Review

Enhancing Underwater SLAM Navigation and Perception: A Comprehensive Review of Deep Learning Integration

Fomekong Fomekong Rachel Merveille [1], Baozhu Jia [1,*], Zhizun Xu [2] and Bissih Fred [3]

1   School of Naval Architecture and Maritime, Guangdong Ocean University, Zhanjiang 524000, China; marvelous@stu.gdou.edu.cn
2   School of Engineering, Newcastle University, Newcastle upon Tyne NE1 7RU, UK; zhizun@gdou.edu.cn
3   College of Fisheries, Guangdong Ocean University, Zhanjiang 524088, China; 1252201246@stu.gdou.edu.cn
*   Correspondence: jiabzh@gdou.edu.cn; Tel.: +86-18641194229

Abstract: Underwater simultaneous localization and mapping (SLAM) is essential for effectively navigating and mapping underwater environments; however, traditional SLAM systems have limitations due to restricted vision and the constantly changing conditions of the underwater environment. This study thoroughly examined the underwater SLAM technology, particularly emphasizing the incorporation of deep learning methods to improve performance. We analyzed the advancements made in underwater SLAM algorithms. We explored the principles behind SLAM and deep learning techniques, examining how these methods tackle the specific difficulties encountered in underwater environments. The main contributions of this work are a thorough assessment of the research into the use of deep learning in underwater image processing and perception and a comparison study of standard and deep learning-based SLAM systems. This paper emphasizes specific deep learning techniques, including generative adversarial networks (GANs), convolutional neural networks (CNNs), long short-term memory (LSTM) networks, and other advanced methods to enhance feature extraction, data fusion, scene understanding, etc. This study highlights the potential of deep learning in overcoming the constraints of traditional underwater SLAM methods, providing fresh opportunities for exploration and industrial use.

Keywords: underwater simultaneous localization and mapping (SLAM); underwater navigation; deep learning; odometry navigation

1. Introduction

The Earth's oceans, which span over 71% of the planet's surface, are precious resources for scientific investigation and environmental understanding [1,2]. Unmanned underwater vehicles (UUVs) [3,4] have a variety of crucial applications, including marine mining and pipeline inspection [5–7]. However, their efficacy is frequently hampered by the limits of conventional navigation systems [8,9]. Inertial sensors and acoustic beacons used for underwater navigation suffer from accumulated errors, limited range, and environmental interference [9,10]. Furthermore, optical problems such as low lighting, turbidity, scattering, and wavelength absorption affect visual dependability [11]. The lack of access to global positioning systems (G.P.S.) [12–14] affects accurate location determination and data collecting in underwater situations.

The difficulties of underwater navigation have increased the demand for more dependable and precise solutions. Recent advances in deep learning [15], particularly in visual simultaneous localization and mapping (SLAM) [16], have revealed intriguing areas for improvement. Significant progress has been achieved in tackling the specific constraints of underwater environments, resulting in enhanced capabilities for unmanned underwater vehicles (UUVs) [17,18]. This study investigates these improvements, evaluating their potential to transform underwater navigation and perception.

This paper reviews crucial studies on how deep learning improves underwater SLAM systems. The literature selection methodology includes comprehensive searches in scientific databases such as IEEE Xplore, SpringerLink, ScienceDirect, Nature, and Google Scholar using keywords such as "underwater SLAM", "deep learning", "UUV navigation", "N.N. underwater" [19–25], "RNN SLAM", "GAN underwater" [26–28], and "V.A.E. mapping". Papers were selected based on their ability to advance the field, their citation count, their publication in reputable journals and conferences, and their relevance to the incorporation of deep learning techniques [29] into underwater SLAM. Recent developments and seminal works were assessed. To preserve the quality and relevance of the review, papers that were not peer-reviewed, lacked experimental validation, or did not explicitly address underwater applications were excluded.

The chosen articles were classified into different groups according to their main emphasis, including conventional primary techniques, CNNs for extracting features, R.N.N.s for modeling temporal aspects, DRL for modeling and mapping, GANs [30] for augmenting data, and V.A.E.s for exploring latent space. Every publication underwent a thorough assessment to identify and extract the main contributions, methodology or processes, results, and implications for the discipline. To enhance the understanding of the status of the field, we emphasized the most influential and significant papers in each area, ensuring a well-organized and comprehensive overview. This study seeks to give a clear roadmap for academics and practitioners interested in advancing the area of underwater simultaneous localization and mapping (SLAM) by integrating deep learning techniques [15,31]. It accomplishes this by categorizing the existing literature and emphasizing the most influential studies.

There are eight main sections in this paper: (1)Introduction, which establishes the study goals and relevance, gives background information on underwater SLAMs, and emphasizes the function of deep learning; (2) progress in algorithmic underwater SLAM with an emphasis on performance in underwater environments, which objectively assesses the most recent advancements in SLAM technologies; (3) mathematical foundations of underwater SLAM and deep learning approaches, which presents the theoretical models underlying these methods; (4) deep learning uses in underwater perception, navigation, and image processing, which examines deep learning techniques for enhancing underwater sensing and navigation; (5) strengths and limitations of deep learning-based underwater SLAM and odometry navigation, which provides a review of the benefits and drawbacks of deep learning for SLAM; (6) comparative analysis of underwater SLAM techniques, which evaluates the resilience and efficiency of several SLAM strategies; (7) deep learning's superiority over conventional methods, which emphasizes the technology's benefits in solving undersea problems; and (8) conclusion and following directions, which emphasizes deep learning's developing importance in underwater SLAM, summarizes findings, and recommends following research directions.

2. Common Underwater SLAM Advancements and Algorithm Performance

SLAM has advanced, especially with the extended Kalman filter (E.K.F.) [32] SLAM for probabilistic robot pose and landmark estimation. Due to its linearization assumptions, EKF SLAM [33] struggles in complicated nonlinear situations despite its popularity. The square root information filter (SRIF) [34] algorithm addresses these issues. By controlling the covariance matrix and resolving numerical instability, SRIF improves underwater UUV navigation stability, numerical reliability, and operational efficiency.

Different SLAM methods address various issues in various contexts. Particle filter SLAM [35] thrives in highly nonlinear settings, while graph-based SLAM [36] optimizes graph representations for configuration detection. Bayesian and Rao–Blackwellized particle filter (RBPF) SLAMs increase mapping and localization using Bayesian [37] estimation and computational efficiency, respectively.

Underwater sensor constraints must be addressed. Innovative technologies like sonar-based [38] mapping and loop closure detection improve SLAM performance.

Visual SLAM [39] algorithms for underwater use have been tested. ORB-SLAM [40] performs best in well-lit and feature-rich situations, while ROVIO [41] uses visual and inertial data to excel in dynamic underwater environments. LSD-SLAM [42] maps perform well in texture-rich underwater environments, and DVO-SLAM [43,44] handles depth changes well. The multi-state constraint Kalman filter (MSCKF) [45,46] uses visual and inertial measurements for precise navigation, and SVO [47,48] is efficient for lightweight underwater vehicles. Deep learning-enhanced [49] VISLAM [50] algorithms can extract and map features in challenging underwater settings. FAB-MAP [51], designed for underwater applications, performs well in varied underwater environments with distinct visual aspects.

Recent advances in underwater SLAM systems, as shown in Figure 1, have improved navigation accuracy. Even though underwater navigation is complex, deep learning and SLAM have improved UUVs for oceanographic investigation [52]. Deep learning improves SLAM by employing neural networks to enhance mapping and navigation, especially with LiDAR and vision sensors [53,54]. Regularly distributed artificial magnetic beacons can provide a repeated regional magnetic field, allowing autonomous cars to navigate with high positioning accuracy by considering landmarks for SLAM algorithms [55]. The VINS-MONO [56] technique uses FAST feature point extraction and inverse optical flow to increase speed and accuracy, addressing underwater picture degradation [57]. End-to-end networks for low-light SLAM preprocessing have also been developed with a low-light improvement branch and a self-supervised feature point detector to increase feature point extraction and reduce re-projection mistakes [58].
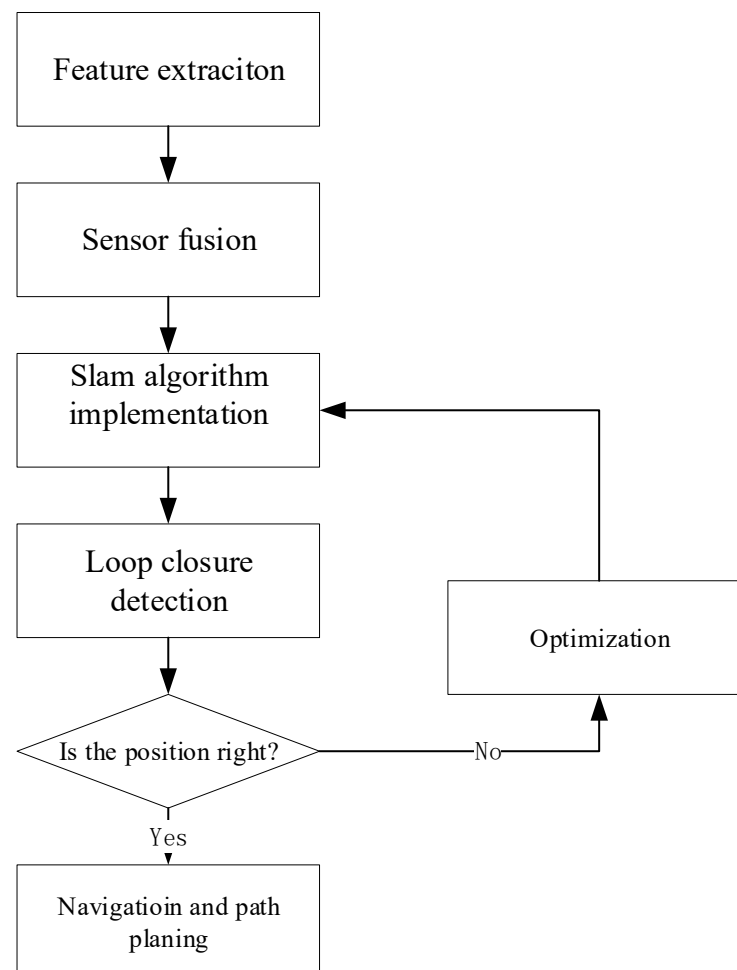

Figure 1. Underwater SLAM process.

Generative adversarial networks (GANs) [27] for real-time underwater image improvement have improved SLAM performance by tackling poor visibility, low contrast, and color distortion, providing robust and accurate monocular SLAM systems. In unknown underwater environments, multi-vehicle collaborative mapping is more efficient due to algorithms like Gaussian mixture robust branch and bound (GMRBnB) that improve map registration accuracy and outlier tolerance [59]. Improved unscented Kalman filter SLAM (IUKF-SLAM) was designed to improve the accuracy, consistency, and convergence of the unscented Kalman filter (U.K.F.) [60] used in SLAM, displaying a better performance compared to existing E.K.F. [33,61] and U.K.F. approaches [62]. Underwater SLAM combines several information sources to overcome the limits of standard navigation systems, particularly in complex and unstructured underwater landscapes, allowing high-precision navigation and placement even without satellite information [63].

These advances in underwater SLAM algorithms, including deep learning integration, magnetic beacon use, image enhancement, collaborative mapping, and improved filtering, have greatly enhanced underwater navigation systems' accuracy and reliability, making systems more autonomous and effective in underwater exploration and operations.

3. Mathematical Formulas for Underwater SLAM and Deep Learning Methods

Underwater SLAM uses a variety of mathematical formulas and algorithms to map and navigate effectively. The PF-backend (particle filter-backend) approach utilizes particle filters for loop closure estimates and map consistency [64]. Variational Bayesian (V.B.) learning estimates UUV route and observation noise using the inverse-gamma distribution. The Gaussian mixture robust branch and bound (GMRBnB) technique improves map registration by extracting features and inliers. For robust SLAM, dual-stage bathymetric data association and Euler-deconvolution algorithms localize magnetic beacons. For real-time location, coalition game theory optimizes artificial magnetic beacon distribution. Neural network-detected semantic landmarks are used in object-level SLAM [65]. The preparation of mechanically scanned imaging sonar (MSIS) data allows for the utilization of traditional 2D laser SLAM frameworks, such as Gmapping and Cartographer, to tackle obstacles encountered underwater [66,67]. These methods jointly enhance the accuracy and dependability of SLAM by tackling problems such as error accumulation and feature extraction in low-texture situations. Below are some derivations of the mathematical formulae used in SLAM.

E.K.F. State Update

$$X_{k|k} = X_{k|k-1} + K_k \left(Z_k - h \left(X_{k|k-1} \right) \right) \tag{1}$$

$X_{k|k}$: Update state estimate, Z_k: Measurement at time k, h: Measurement function.

E.K.F. Covariance Update

$$P_{k|k} = (I - K_k H_k) P_{k|k-1} \tag{2}$$

$P_{k|k}$: Update covariance estimate, I: Identity matrix.

Particle Filter Weight Update

$$W_t[i] \propto W_{t-1}[i].p(z_t| x_t[i], m_{t-1}) \tag{3}$$

$W_t[i]$: Weight of the particle i at time t, z_t: Measurement at time t,
$x_t[i]$: State of particle i at time t, m_{t-1}: Map at time t − 1.

Particle Filter Resampling

$$\{X_t[i], W_t[i]\}_{i=1}^{N} \sim \{X_t[i], \widetilde{W}_t[i]\}_{i=1}^{N} \tag{4}$$

$X_t[i]$: State of particle i, $W_t[i]$: Weight of particle i, $\widetilde{W}_t[i]$: Normalized weight of the particle i.

Convolution Operation

$$(I * K)(i, j) = \sum_m \sum_n I(i + m, j + n). K(m, n) \tag{5}$$

I: Input image, K: Kernel, (i, j): Position in the output image.

ReLU Activation Function

$$f(x) = max(0, x) \tag{6}$$

$f(x)$: Output, x: Input.

R.N.N. Hidden State Update

$$h_t = \sigma_h(W_h x_t + U_h h_{t-1} + b_n) \tag{7}$$

h_t: Hidden state at time t, σ_h: Activation function, W_h: Weight matrix for input, x_t: Input at time t, U_h: Weight matrix for the hidden state, b_n: Bias term.

LSTM Input Gate

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{8}$$

i_t: Input gate at time t, σ: Sigmoid function, W_i: Weight matrix for input, x_t: Input at time t, U_i: Weight matrix for the hidden state, b_i: Bias term.

LSTM Output Gate

$$O_t = \sigma(W_0 x_t + U_0 h_{t-1} + b_0) \tag{9}$$

O_t: Output gate at time t, σ: Sigmoid function, W_0: Weight matrix for input, x_t: Input at time t, U_0: Weight matrix for the hidden state, b_0: Bias term.

