Towards a Real-Time and Energy-Efficient Edge AI Camera Architecture in Mega Warehouse Environment

Yusuke Asai*, Yuki Mori*, Keisuke Higashiura*, Kodai Yokoyama*,
Shin Katayama*, Kenta Urano*, Takuro Yonezawa*†, Nobuo Kawaguchi*†
*Graduate School of Engineering, Nagoya University, Japan
{asayu, ymori, urachan, yokochin, shinsan, vrano}@ucl.nuee.nagoya-u.ac.jp
†Institutes of Innovation for Future Society, Nagoya University, Japan
{takuro, kawaguti}@nagoya-u.jp

***Abstract*—In response to the exponential growth of e-commerce, our paper addresses the transformation of warehouse operations towards mega warehouses, necessitating advanced digitalization. We focus on the integration and optimization of edge AI technologies to enhance operational efficiency, accuracy, and timeliness in logistic warehouses. We propose a novel edge AI architecture tailored to the processes of warehouse operations ensuring that digitalization aligns with operational workflows. Utilizing a real-world warehouse equipped with over 60 cameras as a testbed, we demonstrate the practical application and benefits of edge computing in logistics. Our experiments have shown significant potential improvements in energy efficiency and timeliness, crucial metrics for the successful integration of edge AI technologies in warehouses.**

***Index Terms*—edge AI camera, smart warehouse, warehouse automation, digitalization with cameras**

I. Introduction

In an era marked by the relentless growth of e-commerce, the logistics and supply chain landscape is undergoing a profound transformation. This shift is most visibly manifested in the evolution of warehouse operations, which are transitioning from large-scale warehouses to what are now being termed as mega warehouses. This significant expansion is not merely a change in scale; it necessitates a comprehensive reevaluation and digitalization of their technological infrastructures. Our research seeks to illuminate the critical role of edge AI technologies in facilitating this transition, emphasizing their potential to automate warehouse operations through enhanced efficiency, accuracy, and timeliness.

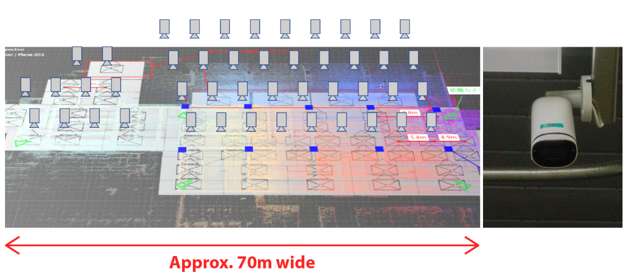

Fig. 1. Schematic of the camera array installation locations and the actual installed camera

Digitalization in this context encompasses a wide array of technologies, from sophisticated robotic systems to advanced data processing capabilities, all aimed at optimizing warehouse operations from receipt to dispatch. In our prior works, we have contributed to the field of warehouse digitalization such as the implementation of Autonomous Mobile Robot (AMR) robotization [1], [2] and the development of a semi-automated image annotation method [3] with digital twin technologies. Building on this foundation, our paper aims to delve into the integration of edge AI technologies within logistic warehouses. By assessing the performance and energy consumption of various deep-learning methods, we seek to provide an understanding of how these technologies can be optimized to meet the challenges of warehouse expansion.

We introduce an edge AI architecture tailored for warehouse operations, addressing inbound activities focusing on energy-efficient design for image recognition. Our testbed warehouse, equipped with over 60 cameras as shown in Fig.1, demonstrates this architecture's practicality and the benefits of edge computing in logistics. Experiments show notable improvements in energy efficiency and timeliness, highlighting the advantages of edge AI in warehouses.

The main contributions of this paper are as follows:

1) We introduce an energy-efficient edge AI architecture designed with the complexities and demands of actual warehouse environments in mind.
2) We implement edge AI technology in a real-world setup to refine the architecture in real operational conditions with theoretical analysis.
3) We demonstrate that we can achieve energy savings up to about 99% in our experiment comparing PC-GPU setups to edge AI processor setups.

The remainder of this paper is organized as follows. First, Section II studies previous work related to our study. Next, Section III explains our edge AI architecture perspective. Section IV presents our preliminary study for warehouse digitalization. Section V describes a warehouse's evolution to explore the implications of technological requirements. Section VI shows the results and evaluations for our experiments. Finally, Section VII concludes our work and presents future research direction.

