



# 第二届人本智造学术会议

暨世界智能制造大会人本智造专题活动

The 2nd Human-Centric Intelligent Manufacturing Academic Conference

南京 · 2024年12月20日-22日



东南大学  
SOUTHEAST UNIVERSITY

## 智能人机交互中的多层级意图识别框架

张瑞升<sup>1</sup>, 裴旭益<sup>2</sup>, 韩己臣<sup>1</sup>, 吴航<sup>1</sup>, 黎明朗<sup>1</sup>, 周小舟<sup>1</sup>, \*

<sup>1</sup>东南大学机械工程学院, 中国南京

<sup>2</sup>中国航空无线电电子研究所, 中国上海

220220426@seu.edu.cn

### 摘要

意图识别是实现智能人机交互的前提和基础。复杂任务情境中, 人的意图常表现出多层级和上下文相关联的特点, 本文提出了一种适用于复杂任务的多层级意图识别框架, 实现由人类行为、机器状态和环境信息等底层数据逐层向上推理各层级意图, 并以直升机和无人机协同山火勘探任务为例情境, 验证该框架可用性。首先, 通过任务分析提取高层级任务意图、低层级交互意图和特征数据之间的意图结构; 之后, 采集16名参与者80个完整试次的飞行仿真实验数据; 最终通过1D-CNN+Bi-LSTM神经网络和动态贝叶斯网络分别实现交互意图和任务意图的可靠识别, 其中交互意图和任务意图最高识别准确率为94.57%和97.33%。本研究的方法和结果为复杂任务中人类意图识别提供了新角度, 并为智能自适应界面和复杂任务情境下的智能人机交互系统提供方法论指导和支撑。

### 介绍

针对人类行为、机器状态和外部环境动态变化的复杂任务和复杂人机交互系统, 本文提出了一种通用的多层级意图识别框架, 如图1所示。框架将意图根据时间颗粒度由大到小划分为任务意图和交互意图, 通过可捕获的实时人、机、环信息输入双层意图识别模型, 来获取具备解释性的双层意图识别结果。本文以具体的直升机无人机协同山火救援任务为例验证该框架的可用性。本研究的主要创新点如下:

- 多层级角度开展意图识别:** 在人类行为所反映交互意图基础上进一步逐层推理更深层的任务目标。
- 人机环系统角度开展意图识别:** 结合机器状态、环境态势等上下文和人类行为共同推理人的意图。
- 专家知识和特征数据双重驱动:** 利用专家知识分析任务和贝叶斯建模, 利用特征数据训练算法模型。

