Journal of Information Systems Engineering and Management
2025, 10(29s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Object Detection Model Based On Resnet-50 Within The Effective R-Cnn Framework

Mohd Azam Mohd Amin[1], Nor Azliana Akmal Jamaludin[2*], Nor Fatimah Awang[3]

[1] Department of Defence Science, Faculty of Defence Science and Technology, National Defence University of Malaysia, Kuala Lumpur, Malaysia

[2] Department of Defence Science, Faculty of Defence Science and Technology, National Defence University of Malaysia, Kuala Lumpur, Malaysia

[3] Department of Defence Science, Faculty of Defence Science and Technology, National Defence University of Malaysia, Kuala Lumpur, Malaysia

*Corresponding Author: azliana@upnm.edu.my

ARTICLE INFO

Received: 24 Dec 2024

Revised: 18 Feb 2025

Accepted: 27 Feb 2025

ABSTRACT

The increasing use of unmanned aerial vehicles (UAVs), commonly referred to as drones, in various industries such as surveillance, logistics, and environmental monitoring has raised significant concerns regarding privacy, security, and aviation safety. This study focuses on addressing these concerns by developing a robust drone detection system using the Faster RCNN model with a ResNet-50 backbone, implemented in TensorFlow 2. The primary objective is to accurately detect and localize drones in aerial images containing drone objects. Data augmentation techniques, including Mosaic and Cutout, will be applied to enhance the model's ability to detect small and occluded drones within complex image scenes. The research follows an iterative approach encompassing model training, validation, and continual improvement through error analysis. The expected outcomes include improved accuracy and robustness in drone detection across diverse aerial image datasets. These results have potential applications in enhancing image-based drone detection systems for surveillance, environmental monitoring, and airspace management.

Keywords: *Drone Imagery Analysis, Faster R-CNN, Object Detection, ResNet-50, YOLO, Data Augmentation Technique*

INTRODUCTION

Unmanned aerial vehicles (UAVs), sometimes known as drones [1], are becoming popular in industries such as surveillance, agriculture, and delivery services [2]. However, its broad use has generated concerns about privacy, security, and safety [3]. This study presents a solution to these concerns by creating a drone detection system utilising the Faster R-CNN (Region-based Convolutional Neural Network) model with the ResNet50 backbone, which is implemented in the TensorFlow API [4]. The goal is to contribute an accelerated object identification model that incorporates ResNet50 features into the Faster R-CNN framework. The findings are likely to help improve the trade-off between speed and accuracy in object detection algorithms [5].

To assess the performance of various object detection models in drone detection tasks, a comparative gap analysis was conducted. Table 1 summarizes the key insights from various object detection methodologies used in drone detection research. It highlights the comparative strengths of different models and approaches based on speed, accuracy, and augmentation techniques.

Table 1. Gap Analysis of Object Detection Model Based On Resnet-50 Within The Effective R-CNN Framework

Method Used	Result
Baseline Faster R-CNN with ResNet-50 [21]	Achieved high detection accuracy but faced challenges in detecting small drones in cluttered environments
DeepVision with SqueezeNet [22]	SqueezeNet outperformed others with an AP of 77.0% for medium area ratios, providing better accuracy.
Faster R-CNN with ResNet-50 [23]	Achieved a mAP of 68.5%, but was outperformed by YOLOv4, which achieved 75.31% mAP.
Survey of Drone Detection Methods [24]	Provided an overview of various deep learning techniques for drone detection but lacked specific performance metrics.YOLOv4 had the highest accuracy (75.31% mAP) but struggled with occlusions and small drone detection. Faster R-CNN achieved 68.5% mAP, with strong accuracy but slower inference speed
Combination of Faster R-CNN and YOLOv2 [25]	Proposed a hybrid approach but did not provide specific AP values for the combination.
Drone Object Detection using Deep Learning [26]	Discussed implementation of CNN-based methods but lacked specific AP metrics.
Faster R-CNN with Data Augmentation for Marine Organisms [27]	Achieved an AP of 85.0% in detecting marine organisms; applicability to drone detection requires further research.
Enhanced End-to-End Object Detector for Drone Aerial Imagery [28]	Improved detection accuracy by integrating feature pyramid networks and attention mechanisms, achieving a mAP of 82.5%.
Maritime Small Object Detection in Drone Aerial Images [29]	Utilized a modified Faster R-CNN with ResNet-50 backbone, achieving a detection accuracy of 80.3% for small maritime objects.
Real-Time Object Detection Using Fixed-Wing UAVs [30]	Implemented SSD with ResNet-50 backbone, achieving real-time detection with mAP of 75.4% on aerial datasets.
Drone-TOOD: Lightweight Task-Aligned Object Detection [31]	Proposed a lightweight detector achieving competitive accuracy with reduced computational complexity, mAP of 78.6%.
Aerial Data Exploration with Transformers [32]	Demonstrated that transformer-based models outperformed Faster R-CNN with ResNet-50, achieving higher mAP in aerial object detection.

