ICCSNT2025

IU-DINO: Improved Detection Method for UAVs Based on DINO

Xuan Tong, Dan Shan, Boyi Duan
Han Li, Rongxi Zhang, Fanyi Kong

Introduction

Unmanned aerial vehicles (UAVs) intrusion detection i s challenged by the small scale, rapid manoeuvrability, an d frequent occlusion of targets in complex environments. Here we present IU-DINO, an enhanced detection frame work built upon DINO, designed to improve the accuracy and robustness of UAVs detection without increasing trai ning cost. IU-DINO incorporates a Reinforced Attention Feedback Module (RAFM), which recycles and enriches s patial features from earlier attention layers, and Unified Query Supervision Loss (UQSL), which jointly optimises one-to-one, one-to-many, and contrastive denoising branc hes. Evaluated on the Ac-UAVs dataset, IU-DINO achieve s an AP of 86.8, surpassing state-of-the-art DETR-based methods by up to 3.0 AP. The design offers consistent gai ns across multiple scales, demonstrating that targeted fea ture reuse and diversified supervision substantially impro ve small-object detection in UAVs surveillance scenarios.

Method

In this work, we present IU-DINO, an enhanced UAV d etection framework that builds upon the DINO architectu re. The key innovations include the Reinforced Attention Feedback Module (RAFM) and the Unified Query Super vision Loss (UQSL). The RAFM recycles low-level attenti on features through a lightweight convolutional pathway, preserving spatial details and improving detection accura cy with minimal computational overhead. The UQSL acce lerates convergence by integrating one-to-one, one-to-ma ny, and contrastive denoising supervision, enhancing rob ustness to occlusion and cluttered backgrounds. Together, these improvements enable IU-DINO to achieve superior performance on UAV detection tasks, particularly in com plex and dynamic environments.

Results

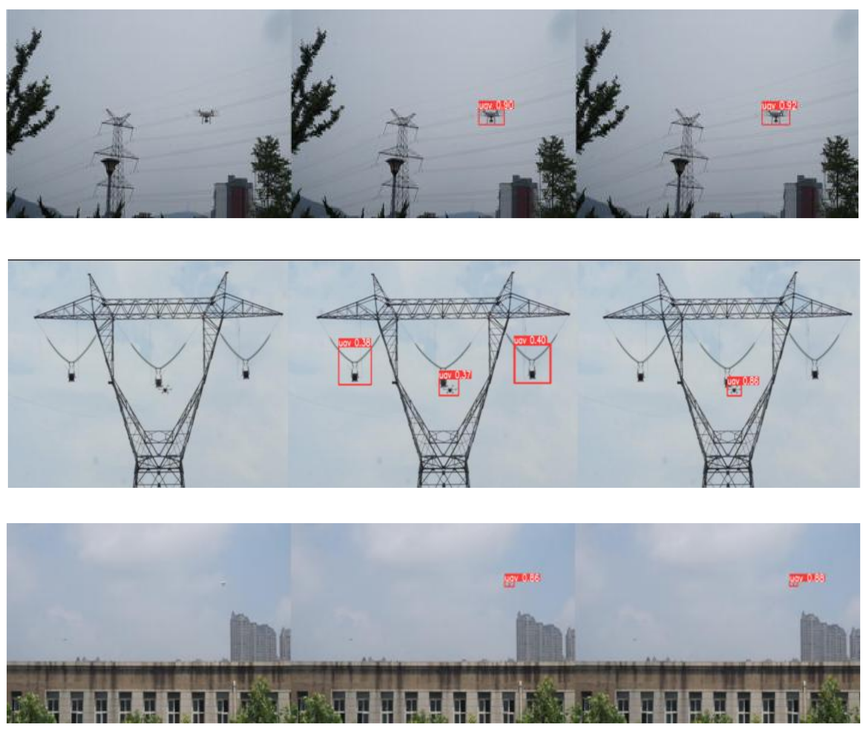

TABLE I
RESULTS OF ABLATION EXPERIMENT

RAFM	UQSL	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
		85.4	91.5	88.2	77.9	84.7	88.8
√		86.5	91.9	88.6	78.1	84.8	88.2
	√	85.7	91.6	87.4	76.9	84.2	87.3
√	√	86.4	92.0	88.3	77.8	84.9	88.9

Conclusion

We have introduced IU-DINO, a UAVs intrusion detectio n framework that advances DINO through the integration o f RAFM and a UQSL. RAFM enhances spatial detail preser vation by reusing refined low-level attention outputs, while the UQSL accelerates convergence and improves robustness through multi-branch, noise-aware supervision. Experimen tal results on the Ac-UAVs dataset confirm notable improve ments in detection accuracy, particularly for small and med ium targets, without additional computational overhead. Th ese findings highlight the value of feature reuse and diversif ied supervisory signals in transformer-based detectors and offer a promising direction for real-time, high-precision UA Vs surveillance in complex operational environments.