Underwater and airborne monitoring of marine ecosystems and debris

Jun-ichiro Watanabe
Yang Shao
Naoto Miura

Underwater and airborne monitoring of marine ecosystems and debris

Jun-ichiro Watanabe,* **Yang Shao, and Naoto Miura**
Hitachi, Ltd., Research and Development Group, Tokyo, Japan

Abstract. Advancing the sustainable use and conservation of marine environments is urgent. Tons of debris including macro- and microplastics generated on land are entering the oceans, marine resources are decreasing, and many species are facing extinction. Though satellite remote sensing techniques are commonly used for global environmental monitoring, it is still difficult to detect small objects such as floating debris on the vast ocean surface, and the ecosystems deep in the oceans where light does not reach are unobservable. An autonomous monitoring system consisting of optimally controlled robots is required for acquiring spatiotemporally rich marine data. However, object detection in marine environments, which is a necessary function the robots should have for underwater and aerial monitoring, has not been extensively studied. Here, we argue that state-of-the-art deep-learning-based object detection works well for monitoring underwater ecosystems and marine debris. We found that by using the deep-learning object-detection algorithm YOLO v3, underwater sea life and debris floating on the ocean surface can be detected with mean average precision of 69.6% and 77.2%, respectively. We anticipate our results to be a starting point for developing tools for enabling safe and precise acquisition of marine data to elucidate and utilize this last frontier. © *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.JRS.13.044509]

1 Introduction

In 2015, the United Nations set 17 goals, known as sustainable development goals (SDGs),[1] to achieve a better and sustainable future for all life on this planet. The goals include roughly two major challenges: one addressing poverty, education, health, energy, and economics, which are directly connected to our society, and the other concerning climate change, conservation, and use of forests and oceans, which affect our lives indirectly through changes in the natural environment that are difficult to control. For the first challenge, we can take measures to improve our society, such as developing new laws or technologies. For the second challenge, however, though we can take any measures to increase resilience against environmental changes and conserve nature, it will take several decades to evaluate their effectiveness. Therefore, continuous and precise environmental monitoring techniques are required.

Satellite remote sensing is commonly used for global environmental monitoring. We can observe wide areas and collect information regarding the atmosphere, land, and ocean surface. The current constellation of satellites has enabled to observe the Earth improving temporal resolution. Furthermore, satellites, such as WorldView-3 and 4, can take images with spatial resolutions on the submeter order enabling the reconstruction of the land surface in three-dimensional. However, spatiotemporal resolution is not precise enough for local observation compared to the amount of information gathered with ordinary surveillance cameras used in cities. Moreover, the satellite imagery products are still expensive, and it is difficult to obtain information from underwater and in dense forests where light cannot reach.

Recently, unmanned aerial vehicles (UAVs) have been used to take high-resolution images for land observation. Studies have focused mainly on using drones for agricultural application[2]

*Address all correspondence to Jun-ichiro Watanabe, E-mail: junichiro.watanabe.ws@hitachi.com

and facility inspections.[3] For underwater observation, autonomous underwater vehicles (AUVs) and remotely operated vehicles (ROVs) have been developed and used for seabed resource exploration,[4] bathymetry analysis,[5] and benthos survey.[6]

Advancing the sustainable use and conservation of marine environments is urgent. Tons of debris, including macro- and microplastics from land, are entering the oceans, marine resources are decreasing, and many species are facing extinction. The oceans have great potential for many industries such as fisheries, resources, logistics, and tourism. However, ocean utilization has been delayed due to inaccessibility and lack of sensing technologies. As UNESCO states in the "Decade of Ocean Science" campaign in 2017, "The ocean covers 71 percent of the globe, but we have explored less than 5 percent," thus, a sufficient amount of marine information has not been collected.

The current study explored marine-monitoring techniques using UAVs, AUVs, and other autonomous robots for contributing to marine-ecosystem conservation and addressing the marine-debris problem stated in SDG 14. Currently, marine-ecosystem surveys are mainly executed by divers and through land-based visual inspections. Marine-debris monitoring also depends on observation from ships or the coast. We report on a preliminary evaluation for applying a state-of-the-art deep-learning algorithm for object detection in marine environments, which include more complex and variable backgrounds compared to the land. We evaluated the performance of the latest deep-learning object-detection algorithm called you only look once (YOLO) v3[7] for identifying underwater sea life and detecting debris on the ocean surface and beaches. We then discuss the possibilities of applying this algorithm to develop a marine-monitoring system involving autonomous robots scattered throughout marine environments.

