



RealDroneVision: Dataset and Architecture Advancements for Small-Object Drone Detection

Arun Kumar Sivapuram, Pranav R. T. Peddinti, Harish Puppala,
Komuravelli Prashanth, Jaladi Sri Harsha,
Rama Krishna Sai Gorthi

Indian Institute of Technology Tirupati
SRM University AP
DeepRaksha Lab

WACV 2026

Problem Statement and Challenges

Drone detection is critical in surveillance and autonomous systems.

- Extremely small targets ($< 0.1\%$ area)
- Motion blur and scale variation
- Distractors (bird vs drone confusion) [1]
- Real-world scenarios exhibit complexities that synthetic datasets often fail to capture [2]

Large benchmarks such as Anti-UAV [3] and UAVData [4] still struggle in tiny-object regimes.

Datasets

- DUT Anti-UAV [3]
- SynDroneVision [2]
- MAV-VID [5]
- Det-Fly [6]
- VisioDECT [7]

Detection Architectures

- TGC-YOLOv5 [8]
- SOD-YOLO [9]
- Improved YOLOv8 [10]
- YOLO-World [11]

Contributions

1. RealDroneVision Dataset

173,023 real-world images constructed using SA-LfV pipeline [12].

2. Nano Object Vision Attention (NOVA)

NOVA integrates depthwise separable convolutions, scale-aware dilated branches, lightweight mixing, and coordinate-aware attention, improving localization of small drones while maintaining computational efficiency..

3. Extensive Benchmarking

Compared against recent works [8, 9, 10].

Semi-Automatic Annotation Pipeline for Dataset Creation

Challenge: High-quality annotation of tiny drones is expensive and error-prone.

Our Solution: SA-LfV-Based Iterative Refinement [12]

1 Keyframe Annotation

Human annotates drone in selected frames.

2 Tracker Propagation

Labels propagated automatically across video frames.

3 Human Validation

False positives removed and tracking drift corrected.

4 False Negative Recovery

Refined detector recovers missed small or low-contrast targets.

5 Iterative Refinement

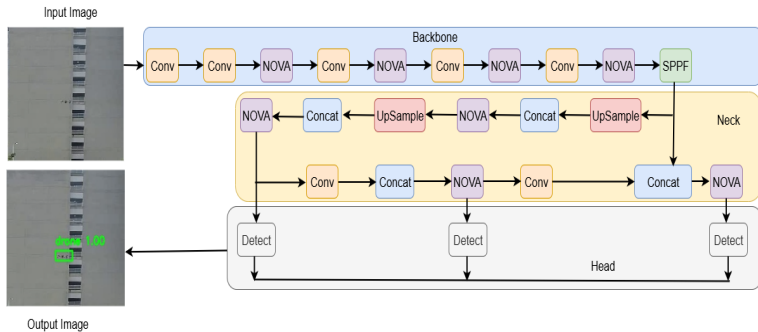
Detection and tracking mutually improve over cycles.

Significantly reduces annotation cost while preserving realism and accuracy.

RealDroneVision Dataset Samples



Proposed Detection Framework



Proposed NOVA Block

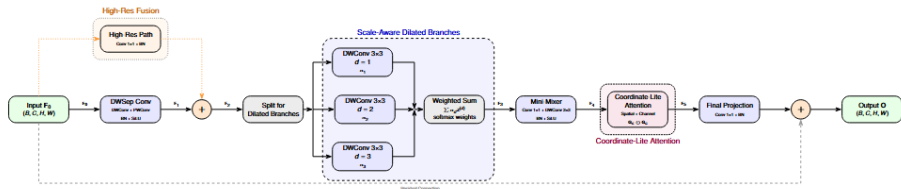


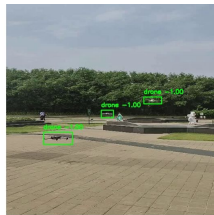
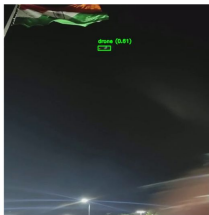
Figure 2. Architecture of the proposed NOVA block

