



OPEN ACCESS

EDITED BY

Jay S. Pearlman,
Institute of Electrical and Electronics
Engineers, France

REVIEWED BY

Néstor E. Ardila,
Rosario University, Colombia
Hao Wang,
China University of Petroleum, China

*CORRESPONDENCE

Pengfei Shi
✉ shipf@hhu.edu.cn

RECEIVED 09 February 2025

ACCEPTED 21 April 2025

PUBLISHED 16 May 2025

CITATION

Fan X, Wang X, Ni H, Xin Y and Shi P (2025)
Water-Adapted 3D Gaussian Splatting for
precise underwater scene reconstruction.
Front. Mar. Sci. 12:1573612.
doi: 10.3389/fmars.2025.1573612

COPYRIGHT

© 2025 Fan, Wang, Ni, Xin and Shi. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Water-Adapted 3D Gaussian Splatting for precise underwater scene reconstruction

Xinnan Fan¹, Xiaotian Wang², Haonan Ni¹, Yuanxue Xin¹
and Pengfei Shi^{3*}

¹The College of Information Science and Engineering, Hohai University, Changzhou, China, ²The College of Computer Science and Software Engineering, Hohai University, Nanjing, China, ³The College of Artificial Intelligence and Automation, Hohai University, Changzhou, China

Underwater 3D reconstruction is essential for marine surveying, ecological protection, and underwater engineering. Traditional methods, designed for air environments, fail to account for underwater optical properties, leading to poor detail retention, color reproduction, and visual consistency. In recent years, 3D Gaussian Splatting (3DGS) has emerged as an efficient alternative, offering improvements in both speed and quality. However, existing 3DGS methods struggle to adaptively adjust point distribution based on scene complexity, often resulting in inadequate detail reconstruction in complex areas and inefficient resource usage in simpler ones. Additionally, depth variations in underwater scenes affect image clarity, and current methods lack adaptive depth-based rendering, leading to inconsistent clarity between near and distant objects. Existing loss functions, primarily designed for air environments, fail to address underwater challenges such as color distortion and structural differences. To address these challenges, we propose an improved underwater 3D Gaussian Splatting method combining complexity-adaptive point distribution, depth-adaptive multi-scale radius rendering, and a tailored loss function for underwater environments. Our method enhances reconstruction accuracy and visual consistency. Experimental results on static and dynamic underwater datasets show significant improvements in detail retention, rendering accuracy, and stability compared to traditional methods, making it suitable for practical underwater 3D reconstruction applications.

KEYWORDS

underwater 3D reconstruction, 3D Gaussian Splatting, water-adapted rendering, dynamic underwater scenes, ocean observation

1 Introduction

In recent years, underwater scene analysis has attracted increasing attention. A variety of methods have been proposed to address the challenges of underwater image processing, such as underwater image captioning (Li et al., 2025) and underwater image enhancement empowered by large foundation models (Wang et al., 2025). However, with the

development of the field, underwater 3D reconstruction has become a crucial research direction due to its ability to provide comprehensive spatial information for underwater scene understanding. High-precision 3D reconstruction technology can support tasks such as seabed topography mapping, marine habitat assessment, and underwater facility inspection, providing accurate spatial data for marine scientific research and engineering (Hu et al., 2023; Cutolo et al., 2024). In underwater environments, realtime 3D reconstruction technology can also assist operators in better identifying and locating target areas during remote monitoring or robotic operations, significantly enhancing operational efficiency.

However, the unique optical characteristics of underwater environments present significant challenges to imaging and 3D reconstruction. The propagation of light in water is strongly influenced by absorption and scattering, leading to image blurring and severe color distortion. Light of different wavelengths attenuates at varying rates, with red and yellow wavelengths disappearing rapidly in shallow water (Lin et al., 2024a), significantly complicating color restoration. As water depth increases, the light attenuation effect becomes more pronounced, further reducing image contrast and clarity, affecting object visibility and detail retention (Jia et al., 2024). In addition, dynamic objects introduce further challenges to 3D reconstruction. Their movement, shape changes, and lighting fluctuations cause variations in the image's geometric structure and lighting, which hinder object recognition and detail preservation (Hua et al., 2023; Rout et al., 2024). Especially around fast-moving or changing objects, traditional underwater reconstruction methods struggle to accurately extract and match the features of these objects, leading to blurred, distorted, or unrealistic reconstruction results. Therefore, effectively handling dynamic objects under the influence of complex optical effects, particularly in preserving their details, is a significant challenge in underwater 3D reconstruction.

In recent years, Neural Radiance Fields (NeRF) methods have made significant progress in 3D reconstruction and novel view synthesis (Mildenhall et al., 2021). NeRF is a deep learning-based method that generates high-quality images from different viewpoints by modeling the scene's density and color distribution. Using a multi-layer perceptron (MLP), NeRF encodes the viewpoint, position, and lighting information of each point in the scene to synthesize realistic images. However, its high computational complexity and training costs limit its use in real-time applications, especially in resource-constrained environments, where its efficiency and flexibility are significant challenges.

To address these inefficiencies, the 3D Gaussian Splatting method was proposed as a more efficient alternative (Kerbl et al., 2023). By representing the scene as a series of Gaussian distribution points, 3DGS eliminates the need for NeRF's implicit representation, resulting in faster rendering and greater editing flexibility. However, current 3DGS methods are less effective in underwater environments. The optical properties of water cause significant imaging differences between near and distant objects, and 3DGS lacks depth-adaptive rendering adjustments, leading to poor visual consistency in underwater scenes with large depth variations. Moreover, these methods, typically based on air

environment imaging assumptions, fail to address color shifts and detail loss in underwater scenes, making it difficult to restore realistic effects in underwater 3D reconstructions. These limitations hinder the effectiveness of 3DGS for high-precision underwater applications, particularly in marine engineering.

To address these issues, this paper proposes an improved underwater 3D Gaussian Splatting method, as shown in Figure 1, which effectively enhances detail retention and visual consistency of underwater scenes by combining complexity-adaptive distribution, depth-adaptive multi-scale rendering, and a loss function specifically designed for underwater environments. Specifically, the main contributions of this paper include:

- **Proposing a complexity-based Gaussian point distribution optimization method:** This method utilizes gradient complexity information from images to adaptively adjust the density distribution of Gaussian points, ensuring sufficient reconstruction in detail rich areas while reducing computational power consumption in simpler regions.
- **Constructing a depth-adaptive multi-scale radius rendering strategy:** This strategy adaptively adjusts the Gaussian point radius factor based on the scene pixel depth, achieving high-resolution rendering for nearby objects and blurring for distant water, thus enhancing the visual consistency of the rendering results.
- **Designing a loss function optimized for underwater reconstruction:** This function combines regularized L1 loss, regularized SSIM loss, and LPIPS perceptual loss, which significantly improves color restoration and structural detail performance in underwater images.

We evaluated the proposed method on the static underwater dataset SeaThru-NeRF (Levy et al., 2023) and the dynamic underwater dataset In-the-wild (Tang et al., 2024). Experimental results demonstrate that, compared to existing NeRF and 3DGS-based methods, the proposed approach offers significant advantages in detail retention, rendering efficiency, and visual consistency, highlighting its effectiveness and superiority in practical underwater 3D reconstruction applications.

