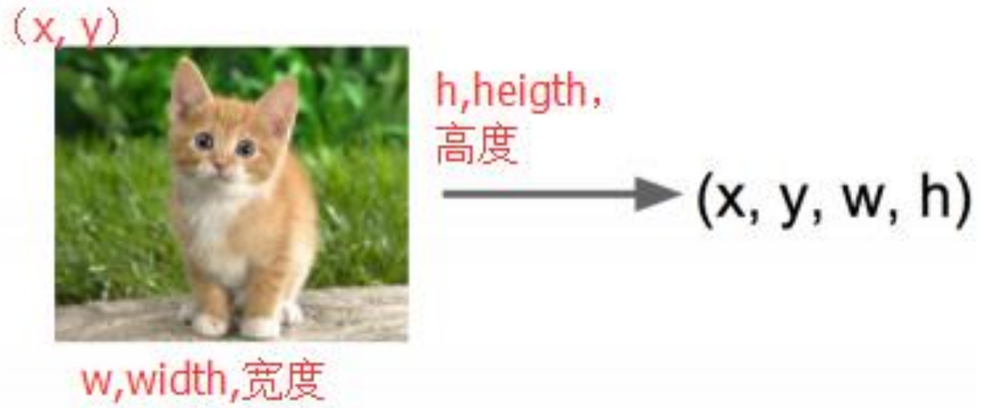
CornerNet——基于关键点的目标检测算法

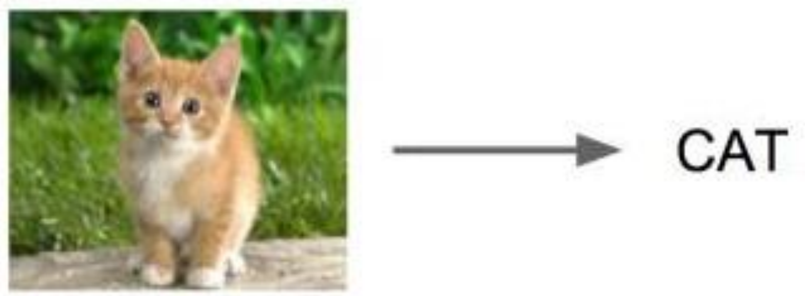
汇报人：范越巧

什么是目标检测？

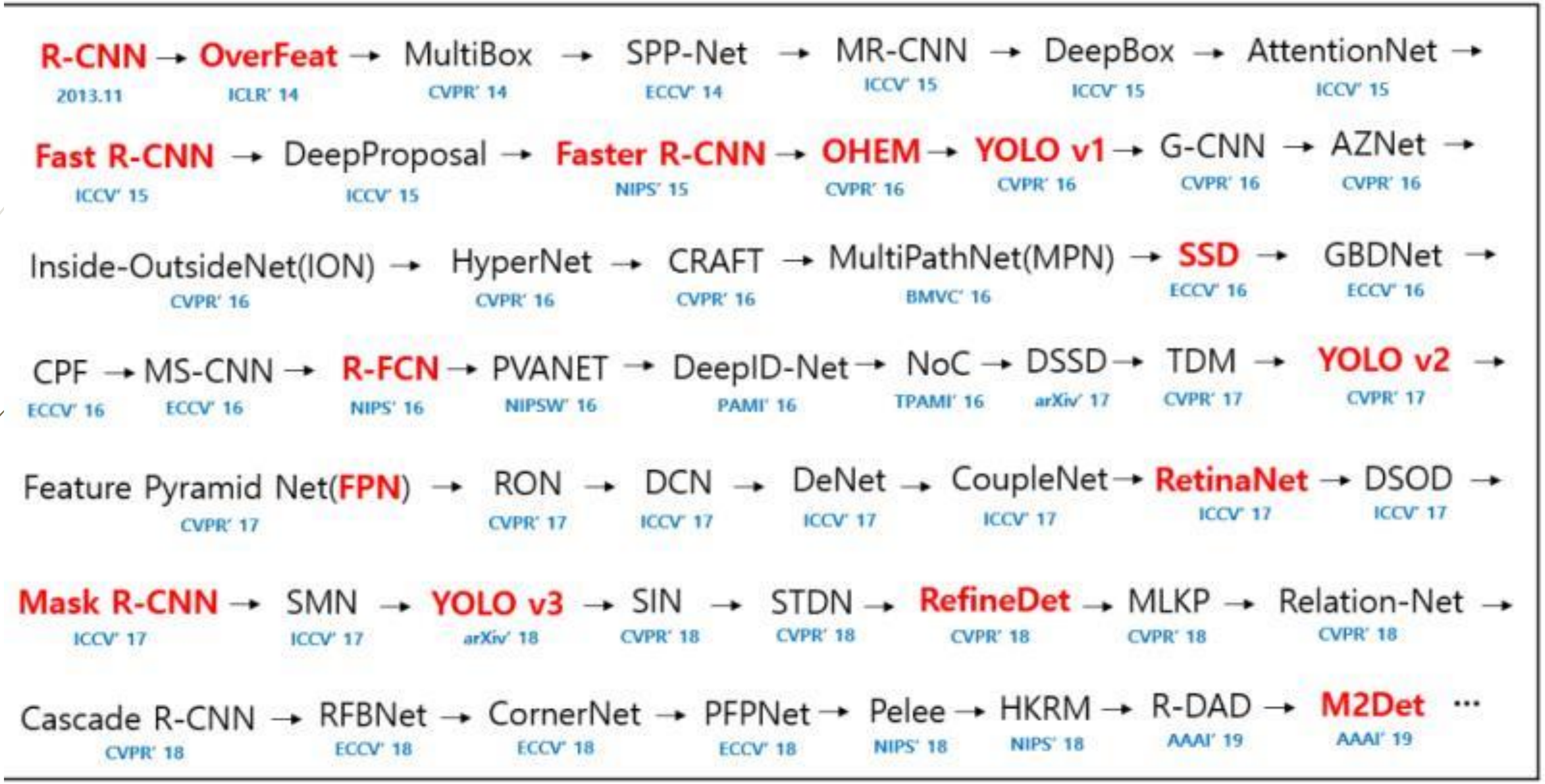
定位（localization）：



图像识别（classification）：



视觉系



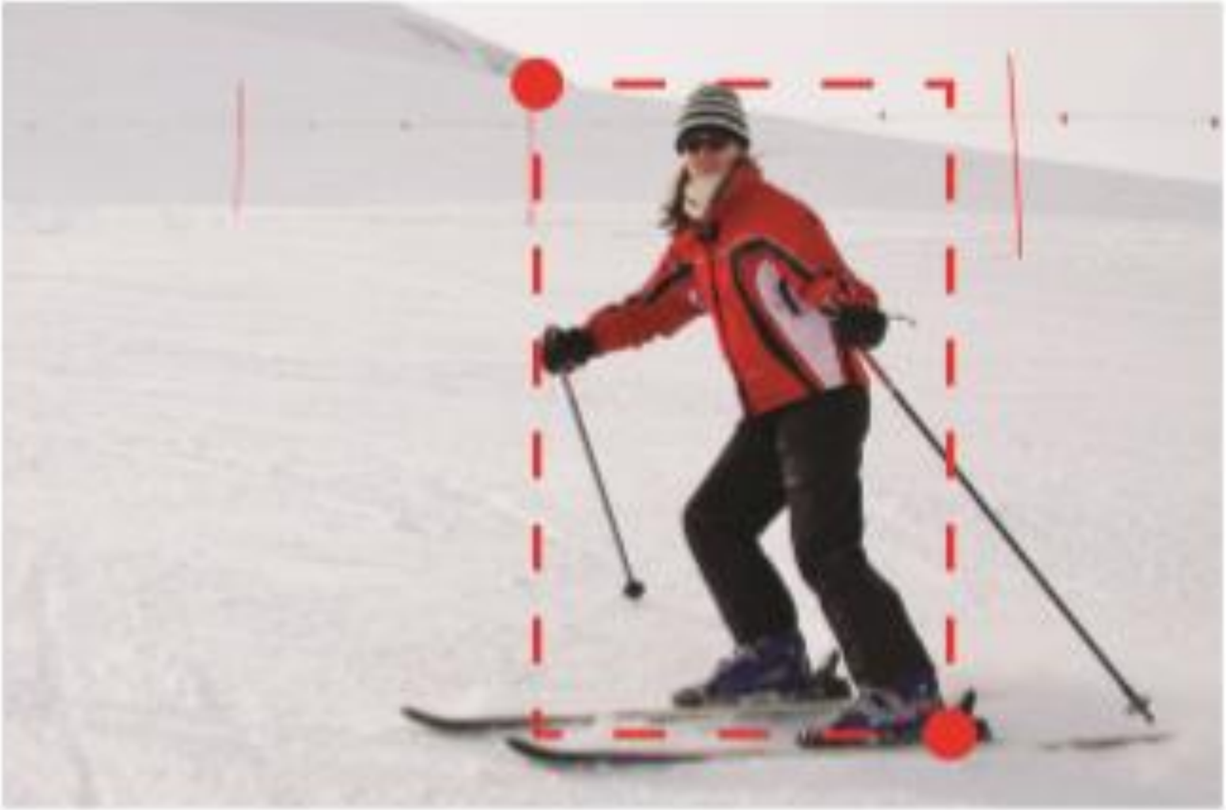
R-CNN →	**OverFeat** →	MultiBox →	SPP-Net →	MR-CNN →	DeepBox →	AttentionNet →		
2013.11	ICLR' 14	CVPR' 14	ECCV' 14	ICCV' 15	ICCV' 15	ICCV' 15		
Fast R-CNN →	DeepProposal →	**Faster R-CNN** →	**OHEM** →	**YOLO v1** →	G-CNN →	AZNet →		
ICCV' 15	ICCV' 15	NIPS' 15	CVPR' 16	CVPR' 16	CVPR' 16	CVPR' 16		
Inside-OutsideNet(ION) →	HyperNet →	CRAFT →	MultiPathNet(MPN) →	**SSD** →	GBDNet →			
CVPR' 16	CVPR' 16	CVPR' 16	BMVC' 16	ECCV' 16	ECCV' 16			
