

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



Large to Mega Warehouse

Warehouses are becoming mega.

- E-commerce Growth
 - Handling large volumes and fast deliveries of goods
- Advanced Automation
 - Use of advanced technologies like robotics and AI

Inbound logistics (receiving, inspection, etc) still requires manual labor

- The forms of incoming shipments are **not standerlized**: different manufacturers and packaging forms
- Demand for capturing movements of workers and shipments for analysing inbound activities



Automated storage / picking robot

Digitalization plays a critical roll for efficiency improvement.

Large to Mega Warehouse



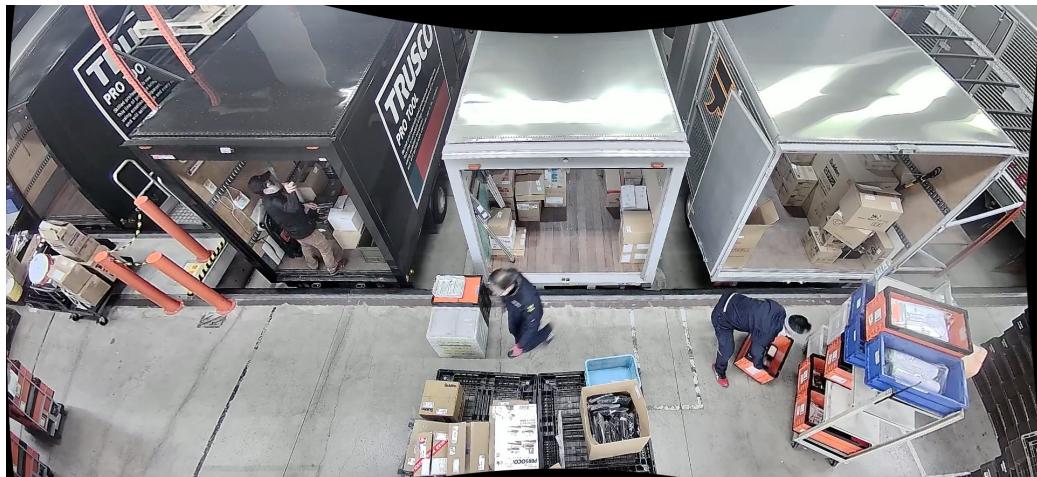
Video in
Inbound Area

Inbound tasks
(Temporary storage,
Inspection,
Allocation,
Relocation(to Lift),
Storage)

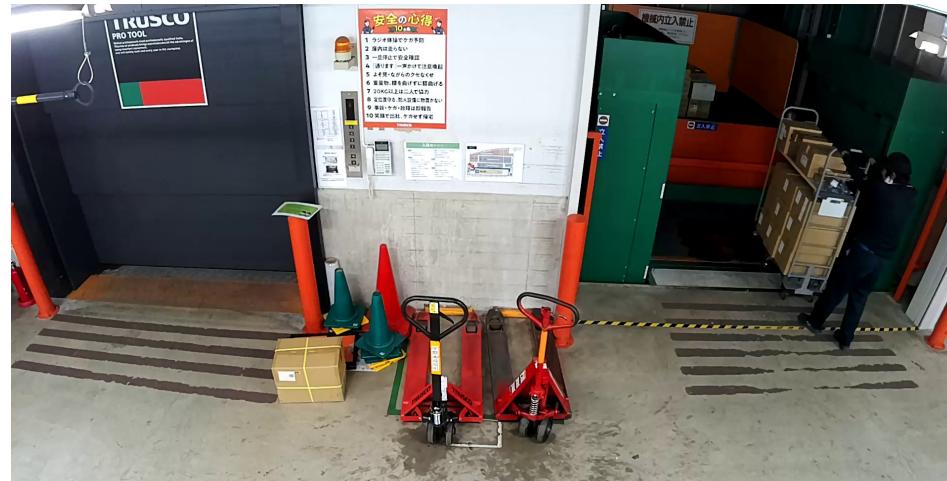
mostly consists
of manual labor.

