

AI-Driven Marine Robotics: Emerging Trends in Underwater Perception and Ecosystem Monitoring

Scarlett Raine, Tobias Fischer

QUT Centre for Robotics, Queensland University of Technology
2 George St, Brisbane, QLD, Australia
{sg.raine, tobias.fischer}@qut.edu.au

Abstract

Marine ecosystems face increasing pressure due to climate change, driving the need for scalable, AI-powered monitoring solutions to inform effective conservation and restoration efforts. This paper examines the rapid emergence of underwater AI as a major research frontier and analyzes the factors that have transformed marine perception from a niche application into a catalyst for AI innovation. We identify three convergent drivers: i) environmental necessity for ecosystem-scale monitoring, ii) democratization of underwater datasets through citizen science platforms, and iii) researcher migration from saturated terrestrial computer vision domains. Our analysis reveals how unique underwater challenges—turbidity, cryptic species detection, expert annotation bottlenecks, and cross-ecosystem generalization—are driving fundamental advances in weakly supervised learning, open-set recognition, and robust perception under degraded conditions. We survey emerging trends in datasets, scene understanding and 3D reconstruction, highlighting the paradigm shift from passive observation toward AI-driven, targeted intervention capabilities. The paper demonstrates how underwater constraints are pushing the boundaries of foundation models, self-supervised learning, and perception, with methodological innovations that extend far beyond marine applications to benefit general computer vision, robotics, and environmental monitoring.

1 Introduction

Marine ecosystems, including coral reefs and seagrass meadows, are under increasing threat from climate change, with rising sea temperatures, ocean acidification, and more frequent extreme weather events driving widespread degradation (Anthony et al. 2020). Effective monitoring and large-scale restoration are urgently needed to safeguard the biodiversity and ecological services these systems provide (Ditria et al. 2022). However, the scale of the challenge is immense: coral reefs span hundreds of thousands of square kilometers globally, while seagrass meadows form vast and often remote coastal habitats that are difficult to access and monitor comprehensively (Lavery et al. 2013).

Traditional monitoring approaches, based on in-water surveys conducted by divers and domain experts, have played a central role in ecological assessment for decades (Murphy and Jenkins 2010). Yet these methods are costly, time-

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

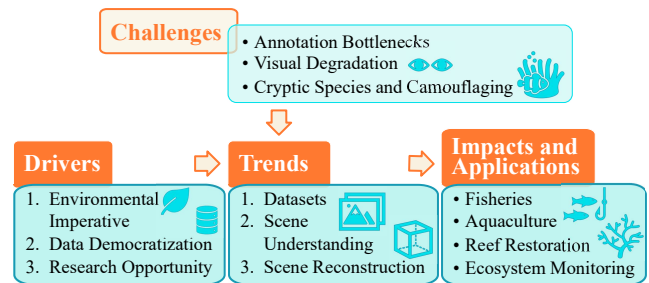


Figure 1: Visual outline. We discuss the three main drivers for underwater AI as a research field of growing importance, identify key trends within underwater perception, placing these in the context of the underwater-specific challenges that have inspired innovations, and finally discuss the real-world impacts and applications of these advances.

intensive, and constrained by human safety limitations, dive duration, and accessibility to remote locations (McKenzie et al. 2020; Raine et al. 2020). The deployment of robotic underwater and surface vehicles, along with towed camera platforms, has begun to mitigate these limitations by enabling the collection of large quantities of image data across previously inaccessible areas. However, this technological shift has created a new bottleneck: the need for automated approaches to analyze the vast quantities of visual data generated (Gonzalez-Rivero et al. 2020; Raine et al. 2024a).

This bottleneck has catalyzed the emergence of AI-driven underwater perception as a rapidly growing research frontier. The convergence of three key drivers is transforming what was once a niche application domain into a frontier for AI innovation (Fig. 1). First, the environmental imperative for scalable marine monitoring has created unprecedented demand for automated solutions that can operate at ecosystem scales. Second, the democratization of underwater data through citizen science initiatives and the release of large-scale datasets has provided the data foundation required for modern AI. Third, as traditional computer vision applications become increasingly saturated, researchers are turning to the

underwater domain as a comparatively understudied environment that offers opportunities for novel methodological contributions while addressing pressing societal needs.

These converging factors have enabled a paradigm shift in underwater perception: from manual species counts and expert-driven habitat assessments toward AI systems capable of real-time scene understanding, automated species detection, and even active intervention for habitat restoration. The unique challenges of the underwater environment, including turbidity, variable lighting conditions, cryptic species, and the need for fine-grained biological expertise, are driving innovations in self-supervised learning, open-set recognition, and robust perception that extend far beyond marine applications.

This paper provides a critical analysis of the emerging trends in AI for underwater perception and marine robotics (Fig. 1). We first identify and motivate the key drivers of underwater perception as an emerging field (Section 2). Next, we discuss recent breakthroughs in key technical areas including datasets, scene understanding and 3D reconstruction (Section 3), and conclude by exploring the broader implications of this emerging field for both marine science and the advancement of AI (Section 4).

2 Why Now? Convergent Drivers

The emergence of AI for underwater perception and marine robotics as a major research frontier is not coincidental, but rather the result of three powerful drivers converging at this moment in time. Understanding these drivers is crucial for appreciating both the urgency and the unprecedented opportunities that define this emerging field.

2.1 Environmental Imperative

The scale and urgency of marine ecosystem degradation has reached a tipping point, with some projections estimating a loss of up to 90% of coral coverage before 2050 (Souter et al. 2021). Ocean warming has accelerated dramatically, with marine heatwaves becoming five times more frequent since the 1980s, triggering mass coral bleaching events that can affect thousands of square kilometers (Anthony et al. 2020; Henley et al. 2024). In the face of this crisis, marine surveys are needed to detect changes early, track ecosystem health over time, and inform management and intervention strategies (Ditria et al. 2022). However, Australia’s Great Barrier Reef alone spans over 344,000 square kilometers, an area larger than the entire United Kingdom, making comprehensive manual monitoring logistically impossible.

