An effective pre-processing pipeline of NeRF 3D reconstruction for benthic target

Lyuyongkang Yuan
College of Literature, Science, and the Arts, University of Michigan,
Ann Arbor, Ann Arbor, 48109, The United States
lyuyong@umich.edu

Abstract. The nearshore oceans, home to rich benthic ecosystems, remain an area of significant research interest. While 2D visual representations have been the mainstay in this area, the intricate, multi-dimensional nature of the seafloor ecosystems underscores the need for 3D modeling to capture their full essence. This research introduces a methodology tailored for static image processing and 3D modeling using Neural Radiance Fields (NeRF), specifically the optimized instantNGP variant. A streamlined pipeline has been developed, focusing on mitigating the visual challenges posed by light interference and seawater in underwater imaging. This pre-processing approach effectively prepares images for the NeRF-based 3D reconstruction without an excessive computational burden. Visual enhancements successfully corrected color imbalances in underwater images, addressing the common blue-green tint caused by light conditions. Furthermore, by dynamically detecting and eliminating seawater borders, the pipeline ensures that the reconstruction models remain concentrated on the seafloor ecosystem. This process does not necessitate extensive datasets or immense computational resources, marking it as an efficient solution for near coast underwater images. Offering a cost-effective and efficient alternative to traditional methods, this research provides marine ecologists with a robust tool for RGB-based 3D modeling of nearshore environments. However, its application might benefit from integration with neural networks for better adaptability across various marine scenarios.

Keywords: Benthic 3D reconstruction, NeRF, Underwater Image Enhancement, Seawater Detection

1. Introduction

The vastness and intricacies of the world's oceans present significant challenges and opportunities for scientific exploration. One particular area of interest lies in the study of benthic organisms and ecosystems situated in the near shore regions [1]. Detailed analyses of these entities, predominantly driven by visual data, offer insights into the marine environment's complex dynamics. However, while advances in underwater imaging tools have provided substantial data, much of the marine research is still constrained to 2D visual representations.

Such two-dimensional portrayals, although informative, may not capture the full complexity of the seafloor, especially in nearshore environments. These areas boast a multitude of lifeforms, geological formations, and ecosystems, all interacting in a three-dimensional space. The limitation of 2D images suggests an urgent need for more comprehensive and immersive 3D models to depict the marine landscape accurately, for example, structured light [2] modeling is laborious to model and susceptible to underwater environments, while sonar modeling does not allow for RGB-based modeling [3] for follow-up studies.

To address this gap, this research introduces a method designed for static image processing, culminating in 3D modeling using Neural Radiance Fields (NeRF) [4]. The popularity of NeRF in recent research stems from its capability to transform 2D images into detailed 3D reconstructions. Unique to the approach proposed in this study is its efficient design, achieving commendable 3D outcomes without an excessive reliance on computational resources. This efficiency ensures that comprehensive 3D modeling remains accessible without being computationally prohibitive.

Central to this approach is a refined pipeline. This streamlined process, optimized for cost-effectiveness and efficiency, prepares images specifically for NeRF, ensuring enhanced 3D reconstruction. Underwater imaging inherently presents challenges, notably the impact of light refraction, diffusion, and absorption in water. Such challenges can result in distorted images with color inaccuracies. The methodology outlined in this study adeptly mitigates these issues, ensuring that 3D models are both accurate and representative of the true underwater environment.

The potential applications and implications of this methodology extend far beyond academic interest. Specifically, marine ecologists, especially those focusing on benthic ecology, can leverage this RGB-based 3D modeling technique. By providing a precise and efficient tool for understanding the nearshore environment, this study contributes to the field of marine ecology, potentially influencing both research methodologies and conservation strategies.

In summary, the oceans, with their vast and uncharted territories, continue to challenge and inspire scientific research. The move from 2D to 3D representations, as facilitated by the present research, promises to enrich the understanding of marine environments, enabling a deeper exploration of the nearshore marine world.

