doi:10.6041/j.issn.1000-1298.2026.23.032

基于多尺度特征和自适应混合多注意力机制的猪只日常行为识别方法

孙立博　沈国栋　杨雨锟　秦文虎
（东南大学仪器科学与工程学院，南京 210096）

摘要：猪只日常行为识别是智慧化养殖技术体系的核心支撑，对精准监测猪群健康状态与生产性能具有重要意义。针对现有识别模型在精度与效率间难以平衡的关键问题，本文提出一种基于多尺度特征和自适应混合多注意力机制的猪只日常行为识别模型 AHMA－PBRNet。该模型首先采用改进的 FasterNet 作为骨干网络，在保留强特征提取能力的基础上显著提升计算效率；其次设计频域注意力模块（FAM）与自适应空间－通道混合注意力机制（ASCAM），通过抑制图像噪声干扰与动态聚焦不同层次特征关键区域，实现对猪只行为本质特征的精准捕捉，有效强化模型学习能力；最后引入加权组合损失函数进一步优化检测精度。基于自主构建的高质量猪只行为数据集开展实验验证，结果表明：AHMA－PBRNet 对站立、坐立、趴卧、侧卧、采食、饮水和爬跨 7 种猪只日常行为的识别精度分别达到 97.3%、91.8%、97.0%、88.8%、98.6%、92.9% 和 99.5%，整体识别精确率、召回率和平均精度均值分别达到 92.7%、90.4% 和 95.1%，而模型参数量仅为 4.07×10^6，浮点运算量为 1.11×10^{10}，综合性能优于当前主流模型。此外，基于该模型搭建的群养生猪行为识别交互平台，显著提升了识别过程的直观性与操作便捷性，为现代化智慧养猪产业发展提供了有力技术支撑。

关键词：猪只；日常行为识别；多尺度特征；自适应混合多注意力；目标检测；行为识别交互平台
中图分类号：TP391　　**文献标识码**：A　　**文章编号**：1000-1298(2026)23-0338-11　　**OSID**：

Pig Behavior Recognition Based on Multi-scale Features and Adaptive Hybrid Multi-attention Mechanism

Sun Libo　Shen Guodong　Yang Yukun　Qin Wenhu
(*School of Instrument Science and Engineering*, *Southeast University*, *Nanjing* 210096, *China*)

Abstract: Pig daily behavior recognition is a core support of the intelligent breeding technology system, which is of great significance for accurately monitoring the health status and production performance of pig herds. Aiming at the key problem that existing recognition models are difficult to balance accuracy and efficiency, a pig daily behavior recognition model AHMA－PBRNet was proposed based on multi-scale features and adaptive hybrid multi-attention. Firstly, the model adopted the improved FasterNet as the backbone network, which significantly improved the computational efficiency while retaining strong feature extraction capability. Secondly, it innovatively designed the frequency-domain attention module (FAM) and adaptive spatial－channel mixed attention mechanism (ASCAM). By suppressing image noise interference and dynamically focusing on key regions of features at different levels, it realized the accurate capture of the essential features of pig behaviors and effectively strengthened the learning ability of the model. Finally, a weighted combination loss function was introduced to further optimize the detection accuracy. Experimental verification based on the self-constructed high-quality pig behavior dataset showed that: AHMA－PBRNet achieved recognition accuracies of 97.3%, 91.8%, 97.0%, 88.8%, 98.6%, 92.9% and 99.5% for seven kinds of pig daily behaviors, including standing, sitting, lying prone, lying on side, eating, drinking and mounting respectively. The overall recognition accuracy, recall rate and mean average precision reached 92.7%, 90.4% and 95.1% respectively,

收稿日期：2025－08－08　修回日期：2025－11－18
基金项目：江苏现代农业产业单项技术研发项目(CX(23)3120)
作者简介：孙立博(1983—)，女，副教授，博士，主要从事计算机视觉与人工智能研究，E-mail：sunlibo@seu.edu.cn
通信作者：秦文虎(1969—)，男，教授，博士，主要从事农业物联网感知技术研究，E-mail：qinwenhu@seu.edu.cn

while the model parameter quantity was only 4.07×10^6 and the computational quantity was 1.11×10^{10}, with comprehensive performance superior to current mainstream models. In addition, the group-housed pig behavior recognition interactive platform built based on this model significantly improved the intuitiveness and operational convenience of the recognition process, providing strong technical support for the development of modern intelligent pig breeding industry.

Key words: pig; daily behavior recognition; multi-scale features; adaptive hybrid multi-attention; object detection; behavior recognition interactive platform

0　引言

随着智慧农业的普及与快速发展,畜牧业的智能化管理已日益成为降低运营成本,提高生产效率、产品价值的重要手段之一[1]。在此背景下,规模化养猪场对猪只日常行为的自动化监测需求日益增长,对猪群的日常行为监测不仅能够提前预防疾病,还可协助养猪场进行日常饲养管理,发现猪只异常行为并做出应对[2-3]。因此,猪只日常行为识别的研究对于猪只的智慧化养殖具有重要意义。

传统的基于传感器的猪只行为识别方法[4-6]复杂费时,需要依靠人工佩戴传感器,易造成猪只损伤和应激,难以满足规模化养猪场高效智能的需求。基于传统计算机视觉技术的猪只行为识别方法[7-9]实时性与自动化程度低,多依赖人工干预调整参数,不利于规模化养猪场对行为识别高效准确的要求。而基于深度学习的方法[10-11]可实现猪只行为识别的自动化、无接触化、高效化和准确化,十分契合当下智慧养殖的要求。

目前基于深度学习的猪只行为识别方法大致可分为两阶段和单阶段这两类。两阶段方法[12-13]存在计算复杂度高、实时性差等问题;单阶段方法[14-17],或者对数据占比小的行为类别识别精度偏低,或者模型的参数量和运算量可能无法满足实时性的要求,或者算法研究的猪只日常行为种类较少。

尽管现有的深度学习模型在猪只行为识别中取得了一定的成果,但往往无法实现识别精度和效率的平衡。一方面由于猪只好堆积、易遮挡和日常行为特征相似且多样,从而使得模型识别困难;另一方面由于在实际养殖场景中实时性是猪只行为识别发挥价值的核心前提,这要求模型具有较低的参数量和计算量。针对上述问题,本研究提出一种基于多尺度特征和自适应混合多注意力的猪只日常行为识别方法 AHMA-PBRNet。该方法首先基于改进的 FasterNet 骨干网络[18],在保证特征提取能力的前提下,提升计算速度和降低延迟。接着设计频域注意力模块(Frequency domain attention module,FAM)和自适应空间-通道混合注意力模块(Adaptive spatial-channel mixed attention module,ASCAM),在抑制图像噪声的基础上根据不同层次的特征自适应强化空间局部区域和特定特征通道,从而从大量视觉信息中精准捕捉猪只日常行为的本质特征,提高模型的学习能力。最后,使用传统检测方法的检测头搭建特征提取-特征处理-检测输出的猪只日常行为识别框架,实现猪只日常行为识别精度和效率的平衡提升。

