Article

An Underwater Salvage Robot for Retrieving Foreign Objects in Nuclear Reactor Pools

Ming Zhong [1] **, Zihan Gao** [1]**, Zhengxiong Mao** [2]**, Ruifei Lyu** [2] **and Yaxin Liu** [1,*]

1 Robotics Institute, Harbin Institute of Technology, Weihai 264200, China; zhongming@hit.edu.cn (M.Z.);
  gzh1119125935@163.com (Z.G.)
2 State Nuclear Power Demonstration Plant Co., Ltd., Rongcheng 264300, China;
  maozhengxiong@spic.com.cn (Z.M.); lvruifei@spic.com.cn (R.L.)
* Correspondence: liuyaxin@hit.edu.cn

Abstract

In this paper, an underwater salvage robot is developed to retrieve foreign objects scattered in nuclear reactor pools. The robot mainly consists of an ROV platform and a 3-DOF Delta robotic arm. Utilizing fused IMU and LED beacon visual data for localization, it achieves pool traversal via six dynamically controlled thrusters. An improved YOLOv8s algorithm is employed to identify foreign objects in underwater environments. During traversal, the robot identifies and retrieves foreign objects along the way. The prototype of the robot was subjected to a series of experiments in an indoor pool. Results show that the improved YOLOv8 algorithm achieves 92.2% mAP, surpassing the original YOLOv8s and Faster-RCNN by 3.7 and 3.3 percentage points, respectively. The robot achieved a foreign-object identification rate of 95.42% and a retrieval success rate of 90.64% under dynamic traversal conditions, indicating that it meets the operational requirements and has significant engineering application value.

Keywords: ROV; nuclear applications; Delta arm; multi-sensor fusion localization; sliding mode control; YOLOv8

1. Introduction

The safe operation and maintenance of nuclear reactor pools is a key link to ensure the stable operation of nuclear power plants. Due to the large size of nuclear power pools, absolute reliability of precautions against foreign objects is hard to achieve, causing foreign objects to fall into nuclear reactor pools from time to time. These foreign objects are mainly bolts, nuts, graphite fragments, steel balls, and wires, which are accidentally dropped during equipment maintenance and fuel replacement. Since foreign objects may fall into the reactor bottom or narrow gaps, it is often required to lift out the whole reactor core, which will result in a delay in the critical path of overhaul and significant economic losses. Additionally, traditional manual salvage operations face significant issues such as high radiation exposure risks and low operational efficiency. Meanwhile, underwater robots have the advantages of long endurance, easy operation, high efficiency, and protection of personal safety, rendering them ideal for replacing manual operations.

Research institutions worldwide have developed several underwater robotic systems addressing nuclear pool environments, covering a wide range of functions such as inspection, welding, and foreign-object retrieval. Kim and his team developed a laser-guided underwater robot called LAROB for inspection of the inner wall of nuclear reactor

pressure vessels [1]; Cho and his team designed an underwater robot KeproVt for visual inspections of reactor vessel interiors [2,3]; Mazumdar and his team implemented a 4-DOF underwater robot employing jet propulsion for pressure vessel and nuclear piping system inspections [4]; Josip and his team developed an ROV system called TARGET to perform non-destructive inspection on bottom-mounted nozzles (BMNs) of reactor pressure vessels [5]; Leon and his team designed a lightweight underwater robotic system based on non-destructive testing sensors and actuators for inspecting circumferential weld joints along the inner wall of nuclear pressure vessels [6]; Luo and his team designed a novel underwater robot that can achieve precise welding of cracks on spent fuel pool walls [7]; Lee and his team developed an operational ROV with a 5-DOF robotic arm for search and retrieval of foreign objects at the bottom of nuclear reactor vessels and the cooling systems [8]; Iborra and his team developed a remotely operated robotic system named TRON for the retrieval of foreign objects inside reactor pools [9]; Zhang and his team proposed a modular cleaning ROV that can endure high-temperature and radioactive environments to remove radioactive sludge and foreign objects in spent fuel pools [10]; and Dong and his team developed a compact underwater robot equipped with cameras and a robotic arm to perform foreign-object retrieval within reactor pressure vessels and spent fuel pools [11].

However, there are still some limitations in the existing studies. Regarding underwater positioning, the dominant methods, including dead-reckoning navigation [12,13], underwater acoustic positioning [14,15], and beacon-based visual localization [16], exhibit inherent error accumulation when deployed individually [17]. In addition, the recognition accuracy of beacon-based visual localization in low-illumination environments needs to be improved. In terms of underwater target detection, conventional algorithms [18–20] suffer from computational inefficiency and inadequate robustness against environmental variability; deep vision techniques employing neural networks such as YOLO [21] and Faster R-CNN [22] demonstrate substantially degraded performance due to chromatic distortion and noise underwater. Consequently, these methods exhibit high miss rates when identifying fine metallic structures such as millimeter-scale wires.

To address the above challenges, this paper proposes a nuclear power pool foreign-object salvage robot based on an ROV platform and a Delta robotic arm, along with a corresponding operational workflow. Combined with fused localization technology, deep vision technology, and motion control technology, the system can achieve full-process autonomous operation from perception and localization to decision-making and execution, significantly enhancing the intelligence level and operational safety of foreign-object-retrieval tasks. The main contributions of this work are summarized as follows:

- A novel underwater salvage robot featuring a modular architecture with mobile, gripping, vision, and control subsystems. Radiation protection is applied to sensitive components to ensure reliable operation in the high-risk environment of nuclear pools.
- A set of LED beacons with distinct color patterns is designed to serve as an information source for underwater visual localization. Combined with IMU-based positioning, this approach improves the system's underwater localization accuracy.
- An enhanced YOLOv8s-based detection network incorporating an added high-resolution detection layer and a CBAM attention mechanism is proposed, which significantly improves the success rate of detecting small underwater targets.

The rest of the paper is organized as follows. Section 2 presents the operational environment, overall system architecture, and the workflow of the underwater salvage robot. Section 3 elaborates on the technical details and algorithms required for the system to achieve the reactor-pool traversal function. Section 4 explains the realization of the foreign-object identification, localization, and grasping function. Section 5 shows the experimental design and results. Section 6 provides conclusions and suggestions for future research.

2. Operational Environment and Overall Design Scheme

2.1. Operational Environment

The main operational zones of the underwater robot comprise a rectangular refueling pool, which measures 10.71 × 8 m, with a centrally positioned reactor pressure vessel (RPV), 4.3 m in diameter and 12.8 m in depth, as shown in Figure 1. During refueling operations, spent fuel will be removed from the RPV, rendering the vessel interior free of internal obstructions. Foreign objects may disperse across the refueling pool floor or accumulate centrally within the hemispherical bottom of the pressure vessel. The common types of foreign objects include graphite fragments and metallic components such as bolts, nuts, steel balls, and metallic wires, with weights primarily ranging from several grams to tens of grams.
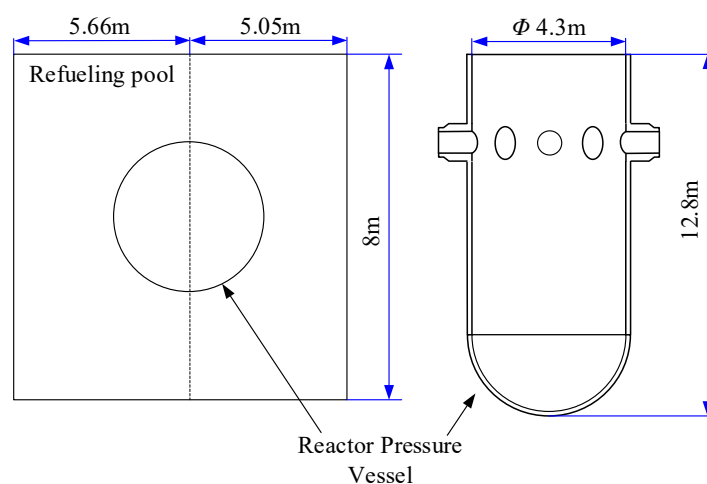

Figure 1. Working environment.

During nuclear power plant operations, personnel access to reactor pressure vessels is strictly prohibited due to radiation hazards. Consequently, foreign-material retrieval operations are exclusively conducted during the refueling process. During this period, the refueling pool will be filled with borated water for radiation shielding, and the radiation dose rate will be reduced to 15 Gy/h. The water may become turbid due to the introduction of suspended particulate matter such as dust, while simultaneously exhibiting a pronounced blue tint as a result of Cherenkov radiation. Both these phenomena adversely affect the performance of visual detection systems.

2.2. Overall Design Scheme of the Underwater Salvage Robot

According to the operational environment, the underwater salvage robot (USR) introduced in this paper must be able to perform comprehensive functions, including autonomous traversal of refueling pools and the bottom of RPVs, foreign-object recognition, and retrieval operations. The USR employs a modular architecture with functional segmentation into four subsystems: mobility module, grasping module, vision module, and control module, as shown in Figure 2. The mobility module provides enough flexibility of movement for free navigation within nuclear reactor pools; the grasping module executes retrieval operations for all detected foreign objects within its operational envelope; the vision module delivers image data for dual localization capabilities, including USR self-positioning and underwater target-positioning; the control module regulates operational states across the rest of the modules.
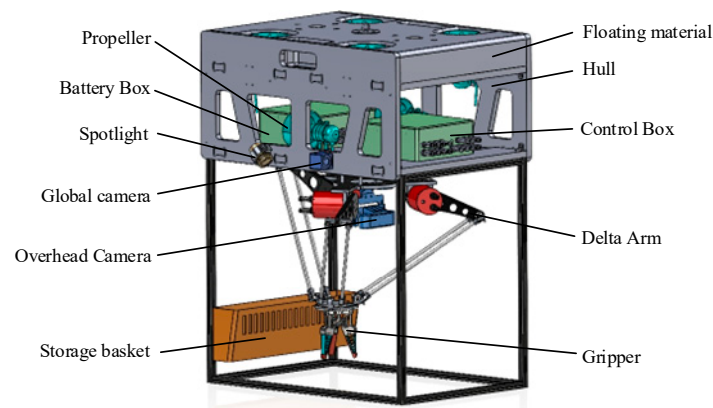
Figure 2. Three-dimensional model of the USR.

The mobile module has a rectangular form with dimensions of 800 mm in length, 590 mm in width, and 380 mm in height. It consists of a hull structure, six propeller thrusters, and the floating material. The hull is made of 20-mm-thick steel plates incorporating pre-installed mounting brackets for the motors and the floating material. Among the six propeller thrusters, two are horizontally oriented to provide propulsion for forward/reverse motion and yaw maneuvering, while the remaining four feature vertical orientation to generate heave-direction thrust for depth control. The floating material, positioned on the upper part of the hull, provides sufficient buoyancy support for the USR to ensure that it can float to the surface of water naturally even if the power supply is exhausted.

The gripping module consists of a 3-DOF Delta arm with a gripper as the end-effector. The Delta mechanism integrates a moving platform, a fixed base, actuated arms, and passive linkage arms. The passive arms are made of carbon fiber composites to reduce inertial mass while enhancing end-effector flexibility. The joint motors are wrapped by waterproof housings and sealed via tight compression of O-rings between the housings and actuated arm interfaces. The gripper is rigidly mounted to the moving platform.

The vision module consists of global cameras, an overhead camera, and underwater spotlights. The global cameras, oriented forward, are mounted on both lateral sides of the USR hull, providing visual localization information for the USR by detecting beacons placed in the environment; the overhead camera, facing downward, is fixed to the fixed base of the Delta arm, operating in an "eye-to-hand" configuration relative to the robotic arm. This arrangement delivers precise 3D perception and depth information for target-retrieval operations. The underwater searchlights are installed behind each global camera, which can flexibly adjust the orientation and brightness to ensure that the two sets of cameras can provide high-quality image output.

The control module consists of a main control box fixed in front of the hull and a battery box fixed at the rear of the hull. The main control box encapsulates a host computer, propulsion control unit, inertial measurement unit (IMU), and other sensors, and is connected to the peripherals of the USR through multiple waterproof cables to realize comprehensive system-wide control. The battery box contains lithium-ion power cells that deliver regulated voltage to the main control box and other equipment through waterproof cables to meet their power requirements.

To enable the USR to operate over extended periods in a radiation-filled aquatic environment, specific radiation-hardening measures have been implemented. The radiation-sensitive components primarily include various electronic elements within the control box. According to the relevant literature [23–25], the network switch exhibits the lowest radiation tolerance, which may fail upon exposure to a radiation dose of approximately

200 Gy. Additionally, waterproof seals made of polymer materials may undergo fracture or cross-linking failure when exposed to radiation beyond a certain dose.

