关于旋转检测中高精度边界框的优化

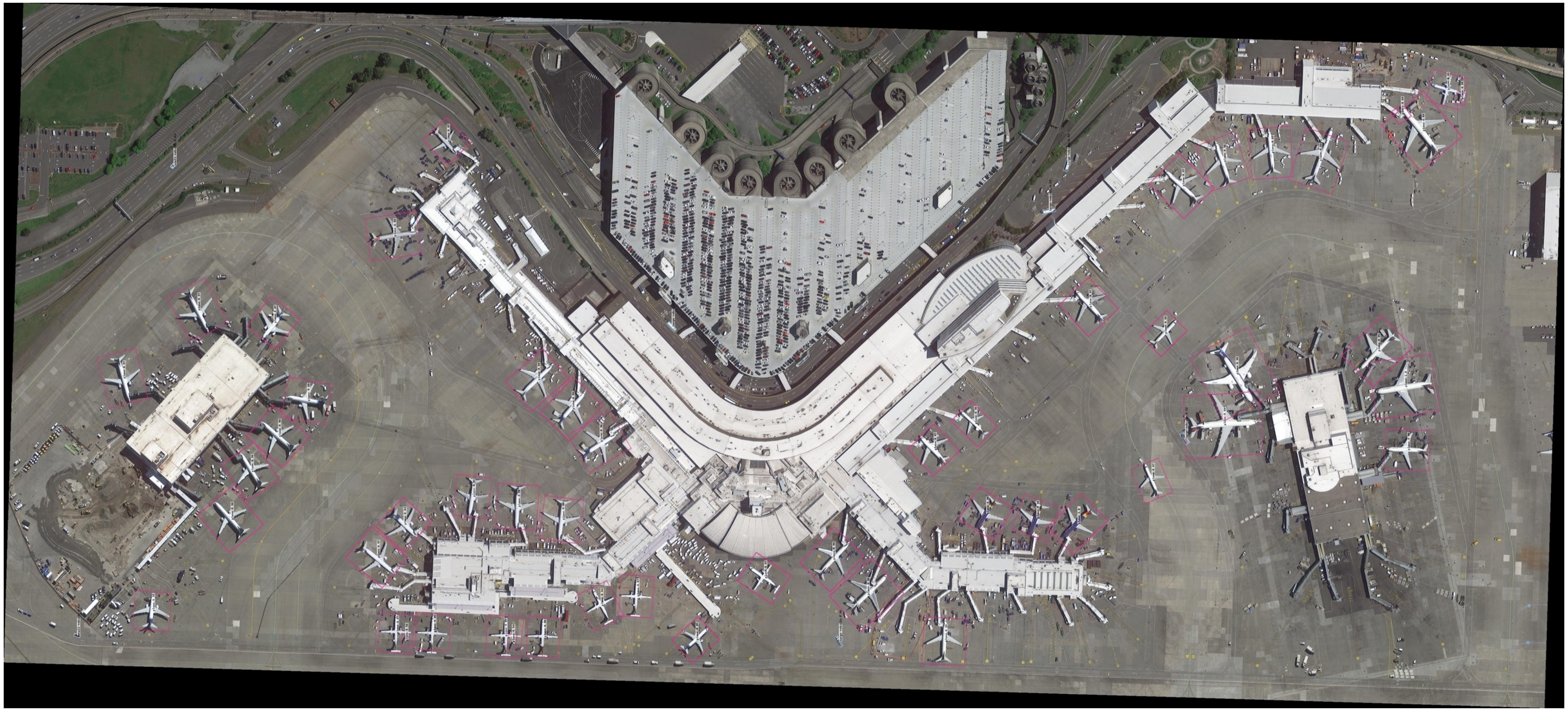
杨学

2021年5月

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

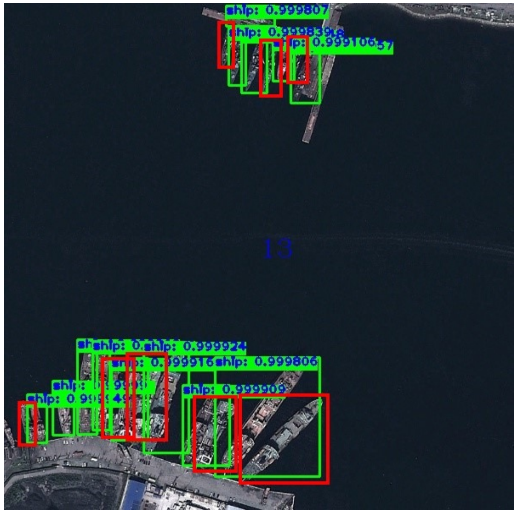
什么是旋转检测？

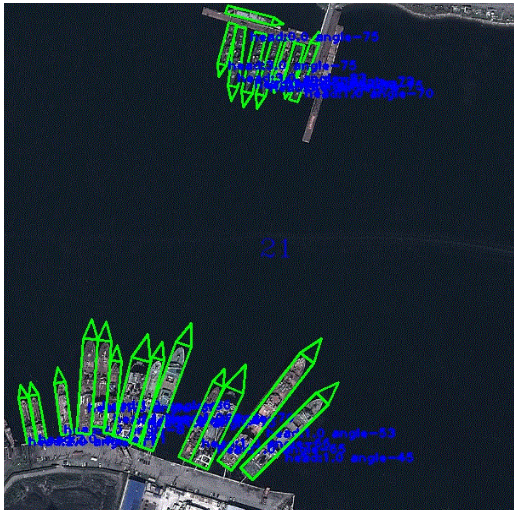
- 旋转检测是找到具有方向的边界框并对目标进行识别。



旋转目标检测的优势

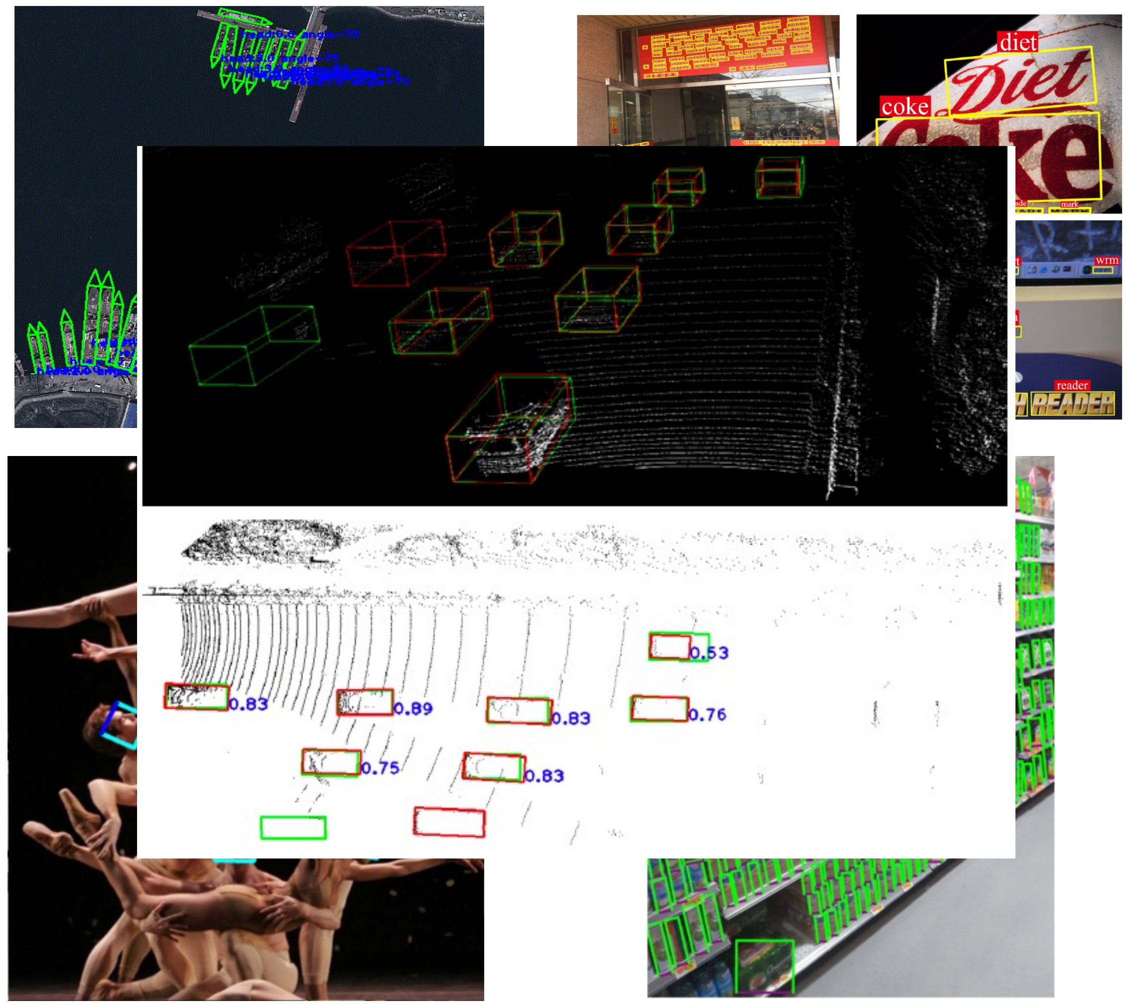
- 保留方向信息
- 适合密集场景，受后处理（NMS）影响小
- 检测结果背景区域占比小
- ……





旋转检测的应用场景

- 遥感检测
- 场景文字检测
- 人脸检测
- 零售场景检测
- 3D目标检测



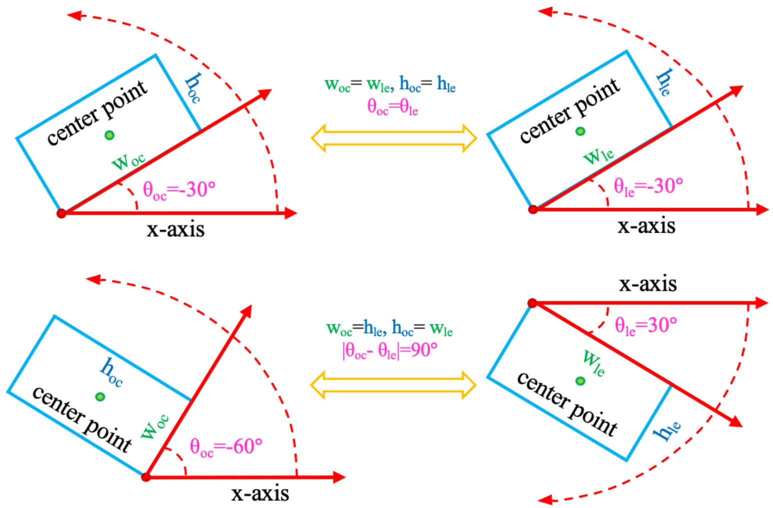
两种常见的旋转框定义

- **旋转框的定义方式：**
  - OpenCV定义法： $(x,y,w_{oc},h_{oc},\theta_{oc})$, $\theta_{oc} \in [-90,0)$
  - 长边定义法： $(x,y,w_{le},h_{le},\theta_{le})$, $\theta_{le} \in [-90,90)$

