Active Robot Vision with Perception-Aware Navigation for Worker Monitoring

Chaewon Yang
M.S. Student
Dept. of Architecture & Architectural Engineering
Seoul National University
chaewony@snu.ac.kr

2025. 02. 24.

Contents

Research Background

▪ Computer Vision Applications for Safety Monitoring

- Vision sensors are used for verifying personal protective equipment (PPE) compliance through object detection models on construction sites.
- As systems like autonomous robots rely on high-confidence to make informed decisions and take appropriate actions (Selvaraj et al., 2020).

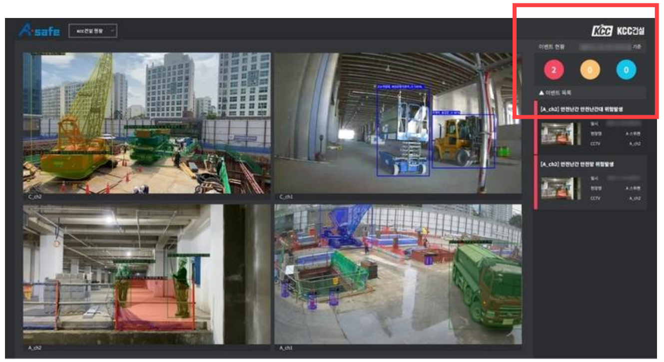
Detection Results
(Worker A)

Safety_Helmet_used 0.9

Safety_Harness_used 0.8

Person 0.9

Safety_Shoe_used 0.6



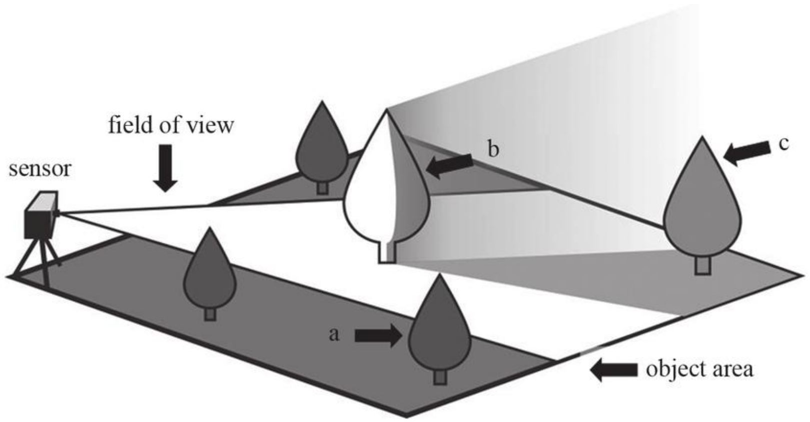
Warning System

- Worker A is properly wearing PPE

Research Background

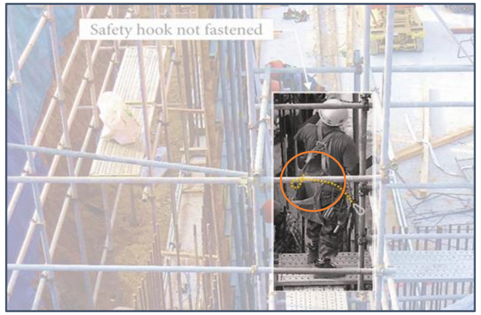
▪ Computer Vision Applications for Safety Monitoring

- Fixed-viewpoint monitoring systems struggle with occlusion and limited Field of View (FoV), which hinder accurate detection and classification(Paneru & Jeelani, 2021).
- These limitations, influenced by object angle, position, and environmental factors, reduce detection reliability and system performance.

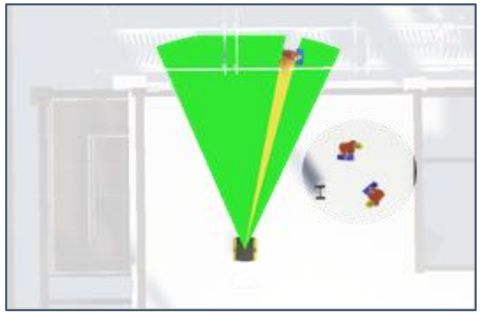


✓ Occlusion

[Occlusion by Obstacle]



[Self-Occlusion]



[View Occlusion]