For the control box, a 20-mm-thick shielding layer made of lead alloy was used to enclose the electronic components. Using the half-value layer method, the formula for the maximum radiation dose I_0 that the system can withstand is given by the following:

$$I_0 = 2^{\frac{d}{\Delta(1/2)}} I \tag{1}$$

where d represents the thickness of the shielding layer; $\Delta(1/2)$ denotes the half-weakening thickness of the shielding material; and I refers to the radiation dose after shielding. All doses are expressed in Gy. When setting d = 20 mm, I = 200 Gy, and $\Delta(1/2)$ = 12.4 mm, the resulting value of I_0 is 611.73 Gy.

For the waterproof seals, the radiation hardening strategy is to use materials with superior radiation resistance. By comparing commonly used seal materials such as natural rubber, silicone rubber, fluororubber, and nitrile rubber [26–28], we ultimately selected nitrile rubber as the base material, which can withstand radiation doses of up to 4 MGy.

In summary, the USR can tolerate a total radiation dose of 611.73 Gy. Assuming a single operational duration of 2 h, corresponding to an exposure of 30 Gy per mission, the system can perform approximately 20 missions.

The autonomous operational workflow of the USR is shown in Figure 3. To begin with, the USR is deployed at the designated location, whereupon multi-sensor fusion localization and motion control algorithms guide the USR to traverse the refueling pool floor and the bottom of the RPV along predetermined paths. Concurrently, foreign-object identification is performed within the traversed regions. Once the target is detected, precise spatial coordinates will be computed to navigate the USR to an optimal manipulation pose for retrieval operations. After the traversal task is finished, the USR returns to the predetermined path for continued searching until completing the area coverage task, subsequently returning to the designated deployment location, thereby concluding the operational cycle.
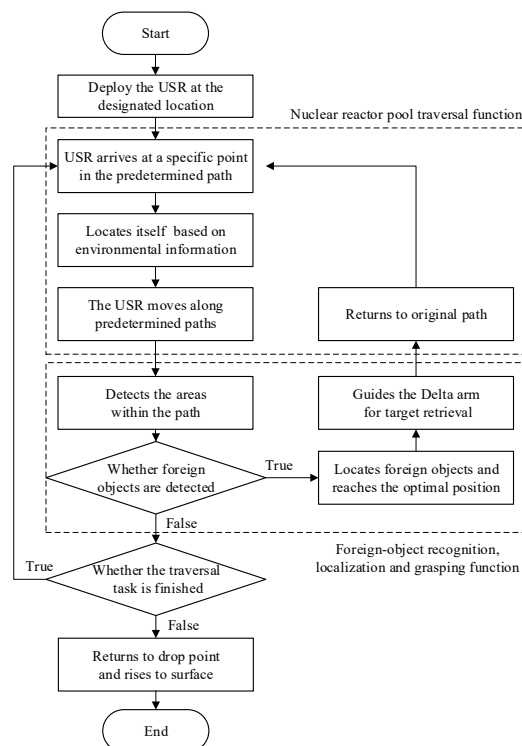

Figure 3. The operational process diagram of the USR.

3. Nuclear Reactor Pool Traversal Function

Comprehensive and precise traversal of the nuclear reactor pool is the basis for the USR to reach a high foreign-object-retrieval rate in a single operation. In this section, the strategies for the USR to realize the accurate traversal function are discussed in detail, focusing on the traversal path planning strategy, underwater localization strategy, and motion control strategy.

3.1. Traversal-Path Planning Strategy

In order to ensure comprehensive traversal and avoid cable entanglement problems, a zigzag-coverage traversal strategy is adopted for the refueling pool, as shown in Figure 4a, while an outward-to-inward circular coverage traversal strategy is implemented for the bottom of the RPV, as shown in Figure 4b. The traversal strategy is based on the zigzag-coverage strategy.
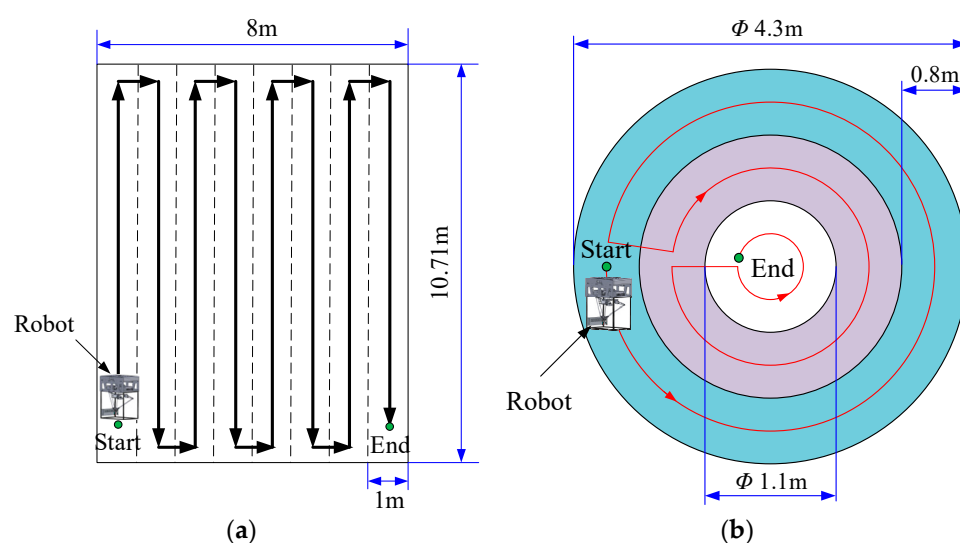

Figure 4. Path planning maps: (**a**) refueling pool; (**b**) bottom of the RPV.

3.2. Underwater Localization Strategy

In this study, artificial beacons are introduced to assist localization on the basis of altimeter, depth gauge, and IMU measurements. The USR acquires environmental images through global cameras and extracts the position, angle, and other geometric features of the beacons according to a specialized beacon recognition algorithm, so as to decode the robot's current coordinates and orientation within the world coordinate system.

To enhance the discernibility of beacons as well as localization accuracy, we adopt the beacon design scheme based on LED lights. The design of the beacon board is shown in Figure 5. The LED lights on the beacon board have a decentralized layout with as many lights as possible, which avoids interference between the lights and makes it relatively easy and accurate to calculate the matching from 2D to 3D points. The beacons are distinguished from each other by an arrangement of LED lights in different colors, as shown in Figure 6. Considering that light with longer wavelengths generally attenuates more rapidly in water, blue and green LEDs were selected to construct the beacons. Arranging the lights in the positions marked from 1 to 7 on the beacon board and sequencing these positions results in a seven-digit binary code, with the blue color representing the "0" in the binary code while the green color representing the "1"; thus, 128 different beacon position codes can be represented. In practical applications, the absolute position of each beacon board in the global reference frame can be determined through decoding.
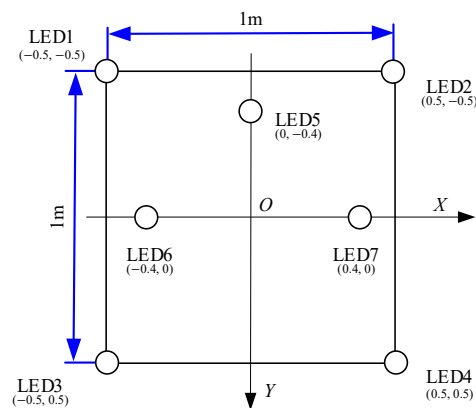
Figure 5. The design of LED beacons.
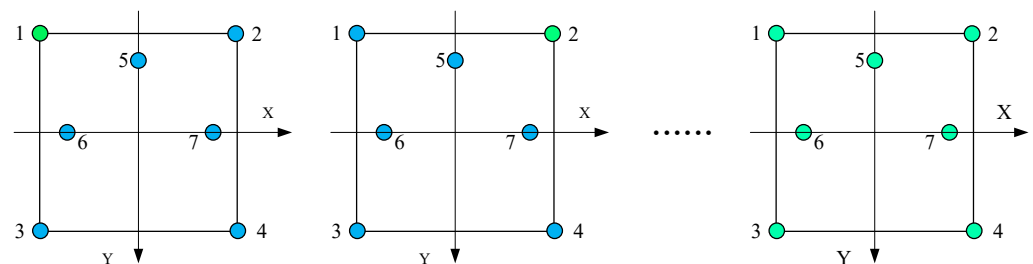
**Figure 6.** Arrangement of LEDs in different colors.

The beacon deployment scheme is illustrated in Figure 7, where the beacons are represented by blue rectangles. For the refueling pool, the beacons are installed at the end of the linear paths, with the global camera maintaining horizontal orientation to detect beacons for robotic localization. For the bottom of the RPV, the beacons are suspended above the reactor pressure vessel, and the angle of the global cameras is adjusted to detect the beacons upward during traversal.

(**a**)   (**b**)
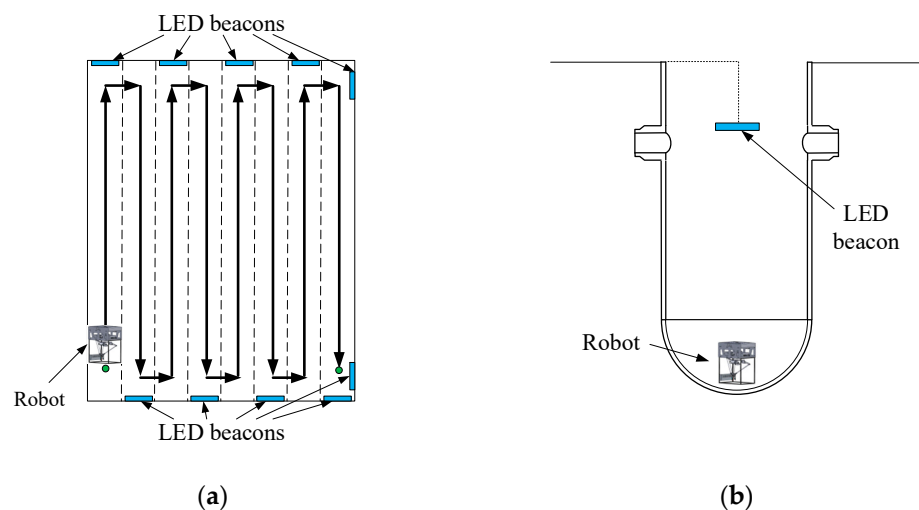
Figure 7. The deployment scheme of LED beacons: (**a**) in the refueling pool; (**b**) at the bottom of the RPV.

The specific process of visual localization is illustrated in Figure 8. First the global camera captures images in the field of view, then the original RGB images are converted into the HSI color space via the OpenCV library, and the Contrast-Limited Adaptive Histogram Equalization (CLAHE) algorithm is employed to enhance the local contrast of the intensity channel [29]. This operation decouples luminance information from color information,

thereby enhancing contrast without altering the intrinsic hue of the image, which improves the discernibility of beacons in water with similar color properties. Following this step, the Canny edge detection algorithm is applied to extract the contours of the LED lights from the enhanced image, which is a widely used technique in various image processing tasks, demonstrating superior noise immunity compared to gradient-based methods [30] and other second-order derivative approaches [31]. After extracting the LED contours, the Hough transform algorithm [32] is used to fit them to circles and calculate the pixel coordinates of their centers. These 2D center coordinate points are then matched to their corresponding 3D spatial coordinates.
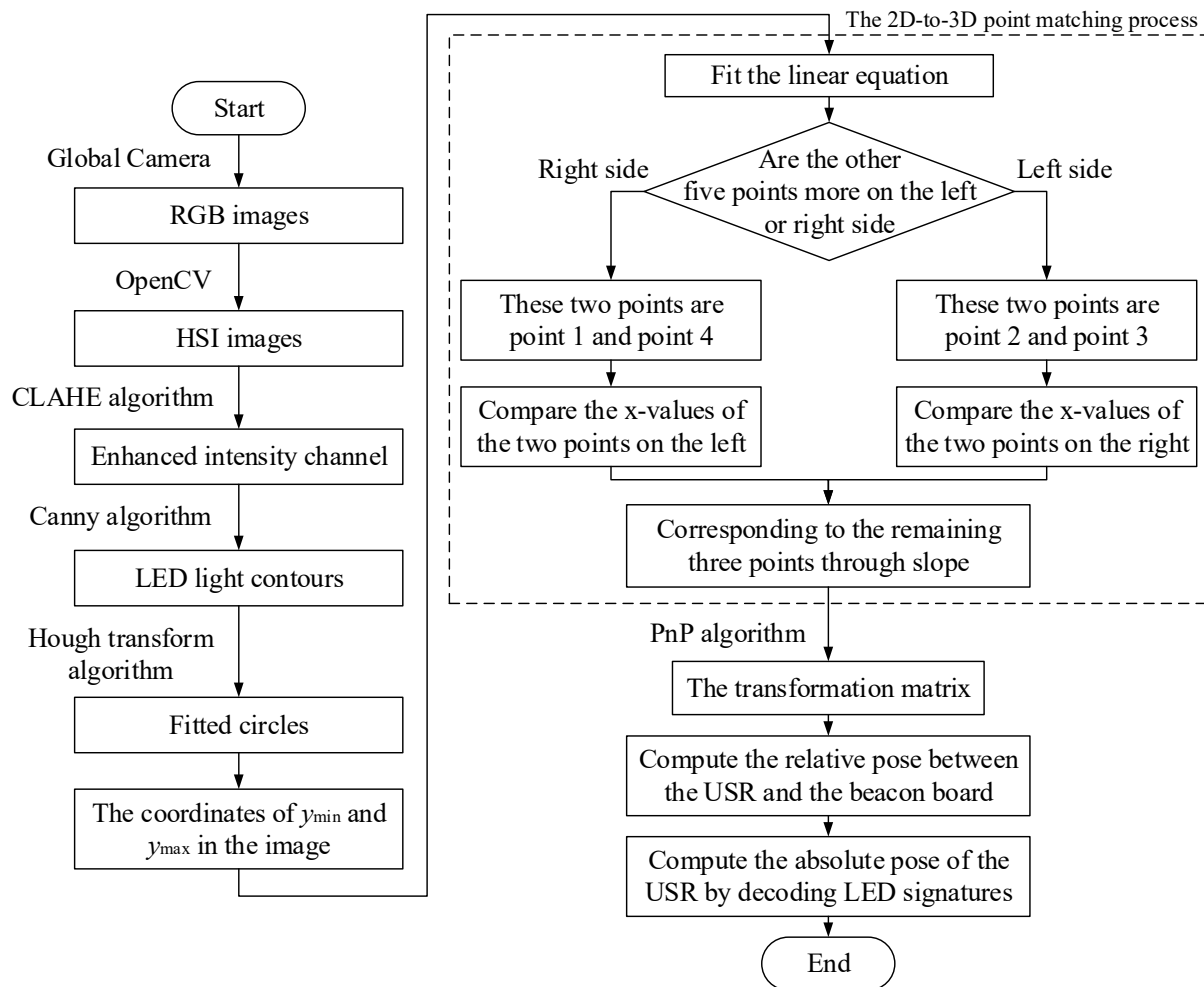
Figure 8. The process of visual localization.

