

用于体质量估测的黄羽鸡姿态关键帧识别与分析

张小敏^{1,2} 徐涛^{1,2} 张延宁^{1,2} 高源^{1,2} 朱逸航^{1,2} 饶秀勤^{1,2}

(1. 浙江大学生物系统工程与食品科学学院, 杭州 310058; 2. 浙江省农业智能装备与机器人重点实验室, 杭州 310058)

摘要: 体质量是评价家禽生长状况的关键指标,但家禽姿态的变化会影响体质量估测精度。本研究提出了一种 SE-ResNet18 + fLoss 网络对平养模式下黄羽鸡姿态关键帧进行识别,融合了注意力机制 SE 模块和残差结构,并改进了损失函数,通过 Focal Loss 监督信号来解决样本不平衡问题,同时引入梯度加权类激活图对末端分类规则的合理性进行解释。利用 4 295 幅鸡只图像构建数据集,测试集中鸡只的站立、低头、展翅、梳理羽毛、坐姿和遮挡 6 类姿态情况识别的 F1 值分别为 94.34%、91.98%、76.92%、93.75%、100% 和 93.68%;黄羽鸡姿态关键帧的识别精确率为 97.38%、召回率为 97.22%、F1 值为 97.26%、识别速度为 19.84 f/s,识别精度、召回率和 F1 值均优于 ResNet18、MobileNet18 V2 和 SE-ResNet18 网络,在提高黄羽鸡姿态关键帧识别精度的同时保证了实时性,为准确估测家禽体质量提供了技术支持。

关键词: 黄羽鸡; 姿态识别; 体质量估测; SE-ResNet; Focal Loss

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2022)12-0254-10 **OSID:** 

Posture Key Frame Recognition and Analysis for Weight Estimation of Yellow-feathered Chickens

ZHANG Xiaomin^{1,2} XU Tao^{1,2} ZHANG Yanning^{1,2} GAO Yuan^{1,2} ZHU Yihang^{1,2} RAO Xiuqin^{1,2}

(1. College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China

2. Key Laboratory of Intelligent Equipment and Robotics for Agriculture of Zhejiang Province, Hangzhou 310058, China)

Abstract: Body weight is a key indicator to evaluate the growth condition of poultry. However, the variation of poultry posture will affect the accuracy of weight estimation. SE-ResNet18 + fLoss network was proposed to detect the posture key frames of floor-reared yellow-feathered chickens. The attention mechanism SE module and residual structure were integrated. And the Focal Loss was added to solve the problem of sample imbalance. In addition, the Gradient-weighted Class Activation Mapping was introduced to explain the rationality of the end classification rule. The dataset was constructed by 4 295 images of yellow-feathered chickens. The F1-score of the SE-ResNet18 + fLoss model on the test set for the chicken situations recognition of six classes: standing, bowing head, spreading wing, grooming feather, sitting and occlusion were 94.34%, 91.98%, 76.92%, 93.75%, 100% and 93.68%, respectively. Towards the detection of key posture frames on chickens, the accuracy, recall, F1-score and detection speed were 97.38%, 97.22%, 97.26% and 19.84 f/s, respectively. And the detection accuracy, recall and F1-score were better than those of ResNet18, MobileNet V2 and SE-ResNet18 networks. The study ensured real-time performance while improving the accuracy of key posture frame recognition, which provided technical support for accurate estimation of poultry weight.

Key words: yellow-feathered chickens; posture recognition; weight estimation; SE-ResNet; Focal Loss

0 引言

体质量是家禽的重要生长指标之一,也是种鸡

选育的重要参考依据。家禽的体质量与经济效益直接相关,体质量信息可用于判断家禽最佳出栏时间、指导日常饲喂、判断是否患病等^[1-2]。目前家禽体

收稿日期: 2022-08-01 修回日期: 2022-09-19
基金项目: 浙江省重点研发计划项目(2021C02026)
作者简介: 张小敏(1996—),女,博士生,主要从事智能农业装备研究,E-mail: 11813014@zju.edu.cn
通信作者: 饶秀勤(1968—),男,教授,博士生导师,主要从事智能农业装备研究,E-mail: xqrao@zju.edu.cn

质量测量可分为人工测量法、称量装置测量法和机器视觉监测法。传统的人工测量法耗时、费力,易造成家禽的应激反应,且获得的体质量信息不全,不具有时间连续性^[3]。称量装置测量法不会引起家禽的应激反应,但此类传感器存在安装复杂、维护困难等问题,且在养殖后期,体质量较大的肉鸡相比其他肉鸡,上秤频次低,造成体质量低估^[4]。相比而言,机器视觉监测法具有实时客观、非侵扰性、可维护性强等优点,可以提高测量效率。

基于机器视觉技术的畜禽体质量估测已取得了一定的进展^[5-10],但目前主要依靠人工选取家禽理想姿态帧,主观性强且耗时费力,难以在家禽体质量自动估测中应用。因此,研究用于家禽体质量估测的理想姿态识别算法具有重要意义。

在动物个体姿态识别方面,LEROY 等^[11]通过点分布模型拟合蛋鸡图像,描述蛋鸡的姿势信息,能够自动识别母鸡的站立、行走和抓挠行为。POURSABERI 等^[12]通过分析牛的背部姿态来识别跛行,通过检测背部轮廓的弧度并通过脊椎线上的选定点拟合圆来检测背部姿态。劳凤丹等^[13]基于 10 个特征量使用贝叶斯分类法对蛋鸡行为进行识别,能够识别单只蛋鸡的拍翅膀、修饰、休息、探索、抖动和举翅膀等行为。劳凤丹等^[14]基于深度图像对蛋鸡行为识别进行了研究,结果表明算法对蛋鸡的采食、躺、站和坐的识别准确率分别为 90.3%、91.5%、87.5% 和 56.2%。司永胜等^[15]提出了一种基于 Kinect 的猪体理想姿态检测算法以提高图像利用率和体尺测量效率,算法利用投影法和差分法识别头部和尾部位置,利用骨骼化算法结合霍夫变换算法检测猪体头部是否歪斜,对于 52 016 帧图像检测出理想姿态 2 592 帧、漏报 432 帧、误报 0 帧。

