

电子科技大学

实验报告

学生姓名：蔡与望 学号：2020010801024 指导教师：徐行

一、实验项目名称：场景识别

二、实验原理：

2.1 Tiny Image 特征提取

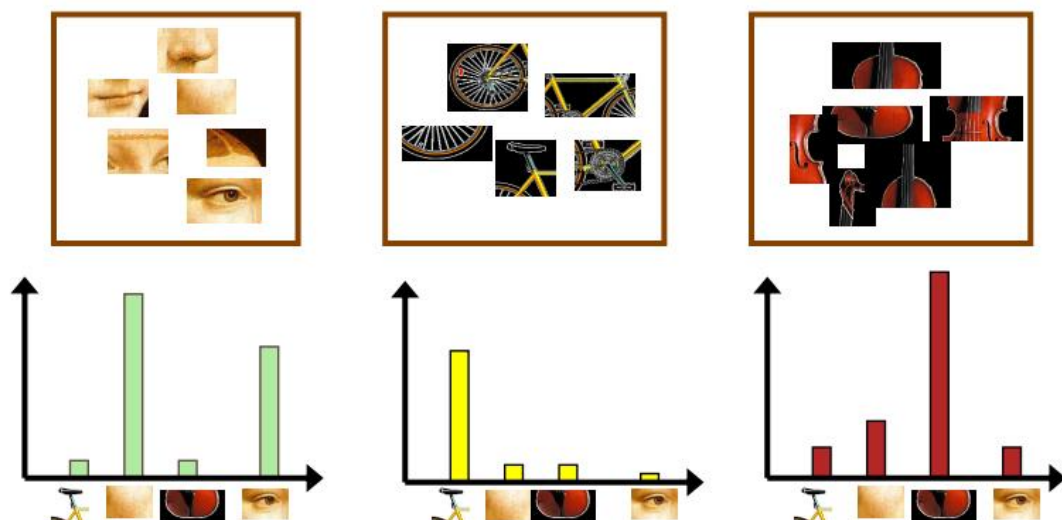
Tiny Image 的基本思路很简单，就是把原图像缩小到一个很小的尺寸，例如 16x16，然后把扁平化的小图像作为它的特征向量。它不需要训练，而是纯粹把图像本身当成了它自己的特征。但也容易想到，这样的提取效果也相对较差。

在缩小尺寸时，如果我们只是修改长宽，那就会造成混叠重影。所以我们在缩小前，需要用高斯模糊来平滑图片。

在缩小尺寸后，标准化图像可以略微提升一些性能。

2.2 词袋特征提取

词袋特征提取的基本思路是，先训练出一个词汇库，即所有场景的各种特征的集合。然后把训练集和测试集的所有图像都转换成它对应的词袋，即这幅图像的特征的直方图。



本实验采用 HOG 算法来提取图片的特征向量。之后，虽然我们可以直接把这些特征向量当成我们的词汇库，但这样形成的词汇库过于庞大，会使得后续的分类性能、效果下降。

所以，我们可以使用 **K-Means** 给这些特征向量聚类，把聚类的中心当作词汇库。这样，就相当于在尽量不影响词汇的特征的情况下，减少了词汇的数量。

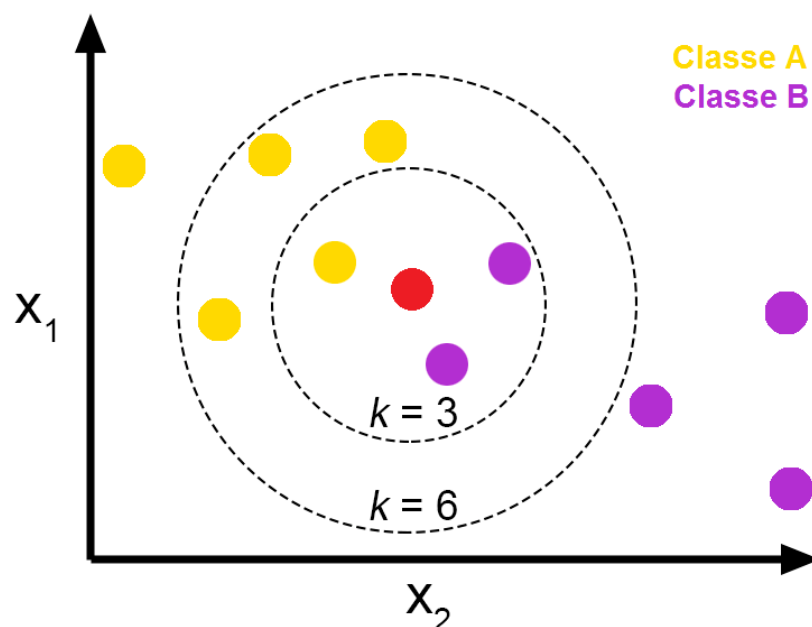
同时，我们没有必要每次都构建一次词汇库。我们可以把词汇库存储进文件，之后每次运行只要从文件中读取现成的词汇库即可。

