

CONFERENCE PRE-PRINT

BB SEGMENT GRASPING PIPELINE WITH VARIABLE ADMITTANCE CONTROL FOR EU DEMO REMOTE MAINTENANCE

H. DUROCHER^{†,1}, X. YANG^{†,1}, D. K. SØRENSEN¹, C. BACHMANN², R. MOZZILLO³, G. JANESCHITZ⁴, and X. ZHANG^{*,1}

[†]: These authors contributed equally.

^{*}: Corresponding author, email: xuzh@mpe.au.dk

1: Aarhus University, Aarhus N, Denmark

2: EUROfusion Consortium, Garching bei München, Germany

3: Basilicata University, Potenza, Italy

4: Max-Planck-Institut für Plasmaphysik, Garching bei München, Germany

Abstract

Secure grasping of the breeding blanket (BB) segments using the vertical transporter (BBVT) is required for remote maintenance via the upper port at EU-DEMO. This entails engaging the folding gripper within the countersunk interface hole of the target BB segment. A detailed grasping pipeline is proposed in this work, split into four discrete actions: interface detection, alignment, insertion and engagement. A hierarchical control framework is also proposed, including a low-level variable admittance controller (VAC) and high-level planner with multisensory action evaluation and recovery. VAC achieves robust contact-rich gripper insertion utilizing the geometry of the interface to passively guide the gripper and allows automatic recovery from some collisions. The proposed pipeline and control strategy were tested using down-scaled prototypes of the gripper and interface together with off-the-shelf components including a UR3e manipulator substitute for the BBVT. Experiments with large random pose errors demonstrated the robustness of the interface detector and VAC, which recovered from 29% of collisions and led to zero “NG” insertions and engagement evaluations. However, the high collision rate (60%) indicated that the performance of alignment evaluator must be improved in the future by changes to the sensor layout and integration of additional modalities. The code for this project is available on GitHub¹.

1. INTRODUCTION

A major technological gap which must be addressed for operating EU-DEMO safely and economically is the availability of robust remote maintenance tools and strategies [1]. Unlike previous tokamaks such as JET and ITER, the breeding blankets (BBs) in DEMO are consolidated into large segments for streamlined maintenance [2]. The resulting crescent-shaped segments, weighing up to 125 and 180 tonnes for the inboard and outboard types, respectively, are maintained via a vertical upper port. The 7-degree of freedom robotic arm known as the BB vertical transporter (BBVT) is under development for this remote handling (RH) task [3], for which safe and reliable grasping of the BB segments will be critical. As a result of the limited exposed grasping area within the upper port and the irregular shapes of the segments, the BBVT must withstand significant moments throughout the RH task while manipulating them with sub-centimeter accuracy to avoid unwanted and potentially costly collisions. It follows that the fit between the conical gripper and the matching standardized BB segment interface hole should be as tight as feasible, and thus a very accurate and precise feedback control architecture will be required for carrying out the grasp. At the same time, the projected gamma radiation dose rate to silicon is approximately 100 Gy/h near the blankets during RH operations [4], leading to potential faults or erroneous behaviors of onboard sensors and microcontrollers [5,6], as well as the possible degradation of the joint servos’ trajectory tracking accuracy [7].

Clearly, a pipeline for robust semi-autonomous BB segment grasping is necessary. It should be composed of simple actions with associated self-correcting recovery actions in case of failure. Evaluation of action success should ideally be based on multiple independent sensor readings to avoid relying on any one sensor which may be faulty. Such a control framework is proposed in this paper. To pair with the fault-tolerant high-level task planning, we propose a low-level control policy based on admittance with variable stiffness and damping to allow the gripper insertion to succeed despite the fact the alignment with the interface could be poor. Finally, gripper alignment is enhanced by applying a state-of-the-art AI-based object detector for interface localization, which should perform well even under challenging environmental conditions.

¹ <https://github.com/AU-DK-Robotics/bb-safe/tree/fec>

The following contributions are presented in this work:

- A hierarchical control architecture for BB segment grasping at EU-DEMO, in which high-level self-correcting action planning based on multi-sensory state evaluation is coupled with a low-level variable-admittance control policy for increasing safety and reliability of contact-rich insertion.
- A proposed action pipeline for BBVT-based grasping of EU-DEMO's BB segments.
- Results of experiments conducted with a down-scaled gripper and interface, which demonstrate the high success rate of the proposed pipeline even when the initial (view) and alignment poses are subject to random disturbances which cause frequent collisions.

