

Vision-Based High-Torque Bolt Fastening for Nuclear Power Plant Maintenance

Hyeokbeom Kwon^{a,b}, Ki Hong Im^b, Jinyi Lee^b, Jongwon Park^{a,b*}

^aUniversity of Science & Technology, Gajeong-ro 217, Yuseong-gu, Daejeon, Republic of Korea

^bKorea Atomic Energy Research Institute, Daedeok-daero 989-111, Yuseong-gu, Daejeon, 04535, Republic of Korea

*Corresponding Author: jwpark@kaeri.re.kr

***Keywords:** Robot, Fastening, Hydraulic Manipulator, Visual Servoing

1. Introduction

Bolt fastening constitutes a critical process across a wide range of industrial domains, including construction, plant installation, and nuclear power plant maintenance. Despite advances in industrial automation, bolt fastening in unstructured environments remains heavily dependent on manual labor, resulting in inefficiencies, elevated manpower requirements, and heightened safety risks. Prior research on robotic automation has largely concentrated on structured settings such as automotive and electronics assembly lines. [1] While these approaches have demonstrated high levels of precision under controlled conditions, they exhibit limited adaptability to dynamic and unstructured field environments.

To overcome these challenges, the present study introduces a vision-based robotic system capable of executing high-torque bolt fastening in unstructured industrial contexts. The system integrates advanced perception algorithms with robust mechanical components and an autonomous fastening strategy to enhance operational efficiency, consistency, and safety.

2. Method

2.1 System Configuration

The robotic platform employed in this study is the ARMstrong Dex dual-arm mobile manipulator developed by KAERI [2]. Each arm has six degrees of freedom, with a two-degree-of-freedom gripper. Bolt fastening is performed using a commercial DeWALT DCF892 impact wrench, capable of delivering 812 Nm of fastening torque and 1084 Nm of loosening torque. The wrench is interfaced with the robot via a custom compliance jig, which mitigates alignment errors and ensures stable fastening operations.

The perception system is realized through a dual-camera configuration. A head-mounted RGB-D camera (Intel RealSense D455) provides a global structural overview and facilitates the detection of approximate bolt locations. A wrist-mounted wide-angle RGB camera (Arducam B0200) captures detailed close-range imagery, thereby supporting refined pose estimation and precise end-effector alignment.

2.2 Perception and Pose Estimation

Bolt detection is achieved through the YOLOv12n [3] Turbo model, trained on a structural bolt dataset and yielding a detection accuracy of $mAP@50 = 0.91$. Three-dimensional pose estimation is accomplished by combining depth data with DBSCAN clustering and singular value decomposition (SVD) to derive surface normals and robust pose information. BoT-SORT [4] is employed to maintain consistent bolt tracking across image sequences.

2.3 Alignment and Fastening Strategy

The fastening procedure is structured into three sequential stages. In the coarse alignment stage, the head-mounted camera defines an initial end-effector pose using inverse kinematics. Subsequently, fine alignment is conducted with the wrist-mounted wide-angle camera, employing SAM2[5]-based segmentation and real-time distance estimation for enhanced precision. Finally, insertion and fastening are performed via a spiral self-aligning insertion strategy, wherein the impact wrench rotates during insertion to exploit mechanical compliance for error compensation.

To enhance efficiency further, the system incorporates a sequential fastening capability whereby bolts are fastened consecutively following an automatically generated zigzag trajectory from left to right, enabling systematic and uninterrupted multi-bolt operations.