✓ Types of Occlusion Problems

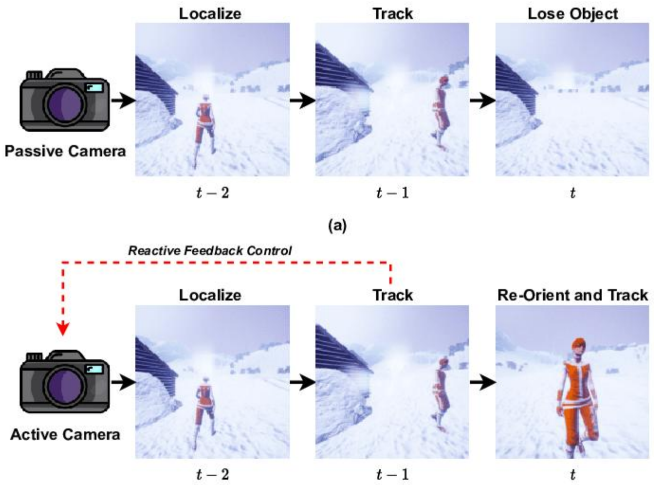
Research Background

▪ Active Vision: Related work

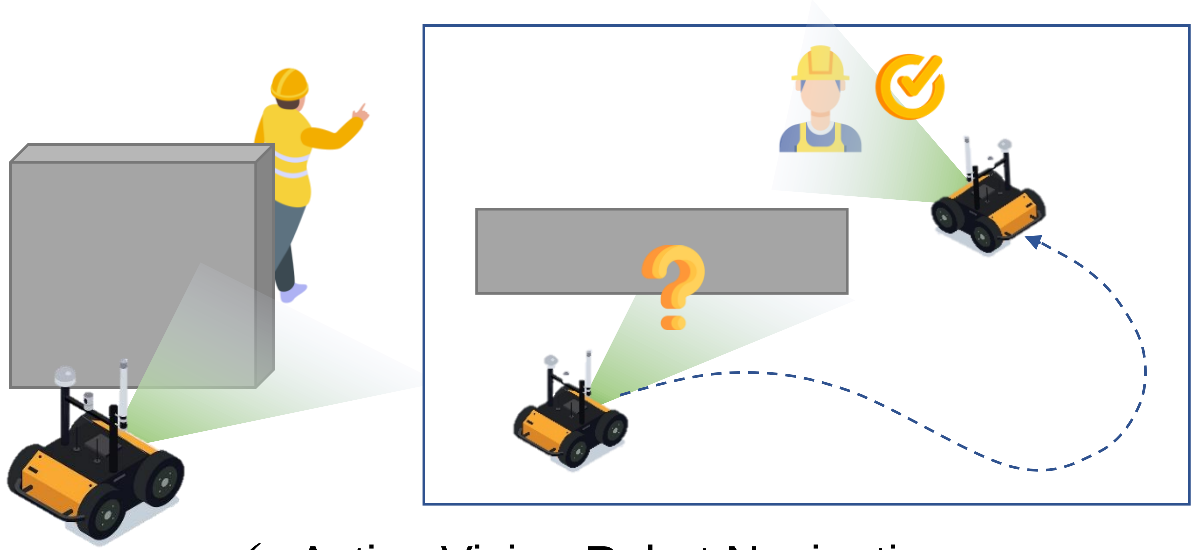
- Mobile robots offer a solution through dynamic repositioning called ‘Active Vision’ (Zhang et al., 2022).

    1) The physical movement of the camera to dynamically collect more information about the environment, allowing it to overcome the limitations of FoV without additional visual devices.

    2) It should be learned to move actively while continuously checking the surrounding environment according to the robot's own goal.



