doi:10.6041/j.issn.1000-1298.2026.23.039

基于惯性动作捕捉的拖拉机驾驶员姿态识别模型研究

张景媛　金晓萍　陈景瑞　张睿妍
（中国农业大学工学院，北京 100083）

摘要：本文基于惯性动作捕捉数据，系统研究拖拉机驾驶员上半身姿态与识别模型。通过采集拖拉机驾驶员作业全过程中的上半身关节角数据，结合用户旅程图调研，提取了能够完整反映驾驶员作业状态的特征标签。基于互信息量计算、人体肌骨功能、姿态差异性，筛选出具有识别价值的关键特征。采用多层感知机验证特征筛选有效性，建立性能基准。随后，建立基于 Transformer 框架的姿态识别模型，对多标签的多类别识别效果良好，最终各标签宏平均 F1 分数在 88% 至 99% 之间。研究结果表明，所提方法能够高效、准确地识别拖拉机驾驶员上半身姿态，为数字孪生虚拟体的构建提供了有效技术方案，为拖拉机驾驶安全监测及智能辅助提供了新的技术路径。

关键词：拖拉机驾驶员；惯性动作捕捉；姿态识别；Transformer 模型
中图分类号：TP312　　**文献标识码**：A　　**文章编号**：1000-1298(2026)23-0418-09　　**OSID**：

Identification Model of Tractor Driver Posture Based on Inertial Motion Capture

Zhang Jingyuan　Jin Xiaoping　Chen Jingrui　Zhang Ruiyan
(*College of Engineering*, *China Agricultural University*, *Beijing* 100083, *China*)

Abstract: It systematically identified and characterized the upper body postures of tractor drivers by constructing corresponding digital twin virtual models based on high-fidelity motion capture data. Throughout the entire operation process, inertial sensor data were continuously collected from drivers and processed by using motion capture software to extract precise joint angle information of the upper body. These biomechanical data were integrated with the drivers' travel trajectory maps to define a comprehensive set of feature labels, enabling a holistic representation of the driver's operational state. To enhance the accuracy and robustness of posture recognition, feature selection was rigorously performed by analyzing mutual information metrics combined with biomechanical function assessments and posture variability evaluations. This process identified key features most relevant to distinguishing complex postural patterns. The effectiveness of the selected features was validated by using a multi-layer perceptron (MLP) model, which established a baseline performance benchmark. The final posture recognition model developed was built upon the Transformer framework. It utilized multi-dimensional temporal features as input and employed a hybrid training strategy combining multi-label joint training with single-label learning. As a result, this model demonstrated comprehensive improvements over the MLP, exhibiting superior and robust posture recognition capabilities. The final recognition system achieved macro F1 scores ranging from 88% to 99% across various posture labels, demonstrating high accuracy, robustness, and stability in real-world conditions. These results confirmed that the proposed methodology effectively captured and discriminated the complex upper body postures of tractor drivers using inertial motion capture data. Moreover, the constructed digital twin virtual body served as a practical and scalable framework for real-time posture monitoring and ergonomic analysis. This approach held considerable promise for advancing safety supervision, enhancing intelligent assistance systems, and ultimately improving the operational safety and comfort of tractor driving in agricultural settings.

Key words: tractor driver; inertial motion capture; posture recognition; Transformer model

收稿日期：2025-09-08　修回日期：2025-10-09
基金项目：军队科研计划项目(202004610711099)
作者简介：张景媛(2001—)，女，硕士生，主要从事人体姿态、人的生理状态与行为研究，E-mail：zjyrl@outlook.com
通信作者：金晓萍(1979—)，女，副教授，博士生导师，主要从事智能驾驶、农业装备等人机系统交互理论研究，E-mail：jinxp@cau.edu.cn

0　引言

数字孪生一般被认为是物理世界的实体、虚拟世界的模型、实体与虚拟体的相互作用关系的集合[1-5]，近年来，数字孪生技术在多个领域展现出广泛的应用潜力[6-7]。作为数字孪生技术的重要分支，人体数字孪生的发展基于传统人体模型的研究，其核心区别在于数字孪生更强调虚拟模型与现实人体之间的交互能力，从而实现对人体状态的动态感知、分析与反馈[5,8-10]。该技术在健康监测、行为分析及人机交互等领域已显示出显著的应用价值[11-14]。

农业作为一个高度复杂且多变的研究领域，近年来数字孪生技术的引入极大促进了农业生产的智能化和精准化发展[15-18]。农业数字孪生通过对农业环境、农机设备及作业过程的全面建模与动态控制，有效提升了农业生产效率和管理水平[17,19]。然而，作为农业作业中关键执行者的拖拉机驾驶员，其数字孪生体的研究尚处于起步阶段[20]，相关研究较为匮乏。

在相关领域中，汽车驾驶员数字孪生体的研究相对成熟，涵盖了驾驶员姿态识别、行为监测及状态评估等多个方面，为提升驾驶安全及辅助驾驶系统的性能提供了重要支持[21-25]。相比之下，拖拉机驾驶不仅涉及车辆行驶的操控，还需操作农机具，导致驾驶员的生物力学变化和姿态表现更加复杂且多样[26-28]，尤其是上半身更明显。这种差异不仅增加了对拖拉机驾驶员状态感知的难度，也凸显了基于人体数字孪生技术开展拖拉机驾驶员虚拟体建模的必要性和现实意义。

基于此，本研究聚焦于拖拉机驾驶员上半身姿态，旨在通过对头部、上肢及上躯干等关键部位的精确建模与识别，提升对驾驶员状态的理解与监测能力，为虚拟体与实体的交互提供坚实的理论和技术基础。姿态作为人体动态表现的重要组成部分，能够反映驾驶员的动作特征和操作习惯[22,26-27]，是实现虚拟体交互及驾驶行为分析的核心要素[28]。通过深入的姿态研究，可有效捕捉驾驶员动作细节，为后续的安全预警和辅助决策提供科学依据[29]。

