Preprints.org

Advancements in Underwater Navigation: Integrating Deep Learning and Sensor Technologies for Unmanned Underwater Vehicles

Fomekong Fomekong Rachel Merveille , Baozhu Jia [*] , Zhizun Xu

Review

Advancements in Underwater Navigation: Integrating Deep Learning and Sensor Technologies for Unmanned Underwater Vehicles

Fomekong Fomekong Rachel Merveille, Baozhu Jia * and Zhizun Xu

Maritime College, Guangdong Ocean University, Zhanjiang 524000 China; marvelous@stu.gdou.edu.cn; zhizun@gdou.edu.cn (Z.X.)

* Correspondence: jiabzh@gdou.edu.cn

Abstract: Unmanned Underwater Vehicles (UUVs) are pivotal in ocean exploration, research, and various industrial activities such as marine mining and offshore engineering. However, traditional methods of navigating these vehicles face significant challenges, mainly due to the inefficacy of electromagnetic waves in water, leading to signal loss. To address these limitations, researchers are increasingly turned learning, an artificial intelligence technique capable of learning from data-enhancing underwater navigation, mainly through visual Simultaneous Localization and Mapping (SLAM). This paper explores integrating deep learning methodologies and sensor technologies to revolutionize underwater navigation for UUVs. Proprioceptive sensors, along with exteroceptive sensors, are crucial in accurately measuring and comprehending the underwater environment. Additionally, the paper provides detailed insights into the processes of underwater SLAM, camera-based underwater positioning systems, sonar systems for underwater navigation, and the utilization of Lidar in underwater navigation. Furthermore, it delves into applying deep learning techniques in underwater SLAM, offering a comprehensive understanding of the innovative processes driving advancements in underwater vehicle navigation. By leveraging these advancements, this research aims to improve underwater navigation systems' precision, reliability, and adaptability, thereby unlocking new frontiers in ocean exploration and industrial applications for UUVs.

Keywords: visual simultaneous localization and mapping (SLAM); unmanned underwater vehicle; deep learning; underwater navigation; sensor integration
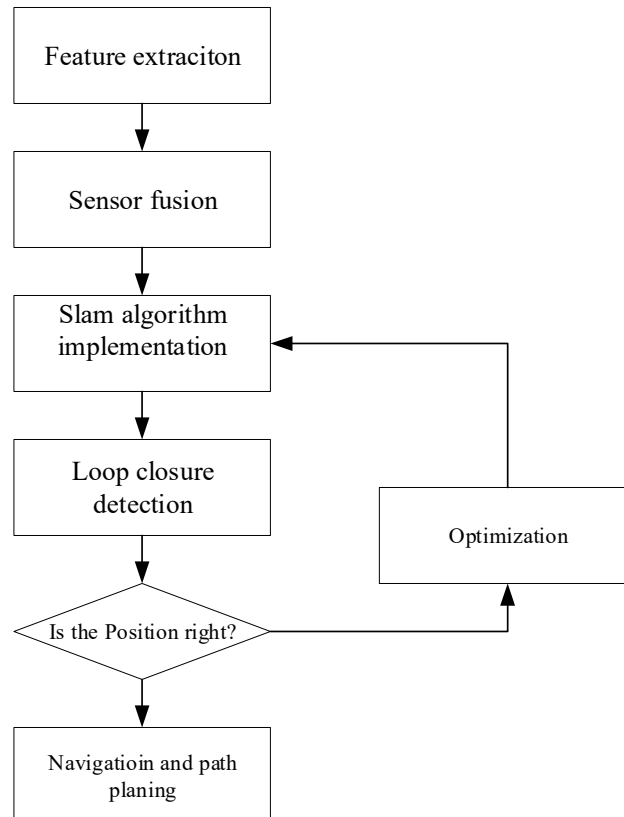

I. Introduction

Navigating the vast and mysterious underwater world is no easy feat, especially when faced with challenges like electromagnetic wave attenuation and limited visibility [1]. How can we improve the precision and reliability of underwater navigation? This research paper delves into integrating multiple sensors and deep learning techniques to enhance underwater navigation and perception, offering innovative solutions to longstanding obstacles.

i)   Background Context

The Earth's oceans cover about 71% of the planet's surface, holding immense value for resources, scientific exploration, and environmental understanding [2,3]. Unmanned Underwater Vehicles (UUVs) [4,5] are instrumental in various applications, from marine mining to pipeline inspection, but their effectiveness is hindered by the limitations of traditional navigation methods [6,7]. These methods, such as inertial sensors and acoustic beacons, struggle in challenging underwater conditions due to cumulative errors, limited range, and environmental interference [7,8]. Unique optical obstacles like low lighting, turbidity scattering, and wavelength absorption affect the quality and reliability of visual data. Moreover, the absence of access to global positioning systems (GPS) [9–11] complicates precise location pinpointing, data collection, etc. The quest for more reliable and

accurate underwater navigation has led to exploring deep learning techniques, particularly in the context of visual Simultaneous Localization and Mapping (SLAM), as a promising avenue for improvement. Underwater navigation poses formidable challenges due to limited visibility, high pressure, harsh conditions, complex terrain, limited communication, sensor integration issues, and the absence of reliable navigation references like GPS. These challenges render traditional SLAM algorithms inadequate for subsea navigation, as they struggle with sensor unsuitability, environmental variability, feature scarcity, and communication constraints. Despite these challenges, the quest for more reliable and accurate underwater navigation has led to exploring deep learning techniques, particularly in visual SLAM, as a promising avenue for improvement along with sensor fusion.

Figure 1. Underwater SLAM process.

ii)   Introduction of standard sensors and methodologies for underwater navigation and perception

Underwater sensors are specialized electronic devices that measure physical and environmental parameters in the ocean, enabling data collection for navigation, research, and monitoring purposes in challenging underwater conditions. There are different types of sensors used for a specific task. Underwater vehicle sensors are classified into proprioceptive and Visual Sensors (a subset of Exteroceptive Sensors).

ii.1 Proprioceptive sensors

Proprioceptive sensors, including accelerometers, gyroscopes, and sometimes magnetometers, provide critical insights into a system's internal state and motion, measuring parameters like position, velocity, acceleration, and orientation. Widely applied in robotics and navigation, these sensors bolster awareness of the device's movements, enabling precise control. Commonly used proprioceptive sensors, such as the depth sensor, Doppler velocity log, inertial measurement unit, and compass, utilize various technologies, including acoustic measurements and magnetic field detection [3,12,13]. This diverse sensor suite ensures accurate measurements of parameters, such as

depth, velocity, acceleration, and orientation, further enhancing awareness for precise control in applications like robotics and navigation.

Compass:

The magnetic compass relies on the magnetic field but is prone to bias. The gyrocompass relies on fast spinning and is unaffected by metal, but it is more expensive.

Pressure Sensors:

Barometers or pressure sensors can be used for depth measurements. They provide essential information for underwater vehicles to determine their depth, aiding navigation, control, and safety.

DVL (Doppler Velocity Log):

Employs acoustic measurements for tracking the seafloor and calculating velocity. Captures Autonomous Underwater Vehicle's (AUV) sway, surge, and heave velocities. Utilizes transmitted acoustic pulses to gauge Doppler shifts from seabed returns [14,15].

IMU (Inertial Motion Unit) Sensors:

Initially developed for aircraft navigation by Ford, IMU sensors now have broad applications, such as in mobile phones and pedometers. The industry standard is MEMS-based IMUs, including those from manufacturers like Analog Devices, EMCORE, Honeywell, and Collins Aerospace. IMUs offer fast data collection and sensitivity but are prone to cumulative errors and have limited runtime. In SLAM, IMUs are often combined with visual and laser sensors to mitigate mistakes by estimating IMU zero bias [16]. To calculate a vehicle's orientation, velocity, and gravitational forces, accelerometers and gyroscopes (sometimes magnetometers) are combined.

Gyroscope:

Measures angular rates using either Ring Laser/Fiber Optic or MEMS technology. The Ring Laser employs mirrors or fiber optic cables to detect angular rates by observing changes in light. MEMS uses an oscillating mass in a spring system, where gyroscope rotation causes a perpendicular Coriolis force on the mass to calculate the angular rate.

Accelerometer:

It measures the force needed to accelerate a proof mass and comes in various designs like a pendulum, Micro-Electro-Mechanical Systems (MEMS), and vibrating beams [17].
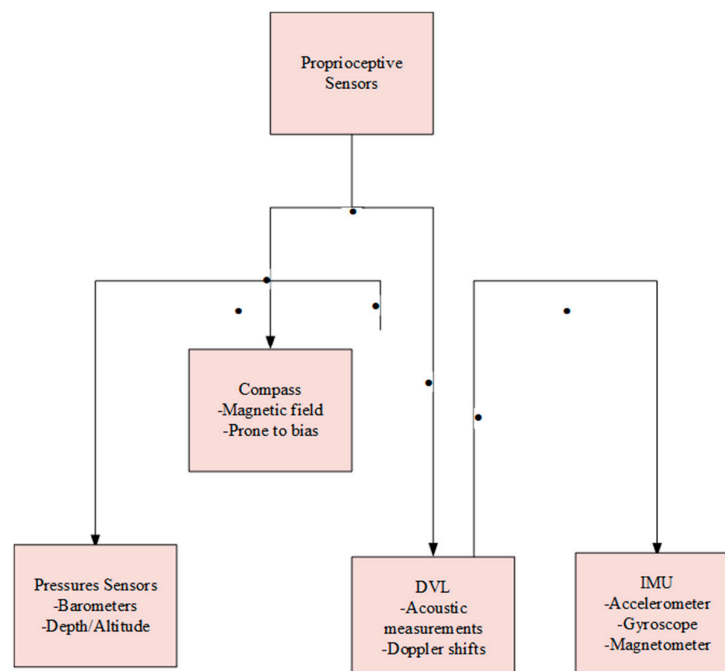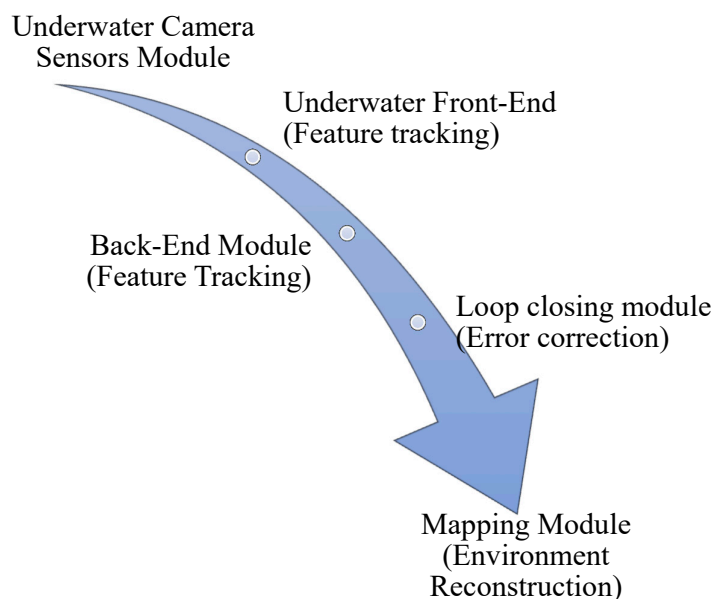

Figure 2. Proprioceptive Sensors Overview.

ii.2 Exteroceptive Sensors

ii.2.1 Visual sensors

Visual SLAM employs cameras to perceive the environment. Mono, stereo, and depth cameras are utilized [18,19]. Visual SLAM employs cameras to perceive the environment. Mono, stereo, and depth cameras are utilized [18,19]. Visual SLAM primarily depends on cameras as exteroceptive sensors to perceive external environmental information. The camera's function is based on optical imaging principles, capturing images using photoreceptors. They come in different types, including monocular, stereo, and depth cameras. SLAM algorithms based on visual inputs are classified into monocular, stereo, and RGB-D categories, contingent on the type of camera employed [20]. Furthermore, specific algorithms, such as ORB-SLAM3, demonstrate adaptability for use with both pinhole and fisheye cameras, broadening their applicability in visual SLAM scenario

Underwater Camera
Sensors Module

Underwater Front-End
(Feature tracking)

Back-End Module
(Feature Tracking)

Loop closing module
(Error correction)

Mapping Module
(Environment
Reconstruction)

Figure 3. The typical architecture of a visual SLAM system.