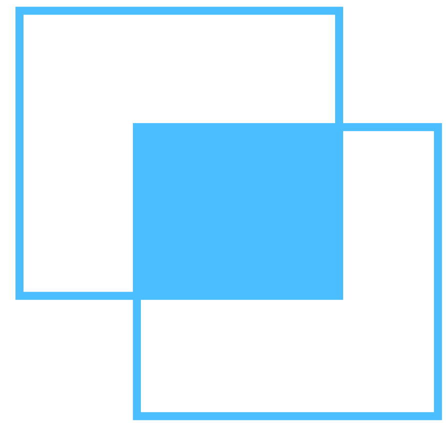
- **转换关系：**

$$D_{le}(w_{le},h_{le},\theta_{le}) = \begin{cases} D_{oc}(w_{oc},h_{oc},\theta_{oc}), & w_{oc} \geq h_{oc} \\ D_{oc}(h_{oc},w_{oc},\theta_{oc}+90^\circ), & otherwise \end{cases}$$

$$D_{oc}(w_{oc},h_{oc},\theta_{oc}) = \begin{cases} D_{le}(w_{le},h_{le},\theta_{le}), & \theta_{le} \in [-90^\circ,0^\circ) \\ D_{le}(h_{le},w_{le},\theta_{le}-90^\circ), & otherwise \end{cases}$$



旋转框IoU的计算

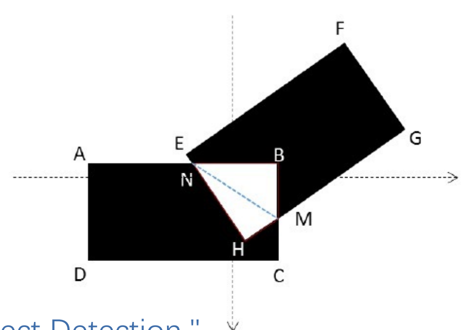


$$\mathrm{IoU} = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

旋转框IoU的计算

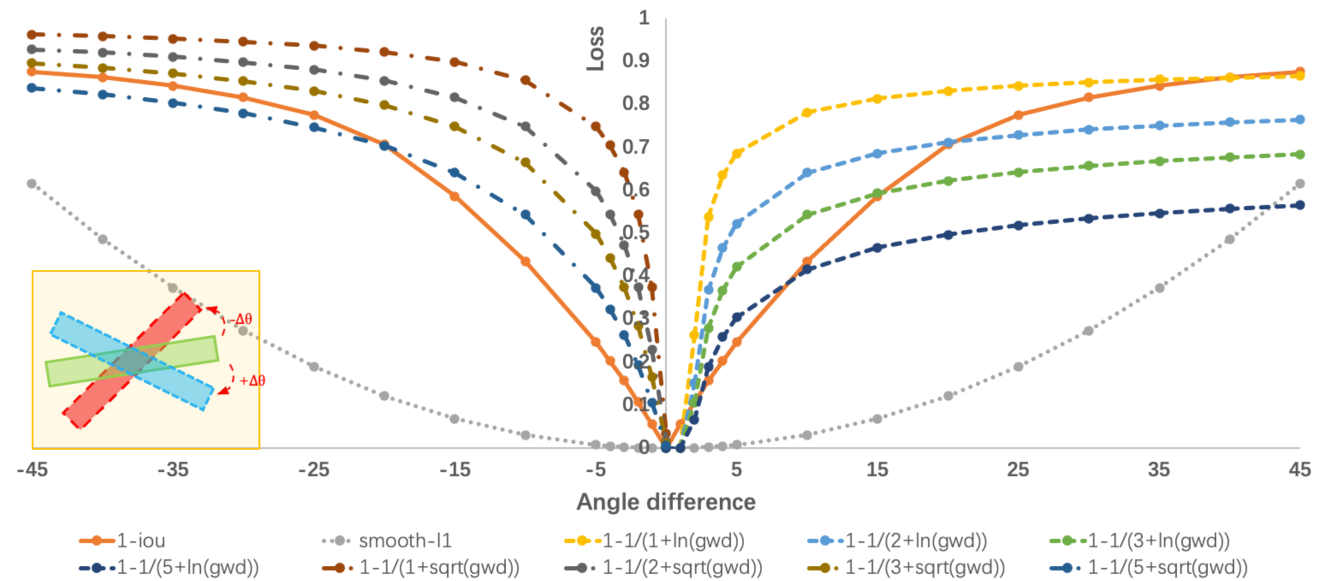
Algorithm 1 IoU computation

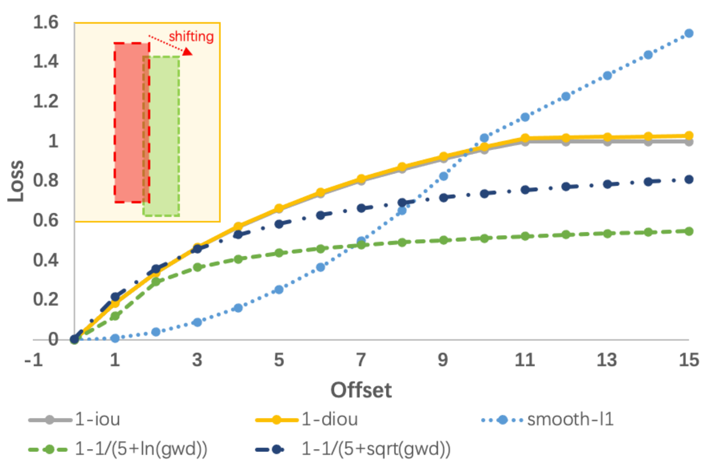
Input: Rectangles $R_1, R_2, ..., R_N$
IoU$[1, N][1, N] \leftarrow 0$
for each pair of R_i, R_j $(i < j)$ **do**
    Point set $PSet \leftarrow \emptyset$
    Add intersection points of R_i and R_j to $PSet$
    Add the vertices of R_i inside R_j into $PSet$
    Add the vertices of R_j inside R_i into $PSet$
    Sort $PSet$ to anti-clockwise order
    Compute intersection I of $PSet$ by triangulation
    IoU$(i, j) \leftarrow$ (Area(R_i) + Area(R_j) - I)/I
end for
return IoU

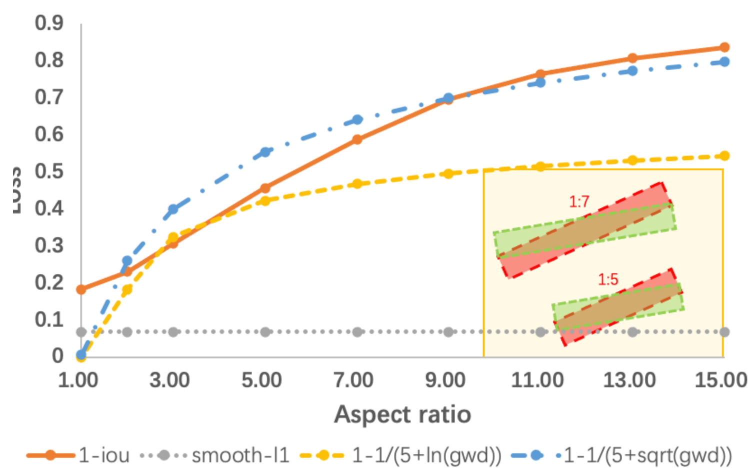


D. Zhou, et al. "IoU Loss for 2D/3D Object Detection.".
https://github.com/lilanxiao/Rotated_IoU

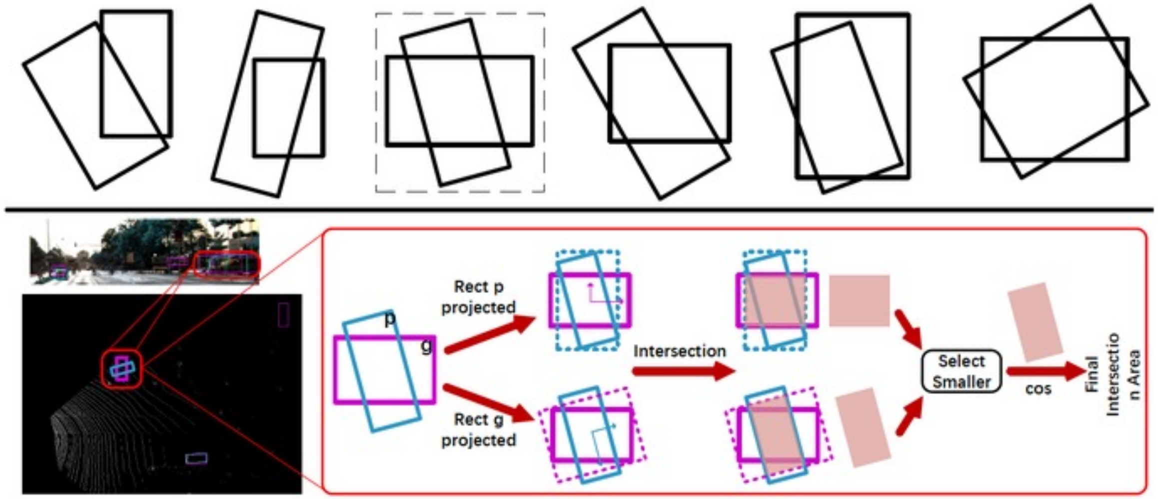
评估与损失不一致问题

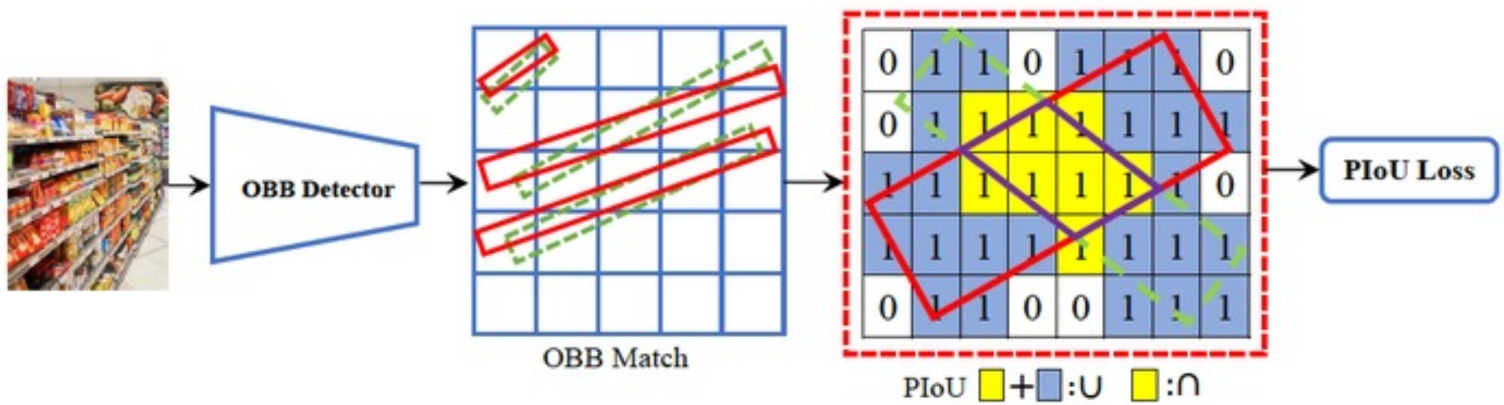

Loss
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-Δθ
+Δθ
-45
-35
-25
-15
-5
5
15
25
35
45
Angle difference
1-iou
smooth-l1
1-1/(1+ln(gwd))
1-1/(2+ln(gwd))
1-1/(3+ln(gwd))
1-1/(5+ln(gwd))
1-1/(1+sqrt(gwd))
1-1/(2+sqrt(gwd))
1-1/(3+sqrt(gwd))
1-1/(5+sqrt(gwd))


Loss
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
shifting
-1
1
3
5
7
9
11
13
15
Offset
1-iou
1-diou
smooth-l1
1-1/(5+ln(gwd))
1-1/(5+sqrt(gwd))


0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1:7
1:5
1.00
3.00
5.00
7.00
9.00
11.00
13.00
15.00
Aspect ratio
1-iou
smooth-l1
1-1/(5+ln(gwd))
1-1/(5+sqrt(gwd))

旋转框IoU的近似计算



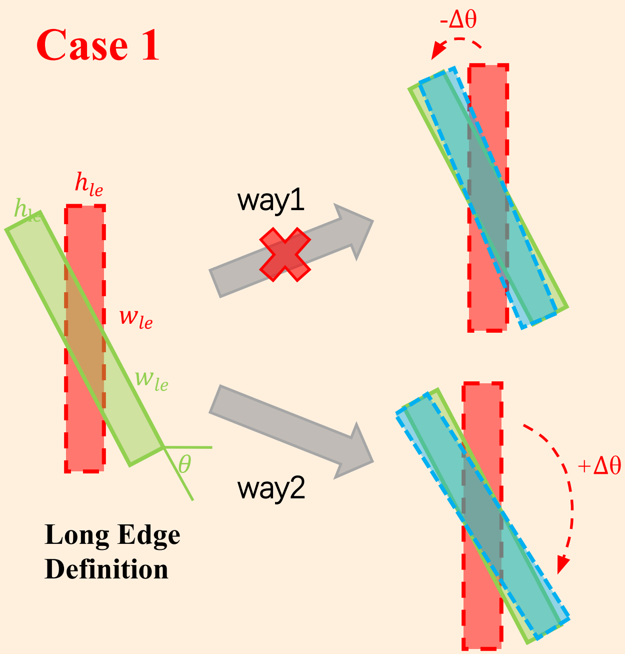


Z. Chen, et al. "PIoU Loss: Towards Accurate Oriented Object Detection in Complex Environments." In ECCV 2020.
Y. Zheng, et al. “Rotation-robust Intersection over Union for 3D Object Detection” In ECCV 2020.

