Graph Neural Networks and Deep Reinforcement Learning in Warehouse Order Picking and Batching - Literature Review

DOI
https://doi.org/
10.18690/um.fov.2.2025.14

ISBN
978-961-286-963-2

Nejc Čelik, Andrej Škraba

University of Maribor, Faculty of Organizational Sciences, Cybernetics & Decision Support Systems Laboratory, Kranj, Slovenia
nejc.celik1@um.si, andrej.skraba@um.si

This paper is a systematic literature review on use of the Deep Reinforcement Learning (DRL) and Graph Neural Networks (GNN) in warehouse. We first explore the use of DRL and GNN for optimization of order picking and batching in warehouse. Because of very little results on use of GNNs in optimization of order picking and batching we extended our search to general use of GNNs in warehouse environment. We identified different topics of research using Latent Dirichlet Allocation (LDA) and identified main problems in use of DRL and GNNs in warehouse environment.

Keywords:
Graph Neural Network,
Deep Reinforcement Learning
warehouse,
order picking,
batching,
optimization

University of Maribor Press

1 Introduction

The order picking process is one of the most labor-intensive processes in warehouses and has been for a long time cited to consist of 55% of all warehouse operating costs and is one of the most researched topic of warehouse research to this day (Bock et al., 2025). Significant research has been done on optimization of order picking paths and batching of orders to minimize path traveled by pickers. Finding shortest order picking path is variation of Traveling Salesman Problem (TSP) which is NP-hard (Bock et al., 2025). In recent years new methods of solving large optimization problems with Deep Reinforcement Learning (DRL) (Arulkumaran et al., 2017; Bello et al., 2017) and with Graph Neural Networks (GNN) (J. Zhou et al., 2020) and combination of both (Munikoti et al., 2022) emerged. In this paper we explore the use of DRL in optimization of order batching and picking tasks. We further explore the use of GNNs in a general warehouse context. We follow PRISMA methodology (Page et al., 2021) for systematic literature Review.

2 Methodology

PRISMA methodology (Page et al., 2021) describes how methods and results in systematic reviews should be reported. We searched 3 different scientific papers databases for papers on topic of use of DRL and GNNs and combination of both in warehouse setting. Databases used were Scopus, Web of Science and ProQuest. We search for the occurrence of certain words in the paper titles and abstracts. Those words include words like "warehouse", "reinforcement learning", "order picking", "order batching" and "graph neural networks". We then excluded duplicate records and non-related papers.

Our goal was to find how GNNs and DRL are used in warehouse settings. We were mainly interested in the topic of order picking and batching optimization. We searched for any papers that included GNNs or DRL for use in order picking and batching. We found only 1 article containing both GNNs and DRL (Begnardi et al., 2024) for order picking or batching problem. So, we extended our search on general use of GNNs in warehouse setting. We were interested in what GNNs were used for and how they were modeled.

To find general topics in research we used Latent Dirichlet Allocation (LDA) (Blei et al., 2003) on papers texts. Our text analysis pipeline (Figure 1) started with extraction of texts from pdf files and cleaning those texts by removing chapters like literature, notes, acknowledgement and all text before the start of abstract. We then used scikit-learn (Pedregosa et al., 2011) implementation of LDA to identify different topics of research. When using LDA we iteratively added new stop words and word replacements and tested them with different numbers of topics to find the most distinguishable topics. In choosing the best number of topics, we used perplexity score of LDA model. With those topics we aimed to find general topics on research in use of GNNs and DRL in warehouse settings.

