CN 51-1781/TN ISSN (Print) 2096-4579 ISSN (Online) 2097-3993

Light-perception-based interactive control of an underwater digital twin hand

Jinlong Lu, Chao Zhang, Hengchang Nong, Dongying Wang, Hongyu Zhou, Junjie Weng, Yuehua Deng, Yang Yu*, Qiang Bian, Jianfa Zhang, Chaofan Zhang, Zhenrong Zhang and Junbo Yang

Citation: Lu JL, Zhang C, Nong HC et al. Light-perception-based interactive control of an underwater digital twin hand. *Opto-Electron Adv* **9**, 250311 (2026).

https://doi.org/10.29026/oea.2026.250311

Received: 11 November 2025; Accepted: 16 March 2026; Published online: 27 April 2026

Opto-Electronic Advances

Article

July 2026, Vol. 9, No. 7

Open Access

Light-perception-based interactive control of an underwater digital twin hand

Jinlong Lu[1,2†], Chao Zhang[1,3†], Hengchang Nong[1,2†], Dongying Wang[4†], Hongyu Zhou[1,5], Junjie Weng[6], Yuehua Deng[1,7], Yang Yu[1]*, Qiang Bian[4], Jianfa Zhang[8], Chaofan Zhang[9], Zhenrong Zhang[2] and Junbo Yang[1]

Abstract: With increasing demand for ocean exploration, underwater tasks become more complex. While soft robotic hands offer superior flexibility and biomimetic advantages over rigid robotic hands, their development has been hindered by low sensing accuracy and poor interactive control. Here, we propose a light-perception-based underwater digital twin hand interactive control system that integrates a multi-degree-of-freedom biomimetic pneumatic soft robotic hand with a wearable data glove based on fiber Bragg grating (FBG) sensors. By fusing data from FBGs, RGB cameras, and inertial measurement units, and optimizing control strategies through reinforcement learning, we established a light-dominated perception-control-feedback closed-loop system that significantly enhances the control performance of the system in complex and dynamic underwater environments. With light as the information carrier, the system integrates a data glove with high strain sensitivity (1.148 pm/με, linearity = 0.9996) and high bending sensitivity (26.667 pm/°, linearity = 0.9791). It achieves four-dimensional collaborative interactive control between the human operator, machine, virtual space, and environment. The system achieves posture monitoring and precise grasping in various underwater environments, offering advantages like low cost, high precision, rapid response, and strong interference resistance. This establishes a new paradigm for underwater intelligent equipment by integrating light sensing with digital twin technology, offering broad prospects for future underwater operations.

Keywords: fiber Bragg gratings (FBGs); human-machine interaction; underwater soft robotic hand; digital twin; deep learning

DOI: 10.29026/oea.2026.250311 | CSTR: 32247.14.oea.2026.250311

Citation: Lu JL, Zhang C, Nong HC et al. Light-perception-based interactive control of an underwater digital twin hand. *Opto-Electron Adv* **9**, 250311 (2026).

1 Introduction

As global resource exploitation accelerates, the ocean is emerging as a critical domain for human activity and development[1]. However, the complexity and harshness of marine environments pose significant challenges for deep-sea exploration and development. Intelligent submersible equipment, typified by remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs)[2,3], has achieved revolutionary breakthroughs in underwater detection and marine engineering due to its high safety and adaptability. As the core executive components of such equipment, biomimetic underwater robotic hands offer new strategies for improving operational precision and reliability in complex marine environments through integrated innovations in multi-degree-of-freedom (DOF) motion control

Received: 11 November 2025

Accepted: 16 March 2026

Published online: 27 April 2026

[1]College of Science, National University of Defense Technology, Changsha 410073, China; [2]Key Laboratory of Multimedia Communication and Network Technology in Guangxi, School of Computer, Electronics and Information, Guangxi University, Nanning 530004, China; [3]School of Automation, Central South University, Changsha 410075, China; [4]College of Meteorology and Oceanography, National University of Defense Technology, Changsha 410073, China; [5]School of Optical and Electronic Information, National Engineering Research Center of Next Generation Internet Access-system, Huazhong University of Science and Technology, Wuhan 430074, China; [6]College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410073, China; [7]School of Physics, Central South University, Changsha 410083, China; [8]Hunan Provincial Key Laboratory of Novel Nano-Optoelectronic, Information Materials and Devices, National University of Defense Technology, Changsha 410073, China; [9]College of Advanced Interdisciplinary Studies, National University of Defense Technology, Changsha 410073, China .

[†]These authors contributed equally to this work.

*Correspondence: Y Yu, E-mail: yuyang08a@nudt.edu.cn

and intelligent perception. Amid accelerating underwater exploration, intelligent robotic hand systems with perception capabilities and rapid responsiveness hold great promise for practical applications and carry significant research value.

Conventional rigid robotic hands struggle to meet the demands of high-precision underwater operations due to limited flexibility and poor environmental adaptability. The emergence of soft robotic hands offers transformative potentia[4−13]: Their compliant structures and continuous deformation capabilities help minimize mechanical damage to marine organisms and adapt effectively to irregular and constrained workspaces, exemplified by Galloway et al. using soft end-effectors for non-destructive benthic sampling[14]. However, the extreme underwater conditions such as currents, high pressure, and low visibility place stringent demands on the actuation, sensing, and human-machine interaction (HMI) of soft robotic hands[11].

Current underwater robotic hands are primarily driven by hydraulic systems, cable mechanisms, shape memory alloys (SMAs), or pneumatic actuation[6,15−17]. Hydraulic actuators provide high output force but are bulky, complex, and costly[6]. Cable-driven actuators offer lightweight and flexible designs suitable for compact ROVs, but suffer from slow dynamic response and dependency on waterproofing. SMA-based actuators are highly integrated, miniaturized, and resistant to electromagnetic interference. However, their limited force output and slow response lead to significant hysteresis[7]. In contrast, pneumatic actuators have attracted increasing attention due to their lightweight structure, environmental compatibility, and rapid response-traits that make them particularly well suited to soft robotic hands[18,19]. For instance, Wu et al. developed a pneumatic biomimetic soft gripper inspired by luminescent octopus suckers, capable of generating substantial gripping force for adaptive grasping of various objects[20]. In another study, they proposed a fluid-driven multifunctional soft robotic gripper with independently controlled arms, which could perform eight distinct grasping modes for different objects and achieve omnidirectional crawling and swimming, enabling dexterous operations beyond the reach of conventional rigid robotic hands[21].

Although underwater soft robotic hands have seen progress in biomimetic structural design, their perception and interaction capabilities remain underdeveloped. This limitation restricts their application in high-precision tasks and intelligent interaction. Therefore, enhancing these capabilities has become a key focus in underwater robotics research.

Achieving precise control and enhanced perception in soft robots often requires the integration of flexible sensors into the system[22−25]. Flexible sensors are designed to minimize constraints on actuator motion while capturing pressure, stress, and curvature to provide critical data such as posture and position[26−28]. Common types of flexible sensors include electrical and optical sensors. These enable the detection of posture changes, structural curvature, and applied pressure or stress in soft robots. For example, S. Frishman et al. proposed a two-finger gripper using a single optical fiber embedded with 17 fiber Bragg gratings (FBGs) to measure normal and biaxial shear stresses on the gripper pad[29]. Zhao et al. reported the use of stretchable optical waveguides as curvature and force sensors integrated into a soft prosthetic hand[24]. Teeple et al. used flexible optical waveguide sensors embedded in deep-sea soft robots to maintain stable signals under pressure changes of 0-4000 psi (1 psi = 6.895 kPa), enabling proprioception and contact force sensing[30]. Electrical sensors typically rely on the deformation of channels filled with conductive liquids (e.g., liquid metal) or piezoresistive materials to change resistance/capacitance; however, under high pressure, they may undergo compression deformation, leading to shifts in baseline resistance/capacitance. Furthermore, electrical sensors require more rigorous encapsulation to prevent short circuits or signal drift. In contrast, optical sensors, such as all-fiber sensors based on FBG, are naturally immune to electromagnetic interference due to their optical signal transmission mechanism. Moreover, sensor arrays can be integrated into a single optical fiber to achieve high-resolution distributed measurement. For example, Wang et al. (2020) designed an FBG array specifically for ocean observation[31], successfully capturing intense frontal structures with temperature gradients of up to 1.2 °C/m during trials-details that traditional methods struggled to distinguish. These characteristics of high reliability and high precision, inherent to optical principles, make "light perception" superior to traditional electrical sensing in underwater environments, rendering it an ideal choice for underwater perception.

Common interaction methods in HMI technologies include vision-based approaches[32,33], electromyography (EMG) signals[34−38] and data gloves[39−42]. For instance, visual features of the human hand can be extracted via camera-based image acquisition and algorithmic interpretation. Rosen et al. employed EMG signals as the primary control input for an exoskeleton system, constructing an HMI interface at the neuromuscular level to predict muscle torques around the elbow joint[43]. Nelson et al. integrated a 4×3 capacitive sensor array into denim fabric using conductive threads, combining Hidden Markov Model and Dynamic Time Warping to enable small-sample training, achieving an average gesture recognition accuracy of 99% across five gestures[44]. In practical HMI, data gloves provide strong environmental adaptability. They offer rapid response and a lightweight experience, enabling more natural and friendly interaction with users.