1　数据集

1.1　数据来源

在生猪的集约化养殖过程中,养殖人员需要依靠对猪只日常行为的观察来掌握其健康状态。为此,对猪的 7 种典型日常行为进行了研究,包括站立、坐立、趴卧、侧卧、采食、饮水和爬跨。这些行为不仅反映猪只的正常生理需求与行为节律,也是健康、应激和环境适应状态的重要表征,为智慧养猪提供了基础依据[19]。上述行为的识别描述如表 1 所示。行为示例如图 1 所示。

表 1　猪只典型日常行为描述

Tab. 1　Descriptions of typical daily behaviors in pigs

日常行为	行为描述	标签
站立	猪只身体直立,腿部伸直,仅四腿接触地面	stand
坐立	猪只前腿伸直,尾部或臀部接触地面,部分身体立起	sit
趴卧	猪只四肢和腹部同时接触地面,四肢呈弯曲姿态,头部向前伸展	lie
侧卧	猪只以躯干一侧为支撑,躯干一侧完全接触地面,四肢伸直	lateral
采食	猪只走到固定食槽,颈部向前伸展,头部伸到食槽中	eat
饮水	猪只走到饮水管,颈部向前伸展,前腿紧靠饮水区域	drink
爬跨	爬跨猪只的前脚攀登上被爬猪只的背部	mount

由于缺乏包含所有 7 种猪只行为类型的综合公共数据集,本研究构建了一个包含多种行为的更全面的数据集,该数据集旨在满足猪只日常行为自动识别研究的需求。使用海康威视 DS-IPC-T13HV3-IA 红外阵列海螺型网络摄像机和 DS-2CD3T66(D)WDV3-I3 筒型网口摄像机安装在江



图 1 猪只日常行为示例

Fig. 1 Examples of pigs' daily behaviors

苏省镇江市丹徒上党绿洲生猪饲养场的 2 个猪栏上。第 1 个摄像机安装在离地高度 2.5 m 处,以 30°的顶角拍摄猪舍内猪只行为视频;第 2 个摄像机安装在离地高度 2.5 m 处,以 90°的顶角拍摄猪舍内猪的行为视频。这 2 个角度的设置一方面是尽可能地避免猪只相互遮挡,另一方面是丰富猪只日常行为识别数据集的采集角度,以增强数据的多样性。为了使数据集更具有代表性和多样性,还分别构建了高密度养殖环境和夜间环境下猪只日常行为识别数据集。由于非接触式数据采集的优势,用于采集数据的猪只均未受到损伤或影响。

标准养殖环境下每个猪栏里有 6 头猪,高密度养殖环境下每个猪栏中有 15 ~ 25 头猪,并配有饲料槽和饮水管。摄像时间是每天 08:00—18:00 中不同的 4 个时间段。夜间环境下摄像时间是每天 21:00 至次日 02:00 中不同的 2 个时间段。摄像机可以不间断地录制视频并直接导出视频数据。猪场猪栏面积为 5.0 m × 4.0 m,平均每头猪的活动面积为 3.33 m^2。每个猪栏配有一个人工供应饲料槽,每个饲料槽之间有 3 个猪只屏障,可同时容纳 4 头猪采食。此外每个猪栏配有 3 个乳头饮水器,猪只可自由取水。猪栏结构如图 2 所示。

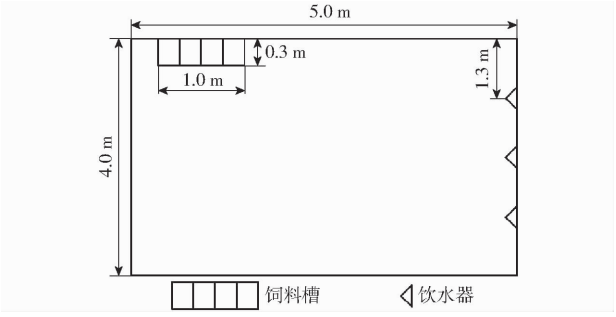


图 2 猪栏结构示意图

Fig. 2 Structure of pig stalls

1.2 数据处理

数据预处理包括图像采集、筛选与滤波、标注及数据增强 4 个过程。

(1)图像采集:使用 2 个在线摄像设备,在群养猪舍环境下获取不同时间段下猪栏中的猪只活动数据视频,并利用 OpenCV 计算机视觉库截取所采集视频中每 10 s 的 jpg 格式图像。

(2)筛选与滤波:在筛除无法辨别猪只日常行为的图像后,最终选取标准养殖环境下 977 幅图像,高密度养殖环境和夜间环境下各 100 幅图像。此外还使用滤波器减少图像中的噪声点,以减少对模型训练的干扰。

(3)标注:使用 LabelImg 软件手动标注图像中的猪只日常行为(https://github.com/tzutalin/labelImg)。按照 COCO 数据集格式生成标准 txt 格式的标注文件。标注完成后,将标准养殖环境、高度密集养殖环境和夜间环境 3 种数据集图像和对应的标注文件分别按比例 8∶1∶1 分为训练集、验证集和测试集。每幅图像中包含多个猪只日常行为实例,该数据集中每个行为的实例数量如表 2 所示。

表 2 猪只日常行为标注数据

Tab. 2 Annotated data of pigs' daily behaviors

日常行为标签	实例数量
站立	2 407
坐立	306
趴卧	1 681
侧卧	804
采食	1 287
饮水	479
爬跨	141

(4)数据增强:数据增强技术可以增加样本的多样性并提高模型的鲁棒性。在本研究中,通过 5 种方法来增强数据,包括图像拼接、镜像、裁剪、随机旋转和 HSV 色调增强,使得用于模型训练数据集的规模扩大,有助于加快模型收敛速度,提升模型的泛化能力。

2 研究方法

2.1 AHMA - PBRNet 网络框架

图 3 所示为本研究提出的 AHMA - PBRNet 端到端的猪只日常行为识别模型结构,图中 H、W 分别表示输入图像的高度和长度,C_1、C_2、C_3、C_4 表示各个阶段的通道数。该模型基于目标检测的经典编解码架构,利用 FasterNet 骨干改进网络系统性地提取猪只图像的多尺度特征和使用自适应空间-通道-频域多注意力混合机制融合和增强特征,显著提升了猪只行为的边界定位能力与精准识别能力。具体而言,编码阶段采用 FasterNet 骨干改进网络来提取猪只图像的多尺度特征;特征融合阶段,一方面设计频域注意力模块对特征图实施快速傅里叶变换与高频

