



# 快速高效小目标检测网络算法研究

**Fast and Efficient Networks for Small Object Detection** 

答辩:张浩

学号: 416114417180

导 师: 洪向共 副教授

日期: 2020年6月6日





01

### 研究背景

摘要、历史

02

### 研究现状

模型、数据集、评价标准

03

### 研究内容

空洞转置、特征提取、模型剪枝

04

### 总结展望

总结、展望







### 研究背景

摘要、历史



### 研究现状

模型、数据集、评价标准



### 研究内容

空洞转置、特征提取、模型剪枝



### 总结展望

总结、展望

# 1.1 目标检测

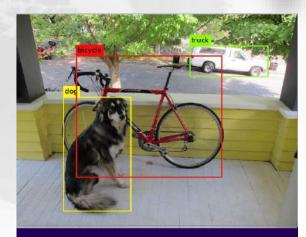
目标检测 (Object Detection) 的目的是识别图像或视频中给定类别的物体及其空间定位。它是分割、场景理解、目标追踪、图像标注及事件检测等高层视觉应用的重要基础。







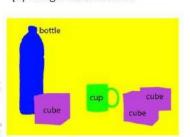




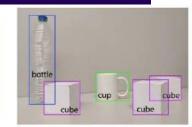
dog: 99% truck: 93% bicycle: 99%



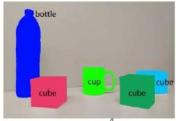
(a) Image classification



(c) Semantic segmentation



(b) Object localization



(d) Instance segmentation

# 1.2 小目标检测

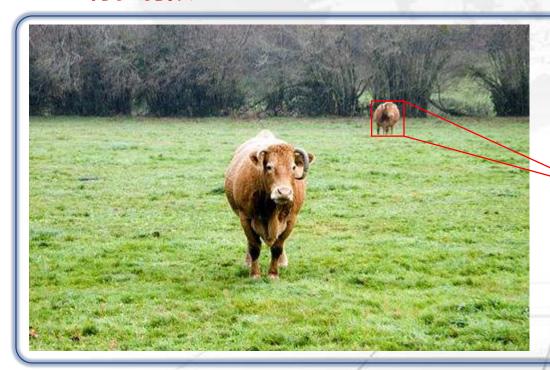
小目标(Small Object):像素小于32×32的

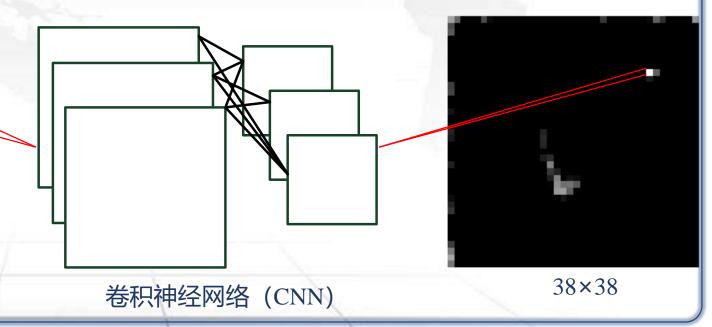
目标区域 (Lin .et al, 2014)

- > 分辨率低
- > 特征有限









# 1.3 目标检测历史

GoogLeNet

(C.S. et al-15)

**VGG** 

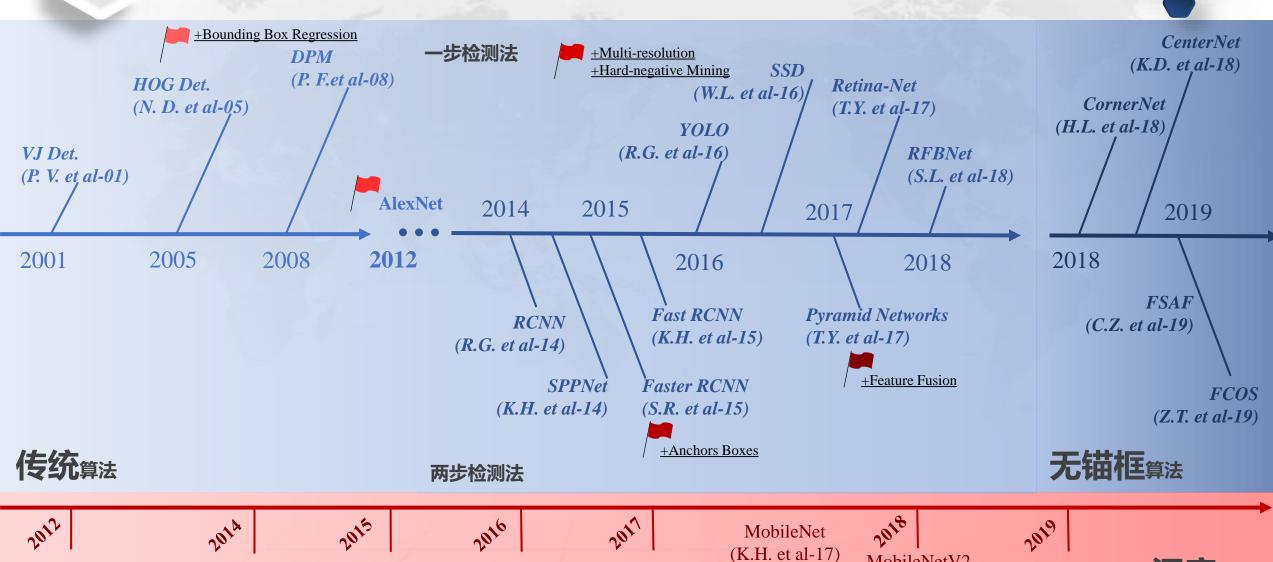
(K.S. et al-14)