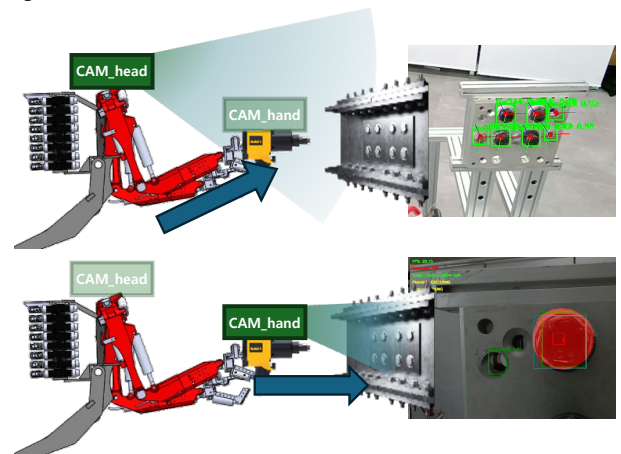


Fig. 1 Concept of vision-based bolt fastening with dual-camera setup

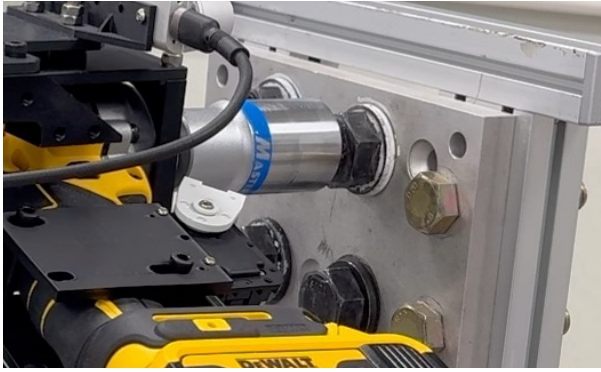


Fig. 2 Close up view of robotic bolt fastening process

2.4 Results

Experimental validation was conducted using M22 hexagonal head bolts positioned at varying orientations. Evaluation metrics included alignment accuracy, fastening success rate, and overall task cycle time.

- Success Rate: 83% (15 successful trials out of 18)
- Average Cycle Time: 15 seconds per bolt

The results substantiate the feasibility of performing stable and efficient fastening exclusively through vision-based perception, without the need for force/torque sensors.

3. Conclusions

This study presents a vision-guided robotic system for high-torque bolt fastening in unstructured industrial environments. By integrating YOLOv12n-Turbo-based detection, SVD-based pose estimation, and SAM2-assisted alignment with a compliance jig, the system achieves reliable fastening performance using a cost-effective sensor configuration.

The main contributions of this work are the demonstration of robust fastening using only vision input, the integration of a wide-angle wrist-mounted camera for improved alignment, and the development of a zigzag-pattern sequential fastening strategy for efficient multi-bolt operations. These outcomes highlight the practical applicability and economic viability of the system in real-world settings.

Future research will focus on broadening system versatility by accommodating diverse bolt specifications, implementing automatic success detection and retry mechanisms, integrating a socket exchange module, and coupling the manipulator with mobile platforms to extend operational coverage.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(Ministry of Science and ICT)(No. RS-2022-00144468).

REFERENCES

- [1] Z. Jia, A. Bhatia, R. M. Aronson, D. Bourne, and M. T. Mason, "A Survey of Automated Threaded Fastening," *IEEE Trans. Automat. Sci. Eng.*, vol. 16, no. 1, pp. 298–310, Jan. 2019, doi: 10.1109/TASE.2018.2835382.
- [2] Park, Jongwon, Jinyi Lee, and Ki Hong Im. "Development and Field Test of a Nuclear Disaster Response Robot, Armstrong." *Transactions of the Korean Nuclear Society Spring Meeting Jeju*, n.d.
- [3] Ravi, Nikhila, Valentin Gabeur, Yuan-Ting Hu, et al. "SAM 2: Segment Anything in Images and Videos." arXiv:2408.00714. Preprint, arXiv, October 28, 2024. <https://doi.org/10.48550/arXiv.2408.00714>.
- [4] Tian, Yunjie, Qixiang Ye, and David Doermann. "YOLOv12: Attention-Centric Real-Time Object Detectors." arXiv:2502.12524. Preprint, arXiv, February 18, 2025. <https://doi.org/10.48550/arXiv.2502.12524>.
- [5] Aharon, Nir, Roy Orfaig, and Ben-Zion Bobrovsky. "BoT-SORT: Robust Associations Multi-Pedestrian Tracking." arXiv:2206.14651. Preprint, arXiv, July 7, 2022. <https://doi.org/10.48550/arXiv.2206.14651>.