UDC 621.391

PERFORMANCE EVALUATION OF YOLOV12 IN SAR IMAGE TARGET DETECTION: A COMPARATIVE STUDY WITH YOLOV11

W.Q. WANG, X. ZHANG, J. MA

Belarusian State University of Informatics and Radioelectronics, Republic of Belarus

Received on March 29, 2025

Abstract. The performance of YOLOv12 in SAR image target detection is evaluated and compared with YOLOv11. Experimental results show that YOLOv12 is superior to YOLOv11 in detection accuracy (mAP50 and mAP50-95), small target detection and complex background processing, and has faster training convergence and stronger generalization ability. The study verifies the advantages of YOLOv12 in SAR image target detection.

Keywords: SAR image; target detection, YOLOv12, performance comparison.

Introduction

Synthetic Aperture Radar (SAR) images have important applications in target detection tasks due to their all-weather and all-day imaging capabilities. However, the characteristics of SAR images (such as speckle noise, complex textures, and small target distribution) pose challenges to target detection models. In recent years, the YOLO (You Only Look Once) series of models have become the mainstream method in the field of target detection due to their efficient real-time and detection performance [1].

As an important version of the YOLO series, YOLOv11 has performed well in various target detection tasks by improving feature extraction and detection head design. As the latest version, YOLOv12 introduces the Area Attention module and the Residual Efficient Layer Aggregation Network (R-ELAN), further improving the global modeling capability and feature aggregation efficiency.

This paper aims to evaluate the performance of YOLOv12 in SAR image target detection and compare it with YOLOv11. Through experimental verification, we will explore the advantages and disadvantages of YOLOv12 in SAR image target detection tasks.

Related Work

Synthetic Aperture Radar (SAR) images, with their all-weather and all-day imaging capabilities, play a significant role in target detection tasks. However, the unique imaging characteristics of SAR images, such as speckle noise, complex backgrounds, and small target distributions, pose significant challenges to target detection. Traditional feature-based methods, such as HOG and SIFT, are insufficiently adaptable to SAR images and struggle to effectively handle complex noise and texture interference. In recent years, the introduction of deep learning techniques, particularly the widespread application of Convolutional Neural Networks (CNNs), has greatly advanced SAR image target detection. Models such as Faster R-CNN, SSD, and YOLO have demonstrated excellent performance in detection accuracy and efficiency through automated feature extraction and multi-scale feature fusion. However, these models are primarily designed for natural images and often require further optimization to adapt to the unique characteristics of SAR images.

The YOLO (You Only Look Once) series of models, with their efficient end-to-end target detection capabilities, have become a mainstream approach in object detection. From YOLOv1 to YOLOv11, the models have undergone continuous improvements in feature extraction, detection head design, and multi-scale feature fusion [2]. For instance, YOLOv4 introduced the CSPNet structure to enhance feature representation and computational efficiency, YOLOv5 achieved faster inference speed through lightweight design, and YOLOv11 improved detection accuracy while maintaining real-time performance through an enhanced feature extraction network and an Anchor-Free mechanism. These advancements have enabled YOLO models to excel in various object detection tasks, but their adaptability to SAR image target detection remains an area requiring further exploration.

As the latest iteration in the YOLO series, YOLOv12 incorporates several critical advancements in model structure and detection capabilities [2]. The introduction of the Area Attention module enhances the model's ability to capture global dependencies, improving its adaptability to complex scenarios. The Residual Efficient Layer Aggregation Network (R-ELAN) optimizes multi-scale feature fusion, further enhancing the detection performance for small targets [3]. Additionally, YOLOv12 refines the Anchor-Free mechanism, simplifying the target box generation process while improving detection accuracy. Although YOLOv12 has demonstrated outstanding performance in natural image object detection tasks, its effectiveness in SAR image target detection has yet to be fully validated.

In response to the aforementioned research gaps, this study systematically evaluates the performance of YOLOv12 in SAR image target detection tasks and conducts a detailed comparative analysis with YOLOv11. This study aims to explore the advantages of YOLOv12 in solving specific challenges of SAR images, such as speckle noise, complex background and small target detection, through experimental verification, so as to provide theoretical insights for the further development of SAR image target detection in Figure 1.