2. RELEVANT WORKS

2.1. BB segment remote maintenance

Both JET [8] and ITER [9] were designed with predominantly horizontal port layouts, though with separate upper, equatorial and divertor ports at ITER. A vertical port was suggested as an option for INTOR-NET [10] in the 80's. Loving et al. [11] and Coleman et al. [2] proposed the adopted vertical remote maintenance architecture for DEMO, while the associated transporter design able to withstand the large static loads plus seismic activity was developed by Bachmann et al. [3] [4]. The self-locking folding gripper and matching interlock in the BB segment upper surface was introduced by Steinbacher et al. [12]. The BBVT spatial kinematics and detailed static loads were investigated in [13]. Meanwhile, the problem of reliable control has not yet been studied in detail.

2.2. Learning-based object detection

Computer vision, a key enabling technology for robotics, will support accurate alignment of the gripper and interface. Through visual input, robots can detect, inspect, and locate objects [14]. Prior to the rise of deep learning, object detection relied on hand-crafted features such as the Histogram of Oriented Gradients (HOG) descriptor [15], which worked well but was sensitive to lighting and background variations. The advent of Convolutional Neural Networks (CNNs) marked a turning point, bringing major gains in detection accuracy, speed, and robustness [16]. CNN-based two-stage detectors such as R-CNN [17], Faster R-CNN [18], SPPNet [19], and Feature Pyramid Networks [20] combine region proposals with feature extraction and deliver high accuracy, particularly on small objects, although they can be too slow for real-time use. To meet real-time requirements, one-stage detectors like YOLO [21,22], SSD [23], and RetinaNet [24] were developed, achieving much faster inference with only a small trade-off in accuracy. Today, learning-based object detection is reliable even in challenging conditions, but the main barriers to deployment lie in camera calibration, sensor fusion, and system integration [25]. Dataset preparation, frame alignment and connecting vision outputs to robot control remain time-consuming and require expert knowledge, making rapid or large-scale deployments difficult. Current research focuses on automating calibration and simplifying integration to make robotic systems faster to set up and more reliable in production.

2.3. Admittance control for contact-rich tasks

As physical human–robot–environment interaction becomes more common, impedance and admittance control have become standard approaches for ensuring safe collaboration and for executing tasks that go beyond pure position, force, or hybrid control schemes. The goal of these methods is to shape the dynamic relationship between the robot and its environment [26,27]:

$$f_{ext} = M_d(\ddot{x} - \ddot{x}_r) + C_d(\dot{x} - \dot{x}_r) + K_d(x - x_r) = M_d\ddot{x}_e + C_d\dot{x}_e + K_dx_e \quad (1)$$

where x_r is the desired trajectory unaffected by the external force f_{ext} , x is the measured trajectory, and M_d , C_d , K_d represent the virtual mass, damping, and stiffness of the controller. These controllers are naturally robust to model errors and external disturbances, making them suitable for interaction with uncertain and changing environments [28]. As contact-rich manipulation tasks become more demanding, research has moved from using fixed impedance gains to approaches that adapt or schedule the virtual impedance online [29]. For instance, in tasks such as valve turning, the required interaction force changes throughout the motion, which calls for varying impedance across different contact phases. This has motivated the development of variable impedance and admittance control schemes to improve adaptability and maintain safety [30,31]. For BB-segment grasping, such adaptive interaction strategies will be crucial for achieving consistent performance and reliable operation.

3. METHODOLOGY

3.1. Overview of the grasping pipeline

An overview of the proposed BB segment grasping architecture and pipeline is shown in Fig. 1. At the task-planning level, the pipeline is split into four simple actions: detection, alignment, insertion and engagement, with the robot state being evaluated based on multiple independent sensor readings after each step. Specifically, we have focused on using an ultrasonic range finder, a wrist force-torque (F-T) sensor and force sensitive resistors (FSRs) on each arm of the gripper locking mechanism for these evaluations. Whether the evaluation outcome is “okay” or “NG” determines if the pipeline should continue nominally or a recovery action should be taken to re-attempt the action. An object detector based on the industry standard YOLO is trained and used for identification and localization of the gripper interface. For gripper insertion, an admittance control policy is activated which varies the stiffness and damping matrices based on the gripper depth, enabling near-perfect insertion accuracy despite misalignment with the interface and even overcoming minor collisions with the BB segment top surface. A flowchart of the pipeline steps with evaluations and re-planning is shown in Fig. 2(a).

