

基于实例分割和光流计算的死兔识别模型研究

段恩泽 王粮局 雷逸群 郝宏运 王红英
(中国农业大学工学院, 北京 100083)

摘要: 为实现自动化识别死兔,提高养殖管理效率,以笼养生长兔为研究对象,以基于优化 Mask RCNN 的实例分割网络和基于 LiteFlowNet 的光流计算网络为研究方法,构建了一种多目标背景下基于视频关键帧的死兔识别模型。该模型的实例分割网络以 ResNet 50 残差网络为主干,结合 PointRend 算法实现目标轮廓边缘的精确提取。视频关键帧同时输入实例分割网络和光流计算网络,获取肉兔掩膜的光流信息和掩膜边界框中心点坐标。利用光流阈值去除活跃肉兔掩膜,通过核密度估计算法获取剩余中心点坐标的密度分布,通过密度分布阈值实现死兔的判别。实验结果表明,肉兔图像分割网络的分类准确率为 96.1%,像素分割精确度为 95.7%,死兔识别模型的识别准确率为 90%。本文提出的死兔识别模型为兔舍死兔识别和筛选工作提供了技术支撑。

关键词: 深度学习;实例分割;光流法;死兔识别

中图分类号: S829.1 文献标识码: A 文章编号: 1000-1298(2022)02-0256-09 OSID: 

Dead Rabbit Recognition Model Based on Instance Segmentation and Optical Flow Computing

DUAN Enze WANG Liangju LEI Yiqun HAO Hongyun WANG Hongying
(College of Engineering, China Agricultural University, Beijing 100083, China)

Abstract: Screening and isolating dead rabbits is one of the important work of meat rabbit farms, which is helpful to build a rabbit breeding safety system. In order to identify dead rabbits automatically and improve the efficiency of breeding management, cage-rearing breeding rabbits was taken as the research object, a dead rabbit recognition model was proposed which was based on the modified Mask RCNN and LiteFlowNet. The instance segmentation part of the model used ResNet 50 residual network as the backbone, used PointRend algorithm as the network head to extract the instance contour accurately. The key frames of the rabbit videos were sent to rabbit instance segmentation network and optical flow calculation network at the same time to obtain the optical flow of the meat rabbit mask and the center point coordinates of the instance boundary boxes. The masks of the active rabbits were removed by the threshold of the optical flow, and then the density distribution of the remaining center point coordinates was obtained by kernel density estimation algorithm, and the dead rabbits were distinguished by density distribution threshold. The experiment results showed that the classification accuracy of the rabbit segmentation network was 96.1%, the pixel segmentation accuracy of the rabbit segmentation network was 95.7%, and the recognition accuracy of the dead rabbit recognition model was 90%. This study provided technical support for dead rabbit recognizing and isolating in rabbit farms.

Key words: deep learning; instance segmentation; optical flow method; dead rabbit recognition

0 引言

随着我国肉兔养殖业向机械化和智能化发展,智能管理装备取代人力劳动进行兔舍的标准化饲养管理是行业的发展趋势。死兔的筛选是兔舍管理的重要工作之一,及时筛选和隔离死兔能够降低疫病

收稿日期: 2021-07-31 修回日期: 2021-09-21
基金项目: 财政部和农业农村部:国家现代农业产业技术体系项目(CARS-43-D-3)
作者简介: 段恩泽(1995—),男,博士生,主要从事畜禽养殖技术与装备研究,E-mail: 952510993@qq.com
通信作者: 王红英(1966—),女,教授,博士生导师,主要从事畜禽养殖技术与装备研究,E-mail: hongyingw@cau.edu.cn

风险。

国内规模化肉兔养殖主要依靠饲养员巡检兔舍完成死兔的识别工作,不仅增加了人力和经济成本,且容易造成漏检或检测不及时。在养殖智能化和精准化的趋势下,死兔的自动化识别亟待解决。

目前国内外有关肉兔健康检测的研究大多集中于兔群体跟踪和病理学分析,而对兔养殖过程中的疾病管理、死兔检测鲜有报道^[1-3]。在养殖动物健康检测领域,现有研究方向集中于通过传感器采集动物图像、声音、温度和光谱等特征,再进行数据分析^[4-6]。相比于上述方法,将机器视觉运用到养殖环节中能降低对动物的干扰,更适合规模化养殖。通过对目标语义分割、实例分割以及热成像等方式,判断动物健康状况、生理活动和福利水平是智能化养殖的热门方向^[7-9]。

目前有关死兔的图像识别和图像信息处理的研究少见报道。文献[10]表明,肉兔死亡后身体迅速僵硬,保持死亡前姿态,因此从静态图像看,死兔与活兔的图像特征差异很小,采用语义分割算法根据图像直接判断死兔准确率和可行性很低。另一方面,由于兔毛隔热性能良好,利用红外相机采集肉兔图像,发现除耳朵、眼球和皮肤裸露部分以外,肉兔红外图像与环境差异很小,采用红外成像技术判断死兔难度较高。

基于上述研究现状,本文采用视频识别方法,将长期保持静止作为死兔判定依据,以肉兔繁育舍中笼养生长兔为研究对象,提出一种基于深度学习的死兔识别方法。该方法结合图像分割网络和光流计算网络,通过处理肉兔视频关键帧,实现准确快速地从多目标环境中判断笼内死兔,从而提高养殖管理效率。构建基于 Mask RCNN + PointRend 网络的肉兔图像分割网络,获取肉兔掩膜图像及边界框中心点,构建基于 LiteFlowNet 的关键帧光流计算网络,获取肉兔掩膜的光流值。利用光流阈值去除肉兔掩膜图中的活跃肉兔。利用核密度估计法,确定边界框中心点的密度分布,进一步确定区分死兔和活兔掩膜的密度分布阈值,实现对死兔的识别。

