

模式识别大作业报告

AlexDawn

2025 年 5 月 28 日

目录

1 引言	2
2 HOG 和 SVM 实验原理简述	2
3 单任务检测水平边界框 (single-task)	2
3.1 问题分析	2
3.1.1 样本集划分	2
3.1.2 标注框的尺寸不固定	3
3.1.3 预测时计算复杂度高	4
3.1.4 输出多个预测框	4
3.2 程序设计思想	4
4 多任务检测水平边界框 (multi-task)	5
4.1 问题分析	5
4.2 程序设计思想	5
5 深度学习车辆信息检测 (选做)	6
5.1 程序设计思想	6
5.1.1 基于目标检测框架的改进方法	6
5.1.2 基于图像分割思想的替代方案	6
6 实验及结果分析	6
6.1 模型评价标准	6
6.2 实验结果	7
6.2.1 单任务检测	7
6.2.2 多任务检测	8
6.2.3 深度学习方法检测(尚未完成)	10
7 总结	10

1 引言

目标检测是模式识别领域的一项基础任务，其主要目标是在输入图像中用标注框标注出待检测的目标。

本实验有三项任务：基于HOG特征提取与SVM分类的单任务检测、多任务检测，以及基于深度学习模型的检测。

2 HOG 和 SVM 实验原理简述

HOG 特征是一种用于目标检测的特征描述子。其主要思想是计算并统计图像局部区域中梯度方向的分布情况。具体来说，首先将图像划分为若干个相连的小单元格（cells），然后在每个单元格内计算像素点的梯度大小和方向。接着，将这些梯度方向量化为固定数量的区间，并构建直方图来表示每个单元格的方向梯度分布。相邻的多个单元格构成一个区块（blocks），并对区块内的直方图进行归一化处理以增强对光照变化的鲁棒性。

SVM 是一种监督学习模型，常用于分类与回归分析。其核心概念是寻找一个最优的超平面，使得不同类别的样本尽可能分开，同时最大化类别之间的间隔。在使用 SVM 进行目标检测时，首先需要利用标记好的正负样本训练 SVM 模型，即通过已知的包含目标（正样本）和不包含目标（负样本）的图像块来优化决策边界。训练完成后，该模型可用于预测新样本属于哪一类，从而实现对目标的有效识别。

3 单任务检测水平边界框（single-task）

参考思路

针对测试集，依次检测图片中的车辆和车牌。

可选思路

区域选择 → 特征提取 → 分类器分类，也可以采用其他方法，需详细说明所用方法的原理和参考文献（针对车辆及车牌区域的处理，建议考虑归一化尺寸：车辆 200×200 ，车牌 100×30 ）。

区域选择

滑动窗口、选择性搜索

特征提取及分类

HOG/DPM/Haar + SVM / Adaboost

评价指标

IoU (Intersection over Union)、AP (Average Precision)、Precision (精确率)、Recall (召回率) 等。

3.1 问题分析

3.1.1 样本集划分

训练数据集中包含1534张训练图片及标注文件，测试集中包含658张测试图片及标注文件。

原始的图片是 2560×1440 的RGB图片，在标注文件中给出了车辆和车牌图片的坐标信息。读取数据，绘制样本图片及标注信息如图 1所示：



图 1: 训练集待检测图片及标注示例

解决思路

1. 读取样本集中的标注数据
2. 根据标注裁切出样本集中的车辆和车牌对应区域，并保存为训练正样本
3. 从标注区域外围随机裁切出一定数量的区域，并保存为样本集中的负样本
4. 为了使模型能够学习到足够的信息，需要让负样本的数量尽可能多于正样本，考虑到计算效率等实际情况，本实验设置车辆检测的负样本数量是正样本的 5 倍，车牌检测的负样本数量是正样本的 10 倍
5. 为了防止模型过拟合，训练集中随机抽取20%的数据作为验证集

