

基于 CNN-LSTM 的欠驱动采摘机械手抓取力控制方法

胡朝晖¹ 董艳珂¹

HU Zhaohui DONG Yanke

摘 要

现有研究因未充分考虑目标特性与动态交互作用,易导致抓取失衡、滑脱和果实损伤等问题。为此,文章提出一种基于 CNN-LSTM 的抓取力控制方法。利用 CNN 提取目标图像特征,解决传统方法对水果特性建模不足的问题;通过 LSTM 捕捉接触力、力矩与扭矩的长期动态平衡关系,优化静力学模型预测能力;结合非线性反馈与逆动力学前馈补偿,提升欠驱动系统的控制鲁棒性。实验结果表明,所提方法在稳定性、角偏差、滑落次数及果实损伤方面均显著优于对比方法,验证了其在高动态采摘场景中的有效性。表明 CNN-LSTM 模型通过多模态数据融合与动态关系学习,为欠驱动机械手的精准抓取控制提供了新思路,兼具理论价值与工程应用潜力。

关键词

CNN-LSTM; 抓取力控制; 长期依赖特征; 非线性补偿项; 欠驱动

doi: 10.3969/j.issn.1672-9528.2025.08.015

0 引言

采摘机械手作为自动化设备的核心,在农业等领域广泛应用,可提升效率、降低成本并减少人工误差。控制力是其性能的关键因素,优化抓取力控制方法能显著提高采摘精度和效率,降低生产成本,改善农产品质量^[1]。不仅可以降低农业生产的成本,提高农作物的质量和产量,还可以为其他

领域的自动化设备提供一种新的控制方法和技术支持,这对于提高我国在机器人技术领域的国际竞争力具有重要意义。

梁栋等人^[2]利用闭环矢量方程建立机械手速度映射模型,并基于拉格朗日方程构建逆刚体动力学模型,设计 PD 补偿力矩控制方法,但未充分考虑复杂形状物体的交互作用,导致抓取稳定性不足。杨明博等人^[3]基于拉格朗日方程建立运动模型,设计自适应滑膜函数调整连杆参数,并通过李雅普诺夫稳定性算法优化抓取力控制,但仍难以完全适应复杂

1. 郑州工商学院信息工程学院 河南郑州 451400

[7] LI C Y, LI L L, JIANG H L, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. (2022-09-07)[2024-06-12]. <https://doi.org/10.48550/arXiv.2209.02976>.

[8] REN S Q, HE K M, SUN J, 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.

[9] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Piscataway: IEEE, 2023: 7464-7475.

[10] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2025-01-22]. <https://doi.org/10.48550/arXiv.1804.02767>.

[11] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4:

optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2024-11-23]. <https://doi.org/10.48550/arXiv.2004.10934>.

[12] ZHANG Y, GUO Z Y, WU J Q, et al. Real-time vehicle detection based on improved YOLOv5[J]. Sustainability, 2022, 14(19): 12274.

[13] HAN K, WANG Y H, TIAN Q, et al. GhostNet: more features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 1580-1589.

[14] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Computer Vision—ECCV 2018. Berlin: Springer, 2018: 3-19.

【作者简介】

宋江民(2001—),男,陕西商洛人,硕士研究生,研究方向:目标检测。

(收稿日期:2025-03-19 修回日期:2025-07-31)

水果的平衡需求。王通等人^[4]通过关节角度约束和扭簧刚度优化静力学模型,结合 ADAMS 动力学仿真计算抓取力,但关节限制位置未考虑水果特殊形状,影响抓取稳定性。Yu 等人^[5]提出增强自适应滑模控制器,利用 TDE 技术减少估计误差,但建模过程忽略水果特性,导致运动状态预测不准确。这些方法在欠驱动环境下均存在对复杂形状水果适应性不足的问题,影响抓取平衡性。