Monocular or single-lens camera:

Singular lenses and monocular cameras provide economical and straightforward imaging solutions. Mono SLAM pioneered in establishing real-time monocular vision SLAM [21,22]. In underwater research, Hidalgo et al. investigated ORB-SLAM through controlled experiments featuring diverse conditions. Their study confirmed ORB-SLAM's effectiveness under adequate illumination, minimal flicker, and abundant scene features. Monocular Visual Odometry (V.O.) computes relative motion and 3D structure from 2D bearing data, establishing the initial distance between the first two camera poses as one due to the unknown absolute scale. The subsequent processing of images infers the relative scale and camera pose of the initial frames using either 3D structure information or the trifocal tensor [23].

Stereo Camera

Stereo cameras, with a baseline affecting measurement range, calculate depth using parallax [24]. Stereo cameras offer a promising solution for accurate underwater robot localization and proximity operations, as they can calculate distance using parallax, unlike monocular cameras that lack depth information. Researchers have developed innovative methods utilizing stereo cameras [8,25–27], such as a relative SLAM approach that employs a topological metric representation for real-time processing. Additionally, some methods fuse visual and inertial data to eliminate noise and achieve precise underwater robot localization, with less than 3% of typical localization errors. These advancements demonstrate the potential of stereo cameras in enhancing underwater robotics [28].

Depth Camera

Depth or RGB-D cameras utilize structured light or time-of-flight mechanisms to measure distances between the camera and objects. These physical methods reduce computational demands compared to software-based distance estimation used by binocular cameras. However, depth cameras face challenges, including limited measurement ranges, high noise levels, restricted fields of view, susceptibility to sunlight interference, and difficulty in measuring translucent materials due to the characteristics of reflected light [18]. Also known as a depth sensor, it captures color (RGB) and depth information, creating a 3D representation of the environment by measuring distances. Various technologies, such as structured light and time-of-flight, are used in depth cameras. Underwater RGB-D cameras, like Kinect v1 and v2, face limitations due to the weakened infrared light. Modern RGB-D sensors use active stereo technology for robust depth acquisition, which is suitable for various applications. Challenges include sensor errors and semiconductor performance, especially in underwater environments [29,30].
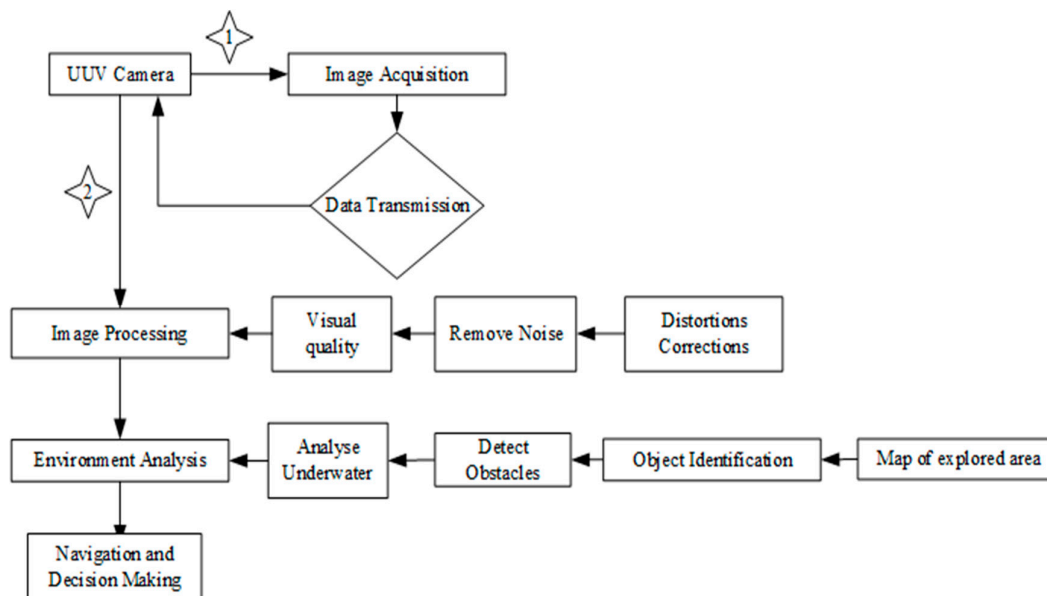

Figure 4. Process of Camera in Underwater Positioning System.

ii.2.2 Sonar Sensors

Sonar sensors, categorized as active and passive, utilize sound waves for underwater detection. Here's a Summary of Sonar Sensor Types and Their Applications in Underwater Environments.

Table 1. Sonar Types and Applications.

| Sonar Type	Description	Reference
Active Sonar	Employed for search and positioning in underwater environments.	[16]
Passive Sonar	Tracks target distance in underwater settings.	[16]
Single-beam Sonar	A single-beam scanning sonar for imaging in low-visibility conditions offers distance information over several meters and is immune to water turbidity.	[31,32]
Multibeam Sonar	Utilizes multiple beams to measure seafloor depth and characteristics rapidly and accurately. Ideal for high-resolution 3D mapping in various underwater applications.	[33,34]
Side-scan Sonar/ forward-looking	They are widely used for detecting underwater objects like wrecks and mines, providing high-resolution acoustic images of seafloor morphology.	[17,35]
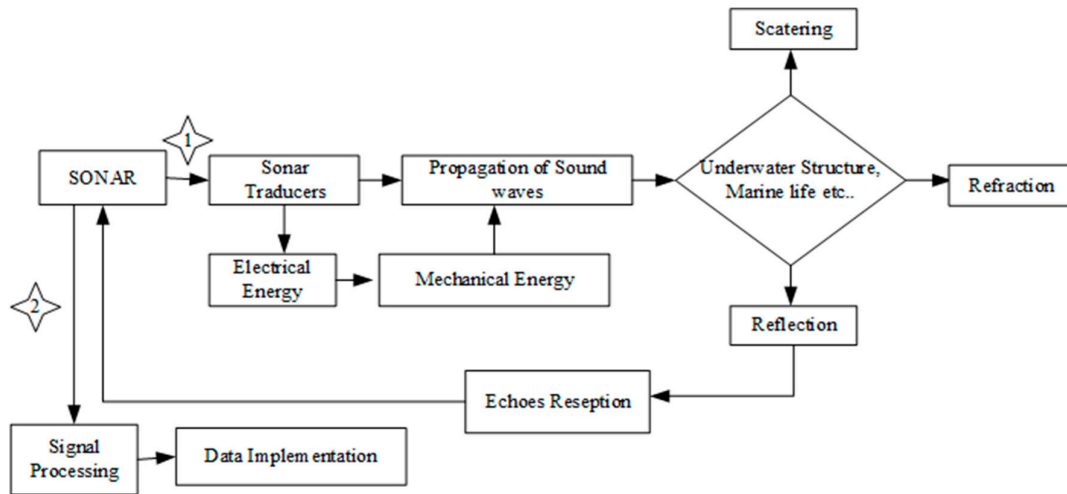
Figure 5. Sonar system process in Underwater navigation.
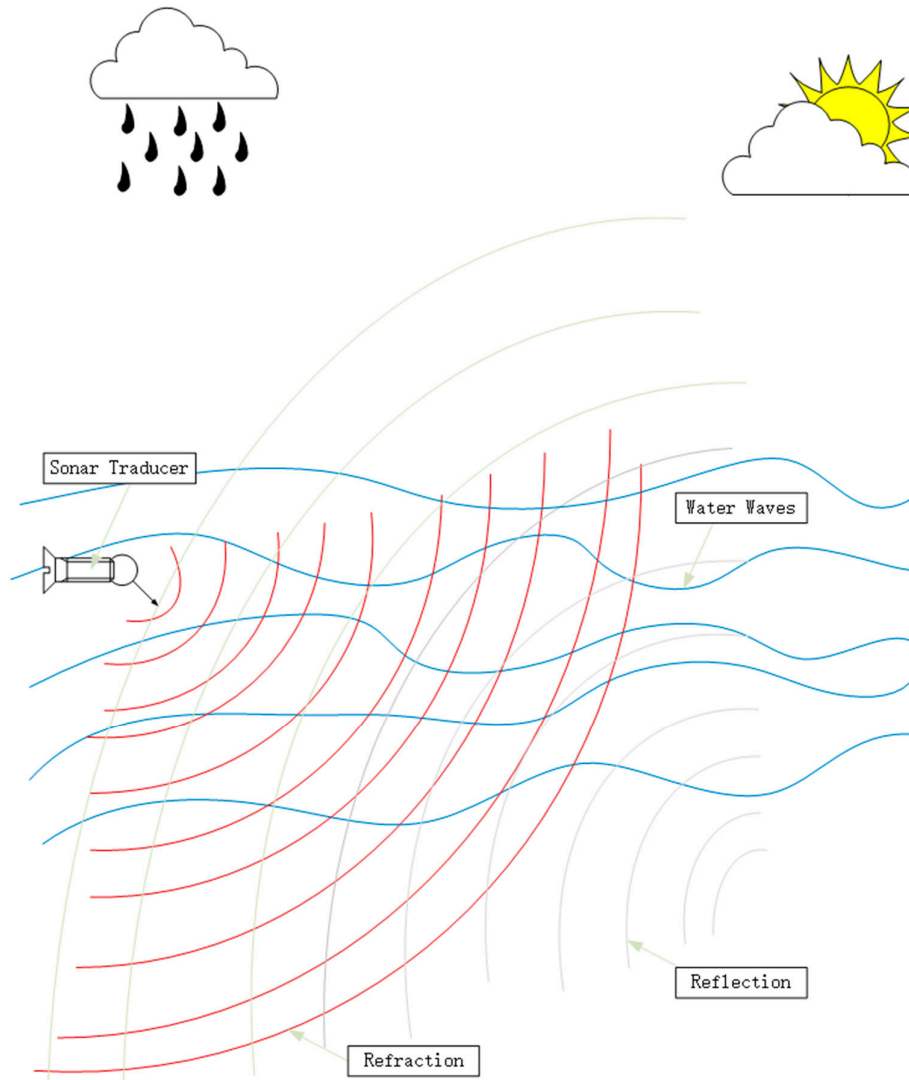

Figure 6. Sonar system underwater.

