ICML
International Conference
On Machine Learning

Rethinking Rotated Object Detection with Gaussian Wasserstein Distance Loss

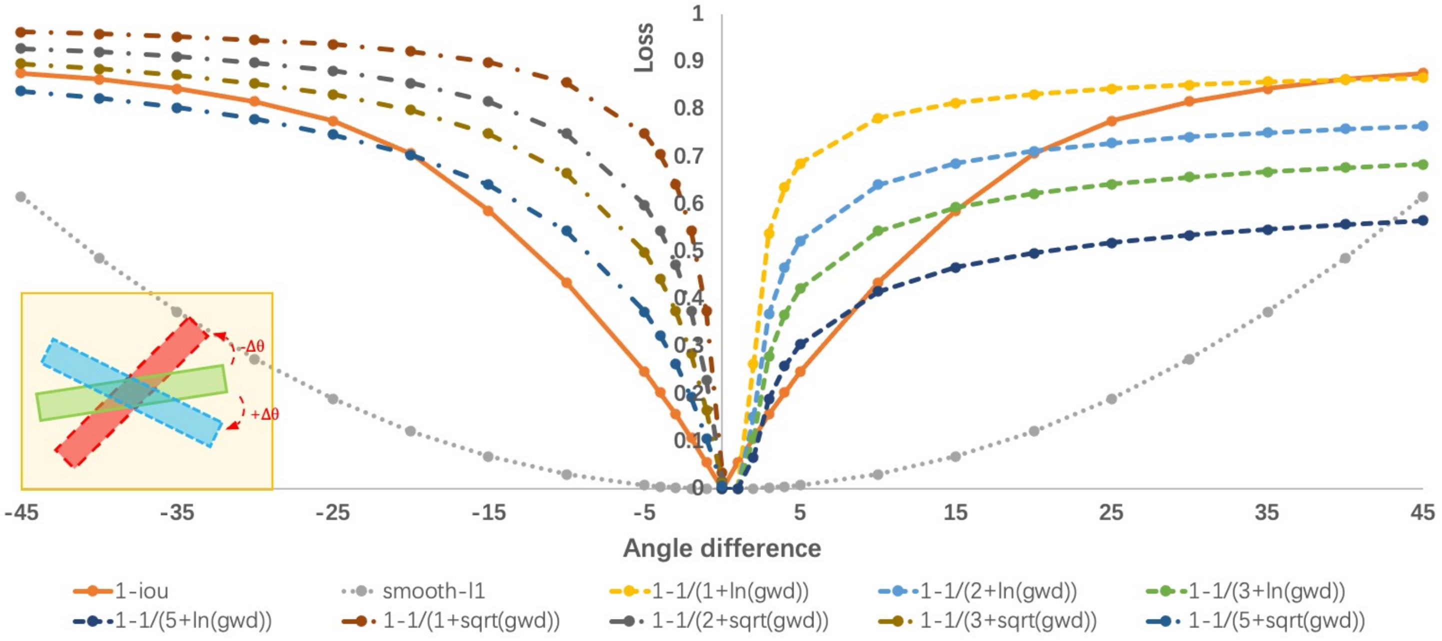
Xue Yang - Shanghai Jiao Tong University