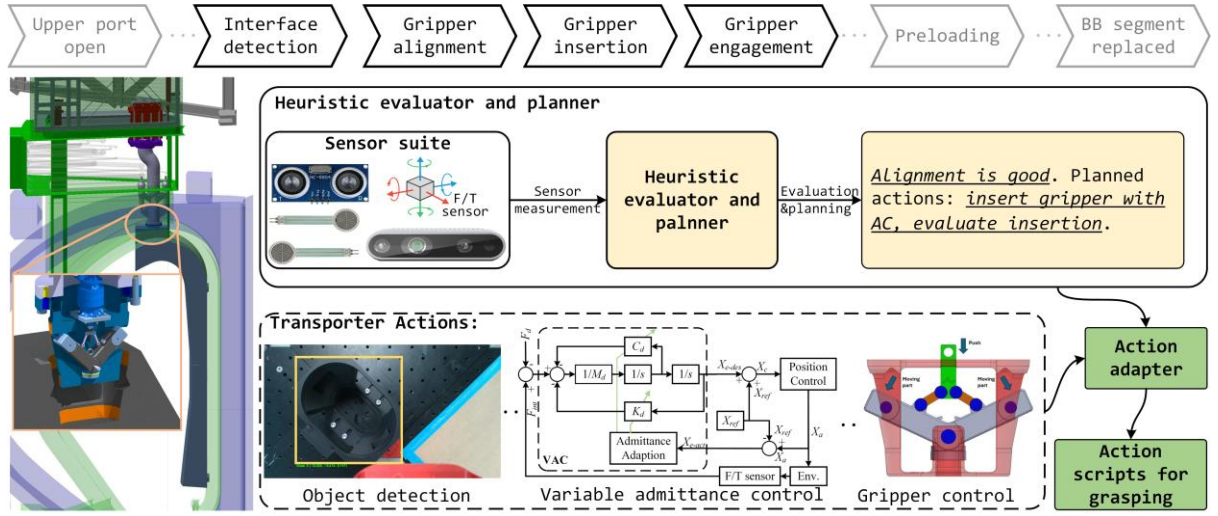


FIG. 1. Overview of the proposed pipeline and associated methods.

3.2. AI-based object detector and eye-in-hand calibration

To increase the accuracy of gripper alignment, the real location of the interface in relation to the gripper should be accurately determined. Localization of the interface within the upper port with environmental conditions unfavorable to CMOS sensors (high gamma radiation and low visibility) must rely on an effective object detection algorithm. Therefore, the state-of-the-art object detector YOLOv8 [22] is adopted in this paper. YOLOv8 is a single-stage, anchor-free detector that directly predicts object classes and bounding box coordinates from feature maps in a single forward pass, enabling real-time performance with high accuracy. Its architecture comprises a CSP-based backbone for feature extraction, a PAN/FPN neck for multi-scale feature fusion and a decoupled head for classification and regression, which together improve convergence and detection precision. To ensure robust performance under varying perspectives and environmental conditions, a diverse and well-labeled training dataset, as shown in Fig. 2(b), is essential. Such a dataset allows the network to learn distinctive object features and remain resilient to noise and disturbances. For this work, two interface geometries with different sizes and features were created to evaluate YOLOv8’s detection performance. A total of 210 images were collected under various orientations, positions, and lighting conditions for the training stage. These images were manually labeled with bounding boxes and class annotations, enabling the detector to learn the features of interface for accurate detection.

The 3D position of the interface obtained from the object detector is expressed in the camera coordinate system, whose origin is located at the optical center of the camera. To enable the robot to manipulate the interface, this position must be transformed into the robot base frame. This requires knowledge of the spatial relationship between the camera and the robot end-effector, a problem known as eye-in-hand calibration.

