

基于深度学习的模糊农田图像中障碍物检测技术

薛金林 李雨晴 曹梓建

(南京农业大学工学院, 南京 210031)

摘要: 针对图像实时采集时,由于镜头缺陷、相机抖动、目标运动等原因造成的模糊图像输入,导致训练完成的深度学习模型检测准确率下降问题,本文提出一种基于改进 Faster R-CNN 和 SSRN-DeblurNet 的两阶段检测方法,用于农田环境模糊图像中的障碍物检测。第1阶段进行锐度评价和去模糊处理,利用简化尺度循环网络(Simplified scale recurrent networks, SSRN-DeblurNet)对模糊农田图像进行去模糊。第2阶段进行障碍物检测,在原有的 Faster R-CNN 网络中添加了候选区域优化网络来提高区域候选网络中的目标区域质量。在自制的模糊数据集上,利用所提出的两阶段检测方法对8种农田障碍物进行检测。与原始 Faster R-CNN 相比,两阶段检测方法的平均精度均值(mAP)提高了12.32个百分点,单幅图像的平均检测时间为0.53 s。所提出的两阶段方法能有效减少模糊农田图像中障碍物的误检和漏检,满足拖拉机低速作业的实时检测需求。

关键词: 农田; 模糊图像; 障碍物检测; 去模糊; Faster R-CNN; SSRN-DeblurNet

中图分类号: S24; TP391.4

文献标识码: A

文章编号: 1000-1298(2022)03-0234-09

OSID:



Obstacle Detection Based on Deep Learning for Blurred Farmland Images

XUE Jinlin LI Yuqing CAO Zijian

(College of Engineering, Nanjing Agricultural University, Nanjing 210031, China)

Abstract: When it is in real-time image acquisition, image blurring caused by lens defects, camera jitter, target movement and so on will result in poor precision of target detection by using the trained deep learning model. Here, a two-stage detection model based on an improved Faster R-CNN and an SSRN-DeblurNet was proposed to perform obstacle detection for blurred farmland images. In the first stage, sharpness evaluation and deblurring were carried out, and the simplified scale recurrent networks (SSRN-DeblurNet) was used for deblurring of blurred farmland images. In the second stage, obstacle detection was implemented by using the improved Faster R-CNN which was added a proposal region optimization network to improve the quality of the regions in the region proposal networks. Then, the proposed two-stage detection model was used to detect eight types of farmland obstacles with self-made blurred dataset. Compared with the original Faster R-CNN, the mean average precision (mAP) value was increased by 12.32 percentage points, and the average detection time of a single image was 0.53 s. The results showed that the proposed two-stage model can not only effectively reduce the false detection and missing detection of obstacles in blurred farmland images, but also can meet the real-time detection requirements of tractors operating at low speeds.

Key words: farmland; blurred images; obstacle detection; deblurring; Faster R-CNN; SSRN-DeblurNet

0 引言

在农业机器人或者无人农业车辆行进路线上不可避免地存在大量障碍物,包括树木、草堆、房屋、电线杆、人、牛、羊和其他农机,如果不通过传感器实时

感知并及时避开,会造成严重的安全问题。因此,农业机器人或者无人农业车辆应该具有检测和识别障碍物的能力^[1]。

针对农田障碍物,传统检测方法采用立体视觉技术^[2]、小波多分辨率分解^[3]、激光扫描定位^[4-5]、

收稿日期: 2021-03-17 修回日期: 2021-05-28

基金项目: 江苏省科技计划项目(BK20151436)和江苏高校“青蓝工程”项目

作者简介: 薛金林(1974—),男,教授,博士,主要从事农业车辆测控技术和智能化研究,E-mail: xuejinlin@njau.edu.cn

HOG 特征与支持向量机^[6]等来进行。传统方法存在的缺陷有：只进行障碍物定位而不对障碍物进行识别分类，这不利于后续精准的路径规划和避障；检测障碍物的种类和数量有限；根据人的经验选取特征，如果选取的特征不足以表征目标障碍物，则会影响识别效果。深度学习的兴起为农田障碍物检测识别带来了新思路。薛金林等^[7]针对农田作业场景中的人和其他农业车辆等动态障碍物提出了一种基于卷积神经网络的农业自主车辆多种类障碍物分类识别方法。刘慧等^[8]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 SSD 方法。杨娟娟^[9]提出了一种结合区域提名与无需区域提名的新型智能目标检测算法。农田环境的地形通常是凹凸不平的，视觉传感器需要在不稳定的状态中采样，以上研究均未考虑不稳定状态采样造成的模糊图像的输入对深度学习检测模型的影响，这势必造成实际检测效果达不到预期水平。

在检测模型的训练阶段，所使用的数据集通常是在一个稳定的、高质量的环境中采集的。然而，当

农机在田间作业时，由于相机镜头缺陷、相机抖动、目标运动等原因，会获得一些模糊图像，它们会改变图像中包含的重要信息，导致训练后的模型在推理阶段检测精度较差。因此，本文研究一种基于改进 Faster R-CNN 和 SSRN-DeblurNet 的两阶段检测方法，来解决模糊图像输入导致的深度学习检测模型检测精度下降、漏检和误检问题。

1 两阶段检测方法原理

两阶段检测模型的网络结构如图 1 所示，整个网络分为 4 部分：去模糊网络、骨架网络、区域候选网络、分类回归网络。去模糊网络由不同尺度的全卷积子网络构成，主要作用是对锐度 S 低于阈值 S_{Thr} 的图像去模糊以恢复图像质量；骨架网络由去除全连接层的 ResNet50 网络构成，主要作用是对图像进行特征提取；区域候选网络由卷积层、池化层和全连接层构成，主要作用是生成优化候选区域；分类回归网络由池化层和全连接层组成，主要作用是进行目标识别和精准定位。