2. Background and Related Work

Underwater environments offer a unique vista that remains largely unexplored. The deep seas and oceans, with diverse terrains and marine life, hold invaluable information essential for fields ranging from marine biology to geology. Accurate mapping and visualization of these environments are pivotal for further scientific discoveries and advancements.

While satellite imaging and sonar technologies have made strides in underwater mapping, visual-based 3D reconstructions provide a richer, more detailed representation. These reconstructions rely heavily on high-quality input images. However, underwater imaging is fraught with challenges. Water absorbs and scatters light, leading to images plagued with non-uniform lighting, low contrast, and color cast, among other issues.

2.1 NeRF, InstantNGP, and Reconstruction

Neural Radiance Fields (NeRF) has established itself as a pivotal method for 3D scene reconstruction from 2D images. Instead of relying on traditional depth maps or voxel grids, NeRF uses a neural network to represent a continuous volumetric scene function. Its performance and adaptability in diverse environments have been emphasized in prior works, but its practicality in underwater settings has been less explored.

Building on the foundational principles of NeRF, the introduction of Instant Neural Graphics Primitives (InstantNGP) [5] marks a notable advancement in the field of 3D reconstruction. InstantNGP streamlines the reconstruction process, enabling near real-time 3D visualization. This accelerated performance is achieved by optimizing the underlying neural structures and leveraging more efficient rendering techniques. Given the time-sensitive nature of many underwater exploration tasks, the adoption of InstantNGP offers potential benefits in terms of rendering speed without compromising on the quality of the reconstruction. The current research aims to tailor and optimizing the capabilities of InstantNGP for underwater environments, exploring its nuances and potential advantages in subaqueous 3D modeling.

2.2 Machine Learning-Based Segmentation

Image segmentation in underwater environments poses unique challenges, primarily due to the scattering and absorption of light rays by water and suspended particles. A prevalent methodology to address this problem is based on machine learning techniques [6][7], which have shown significant capability in segmenting distinct objects submerged in water. Notably, recent efforts in this domain have largely gravitated towards segmenting marine life, often sidestepping the intricacies of the underwater terrain or seabed. While these methods have proven effective in many scenarios, they may have limitation on critical details pertaining to landforms.

2.3 Underwater image process

The realm of underwater image enhancement has seen considerable advancements in recent years. A distinctive method hinged on a conditional generative adversarial network (cGAN) was employed for real-time image betterment. This approach critically assessed perceptual image quality across various metrics, and to support this, introduced the large-scale EUVP dataset of underwater images. Notably, these enhanced images showed improved compatibility with standard underwater object detection models [8]. Another study explored adaptive histogram equalization, introducing regional histogram equalization for real-time enhancement, implemented using the Field Programmable Gate Array (FPGA) [9][10]. Dehazing in underwater images, a challenge due to the pronounced effects of light scattering in turbid waters, was addressed using an algorithm that estimated scene depth by leveraging the differential attenuation across image color channels [11].

3. System Design and Implementation

3.1 Methodology for Underwater Image Color Enhancement

In addressing the challenges of underwater imaging, particularly the non-uniform illumination and degraded contrast, a three-fold approach was adopted: Histogram Stretching, Color Balancing, and Contrast Limited Adaptive Histogram Equalization (CLAHE).

3.1.1 Histogram Stretching

To alleviate the issues of light absorption and scattering, histogram stretching was implemented on individual RGB channels of the image. The histogram of each channel was linearly transformed such that the minimum and maximum intensity values present were mapped to 0 and 255, respectively. For this process, pixels with zero intensity were ignored to ensure that only meaningful intensities were used to determine the stretching limits. This technique amplifies the color contrast, allowing for a clearer differentiation of features within the underwater environment.

3.1.2 Color Balancing

After stretching, the images often displayed a color cast, particularly due to the preferential absorption and scattering of specific wavelengths underwater. To tackle this, the mean of each channel (R, G, B) was computed and compared with the global mean (for all channels). Each channel was then adjusted to match the global mean, ensuring that the final image was free from color bias. This adjustment balanced the color distribution and minimized the inherent blue or green dominance common in underwater images.