x10

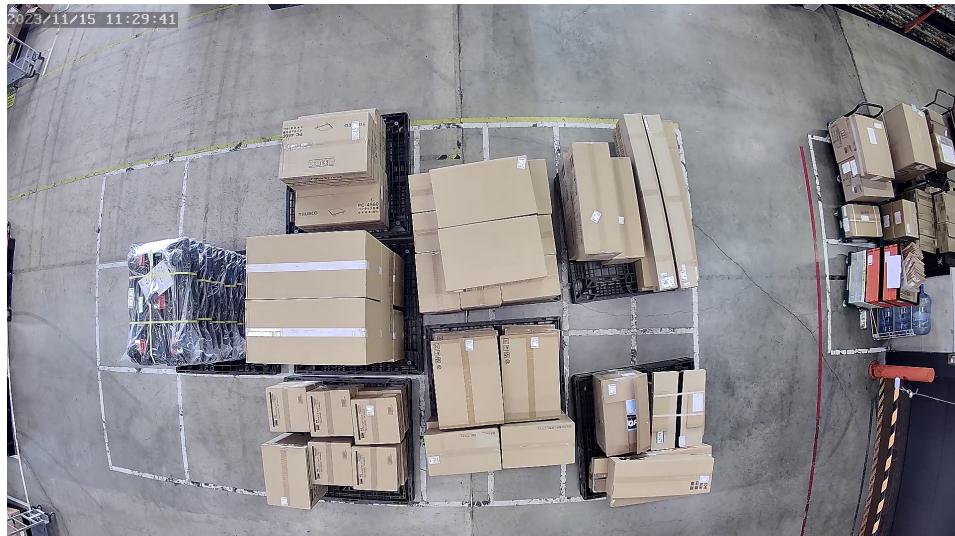
Different Cameras, Different Recognition Models



Recieving



Lift transport



Inspection / Put-away (1F)



Put-away (2F-5F)

Cameras
from
different
target
/ angles.

