

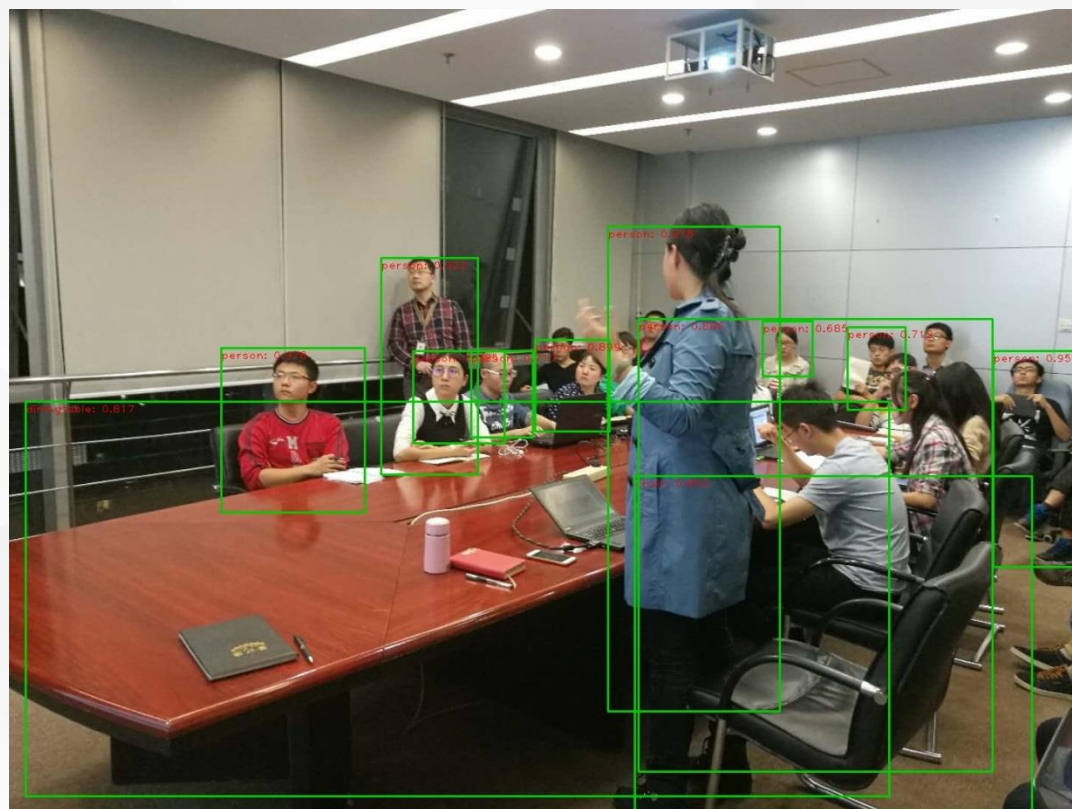
A closer look at the local module in 《Iterative Visual Reasoning Beyond Convolutions》

