Robotic Assistive Smart Touch Inspection of Offshore Oil and Gas Pipeline

Contract #: 140E0123C0007

FINAL REPORT

July 27, 2023 - February 28, 2026

Principal Investigator:

Dr. Zheng Chen

University of Houston

Co-Principal Investigator:

Dr. Gangbing Song

University of Houston

BSEE REPORT COVERSHEET

1. PUBLICATION / CREATION DATE (DD-MM-YYYY):	2. SOURCE / REPORT TYPE (Interim, Final, etc.):	3. START DATE: (DD-MM-YYYY)
28-02-2026	Final Report	27-07-2023
		4. END DATE: (DD-MM-YYYY) 28-02-2026

5. TITLE:	6. DIGITAL OBJECT IDENTIFIER (DOI) NUMBER:
Robotic Assistive Smart Touch Inspection of Offshore Oil and Gas Pipeline	10.70149/27941158
7. CONTRACT/IAA NUMBER: 140E0123C0007	8. RESEARCH PROJECT NUMBER (TAP, TCP, etc.): TAP 812

9. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES):	10. CREATOR / AUTHOR(S) + OPEN RESEARCHER AND CONTRIBUTOR ID (ORCID):
University of Houston 4300 MARTIN LUTHER KING BLVD HOUSTON TX 77204-3067	Zheng Chen, 0000-0002-8918-4066 Gangbing Song, 0000-0001-5135-5555

11. DISTRIBTION / AVAILABILITY / ACCESS RESTRICTIONS (e.g., Unrestricted, restriction limitations, CUI, etc.):	12. USE RESTRICTIONS (e.g. copyright, trademark, internal viewing only, N/A etc.):
Unrestricted	N/A
	13. RIGHTS OWNER (If applicable): N/A

14. FILE RELATIONSHIP (e.g. Has part, Is Part Of):	15. SUPPLEMENTARY NOTES (ex. level of influence, related documents):
N/A	Not influential, 2 journal articles and 3 conference presentations published based on this research, research data disseminated

16. ABSTRACT/ DESCRIPTION (brief, 200-word, factual summary of most significant information):

The demand for offshore oil and gas exploration and production is steadily increasing, driving the expansion of underwater infrastructure. Among the critical components of this infrastructure, pipelines play a vital role in transporting oil and gas efficiently. Flange connections are the most common method for joining pipelines, installing valves, and extending functionality offshore. This research project presents a novel approach for autonomous smart-touch flange inspection using a Remotely Operated Vehicle (ROV). The proposed system is implemented on a BlueROV2 equipped with a gripper that integrates a pair of piezoceramic transducers (PZTs). These PZTs act as smart touch sensors, enabling bolt looseness inspection when the gripper's fingers securely grab a flange.

17. SPONSORING AGENCY NAME(S) / JOINT INDUSTRY PROJECT (JIP):	18. KEY WORDS/PHRASES/CATEGORIES:	19. DIMENSIONS / NUMBER OF PAGES:
BSEE	Bolted structure failure, Boltness inspection, Remotely Operated Vehicle, Control, Machine Learning, Navigation	102

20. CONTRACTING OFFICER REPRESENTATIVE (COR):	21. EMAIL ADDRESS:
Bipin Patel	bsee_researchpubliccomment@bsee.gov

ADMINISTRATIVE RECORDS MANAGEMENT

1. FILE NAME (Computer file name + extension):	2. RECORD ID:
	3. RECORDS SCHEDULE ITEM #:

Privacy Advisory: *This form is for internal BSEE use only. Do not include sensitive personally identifiable information (PII), such as Social Security numbers or dates of birth. Information provided is used solely for official business purposes.*

Table of Contents

LIST OF ACRONYMS

Acronym	Definition
AUV	Autonomous Underwater Vehicle
BCDs	Buoyancy Control Devices
BSEE	Bureau of Safety and Environmental Enforcement
CADA	Centroid Alignment Domain Adaptation
DAQ	Data Acquisition
DANN	Domain-Adversarial Neural Network
DFT	Discrete Fourier Transform
DOI	Department of the Interior
EKF	Extended Kalman Filter
ESKF	Error-State Kalman Filter
FC	Fully Connected (layers)
FFT	Fast Fourier Transform
FRF	Frequency Response Function
GICP	Generalized Iterative Closest Point
IMU	Inertial Measurement Unit
MFCC	Mel-Frequency Cepstral Coefficient
PCB	Printed Circuit Board
PD	Proportional-Derivative (controller)
PEM	Polymer Electrolyte Membrane
PID	Proportional-Integral-Derivative (controller)
PINN	Physics-Informed Neural Network
PZT	Piezoelectric (Lead Zirconate Titanate) Transducer
PWM	Pulse-Width Modulation
RANSAC	Random Sample Consensus
RBF	Radial Basis Function
ReLU	Rectified Linear Unit
ROI	Region of Interest
ROV	Remotely Operated Vehicle
RSS	Received Signal Strength
SHM	Structural Health Monitoring
SVM	Support Vector Machine
YOLO	You Only Look Once

I. DISCLAIMER

Study concept, oversight, and funding were provided by the US Department of the Interior (DOI), Bureau of Safety and Environmental Enforcement (BSEE), Office of Offshore Regulatory Programs (OORP), Sterling, VA, under Contract Number 140E0123C0007. This report has been technically reviewed by BSEE, and it has been approved for publication. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the US Government, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

II. INTRODUCTION

The demand for offshore oil and gas exploration and production is steadily increasing, driving the expansion of underwater infrastructure. Among the critical components of this infrastructure, pipelines play a vital role in transporting oil and gas efficiently. Flange connections are the most common method for joining pipelines, installing valves, and extending functionality offshore. However, subsea pipelines, especially their connection points, are more vulnerable to damage due to harsh underwater conditions such as high pressure, strong currents, and corrosive elements. Bureau of Safety and Environmental Enforcement (BSEE) highlighted connector and bolt failures as a recurring and significant concern in subsea infrastructure systems. Therefore, regular inspection and maintenance of flanges are essential to ensure the safety and reliability of these pipelines, preventing potential leaks, and mitigating environmental and economic risks.

This research presents a novel approach for autonomous smart-touch flange inspection using a Remotely Operated Vehicle (ROV). The proposed system is implemented on a BlueROV2 equipped with a gripper that integrates a pair of piezoceramic transducers (PZTs). These PZTs act as smart touch sensors, enabling bolt looseness inspection when the gripper's fingers securely grab a flange. The specific aims of this research can be summarized as follows:

- Develop **autonomous control** and underwater navigation for ROVs to conduct smart touch inspection.
- Develop a robotics-enabled and **machine learning-based** active sensing approach for smart touch inspection of bolted flange.
- **Integrate** ROVs with all developed components and conduct comprehensive testing with industry partners.

III. SYSTEM DESCRIPTION

The goal of this research is to develop an autonomous system which can detect underwater flanges, navigate a ROV toward them, and securely grasp them to complete the inspection mission. The proposed system is illustrated in Fig. 1. The platform used is BlueROV2, an underwater vehicle from Blue Robotics, equipped with a single monocular camera as the sole sensing component in this study. To facilitate grasping, a Newton Subsea Gripper from Blue Robotics is installed beneath the ROV. The end effector of the gripper has been redesigned and fabricated with 3D printing to integrate a pair of PZT transducers, ensuring compatibility with the flange thickness and enabling effective flange inspection. As an active sensing method, one PZT transducer generates a stress wave that travels across the bolted connection, while the other transducer detects the propagating wave. By analyzing the differences between the excitation and sensing waves, the health condition of the flange can be assessed.

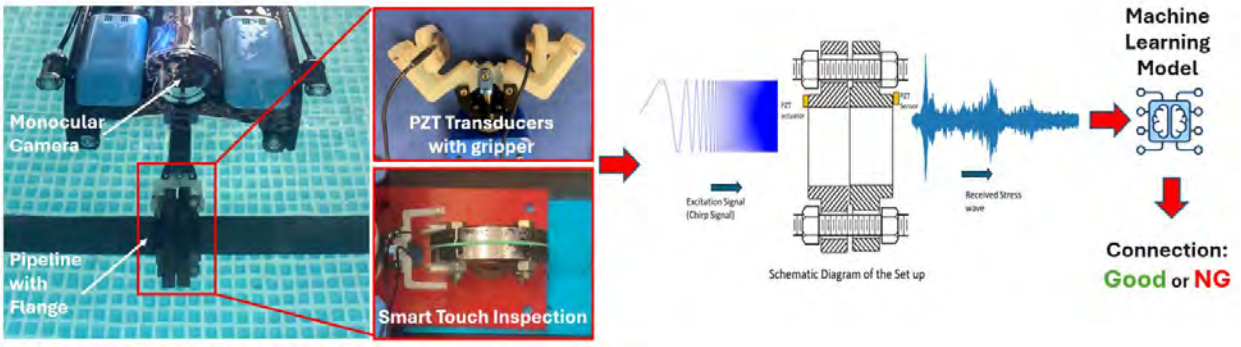


Fig. 1. ROV and Gripper-Based Flange Looseness Inspection System

We have worked on 8 different flanges (4.25-inch, 5-inch, 6-inch, 9-inch, 11-inch, 13-inch, 16-inch, and 21-inch). All the significant results that we report happened after March 25 2024, since we changed the design of the gripper, which made the signal much stronger and helped with the classification task. For the machine learning model, we developed mainly 3 algorithms to do the classification: a support vector machine (SVM) model with MFCC features, a Physics- informed 1D convolutional neural network model with energy of the signal as physical information, and a domain adaptation model.

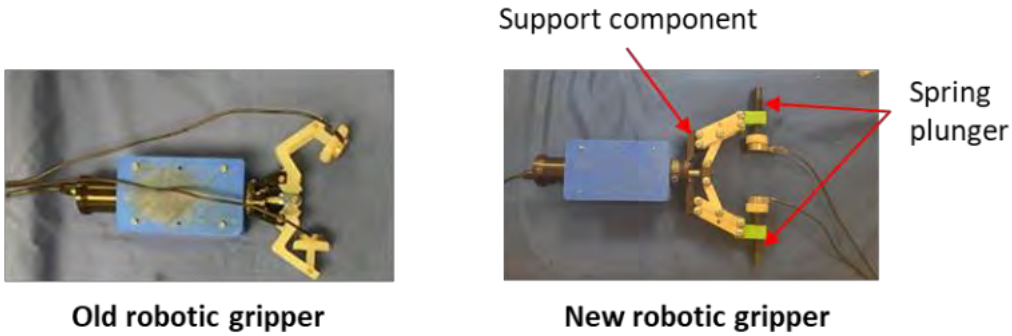
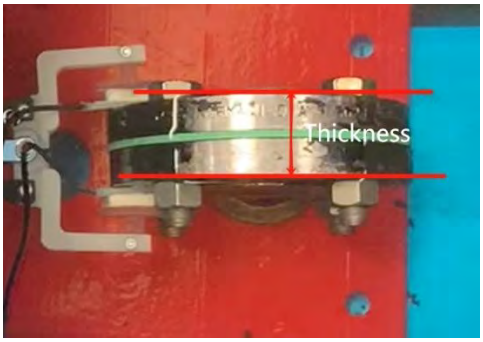


Fig. 2. Old and New Robotic Grippers.

IV. AUTONOMOUS CONTROL OF ROV FOR FLANGE SMART TOUCH INSPECTION

IV.1 Gripper Integrate with Smart Touch Sensor

The system utilizes the Blue Robotics Newtown Gripper as the primary actuator, selected for its ease of ROV integration and robust performance down to a 300-meter depth rating. This actuator offers an adjustable grasping force up to a maximum of 100 Newtons (N). Its internal brushed motor is controlled via a purpose-built controller that accepts a standard servo-style PWM signal, ensuring precise remote operation. The unit is electrically compatible with a 9-18V input and draws a maximum current of 6A. As shown in Fig. 3, we need to redesign the jaw to adapt with different types of flanges.



Flange Type	Nom. Pipe Size(inch)	Thickness of a pair of flanges (mm)	Gripper size
ANSI B16.5 150	1/2	19.30	A: 17mm-30mm
ANSI B16.5 150	1	25.40	
ANSI B16.5 150	1-1/2	31.50	B: 30mm-50mm
ANSI B16.5 150	2	35.05	
ANSI B16.5 150	4	44.70	
ANSI B16.5 150	6	47.75	
ANSI B16.5 150	10	60.4	C: 55mm-75mm
ANSI B16.5 150	14	70.104	

Fig. 3. Flanges to be tested

As shown in Fig .4, firstly, we redesigned and fabricated the jaws to make them more compatible for grasping the flange with different thicknesses. The shaft of the adapter will rotate during the grasping and provide normal force for different thicknesses. Then we integrated the gripper with the smart touch sensor. By designing the adaptor with elastic material (3D printing by Formlabs Flexible 80A), we successfully isolate the wave propagation through the gripper, making sure that the PZT sensor only receives the signal through the flange. We programmed the gripper for different forces and designed the circuit to control the gripper by ESP32.

Fig. 4. 1st version of gripper with adapter redesign

For improvement, we redesigned the gripper to address two major issues: the lack of a spring mechanism, which caused inconsistent grasping force, and the misalignment of the gripper's two axes, leading to inefficiency in data collection. To improve performance, we retained the Blue Robotics gripper's main body for better integration and waterproofing while redesigning the grabbing mechanism. We implemented a double-linkage mechanism that leverages a 10mm linear actuator stroke into a parallel and well-aligned gripping motion. Additionally, we installed a spring plunger at the end effector to provide passive and consistent force, replacing the previous stiff gripping mechanism. The 3D-printed prototype allowed us to test tolerances and evaluate two types of spring plungers (Heavy and Light), resulting in improved grasping quality, efficiency, and more distinct features for PZT sensor data collection. Furthermore, by adjusting the installation depth of the spring plunger, the gripper can accommodate flanges of different thicknesses and adjust the grabbing force. We also upgraded the key parts of the gripper to Aluminum and tested both light

and heavy configurations of spring plunges for better grasping validation. We have successfully collected 2600+ high quality, clearer data points without any mechanical issues.

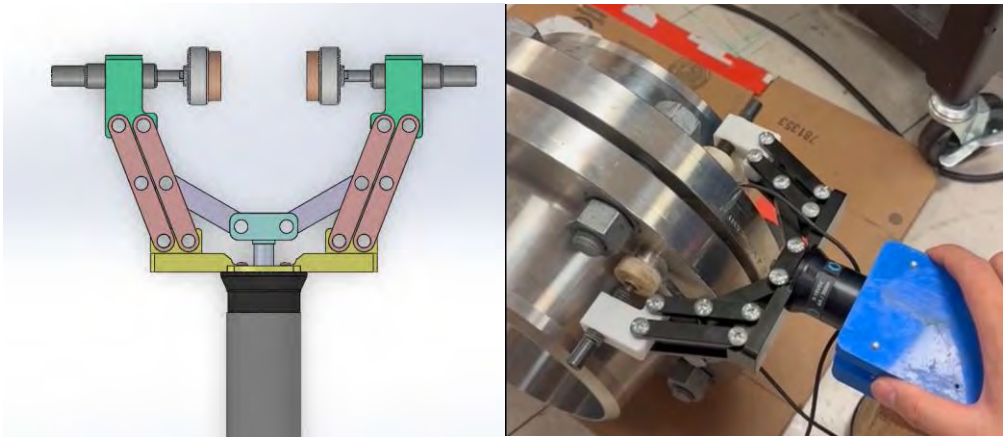


Fig. 5. 2nd version of gripper with adapter redesign

IV.2 Image Processing and Yolo-based Flange Detection

YOLOv8 (You Only Look Once) (Redmon, 2016) is the latest iteration in the YOLO series of real-time object detection models, building on the success of its predecessors by introducing significant improvements in both accuracy and efficiency. Developed to perform object detection tasks swiftly and accurately, YOLOv8 is designed for applications requiring rapid and reliable object recognition in various settings, including autonomous driving, surveillance, and interactive systems where real-time response is crucial. Blue Robotics ROV has a high definition (1080p, 30fps) wide-angle low-light camera, which is suitable for tasks like object detection.

Dataset Prepare: In the development of an object detection model targeting industrial components, specifically flanges and pipelines, a dataset comprising **314 images** extracted from manually controlled videos has been meticulously labeled and utilized. The images are distinctly annotated under two primary categories: "flange" and "pipeline." For developing and validating the model, the dataset was strategically divided following an 80/20 split. Consequently, 251 images were used for training, facilitating the model's learning in accurately identifying and localizing the specified components. The remaining 63 images serve as a validation set, which is crucial for evaluating the model's performance on new, unseen data. This approach ensures the model's robustness and its capability to generalize well across similar tasks in real-world applications, aiming for high reliability in operational environments.

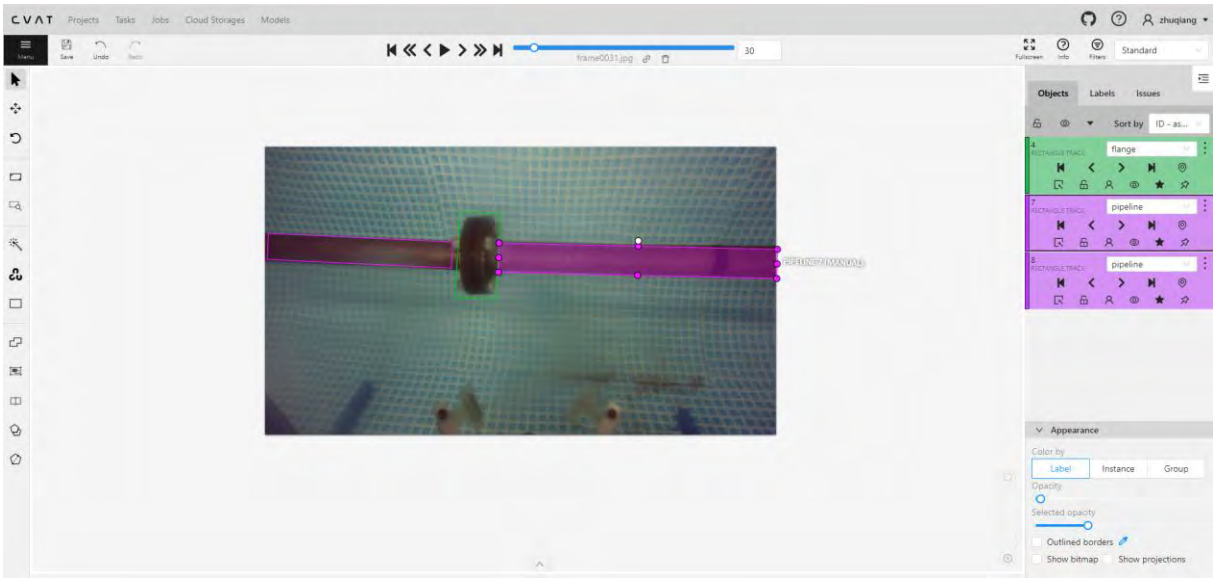


Fig. 6. Annotation of the flanges and pipelines.

Both the training and validation loss metrics for box, class (cls), and distribution-focal (dfl) loss have shown significant reductions early in the training and have plateaued, indicating convergence. The losses stabilize after approximately 50 epochs, suggesting that the model has learned effectively from the training data without showing signs of overfitting, as the validation losses mirror the training losses closely. The object detection model demonstrates promising accuracy when applied to unseen video footage; however, its performance highlights the need for enhanced robustness. The inconsistencies observed suggest the model's training data may not fully encapsulate the variability of real-world conditions.

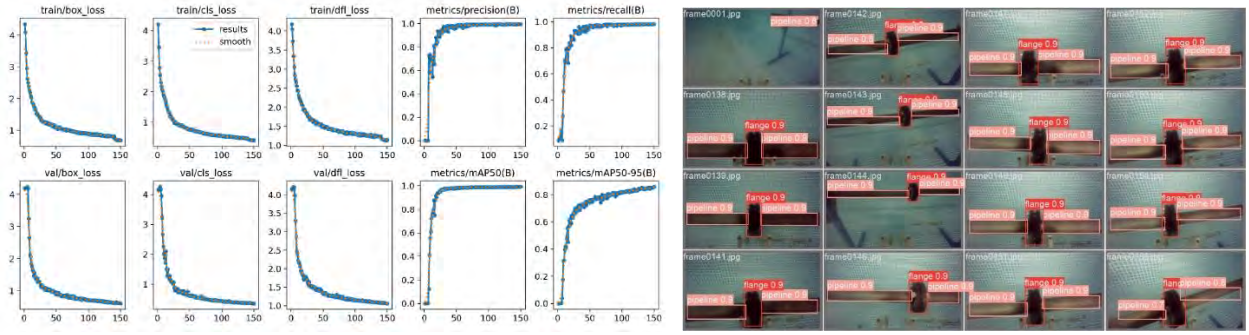


Fig. 7. Results for the flange detection, left detection metrics; right: detection result on validation set.

We expanded our dataset for flange detection, increasing it to 970 labeled images covering 5 different types of flanges. Additionally, we sourced around 100 flange images from online databases and included negative samples in the dataset to enhance the model’s generalization capability. This diverse data set aims to improve the robustness of our detection system. During training, we applied data augmentation and regularization techniques to further improve the model’s performance. The results indicate a significant improvement in detection efficiency. The model is now capable of reliably detecting flanges within a range of 10 cm to 3 m. This increased detection stability provided more consistent target values for our controller, improving overall system performance.

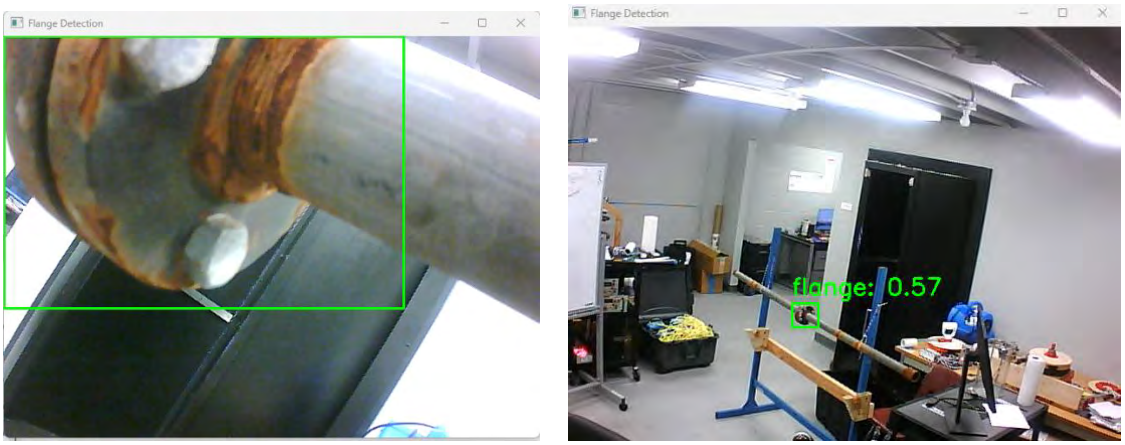


Fig. 8. Flange detection results left: <10cm; right: >3m.

After we have preliminary results of YOLO based flange detection, we processed the onboard recorded video to detect flanges using a YOLO object detection model and identify lines using the Hough Line Transform (Ballard, 1981). It reads an input video, detects flanges in each frame, draws bounding boxes around them, and marks their centers with yellow dots. Concurrently, it performs edge detection and uses the Hough Line Transform to find the boundary of the pipelines. We also calculated the angle between these detected pipelines and a horizontal red reference line, which represents the gripper; this angle serves as a parameter for achieving the autonomous control of the ROV.

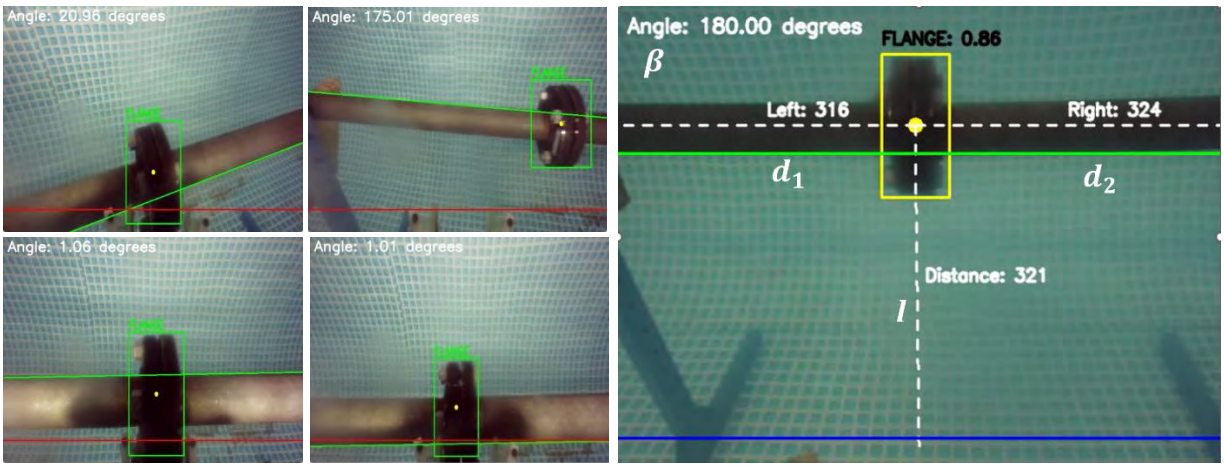


Fig. 9. Image processing for the flange detection.

The image streamed from the on-board camera is processed using a YOLO-based detection algorithm and a Hough Transform-based pipeline detection algorithm to determine the relative orientation between the ROV and the flange. Fig. 9 illustrates the image processing results at the beginning and end of the automation process. In the processed image, the green line represents the detected pipeline, and the yellow box highlights the detected flange, with the accompanying score (Flange: 0.86 in Fig .9) indicating the detection of confidence. The gripper is visible at the mid-bottom of the image, while the fixed blue line serves as a reference line to represent the gripper's location. Four key values are extracted through image processing: (1) β, the angle between the pipeline and the reference horizontal line; (2) d_1, the pixel distance from the center of the flange to the left boundary; (3) d_2, the pixel distance from the center of the flange to the right boundary; and (4) l, the pixel distance from the center of the flange to the gripper. These

parameters provide the necessary feedback for accurate control and alignment of the ROV during the automation process. The errors for control are defined as:

$$e_1 = cos\beta,$$

$$e_2 = |\, d_1 - d_2|,$$

$$e_3 = l - l_{\{threshold\}}.$$

IV.3 Autonomous Control Based on Yolo Detection and Hough Transform

We conducted multiple experiments and tuned the PID parameters to achieve the autonomous smart touch process. Three steps (Parallelize, Center, Approach) are controlled by a PID controller to demonstrate feasibility. The thresholds for gripper activation were also found and tested in the experiments. The autonomous control framework is described in Fig. 10.

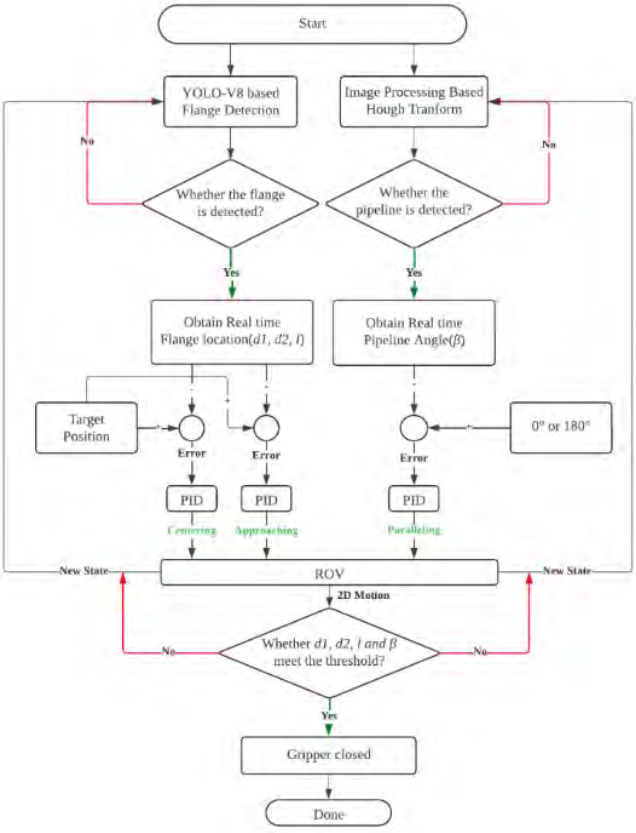


Fig. 10. Autonomous flange gripping process.

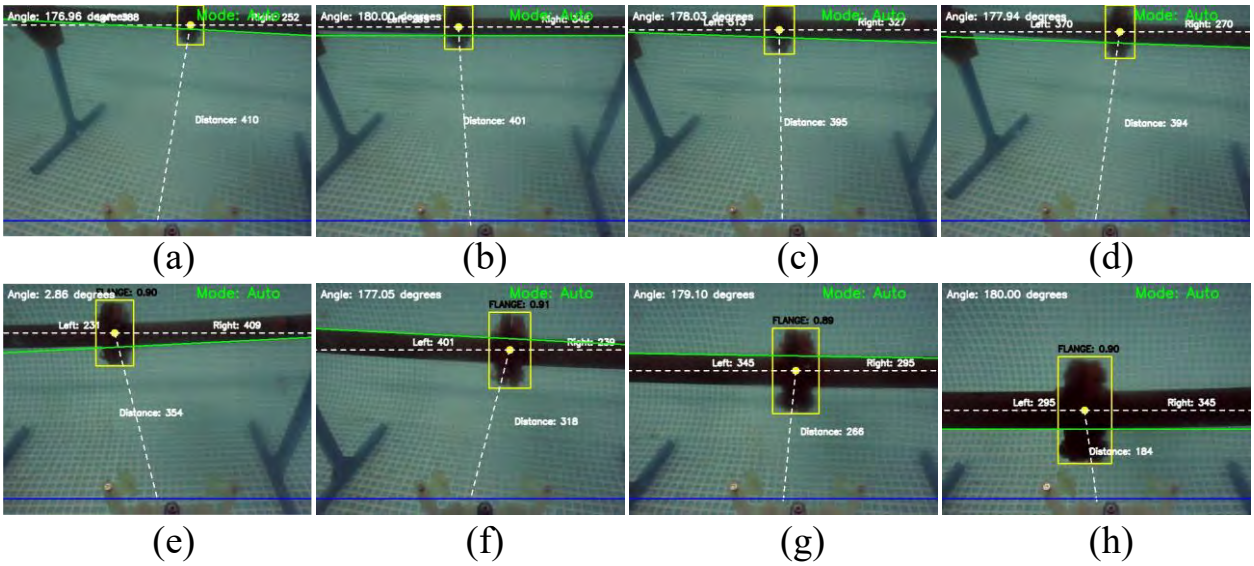


Fig. 11. On-board view of autonomous flange grasping with PID controller.

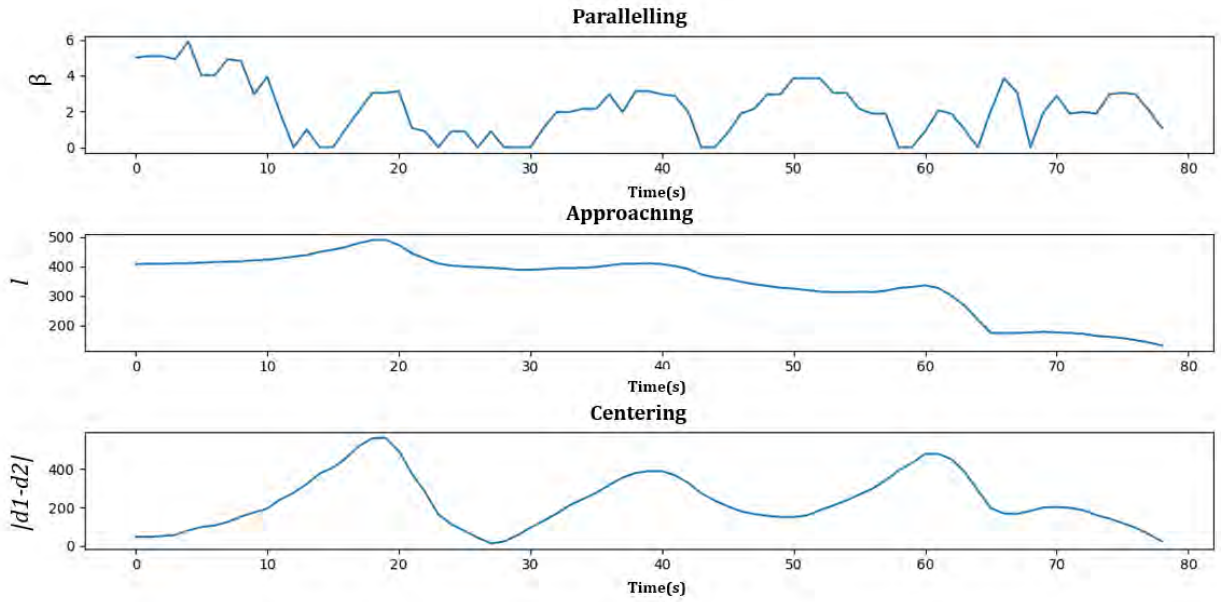


Fig. 12. Parallelling, approaching, centering errors of convergency.

As shown in Fig.12, due to the complexities of the underwater environment, oscillations are observed during the autonomous process. However, clear convergence can still be seen over time, and the system successfully reaches the target at 30s, allowing the ROV to grab the flange with the desired orientation.

IV.4 A Physics-Informed Neural Network Enhanced Kalman Filter Method for ROV localization

Compared to larger vessels, the nonlinearity in BlueROV localization is more pronounced due to its rapidly changing angular velocity (Fossen, 2023). The conventional EKF method, limited by the delayed update of

the covariance matrix, struggles to compensate for transient nonlinear errors in real time (Shaaban, 2024). To address this, we propose a learning-based approach to predict the actual nonlinear state. Building upon the existing EKF framework (Fossen, 2023), our method integrates a novel data-driven approach to enhance robustness against localization of nonlinearity. The proposed method is illustrated in Fig. 13.

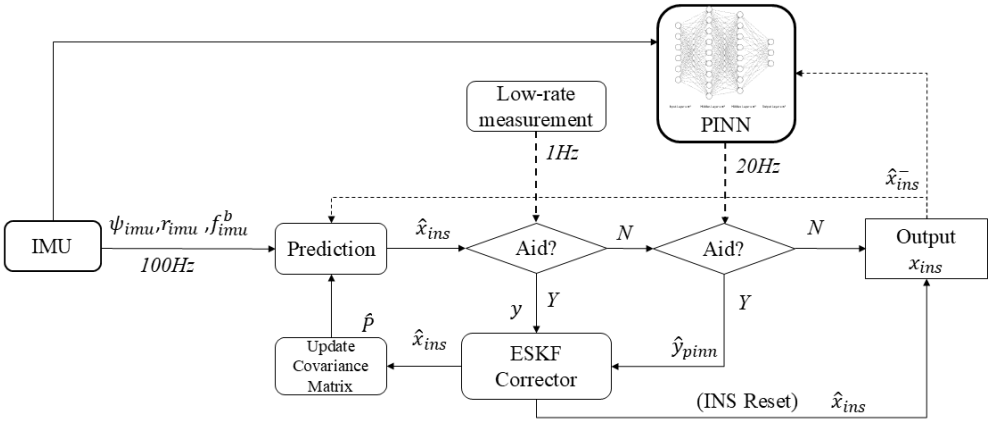


Fig. 13. Data flow of PINN enhanced ESKF.

The proposed method leverages Physics-Informed Neural Networks (PINNs) to predict velocity at a frequency of 20 *Hz*, using inputs such as acceleration f_{imu}, angular velocity r_{imu}, and the state x_{ins} from the previous estimation period. By integrating the nonlinear dynamic model of the ROV into the PINN, the method reduces nonlinear errors (Qian, 2023), enabling the network to provide a reference velocity to support the Kalman filter at a higher update rate, thereby mitigating nonlinearity effects. To test and compare the nonlinearity robustness of the proposed method, the simulation specification is as follows:

Name	sign	Value, noise and bias	Unit
Desired x-axis position	x_p	$x_p \sim \mathcal{N}(0,5.43)$	m
Desired y-axis position	y_p	$y_p \sim \mathcal{N}(0,4)$	m
Desired vertical position	z_p	$z_p \sim \mathcal{N}(1,0.1)$	m
Desired yaw angle	ψ_p	$\psi_p \sim \mathcal{N}(0,0.1\ or\ 1.0)$	rad
Acceleration	f_{imu}	$f \sim \mathcal{N}(0.15,0.2^2)$	m/s^2
Yaw rate	r_{imu}	$r \sim \mathcal{N}(0.1,0.2^2)$	m/s^2
3-D position	y	$y_n \sim \mathcal{N}(0.0,0.1^2)$	m

In the simulation (Michel, O. Cyberbotics ltd. webots™, 2004), this method is evaluated in a set-point control scenario under both high- and low-nonlinearity conditions. Reference points are generated based on a Gaussian distribution, and the BlueROV, controlled by a PID controller, tracks these points. Higher nonlinearity corresponds to more rapid changes in the reference yaw angle.

