



南京大學
NANJING UNIVERSITY

RINC



单类别实时目标检测的算法系统研究

- 答辩人: 董学文
- 导 师: 申富饶



目录

CONTENTS

- 研究背景
- 研究方法
- 实际应用
- 研究生期间工作成果



第一部分

Research Background 研究背景



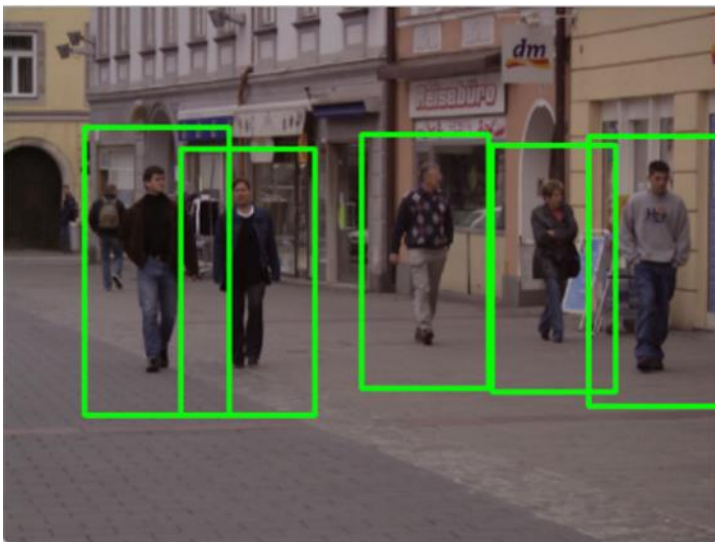
研究背景 | Research Background

目标检测：图像中所有感兴趣的目標，确定它们的类别和位置。

单类别：目标类别个数为1。

实时性：每秒检测图片数量不少于24。

应用场景：



...
S{A} 35ms



B 40ms



研究背景 | Research Background

困难与挑战:

- 实际生活场景下计算机算力往往较低，很多不会使用gpu加速运算；
- 现如今多数网络模型参数量过大，在普通算力计算机上无法实时检测。

网络模型	COCO AP@0.5	FPS
YOLO v2	43.81	12.5464
mobilenet v1 ssd	40.24	16.0845
mobilenet v2 ssd	40.61	15.4497



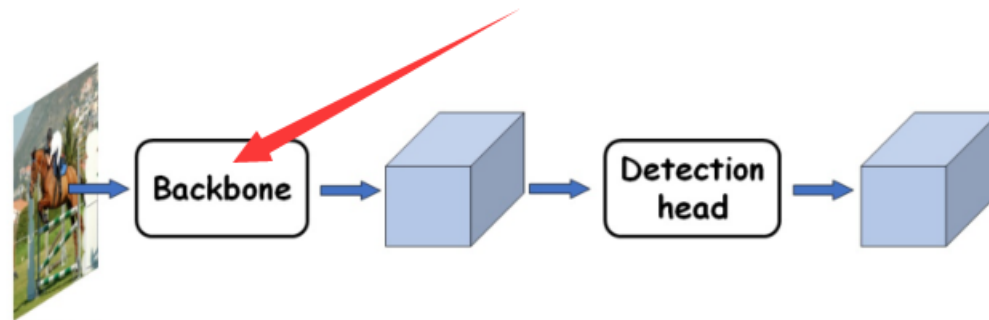
第二部分

Proposed Methods 提出方法



针对上面的问题，从两个方面来分别进行优化，最终达到实时性检测的目的。

- 特征提取模型的架构优化
- 针对单类别目标检测的剪枝算法





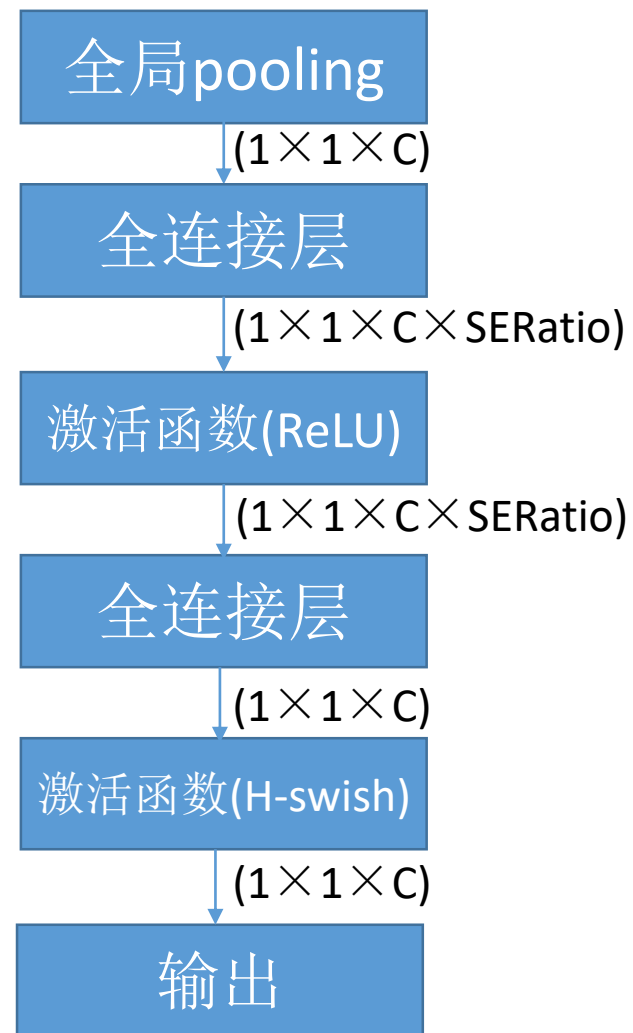
由于本文的目的是做一个可以在普通算力计算机上能够实时检测的单类别目标检测模型，因此模型结构优化基模型选择性能良好参数量极少的mobilenet v3网络模型。