**AlexNet** 

(A.K. et al-12)

ResNet

(K.H. et al-16)



DenseNet

(G.H. et al-17)

深度学习

**EfficientNet** 

(M.T. et al-19)

MobileNetV2

(M.S. et al-18)

ShuffleNetV2

(N.M. et al-18)

ShuffleNet

(X.Z. et al-17)





01

### 研究背景

摘要、历史

02

### 研究现状

模型、数据集、评价标准

03

### 研究内容

空洞转置、特征提取、模型剪枝

04

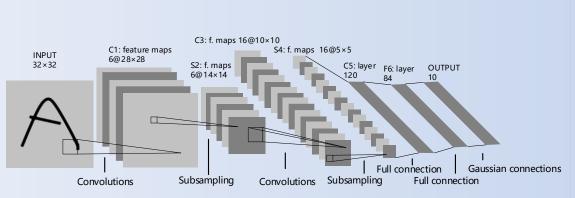
### 总结展望

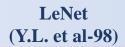
总结、展望

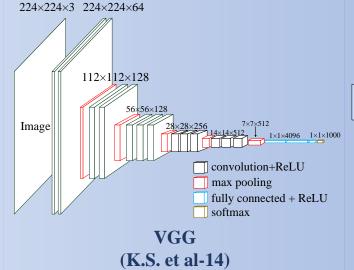


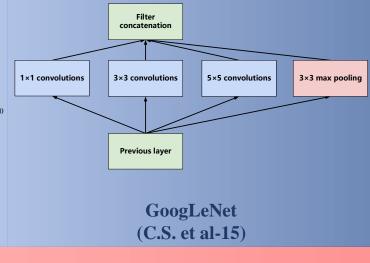
# 2.1 深度学习模型

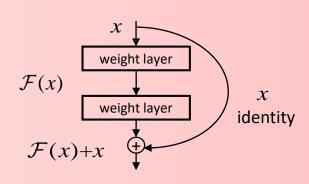




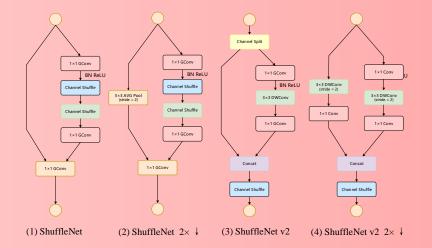




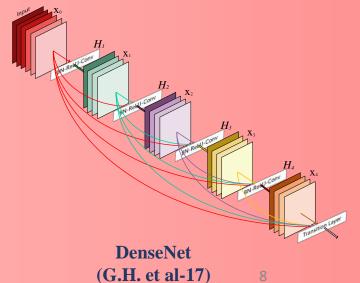




ResNet (K.H. et al-16)



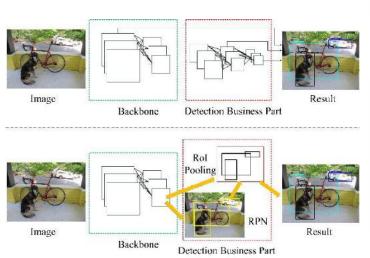
ShuffleNet ShuffleNetV2 (X.Z. et al-17) (N.M. et al-18)



