

〈红外应用〉

基于改进 Faster R-CNN 的车辆乘员数量检测方法

金鑫, 胡英

(大连海事大学 船舶电气工程学院, 辽宁 大连 116026)

摘要: 针对现有以雷达技术和红外热成像技术为代表的 HOV (High occupancy vehicle lane) 车道车辆乘员数量检测方法可靠性差、准确率低等问题, 提出一种基于多光谱红外图像与改进 Faster R-CNN (Region-Convolutional Neural Networks) 的车辆乘员数量检测方法。通过多光谱红外成像系统获得汽车内部空间图像, 结合 Faster R-CNN 深度学习算法实现乘员数量检测, 通过采用全卷积网络结构、多尺度特征预测、使用 ROI-Align 代替 ROI-Pooling 等方式增强网络的泛化能力。通过对样据进行 K-means 聚类得到目标框长宽几何比例先验分布, 提高区域生成 (region proposal network, RPN) 网络训练速度和位置回归准确性。测试结果表明, 获得的汽车内部空间图像较为清晰, 算法可以实现对乘员数量的检测。经过改进, 网络的泛化能力得到增强, 单乘员检测的准确率达到 88.6%, 相比于改进前提高了 13.8%, 能够满足行业规定大于 80% 的要求。

关键词: 多光谱红外图像; Faster-RCNN; 全卷积; K-means 聚类; ROI-Align

中图分类号: TP391

文献标识码: A

文章编号: 1001-8891(2020)11-1103-08

Detection of Vehicle Crews Based on Modified Faster R-CNN

JIN Xin, HU Ying

(Dalian Maritime University, College of Marine Electrical Engineering, Dalian 116026, China)

Abstract: Existing methods for detecting the number of vehicle occupants in a high-occupancy vehicle (HOV) lane, using radar and infrared thermal imaging technology, exhibit low reliability and low accuracy. To address these limitations, a method for detecting the number of vehicle occupants based on multispectral infrared imaging and an improved Faster regions with convolutional neural networks (R-CNN) algorithm is proposed. The vehicle interior space image is obtained using a multispectral infrared imaging system, and the number of passengers is detected by a Faster R-CNN deep learning algorithm. The generalization ability of the network is enhanced using the full convolution network structure and multiscale feature prediction, and ROI-Align is used instead of ROI-Pooling. Through K-means clustering, the prior distribution of the geometric proportion of the length and width of the target frame is obtained, which improves the training speed and the accuracy of position regression of the region proposal network (RPN). The test results showed that the interior space image was clear, and the algorithm could detect the number of passengers. After its improvement, the generalization ability of the network was enhanced, and the accuracy of single occupant detection reached 88.6%, which was 13.8% higher than before its improvement. This meets the requirements of more than 80% of industry regulations.

Key words: multispectral infrared image, faster-RCNN, full convolution, K-means clustering, ROI-Align

0 引言

随着社会经济发展和生活质量的提高, 汽车数量增长而带来的交通拥堵问题已经成为目前亟待解决

收稿日期: 2020-02-19; 修订日期: 2020-09-02.

作者简介: 金鑫 (1996), 男, 硕士研究生, 主要研究方向为计算机视觉、深度学习、目标检测。E-mail: jin_xin@dlmu.edu.cn.

基金项目: 国家自然科学基金 (61973049)。

的问题。HOV 专用车道即高载客率多乘员车道，在规定的时段只对多乘员开放。设置 HOV 专用车道，可以在不增加交通建设成本的情况下，有效利用现有的公共资源提高交通运输的效率，是解决城市拥堵的一种有效措施^[1-2]。

目前 HOV 专用车道执法效率较低，针对汽车乘员数量检测领域的研究比较少，主要是因为汽车车窗贴过车膜后，普通相机很难获得其内部空间图像。当前可参考应用到车辆乘员数量检测的技术主要有雷达法和基于红外图像的目标检测法^[3]。Fadel Adib 等人^[4]提出一种基于射频体反射的多人室内目标定位检测方法，该研究使用 5 对频率在 5.46~7.25 GHz 的信号接收-发射装置，将信号调整为连续波信号后，将这 5 对收发装置的时延-频率图（Time delay-frequency diagram）进行叠加，可以得到目标相对准确的空间位置。该方法虽然满足车辆乘员数量检测的条件和要求，但在实际的检测环境中，路边的行人、行驶的车辆对雷达检测的精度影响较大。在红外图像检测技术方面，马也等人^[5]提出一种复杂背景下红外目标的检测算法。通过多权值高斯背景率除法对人体目标进行分割，然后对得到的候选区域采用融合边缘方向累加和特性的梯度方向直方图进行特征描述，通过对支持向量机进行训练来实现人体目标的判别，但在目标密集、目标遮挡等情况下检测精度还不能达到要求。以应用在成都 HOV 专用车道的红外热成像检测技术为例，只有当车窗摇下时才可获得驾驶室内部空间图像，大大降低了红外热成像检测的适用性，当汽车内部乘员较多时，乘员间相互遮挡会使目标的成像不规律，影响检测结果，红外图像质量的不确定性使得准确率难以保证。