随着深度学习技术的发展,卷积神经网络技术在动物个体姿态识别领域得到了较多应用^[16-18]。文献[19-20]提出了采用单幅图像识别猪只立卧姿态的方法,利用建立的猪只识别模型和头尾识别方法,将待测猪图像调整到水平状态,构建了猪只立卧姿态识别模型,精度达到了 95%。LUO 等^[21]提出了一种轻量级的通道注意力模型,用于实时识别猪只的站立、趴卧、侧卧、坐姿和攀爬 5 种姿势,该算法对上述 5 种姿势的平均识别精度分别为 97.7%、95.2%、95.7%、87.5% 和 84.1%,单幅图像的推理速度约为 63 ms。薛月菊等^[22]提出了一种融合 2D-3D 卷积特征的哺乳母猪姿态转换识别算法,模型姿态转换识别准确率为 97.95%,召回率为 91.67%,测试速度为 14.39 f/s。FANG 等^[23]利用多部位目标检测方法对单只肉鸡的侧面图像进行姿态估计,

利用 ResNet50 和特征金字塔(FPN)结构^[24]构建特征提取器,算法对单个肉鸡姿态估计的准确率和召回率分别为 91.9% 和 86.5%。

基于机器视觉的家禽姿态识别相关研究已取得一定进展,但目前依旧存在以下问题:当鸡只处于低头、展翅、梳理羽毛或坐姿等姿态时,鸡只宽度、体积、背高等特征参数的提取会受到影响,易造成体质量预测的误差较大;而家禽处于站立姿态时,身体直立无多余动作,这样的姿态对于特征的提取更加稳定,因此站立姿态可认为是家禽体质量估计时的最理想姿态^[5,7-8],但目前适用于家禽体质量估计的站立姿态识别方法缺乏,且由于家禽为非合作对象,适于体质量估计的理想姿态在视频中出现较少,样本不均衡问题突出,研究者难以获得较好的理想姿态识别精度。

针对上述问题,本文以 882 黄鸡为试验对象,提出 SE-ResNet18+fLoss 网络,从视频中筛选符合体质量估计的姿态关键帧,模型融合注意力机制 SE 模块和残差结构,并改进损失函数,通过 Focal Loss 监督信号来解决样本不平衡问题,同时引入梯度加权类激活图(Gradient-weighted class activation mapping, Grad-CAM)^[25]对末端分类规则的合理性进行解释,反映每一类姿态所关注的重点区域。在连续视频上进行模型的验证与黄羽鸡姿态特性的分析,为家禽行为研究提供参考依据。通过试验验证本文提出的姿态关键帧识别模型对于提升黄羽鸡体质量估测精度的有效性。

1 试验材料

1.1 数据采集

试验数据采集于浙江省湖州市,共采集 4 个批次,采集时间分别为 2019 年 12 月 5—6 日、2020 年 8 月 22—24 日、2022 年 6 月 4 日和 2022 年 6 月 11 日。将 4 个批次的数据分别标记为 D1、D2、D3 和 D4,具体数据集如表 1 所示。其中,D1 为单只鸡连续 24 h 视频片段;D2 为 20 只鸡视频片段,其中有 18 只鸡平均采集 1.5 h,采集时间为 08:00—17:00,其余 2 只平均采集 15 h,采集时间为 17:00—08:00;D3 和 D4 为同一批 20 只鸡的 2 周次视频片段,每只鸡采集 10 min,采集时间为 08:00—12:00,采集视频的同时,利用电子秤获得其实际体质量,电子秤为大红鹰无线款 32 cm×42 cm,精度为 0.05 kg。本文的试验对象为 882 黄鸡,平养于 4.50 m×2.50 m 的空地区域,内侧配有料槽和饮水器,鸡只可自由采食和饮水。当采集数据时,将待测鸡只转移至试验笼中,试验笼尺寸为 1.45 m×0.97 m×1.30 m,光照时间段为 08:00—17:00。

表 1 试验数据集
Tab.1 Datasets of experiment

采集时间	批次	鸡只数 量/只	周龄/ 周	体质量/ kg	环境温 度/℃
2019 年 12 月 5—6 日	D1	1	16	1.10	3~12
2020 年 8 月 22—24 日	D2	20	18~42	1.20~2.30	24~33
2022 年 6 月 4 日	D3	20	16	0.90~1.55	23~31
2022 年 6 月 11 日	D4	20	17	0.95~1.55	18~21

如图 1 所示,试验笼主要由摄像机、网线、硬盘录像机、饮水器、脚垫、饲料槽和粪槽组成。笼子正上方架设了一台海康威视双光谱摄像机 (DS-2TD2636-10 型),摄像机机芯距离脚垫的高度为 1.45 m,摄像机通过网线连到硬盘录像机 (DS-8616N-I8 型),以 4 f/s 的速度拍摄彩色视频图像,单次采集仅拍摄一只鸡,图像分辨率为 1 920 像素 × 1 080 像素,并将采集到的视频数据以 avi 格式存储至硬盘录像机中,之后导入本地计算机进行视频分析,为了方便后续处理,所有视频都将解码为 jpg 格式的图像。

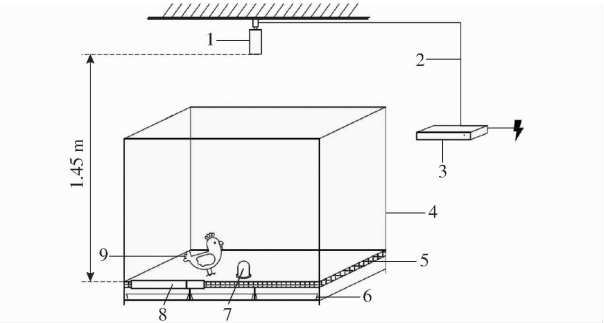


图 1 视频采集示意图
Fig.1 Sketch of video acquisition

1. 摄像机 2. 网线 3. 硬盘录像机 4. 笼子 5. 脚垫 6. 粪槽
7. 饮水器 8. 饲料槽 9. 鸡只

1.2 黄羽鸡姿态关键帧定义

按照家禽的行为特征,可将黄羽鸡姿态分为站立、低头、展翅、梳理羽毛和坐姿。其中,坐姿的定义包含了腹卧情况。此外,当鸡只被遮挡时也无法进行体尺参数提取,故将遮挡情况也加入数据集进行识别,共 6 种姿态情况,表 2 为每种姿态情况定义的详细说明。本文提出只有当鸡只处于站立姿态时,当前帧才被认为是姿态关键帧,方可进行后续的体尺参数提取与体质量估测研究。

1.3 数据集准备

D1、D2 两批次采集的视频文件共 68 个,单个文件时长为 80 min 左右。将所有视频解码为图像后,从 D2 批次中挑选 4 295 幅图像作为数据集,覆盖每一只鸡。最终得到站立、低头、展翅、梳理羽毛、坐姿和遮挡情况的姿态图像数量分别为 1 017、968、

