PolarMask: 一阶段实例分割新思路

报告人：谢恩泽
导师：罗平

前言

- 一阶段的Anchor free的思想在物体检测领域取得了很大的成功…
- FCOS, FoveaBox, RepPoints, CenterNet…

- 问题：
- Anchor free 是否能解决实例分割？
- 实例分割能否有新的建模方式？
- 实例分割能否不依赖检测结果，而且和检测计算量一样？

Revisiting Mask RCNN

- Mask RCNN是先检测再分割，对于RoI区域进行像素级分割

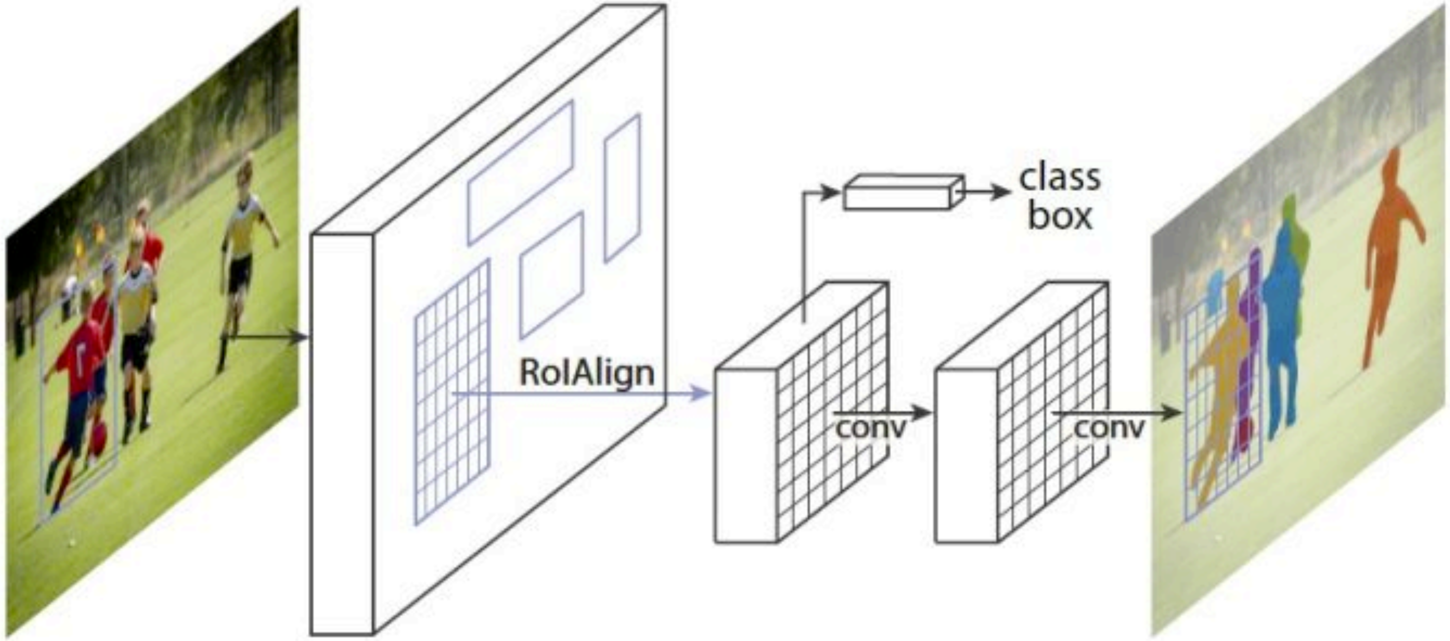


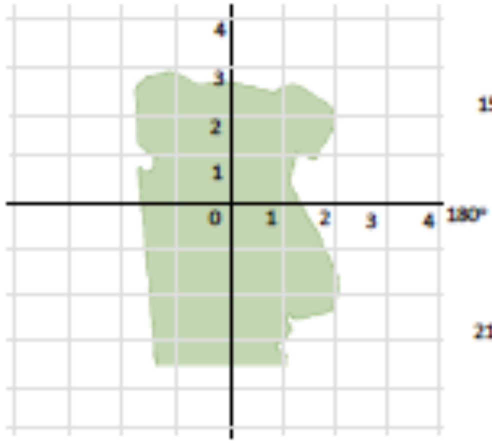
Figure 1. The **Mask R-CNN** framework for instance segmentation.

