doi:10.6041/j.issn.1000-1298.2019.04.003

基于改进 SSD 的果园行人实时检测方法

刘　慧　张礼帅　沈　跃　张　健　吴　边
(江苏大学电气信息工程学院, 镇江 212013)

摘要: 农田障碍物的精确识别是无人农业车辆必不可少的关键技术之一。针对果园环境复杂难以准确检测出障碍物信息的问题,提出了一种改进单次多重检测器(Single shot multibox detector,SSD)深度学习目标检测方法,对田间障碍物中的行人进行识别。使用轻量化网络 MobileNetV2 作为 SSD 模型中的基础网络,以减少提取图像特征过程中所花费的时间及运算量,辅助网络层以反向残差结构结合空洞卷积作为基础结构进行位置预测,在综合多尺度特征的同时避免下采样操作带来的信息损失,基于 Tensorflow 深度学习框架,在卡耐基梅隆大学国家机器人工程中心的果园行人检测开放数据集上进行不同运动状态(运动、静止)、不同姿态(正常、非正常)和不同目标面积(大、中、小)的田间行人识别精度和识别速度的对比试验。试验表明,当 IOU 阈值为 0.4 时,改进的 SSD 模型田间行人检测模型的平均准确率和召回率分别达到了 97.46% 和 91.65%,高于改进前 SSD 模型的 96.87% 和 88.51%,并且参数量减少至原来的 1/7,检测速度提高了 187.5%,检测速度为 62.50 帧/s,模型具有较好的鲁棒性,可以较好地实现田间环境下行人的检测,为无人农机的避障决策提供依据。

关键词: 无人农业车辆; 行人检测; 单次多重检测器; 空洞卷积; MobileNetV2
中图分类号: TP391　　**文献标识码**: A　　**文章编号**: 1000-1298(2019)04-0029-07

Real-time Pedestrian Detection in Orchard Based on Improved SSD

LIU Hui　ZHANG Lishuai　SHEN Yue　ZHANG Jian　WU Bian
(*School of Electrical and Information Engineering*, *Jiangsu University*, *Zhenjiang* 212013, *China*)

Abstract: Reliable pedestrian detection in agriculture field is one of the key technologies for unmanned agricultural vehicles. For the complex environment in the orchard, it is difficult to accurately detect the obstacle information. To solve this problem, an improved single shot multibox detector (SSD) deep learning object detection method was proposed to detect pedestrian in the field obstacles. The lightweight network framework MobileNetV2 was used as the basic network in the SSD model to reduce the time and computational effort of extracting image features. For auxiliary layer of the SSD model, the inverse residual block combined with the dilated convolution was used as the basic structure for position prediction so that the multi-scale features can be integrated and at the same time avoiding the information loss caused by the down sampling operation. Based on the Tensorflow deep learning framework, different motion states (motion and static), different pose states (normal and unnormal) and different object scales (large, medium and small) pedestrian detection experiment in orchard were carried out on the open data set in orchard environment of the National Robotics Engineering Center of Carnegie Mellon University and the accuracy and speed of these different situations were compared. Result showed that the average precision and recall rate of the improved SSD pedestrian detection model in agriculture reached 97.46% and 91.65%, respectively, higher than 96.87% and 88.51% of the original SSD model, and the parameter quantity was decreased by seven times. The speed was accelerated by three times and the detection speed was 62.50 frames per second. The model had good robustness and could detect the pedestrian in the field environment, which could provide a basis for the obstacle avoidance of the unmanned agriculture machinery.

收稿日期: 2018-12-29　修回日期: 2019-02-25
基金项目: 国家自然科学基金项目(5150195)、江苏省国际科技合作项目(BZ2017067)、江苏省重点研发计划项目(BE2018372)、江苏省自然科学基金项目(BK20181443)、镇江市重点研发计划项目(NY2018001)和江苏高校青蓝工程项目
作者简介: 刘慧(1980—),女,副教授,博士,主要从事农业电气化与自动化、智能控制与信号处理研究,E-mail: amity@ujs.edu.cn
通信作者: 沈跃(1978—),男,教授,博士生导师,主要从事农业电气化与自动化、嵌入式系统与控制研究,E-mail: shen@ujs.edu.cn

Key words: unmanned agricultural vehicles; pedestrian detection; SSD; dilated convolution; MobileNetV2

0 引言

随着精准农业理论的提出以及智能化机器人的发展,智能农业车辆的自动导航越来越受到国内外的关注[1]。为了保证智能化车辆在无人工干预时在田间操作的安全性,必须有实时的障碍物检测,当需要人机共同协作完成一些任务时,对田间行人的检测更是首要考虑的问题。在室外田间环境下,自动导航时存在很大的不确定性,常用的障碍物检测方法有激光雷达传感器检测[2]、超声波传感器检测[3]、红外传感器检测[4]和计算机视觉检测[5]等。

田间环境下的障碍物检测,由于其复杂的自然环境、障碍物形态的多变性、光照等外部条件的大范围变化等,实施起来具有一定难度。田间环境下,超声波传感器存在检测障碍物空间位置准确性较差、易受干扰等缺点,激光雷达传感器可以较直观地检测障碍物,但雷达系统的造价昂贵[6]。计算机视觉检测相比于其他障碍物检测方法具有成本低、能够有效利用环境中的颜色与纹理信息等优点。