Underwater environments are often characterized by low visibility and frequent disturbances caused by sediment and water currents. In such complex conditions, posture monitoring of robotic hands becomes particularly important. Traditional camera-based vision systems have limited capability in tracking hand posture under low-visibility and turbid conditions. To improve interactive control of robotic hands in these environments, we introduce FBG sensors in addition to conventional vision monitoring to enable more reliable posture sensing.

In this work, we present a modular integration of an underwater soft robotic hand with a flexible sensing system. The system includes a Biomimetic Pneumatic Soft Manipulator (BPSM), which consists of a three DOF soft robotic arm inspired by an elephant's trunk and a five-finger anthropomorphic soft robotic hand with a rotating wrist. We also introduce an FBG-based sensing glove (FBG-SenseGlove)[45]. Each finger sleeve integrates a single FBG sensor to detect bending and capture hand motion. The wearable glove enables bidirectional HMI. Real-time hand movements are visualized in a virtual environment via digital twin technology, completing the interaction process between the operator and the underwater BPSM. As illustrated in Fig. 1(a), the light-perception-based underwater digital twin hand interactive control system (LP-UDTHS) supports a wide range of applications. The FBG-SenseGlove can be worn by divers, shore- or vessel-based operators, or crew members inside piloted submersibles. The ROV carries a BPSM fitted with an FBG-SenseGlove for underwater tasks such as scientific research, exploration, and cable maintenance in vortex-prone, turbid, or muddy environments. Operators control the soft hand via hand motions, with posture mirrored in real time through the digital twin interface. To enhance control robustness in dynamic underwater environments, we integrate multi-source sensing (FBG-SenseGlove, RGB cameras, inertial measurement units(IMUs)) with a reinforcement learning (RL)-based control optimization framework (Fig. 1(b)). We constructed a custom RL simulation environment based on the unified robot description format (URDF), incorporating defined state spaces, reward functions, and domain randomization. The BPSM performs agile underwater manipulation in real time. The operator monitors hand movements within the virtual environment and issues commands accordingly. The diver wears the FBG-SenseGlove to perform standard underwater gestures. The resulting FBG wavelength shifts are transmitted to the host system, where they are mapped to finger bending angles. These angles are subsequently converted into actuation commands to drive the soft hand in executing the corresponding motions (Fig. 1(c)). Real-time visualization enables comprehensive monitoring and control of human and machine motions, encompassing biomimetic grasping, underwater object retrieval, and other interactive tasks. This integrated perception-control framework establishes a closed-loop coordination among human, machine, environment, and virtual space, opening new possibilities for underwater human-machine collaboration.

2 Results and discussion

2.1 BPSM: Structure and operating principle

The overall structure of the BPSM is shown in Fig. 2(a). It consists of a five-finger anthropomorphic soft hand, a rotary wrist, and a soft arm inspired by an elephant's trunk. The system adopts a modular design to support easy maintenance, component replacement, and structural upgrades. As shown in Fig. 2(b), the dexterous hand comprises pneumatically actuated biomimetic fingers mounted on a palm module. The index, middle, ring, and little fingers each have a three-joint design (105 mm in length), while the thumb uses a two-joint structure (85 mm) with an adduction mechanism for opposable grasping. The palm serves as the assembly base and includes dedicated mounts for each pneumatic actuator. The palm module is 3D printed using thermoplastic polyurethane (TPU) material (dimensions: 98 mm × 89 mm) and provides structural rigidity. A flexible silicone layer coating enhances overall compliance and softness. Pneumatic channels for each finger are embedded within the palm. The thumb is connected to the palm via a dedicated abduction structure, enabling adduction and abduction movements through a predesigned pneumatic chamber.

The rotational wrist module serves as the junction between the palm and the arm, comprising a waterproof servo motor and a cylindrical support housing (51 mm in height, 62 mm in diameter). This configuration enables 180° rotation of the palm, significantly enhancing spatial flexibility at the end effector. The biomimetic flexible arm mimics an elephant trunk and consists of a tri-chamber pneumatic actuator measuring 140 mm in total length. The actuator contains three equally distributed internal chambers. By varying the pressure applied to each chamber, the arm can bend in multiple directions, allowing the dexterous hand at the end to achieve multi-DOF posture adjustments. Both ends of the arm are equipped with standardized interfaces to facilitate integration with the palm module and ROV platforms.

The three pneumatic actuators composing the biomimetic dexterous hand are inspired by the structural design of human finger joints, featuring a rigid-flexible hybrid and multi-layer composite architecture. Each actuator consists of a flexible segment that mimics finger bending joints for controllable deformation and a rigid segment that simulates the finger bones to provide necessary support and stiffness. This design enables natural bending under pneumatic actuation, giving the biomimetic hand both compliance and precision in grasping and manipulation tasks.

All three actuators adopt a multilayer structural design (Fig. 2(c)), comprising an inner pneumatic chamber, a fiber-reinforcement layer, a cotton fabric constraint layer, a biomimetic rigid groove layer, and an outer encapsulation layer. The inner chamber serves as the primary driving unit, undergoing axial elongation and volumetric expansion under

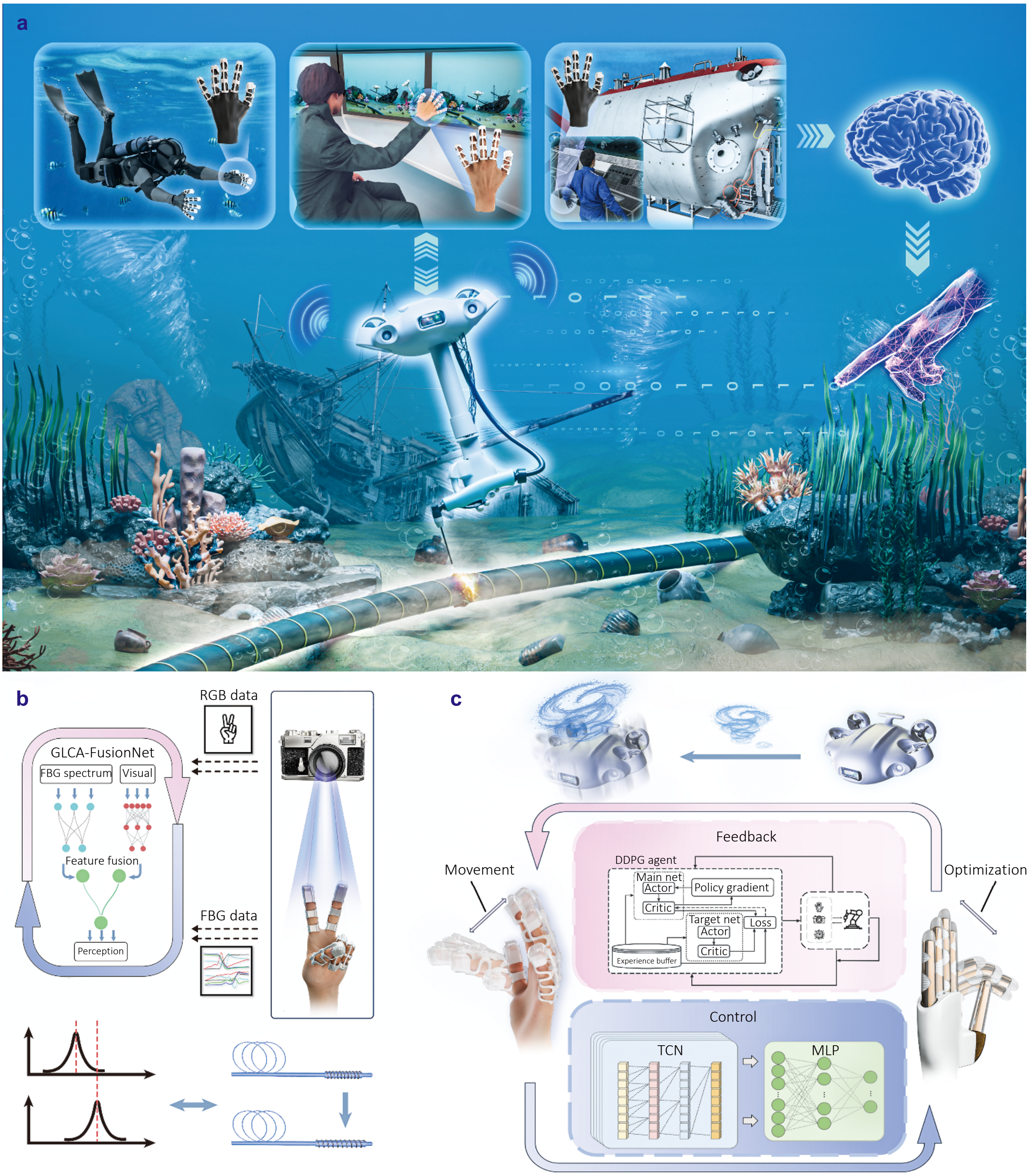

Fig. 1 | LP-UDTHS system overview. (a) Conceptual diagram of the LP-UDTHS system. The FBG-SenseGlove can be worn by divers, shore- or vessel-based operators, or crew members inside piloted submersibles. An ROV equipped with the BPSM integrated with an FBG-SenseGlove performs underwater tasks in complex environments with vortices, turbidity, or sediment disturbance. Operators remotely control the soft hand via hand motions and achieve real-time posture monitoring through digital twin technology. (b) When the FBG-SenseGlove bends, its grating center wavelength shifts. FBG and image data are extracted from the FBG-SenseGlove and RGB cameras respectively. Features are then fused through the multimodal network Global and Local Cross-Attention Multimodal Fusion Network (GLCA-FusionNet), enabling gesture recognition with higher robustness than single-modality methods. (c) We train the system in simulated underwater environments with complex disturbances using a custom simulation model and reinforcement learning parameters. This approach successfully improves multimodal digital twin control accuracy and adaptive posture correction of the BPSM in practical applications.