1　实验设计

1.1　实验方法

(1)实验目的：获取拖拉机驾驶员在作业过程中上半身的关节角数据，用于姿态识别与拖拉机驾驶员上半身虚拟模型的构建。

(2)实验任务：操纵模拟器，控制拖拉机在环形道路上完成播种作业及刮浆作业的完整过程，作业过程中需要及时观察环境并调整农机具。

(3)实验被试：共 32 名，身高均值为 170.7 cm，身体质量指数 BMI 均值为 23.8 kg/m^2。被试均已取得汽车驾驶证或拖拉机驾驶证，无肌骨损伤等疾病，被试的拖拉机熟练度为：有拖拉机的田间驾驶经验或已通过拖拉机驾驶实习专业课程。

(4)实验设备：拖拉机驾驶模拟器、惯性动作捕捉、摄像机。拖拉机驾驶模拟器具有驾驶舱的实体硬件以及拖拉机作业的虚拟环境，虚拟环境模拟了拖拉机驾驶真实场景，具有拖拉机驾驶的声音反馈与触觉反馈，如图 1 所示。XSENS 惯性动作捕捉，具有 17 个身体部位的传感器，能够覆盖驾驶员头部、颈部、上躯干、上肢和手的数据采集，如图 2 所示。摄像机用于获取被试的全流程动作视频。



图 1　拖拉机驾驶模拟器

Fig. 1　Tractor driving simulator

1. 驾驶员视角虚拟图像　2. 方向盘　3. 方向杆　4. 挡杆　5. 钥匙　6. 控制农机具旋钮　7. 控制农机具液压杆

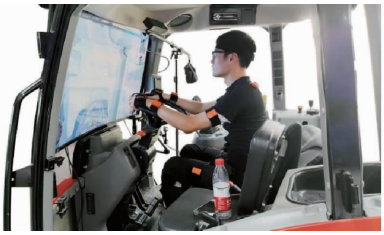

图 2　实验实景

Fig. 2　Experimental scene

(5)实验流程：①被试熟悉实验内容及设备，签署知情同意书。②被试穿戴惯性动捕数据采集装置，在相同初始方向下完成校正。③开始采集，被试启动机器。④被试分别完成一次播种作业与一次刮浆作业，操纵行为包括车辆方向控制、农机具收放、农机具位置微调、观察作业情况等姿态。⑤播种作业时控制拖拉机直线前进，行进至尽头时掉头，并继续直线作业至完成播种。⑥刮浆作业时控制拖拉机螺旋式曲线前进，行进至螺旋中心停止作业。⑦作业间歇拿起水瓶饮水、拿起手机接打电话，模拟真实的非工作状态。⑧完成作业过程，停止采集。

1.2　原始数据及预处理

为提取有效的作业姿态区间并消除数据中的异常值，建立拖拉机驾驶员身体上半身在作业过程中

的姿态数据集。

动捕设备采集的原始传感器数据在仪器配套的 XSENS analyze 软件中完成慢速、重置零点、重置盆骨的再处理以获取更准确的数据。慢速利于获取更精准的数据,重置零点与盆骨便于统一被试在空间中与驾驶舱的相对位置。由于该实验中左脚移动较少,所以选择以左脚脚跟为重置后的零点。处理后导出能够表征人体姿态的关节角数据。

再处理后导出的数据为一个实验流程中的所有数据,为了获取不同姿态对应的数据片段,根据实时录像分割时间轴,并对相应时间序列的数据进行标记。拧钥匙此类姿态完成较快,约 0.5 s,所以分割时间轴的精度选择 1 帧,录像的帧率为 60 f/s,即精度为 1/60 s,满足需求。动捕设备的采样率为 60 Hz,1 s 中的数据量满足精度。

对于数据流中可能出现的短暂缺失,根据上下文采用前向填充、后向填充或插值法进行补充,确保数据序列的连续性。

由于数据具有时序性,所以使用滑动窗口来更好地捕捉特征。滑动窗口将数据切分为不同的时间段,每个窗口代表一个时间序列的片段。滑动窗口能够从连续的数据中提取固定长度的子序列,这些子序列在时间轴上移动,帮助模型捕捉动态变化的特征。

驾驶员姿态产生过程中时间最长的为一次方向盘旋转,最短的为拧一次钥匙,时间区间为 0.5 ~ 3 s,基于此,在采样率为 60 Hz 的情况下,时间窗口的大小选择 5,步长为 1,使时间窗口能够完整的包含姿态的时序特征。取该时间窗口的均值来平滑数据。

因为不同人体的骨骼尺寸不同、肌肉状态不同,即使是同一驾驶环境下也会产生不同的关节角度变化,直接使用这些特征会导致模型过于关注量级大的特征而忽略量级小但重要的特征。Z-score 标准化将每个特征的均值变为 0,标准差变为 1,因此它可以消除特征之间尺度差异的影响。为降低不同人体之间差异产生的影响,对每一个完整动作进行标准化处理,将动作处理至同一尺度下,同时保留一个动作的分布规律。同时,标准化后所有特征尺度变为一致,避免了量级差异对统计结果的干扰,更利于分析各特征在不同组别间的分布差异。标准化的计算如下:

设原始标签数据集为 $y=\{y_1,y_2,\cdots,y_n\}$,其中 n 为样本量,标准化标签 z_i 定义为

$$z_i=\frac{y_i-\mu_y}{\sigma_y}\quad(i=1,2,\cdots,n) \tag{1}$$

其中

$$\sigma_y=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\mu_y)^2} \tag{2}$$

$$\mu_y=\frac{1}{n}\sum_{i=1}^{n}y_i \tag{3}$$

式中 σ_y——标签样本标准差

μ_y——标签样本均值

1.3 标签选择

为了更好地将物理世界的拖拉机驾驶员映射到虚拟世界,从而搭建拖拉机驾驶员身体上半部分的虚拟体,应遵循被识别的姿态尽可能完整地表现拖拉机驾驶员作业的全流程与驾驶员状态的原则。根据该原则开展调研,进行用户访谈与用户旅程地图[30]绘制。