Xiaokai Chen

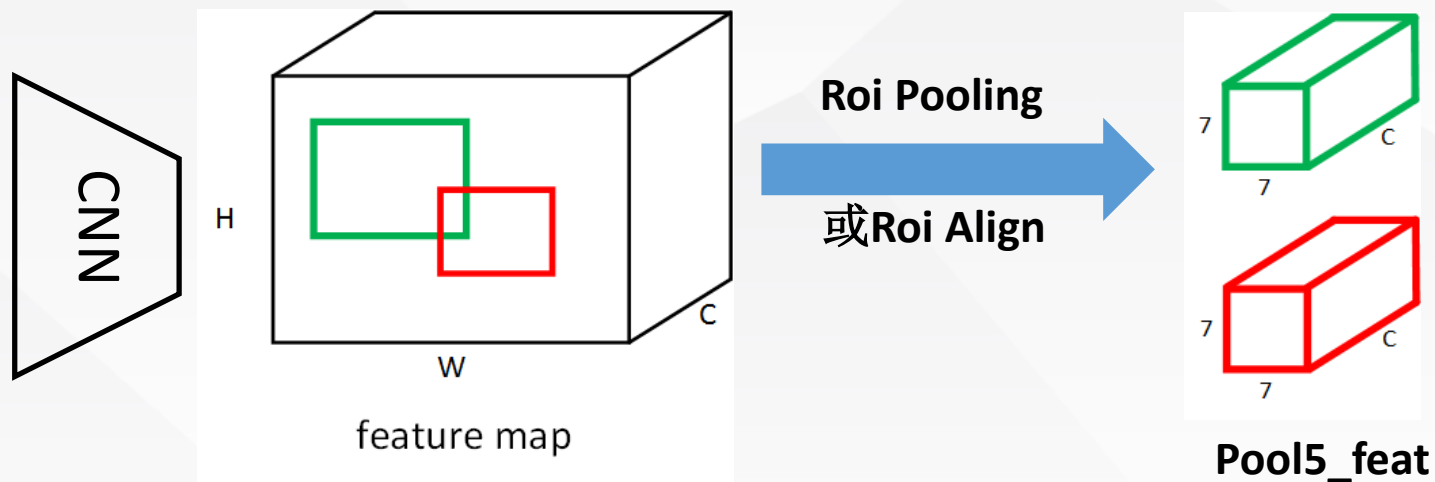
2018-05-03

➤ 进一步理解：局部推理模块（先不管和全局模块的交互）

- 具体任务：区域识别
 - 系统输入：图片，相应的bounding boxes
 - 系统输出：各bounding boxes的类别



➤ 一种直接的思路



- 这其实是文章中的baseline。
- 不足：实体识别是互相独立的^[1], 然而实际上它们之间可能有关联。

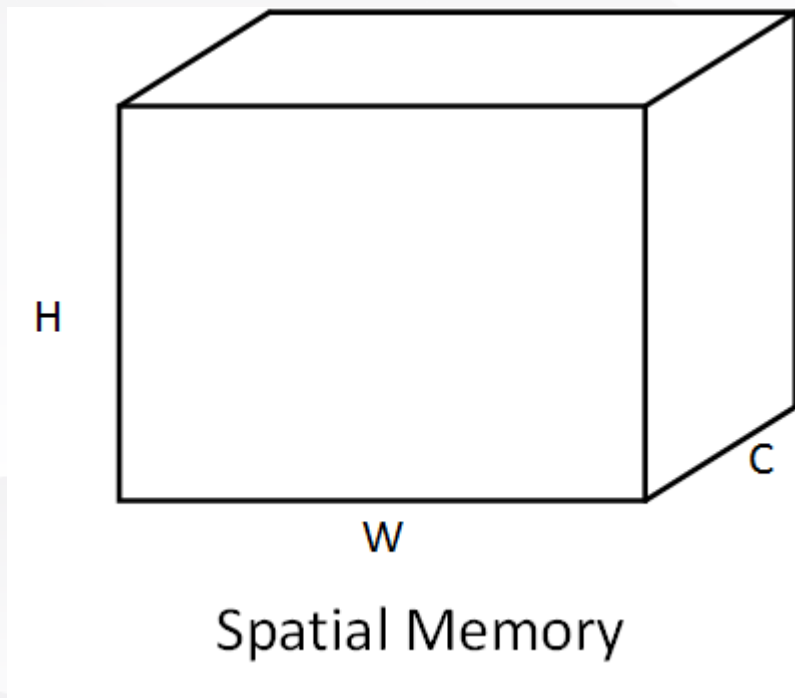
[1]. Spatial Memory for Context Reasoning in Object Detection

