

主席:程明明(南丹大学),王兴刚(华中科技大学),魏云超(UIUC)

报告人: 程明明(南丹大学)

报告题目1:面向开放环境的自适应视觉感知

报告人: 叶会祥(中国科学院大学)

报告题目2:从弱监督到自学习视觉目标建模

报告人: 魏秀叁(旷视)

报告题 53: Weakly-supervised object discovery based on pre-trained deep

CNNs

报告人:黄圣君(南京航空航天大学)

报告题 54: Cost-Sensitive Active Learning

报告人:魏云超(UIUC)

报号题 5: Towards Weakly Supervised Object Recognition and Scene



报告题目1:面向开放环境的自适应视觉感知

- > 报告题目解析: 面向开放环境的自适应感知
- 该题目取自于"科技创新2030"新一代人工智能"重大项目2018年度项目申报指南征求意见稿"。
- 深层含义:针对应用场景变换易导致智能系统性能急剧下问题,发展适应能力强的层次化网络结构、可连续学习的机器学习策略及一般性效能度量方法,突破无监督学习、经验记忆利用、内隐知识发现与引导及注意力选择等难点,推动形成开放环境和变化场景下的通用型感知智能。

讲者简介

- 程明明, 1985年生。南丹大学教授, 国家"万人针剑" 青年拔火人才, 首批天津市杰出青年基金获得者。
- 主要研究方向包括: 计算机图形学、计算机视觉、图像处理等。已在IEEE PAMI, ACM TOG等CCF-A类国际会议及期刊发表论文30余篇。相关研究成果论文他引10,000余次,最高单篇他引2,000余次。
- 研究成果: 华为、腾讯等公司的旗舰产品中应用。显着性物体检测技术被华为Mate 10等旗舰手机作为亮点特性,于产品发布会中展示。其研究工作曾被英国《BBC》、德国《明镜周刊》、美国《赫芬顿邮报》等权威国际媒体撰文报道。
- 游者主页: https://mmcheng.net/zh/cmm/

现有视觉学习技术通常依赖于大规模精确标注的创练数据。在典型实验室环境下设计和创练的人工智能模型,在行业应用场景变换时,容易导致系统性能急剧下降。

本报告将从弱监督视觉理解的角度,介绍在降低模型对特定应用拐景数据依赖方面所开展的一些研究工作,主要包括如何学习一些通用属性,并利用这些通用属性构建弱监督学习模型。

报告插座

研究背景



数据集很重要: 现阶段计算 机视觉CV领域的大多数问题 还是依赖于特定的数据集进 行训练和测试评估的;



标注成本很大: 高质量的图像标注为我们进行图像理解提供了方便, 但获取精确的标注是非常困难和耗时;



研究方法难选:

深度神经网络共性技术 视觉基元属性感知



视觉感知技术的春天

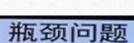
•智能终端、教育、能源、医疗、农业等

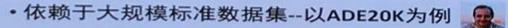












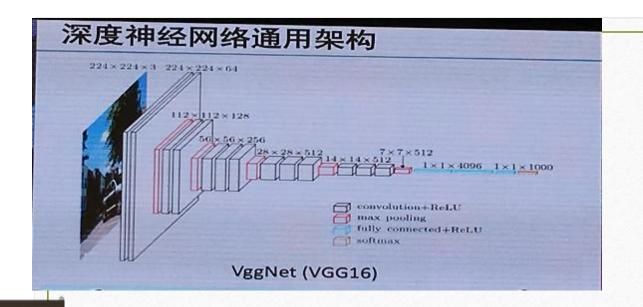


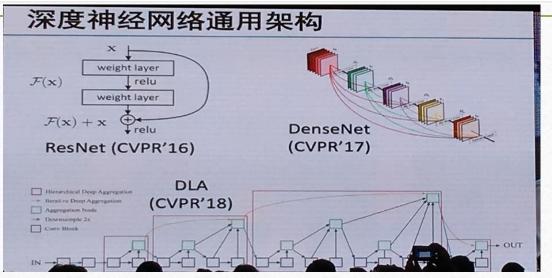
- · Torralba教授的母亲精确标注了213841个物体
- Joke: Job offer: I am looking for more parents
- 开放环境自适应感知能力差
 - 难以满足各种实际应用需求
 - 典型实验室环境下设计和训练的人工智能模型, 在行业应用场景变换时, 容易导致系统性能急 剧下降

