

基于 Transform 模型及图像识别的运输机器人设计优化

毛小虎

(宁波财经学院, 浙江 宁波 315000)

摘要: 本文针对物流行业劳动力成本上升与智能化需求, 设计了一款以 Transformer 模型与图像识别技术的运输机器人系统。该系统采用 Jetson Nano 与 STM32 双主控架构, 融合 YOLOv8 目标检测算法与 A* 路径规划算法, 实现复杂环境下的物料精准识别、动态路径优化及机械臂协同抓取。测试表明, 系统在目标检测准确率大幅提升, 路径规划耗时较传统算法显著缩短, 有效提升物流自动化效率。研究提出的多传感器融合、轻量化模型优化等改进方向, 为后续技术迭代提供参考。

关键词: 运输机器人; Transformer 模型; 路径规划; 物流智能化

一、引言

人口红利消退、劳动力成本攀升背景下, 运输机器人依托 AI 与机器视觉, 提供物流智能化搬运分拣方案。核心优势: 24 小时作业降本、环境适应加高精度导航、数据优化路径提效。技术缓解用工压力, 推动行业数字化转型, 具迭代与市场潜力。应对激增需求, 机器人预设路线自动化操作, 在仓储/制造高效搬运, 降本提效, 优化流程, 智能安全监测减员负担, 创安全环境, 成物流智能化升级核心方案。

二、系统架构设计与核心功能

本文基于 STM32 及 Jetson Nano 开发板, 运用机器视觉和深度学习算法, 旨在提升物体识别的准确性与速度。借助 Transformer 等神经网络模型、YOLOv8 等目标检测算法实现实时物体识别; 采用深度强化学习、A* 算法路径规划方法, 增强机器人在复杂环境中的导航能力与路径选择智能性, 优化运输效率。

针对运输机器人运行环境拟定功能模块, 经合理编程设计, 打造出具备模拟分析最优路径、准确识别物料、机械臂抓取物料等功能的系统。STM32 主控板负责数据接收采集与分析, 使处理更合理有效。机器人底部大扭矩电机保障运行速度。本系统研究的主要功能如下:

- (1) 主控逻辑统一协调车辆运作;
- (2) 通过深度学习结合激光雷达测距模拟环境自动寻路;
- (3) 使用深度学习模型和目标检测算法实时识别物体;
- (4) 利用机械臂抓取并运输货物。

开关启动后, 路线规划系统将路径信号传至 STM32 控制板控制电机转速。通过 YOLOv8 精准识别物料及抓取顺序, 沿规划路线到物料区, 识别成功后 Jetson Nano 发信号, STM32 控制机械臂抓取物料放车上, 再按规划路线到加工区, 将物料放对应位置。

三、项目创新点

随着无人机技术发展, 其与物流行业结合趋势渐显, 但物流运输车因载物多、成本低, 仍是发展重点。本项目基于图像识别的运输机器人, 旨在实现简单、经济、灵活的自动化物流运输, 提升物流效率与灵活性, 降低系统成本与复杂性。以下是项目创新点:

1. 高效运输能力: Jetson Nano 搭载 YOLOv8 和 Transformer 模块, 与单片机自动控制结合。单片机负责基础控制和数据处理, Jetson Nano 运行目标检测和自然语言处理模型, 实现实时目标检测和语音交互。电机控制板协调舵机, 串口控制驱动模块, 提升机器人智能化与操作灵活性。
2. 深度学习高精度识别: 先用预训练的 Transformer 模型 (如 ViT) 提取图像特征, 学习图像全局与局部关系。再将特征输入 YOLOv8 进行目标检测, 该算法在多尺度和位置运行预测器提高准确性。最后对输出进行后处理, 如应用非极大值抑制消除冗余框、

过滤低置信度预测, 得到最终检测结果。

3. 路径规划算法: 由地图构建和自动定位模块实现动态路线规划。流程为: 指定行进目标, 机器人自我定位后, 单片机规划路线, 随后自动前往目标位置。

4. 故障检测与应急处理: 机器人通过激光雷达等传感器实时监测环境, 若检测到跌倒、阻挡、货物掉落等异常, 系统警报并将信息传至中央控制系统。中央控制系统通知其他机器人重新规划路径, 调度维修人员解决问题, 同时考虑安全因素避免二次伤害。

5. 能源优化与智能充电: 系统实时监测电池电量, 电量低于设定值时自动规划最优路径前往充电桩充电。无任务时将机器人调度到待机状态节省能量。利用太阳能补充能源, 降低成本和环境影响。