After matching the corresponding points, we can use the perspective-n-point (PnP) algorithm to solve the transformation matrix. Since the points to be solved are located in the same plane, the PnP problem degenerates into solving a homography matrix. The homography matrix H is solved through the equation outlined below:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} \tag{2}$$

Since H is also a homogeneous matrix, it possesses only eight degrees of freedom. To simplify computation, H is normalized by scaling with $1/h_{33}$ to constrain the last

element. Expanding and rearranging Equation (2) through matrix multiplication yields the following:

$$
\begin{bmatrix}
-x_{w1} & -y_{w1} & -1 & 0 & 0 & 0 & ux_{w1} & uy_{w1} & u \\
0 & 0 & 0 & -x_{w1} & -y_{w1} & -1 & vx_{w1} & vy_{w1} & v \\
-x_{w2} & -y_{w2} & -1 & 0 & 0 & 0 & ux_{w2} & uy_{w2} & u \\
0 & 0 & 0 & -x_{w2} & -y_{w2} & -1 & vx_{w2} & vy_{w2} & v \\
-x_{w3} & -y_{w3} & -1 & 0 & 0 & 0 & ux_{w3} & uy_{w3} & u \\
0 & 0 & 0 & -x_{w3} & -y_{w3} & -1 & vx_{w3} & vy_{w3} & v \\
-x_{w4} & -y_{w4} & -1 & 0 & 0 & 0 & ux_{w4} & uy_{w4} & u \\
0 & 0 & 0 & -x_{w4} & -y_{w4} & -1 & vx_{w4} & vy_{w4} & v
\end{bmatrix}
\begin{bmatrix}
h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32}
\end{bmatrix} = 0
\tag{3}
$$

Apply singular value decomposition (SVD) [33] to Equation (3), and the homography matrix H can be solved. Then, compute the relative pose between the USR and the corresponding beacon board on the basis of H. By integrating this relative pose with the absolute position of the beacon board, which is obtained through decoding the LED signatures, the absolute pose of the USR within the world coordinate system is determined.

During the initial stage of USR motion, visual localization accuracy remains low due to extended distances from beacon boards, whereas IMU positioning provides higher precision. As the USR approaches beacon boards, IMU errors escalate due to cumulative effect and zero-drift phenomena, while visual localization precision improves as the target distance decreases. To maintain consistently high positioning accuracy throughout the traversal task, adaptive weighting between IMU positioning and visual localization must be implemented: IMU data dominate when the beacons are far away, while visual weighting progressively increases during the approaching phase. This adaptive sensor fusion can be realized via a Kalman filter [34], which continuously optimizes the fusion weights between IMU and vision localization inputs based on real-time uncertainty estimation.

During the prediction phase of the Kalman filter, the state transition equation of the USR's pose is estimated based on the information provided by the IMU, as formulated in Equation (4). The prediction yields a rough estimate that can be expressed through the covariance matrix in Equation (5):

$$
x_k = Fx_{k-1} + Bu_{k-1}
\tag{4}
$$

$$
P_k = FP_{k-1}F^T + Q
\tag{5}
$$

where x_{k-1} denotes the state at time step $k-1$; F refers to the state transition matrix; u_{k-1} represents the control input; B refers to the control input matrix; P_k corresponds to the prior error covariance matrix at time step k; and Q denotes the process noise covariance matrix.

During the Kalman filter's update phase, the prior state estimate is corrected using actual measurement data acquired from the vision sensor. The observation equation is formulated as follows:

$$
z_k = H_k x_k + v_k
\tag{6}
$$

where z_k denotes the results of visual localization; H_k represents the observation matrix; and v_k corresponds to stochastic observation noise.

To achieve dynamic weighting adaptation, the Kalman gain K_k governs this correction process by determining the relative confidence weights between predicted states (derived from IMU positioning) and actual observations (derived from vision localization). A higher Kalman gain indicates greater reliance on visual measurements, whereas a lower gain prioritizes the IMU data. The Kalman gain is computed as follows:

$$K_k = P_k H_k^T (H_k P_k H_k^T + R)^{-1} \tag{7}$$

The Kalman filter subsequently updates the USR's state estimate and error covariance using the computed Kalman gain, as formalized in Equations (8) and (9). Through this recursive refinement of position and attitude estimates, the system delivers continuously updated and precise localization solutions in real time.

$$K_k = P_k H_k^T (H_k P_k H_k^T + R)^{-1} \tag{8}$$

$$K_k = P_k H_k^T (H_k P_k H_k^T + R)^{-1} \tag{9}$$

3.3. Motion Control Strategy

Analyzing the motion of the USR should begin with establishing its reference coordinate systems, including the world coordinate system $E\text{-}\xi\eta\zeta$ and the body-fixed coordinate system $O\text{-}xyz$, as illustrated in Figure 9. The world coordinate system $E\text{-}\xi\eta\zeta$ is fixed to the Earth, while the body-fixed coordinate system $O\text{-}xyz$ is attached to the USR body. Both coordinate systems follow the right-hand rule.
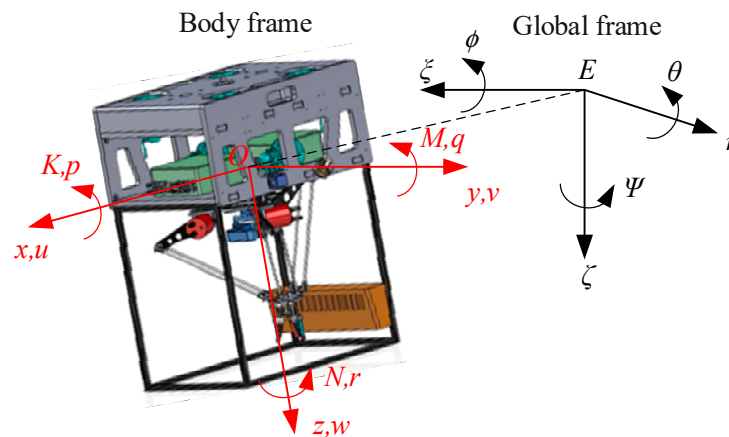
Figure 9. Schematic diagram of the USR coordinate system.

Since the forces and moments acting on the USR are measured and described in the body-fixed coordinate system, while global positioning and trajectory planning are expressed based on the world coordinate system, it is essential to establish the transformation relationship between these two coordinate systems. The position and attitude vector of the USR in the world coordinate system is denoted as $\eta = [\xi, \eta, \zeta, \phi, \theta, \psi]^T$. Differentiating η with respect to time yields the linear and angular velocity vector $\dot{\eta} = [\dot{\xi}, \dot{\eta}, \dot{\zeta}, \dot{\phi}, \dot{\theta}, \dot{\psi}]^T$ of the USR in the world frame. Correspondingly, the velocity and angular velocity vector expressed in the body-fixed coordinate system is denoted as $v = [u, v, w, p, q, r]^T$. The transformation between $\dot{\eta}$ and v is given by the following [35–38]:

$$\dot{\eta} = J(\eta)v \tag{10}$$

The above equation represents the kinematic equation of the USR, where $J(\eta)$ is the Jacobian matrix that maps the velocity in the body-fixed frame to the rate of change in position and attitude in the world coordinate system. $J(\eta)$ can be further expanded as follows:

$$J(\eta) = \begin{bmatrix} J_1(\eta) & O_{3\times3} \\ O_{3\times3} & J_2(\eta) \end{bmatrix} \tag{11}$$

The specific forms of $J_1(\eta)$ and $J_2(\eta)$ are given as follows [35–37]:

$$J_1(\eta) = \begin{bmatrix} \cos\theta\cos\psi & -\cos\phi\sin\psi + \sin\phi\sin\theta\cos\psi & \sin\phi\sin\psi + \cos\varphi\sin\theta\cos\psi \\ \cos\theta\sin\psi & \cos\phi\cos\psi + \sin\phi\sin\theta\sin\psi & -\sin\phi\cos\psi + \cos\varphi\sin\theta\sin\psi \\ -\sin\theta & \sin\phi\cos\theta & \cos\phi\cos\theta \end{bmatrix} \tag{12}$$

$$J_2(\eta) = \begin{bmatrix} 1 & \sin\phi\tan\theta & \cos\phi\tan\theta \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi\sec\theta & \cos\phi\sec\theta \end{bmatrix} \tag{13}$$

After establishing the kinematic equation, the motion trajectory and attitude of the USR can be predicted during operation, thereby providing essential information for the control and navigation of the USR.

To achieve precise motion control of the USR, developing a dynamic model of the system is also required. The following simplifications are adopted in our research [35]: (1) the origin O of the body-fixed coordinate system coincides with the center of gravity of the USR, and the principal axes of inertia are aligned with the three axes of the body frame; (2) the center of buoyancy is located on the Oz-axis of the body-fixed frame during the design phase; (3) by appropriately selecting the dimensions of the buoyancy material during the design phase, the gravity of the USR is set equal to its buoyancy force; (4) the Coriolis and centrifugal forces acting on the USR can be neglected due to its relatively low velocity when traversing the pool. Based on the following simplifications, the mathematical formulations of applied forces are expressed in the following form [38]:

$$\begin{cases} M\dot{v} + D(v)v + g(\eta) = \tau_E + \tau \\ \dot{\eta} = J(\eta)v \end{cases} \tag{14}$$

where v represents the velocity matrix; M refers to the mass matrix; $D(v)$ stands for the damping matrix; $g(\eta)$ represents restoring forces and moments; τ denotes the propulsion force matrix; and τ_E represents cable-induced disturbances.

Due to the strong nonlinearity of the USR's dynamics model, appropriate control algorithms need to be adopted to achieve accurate and stable motion control. Current techniques include PID control, sliding mode control (SMC), adaptive control, and neural network control [39]. Compared to other methods, the variable structure characteristics of SMC endow it with inherent insensitivity to internal parameter variations and external disturbances [40]. This enables SMC to deliver robust performance, even when facing the highly coupled and nonlinear dynamic system of underwater robots [41]. Consequently, this paper employs SMC for the USR's motion attitude control.

In this paper, a first-order sliding surface $s = ce + \dot{e}$ is adopted, and the exponential reaching law $\dot{s} = c\dot{e} + \ddot{e} = -\varepsilon\mathrm{sgn}(s) - qs, \varepsilon > 0, q > 0$ [42] is selected as the reaching law for the sliding mode controller. In practical applications of sliding mode controllers, external disturbances may cause the system state to repeatedly cross the sliding surface, causing rapid switching of the signum function sgn(s), thereby inducing a chattering phenomenon, which may damage actuators. To address this issue, the saturation function sat(s) is employed to substitute sgn(s) [43]. The saturation function can effectively suppress fluctuations in the controller output, so as to reduce chattering while enhancing system robustness and stability simultaneously. The saturation function sat(s) is defined as follows:

$$sat(s) = \begin{cases} 1 & s > \Delta \\ hs & |s| \leq \Delta \\ -1 & s < -\Delta \end{cases} \quad h = \frac{1}{\Delta} > 0 \tag{15}$$

where h represents the slope of the function when $|s| \leq \Delta$; and Δ is the boundary layer thickness.

