doi:10.6041/j.issn.1000-1298.2025.02.016

基于农业机器人本体传感信号的旱田平作与垄作类型识别方法

张伟荣[1,2]　陈学庚[3]　齐江涛[1,2]　周俊博[1,2]　温浩军[3]　刘慧力[2,4]

(1. 吉林大学工程仿生教育部重点实验室, 长春 130022; 2. 吉林大学生物与农业工程学院, 长春 130022;
3. 石河子大学机械电气工程学院, 石河子 832003; 4. 吉林省智慧农业装备与技术重点实验室, 长春 130022)

摘要: 旱田农业耕作模式包括平作与垄作,不同耕作模式的地形起伏差异大,作物行耕作模式的准确识别对机器人行走稳定性具有重要意义,提出一种基于本体传感器信号的平作与垄作类型地形识别方法。首先,采集四足机器人在玉米田间作物行内行走的机身惯性测量单元(Inertial measurement unit, IMU)信号,使用机器人左前腿的足端速度数据作为补充,生成机器人在平作与 2 种不同起垄高度的垄作种植模式下行走的信号数据集。其次,利用卷积神经网络(Convolutional neural networks, CNN)提取信号的空间信息特征,通过双向长短期记忆网络(Bidirectional long short-term memory, BiLSTM)提取时间序列特征,采用注意力机制(Self-attention, SA)提取 CNN 与 BiLSTM 输出特征信息的注意力分值。最后,通过模型对比和田间试验,验证本文模型对平作与垄作类型识别的有效性。结果表明,本文 CNN-BiLSTM-SA 模型 F1 值为 92%,与 CNN、CNN-LSTM、CNN-LSTM-SA 与 CNN-BiLSTM 模型相比,分别提升 10.17、3.51、2.57、1.27 个百分点。内嵌识别模型的田间机器人可在 1.4 s 内实现对当前作物行平作与垄作类型 90% 的识别准确率,在 4.8 s 内达到对作物行类别分类要求,满足机器人面对作物行不同地形的识别快速性、准确性要求。该算法能提供机器人在旱田典型耕作模式下的地形识别能力,为提高四足机器人作业的田间稳定性提供技术支撑。

关键词: 农业机器人; 惯性测量单元; 本体传感信号; 地形识别; 长短期记忆网络

中图分类号: TP29; S24　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)02-0164-11　　**OSID**:

Recognition Method of Flat and Ridged Crop Types in Dry Fields Based on Propriety Sensing Signals of Agricultural Robot

ZHANG Weirong[1,2]　CHEN Xuegeng[3]　QI Jiangtao[1,2]　ZHOU Junbo[1,2]　WEN Haojun[3]　LIU Huili[2,4]

(1. *Key Laboratory of Bionic Engineering*, *Ministry of Education*, *Jilin University*, *Changchun* 130022, *China*
2. *College of Biological and Agricultural Engineering*, *Jilin University*, *Changchun* 130022, *China*
3. *College of Mechanical and Electrical Engineering*, *Shihezi University*, *Shihezi* 832003, *China*
4. *Jilin Provincial Key Laboratory of Smart Agricultural Equipment and Technology*, *Changchun* 130022, *China*)

Abstract: Dryland agricultural cultivation modes include flat cropping and ridge cropping, and the terrain undulation of different cultivation modes varies greatly, so accurate crop row cultivation mode recognition is of great significance to the stability of robot travelling. A methodology for identifying the terrain of crop rows and ridges utilizing appropriate sensor signals was introduced. Initially, the inertial measurement unit (IMU) signals were collected from a quadrupedal robot navigating through the crop rows of a corn field. The velocity data from the robot's left front leg served as supplementary information to compile a comprehensive signal dataset, encompassing the robot's movement in both flat cropping and row cropping modes, with two distinct row heights. Subsequently, spatial information features were extracted from the signals by using convolutional neural networks (CNN), while time series features were derived through bidirectional long short-term memory (BiLSTM) networks. Additionally, self-attention

收稿日期: 2024-11-04　修回日期: 2024-12-26

基金项目: 国家自然科学基金项目(32271988)、吉林省科技发展计划项目(20230508032RC)和吉林省重点研发计划项目(20220202028NC)

作者简介: 张伟荣(1994—),男,博士生,主要从事精准农业与仿生智能农机装备研究,E-mail: zhangwr21@mails.jlu.edu.cn

通信作者: 齐江涛(1981—),男,教授,博士生导师,主要从事精准农业与仿生智能农机装备研究,E-mail: qijiangtao@jlu.edu.cn

(SA) was employed to capture the attention scores of the output feature information from both CNN and BiLSTM. Ultimately, the efficacy of the proposed model in distinguishing between flat and ridge crop types was validated through model comparisons and field experiments. The results indicated that the F1 score of proposed CNN－BiLSTM－SA model reached 92%, marking an improvement of 10.17, 3.51, 2.57 and 1.27 percentage points over that of the CNN, CNN－LSTM, CNN－LSTM－SA, and CNN－BiLSTM models, respectively. When the recognition model was embedded in the field robot, it achieved a 90% accuracy rate in identifying the current crop row tillage type within 1.4 s, and met the classification criteria for flat and ridge categories within 4.8 s. This performance satisfied the robot's requirements for rapid and accurate recognition across various tillage terrains. The algorithm can provide the robot with ability to recognize crop rows under typical tillage patterns in dry fields, and the results can provide technical support for improving the field stability of quadrupedal robots in autonomous operations.

Key words: agricultural robots; inertial measurement units; proprioceptive signals; terrain recognition; long and short-term memory networks

0　引言

玉米是我国主要粮食作物之一,对于保障粮食安全具有重要意义。随着农业装备与信息技术的高度融合,农业机器人自主作业对于提高粮食生产力和农业生产效率具有重要意义[1-2]。利用农业机器人进行大田巡检,实现作物信息监测与病虫害防治,是玉米生长中后期作物管理的良好解决方案[3]。农业机器人需要具备在复杂非/半结构化环境下感知作业环境的能力[4-5],对于农业四足机器人,具备对当前作物行种植模式的地形识别可使机器人及时调整机身姿态,以稳定地进行农田作业。因此,作物行种植模式的地形识别是四足机器人在非结构化环境中进行农情巡检任务的基础功能。

研究人员针对农业地形测量展开了研究[6-7],但没有涉及农业机器人田间自主作业过程中的地形识别。国内外研究人员针对机器人通行环境的地形识别问题展开研究,包括轮式[8]和足式[9]机器人。基于不同形式的感知传感器,可以分为:基于外部传感器和本体传感器的地形感知方法。外部传感器,如基于视觉的传感器、激光雷达(Light detection and ranging, LiDAR)等,常被用于机器人系统的行走环境地形识别[10-12]。有研究基于视觉和激光雷达传感器进行地形识别,提高机器人室外环境中的通过性[13-16]。但相关研究只是识别可通行区域,不能对行走地形进行具体分类。通过视觉和雷达的外部传感器获取的图像和三维点云训练数据不能直接用于地形识别,需要对数据进行手动标记以构建完整的训练数据集,需要大量的计算资源用于数据处理,如CNN 训练,当训练数据中没有足够的差异性,模型分类性能往往会下降。当机器人在农业复杂环境行走时,面对不同光照强度变化、能见度不高及地面杂草覆盖、秸秆干扰的情况时,视觉和雷达传感器的地形识别模型精度会受较大影响[17-18]。