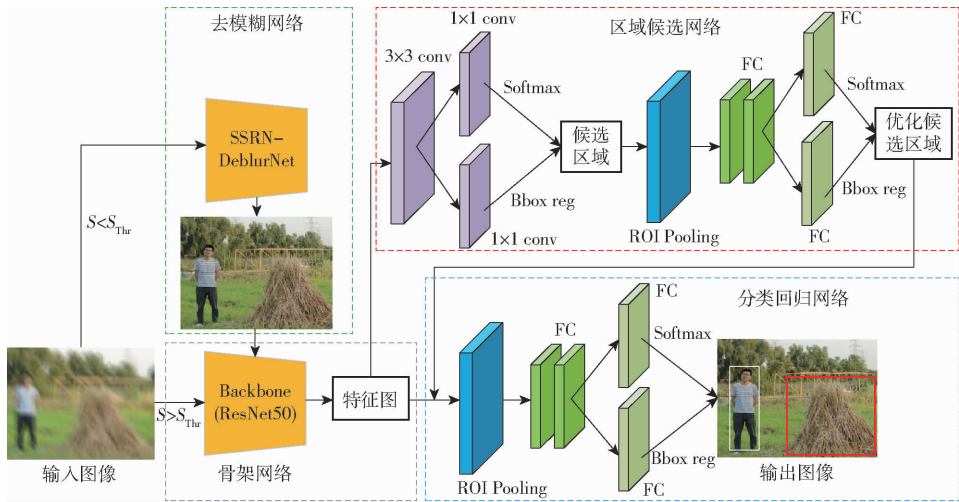


图 1 两阶段检测模型网络结构图

Fig. 1 Network structure of two-stage detection model

整个检测模型的工作流程分为 2 个阶段。第 1 个阶段是基于 SSRN-DeblurNet 去模糊网络进行去模糊，首先对输入图像进行锐度评价，如果输入图像的锐度 S 低于设定的锐度阈值 S_{Thr} ，则图像被送入 SSRN-DeblurNet 网络中进行去模糊处理，去模糊后的图像被送入第 2 阶段进行检测；如果输入图像的锐度 S 高于设定的锐度阈值 S_{Thr} ，则图像直接被送入第 2 阶段进行检测。第 2 阶段是基于改进的 Faster R-CNN 网络检测图像中的农田障碍物。首先利用骨架网络对输入图像进行特征提取，生成的特征图用于区域候选网络 (Region proposal network, RPN) 和分类回归网络。RPN 网络级联 1 个候选区域优化网络用于过滤低质量的候选区域，提升检测

的精度。利用 RPN 网络生成的优化候选区域和骨架网络最后一层得到的特征图，得到固定尺寸的候选区域特征图。最后利用 Softmax 函数进行目标类别的确定并利用边界框回归获得目标的精确位置。

1.1 基于 SSRN-DeblurNet 的去模糊阶段

1.1.1 锐度阈值

清晰图像比模糊图像的边缘更加锐利清晰，边缘像素灰度变化大，因而有更大的梯度。因此，可以利用梯度函数获取图像梯度信息，从而评判图像的锐度。梯度函数主要包括 Brenner 梯度函数^[10]、Tenengrad 梯度函数^[11]、Laplacian 梯度函数^[12]、SMD 函数^[13]、Variance 函数等^[14]。

图像锐度评价算法的性能通过人眼主观评分和算法预测评分的相关性进行评价^[15]。预测准确性和单调一致性是衡量评价算法性能的 2 个常用标准, 通过以下参数体现: 皮尔森线性相关系数 (Pearson linear correlation coefficient, PLCC)、肯德尔秩次相关系数 (Kendallrank order correlation coefficient, KROCC)、斯皮尔曼等级相关系数 (Spearman rank order correlation coefficient, SRCC)、均方根误差 (Root mean square error, RMSE)^[16]。如果以上 3 个相关系数的值越接近于 1, RMSE 越小, 说明锐度评价算法越准确。此外, 本文将单幅图像检测时间也纳入评价参数范围, 单幅图像检测时间是随机选取的 500 幅图像的平均检测时间。基于以上 5 个参数进行图像锐度评价函数的选择。

在 ITU - R BT. 500 图像质量主观评价方法的五级损伤量表^[17]的基础上建立图像锐度的主观评分标准, 本文只需要 2 个尺度对需要去模糊和不需要去模糊的图像进行评价。因此, 主观评分标准为: 当评价人认为图像的模糊程度可以察觉但并不讨厌时, 该图像的主观得分为 2; 当评价人认为图像的模糊程度讨厌时, 该图像的主观得分为 1。

从数据集中随机抽取 500 幅图像, 同时用 4 种不同的梯度函数预测锐度以及进行人眼主观评分。4 种不同的梯度函数性能如图 2 所示。

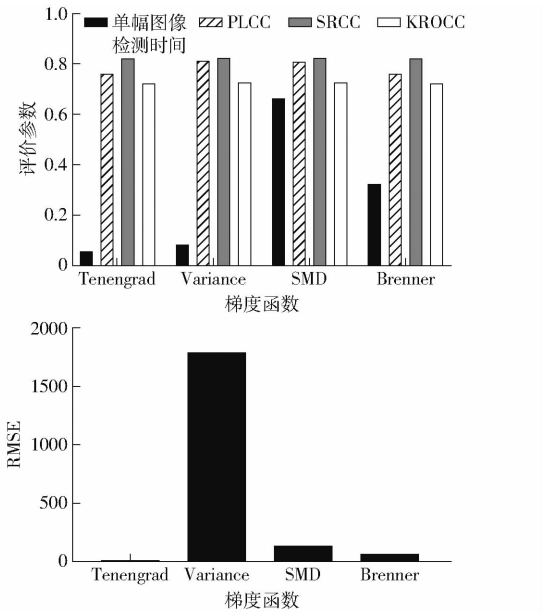


图 2 4 种梯度函数的性能

Fig. 2 Performance of four gradient functions

由图 2 可以看出, 采用 Tenengrad 梯度函数作为锐度评价函数, 检测速度快、均方根误差最小、与主观评价方法相关性较大。Tenengrad 梯度函数采用 Sobel 算子分别提取水平和垂直方向的梯度^[18], 基于 Tenengrad 的图像锐度定义为

$$S = \sum_y \sum_x | \sqrt{G_x^2(x,y) + G_y^2(x,y)} | \quad (G(x,y) > T) \tag{1}$$

