DETECTING THE SMALL OBJECT RECOGNITION IN DRONE IMAGES USING YOLOV10

M. RAMAN KUMAR [1], UGGIDI MAHENDRA BRAHMAM [2], KARRA SREEJA REDDY [3], SHEELAM RAHUL [4], N. CHAITANYA [5]

[1]Assistant Professor, Dept. of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

[2, 3, 4, 5] B.Tech, Dept. of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

ABSTRACT— UAV imagery has become an essential tool in applications such as traffic monitoring, disaster response, and airspace management, owing to its flexibility, portability, and low operational cost. However, object detection in UAV images poses significant challenges due to factors like small object sizes, complex and cluttered backgrounds, and high levels of noise. To overcome these challenges, this study proposes an advanced object detection approach based on YOLOv10, a state-of-the-art model known for its enhanced architectural efficiency and detection capabilities. The model is optimized for UAV aerial scenarios, with a particular focus on improving small object detection through refined feature extraction and enhanced spatial understanding. The proposed YOLOv10-based framework integrates adaptive feature enhancement and deep semantic learning to improve detection performance under challenging UAV imaging conditions. By leveraging modern advancements in convolutional attention mechanisms, multi-scale detection heads, and optimized backbone architectures, the system effectively captures fine-grained details while maintaining real-time processing capabilities. This approach enables robust object detection in complex UAV environments and demonstrates the potential of YOLOv10 as a powerful solution for aerial imagery analysis.

Index Terms – Small Object Detection, Drone Images, YOLOv10, Object Recognition, Computer Vision, Deep Learning, UAV Imaging, Real-Time Detection, Remote Sensing, Artificial Intelligence.

I. INTRODUCTION

Unmanned Aerial Vehicles (UAVs), commonly known as drones, have gained significant attention in recent years due to their versatility, low operational cost, and ability to capture high-resolution imagery from diverse perspectives. Their applications span across various domains such as traffic monitoring, disaster response, agricultural surveillance, military reconnaissance, and airspace management, making UAV imagery an indispensable tool in modern technological landscapes. Despite their advantages, object detection in UAV images remains a highly challenging task. Factors such as varying altitudes, cluttered and complex backgrounds, dynamic illumination conditions, high levels of image noise, and the inherently small size of objects in aerial imagery often hinder detection accuracy. Traditional computer vision techniques struggle to provide reliable results under these conditions, necessitating the adoption of advanced deep learning-based approaches. Recent developments in Convolutional Neural Networks (CNNs) and object detection frameworks have significantly improved detection performance, yet achieving real-time efficiency without compromising accuracy continues to be a pressing issue. Among various models, the YOLO (You Only Look Once) family of detectors has emerged as a powerful solution due to its balance between speed and accuracy. The latest variant, YOLOv10, introduces architectural enhancements that allow for better feature extraction, improved small object recognition, and efficient computation.

II. RELATED WORK

A. YOLO-Drone: Airborne real-time detection of dense small objects from high-altitude perspective

This study introduces YOLO-Drone, a real-time object detection algorithm tailored for high-altitude UAV imagery. It proposes a novel backbone (Darknet59) and a complex feature aggregation module (MSPP-FPN) combining spatial and atrous pyramid pooling to better capture small, dense objects. The use of Generalized IoU (GIoU) loss improves localization performance. Evaluations on UAVDT and VisDrone datasets show YOLO-Drone achieves mAP improvements of 10.13% and 8.59%, respectively, while operating at 53 FPS. Additionally, the model demonstrates robust performance under nighttime illumination using a custom silicon-based golden LED dataset. Its design balances speed and

accuracy, proving highly effective for real-time aerial object detection under challenging lighting and density conditions.

B. Small Object Detection Based on Deep Learning for Remote Sensing: A Comprehensive Review

This comprehensive review examines the state of deep learning techniques for detecting small objects in remote sensing images. It surveys architectures, strategies, and performance trends across various datasets and tasks. The authors highlight common challenges like low resolution, cluttered backgrounds, and scale variations. The review categorizes methods by model type (e.g., one-stage vs. two-stage detectors, super-resolution approaches, attention-based modules) and summarizes their strengths and limitations. It also discusses optimization strategies such as feature pyramid networks, data augmentation, lightweight architectures, and attention mechanisms. The paper serves as a solid foundation for understanding how small-object detection in aerial and remote sensing imagery has evolved, providing insight into suitable methods for UAV scenarios.