为了弥补复杂农业环境下视觉和雷达地形识别方法的局限性,研究者提出了基于本体传感器来实现高精度的地形识别[19],通过使用机器人与地形接触产生的机器人本体感知信息,如振动信号或 IMU 信号来解决地形视觉外观的干扰[20-21]。然而,振动信号易受到机器人电机或者发动机自身振动的影响,影响机器人的感知精度。与轮式机器人不同,四足机器人通过腿部摆动前进,腿部或足端信息可作为本体传感器信号的补充,提高地形分类精度。但相关研究[22-24]只对斜坡地形进行了识别,对于农业环境中的其他地形没有展开研究。研究不能推广到非周期步态、不同速度或不平坦地形下的情形[25],对农业环境不同耕作模式下的地形识别也缺乏研究。

农田地形表面存在起伏差异,尤其是作物行垄作地形会影响机器人通行的稳定性,但目前进行农业机器人农田环境平作与垄作类别的识别研究较少。四足机器人可根据不同耕作模式的地形差异采用不同的行走策略,及时调整机身高度等机身姿态,提高四足机器人农田自主作业的行走效率和稳定性。基于上述问题,本文提出一种适用于玉米田的机器人行走地形类别识别方法,通过机器人本体传感器采集的加速度计和陀螺仪等数据,利用 CNN 和 LSTM 进行数据特征提取,完成玉米大田常见种植模式下机器人行走地形类别识别。

1　数据采集与数据集制作

1.1　设备与作业情形

相比传统的轮式和履带式机器人,四足机器人在移动能力、适应性方面展现出独特优势[26-27]。通过实时监测机器人的 IMU 传感器数据,获取加速度、角速度等机身姿态参数,是通过稳定控制策略调整腿部关节运动、保持机器人平衡和稳定性的前提,以此提高四足机器人在非结构农业环境中作业的通

行能力。本研究使用 Unitree Robotics 公司的 A1 四足机器人平台,在狭窄的玉米冠层下,实现机器人的数据快速采集和处理。如图 1 所示,利用四足机器人采集田间不同地形的加速度计和陀螺仪数据,保存在 NVIDIA JETSON NANO 处理器中,通过处理器分析行走状态下的数据特征,对不同种植模式下的地形类别进行识别,为后续机器人调整机身行走策略提供信息感知能力。



(a) 四足机器人 (b) 平作作物行 (c) 高起垄作物行 (d) 低起垄作物行

图 1 四足机器人 IMU 数据采集典型场景

Fig. 1 Typical scenarios of IMU data acquisition for quadruped robots

1.2 数据集制作

1.2.1 IMU 数据采集

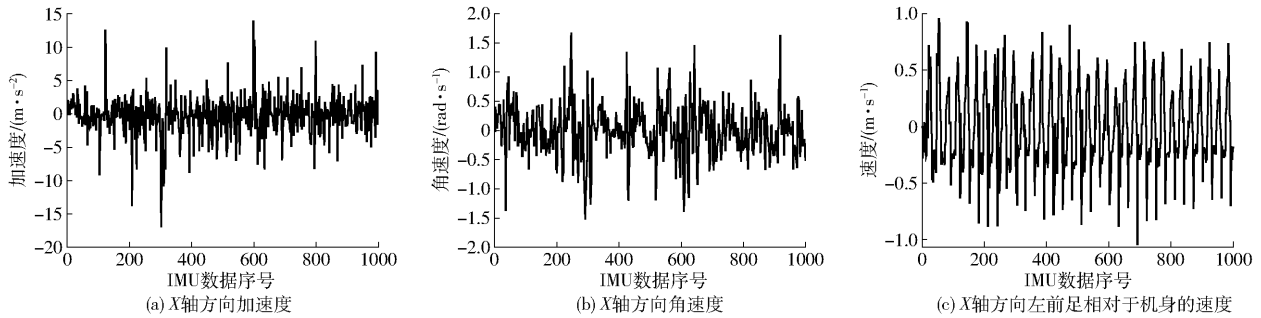
在玉米大田种植的农业场景中,按照是否起垄种植,分为平作、垄作模式,其中垄作模式包含不同高度的起垄种植。IMU 传感器输出频率为 100 Hz。经过前期预试验,在数据采集过程中前进速度设定为 0.3 ~ 0.5 m/s,此时机器人行走稳定性较好,符合机器人在田间巡检时的真实作业速度。主要在吉林省长春市吉林大学农业试验基地进行数据采集,考虑不同土壤环境、种植垄作模式和起垄高度的差异等,在云南省昆明市、海南省三亚市也进行了部分数据采集,采集机器人在不同种植模式下作物行行走的本体传感器信号,采集时间为 2023 年 8 月至 2024 年 7 月。

对于机器人行走的地形,根据玉米田间实际情况和种植模式,本文选定 3 种地形进行机器人行走的数据采集,地形为:田间平作作物行、高起垄作物行(高度 20 cm)、低起垄作物行(高度 10 cm)3 种常见地形。对于采集的四足机器人传感器数据信息,包括来自 IMU 传感器的 X、Y、Z 方向上的加速度计数据和陀螺仪数据,考虑到不同种植模式的地形对四足机器人足端摆动速度影响较大,采集 X、Y、Z 方向上左前腿足端相对于机身的速度作为数据补充,其中左前足的速度基于机器人运动学模型得到,由机器人自带的软件开发工具包(SDK)导出,数据存储在四足机器人机身自带的控制器中。以 IMU 传感器输出数据序号为横坐标,以 X 轴方向为例,机器人采集的原始数据如图 2 所示。

1.2.2 数据降噪与增强

在 IMU 传感器初始化或机器人处于静止状态时,IMU 传感器采集的角速度数据会在某个固定值附近产生波动,信号振荡冗余对后续的地形识别模型精准度产生不利影响。因此,在数据滤波之前首先对机器人的零偏问题进行预处理。具体做法是:在机器人上电后,采集 1 000 帧陀螺仪的原始数据,并计算数据的标准差,该标准差被视为传感器的零偏值。随后,将每一帧的原始数据都减去此零偏值,以消除零偏带来的误差。再对处理后的数据进行高斯滤波,以进一步平滑信号并减少噪声干扰。

在采集试验数据过程时,四足机器人在行走过程中地面或机体会产生噪声,影响数据分析和模型准确性、可靠性。为此本文使用高斯滤波器对数据进行滤波去噪。为了更直观展示滤波前后的数据对比,本文以 X 轴方向上加速度和 X 轴方向左前足相对于机身的速度为例绘制曲线,信号



(a) X轴方向加速度 (b) X轴方向角速度 (c) X轴方向左前足相对于机身的速度

图 2 采集的田间原始 IMU 数据

Fig. 2 Raw IMU data collected in field

在滤波操作前后波形对比见图 3。高斯滤波本质为二维低通滤波，可以根据曲线标准差调整高斯核大小，以调整滤波性能。滤波后曲线更加平滑，减少了原始波形中的尖峰，同时还能还原出信号变化趋势。