3.1.3 Contrast Limited Adaptive Histogram Equalization (CLAHE)

While histogram stretching enhanced the global contrast, it sometimes left local regions of the image under-enhanced. To address this, CLAHE was employed to improve local contrasts. Unlike traditional histogram equalization, CLAHE operates on smaller, overlapping regions of the image, ensuring that the histogram of each tile is equalized independently. This method prevents over-amplification of contrasts and reduces noise. For this application, tiles of size 8 x 8 were used, and a clip limit of 1.0 ensured that the equalization did not produce overly contrasted images. To facilitate the processing of multiple images, an automation routine was implemented, allowing the enhancement of entire folders of images sequentially.

3.2 Seawater Border Detection for Enhanced 3D Reconstruction

For precise seabed 3D reconstruction using NeRF, it is imperative to distinguish between the seabed landforms and the overlaying seawater, which manifests as a progressively dominant blue shade in images. Consequently, NeRF may mistakenly reconstruct the seawater as an integral component of the seabed model. To counteract this effect, a methodology was developed to detect the border between the seabed landforms and the seawater.

3.2.1 Hue-based Image Segmentation

The color bounds for the seawater were dynamically calculated based on the image's height and the previously computed mean hue value. The lower bound was defined by halving the mean value, while the upper bound was set at a predefined maximum hue value, specifically tailored for blue tones typical of seawater.

3.2.2 Region of Interest (ROI) Definition

To avoid erroneous detections and enhance computational efficiency, only the top half of the image, where the seawater's presence is most evident, was considered as the Region of Interest (ROI).

3.2.3 Seawater Masking

Within the defined ROI, a mask was created to isolate the pixels that fell within the seawater color bounds. Contour detection was applied to this mask to identify and filter out smaller noise areas, ensuring only significant blue regions were kept. To further refine the detection, only contours that met a minimum area threshold and touched the top edge of the ROI were preserved.

3.2.4 Post-processing and Smoothing

Post the primary segmentation, contours were merged, and morphological operations were applied to smooth and close gaps in the detected regions. A closing operation using a rectangular kernel efficiently filled small holes and connected nearby contours. The resultant mask was then applied to the original image, blackening the detected seawater regions.

3.2.5 Batch Image Processing

An automation pipeline was designed to process multiple images. Each image from the source was loaded, the mean hue value was calculated, and subsequent seawater border detection was carried out. By identifying and segmenting out the seawater regions, the images were made more suitable for seabed 3D reconstruction with NeRF, ensuring that the resulting model depicted the seabed landforms with increased fidelity.

4. Results

4.1 Underwater Image Color Enhancement Qualitative Evaluation

Visual Enhancement: A series of image pairs, each juxtaposing the original with the enhanced counterpart, showcased the transformation engendered by the proposed techniques. It was discernible that the pervasive blue-green tint in the original images was remarkably alleviated in the enhanced versions, paving the way for a more natural color representation. As shown in Figure 1.

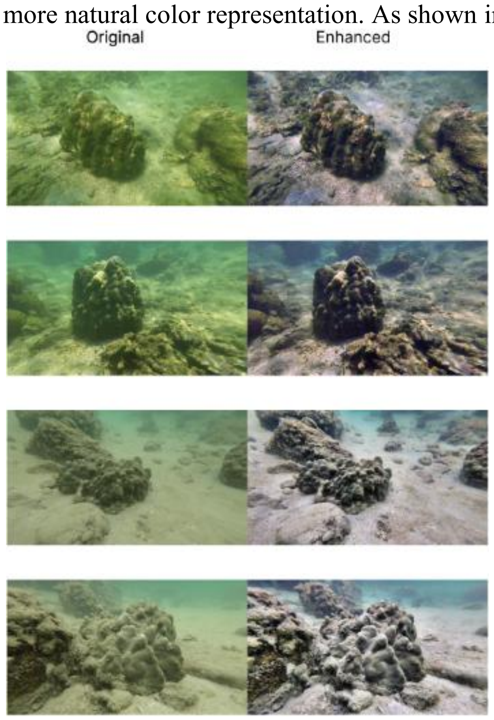

Figure 1. Original vs. Enhanced diagram (Photo/Picture credit: Original)