2 Related Work

Satellite remote sensing techniques are commonly used to obtain marine data. We can observe sea-surface temperature, ocean current, waves, salinity, eutrophication, and/or chlorophyll-a distribution for long periods over wide areas.[8,9] Satellite images have been applied in various fields, including weather forecasting, fishery prediction,[10] seagrass-bed analysis,[11] and even counting whales.[12] However, there is a limitation of spatiotemporal resolution, though satellites capable of taking submeter-order spatial resolution images and their constellation have improved this. It has been pointed out that since coastal ecosystems have high spatial complexity and temporal variability, they frequently have to be observed from both satellite and aircraft to obtain the required spatial, spectral, and temporal resolutions.[13]

To observe under water, where satellites cannot observe, systems, such as those involving buoys with sensors and cameras and underwater vehicles, have been proposed.[14–18] Argo Float[19] is one such system consisting of many buoys scattered throughout the world's oceans to collect salinity and water temperature moving from sea surface to a depth of around 1000 m. Compared to a buoys-based system, autonomous vehicles must provide spatially unrestricted ocean exploration. Actually, Thompson pointed out that future marine monitoring systems would rely heavily on autonomous vehicles to enable persistent and heterogeneous measurements needed to understand the ocean's impact on the climate system.[20] The recently held "Ocean Discovery" is a worldwide competition involving challenges to advance technologies for autonomous ocean exploration.[21] Our study is in line with these challenges to realize autonomous ocean monitoring system.

Plastic debris in marine environments has been widely documented and regarded as a serious problem that affects marine ecosystems and even humanity.[22–27] Jambeck et al.[28] estimated the mass of land-based plastic waste entering the ocean by linking worldwide data on solid waste, population density, and economic status. The Ocean Cleanup[29] developed a system consisting of a 600-m-long floater that sits on the surface of the water and a tapered 3-m-deep skirt attached below to rid the ocean of plastic debris. They also reported that there are areas in the Pacific Ocean that accumulate debris rapidly based on multivessel and aircraft surveys and simulation.[30] However, it is still difficult to estimate debris distribution around the vast ocean and efficiently clean up the area.

Deep neural networks are being successfully used for object recognition in images. A convolutional neural network (CNN)[31] is one such deep network consisting of several convolutional,

pooling, and fully connected layers. Recently developed deep learning algorithms for detecting target regions and categories simultaneously, such as Faster R-CNN,[32] YOLO,[33] and SSD,[34] have shown higher detection accuracy and faster processing time. Previous studies have applied Faster R-CNN or YOLO for detecting fish from underwater images.[35–38] We believe that to realize an autonomous ocean monitoring system, it is very important to evaluate the feasibility of a state-of-the-art algorithm, YOLO v3,[7] for marine ecosystems and debris monitoring.

3 Methods

In this study, we conducted a preliminary study for developing a marine-monitoring system consisting of autonomous robots scattered throughout an environment. We evaluated the possibility of using UAVs and AUVs for such a system.

3.1 Data Acquisition

We first evaluated the visibility of small floating objects from aerial images. To do this, we flew a drone (DJI Matrice 210 RTK) over Ishikari river estuary, Hokkaido, Japan, in September 2018. The drone was equipped with a visible light camera (ZENMUSE X5S) and infrared (IR) camera (ZENMUSE XT). We took images from different altitudes, i.e., 5, 10, and 30 m, targeting a toy fish made of poly vinyl chloride cast into the water by using a fishing rod (Fig. 1). It was cloudy and the wind speed was around 10 m/s that day.

Next, we implemented an object-detection algorithm for underwater ecosystem monitoring and finding marine debris. We collected images for model training and performance evaluation. We took underwater images by scuba diving at Kawana beach in Shizuoka, Japan, in August, 2018. It was cloudy, the average diving depth was around 12 m, visibility in the water was around 6 m, and the camera used was Go Pro Hero 6. The images of marine debris were taken manually at beaches located in the Kanto region (Kanagawa and Chiba prefectures) in Japan in January 2019 using three cameras (Go Pro Hero 6, OLYMPUS Tough TG-5, and iPhone). We also downloaded images including underwater sea life, such as fish, sea turtles, and jelly fish, from Google Open Images[39] to cover the shortage in the variety of images.

