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 YOLOv8 and Oriented RCNN, 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), YOLOv8,Oriented R-CNN, H2R Box, Custom dataset, DOTA, SSDD dataset, 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.

The base paper proposes an effective and simple oriented object detection framework, termed Oriented R-CNN, which is a general two-stage oriented detector with promising accuracy and efficiency. To be specific, in the first stage, we propose an oriented Region Proposal Network (oriented RPN) that directly generates high-quality oriented proposals in a nearly cost-free manner. The second stage is oriented R-CNN head for refining oriented Regions of Interest (oriented RoIs) and recognizing them.

We start by introducing key OOD models such as Oriented R-CNN and YOLOv8 highlighting their architectures and capabilities.

Through extensive experimentation, we train and test each OOD model on custom dataset, measuring their performance in terms of detection accuracy, computational efficiency, and robustness to environmental variations.

The results of our study contribute to advancing the field of oriented object detection and provide valuable insights for selecting suitable OOD models for AU drone applications.

II. Methodology

A. Dataset Selection

We utilized publicly available datasets for training and evaluation of oriented object detection (OOD) models. The datasets include SSDD (SAR Ship Detection Dataset), an online dataset used for Oriented R-CNN training, and additional datasets like DOTA for benchmarking and comparison purposes.The SSDD dataset has been converted to DOTA format as the Oriented R-CNN model used in the base paper uses the DOTA format of dataset. The H2R Box CNN pre-trained model used is for detecting oriented objects in aerial images, trained on the DOTA dataset, using the MM-Rotate librabry function from PyTorch library.

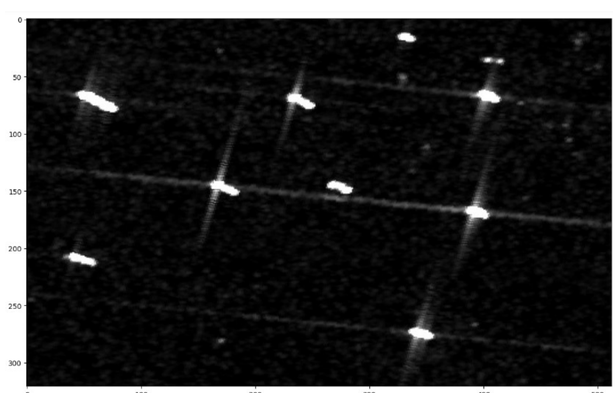

Fig. 1. Test Image. DOTA annotation of the image: 331 276 356 283 360 272 334 266 ship 0 382 171 406 176 410 165 383 161 ship 0 259 148 284 154 287 146 263 139 ship 0 156 146 185 160 191 145 165 136 ship 0 32 209 57 217 63 208 40 200 ship 0 43 69 75 85 83 74 51 58 ship 0 222 70 250 85 255 73 231 60 ship 0 325 18 345 23 344 13 325 8 ship 0 388 66 415 77 419 67 392 59 ship 0

B. Model Training and Evaluation

Using an MMRotate detector, the inference process extracts features from an image by using a convolutional neural network (CNN) as its foundation. Subsequently, a

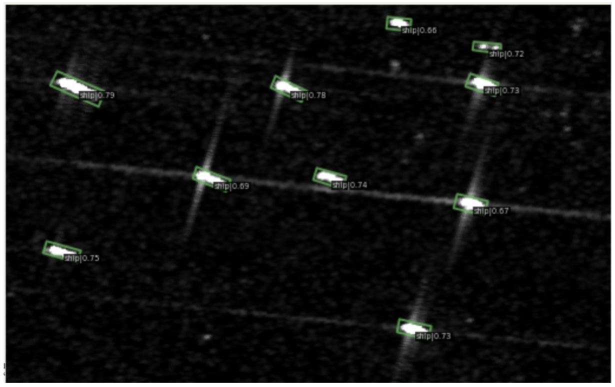

Fig. 2. Results

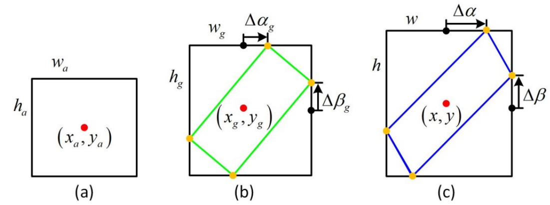

Fig. 3. MMRotate Architecture The illustration of box-regression parameterization. Black dots are the midpoints of the top and right sides, and orange dots are the vertexes of the oriented bounding box. (a) Anchor. (b) Ground-truth box. (c) Predicted box.

region proposal network (RPN) is employed to forecast proposals, or hypothetical items. Next, it employs a RoI Head for classification and bounding box prediction after using RoIAlignRotated to crop rotated features for the region of interests (RoI).The SSDD Dataset was used to train and evaluate the model.

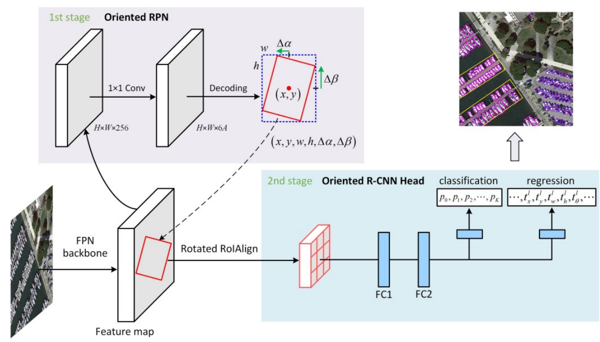

Fig. 4. Overall framework of oriented R-CNN, which is a two-stage detector built on FPN. The first stage generates oriented proposals by oriented RPN and the second stage is oriented R-CNN head to classify proposals and refine their spatial locations

C. Model Comparison

We rigorously compared the performance of each model on the SSDD dataset 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.