2 Related work

2.1 Challenges in underwater imaging and 3D reconstruction

The unique optical characteristics of underwater environments, including light absorption, scattering, and wavelength-dependent attenuation, significantly affect imaging quality, presenting a severe challenge to the accuracy and effectiveness of underwater 3D reconstruction (Li et al., 2024a). Additionally, the heterogeneity of the underwater environment causes significant variations in the complexity of imaging regions, making it difficult for traditional 3D reconstruction methods to balance detail preservation and resource utilization efficiency.

Traditional underwater imaging and 3D reconstruction methods typically rely on physical models, using color correction and depth-dependent light compensation to improve image quality. For example, some studies estimate light attenuation coefficients based on underwater optical models to correct the color of underwater images (Lin et al., 2024b).

However, traditional physics-based methods often require accurate optical parameters for the water body, making it difficult to adapt to dynamic and complex underwater environments. Furthermore, these methods lack applicability in diverse underwater scenes, particularly in environments with continuously changing lighting conditions and water quality, making it challenging to maintain stable imaging quality (Zhang et al., 2024a).

2.2 NeRF-based underwater 3D reconstruction methods

NeRF uses a multi-layer perceptron (MLP) to model the relationship between viewpoint, position, and color density, achieving high-quality view synthesis (Zhang et al., 2024b). However, the base model of NeRF was primarily designed for air environments and does not account for optical characteristics in underwater environments, such as light attenuation and scattering effects, leading to limitations when directly applied to underwater scenes.

To address this issue, Deborah Levy et al. proposed SeaThru-NeRF, which combines underwater optical models with NeRF to compensate for underwater light attenuation characteristics, enabling NeRF to better adapt to color restoration and detail retention in underwater images (Levy et al., 2023). SeaThru-NeRF estimates light attenuation using physical models, improving the blue-green shift in underwater images and making significant progress in color correction and detail restoration.

However, the performance of SeaThru-NeRF is still limited by depth information and optical models, and its effectiveness remains unstable in complex underwater environments. To further improve the applicability of NeRF in underwater environments, WaterNeRF (Sethuraman et al., 2023) combines physical models and data-driven methods, dynamically modeling light attenuation and scattering effects, achieving more stable color restoration and scene detail representation. WaterNeRF not only provides higher visual consistency in underwater scenes with large depth variations but also further optimizes scene clarity through network training.

2.3 3D Gaussian Splatting and its progress in underwater applications

NeRF-based underwater 3D reconstruction methods have made significant progress in detail retention and visual quality. However, their high computational complexity and sensitivity to data limit their efficiency and applicability in practical scenarios. Deep learning-based underwater NeRF methods typically require large

amounts of labeled data, with both training and rendering being time-consuming, creating bottlenecks in real-time underwater monitoring and surveying applications (Zhou et al., 2025). To improve the computational efficiency of 3D reconstruction, the 3DGS method has emerged in recent years. 3DGS explicitly represents the scene as a point cloud of Gaussian distributions, enabling efficient rendering and detail representation. Compared to NeRF's implicit representation, this method significantly reduces computational costs and is suitable for real-time applications (Luo et al., 2024). However, traditional 3DGS methods assume an air environment and lack adaptability to the unique optical characteristics of underwater environments, leading to issues such as color shift and detail blur when applied directly to underwater scenes.

To improve the applicability of 3DGS in underwater environments, Li et al. proposed WaterSplatting (Li et al., 2024b), which integrates 3DGS with volumetric rendering to model geometric structures efficiently and render high quality water media. Experiments show that WaterSplatting improves the detail representation of underwater scenes to some extent, producing more realistic reconstruction results. Additionally, UW-GS (Wang et al., 2024) further enhances the robustness of underwater reconstruction by introducing a dynamic object handling mechanism. The UW-GS method can identify and filter out dynamic objects (e.g., fish), reducing interference from moving objects and improving reconstruction stability in complex environments.

Although 3DGS-based underwater methods have made progress in color restoration, dynamic processing, and detail retention, they still have limitations in adapting to complex underwater environments. Therefore, enhancing the environmental adaptability and rendering accuracy of 3DGS methods has become a key research direction. To address this, this paper proposes an improved underwater 3D Gaussian Splatting method, which enhances detail retention and visual consistency by introducing complexity-adaptive distribution, depth-adaptive multi-scale rendering, and a loss function specifically designed for underwater environments. This method not only retains the high computational efficiency of 3DGS but also optimizes its adaptability in underwater environments, providing a more efficient solution for real-time underwater 3D reconstruction in marine engineering.

3 Materials and methods

This paper proposes an improved underwater 3DGS method, aimed at enhancing the accuracy and resource utilization efficiency of underwater 3D reconstruction. To address the complex optical characteristics of underwater scenes and the varying detail demands across different regions, the proposed method combines complexity-adaptive Gaussian point distribution, depth-based multi-scale Gaussian point radius rendering, and a loss function specifically designed for underwater environments, achieving higher quality reconstruction results. The overall process is illustrated in the Figure 2.

3.1 Complexity-based Gaussian point distribution optimization

To improve detail retention and optimize resource allocation in underwater 3D reconstruction, this paper proposes a complexity-based Gaussian point distribution optimization method. This method dynamically adjusts the distribution density of Gaussian points in different regions by referencing the complexity map of the image, allowing the points to adaptively concentrate in areas with higher detail demands. The specific implementation process is as follows:

Complexity Calculation: The complexity map $C(x, y)$ is generated based on the gradient information of the reference image, where x and y represent pixel positions in the image. Specifically, the Sobel operator is used to calculate the gradient magnitude for each pixel in the image, which measures the complexity of local regions.

The gradient magnitude $G(x, y)$ is defined as Equation 1:

$$G(x, y) = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \quad (1)$$

where $I(x, y)$ represents the pixel intensity of the image, and $\frac{\partial I}{\partial x}$ and $\frac{\partial I}{\partial y}$ are the gradients in the horizontal and vertical directions, respectively. The complexity map $C(x, y)$ is the normalized form of the gradient magnitude $G(x, y)$, used to represent the complexity of different regions in the scene.

Complexity Region Division: High-complexity regions typically include detail-rich parts, such as close-up reef structures or swimming fish, which require complex rendering. Low-complexity regions often correspond to flat or less textured areas, such as large expanses of water. After generating the complexity map, a complexity threshold is set to divide the image into high

complexity and low-complexity regions, effectively distinguishing between regions with high and low detail demands.