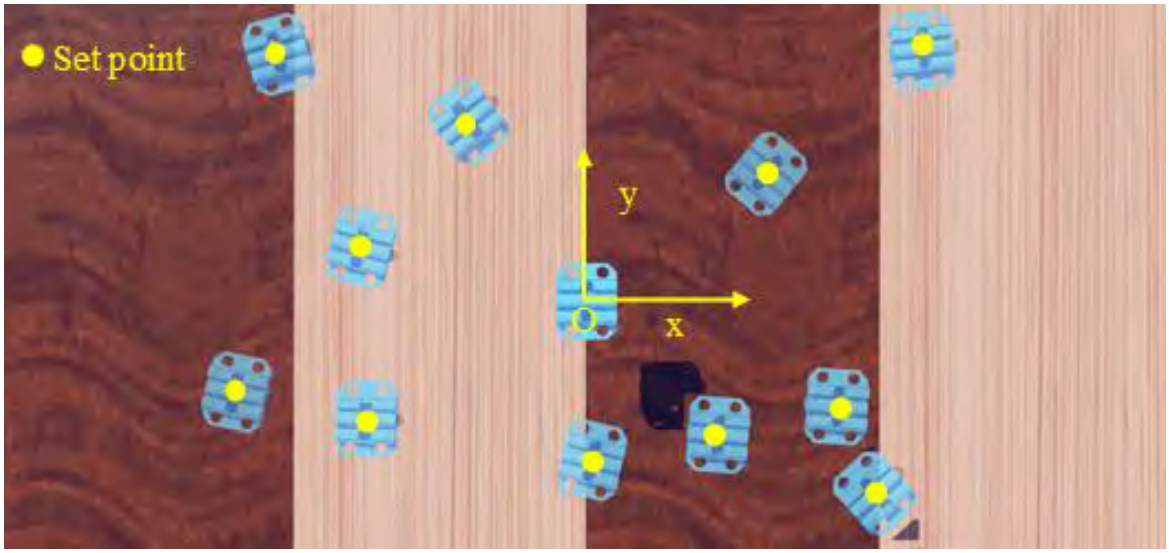


Fig. 14. Simulation for random set-point control scenarios.

As shown in Fig.15 under minimal system nonlinearity, both the traditional EKF and the PINN-enhanced EKF achieve accurate state estimation in the random set-point scenario, with only minor errors.

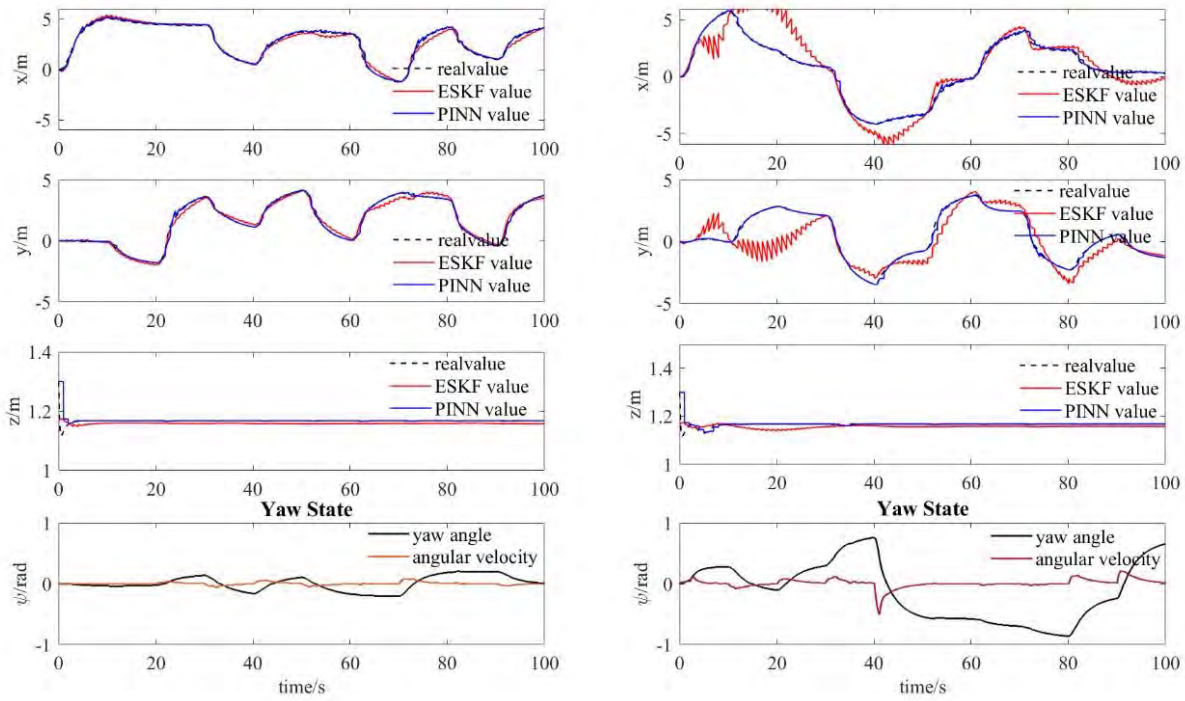


Fig. 15. Simulation for random set-point control scenarios.

As illustrated in Fig. 15, when the nonlinearity is intensified due to larger angular velocity changes in the yaw state, the traditional ESKF method produces significant errors, while the PINN-enhanced ESKF maintains its accuracy in estimation.

IV.5 Scanning Sonar based Long Range Detection

We expanded our research capabilities by acquiring a new ROV. A BlueROV2 Heavy has been assembled and tested, featuring eight thrusters that provide full six degrees of freedom for enhanced maneuverability. The new ROV is equipped with a scanning sonar for extended range exploration and a gripper designed for smart touch inspection, further advancing its potential for autonomous offshore inspection tasks. The system has been tested for waterproof performance and can operate at depths of up to 100 meters, ensuring reliability for further depth control.

Fig. 16. New BlueROV2 Heavy integrated with scanning sonar and gripper.

To enhance the autonomous control system in low-light conditions or at far distances where cameras are ineffective, we began detecting pipelines using sonar imaging. We deployed an ROV equipped with a Ping360 Scanning Sonar in a swimming pool, positioning a pipeline at the far end. With high gain settings, the sonar produced a clear red echo reflection of objects, allowing us to observe the rectangular boundaries of the pool and the straight-line reflection of the pipeline at a 5m distance. However, the high gain also introduced significant background noise, which poses a challenge for accurate detection.

To address this, we applied image processing following the steps shown in Fig. 17. Due to significant noise, directly applying a detection algorithm resulted in false positives, as some noise patterns have similar lengths to the pipeline. To mitigate this, we first detected the swimming pool boundary and then extracted the region inside the pool to isolate relevant sonar reflections. Since the pipeline echo appears red in the scanning sonar, we applied color segmentation to extract red regions, allowing us to enhance the pipeline signal while suppressing noise. After obtaining the red-masked image, we applied Canny edge detection to extract strong edge features corresponding to significant sonar reflections. This edge-detected image was then processed using the Hough Line Transform, which detects straight-line structures in the scene. However, due to noise and multiple small line segments, the raw detection results were fragmented. To improve accuracy, we applied a line merging algorithm, where nearby lines with similar angles and positions were grouped and combined into a single representative line. Once the major line structures were detected, we extended these lines to the image boundaries to ensure full pipeline coverage. However, as

some extensions crossed over unintended areas, we computed intersections between extended lines and clipped them at the nearest intersection points, keeping only the region within the swimming pool.

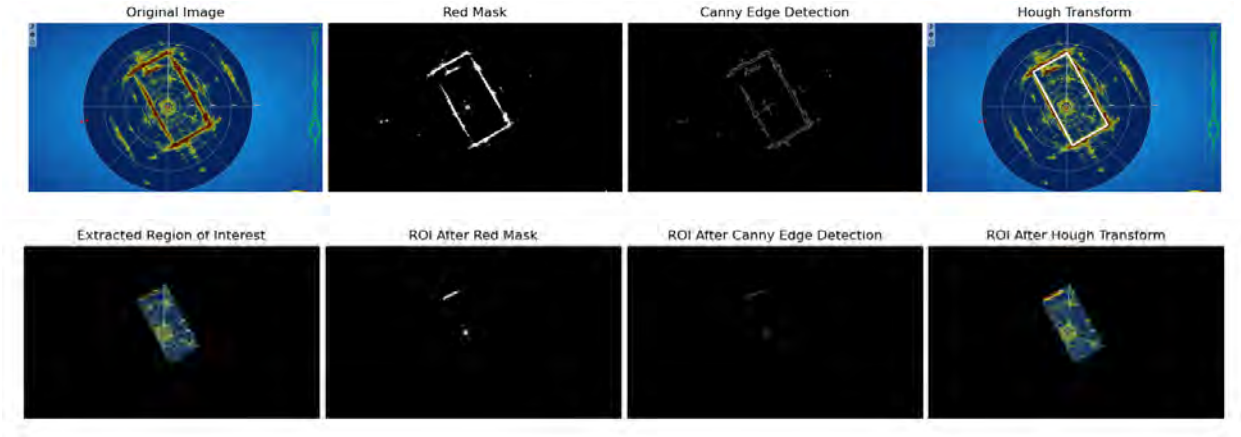


Fig. 17. Image processing method for pipeline detection.

With the pool boundary correctly segmented, we then focused on detecting the pipeline within the extracted region of interest (ROI). Within refined ROI, we re-applied color masking, edge detection, and Hough Transform to detect the pipeline with higher accuracy.

Finally, as shown in Fig. 18, we overlaid the detected pipeline and boundary on the original image to visualize the results. A white boundary was drawn to represent the detected pool edges, while the pipeline was marked in orange for clear identification.

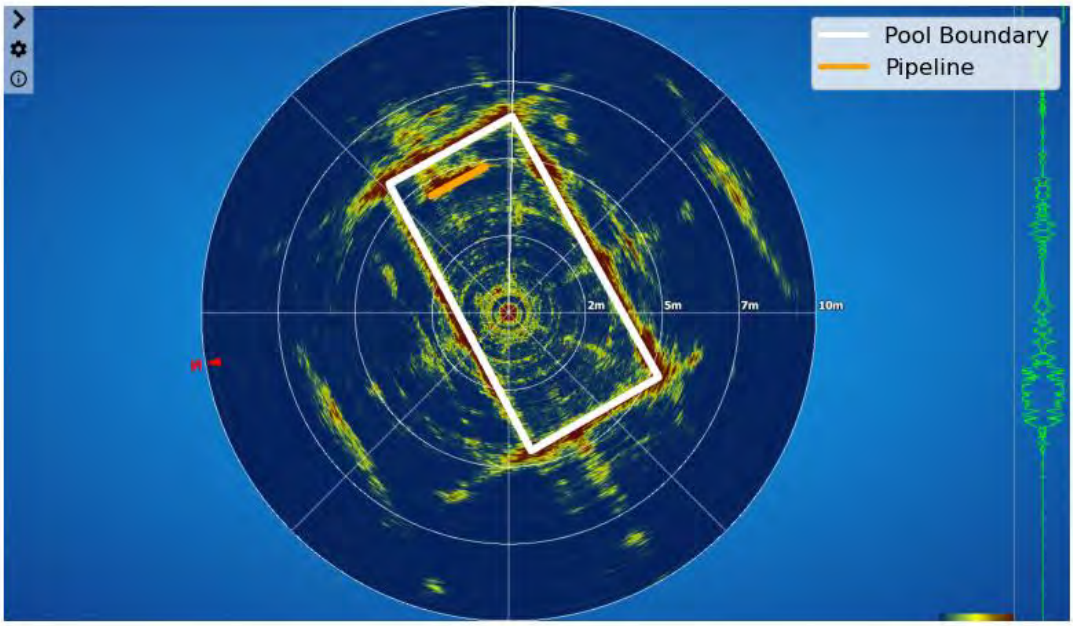


Fig. 18. Detected results from the scanning sonar image.

Figure.19 shows the process of autonomous approaching the pipeline through scanning sonar. To perceive the underwater environment, a Ping360 scanning sonar is used to collect 360-degree intensity data. Each scan angle returns a 1D intensity profile, which is filtered using a median filter and thresholder to reduce

noise and retain strong reflections. The peak intensity at each angle is used to compute the distance of the nearest object. These angle-distance pairs are then converted into Cartesian coordinates to create a 2D point cloud representing the surroundings. A custom (Derpanis, 2010) algorithm is applied to this point cloud to detect straight lines, which represent physical structures such as pool boundaries or pipelines. Lines that persist over multiple scans are promoted to "stable lines." Once five stable lines are detected, a classification step identifies four of them as pool boundaries and the remaining one as the pipeline. The intersection points of the boundary lines are used to compute the pool corners, and the pipeline centerline is extracted for navigation.

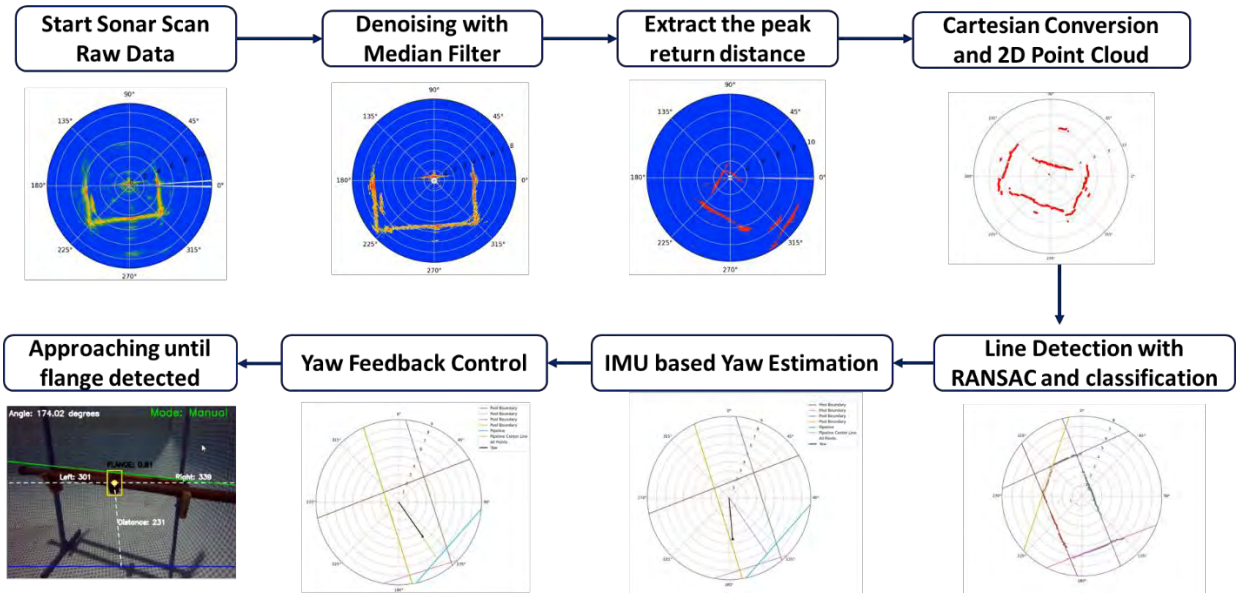


Fig. 19. Scanning sonar data processing steps.

With the pipeline centerline established, the ROV begins its control phase. The system continuously polls the current heading angle (*yaw*) from the onboard Mavlink telemetry. A yaw rectification step is first applied based on the orientation of the pool to ensure alignment with the sonar frame. The control logic calculates the yaw error as the difference between the ROV's current heading and the pipeline centerline direction. A Proportional-Derivative (PD) controller adjusts the ROV's yaw via RC channel override, steering it to align with the pipeline. Once the yaw error falls within a ±10° threshold, the ROV initiates forward movement toward the pipeline. Once the Red, Green, Blue(RGB) camera detects the flange, the thrust is stopped, completing the alignment and approach maneuver, and the control mode will switch to visual servoing through RGB camera.

We successfully integrated long-range sonar-based detection with short-range vision-based detection and grasping, achieving end-to-end autonomous operation for the first time. Due to the limited range of the vision system, we employed single-beam scanning sonar for long-distance detection. To enhance sonar point cloud clarity, we applied denoising and RANSAC algorithms. Simultaneously, the Inertial Measurement Unit (IMU) data was used to estimate the ROV's yaw angle. Upon detecting the pipeline, a path was generated toward the flange center, and a PD controller was used to align the yaw and approach the target. Once the vision system detected the pipeline and flange at close range, control was seamlessly switched to vision-based mode for precise grasping.

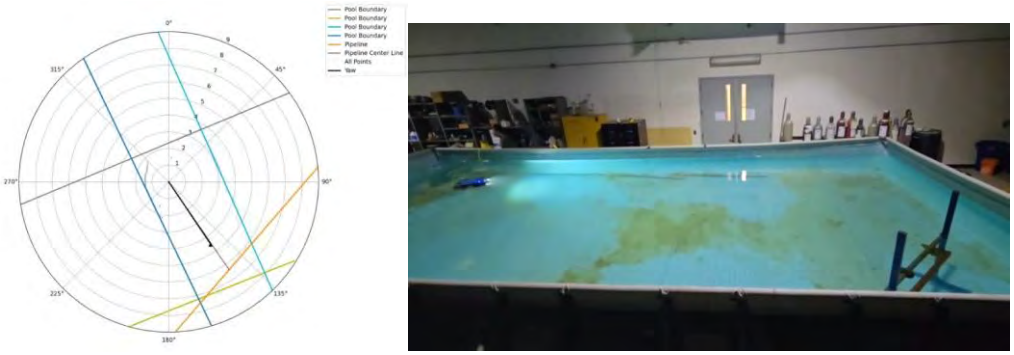


Fig. 20. a) Initial detection phase: Vision is limited, pipeline and boundary detection based on sonar.

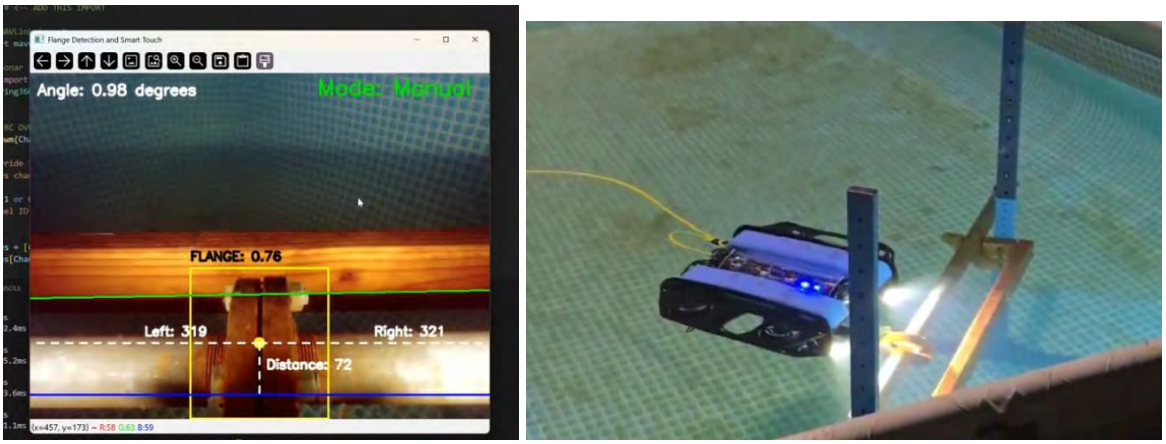


Fig. 20. b). Grasping phase: Once the flange is detected by the vision system, control switches to vision mode for fine alignment — centering, parallel orientation, and precise approach for grasping.

IV.6 ESKF based Sonar Map Reconstruction

In practical deployment scenarios, the ROV experiences continuous motion due to environmental disturbances such as water currents and tether-induced forces, which introduces significant distortion in the sonar scans. These distortions pose challenges for accurate localization and mapping, especially when using low-cost, low-frequency sonar sensors. We addressed the robust underwater localization and map reconstruction for AUVs by fusing sonar and inertial data, enabling reliable operation in GPS-denied and visually degraded environments. We used a tightly coupled, low-cost localization and mapping framework that integrates a single-beam scanning sonar and a low-grade IMU. As shown in Fig. 21, our method begins by feeding raw IMU data into an (Fossen, 2023) to obtain a coarse estimate of velocity and heading. This estimate is then used to compensate for the sonar motion distortion and pre-reconstruct a local sonar map. We then apply the Generalized ICP (GICP) (Segal, 2009) on the reconstructed map to refine the vehicle's position. The GICP output is iteratively fused back into the ESKF to improve estimation accuracy over time.

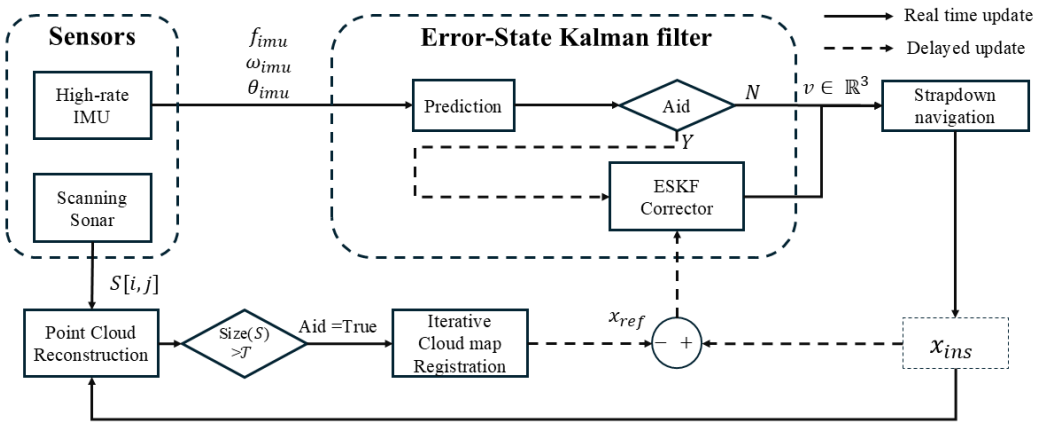
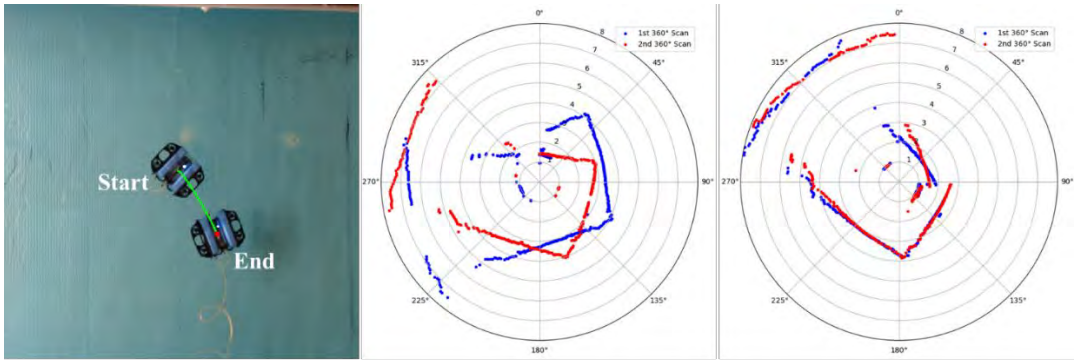


Fig. 21. Flowchart of proposed GICP aided ESK.

As shown in Fig. 22(a), the ROV does not remain stationary after being placed in the pool. Influenced by water currents and the tension of the tether, the ROV drifts from the initial to the final position. Since the single-beam scanning sonar has a low data sampling rate, it takes approximately 35 seconds to complete a full 360-degree scan. As illustrated in Fig. 22(b), this movement causes distortions in the sonar map, resulting in misalignment between successive scanning cycles. From Fig. 22(c), the effectiveness of the proposed GICP-aided ESKF can be observed. Although the ROV experiences drift in both yaw angle and position, the proposed iterative method effectively compensates for this motion. Leveraging the low-frequency but high-accuracy sonar data, the coarse position estimates provided by the ESKF are successfully refined, resulting in more accurate localization. Based on this improved estimation, the sonar map can be compensated and accurately reconstructed. While some residual drift remains due to IMU errors, the overall improvement in map reconstruction is clear.



(a) (b) (c)

Fig. 22. Sonar map reconstruction. (a). Top view of the ROV in a 10m × 5m pool, ROV moves due to the flow and cable's tension; (b). The sonar cloud point map without reconstruction; (c). The sonar cloud map after GCIP aided ESKF processing.

IV.7 Apriltag-Vision based ROV Station-Keeping Control

We successfully completed an end-to-end experiment demonstrating flange detection and grasping using the ROV. A key prerequisite for this experiment was maintaining the ROV in a stationary position, necessitated by the low sampling rate of the single-beam scanning sonar. To address this constraint, we initially implemented a vision-based self-stabilization approach using optical flow. However, the preliminary results fell short of expectations due to the sparse visual features inherent in the underwater environment. To overcome this limitation, we adopted Apriltag (Olson, 2011) markers as a more robust visual reference for stabilization. AprilTags are high-precision, fiducial marker systems (like QR codes) used in robotics, computer vision, and augmented reality to determine the precise 3D position, distance, and orientation of objects relative to a camera. They enable rapid, real-time detection of unique IDs and are widely used for robot localization and alignment. Our method still uses low-cost sensors, the overall system is described in Fig. 23, the ROV is equipped with a Newton Subsea Gripper from Blue Robotics. The end effector of the gripper has been redesigned and 3D-printed to integrate a pair of PZT transducers. The onboard sensors include a Ping360 single-beam scanning sonar for long-range sensing, an IMU for orientation estimation, and a monocular camera for station-keeping and vision-based control.

The detailed methodology and collaboration of sensors are illustrated in Fig. 23, and the process can be classified into three stages. In the first stage of long-range sensing, a Ping360 mechanical single-beam scanning sonar is employed to perceive the unknown environment. Since the sampling frequency is low (e.g., at a 50~m range it takes approximately 33~s for a 360° scan), disturbances such as tether tension and water currents may cause ROV drift; therefore, station-keeping is required to ensure high-quality sonar data. Vision-based control with AprilTags is used as a reference to stabilize the ROV in x-y plane and yaw direction before scanning. In the second stage, the raw sonar data are processed using Random Sample Consensus (RANSAC) with yaw compensation from the IMU to extract features and fit the pipeline and boundaries. Afterward, a path-planning algorithm generates a trajectory for the ROV to approach the pipeline. In the final stage, real-time vision is used for fine perception, where pipelines are identified using the Hough transform and flanges are detected by a YOLOv8 model. Once the flange is detected, the system switches to IBVS mode to perform flange smart touch inspection through centering, paralleling, and approaching.

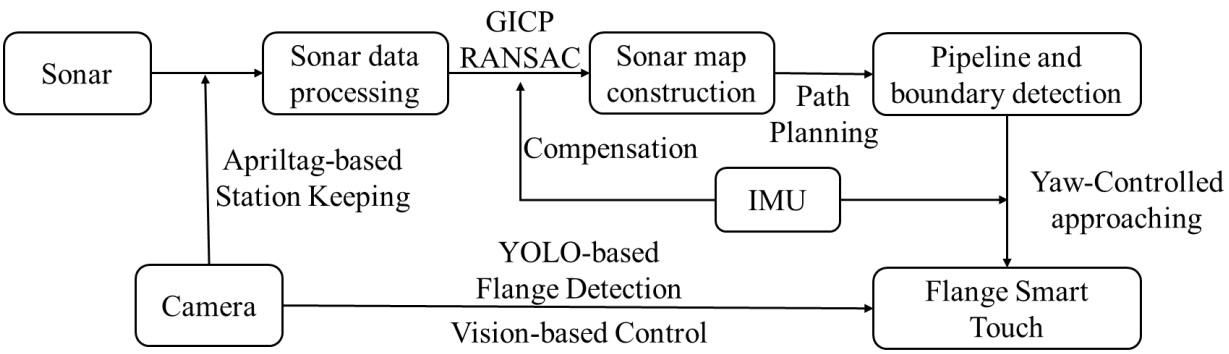
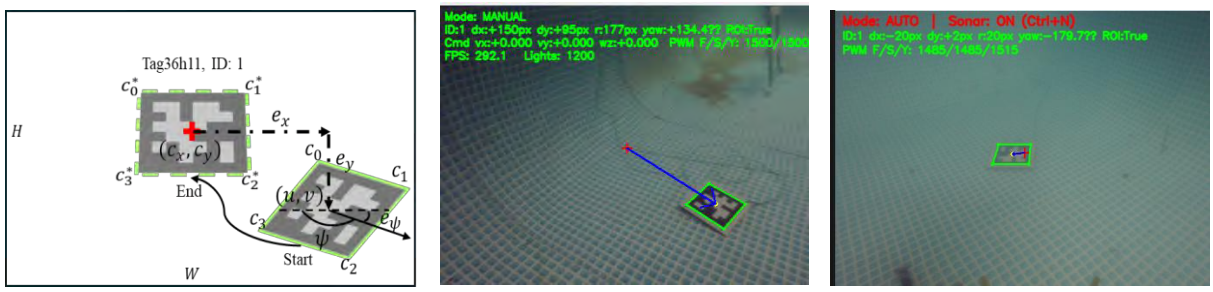


Fig. 23. Flowchart of the proposed autonomous method for pipeline sensing and flange smart touch inspection.

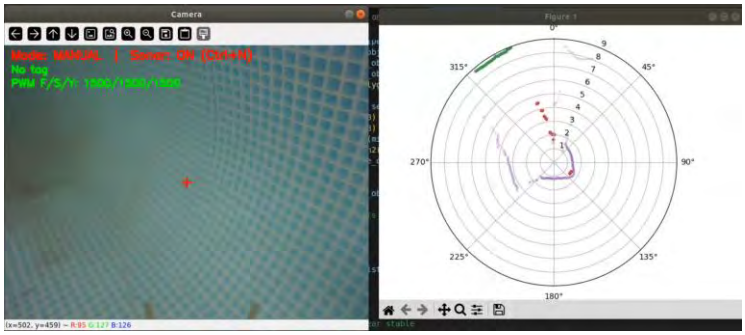
Specifically, for the station-keeping, as shown in Fig. 24, by actively controlling the ROV to keep the tag centered in the camera frame, we achieved significantly improved station-keeping performance.



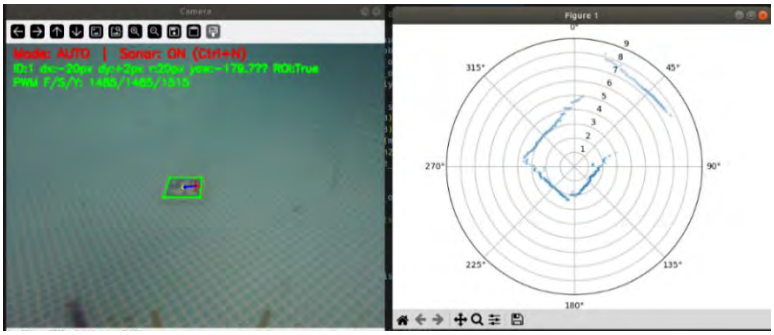
(a) (b) (c)

Fig. 24. Illustration of AprilTag-based station keeping for ROV. (a) The proposed method and notation of the errors, (b) Start of the station keeping, (c) During station keeping.

Fig. 25 demonstrates results of vision-based stabilization for the ROV. (a): Without stabilization, the ROV exhibits noticeable drift caused by cable tension and ambient current, resulting in a distorted and inconsistent sonar map. (b): With vision-based stabilization using Apriltag detection, the controller effectively maintains the tag at the center of the image frame. This leads to significantly improved station-keeping, as evidenced by a clearer and more consistent sonar map with well-defined boundaries. In summary, we preliminarily achieved the goal of end-to-end autonomous pipeline sensing, path tracking, and IBVS based control of ROV for flange smart touch.



(a)



(b)

Fig. 25. Station-Keeping performance comparison.

Case Study 1: Sonar map without stabilization

Fig. 26. Left: Top-down view of ROV motion without active stabilization. Right: Corresponding sonar mapping during free roaming for a swimming pool.

As illustrated in Fig.26, the ROV's uncontrolled roaming significantly impacts sonar map construction. Due to the low update frequency of the single-beam sonar, originally straight lines become distorted into curves, and successive scans exhibit noticeable inconsistency.

Case Study 2: Sonar map construction with stabilization

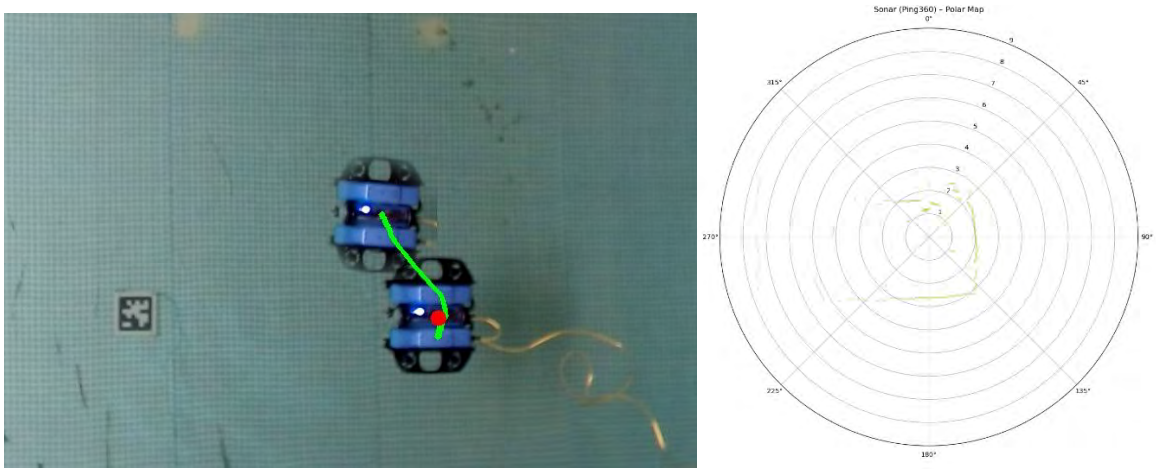


Fig. 27. Left: Top-down view of ROV motion with active stabilization. Right: Sonar mapping under stabilization.

From the top-down perspective in Fig. 27, it is evident that the visual servoing system actively guided the ROV's field of view toward the AprilTag. Upon completion of the alignment, the ROV maintained a station-keeping position to facilitate sonar map construction. The improvement is substantial: the pool's boundary lines now appear straight and well-defined, and each 360° scan yields consistent results. The error of convergence of the stabilization process is shown on Fig. 28, where it is evident that the controller effectively drives the system toward steady-state accuracy.

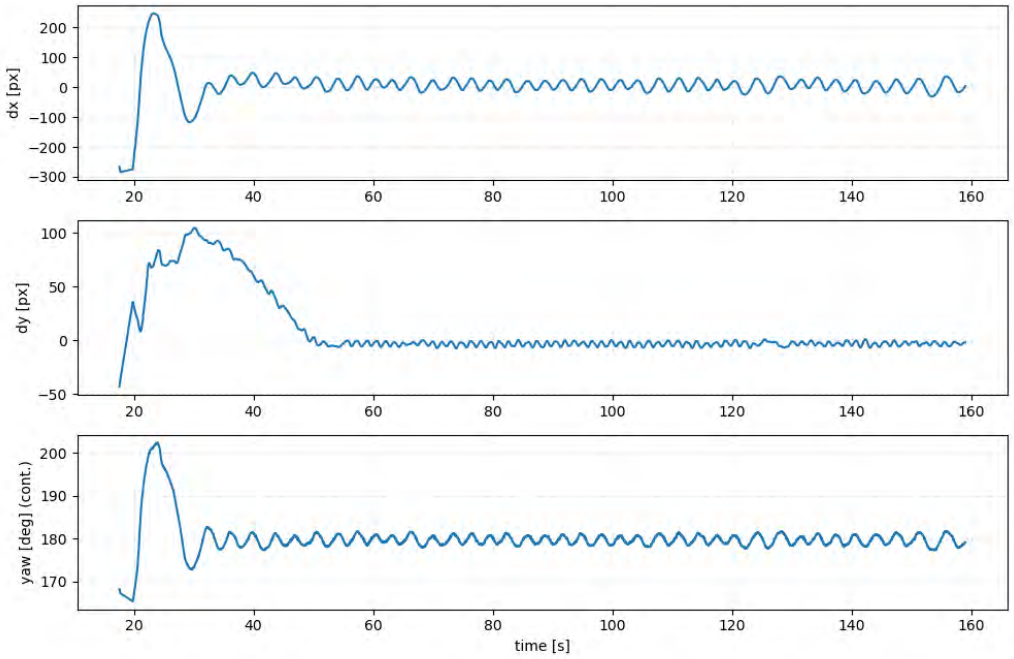


Fig. 28. Error convergence over time during stabilization.

Case Study 3: Sonar map construction with stabilization under outside disturbance

To further validate the performance of the proposed method, we conducted stabilization experiments under external disturbances. As shown in Fig. 29, a water jet was used to generate flow disturbances from top to bottom. Despite the presence of these currents, the ROV successfully maintained station-keeping, demonstrating robust stabilization against directional water flow. As shown in Fig. 30, the steady-state oscillations increase under external disturbance; however, the error of convergence still reflects the robustness of the visual servoing system.