问题和挑战

- **旋转检测中的边界问题**

Case 1



Long Edge Definition

长边定义法

Anchor/Proposal: (0,0,70,10, −90°)
Ground-Truth: (0,0,70,10,65°)
Predict box: (0,0,70,10, −115°)

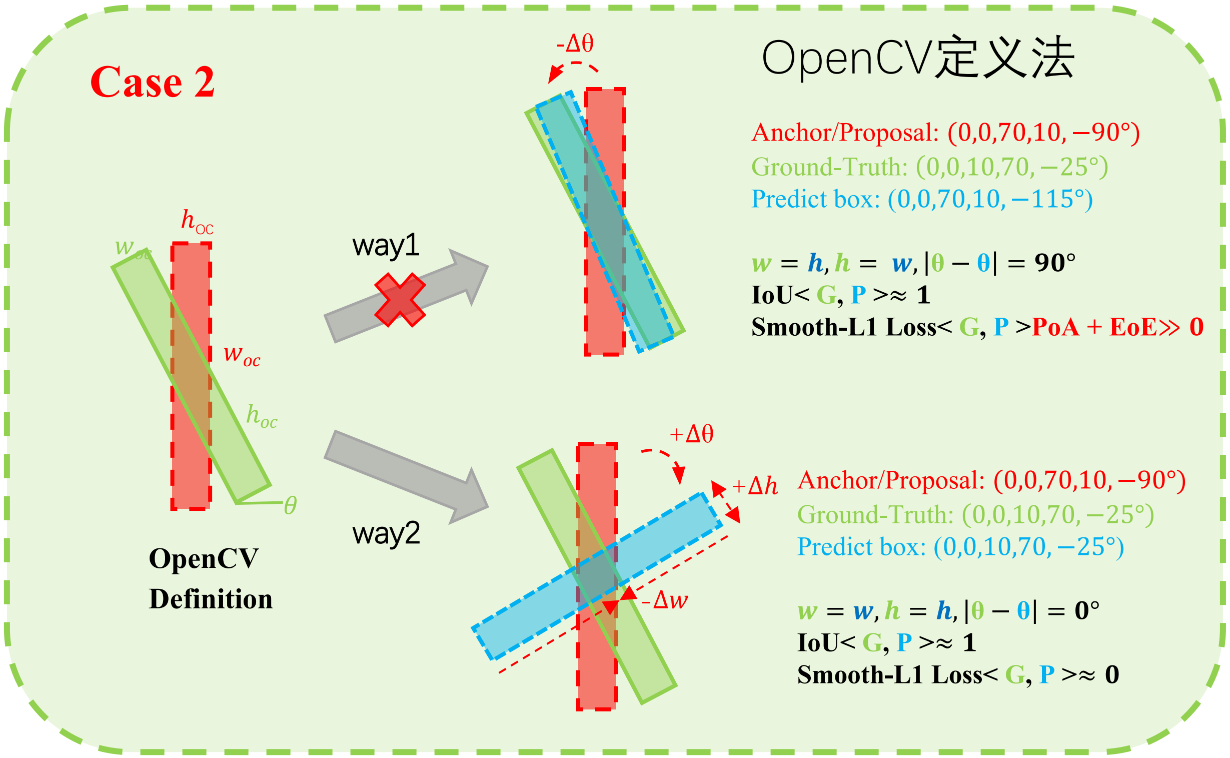
$w = w, h = h, |\theta - \theta| = 180^\circ$
IoU< G, P >≈ 1
Smooth-L1 Loss< G, P >PoA≫ 0

Anchor/Proposal: (0,0,70,10, −90°)
Ground-Truth: (0,0,70,10,65°)
Predict box: (0,0,70,10,65°)

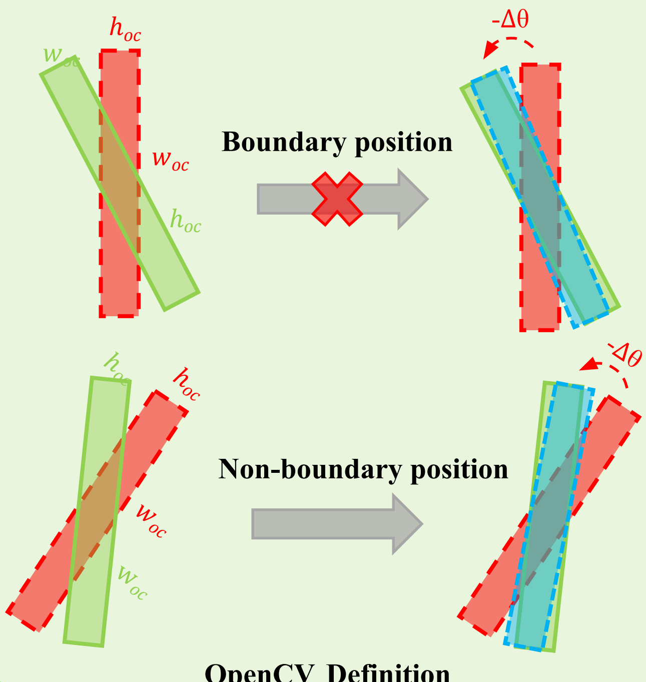
$w = w, h = h, |\theta - \theta| = 0^\circ$
IoU< G, P >≈ 1
Smooth-L1 Loss< G, P > ≈ 0

问题和挑战

- 旋转检测中的边界问题



旋转检测中的边界问题



OpenCV定义法

Anchor/Proposal: (0,0,70,10, −90°)
Ground-Truth: (0,0,10,70, −25°)
Predict box: (0,0,70,10, −115°)

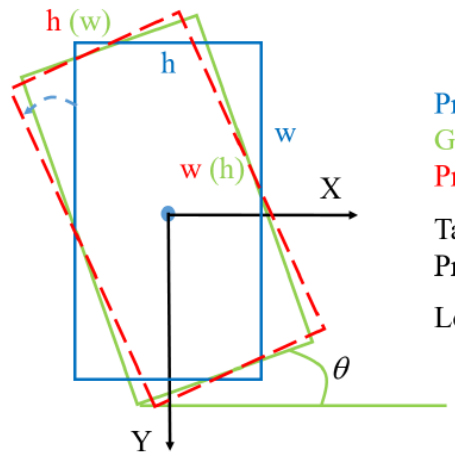
$w = h, h = w, |\theta - \theta| = 90°$
IoU< G, P >≈ 1
Smooth-L1 Loss< G, P > ≫ 0

Anchor/Proposal: (0,0,70,10, −55°)
Ground-Truth: (0,0,70,10, −80°)
Predict box: (0,0,70,10, −80°)

$w = h, h = w, |\theta - \theta| = 90°$
IoU< G, P >≈ 1
Smooth-L1 Loss< G, P > ≈ 0

IoU-Smooth L1 Loss

- 边界不连续性问题通常会使模型的损失值在边界情况下突然增加，主要原因如下：
  - periodicity of angular (PoA)
  - exchangeability of edges (EoE)
- **引入IoU常数因子，让IoU值决定loss的大小，消除边界问题。**

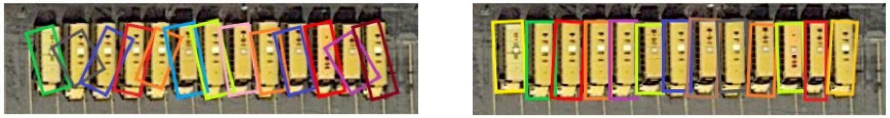


Proposal: (0, 0, 100, 25, -pi/2)
Ground-Truth: (0, 0, 25, 100, -pi/8)
Predict box: (0, 0, 100, 25, -5pi/8)

Target offset: (0, 0, log(1/4), log(4), 3pi/8)
Predict offset: (0, 0, 0, 0, -pi/8)

Loss=Smooth-L1(Predict - Target) >> 0

$$L = \frac{\lambda_1}{N}\sum_{n=1}^{N} t'_n \sum_{j\in\{x,y,w,h,\theta\}} \underbrace{\boxed{\frac{L_{reg}(v'_{nj}, v_{nj})}{|L_{reg}(v'_{nj}, v_{nj})|}}}_{\text{方向}} \underbrace{\boxed{|-\log(IoU)|}}_{\text{幅值}}$$



(a) Smooth L1 loss (b) IoU-smooth L1 loss

X. Yang, et al. "Scrdet: Towards more robust detection for small, cluttered and rotated objects." In ICCV 2019.

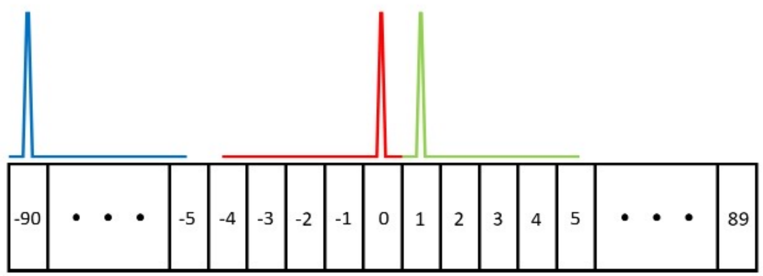
简单的角度分类

- 边界不连续性问题通常会使模型的损失值在边界情况下突然增加。
- 将目标角度的预测视为分类问题，以更好地限制预测结果。解决方案是将目标的角度作为类别标签，类别数与角度范围有关。
- 回归问题转换成分类问题实质是一个连续到离散的过程，中间存在理论精度的损失：

$$Max(loss) = \frac{\omega}{2}, \quad E(loss) = \int_a^b x * \frac{1}{b-a} dx = \int_0^{\omega/2} x * \frac{1}{\omega/2 - 0} dx = \frac{\omega}{4}$$

简单的角度分类

- 简单角度分类存在的问题：
  - 使用OpenCV定义法时EoE问题仍然存在（因此采用长边定义法）
  - 分类损失对于预测标签和标签之间的角度距离是不可感知的。