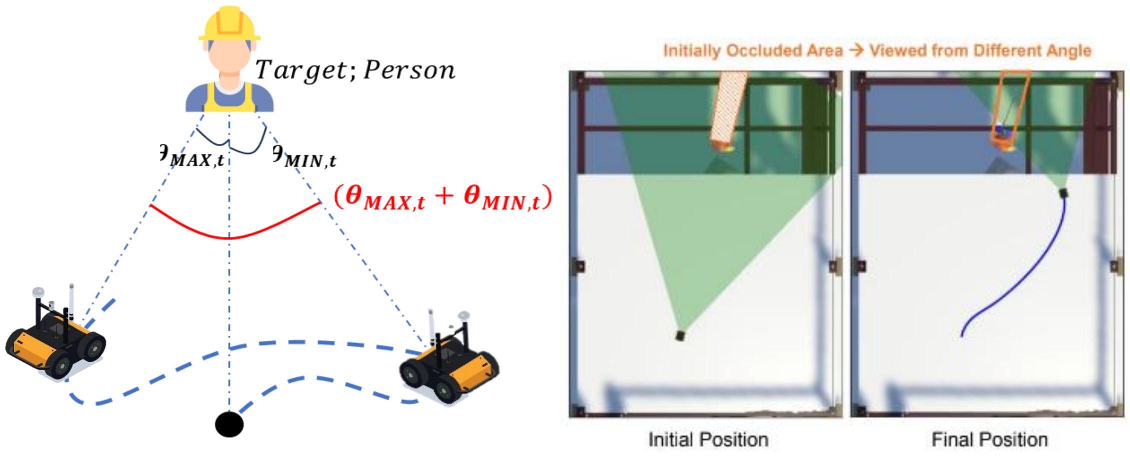
✓ Concept of Active Vision



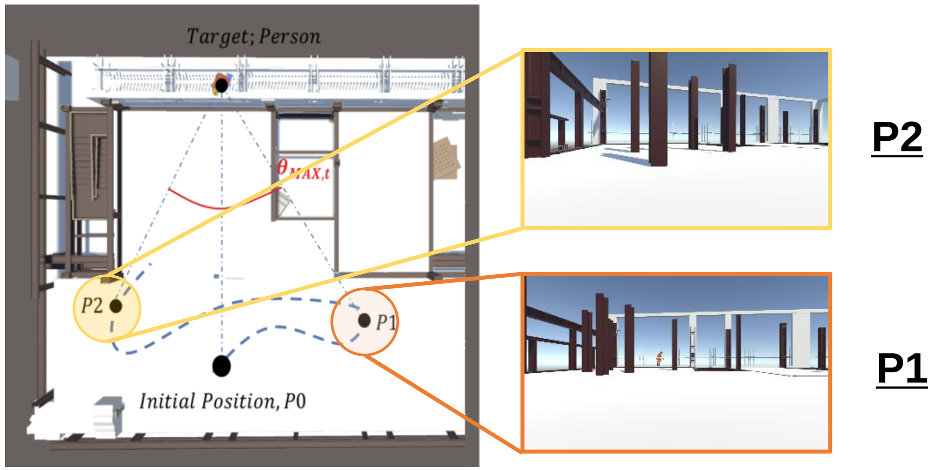
✓ Active Vision Robot Navigation

Research Background

- **Active Vision for Occlusion-Aware Robot Navigation**
  - Previous study has shown that reinforcement learning methods can enhance the proactive capabilities of autonomous monitoring systems in obstacle-rich environment. (Park et al. 2024)
  - Limitations
    1) It is difficult to detect when the distance is too far depending on the performance of the model.
    2) Additional occlusion also occurs in the moved location.



[J Park et al. 2024]



[Limitations]

Research Background

▪ Human-Robot Proxemics.

- The presence of robots introduces potential risks, including physical collisions and perceived threats, which must be mitigated to ensure both physical and psychological safety (Kim et al., 2020)
- Monitoring tasks such as safety management were classified as distant human-robot interaction distances, and were recommended to maintain a certain distance (Sun et al., 2023)

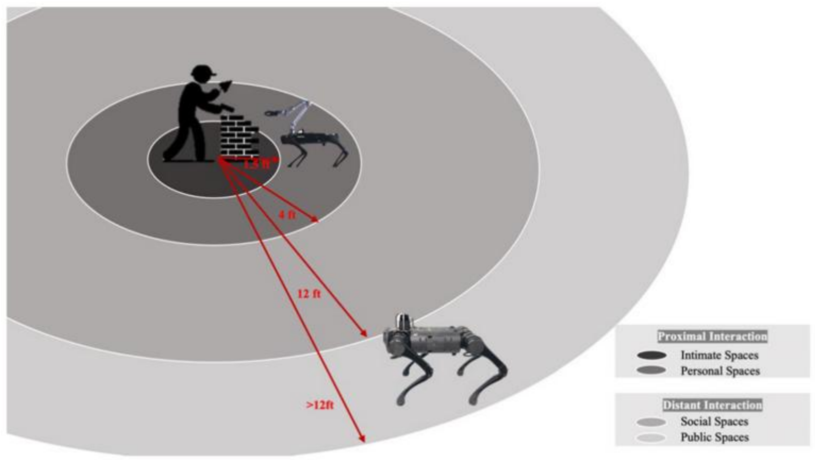


Fig. 1. Proximal and distant interaction spaces between human and robot.

