

基于难分样本挖掘的快速区域卷积神经网络目标检测研究

张 焱 许 艇 冯定忠 蒋美仙* 吴光华

(浙江工业大学机械工程学院 杭州 310023)

摘 要: 针对经典的快速区域卷积神经网络(Faster RCNN)训练过程存在太多难训练样本、召回率低等问题, 该文采用一种基于在线难分样本挖掘技术(OHEM)与负难分样本挖掘(HNEM)技术相结合的方法, 通过训练中实时筛选的最大损失值难分样本进行误差传递, 解决了模型对难分样本检测率低问题, 提高模型训练效率; 为更好地提高模型的召回率和模型的泛化性, 该文改进了非极大值抑制(NMS)算法, 设置了置信度阈值罚函数, 又引入多尺度、数据增强等训练方法。最后通过比较改进前后的结果, 经敏感性实验分析表明, 该算法在VOC2007数据集上取得了较好效果, 平均精度均值从69.9%提升到了74.40%, 在VOC2012上从70.4%提升到79.3%, 验证了该算法的优越性。

关键词: 多目标检测; 在线样本挖掘; 负难分样本挖掘; 深度学习; 非极大值抑制

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2019)06-1496-07

DOI: 10.11999/JEIT180702

Research on Faster RCNN Object Detection Based on Hard Example Mining

ZHANG Ye XU Ting FENG Dingzhong JIANG Meixian WU Guanghua

(College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: Because of the classic Faster RCNN training process with too many difficult training samples and low recall rate problem, a method which combines the techniques of Online Hard Example Mining (OHEM) and Hard Negative Example Mining (HNEM) is adopted, which carries out the error transfer for the difficult samples using its corresponding maximum loss value from real-time filtering. It solves the problem of low detection of hard example and improves the efficiency of the model training. To improve the recall rate and generalization of the model, an improved Non-Maximum Suppression (NMS) algorithm is proposed by setting confidence thresholds penalty function; In addition, multi-scale training and data augmentation are also introduced. Finally, the results before and after improvement are compared: Sensibility experiments show that the algorithm achieves good results in VOC2007 data set and VOC2012 data set, with the mean Average Precision (mAP) increasing from 69.9% to 74.40%, and 70.4% to 79.3% respectively, which demonstrates strongly the superiority of the algorithm.

Key words: Multiple object detection; Online Hard Example Mining (OHEM); Hard Negative Example Mining (HNEM); Deep learning; Non-Maximum Suppression (NMS)

1 引言

近几年随着计算机科学的飞速发展, 基于计算机技术的图像处理^[1,2]、图像目标检测^[3-5]等也获得了空前的快速发展, 其中深度学习通过学习海

量的数字图像特征, 提取关键目标特征, 在目标检测上已超过传统方法。目标识别其中的两大任务是目标定位检测和目标物体分类, 而传统方法的图像特征检测是依靠富有经验的算法工程师设计对应目标的匹配模板, 比如可变形模型^[6-8](deformable part model)、方向梯度直方图(Histogram Of Gradient, HOG)特征提取^[9,10]等。其存在的问题是检测耗时长, 特征匹配精度低, 对特定的目标有效, 导致模型泛化能力弱等一系列问题。而深度学习方法能够针对复杂场景的目标, 进行有效地特征识别, 识别效果远优于传统方法, 但也存在的不足

收稿日期: 2018-07-13; 改回日期: 2019-01-28; 网络出版: 2019-02-18

*通信作者: 蒋美仙 1056294025@qq.com

基金项目: 国家自然科学基金(51605442), 浙江省科技厅公益项目(LGN18G010002)

Foundation Items: The National Natrual Science Foundation of China (51605442), Science Technology Department of Zhejiang Province (LGN18G010002)

之处: (1)数据量大, 深度学习的训练集需要成千上万个样本, 且数据越多, 效果越好, 给数据采集工作带来一定的困难; (2)硬件要求较高, 训练大数据样本, 需要至少4 G的显存, 对硬件提出较高要求; (3)训练技巧性强, 当参数设置不合理时, 样本训练效果较差, 出现难训练的情况。

因此, 本文针对样本问题提供了一种在不增加样本情况下的在线样本挖掘技术(OHEM)和负难分样本挖掘(HNEM)相结合的技术, 使模型能够针对现有的难分样本, 有针对性地学习其特征, 达到模型进一步提升泛化性、鲁棒性的效果。针对训练显存与参数设置问题, 提供了改进的设置方法, 达到较好训练的效果。进一步的, 针对成群的、重叠在一起的多目标检测问题, 提供改进非极大值抑制算法, 降低模型的漏检率, 提高模型的召回率。