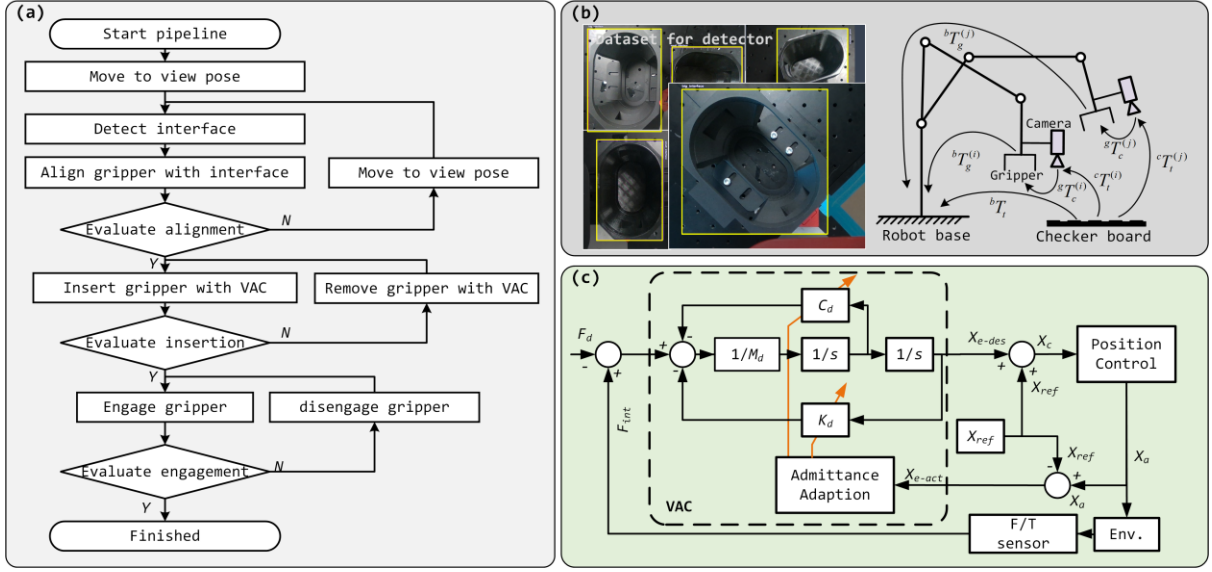


FIG. 2. Technical details: (a) the complete process for the gripper engagement, (b) Dataset for training AI-based object detector and eye-in-hand calibration, and (c) variable admittance control for contact-rich manipulations.

As illustrated in Fig. 2(b), the goal of hand–eye calibration is to determine the homogeneous transformation gT_c from the camera frame c to the gripper (or tool) frame g . Knowing this transformation, along with the robot’s forward kinematics bT_g from the robot base frame b to the gripper frame, allows computation of the camera pose in the base frame ${}^bT_c = {}^bT_g {}^gT_c$. Consequently, any object detected in the camera frame can be mapped into the robot base frame, enabling motion planning and end-effector path generation. To perform the calibration, the camera observes a fixed calibration target (checkerboard) from multiple robot poses. For each pose i , the forward kinematics provide ${}^bT_g^{(i)}$, while the vision system estimates ${}^cT_t^{(i)}$, the pose of the checkerboard t in the camera frame. Because the checkerboard is static in the base frame, the transformation bT_t remains constant across measurements. Using multiple distinct poses, the following relationship holds:

$${}^bT_g^{(i)} {}^gT_c^{(i)} {}^cT_t^{(i)} = {}^bT_t, \quad i = 1, 2, \dots, N \quad (2)$$

Solving this equation for gT_c gives the camera-to-gripper transformation. In practice, over 10 poses are recorded to improve accuracy, and the transformation is estimated using the Tsai–Lenz algorithm, which computes rotation and translation through a least-squares solution over all recorded poses. Once calibrated, the robot can transform any 3D detection from the camera frame into its own base frame, allowing precise motion planning for interface alignment and insertion.

3.3. Variable admittance control for contact-rich manipulation

Variable admittance control (VAC) is adopted to achieve smooth, gap-free insertion of the gripper into the BB segment interface. As illustrated in Fig. 2(c), the VAC module implements a virtual mass–spring–damper system whose input is the desired interaction force F_d and the measured interaction force F_{int} from the wrist-mounted force–torque (F/T) sensor (both specified in the gripper frame). The resulting position offset x_e is computed by solving the admittance dynamics in real time: computed by solving the admittance dynamics in real time:

$$M_d \ddot{x}_e + C_d \dot{x}_e + K_d x_e = F_{int} - F_d \quad (3)$$

where M_d , C_d , K_d are diagonal matrices representing the virtual inertia, damping, and stiffness, respectively. This second-order differential equation is numerically integrated with 4th order Runge-Kutta method at each control cycle to generate the desired end-effector trajectory X_c . The position controller then tracks this trajectory, causing the robot to respond compliantly to external forces rather than strictly holding its commanded pose. To improve performance during insertion, the stiffness and damping gains are continuously scaled as a function of the current insertion depth z according to the relation:

$$s = 1 + (s_m - 1)(1 - z/z_t) \quad (4)$$

where z_t is the total depth of the interface hole, z is the current gripper depth offset and s_m is a maximum scale factor. At shallow depths, lower stiffness and damping allow larger lateral displacements, enabling the gripper to be guided into the hole when contacting the rim. As the gripper moves deeper, the increased gains

reduce compliance, stabilize the motion, and ensure precise alignment as clearance becomes smaller. When the measured vertical contact force F_{int} reaches the expected level F_d , and the final trajectory deviation x_e is smaller than a threshold, the insertion is considered complete. This combination of real-time admittance integration, depth-dependent gain scaling, and termination based on measured contact force provides a robust solution for reliable and repeatable insertion under environmental uncertainties.