其中 $G(x,y) = \sqrt{G_x^2(x,y) + G_y^2(x,y)}$
式中 T ——给定的边缘检测阈值, 对应图像最大梯度的 0.5 倍
 G_x ——像素点 (x,y) 处 Sobel 水平方向边缘检测算子的卷积
 G_y ——像素点 (x,y) 处 Sobel 垂直方向边缘检测算子的卷积

锐度阈值的选择是基于测试集上的人眼主观评分与算法预测分数分布, 需满足检测效率要求和准确性的要求。对测试集进行主观评分后, 利用 Tenengrad 梯度函数对测试集的图像进行锐度预测, 根据人眼主观评价分与算法预测分的分布确定锐度阈值。

如图 3 所示, 当阈值为 15, 有些图像的算法预测得分较高, 但主观感觉严重模糊。当阈值为 20 时, 图像的预测得分与主观得分基本一致。当阈值为 25 时, 对相对清晰的图像去模糊浪费时间和计算资源。因此, 选择锐度阈值为 20 是合理而高效的。

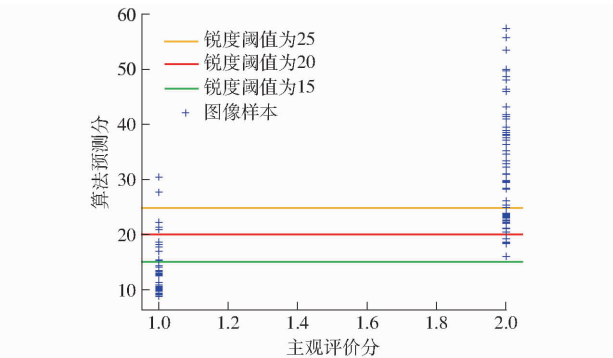


图 3 主观评价分和算法预测分的分布

Fig. 3 Distribution of subjective evaluation score and algorithm prediction score

1.1.2 SSRN - DeblurNet 去模糊网络

引入去模糊网络 (Scale recurrent network, SRN - DeblurNet)^[19], 进行模型精简后得到 SSRN - DeblurNet 进行图像的去模糊操作。SRN - DeblurNet 的网络结构如图 4 所示。

SRN - DeblurNet 由 3 层代表不同尺度输入且结构相同的全卷积子网组成, 3 层子网之间参数共享, 减少可训练参数的数量。每层子网分为编码部分和解码部分, 编码部分包含 3 个编码模块 (EBlock), 如图 4b 所示, 每个 EBlock 由 1 个卷积层 (Conv) 和 3 个残差块 (ResBlock) 组成。解码部分与编码部分结构对称, 对应 3 个解码模块 (DBlock), 如图 4c 所示,

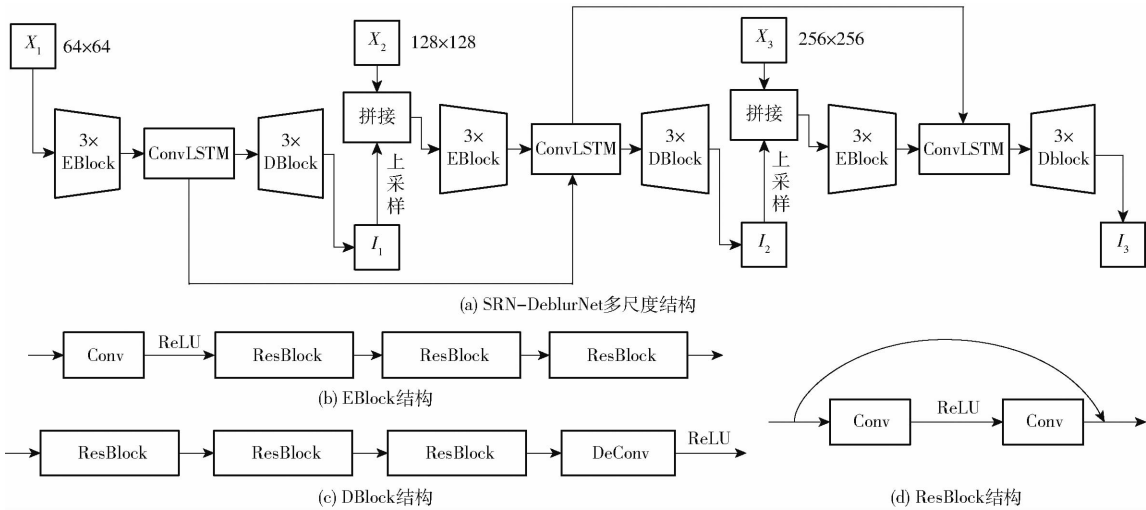


图 4 SRN-DeblurNet 的网络结构

Fig. 4 Structure of SRN-DeblurNet

每个 DBlock 由 3 个 ResBlock 和 1 个转置卷积层 (DeConv) 组成。ResBlock 的结构如图 4d 所示,通过跳连接实现不同层的特征信息共享。此外,每 1 层还有 1 个循环卷积网络 (ConvLSTM),将不同尺度的图像特征看作时序与局部空间均相关的序列进行处理。训练过程中,随机裁剪 256 像素 × 256 像素图像块 (X_3),然后用最近邻插值方法得到 64 像素 × 64 像素分辨率子图 (X_1) 和 128 像素 × 128 像素分辨率子图 (X_2)。上一层网络的输出作为下一层网络的输入,由粗到精,由小到大复原出清晰图像,最后一层网络的输出 I_3 作为最终的复原结果。

