstream models. Compared to YOLOv8, improves the optimal results by 3.1% in AR, 8.3% in AP@0.5, and 3.3% in AP@0.5:0.95.

The detection performance of HFF-YOLO on very small objects is better than that of the baseline model, improving by 2.9%. Compared with the mainstream model, HFF-YOLO not only fuses automatically extracted features of different scales, but also uses attention mechanism to introduce hand-crafted features into the feature fusion process, so that the model has the best detection ability for small objects.

In the experimental results, it is found that HFF-YOLO has fewer false alarms for objects with similar shapes. As shown in the second row of Fig. 2, HFF-YOLO can better distinguish ships from similar objects such as waves and islands. HFF-YOLO also has fewer misses for very small objects, as shown in the third row of Fig. 2, where the object size is about 5×5 pixels.



(a) (b) (c) (d)

Fig. 2. Example results on the AI-TOD. (a) Ground truth. (b) NWD. (c) YOLOv8. (d) HFF-YOLO.

[1] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[2] Z. Cai and N. Vasconcelos, “Cascade R-CNN: Delving Into High Quality Object Detection,” in *Proc. CVPR*, Salt Lake City, UT: IEEE, Jun. 2018, pp. 6154–6162. doi: 10.1109/CVPR.2018.00644.

[3] Z. Zong, G. Song, and Y. Liu, “DETRs with Collaborative Hybrid Assignments Training,” in *Proc. ICCV*, Paris, France: IEEE, Oct. 2023, pp. 6725–6735. doi: 10.1109/ICCV51070.2023.00621.

[4] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, “Focal Loss for Dense Object Detection,” in *Proc. ICCV*, IEEE, 2017, pp. 2999–3007.

[5] X. Yuan, G. Cheng, K. Yan, Q. Zeng, and J. Han, “Small Object Detection via Coarse-to-fine Proposal Generation and Imitation Learning,” in *Proc. ICCV*, Paris, France: IEEE, Oct. 2023, pp. 6294–6304. doi: 10.1109/ICCV51070.2023.00581.

[6] S. Qiao, L.-C. Chen, and A. Yuille, “DetectoRS: Detecting Objects With Recursive Feature Pyramid and Switchable Atrous Convolution,” in *Proc. CVPR*, 2021, pp. 10213–10224.

[7] J. Wang, C. Xu, W. Yang, and L. Yu, “A Normalized Gaussian Wasserstein Distance for Tiny Object Detection,” arXiv.org. Accessed: Feb. 29, 2024. [Online]. Available: https://arxiv.dosf.top/abs/2110.13389v2

[8] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, “Path Aggregation Network for Instance Segmentation,” in *Proc. CVPR*, 2018, pp. 8759–8768.

[9] G. Jocher, A. Chaurasia, and J. Qiu, “Ultralytics YOLO.” Jan. 2023. Accessed: Feb. 29, 2024. [Online]. Available: https://github.com/ultralytics/ultralytics