3.2 Deep Network

We designed two deep-network models for detecting objects from ocean images: one to detect underwater sea life and the other to detect debris floating on the ocean surface and drifting ashore. The sea-life detection model was designed to identify three types of sea life; fishes, sea
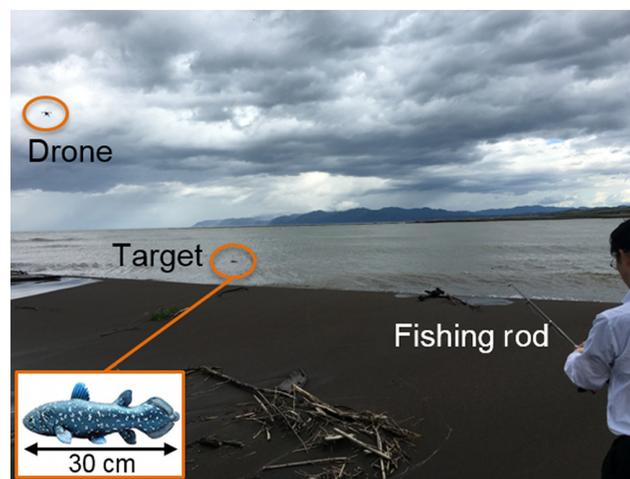
Fig. 1 Field experiment of aerial marine monitoring. Drone and drifting target poly vinyl chloride toy fish.

turtles, and jelly fish. We prepared 8036 images including these targets sourced from Open Images dataset and used 6908 (86%) of them for training the model and 1128 (14%) for evaluation. The debris-detection model was designed to find plastic bottles, plastic bags, drift wood, and other debris. Because it was difficult to obtain images of debris from the Open Images dataset, we used 189 images including the target debris that we took and used 152 (80%) of them for training the model and 37 (20%) for evaluation. Though more than 10,000 images are generally required for developing a robust object detector, we know that it is possible to preliminarily evaluate whether the model will work for the given task.

The computing specifications were as follows: Intel® Core™ i7-7800X CPU 350 GHz, 64 bits, 40-GB RAM, GPU Nvidia GTX 1080, CUDA 9.0, cuDNN 7.0.3, and OS ubuntu 16.04. The programs were implemented using Python 3.5 and OpenCV 3.4. We used YOLO v3 object detection algorithm,[7] which uses a deep network with 53 convolutional layers and easily implemented using a deep-learning framework called Darknet.[40] With YOLO v3, the input images were resized to 608×608 pixels for processing.

3.3 Performance Evaluation

The performance measure commonly used for the object-category-segmentation problem is called intersection-over-union (IoU). The IoU gives the similarity between the predicted and ground-truth regions and is defined as the size of the intersection divided by the union of the two regions. We used 0.5 and 0.75 as the IoU thresholds to examine mean average precision (mAP) for evaluating the model performance.

4 Results

4.1 Visibility of Floating Objects from Air

We let the drone take off from the tip of a spit formed at the Ishikari river estuary and flew it over the area within a few hundred meters from the coast except the no-fly zone designated as a seaside plant protection district. We cast a 30-cm-long toy fish into the water by using a fishing rod to be the target for image shooting (Fig. 1).

The visible light camera had a COMS 4/3 image sensor and 25-mm lens, and its resolution was 5280×3956 pixels (4:3). Table 1 shows the relationship between distance to the target and necessary pixels for representing a 30-cm-long object. According to this estimation, the target may successfully be detected from 5- and 10-m observation, but there may be some cases in which we might fail to detect it from 30 m according to marine conditions such as sun glint, waves, and tide.

Figure 2 shows images taken from different altitudes using the visible light camera. As we estimated, we could find the target visually in the images taken from 5 and 10 m [Figs. 2(a) and 2(b)]. We managed to find it in the image from 30 m [Fig. 2(c)]. However, the target in that image was very small, and it may be difficult to find the same target under different conditions, e.g.,

Table 1 Distances to target and resolutions.

	Distance to target (m)	View area (m)	Resolution (mm/pixel)	Pixels for 30-cm target
Horizontal	5	3.5	0.66	455
Vertical		2.6	0.66	
Horizontal	10	6.9	1.31	229
Vertical		5.2	1.31	
Horizontal	30	20.8	3.93	76
Vertical		15.6	3.94	

(a) 5 m　　　　　　　　(b) 10 m　　　　　　　　(c) 30 m

Fig. 2 Target observed from different altitudes: (a) 5, (b) 10, and (c) 30 m. Orange lines indicate target.

stormy weather, strong sun glint, or different color combinations between background water and target.

