Safety Emergency Science https://mc03.manuscriptcentral.com/safems
https://doi.org/10.26599/SES.2026.9590033

Research Article

Research on key technologies of quadruped robot safety inspection on offshore oil platforms based on embodied intelligence

Xuanzheng Wang[1, 2]✉, Bin Feng[2], Zhipeng Xu[2], Yiwei Cao[2], Xiaonan Zhang[2], Wanru Chen[2], Wenzhao Lin[2], Yunlong Zhao[2], Pingang Wei[2]

[1] Institute of Modern Optics, College of Electronic Information and Optical Engineering, Nankai University, Tianjin 300350, China

[2] CNOOC Safety & Technology Services Co., Ltd., Tianjin 300450, China

✉ Corresponding author. E-mail: 651906173@qq.com.

Received: September 22, 2026; Revised: November 17, 2026; Accepted: November 24, 2026

© The Author(s) 2026.

Abstract: Aiming at the high risks of manual inspection caused by harsh environments, compact spaces, and explosive atmospheres on offshore oil platforms, as well as the insufficient adaptability of traditional wheeled/tracked inspection robots in crossing complex structures such as grating decks and vertical ladders, this paper takes embodied intelligence as the entry point to study key technologies for quadruped robot safety inspection on offshore oil platforms. An integrated "perception–locomotion–decision" inspection system architecture driven by embodied intelligence is proposed. Five key technologies are systematically discussed: multi-modal fusion perception and

offshore-platform-oriented simultaneous localization and mapping (SLAM), reinforcement learning-based all-terrain gait control, task understanding and planning driven by a vision-language-action (VLA) large model, multi-modal anomaly recognition and chain-of-thought disposal decision, and explosion-proof design with human-robot collaboration. The proposed methods are verified on a simulated offshore oil and gas test field. Results show that the proposed methods enable the quadruped robot to achieve an average passing rate exceeding 92% under various typical working conditions including grating decks, vertical ladders, and oily floors, and the comprehensive recognition accuracy for typical anomalies such as leakage, overheating, and personnel violations exceeds 90%, providing effective technical support for intelligent inspection of unmanned offshore platforms.

Keywords: Embodied Intelligence; Quadruped Robot; Offshore Oil Platform; Safety Inspection; Vision-Language-Action Model; Multi-modal Perception

1 Introduction

Offshore oil platforms are a vital component of national energy security systems and are also typical high-risk workplaces. Platforms are long exposed to high temperature, high humidity, high salt spray, strong vibration, and flammable/explosive hydrocarbon gas environments. Routine manual inspection not only involves high labor intensity but also faces multiple occupational hazards such as hydrogen sulfide poisoning, slipping and falling, and falling objects from heights. As China's oil and gas exploration continues to advance into deep and distant seas, the unmanned, reduced-crew, and intelligent operation of platforms has become a clear direction for the industry (Cheng et al., 2023; Sui et al., 2023). Against this background, applying advanced robotics to routine platform safety inspection to gradually replace or assist manual labor has become a key path to improving the inherent safety level of offshore oil and gas production.

Traditionally, auxiliary inspection on offshore platforms has mostly adopted a combination of fixed video surveillance, fixed combustible gas detectors, and a small number of wheeled inspection devices (Shukla and Karki, 2016a, 2016b). Such schemes can play a role in open deck areas, but the platform space is narrow, grating decks are numerous, and there are many ladders and non-planar passages. Wheeled or tracked robots have difficulty crossing grating openings, climbing up/down vertical ladders, and traversing pipe rack gaps, making "the most dangerous areas exactly those that robots cannot reach". Quadruped robots, with their leg-like structures, show significant advantages in traversability over rough terrain, stairs, ladders, and between beams and columns. In recent years, they have been verified in scenarios such as underground pipe galleries, substations, and mines (Tranzatto et al., 2022), and are considered one of the most promising types of legged equipment for entering complex industrial environments instead of humans.

Research on robotic inspection in the oil and gas industry started earlier abroad. Shukla and Karki (2016a, 2016b) systematically reviewed robotic applications in onshore and offshore oil and gas facilities and discussed the challenges of deploying robots in hazardous process areas. Legged robots have since moved from laboratories to industrial sites. The quadruped robot ANYmal was deployed to perform autonomous inspection tasks on an offshore high-voltage direct current (HVDC) converter platform in the North Sea (Gehring et al., 2021), and the CERBERUS team, which relied mainly on ANYmal robots, won the DARPA Subterranean Challenge in GNSS-denied underground environments (Tranzatto et al., 2022). These studies demonstrate the feasibility of quadruped robots in harsh industrial environments; however, they were not conducted on hydrocarbon production facilities with explosive atmospheres.

In parallel, the key enabling technologies have advanced rapidly. For locomotion, reinforcement learning policies trained in simulation have enabled quadruped robots to traverse rough terrain robustly (Hwangbo et al., 2019; Lee et al., 2020). By fusing exteroceptive and proprioceptive perception, Miki et al. (2022) achieved robust perceptive locomotion in the wild, and massively

parallel training has reduced policy training time to minutes (Rudin et al., 2022). Learning-based controllers have also been extended to ladder climbing, a typical obstacle on industrial facilities (Vogel et al., 2025). For localization, tightly coupled LiDAR-inertial odometry methods such as LIO-SAM and FAST-LIO2 provide accurate pose estimation in GNSS-denied environments (Shan et al., 2020; Xu et al., 2022). For task-level intelligence, large language models and vision-language-action models have been used to ground natural language instructions in executable robot skills (Ichter et al., 2023; Driess et al., 2023; Brohan et al., 2023; Kim et al., 2024). However, these technologies have mostly been studied separately and validated in general-purpose scenarios, and their integration under offshore-specific constraints, such as explosion protection, oily floors, and limited communication, has rarely been reported.

In China, research on intelligent offshore oil and gas equipment and unmanned platforms has been accelerating. Cheng et al. (2023) analyzed the development strategy of China's offshore oil and gas equipment, while Sui et al. (2023) and Wu (2024) summarized the technical status and development trends of unmanned offshore platforms and the application of intelligent technologies on offshore oil platforms, respectively. In the field of legged robots, domestic studies have made progress in motion and stability control and in deep reinforcement learning-based locomotion (Liu et al., 2022; Wang, 2025). In the coal mining industry, where explosive gas is also present, a variety of robots including inspection robots have been developed, providing references for explosion-proof design and operation in hazardous environments (Ge et al., 2023). Nevertheless, to the best of our knowledge, studies on quadruped robots that both meet the explosion-proof requirements of offshore platforms and possess embodied intelligence capabilities remain scarce.

However, directly transplanting quadruped robots to offshore platforms does not solve all problems. Offshore platforms impose much more demanding requirements on robots than land-based scenarios: firstly, the hazardous areas on platforms have strict explosion-proof requirements for electrical equipment; secondly, surfaces often contain oil, water, and salt spray, demanding high

robustness in motion control; thirdly, inspection tasks involve numerous semantic-level judgments such as valve status, meter readings, leakage identification, and personnel violations, requiring robots to have the ability to "understand the scene"; fourthly, communication is limited on platforms, requiring robots to make most decisions independently at the edge. These requirements exceed the capabilities of traditional "remote control plus pre-programmed path" robots.