ii.2.3 LiDAR Technology for Underwater Mapping and Navigation

LiDAR sensors excel in providing accurate and high-frequency range measurements, even in challenging underwater conditions [6,36]. They offer superior 3D data resolution in texture-limited underwater scenes, contributing valuable point cloud data for SLAM systems. LiDAR aids in precise seafloor mapping [37], creating detailed 3D models, and detecting objects to enhance navigational maps. Famous for underwater mapping and navigation, Laser SLAM employs 2D or 3D LiDAR sensors. 2D LiDAR provides real-time obstacle scanning in a single plane, while 3D LiDAR offers high accuracy, comprehensive coverage, and 3D imaging for dynamic and static environments. The critical difference lies in 2D LiDAR lacking height information and imaging capabilities, whereas 3D LiDAR excels in generating three-dimensional real-time images and reconstructing spatial data [16].
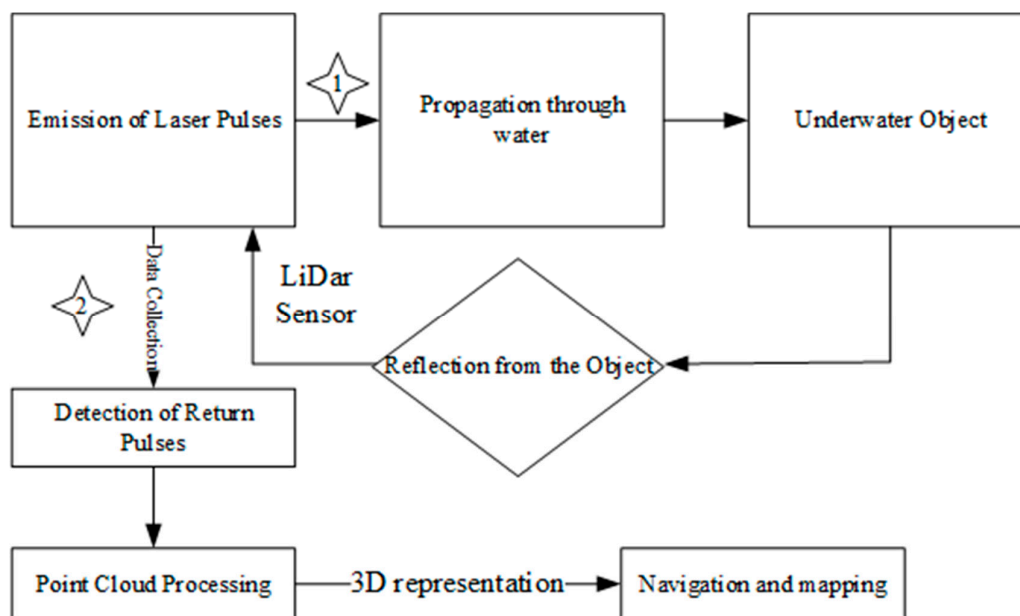

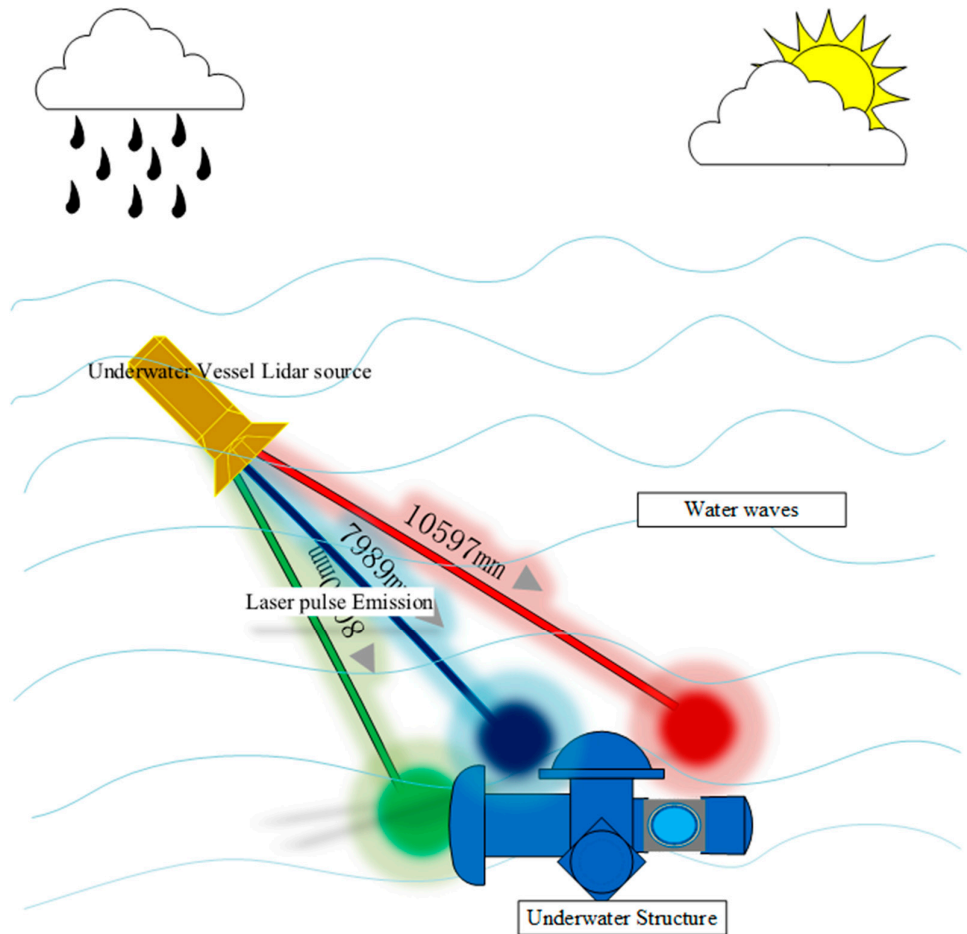**Figure 7.** Process of LiDAR in Underwater navigation and mapping.

Figure 8. Lidar in underwater navigation.
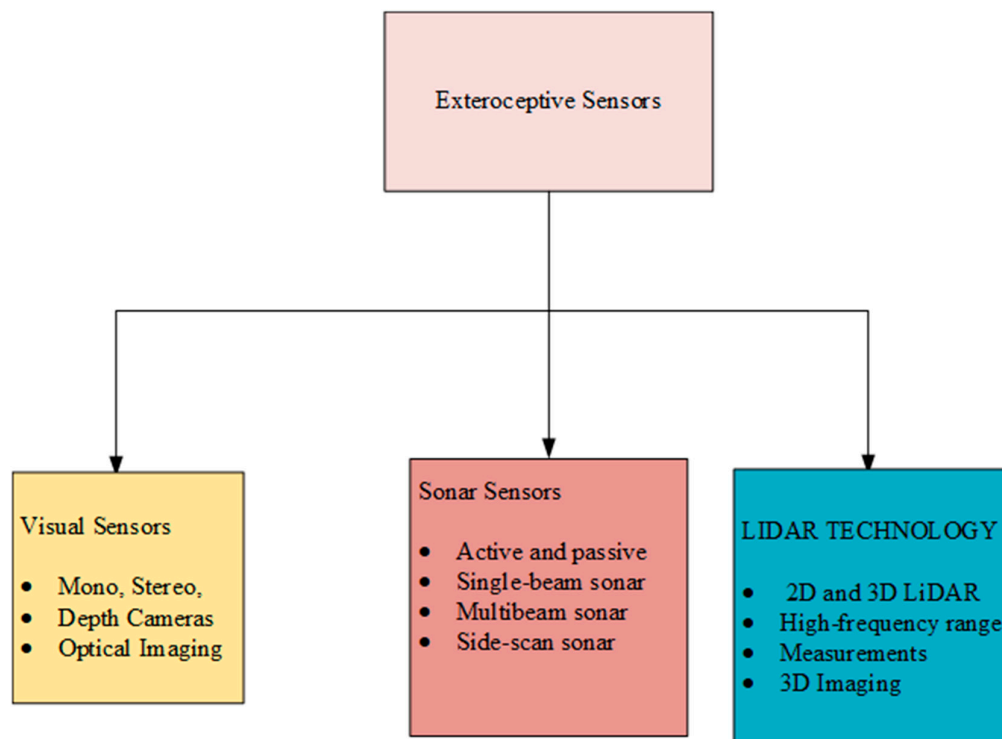
Figure 9. Exteroceptive Sensors Overview.

iii)Trends of visual slam for underwater navigation and mapping

The developments include integrating visual and inertial sensors to enhance accuracy, applying deep learning for robust feature extraction in challenging environments, increasing utilization of 3D vision methodologies for richer depth information, and incorporating adaptive algorithms capable of adjusting to varying underwater conditions. The integration of SLAM in Autonomous Underwater Vehicles (AUVs) for autonomous navigation, the fusion of visual SLAM with underwater LiDAR for comprehensive mapping, and the rising adoption of open-source SLAM frameworks are prominent trends. Additionally, real-time processing, edge computing, collaborative SLAM strategies, and the emergence of standardized benchmarks and datasets contribute to the ongoing efforts to advance the capabilities of underwater SLAM systems, shaping the landscape of autonomous underwater exploration and research.

Figure 10. Trends in visual SLAM for UUV.

iii)   Objective and contribution of our paper

The primary objective of our paper is to investigate the innovative aspects of underwater Simultaneous Localization and Mapping (SLAM) or odometry, focus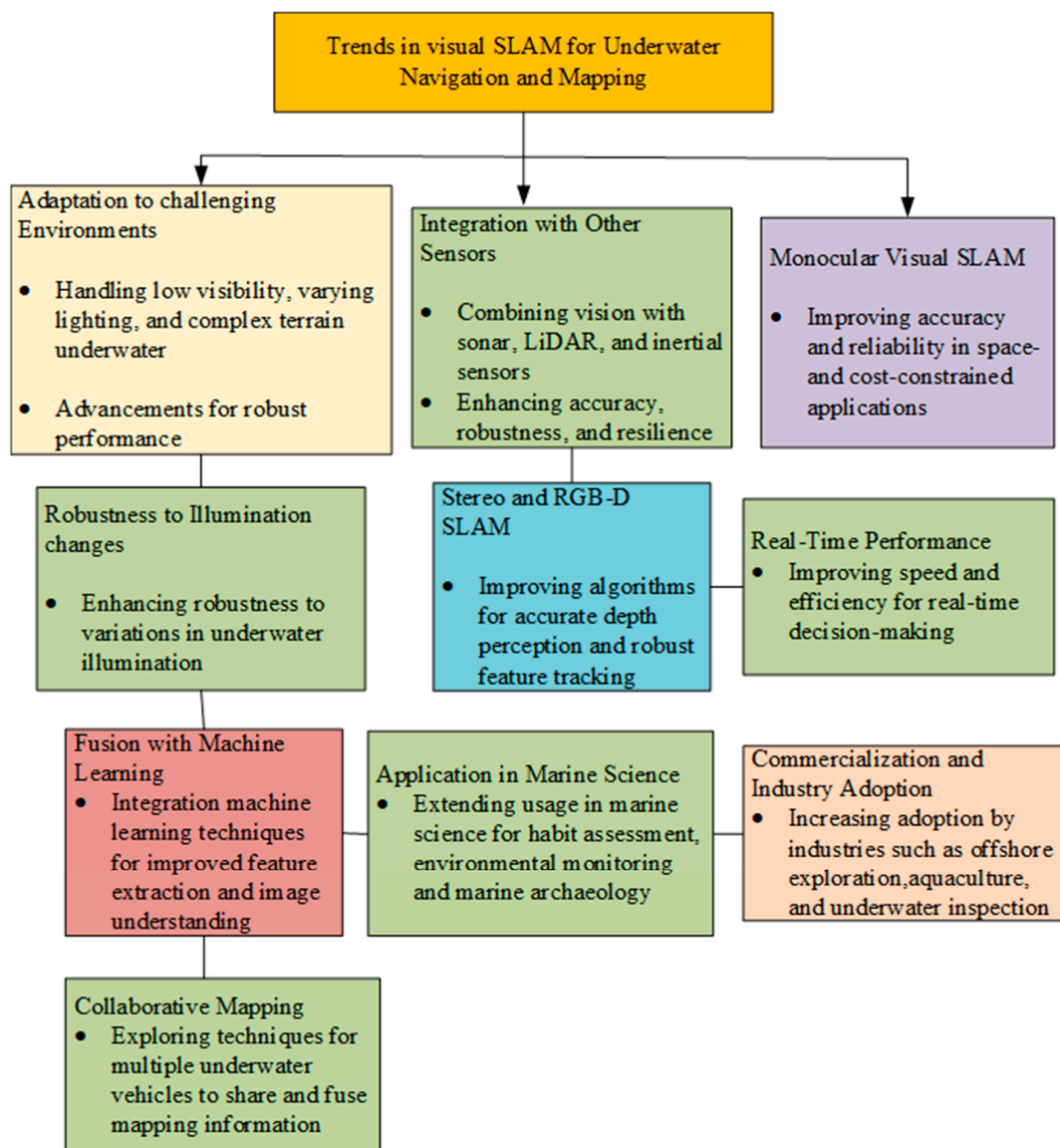ing on integrating multiple sensors and applying deep learning techniques. Specifically, we aim to identify and assess the advancements in SLAM and odometry methodologies that result from integrating various sensors, emphasizing the synergy among them to enhance accuracy and robustness. Additionally, our research delves into the role of deep learning in these underwater systems, evaluating how it contributes to feature extraction, mapping, and navigation in challenging underwater environments. By highlighting these novel approaches, we seek to contribute to the evolving field of underwater robotics and exploration, providing valuable insights for researchers and practitioners working on underwater SLAM and odometry systems.