近年来由于深度学习相关理论的快速发展及计算能力的提升,深度卷积网络在计算机视觉方面取得了很大的成功。在目标检测方面,基于深度学习的方法准确率大大超过了传统的基于 HOG、SIFT 等人工设计特征的检测方法[7]。基于深度学习的目标检测主要包括两类,一类是基于区域生成的卷积网络结构,代表性的网络为 R - CNN[8]、fast R - CNN[9]、faster R - CNN[10];另一类是把目标位置的检测视作回归问题,直接利用 CNN 网络结构对整个图像进行处理,同时预测出目标的类别和位置,代表性的网络有 YOLO[11-13]、SSD[14-15]等。行人检测是目标检测的子问题,基于卷积网络的目标检测模型由于不需要手动设计特征,可通过深层卷积网络结构自动学习图像的高阶特征,从而可生成更加可靠的检测结果[16]。针对深度学习模型参数众多导致其难以部署在嵌入式设备的问题,很多学者提出了各种相应的模型压缩方法[17-19]。

本文在 LIU 等[14]提出的 SSD 目标检测模型基础上对其进行改进,使用 MobileNetV2[19]网络架构作为 SSD 的基础网络进行特征提取,并对辅助层的卷积结构使用反向残差结构并结合空洞卷积进行位置预测,在利用多尺度信息的同时可以减少计算量和参数量,使其能够部署在移动端设备中,通过迁移学习对训练好的网络模型进行调优,从而减少模型训练时间并使得模型更容易收敛。本文采用计算机视觉方法结合深度学习进行农业车辆自动作业过程中的行人障碍物检测,从而为进一步实现田间行人避让提供理论基础。

1 田间行人识别模型

1.1 SSD 网络结构

SSD 目标检测模型由于不需要耗时的区域生成及特征重采样步骤,直接对整个图像进行卷积操作并预测出图像中所包含物体的类别及对应的坐标,从而极大提高了检测速度,同时通过使用小尺寸的卷积核、多尺度预测等使得目标检测的精度得到很大提升。

SSD 网络结构分为基础网络(Base network)和辅助网络(Auxiliary network)两部分:基础网络为在图像分类领域具有很高分类精度且去除其分类层的网络;辅助网络为在基础网络基础上增加的用于目标检测的卷积网络结构,这些层的尺寸逐渐减小从而可以进行多尺度预测。每个添加的辅助网络层都会通过一系列卷积核产生一个固定的预测集,对于一个 $m \times n \times p$(p 为通道数,m、n 为尺寸)的特征层,每个辅助层会使用 $3 \times 3 \times p$ 的卷积核对其进行预测并产生某一类别的得分值,或者是物体相对于默认边界框的位置偏移量,且在 $m \times n$ 个位置都分别预测出相应的值。

SSD 模型在特征图的每个位置预测 k 个边界框,并且同时预测某一物体类别出现在此位置的得分和物体位置相对于边界框的偏移量,从而在每个特征图的位置分别预测 $c \times k$(c 为类别数)个得分和 $4 \times k$ 个位置偏移量,对于一个尺寸为 $m \times n$ 的特征图,总体会预测出 $(c+4)kmn$ 个输出量。最后对输出结果进行非极大值抑制来得到最终的关于图像中物体类别及位置信息的预测值。

1.2 改进的 SSD 田间行人检测模型

SSD 目标检测模型使用 VGG 网络[20]作为基础网络,但 VGG 网络模型参数众多,在特征提取过程中占用了大部分运行时间,且在前向传播过程中由于存在非线性变换导致变换过程中信息的损失。

SANDLER 等[19]在流形学习理论的基础上提出非线性激活函数 ReLU 在高维度下会较好地保留信息,而在低维度下会造成较大的信息丢失,故在输入层应该增加特征维度之后再对其进行非线性变换,而在输出层应该对特征进行降维后使用线性激活函

数以减少信息的丢失，据此提出反向残差结构(Inverted residual block)。

反向残差结构中的下采样操作在增大卷积核感受野的同时会造成特征信息的丢失，所以考虑舍弃卷积结构中的下采样操作并引入空洞卷积[21-22]来解决此问题。空洞卷积是在原始卷积操作的基础上增加一个扩张参数，将卷积核扩张到相应的尺度中，同时在原卷积核中未被占用到的区域填充 0，应用空洞卷积可以在不用下采样操作的情况下增加卷积核的感受野。但空洞卷积的使用会令卷积核对数据操作不连续以及对于小物体不能较好识别，本文考虑使用层级特征融合[23](Hierarchical feature fusion)来解决引入空洞卷积所带来的问题。

层级特征融合是对空洞卷积层的每一个卷积单元的输出依次进行求和，并且把每个求和后的结果都通过连接(Concatenate)操作得到最后的输出结果，如图 1 所示。相比于其他通过使用较小的扩张参数使得学习参数变多的方法，此方法具有操作简单且不增加卷积结构复杂性的优点。空洞卷积层结构表示为(输入通道，感受野，输出通道)，其中空洞卷积核的有效感受野为 $n_k \times n_k$，$n_k = 2^{k-1}(n-1)+1$；$k = 1,2,\cdots,K$。

