

基于 ResNet-18 和数据增强的肉牛姿态识别实现 及其教学应用研究

孙帅¹, 赵永才¹, 李慢¹, 郜静茹¹, 关曼玉¹, 李广博^{1*}

(1.淮北理工学院电子与信息工程学院, 安徽 淮北 235000)

摘要: 肉牛姿态识别是智慧养殖的关键环节, 可为健康监测、行为分析和福利评估提供重要数据支撑, 对提升养殖效率与实现早期疾病预警具有重要意义。然而, 传统人工观察方式存在效率低、主观性强、难以持续监测等不足; 现有深度学习方法则普遍存在模型参数量大、部署成本高、复杂养殖场景下泛化能力有限等问题。针对上述问题, 本文提出一种基于 ResNet-18 的肉牛姿态识别方法, 采用迁移学习加载 ImageNet 预训练权重初始化模型, 结合随机水平翻转、随机旋转与颜色抖动等数据增强策略提升泛化能力, 并使用 SGD 优化器与余弦退火学习率调度完成 100 轮训练。在自建肉牛姿态数据集(含饮水、进食、打斗、舔舐、躺卧、爬跨、探寻、站立、行走共 9 类姿态, 每类 500 张, 总计 4500 张, 按 8:1:1 划分训练集、验证集与测试集)上的实验结果表明, ResNet-18 模型测试集识别准确率达 94.00%, 引入数据增强后进一步提升至 95.33%; 与 AlexNet (92.22%)、ResNet-50 (91.56%) 和 MobileNet-V2 (93.33%) 相比, ResNet-18 仅以 11.2M 参数量取得了具有竞争力的识别精度, 兼具轻量性与高效性。该方法为肉牛姿态的精准、高效识别提供了轻量级解决方案, 对推动智慧养殖中姿态识别技术的实际落地应用具有重要参考价值。

关键词: 肉牛姿态识别; 残差网络; 迁移学习; 数据增强; 智慧养殖

基金项目: 安徽省教育厅优秀青年教师培育项目(YQYB2026117); 安徽省智慧农业技术与装备重点实验室开放基金(AEC2026016; AEC2025009); 安徽省教育厅自然科学研究项目(重大)(2024AH040217); 省级质量工程-教学研究项目(2023jyxm1009; 2023xjzlt117; 2023sdxx145); 基于人工智能下的家禽康养与疾病预警技术研究(S202413620055); 基于自监督学习与对抗增强的小样本信号的智能感知研究(2025SK019); 基于深度学习的生猪识别与康养研究(S202513620083)

DOI: doi.org/10.65436/hssj.v1i9.50

肉牛养殖是我国畜牧业的重要组成部分, 精准掌握肉牛的日常行为状态对提升养殖管理水平具有重要意义。肉牛姿态识别是指通过计算机视觉技术自动判断肉牛当前所处的行为姿态, 如饮水、进食、躺卧、站立、行走等, 其识别结果可为健康监测、行为分析、福利评估及早期疾病预警提供关键数据支撑。传统的肉牛姿态监测主要依赖人工观察与记录, 该方式不仅费时费力、主观性强, 而且难以实现全天候、大规模的持续监测, 在集约化养殖场景下局限性尤为突出。近年来, 随着深度学习技术的快速发展, 基于卷积神经网络的图像识别方法为肉牛姿态识别提供了新的技术路径, 能够实现自动化、客观化、高精度的姿态分类, 有效弥补了传统人工方法的不足。

¹ **作者简介:** 孙 帅 (1997—), 男, 山东烟台人, 硕士, 助理实验师。研究方向: 计算机视觉;

赵永才 (2007—), 男, 安徽蚌埠人, 本科生。研究方向: 人工智能;

李 慢 (2005—), 女, 安徽池州人, 本科生。研究方向: 人工智能;

郜静茹 (2005—), 女, 安徽阜阳人, 本科生。研究方向: 人工智能;

关曼玉 (1998—), 女, 安徽淮北, 硕士, 助教。研究方向: 人工智能与大数据;

