

开题报告

开题报告

本报告为清华大学 2019–2020 春季学期电子工程系媒体与认知课程大作业图像实验开题报告，分别介绍各任务调研结果与目前思路。

任务一 传统机器学习 & 分类任务

相关方法

HOG：方向梯度直方图，通过计算区域内像素点数据的横纵梯度并进行按角度分类，得到区块的梯度直方图，用于刻画图像的边缘信息。

- Histograms of Oriented Gradients for Human Detection, Navneet Dalal, Bill Triggs
- scikit-image https://scikit-image.org/docs/stable/auto_examples/features_detection/plot_hog.html?highlight=hog

SVM：支持向量机，通过构建高维超平面对数据进行划分，从而实现回归。可以使用多个SVM结合实现数据多类的分类。优点在于解是全局最优，理论基础较为完善；缺点在于二次规划问题的求解计算量偏大，需通过多个SVM组合实现多分类问题。

- scikit-learn <https://scikit-learn.org/stable/modules/svm.html>
- opencv https://docs.opencv.org/trunk/d4/db1/tutorial_py_svm_basics.html

随机森林：构建大量决策树，对数据进行多次二分类以分类。优点在于随机性防止过拟合、判断特征的重要程度、训练快；缺点则是对于特殊问题（噪声大的样本或特征取值划分较多的数据）效果较差。

- scikit-learn <https://scikit-learn.org/stable/modules/ensemble.html#forests-of-randomized-trees>

感知机：使用多重感知机对分类函数进行学习。相比SVM和随机森林，更接近深度学习，与前两者相比似乎并无显著优势。

Baseline

目前尝试了单纯使用SVM（默认参数）对图片进行分类，得到了89%的validation准确率，可见对于该数据集的分类还是比较简单的，使用传统机器学习方法就可以获得比较高的准确率，

预计使用深度神经网络会获得更加优秀的效果

思路

相比于多重感知机，HOG+SVM或随机森林的方法所需的计算量更小，训练时间更短，效果跟数据集和超参数设定有关。其中HOG+SVM的方法有较为严格的数学推导，保证了解的存在性；随机森林可能具有更短的训练时间，但训练效果较依赖超参数的设定。

我们的思路是使用openCV或scikit-image这类库进行图片的预处理并生成HOG，然后尝试使用基于scikit-learn的SVM和随机森林进行训练。根据需求可能会加入感知机等方法。

任务二 深度学习 & 分类任务

相关方法

经典文献：

- ResNet <https://arxiv.org/pdf/1512.03385.pdf>
- Xception <https://arxiv.org/pdf/1610.02357.pdf>
- MobileNet <https://arxiv.org/pdf/1801.04381.pdf>

思路

CNN在图像分类这个领域相比传统机器学习方法有很强的优势，其中比较著名的网络模型有ResNet和Xception，且都在经典的图像数据集上取得了不错的效果。从这些经典模型开始着手应是不错的选择。虽然这两个知名网络的效果很好，但是数据量太过庞大，也许借鉴MobileNet这类轻量化网络的经验可以在不降低效果的基础上减轻计算负担。

我们决定先直接使用著名的Xception网络进行训练，再根据结果进行适当地改进。

任务三 小样本分类

相关方法

元学习 Meta-Learning：元学习提出了“学习如何学习”这一概念，即训练网络去学习如何提取特征，从而实现在小样本上进行训练的能力。某种程度上，元学习和迁移学习的思想是一致的。

- Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks <https://arxiv.org/pdf/1703.03400.pdf>
- Prototypical Networks for Few-shot Learning <https://arxiv.org/pdf/1703.05175.pdf>
- Matching Networks for One Shot Learning <https://arxiv.org/pdf/1606.04080.pdf>
- One-shot Learning with Memory-Augmented Neural Networks <https://arxiv.org/pdf/1605.06065.pdf>

除此之外，还有另一种实现one-shot learning的思路——孪生网络Siamese Neural Networks

- Siamese Neural Networks for One-shot Image Recognition <https://www.cs.cmu.edu/~rsalakhu/papers/oneshot1.pdf>

Baseline

目前计划考虑采用1近邻方法。

思路

对于我们来说，这个领域相对陌生，还需要进一步进行文献调研。计划重要文献MAML的方法开始进行深入学习，比较Matching Networks所做工作中各种方式的效果。此外，Siamese Network的想法也不失为一种选择。

任务四 检测并分类

相关方法

- R-CNN 分区域划分和分类两个步骤，作为开山之作有明显的计算过于复杂的缺点。有Fast R-CNN和Faster R-CNN等改进工作。
- Evaluation of Deep Neural Networks for traffic sign detection systems
- HOG Histograms of Oriented Gradients for Human Detection, Navneet Dalal, Bill Triggs
- 随机森林 Traffic sign detection and recognition based on random forests
- yolo <https://arxiv.org/pdf/1506.02640.pdf>
- cornernet <https://arxiv.org/pdf/1808.01244.pdf> <https://arxiv.org/pdf/1904.08900.pdf>

思路

检测并分类主要有两种做法：检测和分类分两步进行，与一次进行检测与分类。本次大作业要求使用前面实现的分类器完成第二步，故应仿照已有的方法实现检测功能。