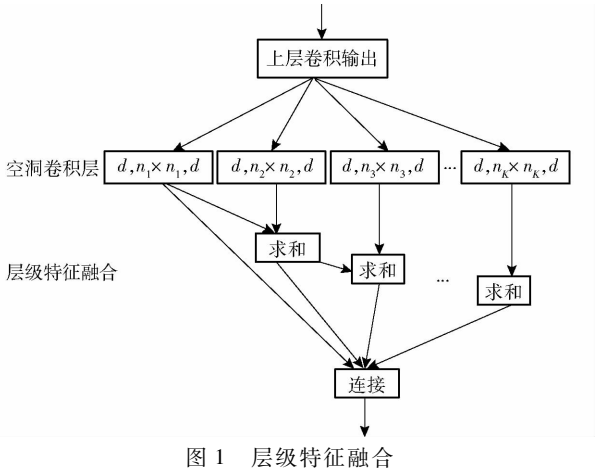

图 1　层级特征融合

Fig. 1　Hierarchical feature fusion

反向残差结构使用 ReLU6 为激活函数，其输出为

$$Y = \min(\max(X,0),6) \tag{1}$$

式中　Y——ReLU6 激活函数的输出

X——输入特征值

ReLU6 相比于 ReLU[24]在低精度运算场景中具有更好的鲁棒性，另外，使用 3×3 的卷积核，并且在训练网络过程中使用 dropout[25] 和 batch normalization[26]以减少训练过程中的过拟合。改进前后的反向残差结构如图 2 所示。其中 Dwise 表示基于深度可分离的卷积结构[27]，Dilated 表示卷积方式是空洞卷积，Linear 为线性激活函数，HFF 表示层级特征融合。

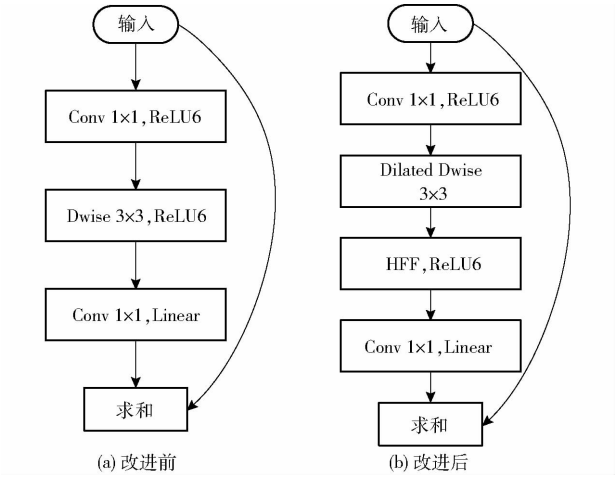

图 2　改进前后反向残差结构

Fig. 2　Inverted residual block before and after improvement

结合改进后的反向残差结构对 SSD 目标检测模型的基础层和辅助层进行改进：①原始 SSD 使用 VGG 网络作为基础层进行特征提取，但 VGG 网络模型不适合在移动设备上部署运行，把其替换为 SANDLER 等[19]在反向残差结构基础上提出的 MobileNetV2 网络，其具有参数少、占用空间小且运行速度较快的优点，以此作为 SSD 的特征提取网络可以加快运行速度并且减少模型的尺寸及运算量。②SSD 辅助层使用传统的卷积网络结构导致其参数量和运算量较大，本文使用改进后的反向残差结构作为辅助层的基本结构，改进后的辅助网络层可以减少学习过程中非线性变换造成的信息损失且卷积核具有多尺度的感受野。改进后的 SSD 网络结构如图 3 所示。基础网络为去除预测层后的 MobileNetV2 网络，辅助网络层使用改进后的反向残差结构进行位置及相应类别概率的预测，其中 19×19×1 280 为输出的特征尺寸，其余类似。

2　目标函数

训练过程中所使用的目标函数为

$$L(x,c,l,g) = \frac{1}{N}(L_{conf}(x,c) + \alpha L_{loc}(x,l,g)) \tag{2}$$

其中

$$L_{loc}(x,l,g) = \sum_{i\in P_{os}}^{N}\sum_{m\in B_{ox}} x_{ij,k} S_{L1}(l_{i,m}(l_{i,m} - \hat{g}_{j,m})) \tag{3}$$

$$L_{conf}(x,c) = -\sum_{i\in P_{os}}^{N} x_{ij,p}\lg\hat{c}_{i,p} - \sum_{i\in N_{eg}}\lg\hat{c}_{i,0} \tag{4}$$