2 本文OHEM和HNEM相结合的方法

基于在线样本挖掘^[11]的思想是每次随机梯度下降的样本由一张图片产生的上百个建议框进行难分样本筛选训练的, 而反向误差传递过程中只针对难分样本进行调整, 原因是罕见的、损失过大的样本在所有样本中较少并难训练, 而训练较好的样本误差值较小但却很难再梯度下降。尽管1次迭代的批量大小等于1, 但是由区域候选网络RPN产生的建议框成百上千, 正适合于随机梯度下降的方法, 并且这种实时筛选机制能够很有针对性地进行学习训练, 因此模型训练调整很快。

另外, 采用的负难分样本挖掘是在难分样本基础上通过调整训练中RPN形成的小批次正负样本比例为1:3, 进行训练。因为一张图片中包含的目标物体信息相对于背景信息而言是相对较小的部分, 如只采用难分的假正例或真正例进行挖掘训练都各有所欠缺。当设置正负样本的比例为1:3既能很好训练正样本, 又在不引入其他背景样本下, 充分学习较多背景信息, 提高模型正负样本的判决力。

负难分样本挖掘的策略为:

(1) 取消负样本的阈值限制。在原先的快速区域卷积神经网络(Faster RCNN)中, 判定为背景的机制是根据RPN产生的候选建议框与真实框之间的交并比IoU的大小进行设定, 当IoU的比例在[0.1, 0.5)时, 被认为是背景的建议框。这忽略了低于0.1以下的罕见、重要的难分负样本, 这些特征没能很好地被学习。因此, 本文设置判定背景的阈值为[0, 0.5)。设置了正样本的阈值为IoU值最大的一个, 或者IoU的阈值在[0.7, 1.0]范围;

(2) 设置难分正负样本的比例1:3。如前所述,

正负样本同时兼顾, 并且更多地学习背景信息, 能够提高模型在特定背景环境下定位检测目标的能力。并且设置了每次的目标图片批次大小为64, 则正样本的数量为16, 负样本的数量为48;

(3) 同时考虑正样本与标签存在多个重复的情况, 采用了非极大值抑制方法, 采用0.7的阈值, 当目标与标签的IoU在0.7以下时, 则予以删除。网络层的参数设置信息如表1所示。

表 1 负难分样本挖掘参数设置

| 参数名称 | 代表含义 | 参数取值 |
|-----------------|----------|------------|
| FG_THRESH | 正样本IoU阈值 | [0.7, 1.0] |
| BG_THRESH_LO | 负样本IoU阈值 | [0, 0.5) |
| HNEM_NMS_THRESH | 非极大值抑制阈值 | 0.7 |
| HNEM_BATCHSIZE | 图片目标批次大小 | 64 |
| RPN_FG_FRACTION | 正样本比例 | 0.25 |
| RPN_BG_FRACTION | 负样本比例 | 0.75 |

基于在线样本挖掘方法, 采用的关键参数为:

(1) 设定难分样本的筛选机制。每次根据当前的总损失 $L(\{p_i\}, \{t_i\})$ 降序排列, 筛选前 B/N 个样本量, 其中 $B = 64$, N 为每次训练的图片张数 $N = 1$, 这样训练的反向传播速度更快;

(2) 提高模型的运算速度。在反向传递时, 通过非难分样本损失置0操作, 并不能减少模型显存量。所以后向传递筛只选出 B/N 难分样本进行梯度传递, 训练的显存容量从3.5G减少到3.0G, 减少了显存的浪费, 提高了训练速度;

(3) 在RPN层和最后全连接层同时增加OHEM模块。由于良好的分类结果取决于精确的目标定位, 即建议框的产生, 所以对RPN层也增加OHEM模块(图1、表2)有利于边框回归出最精确的位置, 提高分类特征提取的有效性, 促进了模型的训练;

(4) 适当地调整分类损失和边框回归损失的权重。在损失函数公式中, 分类损失 L_{cls} 和边框回归损失 L_{reg} 是不均衡的, $N_{cls} = 256$ 是分类的个数, 而 $N_{reg} = 2400$ 是最后特征图滑动的最大维度, 两者取倒数相差约10倍, 因此取 $\lambda = 10$, 这样有利于在回归较好边框的同时, 有针对性地学习相应的目标特征;