两种实例分割建模方式

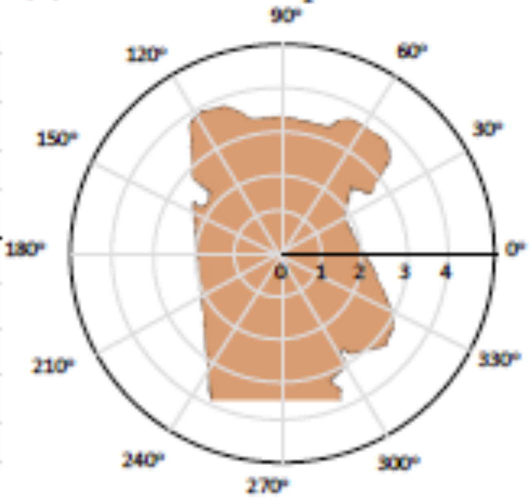
- 1 像素级建模 (b)
- 2 轮廓建模
- 2.1 基于直角坐标系 (c)
- 2.2 基于极坐标系 (d)

(a) Original image

(b) Pixel-wise Representation



(c) Cartesian Representation



(d) Polar Representation

- 在极坐标下，实例分割可以转化为
- (1) 实例中心分类
- (2) 固定角度的距离回归

- 优势：
- (1) anchor free and box free 不依赖检测框
- (2) 把检测和实例分割用统一的建模方式表达
- FCOS可以看成PolarMask的特殊形式，而PolarMask可以看作FCOS的通用形式，因为bbox本质上是最简单的Mask，只有0,90,180,270四个角度回归长度。