X. Yang, et al. “Rethinking Rotated Object Detection with Gaussian Wasserstein Distance Loss.” In ICML21.

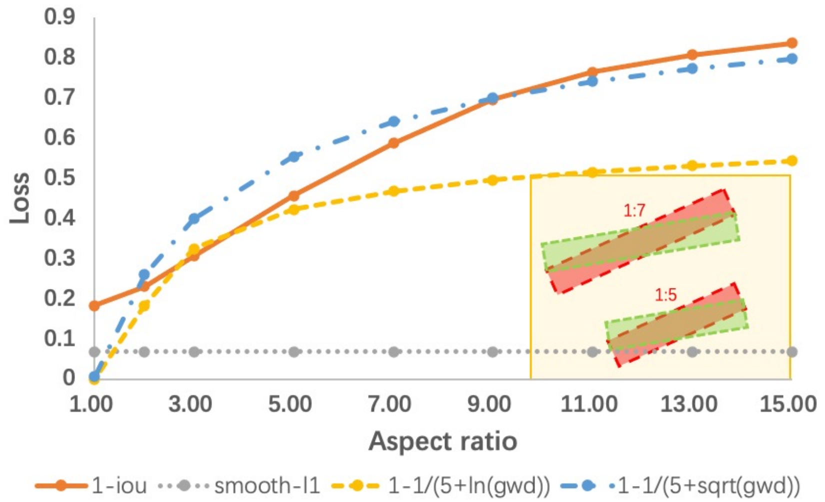
Virtual. 2021

Issues of Current Rotation Detector

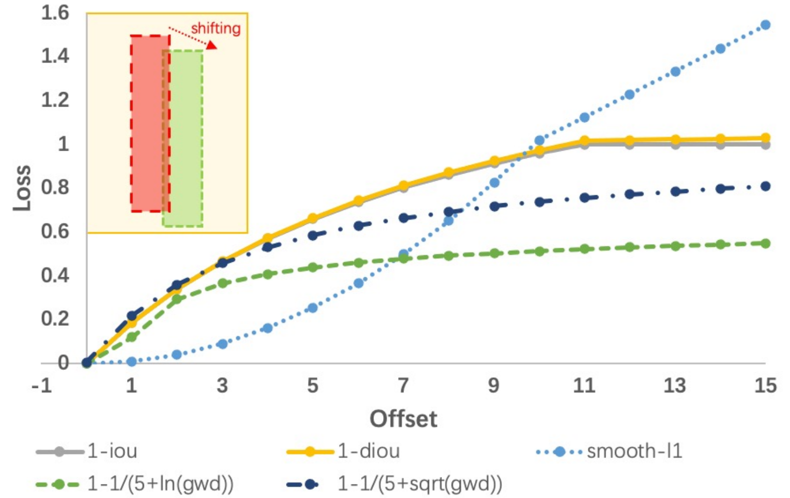
- **Issue 1:** Inconsistency between Metric and Loss