CPF →	MS-CNN →	**R-FCN** →	PVANET →	DeepID-Net →	NoC →	DSSD →	TDM →	**YOLO v2** →
ECCV' 16	ECCV' 16	NIPS' 16	NIPSW' 16	PAMI' 16	TPAMI' 16	arXiv' 17	CVPR' 17	CVPR' 17
Feature Pyramid Net(**FPN**) →	RON →	DCN →	DeNet →	CoupleNet →	**RetinaNet** →	DSOD →		
CVPR' 17	CVPR' 17	ICCV' 17	ICCV' 17	ICCV' 17	ICCV' 17	ICCV' 17		
Mask R-CNN →	SMN →	**YOLO v3** →	SIN →	STDN →	**RefineDet** →	MLKP →	Relation-Net →	
ICCV' 17	ICCV' 17	arXiv' 18	CVPR' 18	CVPR' 18	CVPR' 18	CVPR' 18	CVPR' 18	
Cascade R-CNN →	RFBNet →	CornerNet →	PFPNet →	Pelee →	HKRM →	R-DAD →	**M2Det**	···
CVPR' 18	ECCV' 18	ECCV' 18	ECCV' 18	NIPS' 18	NIPS' 18	AAAI' 19	AAAI' 19	

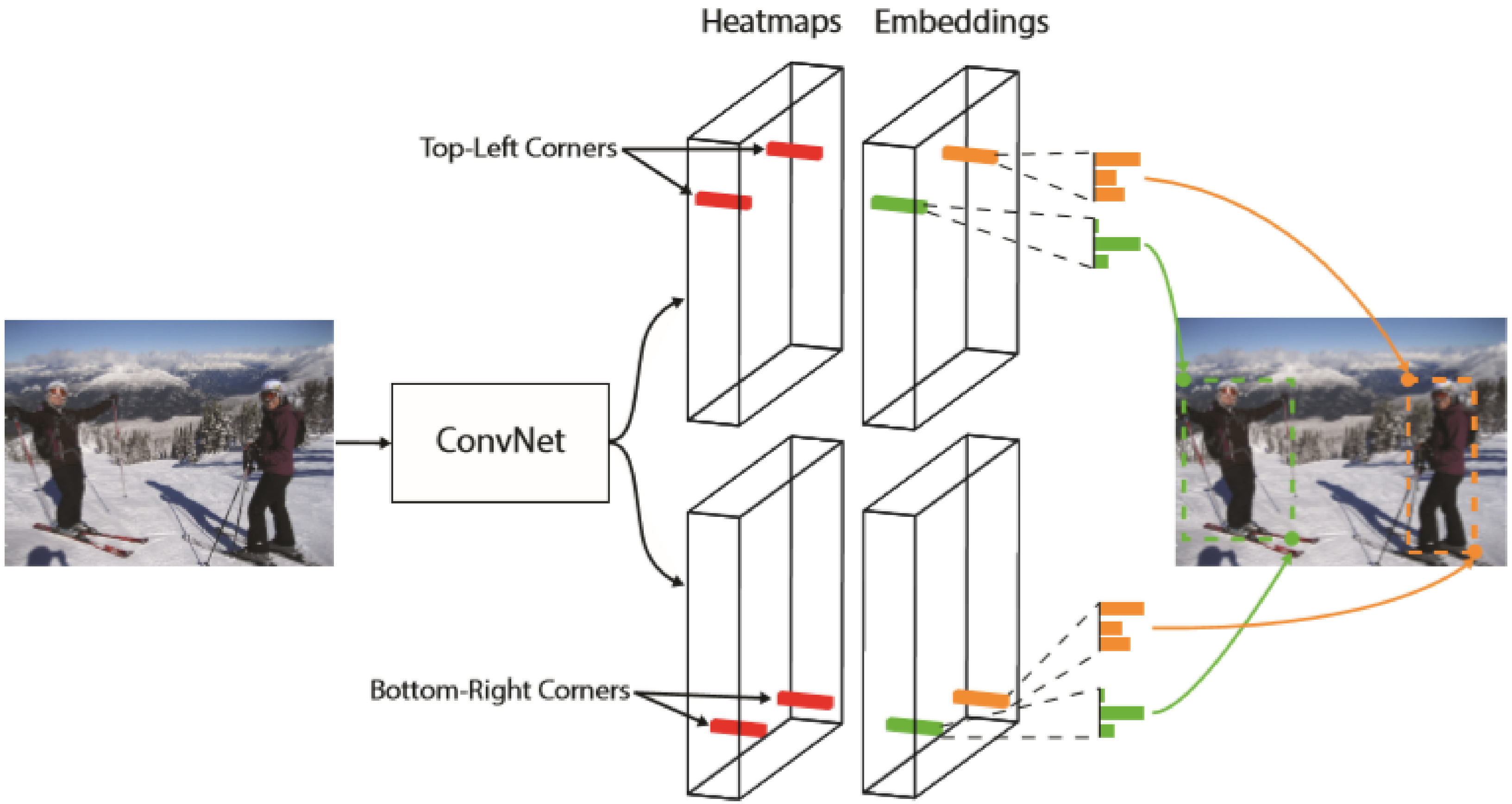
CornerNet: Detecting Objects as Paired Keypoints

