

基于双金字塔网络的 RGB-D 群猪图像分割方法

高云^{1,2} 廖慧敏¹ 黎煊^{1,2} 雷明刚^{2,3} 余梅³ 李小平³

(1. 华中农业大学工学院, 武汉 430070; 2. 生猪健康养殖协同创新中心, 武汉 430070;
3. 华中农业大学动物科技学院, 武汉 430070)

摘要: 为实现群养猪的视觉追踪和行为监测, 针对猪舍中仔猪因拥挤堆叠等习性而导致的目标个体粘连、图像分割困难问题, 提出基于双金字塔网络的 RGB-D 群猪图像分割方法。该方法基于实例分割 Mask R-CNN 框架, 在特征提取网络 (ResNet101) 基础上改进成双金字塔特征提取网络。RGB 图像和 Depth 图像分别提取特征后进行融合, 输入区域生成网络得到预选锚 (ROI) 和共享特征输入 Head 网络, 通过类别、回归和掩模 3 个分支, 输出检测目标的位置和分类结果, 实现猪舍场景下群养仔猪粘连区域的有效个体分割。网络模型训练采用 2 000 组图像样本, 按照 4:1 比例随机划分训练集和验证集。试验结果表明, 双金字塔网络 (Feature pyramid networks, FPN) 能有效解决颜色相近、个体相似的群猪粘连问题, 实现单个仔猪区域的完整分割, 分割准确率达 89.25%, 训练 GPU 占有率为 77.57%, 与 Mask R-CNN 和 PigNet 网络分割结果相比, 分割准确率和分割速度均有较大提高。双金字塔网络模型对于多种行为状态、不同粘连程度的群猪图像中个体分割都取得了良好效果, 模型泛化性和鲁棒性较好, 为群养猪的个体自动追踪提供了新的途径。

关键词: 群养猪; RGB-D; 双金字塔网络; 特征融合; 深度学习

中图分类号: TP391.41 文献标识码: A 文章编号: 1000-1298(2020)07-0036-08 OSID: 

RGB-D Segmentation Method for Group Piglets Images
Based on Double-pyramid Network

GAO Yun^{1,2} LIAO Huimin¹ LI Xuan^{1,2} LEI Minggang^{2,3} YU Mei³ LI Xiaoping³

(1. College of Engineering, Huazhong Agricultural University, Wuhan 430070, China
2. The Cooperative Innovation Center for Sustainable Pig Production, Wuhan 430070, China
3. College of Animal Science and Technology, Huazhong Agricultural University, Wuhan 430070, China)

Abstract: Aiming to achieve automatic individual pig's tracking and monitoring in pig group, an RGB-D image segmentation method based on the double-pyramid network was proposed to solve the segmentation difficulties caused by overlaps and adhesion body areas which were frequently exiting in images because of habits of huddle and crowd in piglets. The method was based on an instance segmentation network Mask R-CNN, modifying its feature extraction network, ResNet101, to a double-pyramid structure. Features were extracted from RGB and Depth images and combined to be inputted into a regional generation network. The network outputted regions of interest (ROI). The combined features and ROIs were then inputted into a head network, which included the classifications and regression and mask branches and outputted the locations of pigs and results of classification. Eventually, the individual pigs were segmented from images according to the outputs. The double-pyramid network was trained using 2 000 groups of images, splitting to a training set and a validation set in a ratio of 4:1 randomly. Experimental results showed that the double-pyramid network (Feature pyramid networks, FPN) can effectively address the segmentation for group pig images of adhesive pigs, and acquire the complete individual pig areas, the segmentation accuracy rate was up to 89.25%. During the training process, the GPU used rate was lower to 77.57%, the FPN outperformed the Mask R-CNN and PigNet networks both in the segmentation accuracy rate and running speed. The double-pyramid network represented its

收稿日期: 2019-10-28 修回日期: 2019-12-26
基金项目: 国家重点研发计划项目(2016YFD0500506)、中央高校自主创新基金项目(2662018JC003、2662018JC010、2662017JC028)和现代农业产业技术体系项目(CARS-35)
作者简介: 高云(1974—),女,副教授,博士,主要从事农业智能检测与控制研究,E-mail: angelclouder@mail.hzau.edu.cn

generalization and robustness on the segmentation for multi-behaviors and diversified adhesions in pig group images, which provided a new approach to automatically track individual pig in group pigs.

