MMRotate: A Rotated Object Detection Benchmark using PyTorch

Yue Zhou*
Department of EE, Shanghai Jiao Tong University
Shanghai, China
sjtu_zy@sjtu.edu.cn

Xue Yang*
MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University
Shanghai, China
yangxue-2019-sjtu@sjtu.edu.cn

Gefan Zhang
MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University
Shanghai, China
lizaozhouke@sjtu.edu.cn

Jiabao Wang
Northwestern Polytechnical University
Xi'an, China
jbwang@mail.nwpu.edu.cn

Yanyi Liu
Northeastern University
Shenyang, China
2001688@stu.neu.edu.cn

Liping Hou
University of Chinese Academy of Sciences
Beijing, China
houliping17@mails.ucas.ac.cn

Xue Jiang †
Department of EE, Shanghai Jiao Tong University
Shanghai, China
xuejiang@sjtu.edu.cn

Xingzhao Liu
Department of EE, Shanghai Jiao Tong University
Shanghai, China
xzhliu@sjtu.edu.cn

Junchi Yan †
MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University
Shanghai, China
yanjunchi@sjtu.edu.cn

Chengqi Lyu
Shanghai AI Laboratory
Shanghai, China
lvchengqi@pjlab.org.cn

Wenwei Zhang
Nanyang Technological University
Singapore, Singapore
wenwei001@ntu.edu.sg

Kai Chen
Shanghai AI Laboratory
SenseTime Research
Shanghai, China
chenkai@sensetime.com

ABSTRACT

We present an open-source toolbox, named MMRotate, which provides a coherent algorithm framework of training, inferring, and evaluation for the popular rotated object detection algorithm based on deep learning. MMRotate implements 18 state-of-the-art algorithms and supports the three most frequently used angle definition methods. To facilitate future research and industrial applications of rotated object detection-related problems, we also provide a large number of trained models and detailed benchmarks to give insights into the performance of rotated object detection. MMRotate is publicly released at https://github.com/open-mmlab/mmrotate.

* Equal contribution. †Correspondence author. Gefan Zhang is also with Cowa Robot, Co Ltd, Wuhu, Anhui, China.

MM '22, October 10–14, 2022, Lisboa, Portugal

ACM ISBN 978-1-4503-9203-7/22/10…$15.00
https://doi.org/10.1145/3503161.3548541

CCS CONCEPTS

• **Computing methodologies → Object detection**.

KEYWORDS

open source; rotation detection; oriented object detection

ACM Reference Format:
Yue Zhou*, Xue Yang*, Gefan Zhang, Jiabao Wang, Yanyi Liu, Liping Hou, Xue Jiang †, Xingzhao Liu, Junchi Yan †, Chengqi Lyu, Wenwei Zhang, and Kai Chen. 2022. MMRotate: A Rotated Object Detection Benchmark using PyTorch. In *Proceedings of the 30th ACM International Conference on Multimedia (MM '22), October 10–14, 2022, Lisboa, Portugal.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3503161.3548541

1 INTRODUCTION

In recent years, deep learning has achieved tremendous success in fundamental computer vision applications such as image recognition [8], object detection [4, 16, 25, 26, 28] and image segmentation [7, 19]. In light of this, deep learning has also been applied to areas such as faces detection [27], text detection [12, 13, 17, 20, 45] and aerial images detection [32, 35, 38]. In these object detection tasks, oriented bounding boxes (OBBs) are widely used instead of horizontal bounding boxes (HBBs) because they can better align the objects for more accurate identification. This kind of special object detection is called rotated object detection, also known as arbitrary-oriented object detection. In addition to the three applications mentioned above, rotated object detection is also widely used in 3D objects detection [44] and retail scenes detection [2, 23].

Different approaches utilize different angle definition methods, optimization strategies (e.g., optimizers, learning rate schedules, epoch numbers, and data augmentation pipelines), and CUDA operators (e.g. IoU and NMS for OBBs). To encompass the diversity of components used in various models, we have proposed the MMRotate toolbox covering recent popular rotated object detection approaches in a unified framework. The toolbox now implements 18 rotated object detection methods, 10 CUDA speed-up operators, and 12 losses. Integrating various algorithms confers code reusability and therefore dramatically simplifies the implementation of algorithms. Moreover, the unified framework allows different approaches to be compared against each other fairly and that their key effective components can be easily investigated. To the best of our knowledge, MMRotate supports most angle definition methods in various open source toolboxes, and it will facilitate future research on rotated object detection.

