Evaluation and Empirical Study of Image Enhancement Strategies for Small Object Detection in Remote Sensing

Runxing Song *

School of Mechanical Engineering, Nanjing University of Science & Technology, Nanjing, Jiangsu, 210094, China

* Corresponding Author Email: 923101170136@njust.edu.cn

Abstract. Optical remote sensing small object detection is limited by the low proportion of target pixels, complex backgrounds, and image degradation. To avoid equating visual enhancement with detection benefits, this paper analyzes degradation perception in image enhancement, combining image degradation, enhancement processing, target area visibility, and detection feedback, and empirically conducts the analysis based on the Northwestern Polytechnical University Very-High-Resolution 10-class (NWPU VHR-10) dataset. This paper combines the Common Objects in Context (COCO) small target area threshold and the relative area of the remote sensing target frame, and proposes a screening rule that the target frame area does not exceed 1% of the image area. Then, 30 positive samples were simulated for downsampling, blurring, noise, and haze degradation, comparing bicubic interpolation, Contrast Limited Adaptive Histogram Equalization (CLAHE), sharpening, dehazing, and fusion enhancement. After further training the You Only Look Once version 8 nano (YOLOv8n) detector, the mAP50 values for the original, degraded, and enhanced images were 0.921, 0.851, and 0.826, respectively, and the recall values were 0.871, 0.789, and 0.753, respectively. The results show that edge texture enhancement cannot be directly equated with detection benefits; enhancement evaluation should focus on full-image fidelity, target area indicators, and authentic detection feedback.

Keywords: Optical remote sensing, small object detection, image enhancement, degradation perception

1. Introduction

Optical remote sensing image target detection is widely used in scenarios such as port supervision, airport management, urban traffic, and disaster emergency response [1]. Compared to natural images, remote sensing images feature top-down viewing, more extensive coverage, significant scale variations, and complex backgrounds. Targets such as aircraft, ships, vehicles, and oil tanks often occupy only a small number of pixels, with edges and textures easily overshadowed by the background, making small object detection more likely to be missed [2].

Previous studies have mainly started from improving detection network structures, such as multi-scale feature fusion, rotary box detection, and attention mechanisms [1-2]. These methods can alleviate issues of insufficient semantic information and unstable localization, but the image quality at the model input still limits the expression of small target features. Low resolution, blur, noise, thin clouds, and fogging compress the difference between the target and the background; weakening a small amount of edge information can affect detection judgment.

This paper discusses whether image enhancement under degradation conditions truly improves visibility in small target areas. To avoid equating visual clarity with detection benefits, this paper analyzes imaging degradation, enhanced processing, target area indicators, and detection feedback in the same logical chain and uses lightweight empirical evidence to illustrate possible trade-offs among different indicators. The purpose of this study is to provide a set of reproducible, enhanced evaluation ideas for remote sensing detection and image preprocessing research, helping readers distinguish between visual improvement, target area visibility improvement, and real detection gains.

2. Image Degradation Mechanism in Optical Remote Sensing Small Target Detection

Fig. 1 summarizes the degradation-aware evaluation idea adopted in this paper. This framework starts from low resolution, blurred noise, thin clouds and fog, and complex backgrounds, connects enhancement methods, full-image fidelity indicators, target area indicators, and real detection feedback, and is finally used to explain whether image enhancement is truly converted into detection benefits.