网络结构

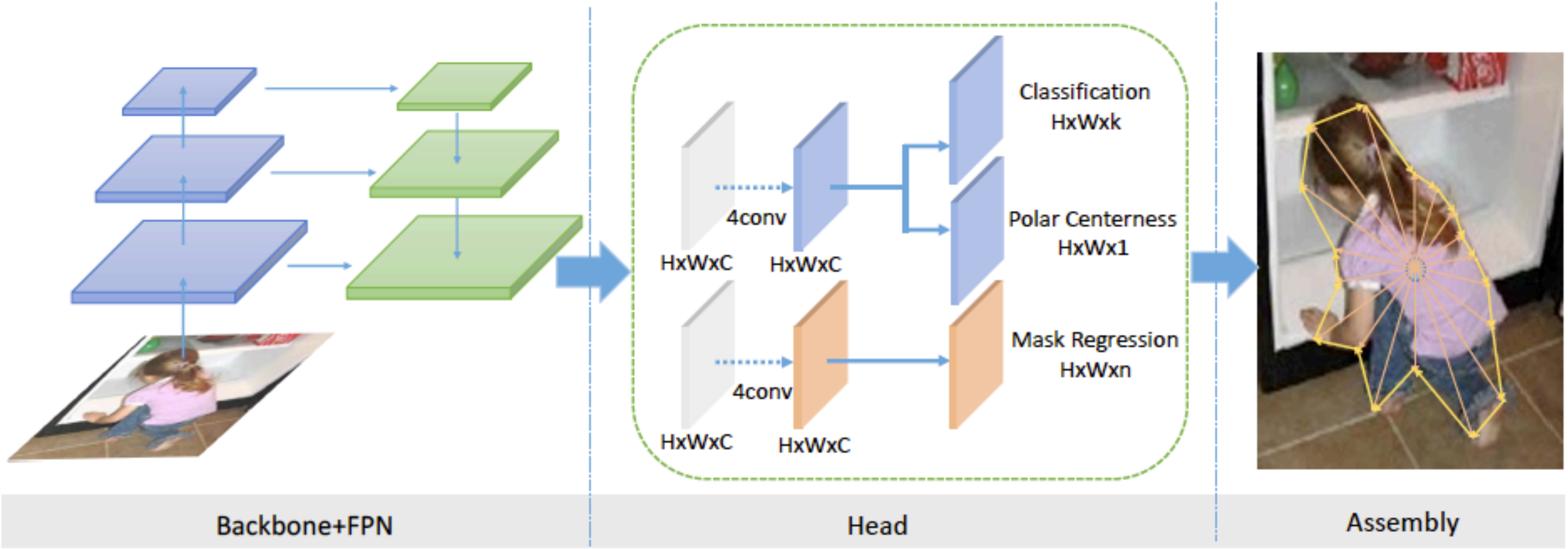
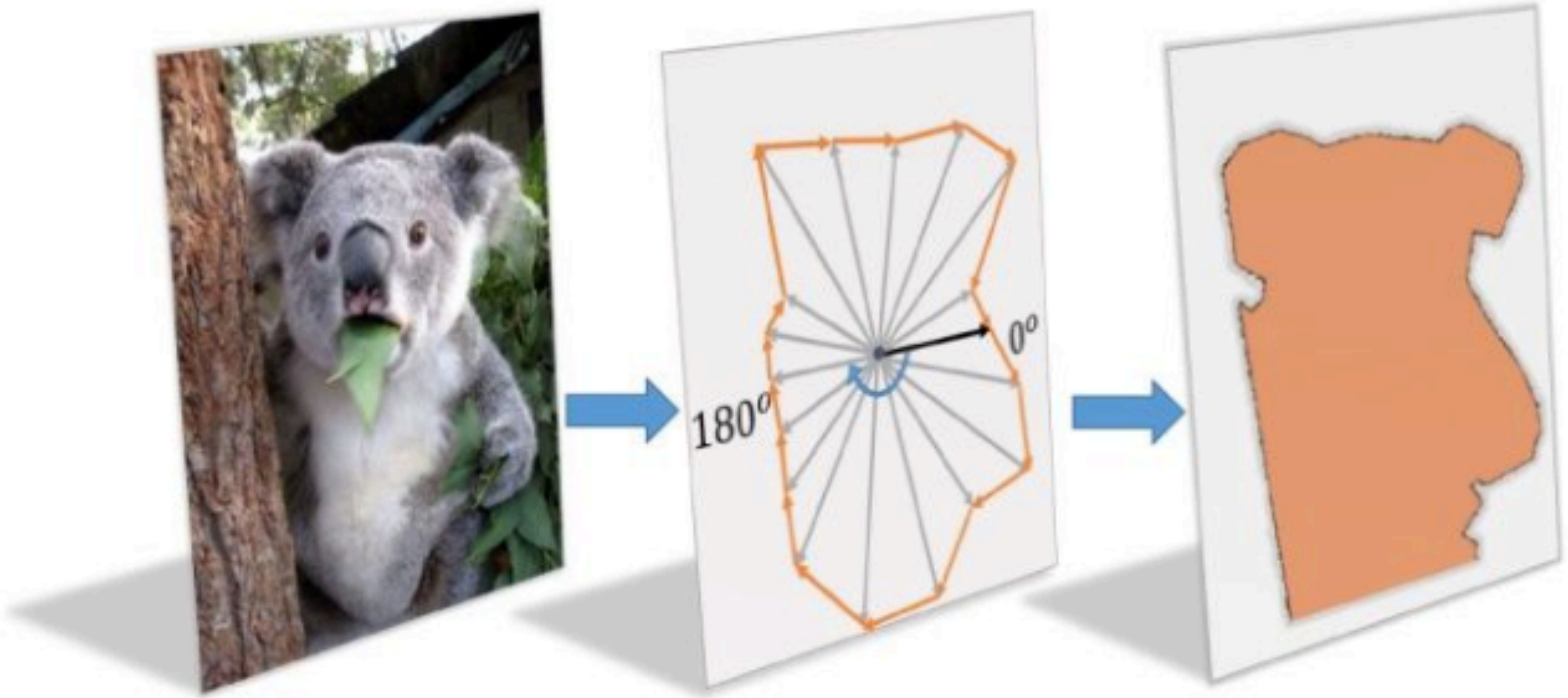