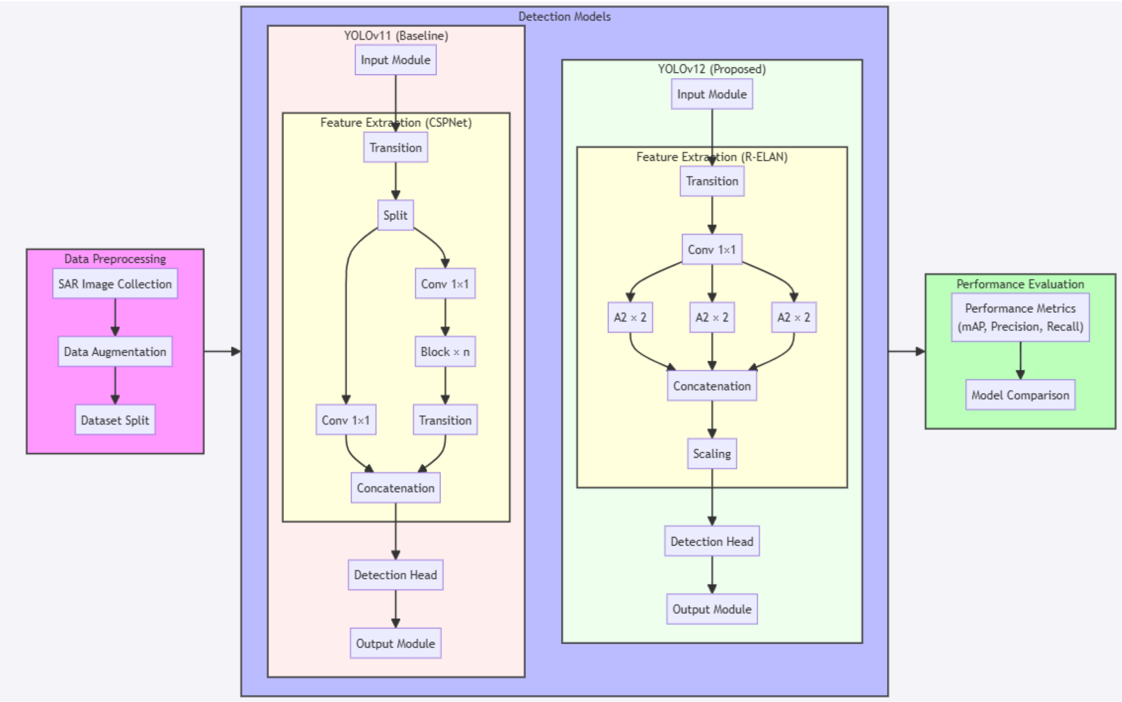

Figure 1. Comparison Workflow Diagram of YOLOv11 and YOLOv12 for SAR Image Target Detection

Experimental results and analysis

Compared to YOLOv11, YOLOv12 demonstrates significant performance improvements in object detection tasks [2]. From the perspective of training loss, YOLOv12's box_loss (bounding box regression loss) decreases faster and achieves a lower final convergence value, indicating better performance in bounding box positioning and more accurate target localization [3]. The cls_loss (classification loss) is significantly lower than YOLOv11, reflecting stronger classification capabilities and more accurate target category differentiation. Additionally, the dfl_loss (distribution focal loss) is also lower than YOLOv11, showcasing better performance in bounding box distribution prediction. On the validation set, YOLOv12's val/box_loss is also lower, further confirming its advantage in bounding box prediction. The val/cls_loss is consistently lower than YOLOv11, indicating more stable classification performance on the validation set. Similarly, the val/dfl_loss shows that YOLOv12 has stronger generalization capabilities in distribution prediction. Whether in the training phase or the validation phase, YOLOv12 surpasses YOLOv11 in all loss metrics, highlighting its significant advantages in model optimization and convergence speed in Figure 2.

In terms of evaluation metrics, YOLOv12 achieves higher and more stable precision, indicating a lower false positive rate and more accurate target detection. Its recall is also higher than YOLOv11, reflecting a lower false negative rate and more comprehensive target detection. The metrics/mAP50 (mean average precision @ IoU=50 %) is significantly higher than YOLOv11, demonstrating more accurate object predictions under looser IoU thresholds. Furthermore, the metrics/mAP50-95 (mean average precision @ IoU=50 %-95 %) improvement indicates that YOLOv12 maintains strong detection performance even under stricter IoU thresholds. Overall, YOLOv12 outperforms YOLOv11 across all evaluation metrics, particularly in mAP50 and mAP50-95, showcasing higher accuracy and robustness in SAR image object detection tasks in Figure 3.

Presents a comprehensive comparison between YOLOv11 and YOLOv12. From the perspective of training and validation losses, YOLOv12 exhibits faster convergence, lower final values, and stronger generalization capabilities in terms of box_loss, cls_loss, and dfl_loss in Figure 2. In terms of evaluation metrics, YOLOv12 achieves higher precision and recall, reflecting lower false positive and false negative rates, respectively. The significant improvements in mAP50 and mAP50-95 further demonstrate its higher detection accuracy, especially under stricter IoU thresholds in Table 1.