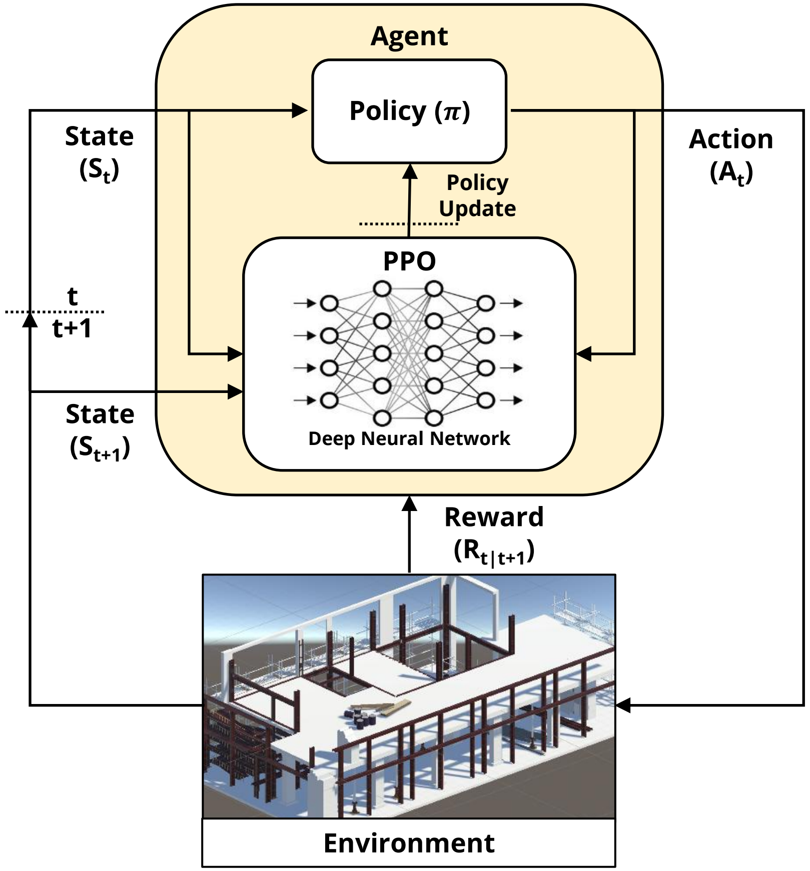
▪ Research Objectives

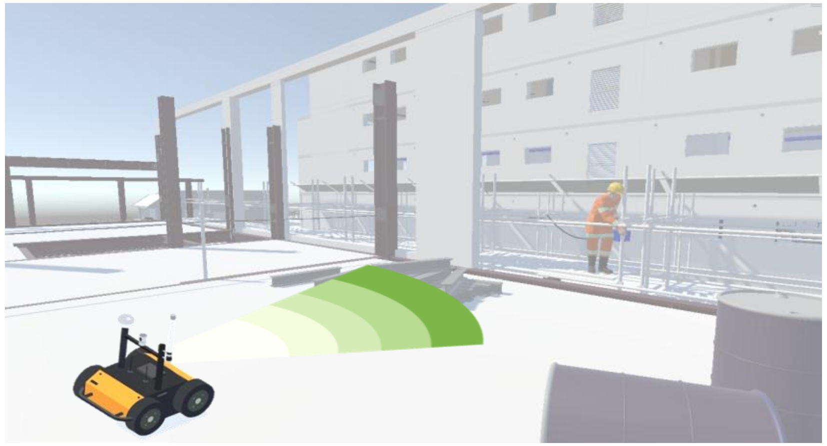
- The robot navigates to achieve the most reliable occlusion-aware monitoring results within its field of view while maintaining a reasonable distance from the worker.

▪ Research Framework

- It focuses on the detection of PPE compliance that may cause self-occlusion during safety monitoring.
- The path planning, in which the robot finds the best viewpoint to locate the target within its detection field, was learned through reinforcement learning DRL.

✓ Reinforcement learning framework





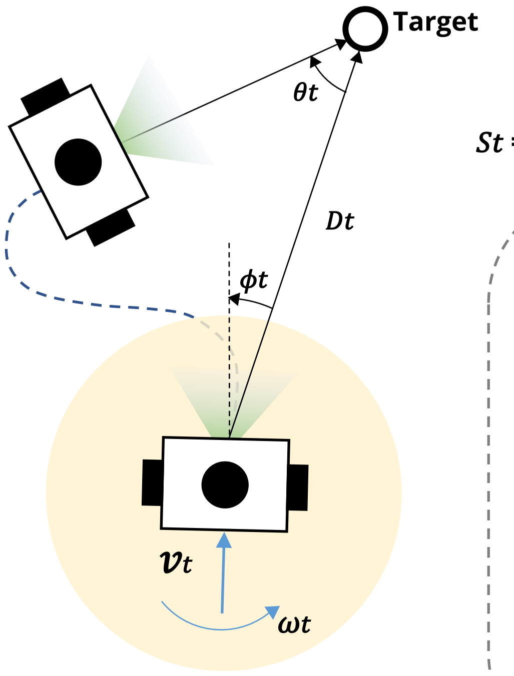
✓ Detection Field

✓ PPE

Methodology

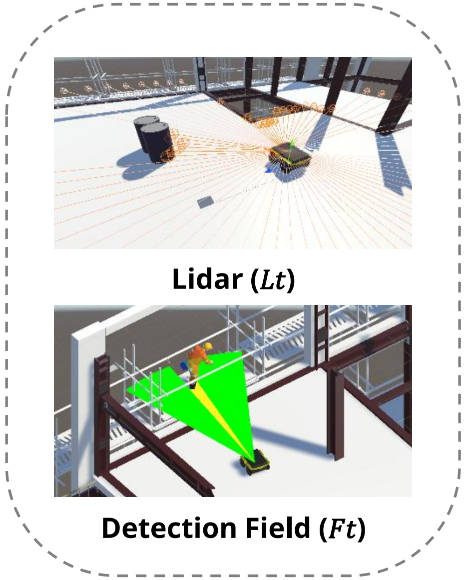
▪ Design of State – Action Space, Reward

