# Project图像识别-3.19-笔记

课程学习地址：<https://www.bilibili.com/video/av76633310>，cs231n计算机视觉经典课程的P1,P2,P4,P5

P1（计算机视觉概述）：视觉数据量爆炸式地增长，大概每秒钟就有五小时的视频上传到Youtube，我们需要电脑帮助我们处理这些数据。但是，视觉数据对于机器来说很难理解。

P2（历史背景）：视觉处理是始于视觉世界的简单结构。目标识别可以先目标分割-把一张图片中的像素点归类到有意义的区域。Adaboost算法进行面部检测。同一类物体的不同个体很难识别，影响因素：相机的角度，遮挡，视角，光线，以及自身的变化。三个问题：目标识别，机器学习。卷形神经网络模型（CNN）让图像识别有很大进步。

P4（图像分类-数据驱动方法）：

1. 作业是用Python和Numpy写的，快速入门教程地址：<http://cs231n.github.io/python-numpy-tutorial/>
2. Numpy可以非常高效地进行向量化操作
3. 机器看到的一张图片是一大堆数字，每一个像素由三个数字表示，给出像素红、绿、蓝三个值。这个巨大的数字矩阵很轻易地完全变化，很难找到共性。鲁棒算法-解决。边缘对于视觉识别很重要，可适用某一类物体，但不能适用许多类物体的识别。数据驱动方法：不设定具体的分类规则来识别，从网上下载许多不同类的图片，一旦拿到数据集。就可以训练机器来分类这些图片，机器会以某种方式总结然后生成一个模型，总结出识别不同对象的核心知识要素，最后就可以用这个模型来识别新的照片并分类。
4. 两个函数：训练函数，接入图片和标签，输入一个模型。预测函数，输入图片和模型，输出标签。
5. 数据驱动类算法是比深度学习更广义的一种理念，也称最近邻算法。训练的速度是O（1），预测的速度是O（n）。会耗费很多时间。
6. 如何比较两个图片：可以使用L1距离（像素相减取绝对值），L2距离（取平方和的平方根）。
7. K的取值一般大于1

P5（图像分类-K最近邻算法）：

1. L1距离取决于选择的坐标系，坐标轴改变，L1也会改变，但是L2距离与坐标轴无关。
2. K最近邻算法不仅可以用于图像分类，还可以用于文字分类，需要改变L1的值。
3. 在使用最近邻算法时，需要选择K值和距离度量（L1\L2），通过尝试选出最适合的超参数。
4. 不要在训练集中尝试不同的K然后选择最好的K（K=1时总是拟合最好，但是在测试集中就不太好了），适当让K大一点，这样虽然在训练集中会出错，但是在测试集中会表现更好
5. 不要把总的分成训练集和测试集，尝试不同的K，选择在测试集中表现最好的K，这个K可能只适用于这个测试集，无法适用于全新的、没见过的集合。
6. 正确的做法是将数据分成训练集、验证集、测试集。在训练集上用不同的K来训练算法，在验证集上进行评估，然后选择最好的一组K，然后再把在测试集中表现最好的分类器拿出来。
7. 必须分割验证集和测试集。