The rise of embodied intelligence provides a new technical path to overcome these challenges. Embodied intelligence emphasizes that "an agent gains understanding and action capabilities through the closed loop of body–perception–action in the physical world". Its core is to deeply integrate the semantic understanding ability of large models with the robot's motion and perception capabilities, enabling the robot not only to "move stably" but also to "see clearly, think correctly, and act properly" (An et al., 2025; Feng et al., 2026). Taking embodied intelligence as the technical thread, this paper systematically studies key technologies for quadruped robot safety inspection on offshore oil platforms, focusing on multi-modal perception and localization, reinforcement learning-driven motion control, large-model-driven task understanding, anomaly recognition and disposal decision-making, intrinsically safe design, and human-robot collaboration. Preliminary verification is carried out through simulation experiments, aiming to provide a reference for the development and engineering application of intelligent inspection equipment for unmanned offshore platforms.

The remainder of this paper is organized as follows. Section 2 analyzes the inspection requirements and technical challenges of offshore oil platforms. Section 3 presents the overall architecture of the embodied-intelligence-based quadruped robot inspection system. Section 4 elaborates on the five key technologies. Section 5 reports the application case and experimental verification. Section 6 concludes the paper and outlines future work.

2 Inspection requirements and technical challenges on offshore oil platforms

2.1 Environmental characteristics of offshore oil platforms

Offshore oil platforms can be classified into fixed platforms, jack-up platforms, semi-submersible

platforms, and floating production storage and offloading (FPSO) units. Regardless of type, their working environments generally share the following common characteristics: first, highly compact spaces – in key areas such as engine rooms, production modules, wellhead areas, and gas turbine rooms, there is a dense arrangement of pipelines, valves, cables, and structural components, with passage widths typically ranging from 0.8 to 1.2 m; second, connections between floors mainly rely on inclined ladders and vertical ladders, and decks largely adopt steel grating structures to facilitate drainage and oil drainage; third, there is a universal risk of hydrocarbon gas release – the wellhead area, crude oil processing area, and flare area are all designated as explosive hazardous zones; fourth, platforms endure long-term wave loads, wind loads, and mechanical vibrations, with sustained low-frequency structural oscillations that impose additional requirements on equipment motion stability.

2.2 Types of inspection tasks and typical objects

The daily safety inspection on offshore platforms covers a wide range of content, which can be roughly summarized into four categories. The first category is equipment status inspection, including valve opening, pressure gauge and liquid level readings, equipment operating sound and vibration status – this forms the main part of inspection tasks. The second category is leakage and abnormal condition inspection, including hydrocarbon gas leakage, oil drip and spill, insulation water ingress, flange leakage, etc., often requiring combined judgment using gas detection, infrared imaging, and visual recognition. The third category is firefighting and emergency facility inspection, including whether fire extinguishers, fire hoses, foam tanks, emergency exit signs, etc., are intact and usable. The fourth category is personnel behavior inspection, focusing on whether workers are wearing safety helmets and safety belts, and whether they are violating regulations in areas such as hot work and confined spaces.

2.3 Limitations of existing inspection methods

Currently, the main inspection methods on offshore platforms can be divided into three types: manual

inspection, fixed monitoring, and mobile robot inspection (Shukla and Karki, 2016b). Their capability comparison is shown in Table 1.

Table 1 Capability comparison of existing offshore oil platform inspection methods

Inspection Method	Coverage	Task	Personnel	Main Limitations
Manual inspection	High	High	High	Exposure to hazardous environments; affected
Fixed monitoring	Medium	Low	Low	Many blind spots; difficult to detect anomalies
Wheeled/tracked robot	Medium	Medium	Low	Difficulty crossing grating decks, vertical
Quadruped robot	High	High	Low	Complex control; demanding intrinsic safety

As seen in Table 1, the quadruped robot performs best overall in terms of coverage capability and task flexibility, but its engineering deployment heavily depends on breakthroughs in key technologies such as control, perception, and intrinsic safety.

2.4 Key technical challenges

Based on the above analysis, applying quadruped robots to safety inspection on offshore platforms mainly faces the following key technical challenges: (1) Stable perception and localization in complex environments – maintaining accurate mapping under global navigation satellite system (GNSS)-denied conditions, structural vibrations, strong reflections, and oil stain occlusions; (2) All-terrain autonomous motion control – ability to cross grating openings, climb up/down vertical ladders, and handle slippery oily floors; (3) Task-level semantic understanding – the robot needs to understand natural language instructions such as "inspect leakage of the No. 3 separator" and decompose them into specific navigation, recognition, and judgment actions; (4) Multi-modal anomaly recognition and decision-making – fusing information from visible light, infrared, gas

concentration, sound, etc., to make comprehensive judgments and provide disposal recommendations; (5) Intrinsic safety and human-robot collaboration – meeting explosion-proof requirements in hazardous areas and forming a reliable collaborative closed loop with on-duty personnel.

3 Overall architecture of quadruped robot inspection system based on embodied intelligence

3.1 Technical connotation of embodied intelligence

Embodied intelligence (Embodied AI) is a concept proposed in contrast to traditional "disembodied" linguistic or visual intelligence. It emphasizes that an intelligent agent must form an understanding of the world through interaction with the physical environment and output actions (Huang et al., 2025; Duan et al., 2022). The fundamental differences from traditional robotics lie in three aspects: first, knowledge representation shifts from rule-based and symbolic to large-model-based unified representation – robots no longer rely on manually coded judgment rules; second, motion control shifts from model-based inverse dynamics solving to learning-based policy networks, enabling robots to acquire robustness against real-world disturbances through simulation; third, task execution shifts from "trajectory replay + conditional branching" to a closed loop of "goal–perception–decision–action", giving robots a certain level of task-level generalization.

3.2 Overall system architecture

The overall architecture of the proposed embodied-intelligence-based quadruped robot safety inspection system for offshore platforms is shown in Fig. 1. The system is divided into five layers: the body layer, perception layer, locomotion layer, intelligence layer, and collaboration layer. The body layer includes an explosion-proof quadruped robot body, pan-tilt camera, infrared thermal imager, light detection and ranging (LiDAR), combustible/toxic gas detection module, sound acquisition module, and communication module. The perception layer is responsible for spatiotemporal alignment, noise suppression, and feature extraction of data from various sensors, outputting a unified

multi-modal environmental representation. The locomotion layer implements gait generation and balance control based on a reinforcement-learning-trained policy network and receives velocity and posture commands from upper layers. The intelligence layer, which is the core of the system, deploys a vision-language-action (VLA) large model and domain knowledge base, responsible for task understanding, inspection planning, anomaly recognition, and disposal reasoning. The collaboration layer interfaces upward with the platform central control system and remote experts to achieve task issuance, alarm reporting, and remote auxiliary control.