This crisis has created unprecedented demand for ecosystem-scale monitoring with minimal human intervention. Blue carbon markets and restoration programs likewise require scalable, quantitative assessments to support policy, carbon verification, and intervention tracking (Gonzalez-Rivero et al. 2020; Murphy and Jenkins 2010). The gap between traditional dive-based surveys and the needs of large-scale conservation is now a major driver for AI-powered automation.

2.2 Data Democratization

The second driver represents a fundamental shift in how underwater data is collected, labeled, and shared. Historically,

underwater imagery was manually collected by divers and snorkelers during expensive research expeditions and labeled by marine biologists, resulting in small, specialized datasets accessible only to domain experts. This paradigm has been disrupted by the convergence of several technological and social trends.

Underwater data collection has shifted from small, expert-driven efforts to large-scale, democratized initiatives powered by affordable imaging, citizen science, and corporate participation. Programs like Virtual Reef Diver, FathomVerse, and the Great Reef Census engage thousands of volunteers in annotation, while platforms such as CoralNet, ReefCloud, and Data Mermaid support semi-automated labeling. Combined with self-supervised learning, these developments create a feedback loop that accelerates data generation, model training, and public engagement (see Section 3).

2.3 Research Opportunity

The third driver reflects a broader shift in the AI landscape. Traditional computer vision tasks reached a level of maturity where incremental improvements require vast computational resources and engineering effort. In these fields, foundation models are increasingly performing multiple computer vision tasks simultaneously (Siméoni et al. 2025). Meanwhile, performance on standard benchmarks has plateaued, and many researchers are seeking new domains where fundamental algorithmic innovations can still yield transformative advances.

The underwater domain offers exactly this opportunity. Unlike terrestrial environments, underwater perception presents unique challenges that cannot be solved by simply applying existing methods: extreme lighting variations, backscatter and turbidity effects, color distortion with depth, and a diversity of marine life that exhibits camouflaging and cryptic behaviors. Foundation models developed for land-based applications typically do not directly translate to underwater applications, due to the sizable domain gap, largely caused by challenging underwater image characteristics, fine-grained classes, and ecological considerations (Raine et al. 2024c), as explained further in Section 3. These challenges require genuine methodological innovation rather than incremental engineering improvements.

2.4 Convergence Effect

The true power of these drivers lies not in their individual impact, but in their convergence. Environmental urgency provides the motivation and funding, democratized data provides the foundation for modern AI approaches, and the research opportunity attracts top talent and computational resources. The underwater domain offers a compelling value proposition: the chance to develop novel AI techniques while addressing urgent environmental challenges. This has already attracted talent and funding from conservation agencies, creating a self-reinforcing cycle where advances enable new applications, generate more data, and drive further breakthroughs.

This convergence has reached a critical mass, transforming underwater perception from a niche application area into one of the most rapidly growing frontiers in AI research, with

implications that extend far beyond marine science into fundamental questions of robust perception, domain adaptation, and human-AI collaboration in challenging environments.

3 Emerging Trends

This section provides a critical analysis of the recent and emerging trends in underwater perception. We do not provide an exhaustive survey of prior approaches, but instead highlight key works that we believe are driving innovation in the field and opening up new research directions. Where relevant, we also identify the specific challenges of the underwater domain driving the technical advances.

3.1 Datasets

As mentioned in Section 2, the transformation of underwater AI has been fundamentally enabled by a dramatic shift in data availability, from small, expert-curated datasets to large-scale, community-driven repositories that rival terrestrial computer vision benchmarks. This data revolution represents more than simple scale increases; it reflects a fundamental change in how underwater imagery is collected, annotated, and shared across research communities.

Citizen Science as a Data Engine: The proliferation of affordable underwater cameras, remotely operated vehicles (ROVs), and autonomous underwater vehicles (AUVs) has dramatically increased the volume of imagery being collected, enabling the emergence of large-scale, open-access datasets. Citizen science has become a powerful mechanism for annotating this data, democratizing participation in marine science at unprecedented scales.

One early initiative, the Virtual Reef Diver (2025), encouraged divers and snorkelers to upload reef images, which were then annotated by everyday citizens. These observations, combined with professional monitoring data, improved predictive maps of coral cover across the Great Barrier Reef (Vercelloni, Santos-Fernández, and Mengersen 2023). Similarly, Citizens of the Reef has shown that citizen scientists can reach expert-level accuracy in coral reef monitoring, with deep learning and crowd-sourced analysis producing 99% accurate benthic cover estimates across 8,086 images (Lawson et al. 2025). The Seatizen Atlas extends this model globally, contributing over 1.6 million geolocated images from the Southwest Indian Ocean collected by both citizens and researchers (Contini et al. 2025a).

Gamification and Corporate Social Responsibility: The rise of Corporate Social Responsibility (CSR) initiatives and gamification has further driven public participation in labeling underwater imagery. Gamified platforms such as FathomVerse (2025), NASA’s Nemo-Net (NASA 2025; Van den Bergh et al. 2021), and NOAA’s Click-a-Coral (Pierce 2025), each with thousands of downloads and users, transform annotation tasks into engaging, interactive experiences. These projects make large-scale annotation efforts more achievable by reaching audiences beyond scientific communities.

Complementing these gamified initiatives, the Citizens of the Reef (2025) Great Reef Census integrates citizen science into education and corporate structures. The Schools Program allows students to learn about reef ecosystems while

contributing data, and the Hour of Power initiative mobilizes CSR by encouraging employees from companies including Atlassian, Dell, Disney, Mars, Cotton On, and Salesforce to contribute one hour of labor for labeling reef imagery. Together, gamification and CSR expand participation, enhance ocean literacy, and strengthen stewardship of marine ecosystems.

