**结合HOG和SVM的目标检测方法**

目录

[**1. 整体实现方案** 1](#_Toc29477382)

[**2. 数据集来源** 2](#_Toc29477383)

[**3. 效果展示** 3](#_Toc29477384)

[**4. 代码展示** 5](#_Toc29477385)

[**5. 实现细节** 6](#_Toc29477386)

**1. 整体实现方案**



方案的大致步骤如下：

1）利用分类样本数据集对SVM进行预训练

2）载入预训练模型，利用检测样本数据集和对应的GroundTruth对SVM进行再训练

3）测试阶段的如上图所示，分别是，输入图像、提取预选边框、提取HOG特征、SVM判别类别和置信度、非极大值抑制，输出检测结果

**2. 数据集来源**

数据集来源于LSI远红外行人数据集

<http://portal.uc3m.es/portal/page/portal/dpto_ing_sistemas_automatica/investigacion/IntelligentSystemsLab/research/InfraredDataset>

该数据库包含从在室外城市场景中驾驶的车辆收集的FIR图像。图像是使用Indigo Omega成像仪采集的，分辨率为164x129像素，灰度级为14位，焦距为318像素。摄像机安装在车辆的外部，以避免挡风玻璃的红外过滤。

手动注释记录的图像，其中每个行人都被标记为边界框。为了防止边界伪影引入偏差，它们的高度随后会按比例增加5%。行人以直立的姿势出现。

数据集分为两部分：

（i）分类数据集：正片和随机采样的负片，高度与宽度的比例固定为（1/2），并重新缩放为64x32像素

（ii）检测数据集：原始正片和负片具有注释。

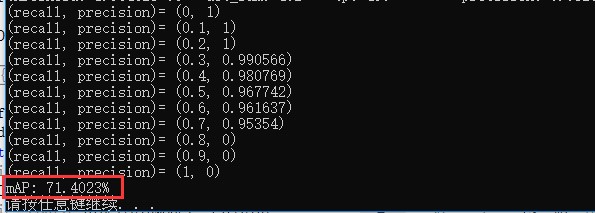
  注意：仅注释高度超过10像素的直立人。注解可能不是100%正确；在某些特殊情况下，部分行人可能会落在边界框之外。未遮挡部分被遮挡的行人或未完全包含在图像内的行人。图像是按顺序获取的，因此，最终两个连续的图像实际上可能是相同的。