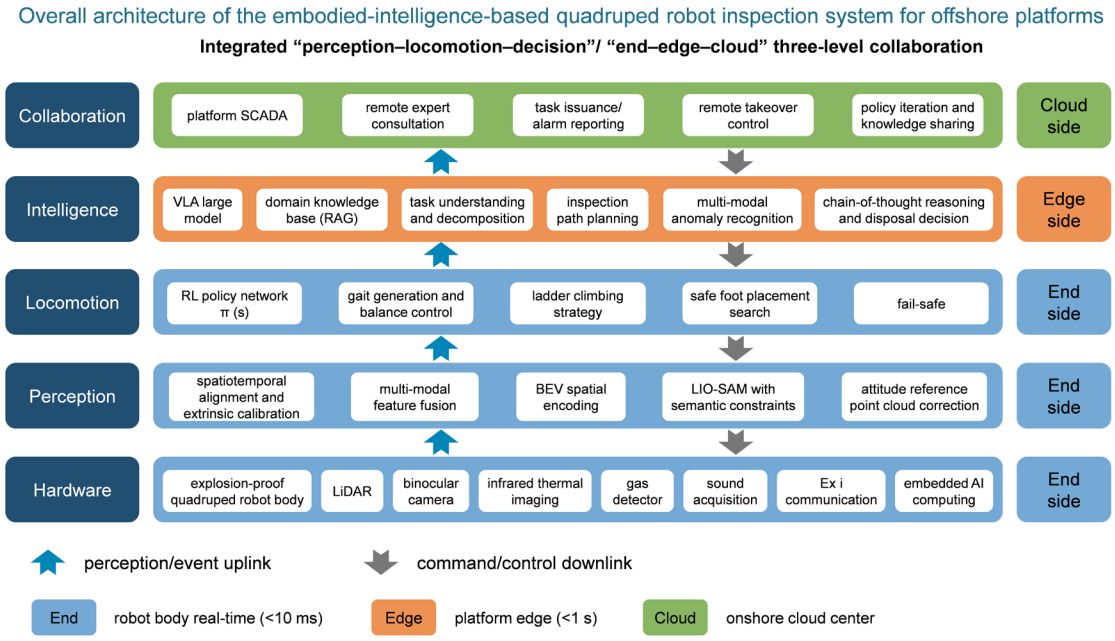


Fig. 1 Overall architecture of the embodied-intelligence-based quadruped robot inspection system for offshore platforms. SCADA, supervisory control and data acquisition.

3.3 Software and hardware collaboration framework

Given the limited communication bandwidth and high real-time requirements on offshore platforms, the system adopts an "end–edge–cloud" three-level software-hardware collaboration framework. On the end side, the robot body is equipped with an embedded AI computing unit, deploying lightweight perception models and motion control strategies to ensure millisecond-level response. On the edge

side, an edge server is deployed in the platform central control, running a medium-scale version of the VLA large model and the inspection domain knowledge base, responsible for task understanding and decision reasoning. On the cloud side, an onshore center retains a full-capability large model and historical database for policy iterative training, complex fault diagnosis, and cross-platform knowledge sharing. By deploying tasks with different latency requirements at appropriate levels, the system maintains real-time performance for low-level control while also achieving high capability for high-level semantic understanding (Macenski et al., 2022).

4 Key technologies

4.1 Autonomous perception and localization in complex environments

4.1.1 Multi-modal sensor fusion perception

Offshore platforms present strong reflections from metal structures, alternating strong light and backlight, and high amounts of oil mist and salt spray in the air, making it difficult for a single sensor to meet perception reliability requirements. This paper adopts a multi-modal fusion scheme mainly based on LiDAR, binocular vision, infrared thermal imaging, and inertial navigation, supplemented by millimeter-wave radar and combustible gas detectors. At the data level, data from various sensors are projected into a unified robot coordinate system through extrinsic calibration and time synchronization. At the feature level, inspired by BEV (Bird's-Eye-View) (Liu et al., 2023), laser point clouds and multi-view visual features are mapped into a top-view grid to facilitate subsequent spatial reasoning. At the decision level, an attention mechanism is introduced to weight the confidence of different modalities – for example, in ground areas with severe oil stain reflections, the visual weight is automatically reduced while the LiDAR and inertial weights are increased, thereby improving the robustness of the perception system under harsh working conditions.

4.1.2 Adaptation of SLAM algorithms for offshore platforms

Offshore platform SLAM faces two special problems: first, highly symmetrical platform structures

(e.g., rows of columns and pipelines) cause traditional point-feature-based SLAM to be prone to false loop closures; second, slight platform oscillations cause periodic drift of LiDAR point clouds relative to the ground. To address the first issue, this paper introduces structural semantic features (pipeline direction, stair orientation) as additional constraints on the basis of a tightly coupled scheme like LIO-SAM (Shan et al., 2020), improving the robustness of loop closure detection. For the second issue, a "pseudo-static reference coordinate system" based on platform attitude reference points is used for point cloud correction, and an industrial-grade inertial navigation system provides long-term constraints. In tests on a simulated platform, this method reduced the cumulative trajectory drift rate of SLAM from about 1.2% in the baseline scheme to below 0.6%.

4.2 Autonomous motion control on complex structures

4.2.1 Reinforcement-learning-driven gait control

Quadruped robots on offshore platforms need to handle various terrains such as grating decks, oily floors, anti-slip patterned steel plates, stairs, and vertical ladders. Traditional model-based whole-body control methods struggle to cover such a wide range of conditions. This paper adopts a deep reinforcement learning-based policy network as the low-level motion controller. Inputs include inertial measurement data, joint states, foot-end forces, and a local depth/height map; outputs are target positions for the twelve joints. During training, a large-scale simulation platform is used to randomize parameters such as ground friction, platform tilt, sensor noise, and external disturbances (domain randomization; Tobin et al., 2017). Within a curriculum learning framework, training transitions gradually from flat walking to complex terrains, enabling sim-to-real transfer (Hwangbo et al., 2019; Lee et al., 2020; Miki et al., 2022; Rudin et al., 2022).

4.2.2 Strategies for climbing ladders and crossing gratings

For the typical scenario of vertical ladders, a "hierarchical strategy" is adopted (Vogel et al., 2025): the upper-level task planning module identifies the ladder structure and issues a high-level "climb ladder" command, while the lower-level motion policy switches to a specially trained ladder-climbing

sub-policy, combined with real-time visual recognition of ladder rungs to perform grasping-type climbing. For grating decks, the focus is on optimizing foot placement and evaluating stability margins to prevent the feet from getting stuck in grating openings. Based on foot force feedback and the height map, a "safe foot placement search" algorithm is introduced to evaluate feasibility among several candidate placements for each step, selecting the foot placement with the highest stability.

4.3 Large-model-driven task understanding and planning

4.3.1 Vision-language-action model architecture

To achieve task-level semantic understanding, a vision-language-action (VLA) large model framework is introduced in the inspection intelligence layer, referencing methods such as RT-2 and OpenVLA (Brohan et al., 2023; Kim et al., 2024; Ichter et al., 2023; Driess et al., 2023), with domain adaptation for offshore platforms. The model takes as input multi-modal images (visible light, infrared) and natural language task descriptions, and outputs a structured sequence of sub-tasks and corresponding robot action commands. During domain adaptation, a multi-modal dataset containing approximately 80,000 images of typical equipment, working conditions, and violation behaviors on offshore platforms is constructed, accompanied by expert-written descriptions and disposal suggestions, serving as data sources for supervised fine-tuning (SFT) and reinforcement learning from human feedback (RLHF).