(5) 调整非极大值抑制(NMS)算法^[12]。经典的非极大值抑制算法不能很好地保留相邻或重叠的多目标检测框, 导致目标检测的召回率降低。因此, 本文改进的非极大值抑制算法采用了分类的置信度惩罚机制, 在去除冗余的建议框的同时还很好地保留了建议框, 并进一步提高了mAP;

(6) 数据增强工作提高了模型的泛化能力。训练时通过随机左右镜像翻转和光照饱和度调节, 增

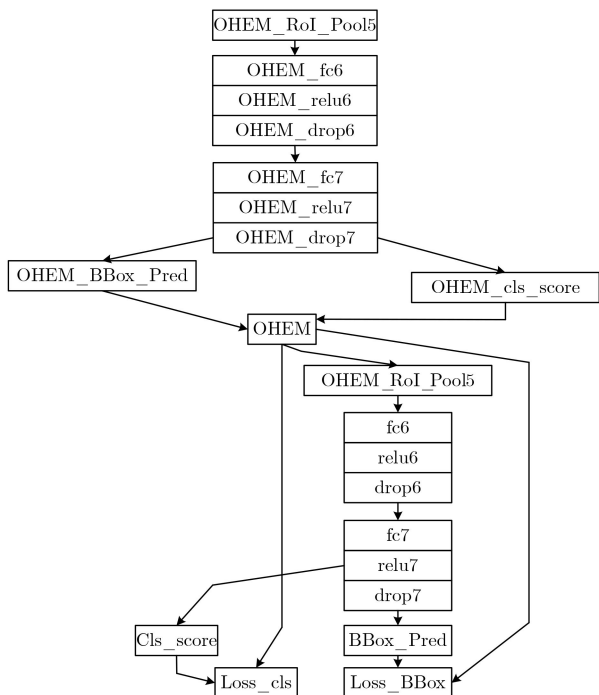


图1 增设的OHEM模块

表2 在线样本挖掘参数设置

| 参数名称 | 代表含义 | 参数取值 |
|----------------|----------|----------|
| ITERS | 每次迭代个数 | 1 |
| OHEM_ROI_POOL5 | 在线样本兴趣池化 | 7×7 |
| OHEM_FC6 | 在线样本全连接层 | 4096 |
| OHEM_RELU6 | 在线样本激活操作 | - |
| OHEM_FC7 | 在线样本全连接层 | 4096 |
| OHEM_RELU7 | 在线样本激活操作 | - |
| OHEM_CLS_SCORE | 在线样本分类数 | 21 |
| OHEM_CLS_PRED | 在线样本边框矩阵 | 84 |
| OHEM | 在线样本处理模块 | OHEMData |

加样本多样性,防止过拟合。为提高模型对不同尺寸图像的检测能力,在训练时使用了多尺度训练^[13,14],将图像的短边设置为{224, 416, 480, 512, 600, 672, 900}的随机尺寸,另一边等比例放缩。样本的增强工作也进一步地提高了mAP;

(7) 调整了学习率的策略。将学习速率的下降

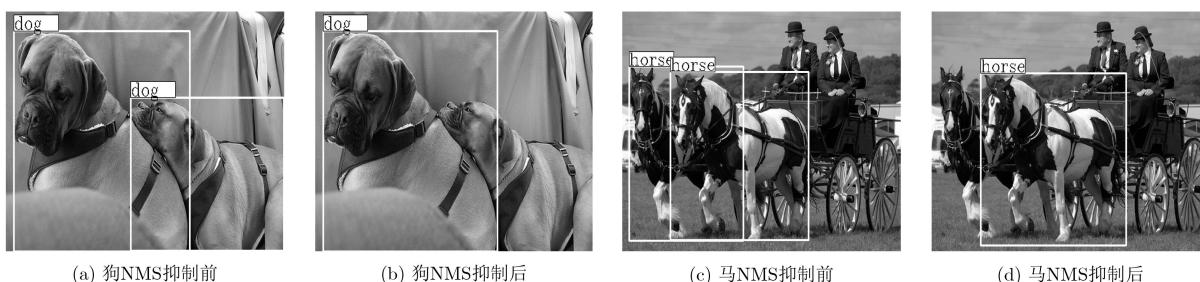


图2 经典非极大值抑制存在的问题

点设置为40k,相当于8个epoch,并且此后隔20k个迭代次数再下降一次学习速率,这样有利于提高模型前期的全局范围搜索能力,避免陷入局部极小值点;同时,后期采用的较小学习速率,进行极小值的精细化调整,有利于损失的一降再降。