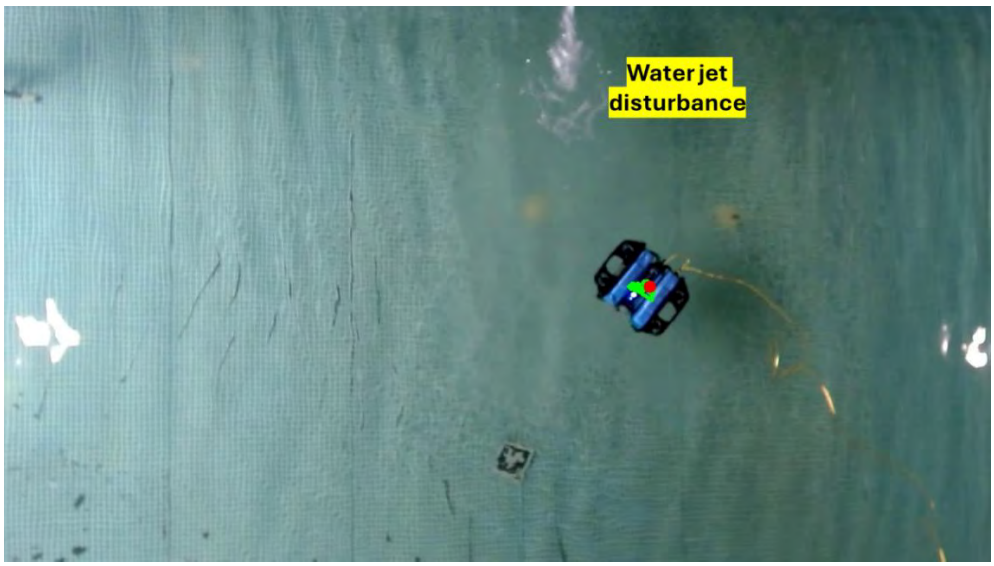


Fig. 29. Top-down trajectory of ROV during stabilization under external disturbance.

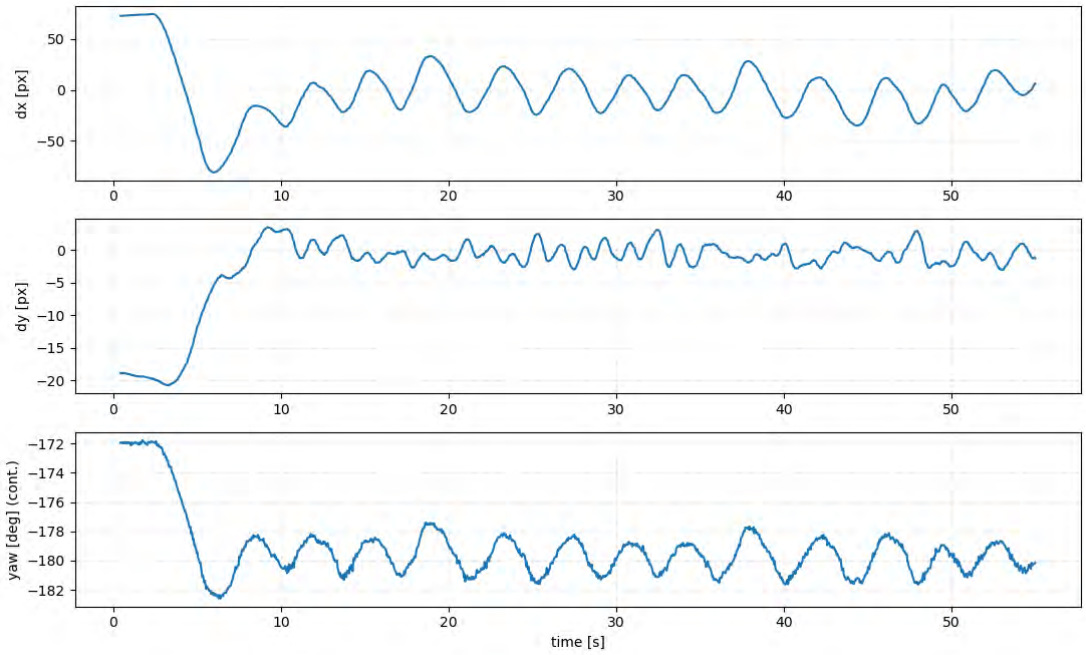


Fig. 30. Error convergence over time during stabilization under external disturbance.

To validate the full autonomous framework, an integrated experiment was conducted combining sonar-based long-range perception, path planning, and vision-based fine manipulation. The objective was to demonstrate that the ROV can autonomously detect the pipeline from sonar data, plan and execute an approach trajectory, switch to vision mode, and perform flange grasping and smart-touch inspection using PZT transducers.

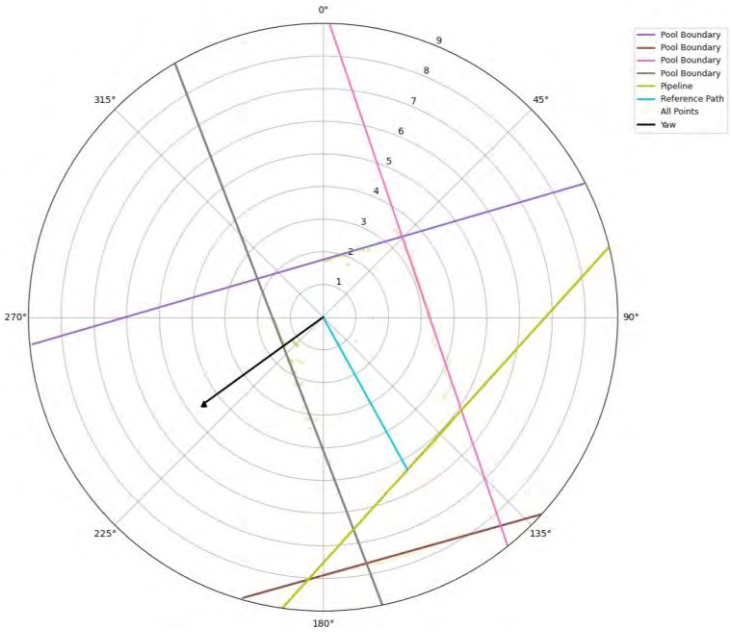


Fig. 31. Results of sonar map construction, pipeline and pool boundaries detection, path planning, and Yaw angle estimation.

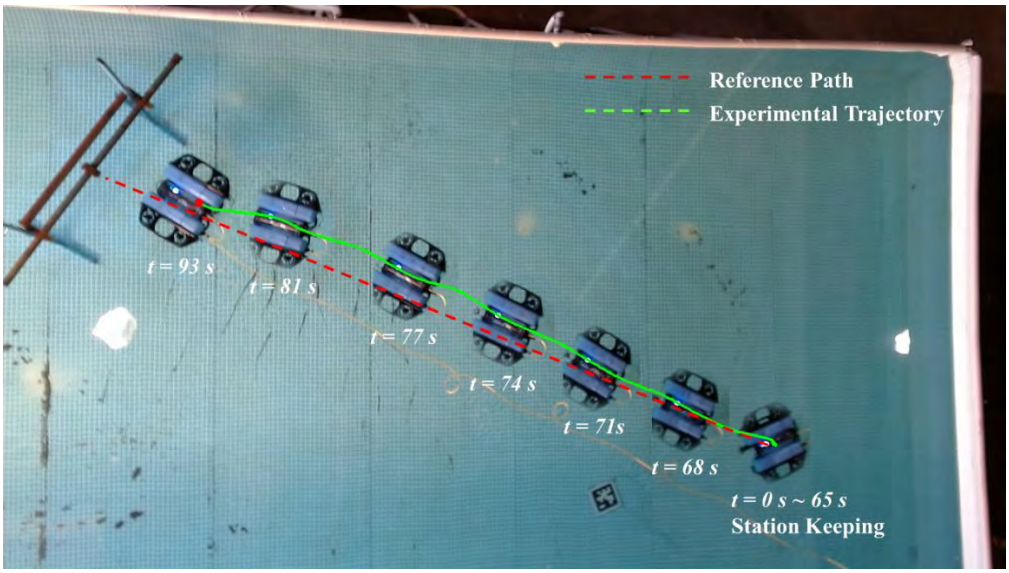


Fig. 32. ROV trajectory during sonar-based path tracking. The red line denotes the reference path toward the pipeline midpoint, and the green line shows the actual ROV trajectory.

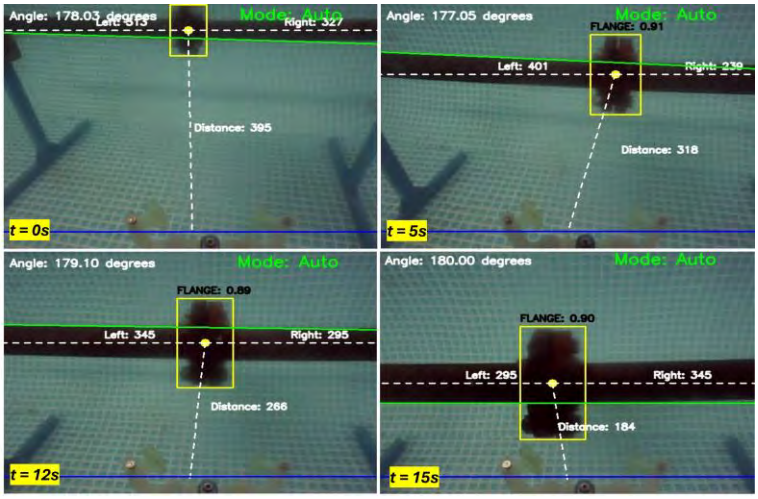


Fig. 33. Results of IBVS-based visual control for flange grasping. The ROV performs centering, alignment, and approach motions for autonomous flange smart-touch inspection.

IV.8 Electrolysis-based Buoyancy Control Devices (BCDs) with ROV for Depth Control

We connected the four BCDs (Keow, 2020) in parallel and applied 3V at a total current of 3.6A for open-loop control. As shown in Fig. 34, the balloons successfully inflated, and the four BCDs collectively lifted the system, achieving near-neutral buoyancy. However, we observed inconsistencies in the gas generation rates among the BCDs and identified minor leakage issues.

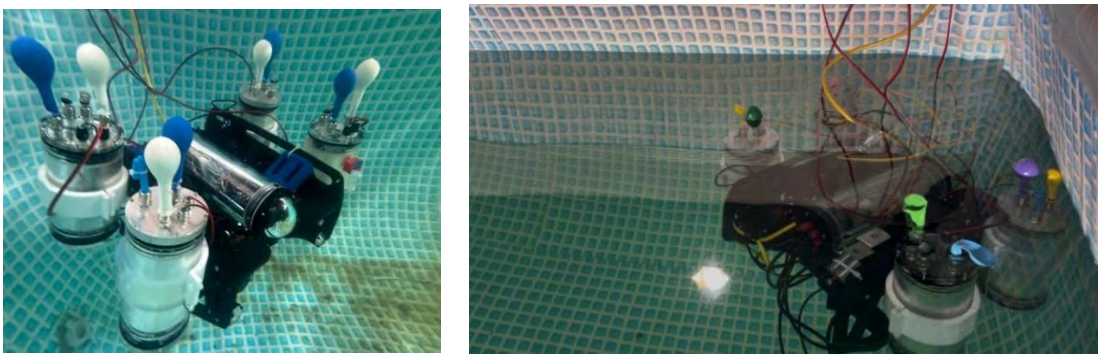


Fig. 34. Open loop control of 4 BCDs with ROV.

Building on these hardware improvements, preliminary open-loop trials of the hybrid actuation system have been conducted in a pool environment. As illustrated in the accompanying figures, these initial results successfully demonstrate the system's vertical transit capability, showing the BCD's ability to reliably move the BlueROV from a submerged position to the pool surface. This outcome confirms the foundational feasibility of using the BCD for controlled buoyancy adjustment and marks a significant operational milestone.

Initial tests have successfully characterized the system's performance, including the identification of minor fluid control challenges within the BCD tanks and a need for enhanced integrity in the ROV's propulsion housing. These findings are crucial for system optimization and highlight the path toward achieving reliable and stable buoyancy control. Furthermore, the analysis of the electrolyzes has revealed a minor asymmetrical performance characteristic, which has already been addressed at the design level. The immediate focus of the project is to implement corrective measures to ensure optimal system performance. This includes resolving the minor fluid control challenges in the BCD tanks and reinforcing the ROV's propulsion housing to prevent fluid ingress. Building on these hardware improvements, preliminary open-loop trials of the hybrid actuation system were conducted in a pool environment. These initial results successfully demonstrate the system's vertical transit capability, showing the BCD's ability to reliably move the BlueROV from a submerged position to the pool surface. This outcome confirms the foundational feasibility of using the BCD for controlled buoyancy adjustment and marks a significant operational milestone.

Concurrently, the primary objective is the development, refinement, and implementation of control logic for the hybrid system. The goal is to fully populate the designed PCB and integrate it with the ESP32 and the four new depth sensors (see Figure 35, illustrating the sensor's placement). This implementation directly addresses the earlier observation of asymmetrical electrolyzer performance by enabling independent and precise control over each BCD unit.

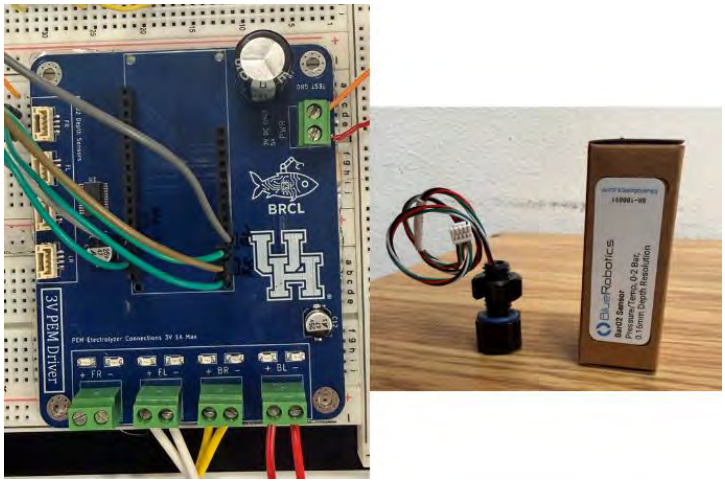


Figure 35. PCB design and Depth sensor for BCD system.

IV.9 Received Signal Strength (RSS) Sensor based Localization of ROV and Flange

The developed system enables localization of an ROV or underwater target flange using a Received Signal Strength (RSS)–based sensing approach (Nguyen & Shin, 2019; Pourkabirian et al., 2023; Wielandner et al., 2021). A central transmitter emits a 433 MHz RF signal (Ryecroft, et al., 2019)Click or tap here to enter text., which is received by an array of four spatially distributed antennas positioned orthogonally in a waterproof enclosure. By measuring the relative signal strength across these antennas, the device infers directional information, enabling coarse bearing estimation between the transmitter and receiver in underwater or air environments.

The antenna array and associated electronics were designed with robustness and modularity in mind. A custom PCB integrates four CC1101 transceivers, an ESP32 microcontroller, and supporting passive components. The layout ensures symmetric antenna spacing for angular resolution and signal fidelity. The electronics are enclosed within a waterproof housing rated for submerged operation, with attention to RF-transparent windowing and cable feedthroughs. The PCB was fabricated, assembled, and successfully tested using air-based trials, as shown in Fig. 36 and Fig. 37.

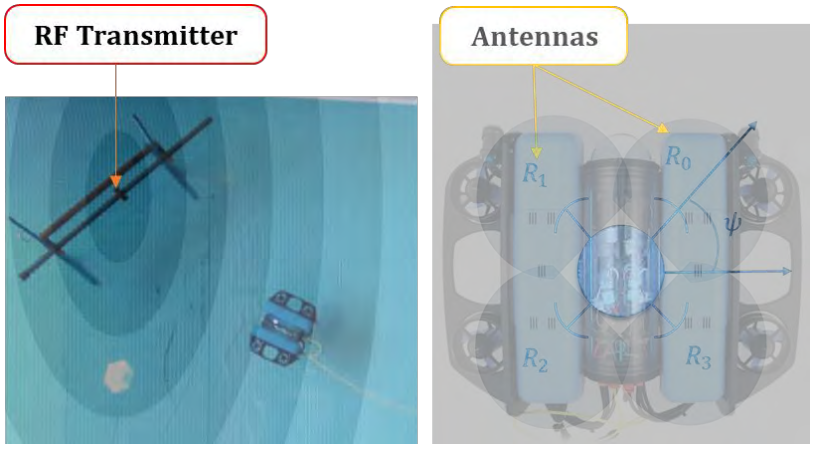


Fig. 36. Illustration of RSS based ROV and flange localization.

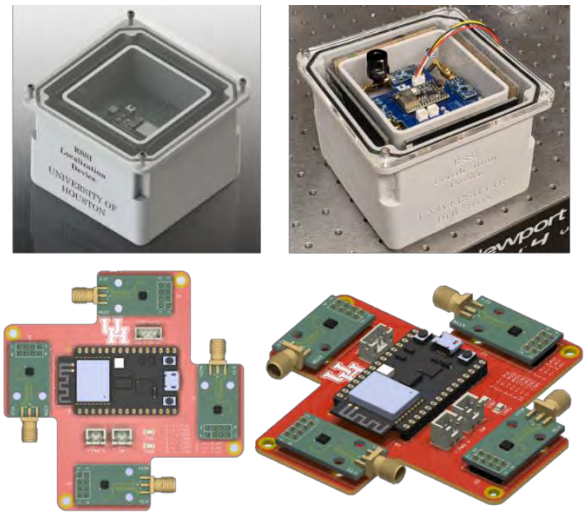


Fig. 37. PCB and waterproof antenna box design.

Initial testing was conducted in air to validate directional sensitivity and signal quality. Fig. 38 presents RSSI measurements as a function of angle for different antenna elements at fixed distances. The plots demonstrate distinguishable variation in signal strength with respect to angular orientation, supporting the viability of RSS-based localization. While these results are preliminary, they confirm that the relative RSSI values from the four receivers encode directional information, laying out the foundation for further underwater experiments and signal fusion strategies.

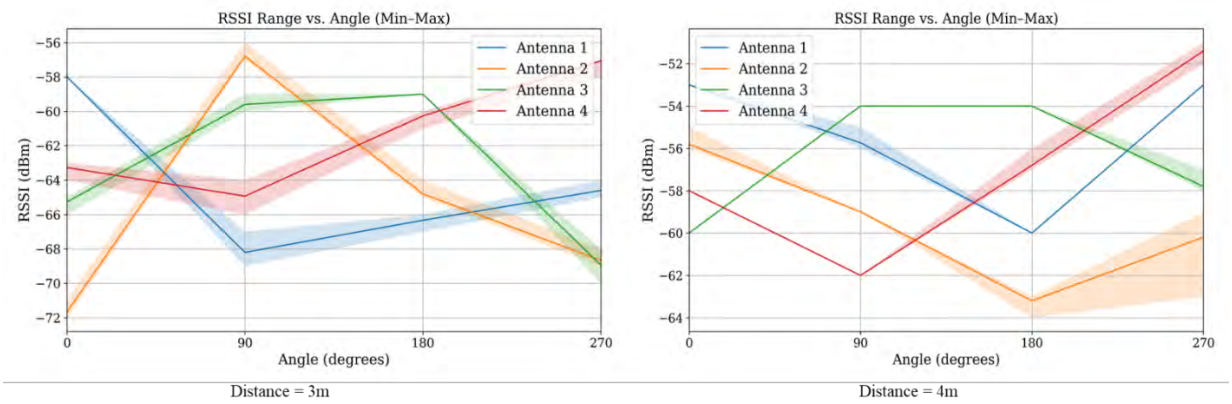


Fig. 38. Signal strength vs angle in air.

Following successful air-based validation, the system was deployed in a submerged environment to evaluate its directional sensing performance under realistic aquatic conditions. Received Signal Strength (RSS) measurements were collected at a fixed distance while varying the angle between the transmitter and receiver across four orthogonally placed antennas.

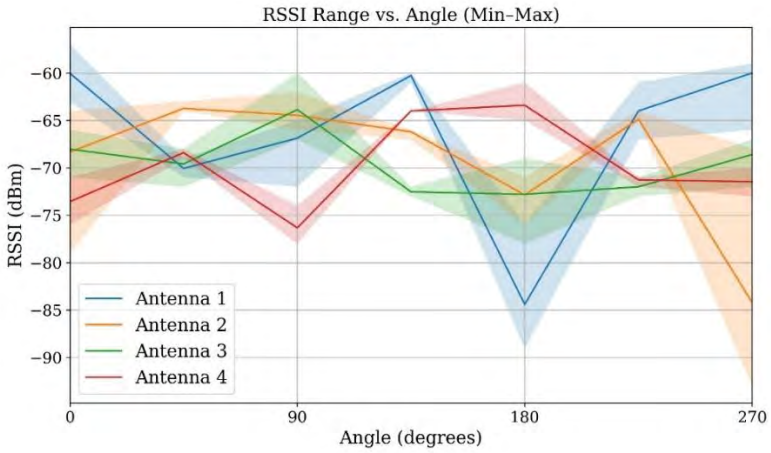


Fig. 39. Measured RSSI range (min–max) vs. angle for all four antennas during underwater testing at a fixed distance of 0.5m. The shaded regions represent the variation across repeated samples.

Fig. 39 illustrates the RSSI response versus angle for each antenna, including the minimum and maximum observed values across repeated samples. Notably, the antenna responses exhibit distinct, angle-dependent variations, confirming that relative RSSI differences persist underwater. However, the data also shows increased variability, likely due to environmental factors such as multipath reflections, absorption by the medium, and the effects of the waterproof housing. These results confirm that angular information remains embedded in the RSSI patterns, although with reduced separation between angle-dependent responses compared to air. This further supports the viability of signal-based bearing estimation, with the caveat that more robust models may be required to handle underwater noise.

Based on the underwater RSSI data, a regression model was developed to estimate bearing angle from the received signal strengths. Despite the structured antenna layout and clear signal variation across angles, the model's performance was limited, showing a tendency to bias toward central values and reduced sensitivity to subtle directional changes. This is attributed to the sparsity of training data, limited angular resolution, and environmental variability inherent in underwater conditions. To address this, additional data collection was carried out in a larger water tank with finer angular sampling.

To enhance the reliability and volume of training data for the RSSI-based localization system, a new experimental framework has been devised. The RSSI receiver device is mounted beneath an ROV, with an AprilTag affixed to the top of the vehicle to enable precise visual tracking. The experiment is conducted in a controlled water tank environment, where an overhead camera system is capturing the AprilTag and computing the ROV’s 2D position and orientation using a calibrated OpenCV-based AprilTag detection algorithm. Simultaneously, RSSI data is streamed from the underwater receiver via a serial interface and logged into a CSV file for post-processing. This synchronized collection of spatial and signal data provides a more robust dataset for training and evaluating machine learning models tasked with inferring bearing estimates from RSSI measurements. An example frame from the top-down visual tracking system is shown in Fig. 40.

Fig. 40. Top camera view showing the estimated position of the AprilTag.

V. FLANGE BOLT LOOSENESS MONITORING WITH MACHINE LEARNING MODEL

V.1 Experimental Setup

The experimental setup is shown in Fig. 41. It consists of a DAQ system (PXI-1042, National Instruments), an amplifier, the robotic gripper (with its associated microcontroller), and a pair of PZT (Lead Zirconate Titanate) transducers. For the 6-inch flange, we used a Gaussian pulse excitation signal, with a central frequency of 6 kHz and a duration of 0.001 seconds, which was generated at a 500 kHz sampling rate by the DAQ system. For the remaining flanges, a chirp signal was used as the excitation input, with an initial frequency of 1 kHz, a final frequency of 60 kHz, a duration of 500 milliseconds, and a sampling rate of 1 MHz. This signal was amplified to drive the PZT transducer (actuator), and the opposing transducer (sensor) recorded the response signal (acoustic signal) as it propagated through the bolted flange. Fig. 42 illustrates the schematic diagram of acoustic signal propagation.

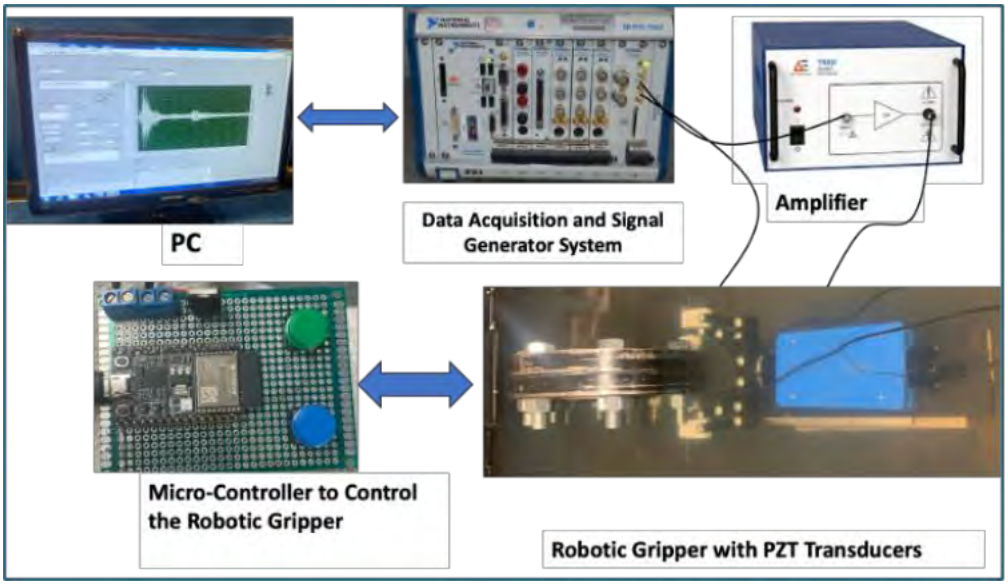


Fig. 41. Underwater Experimental Setup with Robotic Gripper.

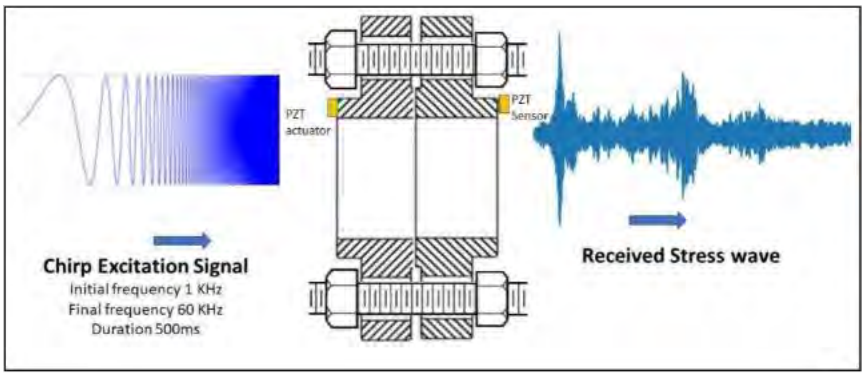


Fig. 42. Schematic Diagram of the Acoustic Signal Propagation. The Left PZT Transducer (Actuator) Outputs the Excitation Signal and the Right PZT Transducer (Sensor) Collects the Response Signal (Acoustic Signal).

V.2 Feature Extraction: Mel-frequency Cepstral Coefficient (MFCC)

We adopted the MFCC as a feature extraction method in this part, and the MFCC extraction processing is shown in Fig. 43. First, according to the specific frame length and step, the signal is split into frames. Next, each frame is transferred into a magnitude spectrum by Discrete Fourier Transform (DFT). Afterwards, passing the DFT result of each frame through a predefined Mel-filter bank, Mel-spectrum (R×M) is obtained. Finally, the MFCC matrix (R×T) can be obtained by applying the discrete cosine transform to the Mel-spectrum.

In our experiments, the MFCC features were extracted from the original signals using the following configuration: a sampling frequency of 1 MHz, a window length of 0.02 seconds with an overlap of 0.015 seconds, and a Hamming window function. A total of 140 cepstral coefficients were computed using 150 Mel filter banks.

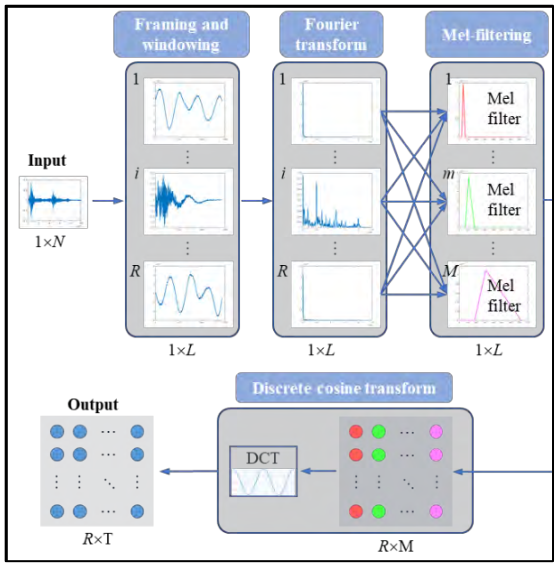


Fig. 43. The Flow Chart of MFCC Processing.

V.3 Classification Method: Support Vector Machine (SVM)

We used SVM as the classification model in this part, and the objective is to minimize the following cost function:

$$\frac{1}{2}\|\omega\|^2 + C\sum_{i}^{N} max\left(0,1 - y_i(\omega^T x_i + b)\right)$$

where, $\frac{1}{2}\|\omega\|^2$ is the regularization, and

$$C\sum_{i}^{N} max\left(0,1 - y_i(\omega^T x_i + b)\right)$$

is the hinge loss.

Our SVM model utilizes the Radial Basis Function (RBF)/Gaussian kernel, which is effective for handling non-linear relationships in the data. The specific parameters for the SVM were fine-tuned as follows:

Kernel: Radial Basis Function (RBF)/Gaussian.

C = 10

Gamma/Kernel Scale = 'auto' (use a heuristic procedure to search the optimal value)

These parameters were adjusted to optimize the classification performance of SVM on validation sets, balancing the trade-off between margin maximization and classification error minimization.

V.4 Dataset Collection, Arrangement and Results

V.4.1 Introduction

In this section, we introduce the dataset collection procedures employed in this work. Specifically, two approaches were utilized: uniform load distribution and non-uniform load distribution. The details of each procedure are provided below.

V.4.1.1 Uniform Load Distribution

In the uniform load distribution approach, a single torque level is applied uniformly to all bolts on the flange. Data are then collected from all bolts at that load level. The torque is subsequently increased to the next level, and the process is repeated until reaching the maximum torque supported by the bolts. Each full progression through the torque levels constitutes one dataset. Multiple such datasets are collected, with several used for training the model and others reserved for independent testing. Fig. 44 illustrates an example of the uniform load distribution procedure for the 9-inch flange.

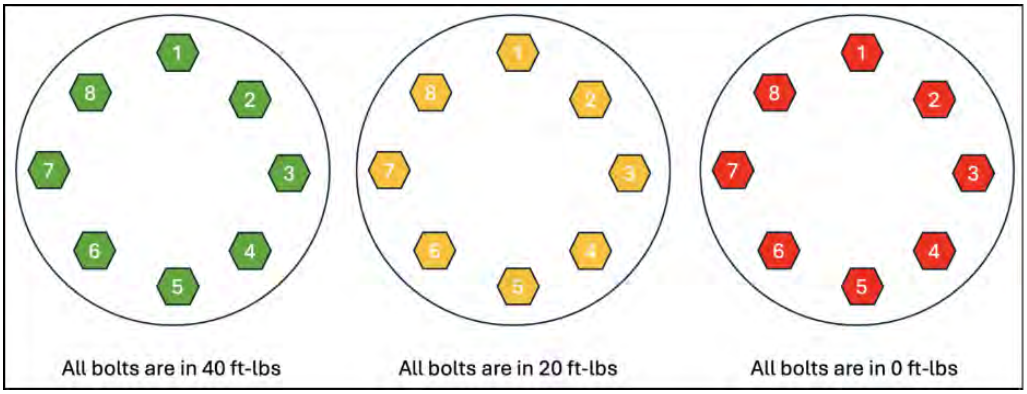


Fig. 44. Uniform Load Distribution for the 9-Inch Flange.

V.4.1.2 Non-Uniform Load Distribution

In the non-uniform load distribution approach, the datasets collected under uniform loading conditions are used to train the model. After developing the SVM model, a selected number of bolts are intentionally loosened while the remaining bolts are kept fully tightened. Data collected under these conditions form the testing set. The developed model is then evaluated across various scenarios involving different combinations of loosened bolts. Fig. 45 presents an example of a non-uniform load distribution for the 9-inch flange.

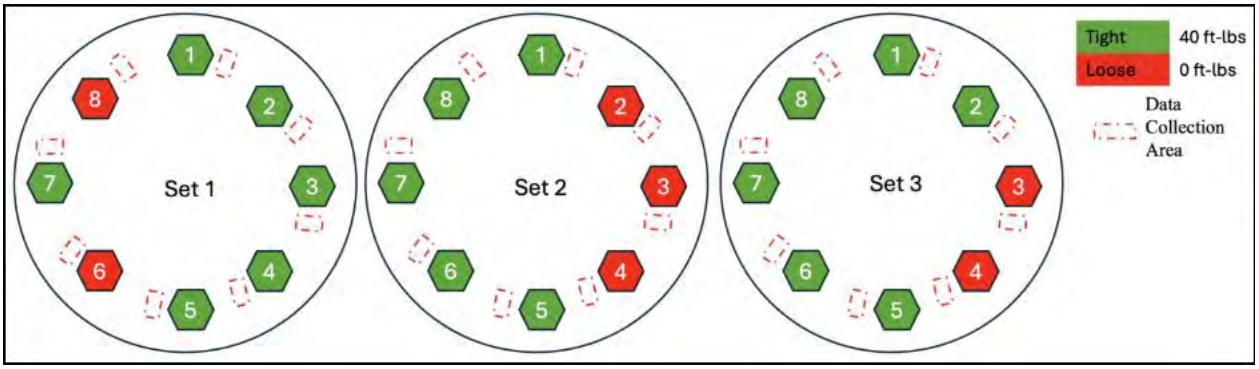


Fig. 45. 9-Inch Flange Joint in Non-uniform Load Distribution Scenario.

The general methodology used for classification is shown in Fig. 46. During the developing stage of the SVM model, datasets collected under the uniformly tightened condition were used to extract signal features through the Mel-Frequency Cepstral Coefficients (MFCCs). These features were then used to develop an SVM model. In the testing stage, the trained models were evaluated on datasets collected under the non-uniform load distribution scenario, where certain bolts were intentionally loosened while others were fully tightened. Signal measurements were obtained near the bolts, converted into MFCC features, and classified by the trained model.

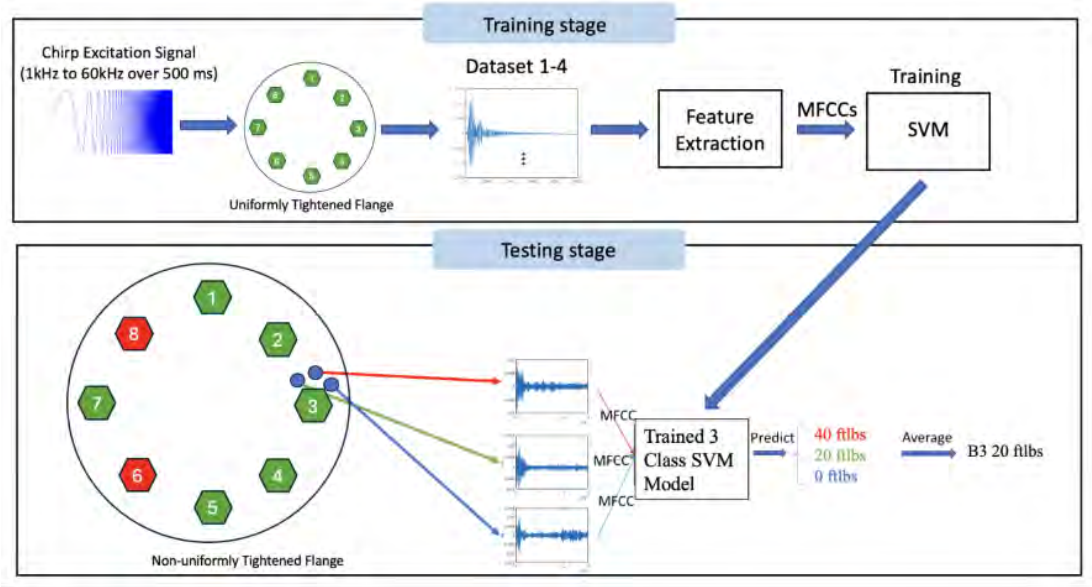


Fig. 46. Example of roughly quantifying the area load next to bolts.

V.4.2 6-inch flange

Fig. 47. 6-inch flange.

Uniform load

The maximum load level for the 6-inch flange is 40 ft-lbs. We collected 10 independent data sets. Each set consists of response signals recorded under four different classes, as indicated in Table 1: Please note that the 10 ft-lb and 20 ft-lb torque levels are combined into a single class in this study. Set 2 contains fewer samples (16 per class) as it was collected during the early experimental stage and was primarily used for preliminary evaluation of data quality. In contrast, sets 8 and 9 include a larger number of samples (120 per class) to increase the dataset size and improve classification accuracy and model reliability.

Table 1. Dataset Arrangement for the 6-inch flange.

Dataset ID	Loose (0 ft-lb, 0%)	Caution (10 & 20 ft-lbs, 25–50%)	Warning (30 ft-lbs, 75%)	Tight (40 ft-lbs, 100%)	Total Number
Set 1	40	80	40	40	200
Set 2	16	32	16	16	80
Set 3	40	80	40	40	200
Set 4	40	80	40	40	200
Set 5	40	80	40	40	200
Set 6	40	80	40	40	200
Set 7	40	80	40	40	200
Set 8	120	240	120	120	600
Set 9	120	240	120	120	600
Set 10	40	80	40	40	200
Total	N/A	N/A	N/A	N/A	**2680**

Table 2 represents the 4-class classification results of the datasets shown in Table 1

Table 2. 4-class Cross-set Validation Results of the 6-Inch Flange.