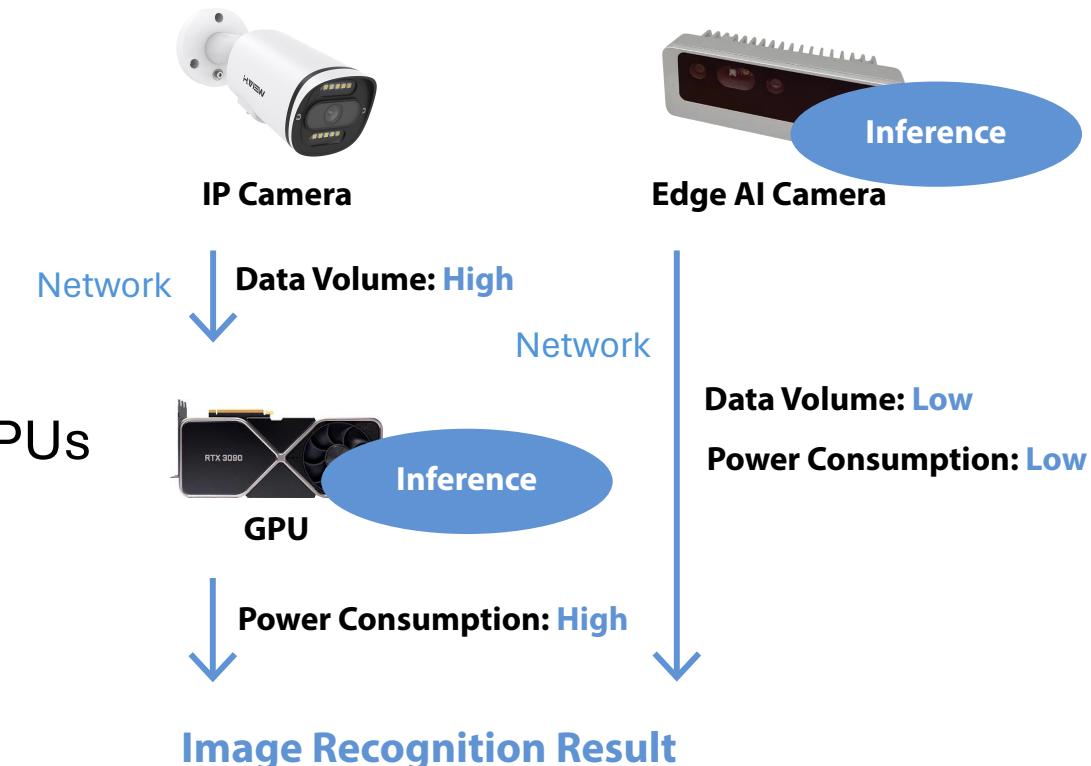
Challenges in Warehouse Digitalization

Digitalization means

recognizing the workplace and extracting insights.

In terms of image recognition tasks, main challenges are

- **Annotation**
 - Manual image annotations are cost intensive
→ Our prior work [1] has worked on semi-auto annotation
- **Data Processing (Fast Processing)**
 - Image recognitions are mainly processed on GPUs
 - Demand for real-time analytics for monitoring in large warehouse environment
- **Energy Efficiency**
 - As warehouses expand in size, maintaining energy efficiency is crucial

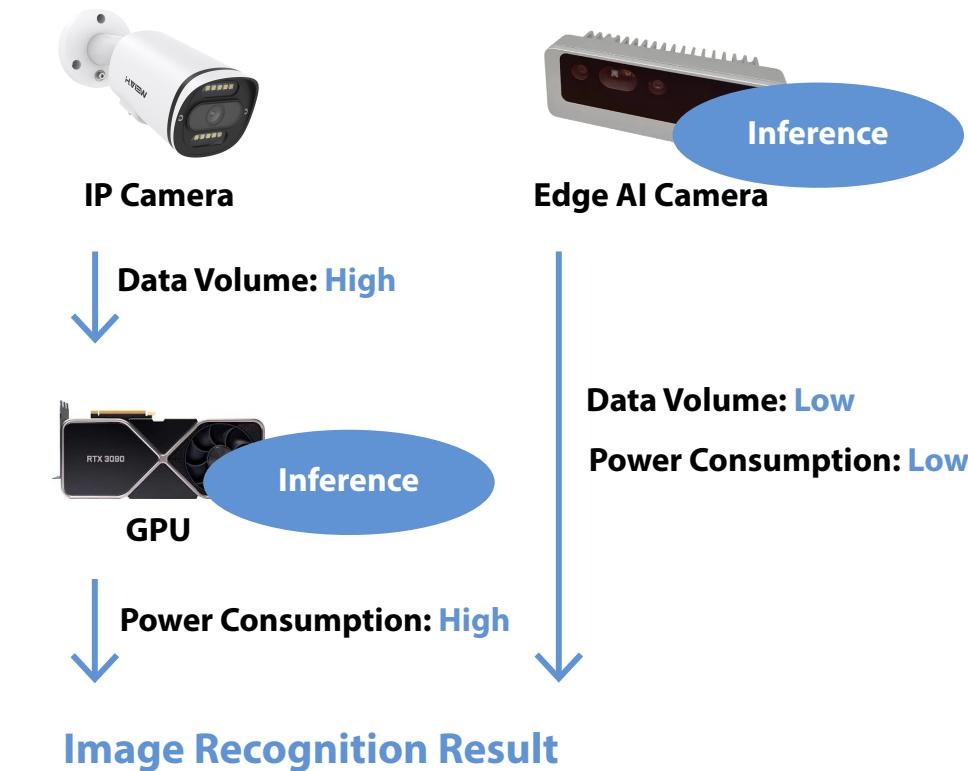


Challenges in Warehouse Digitalization

Digitalization means
recognizing the workplace and extracting insights.