MMRotate is hosted on GitHub under the Apache-2.0 License. The repository contains the compressed archive file of software[1] and documentation, including installation instructions, dataset preparation scripts, API documentation[2], model zoo, tutorials and user manual. MMRotate re-implements 18 state-of-the-art rotated object detection algorithms and provides extensive benchmarks and models trained on popular academic datasets. In addition to (distributed) training and testing scripts, It offers a rich set of utility tools covering visualization and demonstration.

2 RELATED WORK

Text detection. Text detection aims to localize the bounding boxes of text instances. Recent research focus has shifted to challenging arbitrary-shaped text detection [7, 12, 14, 19, 20]. R^2CNN [12] simultaneously predicts the axis-aligned and inclined boxes by adding an inclined box branch and uses an inclined NMS to obtain the detection results. While Mask R-CNN [7] can be used to detect texts, it might fail to detect curved and dense texts due to the rectangle-based ROI proposals. On the other hand, RRPN [20] proposes a Rotation RPN to generate inclined proposals with text orientation angle information and project arbitrary-oriented proposals to the feature map with Rotation ROI pooling. TextSnake [19] describes text instances with a series of ordered, overlapping disks.

Aerial image detection. Aerial image detection plays a vital role in the military and attracts more and more attention in civilian field [3, 33, 34, 37]. It aims to predict more accurate bounding boxes and preserve the direction information of the object on aerial images (including ship, plane, vehicles, bridge, etc.). Although rotated object detection provides more accurate prediction results than horizontal detection, it requires defining a new bounding box representation. The most common is the θ-based representation (x, y, w, h, θ), and it adds an extra angle parameter based on the horizontal box. Depending on the angle range, it can be divided into OpenCV definition ($D_{oc}, \theta \in [-\pi/2, 0)$) [32, 36, 38], long edge 90° definition ($D_{le90}, \theta \in [-\pi/2, \pi/2)$) [3, 6], and long edge 135° definition ($D_{le135}, \theta \in [-\pi/4, 3\pi/4)$) [5]. Recent works used two-dimensional Gaussian distribution [37, 39] and point sets [22–24, 31] to represent objects, which have achieved excellent results. Feature alignment is another research direction of lifting rotated object detection performance. R^3Det [35] proposes a feature refinement module to re-construct the feature map based on the refined bounding box output from the previous stage. S^2A-Net [5] proposes an alignment convolution to alleviate the misalignment between axis-aligned convolutional features and arbitrary oriented objects. Recently, ReDet [6] began to study a novel rotation-equivariant RoI Align to produce rotation-equivariant features. Label Assignment is also a research hotspot. DAL [22] reconsiders whether IoU is a truly credible division basis and defines a new matching degree. SLA [21] proposes a sparse label assignment strategy to achieve training sample selection based on posterior IoU distribution. SASM [9] proposes two novel shape-adaptive strategies which can dynamically select samples and measure the quality of positive samples.

Table 1: Open source rotated object detection benchmarks.

Benchmark	AerialDet	JDet	OBBDet	AlphaRotate	MMRotate
DL library	PyTorch	Jittor	PyTorch	TensorFlow	PyTorch
Inference engine	PyTorch	Jittor	PyTorch	TensorFlow	PyTorch onnx runtime
OS	Linux	Windows Linux	Windows Linux	Linux	Windows Linux
Algorithm	5	8	9	16	18
Dataset	1	4	5	11	4
Doc.	-	-	-	✓	✓
Easy install	-	-	-	-	✓
Maintain	-	✓	✓	✓	✓

Open source toolbox. Several open source rotated object detection toolboxes have been developed over the years to meet the increasing demand from both academia and industry.