3 改进的非极大值抑制算法

3.1 经典的非极大值抑制算法及其缺陷

采用非极大值抑制^[12,15]的目的是剔除冗余框,避免损失的多重计算。经典的非极大值抑制^[15]算法是将RPN层网络生成的建议框去除冗余,因在最后的卷积特征层滑动生成的9种Anchor相互重叠。

经典非极大值抑制算法流程为:首先计算出所有建议框的面积,对所有建议框的分数进行从大到小排序,取出分类置信度最高分数的序号;循环计算当前最高分的窗口到次高分数窗口的交叉面积;再计算交并比IoU,即交叉面积占两个总面积的比例;若比例超过既定阈值,则把次高分的窗口抑制去冗余,否则保留该建议窗口。

这种贪心式算法存在图2的问题,右图因交并比IoU大于既定的阈值而被剔除,造成了目标的漏检。另一方面,NMS的阈值也不太容易确定,设小了会出现以上情况,设置过高又容易增大误检。

3.2 改进的非极大值抑制算法

针对上述阈值存在两难的情形以及阈值选取问题,采取的改进方式是降低其置信度。采取的计算方法有线性加权、高斯加权法、指数加权法。线性加权的计算方法是将IoU大于阈值的降低其置信度,即引入罚函数的思想:

$$s_i = \begin{cases} s_i, & \text{IoU}(b_m, b_i) < N_t \\ s_i \cdot a \cdot (1 - \text{IoU}(b_m, b_i)), & \text{IoU}(b_m, b_i) \geq N_t \end{cases} \quad (1)$$

其中, s_i 表示当前所属类别的置信度得分, a 表示权值系数, $0 < a \leq 1$, b_m 表示置信度得分最高所对应的边框, b_i 表示当前的边框, $\text{IoU}(b_m, b_i)$ 表示两者的交并比, N_t 表示所给定的IoU阈值。

而采用高斯加权法的计算公式为

$$s_i = s_i e^{-\frac{\text{IoU}(b_m, b_i)^2}{\sigma}}, \forall b_i \notin D \quad (2)$$

其中, σ 影响罚函数的惩罚力度, 高斯加权函数相比线性加权具有平滑过渡等特点。

指数加权法同样采用了一定的阈值, 当其IoU大于该阈值时, 对其进行惩罚。计算公式为

$$s_i = \begin{cases} s_i, & \text{IoU}(b_m, b_i) < N_t \\ s_i e^{(N_t - \text{IoU}(b_m, b_i))}, & \text{IoU}(b_m, b_i) \geq N_t \end{cases} \quad (3)$$

其中, N_t 同样表示设定的IoU阈值, s_i 表示当前所属类别的置信度得分, b_m 表示置信度得分最高所对应的边框, b_i 表示当前的边框。

基于以上3种改进的非极大值抑制算法, 归结统一的流程操作如表3所示。

针对每一种改进的非极大值抑制算法, 对应的具体流程和各自的适应情况总结如下:

(1) 针对线性加权方法, 通过置信度降序, 筛选某一类中置信度最高的边框作为最优框, 比较下一边框的IoU与既定阈值 N_t , 当小于该阈值时, 置信度分值不变; 否则置信度阈值降低为 $a(1 - \text{IoU})$ 的比例, 即IoU越大, 惩罚力度越大。循环操作, 直到所有待定边框的置信度小于给定的置信度阈值, 则舍弃, 记录历史最优值。该方法适用于重复边框较多、时间复杂度低、快速筛选结果的情景, 即惩罚的力度较大, 很快筛选完毕, 但得到的效果并不太理想。

(2) 针对高斯加权法, 同样筛选出最优置信度边框, 但无需设定 N_t 的阈值, 通过IoU指数递减方

法, 并且设定的 σ 值控制着递减的程度。该方法适用于预测产生的边框数量中等、时间要求不高的、数量统计准确的情况, 并且平滑过渡的筛选有助于得到较好的回归框, 虽然时间复杂度较线性差, 但得到的结果较优;

(3) 针对指数加权法, 同样比较次优置信度边框与最优边框的IoU, 设定一定的阈值 N_t 。当小于该阈值时, 置信度分值不变; 否则置信度阈值降低为 $e^{(N_t - \text{IoU}(b_m, b_i))}$, 该指数呈现递减的效果。直到所有的框都检测完毕, 保留置信度大于阈值的建议框。该方法适用于保留较多回归框、删选部分冗余框的情景, 可应用于只需检测目标有无, 不需统计数量的情况, 因此效果比线性和高斯加权法差。