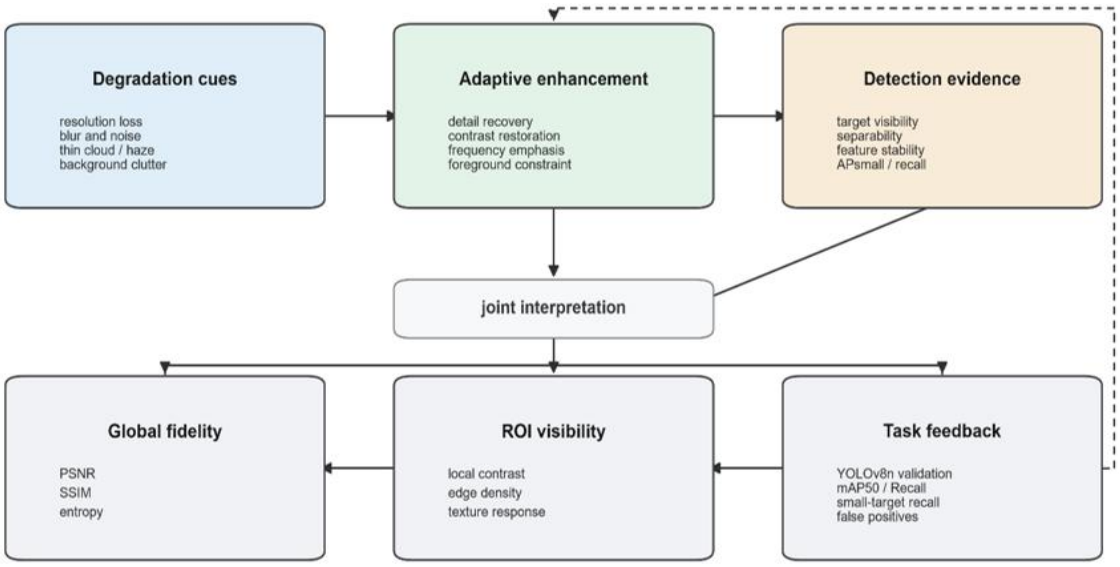


Fig 1. Degradation-Aware Enhancement and Detection-Oriented Evaluation Framework. (Picture credit: Original)

Insufficient spatial resolution is the most direct limitation of small object detection. When the target occupies only a few tens or even fewer pixels, interpolation, compression, and downsampling will further weaken the boundary information. After the depth detection model undergoes multi-level convolution and pooling, small target features are more likely to be covered by the background in deep feature maps. Therefore, although super-resolution and shallow detail enhancement are valuable, if the generated texture deviates from the real target structure, it may also cause false detections.

Blur, noise, and thin cloud fog can affect target visibility in different ways. Blur reduces gradient amplitude; noise may form false edges in the background, and thin clouds and haze cause low-frequency brightness shifts and compress local contrast [3, 4]. For small targets such as ships and vehicles, when water surface ripples, road textures, and building shadows are simultaneously enhanced, background interference will also increase.

In addition, the target categories, geographical areas, and imaging conditions in real open scenes change significantly, and airport detection research also shows that there is still a gap between the closed category settings in public datasets and real large-scale applications [5]. Therefore, image enhancement cannot only pursue improving the contrast of a single image but also needs to maintain spatial structure and background consistency.

3. Overview of Image Quality Enhancement Methods

3.1. Degradation-Driven Signal Repair and Multi-Dimensional Enhancement

Degradation-driven signal repair mainly targets input quality issues such as low resolution, blur, noise, and atmospheric scattering. Super resolution reconstruction aims to restore higher resolution expressions from low resolution images. Traditional bicubic interpolation is stable and fast, but it is difficult to restore true high-frequency details. Wang et al. constructed the Remote Sensing Super-Resolution Object Detection (RSSOD) data set, evaluated remote sensing super-resolution and target

detection under the same benchmark, and reported that their method improved 1.2 dB compared to the Peak Signal-to-Noise Ratio (PSNR) of Non-Local Sparse Network (NLSN), and the five-category detection mean Average Precision (mAP) reached 0.758 [6].

Atmospheric degradation recovery mainly deals with local contrast compression caused by thin clouds, haze, and light offset. Single-image methods often rely on dark channel priors, brightness correction, or neural networks to restore local contrast [7]. Multi-temporal and multi-source methods use other phases or Synthetic Aperture Radar (SAR) information to supplement the occluded area [4,8]. For small object detection, the focus of dehazing is not on obtaining visually more transparent images, but on restoring the boundaries of the target while maintaining background consistency.