In terms of image recognition tasks, main challenges are

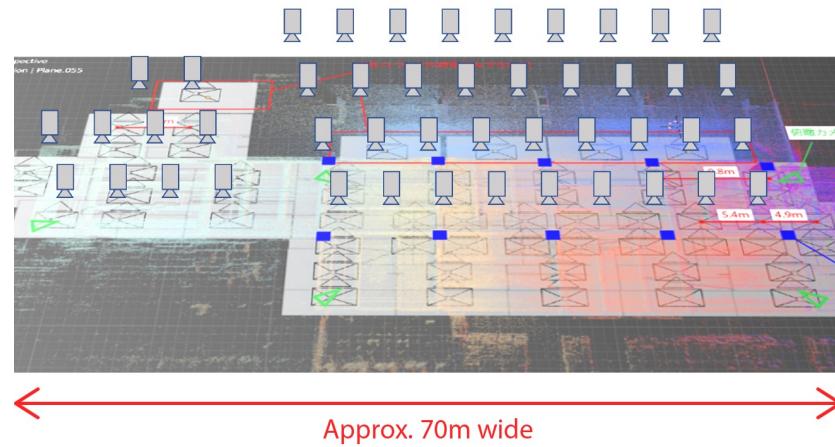
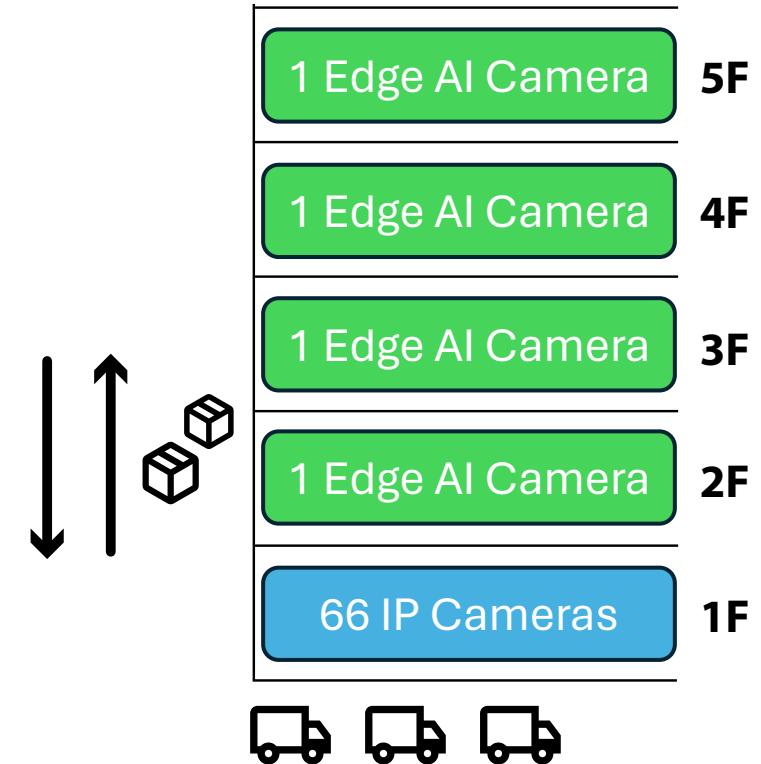
- **Annotation**
 - Manual image annotations are cost intensive
→ Our prior work [1] has worked on semi-auto annotation
- **Data Processing (Fast Processing)**
 - Image recognitions are mainly processed on GPUs
 - Demand for real-time analytics for monitoring in large warehouse environment
- **Energy Efficiency**
 - As warehouses expand in size, maintaining energy efficiency is crucial



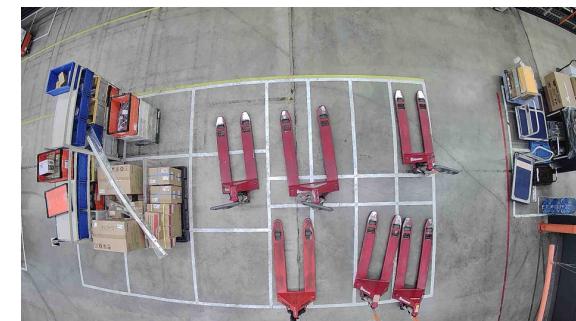
Target Environment

Large-scale camera arrays are deployed.