网络模型	Params(Million)	MAdds(Million)
VGG-16	138	15300
inception v3	23.2	5000
mobilenet v1	4.2	569
mobilenet v2	3.4	326
mobilenet v3 small	2.9	66



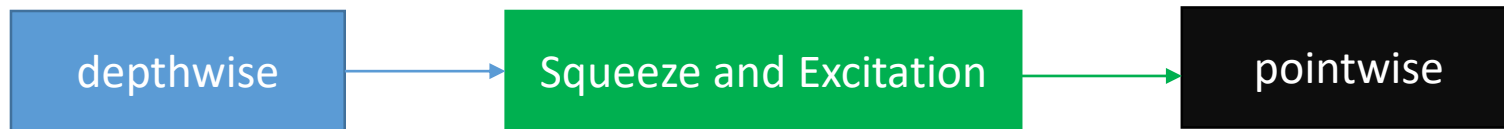
分析mobilenet v3网络可以发现网络参数主要集中在深度可分离卷积部分(depthwise, pointwise)和压缩-激励部分(Squeeze and Excitation), 其中SE块内部逻辑如右侧所示:

实际上, 由于SERadio的引入, SE模块的运算部分主要集中在全局pooling中。因此为了减少这部分的参数量, 优化模型将SE模块放入pointwise之后。





原mobilenet v3网络模型:



更改后的mobilenet v3网络模型:



更改后的mobilenet v3网络模型与原模型相比参数量减少约7%.

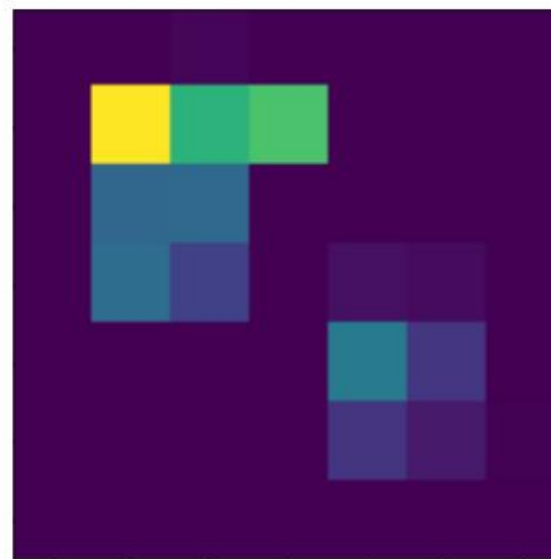
网络模型	Params(Million)	MAdds(Million)
VGG-16	138	15300
inception v3	23.2	5000
mobilenet v1	4.2	569
mobilenet v2	3.4	326
mobilenet v3 small	2.9	66
mobilenet v3 small ₂	2.7	63



在目标检测任务执行过程中，模型往往是通过判断物体的某个特征从而进行分类。同一层中的不同卷积核对该特征的提取有不同程度的影响。可以考虑与目标特征相关性较小卷积核删除从而做到削减网络参数，加快模型运行速度的目的。



(a) 神经网络输入原图



(b) 网络模型最终识别区域



由于卷积层输出与每个卷积核参数已知，因此可以通过偏导判断每个卷积核对目标类别的作用。

$$\gamma_i = \frac{1}{N} \sum_{n=1}^N \frac{\partial Z}{\partial A_i(x_n)}$$

最终将会得到一个长度为通道数的数组，每个值表示该卷积核的激活程度。可以理解成“该卷积核是否同意将结果归类为目标类别”。



规定数值较大的前20%为激活，后20%为抑制。因此剪枝的任务就是如何保留对分类结果影响比较大的卷积核，使其在剪枝前后激活程度变化最小。

- 将实数值最小的卷积核删除(A)
- 将绝对值最小的卷积核删除(B)
- 将平均值附近的卷积核删除(C)

