



DOI:10.12404/j.issn.1671-1815.2404411

引用格式:颜伟,黄冠鹏,高玉萍,等.仓库多AGV路径冲突问题研究综述[J].科学技术与工程,2025,25(18):7465-7474.

Yan Wei, Huang Guanpeng, Gao Yuping, et al. Review on path conflict issues among multiple AGVs in warehouse[J]. Science Technology and Engineering, 2025, 25(18): 7465-7474.

自动化技术、计算机技术

## 仓库多 AGV 路径冲突问题研究综述

颜伟,黄冠鹏,高玉萍,刘阳

(山东科技大学能源与矿业工程学院,青岛 266590)

**摘要** 近年来,中国制造业持续快速发展,越来越多的企业选择无人搬运车(automated guided vehicle,AGV)作为智能物流系统的核心设备。为保证仓储运行效率,解决AGV间运输路径冲突的问题越发得到学者们的关注。从两个角度对仓库多AGV路径冲突问题进行文献综述。首先,从冲突类型的角度,将研究问题分为碰撞问题和死锁问题,分析不同冲突类型下多AGV防撞策略的研究现状。其次,从模型求解算法的角度,将其分为启发式算法和强化学习算法,分析近年来两者在仓库多AGV路径冲突问题中的应用;最后,对现有文献进行总结,并提出未来仓库多AGV路径冲突问题的发展方向。

**关键词** 碰撞;死锁;多AGV;启发式算法;强化学习

**中图分类号** TP23 TP301.6; **文献标志码** A

### Review on Path Conflict Issues among Multiple AGVs in Warehouse

YAN Wei, HUANG Guan-peng, GAO Yu-ping, LIU Yang

(School of Energy and Mining Engineering, Shandong University of Science and Technology, Qingdao 266590, China)

**[Abstract]** In recent years, rapid development has been continuously observed in China's manufacturing industry, where AGVs automated guided vehicles have been increasingly adopted by enterprises as core equipment in intelligent logistics systems. To ensure the efficiency of warehouse operations, addressing the issue of transportation path conflicts among AGVs has garnered growing attention from researchers. A literature review on the issue of multi-AGV path conflicts in warehouses was conducted from two perspectives. First, from the perspective of conflict types, the research problems were categorized into collision problems and deadlock problems, and the current state of research on multi-AGV collision avoidance strategies under different conflict types was analyzed. Second, from the perspective of model-solving algorithms, the study divides the approaches into heuristic algorithms and reinforcement learning algorithms, analyzing their application in multi-AGV path conflict issues in warehouses in recent years. Finally, the existing literature was summarized, and future directions for addressing multi-AGV path conflicts in warehouses were proposed.

**[Keywords]** collision; deadlock; Multiple AGVs; heuristic algorithm; reinforcement learning

近年来,随着经济的快速发展和人民生活水平的提高,个性化的商品逐渐成为新的市场需求。旧有消费需求的发展和新增消费需求的兴起,给仓库的存储及运输带来了巨大的挑战。根据《2023年全国物流运行情况通报》<sup>[1]</sup>,全国社会物流总额352.4万亿元,同比增长5.2%,增速比2022年全年提高1.8个百分点,物流需求规模持续恢复向好,增速稳步回升。为了应对不断增长的新需求,越来越多的企业选择构建新的智能仓储系统。无人搬运车(automated guided vehicle,AGV)作为智能仓储系统中的核心设备,其稳定运行对仓储系统至关重要<sup>[2-3]</sup>。在多AGV运行环境下,一旦AGV间产生路

径冲突问题,将会严重影响整条线路的运行,因此,多AGV运输路径冲突问题近年来逐渐成为研究热点问题。

目前,中国在该领域已经涌现出一些典型团队和重要科学研究项目,例如,浙江工业大学鲁建厦教授团队提出了一种考虑分配规则的多台AGV调度方法(multi-AGV scheduling approach, MASA),并设计了在仓库中有多工作站的情况下,多台AGV生成无冲突工作路线的方法<sup>[4]</sup>,在仓储物流领域产生了广泛影响,为多AGV系统的路径规划和调度提供了新的思路和方法。另外,江苏大学汽车工程研究院与宝胜系统集成科技有限公司联合研

收稿日期:2024-06-13 修订日期:2025-03-17

基金项目:山东省自然科学基金(ZR2023QG073)

第一作者:颜伟(1980—),女,汉族,山东临沂人,博士,副教授。研究方向:工业工程、精益生产。E-mail:skd995501@sdust.edu.cn。

投稿网址:www.stae.com.cn

究项目<sup>[5]</sup>则针对智能车库(robot-based intelligent garage, RIG)中自主引导车(AGV)的任务执行优先级问题,提出了一种基于改进冲突搜索的路径规划模型,该模型主要由任务分配、单AGV路径规划、多AGV冲突检测与解决3个模块组成。其中,多AGV冲突检测与解决模块设计了冲突检测子模块和冲突解决子模块来消解路径冲突,从而生成多AGV无冲突路径集合,最终,有效提升了智能车库存取车效率,为智能车库对AGV路径规划的应用提供了新思路,具有重要的工程应用价值。这些研究都表明,合理的设计能有效保证智能仓储系统的稳定、高效运行。

现选取两个角度对仓库多AGV路径冲突问题进行阐述。一方面,从冲突类型的角度出发,对仓库多AGV路径冲突问题综述;另一方面,从模型求解算法角度,对现有求解算法进行分类。

## 1 多AGV路径冲突类型及影响因素

首先将多AGV路径冲突问题分为碰撞问题和死锁问题两种。统计相关文献,如图1所示。可以看出,近年来对两种冲突类型的研究总体呈上升趋势,说明随着新形势下企业对仓储能力要求的进一步提高,如何有效保证通道内AGV的稳定高效运行越来越受到关注。

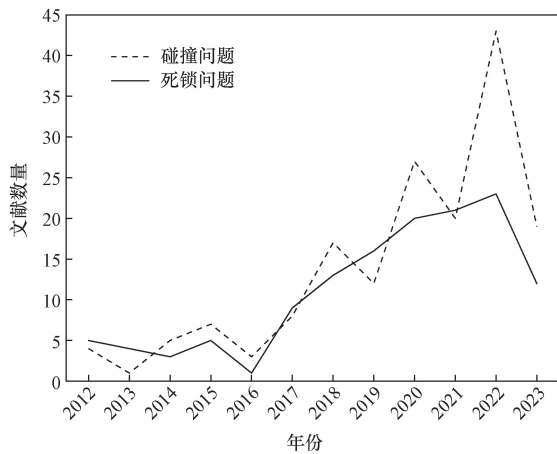


图1 仓库多AGV路径冲突问题文献统计图  
Fig. 1 Literature review chart on path conflict issues among multiple AGVs in warehouse

### 1.1 碰撞问题研究

碰撞是仓库多AGV运行时可能发生的一种事故,碰撞一般出现在路径的交点,也有可能由多AGV间的速度差导致<sup>[6]</sup>。几种常见的碰撞类型如图2所示。近年来,中外学者越发关注到多AGV运行环境下防碰撞的重要意义,并进行了相关研究。