4.3.2 Natural language instruction parsing and task decomposition

Operators can issue natural language instructions, such as "Please go to the No. 3 production module, check whether the propane separator has any leakage, and focus on imaging the bottom flange." The VLA model decomposes this instruction into: (1) invoking a navigation sub-task to reach the designated area; (2) identifying the target device "propane separator" based on visual features in that area; (3) switching to an inspection sub-task, using infrared and visible-light cameras to focus on the flange position; (4) integrating gas detector readings and infrared images to determine whether leakage exists and generating a report. The whole process does not require engineers to pre-program for each

device; instead, the model's semantic understanding ability enables adaptation to new devices and new tasks, significantly reducing deployment costs for new platforms.

4.4 Multi-modal anomaly recognition and disposal decision-making

4.4.1 Multi-modal anomaly event recognition

Common anomaly events on offshore platforms include five major categories: hydrocarbon gas leakage, equipment overheating, flame and heavy smoke, oil leakage, and personnel violations. Each type of event has different most informative sources: leakage is most sensitive to combustible gas detection and infrared optical gas imaging (OGI) (Wang et al., 2020), overheating is most sensitive to infrared thermal imaging, and violations mainly rely on visible light vision (Fang et al., 2018). This paper constructs a multi-branch multi-modal fusion recognition model, where each branch corresponds to one perception modality. Through feature-level attention fusion, the final event classification and localization results are output, achieving a mean average precision (mAP) of 0.89 on a self-built test set.

4.4.2 Chain-of-thought reasoning and disposal decision-making

Simply saying "suspected leakage detected" is insufficient; operators want to know "what substance, how severe it might be, and what action should be taken." This paper provides the inspection knowledge base (including health, safety and environment (HSE) procedures, emergency plans, and typical cases) to the large model via retrieval-augmented generation (RAG) (Lewis et al., 2020), enabling chain-of-thought reasoning (Wei et al., 2022) around anomaly events and outputting structured conclusions including preliminary causes, risk levels, recommended measures, and reporting targets. For example, when the gas concentration near a propane separator exceeds 20% of the lower explosive limit (LEL), the system not only triggers an alarm but also suggests shutting down the upstream feed, coordinating remote cut-off valves, and notifying the shift supervisor, embodying a closed loop of "recognition–judgment–disposal".

4.5 Explosion-proof design and human-robot collaboration safety

4.5.1 Flameproof enclosure design

Most production areas on offshore platforms are classified as Zone 1 or Zone 2 explosive hazardous locations. The robot body and its electrical components must be designed in accordance with the GB/T 3836 series standards. Common solutions include flameproof enclosures (Ex d) combined with intrinsically safe (Ex i) communication links. The core of the flameproof design lies in the control of flameproof joint parameters, mainly consisting of three elements: joint width, gap, and surface roughness. According to GB/T 3836.2-2021 (SAMR and SAC, 2021), the spigot flameproof joint for Group IIC equipment consists of a plane part and a cylindrical part, with a total length $L = c + d$, where the cylindrical part length must exceed half of the total length, the plane part length must be no less than 6 mm, and the gap difference in the cylindrical section shall not exceed 0.1 mm.

For different parts of the robot, differentiated flameproof joint structures are designed in this project (Fig. 2). The top covers of the torso electrical cavity and battery cavity adopt spigot flameproof joints, with the plane part sealed by screw compression. The output rotor of the hip joint adopts a cylindrical flameproof joint, with a radial gap of less than 0.1 mm and a length greater than 18 mm, coupled with a flameproof copper ring to achieve rotary dynamic sealing. The connecting interfaces of each module housing adopt spigot flameproof joints, with a radial gap of less than 0.05 mm and a length greater than 12.5 mm. In terms of structural integration, the thigh and shin joint modules are placed within the same flameproof cavity, and the front and rear modules are connected through circumferentially distributed bolt holes, enabling modular assembly/disassembly and axial space compression. The torso adopts a top-and-bottom split cavity sealing architecture with separation of power and control circuits, achieving ingress protection ratings from IP 65 to IP 67. The core load-bearing components are forged from 7075 aviation aluminum alloy, ensuring the strength of the flameproof enclosure while controlling the total weight of the robot.

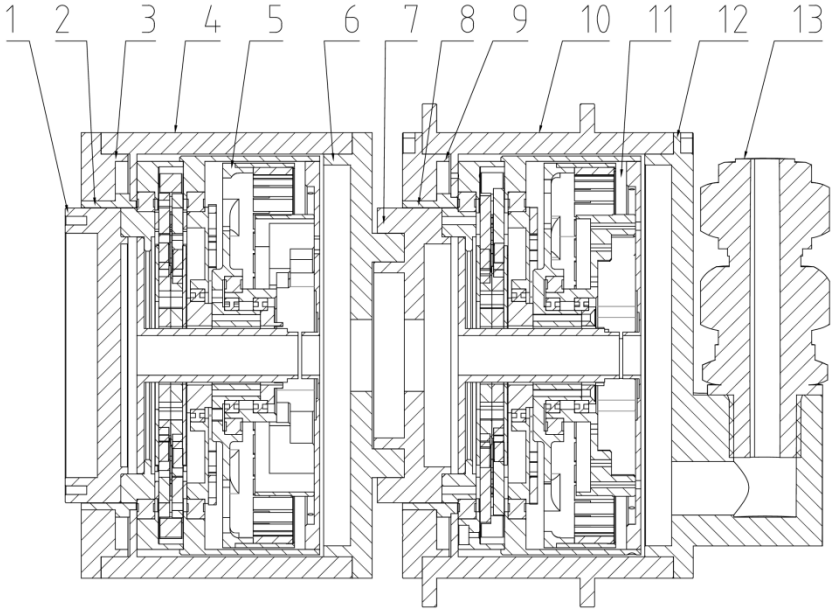


Fig. 2 Cross-sectional view of the double-module flameproof structure

1—front module output rotor; 2—front module flameproof copper ring; 3—front module flameproof upper housing; 4—front module flameproof outer housing; 5—front joint module; 6—front module flameproof rear housing; 7—rear module output rotor; 8—rear module flameproof upper housing; 9—rear module flameproof copper ring; 10—rear module flameproof outer housing; 11—rear joint module; 12—rear module flameproof rear housing; 13—armoured explosion-proof cable gland.

The overall robot structure is shown in Fig. 3. The flameproof design imposes additional constraints on the robot's motion performance. The flameproof enclosure limits heat dissipation conditions, requiring computational units to perform model quantization and pruning at the edge side to reduce power consumption, while the electrical cavity and battery cavity are arranged separately to control temperature rise. In terms of safety protection, the system incorporates multiple layers of fail-safe mechanisms: when perception module anomalies are detected, the system automatically switches to a safe gait and decelerates to a stop; when communication is interrupted, it automatically executes a return-to-base or standby-in-place strategy; the drives are equipped with multiple electrical protections including overvoltage, overcurrent, and overtemperature; and monitoring points are installed at critical flameproof joints to ensure enclosure integrity.