Abstract :We propose CornerNet, a new approach to object detection where we detect an object bounding box as a pair of keypoints, the top-left corner and the bottom-right corner, using a single convolution neural network. By detecting objects as paired keypoints, we eliminate the need for designing a set of anchor boxes commonly used in prior single-stage detectors. In addition to our novel formulation, we introduce corner pooling, a new type of pooling layer that helps the network better localize corners. Experiments show that CornerNet achieves a 42.2% AP on MS COCO, outperforming all existing one-stage detectors.

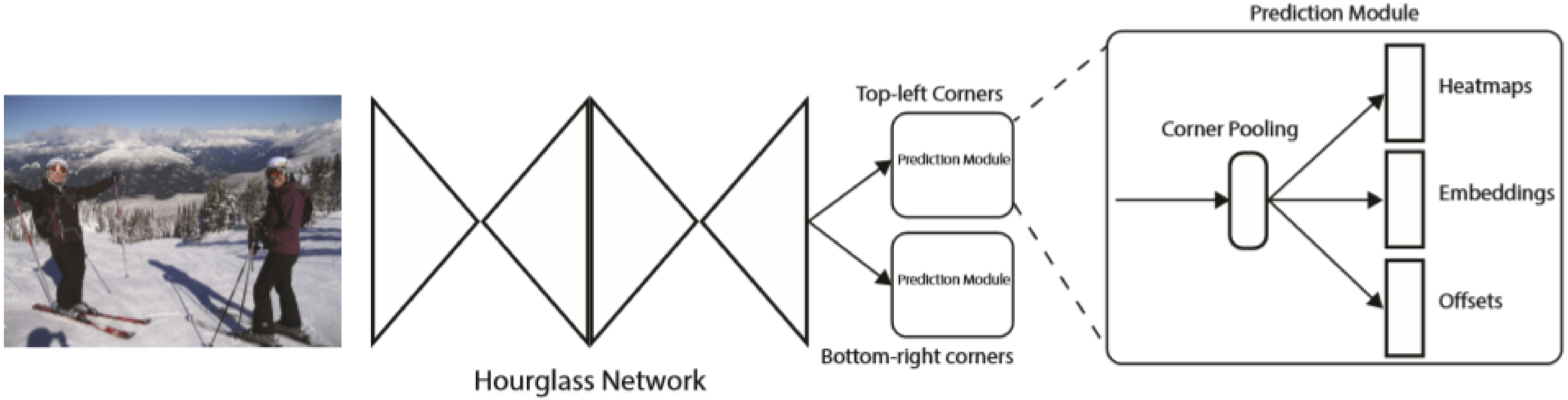
H. Law and J. Deng. Cornernet: Detecting objects as paired keypoints. In Proceedings of the European Conference on Computer Vision (ECCV), pages 734–750, 2018.

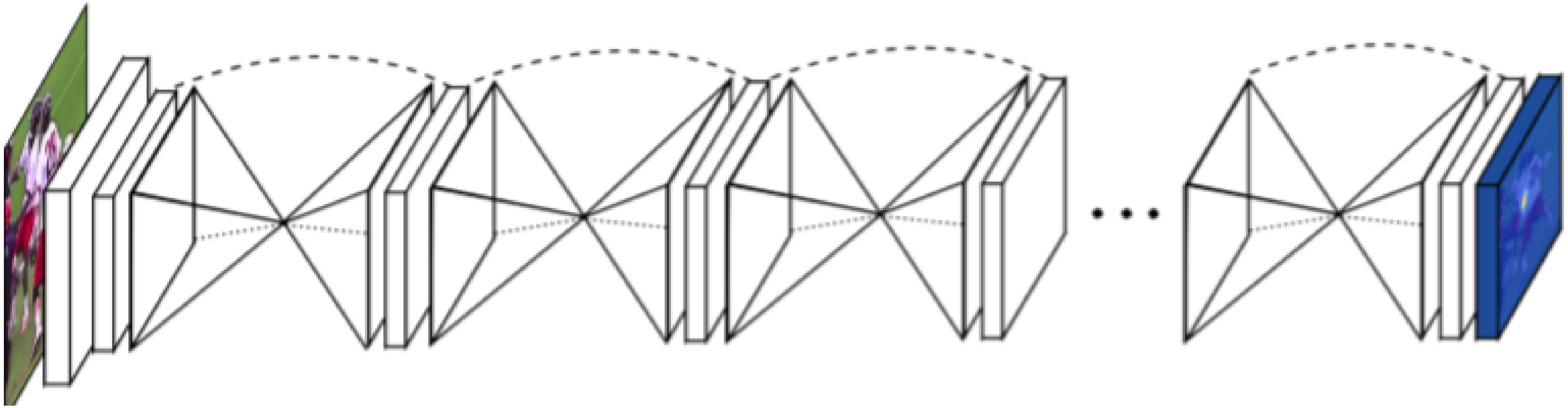
CornerNet: Detecting Objects as Paired Keypoints



