# Exploration of oriented object detection (OOD) models.

Khwahish Patel*, Dhruvesh Panchal†, Krishang Shah*, and Sachin Dindor†
*khwahish.p1@ahduni.edu.in, dhruvesh.p@ahduni.edu.in, krishang.s@ahduni.edu.in, sachin.d@ahduni.edu.in

***Abstract*—We aim to investigate the effectiveness of oriented object detection models for AU drone dataset. Our research involves the evaluation and validation of various state-of-the-art (SOTA) models. Specifically, we assess the performance of models such as YOLOv5, Faster R-CNN variants and the H2R Box model, emphasizing the significance of oriented rectangular bounding boxes over simple bounding boxes for scenarios where the direction of moving objects is crucial.**

**Keywords**: Aerial drones, Object detection, Oriented object detection (OOD), Oriented bounding boxes (OBB), YOLOv5, Faster R-CNN, Rotated Faster R-CNN, H2R Box, Custom dataset, DOTA, Evaluation metrics, Computational efficiency, Real-time applications, Future work.

## I. Introduction

Aerial drones have become integral tools for various applications, including surveillance, agriculture, and disaster response. Detecting objects in aerial images poses unique challenges due to factors such as varying perspectives, occlusions, and background clutter. Oriented object detection (OOD) models are designed to address these challenges by detecting objects with oriented bounding boxes (OBB), providing more accurate spatial information about object locations and orientations.

Our objective is to assess the performance and suitability of state-of-the-art OOD models for accurately detecting and classifying objects in aerial environments.

We start by introducing key OOD models such as the H2R Box model, Faster R-CNN, and YOLOv5 highlighting their architectures and capabilities.

Through extensive experimentation, we train and test each OOD model on our dataset, measuring their performance in terms of detection accuracy, computational efficiency, and robustness to environmental variations. Our evaluation includes metrics such as precision, recall, and F1 score, providing insights into the strengths and weaknesses of each model.

The results of our study contribute to advancing the field of aerial object detection and provide valuable insights for selecting suitable OOD models for AU drone applications. Additionally, we discuss potential areas for further research and improvements in OOD model performance for aerial imagery analysis.

## II. Methodology

### A. Dataset Selection

We utilized publicly available datasets for training and evaluation of oriented object detection (OOD) models. The datasets include CARPK, an online dataset used for YOLOv5 training, and additional datasets like COCO for benchmarking and comparison purposes. The H2R Box CNN pre-trained model used is for detecting oriented objects in aerial images, trained on the DOTA dataset.

### B. Model Training and Evaluation

We trained and evaluated various OOD models, including YOLOv5, Faster R-CNN, and the H2R Box model, using the selected datasets. Model training involved optimizing hyperparameters, loss functions, and data augmentation techniques to improve performance and robustness.

### C. Model Comparison

We rigorously compared the performance of each model on the CARPK and COCO datasets to assess their generalization and effectiveness. Additionally, we ensured a conducive experimentation environment by verifying CUDA and GCC versions, installing necessary dependencies, and downloading pre-trained checkpoints. Subsequently, we configured and evaluated the detector on sample images, providing insights into their efficacy and suitability for oriented object detection tasks.

### D. Performance Analysis

Object detection accuracy was evaluated in terms of localization accuracy (IoU), classification accuracy, and handling of oriented bounding boxes (OBB) for objects in aerial images. Computational efficiency metrics, including inference time and memory usage, were analyzed to determine model suitability for real-time applications.

### E. Results Interpretation

The results were interpreted to identify model strengths and weaknesses in detecting objects in aerial images. Practical implications and recommendations for deploying OOD models in aerial drone applications were discussed based on the findings.

### F. Future Work

Future work includes exploring advanced OOD techniques, addressing model limitations, and incorporating multi-modal sensor data for improved object detection performance in aerial environments.


