[Project Proposal]

Rotated Object Detection: Innovative Applications Across Diverse Fields

Ashu Kumar, Neeraj Menon, Pooja Yadav, Pratik Sanghavi *

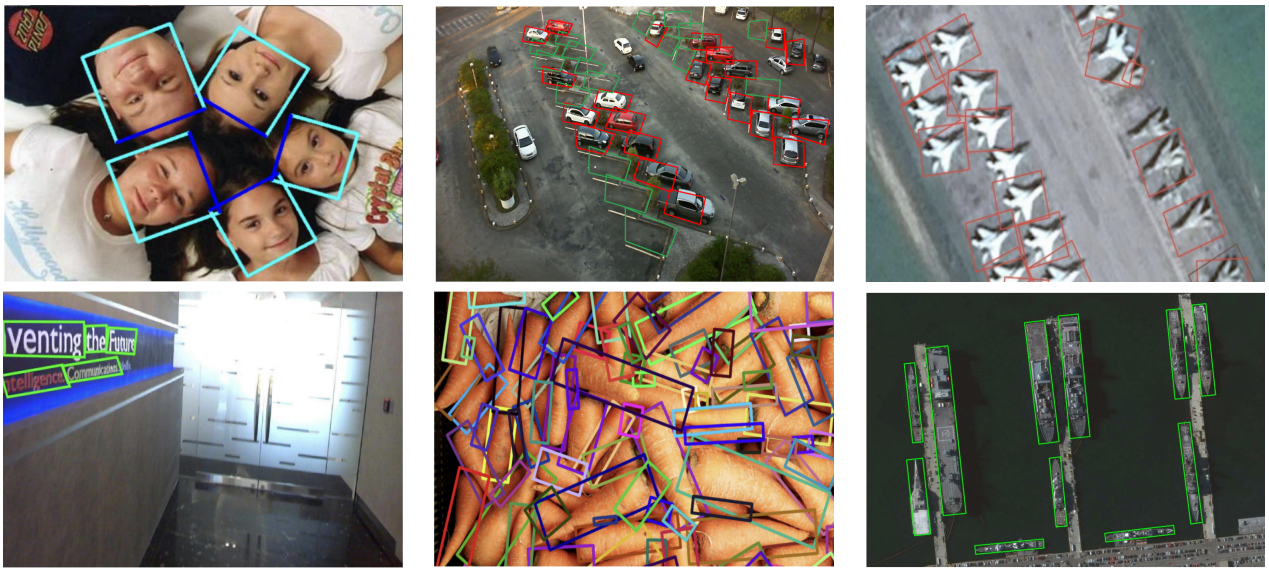

Figure 1: Various application areas of rotated object detection

What problem are you trying to solve?

Rotated object detection, also known as arbitrary-oriented object detection, aims to identify and locate objects in images with arbitrary orientation. In this task, the oriented directions of objects differ significantly across different images, with multiple orientations possible within a single image. This inherent complexity poses a challenge for standard backbone network models to extract high-quality features of these arbitrarily oriented objects. In addition, complex backgrounds, dense distributions, and heavy occlusions in images make accurate identification even more difficult.

In this project, we intend to address the aforementioned challenges by proposing a generalizable approach for accurate and efficient rotated object detection.

Why do you want to solve it?

In contrast to generic object detection, where object instances are assumed to be aligned with the image axes, objects within the natural scenes often exhibit arbitrary orientations. This is commonly observed in the field of remote sensing detection, text detection 'in the wild', face detection, retail scene detection, medical physics, industrial inspection, etc.

Despite the significant differences between the images with generic items and those with oriented objects, the design of conventional visual backbones has largely overlooked the inherent characteristics. As a result, the architecture of these models may be sub-optimal in the rotated object detection task.

*in alphabetical order

Therefore, in this project, we want to address the above challenges since rotated detection will strictly be superior to axis-aligned detection as there is less background within the box and the orientation can potentially convey object pose information.

Lastly, rotated object detection finds its application in various domains (Figure 1) such as autonomous vehicles, sports tracking and analysis, manufacturing and quality control, satellite image analysis, urban development planning, etc. This breadth of application underscores the significance of investing time and effort into addressing this problem.

What are the possible steps to the solution?

We plan to explore the following ways to solve the above problem:

Adaptive Rotated Convolution (ARC)

We take inspiration from [1]. The paper proposes a simple, yet effective module, wherein the convolution kernels rotate adaptively to extract object features with varying orientations in different images, and an efficient conditional computation mechanism is introduced to accommodate the large orientation variations of objects within an image. Experiment results verify that, equipped with the proposed module in the backbone network, the performance of various oriented object detectors improves significantly on commonly used rotated object detection benchmarks while remaining efficient.

In our project, we plan to plug this module into other vision backbones to boost their representation ability and evaluate their performance in detecting oriented objects accurately.

Knowledge Combination Without Rotated Annotation (KCR)

This idea comes from [2]. The paper proposes a framework that allows the model to precisely predict rotated boxes only requiring cheaper axis-aligned annotation of the target dataset. The framework combines task knowledge of an out-of-domain source dataset having stronger annotation with the domain knowledge of the target dataset having weaker annotation. It enables the model to perform on par with a fully-supervised model and works reasonably well under a large domain gap.

In our project, we want to carry out experiments to see if we can incorporate KCR into our approach.

Timetable

Timeline	**Milestone**	**Key Tasks**
Feb 23 - Mar 22	Midterm Report	- Understanding the methodology of shortlisted ideas - Finding resource requirements and coding env setup - Search and finalize suitable datasets - Carry out experiments
Mar 23 - Apr 15	Project Phase 2	- Continue experiments - Carry out performance evaluation - Draft final presentation
Apr 16 - May 3	Project Webpage	- Design webpage layout. - Include overview, methodologies, results, and conclusions. - Embed code snippets, demos, and source code links.

References

[1] Yifan Pu et al. "Adaptive Rotated Convolution for Rotated Object Detection". In: *arXiv preprint arXiv:2303.07820* (2023).

[2] Tianyu Zhu et al. "Knowledge combination to learn rotated detection without rotated annotation". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023, pp. 15518–15527.