Specialized Datasets Driving Technical Innovation: Specialized datasets are also pushing the boundaries of what underwater AI can achieve. CoralVOS introduces dense coral video segmentation, enabling frame-by-frame tracking of coral organisms and advancing beyond sparse point-based analysis (Zheng et al. 2023a). The Sweet Corals dataset demonstrates the integration of 3D reconstruction and species classification, providing more than 90,000 georeferenced images alongside 3D point clouds and Gaussian Splatting models that enable both spatial and biological analysis (Nozdrenkov et al. 2025). CoralScapes establishes the first general-purpose dense semantic segmentation benchmark for coral reefs, analogous to Cityscapes in terrestrial computer vision (Sauder et al. 2025). FathomNet represents a paradigm shift in marine data curation, providing expert-annotated imagery spanning the full diversity of ocean life, from common coastal species to rarely encountered deep-sea organisms that present novel vision challenges (Katija et al. 2022).

Data Consolidation and Standardization: Various efforts consolidate fragmented datasets collected under different labeling schemes and methodologies, unifying resources across diverse seafloor environments. With more than 11.4 million curated images and 3.1 million annotations standardized to consistent taxonomic schemes, BenthicNet enables large-scale model training that generalizes across ecosystems and imaging conditions (Lowe et al. 2025). Marine-Inst20m (Zheng et al. 2024a) combines fifty data sources to yield a dataset with 19.2 million instance masks.

Co-Occurrence With Advances in Self-Supervised Learning: Critically, this data revolution coincides with advances in self-supervised learning, which transform previously unusable unlabeled imagery into valuable training resources. Repositories such as CoralNet (Bejjbom et al. 2012), XL CATLIN (González-Rivero et al. 2014, 2019), and Squidle+ (Squidle 2025) contain millions of images with weak or sparse labels that are increasingly useful under modern frameworks (see Section 3.2). Platforms such as ReefCloud (Australian Institute of Marine Science (AIMS) 2024; ReefCloud 2025), CoralNet, and Data Mermaid (2025) are also integrating semi-automated AI annotation tools, reducing cost and effort while expanding labeled datasets.

The result is a positive feedback loop: better datasets enable stronger models, which attract more researchers and volunteers, leading to further data collection and methodological advances. This interplay of citizen science, CSR, gamification, specialized datasets, and self-supervised learning is fundamentally reshaping underwater AI and creating insights applicable to other domains where expert annotation is scarce and data collection is challenging.



Figure 2: Fine-grained visual similarity of marine species. Left: Coral from the *Stylophora* family. Right: Coral from the *Pocillopora* family. Images taken by Maria Costantini and Tobin Sparling and released under CC BY-NC.

3.2 Scene Understanding

The underwater environment presents a unique combination of perceptual and operational challenges that have historically limited automated scene understanding. However, these same challenges now serve as catalysts for AI innovation, driving the development of novel techniques that advance not only marine robotics but the broader fields of computer vision and machine learning. Rather than viewing underwater conditions as obstacles to overcome, the AI community is increasingly recognizing them as opportunities to push methodological boundaries in ways that benefit numerous application domains. As a result, underwater scene understanding has undergone a fundamental transformation, shifting from narrow, task-specific models to general-purpose foundation models that can handle the full complexity of marine ecosystems. This evolution mirrors broader trends in AI but has been accelerated by the unique demands of underwater environments.

Tackling Annotation Bottlenecks With Weakly Supervised Learning: Marine ecosystems exhibit extraordinary biodiversity, with over 800 hard coral species alone (Dietzel et al. 2021), each requiring specialized taxonomic expertise to identify accurately (Fig. 2). Traditional supervised learning approaches for marine species recognition rely on large, expertly-annotated datasets, a prohibitively expensive and slow process that creates a scalability bottleneck. This challenge has made the underwater domain a natural testbed for weakly supervised and few-shot learning approaches.

The marine biology community’s existing practice of point-based sampling, born from the Coral Point Count method (Kohler and Gill 2006), has given rise to recent advances in weakly supervised learning, creating a rich ecosystem of techniques that extract maximum information from minimal expert input. Point-based annotation in the underwater domain typically describes randomly distributed or grid-spaced sparse points which are labeled by marine scientists (Alonso et al. 2019). This differs from standard computer vision domains, where it is common to have large datasets of images labeled with dense pixel-level labels.

Our recent point label aware superpixel method demonstrated that these sparse expert annotations can be effectively propagated by clustering pixels using learned features while respecting point label constraints, achieving substan-

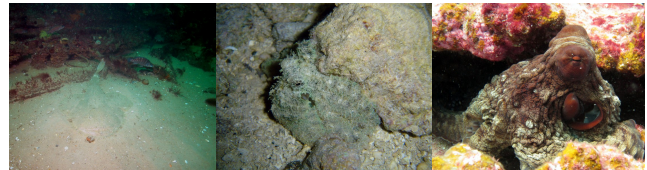


Figure 3: Cryptic marine species. Left: a monkfish, center: a scorpionfish (image taken by David Burdick), and right: an octopus (image taken by Kevin Lino). Images are from the NOAA (2025) image collection, released into the public domain.

tial improvements while reducing computational requirements (Raine et al. 2022). This success has inspired even weaker supervision approaches that operate with image-level labels alone, requiring 25 times fewer annotations than patch-level methods (Raine et al. 2024b). For seagrass meadow monitoring, contrastive pre-training combined with feature similarity has enabled coarse segmentation from image-level supervision, outperforming traditional patch-supervised methods on multi-species classification tasks (Raine et al. 2024b).

The emergence of foundation models has further transformed this landscape, with DINOv2 features now capable of generating multi-species coral masks without any domain-specific pre-training (Raine et al. 2024c). When integrated with human-in-the-loop labeling strategies that strategically leverage expert knowledge, these approaches achieve remarkable gains in efficiency (Raine et al. 2024c). Similarly, Contini et al. (2025b) proposed a framework for weakly supervised semantic segmentation of coral reef aerial imagery, based on DINOv2 for generation of coarse spatial predictions and mask refinement with Segment Anything. This progression from dense supervision to sparse points to image-level labels, and finally to foundation model-assisted annotation, exemplifies how domain constraints can drive methodological innovation with broad applicability across fields requiring expert knowledge integration.

Visual Degradation as a Driver for Robust Perception: Underwater imagery suffers from systematic visual degradation that can cripple traditional computer vision systems: light attenuation causes dramatic color shifts with depth, suspended particles create backscatter and reduce contrast, and turbidity can render scenes nearly inscrutable. These conditions represent an extreme case of the domain shift and image corruption problems that plague real-world AI deployment across many applications.