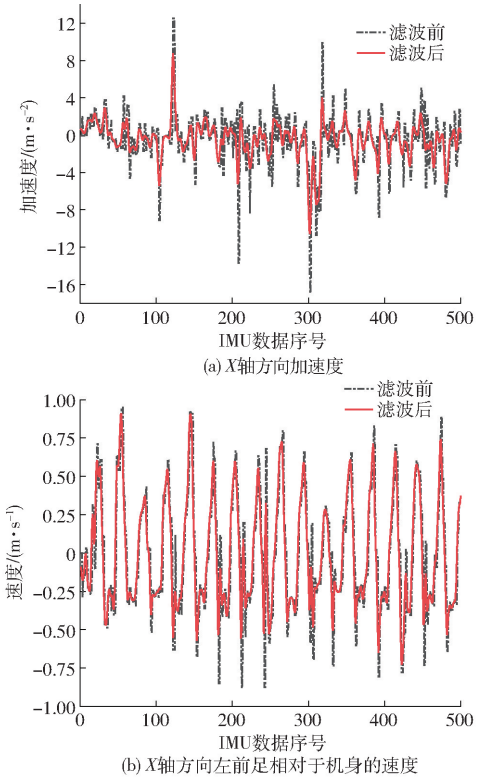


图 3　滤波前后数据波形

Fig. 3　Data wave forms before and after filtering

数据增强是一种利用先验知识在不改变原始数据标签的情况下扩展原始数据的技术，增加样本的多样性，平衡数据集，并提高模型的稳健性。本文在不改变步态类别的基础上，对采集的 IMU 信号进行部分倒置以扩充数据集，生成新的训练样本，增加数据集的多样性和规模。有助于模型更好地学习数据的内在特征，提高泛化能力，并减少过拟合的风险。经上述过程得到数据集，其中训练集、验证集及测试数据集分别包括 15 000、4 500、1 500 组数据。地形类别数据集样本数如表 1 所示。

表 1　不同地形的信号数据集样本数

Tab. 1　Number of samples in signal dataset for different terrains

地形类别	训练集样本数	验证集样本数	测试集样本数
平作作物行	5 000	1 500	500
高起垄作物行	5 000	1 500	500
低起垄作物行	5 000	1 500	500

1.2.3　数据切分及归一化

长序列传感器样本划分为短序列长度的数据集，不仅利于提高模型的训练效率，且能提高模型对数据前后时间序列特征的提取能力。窗口长度包括在一定时间内的所有原始传感器数据，由于机器人足端摆腿频率与前进速度保持固定，当所选滑动窗口包含的数据序号是 128 时，如图 4 所示，包含 4 次完整的摆腿周期。取滑动步长为 64，即取两个滑动窗口 50% 的重叠数据用于训练，更有利于增强网络的上下文理解能力。图 4 显示了在高起垄作物行地形下的一个连续滑动窗口的示例。

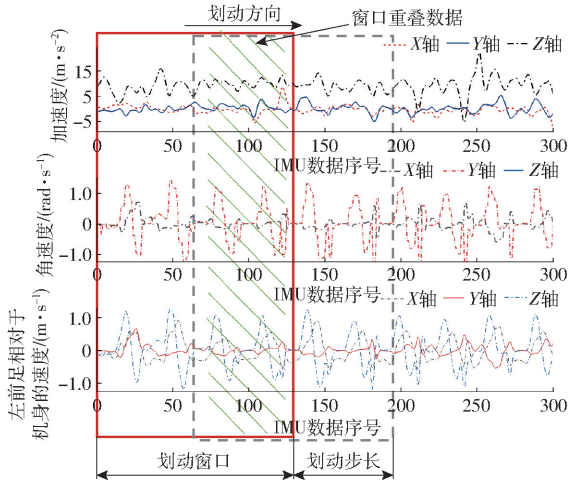


图 4　50% 数据重叠下的滑动窗口数据切分示例

Fig. 4　Example of sliding window data slicing

加速度计和陀螺仪数据的幅值有很大的区别，通道中的数值若直接用来训练模型，可能会导致训练模型不稳定。本文对数据进行归一化，以产生相同数量级的所有数据值。采用零均值归一化方法将加速度和陀螺仪数据归一化为[-1,1]。

2　基于 IMU 数据的地形识别模型

IMU 信号包含丰富的时间序列和空间信息，输出的是时间与幅度信息，属于时间序列数据，每个样本表示为一维向量。一维卷积神经网络（1D - CNN）能捕捉信号的局部信息，常被用作特征提取器，但 1D - CNN 在涉及时间序列相关输入任务的表现相对有限，例如本文中 IMU 信号数据的分析。LSTM 能捕获长距离的序列信息，在进行序列特征建模时具备优势。本文将 1D - CNN 与 LSTM 相结合，添加注意力模块增强网络特征提取能力，提出了一种融合架构的四足机器人行走地形类别识别模型，主要由 1D - CNN、LSTM、注意力模块和输出层组成，借助不同网络的优势来提升机器人地形识别的准确性。

2.1　卷积神经网络

对于 IMU 数据，1D - CNN 优点是可以直接输入简单滤波处理后的原始输入信号，通过滑动窗口机制进行卷积，生成特征图，提取序列数据高维特征。

图 5 为本文构建的卷积神经网络结构图。由卷积层、RelU 激活层、最大池化层和丢弃层组成。

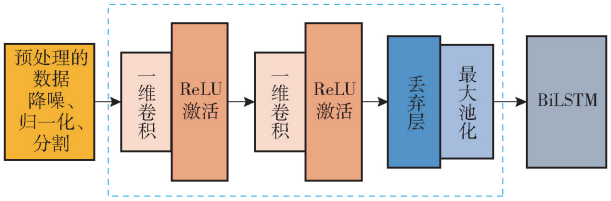


图 5 CNN 网络结构图

Fig. 5 Diagram of convolutional neural network structure

使用 2 层的一维卷积神经网络。第 1 层包含 64 个卷积核,大小为 3,第 2 层包含 64 个卷积核,大小为 5。最大池化层的池大小统一为 2。

2.2 长短期记忆网络

LSTM 是循环神经网络(Recurrent neural network, RNN)的一种,相对于传统神经网络,RNN 网络能更好地处理时间序列问题。LSTM 增加了存储记忆的单元,通过引入门控结构控制信息传递的路径,学习序列数据的长期和短期时间序列特征。

LSTM 每层有 32 个存储单元。输入数据被发送到不同的门,包括输入门、遗忘门和输出门 3 个门控结构,以控制每个记忆细胞的行为。如图 6 所示,双向长短期记忆(Bidirectional long short-term memory, BiLSTM)网络由两个独立的 LSTM 组成,通过前向和后向传播方向的 LSTM 传输序列数据,调整前后状态对当前单元状态的影响。这种网络能够有效增强网络的上下文理解能力,捕捉更长距离的依赖关系,提高序列建模的性能,通过保存信号较长的依赖关系,提高模型的预测精度。

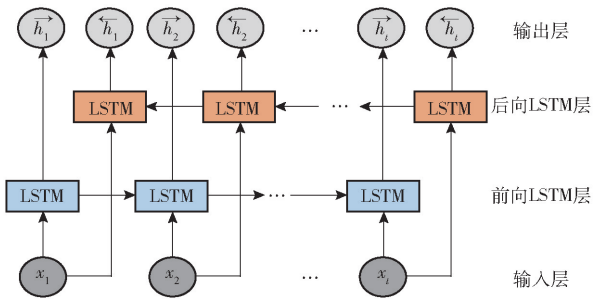


图 6 BiLSTM 网络结构图

Fig. 6 Structure diagram of bidirectional long short-term memory network

在本文中,输入数据首先经过两层 LSTM,以便更好地提取序列数据中的时间特征。BiLSTM 特殊的门控结构可以控制时间序列数据特征的保留与遗忘,但在单独处理长时间序列时,仍然会出现丢失重要特征的可能。因此,本文首先通过 CNN 处理输入数据,然后将输出作为 BiLSTM 网络的输入,从而提高模型预测的精度。