Key words: group piglet; RGB-D; double-pyramid network; feature fusion; deep learning

0 引言

针对猪只个体的行为监测能反映猪只的生长及健康状态,是提高猪只福利与养殖经济效益的主要参考依据。在集约化、高密度养殖的现代猪场中,采用计算机视觉技术对猪只行为^[1-4]进行无接触式的监测较为常见。实现猪只多元行为的监测,需要在猪体颜色和纹理相似、猪体拥挤粘连、猪舍中遮挡严重和猪只之间组合姿态不断变化的图像中准确地分割出猪只个体,这对图像分割的准确性提出了较高的要求。

目前,针对猪舍环境中的猪只分割问题国内外学者已进行了相关研究。文献[5-6]通过提取猪只的时间及空间特征,采用全卷积神经网络(Convolutional neural network, CNN)分割母猪图像,实现了对母猪母性护理行为的识别。文献[7-8]针对粘连猪体分割问题进行研究,提出了基于深度学习的粘连猪体分割方法。大量研究表明,深度学习^[9-11]在视觉领域展现出强大的优势,深度学习通过对低维特征到高维特征的自动特征提取框架的训练,避免了通过人为观察设计特征提取过程,模型普适性强,具有对不同场景的自适应能力。上述猪只图像分割研究围绕彩色三通道图像(RGB image)进行,如果叠加采用深度摄像头采集的深度图像(Depth image),并增加镜头到被拍摄对象间的距离信息,将大大提高分割效果。采用深度摄像头采集的图像包括 RGB 图像和 Depth 图像,统称为 RGB-D 图像。与 RGB 分割方法相比,Depth 图像中可存储空间几何信息,利用几何信息,RGB-D 分割方法能有效改善分割、识别效果。文献[12]将 RGB-D 图像中的深度信息与 RGB 信息进行融合,提出了一种多模态前景-背景分割方法。文献[13]利用 RGB-D 图像的深度信息,融合 RGB 信息,在图像显著性目标检测中获得了比仅采用 RGB 图像更好的效果。文献[14]利用深度图像中的深度信息,从 RGB-D 图像中提取出合理的物体支持特征,实现了良好的场景理解。以上研究表明,RGB-D 图像中包含的深度信息能够提供足够的距离信息特征,这更有利于图像中粘连物体的分割。近年来,RGB-D 图像分割^[15-18]研究侧重于语义分割,同时使用 RGB 和 Depth 图像的实例分割研究较少。在语义分割框架的基础上,实例分割框架将同一类对

象进一步分割为不同的实例个体,既能实现目标物体定位,又能对定位物体进行分类打分。RGB-D 实例分割可充分利用深度图像中的几何结构信息提高分割准确度,为群猪图像中粘连个体的分割提供一种可行的解决方案。

本文采用深度学习的 Tensorflow 框架,对实例分割算法 Mask R-CNN(Regions with CNN features, R-CNN)^[19]进一步优化改进网络模型,以实现群养仔猪 RGB-D 图像中粘连猪只个体区域的有效清晰分割。通过对网络模型的多次训练和优化,得到有效的分割网络模型,通过试验验证本文算法的可行性与有效性,为群养猪猪只图像个体区域的有效分割和个体识别^[20]提供技术参考。

1 材料与方法

1.1 数据采集平台

RGB-D 图像数据于 2018 年 3 月 12 日至 5 月 16 日在华中农业大学试验养猪场 24 h 无间断采集获得。试验每轮养殖时长为 1 个月,每次试验采用 6 头单头质量约为 9.6 kg(试验开始时的质量)的大白保育猪,共采集 2 轮试验数据。试验环境采用本课题组前期研究设计的环境多变量控制养殖箱^[21],如图 1 所示。养殖箱内部尺寸(长×宽)为 2.0 m×2.0 m,内部设有饲料槽和饮水碗,以及 LED 灯全自动照明。LED 灯光照强度根据猪场光照标准控制在 50 lx 左右,照明时间自动设置 07:00—18:00,其余时间 LED 灯关闭;底部为全漏缝地板,下方为粪槽,采用水泡粪方式,粪池容量设计标准为 7 d 抽粪 1 次。将试验图像采集设备深度摄像机(Kinect V2, 微软公司,美国)放置在养殖箱顶部正中间垂直向下,与全漏缝地板垂直距离 2.0 m,数据线外接便携

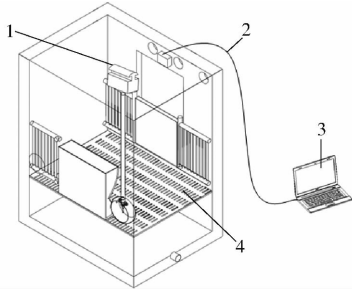


图 1 数据采集平台

Fig. 1 Data acquisition platform

1. Kinect V2 摄像机 2. 数据线 3. 便携式计算机 4. 环境多变量控制养殖箱

式计算机,自动采集存储彩色三通道图像和深度图像,采集速率为3 帧/s。其中 RGB 彩色图像分辨率为1 920 像素×1 080 像素,Depth 深度图像分辨率为512 像素×424 像素,数据存储到硬盘(容量2 TB,希捷公司,美国)。

