

ABSTRACT

In maritime search and rescue (SAR) operations, the timely and accurate detection of small objects, such as boats and people in distress, is critical for effective response. Videos captured by cameras mounted on SAR non-governmental organization (NGO) ships provide a valuable source to support the crew by localizing distant objects through computer vision. This thesis investigates real-time object detection in challenging maritime environments using deep learning-based models, specifically focusing on architectures based on Convolutional Neural Networks and Transformers. Their performances are compared to find the most suitable model for the SAR-CAM system, a camera system for SAR missions. The models are fine-tuned on a maritime image dataset collected during SAR missions, provided by an NGO. The dataset presents unique challenges in complex maritime environments, such as small object sizes, varying weather conditions, waves, and sunlight reflections, which make reliable detection difficult. Various strategies to enhance small object detection performance are evaluated, including hyperparameter tuning, knowledge distillation, COCO class mapping, and architectural modifications. The generalization capabilities of the fine-tuned models are assessed on the benchmark Singapore Maritime Dataset. Results indicate that CNN-based models outperform transformer-based models in terms of average precision and average recall on small objects, while transformer-based models perform better on medium and large objects. YOLO11m emerges as the most accurate model in detecting small objects and meets all requirements of the SAR-CAM system, including real-time inference speed. Knowledge distillation using DINOv3 as a teacher model shows promising results for enhancing small object detection, although it significantly reduces large object detection. COCO class mapping proves to be a simple yet effective method to enhance small object detection for RT-DETR. The evaluation on the Singapore Maritime Dataset reveals that models fine-tuned on SAR-CAM generalize reasonably well to medium and large objects but face difficulties with small object detection, underscoring the challenges in adapting models to different maritime contexts.