AerialDetection[3] is the pioneer of deep learning-based open source rotated object detection toolbox. It was publicly released in 2019, and provides evaluation tools for DOTA data set [29]. It supports five rotated object detection methods, e.g., Rotated RetinaNet [15], Rotated Faster R-CNN [26] and RoI Trans [3]. However, it has not been maintained anymore. OBBDetection[4] [30] is another popular open source oriented object detection toolbox. It supports 9 rotated object detection methods and provides a series of efficient processing tools for huge remote sensing images. AerialDetection and OBBDetection are both modified based on MMDetection[5] [1], which is a state-of-the-art open source object detection toolbox based on PyTorch. Nevertheless, they cannot enjoy the latest technology provided by MMDetection, since they rely on a specific old version of MMDetection. JDet[6] is an open source aerial image object detection toolbox based on a high-performance deep learning library [11]. It can be deployed on multiple platforms such as Linux and Windows, and the 8 detectors it reproduces have faster inference speed than PyTorch. TensorFlow-based rotated object detection toolbox AlphaRotate[7] [40] has been released recently. It

[1] https://github.com/open-mmlab/mmrotate/archive/refs/heads/main.zip
[2] https://mmrotate.readthedocs.io/en/latest/
[3] https://github.com/dingjiansw101/AerialDetection
[4] https://github.com/jbwang1997/OBBDetection
[5] https://github.com/open-mmlab/mmdetection
[6] https://github.com/Jittor/JDet
[7] https://github.com/yangxue0827/RotationDetection

Table 2: Accuracy comparison of rotated object detectors on DOTA v1.0. MS means multiple scale image split. RR means random rotation. All models are inferred with one 2080Ti GPU.

Baseline	**Technique**	fp16	Box Def.	Lr schd.	Mem.(GB)	Inf. time (fps)	Aug.	mAP
RetinaNet-H [15]	-	-	oc	1x	3.38	15.7	-	64.55
	GWD [37]	-	oc	1x	3.39	15.5	-	69.55
	KFIoU [41]	-	le90	1x	3.38	15.1	-	69.60
	KFIoU [41]	-	oc	1x	3.39	15.6	-	69.76
	KFIoU [41]	-	le135	1x	3.38	15.3	-	69.77
	KLD [39]	-	oc	1x	3.39	15.6	-	69.94
RetinaNet-O [15]	-	-	le90	1x	3.38	16.9	-	68.42
	-	✓	le90	1x	2.36	22.4	-	68.79
	CSL [33]	✓	le90	1x	2.60	24.0	-	69.51
	KLD [39]	-	le90	1x	3.35	16.9	-	70.22
	-	-	le135	1x	3.38	17.2	-	69.79
	ATSS [43]	-	le90	1x	3.12	18.2	-	70.64
	ATSS [43]	-	le135	1x	3.19	18.8	-	72.29
	-	-	le90	1x	3.78	17.5	MS+RR	76.50
RepPoints [42]	-	-	oc	1x	3.45	15.6	-	59.44
	SASM [9]	-	oc	1x	3.53	15.7	-	66.45
	G-Rep [10]	-	le135	1x	4.05	8.6	-	69.49
	CFA [23]	-	le135	1x	3.45	16.1	-	69.63
	CFA [23]	-	oc	40e	3.45	16.1	-	73.45
FCOS [28]	-	-	le90	1x	4.18	20.9	-	70.70
	CSL [33]	-	le90	1x	4.23	20.2	-	71.76
	KLD [39]	-	le90	1x	4.18	20.7	-	71.89
R^3Det [35]	-	-	oc	1x	3.54	12.4	-	69.80
	ATSS [43]	-	oc	1x	3.65	13.6	-	70.54
	KLD [39]	-	oc	1x	3.54	12.4	-	71.83
	KFIoU [41]	-	oc	1x	3.62	12.2	-	72.68
S^2ANet [5]	-	-	le135	1x	3.14	15.5	-	73.91
	-	✓	le135	1x	2.17	17.4	-	74.19
Faster RCNN [26]	-	-	le90	1x	8.46	16.5	-	73.40
	Gliding Vertex [31]	-	le90	1x	8.45	16.4	-	73.23
	Oriented RCNN [30]	✓	le90	1x	7.37	21.2	-	75.63
	Oriented RCNN [30]	-	le90	1x	8.46	16.2	-	75.69
	RoI Trans. [3]	✓	le90	1x	7.56	19.3	-	75.75
	ReDet [6]	✓	le90	1x	7.71	13.3	-	75.99
	RoI Trans. [3]	-	le90	1x	8.67	14.4	-	76.08
	ReDet [6]	-	le90	1x	9.32	10.9	-	76.68
	RoI Trans. [3] + Swin-T [18]	-	le90	1x	9.23	10.9	-	77.51
	RoI Trans. [3]	-	le90	1x	8.96	14.4	MS+RR	79.66
	ReDet [6]	-	le90	1x	9.63	10.9	MS+RR	79.87
	RoI Trans. [3] + KLD [39] + Swin-T [18]	-	le90	1x	12.3	10.9	MS+RR	80.90

currently implements 18 rotated object detection methods, including the algorithms that PyTorch and Jittor do not support, e.g., GWD [37], KLD [39], and R^3Det [35]. Comprehensive comparisons among these open source toolboxes are given in Table 1.