与拖拉机驾驶员进行一对一的用户访谈,模式使用半结构式,访谈目的是获取拖拉机驾驶员视角下的作业流程与重点。由于播种与刮浆两种作业场景覆盖了直线前进与螺旋前进的作业路径并且具有精细化操作的特点,其对应的驾驶员姿态具有广度与深度,所以预设这两种场景开展调研。具体内容包括:播种作业的工作流程、刮浆作业的工作流程、高频率发生的姿态、影响作业效率的姿态、有较大负荷的姿态、需要其他人与驾驶员进行交互的情况。用户旅程图如图 3 所示。

根据用户旅程分析获得的姿态重点,以交互主体分类,总结获取的标签如表 1 所示。

2 特征选择

2.1 原始特征提取

关节角作为表征人体姿态的重要特征,能够捕捉身体各节段之间的相对位置关系;通过对连续帧关节角变化的分析,还可进一步反映各节段的运动关系。基于对各关节角度的量化,关节角为刻画复杂的姿态模式提供了一种客观、科学的描述方式。此外,关节角对环境变化和个体差异具有良好的适应性,因此作为模型输入特征,能够显著提升姿态识别和动作分析的准确性与可靠性。在此前提下,根据人体肌骨的基本功能与采集获取的数据,进行特征的初次筛选。

考察的姿态发生在拖拉机驾驶员的上半身,包括头部、颈椎、胸椎、腰椎、肩部、上臂、下臂、手掌。这些身体节段对应的旋转中心点分别是 C1、T1、T9、L1、L4、T4、肩、肘、腕,如图 4 所示。经过用户访谈、用户旅程的前期调研,确定在一个任务流程中,拖拉机驾驶员上半身姿态包括:头部和颈部的扭头;牵涉腰椎和胸椎旋转的推/拉液压杆(右手)、拧旋钮(右手)、接电话、喝水;牵涉胸椎和肩部旋转的扶挡杆

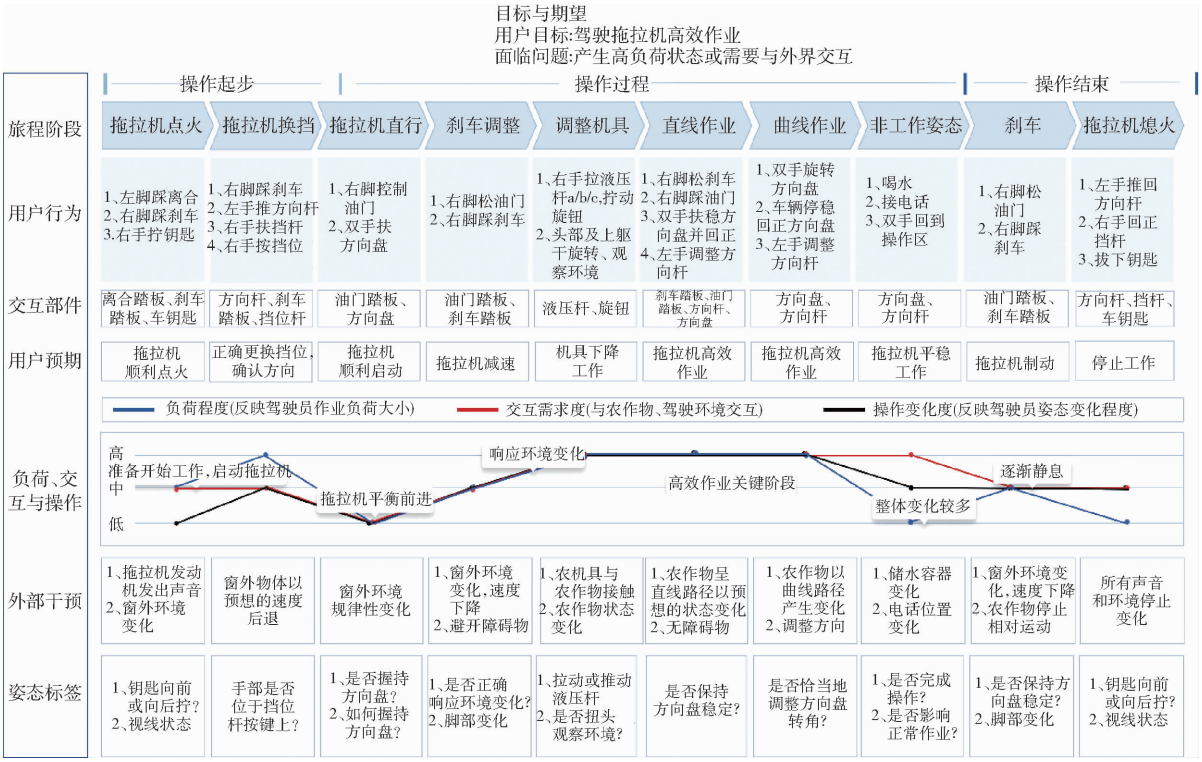

图 3　用户旅程地图

Fig. 3　User journey mapping

表 1　姿态标签及其编号

Tab. 1　Posture labels and numbers

编号	A	B	C	D	E	F	G	H	I	J
姿态标签	钥匙	方向杆	挡杆	方向盘(右手)	方向盘(左手)	液压杆	扭头	旋钮	接电话	喝水
姿态变化	前/后	推/拉	按压	顺/逆时针	顺/逆时针	推/拉	扭头	拧动	接电话	喝水