3.1.2 标注框的尺寸不固定

受限于拍摄角度和车辆的尺寸，难以使用固定高宽的预测框滑动检测。

统计了训练集中标注为“car”和“licence”的标注框的尺寸分布，统计结果如下所示：

总共有 659 个标注为 car 的对象
 宽度的最大值：2009，最小值：143，平均值：954.97，方差：120043.61
 高度的最大值：1097，最小值：67，平均值：651.53，方差：33080.81

总共有 647 个标注为 licence 的对象
 宽度的最大值：181，最小值：40，平均值：96.17，方差：701.94
 高度的最大值：99，最小值：20，平均值：49.57，方差：195.12

可以看出，目标框的方差都很大，说明尺寸分布非常不均匀，因此，不宜使用固定高宽的预测框进行检测。

解决思路

在训练时，把正样本和负样本图像的尺寸全部 Resize 到固定尺寸（车辆Resize到 200×200 ，车牌Resize到 100×30 ），并基于 resize 后的样本训练 SVM 分类器。

3.1.3 预测时计算复杂度高

图像经过HOG特征处理后得到的描述向量的长度取决于图像尺寸和HOG参数，通常维度较高。

此外，在预测阶段，如果直接使用不同大小的预测框滑动计算，那么预测框数量会非常多，计算复杂度也会非常高，计算成本极高。

解决思路

1. 使用PCA对于计算得到的HOG向量降维处理，减小后续的计算量
2. 在测试时，先对原始测试图像进行降采样，然后使用选择性搜索（Selective Search）算法 [2]生成候选框，并通过尺寸限制进一步过滤掉过小和过大的框
3. 尽可能使用并行计算，提高计算效率

3.1.4 输出多个预测框

保留被模型预测为正样的预测框，预测框之间会彼此重叠，因此需要进一步处理。

解决思路

1. 使用非极大值抑制（Non-maximum suppression）算法 [1]对预测框进行过滤
2. 得分综合考虑重叠区域面积，以及SVM输出的Platt Scaling置信度
3. 筛选出预测框中得分最高的预测框，作为模型预测结果

3.2 程序设计思想

综上所述，整理模型训练和推理的计算流程如图 2所示

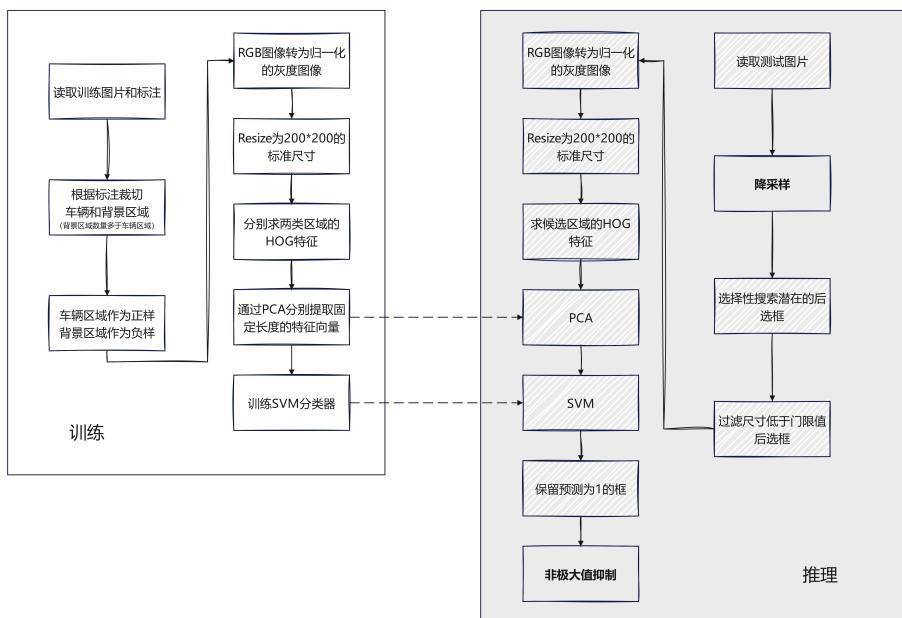


图 2: 模型训练和推理的计算流程

4 多任务检测水平边界框 (multi-task)