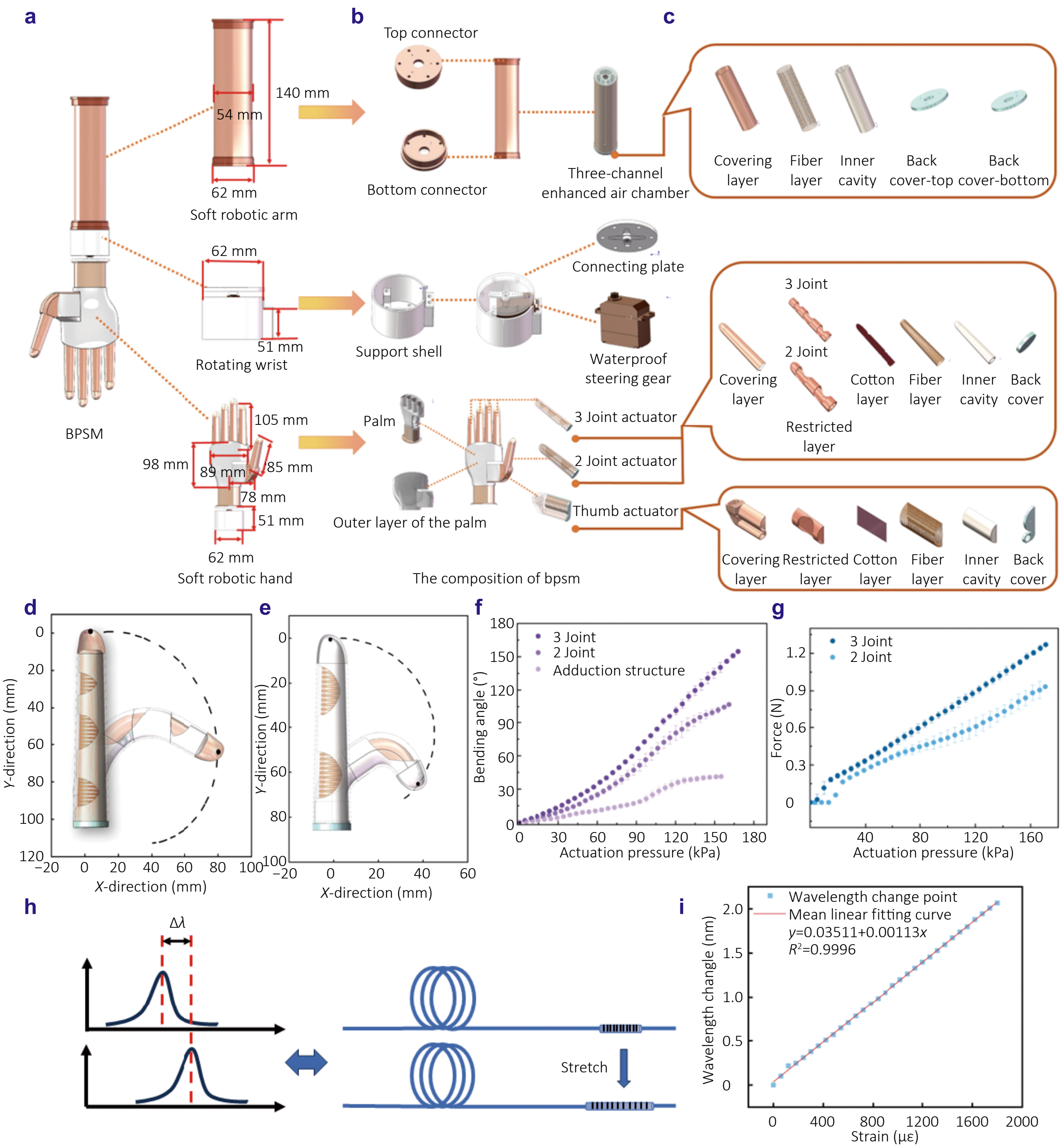

Fig. 2 | Design, working principles, and performance characterization of the BPSM and FBG-SenseGlove. (a) The BPSM consists of a trunk-inspired soft arm, a rotational wrist, and a five-finger anthropomorphic dexterous hand. (b) Modular components and interconnecting parts of the BPSM, designed for structural adaptability. (c) Multilayer construction of the soft arm and the three types of pneumatic actuators in the biomimetic hand. (d) Tracked fingertip trajectory of the three-joint soft finger used to characterize its workspace. (e) Tracked fingertip trajectory of the two-joint soft finger used to characterize its workspace. (f) Bending angles of the three actuator types under different pneumatic input pressures. (g) Fingertip force measurements of the three-joint and two-joint soft fingers at varying input pressures. (h) Schematic of the FBG sensing principle. (i) Experimental results and fitted curve showing the relationship between FBG tensile strain and center wavelength shift.

applied pressure. The fiber-reinforcement layer enhances structural strength while effectively restraining radial expansion. A single-sided cotton fabric layer induces asymmetric deformation and guides the actuator to bend unidirectionally. The rigid groove layer is inspired by human finger joint proportions and defines bending regions and angles to enable precise control of motion. The outermost layer, made of flexible material, encapsulates all functional layers to ensure structural integrity, compliance, and sealing performance. Each actuator is equipped with a standardized pneumatic port at its base which allows seamless connection to a multi-channel pneumatic control system for flexible and coordinated actuation. The detailed fabrication process of the BPSM is provided in Supplementary Section 1.

The deformation mechanism of the pneumatic actuators in the soft fingers relies on the asymmetric response behavior induced by differences in material stiffness across the multilayer structure. When pressurized air is introduced into the inner chamber, the cavity tends to expand in all directions. However, the fiber-reinforcement layer constrains radial expansion, directing the deformation primarily along the longitudinal axis. Simultaneously, the cotton fabric and biomimetic rigid groove layers, located on one side of the structure, impose asymmetric constraints that guide the actuator to bend unilaterally. Bending occurs at pre-defined compliant joint regions and enables the actuator to mimic the natural flexion of human finger joints under pneumatic actuation. Precise posture control and coordinated actuation of multiple fingers in the biomimetic soft hand can be achieved by leveraging this asymmetric deformation strategy and regulating the input pressures of individual pneumatic actuators.

The proposed BPSM is pneumatically driven. We designed a 12-channel pneumatic control system based on proportional-integral-derivative (PID) closed-loop feedback control to achieve precise regulation of positive and negative pressure output. Detailed control methods are provided in Supplementary Section 3.

2.2 Characterization of the elastic pneumaticchambers in soft fingers

We evaluated the system-level performance of the BPSM by experimentally measuring key parameters of its soft finger pneumatic actuators. These include workspace coverage, bending response and fingertip output force. We used a camera-based tracking setup together with OpenCV image processing to determine the workspace. Motion trajectories of both the three-joint and two-joint soft fingers were visualized and analyzed (Fig. 2(d, e)). Distinctive color markers were affixed to the base and fingertip of each actuator, and their positions were extracted using HSV-based color thresholding and centroid localization, calibrated to physical scale via a pre-established pixel-to-length ratio. We tracked the fingertip trajectory under varying input pressures to obtain the two-dimensional reachable workspace of each finger. Simultaneously, the bending angle was estimated by calculating the inclination between the fingertip trajectory vector and the baseline marker. As shown in Fig. 2(d), the workspace of the three-joint index finger exhibits a crescent-like profile with a maximum horizontal span of 79 mm and a vertical reach of 113 mm. Similarly, the thumb (2 Joint configuration) displays a crescent workspace spanning 45 mm horizontally and 73 mm vertically (Fig. 2(e)).

To further evaluate the pressure-dependent actuation performance, we measured the static bending angles of the fingers under varying pneumatic inputs and established pressure-angle response curves. As shown in Fig. 2(f), the bending angle of each actuator increases monotonically with applied pressure, demonstrating a well-behaved, controllable pneumatic response. For the 3 Joint actua tors, a maximum bending angle of $(154.7 \pm 6.0)°$ ($N = 3$) was achieved at 170 kPa. The two-joint actuators reached their peak bending angle of $(106.9 \pm 6.9)°$ ($N = 2$) at 160 kPa. For the thumb adduction actuator, a maximum bending angle of $(41.3 \pm 3.4)°$ ($N = 3$) was recorded at 155 kPa. These results confirm the system's effective modulation of joint angles through pressure variation, enabling precise and repeatable shape adaptation.