LSTM Hidden State Update

$$h_t = O_t \odot tanh(c_t) \tag{10}$$

h_t: Hidden state at time t, O_t: Output gate at time t, *tanh*: Hyperbolic tangent function, c_t: Cell state at time t.

Bellman Equation

$$Q(s, a) = r + \gamma \max_{a'} Q(s', a') \tag{11}$$

$Q(s, a)$: State-action value, r: Reward, γ: Discount factor, a': Next action, s': Next state.

GAN Loss Functions

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(Z)))] \tag{12}$$

G: Generator, D: Discriminator, x: Real data, z: Latent variable, $p_{data}(x)$: Data distribution, $p_z(z)$: Latent distribution.

V.A.E. Loss Function

$$(L(x, z) = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z) - D_{KL}(q_\theta(z|x) \parallel p_\theta(z)) \tag{13}$$

L(x, z): VAE loss, $q_\theta(z|x)$: Approximate posterior, $p_\theta(x|z)$: Likehood, D_{KL}: Kullback-Leibler divergence.

In SLAM, state estimation is vital for accuracy and efficiency. The E.K.F. handles nonlinear systems, predicting and updating states using sensor data. Equations (1) and (2) guide state updating and covariance updates, and Kalman gain could also be added to minimize uncertainty. For more complex non-Gaussian environments, particle filters (Equations (3) and (4)) offer robust alternatives by assigning particle weights and resampling for diversity.

Deep learning models enhance feature extraction in SLAM through convolutions (Equation (5)) and ReLU activation (Equation (6)). Recurrent neural networks (R.N.N.s)

and LSTMs (Equations (7)–(10)) handle temporal dependencies for more stable mapping in dynamic environments. Reinforcement learning (Equation (11)) optimizes exploration, while GANs and V.A.E.s (Equations (12) and (13)) generate environmental models, improving SLAM's robustness in unstructured spaces.

4. Deep Learning Techniques in Underwater Image Processing, Navigation, and Perception

Deep learning techniques are a groundbreaking approach in machine learning that involves using neural networks with numerous layers to analyze and understand intricate patterns in data. Convolutional neural networks (CNNs) [68] are highly effective in image tasks, as they can capture hierarchical features for tasks like image recognition and computer vision applications. On the other hand, recurrent neural networks (R.N.N.s), play a crucial role in processing sequential data, such as language and time-series information, because they can retain context and dependencies. Transfer learning strategies utilize pre-trained models to enhance performance on specific tasks, enabling efficient information transfer. Generative adversarial networks (GANs) [27–30] introduce a new method for generating realistic data. Advanced natural language processing (N.L.P.) models, such as BERT and G.P.T., demonstrate exceptional proficiency in comprehending and producing language that resembles human speech. Deep learning techniques [29,69] have been widely used in underwater applications to improve the accuracy of visual perception systems. These techniques have been applied in several areas, such as image processing, navigation, and perception, and have shown impressive proficiency in object detection, recognition, and segmentation tasks. Transfer learning has successfully addressed the problem of limited data availability, leading to encouraging outcomes in different undersea situations. In addition, integrating sensor data, such as sound and visual inputs, using deep learning architectures improves perceptual abilities, resulting in more precise mapping and better comprehension of the surroundings.

Deep learning has dramatically improved underwater image processing, navigation, and perception, effectively tackling specific issues. Methods such as VDSR and BCFO-based deep CNN enhance the clarity of images and object detection accuracy [21,70]. The FUnIE-GAN architecture enhances visual perception measures such as PSNR (peak signal-to-noise ratio), SSIM (structural similarity index), UCIQE (universal image quality index), and entropy [71]. The combination of VGG16 and visual saliency models improves image clarity and enhances the accuracy of colors [72]. Deep learning enhances underwater perception by using advanced techniques such as RCNN (region convolutional neural network) and CFTA (color filter tensor analysis). These techniques specifically target the challenges of inadequate illumination and limited visibility in underwater environments [73]. Surveys indicate that deep learning performs more remarkably in effectively managing difficulties like underwater turbulence, low contrast, and color distortion [74]. Convolutional neural networks (CNNs) illustrated in Figure 2, are highly proficient in the classification and detection of underwater species and objects, hence assisting in the protection of marine ecosystems [19]. These improvements enable new opportunities for exploring the ocean as well as different industrial uses (Tables 1 and 2 elaborate on the comparative analysis of the traditional deep learning-based underwater SLAM techniques).

The underwater SLAM RNN process using long short-term memory (LSTM) and recurrent neural networks (R.N.N.s) for underwater SLAM are shown in Figure 3. A sequence of retrieved features from consecutive frames is sent into the LSTM network to capture temporal dependencies. Temporal modeling via the LSTM network refines the robot's trajectory. Finally, the LSTM output refines feature-matching pose predictions for accurate and efficient underwater SLAM.

The procedures involved in employing DRL for underwater SLAM are shown in Figure 4. The state representation, which includes the current position, orientation, velocity, observed features, and sensor data, is where the process begins. The action space delineates the range of conceivable motions, depth modifications, and velocity changes. The reward

function promotes goal achievement, energy efficiency, obstacle avoidance, map accuracy, and exploration. Algorithms such as D.Q.N. and P.P.O. are used in policy learning to maximize cumulative rewards. Lastly, a UUV equipped with the trained policy is used for mapping, exploration, and navigation.
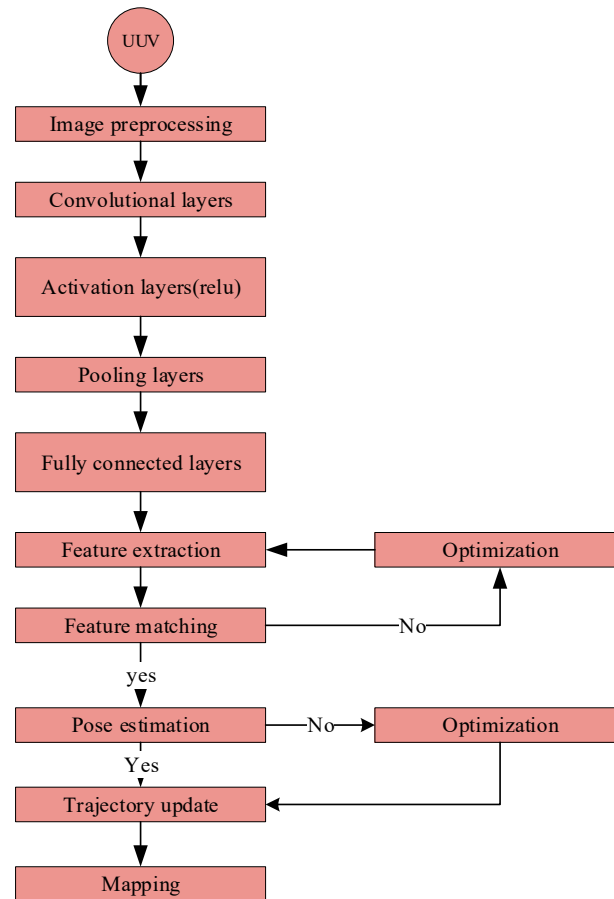

Figure 2. Deep CNNs in underwater SLAM process. This flow chart shows how deep CNNs [70] are used for underwater SLAMs. Images are captured using a UUV camera and preprocessed for visibility. For feature extraction, preprocessed pictures are routed via convolutional, activation, and pooling layers. Fully connected layers provide a high-level visual knowledge and match features between frames. Matching features estimate pose, update the robot's trajectory, and map the environment.

Table 1. A comparison of traditional underwater SLAM techniques caption.

Algorithm	Accuracy	Computational Efficiency	Robustness	Scalability	Sensor Fusion Capability
ORB-SLAM	High	Moderate	Moderate	High	Moderate
ROVIO	High	High	Moderate	High	Moderate
LSD-SLAM	Moderate	High	Moderate	High	Moderate
DVO-SLAM	High	Moderate	High	Moderate	Low
MSCKF	High	Very high	High	High	High
VSO	Moderate	Very high	Low	Low	Low
VSLAM	High	Moderate	High	High	High
RTAB-Map	Moderate	High	Moderate	High	Low
EKF-SLAM	High	Moderate	High	High	High
Graph-Based SLAM	High	Moderate	High	High	High
GMapping	Moderate	High	Moderate	Moderate	Moderate

Table 2. Comparative analysis of deep learning-based underwater SLAM techniques.

| Algorithm	Feature Extraction Quality	Temporal Modeling Accuracy	Data Augmentation Efficiency	Pose Estimation Precision	Mapping Accuracy
DeepVO	High	High	Moderate	High	High
PoseNet	High	High	Low	High	Moderate
DVL-SLAM	Moderate	Moderate	Low	Moderate	Moderate
SuperPoint	High	N/A	N/A	High	High
DeepMatcher	High	N/A	N/A	High	High
GAN-SLAM	High	N/A	High	High	High
VPF SLAM	Moderate	N/A	High	Moderate	High
RTAB-Map	High	High	Moderate	High	High
D.S.O.	High	High	Moderate	High	High
ORB-SLAM2	High	High	High	High	High
LOAM	High	High	Moderate	High	High
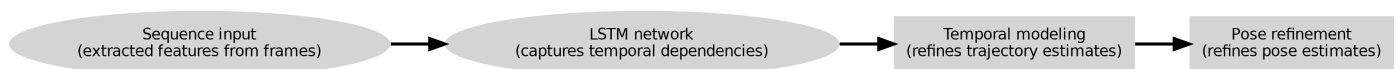
Figure 3. Underwater SLAM RNN process.
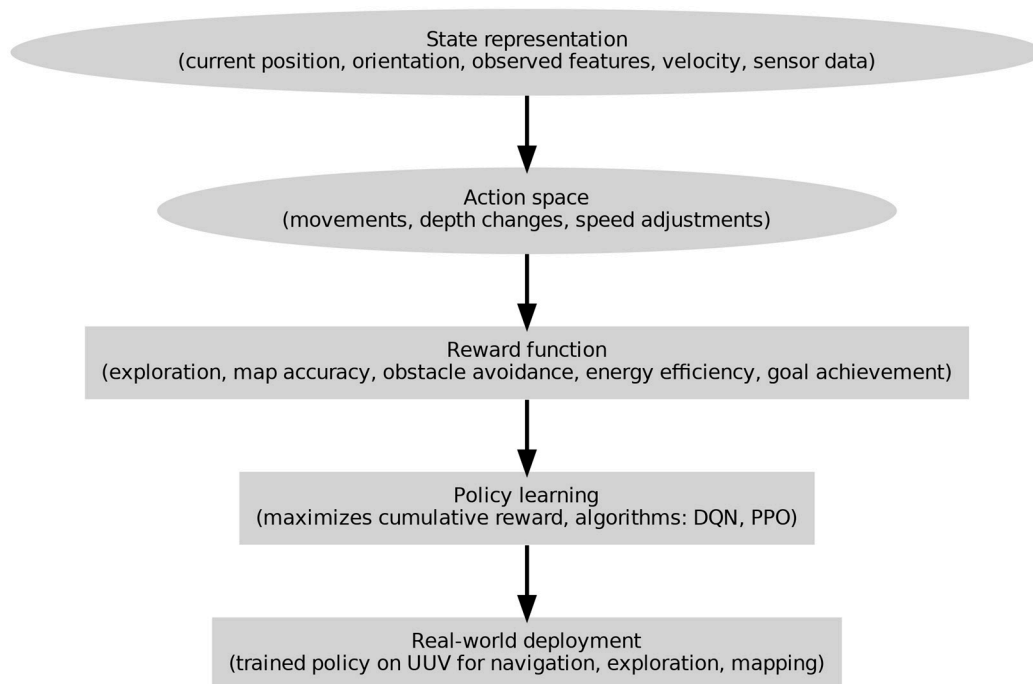
Figure 4. Detailed procedure for underwater SLAM for UUV using deep reinforcement learning (DRL).

Figure 5 shows how GANs generate realistic underwater images for SLAM deep learning model training. Real underwater photos are used to train the GAN. The discriminator network separates real from manufactured underwater images while the generator generates them. Adversarial training loops the generator and discriminator to increase visual realism. Synthetic data is added to the training dataset for SLAM jobs to improve underwater SLAM algorithm performance and robustness.
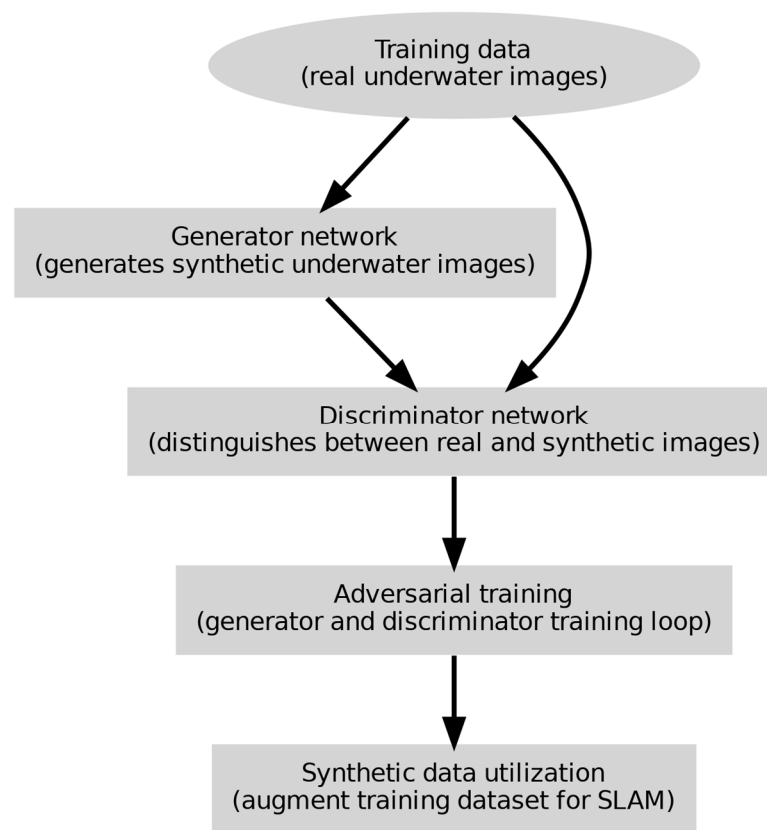
Figure 5. Detailed GAN process in underwater SLAM for UUV.