There have been recent advances to tackle underwater visual degradation through transformer-based image enhancement (Chen et al. 2024; Qiao, Guan, and Zhi 2025). Sauder et al. (2023) introduced self-supervised networks that estimate and correct backscatter and caustics from monocular sequences, improving real-time perception without ground-truth data. Similarly, Wang et al. (2023) and Li, Chen, and Miao (2025) proposed physics-guided transformer architectures to restore contrast and correct color shifts in challenging underwater conditions.

Addressing Cryptic Species and Camouflaging with Fine-Grained Recognition: Marine organisms have evolved sophisticated camouflage and cryptic coloration strategies that represent the ultimate challenge for visual recognition systems (Fig. 3). Fish that perfectly mimic coral textures, octopuses that dynamically change color and pattern, and invertebrates that are virtually indistinguishable from their substrates push the boundaries of what current computer vision can achieve (Hanlon 2007).

Recent investigations into performance disparities across marine species have uncovered that not all detection challenges are created equal, and that data quantity alone cannot explain why certain species consistently underperform in automated detection systems. Analysis of underwater object detection reveals that foreground-background discrimination represents the most critical bottleneck for cryptic species like scallops, regardless of training data availability. Even when class imbalance is artificially corrected, persistent precision gaps indicate that certain marine organisms present intrinsic feature-based challenges that cannot be resolved through data augmentation alone (Wille, Fischer, and Raine 2026). This finding has profound implications for computer vision more broadly, suggesting that some visual recognition problems may be fundamentally algorithmic rather than data-driven.

Foundation and Language Models in the Underwater Domain: The emergence of foundation models like CoralSCOP represents a paradigm shift from treating coral analysis as a collection of separate tasks (classification, detection, segmentation) to unified models capable of dense segmentation across multiple semantic granularities (Zheng et al. 2024b). CoralSCOP addresses the fundamental challenge that coral structures are inherently amorphous and irregular—characteristics that generic segmentation methods fail to capture. Meanwhile, advances in self-supervised learning have enabled these models to leverage the vast quantities of unlabeled underwater imagery. CoralSRT (Zheng et al. 2025) demonstrated that foundation model features can be rectified through self-supervised guidance to achieve significant performance gains without additional human supervision.

Recent work has seen forays into applying vision–language models and large language models to the underwater domain, though this remains an ongoing area of research. For example, MarineGPT demonstrated the integration of language models into marine applications and highlighted the potential for multimodal understanding, where visual perception can be combined with domain-specific knowledge expressed in natural language (Zheng et al. 2023b). Similarly, natural language prompting with the Contrastive Language-Image Pretraining (CLIP) model (Radford et al. 2021) has been demonstrated for image pseudo-labeling, reducing annotation effort (Raine et al. 2024b). MarineInst (Zheng et al. 2024a) builds on these ideas and is a foundation model for marine imagery which outputs both instance masks and captions for object instances, and UW Bench (Zhang et al. 2025) proposes a comprehensive benchmark for underwater vision-language understanding. These studies highlight the early potential for using foundation models in specialized domains.

3.3 Scene Reconstruction

Underwater scene reconstruction has undergone a fundamental paradigm shift: from traditional photogrammetry towards high resolution 3D scene rendering systems, capable of producing photo-realistic digital twins for precise ecosystem monitoring. This transformation represents a significant technical advance in marine robotics, with implications extending beyond underwater applications.

Gaussian Splatting Influence: The emergence of underwater-specific 3D Gaussian Splatting methods has substantially improved reconstruction capabilities, with systems like WaterSplatting (Li et al. 2025), SeaSplat (Yang, Leonard, and Girdhar 2025), and UW-GS (Wang et al. 2025) achieving real-time rendering while incorporating physics-based underwater image formation models. These approaches explicitly model light attenuation, backscatter, and wavelength-dependent absorption; moving beyond generic computer vision techniques toward domain-aware algorithms that account for underwater optics.

Scale and Speed Transformation: Learning-based Structure from Motion has enabled reconstruction of kilometer-length reef surveys in minutes, a process that previously required hours of manual labor (Sauder et al. 2024). These systems also integrate semantic segmentation directly into the reconstruction pipeline, unifying geometric mapping and scene understanding. Recent advances in multi-view stereo combined with physical imaging models (Hu and Liu 2025) further accelerate processing while maintaining geometric accuracy. This represents a shift from post-hoc analysis to rapid, in-situ ecosystem monitoring capabilities that can support active marine conservation efforts.

Integration with Marine Science Workflows: The coupling of 3D reconstruction with semantic segmentation has progressed from technical demonstrations to operational marine science tools. Multi-modal approaches combining geometric reconstruction with automated species classification (Zhong et al. 2023; Remmers et al. 2025) enable direct computation of ecological metrics, transforming how marine biologists approach large-scale surveys and longitudinal studies. Gorry et al. (2025) have demonstrated that visual place recognition can be integrated with feature matching and image segmentation to enable robust identification of revisited underwater areas for ecosystem change analysis. These integrated systems demonstrate how AI-driven reconstruction can simultaneously provide geometric understanding and biological insights at scales previously impractical with traditional survey methods.

Digital Twin Emergence: Underwater reconstruction has evolved toward ecosystem-scale digital twins with demonstrated deployments spanning multi-million dollar research programs (Woods Hole Oceanographic Institution 2022) and into the tens to hundreds of hectares (King Abdullah University of Science and Technology 2025). These systems enable virtual experimentation with management interventions and provide the foundation for predictive ecosystem modeling at unprecedented scales. These underwater reconstruction systems comprise dynamic, semantically aware models which

support active marine conservation and timely ecosystem management.

This evolution from passive documentation to active environmental intelligence represents a broader trend in AI systems moving from observation to intervention, with underwater applications serving as a proving ground for techniques that will likely find application across numerous challenging perception domains.

3.4 From Research to Impact

The emerging trend of AI for underwater environments has already begun to deliver tangible benefits for fisheries, aquaculture, and reef restoration. What was once primarily a research endeavor has transitioned into a suite of operational tools with real-world impact.