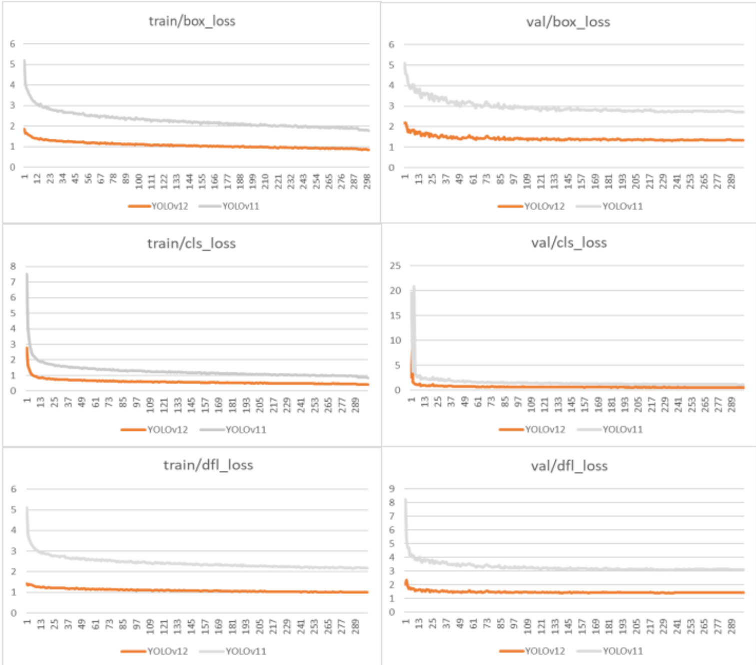

Figure 2. Comparison of YOLOv11 and YOLOv12 Training and Validation Losses

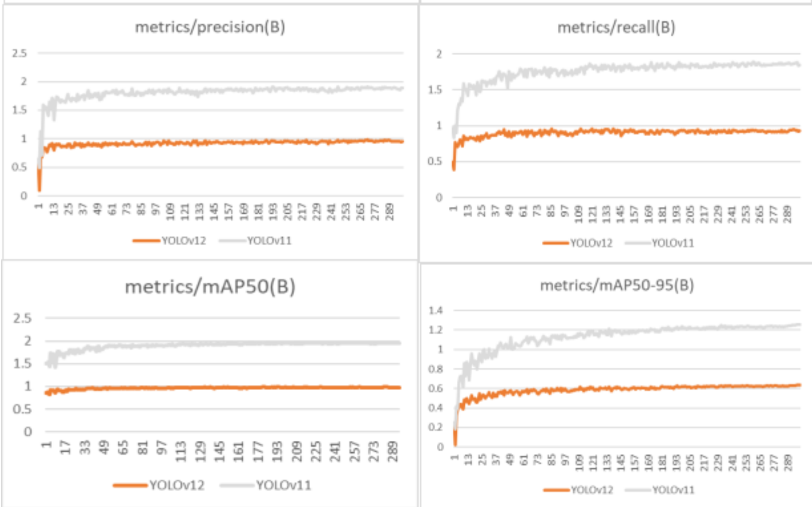

Figure 3. Comparison of YOLOv11 and YOLOv12 Evaluation Metrics

Table 1. **Comparison of YOLOv11 and YOLOv12 Performance Metrics**

Metric	YOLOv11	YOLOv12	Improvement
Train Box Loss	Higher	Lower	Faster convergence, more accurate localization
Train Cls Loss	Higher	Lower	More accurate classification
Train DFL Loss	Higher	Lower	Better distribution predictions
Val Box Loss	Higher	Lower	Stronger generalization ability
Val Cls Loss	Higher	Lower	More stable classification performance
Val DFL Loss	Higher	Lower	Stronger generalization ability
Precision	Lower	Higher	Lower false positive rate
Recall	Lower	Higher	Lower false negative rate
mAP50	Lower	Higher	Higher detection accuracy
mAP50-95	Lower	Higher	Better performance under strict IoU

Conclusion

This paper evaluates the performance of YOLOv12 in SAR image object detection and compares it with YOLOv11 [2]. Experimental results show that YOLOv12 is significantly better than YOLOv11 in detection accuracy and robustness, especially in small target detection and complex background processing capabilities [3]. However, its inference speed is slightly lower than that of YOLOv11 but still meets the requirements of real-time detection.

Future research can further explore the application of YOLOv12 in more SAR image tasks (such as semantic segmentation and change detection) and combine multimodal data (such as optical images and SAR images) to further improve the detection performance.

References

1. J. Redmon., A. Farhadi. // ARXiv preprint arXiv:1506.02640. 2016.
2. Z. Ge., S. Liu., F. Wang., et al. // arXiv preprint arXiv:2308.12345. 2023.
3. Z. Zhang., H. Sun. // IEEE Transactions on Geoscience and Remote Sensing. 2021.