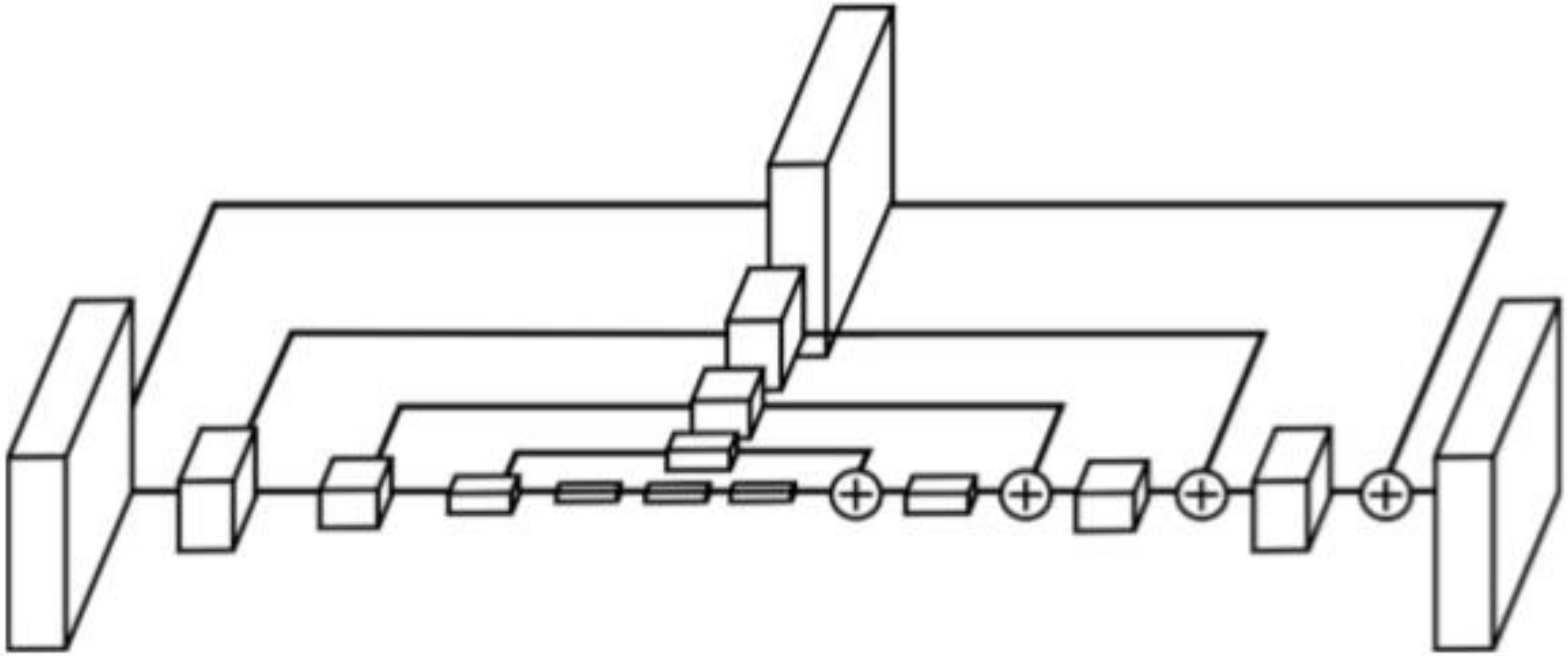

Heatmaps
Embeddings
Top-Left Corners
ConvNet
Bottom-Right Corners

济

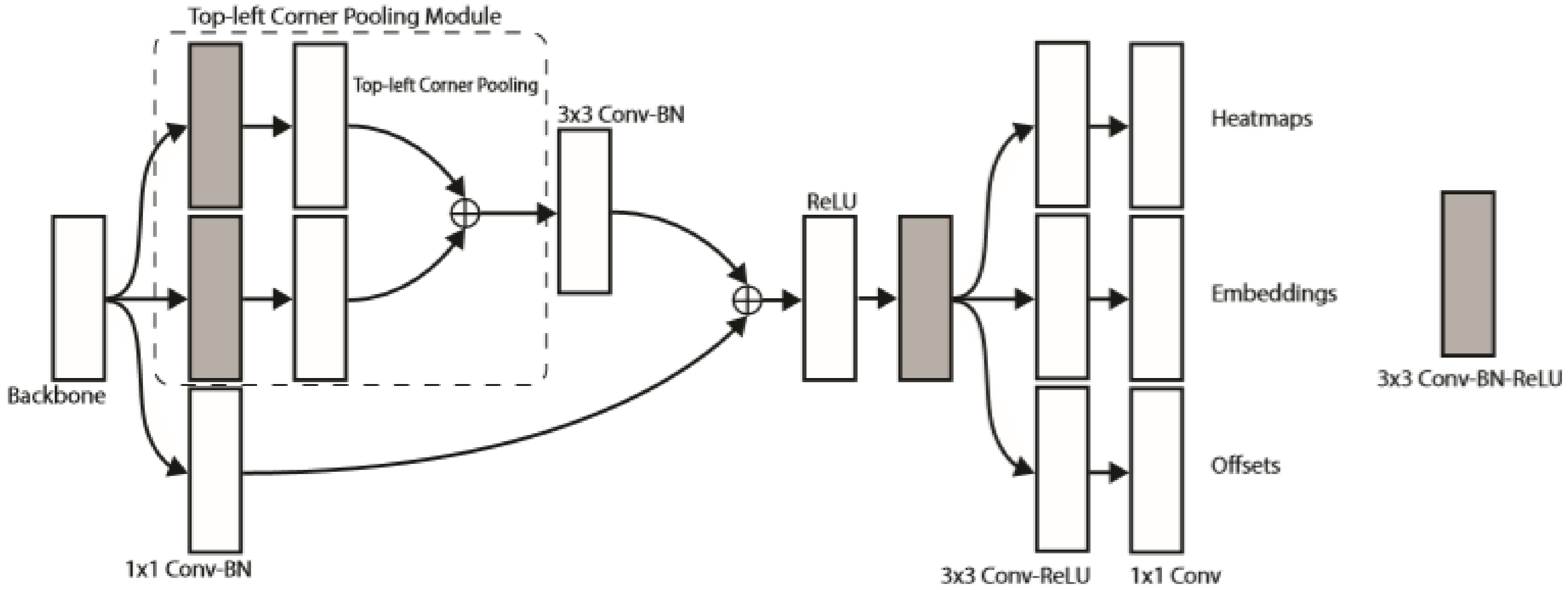


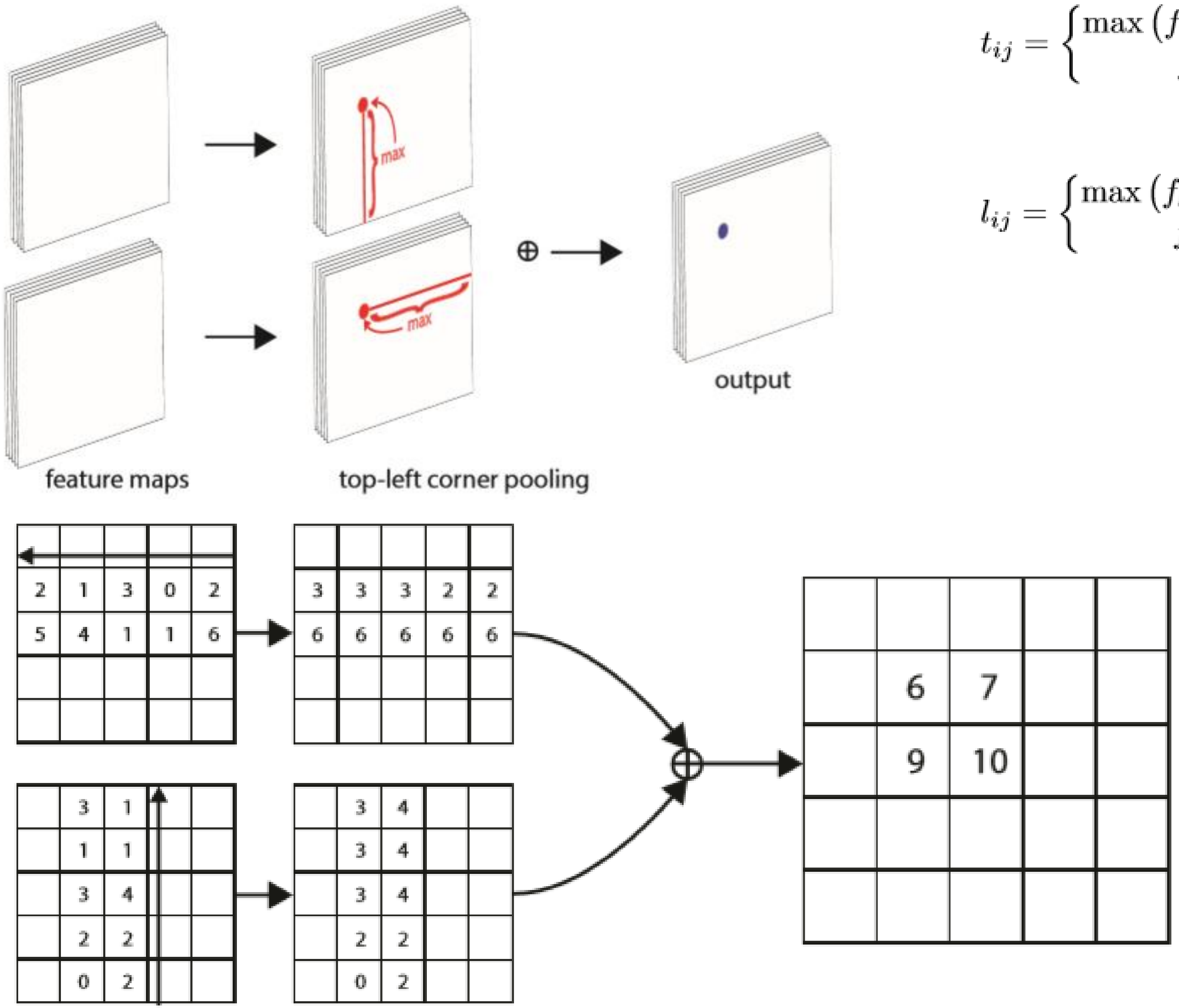


Newell, A., Yang, K., and Deng, J. (2016). Stacked hourglass networks for human pose estimation. In European Conference on Computer Vision, pages 483–499. Springer.



Prediction Module





2	1	3	0	2
5	4	1	1	6

3	3	3	2	2
6	6	6	6	6