II. Related Work

The adoption of edge AI technologies, particularly in configurations involving a substantial number of cameras, remains relatively nascent [4], [5]. One of the primary hurdles to widespread adoption is the inherent complexity of logistic environments [6]. These settings are characterized by a high degree of variability, thus complicating the deployment of standardized solutions.

Despite these challenges, edge AI emerges as a promising paradigm, capable of significantly reducing the strain on central computational resources [7], [8]. By enabling neural network inferences to be processed directly at the source of data collection, edge AI architectures offer a pathway to mitigate bandwidth and latency issues associated with cloud-based processing [9]. This is particularly relevant in scenarios demanding real-time analytics, such as monitoring and management within large warehouse environments [10], [11]. Prior research has extensively covered worker allocation [12], storage optimization [13], and the integration of robotics within operational workflows [14], [15]. Real-time analytics provided by edge AI can dynamically adjust worker assignments, optimize storage locations in response to changing inventory levels, and coordinate robotic systems efficiently. Moreover, addressing operational bottlenecks and standardizing processes emerge as critical considerations [16]. By identifying and managing choke points in real-time, warehouses can significantly enhance operational flow, reducing idle times and improving overall productivity [17].

Traditionally, Warehouse Management Systems (WMS) and Enterprise Resource Planning (ERP) software have been the primary sources of operational data [18], [19]. These systems are invaluable for overarching management and planning but often fall short in capturing the granular details of day-to-day operations [20]. The data collected tend to be sparse, failing to accurately represent the dynamic nature of warehouse activities. This discrepancy highlights a gap in operational intelligence—a gap that edge AI camera is uniquely positioned to fill.

Nonetheless, the task of implementing image processing at the edge, especially in the context of large-scale operations like those found in warehouse settings, is fraught with difficulties [3]. The sheer volume of data generated by numerous cameras, combined with the need for immediate processing and interpretation, presents a challenge due to the high computational power needed to analyze images from numerous cameras [21]. As warehouses expand in size and complexity, the importance of maintaining energy efficiency becomes as crucial as the effectiveness of AI. Balancing the demands for computational intensity with the need for energy conservation is key, necessitating innovative edge AI solutions that are both powerful and sustainable for large-scale applications.

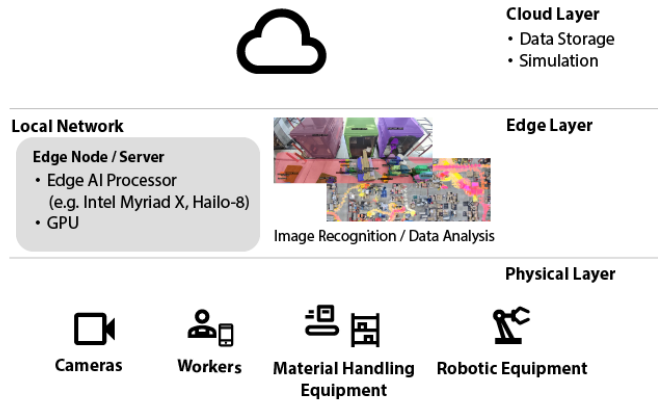

Fig. 2. Our Envisioned Edge AI System Architecture

Fig. 3. Luxonis-OAK-D-Pro-W-PoE Camera

III. Digitalizing Warehouse: Edge AI Architecture Perspective

This section presents our edge AI system architecture designed for warehouse dynamics, discusses camera installation for our testbed warehouse, and describes typical tasks in multi-product warehouses.

A. System Architecture

Fig.2 outlines a cloud-edge architecture for warehouse operations across three layers: Physical, Edge, and Cloud. The Physical Layer features devices like cameras and robots that directly monitor and manage warehouse tasks. The Edge Layer emphasizes localized data processing, with a spotlight on leveraging deep learning techniques for image recognition tasks, for immediate operational improvements and reduced delays. The Cloud Layer serves as the hub for data storage and simulations, enhancing optimization with its computational strength. Together, these layers create a unified system that combines edge-based real-time actions with cloud-level strategic planning, ensuring efficient and adaptive warehouse operations.