表 2 黄羽鸡姿态情况定义
Tab.2 Definition of yellow-feathered chicken posture conditions

姿态	姿态描述	姿态示意图
站立	身体与腿部垂直,背部朝上,腹部朝下,腿部直立,双脚接触地面	
低头	头部向下接近地面,尾部翘起,双脚接触地面	
展翅	有一侧或两侧翅膀张开	
梳理羽毛	用喙部轻轻摩擦、梳理自身羽毛,头部向背部羽毛靠近,呈蜷缩状	
坐姿	腹部坐或卧在地板上,腹部接触地面,身体呈匍匐状放松	
遮挡情况	鸡身被部分或完全遮挡,处于视场外	

74、802、507、927 幅,图像分辨率均为 1 920 像素 × 1 080 像素。将所有图像人工标注姿态类别,每个类别按照 7:2:1 的比例划分为训练集、验证集和测试集,训练集总计 3 002 幅,验证集 861 幅,测试集 432 幅。数据集中各个姿态类别的数量如表 3 所示,观察表 3 可知,数据集中 6 种姿态类别数量相差较大,会带来数据分布不均衡问题。此外,D1 批次视频数据也作为测试集以分析鸡姿态分布特性,视频采样间隔为 1 min/幅,共得到 1 935 幅图像。D3 和 D4 批次视频数据作为 4.5 节中不同姿态下体质量估测结果分析的数据集,分别从 D3 和 D4 批次视频数据中截取 3 min 进行采样,采样间隔为 1 s/幅,共得到 7 200 幅图像,用于后续体质量估测研究。

2 鸡只姿态关键帧识别模型

2.1 模型网络结构

建立 SE-ResNet18 + fLoss 模型,将 SE 模块^[26]

表 3 数据集中各个姿态类别图像数量

Tab.3 Number of different postures in dataset 幅

数据集	站立	低头	展翅	梳理羽毛	坐姿	遮挡情况	总计
训练集	711	677	51	561	354	648	3 002
验证集	204	194	15	160	102	186	861
测试集	102	97	8	81	51	93	432
总计	1 017	968	74	802	507	927	4 295

与 ResNet18^[27] 进行融合,并改进损失函数,通过 Focal Loss^[28] 监督信号来解决样本不平衡问题,网络结构如图 2 所示。SE 模块的核心思想是通过网络根据损失函数去学习特征权重,使得有效的特征图权重高,而无效或效果不显著的特征图权重小,筛选出针对通道的注意力,通过引入 SE 模块,模型能够从网络特征通道层面进一步提升网络性能。

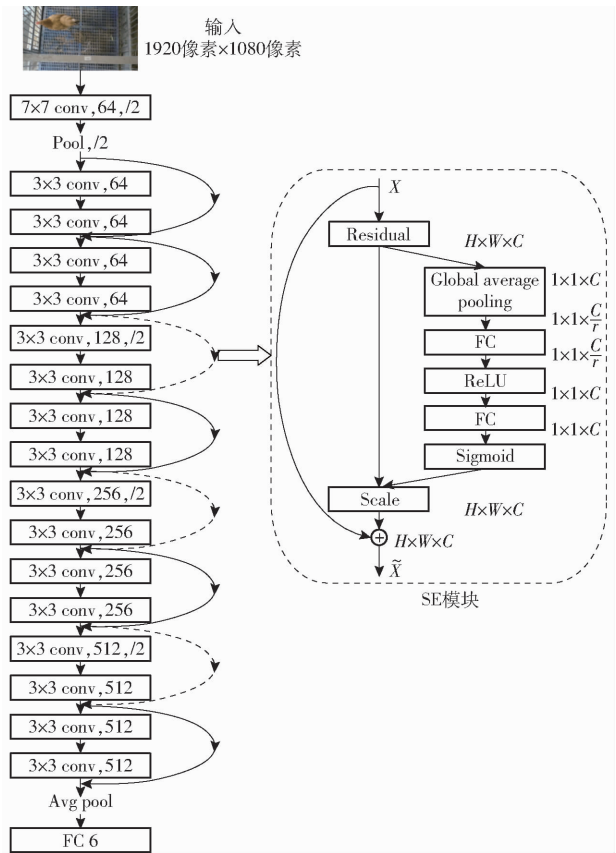


图 2 SE-ResNet18 + fLoss 网络结构图

Fig.2 Network architecture of SE-ResNet18 + fLoss

如图 2 所示,SE-ResNet18 + fLoss 是将 SENet-block 嵌入到 18 层的残差网络结构中,一个完整的 SENet-block 是通过压缩 (Squeeze)、激励 (Excitation) 和重标定 (Reweight) 3 个操作完成。对于 SE-ResNet18 中的残差模块,假设输入的图像为 $H \times W \times C$,使用全局池化 (Global average pooling) 作为 Squeeze 操作,首先经过 FC 层将特征维度降为输入的 $1/r$,此时输入的特征图变为 $1 \times 1 \times \frac{C}{r}$,然后经过 ReLU 激活函数处理后,输入到第 2 个 FC 层,将

特征维度升回原来的维度,接着使用 Sigmoid 函数将权重归一化为概率的形式,完成 Excitation 操作,最后根据每个通道特征的重要程度利用 Scale 操作将权重加权到每个通道的特征上,完成 Reweight 操作,即实现注意力机制。

在连续的鸡只视频数据中,6 类姿态的数量分布差异较大,尤其是展翅姿态,数量较少。针对样本类别不平衡问题,本文通过 Focal Loss 损失函数减少易分类样本的权重,使得模型在训练时更专注于难分类的样本。Focal Loss 损失是基于标准交叉熵损失的基础上修改得到的,对于二分类问题,交叉熵损失函数可以表示为

$$CE(p_i) = -\ln p_i \tag{1}$$

其中

$$p_i = \begin{cases} p & (y = 1) \\ 1 - p & (y \neq 1) \end{cases} \tag{2}$$

式中 p_i ——预测为真实类别的概率

CE——交叉熵损失函数

p ——模型对于类别 $y = 1$ 所得到的预测概率

为了解决类不平衡问题,Focal Loss 首先添加权重因子 α_i ,其次,在交叉熵前添加权重因子 $(1 - p_i)^\gamma$,以减少置信度很高的样本损失在总损失中的比重,最终得到 Focal Loss 为

$$L_F(p_i) = -\alpha_i (1 - p_i)^\gamma \ln p_i \tag{3}$$

式中 α_i ——权重因子,用于抑制正负样本的数量失衡, $\alpha_i \in [0, 1]$

L_F ——Focal Loss 损失函数

γ ——调制因子,用于增加难分样本的损失比例, $\gamma \geq 0$

2.2 模型可解释性

由于本文的模型为图像分类网络,无法像目标检测网络那样给出目标边界框,为了对姿态识别结果进行可视化解释,采用 Grad-CAM 来解释网络进行分类时重点关注的区域。Grad-CAM 相比于 CAM 更具一般性,其利用网络反向传播的梯度计算出特征图每一个通道的权重从而得到热力图,因此 Grad-CAM 可以直接用于各种卷积神经网络,无需重新训练网络。