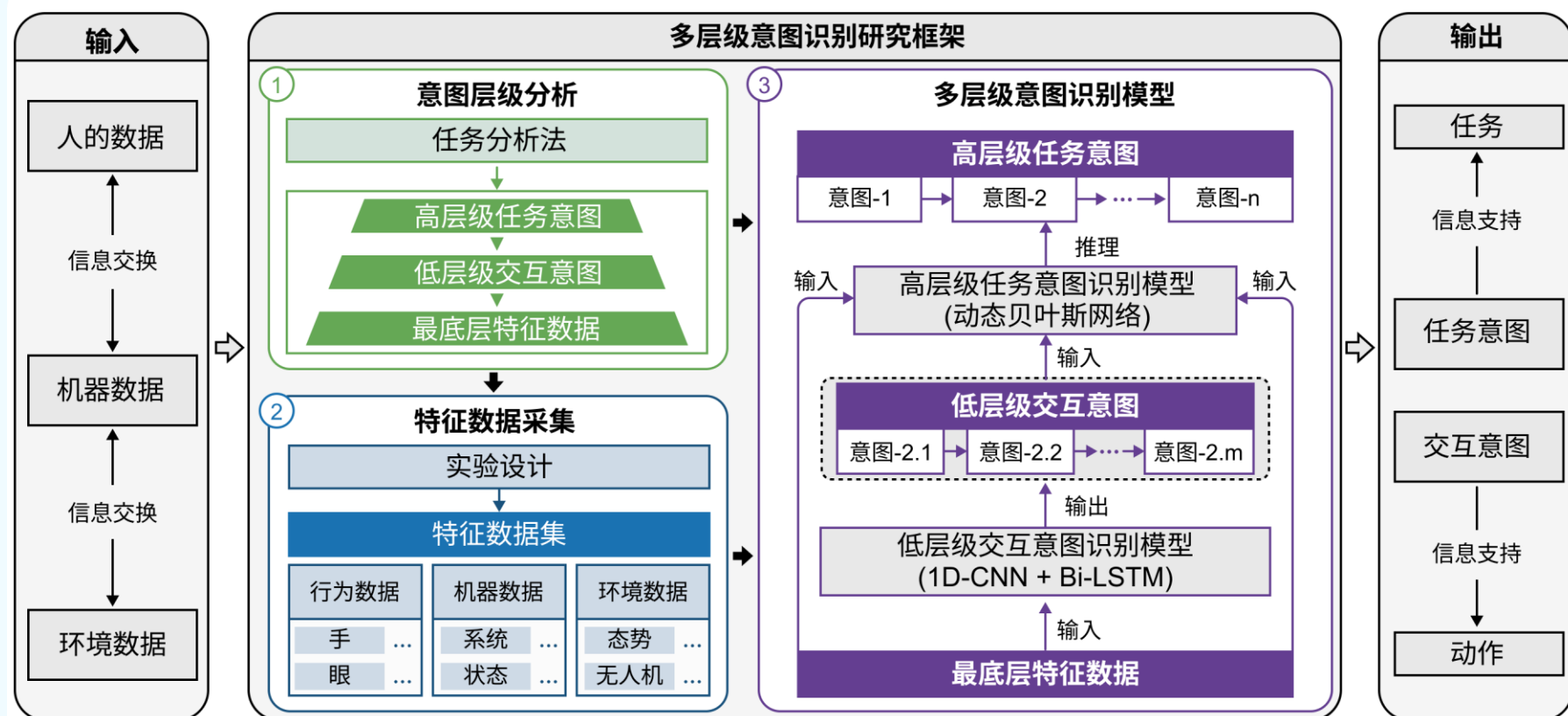


图1 多层级意图识别研究框架

### 方法

#### (1) 意图层级分析

采用层次目标分析(Hierarchical Goal Analysis, HGA)和操作序列图(Operational Sequence Diagrams, OSDs)相结合, 梳理操作员执行复杂任务过程的标准操作程序和多层级的任务目标, 进而提取人的多层级意图、意图关联特征, 如图2所示。