The paper's layout follows Session 1, which delves into the comprehensive description of standard visual SLAM algorithms' performances in underwater applications. In Session 2, we summarise papers that involve multiple sensor integration in SLAM odometry, highlighting their strengths and weaknesses. Shifting the focus to deep learning techniques, Session 3 offers a summary of the applications of deep learning in underwater image processing, navigation, and perception. Additionally, a detailed analysis of their performance and future potential is provided. Session 4 summarises papers involving deep learning-based underwater SLAM or odometry navigation, presenting their respective strengths and weaknesses. Session 5 compiles a list of commonly used datasets for evaluating underwater SLAM algorithms. Lastly, in Session 6, we present our prediction development directions based on the abovementioned content.

II. Common Underwater SLAM Advancements and Algorithm Performance

SLAM has undergone significant advancements, particularly with the Extended Kalman Filter (EKF) SLAM, commonly used for probabilistic robot pose and landmark estimation. Challenges arise in complex, nonlinear environments due to their linearization assumptions, prompting the utilization of the Square Root Information Filter (SRIF) algorithm. SRIF enhances stability, numerical reliability, and operational efficiency for underwater Unmanned Underwater Vehicle (UUV) navigation by effectively managing the covariance matrix and addressing numerical instability.

Diverse SLAM techniques cater to varied challenges in different environments. Graph-based SLAM optimizes a graph representation for configuration identification, Particle Filter SLAM excels in highly nonlinear scenarios, RBPF SLAM enhances computational efficiency, and Bayesian SLAM adopts a Bayesian estimation perspective. In the underwater domain, addressing sensor limitations is crucial, with innovative solutions like sonar-based mapping and loop closure detection playing vital roles. Deep learning integration enhances SLAM solutions, utilizing neural networks to improve mapping accuracy and navigation efficiency in challenging underwater environments, particularly when coupled with lidar and vision sensor advancements.

Various Visual SLAM algorithms have been evaluated for underwater applications. ORB-SLAM offers robust performance in well-lit, feature-rich environments, while ROVIO excels in dynamic underwater settings by leveraging visual and inertial data. LSD-SLAM efficiently maps large, texture-rich underwater environments, and DVO-SLAM handles depth changes effectively. MSCKF combines visual and inertial measurements for precise navigation, while SVO suits lightweight underwater vehicles with real-time efficiency. VISLAM algorithms enhanced with deep learning handle challenging underwater conditions, aiding in feature extraction and mapping. FAB-MAP, adapted for underwater scenarios, excels in environments with distinctive visual features. The choice of algorithm depends on specific environmental conditions and mission requirements, with ongoing research contributing to advancements in the field.

III. Multiple Sensor Integration in Slams Odometry: Strengths and Weaknesses

The accuracy and resilience of underwater SLAM systems through sensor fusion techniques encompasses vision-inertial SLAM, laser-vision SLAM, and multisensor SLAM [37,38]. Multisensor fusion is classified into data layer, feature layer, and decision layer fusion [39,40]. Visual SLAM faces challenges with low-quality images, while IMU-assisted sensors improve

To enhance the accuracy and robustness of underwater SLAM systems, researchers often combine multiple sensors, leveraging sensor fusion techniques. This fusion approach results in a more precise and resilient underwater SLAM. Standard fusion methods include vision-inertial SLAM (utilizing vision and IMU), laser-vision SLAM (combining laser and vision), and multisensor SLAM (incorporating sonar, IMU, vision, etc.). Multisensor fusion can be categorized based on fusion level into data, feature, and decision layer fusion. The complexity of coupling can further be divided into loosely coupled, tightly coupled, and ultralight coupled systems [28]. Visual SLAM algorithms have advanced significantly but struggle with low-quality images caused by rapid camera movements and varying light conditions. IMU-assisted sensors offer improved angular velocity and local position accuracy compared to odometers. They complement each other, with IMUs providing clear images of dynamic objects during fast camera movements and cameras correcting IMU-generated cumulative errors during slower movements. This combination enhances SLAM performance and is cost-effective. Visual-inertial fusion methods can be loosely coupled. IMU and camera motions are estimated separately and then fused or tightly coupled, involving joint construction of motion and observation equations before state estimation [41].

A novel crack assessment technique combines multisensor fusion SLAM and image super-resolution [42]. A modality prediction approach is explored using LiDAR point cloud prediction from 3D acoustic ultrasonic sensor data [39]. Fusion SLAM algorithms also combine 2D LiDAR and RGB-D SLAM, offering a comprehensive visual representation [43]. Considerations in sensor fusion systems include the fusion objective and sensor constraints [44]. An innovative self-localization system uses low-cost sensors and an Extended Kalman Filter [74]. Multi-beam sonar is employed for underwater landmark detection, and an AUV utilizes tightly coupled lidar-visual-inertial SLAM [33].
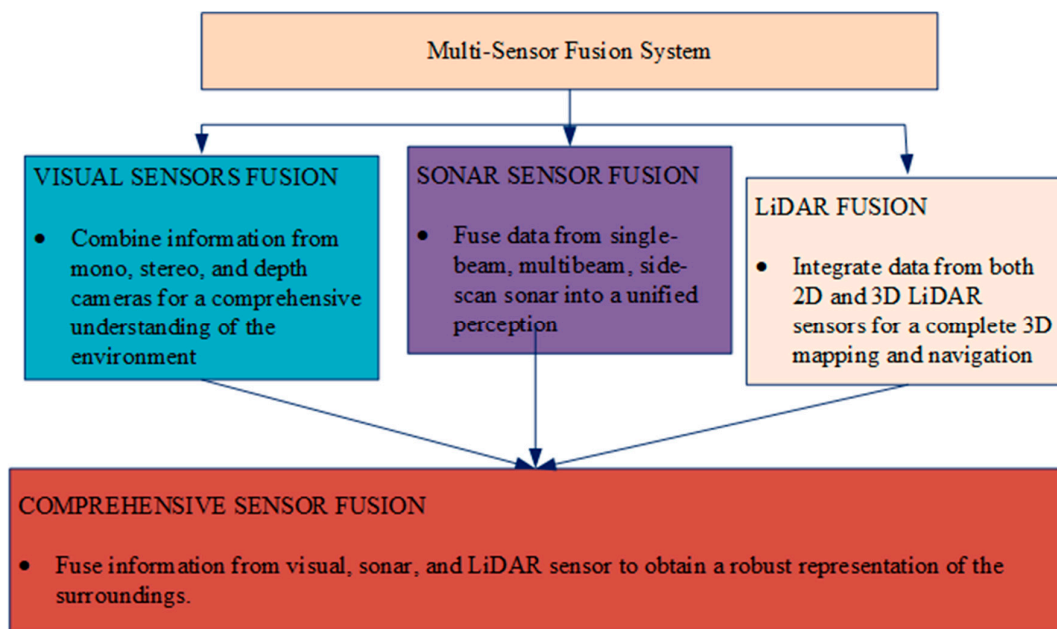
Figure 11. A multisensor fusion overview.

Visual SLAM faces challenges with low-quality images, while IMU-assisted sensors improve accuracy. Visual-inertial fusion methods can be loosely or tightly coupled [41]. Incorporating modalities such as sonar or radar poses a challenge since existing methods are often specialized for conventional sensors. Da Bin Jeon et al. present a Lie theory approach for unmanned underwater vehicle navigation, addressing misalignment issues through Lie algebra operations. It enhances estimation accuracy, stability, and convergence of covariance. However, potential weaknesses may include challenges in real-world implementation, computational complexity, and the need for thorough validation in diverse operational scenarios. Ongoing research aims to refine and validate

the method for broader applicability and to address any identified limitations. Sensor Fusion for Underwater Vehicle Navigation Compensating Misalignment Using Lie Theory.

SVIn2 presents an advanced underwater SLAM system, integrating diverse sensors(Sonar, Visual, Inertial, and water-pressure information) for robust performance in challenging environments. The real-time framework overcomes traditional weaknesses, demonstrating exceptional accuracy and reliability in benchmark datasets and real-world scenarios. However, potential sensor dependencies and generalization across diverse underwater settings pose considerations for further exploration [45]. Chunying Li et al. introduce an innovative Multi-Source Information Fusion (MSIF) model for Spherical Underwater Robots (SURs), enhancing precision and addressing critical issues in Autonomous Underwater Vehicles (AUVs). However, reliance on low-cost sensors may impact accuracy, and performance could vary based on environmental conditions.

Further refinement is needed for robustness in diverse scenarios and adaptability to varying sensor qualities [46]. Researchers proposed a cost-efficient and precise solution for underwater pipeline inspection utilizing an Autonomous Underwater Vehicle (AUV). Successfully navigating the pipeline with minimal sensors, the system exhibits robust performance under varying current velocities, incorporating fuzzy logic for enhanced stability. The ROS/Gazebo-based simulation environment facilitates efficient development and testing. However, challenges in visibility variations and obstacles require further refinement, suggesting potential enhancements through expanding the sensor fusion framework and integrating adaptive parameters in image processing. Future research directions include addressing dynamic surface wave effects through real-world experiments, with consideration given to a down-scaled AUV for pool testing [40]. Di Wang et al. introduce a multisensor fusion method for underwater integrated navigation systems, focusing on SINS/DVL/USBL. It addresses frame system inconsistencies due to velocity errors, demonstrating enhanced accuracy, especially in scenarios with long-distance USBL signal challenges. However, further validation in diverse underwater environments is needed to establish its broader applicability and reliability [47].

IV. Deep Learning Techniques Applied in Underwater Image Processing, Navigation, and Perception, along with Their Performance and Future Potential

Deep learning techniques represent a revolutionary paradigm in machine learning, characterized by using neural networks with multiple layers to model and interpret complex patterns in data. Convolutional Neural Networks (CNNs) excel in image-related tasks, capturing hierarchical features for image recognition and computer vision applications. Recurrent Neural Networks (RNNs) are pivotal in processing sequential data, such as language and time-series information, owing to their ability to retain context and dependencies. Transfer learning strategies leverage pre-trained models to boost performance on specific tasks, facilitating effective knowledge transfer. Generative Adversarial Networks (GANs) introduce a novel approach to realistic data generation. At the same time, advanced natural language processing models like BERT and GPT showcase remarkable capabilities in understanding and generating human-like language.

Deep learning techniques have emerged as powerful tools in underwater applications across image processing, navigation, and perception domains. Convolutional Neural Networks (CNNs) have been extensively applied in underwater image processing, showcasing remarkable proficiency in object detection, recognition, and segmentation tasks. Their performance is notable for robust feature extraction in challenging underwater environments, enhancing the accuracy of visual perception systems. Transfer learning, particularly with pre-trained CNN models, has effectively overcome data scarcity issues, yielding promising results in various underwater scenarios. Additionally, Recurrent Neural Networks (RNNs) contribute to navigation tasks by processing sequential data, aiding in trajectory prediction and underwater vehicle control. The fusion of sensor data, including acoustic and visual inputs, using deep learning architectures enhances perception capabilities, allowing for more accurate mapping and environmental understanding; despite advancements, challenges persist, such as limited labeled underwater datasets and the need for real-time processing.