Case 1



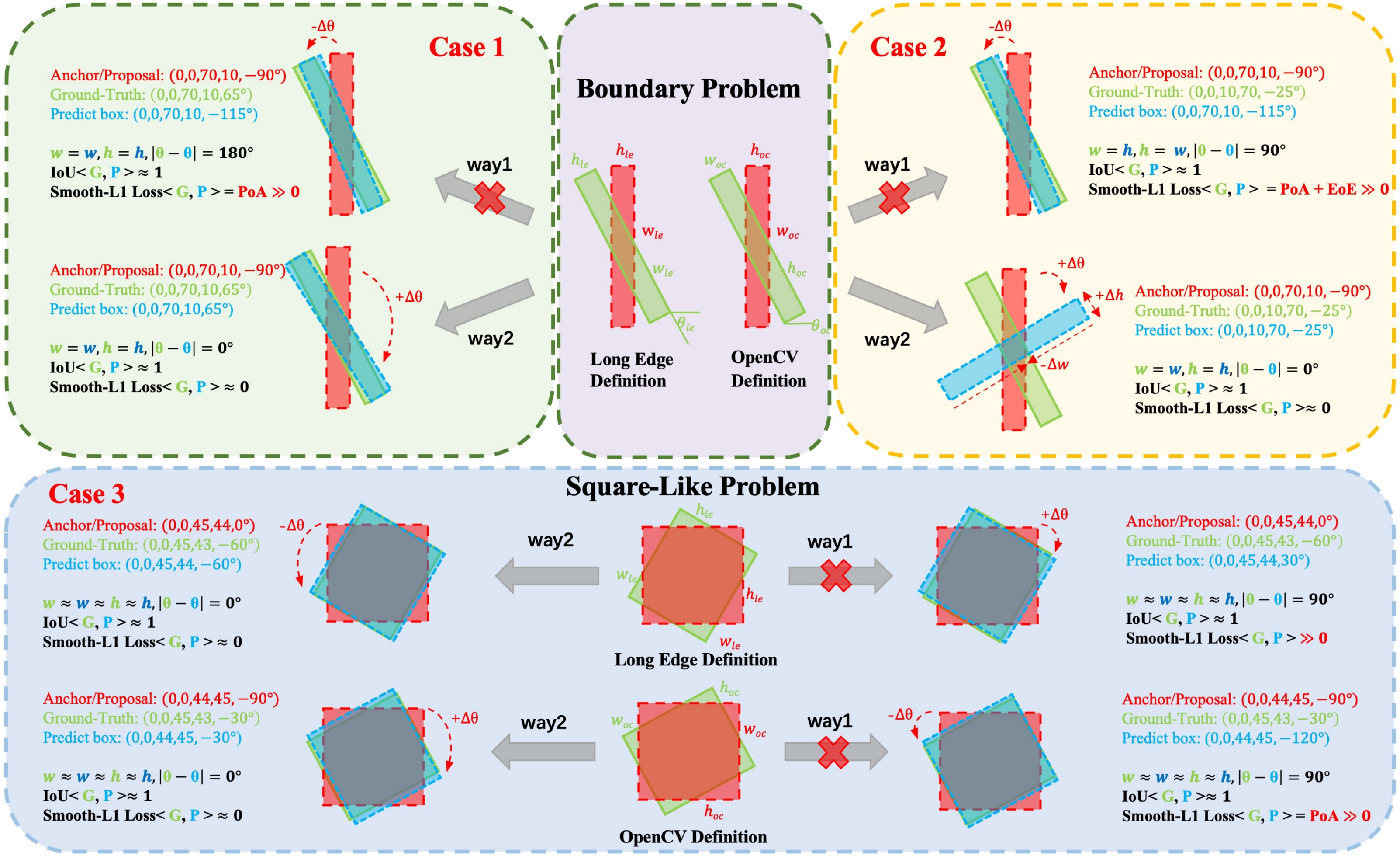
Case 2



Case 3

Issues of Current Rotation Detector

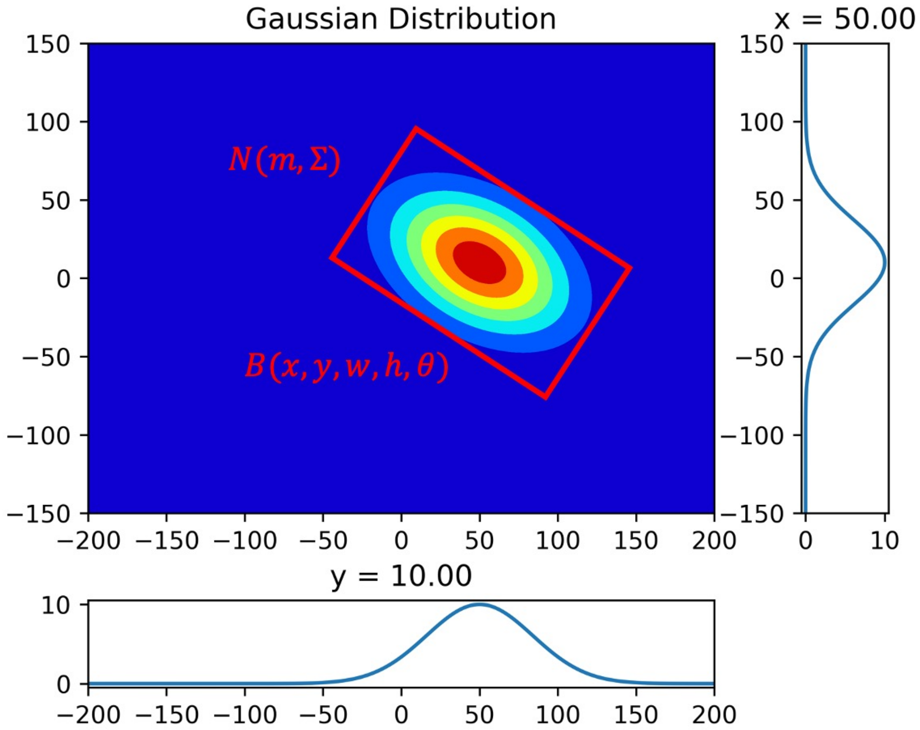
- **Issue 2 and 3:** Boundary Problem and Square-Like Problem



ICML International Conference On Machine Learning

Gaussian Wasserstein Distance

- **Requirement 1:** highly consistent with the IoU induced metrics (which also solves the square-like object problem
- **Requirement 2:** differentiable allowing for direct learning
- **Requirement 3:** smooth at angle boundary case



$$d^2 = \|\mathbf{m}_1 - \mathbf{m}_2\|_2^2 + \mathbf{Tr}\left(\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2 - 2(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2})^{1/2}\right)$$

Property 1: $\mathbf{\Sigma}^{1/2}(w,h,\theta) = \mathbf{\Sigma}^{1/2}(h,w,\theta-\frac{\pi}{2})$;

Property 2: $\mathbf{\Sigma}^{1/2}(w,h,\theta) = \mathbf{\Sigma}^{1/2}(w,h,\theta-\pi)$;