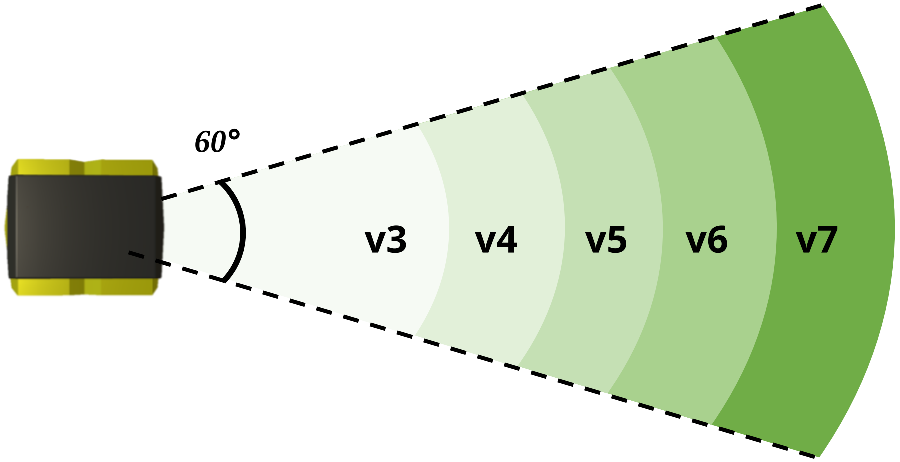
- Develop a dynamic construction site with moving worker in virtual environment.



Action
$At = (v_t, \omega t)$

State
$St = (Dt, \theta t, \phi t, Lt, Ft)$



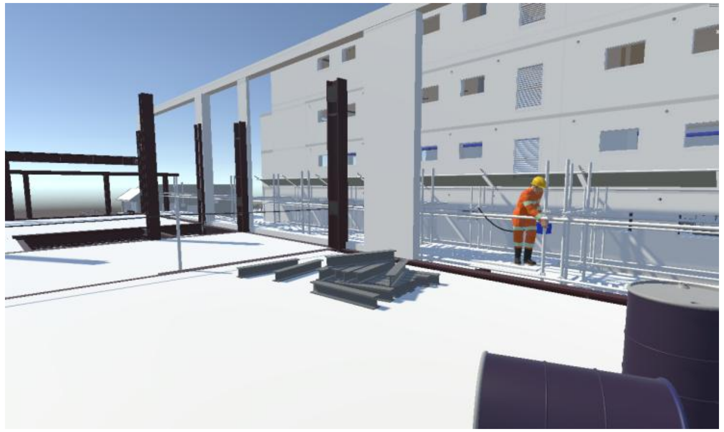


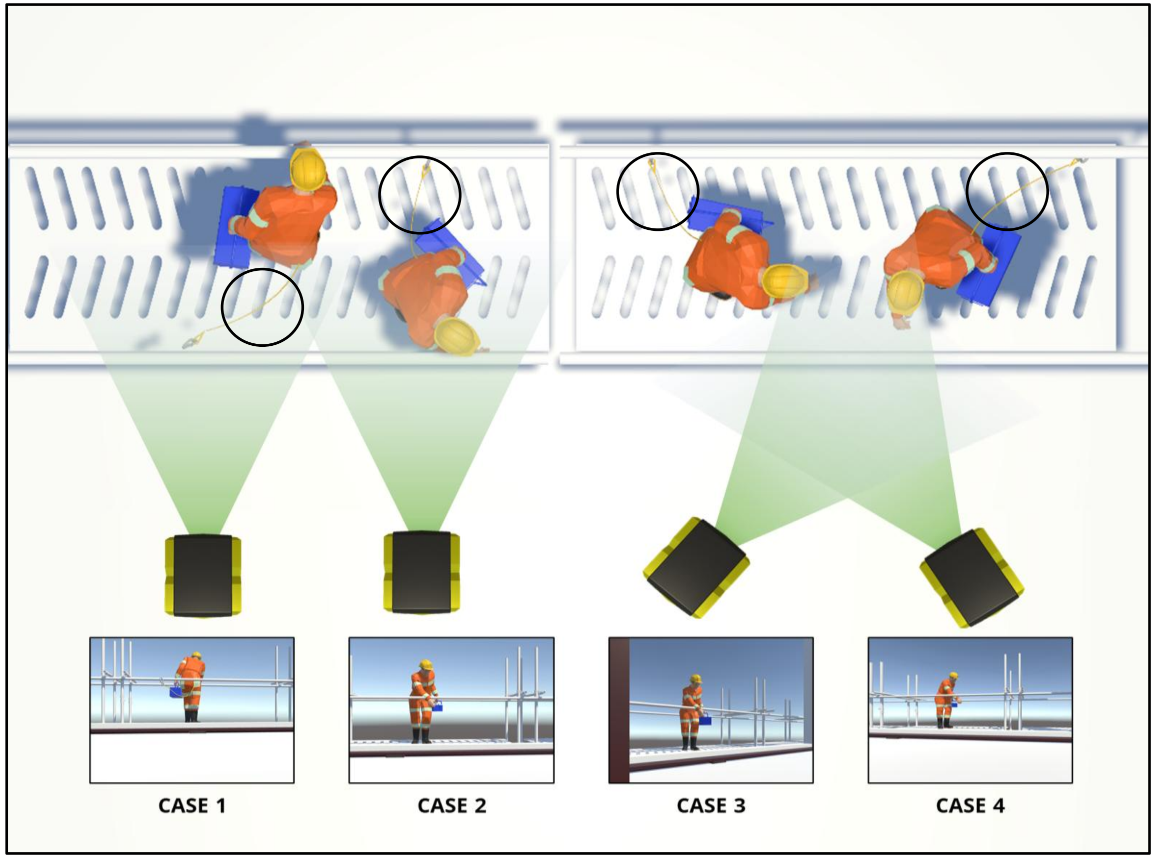
$$r_t(s_t, a_t) = \begin{cases} r_{detected} & \text{if target enters the detection field,} \\ r_{collision} & \text{if collision occurs,} \\ r_{time} \cdot \Delta t & \text{otherwise.} \end{cases}$$

