CenterNet: Keypoint Triplets for Object Detection

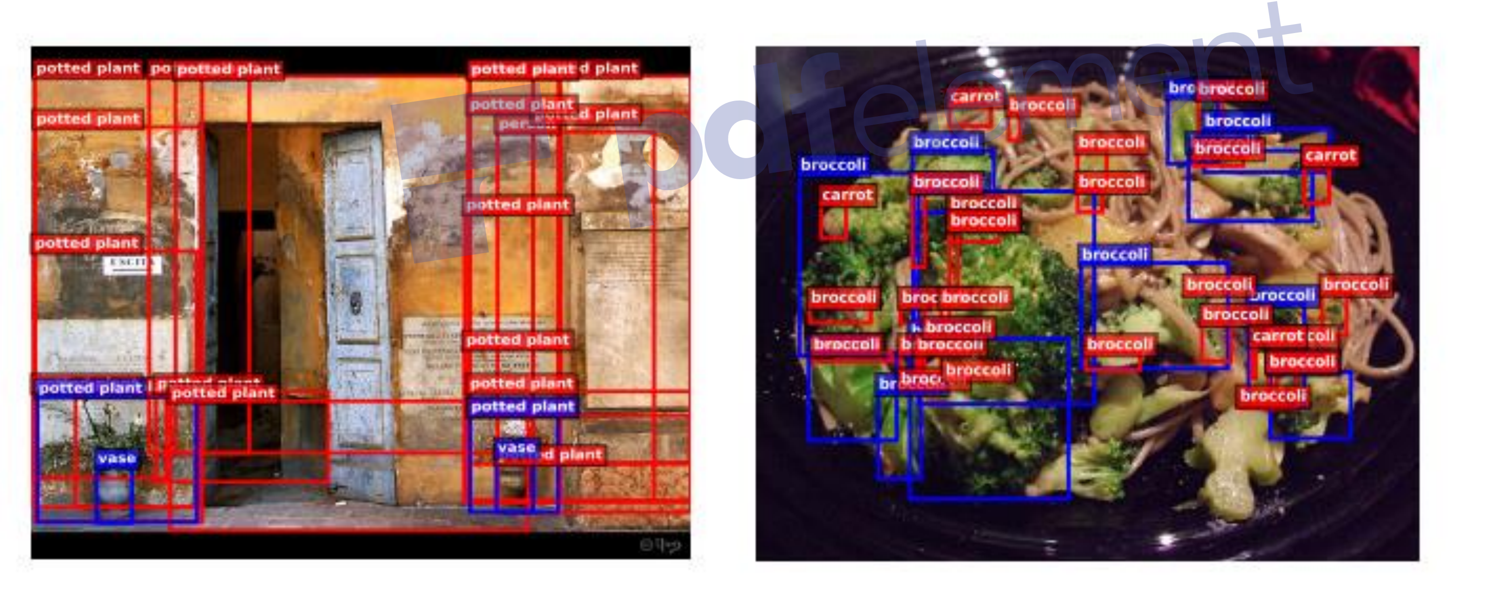
harry
2020.3.6

目录

- 动机
- 方法
  - 整体框架
  - Cascade Pooling And Center Pooling