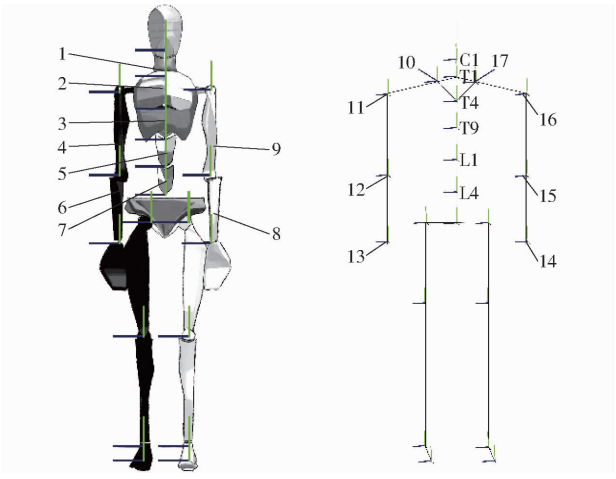

图 4　身体节段与旋转点

Fig. 4　Body segments and rotation point

1. 颈椎　2. 肩部　3. 胸椎　4. 右上臂　5. 腰胸　6. 右下臂　7. 腰椎　8. 左下臂　9. 左上臂　10. 右肩点　11. 右肩　12. 右肘　13. 右手腕　14. 左手腕　15. 左肘　16. 左肩　17. 左肩点

(右手)、顺时针旋转方向盘、逆时针旋转方向盘;牵涉肩部旋转和手部的推拉方向杆(左手);手部的拧钥匙(右手)。

坐标中心、身体部位、上半身身体节段的对应关系如表 2 所示。导出的角度数据中,与上肢动作相关的身体节段有 L4L3、L1T12、T9T8、T1C7、C1 头、T4 右肩、右肩、右肘、右手腕、T4 左肩、左肩、左肘、左手腕,每个身体节段在空间直角坐标系下进行 3 个方向的旋转,坐标原点位于身体节段的一端。

2.2　特征筛选

特征集合完整地包括了上半身身体节段的变化,但是待识别的标签与这些特征并非完全对应,部分特征的变化与某一特定标签的关联性较差,如标签扭头与手部特征没有必然联系。为了剔除高维数据中的冗余特征并识别出与预测目标强相关的关键特征,降低维度灾难风险、优化计算效率,以及为后续建模提供可靠的输入基础,进行特征筛选。

该研究引入互信息量作为特征筛选的依据。相较于皮尔逊相关系数等线性度量,互信息量基于信息熵理论,能够获取任意函数关系[31],适用于关系尚不明晰的身体节段角度与标签,并且对数据尺度不敏感,稳健性好。其概率分布的本质确保了它作为特征筛选依据的可行性。

表 2 坐标中心与身体节段

Tab. 2 Coordinate centers and body segments

坐标中心	L4	L1	T9	T1	C1	T4/肩	肩	肘	腕
关联部位	腰椎	腰胸	胸椎	后颈	脖颈	右肩胛骨	右肩	右上臂	右下臂
身体节段	L4L3	L1T12	T9T8	T1C7	C1 头	T4 右肩	右肩	右肘	右手腕

注:该表格以身体右侧为例,左侧与右侧结构对称。

具体计算方法为:在身体节段角度的数据中根据姿态的时间出入点为标签列赋值,出现标签所表示姿态的时间段赋值 1,未出现的时间段赋值 0,进行互信息量的计算。

连续型特征 X 与标签 Y 的互信息量为

$$I(X;Y)=\iint_{\mathbb{R}^2} f(x,y)\log_2\left(\frac{f(x,y)}{f(x)f(y)}\right)\mathrm{d}x\mathrm{d}y \quad (4)$$

式中 $f(x,y)$——X、Y 的联合概率密度函数

$f(x)$、$f(y)$——X、Y 的边缘概率密度函数

$I(X;Y)$——X、Y 之间的互信息量

由于互信息量的值只表示信息的对应关系,在数值上并没有评判相关性的明确阈值,所以在同一个标签中使用互信息量的相对值评价特征的有效性,当某一特征的互信息量大于或等于与标签紧密相关的特征时,认为该特征的数据与标签具有强相关性。特征与标签的紧密相关性利用人体肌骨功能及姿态的差异性判断。如拧钥匙是右手腕、右肘、上躯干的联合活动过程,虽然拧钥匙发生前的动作会影响各身体节段的活动程度,但是右手腕的旋转是决定拧钥匙发生这一时刻的关键变化,活动程度永不为 0,所以拧钥匙对应的互信息量阈值为右手腕前旋/后旋(P/S,缩写见表 3)互信息量的值。与之类似,标签方向杆对应 T4 左肩内屈/外伸(F/E),标签挡杆对应右肩外展/内收(Ab/Ad),方向盘(右手)对应右肘内屈/外伸(F/E)与右手腕前旋/后旋(P/S)的更高值,方向盘(左手)对应左肘前旋/后旋和左手腕前旋/后旋的更高值,液压杆对应右肩内旋/外旋(IR/ER),扭头对应 T1C7 绕轴旋转,旋钮对应右肘尺侧/桡侧偏移(UD/RD)。标签 A ~ H 特征对应的互信息量的值如表 4 ~ 11 所示。其中表示方向的名称使用如表 3 所示的缩写。

表 3 旋转方向及缩写符号

Tab. 3 Rotation direction and abbreviation symbols

方向	绕轴自转	侧向弯曲	前旋/后旋	内屈/外伸	内旋/外旋	外展/内收	尺侧/桡侧偏移
缩写	AR	LB	P/S	F/E	IR/ER	Ab/Ad	UD/RD

表 4 标签 A 拧钥匙的特征及其互信息量

Tab. 4 Characteristics of label A and value of its mutual information quantity

项目	T1C7 AR	C1 头 AR	L1T12 LB	L4L3 LB	右手腕 P/S
互信息量	0.029 3	0.026 3	0.024 2	0.023 9	0.022 6