1 模型与算法

1.1 数据采集

于2021年2月21—28日在河南省济源市阳光兔业科技有限公司肉兔繁育舍采集了66日龄的伊普吕肉兔视频。实验兔舍为全封闭的阶梯式笼养兔舍,每个兔笼中包含5只肉兔。

采集视频所用工业相机型号为海康威视MV-CA020-10GC型,分辨率为1 624像素×1 240像素。

由于兔笼体积为40 cm×90 cm×35 cm,为使图像包含单笼信息,去除相邻兔笼干扰,视频尺寸设置为1 280像素×720像素,帧率30 f/s。相机安装于立式三脚架上,距离地面1 m,倾斜45°俯拍,如图1所示。

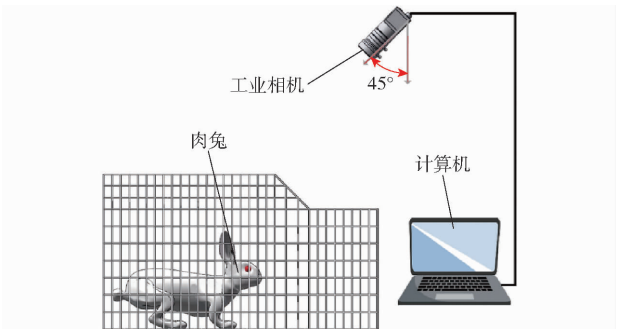


图1 视频采集方法

Fig.1 Method of video acquisition

采集时选取29个存在死亡肉兔的笼位和31个肉兔全部存活的笼位,每个笼位各采集1 min视频,格式为mp4。随机选取20个活兔视频和20个死兔视频作为模型的训练集,剩余20个视频作为模型的验证集。

1.2 死兔识别模型架构

模型通过获取到的笼内肉兔的视频,判别目标是否为死兔,具体流程如图2所示。

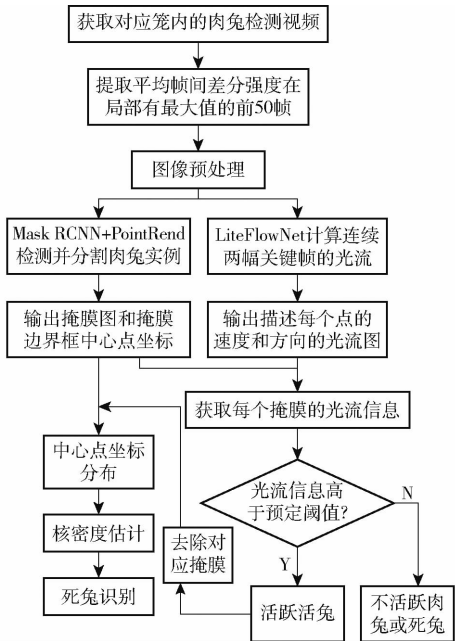


图2 死兔识别模型流程图

Fig.2 Flow chart of dead rabbit identification model

本文提出的死兔识别模型首先提取肉兔视频的关键帧,经图像预处理后,将图像数据集送入 Mask RCNN + PointRend 架构的肉兔图像分割网络进行分割,提取肉兔的掩膜和对应边界框中心点坐标。数据集同时送入 LiteFlowNet 光流计算网络,计算连续

的关键帧的光流信息。通过处理所得掩膜图和光流图,获取每个掩膜的光流,光流高于阈值的掩膜判断为活跃肉兔,去除对应掩膜。得到的剩余掩膜中心点坐标分布进行核密度估计,如高于阈值,该掩膜判定为死兔。

1.3 基于帧间差分的肉兔视频关键帧提取

受活动空间的限制,笼养兔在没有外界干扰的情况下活跃度较低,导致采集到的视频文件中包含大量重复或差别不大的帧,另一方面,本模型更关注短时间内肉兔的姿态变化情况,预测视频中的所有帧会降低模型的识别效率。因此为了提高模型效率,首先将视频中的全部帧分为关键帧和普通帧。关键帧包含了更多活兔的姿势或行为产生变化的信息,利用视频关键帧代替视频进行死兔和活兔的识别分类,可以大大减少运算时间。

视频采集过程中,摄像头静止,环境光稳定,死兔和背景不发生位移,因此视频中的活动目标均为活兔。本模型判别关键帧的依据是连续帧中肉兔的位移程度,采用帧间差分法^[11]得到图像平均像素强度,衡量肉兔产生的位移,计算公式为

$$D = \frac{\sum_{(x,y)} |f_{k+1}(x,y) - f_k(x,y)|}{N}$$

(1)

式中 D ——平均像素强度
 N ——图像列号 k ——帧序号
 $f(x,y)$ ——点 (x,y) 处的像素值

若肉兔在部分时段具有较高活跃度,按平均像素强度选取关键帧会使结果集中在该时段,导致关键帧对视频整体的代表性不足。因此,本模型将视频平均像素强度按时间排序后,选取具有局部最大值的帧作为视频的关键帧,使关键帧均匀分散在视频中,平滑肉兔在图像中的移动。图 3 展示了依据平均像素强度的局部最大值选取部分关键帧的过程。在视频不同时段,具有局部像素强度峰值的帧被选为关键帧。