Fisheries: AI-powered monitoring systems are transforming fishery management with greater transparency, efficiency, and accountability. Deep learning applied to underwater imagery and sonar now enables automated fish counting and biomass estimation, supporting sustainable stock assessments at scale (Connolly et al. 2022, 2023). In parallel, AI fused with satellite and Automatic Identification System data is revolutionizing compliance monitoring, with novel approaches detecting illegal “dark” fishing vessels that disable trackers (Raynor et al. 2025).

Aquaculture: Commercial aquaculture has increasingly adopted AI-enabled Internet-of-Things (AIoT) platforms for operational efficiency. Smart feeding systems combine underwater cameras with behavioral analysis to optimize feed delivery, improving growth while reducing waste (Tidal 2025). Additional developments include water quality monitoring, health diagnostics (e.g., disease detection), and real-time biomass estimation (Connolly et al. 2022; Singh, Sahoo, and Nayyar 2022), greatly reducing labor requirements and costs. Recent surveys highlight that AIoT applications now span feeding, water chemistry, fish health, and even individual animal tracking (Ubina et al. 2023; Hu et al. 2023).

Reef Restoration: Robotics and AI are scaling coral propagation and rehabilitation efforts. A notable example is the Coral Growout Robotic Assessment System (CGRAS), which uses robotic imaging to automatically detect and count millions of baby corals in hatchery tanks (QUT 2024). This replaces thousands of hours of manual labor, providing scientists with real-time growth data and enabling large-scale coral deployment across the Great Barrier Reef.

Complementary initiatives, such as LarvalBots (Mou, Tsai, and Dunbabin 2022) for targeted larval dispersal and Reef Guidance Systems (RGS) for site selection, further accelerate reef gardening and restoration (Raine, Moshirian, and Fischer 2025). The impact of restoring reefs can be established using AI with underwater cameras to quantify fish stock recovery (McAnaney et al. 2025; Connolly et al. 2024). The ReefOS system (CG Labs 2025) combines static cameras, fish detection models, a custom photogrammetry toolkit, and a geospatial visualization dashboard to perform fine-scale monitoring of their restoration sites in French Polynesia, Fiji and Thailand.

Bridging AI Outputs to Ecological Metrics: To ensure uptake, AI outputs are increasingly being mapped onto conventional monitoring protocols (e.g., percent cover from point-intercept). Platforms like ReefCloud and CoralNet operationalize random point sampling and class predictions to yield comparable ecological indicators and dashboards (Australian Institute of Marine Science (AIMS) 2024; Beijbom et al. 2012). For 3D products, standardized co-registration and uncertainty budgets allow statistically defensible temporal change estimates that align with regulatory expectations (Lechene et al. 2024).

Generalization, Adaptation, and Uncertainty: Generalization across regions and sensors remains central. Broad, curated image corpora (e.g., FathomNet) and new multi-region benthic datasets (e.g., BenthicNet) provide the diversity needed for pretraining, open-set handling, and rapid adaptation (Katija et al. 2022; Sauder et al. 2025; Lowe et al. 2025). Early works into out-of-distribution detection (Wyatt et al. 2025) are supporting safer deployment of AI algorithms. In parallel, uncertainty-aware evaluation (model confidence, inter-observer variability, and protocol variance) is increasingly reported alongside accuracy, supporting use in regulation and markets (e.g., blue carbon monitoring, reporting and validation; marine protected area compliance).

4 Conclusions

AI for underwater perception illustrates a bidirectional relationship between environmental necessity and technological innovation. Conservation needs have accelerated research, while underwater constraints have driven advances in computer vision and machine learning that benefit many domains. This dual impact demonstrates that domain-specific engagement can drive both fundamental research and tangible impact.

The convergent drivers, i.e. environmental imperative, data democratization, and research opportunity, have created a self-reinforcing ecosystem where new methods enable applications, generate data, and attract investment—transforming underwater AI into a recognized research frontier with real-world impact. Core challenges such as turbidity, cryptic species detection, and annotation bottlenecks have spurred innovations in weak supervision, open-set recognition, and robust perception in degraded environments. These advances extend beyond marine science to applications like medical imaging, analysis of satellite data, and autonomous systems.

The field exemplifies how AI research can progress from academic novelty to operational capability. Multi-million-dollar deployments in reef monitoring and restoration provide a template for translating innovation into measurable societal outcomes. Underwater AI also serves as a proving ground for emerging techniques in foundation models and human–AI collaboration. The requirements for uncertainty quantification, expert knowledge integration, and robust performance under uncertainty make it an ideal domain to mature reliable AI systems for high-stakes environments.

Acknowledgments

S.R. and T.F. acknowledge continued support from the Queensland University of Technology (QUT) through the Centre for Robotics. T.F. acknowledges funding from an Australian Research Council (ARC) Discovery Early Career Researcher Award (DECRA) Fellowship DE240100149. S.R. and T.F. acknowledge support from the Reef Restoration and Adaptation Program (RRAP) which is funded by a partnership between the Australian Government's Reef Trust and the Great Barrier Reef Foundation.