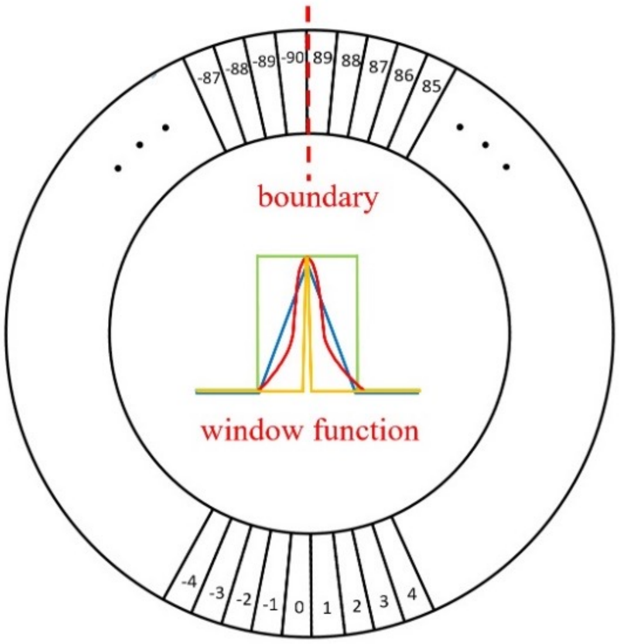


ground truth = one-hot(0)
predict1 ≈ one-hot(1)
predict2 ≈ one-hot(-90)
FL(gt – predict1) ≈ FL(gt – predict2) ✗

Circular Smooth Label (CSL)

- CSL是具有周期性的圆形标签编码，并且分配的标签值平滑且具有一定的容忍性

$$CSL(x) = \begin{cases} g(x), \theta - r < x < \theta + r \\ 0, \quad otherwise \end{cases}$$

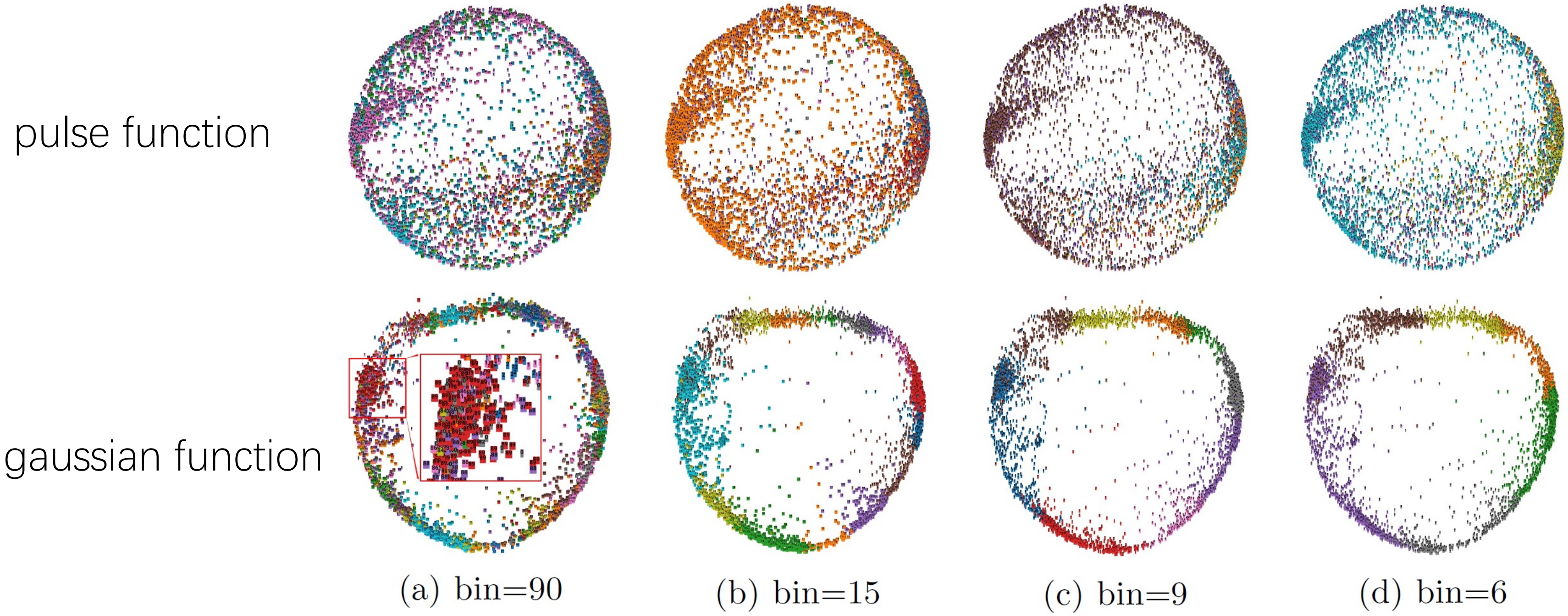


- 性质
  - 周期性 $g(x) = g(x + kT), k \in N. \; T = 180/\omega$
  - 对称性 $0 \le g(\theta + \varepsilon) = g(\theta - \varepsilon) \le 1, |\varepsilon| < r.$
  - 最大值 $g(\theta) = 1$
  - 单调性 $0 \le g(\theta \pm \varepsilon) \le g(\theta \pm \varsigma) \le 1, |\varsigma| < |\varepsilon| < r.$

X. Yang, J. Yan. “Arbitrary-Oriented Object Detection with Circular Smooth Label.” In ECCV 2020.

CSL的可视化

- DOTA数据集上检测器的角度特征可视化。 每个点代表测试集的RoI，以及其所属bin的索引。

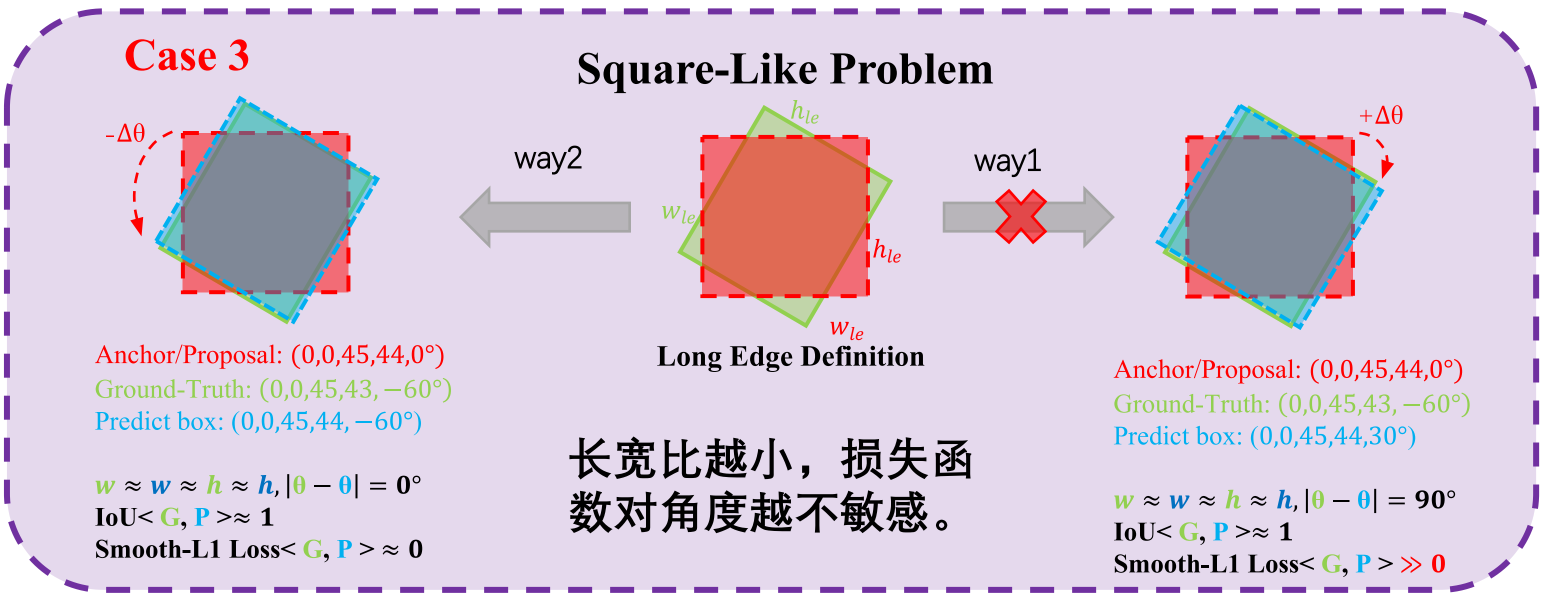


CSL存在的一些问题

- 问题1：厚重的检测头

$$\mathrm{Th}_{reg.} = A$$
$$\mathrm{Th}_{onehot} = \mathrm{Th}_{csl} = A \times AR/\omega$$

- 问题2：长边定义法的使用不利于类正方形目标的检测



Densely Coded Label (DCL)

- 使用密集编码（DCL）取代稀疏编码（SCL） （针对问题1）

$$\underbrace{\mathrm{Th}_{bcl} = \mathrm{Th}_{gcl}}_{\mathrm{TH}_{dcl}} = A \times \lceil \log_2(AR/\omega) \rceil$$

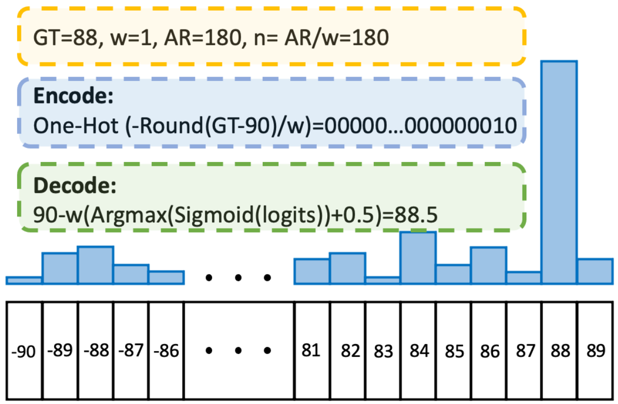
- 一个例子：A=21, AR=180, w=1
  - Th_{reg} = 21, Th_{onehot}=Th_{csl}=3780, Th_{dcl}=168

Base Model	ω	GFlops	ΔGFlops	Params (M)	ΔParams	Training Time
RetinaNet-Reg	-	139.35	-	36.97	-	-
RetinaNet-CSL	1	254.96	+82.96%	45.63	+23.42%	∼3x
RetinaNet-BCL	1	143.87	**+3.24%**	37.31	**+0.92%**	∼**1x**
RetinaNet-GCL	1	143.87	**+3.24%**	37.31	**+0.92%**	∼**1x**

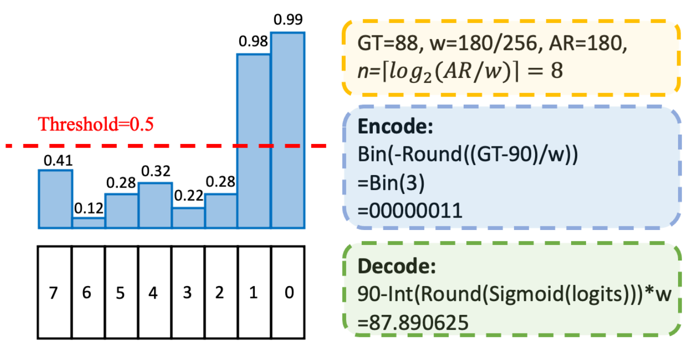
X. Yang, et al. "Dense Label Encoding for Boundary Discontinuity Free Rotation Detection." In CVPR 2021.