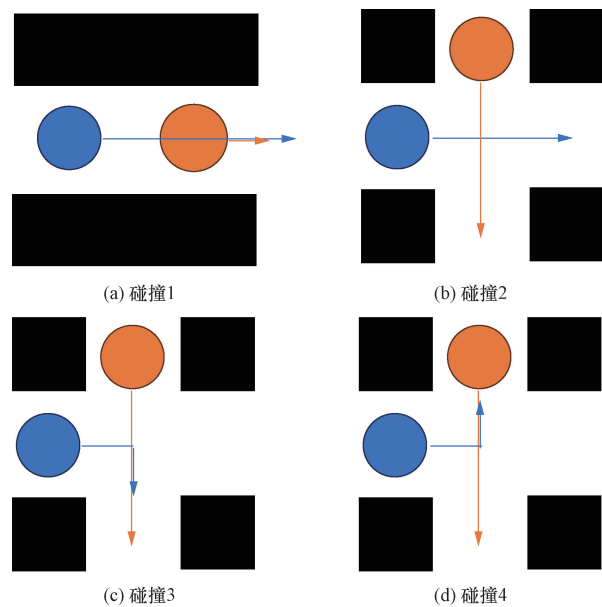


图2 几种常见的碰撞类型

Fig. 2 Common types of collisions

本文研究从不同的影响因素出发,对各项研究进行阐述。

(1)节点连通性、路径代价。杨洁等<sup>[7]</sup>提出了两种和时间推理相结合的多AGV路径规划模型,并使用拓扑建模法对仓储物流中心分拣库区进行数学建模,以模拟实际碰撞环境。陈展等<sup>[8]</sup>构建AGV的拓扑结构地图模型,设计基于全局邻域搜索的禁忌算法,来提高多AGV系统的稳定性和高效性。

(2)碰撞类型、任务分配。李昆鹏等<sup>[9]</sup>提出了两阶段优化算法,首先采用基于货架优先级的任务分配算法得到AGV的货架搬运任务序列,然后设计碰撞检测及避免算法进行主动避撞调度。Liang等<sup>[10]</sup>提出了AGV路径规划的三阶段集成调度算法。通过与岸桥和堆场块的联合优化,在集装箱港口前沿区域建立路网模型,优化AGV在路网中的路径,使其碰撞的概率有所下降。

(3)运动特征、地图环境。Sun等<sup>[11]</sup>提出了一种基于惯性导航系统和超宽带的AGV避撞系统,建立了工厂的电子地图并部署了超宽带锚节点以实现精确定位,利用AGV当前位置及其运动状态数据预测其下一位置,减少AGV控制延迟的影响,避免AGV之间的碰撞。武星等<sup>[12]</sup>针对多载量AGV变长特性,设计了一种变长度AGV路径空间冲突避免方法;针对交叉路口冲突,考虑到多载量AGV变长特性对阻塞环路的影响,提出一种交叉路口通行顺序优化方法。

此外,还有一些学者从其他的角度考虑AGV碰

撞问题。钟佩思等<sup>[13]</sup>根据初始条件和全局路径信息,判断系统中可能出现碰撞的位置,提出基于优先级机制的动态窗口法在相应位置附近完成路径的局部协调,最终在全局最优路径基础上实现局部路径协调避碰。蔡泽等<sup>[14]</sup>提出了一种基于深度强化学习的 AGV 避障方法,将 AGV 的避障问题表示为部分观测马尔可夫决策过程,详细描述了观测空间、动作空间和奖励函数,并采用深度确定性策略梯度算法训练避障策略。孙毛毛等<sup>[15]</sup>建立了多 AGV 避碰模型并结合时间窗模型,将路径规划分为预规划和实时规划两个阶段,使避碰路径规划更适用于复杂动态环境。胡恩泽等<sup>[16]</sup>提出了一种基于分层规划的综合优化调度方法,将调度问题分解为聚合的上层任务定序分配问题和下层路径规划问题。张艳菊等<sup>[17]</sup>利用时空数据设计了一种三维网格冲突检测方法,并根据商品 SKU 数量设定 AGV 的优先级以降低多 AGV 执行任务时的碰撞概率。Cha 等<sup>[18]</sup>提出了一种新的工作站优化分配方法,以根据制造过程中的建筑物布局和运输流程来最小化区域内和区域间的流量,并避免超载区域内多台 AGV 之间的碰撞。

## 1.2 死锁问题研究

死锁问题,在多 AGV 系统中,是指两个或两个以上的 AGV 在运行的过程中,由于争夺同一个节点或者路径,相互等待对方解除对资源的占用而造成无限等待的现象<sup>[19]</sup>。死锁问题会导致仓储能力的严重浪费,针对这个问题,中外学者对其进行了相关研究。

从优化架构的角度出发,Lin 等<sup>[20]</sup>提出了一种多 AGV 路径规划的两层策略,对所有 AGV 进行分组,为每个分组生成一个动态的虚拟领导者,不同组别的车辆被视为动态障碍物,通过快速更新分组车辆的状态减少死锁现象。Samia<sup>[21]</sup>提出了一种基于混合监督的双向 AGV 鲁棒控制架构,基于监督控制理论将问题形式化,以确保 AGV 的路径安全。Jerzy 等<sup>[22]</sup>将运输路网划分为互不重叠的区域,引入一种结构在线控制架构,保证 AGV 基本运输操作的执行不会导致死锁。

从优化策略的角度出发,肖海宁等<sup>[23]</sup>建立了以最小化任务配送路程和最大化待料停产剩余时间为综合优化目标的任务调度数学模型,提出于基于改进带精英策略的非支配排序遗传算法的防死锁任务调度方法。葛志远等<sup>[24]</sup>改进了计算基本蚁群算法,通过用最大最小蚂蚁系统对路径上信息素进行了限制,研究了路径规划中死锁问题的解决方法。李鑫等<sup>[25]</sup>提出一种基于时空冲突约束的 A\* 算

法在拓扑栅格地图的基础上加入时间轴建立时空地图模型,再针对时空地图的特点和冲突约束条件重新设计 A\* 算法的子节点扩展规则和节点评估函数,利用改进后的 A\* 算法按照优先级顺序为各个 AGV 规划路径,避免死锁现象。

综上所述,有关仓库多 AGV 路径冲突问题的研究文献中,以碰撞问题的研究居多,多数文献考虑的因素主要有路径代价、任务分配等,如表 1 所示。对死锁问题的研究,主要集中在优化架构和优化策略两方面,如表 2 所示。

表 1 碰撞问题相关研究

碰撞影响因素	研究方法	文献来源
节点连通性、路径代价	拓扑建模法	[7-8]
碰撞类型、任务分配	多阶段优化算法	[9-10]
运动特征、地图环境	设计避障系统及机制	[11-12]
路径信息	动态窗口法	[13]
奖励函数	深度强化学习	[14]
分层规划	时间窗模型 任务定序	[15-16]

表 2 死锁问题相关研究

研究角度	研究方法	文献来源
优化架构	生成虚拟领导者	[20]
	混合监督	[21]
	结构在线控制	[22]
优化策略	改进任务调度算法	[23]
	信息素限制	[24]
	时空地图模型	[25]