1.2 数据预处理与标注

由于 Kinect V2 相机采集的原始 RGB 图像和 Depth 图像分辨率不一致,为利用 Depth 图像中的距离信息,应将 RGB 图像和 Depth 图像的尺寸转换为一致,RGB 图像和 Depth 图像中像素位置一一对应。深度摄像机拍摄过程中位置固定不变,本文使用 Python 程序计算得到彩色图和深度图的偏移量,通过 OpenCV 的 ROI 裁剪彩色图和深度图重合部分,使用 OpenCV 的 resize 函数和 INTER_LINEAR 函数,实现彩色图和深度图对齐处理。RGB 图像和 Depth 图像经对齐处理后生成 RGB-D 图像用于网络模型的训练和测试,进行群养猪个体分割试验。

采用 Labelme 软件对彩色图像的猪体区域进行手工标注得到标签图。猪体区域依次标注成标签 pig1~pig6,共6 个标签,分别对应群养猪只编号。标注时多只猪体相互粘连位置不交叉重叠。每标注完1 幅彩色图像得到1 个对应相同命名的.json 文件。通过程序批量解析.json 文件得到彩色图 RGB 对应的 Ground Truth 标签文件(标签图 Ground Truth 表示训练样本的真实值)。深度图像预先采用5×5 中值滤波器去噪,为提高对比度增强深度图像质量采用直方图均衡化处理,形成由彩色图、深度图和标

签图组成的网络模型数据集。

1.3 群养猪图像分割的双金字塔网络结构

Mask R-CNN 实例分割对于同类纹理相似物体的分割精准^[22-24],在一些二维图像分割任务中性能表现优异。本文中的 RGB-D 图像实例分割方法将彩色图中的颜色纹理特征和深度图中的空间几何信息结合。根据 RGB-D 图像数据的特点,需要对 Mask R-CNN 的特征提取网络进行改造。Mask R-CNN 是一个灵活的实例分割框架,较多研究关于 Mask R-CNN 网络的空间几何变换,例如膨胀卷积^[25]在保持参数复杂度不变前提下增加接受域的宽度,空间变换 CNN^[26]利用网络学习的全局空间变换实现对特征图的扭曲处理,可变形 CNN^[27]通过网络学习内核偏移量来增加空间采样位置,也可通过增加不同的分支完成如目标分类、目标检测、语义分割、实例分割、姿势识别等任务。本文中分别对 Mask R-CNN 的特征提取网络深度及结构进行改造,对比不同卷积层数和卷积核尺寸的特征提取网络,保持卷积核尺寸不变,依次对比101、54、50、42、36 层多个数量不同的卷积网络,保持卷积层数,依次对比3×3、7×7、9×9 不同卷积核尺寸对网络模型的影响。改造后的网络模型分块结构如图2 所示。图中,Conv2_x、Conv3_x、Conv4_x、Conv5_x 代表4 个独立残差块;Predict3 表示 Conv3_x 和 Conv2_x 特征融合的过程,Predict4 表示 Conv4_x 和 Conv3_x 特征融合的过程,Predict5 表示 Conv5_x 和 Conv4_x 特征融合的过程。

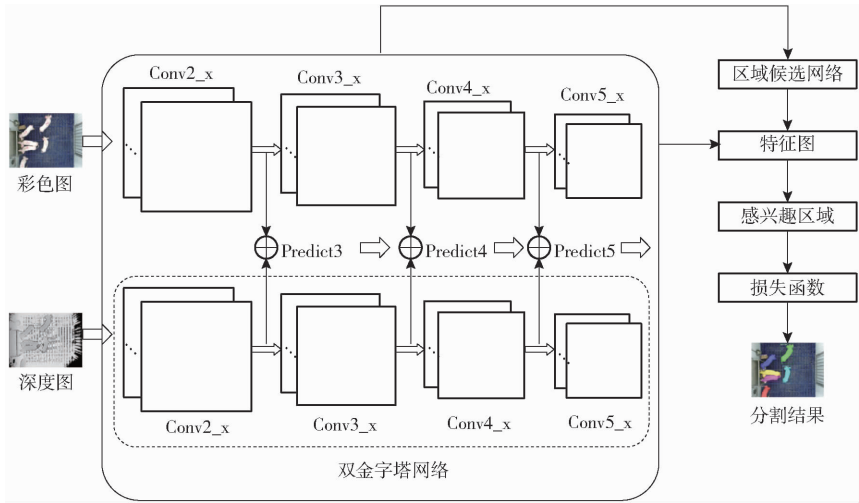


图2 群养猪图像分割的双金字塔网络结构

Fig.2 Double-pyramid networks structure of pig images segmentation

在 Mask R-CNN 特征提取网络101 层 ResNet (Residual network) 基础上,本文方法将101 层网络改变为50 层网络,构建特征提取双金字塔网络结构,分别提取 RGB 和 Depth 图像特征,并进行特征融