2.3 注意力层

注意力机制最初应用于图像识别,使模型专注于识别任务中感兴趣的区域。在 LSTM 层中加入注意力机制,在处理数据时模型可以动态地关注数据的不同部分。注意力层通过学习权重系数,增强模型在 BiLSTM 的垂直和横向传递中独立学习各个参数重要性的能力,并通过学习结果给予重要参数更大的权重。

注意力机制模块如图 7 所示,图中参数为

$$\begin{cases}[\boldsymbol{Q}_h,\boldsymbol{K}_h,\boldsymbol{F}_h]=\boldsymbol{h}_t[\boldsymbol{W}_q,\boldsymbol{W}_k,\boldsymbol{W}_v]\\ \boldsymbol{H}_t=\text{Attention}(\boldsymbol{Q}_h,\boldsymbol{K}_h,\boldsymbol{F}_h)=\text{softmax}\left(\dfrac{\boldsymbol{Q}_h\boldsymbol{K}_h^{\mathrm{T}}}{\sqrt{\boldsymbol{D}_k}}\right)\boldsymbol{V}_h\end{cases} \tag{1}$$

式中 $\boldsymbol{h}_t$——输入向量

$\boldsymbol{Q}_h$——查询向量

$\boldsymbol{K}_h$——键向量

$\boldsymbol{F}_h$、$\boldsymbol{V}_h$——值向量

$\boldsymbol{W}_q$、$\boldsymbol{W}_k$、$\boldsymbol{W}_v$——可训练的参数矩阵

$\boldsymbol{H}_t$——注意力机制模块输出值

$\boldsymbol{D}_k$——K_h的维度,为平滑权重矩阵,保持训练过程梯度的稳定性

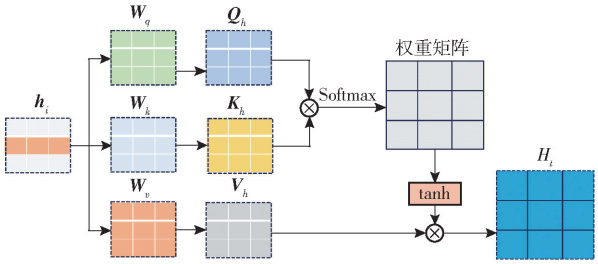


图 7 注意力机制模块结构图

Fig. 7 Structure diagram of attention mechanism module

查询向量表示模型在序列中寻找的内容,键向量计算查询向量与各个输入位置的相似性,即序列中其他单词应该注意的内容,值向量提供输入序列中的实际内容信息。注意力分数 $\boldsymbol{Q}_h$ 和 $\boldsymbol{K}_h$ 计算不同位置之间的相关性,经点积计算相似性得分,然后通过 Softmax 函数归一化得到注意力权重,而后 $\boldsymbol{V}_h$ 与权重矩阵相乘得到输出值 $\boldsymbol{H}_t$。

2.4 输出层

输出层由 1 个全连接层和 1 个 Softmax 分类器组成。全连通层的每个节点都与上层的节点相连,以便从上层提取的特征可以合并。它弥补了全局平均池化层在这方面的不足。全连接层后面为 Softmax 分类器,它将上层的输出转换为概率向量,其值表示当前样本所属地形类别的概率。

2.5 CNN-BiLSTM-SA 模型

CNN-BiLSTM-SA 模型如图 8 所示。

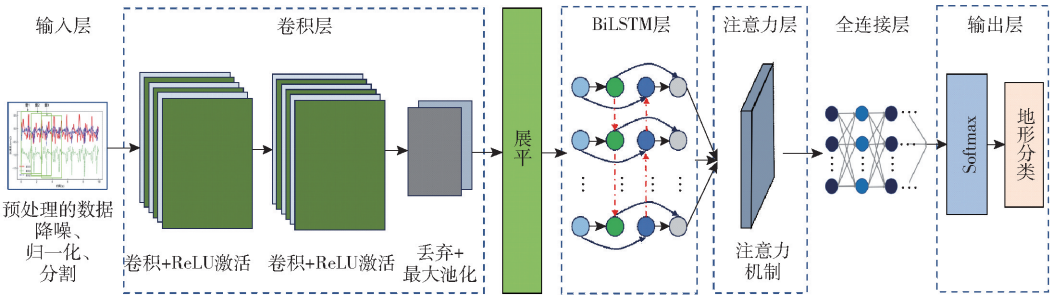


图 8　CNN－BiLSTM－SA 网络结构图

Fig. 8　Structure diagram of CNN－BiLSTM－SA network

CNN－BiLSTM－SA 模型参数如表 2 所示。

表 2　CNN－BiLSTM－SA 模型参数

Tab. 2　Parameters of CNN－BiLSTM－SA model

层类型	核尺寸	核数量	步幅	输出维度
输入层				7×1
Conv1D－1	3	64	2	3×64
最大池化－1	2		1	3×64
Conv1D－2	5	64	2	5×64
最大池化－2	2		1	5×64
展平层				320
BiLSTM				128
全连接			64	64
输出层				3

模型数据处理流程为：

(1)经预处理的数据输入 CNN 网络，通过卷积和池化操作，利用 CNN 提取多模态步态数据的高维特征，实现对数据特征的提取和降维，从 IMU 传感器信号提取抽象信息，在保留数据时间序列特征的同时，扩展了特征的维度，压缩了特征向量的长度。

(2)CNN 层将经过卷积的高维特征向量的抽象信息输入到 BiLSTM 层，从 CNN 网络提取的数据信息中学习数据间的时间拟合关系，对时间序列输入数据进行动态建模，提取长时间特征，BiLSTM 隐藏层的输出会进入注意力层进行下一步处理。

(3)数据输入注意力层，通过学习不同通道的重要性调整权重，选择性地关注序列中关键部分，聚焦更有信息量的特征，增强模型对特征的表达能力。

(4)通过全连接神经网络进行分类，输出预测值。

3　试验结果与分析

3.1　试验平台与训练参数

计算机处理器为 Intel(R) Core(TM) i7－12900K，NVIDIA GeForce RTX1050ti GPU，Ubuntu 操作系统。试验中所有程序基于 Python 3.7 语言，在 PyTorch 深度学习框架下实现。模型训练学习率为 0.01，优化器为 Adam，采用权值衰减为 1×10^{-6} 的 L2 正则化，周期为 200，批大小为 32。如 1.2.3 节所述，模型每次输入 128 个样本数。

3.2　评估指标

为了对提出的基于 CNN－LSTM－SA 特征融合的地形识别方法性能进行客观评价，模型本质为分类模型，使用交叉熵损失函数。本文使用 4 个评估指标对模型性能进行评价，分别为精确率(Precision，P)、召回率(Recall，R)、F1 值(F1-score，F1)和准确率(Accuracy，A)。

3.3　试验结果

CNN－BiLSTM－SA 模型精度和损失值如图 9 所示。图 9a 描述了模型的交叉熵损失值，损失值下降趋势逐渐变缓直至趋于稳定。图 9b 表明，在前 90 个周期中，模型准确率出现较为明显的振荡，随着训练周期的增加，模型实现了稳定收敛。试验结果表明网络学习适当，没有出现过拟合问题。

