Exploration of oriented object detection (OOD) models.

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

***Abstract*—We're looking into how well object detection models designed for orientation perform with the AU drone dataset. Our study focuses on testing and confirming the capabilities of several cutting-edge models. In particular, we're analyzing the effectiveness of Oriented RCNN, highlighting the importance of using oriented rectangular bounding boxes instead of basic ones in situations where the movement direction of objects is vital.**

Keywords: Aerial drones, Object detection, Oriented object detection (OOD), Oriented bounding boxes (OBB), Oriented R-CNN, H2R Box, Mmrotate, Custom dataset, DOTA, SSDD dataset, AU Drone dataset, Computational efficiency, Real-time applications, Future work.

I. Introduction

Aerial drones have become essential tools across various domains like surveillance, agriculture, and disaster response. Detecting objects in aerial images faces unique challenges due to factors such as perspective shifts, obstructions, and background noise. Oriented object detection (OOD) models are specifically crafted to tackle these challenges by using oriented bounding boxes (OBB) to offer more precise spatial details about object positions and orientations.

Our goal is to evaluate the performance and applicability of cutting-edge OOD models in accurately identifying and categorizing objects in aerial settings.

The referenced paper introduces a straightforward yet effective framework for oriented object detection called Oriented R-CNN, a versatile two-stage detector known for its promising accuracy and efficiency. In the initial stage, they propose an oriented Region Proposal Network (oriented RPN) that produces high-quality-oriented proposals with minimal computational cost. The subsequent stage involves an oriented R-CNN head that refines oriented Regions of Interest (oriented RoIs) and performs object recognition.

We start by presenting key OOD models like Oriented R-CNN, outlining their designs and capabilities.

Through rigorous experimentation, we train and evaluate each OOD model on a custom dataset, assessing their detection accuracy, computational performance, and resilience to environmental variations.

The outcomes of our research contribute to pushing the boundaries of oriented object detection and offer valuable insights into selecting appropriate OOD models for AU drone applications.

II. Methodology

A. Dataset Selection

We employed publicly available datasets to train and assess oriented object detection (OOD) models. These datasets include SSDD (SAR Ship Detection Dataset), which was utilized for training Oriented R-CNN, and additional datasets like DOTA, used for benchmarking and comparison. Our primary dataset aligns with the AU Drone dataset.

We converted the SSDD dataset into the DOTA format to ensure compatibility with the Oriented R-CNN model used in the base paper. Additionally, we utilized the H2R Box CNN pre-trained model, which is designed for detecting oriented objects in aerial images. This model was trained on the DOTA dataset and leveraged the MM-Rotate library function from the PyTorch library.

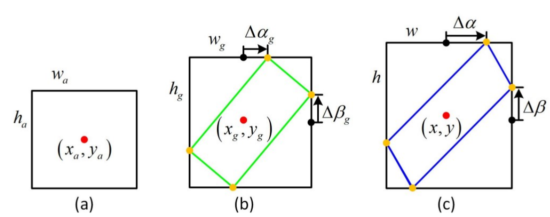

Fig. 1. MMRotate Architecture The illustration of box-regression parameterization. Black dots are the midpoints of the top and right sides, and orange dots are the vertexes of the oriented bounding box. (a) Anchor. (b) Ground-truth box. (c) Predicted box.

B. Model Training and Evaluation

The MMRotate detector operates through several stages during the inference process. It begins by extracting features from an image using a convolutional neural network (CNN) as its core. Following this, a region proposal network (RPN) is utilized to predict proposals or potential objects. Subsequently, it employs a Region of Interest (RoI) Head for classification and bounding

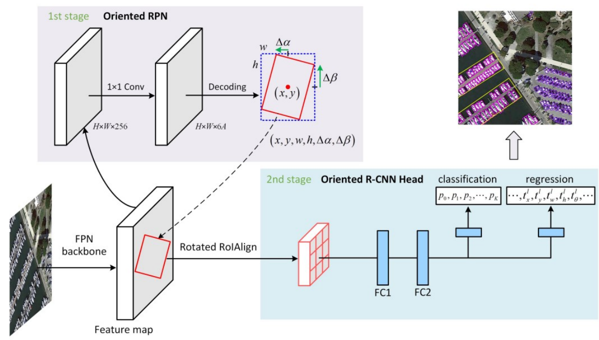

Fig. 2. Overall framework of oriented R-CNN, which is a two-stage detector built on FPN. The first stage generates oriented proposals by oriented RPN and the second stage is oriented R-CNN head to classify proposals and refine their spatial locations