至于把图像转换成词袋，实际上就是：用训练词汇库时的 **HOG** 参数提取图像的特征向量，然后看这个特征向量和哪些词汇离的最近。最近的词汇就会被加入这幅图像的词袋，以直方图的形式存储。

2.3 KNN

KNN 是最简单的几种分类算法之一。它的基本思路是，找到离自己最近的 k 个物体，这些物体里哪个标签最多，那自己也应该属于哪个标签。

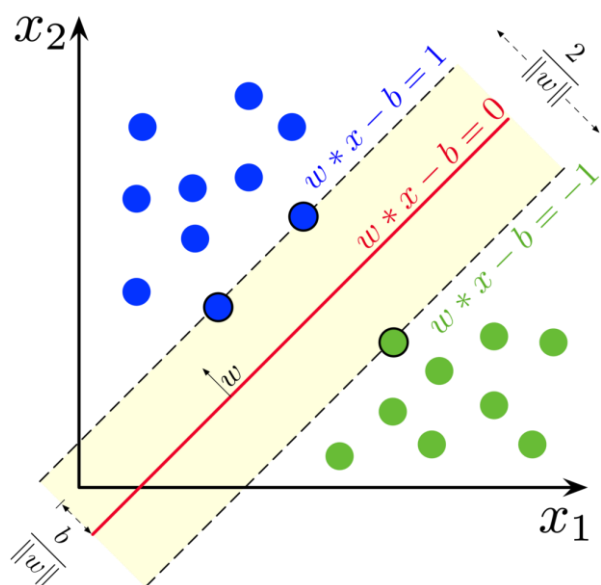
它的优点是不需要训练，可解释性强，而且大多数时候能取得不错的效果。缺点是难以把握 k 的设置，每次测试的计算量大，难以抵抗样本不平衡的情况。



2.4 SVM

相比起 **KNN**，**SVM** 是更可靠的一种分类方法。它的基本思路是，找到一个超平面，尽可能地把属于不同标签的物体完全分开。在预测时，也只要看预测物体落在哪个区域内即可。

它的优点是泛化能力强，特别适合高维问题。缺点是样本多时训练昂贵，对缺失数据敏感。



三、实验目的：

1. 掌握 Tiny Image 和词袋两种图像特征提取的原理和方法。
2. 掌握 KNN 和 SVM 两种分类的原理和方法。

四、实验内容：

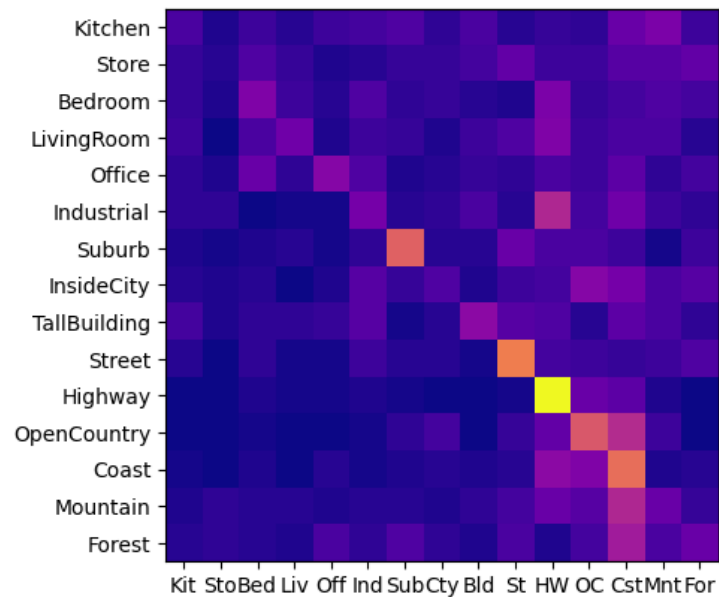
分别利用 Tiny+KNN 和 Bags of Words (SIFT)+SVM 实现对场景的分类，并且比较不同特征和不同分类方法对最终精度的影响。