参考思路

针对测试集，依次检测图片中的车辆和车牌。

可选思路

1. 级联架构：车辆检测（HOG/DPM + Haar/SVM）→ 车牌检测(车辆区域内，利用颜色分割)
2. 联合检测架构：共享特征提取 → 双检测器并行

4.1 问题分析

本实验采用级联检测的思路，即首先检测车辆区域，随后在该区域内进一步检测车牌位置。这种结构能够有效缩小搜索范围，提升检测效率。

然而，由于车辆检测模型存在一定的漏检概率，若第一步未能正确识别出车辆区域，则后续的车牌检测过程将无法正常进行。因此，为提高整体检测的鲁棒性，有必要结合车牌检测的结果对车辆检测结果进行反馈和修正，从而构建一个更具协同性的联合检测流程。

4.2 程序设计思想

基于训练集中提供的车辆标注信息，我们可重新生成车牌检测任务中的负样本区域，并据此训练更为精确的车牌检测模型。

综合上述分析，设计了一个级联式的检测流程，如图 3 所示。该流程包括以下两个主要阶段：

1. 车辆检测：使用训练好的车辆检测模型初步定位图像中的车辆区域；
2. 车牌检测与反馈修正：在检测到的车辆区域内进行车牌识别，并根据车牌检测结果对车辆检测框进行反馈优化，提升整体检测精度与稳定性。

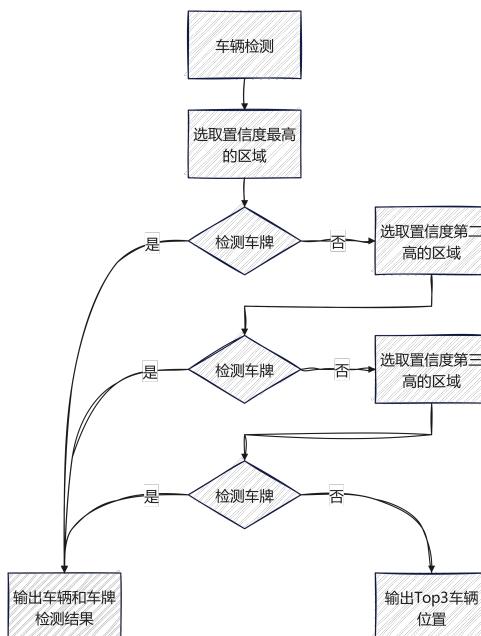


图 3: 级联检测流程

5 深度学习车辆信息检测（选做）

参考思路

采用深度神经网络对训练集数据进行训练，在测试集上验证车辆和车牌检测性能。

可选思路

特征提取 → 分类和bbox回归，也可以使用其他方法，需说明方法原理和参考文献

特征提取及分类

卷积神经网络（backbone+detector）

评价指标

IoU、AP、Precision、Recall等。

5.1 程序设计思想

在本实验中，图像目标检测任务可以从两个不同的角度实现。以下分别介绍这两种思路。

5.1.1 基于目标检测框架的改进方法

一种直观且高效的解决方案是采用已预训练的目标检测模型（如 YOLO 系列），并在本任务的数据集上进行微调。YOLO 类模型具有良好的实时性与检测精度，广泛应用于各类视觉检测任务中。通过迁移学习的方式，在其已有权重基础上针对特定数据集进行进一步训练，有助于快速收敛并提升模型性能。该方法能够直接输出目标的边界框和类别预测，适用于对检测效率和准确性均有要求的场景。

5.1.2 基于图像分割思想的替代方案

另一种可行的思路是将目标检测问题视为图像分割任务的一种简化形式。具体而言，可使用如 U-Net 等经典的语义分割网络对图像中的车辆及车牌区域进行像素级预测。随后，通过对分割结果提取最小外接矩形（Minimum Bounding Rectangle），将其转化为边界框形式的检测结果。这种方法在目标形状不规则或存在遮挡的情况下可能表现出更强的鲁棒性，同时也便于引入更复杂的后处理策略以优化检测框的位置和大小。