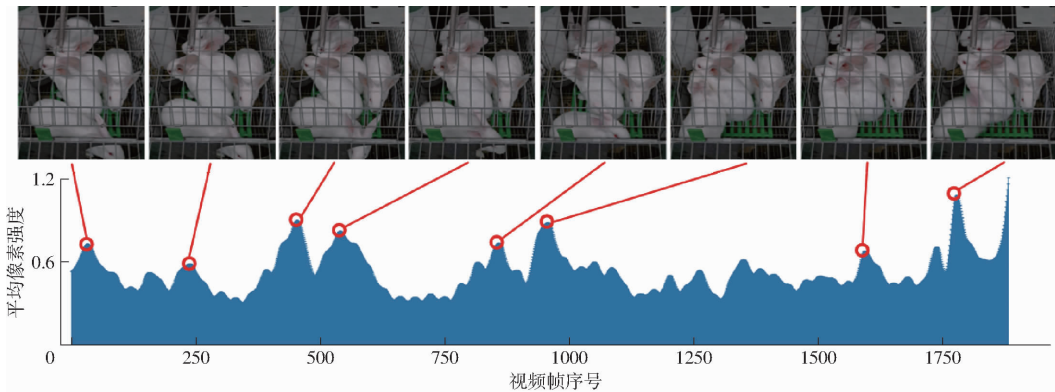


图 3 关键帧选取示例(仅展示部分关键帧)
Fig. 3 Examples of selected key frames

1.4 肉兔实例分割网络构建

肉兔实例分割网络主要分为两部分,采用 Mask RCNN 的基本架构用于识别图像中的肉兔实例,并引入 PointRend 算法取代 Mask RCNN 中原有的网络头以提升网络的分割准确性,尤其是对掩膜边界的分割准确性。

1.4.1 Mask RCNN 的网络结构

Mask RCNN^[12]是在 Faster RCNN^[13]基础上提出的用于图像实例分割的深层卷积神经网络。Mask RCNN 通过额外增加掩膜生成分支,并使用 RoIAlign 算法替换 RoIpooling 算法,保证了目标分割的准确性。基于 Mask RCNN 的肉兔图像分割网络结构如图 4 所示。

本文的研究对象为肉兔,图像特征较为单一,因此采用基于 ResNet 50 的特征金字塔网络(Feature pyramid network, FPN)作为主干网络^[14],通过降低残差网络的层数以缓解梯度消失和训练退化的问题。ResNet 50 通过卷积将原图处理成不同深度的特征图,FPN 将不同深度的特征图融合。特征图输入区域生成网络(Region proposal networks, RPN)中生成大量表达分割物体位置的候选框并进行回归计算,修正候选框,得到感兴趣区域(RoI)。最后特征图被送入头部网络,实现目标分类、边框回归和掩膜生成,得到最终的分割图像。

1.4.2 PointRend 算法结构

Mask RCNN 进行掩膜生成任务时,预测对象为肉兔图像上的规则网格,预测的重点目标是网格上的所有点,降低了对目标边界处分割平滑度的关注,增加了边界预测的损失,使掩膜边界的分割结果不清晰。为了解决这一问题,本文引入 PointRend 算法^[15],通过自适应的点细分策略,针对掩膜边缘的一组非均匀点来计算目标标签。

PointRend 的本质是通过迭代的上采样,判断目标边缘位置,将计算力集中在目标边缘实现边缘细分。共包含 3 个模块:点选取模块、点特征提取模块

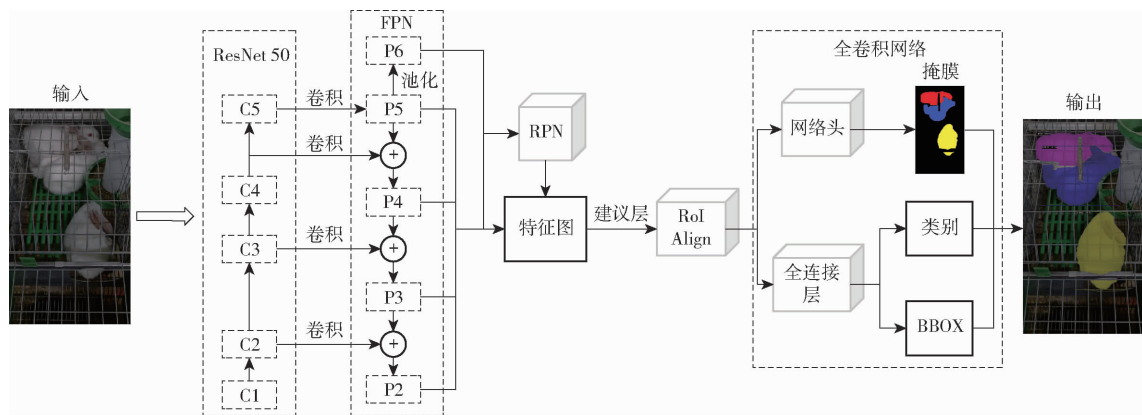


图 4 基于 Mask RCNN 的肉兔图像分割网络结构

Fig. 4 Structure of rabbit segmentation network based on Mask RCNN

和 Point Head 模块。

点选择模块的任务是自适应地筛选预测目标的点,并重点关注处于高频端口附近的点,例如预测对象的边界。在推断过程中,点选择模块的策略是由粗至细的方式迭代地渲染输出图像。在每次迭代过程中,首先预测低空间分辨率的特征图。空间分辨率的网格尺寸最大,预测最粗糙,通过双线性插值法,将特征图上采样扩大 2 倍,从而获得较为精细的高空间分辨率的特征图,使规则网格更密集。在所得的高分辨率的特征图中筛选分割置信度(置信度区间为 $[0,1]$)低于 0.5 的最不确定的点,这些点的选择策略为

$$n_i^* = \arg \min_{n_i} |p(n_i) - 0.5| \quad (2)$$

式中 n_i^* ——被选择的点的集合

$p(n_i)$ ——点 n_i 属于掩膜边缘的概率

通过逐步上采样,对点边缘的预测不断细化,直至特征图被上采样至期望的空间分辨率,点细分预测过程如图 5 所示。

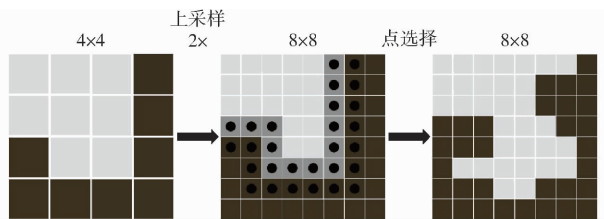


图 5 肉兔图像边缘的点细分预测过程

Fig. 5 Process of one adaptive subdivision step in rabbit image edge points

对于训练过程,点选取模块通过基于随机抽样的非迭代策略找到预测目标的边界。首先从特征图中随机抽取 kN ($k > 1$) 个点,再根据这些点的不确定性估计值对点进行排序,不确定性估计值指所选的点落在前景和背景的概率。从中选择最不确定的 βN ($\beta \in [0,1]$) 个点,则这些点集中在边界处,分布相对均匀。