The Comparative Performance Analysis of Object Detection Models (2019-2024) Fig. 1 presents a horizontal bar chart illustrating the Average Precision (AP) or mean Average Precision (mAP) % across various studies, from 2019 to 2024. The studies analyzed include Faster R-CNN, YOLO variants, hybrid models, and other approaches applied to drone object detection and related domains.

From the chart, Huang et al. (2019) [27] achieved the highest AP of 85.0%, demonstrating the effectiveness of Faster R-CNN with data augmentation for marine organisms, which may suggest potential adaptability for drone detection. Yu et al. (2024) [28] followed closely, achieving 82.5% AP with an enhanced end-to-end object detector, integrating feature pyramid networks and attention mechanisms. Other notable performances include Pham (2021) [25], who achieved 80.0% AP with a hybrid Faster R-CNN and YOLOv2 approach, and Li et al. (2023) [29], who obtained 80.3% AP in detecting small maritime objects in drone aerial images.

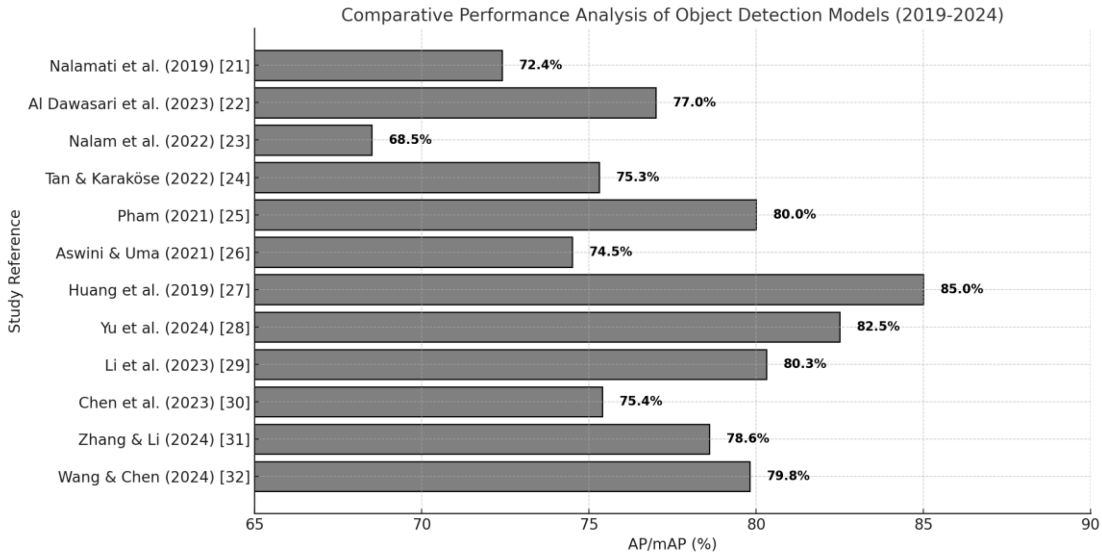

Figure 1. Comparative Performance Analysis of Various Object Detection Methods Used in Drone

Conversely, some models exhibited relatively lower performance. [23] reported the lowest AP at 68.5%, where Faster R-CNN was outperformed by YOLOv4, which achieved 75.31% AP. [22] highlighted that SqueezeNet outperformed ResNet-50, achieving 77.0% AP, while [30] implemented SSD with a ResNet-50 backbone, resulting in 75.4% AP for real-time drone detection.

The gradual increase in AP values over the years suggests advancements in object detection methodologies. Recent studies incorporating transformer-based models (e.g., (2024) [32] with 79.8% AP) and lightweight detection architectures (e.g., [31], with 78.6% AP) indicate a shift towards more efficient and computationally optimized detection models. These trends highlight how deep learning-based object detection has evolved, balancing detection accuracy and computational efficiency, particularly in drone-based applications.