$$\hat{c}_{i,p} = \frac{\exp c_{i,p}}{\sum_{p}\exp c_{i,p}} \tag{5}$$

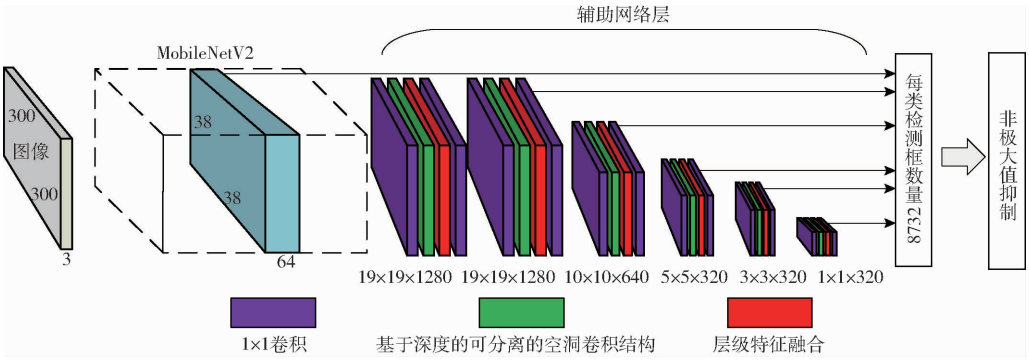

图 3 改进后的 SSD 网络结构

Fig. 3 Improved SSD framework

式中 N——匹配的默认边界框的个数，当 N 为 0 时，设置 L 为 0

x——预测框与真实框是否匹配

c——标注类别

α——位置预测误差与分类预测误差的加权系数

S_{L1}——平滑 L1 误差函数

l——预测的边界框坐标

g——标注的边界框坐标

L_{loc}——位置预测误差，为平滑 L1 误差函数

L_{conf}——对应的 softmax 多分类误差函数

$x_{ij,k}$——第 i 个预测框与第 j 个真实框关于类别 k 是否匹配

$\hat{g}_{j,m}$——真实框 $l_{i,m}$——预测框

P_{os}——样本中的正例

B_{ox}——预测框中心坐标及其宽和高的集合

N_{eg}——样本中的负例

$x_{ij,p}$——预测框 i 与真实框 j 关于类别 p 是否匹配

$\hat{c}_{i,p}$——第 i 个预测框中目标是第 p 个类别的概率

$c_{i,p}$——第 i 个预测框中目标属于类别 p 的得分值

$\hat{c}_{i,0}$——预测框中没有物体的概率

3 苹果园行人识别试验

3.1 试验平台

采用 32GB 内存、NVIDIA GTX1080TI 型 GPU、AMD 锐龙 5－2600X CPU 作为硬件平台，操作系统为 Linux Ubuntu 16.04，并行计算框架版本为 CUDA 8.0，深度神经网络加速库为 CUDNN v7.0。采用 Python 编程语言在 Tensorflow[28] 深度学习框架上实现本文的 SSD 目标检测模型并完成对模型的训练及验证。

3.2 试验数据

针对已开放的行人检测数据集都是在城市道路等结构化环境中，PEZZEMENTI 等[29] 制作了农业环境下行人检测数据集以促进深度学习技术在农业环境下的应用，包括苹果园和橙园环境下的已标注图像。本文使用苹果园环境下的数据，分为训练集、验证集和测试集。训练集包含行人图像有 15 526 幅，不包含行人的图像有 4 570 幅；验证集中包含行人图像有 8 124 幅，不包含行人图像有 1 981 幅；测试集中包含行人图像有 7 691 幅，不包含行人图像有 1 949 幅。其中包含行人的图像分为静止状态和移动状态；按照行人目标所占像素区域面积分为小目标、中等目标、大目标；按照行人所处的姿势分为正常姿势、非正常姿势，其中正常姿势是正常站立姿势，非正常姿势包括躺、坐、蹲和跌落姿势。

3.3 数据预处理

由于样本中各种情形下的图像样本并不一致，为了避免样本分布不均导致训练出的模型泛化性能不佳问题，对样本中数量较少的分类进行水平翻转和平移等操作进行数据增广，使得各种情形下的样本数量基本一致。同时由于苹果园自然光照下，特别是在光照很强时，由于行人被植株遮挡或受植株阴影的影响，使得处于此环境下的行人相对于自然光照射下表面会有较大变化，从而会影响行人检测的精度。本文采用自适应直方图均衡化来对图像进行增强从而提高图像的质量，减少光照变化对图像的影响。

3.4 模型训练

为了节省训练时间及加快收敛速度，本文使用迁移学习来训练深度学习模型。首先加载已经训练好的 MobileNetV2 分类网络的参数，除去最后的分类层，其余参数值赋给 SSD 模型中对应的参数，其余各层参数是以 0 为均值、0.01 为标准差的高斯分布进行随机初始化。

本文使用批量随机梯度下降算法，设置 batch-size 为 128，冲量为 0.9，权值衰减系数为 2×10^{-3}，最大迭代次数为 8×10^{5} 次，初始学习率为 0.004，衰

减率为 0.95，每 10 000 次迭代后衰减一次，每间隔 10 000 次迭代后保存一次模型，最终选取精度最高的模型。训练过程中使用困难样本挖掘（Hard negative mining）策略[30]，即训练过程中先用初始的正负样本训练检测模型，然后使用训练出的模型对样本进行检测分类，把其中检测错误的样本继续放入负样本集合进行训练，从而加强模型判别假阳性的能力。