Densely Coded Label (DCL)

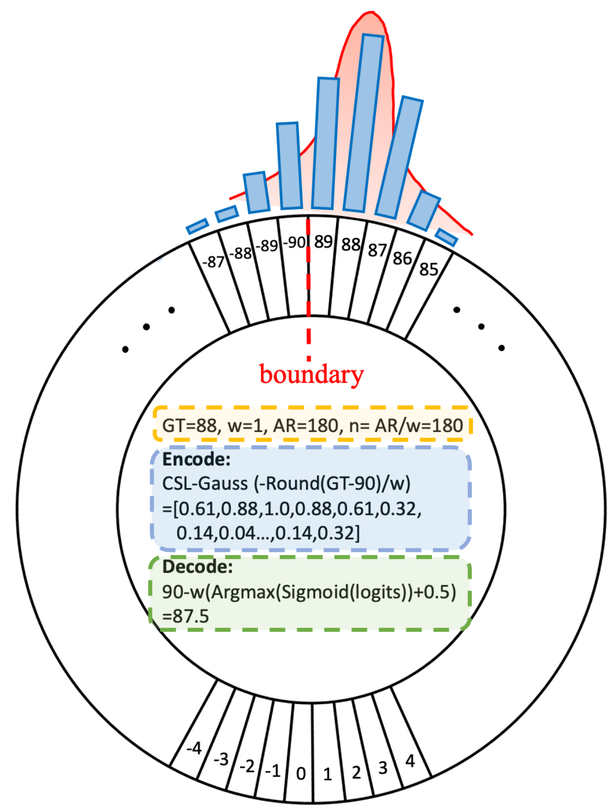
- 使用密集编码（DCL）取代稀疏编码（SCL） （针对问题1）



SCL： One-Hot Label



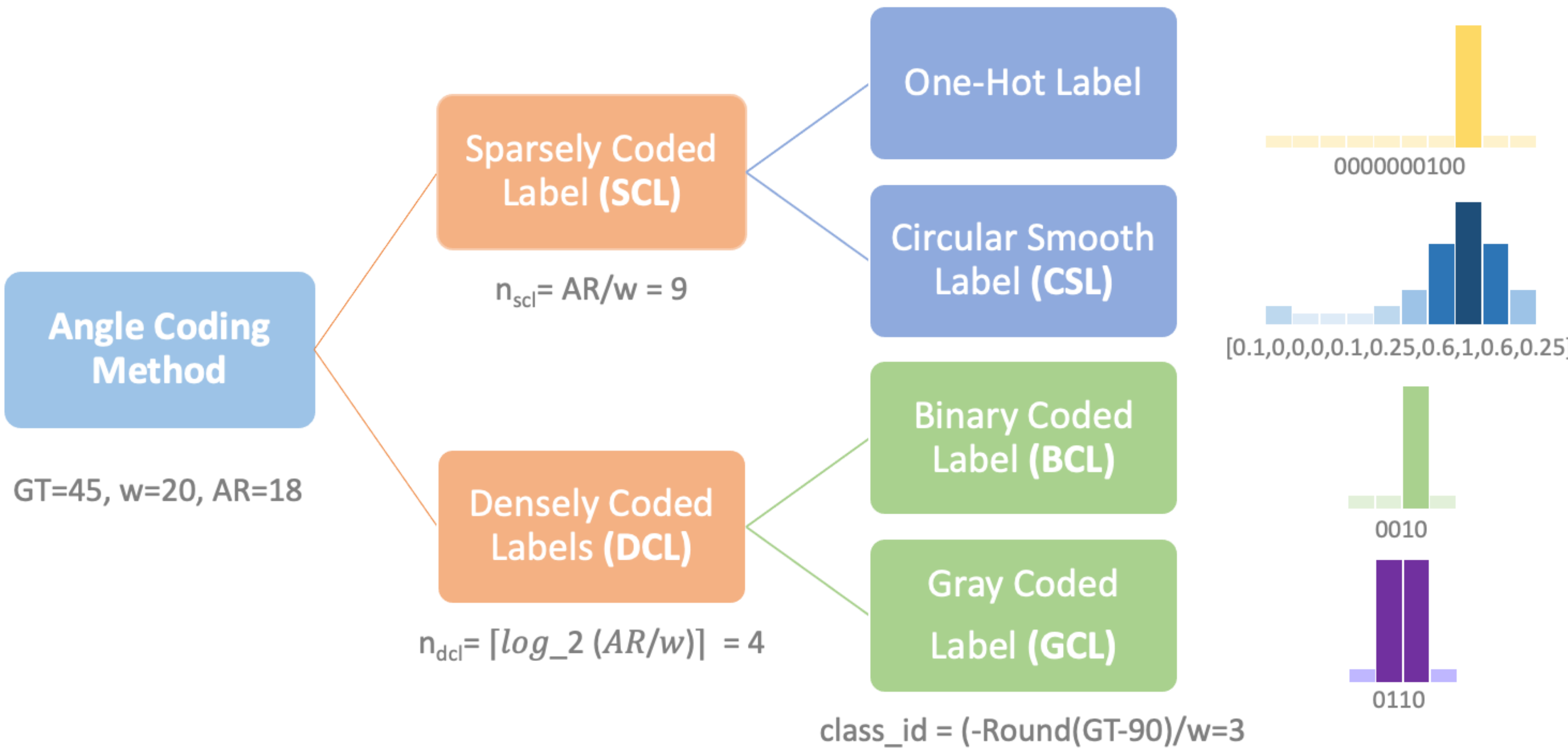
DCL：Binary Coded Labe



SCL: Circular Smooth Label

Densely Coded Label (DCL)

- 使用密集编码（DCL）取代稀疏编码（SCL）（针对问题1）



Densely Coded Label (DCL)

- 角度距离和长宽比感知的权重（针对问题2）

$$W_{ADARSW}(\Delta\theta) = |\sin(\alpha(\Delta\theta))| = |\sin(\alpha(\theta_{gt} - \theta_{pred}))|$$

$$\alpha = \begin{cases} 1, & (h_{gt}/w_{gt}) > r \\ 2, & otherwise \end{cases}$$

(a) Ground Truth

(b) Prediction after using ADARSW

角度离散化粒度w

- 角度离散化粒度w太小，角度类别数太多，分类影响性能
- 角度离散化粒度w太大，理论误差太大，性能上限较低

Method	ω	BR	SV	LV	SH	HA	5-mAP$_{50}$	mAP$_{50}$	mAP$_{75}$	mAP$_{50:95}$
Reg	-	34.52	51.42	50.32	73.37	55.93	53.12	62.21	26.07	31.49
CSL	180/180	35.94	53.42	61.06	81.81	62.14	58.87	64.40	32.58	35.04
BCL	180/4	30.74	40.54	50.98	72.07	59.54	50.77	62.38	24.88	31.01
	180/8	36.65	52.58	60.46	82.24	61.60	58.71	**66.17**	33.14	35.77
	180/32	**39.83**	54.41	60.62	80.81	60.32	**59.20**	65.93	**35.66**	**36.71**
	180/64	38.22	**54.70**	60.16	80.75	60.11	58.79	65.00	34.31	36.00
	180/128	36.76	53.73	**61.35**	**82.52**	58.42	58.56	65.14	34.28	35.69
	180/180	37.42	53.72	58.70	80.73	**63.31**	58.78	65.83	33.94	36.35
	180/256	37.66	53.83	60.66	80.43	60.74	58.66	64.97	33.52	35.21
	180/512	37.93	53.85	58.52	80.04	60.87	58.24	64.88	33.09	34.99
GCL	180/4	30.90	41.20	48.30	72.93	60.16	50.70	62.98	23.83	30.81
	180/8	36.88	51.10	59.81	82.40	61.57	58.35	65.23	33.92	35.29
	180/32	38.04	**54.77**	60.88	**82.75**	61.24	**59.54**	65.11	**34.67**	36.15
	180/64	**38.05**	54.36	60.59	81.84	60.39	59.05	64.78	33.23	35.67
	180/128	37.74	54.36	59.43	81.15	60.51	58.64	**66.13**	33.65	**36.34**
	180/256	35.81	53.78	58.35	81.45	59.84	57.85	64.87	33.77	35.97
	180/512	37.99	54.23	**61.61**	80.84	**62.13**	59.36	64.34	34.08	35.92