合。卷积神经网络前向过程中,特征图的尺寸经过某些卷积层后会改变,而在某些层不会改变,将不改变特征图尺寸的层归为一个阶段,提取每个阶段最后一个残差结构的特征激活输出,对于 ResNet50 的

结构即为 Conv2_x、Conv3_x、Conv4_x、Conv5_x 共 4 个残差块的特征输出。用该特征提取网络分别提取 RGB 图像和 Depth 图像特征,分别通过特征金字塔网络 (Feature pyramid networks, FPN) 的 2 个分支,一个自动提取 RGB 特征,另一个自动提取几何特征,构成双金字塔特征提取网络结构。卷积反向过程进行上采样,采用内插值方法,在原有图像像素之间采用双线性插值算法插入新的元素,实现上采样后的特征图尺寸和下一层的特征图尺寸相同。FPN 提取出的 RGB 特征图和 Depth 特征图在 Predict 过程中实现特征图融合以及上采样结果和前向提取特征图的融合,使图像特征信息更加丰富。由于高于采样频率的高频信号与原来低频信号叠加会导致混叠效应,为消除上采样的混叠效应,采用 3×3 低通滤波器进行处理,即在融合之后采用 3×3 的卷积核,对每个融合结果进行卷积,一层一层依次迭代得到多个新的特征图 (Feature maps)。

上述特征融合过程中产生的分辨率不同的特征图输入区域候选网络 (Region proposal networks, RPN) 中得到一系列预选锚 (Region of interest, ROI),即矩形目标候选框。将 ROIAlign 层取消 2 次 ROI Pooling 操作,采用双线性内插法获得坐标为浮点数像素上的图像采样点数值,实现 RGB 与 Depth 特征图的像素级对齐,从而将整个特征聚集过程转换为一个连续的操作。文献[19]研究表明,该方式可以提升检测模型的准确性。特征图经过对齐后输入 Head 网络结构中,得到图像分割结果。Head 网络结构块包括 Class 分支、Box 分支、Mask 分支,其中 Class 分支和 Box 分支利用全连接层预测每个 ROI 的类别 (Class) 和回归框 (Bounding box);Mask 分支为小型全卷积网络 (Fully convolutional networks, FCN),对每个 ROI 的目标进行分割。损失函数计算包括 Class 分支、Box 分支、Mask 分支分别对应的损失项 L_{cls} 、 L_{bbox} 、 L_{mask} 。设置矩形目标候选框和 Ground Truth 标注框的交并比 (Intersection over union, IOU) 大于 0.5 为有效的 ROI,则 L_{mask} 损失只计算有效的 ROI 项。网络最后的卷积层尺寸为 $28 \times 28 \times 6$,对应标签数为 6 类,每个 ROI 提供其他 2 个输出:类标签和边界框偏移量。

网络充分利用彩色图像和深度图像中的图像特征信息实现 RGB-D 图像实例分割,通过对比分析网络模型确定为 Double FPN 结构,具体对 Mask R-CNN 网络的修改内容包括:

(1) 网络训练数据维度的修改。为实现 RGB-D 图像三维数据的输入,首先需要将 Mask R-CNN 网络数据输入通道数 (Channels) 由 3 个改为 4 个;其

次为保证训练数据维度一致,平均像素矩阵从 3 维改为 4 维,并且重新改写数据读取配置文件中的数据载入函数;最后在训练文件中重新定义 Data 函数,包括标签文件、数据类别、图像尺寸、图像数量等。

(2) 迁移学习的应用。在深度学习领域中,通常使用迁移学习进行模型参数初始化。由于官方 COCO 数据集训练完成的网络模型和本文使用的数据集不一样,COCO 数据集有 80 类,本文构建的数据集是 6 类,为保证迁移学习正常应用,网络模型加载 COCO 数据集官方发布的预训练模型参数时,将其第 1 层网络参数 Conv1 去掉,采用随机初始化方法进行初始化。

(3) 网络模型图像特征提取结构的改进。本文使用的数据为对齐后的 RGB 图像和 Depth 图像,像素尺寸统一为 640 像素 $\times$ 480 像素。随着网络层数加深,学习到的特征信息也越来越丰富,但特征提取优势也逐渐到达瓶颈。经过对比试验,ResNet50 较 ResNet101 模型分割效果好。本文采用 ResNet50 基本结构,对比不同卷积层数和卷积核尺寸的效果,最终确定 42 层卷积和 7×7 的卷积核,并改进成为双金字塔特征提取网络结构,同时提取彩色图和深度图特征,在 ROI 层前实现 RGB 和 Depth 双数据特征信息融合。