剪枝策略	剪枝之前		剪枝之后		stable
	inhi_bound	act_bound	inhi_bound	act_bound	
L1 剪枝策略	-2.04×10^{-7}	2.69×10^{-7}	-4.27×10^{-5}	3.19×10^{-5}	529
L2 剪枝策略			-1.31×10^{-5}	1.27×10^{-5}	527
A 剪枝策略			-1.05×10^{-12}	8.31×10^{-12}	400
B 剪枝策略			-4.28×10^{-7}	5.59×10^{-7}	637
C 剪枝策略			-2.52×10^{-7}	3.40×10^{-7}	664



采用算法C剪枝后的网络模型参数量得到了进一步的减小，结合上一步优化策略后最终模型在cpu i7 6700 gtx 960的普通算力计算机上运行单类别目标检测任务可以达到实时性的要求。

网络模型	VOC2012	COCO	FPS(每秒帧数)
YOLO v2	72.06	42.06	12.9365
原 mobilenet v1 ssd 模型	71.14	40.88	16.1893
原 mobilenet v3 small ₂ ssd 模型	68.09	36.12	22.0683
mobilenet v1 L ₁ 剪枝	33.86	16.49	16.7092
mobilenet v1 L ₂ 剪枝	41.23	20.31	
mobilenet v1 ssd-A 方案	44.29	9.72	
mobilenet v1 ssd-B 方案	68.94	38.61	
mobilenet v1 ssd-C 方案	71.49	39.71	
mobilenet v3 small ₂ ssd-C 方案	70.18	37.41	24.4359

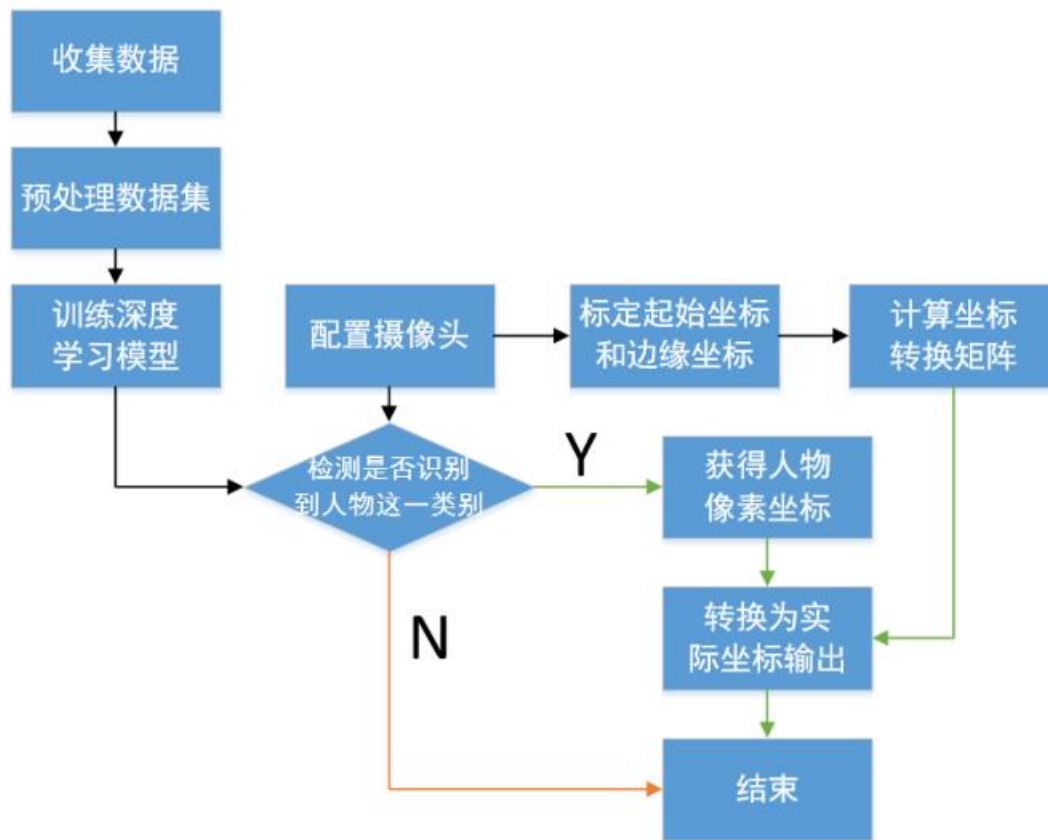


第三部分

Applications 实际应用



基于上述工作制作了一个室内定位系统，该室内定位系统可以实时性的在室内检测人物的位置，支持多人检测。





系统整体运行稳定，目前已经在演示教室成功配置。





第四部分

Work and Research Progress 研究生期间工作成果



专利： 申富饶, 董学文, 赵健, 李俊 “一种基于深度学习的室内人物定位方法” 专利, Dec. 2020.

项目： 国家自然科学基金面上项目 “基于深度感知增量式联想记忆神经网络的信息融合系统研究, Information fusion system based on deep perception and incremental associative memory neural networks” (课题年限 2019.01~2022.12), 负责神经网络模型相关研究。



南京大學
NANJING UNIVERSITY



谢谢!