B. Camera Installation in Our Target Environment

Our testbed warehouse in Aichi Prefecture, Japan, exemplifies the integration of technology in logistics. It uses 66 IP cameras (H.View HV-800G2A5) to monitor the first floor of this 5-storey facility, capturing 1.2TB of video data daily. Six cameras provide broad coverage and five focus on truck berth activities. However, a significant majority, 55 out of the 66 cameras, are deliberately pointed perpendicular to the

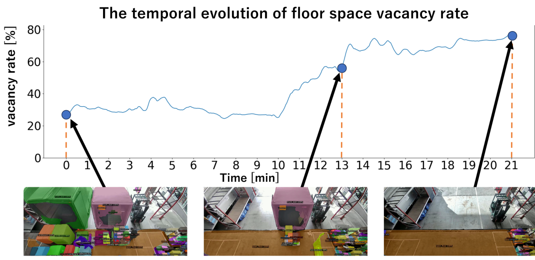

Fig. 4. Temporal evolution of floor vacancy rate in a truck berth during operation hours

Fig. 5. Visualization of moving objects' paths from stitched images on the 1st floor of the warehouse

Fig. 6. Pallet detection and tracking with AI camera (left: elevator right: vertical transport system)

warehouse floor as shown in Fig.1. This unique setup provides a direct downward view, allowing for precise monitoring and analysis of floor activities, offering the ability to track movements and optimize operations with detail.

To increase oversight on floors 2-5, we added 4 edge AI cameras (Luxonis-OAK-D-Pro-W-PoE) shown in Fig.3, enhancing visibility of inbound activities. This expansion, based on discussions with warehouse staff, addresses the greater variability in inbound versus outbound operations due to unpredictable goods arrival times.

C. Operational Workflow: Inbound Activities

Warehouse operations split into inbound and outbound activities, with our study focusing on the former. Shipments are first inspected for quality upon arrival, then sorted by size, shape, and weight for storage. Goods wait in put-away areas before storage, where manual handling is common, despite the presence of automated systems, to ensure organized space use.

Our multi-storey testbed warehouse employs a specialized system where each floor handles specific product types, requiring goods to be sorted and transported to the correct level via an elevator or a vertical transport system for pallets. Standardized pallets are used for moving goods, streamlining handling and storage for improved operational efficiency.

Despite the variance across different types of warehouses, a streamlined sequence of inbound activities – encompassing receiving, inspection, sorting, put-away, and storage – is universally observed [12]. This systematic approach can also be applied to our testbed warehouse's operations, embodying the structured management of inbound goods from arrival to storage.

IV. Preliminary Study For Logistics Warehouse Digitalization

Traditionally, workers scan product barcodes with handheld devices during inspection and storage, ensuring data integration into the Warehouse Management System (WMS). However, WMS data is often sparse, offering snapshots instead of the continuous information flow needed for full optimization. This underscores the need for comprehensive digitalization strategies for detailed, real-time data capture and analysis.

We explore the initial steps taken towards digitalizing the workflow of our testbed warehouse.

At Receiving (Truck Berth)

Utilizing Mask R-CNN [22], trained with the dataset of 1791 annotated images, we employ segmentation methods to assess how effectively truck berth floor space is used. As in our prior work [23], Fig.4 shows the temporal evolution of the floor space vacancy rate that reveals the condition and efficiency of the truck berth. This analysis paves the way for optimizing the scheduling and unloading processes.

From Inspection to Put-Away (1st Floor)

We stitch video from multiple 1st-floor cameras to track goods and workers, analyzing movement by comparing consecutive frames, as shown in Fig.5. This early-stage process enhances space optimization and task monitoring, offering insights into workflow efficiency.

At Put-Away (Upper Floor)

For goods designated for storage on upper floors, palletized items are transported via the elevator or the specialized vertical transport system as shown in Fig.6. At this juncture, we employ YOLOv5 [24] for pallet detection, coupled with Kalman filter and the Hungarian algorithm for tracking, allowing us to accurately count the number of pallets that make their way to the put-away area.

AI cameras track pallets on upper floors, capturing movements near elevators and vertical transport systems. Elsewhere, IP cameras record and stream real-time footage via RTSP for analysis and monitoring, leveraging GPU power. Our experimental setup is divided into two groups: the IP camera-GPU Setup, using standard IP cameras and centralized GPUs for high-volume data, and the AI camera Setup, with AI cameras handling built-in processing.

TABLE I
WAREHOUSE COMPARISON

	W_{target}	$W_{planned}$
Size Category	Large	Mega
Quantity in Stock (items)	Approx. 365,000	Approx. 1,000,000
Total Floor Space (m^2)	23,964	89,864
Number of Floors	5	4
Estimated 1F Space (m^2)	4,792	22,466

V. WAREHOUSE EVOLUTION: LARGE TO MEGA

The shift to larger, or mega, warehouses is fueled by e-commerce growth, automation, and demand for quick delivery. Our study focuses on a real case where our partner company plans to build a mega warehouse, offering a chance to examine the technological needs this expansion entails.