4 实验结果的对比

4.1 性能评价指标的设计

到目前为止, 最具有说服力、通用的模型评价指标(模型的性能)主要有准确率(Precision, P)、召回率(Recall, R)、AP指标(Average Precision)、mAP指标(mean Average Precision)。在文献[13]针对这些评价指标的定量问题, 提出了定量客观标准。文献[14]中也详细介绍了以上指标, 准确率也叫查准率, 召回率也叫查全率, 但两者是一对矛盾。准确率较高时, 召回率较差; 召回率高时, 准确率下降。因此, 引入AP指标, 即 $P-R$ 曲线面积来反映两者均衡的评价效果。其中, 召回率 R 是曲线的横坐标, 准确率 P 是曲线的纵坐标, 两者的计算公式为

$$R = \frac{TP}{TP+FN}, P = \frac{TP}{TP+FP} \quad (4)$$

其中, TP为真正例(True Positive), FP为假正例(False Positive), FN为假反例(False Negative)。

4.2 实验结果的对比

本实验在VOC2007与VOC2012的公共数据集上实验。对比了不同条件下的在线样本挖掘与难分样本挖掘的组合, 实验结果如表4、图3所示。其中, 表4列举了部分, 第1行是原Faster RCNN (FRCNN)的结果; 第2行是只在最后全连接层增设OHEM模块的结果, 实验效果没有提升, 若只在RPN层增设OHEM模块, 实验效果也没有得到较好提升; 第3行是将RPN层和最后全连接层都增设OHEM模块, 并采用随机正负样本比例的在线样本挖掘方法, 分类: 边框回归的损失权重=1:1, 效果提升了; 第4行是在前面的基础上将RPN层分类: 边框回归的损失权重设置为1:10, 即 $\lambda = 10$ 的方法, 采用两类损失平衡后, 进一步提升了1.3%, 相比原来提升了1.8%, 均衡两者的损失有利于模型的训练;

表 3 改进的非极大值抑制算法

输入: 候选边框集合 $B = \{b_1, b_2, \dots, b_N\}$, 置信度集合 $S = \{s_1, s_2, \dots, s_N\}$, IoU阈值 N_t

循环操作:

最优框 $D \leftarrow \{\}$

While $B \neq \text{Null}$ do

$m \leftarrow \arg \text{Max}(S)$

$M \leftarrow b_m$

$D \leftarrow D \cup M; B \leftarrow B - M$

for b_i in B do

If $\text{IoU}(M, b_i) \geq N_t$ then

weight = Method(1 - 3)

$s_i \leftarrow s_i * \text{weight}$

If $s_i \leq \text{threshold}$

$B \leftarrow B - b_i$

End

End

End

End

输出最终结果: D, S

第5行是将正负样本的随机比例去除, 设置为1:3的比例, 且 $\lambda = 10$, 效果又提升了, 而其他比例无显著提升, 说明负难分样本挖掘在模型训练中起到一