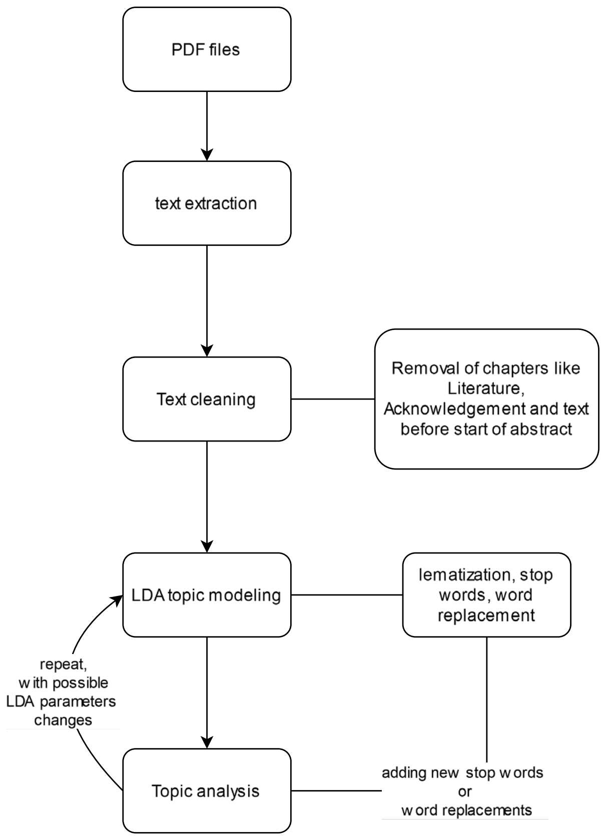

Figure 1: Scientific papers' text analysis pipeline
Source: Own

2 Results

With queries in Table 1 we identified 85 papers. Figure 2 shows identification of new studies via databases and registers. In the process, 31 of papers were excluded as duplicates and we excluded 4 full conference proceedings where combination of words was in full conference abstract. We then screened 50 papers and removed 11 of them as unrelated. Between unrelated papers there were mostly papers that included word warehouse as in context of data-warehouses and some computer vision papers focused on object detection in warehouse settings. We retrieved and analyzed the rest of the papers (31 papers). All but 2 of the included papers were published between the years 2021 and 2025.

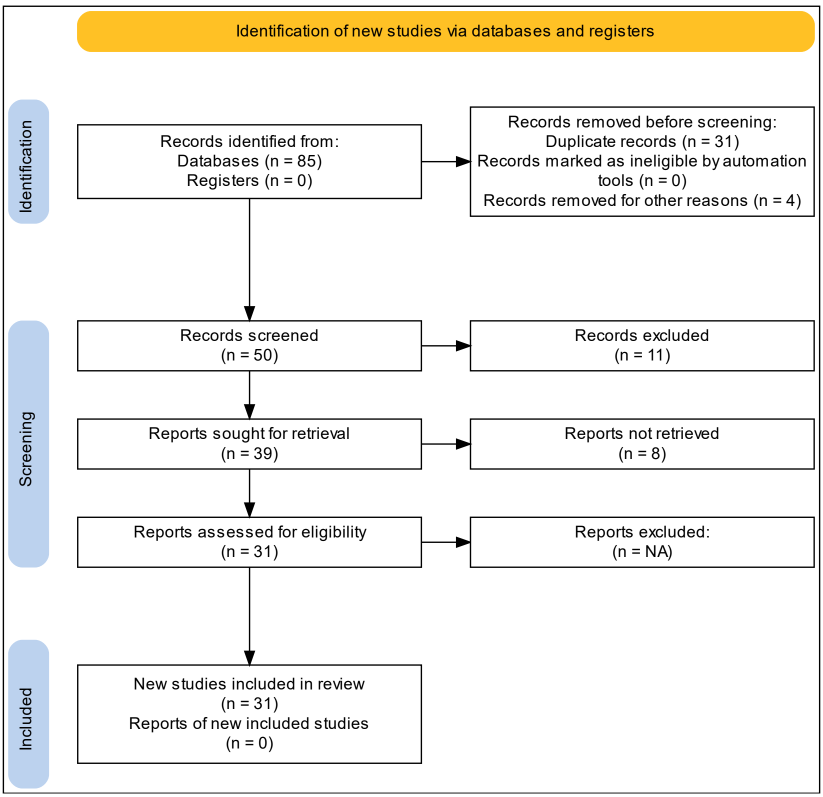

Figure 2: Identification of new studies via databases and registers (Made with PRISMA Flow Diagram app (Haddaway et al., 2022)

