Ahmedabad University

Evaluate performance of various OBB models [UAV videos]

Course Name: CSE641 - Computer Vision
Group Name: CTRL+ALT+DEL
Team Members :

- **Dhruv Patel** (AU2444003)
- **Garima Jain** (AU2444017)
- **Devanshi Rathod** (AU2444023)

Problem Statement

- This study evaluates the performance of various Oriented Bounding Box (OBB) models for object detection in UAV-captured videos.

- UAV-based object detection poses challenges such as occlusions, varying object orientations, and small object sizes, which can affect accuracy and reliability.

- The objective is to analyze and benchmark different OBB models to determine their effectiveness, strengths, and limitations in aerial image processing, enhancing detection performance in real-world scenarios.

Literature Survey

Paper Name & Author(s)	Algorithm Used	Methodology	Key Contribution
An Efficient Instance Segmentation Framework Based on Oriented Bounding Boxes Zhen Zhou et al.	CFNet	Uses box prompt-based segmentation models (e.g., SAM) and an Oriented Bounding Box (OBB) encoder.	Outperforms instance segmentation methods for occluded and dense objects. Reduces computational complexity via knowledge distillation.
Optimization for Arbitrary-Oriented Object Detection via Representation Invariance Loss Qi Ming et al.	Representation Invariance Loss (RIL)	Introduces Hungarian matching-based regression for OBBs. Uses normalized rotation loss to address angle discontinuity.	Improves detection accuracy in remote sensing datasets. Enhances loss function stability.
TricubeNet: 2D Kernel-Based Object Representation for Weakly-Occluded Oriented Object Detection Beomyoung Kim et al.	TricubeNet	Uses 2D Tricube Kernel representation instead of bounding box regression. Employs anchor-free detection for efficiency.	Solves angle discontinuity problem and reduces computational complexity.
Oriented Object Detection in Aerial Images With Box Boundary-Aware Vectors	BBAVectors	Uses keypoint-based detection. Predicts vectors for box boundaries instead of width, height, and angle.	Improves aerial image detection. Reduces imbalance issues between positive and negative anchors.
Gliding Vertex on the Horizontal Bounding Box for Multi-Oriented Object Detection	Gliding Vertex	Uses gliding vertex representation to align bounding boxes with object orientation. Minimal additional regression in Faster R-CNN.	Enhances accuracy for multi-oriented object detection with minimal computational overhead.

Dataset Discussion

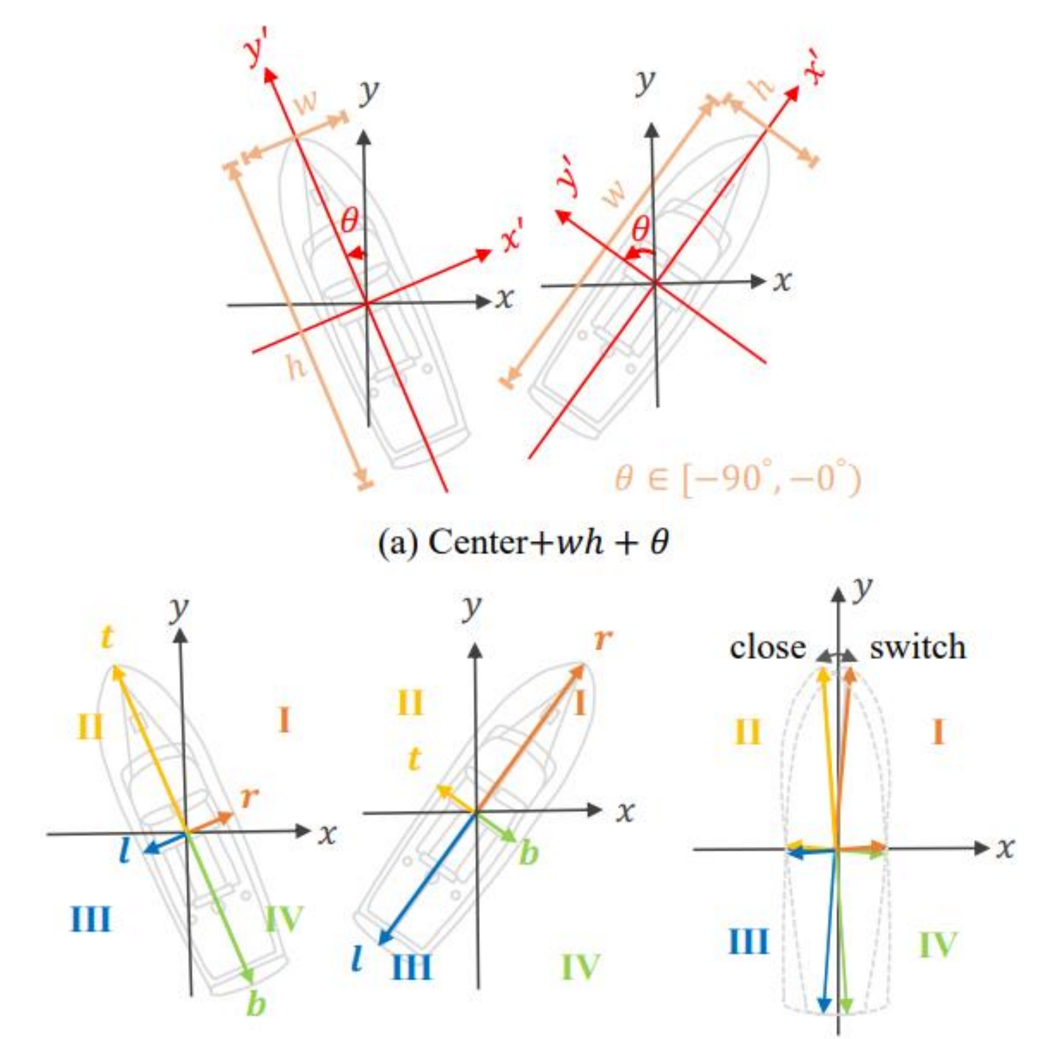
- **Dataset: DOTA v1.5** (aerial images with oriented object annotations)
- **Total images:** ~27,000 high-resolution satellite/UAV images
- **Objects categorized into 16 classes,** including:
  - Vehicles, ships, airplanes, bridges, and buildings
- **Annotation format:** Oriented bounding boxes (OBB) with (x, y, θ)
- **Dataset challenges:**
  - **High-resolution images require strong computational resources**
  - **Objects appear in different scales and orientations**
  - **Crowded object regions increase detection complexity**

