doi:10.6041/j.issn.1000-1298.2024.05.027

基于改进 ConvNeXt 的奶牛行为识别方法

李恩泽[1,2]　王克俭[1,2]　司永胜[1,2]　苑迎春[1,2]　何振学[1,2]

(1. 河北农业大学信息科学与技术学院, 保定 071001;
2. 河北农业大学河北省农业大数据重点实验室, 保定 071001)

摘要: 奶牛的动作行为(进食、躺卧、站立、行走和甩尾)直接或间接地反映了奶牛的健康及生理状况,是奶牛疾病监测及感知奶牛异常的关键,为准确高效地对奶牛行为进行识别,提出了一种融合时间和空间注意信息的多分支并行的 CAFNet(ConvNeXt - ACM - FAM)奶牛行为识别模型,该模型在卷积网络 ConvNeXt 的基础上融合非对称多分支卷积模块(ACM)和特征注意力模块(FAM)。首先,利用 ACM 划分通道分支提取特征并保留一部分原始特征,防止信息过度丢失。其次,FAM 对不同通道的特征进行融合并引入 SimAM 注意力机制,不增加网络参数的同时增强重要特征的有效提取。实验结果表明,该方法对进食、躺卧、站立、行走和甩尾行为识别准确率分别为 95.50%、93.72%、90.26%、86.43%、89.39%,平均准确率为 91.06%,参数量相较于原模型减少了 1.5×10^6,浮点运算量减少了 3×10^8,相较于其他模型,本文模型识别平均准确率平均提升 8.63 个百分点。本文研究成果可为奶牛疾病监测及预防提供技术支持。

关键词: 奶牛; 行为识别; 非对称卷积; SimAM 注意力; ConvNeXt

中图分类号: TP391.4　　**文献标识码**: A　　**文章编号**: 1000-1298(2024)05-0282-08　　**OSID**:

Cow Behavior Recognition Method Based on Improved ConvNeXt

LI Enze[1,2]　WANG Kejian[1,2]　SI Yongsheng[1,2]　YUAN Yingchun[1,2]　HE Zhenxue[1,2]

(1. *College of Information Science and Technology*, *Hebei Agricultural University*, *Baoding* 071001, *China*
2. *Hebei Key Laboratory of Agricultural Big Data*, *Hebei Agricultural University*, *Baoding* 071001, *China*)

Abstract: The behaviors of cows, including eating, lying, standing, walking and tail-flicking, which directly or indirectly reflects the health and physiological condition of the cows. It is necessary to monitor cow diseases and detect anomalies in cow behavior. In order to achieve the goals, a multi-branch parallel CAFNet (ConvNeXt - ACM - FAM) cow behavior recognition model was proposed by combining temporal and spatial attention information. The model combined an asymmetric multi-branch convolutional module (ACM) and a feature attention module (FAM) on the basis of a ConvNeXt convolutional network. Firstly, ACM was utilized to partition channel branches for feature extraction, and retained some original features to prevent excessive information loss. And ACM can improve the running efficiency of the model. Secondly, FAM fused the features from different channels and introduced the SimAM attention mechanism, which enhanced the efficient extraction of important features without increasing network parameters and improved recognition accuracy. The result of experiment demonstrated that the CAFNet achieved recognition accuracy of the method for eating, lying, standing, walking, and tail-flicking was 95.50%, 93.72%, 90.26%, 86.43%, and 89.39%, respectively. And the average recognition accuracy was 91.06%. Compared with the original model, the number of parameters was reduced by 1.5×10^6, the computational complexity was reduced by 3×10^8, and the average recognition accuracy was increased by 8.63 percentage points. The results can provide technical support for cow disease monitoring and prevention.

Key words: cows; behavior recognition; asymmetric convolution; SimAM attention; ConvNeXt

收稿日期: 2023-10-24　修回日期: 2023-12-09

基金项目: 河北省重点研发计划项目(22327404D)、河北农业大学精准畜牧学科群建设项目(1090064)、河北省自然科学基金项目(F2020204003)和国家自然科学基金项目(62102130)

作者简介: 李恩泽(2000—),男,硕士生,主要从事人工智能和图像处理研究,E-mail: 835514120@qq.com

通信作者: 王克俭(1971—),女,教授,主要从事农业领域图像识别、视频处理及智能算法研究,E-mail: wkj71@163.com

0　引言

奶牛的行为是对其内部生理变化或外部环境刺激做出的反应，直接或间接地反映了奶牛的健康及生理状况[1]。例如，奶牛的躺卧行为是奶牛每日所需时间最长的行为，其躺卧时间与产奶量成正比[2]；奶牛在产犊前甩尾次数增加且幅度增大，可以预估其产犊时间；通过监测奶牛的进食行为评估其饲料摄入量，了解其饮食习惯和食欲情况等。因此，在奶牛养殖过程中对奶牛日常行为进行监测，是实现奶牛健康管理的关键[3]。随着奶牛养殖规模的逐渐扩大，传统的人工监测方法因效率低下、准确率难以保证等原因已无法满足养殖要求[4]，因此，利用先进的信息技术实现智能养殖，保障奶牛身体健康、提高奶牛生活福利和养殖场经济效益已经成为奶牛养殖研究领域的重要目标和任务[5]。