视觉感知技术的春天 ADE20K

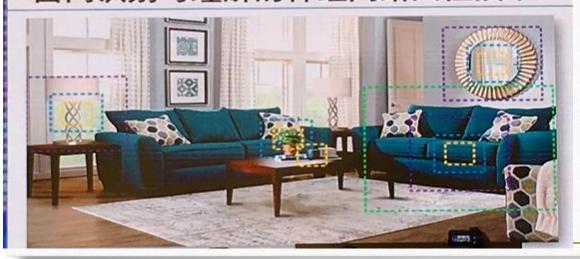
开放环境的自适应感知







面向识别与理解的神经网络共性技术



开放环境的自适应感知

面向行业开放应用场景,而非传统实验环境下的高可靠、高通用性基础算法

关键机器 学习算法

深度神经网 络共性技术 视觉基元

研究内容







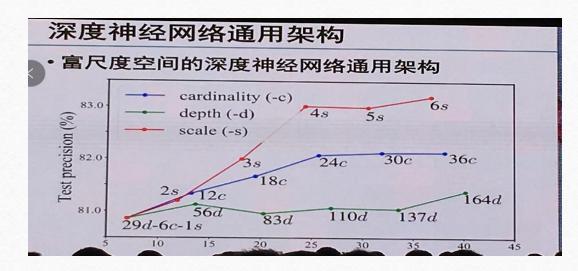
富尺度空向神经网络架构: 多任务协同求解,鲁棒性 提高 显著性物体检测:

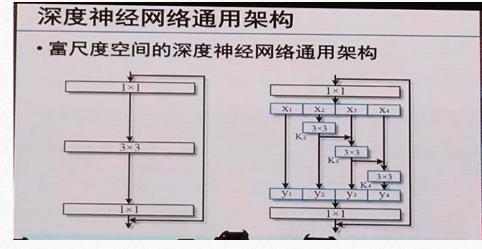
预设基元属性感知能力, 减少数据依赖;

互联网大数据自主学习: 减少人工标注, 自动学习。 1. 富尺度空向的神经网络架构富尺度:位移、尺度、形变不变

网络设计基准:

• 金字俗结构; 空向池化; 残差学习。

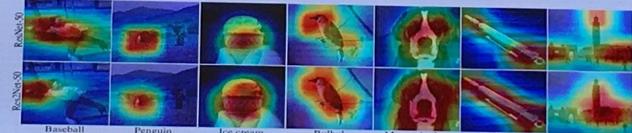




1. 富尺度空向的神经网络的应用 在图像分类、分割任务中效果较好

深度神经网络通用架构

- 富尺度空间的深度神经网络通用架构
 - 应用2: Class Activation Mapping (分类任务)



eball Penguin Ice cream Bulbul Mountain dog Ballpoint Mo Figure 6. Visualization of class activation mapping [31], using ResNet-50 and Res2Net-50 as backbone networks.

深度神经网络通用架构

• 富尺度空间的深度神经网络通用架构

	top-1 err. (%)	top-5 err. (%)
ResNet-50 [13]	23.85	7.13
Res2Net-50	22.83	6.51
InceptionV3 [36]	22.55	6.44
Res2Net-50-299	21.41	5.88
ResNeXt-50 [39]	22.61	6.50
Res2NeXt-50	21.72	6.10
DLA-60 [43]	23.32	6.60
Res2Net-DLA-60	22.67	6.40
DLA-X-60 [43]	22.19	6.13
Res2NeXt-DLA-60	21.72	5.82
SENet-50 [15]	23.24	6.69
SE-Res2Net-50	22.11	6.14
SENeXt-50 [15]	22.16	6.22
SE-Res2NeXt-50	21.70	6.01

深度神经网络通用架构

- 富尺度空间的深度神经网络通用架构
 - •应用3:显著性物体(分割类任务)

Dataset	Backbone	F-measure†	MAE
ECSSD	ResNet-50	0.910	0.065
LCSSD	Res2Net-50	0.924	0.056
PASCAL-S	ResNet-50	0.823	0.105
	Res2Net-50	0.834	0.100
HKU-IS	ResNet-50	0.894	0.058
	Res2Net-50	0.903	0.052
DUT-OMRON	ResNet-50	0.748	0.092
DO1-ONIKON	Res2Net-50	0.784	0.077