1.4 模型训练参数设置

先进行 Head 网络的参数训练,Head 网络结构使用 ResNet-50-C4,代表 50 层残差网络层数,特征提取残差块为 4 个,周期 (epoch) 设置为 40。Head 网络部分训练结束后开始全部网络层继续训练,epoch 设置为 100,输入数据尺寸 640 像素 $\times$ 480 像素,学习率 (Learning rate, LR) 设置为 0.002,在单个 GPU 上训练,动量因子 (Momentum) 和正则化权重衰减系数 (Weight decay) 分别设置为 0.9 和 0.0001。

2 群养猪图像分割试验

2.1 试验方法

视频图像全天无间断连续采集 28 d,采集到的数据涵盖不同时段群猪的多种状态,主要包括饮水^[28-29]、采食^[30]、站立^[31]、俯卧休息等。为确保训练样本的多样性和验证算法的泛化能力,从采集到的 RGB 图像以及相同命名的 Depth 图像中随机选取不同时段、不同状态的群养猪图像 4 000 幅进行标注作为数据集,其中 RGB 图像和 Depth 图像各 2 000 幅,构建本文网络训练模型的数据集 (包括训练集样本和验证集样本)。从数据集中随机抽取

20% 作为验证集,即 800 幅图像。最终用于试验的训练集为 1 600 组图像,验证集为 400 组图像,测试集(未标注图像)为 400 组图像。

2.2 试验设备

试验采用 32 GB 运行内存、NVIDIA GeForce GTX 1080Ti 型号的 GPU、Intel Core i7-8700 型号的 CPU、主频 3.2 GHz 的终端计算机硬件平台,安装 Ubuntu16.04 LTS 64 位操作系统,搭建深度学习框架 Tensorflow,试验编程语言采用 Python 3.6。

2.3 试验流程

进行群养保育仔猪图像分割的双金字塔特征提取网络模型的训练和测试,具体流程如图 3 所示,主要分为 7 个步骤:

(1) Kinect V2 相机采集试验图像数据,随机选取试验数据样本,涵盖不同时段不同群猪状态的视频图像,从中生成训练集、验证集及测试集。

(2) 选取 2 000 组彩色图和深度图,将彩色图与深度图进行对齐处理,尺寸统一为 640 像素 × 480 像素。使用 Labelme 软件标记彩色图,得到其对应的 Ground Truth 标签文件,按照保留法通常按 4:1 比例随机划分得到训练样本集和验证样本集。

(3) 对彩色图像和深度图像同时进行随机尺寸变化、旋转、裁剪 3 种操作实现数据集扩增,数据增加 3 倍。本文数据扩增通过在训练过程中直接调用 Keras 库实现,减少模型的过拟合,提高泛化能力和鲁棒性。

(4) 以 ResNet50 为基础网络构建双金字塔特征提取网络,结合 Mask R-CNN 检测框架和分割框架,进行结构调整优化,形成深度感知群养猪分割网络模型。

(5) 对模型进行训练,对比训练集和验证集图像分割结果评价指标,根据评价结果不断调整网络参数,直至训练集指标与验证集指标接近,得到最优模型。

(6) 利用未标记的测试集验证模型实际效果,微调模型参数,进一步优化模型性能。

(7) 对不同的分割模型方法进行对比试验,确定最优模型。

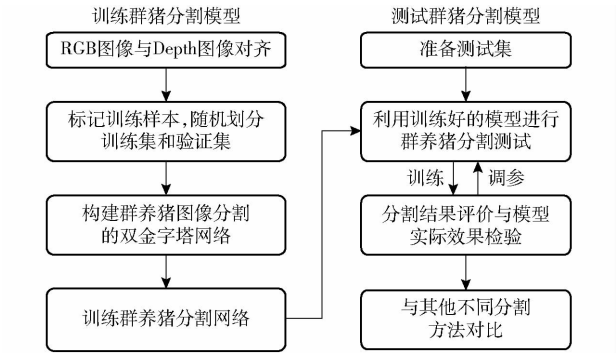


图 3 基于双金字塔网络的群养猪图像分割试验流程图

Fig.3 Flow chart of pig images segmentation based on double-pyramid networks

3 试验结果与分析

3.1 实际效果检验与分析

3.1.1 不同猪体粘连程度及分析

根据图像中猪体粘连程度的不同,定义为一般粘连、困难粘连和极度粘连 3 类。一般粘连表现为猪体间小范围接触,困难粘连定义为 2 头以上的猪只猪体紧密大范围接触,极度粘连表现为多只猪体相互挤压堆叠行为。猪只粘连程度分类定义如表 1 所示。

表 1 猪体粘连程度分类定义

Tab.1 Classification and definition of adhesions in pigs

粘连程度类型	定义
一般粘连	头碰头、头碰身体、身体碰身体等小范围粘连
困难粘连	由于采食饮水运动时群猪间竞争行为引起头部身体大范围粘连
极度粘连	由于群猪猪体相互挤压堆叠导致的极度粘连,多见于躺卧