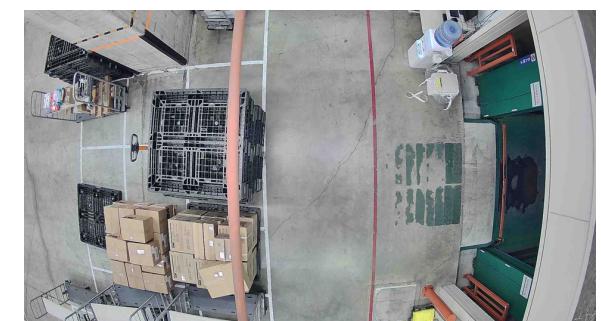
- Our testbed warehouse
 - Total floor space: $23,964 \text{ m}^2$
 - 5-storey (Approx. $70\text{m} \times 50\text{m} \times 5$ floors)
- 1F: 66 IP cameras (5 fps) deployed
 - 330 frames generated per second
- 2F-5F: 1 Edge AI camera deployed in each floor



Camera Arrangement on 1F

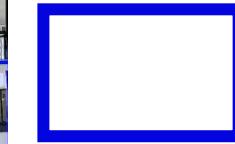


← 7m →



Example Camera Views

Video of Stitched Images in Inbound Area



Captured area of
single camera

Manual labors such as
inspections,
transporting, etc
occur in inbound areas.

x 10

Example of Warehouse Digitalization

Warehouse operations (inbound)

Receiving



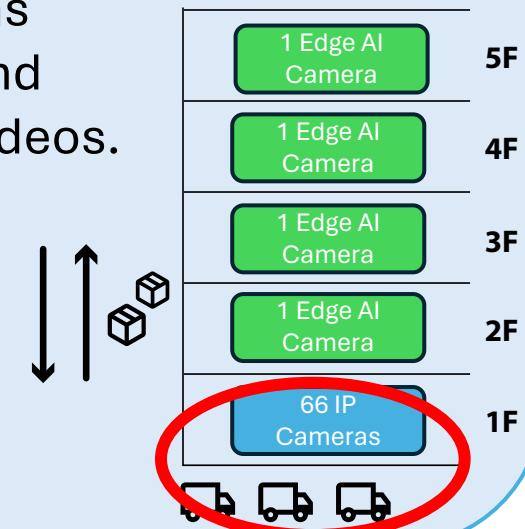
Inspection / Put-away

...



Worker Tracking

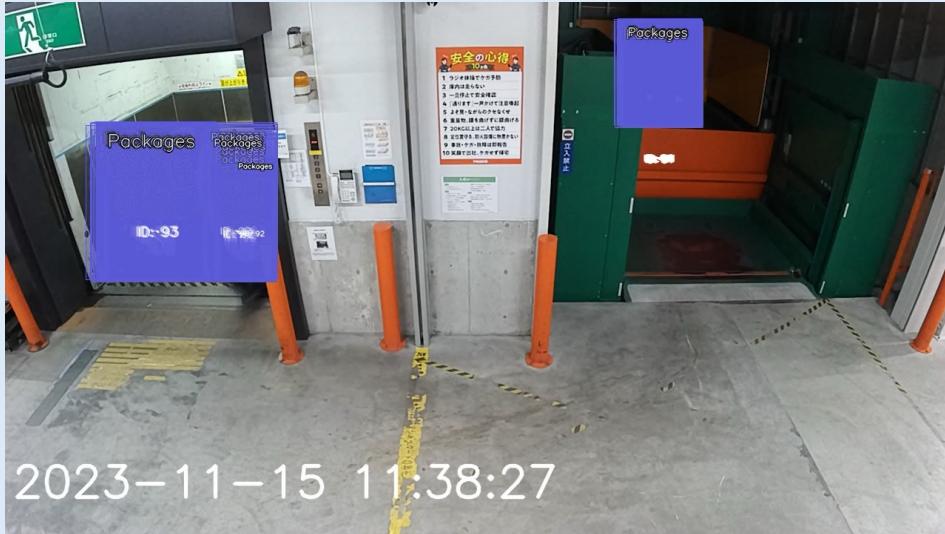
- Multi-camera tracking of workers
- Use recognition results from approx. 40 cameras
- Recognition, tracking and stitching on recorded videos. (post process)