Benchmark Results

Model	Params	GFLOPs	mAP@50	mAP@50–95
YOLOv8m	25.9M	79.1	0.857	0.518
YOLOv8l	43.6M	165.4	0.861	0.526
YOLOv9e	58.1M	192.7	0.863	0.522
YOLOv8-NOVA	2.3M	6.5	0.920	0.543

The proposed YOLOv8-NOVA is lightweight and achieved superior performance on the RealDroneVision dataset.

Qualitative Results






Dataset: [Google Drive Link](#)






Conclusion

- Large-scale real-world drone detection dataset
- Proposed a efficient semi-automatic annotation framework
- Proposed a lightweight small-object-focused NOVA module for improved detection




References I

-  A. Coluccia, A. Fascista, L. Sommer, A. Schumann, A. Dimou, and D. Zarpalas, “The drone-vs-bird detection grand challenge at icassp 2023: A review of methods and results,” *IEEE Open Journal of Signal Processing*, vol. 5, pp. 766–779, 2024.
-  T. R. Lenhard, A. Weinmann, K. Franke, and T. Koch, “Syndronevision: A synthetic dataset for image-based drone detection,” in *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2025.
-  N. Jiang, J. Zhao, Z. Zhang, D. Wang, Y. Sui, D. Li, and N. Yu, “Anti-uav: A large-scale benchmark for vision-based uav tracking,” *IEEE Transactions on Multimedia*, vol. 25, pp. 486–500, 2023.
-  Y. Zeng, Q. Duan, X. Chen, D. Peng, Y. Mao, and K. Yang, “Uavdata: A dataset for unmanned aerial vehicle detection,” *Soft Computing*, vol. 25, no. 7, pp. 5385–5393, 2021.

References II

-  A. Rodriguez-Ramos, J. Rodriguez-Vazquez, C. Sampedro, and P. Campoy, “Adaptive inattentional framework for video object detection with reward-conditional training,” *IEEE Access*, vol. 8, pp. 124 451–124 466, 2020.
-  Y. Zheng, Z. Chen, D. Lv, Z. Li, Z. Lan, and S. Zhao, “Air-to-air visual detection of micro-uavs: An experimental evaluation of deep learning,” *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1020–1027, 2021.
-  S. O. Ajakwe, V. U. Ihekoronye, G. Mohtasin, R. Akter, A. Aouto, D. S. Kim, and J. M. Lee, “Visiodect dataset: An aerial dataset for scenario-based multi-drone detection and identification,” *IEEE Dataport*, 2022.
-  Y. Zhao, Z. Ju, T. Sun, F. Dong, J. Li, R. Yang, Q. Fu, C. Lian, and P. Shan, “Tgc-yolov5: An enhanced yolov5 drone detection model based on transformer, gam & ca attention mechanism,” *Drones*, vol. 7, no. 7, p. 446, 2023.
-  Y. Xiao and N. Di, “Sod-yolo: A lightweight small object detection framework,” *Scientific Reports*, vol. 14, no. 1, p. 25624, 2024.

References III

-  H. Huang, B. Wang, J. Xiao, and T. Zhu, "Improved small-object detection using yolov8: A comparative study," *Applied and Computational Engineering*, vol. 41, pp. 80–88, 2024.
-  T. Cheng, L. Song, Y. Ge, W. Liu, X. Wang, and Y. Shan, "Yolo-world: Real-time open-vocabulary object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.
-  A. K. Sivapuram, P. Komuravelli, and R. K. S. S. Gorthi, "SA-LfV: Self-annotated labeling from videos for object detection," *Machine Learning*, vol. 114, no. 1, p. 21, 2025.

Thank You!
Questions?