To assess the mechanical performance of the soft robotic hand during grasping tasks, we measured the fingertip output force using a force sensor. The sensor was vertically mounted at the fingertip's contact surface to record the output force at maximum bending states for each actuator configuration. As shown in Fig. 2(g), the 3 Joint pneumatic actuator achieved a maximum fingertip output force of (1.26 ± 0.03) N ($N = 3$), while the two-joint actuator reached a maximum output force of (1.04 ± 0.07) N ($N = 3$). To validate the grasping capability of the BPSM, we evaluated its adaptability in underwater environments using the method of reproducing common human grasping postures[46]. As shown in the Supplementary Figure, the soft hand successfully reproduced 33 typical human grasping postures, covering various typical grasping modes including power grasps, lateral pinches, and precision tripod grasps. This fully demonstrates its excellent grasping performance and motion flexibility, capable of meeting the needs of grasping daily lightweight objects and interactive operations. These results indicate that the actuators can meet the force requirements for handling lightweight objects and performing interactive operations.

2.3 FBG-SenseGlove architecture and operating principle

Integrating flexible sensors directly into the body of soft robots can enhance system lightweighting, improve response speed, and often offer superior mechanical compatibility. However, such monolithic integration typically requires more complex fabrication processes and inherently limits system scalability. In the event of localized sensor failure, the entire system may require replacement, increasing maintenance costs. Instead, we decouple the soft robotic hand and the sensing components, integrating the flexible sensors into a lightweight and structurally simple data glove, thereby enhancing modularity, ease of maintenance, and system robustness.

The FBG-SenseGlove detects subtle movements of the wearer's hand based on the unique Bragg wavelength characteristic of FBG that reflects specific wavelengths of light. The reflected light is identified using an FBG interrogator[39,40]. The FBG center wavelength λ_B satisfies the Bragg condition:

$$\lambda_B = 2n_{\text{eff}}\Lambda \ , \tag{1}$$

where λ_B denotes the Bragg wavelength, n_{eff} is the effective refractive index of the optical fiber, and Λ represents the grating period. Strain and temperature variations cause shifts in the Bragg wavelength. The resultant wavelength shift $\Delta\lambda_B$ is expressed as:

$$\Delta\lambda_B = 2n_{\text{eff}}\Lambda\left(1-P_\varepsilon\right)\Delta\varepsilon + 2n_{\text{eff}}\Lambda\left(\alpha+\xi\right)\Delta T \ , \tag{2}$$

where $\Delta\lambda_B$ denotes the wavelength shift and $\Delta\varepsilon, \Delta T$ denote the variations in strain and temperature, respectively. P_ε, α, and ξ correspond to the fiber's photoelastic coefficient, thermal expansion coefficient, and thermo-optic coefficient. Assuming a constant temperature T, the strain ε can be calculated from the measured wavelength shift based on the above equation. When the FBG experiences stress, the Bragg wavelength shifts accordingly, as illustrated in Fig. 2(h). By leveraging the relationship between strain and the Bragg wavelength, high-precision measurements of small physical changes such as finger bending angles can be achieved.

The encapsulation of the FBG is a crucial step to prevent its fracture during finger flexion. The designed data glove structure consists of the finger covers and a silicone protective layer, with the silicone layer serving to encapsulate the FBG and prevent breakage. A non-toxic, odorless silicone material, suitable for direct skin contact, is used as the base material for the glove, with an elongation of up to 672% and tensile strength of 440 psi. Each finger's silicone protective layer houses a single FBG, positioned at the proximal interphalangeal joint (PIP). This location is selected due to its significant contribution to the overall finger movement in most hand gestures, allowing the effective capture of key flexion motions for gesture monitoring.

2.4 FBG-SenseGlove performance characterization

To characterize the strain response of the FBG sensor while eliminating temperature interference —since FBGs are intrinsically sensitive to both strain and temperature —all tests were performed in a temperature-controlled laboratory environment. Tensile forces were applied directly to the grating region of the FBG to induce tensile strain. According to the mechanics of an equal-strength cantilever beam, each 0.2 N of applied force corresponds to a surface strain of approximately 240 με. The strain was progressively increased from 0 to 1800 με, and the corresponding wavelength shifts were monitored. As shown in Fig. 2(i), the obtained results exhibited a strong linear relationship between strain and wavelength shift. A linear regression yielded the equation y=0.03511+0.00113x, corresponding to a sensitivity of 1.148 pm/με and a linearity (R^2) of 0.9996. The maximum observed wavelength shift reached 2.067 nm. The standard deviation across measurements was as low as 4.22 pm which indicates excellent repeatability. These results confirm that the FBG sensor possesses high sensitivity, strong linearity, and stable performance under applied strain.

We then conducted a bending response test on the encapsulated FBG-SenseGlove to evaluate changes in the Bragg wavelength under different bending angles. A 3D-printed biomimetic finger with articulated joints was used to wear the FBG-SenseGlove and was mounted on a precision digital goniometer. The goniometer was employed to calibrate the glove's bending angle. This setup allowed the simulation of mechanical stress induced by finger flexion and enabled accurate measurement of the corresponding Bragg wavelength shifts. The results show that the sensor exhibited a Root Mean Square Error of 0.12148 nm, a linearity of 0.9791, and a bending sensitivity of 26.667 pm/°. These findings confirm that the sensor maintains high stability and sensitivity even after silicone encapsulation.

2.5 Multimodal posture perception

The designed FBG-SenseGlove enables bidirectional pose perception: it can be worn by human operators (e.g., diver) for real-time capture of hand gestures or mounted onto the five fingers of the BPSM to provide real-time feedback on the soft robotic hand's deformation state.

We adopted six common underwater interaction gestures (Fig. 3(a)) to validate the effectiveness, stability, and performance of the FBG-SenseGlove in capturing hand posture information from both the diver and the soft robotic hand in underwater environments. The data acquisition for each gesture was completed within a 3-second period. Each sample consisted of five FBG wavelength channels with 4000 floating-point values per channel. In total, 2400 samples were collected to construct a dataset for underwater gesture recognition under environmental conditions. After shuffling, the dataset was divided into training and validation sets with a ratio of 8∶2. To process the FBG wavelength signals we developed a dedicated deep learning architecture Modular Temporal Convolutional Networks-Underwater AdaptiveNet (MTCN-UANet) optimized for extracting features from spectral time-series data (Fig. 3(b)). The final model achieved a validation accuracy of 100% on the BPSM input set and 99.39% on the diver dataset (Fig. 3(c)). To

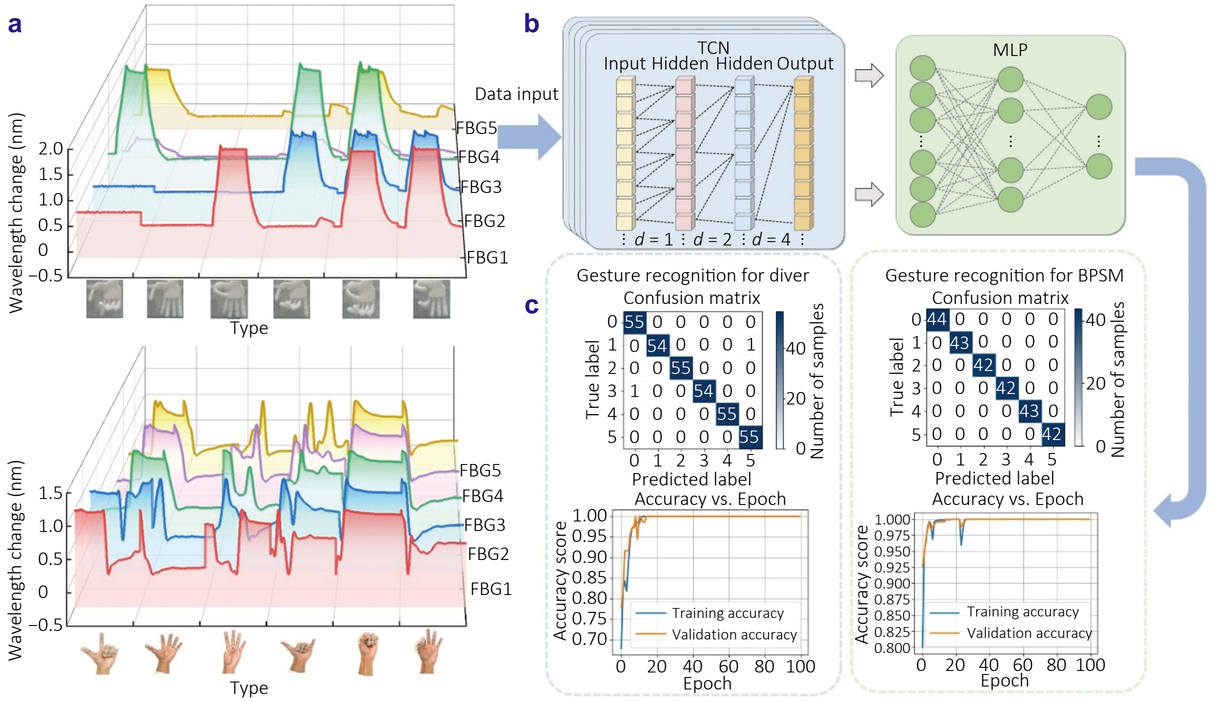

Fig. 3 | Multimodal Pose Perception for Human Hand and BPSM Gestures. (a) FBG data collection of six common underwater interaction gestures for both human hands and BPSM using the FBG-SenseGlove. The graph illustrates how the Bragg wavelength shifts with respect to each gesture, showing distinct patterns for each. FBG1 to FBG5 correspond to the thumb, index, middle, ring, and little fingers of the hand, respectively. (b) The architecture of the deep learning model MTCN-UANet, specifically designed for extracting FBG center wavelength features. (c) Gesture recognition results and classification confusion matrices for six gestures of both diver hands and BPSM.

further assess its generalization, we performed stratified 5-fold cross-validation, which demonstrated strong classification performance and robustness in complex environments (Fig. S2).