随着红外技术与图像融合技术的发展^[6-8]，为设计适用于 HOV 车道执法的检测技术提供了新的思路。本文在前期设计了一套多光谱红外成像装置，采用主动照明技术，由光源系统主动发射多个波长的红外光，形成多光谱人工照明环境。相机内置光学分光系统，通过感光器件形成各个波段的图像，利用图像层叠技术将多张图像合成为一张多光谱红外图像，从而解决了车窗贴膜普通相机不能成像的问题，成像效果如图 1 所示。该视觉系统放置于车道一侧，与高架杆处卡口相机一同触发，卡口相机拍摄汽车的正面图像并识别车牌号码，两相机同时抓拍的照片便形成完整的执法证据。

通过对车内乘员数量的自动检测可以辅助完成 HOV 车道执法工作。在获得汽车内部空间的多光谱红外图像后，车内乘员数量可以通过目标检测算法来实现。深度学习领域出现了一批以 Faster

R-CNN^[9]、YOLO（You Only Look Once）^[10]、SSD（Single Shot MultiBox Detector）^[11]为代表的目标检测算法，与传统算法相比在准确率上有很大的提升。与后两种算法相比，Faster R-CNN 是一种典型的两步目标检测算法，两步法虽然在检测速度上慢了一些，但两步法网络的适用性强、准确率高、结构相对灵活，便于后期有针对性的调整。因此，本文在获得清晰汽车内部空间图像的基础上，以 Faster R-CNN 为基础，提出了基于多尺度特征预测方式、通过 K-means 聚类的方法学习目标框大小的几何先验知识来优化候选框的生成方式、通过用 ROI-Align 代替 ROI-Pooling 完成目标的特征映射，用以解决因模型泛化能力不强而带来的误报、漏报问题，以达到提高检测精度的目的。

1 Faster R-CNN 介绍

Faster R-CNN 网络将目标检测网络中特征提取、候选区域生成、目标分类、目标定位等 4 个步骤融合到一个深度网络中，实现了端到端的训练，算法结构如图 2 所示^[12]。从网络结构上看，所有任务统一到一个任务中完成。在执行顺序上，数据在经过特征提取后，首先经过 RPN 网络生成候选区域，接着判断候选区域内容是否为前景，如果内容是前景，则最后经过 ROI 池化后送入后续的检测进行目标分类和位置回归，否则放弃该候选框。同时当判定为前景后，RPN 网络的另一分支也会对候选框的大小和几何坐标进行调整。所以相比于 YOLO、SSD 等算法而言，Faster R-CNN 是一个两步法的端到端的训练，所以在实际表现中准确率往往会更高一些。

RPN 网络代替了传统的选择性搜索（selective search）方法^[12]，卷积运算使得候选框生成的方式更加科学，提升了候选框生成效率和准确性。ROI 池化操作实现了将不同尺寸的候选框映射成相同尺寸的特征图，可以有效地将 RPN 网络和目标检测网络的权值进行共享，避免了对候选区域进行重复的卷积运算，大大提升了网络的运算速度。

2 网络改进

Faster R-CNN 在 Pascal Voc、Coco 等数据集上实验效果很好，但当同一类别目标特征相差较大时，网络的泛化能力还有很大的提升空间^[13]，特别是在本文中，因为天气、环境、光照等因素使得乘员目标间特征差异较大，前期实验中仍然存在漏报、误报等问题。此外，本章在前期实验的基础上，对本