Training Set	Validation Set	Test Set	Accuracy on test set
Remaining 8 datasets	Dataset 3	Dataset 1	96.5%
Remaining 8 datasets	Dataset 5	Dataset 2	98.75%
Remaining 8 datasets	Dataset 7	Dataset 3	97.5%
Remaining 8 datasets	Dataset 9	Dataset 4	83%
Remaining 8 datasets	Dataset 3	Dataset 5	92%
Remaining 8 datasets	Dataset 6	Dataset 6	96%
Remaining 8 datasets	Dataset 8	Dataset 7	97%
Remaining 8 datasets	Dataset 2	Dataset 8	99.33%
Remaining 8 datasets	Dataset 1	Dataset 9	76.17%
Remaining 8 datasets	Dataset 8	Dataset 10	97.5%
Average	N/A	N/A	**93.38%**

The average accuracy of the 4-class classification is 93.38%, which proves that the model can accurately predict the looseness level of the flange. The confusion matrices of cross-set validation are presented in Fig. A1 in Appendix I. The precision, recall, and F1-score are included in Fig. A2 in Appendix I. In addition, some MFCC matrices are selected and visualized in Fig. A3 in Appendix I.

To reproduce the experimental results, please find the folder '6_inch_underwater_uniform_load_experiment' and follow the readme file to run the script. After running the script, you will get the confusion matrices, classification metrics (precision, recall, and F1-score), and MFCC figures in Appendix I.

V.4.3 9-inch flange

Fig. 48. 9-inch Flange

Uniform load

We collected 3 new independent datasets (Table 3) on the 9-inch flange. Each set consists of response signals recorded under four different flange preload levels.

Table 3. Dataset Arrangement with Robotic Gripper Underwater.

Dataset ID	Loose (0 ft-lb, 0%)	Caution (10 & 20 ft-lbs, 25–50%)	Warning (30 ft-lbs, 75%)	Tight (40 ft-lbs, 100%)	Total Number
Set 1	40	80	40	40	200
Set 2	40	80	40	40	200
Set 3	40	80	40	40	200
Total	N/A	N/A	N/A	N/A	**600**

Table 4 represents the 4-class classification results of the datasets shown in Table 3.

Table 4. 4-class Cross-set Validation Results of the 9-Inch Flange.

Training Set	Validation Set	Test Set	Accuracy on test set
80% of (Dataset 3,2)	20% of (Dataset 3,2)	Dataset 1	94.5%
80% of (Dataset 1,3)	20% of (Dataset 1,3)	Dataset 2	92%
80% of (Dataset 1,2)	20% of (Dataset 1,2)	Dataset 3	99%
Average			95.2%

The average accuracy of the 4-class classification is 95.2%, which proves that the model can accurately predict the looseness level of the flange. The confusion matrices of cross-set validation are presented in Fig. A4 in Appendix I. The precision, recall, and F1-score are included in Fig. A5 in Appendix I. In addition, some MFCC matrices are selected and visualized in Fig. A6 in Appendix I.

To reproduce the experimental results, please find the folder '9_inch_underwater_uniform_load_experiment' and follow the readme file to run the script. After running the script, you will get the confusion matrices, classification metrics (precision, recall, and F1-score), and MFCC figures in Appendix I.

Non-uniform load

In this section, we conducted experiments under a non-uniform load distribution scenario for the 9-inch flange joints. In this setup, certain bolts were intentionally kept fully tightened, while others were left completely loosened, thereby creating an uneven stress distribution around the flange. We collected 3 independent datasets, with torque levels of Tight (40 ft-lbs) and Loose (0 ft-lbs), as shown in Fig. 49. Signal measurements were taken from regions adjacent to the bolts, as indicated in the figure.

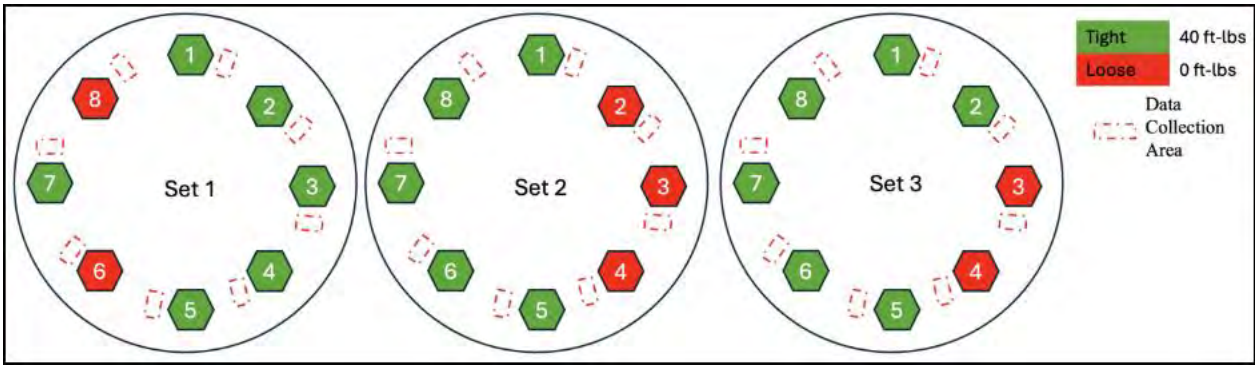


Fig. 49. 9-Inch Flange Joint in Non-uniform Load Distribution Scenario

For the 9-inch flange joint, three non-uniform datasets were tested. As shown in Fig. 50, the ground truth configurations included bolts at different torque levels: Tight (40 ft-lbs.), Warning (26 ft-lbs.), Caution (20 ft-lbs.), and Loose (0 ft-lbs.). The prediction results demonstrate that the trained SVM model was able to reasonably distinguish between the torque categories, although some bolts were occasionally misclassified. Despite these challenges, the model captured the general torque distribution pattern around the flange with good agreement between the predicted and actual bolt states.

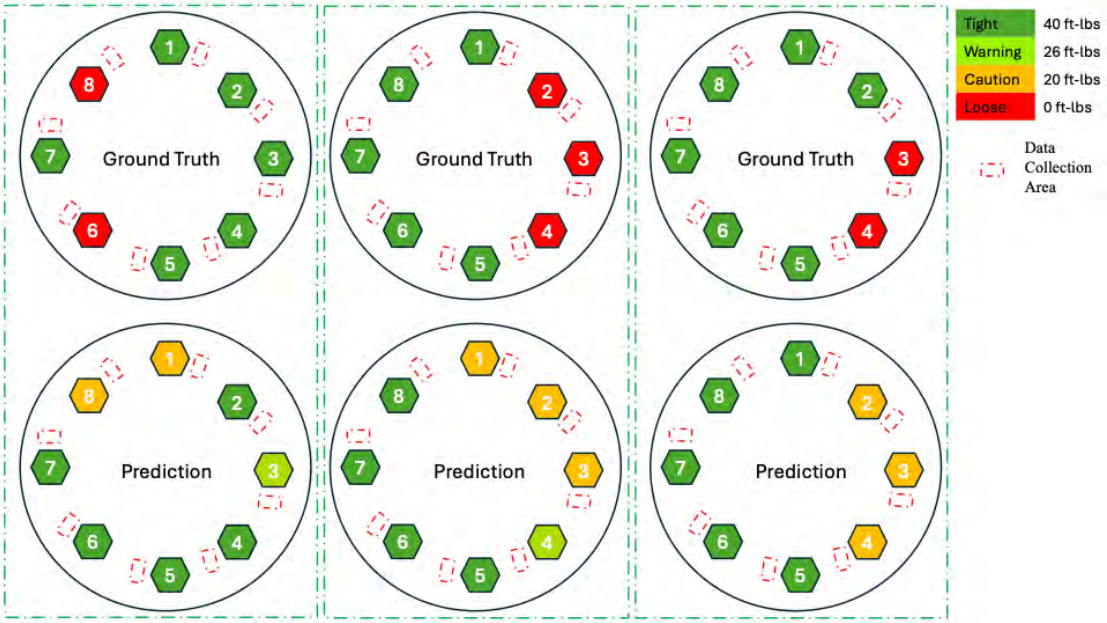


Fig. 50. Prediction results of the 9-inch flange.

V.4.4 11-inch flange

Fig. 51. 11-inch Flange.

Uniform load

For the 11-inch flange, the maximum torque level is 40 ft-lbs. We collected 3 independent datasets (Table 5) from 11-inch flanges. Each set consists of response signals recorded under three different flange preload levels (0, 20, and 40 ft-lbs).

Table 5. Dataset Arrangement of the 11-inch flange.

Dataset ID	Loose (0 ft-lb, 0%)	Warning (20 ft-lbs, 50%)	Tight (40 ft-lbs, 100%)	Total Number
Set 1	40	40	40	120
Set 2	40	40	40	120
Set 3	40	40	40	120
Total	N/A	N/A	N/A	360

The model is trained and validated on 2 different sets and finally tested on a completely unseen test set. Specifically, the parameters of SVM were adjusted based on the accuracy on validation sets. Table 6 represents the 3-class classification results of the datasets shown in Table 5.

Table 6. 3-Class Cross-set Validation Results of the 11-Inch Flange.

Training Set	Validation Set	Test Set	Accuracy on test set
80% of (Dataset 3,2)	20% of (Dataset 3,2)	Dataset 1	85.8.%
80% of (Dataset 1,3)	20% of (Dataset 1,3)	Dataset 2	93.3%
80% of (Dataset 1,2)	20% of (Dataset 1,2)	Dataset 3	95.83%
Average	N/A	N/A	91.7%

The confusion matrices of unseen test sets are shown in the appendix. In the 3-class task, the average classification accuracy is 91.7%. These classification results are promising. The confusion matrices of cross-set validation are presented in Fig. A7 in Appendix I. The precision, recall, and F1-score are included in Fig. A8 in Appendix I. In addition, some MFCC matrices are selected and visualized in Fig. A9 in Appendix I.

To reproduce the experimental results, please find the folder '11_inch_underwater_uniform_load_experiment' and follow the readme file to run the script. After running the script, you will get the confusion matrices, classification metrics (precision, recall, and F1-score), and MFCC figures in Appendix I.

V.4.5 13-inch flange

Fig. 52. 13-inch Flange

Uniform load

For the 13-inch flange, the maximum torque level is 110 ft-lbs. We have collected 2 data sets, each set having 3 classes (0 ft-lbs., 70 ft-lbs. and 110 ft-lbs). Table 7 represents the data distribution for the 13-inch Flange

Table 7. Dataset Arrangement of the 13-inch flange.

Dataset ID	Loose (0 ft-lb, 0%)	Warning (70 ft-lbs, 50%)	Tight (110 ft-lbs, 100%)	Total Number
Set 1	40	40	40	120
Set 2	80	80	80	240
Total	N/A	N/A	N/A	360

For the 13-inch flange joint, we use one of the two sets as the test set and the other one as developing and validation sets. In this training/test setting, the training and test sets are independent of each other. The model is trained and validated on different portions (90% and 10%) of a set from the two datasets and finally tested on the other one. The results of the classification are shown in Table 8.

Table 8. 3-Class Cross-set Validation Results of the 13-Inch Flange.

Training Set	Validation Set	Test Set	Accuracy on test set
90% of Dataset 2	10% of Dataset 2	Dataset 1	88.33%
90% of Dataset 1	10% of Dataset 1	Dataset 2	85%
Average	N/A	N/A	86.7%

For 13-inch flange joints, the average classification accuracy is 86.7%. These classification results are promising and demonstrate that our method (MFCC+SVM) has good generalization ability to work on medium flange joints. These classification results are promising. The confusion matrices of cross-set validation are presented in Fig. A10 in Appendix I. The precision, recall, and F1-score are included in Fig.

A11 in Appendix I. In addition, some MFCC matrices are selected and visualized in Fig. A12 in Appendix I.

To reproduce the experimental results, please find the folder '13_inch_underwater_uniform_load_experiment' and follow the readme file to run the script. After running the script, you will get the confusion matrices, classification metrics (precision, recall, and F1-score), and MFCC figures in Appendix I.

V.4.6 16-inch flange

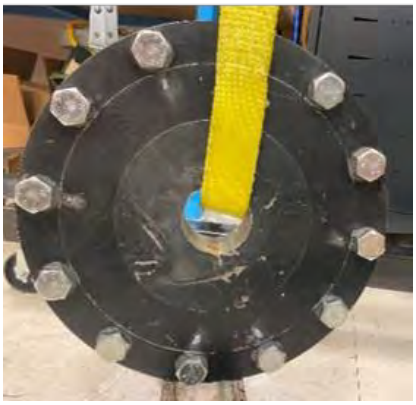


Fig. 53. 16-inch Flange

Uniform load

For the 16-inch flange, the maximum torque level is 150 ft-lbs. We have collected 3 data sets, each set having 4 classes (0 ft-lbs, 50 ft-lbs, 100 ft-lbs and 150 ft-lbs). Table 9 represents the data distribution for the 16-inch flange.

Table 9. Dataset Arrangement with Robotic Gripper for 16-inch Flange Joint.

Dataset ID	Loose (0 ft-lb, 0%)	Caution (50, 33%)	Warning (100 ft-lbs, 66%)	Tight (150 ft-lbs, 100%)	Total Number
Set 1	60	60	60	60	240
Set 2	60	60	60	60	240
Set 3	60	60	60	60	240
Total	N/A	N/A	N/A	N/A	**720**

For the 16-inch flange joint, we used one of 3 sets as the test set and the remaining two as developing and validation sets. In this training/test setting, the developing and test sets are independent of each other. The model is developed and validated on different portions (90% and 10%) of two sets from the three datasets and finally tested on the remaining one. The results of the classification are shown in Table 10.

Table 10. 4-Class Cross-set Validation Results of the 16-Inch Flange.

Training Set	Validation Set	Test Set	Accuracy on test set
90% of (Dataset 3,2)	10% of (Dataset 3,2)	Dataset 1	84.58%
90% of (Dataset 1,3)	10% of (Dataset 1,3)	Dataset 2	100%
90% of (Dataset 1,2)	10% of (Dataset 1,2)	Dataset 3	81.67%
Average	N/A	N/A	88.75%

In the 4-class task, the average classification accuracy is 88.75%. These classification results are promising. The confusion matrices of cross-set validation are presented in Fig. A13 in Appendix I. The precision, recall, and F1-score are included in Fig. A14 in Appendix I. In addition, some MFCC matrices are selected and visualized in Fig. A15 in Appendix I.

To reproduce the experimental results, please find the folder '16_inch_underwater_uniform_load_experiment' and follow the readme file to run the script. After running the script, you will get the confusion matrices, classification metrics (precision, recall, and F1-score), and MFCC figures in Appendix I.

V.4.7 21-inch flange

Fig. 54. 21-inch Flange

Uniform load

For the 21-inch flange, the maximum torque level is 220 ft-lbs. Four datasets were collected. Sets 1 and 2 are five-class datasets (0 ft-lbs, 70 ft-lbs, 120 ft-lbs, 170 ft-lbs, and 220 ft-lbs), whereas Sets 3 and 4 are three-class datasets (0 ft-lbs, 120 ft-lbs, and 220 ft-lbs). The labels in Sets 1 and 2 were transformed as shown in Fig. 19 to convert the task into a three-class classification problem: 0 ft-lbs and 70 ft-lbs were grouped and treated as the "Loose" class, while 170 ft-lbs and 220 ft-lbs were grouped as the "Tight" class.

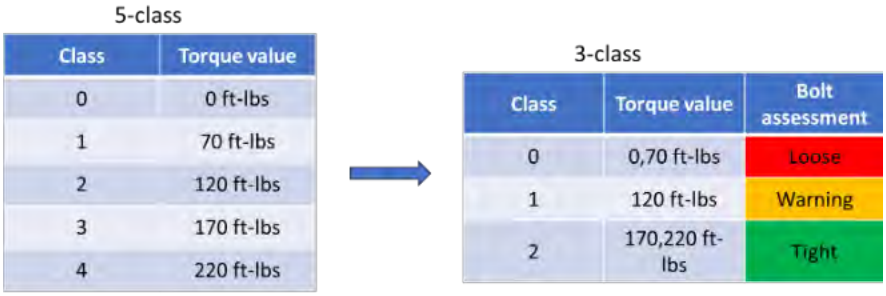


Fig. 55. Class Transformation Strategy.

Table 11 represents the data distribution for the 21-inch Flange.

Table 11. 3-Class Dataset Arrangement for 21-inch Flange Joint.

Dataset ID	Loose (0 & 70 ft-lbs(0 & 31.8%))	Warning (120 ft-lbs (55%))	Tight (170 & 220 ft-lbs (77.2 & 100%)	Total Number
Set 1	120	60	120	300
Set 2	119	60	120	299
Set 3	24	24	24	72
Set 4	24	24	24	72
Total	N/A	N/A	N/A	743

For the 21-inch flange joint, one of the four datasets is designated as the test set, while the remaining three are used for training and validation. Under this training–testing configuration, in which one dataset is reserved for testing and the remaining three are used for developing and validation, the training and test sets are completely independent. The model is developed and validated using a 90%–10% split of the selected three datasets and is subsequently evaluated on the remaining dataset. The classification results are presented in Table 12.

Table 12. 3-Class (0 ft-lbs and 70 ft-lbs,120 ft-lbs,170 ft-lbs and 220 ft-lbs) Cross-set Validation Results of the 21-Inch Flange.

Training Set	Validation Set	Test Set	Accuracy on test set
90% of (Dataset 2,3,4)	10% of (Dataset 2,3,4)	Dataset 1	91.64%
90% of (Dataset 1,3,4)	10% of (Dataset 1,3,4)	Dataset 2	94.67%
90% of (Dataset 1,2,4)	10% of (Dataset 1,2,4)	Dataset 3	94.44%
90% of (Dataset 1,2,3)	10% of (Dataset 1,2,3)	Dataset 4	97.22%
Average	N/A	N/A	94.5%

The confusion matrices of unseen test sets are shown in the appendix. In the 3-class task, the average classification accuracy is 94.5%. These results prove that the model can accurately predict the tightness level of the bolts. The confusion matrices of cross-set validation are presented in Fig. A16 in Appendix I.

The precision, recall, and F1-score are included in Fig. A17 in Appendix I. In addition, some MFCC matrices are selected and visualized in Fig. A18 in Appendix I.

To reproduce the experimental results, please find the folder '21_inch_underwater_uniform_load_experiment' and follow the readme file to run the script. After running the script, you will get the confusion matrices, classification metrics (precision, recall, and F1-score), and MFCC figures in Appendix I.

Non uniform load

In this case, we collected 5 independent datasets (Fig. 55) from a 21-inch flange joint underwater. Each dataset has 36 response signals.

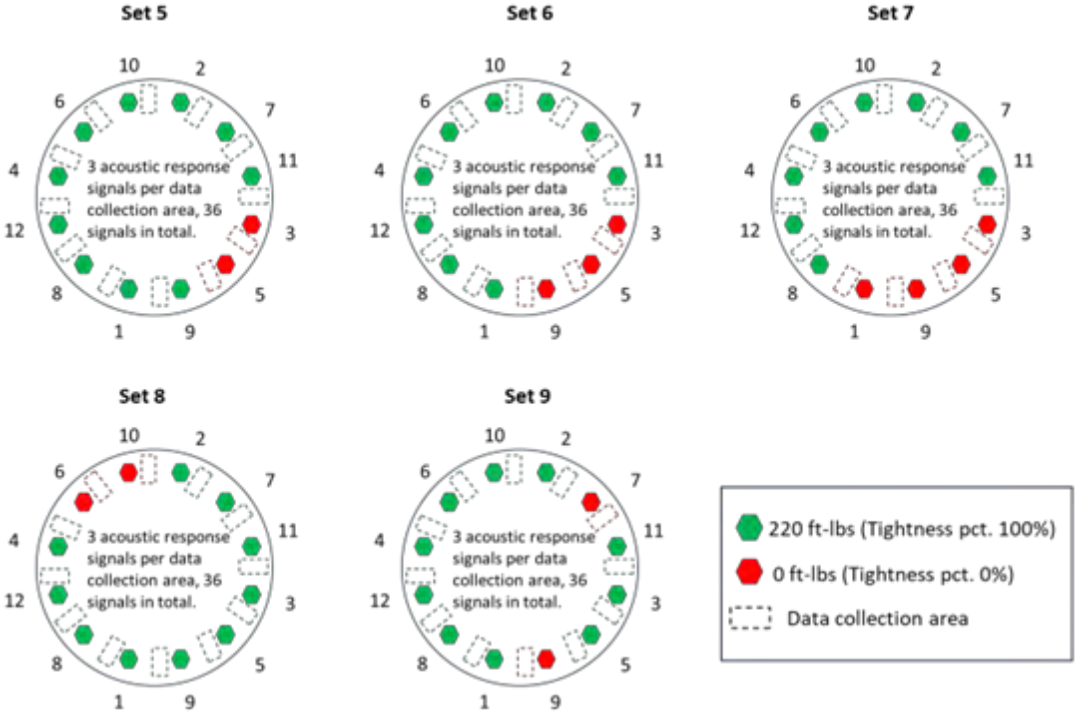


Fig. 56. Datasets of 21-inch Flange Joint in Non-uniform Load Distribution Scenario.

Fig. 57 shows a comparison between the ground truth (left) and prediction (right). A predicted torque value greater than/equal to 170 ft-lbs indicates that the bolt (or the area near the bolt) is in a 'Tight' state. A predicted torque value greater than 70 ft-lbs and smaller than 170 ft-lbs indicates that the bolt (or the area near the bolt) is in a 'Warning' state. A predicted torque value smaller than/equal to 70 ft-lbs indicates that the bolt (or the area near the bolt) is in a 'Loose' state. We studied 5 different sets. The results show that the model can identify the region of the loose bolts. Whenever a bolt is in 'Warning' state representing an intermediate condition in which the bolt is partially loosened and may require maintenance, or 'Loose' state, it needs to be tightened.

To reproduce the experimental results, please find the folder '21_inch_underwater_non_uniform_load_experiment' and follow the readme file to run the script. After

running the script, we will get the true labels and predictions, and we can manually draw thevisualizations of classification results in Fig. 57.

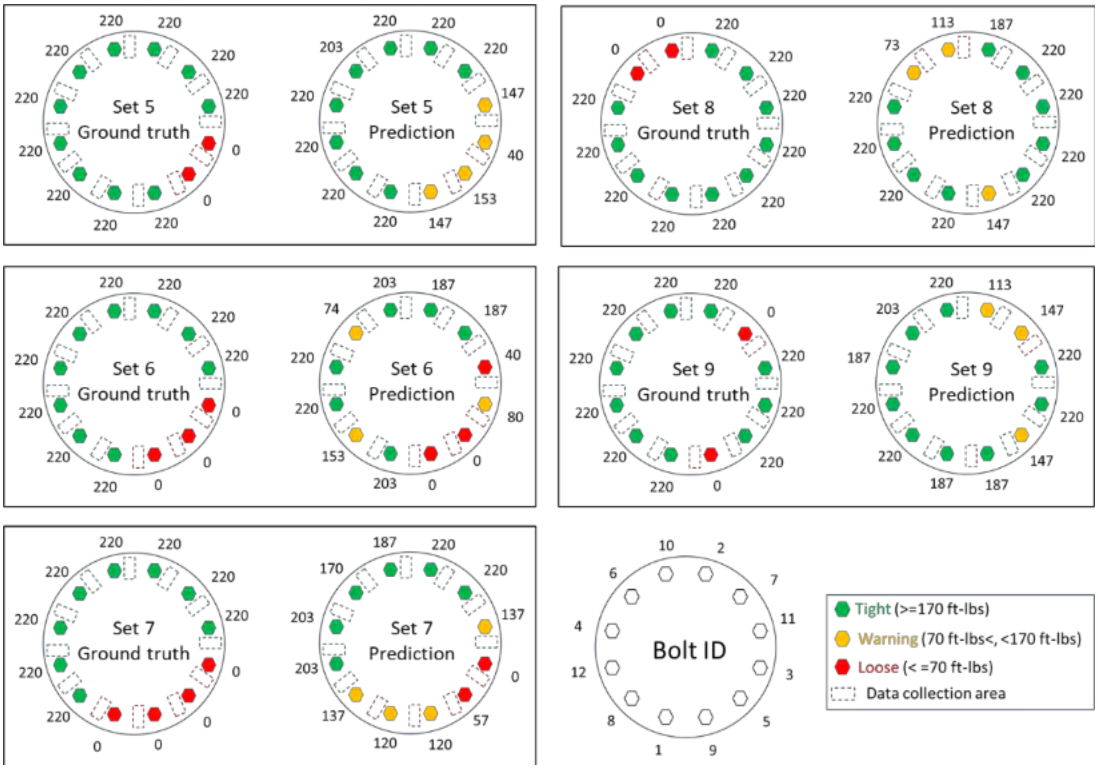


Fig. 57. Prediction results of the 21-inch flange.

V.4.8 5-inch flange

Fig. 58. 5-inch Flange

Non uniform load

The maximum torque level for the 5-inch flange is 30 ft-lbs. We have collected 3 independent data sets, each set consists of 4 classes (0 ft-lb, 10 ft-lbs, 20 ft-lbs and 30 ft-lbs) as shown in Table 13.

Table 13. Dataset Arrangement with Robotic Gripper Underwater For 5-Inch Flange Joint.

Dataset ID	Loose (0 ft-lb, 0%)	Caution (10 ft-lbs, 33%)	Warning (20 ft-lbs, 66%)	Tight (30 ft-lbs, 100%)	Total Number
Set 1	40	40	40	40	160
Set 2	40	40	40	40	160
Set 3	40	40	40	40	160
Total	120	120	120	120	480

These 3 sets are used to train the model, and it was tested on the 3 cases, indicated in Fig. 59. In the first case, only one bolt was loosened. In the second case, two non-adjacent bolts were loosened, and in the last case, two adjacent bolts were loosened.

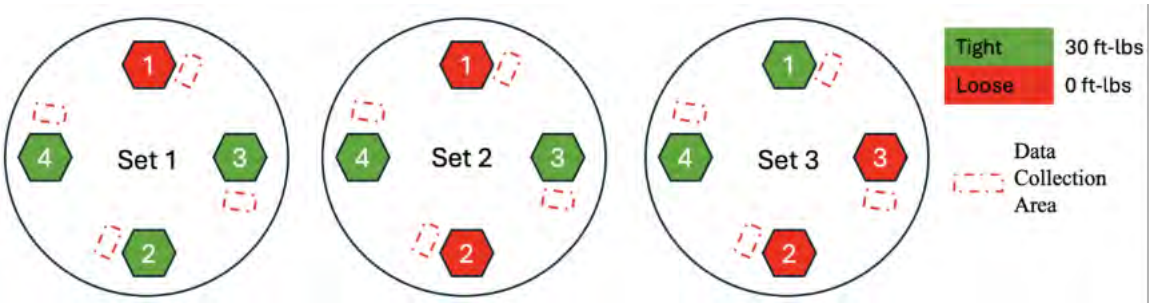


Fig. 59. 5-Inch Flange Joint in Non-uniform Load Distribution Scenario.

Fig. 60 shows the shows a comparison between the ground truth (Top) and prediction (Bottom). The green color means that the bolt is tight, yellow is warning, orange is caution and red is loose. We studied 3 different cases.

Given the low number of bolts, loosening one bolt affects the looseness of the other bolts, which made the results worse than the previous cases, where the number of bolts is higher (8 and 12 bolts)

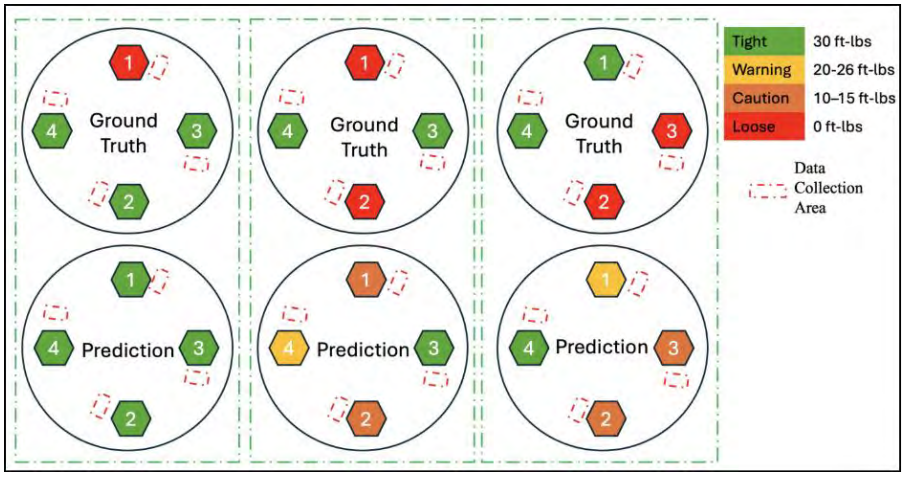


Fig. 60. Ground truth and predicted classification results for the 5-inch flange under non-uniform load distribution.

V.4.9 4.25-inch flange

Fig. 61. 4.25-inch Flange

Non uniform load

The maximum torque level for the 4.25-inch flange is 30 ft-lbs. We have collected 3 independent data sets, each set consists of 4 classes (0 ft-lbs, 10 ft-lbs, 20ft-lbs and 30 ft-lbs) as shown in Table 14.

Table 14. Dataset Arrangement Conducted Underwater For 4.25-Inch Flange Joint.

Dataset ID	Loose (0 ft-lb, 0%)	Caution (10 ft-lbs, 33%)	Warning (20 ft-lbs, 66%)	Tight (30 ft-lbs, 100%)	Total Number
Set 1	40	40	40	40	160
Set 2	40	40	40	40	160
Set 3	40	40	40	40	160
Total	120	120	120	120	480

These 3 sets are used to train the model, and it was tested on the 3 cases, indicated in Fig. 62. In the first case, two adjacent bolts were loosened. In the second case, two non-adjacent bolts were loosened, and in the last case, only one bolt was loosened.

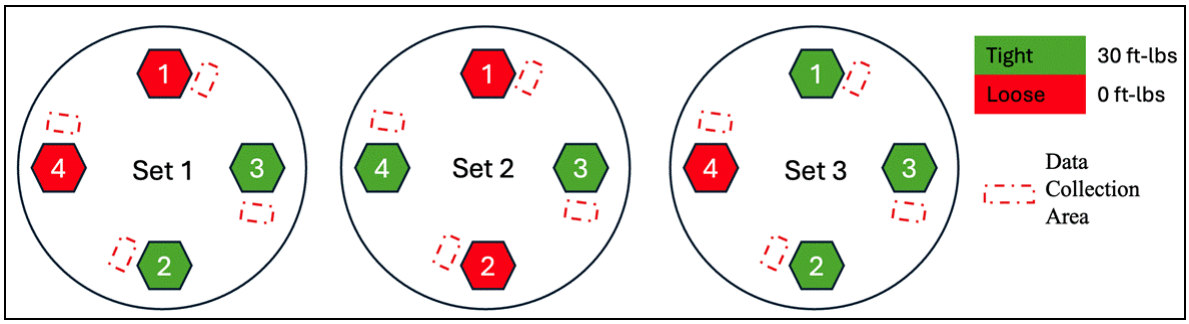


Fig. 62. Non-uniform Load Distribution case for the 4.25-Inch Flange.

Fig. 62 shows the shows a comparison between the ground truth (Top) and prediction (Bottom). The green color means that the bolt is tight, yellow is warning, orange is caution and red is loose.

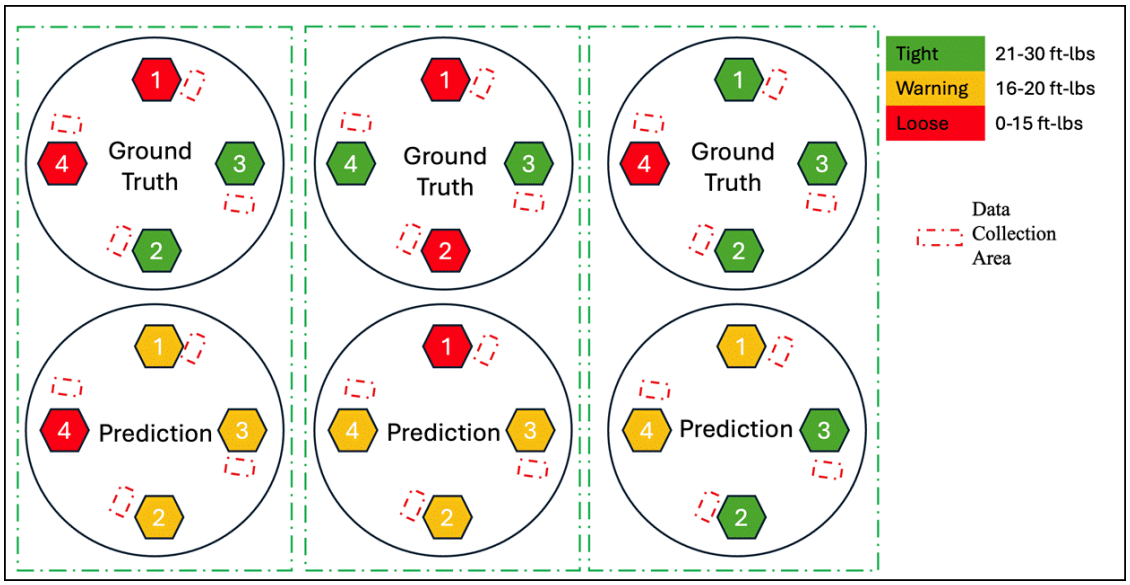


Fig. 63. Ground truth and predicted classification results for the 4.25-inch flange under non-uniform load distribution.

The results obtained for the 4.25-inch flange were consistent with those of the 5-inch flange. In structures with a small number of bolts, loosening even a few bolts significantly affects the responses measured from the remaining tightened bolts.

V.4.10 25-inch flange

Fig. 64. 25-inch Flange

Uniform load

The maximum torque level for the 25-inch flange (shown in the Fig. 64) is 500 ft-lbs. We have collected 3 independent data sets, each set consists of 3 classes (0 ft-lbs, 250 ft-lbs and 500 ft-lbs) as shown in Table 15. Each dataset contains 80 samples per torque level, resulting in 240 samples per dataset and 720 samples in total across all datasets.

Table 15. Dataset Arrangement Conducted Underwater For 25-Inch Flange Joint.

Dataset ID	Class 0 – Loose (0 ft-lbs)	Class 1 – Warning (250 ft-lbs)	Class 2 – Tight (500 ft-lbs)	Total Number
Set 1	80	80	80	240
Set 2	80	80	80	240
Set 3	80	80	80	240
Total	**240**	**240**	**240**	**720**

A 3-class classification task (0, 250, and 500 ft-lbs) was performed using a cross-dataset validation strategy, where two datasets were used for training and the remaining dataset was used for testing. The results are presented in Table 16. Detailed confussion matrices, precision, recall, and F1-scores for each class are provided in the Fig. A1 to A3 in Appendix III.

Table 16. 3-Class (0, 250 and 500 ft-lbs) Classification Results of 25-Inch Flange.

Training Set	Test Set	Accuracy on test set
Set 2 and Set 3	Set 1	75 %
Set 1 and Set 3	Set 2	82 %
Set 1 and Set 2	Set 3	80 %
Average	N/A	**79 %**

The experimental results indicate that while the proposed approach performs reliably for smaller flange sizes, scaling to larger-diameter flanges introduces limitations related to excitation strength. For the 25-inch flange, the increased diameter and width result in higher damping, greater wave dispersion, and reduced signal energy, which collectively limit classification performance when the same actuation strategy is used. Based on these observations, we propose that stronger piezoelectric actuators are required for larger flanges. In particular, the use of higher-energy actuators, such as stack-type piezo actuators, is expected to provide sufficient excitation to improve signal quality. Adequate actuation is necessary to achieve reliable classification for both uniform bolt-loosening conditions and more complex non-uniform loosening cases, where only a subset of bolts is loosened.

V.5 Physics Informed Neural Network (PINN)

V.5.1 Introduction

In recent years, physics-informed neural networks (PINN) have emerged as a promising paradigm for combining the strengths of data-driven models with the interpretability and consistency of physics-based principles. Traditional deep learning models, while effective in capturing complex patterns from large datasets, often lack physical consistency, leading to results that may not generalize well beyond the training domain (Karniadakis, 2021). Physics-informed approaches address this limitation by embedding physical knowledge directly into the learning process, thereby enhancing robustness, interpretability, and data efficiency (Wu, 2024).

In the context of structural health monitoring (SHM), the integration of physical principles into learning frameworks is particularly critical. Purely data-driven methods are prone to overfitting and often lack physical interpretability, which limits their ability to reliably capture underlying structural behaviors. Physics-informed deep learning offers a unique pathway by ensuring that model predictions not only fit the data but also adhere to the underlying physical characteristics of wave propagation (Zhang, 2021) (RAUTELA, 2021).

This study develops a physics-informed three-layered one-dimensional convolutional neural network (1D-CNN) to classify the tightness condition of flanges. The proposed method leverages stress waves generated by piezoceramic (PZT) actuators and captured by PZT sensors during robotic inspection. The recorded signals are processed into Mel-frequency cepstral coefficients (MFCCs). Unlike conventional CNN-based classification approaches, the present framework incorporates physical knowledge through an augmented loss function. Specifically, a penalty term is introduced to enforce consistency with the expected energy distribution of wave responses across torque levels. By embedding this constraint, the model is guided to learn representations that are not only statistically discriminative but also physically meaningful.

V.5.2 Physics Informed Neural Network (PINN) Model Architecture

The proposed model is a three-layer one-dimensional convolutional neural network (1D-CNN) designed to classify flange tightness levels based on stress wave signals. The input to the model is a set of **Mel-frequency cepstral coefficients (MFCCs)** extracted from the recorded waveforms. MFCCs provide a compact and discriminative representation of the frequency content of stress waves, making them suitable for capturing subtle differences associated with varying torque levels.