点特征提取模块负责提取 βN 个点的特征,作为下一阶段的输入。每个点特征由低层特征和高层特征构成,低层特征通过对特征图进行双线性插值方法计算,包含了细节分割结果。高层特征是包含了语义信息的 2 维向量,通过对每层特征图进行 2 倍双线性插值上采样计算。每个点的低层特征和高层特征被送入 Point Head 模块。

Point Head 模块是一个具有 3 个隐含层和 256 个通道的多层感知器 (Multi-layer perceptron, MLP)^[16]。通过迭代对每个点在不同尺寸特征图的特征进行卷积,逐层连接低层特征和高层特征,将目标分辨率最终上采样至期望分辨率,实现对肉兔掩膜边缘部分的精细提取。MLP 的隐含层激活函数为 ReLU^[17],输出层的激活函数为 Sigmoid 函数。

1.5 肉兔图像光流计算网络构建

光流是图像中目标像素点运动的亮度变化,光流计算是从视频图像数据中估计物体的运动。传统的 LK 光流法^[18]和 HS 光流法^[19]利用光流约束方程计算光流场,但存在噪声敏感、运行速度慢等缺点。采用深度神经网络估算图像光流具有更好的鲁棒性和运算速度。

LiteFlowNet^[20]是基于 FlowNet^[21]的用于光流计算的轻量级卷积神经网络,在降低了参数数量的同时提高了对速度场的估算精度。LiteFlowNet 的网络结构分为用于特征提取的编码器 NetC 和实现由粗到精的光流估计的解码器 NetE。LiteFlowNet 的网络结构如图 6 所示。其中,M;S 模块表示亚像素细化,R 模块表示速度场正则化。

NetC 将 2 幅输入的关键帧压缩,形成多尺度的高维特征金字塔 F_k 。输入图像 f_1, f_2 被相同网络结构的卷积层逐层压缩,获得不同空间分辨率的图像特征 $F_k(f_i)$, i 为图像编号, k 为金字塔层级,每层金字塔将图像的分辨率压缩为上层的一半。本文将特征金字塔层数设为 6 层,为便于描述,图 6 只展示了

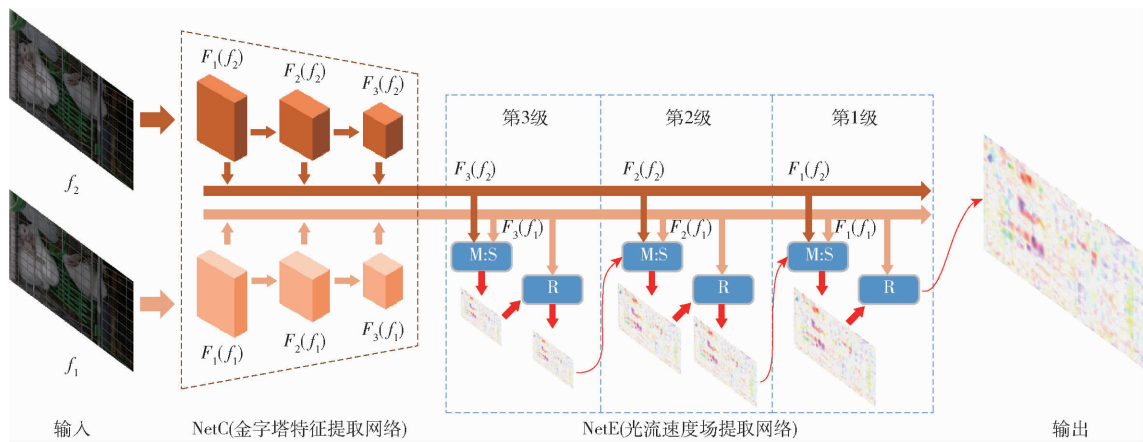


图 6 基于 LiteFlowNet 的肉兔图像光流计算网络结构

Fig. 6 Key frame optical flow estimation network of meat rabbit based on LiteFlowNet

3 层网络。

NetE 通过多级反卷积层处理 NetC 输出的多层图像特征。为估计大位移速度场,图像特征首先经特征变形(f-warp)减小特征间的距离,特征 f_2 通过上层金字塔的预测速度变形到 f_1 ,使大位移速度场的精度更高。随后,通过特征匹配和亚像素细化(M:S),计算特征 f_1 和变形出的特征 $\hat{f}_2$ 之间的相关性,对像素点相匹配,得到粗略的速度场估计,再利用最小特征空间距离,将速度场精细化。最后,为平滑估计的速度场,利用特征驱动的局部卷积实现速度场正则化。

通过特征图在 NetE 中迭代反卷积,最终获取原图尺寸的像素点速度场估计。

2 死兔识别模型实验

2.1 实验平台

目标检测网络和光流计算网络的搭建、训练及调试均基于 Pytorch 深度学习框架,采用 Python 编程语言编写,使用 GPU 加速计算。实验在 Ubuntu 20.04 系统下运行,计算机配置:处理器为 Intel(R) Core(TM) i7-9700K,主频 3.6 GHz,内存 16 GB,显卡为 NVIDIA GeForce RTX 2080(16 GB)。

2.2 数据集制作

通过基于帧间差分的肉兔视频关键帧提取方法处理 40 个训练集视频,每段视频提取 50 幅关键帧,共计 2 000 幅。由于光流法对图像亮度敏感,因此 40 组关键帧图像不作预处理,直接作为肉兔图像光流计算网络的数据集。