4　结果与分析

4.1　评价指标

通过在测试集中分析模型改进前后的检测精度、检测速度和参数量以对比模型的性能。把测试集图像输入训练好的网络，对田间行人位置进行检测并记录检测结果，当模型预测的目标边界与测试集对应的标注数据中的边界框的交并集比（IOU）大于等于设定的阈值时，认为检测结果正确，否则视为检测错误。检测精度的评价指标选取准确率（Precision，P）、召回率（Recall，R）和调和均值 F_1，各个评价指标的定义为

$$P=\frac{T_P}{T_P+F_P}\times 100\% \tag{6}$$

$$R=\frac{T_P}{T_P+F_N}\times 100\% \tag{7}$$

$$F_1=\frac{2PR}{P+R} \tag{8}$$

式中　T_P——正确检测到行人的数量

F_P——误把非行人目标检测为行人目标的数量

F_N——误把行人检测为背景的数量

F_1——对准确率和召回率的调和均值，越接近于 1，表明模型表现越好

4.2　改进 SSD 模型对田间行人的识别结果

在测试集中分别对改进前后的 SSD 目标检测模型进行测试，并统计其检测结果和检测速度，结果见表 1。改进后的网络模型准确率和召回率分别提高了 0.82 个百分点和 1.62 个百分点。通过对卷积网络的改进使得网络模型参数量相比于原始模型减少至原来的 1/7，且其检测速度提高了 187.5%，使得目标检测模型更适合在移动机器人中部署，实时性更好。

4.3　IOU 阈值对检测结果的影响

前述结果分析中默认选择 IOU 阈值为 0.5，当检测结果与标注结果的 IOU 阈值大于等于 0.5 时，认为检测结果正确，否则视为未检测出行人目标。当 IOU 阈值在 0.6 及以下时，行人检测模型的表现

表 1　改进前后 SSD 模型果园行人检测结果比较

Tab.1　Comparison of detection results between improved SSD and original SSD

模型	P/%	R/%	F_1/%	检测速度/（帧·s^{-1}）	参数量
改进后	97.13	89.72	93.28	62.50	5.30×10^6
改进前	96.31	88.10	92.02	21.74	3.61×10^7

对 IOU 阈值的变化并不十分敏感，当 IOU 阈值在 0.6 以上时，行人检测模型会受到很大影响，其漏检率将快速上升[31]。不同于城市环境下对检测到的行人采取的避障策略有很多种选择，在果园田间环境下，由于受到地形的限制，农机所采取的避障策略可能只有减速或者停止，而不能采取规划路径从而避开行人的决策，故而其对田间行人检测的位置精度要求不高。通过比较 IOU 阈值在 0.3～0.7 之间时准确率和召回率，选择出较为合适的 IOU 阈值。不同 IOU 阈值下，改进前后模型检测结果如表 2 所示。根据结果得出改进前后的 SSD 目标检测模型在果园环境下行人检测的 IOU 阈值均在 0.4 最合适，改进后模型的准确率和召回率比阈值在 0.5 时分别提高了 0.33 个百分点和 1.93 个百分点，改进前模型的准确率和召回率比阈值 0.5 时分别提高了 0.56 个百分点和 0.41 个百分点，故而在 IOU 阈值为 0.4 下进行改进前后模型检测性能的对比试验。

表 2　不同 IOU 阈值下改进前后模型的检测结果

Tab.2　Detection statistic results on various IOU thresholds　%

IOU 阈值	模型	P	R	F_1
0.7	改进后	93.40	85.89	89.49
	改进前	93.06	85.48	89.11
0.6	改进后	96.65	89.31	92.84
	改进前	95.76	87.30	91.33
0.5	改进后	97.13	89.72	93.28
	改进前	96.31	88.10	92.02
0.4	改进后	97.46	91.65	94.47
	改进前	96.87	88.51	92.50
0.3	改进后	97.46	91.65	94.47
	改进前	96.87	88.51	92.50

4.4　不同运动状态模型识别结果

NREC 开放田间行人识别数据集中记录了拍摄图像时田间的行人是处于静止状态还是连续运动状态，故而可以根据标注把行人按照运动状态分为静止、运动状态，测试模型在这两种情形下识别性能的变化。当行人处于静止状态时，改进后的 SSD 检测结果的 F_1 为 93.88%，运动状态时为 90.09%，在运动状态下的行人检测性能比静止状态降低了 3.79 个百分点。根据试验结果可知 MobileNetV2 网络架

构相比于原始 SSD 模型中的 VGG 网络在运动目标的特征提取上稍弱,故而改进后的模型在运动目标的检测中性能稍弱,但 MobileNetV2 网络的计算量要远低于 VGG 网络,因此,可大大提高模型的运算速度。检测结果见图 4a。