显式记忆：Spatial Memory

- 使用显式的记忆存储之前所有区域的预测结果，用于下一次迭代。
- 疑问：
 - 显式记忆什么样子？存储的是什么？
 - 如何更新它？

显式记忆什么样子？存储的是什么？

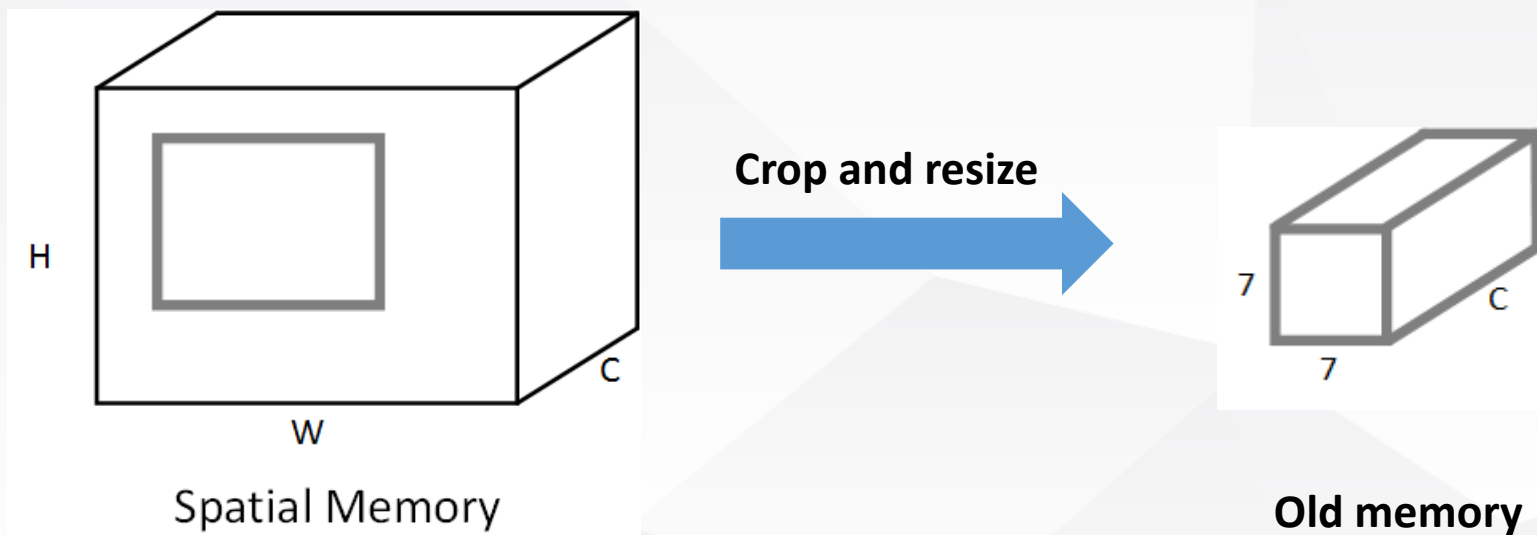
- $H*W*C$ ，和特征图的结构类似
 - 论文中H、W是输入图像尺寸的1/16 (4次pooling)
 - 好处，可以保留原图中各区域的布局，没有破坏空间结构
- 存储内容：各个区域的局部特征以及高层特征。