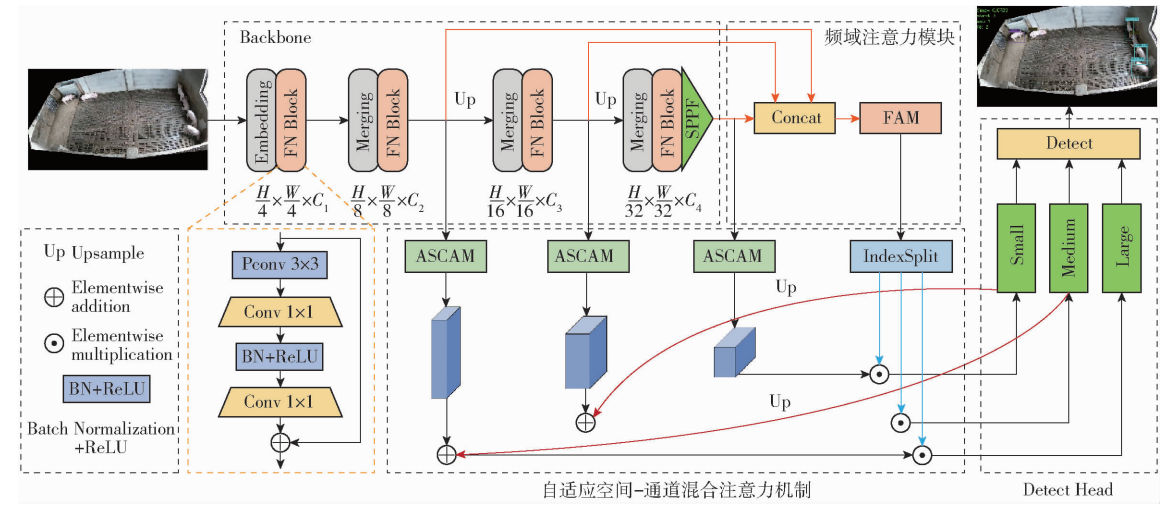


图 3　AHMA－PBRNet 网络模型结构

Fig. 3　Structure of AHMA－PBRNet network model

分量强化，有效地解决了猪只图像中噪声干扰问题。另一方面，为了充分利用图像的浅层和深层特征，设计了自适应空间-通道混合注意力模块，创新地基于图像特征的层次自适应地分配空间注意力和通道注意力的权重，以增强不同层次的图像特征中与自身相关性较大的要素，进而准确捕捉猪只不同日常行为的差异性本质特征；检测输出阶段引入了经典的大、中、小 3 段检测头以满足不同体态的猪只日常行为识别的要求。

AHMA－PBRNet 通过 FasterNet 骨干改进网络、自适应空间-通道-频域多注意力混合机制和目标检测头搭建特征提取-特征处理-检测输出的猪只日常行为识别框架。该模型的骨干网络在利用部分卷积模块（PConv）快速且准确地提取猪只图像行为特征的基础上引入快速的空间金字塔池化模块（SPPF），使得模型可以解决猪只尺度差异问题，从而增强模型对不同尺寸目标的适应性。自适应空间-通道-频域多注意力混合机制分为两部分，一部分是频域注意力模块（FAM），另一部分是自适应空间-通道混合注意力模块（ASCAM）。频域注意力模块从频域角度入手，针对猪只图像中的噪声问题，通过频域特征的引入与处理，有效抑制了噪声和增强了目标特征。自适应空间-通道混合注意力模块针对不同层次猪只图像特征的重点关注信息不同的特点，基于通道信息和空间信息的差异性，利用自适应空间-通道混合注意力机制自适应强化空间局部区域和特定特征通道，准确捕捉猪只日常行为的本质特征，进而提升整体模型的学习能力。

2.2　FasterNet 骨干改进网络

在猪只养殖实际环境中，猪只日常行为识别模型需要具备快速且准确的特点。一方面由于猪只行为具有动态性、时效性，模型必须快速处理并输出结果，否则会失去实际猪只日常行为识别的价值；另一方面，在猪只养殖过程中，准确的猪只日常行为识别可以真正帮助解决猪只养殖管理中的问题，如健康预警、异常干预等。然而当前很多基于深度学习的行为识别模型，在识别的快速性和准确性上很难兼顾。

为了有效应对上述问题，本研究引入 FasterNet 骨干改进网络。如图 3 中 Backbone 所示，该网络主要是由嵌入层（Embedding）、合并层（Merging）、FN Block 和 SPPF 组成。其中的关键特征提取模块 FN Block 主要由 PConv 构建，PConv 通过仅在部分输入通道上应用卷积，同时保持其他通道不变，来减少不必要的计算。由于特征图不同通道间存在相似性，这种方式可在不影响整体特征提取的前提下，降低计算量，使主干网络可在各种设备上更快运行。嵌入层和合并层分别由步长为 4 的基础卷积块（Conv 4×4）和步长为 2 的基础卷积块（Conv 2×2）组成，用于空间下采样和通道数的扩展。此外，通过引入 SPPF 使基础的 FasterNet 骨干网络适应不同尺寸的猪只目标。

AHMA－PBRNet 的 Backbone 首先利用 Embedding 采用基础卷积实现猪只图像的下采样。随后利用 FN Block 实现不同层次特征的提取，Merging 在每一层次的特征提取中发挥了通道数的扩展作用。最后利用 SPPF 实现不同尺度的特征信息的整合，提高模型对不同尺寸目标的识别能力。通过上述过程，FasterNet 骨干改进网络为 AHMA－PBRNet 提供快速且强大的猪只日常行为的特征提取能力，提高 AHMA－PBRNet 在猪只日常行为识别中的实时性和准确性。

2.3　频域注意力模块

为了解决猪只日常养殖图像所存在的噪声问

题,设计了 FAM 模块,从频域角度抑制图像噪声,从而实现特征增强和全局优化。在频域内,不同的频率分量对应于图像中的不同特征。通过对频域的分析,FAM 模块可以有效地区分特征图中的目标和噪声,从而使模型更准确地聚焦目标特征。