Based on Equation (14), the sliding mode controller for USR is designed as outlined below:

$$u = MJ^{-1}(\eta)\left[\ddot{x}_{1d} - c(x_2 - \dot{x}_{1d}) - f(x) - \tau_E - \varepsilon\mathrm{sat}(s) - qs\right], \varepsilon > 0, q > 0 \qquad (16)$$

where u is the control input; x_1 and x_2 are the state vectors of the system; x_{1d} represents the desired pose vector; and $f(x)$ is a nonlinear function of the state variables. Their explicit expressions are given as follows:

$$\begin{cases} u = \begin{bmatrix} \tau_X & \tau_Y & \tau_Z & \tau_K & \tau_M & \tau_N \end{bmatrix}^T \\ x_1 = \begin{bmatrix} \xi & \eta & \zeta & \phi & \theta & \psi \end{bmatrix}^T \\ x_2 = \begin{bmatrix} \dot{\xi} & \dot{\eta} & \dot{\zeta} & \dot{\phi} & \dot{\theta} & \dot{\psi} \end{bmatrix}^T \\ f(x) = \dot{J}(\eta)v - J(\eta)M^{-1}(D(v)v + g(\eta)) \end{cases} \qquad (17)$$

Define the Lyapunov function $V = \frac{1}{2}s^2$. Its derivative is obtained as outlined below:

$$\dot{V} = s\dot{s} = s(-\varepsilon\mathrm{sat}(s) - qs), \varepsilon > 0, q > 0 \qquad (18)$$

When $|s| > \Delta$, the saturation function reduces to the sign function, yielding the following:

$$\dot{V} = s(-\varepsilon \cdot \mathrm{sat}(s) - qs) = -\varepsilon|s| - qs^2 < 0, \varepsilon > 0, q > 0 \qquad (19)$$

When $|s| \leq \Delta$, the saturation function becomes linear, leading to the following:

$$\dot{V} = s(-\varepsilon \cdot hs - qs) = -(\varepsilon h + q)s^2 \leq 0, \varepsilon > 0, q > 0, h > 0 \qquad (20)$$

When $\dot{V} \equiv 0$, $s \equiv 0$. According to LaSalle's invariance principle, the system is asymptotically stable.

The parameters involved in the controller include the adjustable gain c of the first-order sliding surface, the constant velocity term gain q, the exponential term gain ε, and Δ. In accordance with empirical values from existing studies [44], this paper selects $c = 1.5$, $q = 5$, $\varepsilon = 0.5$, and $\Delta = 0.01$.

The control flowchart of the USR is shown in Figure 10. The USR first acquires its position η in the world coordinate frame via the underwater localization function and its velocity v in the body coordinate frame via the velocity sensor. The velocity v is then transformed into the world coordinate frame as $J(\eta)v$ using the kinematic model. The variables η and $J(\eta)v$ are then subtracted from the desired position x_{1d} and the desired velocity $\dot{x}_{1d}$, respectively, to yield the system state vectors x_1 and x_2. Subsequently, the sliding mode controller computes the required control force τ for the system based on x_1, x_2, and $f(x)$, which is derived from the dynamic model. The outputs are allocated to six thrusters via a distribution strategy based on mechanical configuration and robot states, and individual thrust values F_1 to F_6 are calculated. Ultimately, under the thrust generated by the propellers, the USR moves towards the desired position.
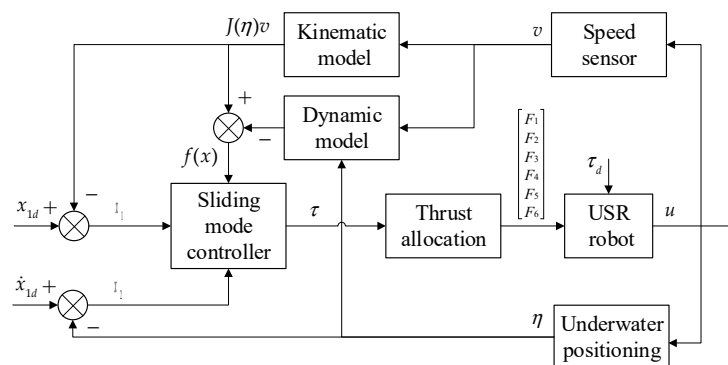
Figure 10. Attitude control flow chart of the USR.

4. Foreign-Object Identification, Localization, and Grasping Function

During autonomous traversal operations in nuclear reactor pools, the USR should retrieve foreign objects detected along its search trajectory. This section details the strategies of foreign-object retrieval, including identification, localization, and grasping procedures. Specifically designed object identification and localization algorithms are applied to identify all debris within the field of view accurately and transform their pixel-coordinate 3D positions into the Delta arm's coordinate frame to provide precise spatial references for grasping. The kinematics of the Delta arm is analyzed, and optimized grasping strategies are developed to enable dexterous and precise object retrieval.

### 4.1. Foreign-Object Identification Algorithm

Traditional object detection methods exhibit significant limitations when facing complex environments, particularly underwater scenarios [45]. In contrast, approaches based on a convolutional neural network (CNN)—especially the YOLO algorithm series—demonstrate superior real-time performance and computational efficiency. This paper adopts the YOLOv8 algorithm [46] for underwater foreign-object identification. Specifically, the YOLOv8 was chosen for its specific model due to its minimal parameter count, faster inference speeds, and lower computational demands while maintaining competitive detection accuracy on many common datasets [47].

Although YOLOv8 demonstrates robust performance in detecting common underwater components such as bolts and nuts, its miss rate increases when identifying tiny fragments and metallic wires. To solve this problem, two structural modifications were adopted to the YOLOv8 detection network. The optimized network architecture is illustrated in Figure 11.

Firstly, the original YOLOv8 model contains three detection layers. When processing 640×640 input images, the output feature maps of these detection layers have dimensions of 80×80, 40×40, and 20×20, designed for detecting small, medium, and large targets, respectively. Due to the high downsampling ratio in YOLOv8s, it is hard for deeper feature maps to capture enough feature information of small targets, resulting in poorer performance in small object detection. To address this, we incorporate a shallower feature map for those smaller foreign objects in our dataset, whose output feature maps have a dimension of 160×160, enabling the network to focus more effectively on small targets and significantly enhance detection accuracy.

Secondly, to overcome the limitations of traditional convolutional neural networks in processing information with varying sizes, shapes, and orientations, the convolutional block attention module (CBAM) is introduced [48]. The CBAM is an attention mechanism that enhances the performance of convolutional neural networks by focusing on critical features. It integrates both channel and spatial attention modules, both of which generate

feature descriptors via global pooling operations, respectively. These descriptors are then processed by a shared multilayer perceptron (MLP) with identical parameters to produce channel and spatial attention maps. The outputs of both modules undergo element-wise multiplication to yield the final attention-enhanced features. These features serve as input to subsequent network layers, preserving critical information while suppressing noise and irrelevant details.
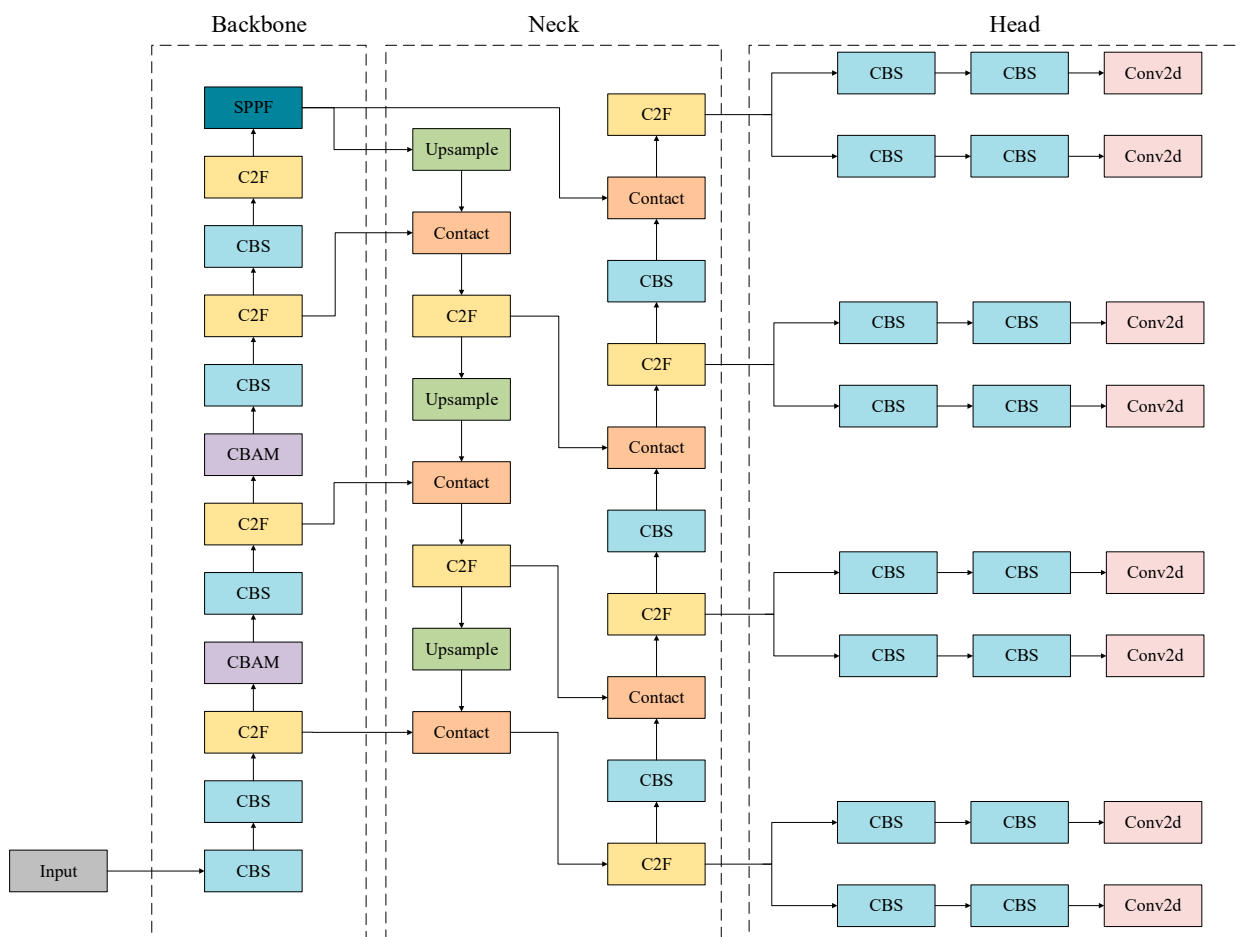
Figure 11. Improved YOLOv8 network.

Based on the practical conditions of the water in nuclear reactor pools, we constructed separate datasets for clear and turbid water conditions, both of which were augmented by introducing an artificial blue color cast to simulate the presence of Cherenkov radiation. Subsequently, the model was trained on these augmented datasets to enhance its robustness against diverse underwater visual conditions.

4.2. Foreign-Object Localization Algorithm

After foreign objects are detected using the foreign-object recognition algorithm, it is necessary to locate these targets, that is, to obtain the accurate coordinates of the targets in 3D space, so that the Delta arm can perform subsequent grasping operations. We adopt a binocular camera to acquire both image and depth information of the target. The three-dimensional coordinates of the target in the world coordinate system are calculated using the pinhole imaging model, the underwater imaging model, and the binocular ranging model.

The pinhole imaging model is responsible for describing the linear projection relationship between the target point and its corresponding pixel point. A schematic diagram of the pinhole imaging model is shown in Figure 12.
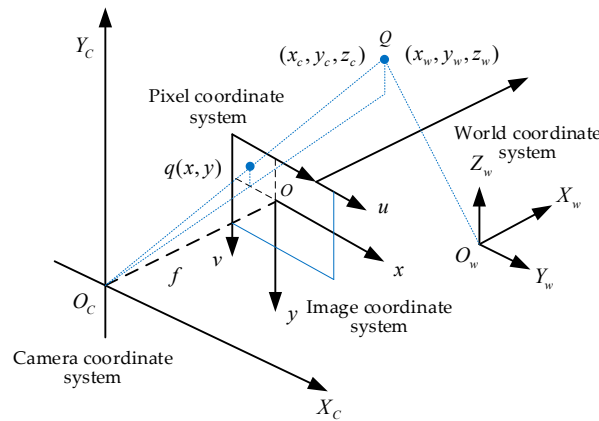
Figure 12. Schematic diagram of the pinhole imaging model.