## 2 求解方法综述

数十年来,随着研究的深入,对路径规划的求解方式不断得到发展。现阶段,最常用的求解方式主要包括:启发式算法,强化学习。常见的启发式算法主要有:蚁群算法<sup>[26-28]</sup>、遗传算法<sup>[29-30]</sup>和 A\*<sup>[31-32]</sup>算法等。Q-learning 算法<sup>[33-34]</sup>和 DQN 算法<sup>[35-37]</sup>等算法作为新兴的机器学习算法也被学者们广泛应用。对近年来多 AGV 路径冲突问题的求解方法进行统计,如图 3 所示。可以看出,在解决多 AGV 路径冲突问题的方法中,启发式算法的应用最为普遍,自 2017 年以来快速增长。机器学习作为新兴的方法,从 2019 年开始也逐渐被应用,并且在稳步发展之中。因此,将从应用最为普遍的启发式算法和机器学习方法对路径冲突问题的求解方法进行阐述。

### 2.1 启发式算法在多 AGV 路径冲突问题上的应用

近年来,多 AGV 路径冲突问题中,启发式算法

的应用情况如图4所示,从图4中可以较为明显地看出,常用的启发式算法主要有:遗传算法、蚁群算法和A\*算法等,它们的优缺点如表3所示。针对启发式算法有时存在的收敛速度慢,陷入局部最优等问题,一些学者提出了自己的改进方案,主要可以分为3个方向:机制优化、因子调节、算法混合。

### 2.1.1 机制优化

针对蚁群算法,胡春阳等<sup>[38]</sup>提出了一种改进的头尾搜索机制,加快了蚁群算法的全局搜索能力和前期收敛速度。刘礼等<sup>[39]</sup>引入多目标路径性能评价指标,在路径长度的单一指标基础上引入路径风

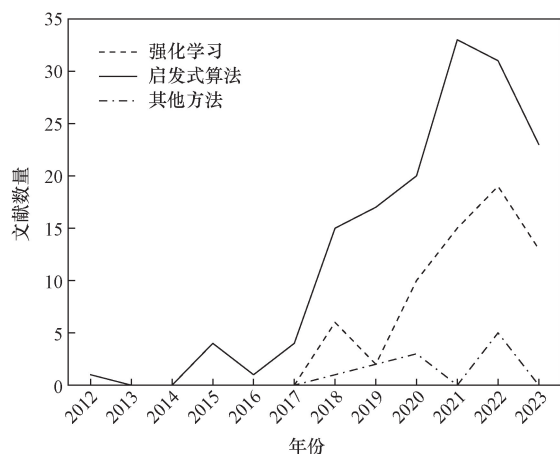


图3 研究方法统计图

Fig. 3 Research method statistics graph

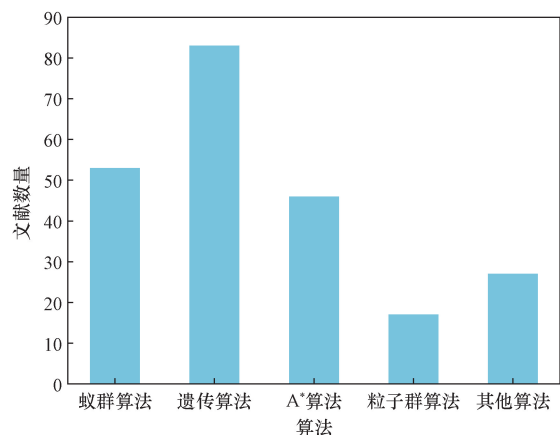


图4 启发式算法应用情况

Fig. 4 The application of heuristic algorithms

表3 常用算法优缺点

Table 3 Advantages and disadvantages of commonly used algorithms

算法	优点	缺点
蚁群算法	鲁棒性较好	收敛较慢
遗传算法	全局搜索能力强	参数设置影响较大
A*算法	对环境反应迅速	评估函数较难选取
粒子群算法	方法简单,效率高	易陷入局部最优

险指标和拐点数目,然后提出一种奖惩机制更新信息素增量,针对不同程度评价指标的路径提供不同的信息素更新机制,避免算法陷入早熟。针对遗传算法,张峰等<sup>[40]</sup>从时间和空间2个维度设计自适应算子,提出了基于二维度自适应遗传算法的联合调度方法,明确了AGV系统组成、工作模式和分配策略,具有较好的鲁棒性。Han等<sup>[41]</sup>采用了基于机器选择和操作排序的两层编码机制,设计了两种解码方法来确定AGV选择,每个种群使用一种解码方法,获得了更好的效率。针对A\*算法,刘威等<sup>[42]</sup>通过调整估价函数,改进估计策略,可在保持路径方向不变的前提下大量减少搜索节点数,有效提高路径规划效率。齐凤莲等<sup>[43]</sup>改进原有八叉树搜索策略提升避障性能,并提高了原有A\*算法的搜索效率。程杰等<sup>[44]</sup>提出一种对A\*算法的搜索优化和平滑优化方法,采用栅格法建立路径搜索空间模型,并引入安全距离约束对A\*算法进行搜索优化。

### 2.1.2 因子调节

针对蚁群算法,Tyler<sup>[45]</sup>引入路径平滑因子和全局信息因子,使AGV的避障能力进一步提升。王晓明等<sup>[46]</sup>引入调节因子来改进信息素更新规则,并对启发函数进行自适应调整,加快了算法搜索效率。杨桂华等<sup>[47]</sup>融合了确定性选择与随机性选择策略的优点,在路径转移概率中引入一个启发性的因子,可以使状态转移概率动态进行调整,从而使算法避免了陷入停滞,并对蚁群算法中的信息素更新时的策略加以改进。针对遗传算法,Ning等<sup>[48]</sup>提出了一种利用间歇分段给物流AGV供电的方法,设计了自适应权重值多目标遗传算法对分段供电电路进行优化设计,提出了3种权重值自适应方法,通过权重值自适应调节算法。刘畅等<sup>[49]</sup>提出了一种引入逆转算子、插入算子以及灾变算子的自适应遗传算法,改善了算法的收敛速度与优化结果。针对A\*算法,郭大林等<sup>[50]</sup>引入选择因子,利用选择因子减少传统A\*算法在路径寻优过程中搜索所有相同代价的冗余节点带来的额外开销,提升路径搜索效率。孔慧芳等<sup>[51]</sup>采用加权曼哈顿距离作为启发函数,使得距离估计成本更接近最短距离,以减少算法遍历节点数,另外,在算法启发函数中引入转弯修正代价因子,从而减少路径转弯次数。

### 2.1.3 算法融合

孟冠军等<sup>[52]</sup>利用A\*算法规划出一条较优路径作为初始路径,再对蚁群算法信息素更新方式以及节点转移概率公式中的启发函数加以改进来提高算法的求解精度。姜鹏等<sup>[53]</sup>将郊狼优化算法的组内郊狼成长、生与死进行改进并与遗传算法结合,提