Frequency domain enhancement focuses on the separation of high-frequency textures and low-frequency masks. Small target edges and corners usually correspond to higher frequencies, while clouds and light shifts correspond more to low-frequency components. Foreground feature enhancement emphasizes highlighting the target area first and then suppressing the background. Tian et al. combined image enhancement with ship detection and achieved maximum Average Precision (AP) improvements of 6.21% and 4.63% on optical and SAR ship datasets, respectively [9]. Recent prospect enhancement research also shows that goal constraints are necessary to enhance evaluation [10].

3.2. Joint Optimization of Task-Oriented Enhancement and Detection

Task-oriented joint optimization of enhancement and detection emphasizes the consistency of enhancement goals and detection goals. The traditional enhancement process usually processes the image first and then inputs it into the detector. The process is simple, but the enhancement goal and detection goal may not be consistent. For example, CLAHE can improve local contrast, but may amplify noise; A more reasonable direction is to introduce target box weights, foreground attention, or detection losses into the enhancement process, so that the enhancement module directly serves the reliability of detection.

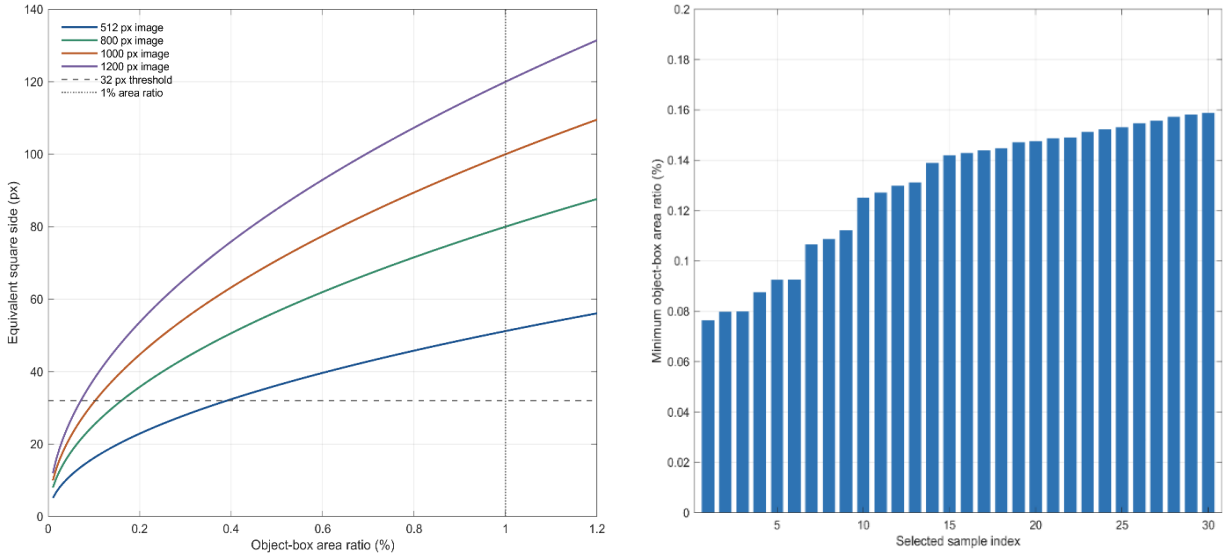
Therefore, this paper does not use visual clarity as the only basis for judgment in the empirical study, but instead compares full image fidelity, target area visibility, and lightweight detector output at the same time. The image index is used to explain the impact of degradation and enhancement on local structure, and the detection index is used to check whether this impact is converted into changes in mAP, Average Precision for small objects (APsmall), Recall, or false detection rate.

4. Datasets, Evaluation Metrics, and Lightweight Empirical Design

The public dataset provides a foundation for algorithm comparison and experimental review. Dataset for Object Detection in Optical Remote Sensing Images (DIOR) and RSSOD cover different image sizes, target categories, and annotation methods [4,10]. This paper uses positive samples from Northwestern Polytechnical University Very-High-Resolution 10-class (NWPU VHR-10) lightweight empirical testing.