表 5 标签 B 方向杆的特征及其互信息量

Tab. 5 Characteristics of label B and value of its mutual information quantity

项目	L4L3 LB	L1T12 LB	T9T8 LB	左手腕 P/S	左 T4 肩 F/E
互信息量	0.035 2	0.035 1	0.033 3	0.020 5	0.019 7

表 6 标签 C 挡杆的特征及其互信息量

Tab. 6 Characteristics of label C and value of its mutual information quantity

项目	右肩			C1 头	T1C7	T9T8	右肘
	IR/ER	F/E	Ab/Ad	AR	AR	F/E	F/E
互信息量	0.018 0	0.013 8	0.010 5	0.016 3	0.016 2	0.015 2	0.010 6

表 7 标签 D 方向盘(右手)的特征及其互信息量

Tab. 7 Characteristics of label D and value of its mutual information quantity

项目	右肩		L4L3	L1T12	T9T8	右肘
	IR/ER	F/E	F/E	F/E	F/E	F/E
互信息量	0.045 6	0.043 2	0.042 8	0.042 6	0.041 8	0.037 1

表 8 标签 E 方向盘(左手)的特征及其互信息量

Tab. 8 Characteristics of label E and value of its mutual information quantity

项目	L4L3 F/E	L1T12 F/E	T9T8 F/E	T1C7 AR	左手腕 P/S	左肘 P/S
互信息量	0.032 7	0.032 7	0.032 3	0.025 3	0.024 3	0.023 8

表 9 标签 F 液压杆的特征及其互信息量

Tab. 9 Characteristics of label F and value of its mutual information quantity

项目	T1C7 LB	C1 头 AR	T1C7 AR	C1 头 LB	右肩	
					Ab/Ad	IR/ER
互信息量	0.068 1	0.061 8	0.061 6	0.061 4	0.058 9	0.058 1

表 10　标签 G 扭头的特征及其互信息量

Tab. 10　Characteristics of label G and value of its mutual information quantity

项目	T9T8	L4L3	L1T12	C1 头		T1C7
	AR	AR	AR	LB	F/E	F/E
互信息量	0.043 9	0.043 5	0.043 4	0.043 1	0.043 1	0.042 2

表 11　标签 H 旋钮的特征及其互信息量

Tab. 11　Characteristics of label H and value of its mutual information quantity

项目	T1C7	C1 头	T9T8	L1T12	右肩	右肘
	AR	AR	AR	AR	Ab/Ad	UD/RD
互信息量	0.088 8	0.087 9	0.086 3	0.084 9	0.073 8	0.067 2

由于接电话和喝水有多种变化策略，难以确定其关键变化，所以根据不同特征的互信息量，选择显著高于其它特征的值。图 5 为与标签 I 互信息量排序前十的特征互信息量柱形图，变化趋势从序列 4 开始逐渐平缓，标签 I 接电话的特征选择为左肘内屈/外伸，左肘前旋/后旋，左手腕前旋/后旋。图 6 为与标签 J 互信息量排序前十的特征互信息量柱形图，变化趋势从序列 6 开始逐渐平缓，标签 J 喝水的特征选择为左手腕屈/伸，左肘前旋/后旋，右肩内收/外展，左肘内屈/外伸，右手腕尺侧/桡侧偏移。

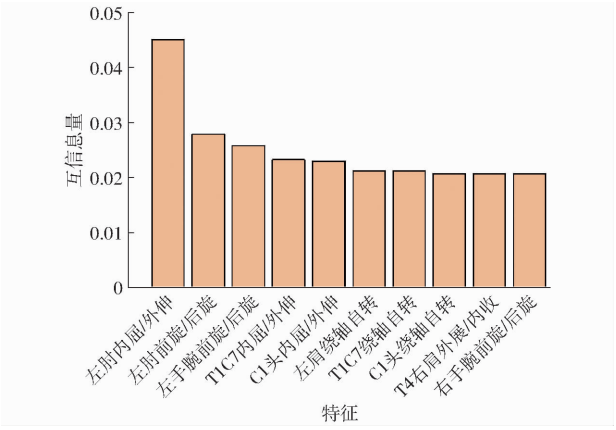

图 5　部分特征与标签 I 互信息量柱形图

Fig. 5　Particular feature and label I mutual information bar chart

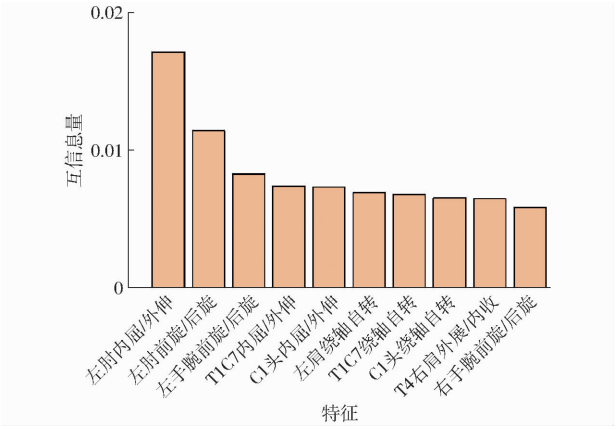

图 6　部分特征与标签 J 互信息量柱形图

Fig. 6　Particular feature and label J mutual information bar chart

2.3　特征有效性验证

为验证特征筛选的有效性并建立性能基准，构建了一个分支式 MLP 模型[32]。该模型为每个标签设计独立分支，输入对应特征子集，包含两层隐藏层和 Dropout 以防止过拟合，其中隐藏层分别为 128、64 个神经元，激活函数采用 ReLU。

采用交叉熵损失和 Adam 优化器，模型经过 30 个周期训练，损失函数从 0.94 平稳降至 0.34，验证集宏平均 F1 分数（Macro-F1）从 49% 提升至 81%（最终周期达到 80%，第 28 周期最高达到 81%），表明模型具备良好的特征学习能力。然而，其性能仍存在明显缺陷，表现为对某些少数类别的识别能力依然不足。各标签详细性能如表 12 所示。

