



基于 CNN-BiLSTM 的可穿戴传感器数据犬类行为识别

杨毅, 柯俊

(浙江理工大学机械工程学院, 杭州 310018)

摘要: 针对犬类行为识别中可穿戴传感器数据特征提取困难、样本类别不平衡以及长时序依赖关系难捕捉等问题, 提出了一种基于数据增强与 CNN-BiLSTM 的混合识别方法。该方法利用鲁棒标准化(Robust normalization, RN)与主成分分析(Principal components analysis, PCA)消除异常值并降低特征维度, 通过重采样策略缓解类别不平衡问题; 在此基础上, 利用卷积神经网络(Convolutional neural network, CNN)提取局部空间特征并抑制高频噪声, 并引入双向长短期记忆网络(Bidirectional long short-term memory, BiLSTM)构建双向时序依赖模型。结果显示: 相较于单向 CNN-LSTM 模型, CNN-BiLSTM 模型的准确率与 F1 分数提升了 2.3% 和 2.1%, 特别是“玩耍”这一复杂行为的 F1 分数提高了 26.0%; 相比于其他主流行为识别算法, CNN-BiLSTM 在处理多达 9 种行为类别的情况下, 仍保持了较高的识别准确率。该研究为基于可穿戴设备的犬类行为监测和识别提供了较为可靠的解决方案。

关键词: 犬类行为识别; 卷积神经网络; 可穿戴传感器; 深度学习; 双向长短期记忆网络; 鲁棒标准化

中图分类号: TP183

文献标志码: A

文章编号: 1673-3851(2026)05-0306-09

引文格式: 杨毅, 柯俊. 基于 CNN-BiLSTM 的可穿戴传感器数据犬类行为识别[J]. 浙江理工大学学报(自然科学), 2026, 55(3): 306-314.

Reference Format: YANG Yi, KE Jun. CNN-BiLSTM-based dog activity detection using wearable sensor data[J]. Journal of Zhejiang Sci-Tech University, 2026, 55(3): 306-314.

CNN-BiLSTM-based dog activity detection using wearable sensor data

YANG Yi, KE Jun

(School of Mechanical Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: A hybrid detection method based on data augmentation and CNN-BiLSTM was proposed to meet the difficulties in feature extraction from wearable sensor data, class imbalance, and the challenge of capturing long-term temporal dependencies in canine behavior recognition. This method used robust normalization (RN) and principal component analysis (PCA) to eliminate outliers and reduce feature dimensions, and mitigated class imbalance through a resampling strategy. On this basis, convolutional neural network (CNN) was used to extract local spatial features and suppress high-frequency noise, bidirectional long short-term memory networks (BiLSTM) were introduced to construct a bidirectional temporal dependency model. The experimental results showed that compared to the unidirectional CNN-LSTM model, the CNN-BiLSTM model proposed in this paper achieved an improvement of 2.3% in accuracy and 2.1% in F1 score, with the F1 score for the complex behavior "playing" being improved by 26.0%. Compared with other mainstream behavior recognition algorithms, CNN-BiLSTM maintained a high recognition accuracy even when handling up to nine types of behavioral categories. It provided a reliable solution for dog behavior monitoring and recognition based on wearable devices.

Key words: dog activity detection; convolutional neural network; wearable sensors; deep learning; bidirectional long short-term memory network; robust normalization

收稿日期: 2025-12-08 网络出版日期: 2026-03-05

基金项目: 国家自然科学基金项目(52102430)

作者简介: 杨毅(2001—), 男, 江西上饶人, 硕士研究生, 主要从事智能信息处理方面的研究。

通信作者: 柯俊, E-mail: jlujun@163.com

0 引言

犬类行为识别旨在基于可穿戴传感器采集的相关数据实现犬类行为的精准识别,该技术对保障犬类健康和建立和谐人犬互动关系方面有重要意义^[1]。犬类行为识别可帮助主人及时发现其健康隐患,进而进行干预治疗;可拓展应用于生态学以及野生动物保护领域,为其他动物的行为模式分析、迁徙路线跟踪和濒危物种的定位保护等提供技术支撑^[2-3]。

犬类行为识别的现有主流方法为人工观察与视频观测,但这类方法效率低、主观性强且准确率低。近年来,随着传感器技术的发展,可穿戴式设备逐渐向集成化、小型化与智能化方向发展^[4],传感器技术的革新也为准确识别犬类行为提供了可能。基于传感器数据的犬类行为识别研究已经历了 2 个阶段。早期研究主要采用机器学习的方法。Ladha 等^[5]通过对加速度数据进行滑动窗口分割与经验累积分布函数归一化操作,将原始的传感器时序数据转化为高判别性的特征,并采用 K 最近邻 (K -nearest neighbor, KNN) 实现了 17 种犬类行为的分类;该方法为犬类行为识别建立了早期框架,但提取统计特征的方式较为简单,不能较好区分相似行为。Gerencsér 等^[6]采用支持向量机 (Support vector machine, SVM) 同步分析了加速度与角速度数据,提高了犬类行为识别的准确率;但人工设计的特征易对训练个体的运动模式产生过拟合,在跨个体应用时识别精度下降。Davoulos 等^[7]对比了随机森林 (Random forest, RF) 等集成模型,发现集成模型能在较低的计算复杂度下保持较高的检测精度;但该方法难以捕捉传感器数据中的时空关联特征,在识别动态行为转换时表现不佳。Kumpulainen 等^[8]对比了传感器不同放置位置的识别表现,发现位置的变动会对识别精度产生显著影响。Marcato 等^[9]将 RF 与孤立森林 (Isolation forest, IF) 进行级联,提高了数据集中少数类行为的识别准确率;但该方法增加了模型的调优难度,且仍依赖复杂的特征工程。Muminov 等^[10]通过为每个数据点计算四元数值,生成少量分类特征,弥补了原始数据在犬只关键姿势判定中的局限性,并结合高斯朴素贝叶斯 (Gaussian naive Bayes, GNB) 提高了识别精度;但该方法研究多在半受控环境 (实验室以及指定的运动场所) 开展,在复杂环境以及传感器噪声干扰下,检测精度易出现急剧下降。