4. EXPERIMENTS AND RESULTS

4.1. Experimental setup

Experiments were conducted to determine the effectiveness of the proposed pipeline and control architecture. The setup for these experiments is shown in Fig. 3. The off-the-shelf UR3e cobot arm was used to stand in for the BBVT as no full-scale or down-scaled BBVT has been completed as of the time of writing. The manipulator is equipped with a kinematically accurate 3:10 scale folding gripper prototype made from PLA and PETG using additive manufacturing, while the complementary interface fixture was bolted within reach of the UR3e. The gripper prototype is fully functional, with a Hitec HSR-2645CRH servo driving the folding arm mechanism via a 3-D printed rack and spur gear. To each folding arm tip is mounted a FSR 402, the output signal of which is wired in series with a 100 k Ω resistor to behave as a switch in accordance with the datasheet. The UR3e integrates the desired wrist F-T sensor, while a Grove v2 ultrasonic range finder mounted near the base of the gripper measures the distance to the BB segment top surface, completing the trio of sensors used for action evaluation. The servo, FSRs and range finder are managed by an Arduino Uno microcontroller, which communicates to the controlling PC via serial. For object detection, a RealSense D435 RGB+IR stereo camera is mounted opposite the ultrasonic range finder near the gripper base and connected directly to the PC. The UR3e is controlled from the PC over Ethernet connection using the Python library “ur_rtde,” which implements a RTDE-based interface.

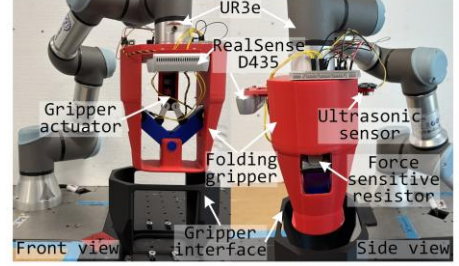


FIG. 3. Experimental setup, including downscaled 3-D printed gripper with folding arm mechanism, BB segment interface, UR3e manipulator and sensor array.

4.2. Heuristics for transporter state evaluation

The outcome of each action is categorized as “okay” if the following criteria are met, otherwise “NG”:

- Alignment: the x - and y -offset of the gripper from the ideal pose are each ≤ 2 cm and the ultrasound distance measurement is ≤ 22 cm.
- Insertion: the F_z force component registered by the UR3e wrist F-T sensor is ≥ 4 N, and the ultrasound distance measurement is in the range $[3, 5]$ cm.
- Engagement: the ultrasound distance measurement is in the range $[3, 5]$ cm and at least one of the analog readings from the arm-tip sensors is ≥ 2.5 V.

4.3. Results and discussion

4.3.1. Object detection

The YOLOv8 detector was trained for 388 epochs until automatic early stopping based on validation performance. As shown in Fig. 4(a), the validation box and classification losses converged smoothly, while mAP50 and mAP50–95 rapidly increased and reached values above 0.95 within the first 100 epochs, remaining steady for the rest of training. This indicates stable convergence and sufficient training duration to capture the object features. Representative results in Fig. 4(b) show that the detector reliably identifies both the small and large 3D-printed interface geometries under a wide range of orientations, positions, and lighting conditions. The predicted bounding boxes align well with the ground truth annotations, and confidence scores are typically higher than 0.9, demonstrating that the trained network is robust and well-suited for integration in the grasping pipeline.