李广博 (1997—), 男, 安徽淮北人, 讲师, 博士。研究方向: 人工智能、智能信号感知、模式识别、计算机视觉等。

通讯作者: 李广博

目前, 基于深度学习的肉牛姿态识别研究主要采用卷积神经网络作为特征提取与分类的核心模型, 涉及 AlexNet、VGG、ResNet、DenseNet 及 EfficientNet 等多种网络架构。然而, 现有方法仍存在以下不足: 其一, 部分研究采用参数量较大的深层网络(如 ResNet-50、VGG-16), 虽然特征表达能力较强, 但模型计算成本高、内存占用大, 不利于在养殖现场的边缘设备上部署; 其二, 复杂养殖环境中存在光照变化、目标遮挡及相似姿态混淆等干扰因素, 部分模型在测试集上的泛化能力有待提升; 其三, 训练过程中数据增强策略的使用不够充分, 容易导致模型在有限样本上发生过拟合。针对上述问题, 本文选择 ResNet-18 作为骨干网络, 其残差连接结构可有效缓解深层网络的梯度消失问题, 同时参数量适中、推理效率较高, 适合构建轻量级的肉牛姿态识别模型。

基于此, 本文提出一种基于 ResNet-18 的肉牛姿态识别方法。该方法首先采用迁移学习策略, 加载在 ImageNet 数据集上预训练的权重对模型进行初始化, 以加速模型收敛并提升特征提取能力; 其次, 在训练阶段引入随机水平翻转、随机旋转($\pm 15^\circ$)和颜色抖动等数据增强策略, 扩充训练样本的多样性, 增强模型对光照变化和姿态差异的鲁棒性; 最后, 使用随机梯度下降(SGD)优化器结合余弦退火学习率调度策略, 在自建肉牛姿态数据集上进行 100 轮训练, 最终实现对 9 类肉牛姿态的精准分类。

本文的主要贡献如下: (1) 构建了一个包含饮水、进食、打斗、舔舐、躺卧、爬跨、探寻、站立、行走共 9 类姿态的肉牛姿态数据集, 每类选取 500 张图像, 总计 4500 张, 并按 8:1:1 的比例划分为训练集、验证集与测试集; (2) 提出一种基于 ResNet-18 与数据增强相结合的轻量级肉牛姿态识别方法, 在测试集上取得了 95.33% 的识别准确率; (3) 通过与 AlexNet、ResNet-50、MobileNet-V2 等模型的对比实验, 验证了 ResNet-18 在识别精度与模型效率之间的最优平衡, 其仅以 11.2M 参数量便取得了具有竞争力的识别性能; (4) 本文方法为智慧养殖场景下肉牛姿态识别技术的轻量化部署与实际应用提供了有益参考。

本文通过以下两方面进行展开:

第一方面为方法理论。首先给出本文肉牛姿态识别方法的总体框架, 阐述从图像输入到姿态类别输出的完整流程; 随后介绍实验所用自建肉牛姿态数据集的采集与预处理方法, 包括 9 类姿态的划分、样本规模、数据集划分方式及图像归一化操作; 在此基础上说明仅对训练集实施的数据增强策略, 以扩充样本多样性并避免数据泄露; 接着详细阐述 ResNet-18 网络结构, 包括残差模块的设计思想、主分支与捷径分支的计算过程, 以及全局平均池化和 Softmax 分类器的作用; 最后说明模型的前向传播与反向传播机制, 包括损失函数、优化器及参数更新过程。

第二方面为实验验证。首先给出实验环境与参数设置, 包括数据集划分、输入尺寸、训练轮次、批大小、优化器及学习率策略等; 其次在统一实验条件下, 将 ResNet-18 与 AlexNet、ResNet-50、MobileNet-V2 三种骨干网络进行横向对比, 从识别准确率、参数量等方面分析不同网络的性能差异; 最后通过训练准确率与损失曲线、混淆矩阵、t-SNE 特征可视化等手段, 对 ResNet-18 在有无数据增强条件下的实验结果进行可视化分析, 验证数据增强策略的有效性 & 本文方法在肉牛姿态识别任务中的可行性。