The CNN consists of three sequential convolutional layers, each followed by batch normalization and rectified linear unit (ReLU) activation. The convolutional layers progressively extract hierarchical features, ranging from local spectral patterns to higher-level abstractions of wave behavior. Max-pooling layers are incorporated to reduce dimensionality and enhance translation invariance. The extracted features are then passed to a fully connected layer with a softmax output, which performs classification across three torque levels.

While a standard CNN minimizes a purely data-driven classification loss (e.g., categorical cross-entropy), this approach may yield solutions that fit the training data but violate known physical behavior of stress waves. To address this limitation, we introduce a physics-informed regularization term into the loss function.

The augmented loss function is expressed as:

$$Ltotal=LCE + \lambda\ Lphy$$

where:

- L_{CE} is the standard cross-entropy classification loss,
- L_{phys} is the physics-informed penalty,
- λ is a weighting parameter that balances the contribution of the physical constraint.

The physics-informed term is constructed based on the energy distribution of stress waves. Since flange tightness affects the attenuation and scattering of stress waves, the expected signal energy should monotonically vary with torque level. L_{phys} penalizes deviations from this physically consistent trend by comparing the predicted class features with their corresponding energy signatures. Fig. 65. represents a flowchart of the proposed method

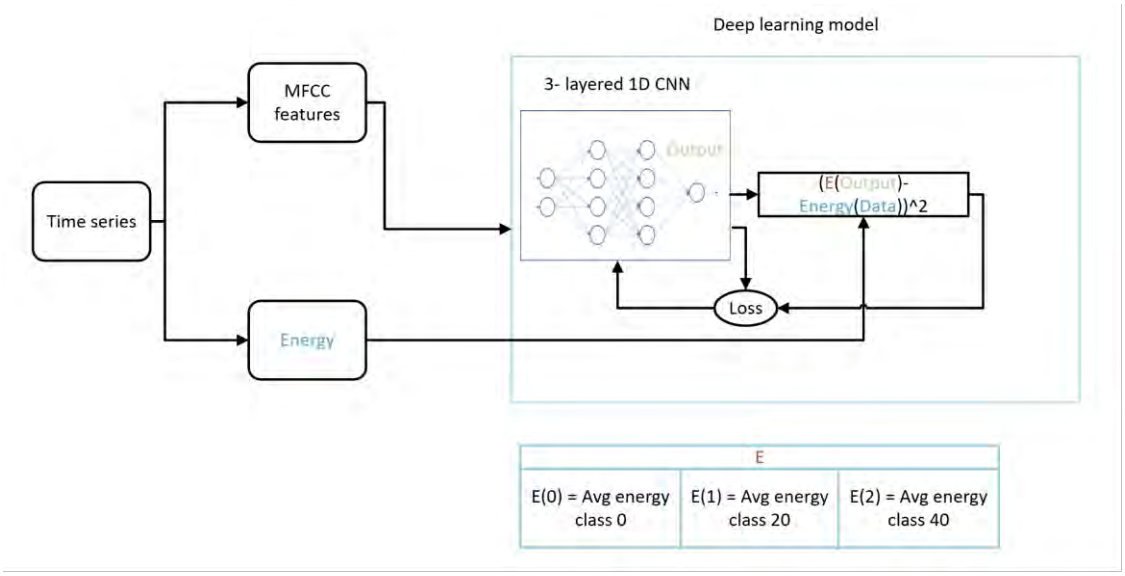


Fig. 65. Flow Chart of The PINN Model.

V.5.3 Data Collection

We have collected 9 independent data sets on the 9-inch flange. Each set had 120 samples for 3 different torque levels (0 ft-lb 20ft-lbs and 40 ft-lbs). Table 17 represents the dataset distribution.

In the case of the PINN, the study was limited to the 9-inch flange, as it provided the largest number of collected data points. Since deep learning methods generally require more data than their shallow learning counterparts, this flange was the most suitable choice for model training and validation.

Table 17. Dataset Arrangement with Robotic Gripper Underwater For 9-Inch Flange Joint.

Dataset ID	**Loose (0 ft-lbs, 0%)**	**Warning (20 ft-lbs, 50%)**	**Tight (40 ft-lbs, 100%)**	**Total Number**
Set 1	**40**	**40**	**40**	**120**

Set 2	**40**	**40**	**40**	**120**
Set 3	**40**	**40**	**40**	**120**
Set 4	**40**	**40**	**40**	**120**
Set 5	**40**	**40**	**40**	**120**
Set 6	**40**	**40**	**40**	**120**
Set 7	**40**	**40**	**40**	**120**
Set 8	**40**	**40**	**40**	**120**
Set 9	**40**	**40**	**40**	**120**
Total	**360**	**360**	**360**	**1080**

V.5.4 Experimental Result and Discussion

The model was developed using seven datasets, validated on one dataset, and evaluated on a completely unseen test set. Model parameters were optimized based on performance accuracy measured on the validation set. The classification accuracy for the 9-inch flange joint in the three-class task is reported in Table 18.

Table 18. 3-Class Classification Results of 9-inch flange using PINN model.

Training Set	Validation Set	Test Set	Accuracy on test set
Remaining 7 datasets	Dataset 9	Dataset 1	95%
Remaining 7 datasets	Dataset 1	Dataset 2	94.17%
Remaining 7 datasets	Dataset 2	Dataset 3	100%
Remaining 7 datasets	Dataset 3	Dataset 4	92.5%
Remaining 7 datasets	Dataset 4	Dataset 5	95.83%
Remaining 7 datasets	Dataset 5	Dataset 6	91.67%
Remaining 7 datasets	Dataset 6	Dataset 7	94.17%
Remaining 7 datasets	Dataset 7	Dataset 8	91.67%
Remaining 7 datasets	Dataset 8	Dataset 9	91.67%

Average	N/A	N/A	94.07%

The proposed model achieved an overall classification accuracy of 94.07%, demonstrating strong predictive performance and reliability. The confusion matrices for each case, presented in the Appendix, provide a detailed breakdown of the classification outcomes. These matrices highlight the consistency of the model across all different datasets, with minimal misclassifications, thereby confirming both the robustness and the practical applicability of the physics-informed CNN framework. The appendix also provides the Training and validation losses and accuracy for each set.

V.6 Domain Adaptation for Flange Looseness Detection

V.6.1 Introduction

Bolted flange joints are critical components in many industrial systems, especially in subsea pipelines, offshore structures, and high-pressure process equipment. Ensuring that flange bolts are properly preloaded is essential for preventing leakage, fatigue damage, and catastrophic failure. Traditional inspection methods are time-consuming, often intrusive, and difficult to deploy in harsh environments such as underwater locations or confined subsea manifolds.

In this work, we use piezoelectric sensing and machine learning to estimate bolt preload levels based on measured vibration responses to a chirp excitation. However, two major challenges arise:

- **Large flange configurations** (e.g., 9-inch and above) are heavy and difficult to handle, making extensive data collection expensive and logistically demanding.
- **Underwater experiments** require specialized facilities, subsea-like conditions, and robust insulation of sensors and cabling, making each additional data point costly and time-intensive.

In both cases, the central problem is **data scarcity** and **setup variability**. We often have relatively abundant data from "easier" configurations (e.g., smaller flanges in air) but limited data from more realistic or demanding configurations (e.g., larger flanges, or flanges underwater). Domain adaptation provides a principled way to bridge this gap by enabling models trained in one domain (source) to generalize effectively to another domain (target) whose data distribution differs.

V.6.1.1 Scope of the Report

This part of the report focuses on two domain adaptation problems for flange looseness detection:

I. **Domain adaptation across flange sizes (6-inch → 9-inch):**
   - The source domain consists of acoustic response signals measured on a 6-inch diameter flange.
   - The target domain consists of response signals measured on a 9-inch flange.
   - A **signal transformation** approach based on frequency response functions is used to transform 6-inch responses so that they resemble 9-inch responses, thereby reducing the amount of data required from the 9-inch configuration.

II. **Domain adaptation across environments (Air → Water):**

- The source domain consists of data collected in air on 6-inch flanges.
- The target domain consists of data collected underwater on the same flange.
- We address severe domain shifts using a combination of Domain-Adversarial Neural Networks (DANN) and a novel Centroid Alignment Domain Adaptation (CADA) method that uses only a small number of labeled underwater samples.

Both parts employ **Mel-frequency Cepstral Coefficient (MFCC)** features and classification model Support Vector Machines (SVM).

V.6.1.2 Contributions

The main contributions of this work are:

- **Frequency Response Function (FRF)–based signal transformation method:**

  We propose a method that uses frequency response functions to transform 6-inch flange responses into equivalent 9-inch responses. This allows the use of abundant 6-inch data to train a model that generalizes the 9-inch flange with minimal additional 9-inch data.

- **Domain-Adversarial Neural Network (DANN) and Centroid Alignment Domain Adaptation (CADA):**
  To handle the dual nature of the air-to-water domain shift, we use DANN to learn domain-invariant yet torque-discriminative features, thereby mitigating covariate shift, and we develop CADA as a few-shot method that uses a small set of labeled underwater samples to align class-wise feature centroids, explicitly correcting the remaining concept shift in the feature–label relationship.

V.6.2 Domain Adaptation Across Flange Sizes (6-inch → 9-inch)

V.6.2.1 Problem Definition

Mechanical differences between flanges of different sizes result in distinct dynamic behaviors. The 6-inch and 9-inch flange systems differ in geometry (thickness and diameter), mass, and stiffness. Even when excited with the same linear chirp signal, their time-domain response signals and corresponding frequency content are not identical.

Collecting large, labeled datasets from the 9-inch flange underwater is costly and time-consuming. In contrast, data from the 6-inch flange can be obtained more easily. Therefore, this study adopts the following approach:

- To **use a small number of representative 9-inch samples** to construct a transformation model.
- To **transform a larger 6-inch dataset** into its 9-inch equivalent using this model.
- To **train an SVM classifier** on transformed signals plus a small real 9-inch dataset, such that it generalizes to unseen 9-inch test sets.

V.6.2.2 Signal Transformation Framework

V.6.2.2.1 System Modeling Using Frequency Response Function (FRF).

In this work, each flange configuration is modeled as a linear dynamic system. A linear chirp excitation spanning a prescribed frequency band is applied to the flange as shown in the Fig. 66, and the resulting vibration response is measured by a piezoelectric sensor. The relationship between the input excitation $x(t)$

and the output response $y(t)$ is described in the frequency domain through a Frequency Response Function (FRF) (Allemang, 2022)

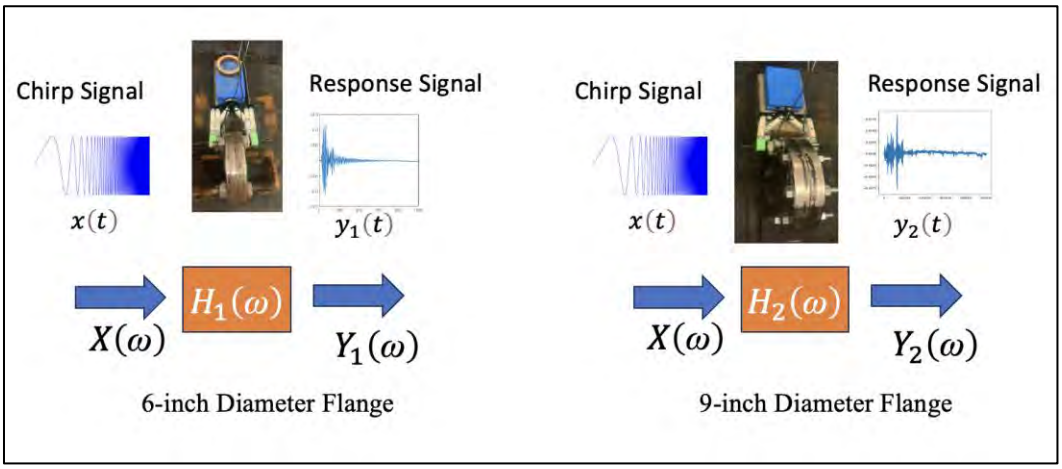


Fig. 66. Illustration of input chirp signal and output response, with the flange systems modeled as a Frequency Response Function.

Let:

- $X(\omega)$ = FFT of the input chirp signal $x(t)$
- $Y(\omega)$ = FFT of the output response $y(t)$
- $H(\omega)$ = Frequency response function of the flange system

The system's behavior is then defined as:

$$H(\omega) = \frac{Y(\omega)}{X(\omega)}$$

where each ω corresponds to a discrete frequency bin in the FFT.

In this study:

- $H_1(\omega)$ denotes FRF of the 6-inch flange (source system).
- $H_2(\omega)$ denotes the FRF of the 9-inch flange (target system).

Given that the input chirp signal is identical for both configurations, the FRF fully captures the mechanical system's frequency-domain response characteristics. Once these FRFs are estimated from input-output data, they can be used to transform signals from one system to another.

V.6.2.2.2 FRF Ratio Method

To estimate what the 9-inch flange response would look like given an observed 6-inch response, we employ a FRF ratio approach. Let:

- $Y_1(\omega)$ = output response of the 6-inch flange

- $Y_2(\omega)$= output response of the 9-inch flange

Using the definitions above:

$$Y_1(\omega) = H_1(\omega)X(\omega), \quad Y_2(\omega) = H_2(\omega)X(\omega)$$

Since $X(\omega)$ is the same for both flanges, we can write the estimated 9-inch response purely in terms of the 6-inch response and the ratio of frequency response functions:

$$Y_2(\omega) = \frac{H_2(\omega)}{H_1(\omega)} Y_1(\omega)$$

V.6.2.2.3 Overall Workflow of Signal Transformation for Domain Adaptation

The overall procedure used to transform 6-inch flange responses into their 9-inch equivalents is illustrated in Fig. 67. During the developing stage, a chirp excitation signal (1 kHz–60 kHz over 500 ms) is applied to the 6-inch flange, and the measured response $y_1(t)$ is processed through the FRF–based signal transformation module to generate the synthetic 9-inch response $y'_2(t)$. MFCC features extracted from these transformed signals form the training input for the SVM classifier. In the test stage, the same chirp excitation is applied to the real 9-inch flange, the resulting response $y_2(t)$ is converted to MFCC features, and these are passed to the trained SVM for preload prediction. The lower part of the figure depicts the core transformation operation: the 6-inch response is converted to the frequency domain using FFT, multiplied by the FRF ratio $\frac{H_2(\omega)}{H_1(\omega)}$, and reconstructed using IFFT to obtain $y'_2(t)$.

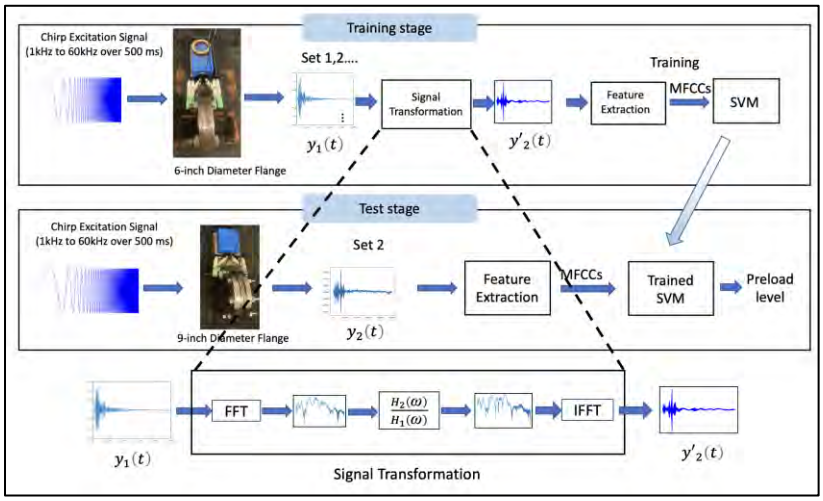


Fig. 67. Overall workflow of the signal-transformation-based domain adaptation from 6-inch to 9-inch flange.

The below Fig. 68. compares the raw 6-inch flange response $y_1(t)$, the transformed 9-inch-equivalent signal $y'_2(t)$. and the actual measured 9-inch response $y_2(t)$. The transformation result $y'_2(t)$. closely resembles the true 9-inch signal in terms of amplitude distribution, oscillatory behavior, and damping characteristics, despite originating from the 6-inch response. The increased complexity and richer low-frequency content present in both $y'_2(t)$ and the real 9-inch response demonstrate that the FRF ratio effectively captures the mechanical differences between the flange sizes.

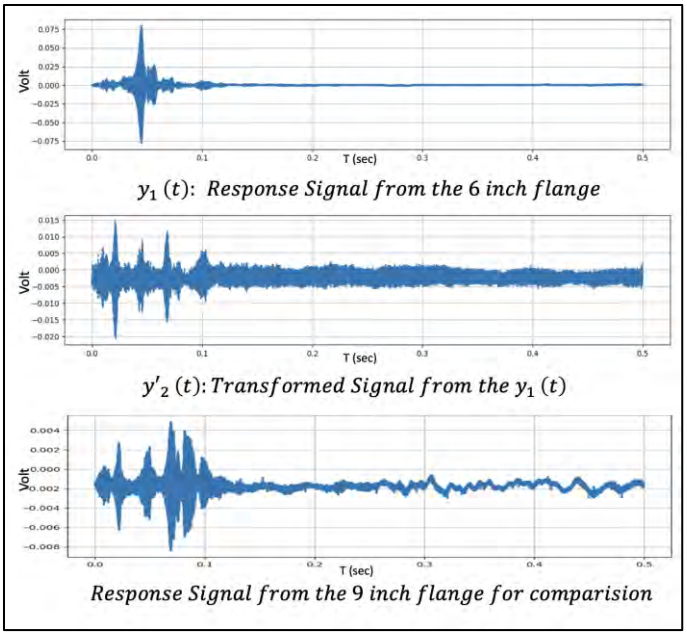


Fig. 68. Comparison of 6-inch response, transformed signal, and real 9-inch response.

Figs. 69–70 show the overlaid Bode plots of the measured FRFs for the 6-inch and 9-inch flanges at 40 ft-lbs preload from four repeated samples each. For both flanges, the magnitude responses exhibit similar resonance locations and overall spectral patterns across samples, indicating good repeatability of the dynamic behavior under identical tightening conditions, with minor amplitude variations due to differences in excitation and measurement conditions. The phase responses show rapid transitions near resonances and discontinuities caused by phase wrapping and reduced reliability in low-magnitude regions.

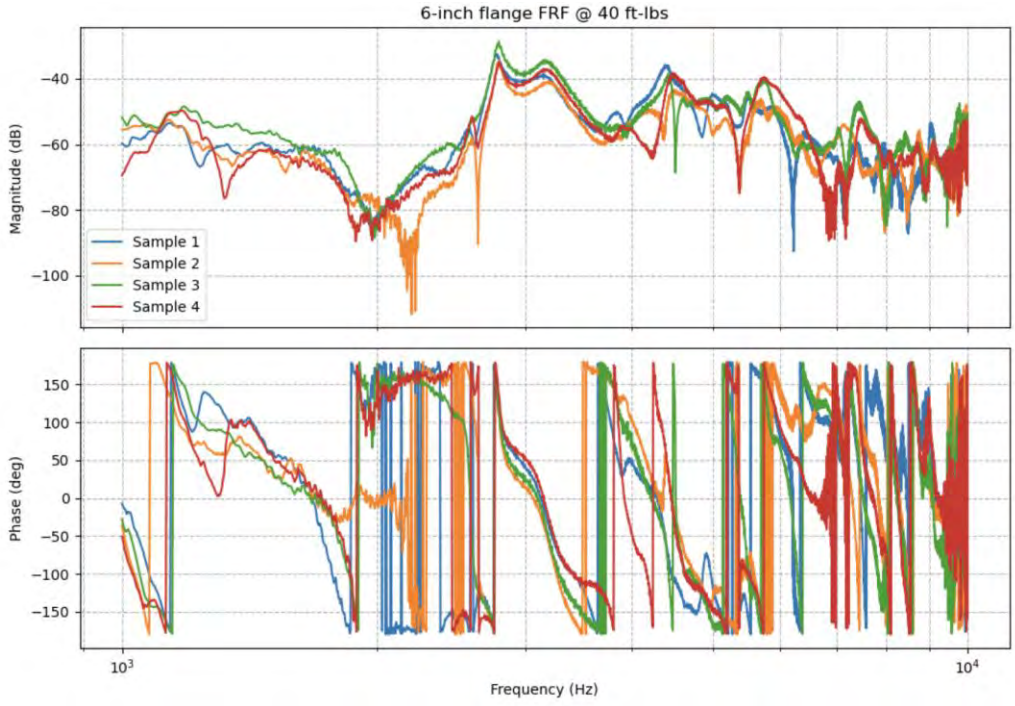


Fig. 69: Frequency Response Functions (FRFs) of a 6-inch flange at 40 ft-lb from four repeated measurements

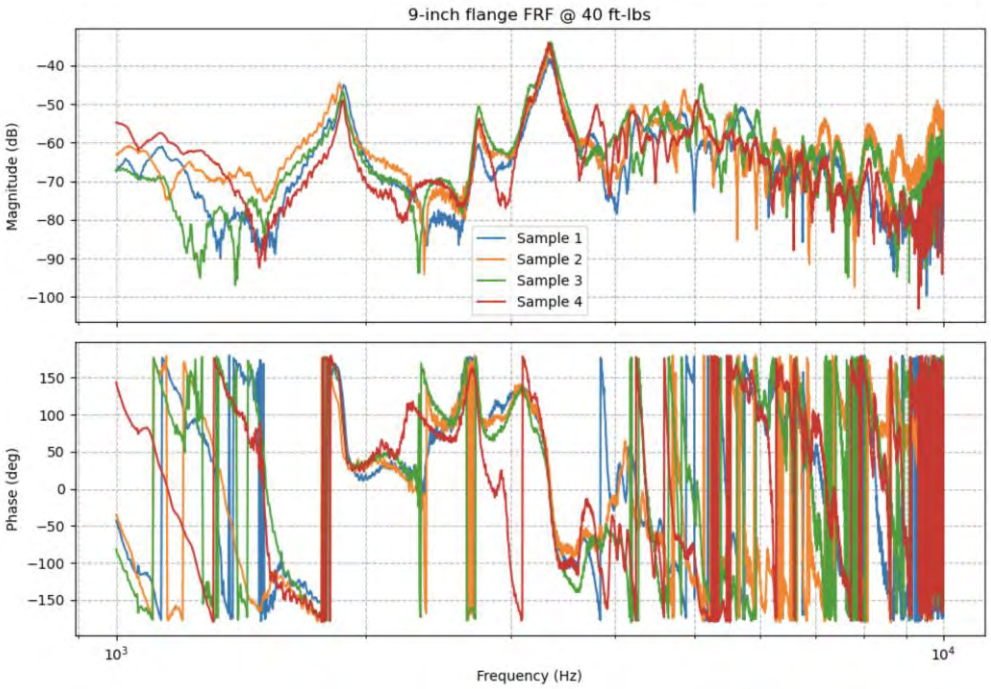


Fig. 70: Frequency Response Functions (FRFs) of a 9-inch flange at 40 ft-lb from four repeated measurements

V.6.3 Dataset Collection, Arrangement and Class Reorganization

We collected data using the robotic gripper setup underwater for the 9-inch diameter flange configuration. Three datasets—Set_9-in 1, Set_9-in 2, and Set_9-in 3—were acquired, each containing 200 response signals captured under five distinct flange preload levels: 0, 10, 20, 30, and 40 ft-lbs (40 samples per level), resulting in a total of 600 samples for the 9-inch flange case (Table 19).

We also collected data for the 6-inch diameter flange configuration. Four datasets—Set_6-in 1 to Set_6-in 4—were acquired under the same five torque levels with 40 samples per level, yielding 200 samples per dataset and a combined total of 800 samples (Table 20).

Table 19. Dataset Arrangement Underwater for **9-inch** Flange.

Dataset ID	0 ft-lb	10 ft-lbs	20 ft-lbs	30 ft-lbs	40 ft-lbs	Total Number
Set 9-in 1	40	40	40	40	40	200
Set 9-in 2	40	40	40	40	40	200
Set 9-in 3	40	40	40	40	40	200
Total	N/A	N/A	N/A	N/A	N/A	**600**

Table 20. Dataset Arrangement Underwater for **6-inch** Flange.

Dataset ID	0 ft-lb	10 ft-lbs	20 ft-lbs	30 ft-lbs	40 ft-lbs	Total Number
Set 6-in 1	**40**	**40**	**40**	**40**	**40**	**200**
Set 6-in 2	**40**	**40**	**40**	**40**	**40**	**200**
Set 6-in 3	**40**	**40**	**40**	**40**	**40**	**200**
Set 6-in 4	**40**	**40**	**40**	**40**	**40**	**200**
Total	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**800**

Although we have five classes (preload levels) for each dataset, sometimes in actual projects, it is unnecessary to have so many categories. Therefore, we not only adopt 5-class, but also simplify 5-class tasks to 2-class, 3-class, and 4-class, respectively (Fig. 71).

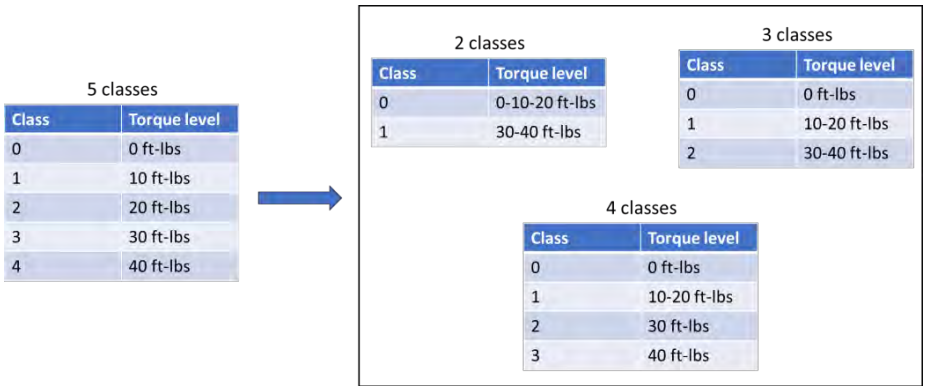


Fig. 71. Class Reorganization.

V.6.4 Signal Transformation from 6-inch Flange to 9-inch Flange

We utilized FRFs to transform the 6-inch system responses into their 9-inch equivalents. Since the dynamic characteristics vary across torque levels, separate FRFs were obtained for each torque label. Table 21 shows the signal transformation arrangement.

The notation $H^{n}_{\{1,x\}}$ and $H^{n}_{\{2,x\}}$ denotes the FRFs used in the transformation process, where the subscript 1 refers to the 6-inch system, 2 refers to the 9-inch system, and x indicates the torque level. The superscript n identifies the sample batch used to compute the FRFs. Sample batches from the Set_6-in 1 and Set_9-in 1 dataset are used for obtaining the FRF. And Set_6-in 2 dataset is transformed to its equivalent 9-inch dataset (**Set_9-in_Tr X**).

Table 21: Summary of Transformed Datasets from Set_6-in 2.

Transformed Dataset $y'_2(t)$	Source Dataset $y_1(t)$	FRFs Involved	Notes
Set_9-in_Tr 1	Set_6-in 2	$H^1_{\{1,x\}}, H^1_{\{2,x\}}$ $x \in \{0,10,20,30,40\}$ ftlbs	FRFs created using first 5 samples per torque level from each flange types
Set_9-in_Tr 2	Set_6-in 2	$H^2_{\{1,x\}}, H^2_{\{2,x\}}$ $x \in \{0,10,20,30,40\}$ ftlbs	Second batch of distinct samples used
Set_9-in_Tr 3	Set_6-in 2	$H^3_{\{1,x\}}, H^3_{\{2,x\}}$ $x \in \{0,10,20,30,40\}$ ftlbs	Third batch of distinct samples used
Set_9-in_Tr 4	Set_6-in 2	$H^4_{\{1,x\}}, H^4_{\{2,x\}}$ $x \in \{0,10,20,30,40\}$ ftlbs	Fourth batch of distinct samples used

V.6.5 Experimental Result and Discussion

The model was tested on 2-class, 3-class, 4-class, and 5-class classification tasks. It performed very well in the 2-class and 3-class cases, with an average accuracy of around 94%. However, the accuracy dropped to 82.5% for the 4-class and 79% for the 5-class tasks. One of reasons to explain the drop is that as the number of classes increases, the differences between torque levels become smaller and harder for the model to pick up using MFCC features. Some torque levels may produce very similar signals, making it challenging to identify them. Also, a limited number of samples per class and the noise in the data may affect the model's performance. Overall, while the model works well for broader categories, improving it for more detailed classification may require more advanced feature extraction methods and more training data.

Table 26 presents the 3-class classification results for the cross-size scenario where the model was trained on data from the 6-inch flange and tested on the 9-inch flange without applying FRF-based signal transformation. The test accuracies range from 30% to 33%, with an average of 31.66%, which is close to random guessing for a three-class problem. This poor performance indicates that the signals from the 6-inch and 9-inch flanges differ significantly due to their distinct dynamic characteristics, leading to a strong domain mismatch. The results highlight the necessity of signal transformation or domain adaptation to enable effective knowledge transfer across flanges of different sizes. Detailed confussion matrices, precision, recall, and F1-scores for each class are provided in the Fig. A4 to A14 in Appendix III.

Table 22. 2 -Class Classification Results.

Training Set	Validation Set	Test Set	Accuracy on Test Set
Set_9-in_Tr 1 to Set_9-in_Tr 4	Set_9-in 3	Set_9-in 2	93%
Set_9-in_Tr 1 to Set_9-in_Tr 4	Set_9-in 2	Set_9-in 3	95%
Average	N/A	N/A	**94%**

Table 23. 3- Class Classification Results.

Training Set	Validation Set	Test Set	Accuracy on Test Set
Set_9-in_Tr 1 to Set_9-in_Tr 4	Set_9-in 3	Set_9-in 2	94%
Set_9-in_Tr 1 to Set_9-in_Tr 4	Set_9-in 2	Set_9-in 3	94%
Average	N/A	N/A	**94%**

Table 24. 4-Class Classification Results.

Training Set	Validation Set	Test Set	Accuracy on Test Set
Set_9-in_Tr 1 to Set_9-in_Tr 4	Set_9-in 3	Set_9-in 2	85%
Set_9-in_Tr 1 to Set_9-in_Tr 4	Set_9-in 2	Set_9-in 3	80%
Average	N/A	N/A	**82.5%**

Table 25. 5-Class Classification Results.

Training Set	Validation Set	Test Set	Accuracy on Test Set
Set_9-in_Tr 1 to Set_9-in_Tr 4	Set_9-in 3	Set_9-in 2	80%
Set_9-in_Tr 1 to Set_9-in_Tr 4	Set_9-in 2	Set_9-in 3	78%
Average	N/A	N/A	**79%**

Table 26. 3-Class Classification Results WITHOUT FRF based signal transformation

Training Set	Validation Set	Test Set	Accuracy on Test Set
Set_6-in 1 to Set_6-in 4	Set_9-in 3	Set_9-in 1	33%
Set_6-in 1 to Set_6-in 4	Set_9-in 2	Set_9-in 2	30%
Set_6-in 1 to Set_6-in 4	Set_9-in 1	Set_9-in 3	32%
Average	N/A	N/A	**31.66%**

V.6.6 Summary of Findings

The FRF-based signal transformation method successfully performs domain adaptation between flanges of different sizes. By using only a small number of 9-inch samples to estimate FRF, we can transform 6-inch responses to closely emulate 9-inch responses and develop an SVM classifier that generalizes real 9-inch data.

Key conclusions:

- The method reduces the need for extensive data collection on the larger 9-inch flange.
- It achieves high accuracy for 2-class and 3-class tasks and acceptable performance for 4-class and 5-class tasks.

This approach demonstrates how physics-based modeling (Frequency Response Functions) can be combined with machine learning to address domain adaptation across mechanical configurations.

V.7 Domain Adaptation from Air to Water

V.7.1 Problem Definition

Underwater experiments are significantly more challenging than air experiments. They require:

- Large water tanks or subsea-like facilities,
- Specialized insulation of piezoelectric sensors and cables,
- Careful handling of robotic grippers and heavy flange assemblies.

As a result, collecting large volumes of labeled underwater data is expensive, time-consuming, and logistically demanding. In contrast, air-domain data on the same flange geometry can be collected more easily and in greater quantities.

The aim of this part of the work is to leverage abundant air-domain data and a small number of underwater samples to build a torque classification model that performs well in water. This is a classic cross-domain generalization problem with a severe distribution shift between source (air) and target (water).

V.7.2 Types of Domain Shift

Domain shift refers to the mismatch between the statistical properties of the source domain and the target domain. In the air-to-water transition, the medium, boundary conditions, damping characteristics, and acoustic propagation all change significantly, causing the input signals and their torque-dependent patterns to differ. Two types of domain shift are relevant:

V.7.2.1 Covariate Shift

Covariate shift (Nair, 2019) occurs when the input feature distributions differ across domains, while the underlying labeling rule remains unchanged:

$$p_s(x) \neq p_t(x), \qquad p_s(y \mid x) = p_t(y \mid x)$$

In this case, the data “looks” different, but if the model could perfectly understand the features, the torque labeling rule would still be valid. For flange signals, this corresponds to the fact that underwater signals

have very different amplitudes, damping, and spectral content compared to air, but the physical meaning of torque does not change.

V.7.2.2 Concept Shift (Conditional Shift)

Concept shift (Kull, 2014)occurs when the mapping between features and labels changes across domains:

$$p_s(x) = p_t(x), \qquad p_s(y \mid x) \neq p_t(y \mid x)$$

Here, the same feature patterns correspond to different labels in the two domains. In the flange experiment, underwater propagation alters how torque signatures appear in the MFCC feature space. The same torque level produces a different pattern underwater than in air, meaning the labeling rule itself shifts.

Our Case: Both Covariate and Concept Shift Occur

In the air-to-water transition, both shifts are present:

$$p_s(x) \neq p_t(x), \qquad p_s(y \mid x) \neq p_t(y \mid x)$$

This is because:

- The waveform distributions change drastically due to fluid–structure coupling, added mass, damping, and insulation effects (covariate shift).
- The relationship between torque levels and MFCC features changes underwater, resulting in altered class-wise feature distributions (concept shift).

A single adaptation method cannot correct both shifts. Therefore, we use two complementary approaches:

- **Domain-Adversarial Neural Network (DANN):** DANN (Ajakan, 2014) is a widely used deep-learning framework for mitigating covariate shifts. It forces the feature extractor to produce domain-invariant representations by using adversarial training with a gradient-reversal layer, ensuring that features from air and water share a common distribution.
- **Centroid Alignment Domain Adaptation (CADA):** We develop CADA to directly address concept shifts, which DANN alone cannot resolve. CADA uses a small number of labeled underwater (target) samples to compute class-wise centroid differences and align the source features accordingly, correcting changes in the feature-to-label relationship.

Together, DANN and CADA form a unified framework capable of handling both covariate and concept shifts in the air-to-water domain adaptation problem.

V.7.3 Domain-Adversarial Neural Network (DANN)

To address the covariate, shift between air and water domains, we employ the Domain-Adversarial Neural Network (DANN), a neural architecture designed to extract features that are both label-discriminative and domain-invariant. This ensures that a model trained on air-domain data can generalize effectively to the underwater domain, despite differences in signal characteristics.

As shown in the DANN architecture diagram in Fig. 72. the network takes MFCC features as input. These features pass through a shared feature extractor, typically composed of fully connected (FC) layers, which

generates a compact representation of the input data. This shared representation is then used by two parallel branches:

- The label predictor learns to classify the input data into torque classes using supervision from the source (air) domain labels.
- The domain classifier, connected via a Gradient Reversal Layer (GRL), learns to distinguish whether the feature originated from the source or target domain. During training, the GRL inverts the gradient from the domain classifier, forcing the feature extractor to learn features that confuse the domain classifier—thus encouraging domain invariance.

The result of this training setup is a feature space where samples from both domains are aligned in distribution, while still retaining meaningful class distinctions.

Fig. 73 illustrates the effect of domain adaptation using DANN and the proposed CADA method on MFCC features. The original MFCC feature space shows overlapping and poorly separated clusters among the three torque classes. After applying DANN, the features become more label-discriminative, resulting in improved separation between classes. The subsequent CADA step primarily aligns the already separated features into a common distribution across domains, reorganizing the feature space without significantly altering class separability. This alignment reduces domain discrepancy while preserving the discriminative structure learned by DANN, enabling more consistent classification across conditions.

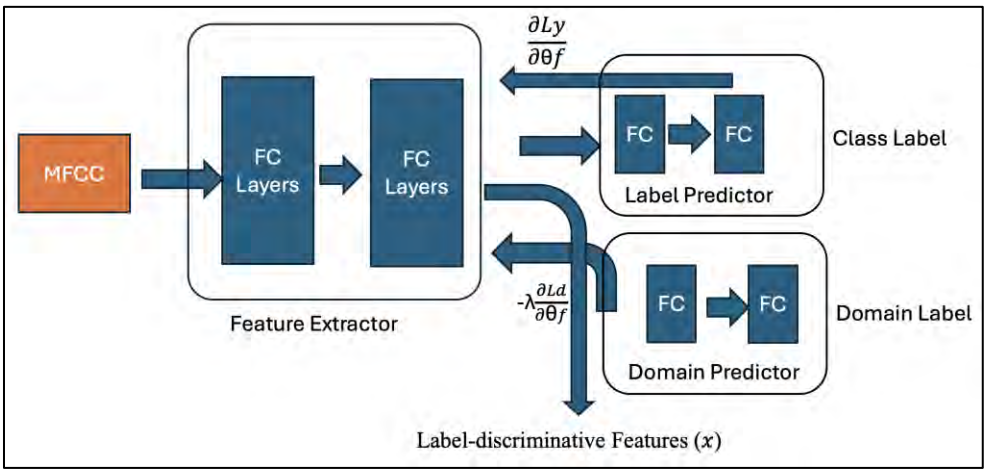


Fig. 72. DANN Architecture Diagram.