  - 后处理
- 实验与结果

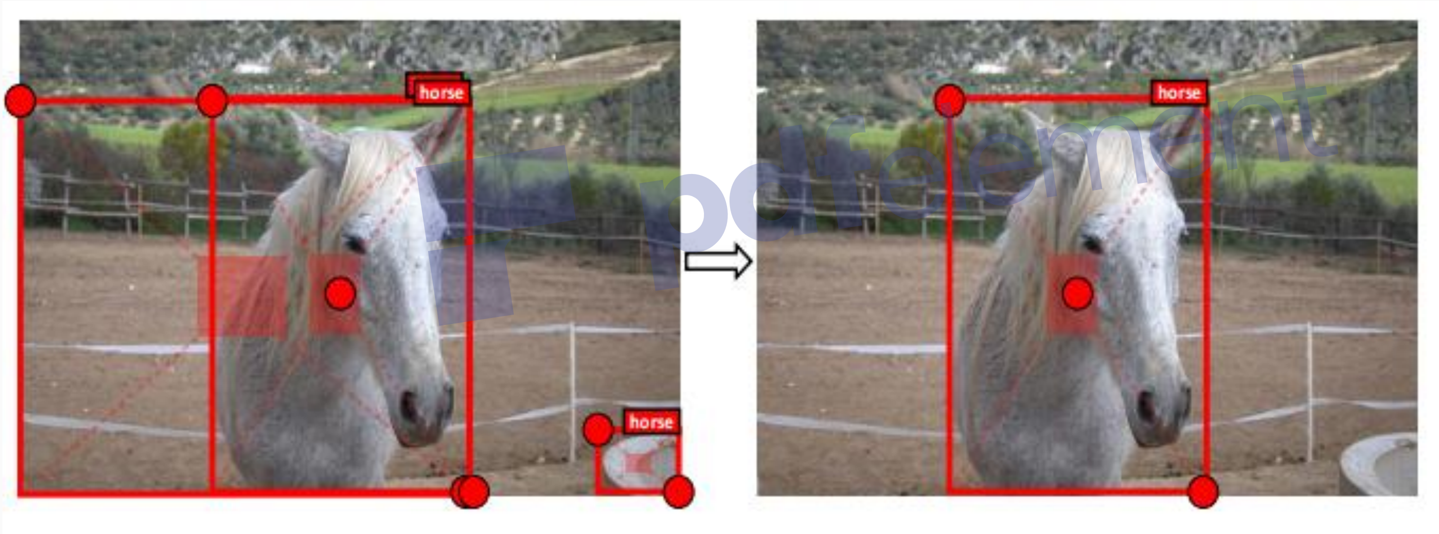
动机

- CornerNet由于是根据角点来检测目标，缺少目标内部信息，容易产生假阳性。



动机

- 新增一个分支用于预测目标的中心点用于预测增强识别信息与过滤假样目标



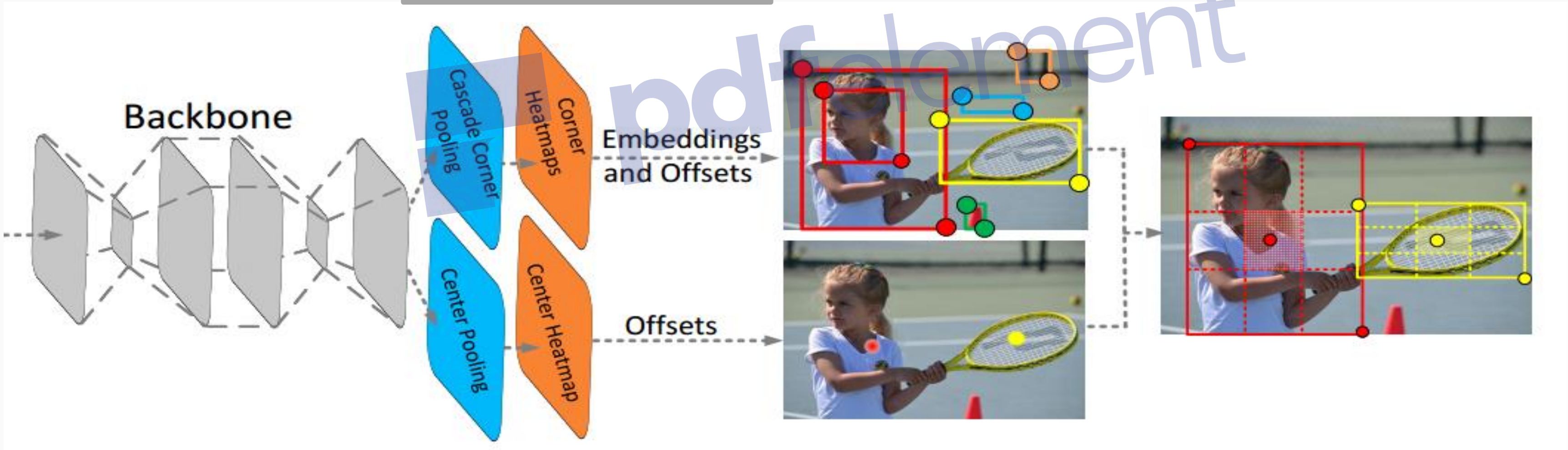
