

CVPR 第五次作业-目标检测

强基数学

吴天阳 2204210460, 马煜璇 2204220461, 申宇环 2201422097,

陈开来 2205110920, 王瑞恒 2202113454

1 实验目的

1. 掌握 HOG 特征描述子和线性 SVM 分类算法.
2. 理解滑动窗口进行目标检测方法.
3. 掌握交并比 (IoU) 的原理与运用.
4. 理解召回率、检验精度的概念, 使用均值平均精度 (mAP) 评估目标检测器.

具体方法: 使用 HOG 特征和 SVM 设计单目标检测器.

2 实验原理

2.1 方向梯度直方图 (Hog)

2.1.1 图像预处理

对原图像进行高斯模糊, 降低噪声, 利用 Gamma 校正公式对原图像亮度进行降低, $f(x) = x^\gamma$, ($\gamma \geq 1$), 当 γ 越大时, 图像亮度越低.

2.1.2 梯度图计算

使用 Sobel 算子分别计算 x 和 y 方向的偏导, 两个方向上的偏导 Sobel 算子如下

$$W_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad W_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

然后计算幅值 (L2 范数) 和梯度方向:

$$g = \sqrt{g_x^2 + g_y^2}, \quad \theta = \left| \arctan \frac{g_y}{g_x} \right|$$

注: 此处的梯度方向 θ 是取过绝对值的结果, 所以角度范围为 $[0, \pi]$.

2.1.3 梯度直方图计算

由于梯度过于离散, 对一个较大区域中全部梯度方向进行统计, 首先假定 cell 大小为图像中 8×8 的一个区域, 将原图划分为多个不交的 8×8 的区域, 期望将 $[0, \pi]$ 平均划分为 9 个部分 (bins), 然后将该区域中的梯度均摊到这 9 个辐角区间中表示这个区域的辐角分布, 如下 10 个端点构成的 9 个连续区间

$$\pi : 0, 20^\circ, 40^\circ, 60^\circ, 80^\circ, 100^\circ, 120^\circ, 140^\circ, 160^\circ, 180^\circ$$

这也就是创建了大小为 9 的直方图数组, 记为 h_1, \dots, h_9 , 然后该 cell 包含的每个像素点处的幅值按该像素点处的方向大小平均分配到最近的区间上即可。

举个例子, 如果某像素点的幅值为 4 辐角为 65° , 那么最近的两个区间点为 $60^\circ, 80^\circ$ 对应数组为 h_4, h_5 , 并且通过距离确定到两点的权重分别为 $1/4, 3/4$, 于是更新直方图数组为 $h_4 \rightarrow h_4 + 1/4 \times 4, h_5 \rightarrow h_5 + 3/4 \times 4$ 。

2.1.4 Block 归一化

计算出每个 Cell 的梯度直方图之后, 再以 3×3 的 cell 作为一组, 称为 block. 由于每个 cell 包含 9 个值, 所以一个 block 具有 $3 \times 3 \times 9 = 81$ 个值, 然后在通过滑动步长为 1 的 block 窗口, 逐步输出每个位置的 block 值, 最后拉成一个行向量, 这就是 **Hog** 特征, 假设原图的大小为 40×100 , 则 cell 个数为 5×12 , 于是最终 Hog 特征维数为 $(5 - 3 + 1) \times (12 - 3 + 1) \times 3 \times 3 \times 9 = 2430$ 。

再分析下每个 block 具体做些什么: 一个 block 是由 $3 \times 3 = 9$ 个 cell 构成的, 每个 cell 有 9 个变量, 所以一共 block 存储的是 81 维向量, 对齐进行归一化处理, 就得到了该 block 所具有的信息。

3 实验步骤与结果分析

本次代码难度较大, 参考老师教程使用 GitHub 中项目完成. 特征提取及 SVM 模型训练: [GitHub-bikz05/object-detector](https://github.com/bikz05/object-detector), mAP 评价目标检测器: [Cartucho/mAP](https://github.com/Cartucho/mAP)

3.1 计算 Hog 特征

首先利用 `extract-features.py` 进行 Hog 特征提取, 使用 `skimage.feature.hog` 对目标进行提取, 正例图像处理的核心代码如下 (处理反例只需将文件路径修改即可):

```

1 for im_path in glob.glob(os.path.join(pos_im_path, "*")): # 读取图片路径
2     im = cv2.imread(im_path, cv2.IMREAD_GRAYSCALE) # 灰度图像读入
3     fd = feature.hog(im, orientations=9, # 将角度分为 9 个区间
4                     pixels_per_cell=(8, 8), # cell 大小
5                     cells_per_block=(3, 3)) # block 大小
6     fd_name = os.path.split(im_path)[1].split(".")[0] + ".feat"
7     fd_path = os.path.join(pos_feat_ph, fd_name) # 设置保存文件路径
8     joblib.dump(fd, fd_path) # 保存图像的 Hog 特征

```

图??是我们自己尝试 Hog 特征提取结果:

4 结论与讨论

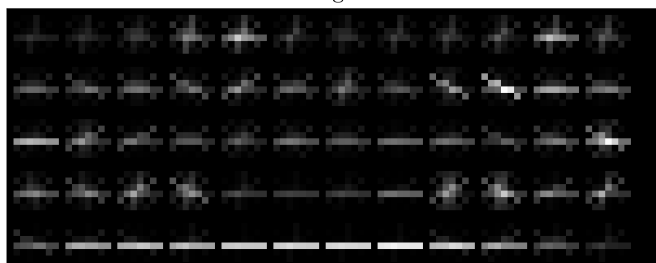
图像1



图像2



Hog1



Hog2