角度离散化粒度w的可视化

(a) $\omega = 180/4$

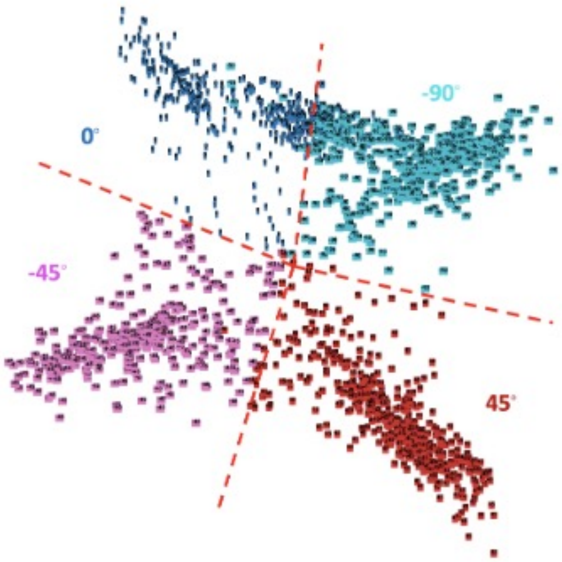
(b) $\omega = 180/32$

(c) $\omega = 180/128$

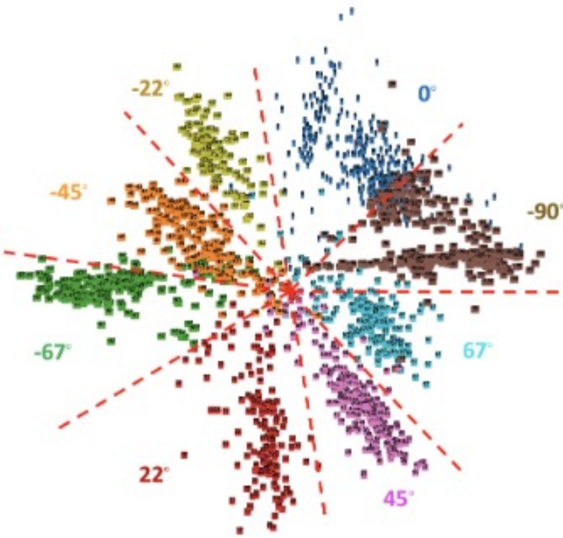
(d) $\omega = 180/256$

DCL的可视化

- 角度特征可视化

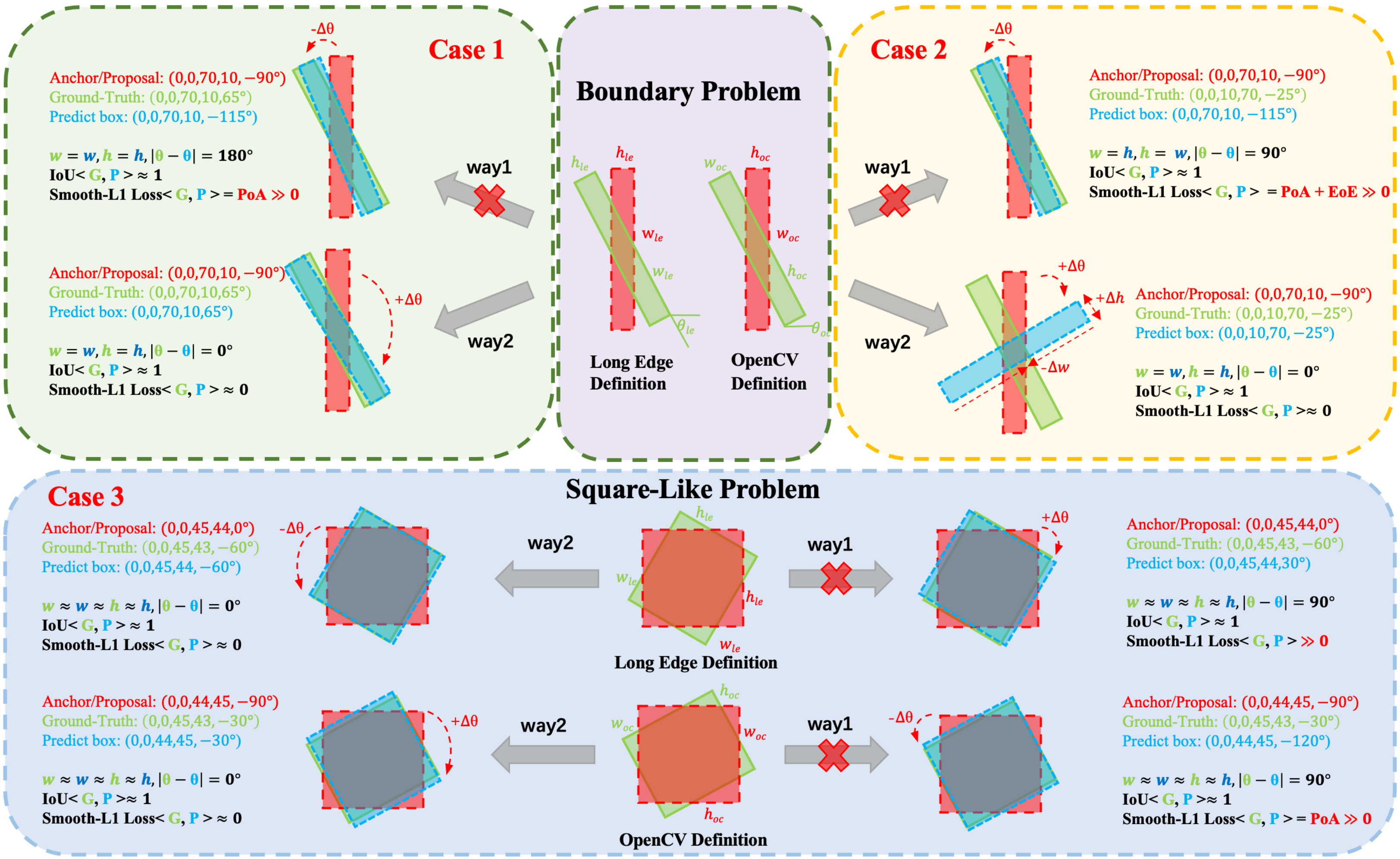


(a) $\omega = 180/4$

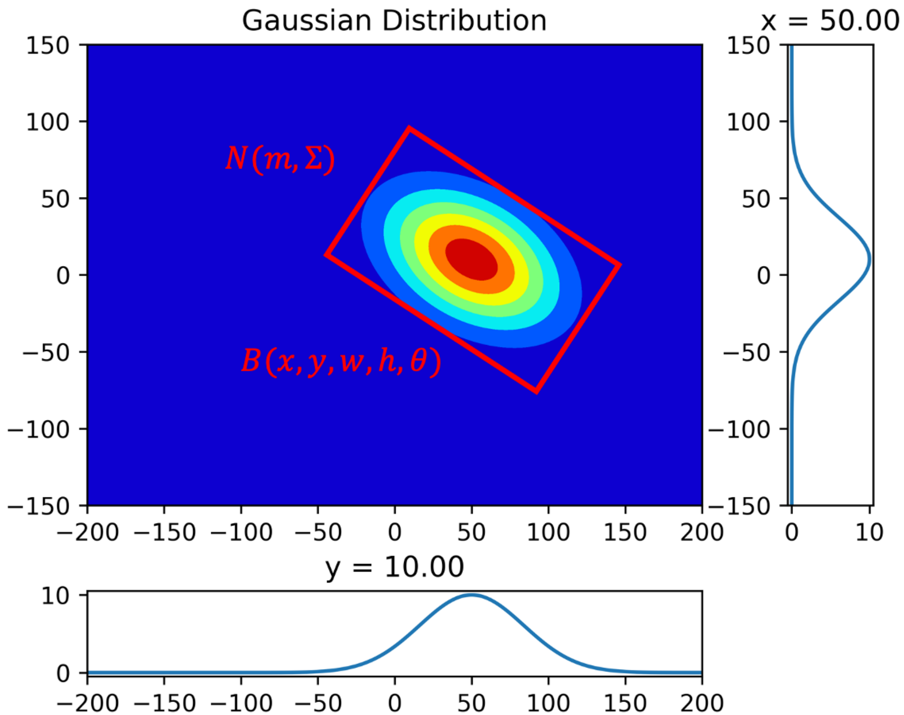


(b) $\omega = 180/8$

边界问题和类正方形检测问题



一种统一的解决方法GWD



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

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 $w \approx h$.

$$
d^2 = \|\mathbf{m}_1 - \mathbf{m}_2\|_2^2 + \mathbf{Tr}\left(\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2 - 2(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2})^{1/2}\right) \tag{3}
$$

$$
L_{gwd} = 1 - \frac{1}{\tau + f(d^2)}, \quad \tau \geq 1
$$

X. Yang, et al. "Rethinking Rotated Object Detection with Gaussian Wasserstein Distance Loss." In ICML 2021.

一种统一的解决方法GWD

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

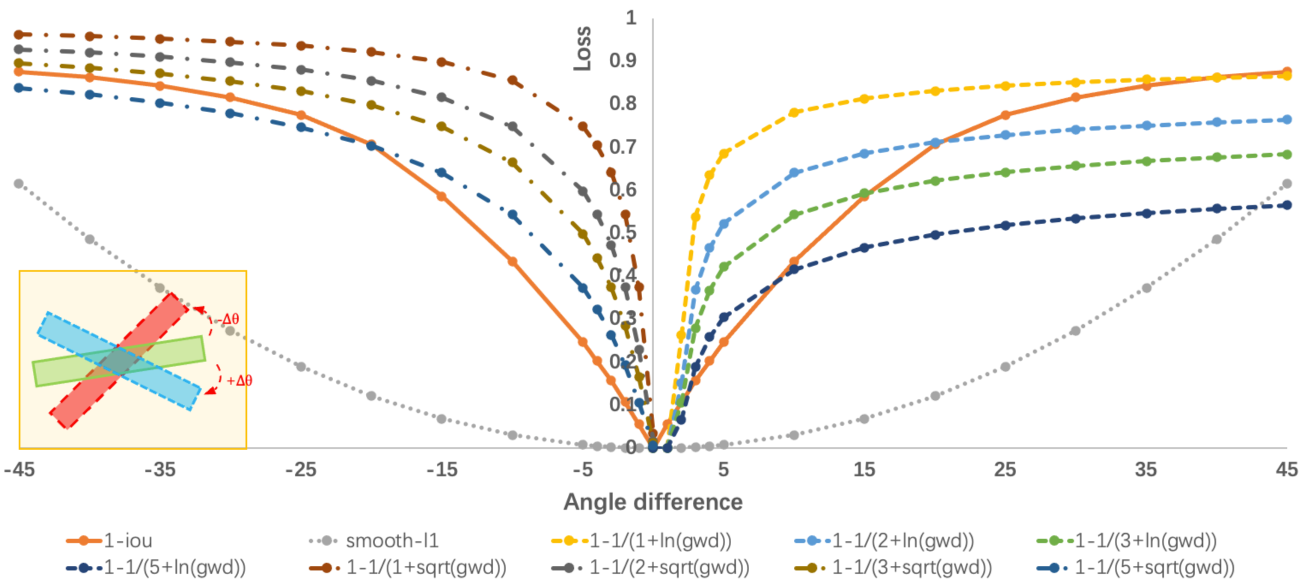
Property 1: $\mathbf{\Sigma}^{1/2}(w,h,\theta) = \mathbf{\Sigma}^{1/2}(h,w,\theta-\frac{\pi}{2})$;

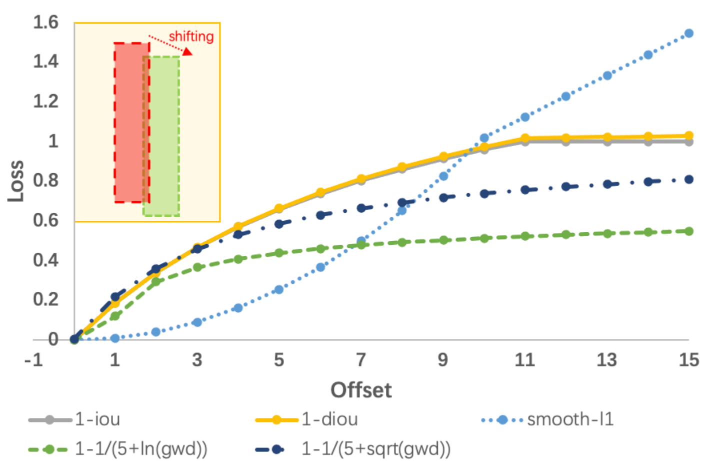
Property 2: $\mathbf{\Sigma}^{1/2}(w,h,\theta) = \mathbf{\Sigma}^{1/2}(w,h,\theta-\pi)$;

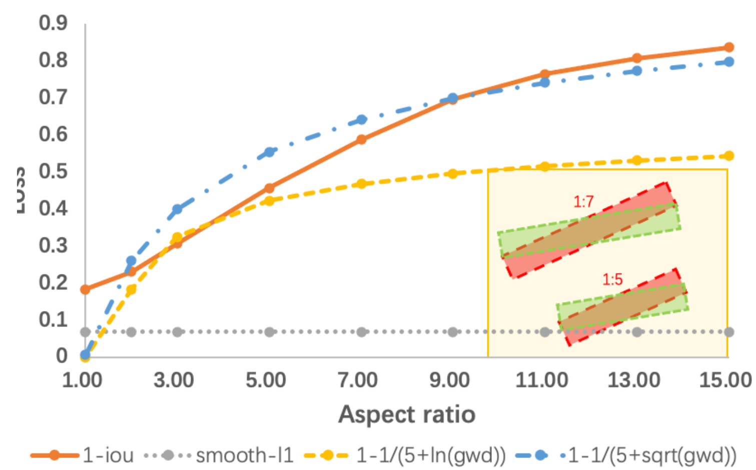
Property 3: $\mathbf{\Sigma}^{1/2}(w,h,\theta) \approx \mathbf{\Sigma}^{1/2}(w,h,\theta-\frac{\pi}{2})$, if $w \approx h$.

```
boxes1 = np.array([    # prediction box
    [50, 50, 10, 70, -35],      # 90 <--> 180
    [50, 50, 70, 10, -90.5],    # 90   PoA + EoE
    [50, 50, 70, 10, -90.5],    # 180  PoA
    [50, 50, 40, 40, -35],      # 180  w=h
], np.float32)

boxes2 = np.array([    # ground truth
    [50, 50, 70, 10, 55],
    [50, 50, 10, 70, -0.5],
    [50, 50, 70, 10, 89.5],
    [50, 50, 40, 40, 55],
], np.float32)

print(iou_rotate_calculate2(boxes1, boxes2).reshape(-1,))  # [0.9999996 0.9999998 0.9999998 1.        ]
print(diou_rotate_calculate(boxes1, boxes2).reshape(-1,))  # [0.9999997  0.99999994 0.99999994 1.        ]
print(gaussian_wasserstein_distance(boxes1, boxes2))     # [6.1035156e-05 3.1062821e-04 3.1062821e-04 0.0000000e+00]
```