表 12　MLP 模型性能

Tab. 12　MLP model performance　%

标签	宏平均 F1 分数	各类别 F1 分数	最差少数类召回率
A	69	100/67/39	30（类别 2）
B	80	100/80/61	48（类别 2）
C	89	100/79	72（类别 1）
D	47	97/42/2	1（类别 2）
E	72	99/67/50	43（类别 2）
F	77	99/62/69	55（类别 2）
G	81	99/63	66（类别 1）
H	92	100/85	85（类别 1）
I	95	100/90	92（类别 1）
J	97	100/94	94（类别 1）

MLP 模型的平均性能（宏观 F1 分数为 80%）良好，证明了特征筛选的有效性与可靠性，确立了一个性能基线。然而，其在标签 D、E 上表现出对少数类的识别困难，标签 D 的类别 2 召回率仅为 1%，F1 分数为 2%，这凸显了简单感知机结构在捕捉复杂特征模式方面的固有局限性。

3　基于 Transformer 框架的识别模型

3.1　模型设计

基于 Transformer 架构[33]设计姿态识别模型。模型输入惯性动作捕捉获取的姿态角数据，提取为 10 个标签 A、B、C、D、E、F、G、H、I、J 的特征值，姿态标签如表 1 所示，标签对应的特征值如表 4 ~ 11 及图 5、6 所示。标签有二分类与三分类，A、B、D、E、F 为三分类标签，其余为二分类标签。基础结构如表 13 所示。该模型能够获取时间序列关节角度数据中的时序特征与空间依赖，以此提升多类别姿态识别的准确性。

表 13 模型框架

Tab. 13 Model framework

标签	输入特征维度(F)	输入形状(batch, 15, F)	分类器类型	输出类别数	类别定义			输出形状(batch, C)
					类别 0	类别 1	类别 2	
A	6	(batch, 15, 6)	单标签	3	未出现	向前拧	向后拧	(batch, 3)
B	5	(batch, 15, 5)	多标签联合	11*	未出现	向前推	向后拉	(batch, 11)
C	7	(batch, 15, 7)	单标签	2	未出现	出现	-	(batch, 2)
D	6	(batch, 15, 6)	多标签联合	11*	未出现	顺时针旋转	逆时针旋转	(batch, 11)
E	6	(batch, 15, 6)	单标签	3	未出现	顺时针旋转	逆时针旋转	(batch, 3)
F	6	(batch, 15, 6)	多标签联合	11*	未出现	拉杆	推杆	(batch, 11)
G	6	(batch, 15, 6)	单标签	2	未出现	出现	-	(batch, 2)
H	6	(batch, 15, 6)	多标签联合	11*	未出现	出现	-	(batch, 11)
I	3	(batch, 15, 3)	单标签	2	未出现	出现	-	(batch, 2)
J	5	(batch, 15, 5)	单标签	2	未出现	出现	-	(batch, 2)

注:标签 B、D、F、H 采用多标签联合分类器,联合输出 11 个类别,涵盖这 4 个标签的所有类别数之和;* 表示其分类器类型为多标签叠加;- 表示无此类别;batch 为批次大小(Batch size),指一次输入模型的样本数量;F 为特征维度;C 为类别数。

将每个滑动窗口内的多维特征序列,即拥有时间序列的各姿态角作为 Transformer 模型的输入。模型核心采用 Transformer 编码器结构,具体包括输入线性投影层、位置编码模块、多层自注意力编码器和分类输出层。输入线性投影层将原始特征维度映射至$d_{model}=64$ 维的嵌入空间,随后加入位置编码以保留序列的时间顺序信息。编码器由两层 Transformer 编码器层组成,每层含 4 头多头自注意力机制和前馈全连接网络,隐藏层维度为 128,使用残差连接与层归一化以促进梯度传播。

针对由右手、右臂和右肩协同完成的 4 个关联性姿态(B、D、F、H),采用了多标签联合训练的设计。该设计基于以下 3 个核心依据:①这些姿态在物理层面共享右手、右臂和右肩的生理结构,且其变化具有固有的互斥性。②标签的发生顺序具有关联性,且时序间蕴含着信息。③右手腕、右肘和右肩等关节角度特征是这 4 个姿态共享的表征,不同姿态的特征之间存在交互作用。为建模这些复杂的关联性,将 4 个标签的特征合并,构成联合输入特征向量,并采用 one-hot 编码构建多标签多类别二值矩阵。模型输出各类别的预测概率,经过阈值(0.5)二值化后,施加互斥约束。该约束源于右手姿态在同一时刻的物理唯一性,即在同一时刻只能执行一个标签下的某一具体类别。此机制确保了预测结果语义一致性与人体工效学的合理性,从而显著提升了准确性。相比之下,其余 6 个标签由于相对独立性,被定义为单标签任务,分别训练独立的 Transformer 模型。每个模型仅输入其对应标签的特征子集,并输出该标签的专属单标签预测结果。

模型输出为时间序列的全局特征表示,通过对时间维度的平均池化获得固定维度向量,最后接入全连接分类层完成姿态类别预测。针对不同标签的类别数量,设计对应的输出维度。单标签任务采用多分类交叉熵损失函数训练,多标签任务采用多标签二值交叉熵损失函数,输出层设计为多标签多类别的二值编码形式。

3.2 模型结果

模型使用 143 160 条数据,按比例 8∶2 划分为训练集(114 528 条)和验证集(28 632 条),训练均进行 30 个 Epoch,模型均实现良好收敛与稳定性能。具体结果如表 14 所示。针对右手 4 个标签设计的联合模型,损失从 0.070 4 迅速下降至 0.005 6,