对于每个尺度的损失 L , 计算式为

$$L = \sum_{i=1}^n \frac{k_i}{N_i} \| I^i - I_*^i \|^2 \quad (2)$$

式中 I^i ——第 i 个尺度上的网络输出
 I_*^i ——第 i 个尺度上的真值
 k_i ——每个尺度的权重,初始权重设置为 1
 N_i —— I^i 中需标准化的元素个数

SRN-DeblurNet 存在参数过多,在推理阶段运行速度较慢的问题,在保证模型精度的情况下,希望该模型参数量少、占用内存少、速度快、计算量低,所以本文对 SRN-DeblurNet 去模糊模型进行了精简。深度学习模型的精简方式一般有网络剪枝、知识蒸馏、参数量化等^[20]。本文采用参数量化对 SRN-DeblurNet 进行模型压缩加速,得到精简模型 SSRN-DeblurNet。

对于参数量化对象,主要有:权重、激活函数输出、梯度^[21]。量化位数可分为:16 位、8 位、8 位以下。本文采用的 8 位量化可带来 4 倍的模型压缩、4 倍的内存带宽提升以及更高效的缓存利用。此外,计算速度也通常有 2 ~ 3 倍的提升,且对精度不会造成较大损失。按量化过程中是否需要进行训练

可分为以下 2 类^[22]:在模型训练后做量化 (Post training quantization, PTQ)、在模型训练时做量化 (Quantization aware training, QAT)。一般情况下, QAT 可以得到更高的准确率,但同时需要有训练数据和训练环境。

深度学习模型压缩流程如图 5 所示。基于 TensorFlow 工具箱中的 graph_transforms, 本文对训练好的 SRN-DeblurNet 网络的权重进行 32 位转 8 位的 PTQ 量化操作,并删除一些在推理阶段不使用的计算节点,合并一些重复计算节点,得到一个压缩率与准确率都满足实际需求的精简模型 SSRN-DeblurNet。

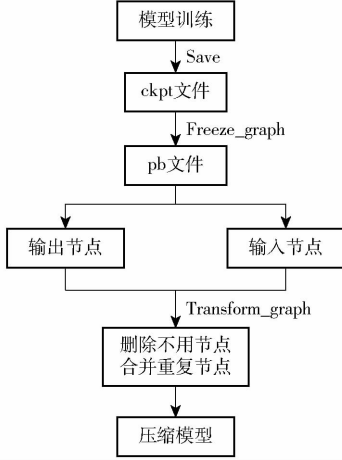


图 5 深度学习模型压缩流程图

Fig. 5 Compression process of deep learning model

1.2 基于改进 Faster R-CNN 的目标检测阶段

ROSS 在 R-CNN 和 Fast R-CNN 基础上提出了 Faster R-CNN^[23],其主要由骨架网络、区域候选网络、分类回归网络组成。区域候选网络为每幅输入图像输出大约 300 个候选区域,其中存在大量与障碍物真值框重叠率较低的低质量候选区域,这给

Faster R-CNN 网络增加了许多无用的计算。为了减少低质量候选区域,本文在区域候选网络后级联了一个候选区域优化网络。该优化网络主要由 ROI 池化层和 4 个全连接层组成。候选区域优化网络以候选区域生成网络输出的候选区域作为输入;对候选区域所包含的特征进行池化,生成池化特征图;利用全连接层 FC1 和 FC2 对池化特征图进行特征映射;利用全连接层 FC3 根据映射后的特征输出该候选区域内是否包含待测目标的分数,利用 Softmax 函数对分数进行归一化,从而得到候选区域内是否包含待测目标的置信度;全连接层 FC4 通过边界框回归输出边界框的坐标偏移量,从而修正候选区域的位置。将候选区域内包含待测目标的置信度与设置的前景过滤阈值(0.8)和背景过滤阈值(0.2)进行比较,输出置信度大于前景阈值的正样本和置信度小于背景阈值的负样本。候选区域优化网络的结构如图 6 所示。

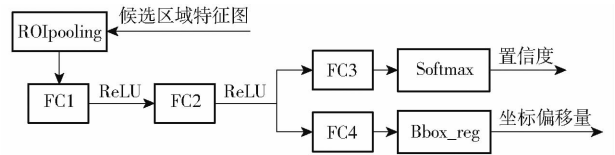


图 6 候选区域优化网络
Fig. 6 Region proposal optimization network

候选区域优化网络的损失函数为

$$L(p,u)=\begin{cases}-\lambda\lg(1-p) & (u=0) \\ -\lambda\lg p & (u=1)\end{cases}\quad (3)$$

式中 p ——候选区域内包含待测目标的置信度
 u ——候选区域的标签,正样本时为 1,否则为 0
 λ ——损失函数的损失权重,初始设置为 1

2 试验

2.1 图像采集

图像采集设备为 QR-USB500W02M-V60 型 USB 高清数码摄像头,分辨率为 640 像素×480 像素时,帧率为 30 f/s,实际使用时为 20 f/s 左右。将摄像头安装在茂源 MY250 型轮式拖拉机上,摄像头与

地面的垂直距离为 1.33 m。摄像头通过以太网将接收到的图像传送到计算机上。摄像头的安装位置如图 7 所示。

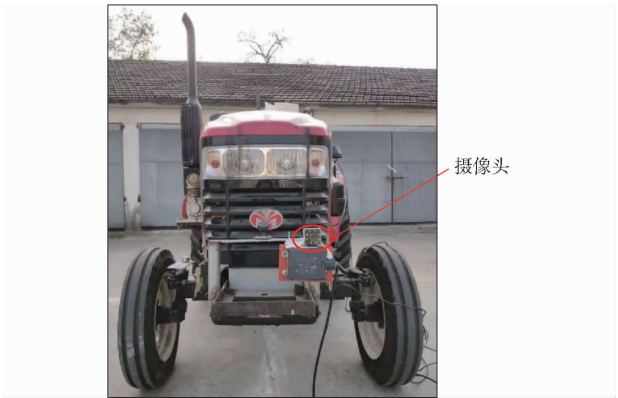


图 7 摄像头安装位置
Fig. 7 Installation location of camera

将收集到的 2 370 幅图像按比例 7:2:1 分为训练集、验证集、测试集。利用标注软件对图像中的人、拖拉机、草堆、电线杆、树、房子、羊、牛 8 种农田障碍物进行标注。为最大限度模拟农田环境下图像采集可能遇到的模糊类型,模糊数据集由 GOPRO 数据集与自制模糊数据集组成,共 5 347 个清晰和模糊图像对。其中自制模糊数据集基于马尔可夫过程随机生成相机运动轨迹和模糊内核^[24],对图像进行处理得到,如图 8 所示。