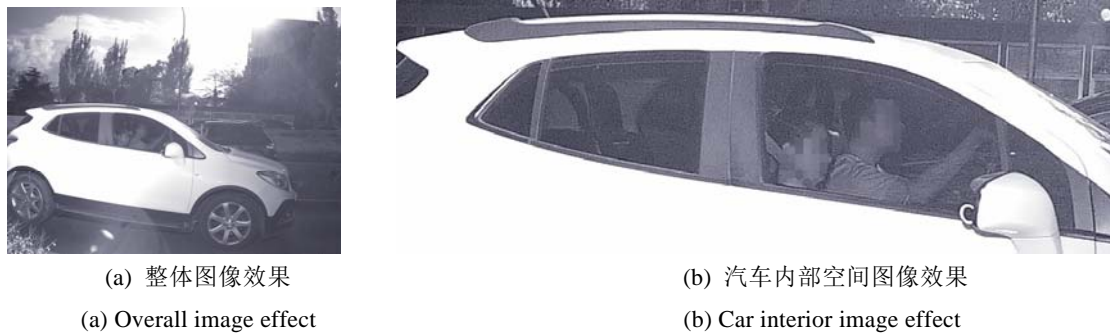


图1 多光谱红外成像效果

Fig.1 Multispectral visual imaging effect

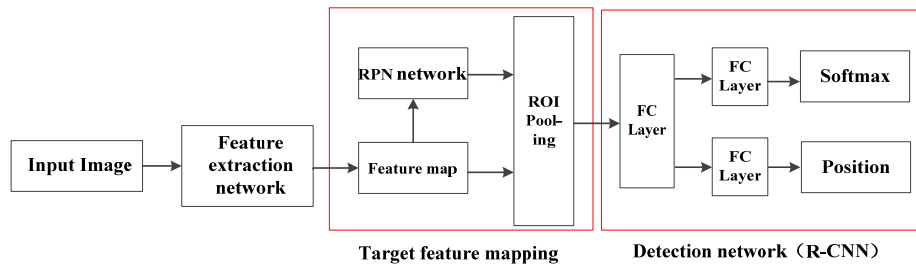


图2 Faster R-CNN 网络结构图

Fig.2 Faster R-CNN network structure diagram

文研究图像的目标检测做出分析：①应调整网络结构使网络泛化能力得到提高，减少漏报、误报情况的出现。②原始RPN网络中Anchor的设置是针对一般目标的，不适合本文的乘员目标检测。③RPN网络中将对应于原图上的ROI区域映射到特征图上，ROI-Pooling经历两次取整，最终提取的特征和原图的ROI不再对齐，导致目标的细节信息丢失。

2.1 全卷积网络结构

全卷积神经网络(Fully Convolutional Networks, FCN)^[14]是Jonathan Long等人在2015年提出用于语义分割的一种网络结构，已经被广泛地应用于计算机视觉研究领域。全卷积网络的3个核心思想：卷积化、上采样、跳跃结构，本文基于预训练的VGG-16^[15]网络结合上述3个方面做出改进。在原有VGG-16参数的基础上通过调整原有卷积核的步长代替池化层实现数据的降维，通过Loss值的反向传播来学习采样区域的最佳表达信息，使得图像更多的细节信息得以保留。用1×1的卷积核代替全连接层，通过控制卷积核的数量实现通道数大小的放缩，同时非线性激活函数的引入使得网络的非线性得到增强。

图3是本文采用的基于VGG-16特征提取网络结构，与改进前相比，本文采用了两种特征尺度输出，在两种特征尺度基础上分别连接RPN网络并分别进行目标特征映射、目标分类与位置回归，有效弥补

了单一尺度特征预测的局限性，使得网络在提高检测精度的同时提高网络对小目标的检测能力。本文将Conv5_3输出的特征进行上采样并与Conv4_3输出的特征进行融合，由于反卷积会随着网络深度增加出现训练不稳定、难收敛的问题，所以本文采用双线性插值的方法进行上采样。融合后特征的通道数发生了改变，所以再将融合后的特征通过1×1卷积降维到与Conv5_3相同的通道数。跳跃结构的特征融合可以有效缓解原始网络在数据降维时细节特征丢失严重的问题，同时深层特征图中包含了目标丰富的语义信息，浅层特征图中包含了目标的位置信息，二者融合将有效提高后续的目标分类与位置回归的特征基础。