A. Mega Warehouse Planning for Digitalization

Our analysis highlights the growing significance of edge devices in large warehouses, where edge computing is key to reducing latency and improving energy efficiency, vital for real-time processing and sustainable operations.

As shown in Table.I, we compare our current target warehouse (W_{target}) with the proposed mega warehouse ($W_{planned}$) categorized by floor space [25]. We examine key parameters such as total floor space, stock capacity, and estimated first-floor (1F) space relative to the overall floor space and number of floors.

To digitalize the mega warehouse, its first-floor area will be 4.7 times larger than our current warehouse. We'll increase our camera count from 66 to 310, aiming for 5 fps per camera. This change raises our processing needs from 330 fps to 1560 fps, highlighting the need to boost our computational power to handle the increased data flow efficiently.

B. Latency and Energy Consumption

We compare the IP camera-GPU architecture and AI camera architecture within our target warehouse and the planned mega warehouse. This analysis is grounded in the following key assumptions:

- The power consumption for each GPU is established at 450W, specifically using NVIDIA RTX 3090 units.
- Our model is designed to process 30 frames per second (fps) per GPU [26].
- Each GPU is capable of processing the video streams (5fps per stream) from 6 cameras concurrently.
- IP cameras (H.View HV-800G2A5) in our system consume 7.5W each, whereas AI cameras (Luxonis-OAK-D-Pro-W-PoE) operate at a slightly lower consumption rate of 7.0W.

Our target warehouse requires 11 GPUs, while the expansive mega warehouse would ideally need 52 GPUs to ensure real-time processing. This setup corresponds to 66 cameras in the target environment and 310 cameras for the first floor of the mega warehouse.

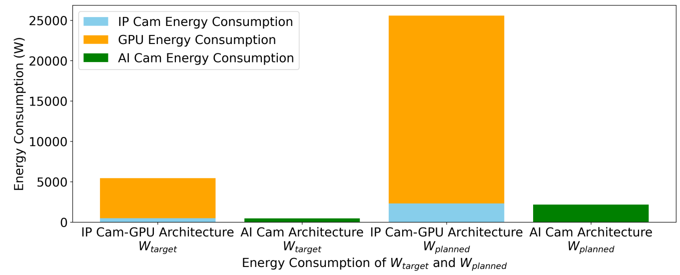

Fig. 7. Energy Consumption of Our Target Warehouse and Planned Mega Warehouse

1) Power Consumption: Fig.7 presents a stark contrast between the two architectural setups. Utilizing the IP camera-GPU architecture, the energy demands amount to 5445W for the target warehouse and escalate to 25575W for the mega warehouse. Conversely, adopting an AI camera architecture significantly reduces these figures to 462W and 2170W, respectively, highlighting the superior energy efficiency of AI cameras.

2) Latency and Energy Considerations for IP camera-GPU Architecture: Our analysis also introduces a metric, the "GPU Utilization Ratio (R_{util})", which quantifies the actual number of GPUs utilized versus the ideal number required given by

$$R_{util} = N_{actual}/N_{needed}$$

where N_{actual} and N_{needed} are process frames that the system can actually process and need to process by GPUs, respectively. This measure captures the trade-off between achieving timeliness in processing and optimizing resource allocation. The latency per frame can be calculated as follows: $\frac{1}{N_{actual}} - \frac{1}{N_{needed}}$. Hence, we can multiply the per-frame latency by the total number of frames processed to calculate the estimated latency.

We show the figures outlining the estimated latency and energy consumption for W_{target} and $W_{planned}$ in Fig.8 and 9, respectively. The figures show how GPU use affects computational efficiency, revealing the balance between reducing latency and higher energy use. As our metric increases, latency significantly drops but plateaus near a value of one, showing diminishing returns. Meanwhile, adding more GPUs steadily raises energy consumption, underlining a clear link between GPU use and energy needs.

The number of GPUs used depends on processing needs and balancing speed with resource efficiency. Despite limits, edge AI cameras processing on-device excel in speed and energy use, highlighting edge AI's value in reducing latency and energy consumption.