由于最后一层特征图具有丰富的高维语义信息和空间信息,对卷积神经网络最后一层的所有特征图求偏导,得到特征层 A 中任一通道 k 的权重 α_k^c ,其计算公式为

$$\alpha_k^c = \frac{1}{Z} \left(\sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \right) \tag{4}$$

式中 Z ——特征图的像素个数

y^c ——网络针对类别 c 预测的得分,这里没有经过 softmax 激活

A_{ij}^k ——特征层 A 第 k 通道在坐标 (i, j) 的值

接着通过权重 α_k^c 对特征层 A 每个通道进行加权求和得到热力图,最后通过 ReLU 激活函数处理得到类别激活热力图 $L_{Grad-CAM}^c$,只保留对于类别 c 具有正作用的区域, $L_{Grad-CAM}^c$ 计算公式为

$$L_{Grad-CAM}^c = \text{ReLU} \left(\sum_k \alpha_k^c A^k \right) \tag{5}$$

式中 A^k ——特征层 A 的第 k 通道

2.3 模型训练环境

模型训练平台为超威服务器,CPU 为 Intel(R) Xeon(R) Gold 6147M 2.50 GHz, GPU 为 NVIDIA TITAN RTX,内存 256 GB,操作系统为 Ubuntu 18.04.1,深度学习框架为 Pytorch, CUDA 版本为 10.0。

2.4 模型评价指标

采用精确率(Precision, P)、召回率(Recall, R)、F1 值和识别速度 4 个指标来评价模型的识别性能。精确率表示正确预测为正样本占全部预测为正样本的比例,召回率表示正确预测为正样本占全部实际为正样本的比例,F1 值为精确率和召回率的调和平均值,识别速度为每秒内处理的图像帧数。考虑到每个类别样本数量不均衡,最终模型的总体评价指标均采用加权平均值(weighted avg)。

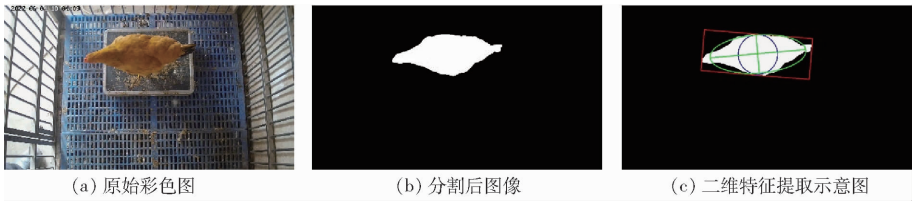


图3 图像分割与二维特征提取示意图

Fig. 3 Schematics of image segmentation and two-dimensional feature extraction

3.2 体质量估测模型建立

体质量估测模型采用随机森林(Random forest, RF)^[30]算法进行建立,RF 是一种集成学习方法,在以决策树为基学习器构建 Bagging 集成的基础上,进一步引入了随机属性选择,算法简单、高效且计算开销小。RF 回归算法的具体步骤为:①从原始数据集中随机有放回地抽取 n 个训练样本,重复 T 次。②对于 T 个子数据集,分别训练一个 CART 回归树模型。在训练过程中,每个结点的划分规则是先随机从所有特征中选取 k 个特征,再从 k 个特征中选择最优属性用于划分。③将生成的 T 棵决策树组成随机森林,随机森林最终的预测结果为所有 T 棵 CART 回归树预测结果的均值。

将 3.1 节得到的特征参数(周龄 O 、投影面积 A 、周长 P 、宽度 D 、长度 L 、最大内切圆半径 R 和偏心率 E)作为输入,实际体质量作为真实标签,建立

3 体质量估测方法

3.1 图像特征提取

为了验证本文提出的姿态关键帧对于提升黄羽鸡体质量估测精度的有效性,进一步设计了称量试验,对 D3 和 D4 批次视频数据采样得到的数据集进行分析,比较不同姿态下黄羽鸡体质量估测的精度。

首先采用 Unet 网络^[29]对 D3 和 D4 批次视频数据采样得到的 7 200 幅数据进行分割,Unet 模型采用 D1 和 D2 两批次数据中采样得到的 1 000 幅图像按照 7:2:1 的比例划分训练集、验证集和测试集进行训练得到,对原始鸡只彩色图进行分割后的图像如图 3b 所示。接着对分割后的图像进行一维和二维特征提取,用于后续黄羽鸡体质量估测模型的建立。参照文献[6-9],本文提取的一维特征包括周龄 O ,二维特征包括投影面积 A 、周长 P 、宽度 D 、长度 L 、最大内切圆半径 R 和偏心率 E ,二维特征提取的示意图如图 3c 所示。其中,投影面积 A 为分割前景的像素面积,周长 P 为分割前景的轮廓像素数,宽度 D 和长度 L 分别为分割前景的最小外接矩形的宽度和长度,最大内切圆半径 R 为分割前景的最大内切圆半径,偏心率 E 为椭圆拟合分割前景的偏心率。

基于随机森林的黄羽鸡体质量估测模型。数据集按照 7:3 的比例划分训练集和测试集,并采用均方根误差(RMSE)以及相对误差(MRE)对测试集中不同姿态下体质量估测的效果进行评价。

4 结果与分析

4.1 模型训练与参数优化

本文采用小批量的随机梯度下降法,对 SE-ResNet18 + fLoss 模型进行训练。训练过程中,迭代次数(Epochs)设置为 200 次,最小批处理大小(MiniBatch size)为 16。学习率初始化为 0.01,并采用学习率衰减训练策略调整学习率,学习率调整如表 4 所示。

为了解决数据分布不平衡的问题,使用 Focal Loss 损失函数提升分类精度,训练过程中 SE-ResNet18 + fLoss 模型的损失函数曲线和验证集准

表 4 学习率调整策略

Tab. 4 Learning rate adjustment strategy				
迭代次数	0 ~ 5	5 ~ 25	25 ~ 50	50 ~ 200
学习率	0.01	0.001	0.000 1	0.000 01