一、方法

(一) 总体框架

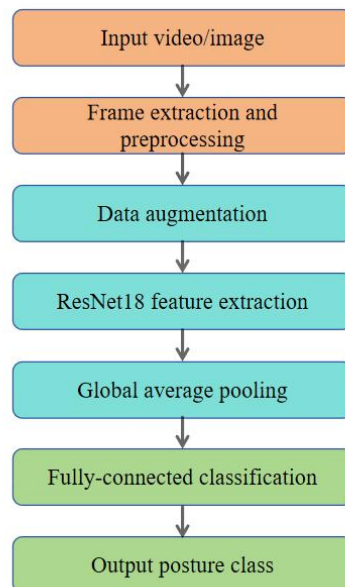


图 1: ResNet-18 的肉牛姿态识别整体框架

本文构建了基于 ResNet18 的肉牛姿态识别整体框架如图 1 所示, 整体流程严格遵循“数据预处理—特征提取器—分类器—损失约束”的深度学习识别范式, 完整链路为: 输入牧场监控视频与图像数据, 经视频抽帧、图像尺寸归一化等预处理操作统一样本维度, 仅对训练集样本执行多策略数据增强以丰富特征空间、规避数据泄露问题。预处理与增强后的图像送入 ResNet18 骨干网络完成深度特征提取, 该网络依靠残差学习机制解决深层模型梯度消失问题, 残差模块核心映射关系如公式所示:

$$H(x)=F(x)+x \quad (1)$$

式中 x 为残差模块输入特征, $F(x)$ 为网络学习的残差特征, $H(x)$ 为模块输出特征。当网络输入与输出特征维度不一致时, 通过捷径分支卷积进行维度匹配, 修正公式为:

$$H(x)=F(x)+W_s x \quad (2)$$

ResNet18 残差模块主分支通过双层 3×3 卷积完成特征变换, 特征变换过程可表示为: $F_l = \text{ReLU}(\text{BN}(\text{Conv}_{3 \times 3}(x)))$

、 $F(x) = \text{BN}(\text{Conv}_{3 \times 3}(F_l))$ 网络通过 ReLU 激活函数引入非线性特征表达, 其计算公式为:

$$\text{ReLU}(z)=\max(0,z) \quad (3)$$

为加速模型收敛、优化特征分布, 网络引入批归一化操作, 计算公式为:

$$z=\frac{z-\mu}{\sqrt{\sigma^2+\epsilon}}, \text{BN}(z)=\gamma z+\beta \quad (4)$$

图像经过多层残差网络完成深度特征提取后, 采用全局平均池化对特征图进行降维压缩, 池化运算公式为:

$$y_c=\frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F_{c,i,j} \quad (5)$$

降维后的一维特征向量输入全连接层进行特征映射, 最后通过 Softmax 函数完成类别概率归一化, 实现 9 类肉牛姿态分类输出, 其计算公式为:

$$p_k=\frac{e^{\tilde{z}_k}}{\sum_{i=1}^K e^{\tilde{z}_i}} \quad (6)$$

在模型训练阶段, 本文采用交叉熵损失函数约束模型优化方向, 计算预测值与真实标签的误差, 损失函数公式为:

$$L=-1N \sum_{i=1}^N \sum_{k=1}^K y_{i,k} \log(p_{i,k}) \quad (7)$$

模型通过反向传播不断迭代更新网络参数, 最小化损失误差, 最终实现对肉牛站立、行走、躺卧、采食、饮水等九类姿态的精准识别。该总体框架结构清晰、层级完整, 兼顾模型轻量化与识别精度, 能够有效适配智慧养殖场景下的肉牛姿态识别任务。

(二) 数据采集与预处理

本文采用真实牧场现场拍摄图像构建自建数据集, 共划分 9 类肉牛姿态: 饮水、采食、打斗、舔舐、躺卧、爬跨、搜寻、站立、行走, 每类样本约 500 张。数据集按 8:1:1 比例划分为训练集、验证集、测试集; 预处理阶段将图像统一缩放至 224×224 并完成归一化, 人工标注图像姿态标签, 用于模型训练。

(三) 数据增强策略

为扩充样本多样性并抑制模型过拟合, 本实验仅针对训练集实施数据增强, 验证集与测试集不做增强处理, 防止发生数据泄露。所使用的数据增强手段包括随机水平翻转、 $\pm 15^\circ$ 随机旋转、随机裁剪、亮度 / 对比度 / 饱和度调整、高斯噪声添加与随机擦除, 通过多样化图像变换丰富样本特征, 提升模型对肉牛姿态识别的泛化能力。