本文方法对群猪多种姿态、各种粘连情况均能表现出良好的分割效果,不同粘连程度的分割示意图如图 4 所示。群猪头尾相连、多只猪体间紧密接触以及猪群躺卧时,多表现为相互挤压堆叠易导致部分猪体被遮挡,给猪只个体的分割带来较大影响。

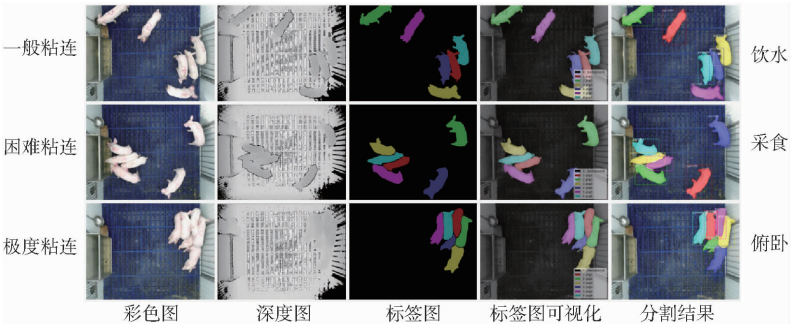


图 4 部分测试集的分割效果

Fig.4 Segmentation effect of partial test sets

由图 4 可知,图像中群猪中的猪只首尾不同方向,猪体间不同的曲直程度,以及不同粘连程度的猪体均能被有效分割。

3.1.2 不同猪只行为分析及个体识别

相比较于单只猪体的行为状态,群猪的行为具有更多的复杂性和多样性。图 4 中群猪饮水、采食和俯卧 3 个典型行为状态中,饮水有单只猪饮水和猪群中有其他猪只站立及俯卧休息等多种行为掺杂,猪体间表现为一般粘连;采食行为包括多猪只竞争采食,身体有大量触碰,猪体间表现为困难粘连,同时猪群中有其他猪只正在行走;俯卧行为中群猪猪体粘连严重,猪体间表现为极度粘连。群猪群体活动时猪只各种行为会相互参杂,猪体粘连情况不一。对于常有猪体堆叠或猪只超出拍摄范围造成的部分猪体不在视野中的现象,本文方法也都可得到准确的分割结果。

文献[32-33]采用记号笔在猪只背上书写标记来标识不同的猪只,算法会将标记认为是猪只特征的一部分,当标记褪色和标记变形会影响猪只个体识别结果。本文中对猪只进行标识时,根据猪只背部试验初期的人工标记,依次按照相同序号命名,因试验时间较长,试验过程中标记逐步褪色至不明显。图 4 中显示了算法对猪只个体均实现身份识别,当背部标识部分在镜头内不完整以及部分猪体俯卧背部标识被遮挡亦能实现猪只个体的正确识别。但少量图像样本仍然存在错误识别的现象,如图 5 所示,图 5a 将爬跨猪只粘连在一起的两头猪均识别为 pig2,图 5b 侧卧猪只未被正确分割出来。

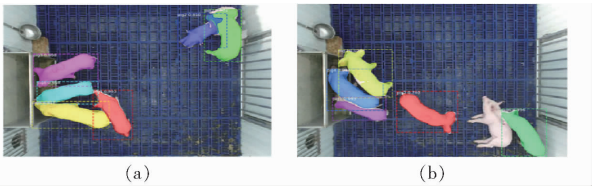


图 5 误分割示意图

Fig.5 Error segmentation diagrams

3.2 不同模型性能分析与比较

本研究主要目标是准确分割出群猪图像中的猪只个体区域位置和数量,为猪只的个体行为追踪提供途径,因此分析以图像分割结果和猪只个体边界区域基本相符即认为分割正确,即算法将图像中群猪猪只个体分割成独立的 6 个连通域,且 6 个连通域正确覆盖猪体达 80%,即判定该图像为分割正确样本。分割准确率为正样本数占总样本数的百分比,表 2 为不同分割方法的结果对比。Mask R-CNN 和 PigNet 的样本对像为 RGB 彩色图像,本文

方法准确率最高达到 89.25%,分割时间也最短,为 1.45 s,相比其他 2 种算法大幅提高了分割的准确率,减少了运算时间。

表 2 不同分割方法的性能比较

Tab.2 Performance comparison of different segmentation methods

方法	准确率/%	平均分割时间/s	GPU 占用率/%
PigNet	85.40	2.12	
Mask R-CNN	75.00	1.53	83.86
本文方法	89.25	1.45	77.57