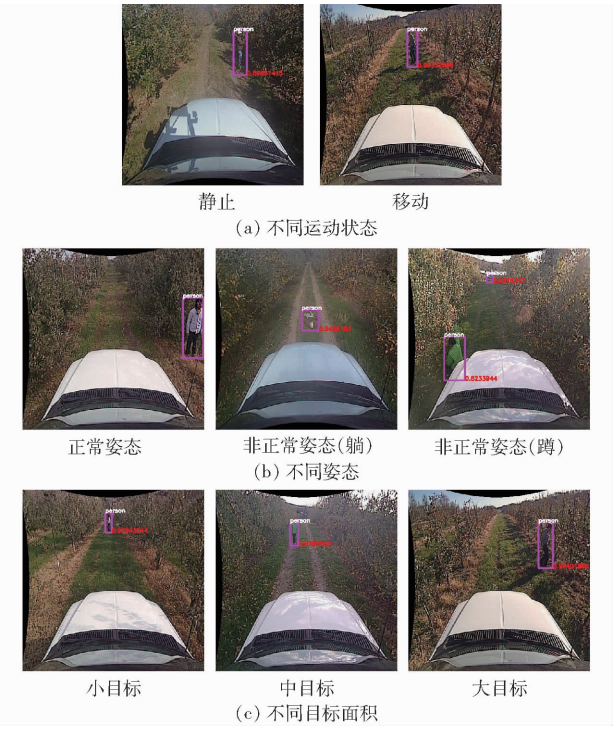

图 4 改进 SSD 模型对不同情形下的行人识别结果

Fig. 4 Detection results of improved SSD model on pedestrian in various conditions

4.5 不同行人姿态下模型识别结果

数据集中包含果园行人处于多种姿态下的图像,把其分为两类:一类是正常姿态,即行人处于正常站立状态;另一类为非正常姿态,包括行人处于躺、坐、蹲和跌落的状态。在正常姿态下改进后 SSD 模型识别结果的 F_1 为 92.66%,非正常姿态下 F_1 为 89.88%。行人不同姿态对模型检测结果有较大影响,检测结果见图 4b。

4.6 不同目标面积下模型识别结果

按照田间行人目标占区域像素面积的比例分为大目标、中目标、小目标,其中像素面积在 3 501 像素以上的为大目标,面积在 1 301 ~ 3 500 像素之间的为中目标,面积在 1 300 像素以下的为小目标。改进后 SSD 模型在目标为大、中、小情况下的 F_1 分别为 95.07%、92.64%、84.09%。由于行人所占区域像素面积一定程度上反映了行人与摄像机的距离,故而可以分析出当行人处于近、中距离时检测性能较高,而在小目标情况下即行人距离较远时检测性能较改进前稍弱,这是由于层级特征融合并不能完全消除空洞卷积带来的影响,从而影响了小目标的检测性能。由此可见,空洞卷积的引入提高了模型检测近、中距离目标的性能,检测远处目标的性能稍微下降,符合障碍物检测任务中不同障碍物检测的优先顺序,即近、中距离处的障碍物是优先需要处理的目标。检测结果见图 4c。

表 3 为不同情形下改进前 SSD 模型与改进后 SSD 模型的准确率、召回率和 F_1 的统计结果,其中 IOU 阈值取为 0.4。改进后的 SSD 模型在除行人处于运动状态下检测准确率和召回率有下降外,其余情形均有所提高。

表 3 不同行人状态下的检测结果

Tab. 3 Detection statistic result on pedestrian in various status %

	类别	模型	P	R	F_1
不同运动状态	静止	改进后	97.73	90.32	93.88
		改进前	96.92	89.44	93.03
	运动	改进后	93.13	87.25	90.09
		改进前	95.46	88.51	91.85
不同姿态	正常姿态	改进后	95.83	89.69	92.66
		改进前	97.19	85.48	90.96
	非正常姿态	改进后	93.56	86.47	89.88
		改进前	93.54	83.63	88.31
不同目标面积	大目标	改进后	97.75	92.54	95.07
		改进前	97.76	91.53	94.54
	中目标	改进后	96.23	89.31	92.64
		改进前	97.73	85.48	91.20
	小目标	改进后	95.37	75.20	84.09
		改进前	94.10	77.62	85.07

5 结论

(1)基于卡耐基梅隆大学国家机器人工程中心开放的用于农业环境下行人检测的数据集,采用改进的 SSD 目标检测模型,进行田间环境下的行人障碍物检测,模型占用空间较小且轻量化,适合于在移动设备上部署,试验结果表明,模型具有较高的准确性,准确率和召回率分别达到了 97.13% 和 89.72%,每幅图像的平均检测速度为 62.50 帧/s。在对 IOU 阈值进行合理调整后最终平均准确率和召回率分别达到了 97.46% 和 91.65%。

(2)通过迁移学习技术,把 MobileNetV2 在 Imagenet 分类表现较好的参数移植到 SSD 的特征提取网络模型中,从而简化了目标检测模型的训练过程并缩短了训练时间。

(3)通过改进原始 SSD 的特征提取网络,使用更加轻量化的 MobileNetV2 网络模型进行特征提取,辅助层使用改进后的反向残差结构进行卷积运算,从而可以利用多特征信息并且减少运算量,当