(四) ResNet18 网络结构

本文选用 ResNet18 作为肉牛姿态识别骨干网络, 网络结构如图 2 所示。肉牛姿态图像输入网络后, 先通过 Conv1 与最大池化进行浅层特征提取, 再依次经过 Layer1、Layer2、Layer3 残差模块提取深度特征。残差模块包含主分支与 shortcut 捷径分支: 主分支通过两组 3×3 卷积、BN 与 ReLU 激活完成特征变换; 捷径分支在

需要时利用 1×1 卷积调整特征维度，两路特征相加后经 ReLU 输出。特征提取完成后，采用全局平均池化压缩特征，最终通过 Softmax 层输出肉牛姿态类别的预测结果。

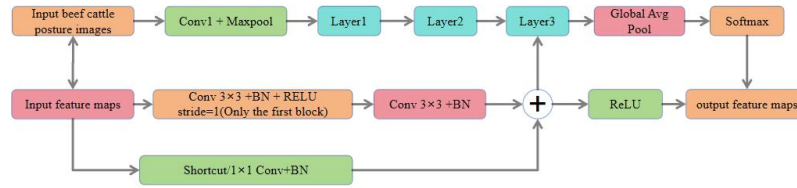


图 2: ResNet-18 骨干网络结构图

(五) 前向传播与反向传播

模型训练包含前向传播推理与反向传播参数更新两个阶段，整体流程如图 3 所示。在前向传播阶段，图像输入网络后，依次经过卷积层 Conv、批归一化 BN、ReLU 激活函数与最大池化层，再送入残差块模块完成深度特征提取；特征提取完成后经过全局平均池化 GAP、全连接层 FC，最后由 Softmax 分类器输出各类别的预测概率。完成前向推理后，采用交叉熵损失函数计算预测结果与真实标签之间的误差，随后进入反向传播阶段。反向传播过程依据损失梯度，利用优化器完成网络参数更新，本文实验选用 Adam 优化器，学习率设置为 0.001，批次大小 Batch size 取 32，总训练轮数设置为 100 轮；模型采用 ImageNet 预训练权重进行初始化，再针对肉牛姿态数据集开展微调训练。反向传播中参数更新公式为：

$$(W \leftarrow W - \eta \frac{\partial L}{\partial W}) \quad (8)$$

式中， W 代表网络权重参数， (η) 为学习率， $(\frac{\partial L}{\partial W})$ 为损失对权重的梯度，通过不断迭代最小化交叉熵损失，持续优化网络的特征提取与分类能力。

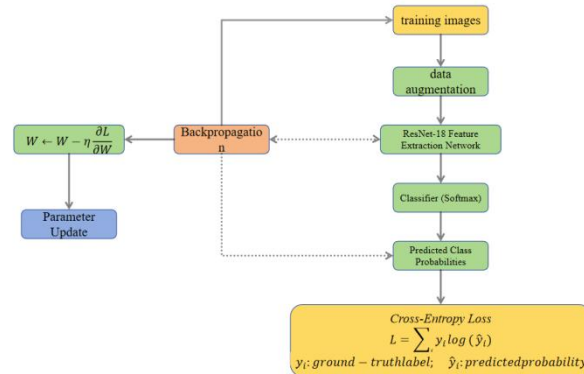


图 3: ResNet-18 的肉牛姿态识别前向与反向传播训练框架

二、实验

(一) 实验设置

本实验采用自建肉牛姿态数据集，涵盖饮水、采食、打斗、舔舐、躺卧、爬跨、搜寻、站立、行走共 9 类姿态，每类 500 张图像，按 8:1:1 比例划分为训练集、验证集与测试集。模型输入统一缩放至 224×224 像素，基于 PyTorch 2.5.1 框架实现，采用 ImageNet 预训练权重初始化并微调。训练共 100 轮，批大小为 32，优化器选用 SGD (momentum=0.9, weight_decay=1e-4)，初始学习率 0.001 并采用余弦退火策略调整，随机种子固定为 42，实验设置如表 1 所示。

表 1 实验设置的基本内容

项目	设置
数据集	自建肉牛姿态数据集
姿态类别	饮水、采食、打斗、舔舐、躺卧、爬跨、搜寻、站立、行走（共 9 类）
每类样本	500 张
训练集：验证集：测试集	8:1:1
输入尺寸	224×224