随着深度学习方法的发展,卷积神经网络

(Convolutional neural network, CNN) 因其优秀的局部特征提取能力,在行为识别技术领域展现出突出优势^[11-12]。Amano 等^[13]引入 CNN 直接从原始信号中自动学习特征,取代了繁琐的手工设计特征,但该方法无法捕捉长时序依赖关系;Hussain 等^[14]针对时间序列数据的时序依赖问题,探索了长短期记忆网络 (Long short-term memory, LSTM) 的应用,通过引入门控机制记录历史信息提高了识别准确率,但单向的网络结构并不能预知未来信息,无法结合上下文信息对识别结果进行修正;Kim 等^[15]进一步构建了 CNN-LSTM 混合架构,通过 CNN 提取空间局部特征, LSTM 提取时序特征,提高了识别精度。深度学习方法在行为识别领域已取得明显成果,但在实际应用中仍面临许多挑战,传感器数据噪声大、特征冗余一直是亟待解决的难点。目前的研究虽利用 CNN 提取特征来增强表征能力,但常常忽视了原始信号中的异常值干扰问题,导致模型无法对静态动作作出准确的判断。针对行为样本类别严重不平衡的问题,现有大部分研究缺乏有效的数据集重采样策略,使得模型对于少数类动作存在识别倾向性;针对复杂动态行为识别难的问题,现有的 LSTM 模型多是单向的,只能利用历史信息,并没有考虑未来上下文信息,在捕捉动作转换过程中会存在一些盲区,无法对“玩耍”等不平稳行为进行准确地识别。

针对犬类行为识别中可穿戴传感器数据特征提取困难、样本类别不平衡以及长时序依赖关系难捕捉等问题,本文融合 CNN 与 BiLSTM 的优势,提出了一种基于 CNN 和双向长短期记忆网络 (Bidirectional long short-term memory, BiLSTM) 的混合深度学习模型 CNN-BiLSTM,以精确识别犬类行为。综合采用鲁棒标准化 (Robust normalization, RN) 与主成分分析 (Principal components analysis, PCA) 消除异常值并降低特征维度,通过重采样策略缓解类别不平衡问题;在此基础上,利用 CNN 提取局部空间特征并抑制高频噪声,并引入 BiLSTM 构建双向时序依赖模型。本文在公开的犬类行为数据集上进行了实验验证,旨在挖掘可穿戴感知技术的应用潜力,为动物行为学的智能化分析提供新的研究视角。

1 方法

本文提出的 CNN-BiLSTM 模型,在数据预处理环节使用 RN 和 PCA 及重采样技术来改善原始输入的质量,并缓解各类别样本数量不平衡的问题;

在算法设计上采取串联式神经网络的形式,结合 CNN 的空间特征提取能力和 BiLSTM 的双向时序建模能力,以提升模型泛化性能及复杂环境下的行为识别准确率。

1.1 数据集描述与预处理

1.1.1 数据集描述

使用的数据集源自 Vehkaoja 等^[16]在 2022 年公开发布的犬类运动传感器行为分类数据集,目前托管于 Kaggle 平台。该数据集由佩戴在犬类背部及颈部的运动传感器(即加速度计和陀螺仪)采集的测量数据组成。实验对象由 45 只健康的中大型犬组成,涵盖 27 个不同品种,年龄为 1~9 岁,体重为 13~41 kg。所有数据均采用 ActiGraph GT9X Link 惯性传感器记录,其中包含三轴加速度计与三轴陀螺仪,并以 100 Hz 的频率进行采样。实验人员在犬只颈部与背部各安装一个传感器,同步采集不同身体部位的运动信号。实验地点是一个铺设了人工草坪的室内场地,犬只在主人的引导下完成静态动作与动态动作,两类动作交替进行。随后对照视频将狗在做指定动作时的实际行为进行标记,且仅标记持续时间为 1 s 且含义清晰的动作。最后将得到的数据与传感器信息准确对应起来,以提供每只狗行为的详细且同步的描述。

1.1.2 数据集预处理

本文只保留原始数据集中能直接反映犬类行为的样本,犬只的过渡与准备性活动不在分析范围内,数据集中未标注行为类别的样本同样也被移除。最后数据集的大小为 209753 行,包含 9 种行为类别:进食、奔跑、踱步、玩耍、甩动、嗅闻、站立、拖拽、行走。不同行为下的样本分布直方图如图 1 所示,每个数据包含一个 1.0 s 的运动传感器信号和指示犬类行为的相应标签。