FAM 网络框架如图 4 所示。模块内部首先对输入的特征图使用傅里叶变换,接着使用二维高斯滤波器,用于频率域的平滑处理。随后计算注意力权重并应用到滤波后的特征,其具体表达式为

$$x = \mathrm{FFT}(R) \tag{1}$$

$$w = \mathrm{S}(\mathrm{GAP}(\mathrm{G}(x))) \tag{2}$$

式中 R——FasterNet 骨干改进网络输出特征的拼接特征

FFT——傅里叶变换

x——经傅里叶变换后的特征

G——二维高斯滤波器

GAP——全局平均池化

S——Sigmoid 激活函数

w——注意力权重

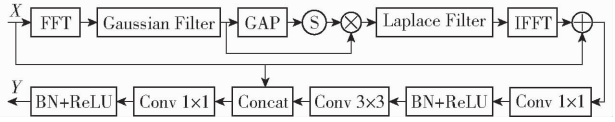


图 4 FAM 的网络框架

Fig. 4 Network framework of FAM

再利用拉普拉斯高通滤波器增强特征频率域的边缘,最后通过逆傅里叶变换将特征转换到空间域,并通过多个卷积层和 ReLU 激活函数输出最终结果,进而从频域角度抑制猪只图像噪声,实现特征增强和全局优化。其具体公式为

$$T = \mathrm{IFFT}(\mathrm{Laplace}(w \otimes x)) \oplus R \tag{3}$$

$$y = \mathrm{Concat}(\mathrm{Conv3}(\mathrm{ReLU}(\mathrm{BN}(\mathrm{Conv1}(T)))), R) \tag{4}$$

$$Y = \mathrm{Conv1}(\mathrm{BN}(\mathrm{ReLU}(y))) \tag{5}$$

式中 Laplace——拉普拉斯高通滤波器

IFFT——逆傅里叶变换

T——经过变换且转换到空间域的特征

Conv1——1 × 1 卷积

ReLU——ReLU 激活函数

BN——批归一化函数

Conv3——3 × 3 卷积

Concat——拼接操作

y——FAM 模块的前置输出特征

Y——FAM 模块的输出特征

2.4 自适应空间-通道混合注意力模块

为了解决猪只日常行为目标特征相似和关键行为特征分散所导致的日常行为识别精度低的问题,设计了自适应空间-通道混合注意力模块(ASCAM)。该模块从空间和通道 2 个维度动态地聚焦不同层次的有效特征,从而实现关键信息强化和冗余信息抑制。在图像特征中,经由骨干网络提取的不同层次特征图关注猪只行为的不同关键信息。通过对不同层次特征图中空间和通道特征的动态调整,ASCAM 可以有效地区分特征图中的行为与环境干扰以及猪只不同日常行为的特征,从而使模型更精准地捕捉猪只不同日常行为所对应的核心特征。

ASCAM 网络框架如图 5 所示。经由骨干网络提取的特征图分别输入到空间注意力机制 SAM 和通道注意力机制 CAM。SAM 和 CAM 可表示为

$$S_i = X_i \otimes \mathrm{S}(\mathrm{Conv7}(\mathrm{GMP}(X_i))) \tag{6}$$

$$C_i = X_i \otimes \mathrm{S}(\mathrm{FC2}(\mathrm{ReLU}(\mathrm{FC1}(\mathrm{GMP}(X_i))))) \tag{7}$$

式中 X_i——骨干网络不同层级的输出

GMP——全局最大池化

Conv7——7 × 7 卷积

FC1、FC2——全连接

S_i、C_i——SAM 和 CAM 的输出特征

随后通过可学习的权重矩阵 $\boldsymbol{\omega}$ 和 $\boldsymbol{I}-\boldsymbol{\omega}$ 分别点乘 SAM 和 CAM 的输出以实现从空间和通道 2 个维度动态地聚焦不同层次的有效特征。最后通过卷积

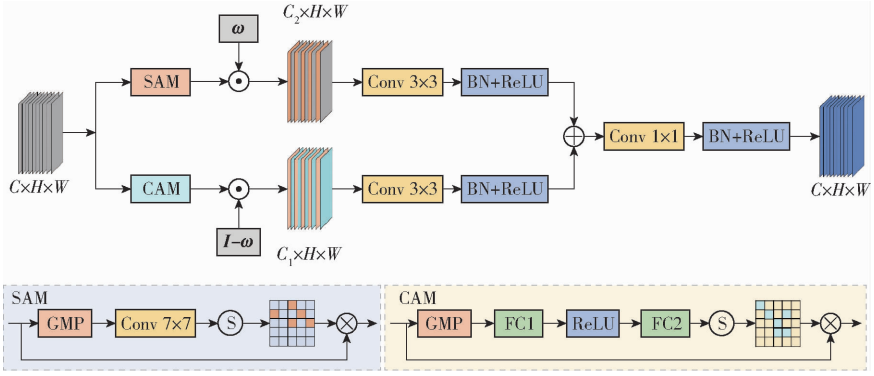


图 5 ASCAM 的网络框架

Fig. 5 Network framework of ASCAM

层、批归一化和 ReLU 激活函数实现 2 个特征图的加和，进而得到 ASCAM 的输出。该过程可表示为

$$S_i' = \mathrm{ReLU}(\mathrm{BN}(\mathrm{Conv3}(\boldsymbol{\omega} \odot S_i))) \quad (8)$$

$$C_i' = \mathrm{ReLU}(\mathrm{BN}(\mathrm{Conv3}((\boldsymbol{I} - \boldsymbol{\omega}) \odot C_i))) \quad (9)$$

$$Z_i = \mathrm{ReLU}(\mathrm{BN}(\mathrm{Conv1}(S_i' \oplus C_i'))) \quad (10)$$

式中　$\boldsymbol{I}$——全一矩阵

S_i'——空间维度的输出特征

C_i'——通道维度的输出特征

Z_i——ASCAM 的输出特征

2.5　损失函数

在猪只行为识别中，损失函数直接影响模型对猪只目标边界框、行为类别信息的学习效果，其设计需兼顾定位精度、分类准确性及训练稳定性。针对本研究提出的猪只行为识别网络 AHMA－PBRNet，采用边界框损失 L_{box}、置信度损失 L_{obj} 和分类损失 L_{cls} 的加权组合损失。总损失函数 L_{total} 定义为

$$L_{total} = \lambda_{box} L_{box} + \lambda_{obj} L_{obj} + \lambda_{cls} L_{cls} \quad (11)$$

式中　λ_{box}——边界框损失权重，常取 7.5

λ_{obj}——置信度损失权重，常取 10

λ_{cls}——分类损失权重，常取 0.5

3　实验与结果分析

3.1　实验环境

本文将具有 24 GB 显存的 NVIDIA GeForce RTX 4090 显卡用于实验。Python 包管理采用 Anaconda，并使用 PyTorch 深度学习框架。所有实验均在 64 位 Ubuntu 20.04.6 LTS 系统上进行，配备 CUDA 12.1、Python 3.8 和 PyTorch 2.2.2。合理地选择超参数是模型训练过程的关键，本研究实验中使用的超参数设置详细信息如表 3 所示。

表 3　实验中使用的超参数

Tab.3　Hyperparameters used in experiments

超参数	数值/类型
优化器	SGD
学习率	0.01
学习率动量	0.937
权重衰减数	0.000 5
学习率预热轮数	3
预热动量	0.8
训练图像尺寸($H \times W \times C$)/(像素×像素×通道)	640×640×3
批大小	64
训练轮次	300