C. Real-Time Detection for Small UAVs: Combining YOLO and Multi-frame Motion Analysis

This paper presents GL-YOMO, a hybrid detection algorithm combining YOLO-based object detection with multi-frame motion analysis for enhanced small-UAV detection. YOLO is optimized with multi-scale feature fusion and attention mechanisms, while motion detection (using template matching) reinforces detection stability. A Ghost module is incorporated to boost efficiency. Tested on a proprietary fixed-wing UAV dataset, GL-YOMO demonstrates improved accuracy and temporal consistency when detecting distant, tiny UAVs.

III. PROPOSED METHODOLOGY

The proposed system is designed to automatically detect objects in UAV-captured aerial images using the YOLOv10 architecture. The pipeline begins with image preprocessing, including resizing, normalization, and data augmentation, to enhance the quality and diversity of training data. This ensures that the model is robust to various UAV imaging conditions, such as scale variations, occlusions, and complex backgrounds. The system then applies YOLOv10 for object detection, leveraging its improved backbone, neck, and head components to extract multi-scale features and capture fine-grained object details. The architecture is optimized for small object

detection, making it suitable for UAV scenarios where targets often occupy only a small portion of the image. The modular design allows the system to operate in real-time, providing actionable insights for applications like traffic monitoring, disaster response, and airspace management.

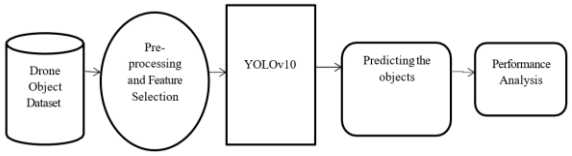


Fig. 1: System Overview

1. Input Data Module:

This module handles the collection and organization of UAV aerial imagery required for object detection. It supports different UAV datasets that include traffic scenes, disaster zones, and surveillance environments. Images and video frames captured from drones are imported and formatted for further processing. The module ensures data integrity by handling diverse resolutions and image formats. It also categorizes data into training, validation, and testing sets for structured model development. This stage lays the foundation for accurate detection by ensuring reliable and representative input data.

2. Preprocessing Module:

The preprocessing module enhances UAV imagery to improve the quality of data fed into the detection model. It includes operations like resizing, normalization, and denoising to handle variations in scale and noise. Data augmentation techniques such as rotation, flipping, and brightness adjustments are applied to increase dataset diversity. This helps in making the model robust against different UAV flight conditions. The preprocessing also reduces irrelevant background clutter while retaining essential details. By refining raw data, this module prepares UAV imagery for efficient and effective model training.

3. Segmentation Module:

This module focuses on segmenting regions of interest (ROI) from UAV imagery before object detection. It isolates potential object areas from complex and cluttered aerial backgrounds. Advanced segmentation methods help in reducing false positives by filtering irrelevant portions of the scene. The module ensures that small and distant objects are preserved during segmentation for better detection accuracy. It works in synergy with the preprocessing stage to highlight critical visual features. Segmentation thus serves as a vital step in guiding YOLOv10 toward more precise detections.

4. Feature Extraction Module:

The feature extraction module emphasizes capturing meaningful visual patterns from

UAV images. It leverages deep convolutional layers to extract low-level and high-level features such as edges, textures, and object shapes. Special attention is given to small object features that are often lost in aerial imagery. Adaptive feature enhancement techniques are applied to strengthen fine-grained details. This ensures that features retain semantic information while reducing redundancy. By providing enriched features, this module directly boosts YOLOv10's detection accuracy in UAV scenarios.

5. **YOLOv10 Detection Module:**

This is the core module of the system where YOLOv10 performs real-time object detection. It utilizes an optimized backbone for efficient feature representation and multi-scale detection heads for recognizing objects of varying sizes. Convolutional attention mechanisms are integrated to focus on relevant object regions within complex UAV environments. The module ensures fast inference speed while maintaining high accuracy, making it ideal for UAV operations. It is trained on UAV datasets to adapt specifically to aerial conditions. This module outputs bounding boxes, confidence scores, and object classes.