	3	1		
	1	1		
	3	4		
	2	2		
	0	2		

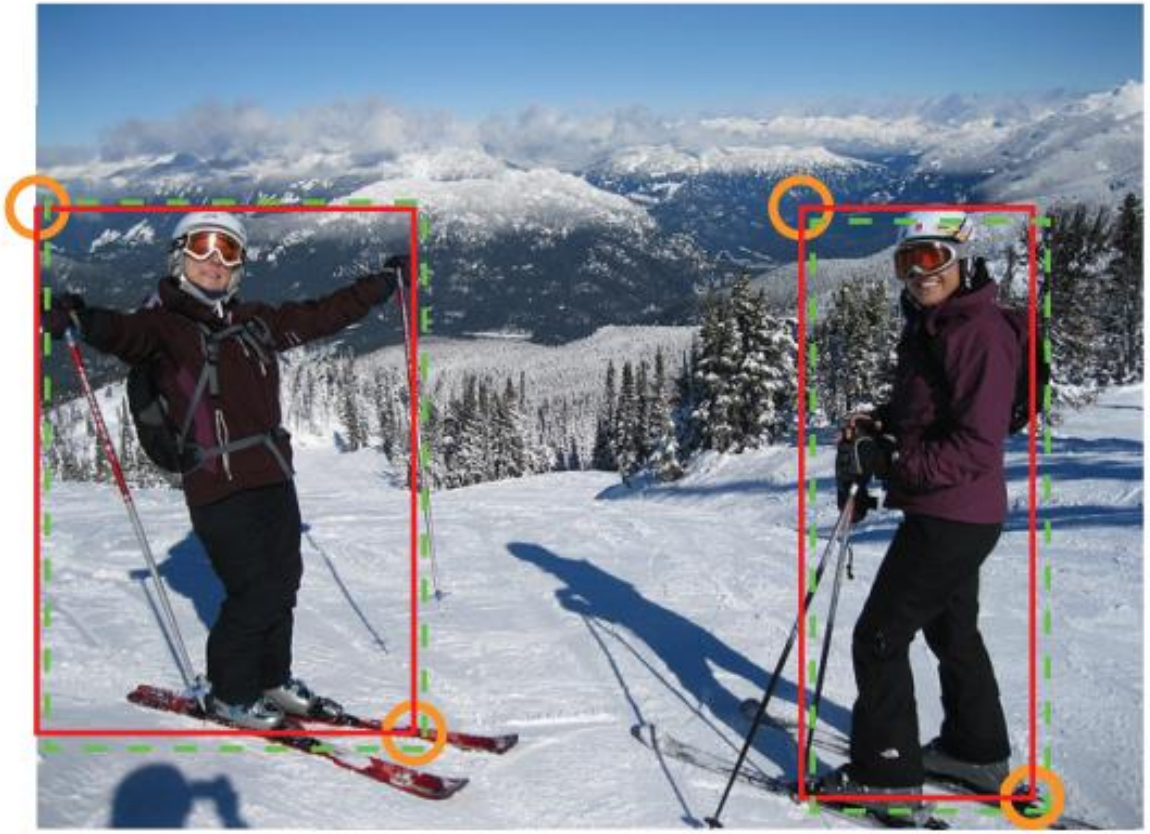
	3	4		
	3	4		
	3	4		
	2	2		
	0	2		

	6	7		
	9	10		

$$t_{ij} = \begin{cases} \max\left(f_{t_{ij}}, t_{(i+1)j}\right) & \text{if } i < H \\ f_{t_{Hj}} & \text{otherwise} \end{cases}$$

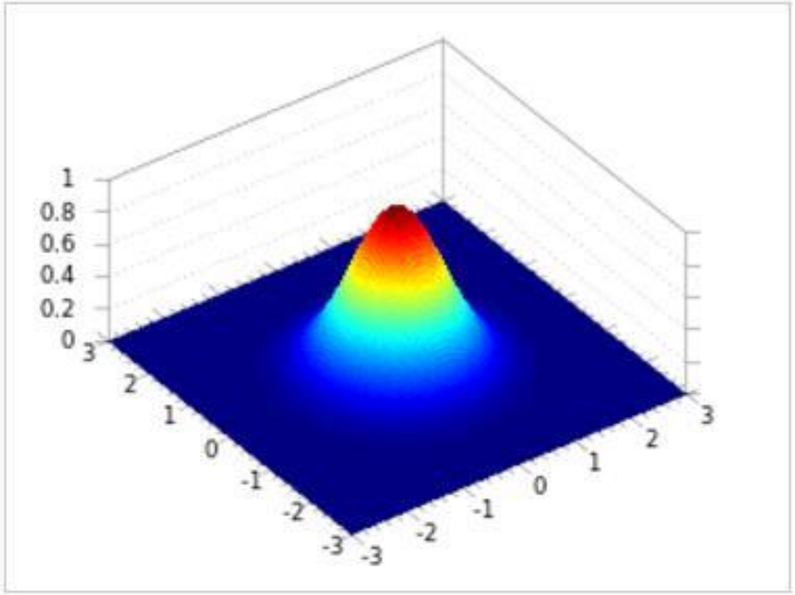
$$l_{ij} = \begin{cases} \max\left(f_{l_{ij}}, l_{i(j+1)}\right) & \text{if } j < W \\ f_{l_{iW}} & \text{otherwise} \end{cases}$$