Against the backdrop of increasingly refined intelligent interaction control relying solely on a single sensor modality is no longer sufficient to meet the demands for high precision and robustness in manipulation tasks. In our posture perception tests, specifically within the complex environment test set simulating turbid water and flow disturbances, our quantitative benchmark results are as follows: Under the single vision modality, traditional models (e.g., CNN, MobileNetV2) exhibited low test accuracy in complex environments, at 21.25% and 28.75%, respectively. Under the single FBG modality, MTCN-UANet remained robust due to the advantages of optical fiber sensing, achieving 96.25% accuracy in this test set.

By introducing four-camera multi-view fusion and global and local cross-attention (GL attention), the model leverages the cross-attention mechanism to deeply mine the spatiotemporal complementary characteristics of FBG and images, achieving superior fusion effects. Under full-modality input (five FBG sensors + four-view images), its accuracy reached 99.2%. This indicates that although MTCN-UANet itself already possesses extremely high precision, GLCA-FusionNet can effectively correct subtle misjudgments (an improvement of approximately 2.70%) that may occur in a single modality under extreme environmental interference, which is crucial for high-reliability underwater interaction. Remarkably, even with only a single FBG sensor and one camera view it still attains a test accuracy of 86.93% significantly outperforming unimodal approaches. This demonstrates how multimodal data fusion can fully exploit the complementary strengths of heterogeneous sources and thereby enhancing the system's overall perception capabilities and substantially improving the accuracy and fidelity of complex hand gesture recognition.

2.6 Light-perception-based interactive control system framework and underwater operations

We selected Python as the development language to construct a fully functional, modular, and highly extensible multimodal interactive system. We adopted a modular design approach where each sensor or control device has its own dedicated interface package to enhance system cohesion and maintainability. In terms of communication, the system acquires FBG sensor data from the FBG-SenseGlove via the UDP Ethernet protocol, while RGB image data is retrieved using the OpenCV library through UVC protocol-based video streams. After temporal synchronization and spatial alignment of all multimodal data, we fuse and process them to control multiple output devices including BPSM controlled via serial communication and a 3D virtual hand model interacting through UDP protocol. This setup facilitates end-to-end real-time control and feedback loops.

The architecture of the proposed interactive control system is illustrated in Fig. 4(a). The BPSM is actuated by a pneumatic control system based on closed-loop PID feedback that enables precise and adjustable bidirectional pressure output. Details of the implementation are provided in Supplementary Section 3. The control system also includes a host computer responsible for input signal mapping, control command transmission, and virtual environment rendering. Within this interactive system the FBG-SenseGlove functions as the gesture input interface. The Bragg wavelengths from the FBG sensors are acquired by an FBG interrogator and transmitted to the host computer, where they are mapped to joint angle variations using the calibrated strain–wavelength relationship. These joint angles are further mapped to target pneumatic pressures using a predefined pressure-angle mapping model. The resulting pressure commands are sent to the pneumatic control system where a microcontroller unit (MCU) receives and decodes the instructions via a dedicated parsing algorithm and generates control signals for the solenoid valves. This allows accurate regulation of air pressure to drive the soft hand and achieve the desired bending motions.

We devised a practical deployment scenario in which the entire BPSM is mounted onto an ROV via custom connectors and operated within an underwater test tank. One FBG-SenseGlove is attached to the BPSM to monitor its real-time bending and motion posture, while another FBG-SenseGlove is worn by a human operator to capture hand gestures and control the soft robotic hand accordingly, as illustrated in Fig. 4(b).

Underwater soft hand posture monitoring. When the operator moves their index and middle fingers, the FBG sensors embedded in the wearable glove detect shifts in the Bragg wavelength which are mapped to corresponding finger bending angles. These angles are transmitted as control commands to the host computer which then translates them into the target bending angles of the underwater soft hand and calculates the required actuation pressure. The control signals are subsequently sent to the slave system, which actuates the corresponding soft fingers underwater to achieve the desired bending posture. Simultaneously, the FBG sensors integrated into the glove mounted on the soft hand respond to the resulting deformation and provide real-time feedback by registering changes in Bragg wavelength. This process enables continuous monitoring of the soft hand's posture as illustrated in Fig. S4, thereby achieving closed-loop posture tracking and interactive control between the FBG-SenseGlove and the soft robotic hand.

Underwater Biological Model Grasping. We demonstrate how an operator remotely controls finger flexion to interactively guide the BPSM for underwater retrieval tasks (Fig. 4(c) and Movie S1). The operator drives the ROV into the underwater environment while monitoring the system camera. When the ROV approaches the target biological

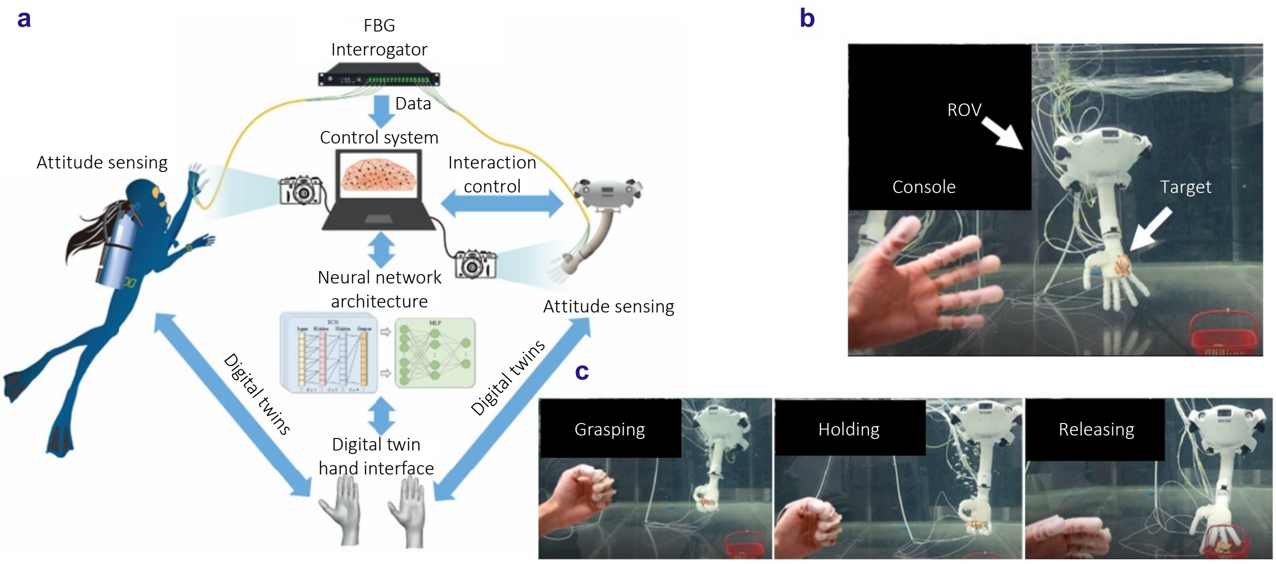

Fig. 4 | Interactive control system based on optical sensing. (a) Schematic of the optical sensing-based interactive control system. Both the BPSM and the simulated underwater diver wear the FBG-SenseGlove, which collects FBG data and RGB image data. These data are then fused and processed to control the motion of the BPSM and dynamically map the entire hand posture onto a 3D virtual hand model, enabling digital twin visualization. (b) Illustration of the simulated underwater grasping of a clownfish model. (c) Dynamic control of the underwater grasping experiment, including the states of grasping, holding, and releasing the clownfish model.

model (e.g., clownfish, hermit crab, octopus model, etc.), it halts. At this point, the land-based operator begins controlling the soft glove and can adjust the soft arm using the joystick and the wrist rotation using the potentiometer. Together, these controls allow the operator to perform the grasping operation on the underwater model. The underwater soft hand bends to the mapped angle and successfully grasps the biological model. The operator then lifts the model, drives the ROV back to the designated basket, and releases the model, completing the underwater retrieval task.

Underwater retrieval. The operator guided the ROV through the underwater environment while monitoring the onboard camera feed. Upon approaching the designated retrieval target the ROV was brought to a halt. The land-based operator then engaged the wearable glove to control the soft robotic hand. The underwater soft hand bent to the mapped angle and securely grasped the target basket. Subsequently, the ROV transported the retrieved item to a predefined location, thereby completing the underwater retrieval task, as demonstrated in Movie S1.