Source: Own

Table 1: Queries

Topic	query	Database	results
DRL	TITLE-ABS-KEY ((warehouse*) AND ("order picking" OR "order batching") AND ("reinforcement learning"))	Scopus	21
DRL	TS=((warehouse*) AND ("order picking" OR "order batching") AND ("reinforcement learning"))	WOS	17
DRL	ab(warehouse* AND ("order picking" OR "order batching") AND ("reinforcement learning"))	ProQuest	7
DRL	Left after leaving non accessible, duplicates and unrelated		23
GNN	TITLE-ABS-KEY (warehouse* AND ("Graph Neural Networks" OR gnn*))	Scopus	23
GNN	TS=(warehouse* AND ("Graph Neural Networks" OR gnn*))	WOS	11
GNN	ab((warehouse* AND ("Graph Neural Networks" OR gnn*)))	ProQuest	6
GNN	Left after leaving non accessible, duplicates and unrelated		9
Combined	Both topics after removing duplicates		31

We analyzed both topics (DRL and GNN) separately. First, we analyzed papers on DRL. Research highlights the growing role of Deep Reinforcement Learning (DRL) in warehouse logistics and order fulfillment. Several studies demonstrate the superior performance of DRL-based methods compared to traditional heuristics (Begnardi et al., 2024; D. Wang et al., 2022; X. Wang et al., 2025). However, challenges remain, including the difficulty of modeling DRL agents effectively (Cals et al., 2021), the need for more realistic problem formulations (Dehghan et al., 2023; Neves-Moreira & Amorim, 2024), and issues related to data collection and model complexity (Cheng, Wang, et al., 2024). Several studies emphasize the importance of incorporating additional constraints and real-world factors such as worker ergonomics (Niu et al., 2021), congestion, and system breakdowns (Perumaal Subramanian & Kumar Chandrasekar, 2024).

Use of DRL for optimization of order batching and picking can be separated into five topics by using LDA (Table 2). In all topics use of DRL is often leveraged to offer way of joint optimization for multiple goals in real time (dynamic decision making). The first topic explores the use of DRL for optimization with focus on picking task. The second topic is more focused on external logistics (customers and delivery). The third topic adds more focus on order batching. The fourth topic is addressing problems with the use of automated guided vehicles (AGV). And the last fifth topic has more human centric focus (ergonomic, human-robot cooperation and transition to automation).

Table 2: DRL papers

Reference	Lda Topic
(Begnardi et al., 2024) (X. Wang et al., 2025) (Mahmoudinazlou et al., 2024) (Smit et al., 2024)	1
(Dehghan et al., 2023) (Kamoshida & Kazama, 2017) (Neves-Moreira & Amorim, 2024)	2
(D. Wang et al., 2022) (Cals et al., 2021) (H. Li et al., 2022) (L. Zhou et al., 2022) (Drakaki & Tzionas, 2017) (Cheng, Wang, et al., 2024)	3
(Kaiser et al., 2023) (K. Li et al., 2024) (Cheng, Xie, et al., 2024) (Tang et al., 2021) (Perumaal Subramanian & Kumar Chandrasekar, 2024) (Chen et al., 2024) (Wu et al., 2024)	4
(Krnjaic et al., 2024) (Yoshitake & Abbeel, 2023) (Niu et al., 2021)	5

Figure 3 shows LDA Topics recommendation for selected papers.