项目	设置
训练轮数	100
批大小 (Batch size)	32
优化器	SGD (momentum=0.9, weight_decay=1e-4)
学习率调度	余弦退火 (CosineAnnealing)
初始学习率	0.001
随机种子	42
深度学习框架	PyTorch 2.5.1 (CUDA 12.1)
预训练	ImageNet 预训练权重初始化并微调

(二) 比较方法及识别性能

本文对比 AlexNet、ResNet-18、ResNet-50、MobileNet-V2 四种骨干网络，实验统一设置为 100 轮训练、每类样本 500 个，数据集按 8:1:1 划分，仅使用基础预处理，结果见表 2、表 3。↑ 代表指标越大性能越好，↓ 代表指标越小性能越好。实验表明 ResNet-18 综合性能最优，其测试准确率 94.00%，高于 MobileNet-V2、AlexNet、ResNet-50；参数量仅 11.2M，远小于其余网络，推理稳定，便于在养殖边缘设备部署。实验验证了 ResNet-18 用于肉牛姿态识别任务的可行性与优越性。

表 2 无数据增强时不同骨干网络的肉牛姿态识别性能对比

	参数量 / M ↓	最佳验证准确率 / % ↑	最佳验证轮次	测试集准确率 / % ↑
AlexNet	57.0	94.67	55	92.22
MobileNet-V2	3.4M	95.33	12	93.33
ResNet-50	25.6	92.22	50	91.56
ResNet-18	11.2	94.67	59	94.00

表 3 有数据增强时不同骨干网络的肉牛姿态识别性能对比

骨干网络	参数量 / M ↓	最佳验证准确率 / % ↑	最佳验证轮次	测试集准确率 / % ↑
AlexNet	57.0	96.00	33	95.78
MobileNet-V2	3.5	96.00	36	94.33
ResNet-50	25.6	96.22	73	96.22
ResNet-18	11.2	96.89	56	95.33

(三) 训练曲线

图中展示了 ResNet-18 在有无数据增强两种条件下，100 轮训练过程中训练集与验证集的准确率和损失变化曲线（图 4、图 5）。由图 4 可见，在无数据增强条件下，训练集准确率迅速趋近于 100%，训练集损失快速下降至接近 0，而验证集准确率在 94.67% 附近波动、验证集损失在 0.25 附近趋于平稳，训练集与验证集之间存在较明显的性能差距，表明模型出现了一定程度的过拟合。引入数据增强后（图 5），训练集准确率的上升趋势更为平缓，训练集与验证集准确率之间的差距明显缩小，验证集最佳准确率提升至 96.89%，验证集损失进一步降低至 0.18 左右，训练过程更加稳定。上述结果表明，数据增强策略有效扩充了训练样本的特征空间分布，缓解了模型过拟合现象，提升了 ResNet-18 在肉牛姿态识别任务上的泛化能力。

