基于静态图像的教室人数识别算法设计

201800820139 王福临

**摘要：**为了解决自习室静态背景下的人数识别问题，本文使用Python的opencv包，对固定场景的图像进行目标区域选取，再对这些大小归一化后的目标区域进行预处理和HOG特征的提取，来训练各种常用分类器。在选定效果最好的SVM算法后，检验得知该算法能够较好地实现人数识别功能。

**关键词：人数识别 HOG特征 SVM分类器**

**一、引言**

为了进一步提高考试月学校自习教室的使用率，让同学们避开人数较多的自习室，本项目计划从自习室的监控视频中提取的图像入手，对图像进行人数识别，从而节省大学生寻找自习教室的时间，并提高自习教室的利用率。

**二、方案与设计思路**

**1. 算法方案**

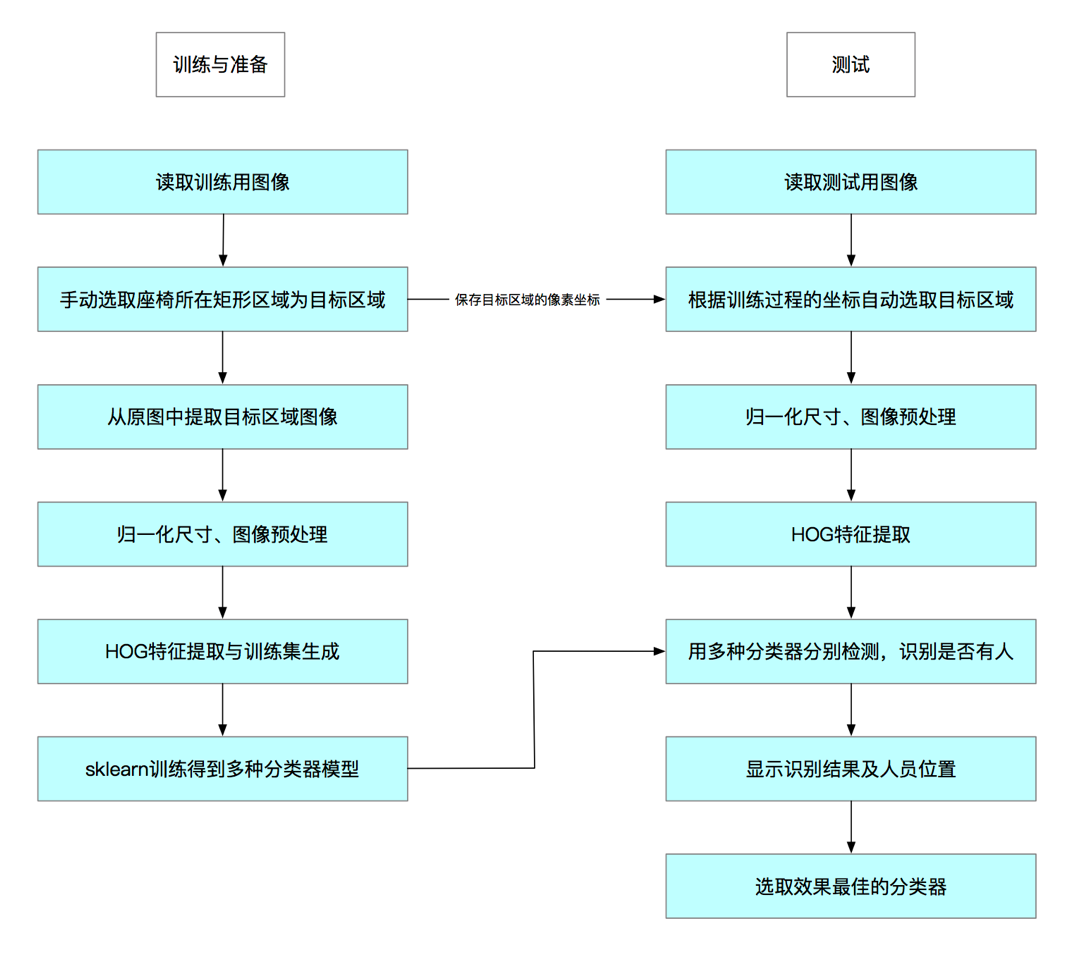
在自习室场景下，监控设备获取的图像往往较为模糊，每个人的特征并不容易提取，因此通过个体的特征对特定目标进行跟踪与检测的难度较大。由于监控设备的位置一般是固定的，监控视频的图像中每张座椅的像素坐标也是固定的，因此，对于教室这一特定场景，可以通过将识别区域固定在每个座位上，对这些区域做相同预处理并统一大小后进行特征提取，利用作为训练集的图片结合机器学习分类算法生成分类器；最后将训练好的分类器通过对座椅区域“是否有人”进行分类的方式，实现图像中人数的识别和人员所在位置的确定，从而达到既定目的。算法流程图及目标区域提取效果图如下所示。

图1 人数识别算法流程图

图2 目标区域提取效果图



图2 目标区域选取的效果图

在上述算法流程中，由于监控设备放置的位置和角度不同，各个目标区域的大小也不同。为了使各个区域提取出的特征数相同，将每个区域的大小归一化，从而生成可用于常见分类器的训练集。