Figure 2 – The overall pipeline of PolarMask. The left part contains the backbone and feature pyramid to extract features of different levels. The middle part is the two heads for classification and polar mask regression. H, W, C are the height, width, channels of feature maps, respectively, and k is the number of categories (e.g., $k = 80$ on the COCO dataset), n is the number of rays (e.g., $n = 36$)

Polar Segmentation建模



$$x_i = \cos\theta_i \times d_i + x_c$$

$$y_i = \sin\theta_i \times d_i + y_c.$$

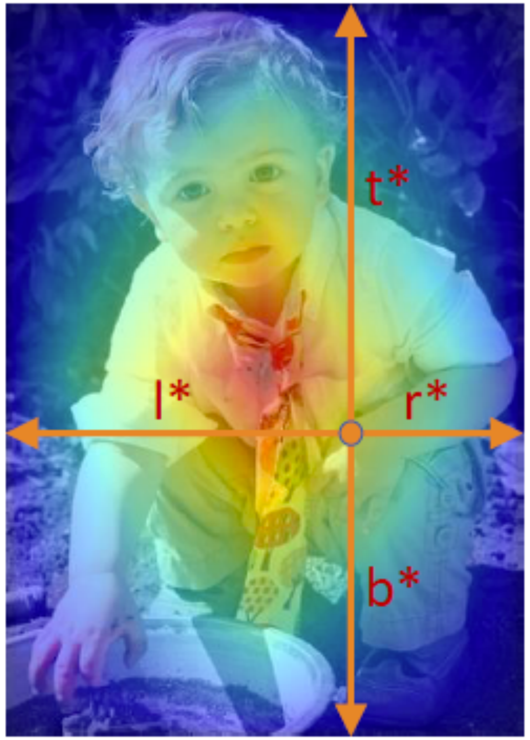
两个重要的优化方法

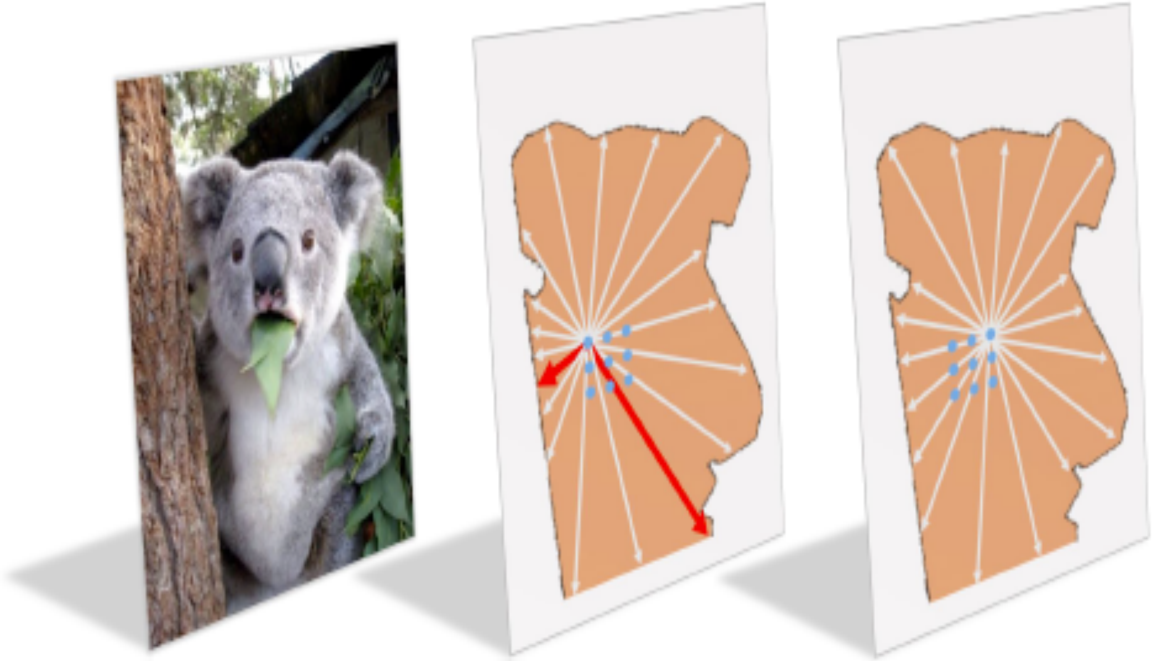
- (1) Polar Centerness
- (2) Polar IoU Loss