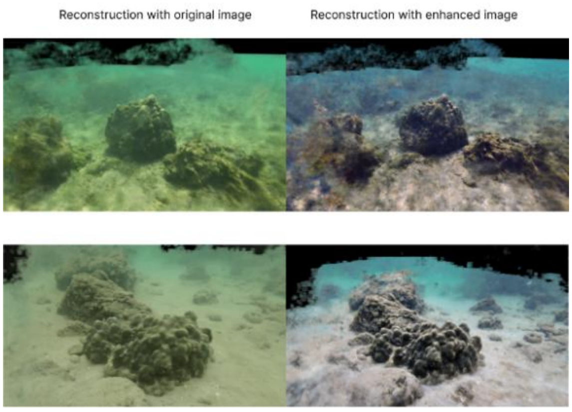

Figure 2. 3D reconstruction based on original / enhanced images (Photo/Picture credit: Original)

Beyond color rectification, a notable enhancement in the visibility of submerged features was observed. Details, previously obfuscated due to unfavorable lighting conditions, were now discernible, indicating the effectiveness of the methodology in mitigating scattering effects and improving image clarity. As shown in Figure 2. Histogram Analysis: In the evaluation of the image enhancement methodology, the histograms of both the original and enhanced images were analysed. The histograms for each BGR channel of the original image exhibited a pronounced drop, with the green and red channels falling to zero at an intensity of 200 and the blue channel even earlier at 150. This indicates an absence of higher intensities, confirming the visual observation of subdued brightness and potential loss of details in the brighter regions, a common trait in underwater imagery due to selective color absorption. In contrast, the histograms of the enhanced images present a more continuous distribution, stretching further along the intensity spectrum. This behavior signifies the successful restoration of the suppressed pixel intensities, resulting from the applied enhancement techniques. As shown in Figure 3.

Figure 3. Original vs. Enhanced image histogram subjects (Photo/Picture credit: Original).

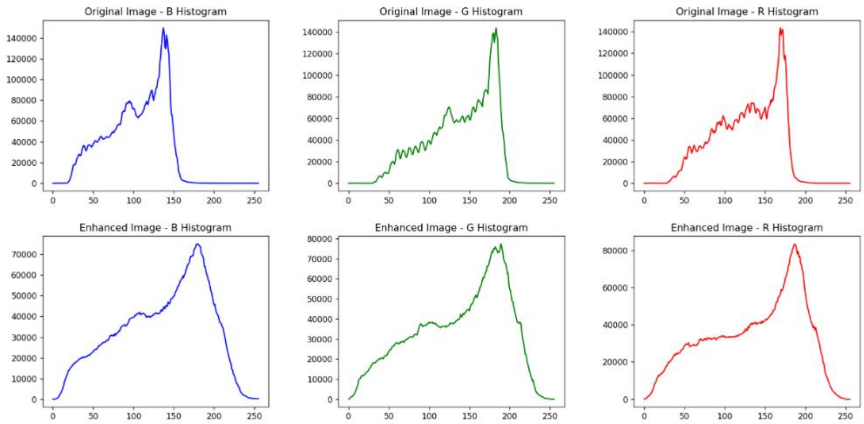

Figure 4. Experimental results (Photo/Picture credit: Original)

Border Detection (Seawater vs. Seabed): In the endeavour to enhance the accuracy of seabed 3D reconstruction using InstantNGP, an essential step involves detecting and isolating the boundary between the seawater and the seabed. The water, especially seawater, presents a significant challenge as it tends to impose a blue tint on photographs due to its light scattering properties. When reconstructing a 3D model, this tint can lead to erroneous inclusion of the water as an actual instance in the output model. As shown in Figure 4. Contour Analysis: On applying the proposed border detection algorithm, contour maps were generated delineating the regions of seawater from the seabed. The methodology employed hue-based detection targeting the characteristic color of the seawater. The resultant contours were smooth, indicating an acceptable border establishment between the seawater and seabed regions.