确率曲线如图 4 所示。由图 4 可以看出,训练集的损失函数随迭代次数的增加而下降,而准确率随迭代次数的增加而逐渐增加,当迭代次数达到 100 次左右时,模型基本收敛,训练损失函数曲线和准确率曲线都趋于平稳。

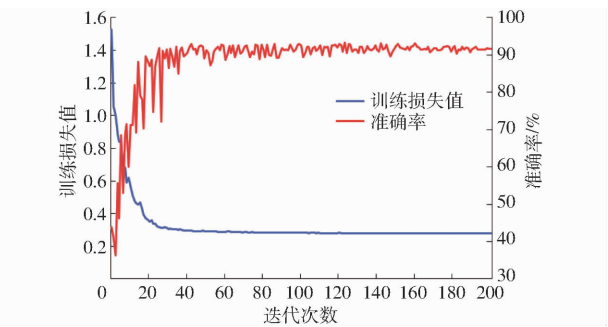


图 4 模型训练损失函数曲线和验证集上准确率曲线

Fig. 4 Model training loss function curve and accuracy curve

Focal Loss 损失函数最重要的两个参数是调制因子 γ 和权重因子 α_i ,本文对于调制因子 γ 和权

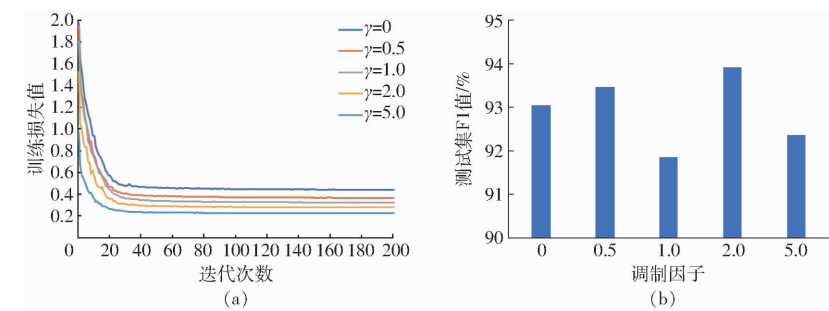


图 5 不同调制因子 γ 和权重因子 α_i 的模型性能

Fig. 5 Model performance with different modulation factors γ and weighting factors α_i

4.2 模型姿态识别结果分析

图 6 为 SE - ResNet18 + fLoss 模型在 432 幅测试集图像上测试结果的混淆矩阵,表 5 为各姿态情况分类的精确率、召回率及 F1 值。结果表明模型对于站立、低头、展翅、梳理羽毛、坐姿和遮挡情况 6 种类别识别的 F1 值分别为 94.34%、91.98%、76.92%、93.75%、100% 和 93.68%,6 类姿态的加权平均 F1 值达到 93.90%,具有较高的识别精度和较强的鲁棒性,模型对于夜间昏暗光线场景下的鸡只姿态也能进行很好的识别。进一步发现,模型对于坐姿的识别效果最好,主要是由于鸡只坐姿通常发生在晚上,光照与白天有明显不同,而对于展翅姿态的识别性能有待提高,这可能与展翅数据本身数量较少且展翅幅度不定有关。

重因子 α_i 的选取进行讨论。图 5a 为调制因子 γ 分别取 0、0.5、1.0、2.0、5.0 时模型的训练损失函数曲线,图 5b 为不同调制因子对应测试集 F1 值,当 $\gamma = 0$ 时,Focal Loss 等于标准交叉熵函数,可以发现随着 γ 的增大,损失曲线收敛更快,且越趋近于 0,通过图 5b 发现,当 $\gamma = 2.0$ 时,F1 值最高,故模型取 $\gamma = 2.0$ 。图 5c 为不同权重因子 α_i 对应在测试集上测试得到的 F1 值,固定调制因子 $\gamma = 2.0$,对于站立、低头、梳理羽毛、展翅、坐姿和遮挡情况 6 种类别的权重比例分别设置为: α_1 为 1:1:1:1:1:1, α_2 为 1:1:1:2:1:1, α_3 为 1:1:1:5:2:1, α_4 为 1:1:1:10:2:1, α_5 为 1:1:1:15:3:1, α_6 为 1:1:1:20:4:1,由于展翅和坐姿姿态的图像数量明显较少,本文将展翅和坐姿两种姿态的权重进行了增大,当权重因子为 α_i 时,表示未对正负样本的数量进行平衡。从图 5c 中可以发现,当各姿态类别的权重因子比例设置为 α_4 ,即根据样本的分布进行设置时,F1 值最高,将比值归一化到 $[0,1]$,可得 $\alpha_i = (0.0625, 0.0625, 0.0625, 0.625, 0.125, 0.0625)$ 时模型最优。此外,还可发现相比于 α_1 ,其他权重因子对应在测试集上的 F1 值都有显著提高,说明 Focal Loss 损失函数对模型分类效果有明显提升。

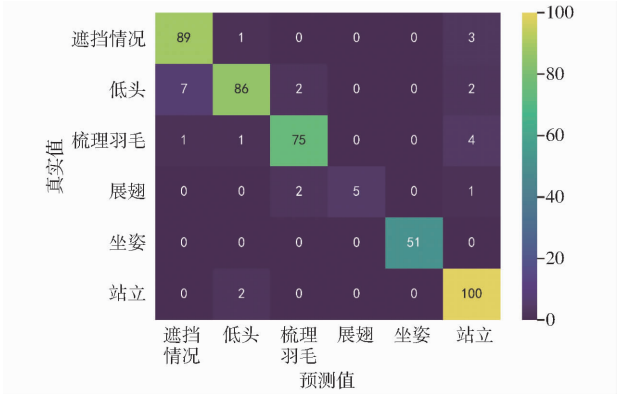
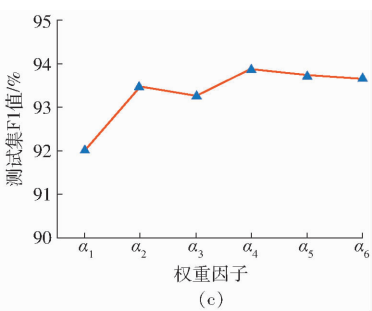


图 6 SE - ResNet18 + fLoss 模型对鸡只姿态情况分类的混淆矩阵

Fig. 6 Confusion matrix of chicken classification based on SE - ResNet18 + fLoss model