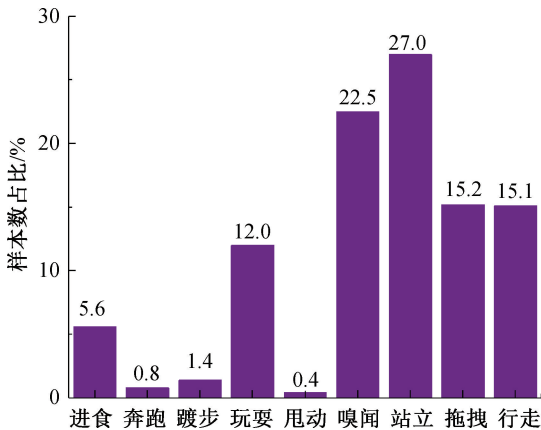


图 1 不同行为下的样本分布直方图

由于信号的采样率为 100 Hz,本文采用 1.0 s 作为固定窗口长度,时间步数为 100,并设置 0.2 s 偏移。这样保证了时间动态信息的连续性,同时增加了样本数量。此外,为了保留物理意义上的加速度与角速度比率,本文对加速度通道与陀螺仪通道分别独立缩放。对于每个原始特征 x ,RN 的变换公式如式(1)~(2)所示:

$$x_s = \frac{x - Q_2}{R_{Iq}} \quad (1)$$

$$R_{Iq} = Q_3 - Q_1 \quad (2)$$

其中: x_s 表示缩放后的特征值, Q_2 为样本中位数, R_{Iq} 为四分位数范围, Q_3 、 Q_1 分别为样本的 75% 与 25% 分位数。

该方法在存在异常值时更加稳健,并能保持不同传感器通道之间的相对幅值关系,为避免数据泄露,缩放器参数仅在训练集上拟合,并应用到测试集。

原始数据集中存在多维数据(包括犬只的生理特征、角速度和加速度),本文对月龄、体重、性别以及是否被绝育等生理特征与平均运动数据(加速度、角速度)的关系进行了探究。犬类生理特征与平均运动数据的散点气泡图如图 2 所示,从图中可以看出犬只的生理特征与其运动特性无明显的线性或者非线性的关系,这说明运动模式在个体生理层面上具有较高的相似性。因此,为提升模型的泛化能力,本文排除了静态生物维度。

考虑到在犬只活动过程中采集角速度信号所产生的噪声明显高于加速度信号,本文剔除了角速度这一维度,并选取了背部和颈部方向上的 6 轴加速度与对应的 2 个合成加速度幅值,共 8 个关键通道。这些特征对肢体躯干的机械移动最为敏感,是动物行为分析中的高相关性特征。为进一步降低通道维度的冗余性和噪声影响,本文在通道维度上应用了 PCA,且不破坏时间维度的结构,以保留完整的时间序列信息,供后续 CNN 捕捉。设一个窗口样本矩阵为 $\mathbf{X} \in \mathbf{R}^{T \times C}$,其中: T 为时间步数, C 为通道数,则 PCA 的处理步骤如下:

对每个窗口样本矩阵进行中心化处理,并计算其协方差矩阵 \mathbf{S} ,以表征各通道间的相关性,如式(3)所示:

$$\mathbf{S} = \frac{1}{N} \sum_{n=1}^N (\mathbf{X}_n - \bar{\mathbf{X}})^T (\mathbf{X}_n - \bar{\mathbf{X}}) \quad (3)$$

其中: \mathbf{X}_n 表示第 n 个样本矩阵; $\bar{\mathbf{X}}$ 为所有样本在通道维度上的均值矩阵; N 为样本总数。

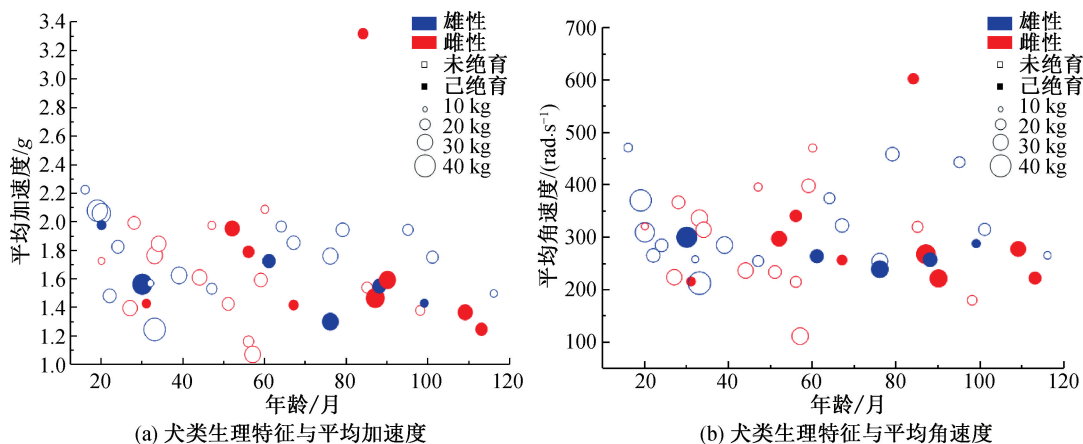


图 2 犬类生理特征与平均运动数据的散点气泡图

对协方差矩阵 S 进行特征值分解,得到特征向量矩阵与特征值对角矩阵,如式(4)所示:

$$S = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T \quad (4)$$

其中: \mathbf{V} 为特征向量矩阵; $\mathbf{\Lambda}$ 为特征值对角矩阵。

选取前 k 个最大特征值对应的特征向量构建投影矩阵,将原始数据投影至低维空间,得到降维后的特征矩阵,如式(5)所示:

$$\mathbf{Z} = \mathbf{X}\mathbf{V}_k \quad (5)$$

其中: \mathbf{V}_k 为投影矩阵; \mathbf{Z} 为降维后的特征矩阵。

特征数量与累计方差解释率点线图如图 3 所示,从图中可以看出前 5 个特征已解释超过 80% 的总方差。因此选择降维后的特征矩阵为 $\mathbf{Z} \in \mathbf{R}^{100 \times 5}$,在保证大部分信息量的情况下降低了模型参数量以及运算复杂度。

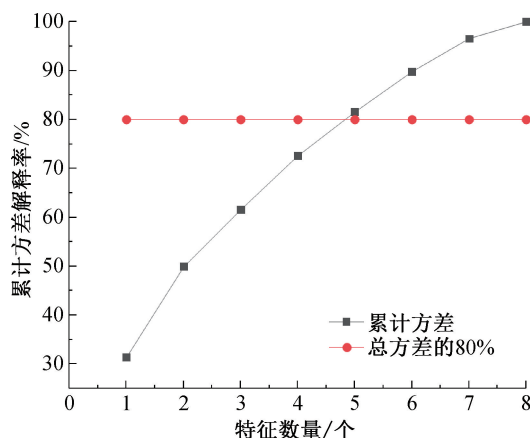


图 3 特征数量与累计方差解释率点线图

在数据集构建过程中,考虑到数据集中存在类别间样本不平衡的情况,若直接用于训练,可能导致模型倾向于判断为多数类而降低少数类的识别能力。为缓和这一问题,本文在划分数据集后将训练集进行重采样,对少数类进行过采样来增加样本数量,对多数类进行欠采样以减少样本数量。与 RN

和 PCA 相同,重采样的过程只在训练集中进行,测试集保持不变,以避免测试样本泄露,保证模型评估的公正性和泛化性。经重采样处理之后,训练集中各行为类别的样本数占比更加平衡,处理后的样本分布直方图如图 4 所示。

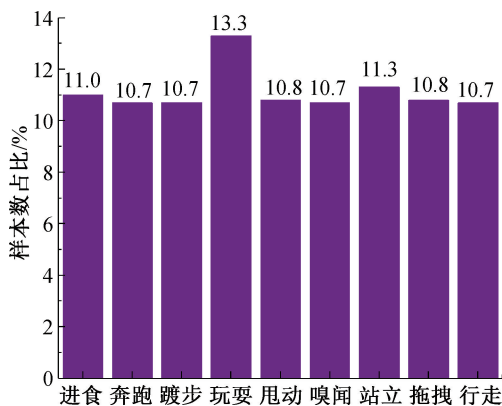


图 4 重采样处理后的样本分布直方图

1.2 本文方法设计

本文提出了一种 CNN-BiLSTM 的混合时空特征融合网络,同时考虑局部感知和全局时序推理,即在前端使用 CNN 自适应地提取传感器信号的局部空间特征,并抑制其高频噪声,在后端则采用 BiLSTM 来捕捉“过去-未来”全时域范围内的双向上下文依赖,跟踪行为序列的变化趋势。该网络通过对时空两个维度进行综合建模来提高模型在“玩耍”这类复杂不规则动作上的识别精度和稳定性。

1.2.1 CNN-BiLSTM 网络模型

针对传感器原始信号噪声干扰大及语义层次浅等问题,本文设计了一个端对端深度学习网络结构,即把行为识别分解为“空域特征编码”和“双向时序建模”。模型整体结构如图 5 所示,包括 CNN 层与 BiLSTM 层。前端采用 CNN 作为特征编码器,在高维、冗余的原始传感器数据中提取出具有平移不