Assume that the spatial coordinates of the target point Q in the world coordinate system O_W-$X_W Y_W Z_W$ are (X_W, Y_W, Z_W), and its projection point q in the visual image captured by the binocular camera has pixel coordinates (x,y) in the pixel coordinate system O-xyz. To obtain the 3D coordinates of the target point from the visual image, the following coordinate transformations must be performed sequentially: from the pixel coordinate system to the image coordinate system, then to the camera coordinate system, and finally to the world coordinate system. This process establishes the mapping between the world coordinate system and the pixel coordinate system as follows:

$$Z_C \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & x_0 & 0 \\ 0 & f_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Z_W \\ Y_W \\ 1 \end{bmatrix} = KM \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \tag{21}$$

where Z_C denotes the z-axis coordinate of point Q in the camera coordinate system; f_x and f_y represent the focal lengths of the image in the horizontal and vertical directions, respectively; x_0 and y_0 are the offsets of the image coordinate system's origin relative to the pixel coordinate system's origin; R is the rotation matrix between the camera and world coordinate systems; T is the translation vector between these two systems; K is the camera intrinsic matrix; and M denotes the camera extrinsic matrix.

In this study, the camera is housed within a waterproof enclosure. Light passes sequentially through water, glass, air, and the lens, resulting in refraction that displaces the actual imaging point. To correct this deviation, the underwater imaging model is introduced. The schematic diagram of the model is shown in Figure 13, where Z_C denotes the distance from the point in the world coordinate system to the camera; q_a represents the corresponding image point formed by light traveling through air and incident on the CCD; q_w denotes the image point formed when light passes through water, glass, and air sequentially before reaching the CCD; and δ is the distance between the camera's focal point and the waterproof glass.

The model assumes that the waterproof glass is homogeneous and sufficiently thin, and so the refraction at both interfaces can be simplified as an equivalent single refraction occurring at the transition from water to air, thereby simplifying the optical path analysis. Based on the light propagation path, the following relationship can be established:
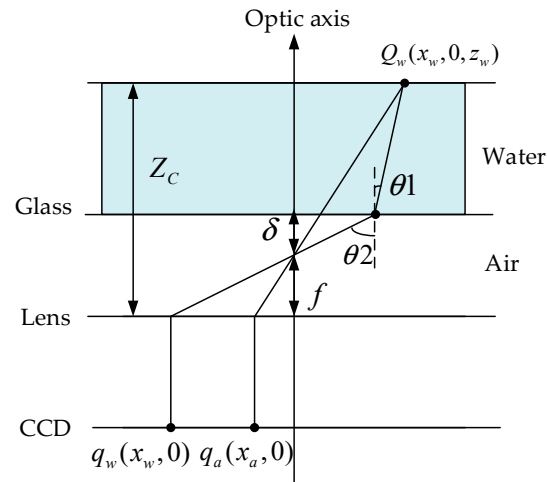
$$x_w = f \tan \theta_2 \tag{22}$$

$$x_w = \delta \tan \theta_2 + (z - \delta - f) \tan \theta_1 \tag{23}$$

$$n_w \sin\theta_1 = n_a \sin\theta_2 \tag{24}$$

where n_w and n_a are the refractive indices of water and air, respectively.

Figure 13. Schematic diagram of the underwater imaging model.

Given that the refraction angles are small, the sine and tangent values can be approximated by the angles. Combining Equations (22)–(24) yields the following:

$$x_{water} = \frac{f x_w}{\delta + (Z - \delta - f)\dfrac{n_a}{n_w}} \tag{25}$$

The z-axis coordinate Z_C of the target point Q in the camera coordinate system is obtained using the binocular ranging model. Figure 14 illustrates the schematic diagram of binocular vision-based ranging, where O_L and O_R represent the optical centers of the left and right camera lenses, respectively.
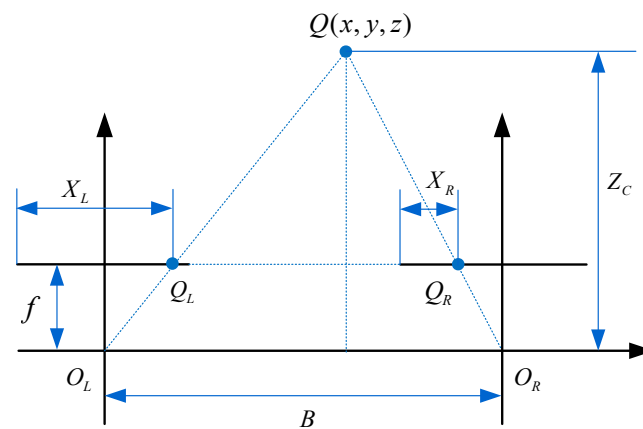
Figure 14. Schematic diagram of the binocular ranging model.

Based on the geometric relationship, the formula for calculating Z_C can be derived as follows:

$$Z_C = \frac{Bf}{X_L - X_R} = \frac{Bf}{d} \tag{26}$$

where Z_C is the actual distance from the target point to the camera; B is the baseline length of the binocular camera; f is the focal length; and d is the horizontal disparity of the target point between the left and right images.

The workflow for locating the target point is shown in Figure 15.
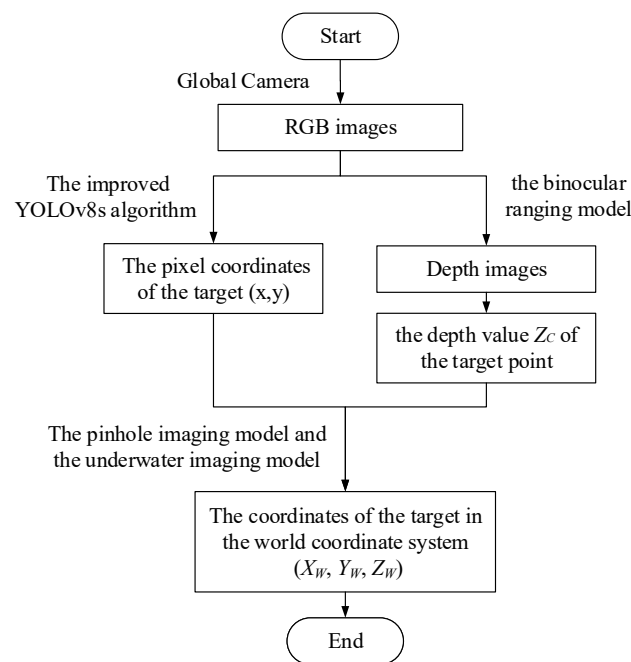
Figure 15. The process of foreign-object localization.

First, the depth camera captures a color image of the environment. This image is processed by the improved YOLOv8s algorithm, which outputs a detection bounding box containing the pixel coordinates (x,y) of its center. Simultaneously, the camera generates a depth image containing depth information based on the binocular ranging model, and the depth value Z_C of the target point can be retrieved. Finally, the world coordinate values of the target point are computed from x, y, and Z_C using the pinhole imaging model and the underwater imaging model.

4.3. Foreign-Object Grasping Strategy

After acquiring the 3D spatial coordinates of the target foreign object, the system controls the Delta arm to perform grasping and retrieval operations. To drive the gripper to the target location and perform the grasping operation by controlling the three motors of the Delta arm, kinematic analysis and path planning of the parallel manipulator are performed. The kinematic model of the Delta arm is established, as shown in Figure 16.
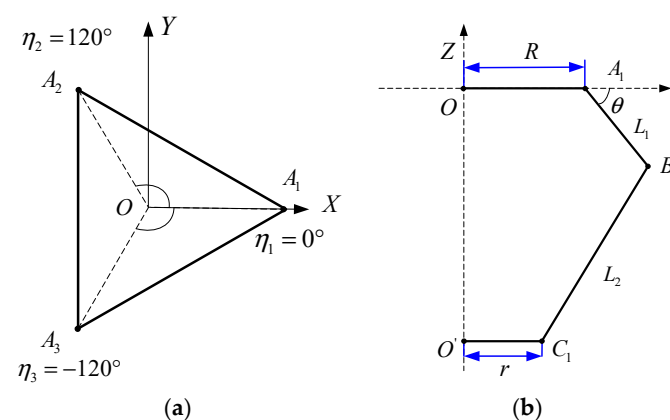
Figure 16. The kinematic model of the Delta arm: (**a**) top view model; (**b**) front view model.

The forward kinematics analysis of the Delta arm is to determine the position of the moving platform relative to the fixed platform's coordinate system based on known rotational angles of its three actuated arms.

The kinematic equation for a single chain of the parallel mechanism can be written as follows:

$$[x - (R + L_1 \cos\theta_i - r)\cos\eta_i]^2 + [y - (R + L_1 \cos\theta_i - r)\sin\eta_i]^2 + (z + L_1 \sin\theta_i)^2 = L_2^2 \tag{27}$$

By substituting $\eta_1 = 0$, $\eta_1 = 2\pi/3$, $\eta_1 = -2\pi/3$ sequentially into Equation (27), the forward kinematics solution for the Delta manipulator can be obtained, as outlined below:

$$\begin{cases} [x - (R + L_1 \cos\theta_i - r)]^2 + y^2 + (z + L_1 \sin\theta_i)^2 = L_2^2 \\ [x + (R + L_1 \cos\theta_i - r)/2]^2 + \sqrt{3}[y - (R + L_1 \cos\theta_i - r)/2]^2 + (z + L_1 \sin\theta_i)^2 = L_2^2 \\ [x + (R + L_1 \cos\theta_i - r)/2]^2 + \sqrt{3}[y + (R + L_1 \cos\theta_i - r)/2]^2 + (z + L_1 \sin\theta_i)^2 = L_2^2 \end{cases} \tag{28}$$

The inverse kinematics analysis of the Delta arm is to determine the rotational angles of its three actuated arms on the basis of given position coordinates of the end-effector. By setting $u = \tan(\theta/2)$, $\cos\theta = (1 - u^2)/(1 + u^2)$, $\sin\theta = 2u/(1 + u^2)$ and transforming the kinematic equation shown in Equation (27), we derive the following polynomial equations with respect to a single variable:

$$\begin{cases} A_i \cos\theta_i + B_i \sin\theta_i = C_i \\ A_i = 2L_1(R - r - x\cos\eta_i - y\sin\eta_i) \\ B_i = -2L_1 z \\ C_i = L_2^2 - L_1^2 + x^2 + y^2 + z^2 + (R - r)^2 + 2(R - r)(x\cos\eta_i + y\sin\eta_i) \end{cases} \tag{29}$$

The solution to the equation system above provides the rotational angle θ_i ($i = 1,2,3$) of the i-th joint for a given end-effector position (x, y, z):

$$\theta_i = 2\tan^{-1}\frac{B_i \pm \sqrt{B_i^2 + A_i^2 - C_i^2}}{A_i + C_i} \tag{30}$$

The path planning of the Delta arm is to determine the spatial coordinates (x, y, z) of the end-effector as a function of time t. To prevent excessive vibration and ensure motion stability during the retrieval operation, this paper employs fifth-order polynomial motion laws for trajectory planning of the Delta arm. The displacement expression under the fifth-order polynomial motion law is defined as follows:

$$s(t) = c_0 + c_1 t + c_2 t^2 + c_3 t^3 + c_4 t^4 + c_5 t^5 \tag{31}$$

Let t_0 denote the initial time when the end-effector is at the starting point, while t_1 represents the terminal time when it reaches the target point, the boundary conditions for displacement, velocity, and acceleration are defined as follows:

$$\begin{cases} s(t_0) = \dot{s}(t_0) = \ddot{s}(t_0) \\ s(t_1) = d \\ \dot{s}(t_1) = \ddot{s}(t_1) = 0 \end{cases} \tag{32}$$

Substituting Equation (32) into Equation (31) yields the solution for $s(t)$, that is, the distance traveled by the end-effector along the predefined trajectory at time t.

In this study, we employ a gate-shaped trajectory with fundamental parameters set to a height of 100 mm and a variable length of w. Corner transitions are smoothed using circular arcs with radii equal to one-fifth of the horizontal path length, as illustrated in Figure 17. Within the robot coordinate frame, the starting point is set at (x_1, y_1, z_1), corresponding to the foreign-object storage basket. The coordinates of target points, distributed

within a 250 mm radius circle centered at (0,0), are determined via the foreign-object localization algorithm.
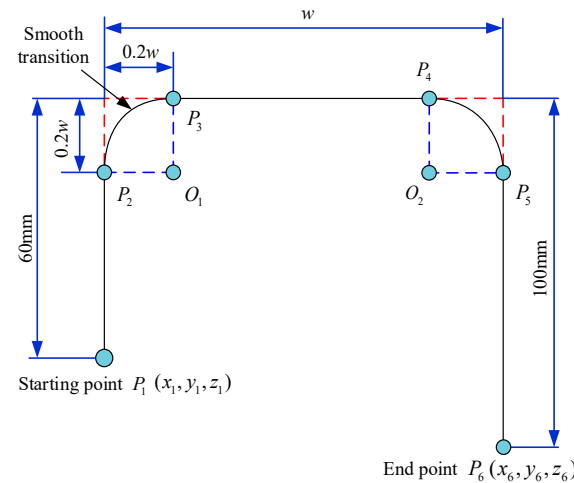
Figure 17. Trajectory planning of Delta arm end-effector.