表 14 模型结果

Tab. 14 Model results

标签	任务类型	类别数	宏平均			支持样本数(总计)	验证集宏平均 F1 分数(最终 Epoch)/%
			精确率/%	召回率/%	F1 分数/%		
B、D、F、H	多标签联合	12	88	90	89	114 528	88
A	单标签	3	97	97	97	28 632	97
C	单标签	2	98	92	95	28 632	94
E	单标签	3	97	92	94	28 632	84
G	单标签	2	92	96	94	28 632	92
I	单标签	2	99	100	99	28 632	99
J	单标签	2	96	100	98	28 632	97

12 个类别上的宏平均 F1 分数在 30 个周期内从 45% 稳步提升至 80%，充分体现了联合学习对捕捉标签间内在关联性的有效性。其余 6 个标签分别通过单标签 Transformer 模型独立训练，均取得了较高的识别精确率和稳定的宏平均 F1 分数，尤其是三分类标签 A 和二分类标签 I，分别达到了 97% 和 99% 的宏平均 F1 分数，显示了模型对不同任务类型的良好适应能力。整体来看，模型在多类别、多任务场景下均表现出良好的泛化能力和收敛稳定性，验证了 Transformer 结构在惯性动作捕捉数据姿态角分类任务中的有效性与平稳性。

3.3　模型验证

利用 MLP 基准验证模型有效性。结果表明，Transformer 模型不仅解决了标签 D 的识别问题，而且在性能方面显著超越 MLP 基准，具体对比如表 15 所示。Transformer 在标签 A、E 上宏平均 F1 分数实现了 28、12 个百分点的显著提升，对于 MLP 表现极差的标签 D（宏平均 F1 分数 47%），其作为多标签模型的一部分，性能被提升至 88% 的水平，彻底解决了 MLP 在特定标签上的识别瓶颈。Transformer 模型凭借其强大的时序建模和特征交互能力，依然展现出显著且一致的性能优势。这不仅证明了其架构的优越性，更重要的是验证了其在真实、复杂且不均衡的运动学数据分类任务中的有效性和可靠性。

表 15　MLP 与 Transformer 模型性能对比

Tab. 15　Comparison of performance between MLP and Transformer models

标签	宏平均 F1 分数/%		性能提升/个百分点
	MLP	Transformer	
A	69	97	28
B	80	（多标签联合）	–
C	89	94	5
D	47	（多标签联合）	–
E	72	84	12
F	77	（多标签联合）	–
G	81	92	11
H	92	（多标签联合）	–
I	95	99	4
J	97	98	1
D/E/F/H	72	88	16

相较于 MLP 在处理某些复杂模式（如标签 D）时存在瓶颈，Transformer 模型展现出了强大的适应性和优越性。其自注意力机制能够自适应地捕捉不同特征维度间的全局依赖关系，为分类决策提供更丰富的上下文信息，这是 MLP 等局部连接模型所不具备的。同时，姿态编码的引入有效建模了运动学特征的时序特性，将静态特征点转化为动态序列，从而能够更好地理解动作的连续性与动态变化。而多标签联合学习架构通过参数共享和特异性优化，显著提升了模型的学习效率和泛化性能，尤其适用于存在相关性的任务组。

总的来说，模型实验验证了 Transformer 架构在人体姿态分析任务中的巨大潜力和可靠性。为基于可穿戴传感器的高精度、高稳定性姿态识别与评估提供了一种有效的解决方案。

4　结论

（1）针对拖拉机驾驶员上半身姿态的数字孪生虚拟体建模问题，提出了一种基于动捕数据的姿态识别与模型构建方法。通过采集拖拉机驾驶员在作业全过程中的上半身动捕数据，结合用户旅程图调研获取的作业全流程及驾驶员状态特征标签，系统分析和筛选了具有识别价值的特征标签。基于互信息量计算以及人体肌骨功能与姿态差异性的综合判断，选取了关键特征以提升模型的识别准确性。

（2）在模型构建方面，首先采用多层感知机（MLP）建立基准模型，验证了特征筛选方法的有效性。随后引入 Transformer 模型，进一步提升了姿态识别性能，最终实现了各标签宏平均 F1 分数在 88% 至 99% 之间的优异表现。该结果表明所提方法能够高效且准确地识别拖拉机驾驶员的上半身姿态，为数字孪生虚拟体的构建提供了坚实的技术支撑。

参　考　文　献

[1] Singh M, Fuenmayor E, Hinchy E P, et al. Digital twin: origin to future[J]. Applied System Innovation, 2021, 4(2):36.

[2] Tiwari S, Shukla A. Digital twins in industry: real-world applications and innovations[J]. Springer Nature Switzerland, 2024, 4(2):1 – 18.

[3] Botín-Sanabria D M, Mihaita A, Peimbert-García R E, et al. Digital twin technology challenges and applications: a comprehensive review[J]. Remote Sensing, 2022, 14(6):1335.

[4] Zhang Yanan, Du Yuefeng, Yang Zihan, et al. Construction method of high-horsepower tractor digital twin[J]. Digital Twin, 2022, 1(4): 12.

[5] He Qiqi, Li Li, Li Dai, et al. From digital human modeling to human digital twin: framework and perspectives in human factors [J]. Chinese Journal of Mechanical Engineering, 2024, 37(1): 1 – 14.

[6] Yao Junfeng, Yang Yong, Wang Xuecheng, et al. Systematic review of digital twin technology and applications[J]. Visual Computing for Industry, Biomedicine, and Art, 2023, 6(1): 10.