五、实验步骤：

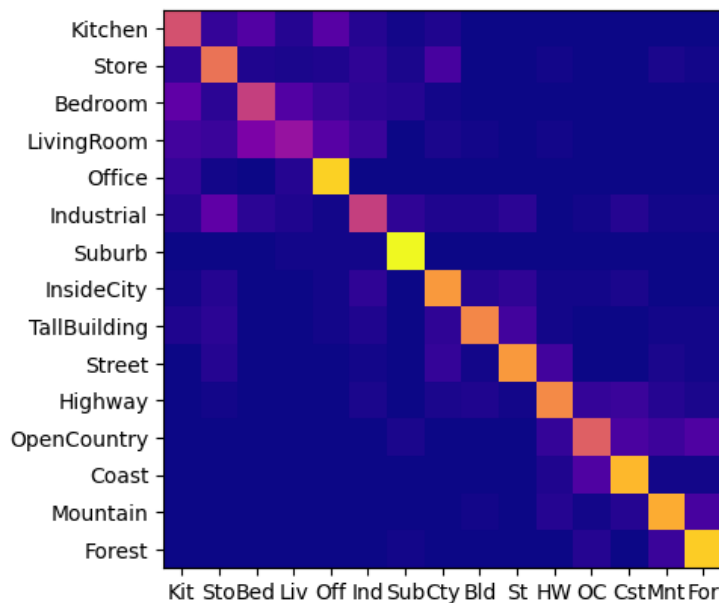
1. 编写函数，使用 Tiny Image 提取图像特征。
2. 编写函数，构建词汇库。
3. 编写函数，使用词袋提取图像特征。
4. 编写函数，使用 KNN 预测图像标签。
5. 编写函数，使用 SVM 预测图像标签。
6. 分别使用 Tiny+KNN 和词袋+SVM 在测试集上测试，研究准确率的区别。

六、实验数据及结果分析：

首先使用 Tiny+KNN 的组合。Tiny Image 将图像缩小成 16x16，KNN 寻找最近的 3 个词汇。此时的混淆矩阵如下图所示，准确率 23.867%。



然后使用词袋+SVM 的组合。在提取特征向量时，把图片放大为 256x256，每个 block 有 2x2 个 cell，每个 cell 有 4x4 个像素。300 次迭代后，聚类出 200 个词汇。此时的混淆矩阵如下图所示，准确率 67.667%。



七、实验结论：

很明显地看到，词袋+SVM 组合的准确率比 Tiny+KNN 组合的准确率提升了不少。这第一是由于，Tiny Image 并不能很好地描述图像的特征，比不上专业的 HOG 算法；第二是由于 SVM 特别擅长解决高维问题，而且准确率一般来说都要比 KNN 高。

实验最后的准确率为 67.667%。按照示例代码给的标准，这说明实验代码正确地实现了词袋和 SVM 算法，并且已经将参数调优。实验成功。

八、总结及心得体会：

通过本次实验，我深刻理解了 Tiny Image 和词袋两种特征提取、KNN 和 SVM 两种分类的原理和方法。同时，我将他们的组合进行对比，直观地感受到了不同算法，对于场景识别问题，在性能和效果上的优劣，理解了每种算法各自的优缺点。

九、对本实验过程及方法的改进建议：

无。

报告评分：

指导教师签字：