Given the coordinates of the starting point (x_1, y_1, z_1) and the target point (x_6, y_6, z_6), the corresponding gate-shaped trajectory is uniquely determined. Therefore, $s(t)$ can be converted into the three-dimensional coordinates $(x(t), y(t), z(t))$ of the gripper at time t based on the geometric relationship illustrated in Figure 17. By combining $x(t), y(t), z(t)$ with Equation (30), the angular displacement curves $\theta_i(t)$ for the three joint motors of the parallel manipulator in the joint space can be obtained. Differentiating these curves yields the corresponding angular velocity and angular acceleration profiles, thereby enabling precise control of the end-effector's motion. Figure 18 illustrates how the gripper moves from the initial position to the target position under the control of the three motors.
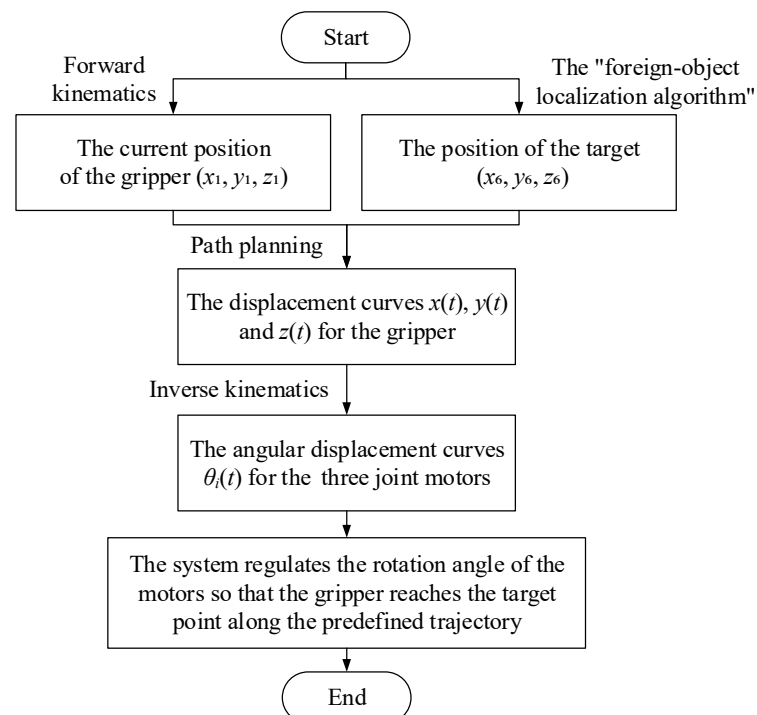
**Figure 18.** Flowchart of the gripper moving from the initial point to the target point.

The gripper is driven by a servo motor. When receiving different servo-encoded signals, the servo's output shaft rotates to specific angular positions and maintains that

position as long as the encoded signal persists. In this study, a 50 Hz PWM signal is employed as the control input for the servo, with the rotation angle regulated by varying the PWM duty cycle to achieve gripper opening/closing. The duty cycle is defined as the ratio of high-level duration to the total PWM period within a single pulse cycle. The servo's rotation range from 0° to 180° corresponds to high-level durations of 500 μs to 2500 μs. At a 50 Hz PWM frequency, the corresponding duty cycle is 2.5% to 12.5%. When the duty cycle is at 2.5%, the servo rotates to the 0° position and closes the gripper; when the duty cycle is at 12.5%, it rotates to 180° and makes the gripper fully open.

Upon reaching the target's location along the planned trajectory, the gripper closes via PWM signal control to grasp the foreign object. Subsequently, the system generates a retrieval trajectory to the storage basket using the coordinates stored before. When positioned above the storage basket, the gripper opens to release the object into the basket, thereby completing the foreign-object retrieval operation.

5. Experiments

With the analyses of the mobility module, grasping module, vision module, control module, and each function, the prototype of the USR has been fabricated and integrated, as shown in Figure 19. The main controller employs an Orange Pi core processing unit for handling vision, IMU, and other sensor data, generating control commands that are transmitted to the Pixhawk flight controller and C2200 motion controller. The Pixhawk system regulates six thrusters via electronic regulators, while the C2200 controller drives three joint motors of the Delta arm and the gripper's servo. Two lithium-ion batteries rated at 25.2 V and 16 Ah are used to power the main control box and other equipment via waterproof cables. For sensors, the global camera employs a 1080 p USB waterproof monocular camera, while the overhead camera utilizes an Intel RealSense D435i depth camera. The inertial measurement unit employs an MPU6500 6-axis sensor. The altimeter offers a measurement range of 20 m, and the depth sensor features a 3 MPa pressure rating, satisfying all operational requirements.
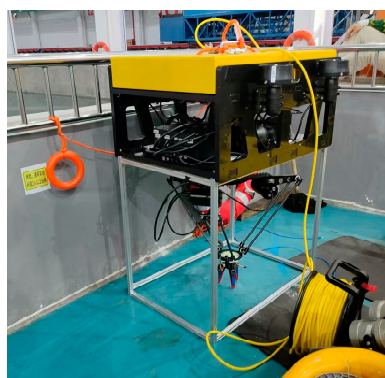
Figure 19. The prototype of the USR.

In terms of sensor calibration, we adopted a multi-sensor joint calibration scheme based on the Kalibr toolbox. During the calibration process, the camera–IMU system is moved in front of a calibration board at varying speeds and orientations, while simultaneously capturing image data from the camera and inertial measurements from the IMU. Kalibr then automatically performs the calibration, simplifying the complex multi-sensor calibration task into a standardized procedure that can be completed within a relatively short time.

In terms of algorithm integration, the Robot Operating System (ROS) architecture is implemented. To meet the functional requirements of the USR, the ROS framework

on the main controller incorporates dedicated nodes for environmental perception, attitude control, Delta arm control, and human–machine interaction. The environmental perception node processes data provided by cameras and other sensors; the attitude control node regulates robotic motion based on sensor feedback; the Delta arm control node coordinates the Delta arm and gripper for object-retrieval operations; the human–machine interaction node manages the operator interfaces. Each node executes specialized tasks independently while exchanging data through ROS communication mechanisms, ensuring integrated functionality.

The physical prototype of the LED beacon is shown in Figure 20. The frame is constructed from aluminum profiles. The LED units are potted with epoxy resin and anchored at designated positions on the frame. The LEDs are connected in parallel and powered by a 5 V DC power supply.
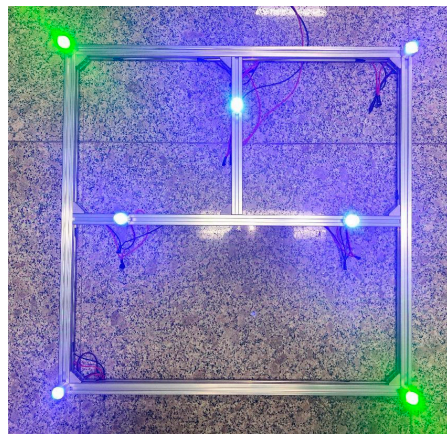
Figure 20. The prototype of the LED beacon.

To validate the algorithmic performance and operational efficacy of the USR, a series of experimental tests was conducted in an indoor water pool, measuring 10 m (L) × 5 m (W) × 2 m (D), which provided sufficient underwater space to simulate traversal and foreign-object-retrieval tasks. The experimental site is shown in Figure 21. The experiments mainly contain three aspects: (1) underwater localization and navigation experiments; (2) foreign-object identification, localization, and grasping experiments; (3) autonomous underwater object-retrieval experiments.


Figure 21. Experimental site for the USR.

5.1. Underwater Localization and Navigation Experiments

To validate the efficacy of the beacon-based vision localization theory, systematic experiments were conducted across varying distance ranges. The beacons were mounted on

the pool walls or suspended from the water surface following the configuration illustrated in Figure 7, as shown in Figure 22.

(a)   (b)
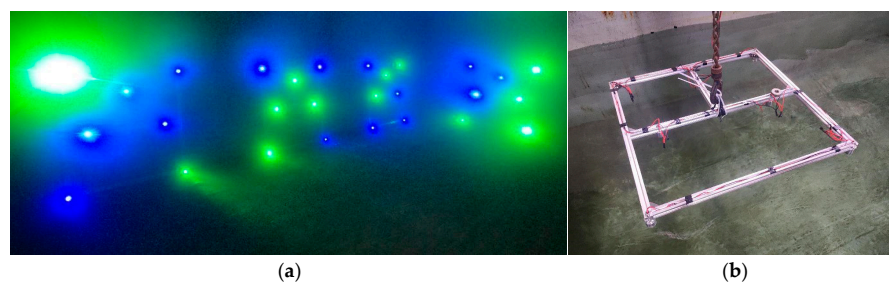
Figure 22. The deployment of LED beacons: (**a**) on the pools; (**b**) above the water surface.

In the first set of experiments, the USR was placed at the corner of the pool and was commanded to move along the first straight segment of the zigzag trajectory at a speed of 0.5 m/s. The USR sent back positioning data every 2 s. Tests were performed separately for the pure visual localization scheme and the multi-sensor fusion localization scheme, and each scheme was tested over 10 independent experimental runs. Finally, the average absolute error and relative error at different positions of the USR were recorded to evaluate the positioning accuracy of the underwater robot under different schemes. The formulas for calculating the absolute error and relative error are as follows:

$$\varepsilon = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2} \tag{33}$$

$$\lambda = \varepsilon / L \tag{34}$$

where ε represents the absolute error; λ represents the relative error; Δx, Δy, and Δz denote the error values in the *X*, *Y*, and *Z* directions, respectively; and *L* denotes the distance between the USR and the beacon.

The experimental results are presented in Table 1 and Figure 23. The data indicate that the pure visual localization scheme exhibits relatively large absolute and relative errors when the USR is far from the beacon board. This is primarily due to the more pronounced interference between light sources at long ranges, which complicates the pixel-level localization of points on the beacon board. In contrast, the multi-sensor fusion localization scheme demonstrates significantly reduced absolute and relative positioning errors within the 5–10 m range, with the absolute error consistently remaining below 8 cm and the relative error controlled within 2%.

Table 1. Underwater localization experiment results.

Time/s	Distance/m	Pure Visual Localization		Multi-Sensor Fusion Localization	
		Average Absolute Error/mm	Relative Error/%	Average Absolute Error/mm	Relative Error/%
0	10	421.2	4.21	84.7	0.84
2	9	394.6	4.38	70.3	0.78
4	8	352.6	4.41	63.1	0.79
6	7	267.2	3.82	58.1	0.83
8	6	171.6	2.86	53.2	0.87
10	5	86.2	1.72	37.1	0.74
12	4	46.1	1.15	31.6	0.79
14	3	26.1	0.87	22.5	0.75
16	2	21.2	1.06	19.4	0.97
18	1	16.9	1.69	16.6	1.66
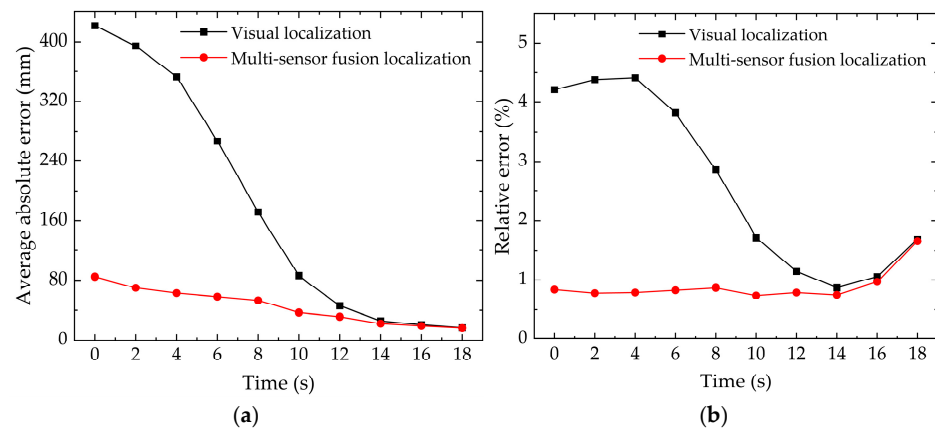
Figure 23. (**a**) Average absolute error for visual and multi-sensor fusion localization; (**b**) relative error for visual and multi-sensor fusion localization.