变性(如震颤、瞬时冲击等)的局部波形特征,达到信号的空间降噪与抽象的效果;后端采用 BiLSTM,利用其特有的门控单元以及双向传播特性,对 CNN

提取的特征序列进行上下文推理。输出的时序特征向量通过全连接层映射到 9 类动作上,获得归一化后的预测结果。

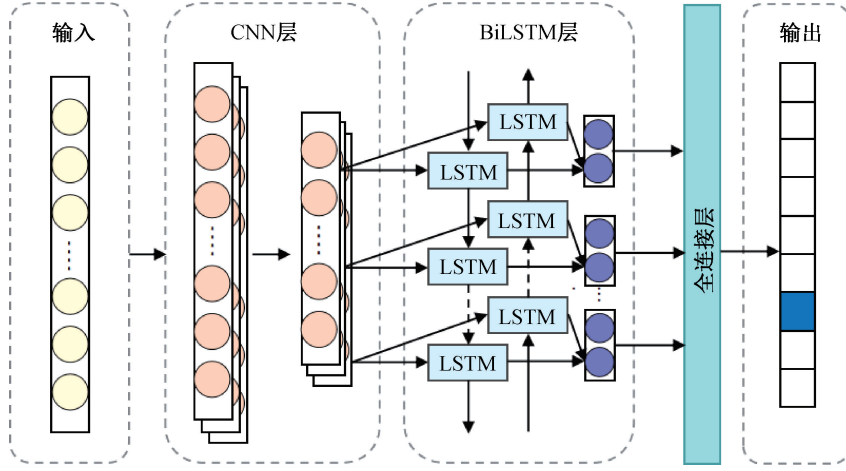


图 5 CNN-BiLSTM 网络结构图

1.2.2 卷积神经网络层

不同犬类行为在特定频段的传感器信号中,往往表现出独特的波形分布。因此本文引入 CNN 进行局部空间特性的提取,采用大感受野与降采样相结合的方式替代传统逐点输的方式,以挖掘高维传感器信号中的短时局部相关性。考虑到犬类动作存在具有一定持续宽度的物理波形(例如爪子接触地面),为捕获该物理过程的动态变化,本文首先采用卷积核大小为 7、步长为 2、填充为 3 的一维卷积层(Conv1d)。大卷积核可起到滤除高频噪声的作用,

并将原始传感信号映射到 64 维的高维空间里,以增加特征的稀疏表达能力。然后引入批归一化层(BatchNorm1d)与 ReLU 激活函数进行非线性变换,消除协变量偏移并加快网络收敛速度。最后利用核大小为 3、步长为 2、填充为 1 的最大池化层(MaxPool1d)进行二次降采样,在尽可能保留最大局部活跃度信息的基础上,进一步将序列长度缩减为输入的 1/4,从而实现从原始信号到初级语义特征的高效编码。CNN 网络架构流程图如图 6 所示。

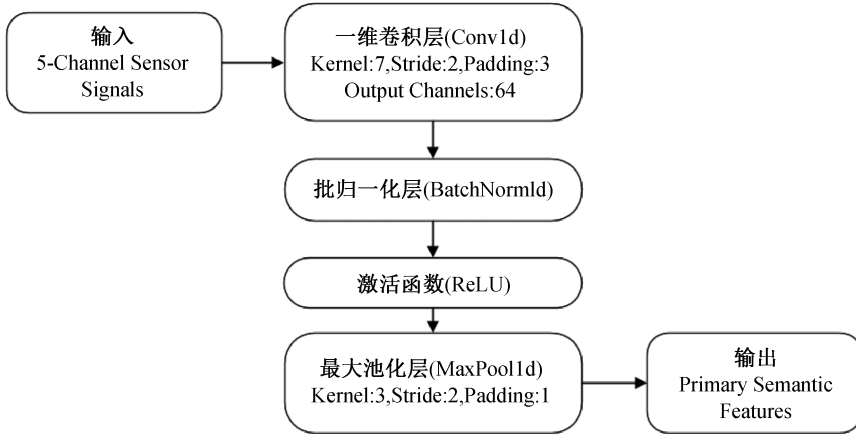


图 6 CNN 网络架构流程图

1.2.3 双向长短期记忆网络层

由于犬类行为具有时序连续性和多义性的特点,在某一时刻提取到的表观信息可能不足以确定其对应的行为动作。因此,本文采用深层堆叠的双向时序编码器来挖掘“前-后”时序的语义关系。

Bi-LSTM 网络架构示意图如图 7 所示。前向层进行前向传播,并在每一步保存前向隐藏层的输出;同时后向层进行反向传播,在每一步中储存后向隐藏层的输出。经过以上两个步骤后,将前向层与后向层存储的输出合并,得到最终输出结果。Bi-LSTM 计算公式如式(6)所示:

$$\begin{cases} \vec{h}t = f(\omega_1 x_t + \omega_2 \vec{h}t - 1) \\ \overleftarrow{h}t = f(\omega_3 x_t + \omega_4 \overleftarrow{h}t + 1) \\ y_t = g(\omega_5 \vec{h}t + \omega_6 \overleftarrow{h}t) \end{cases} \quad (6)$$

其中: $\vec{h}t$ 和 $\overleftarrow{h}t$ 分别是正向和反向 LSTM 的输出; y_t 是隐藏层的最终输出; $f(\cdot)$ 与 $g(\cdot)$ 是对应的激活函数。

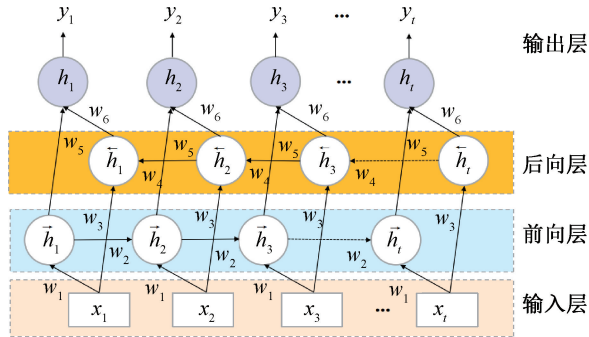


图 7 Bi-LSTM 网络架构示意图

Bi-LSTM 作为横向特征提取模块的核心,处理 CNN 维度变换后的特征序列。考虑到模型尺寸与计算效率是相互制约的关系,每层 LSTM 的隐藏单元数(Hidden size)都设定为 64。该模块使用前 2 个隐藏层从犬只行为特征数据中提取信息,并且两者都连接到同一输出层中,输出维度扩展至 128 (64×2) 维。这样可以使每个行为样本的分类同时考虑了过去和将来的行为信息,克服动作间边界不清晰的问题。另外,在 LSTM 层中使用 Dropout 来避免过拟合,设置其比例为 0.2,以抑制过拟合。最后,128 维时序特征向量经过全连接层映射至 9 类行为上,利用 Softmax 函数获得归一化后的预测结果。