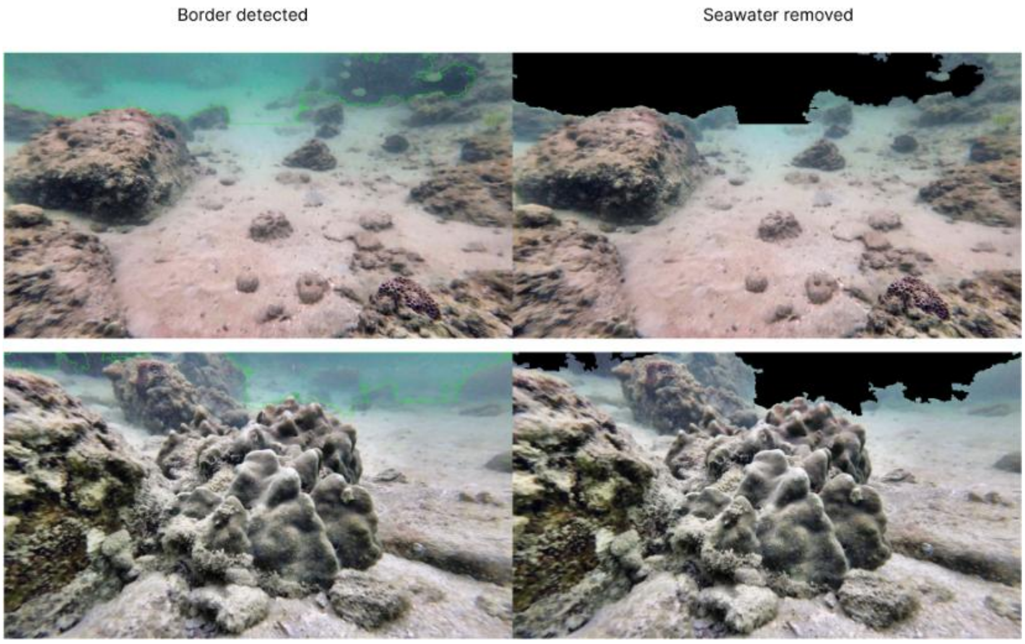

Figure 5. Border detection examples (Photo/Picture credit: Original)

Following the enhanced detection, preliminary evaluations of 3D reconstructions using NeRF were conducted. It was observed that by effectively mitigating the seawater influence in the input images, the reconstructed 3D models exhibited significantly fewer water instances. As shown in Figure 5.