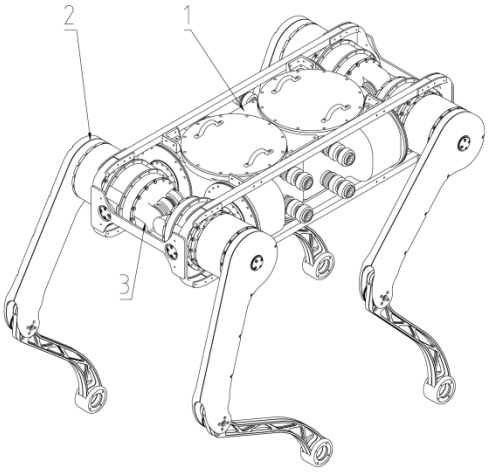


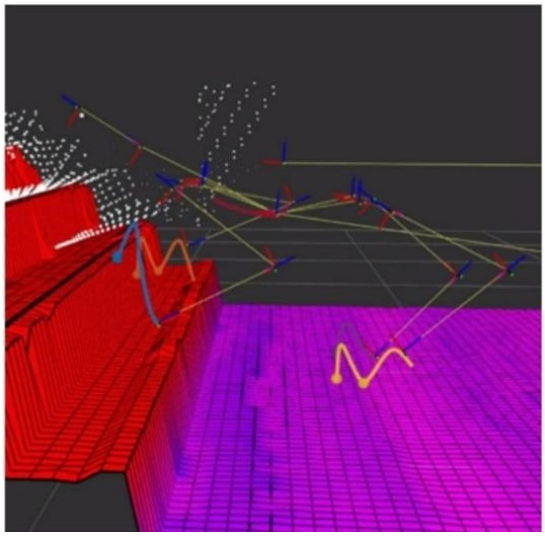
Fig. 3 Overall schematic of the explosion-proof quadruped robot

1—explosion-proof torso; 2—leg; 3—hip.

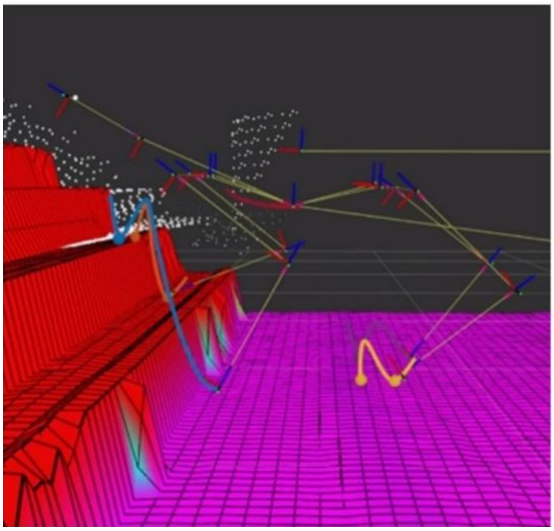
4.5.2 Simulation and physical experimental verification

In terms of human-robot collaboration, the system designs three working modes: autonomous inspection mode, where the robot works independently according to a predetermined schedule; collaborative assistance mode, where on-duty personnel remotely view the robot's perspective via a central control screen and provide instructions; and remote takeover mode, where onshore experts directly take over control in case of anomalies. The three modes switch through a clear authority management and status feedback mechanism to avoid control conflicts, complying with the traceability requirements of offshore platform operations.

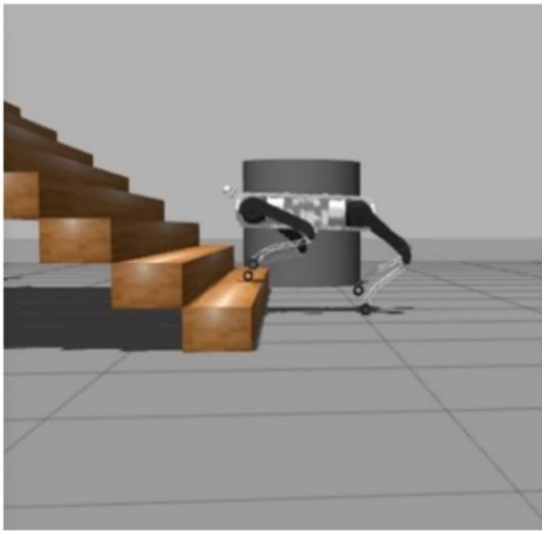
This project builds a simulation and physical verification platform based on the Robot Operating System (ROS) architecture. For simulation modeling, xacro files are used to construct the robot Unified Robot Description Format (URDF) model, which divides the robot into five parts: the torso and four legs. Each leg is further subdivided into the hip module, thigh module, shin module, and foot module. The Pinocchio library is employed to automatically solve dynamic parameters such as inertia matrices based on the URDF file. In the Gazebo simulation environment, real-time terrain elevation mapping provides prior environmental information for gait planning on unstructured terrain.



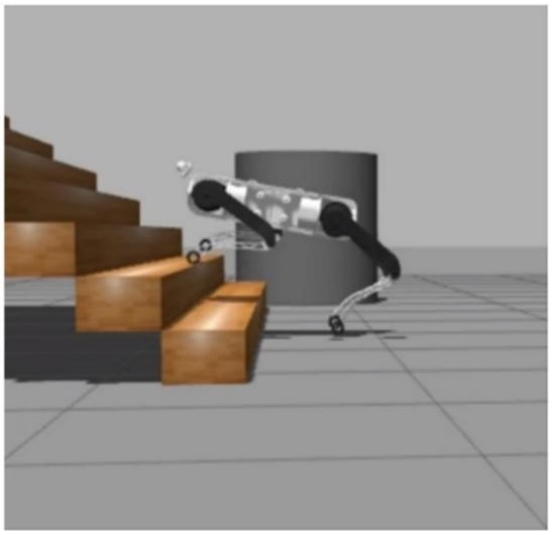
(a) Elevation map and planned foot swing trajectories for stepping onto the first step



(b) Elevation map and planned foot swing trajectories for stepping onto the second step



(c) Front foot placed on the first step in Gazebo



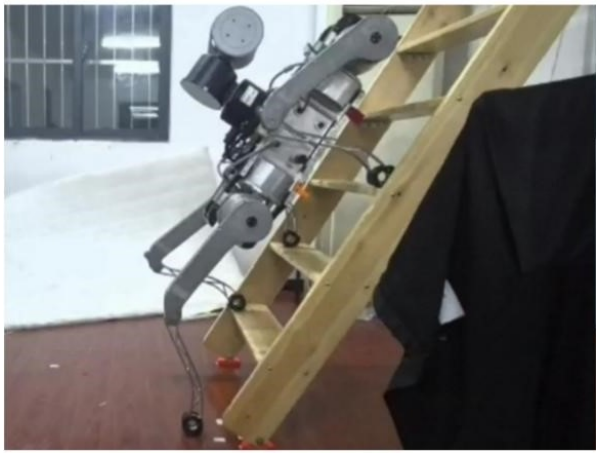
(d) Front foot placed on the second step in Gazebo

Fig. 4 Simulation elevation maps of the explosion-proof quadruped robot

In Fig. 4, (a) and (b) show the state of stepping up onto a stair, where the curves indicate the desired swing trajectories of the foot ends. (c) and (d) are simulation images of subsequent locomotion.