3.2　评价指标

本研究采用精确率(Precision，P)、召回率(Recall，R)、平均精确率均值(Mean average precision，mAP)、参数量和浮点运算量(FLOPs)作为评估指标来评估模型性能。

精确率越高，模型的错误预测就越少，能直接减轻因判断失误给农场带来的经济负担；召回率越高，对猪只日常行为的识别越全面，可有效控制因遗漏识别而产生的各类损失。更高的 mAP 意味着模型对猪只各类行为的识别精度和定位准确性更优，能减少因目标误判或定位偏差导致的决策失误，降低农场在行为管理上的额外成本。更合理的参数规模表明模型在保证性能的同时避免了冗余计算，节省硬件资源消耗的同时提高运行效率，为农场实时监测提供稳定支撑。更低的浮点运算量代表模型计算量更小，可在有限设备算力下快速响应，减少因处理延迟造成的信息滞后，降低因未能及时干预产生的潜在损失。这 5 个指标相互配合，从不同维度证明 AHMA－PBRNet 中各关键组件对猪只日常行为识别性能的积极影响。

3.3　AHMA－PBRNet 消融实验分析

本研究通过消融实验证实了 AHMA－PBRNet 中各关键组件对猪只日常行为识别性能的影响。首先为了证明 AHMA－PBRNet 网络中 FasterNet 骨干改进网络在猪只日常行为识别中的实时性和准确性，与传统或者经典的骨干网络进行了对比实验。接着基于 FasterNet 骨干改进网络开展了各关键组件的消融实验，以证明 FAM 和 ASCAM 的有效性。

骨干网络的对比实验结果如表 4 所示。传统骨干网络 ResNet50[20] 与 ResNet101[20] 尽管能达到较高的 mAP50，但在参数量与浮点运算量方面均较 FasterNet 骨干改进网络高出数十倍，这对猪只行为识别的实时性构成显著制约。当前常用的 YOLO v8[21] 骨干网络虽参数量与浮点运算量少于 FasterNet 骨干改进网络，但其 mAP50 表现偏低，难以满足猪只行为识别对精确性的要求。综合来看，选择 FasterNet 骨干改进网络作为 AHMA－PBRNet 的骨干网络，能够在猪只行为识别的实时性与精确性之间取得最佳平衡，是最优选择。

表 4　骨干网络的对比实验

Tab.4　Comparative experiments on backbone

骨干网络	参数量	浮点运算量	mAP50/%
ResNet50	51 684 965	1.751×10^{11}	92.7
ResNet101	70 650 981	2.357×10^{11}	93.0
YOLO v8	2 022 053	6.1×10^{9}	91.5
FasterNet	3 374 105	9.7×10^{9}	92.3

消融实验结果如表 5 所示。数据分析表明，仅将 FAM 集成到 FasterNet 骨干改进网络中后，精确率、召回率和 mAP50 分别提高了 1.1、0.6、0.9 个百分点。这表明 FAM 可以有效地区分特征图中的目

标和噪声,从而使模型更准确地聚焦目标特征,进而成功地提高了模型的识别精度和 mAP50。虽然参数量、浮点运算量有较小幅度的上升,但增幅在可以接受的范围内。此外,仅将 ASCAM 集成到 FasterNet 骨干改进网络中后,通过对不同层次特征图中空间和通道特征的动态调整,ASCAM 可以指导模型有效地区分猪只不同日常行为的特征,从而使模型的精确率、召回率和 mAP50 分别提高 2.7、1.8、1.9 个百分点,这表明模型整体精度显著提升,大大提升了猪只日常行为识别的准确性。通过结合 FasterNet 骨干改进网络、FAM 和 ASCAM 得到本研究所提出的网络 AHMA-PBRNet,性能最佳,精确率、召回率和 mAP50 分别提高了 3.5、2.4、2.8 个百分点,所有数值均优于基线模型。虽然参数量和浮点运算量有较小的增加,但可以满足猪只日常行为识别实时性的要求。综上,模型综合性能达到最优。

表 5 基于 AHMA-PBRNet 网络的消融实验结果

Tab. 5 Results of ablation experiments based on AHMA-PBRNet network

模型	FasterNet 改进	FAM	ASCAM	精确率/%	召回率/%	mAP50/%	参数量	浮点运算量
基线	√	×	×	89.2	88.0	92.3	3.37×10^6	9.70×10^9
模型 1	√	√	×	90.3	88.6	93.2	3.41×10^6	1.07×10^{10}
模型 2	√	×	√	91.9	89.8	94.2	4.18×10^6	1.12×10^{10}
模型 3	√	√	√	92.7	90.4	95.1	4.07×10^6	1.11×10^{10}

注:“√”表示添加相应模块,“×”表示未添加相应模块。

消融实验各模型的精确率-召回率(P-R)曲线如图 6 所示。P-R 曲线的对比结果直观地展示了各模块对 AHMA-PBRNet 整体性能的贡献,也直观地展示了各模块对各种猪只日常行为识别精确率的贡献。

3.4 AHMA-PBRNet 实验分析

为分析 AHMA-PBRNet 模型在不同猪只日常行为上的提升效果,对基线模型与 AHMA-PBRNet 模型进行不同行为的数据对比,实验结果如表 6 所示。

对表 6 数据分析表明,AHMA-PBRNet 对猪只日常行为识别的精确率、召回率和 mAP50 达到了 92.7%、90.4% 和 95.1%,实现了对猪只日常行为的准确识别。具体细化到 7 种日常行为的分析,可