Bounding box representation and challenges

(a) **Center + Width-Height + Angle Representation**
Represents an **oriented bounding box (OBB)** using **(x, y, w, h, θ)**, where **θ ∈ [-90°, 0°]** ensures consistency. Rotation transforms **(x, y)** ↔ **(x', y')**.

(b) **Box Boundary-Aware Vectors (BBAVectors)**
Defines an OBB using **four boundary-aware vectors (l, r, t, b)** for precise localization. Each vector corresponds to quadrants **I–IV**, improving accuracy over fixed bounding boxes.

(c) **Corner Cases Handling**
Addresses **extreme aspect ratios** and **boundary switching**. The **"close"** case manages near-overlapping boundaries, while the **"switch"** case corrects quadrant-flipping issues.



(a) Center+wh + θ

(b) Box boundary-aware vectors

(c) Corner cases

Ahmedabad University

Approaches

1. **Image Segmentation**

   Used **Mask R-CNN** & **YOLACT** for object segmentation in dataset.

2. **Oriented Bounding Box (OBB) Integration**

   Apply **OBBs** after segmentation to enhance localization.

   Improve detection under occlusion, varying orientations, and small object sizes.

3. **Model Evaluation**

   Test detection models: **TricubeNet, RetinaNet, Detectron2, EfficientDet**.

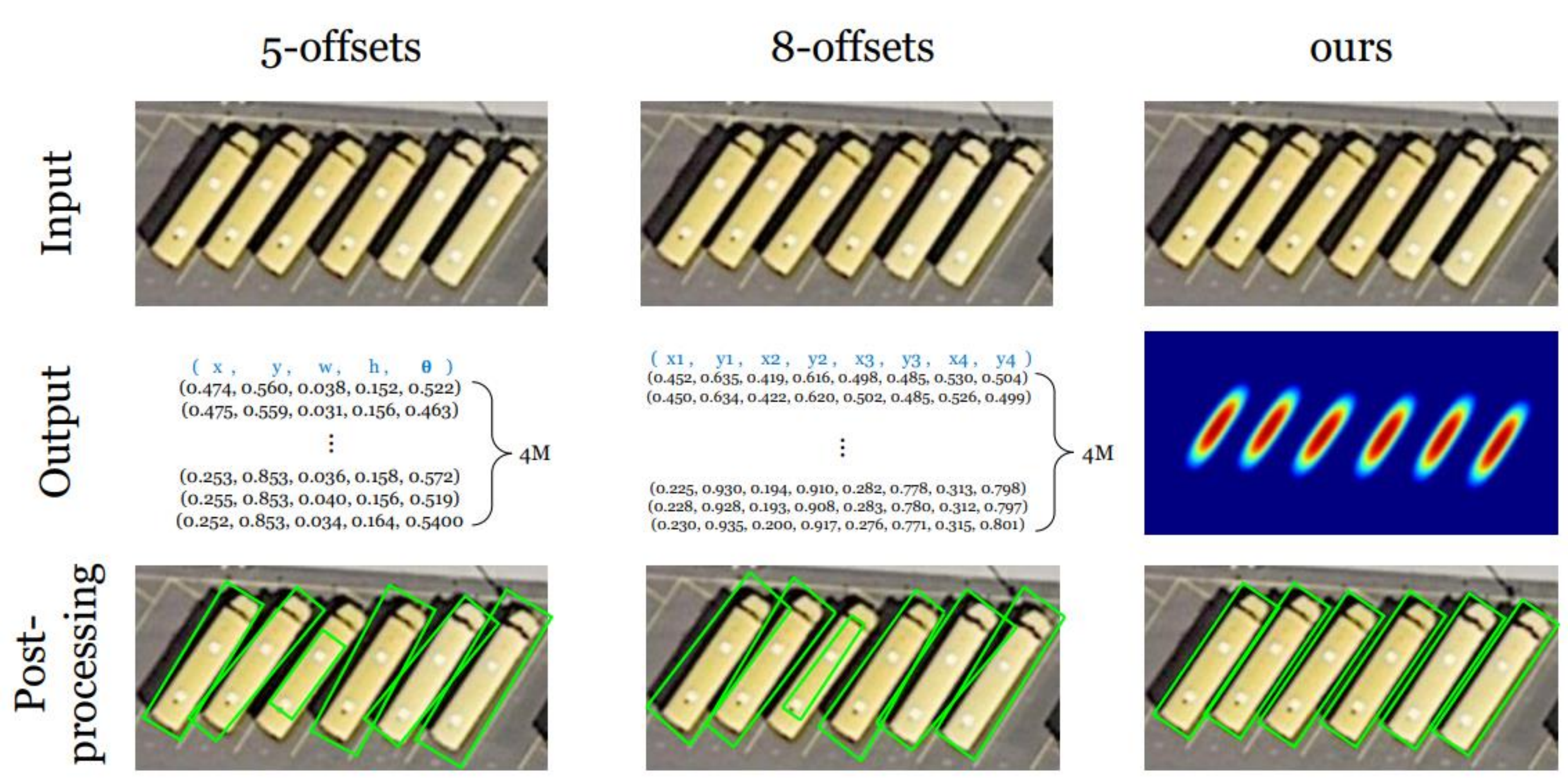
   Compare performance on **DOTA v1.5** dataset.

