文章编号:1001-5078(2021)01-0052-07

·红外技术及应用·

基于改进深度网络的车辆乘员数量检测研究

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:车辆乘员数量检测研究对推进 HOV 车道建设具有重要意义,本文以 Faster RCNN 网络模型为基础,结合多光谱红外系统获得的汽车驾驶室图像来展开研究。现有数据中因过曝、曝光不足等问题导致图像中目标特征差异大,网络检测的准确率不高,对此采用变形结构的卷积计算提高特征单元的感受野和目标边缘信息的表述能力,采用变形的 ROI-Pooling 来增强 ROI 区域特征映射后的特征表述,增强网络的泛化能力。针对在多乘员情况下,因乘员间遮挡导致的错检、漏检等问题引入 KL 损失,同时采用 Soft-NMS 与方差投票相结合的方式来改善 NMS 滤除重复目标框过程的合理性,提高了位置回归的合理性和重叠目标的预测能力,整体检测的准确率得到提高。实验结果表明,本文网络在不同数量下算法检测的准确率得到提高,基本可以满足行业规定大于 80 % 的要求。

关键词:HOV 车道;Faster RCNN;变形卷积;KL 损失;Soft-NMS
中图分类号:TN219;TP391.41　　**文献标识码:**A　　**DOI:**10.3969/j.issn.1001-5078.2021.01.009

Research on vehicle occupant number detection based on improved depth network

JIN Xin, HU Ying
(College of Marine Electrical Engineering, Dalian Maritime University, Dalian 116026, China)

Abstract: The research on vehicle occupant number detection is of great significance to promote the construction of HOV lanes. Based on the Faster RCNN network model, this paper combines the multi-spectral infrared system to obtain the image of the car cab. In the existing data, due to overexposure and underexposure problems, the target feature differences in the image are large, and the accuracy of network detection is not high. For this, the convolution calculation of the deformed structure is used to improve the ability to express the receptive field of the feature unit and the target edge information. Use deformed ROI-Pooling to enhance the feature expression after ROI feature mapping and enhance the generalization ability of the network. In the case of multiple occupants, the KL loss is introduced due to misdetection and missed detection caused by occlusion among the occupants. At the same time, the combination of Soft-NMS and variance voting is used to improve the rationality of the NMS filtering process of repeated target frames and improve in addition to the rationality of position regression and the prediction ability of overlapping targets, the overall detection accuracy has been improved. The experimental results show that the accuracy of the algorithm detection of the network under different numbers in this paper has been improved, and can basically meet the requirements of the industry regulations greater than 80 %.

Keywords: HOV lane; Faster RCNN; deformed convolution; KL loss; Soft-NMS

基金项目:国家自然科学基金项目(No. 61973049)资助。
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
收稿日期:2020-03-25;**修订日期:**2020-05-18

1 引 言

HOV(High-Occupancy Vehicle)车道又称多乘员车道、高载客率车道,指规定时间对多乘员车辆开放的车道,可以有效缓解交通拥堵问题,提高车道的出行效率[1]。但是目前国内 HOV 车道的执法主要通过交警现场人工排检来完成,执法效率较低,因此设计一套自动化的车辆乘员数量检测系统就显得十分必要。本文采用光谱红外技术与图像融合技术设计了一套多光谱红外成像系统,可以获得较为清晰的驾驶室内部空间图像,成像效果如图 1 所示。车内乘员数量可以通过目标检测算法来实现,在前期实验过程中发现,将乘员面部作为检测目标时检测效果相对较好,但目标尺寸较小,当汽车后排人员较多时容易发生遮挡,在预测过程中容易在 NMS 过程当做相同目标被滤除掉,同时受天气、曝光程度、光照等因素的影响,目标在成像效果上特征差异较大,因此在检测结果中经常出现漏报和误报。



(a)整体图像效果 (b)汽车内部空间图像效果

图 1 多光谱红外成像效果

Fig. 1 Multi-spectral infrared imaging effect

近年来,国内外学者针外弱小多目标检测问题进行了深入研究,同时随着深度学习技术的成熟发展,红外目标检测精度已经有了很大的提升[2]。张骏等人[3]提出了一种基于 U-Net[4] 网络的消防红外图像的人体检测算法,通过该算法解决了人体姿态复杂、多障碍物遮蔽等情况下的目标识别,但是在多目标场景下检测效果还不够理想。李承昊等人[5]提出一种可变锚框候选区域网络的目标检测方法,通过注意力机制引导模块对 Faster RCNN[6] 的 RPN 进行改进,并且引入了一种置信因子的 NMS 算法,在增强候选框生成合理性的同时提高了检测精度,但是对于目标位置回归的精度并没有起到很好的引导作用,多目标聚集下效果仍然不够理想。