Softmax之前的
得分向量

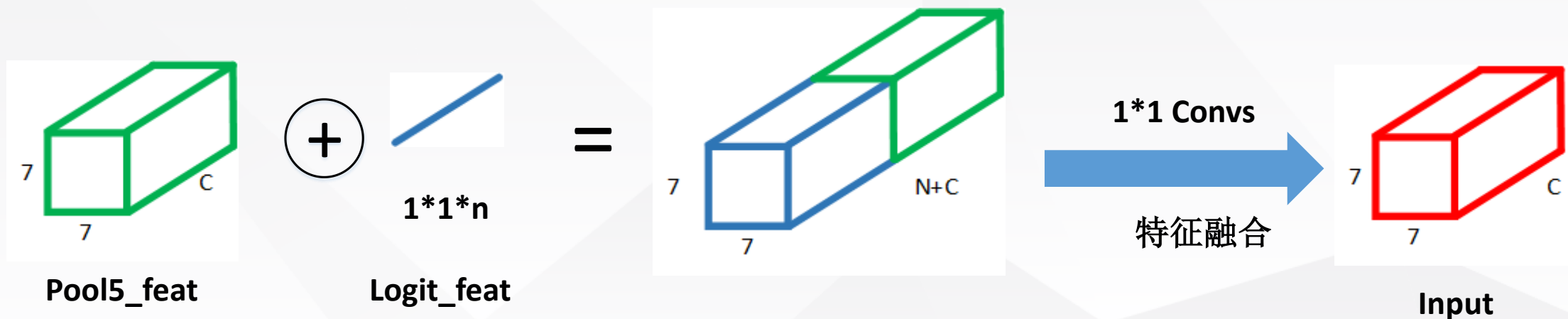
➤ 如何更新显式记忆？

- 以区域为单位更新，对于每个区域，我们需要：
 - 获取“旧”记忆



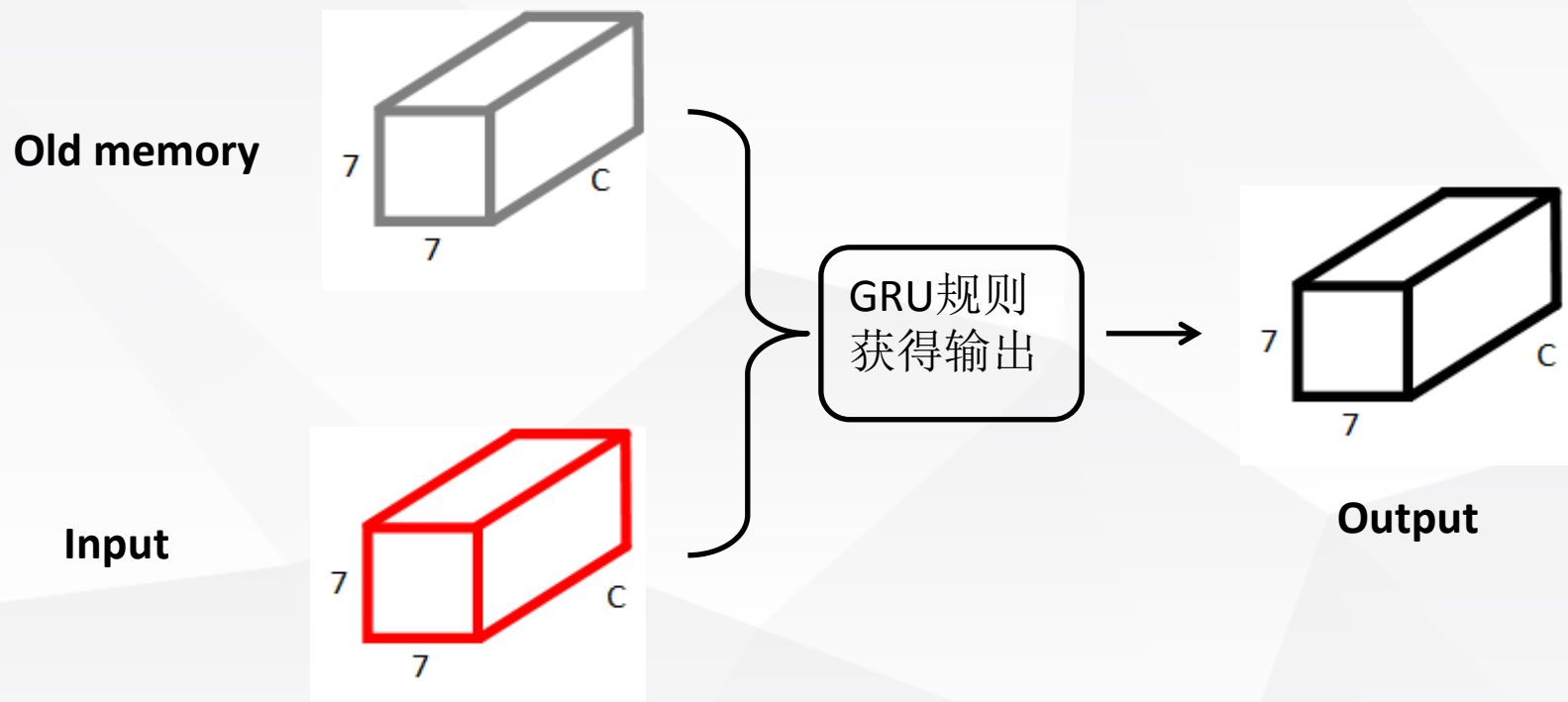
➤ 如何更新显式记忆？

- 确定输入，pool5_feat与分类得分向量logit融合
- 一个样本的迭代过程中，Pool5_feat不动，改变的是Logit_feat



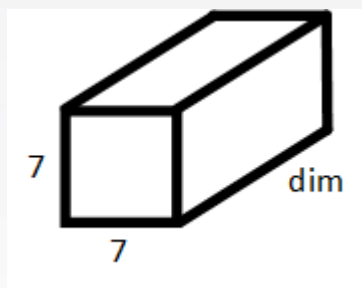
➤ 如何更新显式记忆？