根据表 2 的分割结果,对几种算法模型进行具体分析。

(1)本文方法和 PigNet 方法对比。本文网络模型采用的 RGB-D 图像尺寸为 640 像素×480 像素,包括 RGB 的 3 通道及 Depth 图像,对齐处理后,相当于 4 维矩阵,单个样本总输入数据个数为 1 228 800。PigNet 网络模型采用的 RGB 图像尺寸为 1 920 像素×1 080 像素,通道为 3 维,单个样本的输入数据个数为 6 220 800,远大于本文模型的输入数据个数,因此本文方法的平均分割时间减少 0.67 s。另外因为增加 Depth 图像信息,虽然 RGB 图像尺寸小于 PigNet 输入图像,但是输入信息的多样性方面大幅增强,使分割准确率提高了 3.85 个百分点。

(2)本文方法和 Mask R-CNN 对比。本文方法在 ResNet50 的基础上进行改进,采用双金字塔并行结构,在单向层数上远少于 Mask R-CNN 特征提取网络采用的 ResNet101,使平均分割时间减少了 0.08 s,GPU 占用率也减少了 6.29 个百分点。由于采用 RGB-D 的图像输入,输入信息量的增加,使准确率提高 14.25 个百分点。

通过上述分析可以看出改进后的网络模型明显优于其他 2 种网络模型,同时使用 RGB 图像和 Depth 图像深度信息训练的双金字塔网络模型能够在群猪图像分割上取得更好的效果。

4 结论

(1)通过对 Mask R-CNN 实例分割网络结构进行改进,得到双金字塔特征提取网络的群养猪 RGB-D 图像分割网络模型。以 Mask R-CNN 分割框架和检测框架为基础,将特征提取网络 ResNet101 减少至 ResNet50 层,在此基础上构建双金字塔特征提取网络结构,同时提取 RGB 图像和 Depth 图像特征信息,并在 ROI 层前进行特征融合,实现了群猪猪只个体识别及分割。

(2)与 PigNet 和 Mask R-CNN 分割方法相比,本文方法的分割准确率分别提高 3.85、14.25 个百

分点,单帧平均分割时间为 1.45 s,在分割准确率和分割速度上均优于 PigNet 和 Mask R - CNN。本文方法能够有效解决群猪图像中猪只粘连问题,分割

出的猪只轮廓清晰、猪只个体区域边界基本吻合,分割准确率可达 89.25% ,并有很好的泛化性和鲁棒性。

参 考 文 献

[1] 何东健,刘冬,赵凯旋.精准畜牧业中动物信息智能感知与行为检测研究进展[J/OL].农业机械学报,2016,47(5):231 - 244. HE Dongjian, LIU Dong, ZHAO Kaixuan. Review of perceiving animal information and behavior in precision livestock farming [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(5):231 - 244. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160532&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2016.05.032. (in Chinese)

[2] NASIRAHMADI A, HENSEL O, EDWARDS S A, et al. Automatic detection of mounting behaviours among pigs using image analysis[J]. Computers and Electronics in Agriculture, 2016, 124:295 - 302.

[3] CHEN C, ZHU W, MA C, et al. Image motion feature extraction for recognition of aggressive behaviors among group-housed pigs[J]. Computers and Electronics in Agriculture, 2017, 142:380 - 387.

[4] VALLETTA J J, TORNEY C, KINGS M, et al. Applications of machine learning in animal behaviour studies[J]. Animal Behaviour, 2017, 124:203 - 220.

[5] YANG A Q, HUANG H S, ZHU X M, et al. Automatic recognition of sow nursing behaviour using deep learning-based segmentation and spatial and temporal features[J]. Biosystems Engineering, 2018, 175:133 - 145.

[6] YANG A Q, HUANG H S, ZHENG C, et al. High-accuracy image segmentation for lactating sows using a fully convolutional network[J]. Biosystems Engineering, 2018, 176:36 - 47.

[7] JU M, CHOI Y, SEO J, et al. A Kinect-based segmentation of touching-pigs for real-time monitoring[J]. Sensors, 2018, 18(6):1746.

[8] 高云,郭继亮,黎焯,等.基于深度学习的群猪图像实例分割方法[J/OL].农业机械学报,2019,50(4):179 - 187. GAO Yun, GUO Jiliang, LI Xuan, et al. Instance-level segmentation method for group pig images based on deep learning[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(4):179 - 187. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190420&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.04.020. (in Chinese)

[9] KAMILARIS A, PRENAFETA-BOLDÜ F X. Deep learning in agriculture: a survey[J]. Computers and Electronics in Agriculture, 2018, 147:70 - 90.

[10] 孙红,李松,李民赞,等.农业信息成像感知与深度学习应用研究进展[J/OL].农业机械学报,2020,51(5):1 - 17. SUN Hong, LI Song, LI Minzan, et al. Research progress of image sensing and deep learning in agriculture [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(5):1 - 17. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20200501&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2020.05.001. (in Chinese)

[11] ZHENG C, ZHU X, YANG X, et al. Automatic recognition of lactating sow postures from depth images by deep learning detector[J]. Computers and Electronics in Agriculture, 2018, 147:51 - 63.