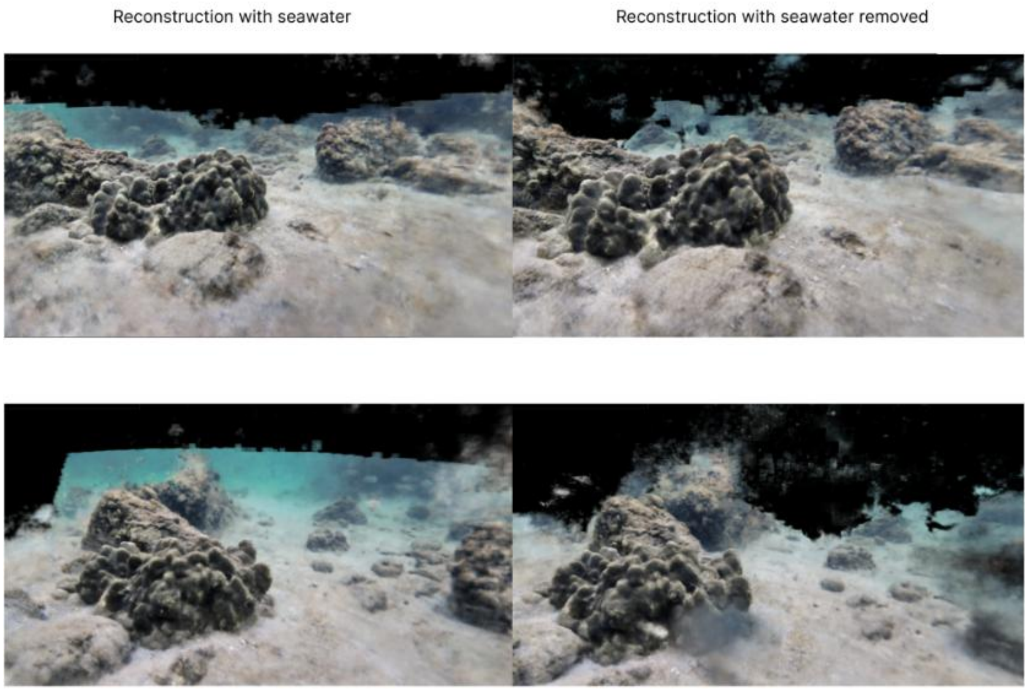

Figure 6. 3D reconstruction with seawater removed (Photo/Picture credit: Original).

One of the distinctive benefits of this pipeline is its inherent computational efficiency. Unlike deep learning-based approaches, which often require extensive computational resources a training time, this pipeline operates a lower level of image processing, thus necessitating fewer computational resources. This makes the approach more accessible, especially in settings where high-end computational capabilities are a constraint. Furthermore, a significant advantage of this pipeline is the absence of a dependency on training datasets. Acquiring data, particularly for specific scenarios like near coast underwater environments, can be challenging. In contrast, the developed pipeline can be applied directly to new underwater images without the need for prior training or dataset availability.

5. Challenges and Limitations

While the proposed methodology offers several advantages, there are also challenges and limitations inherent to this approach.

5.1 Limited Generalization Capability

One of the primary challenges is the pipeline's specificity to the 1–3-meter foreshore area. This focus makes its application in broader underwater settings questionable. While the current process can handle certain coastal environments efficiently, its performance may be compromised in diverse underwater conditions.

5.2 Dynamic Seawater Boundary Detection

Accurately detecting and handling the variable nature of seawater across images remains a challenge. The appearance and effects of seawater in imagery are contingent on the scene's distance from the camera due to light attenuation over distance. The method would benefit from a more dynamic algorithm that understands the relationship between hue attenuation and distance, optimizing boundary detection accordingly.

5.3 Geographical Limitations

The method's design assumes a specific orientation of seawater – tapering from top to bottom within the frame, owing to its focus on offshore seabed modeling. This orientation may not hold true for various marine regions. Incorporating more adaptable techniques, potentially involving deep learning networks, could assist in identifying the modeling seawater components across different marine settings.

6. Conclusion

The research has developed a comprehensive framework tailored for the challenges posed by underwater imaging, particularly in the nearshore regions. By emphasizing the importance of pre-processing techniques, the pipeline has effectively mitigated issues related to light interference and seawater presence, creating an idea input for the InstantNGP-based NeRF 3D reconstruction. As visual assessments have affirmed, the outcome is a more accurate representation of the near-coast underwater environment, providing invaluable tools for marine ecologists.

While the pipeline offers numerous advantages, including cost-effectiveness and computational efficiency, its specificity to certain marine contexts underscores the need for adaptable methods in broader scenarios.

7. Future Work

7.1 Integration with Neural Networks for Enhanced Segmentation

The methodology's core, while efficient, is ripe for augmentation with neural networks. By combining with simple network architectures, the segmentation processes can be further optimized [12][13]. Notably, this could open avenues for extending the system's capability to address more specific challenges like measuring areas affected by coral bleaching. Such an enhancement, while retaining the pipeline's efficiency, should significantly bolster its generalization capabilities across varied marine environments.