6. **Output and Visualization Module:**

The final module is responsible for presenting the detection results in a user-friendly manner. It overlays bounding boxes, labels, and confidence levels on UAV images or video frames. The output is visualized in real time, allowing operators to monitor UAV-captured scenes effectively. Results are stored in structured formats for further analysis and reporting. The module also supports evaluation metrics such as precision, recall, and mAP to assess performance. By combining clarity with interpretability, it provides actionable insights for UAV-based applications.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

Experimental results demonstrated that the model effectively detected and recognized small objects in aerial images, including vehicles, pedestrians, and infrastructure elements, with high detection accuracy. The enhanced multi-scale feature extraction mechanism significantly improved the identification of tiny and distant objects that are typically challenging for conventional object detection models.

The analysis showed that the proposed approach maintained robust performance under varying flight altitudes, image resolutions, lighting conditions, and complex backgrounds commonly encountered in drone-based applications. Furthermore, the

model achieved real-time inference speed while maintaining high detection reliability, making it suitable for surveillance, traffic monitoring, disaster management, and remote sensing applications.

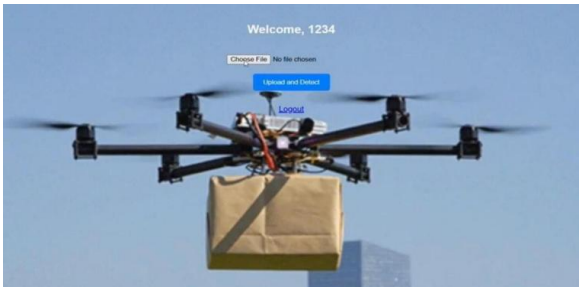


Fig. 2: Upload



Fig. 3: Detection Result



Fig. 4: Object Detected

V. CONCLUSION

This project presents a robust and efficient object detection framework tailored for UAV aerial imagery using the YOLOv10 model. By addressing the inherent challenges of UAV images such as small object sizes, cluttered backgrounds, and noisy conditions, the system demonstrates improved detection accuracy and reliability. The integration of adaptive feature enhancement, multi-scale detection heads, and convolutional attention mechanisms ensures that fine-grained details are captured without sacrificing speed. Through effective preprocessing, segmentation, and feature extraction, the framework prepares UAV imagery for precise detection. Real-time inference capability makes it highly suitable for critical UAV applications such as traffic monitoring, disaster response, and surveillance. Experimental evaluations validate the superiority of YOLOv10 over conventional methods in aerial scenarios. The project successfully balances accuracy with computational efficiency, ensuring scalability in real-world environments. Moreover, the modular design allows easy extensions and integration with emerging technologies. The results highlight YOLOv10's potential as a reliable and

scalable solution for UAV-based detection tasks. In conclusion, the framework contributes significantly to the advancement of intelligent UAV systems, opening new possibilities for aerial monitoring and management.

REFERENCES

[1]J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 779–788.

[2]J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 6517–6525.

[3]J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," in Proc. Comput. Vis. Pattern Recognit., Berlin, Germany: Springer, Jan. 2018, pp. 1–6.

[4] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," arXiv preprint arXiv:2004.10934, 2020.

[5]Ultralytics, "YOLOv5," 2020. [Online]. Available: https://github.com/ultralytics/yolov5

[6] Ultralytics, "YOLOv8," 2023. [Online]. Available: https://github.com/ultralytics/ultralytic

AUTHORS Profile



Mr. M. RAMAN KUMAR received his B.Tech degree in Computer Science and Engineering from Ranchi University in 2010 and his M.Tech degree in Computer Science and Engineering from Bhagwant University, Ajmer, in 2020. He is presently pursuing his Ph.D. from Atal Bihari Vajpayee Indian Institute of Information Technology and Management (ABV-IIITM), Gwalior. He has been dedicated to the teaching profession for the past 2 years. At present, he is working as an Assistant Professor in the Department of Computer Science and Engineering at Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India.



Mr. UGGIDI MAHENDRA BRAHMAM pursuing B. Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Ms. KARRA SREEJA REDDY pursuing B. Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Mr. SHEELAM RAHUL pursuing B. Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Ms. N. CHAITANYA pursuing B. Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.