[12] TRABELSI R, JABRI I, SMACH F, et al. Efficient and fast multi-modal foreground-background segmentation using RGBD data[J]. Pattern Recognition Letters, 2017, 97:13 - 20.

[13] XU H, ZHANG G, ZHANG Q. An iterative propagation based co-saliency framework for RGB - D images[J]. Journal of Visual Communication and Image Representation, 2019, 59:186 - 194.

[14] XUE F, XU S, HE C, et al. Towards efficient support relation extraction from RGB - D images[J]. Information Sciences, 2015, 320:320 - 332.

[15] GUPTA S, GIRSHICK R, ARBELÁEZ P, et al. Learning rich features from RGB - D images for object detection and segmentation[C]//European Conference on Computer Vision. Zurich: Springer, Cham, 2014: 345 - 360.

[16] CHENG Y, CAI R, LI Z, et al. Locality-sensitive deconvolution networks with gated fusion for RGB - D indoor semantic segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 1475 - 1483.

[17] PARK S J, HONG K S, LEE S. Rdfnet: RGB - D multi-level residual feature fusion for indoor semantic segmentation[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 4980 - 4989.

[18] QI X, LIAO R, JIA J, et al. 3D graph neural networks for RGB - D semantic segmentation [C] // IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 5209 - 5218.

[19] HE K M, GEORGIA G, PIOTR D, et al. Mask R - CNN [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 99:1 - 12.

[20] HANSEN M F, SMITH M L, SMITH L N, et al. Towards on-farm pig face recognition using convolutional neural networks [J]. Computers in Industry, 2018, 98:145 - 152.

[21] 高云,陈震撼,王瑜,等.多环境参数控制的猪养殖箱设计及箱内气流场分析[J].农业工程学报,2019,35(2):203 - 212.

- GAO Yun, CHEN Zhenhan, WANG Yu, et al. Design for pig breeding chamber under multiple environment variable control and analysis of internal flow field[J]. Transactions of the CSAE, 2019, 35(2):203–212. (in Chinese)
- [22] LI Y, QI H Z, DAI J F, et al. Fully convolutional instance-aware semantic segmentation[C] // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017:2359–2367.
- [23] 李丹, 张凯锋, 李行健, 等. 基于 Mask R-CNN 的猪只爬跨行为识别[J/OL]. 农业机械学报, 2019, 50(增刊):261–266, 275. LI Dan, ZHANG Kaifeng, LI Xingjian, et al. Mounting behavior recognition for pigs based on Mask R-CNN[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(Supp.):261–266, 275. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=2019s041&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.S0.041. (in Chinese)
- [24] 林相泽, 朱赛华, 张俊媛, 等. 基于迁移学习和 Mask R-CNN 的稻飞虱图像分类方法[J/OL]. 农业机械学报, 2019, 50(7):201–207. LIN Xiangze, ZHU Saihua, ZHANG Junyuan, et al. Rice planthopper image classification method based on transfer learning and Mask R-CNN[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(7):201–207. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190721&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.07.021. (in Chinese)
- [25] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[C] // International Conference on Learning Representations (ICLR), 2016.
- [26] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial transformer networks[C] // Neural Information Processing Systems (NIPS), 2015.
- [27] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks[C] // The IEEE International Conference on Computer Vision (ICCV), 2017:764–773.
- [28] LAO F, BROWN-BRANDL T, STINN J P, et al. Automatic recognition of lactating sow behaviors through depth image processing[J]. Computers and Electronics in Agriculture, 2016, 125:56–62.
- [29] 杨秋妹, 肖德琴, 张根兴. 猪只饮水行为机器视觉自动识别[J/OL]. 农业机械学报, 2018, 49(6):232–238. YANG Qiumei, XIAO Deqin, ZHANG Genxing. Automatic pig drinking behavior recognition with machine vision[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(6):232–238. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20180627&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2018.06.027. (in Chinese)
- [30] YANG Q M, XIAO D Q, LIN S C. Feeding behavior recognition for group-housed pigs with the Faster R-CNN[J]. Computers and Electronics in Agriculture, 2018, 144:453–460.
- [31] JUNHEE L, JONGUK L, DAIHEE P, et al. Individual pig detection using Kinect depth information and convolutional neural network[J]. Sensor, 2018, 18(2):1–10.
- [32] KASHIHA M A, BAHR C, OTT S, et al. Automatic identification of marked pigs in a pen using image pattern recognition[J]. Computers & Electronics in Agriculture, 2013, 93(2):111–120.
- [33] ZHANG Y Q, CAI J H, XIAO D Q, et al. Real-time sow behavior detection based on deep learning[J]. Computers and Electronics in Agriculture, 2019, 163:104884.