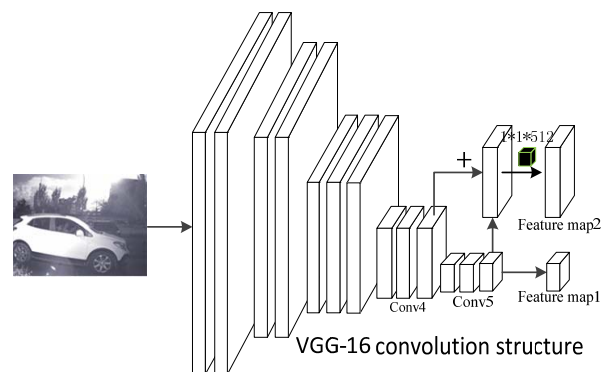


图3 基于VGG-16的特征提取网络

Fig.3 Feature extraction network based on VGG-16

2.2 K-means 聚类

RPN 网络使用了候选框滑动窗口生成机制，实现了多个候选区域的同时预测。RPN 网络结构如图4所示，图右侧为在每个滑动窗口所产生的 k 个候选框，由于待测目标的尺寸、宽高比的差异，往往需要设定多种几何尺度的候选框。在原始 RPN 网络中，候选框大小由3种尺度（128, 256, 512）、3种宽高比（1:1, 1:2, 2:1）组合而成，共9种，分别用于检测不同大小的目标。

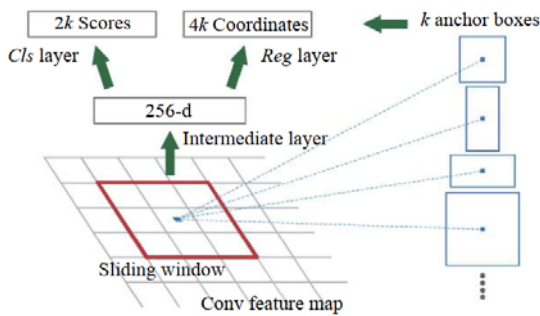


图4 RPN 网络结构

Fig.4 RPN network structure diagram

本文识别目标为乘员脸部，目标大小总体上差异并不大，原始尺寸的候选框并不适合本文应用，因此本文选取 $K=6$ ，避免生成多余无用尺寸的候选框，同时对 RPN 网络中生成候选框的方式作以改进。通过一个基于交并比的 K-means 聚类算法，对大量训练数据进行目标框几何大小的先验知识进行学习^[16]，得到适合本文两种尺度特征预测的 6 种尺寸的候选框。以此为参考来生成候选框，可以降低 RPN 网络训练的初始误差，提升网络训练的速度和目标定位精度。

在将数据集中所有边界框聚集在一起之后，从所有标注边界框中随机选取 K 个作为初始质心框，采用 K-means 聚类算法对 n 个标注框和 K 个初始质心框在宽高平面上进行聚类，其中第 l 个质心边界框到第 i 个标注边界框的距离 $d(\bar{B}_l - B_i)$ ，如(1)所示：

$$d(\bar{B}_l - B_i) = 1 - \text{IoU}(\bar{B}_l, B_i) \quad (1)$$

其中：

$$\text{IoU}(\bar{B}_l, B_i) = \frac{S(\bar{B}_l \cap B_i)}{S(\bar{B}_l \cup B_i)} \quad (2)$$

$S(\bar{B}_l \cap B_i)$ 与 $S(\bar{B}_l \cup B_i)$ 分别表示第 l 个质心框与第 i 个标注框的相交面积与相并面积，则公式(2)表示第 l 个质心框与第 i 个标注框面积的交并。通过公式(1)(2)定义的距离，可生成 K 个簇，分别求出每个簇

中所有边界框的宽中值与高中值来作为这一簇中新质心框的宽与高。不断重复上述过程，当每个簇的质心框的更新误差小于给定误差时，停止聚类。本文使用 8000 张数据集 15000 左右个目标进行候选框大小的聚类，聚类结果如图 5 所示， $c_0 \sim c_5$ 为最终获得按照从小到大顺序排列的 6 种候选框尺寸，横纵坐标分别代表候选框与图片宽高长度的比值。本文的特征提取网络有两种尺度特征图输出，小尺度特征图具有较大的感受野，适合大目标的预测，因此 $c_3 \sim c_5$ 为 feature map1 连接的 RPN 网络生成候选框的尺寸。大尺度特征图相对感受野较小，适合小目标的预测，所以 $c_0 \sim c_2$ 为 feature map2 连接的 RPN 网络生成候选框的尺寸。