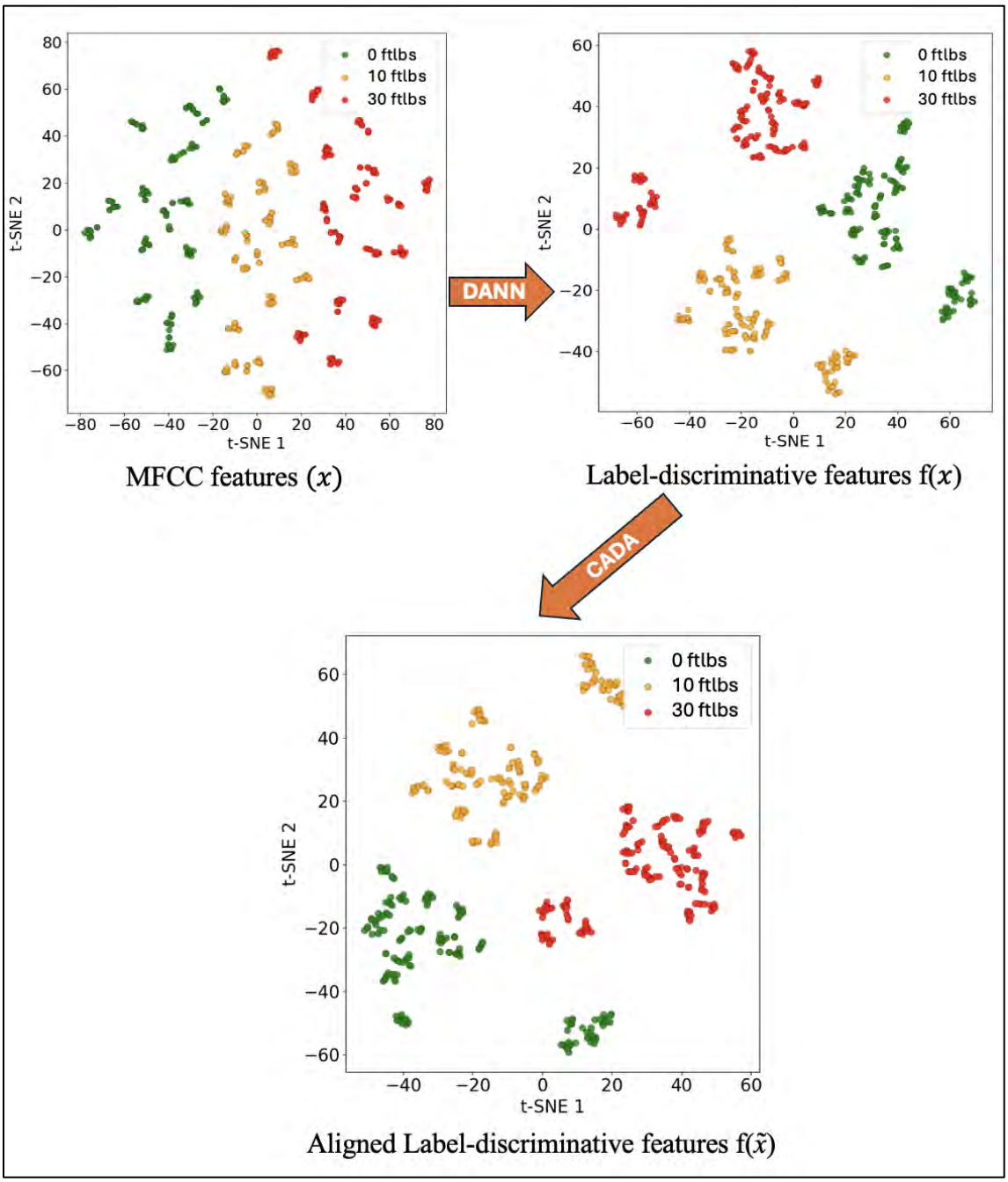


Fig. 73: Effect of DANN and CADA on MFCC Features.

V.7.4 Centroid Alignment Domain Adaptation (CADA) Using Few-Shot Learning

In the presence of concept shift, where the label distribution given input features $P(Y \mid X)$ differs between the source and target domains, traditional domain adaptation methods like DANN fall short. To overcome this, we propose a few-shot learning-based approach called Centroid Alignment Domain Adaptation (CADA). Instead of adapting the model directly to the target domain, CADA modifies the distribution of

the source training features by aligning them with the distribution observed in the few-shot labeled samples from the target domain.

V.7.4.1 Few-Shot Learning: Motivation and Intuition

Few-shot learning (Parnami, 2022) refers to a learning paradigm where the model is required to adapt to a new task or domain using only a small number of labeled samples. The underlying assumption is that by leveraging prior knowledge learned from the source domain, the model can generalize to new, low-resource settings with minimal supervision.

A relatable analogy is learning to drive a car in the U.S. and then quickly adapting to driving in Japan. Although the traffic rules and orientation change, prior to driving knowledge allows for fast adaptation after observing just a few signs and traffic patterns.

In our case, few-shot learning is used to incorporate a small number of labeled samples from the target (water) domain to help the model understand how class distributions shift under underwater conditions.

V.7.4.2 Visualizing Concept Shift in Source and Target Domains

This schematic t-SNE plot in Fig. 74 illustrates the distribution of label-discriminative features x extracted using DANN, emphasizing the class-wise centroids for both the source domain (large circles) and the target domain (few-shot samples, shown as smaller squares). Each color represents a different torque class label $c \in \{0,10,20,30,40\}$ ft-lbs. Although DANN effectively separates features by class, this visualization highlights the misalignment between the source centroids $\boldsymbol{\mu_S^c}$ and their corresponding target centroids $\boldsymbol{\mu_T^c}$. The separation between centroids across domains reflects the presence of concept shift, where the conditional distribution $P(Y \mid X)$ varies. This discrepancy motivates the use of **Centroid Alignment Domain Adaptation (CADA)** to correct the alignment by adjusting source domain features based on the centroid shifts, thereby improving classification performance on the target domain.

For each class c, we compute the **centroid** of the source domain features $\boldsymbol{\mu_S^c}$ and the centroid of the corresponding few-shot target domain features $\boldsymbol{\mu_T^c}$ using:

$$\mu_S^c = \frac{1}{|D_S^c|} \sum_{x \in D_S^c} x$$

$$\mu_T^c = \frac{1}{|D_T^c|} \sum_{x \in D_T^c} x$$

$| D_S^c |$ and $| D_T^c |$ denote the number of feature vectors (i.e., sample count) in the respective domain sets.

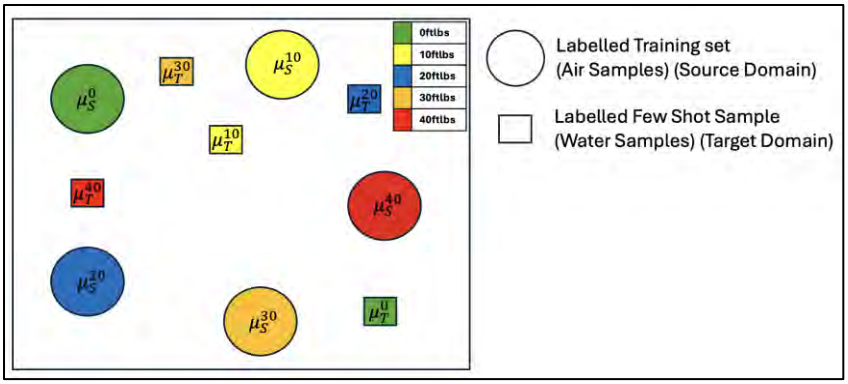


Fig. 74: Class-wise Centroid Misalignment in Source and Target Domains.

V.7.4.3 Class-wise Shift Vector Calculation

To effectively align the class-wise distributions of the source and target domains, we apply Centroid Alignment Domain Adaptation (CADA) by calculating and utilizing shift vectors. These shift vectors are derived from the difference between the class centroids of the source domain (air samples) and those of the few-shot labeled target domain (water samples).

The shift vector δ^c is then defined as:

$$\delta^c = \mu_T^c - \mu_S^c$$

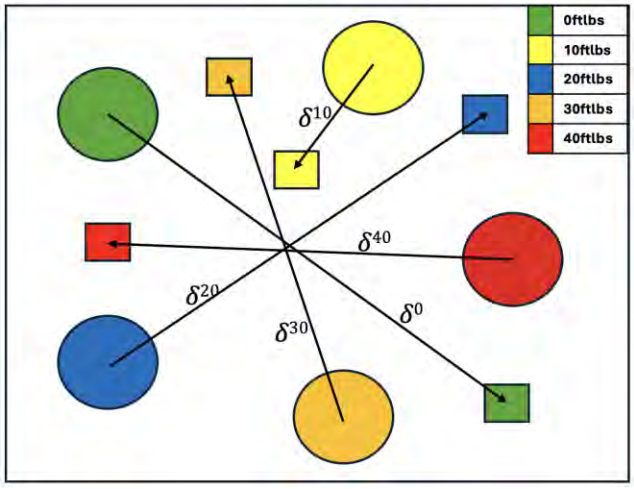


Fig. 75： Class-wise Shift Vectors.

This vector quantifies the offset between the two domains for class c. as shown in Fig. 75, arrows indicate the direction and magnitude of δ^c for each class. These vectors are used to adjust the source domain features or the target domain features to reduce misalignment in the feature space.

V.7.4.4 Adaptation of the Source Training Set to Target Domain

To bridge the distribution gap between the source and target domains under concept shift, we apply Centroid Alignment Domain Adaptation (CADA) by transforming the source training set using class-wise shift

vectors. This transformation ensures that the source features better reflect the distribution of the few-shot labeled samples from the target domain.

For each class $c \in \{0,10,20,30,40\}$ ft-lbs, let D_S^c be the set of source domain feature vectors belonging to class c, and let $x^c \in D_S^c$ represent a sample from that set. Using the previously computed centroids μ_S^c and μ_T^c, the class-wise shift vector is given by:

$$\delta^c = \mu_T^c - \mu_S^c$$

Each source sample x^c is then shifted by this vector to produce an aligned feature:

$$\tilde{x}^c = x^c + \delta^c$$

This operation effectively translates the source domain features, so they are aligned with the centroid of the corresponding class in the target domain. The aligned source training set for class c is then defined as:

$$\widetilde{D}_S^c = \{x^c + \delta^c \mid x^c \in D_S^c\}$$

The union of all such class-wise aligned sets is used to train the final classifier (e.g., SVM)

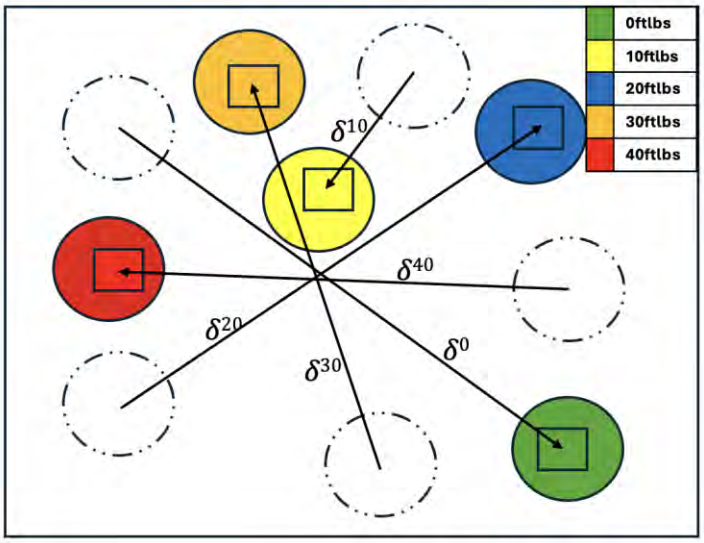


Fig. 76 Alignment of Source Training Samples Using Centroid Shifts.

This schematic plot shown in Fig. 76 illustrates the effect of Centroid Alignment Domain Adaptation (CADA) by visually representing the transformation of source domain samples. The colored squares represent the few-shot labeled samples from the target domain (water) for each torque class $c \in \{0,10,20,30,40\}$ ft-lbs. The colored circles indicate the source domain samples after alignment, where each source sample has been shifted by its corresponding shift vector δ^c. The dashed circles denote the original positions of the source domain samples before alignment. The arrows represent the shift direction for each class, showing how the source samples are moved toward the target distribution, thus improving domain alignment under concept shift.

V.7.5 Domain Adaptation Framework Combining DANN and CADA

The overall model architecture for torque level classification combines both the Domain-Adversarial Neural Network (DANN) and our proposed Centroid Alignment Domain Adaptation (CADA) method is shown in Fig. 77.

The process begins by extracting MFCC features from the source domain training set (air) and the few-shot labeled samples from the target domain (water). These features are passed through a shared DANN network that learns label-discriminative but domain-invariant representations, denoted as x. Simultaneously, MFCC features from 25 target-domain few-shot samples are also passed through the same DANN to obtain corresponding target-domain features.

Using both sets of DANN-extracted features, CADA computes class-wise shift vectors δ^c , which are used to align the source features toward the few-shot target distribution, resulting in the aligned training set. This aligned dataset is then used to train a Support Vector Machine (SVM) classifier.

MFCC features from the test set (target domain) are processed through the same DANN network to ensure feature space consistency. These features are then passed to the trained SVM for torque label prediction.

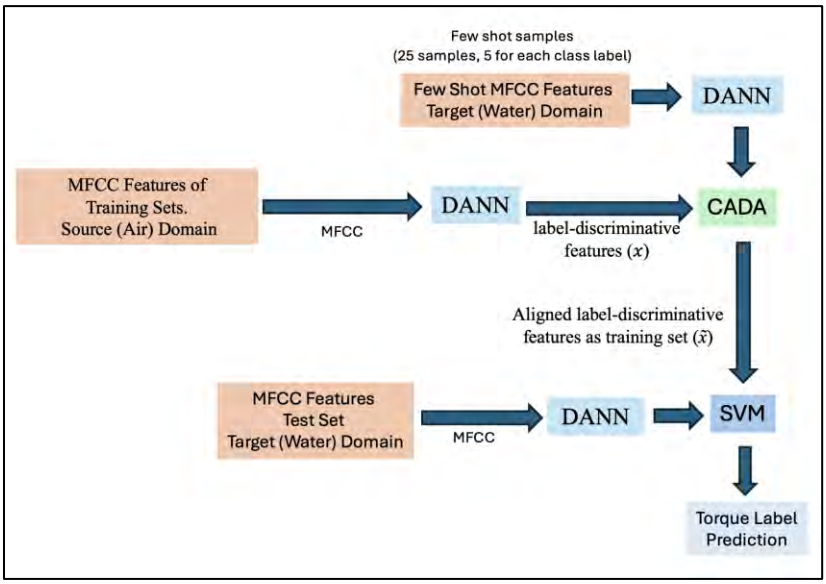


Fig. 77. Proposed Air-to-Water Domain Adaptation Architecture.

V.7.6 **Dataset Collection, Arrangement Class Reorganization.**

Datasets were collected under both air and water conditions using a pair of 6-inch flanges and five torque levels: 0, 10, 20, 30, and 40 ft-lbs.

Five datasets (Set 1–Set 5) were collected in air, each containing 40 samples per torque class, totaling 200 samples per set (Table 27).

Six datasets (Set 1–Set 6) were collected underwater, also with 40 samples per torque level, totaling 200 samples per set (Table 28).

Table 27. Dataset Arrangement in Air for 6-inch Flange

Dataset ID	0 ft-lbs	10 ft-lbs	20 ft-lbs	30 ft-lbs	40 ft-lbs	Sample count	Medium
Set 1	40	40	40	40	40	200	Air
Set 2	40	40	40	40	40	200	Air
Set 3	40	40	40	40	40	200	Air
Set 4	40	40	40	40	40	200	Air
Set 5	40	40	40	40	40	200	Air

Table 28. Dataset Arrangement Underwater for 6-inch Flange.

Dataset ID	0 ft-lbs	10 ft-lbs	20 ft-lbs	30 ft-lbs	40 ft-lbs	Sample count	Medium
Set 1	40	40	40	40	40	200	Water
Set 2	40	40	40	40	40	200	Water
Set 3	40	40	40	40	40	200	Water
Set 4	40	40	40	40	40	200	Water
Set 5	40	40	40	40	40	200	Water
Set 6	40	40	40	40	40	200	Water

Although we have five classes (preload levels) for each dataset, sometimes in actual projects, it is unnecessary to have so many categories. Therefore, we simplified the 5-class task to 2-class, 3-class, and 4-class. (Fig. 78).

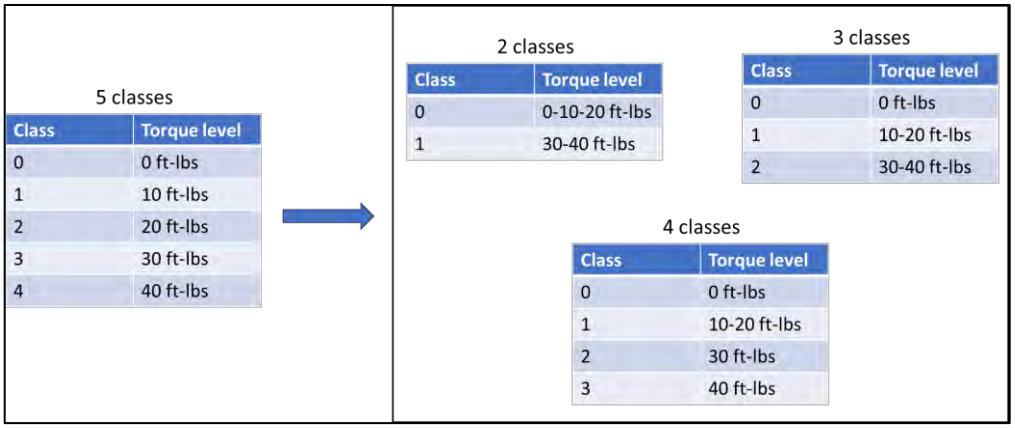


Fig. 78. Class Reorganization.

V.7.7 Experimental Result and Discussion

The performance of the proposed domain adaptation pipeline—combining DANN and CADA—was evaluated using the test datasets from the target (water) domain. The model was trained using five datasets from the source (air) domain (Set 1 to Set 5), while 25 few-shot samples from the target domain were used to perform centroid alignment (CADA). The evaluation was carried out using various class groupings: 4-class, 3-class, and 2-class configurations.

Table 29 presents classification accuracy results under different label configurations. Here are our observations:

- The model achieved an average accuracy of 87% in the 4-class classification setting, indicating strong generalization after domain adaptation.
- Performance improved further as the number of classes decreased: 88% average accuracy for the 3-class case and 98% for the binary (2-class) classification.
- Notably, in the 2-class setting, five out of six test configurations achieved near or full 100% accuracy.

Detailed confussion matrices, precision, recall, and F1-scores for each class are provided in the Fig. A15 to A32 in Appendix III.

Table 29. Classification Results After Domain Adaptation from Air to Water

Training Set (Air)	Few-Shot Set (Water, 25 samples)	Validation Set (Water)	Test Set (Water)	Accuracy – 4 Class	Accuracy – 3 Class	Accuracy – 2 Class
Set 1 to Set 5	Set 6	Set 6	Set 1	85%	85%	100%
Set 1 to Set 5	Set 6	Set 6	Set 2	97%	98%	100%
Set 1 to Set 5	Set 2	Set 2	Set 3	87%	92%	100%
Set 1 to Set 5	Set 2	Set 2	Set 4	91%	92%	100%
Set 1 to Set 5	Set 3	Set 3	Set 5	77%	78%	99.5%
Set 1 to Set 5	Set 3	Set 3	Set 6	83%	83%	86%
Average ± Std				86.7 ± 6.2 %	88.0 ± 6.6 %	97.5 ± 5.1 %

V.7.8 Traning and validation loss curve of the DANN model.

Fig. 79 and Fig. 80 show the training and validation performance of the domain-adaptation model in terms of classification and domain discrimination. The class (task) accuracy on the training set increases steadily and converges close to 1.0, while the corresponding loss decreases to near zero, indicating effective learning of the classification task. In contrast, the domain classifier exhibits oscillatory accuracy and relatively high loss, which is characteristic of adversarial training where the feature extractor attempts to confuse the domain discriminator and promote domain-invariant features. On the validation data, the class accuracy stabilizes around 0.75–0.85 with moderate loss values, suggesting reasonable generalization to unseen underwater datasets despite a noticeable gap from training performance. Overall, the curves indicate stable convergence of the task network and successful adversarial interaction between the feature extractor and domain classifier, leading to partial alignment between air and water domains.

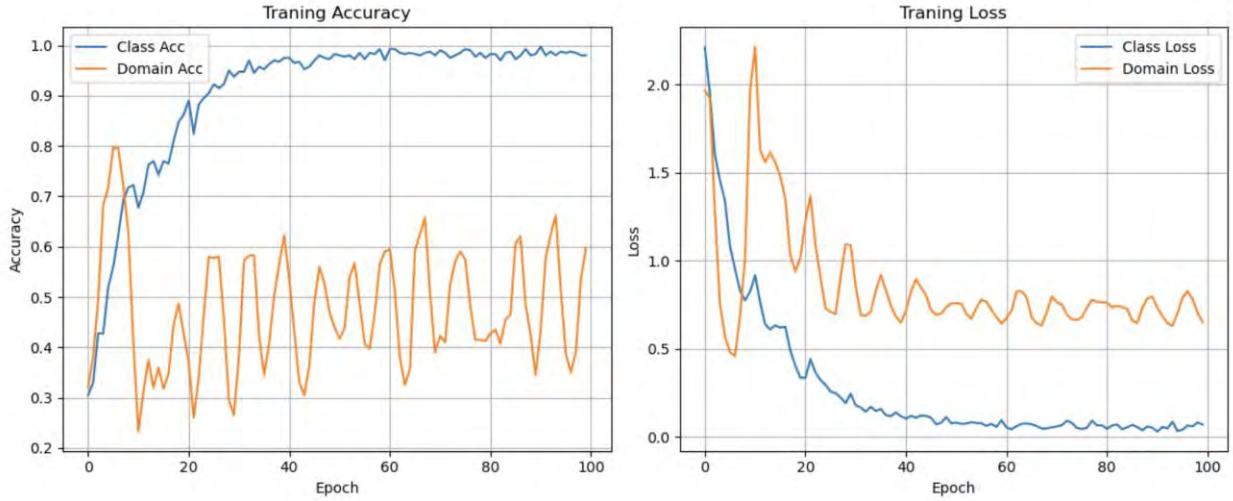


Fig. 79. Traning Curves of DANN

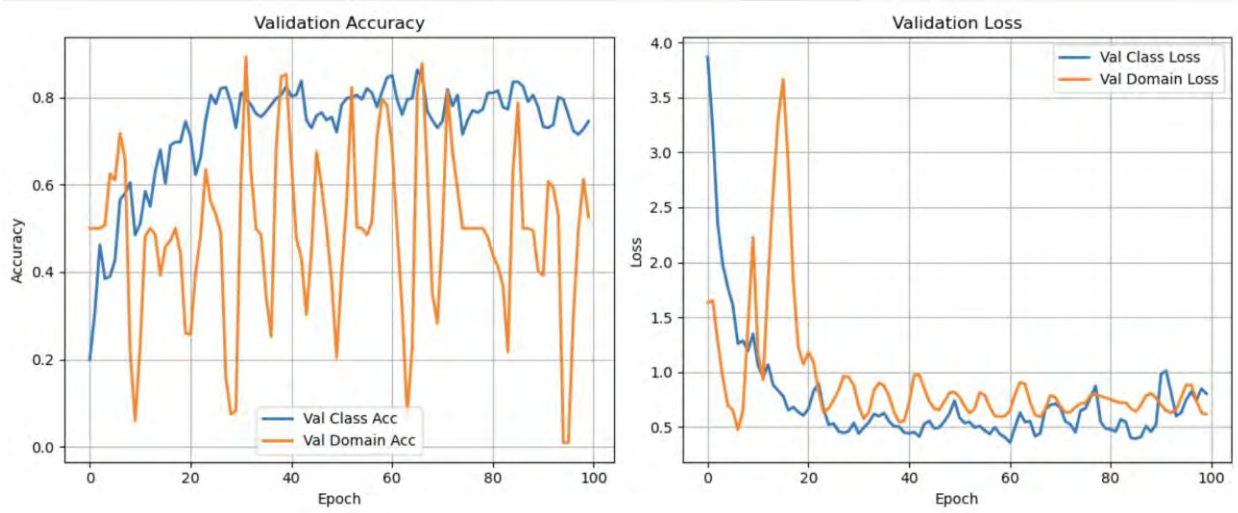


Fig. 80. Validation Curves of DANN

V.7.9 Summary of Findings

The proposed air to water domain adaptation method combining DANN and CADA achieves the following:

- **Substantial reduction in underwater data requirements:** The model leverages abundant air-domain data and only a small number of labeled water-domain samples (few-shot) to achieve high classification accuracy in water.
- **Effective handling of both covariate and concept shift:**
  - DANN provides domain-invariant yet label-discriminative features, mitigating covariate shifts.
  - CADA aligns class-wise centroids, directly addressing concept shift under few-shot supervision.
- **High performance across class configurations:**
  - Average accuracies of **87% (4-class)** and **88% (3-class)** show good robustness for moderate label granularity.
  - Near-perfect **98% (2-class)** accuracy demonstrates the method's suitability for practical "loose vs tight" monitoring scenarios.

Overall, this framework provides a **practical and efficient approach** for adapting torque classification models trained in air to operate reliably underwater.

VI. CONCLUSION

The autonomous control system successfully integrates advanced computer vision techniques to achieve precise alignment and approach for subsea intervention tasks, specifically targeting flange grasping on pipelines. By utilizing the Hough Transform for robust pipeline orientation and YOLOv8 for accurate flange localization, the system reliably derives the necessary positional data (angle and distance errors). The three-step control strategy—Parallelization (yaw correction), Centering (lateral correction), and Approach (forward movement)—ensures that the ROV achieves optimal spatial configuration, minimizing angular and translational errors before initiating the final grasp. This structured, vision-based feedback loop represents a critical step toward fully autonomous, highly accurate subsea manipulation, significantly reducing operational time and dependence on manual piloting in complex environments.

We proposed an end-to-end autonomous underwater inspection framework that integrates vision-based stabilization, sonar-based perception, and image-based manipulation for flange smart-touch inspection. During long-range sensing, an AprilTag-based visual feedback controller was employed to maintain precise station-keeping while the Ping360 single-beam scanning sonar performed a full 360° sweep. The stabilized platform enabled high-quality sonar data acquisition and reliable map construction using RANSAC-based line fitting for pipeline detection and orientation estimation. A LOS guidance strategy was then implemented for coarse path planning and tracking, allowing the ROV to approach the pipeline smoothly under sonar guidance. Once the pipeline entered the visual sensing range, the system automatically switched to an IBVS controller for fine alignment, approach, and flange grasping. The integrated PZT sensors embedded in the gripper successfully performed active smart-touch inspection upon contact. Experimental results demonstrated that the proposed system achieved stable AprilTag-based station-keeping, consistent sonar map reconstruction, accurate sonar-guided navigation, and robust vision-based flange manipulation in a controlled underwater environment.

The redesigned robotic gripper fundamentally improved the quality and reliability of vibration-based flange inspection. Compared to the earlier design, the new configuration delivered significantly higher signal energy, better coupling between the PZT transducers and the flange surface, and more stable repeatability across datasets. These improvements directly enabled stronger feature separability in MFCC space and noticeably higher classification accuracy across all flange sizes. In short, the new gripper design eliminated the major hardware bottleneck that had previously limited model performance, allowing subsequent machine-learning methods (SVM, PINN, and domain-adaptation models) to fully leverage the physical information embedded in the stress-wave responses.

The MFCC + SVM framework proved to be a fast, lightweight, and highly reliable baseline for bolt-tightness classification across a broad range of flange sizes (4.25-inch to 25-inch) and load conditions. Despite its simplicity, the method consistently achieved high accuracy—typically above 90% for most flanges—and generalized well to non-uniform load distributions. The results demonstrate that MFCCs contain sufficient discriminative information for separating preload classes, and that SVMs with RBF kernels can exploit these patterns even under moderate noise. The main limitation is reducing performance when the number of torque classes increases or when bolt count is small, where structural interactions cause neighboring bolts to influence each other. Nevertheless, MFCC + SVM remains a strong, robust baseline that sets a clear performance benchmark for more advanced deep-learning methods.

The domain-adaptation studies demonstrated that both mechanical-model‑based transformation and data-driven feature alignment can substantially improve bolt-looseness classification when target-domain data are scarce. For size-based adaptation (6-inch → 9-inch), the frequency-response-function (FRF)-based signal transformation accurately converted 6-inch responses into their 9-inch equivalents, enabling the SVM classifier to achieve approximately 94% accuracy in the 2-class and 3-class tasks and 82.5% and 79% accuracy in the 4-class and 5-class tasks, respectively. For environment adaptation (Air → Water), the combined DANN + CADA framework successfully addressed both covariate and concept shift, reaching 87% (4-class), 88% (3-class), and 98% (2-class) mean accuracy while using only 25 labeled underwater samples. Collectively, these methods demonstrate that effective domain adaptation—whether through physics-based transformations or few-shot learning—can dramatically reduce the volume of required target-domain data while maintaining strong torque-classification performance across different flange sizes and sensing environments.

VII. PUBLICATIONS

Journal Papers:

1. D. Koc, Q. Zhu, F. Ghorbel, and Z. Chen, "Experimental Validation of Underwater Depth and Orientation Control Using Reversible Fuel Cell Electrolysis", *IEEE/ASME Transactions on Mechatronics*, vol. 30, no. 6, pp. 7649-7658, Dec. 2025, doi:10.1109/TMECH.2025.3579654.
2. Q. Zhu, and Z. Chen, "Visual Servoing Control of Remotely Operated Vehicle for Autonomous Flange Bolts Looseness Inspection", submitted to *IEEE Journal on Oceanic Engineering*, under review, 2026.
3. Q. Zhu, Y Ying, and Z. Chen, "Underwater Flange Detection and Grasping Using Sonar-to-Vision Sensing", submitted to *IEEE/ASME Transactions on Mechatronics*, under review, 2026.
4. T. Kaaya, Y. Ying, W. Zuo, F. Ghorbel, and Z. Chen, " Control of AUVs using Buoyancy-Induced Soft Actuation Enabled by Reversible Fuel Cells", submitted to *IEEE Journal of Oceanic Engineering*, Special Issue on Autonomous Underwater Vehicle, under review, 2026.

Conference Papers:

5. Qiang Zhu, Yuhang Ying, Zheng Chen, "Low-Cost Underwater Mapping via Single-Beam Sonar and Inertial Fusion", *Proc. Of the 5th Modeling, Estimation and Control Conference (MECC)*, Pittsburg, PA, October 5-8, pp. 1000-1005, 2025.
6. Y. Ying, Q. Zhu and Z. Chen, "A Physics-Informed Neural Network Enhanced Kalman Filter Method for Inertial Navigation of Remotely Operated Vehicles", *Proc. Of the 5th Modeling, Estimation and Control Conference (MECC),* Pittsburg, PA, October 5-8, pp. 485-490, 2025.
7. Denizcan Koc, Fathi H. Ghorbel, Zheng Chen, "Modeling and Control of Jellyfish Inspired Robot Enabled by Soft and Hard Actuators", *Proc. Of the 5th Modeling, Estimation and Control Conference (MECC)*, Pittsburg, PA, October 5-8, pp. 126-131,2025.
8. Muhammad Umar Masood, Zheng Chen, "Towards Underwater Swarm: A Relative Localization Framework", *Proc. Of the 5th Modeling, Estimation and Control Conference (MECC)*, Pittsburg, PA, October 5-8, pp. 251-256, 2025.

Patents:

- Gangbing Song, Jinwei Jiang, Siu Chun Michael Ho, Zheng Chen, "Portable system for PZT-based inspection of bolted connections", US Patent, 11940418, March 26, 2024.

REFERENCES

Ajakan, H. e. (2014). Domain-adversarial neural networks.

Allemang, R. J. (2022). Frequency response function estimation techniques and the corresponding coherence functions. *Mechanical Systems and Signal Processing.*

Ballard, D. H. (1981). Generalizing the Hough transform to detect arbitrary shapes. *Pattern recognition*, 111-122.

Derpanis, K. G. (2010). Overview of the RANSAC Algorithm. *Image Rochester NY*, 2-3.

Fossen, T. I. (2023). Feedback error-state Kalman filter with time-delay compensation for hydroacoustic-aided inertial navigation of underwater vehicles. *Control Engineering Practice*, 105603.

Karniadakis, G. E. (2021). Physics-informed machine learning. *Nature Reviews Physics*, 422-440.

Keow, A. a.-S. (2020). PIDA control of buoyancy device enabled by water electrolysis. *IEEE/ASME Transactions on Mechatronics*, 1202-1210.

Kull, M. a. (2014). Patterns of dataset shift. *First international workshop on learning over multiple contexts (LMCE) at ECML-PKDD.*

Michel, O. Cyberbotics ltd. webots™. (2004). *professional mobile robot simulation. International Journal of Advanced Robotic Systems.* sagepub.

Nair, N. G. (2019). Covariate shift: A review and analysis on classifiers. *global conference for advancement in technology (GCAT).*

Nguyen, T., & Shin, Y. (2019). An efficient RSS localization for underwater wireless sensor networks. *Sensors, 19*(14). doi:10.3390/s19143105

Olson, E. (2011). AprilTag: A robust and flexible visual fiducial system. *2011 IEEE international conference on robotics and automation* (pp. 3400-3407). IEEE.

Parnami, A. a. (2022). Learning from few examples: A summary of approaches to few-shot learning.

Pourkabirian, A., Kooshki, F., Anisi, M., & Jindal, A. (2023). An accurate RSS/AoA-based localization method for internet of underwater things. *Ad Hoc Networks, 145*, 103177. doi:10.1016/j.adhoc.2023.103177

Qian, K. T. (2023). *Frequency-domain physical constrained neural network for nonlinear system dynamic prediction. Engineering Applications of Artificial Intelligence.*

RAUTELA, M. R. (2021). Simulation of guided waves for structural health monitoring using physics-informed neural networks. *Structural health monitoring* .

Redmon, J. a. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779--788). IEEE.

Ryecroft, S., Shaw, A., Fergus, P., Kot, P., Hashim, K., Moody, A., & Conway, L. (2019). A first implementation of underwater communications in raw water using the 433 MHz frequency combined with a bowtie antenna. *Sensors, 19*(8). doi:10.3390/s19081813

Segal, A. a. (2009). Generalized-icp. *Robotics: science and systems*, 435.

Shaaban, G. F. (2024). *Q-learning-based noise covariance matrices adaptation in kalman filter for inertial navigation.* IFAC.

Wielandner, L., Leitinger, E., & Witrisal, K. (2021). RSS-Based Cooperative Localization and Orientation Estimation Exploiting Antenna Directivity. *IEEE Access, 9,* 53046–53060. doi:10.1109/ACCESS.2021.3070441

Wu, Y. S. (2024). Physics-informed machine learning: A comprehensive review on applications in anomaly detection and condition monitoring. *Expert Systems with Applications*.

Zhang, X. C. (2021). A physics-informed neural network for creep-fatigue life prediction of components at elevated temperatures. *Engineering Fracture Mechanics*.

APPENDIX I

Fig. A1. Confusion matrices of the 6-inch flange.

Dataset 1

	Precision	Recall	F1Score
Class1	1	0.95	0.974
Class2	0.976	1	0.988
Class3	0.889	1	0.941
Class4	1	0.875	0.933
Macro	0.96612	0.95625	0.95913

Dataset 2

	Precision	Recall	F1Score
Class1	1	1	1
Class2	1	1	1
Class3	0.941	1	0.97
Class4	1	0.938	0.968
Macro	0.98529	0.98438	0.98436

Dataset 3

	Precision	Recall	F1Score
Class1	0.976	1	0.988
Class2	1	0.988	0.994
Class3	0.909	1	0.952
Class4	1	0.9	0.947
Macro	0.97118	0.97187	0.97028

Dataset 4

	Precision	Recall	F1Score
Class1	0.889	1	0.941
Class2	1	0.938	0.968
Class3	1	0.293	0.453
Class4	0.574	1	0.729
Macro	0.8656	0.80755	0.77268

Dataset 5

	Precision	Recall	F1Score
Class1	0.816	1	0.899
Class2	1	0.888	0.94
Class3	0.851	1	0.92
Class4	1	0.825	0.904
Macro	0.91685	0.92813	0.91573

Dataset 6

	Precision	Recall	F1Score
Class1	1	0.8	0.889
Class2	0.909	1	0.952
Class3	1	1	1
Class4	1	1	1
Macro	0.97727	0.95	0.96032

Dataset 7

	Precision	Recall	F1Score
Class1	0.909	1	0.952
Class2	1	0.938	0.968
Class3	0.952	1	0.976
Class4	1	0.975	0.987
Macro	0.96537	0.97813	0.97077

Dataset 8

	Precision	Recall	F1Score
Class1	0.968	1	0.984
Class2	1	0.983	0.992
Class3	1	1	1
Class4	1	1	1
Macro	0.99194	0.99583	0.9938

Dataset 9

	Precision	Recall	F1Score
Class1	0.975	0.642	0.774
Class2	0.847	0.992	0.914
Class3	1	0.183	0.31
Class4	0.55	1	0.71
Macro	0.84303	0.70417	0.67685

Dataset 10

	Precision	Recall	F1Score
Class1	1	0.9	0.947
Class2	0.941	1	0.97
Class3	1	0.975	0.987
Class4	1	1	1
Macro	0.98529	0.96875	0.9761

Fig. A2. Precision, recall, and F1-score of Confusion matrices of the 6-inch flange. The classes 1, 2, 3, and 4 correspond to the classes 0, 1, 2, and 3 in Fig. A1, respectively.