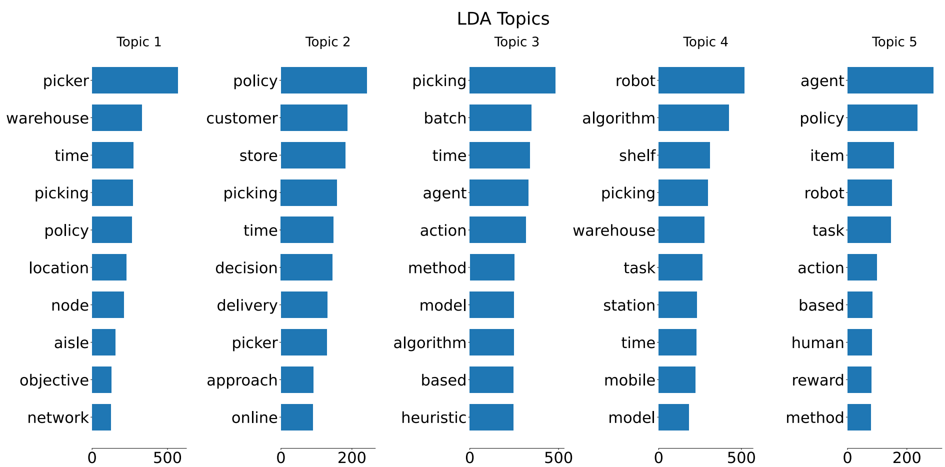

Figure 3: LDA Topics recommendation
Source: Own

The biggest problem highlighted in papers was need for adding more optimization constraints and difficulty of modeling problems for DRL (Cals et al., 2021; Cheng, Wang, et al., 2024; Dehghan et al., 2023; Krnjaic et al., 2024; K. Li et al., 2024; Neves-Moreira & Amorim, 2024; Tang et al., 2021; Wu et al., 2024; Yoshitake & Abbeel, 2023).

Use of GNNs in warehouse settings can be separated into two topics by using LDA. The first topic explores finding the best action for agent (AGV, human worker) with emphasis on RL and the second topic explores scheduling and allocation/assignment of tasks between agents. Table 3 shows the set of GNN papers.

Table 3: GNN papers

Citation	LDA Topic
(Begnardi et al., 2024) (Knippenberg et al., 2021) (X. Li et al., 2023) (Xiao et al., 2024) (Pham & Bera, 2024) (Shelke et al., 2023)	1
(Paul et al., 2023) (Cho et al., 2024) (Z. Wang et al., 2022)	2

Figure 4 shows LDA Topics recommendation for selected papers.

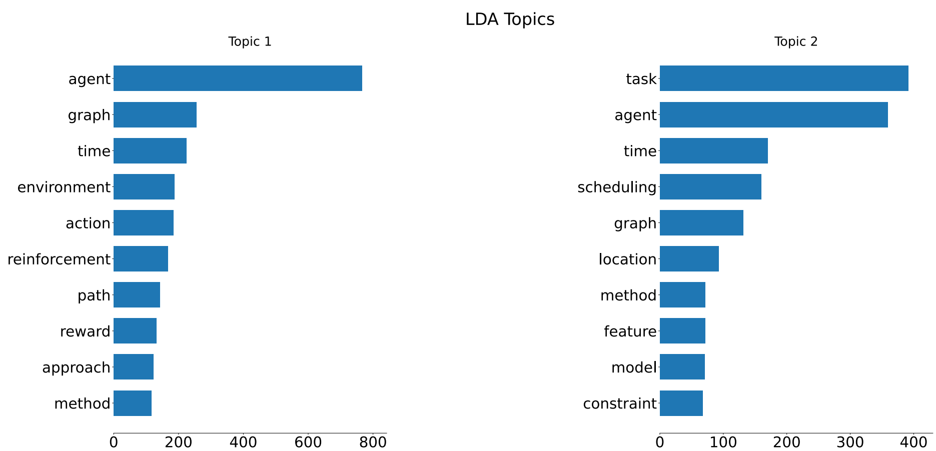

Figure 4: LDA Topics
Source: Own

Analyzing the papers we find that there is no clear and unified way to model warehouses as graphs for GNN. Every paper has different hand-picked nodes, edges and features of graphs with while many papers highlight problem of not considering enough constraints and factors to generalize well to real-word environments (Begnardi et al., 2024; Paul et al., 2023; Pham & Bera, 2024; Z. Wang et al., 2022; Xiao et al., 2024). There is an idea of foundation models (like for large language models) for DRL-GNN paradigm (Munikoti et al., 2022). Such model in warehouse setting (foundational warehouse DRL-GNN model) would greatly decrease the both computational and modeling difficulty of training DRL-GNN models for any optimization or decision-making task in warehouse environment while possibly having more wholistic view on warehouse environment which could in term give better performance in real word applications.

3 Discussion

In this paper we identified general research topics in use of GNNs and DRL in warehouse settings. We found a lack of research on the use of GNN in warehouse settings which implies possibilities for future research on modeling warehouse as graphs for use on any downstream optimization task or decision making. Use of GNNs with combination with DRL provide promising results compared to other optimization methods. The use of both GNNs and DRL together for optimizing order picking paths is only explored in one conference paper. Further research is needed to explore the use of GNNs and DRL in optimization of warehouse processes such as optimizing order picking paths.