There is no single absolute definition of small targets. Common Objects in Context (COCO) evaluation classifies objects with an area smaller than 32×32 pixels as small objects [11]. RSSOD research also pointed out that some remote sensing small targets account for less than 1% of the entire image area [6]. Based on these two types of basis, this article adopts operational screening rules: if the target frame area is less than 32×32 pixels, or the target frame area does not account for more than 1% of the image area, it can be regarded as a small target candidate. In the actual experiment, 30 samples were selected from small to large according to the relative area of the minimum target frame.

Fig. 2 shows the quantification rule and the sample distribution of this article. Fig. 2(a) illustrates that the equivalent side lengths corresponding to the same relative area are different under different image sizes, and the fixed pixel threshold cannot fully adapt to remote sensing images. Fig. 2(b) shows that the minimum target frame area ratio of the 30 selected samples is less than 1%, so it can support the discussion of small target enhancement.



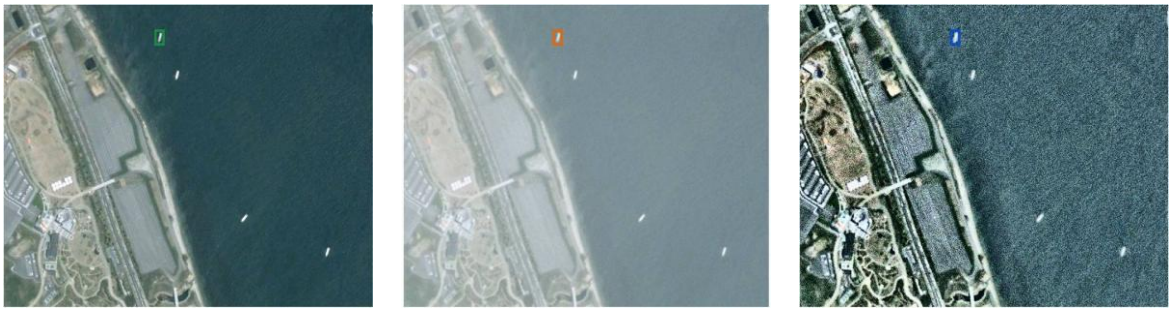
(a) (b)

Fig 2. Quantitative Screening Rule for Small Targets and Sample Distribution. Fig. 2(a) shows the relationship between target bounding-box area ratio and equivalent side length; Fig. 2(b) shows the distribution of the minimum target bounding-box area ratios for 30 samples. (Picture credit: Original).

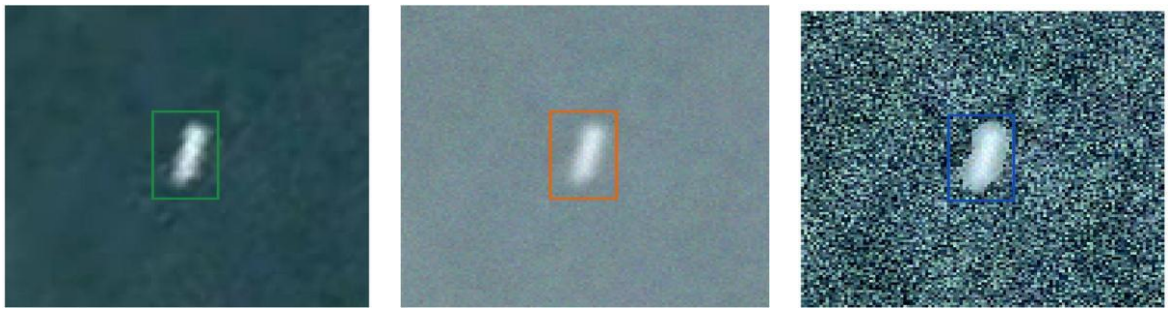
The experiment first simulates downsampling, blur, noise, and haze degradation, and then compares bicubic interpolation, Contrast Limited Adaptive Histogram Equalization (CLAHE), sharpening, dark channel approximate dehazing, and fusion enhancement [7]. Evaluation indicators include PSNR, Structural Similarity (SSIM), Region of Interest (ROI) local contrast, and ROI edge density. To supplement real detection feedback, this paper divides the 650 positive samples of NWPU VHR-10 into 454 training images, 97 verification images, and 99 test images, and trains 20 rounds of lightweight detectors based on You Only Look Once version 8 nano (YOLOv8n) implemented by Ultralytics [12]. During the test phase, three sets of inputs were constructed: original image, degraded image, and enhanced image, and compared mAP50, Recall, small target Recall, and the number of single image false detections at a confidence level of 0.25.