评估与损失不一致问题


Loss
Angle difference
-Δθ
+Δθ
1-iou
smooth-l1
1-1/(1+ln(gwd))
1-1/(2+ln(gwd))
1-1/(3+ln(gwd))
1-1/(5+ln(gwd))
1-1/(1+sqrt(gwd))
1-1/(2+sqrt(gwd))
1-1/(3+sqrt(gwd))
1-1/(5+sqrt(gwd))


Loss
Offset
shifting
1-iou
1-diou
smooth-l1
1-1/(5+ln(gwd))
1-1/(5+sqrt(gwd))


Aspect ratio
1:7
1:5
1-iou
smooth-l1
1-1/(5+ln(gwd))
1-1/(5+sqrt(gwd))

一种统一的解决方法GWD

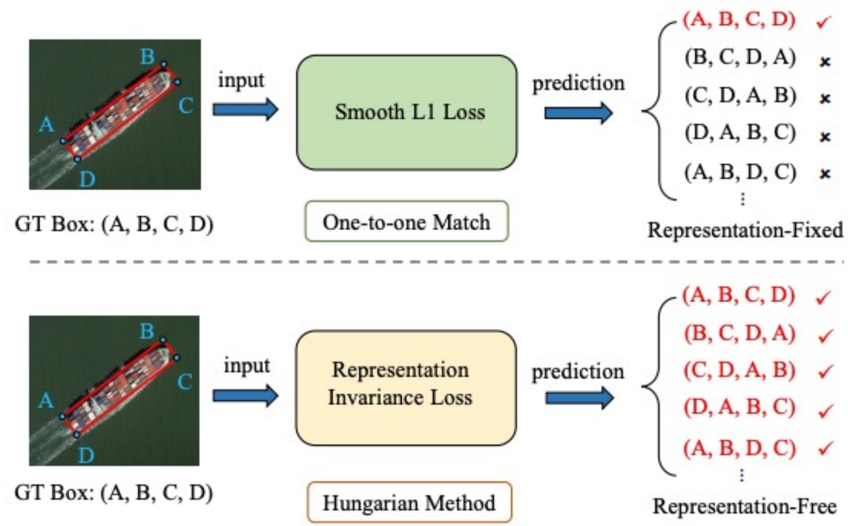
	Method	Backbone	MS	PL	BD	BR	GTF	SV	LV	SH	TC	BC	ST	SBF	RA	HA	SP	HC	mAP$_{50}$
Two-stage methods	FR-O (Xia et al., 2018)	R-101		79.09	69.12	17.17	63.49	34.20	37.16	36.20	89.19	69.60	58.96	49.4	52.52	46.69	44.80	46.30	52.93
	ICN (Azimi et al., 2018)	R-101	✓	81.40	74.30	47.70	70.30	64.90	67.80	70.00	90.80	79.10	78.20	53.60	62.90	67.00	64.20	50.20	68.20
	KARNET (Tang et al., 2020)	R-50		89.33	83.55	44.79	71.61	63.05	67.06	69.53	90.47	79.46	77.84	51.04	60.97	65.38	69.46	49.53	68.87
	RADet (Li et al., 2020b)	RX-101		79.45	76.99	48.05	65.83	65.46	74.40	68.86	89.70	78.14	74.97	49.92	64.63	66.14	71.58	62.16	69.09
	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
	AOOD (Zou et al., 2020)	DPN-92	✓	89.99	81.25	44.50	73.20	68.90	60.33	66.86	90.89	80.99	86.23	64.98	63.88	65.24	68.36	62.13	71.18
	Cascade-FF (Hou et al., 2020)	R-152		89.9	80.4	51.7	77.4	68.2	75.2	75.6	90.8	78.8	84.4	62.3	64.6	57.7	69.4	50.1	71.8
	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
	SARD (Wang et al., 2019b)	R-101		89.93	84.11	54.19	72.04	68.41	61.18	66.00	90.82	87.79	86.59	65.65	64.04	66.68	68.84	68.03	72.95
	GLS-Net (Li et al., 2020a)	R-101		88.65	77.40	51.20	71.03	73.30	72.16	84.68	90.87	80.43	85.38	58.33	62.27	67.58	70.69	60.42	72.96
	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
	MFIAR-Net (Yang et al., 2020a)	R-152	✓	89.62	84.03	52.41	70.30	70.13	67.64	77.81	90.85	85.40	86.22	63.21	64.14	68.31	70.21	62.11	73.49
	Gliding Vertex (Xu et al., 2020b)	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
	SAR (Lu et al., 2020)	R-152		89.67	79.78	54.17	68.29	71.70	77.90	84.63	90.91	88.22	87.07	60.49	66.95	75.13	75.28	64.29	75.28
	Mask OBB (Wang et al., 2019a)	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., 2020b)	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
	F^3-Net (Ye et al., 2020)	R-152	✓	88.89	78.48	54.62	74.43	72.80	77.52	87.54	90.78	87.64	85.63	63.80	64.53	78.06	72.36	63.19	76.02
	CenterMap (Wang et al., 2020b)	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
	MRDet (Qin et al., 2020)	R-101		89.49	84.29	55.40	66.68	76.27	82.13	87.86	90.81	86.92	85.00	52.34	65.98	76.22	76.78	67.49	76.24
	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
	OPLD (Song et al., 2020)	R-101	✓	89.37	85.82	54.10	79.58	75.00	75.13	86.92	90.88	86.42	86.62	62.46	68.41	73.98	68.11	63.69	76.43
	SCRDet++ (Yang et al., 2020c)	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
	HSP (Xu et al., 2020a)	R-101	✓	90.39	86.23	56.12	80.59	77.52	73.26	83.78	90.80	87.19	85.67	69.08	72.02	76.98	72.50	67.96	78.01
	FR-Est (Fu et al., 2020a)	R-101-DCN	✓	89.78	85.21	55.40	77.70	80.26	83.78	87.59	90.81	87.66	86.93	65.60	68.74	71.64	79.99	66.20	78.49
Single-stage methods	IENet (Lin et al., 2019)	R-101	✓	80.20	64.54	39.82	32.07	49.71	65.01	52.58	81.45	44.66	78.51	46.54	56.73	64.40	64.24	36.75	57.14
	TOSO (Feng et al., 2020)	R-101	✓	80.17	65.59	39.82	39.95	49.71	65.01	53.58	81.45	44.66	78.51	48.85	56.73	64.40	64.24	36.75	57.92
	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
	Axis Learning (Xiao et al., 2020)	R-101		79.53	77.15	38.59	61.15	67.53	70.49	76.30	89.66	79.07	83.53	47.27	61.01	56.28	66.06	36.05	65.98
	A^{2}S-Det (Xiao et al., 2021)	R-101		89.59	77.89	46.37	56.47	75.86	74.83	86.07	90.58	81.09	83.71	50.21	60.94	65.29	69.77	50.93	70.64
	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
	ROPDet (Yang et al., 2020d)	R-101-DCN	✓	90.01	82.82	54.47	69.65	69.23	70.78	75.78	90.84	86.13	84.76	66.52	63.71	67.13	68.38	46.09	72.42
	HRP-Net (He et al., 2020)	HRNet-W48		89.33	81.64	48.33	75.21	71.39	74.82	77.62	90.86	81.23	81.96	62.93	62.17	66.27	66.98	62.13	72.83
	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
	CFC-Net (Ming et al., 2021)	R-101	✓	89.08	80.41	52.41	70.02	76.28	78.11	87.21	90.89	84.47	85.64	60.51	61.52	67.82	68.02	50.09	73.50
	R^{4}Det (Sun et al., 2020)	R-152		88.96	85.42	52.91	73.84	74.86	81.52	80.29	90.79	86.95	85.25	64.05	60.93	69.00	70.55	67.76	75.84
	R^{3}Det (Yang et al., 2021)	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., 2020b)	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., 2020)	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

一种统一的解决方法GWD

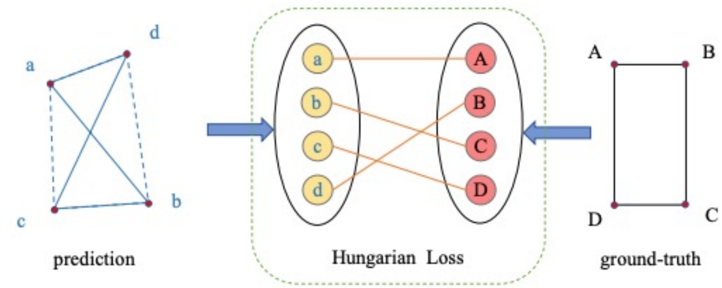
Method	Dataset	Data Aug.	Reg. Loss	$Hmean_{50}/AP_{50}$	$Hmean_{60}/AP_{60}$	$Hmean_{75}/AP_{75}$	$Hmean_{85}/AP_{85}$	$Hmean_{50:95}/AP_{50:95}$
RetinaNet	HRSC2016	R+F+G	Smooth L1	84.28	74.74	48.42	12.56	47.76
			GWD	85.56 (+1.28)	84.04 (+9.30)	60.31 (+11.89)	17.14 (+4.58)	52.89 (+5.13)
			-	**87.45** (+3.17)	**86.72** (+11.98)	**72.39** (+23.97)	**27.68** (+15.12)	**57.80** (+10.04)
R^3Det			Smooth L1	88.52	79.01	43.42	4.58	46.18
			GWD	89.43 (+0.91)	88.89 (+9.88)	65.88 (+22.46)	15.02 (+10.44)	56.07 (+9.89)
			-	**89.97** (+1.45)	**89.73** (+10.72)	**77.38** (+33.96)	**25.12** (+20.54)	**61.40** (+15.22)
RetinaNet	MSRA-TD500	R+F+G	Smooth L1	70.98	62.42	36.73	12.56	37.89
			GWD	76.76 (+5.78)	68.58 (+6.16)	44.21 (+7.48)	17.75 (+5.19)	43.62 (+5.73)
			-	**76.96** (+5.98)	**70.08** (+7.66)	**46.95** (+10.22)	**19.59** (+7.03)	**45.24** (+7.35)
	ICDAR2015	F	Smooth L1	69.78	64.15	36.97	8.71	37.73
			GWD	74.29 (+4.51)	68.34 (+4.19)	43.39 (+6.42)	10.50 (+1.79)	41.68 (+3.95)
			-	**75.32** (+5.54)	**69.94** (+5.79)	**44.46** (+7.49)	**10.70** (+1.99)	**42.68** (+4.95)
		R+F	Smooth L1	74.83	69.46	42.02	11.59	41.98
			GWD	76.15 (+1.32)	71.26 (+1.80)	**45.59** (+3.57)	**11.65** (+0.06)	43.58 (+1.60)
			-	**77.92** (+3.09)	**72.77** (+3.31)	43.27 (+1.25)	11.09 (-0.50)	**43.65** (+1.67)
R^3Det		F	Smooth L1	74.28	68.12	35.73	8.01	39.10
			GWD	75.59 (+1.31)	68.36 (+0.24)	40.24 (+4.51)	9.15 (+1.14)	40.80 (+1.70)
			-	**77.72** (+2.43)	**71.99** (+3.87)	**43.95** (+8.22)	**10.43** (+2.42)	**43.29** (+4.19)
		R+F	Smooth L1	75.53	69.69	37.69	9.03	40.56
			GWD	77.09 (+1.56)	71.52 (+1.83)	41.08 (+3.39)	10.10 (+1.07)	42.17 (+1.61)
			-	**79.63** (+4.63)	**73.30** (+3.61)	**43.51** (+5.82)	**10.61** (+1.58)	**43.61** (+3.05)