2.2 试验平台

试验平台为台式计算机,处理器为 Intel Core i7-8700k,主频为 3.70 GHz,32 GB 内存,250 GB 固态硬盘,显卡为 GeForce GTX1060。运行环境在 Ubuntu 16.04 LTS 下,编程语言为 Python 3.6,模型训练均在 TensorFlow 深度学习框架下进行,并行计算框架和深度学习 GPU 加速库分别为 CUDA 9.0 和 cuDNN 7.4.2。

2.3 SSRN-DeblurNet 去模糊网络训练

训练去模糊网络 SRN-DeblurNet 时,在每次迭代过程中,随机剪裁 256 像素×256 像素的模糊图像块作为输入,模糊图像块每个通道除以 255 进行数据归一化。尺度设为 3,批量大小设为 8,训练轮



图 8 自制模糊数据集
Fig. 8 Self-made blurring dataset

数设为 2 000 轮,动量因子设置为 0.9。采用 Adam 优化器,参数 $\beta_1=0.9,\beta_2=0.999,\varepsilon=10^{-4}$ 。利用随机梯度下降法训练网络,初始学习率设置为 0.001。训练参数初始化方式采用 Xavier 方法。

2.3.1 去模糊结果评价指标

以峰值信噪比(Peak signal to noise ratio,PSNR)和结构相似度(Structural similarity,SSIM)作为去模糊网络输出结果的评价指标^[25]。

PSNR 是一个表示信号最大可能功率和噪声功率比值的工程术语,单位是 dB,一般取值范围为 20~40 dB,数值越大表示失真越小。SSIM 分别从亮度、对比度、结构三方面度量图像相似性,取值范

围为 0~1,其值越大,表示图像失真越小。

2.3.2 去模糊结果

SSRN-DeblurNet 与 SRN-DeblurNet 的参数对比如表 1 所示,包括参数量、单幅图像去模糊时间、PSNR 和 SSIM。图 9 为 SSRN-DeblurNet 与 SRN-DeblurNet 的去模糊效果对比。

表 1 SSRN-DeblurNet 与 SRN-DeblurNet 的参数对比
Tab.1 Comparison of indicators between SSRN-DeblurNet and SRN-DeblurNet

模型	参数量/MB	时间/s	PSNR/dB	SSIM
SRN-DeblurNet	27.60	0.85	28.59	0.83
SSRN-DeblurNet	7.00	0.27	26.42	0.75



图 9 SSRN-DeblurNet 与 SRN-DeblurNet 去模糊效果对比

Fig.9 Comparison of deblurring between SSRN-DeblurNet and SRN-DeblurNet

从图 9 可以看出,相比于原始模糊图,SSRN-DeblurNet 恢复了部分图像质量。但恢复效果比 SRN-DeblurNet 差。如表 1 所示,SSRN-DeblurNet 与 SRN-DeblurNet 相比,PSNR 下降 7.59%,SSIM 下降 9.64%。PSNR 和 SSIM 小幅度下降是因为 SSRN-DeblurNet 的权值数值类型是 8 位而不是 32 位,权重值更稀疏必然会导致量化过程中数值精度的损失,进而影响 SSRN-DeblurNet 的去模糊效果。虽然去模糊的效果不及 SRN-DeblurNet,但 SSRN-DeblurNet 在推理速度上有显著提高,不仅参数量减少到原模型的 25.36%,检测时间也减少到原模型的 31.76%,实现了精度和推理速度之间的平衡。

2.4 改进 Faster R-CNN 网络训练

在训练阶段,Faster R-CNN 网络与 SSRN-DeblurNet 网络相互独立进行训练。为了检验候选区域优化网络的效果,将所有损失函数的权值设为 1。使用 ImageNet 预训练模型对改进 Faster R-CNN 的权重进行初始化,训练 50 000 轮,初始学习率为 0.002 5,权重衰减为 0.000 1。

2.4.1 目标检测结果评价指标

本文以平均精度均值(mAP)和平均精度(AP)作为目标检测结果评价指标。

当拖拉机在水泥平地处于启动状态时,在相机采集到的连续 7 帧图像中,如图 10a 所示,某帧清晰图像的锐度约为 42.73,如图 10b 所示,某帧轻微模糊图像的锐度约为 38.96,说明拖拉机的振动对相机成像质量有一定影响。假设拖拉机处于以下状态:在农田作业时遇到起伏较大的坑洼地带;镜头中的目标障碍物快速运动;相机未能成功对焦时,轻微模糊图像出现的时间点将有更大的概率生成锐度低于 20 的模糊图像。因此以 7 帧图像作为 1 个周期计算单幅图像的推理时间,计算方式为