边框对应的中心点：若存在相应类别的响应，则保留。

方法：整体框架

Hourglass-52 or 104

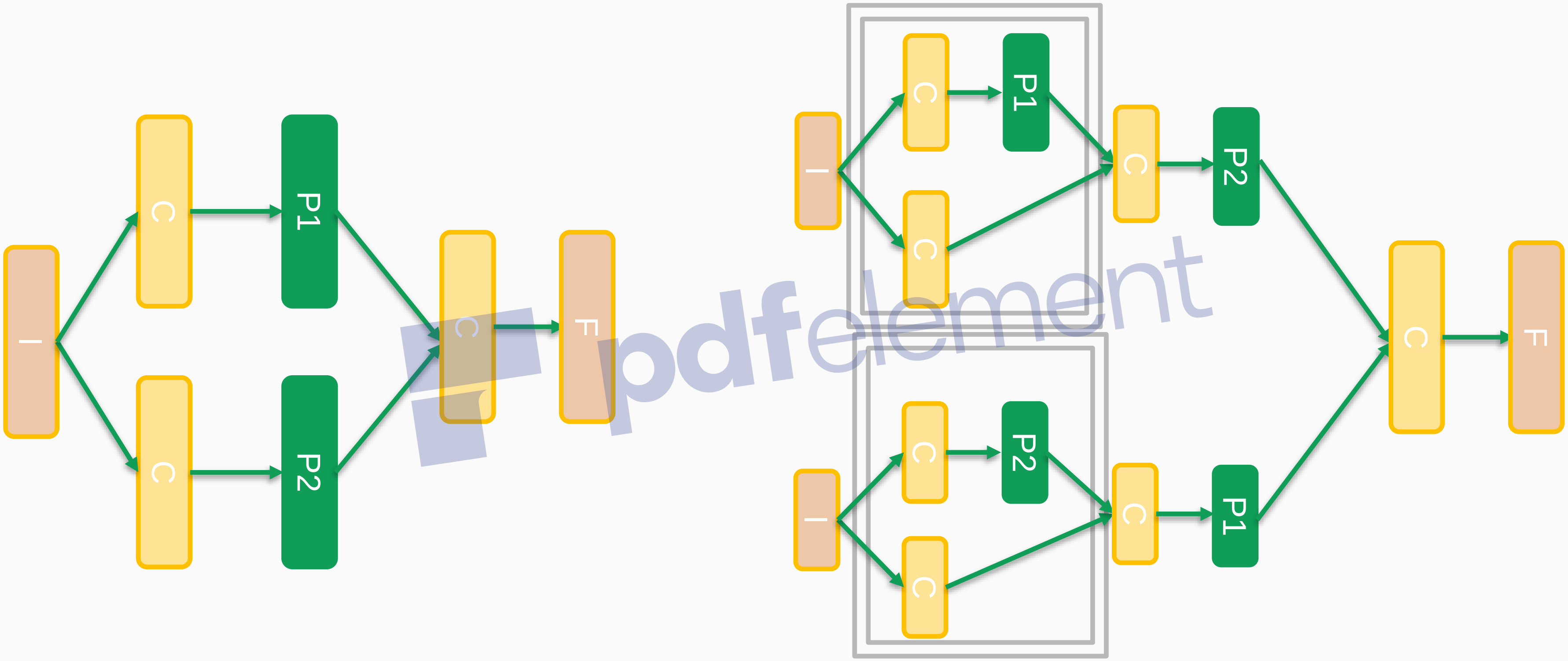
top-left, bottom-right:
Classes
2(reg)
1(embedding feature)