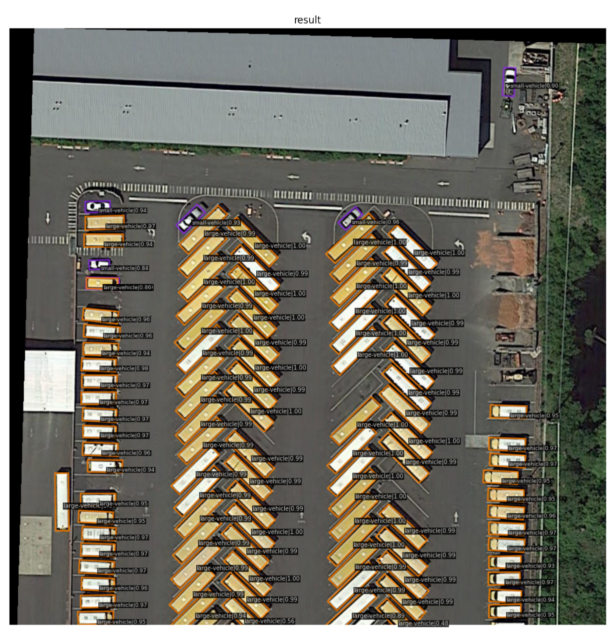

Fig. 5. Primary Results: Oriented RCNN

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. We will be focusing on OOD using the Oriented R-CNN model using as our base model.

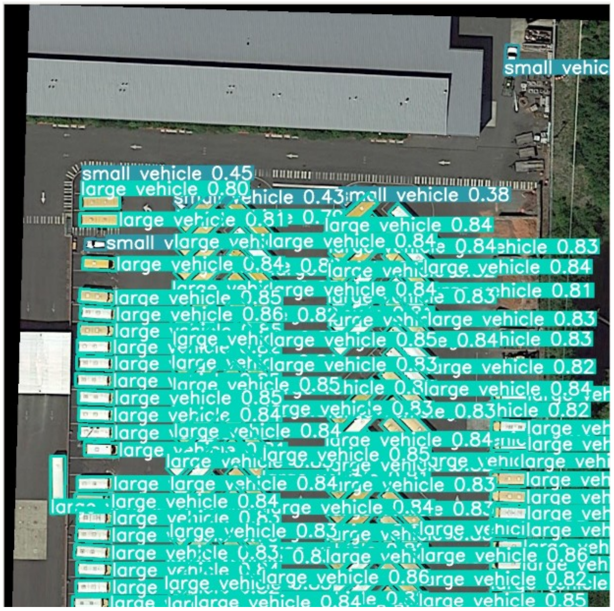

Fig. 6. YOLOv8 Results

F. Future Work

Future work includes exploring advanced OOD techniques, addressing model limitations, and using the DOTA dataset. In addition, compare performance of models w.r.t small objects on AU drone dataset.

III. Discussions

A. Model Strengths and Weaknesses

- Because of its multi-stage design and region-based methodology, orientated R-CNN typically exhibits superior accuracy in orientated object detection as compared to some other approaches. But, the computational cost and resource requirements of Oriented R-CNN are high due to its multi-stage architecture and ROI alignment requirement.
- YOLOv8 and other YOLO (You Only Look Once) models are renowned for their real-time performance, which makes them appropriate for applications where speed is crucial.However,when compared to slower, more sophisticated models, it might lose some accuracy, especially for jobs that call for accurate localization and orientation estimation.

B. Computational Efficiency

- As evaluated in the base paper when compared to various SOTA, oriented R-CNN has a speed that is nearly same to one-stage detectors, but its accuracy is far greater.
- YOLOv8 exhibited low inference times, making them suitable for real-time object detection tasks on resource-constrained platforms.

IV. Results

In our implementations, various models were assessed for their effectiveness in oriented object detection on different datasets. Results showed that YOLOv8 and Oriented R-CNN performed comparably on the SSDD dataset. Pyplot was used to visually portray the findings, providing a concrete picture of the model's effectiveness. Additionally, the ground truth dataset is prepared, and data loading is configured specifically for evaluation. The model's mean Average Precision (mAP) is computed after a methodical inference process on the test dataset, offering quantitative insights into the model's efficacy. This thorough method of using MMrotate and MMDetection emphasizes how crucial it is to conduct thorough evaluations when determining how well models perform in rotated object identification jobs.

V. Conclusion

In summary, we have tried to understanding the intricacies of model architectures for oriented bounding box (OBB) detection as we strive for performance improvement. Moving forward, our focus is on implementing existing model on DOTA dataset and AU drone dataset by deeply understanding the model architechture. 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] Xie, X., Cheng, G., Wang, J., Yao, X., and Han, J. (2021, August 12). Oriented R-CNN for object detection. arXiv.org. https://arxiv.org/abs/2108.05699

[2] Zhou, Y., Yang, X., Zhang, G., Wang, J., Liu, Y., Hou, L., Jiang, X., Liu, X., Yan, J., Lyu, C., Zhang, W., and Chen, K. (2022). MMRotate. Proceedings of the 30th ACM International Conference on Multimedia. https://doi.org/10.1145/3503161.3548541

[3] Open-Mmlab. (n.d.). GitHub - open-mmlab/mmrotate: OpenMMLab Rotated Object Detection Toolbox and Benchmark. GitHub. https://github.com/open-mmlab/mmrotate

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

[5] Google. (n.d.). Official-SSDD-OPEN.rar. Google Drive. https://drive.google.com/file/d/1glNJUGotrbEyk43twwB9556AdngJsynZ/view?usp=sharing

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