3 ROTATED OBJECT DETECTION STUDIES

Many important factors can affect the performance of deep learning-based detectors. This section investigates the angle definition method, backbones, and loss of network architectures. We exchange the above-mentioned components between different rotated object detection approaches to measure the performance, memory usage, and inference time.

Angle definition method. OpenCV definition, long edge 90° definition, and long edge 135° definition are all supported by MMRotate. All rotated object detection algorithms can easily switch between these three angle definition methods by modifying the configuration file. Meanwhile, we established an angle conversion API to facilitate other angle definition methods.

Backbone. ResNet50 [8] are commonly used in object detection approaches. To improve the accuracy, we also introduce a transformer-based backbone Swin transformer [18]. Table 2 compares ResNet50 and Swin-T in terms of memory, inference time and mAP by plugging them in RoI Trans. It has been shown that Swin-T significantly outperforms ResNet50, although its inference speed is 21% slower than that of ResNet50.

Loss. GWD, KLD, and KFIoU propose different loss to train the rotated object detector. Our experimental results in Table 2 show that the KLD loss achieved the best mAP in the OpenCV definition method when using RetinaNet as the baseline. However, when using the R^3Det as the baseline, the KFIoU loss achieved the best mAP in the OpenCV definition method.

Mixed precision training and Useful tools. All detectors in MMRotate support mixed precision training. Our experimental results in Table 2 show that the model trained with fp16 has a similar mAP as the original model. MMRotate also provides a range of efficient and convenient tools (including visualization, confusion matrix analysis, huge image inference), allowing researchers to focus on the rotated object detection algorithm itself.

4 CONCLUSIONS

With the practical importance and academic emergence for visual rotation detection, MMRotate is a deep learning benchmark for visual object rotation detection in PyTorch under the Apache-2.0 license. The architecture is designated for flexibility and ease of use to facilitate the deployment of rotated object detection in diverse domains, both in industrial applications and academic research. We will continue to improve the entire optimized benchmark and support representative detection methods in the future. We also welcome the community to participate in the development.

ACKNOWLEDGMENTS

This work was supported by National Key Research and Development Program of China (2020AAA0107600), and National Natural Science Foundation of China (61971279, 61972250, 62022054).

REFERENCES

[1] Kai Chen, Jiaqi Wang, Jiangmiao Pang, Yuhang Cao, Yu Xiong, Xiaoxiao Li, Shuyang Sun, Wansen Feng, Ziwei Liu, Jiarui Xu, Zheng Zhang, Dazhi Cheng, Chenchen Zhu, Tianheng Cheng, Qijie Zhao, Buyu Li, Xin Lu, Rui Zhu, Yue Wu, Jifeng Dai, Jingdong Wang, Jianping Shi, Wanli Ouyang, Chen Change Loy, and Dahua Lin. 2019. MMDetection: Open MMLab Detection Toolbox and Benchmark. *arXiv preprint arXiv:1906.07155* (2019).

[2] Zhiming Chen, Kean Chen, Weiyao Lin, John See, Hui Yu, Yan Ke, and Cong Yang. 2020. Piou loss: Towards accurate oriented object detection in complex environments. In *European Conference on Computer Vision*. Springer, 195–211.

[3] Jian Ding, Nan Xue, Yang Long, Gui-Song Xia, and Qikai Lu. 2019. Learning RoI Transformer for Oriented Object Detection in Aerial Images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2849–2858.

[4] Ross Girshick. 2015. Fast r-cnn. In *Proceedings of the IEEE International Conference on Computer Vision*. 1440–1448.

[5] Jiaming Han, Jian Ding, Jie Li, and Gui-Song Xia. 2021. Align deep features for oriented object detection. *IEEE Transactions on Geoscience and Remote Sensing* (2021).