从 2 000 幅关键帧中随机选取 1 000 幅图像进行预处理。为使数据集符合肉兔实例分割网络的训练要求,并最大程度保留图像信息,利用 OpenCV 扩充图像边界并将图像分辨率降低至 1 024 像素 × 1 024 像素。使用 Labelme 图像标注工具标注肉兔

轮廓,死兔和活兔均定义为“rabbit”类,共标注目标 4 873 个,制作为 COCO(Common objects in context)格式的数据集。随机选用图像 70%(700 幅)作为肉兔图像分割网络的训练集,30%(300 幅图像)作为肉兔图像分割网络的验证集,用于验证和优化网络。

2.3 肉兔图像分割网络训练和推断

使用 COCO 数据集上的 ResNet 预训练权重作为初始输入权重对肉兔实例分割网络进行迁移训练。引入 ResNet 50 和 PointRend 分别取代原有的残差网络和网络头,实现每幅图像的前景精细分割。采用随机梯度下降法进行训练,初始学习率为 0.001,采用线性增加策略,权重衰减系数为 0.000 1,优化器动能 0.9,共训练 800 个 epoch。选取精度最高的权重作为预测权重。

使用训练好的权重推断目标时,输出每幅图像的推断图像和图像中每个掩膜的边界框角点坐标,为后续死兔识别提供基础。

2.4 基于目标实例分割的光流提取

肉兔图像光流计算网络的训练集中包含 2 000 幅图像,可以按视频分为 40 组。使用肉兔实例分割网络和肉兔图像光流计算网络推断各组图像,最终输出每个肉兔掩膜的光流。图 7 展示了光流提取的过程。

在肉兔实例分割网络的推断中,使用相同颜色填充掩膜,即每幅图像推断出的第 1 个掩膜均为红色,第 2 个掩膜均为蓝色,以此类推。输出每幅图像的检测结果,输出过程如图 7a~7c 所示。每个视频输出 50 幅掩膜提取图。

在肉兔图像光流计算网络的计算过程中,对于每个视频,以每幅关键帧与后一幅关键帧作为输入,可以得到两帧之间的光流输出,如图 7d 所示。由于最后一帧没有下一幅图像与之对应,因此每组图像

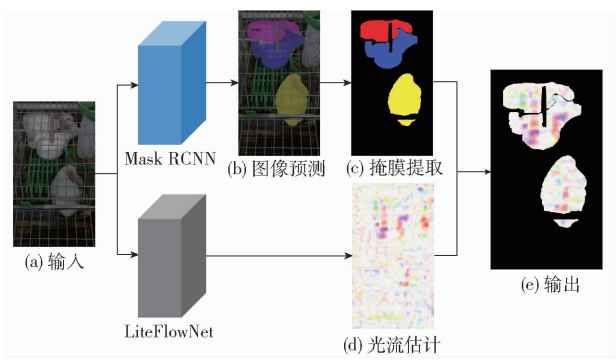


图 7 掩膜光流提取
Fig. 7 Extraction of mask optical flow

输出 49 幅光流估计图。

去除每组图像的最后一幅掩膜提取图,使掩膜提取图均和光流估计图一一对应。按掩膜的颜色分割光流估计图,可以输出图像中每个实例的光流,如图 7e 所示。

光流图中,RGB 颜色空间下,每个像素点的 RGB 通道值代表其运动方向和运动速度,则每个像素点的灰度可以表示像素点的运动,以此作为判断死兔的依据。

2.5 基于掩膜中心点核密度估计的死兔识别方法

本文提出的 Mask RCNN + PointRend 只针对图像进行实例分割,无法对肉兔进行多目标跟踪 (MOT),导致前后帧中所识别的肉兔掩膜不能代表同一只肉兔,单纯采用光流不能判断肉兔死活。目前在 MOT 领域中,SOTA 模型的目标跟踪精度大多低于 70% [22-24],且对目标类别和视频帧率均有要求,实用性差。因此本文提出了一种以掩膜中心点坐标为判断依据的死兔识别方法,并利用光流去除活兔干扰,原理为在 49 幅关键帧中,由于死兔一直保持静止,因此其每个掩膜的边界框中心点会集中在一处,而活兔在部分时间段保持静止,在部分时间段活跃,将活跃时刻的掩膜去除后,剩余时刻的掩膜边界框中心点会分布松散,点的数量少于死兔。

肉兔实例分割模型输出如图 7c 所示的掩膜提取图、每个掩膜边界框的角点坐标和对应的掩膜序号。根据角点坐标计算边界框的中心点坐标,与掩膜序号相对应。计算图 7e 中每个掩膜处的光流和



图 8 关键帧中掩膜的光流结果
Fig.8 Mask optical flow result in key frames

掩膜序号,若掩膜的光流高于阈值,将对应的掩膜判断为活兔,去除对应序号的掩膜边界框中心点;若低于阈值,表明对应掩膜为非活跃的活兔或死兔。图 8 截取了某段视频中 20 幅关键帧中掩膜的光流结果,为便于展示,图像有剪裁。

图 8 中每幅图像底部的掩膜为死兔,其余均为活兔。第 1 幅图像和最后 1 幅图像分别为关键帧的首尾图像,其余图像从 47 幅关键帧中随机抽取。由图 8 可知,代表死兔的掩膜几乎全为白色,表明死兔掩膜的光流很小。而对于活兔,大部分掩膜具有鲜明的色彩,仅有少数帧中的少数掩膜为白色,这表明活兔掩膜的光流较大,而不活跃的肉兔在一段视频中仅占少数时间段。采用光流法去除活兔的干扰有可行性。

对每段视频统计 49 幅关键帧中,光流低于阈值的掩膜所对应的边界框中心点坐标,随后采用无参核密度估计 [25] (Kernel density estimation, KDE) 的方法分析中心点坐标密度。