Post-Processing



方法：整体框架

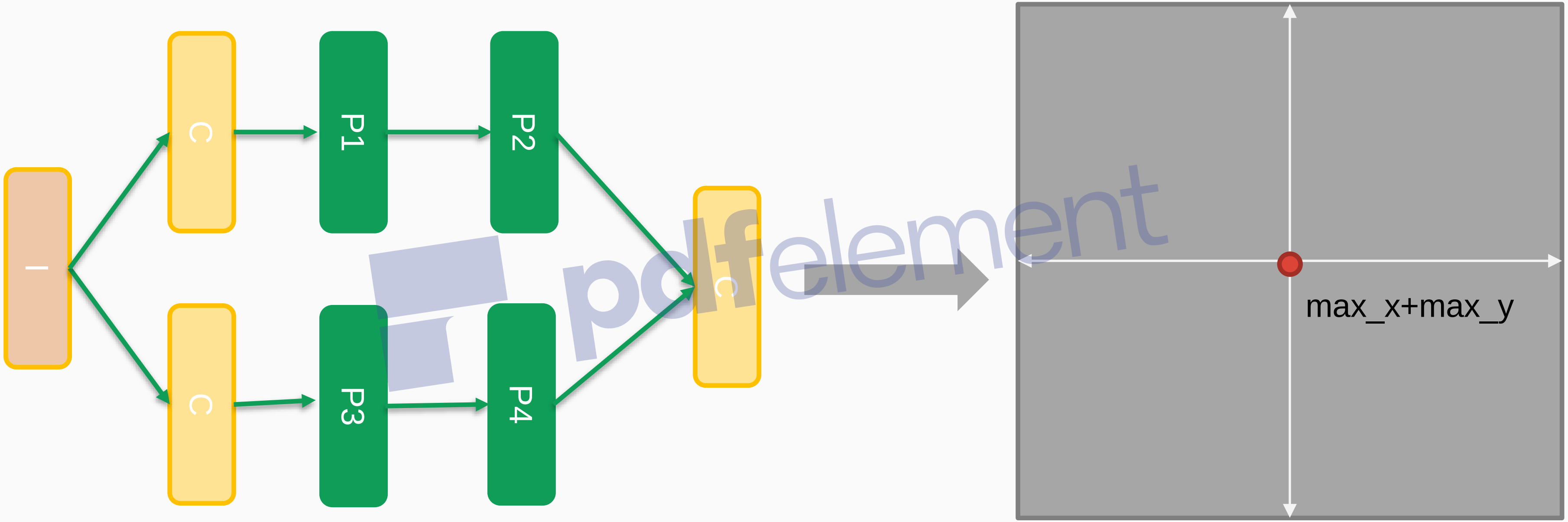
方法：Corner pooling Vs Cascade Corner Pooling



增强了边缘特性，同时减少了false posivite。

方法：Corner pooling Vs Cascade Corner Pooling

方法：Center pooling



方法：后处理

从tl-heatmap and br-heatmap 选取top-70个点，同时使用offset来remap

通过embedding features来group

根据bbox中心位置的响应来决定其是否保留，分数取三者平均

做NMS，取分数最高的前100个bbox，

当bbox的scale小于150时，n取3；大于150时，n取5

$$\begin{cases} \mathrm{ctl_x} = \dfrac{(n+1)\mathrm{tl_x} + (n-1)\mathrm{br_x}}{2n} \\ \mathrm{ctl_y} = \dfrac{(n+1)\mathrm{tl_y} + (n-1)\mathrm{br_y}}{2n} \\ \mathrm{cbr_x} = \dfrac{(n-1)\mathrm{tl_x} + (n+1)\mathrm{br_x}}{2n} \\ \mathrm{cbr_y} = \dfrac{(n-1)\mathrm{tl_y} + (n+1)\mathrm{br_y}}{2n} \end{cases}$$

实验与结果

消融实验：

- CRE: 加入center分支
- CTP: center pooling
- CCP: cascade corner pooling

CRE	CTP	CCP	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L	AR_1	AR_{10}	AR_{100}	AR_S	AR_M	AR_L
			37.6	53.3	40.0	18.5	39.6	52.2	33.7	52.2	56.7	37.2	60.0	74.0
		✓	38.3	54.2	40.5	18.6	40.5	52.2	34.0	53.0	57.9	36.6	60.8	75.8
✓			39.9	57.7	42.3	23.1	42.3	52.3	33.8	54.2	58.5	38.7	62.4	74.4
✓	✓		40.8	58.6	43.6	23.6	43.6	53.6	33.9	54.5	59.0	39.0	63.2	74.7
✓	✓	✓	**41.3**	**59.2**	**43.9**	**23.6**	**43.8**	**55.8**	**34.5**	**55.0**	**59.2**	**39.1**	**63.5**	**75.1**

Table 4: Ablation study on the major components of CenterNet511-52 on the MS-COCO validation dataset. The CRE denotes central region exploration, the CTP denotes center pooling, and the CCP denotes cascade corner pooling.