Adaptive Gaussian Point Distribution: After dividing the regions, Gaussian point density parameters D_{high} and D_{low} are defined. The Gaussian point positions are adaptively adjusted based on the region's complexity. The final set of Gaussian point positions is as Equation 2:

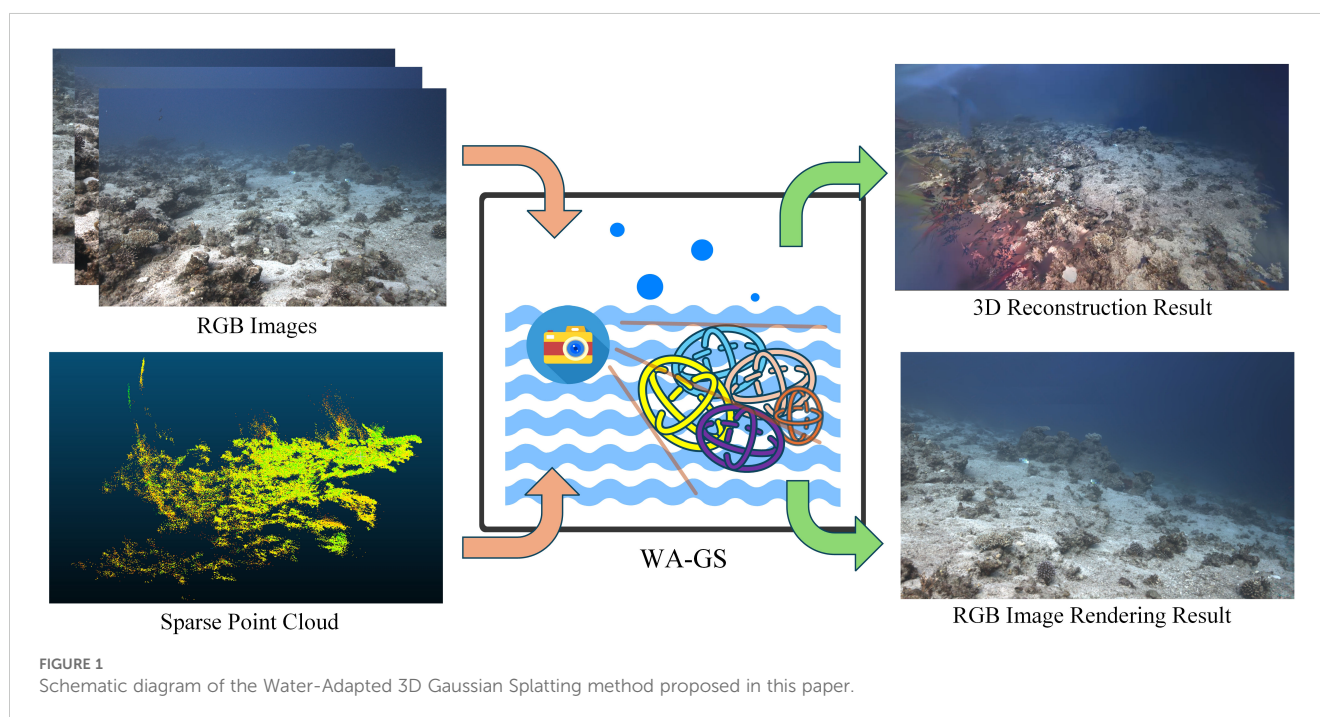
$$P = \{p_i | p_i \in \text{High_complexity_region}, i = 1, \dots, D_{high}\} \cup \{p_j | p_j \in \text{Low_complexity_region}, j = 1, \dots, D_{low}\} \quad (2)$$

Through this complexity-based adaptive Gaussian point distribution optimization strategy, this paper allocates more Gaussian points in detail-rich areas to improve reconstruction accuracy, while reducing the distribution of Gaussian points in simpler regions, thereby saving computational resources and optimizing the overall effect of underwater 3D reconstruction.

3.2 Depth-adaptive multi-scale radius rendering strategy

In the 3D Gaussian Splatting rasterization process, the scale parameter Radii is used to control the spread of each 2D Gaussian distribution, thus affecting the blurriness and coverage scale of the Gaussian points on the image plane. During rendering, larger scale values make the coverage area of the Gaussian points wider, reducing computational load and weakening boundary details, while smaller scale values limit the spread of Gaussian points, enhancing detail preservation.

Furthermore, the scale parameter plays a key role in the fusion effect between different depth layers. In the 2D projection of multi-layer depth information, larger scale values help increase the



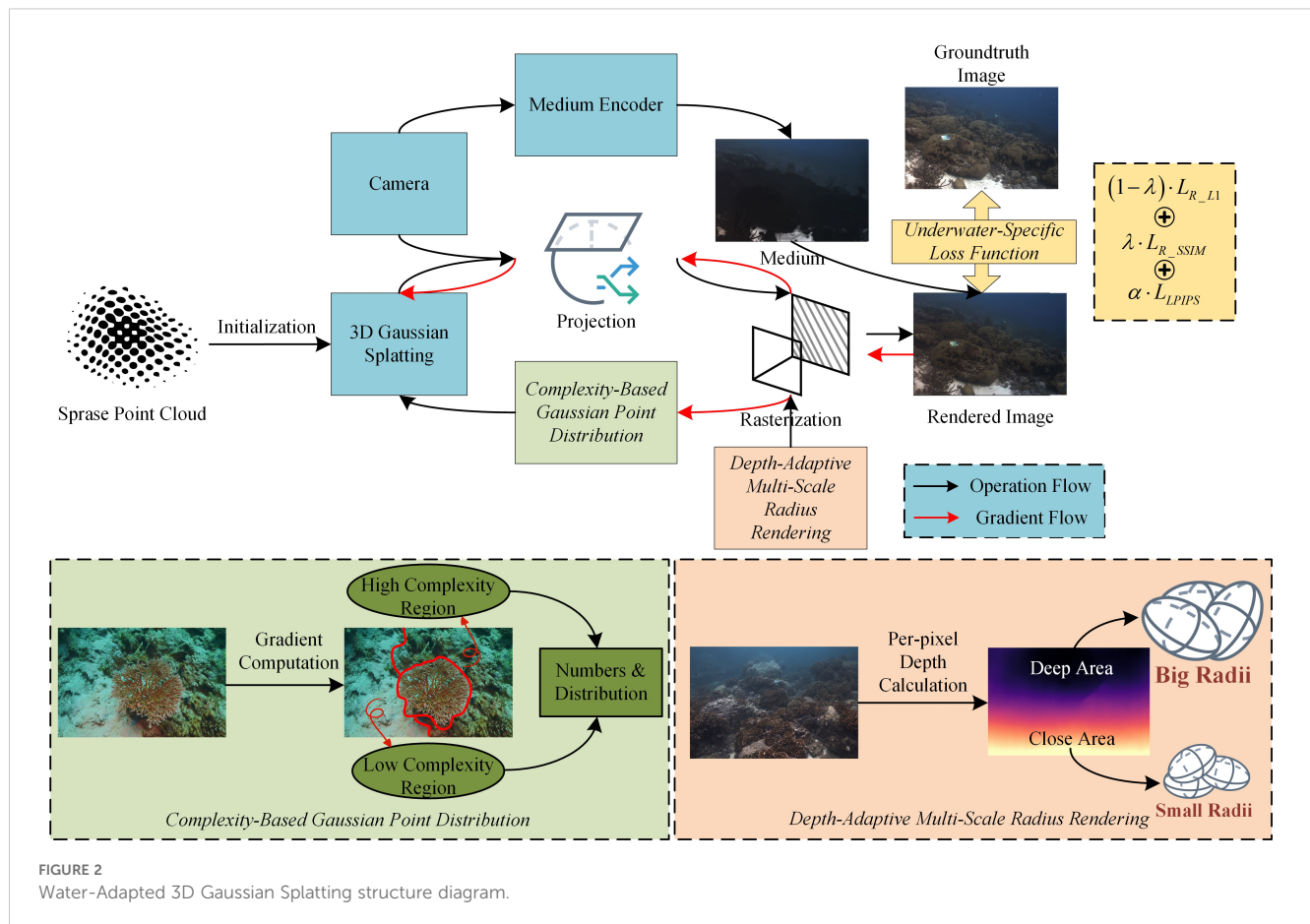


FIGURE 2
Water-Adapted 3D Gaussian Splatting structure diagram.