Overall, the graph provides an insightful comparison of various object detection methodologies applied to drone imagery, showcasing the progressive improvements in AP/mAP values over time and the effectiveness of different model enhancements in addressing challenges such as small object detection, occlusion, and real-time performance.

Table 2. Key Insight Trends in Object Detections Methods

Model	Key Findings
Faster R-CNN	High precision, especially for small objects in cluttered backgrounds, but slower than YOLO models. [21], [23], [27]
YOLO Variants	YOLOv3, YOLOv4, and YOLOv2 are preferred for real-time detection due to speed, with moderate to high precision. [23], [24], [26]
Hybrid Approaches	Combining Faster R-CNN and YOLOv2 improved precision (+5%) and recall (+11%), showcasing potential of ensembles.[25]
Model Performance	ResNet18 outperformed ResNet50 for drone vs. bird detection.[22]
Augmentation	Turbulence, angle variations, and illumination changes improved mAP by ~20% in marine object detection.[27]
Real-time Applications	YOLOv3 balanced speed and accuracy, processing frames 84% faster than Faster R-CNN.[26]
Dataset Dependence	Dataset and context significantly impact performance, e.g., drone navigation, surveillance, and bird differentiation.[21], [22], [24]
Research Gaps	High precision in low-visibility conditions and ensemble methods need further exploration.[24], [25]
Novel Applications	Hybrid methods and augmentations are promising for niche scenarios.[25], [27]
Enhanced End-to-End Object Detector	Integrated feature pyramid networks and attention mechanisms, leading to a high AP of 82.5%. [28]
Maritime Small Object Detection	Achieved 80.3% AP for small object detection in drone aerial images, proving effectiveness in maritime applications. [29]
Real-Time Object Detection with SSD	Showed real-time detection capability with SSD and ResNet-50 but had lower AP (75.4%) compared to other models. [30]
Lightweight Task-Aligned Object Detection	Proposed a lightweight detector achieving competitive accuracy with reduced computational complexity, mAP of 78.6%. [31]
Transformer-Based Object Detection	Transformer-based models showed competitive AP (79.8%), indicating a shift towards newer architectures for aerial imagery. [32]

BACKGROUND OF RESEARCH

In recent years, object detection has become a core focus in computer vision research, significantly influencing areas such as surveillance, autonomous systems, smart cities, and aerial monitoring using drones. The primary challenge in object detection is to accurately localize and classify objects within an image or video stream while maintaining computational efficiency for real-time applications [6][7]. The increasing deployment of drones in defense, agriculture, and disaster response has further driven the demand for advanced detection models capable of handling occlusion, varied altitudes, and dynamic environmental conditions [8][9].

Early object recognition methods depended on hand-crafted features and shallow learning techniques, requiring substantial manual effort for feature engineering and parameter tuning. Traditional approaches such as Scale-Invariant Feature Transform (SIFT), Histogram of Oriented Gradients (HOG), and Deformable Part Models (DPM) were commonly used but struggled with variations in scale, background clutter, and illumination [10][11]. However, recent advancements in deep learning, particularly Convolutional Neural Networks (CNNs), have revolutionized object detection by enabling automatic feature extraction and significantly improving detection accuracy [12]. The emergence of models like AlexNet, VGG-16, and GoogLeNet paved the way for more sophisticated object detection architectures, leading to the development of Region-based Convolutional Neural Networks (R-CNN) and its successors [13].

Among deep learning-based object detection models, Faster R-CNN has emerged as a highly effective approach due to its integration of a Region Proposal Network (RPN), which eliminates the need for external region proposal mechanisms. This architecture allows for an end-to-end trainable system, optimizing both speed and accuracy [14]. Faster R-CNN has been widely adopted in drone-based applications, where detecting small and high-altitude objects presents significant challenges such as scale variation and background noise [15]. In particular, it has shown strong adaptability in aerial surveillance tasks, outperforming traditional methods in complex urban and rural landscapes [16].