Simulation experiments verify the robot's motion performance under two typical conditions: flat

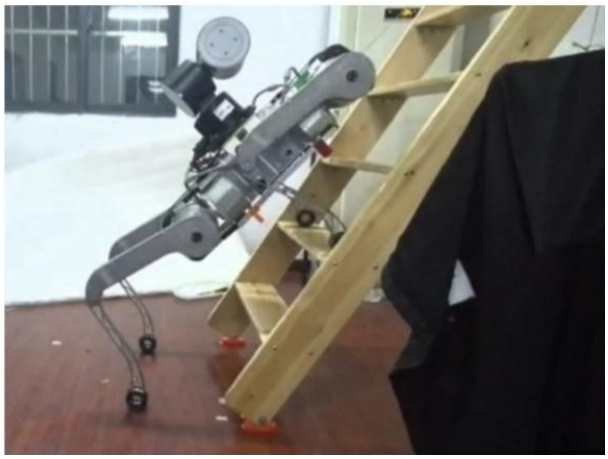
ground and stairs. In the flat-ground walking test, the robot completes a full gait cycle, with smooth attitude quaternion curves and a stable center-of-mass height. In the stair-climbing test, the foothold planning algorithm based on the elevation map modifies the foot-end trajectory in real time according to terrain information, enabling the robot to step onto the stairs and continue moving forward. The physical prototype is shown in Fig. 5. For physical validation, through adaptation and rewriting of low-level control interfaces, the control effectiveness under both translation and rotation basic motion modes is verified. The combined simulation and physical experimental results validate the effectiveness and engineering feasibility of the designed control system.



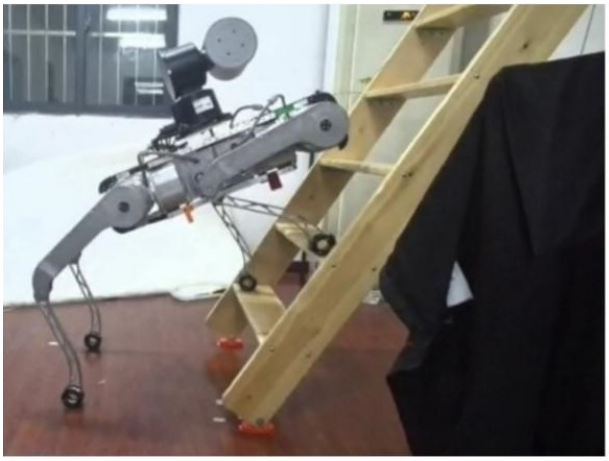
(a) Posture 1 on the ladder



(b) Posture 2 on the ladder



(c) Posture 3 on the ladder



(d) Posture 4 on the ladder

Fig. 5 Physical prototype of the explosion-proof quadruped robot

5 Application case and experimental verification

5.1 Test platform and typical scenarios

To verify the effectiveness of the proposed key technologies, a 1:1 test scenario with the same main structure as an actual platform was built using a marine oil engineering simulation test field. The scenario includes four typical functional areas: wellhead area, production module, engine room, and flare area, equipped with process pipelines of various diameters, vertical ladders, grating decks, and stairs. Oily floors, water accumulation areas, and salt spray environments were deliberately created to simulate extreme conditions. The quadruped robot used in the tests has a body weight of approximately 45 kg, a single-charge endurance of 3 hours, and is equipped with an embedded AI computing unit, LiDAR, binocular camera, combustible gas detection module, and infrared thermal imager, overall meeting the Ex d IIB T4 Gb explosion-proof requirement.

5.2 Key performance indicator testing

Test schedule. The tests were conducted in the offshore oil and gas simulation test field described in Section 5.1 over approximately two weeks of consecutive working days, including equipment commissioning and formal testing, during the day-shift period. Each paired comparison round, comprising one trial of the proposed method and one corresponding trial of the baseline method under the same test conditions, lasted approximately 8–15 min in total. Multiple comparison rounds were conducted within each test batch, with an interval of approximately 30–45 min between successive batches for charging, data download, scene reset, and safety inspection.

Test procedure. Each test day started with a pre-test check of the robot body, sensors, battery level, and joint temperatures, followed by scene reset (reapplying the oil film and checking the gratings and ladder) and communication and time synchronization with the central control system. In formal testing, the robot traversed a continuous grating area along a predefined path with turns and local oily patches

(50 trials), climbed up and down a vertical ladder using the ladder-climbing sub-policy (30 trials), and walked along straight and curved paths on a continuous oil film (50 trials); the number of successful traversals, completion time, foot entrapments, and slips were recorded. The instruction understanding test (100 randomly issued natural language instructions) and the anomaly recognition test (500 real-shot and simulated samples) evaluated the correctness of task decomposition and the multi-modal recognition accuracy, respectively. After each test, the robot returned to the charging station, the logs and sensor data were uploaded, and the joints and feet were inspected before the next round.

Contingency plan for emergencies. A three-level safety response mechanism consistent with Section 4.5 was applied during the tests. For perception or communication anomalies, the robot switched to a safe gait, stopped, and reported to the central control system. For motion or stability anomalies, such as foot entrapment, severe slipping, or an excessive attitude angle, the current task was aborted and the robot attempted self-recovery; if recovery failed, it returned to base and raised an alarm. For a combustible gas concentration exceeding 20% LEL or signs of fire, the robot stopped all motion, cut off non-essential power, sent a highest-priority alarm, and activated the remote takeover mode. In addition, an on-site safety officer supervised the tests and could cut off the robot's power at any time with an emergency stop button.

The tests mainly examine three aspects of performance: (1) Complex structure pass rate – statistics of the robot's success rate in completing full traversal under typical working conditions such as grating decks, vertical ladders, and oily floors; (2) Task-level instruction understanding accuracy – correctness rate of task decomposition for 100 natural language inspection instructions; (3) Anomaly recognition accuracy – the proportion of correctly recognized samples among 500 anomaly event samples, including both real-shot and simulated cases. The results are summarized in Table 2. Table 2 lists both the rates and the underlying counts (number of successful trials or correct results / sample size), so that each percentage can be traced back to the raw test records.

Table 2 Key performance indicator test results (values in parentheses: successes / sample size)

Indicator	Sample Size	Proposed Method	Baseline Method
Grating deck pass rate	50 trials	96.0% (48/50)	78.0% (39/50)
Vertical ladder pass rate	30 trials	90.0% (27/30)	53.3% (16/30)
Oily floor pass rate	50 trials	92.0% (46/50)	80.0% (40/50)
Task instruction understanding accuracy	100 instructions	94.0% (94/100)	76.0% (76/100)
Overall anomaly recognition accuracy	500 samples	90.2% (451/500)	81.6% (408/500)

As shown in Table 2, after introducing the embodied-intelligence-related key technologies, the robot's pass rates and recognition accuracy under various typical offshore platform conditions are significantly improved compared to the baseline method. Particularly for vertical ladder passing and task instruction understanding, the improvements exceed 15 percentage points, verifying the advantages of the proposed methods on complex structures and semantic tasks.