2.7 Interactive control of underwater digital twins

We employed Python scripts to precisely bind the joints of the virtual soft robotic hand, enabling real-time angular mapping of all FBG signals from the data glove within the virtual environment. This setup allowed immediate visual feedback of the 3D virtual hand model in response to glove input signals. The BPSM integrated with the FBG-SenseGlove exhibited high-fidelity synchronization with human hand motions thereby accomplishing digital twin-based visual representation.

To enhance the control performance of the digital twin system in dynamically complex underwater environments, we integrated the FBG-SenseGlove with an RGB camera and IMU sensors, and adopted reinforcement learning to optimize the control strategy. A custom simulation environment was developed using URDF-based modeling, incorporating specifically designed state spaces, reward functions, domain randomization, and other parameter optimization schemes. We specifically compared the motion accuracy and dynamic performance of the traditional PID controller, the Model Predictive Control (MPC) method, and the Deep Deterministic Policy Gradient (DDPG) reinforcement learning algorithm in executing complex tasks. The final results show that compared to the traditional PID controller and MPC method, the DDPG algorithm reduced the position error (Epos) by 5.4 mm and 2.1 mm, and the velocity error (Evel) by 8.1 mm/frame and 4.6 mm/frame, respectively. This successfully improved the response speed and accuracy of the digital twin hand for complex gestures and actions. Detailed training environment and parameter designs are shown in Supplementary Section 5. This framework significantly improved the response speed and accuracy of the digital twin hand in replicating complex gestures and motions.

As illustrated in Fig. 5(a), the system diagram depicts the underwater human-machine interactive control framework based on a digital twin. The BPSM equipped with an FBG-SenseGlove is mounted onto the ROV via a custom connector. The ROV serves as an underwater submersible platform, while the BPSM functions as the execution unit for deep-sea operations. The underwater data glove captures the BPSM's bending and posture dynamics. Meanwhile a human operator located in a shallow-water environment wears a second FBG-SenseGlove to perform underwater operations. Multiple shallow-water scenarios are configured including calm water, flow disturbances, turbid conditions, and combined turbulence with turbidity to simulate underwater operations performed by divers under varying environmental complexities. The glove captures the operator's finger posture and bending motions and uses them to control the movements of the underwater actuator. A PC-based remote monitoring system constructs two parallel 3D virtual hands that dynamically map FBG sensor data into a virtual space. One represents the underwater BPSM, and the other represents the diver's hand, enabling real-time interaction and posture monitoring of both physical entities. Human operators at the remote end can monitor the operational status of both the shallow-water diver and the deep-sea BPSM in real time. Upon receiving commands from the remote operator, the diver performs the instructed motion, which is then mirrored by the BPSM operating in the deep-sea zone. Both the BPSM and the diver's hand postures are fed back to the monitoring system allowing the supervisor to assess the operational state of the underwater system including motion trajectories and potential faults. This feedback loop supports further command issuance and enables an interactive sensing link among the human operator, the machine, the physical underwater environment and the virtual space. Ultimately, the system achieves intelligent interaction and control across human, machine, environment, and virtual space.

We demonstrated that a human operator can achieve real-time control of the underwater BPSM through finger movements, enabling both gesture-based control and underwater grasping, with continuous posture feedback and monitoring at the remote terminal. As shown in Movie S2, under normal shallow-water conditions the underwater operator performs a series of single commonly used gestures progressing from individual finger movements to dual-finger gestures. During this process, both the BPSM and the virtual representations of the operator's hand and the BPSM respond synchronously replicating the intended gestures. The FBG wavelength data, corresponding dual-hand postures, and digital twin monitoring terminals during the entire action process I-IX are recorded in Fig. 5(b). It can be seen that the five FBGs on each glove successfully completed the posture monitoring of individual fingers during actions I-IX without mutual interference, with accurate and rapid

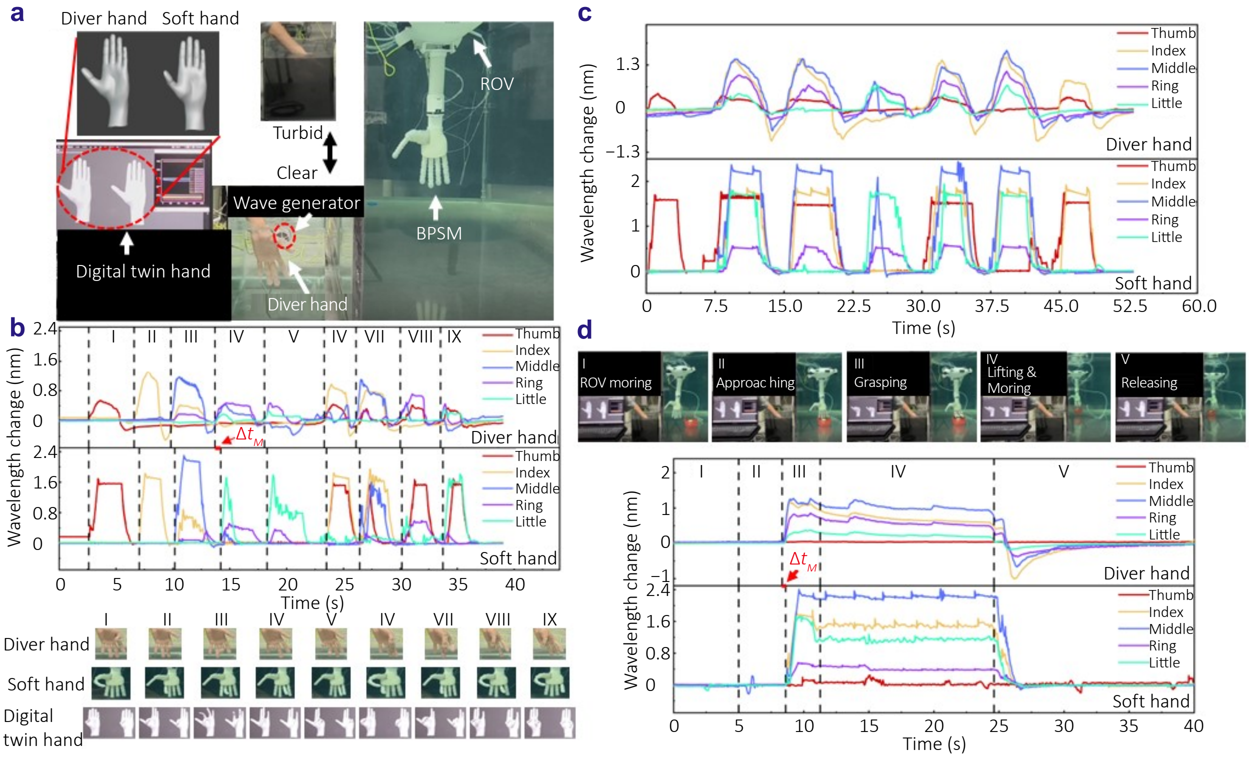

Fig. 5 | Underwater digital twin interactive control. (a) System diagram of the underwater digital twin HMI control. Both the simulated underwater diver and the BPSM wear FBG-SenseGloves for interactive control in the underwater environment. Remote digital twin monitoring of both hands is performed on the PC. The simulated diver's environment includes normal underwater conditions, flow disturbances, turbid water, and combined turbulence with turbidity. (b) Under normal underwater conditions, wavelength shifts of each FBG sensor corresponding to individual fingers vary across nine distinct gestures performed by the simulated diver, alongside the associated gesture motions and digital twin monitoring. (c) In a turbid underwater environment with flow disturbances, wavelength shifts of each FBG sensor corresponding to individual fingers vary according to different gestures performed by the simulated diver. (d) Dynamic underwater retrieval experiment of the BPSM grasping a basket under turbid and flow-disturbed conditions, with temporal variations of wavelength shifts for each finger's FBG sensor.

wavelength change responses. During the full actuation of the BPSM palm, the maximum response time Δt_{m} remains below 0.5 s. Once the BPSM fingers reach their target bending angles, the system maintains stable posture, with minimal fluctuation in the FBG wavelength shifts.

We further evaluated the system under more challenging conditions, including a shallow-water environment with turbidity and flow disturbances. In this setting, the underwater operator performed a series of more complex commonly used underwater gestures. The FBG sensor readings, shown in Fig. 5(c), demonstrate consistent wavelength shifts corresponding to finger movements. As before, all five FBG sensors embedded in the glove effectively tracked individual finger postures, maintaining high responsiveness and signal integrity despite the presence of flowing water and low-visibility conditions. The digital twin system likewise captured the full range of motion of both the human hand and the BPSM as visualized in Movie S2. Movie S2 also documents the underwater dynamic retrieval task performed by the BPSM. As the ROV navigated through the underwater environment and approached the target object the diver received a remote instruction and initiated the grasping action. Bending signals detected by the diver's FBG-SenseGlove were transmitted to the host computer and the host computer in turn issued commands to the pneumatic control system. The control system then actuated the underwater BPSM to execute the corresponding grasping motion. These bending signals were fully mapped onto the remote 3D virtual hand model, allowing the monitoring terminal to verify task completion once the hand movements of both the diver and the BPSM were in agreement. Upon successful grasp, the ROV transported the object to a designated location and released it. The complete operation process, including ROV motion, object approach, diver grasping, BPSM actuation, lifting, and final release, was continuously monitored at the remote end. The FBG data collected from both the diver's glove and the BPSM glove during these key stages is shown in Fig. 5(d). Throughout the entire motion and control sequence the system exhibited minimal latency, with the maximum response delay Δt_{m} remaining below 0.4 seconds.