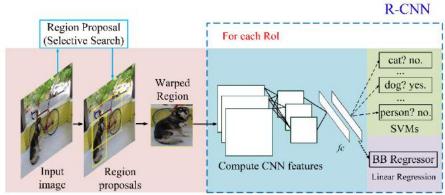


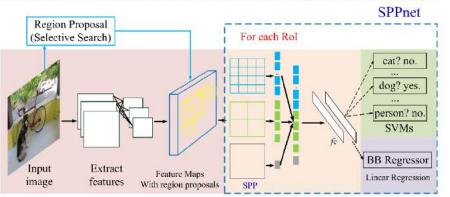
### 2.2 目标检测模型

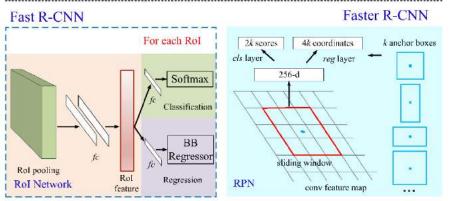


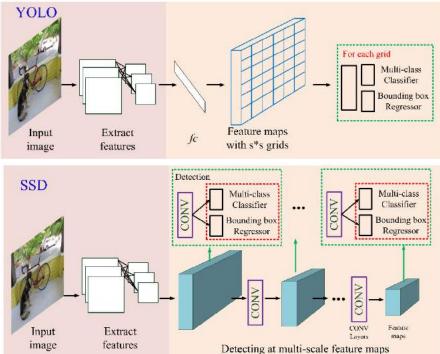


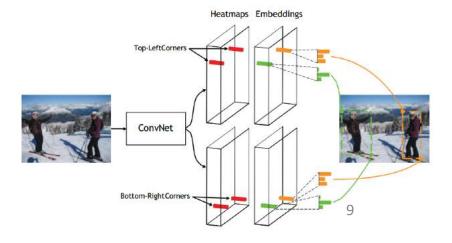








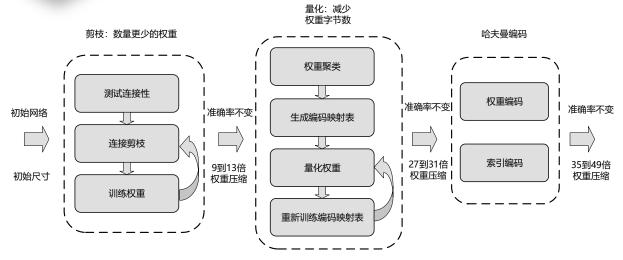




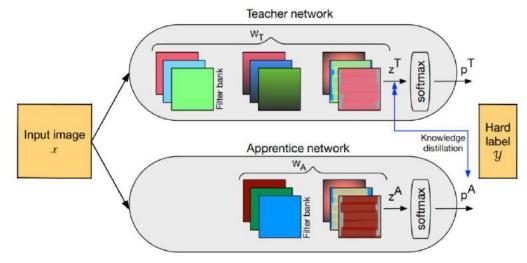


### 2.3 模型压缩

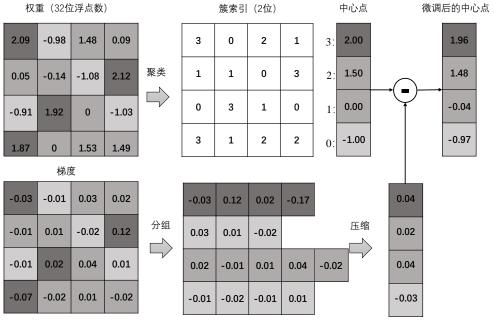




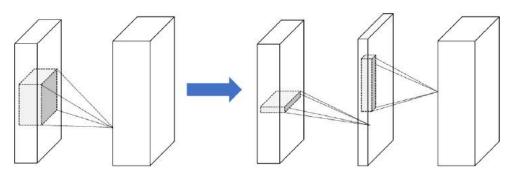
深度压缩算法(S.H. et al-15)



知识蒸馏(G.H. et al-17)



量化

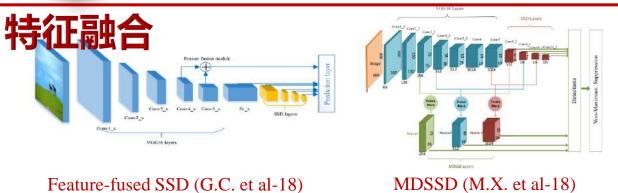


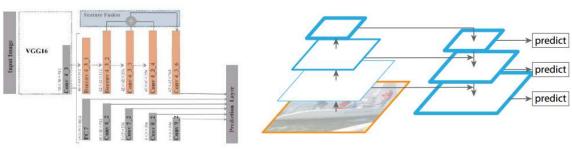
低秩近似



# 2.4 小目标检测模型

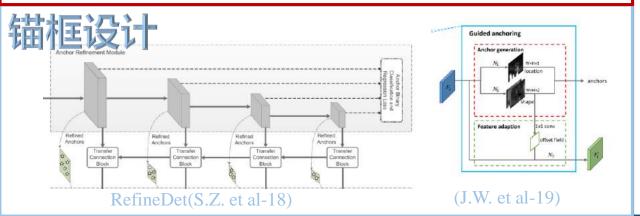


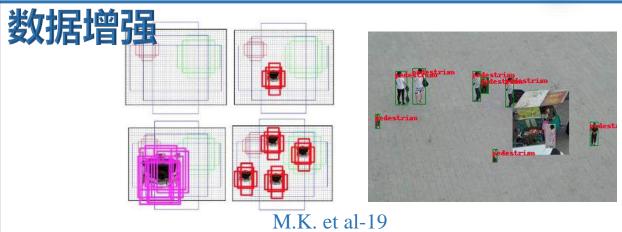


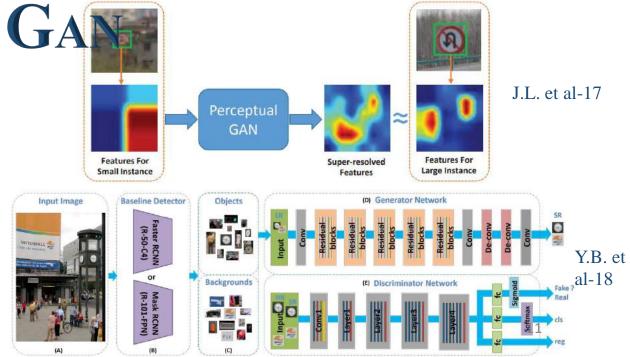


SFE-SSD(H.T. et al-19)