出了一种带随机动态分组的遗传-郊狼混合算法,解决了遗传算法收敛速度慢、易陷入局部最优的问题。彭斌等<sup>[54]</sup>提出了一种结合改进 A\* 算法和动态窗口算法的混合算法,使用加权欧氏距离优化传统 A\* 算法的启发函数,并提出一种拐点识别算法去除路径冗余节点。邓希等<sup>[55]</sup>提出基于时间表和 A\* 算法的混合遗传算法,设计了基于任务单元的染色体编码方式,改进了种群初始化方案,交叉变异算子和精英保留策略。李西兴等<sup>[56]</sup>对鲸鱼优化算法进行离散化改进,针对性地设计了多种种群初始化策略,引入遗传算法的交叉、变异操作以提升鲸鱼优化算法的全局搜索能力,并嵌入局部搜索算法以达到全局搜索和局部搜索的平衡。Mousavi 等<sup>[57]</sup>将遗传算法与粒子群算法混合,并与单独的算法进行比较,证明混合算法具有更高的运行效率。Gao 等<sup>[58]</sup>设计了一种考虑最远插入启发式的大邻域搜索混合遗传算法,通过增加遗传算法的局部搜索能力来提高解的最优质量。

通过对以上 3 个方面的文献进行分析,可以发现机制优化、因子调节、算法混合分别适用于不同的场景,根据具体的应用场景和需求,选择合适的改进方式,可以显著提升路径规划的效率和解的质量,具体分析如下。

机制优化通常侧重于改进现有算法的核心操作机制,以提高算法的效率和性能。它的目标是解决原始算法中存在的收敛速度慢、易陷入局部最优等问题。因此,该方法适用于复杂、动态的环境,尤其是当路径规划问题规模较大时。例如,在大规模环境下对多 AGV 系统进行路径规划,环境复杂度较高,障碍物多,可能会导致传统算法(如蚁群算法、遗传算法)在搜索过程中收敛速度慢或陷入局部最优。机制优化通过改进搜索机制,能够有效地提升全局搜索能力,加速收敛过程,减少计算量。而且,改进的蚁群算法头尾搜索机制可以在大规模环境中更好地探索解空间。在面对不断变化的环境时,算法需要具备灵活性和快速适应能力。机制优化通过调整算法的搜索策略或信息更新方式,能够更好地应对动态障碍物或 AGV 状态的变化。另外,基于多目标路径评价的遗传算法可以实时考虑路径的风险、拐点数目等因素,避免路径设计过于局限。

因子调节是通过引入不同的调节因子来调整算法中的关键参数,以改善搜索过程的稳定性和效率。因此适用于那些需要精确控制搜索行为的场景,特别是在面对多变的条件和约束条件时,能够通过动态调整算法参数来达到优化效果。在不确定性较高的环境,如 AGV 系统在复杂的物流环境中

运行时,可能存在动态变化的任务优先级、交通状况或障碍物位置。此时,因子调节可以通过引入动态的调节因子(如启发式因子、路径平滑因子等)来实时调整算法的搜索策略,提高系统的灵活性和适应能力。在多目标优化问题中,多 AGV 协作路径规划往往需要同时考虑多个目标(如最短路径、最小能耗、最大避障等)。因子调节能够通过调整算法中的权重因子或参数,动态平衡不同目标之间的关系,找到最合适解。

算法混合通过结合多种算法的优势,以弥补单一算法的不足,通常用于复杂的、多约束的路径规划问题。它适用于高复杂度的路径规划问题,例如,在多 AGV 调度和路径规划中,单一算法可能无法高效地处理复杂约束(如避障、优先级、时间窗口等)。而算法混合可以结合不同算法(如 A\* 算法与蚁群算法、遗传算法与粒子群算法等)的优点,提高搜索效率和结果质量。除此之外,A\* 算法可以作为路径规划的初步解法,而蚁群算法则用于局部优化,从而提高整体路径质量。在一些解的质量和效率有要求的场景,当路径规划不仅要求计算速度,还要求路径质量(如最短路径、低能耗路径等)时,混合算法的效果尤为突出。通过结合全局搜索算法(如遗传算法、粒子群算法)与局部搜索算法(如局部优化的 A\* 算法、动态窗口算法等),能够在保持全局视野的同时,细化局部解的质量,适应复杂约束条件。

对上述研究进行总结,如表 4 所示。

表 4 启发式算法相关研究

Table 4 Related research on heuristic algorithms

研究角度	研究方法	文献来源
机制优化	改进头尾搜索机制	[38]
	奖惩机制	[39]
	两层编码机制	[41]
因子调节	路径平滑因子和全局信息因子	[45]
	启发性因子	[47]
算法混合	选择因子	[50]
	A* 算法与蚁群算法混合	[52]
	郊狼优化算法与遗传算法混合	[53]
	遗传算法与粒子群算法混合	[57]

## 2.2 机器学习在多 AGV 路径冲突问题上的应用

近年来,强化学习作为机器学习的重要分支,在 AGV 路径问题上逐渐得到了重视与运用。强化学习是不依赖环境的先验模型,且无需建立模型,相比于启发式算法,它在 AGV 路径冲突问题上的应用更加灵活。Q-learning 算法<sup>[59]</sup>在强化学习的发展中占有重要地位,是强化学习算法中最成熟、应用最广泛的算法之一。同时,Q-learning 算法也存在着

一些缺陷,针对这些缺陷,学者们也提出了相应的改进措施。

(1)前期收敛较慢。强化学习本质上是试错的过程,前期需要通过不断的学习来积累经验,这其中漫无目的地搜索会导致前期收敛速度较慢的问题。针对这个问题,王志伟等<sup>[60]</sup>改变传统  $Q$ -learning 算法的搜索方式,由原先的 8 方向变成 16 方向,利用模拟退火算法对  $Q$ -learning 进行迭代优化实现快速收敛。周琴等<sup>[61]</sup>考虑随机环境中存在的自循环结构,利用逐次超松弛技术对 Double Speedy  $Q$ -learning 算法的 Bellman 算子进行改进,加快了收敛速度。Hassen 等<sup>[62]</sup>在  $Q$ -learning 算法的搜索阶段融入了拥塞避免机制,并提出了基于 Max-Boltzman 的新策略,使得算法在平均延迟、收敛时间和计算时间等方面的性能得到了有效的提升。

(2)过估计问题。在  $Q$ -learning 算法中,更新  $Q$  函数时,通常会采用下一时刻中最优的  $Q$ ,但这个值存在不准确性,因此可能会引发过估问题。为改善这一现象,郑帅等<sup>[63]</sup>基于 Double  $Q$ -learning 算法中双估计器可以改善智能体收敛速度的特性,提出了一种改进算法 Double Speedy  $Q$ -learning,其通过双估计器,分离最优动作和最大  $Q$  值的选择。董永峰等<sup>[64]</sup>提出一种动态融合深度双  $Q$  算法,通过动态融合 DDQN 和平均 DQN 的先验知识进行网络参数训练,使网络输出的  $Q$  更加接近真实  $Q$  值。张峰等<sup>[65]</sup>提出一种基于树的经验存储结构来存储探索过程中的状态转移概率,并根据该存储方式,提出基于期望经验回放的  $Q$  学习算法在保证算法复杂度较低的情况下,可实现对环境状态转移的无偏估计,减少  $Q$  学习算法的过估计问题。