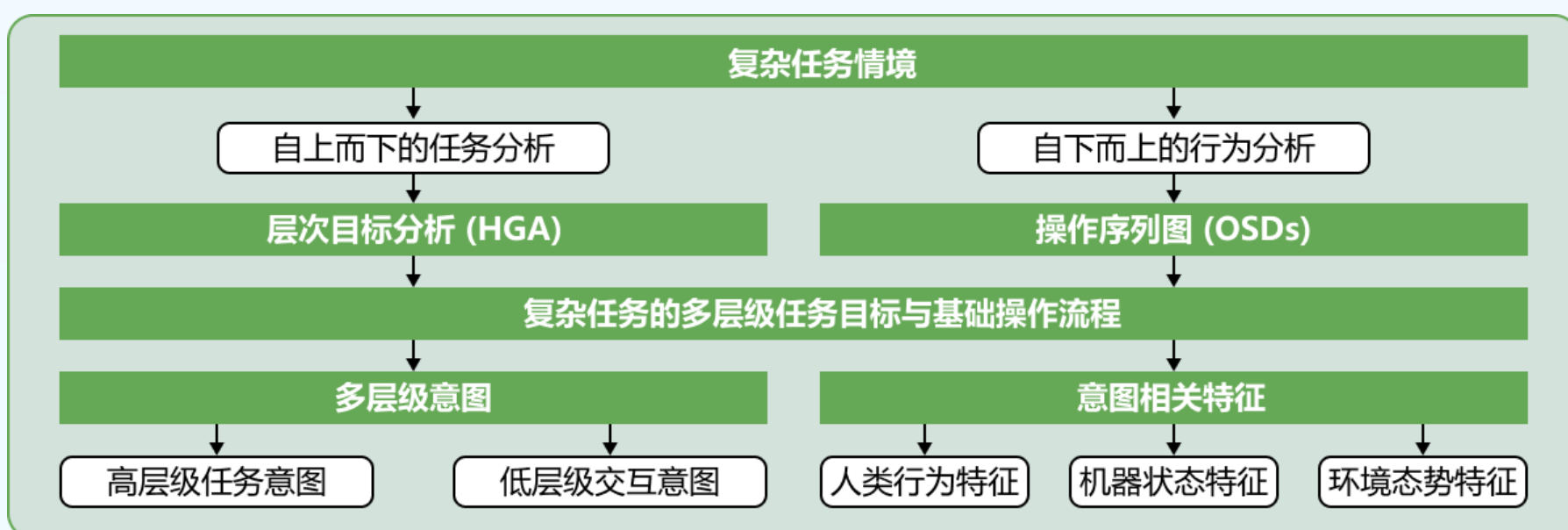


图2 多层级意图和关联特征分析流程

#### (2) 特征数据收集

邀请16名研究生被试参与飞行仿真驾驶实验, 实验界面和场景如图3所示。实验前, 被试接受专业飞行驾驶训练。实验中, 被试在专家监督下独立执行任务, 后台程序实时收集注视位置、AOI、手部按键、系统信息等数据; 此外, 被试被要求在满足基础任务目标前提下自主决定操作流程来保证数据的代表性和多样性。实验后, 根据实验视频回放标注各类意图的起始点和终止点。实验数据经数据清洗后汇总为结构化数据集。



图3 特征数据采集实验相关设置

#### (3) 多层级意图识别模型

融合专家建模和深度学习手段实现各层级意图识别, 如图4所示。采用1DCNN+Bi-LSTM神经网络识别交互层级意图, 并将交互意图识别结果、人类行为、机器和环境信息共同输入动态贝叶斯网络来识别任务层级意图。