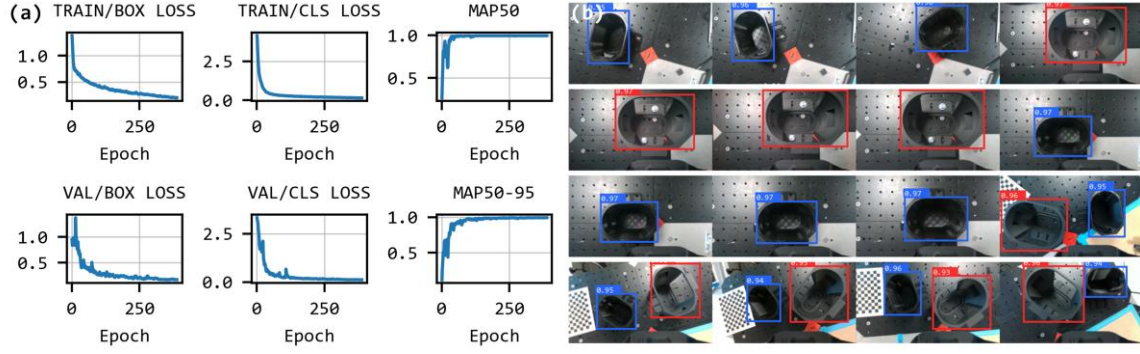


FIG. 4. (a): Performance of object detector on validation dataset. (b): sample detection results using two different 3-D printed interface geometries (small and big).

4.3.2. Full pipeline with admittance control and rule-based action evaluator

To verify the robustness of the proposed pipeline, it was implemented and executed 35 times using the downscaled test system (Fig. 3) while introducing uniformly random x and y offsets of up to ± 4 cm for both the detection gripper alignment poses. The possible error in interface localization due to detecting the partially cut off interface in the images combined with the artificial offset to result in cumulative alignment pose errors that could be significantly greater than ± 4 cm. For comparison, the width and height of the interface hole are 18.5 and 12.4 cm, respectively. These random pose offsets simulated the presence of large positioning errors caused by sensor or actuator faults or inaccuracies in object detection due to noise.

Of the 35 experiments conducted, only 14 (40%) resulted in the gripper avoiding collision with the top surface of the BB segment. However, the overall grasping pipeline success rate was 57%, as the VAC was able to recover in 6 experiments (29% of collisions), inserting the gripper after one or more bounces against the BB segment upper surface. Trajectory and force data from one such experiment is shown in Fig. 5.

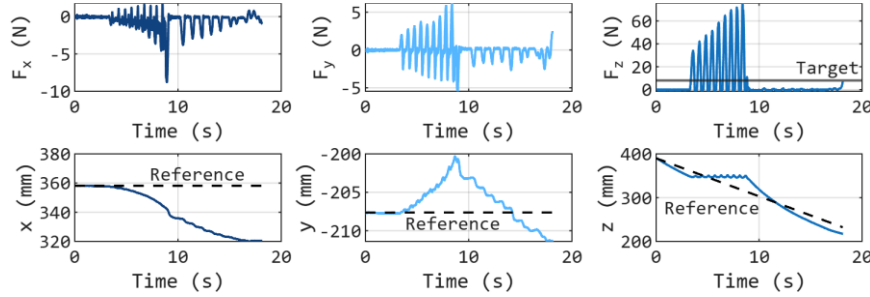


FIG. 5. Gripper position and wrist F-T sensor data for an insertion attempt in which the gripper collided with the top BB segment interface surface, yet the VAC was able to successfully complete the grasp.

The 15 unsuccessful grasps all failed during the insertion step of the pipeline due to collision with the BB segment upper surface. In other words, every grasp in which the admittance control recovered from collision or never collided at all resulted in successful engagement of the locking mechanism, demonstrating the reliability of the variable admittance policy at inserting the gripper completely and correctly with a tight fit.

On the other hand, the high number of insertion failures points to a flawed implementation of the alignment evaluator. Indeed, only a single “NG” alignment evaluation was recorded across 20 successful grasps. This can be largely explained by suboptimal integration of the ultrasonic sensor into the gripper prototype. The significant beam spreading in the cheap off-the-shelf model combined with the mounting location near the gripper base caused it to be insensitive to small or medium interface misalignments. Also, the use of one sensor on the side of the gripper meant that the evaluator was only sensitive to misalignments in a single direction which would cause the sensor to be placed above the interface hole. Thus, for reliable alignment evaluation using this sensor type, they should be integrated into the bottom surface of the gripper and at least four should be used (one per edge).

Fig. 6 illustrates how the implemented alignment, insertion and engagement evaluators are meant to help reduce failures by using simple heuristics like the distance measured from the ultrasonic sensor to trigger re-planning. This is true in principle, however, given the observed results, improvements such as incorporate more sensor modalities are required to improve reliability. Early trials indicate that integrating data from the RGB camera image can significantly improve the reliability of the alignment evaluator, for example.