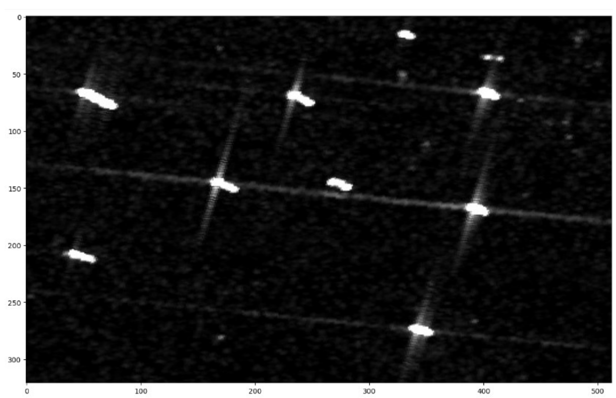

Fig. 3. Test image from SSDD dataset. DOTA annotation of the image: 331 276 356 283 360 272 334 266 ship 0 382 171 406 176 410 165 383 161 ship 0 259 148 284 154 287 146 263 139 ship 0 156 146 185 160 191 145 165 136 ship 0 32 209 57 217 63 208 40 200 ship 0 43 69 75 85 83 74 51 58 ship 0 222 70 250 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

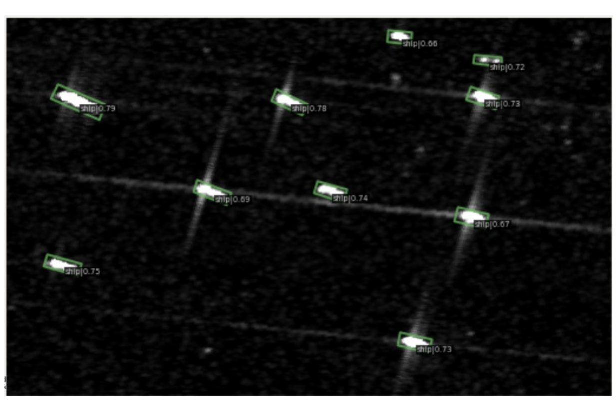

Fig. 4. Results of the test image: Oriented R-CNN

box prediction, using RoIAlignRotated to crop rotated features specifically for the region of interest (RoI).

Initially, we trained and evaluated the model using the SSDD Dataset. However, for our final proposal, we applied the Oriented RCNN model to the AU Drone Dataset.

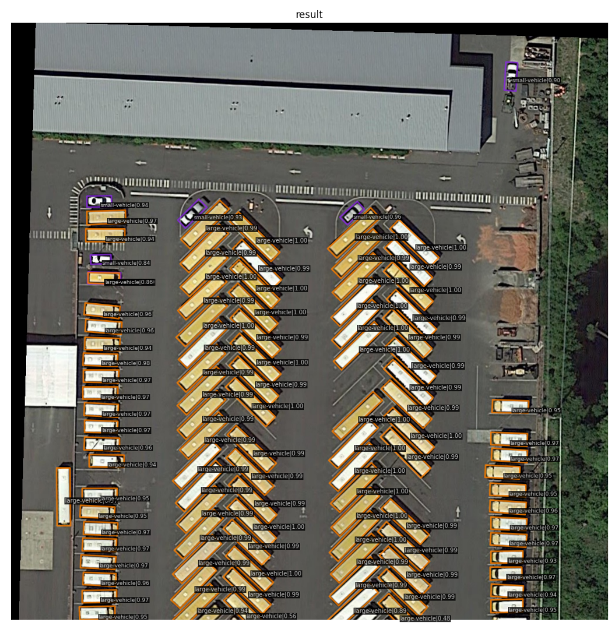

Fig. 5. Primary Results: Oriented RCNN

C. Model Comparison

We conducted a thorough comparison of model's performance on the SSDD dataset to gauge their generalization and effectiveness. To ensure a conducive experimental setup, we verified CUDA and GCC versions, installed required dependencies, and obtained pre-trained checkpoints. We then configured and evaluated the detector on sample images, offering insights into their effectiveness and suitability for oriented object detection tasks.