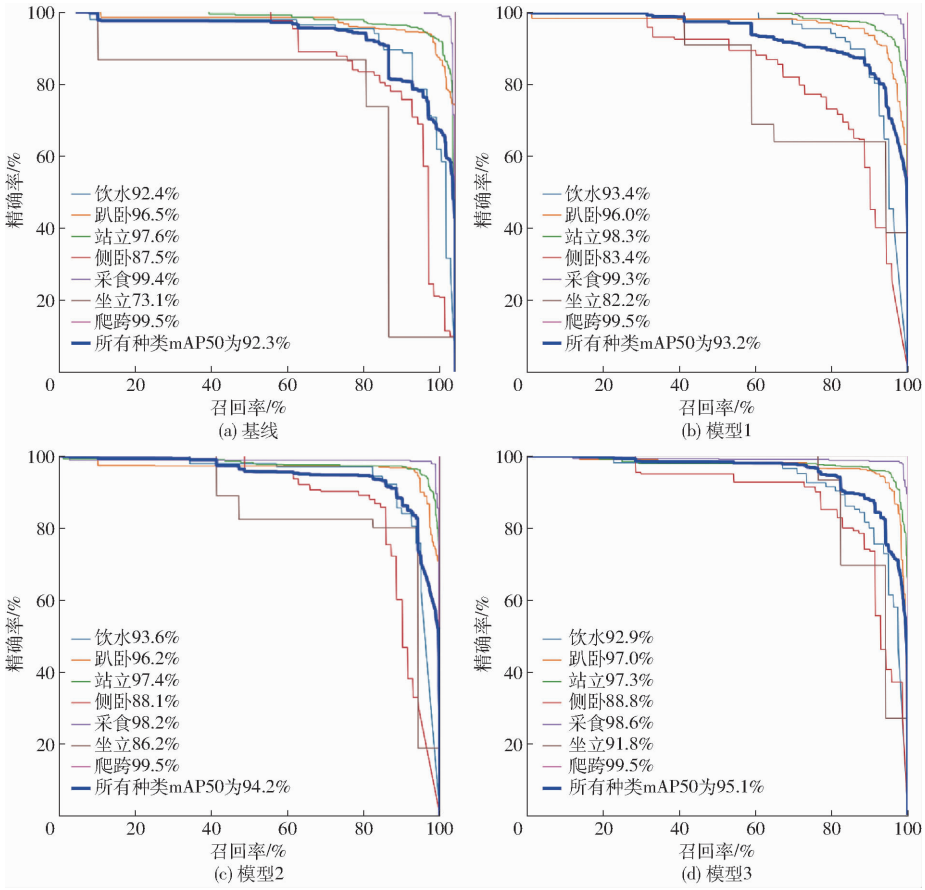


图 6 各模型 P-R 曲线

Fig. 6 P-R curves of model

表 6　猪只不同日常行为的识别实验对比

Tab. 6　Comparison of recognition experiments on different daily behaviors of pigs　%

模型	mAP50							精确率	召回率	mAP50
	站立	坐立	趴卧	侧卧	采食	饮水	爬跨			
基线	97.6	73.1	96.5	87.5	99.4	92.4	99.5	89.2	88.0	92.3
AHMA－PBRNet	97.3	91.8	97.0	88.8	98.6	92.9	99.5	92.7	90.4	95.1

以发现与基线模型相比，通过设计 FAM 和 ASCAM 模块，模型在较难识别的坐立和侧卧行为上均有较显著提升。虽然在站立和采食行为上有细微的降低（这是因为两种行为的识别比较简单，模型为了实现整体性能提升而进行合理权衡），属于可接受的范围，并不会影响模型在猪只日常行为识别任务中的有效性和优越性。

3.5　与其他模型的对比实验

为了验证 AHMA－PBRNet 在猪只日常行为识别中的有效性和优越性，选择了家畜行为识别中使用的主要识别模型进行比较分析。评估包括 Faster R－CNN[22]、EfficientDet[23]、YOLO v5[24]、YOLO v8 和 YOLO v10[25]等知名模型，所有模型都采用相同的数据集和评估标准，结果如表 7 所示。

表 7　不同模型的猪只日常行为识别对比结果

Tab. 7　Comparison results of pigs' daily behavior recognition among different models

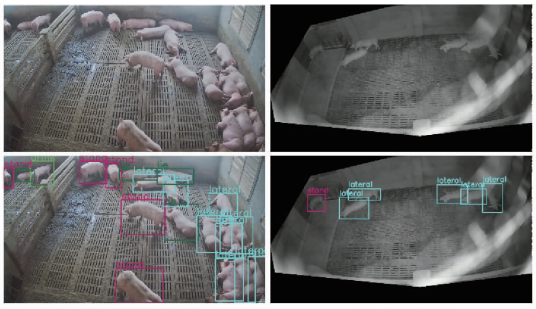
模型	精确率/%	召回率/%	mAP50/%	参数量	浮点运算量
Faster R－CNN	78.4	90.4	90.5	136 795 237	4.018×10^{11}
EfficientDet	90.5	89.3	92.7	6 624 784	1.16×10^{10}
YOLO v5n	88.3	88.5	92.1	2 509 829	7.2×10^{9}
YOLO v8n	88.5	88.2	92.2	3 007 013	8.1×10^{9}
YOLO v10n	86.3	89.6	92.7	2 709 770	8.4×10^{9}
YOLO v10m	92.3	88.6	93.4	16 492 234	6.4×10^{10}
AHMA－PBRNet	92.7	90.4	95.1	4 073 297	1.11×10^{10}

表 7 所示的数据表明，从精度指标来看，AHMA－PBRNet 表现最为出色。其精确率达到 92.7%，召回率为 90.4%，mAP50 高达 95.1%。与 Faster R－CNN 相比，AHMA－PBRNet 的精确率提升了 14.3 个百分点，mAP50 提升了 4.6 个百分点，且召回率相同，这表明在识别准确性和综合性能上，AHMA－PBRNet 远超两阶段的 Faster R－CNN。在 YOLO 系列模型中，YOLO v10m 的精确率为 92.3%，略低于 AHMA－PBRNet，而 mAP50 为 93.4%，比 AHMA－PBRNet 低 1.7 个百分点，可见 AHMA－PBRNet 在精度上优势明显。EfficientDet 和 YOLO v10n 的 mAP50 均为 92.7%，但 AHMA－PBRNet 的精确率较其分别高出 2.2、6.4 个百分点，召回率分别高出 1.1 个百分点和 0.8 个百分点，进一步体现了 AHMA－PBRNet 在识别精度上的优越性。在效率指标方面，Faster R－CNN 的参数量为 136 795 237，浮点运算量为 4.018×10^{11}，是所有模型中最高的，这意味着其计算成本高，不利于在资源有限的场景中部署。EfficientDet 参数量为 6 624 784，浮点计算量为 1.16×10^{10}，具有一定的效率优势。YOLO 系列的 n 版本（YOLO v5n、YOLO v8n、YOLO v10n）参数量均在 3×10^{6} 左右，计算量在 $7.2\times10^{9}\sim8.4\times10^{9}$ 之间，属于轻量级模型，效率较高。AHMA－PBRNet 的参数量为 4 073 297，浮点运算量为 1.11×10^{10}，与 EfficientDet 相比，参数量更低，浮点运算量更优；与 YOLO v10m 相比，参数量仅为其 24.7%，浮点运算量仅为其 17.3%，在保证高精度的同时，具有较好的轻量化程度和计算效率。

综合来看，AHMA－PBRNet 在猪只日常行为识别任务中实现了精度和效率的良好平衡，不仅识别精度最高，而且在模型轻量化和计算效率方面也表现出色，适用于实际的养殖场景监测。