Centerness -> Polar Centerness

$$\text{centerness}^* = \sqrt{\frac{\min(l^*, r^*)}{\max(l^*, r^*)} \times \frac{\min(t^*, b^*)}{\max(t^*, b^*)}}.$$

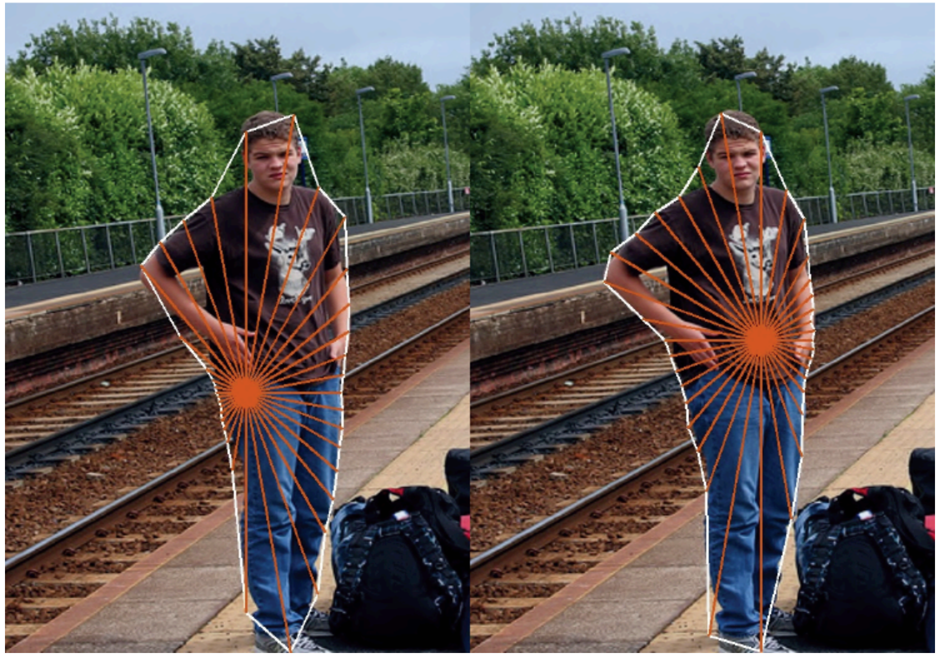
$$\text{Polar Centerness} = \sqrt{\frac{\min(\{d_1, d_2, \ldots, d_n\})}{\max(\{d_1, d_2, \ldots, d_n\})}}$$