- 确定更新规则，参考GRU方程，卷积替代矩阵乘法



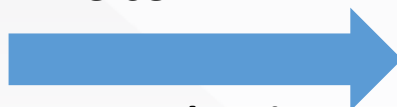
>> 如何更新显式记忆？

■ 填回去

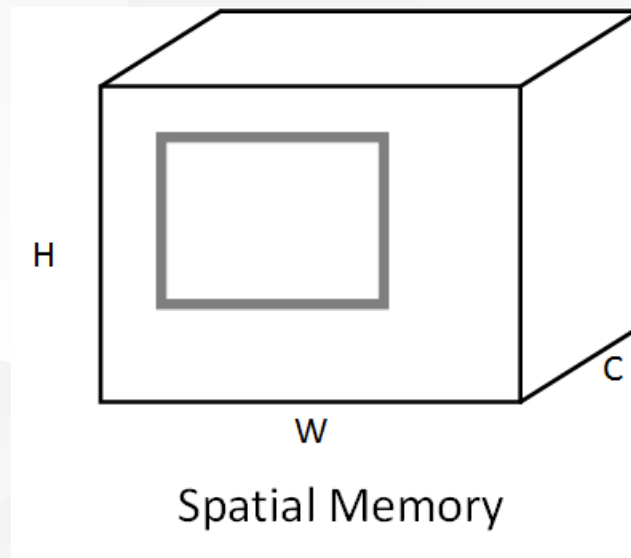


Output

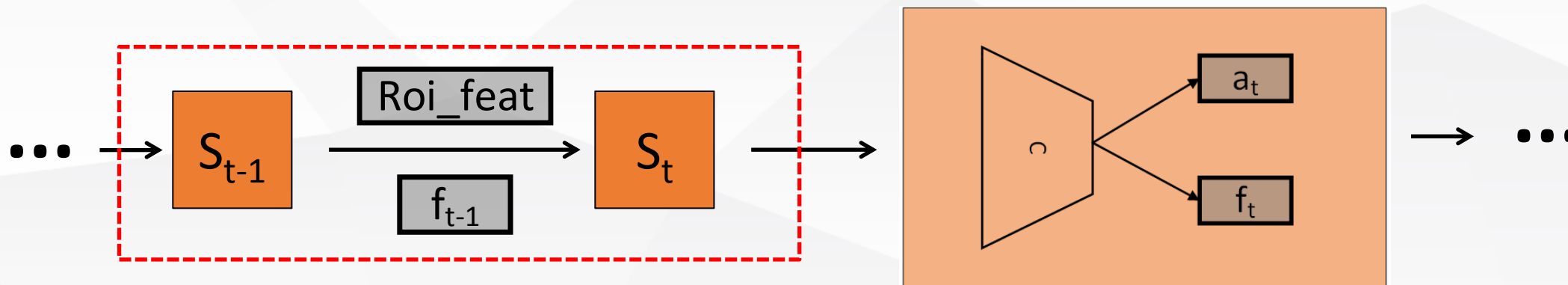
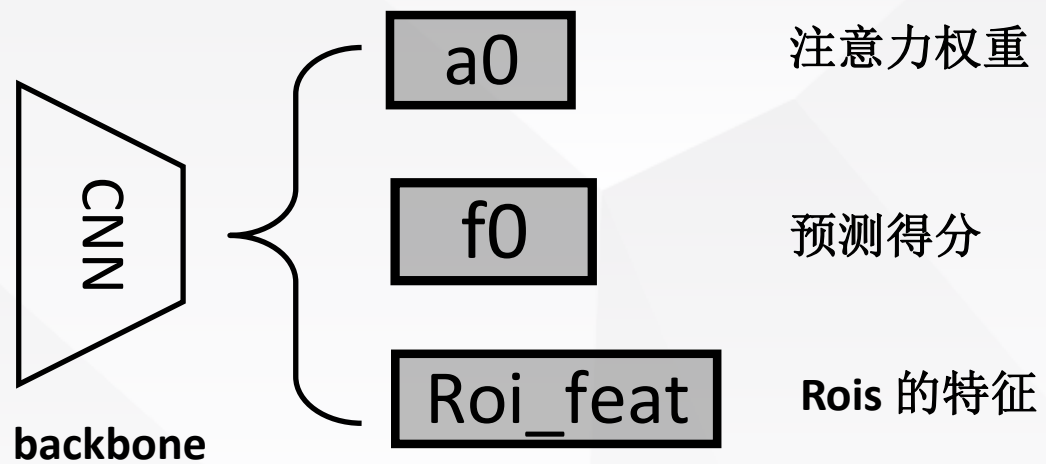
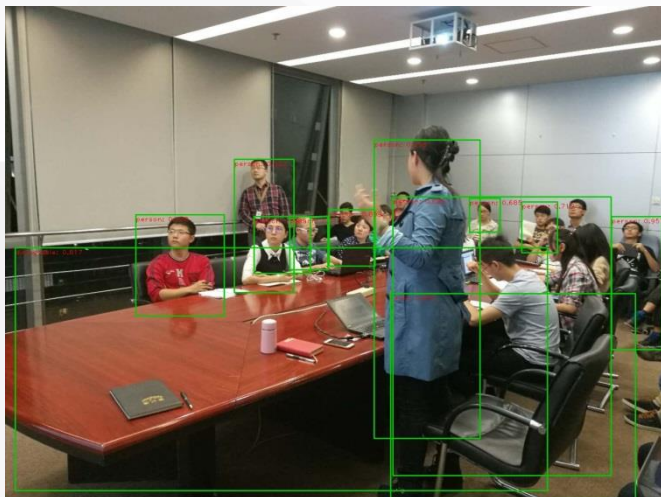
Inverse



Crop and resize



整体结构



The End
Thanks