(3)维数灾难。在传统  $Q$ -learning 算法中,会使用  $Q$  表来储存  $Q$  值信息。当智能体每进行一个动作, $Q$  表都会更新一次。当环境十分复杂时,智能体的动作量将十分庞大,导致  $Q$  表无法完全储存所有的  $Q$  值。为解决该问题,Christion 等<sup>[66]</sup>提出了一种允许基于时间差异的更新来调整  $Q$  值 DQL 算法,损失函数计算为当前估计值与计算目标之间的均方误差,克服了维数灾难。赵德京等<sup>[67]</sup>在 Speedy  $Q$ -learning 算法的基础上提出了一种基于动作采样的 ASSQ 算法,采用集中训练-分散执行的框架,将上一迭代步更新后的  $Q$  作为下一状态的最大  $Q$ ,有效降低了  $Q$  的比较次数。唐昊等<sup>[68]</sup>利用其分层结构特点或引入分层控制方式,借助分层强化学习解决维数灾难问题。

对上述研究进行总结,如表 5 所示。相比于启发式算法,强化学习无需数学模型,也不依赖环境,

表 5  $Q$ -learning 算法相关研究

算法缺陷	研究方法	文献来源
前期收敛较慢	搜索方式由 8 方向变为 16 方向	[60]
	逐次超松弛技术	[61]
	拥塞避免机制	[62]
过估计	通过双估计器,分离最优动作和最大 $Q$	[63]
	动态融合 DDQN 和平均 DQN 的先验知识进行网络参数训练	[64]
	基于树的经验存储结构	[65]
维数灾难	允许基于时间差异的更新来调整 $Q$	[66]
	采用集中训练-分散执行的框架	[67]
	分层强化学习	[68]

但传统的强化学习算法也存在收敛较慢的问题。随着强化学习的发展,DQN 算法等深度学习算法被逐渐应用到更加复杂的环境当中。

### 3 结论与展望

对仓库多 AGV 路径冲突问题的研究,主要从冲突类型和研究方法进行了总结。从其研究现状以及发展前景来看,未来的研究主要集中在以下 3 个方面。

(1)更智能、高效的路径规划算法。目前,路径规划算法在仓库多 AGV 路径冲突问题中发挥着关键作用。然而,现有启发式算法在解决效率和实时性上仍存在不足。例如,在复杂任务环境中,路径规划的收敛速度可能过慢,且算法对动态变化的场景缺乏适应性。此外,强化学习作为一种新兴的解决方案,虽然在小规模实验中展现了一定潜力,但受制于训练样本的可靠性、算法的泛化能力以及“维数灾难”等问题,其实际应用仍然有限。因此,未来研究的重点可以放在:开发能够快速适应动态环境的在线路径规划算法、优化现有强化学习模型以适应大规模多 AGV 场景、探索深度学习与传统算法结合的混合式算法,以提升算法的求解质量和效率这几个方面。

(2)传感器与环境感知技术的创新。现阶段,解决路径冲突问题主要依赖路径规划和防冲突策略,而环境感知技术的发展尚未充分融入多 AGV 系统。未来,基于激光雷达、深度摄像头、超声波等先进传感器技术,实现对仓库环境的精准感知和动态建模,将显著提升 AGV 的自主避障能力。例如,通过多传感器融合技术,构建实时更新的高分辨率仓库地图,可以更精确地预测 AGV 之间的可能冲突。而这一领域的难点则在于如何将传感器感知与实时路径规划无缝结合,以实现从感知到行动的闭环控制。

(3)多 AGV 间的通信与协作。在仓库多 AGV 运行的环境下,现有研究较少涉及多 AGV 间的通信与协作,但是这种通信和协作在解决冲突时也可以发挥不小的作用。例如,基于物联网和 5G 等通信技术的 AGV 之间的实时通信机制,可以实现即时的路径调整和冲突解决,亦可以通过集体智能算法或者分布式协商机制,使得多个 AGV 能够共同协作解决路径冲突问题,提高整体效率和安全性。关于这类问题,目前的难点主要在于如何实现实时性和低延迟通信以及建立高效的信息共享和协调机制。多 AGV 间的通信与协作,将是未来此类问题研究的新方向之一。