Polar Centerness

- Polar Centerness可以有效提高1.4的性能，同时不增加网络复杂度。



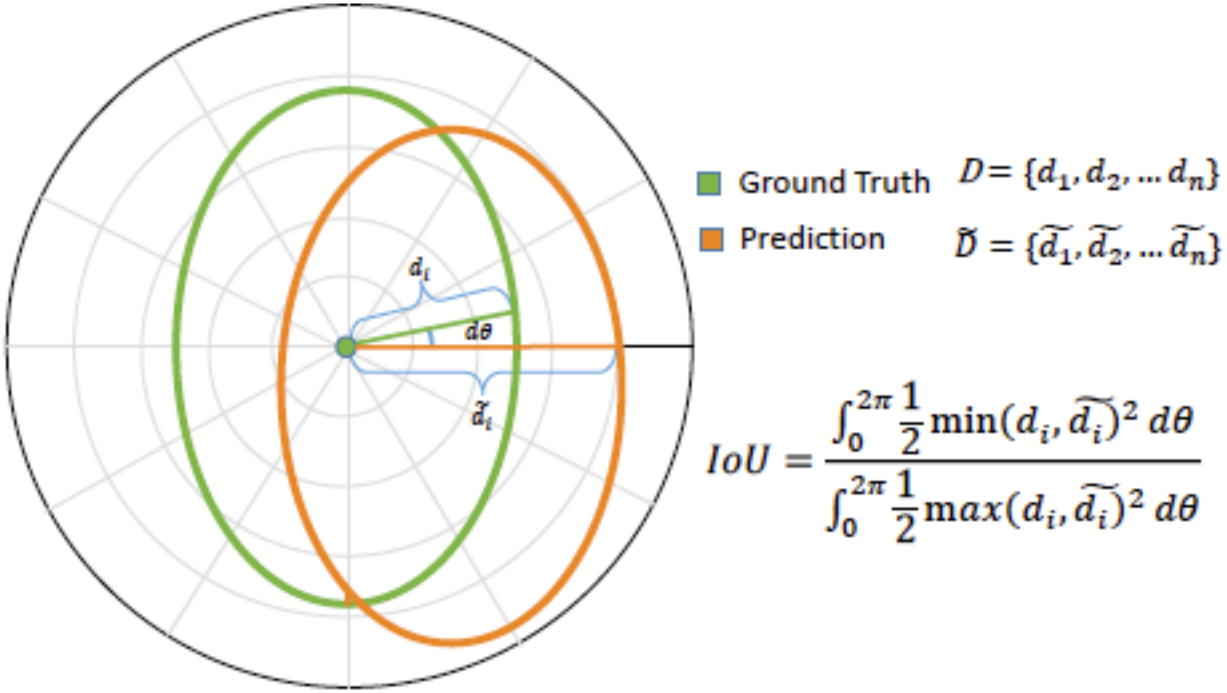
centerness	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
Original	27.7	49.6	27.4	12.6	30.2	39.7
Polar	**29.1**	49.5	29.7	12.6	31.8	42.3

(c) **Polar Centerness *vs.* Centerness**: Polar Centerness bring a large gain, especially high IoU AP_{75} and large instance AP_L.

Polar IoU Loss

- 在PolarMask中，需要回归k(k=36)根射线的距离，这相比目标检测更为复杂
- 大多数anchor free检测器采用IoU Loss, 训练更加稳定…
- 如何在极坐标下巧妙的计算出Mask 的近似IoU?

Polar IoU Loss



假设两个射线的夹角无穷小，
即可用微积分来求Mask 面积

$$\mathrm{IoU} = \frac{\int_0^{2\pi} \frac{1}{2} \min(d, d^*)^2 d\theta}{\int_0^{2\pi} \frac{1}{2} \max(d, d^*)^2 d\theta}$$

$$\mathrm{IoU} = \lim_{N \to \infty} \frac{\sum_{i=1}^{N} \frac{1}{2} d_{\min}^2 \Delta\theta_i}{\sum_{i=1}^{N} \frac{1}{2} d_{\max}^2 \Delta\theta_i}$$

$$\text{Polar IoU} = \frac{\sum_{i=1}^{n} d_{\min}}{\sum_{i=1}^{n} d_{\max}}$$

$$\text{Polar IoU Loss} = \log \frac{\sum_{i=1}^{n} d_{\max}}{\sum_{i=1}^{n} d_{\min}}$$