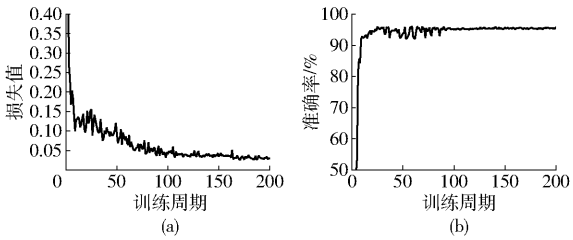


图 9　训练精度与损失值曲线

Fig. 9　Curves of training accuracy and loss value

为了更全面地评价本文提出方法的性能，图 10 显示了 CNN－BiLSTM－SA 模型的数据集混

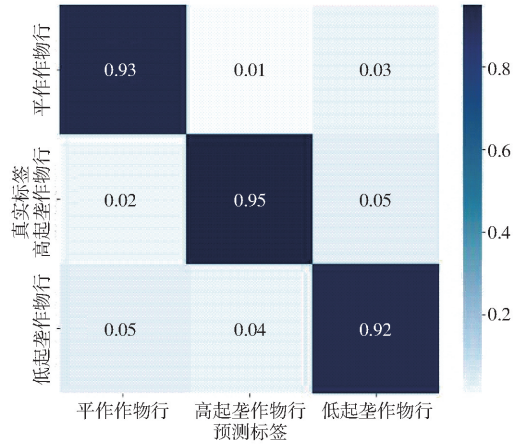


图 10　混淆矩阵

Fig. 10　Confusion matrix

淆矩阵。矩阵的对角线元素对应于模型地形分类的准确性。研究结果表明,本文模型能够有效捕获信号的时空特征。具体来说,该模型在识别高起垄地形方面达到的准确率最高,对于低起垄的地形识别率较低,主要原因是由于低起垄地形与平作及高起垄地形的数据具有一定相似性,容易出现分类混淆。

3.4 模型性能比较

为了验证本文提出的作物行平作与垄作识别模型的有效性,与其他模型进行比较。

3.4.1 参数设置

对 CNN 和 LSTM 进行了层数、卷积核尺寸和隐藏层尺寸参数试验,设置了各模型的合适参数,如表 3 所示。

表 3 模型参数设置

Tab. 3 Parameter settings of model

模型	CNN	LSTM
CNN	层数:2;卷积核尺寸:5×1×64、3×1×64	
CNN-LSTM CNN-LSTM-SA	层数:2;卷积核尺寸:5×1×256、3×1×256	层数:1;隐藏层尺寸:128
CNN-BiLSTM CNN-BiLSTM-SA	层数:2;卷积核尺寸:5×1×256、3×1×256	层数:2;隐藏层尺寸:128、128

3.4.2 模型精度对比结果

基于双向 LSTM 特征融合与注意力机制的 CNN-BiLSTM-SA 地形识别模型,包含 2 个消融因素,即 BiLSTM 与 SA。为了验证模型参数对于分类精度的影响,改变 CNN 的卷积核尺寸和 LSTM 的隐藏层尺寸,构建 5 个不同模型参数的网络模型进行试验。与其他模型识别效果对比分析,验证结合 BiLSTM 与 SA 的地形识别模型有效性,试验结果如表 4 所示,其中,表 4 中的默认参数如表 3 模型参数设置所示。

表 4 模型精度结果

Tab. 4 Results of accuracy of model %

模型	参数	P	R	F1 值	A
CNN	默认参数	81.17	82.51	81.83	87.25
	卷积核尺寸:3×1×128×2	79.12	80.13	79.62	86.23
CNN-LSTM	默认参数	87.30	89.72	88.49	88.79
	LSTM:128	79.81	80.20	80.05	84.34
CNN-LSTM-SA	默认参数	88.67	90.21	89.43	90.87
	LSTM:128	79.21	81.76	80.46	83.35
CNN-BiLSTM	默认参数	89.63	91.86	90.73	90.33
	LSTM:128	81.13	82.57	81.84	83.76
CNN-BiLSTM-SA	默认参数	91.73	92.28	92.00	92.37
	LSTM:128	84.41	85.15	84.77	85.49

(1)采用本文默认参数模型精确率、召回率和 F1 值精度都高于对比模型,表明 CNN 的卷积核尺寸和 LSTM 的隐藏层尺寸的默认参数配置对于地形识别的有效性,本文后续都以默认参数的网络进行阐述。

(2)CNN-LSTM 模型精确率、召回率和 F1 值分别为 87.3%、89.72%、88.49%,比 CNN 模型分别高 6.13、7.21、6.66 个百分点,各项评价指标均优于 CNN 模型。这是因为在 IMU 信号特征中时间序列非常重要,和 CNN 模型相比,CNN-LSTM 模型具备了对信号时间序列特征的提取能力,而单 CNN 模型对信号的局部特征提取不全面,因此 LSTM 模块的加入能更精准的分析 IMU 信号时间序列特征,模型分类能力更强。

(3)CNN-BiLSTM 模型精确率、召回率和 F1 值分别为 89.63%、91.86%、90.73%,各项评价指标均优于 CNN-LSTM 模型,精确率、召回率、F1 值相比分别高 2.33、2.14、2.24 个百分点。BiLSTM 模块通过双向 LSTM 结构动态增强了信号特征的上下文语义表示,与 CNN-LSTM 算法相比,基于双向 LSTM 的结构更易捕捉分段信号之间的联系特征并进行互补,因此,BiLSTM 的特征提取能力比 LSTM 更加优秀。

(4)加入注意力机制的 CNN-BiLSTM-SA 模型各评价指标值最高,与不带注意力机制的 CNN-BiLSTM 模型相比,F1 值提高 1.27 个百分点。这是因为注意力机制的引入有效利用了信号特征中的权重信息,通过注意力机制的改进能提升训练效果,提升地形分类能力。

上述结果表明,BiLSTM 模型比 LSTM 模型具有更高的分类和识别精度,有注意力机制的模型比没有注意机制的模型具有更高的准确率。结合双向 LSTM 融合与注意力机制的 CNN-BiLSTM-SA 模型在地形类别识别任务上具有更高的准确率,与 CNN 模型、CNN-LSTM 模型、CNN-LSTM-SA 模型、CNN-BiLSTM 模型相比,F1 值分别提升 10.17、3.51、2.57、1.27 个百分点,验证了本文模型的有效性。

3.4.3 注意力机制

深度学习技术具备对复杂数据进行特征提取的

优势，但训练过程中不能很好对重要特征赋予更多的权重。本文中，通过引入注意力机制来增强分类算法，注意力机制方法增强了可解释的表示，突出了模型中单个输入数据段的重要性。如图 11a 所示，注意力机制的加入提高了所有场景下的识别效率。尤其是对于数据具有相似性的低起垄和高起垄的地形，引入注意力机制的 CNN－SA 和 LSTM－SA 模型表现出更为显著的性能，说明注意力机制的引入对于相似地形的分类作用明显。