实验与结果

Method	Backbone	Train input	Test input	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L	AR_1	AR_{10}	AR_{100}	AR_S	AR_M	AR_L
Mask R-CNN [12]	ResNeXt-101	$\sim 1300\times800$	$\sim 1300\times800$	39.8	62.3	43.4	22.1	43.2	51.2	-	-	-	-	-	-
Soft-NMS [2]	Aligned-Inception-ResNet	$\sim 1300\times800$	$\sim 1300\times800$	40.9	62.8	-	23.3	43.6	53.3	-	-	-	-	-	-
Fitness R-CNN [41]	ResNet-101	512×512	1024×1024	41.8	60.9	44.9	21.5	45.0	57.5	-	-	-	-	-	-
Cascade R-CNN [4]	ResNet-101	-	-	42.8	62.1	46.3	23.7	45.5	55.2	-	-	-	-	-	-
Grid R-CNN w/ FPN [28]	ResNeXt-101	$\sim 1300\times800$	$\sim 1300\times800$	43.2	63.0	46.6	25.1	46.5	55.2	-	-	-	-	-	-
D-RFCN + SNIP (multi-scale) [38]	DPN-98 [5]	$\sim 2000\times1200$	$\sim 2000\times1200$	45.7	**67.3**	51.1	29.3	48.8	57.1	-	-	-	-	-	-
PANet (multi-scale) [26]	ResNeXt-101	$\sim 1400\times840$	$\sim 1400\times840$	**47.4**	67.2	**51.8**	**30.1**	**51.7**	**60.0**	-	-	-	-	-	-
CornerNet511 (multi-scale) [20]	Hourglass-52	511×511	$\leq1.5\times$	39.4	54.9	42.3	18.9	41.2	52.7	35.0	53.5	57.7	36.1	60.1	75.1
CornerNet511 (single-scale) [20]	Hourglass-104	511×511	ori.	40.5	56.5	43.1	19.4	42.7	53.9	35.3	54.3	59.1	37.4	61.9	76.9
RefineDet512 (multi-scale) [45]	ResNet-101	512×512	$\leq2.25\times$	41.8	62.9	45.7	25.6	45.1	54.1						
CornerNet511 (multi-scale) [20]	Hourglass-104	511×511	$\leq1.5\times$	42.1	57.8	45.3	20.8	44.8	56.7	36.4	55.7	60.0	38.5	62.7	77.4
CenterNet511 (single-scale)	Hourglass-52	511×511	ori.	41.6	59.4	44.2	22.5	43.1	54.1	34.8	55.7	60.1	38.6	63.3	76.9
CenterNet511 (single-scale)	Hourglass-104	511×511	ori.	44.9	62.4	48.1	25.6	47.4	57.4	36.1	58.4	63.3	41.3	67.1	80.2
CenterNet511 (multi-scale)	Hourglass-52	511×511	$\leq1.8\times$	43.5	61.3	46.7	25.3	45.3	55.0	36.0	57.2	61.3	41.4	64.0	76.3
CenterNet511 (multi-scale)	Hourglass-104	511×511	$\leq1.8\times$	**47.0**	**64.5**	**50.7**	**28.9**	**49.9**	**58.9**	**37.5**	**60.3**	**64.8**	**45.1**	**68.3**	**79.7**

