



微调



标注一个数据集很贵



我的数据集

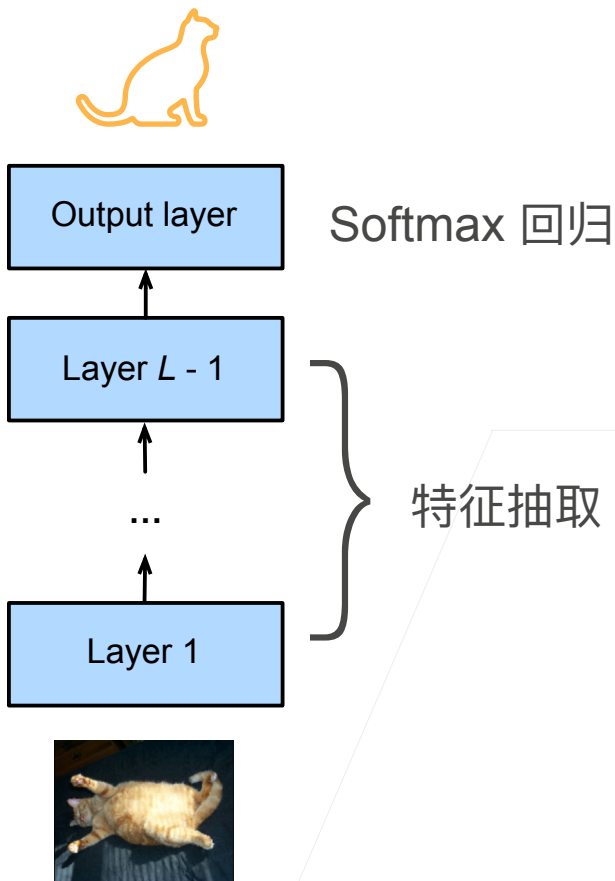


样本数	1.2 M	50K	60 K
类别数	1,000	100	10

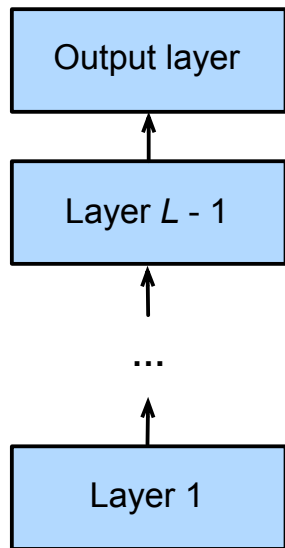
网络架构



- 一个神经网络一般可以分成两块
 - 特征抽取将原始像素变成容易线性分割的特征
 - 线性分类器来做分类



微调



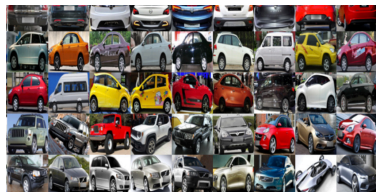
不能直接使用，因为标号可能变了



可能仍然可以对我数据集做特征抽取



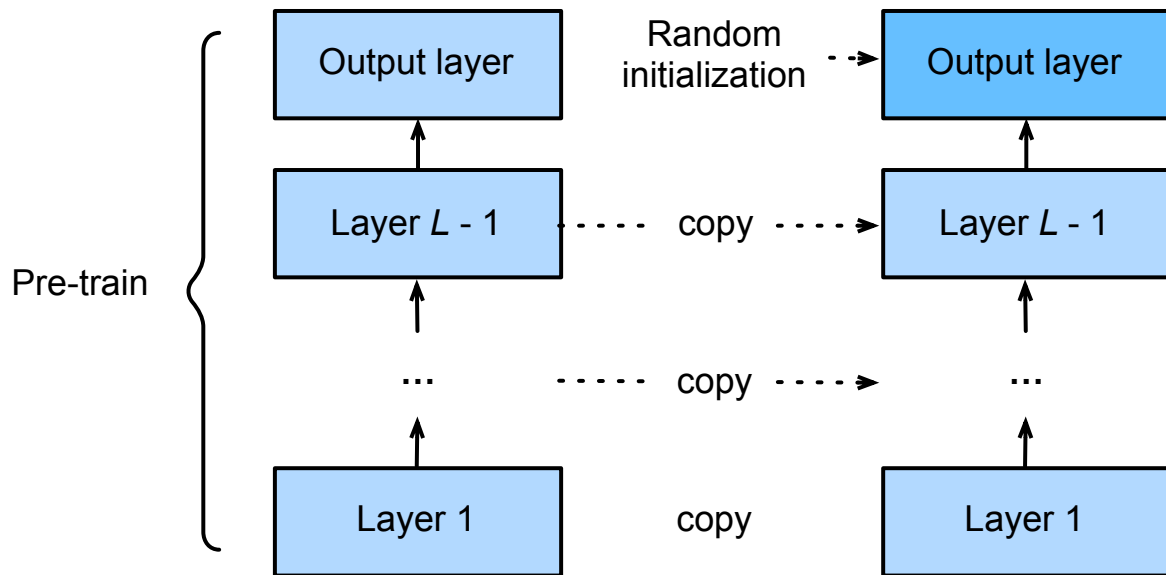
源数据集



目标数据集



微调中的权重初始化



源数据集



目标数据集





训练

- 是一个目标数据集上的正常训练任务，但使用更强的正则化
 - 使用更小的学习率
 - 使用更少的数据迭代
- 源数据集远复杂于目标数据，通常微调效果更好



重用分类器权重

- 源数据集可能也有目标数据中的部分标号
- 可以使用预训练好模型分类器中对应标号对应的向量来做初始化



Racer, race car, racing car

A fast car that competes in races

Treemap Visualization

Images of the Synset

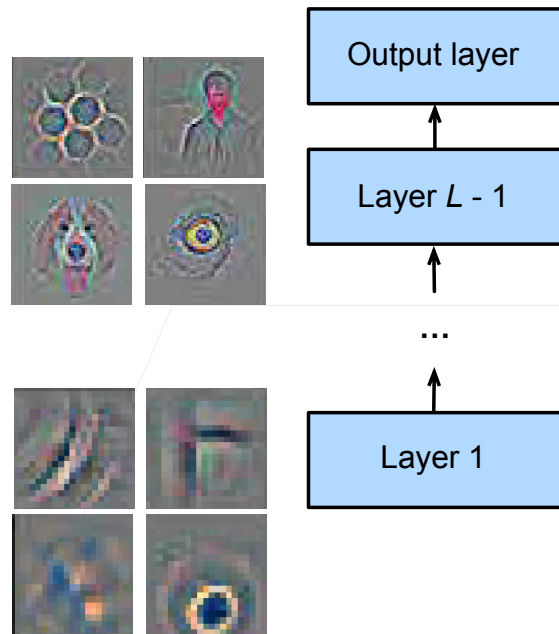
Downloads



固定一些层



- 神经网络通常学习有层次的特征表示
 - 低层次的特征更加通用
 - 高层次的特征则更跟数据集相关
- 可以固定底部一些层的参数，不参与更新
 - 更强的正则



总结



- 微调通过使用在大数据上得到的预训练好的模型来初始化模型权重来完成提升精度
- 预训练模型质量很重要
- 微调通常速度更快、精度更高