Baseline	Method	Box Def.	v1.0 tranval/test									v1.0 train/val			v1.5	v2.0
			BR†	SV†	LV†	SH†	HA†	ST‡	RA‡	7-AP_{50}	AP_{50}	AP_{50}	AP_{75}	$AP_{50:95}$	AP_{50}	AP_{50}
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	58.87	44.16
	-	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	56.10	43.06
	IoU-Smooth L1	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	59.16	46.31
	Modulated Loss	Quad.	43.21	70.78	54.70	72.68	**60.99**	79.72	62.08	63.45	67.20	65.15	40.59	**39.12**	**61.42**	**46.71**
	RIDet	Quad.	40.81	67.63	55.45	72.42	55.49	78.09	64.75	62.09	66.06	64.07	**40.98**	39.05	58.91	45.35
	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	58.55	43.34
	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	59.38	45.46
	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	60.03	46.65
	-	D_{oc}	**44.00**	**74.45**	**72.48**	**84.30**	**65.54**	**80.03**	**65.05**	**69.41**	**71.28**	**68.14**	**44.48**	**42.15**	**62.50**	**47.69**
R^3Det	-	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	62.91	48.43
	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	61.98	48.71
	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**	**63.22**	**49.25**
	-	D_{oc}	**48.34**	**75.09**	**78.88**	86.52	**65.48**	82.08	**61.51**	**71.13**	**71.73**	**68.87**	**44.48**	**42.11**	**65.18**	**50.90**

无序四边形检测扩展



- 无序四边形检测的常常遭受表示歧义的困扰。
- 通过匈牙利算法对预测框和Ground Truth 进行点的配对，将损失最小的匹配结果来优化模型。

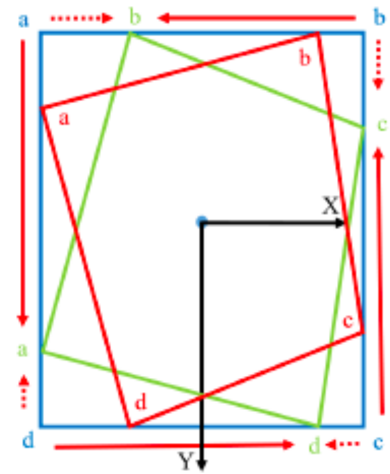


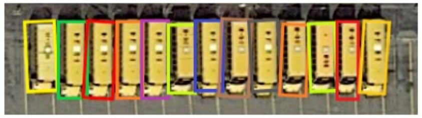
Q. Ming, Z. Zhou, L. Miao, X. Yang, et al. " Optimization for Oriented Object Detection via Representation Invariance Loss."

有序四边形检测扩展

- 四边形检测也存在边界问题
- RSDet将排好序的角点向前或者向后都挪一位，然后各自计算损失，取最小的的那个损失值，具体公式如下：

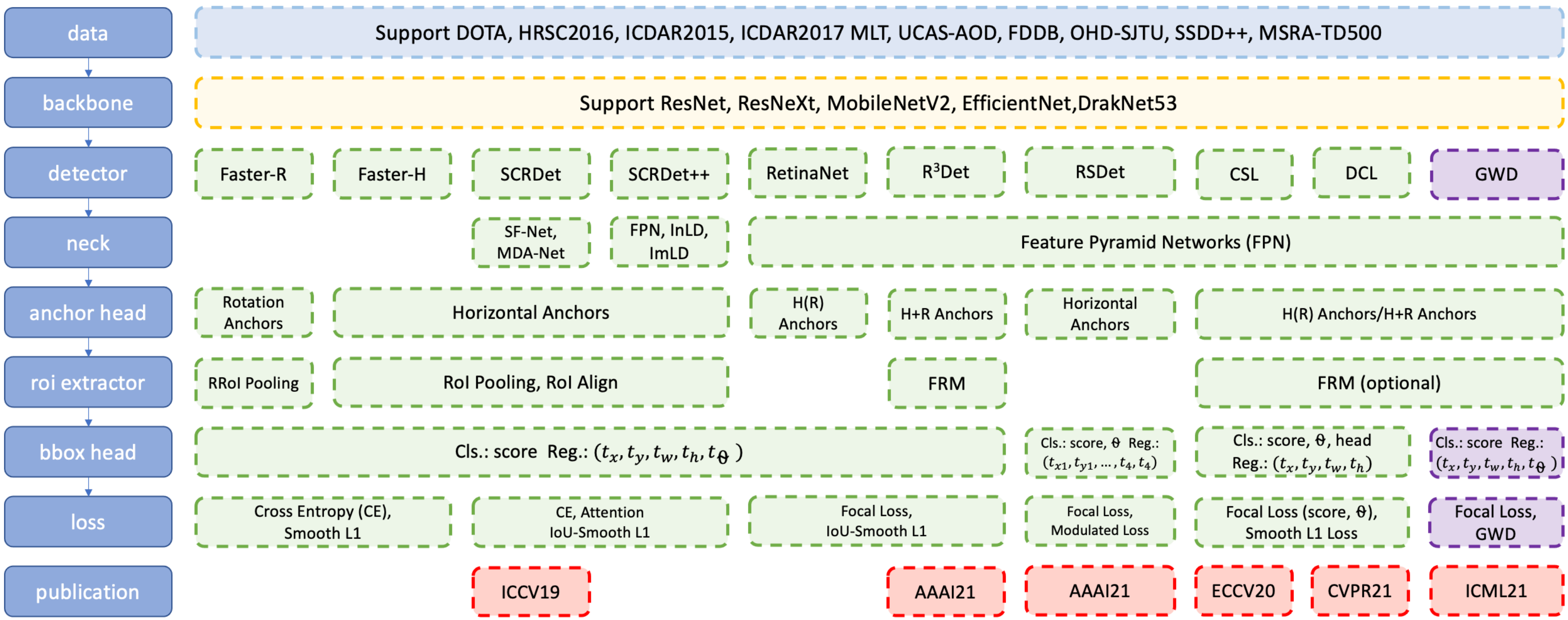
$$
\ell_{mr}^{8p} = \min \begin{cases} \sum_{i=0}^{3} \left(\frac{|x_{(i+3)\%4} - x_i^*|}{w_a} + \frac{|y_{(i+3)\%4} - y_i^*|}{h_a} \right) \\ \sum_{i=0}^{3} \left(\frac{|x_i - x_i^*|}{w_a} + \frac{|y_i - y_i^*|}{h_a} \right) \\ \sum_{i=0}^{3} \left(\frac{|x_{(i+1)\%4} - x_i^*|}{w_a} + \frac{|y_{(i+1)\%4} - y_i^*|}{h_a} \right) \end{cases}
$$





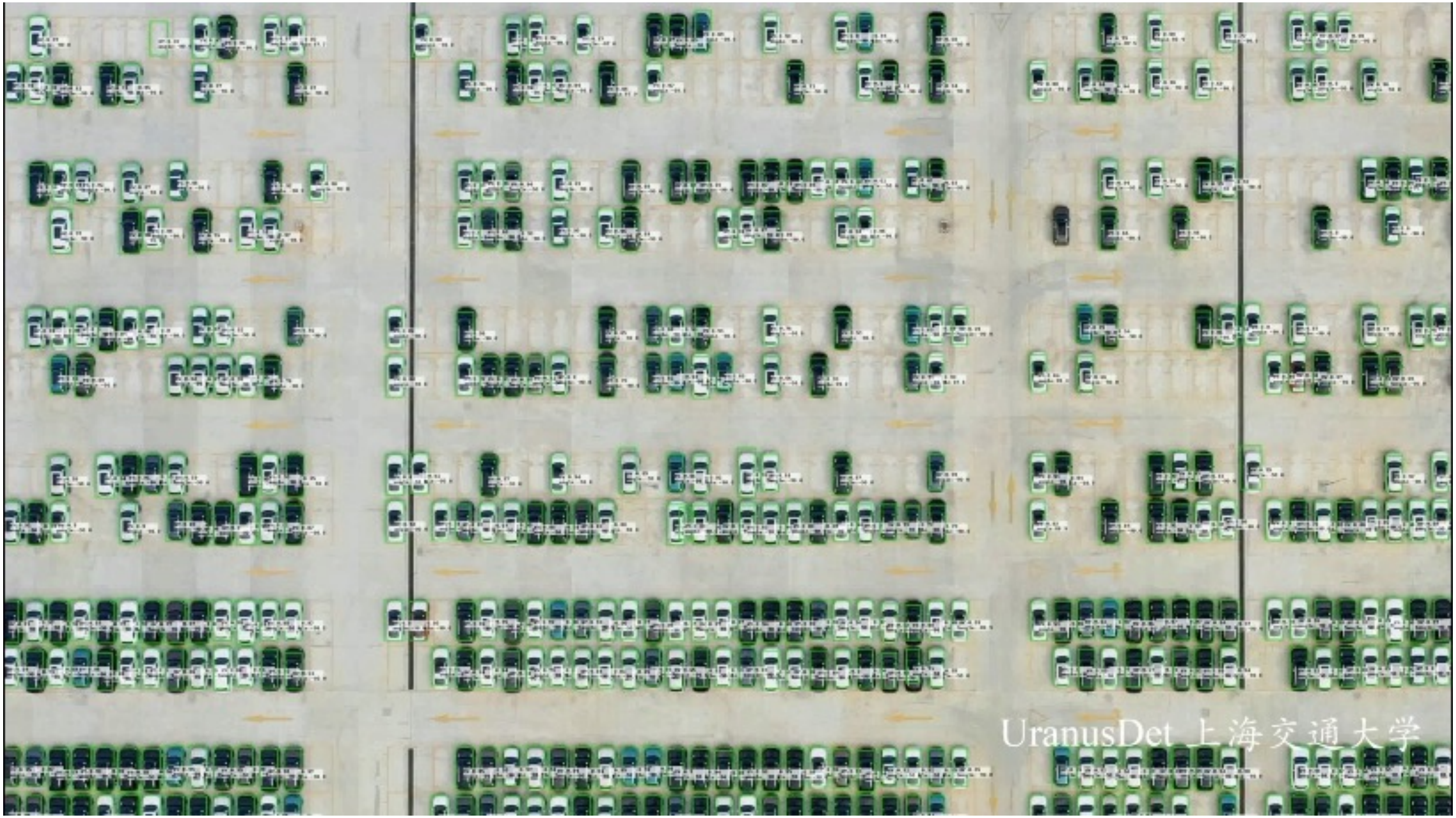
W. Qian, X. Yang, et al. "Learning modulated loss for rotated object detection." In AAAI 2021

旋转检测benchmark



https://github.com/yangxue0827/RotationDetection

Demo



相关资料

- 个人主页：https://yangxue0827.github.io/
- 参考代码：https://github.com/yangxue0827/RotationDetection
- 参考文献

  - X. Yang, et al. "Scrdet: Towards more robust detection for small, cluttered and rotated objects." In ICCV. 2019.
  - X. Yang, J. Yan. “Arbitrary-Oriented Object Detection with Circular Smooth Label.” In ECCV 2020.
  - W. Qian, X. Yang, et al. "Learning modulated loss for rotated object detection." In AAAI 2021.
  - X. Yang, et al. "Dense Label Encoding for Boundary Discontinuity Free Rotation Detection." In CVPR 2021.
  - X. Yang, et al. "Rethinking Rotated Object Detection with Gaussian Wasserstein Distance Loss." In ICML 2021.
  - X. Yang, et al. "R3Det: Refined Single-Stage Detector with Feature Refinement for Rotating Object." In AAAI 2021.
  - Q. Ming, Z. Zhou, L. Miao, X. Yang, et al. " Optimization for Oriented Object Detection via Representation Invariance Loss."

Q&A

谢谢！