模型性能很好，几乎于二阶段持平（同样有很多trick）；但是目前由于主干网络与Corner Pooling，速度很慢

补充实验

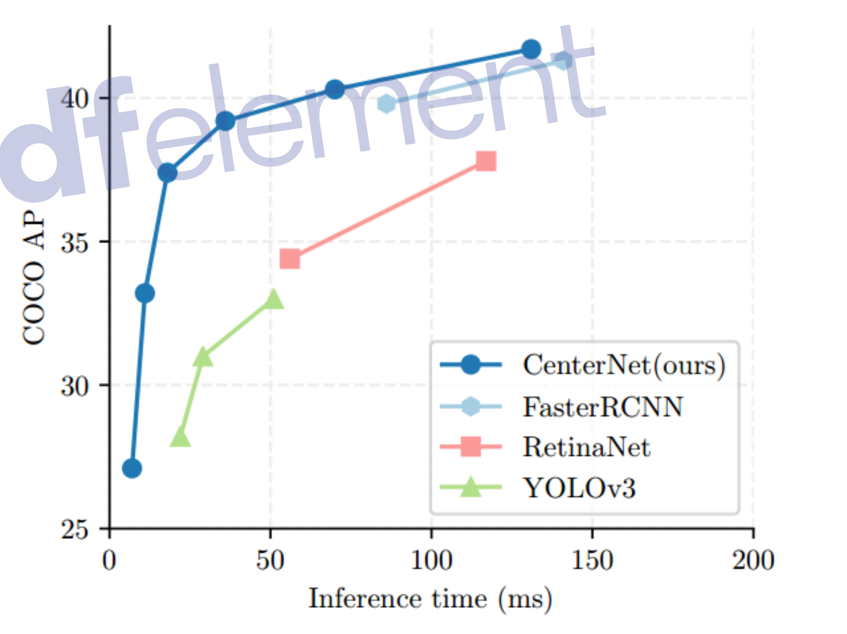
将预测的Center Heatmap换成真实的Heatmap模型性能的提升。

Method	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
CenterNet511-52 w/o GT	41.3	59.2	43.9	23.6	43.8	55.8
CenterNet511-52 w/ GT	**56.5**	**78.3**	**61.4**	**39.1**	**60.3**	**70.3**
CenterNet511-104 w/o GT	44.8	62.4	48.2	25.9	48.9	58.8
CenterNet511-104 w/ GT	**58.1**	**78.4**	**63.9**	**40.4**	**63.0**	**72.1**