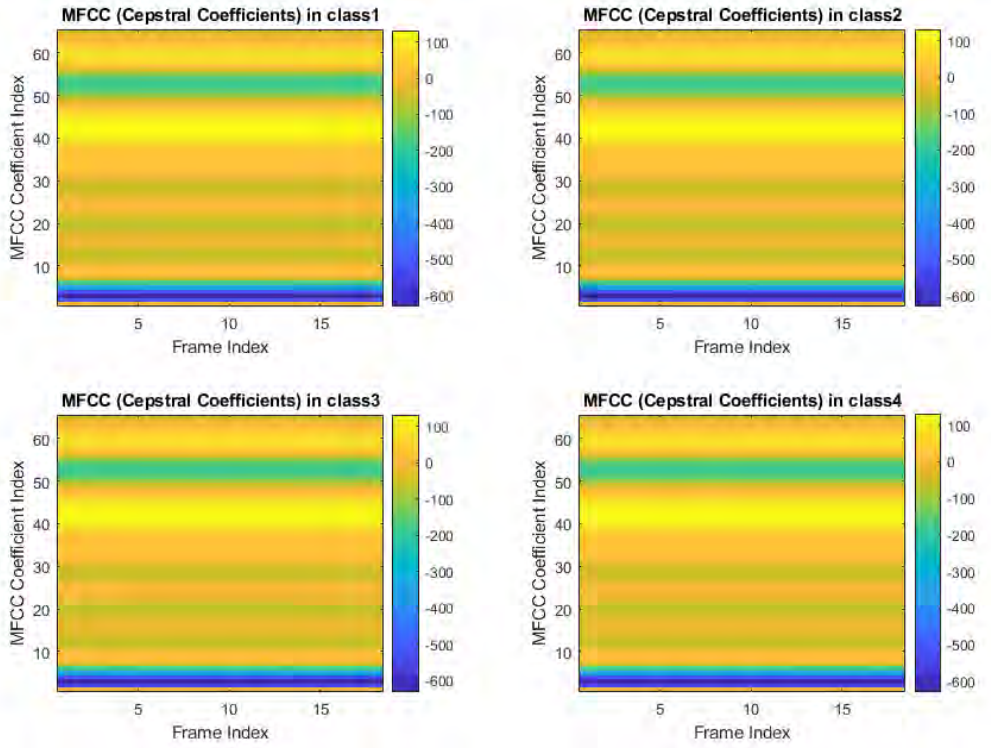


Fig. A3. Mel-frequency Cepstral Coefficients of signals (from dataset 1) in different preload levels. The classes 1, 2, 3, and 4 correspond to the classes 0, 1, 2, and 3 in Fig. A1, respectively.

Fig. A4. Confusion matrices of the 9-inch flange.

Dataset 1

	Precision	Recall	F1Score
Class1	1	1	1
Class2	0.952	1	0.976
Class3	1	0.725	0.841
Class4	0.851	1	0.92
Macro	0.95086	0.93125	0.93393

Dataset 2

	Precision	Recall	F1Score
Class1	1	1	1
Class2	1	1	1
Class3	0.714	1	0.833
Class4	1	0.6	0.75
Macro	0.92857	0.9	0.89583

Dataset 3

	Precision	Recall	F1Score
Class1	1	1	1
Class2	1	1	1
Class3	1	0.95	0.974
Class4	0.952	1	0.976
Macro	0.9881	0.9875	0.98749

Fig. A5. Precision, recall, and F1-score of Confusion matrices of the 9-inch flange. The classes 1, 2, 3, and 4 correspond to the classes 0, 1, 2, and 3 in Fig. A4, respectively.

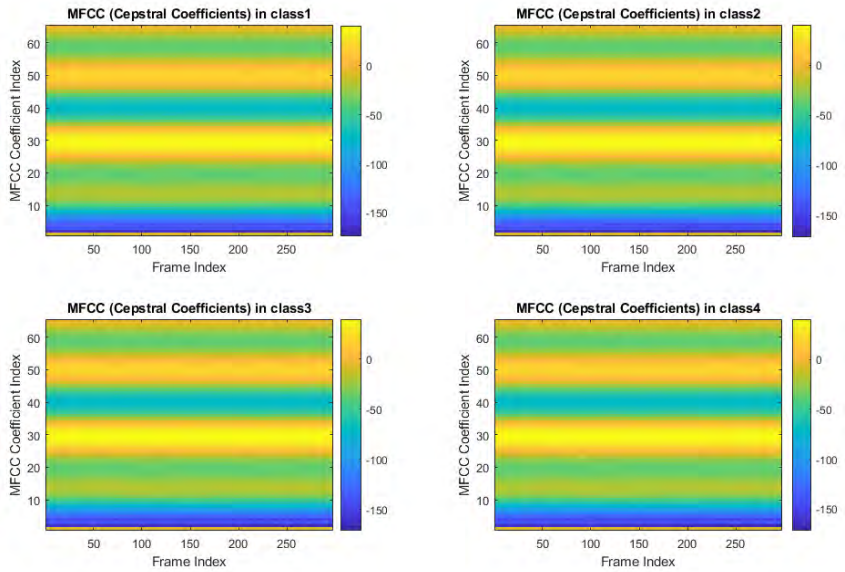


Fig. A6. Mel-frequency Cepstral Coefficients of signals (from dataset 1) in different preload levels. The classes 1, 2, 3, and 4 correspond to the classes 0, 1, 2, and 3 in Fig. A4, respectively.

Fig. A7. Confusion matrices of the 11-inch flange.

Dataset 1

	Precision	Recall	F1Score
Class1	1	1	1
Class2	0.702	1	0.825
Class3	1	0.575	0.73
Macro	0.90058	0.85833	0.85163

Dataset 2

	Precision	Recall	F1Score
Class1	1	1	1
Class2	1	0.8	0.889
Class3	0.833	1	0.909
Macro	0.94444	0.93333	0.93266

Dataset 3

	Precision	Recall	F1Score
Class1	1	0.875	0.933
Class2	0.889	1	0.941
Class3	1	1	1
Macro	0.96296	0.95833	0.95817

Fig. A8. Precision, recall, and F1-score of Confusion matrices of the 11-inch flange. The Classes 1, 2, and 3 correspond to the classes 0, 1, and 2 in Fig. A7, respectively.

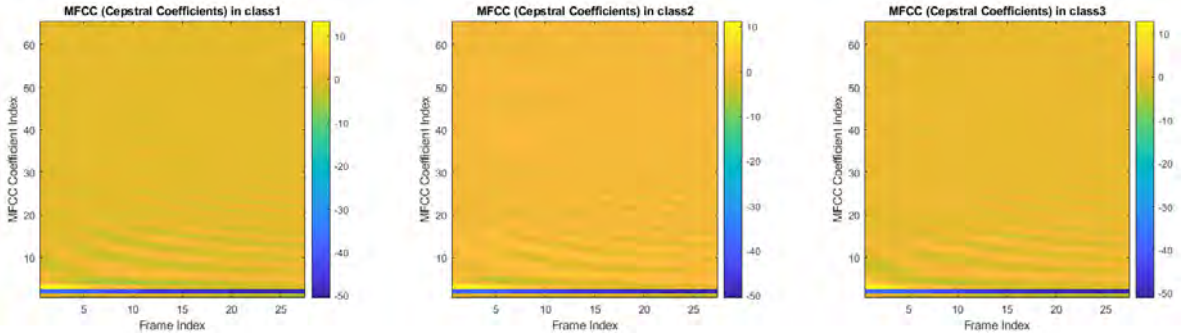


Fig. A9. Mel-frequency Cepstral Coefficients of signals (from dataset 1) in different preload levels. The Classes 1, 2, and 3 correspond to the classes 0, 1, and 2 in Fig. A7, respectively.

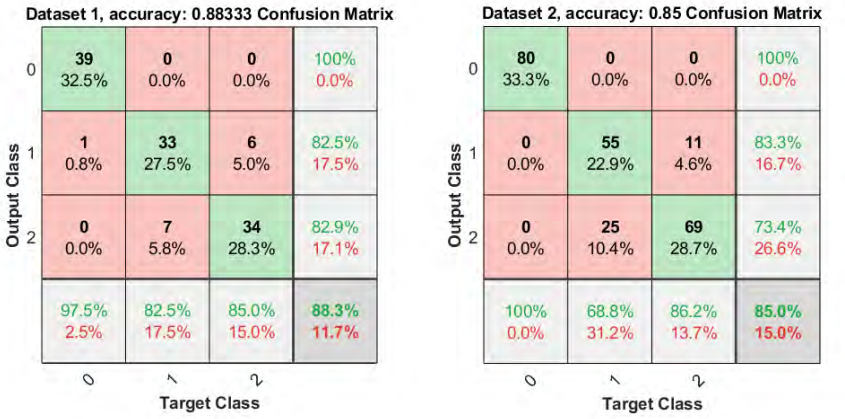


Fig. A10. Confusion matrices of the 13-inch flange.

Dataset 1

	Precision	Recall	F1Score
Class1	1	0.975	0.987
Class2	0.825	0.825	0.825
Class3	0.829	0.85	0.84
Macro	0.88476	0.88333	0.88395

Dataset 2

	Precision	Recall	F1Score
Class1	1	1	1
Class2	0.833	0.688	0.753
Class3	0.734	0.863	0.793
Macro	0.85579	0.85	0.84884

Fig. A11. Precision, recall, and F1-score of Confusion matrices of the 13-inch flange. The Classes 1, 2, and 3 correspond to the classes 0, 1, and 2 in Fig. A10, respectively.

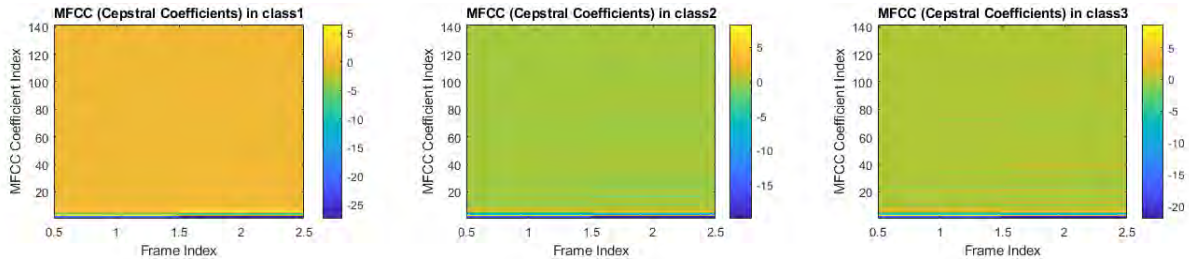


Fig. A12. Mel-frequency Cepstral Coefficients of signals (from dataset 1) in different preload levels. The Classes 1, 2, and 3 correspond to the classes 0, 1, and 2 in Fig. A10, respectively.

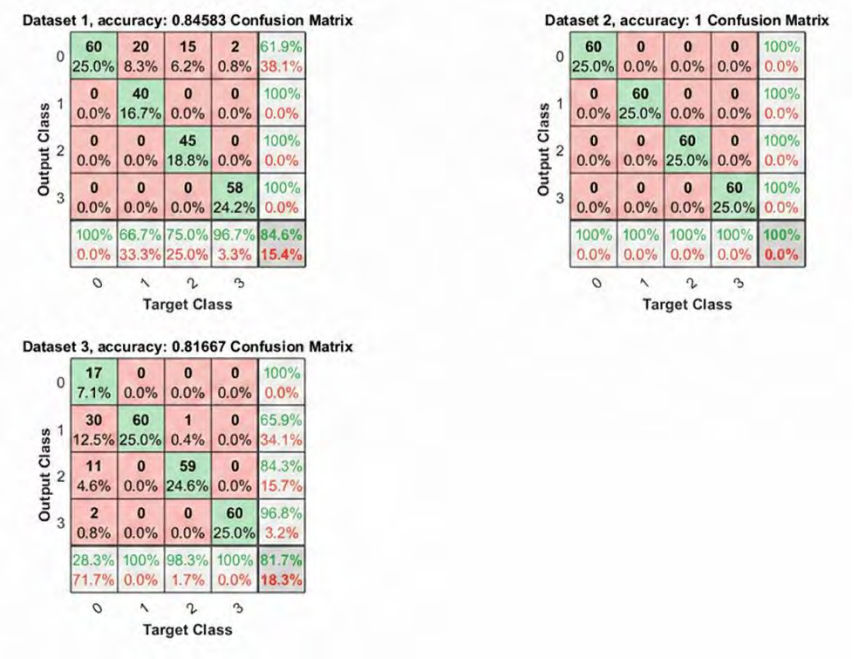


Fig. A13. Confusion matrices of the 16-inch flange.

Dataset 1

	Precision	Recall	F1Score
Class1	0.619	1	0.764
Class2	1	0.667	0.8
Class3	1	0.75	0.857
Class4	1	0.967	0.983
Macro	0.90464	0.84583	0.85113

Dataset 2

	Precision	Recall	F1Score
Class1	1	1	1
Class2	1	1	1
Class3	1	1	1
Class4	1	1	1
Macro	1	1	1

Dataset 3

	Precision	Recall	F1Score
Class1	1	0.283	0.442
Class2	0.659	1	0.795
Class3	0.843	0.983	0.908
Class4	0.968	1	0.984
Macro	0.86748	0.81667	0.78189

Fig. A14. Precision, recall, and F1-score of Confusion matrices of the 16-inch flange. The classes 1, 2, 3, and 4 correspond to the classes 0, 1, 2, and 3 in Fig. A13, respectively.

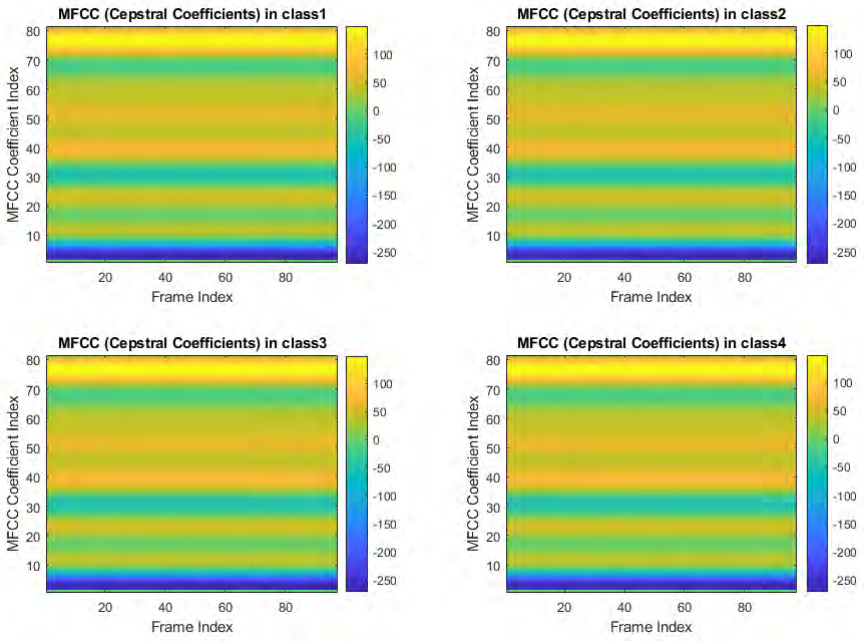


Fig. A15. Mel-frequency Cepstral Coefficients of signals (from dataset 1) in different preload levels. The classes 1, 2, 3, and 4 correspond to the classes 0, 1, 2, and 3 in Fig. A13, respectively.

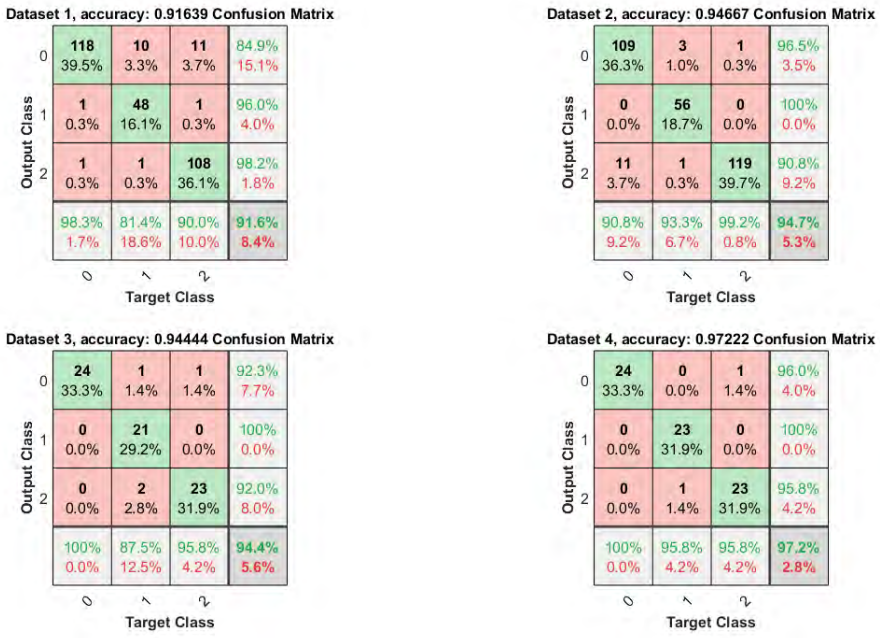


Fig. A16. Confusion matrices of the 21-inch flange.

Dataset 1

	Precision	Recall	F1Score
Class0	0.849	0.983	0.911
Class1	0.96	0.814	0.881
Class2	0.982	0.9	0.939
Macro	0.93025	0.89896	0.91035

Dataset 2

	Precision	Recall	F1Score
Class0	0.965	0.908	0.936
Class1	1	0.933	0.966
Class2	0.908	0.992	0.948
Macro	0.95767	0.94444	0.94978

Dataset 3

	Precision	Recall	F1Score
Class0	0.923	1	0.96
Class1	1	0.875	0.933
Class2	0.92	0.958	0.939
Macro	0.94769	0.94444	0.94404

Dataset 4

	Precision	Recall	F1Score
Class0	0.96	1	0.98
Class1	1	0.958	0.979
Class2	0.958	0.958	0.958
Macro	0.97278	0.97222	0.97222

Fig. A17. Precision, recall, and F1-score of Confusion matrices of the 21-inch flange.

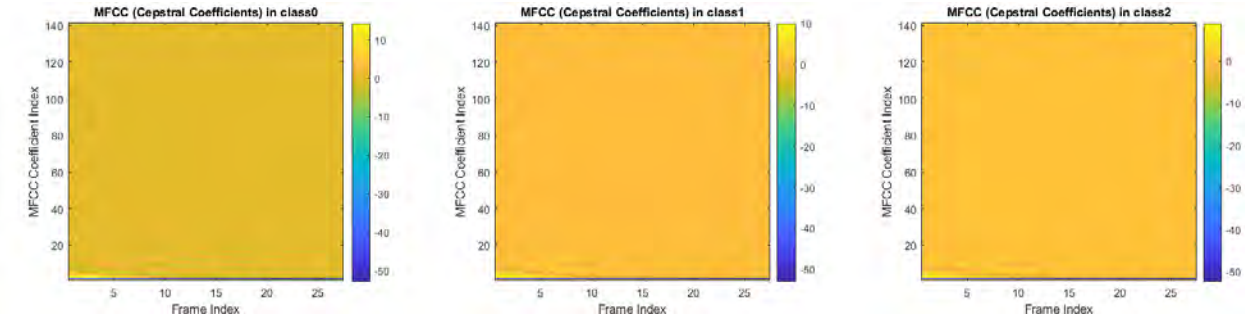


Fig. A18. Mel-frequency Cepstral Coefficients of signals (from dataset 1) in different preload levels.

APPENDIX II

Confusion matrices of the physics-informed neural network model

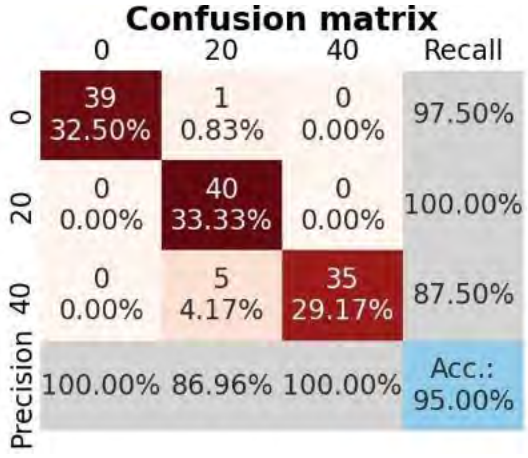


Fig. A1: Confusion matrix of set 1.

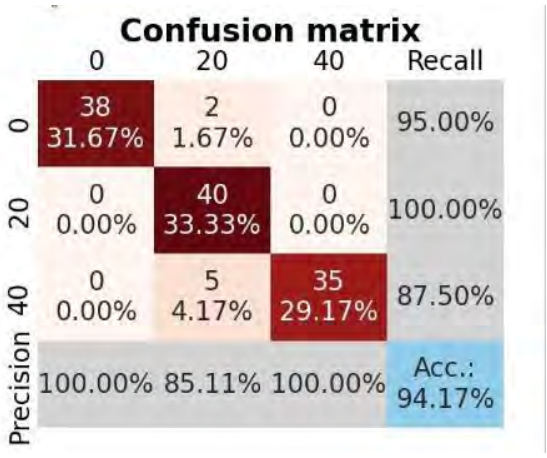


Fig. A2: Confusion matrix of set 2.

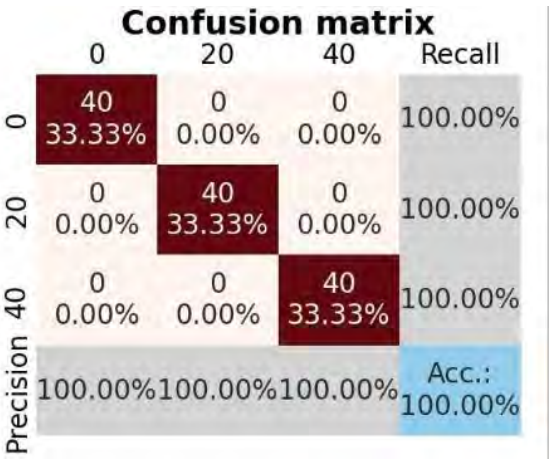


Fig. A3: Confusion matrix of set 3.

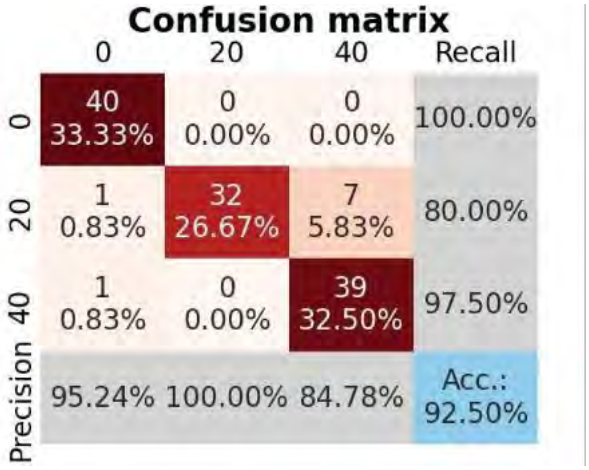
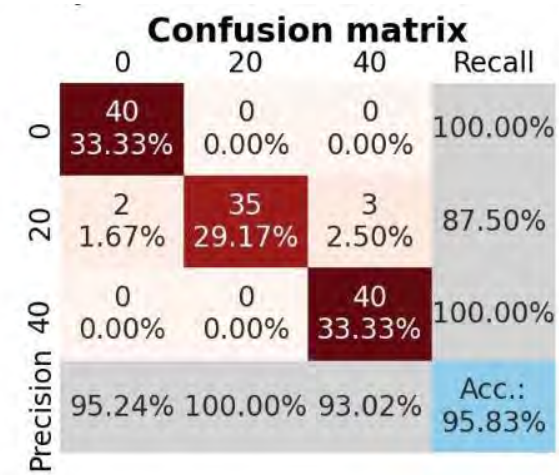


Fig. A4: Confusion matrix of set 4.



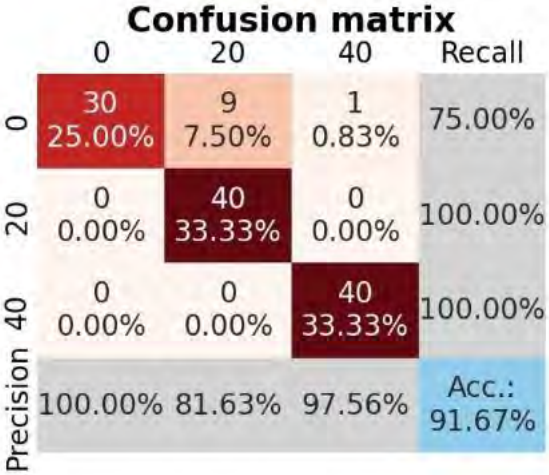


Fig. A5: Confusion matrix of set 5.

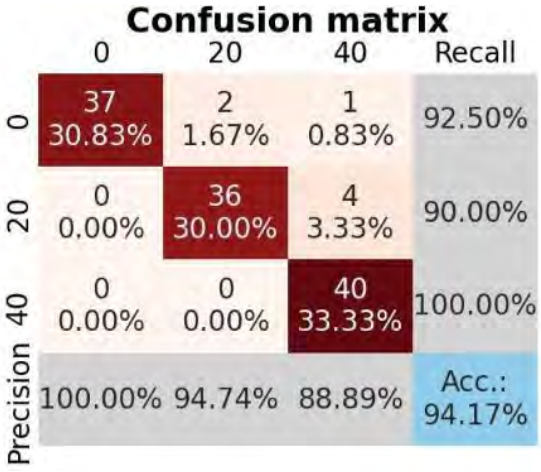


Fig. A6: Confusion matrix of set 6.

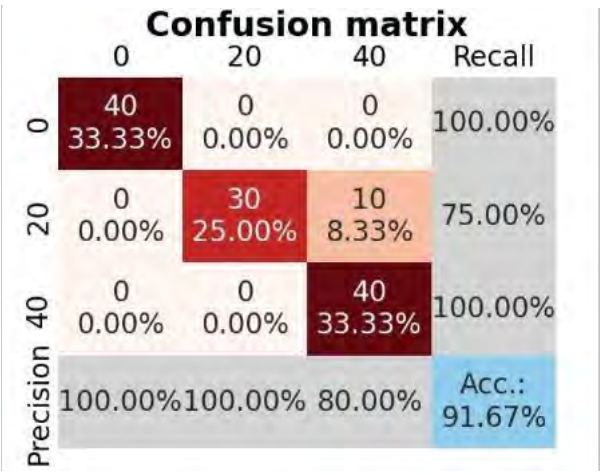


Fig. A7: Confusion matrix of set 7

Fig. A8: Confusion matrix of set 8

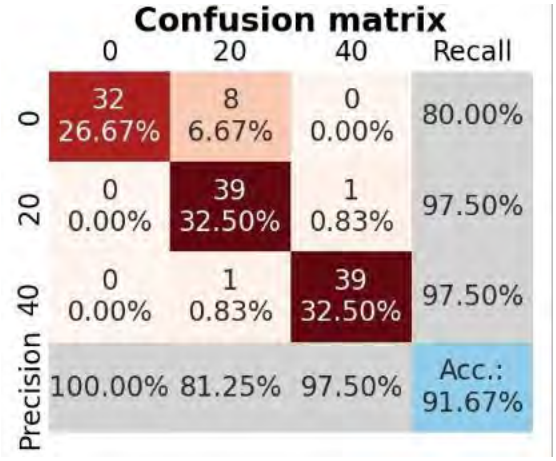


Fig. A9: Confusion matrix of set 9

Training and validation losses of the physics-informed neural network model

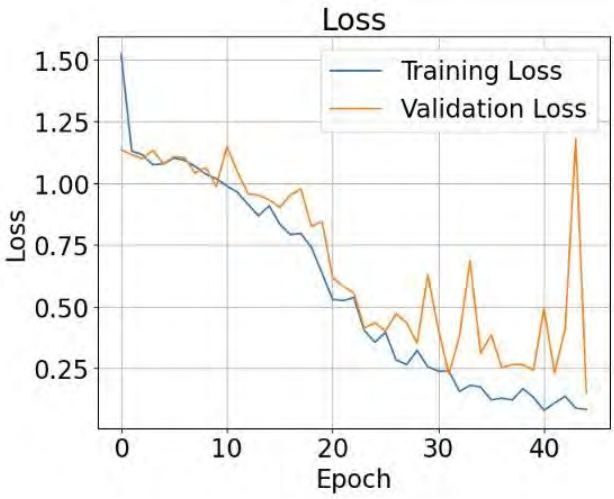


Fig. A10: Training & validation Loss of set 1

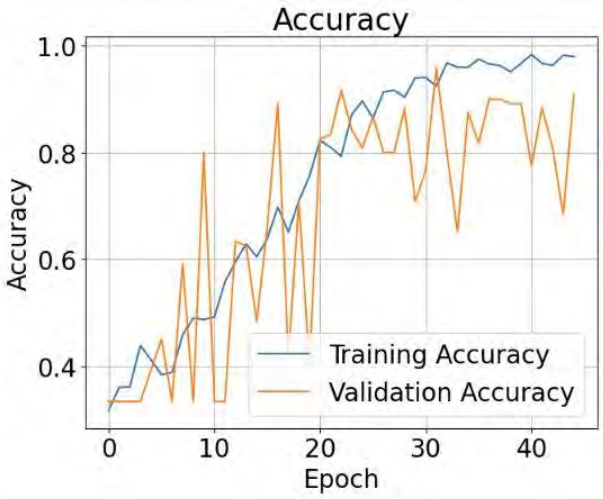


Fig. A11: Training & validation accuracy of set 1

Fig. A12: Training & validation Loss of set 2

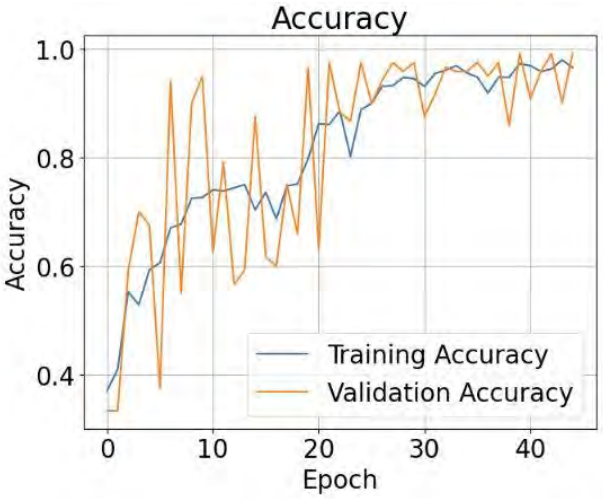


Fig. A13: Training & validation Accuracy of set 2

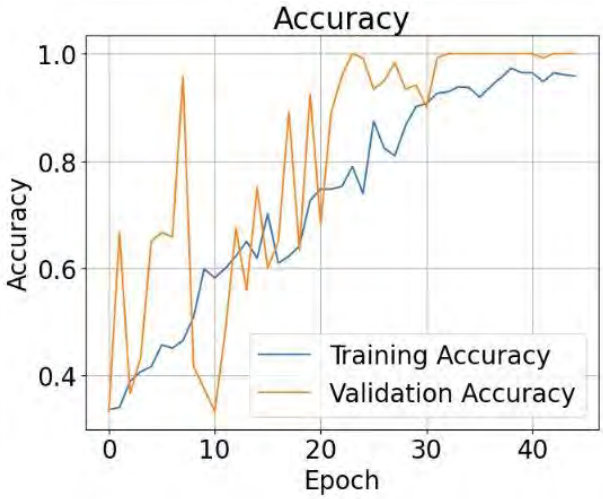


Fig. A14: Training & validation Loss of set 3

Fig. A15: Training & validation Accuracy of set 3

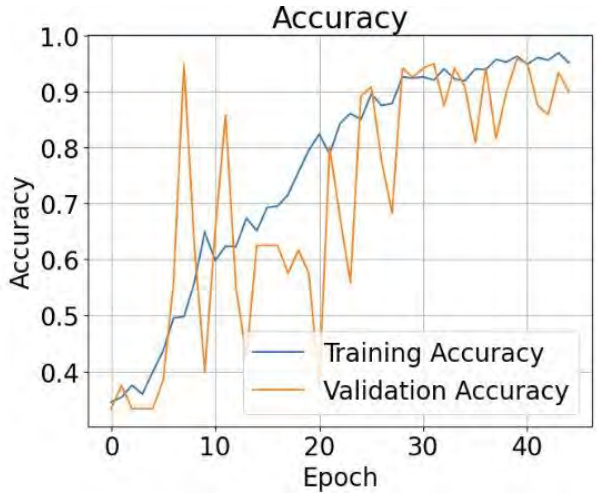
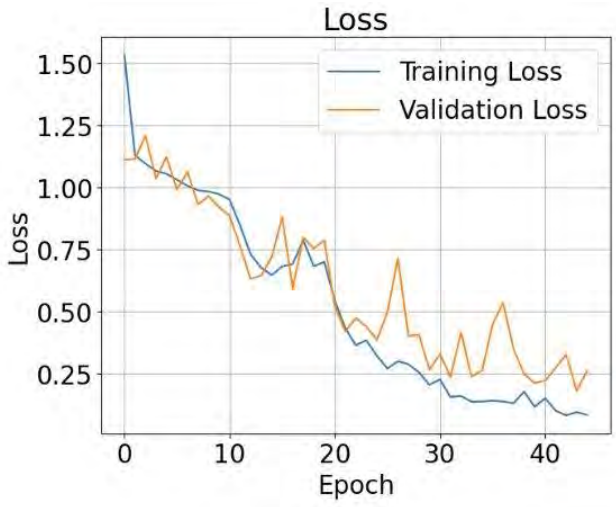