Figure 6 shows a detailed approach of V.A.E.s in underwater SLAM for UUVs.
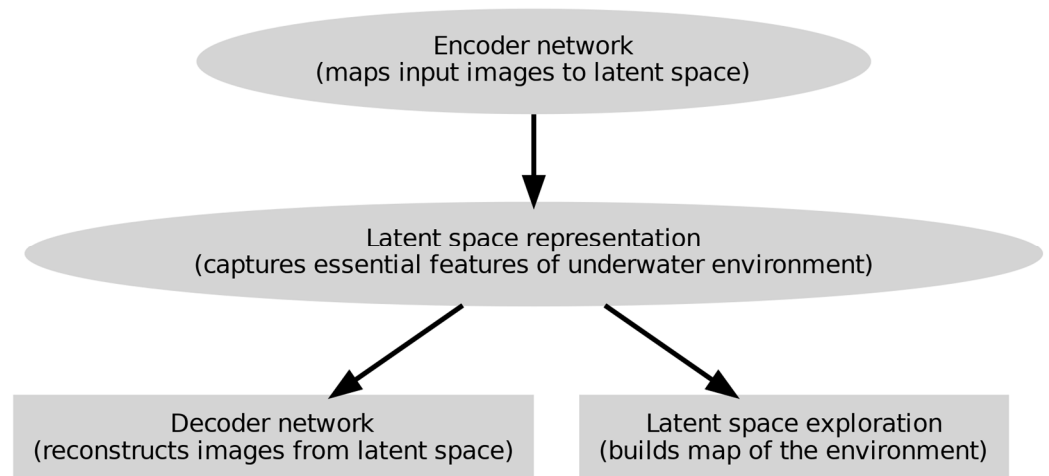
Figure 6. Comprehensive variational autoencoders.

The procedures for applying V.A.E.s for unsupervised learning of latent representations of underwater environments are shown in Figure 6. Starting with the encoder network, which maps input images to a latent space capturing the salient characteristics of the underwater environment, the process moves from the latent space and the decoder network then rebuilds the photos. Then, a map of the surroundings is created using the latent space representations, enabling efficient underwater SLAM.

Figure 7 aggregates the whole underwater SLAM in the UUVs process. Data collecting starts the process; underwater photos and sensor data are obtained. Image quality is

improved, and sensor data is preprocessed in the preprocessing phase. CNNs [75] are used in feature extraction to extract important picture features. Recurrent neural networks (R.N.N.s) or long short-term memory (LSTM) networks represent temporal dependencies in temporal modeling. Pose estimation computes the robot's posture with temporal modeling features. With generative adversarial networks (GANs) [27] employed for data augmentation and latent space exploration, deep reinforcement learning (DRL), shown in, helps navigate and generate a map, facilitating navigation and mapping.
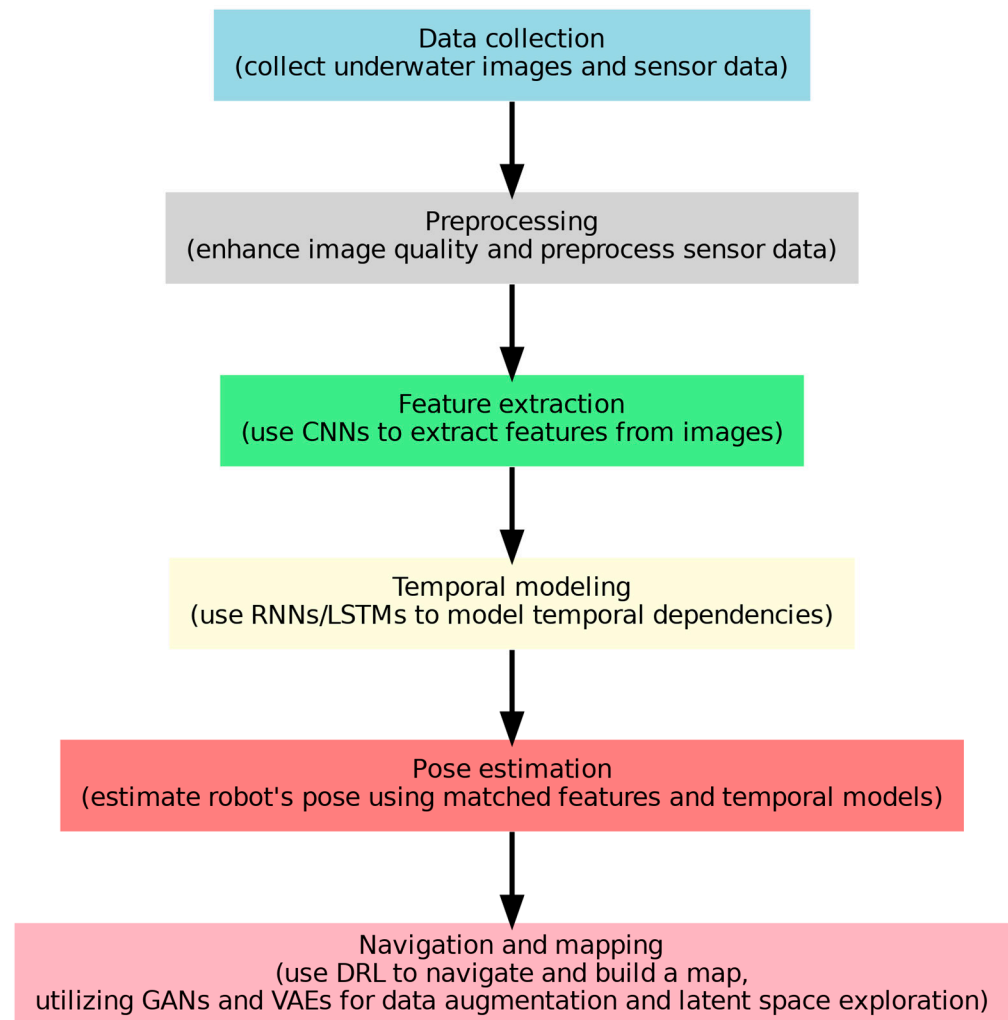
```
┌──────────────────────────────────────────────┐
│             Data collection                    │
│  (collect underwater images and sensor data)   │
└──────────────────────────────────────────────┘
                      │
                      ▼
┌──────────────────────────────────────────────┐
│               Preprocessing                    │
│  (enhance image quality and preprocess sensor  │
│                  data)                         │
└──────────────────────────────────────────────┘
                      │
                      ▼
┌──────────────────────────────────────────────┐
│             Feature extraction                 │
│  (use CNNs to extract features from images)    │
└──────────────────────────────────────────────┘
                      │
                      ▼
┌──────────────────────────────────────────────┐
│             Temporal modeling                  │
│  (use RNNs/LSTMs to model temporal             │
│              dependencies)                     │
└──────────────────────────────────────────────┘
                      │
                      ▼
┌──────────────────────────────────────────────┐
│               Pose estimation                  │
│  (estimate robot's pose using matched features │
│           and temporal models)                 │
└──────────────────────────────────────────────┘
                      │
                      ▼
┌──────────────────────────────────────────────┐
│           Navigation and mapping               │
│       (use DRL to navigate and build a map,    │
│ utilizing GANs and VAEs for data augmentation  │
│        and latent space exploration)           │
└──────────────────────────────────────────────┘
```

Figure 7. Workflow integration for UUV underwater SLAM.

Underwater SLAM systems have performed much better when advanced neural network designs like CNNs, GANs, and LSTMs were incorporated to handle issues like low visibility and changing conditions. Teixeira et al. found that deep learning models outperform classical approaches in underwater settings, improving visual odometry accuracy [76]. CNNs are particularly good at extracting features from underwater data, which makes them useful for tasks like multi-target tracking and habitat mapping [77–79]. Neural networks improve AUV (Autonomous underwater vehicle) navigation and obstacle avoidance, enabling them to adapt to changes in their surroundings and system degradation [80,81].

To enable real-time navigation by processing sequential input, LSTM networks process dynamic planning [82,83]. Also, deep networks make possible model-free localization, efficiently managing noise and variability [84,85]. According to some researchers, unsupervised neural networks increase mapping efficiency by improving loop detection. However, problems still exist, such as the requirement for massive databases and process-

ing power. The DM-GAN model, which improves depth map accuracy, is an example of how GANs support dense mapping in monocular SLAM [86]. Furthermore, GANs, aid in path planning by producing tenable routes for self-navigating systems [87].

Adding deep learning to SLAM increases accuracy by enhancing feature extraction and loop closure detection [88,89]. Spatial maps are improved by semantic mapping using CNNs. To demonstrate the efficacy of generative models in SLAM tasks, systems such as GEN-SLAM utilize generative modeling for monocular localization [90]. Even with the progress made, there are still difficulties in smoothly combining these models with conventional SLAM frameworks [91] and showing the improvements of deep learning methods in underwater SLAM. The next part of our dissertation elaborates on the strengths and weaknesses of deep learning UUV SLAM (with the workflow illustrated in Figure 7) and visual odometry navigation.

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

Deep learning-based SLAM and odometry [92,93] navigation systems exhibit several strengths. These include advanced feature extraction, where deep learning models identify intricate patterns from sensor data, thus enhancing mapping accuracy and navigation, and the use of end-to-end learning, simplifying navigation and boosting efficiency [3]. Additionally, deep learning [94,95] models can seamlessly integrate information from various sensors, handling diverse data sources coherently. They are capable of modeling nonlinear relationships and adapting to dynamic changes, making them suitable for unpredictable underwater scenarios [24,96,97].

Moreover, transfer learning allows pre-trained models to be adapted for underwater navigation, accelerating the training process and enabling quicker deployment in real-world scenarios [25]. However, these systems face challenges such as high computational demands and the 'black box' nature of neural networks, which may raise interpretability concerns [38,98]. Despite these challenges, the advantages presented by deep learning methods, make it a transformative force in advancing underwater SLAM navigation, offering promising avenues for further research and development.

Underwater SLAM and odometry [93,99–101] navigation systems that utilize deep learning have notable benefits, as shown in Figures 15 and 16, but they also encounter certain constraints [102]. These systems demonstrate exceptional performance in intricate and disorganized contexts where conventional approaches frequently prove ineffective. Deep neural networks (N.N.s) improve the precision of underwater acoustic localization (U.A.L.) in environments with reverberation [103]. Generative adversarial networks (GANs) enhance monocular visual SLAM by mitigating issues related to limited visibility and color aberration [26]. Systems like SVIn2, which combine sonar, visual, inertial, and water-pressure data, offer strong performance and dependable initialization [104]. Models with low computational requirements, such as TinyOdom, allow its immediate implementation on devices with few resources [105]. Nevertheless, the effectiveness of these systems relies on acquiring high-caliber training data, a challenging task in underwater environments, impacting accuracy [106]. Additionally, they need substantial processing resources, which pose difficulties for real-time applications on constrained systems. Integrating many sensors becomes more expensive and consumes more power as the complexity increases [107]. Additionally, the performance of neural inertial odometry frameworks might be negatively affected by environmental fluctuations and disturbances [108]. Notwithstanding these difficulties, progress in deep learning and sensor fusion is enhancing UUVs and applications of underwater robotics [109–111].

Figure 8 illustrates the sequential process from gathering data to implementing learning techniques, including supervised and unsupervised methods. The main steps involved in this process include data preprocessing, model training using CNN/RNN [20,21,23–25] architectures, and integrating input from visual, auditory, and geometric sensors in different ways. The process ends with the implementation of localization and mapping techniques, the

utilization of optimization strategies, and the use of advanced machine learning techniques such as reinforcement learning and ensemble approaches. Figure 8 offers a thorough outline of the standard procedures used to develop autonomous robotic systems.
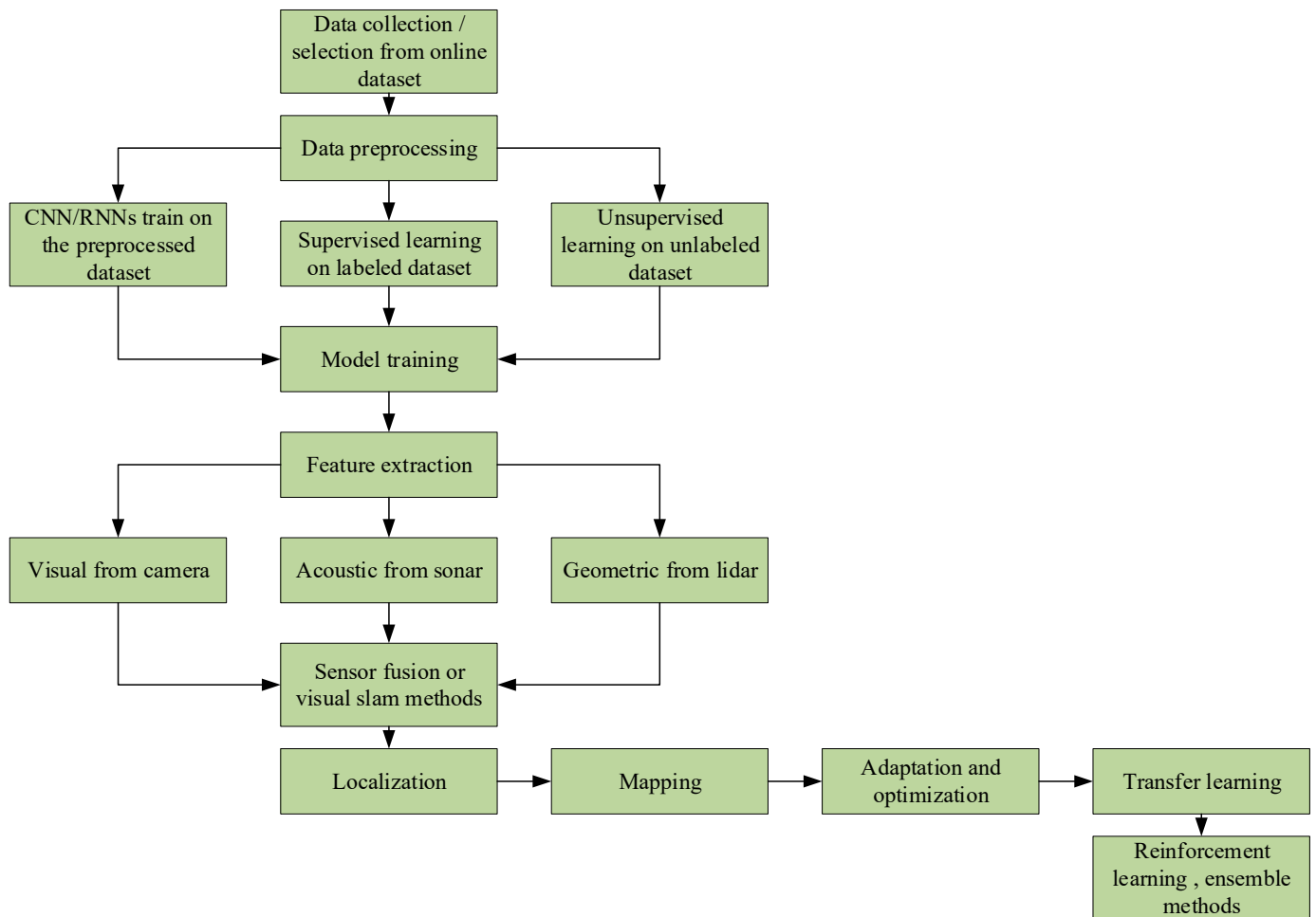

Figure 8. A flowchart depicting the machine learning pipeline for autonomous systems.