CenterNet: Objects as Points

目录

- 动机
- 方法
- 实验与结果



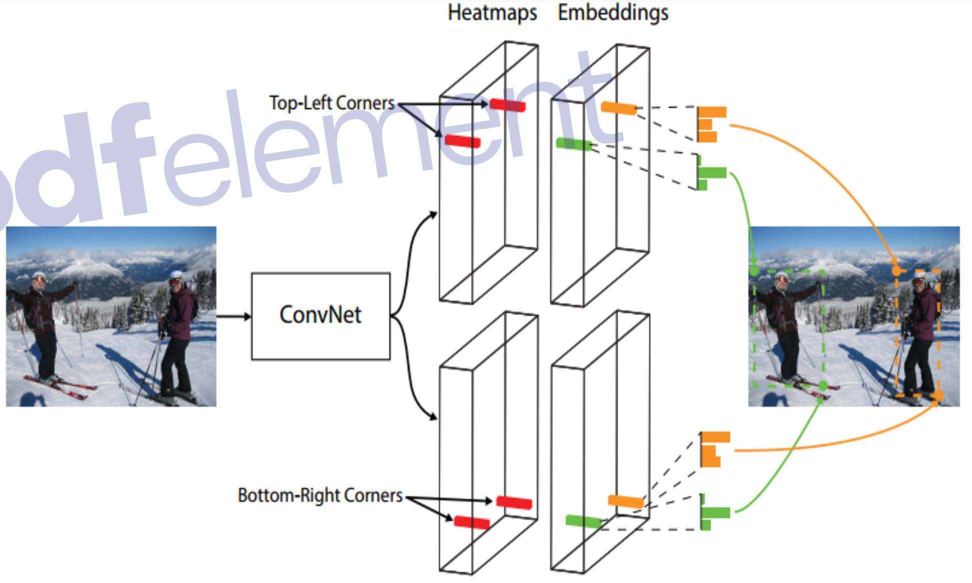
动机

- CornerNet由于是根据角点来检测目标，后处理中会涉及配对操作，处理流程复杂，效率低。

从tl-heatmap and br-heatmap选取top-70个点，同时使用offset来remap

通过embedding features来group

做NMS，取分数最高的前100个bbox，



方法

- CenterNet将检测任务分解成中心点检测与长宽回归。

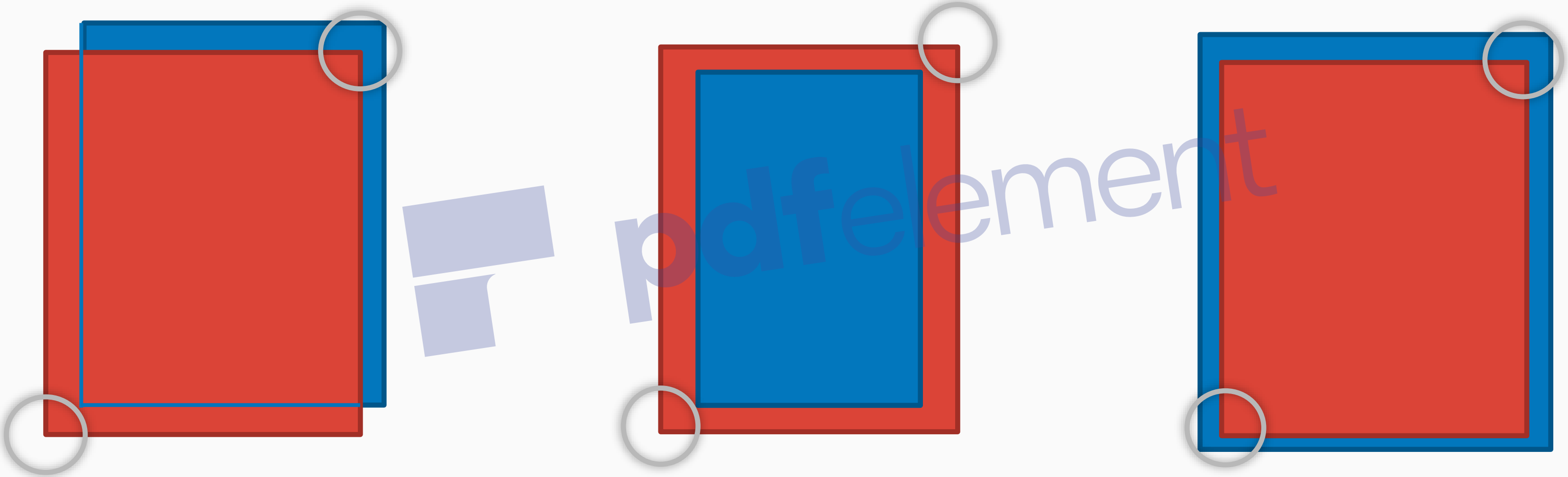
因此网络最终输出categories+4(offsets, height, width)个heatmaps。同时没有NMS, 没有复杂的后处理。