References

- Alonso, I.; Yuval, M.; Eyal, G.; Treibitz, T.; and Murillo, A. C. 2019. CoralSeg: Learning coral segmentation from sparse annotations. *Journal of Field Robotics*, 36(8): 1456–1477.
- Anthony, K. R.; Helmstedt, K. J.; Bay, L. K.; Fidelman, P.; Hussey, K. E.; Lundgren, P.; Mead, D.; McLeod, I. M.; Mumby, P. J.; Newlands, M.; et al. 2020. Interventions to help coral reefs under global change—A complex decision challenge. *PLOS One*, 15(8): e0236399.
- Australian Institute of Marine Science (AIMS). 2024. Reef-Cloud. <https://doi.org/10.25845/g5gk-ty57>. Accessed: 2025-08-27.
- Beijbom, O.; Edmunds, P. J.; Kline, D. I.; Mitchell, B. G.; and Kriegman, D. 2012. Automated annotation of coral reef survey images. In *IEEE Conference on Computer Vision and Pattern Recognition*, 1170–1177.
- CG Labs. 2025. ReefOS: The most advanced AI for the reef. <https://cglabs.coralgardeners.org/>. Accessed: August 2025.
- Chen, W.; Lei, Y.; Luo, S.; Zhou, Z.; Li, M.; and Pun, C.-M. 2024. UWFormer: Underwater image enhancement via a semi-supervised multi-scale transformer. In *IEEE International Joint Conference on Neural Networks*, 1–8.
- Citizens of the Reef. 2025. Great Reef Census: Join a new era of reef conservation. <https://greatreefcensus.org/>. Accessed: August 2025.
- Connolly, R.; Herrera, C.; Rasmussen, J.; Buelow, C.; Sievers, M.; Jinks, K.; Brown, C.; Lopez-Marcano, S.; Sherman, C.; Martínez-Baena, F.; et al. 2024. Estimating enhanced fish production on restored shellfish reefs using automated data collection from underwater videos. *Journal of Applied Ecology*, 61(4): 633–646.
- Connolly, R.; Jinks, K.; Shand, A.; Taylor, M.; Gaston, T.; Becker, A.; and Jinks, E. 2023. Out of the shadows: Automatic fish detection from acoustic cameras. *Aquatic Ecology*, 57(4): 833–844.
- Connolly, R. M.; Jinks, K. I.; Herrera, C.; and Lopez-Marcano, S. 2022. Fish surveys on the move: Adapting automated fish detection and classification frameworks for videos on a remotely operated vehicle in shallow marine waters. *Frontiers in Marine Science*, 9: 918504.
- Contini, M.; Illien, V.; Julien, M.; Ravitchandirane, M.; Rusias, V.; Lazennec, A.; Chevrier, T.; Rintz, C. L.; Carpentier, L.; Gogondeau, P.; et al. 2025a. Seatizen Atlas: A collaborative dataset of underwater and aerial marine imagery. *Scientific Data*, 12(1): 67.
- Contini, M.; Illien, V.; Poulain, S.; Bernard, S.; Barde, J.; Bonhommeau, S.; and Joly, A. 2025b. The point is the mask: Scaling coral reef segmentation with weak supervision. In *IEEE/CVF International Conference on Computer Vision*, 2056–2065.
- Data Mermaid. 2025. The Data Aid your reef conservation deserves. <https://datamermaid.org/>. Accessed: August 2025.
- Dietzel, A.; Bode, M.; Connolly, S. R.; and Hughes, T. P. 2021. The population sizes and global extinction risk of reef-building coral species at biogeographic scales. *Nature Ecology & Evolution*, 5(5): 663–669.
- Ditria, E. M.; Buelow, C. A.; Gonzalez-Rivero, M.; and Connolly, R. M. 2022. Artificial intelligence and automated monitoring for assisting conservation of marine ecosystems: A perspective. *Frontiers in Marine Science*, 9: 918104.
- FathomVerse. 2025. FathomVerse. <https://www.fathomverse.game/>. Accessed: August 2025.
- Gonzalez-Rivero, M.; Beijbom, O.; Rodriguez-Ramirez, A.; Bryant, D. E.; Ganase, A.; Gonzalez-Marrero, Y.; Herrera-Reveles, A.; Kennedy, E. V.; Kim, C. J.; Lopez-Marcano, S.; et al. 2020. Monitoring of coral reefs using artificial intelligence: A feasible and cost-effective approach. *Remote Sensing*, 12(3): 489.
- González-Rivero, M.; Bongaerts, P.; Beijbom, O.; Pizarro, O.; Friedman, A.; Rodriguez-Ramirez, A.; Upcroft, B.; Laffoley, D.; Kline, D.; Bailhache, C.; et al. 2014. The Catlin Seaview Survey—kilometre-scale seascape assessment, and monitoring of coral reef ecosystems. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 24(S2): 184–198.
- González-Rivero, M.; Rodriguez-Ramirez, A.; Beijbom, O.; Dalton, P.; Kennedy, E. V.; Neal, B. P.; Vercelloni, J.; Bongaerts, P.; Ganase, A.; Bryant, D. E.; et al. 2019. Seaview survey photo-quadrat and image classification dataset.
- Gorry, B.; Fischer, T.; Milford, M.; and Fontan, A. 2025. Image-based relocalization and alignment for long-term monitoring of dynamic underwater environments. In *IEEE/RSJ Conference on Intelligent Robots and Systems*.
- Hanlon, R. 2007. Cephalopod dynamic camouflage. *Current biology*, 17(11): R400–R404.
- Henley, B. J.; McGregor, H. V.; King, A. D.; Hoegh-Guldberg, O.; Arzey, A. K.; Karoly, D. J.; Lough, J. M.; DeCarlo, T. M.; and Linsley, B. K. 2024. Highest ocean heat in four centuries places Great Barrier Reef in danger. *Nature*, 632(8024): 320–326.
- Hu, S.; and Liu, Q. 2025. Fast underwater scene reconstruction using multi-view stereo and physical imaging. *Neural Networks*, 107568.
- Hu, W.-C.; Chen, L.-B.; Wang, B.-H.; Li, G.-W.; and Huang, X.-R. 2023. Design and implementation of a full-time artificial intelligence of things-based water quality inspection and prediction system for intelligent aquaculture. *IEEE Sensors Journal*, 24(3): 3811–3821.
- Katija, K.; Orenstein, E.; Schlining, B.; Lundsten, L.; Barnard, K.; Sainz, G.; Boulais, O.; Cromwell, M.; Butler, E.; Woodward, B.; et al. 2022. FathomNet: A global image database for enabling artificial intelligence in the ocean. *Scientific Reports*, 12(1): 15914.