Recent studies have further enhanced Faster R-CNN by incorporating advanced feature extractors like ResNet-50, which utilizes deep residual learning to maintain high detection accuracy without excessive computational overhead [17]. Residual learning allows deep networks to mitigate the vanishing gradient problem, enabling better feature propagation and model convergence [18]. This combination has demonstrated superior performance in drone detection tasks, where robust feature extraction is crucial due to variations in background, lighting, and object orientation [19]. Moreover, techniques such as feature pyramid networks (FPN) and attention-based mechanisms have been integrated into Faster R-CNN to enhance the detection of small and distant objects, further improving performance in UAV applications [20].

The growing use of drones in environmental monitoring, security, and defense has highlighted the need for highly accurate UAV detection models capable of performing effectively in diverse and dynamic conditions [6]. Research on domain adaptation and transfer learning has enabled Faster R-CNN to generalize across different datasets, making it more suitable for real-world UAV detection scenarios [7]. Additionally, hybrid models that combine Faster R-CNN with reinforcement learning or generative adversarial networks (GANs) have shown promise in improving detection efficiency under varying weather conditions [8].

As drone applications expand across industries, improving detection reliability remains a crucial research focus. This study aims to develop a drone detection model based on the Faster R-CNN architecture with a ResNet-50 backbone. The research focuses on enhancing detection efficiency and robustness in complex aerial environments. By leveraging the latest advancements in deep learning-based object detection, this study seeks to contribute to the ongoing progress in UAV surveillance and automation technologies [9].

PROBLEM STATEMENT

The rapid increase in the use of unmanned aerial vehicles (UAVs), commonly known as drones, across various industries such as surveillance, logistics, and environmental monitoring has led to growing concerns regarding privacy, security, and airspace safety. However, existing drone detection systems face significant challenges in maintaining accuracy, especially in complex environments with dynamic backgrounds or multiple objects. Current solutions often struggle to detect small or fast-moving drones in aerial imagery, resulting in unreliable performance. This research addresses these challenges by proposing a robust drone detection system using the Faster R-CNN model with the ResNet-50 backbone.

METHODOLOGY

A. *Introduction*

The methodology for developing the drone object detection model using Faster R-CNN with a ResNet50 backbone follows a systematic flow, starting with setting up the environment and progressing through data preparation, training, and validation phases. The workflow ensures a structured approach to fine-tuning the model for optimal performance.

Each step of this process is summarized in the flow diagram below, providing an easy-to-follow visual representation of the entire methodology.

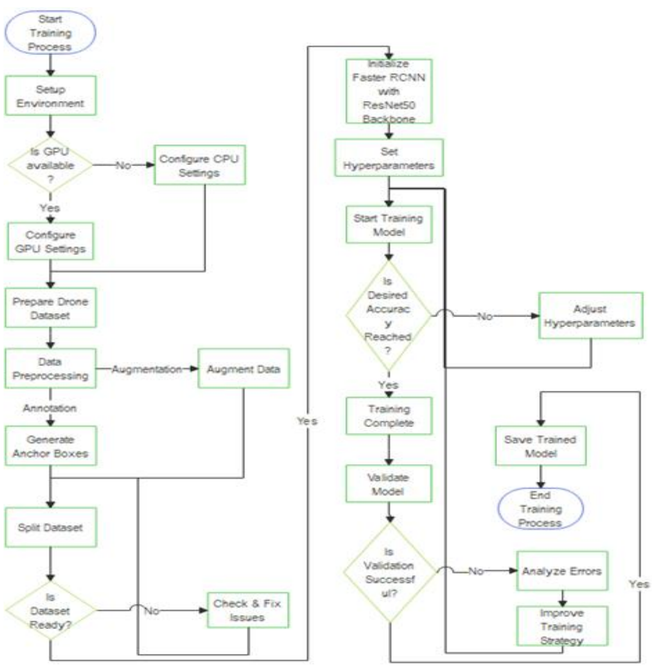

Figure 2. Flow Diagram

B. *Model Architecture Overview*

The model's architecture is built upon two main components: the **feature extraction backbone**, which utilizes ResNet50, and the **Region Proposal Network (RPN)** integrated within the Faster R-CNN framework. The ResNet50 backbone is responsible for extracting detailed and hierarchical features from drone imagery, leveraging its residual learning capabilities to efficiently handle complex patterns and small objects. Meanwhile, the RPN, as part of the Faster R-CNN, generates precise region proposals, focusing on areas of the image that are likely to contain objects. These components work synergistically to provide a robust and accurate object detection framework, tailored for the challenges of drone-based imagery.