In addition to the statistical indicators in Table 2, the raw motion data obtained in the simulation and physical prototype verification described in Section 4.5.2 are provided in Fig. 6 and Table 3 as directly readable intermediate data for the basic locomotion capability underlying these indicators. The simulation data were recorded in Gazebo through ROS at 500 Hz (IMU), 100 Hz (joint states), and approximately 167 Hz (body pose), covering standing up (0–4.3 s), in-place turning (4.3–8.2 s), and straight trotting (10.5–18.3 s). The physical data were collected from the hip, thigh, and shank joints of the right front leg of the prototype during forward, backward, and lateral translation, including the commanded position, measured position, and motor current of each joint (3291 samples). During in-place turning, the robot rotated by approximately 86° at an average yaw rate of 22.3°/s, with a body position drift of 0.09 m. During straight trotting, it travelled 2.17 m in 7.8 s (average speed 0.28 m/s) at a gait frequency of approximately 1.67 Hz; the standard deviation of the body height was 1.5 mm, the roll and pitch angles remained within ±1.3°, and the standard deviation of the heading was 0.51°. On the physical prototype, the root-mean-square tracking errors of the hip, thigh, and shank joints were

0.007, 0.021, and 0.033 rad, respectively, and the maximum tracking error did not exceed 0.081 rad, indicating that the rewritten low-level control interface closely tracks the commanded joint trajectories.

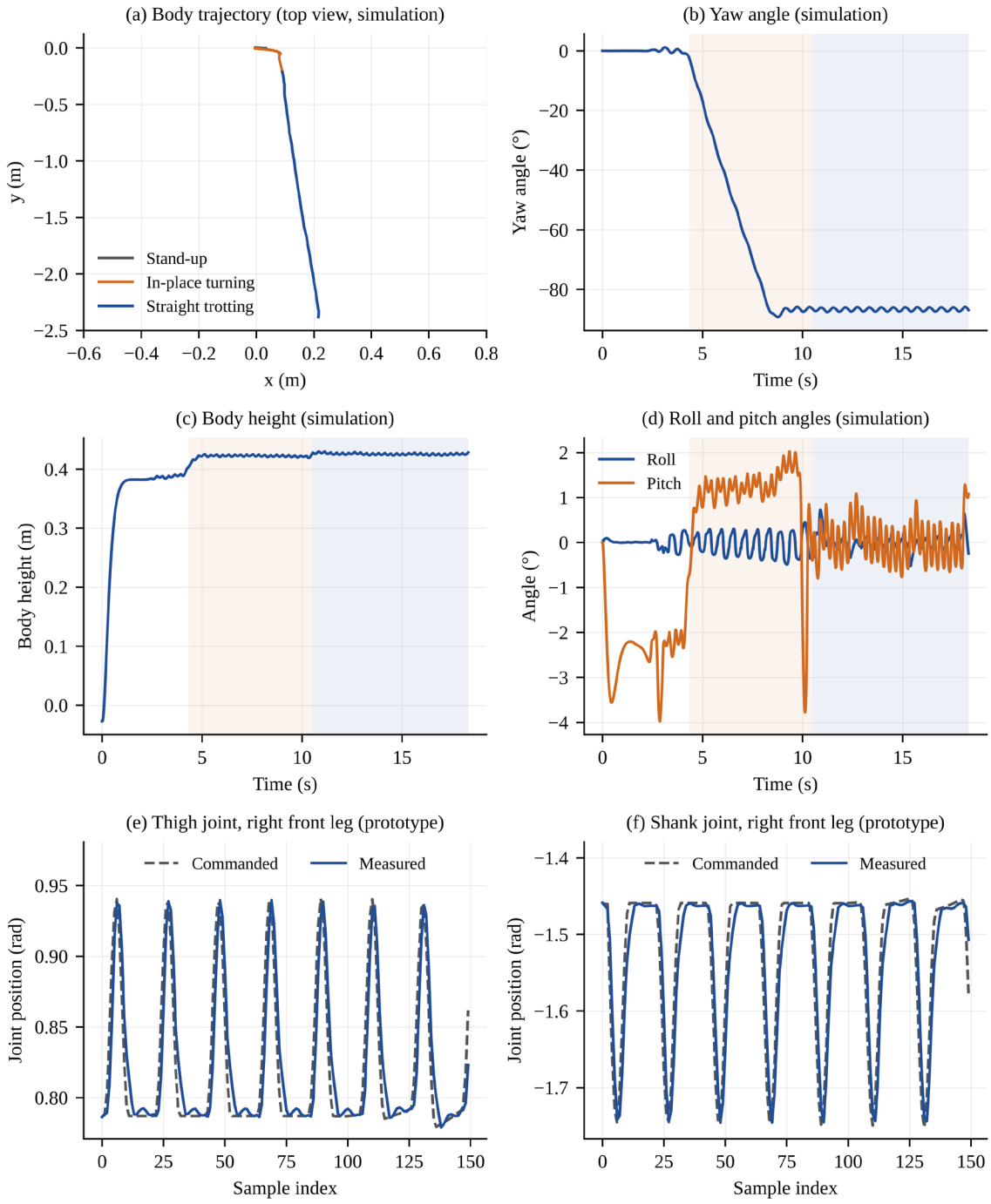


Fig. 6 Motion test data of the simulation model and the physical prototype

Table 3 Basic locomotion test data

Test condition	Indicator	Measured value	Unit
Simulation: in-place turning	Turning angle	86	°
	Average yaw rate	22.3	°/s
	Body position drift	0.09	m
Simulation: straight trotting	Distance / duration	2.17 / 7.8	m / s
	Average speed	0.28	m/s
	Gait frequency	1.67	Hz
	Body height (mean ± SD)	0.426 ± 0.0015	m
	Maximum |roll| / |pitch|	0.73 / 1.29	°
	Heading standard deviation	0.51	°
Prototype: right front leg	Tracking RMSE (hip / thigh / shank)	0.007 / 0.021 / 0.033	rad
	Maximum tracking error (hip / thigh / shank)	0.036 / 0.081 / 0.081	rad
	Peak motor current (hip / thigh / shank)	22.0 / 30.4 / 29.6	A

5.3 Discussion of results

The test results show that the proposed key technologies are engineering-feasible in typical offshore platform scenarios, but several issues remain for further study. First, under heavily oily conditions, the foot friction characteristics differ significantly from the training domain, leading to occasional slipping events – further optimization of foot-end materials and slip detection strategies is needed. Second, the VLA model has limited generalization ability for extreme conditions (e.g., small target recognition under severe backlight) – expanding extreme samples and introducing active illumination assistance are required. Third, under long-term continuous operation (>4 h), aspects such as heat dissipation, energy consumption, and model inference stability need further verification. Fourth, deep integration

with platform supervisory control and data acquisition (SCADA) and safety instrumented systems is still in a semi-offline state – closed-loop integration testing should be carried out in the next engineering demonstration phase.

6 Conclusions and outlook

(1) Focusing on the practical needs of safety inspection on offshore oil platforms, and addressing the limitations of traditional inspection methods and wheeled/tracked robots in complex structures, intrinsic safety, and semantic tasks, this paper proposes an overall architecture for a quadruped robot safety inspection system based on embodied intelligence.

(2) For the five key technologies – multi-modal fusion perception and offshore-platform-oriented SLAM, reinforcement-learning-driven all-terrain motion control, VLA-large-model-driven task understanding and planning, multi-modal anomaly recognition and chain-of-thought disposal decision, and explosion-proof design with human-robot collaboration – corresponding technical solutions adapted to offshore platforms are provided.

(3) Preliminary verification on a simulated test field shows that the proposed methods enable the quadruped robot to achieve an average pass rate exceeding 92% under working conditions such as grating decks, vertical ladders, and oily floors, and a comprehensive recognition accuracy exceeding 90% for anomalies such as leakage, overheating, and personnel violations, effectively replacing part of manual inspection tasks.