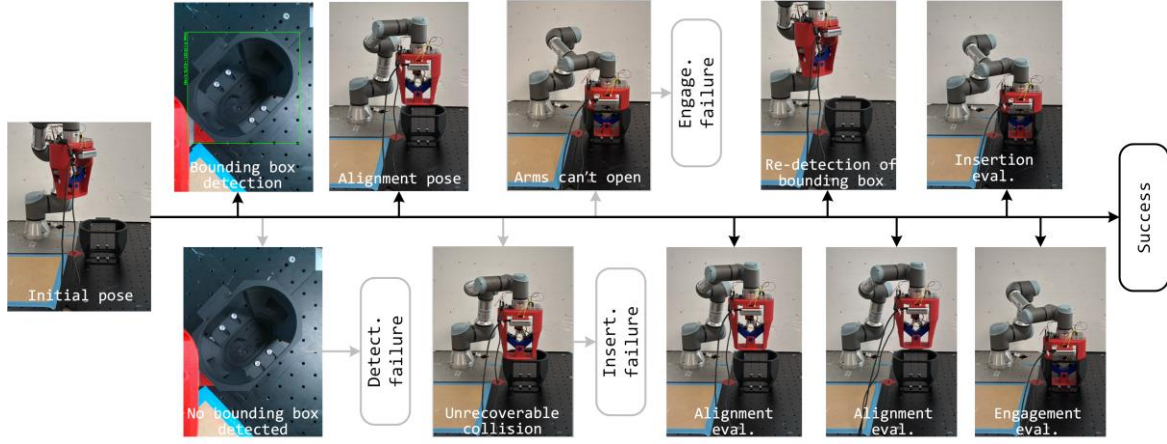


FIG. 6. Demonstration of the pipeline as implemented. Gray arrows show some of the failure types which are more likely to occur without action evaluation, due to the lack of state awareness and re-planning.

5. CONCLUSION AND FUTURE WORK

This work introduced a pipeline consisting of four steps for BB segment grasping at EU-DEMO using the BBVT robot and its actuated gripper. A control framework for BB grasping was also proposed which aimed to be robust to potential sensor and actuator faults resulting from gamma radiation while achieving the tight fit needed for BB manipulation. This included fault-tolerant action planning with multisensory robot state evaluation together with reliable interface detection using YOLO and robust contact-rich insertion using VAC. Experiments were carried out using a 3:10 scale gripper prototype complete with actuated locking mechanism and sensor array including an F-T sensor, rangefinder and two FSRs. With a successful grasp rate of 57% despite large random artificial pose errors leading to collisions in 60% of trials, the experiments showed that the VAC was fault-tolerant and reliable, but the alignment evaluator was flawed and failed to ensure gripper alignment that would prevent collision. Thus, a major focus of future work will be improving the evaluators by design changes or integrating more modalities, e.g., using transformer or diffusion-based models to incorporate RGB image data. An accurate downscaled 7-degree of freedom BBVT prototype with which the gripper can be integrated is also planned. Also, the gripper and BB segment interface will be manufactured with better tolerances and more accurate materials in the future.

ACKNOWLEDGEMENTS

H. Durocher would like to thank EUROfusion for financial support for his Ph.D. thesis project. X. Yang would also like to thank the EUROfusion Engineer Grant for financial support for his postdoc project.

REFERENCES

- [1] FEDERICI, G., Testing needs for the development and qualification of a breeding blanket for DEMO, Nucl. Fusion **63** 12 (2023) 125002.
- [2] COLEMAN, M. et al., Concept for a vertical maintenance remote handling system for multi module blanket segments in DEMO, Fusion Engineering and Design **89** 9 (2014) 2347.
- [3] BACHMANN, C. et al., Progress in the development of the in-vessel transporter and the upper port cask for the remote replacement of the DEMO breeding blanket, Fusion Engineering and Design **194** (2023) 113715.
- [4] BACHMANN, C., GLISS, C., JANESCHITZ, G., STEINBACHER, T., MOZZILLO, R., Conceptual study of the remote maintenance of the DEMO breeding blanket, Fusion Engineering and Design **177** (2022) 113077.