FPN (T.L. et al-17)







# 2.5 小目标检测模型存在问题



特征 融合 数据 增强 错框 设计 GAN

数据增强、锚框设计、GAN等可以看作 训练策略上的提升 故本文主要研究小目标检测在特征融合方面的改进

小目标的主要特点有: 低分辨率、特征难提取

1. SSD等目标检测算法浅层信息用于检测小目标, 但浅层语义信息不够,无法有效识别小目标

基于空洞、转置卷积的小目标检测

2.多次下采样导致小目标特征丢失

基于目标检测特征提取的小目标检测

3. 特征提取、检测复杂度深,导致小目标检测模

型无法实时运行

基于模型剪枝的小目标检测



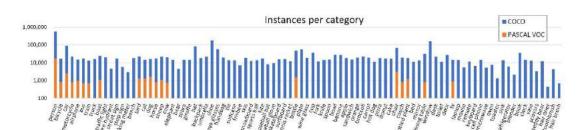
# 2.6 数据集

















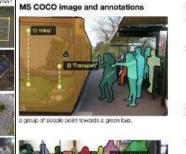




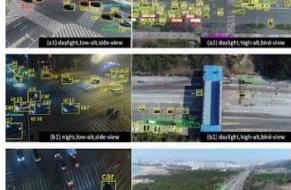












OCR results:



2) '2.20' legible handwitten English 3) '1.50'

3) '1.50' legible handwritten English

OCR results

# 2.7 评价标准



IoU



Positive

#### 混淆矩阵

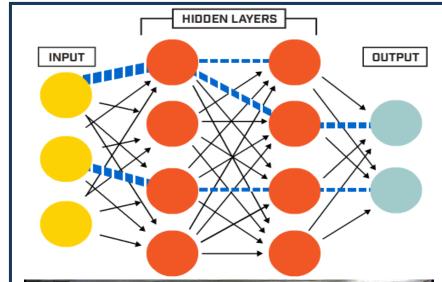
TRUE	True Positive (TP)	True Negative (TP)				
FALSE	False Positive (FP)	False Negative (FN)				
Precisio	$n = \frac{TP}{TP + FP}$	$Recall = \frac{TP}{TP + FN}$				

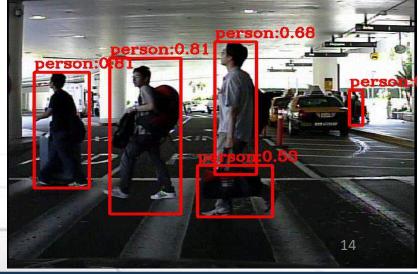
Negative

#### PSACAL VOC 与 MS COCO评价标 准

	VOC	在 $IoU = 0.5$ 时,所有类别下的 AP
		● <i>AP<sub>coco</sub></i> : <i>IoUs</i> : {0.5:0.05:0.95} 下的 mAP
	• $AP_{COCO}^{IoU=0.5}$ : $IoU=0.5$ 时的 mAP	
mAP		• $AP_{COCO}^{IoU=0.75}$ : $IoU=0.75$ 时的 mAP
MS COCO	● <i>AP</i> <sup>small</sup> <sub>coco</sub> : 小目标(像素面积 area < 32 <sup>2</sup> )的 mAP	
		• $AP_{coco}^{\text{medium}}$ : 中等目标( $32^2 < area < 96^2$ )的 mAP
		• $AP_{COCO}^{\text{large}}$ : 大目标 ( $area > 96^2$ ) 的 mAP

#### 帧率 (FPS, Frame Per Second)











### 研究背景

摘要、历史



### 研究现状

模型、数据集、评价标准



### 研究内容

空洞转置、特征提取、模型剪枝

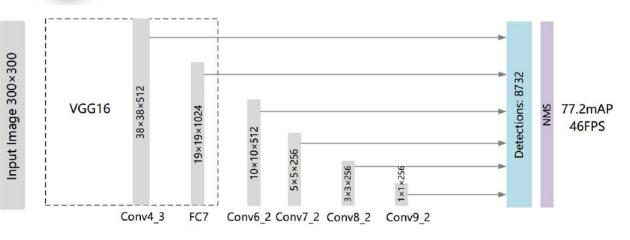


### 总结展望

总结、展望



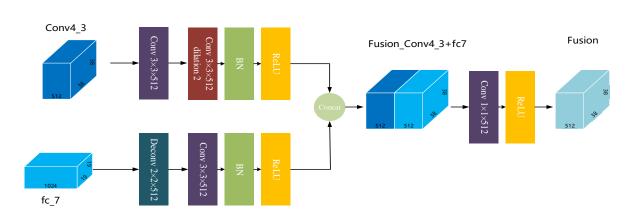




Single-Shot Multi-Box Detector (*W.L. et al-16*)