图 10 拖拉机振动对图像质量的影响

Fig.10 Influence of tractor vibration on image quality

$$t \approx \frac{t_i + t_{i+1} + \cdots + t_{i+6}}{7} \quad (4)$$

式中 $t_i、t_{i+1}、\cdots、t_{i+6}$ ——第 $i、i+1、\cdots、i+6$ 帧图像的推理时间

2.4.2 改进 Faster R - CNN 检测结果

对无优化网络和有优化网络时的候选区域生成结果进行可视化,如图 11 所示。

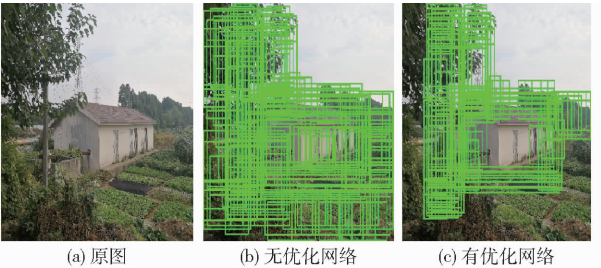


图 11 生成候选区域比较图
Fig. 11 Comparison of proposal regions

图 11a 为输入网络前的原图,图 11b 为原图输入原始 Faster R - CNN 网络时 RPN 网络输出的候选区域。从图 11b 可以看出,RPN 网络将房屋前的农作物、房屋左侧的杂草和树叶都作为目标障碍物生成了候选区域。但图像中只有房屋和电线杆是目标障碍物,因此无优化网络的 RPN 网络生成的这些候选区域是低质量的,这些低质量的候选区域将为分类回归网络增加大量无用计算,降低整体网络的收敛速度。图 11c 为在 RPN 网络后级联了一个优化网络后原始图像输入网络生成的优化候选区域。由于设置了更高的正样本过滤阈值 0.8,优化后的 RPN 网络将房屋周围的杂草、农作物和树叶等作为目标障碍物生成的候选区域进行了滤除,保留的候选区域边界框更集中于目标障碍物的真值边界框附近,这有利于网络快速收敛和提升网络的检测精度。从表 2 可以看出,改进后的 Faster R - CNN 网络对各种农田障碍物的平均精度平均提高了 1.93 个百分点,这说明候选区域优化网络的加入能提升检测精度。

表 2 8 种农田障碍物的平均精度

Tab. 2 Average precision of eight kinds of farmland obstacles		
	obstacles	
	%	
障碍物类别	无优化网络	有优化网络
人	80.34	82.11
草堆	80.26	81.29
电线杆	74.56	75.90
拖拉机	88.41	90.85
羊	66.55	69.46
房屋	54.11	56.07
牛	65.98	67.62
树	60.18	62.53

2.5 推理阶段检测结果

为了验证改进检测模型的性能,在推理阶段比较以下 5 种目标检测方法:原始 Faster R - CNN 网络、改进 Faster R - CNN 网络、YOLO v5s 网络、改进 Faster R - CNN 与 SRN - DeblurNet 的两阶段检测模型和改进 Faster R - CNN 与 SSRN - DeblurNet 的两阶段检测模型。利用 AP、mAP、单幅图像推理时间对检测结果进行综合评价。

IoU 阈值为 0.5 时,不同方法的各类障碍物的平均精度如表 3 所示。

表 3 不同方法的各类障碍物平均精度
Tab. 3 Average precision of different algorithms for eight kinds of farmland obstacles %

障碍物类别	原始	改进		改进 Faster	改进 Faster
	Faster	Faster	YOLO v5s	R - CNN 与	R - CNN 与
	R - CNN	R - CNN	网络	SRN	SSRN
	网络	网络		DeblurNet	DeblurNet
人	60.46	63.08	54.17	80.33	76.19
草堆	59.01	60.45	55.34	76.04	73.42
电线杆	55.27	55.19	50.09	71.57	69.58
拖拉机	61.62	64.83	52.50	82.19	81.05
羊	47.33	50.14	43.26	57.44	56.58
房屋	42.76	43.42	39.78	56.89	51.77
牛	45.40	46.46	40.07	55.72	53.97
树	42.21	45.82	39.28	54.50	50.11

如表 3 所示,不同方法对人、草堆、电线杆、拖拉机这 4 类目标障碍物的检测效果较好,检测的平均精度在 50.00% 以上;不同方法对羊、房屋、牛、树这 4 类目标障碍物的检测效果较差。主要原因在于数据集的样本分布不均,后 4 种目标障碍物的训练样本数量偏少,出现过拟合现象,在推理阶段表现为训练样本偏少的目标障碍物的平均检测精度较低。由于采用的测试集不同,与表 2 相比, Faster R - CNN 与改进 Faster R - CNN 的平均精度均大幅降低,这进一步证实了模糊图像的输入会导致训练好的深度学习检测模型的检测精度降低。

IoU 阈值为 0.5 时,改进 Faster R - CNN 由于增加了一个候选区域优化网络,其 mAP 较原始 Faster R - CNN 提升了 1.91 个百分点。对于结合了去模糊网络的两阶段检测模型, mAP 都有了很大的提升,其中改进 Faster R - CNN 与 SRN - DeblurNet 的两阶段检测模型的 mAP 最高,为 66.84%。两阶段检测方法由于添加了去模糊网络,提高了严重模糊图像的质量,因此大大减少了漏检和误检的情况,从而提高了 mAP,与原始 Faster R - CNN 相比,改进

Faster R - CNN 与 SSRN - DeblurNet 的 mAP 提高了 12.32 个百分点。出于实时性因素的考虑,对 SRN - DeblurNet 网络进行了量化压缩,以提升检测速度,但也因此造成了改进 Faster R - CNN 与 SSRN - DeblurNet 的两阶段检测模型的 mAP 降低了 2.76 个百分点。

由于增加了候选区域优化网络,改进 Faster R - CNN 相比于原始 Faster R - CNN 的单幅图像推理时间增加了 0.19 s。不同组合的两阶段检测模型的单幅图像推理时间均较长,主要原因是去模糊操作需要占用大量时间和计算资源。为此对去模糊网络 SRN - DeblurNet 进行了模型压缩,降低去模糊操作占用的时间,从而改进 Faster R - CNN 与 SSRN - DeblurNet 的单幅图像推理时间较改进 Faster R - CNN 与 SRN - DeblurNet 的时间降低了 0.09 s,单幅图像平均推理时间为 0.53 s。