定作用; 第6行在OHEM和HNEM基础上, 通过数据增强手段, 进一步提升mAP; 第7行至第9行采用了改进的NMS算法, 由敏感性分析(图4)得最佳

表4 在线样本挖掘等实验mAP指标结果

| 类别 | bird | boat | bottle | bus | car | chair | cow | table | dog | horse | person | plant | sheep | sofa | train | mAP |
|-----------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| FRCNN | 68.5 | 54.7 | 50.6 | 78.1 | 80.2 | 50.7 | 74.6 | 65.5 | 81.3 | 83.7 | 75.7 | 38.3 | 70.6 | 67.1 | 80.7 | 69.9 |
| ohem_fc | 69.2 | 57.9 | 46.5 | 81.8 | 79.1 | 47.9 | 76.2 | 68.9 | 83.2 | 80.8 | 72.7 | 39.9 | 67.5 | 66.2 | 75.6 | 69.9 |
| ohem1: 1 | 71.1 | 54.6 | 52.3 | 79.7 | 81.3 | 50.3 | 74.3 | 66.8 | 80.7 | 83.7 | 76.7 | 40.9 | 70.0 | 68.2 | 77.6 | 70.4 |
| ohem1: 10 | 71.8 | 58.5 | 53.2 | 79.3 | 82.9 | 52.2 | 81.2 | 70.0 | 81.4 | 83.2 | 77.9 | 43.7 | 71.9 | 67.1 | 75.0 | 71.7 |
| ohem1: 3 | 72.2 | 57.8 | 56.6 | 80.8 | 84.0 | 53.8 | 77.5 | 68.0 | 82.2 | 84.0 | 77.6 | 43.2 | 70.9 | 68.4 | 79.4 | 72.1 |
| 数据增强 | 69.8 | 62.0 | 55.2 | 80.2 | 83.6 | 54.5 | 80.3 | 67.2 | 80.7 | 85.0 | 78.0 | 44.6 | 70.8 | 69.4 | 79.0 | 72.5 |
| NMS-线 | 74.5 | 64.4 | 57.8 | 80.0 | 84.3 | 57.4 | 80.8 | 70.1 | 83.2 | 83.7 | 81.3 | 48.3 | 71.9 | 68.4 | 79.4 | 74.1 |
| NMS-高 | 74.7 | 64.0 | 58.5 | 80.5 | 84.5 | 56.9 | 81.5 | 70.1 | 83.8 | 84.2 | 81.5 | 47.8 | 71.5 | 69.1 | 79.6 | 74.3 |
| NMS-指 | 73.7 | 63.7 | 56.9 | 79.6 | 83.9 | 56.5 | 80.7 | 69.4 | 82.8 | 82.7 | 80.8 | 48.0 | 70.5 | 66.8 | 79.2 | 73.3 |
| Lr-调整 | 75.8 | 63.3 | 57.6 | 81.1 | 84.7 | 56.5 | 83.1 | 70.6 | 84.8 | 85.2 | 81.2 | 47.8 | 71.6 | 68.6 | 79.1 | 74.4 |
| 12+ohem | 76.8 | 64.8 | 61.4 | 85.0 | 84.1 | 59.9 | 82.6 | 61.9 | 88.5 | 85.2 | 86.9 | 56.7 | 79.5 | 67.5 | 85.4 | 77.5 |
| 12+ohem* | 78.1 | 65.0 | 55.4 | 84.9 | 84.0 | 62.1 | 83.6 | 67.3 | 91.3 | 88.9 | 85.6 | 54.7 | 83.8 | 77.3 | 88.3 | 79.3 |