2 实验设计与结果分析

2.1 实验设计

2.1.1 实验环境

实验所用的操作系统是 Windows 10 专业版(64 位),深度学习框架是 PyTorch 1.10,CPU 是 Intel(R) Core(TM) i5-10400F,2.90 GHz, RAM 为 16.0 GiB,GPU 是 NVIDIA GeForce RTX 3060,显存 12 GiB,编程语言为 Python 3.10,软件运行环境为 Pycharm。在本文中,从运动传感器的时间序列数据建立了预处理过的数据集,共包括 209753 个样本。将这些数据严格地分成训练集和测试集,比例为 8:2,其中:训练集中有 167802 个样本,在训练集上对模型进行迭代学习;测试集中有 41951 个样本,用于最终性能评估。

2.1.2 实验参数设置

本文使用 Adam 优化器中的自适应下降算法对网络权值进行迭代,综合考虑 GPU 显存限制及运算性能,批处理量(Batch size)设为 128,以达到较好的训练稳定性与收敛速度。由于本文为分类问题,故选用交叉熵损失函数来指导网络的学习。其中,设置初始学习率为 0.002,网络训练迭代次数为 200 Epoch。为了避免模型在训练集上过拟合,并且不丢失最优参数,在每一轮训练结束之后立即在测试集上进行一次前向运算,得到当前的准确率与损失值,仅当准确率高于以往的最佳纪录时,才把当前的模型参数作为最后测试模型的参数,以保证最终模型具有最优的泛化性能。

2.1.3 实验评价指标

本文选用的算法评价指标有精确率(Precision)、召回率(Recall)、F1 分数(F1 Score)、准确率(Accuracy),其公式分别如式(7)–(10)所示:

$$P_p = \frac{T_p}{T_p + F_p} \quad (7)$$

$$P_r = \frac{T_p}{T_p + F_N} \quad (8)$$

$$P_f = \frac{2 \times T_p}{2 \times T_p + F_N + F_p} \quad (9)$$

$$P_a = \frac{T_p + T_N}{T_p + T_N + F_N + F_p} \quad (10)$$

其中: P_p 为精确率; P_r 为召回率; P_f 为 F1 分数; P_a 为准确率; T_p (True positive) 为模型正确预测为正例的样本数; F_p (False positive) 为模型错误预测为正例的样本数; T_N (True negative) 为模型正确预测为负例的样本数; F_N (False negative) 为模型错误预测为负例的样本数。

以上指标分别从查准、查全、平衡以及全局判别正确率等方面对模型进行评价,以避免在类别不平衡情形下因依赖单一指标得出的片面结果。

2.2 实验结果

对本文中实验涉及的所有犬类行为分类模型进行总体性能分析,决策树(Decision tree, DT)、RF、CNN、CNN-LSTM 和 CNN-BiLSTM 的性能如表 1 所示。从表 1 中可以看出:本文提出的 CNN-BiLSTM 模型在测试集上达到了 95.5% 的准确率与 97.2% 的 F1 分数,并且各评估指标都优于其他模型;仅利用纯卷积网络提取特征效果较差(F1 分数为 90.5%),加入单方向时间信息后(LSTM),结果有所提升(F1 分数为 95.1%),而本文所提模型

考虑了双向时间信息,进一步提升了识别精度。这也证明犬类行为识别不仅仅取决于某一帧的姿态信息,更大程度依赖于行为的时序发展规律。

表 1 犬类行为预测模型的总体性能评价指标 %

模型	准确率	精确率	召回率	F1 分数
DT	67.3	74.1	72.4	72.6
RF	73.6	80.5	77.6	78.5
CNN	84.6	94.1	87.5	90.5
CNN-LSTM	93.2	95.6	92.7	95.1
本文模型	95.5	98.5	96.1	97.2

注:加粗字体代表当前评价指标的最好结果。

各模型预测效果最佳与最差的行为特征如表 2 所示,该表汇总了各模型识别效果最好的前两名行为及对应的 F1 分数、识别效果最差的后两名行为及对应的 F1 分数。不同模型结构对于不同行为类型的敏感程度存在差别,如所有模型都容易识别“甩

表 2 各模型预测效果最佳与最差的行为

模型	最佳特征	次佳特征	最差特征	次差特征
DT	甩动 (96.2%)	嗅闻 (85.4%)	踱步 (44.3%)	进食 (46.1%)
RF	甩动 (96.4%)	嗅闻 (93.8%)	踱步 (31.3%)	进食 (70.1%)
CNN	甩动 (98.1%)	嗅闻 (96.3%)	玩耍 (37.9%)	站立 (67.4%)
CNN-LSTM	甩动 (98.3%)	嗅闻 (99.2%)	玩耍 (60.2%)	进食 (86.2%)
本文模型	甩动 (100.0%)	嗅闻 (99.4%)	玩耍 (86.2%)	进食 (87.8%)