3.6　泛化性实验

为了验证 AHMA－PBRNet 在高密度养殖和夜间环境下猪只日常行为识别中的有效性，分别补充了高密度养殖环境和夜间环境下猪只日常行为识别数据集。这 2 个数据集共由 200 幅图像和对应的标注文件组成，高密度养殖环境和夜间环境各 100 幅，其中站立、坐立、趴卧、侧卧、采食、饮水和爬跨 7 种猪只日常行为实例数量分别为 672、54、482、292、152、122 和 20，示意图如图 7 所示。按常规猪只日常行为识别数据集的划分标准，将 2 个数据集分别划分为训练集、验证集和测试集。随后将之前常规数据集的训练集和这 2 个数据集的训练集混合并训练，进而测试 AHMA－PBRNet 在高密度养殖和夜间环境下猪只日常行为识别效果。其结果如表 8 所示。



(a) 高密度场景 (b) 夜间场景

图 7 高密度和夜间养殖场景下数据集示意图

Fig. 7 Schematic of dataset under high-density and nighttime farming scenarios

表 8 不同场景的猪只日常行为识别结果

Tab. 8 Recognition results of pigs' daily behaviors in different scenarios %

场景	精确率	召回率	mAP50
常规	92.7	90.4	95.1
高密度	82.7	82.2	86.6
夜间	87.4	88.2	91.2

由表 8 可知,AHMA - PBRNet 在常规、高密度、夜间 3 种猪只养殖场景下均展现出有效但有差异的日常行为识别能力。常规场景中下精确率、召回率、mAP50 分别达 92.7%、90.4%、95.1%,验证了模型在理想环境下的高识别精度;高密度场景下 3 项指标均显著下降,分别为 82.7%、82.2% 和 86.6%,这主要因为猪只身体遮挡导致特征重叠干扰,以及混合训练中高密度样本代表性不足,但指标仍保持在 80% 以上,体现模型应对高密度场景的实用价值。夜间场景下指标相较于常规场景有所下降,但光照不足仅导致猪只细节特征弱化而非特征干扰,使召回率下降幅度最小,且 mAP50 超 91%,表明模型具备较强低光照鲁棒性。综合来看,模型场景适应性排序是常规、夜间和高密度,能覆盖养殖场多场景监测需求,但模型对于高密度场景和夜间场景仍有提升空间。后续可通过继续增加高密度遮挡和夜间训练样本进一步优化模型性能。

4 可视化部署

为实现猪只日常行为识别模型的落地应用,提升检测过程的直观性与操作便捷性,基于训练完成的模型构建了群养生猪日常行为识别交互平台,具体实现过程与功能如图 8 所示。

所构建的交互平台以 AHMA - PBRNet 模型为核心检测框架,该模型基于 PyTorch 框架训练完成,其最优权重文件为 AHMA - PBRNet. pt。为适配跨框架部署需求并提升平台兼容性,通过 ONNX(Open neural network exchange)[26] 格式转换技术,将 . pt 权重文件转换为 AHMA - PBRNet. onnx 格式。ONNX 作为一种开放式模型格式,可实现不同深度学习框架间的模型互操作性,为后续在交互平台中调用模型提供了统一的接口支持,确保了模型推理过程的稳定性与高效性。

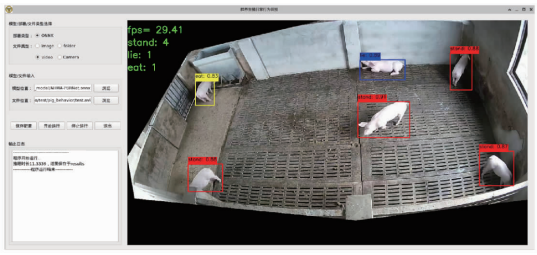


图 8 群养生猪日常行为识别交互平台

Fig. 8 Interactive platform for daily behavior recognition of group-housed pigs

群养生猪日常行为识别系统基于 Qt 框架搭建人机交互(UI)界面,采用 Python 语言实现核心功能开发。平台支持通过导入 AHMA - PBRNet. onnx 模型文件,实现对多种输入类型的检测。这包括 4 种类型,第 1 种是单图检测,支持导入单幅图像文件,实时输出猪只行为识别结果(如站立、趴卧、侧卧、采食等)及目标框标注。第 2 种是批量图像检测,可对文件夹内所有图像进行批量处理,自动生成带识别结果的输出文件夹,适用于大规模数据的批量分析。第 3 种是视频检测,支持导入本地视频文件,逐帧检测并生成带实时标注的输出视频,便于观察猪只行为的动态变化。第 4 种是摄像头在线检测,可调用外接摄像头进行实时流检测,实现养殖场场景的在线监测与行为记录。平台采用模块化设计,将模型推理、数据输入、结果可视化等功能封装为独立模块,通过 Qt 界面实现用户友好的交互操作,降低了非专业人员的使用门槛。

综上,该交互平台实现了猪只日常行为识别从模型到应用的转变,为养殖场的智能化管理提供了直观、高效的技术工具。

5 结论

(1)本研究旨在识别自然养殖环境中猪只各种日常行为,为此提出了一种基于多尺度特征和自适应混合多注意力的猪只日常行为识别框架 AHMA - PBRNet。首先,针对当前很多基于深度学习的家畜行为识别模型,在识别的快速性和准确性上很难兼顾的问题,引入 FasterNet 骨干改进网络,经过骨干网络的对比实验证明了 FasterNet 骨干改进网络能够在猪只日常行为识别的实时性与精确性之间取得最佳平衡。

(2)在特征融合阶段,分别创新地设计了频域

注意力模块和自适应空间-通道混合注意力模块。通过对频域的分析，FAM 可以有效地区分猪只特征图中的目标和噪声，从而使模型更准确地聚焦目标特征；通过对不同层次特征图中空间和通道特征的动态调整，ASCAM 可以有效地区分特征图中的行为与环境干扰以及猪只不同日常行为的特征，从而使模型更精准地捕捉猪只不同日常行为所对应的核心特征。两个模块的有效结合大大提升了 AHMA - PBRNet 识别猪只日常行为的精确率。

(3)在自建的猪只日常行为识别数据集上进行了对比实验，实验结果表明，AHMA - PBRNet 对站立、坐立、趴卧、侧卧、采食、饮水和爬跨行为的 mAP50 达到了 97.3%、91.8%、97.0%、88.8%、98.6%、92.9% 和 99.5%。模型在整体猪只日常行为识别的精确率、召回率和 mAP50 方面均优于现有经典和先进的家畜行为识别模型，分别达到了 92.7%、90.4% 和 95.1%，而参数量和浮点运算量仅为 4 073 297 和 1.11×10^{10}。

(4)为了证实 AHMA - PBRNet 中各关键组件对猪只日常行为识别性能的积极影响，开展了消融实验。实验结果表明，FasterNet 骨干改进网络、FAM 和 ASCAM 共同促进了 AHMA - PBRNet 网络在猪只日常行为识别中取得最佳效果。为了验证 AHMA - PBRNet 的泛化性，分别构建了高密度养殖和夜间环境下猪只日常行为识别数据集，经过实验对比本研究所提出的模型能基本覆盖养殖场多场景监测需求，但模型对于高密度和夜间场景下猪只日常行为识别精度仍有提升空间。后续考虑增加高密度遮挡和夜间训练样本进一步优化模型性能。