The future potential of deep learning in underwater applications hinges on several critical areas of development. Refining model architectures is crucial for greater precision and adaptability to diverse underwater environments. Addressing domain adaptation challenges will further enhance the robustness of these techniques across varying conditions. Optimizing algorithms for efficiency ensures real-time processing capabilities while expanding the application scope, which broadens the utility of deep learning in underwater systems. Additionally, exploring unsupervised learning methods holds promise for advancing autonomy in underwater applications. Continued research and innovation in these areas are essential to unlock further advancements in underwater image processing, navigation, and perception applications.

V. V. Deep Learning-Based Underwater SLAM and Odometry Navigation: Strengths and Weaknesses

Jayashree Rajesh et al. propose a system to properly recognize and classify underwater life and objects in underwater images. It provides a new way to identify and categorize many classes, broadening the model's uses. The study uses statistical, hardware, and software solutions, focusing on deep learning approaches like YOLOv4 and CNN to accurately classify underwater objects [48]. ANWAR KHAN et al. review recent underwater target detection algorithms for wireless sensor networks. It categorizes and assesses these algorithms, discussing their applications, strengths, and weaknesses. A comparative analysis and trend evaluation over the last decade is provided [49]. Ali Khandouzi et al. use deep learning and classical image processing to enhance underwater images. A lightweight colour retrieval network updated histogram equalization for contrast improvement, and an attention module for synergistic integration comprise the three-module framework. The approach solves underwater picture problems with minimum computational load. However, potential drawbacks include picture augmentation, dataset-specific effectiveness, algorithm complexity, generalization across various contexts, and overfitting to specific conditions [50]. A CNN and intensity changes improve underwater image quality in two steps, according to Laura A. Martinho1 et al. 2024. The approach performs well on various datasets, including a new Amazon dataset. Effective deep learning-based augmentation and dataset development are strengths. Lack of extensive comparison insights and explicit future inquiry objectives are weaknesses [51]. A deep learning model for underwater image restoration, the Combining Attention and Brightness Adjustment Network (CABA-Net), mitigates colour-cast, low brightness, and low contrast. Ablation investigations confirm the efficiency of individual network components, and the approach adapts to underwater settings, boosting image contrast and color and aligning with human visual system properties. However, the study lacks discussions on constraints, computing efficiency, and broader applicability [52]. A unique integrated system for underwater object and temporal signal detection employing 3D integral imagery in degraded settings is proposed. Deep learning improves 2D imaging performance. 3D integral imaging improves image reconstruction and segmentation, enhancing detection accuracy. The method's weaknesses include undiscovered computational complexity and color distortion. Future directions include optimal configuration research and resolving issues in increasingly complicated underwater environments [53]. Researchers use convolutional neural networks (CNN) and recurrent CNN to estimate ego-motion in autonomous underwater robots' forward-looking sonar (FLS). Both models can learn from synthetic and field data, but the recurrent model predicts synthetic data better. FLS sensor configurations imaging terrains need further study, according to the study. It advises using larger field datasets and diverse sensor features to improve model performance [54]. Yelena Randall1 et al. present unique forward-looking underwater stereo-vision and visual-inertial datasets essential for testing autonomous systems and algorithms in challenging underwater conditions. The datasets cover various scenarios, providing synchronized images, ground truth depth maps, calibrations, and known object measurements [55]. An autonomous underwater vehicle (AUV) with an intelligence system recognizes and tracks underwater objects. Semi-Global Block Matching (SGBM) methods forecast depth maps, and Deep Q-Network (DQN) localizes disparity map objects. The system detects objects using a Faster Region-based Convolutional Neural Network (R-CNN). DQN optimization of SGBM parameters, 3D point

cloud images for object information calculation, object 3D information convergence with increasing learning episodes, and wave height's effect on object size estimation and AUV maneuvering performance are notable findings [56]. A new neural network uses autoencoder architecture and SIFT-based descriptors to detect underwater visual loops quickly and reliably. Its unsupervised training method beats others, making it suited for AUVs with limited computational resources [57]. The underwater visual simultaneous localization and mapping (VSLAM) system ULL-SLAM was developed by Zhichao Xin et al. to handle low-light problems. The model's end-to-end design includes a low-light enhancement branch with a non-reference loss function, allowing image augmentation without paired low-light data. A self-supervised feature point detector and descriptor extraction branch improves matching without pseudo-ground truth. The suggested method ensures trajectory continuity, stability, and accurate state estimation under demanding underwater environments to improve VSLAM performance. The research mentions features including better feature point extraction in low-light circumstances but does not examine limitations, computing efficiency, or the approach's generalizability to varied underwater exploration settings [58]. Researchers suggest computer vision-based AUV position estimates to alleviate navigation errors. The method uses deep learning and computer vision to analyze real-time environmental photos to a Digital Surface Model map. The approach lowers positioning errors (30–60 m) and works with incomplete land representations. It can extract land features accurately, reduce dead reckoning errors, and adapt to difficult sea situations. The technique could enable fully autonomous AUV navigation in GNSS-denied conditions, improving low-cost AUV technology [59]. A comprehensive dataset from a controllable AUV with high-precision fiber-optic inertial sensors, a Doppler Velocity Log (DVL), and depth sensors by Can Wang et al. advances autonomous underwater vehicle (AUV) navigation. The dataset includes numerous natural scenarios from multiple locations and timelines, both beneath and on the surface, to address the lack of publically accessible data for training machine learning algorithms in underwater navigation. Rigorous testing and algorithmic evaluations of real and calculated positions prove the dataset's usefulness. Limitations and use cases of the dataset are not discussed in the study. Its influence on autonomous exploration in limited underwater habitats needs additional study [60].
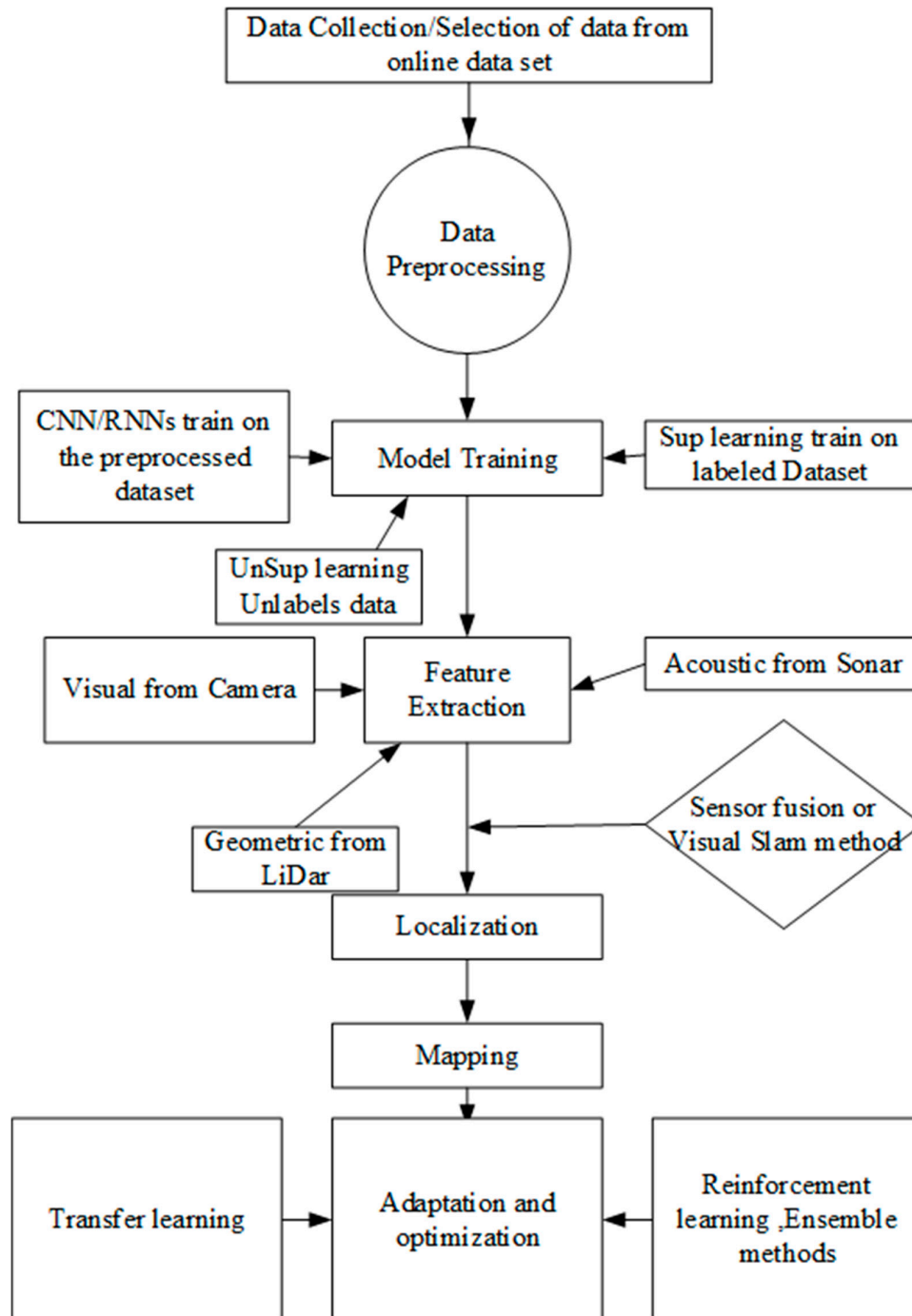
Figure 12. Process of deep learning in underwater SLAM.

An improved visual-inertial odometry system called Semantic SLAM uses semantic characteristics from an RGB-D sensor to improve camera localization in Visual Simultaneous Localization and Mapping (VSLAM). It excels in indoor conditions with little camera input and is scene-agnostic. A convolutional short-term (ConvLSTM) network refines the semantic map, improving pose estimation by 17% over VSLAM. The semantic map provides interpretable information for robot navigation tasks, including path planning and obstacle avoidance. The public code shows that semantic aspects in SLAM systems are feasible [65]. SplaTAM, a pioneering SLAM system for a single unposed monocular RGB-D camera, uses a 3D Gaussian Splatting radiance field for map representation. It suggests ways for more advanced and efficient SLAM systems to analyze

scenes [61]. GO-SLAM, a real-time dense visual SLAM system, optimizes camera postures and 3D reconstruction using neural implicit representations. It outperforms state-of-the-art algorithms in robust posture prediction, loop closing, and online total bundle adjustment. The versatile approach supports monocular, stereo, and RGB-D inputs and dynamically changes the continuous surface representation for global consistency. It excels on varied datasets of lengthy monocular trajectories without depth information [67]. An uncertainty learning method for dense neural SLAM estimates pixel-wise depth uncertainties without ground truth data, improving mapping and tracking accuracy. The method outperforms alternatives on many datasets, demonstrating its multisensor input flexibility [62]. Researchers introduced NICE-SLAM, a dense visual SLAM system that improves scalability, efficiency, and resilience by combining neural implicit representations with hierarchical grid-based scene representation. The method enhances mapping detail, tracking accuracy, and speed with less processing. NICE-SLAM outperforms neural implicit SLAM methods in mapping and tracking tough datasets without over-smoothing [63]. Point-SLAM, a dense neural SLAM system for monocular RGBD input, uses a dynamically produced neural point cloud to adapt density to input information. It performs better than existing tracking, mapping, and rendering algorithms on numerous datasets, improving resource utilization and 3D scene representation accuracy [64].

Table 2. Deep Learning-Based Underwater SLAM Strengths and Weaknesses.