The diagram in part 1 of Figure 9 depicts the main techniques used in underwater simultaneous localization and mapping (SLAM) using deep learning. These techniques can be classified into the following three basic categories: (1) Visual SLAM, which refers to a collection of methods used for tasks such as visual odometry (DeepVO) [112], pose estimation (PoseNet) [113], and integrating Doppler velocity log data with visual SLAM (DVL-SLAM); (2) feature extraction and matching, which involves methods like CNN-based feature extractors, self-supervised [69] interest point detection and description (SuperPoint) [114], and feature matching in images (DeepMatcher); and (3) loop closure detection, which includes deep loop closure detection using deep learning [115] and probabilistic place recognition based on appearance (FAB-MAP). These technologies jointly improve the precision and durability of underwater SLAM systems, enabling better navigation and mapping in intricate underwater settings.
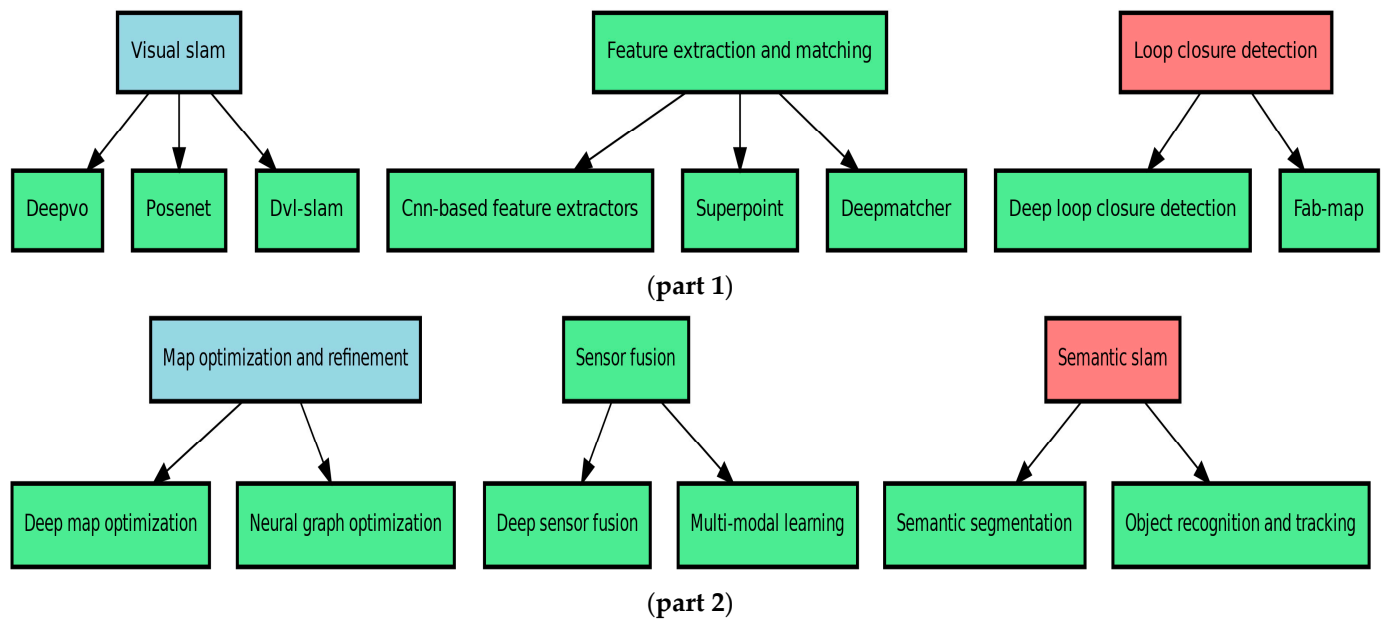
(**part 1**)

(**part 2**)

Figure 9. Introduction to deep learning methods for underwater simultaneous localization and mapping (SLAM) (**part 1**, **part 2**).

Figure 9 part 2 depicts the supplemental deep learning approaches utilized in underwater simultaneous localization and mapping (SLAM), categorized into the following three key domains: (1) Map optimization and refinement, which uses sophisticated methods like deep map optimization and neural graph optimization to enhance and streamline underwater maps. Sensor fusion combines data from various sensors, such as deep sensor fusion [116] and multi-modal learning, to improve the performance of SLAM (simultaneous localization and mapping). Semantic SLAM employs techniques like semantic segmentation to comprehend underwater scenes and object recognition and tracking to detect and monitor things in the surroundings. These strategies are crucial for improving underwater SLAM systems' accuracy, durability, and efficiency.

Figure 10 demonstrates the crucial role of transfer learning [117] in various deep learning techniques used in underwater environments. Transfer learning uses pre-trained models that are subsequently adjusted for specific underwater tasks, including 3D reconstruction, anomaly detection, image classification, acoustic signal processing, obstacle avoidance, SLAM (simultaneous localization and mapping), semantic segmentation, image enhancement, and object detection. Every category uses transfer learning to modify general models for the specific difficulties of undersea applications, ultimately improving performance and efficiency.

The graphic in Figure 11 depicts a range of deep learning techniques used in underwater applications. A concise explanation and practical use in underwater settings accompany every method. Techniques such as data augmentation, self-supervised learning [69,118], unsupervised learning [119], reinforcement learning [120,121], semi-supervised learning [122], domain adaptation, generative adversarial networks (GANs), hybrid models, few-shot learning, active learning, and model ensemble are used to improve the effectiveness and efficiency in underwater environments. From here, our next focus is the comparative analysis of underwater SLAM techniques.
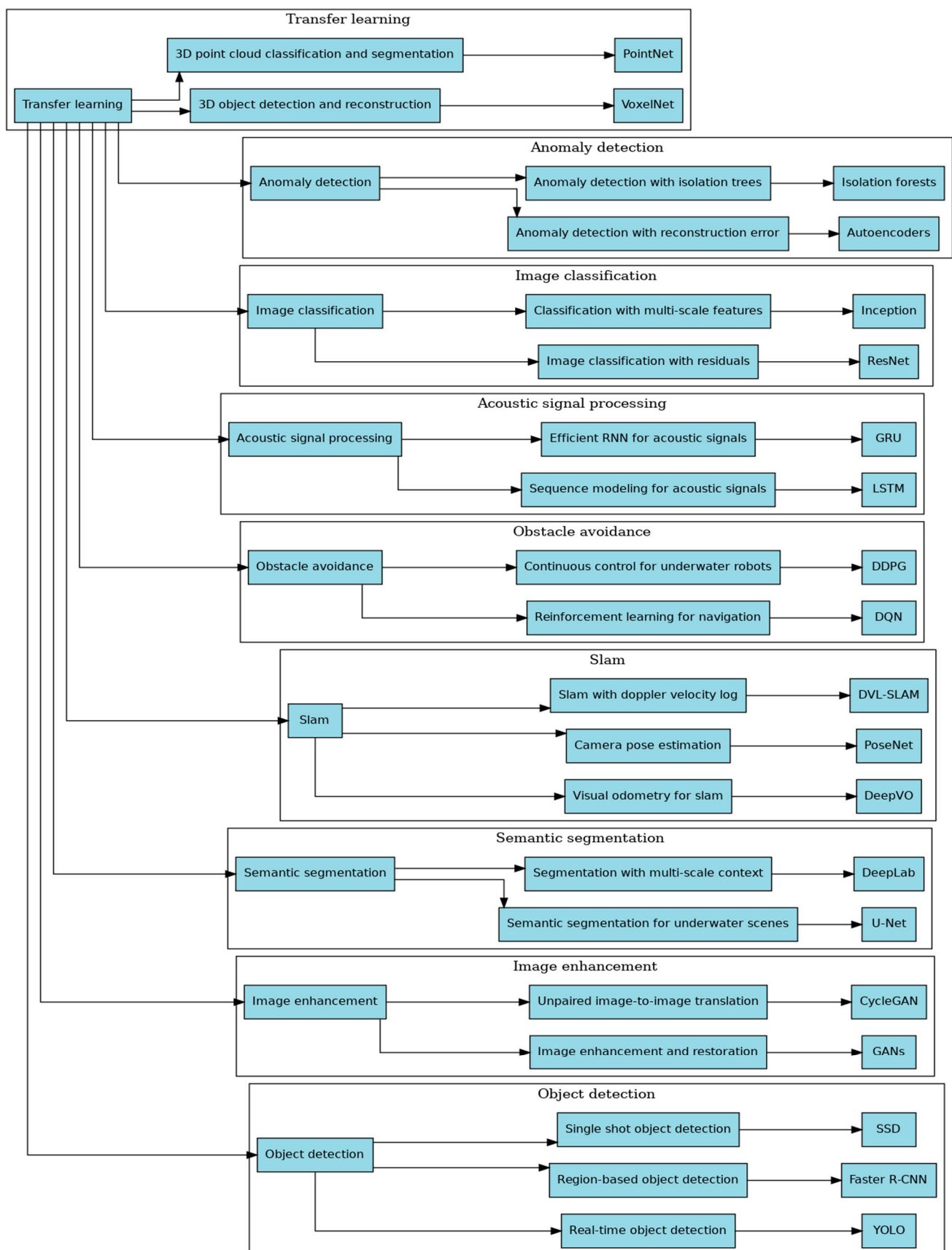
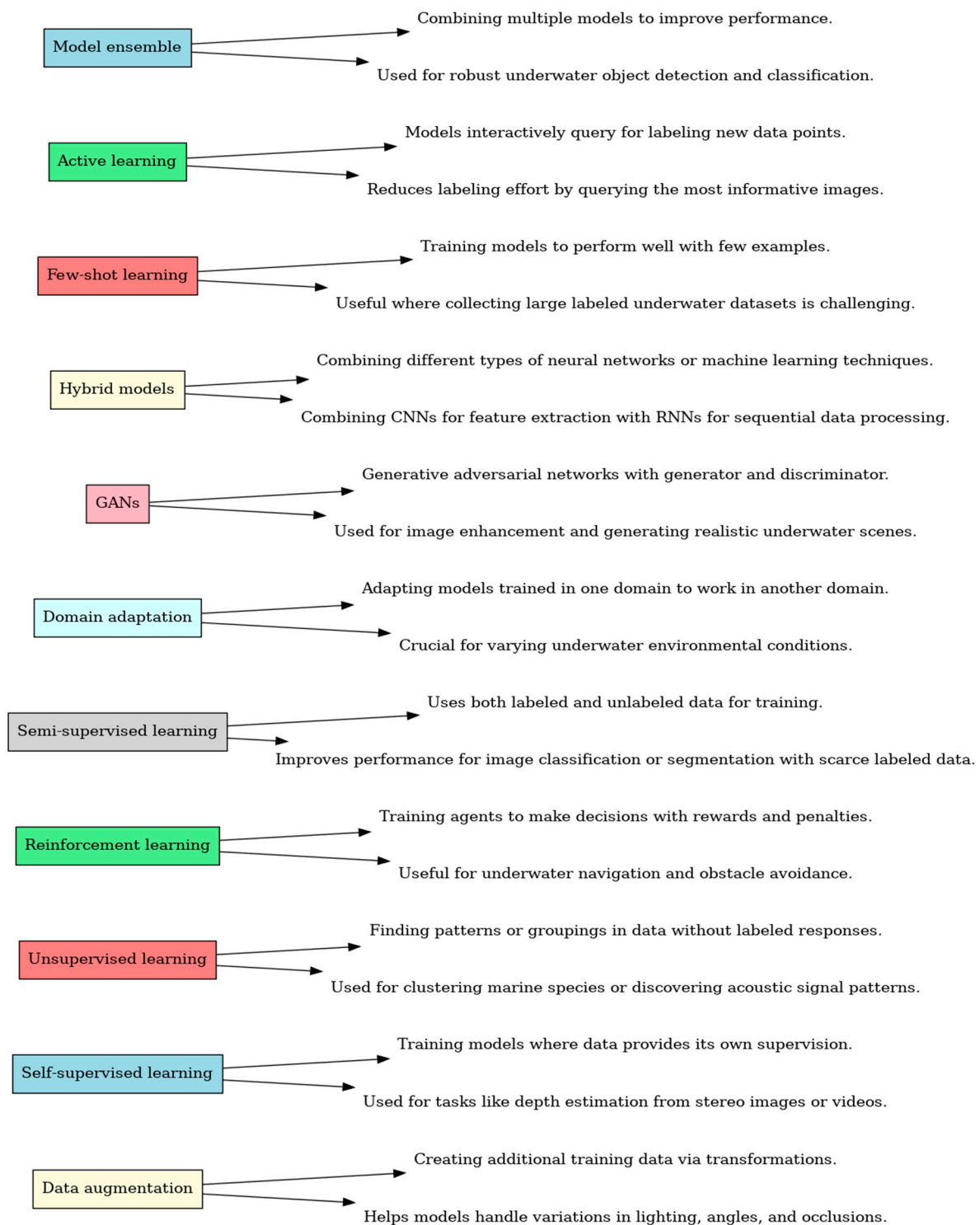**Figure 10.** Utilizing transfer learning to leverage deep learning methods [75] for underwater applications.

Figure 11. Exploring deep learning techniques for underwater applications, excluding transfer learning.

6. Comparative Analysis of Underwater SLAM Techniques

Underwater simultaneous localization and mapping (SLAM) is essential for unmanned underwater vehicles (UUVs) [92], allowing navigation in the absence of a global positioning system (G.P.S.). While promising, ORB-SLAM3 [123] and DF-VO encounter difficulties facing underwater obstacles, such as limited visibility and light absorption. ORB-SLAM3 [123] performs exceptionally well in low-light conditions but lacks robustness in complicated situations, whereas DF-VO provides superior robustness but requires more extensive pro-

cessing resources [124,125]. Dark channel prior (D.C.P.), an image processing technique, enhances ORB-SLAM2 [126] by reducing distortions and improving feature matching [127]. Multi-vehicle mapping, utilizing advanced algorithms such as Gaussian mixture robust branch and bound (GMRBnB), surpasses conventional approaches in map registration. Acoustic-based simultaneous localization and mapping (SLAM) utilizes inertial sensors and sonar to accurately navigate conditions with limited vision. This approach incorporates dead reckoning and Bayesian–Gaussian mixtures to create real-time maps [128]. The limitations of VSLAM [129] are addressed by advancements in low-light picture augmentation and self-supervised learning for feature detection [130]. There are still difficulties in accurately aligning maps and dealing with abnormal data points in multi-vehicle SLAM. Integrating various sensors and complex algorithms is crucial for enhancing accuracy and resilience.