3 Conclusions

In this work, we developed a multimodal fusion-based LP-UDTHS. The system comprises an underwater soft robotic hand, which is capable of dexterous grasping and movement when mounted on an ROV, a data glove based on FBG for capturing finger posture information. By integrating data from the FBG sensors on the glove, RGB cameras, and IMU sensors, we enable real-time monitoring and visualization of the diver's and soft robotic hand's motions. The data is transmitted into a digital twin environment on the terminal, facilitating the mapping and control feedback of hand postures to validate HMI performance. This system relies on optical perception, enabling a four-dimensional (human-machine-virtual space-environment) intelligent interactive control system. This supports real-time remote operation of the underwater robotic hand and feedback from the virtual environment. The key advantages of our system include simplicity, low cost, fast response, non-destructive interaction, and environmental friendliness. Supplementary Table S6 summarizes the sensing capabilities, interaction capabilities, algorithmic control, and digital twin capabilities of recent intelligent underwater manipulator systems. These characteristics provide a new paradigm for underwater operations such as scientific exploration and underwater retrieval.

Future work for improvement can be done including upgrading the soft finger structure by incorporating variable stiffness flexible modules to enhance grasping strength, adopting more advanced pneumatic control algorithms to further improve control accuracy, integrating additional flexible sensors into the soft hand to enhance its perceptual capabilities, allowing for more environmental information (e.g., pressure, tactile information, etc.) to be gathered, and exploring the use of brain-machine interfaces for direct control of the underwater robotic hand. In addition, deploying FBG sensing systems in deep-sea environments often requires consideration of temperature-pressure cross-sensitivity. Future work could combine the high-sensitivity optical fiber seawater temperature and pressure sensor proposed by Xu et al. to achieve 'decoupling' of temperature and pressure sensing characteristics at the hardware level[47], thereby providing a clear technical path for the deep-sea evolution of this system.

4 Materials and methods

4.1 Posture acquisition method of FBG-SenseGlove

After donning the silicone-based data glove, the five FBGs are connected to a multi-channel FBG interrogator (TV-1600-1000 Hz-16-channel, Tongwei Technology, Beijing, China) via FC/APC connectors. The FBG interrogator collects the Bragg wavelength data of the five FBGs at a sampling rate of 1000 Hz. The working principle of the FBG is based on strain-induced shifts in the fiber Bragg grating's spectrum, which vary with changes in finger joint angles. Therefore, the key to the motion tracking algorithm is to develop an accurate and reliable model that maps the relationship between finger joint angle changes and fiber strain. Before posture data collection begins, the initial mapping parameters need to be adjusted according to the wearer's hand shape. Bragg wavelength data are collected for both fully extended and fist positions, after which hand posture information can be captured. The mapping transformation for a single finger can be expressed as:

$$F = (W - W_{\min}) * \frac{k}{(W_{\max} - W_{\min})} \, . \tag{3}$$

In this expression, W denotes the currently acquired wavelength value of the finger, while $W_{\max}$ and $W_{\min}$ represent the wavelength values recorded during full fist closure and full extension, respectively. The parameter k is a normalization coefficient, typically set to 90. The final output F corresponds to the mapped flexion value of the individual finger.

4.2 Configuration of diverse underwater environments

The shallow-water environment was established in a tank measuring 60 cm × 30 cm × 35 cm, with water filled to a height of 30 cm to simulate standard shallow-water conditions. Flow disturbances were introduced by activating a water pump with a maximum head of 2 m H_2O, while water turbidity was simulated by adding 10 mL of black ink to the tank. These different shallow-water scenarios were designed to emulate environmental challenges that may be encountered by divers operating near the surface, such as surface wave-induced disturbances.

The underwater ROV (FIFISH V-EVO, Finware Technology, Guangdong, China) was deployed in a large tank measuring 2.4 m × 0.9 m × 0.9 m, filled to a depth of 0.85 m to replicate a deep-water operational environment.

4.3 Construction of the MTCN-UANet architecture

To facilitate a horizontal comparison of the performance of various backbone models in extracting FBG features, we trained multiple models using combined data collected from a single thumb sensor. The training results revealed significant differences in recognition accuracy across validation sets, as detailed in Supplementary Section 2. Based on a comprehensive assessment of recognition accuracy and model complexity, we selected the Temporal Convolutional Network (TCN) for extracting temporal features of the FBG wavelength signals. TCN is a convolution-based deep neural architecture that replaces recurrent structures with temporal convolutions, enabling efficient modeling of long-range temporal dependencies. It also offers improved parallelism and more stable gradient propagation. Compared to traditional recurrent architectures such as GRUs and LSTMs, TCN demonstrates superior expressiveness in time-series modeling due to its causality, dilated convolutions, and residual connections.

The core architecture of the TCN consists of multiple layers of dilated causal convolutions. Causal convolution ensures that the output at time step t depends only on the current and past inputs $x \leq t$, thereby preventing information leakage from future time steps and preserving the temporal integrity of the sequence. Dilated convolution, on the other hand, introduces spacing between kernel elements via a dilation factor, allowing the model to achieve a large receptive field without increasing the number of parameters, thereby capturing long-range temporal dependencies.

The convolution operation can be formulated as follows:

$$y(t) = \sum_{i=0}^{k-1} f(i) \cdot x(t - d \cdot i) \ , \tag{4}$$

where k denotes the kernel size, d is the dilation rate (e.g., 1, 2, 4, 8), and $f(i)$ represents the i-th convolution kernel weight. By progressively increasing the dilation rate across layers, the model is able to learn long-term temporal dependencies at deeper hierarchical levels.

To enhance training stability, the TCN architecture incorporates residual connection structures. Each residual block consists of two layers of dilated causal convolutions, each followed by a ReLU activation, normalization, and dropout. A skip connection is then applied, directly adding the block' s input to its output:

$$Output = ReLU(x + F(x)) \ , \tag{5}$$

where $F(x)$ represents the nonlinear transformation output of the residual block. This structure helps alleviate the vanishing gradient problem and enhances the trainability and generalization ability of deep networks.

In the actual model design, the time-series data of the Bragg wavelength from each FBG channel (original length of 4000) was segmented and reconstructed into a two-dimensional sequence input with a shape of (20,200), dividing the entire time series into 20 time segments with 200 feature points each. The corresponding input tensor dimensions are:

$$X \in \mathbb{R}^{B \times 20 \times 200} \ , \tag{6}$$

where B is the batch size. For each FBG channel, we construct an independent sub-network with three stacked residual convolution modules (nb_stacks=3), where each module contains four layers of dilated convolution with dilation rates of 1, 2, 4, and 8, respectively. The kernel size is 2, and the number of channels (nb_filters) is 64. The causal padding strategy ensures that the time-series output is aligned with the input, preventing any future information leakage.

With the above parameter configuration, the final network can achieve a receptive field of:

$$ReceptiveField = (k-1) \cdot \sum_{l=0}^{L-1} d_l + 1 \ . \tag{7}$$

For this model, with a kernel size of $k = 2$, dilation rates of [1, 2, 4, 8], and three residual stacks (a total of 12 convolutional layers), the overall receptive field can sufficiently cover the entire input sequence.

In the overall model, the five-channel TCN feature extraction modules operate in parallel, independently modeling the dynamic characteristics of each FBG channel. These extracted features are then fused to enable collaborative perception and discrimination of multi-source heterogeneous temporal information. The proposed MTCN-UANet architecture optimizes the processes of feature extraction, fusion, and classification in an end-to-end manner. By preserving the integrity of temporal dependencies, it significantly enhances the robustness and accuracy of multi-channel FBG-based gesture recognition. The detailed network architecture is illustrated in Fig. 3(b).

To evaluate the discrepancy between the model's predictions and the ground truth, the loss function used during training is Sparse Categorical Cross entropy (SCCE). The model is optimized using the Adam optimizer with an initial learning rate of 0.001. The batch size is set to 32, and the model is trained for 100 epochs.

