



面向真实场景的违禁物品检测

Towards Real-World Prohibited Item Detection: A Large-Scale X-ray Benchmark (ICCV2021)

王伯英, 张立波, 文龙银, 刘祥龙, 武延军
通讯作者: 张立波 (libo@iscas.ac.cn)

挑战及贡献

1. 挑战

- X射线强大的穿透力使得物体丢失很多重要的特征, 并且物体重叠增加了检测的难度。
- 包裹中存在许多类别的物体, 但是并不能进行完全标注, 因此非违禁物品对违禁物品的检测造成干扰。
- 自然图像中存在的类内差异、类间不平衡问题依然存在。

2. 贡献

- 针对真实场景中的违禁品检测问题, 提出一个大规模的数据集PIDray。它是目前**规模最大、首个针对蓄意隐藏场景的违禁品检测基准**的违禁品检测数据集
- 为PIDray提出一个选择性密集注意力网络(SDANet)。

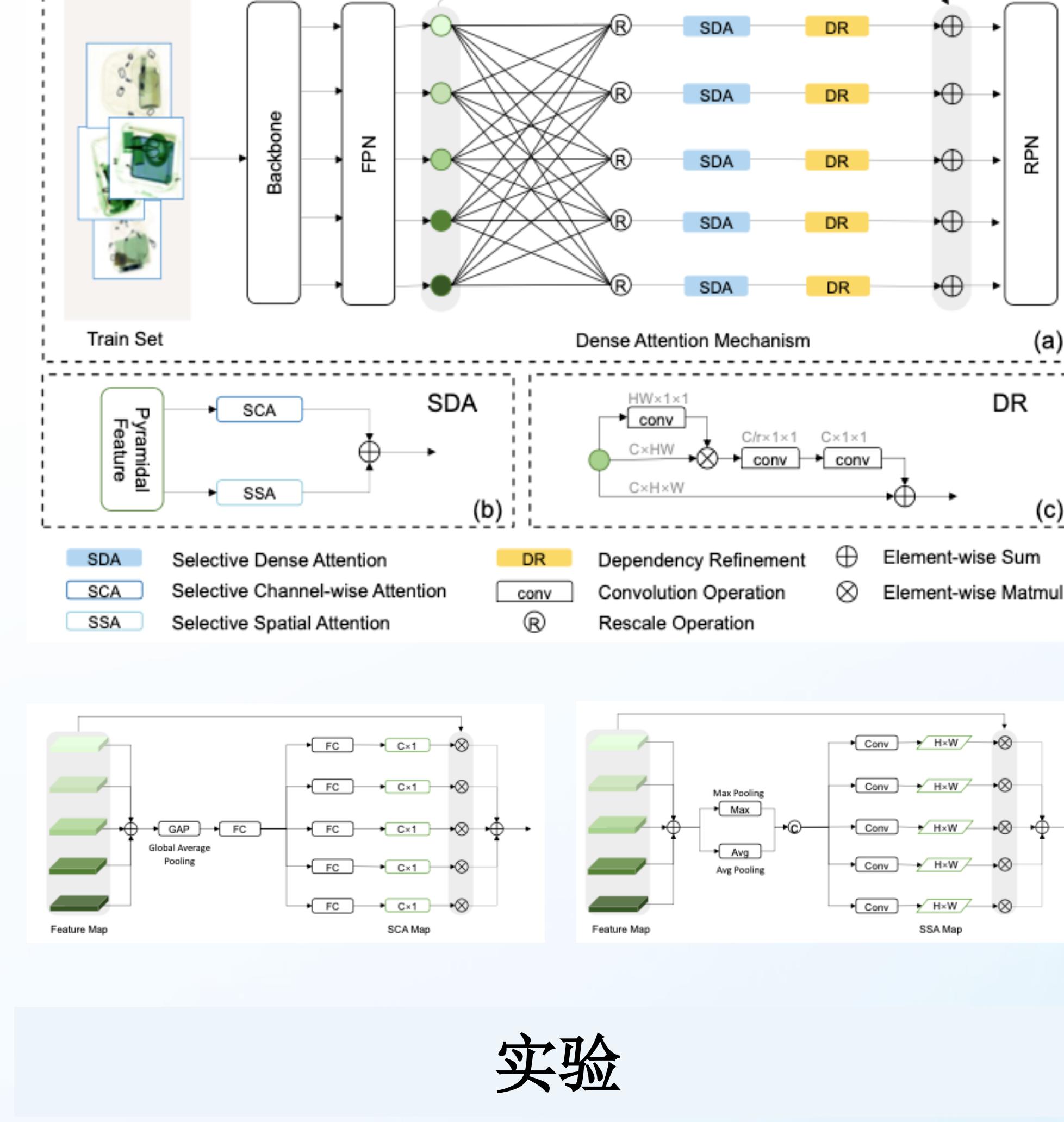
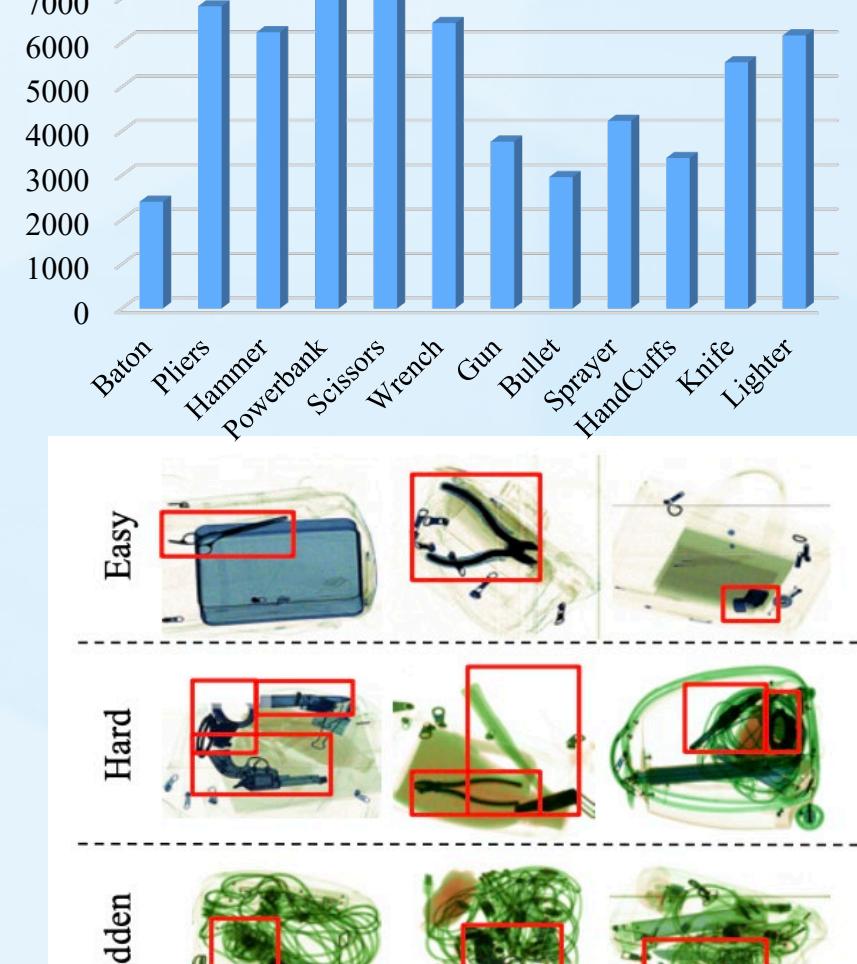


违禁品检测数据集

- 目前的大部分数据集不能满足现实需求, 原因有三个:1)现有数据集只包含少量和极少数类别的违禁物品。2)一些真实场景需要更高级别的标注信息。3)现有数据集很少对蓄意隐藏数据进行研究。
- PIDray包含**47677**张图片, 其中有**12**类违禁物品, 分别为:枪、刀、扳手、钳子、剪刀、锤子、手铐、电警棍、喷雾器、充电宝、打火机和子弹。
- PIDray划分为训练集和测试集, 分别包含**29457**张(约60%)和**18220**张(剩余40%)图像。
- 根据违禁物体检测的难易程度, 测试集分为**容易、困难、隐藏**三个子集。