Experiments

▪ Experimental Setup

- Develop a dynamic construction site with moving worker in virtual environment.





Experiments

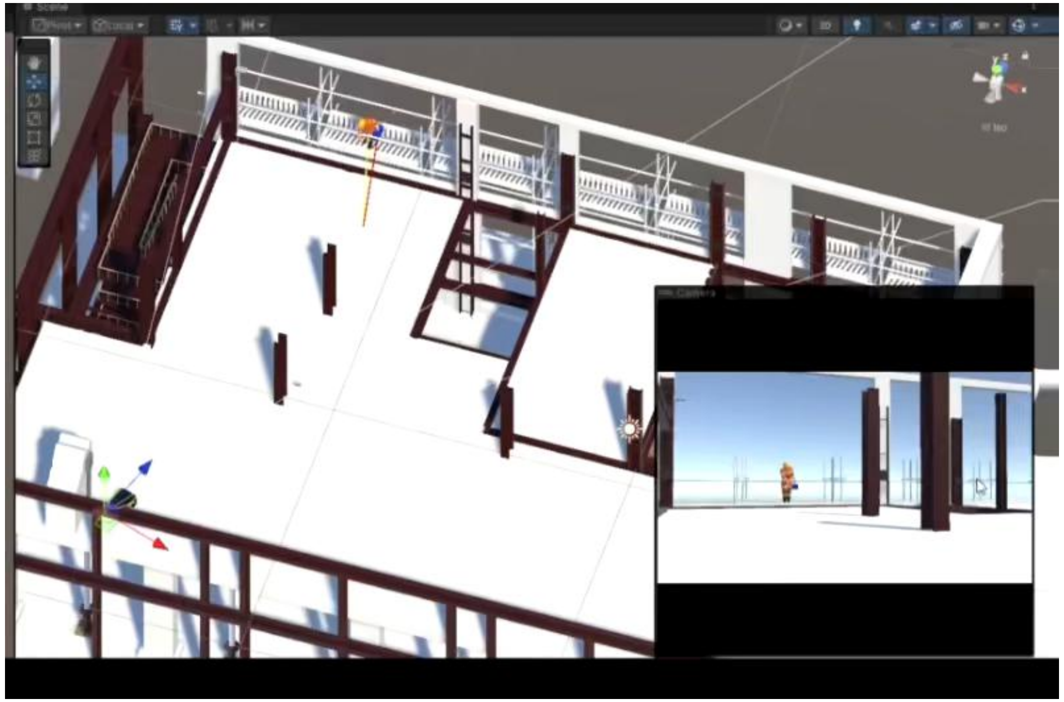
▪ Evaluation Metrics

- Success Rate (SR): the proportion of episodes in which object detection exceeds a predefined confidence threshold 0.8.

$$SR = \frac{1}{N}\sum_{i=1}^{N} S_i$$

Success Rate

- Trajectory Length (TL)
- Minimum Distance (D_{min}): the target along the path to assess the robot's proximity to workers