Fig. A16: Training & validation Loss of set 4

Fig. A17: Training & validation Accuracy of set 4



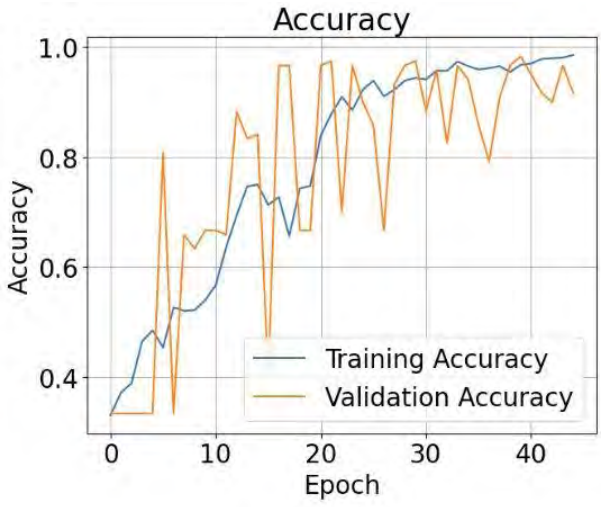


Fig. A18: Training & validation Loss of set 5

Fig. A19: Training & validation Accuracy of set 5

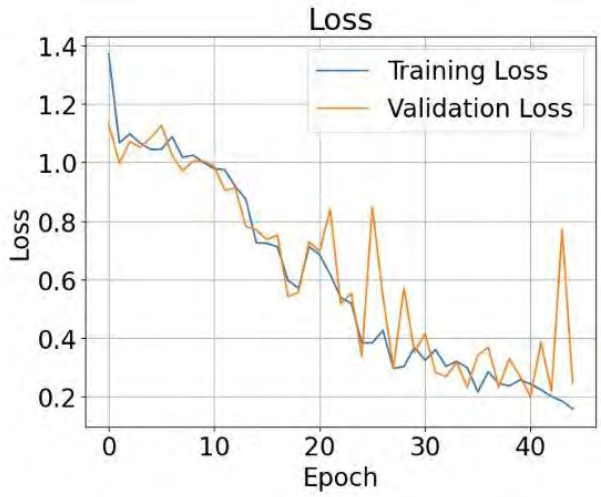


Fig. A20: Training & validation Loss of set 6

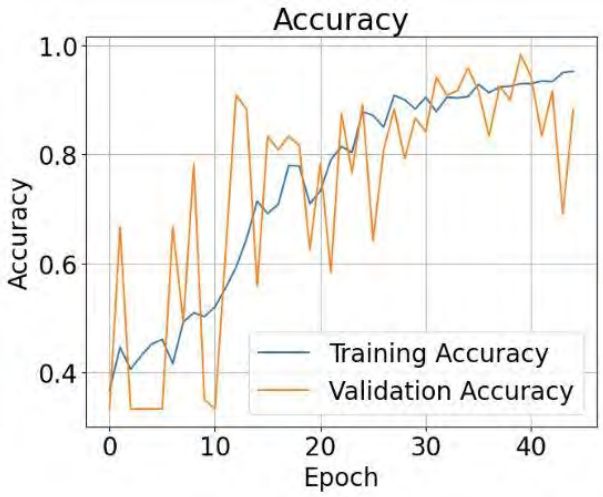


Fig. A21: Training & validation Accuracy of set 6

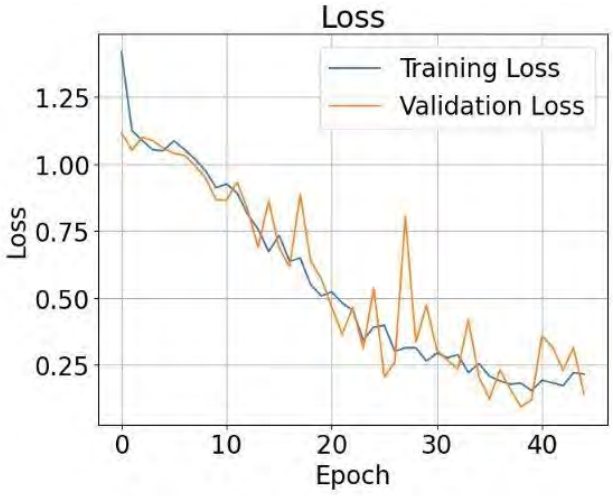


Fig. A22: Training & validation Loss of set 7

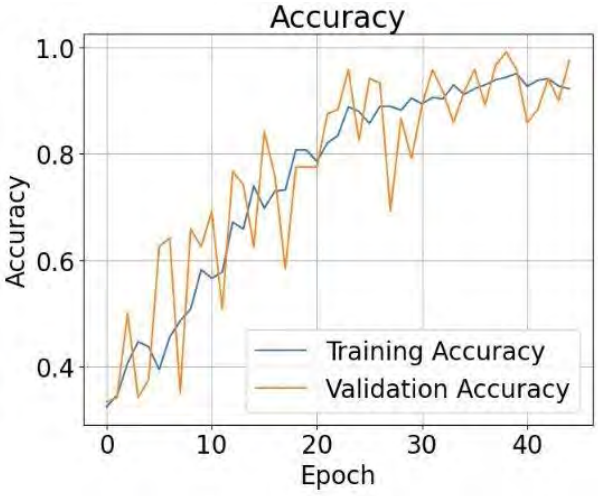


Fig. A23: Training & validation Accuracy of set 7

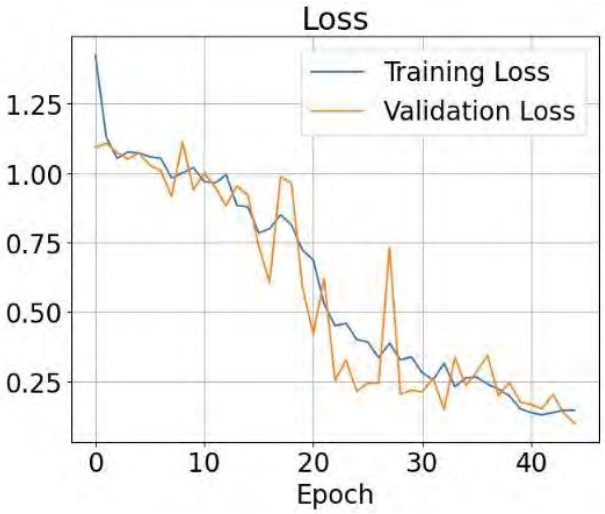


Fig. A24: Training & validation Loss of set 8

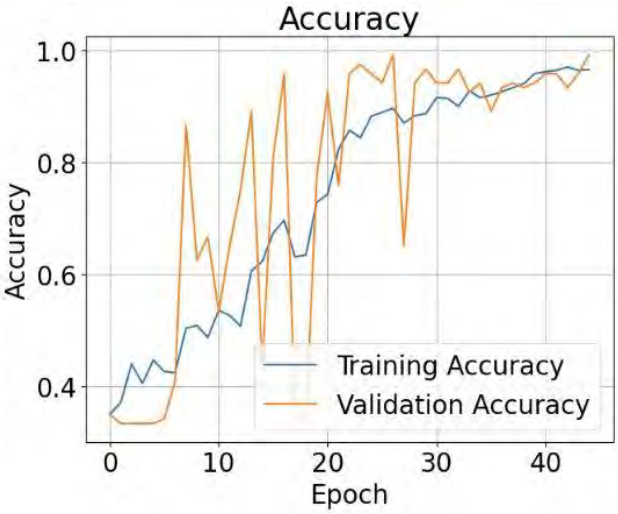


Fig. A25: Training & validation Accuracy of set 8

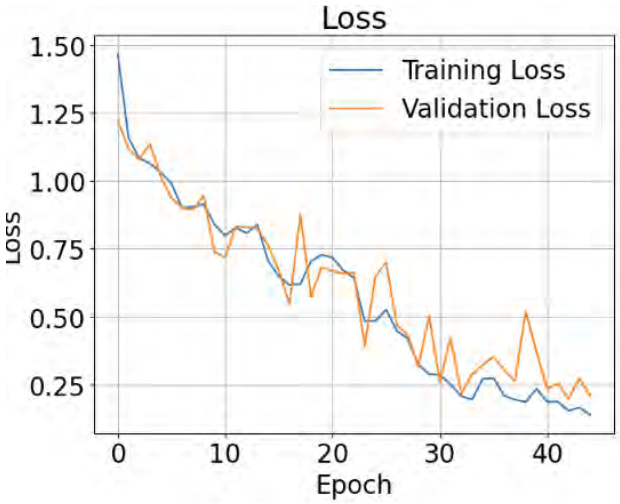


Fig. A26: Training & validation Loss of set 9

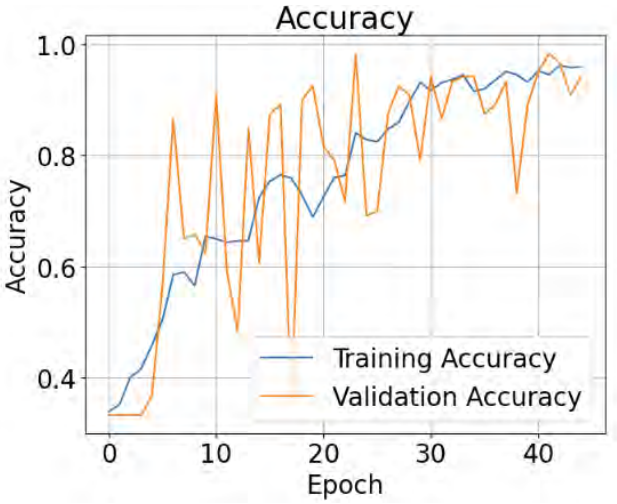


Fig. A27: Training & validation Accuracy of set 9

APPENDIX III

Confusion matrices of the 25-inch flange

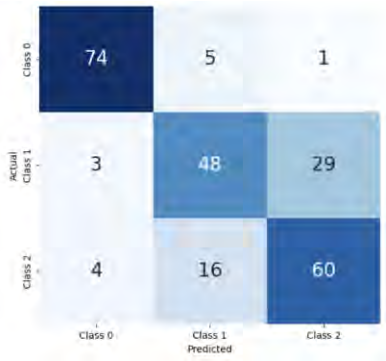


Fig. A1: Set 1

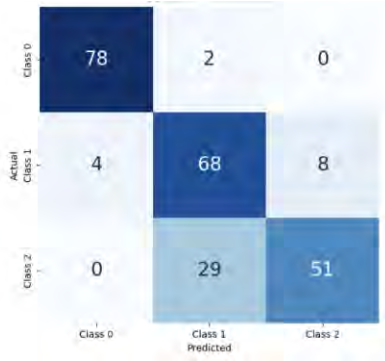


Fig. A2: Set 2

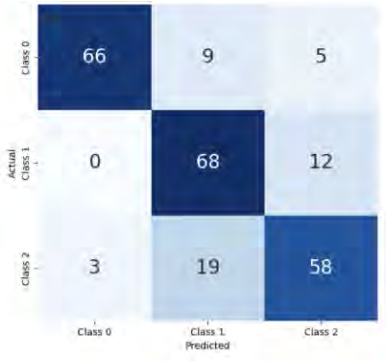


Fig. A3: Set 3

Set 1:

Class	Precision	Recall	F1-Score	Support
0	0.91	0.93	0.92	80
1	0.70	0.60	0.64	80
2	0.67	0.75	0.71	80
Accuracy	—	—	**0.75**	**240**
Macro Avg	0.76	0.76	0.76	240
Weighted Avg	0.76	0.76	0.76	240

Set 2:

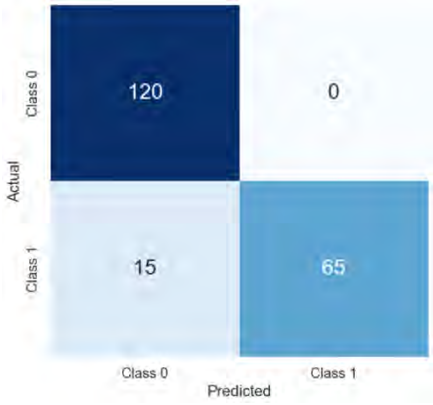
Class	Precision	Recall	F1-Score	Support
0	0.95	0.97	0.96	80
1	0.69	0.85	0.76	80
2	0.86	0.64	0.73	80
Accuracy	—	—	**0.82**	**240**
Macro Avg	0.83	0.82	0.82	240
Weighted Avg	0.83	0.82	0.82	240

Set 3:

Class	Precision	Recall	F1-Score	Support
0	0.96	0.82	0.89	80
1	0.71	0.85	0.77	80
2	0.77	0.72	0.75	80
Accuracy	—	—	**0.80**	**240**
Macro Avg	0.81	0.80	0.80	240
Weighted Avg	0.81	0.80	0.80	240

Experimental Results Across Flange Sizes (6-inch → 9-inch)

Fig. A4 to A14 show the confusion matrices of 2-class, 3-class, 4-class and 5-class tasks. .



Class	Precision	Recall	F1-Score	Support
0	0.89	1.00	0.94	120
1	1.00	0.81	0.90	80
Accuracy	—	—	0.93	200
Macro Avg	0.94	0.91	0.92	200
Weighted Avg	0.93	0.93	0.92	200

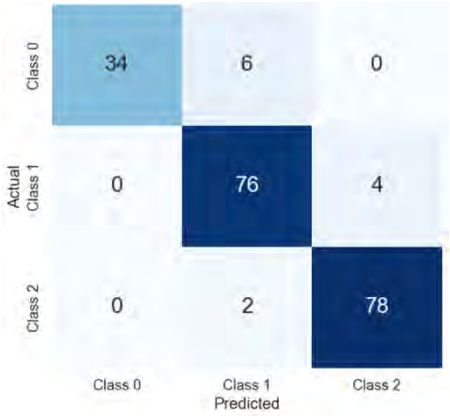
Fig. A4: Set_9-in 2

Class	Precision	Recall	F1-Score	Support
0	0.97	0.95	0.96	120
1	0.93	0.96	0.94	80
Accuracy	—	—	0.95	200
Macro Avg	0.95	0.96	0.95	200
Weighted Avg	0.96	0.95	0.96	200

Fig. A5: Set_9-in 3

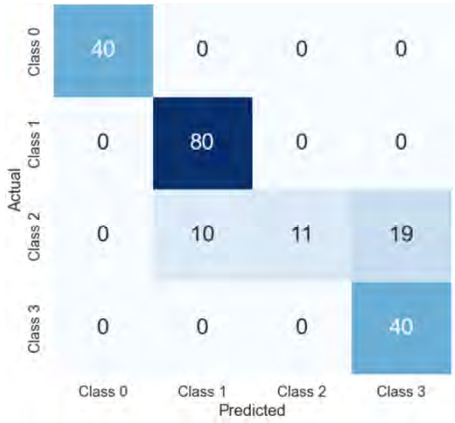
Class	Precision	Recall	F1-Score	Support
0	1.00	1.00	1.00	40
1	0.86	1.00	0.92	80
2	1.00	0.84	0.91	80
Accuracy	—	—	0.94	200
Macro Avg	0.95	0.95	0.95	200
Weighted Avg	0.94	0.94	0.93	200

Fig. A6: Set_9-in 2



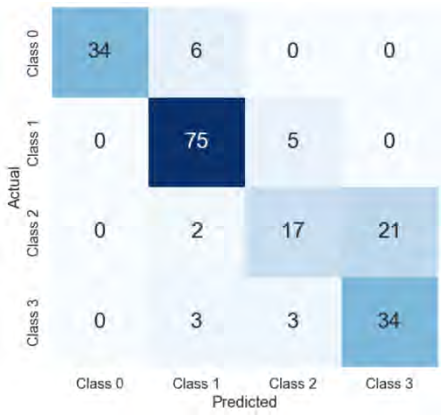
Class	Precision	Recall	F1-Score	Support
0	1.00	0.85	0.92	40
1	0.90	0.95	0.93	80
2	0.95	0.97	0.96	80
Accuracy	—	—	0.94	200
Macro Avg	0.95	0.92	0.94	200
Weighted Avg	0.94	0.94	0.94	200

Fig. A7: Set_9-in 3



Class	Precision	Recall	F1-Score	Support
0	1.00	1.00	1.00	40
1	0.89	1.00	0.94	80
2	1.00	0.28	0.43	40
3	0.68	1.00	0.81	40
Accuracy	—	—	0.85	200
Macro Avg	0.89	0.82	0.80	200
Weighted Avg	0.89	0.85	0.82	200

Fig. A8: Set_9-in 2



Class	Precision	Recall	F1-Score	Support
0	1.00	0.85	0.92	40
1	0.87	0.94	0.90	80
2	0.68	0.42	0.52	40
3	0.62	0.85	0.72	40
Accuracy	—	—	0.80	200
Macro Avg	0.79	0.77	0.77	200
Weighted Avg	0.81	0.80	0.79	200

Fig. A9: Set_9-in 3

Class	Precision	Recall	F1-Score	Support
0	1.00	1.00	1.00	40
1	0.96	0.65	0.78	40
2	0.67	1.00	0.80	40
3	1.00	0.40	0.57	40
4	0.68	0.97	0.80	40
Accuracy	—	—	0.81	200
Macro Avg	0.86	0.80	0.79	200
Weighted Avg	0.86	0.81	0.79	200

Fig. A10: Set_9-in 2

Class	Precision	Recall	F1-Score	Support
0	1.00	0.82	0.90	40
1	0.82	0.78	0.79	40
2	0.81	0.97	0.89	40
3	0.73	0.47	0.58	40
4	0.62	0.85	0.72	40
Accuracy	—	—	0.78	200
Macro Avg	0.80	0.78	0.78	200
Weighted Avg	0.80	0.78	0.78	200

Fig. A11: Set_9-in 3

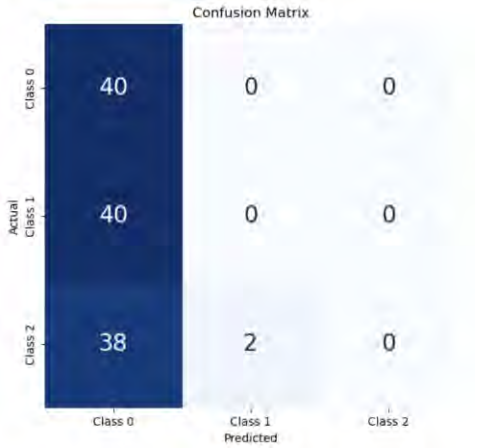


Fig. A12: Set_9-in 1

Class	Precision	Recall	F1-Score	Support
0	0.34	1.00	0.51	40
1	0.00	0.00	0.00	40
2	0.00	0.00	0.00	40
Accuracy	—	—	**0.33**	**120**
Macro Avg	0.11	0.33	0.17	120
Weighted Avg	0.11	0.33	0.17	120

Fig. A13: Set_9-in 2

Class	Precision	Recall	F1-Score	Support
0	0.33	0.90	0.49	40
1	0.00	0.00	0.00	40
2	0.00	0.00	0.00	40
Accuracy	—	—	**0.30**	**120**
Macro Avg	0.11	0.30	0.16	120
Weighted Avg	0.11	0.30	0.16	120

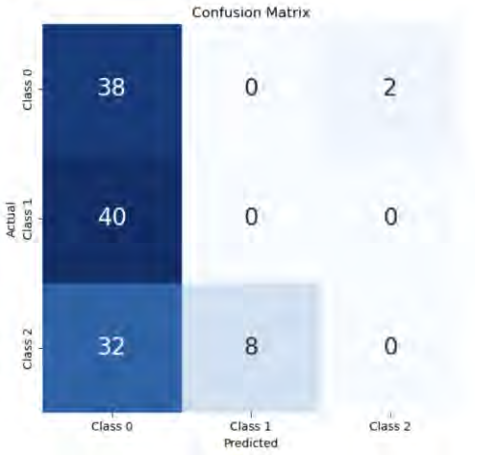


Fig. A14: Set_9-in 3

Class	Precision	Recall	F1-Score	Support
0	0.35	0.95	0.51	40
1	0.00	0.00	0.00	40
2	0.00	0.00	0.00	40
Accuracy	—	—	**0.32**	**120**
Macro Avg	0.12	0.32	0.17	120
Weighted Avg	0.12	0.32	0.17	120

Experimental Results on Domain Adaptation from Air to Water

Fig. A15 to A32 present the confusion matrices for water Set 1 through Set 6, respectively, across the 4-class, 3-class, and 2-class classification tasks.

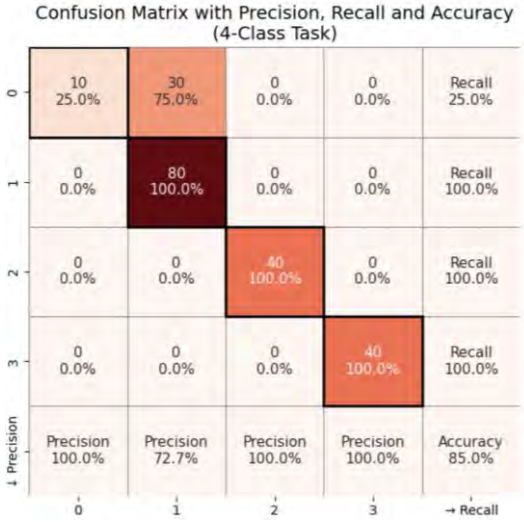


Fig. A15: Set 1

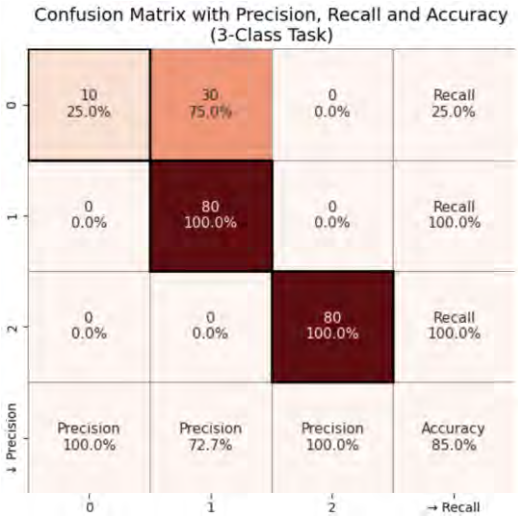


Fig. A16: Set 1

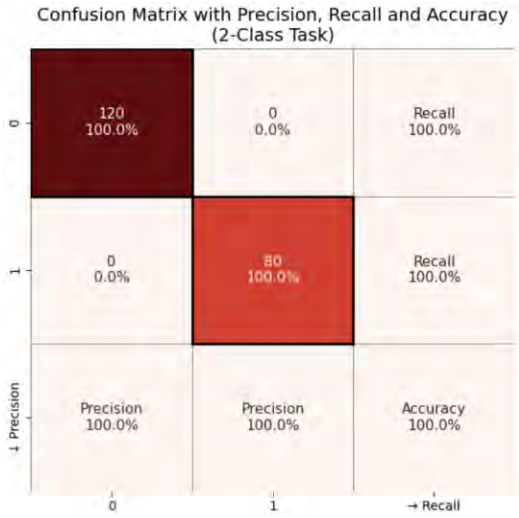


Fig. A17: Set 1

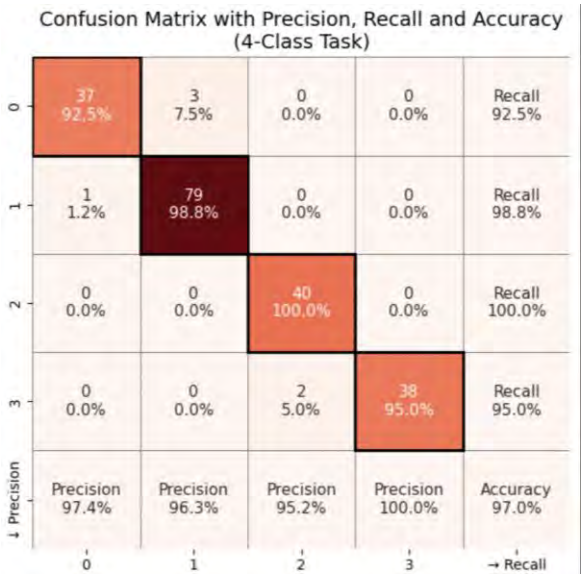


Fig. A18: Set 2

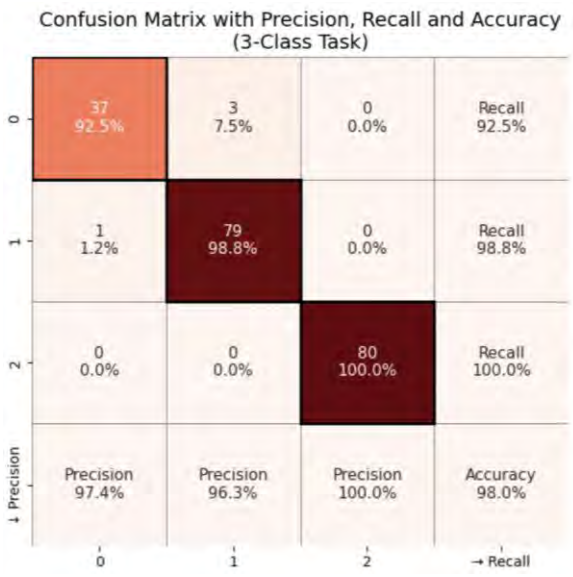


Fig. A19: Set 2

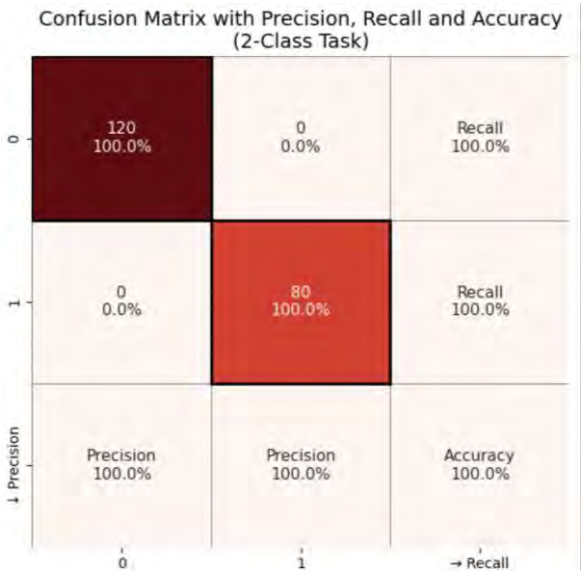


Fig. A20: Set 2

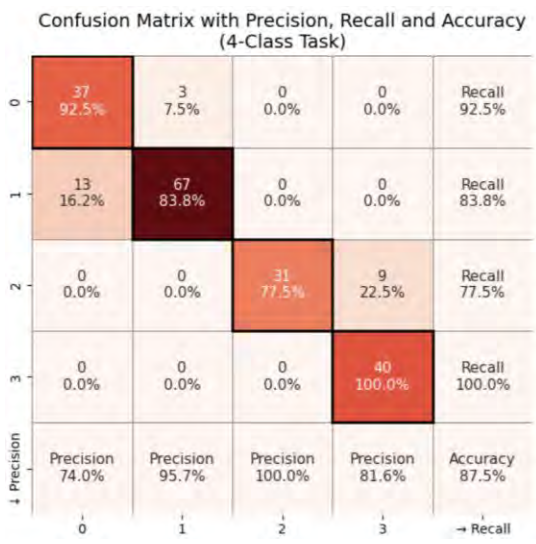


Fig. A21: Set 3

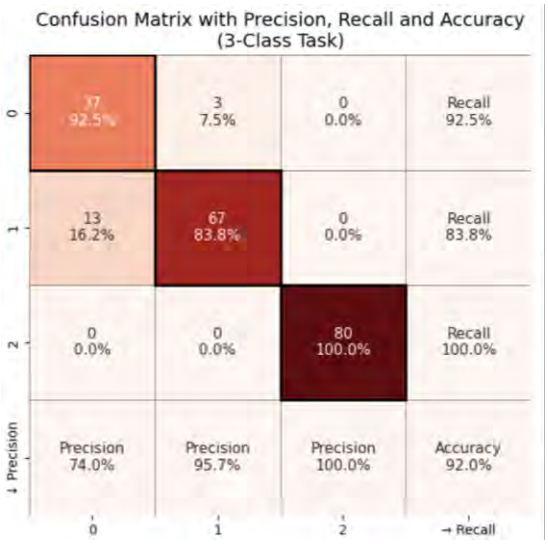


Fig. A22: Set 3

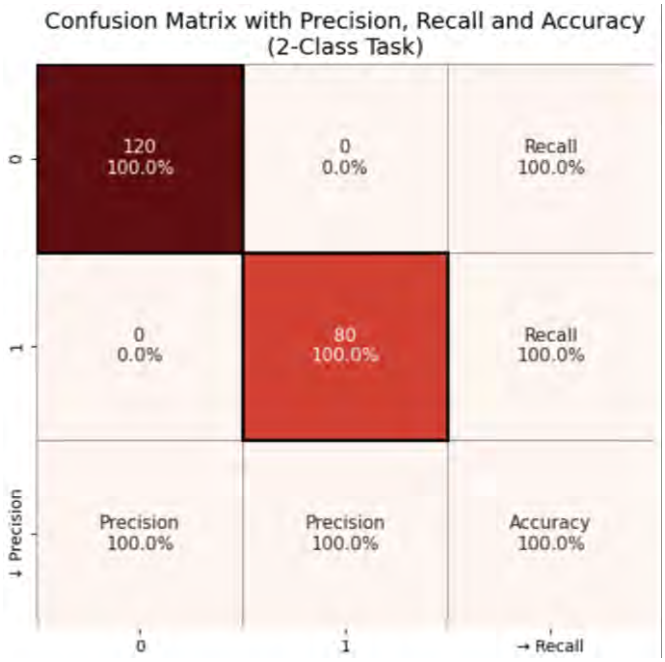


Fig. A23: Set 3

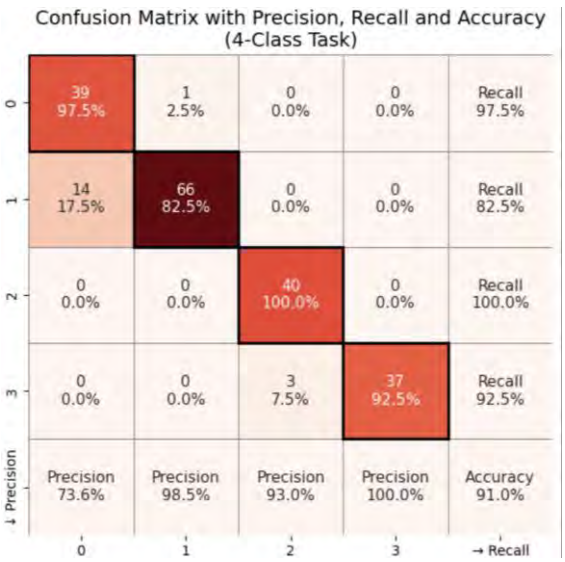


Fig. A24: Set 4

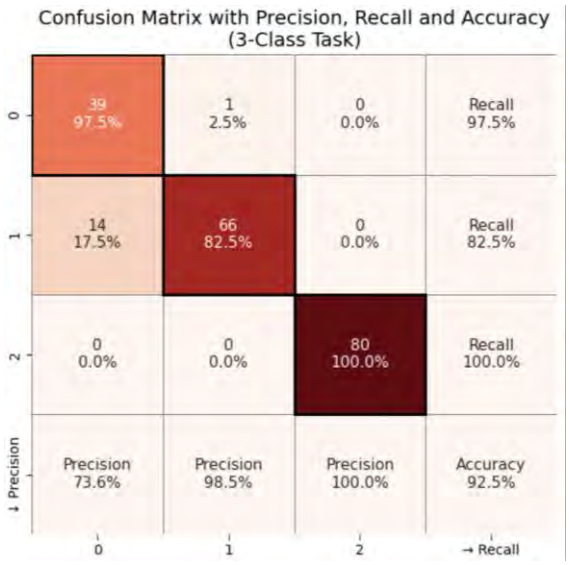


Fig. A25: Set 4

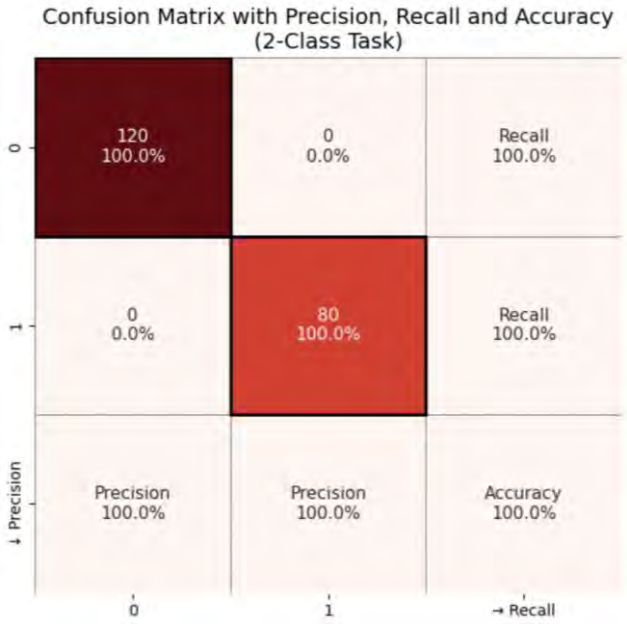


Fig. A26: Set 4

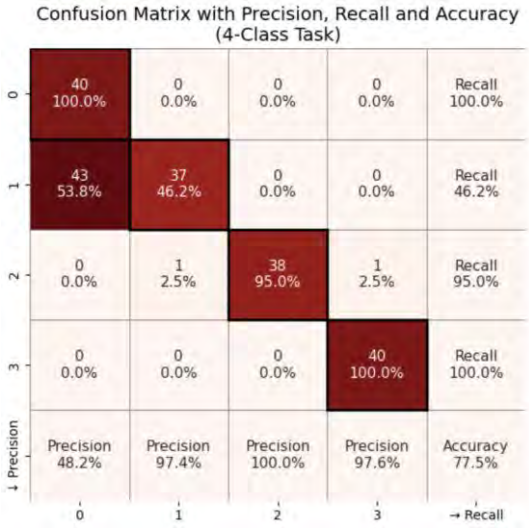


Fig. A27: Set 5

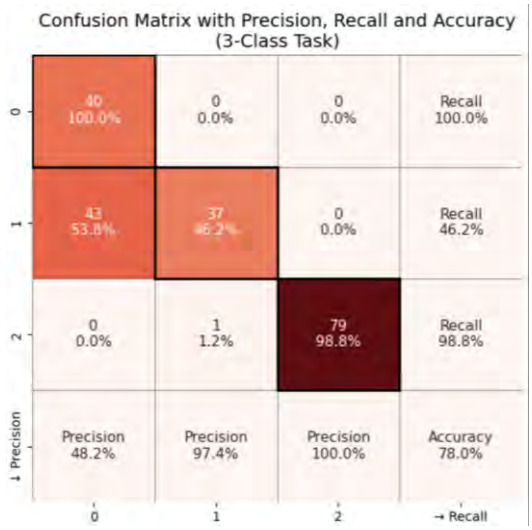


Fig. A28: Set 5

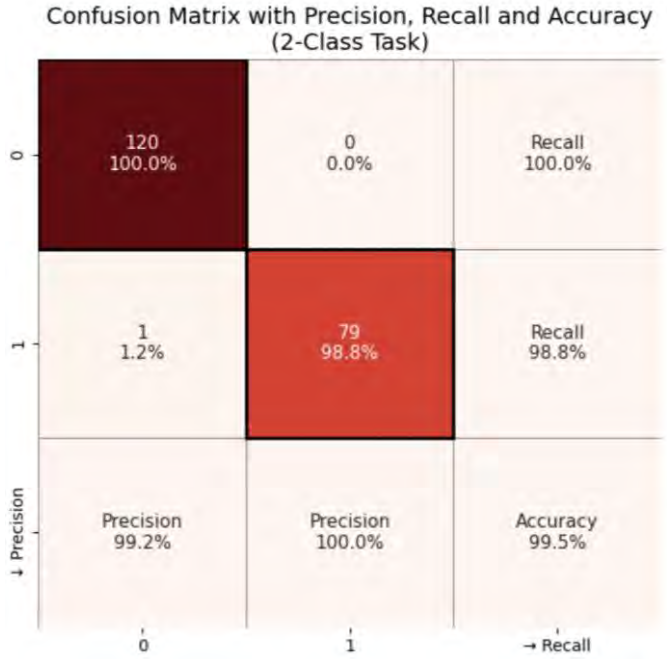


Fig. A29: Set 5

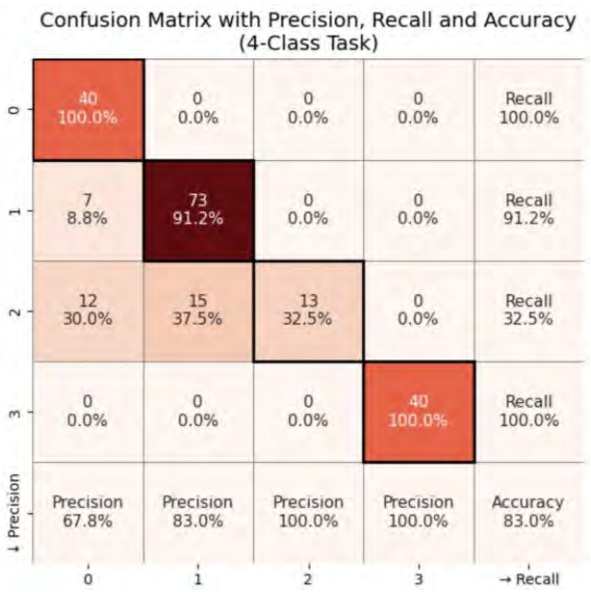


Fig. A30: Set 6

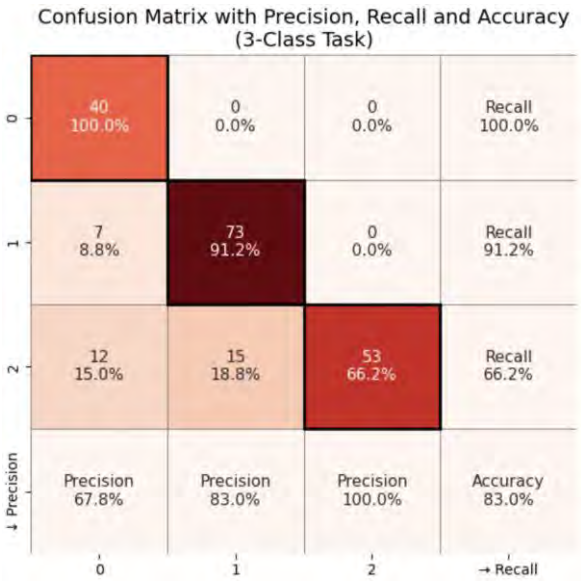


Fig. A31: Set 6

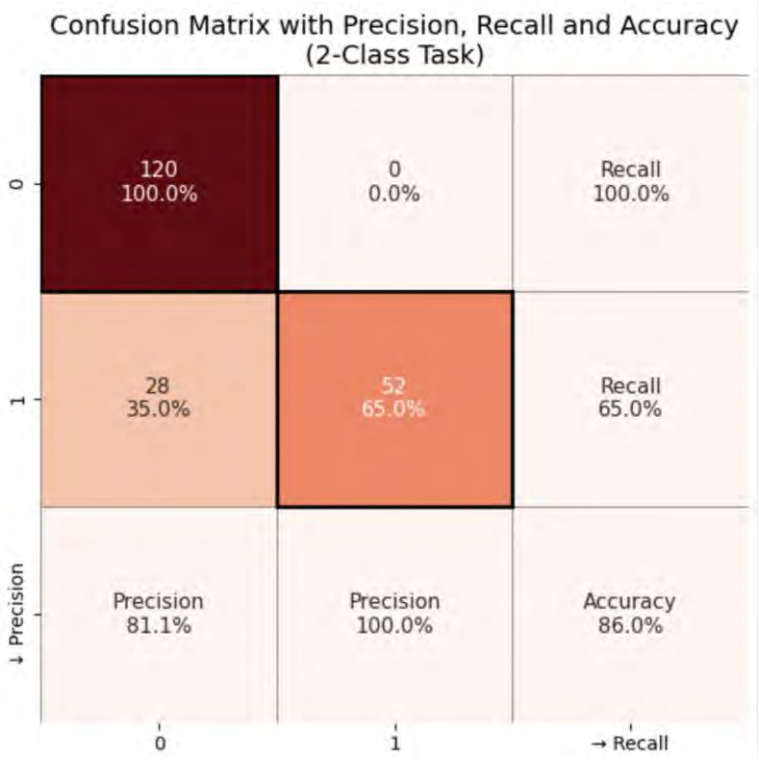


Fig. A32: Set 6