### 参 考 文 献

- [1] 中国物流与采购联合会. 2023 年全国物流运行情况通报[EB/OL]. (2024-02-07) [2024-04-01]. <http://www.chinawuliu.com.cn/xsyj/202402/07/626457.shtml>.  
China Federation of Logistics and Purchasing. The 2023 national logistics operation bulletin 2023[EB/OL]. (2024-02-07) [2024-04-01]. <http://www.chinawuliu.com.cn/xsyj/202402/07/626457.shtml>.
- [2] 王晓军, 杨春霞, 陈航慧, 等. 考虑倒箱操作的 AutoStore 魔方型仓储系统自动引导车双层路径规划[J]. 科学技术与工程, 2021, 21(36): 15546-15554.  
Wang Xiaojun, Yang Chunxia, Chen Hanghui, et al. Double layer path planning for AutoStore cube type storage system automatic guided vehicle considering container dumping operation[J]. Science Technology and Engineering, 2021, 21(36): 15546-15554.
- [3] 诸葛沁沁, 许钢焱, 周耀明. 基于图论的 AGV 数量配置与调度优化方法[J]. 工业工程与管理, 2024, 29(2): 217-226.  
Zhuge Qin, Xu Gangyan, Zhou Yaoming. Optimization method for AGV quantity configuration and scheduling based on graph theory[J]. Industrial Engineering and Management, 2024, 29(2): 217-226.
- [4] Lu J S, Ren C H, Shao Y P, et al. An automated guided vehicle conflict-free scheduling approach considering assignment rules in a robotic mobile fulfillment system[J]. Computers & Industrial Engineering, 2023, 176. DOI: 10.1016/J.CIE.2022.108932.
- [5] 任明辉, 梁军, 陈龙, 等. 基于改进冲突搜索的智能车库多 AGV 路径规划[J]. 汽车工程, 2023, 45(10): 1933-1943.  
Ren Minghui, Liang Jun, Chen Long, et al. Intelligent garage multi AGV path planning based on improved conflict search [J]. Automotive Engineering, 2023, 45(10): 1933-1943.
- [6] Zhang Z, Guo Q, Chen J, et al. Collision-free route planning for multiple agvs in an automated warehouse based on collision classification[J]. IEEE Access, 2018, 6: 26022-26035.
- [7] 杨洁, 田钊, 刘国宁. 物流中心多 AGV 无碰撞路径规划[J]. 机械设计与制造, 2022(8): 271-277, 281.  
Yang Jie, Tian Zhao, Liu Guoning. Collision-free path planning for multiple AGVs in logistics center [J]. Mechanical Design and Manufacturing, 2022(8): 271-277, 281.
- [8] 陈展, 公建宁, 刘媛媛, 等. 基于禁忌搜索的多 AGV 系统路径优化算法[J]. 计算机工程与应用, 2021, 57(10): 273-278.  
Chen Zhan, Gong Jianing, Liu Yuanyuan, et al. Path optimization algorithm for multi-AGV system based on taboo search [J]. Computer Engineering and Applications, 2021, 57(10): 273-278.
- [9] 李昆鹏, 刘腾博, 贺冰倩, 等. “货到人”拣选系统中 AGV 路径规划与调度研究[J]. 中国管理科学, 2022, 30(4): 240-251.  
Li Kunpeng, Liu Tengbo, He Bingqian, et al. Research on AGV path planning and scheduling in “goods-to-person” picking system [J]. Chinese Journal of Management Science, 2022, 30(4): 240-251.
- [10] Liang C J, Zhang Y, Dong L. A three stage optimal scheduling algorithm for AGV route planning considering collision avoidance under speed control strategy [J]. Mathematics, 2022, 11(1): 138-138.
- [11] Sun S K, Hu J P, Li J, et al. An INS-UWB based collision avoidance system for AGV [J]. Algorithms, 2019, 12(2). DOI: 10.3390/a12020040.
- [12] 武星, 翟晶晶, 肖海宁, 等. 多载量 AGV 系统防死锁路口通行顺序优化及避碰[J]. 计算机集成制造系统, 2022, 28(4): 979-989.  
Wu Xing, Zhai Jingjing, Xiao Haining, et al. Optimization of passing sequence and collision avoidance for deadlock-free intersections in multiloader AGV systems[J]. Computer Integrated Manufacturing Systems, 2022, 28(4): 979-989.
- [13] 钟佩思, 徐东方, 李东民, 等. 基于双层协调体系的多移动机器人路径规划方法[J]. 科学技术与工程, 2020, 20(29): 12000-12006.  
Zhong Peisi, Xu Dongfang, Li Dongmin, et al. Path planning method for multiple mobile robots based on a dual-layer coordination system [J]. Science Technology and Engineering, 2020, 20(29): 12000-12006.
- [14] 蔡泽, 胡耀光, 闻敬谦, 等. 复杂动态环境下基于深度强化学习的 AGV 避障方法[J]. 计算机集成制造系统, 2023, 29(1): 236-245.  
Cai Ze, Hu Yaoguang, Wen Jingqian, et al. Obstacle avoidance method for AGV based on deep reinforcement learning in complex dynamic environments [J]. Computer Integrated Manufacturing Systems, 2023, 29(1): 236-245.
- [15] 孙毛毛, 匡兵. 基于时间窗的 AGV 动态避碰路径规划方法[J]. 计算机集成制造系统, 2023, 29(1): 236-245.  
Sun Maomao, Kuang Bing. Time window-based dynamic collision avoidance path planning method for AGV [J]. Journal of Computer Integrated Manufacturing Systems, 2023, 29(1): 236-245.
- [16] 胡恩泽, 贺建军, 申帅, 等. 面向大型工业仓库的多 AGV 综合优化调度方法[J]. 中南大学学报(自然科学版), 2023, 54(5): 1779-1790.  
Hu Enze, He Jianjun, Shen Shuai, et al. Comprehensive optimization scheduling method for multi-AGV in large-scale industrial warehouses[J]. Journal of Central South University (Natural Science Edition), 2023, 54(5): 1779-1790.
- [17] 张艳菊, 吴俊, 程锦倩, 等. 多搬运任务下考虑碰撞避免的 AGV 路径规划[J]. 计算机应用研究, 2024, 41(5): 1462-1469.  
Zhang Yanju, Wu Jun, Cheng Jinqian, et al. AGV path planning considering collision avoidance under multiple handling tasks[J]. Computer Application Research, 2024, 41(5): 1462-1469.