为了解决上述方法存在的问题,提出一种基于 CNN-LSTM 的欠驱动采摘机械手抓取力控制方法。

1 基于 CNN-LSTM 抓取力控制方法设计

为了在采摘任务中确保机械手施加足够的抓取力以防止物体滑脱,同时平衡力矩和扭矩,实现可靠而稳定的抓取,对抓取力的准确控制至关重要^[6-7]。因为过小或过大的抓取力都可能导致失败或损坏,过小的抓取力可能无法确保牢固抓住物体,导致物体滑脱,过大的抓取力则可能造成物体变形或损坏。设计 CNN-LSTM 网络模型是一种有效的方法, CNN-LSTM 网络模型结合了 CNN 和 LSTM 的优势,利用 CNN 提取采摘目标图像的特征,并利用 LSTM 捕捉接触力、力矩和扭矩之间的静力学平衡关系的长期依赖。这种设计可以使抓取力的控制更加精确和稳定,并能够适应不同的工作环境和物体特性^[8-9]。

在 CNN 模型中,通过使用 ReLU 函数和 sigmoid 函数作为激活函数的两层结构,实现抓取目标图像特征的提取。经过卷积操作后,获得长度为 j_c 的特征序列,对其展开最大池化操作,使特征序列长度缩短,完成抓取目标图像特征向量 $\mathbf{b} = (v_1, v_2, \dots, v_n)$ 的提取。

将这些特征向量的输出结果再次输入到 LSTM 模型中, LSTM 模型能够捕捉时间序列中的长期依赖关系,用于预测完成抓取过程中持续时间内的抓取力。LSTM 模型的核心思想是利用上一时刻 $g-1$ 的细胞单元状态对当前时刻 g 的细胞单元状态作出决定。欠驱动采摘机械手抓取力预测,需要对当前时刻 LSTM 模型隐藏层的输出单元状态 $q(g)$ 展开计算:

$$q(g) = [\log q(g-1) \cdot d(g)] \times j_e \mathbf{b} \quad (1)$$

式中: $q(g-1)$ 代表在 $g-1$ 时的输出单元状态; $d(g)$ 代表 g 时输入单元的特征向量。

对输入门限 $u(g)$ 、遗忘门限 $f(g)$ 与输出门限 $A(g)$ 的值展开计算,公式为:

$$\begin{cases} f(g) = \sigma(E_{(f)} \cdot d(g) + I_{(f)} \cdot q(g-1) + N_{(f)}) \\ u(g) = \sigma(E_{(u)} \cdot d(g) + I_{(u)} \cdot q(g-1) + N_{(u)}) \\ A(g) = \sigma(E_{(A)} \cdot d(g) + I_{(A)} \cdot q(g-1) + N_{(A)}) \end{cases} \quad (2)$$

式中: $N_{(A)}$ 代表输出门的偏置; $N_{(u)}$ 代表遗忘门的偏置; $N_{(f)}$ 代表输入门的偏置; $E_{(f)}$ 与 $I_{(f)}$ 代表输入门的权重; $E_{(u)}$ 与 $I_{(u)}$

代表遗忘门的权重; $E_{(A)}$ 与 $I_{(A)}$ 代表输出门的权重。

在输入门中,首先通过 sigmoid 激活函数对需要更新的细胞单元状态作出判断,然后在 tanh 层对原始单元状态展开创建:

$$\tilde{e}(g) = \tanh(E(e) \cdot d(g) + I(e) \cdot q(g-1) + N(e)) \quad (3)$$

在此基础上,对原始的单元状态展开更新,获得更新后的状态:

$$e(g) = f(g) \cdot e(g-1) + u(g) \cdot \tilde{e}(g) \quad (4)$$

在输出门中,首先通过 sigmoid 函数对需要输出的细胞信息展开判断,然后利用 tanh 函数获得当前时刻隐藏层的输出:

$$q(g) = A(g) \cdot \tanh(e(g-1)) \quad (5)$$

在上述基础上,利用 softmax 激活函数对隐藏层的权重向量 α 展开归一化处理,获得处理后的权重向量 $\alpha(u)$,计算 $\alpha(u)$ 与 $q(g)$ 的乘积,即平稳时间序列内完成抓取的长期依赖抓取力的预测结果:

$$w(g) = \sum \alpha_u q(g) \quad (6)$$

根据预测的抓取力和目标抓取力之间的差异,可以计算出抓取力误差 $i(t)$,为了对抓取力误差进行更进一步的分析,可以将其进行时间求导:

$$\begin{cases} i(t) = (\theta_c(t) - \theta(t)) w_g \\ \dot{e} = (\theta_c - \theta) w_g \end{cases} \quad (7)$$