后处理：保留比其8邻域大的响应。（使用maxpool就可以解决）

方法：整体框架

方法：Radius

$r = \mathrm{argmax}(iou>t)(t=0.7)$

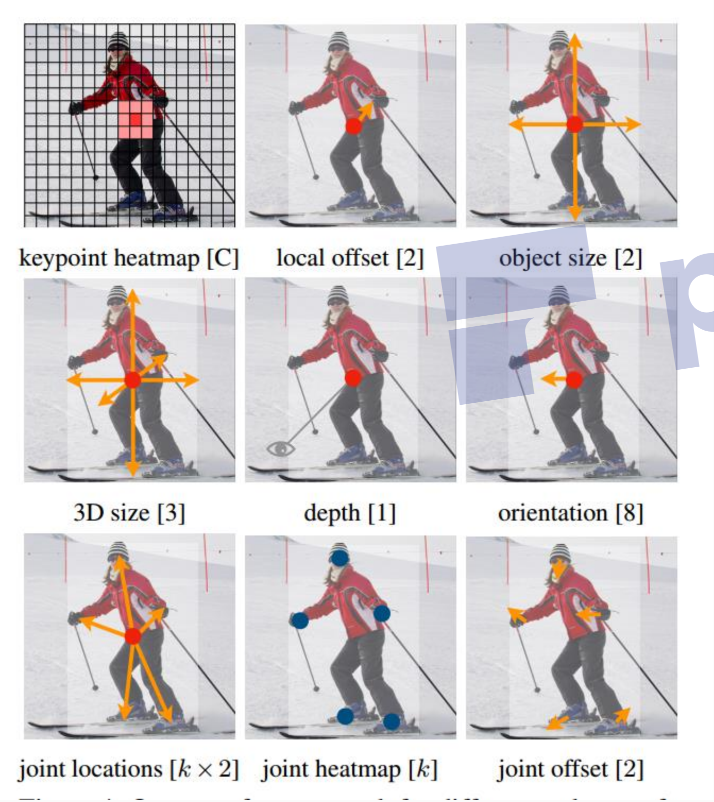


r1 r2 r3

$r = \min(r1, r2, r3)$

方法

- 作者将模型泛化至3D检测，关键点检测。将任务分解成中心点与其属性。



目标检测：目标中心+尺寸

3D检测：目标中心+尺寸+深度信息+方向

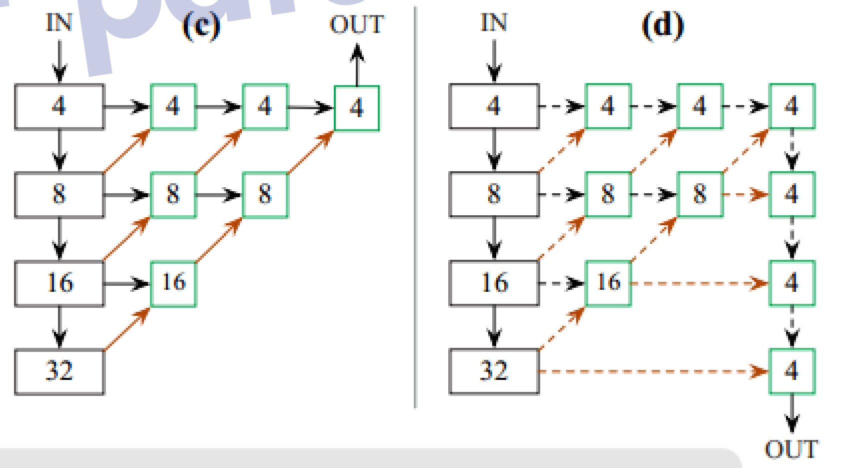
关键点检测：目标中心+k个节点的offset

实验

1.使用不同backbone的结果（其他模型使用了预训练模型）

	AP			AP_{50}			AP_{75}			Time (ms)			FPS		
	N.A.	F	MS	N.A.	F	MS	N.A.	F	MS	N.A.	F	MS	N.A.	F	MS
Hourglass-104	**40.3**	**42.2**	**45.1**	**59.1**	**61.1**	**63.5**	**44.0**	**46.0**	**49.3**	71	129	672	14	7.8	1.4
DLA-34	37.4	39.2	41.7	55.1	57.0	60.1	40.8	42.7	44.9	19	36	248	52	28	4
ResNet-101	34.6	36.2	39.3	53.0	54.8	58.5	36.9	38.7	42.0	22	40	259	45	25	4
ResNet-18	28.1	30.0	33.2	44.9	47.5	51.5	29.6	31.6	35.1	**7**	**14**	**81**	**142**	**71**	**12**

左为原始的DLA;
右为作者修改后的DLA。



实验

1.其他消融实验

Resolution	AP	AP_{50}	AP_{75}	Time
Original	**36.3**	54.0	**39.6**	19
512	36.2	**54.3**	38.7	16
384	33.2	50.5	35.0	**11**

测试时输入尺寸的影响

λ_{size}	AP	AP_{50}	AP_{75}
0.2	33.5	49.9	36.2
0.1	**36.3**	54.0	**39.6**
0.02	35.4	**54.6**	37.9

平衡系数的影响

Loss	AP	AP_{50}	AP_{75}
l1	**36.3**	**54.0**	**39.6**
smooth l1	33.9	50.9	36.8

回归损失的对比

Epoch	AP	AP_{50}	AP_{75}
140	36.3	54.0	39.6
230	**37.4**	**55.1**	**40.8**

训练时长的影响

与YOLOV1的比较

	YOLOV1	CenterNet	Diff
输入与输出的设置	结果为7*7的Heatmaps 5*2+20 (x, y, w, h, c)，其中x,y是相对于cell左上点的。w, h是相对于整张图的。	结果为128*128的Heatmaps 80+2+2	softmax vs sigmoid points and Gaussian map
网络与损失的设置	相当于VGG19，没有BN	HourglassNet,点检测领域的标配。	Cross entropy loss vs focal loss

错误分析

	AP	AP_{50}	AP_{75}
	36.3	54.0	39.6
w/ gt size	41.9	56.6	45.4
w/ gt heatmap	54.2	82.6	58.1
w/ gt heatmap+size	83.1	97.9	90.1
w/ gt hm.+size+offset	99.5	99.7	99.6

- 中心点预测的很差。
- 中心点预测准的前提下，谈论size才更加具有意义。