KDE 是一种从数据样本出发研究数据分布特征、估计密度函数的方法,利用平滑峰值函数(核)拟合数据点模拟真实的概率分布曲线,核密度函数为

$$\hat{f}_h(p) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{p - p_i}{h}\right) \tag{3}$$

式中 n ——样本数量
 h ——平滑参数,带宽
 K ——核函数
 p ——所求点的坐标
 p_i ——第 i 个点的坐标

本文选用高斯函数作为核函数,反映随机变量的概率分布。核函数能够反映每个样本对采样点概率密度分布的贡献度,在本文中,死兔掩膜的边界框中心点密集,每个中心点对采样点的概率密度分布贡献度会显著高于活兔掩膜。通过统计点的概率密度,可以判断视频中是否存在死兔。

3 结果与分析

3.1 肉兔实例分割网络性能分析

采用准确率 (Accuracy, A)、平均精确率

(Average precision, AP) 和平均像素精度 (Mean pixel accuracy, MPA) 评价肉兔实例分割网络性能。

为表明 PointRend 网络结构对分割性能的影响,以相同数据集和超参数训练了 Mask RCNN,图 9 为在训练过程网络损失值和验证集准确率与迭代次数之间的关系曲线。

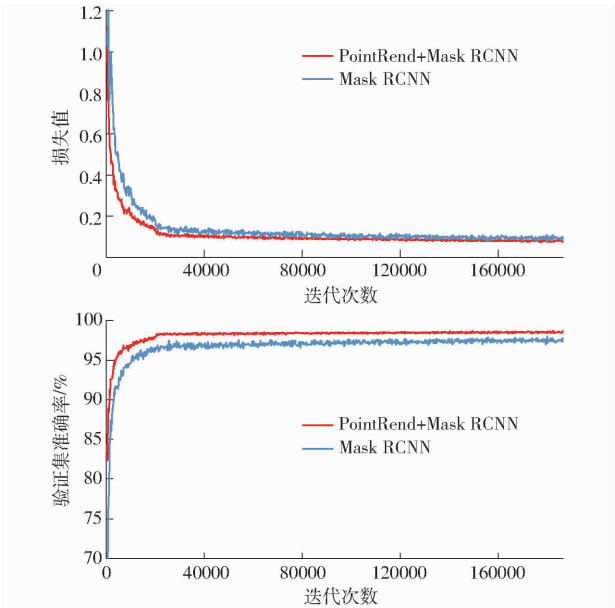


图 9 网络损失值与准确率对比

Fig.9 Total loss and accuracy comparison of two networks

由图 9 可知,两种网络的损失值随着训练迭代次数增加而降低,PointRend + Mask RCNN 的损失值收敛于 0.077,Mask RCNN 的损失值收敛于 0.090;两种模型的准确率随着训练迭代次数增加而增加,Mask RCNN + PointRend 的准确率收敛于 98.5%,Mask RCNN 的准确率收敛于 97.3%,这表明采用 PointRend 取代原网络头后,肉兔图像分割网络的准确率更高,降低了 Mask 部分的损失。

图 10 展示了 2 个网络对同一幅图像的部分掩膜边缘分割结果。PointRend + Mask RCNN 网络对掩膜边缘的分割更为平滑,相比于 Mask RCNN,分割出的掩膜更能代表目标。

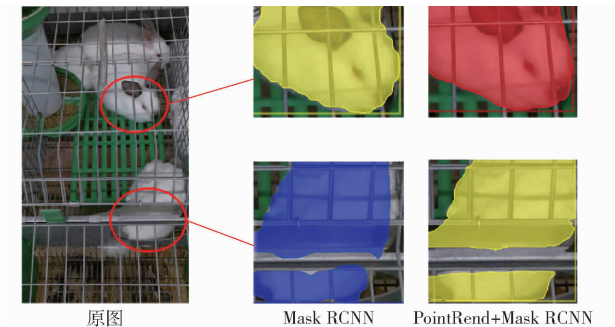


图 10 两种网络对掩膜边缘的分割结果

Fig.10 Mask boundary segmentation results of two networks

利用 COCOAPI 对验证集的推断结果进行性能

评估,结果如表 1 所示。

表 1 两种网络对肉兔的分割性能对比
Tab.1 Comparison of rabbit segmentation performance between two networks

网络	平均精确率/%	平均像素精度/%	运算时间/s
Mask RCNN	93.1	91.6	0.144
PointRend + Mask RCNN	96.1	95.7	0.121

平均精确率反映网络在目标识别中的分类准确度,表示分类为正样本中分类正确的数量占比。平均像素精度反映网络在实例分割中像素分类的准确度,表示正确分割的像素数量在图像像素数量中的占比。根据表 1 可知,PointRend + Mask RCNN 对验证集的分类性能比 Mask RCNN 提高了 3.0 个百分点,实例分割性能提高了 4.1 个百分点,算法推断 1 幅图像的速度提高了 16.0%。

3.2 死兔识别模型的识别结果与分析

基于肉兔图像分割网络和光流计算网络,识别并去除大部分活兔的干扰。首先提取每组关键帧中肉兔掩膜的光流,输出每个掩膜的总灰度。由于灰度越高,亮度越接近白色,对应像素点的运动速度越低,因此对灰度进行反运算,以反运算灰度均值表示光流值。掩膜的反运算灰度均值计算公式为

G_m = 255 - G_p / N_p (4)

式中 G_m——掩膜的平均灰度
G_p——掩膜的总灰度
N_p——掩膜的像素点数量

掩膜的反运算灰度均值表示掩膜像素点的运动程度,均值越高,对应的掩膜越可能是活兔。利用阈值筛除活兔,以包含 1 只死兔、3 只活兔的某段视频为例,图 11 展示了该段视频中 50 幅关键帧所提取的掩膜反运算灰度均值。