式中: θ_c 代表机械手关节的角速度; t 代表时间。

基于 CNN-LSTM 的欠驱动采摘机械手抓取力控制需要结合非线性反馈与逆动力学前馈对机械手驱动电机进行调整,完成误差补偿。抓取力控制框图如图 1 所示。

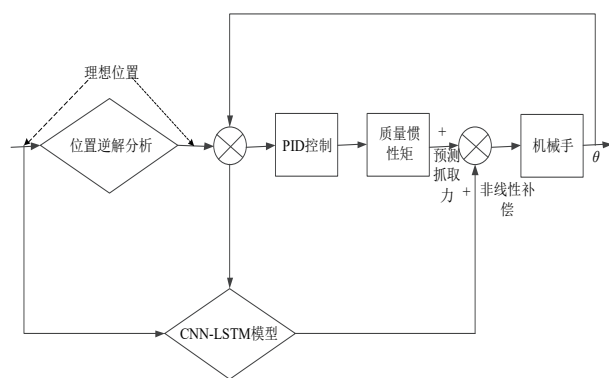


图 1 抓取力控制框图

对抓取力的非线性补偿项 τ_e 展开计算:

$$\tau_e = B(R_c i(t) + \dot{e} R_i i + R_o \int e dt) \quad (8)$$

式中: B 代表质量惯性矩; R_c 、 R_i 、 R_o 代表控制器的常数增益矩阵。

当增益矩阵设置合理时,通过多次迭代和搜索,找到能

够获得最优抓取力控制状态的控制参数,即可实现抓取力的最优控制^[10]。

2 实验操作与结果分析

为验证基于 CNN-LSTM 的欠驱动采摘机械手抓取力控制方法的有效性,进行实验测试。实验环境如图 2 所示,在实验现场中进行采摘机械手运动仿真实验。使用 Solidworks 软件建立采摘机械手的三维实体模型,在 ADAMS 中新建一个工程,将从 Solidworks 导出的机械手模型转换导入 ADAMS 中,构建虚拟样机模型。根据机械手的运动自由度和约束条件,在 ADAMS 中对机械手中的关节添加驱动,为更好地体现欠驱动性质,将驱动限制在少于关节数目的执行器上。进行仿真前,需要设定各构件属性:腰关节设定为 2.78 kg/m、肩关节设定为 1.48 kg/m、腕关节设定为 0.58 kg/m。设置单位模式为 MKS,重力加速度项为 9.8 N/kg。此外,还需设定所有关节最大接触力控制目标值为 15 N。在 ADAMS 仿真中运行实验,观察机械手在仿真环境中的运动行为。通过分析仿真结果,可以评估采摘机械手在实际采摘环境中的性能表现和抓取力控制效果。



图 2 采摘机械手现场采摘环境

本实验旨在评估基于 CNN-LSTM 的欠驱动采摘机械手抓取力控制方法在抓取平衡状态下的性能,并将其与文献[2]方法和文献[3]方法进行对比测试。

2.1 稳定性

采用所提方法、文献[2]方法和文献[3]方法对机械手力抓取力稳定性展开测试,对比结果如图 3 所示。

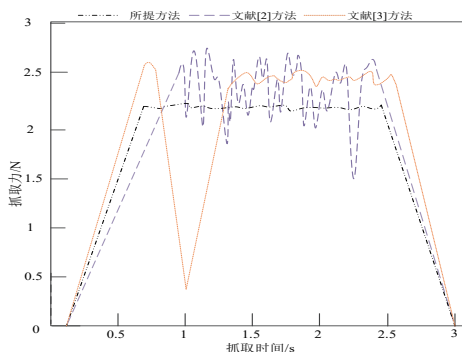


图 3 稳定性对比结果

由图 3 可知,文献[2]方法的抓取力波动较大,文献[3]方法出现控制中断,效果不稳定,而所提方法的稳定性最高,这是因为 LSTM 有助于准确估计机械手静力学平衡关系的抓取力,并捕捉到可能存在的稳定性问题。通过使用 LSTM 网络模型,可以更好地控制和调整机械手的抓取力,提高抓取过程中的稳定性。