[7] Jia Wenjie, Wang Wei, Zhang Zhenzu. From simple digital twin to complex digital twin Part I: a novel modeling method for

multi-scale and multi-scenario digital twin[J]. Advanced Engineering Informatics, 2022, 53: 101706.
[8] Scataglini S, Truijen S. Critical appraisal of using digital human model, virtual human, human digital twin and digital twin[C]// International Conference on Digital Human Modeling, 2023: 154 - 158.
[9] Paul G, Abele N D, Kluth K, et al. A review and qualitative meta-analysis of digital human modeling and cyber-physical-systems in Ergonomics 4.0[J]. IISE Transactions on Occupational Ergonomics and Human Factors, 2021, 9(34):111 - 123.
[10] Fei Tao, He Zhang, Ang Liu, et al. Digital twin in Industry: state-of-the-art[J]. IEEE Trans. Industrial Informatics, 2019, 15(4):2405 - 2415.
[11] Elif Bozkaya A, Tolga O, Levent E, et al. Intelligent human activity recognition for healthcare digital twin[J]. Internet of Things, 2025, 30:101497.
[12] Saul Davila G, Sergio M. Human digital twin in Industry 5.0: a holistic approach to worker safety and well-being through advanced AI and emotional analytics[J]. Sensors, 2024, 24(2): 655.
[13] Lu Weisheng, Chen Junjie, Fu Yonglin, et al. Digital twin-enabled human-robot collaborative teaming towards sustainable and healthy built environments[J]. Journal of Cleaner Production, 2023, 412: 137412.
[14] Maruyama T, Ueshiba T, Tada M, et al. Digital twin-driven human robot collaboration using a digital human[J]. Sensors, 2021, 21(24):8266.
[15] Nasirahmadi A, Hensel O. Toward the next generation of digitalization in agriculture based on digital twin paradigm[J]. Sensors, 2022, 22(2):498.
[16] Nikolaos P, Dimitrios P, Vasileios C, et al. Enhancing smart agriculture by implementing digital twins: a comprehensive review[J]. Sensors, 2023, 23(16): 7128.
[17] Cesco S, Sambo P, Borin M, et al. Smart agriculture and digital twins: applications and challenges in a vision of sustainability [J]. European Journal of Agronomy, 2023, 146:126809.
[18] Zain A A, Mahreen Z, Raza H, et al. Circular economy advances with artificial intelligence and digital twin: multiple-case study of Chinese industries in agriculture[J]. Journal of the Knowledge Economy, 2024, 16(1):1 - 37.
[19] Chaux J D, Sanchez L D, Barbieri G. A digital twin architecture to optimize productivity within controlled environment agriculture[J]. Applied Sciences, 2021, 11(19):8875.
[20] 郭博健,王娜,王君姣.数字孪生系统在农业领域的建构路径探究[J].现代农业装备,2024,45(2):29 - 33,79.
Guo Bojian, Wang Na, Wang Junjiao. Exploration of the construction path of digital twin system in the field of agriculture [J]. Modern Agricultural Equipment, 2024, 45(2): 29 - 33, 79. (in Chinese)
[21] Ma Y, Du R, Abdelraouf A, et al. Driver digital twin for online recognition of distracted driving behaviors[J]. IEEE Transactions on Intelligent Vehicles, 2024, 9(2): 3168 - 3180.
[22] Hu Z, Lou S, Xing Y, et al. Review and perspectives on driver digital twin and its enabling technologies for intelligent vehicles[J]. IEEE Transactions on Intelligent Vehicles, 2022, 7(3): 417 - 440.
[23] Mahomed A S, Saha A K. Driver posture recognition: a review[J]. IEEE Access, 2024, 12: 176301 - 176345.
[24] Liao X, Zhao X, Wang Z, et al. Driver digital twin for online prediction of personalized lane-change behavior[J]. IEEE Internet of Things Journal, 2023, 10(15): 13235 - 13246.
[25] Shoukat M U, Yan L, Yan Y, et al. Autonomous driving test system under hybrid reality: the role of digital twin technology [J]. Internet of Things, 2024, 27: 101301.
[26] Zhang J, Jin X,Huang Q, et al. Human sitting posture changes under different backrest tilt angles and headrest conditions in autonomous vehicle[J]. International Journal of Precision Engineering and Manufacturing,2025,26(8):1 - 14.
[27] 宋怀波,韩梦璇,王云飞,等. 基于人体姿态估计与场景交互的果园喷施行为检测方法[J]. 农业机械学报, 2023, 54(2): 63 - 72.
Song Huaibo, Han Mengxuan, Wang Yunfei, et al. A method for detecting spraying behavior in orchards based on human pose estimation and scene interaction[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(2): 63 - 72. (in Chinese)
[28] 任德良,徐书雷,王兴伟,等. 基于 RAMSIS 的拖拉机驾驶室人机工程布置设计与测试[J]. 农业机械学报, 2022, 53(增刊 2): 303 - 309.
Ren Deliang, Xu Shulei, Wang Xingwei, et al. Design and testing of an ergonomic layout for a tractor cab based on RAMSIS [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(S2): 303 - 309. (in Chinese)
[29] 蒲六如,赵永杰,杨广元,等. 基于 Gait - AVG 的设施火龙果园喷施作业人员重识别方法[J]. 农业机械学报, 2025, 56(8): 438 - 446.
Pu Liuru, Zhao Yongjie, Yang Guangyuan, et al. A Gait-AVG-based method for identifying operators during spraying operations in protected dragon fruit orchards[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(8): 438 - 446. (in Chinese)
[30] Ding Y, Lattin J M, Buahin A M, et al. A customer journey map approach with mobile application[J]. Journal of Mechanical Design, 2025, 147(7): 071404.
[31] Zhou H, Wang X, Zhu R. Feature selection based on mutual information with correlation coefficient[J]. Applied Intelligence, 2022, 52(5): 5457 - 5474.
[32] Lewicki G, Marino G. Approximation by superpositions of a sigmoidal function [J]. Zeitschrift für Analysis und ihre Anwendungen,2003,22(2):463 - 470.
[33] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[PP/OL]. V7. arXiv (2023 - 08 - 02) [2025 - 09 - 07]. https://arxiv.org/abs/1706.03762.