References

1. Webb TJ, Berghe EV, O'Dor R. Biodiversity's big wet secret: the global distribution of marine biological records reveals chronic under-exploration of the deep pelagic ocean. *PLoS One* **5**, e10223 (2010).
2. Li ZJ, Yang CG, Ding N et al. Robust adaptive motion control for underwater remotely operated vehicles with velocity constraints. *Int J Control Autom Syst* **10**, 421–429 (2012).
3. Sivčev S, Coleman J, Omerdić E et al. Underwater manipulators: a review. *Ocean Eng* **163**, 431–450 (2018).
4. Frey ST, Haque ABMT, Tutika R et al. Octopus-inspired adhesive skins for intelligent and rapidly switchable underwater adhesion. *Sci Adv* **8**, eabq1905 (2022).
5. Rus D, Tolley MT. Design, fabrication and control of soft robots. *Nature* **521**, 467–475 (2015).
6. Dou WQ, Zhong GL, Cao JL et al. Soft robotic manipulators: designs, actuation, stiffness tuning, and sensing. *Adv Mater Technol* **6**, 2100018 (2021).
7. Shintake J, Cacucciolo V, Floreano D et al. Soft robotic grippers. *Adv Mater* **30**, 1707035 (2018).
8. El-Atab N, Mishra RB, Al-Modaf F et al. Soft actuators for soft robotic applications: a review. *Adv Intell Syst* **2**, 2000128 (2020).
9. Xie ZX, Yuan FY, Liu JQ et al. Octopus-inspired sensorized soft arm for environmental interaction. *Sci Robot* **8**, eadh7852 (2023).
10. Zhang NB, Ren JJ, Dong YS et al. Soft robotic hand with tactile palm-finger coordination. *Nat Commun* **16**, 2395 (2025).
11. Li GR, Wong TW, Shih B et al. Bioinspired soft robots for deep-sea exploration. *Nat Commun* **14**, 7097 (2023).
12. Li GR, Chen XP, Zhou FH et al. Self-powered soft robot in the Mariana Trench. *Nature* **591**, 66–71 (2021).

13. Baines R, Patiballa SK, Booth J et al. Multi-environment robotic transitions through adaptive morphogenesis. *Nature* **610**, 283–289 (2022).
14. Galloway KC, Becker KP, Phillips B et al. Soft robotic grippers for biological sampling on deep reefs. *Soft Robot* **3**, 23–33 (2016).
15. Zaidi S, Maselli M, Laschi C et al. Actuation technologies for soft robot grippers and manipulators: a review. *Curr Robot Rep* **2**, 355–369 (2021).
16. Chen XQ, Zhang X, Huang YY et al. A review of soft manipulator research, applications, and opportunities. *J Field Robot* **39**, 281–311 (2022).
17. Stuart H, Wang SQ, Khatib O et al. The ocean one hands: an adaptive design for robust marine manipulation. *Int J Robot Res* **36**, 150–166 (2017).
18. Chen P, Ding Q, Liu YW et al. Programmable pressure control in pneumatic soft robots with 2-way 2-state solenoid valves. *IEEE Robot Autom Lett* **9**, 6448–6455 (2024).
19. Puhlmann S, Harris J, Brock O. *RBO hand 3*: a platform for soft dexterous manipulation. *IEEE Trans Robot* **38**, 3434–3449 (2022).
20. Wu MX, Zheng XW, Liu RS et al. Glowing sucker octopus (*Stauroteuthis syrtensis*)-inspired soft robotic gripper for underwater self-adaptive grasping and sensing. *Adv Sci* **9**, 2104382 (2022).
21. Wu MX, Afridi WH, Wu JX et al. Octopus-inspired underwater soft robotic gripper with crawling and swimming capabilities. *Research* **7**, 0456 (2024).
22. Zhou S, Li YH, Wang QQ et al. Integrated actuation and sensing: toward intelligent soft robots. *Cyborg Bionic Syst* **5**, 0105 (2024).
23. Liu SY, Zhang DH, Fu X et al. Tactile sensing for soft robotic manipulators in 50 MPa hydrostatic pressure environments. *Adv Intell Syst* **5**, 2300296 (2023).
24. Zhao HC, O'Brien K, Li S et al. Optoelectronically innervated soft prosthetic hand via stretchable optical waveguides. *Sci Robot* **1**, eaai7529 (2016).
25. Zhou HY, Zhang C, Nong HC et al. Multi-photon neuron embedded bionic skin for high-precision complex texture and object reconstruction perception research. *Opto-Electronic Adv* **8**, 240152 (2025).
26. Galloway KC, Chen Y, Templeton E et al. Fiber Optic shape sensing for soft robotics. *Soft Robot* **6**, 671–684 (2019).
27. Sun ZJ, Wang SY, Zhao YL et al. Discriminating soft actuators' thermal stimuli and mechanical deformation by hydrogel sensors and machine learning. *Adv Intell Syst* **4**, 2200089 (2022).
28. Tavakoli M, Lopes P, Lourenco J et al. Autonomous selection of closing posture of a robotic hand through embodied soft matter capacitive sensors. *IEEE Sens J* **17**, 5669–5677 (2017).
29. Frishman S, Di J, Karachiwalla Z et al. A multi-axis FBG-based tactile sensor for gripping in space. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* 1794–1799 (IEEE, 2021). http://doi.org/10.1109/IROS51168.2021.9635998.
30. Teeple CB, Becker KP, Wood RJ. Soft curvature and contact force sensors for deep-sea grasping via soft optical waveguides. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* 1621–1627 (IEEE, 2018). http://doi.org/10.1109/IROS.2018.8594270.
31. Wang L, Wang YJ, Wang JF et al. A high spatial resolution FBG sensor array for measuring ocean temperature and depth. *Photonic Sens* **10**, 57–66 (2020).
32. Guo N, Han XD, Liu XB et al. Autoencoding a soft touch to learn grasping from on-land to underwater. *Adv Intell Syst* **6**, 2300382 (2024).
33. Zhang YJ, Jiang Y, Qi H et al. An Underwater human-robot interaction using a visual-textual model for autonomous underwater vehicles. *Sensors* **23**, 197 (2022).
34. Khushaba RN, Al-Timemy A, Kodagoda S et al. Combined influence of forearm orientation and muscular contraction on EMG pattern recognition. *Expert Syst Appl* **61**, 154–161 (2016).
35. Sundaram, Sahana BC. Multivariate EMG signal based automated hand gestures recognition framework for elder care. *Int J Precis Eng Manuf* **26**, 451–466 (2025).
36. Cao CH, Ma G, Chen ZL et al. A high-precision dynamic movement recognition algorithm using multimodal biological signals for human-machine interaction. *Adv Intell Syst* **7**, 2400483 (2025).
37. Bao TZ, Xie SQ, Yang PF et al. Toward robust, adaptiveand reliable upper-limb motion estimation using machine learning and deep learning-a survey in myoelectric control. *IEEE J Biomed Health Inform* **26**, 3822–3835 (2022).
38. Ge ZJ, Wu ZL, Han X et al. Gesture recognition and master-slave control of a manipulator based on sEMG and convolutional neural network-gated recurrent unit. *J Eng Sci Med Diagn Ther* **6**, 021004 (2023).
39. Connolly J, Condell J, O'Flynn B et al. IMU sensor-based electronic goniometric glove for clinical finger movement analysis. *IEEE Sens J* **18**, 1273–1281 (2017).
40. Zhao L, Wu B, Niu Y et al. Soft optoelectronic sensors with deep learning for gesture recognition. *Adv Mater Technol* **7**, 2101698 (2022).
41. Zhu RJ, Fan DL, Lin J et al. Machine-learning-assisted soft fiber optic glove system for sign language recognition. *IEEE Robot Autom Lett* **9**, 1540–1547 (2024).
42. Xu Z, Yang CJ, Wu WT et al. Design of underwater humanoid flexible manipulator motion control system based on data glove. In *2020 6th International Conference on Mechatronics and Robotics Engineering (ICMRE)* 120–124 (IEEE, 2020). http://doi.org/10.1109/ICMRE49073.2020.9065105.
43. Rosen J, Brand M, Fuchs MB et al. A myosignal-based powered exoskeleton system. *IEEE Trans Syst, Man, Cybern A* **31**, 210–222 (2001).
44. Nelson A, Singh G, Robucci R et al. Adaptive and personalized gesture recognition using textile capacitive sensor arrays. *IEEE Trans Multi-Scale Comp Syst* **1**, 62–75 (2015).
45. Zhang C, Zhou HY, Nong HC et al. FBG-SenseGlove: a digital twin data glove for underwater gesture recognition using deep learning. *IEEE Sens J* **25**, 8347–8354 (2025).
46. Feix T, Romero J, Schmiedmayer H B et al. The GRASP taxonomy of human grasp types. *IEEE Trans Human-Mach Syst* **46**, 66–77 (2016).
47. Xu Y, Huang XY, Chen ZK et al. High-sensitivity optical fiber seawater temperature and pressure sensor with a low-error demodulation method. *Opt Express* **33**, 1909–1931 (2025).

Acknowledgements

We are grateful for financial supports from National Natural Sciences Foundation of China grant No.62275269; the Key Research & Development Program of Guangxi grant GuiKeAB23075155; the Key Research and Development Program of Guangxi grant GuiKeAB24010203.

Author contributions

J.L., C.Z. (Chao Zhang), H.N. and Y.Y. conceived the idea, designed the experiments. J.L., C.Z. (Chao Zhang), H.N. and D.W. performed the experiments. J.L., C.Z. (Chao Zhang), D.W. and Y.D. performed data post-processing. J.L., C.Z. (Chao Zhang), H.N. wrote the original manuscript. D.W., H.Z., J.W., Q.B. and J.Z. revised original draft. C.Z(Chaofan Zhang), Z.Z. and Y.Y. discussed the manuscript. Y.Y., J.Y. and Z.Z. supervised the study.

Competing interests

The authors declare no competing financial interests.

Supplementary information

Supplementary information for this paper is available at https://doi.org/10.29026/oea.2026.250311

Published by Editorial Office of *Opto-Electronic Advance*, Institute of Optics and Electronics, Chinese Academy of Sciences.