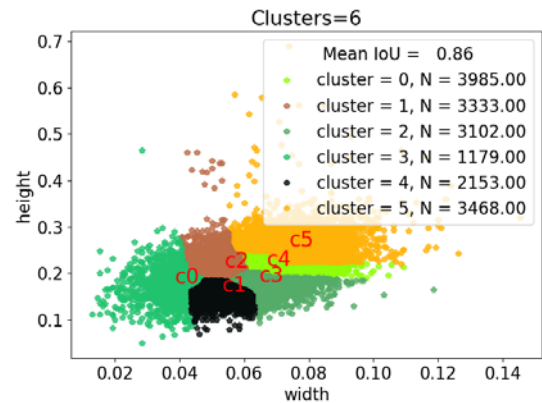


图5 聚类结果

Fig.5 Clustering results

2.3 ROI-Align

RPN 网络会产生大小不同的候选框，并在 ROI-Pooling 层中映射为固定大小的特征图，具体过程如图 6 所示。候选目标框首先映射到深层特征图获得该候选目标的特征图，为了方便后续目标分类与坐标回归网络的计算，还会继续映射成固定大小的输入尺寸，这两次量化过程中，会存在非整数倍的缩放。以将目标特征图映射成固定尺寸输入为例，在 ROI-Pooling 中采用的处理是浮点数取整，如图 6(a)所示，采样网格外的阴影部分便是取整后省略掉的特征信息，正是这一原因使得 Faster R-CNN 对小目标的检测能力不强，同时减少特征信息的丢失有利于提高网络的泛化能力^[17-18]。本文检测目标为乘员脸部，几何面积较小、目标特征差异较大，因此引用了 ROI-Align 来对此改进，提升网络的泛化能力和检测精度。参考文献 18 对于此部分的研究内容，ROI-Align 完成特征映射的过程如图 6(b)所示。

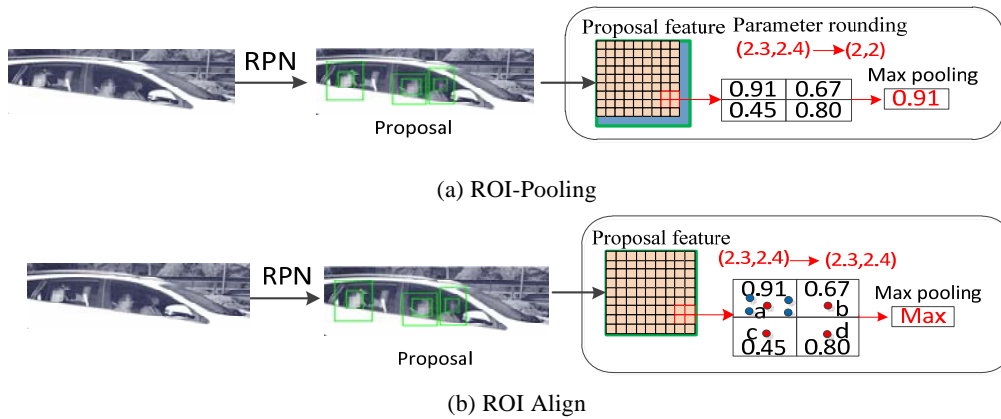


图6 ROI-Pooling 和 ROI Align 过程图

Fig.6 ROI-Pooling and ROI Alignment process diagram

ROI-Align 是在 Mask RCNN^[19]中使用以将候选目标的特征图映射产生固定大小的 feature map 时提出的，它增加了额外的网络分支用于语义分割任务，虽然目标检测的准确率得到了提升，但网络的检测速度却大大降低，因此本文只借鉴了它 ROI-Align 部分的改进。它与 ROI-Pooling 的区别是保留了非整数倍的缩放，对于缩放后坐标不能刚好为整数的候选框，最近邻插值法直接选择离目标点最近的整数点坐标去代替原始坐标。如图 6(b)所示，每一池化单元分为 4 等份，假设每一等份的中心点分别为 a 、 b 、 c 、 d 。

假设 a 点坐标为 $(i+u, j+v)$ ，其中 i, j 均为非负整数， (u, v) 为 $[0, 1]$ 区间的浮点数， a 点附近的 4 个圆点分别表示离 a 点最近且坐标为整数的点，它们的坐标分别为 (i, j) ， $(i+1, j)$ ， $(i, j+1)$ ， $(i+1, j+1)$ ， $f(i, j)$ 表示坐标为 $(i+1, j)$ 的点的像素值。则 a 点的像素值的计算公式可以表示为：

$$f(i+u, j+v) = (1-u) \times (1-v) \times f(i, j) + (1-u) \times v \times f(i, j+1) + u \times (1-v) \times f(i+1, j) + u \times v \times f(i+1, j+1) \quad (3)$$

其他 3 个点 b ， c ， d 处的像素值计算类似。

分别求得各个等分中的像素值后再进行最大化采样，使得候选区域的特征信息全部都得以利用，增大了特征图单元的感受野。大量的实验表明，在检测大目标时，两者方案差别不大；当检测小目标时，RoI-Align 更精准。