In the second set of experiments, the USR was commanded to follow the zigzag trajectory and the circular trajectory at a speed of 0.1 m/s, with fused localization results recorded every 2 s. Ten experimental trials were conducted for each trajectory, and the average absolute errors at multiple points were recorded to evaluate the control performance of the robot during autonomous movement. Figure 24 illustrates the USR moving along the zigzag trajectory within the test pool.
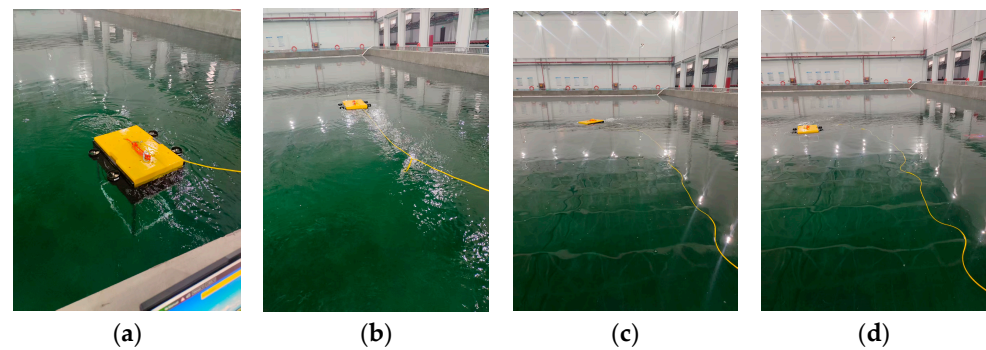
Figure 24. Process of the USR moving along the zigzag trajectory: (**a**) release stage; (**b**) initial straight-line traversal stage; (**c**) left-turn maneuver stage; (**d**) second straight-line traversal stage.

As shown in Figure 25, the experimental results demonstrate that for both paths, the absolute error of the USR's motion control remained within 9 cm. Furthermore, the system exhibited excellent responsiveness and stability, enabling rapid adaptation to environmental changes and efficient task execution.
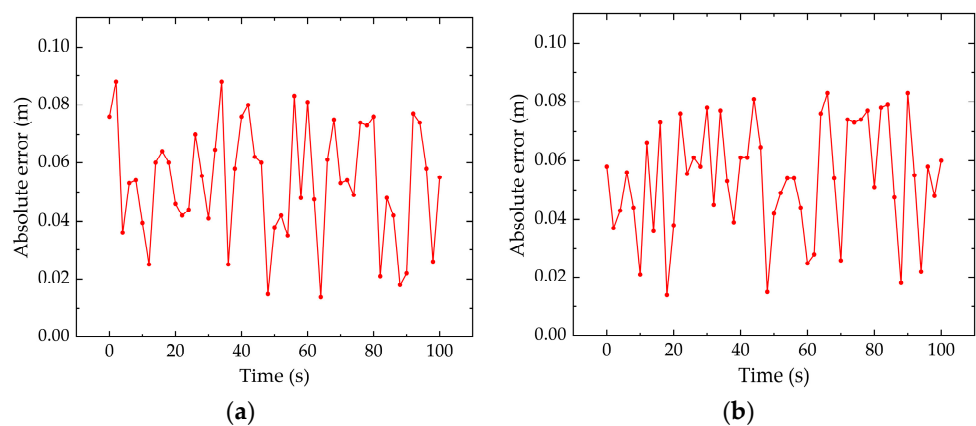
Figure 25. Motion error analysis curves: (**a**) zigzag trajectory; (**b**) circular trajectory.

The position error of the USR comprises the localization error and the motion control error. Given that the localization error remains within 8.5 cm and the motion control error remains within 9 cm, the maximum position error of the USR is 17.5 cm.

During the traversal operation, the USR maintains a suspended height of $H = 750$ mm, with the selected global camera possessing a horizontal FOV of $\theta_x = 86°$. The width of the camera's FOV is calculated to be $D = 12$ cm. According to the path planning design illustrated in Figure 4, when under the maximum positioning error condition, the FOV coverage on one side of the camera's optical axis remains greater than half the width of the planned path (0.5 m), as visually demonstrated in Figure 26. This result indicates that the localization and motion control algorithms adopted in this study ensure the camera's FOV fully covers the underwater region, thereby enabling the USR to effectively accomplish comprehensive pool-floor inspection.
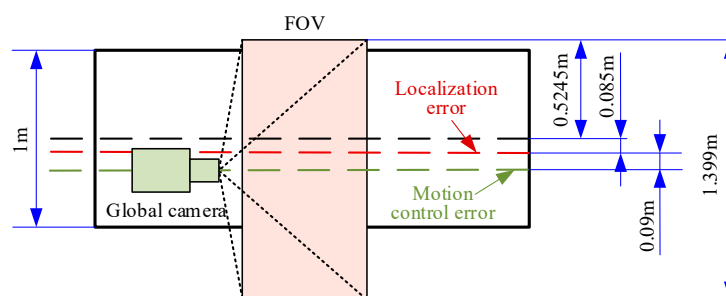
Figure 26. Schematic diagram of the camera's FOV range under maximum error condition.

5.2. Foreign-Object Identification, Localization, and Grasping Experiments

A total of 20 experimental trials were designed, each deploying a variable quantity of foreign objects including screws, nuts, washers, and metallic wires in different poses within the overhead camera's field of view. The cumulative identification count reached 112 objects. To further analyze the performance of the foreign-object identification algorithm, experimental results of the improved YOLOv8s network were compared against other models, including the original YOLOv8s, YOLOv5s, and Faster-RCNN. After identifying objects, localization and grasping operations were executed, and the success rates were recorded. The autonomous grasping sequence of the USR is depicted in Figure 27. Table 2 compares the recognition performance and parameters of various network models. The improved YOLOv8s model achieves mAP improvements of 3.7%, 5.3%, and 3.9% over original YOLOv8s, YOLOv5s, and Faster R-CNN, respectively. Notably, as a two-stage detection algorithm, Faster R-CNN delivers higher accuracy but slower inference speeds. Among single-stage detectors, the improved YOLOv8s exhibits a slight increase in inference time compared to the original architecture while still meeting real-time requirements and delivering superior detection accuracy.

(**a**)   (**b**)

Figure 27. *Cont.*

(**c**)   (**d**)

Figure 27. The process of the USR retrieving foreign objects: (**a**) the USR identifies the target and reaches the optimal position for retrieval; (**b**) the gripper reaches the target's location; (**c**) the gripper grasps the target; (**d**) the gripper moves above the storage basket and releases the target.

Table 2. Comparison of recognition effects between different algorithms.

Model	mAP (%)	Detection Speed (s/pic)
Improved YOLOv8s	92.2	0.088
YOLOv8s	88.5	0.076
YOLOv5s	86.9	0.094
Faster-RCNN	88.9	0.599

Among the 112 foreign-object identification trials, 108 objects were successfully detected, representing a 96.4% success rate in identification; 107 were accurately localized, representing a 95.5% success rate in localization. The localization failures primarily occurred with elongated objects like metallic wires, where deviations between the positioning center and the object's physical center adversely impact grasping precision.

5.3. Autonomous Underwater Object-Retrieval Experiments

This set of experiments required the USR to complete full-cycle operations from traversal detection to foreign object retrieval autonomously. A total of 50 experimental trials were designed, each deploying 5–20 randomly distributed foreign objects to simulate authentic underwater working environments, with the cumulative number of objects amounting to 534. In each trial, the USR first scanned the target area comprehensively. Upon successful detection, it stored the current position and calculated the 3D coordinates of targets via the foreign-object localization algorithm. Based on these coordinates, the USR navigated autonomously to position itself directly above the target, and then activated vertical thrusters to land on the pool floor. After that, the vision system reconfirmed the updated spatial coordinates of the target and transmitted data to the Delta arm controller for grasping and retrieval until the vision system could no longer detect the target. The USR then returned to its initial stored position to resume area traversal. This cycle repeated until all targets were recovered. Critical metrics, including the total number of objects involved and the times of successful retrievals, were recorded. Due to space limitations, 10 groups of data are randomly selected from 50 trials for demonstration, as shown in Table 3.

Among 534 retrieval operations, the target was successfully identified in 506 cases, achieving an average detection rate of 94.76% (95% CI: 92.5–96.5%). Successful retrieval was achieved 472 times, achieving an average foreign-object-retrieval rate of 88.39% (95% CI: 85.6–90.9%). Though slightly lower than those under static conditions, they effectively meet real-time operational requirements.

Table 3. Results of foreign-object-retrieving experiments.

No.	Number of Objects	Successful Detection	Successful Retrieval	Detection % (95% CI)	Retrieval % (95% CI)
1	10	9	9	90.0 (59.59–98.21)	90.0 (59.59–98.21)
2	13	12	11	92.3 (66.69–98.63)	84.6 (57.77–95.68)
3	8	8	8	100 (67.56–100)	100 (67.56–100)
4	12	12	10	100 (75.75–100)	83.3 (55.19–95.3)
5	11	11	10	100 (74.12–100)	90.9 (62.27–90.91)
6	14	13	11	92.9 (68.53–98.73)	78.6 (52.41–92.43)
7	5	5	5	100 (56.55–100)	100 (56.55–100)
8	7	6	6	85.7 (48.68–97.43)	85.7 (48.68–97.43)
9	15	14	14	93.3 (70.18–98.81)	93.3 (70.18–98.81)
10	13	13	13	100 (77.19–100)	100 (77.19–100)

The failure of autonomous retrieval operations primarily stems from the following factors: (1) During traversal, dynamic images are more blurred than static images, resulting in unsuccessful detection of foreign objects within the visual field. (2) The limited positioning accuracy led to failed target grasping. (3) The adopted flexible three-finger gripper cannot achieve a stable grasp for certain shapes like curved metallic wires. (4) In order to improve operational efficiency, the movement speed of the Delta arm is set to a higher level, which may induce vibrations and inertial forces, causing objects to drop during retrieval operations. (5) The movement of the Delta arm and the gripper generates water turbulence, displacing lightweight objects and invalidating their calculated positioning coordinates, which affects the subsequent gripping operations.

6. Conclusions

To address the demand for foreign-object retrieval in nuclear pool environments, we propose an ROV system incorporating a Delta robotic arm. The system achieves autonomous underwater object retrieval through multi-sensor fusion positioning, enhanced sliding mode control, and an improved YOLOv8s detection algorithm. Experimental validation demonstrates that compared to existing research, the multi-sensor fusion positioning method reduces positioning errors by 50% within 4–8 m and 80% within 10 m compared to vision-only methods. The improved YOLOv8s algorithm achieves a 92.2% mAP in small object detection, surpassing the original YOLOv8s by 3.7 percentage points and Faster-RCNN by 3.3 percentage points. The average foreign-object-retrieval success rate reaches 90.64%, which is a high level compared to other salvage robots for nuclear power pools. These results validate the reliability of the robot system and its engineering value in nuclear pool maintenance. It is also worth noting that the proposed system was designed with cost-effectiveness in mind. The USR utilizes Commercial Off-The-Shelf (COTS) components for the ROV hardware platforms and adopts open-source ROS architecture for its framework. Furthermore, the modular architecture allows for easy replacement of individual subsystems in case of failure, avoiding the need to replace the entire system. These approaches significantly reduce the development and maintenance costs.

However, this study still exhibits notable limitations. For example, the life span of the battery system is limited under continuous operating conditions; the system demonstrates significantly lower retrieval success rates for slender-shaped foreign objects due to positioning inaccuracies, gripper limitations, and water disturbances; due to the extended maintenance cycles of nuclear power plants, the current experiments were conducted in a laboratory water pool environment, so the radiation tolerance of the system and the robustness of the visual inspection algorithm against a nuclear environment have not been tested.

To overcome these limitations, the following directions could be investigated:

1. A power supply scheme that steps down the high-voltage direct current and transmits it through cables could be adopted to support prolonged operation of the USR;
2. Dynamic images could be preprocessed using deep learning-based dynamic scene deblurring algorithms to enhance the system's recognition accuracy and localization precision of target objects;
3. The triple-loop control strategy could be employed to drive the motors of the Delta arm, thereby enhancing both the absolute positioning accuracy of the end-effector and its resistance to water flow disturbances;
4. An automatic quick-change end-effector system with multiple grasping tools, including three-finger grippers, two-finger grippers, and electromagnetic actuators, could be developed to handle targets of different shapes;
5. The system's radiation tolerance could be improved, and experiments within a real nuclear reactor pool need to be conducted.

Author Contributions: Conceptualization, M.Z.; Funding acquisition, M.Z.; Investigation, Z.M., R.L. and Y.L.; Methodology, Z.G., Z.M. and R.L.; Resources, Z.M. and R.L.; Software, Z.G.; Supervision, Y.L. and M.Z.; Validation, Z.M. and R.L.; Writing—original draft, Z.G.; Writing—review & editing, M.Z. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The original contributions presented in this study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author(s).