Methods	Strength	Weaknesses/limitations	Framework	Code available	Year
Semantic SLAM	The SemanticSLAM system introduces innovation with scene-agnostic functionality across diverse environments, constructing a semantic map for interpretability and utilizing a ConvLSTM network to correct errors and enhance pose estimation.	The paper exhibits limitations, including a restricted performance evaluation, sparse details on system implementation, unclear generalization to outdoor environments, and a lack of concrete plans to address identified limitations in future work.	Pytorch	Yes	2024
SplaTAM	SplaTAM demonstrates remarkable performance, achieving up to 2× state-of-the-art results in camera pose estimation and scene reconstruction, leveraging an innovative 3D Gaussian Splatting representation for fast rendering, optimization, and explicit spatial awareness in a single unposed monocular RGB-D camera setup with structured map expansion capabilities.	The paper lacks comprehensive insights into the system's generalization across diverse environments, fails to address the computational requirements for real-time applications thoroughly, relies on assumptions about the universal suitability of Gaussian Splatting, and would benefit from a more in-depth comparative analysis with existing SLAM methods to enhance credibility.	Pytorch	Yes	2023
GO-SLAM	GO-SLAM introduces global optimization for camera poses and 3D reconstruction, ensuring improved tracking and accuracy across versatile inputs, including monocular, stereo, and RGB-D setups, while maintaining real-time performance for dynamic environments and continuous adaptation for global consistency.	GO-SLAM exhibits potential concerns, including the risk of error accumulation over time in complex scenarios, challenges on resource-constrained devices due to computational demands, potential hindrance in understanding and implementation due to algorithmic complexity, and the need for further investigation into its performance under highly variable real-world conditions.	Pytorch	Yes	2023

UncLe-SLAM	Innovative uncertainty learning for dense neural SLAM demonstrated performance improvement adaptability to multisensor inputs.	Limited depth sensor comparison, reliance on self-supervised training.	Pytorch	2023
NICE-SLAM:	NICE-SLAM excels with a hierarchical scene representation, ensuring detailed reconstruction and scalability. It achieves efficiency, competitive mapping, and tracking quality, overcoming over-smoothing challenges. The model adeptly fills small holes, extrapolates scene geometry, and benefits from geometric priors for enhanced reconstruction in large indoor scenes.	The method's predictive capability is confined to the scale of the coarse representation, and loop closures are not currently incorporated. Exploring loop closures presents an intriguing avenue for future research. While traditional methods lack certain features, a gap remains to be bridged with learning-based approaches.	Anaconda. YES	2022

Table 3. List commonly used data sets to evaluate underwater slams.

MARAS Dataset	It was collected in the Mediterranean Sea, providing acoustic, visual, and inertial sensing data.	MARAS Dataset	MARAS: A Dataset for Marine Robot Assistance Systems
UW-ETH-ASL Dataset	Captured in various environments, this dataset from ETH Zurich includes RGB-D data and ground truth for benchmarking visual and inertial SLAM algorithms.	UW-ETH-ASL Dataset	NICE-SLAM
SAUVC Dataset	The Singapore AUV Challenge (SAUVC) dataset, collected in swimming pool conditions, includes visual and inertial sensor data.	SAUVC Dataset	SLAM-Based Navigation for an AUV in Indoor Pools
URB Dataset	Developed by the UUST (Underwater Robotics and Imaging) group, URB provides datasets for visual and acoustic SLAM under challenging conditions.	URB Dataset	Multi-Modal Underwater Simultaneous Localization and Mapping
LIU-UW Dataset	The Linköping University Underwater (LIU-UW) dataset includes data from various underwater environments, providing visual and inertial measurements.	LIU-UW Dataset	GraphSLAM for Underwater 3D Reconstruction with Stereo Camera and Inertial Measurement Unit
UW3D Dataset	A benchmark dataset for underwater 3D reconstruction, UW3D includes RGB-D images and is designed to evaluate SLAM algorithms.	UW3D Dataset	An Underwater Stereo Camera System for 3D Reconstruction and Object Identification
ScanNet	ScanNet is a dataset of annotated 3D reconstructions of indoor scenes used for research in scene understanding and robotics.	ScanNet Dataset	GO-SLAM, NICE-SLAM
Replica	Dataset featuring photorealistic 3D reconstructions of indoor scenes, commonly used for research in computer vision and virtual reality.	Replica dataset	UncLe-SLAM, NICE-SLAM
TUM-RGBD	A dataset containing synchronized RGB and depth images captured from indoor scenes is often used to research visual SLAM		Point-SLAM

(Simultaneous Localization and Mapping), 3D reconstruction, and scene understanding.

NICE-SLAM: Neural Implicit Scalable Encoding for SLAM. FIDCE: Filter-Guided Inverse Dark Channel Inversion Exposure Compensation/ MFONet: MobileNetV2 Feature Extraction Network.

VI. Advantage of Deep Learning Relative to the Conventional Method

Deep learning, a transformative paradigm in artificial intelligence, brings several advantages to underwater Simultaneous Localization and Mapping (SLAM) navigation compared to conventional methods. One significant strength is its ability to comprehend complex underwater scenes, as neural networks excel at discerning intricate patterns within underwater data, providing a nuanced understanding of the environment, crucial for navigating challenging conditions characterized by low visibility or uneven terrain [65,66]. Matias Valdenegro-Toro1 introduces a CNN-based approach for accurate sonar image matching in AUV applications, outperforming traditional methods. The study anticipates improvements with more significant, diverse datasets. Despite constraints, the proposed method holds promise for enhancing AUV perception, with future work aiming to develop unsupervised learning for sonar image similarity functions [67]. Researchers introduce an underwater loop-closure detection method using an unsupervised UVAE network, achieving a 92.31% recall rate in dynamic underwater scenarios. It addresses challenges like viewpoint changes, textureless images, and fast-moving objects. The system includes semantic object segmentation and an image description scheme for efficient information access. Real-world testing demonstrates robustness and real-time performance. Future work aims to enhance accuracy in complex underwater environments and explore decentralized visual SLAM for multiple AUVs in more significant scenarios [68]. Bryan Pedraza and Dimah Dera present a Bayesian Actor-Critic (A2C) reinforcement learning approach for robust simultaneous localization and mapping (SLAM) in noisy environments. Leveraging Bayesian inference, the model generates robot actions while quantifying uncertainty. The proposed framework has broad applications in underwater robots, biomedical devices, micro-robots, and drones, emphasizing its adaptability and reliability in uncertain environments [69]. Researchers proposed an article that assesses Visual Odometry (V.O.) in challenging underwater conditions, comparing classical and deep learning methods. Traditional systems struggle with initialization and tracking, while deep learning architectures exhibit superior performance, providing continuous pose estimation in complex scenarios. The study emphasizes the potential of data-driven approaches for robust underwater robot navigation [70]. Zhengyu Xing et al. introduce an enhanced underwater image enhancement model based on ShallowUWnet, utilizing convolutional blocks, batch normalization, and LeakyReLU activation. The model, incorporating various loss functions, outperforms advanced methods in evaluation metrics, showcasing superior performance and generalization. Practical testing on engineering cases highlights its effectiveness, offering a faster processing alternative to deep neural network methods for underwater image enhancement [71]. In the dynamic realm of underwater navigation, deep learning's holistic approach stands out. Traditional SLAM systems often involve separate modules for localization and mapping, requiring intricate integration. Deep learning models employ end-to-end learning, enabling the system to grasp its location and construct a map simultaneously, simplifying the navigation process for more efficient underwater exploration [72,73]. Another advantage is the flexibility of deep learning in handling different sensors commonly used in underwater navigation, such as sonar, LiDAR, and cameras. Deep learning models can seamlessly integrate information from these diverse sensors, learning to interpret varied data sources coherently. This adaptability contrasts with traditional SLAM systems, which may require complex calibration and synchronization processes when dealing with multiple sensors [74,75]. Thanks to neural networks' ability to model nonlinear relationships and adapt to dynamic changes, deep learning's prowess becomes evident in navigating tricky underwater situations. This makes them well-suited for handling unpredictable underwater scenarios, where traditional methods may struggle without sophisticated filtering techniques [76]. Transfer learning, a key feature of deep learning, introduces another layer of advantage. Pre-trained models can be adapted for underwater navigation, where obtaining large labeled datasets can be challenging. This

capability significantly accelerates the training process, allowing for quicker deployment of models in real-world underwater exploration scenarios [77]. Moreover, the ability to effectively fuse information from different sensors is a distinctive strength of deep learning [74]. Deep learning models can harmoniously integrate these disparate data sources in underwater environments, where a combination of sonar, LiDAR, and optical sensors is common. Traditional SLAM systems may face challenges in achieving such seamless integration, requiring intricate adjustments and coordination. While acknowledging these advantages, it's essential to consider practical factors such as computational requirements and interpretability. Deep learning's computational demands can be significant, and the 'black-box' nature of neural networks may raise interpretability concerns. Nevertheless, the suite of advantages presented by deep learning positions it as a transformative force in advancing the capabilities of underwater SLAM navigation, offering a promising avenue for further exploration and research in this dynamic field. Integrating Deep Learning methods into underwater navigation represents a revolutionary stride in enhancing the capabilities of Unmanned Underwater Vehicles (UUVs). Under the umbrella of Artificial Intelligence (AI), these methods enable UUVs to delve deeper into the intricacies of underwater environments through sophisticated data processing. Deep Learning-based SLAM algorithms [78] empower UUVs with advanced cognitive abilities to make real-time decisions, adapt to dynamic underwater conditions, and navigate with unparalleled accuracy [39]. Deep Learning algorithms excel at extracting intricate patterns and representations from sensor data, encompassing visual, inertial, and acoustic inputs. This enables UUVs to construct highly detailed maps of their surroundings and concurrently estimate their precise positions within the underwater landscape. The adaptive learning capabilities of Deep Learning methods empower UUVs to continually refine their navigation strategies based on accumulated experiences [8]. The fusion of Deep Learning methodologies with UUVs elevates the accuracy and reliability of underwater operations. It propels these vehicles to new frontiers of exploration and research in the marine domain [4]. It positions UUVs as intelligent entities capable of autonomously navigating through challenging underwater terrains, leveraging the power of advanced neural network architectures for unparalleled adaptability and performance. Researchers have demonstrated the application of deep neural networks [78] to predict interframe poses, replacing traditional visual odometry. A keypoint rejection system is used to supervise neural network training, improving the reliability of visual ego-motion estimation by filtering out unsuitable vital points [43]. Dr J. Priscilla Sasi et al. explore Convolutional Neural Networks (CNNs) in autonomous underwater robot navigation, emphasizing their effectiveness in overcoming challenges like low visibility and object detection. Case studies showcase CNNs' potential for transforming underwater robotics, highlighting the need for ongoing research to enhance adaptability in challenging environments [79]. Olaya Alvarez-Tu' n˜on et al. survey the landscape of visual simultaneous localization and mapping (SLAM) algorithms in geometry-based and learning-based frameworks. It introduces a comprehensive SLAM pipeline formulation, categorizes implementations, and evaluates their performance under varying environmental challenges. The study emphasizes the shift towards end-to-end pipelines driven by deep learning, addressing efficiency limitations and the need for generalizability in diverse deployment conditions. The findings highlight the potential of merging geometry and learning-based approaches for future advancements in visual SLAM [80]. Self-organizing maps (SOMs), another neural network approach, are employed for multi-robot SLAM, offering unsupervised training capabilities [17]. A refined super-resolution reconstruction method enhances and recovers underwater images by decomposing RGB attenuation to calculate transmission maps and improve image quality [78]. All these methods show how deep learning SLAM is in the field of underwater SLAM. Below are some practical examples implemented in the domain, starting with a proposed deep-learning sensor fusion algorithm. The research introduces a novel FIDCE algorithm for precise biofouling identification in underwater images alongside the MFONet model for pixel-level segmentation. FIDCE enhances image quality and accurately identifies biofouling, which is vital for ship maintenance. MFONet outperforms classical algorithms, offering superior speed and accuracy, enabling automated cleaning and maintenance planning for underwater vehicles [81]. Laura A. Martinho et al. propose a learning-based approach for enhancing

the quality of underwater images. It involves two main steps: Firstly, a Convolutional Neural Network (CNN) Regression model learns optimal parameters for enhancing different types of underwater images. Second, intensity transformation techniques are applied to raw underwater images to compensate for the loss of visual information [51].