VI. EXPERIMENT

In this section, we describe our experiments for assessing the performance and energy efficiency of image recognition tasks by five YOLOv8 network variants using a high-spec PC, an AI camera, and a Hailo-8 processor [27] that can execute deep learning applications at the edge with high efficiency and speed.

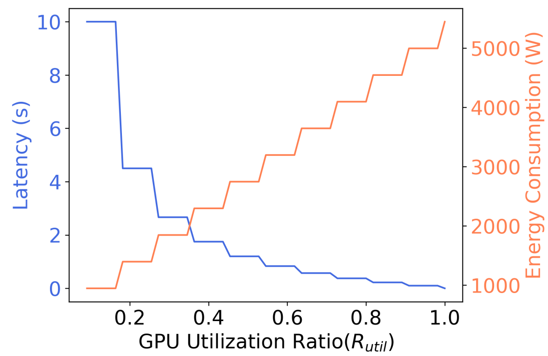

Fig. 8. Latency and Energy Consumption Based on R_{util} of Our Target Warehouse

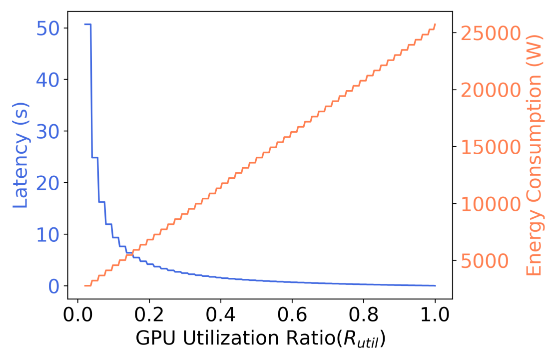

Fig. 9. Latency and Energy Consumption Based on R_{util} of Planned Mega Warehouse

A. Experimental Setup

We evaluate the performance of five distinct YOLOv8 [24] networks with the input resolution (HxWxC) of 640x640x3: YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x, focusing on their processing speed and power consumption. Table.II details the number of network layers and network parameters, providing a clear overview of each model's complexity and computational demands.

Our experimental apparatus comprises a PC outfitted with a Ryzen 9 5950X CPU and an NVIDIA RTX 3090 GPU, an AI camera (Luxonis-OAK-D-Pro-W-PoE) featuring a Myriad X Vision Processing Unit (VPU), and a Hailo-8 processor with Rock5B (Octa Core ARM Processor RK3588) shown in Fig.10. The Hailo-8 is an advanced and specialized artificial intelligence (AI) processor designed to deliver high performance in edge devices.

For controlled testing, our PC setup uses a prerecorded Full HD video at 5 fps, eliminating latency from network delays or streaming issues and focusing on hardware and software processing. The Hailo-8 processor tests also use this video. The AI camera setup is evaluated using a live feed, with the frame rate adjusted according to its processing capacity.

The REVEX ET30D meter tracks real-time power consumption in PC and AI camera setups by connecting to an outlet. Power use for inference tasks is calculated by subtracting the average idle energy, measured three times over 15 seconds, from total use. For the Hailo-8, power data comes directly from the processor's API.

TABLE II
YOLOv8 Variants Comparison

Model	Layers	Prams (Million)
YOLOv8n	225	3.2
YOLOv8s	225	11.2
YOLOv8m	295	25.9
YOLOv8l	365	43.7
YOLOv8x	365	68.2

TABLE III
Performance Metrics for YOLOv8 Models Across Different Setups

	PC		Luxionis OAK		Hailo-8	
Model	fps	Power (W)	fps	Power (W)	fps	Power (W)
YOLOv8n	56.8	128.3	8.0	7.0	191.0	2.4
YOLOv8s	57.7	146.2	3.5	7.0	106.2	2.9
YOLOv8m	45.7	210.8	1.6	7.0	55.0	3.1
YOLOv8l	40.3	242.3	0.9	7.0	29.5	3.3
YOLOv8x	34.4	276.8	0.5	7.0	12.6	2.9

B. Results and Discussions

We present a detailed analysis of the processing speed and power consumption for each YOLO model across three different setups: PC, AI camera, and Hailo processor. This comprehensive evaluation is encapsulated in Table.III.

1) PC Setup: The PC setup showcases superior computing performance across the board, except for YOLOv8n, YOLOv8s, and YOLOv8m models, where the Hailo setup excels. However, this high performance comes at a cost, with significantly higher energy consumption compared to the AI camera and Hailo setups. Notably, as the complexity of the network model increases, there is a marked increase in energy consumption, highlighting the trade-off between computational power and energy efficiency.