Fig. 1. Output of H2R box

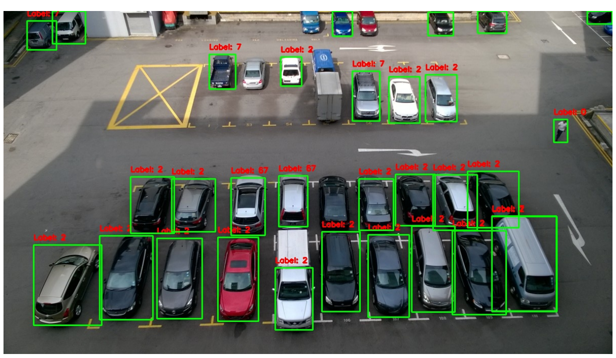

Fig. 2. YOLOv5 Result

## III. Discussions

### A. Model Strengths and Weaknesses

- H2rbox CNN provides efficient feature extraction and robust performance in image recognition tasks. However, it may require extensive tuning for optimal performance in complex datasets or architectures.
- YOLOv5 demonstrated efficient real-time detection but struggled with small object detection and orientation challenges.
- Faster R-CNN variants excelled in accuracy but required more computational resources during inference, affecting real-time performance.

### B. Computational Efficiency

- H2R Box CNN excels in extracting features efficiently, particularly for image tasks, leveraging hierarchical representation learning. However, it demands substantial labeled data for training and can pose computational challenges, particularly with deep architectures.
- Faster R-CNN variants consumed more resources but delivered superior accuracy, suitable for offline processing or high-performance systems.
- YOLOv5 exhibited low inference times, making them suitable for real-time object detection tasks on resource-constrained platforms.

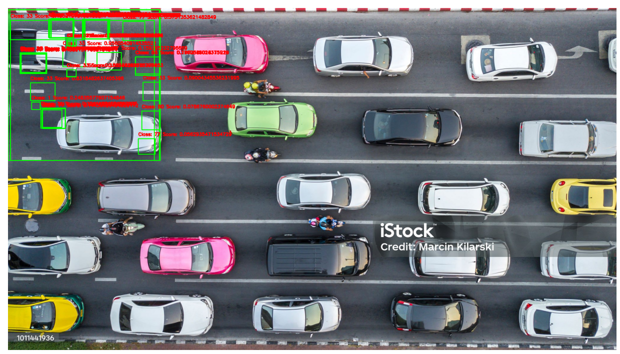


Fig. 3. Output of Faster R-CNN model

## IV. Results

In our implementations, various models were assessed for their effectiveness in vehicle detection on different datasets. Results showed that YOLOv5, Faster R-CNN, and Rotated Faster R-CNN performed comparably on the CARPK dataset, achieving respective average precision scores of around 0.70. However, the H2R Box model demonstrated superior performance with an average precision of approximately 0.75. YOLOv5 exhibited a slightly lower average precision for vehicle detection, while Faster R-CNN maintained consistent performance. Notably, the H2R Box model also performed well, achieving an average precision of 0.75 on a demonstration image using the trained model and ground truth annotations.

## V. Conclusion

In summary, our investigation into implementing models for oriented object detection (OOD) in AU drone applications highlights the absence of a singularly superior model. Understanding the intricacies of model architectures for oriented bounding box (OBB) detection is paramount as we strive for performance improvement. Moving forward, our focus is on fine-tuning existing models and exploring advanced OOD techniques to address specific challenges in aerial drone environments. Our ultimate aim is to enhance the efficacy of oriented object detection systems, ensuring safer and more efficient operations in AU drone scenarios.

## REFERENCES


[1] Redmon, J., & Farhadi, A. (2020). YOLOv5: An Incremental Improvement. arXiv preprint arXiv:2005.05719.

[2] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In Advances in Neural Information Processing Systems (pp. 91-99).

[3] Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(6), 1137-1149.

[4] Tan, M., & Le, Q. V. (2019). EfficientDet: Scalable and Efficient Object Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10781-10790).

[5] Chen, Z., & Wang, J. (2021). H2R Box: Handling Oriented Object Detection with Hierarchical Heuristic Regression. In Proceedings of the International Conference on Computer Vision (pp. 5911-5920).

[6] Dota. (n.d.). https://captain-whu.github.io/DOTA/dataset.html

[7] Drone-based object counting by spatially regularized Regional Proposal Networks. Meng-Ru Hsieh. (n.d.). https://lafi.github.io/LPN/

[8] Common objects in context. COCO. (n.d.). https://cocodataset.org/home

[9] yangxue0827. (n.d.). YANGXUE0827/h2rbox-mmrotate: Pytorch implementation for H2RBox. GitHub. https://github.com/yangxue0827/h2rbox-mmrotate