(4) Future work will focus on robustness under extreme conditions, long-term reliability, deep integration with platform SCADA and safety instrumented systems, and carrying out engineering demonstrations on more actual offshore platforms to promote the large-scale application of embodied intelligence inspection robots in the offshore oil safety product system.

Data availability statements

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was financially supported by Tianjin Science and Technology Bureau, Tianjin, China (No. 22YDLQCG00110).

References

An, Z., Yu, X., Wang, C., et al. 2025. Embodied intelligence: Recent advances and future perspectives. *The Innovation Informatics*, 1: 100008.

Brohan, A., Brown, N., Carbajal, J., et al. 2023. RT-2: Vision-language-action models transfer web knowledge to robotic control. In: Proceedings of the 7th Conference on Robot Learning. Proceedings of Machine Learning Research, 229: 2165–2183.

Cheng, B., Fu, Q., Li, Q., et al. 2023. Development strategy of China's offshore oil and gas equipment. *Strategic Study of CAE*, 25: 13–21. (in Chinese)

Driess, D., Xia, F., Sajjadi, M. S. M., et al. 2023. PaLM-E: An embodied multimodal language model. In: Proceedings of the 40th International Conference on Machine Learning. Proceedings of Machine Learning Research, 202: 8469–8488.

Duan, J., Yu, S., Tan, H. L., et al. 2022. A survey of embodied AI: From simulators to research tasks. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 6: 230–244.

Fang, Q., Li, H., Luo, X., et al. 2018. Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. *Automation in Construction*, 85: 1–9.

Feng, Z., Xue, R., Yuan, L., et al. 2026. Multi-agent embodied AI: advances and future directions. *Science China Information Sciences*, 69: 151202.

Ge, S., Hu, E., Li, Y. 2023. New progress and direction of robot technology in coal mine. *Journal of China Coal Society*, 48: 54–73. (in Chinese)

Gehring, C., Fankhauser, P., Isler, L., et al. 2021. ANYmal in the field: Solving industrial inspection of an offshore HVDC platform with a quadrupedal robot. In: Field and Service Robotics. Springer Proceedings in Advanced Robotics, 16: 247–260.

Huang, J., Xu, Y., Wang, Q., et al. 2025. Foundation models and intelligent decision-making: Progress, challenges, and perspectives. *The Innovation*, 6: 100948.

Hwangbo, J., Lee, J., Dosovitskiy, A., et al. 2019. Learning agile and dynamic motor skills for legged robots. *Science Robotics*, 4: eaau5872.

Ichter, B., Brohan, A., Chebotar, Y., et al. 2023. Do as I can, not as I say: Grounding language in robotic affordances. In: Proceedings of the 6th Conference on Robot Learning. Proceedings of Machine Learning Research, 205: 287–318.

Kim, M. J., Pertsch, K., Karamcheti, S., et al. 2024. OpenVLA: An open-source vision-language-action model. In: Proceedings of the 8th Conference on Robot Learning. Proceedings of Machine Learning Research, 270: 2679–2713.

Lee, J., Hwangbo, J., Wellhausen, L., et al. 2020. Learning quadrupedal locomotion over challenging terrain. *Science Robotics*, 5: eabc5986.

Lewis, P., Perez, E., Piktus, A., et al. 2020. Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33: 9459–9474.

Liu, W., Li, B., Hou, L., et al. 2022. Review of quadruped robot research based on deep reinforcement learning. *Journal of Qilu University of Technology*, 36: 67–74. (in Chinese)

Liu, Z., Tang, H., Amini, A., et al. 2023. BEVFusion: Multi-task multi-sensor fusion with unified bird's-eye view representation. In: 2023 IEEE International Conference on Robotics and Automation (ICRA), London, UK, 2774–2781.

Macenski, S., Foote, T., Gerkey, B., et al. 2022. Robot Operating System 2: Design, architecture, and uses in the wild. *Science Robotics*, 7: eabm6074.

Miki, T., Lee, J., Hwangbo, J., et al. 2022. Learning robust perceptive locomotion for quadrupedal robots in the wild. *Science Robotics*, 7: eabk2822.

Rudin, N., Hoeller, D., Reist, P., et al. 2022. Learning to walk in minutes using massively parallel deep reinforcement learning. In: Proceedings of the 5th Conference on Robot Learning. Proceedings of Machine Learning Research, 164: 91–100.

Shan, T., Englot, B., Meyers, D., et al. 2020. LIO-SAM: Tightly-coupled lidar inertial odometry via smoothing and mapping. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 5135–5142.

Shukla, A., Karki, H. 2016a. Application of robotics in onshore oil and gas industry—A review Part I. *Robotics and Autonomous Systems*, 75: 490–507.

Shukla, A., Karki, H. 2016b. Application of robotics in offshore oil and gas industry—A review Part II. *Robotics and Autonomous Systems*, 75: 508–524.

State Administration for Market Regulation, Standardization Administration of China (SAMR, SAC). 2021. Explosive atmospheres—Part 2: Equipment protection by flameproof enclosures “d” (GB/T 3836.2—2021). Beijing: Standards Press of China. (in Chinese)

Sui, X., Qiu, H., Zhang, F., et al. 2023. Current situation and development trend of unmanned platform technology in offshore oilfield. *Ocean Engineering Equipment and Technology*, 10: 111–116. (in Chinese)

Tobin, J., Fong, R., Ray, A., et al. 2017. Domain randomization for transferring deep neural networks from simulation to the real world. In: 2017 IEEE/RSJ International Conference on Intelligent

Robots and Systems (IROS), 23–30.

Tranzatto, M., Miki, T., Dharmadhikari, M., et al. 2022. CERBERUS in the DARPA Subterranean Challenge. *Science Robotics*, 7: eabp9742.

Vogel, D., Baines, R., Church, J., et al. 2025. Robust ladder climbing with a quadrupedal robot. In: 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).

Wang, J., Tchapmi, L. P., Ravikumar, A. P., et al. 2020. Machine vision for natural gas methane emissions detection using an infrared camera. *Applied Energy*, 257: 113998.

Wang, X. 2025. An overview of key technologies for quadruped robot motion and stability control. *Journal of Jilin University (Engineering and Technology Edition)*, 55: 1483–1496. (in Chinese)

Wei, J., Wang, X., Schuurmans, D., et al. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35: 24824–24837.

Wu, B. 2024. Application research of intelligent technology in improving production stability and efficiency on offshore oil platforms. *Ocean Engineering Equipment and Technology*, 11: 29–33. (in Chinese)

Xu, W., Cai, Y., He, D., et al. 2022. FAST-LIO2: Fast direct LiDAR-inertial odometry. *IEEE Transactions on Robotics*, 38: 2053–2073.

Author biography



WANG Xuanzheng, male, born in 1986, earned his Master's degree from Hoseo University, South Korea, in 2012. He is currently a PhD candidate at Nankai University. He joined CNOOC Safety & Technology Services Co., Ltd. in 2012. He has presided over nearly 20 scientific research projects, with multiple achievements reaching internationally-advanced levels. He spearheaded digital

benchmark projects and developed digital-intelligent safety products, contributing to the high-quality development of the company.