注:括号中的数值为对应行为类别预测结果的 F1 分数。

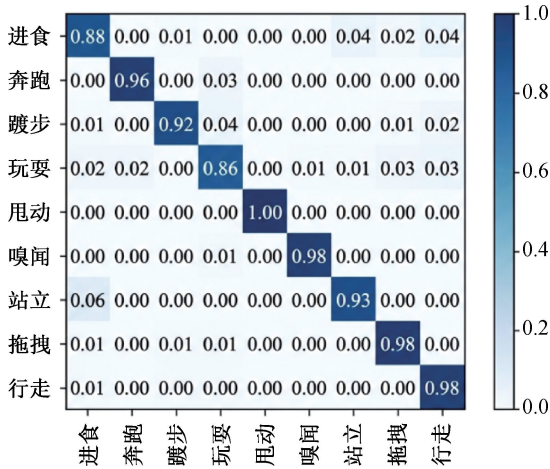


图 8 基于 CNN-BiLSTM 的犬类行为识别混淆矩阵图

本文设置了 4 组实验以确定 RN、PCA 以及重采样对犬类行为识别效果的单独影响,且均采用相同的 CNN-BiLSTM 混合架构,实验结果如表 3 所示。从表 3 的方案 1 与方案 2 可以看出,在引入 PCA 进行降维后,准确率提升了 1.6%,这说明经过 PCA 降维得到的前 5 个主要通道能够降低不同传感器之间的耦合噪声,突出最能代表行为特征的维度信息。方案 3 在方案 1 的基础上增加了重采样,

而较难识别出“玩耍”和“踱步”等行为。

不同模型对于识别“玩耍”这一复杂行为的表现效果差异明显。从表 2 可以看出,在 CNN 对“玩耍”这一类别的 F1 分数只有 37.9%,表现最差。这是因为“玩耍”具有高度的非平稳性,并且含义多样化,它往往是奔跑、跳跃、拉扯等一系列子动作穿插而成的行为。而对 CNN 模型,它极容易根据局部窗口内特征的加速度值,片面地将玩耍过程的加速阶段错识别成“奔跑”,将停顿阶段错识别成“站立”。相比 CNN,本文模型将“玩耍”的 F1 分数提升了 48.3%,对比 CNN-LSTM,提升了 26.0%。这也验证了双向机制的核心作用,模型可以挖掘历史与未来信息,把碎片化的动作片段“跑”“停”“跳”建立成连续的“玩耍”行为。混淆矩阵图如图 8 所示,从图中可以看出本文模型减少了“玩耍”与“奔跑/行走”之间的混淆。

使得样本各行为类别更加平衡,F1 分数增加了 2.1%,因为原数据集中的“奔跑”“踱步”等行为类别样本量少,重采样避免模型对大多数类别的过拟合,并提高对小类别的识别能力。方案 4(本文最终方案)结合了统计降噪与分布平衡化处理,达到了最优的准确率(95.5%)与 F1 分数(97.2%)。实验证明,预处理的每个环节对于构建稳健的犬类行为识别模型均具有重要作用。

表 3 预处理模块有效性消融实验结果 %

方案	预处理设置	准确率	F1 分数
1	基准模型(仅 RN)	91.2	92.5
2	基准+PCA(降噪与降维)	92.8	93.9
3	基准+重采样(处理类别失衡)	93.5	94.6
4	本文完整方案(PCA+重采样)	95.5	97.2

本文在同一硬件环境下比较了 CNN-LSTM 与本文 CNN-BiLSTM 模型的复杂度与性能,以验证双向时序建模方法在资源消耗与性能提升上的关系,具体数据如表 4 所示。从表 4 中可以看出:在模型参数量增加至 0.16 MiB 和 FLOPs 为 7.80 M 的情况下,单样本延迟时间仅增加了 1.1 ms,仍能满足可穿戴设备的实时处理需求;在保证实时

性的前提下,“玩耍”行为的 F1 分数提升了 26.0%,说明本文模型能够有效地克服单向模型

在复杂行为上存在上下文盲区的问题,有较高的实际应用价值。

表4 单向与双向 LSTM 模型复杂度与性能

模型	参数量/MiB	单轮训练耗时/s	延迟时间/ms	FLOPs/M	玩耍行为 F1 分数/%
CNN-LSTM	0.07	18	3.6	4.2	60.2
CNN-BiLSTM	0.16	30	4.7	7.8	86.2

2.3 与现有研究结果对比

本文模型与其他犬类行为识别方法在数据集、行为类别数以及网络结构进行了对比,结果如表5所示。本文所选数据来源于 Kaggle 上公开的数据集,数据质量得以保证的同时,且方便方法的复现以及与其他研究的对比分析。从上述方法来看,Hussain 等^[14]采用的深层 LSTM 网络在处理 10 类行为的准确率为 94.3%,说明时间序列分析对于动物动作分类有一定的效果,但单向的网络结构无法预知未来信息,从而对判断进行修正。在多数研究仅开展少量行为类别的情况下,本文模型在 9 种犬类行为的识别任务中,依然达到了 95.5%的最优识

别准确率。随着类别数量的增加,特征空间边界的重合率以指数级递增,且对于区分生物力学特征高度相似的行为(如“行走”与“踱步”)而言难度更大,模型性能通常会下降。

本文在公开数据集上构建 CNN-BiLSTM 模型,有效融合了局部特征提取与双向时序建模能力,在 9 类犬类行为识别任务中取得了 95.5%的准确率。相较于单向 LSTM 方法,本文构建了双向时序依赖模型,增强了对行为序列前后关联的利用,准确率较现有最优结果提高了 1.2%。这表明本文模型在复杂动物行为识别任务中的有效性与先进性,同时也体现了采用公开数据促进了研究的可比性与复现性。