Object Detection / Tracking by Edge AI Camera

Warehouse operations (inbound)

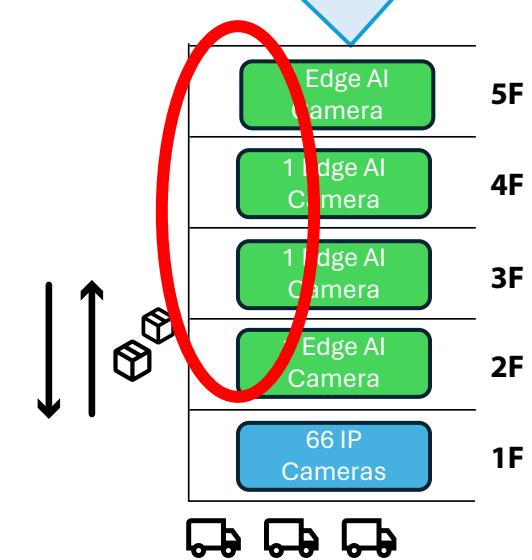
Put-away (2F-5F) → Storage



Pallet Tracking and Counting

- Pallet detection and tracking
- Monitor daily inbound activities on upper floors
- Real-time processing using edge AI camera

Deployed Luxonis AI Cam



Warehouse Evolution: Impact on Energy Efficiency

Assumption

IP Camera - GPU



Power Consumption

7.5W each

IP Cam (5fps)



NVIDIA RTX 3090

AI Camera



450W / 6 cams
(30 frames per second)

7.0W each

Luxonis AI Cam



Our testbed warehouse

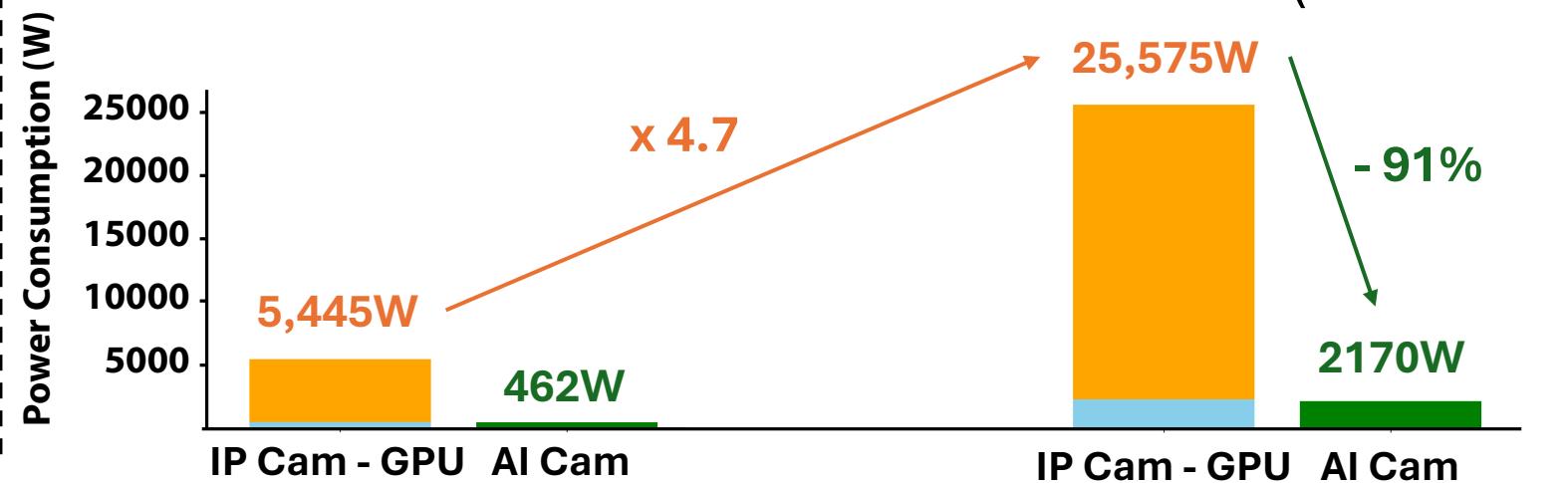
- Category: Large
- Floor space: $23,964 \text{ m}^2$
- Num of cameras (1F):