#### 多尺度检测

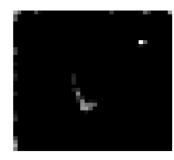
浅层信息用于检测小目标 (缺乏深层语义信息, 且感受野大小不匹配)

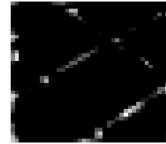


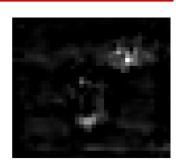
$$X_f = \mathcal{C}\{\mathcal{D}(X_{conv4_3}), \mathcal{T}(X_{fc7})\}\$$

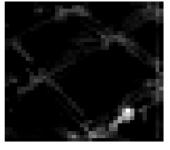












Input images SSD DDSSD 16

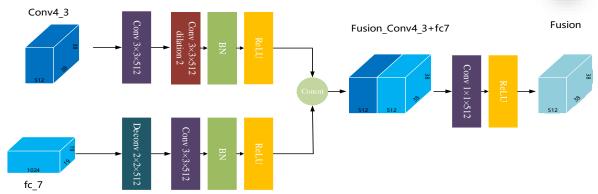




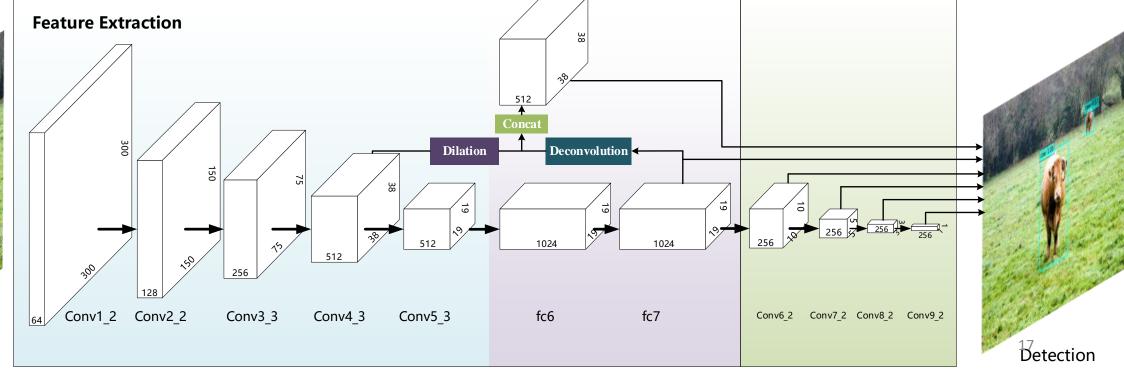
$$X_f = \mathcal{C}\{\mathcal{D}(X_{conv4_3}), \mathcal{T}(X_{fc7})\}\$$

空洞卷积: 提升感受野, 保持分辨率大小

转置卷积: 提升特征分辨率











#### 测试结果分析

本章所提出的算法DDSSD超越了众多基于SSD的小目标检测算法,小目标检测性能比SSD高出了3.9%,上达到了与多尺度特征融合的MDSSD接近的效果,但由此可见,深层语义特征有助于小目标检测,但会一定程度影响大目标的性能。消融实验证明了空洞转置卷积模块的有效性。

Method	SSD				DDSSD
Conv4_3 + Conv		$\sqrt{}$	$\checkmark$		
Conv4_3 + Dilation				$\sqrt{}$	
Conv4_3 + Dilation Module					$\sqrt{}$
Conv5_3 + Deconv Module		$\sqrt{}$			
FC_7 + Deconv Module			$\checkmark$	$\sqrt{}$	$\sqrt{}$
mAP	77.5	78.8	79.2	79.4	79.7

Method	backbone	VOC07	coco	COCO_s	FPS	Proposals	GPU	resolution
Faster RCNN	VGG	73.2	21.9	-	7	6000	Titan X	~1000×60 0
R-FCN	ResNet	80.5	29.2	10.3	9	300	Titan X	~1000×60 0
SSD	VGG	77.5	25.1	6.6	46	8732	Titan X	300
DSSD	ResNet	78.6	28	7.4	11.8	17080	Titan X	321
ESSD	VGG	79.4	-	-	25	8732	Titan X	300
DiSSD	VGG	78.1	-	-	40.8	8732	Titan X	300
FSSD	VGG	78.8	27.1	8.7	65.8	8732	1080Ti	300
MDSSD	VGG	78.6	26.8	10.8	38.5	8732	1080Ti	300
SSD	VGG	77.4	25.1	6.6	72	8732	1080Ti	300
FA-SSD	ResNet-50	78.3	-	-	34.7	8732	Titan Xp	300
SFE-SSD	VGG	78.4	-	-	68	8732	1080Ti	300
RefineDet	VGG	80	29.4	10	40.3	6375	Titan Xp	320
DDSSD	VGG	79.7	28.3	10.5	45.4	8732	1080Ti	300