拖拉机根据其负载水平可将其作业工况分为重负荷作业、中等负荷作业、轻负荷作业及道路运输作业等^[26]。其中,拖拉机处于重负荷和中等负荷作业工况的速度覆盖 2 ~ 14 km/h,且拖拉机大部分时间都在此速度范围作业;拖拉机处于轻负荷和道路运输作业工况的速度覆盖 14 ~ 30 km/h。因此,改进 Faster R - CNN 与 SSRN - DeblurNet 的两阶段检测方法基本满足拖拉机低速作业的实时性要求。

当输入锐度低于 20 的模糊图像时,原始 Faster R - CNN 与两阶段检测方法检测结果如图 12 所示。

利用原始 Faster R - CNN 方法对 4 幅锐度低于 20 的测试图像进行检测(图 12a),这 4 幅图像的前 2 幅是人工生成的,后 2 幅是在农田中实际采集的。在第 1 幅测试图像中,草堆的置信度为 0.95,草堆左边的 2 人均被漏检;在第 2 幅测试图像中,在草堆的位置同时检测出牛和草堆这 2 种目标障碍物,其中,牛的置信度为 0.63,草堆的置信度为 0.61;在第 3 幅和第 4 幅实际采集的测试图像中,人作为目标障碍物均被漏检。

采用改进 Faster R - CNN 与 SSRN - DeblurNet 的两阶段检测方法时(图 12b),锐度均低于 20 的图像被输入去模糊网络中进行去模糊后再进行检测。第 1 幅测试图像中的 2 人均被检出,且置信度均为 1;第 2 幅测试图像中草堆被正确检出,置信度为 0.99;第 3 幅和第 4 幅测试图像中的人均被检出,

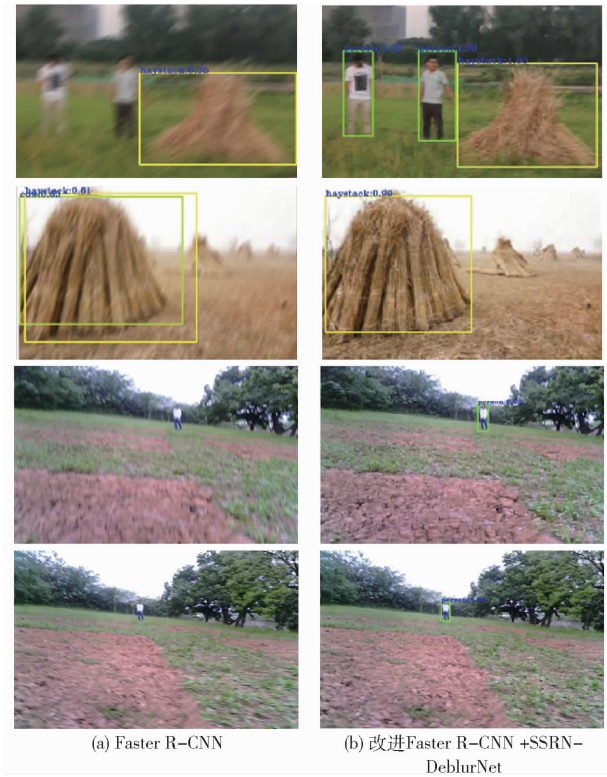


图 12 Faster R - CNN 与两阶段检测方法检测结果对比
Fig. 12 Comparison of results between Faster R - CNN and two-stage detection method

置信度均为 0.99。因此,本文提出的两阶段检测方法可以很好地解决由于模糊图像的输入造成的漏检和误检。

3 结论

(1)为实现模糊农田图像中的障碍物检测,提出了一种基于改进 Faster R - CNN 和 SSRN - DeblurNet 的两阶段农田障碍物检测方法。试验表明,与原始 Faster R - CNN 相比,两阶段检测方法的 mAP 提高了 12.32 个百分点,在模糊图像输入情况下很好地缓解了漏检和误检情况;单幅图像平均推理时间为 0.53 s,满足拖拉机在重负荷和中等负荷作业工况的实时性需求。

(2)由于现有的数据集样本不均匀,进一步需建立更完整、数据量更大的农田障碍物数据集。为满足拖拉机在轻负荷以及道路运输工况时的实时性需求,进一步需将 YOLO v5s 和 SSRN - DeblurNet 融合进一个框架,进行端到端训练以进一步降低模型复杂度和提升效率。

参 考 文 献

[1] MWALUPASO G E, WANG S, RAHMAN S, et al. Agricultural informatization and technical efficiency in maize production in Zambia[J]. Sustainability, 2019, 11(8): 2451.
[2] WANG Qiao, MENG Zhijun, LIU Hui. Review on application of binocular vision technology in field obstacle detection[C]// Proceedings of International Conference on AI and Big Data Application(AIBDA 2019) , 2019: 13.