4. **Performance Comparison**

   Evaluate using mean average precision(**mAP),** Intersection of Union **(IoU), and inference speed**.

   Analyze trade-offs between accuracy & computational efficiency.

TricubeNet


5-offsets
8-offsets
ours
Input
Output
Post-processing
(x , y , w , h , θ)
(0.474, 0.560, 0.038, 0.152, 0.522)
(0.475, 0.559, 0.031, 0.156, 0.463)
(0.253, 0.853, 0.036, 0.158, 0.572)
(0.255, 0.853, 0.040, 0.156, 0.519)
(0.252, 0.853, 0.034, 0.164, 0.5400
4M
(x1 , y1 , x2 , y2 , x3 , y3 , x4 , y4)
(0.452, 0.635, 0.419, 0.616, 0.498, 0.485, 0.530, 0.504)
(0.450, 0.634, 0.422, 0.620, 0.502, 0.485, 0.526, 0.499)
(0.225, 0.930, 0.194, 0.910, 0.282, 0.778, 0.313, 0.798)
(0.228, 0.928, 0.193, 0.908, 0.283, 0.780, 0.312, 0.797)
(0.230, 0.935, 0.200, 0.917, 0.276, 0.771, 0.315, 0.801)
4M

Future Work

- **Framework Development**

  Build an open-source Python-based benchmarking framework for evaluating OBB models on UAV videos.

- **Enhanced Object Detection**

  Improve detection accuracy for small, occluded, and rotated objects in dataset. Further we will use **CFNet** and **SAM** for enhancing object segmentation and detection.

- **Model Optimization**

  Fine-tune models on DOTA v1.5/DOTA v2 datasets for better real-world performance.

- **Real-World Testing**

  Evaluate the models on diverse UAV video datasets to ensure robustness and scalability.

References

- Beomyoung Kim, Hyunseok Kim, Jeonghee Kim, Beomki Lee, and Youngjoon Yoo. **TricubeNet: 2D Kernel-Based Object Representation for Weakly-Occluded Oriented Object Detection.** *arXiv preprint*, arXiv:2104.11435, 2021.
- Qi Ming, Xuepeng Shi, Pei Xu, Jianqiang Huang, and Xiaolin Hu. **Optimization for Arbitrary-Oriented Object Detection via Representation Invariance Loss.** *arXiv preprint*, arXiv:2103.11636, 2021.
- Zhen Zhou, Qiyu Kang, Zizhao Shen, Fan Jin, and Song Bai. **CFNet: An Efficient Instance Segmentation Framework Based on Oriented Bounding Boxes.** *arXiv preprint*, arXiv:2401.08174, 2024.
- Jingru Yi, Wenjing Jia, and Dacheng Tao. **Oriented Object Detection in Aerial Images with Box Boundary-Aware Vectors.** In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 1406–1415, 2021.
- Yongchao Xu, Mingtao Fu, Qimeng Wang, Yukang Wang, Gui-Song Xia, Kai Chen, and Xiang Bai. **Gliding Vertex on the Horizontal Bounding Box for Multi-Oriented Object Detection.** *IEEE Transactions*, April 2021.