Property 3: $\mathbf{\Sigma}^{1/2}(w,h,\theta) \approx \mathbf{\Sigma}^{1/2}(w,h,\theta-\frac{\pi}{2})$, if $w \approx h$.

$$\begin{aligned}
\mathbf{\Sigma}^{1/2} =& \mathbf{RSR}^\top \\
=& \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \frac{w}{2} & 0 \\ 0 & \frac{h}{2} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\
=& \begin{pmatrix} \frac{w}{2}\cos^2\theta + \frac{h}{2}\sin^2\theta & \frac{w-h}{2}\cos\theta\sin\theta \\ \frac{w-h}{2}\cos\theta\sin\theta & \frac{w}{2}\sin^2\theta + \frac{h}{2}\cos^2\theta \end{pmatrix} \\
\mathbf{m} =& (x,y)^\top
\end{aligned}$$

ICML International Conference On Machine Learning

Experiments

Base Detector	Method	Box Def.	IML	BD EoE	BD PoA	SLP	TRANVAL/TEST BR†	SV†	LV†	SH†	HA†	ST‡	RA‡	7-mAP$_{50}$	mAP$_{50}$	TRAIN/VAL mAP$_{50}$	mAP$_{75}$	mAP$_{50:95}$
RetinaNet	-	D_{oc}	✓	✓	✓	×	42.17	65.93	51.11	72.61	53.24	78.38	62.00	60.78	65.73	64.70	32.31	34.50
	-	D_{le}	✓	✓	✓	✓	38.31	60.48	49.77	68.29	51.28	78.60	60.02	58.11	64.17	62.21	26.06	31.49
	IoU-Smooth L1 Loss	D_{oc}	✓	×	×	×	44.32	63.03	51.25	72.78	56.21	77.98	63.22	61.26	66.99	64.61	34.17	36.23
	Modulated Loss	D_{oc}	✓	×	×	×	42.92	67.92	52.91	72.67	53.64	80.22	58.21	61.21	66.05	63.50	33.32	34.61
	CSL	D_{le}	✓	×	×	✓	42.25	68.28	54.51	72.85	53.10	75.59	58.99	60.80	67.38	64.40	32.58	35.04
	DCL (BCL)	D_{le}	✓	×	×	×	41.40	65.82	56.27	73.80	54.30	79.02	60.25	61.55	67.39	65.93	35.66	36.71
	GWD	D_{oc}	×	×	×	×	44.07	71.92	62.56	77.94	60.25	79.64	63.52	65.70	68.93	65.44	38.68	38.71
R^{3}Det	-	D_{oc}	✓	✓	✓	×	44.15	75.09	72.88	86.04	56.49	82.53	61.01	68.31	70.66	67.18	38.41	38.46
	DCL (BCL)	D_{le}	✓	×	×	×	46.84	74.87	74.96	85.70	57.72	84.06	63.77	69.70	71.21	67.45	35.44	37.54
	GWD	D_{oc}	×	×	×	×	46.73	75.84	78.00	86.71	62.69	83.09	61.12	70.60	71.56	69.28	43.35	41.56