Navigation and mapping in underwater environments are difficult; hence, SLAM techniques must be tested. This comparison compares standard and deep learning-based SLAM systems based on accuracy, computing efficiency, resilience, scalability, and sensor fusion. These criteria were chosen because they are crucial to underwater SLAM performance. After a thorough literature analysis, each algorithm was evaluated using RMSE, processing time, and robustness tests. Deep learning-based DeepVO [112] and GAN-SLAM [27] provide superior feature extraction and mapping accuracy, whereas traditional ORB-SLAM [40] and EKF-SLAM are resilient and accurate in underwater applications. The tables below show that deep learning algorithms are better at feature extraction and posture estimation than standard methods, making them appropriate for precise underwater mapping. This comprehensive comparison shows the present status of SLAM technologies and highlights topics for future research, emphasizing the need for resilient, efficient, and accurate underwater SLAM systems.

The above sections provide a qualitative summary of SLAM techniques and their improvements, but a quantitative comparison is needed to reach a complete understanding. Following are the criteria used to compare SLAM methods in each table:

Table 1 compares underwater-applicable classical SLAM algorithms [36,40,44,48,130] based on accuracy, computational efficiency, resilience, scalability, and sensor fusion.

Table 2 compares underwater-applicable deep learning-based SLAM algorithms based on feature extraction, temporal modeling, data augmentation, pose prediction, and mapping accuracy [131,132].

After this clear comparison, the next interest of our dissertation is to elaborate on the advantages and superiority of deep learning relative to conventional methods.

7. Advantage and Superiority of Deep Learning Relative to the Conventional Method

Figure 12 illustrates the differences in feature extraction performance between traditional methods (solid blue line) and deep learning methods [133,134] (dashed orange line) using hypothetical data. The X-axis represents different data points or feature indices, while the Y-axis indicates the normalized value or significance of the extracted features. Traditional methods show significant variability with pronounced peaks and troughs, suggesting inconsistent feature extraction performance. In contrast, deep learning methods exhibit a more balanced and consistent extraction pattern, highlighting their potential advantage in capturing a broader range of features effectively [135]. The data is hypothetical and intended for conceptual demonstration.

Figure 13 depicts the disparities in the effectiveness of data fusion between conventional techniques (represented by a solid blue line) and deep learning techniques [99,136–140] (represented by a dashed orange line) using hypothetical data. The X-axis depicts distinct data points, while the Y-axis indicates the standardized value of the combined data from two sensors. Conventional techniques employ a basic mean calculation for data fusion, leading to a direct but maybe less precise amalgamation of sensor data. Deep learning [141,142] approaches utilize a weighted average, showcasing a more complex and presumably

more efficient strategy for combining sensor data. The data is fictitious and provided for conceptual purposes.
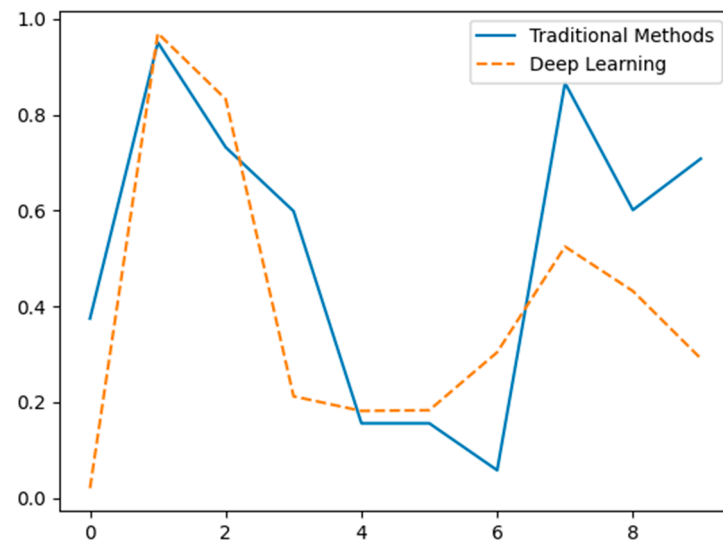

Figure 12. Comparison of feature extraction capabilities between traditional methods and deep learning (hypothetical data).
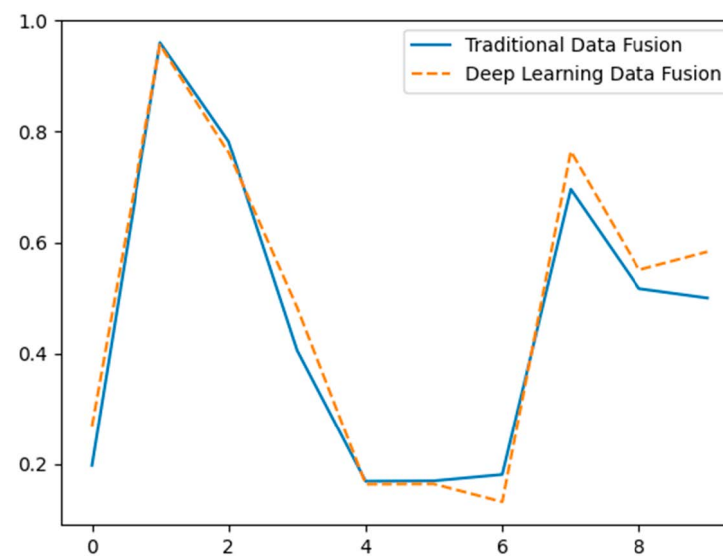

Figure 13. Comparative evaluation of data fusion performance: traditional methods vs. deep learning (hypothetical data).

Figure 14 contrasts the adaptation of classical approaches (solid blue line) versus deep learning methods (dashed orange line) to new contexts using hypothetical data. The X-axis depicts various data points or environmental variables, while the Y-axis shows the normalized adaptation performance statistic. Traditional approaches exhibit heterogeneity in adaptation, indicating possible difficulties in changing to new environments without retraining. On the other hand, deep learning approaches show greater consistency and adaptability, emphasizing their ability to learn and respond more efficiently to changing settings. The data is fictitious and designed for conceptual presentation.
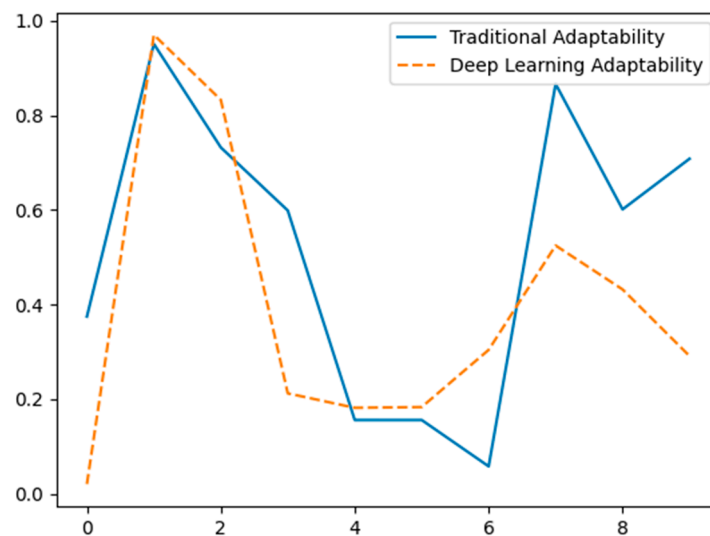
Figure 14. Adaptation to new environments comparison of traditional methods vs. deep learning (hypothetical data).

Figure 15 compares the real-time processing performance between classical approaches (represented by a solid blue line) and deep learning methods (represented by a dashed orange line) using hypothetical data. The X-axis shows discrete time intervals or occurrences, whereas the Y-axis indicates the standardized processing durations. Conventional techniques demonstrate longer processing times, indicating a slower performance. On the other hand, deep learning techniques have considerably shorter processing durations, suggesting their capacity to handle data more effectively in real time. The data provided is fictitious and is intended solely for conceptual demonstration purposes.
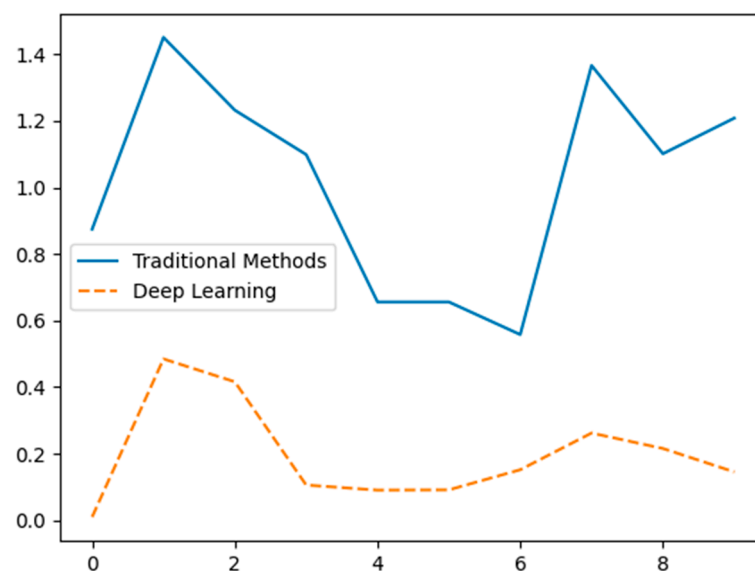

Figure 15. Comparison of real-time processing performance between traditional methods and deep learning using hypothetical data.

Figure 16 illustrates the comparative capacities of classical approaches (represented by a solid blue line) and deep learning methods [143,144] (represented by a dashed orange line). The data used in this comparison is hypothetical. The X-axis depicts varying degrees of scene complexity, while the Y-axis indicates the standardized performance in comprehending and analyzing the situations. Conventional approaches exhibit inconsistent

results, frequently encountering difficulties when dealing with more complicated scenes. Deep learning [112,114,145] approaches, on the other hand, provide exceptional performance [146], especially in intricate environments, emphasizing their capacity to capture and comprehend advanced semantic information efficiently. The data provided is fictitious and is intended solely for conceptual demonstration purposes.
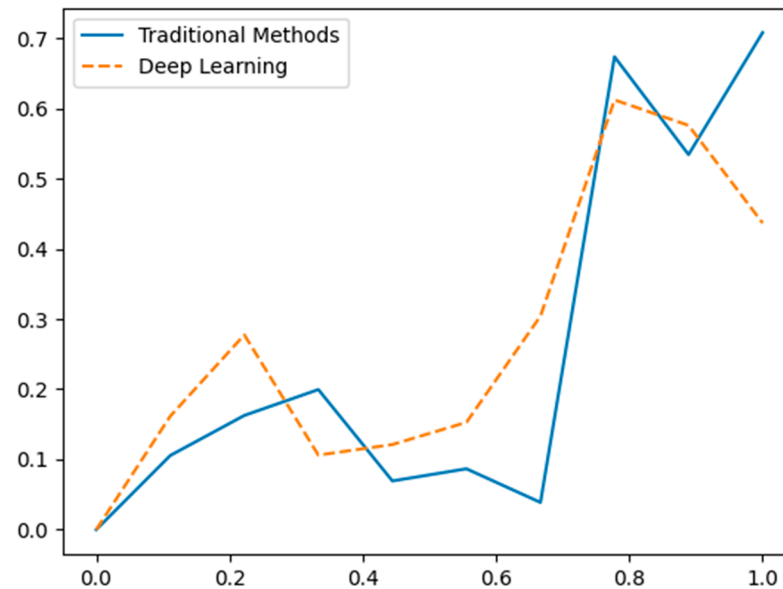

Figure 16. Comparison of the performance in high-level scene understanding between traditional methods and deep learning using hypothetical data.

The hypothetical data results reveal various advantages of deep learning over previous approaches in underwater SLAM navigation. First, deep learning approaches extract features more reliably and precisely than classical methods, which have abrupt peaks and troughs. Understanding underwater settings requires reliable feature extraction. Additionally, deep learning algorithms are more noise-resistant. Traditional approaches [147] are more noise-sensitive, causing more significant feature value variations. In contrast, deep learning approaches, as shown in Figures 14–16, preserve more stable and consistent values in noise, improving performance in noisy underwater settings. While straightforward, traditional data fusion approaches employ averaging, which can be inaccurate. Deep learning approaches use weighted averaging to integrate sensor data sources and improve SLAM system accuracy and reliability.

Deep learning also excels at adaptability to hypothetical data, as illustrated in Figures 14–16 [148]. Deep learning algorithms [131] adapt to new surroundings more reliably than traditional methods. Due to its higher learning capabilities, deep learning is ideal for dynamic underwater situations. Deep learning algorithms outperform standard methods in complex scenarios for high-level scene understanding. Deep learning algorithms improve interpretation and decision-making in complicated scenes, while traditional methods do poorly in such cases.

Another benefit of deep learning is real-time processing. Traditional methods take longer to process, indicating slower performance. Deep learning algorithms [20] have far reduced processing times, indicating more efficient real-time processing, which is crucial for underwater SLAM applications. Hypothetical data illustrates this in Figures 14–16.

Finally, traditional methods vary in retaining excellent perception and mapping quality. However, deep learning algorithms provide more accurate and detailed underwater maps across perceptual quality levels.

Therefore, deep learning in underwater SLAM navigation could increase autonomous underwater vehicle accuracy, reliability, and efficiency. Traditional methods have draw-

backs, but deep learning uses complicated models and data-driven approaches to overcome them. These algorithms can learn from big datasets, adapt to changing conditions, and integrate numerous sensor data sources better, making them ideal for underwater SLAM navigation. Further research should address constraints like computing requirements and vast training datasets to fully realize these strategies' benefits in real-world applications.

Deep learning brings several critical advantages to underwater SLAM (simultaneous localization and mapping) navigation, addressing many limitations of traditional methods.

Firstly, feature extraction is a significant benefit, as deep learning models excel at extracting complex and high-level features from raw sensor data [58]. This is particularly useful in underwater environments, where traditional methods often struggle with noise, low visibility, and the lack of distinctive features. Deep learning allows the extraction of more informative and reliable features, even in challenging conditions.

Secondly, deep learning [136,137] algorithms show robustness to noise. Deep learning models can maintain accurate positioning and mapping in underwater settings, where environmental noise and uncertainties often affect sensors. This results in improved SLAM accuracy and system reliability in situations where traditional methods would falter.

Another critical advantage is data fusion. Deep learning [149,150] models can integrate data from multiple types of sensors, such as sonar, LiDAR, and cameras. This ability enhances the SLAM system's overall performance by providing more robust positioning and mapping, even in cases where individual sensors may give incomplete or erroneous data.

In addition, deep learning [136–139] enables adaptability and learning. Unlike traditional SLAM approaches that rely on predefined models and assumptions about the environment, deep learning models can learn from the data and adapt to new environments over time. This makes them more flexible and scalable in dynamic underwater conditions where the environment can change unpredictably.