数据集	年份	类别	图片		标签			类型	场景	应用	是否公开
			总共	正样本	图像	实例	掩码				
GDxRay	2015	3	8,150	8,150	✓	✓	Real	-	C+O	✓	
Dbf6	2017	6	11,627	11,627	✓	✓	Real	-	C+O	✗	
Dbf3	2018	3	7,603	7,603	✓	✓	Real	-	C+O	✗	
Liu et al.	2019	6	32,253	12,683	✓	✓	Real	S	C+O	✗	
SIxRay	2019	6	1,059,231	8,929	✓	✓	Real	S	C+O	✓	
OPIXray	2020	5	8,885	8,885	✓	✓	Synth	A	C+O	✓	
PIDray	2021	12	47,677	47,677	✓	✓	Real	S+A+R	C+O+I	✓	

Mode	Train	Test		
		Easy	Hard	Hidden
Count		29,457	9,482	3,733
Total			47,677	



实验

Table 3. The evaluation results on the proposed PIDray dataset. COCO mAP (%) is used to evaluate performance of all methods.

Method	Backbone	Detection AP			Segmentation AP				
		Easy	Hard	Hidden	Overall	Easy	Hard	Hidden	Overall
FCOS	ResNet-101-FPN	61.8	51.7	37.5	50.3	-	-	-	-
RetinaNet	ResNet-101-FPN	61.8	52.2	40.6	51.5	-	-	-	-
Faster R-CNN	ResNet-101-FPN	63.3	57.2	42.1	54.2	-	-	-	-
Libra R-CNN	ResNet-101-FPN	64.7	58.8	42.9	55.5	-	-	-	-
Mask R-CNN	ResNet-101-FPN	64.7	59.0	43.8	55.8	57.6	50.2	35.2	47.7
SSD512	VGG16	68.1	58.9	45.7	57.6	-	-	-	-
Cascade R-CNN	ResNet-101-FPN	69.3	62.8	48.0	60.0	-	-	-	-
Cascade Mask R-CNN	ResNet-101-FPN	70.9	64.0	48.0	61.0	59.2	51.5	36.1	48.9
SDANet(ours)	ResNet-101-FPN	71.2	64.2	49.5	61.6	59.9	52.0	37.4	49.8
Cascade Mask R-CNN	ResNet-101-BiFPN	68.0	61.1	46.9	58.7	58.0	49.8	35.3	47.7
Cascade R-CNN	ResNet-101-FPN	70.4	63.4	46.7	60.2	59.2	51.4	35.0	48.5
Cascade Mask R-CNN	ResNet-101-FPN	70.9	64.0	48.0	61.0	59.2	51.5	36.1	48.9
SDANet(ours)	ResNet-101-FPN	71.2	64.2	49.5	61.6	59.9	52.0	37.4	49.8

Table 4. Effectiveness of various designs. All models are trained on the PIDray training subset and tested on the PIDray hidden test set. The accuracies are indicated by "detection AP/segmentation AP".

SCA	SSA	DR	AP	AP ₀	AP _S	AP _S	AR ₁	AR ₁₀	AR ₁₀₀	AR _S
✓			48.0/36.1	62.7/59.8	54.0/40.4	57.0/43.5	56.0/42.9	57.6/44.0	57.6/44.0	57.6/44.0
	✓		48.3/36.5	63.5/59.3	54.3/41.2	57.2/43.9	56.2/43.4	57.9/44.4	57.9/44.4	57.9/44.4
✓	✓		48.3/36.2	63.2/59.6	54.6/40.1	57.4/43.8	56.6/43.3	58.1/44.3	58.1/44.3	58.1/44.3
✓	✓	✓	48.9/36.7	63.8/60.0	55.4/40.8	58.3/44.3	57.4/43.8	59.3/45.0	59.3/45.0	59.3/45.0
✓	✓	✓	49.5/37.4	64.5/60.6	55.7/42.2	58.5/44.8	57.2/44.1	59.5/45.5	59.5/45.5	59.5/45.5

Table 5. Comparison of dependency refinement (DR) and other attention mechanisms on the hidden test set.

Method	Det AP	Seg AP
ours w/o DR	48.9	36.7
+SE	49.1	36.7
+CBAM	47.0	35.8
+DR	49.5	37.4

Table 6. Experimental results on general detection datasets.

Method	MS COCO	PASCAL VOC
baseline	42.9	81.5
SDANet	43.5[+0.6]	82.5[+1.0]