6 实验及结果分析

本节主要简要介绍该实验采用的模型的评价标准，描述三部分实验的结果，并对实验结果进行分析。

6.1 模型评价标准

考虑到目标检测任务的核心在于定位精度，而 IoU 指标能够直接反映预测边界框与真实边界框之间的重叠程度，因此具有更强的代表性，本实验后续将**IoU作为重点评价指标**。此外，通常情况下，若模型预测结果的 IoU 大于 0.5，则认为该预测是正确的。

表 1: 混淆矩阵定义

真实值/预测值	正例	负例
正例	TP	FN
负例	FP	TN

交并比 (IoU) 图像预测区域重合度:

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}$$

A 为真实标签面积, B 为预测标签面积。

6.2 实验结果

6.2.1 单任务检测

实验结果

从实验结果来看, 单任务检测模型的整体性能仍有较大提升空间。在包含658张图像的测试集中, 模型共检出513辆车 (真实标注数为659), 以及225个车牌 (真实标注数为647)。各项具体评估指标详见表 2。

表 2: 车辆与车牌检测模型性能对比

指标	车辆检测		车牌检测	
	值	数量统计	值	数量统计
TP	—	513	—	225
FP	—	103	—	433
FN	—	145	—	422
Precision	0.8328	—	0.3419	—
Recall	0.7796	—	0.3478	—
F1 Score	0.8053	—	0.3448	—
Mean IoU	0.7143	—	0.2848	—

结果分析

分类错误的样本可以大致分为六类, 在表 3中进行了简要的分析, 错误样本示例如图4 所示。

表 3: 任务1识别错误类型的名称及原因分析

编号	错误类型	原因分析
类型1	多个目标车辆，只检测到一个目标车辆	模型设置了只取置信度最高的1个目标，当图片中存在多个目标时，会漏掉其他目标。
类型2	多个目标车辆，误识别为一个目标车辆	类似于“类型1”，可能由于目标之间的距离较近所导致。
类型3	完全误识别为车辆	背景中的某些特征被误认为是车辆，可能由于用于训练的负样本多样性不足所导致。
类型4	未检测到车辆	目标车辆存在，但未能成功识别，可能由车辆画面不完整所导致。
类型5	将车身其他位置误识别为车牌	可能由于用于训练的负样本多样性不足所导致。
类型6	将车身外区域误识别为车牌	类似于“类型5”，可能也是由用于训练的负样本多样性不足所导致，但错误发生在车身外部，可以通过联合检测来解决。



图 4: 单任务检测错误示例

6.2.2 多任务检测

实验结果

从实验结果来看，多任务检测模型在车牌识别任务中相较于单任务模型表现出更优的性能，但在车辆检测任务中却出现了性能下降。该现象可能归因于模型在预测得到的车辆区域内未能成功定位车牌区域，从而对整体检测流程产生了负面影响。

在由 658 张图像组成的测试集中，模型共识别出 508 辆车辆（真实标注数为 659），并检测到 281 个车牌（真实标注数为 647）。和单任务检测的具体性能指标对比详见表 4。

表 4: 单任务与多任务检测模型在车辆与车牌检测中的性能对比

指标	单任务检测		多任务综合检测	
	车辆检测值	车牌检测值	车辆检测值	车牌检测值
TP	513	225	508	281
FP	103	433	188	286
FN	145	422	150	366
Precision	0.8328	0.3419	0.7299	0.4956
Recall	0.7796	0.3478	0.7720	0.4343
F1 Score	0.8053	0.3448	0.7504	0.4629
Mean IoU	0.7143	0.2848	0.6285	0.4338