(5)为了实现猪只日常行为识别模型的落地应用，提升检测过程的直观性与操作便捷性，基于训练完成的模型 AHMA - PBRNet. pt 构建了群养生猪日常行为识别交互平台，实现了猪只日常行为识别从模型到应用的转变。

参 考 文 献

[1] 代东亮，刘志红，赵存，等. 人工智能技术在畜牧业中的应用进展[J]. 畜牧与饲料科学，2021，42(5)：112 - 119.
Dai Dongliang，Liu Zhihong，Zhao Cun，et al. Research advances on application of artificial intelligence technology in animal husbandry[J]. Animal Husbandry and Feed Science，2021，42(5)：112 - 119.（in Chinese）

[2] Kuster S，Kardel M，Ammer S，et al. Usage of computer vision analysis for automatic detection of activity changes in sows during final gestation[J]. Computers and Electronics in Agriculture，2020，169：105177.

[3] Kashiha A M，Bahr C，Ott S，et al. Automatic monitoring of pig locomotion using image analysis[J]. Livestock Science，2014，159：141 - 148.

[4] 何金成，杨万林，刘涛，等. 基于可穿戴传感器组合部署的猪只行为识别研究[J]. 中国农机化学报，2025，46(4)：42 - 49.
He Jincheng，Yang Wanlin，Liu Tao，et al. Research on pig behavior recognition based on the combination of wearable sensors[J]. Journal of Chinese Agricultural Mechanization，2025，46(4)：42 - 49.（in Chinese）

[5] 闫丽，沈明霞，姚文，等. 基于 MPU6050 传感器的哺乳期母猪姿态识别方法[J]. 农业机械学报，2015，46(5)：279 - 285.
Yan Li，Shen Mingxia，Yao Wen，et al. Recognition method of lactating sows' posture based on sensor MPU6050[J]. Transactions of the Chinese Society for Agricultural Machinery，2015，46(5)：279 - 285.（in Chinese）

[6] Sun G，Shi C，Liu J，et al. Behavior recognition and maternal ability evaluation for sows based on triaxial acceleration and video sensors[J]. IEEE Access，2021，9：65346 - 65360.

[7] 刘冬，何东健，陈晨，等. 基于 ALR - GMM 的群养猪攻击行为识别算法研究[J]. 农业机械学报，2021，52(1)：201 - 208.
Liu Dong，He Dongjian，Chen Chen，et al. Recognition of aggressive behavior in group-housed pigs based on ALR - GMM[J]. Transactions of the Chinese Society for Agricultural Machinery，2021，52(1)：201 - 208.（in Chinese）

[8] Lee J，Jin L，Park D，et al. Automatic recognition of aggressive behavior in pigs using a kinect depth sensor[J]. Sensors，2016，16(5)：631.

[9] 靳敏，杨博文，王春光. 特征提取和特征降维在猪只行为分类识别中的应用[J]. 黑龙江畜牧兽医，2022，659(23)：53 - 59，134.

[10] Mao R，Shen D，Wang R，et al. An integrated gather-and-distribute mechanism and attention-enhanced deformable convolution model for pig behavior recognition[J]. Animals，2024，14(9)：1316.

[11] Cui Y，Wang R，Shen D，et al. DACN - YOLO：an improved abnormal behavior recognition model for pigs[C]// 2024 5th International Conference on Computer Vision，Image and Deep Learning (CVIDL). IEEE，2024：487 - 495.

[12] 薛月菊，朱勋沐，郑婵，等. 基于改进 Faster R - CNN 识别深度视频图像哺乳母猪姿态[J]. 农业工程学报，2018，34(9)：189 - 196.
Xue Yueju，Zhu Xunmu，Zheng Chan，et al. Lactating sow postures recognition from depth image of videos based on improved Faster R - CNN[J]. Transactions of the CSAE，2018，34(9)：189 - 196.（in Chinese）

[13] 李丹, 张凯锋, 李行健, 等. 基于 Mask R-CNN 的猪只爬跨行为识别[J]. 农业机械学报, 2019, 50(增刊 1): 261-266,275.
Li Dan, Zhang Kaifeng, Li Xingjian, et al. Mounting behavior recognition for pigs based on Mask R-CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(S1): 261-266,275. (in Chinese)

[14] 董力中, 孟祥宝, 潘明, 等. 基于姿态与时序特征的猪只行为识别方法[J]. 农业工程学报, 2022, 38(5): 148-157.
Dong Lizhong, Meng Xiangbao, Pan Ming, et al. Recognizing pig behavior on posture and temporal features using computer vision[J]. Transactions of the CSAE, 2022, 38(5): 148-157. (in Chinese)

[15] Luo Y, Lin K, Xiao Z, et al. PBR-YOLO: a lightweight piglet multi-behavior recognition algorithm based on improved YOLOv8[J]. Smart Agricultural Technology, 2025, 10: 100785.

[16] 仝志民, 徐天哲, 石传淼, 等. 基于 CBCW-YOLO v8 的猪只行为识别方法研究[J]. 农业机械学报, 2025, 56(2): 411-419.
Tong Zhimin, Xu Tianzhe, Shi Chuanmiao, et al. Pig behavior recognition based on CBCW-YOLO v8 mode[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(2): 411-419. (in Chinese)

[17] 漆海霞, 冯发生, 尹选春, 等. 轻量级多场景群养猪只行为识别模型研究[J]. 农业机械学报, 2024, 55(10): 306-317.
Qi Haixia, Feng Fasheng, Yin Xuanchun, et al. Research of lightweight multi-scene group pig behavior recognition model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10): 306-317. (in Chinese)

[18] Chen J, Kao S, He H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.

[19] Li D, Zhang K, Li Z, et al. A spatiotemporal convolutional network for multi-behavior recognition of pigs[J]. Sensors, 2020, 20(8): 2381.

[20] 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.

[21] Yaseen M. What is YOLOv8: an in-depth exploration of the internal features of the next-generation object detector[PP/OL]. V1. arXiv (2024-08-28)[2026-08-22]. https://arxiv.org/abs/2408.15857.

[22] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6):1137-1149.

[23] Tan M, Pang R, Le Q V. Efficientdet: scalable and efficient object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10778-10787.

[24] Yu Y, Liu Y, Li Y, et al. Object detection algorithm for citrus fruits based on improved YOLOv5 model[J]. Agriculture, 2024, 14(10): 1798.

[25] Wang A, Chen H, Liu L, et al. YOLO v10: real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107984-108011.

[26] Ren D, Li W, Ding T, et al. Onnxpruner: ONNX-based general model pruning adapter[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(7): 5806-5817.