smoothness of inter-layer transitions, while smaller scale values preserve sharper boundary delineations. Therefore, we design an adaptive scale adjustment strategy based on scene depth, allowing closer objects to have higher clarity, while distant water exhibits a more natural blur, thus achieving visual consistency and realism in the rendered scene. To implement depth-adaptive adjustment, this paper calculates the corresponding scale factor based on the depth information of each pixel, which is used to dynamically control the spread of Gaussian points. Specifically, the formula for calculating the scale factor we designed is as Equation 3:

$$S(x, y) = 0.5 + \frac{D(x, y)}{D_{\max}} \quad (3)$$

Where: $S(x, y)$ is the scale factor at position (x, y) , controlling the spread of Gaussian points at that location. $D(x, y)$ is the depth value at position (x, y) . $D_{\max}$ is the maximum depth value in the scene, used for normalizing the depth range. According to this formula, as the depth increases, the scale factor is bounded within $(0.5, 1.5)$. In this way, the scale factor for closer objects is smaller, leading to higher resolution and clarity. For distant objects, the scale factor is larger, which, for large water areas, results in an imaging effect that aligns with natural visual depth perception without significantly increasing computational cost.

Through this strategy, we dynamically adjust the spread of Gaussian points based on depth information, achieving visual

consistency in underwater scenes with significant depth variations, thereby enhancing the realism and detail representation of the reconstruction.

3.3 Underwater loss function design

To enhance color restoration, structural consistency, and detail representation in underwater 3D reconstruction, this paper designs a loss function specifically for underwater environments, combining regularized L1 loss, regularized SSIM loss, and LPIPS perceptual loss. This combined loss function helps suppress overfitting to specific pixels or structures while capturing subtle perceptual differences and structural features in underwater images, leading to higher quality reconstruction results.

Regularized L1 loss: The regularized L1 loss calculates the pixel-wise error between the predicted and ground truth images and normalizes it in high-brightness areas to suppress overfitting in specific pixel regions, enhancing overall color and brightness consistency. Its formula is shown as Equation 4:

$$L_{R_L1} = \frac{1}{N} \sum_{i=1}^N \left| \frac{I_{gt}(i) - I_{pred}(i)}{I_{gt}(i) + \epsilon} \right| \quad (4)$$

Here, I_{gt} and I_{pred} represent the pixel values of the ground truth image and the predicted image, respectively. N is the total number

of pixels, and ε is a small constant used to avoid division by zero errors.

Regularized SSIM Loss: The regularized SSIM loss is used to assess the local structural similarity of images and controls the structural consistency of different pixel regions by normalization. It reduces the influence of local high-contrast areas and enhances the structural fidelity of the image. It is defined as Equation 5:

$$L_{R_SSIM} = 1 - \text{SSIM}\left(\frac{I_{gt}}{I_{pred} + \varepsilon}, \frac{I_{pred}}{I_{pred} + \varepsilon}\right) \quad (5)$$

LPIPS Perceptual Loss: To further enhance the perceptual realism of the image, this paper introduces the Learned Perceptual Image Patch Similarity (LPIPS) perceptual loss. The LPIPS perceptual loss relies on a pre-trained deep feature network, which calculates the high-level feature differences between the predicted and ground truth images to capture subtle structural and detail differences perceived by the human eye. It is defined as Equation 6:

$$L_{LPIPS} = \frac{1}{N} \sum_{i=1}^N \|\phi(I_{gt})(i) - \phi(I_{pred})(i)\|_2 \quad (6)$$

Where ϕ represents the feature maps of the deep feature network.

The combined loss function in this paper is a weighted combination of the three aforementioned losses, balancing each loss's contribution to the total loss, ensuring color, structure, and detail representation during the reconstruction process. The combined loss function is defined as Equation 7:

$$L = (1 - \lambda) \cdot L_{R_L1} + \lambda \cdot L_{R_SSIM} + \alpha \cdot L_{LPIPS} \quad (7)$$

Where the parameters λ and α control the weights of the individual losses, ensuring the balance of color, structure, and detail representation during the reconstruction process.

4 Results

4.1 Experimental settings

Our code is based on WaterSplatting (Li et al., 2024b), which is built upon the reconstructed version of 3DGS provided by NeRFStudio. The experiments were run on a workstation equipped with an NVIDIA RTX 4090 GPU, an Intel Core i5-13490K processor, and 64GB of RAM. During the actual training, the GPU memory usage was approximately 10GB. In the loss function parameter settings, we set $\lambda = 0.2$ and $\alpha = 0.05$, and used the pre-trained AlexNet network as the feature extractor for LPIPS perceptual loss. Regarding the training parameters, the total number of iterations was set to 15,000. The first 500 iterations were used as a warm-up phase, primarily optimizing the model using regularized L1 loss and regularized SSIM loss, adjusting the initialized point cloud, and pre-training the MLP.

4.2 Datasets and baseline methods

We conducted experimental evaluations on two datasets: SeaThru-NeRF (Levy et al., 2023) and In-the-Wild (Tang et al., 2024). The SeaThruNeRF dataset includes three different marine scenes, covering various water quality conditions and imaging environments, ensuring the richness and diversity of the data, which effectively validates the adaptability of our method in different underwater environments. The In-the-Wild dataset focuses on underwater scene reconstruction, encompassing complex light attenuation effects, unstable scattering phenomena, and the motion of underwater dynamic organisms. It aims to address the unique challenges of underwater environments and provides valuable test data for high-fidelity underwater scene reconstruction.

In the comparative experiments, we selected several baseline methods, including SeaThruNeRF (Levy et al., 2023), ZipNeRF (Barron et al., 2023), DynamicNeRF (Gao et al., 2021), MIP-360 (Barron et al., 2022), Instant-NGP (Müller et al., 2022), 3DGS (Luo et al., 2024), Tang et al (Tang et al., 2024), SeaSplat (Yang et al., 2024), and WaterSplatting (Li et al., 2024b). Due to the innovative nature of our research direction, some of the baseline methods come from preprint papers, which are part of the cutting-edge work currently under development in the field. Through these baseline methods, we are able to comprehensively evaluate the performance and advantages of the WA-GS method in underwater scene reconstruction.

4.3 Experimental results and analysis

The Table 1 shows the results of experiments on the SeaThru-NeRF dataset. We compared the performance of the WA-GS method with several baseline methods across multiple sequences. PSNR is a traditional metric for measuring image quality, with higher values indicating better reconstruction quality. WA-GS shows excellent PSNR values of 30.43 on IUI3-RedSea and 30.07 on Panama. SSIM measures the structural similarity of images, with higher values indicating stronger structural consistency and better detail retention. WAGS achieves very high SSIM values across all datasets, especially 0.938 on Panama and 0.891 on IUI3-RedSea, demonstrating excellent structural consistency and detail retention. LPIPS measures perceptual differences between images, with lower values indicating that the image is closer to the real image in terms of visual effect. Since we incorporated LPIPS loss into our loss function, our method achieves the best LPIPS results across all sequences. In the Panama scene, our LPIPS value (0.084) is much lower than those of 3D GS (0.152) and SeaSplat (0.154), proving that its underwater perception loss function can better suppress vortex artifacts. Figure 3 shows a comparison of rendering effects for three typical underwater scenes: Curaçao, IUI3 RedSea, and J. Gradens RedSea. Figure 4 compares the output effects of different rendering modes in the Panama scene.