目前的行为识别方法主要分为接触式[6-9]和非接触式[10-14]两种。接触式监测方法主要依靠在奶牛不同部位固定的传感器收集不同行为产生的运动数值，并根据数值的差异进行识别[9]。非接触式监测主要依靠视频图像分析，其中深度学习[15]作为奶牛行为识别的计算方法被广泛应用。

WANG 等[16]在 YOLO v5 模型上引入了空洞空间金字塔池化模块和 C3GC－3 特征增强模块，实现了高层语义信息的多尺度融合，提高了对奶牛爬跨行为的识别准确率。BLOCH 等[17]考虑到训练数据集和迁移学习的使用，构建了基于卷积神经网络（CNN）的奶牛喂养行为分类模型，对不同的进食行为进行分类，平均准确率达到 90%。MA 等[18]通过将 Rexnet 网络扩展为秩扩展三维网络（Rexnet 3D），识别单头奶牛的躺卧、站立和行走行为，识别准确率达到 95.00%，但行为类别较少。王克俭等[19]提出一种基于时空特征（TFL－TSN）的奶牛行为识别算法，对奶牛行走、站立、进食、躺卧 4 种行为的识别平均准确率达到 82.60%。

虽然接触式传感器具有较高识别精度，但随着使用时间的延长，传感器会产生老化或水分渗透等问题，影响记录数值的准确性，从而对行为产生误判。而在非接触式的视频图像分析方法中，虽避免了传感器的弊端，但仍存在识别行为类别较少，个别行为准确率较低的缺点。

针对以上问题，本文以牧场自然环境下群体养殖的奶牛为监测对象，使用高清摄像头实时获取奶牛在牛棚中不同时段的运动行为，为实现对奶牛行为的快速监测和及时干预，选择以图像帧的方式进行处理，构建一种多分支并行且融合空间和通道注意信息的结构（CAFNet 模型），对奶牛多行为识别分类，进而提高模型行为识别准确率与运行效率。

1　实验数据

原始视频数据集采集自河北省保定市满城区宏达牧场，牛舍总长 50 m、宽 30 m，4 个广角摄像机（DH－IPC－HFW8449K－ZFR 型，大华）安装在牛舍中部的左右两侧，安装高度距地面 6 m，相对的 2 台摄像机之间距离为 12 m，俯视角约为 15°，摄像机焦距为 3.6 mm，分辨率为 2 688 像素×1 520 像素。图像采集现场环境如图 1 所示。视频采集时段为 10:00—03:00，由于在实验过程中发现不同光照条件下各行为识别准确率存在差异，将其划分为白天、傍晚和夜间 3 个时间段，其中定义 10:00—17:00 为白天，定义 17:00—19:00 为傍晚，定义 19:00—03:00 为夜间，旨在研究不同时间段的光照条件对奶牛行为识别的影响。本文选取的 5 种基本行为及描述分别为进食（站立时头部伸出栏杆）、躺卧（腹部贴近地面）、行走（蹄部抬起或四肢不在同一水平线）、站立（蹄部贴地且至少两只平行）、甩尾（其他基本行为状态下牛尾非自然下垂），识别时进食主要区分头部及附近特征，躺卧区分腹部附近特征，站立与行走主要区分蹄部是否抬起或腿部是否交叉，甩尾与站立、行走行为主要区分在尾部特征。图 2 为实验所用数据集中的不同行为示例。



图 1　图像采集现场环境

Fig. 1　Image acquisition site environment

对采集到的原始视频进行筛选，剔除掉因拍摄角度或其他因素引起的视频无奶牛主体或视频模糊片段，且截取的视频片段中只包含单头或多头奶牛的同一种行为。考虑到不同时间段的光照条件可能会对识别结果产生影响，因此每种行为均包括白天、傍晚和夜间采集的视频，裁剪得到 924 个视频片段，每个视频片段时长 3～10 s 不等，其中白天、傍晚和夜间的时间比约为 4:3:3，由于奶牛夜间进食行为较少，本文仅对其他 4 种行为的夜间情况进行研究。利用稀疏采样[20]的思想将 924 个视频片段按秒划分为若干个子视频，依次从每个子视频中随机抽取一帧图像，共得到 5 708 幅视频帧图像，按 8:2 划分



图 2 奶牛不同行为示例图

Fig. 2 Example diagrams of different behaviors of cows

为训练集和测试集。图像由“行为名称_编号”命名,且每种行为的样本数据集相对平衡。

2 研究方法

奶牛行为识别可定义为测试行为图像与预先校准的行为参考特征之间的匹配度[21]。由于现阶段奶牛行为识别准确率及效率较低,本文提出了融合时间和空间注意信息的多分支并行卷积识别模型 CAFNet,该模型在卷积网络 ConvNeXt 基础上融合了非对称多分支卷积模块(ACM)和特征注意力模块(FAM),并经过 Layer Scale 与 Dorp Path 操作,如图 3 所示。