✓ Sample video of Experiment

Results

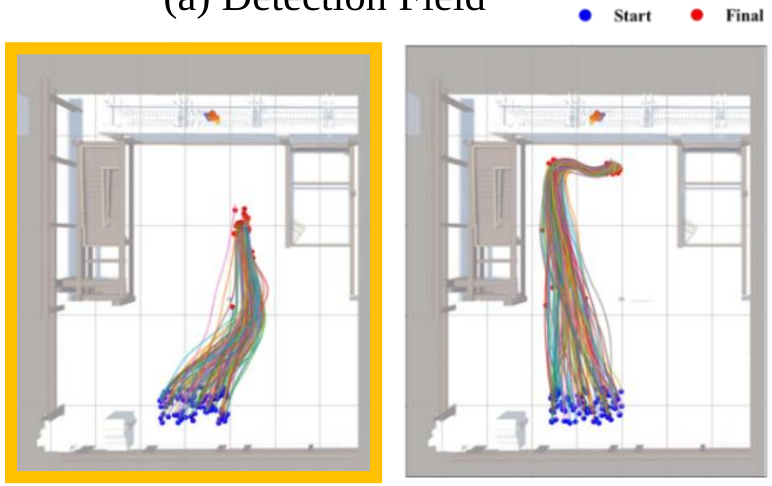
▪ Experiment Results

- The results of this study highlight the significant impact of detection field configurations on the robot's performance.

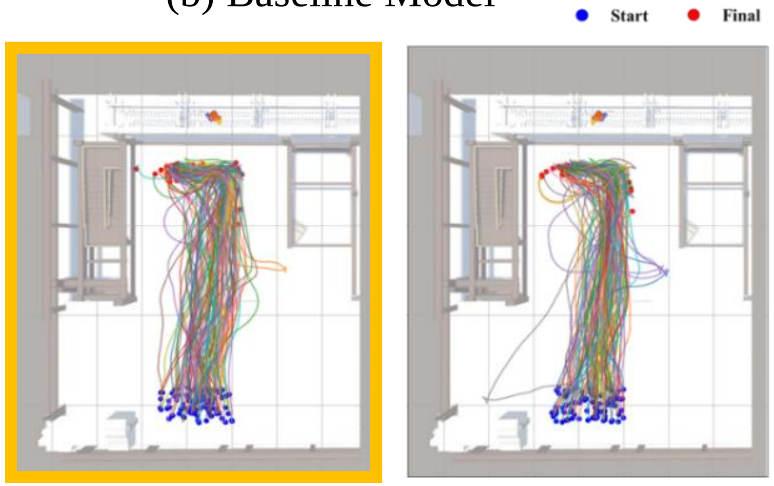
	Experimental Setups								
	w/ occlusion (n=300)			wo/occlusion (n=100)			Total Scenario		
Method	**SR**	**TL**	D_{min}	**SR**	**TL**	D_{min}	**SR**	**TL**	D_{min}
Baseline	0.42	39.53	2.25	0.92	40.13	2.10	0.55	39.68	2.22
Detection Field v3	0.80	35.68	1.52	0.98	28.91	2.65	0.85	33.99	1.80
Detection Field v4	0.77	36.00	2.98	0.95	32.01	4.22	0.82	35.00	3.29
Detection Field v5	0.78	40.00	3.5	0.96	24.45	4.99	0.83	36.11	3.68
Detection Field v6	0.59	37.12	3.62	0.91	26.86	5.20	0.67	34.56	4.02
Detection Field v7	0.56	40.90	4.95	0.89	25.46	5.38	0.65	36.94	5.06

- This highlights the trade-off between maintaining worker safety by increasing distance and ensuring high detection reliability.

(a) Detection Field



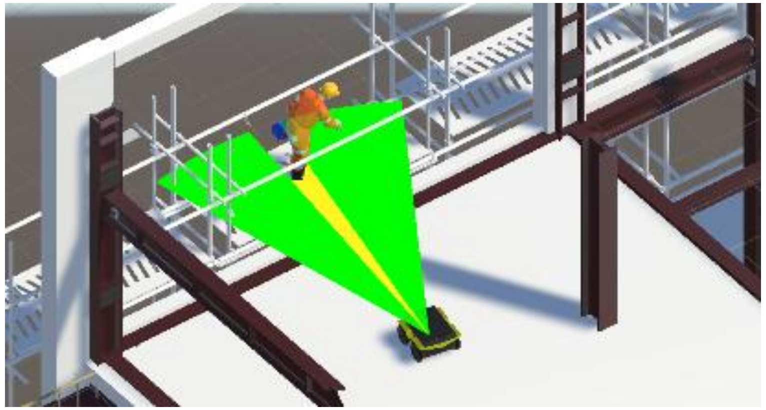
(b) Baseline Model



Conclusion

▪ Impact of Detection Field

- The robot's detection range significantly influences its navigation policy. Expanding the range allows for non-intrusive monitoring of workers, but long-distance detection reduces reliability.
- In disocclusions, detection is possible without additional movement. In occluded conditions, the robot must move closer to the target, reducing the minimum detection distance.



▪ Further Research

- Further research should focus on multi-worker monitoring and addressing interference between robots.
- Test the system in real construction environments to verify its performance under diverse conditions.