Conflicts of Interest: Author Zhengxiong Mao and Ruifei Lyu were employed by the company State Nuclear Power Demonstration Plant Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Kim, J.H.; Lee, J.C.; Choi, Y.R. LAROB: Laser-Guided Underwater Mobile Robot for Reactor Vessel Inspection. *IEEE/ASME Trans. Mechatron.* **2014**, *19*, 1216–1225. [CrossRef]
2. Cho, B.H.; Byun, S.H.; Shin, C.H.; Yang, J.B.; Song, S.I.; Oh, J.M. KeproVt: Underwater Robotic System for Visual Inspection of Nuclear Reactor Internals. *Nucl. Eng. Des.* **2004**, *231*, 327–335. [CrossRef]
3. Park, J.Y.; Cho, B.H.; Lee, J.K. Trajectory-tracking Control of Underwater Inspection Robot for Nuclear Reactor Internals Using Time Delay Control. *Nucl. Eng. Des.* **2009**, *239*, 2543–2550. [CrossRef]
4. Mazumdar, A.; Lozano, M.; Fittery, A.; Asada, H.H. A Compact, Maneuverable, Underwater Robot for Direct Inspection of Nuclear Power Piping systems. In Proceedings of the 2012 IEEE International Conference on Robotics and Automation (RAS), Saint Paul, MN, USA, 14–18 May 2012; pp. 2818–2823.
5. Josip, J.; Markulin, K.; Pavlović, N. TARGET–Development of Submersible ROV System for BMN Inspection. *J. Energy Energ.* **2022**, *71*, 24–28.
6. Leon-Rodriguez, H.E.; Sattar, T.; Shang, J.; Bouloubasis, A.K.; Markopoulos, Y.P. Wall Climbing and Pipe Crawler Robots for Nozzle Weld Inspection Inside Nuclear Pressure Vessels. In Proceedings of the 9th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines (CLAWAR), Brussels, Belgium, 12–14 September 2006; pp. 545–551.
7. Luo, Y.; Tao, J.; Sun, Q.; Deng, L.; Deng, Z. A New Underwater Robot for Crack Welding in Nuclear Power Plants. In Proceedings of the 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO), Kuala Lumpur, Malaysia, 12–15 December 2018; pp. 77–82.
8. Lee, S.U.; Choi, Y.S.; Jeong, K.M.; Jung, S. Development of a Tele-operated Underwater Robotic System for maintaining a light-water type power reactor. In Proceedings of the 2006 SICE-ICASE International Joint Conference, Busan, Republic of Korea, 18–21 October 2006; pp. 3017–3021.

9. Iborra, A.; Alvarez, B.; Navarro, P.J.; Fernandez, J.M.; Pastor-Franco, J.A. Robotized system for retrieving fallen objects within the reactor vessel of a nuclear power plant (PWR). In Proceedings of the 2000 IEEE International Symposium on Industrial Electronics (ISIE), Cholula, Puebla, Mexico, 4–8 December 2000; pp. 529–534.

10. Zhang, L.; Lan, Q.; Yu, L.; Wang, C.; Chen, W.; Lei, Y. Development of a Cleaning ROV for High-Temperature Radioactive Environment. In Proceedings of the International Conference on Intelligent Robotics and Applications (ICRA), Singapore, 31 July 2024; pp. 116–130.

11. Dong, M.J.; Chou, W.S.; Yao, G.D. A new navigation strategy for underwater robot in reactor pool combined propeller speed detection and dynamics analysis with sonar data correction. *J. Nucl. Sci. Technol.* **2018**, *55*, 1–10. [CrossRef]

12. Martinez, A.; Hernandez, L.; Sahli, H.; Valeriano-Medina, Y.; Orozco-Monteagudo, M.; Garcia-Garcia, D. Model-aided Navigation with Sea Current Estimation for an Autonomous Underwater Vehicle. *Int. J. Adv. Robot. Syst.* **2015**, *12*, 103. [CrossRef]

13. Zhai, Y.Y.; Gong, Z.B.; Wang, L.; Zhang, R.Y.; Luo, H.X. Study of Underwater Positoning Based on Short Baseline Sonar System. In Proceedings of the 2009 International Conference on Artificial Intelligence and Computational Intelligence (AICI), Shanghai, China, 7–8 November 2009; pp. 343–346.

14. Bahr, A.; Leonard, J.J.; Fallon, M.F. Cooperative Localization for Autonomous Underwater Vehicles. *Int. J. Robot. Res.* **2009**, *28*, 714–728. [CrossRef]

15. Sato, Y.; Maki, T.; Mizushima, H.; Matsuda, T.; Sakamaki, T. Evaluation of Position Estimation of AUV Tri-TON 2 in Real Sea Experiments. In Proceedings of the OCEANS 2015—Genova, Genova, Italy, 18–21 May 2015; pp. 1–6.

16. Jung, J.; Li, J.H.; Choi, H.T.; Myung, H. Localization of AUVs using visual information of underwater structures and artificial landmarks. *Intell. Serv. Robot.* **2017**, *10*, 67–76. [CrossRef]

17. Zhang, S.; Zhao, S.; An, D.; Liu, J.; Wang, H.; Feng, Y.; Li, D.; Zhao, R. Visual SLAM for underwater vehicles: A survey. *Comput. Sci. Rev.* **2022**, *46*, 100510. [CrossRef]

18. Beijbom, O.; Edmunds, P.J.; Kline, D.I.; Mitchell, B.G.; Kriegman, D. Automated annotation of coral reef survey images. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, USA, 16–21 June 2012; pp. 1170–1177.

19. Kim, D.; Lee, D.; Myung, H.; Choi, H.T. Artificial landmark-based underwater localization for AUVs using weighted template matching. *Intel. Serv. Robot.* **2014**, *7*, 175–184. [CrossRef]

20. Barat, C.; Phlypo, R. A Fully Automated Method to Detect and Segment a Manufactured Object in an Underwater Color Image. *EURASIP J. Adv. Signal Process.* **2010**, *1*, 568092. [CrossRef]

21. Terven, J.; Córdova-Esparza, D.M.; Romero-González, J.A. A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Mach. Learn. Knowl. Extr.* **2023**, *20*, 1680–1716. [CrossRef]

22. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 1137–1149. [CrossRef]

23. Hanson, S.C.; Hayes, R.B. Radiation Hardness Assurance by Redundancy in Raspberry Pi Zero W Computation Metrics via Total Ionizing Dose 60Co Testing for Spacecraft Applications. *Health Phys.* **2025**, *128*, 457–466. [CrossRef]

24. Abubakkar, S.F.O.; Zabah, N.F.; Abdullah, Y.; Fauzi, D.A.; Muridan, N.; Hasbullah, N.F. Effects of electron radiation on commercial power MOSFET with buck converter application. *Nucl. Sci. Tech.* **2017**, *28*, 31. [CrossRef]

25. Huang, Q. Investigation of Radiation-Hardened Design of Electronic Systems with Applications to Post-Accident Monitoring for Nuclear Power Plants. Ph.D. Thesis, The University of Western Ontario, London, ON, Canada, 2019.

26. Luo, R.; Kang, D.; Huang, C.; Yan, T.; Li, P.; Ren, H.; Zhang, Z. Mechanical properties, radiation resistance performances, and mechanism insights of nitrile butadiene rubber irradiated with high-dose gamma rays. *Polymers* **2023**, *15*, 3723. [CrossRef]

27. Porter, C.P.; Edge, R.; Ogden, M.D. Polymeric seal degradation in nuclear power plants: Effect of gamma radiation on sealing properties. *J. Appl. Polym. Sci.* **2017**, *134*, 44618. [CrossRef]

28. Zenoni, A.; Bignotti, F.; Donzella, A.; Donzella, G.; Ferrari, M.; Pandini, S.; Andrighetto, A.; Ballan, M.; Corradetti, S.; Manzolaro, M.; et al. Radiation resistance of elastomeric O-rings in mixed neutron and gamma fields: Testing methodology and experimental results. *Rev. Sci. Instrum.* **2017**, *88*, 113304. [CrossRef]

29. Kaur, R.; Saini, D. Image enhancement of underwater digital images by utilizing L* A* B* color space on gradient and CLAHE based smoothing. *Image* **2016**, *4*, 22–30. [CrossRef]

30. Kaur, S.; Singh, I. Comparison between edge detection techniques. *Int. J. Comput. Appl.* **2016**, *145*, 15–18. [CrossRef]

31. Liang, L.; Chen, J.; Shi, J.; Zhang, K.; Zheng, X. Noise-Robust image edge detection based on multi-scale automatic anisotropic morphological Gaussian Kernels. *PLoS ONE* **2025**, *20*, e0319852. [CrossRef]

32. Kierkegaard, P. A method for detection of circular arcs based on the Hough transform. *Mach. Vis. Appl.* **1992**, *5*, 249–263. [CrossRef]

33. Leng, D.; Sun, W. Finding all the solutions of PnP problem. In Proceedings of the 2009 IEEE International Workshop on Imaging Systems and Techniques (IST), Shenzhen, China, 11–12 May 2009; pp. 348–352.

34. Li, Q.; Li, R.; Ji, K.; Dai, W. Kalman Filter and Its Application. In Proceedings of the 2015 8th International Conference on Intelligent Networks and Intelligent Systems (ICINIS), Tianjin, China, 1–3 November 2015; pp. 74–77.

35. Ren, F.; Hu, Q. ROV sliding mode controller design and simulation. *Processes* **2023**, *11*, 2359. [CrossRef]

36. Garcia-Nava, S.; García-Rangel, M.A.; Zamora-Suárez, Á.E.; Manzanilla-Magallanes, A.; Muñoz, F.; Lozano, R.; Serrano-Almeida, A. Development of a 6 Degree of Freedom Unmanned Underwater Vehicle: Design, Construction and Real-Time Experiments. *J. Mar. Sci. Eng.* **2023**, *11*, 1744. [CrossRef]

37. González-García, J.; Gómez-Espinosa, A.; García-Valdovinos, L.G.; Salgado-Jiménez, T.; Cuan-Urquizo, E.; Escobedo Cabello, J.A. Experimental validation of a model-free high-order sliding mode controller with finite-time convergence for trajectory tracking of autonomous underwater vehicles. *Sensors* **2022**, *22*, 488. [CrossRef] [PubMed]

38. Maalouf, D.; Chemori, A.; Creuze, V. L1 adaptive depth and pitch control of an underwater vehicle with real-time experiments. *Ocean Eng.* **2015**, *98*, 66–77. [CrossRef]

39. Velazquez, M.; Cruz, D.; Garcia, S.; Bandala, M. Velocity and Motion Control of a Self-Balancing Vehicle Based on a Cascade Control Strategy. *Int. J. Adv. Robot. Syst.* **2016**, *13*, 106. [CrossRef]

40. Akçakaya, H.; Yildiz, H.A.; Sağlam, G.; Gürleyen, F. Sliding mode control of autonomous underwater vehicle. In Proceedings of the 2009 International Conference on Electrical and Electronics Engineering (ICEEE), Bursa, Turkey, 5–8 November 2009; pp. II–332–II–336.

41. Guo, L.; Liu, W.; Li, L.; Lou, Y.; Wang, X.; Liu, Z. Neural network non-singular terminal sliding mode control for target tracking of underactuated underwater robots with prescribed performance. *J. Mar. Sci. Eng.* **2022**, *10*, 252. [CrossRef]

42. Bouzerzour, H.; Guiatni, M.; Hamerlain, M.; Allam, A. Vision-based Sliding Mode Control with Exponential Reaching Law for Uncooperative Ground Target Searching and Tracking by Quadcopter. In Proceedings of the 19th International Conference on Informatics in Control, Automation and Robotics (ICINCO), Lisbon, Portugal, 14–16 July 2022; pp. 555–564.

43. Lee, H.; Utkin, V.I. Chattering suppression methods in sliding mode control systems. *Annu. Rev. Control* **2007**, *31*, 179–188.

44. Lu, Y.; Wu, J.; Zhou, H. Trajectory tracking control of underwater vehicle based on hydrodynamic parameters calculated by CFD. *Chin. J. Ship Res.* **2022**, *17*, 237–245.

45. Ma, Y.; Cheng, Y.; Zhang, D. Comparative analysis of traditional and deep learning approaches for underwater remote sensing image enhancement: A quantitative study. *J. Mar. Sci. Eng.* **2025**, *13*, 899. [CrossRef]

46. Sohan, M.; Sai Ram, T.; Rami Reddy, C.V. A review on yolov8 and its advancements. In Proceedings of the International Conference on Data Intelligence and Cognitive Informatics (ICDICI), Singapore, 7 January 2024; pp. 529–545.

47. Zhong, M.; Ma, Y.; Li, Z.; He, J.; Liu, Y. Facade Protrusion Recognition and Operation Effect Inspection Methods Based on Binocular Vision for Wall-Climbing Robots. *Appl. Sci.* **2023**, *13*, 5721. [CrossRef]

48. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.