#### 可视化结果

与SSD(1)相比 更多的检测 更高的置信度



















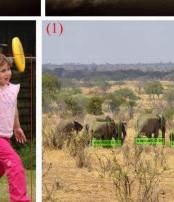




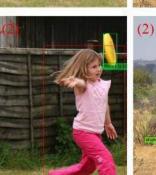






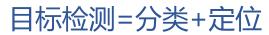














#### 目标检测任务=分类任务+回归任务



Classification

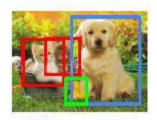


CAT



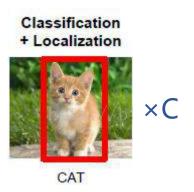


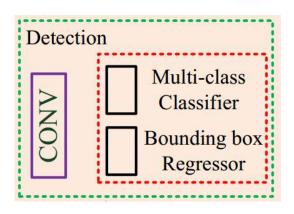
**Object Detection** 



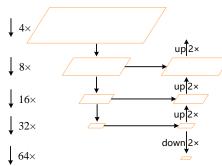
CAT, DOG, DUCK

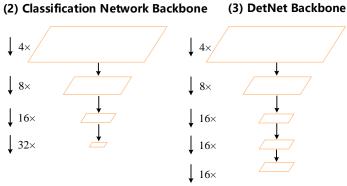
语义+细节信息







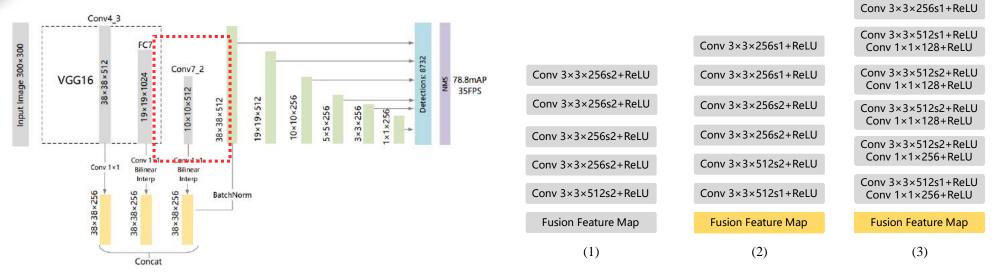


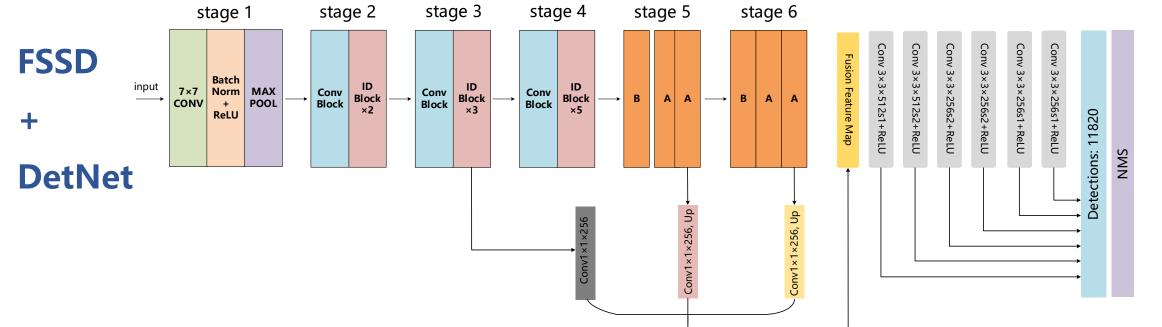


不同任务的下采样示意图



**FSSD** 









#### 测试结果分析

SSD\_DetNet可以仅靠59层达到与101层的 SSD\_ResNet的相同的效果,证明了DetNet目 标检测特征提取网络的有效性。

本章提出的算法FSSD\_DetNet在VOC 2007上取得了79.7%的效果,在小目标检测上超越了众多参与对比的算法,以微弱的速度损失比原FSSD高出了2.1%。

从与第三章算法DDSSD(1)可视化结果看,本章算法都取得了较好的效果。

Method	backbone	VOC07	coco	COCO_s	FPS	Proposals	GPU	resolutio n
FSSD	VGG	78.8	27.1	8.7	65.8	8732	GTX 1080Ti	300
MDSSD	VGG	78.6	26.8	10.8	38.5	8732	GTX 1080Ti	300
SSD	VGG	77.4	25.1	6.6	72	8732	GTX 1080Ti	300
SSD	ResNet-	77.1	28	6.2	-	11620	-	321
SSD	DetNet	77.1	-	-	61.7	11620	GTX 1080Ti	300
FSSD	DetNet	79	-	-	57.2	11620	GTX 1080Ti	300
RefineDet	VGG	80	29.4	10	40.3	6375	Titan Xp	320
DDSSD	VGG	79.7	28.3	10.5	45.4	8732	GTX 1080Ti	300
FSSD_De tNet	DetNet	79.7	29.2	10.9	58.5	11620	GTX 1080Ti	300







