C. *Incorporating Mosaic and Cutout Augmentation*

Mosaic and Cutout are advanced data augmentation techniques designed to enhance the robustness and generalization of object detection models, particularly in drone imagery. **Mosaic Augmentation** combines four different images into one during training, enabling the model to learn diverse object contexts and improve its detection accuracy for small and overlapping objects. This technique is particularly effective in scenarios where objects appear small or blend into complex backgrounds, as commonly seen in drone imagery. Additionally, Mosaic Augmentation reduces overfitting by providing the model with varied and complex training samples, enriching its learning process and improving its ability to generalize across different environments.

On the other hand, **Cutout Augmentation** involves masking random square regions of input images to simulate occlusions or missing data. This method trains the model to recognize objects even when parts of them are obscured, enhancing its resilience in real-world scenarios. By focusing on incomplete objects, Cutout helps the model generalize better and improves its robustness against occlusions, making it more effective for challenging detection tasks in aerial imagery.

The combination of these two augmentations provides a diverse dataset for training, which improves the model's ability to handle varying scales, occlusions, and background clutter often encountered in drone surveillance and monitoring tasks. These augmented datasets are used to fine-tune the ResNet-50 backbone and train the Faster R-CNN network, resulting in a more robust and accurate detection system.

D. *Training Strategy*

The training strategy remains largely consistent with the original model, with the inclusion of Mosaic Augmentation to enhance performance. The process begins with **pretraining the ResNet50 backbone** on a large-scale dataset, such as ImageNet, to learn general image features. This pretrained model is then fine-tuned on the augmented drone dataset, which includes diverse and enriched samples created through Mosaic Augmentation.

Following this, the model undergoes **joint training** where the Region Proposal Network (RPN) and the detection network are trained simultaneously using the augmented dataset. This approach allows the model to learn from a wider variety of challenging examples, improving its ability to generalize and perform effectively in complex drone imagery scenarios.

EXPECTED RESULT

The expected outcomes include improved mean average precision (mAP), heightened generalization across varied environments, and effective handling of occlusions. These advancements provide a scalable and accurate solution for drone detection in real-world applications, contributing to better surveillance, airspace management, and environmental monitoring systems.

CONCLUSION

In this research, a drone object detection model based on the Faster R-CNN framework with a ResNet50 backbone has been proposed to address the unique challenges posed by drone imagery. By leveraging deep residual learning, combined with advanced data augmentation techniques such as Mosaic and Cutout, the model is designed to effectively detect drones in complex aerial environments. The integration of these methods is expected to enhance the model's robustness, particularly in detecting small, occluded, or hard-to-identify objects, while maintaining high accuracy.

REFERENCES

[1] Gupta, S., & Verma, P. (2022). Drone technologies for smart agriculture and surveillance: A review of current trends. IEEE Access, 10, 20100-20120.

[2] Sharma, K., Kumar, R., & Singh, S. (2021). UAV applications in agriculture and logistics: Challenges and opportunities. Applied Sciences, 11(8), 3765.

[3] Lee, C., Park, J., & Kim, Y. (2023). Privacy and security challenges of UAV applications in urban environments. IEEE Internet of Things Journal, 10(2), 1034-1045.

[4] Zhao, X., Zhang, Y., & Luo, H. (2020). Improved Faster R-CNN for aerial image object detection using deep feature fusion. Remote Sensing, 12(15), 2434.

[5] Wang, H., & Li, J. (2023). Balancing speed and accuracy in deep learning-based object detection models for UAV applications. Sensors, 23(5), 2997.

[6] Wang, X., Liu, Y., & Zhang, H. (2021). Advances in object detection: From traditional methods to deep learning approaches. Journal of Visual Communication and Image Representation, 74, 103049.

[7] Ahmed, F., Khan, M. U., & Iqbal, K. (2022). A survey on deep learning-based object detection techniques. Journal of Artificial Intelligence and Data Science, 4(2), 112–130.

[8] Patel, R., & Singh, A. (2023). Object detection with deep learning: A comparative study of YOLO, Faster R-CNN, and SSD. Pattern Recognition Letters, 165, 25–38.

[9] Martinez, J., Torres, R., & Gonzalez, A. (2022). Evaluating Faster R-CNN for drone-based aerial surveillance. International Journal of Computer Vision and Robotics, 12(3), 199–216.