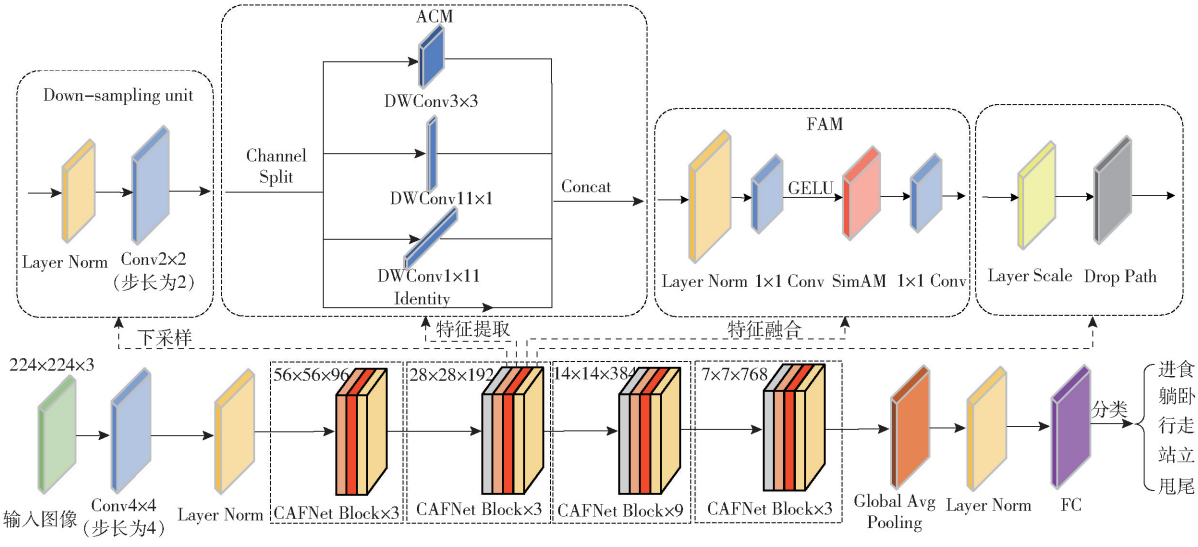

图 3 CAFNet 模型结构图

Fig. 3 CAFNet model structure diagram

首先经过下采样单元进行特征降维,使模型在处理较大尺寸图像时显著降低浮点运算量;然后在 ACM 进行特征提取的过程中保留一部分原始特征,防止信息过度丢失;之后通过 FAM 使模型在空间和通道维度学习更多有辨性神经元并完成不同通道之间的特征融合;再通过 Layer Scale 使模型自适应地缩放每个通道,更好地匹配数据的特性,经过 Dorp Path 随机丢弃网络中的某些通道,使网络不过度依赖特定的权重连接,有助于模型更加适应不同数据样本,最后将结果输入到全连接层得到最终的分类结果。

奶牛行为识别流程图如图 4 所示。首先将采集到的视频序列通过稀疏采样策略,得到本实验所需图像帧数据集并输入到 CAFNet 中;然后分别通过 ACM、FAM 进行特征提取与特征融合;最后,将融合后的特征输入全连接层得到分类结果。

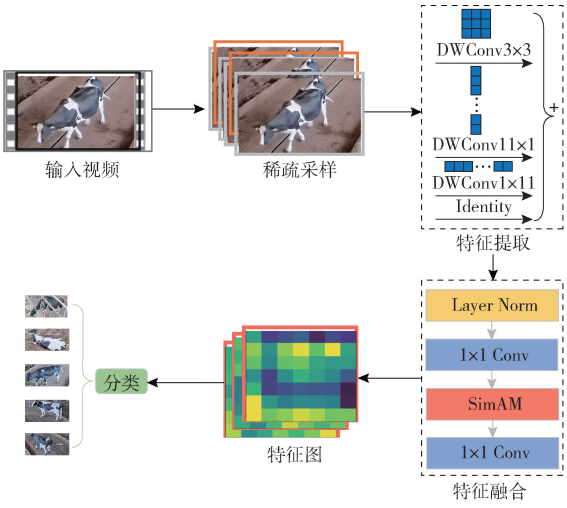

图 4　奶牛行为识别流程图

Fig. 4　Cow behavior recognition flowchart

2.1　奶牛图像特征提取

因传统卷积相关模型通常采用较大卷积核进行深层堆叠，导致参数量和浮点运算量急剧增加，在计算资源有限的情况下无法高效运行，且容易受到训练和推理速度的影响。因此在逐通道卷积的基础上引入了 ACM 模块。采用小核逐通道卷积与非对称卷积的组合，并保持部分通道不变，将其标识为恒等映射的一个分支。通过保留部分通道，在特征提取过程中保持一些原始输入信息，有助于防止信息过度丢失，并提供一种路径，使模型能够利用原始输入的一些低级特征或低频信息。对于输入 X，沿着通道维度分为 4 组并行分支，其数学表达式为