该算法在训练时同时选取了随机森林、SVM、多层感知机、朴素贝叶斯共4种较为常见的分类算法，使用Python的sklearn包进行训练，并在测试时对每种算法的准确率进行对比，最终选定效果最好的SVM作为分类器。

1. **特征提取**

为完成分类的目的，选取每个目标区域的HOG（方向梯度直方图）特征作为分类器判断该区域是否有人的依据。相比其他常见的图像特征特征而言，HOG特征对图像几何和光学的形变都能保持很好的不变性，且它没有旋转和尺度不变性，因此计算量较小。

HOG特征提取主要分为以下几步：

（1）图像预处理：

图像预处理分为图像灰度化和Gamma校正两步。

图3 灰度化后的图像

对于一个三维图像，HOG特征提取的过程中颜色信息的作用通常较小，因此将图像先转化为灰度图。灰度化后的图像如下图所示：

Gamma校正对输入图像进行颜色空间的标准化，将像素值转化为0-1之间的实数，从而调节图像的对比度，降低图像局部的阴影和光度变化对实验的影响，也可抑制噪声的干扰。Gamma校正的公式如下所示：（一般取）

Gamma校正后的图像效果如下图：

图4 Gamma校正后的图像

（2）计算每一个像素点的梯度值和方向：

将图像中每一个像素点所在的3\*3像素方格与Sobel算子做卷积运算，得到某点处梯度值和方向的计算公式如下：

其中表示像素点处的像素值，和表示处的水平、垂直梯度,表示梯度的方向与x轴夹角。

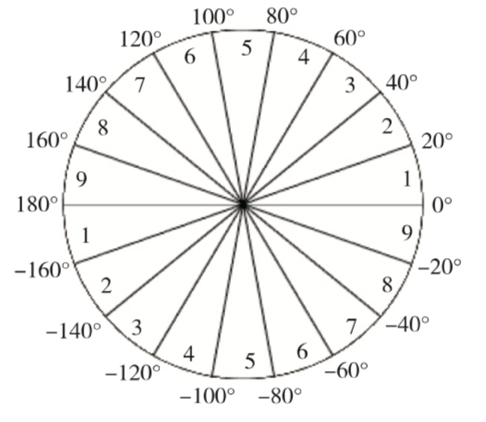
（3）每个cell的HOG特征的生成：将图像划分成许多cells(细胞单元)，每个像素作为一个cell，且相邻的cell之间没有重合部分。在每个cell中，将所有梯度方向放入划分的9个部分（这9个部分由0-360度均分得到），作为直方图的横轴，且对于每个处在角度范围内的点，把它们的梯度值累加作为直方图的纵轴。这样生成的直方图对应着一个9维向量。

图5 9个bin的方向梯度的划分

（4）每个cell合成大的可重叠的block：

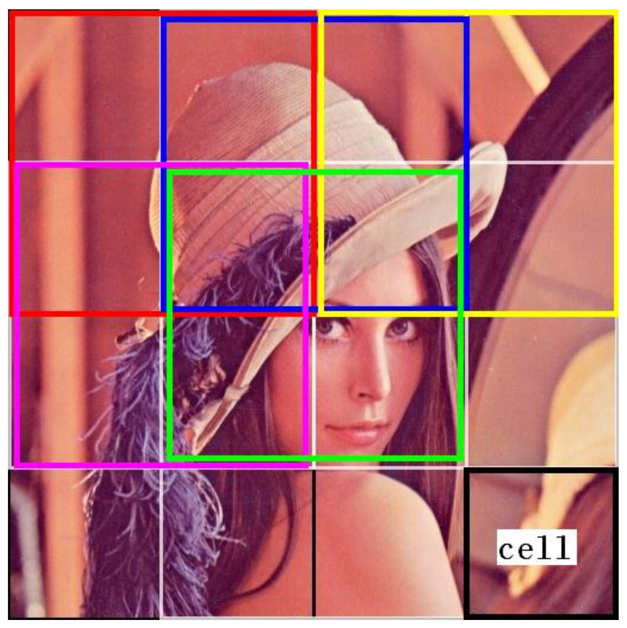
在HOG特征提取前图像的大小归一化时，为了使图像能够被完全划分为的cell（也就是划分后没有多余像素），预先将图像归一化为像素大小，从而图像可以被划分为个的cell。将左上角方格的cell构成一个block，再将block每次向右、向下平移1个cell的步长，直到覆盖整张图像，如此生成了个block。注意到相邻的block之间是有重叠的，这样通过更全面地分析各个相邻像素之间的共同特征，提高了算法的准确率。

图6 block与cell的划分示例（其中每个彩色框为一个block，每个小方格为一个cell）

（5）生成特征向量：

将一个block内所有的cell的特征向量合并为一个新的向量，再将每个block对应的特征向量归一化，得到该block的HOG特征。最后将每张图片中所有的block的特征向量合并为一个新的向量，得到图片的HOG特征向量。在本文中，特征向量的维度为。其中为block数量，为每个block中cell的数量，为每个cell的特征向量的维度，也就是将360度划分成几个部分的数目。