表 5 SE-ResNet + fLoss 模型对鸡只不同姿态分类结果

Tab. 5 Classification results of SE-ResNet + fLoss model for different postures of chickens %			
类别	精确率	召回率	F1 值
站立	90. 91	98. 04	94. 34
低头	95. 56	88. 66	91. 98
展翅	100	62. 50	76. 92
梳理羽毛	94. 94	92. 59	93. 75
坐姿	100	100	100
遮挡情况	91. 75	95. 70	93. 68
平均值	94. 13	93. 98	93. 90

由于本文最终旨在识别出姿态关键帧(站立姿态)用于后续体质量估测,故将非站立姿态看作一类,可得到表 6 所示模型对于姿态关键帧的识别性能,可以看出若仅识别是否为姿态关键帧,模型的识别精度为 97. 38%、召回率为 97. 22%、F1 值为 97. 26%,识别效果较优。

表 6 SE-ResNet + fLoss 模型对于鸡只姿态关键帧的识别结果

Tab. 6 Detection results of SE-ResNet + fLoss model for the posture key frame of chicken %			
类别	精确率	召回率	F1 值
站立(姿态关键帧)	90. 91	98. 04	94. 34
其他姿态情况	99. 38	96. 97	98. 16
加权平均值	97. 38	97. 22	97. 26

为了更好地解释本文提出的 SE-ResNet + fLoss 模型,对模型的最后一层卷积层的特征图利用 Grad-CAM 进行可视化,获得 6 种姿态的类激活图如图 7 所示,每一类姿态包括第 1 行的原始 RGB 图像和第 2 行的 Grad-CAM 可视化效果。通过可视化结果可以看出,不同姿态下模型的类激活图存在差异,反映了每一类所关注的重点区域不一致。其中,站立姿态重点关注于鸡头部位,而低头姿态关注于鸡的尾部和颈部,梳理羽毛姿态更关注于扭头时鸡的头部与身体部分,展翅姿态关注于身体翅膀展开部分,坐姿重点关注于鸡坐立时突出的侧面部分,遮挡情况则更关注于鸡在视场内的身体部分,此部分通常处于图像的边缘。

通过 Grad-CAM 可以很好地解释本文方法分类的依据,即使不采用目标检测网络,模型亦能准确定位鸡的位置,尤其是正常姿态下,如图 8 所示,Guided Grad-CAM 表现了细粒度的重要性,可以看出类激活图更关注于鸡头部位,这也为弱监督的目标检测提供了参考。

4.3 不同方法比较

本文将 SE-ResNet18 + fLoss 与 ResNet18、

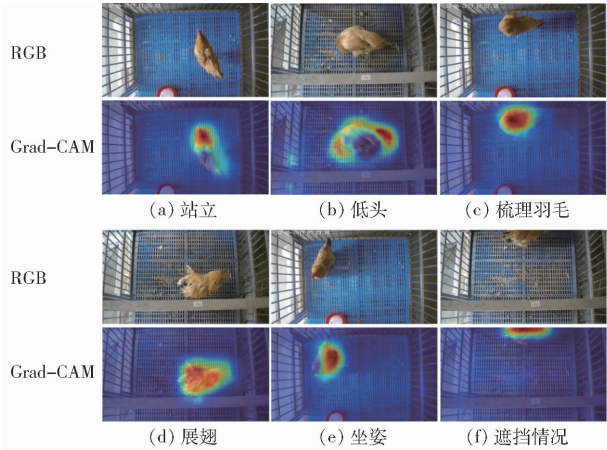


图 7 黄羽鸡不同姿态下分类的类激活图
Fig. 7 Class activation maps for classification of yellow-feathered chickens in different postures

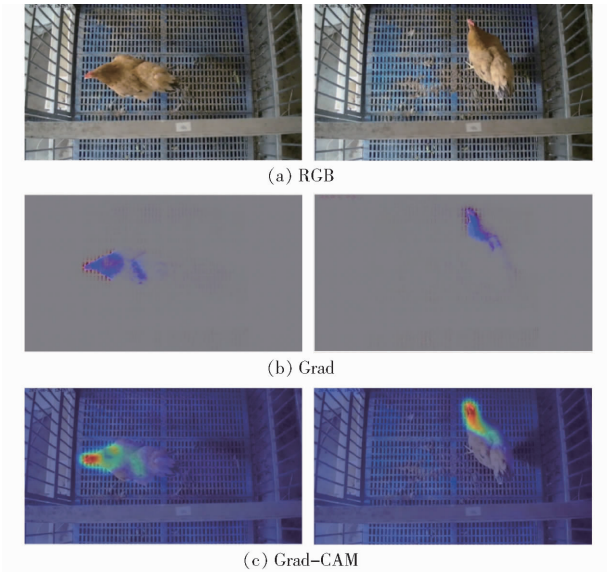


图 8 站立姿态下鸡只姿态识别的类激活图
Fig. 8 Class activation maps for posture detection of chicken in ‘stand’ pose

MobileNet V2、SE-ResNet18 进行了姿态识别对比试验。其中,MobileNet V2^[31]是 2018 年 Google 团队提出的轻量级网络。不同模型均采用相同的试验训练参数,并在相同测试集上验证模型的性能,识别速度是在批处理大小为 16 的情况下进行推理得到的,试验结果如表 7 所示。从表 7 可以看出,SE-ResNet + Focal Loss 在基础网络中加入了注意力机制 SE 模块和 Focal Loss,模型在精确率、召回率和 F1 值方面相比其他网络均最优,在识别速度方面 ResNet18 模型最优,但 SE-ResNet + fLoss 的识别速度与其相差不大。从结果可以看出本文改进的 SE-ResNet 能较为理想地兼顾识别性能与识别速度。

为了进一步说明不同方法姿态识别的结果,本文可视化 ResNet18、MobileNet V2、SE-ResNet18 和 SE-ResNet18 + fLoss 在 2 h 的视频段上的测试结

表 7 不同方法结果比较

Tab. 7 Comparison of results using different methods				
方法	精确率/	召回率/	F1 值/	识别速度/
	%	%	%	(f·s ⁻¹)
ResNet18	96.82	96.76	96.78	20.02
MobileNet V2	95.77	95.60	95.65	18.63
SE-ResNet18	96.05	96.06	96.06	17.79
SE-ResNet18 + fLoss	97.38	97.22	97.26	19.84