本文在总结现有研究的基础上,以 Faster RCNN 为基础,通过使用变形结构的卷积计算和 ROI-Pooling 来提高网络的泛化能力[7],引入 KL 损失函数[8]来改进目标框位置回归过程,同时采用了 Soft-NMS 与方差投票相结合的方式提高 NMS 过程的合理性,提高目标框定位精度,避免相邻目标间因目标框重叠度大而产生误检,多乘员检测的准确率得到明显提高。

2 Faster RCNN 介绍

Faster RCNN 网络将目标检测网络中特征提取、候选区域生成、目标分类、位置回归等四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[9]。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-Pooling 将其映射成固定尺寸的特征图进行目标分类和位置回归,否则放弃该候选框,RPN 网络也会对候选框的几何坐标进行调整。所以相比于 YOLO、SSD 等算法而言,Faster RCNN 是一个两步法的端到端训练,所以在实际表现中准确率会更高一些。

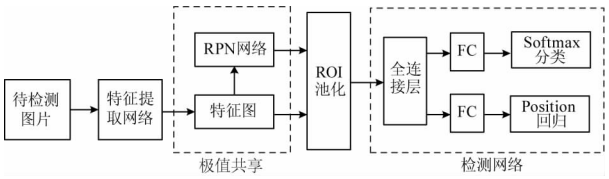


图 2 Faster RCNN 网络结构图

Fig. 2 Faster RCNN network structure

RPN 网络通过卷积运算使得候选框生成的方式更加科学,可以有效地的将 RPN 网络和特征提取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。在 RPN 网络的训练上,一般将与任意标注框的 IOU 值最大的或者与标注框的 IOU 值大于 0.7 的候选框分配正样本,为 IOU 值低于 0.3 的候选框分配负样本,RPN 的损失包括分类损失和位置回归损失,函数定义为:

$$L(\{p_i\},\{t_i\}) = \frac{1}{n_{cls}}\sum_i L_{cls}(p_i,p_i^*) + \lambda\frac{1}{n_{reg}}\sum_i p_i^* L_{reg}(t_i,t_i^*) \quad (1)$$

3 模型的改进

3.1 变形卷积

卷积神经网络(CNN)是一种重要的深度学习计算,在图像分类、目标检测、语义分割等方面都有着很成功的应用。变形卷积是指通过将可学习的 2D 偏移参数添加到标准卷积中的网格采样位置,实现了采样网格的自由变形,使得根据目标尺寸、形状

来自适应地确定卷积区域。卷积变形前后的效果如图 3 所示。从图中可以看出,变形后的卷积计算避免了对背景区域的无用计算,计算效率得到提高,特征单元的感受野也随之增大。



(a) 卷积变形前 (b) 卷积变形后

图 3 采样网格的变换

Fig. 3 Transformation of sampling grid

卷积过程主要分为两个步骤:1)在输入特征图 x 上使用规则网格 R 进行采样;2)对采样值按权重加权求和。只要我们改变采样网格的规则性,便可实现变形卷积。规则网格 R 定义了感受野大小和扩张尺度。比如定义一个 3×3 的卷积核,扩张大小为 1。

$$R=\{(-1,-1),(-1,0),\cdots,(0,1),(1,1)\}$$

对于输出特征图 y 上每个位置 p_0 有:

$$y(p_0)=\sum_{p_n\in R}w(p_n)\cdot X(p_0+p_n) \tag{2}$$

其中,p_n 是网格 R 中的位置。

在变形卷积中,标准的格子 R 用 $\{\Delta p_n\mid n=1,\cdots,N\}$ 增加了偏移量,其中 $N=|R|$,则此时公式(2)变成:

$$y(p_0)=\sum_{p_n\in R}w(p_n)\cdot x(p_0+p_n+\Delta p_n) \tag{3}$$

现在采样是在不规则、有偏移位置($p_n+\Delta p_n$)上进行的。由于偏移量 Δp_n 通常是小数,则公式(3)通过双线性差值变成:

$$x(p)=\sum_q G(q,p)\cdot x(q) \tag{4}$$

其中,p 表示任意小数位置($p=p_0+p_n+\Delta p_n$);q 表示特征图 x 中所有整体空间位置;G 代表双线性插值核。通过双线性插值变换便可以获取到变形后每一个采样网格单元对应的像素点。G 是二维的,可表示成:

$$G(q,p)=g(q_x,p_x)\cdot g(q_y,p_y) \tag{5}$$

其中,$g(a,b)=\max(0,1-|a-b|)$。

如图 4 所示,某一通道数为 N 的特征图先经过一层卷积获得偏移量,卷积层的输出与输入的特征图具有相同的尺寸,但通道维数为 $2N$,对应于 x、y 两个方向的偏移量。在训练期间,提取特征的卷积核与获得偏移量的卷积核参数同时学习。

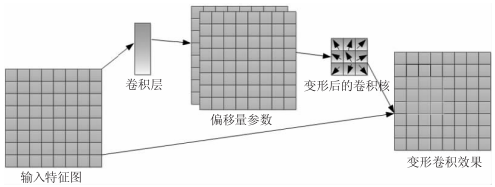


图 4 3×3 的变形卷积过程

Fig. 4 3×3 deformation convolution process

3.2 变形的 ROI-Pooling

ROI-Pooling 是 RPN 网络的中间环节,它可以将任意大小的候选目标区域缩放成固定尺寸特征。数据降维的池化过程会导致细节特征丢失,同时在完成候选目标区域和深层特征的映射时还会存在非整数倍的取整,使得采样网格是非均匀划分的,也会造成一定的细节特征丢失,对小目标的检测精度影响很大[10],对此本文展开了深入的研究。受变形卷积的启发,本文设计了一种变形的 ROI-Pooling,过程与之相似,唯一有区别的是通过全连接层求取参数,这样是为了增强参数求取的针对性,将其与变形卷积学习过程区别开来,偏移量表示为 $\{\Delta p_{ij}\mid 0\leqslant i,j<k\}$。

变形池化的过程如图 5 所示,偏移的采样方式使得可以将采样区域重点关注在目标上,使得完成映射的特征图尽可能多的保留目标特征,减少缩放过程的信息丢失,提高了位置回归和目标分类的特征质量。假设给定输入特征图 x 和大小为 $w\times h$ 的 RoI,并且左上角为 p_0,ROI-Pooling 将 RoI 划分为 $k\times k$(k 为自由参数)个 bin(粗实线框),并输出 $k\times k$ 特征图 y。对于 y 的特征值:

$$y(i,j)=f_{\text{maxpool}}x(p_0+p)p\in\text{bin}(i,j) \tag{6}$$

其中,p 代表采样区域,最终取得的 y 值为该采样区域的最大值。变形池化可以表示:

$$y(i,j)=f_{\text{maxpool}}x(p_0+p+\Delta p_{ij})p\in\text{bin}(i,j) \tag{7}$$

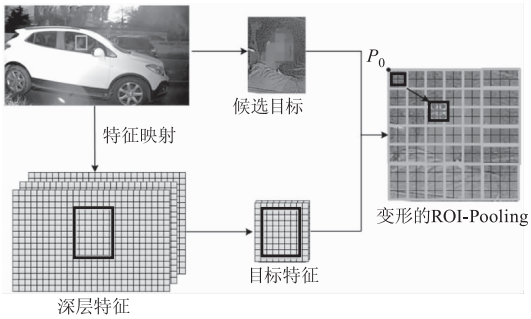


图 5 变形池化过程

Fig. 5 Deformed pooling process

候选区域完成映射得到目标特征后，紧接着将其均分成 $k \times k$ 的 bin(粗实线框)，第 (i,j) 个 bin 的几何尺度可以表示为 $\left| i\frac{w}{k} \right| \leqslant p_x < \left| (i+1)\frac{w}{k} \right|$，则变形后的可表示为 $\left| (i+\Delta i)\frac{w}{k} \right| \leqslant p_x < \left| (i+\Delta i+1)\frac{w}{k} \right|$，变形后的 bin 为图中方框区域，由于 Δp_{ij} 为小数，导致其不会与原特征图的网格重合，因此新 bin 中的特征值要通过双线性插值的方法求取。

4 损失函数改进

在本节首先介绍了修改后边界框的预测方式。然后提出了 KL Loss 用于训练检测网络的定位置信度。最后，采用了 Soft-NMS 与方差投票结合的方式提高目标框回归的准确性。