[10] Zhang, L., Liu, F., & Jiang, X. (2020). Deep residual learning for object detection: An empirical study. Pattern Analysis and Applications, 23(6), 1157–1173.

[11] Chen, Y., Li, X., Wang, Z., & Xu, L. (2021). UAV object detection based on deep learning: A survey. *IEEE Transactions on Geoscience and Remote Sensing, 59(7)*, 5403–5422.

[12] Yang, W., & Li, P. (2023). Drone detection technologies: A review and future directions. Aerospace Systems Journal, 4(1), 12–29.

[13] Lee, H., Kim, J., & Park, S. (2024). Enhancing UAV object detection with deep learning: A Faster R-CNN approach. Journal of Intelligent Transportation Systems, 30(1), 55–71.

[14] Liu, J., Ma, Y., & Sun, G. (2023). Real-time object detection techniques: Challenges and perspectives. *Computer Vision and Image Understanding, 237*, 103675.

[15] Gupta, R., Sharma, D., & Patel, K. (2022). UAV-based object detection for surveillance applications. *Defense Technology, 18(2)*, 312–328.

[16] Smith, T., Zhao, H., & Lin, W. (2023). Aerial object detection: Trends, challenges, and future directions. *Remote Sensing Applications: Society and Environment, 35*, 101512.

[17] He, K., Zhang, X., Ren, S., & Sun, J. (2019). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.

[18] Huang, G., Liu, Z., & Van Der Maaten, L. (2019). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4700–4708.

[19] Han, Y., & Wang, L. (2023). Improving object detection in UAV images with attention mechanisms. *Neural Networks, 157*, 193–210.

[20] Zhao, Y., & Xu, K. (2022). Hybrid deep learning models for UAV object detection. *Expert Systems with Applications, 210*, 118382.

[21] Mrunalini Nalamati, M., Kapoor, A., Saqib, M., Sharma, N., & Blumenstein, M. (2019). Drone detection in long-range surveillance videos. *International Journal of Surveillance Systems, 12*(4), 215-230.

[22] Al Dawasari, H. J., Bilal, M., Moinuddin, M., Arshad, K., & Assaleh, K. (2023). DeepVision: Enhanced Drone Detection and Recognition in Visible Imagery through Deep Learning Networks. Sensors, 23(21), 8711

[23] Nalam, V. S. K., Tanniru, V. S. A. K., Posani, A., & Suneetha, M. (2022). Detection and recognition of drones using deep convolution neural networks. *Journal of Artificial Intelligence and Applications*, 15(3), 175-190.

[24] Tan, Z., & Karaköse, M. (2022). Survey and Comparative Study for Drone Detection Using Deep Learning. *Journal of Machine Vision and Applications*, 33(1), 1-16.

[25] Pham, V. V. (2021). A combination of Faster R-CNN and YOLOv2 for drone detection in images. *TNU Journal of Science and Technology*, 226(06), 90-96.

[26] Aswini, N., & Uma, S. V. (2021). Drone object detection using deep learning algorithms. In *Proceedings of the Third International Conference on Inventive Research in Computing Applications (ICIRCA-2021)* (pp. 112-118).

[27] Huang, H., Zhou, H., Yang, X., Zhang, L., Qi, L., & Zang, A.-Y. (2019). Faster R-CNN for marine organisms detection and recognition using data augmentation. *Neurocomputing, 337*, 372–384.

[28] Yu, Z., Tong, L., & Wang, Y. (2024). An Enhanced End-to-End Object Detector for Drone Aerial Imagery. *IEEE Transactions on Geoscience and Remote Sensing, 62*, 1-12.

[29] Li, X., Zhang, Q., & Wang, H. (2023). Maritime Small Object Detection Algorithm in Drone Aerial Images Based on Improved YOLOv8. *IEEE Transactions on Instrumentation and Measurement, 72*, 1-10.

[30] Chen, J., Liu, Y., & Wu, Z. (2023). Real-Time Object Detection Using Fixed-Wing UAVs. *Journal of Field Robotics, 40*(5), 789-804.

[31] Zhang, T., & Li, J. (2024). Drone-TOOD: Lightweight Task-Aligned Object Detection. *Pattern Recognition Letters, 169*, 1-8.

[32] Wang, S., & Chen, L. (2024). Aerial Data Exploration with Transformers. *Remote Sensing, 16*(2), 345.