- King Abdullah University of Science and Technology. 2025. Adoption of AI to accelerate world's largest coral restoration project. <https://www.kaust.edu.sa/en/news/adoption-of-ai-to-accelerate-world's-largest-coral-restoration-project/>. Accessed: August 2025.
- Kohler, K. E.; and Gill, S. M. 2006. Coral Point Count with Excel extensions (CPCe): A Visual Basic program for the determination of coral and substrate coverage using random point count methodology. *Computers & Geosciences*, 32(9): 1259–1269.
- Lavery, P. S.; Mateo, M.-Á.; Serrano, O.; and Rozaimi, M. 2013. Variability in the carbon storage of seagrass habitats and its implications for global estimates of blue carbon ecosystem service. *PLOS One*, 8(9): e73748.
- Lawson, C. L.; Chartrand, K. M.; Roelfsema, C. M.; Kolluru, A.; and Mumby, P. J. 2025. Broad-scale reconnaissance of coral reefs from citizen science and deep learning. *Environmental Monitoring and Assessment*, 197(7): 1–21.
- Lechene, M. A.; Figueira, W. F.; Murray, N. J.; Aston, E. A.; Gordon, S. E.; and Ferrari, R. 2024. Evaluating error sources to improve precision in the co-registration of underwater 3D models. *Ecological Informatics*, 81: 102632.
- Li, H.; Song, W.; Xu, T.; Elsig, A.; and Kulhanek, J. 2025. WaterSplatting: Fast underwater 3D scene reconstruction using Gaussian splatting. In *International Conference on 3D Vision*.
- Li, Z.; Chen, Q.; and Miao, J. 2025. SPMFormer: Simplified Physical Model-based transformer with cross-space loss for underwater image enhancement. *Knowledge-Based Systems*, 113694.
- Lowe, S. C.; Misiuk, B.; Xu, I.; Abdulazizov, S.; Baroi, A. R.; Bastos, A. C.; Best, M.; Ferrini, V.; Friedman, A.; Hart, D.; et al. 2025. BenthicNet: A global compilation of seafloor images for deep learning applications. *Scientific Data*, 12(1): 230.
- McAneney, C. A.; Sievers, M.; Reeves, S. E.; Rasmussen, J. A.; Leaning, A.; Lopez-Marcano, S.; and Connolly, R. M. 2025. Fish use of restored shellfish reefs extends beyond the reef edge. *Restoration Ecology*, e70064.
- McKenzie, L. J.; Nordlund, L. M.; Jones, B. L.; Cullen-Unsworth, L. C.; Roelfsema, C.; and Unsworth, R. K. 2020. The global distribution of seagrass meadows. *Environmental Research Letters*, 15(7): 074041.
- Mou, S.; Tsai, D.; and Dunbabin, M. 2022. Reconfigurable robots for scaling reef restoration. *arXiv preprint arXiv:2205.04612*.
- Murphy, H. M.; and Jenkins, G. P. 2010. Observational methods used in marine spatial monitoring of fishes and associated habitats: A review. *Marine and Freshwater Research*, 61(2): 236–252.
- NASA. 2025. Nemo-Net. <https://nemonet.info/>. Accessed: August 2025.
- NOAA. 2025. NOAA Digital Photo Collection. <https://www.noaa.gov/digital-collections/photo-library>. Accessed: August 2025.
- Nozdrenkov, S.; Mujiyanto; Zedta, R. R.; Fakhurrozi; Barker, T.; Rahman, M. D.; Samusamu, A. S.; Rochman, F.; Peters, L.; Rachmawati, D. R.; Johan, D. O.; Craggs, D. J.; and Sweet, P. M. 2025. sweet-corals (Revision bd316dc).
- Pierce, J. 2025. Click-a-Coral. <https://www.zooniverse.org/projects/jordan-pierce/click-a-coral>. Accessed: August 2025.
- Qiao, J. H.; Guan, Y. T.; and Zhi, J. N. 2025. Transformer-CNN hybrid network for underwater image enhancement. *The Journal of Supercomputing*, 81(10): 1–17.
- QUT. 2024. Robot designed to boost coral regeneration on the Great Barrier Reef. <https://www.qut.edu.au/about/faculty-of-science/news-and-events?id=194074>. Accessed: August 2025.
- Radford, A.; Kim, J. W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; et al. 2021. Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning*, 8748–8763. PmLR.
- Raine, S.; Maire, F.; Suenderhauf, N.; and Fischer, T. 2024a. Reducing label dependency for underwater scene understanding: A survey of datasets, techniques and applications. *arXiv preprint arXiv:2411.11287*.
- Raine, S.; Marchant, R.; Kusy, B.; Maire, F.; and Fischer, T. 2022. Point label aware superpixels for multi-species segmentation of underwater imagery. *IEEE Robotics and Automation Letters*, 7(3): 8291–8298.
- Raine, S.; Marchant, R.; Kusy, B.; Maire, F.; and Fischer, T. 2024b. Image labels are all you need for coarse seagrass segmentation. In *IEEE/CVF Winter Conference on Applications of Computer Vision*, 5943–5952.
- Raine, S.; Marchant, R.; Kusy, B.; Maire, F.; Suenderhauf, N.; and Fischer, T. 2024c. Human-in-the-loop segmentation of multi-species coral imagery. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2723–2732.
- Raine, S.; Marchant, R.; Moghadam, P.; Maire, F.; Kettle, B.; and Kusy, B. 2020. Multi-species seagrass detection and classification from underwater images. In *IEEE Digital Image Computing: Techniques and Applications*.
- Raine, S.; Moshirian, B.; and Fischer, T. 2025. AI-driven dispensing of coral reseeding devices for broad-scale restoration of the Great Barrier Reef. *arXiv preprint*.
- Raynor, J.; Orofino, S.; Costello, C.; McDonald, G.; Mayorga, J.; and Sala, E. 2025. Little-to-no industrial fishing occurs in fully and highly protected marine areas. *Science*, 389(6758): 392–395.
- ReefCloud. 2025. Protecting the Future of Coral Reefs. <https://reefcloud.ai/>. Accessed: August 2025.
- Remmers, T.; Boutros, N.; Wyatt, M.; Gordon, S.; Toor, M.; Roelfsema, C.; Fabricius, K.; Grech, A.; Lechene, M.; and Ferrari, R. 2025. RapidBenthos: Automated segmentation and multi-view classification of coral reef communities from photogrammetric reconstruction. *Methods in Ecology and Evolution*, 16(2): 427–441.
- Sauder, J.; Banc-Prandi, G.; Meibom, A.; and Tuia, D. 2024. Scalable semantic 3D mapping of coral reefs with deep learning. *Methods in Ecology and Evolution*, 15(5): 916–934.

