因果推断在机器学习上的应用——以WSSS上的应用为例

郑东森 202010799 计23

机械学习里，特别是计算十分复杂的计算机视觉领域，工程师往往需要通过在极大的样本集合上进行训练，从而把问题转化到i.i.d领域。但现实情况下，这往往是不被允许的，数据获取成本、分布变化的难以控制，都很可能会导向失败。近年来，学者们尝试整合了因果推断与机械学习，很好地处理了以上问题。本次的大作业我会从传统因果推断讲起，以因果学习在WSSS上的应用为例，说明因果推断在机器学习上的应用。

1. **传统因果推断[[1]](#footnote-1)**

Notation:

T: 所观察的治疗/政策…

Y: 观察的结果

C: 后缀修改集合（sufficient adjustment set）

:在治疗1下的Potential Outcome

传统因果推断的目标就是通过计算Potential Outcome 来衡量一个治疗/政策…对某个结果的因果性，对于个人来讲，这个值被称为ITE (Individual Treatment Effect) ,通过将治疗前后的效果相减进行计算：。由于不可能同时知道同一个个体的两种结果，传统因果推断中，更多使用的是ATE (Average Treatment Effect)，将个体转化为期望值来计算 。一般情况下使用随机对照试验来统计ATE。

因果推断中最重要的一点便是阐明了因果性不等于相关性，在计算ATE时，要么要求数据是 i.i.d 的，要么就要用Unconfoundedness假设。Unconfoundedness假设是通过给每条因果图上的后缀边的后缀集合添加不同权重，从而让这条边在因果图上断开的方法，具体即以下公式（以图1为例）：

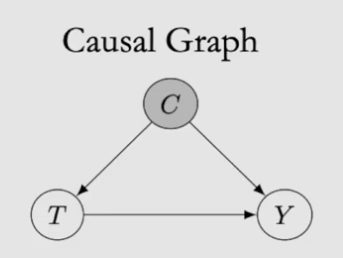


图1.

这样我们就相当于断开了后缀边 C->T

1. **因果推断在WSSS上的应用[[2]](#footnote-2)**

WSSS指弱监督语义分割Weakly-Supervised Semantic Segmentation。语义分割就是把图片中每个元素的名称和范围标记出来（如图2），监督语义分割的数据成本极高，大约要1.5天/人才能处理好一张图片，而弱语义分割则是人为指出图片中元素，再结合分割模型来实现类似功能（如图3）。

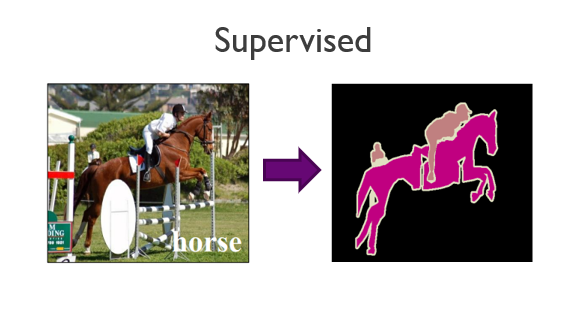


图2.

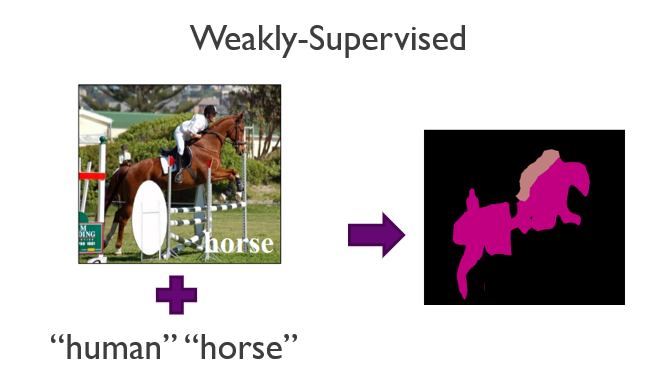


图3.

虽然WSSS很节省成本，但它有一个致命问题就是常常有虚假关联（如图5），如（一）中介绍的，因果推断可以用Unconfoundedness假设去掉虚假关联，于是这里可以应用因果推断来解除虚假关联。

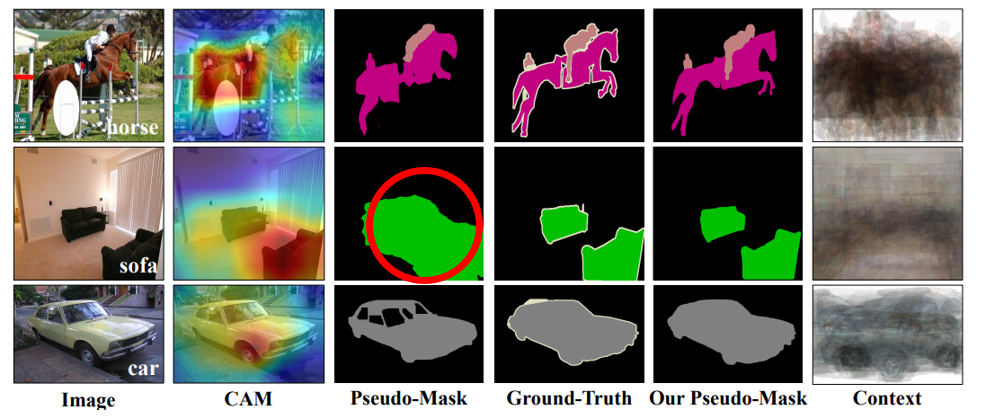


图5.