Acknowledgment

The authors acknowledge the financial support from the Slovenian Research and Innovation Agency (research core funding No. P5-0018), the Ministry of Higher Education, Science, and Innovation of the Republic of Slovenia as part of the Next Generation EU National Recovery and Resilience Plan (Grants No. 3330-22-3515; NOO No: C3330-22-953012).

Notes

This systematic literature review follows the PRISMA methodology (Page et al., 2021) to ensure transparency in selecting and analyzing studies on DRL and GNN for warehouse optimization. The findings highlight a research gap in applying GNNs to order picking and batching, with only one study integrating both approaches (Begnardi et al., 2024). DRL is commonly used for real-time joint optimization in areas like order batching, AGV routing, and human-robot collaboration. However, the lack of a standardized warehouse graph model limits the applicability of GNNs. Future research should focus on developing a foundational DRL-GNN model to enhance computational efficiency and real-world adoption.

References

Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep Reinforcement Learning: A Brief Survey. *IEEE Signal Processing Magazine*, *34*(6), 26–38. IEEE Signal Processing Magazine. https://doi.org/10.1109/MSP.2017.2743240

Begnardi, L., Baier, H., Jaarsveld, W. van, & Zhang, Y. (2024). Deep Reinforcement Learning for Two-sided Online Bipartite Matching in Collaborative Order Picking. *Proceedings of the 15th*

Asian Conference on Machine Learning, 121–136. https://proceedings.mlr.press/v222/begnardi24a.html

Bello, I., Pham, H., Le, Q. V., Norouzi, M., & Bengio, S. (2017). *Neural Combinatorial Optimization with Reinforcement Learning* (No. arXiv:1611.09940). arXiv. https://doi.org/10.48550/arXiv.1611.09940

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, *3*(null), 993–1022.

Bock, S., Bomsdorf, S., Boysen, N., & Schneider, M. (2025). A survey on the Traveling Salesman Problem and its variants in a warehousing context. *European Journal of Operational Research*, *322*(1), 1–14. https://doi.org/10.1016/j.ejor.2024.04.014

Cals, B., Zhang, Y., Dijkman, R., & van Dorst, C. (2021). Solving the online batching problem using deep reinforcement learning. *Computers & Industrial Engineering*, *156*, 107221. https://doi.org/10.1016/j.cie.2021.107221

Chen, F., Wang, L., Zhang, J., Sun, Q., & Lai, P. (2024). An Attention Model Based Heuristic Algorithm for Storage Selection and Order Picking Path Joint Optimization. *2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 1159–1164. https://doi.org/10.1109/CSCWD61410.2024.10580420

Cheng, B., Wang, L., Tan, Q., & Zhou, M. (2024). A deep reinforcement learning hyper-heuristic to solve order batching problem with mobile robots. *Applied Intelligence*, *54*(9), 6865–6887. https://doi.org/10.1007/s10489-024-05532-9

Cheng, B., Xie, T., Wang, L., Tan, Q., & Cao, X. (2024). Deep reinforcement learning driven cost minimization for batch order scheduling in robotic mobile fulfillment systems. *Expert Systems with Applications*, *255*, 124589. https://doi.org/10.1016/j.eswa.2024.124589

Cho, S.-H., Shin, W.-J., Ahn, J., Joo, S., & Kim, H.-J. (2024). Dynamic Crane Scheduling with Reinforcement Learning for a Steel Coil Warehouse. *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 16545–16552. https://doi.org/10.1109/ICRA57147.2024.10610858

Dehghan, A., Cevik, M., & Bodur, M. (2023). *Neural Approximate Dynamic Programming for the Ultra-fast Order Dispatching Problem* (No. arXiv:2311.12975). arXiv. https://doi.org/10.48550/arXiv.2311.12975

Drakaki, M., & Tzionas, P. (2017). Manufacturing Scheduling Using Colored Petri Nets and Reinforcement Learning. *Applied Sciences*, *7*(2), Article 2. https://doi.org/10.3390/app7020136

Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Systematic Reviews*, *18*(2), e1230. https://doi.org/10.1002/cl2.1230

Kaiser, P., Thevapalan, A., Roidl, M., Kern-Isberner, G., & Wilhelm, M. (2023). *Towards Finding Optimal Solutions For Constrained Warehouse Layouts Using Answer Set Programming.* https://repo.uni-hannover.de/handle/123456789/13590

Kamoshida, R., & Kazama, Y. (2017). Acquisition of Automated Guided Vehicle Route Planning Policy Using Deep Reinforcement Learning. *2017 6th IEEE International Conference on Advanced Logistics and Transport (ICALT)*, 1–6. https://doi.org/10.1109/ICAdLT.2017.8547000

Knippenberg, M. van, Holenderski, M., & Menkovski, V. (2021). Time-Constrained Multi-Agent Path Finding in Non-Lattice Graphs with Deep Reinforcement Learning. *Proceedings of The 13th Asian Conference on Machine Learning*, 1317–1332. https://proceedings.mlr.press/v157/knippenberg21a.html

Krnjaic, A., Steleac, R. D., Thomas, J. D., Papoudakis, G., Schäfer, L., Keung To, A. W., Lao, K.-H., Cubuktepe, M., Haley, M., Börsting, P., & Albrecht, S. V. (2024). Scalable Multi-Agent Reinforcement Learning for Warehouse Logistics with Robotic and Human Co-Workers. *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 677–684. https://doi.org/10.1109/IROS58592.2024.10802813

Li, H., Zhang, Y., Xiao, Z., & Li, D. (2022). Order Sequencing Problem in a Robotic Mobile Fulfillment System. *2022 34th Chinese Control and Decision Conference (CCDC)*, 74–79. https://doi.org/10.1109/CCDC55256.2022.10033985

Li, K., Liu, T., Ram Kumar, P. N., & Han, X. (2024). A reinforcement learning-based hyper-heuristic for AGV task assignment and route planning in parts-to-picker warehouses. *Transportation Research Part E: Logistics and Transportation Review*, *185*, 103518. https://doi.org/10.1016/j.tre.2024.103518

Li, X., Gao, J., & Li, Y. (2023). Multi-Agent Path Finding Based on Graph Neural Network. *2023 42nd Chinese Control Conference (CCC)*, 5458–5463. https://doi.org/10.23919/CCC58697.2023.10240345

Mahmoudinazlou, S., Sobhanan, A., Charkhgard, H., Eshragh, A., & Dunn, G. (2024). *Deep Reinforcement Learning for Dynamic Order Picking in Warehouse Operations* (No. arXiv:2408.01656). arXiv. https://doi.org/10.48550/arXiv.2408.01656

Munikoti, S., Agarwal, D., Das, L., Halappanavar, M., & Natarajan, B. (2022). *Challenges and Opportunities in Deep Reinforcement Learning with Graph Neural Networks: A Comprehensive review of Algorithms and Applications* (No. arXiv:2206.07922). arXiv. https://doi.org/10.48550/arXiv.2206.07922

Neves-Moreira, F., & Amorim, P. (2024). Learning efficient in-store picking strategies to reduce customer encounters in omnichannel retail. *International Journal of Production Economics*, *267*, 109074. https://doi.org/10.1016/j.ijpe.2023.109074

Niu, Y., Schulte, F., & Negenborn, R. R. (2021). Human Aspects in Collaborative Order Picking – Letting Robotic Agents Learn About Human Discomfort. *Procedia Computer Science*, *180*, 877–886. https://doi.org/10.1016/j.procs.2021.01.338

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Paul, S., Li, W., Smyth, B., Chen, Y., Gel, Y., & Chowdhury, S. (2023). Efficient Planning of Multi-Robot Collective Transport using Graph Reinforcement Learning with Higher Order Topological Abstraction. *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 5779–5785. https://doi.org/10.1109/ICRA48891.2023.10161517

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830.