$$Split(X_{hw},X_w,X_h,X_{id})=(X_{0:g},X_{g:2g},X_{2g:3g},X_{3g:4g}) \tag{1}$$

式中　$Split(\cdot)$——沿通道维度划分操作

$X_{0:g}$——输入 X 从 0 到 g 的分支

$X_{g:2g}$——输入 X 从 g 到 $2g$ 的分支

$X_{2g:3g}$——输入 X 从 $2g$ 到 $3g$ 的分支

$X_{3g:4g}$——输入 X 从 $3g$ 到 $4g$ 的分支

X_{hw}——方形卷积核分支

X_w——宽度方向非对称卷积分支

X_h——高度方向非对称卷积分支

X_{id}——恒等映射分支

g——卷积分支的通道数

下一步划分输入到不同的并行分支，采用滑动窗口机制[22]对得到的特征图进行特征提取，最后将每个分支的输出连接合并。

$$X'_{hw}=\mathrm{DWConv}^{g\to g}_{k_s\times k_s}g(X_{hw}) \tag{2}$$

$$X'_{w}=\mathrm{DWConv}^{g\to g}_{1\times k_b}g(X_{w}) \tag{3}$$

$$X'_{h}=\mathrm{DWConv}^{g\to g}_{k_b\times 1}g(X_{h}) \tag{4}$$

$$X'_{id}=X_{id} \tag{5}$$

$$X'=\mathrm{Concat}(X'_{hw},X'_{w},X'_{h},X'_{id}) \tag{6}$$

式中　DWConv——深度可分离卷积操作

Concat——拼接操作

X'_{hw}——方形卷积核分支输出

X'_{h}——高度方向非对称卷积分支输出

X'_{w}——宽度方向非对称卷积分支输出

X'_{id}——恒等映射分支

X'——不同分支输出拼接后的结果

k_s——方形卷积核大小

k_b——非对称卷积核大小

非对称卷积核大小决定了模型在输入数据中能够捕获的局部信息的最大范围，较大的卷积核可以覆盖更大的感受野，能够捕获更广泛的上下文信息。但形状过大需要更多参数和计算量，且在处理局部细节和细粒度特征时效果较差；形状过小在感受野范围内能够捕获的局部信息相对有限，特征表达能力较低。本文采取大小为 11 的非对称卷积核，如图 5 所示。

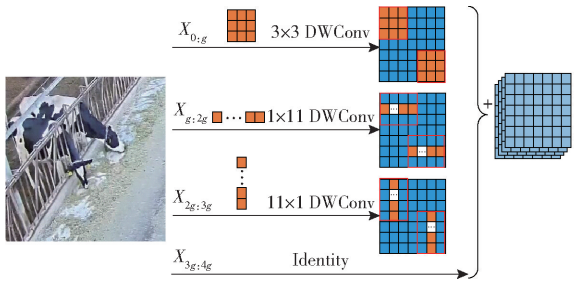

图 5　非对称多分支卷积模块

Fig. 5　Asymmetric multi branch convolutional module

2.2　奶牛图像特征融合

因现有注意力模块只能沿着通道或空间维度细化特征，限制了模型学习跨通道和空间变化注意力权重的灵活性。为从全局视图中捕获更多的上下文线索，同时针对因拍摄角度导致的奶牛主体较小，模型存在将奶牛主要部位推理为背景，无法正常检测奶牛行为的问题，对于特征提取后的结果，本文采用图 6 所示的特征注意力模块。FAM 通过不同通道之间的特征映射增强特征之间的关联性，获取更多有关小目标物体的特征。实验使用 3D 无参注意力模块(SimAM)[23]，其独特优势在于无需增加额外参数，即可为特征图推导出三维权值(空间和通道维度)，使得网络学习更多有辨性神经元以增强重要特征的有效提取，强化奶牛主体特征并抑制背景特征。

针对每层提取的特征，通过逐点卷积对不同通道之间的特征进行融合，再经过非线性、计算复杂度低的 GELU 激活函数得到初步融合的特征图；随后对该特征图采用 SimAM 注意力机制直接提取空间和通道上的重要特征并通过逐点卷积进行融合，得

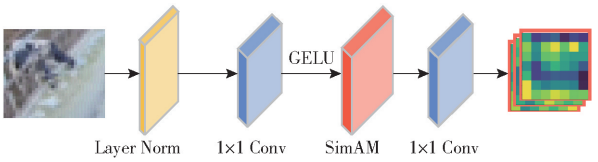

图 6 特征注意力模块

Fig. 6 Feature attention module

到下一阶段的输入。逐点卷积融合计算公式为

$$T_{out}(i,j)=\sum(T_{in}(i,j,s)W_s)\quad(s\in[0,K-1]) \tag{7}$$

式中 T_{out}——输出特征映射

T_{in}——输入特征映射

(i,j)——输出特征映射中元素位置

s——输入特征映射通道索引

W_s——滤波器

K——输入特征映射通道数

其中,SimAM 注意力机制直接对特征图推导出三维权值,寻找表现出明显空间抑制效应的神经元并将其赋予更高的优先权,其能量函数计算得到的能量越低,说明神经元 t 与周围神经元区别越大,所以为神经元 t 赋予更高的优先权,从而增强奶牛主要特征提取,有效抑制了周围环境特征干扰。