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. https://doi.org/10.1109/msp.2017.2743240

Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., & Leonard, J. J. (2016). Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age. *IEEE Transactions on Robotics*, *32*(6), 1309-1332. https://doi.org/10.1109/tro.2016.2624754

Chen, H., Hou, L., Zhang, G., & Wu, S. (2023). Using Context-Guided data Augmentation, lightweight CNN, and proximity detection techniques to improve site safety monitoring under occlusion conditions. *Safety Science*, *158*. https://doi.org/10.1016/j.ssci.2022.105958

Hall, E. T. (1966). The hidden dimension. *Garden City*.

Kim, D., Lee, S., & Kamat, V. R. (2020). Proximity Prediction of Mobile Objects to Prevent Contact-Driven Accidents in Co-Robotic Construction. *Journal of Computing in Civil Engineering*, *34*(4). https://doi.org/10.1061/(asce)cp.1943-5487.0000899

Kim, Y., Kim, S., Chen, Y., Yang, H., Kim, S., Ha, S., Gombolay, M., Ahn, Y., & Cho, Y. K. (2024). Understanding human-robot proxemic norms in construction: How do humans navigate around robots? *Automation in Construction*, *164*. https://doi.org/10.1016/j.autcon.2024.105455

Li, Z., & Li, D. (2022). Action recognition of construction workers under occlusion. *Journal of Building Engineering*, *45*. https://doi.org/10.1016/j.jobe.2021.103352

Luo, H., Liu, J., Fang, W., Love, P. E. D., Yu, Q., & Lu, Z. (2020). Real-time smart video surveillance to manage safety: A case study of a transport mega-project. *Advanced Engineering Informatics*, *45*. https://doi.org/10.1016/j.aei.2020.101100

Mammadov, M. (2023). End-to-end Lidar-Driven Reinforcement Learning for Autonomous Racing. *arXiv preprint arXiv:2309.00296*. https://doi.org/10.48550/arXiv.2309.00296

Nath, N. D., Behzadan, A. H., & Paal, S. G. (2020). Deep learning for site safety: Real-time detection of personal protective equipment. *Automation in Construction*, *112*. https://doi.org/10.1016/j.autcon.2020.103085

Pan, L., Li, A., Ma, J., & Ji, J. (2021). Learning Navigation Policies for Mobile Robots in Deep Reinforcement Learning with Random Network Distillation. *2021 the 5th International Conference on Innovation in Artificial Intelligence*, 151-157. https://doi.org/10.1145/3461353.3461365

Paneru, S., & Jeelani, I. (2021). Computer vision applications in construction: Current state, opportunities & challenges. *Automation in Construction*, *132*. https://doi.org/10.1016/j.autcon.2021.103940

References

Park, J., Kim, S., Park, M., & Ahn, C. R. A. (2024). Occlusion-Aware Object Detection for Worker Monitoring Using a Reinforcement Learning Approach with a Mobile Robot. *Proceedings of the 2023 ASCE International Conference on Computing in Civil Engineering (i3CE)*.

Samarakoon, S. M. B. P., Muthugala, M. A. V. J., & Jayasekara, A. G. B. P. (2022). A Review on Human–Robot Proxemics. *Electronics*, *11*(16). https://doi.org/10.3390/electronics11162490

Seits, F., Kurmi, I., Nathan, R. J. A. A., Ortner, R., & Bimber, O. (2022). On the Role of Field of View for Occlusion Removal with Airborne Optical Sectioning. *arXiv preprint arXiv:2204.13371*. https://doi.org/10.48550/arXiv.2204.13371

Selvaraj, N. M., Muhammad, I., & Cheah, C. C. (2020). Towards dependable object detection. *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, 523-528. https://doi.org/10.1109/IECON43393.2020.9255076

Shang, J., & Ryoo, M. S. (2024). Active vision reinforcement learning under limited visual observability. *Advances in Neural Information Processing Systems*, *36*.

Sun, Y., Jeelani, I., & Gheisari, M. (2023). Safe human-robot collaboration in construction: A conceptual perspective. *J Safety Res*, *86*, 39-51. https://doi.org/10.1016/j.jsr.2023.06.006

Taheri, H., Hosseini, S. R., & Nekoui, M. A. (2024). Deep reinforcement learning with enhanced ppo for safe mobile robot navigation. *arXiv preprint arXiv:2405.16266*. https://doi.org/10.48550/arXiv.2405.16266

Wang, H., Wang, W., Shu, T., Liang, W., & Shen, J. (2020). Active visual information gathering for vision-language navigation. *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXII 16*, 307-322. https://doi.org/10.48550/arXiv.2007.08037

Kim, M. H., Shin, S. W., & Suh, Y. Y. (2019). Application of Deep Learning Algorithm for Detecting Construction Workers Wearing Safety Helmet Using Computer Vision. Journal of the Korean Society of Safety, 34(6), 29–37. https://doi.org/10.14346/JKOSOS.2019.34.6.29

Thank you!

Chaewon Yang
M.S. Student
Dept. of Architecture & Architectural Engineering
Seoul National University
chaewony@snu.ac.kr