TABLE 1 Experimental results of SeaThru-NeRF dataset.

Dataset	Curaçao			IUI3-RedSea			J.Gradens-RedSea			Panama			Time
Methods\ Metrics	PSNR↑	SSIM↑	LPIPS↓	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	
SeaThru-NeRF	30.48	0.873	0.210	25.91	0.785	0.304	21.84	0.767	0.249	27.85	0.834	0.224	10h
ZipNeRF	19.96	0.442	0.421	16.94	0.474	0.412	19.022	0.349	0.483	19.01	0.349	0.482	6h
3D Gaussian	28.31	0.873	0.221	22.98	0.843	0.246	21.49	0.854	0.216	29.20	0.893	0.152	18min
MIP-360	28.23	0.683	0.571	19.55	0.510	0.520	19.62	0.624	0.492	18.32	0.556	0.600	7h
Instant-NGP	27.66	0.684	0.606	20.85	0.519	0.623	23.19	0.726	0.459	21.85	0.604	0.595	5min
Tang et al.	30.03	0.828	0.238	22.70	0.624	0.348	25.81	0.853	0.183	23.75	0.687	0.263	45min
SeaSplat	30.30	0.900	0.194	26.67	0.872	0.208	22.70	0.873	0.179	28.76	0.902	0.154	1h25min
WA-GS(ours)	28.29	0.900	0.158	30.43	0.891	0.186	23.17	0.864	0.153	30.07	0.938	0.084	20min

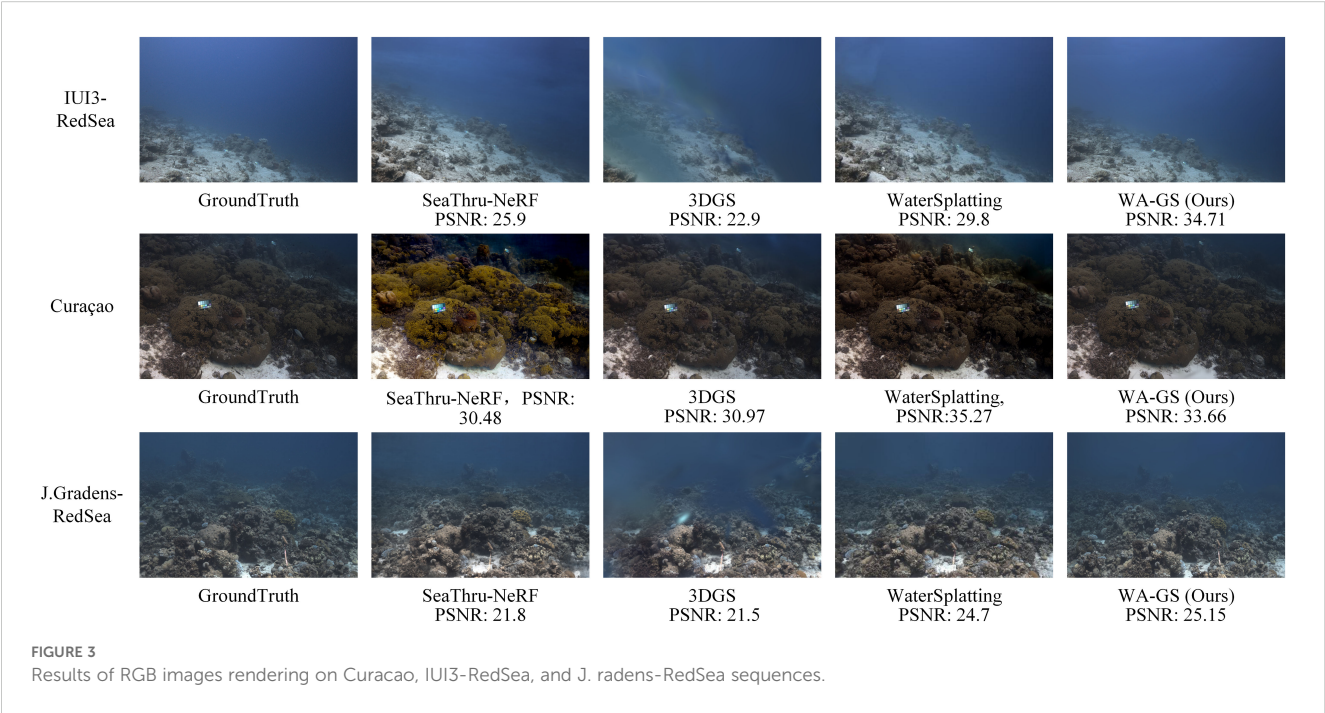
The bold values in the table represent the optimal values.
The meaning of ↑ is that the larger the value, the better the effect. The meaning of ↓ is that the smaller the value, the better the effect.

Additionally, our proposed method belongs to the 3DGS family, offering a significant improvement in computational efficiency compared to NeRF-based methods. In addition to static datasets, we also conducted comparative experiments on the underwater dataset In-the-Wild, which contains dynamic targets, and the results are shown in the Table 2.

In the second set of experiments, the dataset includes dynamic scenes, which involve motion and morphological changes of dynamic objects, thereby presenting higher reconstruction difficulty. The challenge of the Sardine sequence lies in the fact that each frame contains a large number of swimming sardines, resulting in significant variations in pixel depth and increased fitting complexity for Gaussian-based algorithms. Benefiting from our designed Gaussian point distribution optimization method and depthadaptive rendering

strategy, our algorithm partially eliminates interference from these dynamic objects, achieving superior rendering results. The difficulty of the Turtle sequence stems from the lack of sufficient varied viewing angles for reconstruction, which causes most existing methods to fail on this sequence. Although our approach cannot overcome limitations in model completeness, it achieves better underwater visual consistency through improved loss function design.