|  |  |
| --- | --- |
| (a) 正样本 | (b) 负样本 |
| 分类数据集 | |

|  |  |
| --- | --- |
| (a) 检测数据集中的训练样本 | (b) 检测数据集中的测试样本 |
| 检测数据集 | |

**3. 效果展示**

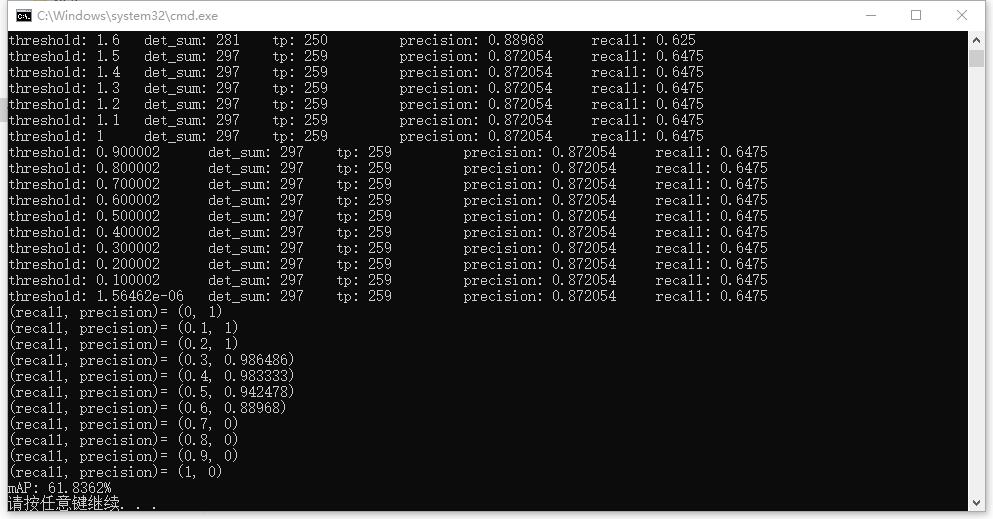
训练的结果，训练集上测得mAP为71.4%



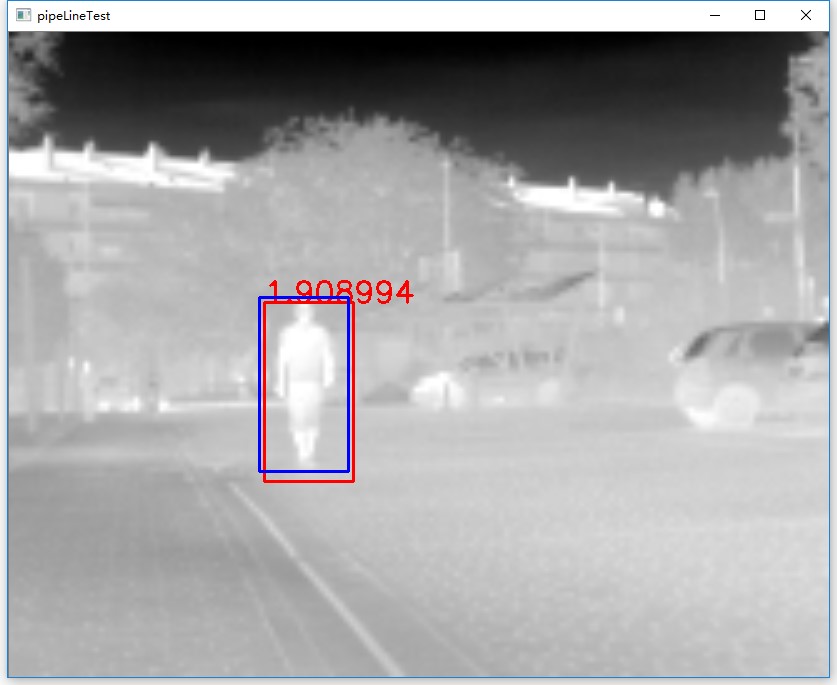
训练集上的检测结果



测试的结果，测试集上测得的mAP为61.8%



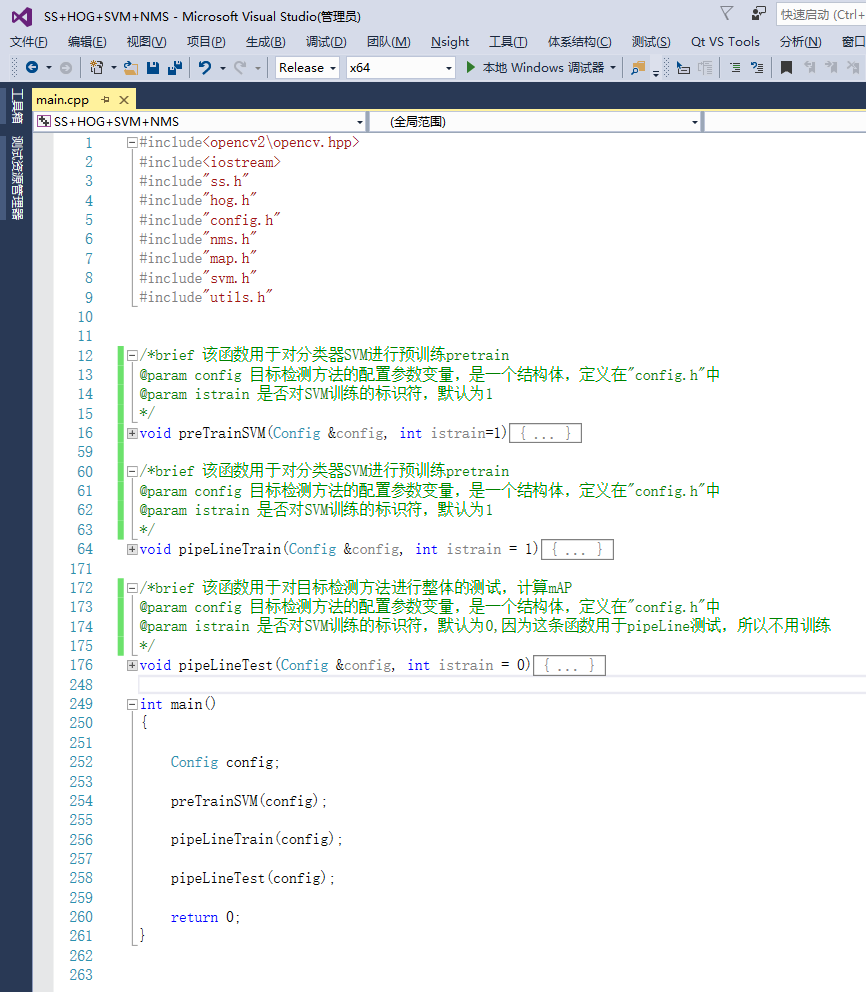
测试集上的检测结果

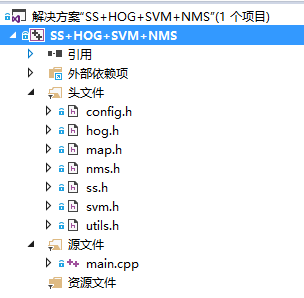


**4. 代码展示**

详细代码：<https://github.com/LeonJinC/SS-HOG-SVM-NMS>

源码包含8个核心文件，分别为ss.h、 nms.h、 hog.h、 svm.h、 utils.h、 map.h、config.h和main.cpp





**5. 实现细节**

5.1训练阶段的SVM训练

所有的预选边框与对应的GroundTruth计算iou值，iou≥0.7为正样本，iou<0.1为负样本，从而构建正负样本数据集用于训练SVM分类器。

5.2测试阶段mAP计算

iou阈值≥0.5为TP，iou<0.5为FP



计算mAP的具体步骤如下，

1）遍历confidence threshold（比如，0.0~1.0）可以得到n组（precision，recall）数据点

2）选取当Recall >= 0, 0.1, 0.2, ..., 1共11个点时的Precision最大值

3）AP就是这11个Precision的平均值。

4）mAP就是对每一个类别都计算出AP，然后再计算所有类别AP平均值

mAP计算方法参考了知乎上的一个回答：<https://www.zhihu.com/question/53405779/answer/419532990?utm_source=wechat_session&utm_medium=social&utm_oi=600695103581458432>