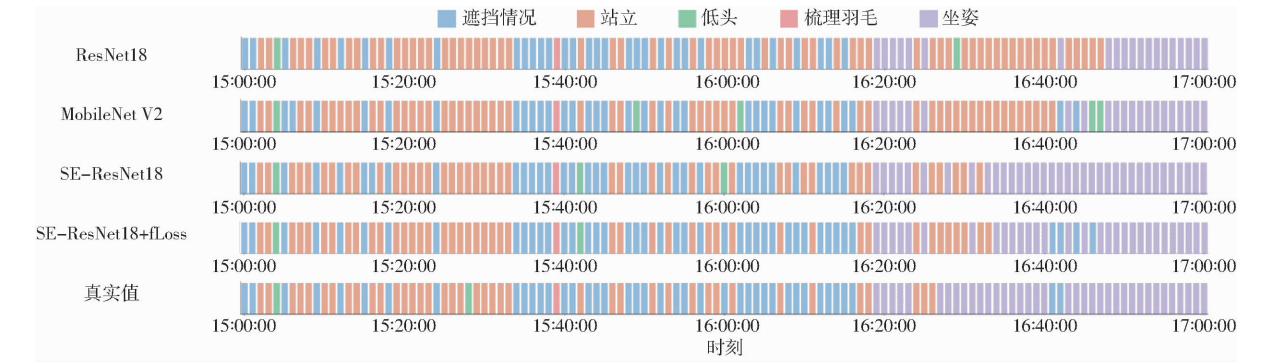


图 9 不同方法部分测试结果可视化
Fig. 9 Visualization of some test results of different methods

4.4 模型测试

图 10 为 SE-ResNet18 + fLoss 在连续 24 h 视频段上(2019 年 12 月 5 日 08:00—2019 年 12 月 6 日 08:00)的姿态自动识别结果。从图 10 可以看出,黄羽鸡白天多为站立姿态,而夜间多为坐姿,白天姿态转换频率显著高于夜间姿态转换频率。此外,2019 年 12 月 5 日 08:00—16:00 期间光照较为明亮,有利于提取鸡只的颜色和纹理特征,通过统计 SE-ResNet18 + fLoss 在 2019 年 12 月 5 日 08:00—16:00 视频段上的姿态占比结果(图 11),发现白天视频段中有超过 40% 的时间鸡只处于站立姿态,这意味着有相当一部分视频数据是姿态关键帧,可以用于体质量预测。

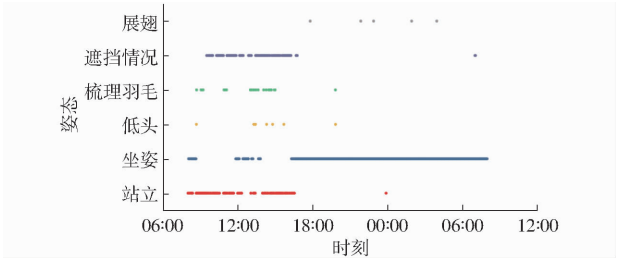


图 10 SE-ResNet18 + fLoss 在连续 24 h 视频段上的姿态自动识别结果

果,如图 9 所示,视频采样间隔为 1 min/幅。从图 9 可以看出,ResNet18 和 MobileNet V2 有较多图像将其他姿态误判为站立姿态,这对于本文进行姿态关键帧识别是不利的,因为这种误判将直接导致后续体质量预测的误差较大。与其他 3 种方法相比,SE-ResNet18 + fLoss 对于姿态的识别更加准确,能够较为全面地覆盖姿态转换发生的时间区间。

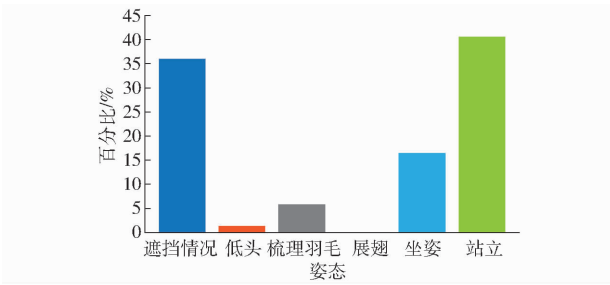


图 11 SE-ResNet18 + fLoss 在 2019 年 12 月 5 日 08:00—16:00 时间段的姿态占比统计结果

Fig. 11 Share statistics results of posture at December 5, 2019, 08:00—16:00 by SE-ResNet18 + fLoss model

表 8 不同姿态下黄羽鸡体质量估测结果比较

Tab. 8 Comparison of yellow-feathered chicken weight estimation results in different postures		
类别	均方根误差/kg	平均相对误差/%
站立	0.059	3.725
低头	0.135	8.311
展翅	0.123	6.810
梳理羽毛	0.145	8.906
坐姿	0.126	7.124
遮挡情况	0.144	9.144
全部数据集	0.128	7.481

4.5 不同姿态下黄羽鸡体质量估测结果

表 8 为不同姿态下黄羽鸡的体质量估测结果,可以发现当使用全部数据集时,均方根误差为 0.128 kg,平均相对误差为 7.481%,而经过 SE-ResNet18 + fLoss 模型对黄羽鸡姿态关键帧(站立姿

态)进行筛选后,均方根误差和平均相对误差均有明显降低,均方根误差为 0.059 kg,平均相对误差为 3.725%,且站立姿态相比于其他姿态的体质量估测精度更高,这也证明了 SE-ResNet18 + fLoss 模型对于鸡只姿态关键帧识别的有效性,能够最终提高体质量估测的精度。由于本文仅提取一维特征和二维特征进行体质量估测模型的建立,若在此基础上,融

合三维特征,将会进一步提高鸡只体质量估测的精度。

5 结束语

本研究提出一种用于平养黄羽鸡体质量估测的姿态关键帧识别模型 SE - ResNet18 + fLoss,模型融合了注意力机制 SE 模块和残差结构,并改进了损失函数,通过 Focal Loss 监督信号来解决样本不平衡问题,同时引入梯度加权类激活图对末端分类规则的合理性进行解释。模型在测试集上对于黄羽鸡的站立、低头、展翅、梳理羽毛、坐姿和遮挡 6 类姿态情况识别的 F1 值分别为 94.34%、91.98%、76.92%、93.75%、100% 和 93.68%;若仅识别是否

为鸡只姿态关键帧,识别精确率为 97.38%、召回率为 97.22%、F1 值为 97.26%、识别速度为 19.84 f/s,其识别精确率、召回率和 F1 值均优于 ResNet18、MobileNet V2 和 SE - ResNet18 网络。结果发现通过 Grad - CAM 可以很好地解释本文方法分类的依据,反映了每一类所关注的重点区域,其中,站立姿态重点关注于鸡头部位。其次,在连续视频上进行了模型的验证与鸡姿态特性的分析,发现白天有较多时刻黄羽鸡被识别为姿态关键帧,可用于后续体质量估测分析。进一步通过试验验证了 SE - ResNet18 + fLoss 模型对于黄羽鸡姿态关键帧识别的有效性,能够提高体质量估测的精度。

参 考 文 献

[1] NYALALA I, OKINDA C, CHEN K, et al. Weight and volume estimation of poultry and products based on computer vision systems: a review[J]. Poultry Science, 2021, 100(5): 101072.