4.1 函数计算基础

本文与原 Faster RCNN 网络中位置预测方式不同，为了探索学习到的坐标方差与坐标位置的内在关系，本文在直接预测候选框坐标的同时预测和估计坐标预测的置信度来提升位置回归过程的准确性。本文采取预测其概率分布的形式，尽管分布可能很复杂，例如多元高斯分布或高斯混合分布，但在本文中，由于候选框的各个坐标相互独立，可以将其数学模型表示为：

$$P_e(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x_e)^2}{2\sigma^2}} \tag{8}$$

其中，x_e 是预测的边框位置。标准差 σ 代表预测的不确定度，当 $\sigma \to 0$ 时，这意味着我们的网络对目标框位置预测的置信度很高，它是在原网络的基础上添加一个全连接层分支产生的，参考文献[8]，可表示为如图 6 所示。

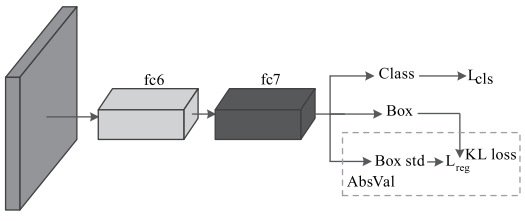


图 6 改进后的网络结构

Fig. 6 The improved network structure

同样的，我们也需要为数据中真实目标框的概率分布建立一个数学函数，按照之前的分析，将式(8)取 $\sigma = 0$ 便转换为狄拉克三角函数；

$$P_D(x) = \delta(x - x_g) \tag{9}$$

其中，x_g 代表标签中真实的坐标值。

4.2 KL 损失函数

KL 又称 KL 散度、相对熵，是用来描述相同时间空间里某一概率分布像另一概率分布的拟合情况，当 KL 散度越小，两个概率分布的差异也就越小。在本文中为了捕捉目标框预测的不确定性，本文将高斯分布函数向狄拉克三角函数拟合的 KL 散度作为损失函数，通过训练使得 KL 散度最小。

用 L_{reg} 记为目标框的误差函数：

$$\begin{aligned} L_{reg} &= D_{KL}(P_D(x) \parallel P_e(x)) \\ &= \int P_D(x)\log P_D(x)dx - \int P_D(x)\log P_e(x)dx \\ &= \frac{(x_g - x_e)^2}{2\sigma^2} + \frac{\log(\sigma^2)}{2} + \frac{\log(2\pi)}{2} - \\ &\quad H(P_D(x)) \end{aligned} \tag{10}$$

$\frac{\log(2\pi)}{2}$ 和 $H(P_D(x))$ 为常量，因此：

$$L_{reg} \propto \frac{(x_g - x_e)^2}{2\sigma^2} + \frac{\log(\sigma^2)}{2} \tag{11}$$

因为 σ 在分母中，会在训练时存在梯度爆炸的现象。为了避免梯度爆炸，我们的网络实际上采用 $\alpha = \log(\sigma^2)$ 的方式，而不是直接预测 σ，对于 $|x_g - x_e| > 1$，采用类似 Smooth L1 损失函数的形式，表示为：

$$\begin{cases} L_{reg} \propto \frac{e^{-\alpha}}{2}(x_g - x_e)^2 + \frac{1}{2}\alpha, \text{otherwise} \\ L_{reg} \propto e^{-\alpha}\left(|x_g - x_e| - \frac{1}{2} \right) + \frac{1}{2}\alpha, |x_g - x_e| > 1 \end{cases} \tag{12}$$

4.3 NMS 改进

本小节采用了将 Soft-NMS 与方差投票结合的方式对目标框筛选、目标定位等过程做出改进，前者将 NMS 中 IOU 值较高的预测框重新赋值一个较低的分数，而不是直接滤除掉，避免删除堆叠目标。由于线性加权方式不是连续函数，这样会使得结果中分数出现间断，因此采用高斯加权的方式：

$$s_i = s_i e^{-\frac{IOU(M,b_i)^2}{\sigma}}, \forall b_i \notin D \tag{13}$$

式中，s_i 为第 i 个目标框的分类得分；M 表示为当前选定得分最高的目标框；b_i 为当前与 M 进行分值抑制计算的目标框；D 为分值抑制计算完成的候选框集合。