3 实验分析

实验部分主要从 RPN 网络的训练曲线、改进前后网络泛化能力和识别准确率、检测速度等方面作

以分析，同时随着乘员人数的增多，检测过程中漏报和误报的概率增大，所以将不同人数下的准确率作以统计对比。本文实验的硬件环境为 Intel i7-7600k 处理器，12 GB 内存，Nvidia GeForce GTX 1080Ti 显卡。软件上采用 PyCharm 编译平台，TensorFlow 深度学习框架，CUDA 版本为 9.0。

3.1 RPN 网络训练曲线

本文特征提取网络输出的两种尺度特征分别连接了 RPN 网络，并分别进行训练，对于两个网络生成的候选框使用非极大值抑制的方法滤除相同目标框。本文采用学习率动态更替的训练方式，batch_size 大小为 128，对原始的 RPN 网络及本文的两种特征尺度为预测基础的 RPN 网络绘制了位置回归 Loss 曲线，如图 7 所示。从图中可以看出，改进后两个 RPN 网络的初始误差与改进前相比缩小了将近 2 倍，说明聚类后得到的候选框大小与实际的候选框大小更加接近。改进后的两个 RPN 网络在迭代六千次以后趋于收敛，网络收敛速度有明显提升，并且在训练过程中，震荡较小，Loss 值下降明显。

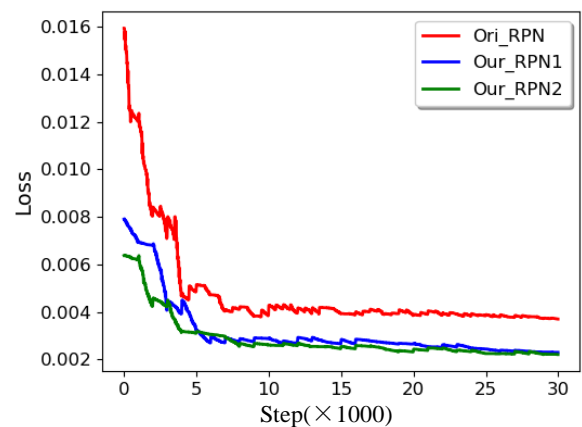


图7 RPN loss 训练曲线对比

Fig.7 Comparison of RPN loss training curves
为了直观验证 RPN 网络改进的效果，本文对改

进前后的网络在 1000 张数据集上测试生成候选框的准确率，统计结果如表 1 所示。同时为了保证测试结果的可靠性，本文也对网络改进（ 1×1 卷积代替全连接等）部分进行试验。

表 1 RPN 改进前后对比

Table 1 RPN Comparison before and after improvement

Method	Regions	AP/%	Times/ms
Ori_RPN	2000	69.00	235
Our_RPN1	2000	74.43	240
Our_RPN2	2000	75.52	241
Our_RPN1+Net-Improved	2000	74.91	237
Our_RPN2+Net-Improved	2000	76.17	237

从表中分步试验的结果可以看出，本文改进后的两个 RPN 网络相比于改进前分别提高了 5.91% 和 7.17% 的精度，网络结构的改进分别使得网络提高了 0.48% 和 0.65% 的精度，这说明改进后的候选框的质量更高。此外从图 7 和表 1 也可以看到，RPN2 网络无论从位置回归 Loss 曲线的训练过程还是候选框的生成质量都具有很好的表现，说明特征融合后提高了特征质量，特征尺度变大后更有利于本文人脸目标的检测。

3.2 检测结果对比



(a) 改进前识别效果
(a) Recognition effect before improvement



(b) 改进后识别效果
(b) Improved recognition effect

图 8 改进前后检测效果对比

首先测试模型的泛化能力，选取了一些典型的乘员目标特征差异较大的数据进行测试，特征差异主要表现在乘员脸部曝光程度、明暗差异等等。改进前后的检测效果如图 8 所示。

由图 8 中可以看出，即使是当曝光程度较高导致面部信息缺失严重时、人脸纹理特征不明显时，改进后的网络也可以对乘员面部进行准确标记，并保持较高的识别分数。当车膜较厚时，相机的透光率不足导致整体成像灰暗，主背景层次不明显，改进后的网络也可以准确标记。此外，抓拍过程的随机性与乘员目标状态的不确定性，乘员面部容易被车窗中间的壁柱遮挡，并且乘员目标的帽子、口罩等因素都会对脸部发生遮挡，从实验中可以看出改进后的网络可以实现对存在遮挡的目标准确判别。从实际检测效果来看，在可能存在漏报的几种情况下，改进后的网络的可以很好的完成检测任务，同时也能避免对靠枕等非检测目标的误报。