[6] Jiaming Han, Jian Ding, Nan Xue, and Gui-Song Xia. 2021. Redet: A rotation-equivariant detector for aerial object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2786–2795.

[7] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. 2017. Mask r-cnn. In *Proceedings of the IEEE International Conference on Computer Vision*. 2961–2969.

[8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 770–778.

[9] Liping Hou, Ke Lu, Jian Xue, and Yuqiu Li. 2022. Shape-Adaptive Selection and Measurement for Oriented Object Detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

[10] Liping Hou, Ke Lu, Xue Yang, Yuqiu Li, and Jian Xue. 2022. G-Rep: Gaussian Representation for Arbitrary-Oriented Object Detection. *arXiv preprint arXiv:2205.11796* (2022).

[11] Shi-Min Hu, Dun Liang, Guo-Ye Yang, Guo-Wei Yang, and Wen-Yang Zhou. 2020. Jittor: a novel deep learning framework with meta-operators and unified graph execution. *Science China Information Sciences* 63, 222103 (2020), 1–21.

[12] Yingying Jiang, Xiangyu Zhu, Xiaobing Wang, Shuli Yang, Wei Li, Hua Wang, Pei Fu, and Zhenbo Luo. 2017. R2CNN: rotational region CNN for orientation robust scene text detection. *arXiv preprint arXiv:1706.09579* (2017).

[13] Minghui Liao, Baoguang Shi, and Xiang Bai. 2018. Textboxes++: A single-shot oriented scene text detector. *IEEE Transactions on Image Processing* 27, 8 (2018), 3676–3690.

[14] Minghui Liao, Zhaoyi Wan, Cong Yao, Kai Chen, and Xiang Bai. 2020. Real-time scene text detection with differentiable binarization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 11474–11481.

[15] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. 2017. Focal loss for dense object detection. In *Proceedings of the IEEE International Conference on Computer Vision*. 2980–2988.

[16] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. 2016. Ssd: Single shot multibox detector. In *European Conference on Computer Vision*. Springer, 21–37.

[17] Xuebo Liu, Ding Liang, Shi Yan, Dagui Chen, Yu Qiao, and Junjie Yan. 2018. Fots: Fast oriented text spotting with a unified network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 5676–5685.

[18] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. 2021. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE International Conference on Computer Vision*.

[19] Shangbang Long, Jiaqiang Ruan, Wenjie Zhang, Xin He, Wenhao Wu, and Cong Yao. 2018. Textsnake: A flexible representation for detecting text of arbitrary shapes. In *European Conference on Computer Vision*. 20–36.

[20] Jianqi Ma, Weiyuan Shao, Hao Ye, Li Wang, Hong Wang, Yingbin Zheng, and Xiangyang Xue. 2018. Arbitrary-oriented scene text detection via rotation proposals. *IEEE Transactions on Multimedia* 20, 11 (2018), 3111–3122.

[21] Qi Ming, Lingjuan Miao, Zhiqiang Zhou, Junjie Song, and Xue Yang. 2021. Sparse Label Assignment for Oriented Object Detection in Aerial Images. *Remote Sensing* 13, 14 (2021), 2664.

[22] Qi Ming, Zhiqiang Zhou, Lingjuan Miao, Hongwei Zhang, and Linhao Li. 2021. Dynamic Anchor Learning for Arbitrary-Oriented Object Detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 2355–2363.

[23] Xingjia Pan, Yuqiang Ren, Kekai Sheng, Weiming Dong, Haolei Yuan, Xiaowei Guo, Chongyang Ma, and Changsheng Xu. 2020. Dynamic Refinement Network for Oriented and Densely Packed Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 11207–11216.

[24] Wen Qian, Xue Yang, Silong Peng, Junchi Yan, and Yue Guo. 2021. Learning Modulated Loss for Rotated Object Detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 2458–2466.

[25] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 779–788.

[26] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems*. 91–99.

[27] Xuepeng Shi, Shiguang Shan, Meina Kan, Shuzhe Wu, and Xilin Chen. 2018. Real-time rotation-invariant face detection with progressive calibration networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2295–2303.

[28] Zhi Tian, Chunhua Shen, Hao Chen, and Tong He. 2019. Fcos: Fully convolutional one-stage object detection. In *IEEE International Conference on Computer Vision*. 9627–9636.