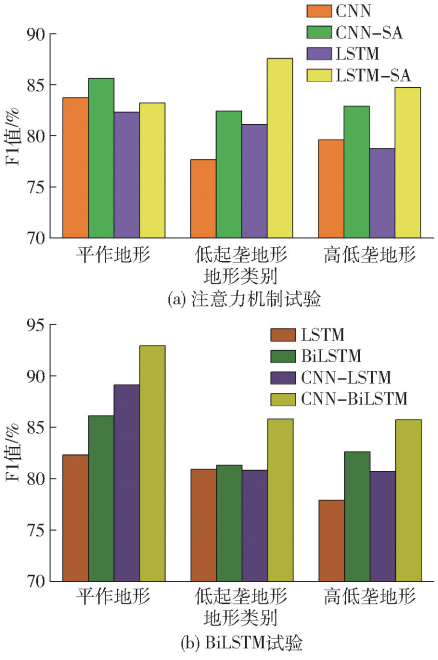


图 11　对比试验结果

Fig. 11　Comparison of experimental results

3.4.4　双向长短期记忆网络

使用 LSTM 模型可以更好地捕捉到较长距离的依赖关系，这是由于 LSTM 在训练过程可以进行记忆和遗忘信息学习。但是利用 LSTM 进行的地形分类存在一个问题：无法编码从后到前的信息。在更细粒度的分类时，如对于平作和起垄这两个类别及低起垄和高起垄两个类别来说差别均不明显，因此需对 IMU 数据的时间轴前向和后向进行特征提取。

BiLSTM 通过将 LSTM 层沿着时间轴前向和后向运行来计算隐藏状态。两个方向的隐藏状态被连接在一起形成最终的双向隐藏状态。通过 BiLSTM 可以更好地捕捉双向的语义依赖，能捕捉到序列中每个时间步之前和之后的上下文信息，从而提供更全面的特征表示。如图 11b 所示，研究结果表明，BiLSTM 的加入提高了所有场景下的识别效率。BiLSTM 和 CNN－BiLSTM 模型在信号数据集上表现出更好的性能改进。

3.4.5　模型性能对比

除了上述识别精度评价指标外，重要的性能评估指标还包括样本识别时间、平均识别精度及总体参数量等，表示模型进行训练和测试所花费的时间。表 5 展示了模型性能比较，其中包括表 4 中默认参数模型。

表 5　模型性能比较

Tab. 5　Comparison of model performance

模型	样本识别时间/s	平均精度/%	总体参数量
CNN	0.014	87.25	19 715
LSTM	0.012	79.23	70 019
BiLSTM	0.025	83.85	188 419
ABLSTM[28]	0.037	86.74	86 659
CNN－LSTM	0.006	88.79	1 063 043
CNN－LSTM－SA	0.004	90.87	975 427
CNN－BiLSTM	0.011	90.33	341 827
CNN－BiLSTM－SA	0.009	92.37	211 843
CNN－GRU[29]	0.007	89.26	88 547
CNN－GRU－SA[29]	0.006	90.75	80 811

如表 5 所示，对于样本识别时间，ABLLSTM 模型的识别时间最长，由于采用双向 LSTM 和注意力机制，在没有 1D－CNN 网络对 IMU 样本进行处理的情况下，庞大的计算量使其前后向的传播耗费了时间。虽然本文模型在单样本的处理时间上不具备优势，但本文 CNN－BiLSTM－SA 模型在平均精度上最高，其次是 CNN－LSTM－SA 和 CNN－GRU－SA，这表明模型在处理田间复杂情况下的信号时，效果最佳。对于模型总体参数量，CNN 与 LSTM 模型的参数量较小，这表明单模型运行的内存需求小于其他模型。虽然本文模型总体参数量较大，但模型最终要部署在田间机器人上进行田间实际试验，其配备的处理器为 NVIDIA JETSON NANO，能够适配本文模型的运行需求。

综上，综合单样本处理时间、识别精度和参数量，由于要将模型部署在田间机器人的控制板，面对田间的复杂环境，识别精度对于机器人的田间作物行分类更为重要。本文使用模型对当前作物行地形进行分类时，当模型连续维持 3 s 以上 90% 识别准确率时，此时判定当前地形类别，因此，相比于识别精度，实时性问题非本文重点，田间平作与垄作类别的识别精度指标对于本文模型至关重要。

3.5　不同地形下机器人姿态参数优选试验

3.5.1　姿态参数优选试验方法

为进行玉米冠下四足机器人在不同种植模式地形条件下的行走参数（包括机身高度和抬腿高度）优选，本节改变机器人行走参数，在冠下作物行中采集数据，保存 IMU 和足端速度数据作后续标准差分析，旨在不同行走参数组合下优化机器人行走过程

中的稳定性和效率。选择机器人机身高度和抬腿高度作为核心调节参数,以确保机器人在不同作物行地形条件下保持行走稳定性。较低的机身高度有助于降低重心并减少摔倒风险,机身高度设定为低(26 cm)、中(28 cm)2 个等级。较低的抬腿高度有助于保持步态稳定性和减少能耗,适用于平坦地形,但在起垄等复杂地形中,抬腿高度过低可能使机器人受起垄斜坡或地面障碍物影响无法完成足端正常摆动。因此,抬腿高度设定为低(6 cm)、中(8 cm)和高(10 cm)3 个等级。机器人的默认行走参数为机身高度 28 cm 和抬腿高度 8 cm。

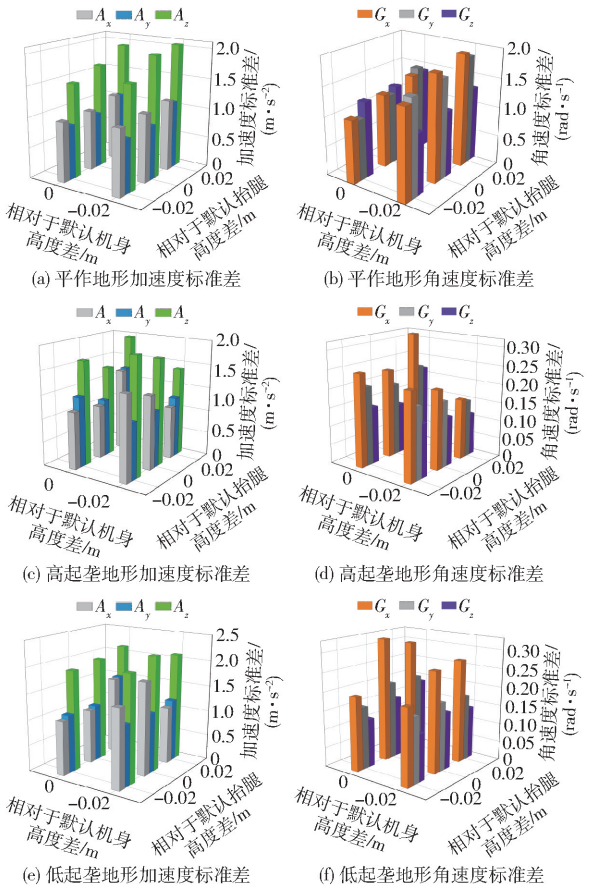
对 IMU 信号进行时域特征提取,引入信号的标准差作为机器人稳定性描述参数。标准差主要衡量信号离散程度,是反映数据波动程度的一个重要指标。当信号稳定性差时,该信号的离散程度往往会很高,标准值越大,表示偏离信号均值越远,使用姿态角数值标准差表示机器人行走时稳定性描述。