实验结论

- 1) smooth l1 loss很大，需要对weight进行精细调整
- 2) 大物体l1 loss倾向于>小物体的l1 loss
- 3) l1 loss只能对射线进行单独优化，而iou loss对整体进行优化

loss	α	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
Smooth-l_1	0.05	24.7	47.1	23.7	11.3	26.7	36.8
	0.30	25.1	46.4	24.5	10.6	27.3	37.3
	1.00	20.2	37.9	19.6	8.6	20.6	31.1
Polar IoU	1.00	**27.7**	49.6	27.4	12.6	30.2	39.7

其他消融实验

rays	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
18	26.2	48.7	25.4	11.8	28.2	38.0
24	27.3	49.5	26.9	12.4	29.5	40.1
36	**27.7**	49.6	27.4	12.6	30.2	39.7
72	27.6	49.7	27.2	12.9	30.0	39.7

(a) **Number of Rays**: More rays bring a large gain, while too many rays saturate since it already depicts the mask ground-truth well.

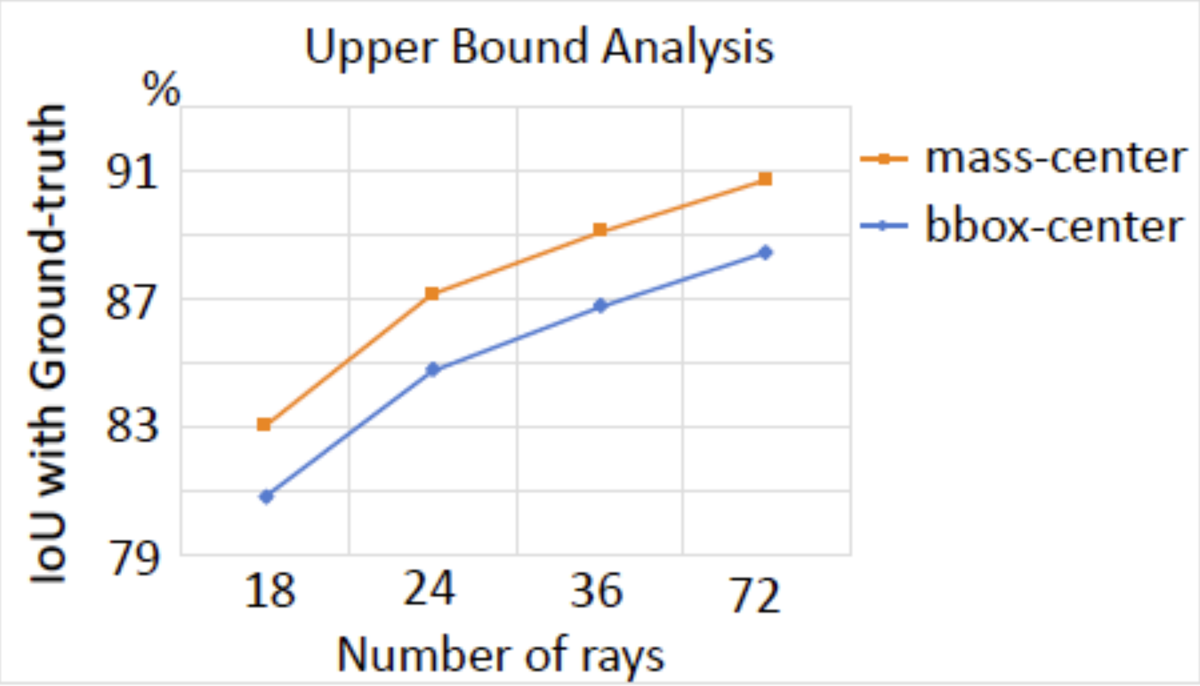
box branch	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
w	27.7	49.6	27.4	12.6	30.2	39.7
w/o	27.5	49.8	27.0	13.0	30.0	40.0

(d) **Box Branch**: Box branch makes no difference to performance of mask prediction.