CornerNet——corners



$$e^{\frac{-(x^2+y^2)}{2\sigma^2}}$$

$$\sigma = \frac{1}{3}r$$



$$L_{det} = \frac{-1}{N} \sum_{c=1}^{C} \sum_{i=1}^{H} \sum_{j=1}^{W} \begin{cases} (1 - p_{cij})^{\alpha} \log (p_{cij}) & \text{if } y_{cij} = 1 \\ \underbrace{(1 - y_{cij})^{\beta}}_{\text{平衡因子}} (p_{cij})^{\alpha} \log (1 - p_{cij}) & \text{otherwise} \end{cases}$$

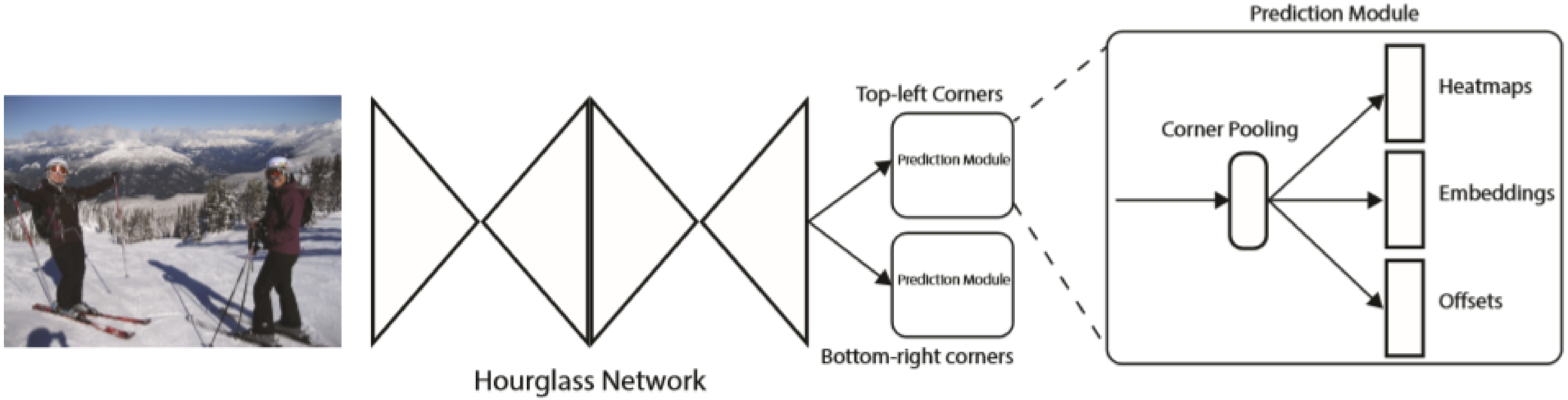
focal loss

- 样本类别分类不均衡
- 样本正负案例不均衡

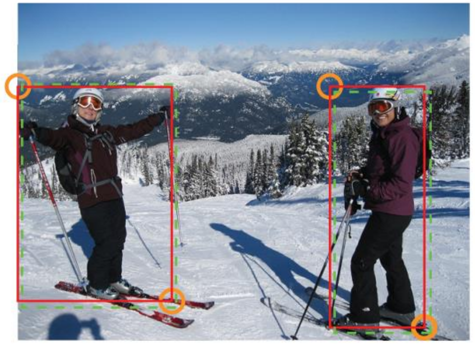
α=2,β=4

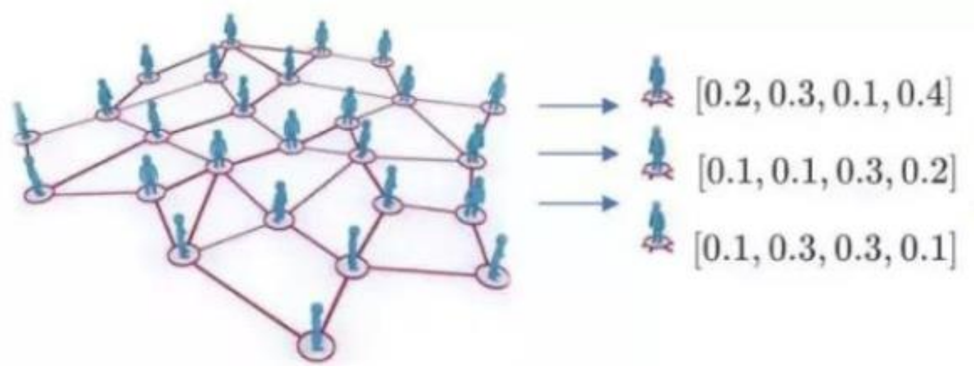
y_{cij} 是用非标准化高斯增强的ground-truth热图

p_{cij} 是预测热图中C类位置(i , j)的得分


Hourglass Network
Top-left Corners
Prediction Module
Prediction Module
Bottom-right corners
Prediction Module
Corner Pooling
Heatmaps
Embeddings
Offsets

Embedding and vector embedding

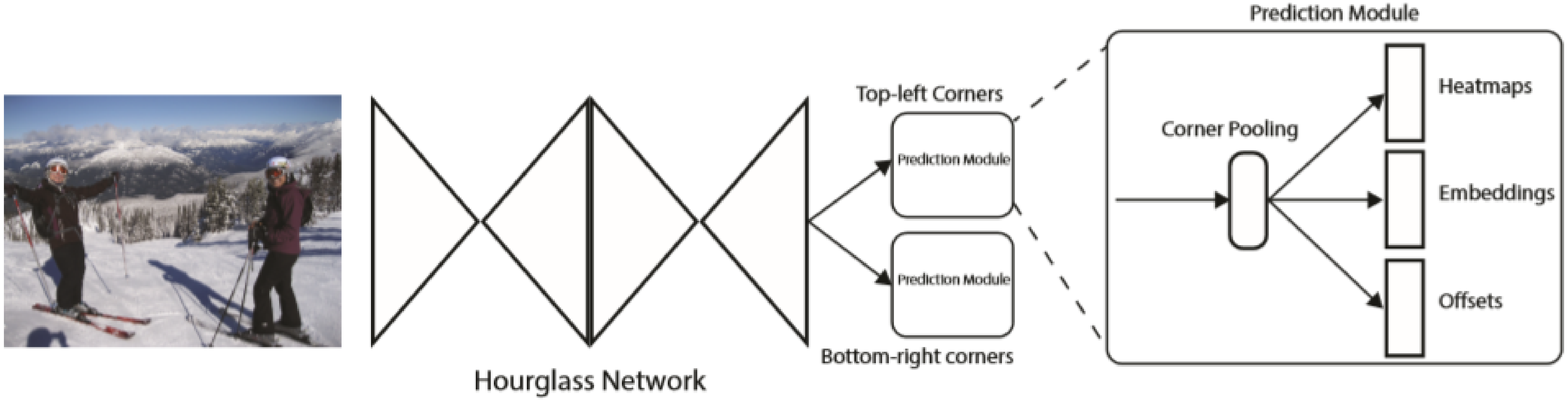




$$L_{pull} = \frac{1}{N} \sum_{k=1}^{N} \left[(e_{t_k} - e_k)^2 + (e_{b_k} - e_k)^2 \right]$$