Subsequently, we aimed to enhance the performance of the Oriented RCNN model on our AU Drone dataset through rigorous training and detailed result analysis.

D. Performance Analysis

The evaluation of object detection accuracy involved assessing localization accuracy (Intersection over Union, IoU), classification accuracy, and the model's ability to handle oriented bounding boxes (OBB) in aerial images. Computational efficiency metrics such as inference time and memory usage were also analyzed to gauge the model's suitability for real-time applications.

A table was created to display performance metrics including mean Average Precision (mAP), accuracy,

Fig. 6. AU Drone dataset day 9 result

Fig. 7. AU Drone dataset day 10 result

and other relevant measures. The final outcome of the Oriented RCNN model on the AU Drone dataset is summarized in this table, showcasing an impressive accuracy of 94.6289 (nearly 95%).

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

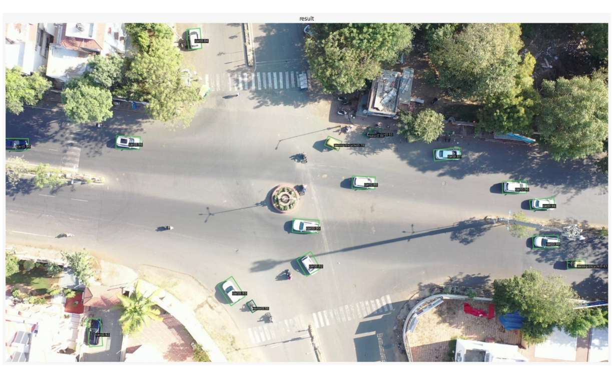

Fig. 8. AU Drone dataset day 21 result

Class	**GTS**	**DETS**	**Recall**	**AP**
awning-tricycle	89	323	0.989	0.836
motor	334	493	0.731	0.635
people	4	0	0.000	0.000
truck	1	0	0.000	0.000
car	177	286	1.000	0.993
bicycle	10	0	0.000	0.000
van	6	0	0.000	0.000
tricycle	4	0	0.000	0.000
mAP				0.308

TABLE I
RESULT OF ORIENTED RCNN ON AU DRONE DATASET

discussed based on the findings. We focused on using the Oriented R-CNN model as our base model and applied it over the AU Drone Dataset to yield final results.

F. Future Work

Future work includes exploring advanced OOD techniques, addressing model limitations, and improving the model performance over real-time performance.

III. DISCUSSIONS

A. Model Strengths and Weaknesses

- Oriented R-CNN typically offers superior accuracy in oriented object detection compared to some other approaches due to its multi-stage design and region-based methodology. However, its computational cost and resource requirements are high because of its multi-stage architecture and ROI alignment necessity.

B. Computational Efficiency

- As assessed in the base paper, Oriented R-CNN shows a speed that is comparable to one-stage detectors, yet its accuracy significantly surpasses them, marking it as a promising choice for oriented object detection.
- The Oriented RCNN model showcased strong performance on the AU Drone dataset and demonstrated higher accuracy compared to other models, establishing its effectiveness for aerial object detection tasks.

IV. RESULTS

In our implementations, we evaluated various models to gauge their effectiveness in oriented object detection across different datasets. The results indicated that Oriented R-CNN performed comparably on the SSDD dataset and more efficiently on the AU Drone Dataset. We used Pyplot to visually represent these findings, providing a clear depiction of the model's effectiveness. Additionally, we meticulously prepared the ground truth dataset and configured data loading specifically for evaluation purposes. Following a systematic inference process on the test dataset using MMrotate and MMDetection, we computed the model's mean Average Precision (mAP), providing quantitative insights into its efficacy. This comprehensive approach underscores the importance of conducting thorough evaluations when assessing how well models perform in tasks involving rotated object identification.

V. Conclusion

In summary, our efforts have revolved around delving into the complexities of model architectures for oriented bounding box (OBB) detection, aiming for performance enhancement. We concentrated on implementing existing models on both the DOTA dataset and the AU Drone dataset, gaining a deep understanding of their architectures. Our ultimate goal is to improve the effectiveness of oriented object detection systems, thereby ensuring safer and more efficient operations in real-time 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