As shown in Table 2, WA-GS outperforms baselines across metrics including PSNR, SSIM, and LPIPS, particularly on the Composite, Coral, and Sardine datasets. In the Composite scene, WA-GS has an LPIPS of 0.129, which is 25.4% lower than that of the second - best method, WaterSplatting (0.173). This indicates WA-GS’s superior ability to model dynamic suspended particles and light interactions. The SSIM of WAGS is 0.867, close to the theoretical



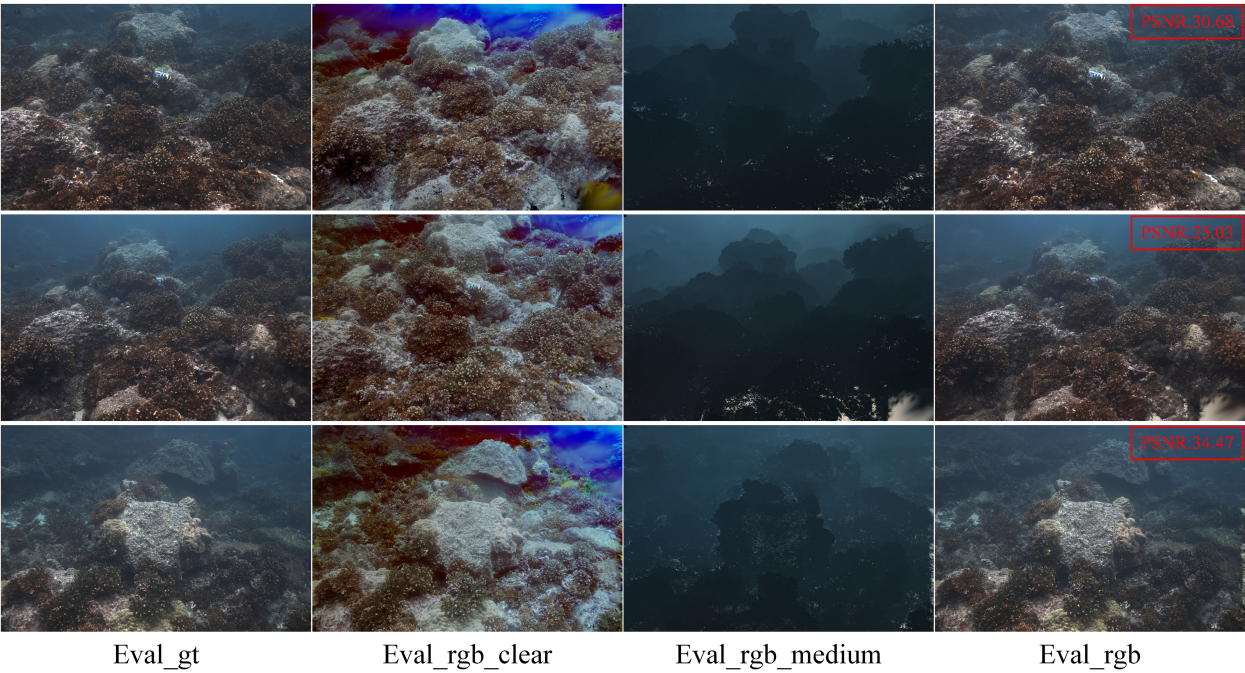


FIGURE 4
RGB rendering results on the Panama sequence. Eval gt indicates the real image, Eval rgb clear indicates the RGB image after underwater media is removed, Eval rgb medium indicates the RGB image after considering the influence of underwater media, and Eval rgb indicates the standard RGB image generated by the model.

maximum, which confirms the effectiveness of the multi scale radius strategy in reducing motion blur. WAGS also achieves the best results in the Coral and Sardine scenes. Figure 5 demonstrates that WA-GS significantly enhances underwater reconstruction quality, providing finer detail recovery and better visual consistency, establishing itself as a robust solution for dynamic underwater scenes. Figure 6 highlights the robustness of our method on highly challenging sequences, with superior performance in perceptual quality (LPIPS) and visual consistency (SSIM).

4.4 Ablation study and computational cost

To further validate the independent contribution of each innovation to the reconstruction performance, an ablation study was designed to analyze the role of each innovative component. The

experiments were conducted on the Coral sequence of the In-the-wild dataset, using WaterSplatting as the baseline method. The results of the ablation study are shown in the Table 3.

The results in Table 3 demonstrate that the synergistic effect of the components plays a crucial role in detail retention and visual consistency, leading to a significant improvement in these aspects for the WA-GS method compared to the baseline method. Although some improvements have increased computation time, in underwater reconstruction tasks, WA-GS has an LPIPS of 0.094, which is 48.9% lower than the baseline (0.184), and an SSIM of 0.904, which is 5.2% higher. Experiments show joint module optimization is key for WA-GS high-fidelity reconstruction in complex underwater scenes. The performance-cost tradeoff has significant engineering value. Complexity based Gaussian point distribution optimization and the new loss function slightly affect overhead. However, depth-adaptive

TABLE 2 Experimental results of In-the-Wild dataset.

Dataset	Composite			Coral			Turtle			Sardine		
Methods\Metrics	PSNR↑	SSIM↑	LPIPS↓	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
SeaThru-NeRF	16.21	0.406	0.829	23.89	0.649	0.405	27.06	0.882	0.192	21.37	0.578	0.606
Instant-NGP	22.81	0.597	0.572	20.87	0.439	0.731	26.42	0.874	0.225	21.73	0.649	0.467
DynamicNeRF	16.27	0.739	0.476	17.77	0.542	0.826	23.31	0.837	0.426	19.70	0.676	0.690
Tang et al.	25.09	0.799	0.239	26.17	0.828	0.157	28.10	0.900	0.217	21.58	0.723	0.454
WaterSplatting	25.11	0.840	0.173	27.48	0.859	0.184	21.77	0.845	0.350	19.14	0.691	0.622
WA-GS(ours)	25.45	0.867	0.129	28.66	0.904	0.094	24.58	0.886	0.229	22.81	0.857	0.212

The bold values in the table represent the optimal values.
The meaning of ↑ is that the larger the value, the better the effect. The meaning of ↓ is that the smaller the value, the better the effect.

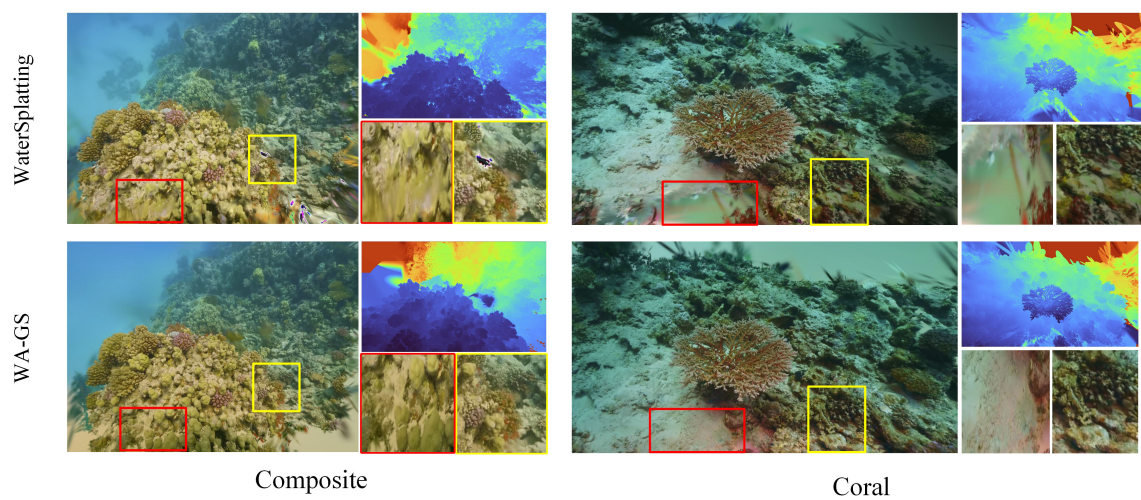


FIGURE 5
3D reconstruction results and depth results on the ‘Composite’ and ‘Coral’ sequences. Our proposed WA-GS achieves better visual consistency and reconstruction quality in complex underwater reconstruction tasks.