- [5] DIGGINS, Z.J. et al., Range-Finding Sensor Degradation in Gamma Radiation Environments, *IEEE Sensors Journal* **15** 3 (2015) 1864.
- [6] KHANAM, Z. et al., Gamma-Induced Image Degradation Analysis of Robot Vision Sensor for Autonomous Inspection of Nuclear Sites, *IEEE Sensors Journal* **22** 18 (2022) 17378.
- [7] COLOMA, S. et al., The effect of ionizing radiation on robotic trajectory movement and electronic components, *Nuclear Engineering and Technology* **55** 11 (2023) 4191.
- [8] ROLFE, A.C. et al., A report on the first remote handling operations at JET, *Fusion Engineering and Design* **46** 2 (1999) 299.
- [9] HONDA, T. et al., Remote handling systems for ITER, *Fusion Engineering and Design* **63–64** (2002) 507.
- [10] FARFALETTI-CASALI, F. et al., The interaction of systems integration, assembly, disassembly and maintenance in developing the INTOR-NET mechanical configuration, *Nuclear Engineering and Design. Fusion* **1** 2 (1984) 115.
- [11] LOVING, A. et al., Pre-conceptual design assessment of DEMO remote maintenance, *Fusion Engineering and Design* **89** 9 (2014) 2246.
- [12] STEINBACHER, T., BACHMANN, C., GLISS, C., JANESCHITZ, G., MOZZILLO, R., Design of the gripper interlock that engages with the DEMO breeding blanket during remote maintenance, *Fusion Engineering and Design* **193** (2023) 113641.
- [13] DUROCHER, H., BACHMANN, C., MOZZILLO, R., JANESCHITZ, G., ZHANG, X., Motion planning with inverse kinematics and statics of a breeding blanket transporter for robotic remote maintenance of the EU DEMO tokamak, *Research Square* (preprint) (2025).
- [14] AGRAWAL, A., SUN, Y., BARNWELL, J., RASKAR, R., Vision-guided robot system for picking objects by casting shadows, *The International Journal of Robotics Research* **29** 2–3 (2010) 155.
- [15] DALAL, N., TRIGGS, B., Histograms of Oriented Gradients for Human Detection, 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 1, IEEE (2005) 886–893.
- [16] ZOU, Z., CHEN, K., SHI, Z., GUO, Y., YE, J., Object detection in 20 years: A survey, *Proceedings of the IEEE* **111** 3 (2023) 257.
- [17] VAN DE SANDE, K.E., UIJLINGS, J.R., GEVERS, T., SMEULDERS, A.W., Segmentation as Selective Search for Object Recognition, 2011 International Conference on Computer Vision, IEEE (2011) 1879–1886.
- [18] REN, S., HE, K., GIRSHICK, R., SUN, J., Faster R-CNN: Towards real-time object detection with region proposal networks, *IEEE transactions on pattern analysis and machine intelligence* **39** 6 (2016) 1137.
- [19] HE, K., ZHANG, X., REN, S., SUN, J., Spatial pyramid pooling in deep convolutional networks for visual recognition, *IEEE transactions on pattern analysis and machine intelligence* **37** 9 (2015) 1904.
- [20] LIN, T.-Y. et al., Feature Pyramid Networks for Object Detection, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2017) 2117–2125.
- [21] REDMON, J., DIVVALA, S., GIRSHICK, R., FARHADI, A., You Only Look Once: Unified, Real-Time Object Detection, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2016) 779–788.
- [22] VARGHESE, R., SAMBATH, M., YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness, 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), IEEE (2024) 1–6.
- [23] LIU, W. et al., Ssd: Single Shot Multibox Detector, *European Conference on Computer Vision*, Springer (2016) 21–37.
- [24] LIN, T.-Y., GOYAL, P., GIRSHICK, R., HE, K., DOLLÁR, P., Focal Loss for Dense Object Detection, *Proceedings of the IEEE International Conference on Computer Vision*, (2017) 2980–2988.
- [25] NI, J. et al., Deep learning-based scene understanding for autonomous robots: A survey, *Intelligence & Robotics* **3** 3 (2023) 374.
- [26] SONG, P., YU, Y., ZHANG, X., A tutorial survey and comparison of impedance control on robotic manipulation, *Robotica* **37** 5 (2019) 801.
- [27] ZHOU, Z., YANG, X., ZHANG, X., Variable impedance control on contact-rich manipulation of a collaborative industrial mobile manipulator: An imitation learning approach, *Robotics and Computer-Integrated Manufacturing* **92** (2025) 102896.
- [28] HOGAN, N., On the stability of manipulators performing contact tasks, *IEEE Journal on Robotics and Automation* **4** 6 (2002) 677.
- [29] SCHUMACHER, M., WOJTUSCH, J., BECKERLE, P., VON STRYK, O., An introductory review of active compliant control, *Robotics and Autonomous Systems* **119** (2019) 185.
- [30] ABU-DAKKA, F.J., SAVERIANO, M., Variable impedance control and learning—a review, *Frontiers in Robotics and AI* **7** (2020) 590681.
- [31] FERRAGUTI, F. et al., A variable admittance control strategy for stable physical human–robot interaction, *The International Journal of Robotics Research* **38** 6 (2019) 747.