方差投票是指根据学习到的有重叠边界框的方

差对候选框位置进行权重分配，受 Soft-NMS 的启发，我们为较接近且不确定性较低的候选框分配较高的权重。设 x 为坐标，x_i 为第 i 个候选框的坐标。新坐标的计算如下：

$$p_i = e^{\frac{-(1-\mathrm{IOU}(b_i,b))^2}{\sigma_t}}$$

$$x = \frac{\sum_i p_i x_i/\sigma_{x_i}^2}{\sum_i p_i/\sigma_{x_i}^2} \quad (14)$$

$$\text{subject to } \mathrm{IOU}(b_i,b) > 0$$

其中，σ_t 是方差投票过程可手动调整的参数。在投票期间，两种类型的预测框将具有较低的权重：①已学习方差较高的候选框；②与重新计算坐标的预测框交并比较小的候选框。类别分数不参与投票，分数较低的预测框可能具有较高的位置置信度。

5 实验分析

实验部分首先从 RPN 网络的位置 loss 回归曲线、改进前后网络泛化能力、检测速度等方面作以分析，然后重点在多乘员、目标存在遮挡的情况下测试检测效果，并进行准确率统计做出定量分析，最后将其和改进前、其他主流的目标检测网络进行对比实验，并对实验结果进行总结。实验平台配置及相关训练参数如表 1 所示。

表 1 实验平台配置及相关训练参数

Tab. 1 Experimental platform configuration and related training parameters

项目	名称/参数
GPU	Nvidia GeForce GTX 1080Ti
CPU	Intel i7 - 7600k
RAM	12 GB
学习框架	Tensorflow
CUDA	9.0
训练数据量	10000 张
Batch_size	128
Learning_rate	0.001

5.1 RPN 网络训练过程

RPN 网络的训练效果将直接影响目标检测的准确率，同时为了验证损失函数改进的可行性，本文绘制了 RPN 网络改进前后的位置 loss 回归曲线，如图 7 所示。从图中可以看出，改进后的 RPN 网络在经过六千次迭代以后开始平稳收敛，训练过程相对平稳，证明损失函数改进可行。此外改进后网络训练误差的初始值明显降低，收敛速度更快，loss 值收敛更加理想，这说明网络结构改进对网络整体性能的提升作用很大。

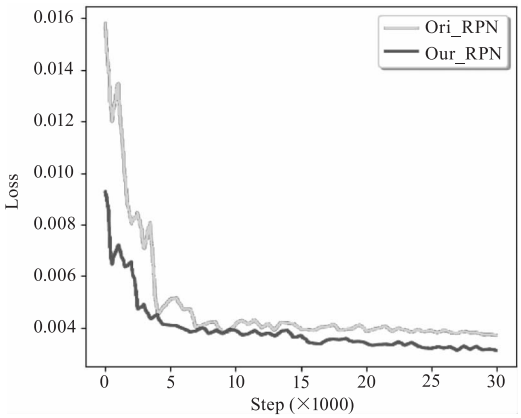


图 7 RPN loss 训练曲线对比

Fig. 7 RPN loss training curve comparison

5.2 NMS 改进实验

NMS 改进主要是为了使网络能够更好应对在 4 + 数量乘员情况下，因目标遮挡而造成的误报问题。一部分原因是目标遮挡会造成目标有效识别面积变小，还有一部分原因是原 NMS 中固定的 IOU 阈值很难准确适应重复目标的滤除需要。如图 8 所示，当阈值取为 0.3 时，原网络由于位置预测的不够准确使得重叠目的交并比过大，造成了目标漏报，当阈值增加到 0.4 时，又会造成重复目标框滤除效果不理想，造成误报。



(a)IOU=0.3



(b)IOU=0.4

图 8 原网络测试效果

Fig. 8 The original network test effect

由于 IOU 阈值偏大会影响重复目标框的滤除效果，因此选取 IOU = 0.3。图 9 的实验结果表明，本文算法可以满足重叠目标的预测，目标框位置预测的精度更高，加入方差投票改进后，目标框标记的视觉效果更好，重叠区域变小，提高了在执法时人工复检过程中判断证据的可靠性。



(a)Soft-NMS



(b)Soft NMS+方差投票

图 9 NMS 改进效果(IOU = 0.3)

Fig. 9 NMS improvement effect(IOU = 0.3)