计算每个神经元能量函数公式为

$$e_t(w_t,b_t,y,x_i)=\frac{1}{M-1}\sum_{i=1}^{M-1}[-1-(w_tx_i+b_t)]^2+[1-(w_tt+b_t)]^2+\lambda w_t^2 \tag{8}$$

式中 e_t——神经元能量函数

w_t——权重系数 b_t——偏置系数

y——标签值

x_i——输入特征 X 其他神经元

M——某通道上所有神经元个数

t——输入特征 X 目标神经元

λ——正则化系数

2.3 奶牛行为分类

为保证奶牛行为分类准确率及效率,在提取并融合图像的特征后,使用全局平均池化层[24](Global average pooling, GAP)、层归一化[25-26](Layer normalization,LN)与全连接层的组合。其中替代平均池化层(Average pooling,AP)的 GAP 对整个特征图进行平均,不需要额外参数,得到固定长度的特征向量且具有位置不变性,不同位置的相同目标都会生成相似的特征表示,使得模型对于目标物体在图像中的位置变化更加鲁棒,GAP 通过将整个特征图平均化,捕捉整体图像的全局信息,从而更好地保留特征图中的重要特征,使模型更加有效地进行分类。其次对于 GAP 操作后得到的特征向量,使用 LN 替代传统卷积网络中的批归一化(Batch normalization, BN)进行归一化,使得每个维度上的特征具有相似的均值和方差,规范化特征向量,减少特征间协变量偏移,且 LN 在每个样本上独立计算均值和方差,能够更好地保持输出特征通道之间的独立性。最后将学习到的特征作为全连接层的输入进行分类并输出。

3 实验与结果分析

3.1 实验设置

实验在 Linux 4.15.0 系统下,基于 PyTorch 1.13.1 + CUDA 11.4 实现。计算机配置为 CPU Intel(R) Xeon(R) Gold 5220,NVIDIA Quadro RTX 5000 显卡。激活函数为 GELU[27] 函数,批处理量设置为 16,迭代次数为 300。为了使模型更好地拟合数据,采用余弦退火(Cosine annealing)策略[28],初始学习率设为 0.000 1,最大倍率因子为 0.001,最小倍率因子为 0.000 001,学习率在训练过程中以余弦函数的形状逐渐减小,使得网络在训练初期可以更快地收敛,而在后期可以更加细致地调整参数。

3.2 评价指标

选取参数量(Parameters)、浮点运算量(FLoPs)、帧速率(Frame per second,FPS)、平均准确率(Average accuracy,Acc)作为评价指标衡量模型性能,其中参数量和计算复杂度越小说明模型具有更快的推理速度和更低的资源消耗且易于部署;帧速率为每秒传输的帧数,表示模型每秒检测的图像帧数,高帧速率表示模型能够更快地处理图像数据;高平均准确率表明模型能更加正确地对不同类别的样本进行分类。

3.3 结果与分析

3.3.1 不同大小非对称卷积核对模型的影响

为寻找本模型 ACM 模块中非对称卷积核的最优大小,采集白天(2 280 幅)、傍晚(1 640 幅)和夜间(1 788 幅)的视频帧图像,对不同大小非对称卷积核进行测试,实验证明当非对称卷积核大小为 11 时效果最好,如表 1 所示。

表 1 不同大小非对称卷积核对模型的影响

Tab. 1 Impact of different sizes of asymmetric convolution kernels on model

非对称卷积核大小	帧速率/(f·s^{-1})	平均准确率/%
3	83.56	87.40
5	94.62	88.90
7	94.30	88.78
9	96.95	89.08
11	122.78	91.06
13	105.11	89.27

在测试过程中，非对称卷积核初始大小设定为3，随着卷积核大小的逐渐增加，本模型参数量及浮点运算量变化可忽略不计，而当卷积核大小为 11 时，模型的帧速率达到 122.78 f/s，平均准确率达到 91.06%，均为本次测试最大值，故而本文采用大小为 11 的非对称卷积核。

3.3.2　不同模型对比

在上述相同实验条件下将 CAFNet 模型与 ResNet 50、ResNeXt50、TFL－TSN、Swin－Transformer 模型进行对比实验，并与图像序列输入方式进行对比，由表 2 可知，本文模型识别平均准确率分别提升 13.03、7.95、8.46、5.06 个百分点。相较于 ResNet 50、ResNeXt50、TFL－TSN，参数量略微增加，帧速率与识别平均准确率提升效果明显，其原因在于本文提出的 CAFNet 能有效利用小核多分支的特点提升模型推理速度，扩大感受野并增强学习局部特征的能力，结合空间和通道下的不同特征信息，从而提高不同行为的辨识度与识别准确率。