Moreover, deep learning models excel at handling complex dynamics. Underwater environments are highly dynamic and nonlinear, often making it difficult for traditional methods to cope. Deep learning models, with their ability to model complex patterns, improve the accuracy of trajectory estimation and mapping in such settings. Because of these operational advantages, deep learning is superior in understanding complex underwater scenes.

For example, deep learning models such as CNNs (convolutional neural networks) can highly understand scenes. This allows them to interpret and analyze semantic information from underwater scenes that may be too complex for traditional methods, leading to more accurate scene interpretation.

Deep learning [69] also enables end-to-end learning, where raw sensor inputs are processed directly into SLAM outputs without needing handcrafted feature extraction. This reduces complexity in the system design and allows for a more seamless, automated approach to SLAM in underwater settings.

Additionally, deep learning models demonstrate a high level of generalization across environments. Models trained on large and diverse datasets are more likely to adapt to new environments with minimal tuning than traditional methods requiring extensive adjustments to function effectively [151] in different underwater terrains.

Another key benefit is the capability for real-time processing. Deep learning algorithms can process complex underwater scenes in real time, making them suitable for dynamic and time-sensitive SLAM applications, such as autonomous underwater vehicle (AUV) navigation.

Finally, deep learning enhances perception and mapping accuracy. By using deep learning algorithms, SLAM systems can produce more detailed and accurate underwater maps improving perception and navigation [152].

8. Conclusions/Significance

While our work summarizes the advances made possible by deep learning in underwater SLAM, several issues still need to be further explored. Deep learning methods have

a limited real-time application in dynamic underwater environments because of their high computational resource requirements, even though they are resilient in feature extraction and noise treatment. The requirement for extensive and varied datasets for training presents another difficulty because gathering such data in underwater environments is expensive and challenging. Even with sophisticated algorithms, environmental disturbances, including shifting circumstances, visibility problems, and sensor noise, still impact performance. Furthermore, unmanned underwater vehicles (UUVs) present particular concerns about the high energy consumption linked to operating complicated deep learning models, as energy economy is crucial for prolonged missions. Finally, despite the significant advances in mapping precision and flexibility that deep learning provides, hardware constraints still limit real-time processing. These difficulties show that to fully realize the potential of deep learning in underwater SLAM applications, more research is required in model optimization, effective data gathering, improving data availability through simulations, and exploring hybrid approaches combining deep learning with traditional underwater systems and hardware solutions.

This paper lays a solid groundwork for the progress of unmanned underwater vehicle (UUV) navigation, with a specific emphasis on enhancing AI-SLAM algorithms, specifically those powered by deep learning. Future research will focus on improving various sensor fusion approaches and integrating advanced technologies such as multibeam sonar, stereo cameras, LiDAR, and Imu, as well as methodologies like SBL/USBL [153,154] and DLV [155]. The goal should be to enhance the accuracy of UUV navigation in challenging underwater situations.

Investigating the incorporation of cutting-edge technology like machine learning and enhanced computer vision shows potential for improving the reliability of UUV navigation systems. Ensuring the capacity to adapt to various unmanned underwater vehicles (UUVs) and mission needs is critical. It is imperative to have cooperation between researchers, industry professionals, and policymakers to establish standards and effectively apply these improvements in real-world scenarios.

This work has thoroughly examined underwater simultaneous localization and mapping (SLAM) technologies, specifically focusing on incorporating deep learning methods. We have emphasized their crucial function in improving underwater navigation and perception by analyzing different sensors and their respective uses, such as vision sensors, sonar, and LiDAR. The assessment of various simultaneous localization and mapping (SLAM) algorithms highlighted the progress and difficulties in the field, namely the advantages of integrating deep learning techniques such as convolutional neural networks (CNNs), long short-term memory (LSTM) networks, generative adversarial networks (GANs) and other deep learning methods.

We have shown that the utilization of deep learning greatly enhances the process of extracting features, estimating poses, and fusing data, resulting in underwater SLAM systems that are more precise and resilient. Nevertheless, these systems have difficulties, such as intense computational requirements and the opaque nature of neural networks, which can affect the capacity to interpret and apply them in real time.

Author Contributions: Author Contributions: Conceptualization F.F.R.M. and B.J.; methodology, F.F.R.M. and B.J.; software, F.F.R.M.; validation, B.J., B.F. and Z.X.; formal analysis, F.F.R.M. and B.J.; investigation, B.F.; resources, B.J.; data curation, Z.X.; writing—original draft preparation, B.F.; writing—review and editing, Z.X.; visualization, F.F.R.M.; supervision, B.J. and Z.X; project administration, F.F.R.M.; funding acquisition, B.J. All authors have read and agreed to the published version of the manuscript.

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

References

1. Xu, Z.; Haroutunian, M.; Murphy, A.J.; Neasham, J.; Norman, R. An underwater visual navigation method based on multiple aruco markers. *J. Mar. Sci. Eng.* **2021**, *9*, 1432. [CrossRef]
2. Sun, K.; Cui, W.; Chen, C. Review of underwater sensing technologies and applications. *Sensors* **2021**, *21*, 7849. [CrossRef] [PubMed]
3. Zhang, Y.; Wu, Y.; Tong, K.; Chen, H.; Yuan, Y. Review of Visual Simultaneous Localization and Mapping Based on Deep Learning. *Remote Sens.* **2023**, *15*, 2740. [CrossRef]
4. Azam, A.B.; Kong, Z.J.; Ng, S.Y.; Florian, M.S.; Elhadidi, B.; Seet, G.; Zheng, J.; Cai, Y. Low-cost Underwater Localisation Using Single-Beam Echosounders and Inertial Measurement Units. In Proceedings of the OCEANS 2023—Limerick, Limerick, Ireland, 5–8 June 2023; Institute of Electrical and Electronics Engineers (IEEE): New York, NY, USA, 2023; pp. 1–7. [CrossRef]
5. Nauert, F.; Kampmann, P. Inspection and maintenance of industrial infrastructure with autonomous underwater robots. *Front. Robot. AI* **2023**, *10*, 1240276. [CrossRef] [PubMed]
6. Orinaitė, U.; Karaliūtė, V.; Pal, M.; Ragulskis, M. Detecting Underwater Concrete Cracks with Machine Learning: A Clear Vision of a Murky Problem. *Appl. Sci.* **2023**, *13*, 7335. [CrossRef]
7. Zhang, E.; Jiang, T.; Duan, J. A Multi-Stage Feature Aggregation and Structure Awareness Network for Concrete Bridge Crack Detection. *Sensors* **2024**, *24*, 1542. [CrossRef]
8. Yang, H.; Xu, Z.; Jia, B. An Underwater Positioning System for UUVs Based on LiDAR Camera and Inertial Measurement Unit. *Sensors* **2022**, *22*, 5418. [CrossRef]
9. Tani, S.; Ruscio, F.; Bresciani, M.; Nordfeldt, B.M.; Bonin-Font, F.; Costanzi, R. Development and testing of a navigation solution for Autonomous Underwater Vehicles based on stereo vision. *Ocean Eng.* **2023**, *280*, 114757. [CrossRef]
10. Hou, J.; Ye, X. Real-time Underwater 3D Reconstruction Method Based on Stereo Camera. In Proceedings of the 2022 IEEE International Conference on Mechatronics and Automation, ICMA 2022, Guilin, China, 7–10 August 2022; Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA, 2022; pp. 1204–1209. [CrossRef]
11. Vargas, E.; Scona, R.; Willners, J.S.; Luczynski, T.; Cao, Y.; Wang, S.; Petillot, Y.R. Robust Underwater Visual SLAM Fusing Acoustic Sensing. In Proceedings of the IEEE International Conference on Robotics and Automation, Xi'an, China, 30 May–5 June 2021; Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA, 2021; pp. 2140–2146. [CrossRef]
12. Aravind, J.V.; Ganesh, K.V.S.S.; Prince, S. Real-Time Appearance Based Mapping using Visual Sensor for Unknown Environment. *J. Phys. Conf. Ser.* **2022**, *2335*, 012057. [CrossRef]
13. Mo, J. Towards a Fast, Robust and Accurate Visual-Inertial Simultaneous Localization and Mapping System. Doctoral Dissertation, University of Minnesota, Minneapolis, MN, USA, 2022.
14. Li, S.; Li, Z.; Liu, X.; Shan, C.; Zhao, Y.; Cheng, H. Research on Map-SLAM Fusion Localization Algorithm for Unmanned Vehicle. *Appl. Sci.* **2022**, *12*, 8670. [CrossRef]
15. Sasi, J.P.; Pandagre, K.N.; Royappa, A.; Walke, S.; Pavithra, G.; Natrayan, L. Deep Learning Techniques for Autonomous Navigation of Underwater Robots. In Proceedings of the 2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), Gautam Buddha Nagar, India, 1–3 December 2023; IEEE: New York, NY, USA, 2023; pp. 1630–1635. [CrossRef]
16. Al-Tawil, B.; Hempel, T.; Abdelrahman, A.; Al-Hamadi, A. A review of visual SLAM for robotics: Evolution, properties, and future applications. *Front. Robot. AI* **2024**, *11*, 1347985. [CrossRef] [PubMed]
17. Yan, T.; Xu, Z.; Yang, S.X.; Gadsden, S.A. Formation control of multiple autonomous underwater vehicles: A review. *Intell. Robot.* **2023**, *3*, 1–22. [CrossRef]
18. Chen, D.; Huang, B.; Kang, F. A Review of Detection Technologies for Underwater Cracks on Concrete Dam Surfaces. *Appl. Sci.* **2023**, *13*, 3564. [CrossRef]
19. Noor, A.; Ruhaiyem, N.I. Underwater image processing based on CNN applications: A review. In Proceedings of the Cognitive Models and Artificial Intelligence Conference, İstanbul, Turkey, 25–26 May 2024; ACM: New York, NY, USA, 2024; pp. 75–84. [CrossRef]
20. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [CrossRef]
21. Lyernisha, S.R.; Christopher, C.S.; Fernisha, S.R. Object recognition from enhanced underwater image using optimized deep-CNN. *Int. J. Wavelets Multiresolution Inf. Process* **2023**, *21*, 2350007. [CrossRef]
22. Amarasinghe, C.; Ratnaweera, A.; Maitripala, S. UW Deep SLAM-CNN Assisted Underwater SLAM. *Appl. Comput. Syst.* **2023**, *28*, 100–113. [CrossRef]
23. Qi, H.; Wang, C.; Li, J.; Shi, L. Loop Closure Detection with CNN in RGB-D SLAM for Intelligent Agricultural Equipment. *Agriculture* **2024**, *14*, 949. [CrossRef]
24. Khandouzi, A.; Ezoji, M. Coarse-to-fine underwater image enhancement with lightweight CNN and attention-based refinement. *J. Vis. Commun. Image Represent.* **2024**, *99*, 104068. [CrossRef]
25. Munoz, B.; Troni, G. Learning the Ego-Motion of an Underwater Imaging Sonar: A Comparative Experimental Evaluation of Novel CNN and RCNN Approaches. *IEEE Robot. Autom. Lett.* **2024**, *9*, 2072–2079. [CrossRef]

26. Zheng, Z.; Xin, Z.; Yu, Z.; Yeung, S.-K. Real-time GAN-based image enhancement for robust underwater monocular SLAM. *Front. Mar. Sci.* **2023**, *10*, 1161399. [CrossRef]

27. Savinykh, A.; Kurenkov, M.; Kruzhkov, E.; Yudin, E.; Potapov, A.; Karpyshev, P.; Tsetserukou, D. DarkSLAM: GAN-Assisted Visual SLAM for Reliable Operation in Low-Light Conditions. *arXiv* **2022**, arXiv:2206.02199.

28. Estrada, D.C.; Dalgleish, F.R.; Den Ouden, C.J.; Ramos, B.; Li, Y.; Ouyang, B. Underwater LiDAR Image Enhancement Using a GAN Based Machine Learning Technique. *IEEE Sens. J.* **2022**, *22*, 4438–4451. [CrossRef]

29. Shorten, C.; Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*, 60. [CrossRef]

30. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengiot, Y. Generative Adversarial Nets. Available online: http://www.github.com/goodfeli/adversarial (accessed on 6 October 2024).

31. Wang, M.; Deng, W. Deep Visual Domain Adaptation: A Survey. *Neurocomputing* **2018**, *312*, 135–153. [CrossRef]

32. Eitel, A.; Springenberg, J.Y.; Spinello, L.; Riedmiller, M.; Burgard, W. Multimodal deep learning for robust RGB-D object recognition. In Proceedings of the IROS Hamburg 2015 Conference Digest: IEEE/RSJ International Conference on Intelligent Robots and Systems, Hamburg, Germany, 28 September–2 October 2015.

33. Davison, A.J. Real-Time Simultaneous Localisation and Mapping with a Single Camera. 2003. Available online: http://www.robots.ox.ac.uk/~ajd/ (accessed on 6 October 2024).

34. Zuo, X.; Jiang, X.; Li, P.; Wang, J.; Ge, M.; Schuh, H. A square root information filter for multi-GNSS real-time precise clock estimation. *Satell. Navig.* **2021**, *2*, 28. [CrossRef]

35. Liu, T.; Xv, C.; Qiao, Y.; Jiang, C.; Yu, J. Particle Filter SLAM for Vehicle Localization. *J. Ind. Eng. Appl. Sci.* **2024**, *2*, 27–31. [CrossRef]

36. Grisetti, G.; Kummerle, R.; Stachniss, C.; Burgard, W. A tutorial on graph-based SLAM. *IEEE Intell. Transp. Syst. Mag.* **2010**, *2*, 31–43. [CrossRef]

37. Wang, Q.; Yuan, X.; Xu, C.; Wang, X. A Bayesian Approach to Communication-Driven SLAM Based on Diffuse Reflection Model. *IEEE Wirel. Commun. Lett.* **2023**, *12*, 1279–1283. [CrossRef]

38. Balemans, N.; Hellinckx, P.; Latre, S.; Reiter, P.; Steckel, J. S2L-SLAM: Sensor Fusion Driven SLAM using Sonar, LiDAR and Deep Neural Networks. In Proceedings of the IEEE Sensors, Sydney, Australia, 31 October–3 November 2021; Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA, 2021. [CrossRef]

39. Hu, C.; Zhu, S.; Song, W. Real-time Underwater 3D Reconstruction Based on Monocular Image. In Proceedings of the 2022 IEEE International Conference on Robotics and Biomimetics (ROBIO), Jinghong, China, 5–9 December 2022; pp. 1239–1244.