表5 本文模型与现有研究的综合情况

研究来源	数据集	行为类别数/种	方法主要模块	准确率/%
Amano 等 ^[13]	自有数据集	5	CNN(2 卷积+2 池化+2Dropout+1 展平+3 全连接)	92.6
Kumpulainen 等 ^[8]	公开数据集 ^[16]	7	SVM	91.4
Muminov 等 ^[10]	自有数据集	6	GNB	88.0
Hussain 等 ^[14]	自有数据集	10	LSTM(6LSTM+3 Dropout+3 全连接)	94.3
Marcato 等 ^[9]	自有数据集	5	RF	90.0
本文模型	公开数据集 ^[16]	9	CNN-BiLSTM	95.5

3 结论

本文设计了一种基于数据增强与 CNN-BiLSTM 混合架构的犬类行为识别方法,主要研究结论如下:

a) 构建了包含 RN、PCA 及重采样的数据处理策略。实验表明,该策略能消除传感器数据异常值,降低特征维度及平衡样本分布,为模型提供了高质量的输入数据。

b) 设计了 CNN-BiLSTM 混合识别方法。利用 CNN 提取传感器信号局局部空间特征,以及引入 BiLSTM 构建双向时序依赖模型,解决了单向模型在处理复杂行为时存在的上下文缺失问题。

c) 验证了本文模型的有效性与先进性。实验结果显示,该模型的识别准确率达到 95.5%,F1 分数达到 97.2%。特别是在高复杂度的“玩耍”行为识别上,相较于单向 CNN-LSTM 模型,F1 分数提升了 26.0%,证明了双向时序建模在复杂行为识别

中的关键作用。

本文提出的模型虽整体表现优异,但在识别“踱步”与“行走”等高相似度行为时仍有少量混淆,后续研究可考虑引入注意力机制,以提高对局部信息的关注度,从而获得更丰富的特征信息。

参考文献:

- [1] Deng R, Zhou G, Tang L, et al. E-DOCRNet: a multi-feature fusion network for dog bark identification [J]. Applied Acoustics, 2024, 220: 109950.
- [2] 刘艳秋,宣传忠,武佩,等. 基于 K-means-BP 神经网络的舍饲环境母羊产前运动行为分类识别[J]. 中国农业大学学报,2021,26(3): 86-95.
- [3] Väätäjä H, Majaranta P, Isokoski P, et al. Happy dogs and happy owners: Using dog activity monitoring technology in everyday life [C] // Proceedings of the 5th International Conference on Animal-Computer Interaction. Atlanta Georgia USA. ACM, 2018: 1-12.
- [4] 李玲. 宠物狗可穿戴设备产品商业模式研究[D]. 北京:北京邮电大学,2017: 13-17.

- [5] Ladha C, Hammerla N, Hughes E, et al. Dog's life: wearable activity recognition for dogs[C]//Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing. Zurich Switzerland. ACM, 2013: 415-418.
- [6] Gerencsér L, Vásárhelyi G, Nagy M, et al. Identification of behaviour in freely moving dogs (*Canis familiaris*) using inertial sensors[J]. *PLoS One*, 2013, 8(10): e77814.
- [7] Davoulos G, Lalakou I, Hatzilygeroudis I. Recognition of dog motion states: Ensemble vs deep learning models[C]//2024 15th International Conference on Information, Intelligence, Systems & Applications (IISA). July 17-19, 2024, Chania Crete, Greece. IEEE, 2024: 1-8.
- [8] Kumpulainen P, Cardó A V, Somppi S, et al. Dog behaviour classification with movement sensors placed on the harness and the collar[J]. *Applied Animal Behaviour Science*, 2021, 241: 105393.
- [9] Marcato M, Tedesco S, O'Mahony C, et al. Machine learning based canine posture estimation using inertial data[J]. *PLoS One*, 2023, 18(6): e0286311.
- [10] Muminov A, Mukhiddinov M, Cho. Enhanced classification of dog activities with quaternion-based fusion approach on high-dimensional raw data from wearable sensors[J]. *Sensors*, 2022, 22(23): 9471.
- [11] 翟明欣. 基于三维卷积神经网络的奶山羊行为识别方法研究[D]. 杨凌: 西北农林科技大学, 2023: 28-31.
- [12] 李晓莉, 韩鹏, 李晓光. 基于典型样本的卷积神经网络技术[J]. *计算机工程与设计*, 2020, 41(4): 1113-1117.
- [13] Amano R, Ma J. Recognition and change point detection of dogs' activities of daily living using wearable devices[C]//2021 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech). October 25-28, 2021. AB, Canada. IEEE, 2021: 693-699.
- [14] Hussain A, Begum K, Armand T P T, et al. Long short-term memory (LSTM)-based dog activity detection using accelerometer and gyroscope[J]. *Applied Sciences*, 2022, 12(19): 9427.
- [15] Kim J, Moon N. Dog behavior recognition based on multimodal data from a camera and wearable device [J]. *Applied Sciences*, 2022, 12(6) 3199.
- [16] Vehkaoja A, Somppi S, Törnqvist H, et al. Description of movement sensor dataset for dog behavior classification[J]. *Data in Brief*, 2022, 40: 107822.

(责任编辑:康 锋)