如图 11 所示,黄色三角点表示死兔掩膜的反运算灰度均值。在 48 个黄色三角点中,仅有 3 个

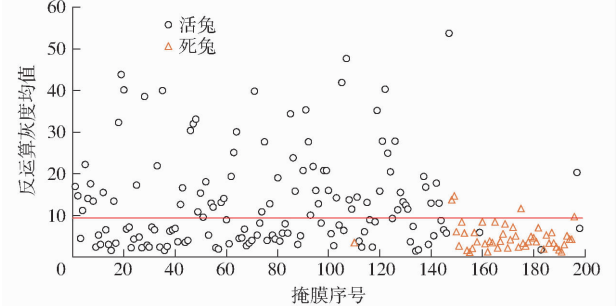


图 11 掩膜反运算灰度均值分布

Fig.11 Distribution diagrams of mean inverse operation gray value

点的反运算灰度均值高于 10。而活兔掩膜的反运算灰度均值在 0~55 之间均有分布。若将阈值选为 15, 会多出大量活兔掩膜形成干扰。对训练集中 40 个视频进行处理和统计, 最终将阈值确定为 10。

统计全部的掩膜边界框中心点坐标, 将掩膜

的反运算灰度均值高于 10 所对应的中心点坐标去除, 图 12 以两段视频为例, 描述掩膜边界框中心点的分布规律, 图中横纵坐标表示像素点的列序号、行序号。其中, video_01 为包含 1 只死兔、3 只活兔的视频, video_03 为包含 0 只死兔、4 只活兔的视频。

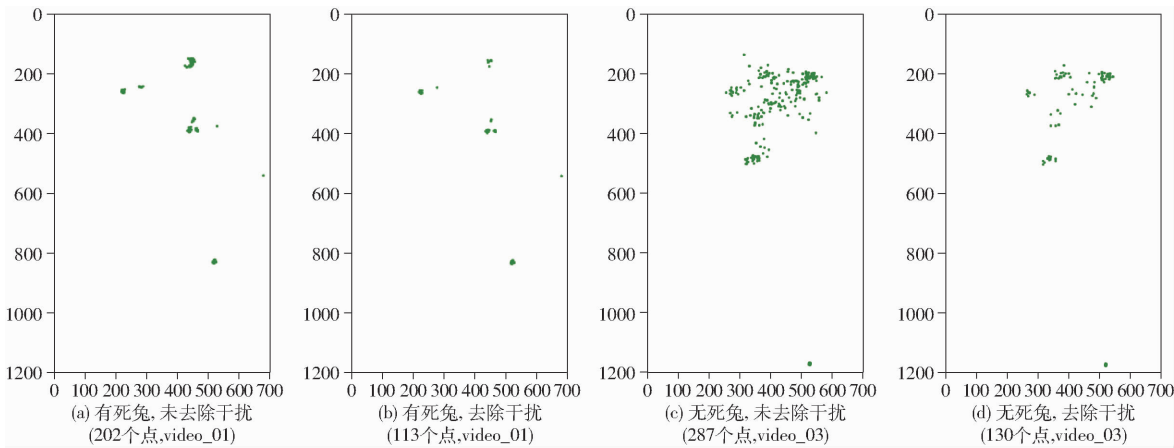


图 12 掩膜边界框中心点分布图

Fig. 12 Center point distributions of mask bounding box

如图 12a、12b 所示, 阈值将原本 202 个坐标点减少至 113 个, 死兔的坐标点基本重叠在同一位置, 剩余的活兔坐标点较为松散。如图 12c、12d 所示, 光流信息将原本 287 个坐标点减少至 130 个, 效果较为显著。

通过 KDE 对去除干扰后的中心点做密度分析, 其可视化结果如图 13 所示, 图中横、纵坐标表示像素点的列序号、行序号。

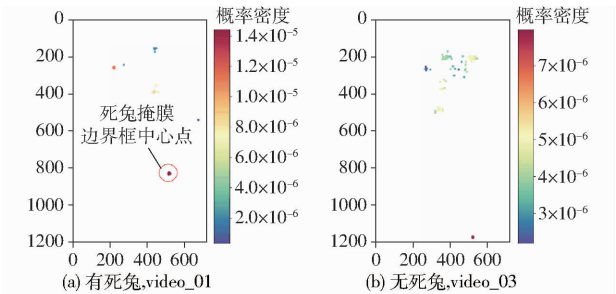


图 13 中心点核密度估计可视化结果

Fig. 13 Visualization results of center point kernel density estimation

图 13a 中, 代表死兔掩膜的中心点概率密度高于 1.4×10^{-5} , 代表活兔掩膜的中心点概率密度低于 1.1×10^{-5} 。图 13b 中, 所有活兔的中心点概率密度低于 8×10^{-6} 。

通过对训练集的 40 个视频进行中心点核密度估计, 最终确定了死兔的概率密度阈值为 1.25×10^{-5} , 当存在中心点概率密度高于 1.25×10^{-5} 时, 视频中包含死兔。

使用死兔识别模型对验证集的 20 个视频进行

死兔识别, 其中有 9 个视频存在死兔, 11 个视频全为活兔。结果表明, 存在死兔且模型检测到死兔的视频数为 9 个, 全为活兔且模型未检测到死兔的视频数为 9 个, 存在死兔但模型未检测到死兔的视频数为 0 个, 全为活兔但模型误检测到死兔的视频数为 2 个, 因此模型判断死兔的准确率为 90%。

检测结果表明, 对于存在死兔的视频, 死兔识别模型能够全部检出。对于全为活兔的视频, 死兔识别模型检测存在错误。人工检验后发现, 这是由于视频中的肉兔活跃性较低, 在 1 min 内的移动过少。在实际养殖过程中, 可以适当延长采集视频的时长, 或选择肉兔采食期间采集视频, 以提高模型的识别准确度。