[2] LOTT B D, REECE F N, MCNAUGHTON J L. An automated weighing system for use in poultry research[J]. Poultry Science, 1982, 61(2): 236 - 238.

[3] 沈明霞,刘龙申,闫丽,等. 畜禽养殖个体信息监测技术研究进展[J]. 农业机械学报, 2014, 45(10): 245 - 251. SHEN Mingxia, LIU Longshen, YAN Li, et al. Review of monitoring technology for animal individual in animal husbandry[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(10): 245 - 251. (in Chinese)

[4] CHEDAD A, AERTS J M, VRANKEN E, et al. Do heavy broiler chickens visit automatic weighing systems less than lighter birds? [J]. British Poultry Science, 2003, 44(5): 663 - 668.

[5] WET L D, VRANKEN E, CHEDAD A, et al. Computer-assisted image analysis to quantify daily growth rates of broiler chickens[J]. British Poultry Science, 2003, 44(4): 524 - 532.

[6] MOLLAH M, HASAN M A, SALAM M A, et al. Digital image analysis to estimate the live weight of broiler[J]. Computers and Electronics in Agriculture, 2010, 72(1): 48 - 52.

[7] MORTENSEN A K, LISOUSKI P, AHRENDT P. Weight prediction of broiler chickens using 3D computer vision[J]. Computers and Electronics in Agriculture, 2016, 123: 319 - 326.

[8] AMRAEI S, ABDANAN M S, SALARI S. Broiler weight estimation based on machine vision and artificial neural network[J]. British Poultry Science, 2017, 58(2): 200 - 205.

[9] 谢秋菊,周红,包军,等. 基于机器视觉的畜禽体质量评估研究进展[J]. 农业机械学报, 2022, 53(10): 1 - 15. XIE Qiuju, ZHOU Hong, BAO Jun, et al. Review on machine vision-based weight assessment for livestock and poultry[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(10): 1 - 15. (in Chinese)

[10] AMRAEI S, MEHDIZADEH S A, NS I D A. Development of a transfer function for weight prediction of live broiler chicken using machine vision[J]. Engenharia Agrícola, 2018, 38(5): 776 - 782.

[11] LEROY T, VRANKEN E, VAN BRECHT A, et al. A computer vision method for on-line behavioral quantification of individually caged poultry[J]. Transactions of the American Society of Agricultural and Biological Engineers, 2006, 49(3): 795 - 802.

[12] POURSABERI A, BAHR C, PLUK A, et al. Real-time automatic lameness detection based on back posture extraction in dairy cattle: shape analysis of cow with image processing techniques[J]. Computers and Electronics in Agriculture, 2020, 74(1): 110 - 119.

[13] 劳凤丹,滕光辉,李军,等. 机器视觉识别单只蛋鸡行为的方法[J]. 农业工程学报, 2012, 28(24): 157 - 163. LAO Fengdan, TENG Guanghui, LI Jun, et al. Behavior recognition method for individual laying hen based on computer vision [J]. Transactions of the CSAE, 2012, 28(24): 157 - 163. (in Chinese)

[14] 劳凤丹,杜晓冬,滕光辉. 基于深度图像的蛋鸡行为识别方法[J]. 农业机械学报, 2017, 48(1): 155 - 162. LAO Fengdan, DU Xiaodong, TENG Guanghui. Automatic recognition method of laying hen behaviors based on depth image processing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(1): 155 - 162. (in Chinese)

[15] 司永胜,安露露,刘刚,等. 基于 Kinect 相机的猪体理想姿态检测与体尺测量[J]. 农业机械学报, 2019, 50(1): 58 - 65. SI Yongsheng, AN Lulu, LIU Gang, et al. Ideal posture detection and body size measurement of pig based on Kinect[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(1): 58 - 65. (in Chinese)

[16] NASIRAHMADI A, STURM B, EDWARDS S, et al. Deep learning and machine vision approaches for posture detection of individual pigs[J]. Sensors, 2019, 19(17): 3738.

[17] RIEKERT M, KLEIN A, ADRIAN F, et al. Automatically detecting pig position and posture by 2D camera imaging and deep learning[J]. Computers and Electronics in Agriculture, 2020, 174: 105391.

[18] YANG X, ZHENG C, ZOU C, et al. A CNN-based posture change detection for lactating sow in untrimmed depth videos[J]. Computers and Electronics in Agriculture, 2021, 185: 106139.

[19] 宋晨波. 基于深度学习的母猪行为检测方法研究[D]. 杭州: 浙江大学, 2019.

[20] SONG Chenbo. Detection method for behaviors of sows based on deep learning[D]. Hangzhou: Zhejiang University, 2019. (in Chinese)

[21] 浙江大学. 一种采用单幅图像检测猪站立和躺卧姿态的方法: 201910060130.7[P]. 2020-12-11.

[22] LUO Y, ZENG Z, LU H, et al. Posture detection of individual pigs based on lightweight convolution neural networks and efficient channel-wise attention[J]. Sensors, 2021, 21(24): 8369.

[23] 薛月菊, 李诗梅, 郑婵, 等. 融合 2D-3D 卷积特征识别哺乳母猪姿态转换[J]. 农业工程学报, 2021, 37(9): 230-237.

[24] XUE Yueju, LI Shimei, ZHENG Chan, et al. Posture change recognition of lactating sow by using 2D-3D convolution feature fusion[J]. Transactions of the CSAE, 2021, 37(9): 230-237. (in Chinese)

[25] FANG C, ZHENG H, YANG J, et al. Study on poultry pose estimation based on multi-parts detection[J]. Animals, 2022, 12(10): 1322.

[26] LIN T Y, 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.

[27] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 618-626.

[28] HU J, LI S, GANG S. Squeeze-and-excitation networks[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.

[29] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.

[30] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal Loss for dense object detection[C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.

[31] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[C] // International Conference on Medical Image Computing and Computer-assisted Intervention, 2015: 234-241.

[32] BREIMAN L. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.

[33] 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.