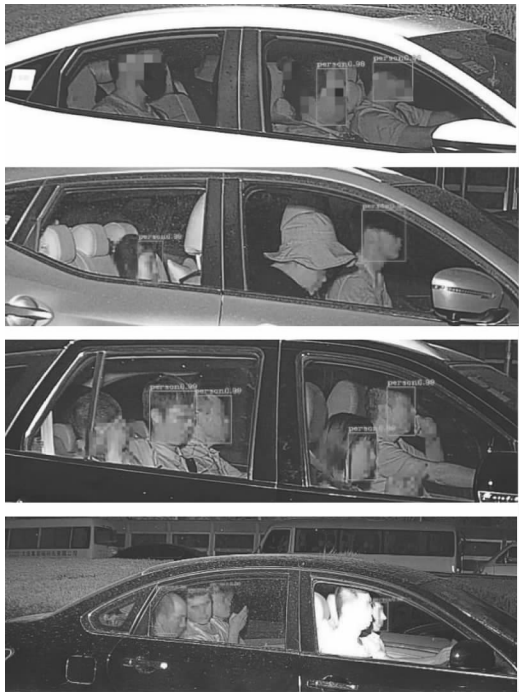
5.3 泛化能力测试

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、有口罩、手掌、帽子遮挡等等。改进前后的检测效果如图 10 所示。

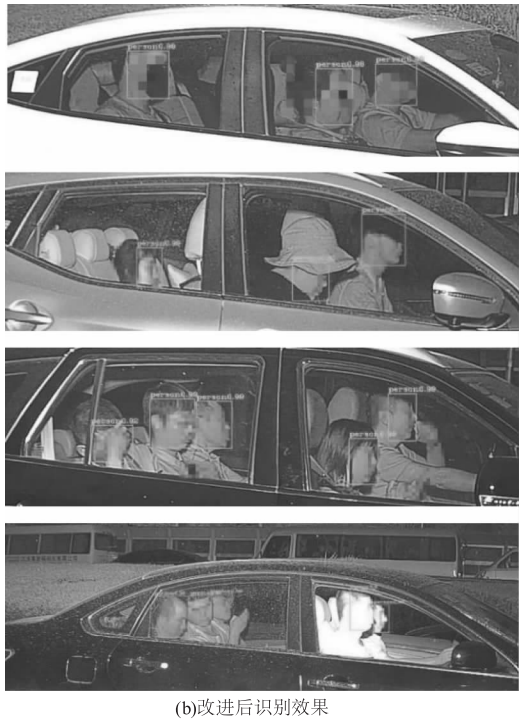
由检测结果可以看出,当目标的曝光程度较高导致面部信息缺失严重时,改进后的网络也可以对乘员面部准确识别,并保持较高的识别分数。此外由于照片抓拍过程的不确定性,乘员面部容易被车窗、帽子、口罩等物体遮挡,也因此对网络的泛化能力提出了更高的要求。从实验中可以看出原网络可以实现对理想目标的准确识别,但当存在上述干扰时还不能很好的胜任检测任务,在目标存在外部遮挡导致同一类别目标特征的差异较大时,改进后的网络的可以很好的完成检测任务,这说明变形改进的网络计算可以尽可能多的保留目标细节特征,与预期设想基本相符,同时目标框的位置预测更加精准,避免了相邻目标发生重叠。

同时为了量化验证本文的改进效果,选取了多组数据,每组 2500 张图片,每类乘员数量(1 ~ 5)各 500 张,将改进后的网络和目前主流的 YOLOv3、Mask RCNN 等检测算法分别进行实验,记录各类乘员数量下检测的平均准确率和检测速度,统计结果如表 2 所示。

从表 2 中数据可以看出,本文模型在各类乘员数量下识别准确率较高,基本保证在 80 % 以上,检测的速度与改进前略微降低,但仍在可承受范围之内。与改进前相比,检测乘员越多,本文模型准确率提升幅度越大,说明本文的 NMS 改进能够很好应对多乘员情况下重叠目标的检测。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,本文数据经过裁剪多余背景信息的预处理后图像多为长方形,压缩为正方形后再输入网络训练,这样会使得图像质量降低,网络预测的准确率不佳。Mask RCNN 的实验结果表明实例分割的