4 结论

(1) 以笼养肉兔为研究对象, 通过帧间差分方法提取肉兔视频的关键帧, 并建立了一种基于图像分割和光流计算的死兔识别模型, 实现多目标背景下的肉兔图像分割和死兔检测。

(2) 构建了基于 PointRend + Mask RCNN 的肉兔图像分割网络, 以 ResNet50 和特征金字塔作为特征提取网络, 采用 PointRend 算法替换原有全卷积网络头, 通过自适应的点细分策略实现对肉兔掩膜边缘的精细提取。构建的肉兔图像分割网络在验证集上的平均精确率和平均像素精度分别为 96.1% 和 95.7%, 相比于 Mask RCNN 网络分别提升了 3.0 个百分点和 4.1 个百分点, 单幅图像检测速度

提升了 16.0%。

(3)构建了基于 LiteFlowNet 的肉兔光流计算网络,以连续关键帧为输入,通过特征金字塔网络和多级反卷积网络获取基于关键帧的肉兔掩膜光流信息。

(4)以图像分割结果和光流计算结果为基础,提出了一种基于掩膜中心点核密度估计的死兔识别

方法,确定了活跃肉兔的掩膜反运算灰度均值的阈值为 10,确定了死兔的掩膜边界框中心点的概率密度阈值为 1.25×10^{-5} 。

(5)死兔识别模型对验证集视频的准确率为 90%,误差原因是肉兔活跃度过低,通过延长视频时长和选择肉兔活跃时刻拍摄视频能提高模型准确性。

参 考 文 献

[1] COOKE B. Long-term monitoring of disease impact: rabbit haemorrhagic disease as a biological control case study [J]. Veterinary Record, 2018,182(20): 571 – 572.

[2] COCCHI M, DRIGO I, BACCHIN C, et al. Toxin-genotyping of Clostridium perfringens strains isolated from rabbits with enteric disease [C] // World Rabbit Congress, 2008:921 – 924.

[3] AGNOLETTI F. Update on rabbit enteric diseases: despite improved diagnostic capacity, where does disease control and prevention stand [C] // Proceedings 10th World Rabbit Congress, World Rabbit Science Association, 2012: 3 – 6.

[4] CUAN K, ZHANG T, HUANG J, et al. Detection of avian influenza-infected chickens based on a chicken sound convolutional neural network [J]. Computers and Electronics in Agriculture, 2020,178:105688.

[5] 张春慧,宣传忠,于文波,等. 基于三轴加速度传感器的放牧羊只牧食行为研究 [J]. 农业机械学报, 2021,52(10):307 – 313. ZHANG Chunhui, XUAN Chuanzhong, YU Wenbo, et al. Grazing behavior of herding sheep based on three-axis acceleration sensor [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(10):307 – 313. (in Chinese)

[6] 杜晓冬,滕光辉,TOMAS Norton,等. 基于声谱图纹理特征的蛋鸡发声分类识别 [J]. 农业机械学报, 2019,50(9):215 – 220. DU Xiaodong, TENG Guanghui, TOMAS Norton, et al. Classification and recognition of laying hens' vocalization based on texture features of spectrogram [J]. Transactions of the Chinese Society for Agricultural Machinery, 2019,50(9):215 – 220. (in Chinese)

[7] 何东健,宋子琪. 基于热红外成像与骨架树模型的奶牛眼温自动检测 [J]. 农业机械学报, 2021,52(3):243 – 250. HE Dongjian, SONG Ziqi. Automatic detection of dairy cow's eye temperature based on thermal infrared imaging technology and skeleton tree model [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(3):243 – 250. (in Chinese)

[8] 王少华,何东健. 基于改进 YOLO v3 模型的奶牛发情行为识别研究 [J]. 农业机械学报, 2021,52(7):141 – 150. WANG Shaohua, HE Dongjian. Estrus behavior recognition of dairy cows based on improved YOLO v3 model [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(7):141 – 150. (in Chinese)

[9] 方鹏,郝宏运,李腾飞,等. 基于注意力机制和可变形卷积的鸡只图像实例分割提取 [J]. 农业机械学报, 2021,52(4):257 – 265. FANG Peng, HAO Hongyun, LI Tengfei, et al. Instance segmentation of broiler image based on attention mechanism and deformable convolution [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(4):257 – 265. (in Chinese)

[10] 谷子林,秦应和,任克良. 中国养兔学 [M]. 北京:中国农业出版社, 2013.

[11] 罗敏,刘洞波,文浩轩,等. 基于背景差分法和帧间差分法的车辆运动目标检测 [J]. 湖南工程学院学报(自然科学版), 2019,29(4): 58 – 61. LUO Min, LIU Dongbo, WEN Haoxuan, et al. A new vehicle moving detection method based on background difference and frame difference [J]. Journal of Hunan Institute of Engineering (Natural Science Edition), 2019,29(4): 58 – 61. (in Chinese)

[12] HE K, GEORGIA G, PIOTR F, et al. Mask R – CNN [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017,36(9): 1137 – 1149.

[13] REN S, HE K. Faster R – CNN towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis Machine Intelligence, 2017,36(9): 1137 – 1149.

[14] LIN T, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:2117 – 2125. .

[15] KIRILLOV A, WU Y, HE K, et al. PointRend: image segmentation as rendering [C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020:9799 – 9808.

[16] ZHANG Z, LYONS M, SCHUSTER M, et al. Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron [C] // Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition. IEEE, 1998:454 – 459.

[17] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines [C] // ICML, 2010:807 – 814.

[18] LUCAS B D, KANADE T. An iterative image registration technique with an application to stereo vision (DARPA) [J]. Proc. IJCAI, 1981,81(3): 674 – 679.

[19] HORN B K, SCHUNCK B G. Determining optical flow [J]. Artificial Intelligence, 1981,17(1 – 3): 185 – 203.