Perumaal Subramanian, S., & Kumar Chandrasekar, S. (2024). Simultaneous allocation and sequencing of orders for robotic mobile fulfillment system using reinforcement learning algorithm. *Expert Systems with Applications*, *239*, 122262. https://doi.org/10.1016/j.eswa.2023.122262

Pham, P., & Bera, A. (2024). Optimizing Crowd-Aware Multi-Agent Path Finding through Local Communication with Graph Neural Networks. *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 8913–8913. https://doi.org/10.1109/IROS58592.2024.10801969

Shelke, O., Pathakota, P., Chauhan, A., Khadilkar, H., Meisheri, H., & Ravindran, B. (2023). *Multi-Agent Learning of Efficient Fulfilment and Routing Strategies in E-Commerce* (No. arXiv:2311.16171). arXiv. https://doi.org/10.48550/arXiv.2311.16171

Smit, I. G., Bukhsh, Z., Pechenizkiy, M., Alogariastos, K., Hendriks, K., & Zhang, Y. (2024). *Learning Efficient and Fair Policies for Uncertainty-Aware Collaborative Human-Robot Order Picking* (No. arXiv:2404.08006). arXiv. https://doi.org/10.48550/arXiv.2404.08006

Tang, H., Wang, A., Xue, F., Yang, J., & Cao, Y. (2021). A Novel Hierarchical Soft Actor-Critic Algorithm for Multi-Logistics Robots Task Allocation. *IEEE Access*, *9*, 42568–42582. IEEE Access. https://doi.org/10.1109/ACCESS.2021.3062457

Wang, D., Jiang, J., Ma, R., & Shen, G. (2022). Research on Hybrid Real-Time Picking Routing Optimization Based on Multiple Picking Stations. *Mathematical Problems in Engineering*, *2022*(1), 5510749. https://doi.org/10.1155/2022/5510749

Wang, X., Zhang, L., Wang, L., Zuñiga, E. R., Wang, X. V., & Flores-García, E. (2025). Dynamic multi-tour order picking in an automotive-part warehouse based on attention-aware deep reinforcement learning. *Robotics and Computer-Integrated Manufacturing*, *94*, 102959. https://doi.org/10.1016/j.rcim.2025.102959

Wang, Z., Liu, C., & Gombolay, M. (2022). Heterogeneous graph attention networks for scalable multi-robot scheduling with temporospatial constraints. *Autonomous Robots*, *46*(1), 249–268. https://doi.org/10.1007/s10514-021-09997-2

Wu, S.-C., Chiu, W.-Y., & Wu, C.-F. (2024). Deep Reinforcement Learning for Task Assignment and Shelf Reallocation in Smart Warehouses. *IEEE Access*, *12*, 58915–58926. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3392752

Xiao, Z., Li, P., Liu, C., Gao, H., & Wang, X. (2024). MACNS: A generic graph neural network integrated deep reinforcement learning based multi-agent collaborative navigation system for dynamic trajectory planning. *Information Fusion*, *105*, 102250. https://doi.org/10.1016/j.inffus.2024.102250

Yoshitake, H., & Abbeel, P. (2023). The Impact of Overall Optimization on Warehouse Automation. *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 1621–1628. https://doi.org/10.1109/IROS55552.2023.10342333

Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., & Sun, M. (2020). Graph neural networks: A review of methods and applications. *AI Open*, *1*, 57–81. https://doi.org/10.1016/j.aiopen.2021.01.001

Zhou, L., Lin, C., Ma, Q., & Cao, Z. (2022). A Learning-based Iterated Local Search Algorithm for Order Batching and Sequencing Problems. *2022 IEEE 18th International Conference on Automation Science and Engineering (CASE)*, 1741–1746. https://doi.org/10.1109/CASE49997.2022.9926486

About the authors

Nejc Čelik is teaching assistant of Information Systems at University of Maribor, Faculty of Organizational Sciences. His research interests include the use of artificial intelligence in information systems.

Andrej Škraba is Professor of Information Systems at the University of Maribor, Faculty of Organizational Sciences. His research interests cover systems theory, modeling and simulation, cyber-physical systems, and decision processes. Prof. Škraba has received a Bronze Medal from the University of Maribor for successful research and pedagogical work in the field of Systems Modeling and Simulation and serves as a Vice Dean for Development at the Faculty of Organizational Sciences. A. Škraba is a board member of the EUROMICRO Association for Research and Development of Information Technology and Microelectronics.