VII. Predictions about Future Development Directions Based on the Above Content

This study establishes a strong foundation for advancing Unmanned Underwater Vehicle (UUV) navigation, focusing on refining AI-SLAM algorithms, particularly those driven by deep learning. Future research endeavors will involve optimizing multiple sensor fusion techniques, incorporating technologies like multi-beam sonar, stereo cameras, Lidar, IMU (INS), and methods such as SBL/USBL and DLV to enhance UUV navigation accuracy in complex underwater environments. Exploring the integration of emerging technologies, such as machine learning and advanced computer vision, holds promise for developing even more robust UUV navigation systems. Scalability for different UUV types and mission requirements is a crucial consideration, and collaborative efforts among researchers, industry experts, and policymakers are essential for standardizing and implementing these advancements in practical applications. As technology evolves, the future of underwater navigation and Simultaneous Localization and Mapping (SLAM) is poised for significant growth, driven by deep learning applications. Anticipated developments include refining model architectures, addressing domain adaptation challenges, optimizing algorithms for real-time processing efficiency, expanding application scopes, integrating deep learning with sensor advancements, exploring unsupervised learning methods, and fostering interdisciplinary collaborations. These efforts aim to propel the field towards more precise, adaptable, and robust underwater navigation systems, harnessing the transformative capabilities of deep learning in navigating complex and dynamic underwater environments. Future research may also explore other Deep Reinforcement Learning (DRL) algorithms like Deep Deterministic Policy Gradient (DDPG), Soft Actor-Critic (SAC), and Proximal Policy Optimization (PPO) for further optimization of the 3D image model reconstruction.

Conclusion/Significance

Integrating multiple sensors and deep learning techniques represents a significant leap forward in improving underwater navigation. By overcoming the limitations of traditional methods through innovative approaches like visual SLAM and sensor fusion, this research opens new possibilities for using Unmanned Underwater Vehicles in a wide range of applications, from scientific exploration to marine resource management. Advancements in precision and reliability enhance the capabilities of UUVs and contribute to our understanding and stewardship of the Earth's oceans, making this work a crucial step in the ongoing evolution of underwater technology.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. E. Vargas *et al.*, "Robust Underwater Visual SLAM Fusing Acoustic Sensing," in *Proceedings - IEEE International Conference on Robotics and Automation*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 2140–2146. doi: 10.1109/ICRA48506.2021.9561537.
2. Z. Xu, M. Haroutunian, A. J. Murphy, J. Neasham, and R. Norman, "An underwater visual navigation method based on multiple aruco markers," *J Mar Sci Eng*, vol. 9, no. 12, Dec. 2021, doi: 10.3390/jmse9121432.
3. K. Sun, W. Cui, and C. Chen, "Review of underwater sensing technologies and applications," *Sensors*, vol. 21, no. 23. MDPI, Dec. 01, 2021. doi: 10.3390/s21237849.

4.   Y. Zhang, Y. Wu, K. Tong, H. Chen, and Y. Yuan, "Review of Visual Simultaneous Localization and Mapping Based on Deep Learning," *Remote Sensing*, vol. 15, no. 11. MDPI, Jun. 01, 2023. doi: 10.3390/rs15112740.

5.   A. B. Azam *et al.*, "Low-cost Underwater Localisation Using Single-Beam Echosounders and Inertial Measurement Units," Institute of Electrical and Electronics Engineers (IEEE), Sep. 2023, pp. 1–7. doi: 10.1109/oceanslimerick52467.2023.10244255.

6.   H. Yang, Z. Xu, and B. Jia, "An Underwater Positioning System for UUVs Based on LiDAR Camera and Inertial Measurement Unit," *Sensors*, vol. 22, no. 14, Jul. 2022, doi: 10.3390/s22145418.

7.   S. Tani, F. Ruscio, M. Bresciani, B. M. Nordfeldt, F. Bonin-Font, and R. Costanzi, "Development and testing of a navigation solution for Autonomous Underwater Vehicles based on stereo vision," *Ocean Engineering*, vol. 280, Jul. 2023, doi: 10.1016/j.oceaneng.2023.114757.

8.   J. Hou and X. Ye, "Real-time Underwater 3D Reconstruction Method Based on Stereo Camera," in *2022 IEEE International Conference on Mechatronics and Automation, ICMA 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 1204–1209. doi: 10.1109/ICMA54519.2022.9855905.

9.   J. V. Aravind, K. V. S. Sai Ganesh, and S. Prince, "Real-Time Appearance Based Mapping using Visual Sensor for Unknown Environment," in *Journal of Physics: Conference Series*, Institute of Physics, 2022. doi: 10.1088/1742-6596/2335/1/012057.

10. Mo J. Towards a fast, robust and accurate visual-inertial simultaneous localization and mapping system (Doctoral dissertation, University of Minnesota).

11. S. Li, Z. Li, X. Liu, C. Shan, Y. Zhao, and H. Cheng, "Research on Map-SLAM Fusion Localization Algorithm for Unmanned Vehicle," *Applied Sciences (Switzerland)*, vol. 12, no. 17, Sep. 2022, doi: 10.3390/app12178670.

12. J. Yin, Y. Wang, J. Lv, and J. Ma, "Study on Underwater Simultaneous Localization and Mapping Based on Different Sensors," in *Proceedings of 2021 IEEE 10th Data Driven Control and Learning Systems Conference, DDCLS 2021*, Institute of Electrical and Electronics Engineers Inc., May 2021, pp. 728–733. doi: 10.1109/DDCLS52934.2021.9455508.

13. D. Q. Huy, N. Sadjoli, A. B. Azam, B. Elhadidi, Y. Cai, and G. Seet, "Object perception in underwater environments: a survey on sensors and sensing methodologies," *Ocean Engineering*, vol. 267. Elsevier Ltd, Jan. 01, 2023. doi: 10.1016/j.oceaneng.2022.113202.

14. A. Kim and R. M. Eustice, "Real-time visual SLAM for autonomous underwater hull inspection using visual saliency," *IEEE Transactions on Robotics*, vol. 29, no. 3, pp. 719–733, 2013, doi: 10.1109/TRO.2012.2235699.

15. D. Scaramuzza and F. Fraundorfer, "Tutorial: Visual odometry," *IEEE Robot Autom Mag*, vol. 18, no. 4, pp. 80–92, Dec. 2011, doi: 10.1109/MRA.2011.943233.

16. W. Chen *et al.*, "Overview of Multi-Robot Collaborative SLAM from the Perspective of Data Fusion," *Machines*, vol. 11, no. 6. MDPI, Jun. 01, 2023. doi: 10.3390/machines11060653.

17. L. Paull, S. Saeedi, M. Seto, and H. Li, "AUV navigation and localization: A review," *IEEE Journal of Oceanic Engineering*, vol. 39, no. 1. pp. 131–149, Jan. 2014. doi: 10.1109/JOE.2013.2278891.

18. S. Zhang *et al.*, "Visual SLAM for underwater vehicles: A survey," *Computer Science Review*, vol. 46. Elsevier Ireland Ltd, Nov. 01, 2022. doi: 10.1016/j.cosrev.2022.100510.

19. I. Abaspur Kazerouni, L. Fitzgerald, G. Dooly, and D. Toal, "A survey of state-of-the-art on visual SLAM," *Expert Systems with Applications*, vol. 205. Elsevier Ltd, Nov. 01, 2022. doi: 10.1016/j.eswa.2022.117734.

20. J. Cheng, L. Zhang, Q. Chen, X. Hu, and J. Cai, "A review of visual SLAM methods for autonomous driving vehicles," *Engineering Applications of Artificial Intelligence*, vol. 114. Elsevier Ltd, Sep. 01, 2022. doi: 10.1016/j.engappai.2022.104992.

21. Y. Zhang, L. Zhou, H. Li, J. Zhu, and W. Du, "Marine Application Evaluation of Monocular SLAM for Underwater Robots," *Sensors*, vol. 22, no. 13, Jul. 2022, doi: 10.3390/s22134657.

22. Z. Zheng, Z. Xin, Z. Yu, and S. K. Yeung, "Real-time GAN-based image enhancement for robust underwater monocular SLAM," *Front Mar Sci*, vol. 10, 2023, doi: 10.3389/fmars.2023.1161399.

23. D. Scaramuzza and F. Fraundorfer, "Tutorial: Visual odometry," *IEEE Robot Autom Mag*, vol. 18, no. 4, pp. 80–92, Dec. 2011, doi: 10.1109/MRA.2011.943233.

24. Z. Javed and G. W. Kim, "PanoVILD: a challenging panoramic vision, inertial and LiDAR dataset for simultaneous localization and mapping," *Journal of Supercomputing*, vol. 78, no. 6, pp. 8247–8267, Apr. 2022, doi: 10.1007/s11227-021-04198-1.

25. J. Hou and X. Ye, "Real-time Underwater 3D Reconstruction Method Based on Stereo Camera," in *2022 IEEE International Conference on Mechatronics and Automation, ICMA 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 1204–1209. doi: 10.1109/ICMA54519.2022.9855905.

26. J. Hou and X. Ye, "Real-time Underwater 3D Reconstruction Method Based on Stereo Camera," in *2022 IEEE International Conference on Mechatronics and Automation, ICMA 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 1204–1209. doi: 10.1109/ICMA54519.2022.9855905.

27. J. Hou and X. Ye, "Real-time Underwater 3D Reconstruction Method Based on Stereo Camera," in *2022 IEEE International Conference on Mechatronics and Automation, ICMA 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 1204–1209. doi: 10.1109/ICMA54519.2022.9855905.

28. X. Wang, X. Fan, P. Shi, J. Ni, and Z. Zhou, "An Overview of Key SLAM Technologies for Underwater Scenes," *Remote Sensing*, vol. 15, no. 10. MDPI, May 01, 2023. doi: 10.3390/rs15102496.

29. K. Köser and U. Frese, "Challenges in Underwater Visual Navigation and SLAM," in *Intelligent Systems, Control and Automation: Science and Engineering*, vol. 96, Springer Netherlands, 2020, pp. 125–135. doi: 10.1007/978-3-030-30683-0_11.

30. T. J. Chong, X. J. Tang, C. H. Leng, M. Yogeswaran, O. E. Ng, and Y. Z. Chong, "Sensor Technologies and Simultaneous Localization and Mapping (SLAM)," in *Procedia Computer Science*, Elsevier B.V., 2015, pp. 174–179. doi: 10.1016/j.procs.2015.12.336.

31. A. B. Azam *et al.*, "Low-cost Underwater Localisation Using Single-Beam Echosounders and Inertial Measurement Units," Institute of Electrical and Electronics Engineers (IEEE), Sep. 2023, pp. 1–7. doi: 10.1109/oceanslimerick52467.2023.10244255.

32. H. Horimoto, T. Maki, and K. Kofuji, "XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE Autonomous Sea Turtle Detection Using Multi-beam Imaging Sonar: Toward Autonomous Tracking."

33. J. Pyo, S. Song, and S.-C. Yu, "Acoustic beam-based man-made underwater landmark detection method for multi-beam sonar."

34. Y. Lim, Y. Lee, T. K. Yeu, and S. Lee, "Underwater Terrain Map Building Based on Depth Image Using Multi-Beam Sonar Sensor," in *2023 20th International Conference on Ubiquitous Robots, U.R. 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 54–58. doi: 10.1109/UR57808.2023.10202363.

35. M. Xiang-Jian, X. Wen, S. Bin-Jian, G. Xinxin, L. Xin-Yu, and L. Yushi, "An Imaging Algorithm for High-speed Side-scan Sonar Based on Multi-beam Forming Technology," in *2020 Global Oceans 2020: Singapore - U.S. Gulf Coast*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020. doi: 10.1109/IEEECONF38699.2020.9389346.

36. D. C. Estrada, F. R. Dalgleish, C. J. Den Ouden, B. Ramos, Y. Li, and B. Ouyang, "Underwater LiDAR Image Enhancement Using a GAN Based Machine Learning Technique," *IEEE Sens J*, vol. 22, no. 5, pp. 4438–4451, Mar. 2022, doi: 10.1109/JSEN.2022.3146133.