每张训练集、测试集图片经过上述算法，都可以生成一个864维的特征向量用于训练和预测。

1. **分类器的选取**

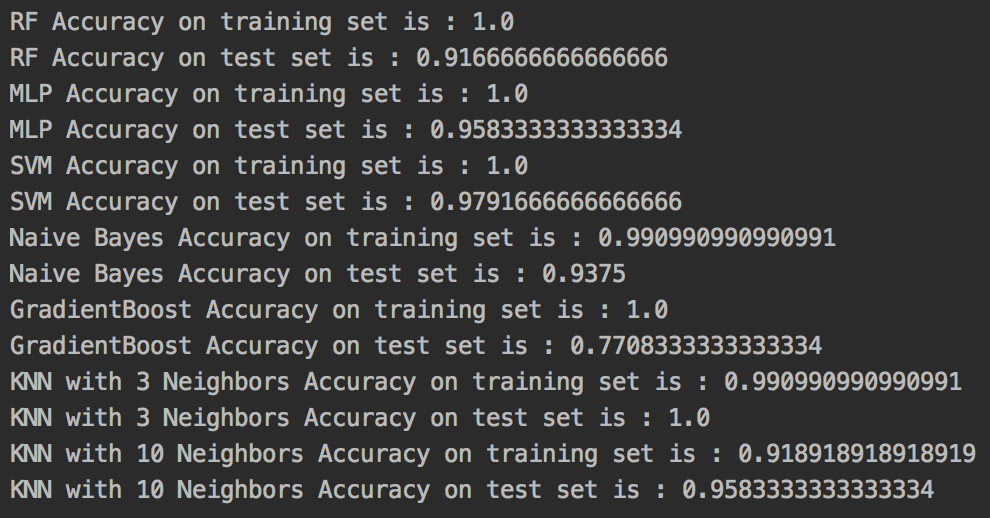
在已有图片中分别截取有人、无人状态下的座位图片作为训练集，将训练集随机打乱后取30%的数据作为训练集，用随机森林、多层感知机、SVM、朴素贝叶斯、梯度提升、KNN这几种常用算法训练的结果如下图。在选取训练算法的过程中，KNN算法在邻居数小于4时的表现最好，但是当KNN的邻居数过少时，算法的鲁棒性会严重降低，因此最终选取SVM作为分类器。

图7 各常见算法的准确率测试结果

**三、实验与结果分析**

**1. 训练集的生成**

首先在教室对应区域的Mask中，选取各80张有人、无人区域的图像作为训练集的正负样本，这些样本均已经被归一化为大小。部分样本如下图所示：

****

图8.1 负样本图片示例

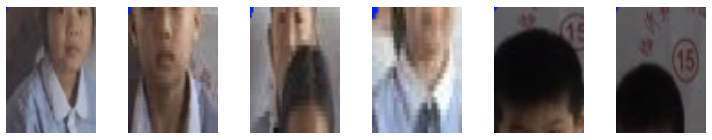
****

图8.2 正样本图片示例

**2. 测试集的生成**

在教室的监控视频提取的图像中，由于监控设备的位置固定，不同时间点相同的像素坐标仍然代表相同的目标区域，因此可以直接通过保存训练集目标区域的像素点来进行测试集目标区域的选择。（注：在本次实验中，由于条件限制，无法获取教室的监控视频，因此用拍照的方式获取了训练集，这可能会导致不同图片之间的目标区域像素坐标不同。因此在本文的测试中，仍然选取手动标定测试集目标区域来进行识别的方式。）

1. **测试结果展示**

人数识别的测试结果如下图所示：（识别结果为有人则将目标区域用红框标定，识别结果为无人则将目标区域用黄框标定）



图9 人数识别结果

**四、结论**

本文设计了一种针对自习室环境的人数识别算法，通过对预先划定的目标区域进行图像预处理和HOG特征提取，生成可用于常见分类器的数据集，从而训练得到合适的SVM分类器。该方法生成的分类器在遮挡等较复杂的现实情形下，也有着较好的效果和准确率，在静态图像人数识别的应用场景下有着较强的性能。

**五、参考文献**

【1】周正宽,田小平,贺玲玲,任继伟.基于贝叶斯分类的人数识别算法研究[J].北京石油化工学院学报,2019,27(04):47-53.

【2】李志刚. 基于图像的公共汽车人数自动统计技术研究[D]. 硕士学位论文, 太原: 中北大学, 2008.

【3】应俊. 基于计算机视觉的电梯轿厢内人数统计研究[D]. 硕士学位论文, 杭州: 杭州电子科技大学, 2013.

【4】张海潮,马风玲,郑伯明,蔡占秀,邱召运.一套教室人数智能统计与显示系统的研制[J].中国教育技术装备,2016(24):40-42.

【5】Irida Shallari,Silvia Krug,Mattias O’Nils. Communication and computation inter-effects in people counting using intelligence partitioning[J]. Journal of Real-Time Image Processing,2020(prepublish).

【6】刘洋.基于开运算与区域数统计的教室人数估计[J].电子技术与软件工程,2020(07):152-154.