Table 7. Salient object detection results on different datasets, me-

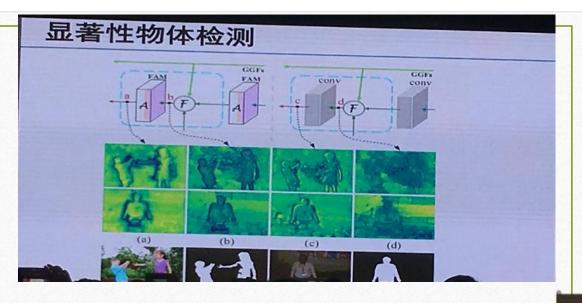
2. 通用视觉基元属性感知 1) 显著性物体检测技术 2) 边缘检测技术

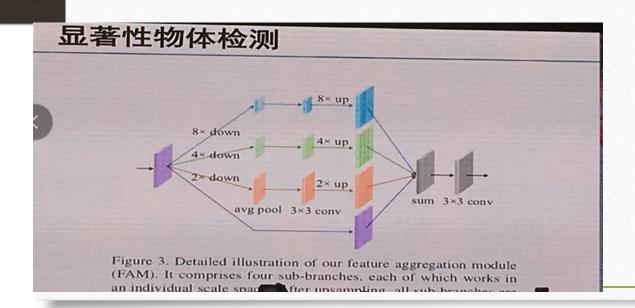


2. 通用视觉基元属性感知方法分类

- > 显著性物体检测
- > RGBD显著性物体检测
- > 边缘检测
-) 视觉注意力机制若监督语义分割





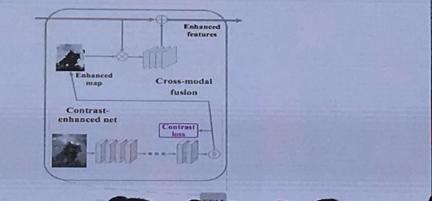


	Training		ECSSD [46] PASCAL		L-S [23] DUT-O [) [47] HKU-		IS [20]	SOD [32, 34]		DUTS-TE [40]		
Model	#Images	Dataset	MaxF†	MAE 4	MaxF †	MAEL	MaxF†	MAEL			MaxF †		MaxP 1	
VGG-16 backbo	ne		No.									1-10-41-4	INTERES	MALE
DCL [21]	2,500	MB	0.896	0.080	0.805	0.115	0.733	0.094	0.893	0.063	0.831	0.111	0.000	
RFCN [41]	10,000	MK	0.898	0.097	0.827	0.118	0.747	0.094	0.895	0.079	0.805	0.131	0.786	0.08
DHS [25]	9,500	MK+DTO	0.905	0.062	0.825	0.092		0.034	0.892	0.079	0.805	0.161	0.786	0.090
ELD [9]	9,000	MK	0.865	0.082	0.772	0.122	0.738	0.093	0.843	0.052	0.823	0.128	0.815	0.065
MSR [19]	5,000	MB + H	0.903	0.059	0.839	0.083	0.790	0.073	0.907	0.043	0.841		0.747	0.093
DSS [12]	2,500	MB	0.906	0.064	0.821	0.101	0.760	0.074	0.900	0.050	0.834	0.111	0.824	0.063
NLDF [30]	3,000	MB	0.903	0.065	0.822	0.098	0.753	0.079	0.902	0.036	0.837	0.123	0.815	0.065
UCF [51]	10,000	MK	0.908	0.080	0.820	0.127	0.735	0.131	O.BER	0.073	0.798	0.164	0.816	0.063
Amulet [50]	10,000	MK	0.911	0.062	0.826	0.092	0.737	0.083	0.889	0.052	0.799	0.146	0.771	0.110
PAGR [52]	10.553	DTS	0.924	0.064	0.847	0.089	0.771	0.071	0.919	0.032	0.755	0.140	0.854	0.075
PiCANet [26]	10,553	DTS	0.930	0.049	0.858	0.078	0.815	0.067	0.921	0.042	0.863	0.102	0.855	0.05
PoolNet (Ours)	2,500	MH	0.918	0.057	0.828	0.098	0.783	0.065	0.908	0.044	0.846	0.124	0.819	0.06
PoolNet (Ours)	5,000	MB + H	0.930	0.053	0.838	0.093	0.806	0.063	0.936	0.032	0.841	0.118	0.855	0.05
PoolNet (Ours)	10,553	DTS	0.936	0.047	0.857	0.078	0.817	0.058	0.928	0.035	0.859	0.115	0.876	0.04
PoolNet† (Ours)	10.553	DTS	0.937	0.044	0.865	0.072	0.821	0.056	0.931	0.033	0.866	0.105	0.880	0.04
ResNet-50 backb	one													
SRM [42]	10,553	DTS	0.916	0.056	0.838	0.084	0.769	0.069	0.906	0.046	0.840	0.126	non	
DGRL [43]	10,553	DTS	0.921	0.043	0.844	0.072	0.774	0.062	0.910	0.036	0.843	0.103	0.826	0.05
PICANet [26]	10,553	DTS	0.932	0.048	0.864	0.075	0.820	0.064	0.920	0.044	0.843	0.103		0.04
PoolNet (Ours)	10,553	DTS	0.940	0.042	0.863	0.075	0.830	0.055	0.934	0.032	0.867	0.103	0.863	0.05
PoolNet† (Ours)	10,553	DTS	0.945	0.038	0.880	0.065	0.833	0.053	0.935	0.032	0.867	0.100	0.886	0.04