40. Mur-Artal, R.; Montiel, J.M.M.; Tardos, J.D. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *IEEE Trans. Robot.* **2015**, *31*, 1147–1163. [CrossRef]

41. Bloesch, M.; Omari, S.; Hutter, M.; Siegwart, R. Robust Visual Inertial Odometry Using a Direct EKF-Based Approach. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 28 September–2 October 2015. [CrossRef]

42. Engel, J.; Schöps, T.; Cremers, D. LNCS 8690—LSD-SLAM: Large-Scale Direct Monocular SLAM. In *Computer Vision—ECCV 2014, Proceedings of the 13th European Conference, Zurich, Switzerland, 6–12 September 2014*; Springer: Berlin/Heidelberg, Germany, 2014.

43. Kerl, C.; Sturm, J.; Cremers, D. Dense Visual SLAM for RGB-D Cameras. In Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, Tokyo, Japan, 3–7 November 2013.

44. Engel, J.; Koltun, V.; Cremers, D. Direct Sparse Odometry. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 611–625. [CrossRef]

45. Mourikis, A.I.; Roumeliotis, S.I. A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation. In Proceedings of the 2007 IEEE International Conference on Robotics and Automation, Rome, Italy, 10–14 April 2007.

46. Abdollahi, M.R.; Pourtakdoust, S.H.; Nooshabadi, M.H.Y.; Pishkenari, H.N. An Improved Multi-State Constraint Kalman Filter for Visual-Inertial Odometry. Available online: https://youtu.be/aE6_av59QXw (accessed on 6 October 2024).

47. Forster, C.; Pizzoli, M.; Scaramuzza, D. SVO: Fast Semi-Direct Monocular Visual Odometry. Available online: http://rpg.ifi.uzh.ch (accessed on 6 October 2024).

48. Forster, C.; Zhang, Z.; Gassner, M.; Werlberger, M.; Scaramuzza, D. SVO: Semidirect visual odometry for monocular and multicamera systems. *IEEE Trans. Robot.* **2016**, *33*, 249–265. [CrossRef]

49. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]

50. Li, M.; Mourikis, A.I. High-Precision, Consistent EKF-based Visual-Inertial Odometry. *Int. J. Robot. Res.* **2013**, *32*, 690–711. [CrossRef]

51. Cummins, M.; Newman, P. FAB-MAP: Probabilistic Localization and Mapping in the Space of Appearance. *Int. J. Robot. Res.* **2008**, *27*, 647–665. [CrossRef]

52. Yang, W. Researches Advanced in Autonomous Underwater Robots based on SLAM. In Proceedings of the 2022 4th International Conference on Robotics, Intelligent Control and Artificial Intelligence, Dongguan, China, 16–18 December 2022; ACM: New York, NY, USA, 2022; pp. 200–203. [CrossRef]

53. He, W.; Li, R.; Liu, T.; Yu, Y. LiDAR-based SLAM pose estimation via GNSS graph optimization algorithm. *Meas. Sci. Technol.* **2024**, *35*, 096304. [CrossRef]

54. Chen, X.; Läbe, T.; Milioto, A.; Röhling, T.; Vysotska, O.; Haag, A.; Behley, J.; Stachniss, C. OverlapNet: Loop Closing for LiDAR-based SLAM. In Proceedings of the Robotics: Science and Systems 2020, Corvalis, OR, USA, 12–16 July 2020. [CrossRef]

55. Chang, S.; Wan, C.; Zhang, D.; Li, H.; Lin, Y. An Underwater SLAM Approach Using Regularly Distributed Magnetic Beacons. In *Advances in Guidance, Navigation and Control, Proceedings of the 2022 International Conference on Guidance, Navigation and Control, Tianjin, China, 5–7 August 2022*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 300–308. [CrossRef]

56. Qin, T.; Li, P.; Shen, S. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator. *IEEE Trans. Robot.* **2018**, *34*, 1004–1020. [CrossRef]

57. Wang, X.; Fan, X.; Shi, P.; Ni, J.; Zhou, Z. An Overview of Key SLAM Technologies for Underwater Scenes. *Remote Sens.* **2023**, *15*, 2496. [CrossRef]

58. Xin, Z.; Wang, Z.; Yu, Z.; Zheng, B. ULL-SLAM: Underwater low-light enhancement for the front-end of visual SLAM. *Front. Mar. Sci.* **2023**, *10*, 1133881. [CrossRef]

59. Zhang, F.; Xu, D.; Cheng, C. An Underwater Distributed SLAM Approach Based on Improved GMRBnB Framework. *J. Mar. Sci. Eng.* **2023**, *11*, 2271. [CrossRef]

60. Bucci, A.; Franchi, M.; Ridolfi, A.; Secciani, N.; Allotta, B. Evaluation of UKF-Based Fusion Strategies for Autonomous Underwater Vehicles Multisensor Navigation. *IEEE J. Ocean. Eng.* **2023**, *48*, 1–26. [CrossRef]

61. Viset, F.; Helmons, R.; Kok, M. An Extended Kalman Filter for Magnetic Field SLAM Using Gaussian Process Regression. *Sensors* **2022**, *22*, 2833. [CrossRef] [PubMed]

62. Eraghi, H.E.; Taban, M.R.; Bahreinian, S.F. Improved Unscented Kalman Filter Algorithm to Increase the SLAM Accuracy. In Proceedings of the 2023 9th International Conference on Control, Instrumentation and Automation (ICCIA), Tehran, Iran, 20–21 December 2023; IEEE: New York, NY, USA, 2023; pp. 1–5. [CrossRef]

63. Li, H.; Wang, G.; Li, X.; Lian, Y. A Review of Underwater SLAM Technologies. In Proceedings of the 2023 5th International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI), Hangzhou, China, 1–3 December 2023; IEEE: New York, NY, USA, 2023; pp. 215–225. [CrossRef]

64. Chang, S.; Zhang, D.; Zhang, L.; Zou, G.; Wan, C.; Ma, W.; Zhou, Q. A Joint Graph-Based Approach for Simultaneous Underwater Localization and Mapping for AUV Navigation Fusing Bathymetric and Magnetic-Beacon-Observation Data. *J. Mar. Sci. Eng.* **2024**, *12*, 954. [CrossRef]

65. Teng, M.; Ye, L.; Yuxin, Z.; Yanqing, J.; Qianyi, Z.; Pascoal, A.M. Efficient Bathymetric SLAM with Invalid Loop Closure Identification. *IEEE/ASME Trans. Mechatron.* **2021**, *26*, 2570–2580. [CrossRef]

66. Xu, Z.; Qiu, H.; Dong, M.; Wang, H.; Wang, C. Underwater Simultaneous Localization and Mapping Based on 2D-SLAM Framework. In Proceedings of the 2022 IEEE International Conference on Unmanned Systems (ICUS), Guangzhou, China, 28–30 October 2022; IEEE: New York, NY, USA, 2022; pp. 184–189. [CrossRef]

67. Xu, S.; Ma, T.; Li, Y.; Ding, S.; Gao, J.; Xia, J.; Qi, H.; Gu, H. An effective stereo SLAM with high-level primitives in underwater environment. *Meas. Sci. Technol.* **2023**, *34*, 105405. [CrossRef]

68. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv* **2017**, arXiv:1704.04861.

69. Jing, L.; Tian, Y. Self-Supervised Visual Feature Learning with Deep Neural Networks: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 4037–4058. [CrossRef]

70. Martinho, L.A.; Calvalcanti, J.M.B.; Pio, J.L.S.; Oliveira, F.G. Diving into Clarity: Restoring Underwater Images using Deep Learning. *J. Intell. Robot. Syst.* **2024**, *110*, 32. [CrossRef]

71. Maitre, A.R.; Abin, D. Improved Deep Learning Approach for Underwater Image Enhancement. In *Information and Communication Technology for Competitive Strategies (ICTCS 2022)*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 41–53. [CrossRef]

72. Yao, X.; He, F.; Wang, B. Deep learning-based recurrent neural network for underwater image enhancement. In Proceedings of the Sixth Conference on Frontiers in Optical Imaging and Technology: Imaging Detection and Target Recognition, Nanjing, China, 22–24 October 2023; Xu, J., Zuo, C., Eds.; SPIE: Philadelphia, PA, USA, 2024; p. 61. [CrossRef]

73. Sangari, M.S.; Thangaraj, K.; Vanitha, U.; Srikanth, N.; Sathyamoorthy, J.; Renu, K. Deep learning-based Object Detection in Underwater Communications System. In Proceedings of the 2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichirappalli, India, 5–7 April 2023; IEEE: New York, NY, USA, 2023; pp. 1–6. [CrossRef]

74. Sree Vidhya, K.S.; Deepthi, P.S. A Comprehensive Analysis of Underwater Image Processing based on Deep Learning Techniques. In Proceedings of the 2023 International Conference on Control, Communication and Computing (ICCC), Thiruvananthapuram, India, 19–21 May 2023; IEEE: New York, NY, USA, 2023; pp. 1–6. [CrossRef]

75. Khan, S.; Rahmani, H.; Shah, S.A.A.; Bennamoun, M.; Medioni, G.; Dickinson, S. *A Guide to Convolutional Neural Networks for Computer Vision*; Springer: Cham, Switzerland, 2018.

76. Teixeira, B.; Silva, H.; Matos, A.; Silva, E. Deep Learning for Underwater Visual Odometry Estimation. *IEEE Access* **2020**, *8*, 44687–44701. [CrossRef]

77. Wang, G.; Lin, H.; Wang, Q. Research on underwater target tracking method combining deep learning and kernel correlation filtering. In Proceedings of the 2024 5th International Conference on Computer Vision, Image and Deep Learning (CVIDL), Zhuhai, China, 19–21 April 2024; IEEE: New York, NY, USA, 2024; pp. 1458–1461. [CrossRef]

78. Yasir, M.; Rahman, A.U.; Gohar, M. Habitat mapping using deep neural networks. *Multimed. Syst.* **2021**, *27*, 679–690. [CrossRef]

79. Huang, R.-J.; Lai, Y.-C.; Tsao, C.-Y.; Kuo, Y.-P.; Wang, J.-H.; Chang, C.-C. Applying convolutional networks to underwater tracking without training. In Proceedings of the 2018 IEEE International Conference on Applied System Invention (ICASI), Chiba, Japan, 13–17 April 2018; IEEE: New York, NY, USA, 2018; pp. 342–345. [CrossRef]

80. Guerrero-Gonzalez, A.; Garcia-Cordova, F.; Gilabert, J. A biologically inspired neural network for navigation with obstacle avoidance in autonomous underwater and surface vehicles. In Proceedings of the OCEANS 2011 IEEE—Spain, Santander, Spain, 6–9 June 2011; IEEE: New York, NY, USA, 2011; pp. 1–8. [CrossRef]

81. Santora, M.; Alberts, J.; Edwards, D. Control of Underwater Autonomous Vehicles Using Neural Networks. In Proceedings of the OCEANS 2006, Boston, MA, USA, 18–21 September 2006; IEEE: New York, NY, USA, 2006; pp. 1–5. [CrossRef]

82. Wang, H.; Lu, B.; Li, J.; Liu, T.; Xing, Y.; Lv, C.; Cao, D.; Li, J.; Zhang, J.; Hashemi, E. Risk Assessment and Mitigation in Local Path Planning for Autonomous Vehicles with LSTM Based Predictive Model. *IEEE Trans. Autom. Sci. Eng.* **2022**, *19*, 2738–2749. [CrossRef]

83. Villacorta-Atienza, J.A.; Makarov, V.A. Neural Network Architecture for Cognitive Navigation in Dynamic Environments. *IEEE Trans. Neural Netw. Learn. Syst.* **2013**, *24*, 2075–2087. [CrossRef] [PubMed]

84. Pan, W.; Lv, B.; Peng, L. Research on AUV navigation state prediction method using multihead attention mechanism in a CNN-BiLSTM model. In Proceedings of the Seventh International Conference on Advanced Electronic Materials, Computers, and Software Engineering (AEMCSE 2024), Nanchang, China, 10–12 May 2024; Yang, L., Ed.; SPIE: New York, NY, USA, 2024; p. 105. [CrossRef]

85. Liang, Z.; Wang, K.; Zhang, J.; Zhang, F. An Underwater Multisensor Fusion Simultaneous Localization and Mapping System Based on Image Enhancement. *J. Mar. Sci. Eng.* **2024**, *12*, 1170. [CrossRef]

86. Yang, X.; Chen, J.; Wang, Z.; Zhang, Q.; Liu, W.; Liao, C.; Cheng, K.-T. Monocular Camera Based Real-Time Dense Mapping Using Generative Adversarial Network. In Proceedings of the 26th ACM international conference on Multimedia, Seoul, Republic of Korea, 22–26 October 2018; ACM: New York, NY, USA, 2018; pp. 896–904. [CrossRef]

87. Mohammadi, M.; Al-Fuqaha, A.; Oh, J.-S. Path Planning in Support of Smart Mobility Applications Using Generative Adversarial Networks. In Proceedings of the 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Halifax, NS, Canada, 30 July–3 August 2018; IEEE: New York, NY, USA, 2018; pp. 878–885. [CrossRef]

88. Li, D.; Shi, X.; Long, Q.; Liu, S.; Yang, W.; Wang, F.; Wei, Q.; Qiao, F. DXSLAM: A Robust and Efficient Visual SLAM System with Deep Features. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 24 October 2020–24 January 2021; IEEE: New York, NY, USA, 2020; pp. 4958–4965. [CrossRef]

89. Zuo, L.; Zhang, C.H.; Liu, F.L.; Wu, Y.F. Performance Evaluation of Deep Neural Networks in Detecting Loop Closure of Visual SLAM. In Proceedings of the 2019 11th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Hangzhou, China, 24–25 August 2019; IEEE: New York, NY, USA, 2019; pp. 171–175. [CrossRef]

90. Chakravarty, P.; Narayanan, P.; Roussel, T. GEN-SLAM: Generative Modeling for Monocular Simultaneous Localization and Mapping. In Proceedings of the 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; IEEE: New York, NY, USA, 2019; pp. 147–153. [CrossRef]

91. Jatavallabhula, K.M.; Iyer, G.; Paull, L. ∇slam: Dense slam meets automatic differentiation. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May–31 August 2020; pp. 2130–2137.