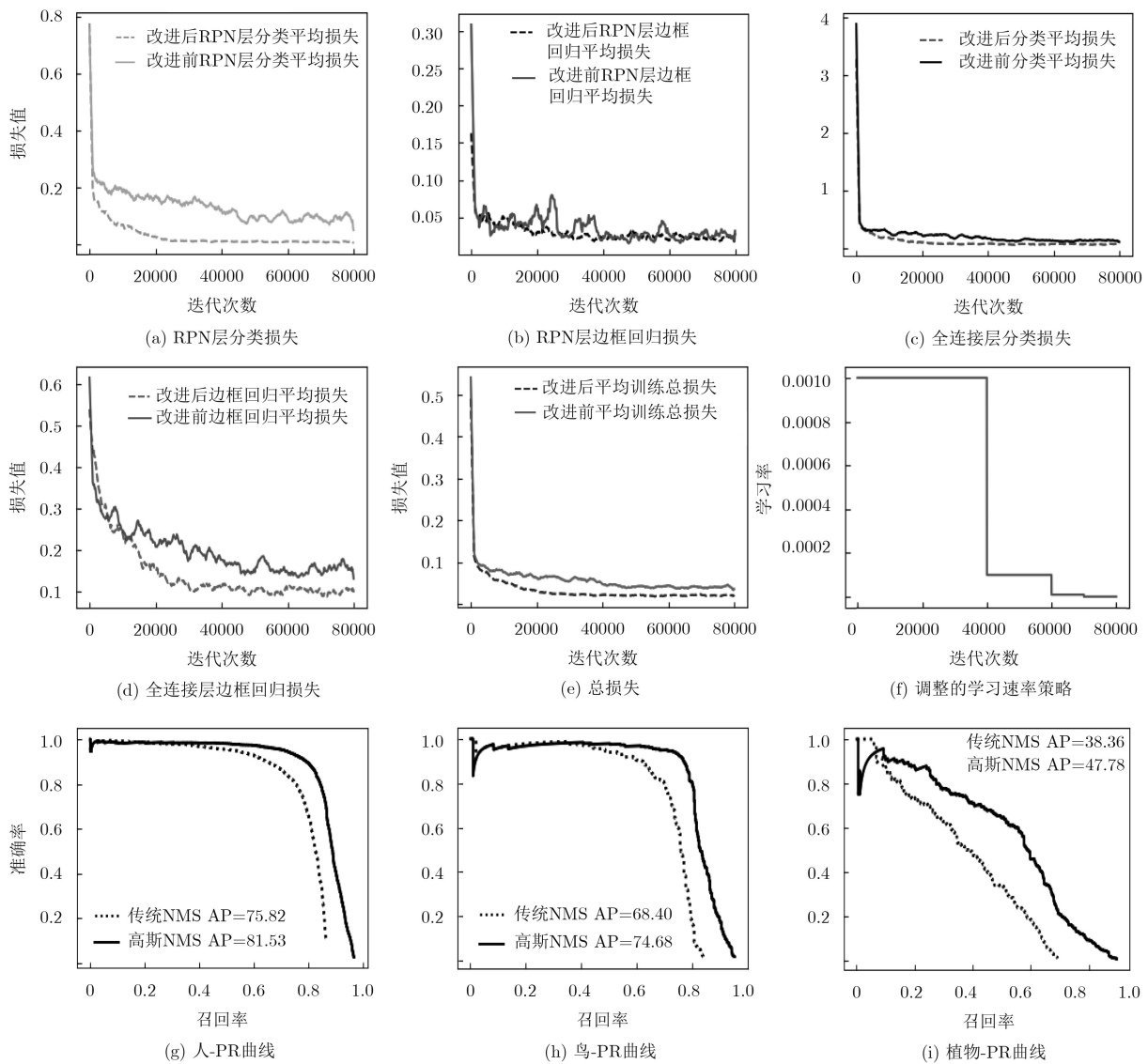


图3 改进前后损失曲线与召回率的表现

为高斯加权法, 设置 $\sigma = 0.3$, 阈值为0.003, 提升的目标物体大多为集群物体, 如狗、人、鸟、羊、植物, 分别提升了3.1%, 7.6%, 9.1%, 1.3%, 24.8% (图3(g)—图3(i)); 而第10行, NMS高斯改进基础上调整学习速率策略, 对损失进行再次下降(图3(f)), mAP相比原来提升了4.5。第11行, 采用VOC2012的样本测试, 相当于增加训练样本, 得到3.1点的提升; 第12行, 进一步采用高斯方法测试, mAP又提升1.8。由损失收敛曲线分析, 调整后的学习速率策略使边框回归和分类的损失不断下降(图3(a)—图3(e))。实验中发现增加前期搜索的时长, 将学习速率下降点设置为40k, 有利于避免陷入局部极小值点。

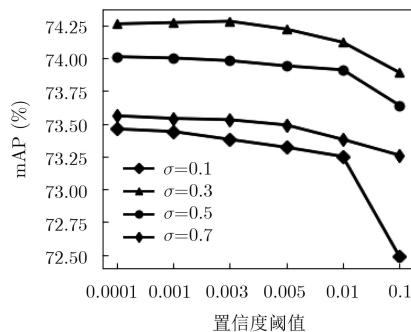


图4 敏感性分析实验

5 结束语

本文基于在线样本挖掘与负难分样本挖掘相结合的技术, 最突出的特点是在不扩充样本的情况下, 放宽负样本的定义, 通过样本本身在线挖掘出更多的难训练样本; 设置了正负样本比例, 合理而简便地计算损失最大、难训练的、罕见的样本; 对分类、边框回归的损失进行了平衡化处理, 能够满足训练损失的持续降低。其次针对多目标重叠漏检问题, 通过改进的非极大值抑制, 采用3种不同方式引入罚函数思想减低其置信度; 经敏感性分析得出采用高斯加权法对模型效果最好, 提升了召回率, 解决了多目标漏检问题。