5. Result Description and Discussion

Fig. 3 shows the original image, the degradation image, the enhancement image, and the enlargement results of the target area of a ship sample. Fig. 3(a)-Fig. 3(c) corresponds to the original image, degraded image, and enhanced image from left to right; Fig. 3(d)-Fig. 3(f) showsw the enlargement of the corresponding target area. After degradation, the grayscale difference between the water surface and the ship decreases, and after enhancement, the contour of the ship becomes more prominent, but the texture of the water surface is also significantly enhanced.



(a) (b) (c)



(d) (e) (f)

Fig 3. Comparison of Sample Degradation and Enhancement.
Fig. 3(a)-Fig. 3(c) shows the original image from the NWPU VHR-10 dataset, the simulated degraded image, and the enhanced image; Fig. 3(d)-Fig. 3(f) shows the corresponding enlarged target regions. Source image: NWPU VHR-10 dataset [13]; degradation simulation, enhancement processing, region cropping, and visualization (Original).

Fig. 4 further compares the indicator changes of various enhancement strategies. Fig. 4(a) shows that the SSIM of bicubic interpolation is higher, while the ROI edge density of fusion enhancement is higher. Fig. 4(b) shows the relative changes based on the degraded image. Fusion enhancement increases the ROI edge density by about 15.6% but reduces SSIM by about 53.8% and ROI contrast by about 50.3%.