IMU 传感器可测得三轴的加速度(A_x,A_y,A_z)和角速度(G_x,G_y,G_z),由图 1a 可知机器人坐标轴方向。对同一姿态参数下采用标准差平均值 A_{ag} 表示机器人的稳定性,即

$$A_{ag}=\frac{A_x+A_y+A_z+G_x+G_y+G_z}{6} \tag{2}$$

3.5.2 姿态参数优选结果

进行不同机身行走姿态参数的行走试验,计算不同地形下的加速度与角速度标准差,如图 12 所示。与垄作的两种地形相比,机器人在平作地形行走的稳定性较高,加速度和角速度标准差都较小;在高起垄地形下,机器人 IMU 传感器数据抖动程度最高,这是由于机器人在高垄坡产生滑移的概率变大,增加了机身和传感器的数据抖动;由图 12f 可知,G_x 表示机身沿 X 轴的角速度,相比于其他地形,低起垄地形的 G_x 标准差明显偏大,相比于高起垄地形,低起垄地形对机器人的行进区间约束较小,机器人在沿作物行前进时,机器人出现沿 X 轴方向的更大程度摇晃,对机器人的行走稳定性产生了较大影响。



(a) 平作地形加速度标准差 (b) 平作地形角速度标准差
(c) 高起垄地形加速度标准差 (d) 高起垄地形角速度标准差
(e) 低起垄地形加速度标准差 (f) 低起垄地形角速度标准差

图 12 机器人加速度与角速度标准差柱状图

Fig. 12 Histograms of standard deviation of robot acceleration versus angular velocity

由图 12 和式(2)计算得到的标准差平均值如表 6 所示,其中机身高度表示为:26 cm(Ha)、28 cm(Hb);抬腿高度表示为:6 cm(Fha)、8 cm(Fhb)、10 cm(Fhc)。

表 6 不同地形标准差平均值

Tab. 6 Mean value of standard deviation in different terrains

平作地形						高起垄地形						低起垄地形					
Ha			Hb			Ha			Hb			Ha			Hb		
Fha	Fhb	Fhc	Fha	Fhb	Fhc	Fha	Fhb	Fhc	Fha	Fhb	Fhc	Fha	Fhb	Fhc	Fha	Fhb	Fhc
0.674	0.744	0.806	0.625	0.664	0.758	0.995	0.921	0.787	0.917	0.808	1.123	0.874	0.951	0.838	0.777	0.810	0.978

综合 3 种地形试验结果表明,确定机器人行走姿态参数优选,对于平作和低起垄地形,默认机身高度(28 cm)和较低的抬腿高度(6 cm)增加了机身稳定性,尤其对于低起垄地形,较低的抬腿高度能使机器人更稳定地在垄谷行走,减少机身沿 X 轴方向摇晃程度;对于高起垄地形,由于地形起伏较大,较低的机身高度(26 cm)能降低机身重心,较高的抬腿高度(10 cm)能减少机器人行走过程中与垄坡斜面的碰撞几率。起垄地形中,较低机身高度和适中抬腿高度是确保稳定性的关键。

3.6 机器人地形识别试验

基于以上研究基础,开展基于平作与起垄地形识别的机器人行走试验。设定模型对作物行垄作的识别精度阈值为 90%,当模型达到此作物行

地形识别精度且持续时间大于 3 s 时，完成机器人行走地形类别确认。根据具体地形分类结果，由运动控制器进行与当前地形匹配的机器人机身行走参数调整，在每种地形下进行 50 次重复试验，确保试验数据的可靠性。机器人识别与控制流程如图 13 所示。

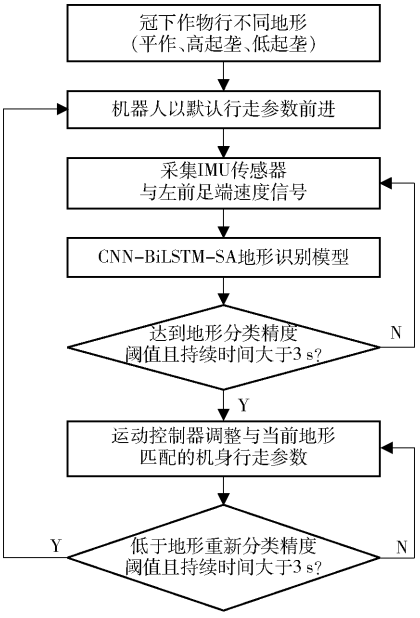


图 13　机器人地形识别与控制流程图

Fig. 13　Flowchart for terrain recognition and control of robot

将训练好的模型布置在机器人的 NVIDIA JETSON NANO 处理器中，进行田间试验。表 7 为机器人基于地形识别模型的田间性能结果，其中 T_R 表示模型推断地形结果所需的平均时间；T_{FR} 表示首次达到 90% 准确率平均时间；T_{ER} 表示维持 3 s 以上 90% 识别准确率平均时间。

表 7　机器人地形识别时间

Tab. 7　Terrain recognition time of robot　　s

地形类型	T_R	T_{FR}	T_{ER}
平作地形	0. 133	1. 184	4. 616
高起垄地形	0. 126	0. 784	4. 245
低起垄地形	0. 138	1. 359	4. 780

结果表明，模型部署在机器人控制板的运行速度能保证地形识别的实时性，在 0. 138 s 内可实现当前行走地形的识别。在机器人行走稳定后，模型能在 1. 4 s 内实现对当前地形 90% 的识别准确率，在 4. 8 s 内达到本文维持 3 s 以上 90% 识别准确率的地形分类要求，此后机器人调整行走机身参数并实现当前作物行的作业或巡检。

4　结论

（1）农田种植方式地形类别的识别对农业机器人的行走稳定性具有重要意义。为降低采集的 IMU 传感器和足端速度信号中噪声对于模型输入数据的干扰，针对不同信号数据噪声幅值不同的问题，采用不同大小的高斯核对信号进行滤波处理；根据信号波动规律对数据信号进行了窗口切分，并对信号数据进行了 50% 的重叠窗口滑动，以增强模型训练效率和时间序列信号关联性。

（2）为提高采集传感器信号数据的特征提取能力，提出了结合双向 LSTM 融合与注意力机制的 CNN－BiLSTM－SA 模型。注意力层添加了权重系数的学习，增强模型在 LSTM 中通过学习结果给予重要参数更大权重的能力；双向 LSTM 沿着时间轴前向和后向计算双向隐藏状态，能够有效增强网络的上下文理解能力，捕捉更长距离的依赖关系，提高序列建模的性能。本文 CNN－BiLSTM－SA 模型在地形类别识别任务上具有更高的准确率，与 CNN 模型、CNN－LSTM 模型、CNN－LSTM－SA 模型、CNN－BiLSTM 模型相比，本文模型 F1 值分别提升 10. 17、3. 51、2. 57、1. 27 个百分点，验证了模型的有效性。

（3）搭载平作与垄作识别模型的机器人在行走状态稳定后，能在 1. 4 s 内实现对当前作物行地形 90% 的识别准确率，在 4. 8 s 内达到对当前作物行的地形分类确认，此后机器人调整行走机身参数并实现当前作物行的作业或巡检。

参　考　文　献

[1]　陈学庚，温浩军，张伟荣，等. 农业机械与信息技术融合发展现状与方向[J]. 智慧农业（中英文），2020，2(4)：1－16.
CHEN Xuegeng，WEN Haojun，ZHANG Weirong，et al. Advances and progress of agricultural machinery and sensing technology fusion[J]. Smart Agriculture，2020，2(4)：1－16. (in Chinese)

[2]　韩佳伟，朱文颖，张博，等. 装备与信息协同促进现代智慧农业发展研究[J]. 中国工程科学，2022，24(1)：55－63.
HAN Jiawei，ZHU Wenying，ZHANG Bo，et al. Equipment and information collaboration to promote development of modern smart agriculture[J]. Strategic Study of CAE，2022，24(1)：55－63. (in Chinese)

[3]　张伟荣，陈学庚，齐江涛，等. 基于深度学习和高斯过程回归的玉米冠下视觉导航路径提取方法[J]. 农业机械学报，2024，55(7)：15－26.
ZHANG Weirong，CHEN Xuegeng，QI Jiangtao，et al. Deep learning and Gaussian process regression based path extraction for visual navigation under canopy[J]. Transactions of the Chinese Society for Agricultural Machinery，2024，55(7)：15－26. (in Chinese)

[4]　张庆，潘烤鑫，王振宇，等. 复杂地形下仿生轮腿式机器人位姿控制研究[J]. 农业机械学报，2024，55(6)：380－391，403.