7.2 Deployment on Micro Autonomous Underwater Vehicles (AUVs)

A future direction worth exploring is the application of this pipeline on micro AUVs [14]. The potential for real-time or near-real-time underwater environment reconstructions presents possibilities. By being equipped with this tool, AUVs could automatically investigate near-coast regions, planning routes or gathering crucial environmental information, thereby furthering our understanding and conservation efforts for these vital ecosystems.

Reference:

[1] de Juan, S., Bremner, J., Hewitt, J., Törnroos, A., Mangano, M. C., Thrush, S., & Hinz, H. (2022). Biological traits approaches in benthic marine ecology: Dead ends and new paths. Ecology and Evolution, https://doi.org/10.1002/ece3.9001

[2] Zhao, J., Cheng, Y., Cai, G., Feng, C., Liao, L., & Xu, B. (2022). Correction model of linear structured light sensor in underwater environment. Optics and Lasers in Engineering. https://doi.org/10.1016/j.optlaseng.2022.107013

[3] Kwon, S., Park, J., & Kim, J. (2017). 3D reconstruction of underwater objects using a wide-beam imaging sonar. 2017 IEEE Underwater Technology (UT) (pp. 1-4). Busan, Korea (South). https://doi.org/10.1109/UT.2017.7890306

[4] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2020). NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. arXiv. https://doi.org/10.48550/arXiv.2003.08934

[5] Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. arXiv. https://doi.org/10.1145/3528223.3530127

[6] Sarma, K., & Vigneshwaran, P. (2022). Underwater Image Enhancement Using Deep Learning. In J.IZ. Chen, J.M.R.S. Tavares, A.M. Iliyasu, & KL. Du (Eds.), Second International Conference on Image Processing and Capsule Networks. ICIPCN 2021. Lecture Notes in Networks and Systems, vol 300. Springer, Cham. https://doi-org.proxy.lib.umich.edu/10.1007/978-3-030-84760-9_38

[7] Venkatramanan Sethuraman, A., Ramanagopal, M. S., & Skinner, K. A. (2022). WaterNeRF: Neural Radiance Fields for Underwater Scenes. arXiv. https://doi.org/10.48550/arXiv.2209.13091

[8] Islam, Md J., Xia, Y., & Sattar, J. (2020). Fast Underwater Image Enhancement for Improved Visual Perception. arXiv. https://doi.org/10.48550/arXiv.1903.09766

[9] Alex, R. S. M., Deepa, S., & Supriya, M. H. (2016). Underwater image enhancement using CLAHE in a reconfigurable platform. OCEANS 2016 MTS/IEEE Monterey (pp. 1-5). Monterey, CA, USA. https://doi.org/10.1109/OCEANS.2016.7761194

[10] Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J. B., & Zuiderveld, K. (1987). Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing, 39(3), 355–368.

[11] Carlevaris-Bianco, N., Mohan, A., & Eustice, R. M. (2010). Initial results in underwater single image dehazing. OCEANS 2010 MTS/IEEE SEATTLE (pp. 1-8). Seattle, WA, USA. https://doi.org/10.1109/OCEANS.2010.5664428

[12] Zhang, J., Gao, J., Liang, J., Wu, Y., Li, B., Zhai, Y., & Li, X. (2023). Efficient Water Segmentation with Transformer and Knowledge Distillation for USVs. J. Mar. Sci. Eng., 11, 901. https://doi.org/10.3390/jmse11050901

[13] Drews-Jr, P., Souza, I.d., Maurell, I.P. et al. (2021). Underwater image segmentation in the wild using deep learning. J Braz Comput Soc, 27, 12. https://doi.org/10.1186/s13173-021-00117-7

[14] Billings, Gideon. (n.d.). Visual Methods Towards Autonomous Underwater Manipulation.