$$L_{push} = \frac{1}{N(N-1)} \sum_{k=1}^{N} \sum_{\substack{j=1 \\ j \neq k}}^{N} \max\left(0, \Delta - |e_k - e_j|\right)$$

Newell, A., Huang, Z., and Deng, J. (2017). Associative embedding: End-to-end learning for joint detection and grouping. In Advances in Neural Information Processing Systems, pages 2274–2284.


Prediction Module
Top-left Corners
Prediction Module
Prediction Module
Bottom-right corners
Hourglass Network
Corner Pooling
Heatmaps
Embeddings
Offsets

$$(x,y) \longrightarrow \left(\left\lfloor \frac{x}{n} \right\rfloor, \left\lfloor \frac{y}{n} \right\rfloor\right)$$

$$\boldsymbol{o}_k = \left(\frac{x_k}{n} - \left\lfloor \frac{x_k}{n} \right\rfloor, \frac{y_k}{n} - \left\lfloor \frac{y_k}{n} \right\rfloor\right)$$

$$L_{off} = \frac{1}{N} \sum_{k=1}^{N} \mathrm{SmoothL1Loss}\left(\boldsymbol{o}_k, \hat{\boldsymbol{o}}_k\right)$$

$$L = L_{det} + \alpha L_{pull} + \beta L_{push} + \gamma L_{off}$$

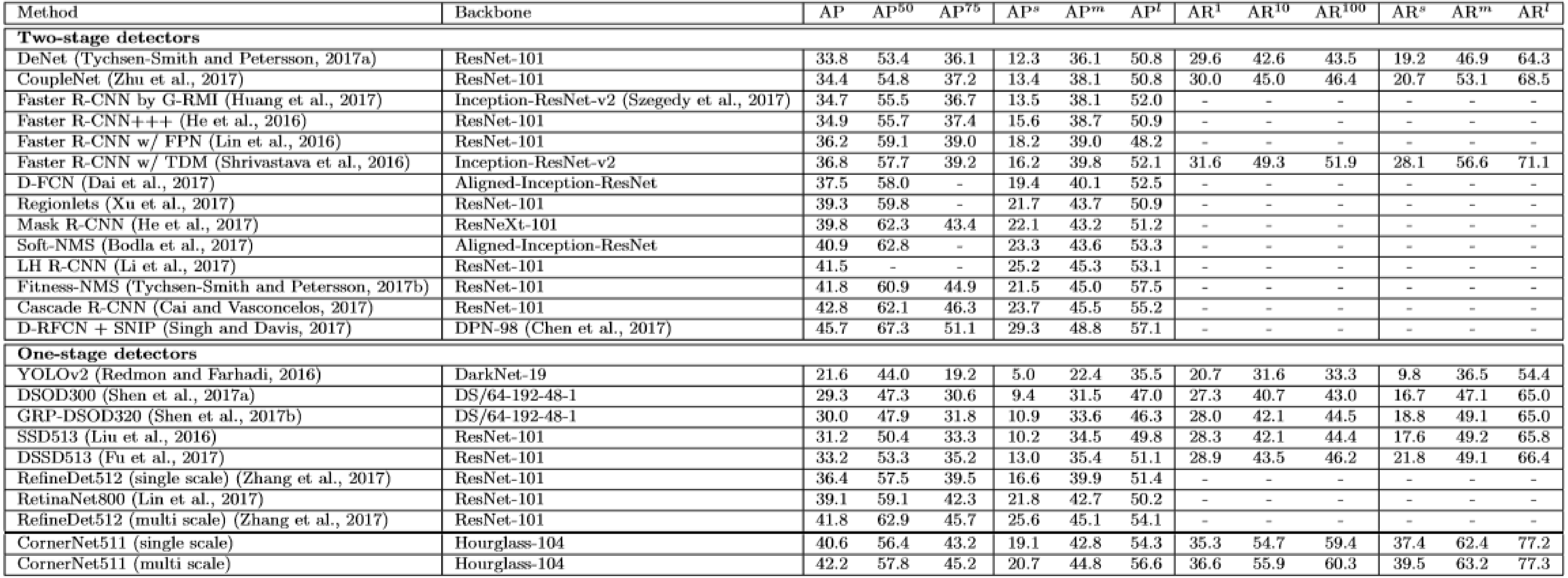
α=0.1,β=0.1,γ=1

$$L_{det} = \frac{-1}{N} \sum_{c=1}^{C} \sum_{i=1}^{H} \sum_{j=1}^{W} \begin{cases} \left(1 - p_{cij}\right)^{\alpha} \log\left(p_{cij}\right) & \text{if } y_{cij} = 1 \\ \left(1 - y_{cij}\right)^{\beta} \left(p_{cij}\right)^{\alpha} \log\left(1 - p_{cij}\right) & \text{otherwise} \end{cases}$$