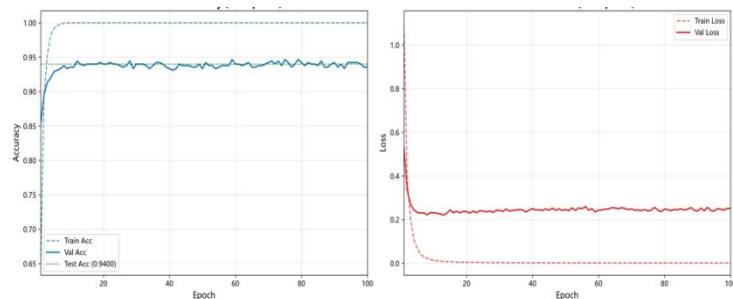


图 4：无数据增强条件下 ResNet-18 ACC、Loss 变化曲线

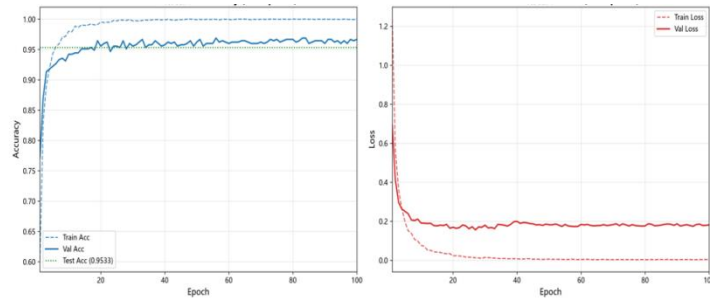


图 5：有数据增强条件下 ResNet-18 ACC、Loss 变化曲线

(四) 可视化分析

为直观分析数据增强对 ResNet-18 模型特征学习与分类性能的影响，本文对比有无数据增强条件下模型的 t-SNE 特征可视化结果（图 6）与混淆矩阵（图 7）。由图 6 的 t-SNE 可视化结果可见，引入数据增强后，各类肉牛姿态样本的特征聚类效果更加紧凑，不同类别之间的特征边界更加清晰，类别区分度得到提升。从图 7 混淆矩阵可以看出，站立与行走、躺卧等动作视觉相似度较高，属于易混淆姿态；在添加数据增强后，上述易混淆类别的误判样本数量明显减少，模型对肉牛姿态的识别稳定性进一步提升。

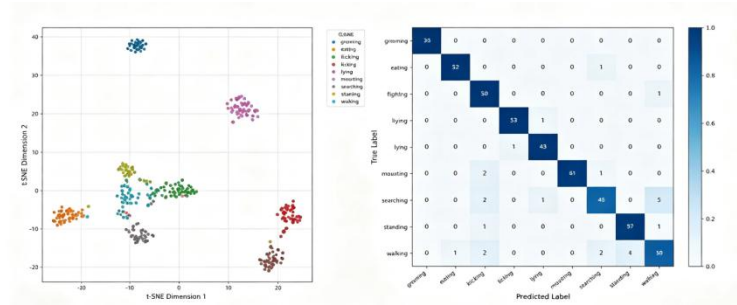


图 6：ResNet-18 无数据增强模型 t-SNE 特征可视化与混淆矩阵

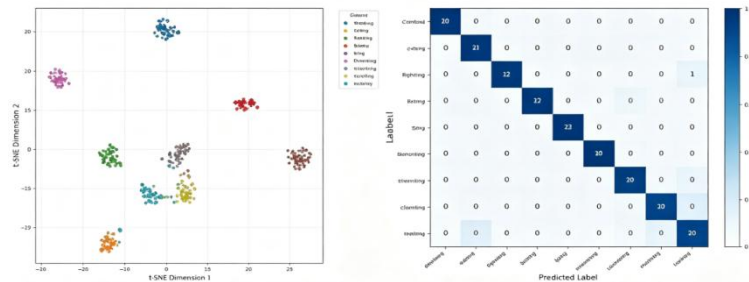


图 7：ResNet-18 有数据增强模型 t-SNE 特征可视化与混淆矩阵

三、结论

本文针对肉牛姿态识别中传统人工观察费时费力、现有深度学习方法模型参数量大且复杂场景下泛化能力不足的关键问题，本文提出了基于 ResNet-18 的肉牛姿态识别方法，结合迁移学习与数据增强策略，通过四种骨干网络的系统性对比实验和多维度可视化验证，得出以下主要结论：

1. ResNet-18 在肉牛姿态识别任务中展现出综合性能优势：相较于参数量庞大的 AlexNet（57.0M）和易发生过拟合的 ResNet-50（25.6M），以及网络结构较为简单的 MobileNet-V2，本文采用的 ResNet-18 通过残差连接结构有效缓解了深层网络训练中的梯度消失问题，仅以 11.2M 的参数量便实现了较强的多尺度特征提取能力，在识别精度与模型推理效率之间取得了最优平衡，为养殖现场边缘设备的轻量化部署提供了可行的网络架构选择。

2. 数据增强策略显著提升了模型的泛化能力与识别精度：实验结果表明，在引入随机水平翻转、随机旋转（ $\pm 15^\circ$ ）和颜色抖动等数据增强策略后，四种骨干网络的测试集准确率均得到不同程度的提升，其中 ResNet-50 提升最为显著（由 91.56% 提升至 96.22%，提升幅度达 4.66%），AlexNet 提升 3.56%，ResNet-18 提升 1.33%，MobileNet-V2 提升 1.00%。混淆矩阵、t-SNE 特征可视化、训练准确率与损失曲线等多维度结果一致证实，数据增强有效扩充了训练样本的特征空间分布，缓解了小样本条件下的过拟合现象，增强了模型对养殖现场光照波动、拍摄角度差异及肉牛个体形态变化的鲁棒性。