92. Xu, Z.; Haroutunian, M.; Murphy, A.J.; Neasham, J.; Norman, R. An Integrated Visual Odometry System with Stereo Camera for Unmanned Underwater Vehicles. *IEEE Access* **2022**, *10*, 71329–71343. [CrossRef]

93. Xu, Z.; Haroutunian, M.; Murphy, A.J.; Neasham, J.; Norman, R. An Integrated Visual Odometry System for Underwater Vehicles. *IEEE J. Ocean. Eng.* **2021**, *46*, 848–863. [CrossRef]

94. Khan, A.; Fouda, M.M.; Do, D.T.; Almaleh, A.; Alqahtani, A.M.; Rahman, A.U. Underwater Target Detection Using Deep Learning: Methodologies, Challenges, Applications, and Future Evolution. *IEEE Access* **2024**, *12*, 12618–12635. [CrossRef]

95. Xu, Z.; Haroutunian, M.; Murphy, A.J.; Neasham, J.; Norman, R. A Low-Cost Visual Inertial Odometry System for Underwater Vehicles. In Proceedings of the 2021 4th International Conference on Mechatronics, Robotics and Automation (ICMRA), Zhanjiang, China, 22–24 October 2021; IEEE: New York, NY, USA, 2021; pp. 139–143. [CrossRef]

96. Zheng, J.; Zhao, R.; Yang, G.; Liu, S.; Zhang, Z.; Fu, Y.; Lu, J. An Underwater Image Restoration Deep Learning Network Combining Attention Mechanism and Brightness Adjustment. *J. Mar. Sci. Eng.* **2024**, *12*, 7. [CrossRef]

97. Joshi, R.; Usmani, K.; Krishnan, G.; Blackmon, F.; Javidi, B. Underwater object detection and temporal signal detection in turbid water using 3D-integral imaging and deep learning. *Opt. Express* **2024**, *32*, 1789. [CrossRef]

98. Ashwini, A.; Purushothaman, K.E.; Gnanaprakash, V.; Shahila, D.F.D.; Vaishnavi, T.; Rosi, A. 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, Kollam, India, 10–11 August 2023; Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA, 2023; pp. 1545–1549. [CrossRef]

99. Yu, L.; Yang, E.; Yang, B.; Fei, Z.; Niu, C. A robust learned feature-based visual odometry system for UAV pose estimation in challenging indoor environments. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 5015411. [CrossRef]

100. Fraundorfer, F.; Scaramuzza, D. Visual Odometry: Part II: Matching, Robustness, Optimization, and Applications. *IEEE Robot. Autom. Mag.* **2012**, *19*, 78–90. [CrossRef]

101. Scaramuzza, D.; Fraundorfer, F. Tutorial: Visual odometry. *IEEE Robot. Autom. Mag.* **2011**, *18*, 80–92. [CrossRef]

102. Hoda, M.N. Bharati Vidyapeeth's Institute of Computers Applications and Management Delhi, and Institute of Electrical and Electronics Engineers Delhi Section. In Proceedings of the 17th INDIACom-2023 10th International Conference on Computing for Sustainable Global Development, New Delhi, India, 15–17 March 2023.

103. Weiss, A.; Deane, G.B.; Singer, A.C.; Wornell, G. On data-driven underwater acoustic direct localization: Design considerations of a deep neural network-based solution. *J. Acoust. Soc. Am.* **2023**, *153*, A177. [CrossRef]

104. Rahman, S.; Li, A.Q.; Rekleitis, I. SVIn2: A multi-sensor fusion-based underwater SLAM system. *Int. J. Rob. Res.* **2022**, *41*, 1022–1042. [CrossRef]

105. Saha, S.S.; Sandha, S.S.; Garcia, L.A.; Srivastava, M. Tinyodom: Hardware-aware efficient neural inertial navigation. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2022**, *6*, 1–32. [CrossRef]

106. Liu, Z.; Malis, E.; Martinet, P. A New Dense Hybrid Stereo Visual Odometry Approach. In Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Kyoto, Japan, 23–27 October 2022; IEEE: New York, NY, USA, 2022; pp. 6998–7003. [CrossRef]

107. Liu, Y.; Sun, Y.; Li, B.; Wang, X.; Yang, L. Experimental Analysis of Deep-Sea AUV Based on Multi-Sensor Integrated Navigation and Positioning. *Remote Sens.* **2024**, *16*, 199. [CrossRef]

108. Chen, C.; Lu, X.; Markham, A.; Trigoni, N. Ionet: Learning to cure the curse of drift in inertial odometry. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32. [CrossRef]

109. Bajwa, A.; Cossette, C.C.; Shalaby, M.A.; Forbes, J.R. DIVE: Deep Inertial-Only Velocity Aided Estimation for Quadrotors. *IEEE Robot. Autom. Lett.* **2024**, *9*, 3728–3734. [CrossRef]

110. Abdelaziz, N.; El-Rabbany, A. Deep Learning-Aided Inertial/Visual/LiDAR Integration for GNSS-Challenging Environments. *Sensors* **2023**, *23*, 6019. [CrossRef]

111. Kabanov, A.; Kramar, V.; Lipko, I.; Dementiev, K. Cooperative Control of Underwater Vehicle–Manipulator Systems Based on the SDC Method. *Sensors* **2022**, *22*, 5038. [CrossRef]

112. Wang, S.; Clark, R.; Wen, H.; Trigoni, N. DeepVO: Towards end-to-end visual odometry with deep Recurrent Convolutional Neural Networks. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017; IEEE: New York, NY, USA, 2017; pp. 2043–2050. [CrossRef]

113. Kendall, A.; Grimes, M.; Cipolla, R. PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 11–18 December 2015; IEEE: New York, NY, USA, 2015; pp. 2938–2946. [CrossRef]

114. Detone, D.; Malisiewicz, T.; Rabinovich, A. SuperPoint: Self-supervised interest point detection and description. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; IEEE Computer Society: Washington, DC, USA, 2018; pp. 337–349. [CrossRef]

115. Song, S.; Yu, F.; Jiang, X.; Zhu, J.; Cheng, W.; Fang, X. Loop closure detection of visual SLAM based on variational autoencoder. *Front. Neurorobot* **2023**, *17*, 1301785. [CrossRef]

116. Bescos, B.; Fácil, J.M.; Civera, J.; Neira, J. DynaSLAM: Tracking, Mapping and Inpainting in Dynamic Scenes. *IEEE Robot. Autom. Lett.* **2018**, *3*, 4076–4083. [CrossRef]

117. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How Transferable Are Features in Deep Neural Networks? *arXiv* **2014**, arXiv:1411.1792.

118. Ohri, K.; Kumar, M. Review on self-supervised image recognition using deep neural networks. *Knowl.-Based Syst.* **2021**, *224*, 107090. [CrossRef]

119. Memon, A.R.; Wang, H.; Hussain, A. Loop closure detection using supervised and unsupervised deep neural networks for monocular SLAM systems. *Robot. Auton. Syst.* **2020**, *126*, 103470. [CrossRef]

120. Pedraza, B.; Dera, D. Robust Active Simultaneous Localization and Mapping Based on Bayesian Actor-Critic Reinforcement Learning. In Proceedings of the 2023 IEEE Conference on Artificial Intelligence, CAI 2023, Santa Clara, CA, USA, 5–6 June 2023; Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA, 2023; pp. 63–66. [CrossRef]

121. Lillicrap, T.P.; Hunt, J.J.; Pritzel, A.; Heess, N.; Erez, T.; Tassa, Y.; Silver, D.; Wierstra, D. Continuous Control with Deep Reinforcement Learning. *arXiv* **2015**, arXiv:1509.02971.

122. Zhao, Z.; Alzubaidi, L.; Zhang, J.; Duan, Y.; Gu, Y. A comparison review of transfer learning and self-supervised learning: Definitions, applications, advantages and limitations. *Expert Syst. Appl.* **2024**, *242*, 122807. [CrossRef]

123. Campos, C.; Elvira, R.; Rodriguez, J.J.G.; Montiel, J.M.M.; Tardos, J.D. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multimap SLAM. *IEEE Trans. Robot.* **2021**, *37*, 1874–1890. [CrossRef]

124. Drupt, J.; Dune, C.; Comport, A.I.; Hugel, V. Qualitative evaluation of state-of-the-art DSO and ORB-SLAM-based monocular visual SLAM algorithms for underwater applications. In Proceedings of the OCEANS 2023—Limerick, Limerick, Ireland, 5–8 June 2023; IEEE: New York, NY, USA, 2023; pp. 1–7. [CrossRef]

125. Li, M.; Yang, K.; Qin, J.; Zhong, J.; Jiang, Z.; Su, Q. Comparative study on real-time pose estimation of vision-based unmanned underwater vehicles. *Cobot* **2023**, *2*, 2. [CrossRef]

126. Mur-Artal, R.; Tardos, J.D. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. *IEEE Trans. Robot.* **2017**, *33*, 1255–1262. [CrossRef]

127. Zhang, Y.; Zhou, L.; Li, H.; Zhu, J.; Du, W. Marine Application Evaluation of Monocular SLAM for Underwater Robots. *Sensors* **2022**, *22*, 4657. [CrossRef]

128. Vial, P.; Palomeras, N.; Solà, J.; Carreras, M. Underwater Pose SLAM using GMM scan matching for a mechanical profiling sonar. *J. Field Robot.* **2024**, *41*, 511–538. [CrossRef]

129. Eustice, R.; Pizarro, O.; Singh, H.; Howland, J. UWIE underwater image toolbox for optical image processing and mosaicking in MATLAB. In Proceedings of the 2002 Interntional Symposium on Underwater Technology, Tokyo, Japan, 16–19 April 2002; pp. 141–145. [CrossRef]

130. Grisetti, G.; Stachniss, C.; Burgard, W. Improved techniques for grid mapping with Rao-Blackwellized particle filters. *IEEE Trans. Robot.* **2007**, *23*, 34–46. [CrossRef]

131. Gao, X.; Wang, R.; Demmel, N.; Cremers, D. LDSO: Direct Sparse Odometry with Loop Closure. In Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018. [CrossRef]

132. Zhu, Z.; Peng, S.; Larsson, V.; Xu, W.; Bao, H.; Cui, Z.; Oswald, M.R.; Pollefeys, M. NICE-SLAM: Neural Implicit Scalable Encoding for SLAM. Available online: https://github.com/cvg/nice-slam (accessed on 6 October 2024).

133. Deng, X.; Liu, T.; He, S.; Xiao, X.; Li, P.; Gu, Y. An underwater image enhancement model for domain adaptation. *Front. Mar. Sci.* **2023**, *10*, 1138013. [CrossRef]

134. Wang, Y.; Ma, X.; Wang, J.; Hou, S.; Dai, J.; Gu, D.; Wang, H. Robust AUV Visual Loop-Closure Detection Based on Variational Autoencoder Network. *IEEE Trans. Ind. Inform.* **2022**, *18*, 8829–8838. [CrossRef]

135. Jin, A.; Zeng, X. A Novel Deep Learning Method for Underwater Target Recognition Based on Res-Dense Convolutional Neural Network with Attention Mechanism. *J. Mar. Sci. Eng.* **2023**, *11*, 69. [CrossRef]

136. Valdenegro-Toro, M. Improving Sonar Image Patch Matching via Deep Learning. *arXiv* **2017**, arXiv:1709.02150.

137. Qin, J.; Li, M.; Li, D.; Zhong, J.; Yang, K. A Survey on Visual Navigation and Positioning for Autonomous UUVs. *Remote Sens.* **2022**, *14*, 3794. [CrossRef]

138. Favorskaya, M.N. Deep Learning for Visual SLAM: The State-of-the-Art and Future Trends. *Electronics* **2023**, *12*, 2006. [CrossRef]

139. Chen, C.; Wang, B.; Lu, C.X.; Trigoni, N.; Markham, A. Deep Learning for Visual Localization and Mapping: A Survey. *arXiv* **2023**, arXiv:2308.14039. [CrossRef]

140. Wright, R.G. Intelligent autonomous ship navigation using multisensor modalities. *TransNav* **2019**, *13*, 503–510. [CrossRef]

141. Li, J.; Xiang, X.; Yang, S. Robust adaptive neural network control for dynamic positioning of marine vessels with prescribed performance under model uncertainties and input saturation. *Neurocomputing* **2022**, *484*, 1–12. [CrossRef]

142. Iman, M.; Rasheed, K.; Arabnia, H.R. A Review of Deep Transfer Learning and Recent Advancements. *Technologies* **2023**, *11*, 40. [CrossRef]

143. Sarlin, P.-E.; DeTone, D.; Malisiewicz, T.; Rabinovich, A. SuperGlue: Learning Feature Matchin g with Graph Neural Networks. *arXiv* **2019**, arXiv:1911.11763.

144. Lyu, Y.; Huang, X.; Zhang, Z. Learning to Segment 3D Point Clouds in 2D Image Space. Available online: https://github.com/Zhang-VISLab (accessed on 6 October 2024).

145. Chaudhuri, R.; Deb, S.; Das, H. Noble Approach on Sensor Fused Bio Intelligent Path Optimisation and Single Stage Obstacle Recognition in Customized Mobile Agent. *Procedia Comput. Sci.* **2022**, *218*, 778–787. [CrossRef]

146. Zhou, B.; Lapedriza, A.; Khosla, A.; Oliva, A.; Torralba, A. Places: A 10 Million Image Database for Scene Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 1452–1464. [CrossRef]

147. Kim, A.; Eustice, R.M. Real-time visual SLAM for autonomous underwater hull inspection using visual saliency. *IEEE Trans. Robot.* **2013**, *29*, 719–733. [CrossRef]

148. Chen, L.; Liu, Y.; Dong, P.; Liang, J.; Wang, A. An Intelligent Navigation Control Approach for Autonomous Unmanned Vehicles via Deep Learning-Enhanced Visual SLAM Framework. *IEEE Access* **2023**, *11*, 119067–119077. [CrossRef]

149. Lin, Y.H.; Wu, T.L.; Yu, C.M.; Wu, I.C. Development of an intelligent underwater recognition system based on the deep reinforcement learning algorithm in an autonomous underwater vehicle. *Measurement* **2023**, *214*, 112844. [CrossRef]

150. Burguera, A.; Bonin-Font, F. An Unsupervised Neural Network for Loop Detection in Underwater Visual SLAM. *J. Intell. Robot. Syst. Theory Appl.* **2020**, *100*, 1157–1177. [CrossRef]

151. Zalewski, J.; Hożyń, S. Computer Vision-Based Position Estimation for an Autonomous Underwater Vehicle. *Remote Sens.* **2024**, *16*, 741. [CrossRef]

152. Wang, C.; Cheng, C.; Yang, D.; Pan, G.; Zhang, F. Underwater AUV Navigation Dataset in Natural Scenarios. *Electronics* **2023**, *12*, 3788. [CrossRef]

153. Yang, Y.; Xiao, Y.; Li, T. A Survey of Autonomous Underwater Vehicle Formation: Performance, Formation Control, and Communication Capability. *IEEE Commun. Surv. Tutorials* **2021**, *23*, 815–841. [CrossRef]

154. Shaukat, N.; Ali, A.; Iqbal, M.J.; Moinuddin, M.; Otero, P. Multi-Sensor Fusion for Underwater Vehicle Localization by Augmentation of RBF Neural Network and Error-State Kalman Filter. *Sensors* **2021**, *21*, 1149. [CrossRef]

155. Wang, D.; Wang, B.; Huang, H.; Zhang, H. A Multisensor Fusion Method Based on Strict Velocity for Underwater Navigation System. *IEEE Sens. J.* **2023**, *23*, 18587–18598. [CrossRef]