同时为了量化验证本文的改进效果，选取了 2500 张图片，每类乘员数量（1~5）各 500 张，将改进后网络和目前主流的 YOLOv3^[20]、Mask R-CNN 等检测算法分别进行实验，记录各类乘员数量下检测的准确率和检测速度，统计结果如表 2 所示。

Fig.8 Comparison of detection effect before and after improvement

表2 不同算法检测效果对比

Table 2 Comparison of detection results of different algorithms

Network	Accuracy/%					Fps
	1	2	3	4	5	
Faster R-cnn	74.8	72.6	69.8	65.4	58.8	17
Ours	88.6	86.2	83.0	78.8	73.2	15
YOLOv3	68.6	65.2	61.8	57.4	51.2	30
Mask R-CNN	82.4	78.8	76.2	72.4	66.8	5

从表2的数据可以看出，在准确率上，本文的模型比其它模型表现更好，原因在于本文结合了其它几种模型的优点，准确率提高的同时增强了模型的泛化能力，检测的速度比改进前略微降低，但并不影响检测的实时性。相比于改进前的网络，单乘员检测的准确率提高了13.8%，多乘员检测的准确率也随之有明显提高，基本能保持在70%以上。YOLOv3是目前较为优秀的目标检测网络，但其算法计算的基础是针对正方形的图像输入，图像识别前会经预处理裁减掉大部分的背景信息，得到的驾驶室图像多为长方形，需要压缩成正方形后再输入网络，压缩后图像质量降低，网络预测的准确率不佳。Mask R-CNN的实验结果表明实例分割的效果优于一般的目标检测算法，但准确率略低于本文，说明虽然都采用了ROI-Align，但本文网络结构的改进提升了检测精度，此外，单目标的准确率也可以达到行业内的规定标准，但是语义分割网络的加入使得网络整体的计算量增大，检测速度只有本文的1/3。从实验数据可以总结得知，随着人数的增加，图像成像的不确定因素增大，检测的准确率随检测人数增加呈下降趋势，当检测人数不超过3人时，本文可以基本满足行业规定标准。

4 结语

结合已获得的多光谱红外图像，本文以Faster R-CNN网络模型为基础，通过对网络结构、候选框生成方式、目标特征图的池化方式等方面做出改进，网络模型的泛化能力得到明显的提高，对图像中目标曝光度过高、目标纹理不清晰、面部遮挡严重等可能存在误报、漏报的情况也可以准确识别。相比于改进之前，本文单乘员的检测准确率达到88.6%，提升了13.8%，满足了行业内的规定标准。但随着乘员人数的增加，检测的不确定因素也随之增加，多乘员之间容易发生相互遮挡，当检测人数大于3时，准确率还不能达到行业内的规定标准，也就不能够满足乘员超载情况的检测要求。此外，在检测速度

上相比于YOLOv3还有很大差距，提高各类数量乘员检测精度的同时提升检测速度将作为以后主要的研究方向。

参考文献：

- [1] 曾炎盛. HOV 车道实施效果评价体系研究——以深圳市为例[J]. 交通科技与经济, 2019, 21(5): 39-43.
ZENG Yansheng. Study on the evaluation system of the implementation effect of HOV Lane -- Taking Shenzhen as an example[J]. *Transportation Technology and Economy*, 2019, 21(5): 39-43.
- [2] 陈鲁峰, 叶鹏飞, 余佳欢, 等. HOV 在城市交通管理中的应用可行性分析[J]. 交通企业管理, 2016, 31(4): 38-41.
CHEN Lufeng, YE Pengfei, YU Jiahuan, et al. Feasibility analysis of HOV application in urban traffic management [J]. *Traffic Enterprise Management*, 2016, 31(4): 38-41.
- [3] 程建梅. HOV 专用车道管理思路探索[J]. 行业管理, 2018(4): 34-35.
CHENG Jianmei. Exploration of management ideas of HOV special lane[J]. *Industry Management*, 2018(4): 34-35.
- [4] Fadel Adib, Zachary Kabelac, Dina Katabi. Multi-person Localization via RF Body Reflections[C]//*Proceedings of the 12th USENIX Conference on Networked Systems Design and Implementation (NSDI'15)*, 2015: 279-292.
- [5] 马也, 常青, 胡谋法. 复杂背景下红外人体目标检测算法研究[J]. 红外技术, 2017, 39(11): 1038-1053.
MA Ye, CHANG Qing, HU Moufa. Research on infrared human target detection algorithm in complex background[J]. *Infrared Technology*, 2017, 39(11): 1038-1053.
- [6] 马翰飞, 范海震, 李强, 等. 基于多光谱融合图像的飞机导航系统设计[J]. 电子设计工程, 2019, 27(24): 161-166.
MA Hanfei, FAN Haizhen, LI Qiang, et al. Design of aircraft navigation system based on multispectral fusion image[J]. *Electronic Design Engineering*, 2019, 27(24): 161-166.
- [7] 李雪欣, 马保东, 张嵩, 等. 融合多光谱与 SAR 影像的地物分类研究[J]. 测绘与空间地理信息, 2019, 42(12): 55-58.
LI Xuexin, MA Baodong, ZHANG Song, et al. Study on the classification of ground features using multispectral and SAR images[J]. *Mapping and Spatial Geographic Information*, 2019, 42(12): 55-58.
- [8] 赵庆展, 刘伟, 尹小君, 等. 基于无人机多光谱影像特征的最佳波段组合研究[J]. 农业机械学报, 2016(3): 242-248.
ZHAO Qingzhan, LIU Wei, YIN Xiaojun, et al. Research on optimal band combination based on multi spectral image characteristics of