- [18] Cha R G, Ji C, Nam G H. Research on workstation-to-zone assignment for tandem automated guided vehicle system with multiple AGVs in overload zone of transport[J]. The International Journal of Advanced Manufacturing Technology, 2024, 135 (12): 4653-4668.
- [19] 陈果, 何代钰, 欧阳博, 等. 基于链状结构保持的多 AGV 系统灵活高效死锁避免控制策略[J]. 机器人, 2023, 45(5): 591-602.  
Chen Guo, He Daiyu, Ouyang Bo, et al. Flexible and efficient deadlock avoidance control strategy for multi-AGV systems based on chain-like structure preservation[J]. Robot, 2023, 45(5): 591-602.
- [20] Lin S W, Liu A, Wang J G. A dual-layer weight-leader-vicsek model for multi-AGV path planning in warehouse [J]. Biomimetics, 2023, 8(7). DOI: 10.3390/BIOMIMET-ICS8070549.
- [21] Samia M. Hybrid supervisory-based architecture for robust control of Bi-directional AG-Vs[J]. Computers in Industry, 2023, 144. DOI: 10.1016/J.COMPIND.2022.103797.
- [22] Jerzy Z, Waldemar M. Structural on-line control policy for collision and deadlock resolution in multi-AGV systems[J]. Journal of Manufacturing Systems, 2021, 60: 80-92.
- [23] 肖海宁, 楼佩煌, 武星, 等. 多载量自动导引车系统防死锁任务调度方法[J]. 计算机集成制造系统, 2022, 28(11): 3325-3339.  
Xiao Haining, Lou Pei Huang, Wu Xing, et al. Deadlock prevention task scheduling method for multi-load automated guided vehicle systems [J]. Computer Integrated Manufacturing Systems, 2022, 28(11): 3325-3339.
- [24] 葛志远, 肖本贤. 使用改进蚁群算法的 AGV 路径规划研究[J]. 机械设计与制造, 2020(6): 241-244, 248.  
Ge Zhiyuan, Xiao Benxian. Research on AG-V path planning using improved ant colony algorithm[J]. Mechanical Design and Manufacturing, 2020(6): 241-244, 248.
- [25] 李鑫, 廖凯文, 陈薇, 等. 多自动导引车路径规划的时空冲突约束 A\* 算法[J]. 计算机集成制造系统, 2021, 27(11): 3219-3226.  
Li Xin, Liao Kaiwen, Chen Wei, et al. A\* algorithm with spatio-temporal conflict constraints for multi-automated guided vehicle path planning[J]. Computer Integrated Manufacturing Systems, 2021, 27(11): 3219-3226.
- [26] 金鑫鑫. 一种基于改进蚁群算法的 AGV 小车三维路径规划研究[J]. 计算机应用与软件, 2022, 39(7): 275-280.  
Jin Xinxin. Research on three-dimensional path planning of AGV based on improved ant colony algorithm[J]. Computer Applications and Software, 2022, 39(7): 275-280.
- [27] 于军琪, 李若琳, 赵安军, 等. 并行排序蚁群算法规划仓库 AGV 路径研究[J]. 机械科学与技术, 2021, 40(4): 609-618.  
Yu Junqi, Li Ruolin, Zhao Anjun, et al. Research on warehouse AGV path planning using parallel sorting ant colony algorithm[J]. Mechanical Science and Technology, 2021, 40(4): 609-618.
- [28] Yu J Q, Li R L, Feng Z X, et al. A novel parallel ant colony optimization algorithm for warehouse path planning[J]. Journal of Control Science and Engineering, 2020. DOI: 10.1155/2020/5287189.
- [29] 吴立辉, 胡文博, 周秀, 等. 面向多搬运任务的柔性制造车间多载具 AGV 节能路径规划[J]. 计算机应用研究, 2023, 40(1): 57-62.  
Wu Lihui, Hu Wenbo, Zhou Xiu, et al. Energy-saving path planning for multi-vehicle AGV in flexible manufacturing workshop with multiple handling tasks[J]. Computer Application Research, 2023, 40(1): 57-62.
- [30] Wang J J, Pan J N, Huo J K, et al. Research on optimization of multi-AGV path Based on genetic algorithm considering charge utilization[J]. Journal of Physics: Conference Series, 2021, 1769(1). DOI: 10.1088/1742-6596/1769/1/012052.
- [31] 陆舜路, 熊峰, 孔维畅. 基于时间窗 A\* 算法的船体焊接车间调度研究[J]. 计算机仿真, 2023, 40(11): 212-217.  
Lu Shunlu, Xiong Feng, Kong Weichang. Research on scheduling of ship welding workshop based on time window A\* algorithm[J]. Computer Simulation, 2023, 40(11): 212-217.
- [32] 张建光, 张方, 陈良港, 等. 基于改进 A\* 算法的自动导引车的路径规划[J]. 国外电子测量技术, 2022, 41(1): 123-128.  
Zhang Jianguang, Zhang Fang, Chen Liangang, et al. Path planning of automatic guided vehicles based on improved A\* algorithm [J]. Foreign Electronic Measurement Technology, 2022, 41(1): 123-128.
- [33] Zhou P F; Lin L; Kim K H. Anisotropic Q-learning and waiting estimation based real-time routing for automated guided vehicles at container terminals [J]. Journal of Heuristics, 2021, 29(2): 1-22.
- [34] Bai Y F, Ding X F, Hu D S, et al. Research on dynamic path planning of Multi-AGVs based on reinforcement learning[J]. Applied Sciences, 2022, 12(16). DOI: 10.3390/AP12168166.
- [35] 孙辉, 袁维. 基于深度强化学习的多自动导引车运动规划[J]. 计算机集成制造系统, 2024, 30(2): 708-716.  
Sun Hui, Yuan Wei. Motion planning for multiple automated guided vehicles based on deep reinforcement learning[J]. Computer Integrated Manufacturing Systems, 2024, 30(2): 708-716.
- [36] 谢勇, 郑绥君, 程念胜, 等. 基于改进 DQN 算法的无人仓多 AGV 路径规划[J]. 工业工程, 2024, 27(1): 36-44, 53.  
Xie Yong, Zheng Suijun, Cheng Niansheng, et al. Multi-AGV path planning in unmanned warehouse based on improved DQN algorithm[J]. Industrial Engineering, 2024, 27(1): 36-44, 53.
- [37] 黄岩松, 姚锡凡, 景轩, 等. 基于深度 Q 网络的多起点多终点 AGV 路径规划[J]. 计算机集成制造系统, 2023, 29(8): 2550-2562.  
Huang Yansong, Yao Xifan, Jing Xuan, et al. Multi-start multi-end AGV path planning based on deep Q network[J]. Computer Integrated Manufacturing Systems, 2023, 29(8): 2550-2562.
- [38] 胡春阳, 姜平, 周根荣. 改进蚁群算法在 AGV 路径规划中的应用[J]. 计算机工程与应用, 2020, 56(8): 270-278.  
Hu Chunyang, Jiang Ping, Zhou Genrong. Application of improved ant colony algorithm in AGV path planning[J]. Computer Engineering and Applications, 2020, 56(8): 270-278.
- [39] 刘礼, 刘勇, 孙云权, 等. 基于自适应蚁群算法的 AGV 路径规划优化[J]. 电子测量技术, 2023, 46(18): 100-107.  
Liu Li, Liu Yong, Sun Yunquan, et al. AGV path planning optimization based on adaptive ant colony algorithm [J]. Electronic Measurement Technology, 2023, 46(18): 100-107.
- [40] 张峰, 陈乃超, 邢海燕. 柔性车间生产资源与 AGV 物流资源联合优化调度[J]. 现代制造工程, 2023(10): 1-7.  
Zhang Feng, Chen Naichao, Xing Haiyan. Joint optimization