(a)改进前识别效果



(b)改进后识别效果

图 10 改进前后检测效果对比

Fig. 10 Comparison of detection effects before and after improvement

效果优于一般的目标检测算法,但是语义分割网络的加入使得网络整体的计算量增大,检测速度牺牲较大。基于以上分析总结得知,检测精度随乘员数量增加呈下降趋势,本文在 3 + 数量上准确较改进前提高幅度更大,与其他算法相比综合检测精度、检测速度等指标,本文算法仍是最优选择。

表 2 不同算法检测效果对比

Tab. 2 Comparison of detection effects of different algorithms

网络	平均准确率/%					FPS
	1	2	3	4	5	帧/秒
Faster RCNN	74.8	72.6	69.8	65.4	58.8	17
本文	86.8	83.4	81.4	79.8	75.2	12
YOLOv3	68.6	65.2	61.8	57.4	51.2	30
Mask RCNN	82.4	78.8	76.2	72.4	66.8	5

6 结 论

结合已获得的多光谱红外图像,本文以 Faster RCNN 网络模型为基础,通过变形卷积网络结构、在 RPN 网络中目标特征图映射方式中提出变形的 ROI-Pooling,以提高 RPN 网络中特征映射的质量,使其可以更多的保留目标的细节特征,网络模型的泛化能力得到明显的提高。在位置回归过程中,KL 损失与方差投票相结合的改进方式,使得网络对于重叠目标的识别能力明显增强。本文算法可以很好应对在可能存在误报、漏报的情况下的目标识别。相比于改进之前,3 + 乘员数量检测准确率提升幅度更大,虽然 4 + 数量乘员检测准确率还不能稳定居于 80 % 以上,但在 4 + 数量乘员以上允许 1 ~ 2 名漏报的原则下,本文算法完全可以满足实际的执法和检测需求。但是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] Chen Lufeng, Ye Pengfei, Yu Jiahuan, et al. Feasibility analysis of HOV application in urban traffic management [J]. Transportation Enterprise Management, 2016, (4): 38 – 41. (in Chinese)
陈鲁锋,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,(4):38 – 41.

[2] Yang Pi, Xu Changbin, Ma Yuying, et al. A review of infrared dim target detection algorithm under low SNR[J]. Laser & Infrared. 2019, 49(6): 643 – 649. (in Chinese)
杨昳,徐长彬,马玉莹,等. 低信噪比下的红外弱小目标检测算法研究综述[J]. 激光与红外,2019,49(6):643 – 649.

[3] Zhang Jun, Fan bin, Yang Xinjun. Human body detection algorithm of fire infrared image based on unet network [J]. Laser & Infrared, 2019, 49(12): 1483 – 1489. (in Chinese)
张骏,范彬,杨新军. 基于 U-Net 网络的消防红外图像的人体检测算法[J]. 激光与红外,2019,49(12):1483 – 1489.

[4] Ronneberger, Olaf, Fischer, et al. U-Net: convolutional networks for biomedical image segmentation[C]//MICCAI, 2015.

[5] Li Chenghao, Ru le, He Linyuan, et al. A target detection method for variable anchor frame candidate area networks [J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(8): 1610 – 1617.
李承昊,茹乐,何林远,等. 一种可变锚框候选区域网络的目标检测方法[J]. 北京航空航天大学学报,2020,46(8):1610 – 1617.

[6] Ren Shaoqing, He Kaiming, Girshick, Ross, et al. Faster RCNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137 – 1149.

[7] Jifeng Dai, Haozhi Qi, Yuwen Xiong, et al. Deformable convolutional network[C]//International Conference on Computer Vision, 2016.

[8] Yihui He, Chenchen Zhu, Jianren Wang, et al. Bounding box regression with uncertainty for Accur-ate object detection[J]. arXiv preprint arXiv: 1809.08545, 2018, 9.

[9] XU Dai, YUE Zhang, YANG Wenxia, et al. Tamper image recognition based on improved three – way Fasert RCNN [J/OL]. Journal of Computer Applications, http://www.joca.cn/CN/10.11772/j.issn.1001 – 9081.2019081515. (in Chinese)
徐代,岳章,杨文霞,等. 基于改进的三向流 Faster RCNN 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.

[10] Sun Qing, Li Ling, Xin Yunhong. Infrared small target detection algorithm based on local multiscale low rank decomposition[J]. Laser & Infrared, 2019, 49(3): 370 – 375. (in Chinese)
孙青,李玲,辛云宏. 基于局部多尺度低秩分解的红外小目标检测算法[J]. 激光与红外,2019,49(3):370 – 375.