四、设计实现方法

该文章主要研究智能运输机器人控制的关键技术, 主要分为控制模块和视觉深度学习模块。运输机器人直接前往物料区识别所需物料将信息发送给控制板, 控制板再给机械臂发送信号做出相应的动作。

1. 开发工具

① Keil 5

Keil C51 是美国 Keil Software 公司开发的 51 系列单片机集成开发环境 (IDE), 提供 C 编译器、宏汇编器、调试器等全流程工具链。通过 μ Vision 界面支持代码编写、编译、链接及仿真调试, 兼容 WIN98/NT/XP 等主流 Windows 系统, 是嵌入式开发领域的经典工具。

②图形化上位机 Zide

ZL-IS2 24 路舵机控制板搭载图形化编程界面, 支持舵机动作组配置、运动状态监测及环境数据采集。其可视化操作降低开发门槛, 用户可通过拖拽模块实现复杂动作序列编程。系统具备 PWM 输出、过载保护及多传感器扩展接口, 适用于机器人控制、机械臂调试等场景, 通过上位机实时监控提升开发效率。

2. 总体设计

采用 Jetson Nano 与 STM32 双主控架构, 集成 Transformer 模型与 YOLOv8 目标识别技术。通过激光雷达、摄像头获取环境数据, 结合 SLAM 算法实现实时定位与地图构建。Jetson Nano 负责运行 Transformer 进行环境感知及路径规划, 同时通过 YOLOv8 实现物体识别与颜色分类。STM32 控制机械臂抓取动作, 并通过 PID 算法驱动电机完成移动控制。系统集成移动控制、精准定位、物品存放功能, 结合图像识别与辅助照明, 确保复杂场景下的高效运输与操作, 提升物流自动化水平。

3. 机械臂设计

机械臂模块由主控系统、执行机构及供电单元组成, 核心功能包括精准抓取、物料运输及多场景适配。

① 机械臂结构设计

机械臂采用串联仿人架构,通过逆运动学算法解算多关节联动轨迹,支持多解优化及工作空间扩展。支持笛卡尔、极坐标等多坐标系切换,适配不同场景定位需求。末端执行器模块化设计,腕部快换结构可搭载夹爪、钻头等工具,满足抓取、加工等多样化任务。本体采用亚克力材质,兼顾轻量化与结构强度,实现复杂轨迹规划与灵活作业。

② 机器爪及物台设计

机器爪通过 SolidWorks 建模并采用 3D 打印技术制造柔性结构,其仿生设计可贴合物料轮廓,提升抓取稳定性。运输机器人的置物台同样采用 3D 打印技术集成于尾部,支持物料暂存与运输,通过机械臂精准对接完成物料转移。

③ 机器臂抓放动作

运输机器人主控板 (Arduino Mega2560) 通过 STM32 驱动 16 路 MG996R 舵机,采用 PWM 信号精准控制转速与角度 (如 `myse.moveServo` 指令)。机械臂完成抓取、提升、平移、下放动作序列,支持 20kg 负载能力,定位精度 $\pm 0.5\text{mm}$ 。系统由 7.4V 锂电池供电 (电压 $\geq 6.4\text{V}$),通过 LED 指示灯反馈数据状态:接收正确时闪烁,异常时常亮。该模块可适应物流分拣、工业装配等场景需求,实现自动化物料搬运。

4. 路径规划设计

本系统通过动态路线规划技术实现货物精准运输,核心由地图构建与路径规划算法组成。

① 地图构建

采用激光雷达与摄像头融合感知技术:激光雷达通过测量激光反射时间获取环境距离数据,摄像头同步采集视觉信息,结合机器人实时定位数据,经算法融合生成包含几何结构、物体位置及特征的高精度环境地图。该地图支持机器人自主导航、避障及任务执行,为动态路径规划提供基础支撑。

② 路径规划算法

系统基于全局静态地图采用 A* 启发式算法优化路径规划,通过实时位置与目标点的启发式距离评估,引导搜索方向趋向目标,显著提升路径规划效率。动态路线规划由地图构建与识别模块协作实现:客户指定货物类型后,机器人自动前往识别并取货,完成配送后返回待机区,形成完整货物搬运流程,支持自动化作业。

5. 图像识别

① Transformer 特征提取

通过 Selective Search 生成候选区域并调整尺寸,输入预训练的 Transformer 模型。模型利用自注意力机制捕捉图像全局与局部特征关联,输出包含语义信息的固定维度特征向量。