ZHANG Qing, PAN Kaoxin, WANG Zhenyu, et al. Pose control of biomimetic wheel-legged robots in complex terrain[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(6): 380 - 391, 403. (in Chinese)

[5] 张春燕,刘玉航,丁兵,等. 可重构仿生四足机器人倾覆后恢复机理与特性研究[J]. 农业机械学报,2024,55(2):433 - 441,458.
ZHANG Chunyan, LIU Yuhang, DING Bing, et al. Recovery mechanism and characteristics of reconfigurable bionic quadruped robot after overturning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 433 - 441, 458. (in Chinese)

[6] 张宏鸣,常毅,杨勤科,等. 基于 SRTM 的地形因子提取方法研究[J]. 农业机械学报,2022,53(1):205 - 214.
ZHANG Hongming, CHANG Yi, YANG Qinke, et al. Topographic factor extraction method based on SRTM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(1): 205 - 214. (in Chinese)

[7] 汪沛,冯达文,陈高隆,等. 农田精准平整过程中三维地形实时测量方法研究[J]. 农业机械学报,2023,54(3):41 - 48.
WANG Pei, FENG Dawen, CHEN Gaolong, et al. Real-time 3D terrain measurement method and experiment in farmland leveling [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(3): 41 - 48. (in Chinese)

[8] MEI M, CHANG J, LI Y, et al. Comparative study of different methods in vibration-based terrain classification for wheeled robots with shock absorbers[J]. Sensors, 2019, 19(5): 1 - 18.

[9] KERTESZ C. Rigidity-based surface recognition for a domestic legged robot[J]. IEEE Robotics and Automation Letters, 2017, 1(1): 309 - 315.

[10] CHEN M, DAI X, LV Z, et al. Terrain classification in field environment with grid-based SURF feature[C]//2014 33rd Chinese Control Conference. IEEE, 2014.

[11] WANG W, ZHANG B, WU K, et al. A visual terrain classification method for mobile robots' navigation based on convolutional neural network and support vector machine[J]. Transactions of the Institute of Measurement and Control, 2022, 44(4): 744 - 753.

[12] GUAN T, HE Z, SONG R, et al. TNES: terrain traversability mapping, navigation and excavation system for autonomous excavators on worksite[J]. Autonomous Robots, 2023, 47(6): 695 - 714.

[13] 王树泽. 搜救机器人的环境建图与路径规划技术研究[D]. 武汉:华中科技大学,2022.
WANG Shuze. Research on environment mapping and path planning of search and rescue robot[D]. Wuhan: Huazhong University of Science and Technology, 2022. (in Chinese)

[14] MCDANIEL M W, NISHIHATA T, BROOKS C A, et al. Ground plane identification using LiDAR in forested environments[C]// IEEE International Conference on Robotics and Automation. IEEE, 2010.

[15] LAIBLE S, KHAN Y N, BOHLMANN K, et al. 3D LiDAR-and camera-based terrain classification under different lighting conditions[J]. Autonomous Mobile Systems, 2012(1): 21 - 29.

[16] ZHOU S, XI J, MCDANIEL M W, et al. Self-supervised learning to visually detect terrain surfaces for autonomous robots operating in forested terrain[J]. Journal of Field Robotics, 2012, 29(2): 277 - 297.

[17] GAO H. Investigation of context determination for advanced navigation using smartphone sensors[D]. London: University College London, 2019.

[18] KOLVENBACH H, WISTH D, BUCHANAN R, et al. Towards autonomous inspection of concrete deterioration in sewers with legged robots[J]. Journal of Field Robotics, 2020, 37(8): 1314 - 1327.

[19] 陈青,刘晓东,周寒,等. 基于多维全局特征融合的移动机器人地形识别[J]. 计算技术与自动化,2023,42(2):20 - 24.
CHEN Qing, LIU Xiaodong, ZHOU Han, et al. Research on terrain recognition of mobile robot based on multidimensional global feature fusion[J]. Computing Technology and Automation, 2023, 42(2): 20 - 24. (in Chinese)

[20] DUPONT E M, MOORE C A, COLLINS E G, et al. Frequency response method for terrain classification in autonomous ground vehicles[J]. Autonomous Robots, 2008, 24(4): 337 - 347.

[21] LV F, GAO H, BAI Y, et al. Extraction of speed-independent vibration features for terrain classification in lugged-wheel rovers [C]//2018 IEEE International Conference on Robotics and Biomimetics. IEEE, 2018: 1580 - 1585.

[22] HOEPFLINGER M A, REMY C D, HUTTER M, et al. Haptic terrain classification on natural terrains for legged robots[J]. Emerging Trends in Mobile Robotics, 2010, 1(1): 785 - 792.

[23] ZHU X, WANG M, RUAN X, et al. Adaptive motion skill learning of quadruped robot on slopes based on augmented random search algorithm[J]. Electronics, 2022, 11(6): 842.

[24] ZHANG Z, AN H, WEI X, et al. Unknown slope-oriented research on model predictive control for quadruped robot[J]. Machines, 2023, 11(2): 133.

[25] BEDNAREK J, BEDNAREK M, KICKI P, et al. Robotic touch: classification of materials for manipulation and walking[C]// 2019 2nd IEEE International Conference on Soft Robotics. IEEE, 2019: 527 - 533.

[26] 王晓磊,金振林,李晓丹,等. 串并混联四足仿生机器人动力学建模与分析[J]. 农业机械学报,2019,50(4):401 - 412.
WANG Xiaolei, JIN Zhenlin, LI Xiaodan, et al. Dynamic modeling and analysis of serial-parallel hybrid quadruped bionic robot [J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(4): 401 - 412. (in Chinese)

[27] 陈久朋,李春磊,伞红军,等. 基于模型的四足机器人步态转换控制研究[J]. 农业机械学报,2024,55(3):431 - 440,451.
CHEN Jiupeng, LI Chunlei, SAN Hongjun, et al. Model based gait transition control for quadruped robots[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(3): 431 - 440, 451. (in Chinese)

[28] CHEN Z, ZHANG L, JIANG C, et al. WiFi CSI based passive human activity recognition using attention based BLSTM[J]. IEEE Transactions on Mobile Computing, 2018, 18: 2714 - 2724.

[29] SHALABY E, ELSHENNAWY N, SARHAN A. Utilizing deep learning models in CSI-based human activity recognition[J]. Neural Computing and Applications, 2022, 34(8): 5993 - 6010.