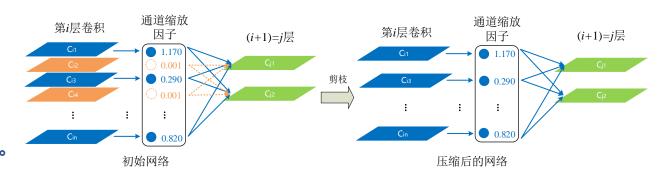
#### 模型剪枝

神经网络模型的参数通常数以亿计,存在着大量的冗余,通过模型剪枝可以有效地减去不重要的参数,使得模型更"瘦",从而达到提高模型推理速度的效果。

模型剪枝的一般流程通常较为复杂,大致可以分为: 稀疏化训练、剪枝、微调等几个流程

$$L = \sum_{(x,y)} l(f(x,W), y) + \lambda \sum_{\gamma \in f} g(\gamma)$$

其中 W 是网络权重, $\mathcal{Y}$  是设置的缩放系数,为一个超参数,用于衡量网络权重产生的损失和缩放系数引入的损失,一般设置为1. L1 正则化引入了稀疏性,



Net Slimming剪枝算法



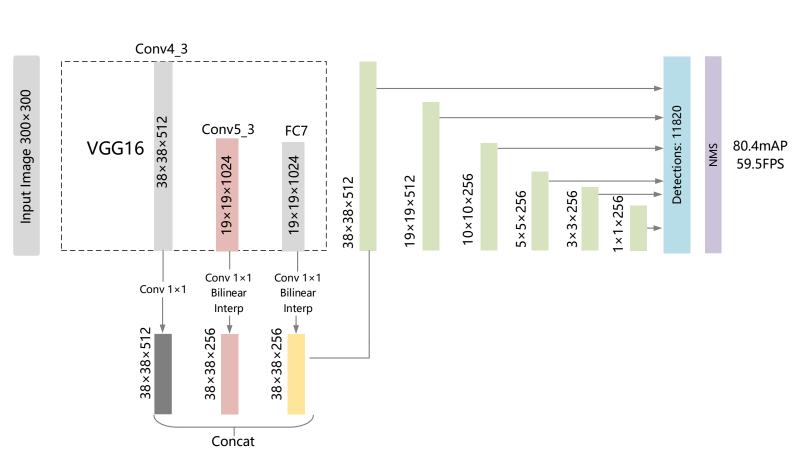
模型剪枝的一般流程





#### FSSDv2

- 选用Conv4\_3、Conv5\_3 与 FC7 进 行特征融合, Conv5\_3 的特征更有 利于小目标的检测;
- 2. 提高小目标层Conv4\_3的通道数, 即偏向于小目标的检测;
- 3. 加入BN层后,采用较大批量数 (batch size) 进行训练,可以显著 提高目标检测性能;
- 4. 对模型进行20%的剪枝,以提高其运行速度。







#### 测试结果分析

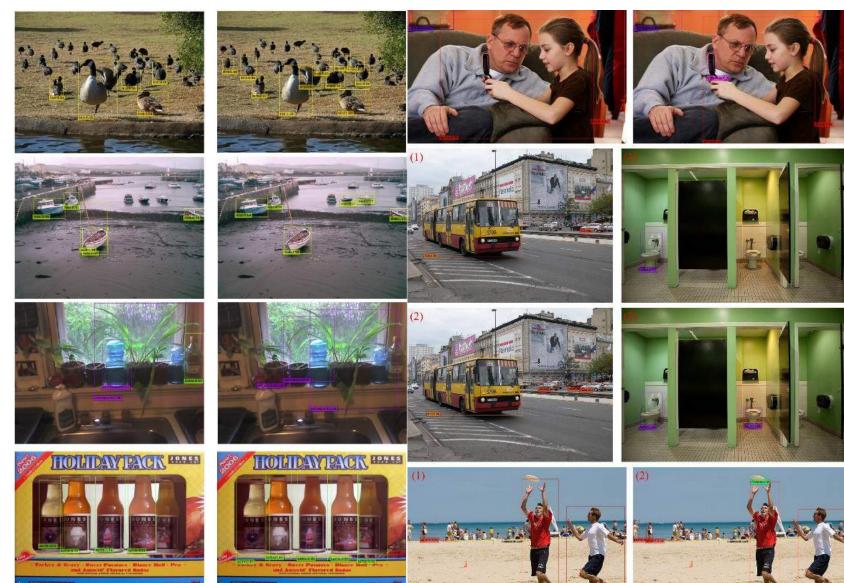
通过修改的模型FSSDv2配合BN层进行训练,可以在VOC 2007上达到80.4%,实现几乎与RFB-Net等优秀的SSD模型相同的效果,而在小目标检测上的性能可以达到12.1%的性能,比RFB-Net等模型高了2.1个百分点。