上限分析

- 相比pixel建模，polar建模对于凹形物体有一定损失(留给futher research…)
- 但是目前stoa的算法距离100AP差很远，因此讨论理论上限意义不大。
- 基于重心和box中心进行

上限分析，发现和gt的平均iou可以到90左右(not bad)



和stoa对比

method	backbone	epochs	aug	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
two-stage									
MNC [7]	ResNet-101-C4	12	◦	24.6	44.3	24.8	4.7	25.9	43.6
FCIS [21]	ResNet-101-C5-dilated	12	◦	29.2	49.5	-	7.1	31.3	50.0
Mask R-CNN [15]	ResNeXt-101-FPN	12	◦	37.1	60.0	39.4	16.9	39.9	53.5
one-stage									
ExtremeNet [34]	Hourglass-104	100	✓	18.9	44.5	13.7	10.4	20.4	28.3
TensorMask [4]	ResNet-101-FPN	72	✓	37.1	59.3	39.4	17.1	39.1	51.6
YOLACT [2]	ResNet-101-FPN	48	✓	31.2	50.6	32.8	12.1	33.3	47.1
PolarMask	ResNet-101-FPN	12	◦	30.4	51.9	31.0	13.4	32.4	42.8
PolarMask	ResNet-101-FPN	24	✓	32.1	53.7	33.1	14.7	33.8	45.3
PolarMask	ResNeXt-101-FPN	12	◦	32.9	55.4	33.8	15.5	35.1	46.3
PolarMask	ResNeXt-101-FPN-DCN	24	✓	36.2	59.4	37.7	17.8	37.7	51.5

Speed, Gflops and more results

method	backbone	size	GFLOPs	Parameters
FCOS [8]	R-50	800 × 1280	200.4	32.02M
PolarMask w/o box	R-50	800 × 1280	202.3	32.09M
PolarMask w/ box	R-50	800 × 1280	252.7	34.46M
YOLACT [1]	R-50	800 × 1280	196.8	31.16M
Mask R-CNN [4]	R-50	800 × 1280	275.5	44.18M

method	backbone	Device	FPS	Time(ms)
PolarMask	R-101	V100	**12.3**	81
PolarMask	R-101-DCN	V100	8.18	122
TensorMask [2]	R-101	V100	2.63	380
Mask R-CNN [4]	R-101	M40	5.12	195

method	backbone	epochs	ms-train	AP
PolarMask	R-50	12	○	29.1
	R-50-DCN	12	○	32.0
	R-50	24	✓	30.5
	R-50-DCN	24	✓	33.3
	R-101	12	○	30.4
	R-101-DCN	12	○	33.5
	R-101	24	✓	31.9
	R-101-DCN	24	✓	34.3
	X-101	12	○	32.6
	X-101-DCN	12	○	34.9
	X-101	24	✓	33.5
	X-101-DCN	24	✓	35.9

总结和展望

总结：

(1) PolarMask 建模比较简单，不需要复杂的操作(DCN, RoIAlign), 对于工业界应用比较友好

(2)找到了一种表达方式，把bbox detection和mask segmentation建模统一了起来，PolarMask本质上可以看成一个目标检测和实例分割统一的框架。只需要简单修改就可以退化到FCOS。

展望：

(1) 通过更好的表达 解决目前建模损失的问题
(2) 设计更好的采样和监督策略
(3) 统一检测分割和人体关键点任务

谢谢观看

Arxiv: https://arxiv.org/abs/1909.13226

Code: https://github.com/xieenze/PolarMask