- [3] 韩永华, 汪亚明, 康锋, 等. 基于小波多分辨率分解的农田障碍物检测[J]. 农业机械学报, 2013, 44(6): 215–221.
HAN Yonghua, WANG Yaming, KANG Feng, et al. Farmland obstacle detection based on wavelet multiresolution decomposition[J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(6): 215–221. (in Chinese)
- [4] 刘美辰. 农机作业障碍物检测与避障方法研究[D]. 杨凌: 西北农林科技大学, 2018.
LIU Meichen. Research on obstacle detection and obstacle avoidance method in agricultural machinery operation[D]. Yangling: Northwest A&F University, 2018. (in Chinese)
- [5] 袁文涛, 刘卉, 胡书鹏. 面向自动导航拖拉机的农田障碍物识别研究[J]. 农机化研究, 2018, 40(10): 247–251.
YUAN Wentao, LIU Hui, HU Shupeng. Research on farmland obstacle recognition for automatic navigation tractor[J]. Journal of Agricultural Mechanization Research, 2018, 40(10): 247–251. (in Chinese)
- [6] USPENSKIY I, REMBALOVICH G, YUKHIN I, et al. Development and testing of a conveyor for detecting various types of vehicles when transporting agricultural products from the field[J]. IOP Conference Series: Materials Science and Engineering, 2020, 832(1): 012059.
- [7] 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]. 农业机械学报, 2018, 49(增刊): 42–48.
XUE Jinlin, YAN Jia, FAN Bowen. Convolution neural network classification and recognition method for multi class farmland obstacles[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(Supp.): 42–48. (in Chinese)
- [8] 刘慧, 张礼帅, 沈跃, 等. 基于改进 SSD 的果园行人实时检测方法[J]. 农业机械学报, 2019, 50(4): 29–35, 101.
LIU Hui, ZHANG Lishuai, SHEN Yue, et al. Real time pedestrian detection method in orchard based on improved SSD[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(4): 29–35, 101. (in Chinese)
- [9] 杨娟娟. 基于深度学习的农用无人机避障系统研究[D]. 兰州: 甘肃农业大学, 2019.
YANG Juanjuan. Research on obstacle avoidance system of agricultural UAV based on deep learning[D]. Lanzhou: Gansu Agricultural University, 2019. (in Chinese)
- [10] 毕天华, 杜文华. 一种改进的 Brenner 清晰度评价函数[J]. 电子测量技术, 2019, 42(9): 80–84.
BI Tianhua, DU Wenhua. An improved Brenner clarity evaluation function[J]. Electronic Measurement Technology, 2019, 42(9): 80–84. (in Chinese)
- [11] 黄涛, 方梦瑞, 夏华鹏, 等. 基于清晰度的茶叶嫩芽聚类分割方法[J]. 湖北农业科学, 2020, 59(8): 154–157.
HUANG Tao, FANG Mengrui, XIA Huakun, et al. A cluster segmentation method for tea buds based on definition[J]. Hubei Agricultural Sciences, 2020, 59(8): 154–157. (in Chinese)
- [12] 毕超, 郝雪, 李剑飞, 等. 气膜孔图像对焦评价函数的实验研究[J]. 宇航计测技术, 2019, 39(6): 77–83.
BI Chao, HAO Xue, LI Jianfei, et al. Experimental study on focusing evaluation function of gas film hole image[J]. Astronautical Metrology and Measurement Technology, 2019, 39(6): 77–83. (in Chinese)
- [13] 王竑舒, 徐熙平. 广角相机图像的调焦评价函数研究[J]. 长春理工大学学报(自然科学版), 2019, 42(3): 43–46, 55.
WANG Hongshu, XU Xiping. Research on focusing evaluation function of wide-angle camera image[J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2019, 42(3): 43–46, 55. (in Chinese)
- [14] 王灿芳, 崔良玉, 阎兵. 基于 Variance–Brenner 函数的显微图像清晰度评价算法研究[J]. 装备制造技术, 2020(10): 78–82.
WANG Canfang, CUI Liangyu, YAN Bing. Research on clarity evaluation algorithm of microscopic image based on Variance–Brenner function[J]. Equipment Manufacturing Technology, 2020(10): 78–82. (in Chinese)
- [15] 曹欣, 李战明, 胡文瑾. 一种彩色图像质量评价方法[J]. 应用科学学报, 2019, 37(3): 398–406.
CAO Xin, LI Zhanming, HU Wenjin. A method of color image quality evaluation[J]. Journal of Applied Sciences, 2019, 37(3): 398–406. (in Chinese)
- [16] 史晨阳, 林燕丹. 基于图像色貌和梯度特征的图像质量客观评价[J]. 物理学报, 2020, 69(22): 401–412.
SHI Chenyang, LIN Yandan. Objective evaluation of image quality based on color appearance and gradient features[J]. Acta Physical Sinica, 2020, 69(22): 401–412. (in Chinese)
- [17] UNION I T. Methodology for the subjective assessment of the quality of television pictures[S]. ITU–R Recommendation BT, 2002.
- [18] 郭晨龙, 仇振安. 线列红外无热化检测系统及其图像清晰度评价方法[J]. 电光与控制, 2019, 26(12): 59–63.
GUO Chenlong, QIU Zhen'an. Linear infrared athermalization detection system and its image definition evaluation method[J]. Electro Optics and Control, 2019, 26(12): 59–63. (in Chinese)
- [19] TAO Xin, GAO Hongyun, WANG Jue, et al. Scale-recurrent network for deep image deblurring[C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018.
- [20] 耿丽丽, 牛保宁. 深度神经网络模型压缩综述[J]. 计算机科学与探索, 2020(9): 1441–1455.
GENG Lili, NIU Baoning. Review of deep neural network model compression[J]. Computer Science and Exploration, 2020(9): 1441–1455. (in Chinese)
- [21] CHOUDHARY T, MISHRA V, GOSWAMI A, et al. A comprehensive survey on model compression and acceleration[J]. Artificial Intelligence Review, 2020, 53(3): 5113–5155.
- [22] 陆晔, 祝涵珂, 黄海涛. 深度学习模型压缩方法及产品研究[J]. 电信科学, 2020, 36(增刊1): 185–192.
LU Ye, ZHU Hanke, HUANG Haitao. Research on compression methods and products of deep learning model[J]. Telecommunications Science, 2020, 36(Supp.1): 185–192. (in Chinese)
- [23] REN S, HE K, GIRSHICK R, 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.
- [24] BORACCHI G, FOI A. Modeling the performance of image restoration from motion blur[J]. IEEE Transactions on Image Processing, 2012, 21(8): 3502–3517.
- [25] SETIADI D. PSNR vs SSIM: imperceptibility quality assessment for image steganography[J]. Multimedia Tools and Applications, 2020, 80(6): 8423–8444.
- [26] 宋悦. 单排多区段液压功率分流无级变速箱的设计与控制[D]. 泰安: 山东农业大学, 2020.
SONG Yue. Design and control of single row multi-section hydraulic power split stepless transmission[D]. Taian: Shandong Agricultural University, 2020. (in Chinese)