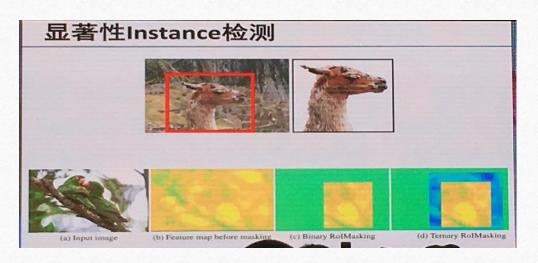
RGBD显著性物体检测 • 难点: 深度图质量, 多模态融合机制 Concatenate (a) Predict (b) CNNs CNNs Model Predict fusion CNNs



RGBD显著性物体检测 • 利用对比度先验





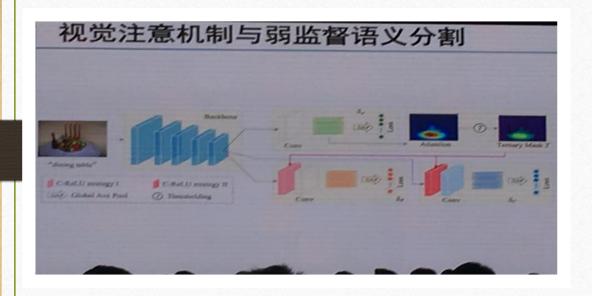


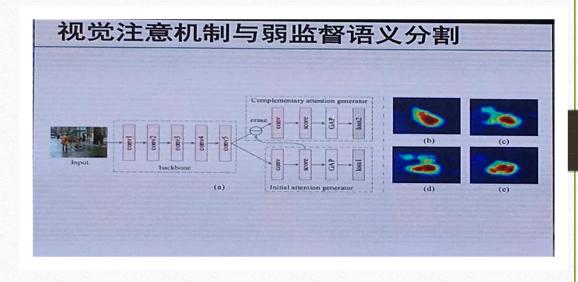


Methods	mAP ^{0.5}	mAP ^{0.7}	mAP _O .5	$mAP_O^{0.7}$
MSRNet [33]	65.3%	52.3%	-	-
S4Net	92.1%	69.6%	87.5%	60.2%

Table 4: Quantitative comparisons with existing methods on the 'test' set. As the instance segmentation maps of [33] and related code are not available, thus we use '-' to fill the blank cells.