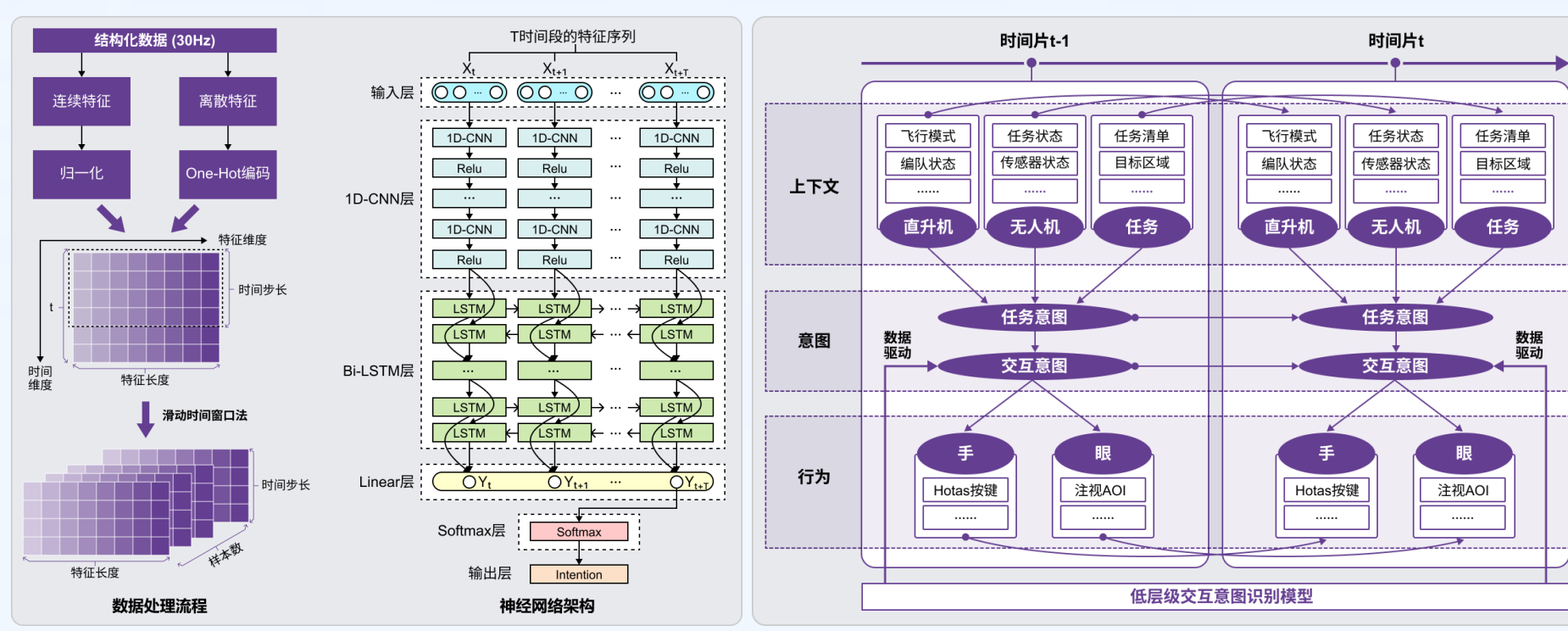


图4 多层级意图识别模型

### 结果

本框架在飞行仿真实验采集的结构化数据集基础上开展验证。结果显示, 在直升机无人机协同山火救援任务中, 本框架最终实现94.57%的低层级交互意图识别准确率和97.33%的高层级任务意图识别准确率, 如图5所示。具体结果如下:

#### (1) 低层级交互意图识别结果

针对16类交互意图, 模型最终取得94.57%总体准确率、93.21%精确率、81.27%召回率和84.47%F1分数。

#### (2) 高层级任务意图识别结果

针对5类任务意图, 模型最终取得97.33%总体准确率、96.89%精确率、94.44%召回率和95.59%F1分数。



图5 多层级意图识别混淆矩阵

### 讨论

上述结果表明, 本文提出的多层级意图识别算法框架能够在人、机、环信息动态变化的复杂任务中实现同时处理复杂多维时间序列数据和保持模型的可解释能力。本研究有如下发现:

- 适用于意图建模的动态贝叶斯网络:** 利用专家定义建立上下文-意图-行为的贝叶斯因果关系, 完成人类意图的可解释性建模; 通过输入交互意图的识别结果, 可进一步增强概率图模型识别效果。
- 机器状态和环境态势等上下文的关键作用:** 同一意图在不同上下文中可能传达不同任务目标, 综合上下文和行为才能保证建立机器对人的全面理解。

本研究的未来工作如下:

- 意图预测技术:** 考虑复杂系统信息处理延时, 未来将开发意图预测技术提前判别, 保证实时应用效果。
- 丰富输入特征:** 引入更丰富的眼动特征或脑电、肌电、心率等生理特征实现更精确的意图识别。

### 结论

本文研究成果为复杂任务情境意图识别提供了一种思路, 并为构建复杂任务情境下智能人机交互系统提供技术支撑。未来将在此基础上进一步开展意图预测研究, 实现对人类意图提前判别, 并给予实时智能信息支持, 实现实时意图驱动的智能人机交互系统应用。

### 致谢

本研究由国家自然科学基金资助[基金号: 52275238, 71901061]。特别感谢东南大学机械工程学院工业设计系团队老师和同学在研究过程的建议与帮助。

第二届人本智造学术会议

2024年·东南大学