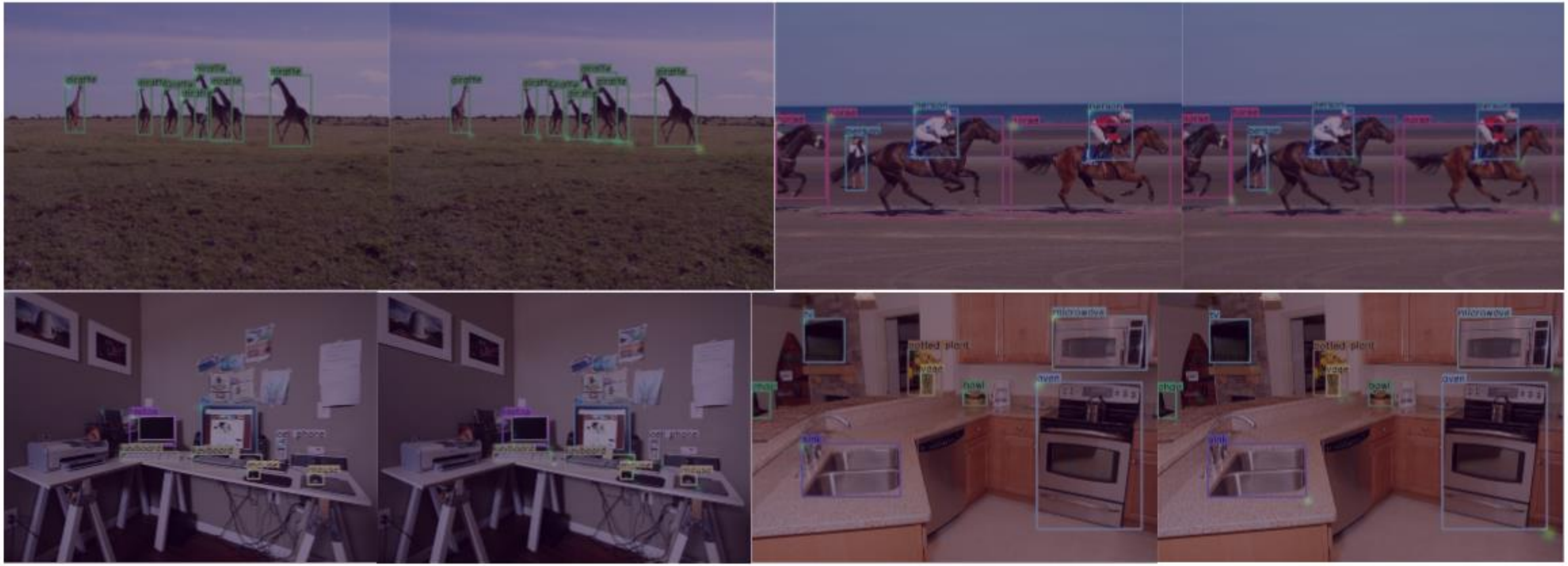
$$L_{pull} = \frac{1}{N} \sum_{k=1}^{N} \left[\left(e_{t_k} - e_k\right)^2 + \left(e_{b_k} - e_k\right)^2 \right]$$

$$L_{push} = \frac{1}{N(N-1)} \sum_{k=1}^{N} \sum_{\substack{j=1 \\ j \neq k}}^{N} \max\left(0, \Delta - \left|e_k - e_j\right|\right)$$

$$L_{off} = \frac{1}{N} \sum_{k=1}^{N} \text{SmoothL1Loss}\left(\boldsymbol{o}_k, \hat{\boldsymbol{o}}_k\right)$$



Method	Backbone	AP	AP^{50}	AP^{75}	AP^{s}	AP^{m}	AP^{l}	AR^{1}	AR^{10}	AR^{100}	AR^{s}	AR^{m}	AR^{l}
Two-stage detectors													
DeNet (Tychsen-Smith and Petersson, 2017a)	ResNet-101	33.8	53.4	36.1	12.3	36.1	50.8	29.6	42.6	43.5	19.2	46.9	64.3
CoupleNet (Zhu et al., 2017)	ResNet-101	34.4	54.8	37.2	13.4	38.1	50.8	30.0	45.0	46.4	20.7	53.1	68.5
Faster R-CNN by G-RMI (Huang et al., 2017)	Inception-ResNet-v2 (Szegedy et al., 2017)	34.7	55.5	36.7	13.5	38.1	52.0	-	-	-	-	-	-
Faster R-CNN+++ (He et al., 2016)	ResNet-101	34.9	55.7	37.4	15.6	38.7	50.9	-	-	-	-	-	-
Faster R-CNN w/ FPN (Lin et al., 2016)	ResNet-101	36.2	59.1	39.0	18.2	39.0	48.2	-	-	-	-	-	-
Faster R-CNN w/ TDM (Shrivastava et al., 2016)	Inception-ResNet-v2	36.8	57.7	39.2	16.2	39.8	52.1	31.6	49.3	51.9	28.1	56.6	71.1
D-FCN (Dai et al., 2017)	Aligned-Inception-ResNet	37.5	58.0	-	19.4	40.1	52.5	-	-	-	-	-	-
Regionlets (Xu et al., 2017)	ResNet-101	39.3	59.8	-	21.7	43.7	50.9	-	-	-	-	-	-
Mask R-CNN (He et al., 2017)	ResNeXt-101	39.8	62.3	43.4	22.1	43.2	51.2	-	-	-	-	-	-
Soft-NMS (Bodla et al., 2017)	Aligned-Inception-ResNet	40.9	62.8	-	23.3	43.6	53.3	-	-	-	-	-	-
LH R-CNN (Li et al., 2017)	ResNet-101	41.5	-	-	25.2	45.3	53.1	-	-	-	-	-	-
Fitness-NMS (Tychsen-Smith and Petersson, 2017b)	ResNet-101	41.8	60.9	44.9	21.5	45.0	57.5	-	-	-	-	-	-
Cascade R-CNN (Cai and Vasconcelos, 2017)	ResNet-101	42.8	62.1	46.3	23.7	45.5	55.2	-	-	-	-	-	-
D-RFCN + SNIP (Singh and Davis, 2017)	DPN-98 (Chen et al., 2017)	45.7	67.3	51.1	29.3	48.8	57.1	-	-	-	-	-	-
One-stage detectors													
YOLOv2 (Redmon and Farhadi, 2016)	DarkNet-19	21.6	44.0	19.2	5.0	22.4	35.5	20.7	31.6	33.3	9.8	36.5	54.4
DSOD300 (Shen et al., 2017a)	DS/64-192-48-1	29.3	47.3	30.6	9.4	31.5	47.0	27.3	40.7	43.0	16.7	47.1	65.0
GRP-DSOD320 (Shen et al., 2017b)	DS/64-192-48-1	30.0	47.9	31.8	10.9	33.6	46.3	28.0	42.1	44.5	18.8	49.1	65.0
SSD513 (Liu et al., 2016)	ResNet-101	31.2	50.4	33.3	10.2	34.5	49.8	28.3	42.1	44.4	17.6	49.2	65.8
DSSD513 (Fu et al., 2017)	ResNet-101	33.2	53.3	35.2	13.0	35.4	51.1	28.9	43.5	46.2	21.8	49.1	66.4
RefineDet512 (single scale) (Zhang et al., 2017)	ResNet-101	36.4	57.5	39.5	16.6	39.9	51.4	-	-	-	-	-	-
RetinaNet800 (Lin et al., 2017)	ResNet-101	39.1	59.1	42.3	21.8	42.7	50.2	-	-	-	-	-	-
RefineDet512 (multi scale) (Zhang et al., 2017)	ResNet-101	41.8	62.9	45.7	25.6	45.1	54.1	-	-	-	-	-	-
CornerNet511 (single scale)	Hourglass-104	40.6	56.4	43.2	19.1	42.8	54.3	35.3	54.7	59.4	37.4	62.4	77.2
CornerNet511 (multi scale)	Hourglass-104	42.2	57.8	45.2	20.7	44.8	56.6	36.6	55.9	60.3	39.5	63.2	77.3


giraffe
person
horse
tv
potted plant
vase
microwave
bowl
oven
chair
sink
keyboard
cell phone
mouse

CornerNet 实现

https://blog.csdn.net/qq_36492210/article/details/84993195
https://github.com/princeton-vl/CornerNet-Lite

https://github.com/princeton-vl/CornerNet

CornerNet 代码

链接：https://pan.baidu.com/s/16QreUNVwGK7E95oDfAsvtg
提取码：9es