

Self-Supervision learning

什么是自监督学习

- ▶ 在我的理解中，自监督学习就是在没有标注数据的情况下，找寻数据内部潜在的信息作为监督信息进行学习，本质上就借助于利用特征的匹配完成判断任务。
- ▶ 这个过程就像婴儿的学习，婴儿的学习初期是无法理解所谓的“label”的，他的学习就是一种自监督的学习，通过对特征的匹配做出自己判断。
- ▶ 我们首先对无标签的数据集进行学习，在对比不同类别之后提取同类别内的特征，这些特征足以让模型在面对测试中同类型输入时可以将其归为同一类。当模型得到了足够多的可以用于区分类别的特征后，再将模型迁移到其他判断任务中。这也就是为什么说，一般情况下数据越多，训练时间越长，自监督学习模型的效果往往越好。

自监督应用于video object tracking

- ▶ 利用需要追踪的目标的数据连贯性，自监督学习可以很好的找到需要关注的物体的特征，并找到连续的frame之间的联系，从而很好的实现对目标的追踪。比如目标在移动时在时间层面是连续的 t 和 $t+1$ ，在图像层面也是连续的变化，而不会产生离散的突变。利用这样的内在联系和目标的特征就可以很好的实现对目标的追踪。

声音与影像的匹配

- ▶ 声音与影像的匹配与论文中相同，分别将声音与影像经过卷积、池化和全连接层得到相同纬度的向量，串联到一起后输入全连接层，最后将卷积层的结果通过softmax得到视听匹配程度的输出。后面介绍了一个在声音与影像匹配的基础上，还要做到对声音的定位，网络的结构与前者不同，在对视听的匹配程度输出的同时还输入了对声音的定位。

phases

- Talk的最后提出了三个phases, 分别是classic、expansion和uncurated。分别告诉我们: 1、用自监督学习的方式进行表征学习的效果比强监督学习更好; 2、更多的数据、更长的训练、更复杂的网络, 对于自监督学习来说都是好事, 而且自监督学习可以更好的应对多模态问题。3、使用没有标注的数据做自监督学习可以得到更通用的模型, 可以更好的胜任多模态任务。但是标注数据也不是一无是处, 他可以用于检验和评估模型。

Curiosity-driven Exploration by Self-supervised Prediction

- 提出的原因：强化学习需要收到agent输出当前的return的最大期望，依次作为更新网络的依据。然而有时agent给出的reward是稀疏甚至是无法给出的，只有当代理达到了最终的目标状态之后才会返回一个return，进行网络的加强，这是此类强化学习任务的最大问题。

- ▶ 与传统的强化学习不同的地方在于：只研究由于agent做出相应动作后对环境产生的影响，而忽略环境的其他因素。处理的不再是输出的所有环境和action，而是在t下的与agent做出action有关的环境变化和t的action，将两者转化到feature-space内，预测feature-space的变化，以预测的误差作为驱动模型学习的内在“reward”，再将内在的“reward”与实际agent产生的外在action加和，作为当前时刻的总reward。
- ▶ 这样的设计思路可以潜在的应用在不同的agent内（因为他其实不妨碍通常的策略网络更新，只是对reward添加了一项可以多方面评估策略的curiosity项）
- ▶ 为了学习生成feature-space的网络，将 s_t 和 s_{t+1} 作为输入，预测 a_t ，通过误差更新网络；为了学习生成curiosity-reward，通过输入 s_t 和 a_t ，预测 s_{t+1} ，通过误差更新网络，其中curiosity-reward_t为 $||\text{feature-space}(s_t) - \text{feature-space}(s_{t+1})||$

测试结果

- ▶ 在不加入外部reward的情况下，仅加入curiosity-reward的模型可以闯过第一关的30%。
- ▶ 添加curiosity-reward和外部reward的模型在VizDoom的测试中可以更快的完成任务。
- ▶ 仅使用curiosity-reward训练的模型在VizDoom中学习更快，测试效果更好。

综合感想

- ▶ 自监督学习往往比有监督的学习效果好，可以更好的胜任多模态的任务。
- ▶ 自监督学习渴望更多的数据集，更多的训练轮次，more is better。
- ▶ 不能局限于现有的数据集，模型不能停止学习，特别是自监督学习的模型，需要学习越来越新颖的特征。比如我们在day1是模型1，得到day2的数据后可以进行验证和测试，但是最后要将day2的数据也学习进来，加入到训练集中，得到更新的模型。这样外部的数据在逐步更新，我们的模型也随之适应。
- ▶ 当时印象中在下半场的talk1和talk3中提到了两个不同的说法，印象里有一个内容是如果我们将图像翻转，那么图像的内涵可能发生改变（当时是猎人与熊，追与被追在翻转后发生变化），但另一个talk内说我们渴望更多更新的数据集，当得到新的数据是，我们可以通过对他翻转等处理，获得更多的信息，比如一只鸟的头部，可以翻转不同角度。我有一些不理解这样的说法，可能区别在于局部的匹配和全局的匹配？