	Method	Backbone	MS	PL	BD	BR	GTF	SV	LV	SH	TC	BC	ST	SBF	RA	HA	SP	HC	mAP$_{50}$
Two-stage methods	ICN (Azimi et al., 2018)	R-101	✓	81.40	74.30	47.70	70.30	64.90	67.80	70.00	90.80	79.10	78.20	53.60	62.90	67.00	64.20	50.20	68.20
	RoI-Trans. (Ding et al., 2019)	R-101	✓	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	69.56
	CAD-Net (Zhang et al., 2019)	R-101		87.8	82.4	49.4	73.5	71.1	63.5	76.7	90.9	79.2	73.3	48.4	60.9	62.0	67.0	62.2	69.9
	SCRDet (Yang et al., 2019)	R-101	✓	89.98	80.65	52.09	68.36	68.36	60.32	72.41	90.85	87.94	86.86	65.02	66.68	66.25	68.24	65.21	72.61
	FADet (Li et al., 2019)	R-101	✓	90.21	79.58	45.49	76.41	73.18	68.27	79.56	90.83	83.40	84.68	53.40	65.42	74.17	69.69	64.86	73.28
	Gliding Vertex (Xu et al., 2020)	R-101		89.64	85.00	52.26	77.34	73.01	73.14	86.82	90.74	79.02	86.81	59.55	70.91	72.94	70.86	57.32	75.02
	Mask OBB (Wang et al., 2019)	RX-101	✓	89.56	85.95	54.21	72.90	76.52	74.16	85.63	89.85	83.81	86.48	54.89	69.64	73.94	69.06	63.32	75.33
	FFA (Fu et al., 2020)	R-101	✓	90.1	82.7	54.2	75.2	71.0	79.9	83.5	90.7	83.9	84.6	61.2	68.0	70.7	76.0	63.7	75.7
	APE (Zhu et al., 2020)	RX-101		89.96	83.62	53.42	76.03	74.01	77.16	79.45	90.83	87.15	84.51	67.72	60.33	74.61	71.84	65.55	75.75
	CenterMap (Wang et al., 2020a)	R-101	✓	89.83	84.41	54.60	70.25	77.66	78.32	87.19	90.66	84.89	85.27	56.46	69.23	74.13	71.56	66.06	76.03
	CSL (Yang & Yan, 2020)	R-152	✓	90.25	85.53	54.64	75.31	70.44	73.51	77.62	90.84	86.15	86.69	69.60	68.04	73.83	71.10	68.93	76.17
	RSDet-II (Qian et al., 2021)	R-152	✓	89.93	84.45	53.77	74.35	71.52	78.31	78.12	91.14	87.35	86.93	65.64	65.17	75.35	79.74	63.31	76.34
	SCRDet++ (Yang et al., 2020)	R-101	✓	90.05	84.39	55.44	73.99	77.54	71.11	86.05	90.67	87.32	87.08	69.62	68.90	73.74	71.29	65.08	76.81
Single-stage methods	PIoU (Chen et al., 2020)	DLA-34		80.9	69.7	24.1	60.2	38.3	64.4	64.8	90.9	77.2	70.4	46.5	37.1	57.1	61.9	64.0	60.5
	O^2-DNet (Wei et al., 2020)	H-104	✓	89.31	82.14	47.33	61.21	71.32	74.03	78.62	90.76	82.23	81.36	60.93	60.17	58.21	66.98	61.03	71.04
	P-RSDet (Zhou et al., 2020)	R-101	✓	88.58	77.83	50.44	69.29	71.10	75.79	78.66	90.88	80.10	81.71	57.92	63.03	66.30	69.77	63.13	72.30
	BBAVectors (Yi et al., 2020)	R-101	✓	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	72.32
	DRN (Pan et al., 2020)	H-104	✓	89.71	82.34	47.22	64.10	76.22	74.43	85.84	90.57	86.18	84.89	57.65	61.93	69.30	69.63	58.48	73.23
	R^{3}Det (Yang et al., 2021b)	R-152	✓	89.80	83.77	48.11	66.77	78.76	83.27	87.84	90.82	85.38	85.51	65.67	62.68	67.53	78.56	72.62	76.47
	PolarDet (Zhao et al., 2020)	R-101	✓	89.65	87.07	48.14	70.97	78.53	80.34	87.45	90.76	85.63	86.87	61.64	70.32	71.92	73.09	67.15	76.64
	S^{2}A-Net-DAL (Ming et al., 2020)	R-50	✓	89.69	83.11	55.03	71.00	78.30	81.90	88.46	90.89	84.97	87.46	64.41	65.65	76.86	72.09	64.35	76.95
	R^{3}Det-DCL (Yang et al., 2021a)	R-152	✓	89.26	83.60	53.54	72.76	79.04	82.56	87.31	90.67	86.59	86.98	67.49	66.88	73.29	70.56	69.99	77.37
	RDD (Zhong & Ao, 2020)	R-101	✓	89.15	83.92	52.51	73.06	77.81	79.00	87.08	90.62	86.72	87.15	63.96	70.29	76.98	75.79	72.15	77.75
	S^{2}A-Net (Han et al., 2021)	R-101	✓	89.28	84.11	56.95	79.21	80.18	82.93	89.21	90.86	84.66	87.61	71.66	68.23	78.58	78.20	65.55	79.15
	GWD (Ours)	R-152	✓	89.66	84.99	59.26	82.19	78.97	84.83	87.70	90.21	86.54	86.85	73.47	67.77	76.92	79.22	74.92	80.23

Thank You!

- Paper: https://arxiv.org/abs/2101.11952
- Code: https://github.com/yangxue0827/RotationDetection
- Contact:
  - Xue Yang: yangxue-2019-sjtu@sjtu.edu.cn
  - Junchi Yan: yanjunchi@sjtu.edu.cn
- Homepage of our Lab:
  - http://thinklab.sjtu.edu.cn/