2) AI Camera Setup: In contrast, the AI camera setup demonstrates relatively lower computing performance, which diminishes further as the complexity of the network model escalates. Despite this limitation in processing capacity, the AI camera maintains a consistent energy consumption rate across all YOLO models. This steadiness in power usage, regardless of model complexity, underscores the AI camera's advantage in energy efficiency.

3) Hailo Setup: The Hailo processor stands out for its robust performance, maintaining relatively high processing speeds even as the complexity of the network models increases. Similar to the AI camera setup, the Hailo processor exhibits relatively stable energy consumption across different YOLO models.

Using the most simplest model in our experiments, we achieved a reduction in energy usage of 99.45% per processing frame when compared to the PC-GPU setup and the Hailo-

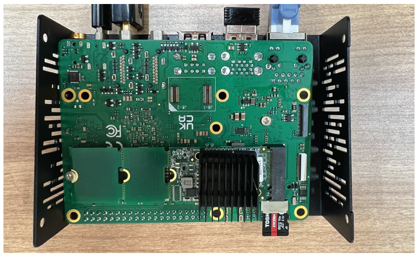

Fig. 10. Hailo-8 processor with Rock5B (Octa Core ARM Processor RK3588)

8 processor configuration. This demonstrates the potential for substantial energy savings in warehouse digitalization. Moving forward, additional testing, especially with the Hailo processor, will be crucial to understanding the full potential of these configurations in practical applications.

VII. Conclusion

In this paper, we explore the use of edge AI technologies within warehouse operations revealing a promising pathway toward addressing the challenges posed by the transition to mega warehouses. By proposing a novel edge AI architecture for the dynamic environment of warehouses, we have demonstrated significant potential for improving computational efficiency and timeliness. These findings underscore the critical role of edge AI in the digital transformation of warehouses, offering scalable solutions that meet the growing demands of the logistics industry. As we move forward, our research lays the groundwork for further advancements in warehouse digitalization, promising a more efficient, responsive, and sustainable future for the sector.

Acknowledgment

This paper is based on results obtained from projects JPNP23003 and JPNP23025 commissioned by the New Energy and Industrial Technology Development Organization (NEDO). This paper is partially supported by TRUSCO Nakayama Corp, JST KAKENHI(22K18422) and JST CREST(JPMJCR22M4).

References

[1] Y. Fukushima, Y. Asai, S. Aoki, T. Yonezawa, and N. Kawaguchi, "Digimobot: Digital twin for human-robot collaboration in indoor environments," in *2021 IEEE Intelligent Vehicles Symposium (IV)*, pp. 55–62, IEEE, 2021.

[2] S. Aoki, T. Yonezawa, and N. Kawaguchi, "Robotnest: Toward a viable testbed for iot-enabled environments and connected and autonomous robots," *IEEE Sensors Letters*, vol. 6, no. 2, pp. 1–4, 2022.

[3] K. Higashiura, K. Yokoyama, Y. Asai, H. Shimosato, K. Kano, S. Katayama, K. Urano, T. Yonezawa, and N. Kawaguchi, "Semi-automated framework for digitalizing multi-product warehouses with large scale camera arrays," in *2024 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, IEEE, 2024.

[4] G. Ananthanarayanan, P. Bahl, P. Bodík, K. Chintalapudi, M. Philipose, L. Ravindranath, and S. Sinha, "Real-time video analytics: The killer app for edge computing," *computer*, vol. 50, no. 10, pp. 58–67, 2017.

[5] S. Wan, S. Ding, and C. Chen, "Edge computing enabled video segmentation for real-time traffic monitoring in internet of vehicles," *Pattern Recognition*, vol. 121, p. 108146, 2022.

[6] R. Singh and S. S. Gill, "Edge ai: a survey," *Internet of Things and Cyber-Physical Systems*, 2023.

[7] L. Lin, X. Liao, H. Jin, and P. Li, "Computation offloading toward edge computing," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1584–1607, 2019.

[8] H. Lin, S. Zeadally, Z. Chen, H. Labiod, and L. Wang, "A survey on computation offloading modeling for edge computing," *Journal of Network and Computer Applications*, vol. 169, p. 102781, 2020.