WSSS中的虚假关联主要是由上下文关系（注：视觉场景中物体一般的关系）产生的，可以分析到，是上下文关系影响了模型对特定沙发地板排布的表示和原始图像的判断，故可以有以下因果图（图6）：

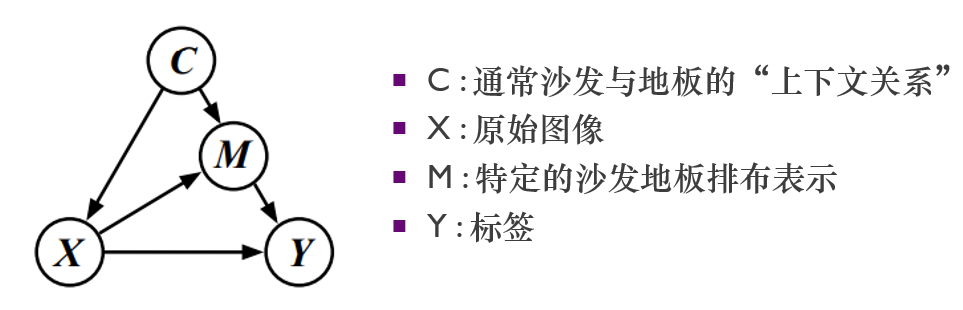


图6.

这里我们可以通过将上下文后缀集合C分成若干块再施加权重进行Unconfoundedness假设：

这样我们就可以把原本模型中的统计量P(Y|X)转化为因果量P(Y|do(X))，从解除虚假关联。从整体训练流程上，就做了如图7到图8的改变：

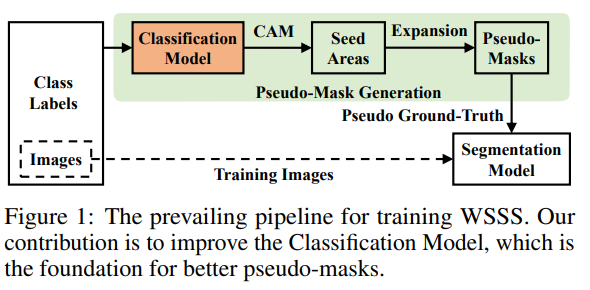


图7. 原本训练流程

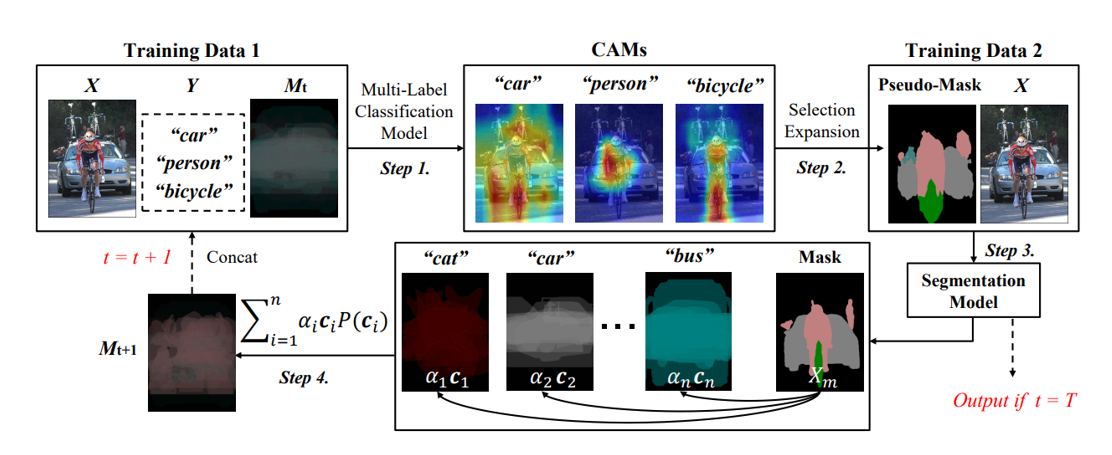


图8.改善后的训练流程

最终成果如图9，因果学习很好地处理了虚拟关联问题：

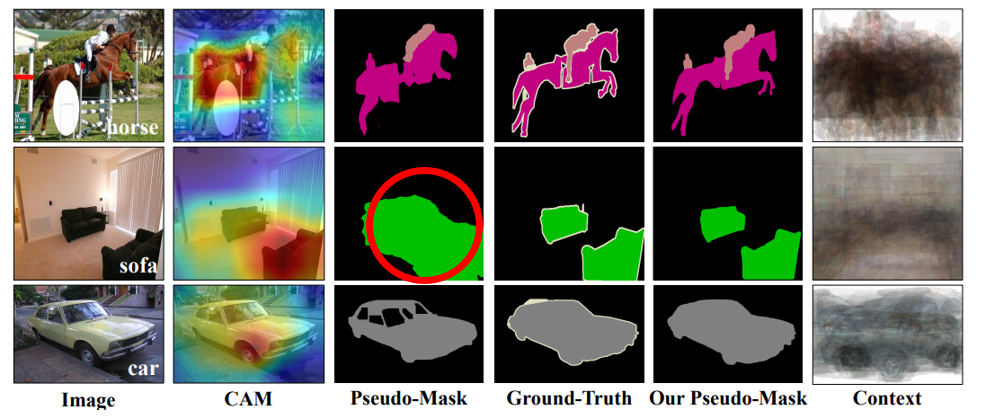


图9. 最终成果

1. Brady Neal: *Introduction to Causal Inference*, https://www.bradyneal.com/causal-inference-course [↑](#footnote-ref-1)
2. Zhang, Dong, et al. "Causal intervention for weakly-supervised semantic segmentation." Advances in Neural Information Processing Systems 33 (2020): 655-666 [↑](#footnote-ref-2)