66

Future target warehouse

- Category: Mega
- Floor space: $89,864 \text{ m}^2$
- Num of cameras (1F):

310 (estimated)



Edge AI Experiments

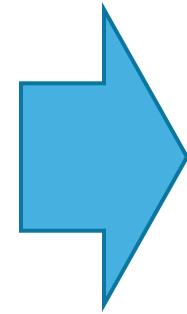
Evaluate performance of YOLO networks on PC, Edge AI Camera and AI accelerator



Recorded Video



Inference



GPU (RTX 3090)
with Ryzen 9 5950X



Luxonis AI Cam
(OAK-D-Pro-W-PoE)

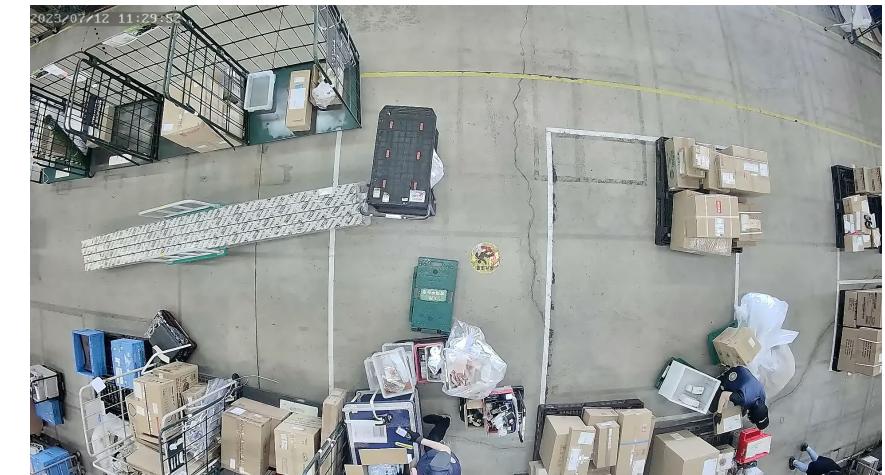


Hailo-8 AI accelerator
with Rock5B

Evaluate



- Processing Speed (fps)
- Power Consumption (W)



- Pre-trained YOLO models
 - v8n, v8s, v8m, v8l and v8x
- Input (H x W x C): 640x640x3
- Targeting 1 class
 - person

Results of Edge AI Experiments

- Hailo setup shows superior computing performance
 - fps: in less complex models (YOLO v8n, v8s and v8m)
 - Power: All models
- Luxonis Edge AI Camera: Power comsumption is steadily low, but its performance is also low

Model	PC		Luxionis OAK		Hailo-8	
	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

Power comsumption per processing frame is reduced > 98% when compared to PC-GPU setup

Low power consumption, but low performance

Learnings from Real-World Digitalization

- Network bandwidth:
 - More cameras, more packet loss. → Frame discrepancies of recorded videos
 - Switch RTSP transport (video stream) from UDP to TCP. → More packet loss
 - Increased latency
- Storage:
 - Video files of 66 IP cameras requires 1.2TB of storage daily.

Future work

- Reduce data transmission
 - Sending object recognition or segmentation results processed by edge devices for post processes

Our preliminary work suggests over 57% data reduction.

Conclusion

- Explored the use of edge AI technologies within warehouse operations
- Addressed the challenges posed by transition to mega warehouse.
- Proposed the groundwork for further advancements in warehouse digitalization with edge AI technologies.

Future Works

- We'd like to propose an image processing pipeline for edge AI architecture targeting warehouse environments.
- We'd like to deal with the privacy concerns related to image processing.
 - e.g. Data processing levels on edge devices.
 - Only necessary data for specific task should be processed and sent to the central server.

Thank you for listening!