3. 为智慧养殖场场景下肉牛姿态识别提供了轻量级、高精度的有效解决方案：本文方法的核心价值在于以较小的模型参数量实现了较高的识别精度，ResNet-18 结合数据增强后测试集准确率达 95.33%，能够满足集约化养殖现场对姿态识别精度与实时性的双重需求，为肉牛健康监测、日常行为分析和早期疾病预警提供了坚实的技术支撑，在推动智慧养殖技术落地应用方面具有重要的实际意义和应用潜力。

展望未来，本肉牛姿态识别方法在复杂养殖环境下的识别鲁棒性、细粒度姿态类别体系扩展及边缘设备实时推理优化等方面仍有提升空间，后续将围绕多目标检测与跟踪、异常行为类别扩展、模型轻量化技术三个方向展开研究，推动该技术在养殖现场的实际部署与应用。

参考文献：

- [1] LI G B, LIU W X. Overcoming data scarcity: few-shot pig vocalization recognition via domain expansion, knowledge transfer, and feature alignment[J]. *Animals*, 2026, 16(13): 2074.
- [2] LI G B, SHI G L, ZHU C J. Dynamic serpentine convolution with attention mechanism enhancement for beef cattle behavior recognition[J]. *Animals*, 2024, 14(3): 466.
- [3] GAO S, YANG J F, XU D L. Beef cattle species and behavior recognition based on improved YOLO v10[J]. *Journal of South China Agricultural University*, 2025, 46(6): 832-842.
- [4] NI P F, ZHOU S Y, YE J H, et al. Pig behavior recognition method based on feature point detection[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2025, 41(6): 173-184.
- [5] LI G B, SUN J Y, GUAN M Y, et al. A new method for non-destructive identification and tracking of multi-object behaviors in beef cattle based on deep learning[J]. *Animals*, 2024, 17: 2464.
- [6] 付辰伏, 任力生, 王芳. 自动化场景区分下 FABF-YOLOv8s 轻量化肉牛行为识别方法[J]. *农业工程学报*, 2024, 40(15): 152-163.

Beef Cattle Posture Recognition Based on ResNet-18 and Data Augmentation: Implementation and Teaching Application

SUN Shuai, ZHAO Yongcai, LI Man, GAO Jingru, GUAN Manyu, LI Guangbo

(College of Electronic and Information Engineering, Huaibei Institute of Technology, Huaibei 235000, China)

Abstract: Beef cattle posture recognition is a critical component of intelligent livestock farming. It provides essential data support for health monitoring, behavior analysis and welfare assessment, and is of great significance for improving breeding efficiency and enabling early disease warning. However, traditional manual observation methods suffer from shortcomings such as low efficiency, strong subjectivity and difficulty in continuous monitoring. Existing deep learning methods generally face problems including large model parameters, high deployment costs and limited generalization ability in complex breeding scenarios. To address the above issues, this paper proposes a beef cattle posture recognition method based on ResNet-18. Transfer learning is adopted to initialize the model with ImageNet pre-trained weights. Data augmentation strategies including random horizontal flipping, random rotation and color jitter are combined to improve generalization performance, and the model is trained for 100 epochs using the SGD optimizer with cosine annealing learning rate scheduling. Experiments are carried out on a self-built beef cattle posture dataset, which contains 9 posture categories: drinking, feeding, fighting, licking, lying, mounting, exploring, standing and walking. There are 500 images per category with a total of 4500 images, and the dataset is divided into training set, validation set and test set at a ratio of 8:1:1. The results show that the ResNet-18 model achieves a test recognition accuracy of 94.00%, which is further improved to 95.33% with the introduction of data augmentation. Compared with AlexNet (92.22%), ResNet-50 (91.56%) and MobileNet-V2 (93.33%), ResNet-18 delivers competitive recognition accuracy with only 11.2M parameters, featuring both light weight and high efficiency. This method provides a lightweight solution for accurate and efficient beef cattle posture recognition, and holds important reference value for promoting the practical application of posture recognition technology in intelligent livestock farming.

Keywords: Beef cattle pose recognition; Residual network; Transfer learning; Data augmentation; Smart livestock farming