- UAV[J]. *Journal of Agricultural Machinery*, 2016(3): 242-248.
- [9] REN Shaoqing, HE Kaiming, Girshick Ross, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, **39**(6): 1137-1149.
- [10] 郑国书. 基于深度学习 SSD 模型的视频室内人数统计[J]. *工业控制计算机*, 2017, **30**(11): 48-50.
- ZHENG Guoshu. Statistics of indoor video population based on deep learning SSD model[J]. *Industrial Control Computer*, 2017, **30**(11): 48-50.
- [11] Oseph Redmon, Santosh Diwala, Ross Girshick, et al. You Only Look Once: Unified, Real-Time Object Detection[C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779-788.
- [12] 徐代, 岳章, 杨文霞, 等. 基于改进的三向流 Faster R-CNN 篡改图像识别[J/OL]. *计算机应用*, <http://www.joca.cn/CN/10.11772/j.issn.1001-9081.2019081515>.
- XU Dai, YUE Zhang, YANG Wenxia, et al. Tamper image recognition based on improved three-way Fasert R-CNN[J/OL]. *Journal of Computer Applications*, <http://www.joca.cn/CN/10.11772/j.issn.1001-9081.2019081515>.
- [13] 孙雄峰, 林浒, 王诗宇, 等. 基于改进 Faster RCNN 的工业机器人分拣系统[J]. *计算机系统应用*, 2019, **28**(9): 258-263
- SUN Xiongfeng, LIN Hu, WANG Shiyu, et al. Industrial robot sorting system based on improved fast RCNN[J]. *Computer System Application*, 2019, **28**(9): 258-263.
- [14] LONG J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//*Computer vision and Pattern Recognition*, 2015: 3431-3440.
- [15] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. *Computer Science*, 2014, **42**(5): 324-328.
- [16] 张琦, 丁新涛, 王万军, 等. 基于 Faster R-cnn 的交通目标检测方法[J]. *皖西学院学报*, 2019, **35**(5): 50-55.
- ZHANG Qi, DING Xintao, WANG Wanjun, et al. Traffic target detection method based on fast R-cnn[J]. *Journal of West Anhui University*, 2019, **35**(5): 50-55.
- [17] 陈泽, 叶学义, 钱丁炜, 等. 基于改进的 Faster R-CNN 小尺度行人检测[J]. *计算机工程*, 2020, **46**(9): 226-232, 241.
- CHEN Ze, YE Xueyi, QIAN Dingwei, et al. Small scale pedestrian detection based on improved Faster R-CNN[J]. *Computer Engineering*, 2020, **46**(9): 226-232, 241.
- [18] 侯志强, 刘晓义, 余旺盛, 等. 基于双阈值-非极大值抑制的 Faster R-CNN 改进算法[J]. *光电工程*, 2019, **46**(12): 1-11.
- HOU Zhiqiang, LIU Xiaoyi, YU Wangsheng, et al. Improved Faster R-CNN algorithm based on double threshold-non-maximum suppression[J]. *Optoelectronic Engineering*, 2019, **46**(12): 1-11.
- [19] HE K M, Gkioxari G, Dollár P, et al. Mask R-CNN[C]//*Proceedings of IEEE International Conference on Computer Vision*, 2017: 1026-1032.
- [20] Redmon J, Farhadi A. YOLOv3: An incremental improvement[J/OL]. *arXiv*: 1804.02767, 2018.