

作者：小哲

微信公众号：小哲AI

原文来自微信公众号

github: <https://github.com/lxztju>

自监督学习之SimCLR

1. 论文摘要

本文提出了SimCLR：一个生成视觉表达（representation）的简单对抗学习框架。我们简化了近些年提出的对抗学习的框架，我们的算法不需要特殊的结构和庞大的存储体（memory bank），为了理解到底什么使得这种对抗学习能够学得有用3的表达，我们系统的研究了我们框架中的基本的组件：**1. 数据增广在高效的视觉预测任务中起很大的作用。2. 引入了在表达的非线性转换与一个对抗损失，这种非线性转换与损失函数能够提升可学习表达的质量。3. 相比与监督学习更大的batchsize与更多的训练步骤对于对抗学习更加有用。**混合上边介绍的三个发现，我们的在ImageNet上实现的最优结果超过了目前最优的半监督与自监督学习方案。利用SimCLR得到的表达训练一个简单的线性分类器在ImageNet上可以得到76.5的TOP1正确率，这个结果超过了之前resnet50上的最优结果7%。我们使用1%的label数据机型微调在ImageNet上可以得到85.8的top5准确率，这使用了比AlexNet少100倍的label数据。

总结说来论文提出了三点：

1. 数据增广有很大的作用
2. 提出了一个非线性转换
3. 更多的训练步骤与更大的batchsize对于对抗训练的作用相较于有监督训练更大

2. 论文主要思想

算法的主要操作就是无监督的预训练与有监督的微调实现自监督学习。

其中最主要的就是无监督的预训练。采用数据增广与非线性转换与对抗损失函数。

1. 数据增广经过论证后证明采用随机裁剪与色彩变换还有高斯随机噪声的配合能够更加有效
2. 在resnet50得到表达之后，采用带有隐层的MLP网络进行非线性转换
3. 引入一种对抗损失函数。

$$\ell_{i,j} = -\log \frac{\exp(\text{sim}(z_i, z_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{[k \neq i]} \exp(\text{sim}(z_i, z_k)/\tau)},$$

这个公式类似于KL散度，分子表示同样类别的相似性（这里的同类别指的是同一张图像的数据增广出来的图像），分母表示不同类别之间的相似性（表示不同图像增广出来的数据样本）。这里的相似性度量采用cosine相似性。

论文地址: [A Simple Framework for Contrastive Learning of Visual Representations](#)