|  |  |  |
| --- | --- | --- |
| **文章信息** | **摘要及贡献** | **动机方法概述** |
| **阅读程度**：  仅浏览  **文章标题：**  Unsupervised Feature Learning via Non-Parametric Instance Discrimination  **中文标题：**  通过非参数实例区分进行无监督特征学习  **发表于：**  CVPR2018  **作者：**  Zhirong Wu  **单位：**  CUHK  UC Berkeley / ICSI | 在带有注释类标签的数据上训练的神经网络分类器也可以捕获类别之间明显的视觉相似性，而不被指导这样做。  我们研究这一观察结果是否可以扩展到传统的监督学习领域之外：我们能否学习一个好的特征表示，通过要求特征对个体实例的区别来捕获实例之间而不是类之间的明显相似性？  我们将这种直觉表述为实例级的非参数分类问题，并使用噪声对比估计（NCE）来解决大量实例类所带来的计算挑战。  我们的实验结果表明，在无监督学习设置下，我们的方法大大超过了最先进的ImageNet分类。（先进性）  我们的方法在使用更多的训练数据和更好的网络架构来持续提高测试性能方面也很显著。（具有潜力）  通过对学习特征的微调，我们进一步获得了半监督学习和目标检测任务的竞争结果。（向分类任务以外的其他任务扩展）  我们的非参数模型是非常紧凑的：每张图像有128个特征，我们的方法只需要100万张图像的600MB的存储空间，能够在运行时快速进行最近邻检索。（存储空间不大且运行快） | 本文的动机来自于监督学习，由图可以观察到一个现象，监督学习只教会网络将图片归类到其对应的类别中，并没有主动告诉网络那些类别较为相似，那些类别不想似。如雪豹和豹这两类，比雪豹和车这两类更为相似，网络输出的关于雪豹与豹这两类的概率值较为接近，事实上我们并没有告诉网络雪豹和豹比较相近，网络是在学习过程中自发发现雪豹与豹这两类较为接近。那我们可以将这一发现进行推广，如果我们将每张图片视为一个类进行区分，网络将能够自发地学会那些图片看上去比较接近。这就是instance discrimination这个代理任务的设计动机。    图片经过CNN特征提取得到128维的特征向量，将特征向量的欧式长度进行归一化即得到归一化特征向量。  Memory bank存储了所有图片的归一化特征向量，图片数量记为n，我们可以将Memory Bank视为一个全连接层，输入为128，输出为n（n中第i个值越大，意味着CNN提取到的特征f与Memory Bank中的第i个特征向量的余弦相似度最高，就认为输入样本被分为第i类，即输入样本是训练集中的第i张样本），这个全连接层中的权重是所有训练样本的归一化特征向量，并不是可训练参数w，这就是标题Non-Parametric的意思。  Memory Bank最开始是随机初始化的，每个minibatch训练结束后，用最新CNN提取到的归一化特征向量更新Memory Bank中对应的特征向量，即每个batch只更新个别特征向量，每个epoch后Memory Bank中的所有特征向量被完整更新一次。  理论上Memory bank视为全连接层，128维的特征编码输入后得到n维的输出，n维的输出应该与真值（即输入图片是训练集中第几张图片，这个信息实际上是不用进行人工标注的）进行交叉熵损失函数计算，但此时类别数为n，n等于训练集中的样本数，显然这是极大的开销，因此引入NCEloss。  NCEloss的基本思路是从MB中采样N+1个特征向量，其中必有一个是输入图片实际对应的特征向量认为是正样本，而其他N个认为是负样本，此时相当于把N+1个特征向量分为两类，正类和负类，只需要增大f与正类特征向量之间的余弦相似度，减小f与负类样本之间的余弦相似度即可。（其实并没有看懂这里的数学推导，太鸡儿复杂了） |