表 2　不同模型实验结果对比

Tab. 2　Comparison of experimental results of different models

模型	输入方式	参数量	浮点运算量	帧速率/(f·s^{-1})	平均准确率/%
ResNet 50	图像帧	2.352×10^7	4.14×10^9	85.05	78.03
	图像序列	2.431×10^7	4.21×10^9	78.35	76.88
ResNeXt50	图像帧	2.299×10^7	4.27×10^9	81.78	83.11
	图像序列	2.356×10^7	4.38×10^9	76.65	82.96
TFL－TSN	图像帧	2.556×10^7	4.33×10^9	48.46	82.60
	图像序列	2.784×10^7	4.41×10^9	40.67	83.78
Swin－Transformer	图像帧	2.829×10^7	4.36×10^9	60.83	86.00
	图像序列	2.998×10^7	4.45×10^9	50.73	89.27
CAFNet	图像帧	2.752×10^7	4.17×10^9	122.78	91.06
	图像序列	2.859×10^7	4.47×10^9	96.09	90.81

通过表 2 中不同输入方式对比可见，ResNet 50、ResNeXt50、TFL－TSN、Swin－Transformer、CAFNet 的图像帧输入方式在参数量、浮点运算量、帧速率方面均优于序列输入，这是由于序列输入需要更多的参数和更高的浮点运算量来捕获序列中的时间依赖关系，由此导致处理速度降低，帧速率随之下降。本文模型识别平均准确率变化不大，而 TFL－TSN、Swin－Transformer 的识别平均准确率虽然有所增加，但均低于本文 CAFNet 的图像帧输入方式，因此本文采用图像帧输入方式的 CFANet 模型对奶牛行为进行识别分类。

3.3.3　消融实验

为验证 ACM 模块和 FAM 模块对改进的奶牛行为识别模型的影响，在上述相同实验条件下，利用该方法对建立的奶牛行为数据集进行识别，参考控制变量法的思想，保证学习率、优化器等参数一致，通过消融实验验证本文模型的效果，结果如表 3 所示。

表 3　模型优化消融实验结果对比

Tab. 3　Comparison of model optimization ablation experimental results

模块		参数量	浮点运算量	帧速率/	准确率/%					平均
ACM	FAM			(f·s^{-1})	进食	躺卧	站立	行走	甩尾	准确率/%
×	×	2.902×10^7	4.47×10^9	127.46	87.71	88.70	82.33	78.80	81.21	83.75
√	×	2.752×10^7	4.17×10^9	144.74	90.36	90.20	85.60	82.45	77.19	85.16
×	√	2.902×10^7	4.57×10^9	109.68	91.50	93.65	86.30	82.67	89.13	88.65
√	√	2.752×10^7	4.17×10^9	122.78	95.50	93.72	89.26	85.43	91.39	91.06

注：“×”表示不使用该模块，“√”表示使用该模块。

相较于原模型，对比引入 ACM 模块前后，参数量减少、浮点运算量减少、帧速率增加，这是由于原模型 7×7 方形卷积被替换后参数量减少，且每个位置所需要进行乘法和累加操作的总数减少，加快了模型的推理速度。在 ACM 提取特征时，更大的感受野增加了网络捕获上下文的能力，进食、躺卧、站立、行走 4 种行为识别准确率均有所提高。考虑到甩尾与站立、行走两种行为在识别时相似度过大，且牛尾、牛蹄在图像帧中占比过小，通过引入 FAM 模块，对各通道添加注意力机制并融合各通道特征，聚焦当前图像的重要特征，不同时间段光照下的效果如图 7 所示。