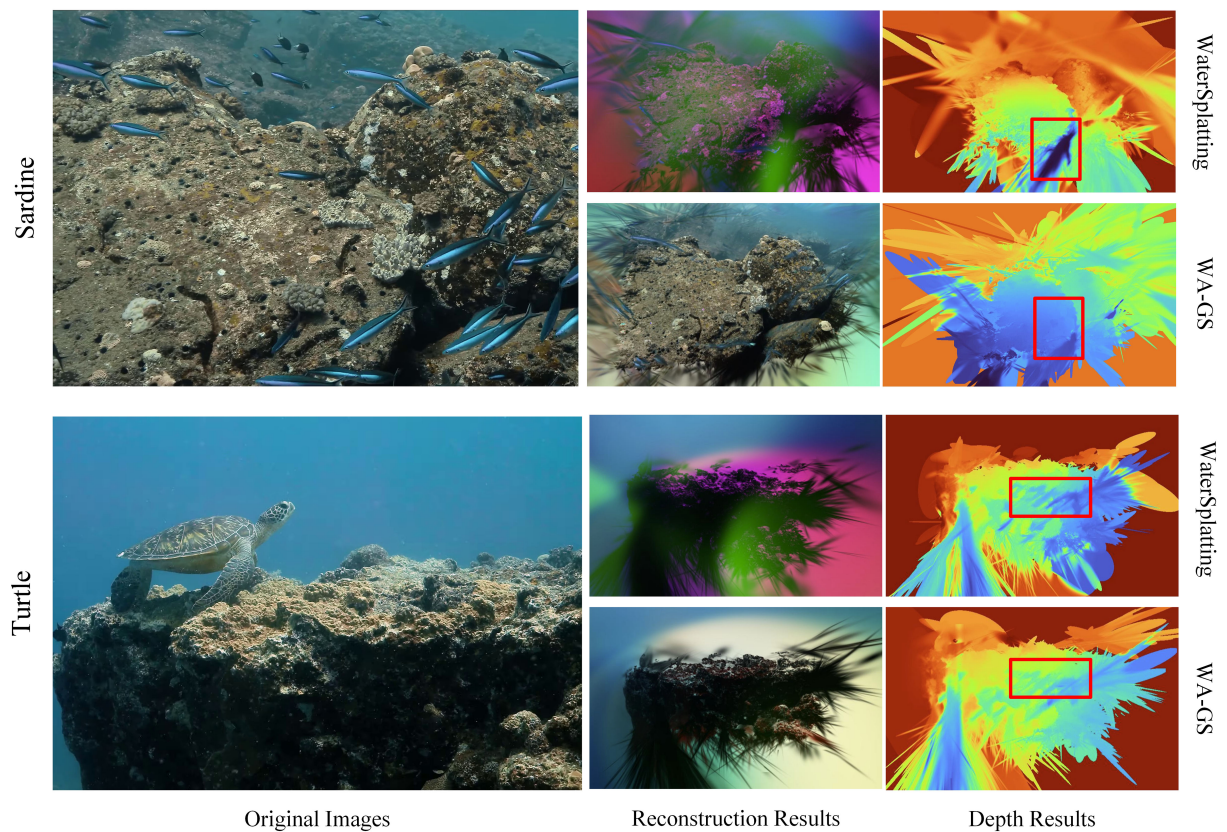


FIGURE 6
3D reconstruction results and depth results on the ‘Sardine’ and ‘Turtle’ sequences. Almost all existing methods fail to achieve good results on “Turtle” sequence. Thanks to our improvements, results that should have been judged as ‘failures’ were successfully rendered.

TABLE 3 Ablation study and computational cost statistics.

Sequence	Coral			Time
Ablation	PSNR↑	SSIM↑	LPIPS↓	
Basic	27.48	0.859	0.184	7min4s
WithoutA	27.70	0.870	0.138	18min43s
WithoutB	28.54	0.903	0.094	9min55s
WithoutC	28.62	0.903	0.107	13min58s
WA-GS(ours)	28.66	0.904	0.094	19min2s

A represents the complexity-based Gaussian point distribution optimization; B represents the depth-adaptive multi-scale rendering; C represents the underwater-specific loss function we proposed. The ablation study uses the common loss function of 3DGS: $L = (1 - \lambda) \cdot L_{L1} + \lambda \cdot L_{SSIM}$. The bold values in the table represent the optimal values. The meaning of ↑ is that the larger the value, the better the effect. The meaning of ↓ is that the smaller the value, the better the effect.

multi-scale rendering, which requires per-point depth-based scale factor calculation, increases overhead and training time.

5 Discussion

This paper addresses the challenges in underwater 3D scene reconstruction and proposes an innovative Water-Adapted 3D Gaussian Splatting method, which effectively handles the geometric structure of underwater scenes and the light scattering effects of the water medium. Our method solves issues such as blurring, color distortion, and detail loss in underwater imaging through complexity-based Gaussian point distribution optimization and depth-adaptive multi-scale rendering strategies, while providing a significant improvement in computational efficiency compared to traditional NeRF methods.

Experimental results show that the proposed method achieves higher rendering quality than existing underwater reconstruction methods on both static and dynamic scene datasets, and provides real-time rendering performance. Furthermore, the innovative strategies proposed in this paper have strong generalizability and can be easily transferred to other 3DGS-based reconstruction frameworks.

Despite achieving good visual consistency and reconstruction quality in underwater scenes, there are still many areas that require further work in the future. The depth-adaptive multi-scale rendering strategy introduces certain computational overhead. In practical applications, especially when handling large-scale scenes or complex dynamic environments, the computational burden may increase significantly. We plan to optimize this strategy and enhance its adaptability in realtime systems. Moreover, like most 3D Gaussian-based reconstruction methods, our approach relies on initial sparse point cloud information, which means it cannot directly generate complete 3D reconstruction results from RGB images. To further improve the flexibility of the system, we plan to explore ways to generate high-quality reconstruction results directly from RGB images without relying on COLMAP for sparse point cloud generation.

In view of growing ocean resource development and ecological conservation needs, future underwater 3D reconstruction research will center on:

- Multimodal intelligent sensing: In complex underwater conditions, single sensors are limited. Research will focus on building a sound-light-magnetism cross-physical-field sensing framework. Quantum sonar for lowlight distance measurement, polarized light imaging for suppressing scattering, and multispectral lidar for material structure identification will be explored. Integration of these will obtain comprehensive underwater scene data for 3D reconstruction, enhancing accuracy and integrity.
- Dynamic scene generalization ability: Given the dynamic nature of underwater scenes, work will involve using diffusion models and engines like PhyNex to create a dynamic light - field simulation platform. This will generate pretraining data for complex scenarios. Deep learning-based 3D reconstruction methods will be developed to learn dynamic patterns, enabling accurate real - time reconstruction for marine monitoring and facility assessment.
- Standardization and interdisciplinary: Future research will include formulating unified evaluation criteria for underwater 3D reconstruction to standardize work and accelerate iteration. As it spans multiple disciplines, efforts will be made to develop a seabed digital twin platform integrating multi-disciplinary knowledge for real-time seabed health management, problem detection, and earth science research support.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

