

实验三 与无监督学习相关论文的学习笔记

MIL2019 袁富

1. 论文简介

在本论文中，作者试图使用简单的学习算法，通过仔细的选择网络参数以寻求更高的性能。结果表明，无论选择何种无监督学习算法，如降维（白化），选择大量特征和选择小步长等都可以使性能得到一定程度的提高。而论文在数据集上的实践表明，当我们将这些参数推到极限时，可以实现比很多最先进和复杂算法更好的性能，而在几种无监督算法之中，K-均值取得了最佳的性能。鉴于此，我使用了K-均值对论文中提到的特征提取方法在MNIST上进行实践。

2. 无监督特征学习框架

论文在这个部分，描述了一个通用的特征学习框架。并且专注于将这些算法应用于图片的特征提取。在一般情况下，我们提取特征的步骤如下；

1. 在没有标签的图像上随机提取一些小的图形块作为特征
2. 对上面提取到的图形块进行自己所需要的预处理
3. 利用一个无监督学习方法去学得关于这些图形块特征的一个映射

在学习到特征映射后，利用一个有标签的训练集，我们就可以实现特征提取和图片分类了，步骤如下：

1. 在输入图像上等间距的提取特征，并可以把图像大体上覆盖住
2. 对特征进行池化处理以减少特征数量
3. 训练一个线性分类器来预测给定特征向量的标签

2.1 预处理

通常作法是让每个图像分块减去平均值并除以其元素的标准差来标准化。在MNIST数据集上，通常情况调用函数得到的数据集是已经被标准化后的，各元素值处于0到1之间。鉴于MNIST数据集表示图像是黑底白字的，所以在图像取块前对数值进行0-1二值化。在实际调试过程中，发现降维对于该图像数据集的分类有副作用，所以弃用降维操作。

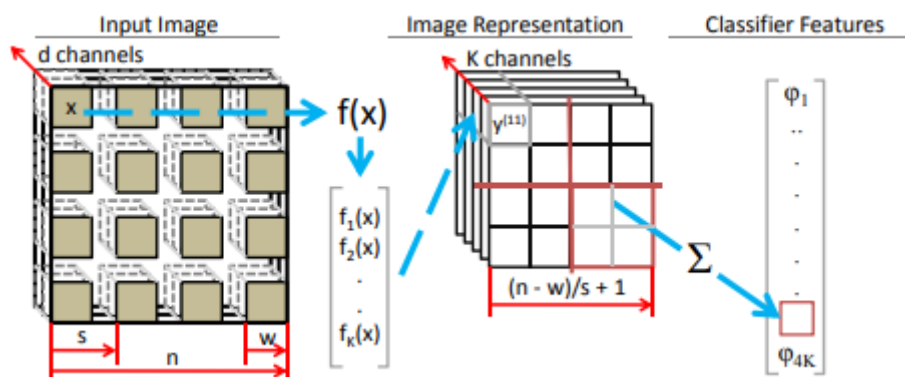
2.2 无监督学习具体过程

1. 对每个训练图像进行随机取块。取块完成后要保证所以子块可以凑成一个正方形（假设边长为a），这就要求我们在取块时需要结合图像大小进行综合考量。

2. 对所有训练图像的子块进行K-均值聚类时，选择K-means(hard)方式，如下公式所示，得到k个聚类中心，而具体的中心个数需要不断调整观察结果，选取最优的值。

$$f_k(x) = \begin{cases} 1 & \text{if } k = \arg \min_j \|c^{(j)} - x\|_2^2 \\ 0 & \text{otherwise.} \end{cases}$$

3. 对训练图像重新进行子块提取，子块大小与随机提取时的大小保持一致，且这次提取时每张图像的子块提取位置和个数要相互一致，但具体提取个数可以与随机取块时有变化。
4. 利用上面得到的k-均值模型，对第一张图像的所有子块进行分类，得到每个子块的特征向量之后，将所有特征向量组合得到一个(a, a, k)的三维数组。如下图所示，沿着K channels 方向将三维数组切分为k层，并对每层进行池化操作，每层最终只有四个值；得到(4, 4, k)的三维数组，再将其展开为(4k, 1)的一维向量，这便是第一张图像的特征向量。对所有输入图片进行上述操作，并组合得到训练集的特征矩阵。对测试集进行相同操作得到测试集的特征矩阵。



5. 以训练集的特征矩阵和标签作为输入，使用SVM进行训练得到分类模型；再以测试集特征矩阵为输入，利用训练得到的模型进行分类预测，输出正确率。

3. 由论文和高飞老师所讲词向量联想到的关于图像的特征向量的一种生成方法

关于数据集的特征向量生成，把2.2中第4步的内容稍作修改。直观上说就是将(4, 4, k)的数组分成k层，每层16个值相加的和作为新的值，最终得到(1, 4k)的特征向量，运用这个方法在MNIST数据集上训练样本为15000的情况下得到了0.85的预测准确率，可以推测继续追加训练样本时准确率有望达到0.9之上。或者换个思路也可以这么说，首先定义一个长度为k的行向量，用A表示；假设第2步中生成了100个聚类中心，按照生成的用来表示聚类中心的矩阵中各中心向量所在行的下标的顺序，对100个中心向量进行编号。对第一张图中按照固定位置截取到的第一个子块，我们得到表示该子块的一维行向量，将这个行向量扔进我们得到的K-均值聚类模型中，得到它属于的中心的编号，假定为j，则A[j]值加1；对第一张图的所有子块进行同样操作，得到最终的关于第一张图片的向量A。下图表示论文方法与修改后方法，各自得到的关于每张图片的特征向量的样子。

$\begin{bmatrix} 0 & 1 & 1 & 0 & \dots & 1 & 0 & 0 & 0 \end{bmatrix}$ 向量长度为4k，非0元素只有1，且所有1相加为16，表示每张图固定截取到的16个子块（某下标处的非0值表示这16个子块中有几个属于下标所表示类别）。

$\begin{bmatrix} 0 & 1 & 2 & 0 & \dots & 1 & 4 & 0 & 0 \end{bmatrix}$ 向量长度为k，非0元可以为1或其它，但所有元素的和仍为16，表示每张图固定截取到的16个子块。