- Sauder, J.; Domazetoski, V.; Banc-Prandi, G.; Perna, G.; Meibom, A.; and Tuia, D. 2025. The Coralscapes dataset: Semantic scene understanding in coral reefs. *arXiv preprint arXiv:2503.20000*.
- Siméoni, O.; Vo, H. V.; Seitzer, M.; Baldassarre, F.; Oquab, M.; Jose, C.; Khalidov, V.; Szafraniec, M.; Yi, S.; Ramamonjisoa, M.; et al. 2025. DINOv3. *arXiv preprint arXiv:2508.10104*.
- Singh, M.; Sahoo, K. S.; and Nayyar, A. 2022. Sustainable IoT solution for freshwater aquaculture management. *IEEE Sensors Journal*, 22(16): 16563–16572.
- Souter, D.; Planes, S.; Wicquart, J.; Logan, M.; Obura, D.; and Staub, F., eds. 2021. *Status of Coral Reefs of the World: 2020 Report*. Global Coral Reef Monitoring Network (GCRMN) and International Coral Reef Initiative (ICRI). DOI: 10.59387/WOTJ9184.
- Squidle. 2025. IMOS-UMI Squidle+ A tool for managing, exploring & annotating images, video & large-scale mosaics. <https://squidle.org/>. Accessed: August 2025.
- Tidal. 2025. AI for the next generation of sustainable aquaculture. <https://www.tidalx.ai/en>. Accessed: August 2025.
- Ubina, N. A.; Lan, H.-Y.; Cheng, S.-C.; Chang, C.-C.; Lin, S.-S.; Zhang, K.-X.; Lu, H.-Y.; Cheng, C.-Y.; and Hsieh, Y.-Z. 2023. Digital twin-based intelligent fish farming with Artificial Intelligence Internet of Things (AIoT). *Smart Agricultural Technology*, 5: 100285.
- Van den Bergh, J.; Chirayath, V.; Li, A.; Torres-Pérez, J. L.; and Segal-Rozenhaimer, M. 2021. NeMO-net—gamifying 3D labeling of multi-modal reference datasets to support automated marine habitat mapping. *Frontiers in Marine Science*, 8: 645408.
- Vercelloni, J.; Santos-Fernández, E.; and Mengersen, K. 2023. Virtual Reef Diver: Enabling People to Help Protect the Great Barrier Reef. *Citizen Science: Theory and Practice*, 8(1).
- Virtual Reef Diver. 2025. Virtual Reef Diver: Become a Citizen Scientist. <https://www.virtualreef.org.au/>. Accessed: August 2025.
- Wang, B.; Xu, H.; Jiang, G.; Yu, M.; Chen, Y.; Ding, L.; Zhang, X.; and Luo, T. 2023. Underwater image co-enhancement based on physical-guided transformer interaction. *Displays*, 79: 102505.
- Wang, H.; Anantrasirichai, N.; Zhang, F.; and Bull, D. 2025. UW-GS: Distractor-aware 3D Gaussian splatting for enhanced underwater scene reconstruction. In *IEEE/CVF Winter Conference on Applications of Computer Vision*, 3280–3289.
- Wille, M.; Fischer, T.; and Raine, S. 2026. Are all marine species created equal? Performance disparities in underwater object detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*.
- Woods Hole Oceanographic Institution. 2022. Digital Reefs awarded \$5 million. <https://www.whoi.edu/press-room/news-release/digital-reefs-awarded-5-million/>. Accessed: August 2025.
- Wyatt, M.; Hickey, S.; Radford, B.; Gonzalez-Rivero, M.; Boutros, N.; Callow, N.; Ryan, N.; Chennu, A.; Bennamoun, M.; and Gilmour, J. 2025. Safe AI for coral reefs: Benchmarking out-of-distribution detection algorithms for coral reef image surveys. *Ecological Informatics*, 103207.
- Yang, D.; Leonard, J. J.; and Girdhar, Y. 2025. SeaSplat: Representing underwater scenes with 3D Gaussian splatting and a physically grounded image formation model. In *IEEE International Conference on Robotics and Automation*.
- Zhang, D.; Rong, C.; Li, B.; Wang, F.; Zhao, Z.; Gao, J.; and Li, X. 2025. UW-Bench: A Comprehensive Vision-Language Benchmark for Underwater Understanding. *arXiv preprint arXiv:2510.18262*.
- Zheng, Z.; Chen, Y.; Zeng, H.; Vu, T.-A.; Hua, B.-S.; and Yeung, S.-K. 2024a. MarineInst: A foundation model for marine image analysis with instance visual description. *European Conference on Computer Vision*.
- Zheng, Z.; Liang, H.; Hua, B.-S.; Wong, Y. H.; Ang, P.; Chui, A. P. Y.; and Yeung, S.-K. 2024b. CoralSCOP: Segment any coral image on this planet. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 28170–28180.
- Zheng, Z.; Wong, K.; Binh-Son, H.; Jianbo, S.; and Sai-Kit, Y. 2025. CoralSRT: Revisiting coral reef semantic segmentation by feature rectifying via self-supervised guidance. In *Proceedings of the IEEE International Conference on Computer Vision*.
- Zheng, Z.; Xie, Y.; Liang, H.; Yu, Z.; and Yeung, S.-K. 2023a. CoralVOS: Dataset and benchmark for coral video segmentation. *arXiv preprint arXiv:2310.01946*.
- Zheng, Z.; Zhang, J.; Vu, T.-A.; Diao, S.; Tim, Y. H. W.; and Yeung, S.-K. 2023b. MarineGPT: Unlocking secrets of ocean to the public. *arXiv preprint arXiv:2310.13596*.
- Zhong, J.; Li, M.; Zhang, H.; and Qin, J. 2023. Combining photogrammetric computer vision and semantic segmentation for fine-grained understanding of coral reef growth under climate change. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 186–195.