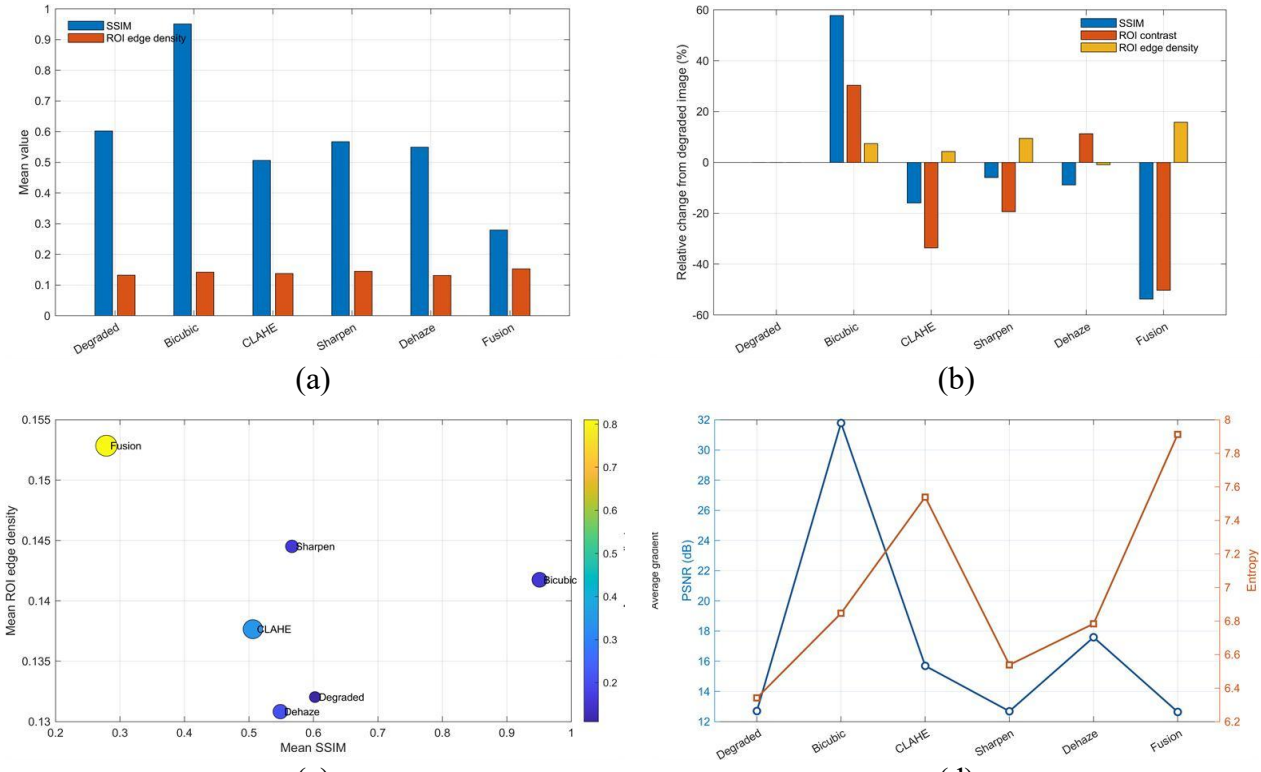


Fig. 4. Multi-Method Metric Comparison and Trade-Off Relationships.
Fig. 4(a) shows the mean SSIM and ROI edge density; Fig. 4(b) shows the relative changes; Fig. 4(c) shows the trade-off between SSIM and ROI edge density; Fig. 4(d) shows the changes in PSNR and information entropy. (Picture credit: Original)

Table 1 lists the average results of the six treatments. Bicubic interpolation is closer to the original clear image, and fusion enhancement makes it easier to strengthen edge textures, but at the same time sacrifices structural similarity. This result shows that, only from the perspective of image indicators, fusion enhancement is more biased towards local edge response rather than full image structure recovery.

Table 1. Average Metric Comparison of Different Enhancement Strategies

Method	PSNR/dB	SSIM	ROI Contrast	ROI Edge Density	Information entropy	Average gradient
degraded image	12.699	0.6024	0.9412	0.1320	6.3412	0.1099
bicubic interpolation	31.787	0.9507	1.2264	0.1417	6.8463	0.1570
CLAHE	15.693	0.5061	0.6243	0.1377	7.5393	0.3427
sharpening	12.684	0.5668	0.7589	0.1445	6.5385	0.1613
Dehazing	17.586	0.5488	1.0473	0.1308	6.7837	0.2101
Fusion enhancement	12.639	0.2789	0.4681	0.1527	7.9134	0.8101

Fig. 5 shows the real detector verification results. For the original images, mAP50, Recall, and small-target Recall were 0.921, 0.871, and 0.898, respectively; for the degraded images, they were 0.851, 0.789, and 0.797; for the enhanced images, they were 0.826, 0.753, and 0.780. The number of single-image false detections for the enhanced image increased from 0.90 for the degraded image to 1.28, indicating that texture enhancement caused by sharpening and local equalization may increase background interference.

