Project 6: Oriented Object Detection Model for Aerial Images

Hrishikesh Rana, Divya Patel, Aditya Chaudhari, Preksha Morbia
Ahmedabad University

Abstract

Oriented Object Detection is challenging, as the objects in aerial images are displayed in arbitrary directions. Current methods rely on two-stage anchor-based detectors, which typically suffer from severe imbalances between positive and negative anchor boxes. This model detects centre key points and regresses box boundary-aware vectors (BBAVectors) to capture oriented bounding boxes across the cartesian plane. This approach outperforms two-stage anchor-based detectors by directly predicting box parameters.

1. Keywords

Oriented object detection, Box boundary-aware vectors, Object bounding boxes, Key point-based detection, Anchor-free detection, Single-stage detector, Rotational object detection, Remote sensing, Arbitrary orientation, Center key point prediction, Heatmap, Offset prediction, Orientation classification, Corner case handling, DOTA dataset, Convolutional neural networks, Aerial imagery analysis

2. Introduction

Object detection in aerial images is crucial. However, it is difficult to do so due to the variance in textures, scales, complex backgrounds, and arbitrary object orientations. Horizontal bounding boxes don't align well with objects, prompting a shift to oriented bounding boxes. Current methods use two-stage anchor-based detectors, with the first stage proposing regions via anchor boxes and the second refining them. However, these methods need more design complexity, imbalance issues, and computational costs. Zhou's CenterNet proposes directly regressing width and height but faces challenges in jointly learning parameters for arbitrarily oriented objects.

3. Methodology

The paper uses pre-trained ResNet101 Conv1-5 as the foundation for this neural network that helps extract features from input images for accurate predictions.

First, we up-sample the image by increasing the size of extracted features to make it four times smaller, which is done to retain finer details.

The up-sampled feature map is refined through a 3 X 3 Conv layer. This refined feature map is concatenated with the shallow layer, followed by a 1x1 convolutional layer to refine channel-wise features.

The Output feature map is transformed into four branches:

I. Heat maps are used here to detect oriented objects' centre points, training via focal loss using 2D Gaussian from ground truth, which addresses imbalance.

II. Offset: An offset map is predicted to compensate for the difference between the quantified floating centre point (generated from a downscaling image)

Method	mAP	Plane	BD	Bridge	GTF	SV	LV	Ship	TC	BC	ST	SBF	RA	Harbor	SP	HC
YOLOv2 [18]	25.49	52.75	24.24	10.6	35.5	14.36	2.41	7.37	51.79	43.98	31.35	22.3	36.68	14.61	22.55	11.89
FR-O [22]	54.13	79.42	77.13	17.7	64.05	35.3	38.02	37.16	89.41	69.64	59.28	50.3	52.91	47.89	47.4	46.3
R-DFPN [24]	57.94	80.92	65.82	33.77	58.94	55.77	50.94	54.78	90.33	66.34	68.66	48.73	51.76	55.10	51.32	35.88
R^2CNN [7]	60.67	80.94	65.75	35.34	67.44	59.92	50.91	55.81	90.67	66.92	72.39	55.06	52.23	55.14	53.35	48.22
Yang *et al.* [25]	62.29	81.25	71.41	36.53	67.44	61.16	50.91	56.60	90.67	68.09	72.39	55.06	55.60	62.44	53.35	51.47
ICN [1]	68.16	81.36	74.30	47.70	70.32	64.89	67.82	69.98	90.76	79.06	78.20	53.64	**62.90**	**67.02**	64.17	50.23
ROI Trans. [2]	67.74	88.53	77.91	37.63	**74.08**	66.53	62.97	66.57	90.5	79.46	76.75	**59.04**	56.73	62.54	61.29	55.56
ROI Trans.+FPN [2]	69.56	**88.64**	78.52	43.44	**75.92**	68.81	73.68	83.59	90.74	77.27	81.46	**58.39**	53.54	62.83	58.93	47.67
BBAVectors+r	71.61	88.54	76.72	49.67	65.22	75.58	**80.28**	87.18	90.62	**84.94**	**84.89**	47.17	60.59	65.31	63.91	53.52
BBAVectors+rh	**72.32**	88.35	**79.96**	**50.69**	62.18	**78.43**	78.98	**87.94**	**90.85**	83.58	84.35	54.13	60.24	65.22	**64.28**	**55.70**
BBAVectors+rh^*	**75.36**	**88.63**	**84.06**	**52.13**	69.56	**78.26**	**80.40**	**88.06**	**90.87**	**87.23**	**86.39**	56.11	**65.62**	**67.10**	**72.08**	**63.96**

Table 1. Detection results on the testing set of DOTA-v1.0. The performances are evaluated through the online server. Symbol $*$ shows the result with a larger training batch size (i.e., 48 on 4 Quadro RTX 6000 GPUs). Red and Blue colors label the best and second best detection results in each column.

and the integer centre point (generated from heatmaps). This is done using smooth L1 loss with ground-truth offsets to ensure correct localisation.

III. Box-Parameters: The parameters for box boundary aware (BBA) vectors contain the top t, correct r, bottom b, and left l vectors, which are from the centre points of the objects. These four vectors are distributed in four quadrants of the cartesian coordinate system. This helps enhance mutual information sharing; it is trained via smooth L1 loss to regress BBA vectors and external size parameters for improved results.

IV. Orientation: Objects are categorised into two types of bounding boxes, HBB and RBB, which help address detection failures in corner cases by transforming them into horizontal ones that are easy to deal with.

4. Results

With the orientation classification and additional external Oriented Bounding Box (OBB) size parameters, our method achieves 72.32% accuracy. With the larger training batch size, it achieves 75.36% accuracy. Moreover, this method is robust enough to capture objects even in tiny, crowded vehicles.

Refer to Table 1.

5. Discussion:

We are facing issues with the accuracy of the bounding box formation, and our model cannot detect objects that are very small in size and very close to each other. Due to incompatible versions of dependencies in the model, it cannot run in the local environment.

6. Conclusion

The proposed oriented object detection method using box boundary-aware vectors and centre points detection is single-stage and anchor box-free, outperforming baseline methods on the DOTA dataset. Its simplicity and effectiveness make it a promising approach for future-oriented object detection tasks.

7. Reference:

- J. Yi, P. Wu, B. Liu, Q. Huang, H. Qu, and D. Metaxas, “Oriented Object Detection in Aerial Images with Box Boundary-Aware Vectors.” Accessed: Mar. 16, 2024. [Online]. Available: https://arxiv.org/pdf/2008.07043v2.pdf