2.2 角偏差

采用所提方法、文献[2]方法和文献[3]方法对机械手的角偏差展开测试,对比结果如图 4 所示。

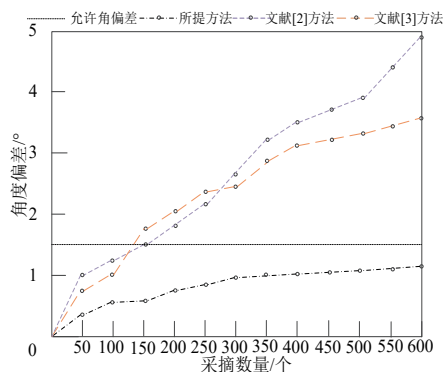


图 4 角偏差对比结果

根据图 4 可知,文献[2]方法与文献[3]方法在采摘时角偏差过大,易出现无效抓取现象,而所提方法通过考虑非线性因素对控制的影响,能够有效地纠正系统的偏差,并实现精准控制。这一步骤中,结合了静力学模型、CNN-LSTM 网络模型的输出和其他相关信息,综合考虑了各种因素,使得控制下的角偏差始终低于允许的角偏差,证明所提方法的采摘较为精准。

2.3 滑落次数

采用所提方法、文献[2]方法和文献[3]方法对机械手抓物体的滑落次数展开测试,对比结果如表 1 所示。

表 1 三种方法被抓物体的滑落次数

抓取物体质量 /g	滑落次数 /s		
	所提方法	文献[2]方法	文献[3]方法
100	0	3	0
110	0	3	0
120	0	1	0
130	0	1	0
140	0	1	0
150	0	0	1
160	0	0	1
170	0	0	1
180	0	0	1
190	0	0	2
200	0	1	3
210	0	3	4

根据表 1 可知, 所提方法抓取时无滑落次数, 这是因为设计的 CNN-LSTM 网络模型中的 LSTM 网络能够适应动态抓取场景, 提高了抓取力的稳定性。相比之下, 文献 [2] 方法在被抓物体质量较轻与较重时皆有滑落, 文献 [3] 方法在被抓物体较重时滑落次数过多。