参考文献

- [1] 吕博云. 数字图像处理技术及应用研究[J]. 科技与创新, 2018(2): 146–147. doi: [10.15913/j.cnki.kjyxc.2018.02.146](https://doi.org/10.15913/j.cnki.kjyxc.2018.02.146).
LÜ Boyun. Research on the technology and application of digital image processing[J]. *Science and Technology & Innovation*, 2018(2): 146–147. doi: [10.15913/j.cnki.kjyxc.2018.02.146](https://doi.org/10.15913/j.cnki.kjyxc.2018.02.146).
- [2] 王湘新, 时洋, 文梅. CNN卷积计算在移动GPU上的加速研究[J]. 计算机工程与科学, 2018, 40(1): 34–39. doi: [10.3969/j.issn.1007-130X.2018.01.005](https://doi.org/10.3969/j.issn.1007-130X.2018.01.005).
WANG Xiangxin, SHI Yang, and WEN Mei. Accelerating CNN on mobile GPU[J]. *Computer Engineering & Science*, 2018, 40(1): 34–39. doi: [10.3969/j.issn.1007-130X.2018.01.005](https://doi.org/10.3969/j.issn.1007-130X.2018.01.005).
- [3] 胡炎, 单子力, 高峰. 基于Faster-RCNN和多分辨率SAR的海上舰船目标检测[J]. 无线电工程, 2018, 48(2): 96–100. doi: [10.3969/j.issn.1003-3106.2018.02.04](https://doi.org/10.3969/j.issn.1003-3106.2018.02.04).
HU Yan, SHAN Zili, and GAO Feng. Ship detection based on faster-RCNN and multiresolution SAR[J]. *Radio Engineering*, 2018, 48(2): 96–100. doi: [10.3969/j.issn.1003-3106.2018.02.04](https://doi.org/10.3969/j.issn.1003-3106.2018.02.04).
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2014: 580–587. doi: [10.1109/CVPR.2014.81](https://doi.org/10.1109/CVPR.2014.81).
- [5] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137–1149. doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031).
- [6] FELZENSZWALB P, MCALLESTER D, and RAMANAN D. A discriminatively trained, multiscale, deformable part model[C]. Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, USA, 2008: 1–8. doi: [10.1109/CVPR.2008.4587597](https://doi.org/10.1109/CVPR.2008.4587597).
- [7] YAN Junjie, LEI Zhen, WEN Longyin, et al. The fastest deformable part model for object detection[C]. Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2014: 2497–2504.
- [8] FORSYTH D. Object detection with discriminatively trained part-based models[J]. *Computer*, 2014, 47(2): 6–7. doi: [10.1109/MC.2014.42](https://doi.org/10.1109/MC.2014.42).
- [9] DALAL N and TRIGGS B. Histograms of oriented gradients for human detection[C]. Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, USA, 2005: 886–893. doi: [10.1109/CVPR.2005.177](https://doi.org/10.1109/CVPR.2005.177).
- [10] WANG Xiaoyu, HAN T X, and YAN Shuicheng. An HOG-LBP human detector with partial occlusion handling[C]. Proceedings of 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 2009: 32–39. doi: [10.1109/ICCV.2009.5459207](https://doi.org/10.1109/ICCV.2009.5459207).
- [11] ERHAN D, SZEGEDY C, TOSHEV A, et al. Scalable object detection using deep neural networks[C]. Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2014: 2155–2162. doi: [10.1109/CVPR.2014.273](https://doi.org/10.1109/CVPR.2014.273).

- [10.1109/CVPR.2014.276](https://doi.org/10.1109/CVPR.2014.276).
- [12] NEUBECK A and VAN GOOL L. Efficient non-maximum suppression[C]. Proceedings of the 18th International Conference on Pattern Recognition, Hongkong, China, 2006: 850–855. doi: [10.1109/ICPR.2006.479](https://doi.org/10.1109/ICPR.2006.479).
- [13] 李航. 统计学习方法[M]. 北京: 清华大学出版社, 2012: 18–23. LI Hang. Statistical Learning Method[M]. Beijing: Tsinghua University Press, 2012: 18–23.
- [14] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 23–35. ZHOU Zhihua. Machine Learning[M]. Beijing: Tsinghua University Press, 2016: 23–35.
- [15] SUN Changming and VALLOTTON P. Fast linear feature detection using multiple directional non-maximum suppression[J]. *Journal of Microscopy*, 2009, 234(2): 147–157. doi: [10.1111/jmi.2009.234.issue-2](https://doi.org/10.1111/jmi.2009.234.issue-2).
- 张 焯: 男, 1973年生, 副教授, 硕士生导师, 研究方向为物联网、深度学习、无线传感器网络的设计与仿真等.
- 许 艇: 男, 1993年生, 硕士生, 研究方向为计算机视觉、深度学习、物联网技术等.
- 冯定忠: 男, 1963年生, 教授, 博士生导师, 研究方向为企业智能物流、工业工程技术及应用等.
- 蒋美仙: 女, 1973年生, 副教授, 硕士生导师, 研究方向为企业物流、系统工程等.
- 吴光华: 男, 1983年生, 讲师, 博士, 研究方向为智能物流、物联网技术等.