XF: Formal analysis, Funding acquisition, Methodology, Software, Writing – review & editing. XW: Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. HN: Data curation, Validation, Visualization, Writing – review & editing. YX: Validation, Visualization, Writing – review & editing. PS: Funding acquisition, Project administration, Resources, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. National Key R&D Program of China, Grant/Award Number: 2022YFB4703400; the National Natural Science Foundation of China (grant numbers 62476080); Jiangsu Province Natural Science Foundation, grant number: BK20231186; Key Laboratory about Maritime Intelligent

Network Information Technology of the Ministry of Education (EKLMIC202405).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. During the preparation of this work,

the authors used the artificial intelligence language model ChatGPT to assist with translation, thereby improving the clarity and readability of the language. After using this tool, the authors reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Barron, J. T., Mildenhall, B., Verbin, D., Srinivasan, P. P., and Hedman, P. (2022). "Mip-nerf 360: Unbounded anti-aliased neural radiance fields," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. USA: IEEE COMPUTER SOC. 5470–5479.
- Barron, J. T., Mildenhall, B., Verbin, D., Srinivasan, P. P., and Hedman, P. (2023). "Zip-nerf: Antialiased grid-based neural radiance fields," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*. USA: IEEE COMPUTER SOC. 19697–19705.
- Cutolo, E., Pascual, A., Ruiz, S., Zarokanellos, N. D., and Fablet, R. (2024). Cloinet: ocean state reconstructions through remote-sensing, *in-situ* sparse observations and deep learning. *Front. Marine Sci.* 11, 1151868. doi: 10.3389/fmars.2024.1151868
- Gao, C., Saraf, A., Kopf, J., and Huang, J. B. (2021). "Dynamic view synthesis from dynamic monocular video," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*. New York, NY, USA: IEEE. 5712–5721.
- Hu, K., Wang, T., Shen, C., Weng, C., Zhou, F., Xia, M., et al. (2023). Overview of underwater 3d reconstruction technology based on optical images. *J. Marine Sci. Eng.* 11, 949. doi: 10.3390/jmse11050949
- Hua, X., Cui, X., Xu, X., Qiu, S., Liang, Y., Bao, X., et al. (2023). Underwater object detection algorithm based on feature enhancement and progressive dynamic aggregation strategy. *Pattern Recognition* 139, 109511. doi: 10.1016/j.patcog.2023.109511
- Jia, Y., Wang, Z., and Zhao, L. (2024). An unsupervised underwater image enhancement method based on generative adversarial networks with edge extraction. *Front. Marine Sci.* 11, 1471014. doi: 10.3389/fmars.2024.1471014
- Kerbl, B., Kopanas, G., Leimkühler, T., and Drettakis, G. (2023). 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.* 42, 139–131. doi: 10.1145/3592433
- Levy, D., Peleg, A., Pearl, N., Rosenbaum, D., Akkaynak, D., Korman, S., et al. (2023). "Seathrunker: Neural radiance fields in scattering media," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. USA: IEEE COMPUTER SOC. 56–65.
- Li, B., Chen, Z., Lu, L., Qi, P., Zhang, L., Ma, Q., et al. (2024a). Cascaded frameworks in underwater optical image restoration. *Inf. Fusion* 117, 102809. doi: 10.1016/j.inffus.2024.102809
- Li, H., Song, W., Xu, T., Elsig, A., and Kulhanek, J. (2024b). Watersplatting: Fast underwater 3d scene reconstruction using gaussian splatting. *arXiv preprint arXiv:2408.08206*.
- Li, H., Wang, H., Zhang, Y., Li, L., and Ren, P. (2025). Underwater image captioning: Challenges, models, and datasets. *ISPRS J. Photogrammetry Remote Sens.* 220, 440–453. doi: 10.1016/j.isprsjprs.2024.12.002
- Lin, Q., Feng, Z., Wang, Y., Wang, X., Bian, Z., Zhang, F., et al. (2024a). Light attenuation parameterization in a highly turbid mega estuary and its impact on the coastal planktonic ecosystem. *Front. Marine Sci.* 11, 1486261. doi: 10.3389/fmars.2024.1486261
- Lin, S., Sun, Y., and Ye, N. (2024b). Underwater image restoration via attenuated incident optical model and background segmentation. *Front. Marine Sci.* 11, 1457190. doi: 10.3389/fmars.2024.1457190
- Luo, J., Huang, T., Wang, W., and Feng, W. (2024). A review of recent advances in 3d gaussian splatting for optimization and reconstruction. *Image Vision Computing* 151, 105304. doi: 10.1016/j.imavis.2024.105304
- Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R. (2021). Nerf: Representing scenes as neural radiance fields for view synthesis. *Commun. ACM* 65, 99–106. doi: 10.1145/3503250
- Müller, T., Evans, A., Schied, C., and Keller, A. (2022). Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graphics (TOG)* 41, 1–15. doi: 10.1145/3528223.3530127
- Rout, D. K., Kapoor, M., Subudhi, B. N., Thangaraj, V., Jakhethiya, V., and Bansal, A. (2024). Underwater visual surveillance: A comprehensive survey. *Ocean Eng.* 309, 118367. doi: 10.1016/j.oceaneng.2024.118367
- Sethuraman, A. V., Ramanagopal, M. S., and Skinner, K. A. (2023). "Waternerf: Neural radiance fields for underwater scenes," in *OCEANS 2023-MTS/IEEE US Gulf Coast (IEEE)*. USA: IEEE Biloxi. 1–7.
- Tang, Y., Zhu, C., Wan, R., Xu, C., and Shi, B. (2024). "Neural underwater scene representation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. USA: IEEE COMPUTER SOC. 11780–11789.
- Wang, H., Anantrasirichai, N., Zhang, F., and Bull, D. (2024). Uw-gs: Distractor-aware 3d gaussian splatting for enhanced underwater scene reconstruction. *arXiv preprint arXiv:2410.01517*.
- Wang, H., Köser, K., and Ren, P. (2025). Large foundation model empowered discriminative underwater image enhancement. *IEEE Trans. Geosci. Remote Sens.* 63, 1–17. doi: 10.1109/TGRS.2025.3525962
- Yang, D., Leonard, J. J., and Girdhar, Y. (2024). Seasplat: Representing underwater scenes with 3d gaussian splatting and a physically grounded image formation model. *arXiv preprint arXiv:2409.17345*.
- Zhang, X., Chang, D., Ma, Y., and Du, J. (2024b). Precision detection method for ship shell plate molding based on neural radiance field. *Ocean Eng.* 309, 118459. doi: 10.1016/j.oceaneng.2024.118459
- Zhang, P., Yang, Z., Yu, H., Tu, W., Gao, C., and Wang, Y. (2024a). Rusnet: Robust fish segmentation in underwater videos based on adaptive selection of optical flow. *Front. Marine Sci.* 11, 1471312. doi: 10.3389/fmars.2024.1471312
- Zhou, J., Liang, T., Zhang, D., Liu, S., Wang, J., and Wu, E. Q. (2025). Waterhe-nerf: Water-ray matching neural radiance fields for underwater scene reconstruction. *Inf. Fusion* 115, 102770. doi: 10.1016/j.inffus.2024.102770