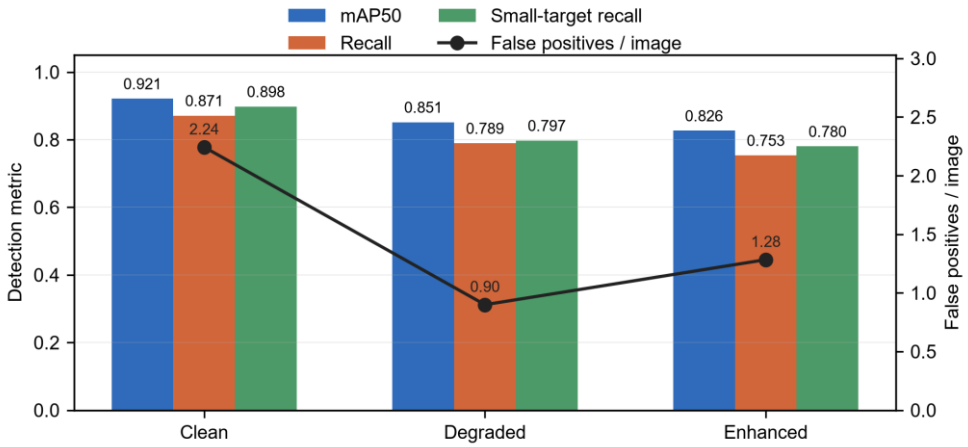


Fig. 5. YOLOv8n Detection Performance Under Clean, Degraded, and Enhanced Test Inputs. The detector was trained on 454 NWPU VHR-10 positive samples and tested on 99 held-out images. (Picture credit: Original).

Fig. 6 shows the detection frame visualization of the same test image under three types of inputs. The degradation map can still detect the main aircraft target, but the confidence level is partially reduced; the local texture in the enhanced image is clearer, and the confidence of individual targets has rebounded, but the overall Recall and mAP50 have not improved. This is consistent with the indicator trade-off in Fig. 4, that is, increased ROI edge density does not necessarily lead to improved detection accuracy.



(a) (b) (c)

Fig. 6. Qualitative YOLOv8n Detection Examples for the Same NWPU VHR-10 Test Image. Fig. 6(a)-Fig. 6(c) corresponds to clean, degraded, and enhanced inputs. Source image: NWPU VHR-10 dataset [13]; degradation simulation, enhancement processing, and detection visualization (Original).

Taken together, enhancement for remote sensing small target detection cannot just follow the visual standards of natural image enhancement. If the enhancement method simultaneously amplifies the target edge and background texture, the detector may obtain a stronger edge response and may also face more false edge interference. The YOLOv8n results of this article show that the visual sharpening after image enhancement does not translate into mAP50 and Recall improvements, so degradation perception, target area constraints, and real detection feedback should jointly enter the evaluation framework.

6. Limitations and Future Research Directions

This article still has limitations. Firstly, the experimental samples mainly come from NWPU VHR-10, and only 30 positive samples were selected for the image index experiment. The training and testing of the detector did not cover more sensors, geographical areas, or target categories. In the future, public data sets such as Dataset for Object Detection in Aerial Images (DOTA) and DIOR should be introduced in addition to NWPU VHR-10 to compare the stability of different categories, different target scales, and different imaging conditions [14,15]. Second, this article only uses a single lightweight detector, YOLOv8n, for preliminary closed-loop verification and has not yet compared real-time Transformer detectors such as rotating frame detectors, multi-scale small target detectors, and Real-Time Detection Transformer (RT-DETR) [2,16]. Therefore, the results of this article can show that the current enhancement strategy does not bring stable detection gains, but it cannot yet represent the response of all detectors to enhanced images. Third, the degradation simulation only covers downsampling, blur, noise, and haze, and has not yet dealt with complex situations such as thick cloud occlusion, multi-temporal phase loss, and SAR-optical fusion recovery. Multi-source fusion methods also suggest that SAR and adjacent optical images can supplement occlusion area information but may introduce registration errors and cross-modal differences [8]. Future work should verify the enhancement strategy under larger samples, more realistic degradation, and multi-detector feedback constraints, rather than judging its effectiveness solely based on visual clarity or a single image quality metric.