此外,剪枝20%后的模型FSSDv2\_Prune 仅以微弱的性能的损失实现了33%的加速, 取得了速度与精度之间较好的平衡。

Method	backbone	VOC2007	COCO	COCO_s	FPS	#Proposals	GPU	Input resolution
FSSD	VGG	78.8	27.1	8.7	65.8	8732	GTX 1080Ti	300×300
MDSSD	VGG	78.6	26.8	10.8	38.5	8732	GTX 1080Ti	300×300
SSD	VGG	77.4	25.1	6.6	72	8732	GTX 1080Ti	300×300
SSD	DetNet	77.1	-	-	61.7	11620	GTX 1080Ti	300×300
FSSD	DetNet	79	-	-	57.2	11620	GTX 1080Ti	300×300
RefineDet	VGG	80	29.4	10	40.3	6375	Titan Xp	320×320
RFB-Net	VGG	80.5	30.3	10	49	11620	GTX 1080Ti	300×300
DDSSD	VGG	79.7	28.3	10.5	45.4	8732	GTX 1080Ti	300×300
FSSD_Det Net	DetNet	79.7	29.2	10.9	58.5	11620	GTX 1080Ti	300×300
FSSDv2	VGG	80.4	29.4	12.1	59.5	11620	GTX 1080Ti	300×300
FSSDv2_P rune	VGG	79.9	-	-	79.5	11620	GTX₂6 1080Ti	300×300







(2)





01

### 研究背景

摘要、历史

02

### 研究现状

模型、数据集、评价标准

03

### 研究内容

空洞转置、特征提取、模型剪枝

04

### 总结展望

总结、展望

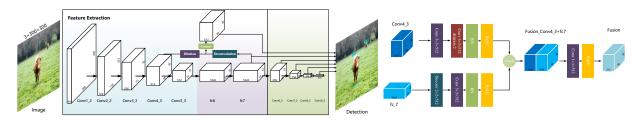


# 总结与展望

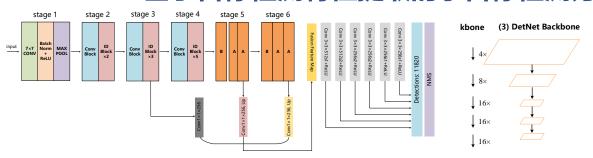


#### ✓ 总结

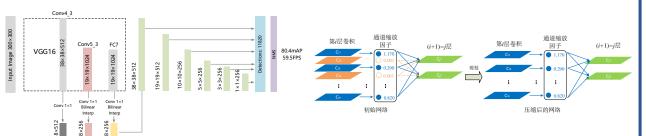
#### 1.基于空洞、转置卷积的小目标检测方法



#### 2.基于目标检测特征提取的小目标检测方法



#### 3.基于模型剪枝的小目标检测方法



#### > 展望

#### > 复杂场景下的小目标检测



#### > 资源受限的硬件环境



**Jeston Nano** 



Raspberry Pi





#### 专家一

论文结构合理、条理清晰、图表齐全、写作规范。论文反应出作者具有扎实的专业理论基础和分析解决实际问题的能力。同意进行答辩。

#### 专家二

- 英文摘要有语法错误,英文需提高。
- ▶ 研究现状的参考文献需要补充近1-2年的研究成果。
- > 对比的现参考文献格式不规范。
- > 对比的现有方法不全面,实验部分应分析论文采用的方法为什么不如现有的一些方法。

#### 答:

- 1. 已修改英文摘要中的语法错误。
- 2. 已添加近1-2年的研究成果。
- [53] H. Law, J. Deng. Cornernet: Detecting objects as paired keypoints[C]//Proceedings of the European Conference on Computer Vision (ECCV); Springer. 2018: 734-50.
- [54] K. Duan, S. Bai, L. Xie, et al. Centernet: Keypoint triplets for object detection[C]//Proceedings of the IEEE International Conference on Computer Vision; IEEE. 2019: 6569-78.
- [55] X. Zhou, D. Wang, P. Krähenbühl. Objects as points [J]. arXiv preprint arXiv:1904.07850, 2019.
- [56] C. Zhu, Y. He, M. Savvides. Feature selective anchor-free module for single-shot object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; IEEE. 2019: 840-9.
- 3. 已按照GBT7714-2015严格修改参考文献格式。
- 4. 分析部分已添加部分现有方法,并分析原因。
- [67] H. TAN, J. ZHOU, S. TIAN, et al. SFE-SSD: Shallow Feature Enhancement SSD for Small Object Detection [J]. Journal of Mathematical Research with Applications, 2019, 39(6): 733-44.
- [68] J.-S. Lim, M. Astrid, H.-J. Yoon, et al. Small Object Detection using Context and Attention [J]. arXiv preprint arXiv:1912.06319, 2019.





# Thank you!