结果分析

分类错误的样本主要有五类，在表 5 中进行了简要的分析，错误样本示例如图 5 所示。其中部分任务可以通过改进训练集组成改善。

表 5: 任务 2 识别错误类型的名称及原因分析

编号	错误类型	原因分析
类型1	车辆识别正确，车牌识别错误	车牌检测模型仅选取置信度最高的一个目标，导致漏检真实车牌；同时，由于训练中负样本多样性不足，模型易将车身其他区域误识别为车牌。
类型2	车辆识别区域不完整，车牌暴露在车辆区域外	车辆检测器因正样本不足，导致截取的车辆区域不完整，使车牌位于预测框之外。可通过数据增强扩充训练集以提升检测完整性。
类型3	车辆识别错误	车辆存在但被错误识别，且在错误区域内仍检测出车牌，可能是由于错误区域与真实目标特征相似所致。需通过优化训练集来提升模型鲁棒性。
类型4	未检测到车辆	车辆确实存在，但未能被检测到，可能由于训练样本多样性不足，导致模型对某些外观或姿态的目标泛化能力较弱。
类型5	检测到车辆，但未能检测到车牌	车辆区域已正确识别，但未检测到其中的车牌，同样可能是由于车牌检测器训练样本多样性不足所导致。

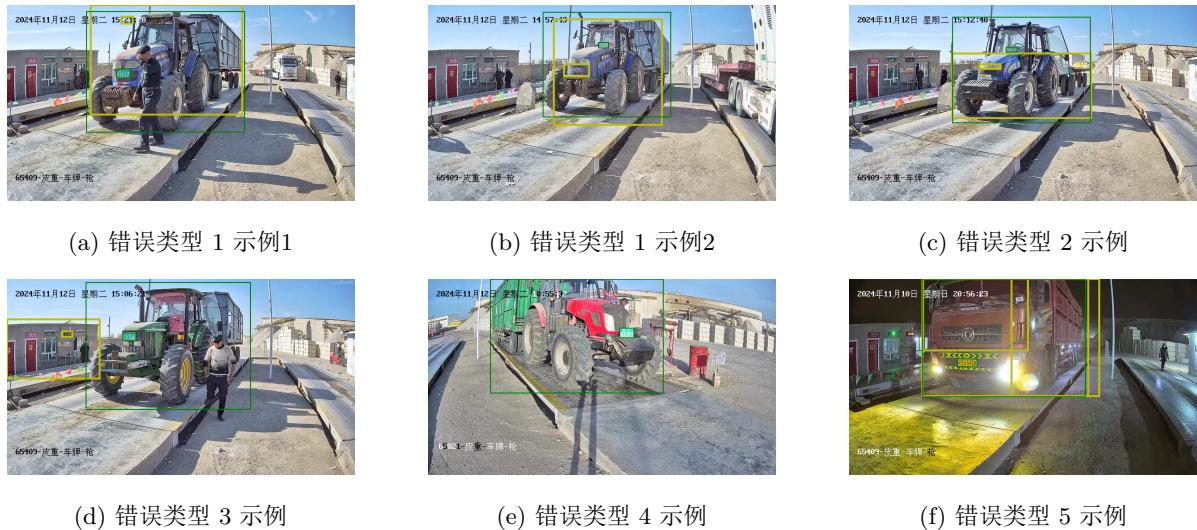


图 5: 多任务检测错误示例

6.2.3 深度学习方法检测(尚未完成)

实验结果

结果分析

7 总结

通过本次实验，实现了使用HOG+SVM进行车辆车牌单任务检测以及多任务联合检测。对非深度学习的目标检测方法有了更深入的理解。

通过本次实验，实现了基于 HOG 特征提取与 SVM 分类器的车辆及车牌目标检测，完成了单任务检测与多任务联合检测的算法实现。在实验过程中，不仅掌握了传统非深度学习方法在目标检测任务中的基本流程，加深了对 HOG 特征提取SVM 分类器的理解。

实验结果表明，虽然 HOG+SVM 方法计算资源不充沛的在场景（如边缘计算等）下仍具有一定的实用性，但训练一个效果很好的模型需要大量的训练数据和优化策略。

本次实验加深了我对传统目标检测方法的理解，可续可以完善使用深度学习方法进行目标检测的部分，加深各方法的理解。

参考文献

- [1] Alexander Neubeck and Luc J. Van Gool. Efficient non-maximum suppression. In *International Conference on Pattern Recognition*, 2006.
- [2] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. M. Smeulders. Selective search for object recognition. *International Journal of Computer Vision*, 104(2):154–171, 2013.