37. C. Debeunne and D. Vivet, "A review of visual-lidar fusion based simultaneous localization and mapping," *Sensors (Switzerland)*, vol. 20, no. 7. MDPI AG, Apr. 01, 2020. doi: 10.3390/s20072068.

38. H. Xing *et al.*, "A Multisensor Fusion Self-Localization System of a Miniature Underwater Robot in Structured and GPS-Denied Environments," *IEEE Sens J*, vol. 21, no. 23, pp. 27136–27146, Dec. 2021, doi: 10.1109/JSEN.2021.3120663.

39. N. Balemans, P. Hellinckx, S. Latre, P. Reiter, and J. Steckel, "S2L-SLAM: Sensor Fusion Driven SLAM using Sonar, LiDAR and Deep Neural Networks," in *Proceedings of IEEE Sensors*, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/SENSORS47087.2021.9639772.

40. I. C. Sang and W. R. Norris, "An Autonomous Underwater Vehicle Simulation With Fuzzy Sensor Fusion for Pipeline Inspection," *IEEE Sens J*, vol. 23, no. 8, pp. 8941–8951, Apr. 2023, doi: 10.1109/JSEN.2023.3250721.

41. W. Chen *et al.*, "SLAM Overview: From Single Sensor to Heterogeneous Fusion," *Remote Sensing*, vol. 14, no. 23. MDPI, Dec. 01, 2022. doi: 10.3390/rs14236033.

42. C. Q. Feng, B. L. Li, Y. F. Liu, F. Zhang, Y. Yue, and J. S. Fan, "Crack assessment using multisensor fusion simultaneous localization and mapping (SLAM) and image super-resolution for bridge inspection," *Autom Constr*, vol. 155, Nov. 2023, doi: 10.1016/j.autcon.2023.105047.

43. R. Chaudhuri, S. Deb, and H. Das, "Noble Approach on Sensor Fused Bio Intelligent Path Optimisation and Single Stage Obstacle Recognition in Customized Mobile Agent," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 778–787. doi: 10.1016/j.procs.2023.01.058.

44. T. Nicosevici, R. Garcia, M. Carreras, and M. Villanueva, "A Review of Sensor Fusion Techniques for Underwater Vehicle Navigation."

45. S. Rahman, A. Quattrini Li, and I. Rekleitis, "SVIn2: A multisensor fusion-based underwater SLAM system."

46. Li, C. and Guo, S., 2023. Characteristic evaluation via multisensor information fusion strategy for spherical underwater robots. *Information Fusion*, *95*, pp.199-214.

47. D. Wang, B. Wang, H. Huang, and H. Zhang, "A Multisensor Fusion Method Based on Strict Velocity for Underwater Navigation System," *IEEE Sens J*, vol. 23, no. 16, pp. 18587–18598, Aug. 2023, doi: 10.1109/JSEN.2023.3287600.

48. M. N. Hoda, Bharati Vidyapeeth's Institute of Computers Applications and Management Delhi, and Institute of Electrical and Electronics Engineers Delhi Section, *Proceedings of the 17th INDIACom; 2023 10th International Conference on Computing for Sustainable Global Development (15th-17th March, 2023) INDIACom-2023*.

49. A. Khan, M. M. Fouda, D. T. Do, A. Almaleh, A. M. Alqahtani, and A. U. Rahman, "Underwater Target Detection Using Deep Learning: Methodologies, Challenges, Applications, and Future Evolution," *IEEE Access*, vol. 12, pp. 12618–12635, 2024, doi: 10.1109/ACCESS.2024.3353688.

50. A. Khandouzi and M. Ezoji, "Coarse-to-fine underwater image enhancement with lightweight CNN and attention-based refinement," *J Vis Commun Image Represent*, vol. 99, Mar. 2024, doi: 10.1016/j.jvcir.2024.104068.

51. L. A. Martinho, J. M. B. Calvalcanti, J. L. S. Pio, and F. G. Oliveira, "Diving into Clarity: Restoring Underwater Images using Deep Learning," *J Intell Robot Syst*, vol. 110, no. 1, Mar. 2024, doi: 10.1007/s10846-024-02065-8.

52. J. Zheng *et al.*, "An Underwater Image Restoration Deep Learning Network Combining Attention Mechanism and Brightness Adjustment," *J Mar Sci Eng*, vol. 12, no. 1, Jan. 2024, doi: 10.3390/jmse12010007.

53. R. Joshi, K. Usmani, G. Krishnan, F. Blackmon, and B. Javidi, "Underwater object detection and temporal signal detection in turbid water using 3D-integral imaging and deep learning," *Opt Express*, vol. 32, no. 2, p. 1789, Jan. 2024, doi: 10.1364/oe.510681.

54. B. Munoz and G. Troni, "Learning the Ego-Motion of an Underwater Imaging Sonar: A Comparative Experimental Evaluation of Novel CNN and RCNN Approaches," *IEEE Robot Autom Lett*, vol. 9, no. 3, pp. 2072–2079, Mar. 2024, doi: 10.1109/LRA.2024.3352357.

55. Y. Randall and T. Treibitz, "FLSea: Underwater Visual-Inertial and Stereo-Vision Forward-Looking Datasets," Feb. 2023, [Online]. Available: http://arxiv.org/abs/2302.12772

56. Y. H. Lin, T. L. Wu, C. M. Yu, and I. C. Wu, "Development of an intelligent underwater recognition system based on the deep reinforcement learning algorithm in an autonomous underwater vehicle," *Measurement (Lond)*, vol. 214, Jun. 2023, doi: 10.1016/j.measurement.2023.112844.

57. A. Burguera and F. Bonin-Font, "An Unsupervised Neural Network for Loop Detection in Underwater Visual SLAM," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 100, no. 3–4, pp. 1157–1177, Dec. 2020, doi: 10.1007/s10846-020-01235-8.

58. Z. Xin, Z. Wang, Z. Yu, and B. Zheng, "ULL-SLAM: underwater low-light enhancement for the front-end of visual SLAM," *Front Mar Sci*, vol. 10, 2023, doi: 10.3389/fmars.2023.1133881.

59. J. Zalewski and S. Hożyń, "Computer Vision-Based Position Estimation for an Autonomous Underwater Vehicle," *Remote Sens (Basel)*, vol. 16, no. 5, p. 741, Feb. 2024, doi: 10.3390/rs16050741.

60. C. Wang, C. Cheng, D. Yang, G. Pan, and F. Zhang, "Underwater AUV Navigation Dataset in Natural Scenarios," *Electronics (Switzerland)*, vol. 12, no. 18, Sep. 2023, doi: 10.3390/electronics12183788.

61. N. Keetha *et al.*, "SplaTAM: Splat, Track & Map 3D Gaussians for Dense RGB-D SLAM," Dec. 2023, [Online]. Available: http://arxiv.org/abs/2312.02126

62. E. Sandström, K. Ta, L. Van Gool, and M. R. Oswald, "UncLe-SLAM: Uncertainty Learning for Dense Neural SLAM," Jun. 2023, [Online]. Available: http://arxiv.org/abs/2306.11048

63. Z. Zhu *et al.*, "NICE-SLAM: Neural Implicit Scalable Encoding for SLAM." [Online]. Available: https://github.com/cvg/nice-slam.

64. E. Sandström, Y. Li, L. Van Gool, M. R. Oswald, E. Zürich, and K. Leuven, "Point-SLAM: Dense Neural Point Cloud-based SLAM."

65. B. Teixeira, H. Silva, A. Matos, and E. Silva, "Deep Learning for Underwater Visual Odometry Estimation," *IEEE Access*, vol. 8, pp. 44687–44701, 2020, doi: 10.1109/ACCESS.2020.2978406.

66. A. Jin and X. Zeng, "A Novel Deep Learning Method for Underwater Target Recognition Based on Res-Dense Convolutional Neural Network with Attention Mechanism," *J Mar Sci Eng*, vol. 11, no. 1, Jan. 2023, doi: 10.3390/jmse11010069.

67. M. Valdenegro-Toro, "Improving Sonar Image Patch Matching via Deep Learning."

68. Y. Wang *et al.*, "Robust AUV Visual Loop-Closure Detection Based on Variational Autoencoder Network," *IEEE Trans Industr Inform*, vol. 18, no. 12, pp. 8829–8838, Dec. 2022, doi: 10.1109/TII.2022.3145860.

69. B. Pedraza and D. Dera, "Robust Active Simultaneous Localization and Mapping Based on Bayesian Actor-Critic Reinforcement Learning," in *Proceedings - 2023 IEEE Conference on Artificial Intelligence, CAI 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 63–66. doi: 10.1109/CAI54212.2023.00035.

70. B. Teixeira, H. Silva, A. Matos, and E. Silva, "Deep Learning for Underwater Visual Odometry Estimation," *IEEE Access*, vol. 8, pp. 44687–44701, 2020, doi: 10.1109/ACCESS.2020.2978406.

71. Z. Xing, M. Cai, and J. Li, "Improved Shallow-UWnet for Underwater Image Enhancement," in *Proceedings of 2022 IEEE International Conference on Unmanned Systems, ICUS 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 1191–1196. doi: 10.1109/ICUS55513.2022.9986534.

72. M. N. Favorskaya, "Deep Learning for Visual SLAM: The State-of-the-Art and Future Trends," *Electronics (Switzerland)*, vol. 12, no. 9. MDPI, May 01, 2023. doi: 10.3390/electronics12092006.

73. C. Chen, B. Wang, C. X. Lu, N. Trigoni, and A. Markham, "Deep Learning for Visual Localization and Mapping: A Survey," Aug. 2023, [Online]. Available: http://arxiv.org/abs/2308.14039

74. J. Qin, M. Li, D. Li, J. Zhong, and K. Yang, "A Survey on Visual Navigation and Positioning for Autonomous UUVs," *Remote Sensing*, vol. 14, no. 15. MDPI, Aug. 01, 2022. doi: 10.3390/rs14153794.

75. R. Glenn Wright, "Intelligent autonomous ship navigation using multisensor modalities," *TransNav*, vol. 13, no. 3, pp. 503–510, Sep. 2019, doi: 10.12716/1001.13.03.03.

76. J. Li, X. Xiang, and S. Yang, "Robust adaptive neural network control for dynamic positioning of marine vessels with prescribed performance under model uncertainties and input saturation," *Neurocomputing*, vol. 484, pp. 1–12, May 2022, doi: 10.1016/j.neucom.2021.03.136.

77. M. Iman, K. Rasheed, and H. R. Arabnia, "A Review of Deep Transfer Learning and Recent Advancements template 2 A Review of Deep Transfer Learning and Recent Advancements."

78. A. Ashwini, K. E. Purushothaman, V. Gnanaprakash, D. F. Deva Shahila, T. Vaishnavi, and A. Rosi, "Transmission Binary Mapping Algorithm with Deep Learning for Underwater Scene Restoration," in *Proceedings of the International Conference on Circuit Power and Computing Technologies, ICCPCT 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 1545–1549. doi: 10.1109/ICCPCT58313.2023.10245055.

79. J. P. Sasi, K. Nidhi Pandagre, A. Royappa, S. Walke, P. G, and N. L, "Deep Learning Techniques for Autonomous Navigation of Underwater Robots," in *2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, IEEE, Dec. 2023, pp. 1630–1635. doi: 10.1109/UPCON59197.2023.10434865.

80. O. Alvarez-Tunon, Y. Brodskiy, and E. Kayacan, "Monocular visual simultaneous localization and mapping: (r)evolution from geometry to deep learning-based pipelines," *IEEE Transactions on Artificial Intelligence*, 2023, doi: 10.1109/TAI.2023.3321032.

81. Zhao, W., Han, F., Qiu, X., Peng, X., Zhao, Y. and Zhang, J., 2023. Research on the identification and distribution of biofouling using underwater cleaning robot based on deep learning. *Ocean Engineering, 273*, p.113909.