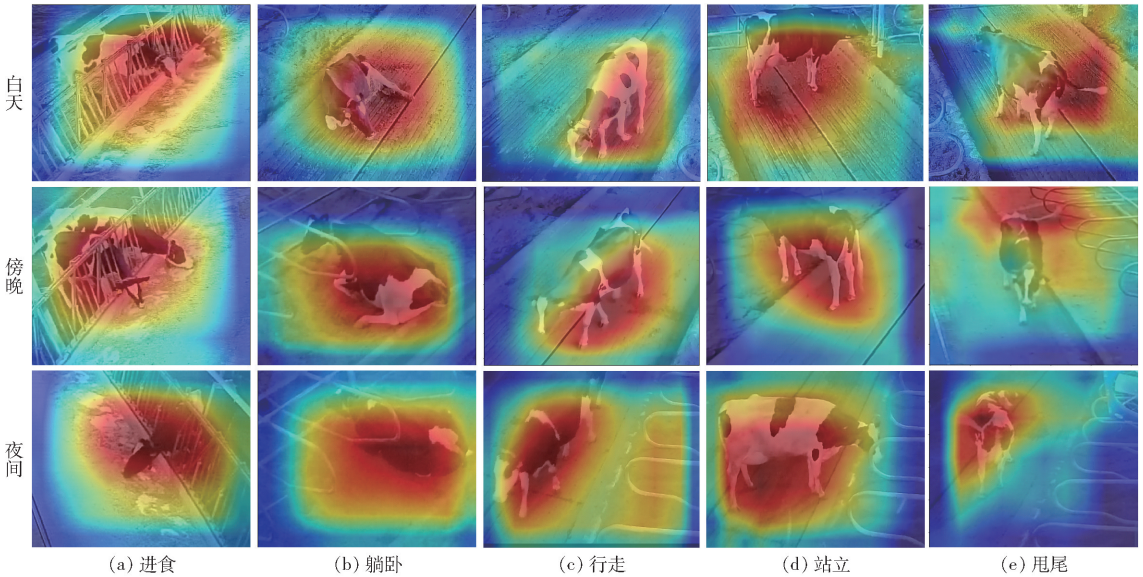

图 7 不同光照条件下 SimAM 注意力机制识别效果可视化

Fig. 7 Visualization of SimAM attention mechanism under different lighting conditions

进食行为集中学习头部及附近食槽,躺卧行为聚焦在腹部及贴近地面部分,行走及站立行为更加关注腿部和蹄部特征,而甩尾作为在其他基本行为状态下发生的行为,模型更注重学习尾部附近特征。在引入 FAM 后,模型识别的 5 种行为准确率均有提升,原因在于 FAM 从全局视图中捕获了更多上下文线索,降低了模型将奶牛主体或部位(牛尾、牛蹄)识别为背景的概率。SimAM 注意力机制的添加增强了模型学习细化特征的能力,使得模型对于一些细小特征的变化更加敏感,尤其在甩尾、行走行为发生时,牛尾和牛蹄附近神经元表现出明显空间抑制效应,FAM 为其赋予更高的优先权,抑制了周围环境特征的干扰,从而有效区分甩尾、站立、行走 3 种行为,其中甩尾行为识别效果提升最为明显。而计算复杂度略微增加,帧速率有所下降,主要由于逐点卷积在融合不同通道信息的过程中增加了模型的计算负载,尽管 SimAM 注意力机制无参数增加,但在计算注意力权重时,仍需要根据特征相似度进行加权以及对特征进行计算和比较操作,行为识别准确率的提升证明该方法是可行的。总体结果表明,本文方法可以实现高精度奶牛基本行为识别。

3.3.4 不同光照条件下模型鲁棒性分析

考虑到奶牛养殖场随时间变化,不同光照强度会干扰识别准确率,为保证选取的数据存在光照差异较大,选取 10:00—03:00 拍摄的视频,时间段包括白天、傍晚、夜间。分析不同时段光照下本文方法的实用性。由于夜间奶牛进食行为较少,本实验在夜间仅对其余 4 种行为进行识别。

验证结果如表 4 所示,同一行为识别准确率随着时间的推移而略微降低,可能原因是光照强度的降低使模型将奶牛主体识别为背景,但总体差距并不大。因非对称卷积核与 SimAM 的组合贯穿整个模型,ACM 提取的较大感受野特征通过 FAM 模块融合空间和通道中包含的光照信息,并传递给下一模块,大大降低了模型将奶牛主体推理为背景的概率。实验结果表明,不同时间段光照条件的变化会对奶牛行为识别准确率产生一定的影响,但本文模型所受光照影响较小。

表 4 不同光照条件下本文模型识别效果对比

Tab. 4 Comparison of lighting results at different time periods %

光照条件	准确率					平均准确率
	进食	躺卧	站立	行走	甩尾	
白天	96.80	94.72	91.26	88.43	92.39	92.72
傍晚	96.00	94.10	91.07	87.40	89.42	91.60
夜间		93.88	89.37	86.69	85.50	88.86

4 结论

(1) 为实现养殖场奶牛行为的识别,在 ConvNeXt 模型的基础上,引入非对称多分支卷积模块和特征注意力模块,提出了一种多分支并行结构且融合空间和通道注意力信息的奶牛行为识别模型。

(2)实验结果表明,CAFNet 模型对进食、躺卧、站立、行走、甩尾行为的识别准确率分别达到 95.50%、93.72%、89.26%、85.43%、91.39%。本文模型相较于 ResNet 50、ResNeXt50、TFL - TSN、Swin - Transformer 平均准确率分别提高 13.03、7.95、8.46、5.06 个百分点。

(3)不同时间段光照条件下的实验结果表明,

光照强度的变化对奶牛行为识别准确率产生一定的影响，但本文模型所受影响较小。本文模型可以满足奶牛养殖场环境下奶牛行为识别的需要，为奶牛疾病监测及感知奶牛异常提供了技术支持。

参 考 文 献

[1] 程国栋, 吴建寨, 邢丽玮,等. 基于 IMU 的细粒度奶牛行为判别[J]. 中国农业大学学报, 2022, 27(4): 179 - 186.
CHENG Guodong, WU Jianzhai, XING Liwei, et al. Fine-grained cow behavior discrimination based on IMU[J]. Journal of China Agricultural University, 2022, 27(4): 179 - 186. (in Chinese)

