

小样本学习背景

专业性较强领域数据标注成本高昂
由于数据隐私导致可用样本少
部分领域本身就是小样本问题

小样本学习基本要素

support set 支持集
query 查询样本
training set 训练集

有监督学习 测试样本从未见过
小样本学习 测试样本是从已知类里来的
小样本学习 值得查询样本从未见过
查询样本是从自己也见过

k-way: 支持集里的类别数
n-shot: 支持集一个类里有多少样本



相似度函数学习

学习一个相似度函数 $\text{sim}(x, x')$



理想 $\text{sim}(x_1, x_2) = 1$, $\text{sim}(x_1, x_3) = 0$

使用相似度函数进行预测，对比查询样本与支持集每个样本的相似度找到得最高者

元学习 meta learning: 学习如何学习 (learn to learn)

小样本学习是一种元学习

- 利用多个任务的先验信息辅助实现新的任务
- 使网络具备学习的能力
- 需要准备许多任务来进行学习

Goal

Machine Learning = find a function f

Dog-Cat Classification $f(\text{dog}) = \text{"cat"}$

Meta Learning ≈ find a function F that finds a function f

Learning Algorithm $F(\text{cat}, \text{dog}, \text{cat}, \text{dog}) = f$
Training Examples

Machine Learning

Within-task Training

Train cat dog → f → f^{*}

Hand-crafted

Meta Learning

Training Tasks

Task 1: Train apple orange Test apple orange

Task 2: Train bike car Test bike car

Learning Algorithm

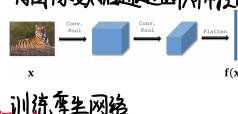
Across-task Training

李生网络

提取训练数据



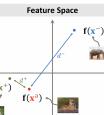
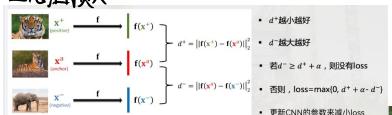
将图像数据通过卷积神经网络提取特征转化为向量



训练李生网络



三元组损失



从 dist(query, shot) 最高者分类

小样本学习流程

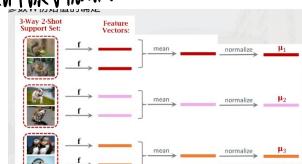
▶ 在大规模数据集上预训练一个李生网络

▶ 给定一个k-way, n-shot的支持集

▶ 利用支持集数据进行微调

▶ 给定查询样本，预测其属于的分类

预训练与微调



不进行微调对query样本做预测

Normalized mean feature vectors extracted from the support set.

余弦相似度：若 $\|x\|_2 = \|y\|_2 = 1$ ，则相似度为 $\text{sim}(x, y) = \frac{x^T y}{\|x\|_2 \|y\|_2}$

$\Phi = [\phi_1, \phi_2, \dots, \phi_L] \in \mathbb{R}^L$

$p = \text{normalize}([\phi_1^T, \phi_2^T, \dots, \phi_L^T]) \in \mathbb{R}^L$

$p_i \in \text{Softmax}(\Phi)$

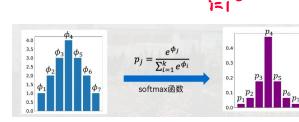
$p_i \geq 0 \text{ for } i=1, \dots, L$

$p_1 + p_2 + \dots + p_L = 1$

$$P = \text{softmax}(\Phi)$$

$$P_j = \frac{e^{\phi_j}}{\sum_{i=1}^L e^{\phi_i}}$$

$$\text{softmax} \text{ 函数}$$



初始化 $W = M$ 和 $b = 0$

$\min \sum_i \text{CrossEntropy}(y_i, p_i) + \text{Regularization}$.

Sum over all the samples in the support set.