7. Conclusion

This paper focuses on the issue of image quality enhancement in optical remote sensing small target detection, forms an analysis framework for degradation perception and task orientation, and provides reviewable small target screening rules based on the relative area of the target frame. The minimum target frame area ratio of the 30 NWPU VHR-10 samples is less than 1%, and the average value is about 0.1299%, which represents the small target scene that this article focuses on. Image indicator empirical results show that after fusion enhancement, the PSNR changes from 12.699 dB to 12.629 dB, the SSIM decreases from 0.6024 to 0.2783, the ROI local contrast decreases from 0.9412 to 0.4682, and the ROI edge density increases from 0.1320 to 0.1527, an increase of approximately 15.6%. Multi-method comparison further shows that the PSNR and SSIM of bicubic interpolation reach 31.787 dB and 0.9507, respectively, which are closer to the original image structure; The information entropy and average gradient of fusion enhancement reach 7.9134 and 0.8101, respectively, which are more inclined to strengthen texture and edges. After further introducing the YOLOv8n detector, the mAP50 of the original image, degraded image, and enhanced image tests are 0.921, 0.851, and 0.826, respectively; the Recall is 0.871, 0.789, and 0.753, respectively, and the small target Recall is 0.898, 0.797, and 0.780, respectively. This shows that although the current fusion enhancement improves the ROI edge density, it does not improve the real detection results. From this, it can be seen that image enhancement for small object detection cannot only pursue visual sharpening or a single full image metric, but should also consider full image fidelity, target area visibility, and detector feedback simultaneously. The quantitative results of this article also show that

the same enhancement strategy may improve on edge indicators and degrade on structural indicators and detection indicators, so it needs to be interpreted according to task goals.

References

[1] Gui S, Song S, Qin R, Tang Y. Remote sensing object detection in the deep learning era-a review. Remote Sensing, 2024, 16(2): 327.

[2] Hua W, Chen Q. A survey of small object detection based on deep learning in aerial images. Artificial Intelligence Review, 2025, 58(6): 162.

[3] Liu J, Pan B, Shi Z. CR-Famba: A frequency-domain assisted Mamba for thin cloud removal in optical remote sensing imagery. IEEE Transactions on Multimedia, 2025, 27: 5659-5668.

[4] Ning J, Xie L, Yin J, Liu Y. Cloud removal advances: a comprehensive review and analysis for optical remote sensing images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2025, 18: 15914-15930.

[5] Li N, Cheng L, Ji C, Chen H, Geng W, Yang W. Airport detection in remote sensing real-open world using deep learning. Engineering Applications of Artificial Intelligence, 2023, 122: 106083.

[6] Wang Y, Bashir S M A, Khan M, Ullah Q, Wang R, Song Y, Guo Z, Niu Y. Remote sensing image super-resolution and object detection: benchmark and state of the art. Expert Systems with Applications, 2022, 197: 116793.

[7] He K, Sun J, Tang X. Single image haze removal using dark channel prior. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12): 2341-2353.

[8] Wang Y, Cheng J. A novel cloud removal method by fusing features from SAR and neighboring optical remote sensing images. PLOS ONE, 2025, 20(9): e0332898.

[9] Tian L, Cao Y, He B, Zhang Y, He C, Li D. Image enhancement driven by object characteristics and dense feature reuse network for ship target detection in remote sensing imagery. Remote Sensing, 2021, 13(7): 1327.

[10] Meng L, Li H, Han H, Xu M, Wu J, Hou S, Duan W. Progressive enhancement of foreground features for salient object detection in optical remote sensing images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2025, 18: 7572-7591.

[11] Lin T-Y, Maire M, Belongie S, Hays J, Perona P, Dollar P, Zitnick C L. Microsoft COCO: common objects in context. Computer Vision-ECCV 2014, Lecture Notes in Computer Science, 2014, 8693: 740-755.

[12] Jocher G, Qiu J, Chaurasia A. Ultralytics YOLO. GitHub repository, 2023. Available: https://github.com/ultralytics/ultralytics.

[13] Cheng G, Han J, Zhou P, Guo L. Multi-class geospatial object detection and geographic image classification based on collection of part detectors. ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 98: 119-132.

[14] Li K, Wan G, Cheng G, Meng L, Han J. Object detection in optical remote sensing images: a survey and a new benchmark. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 159: 296-307.

[15] Xia G-S, Bai X, Ding J, Zhu Z, Belongie S, Luo J, Datcu M, Pelillo M, Zhang L. DOTA: A large-scale dataset for object detection in aerial images. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 3974-3983.

[16] Zhao Y, Lv W, Xu S, Wei J, Wang G, Dang Q, Liu Y, Chen J. DETRs beat YOLOs on real-time object detection. arXiv:2304.08069v3, 2024.