[2] 张宏鸣, 孙扬, 赵春平,等. 反刍家畜典型行为监测与生理状况识别方法研究综述[J]. 农业机械学报, 2023, 54(3): 1 - 21.
ZHANG Hongming, SUN Yang, ZHAO Chunping, et al. Review on typical behavior monitoring and physiological status identification methods of ruminants[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(3): 1 - 21. (in Chinese)

[3] 赵凯旋, 何东健. 基于卷积神经网络的奶牛个体身份识别方法[J]. 农业工程学报, 2015, 31(5): 181 - 187.
ZHAO Kaixuan, HE Dongjian. Individual cow identification method based on convolutional neural network[J]. Transactions of the CSAE, 2015, 31(5): 181 - 187. (in Chinese)

[4] REITH S, HOY S. Review: behavioral signs of estrus and the potential of fully automated systems for detection of estrus in dairy cattle[J]. Animal, 2018, 12(2): 398 - 407.

[5] 康熙, 刘刚, 初梦苑,等. 基于计算机视觉的奶牛生理参数监测与疾病诊断研究进展及挑战[J]. 智慧农业(中英文), 2022, 4(2): 1 - 18.
KANG Xi, LIU Gang, CHU Mengyuan, et al. Research progress and challenges of physiological parameter monitoring and disease diagnosis of dairy cows based on computer vision[J]. Smart Agriculture, 2022, 4(2): 1 - 18. (in Chinese)

[6] RIABOFF L, POGGI S, MADOUASSE A,et al. Development of a methodological framework for a robust prediction of the main behaviours of dairy cows using a combination of machine learning algorithms on accelerometer data[J]. Computers and Electronics in Agriculture, 2020, 169:105179.

[7] 王俊, 张海洋, 赵凯旋,等. 基于最优二叉决策树分类模型的奶牛运动行为识别[J]. 农业工程学报, 2018, 34(18): 202 - 210.
WANG Jun, ZHANG Haiyang, ZHAO Kaixuan, et al. Dairy cow behavior recognition based on optimal binary decision tree classification model[J]. Transactions of the CSAE, 2018, 34(18): 202 - 210. (in Chinese)

[8] PENG Y, KONDO N, FUJIURA T, et al. Classification of multiple cattle behavior patterns using a recurrent neural network with long short-term memory and inertial measurement units[J]. Computers and Electronics in Agriculture, 2019, 157: 247 - 253.

[9] HOANG Q T, PHUNG C P K, BUI T N, et al. Cow behavior monitoring using a multidimensional acceleration sensor and multiclass SVM[J]. International Journal of Machine Learning and Networked Collaborative Engineering, 2018, 2(3): 110 - 118.

[10] WU D, WANG Y, HAN M, et al. Using a CNN - LSTM for basic behaviors detection of a single dairy cow in a complex environment[J]. Computers and Electronics in Agriculture, 2021, 182: 106016.

[11] YIN X, WU D, SHANG Y, et al. Using an EfficientNet - LSTM for the recognition of single cow's motion behaviours in a complicated environment[J]. Computers and Electronics in Agriculture, 2020, 177: 105707.

[12] FUENTES A, YOON S, PARK J, et al. Deep learning-based hierarchical cattle behavior recognition with spatio-temporal information[J]. Computers and Electronics in Agriculture, 2020, 177: 105627.

[13] 王少华, 何东健, 刘冬. 基于机器视觉的奶牛发情行为自动识别方法[J]. 农业机械学报, 2020, 51(4): 241 - 249.
WANG Shaohua, HE Dongjian, LIU Dong. Automatic recognition method of dairy cow estrus behavior based on machine vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(4): 241 - 249. (in Chinese)

[14] YANG Q, XIAO D, CAI J. Pig mounting behaviour recognition based on video spatial - temporal features[J]. Biosystems Engineering, 2021, 206: 55 - 66.

[15] 刘忠超, 何东健. 基于卷积神经网络的奶牛发情行为识别方法[J]. 农业机械学报, 2019, 50(7): 186 - 193.
LIU Zhongchao, HE Dongjian. Recognition method of cow estrus behavior based on convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(7): 186 - 193. (in Chinese)

[16] WANG R, GAO Z, LI Q, et al. Detection Method of cow estrus behavior in natural scenes based on improved YOLO v5[J]. Agriculture, 2022, 12(9): 1339.

[17] BLOCH V, FRONDELIUS L, ARCIDIACONO C, et al. Development and analysis of a CNN-and transfer-learning-based classification model for automated dairy cow feeding behavior recognition from accelerometer data[J]. Sensors, 2023, 23(5): 2611.

[18] MA S, ZHANG Q, LI T, et al. Basic motion behavior recognition of single dairy cow based on improved Rexnet 3D network [J]. Computers and Electronics in Agriculture, 2022, 194: 106772.

[19] 王克俭,孙奕飞,司永胜,等. 基于时空特征的奶牛视频行为识别[J]. 农业机械学报,2023,54(5):261 - 267,358.
WANG Kejian, SUN Yifei, SI Yongsheng, et al. Video behavior recognition of dairy cows based on spatio-temporal features [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(5): 261 - 267,358. (in Chinese)

(下转第 404 页)