IOU 阈值为 0.4 时，较原始 SSD 网络模型的准确率和召回率分别提高了 0.59 个百分点和 3.14 个百分点，参数量减少至原来的 1/7，检测速度提高了 187.5%。

(4)根据果园环境下和城市环境下对避开行人所采取的决策不同，提出在果园环境下合适的 IOU 阈值为 0.4 左右，在此阈值下，模型的准确率和召回率比阈值在 0.5 时分别提高了 0.33 个百分点和 1.93 个百分点。

参 考 文 献

[1] 钱晓明，朱立群，楼佩煌，等. 基于位姿状态的全向运行型 AGV 路径跟踪优化控制方法[J/OL]. 农业机械学报，2018，49(4)：20－26.
QIAN Xiaoming，ZHU Liqun，LOU Peihuang，et al. Optimal path tracking control method of omni-directional mobile AGV based on pose state[J/OL]. Transactions of the Chinese Society for Agricultural Machinery，2018，49(4)：20－26. http：//www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20180402&journal_id=jcsam. DOI：10.6041/j.issn.1000-1298.2018.04.002.(in Chinese)

[2] KISE M，ZHANG Q，NOGUCHI N. Obstacle detection using laser range finder for autonomous agricultural machinery[J]. Transactions of the ASABE，2005，48(3)：1269－1278.

[3] GUO Linsong，ZHANG Qin，HAN Shufeng. Agricultural machinery safety alert system using ultrasonic sensors[J]. Journal of Agricultural Safety and Health，2002，8(4)：385－396.

[4] 薛金林，董淑娴，范博文. 基于信息融合的农业自主车辆障碍物检测方法[J/OL]. 农业机械学报，2018，49(增刊)：29－34.
XUE Jinlin，DONG Shuxian，FAN Bowen. Detection of obstacles based on information fusion for autonomous agricultural vehicles[J/OL]. Transactions of the Chinese Society for Agricultural Machinery，2018，49(Supp.)：29－34. http：//www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=2018s005&journal_id=jcsam. DOI：10.6041/j.issn.1000-1298.2018.S0.005.(in Chinese)

[5] 薛金林，闫嘉，范博文. 多类农田障碍物卷积神经网络分类识别方法[J/OL]. 农业机械学报，2018，49(增刊)：35－41.
XUE Jinlin，YAN Jia，FAN Bowen. Classification and identification method of multiple kinds of farm obstacles based on convolutional neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery，2018，49(Supp.)：35－41. http：//www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=2018s006&journal_id=jcsam. DOI：10.6041/j.issn.1000-1298.2018.S0.006.(in Chinese)

[6] 何勇，蒋浩，方慧，等. 车辆智能障碍物检测方法及其农业应用研究进展[J]. 农业工程学报，2018，34(9)：21－32.
HE Yong，JIANG Hao，FANG Hui，et al. Research progress of intelligent obstacle detection methods of vehicles and their application on agriculture[J]. Transactions of the CSAE，2018，34(9)：21－32.(in Chinese)

[7] HU Q，WANG P，SHEN C，et al. Pushing the limits of deep cnns for pedestrian detection[J]. IEEE Transactions on Circuits and Systems for Video Technology，2018，28(6)：1358－1368.

[8] GIRSHICK R，DONAHUE J，DARRELL T，et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition，2014：580－587.

[9] GIRSHICK R. Fast R-CNN[C]// Proceedings of the IEEE International Conference on Computer Vision，2015：1440－1448.

[10] REN S，HE K，GIRSHICK R，et al. Faster R－CNN：towards real-time object detection with region proposal networks[C]// International Conference on Neural Information Processing Systems，2015：91－99.

[11] REDMON J，DIVVALA S，GIRSHICK R，et al. You only look once：unified，real-time object detection[C]// Computer Vision and Pattern Recognition. IEEE，2016：779－788.

[12] REDMON J，FARHADI A. YOLO9000：better，faster，stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition，2017：6517－6525.

[13] 熊俊涛，刘振，林睿，等. 自然环境下树上绿色芒果的无人机视觉检测技术[J/OL]. 农业机械学报，2018，49(11)：23－29.
XIONG Juntao，LIU Zhen，LIN Rui，et al. Unmanned aerial vehicle vision detection technology of green mango on tree in natural environment[J/OL]. Transactions of the Chinese Society for Agricultural Machinery，2018，49(11)：23－29. http：//www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20181103&journal_id=jcsam. DOI：10.6041/j.issn.1000-1298.2018.11.003.(in Chinese)

[14] LIU W，ANGUELOV D，ERHAN D，et al. SSD：single shot multibox detector[C]// European Conference on Computer Vision，2016：21－37.

[15] 彭红星，黄博，邵园园，等. 自然环境下多类水果采摘目标识别的通用改进 SSD 模型[J]. 农业工程学报，2018，34(16)：155－162.
PENG Hongxing，HUANG Bo，SHAO Yuanyuan，et al. General improved SSD model for picking object recognition of multiple fruits in natural environment[J]. Transactions of the CSAE，2018，34(16)：155－162.(in Chinese)

[16] ZHANG S，BENENSON R，OMRAN M，et al. Towards reaching human performance in pedestrian detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2018，40(4)：973－986.