[29] Gui-Song Xia, Xiang Bai, Jian Ding, Zhen Zhu, Serge Belongie, Jiebo Luo, Mihai Datcu, Marcello Pelillo, and Liangpei Zhang. 2018. DOTA: A large-scale dataset for object detection in aerial images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3974–3983.

[30] Xingxing Xie, Gong Cheng, Jiabao Wang, Xiwen Yao, and Junwei Han. 2021. Oriented R-CNN for Object Detection. In *Proceedings of the IEEE International Conference on Computer Vision*. 3520–3529.

[31] Yongchao Xu, Mingtao Fu, Qimeng Wang, Yukang Wang, Kai Chen, Gui-Song Xia, and Xiang Bai. 2020. Gliding vertex on the horizontal bounding box for multi-oriented object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, 4 (2020), 1452–1459.

[32] Xue Yang, Hao Sun, Kun Fu, Jirui Yang, Xian Sun, Menglong Yan, and Zhi Guo. 2018. Automatic ship detection in remote sensing images from google earth of complex scenes based on multiscale rotation dense feature pyramid networks. *Remote Sensing* 10, 1 (2018), 132.

[33] Xue Yang and Junchi Yan. 2020. Arbitrary-Oriented Object Detection with Circular Smooth Label. In *European Conference on Computer Vision*. Springer, 677–694.

[34] Xue Yang and Junchi Yan. 2022. On the arbitrary-oriented object detection: Classification based approaches revisited. *International Journal of Computer Vision* (2022), 1–26.

[35] Xue Yang, Junchi Yan, Ziming Feng, and Tao He. 2021. R3Det: Refined Single-Stage Detector with Feature Refinement for Rotating Object. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 3163–3171.

[36] Xue Yang, Junchi Yan, Wenlong Liao, Xiaokang Yang, Jin Tang, and Tao He. 2022. SCRDet++: Detecting Small, Cluttered and Rotated Objects via Instance-Level Feature Denoising and Rotation Loss Smoothing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022).

[37] Xue Yang, Junchi Yan, Qi Ming, Wentao Wang, Xiaopeng Zhang, and Qi Tian. 2021. Rethinking rotated object detection with gaussian wasserstein distance loss. In *International Conference on Machine Learning*. PMLR, 11830–11841.

[38] Xue Yang, Jirui Yang, Junchi Yan, Yue Zhang, Tengfei Zhang, Zhi Guo, Xian Sun, and Kun Fu. 2019. Scrdet: Towards more robust detection for small, cluttered and rotated objects. In *Proceedings of the IEEE International Conference on Computer Vision*. 8232–8241.

[39] Xue Yang, Xiaojiang Yang, Jirui Yang, Qi Ming, Wentao Wang, Qi Tian, and Junchi Yan. 2021. Learning high-precision bounding box for rotated object detection via kullback-leibler divergence. *Advances in Neural Information Processing Systems* 34 (2021).

[40] Xue Yang, Yue Zhou, and Junchi Yan. 2021. AlphaRotate: A Rotation Detection Benchmark using TensorFlow. *arXiv preprint arXiv:2111.06677* (2021).

[41] Xue Yang, Yue Zhou, Gefan Zhang, Jitui Yang, Wentao Wang, Junchi Yan, Xiaopeng Zhang, and Qi Tian. 2022. The KFIoU Loss for Rotated Object Detection. *arXiv preprint arXiv:2201.12558* (2022).

[42] Ze Yang, Shaohui Liu, Han Hu, Liwei Wang, and Stephen Lin. 2019. Reppoints: Point set representation for object detection. In *Proceedings of the IEEE International Conference on Computer Vision*. 9657–9666.

[43] Shifeng Zhang, Cheng Chi, Yongqiang Yao, Zhen Lei, and Stan Z Li. 2020. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection. In *IEEE Conference on Computer Vision and Pattern Recognition*. 9759–9768.

[44] Yu Zheng, Danyang Zhang, Sinan Xie, Jiwen Lu, and Jie Zhou. 2020. Rotation-Robust Intersection over Union for 3D Object Detection. In *European Conference on Computer Vision*. Springer, 464–480.

[45] Xinyu Zhou, Cong Yao, He Wen, Yuzhi Wang, Shuchang Zhou, Weiran He, and Jiajun Liang. 2017. East: an efficient and accurate scene text detector. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 5551–5560.