[9] X. Chen, J. Zhang, B. Lin, Z. Chen, K. Wolter, and G. Min, "Energy-efficient offloading for dnn-based smart iot systems in cloud-edge environments," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 3, pp. 683–697, 2021.

[10] F. Halawa, H. Dauod, I. G. Lee, Y. Li, S. W. Yoon, and S. H. Chung, "Introduction of a real time location system to enhance the warehouse safety and operational efficiency," *International Journal of Production Economics*, vol. 224, p. 107541, 2020.

[11] G. Zhang, X. Shang, F. Alawneh, Y. Yang, and T. Nishi, "Integrated production planning and warehouse storage assignment problem: An iot assisted case," *International Journal of Production Economics*, vol. 234, p. 108058, 2021.

[12] O. Ganbold, K. Kundu, H. Li, and W. Zhang, "A simulation-based optimization method for warehouse worker assignment," *Algorithms*, vol. 13, no. 12, p. 326, 2020.

[13] F. N. Arif, B. Irawan, and C. Setianingsih, "Optimization of storage allocation for final goods warehouse using genetic algorithm," in *2020 2nd International Conference on Electrical, Control and Instrumentation Engineering (ICECIE)*, pp. 1–6, IEEE, 2020.

[14] T. Petković, J. Hvězda, T. Rybeckỳ, I. Marković, M. Kulich, L. Přeučil, and I. Petrović, "Human intention recognition for human aware planning in integrated warehouse systems," in *2020 28th Mediterranean Conference on Control and Automation (MED)*, pp. 586–591, IEEE, 2020.

[15] S. Srinivas and S. Yu, "Collaborative order picking with multiple pickers and robots: Integrated approach for order batching, sequencing and picker-robot routing," *International Journal of Production Economics*, vol. 254, p. 108634, 2022.

[16] W. Garn, J. Aitken, and R. Schmenner, "Smoothly pass the parcel: Implementing the theory of swift, even flow," *ResearchGate*, 2020.

[17] L. D. Fredendall, J. B. Craig, P. J. Fowler, and U. Damali, "Barriers to swift, even flow in the internal supply chain of perioperative surgical services department: a case study," *Decision Sciences*, vol. 40, no. 2, pp. 327–349, 2009.

[18] D. Minashkina and A. Happonen, "Decarbonizing warehousing activities through digitalization and automatization with wms integration for sustainability supporting operations," in *E3S Web of Conferences*, vol. 158, p. 03002, EDP Sciences, 2020.

[19] T. Kato and R. Kamoshida, "Multi-agent simulation environment for logistics warehouse design based on self-contained agents," *Applied Sciences*, vol. 10, no. 21, p. 7552, 2020.

[20] M. G. Khan, N. U. Huda, and U. K. U. Zaman, "Smart warehouse management system: Architecture, real-time implementation and prototype design," *Machines*, vol. 10, no. 2, p. 150, 2022.

[21] M.-A. Maleki, A. Nabipour-Meybodi, M. Kamal, A. Afzali-Kusha, and M. Pedram, "An energy-efficient inference method in convolutional neural networks based on dynamic adjustment of the pruning level," *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, vol. 26, no. 6, pp. 1–20, 2021.

[22] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969, 2017.

[23] K. Yokoyama, S. Katayama, K. Urano, T. Yonezawa, and N. Kawaguchi, "Digitization and analysis framework for warehouse truck berth," in *2023 Fourteenth International Conference on Mobile Computing and Ubiquitous Network (ICMU)*, pp. 1–4, IEEE, 2023.

[24] G. Jocher, A. Chaurasia, and J. Qiu, "Yolo by ultralytics," *URL: https://github. com/ultralytics/ultralytics*, 2023.

[25] D. Andreoli, A. Goodchild, and K. Vitasek, "The rise of mega distribution centers and the impact on logistical uncertainty," *Transportation Letters*, vol. 2, no. 2, pp. 75–88, 2010.

[26] S. Jung, H. Heo, S. Park, S.-U. Jung, and K. Lee, "Benchmarking deep learning models for instance segmentation," *Applied Sciences*, vol. 12, no. 17, p. 8856, 2022.

[27] R. Griessl, F. Porrmann, N. Kucza, K. Mika, J. Hagemeyer, M. Kaiser, M. Porrmann, M. Tassemeier, M. Flottmann, F. Qararyah, *et al.*, "Evaluation of heterogeneous aiot accelerators within vedliot," in *2023 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 1–6, IEEE, 2023.