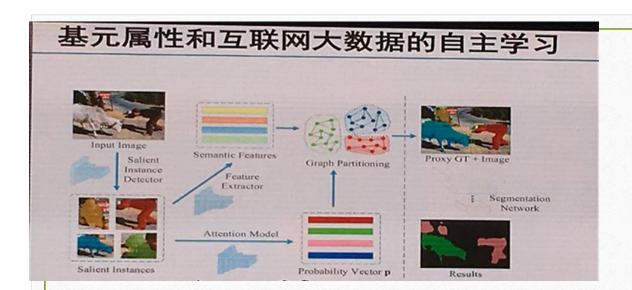


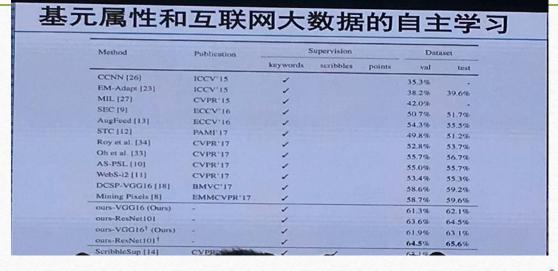


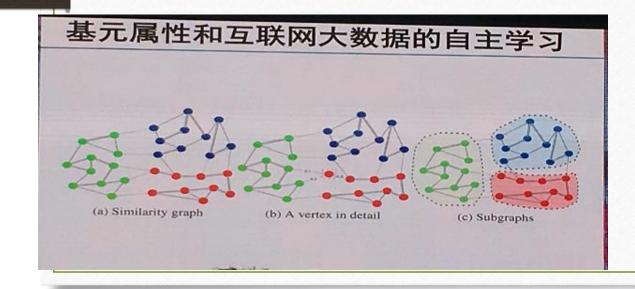


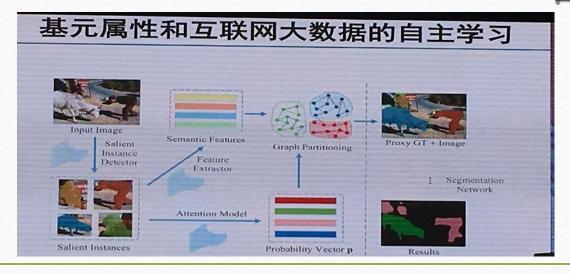
3. 互联网大数据知何利用

- > 基元属性与互联网大数据的关系
- 户细何利用互联网数据自主学习





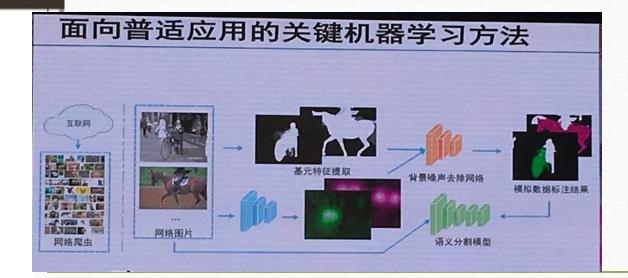




利用互联网大数据的自主学习

- 现有基于互联网数据的弱监督语义分割方法
 - 依赖于 (图像关键字级) 人工标注来去燥
 - · 通过挑选简单图像来提高Proxy-GT的精度





利用互联网大数据的自主学习

Methods	Publication	Training Set	Supervision	Backbone	mIoU val (%)	mIoU test (%)
SEC [20]	ECCV'16	$\mathcal{D}(V)$	weak	VGGNet	50.7	51.7
AugFeed [32]	ECCV'16	$\mathcal{D}(V)$	weak	VGGNct	54.3	55.5
Oh et al. [28]	CVPR'17	$\mathcal{D}(V)$	weak	VGGNet	55.7	56.7
AE-PSL [37]	CVPR'17	$\mathcal{D}(V)$	weak	VGGNet	55.0	55.7
DCSP [3]	BMVC'17	$\mathcal{D}(V)$	weak	VGGNet	58.6	59.2
DCSP [3]	BMVC'17	$\mathcal{D}(V)$	weak	ResNets	60.8	61.9
DSRG [18]	CVPR'18	$\mathcal{D}(V)$	weak	VGGNet	59.0	60.4
DSRG [18]	CVPR'18	$\mathcal{D}(V)$	weak	ResNet	61.4	63.2
MCOF [36]	CVPR'18	$\mathcal{D}(V)$	weak	VGGNet	56.2	57.6
Ahn et al [1]	CVPR'18	$\mathcal{D}(V)$	weak	VGGNet	58.4	60.5
Wei et al [39]	CVPR'18	$\mathcal{D}(V)$	weak	VGGNet	60.4	60.8
GAIN [21]	CVPR'18	1464 pixel + D(V)	semi	VGGNet	60.5	62 1
Fan et al. [10]	ECCV'18	$\mathcal{D}(V)$	weak	ResNet	63.6	64.5
Semantic segmentation	on using web d	ata				
STC [38]	PAMI'17	40K web + D(V)	weak	VGGNet	49.8	51.2
WebS-i2 [19]	CVPR'17	19K web + D(V)	weak	VGGNet	53,4	55.3
Hong et al. [14]	CVPR'17	video + D(V)	weak	VGGNet	58.1	58.7
Shen et al. [34]	CVPR'18	80K web + $\mathcal{D}(V)$	weak	VGGNet	58.8	60.2
Shen et al. [34]	CVPR'18	80K web + D(V)	weak	ResNet	63.0	63.9
WebSearch (Ours)		$33K \text{ web} + \mathcal{D}(V)$	weak	VGGNet	62.5	62.2
WebSearch (Ours)		33K web + $\mathcal{D}(V)$	weak	ResNets	65.8	66.1
menoral til (Ours)					011	

本分享来自参加VALSE大会的一位同学,严禁用作任何商业性的分享及大量传播