- scheduling of flexible workshop production resources and AGV logistics resources [J]. *Modern Manufacturing Engineering*, 2023 (10): 1-7.
- [41] Han X Q, Cheng W Y, Meng L L, et al. A dual population collaborative genetic algorithm for solving flexible job shop scheduling problem with AGV [J]. *Swarm and Evolutionary Computation*, 2024, 86. DOI: 10.1016/j.swevo.2024.101538.
- [42] 刘威, 储春华, 肖明伟. 一种基于时间和路径双重优化的改进 A\* 算法 [J]. *制造业自动化*, 2023, 45(12): 173-177.  
Liu Wei, Chu Chunhua, Xiao Mingwei. An improved A\* algorithm based on dual optimization of time and path [J]. *Manufacturing Automation*, 2023, 45(12): 173-177.
- [43] 齐凤莲, 王晓庆, 张帼英. 改进 A\* 算法的 AGV 避障路径规划研究 [J]. *机床与液压*, 2023, 51(9): 34-39.  
Qi Fenglian, Wang Xiaoqing, Zhang Guoying. Research on obstacle avoidance path planning of AGV based on improved A\* algorithm [J]. *Machine Tool & Hydraulic*, 2023, 51(9): 34-39.
- [44] 程杰, 陈姚节, 刘志峰. 基于 A\* 算法优化的无人水面艇路径规划 [J]. *科学技术与工程*, 2019, 19(29): 201-206.  
Cheng Jie, Chen Yaojie, Liu Zhifeng. Path planning for unmanned surface vehicles based on optimized A\* algorithm [J]. *Science Technology and Engineering*, 2019, 19(29): 201-206.
- [45] Tyler B. Research on obstacle avoidance path selection of AGV based on improved ant colony algorithm [J]. *Computer Informatization and Mechanical System*, 2023, 6(2): 1-5.
- [46] 王晓明, 尚硕, 梁超, 等. 改进蚁群算法的 AGV 路径规划 [J]. *组合机床与自动化加工技术*, 2023(3): 63-65.  
Wang Xiaoming, Shang Shuo, Liang Chao, et al. AGV path planning based on improved ant colony algorithm [J]. *Combined Machine Tools and Automatic Processing Technology*, 2023 (3): 63-65.
- [47] 杨桂华, 符士宾, 刘志毅, 等. 基于改进蚁群算法的室内移动机器人路径规划 [J]. *科学技术与工程*, 2019, 19(19): 175-179.  
Yang Guihua, Fu Shibin, Liu Zhiyi, et al. Indoor Mobile robot path planning based on improved ant colony algorithm [J]. *Science Technology and Engineering*, 2019, 19(19): 175-179.
- [48] Ning P F, Li T T, Niu P J. The design of AGV's dynamic wireless power transmission power supply based on adaptive weight value multi-objective genetic algorithm [J]. *IOP Conference Series: Materials Science and Engineering*, 2019, 688(3). DOI: 10.1088/1757-899X/688/3/033020.
- [49] 刘畅, 张承瑞, 孙玉玺. 改进自适应遗传算法在多载 AGV 调度的应用研究 [J]. *小型微型计算机系统*, 2021, 42(11): 2241-2245.  
Liu Chang, Zhang Chengrui, Sun Yuxi. Application research of improved adaptive genetic algorithm in multi-load AGV scheduling [J]. *Journal of Miniaturization of Computer Systems*, 2021, 42(11): 2241-2245.
- [50] 郭大林, 王淑莹, 曾文驱, 等. 改进 A\* 算法的路径规划在列检环境中的应用 [J]. *计算机工程与设计*, 2023, 44(1): 299-306.  
Guo Dalin, Wang Shuying, Zeng Wenqu, et al. Application of improved A\* algorithm in path planning in column inspection environment [J]. *Computer Engineering and Design*, 2023, 44(1): 299-306.
- [51] 孔慧芳, 盛阳. 基于改进 A\* 算法的 AGV 路径规划研究 [J]. *现代制造工程*, 2021(10): 60-64.  
Kong Huifang, Sheng Yang. Research on AGV path planning based on improved A\* algorithm [J]. *Modern Manufacturing Engineering*, 2021(10): 60-64.
- [52] 孟冠军, 陈信华, 陶细佩, 等. 基于混合蚁群算法的 AGV 路径规划 [J]. *组合机床与自动化加工技术*, 2021(1): 70-73.  
Meng Guanjuan, Chen Xinhua, Tao Xipei, et al. AGV path planning based on hybrid ant colony algorithm [J]. *Combined Machine Tools and Automatic Processing Technology*, 2021(1): 70-73.
- [53] 姜鹏, 方成刚, 杨帆. GA-COA 求解柔性作业车间多资源调度问题 [J]. *机械设计与制造*, 2023(3): 156-159, 164.  
Jiang Peng, Fang Chenggang, Yang Fan. Solving multi-resource scheduling problem in flexible job shop with GA-COA [J]. *Machinery Design and Manufacturing*, 2023(3): 156-159, 164.
- [54] 彭斌, 王力, 杨思霖. 基于改进 A\* 算法和动态窗口算法的自动引导小车轨迹规划 [J]. *计算机应用*, 2022, 42(S1): 347-352.  
Peng Bin, Wang Li, Yang Silin. Trajectory planning of automatic guided vehicle based on improved A\* algorithm and dynamic window approach [J]. *Computer Applications*, 2022, 42(S1): 347-352.
- [55] 邓希, 胡晓兵, 江代渝, 等. 基于混合遗传算法的柔性作业车间机器人 AGV 规划 [J]. *四川大学学报(自然科学版)*, 2021, 58(2): 73-82.  
Deng Xi, Hu Xiaobing, Jiang Daiyu, et al. Machine and AGV planning in flexible job shop based on hybrid genetic algorithm [J]. *Journal of Sichuan University (Natural Science Edition)*, 2021, 58(2): 73-82.
- [56] 李西兴, 杨道明, 李鑫, 等. 基于混合遗传鲸鱼优化算法的柔性作业车间自动引导车融合调度方法 [J]. *中国机械工程*, 2021, 32(8): 938-950, 986.  
Li Xixing, Yang Daoming, Li Xin, et al. Fusion scheduling method of automatic guided vehicles in flexible job shop based on hybrid genetic whale optimization algorithm [J]. *China Mechanical Engineering*, 2021, 32(8): 938-950, 986.
- [57] Mousavi M, Yap H J, Musa S N, et al. Multi-objective AGV scheduling in an FMS using a hybrid of genetic algorithm and particle swarm optimization [J]. *PLoS ONE*, 2017, 12(3): e0169817.
- [58] Gao J, Zheng X J, Gao F. Heterogeneous multitype fleet green vehicle path planning of automated guided vehicle with timewindows in flexible manufacturing system [J]. *Machines*, 2022, 10(3). DOI: 10.3390/MACHINES10030197.
- [59] Christopher J C H, Watkins P D. Technical note: Q-learning [J]. *Machine Learning*, 1992, 8(3): 279-292.
- [60] 王志伟, 邹艳丽, 刘唐慧美, 等. 基于改进 Q-learning 算法和 DWA 的路径规划 [J]. *传感器与微系统*, 2023, 42(9): 148-152.  
Wang Zhiwei, Zou Yanli, Liu Tanghuimei, et al. Path planning based on improved Q-learning algorithm and DWA [J]. *Sensors and Microsystems*, 2023, 42(9): 148-152.
- [61] 周琴, 罗飞, 丁炜超, 等. 基于逐次超松弛技术的 Double Speedy Q-Learning 算法 [J]. *计算机科学*, 2022, 49(3): 239-245.  
Zhou Qin, Luo Fei, Ding Weichao, et al. Double speedy Q-Learning algorithm based on successive over-relaxation Technique [J]. *Computer Science*, 2022, 49(3): 239-245.
- [62] Hassen H, Meherzi S, Jemaa Z B. Improved exploration strategy

- for  $Q$ -learning based multipath routing in SDN networks[J]. Journal of Network and Systems Management, 2024, 32(2). DOI: 10.1007/s10922-024-09804-0.
- [63] 郑帅, 罗飞, 顾春华, 等. 基于双估计器的改进 Speedy  $Q$ -learning 算法[J]. 计算机科学, 2020, 47(7): 179-185.  
Zheng Shuai, Luo Fei, Gu Chunhua, et al. Improved speedy  $Q$ -learning algorithm based on dual estimators[J]. Computer Science, 2020, 47(7): 179-185.
- [64] 董永峰, 杨琛, 董瑶, 等. 基于改进的 DQN 机器人路径规划[J]. 计算机工程与设计, 2021, 42(2): 552-558.  
Dong Yongfeng, Yang Chen, Dong Yao, et al. Path planning of robots based on improved DQN[J]. Computer Engineering and Design, 2021, 42(2): 552-558.
- [65] 张峰, 钱辉, 董春茹, 等. 随机状态下基于期望经验回放的  $Q$  学习算法[J]. 深圳大学学报(理工版), 2020, 37(2): 202-207.  
Zhang Feng, Qian Hui, Dong Chunru, et al.  $Q$ -learning algorithm based on expected experience replay under random states[J]. Journal of Shenzhen University Science and Engineering, 2020, 37(2): 202-207.
- [66] Christian M, Paul A, Jimmy G, et al. Enhancing energy management strategies for extended-range electric vehicles through deep  $Q$ -learning and continuous state representation[J]. Energies, 2024, 17(2). DOI: 10.3390/EN17020514.
- [67] 赵德京, 马洪聪, 王家曜, 等. 改进  $Q$  学习算法在多智能体强化学习中的应用[J]. 自动化与仪器仪表, 2022(6): 13-16, 22.  
Zhao Dejing, Ma Hongcong, Wang Jiayao, et al. Application of improved  $Q$ -learning algorithm in multi-agent reinforcement learning[J]. Automation and Instrumentation, 2022(6): 13-16, 22.
- [68] 唐昊, 张晓艳, 韩江洪, 等. 基于连续时间半马尔可夫决策过程的 Option 算法[J]. 计算机学报, 2014, 37(9): 2027-2037.  
Tang Hao, Zhang Xiaoyan, Han Jianghong, et al. Option algorithm based on continuous-time semi-Markov decision process[J]. Chinese Journal of Computers, 2014, 37(9): 2027-2037.