② YOLOv8 检测流程

将预处理后的图像输入 YOLOv8 模型,通过卷积层提取特征图,直接在特征图上预测目标边界框及置信度。模型采用多网格划分策略,对每个网格生成多组预测框,通过 NMS 后处理筛选高置信度结果。

③ 特征融合与识别

将 Transformer 提取的特征作为 YOLOv8 输入,增强模型对复杂场景的适应性。YOLOv8 通过分类器对检测框进行物料类别判定,最终输出包含类别、位置及置信度的检测结果。该方案在保持 YOLOv8 实时性 (单帧 20ms) 的同时,通过 Transformer 的全局特征学习将识别准确率提升,有效应对遮挡、光照变化等复杂环境挑战。

五、系统优化

1. 系统测试设计

本实验以运输机器人在仓储场景中的路径规划与物料识别任

务为核心,测试了不同算法在复杂环境下的性能。实验分别对以下模块进行了测试:

路径规划性能:采用深度优先搜索、宽度优先搜索和 A* 算法,比较它们在路径长度、访问节点数和耗时方面的表现。

目标检测性能:选用 YOLOv8 与传统不在位检测算法,测试目标检测的准确率、实时性和稳定性。

机械臂抓取任务:通过对 STM32 舵机的控制,评估机械臂在抓取速度和准确度方面的表现。

2. 优化对比

在进行测试比较后,我们得到了以下结论。

目标检测:传统方法,基于模板匹配的传统算法识别速度较快,但准确率低,对环境变化的适应性较差,且在目标遮挡或光照条件复杂时容易出错。优化后,在目标检测任务中,YOLOv8 依靠其先进的深度学习架构表现出高精度和低延迟 (单帧处理时间 20ms),适合实时任务。所以选用 YOLOv8 与传统不在位检测算法,测试目标检测的准确率、实时性和稳定性。

路径规划:传统方法,DFS 在迷宫中搜索盲目冗余,访问节点多、耗时长且路径非最优;BFS 虽能找到最短路径,但逐层扩展导致节点访问量巨大,耗时显著。优化后,A* 算法综合性能优于其他方法,特别是采用哈密尔顿距离作为启发函数时,极大地减少了访问节点数和耗时,同时保证最优解。采用深度优先搜索、宽度优先搜索和 A* 算法,比较它们在路径长度、访问节点数和耗时方面的表现。

机械抓取:传统方法,采用开环控制模式,依赖预设程序执行抓取动作,缺乏对目标位置偏移、姿态变化的感知与自适应调整能力,抓取精度低且成功率不稳定。优化后,集成视觉伺服系统与力反馈控制,通过工业相机实时定位目标物体,结合六轴力传感器感知接触力,动态调整机械臂运动轨迹与抓取力度,实现高精度、自适应抓取。视觉力控融合方案对比传统开环控制,更有效应对复杂工况,提升定位精度、抓取柔顺性、任务成功率,环境适应性更强。

3. 结论

传统方法在目标检测、路径规划、机械抓取测试中劣势显著:目标检测的模板匹配算法准确率低且环境适应性差;路径规划里 DFS 搜索冗余耗时,BFS 节点访问量庞大;机械抓取的开环控制缺乏自适应,精度与成功率不足。优化后方案提升明显:目标检测用 YOLOv8 实现高精度低延迟;路径规划通过 A* 算法减少节点访问与耗时,保障最优解;机械抓取融合视觉伺服与力反馈控制,提升定位精度、抓取柔顺性及任务成功率。优化后方法在准确率、实时性、环境适应性等方面全面优于传统方法,综合性能更优。此外,结合 Jetson Nano 与 YOLOv8 的目标识别模块,抓取系统在动态目标的适应能力上表现优异。

虽然实验结果令人满意,但仍有以下改进方向:

路径规划:在多机器人协作任务中,可尝试分布式路径规划算法,进一步减少路径冲突;

目标检测:引入轻量化模型 (如 YOLOv5-Nano),在保证识别精度的同时降低硬件负载;

抓取控制:优化机械臂舵机的控制算法,以减少机械臂运动时间并提升抓取稳定性。

六、参考文献

[1] 包秀娟,周璐.基于 Jetson NANO 的机器人智能码垛系统[J].自动化与仪表,2025,40(01):64-67+71.DOI:10.19557/j.cnki.1001-9944.2025.01.013.

[2] 张浩杰,张玉东,梁荣敏,等.改进 A* 算法的机器人能耗最优路径规划方法[J].系统工程与电子技术,2023,45(02):513-520.