2.4 损伤实验

将表皮未损伤的苹果作为采摘物体, 用于对比所提方法、文献 [2] 方法和文献 [3] 方法施力造成的苹果表面损伤, 结果如图 5 所示。

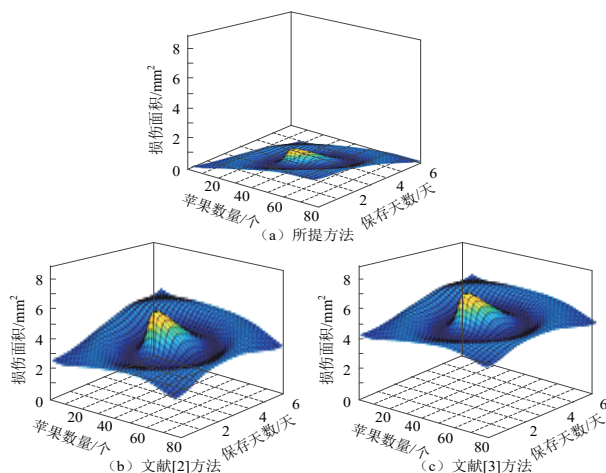


图 5 苹果损伤面积统计结果

采摘机械手施力造成的苹果内部挤压损伤, 随着时间流逝会表现在苹果表面。根据图 5 的对比结果可知, 文献 [2] 方法和文献 [3] 方法的苹果表皮出现大面积损伤, 而所提方法造成的损伤很小, 因为所提方法通过 CNN-LSTM 网络模型对抓取力进行精确控制。该模型中 CNN 网络可以准确感知和理解目标水果的视觉信息, 包括大小、形状等特征。这有助于执行精确的抓取动作, 以避免对苹果造成过大的外部压力和损伤。在控制抓取力时, LSTM 网络能够学习和预测接触力、力矩和扭矩之间的静力学平衡关系。通过对这些关系的准确把握和调整, 可以更好地控制抓取力的大小和分布, 从而减小苹果损伤面积。

3 结论

采摘机械手作为一种先进的采摘设备, 已在采摘领域广泛应用。然而, 欠驱动环境下, 机械手控制所需的自由度 (即可控制的方向数) 多于可供控制的驱动器数量。传统控制方法无法实现全面的位置和力的控制, 这导致机械手在抓取过程中无法充分适应并保持水果的平衡性, 容易出现滑落或损坏的情况。为此, 本研究提出一种基于 CNN-LSTM 的欠驱动采摘机械手抓取力控制方法。实验结果表明, 通过使用 CNN-LSTM 网络模型, 本文方法在稳定性、角偏差、滑落次

数和损伤实验等方面都表现出优异的性能。本文方法能够很好地控制和调整机械手的抓取力, 以提高抓取过程中的稳定性。同时, 本文方法还能够校正系统在抓取过程中的偏差, 实现精准控制, 减小苹果损伤面积, 几乎不造成苹果内部挤压损伤, 从而保护物体的完整性。为日后实现根据物体形状进行自主抓取奠定了基础。

参考文献:

- [1] 张凤英. 物联网下并联机械手级联式控制仿真 [J]. 计算机仿真, 2020, 37(11): 348-351.
- [2] 梁栋, 梁正宇, 杨建成, 等. 提综臂辅助旋铆机械手动力学建模及控制仿真 [J]. 天津工业大学学报, 2022, 41(1): 81-88.
- [3] 杨明博, 周紫阳. 基于滑模自适应控制的双关节机械手轨迹跟踪 [J]. 组合机床与自动化加工技术, 2021(6): 81-84.
- [4] 王通, 尹建军. 一种抓取果实的欠驱动手指机构设计与静力学分析 [J]. 农机化研究, 2016, 38(3): 110-114.
- [5] YU X Y, GUO J Q, ZHANG J Y. Time delay estimation-based reactionless augmented adaptive sliding mode control of a space manipulator's pregrasping a target [J]. Robotica: international journal of information, education and research in robotics and artificial intelligence, 2022, 40(9): 3136-3156.
- [6] 陈金明, 钟盛华, 刘逸平, 等. 地铁扣件 III 型弹条的静力学模型与参数优化分析 [J]. 科学技术与工程, 2023, 23(22): 9650-9658.
- [7] 杨亚芳, 梁茂林. 非奇异 H 矩阵判定的充分条件 [J]. 宝鸡文理学院学报 (自然科学版), 2021, 41(3): 20-25.
- [8] 凌家源, 彭勇刚. 基于事件检测与 CNN 模型的非侵入式负荷识别方法及实现 [J]. 电工电能新技术, 2021, 40(3): 46-54.
- [9] 杨婷婷, 蓝流剑, 庄志宝, 等. 基于 LSTM 模型的直接空冷机组多目标分布式最优控制研究 [J]. 中国电机工程学报, 2022, 42(S1): 204-214.
- [10] 陈华, 刘新宇, 赵学峰. 基于微分法串联机器人的误差敏感度分析 [J]. 江苏大学学报 (自然科学版), 2020, 41(3): 353-358.

【作者简介】

胡朝晖 (1995—), 女, 河南商丘人, 硕士, 助教, 研究方向: 自然语言处理、计算机视觉。

董艳珂 (1996—), 女, 河南周口人, 硕士, 助教, 研究方向: 计算机视觉、计算机网络。

(收稿日期: 2025-04-14 修回日期: 2025-08-05)