[17] HAN S，POOL J，TRAN J，et al. Learning both weights and connections for efficient neural network[C]// Advances in Neural Information Processing Systems，2015：1135－1143.

[18] COURBARIAUX M，HUBARA I，SOUDRY D，et al. Binarized neural networks：training deep neural networks with weights and activations constrained to +1 or －1[J]. arXiv preprint arXiv：1602.02830，2016.

(下转第 101 页)

[20] 张瑞. 一年两熟地区麦茬地玉米免耕播种播深控制机构的研究[D]. 北京:中国农业大学, 2016.
ZHANG Rui. Study on precision depth-control mechanism of corn no-till planter in double-cropping area[D]. Beijing: China Agricultural University, 2016. (in Chinese)
[21] 贾洪雷,朱龙图,黄东岩,等. 基于 Flex 传感器的免耕播种机播深自动控制系统[J]. 吉林大学学报,2012,42(1):166 - 175.
JIA Honglei, ZHU Longtu, HUANG Dongyan, et al. Automatic control system of sowing depth for no-till planter based on Flex sensor[J]. Journal of Jilin University, 2012, 42(1): 166 - 175. (in Chinese)
[22] 王汉羊,陈海涛,纪文义. 2BMFJ - 3 型麦茬地免耕精播机防堵装置[J/OL]. 农业机械学报,2013,44(4):64 - 70.
WANG Hanyang, CHEN Haitao, JI Wenyi. Anti-locking mechanism of type 2BMFJ - 3 no-ill precision planter for wheat stubble fields[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(4): 64 - 70. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no = 20130412&flag = 1. DOI: 10.6041/j.issn.1000-1298.2013.04.012. (in Chinese)
[23] 陈海涛,查韶辉,顿国强,等. 2BMFJ 系列免耕精量播种机清秸装置优化与试验[J/OL]. 农业机械学报,2016,47(7):96 - 102.
CHEN Haitao, ZHA Shaohui, DUN Guoqiang, et al. Optimization and experiment of cleaning device of 2BMFJ type no-till precision planter[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(7): 96 - 102. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no = 20160714&flag = 1. DOI:10.6041/j.issn.1000-1298.2016.07.014. (in Chinese)
[24] 顿国强. 大垄密植原茬地免耕精量播种机关键技术研究[D]. 哈尔滨:东北农业大学,2013.
DUN Guoqiang. Key technology research of no-tillage precision planter on grand ridge planting with original stubble[D]. Harbin: Northeast Agricultural University, 2013. (in Chinese)
[25] 哈尔滨工业大学理论力学教研室. 理论力学(Ⅱ)[M]. 北京:高等教育出版社,2009.
[26] 侯守印,陈海涛. 立式轴流大豆育种脱粒机参数优化[J]. 农业工程学报,2012,28(5):19 - 25.
HOU Shouyin, CHEN Haitao. Parameters optimization of vertical axial flow thresher for soybean breeding[J]. Transactions of the CSAE, 2012, 28(5): 19 - 25. (in Chinese)
[27] GB/T 20865—2007　免耕施肥播种机[S]. 2007.

(上接第 35 页)

[19] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510 - 4520.
[20] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv: 1409.1556, 2014.
[21] YU F, KOLTUN V, FUNKHOUSER T A. Dilated residual networks[C]// CVPR, 2017, 2: 3.
[22] 孙俊,何小飞,谭文军,等. 空洞卷积结合全局池化的卷积神经网络识别作物幼苗与杂草[J]. 农业工程学报,2018,34(11):159 - 165.
SUN Jun, HE Xiaofei, TAN Wenjun, et al. Recognition of crop seedling and weed recognition based on dilated convolution and global pooling in CNN[J]. Transactions of the CSAE, 2018, 34(11): 159 - 165. (in Chinese)
[23] MEHTA S, RASTEGARI M, CASPI A, et al. ESPNet: efficient spatial pyramid of dilated convolutions for semantic segmentation[J]. arXiv preprint arXiv:1803.06815, 2018.
[24] XU B, WANG N, CHEN T, et al. Empirical evaluation of rectified activations in convolutional network[J]. arXiv preprint arXiv:1505.00853, 2015.
[25] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. Computer Science, 2012, 3(4): 212 - 223.
[26] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[J]. arXiv preprint arXiv:1502.03167, 2015.
[27] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint arXiv:1704.04861, 2017.
[28] ABADI M, BARHAM P, CHEN J, et al. Tensorflow: a system for large-scale machine learning[C]// OSDI, 2016, 16: 265 - 283.
[29] PEZZEMENTI Z, TABOR T, HU P, et al. Comparing apples and oranges: off-road pedestrian detection on the national robotics engineering center agricultural person-detection dataset[J]. Journal of Field Robotics, 2018, 35(4): 545 - 563.
[30] SHRIVASTAVA A, GUPTA A, GIRSHICK R. Training region-based object detectors with online hard example mining[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 761 - 769.
[31] DOLLÁR P, WOJEK C, SCHIELE B, et al. Pedestrian detection: an evaluation of the state of the art[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(4): 743 - 761.