4.2 *Object Detection in Marine Environments*

To train the YOLO v3 object detection model, data gathered should be annotated to provide information of the target region and label. Since the annotations were automatically created with Open Images dataset, we just converted them to the YOLO format to train the sea-life detection model. On the contrary, we had to annotate debris images manually to train the debris-detection model.

Table 2 shows the model performances of object detection in marine environments. The developed models achieved mAP values of 69.6% and 77.2% for sea-life and debris detection, respectively, if we set IoU threshold as 0.5 as commonly used. These are sufficiently high compared to the one reported in the original YOLO v3 paper (57.9% for YOLO v3-608 applied to COCO dataset[41]). We found that the mAP values decreased when we set stricter IoU as 0.75.

Figure 3 shows example results of underwater fish detection for video input we took by scuba diving. Schools of mid-sized fishes [Fig. 3(a)], small blue ones clearly contrasted with background rocks [Fig. 3(b)], and swimming in low transparency water [Fig. 3(c)] were successfully detected (denoted with pink bounding boxes). Surprisingly, we found that our sea-life detection model could identify small fishes, e.g., the one at the left far end in Fig. 3(a) and ones above the rock in Fig. 3(c), which humans may fail to detect visually.

Figure 4 shows example results of marine-debris detection for video input we took manually at the beaches. We found that our debris-detection model could successfully detect plastic bottles (pink bounding boxes) and other debris such as plastic trays (light blue bounding boxes) floating on the water [Fig. 4(a)]. Furthermore, not only plastic bottles but also a plastic bag (orange bounding box) and drift wood (light green bounding boxes) were successfully detected on the beach [Figs. 4(b) and 4(c)].

We confirmed that when inputting video images, the object detection process runs in real time. Thus, we can conclude that the YOLO v3-based object detector works well for both underwater and beach environments.

5 Discussion

We explored possibilities of developing a marine-monitoring system consisting of autonomous robots, including UAVs and AUVs, for object detection.

Table 2 Performances of sea-life and marine debris detection.

		mAP	
		Sea-life detection	Debris detection
IoU	0.50	69.6%	77.2%
	0.75	50.8%	62.6%

Input images   Detection results
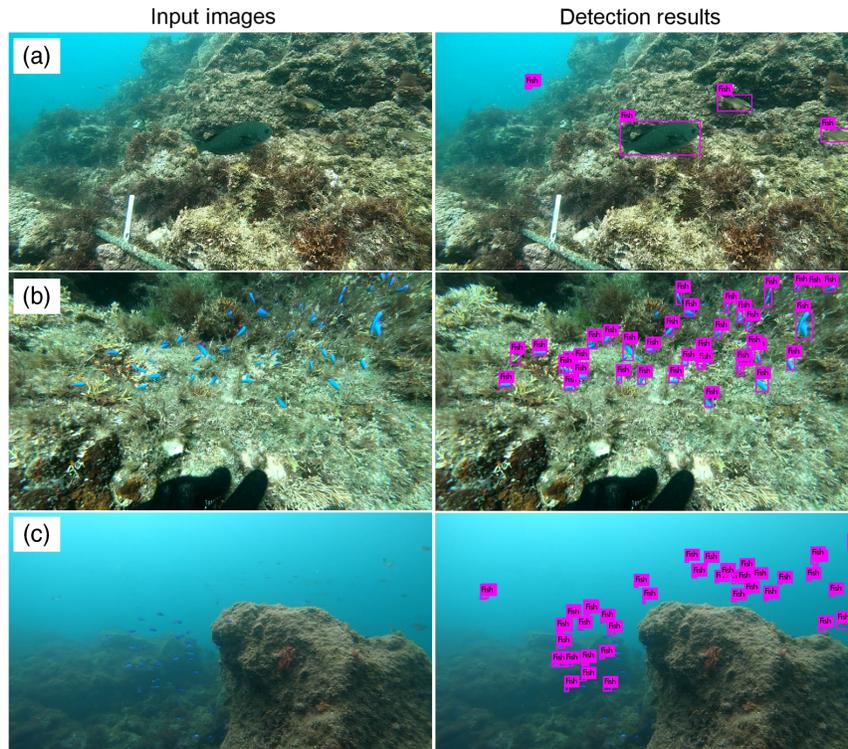
Fig. 3 Example detection results of schools of (a) mid-sized fishes, (b) small blue fishes clearly contrasted with background, and (c) swimming in low transparency water. Video taken by scuba diving was input.

Input images   Detection results
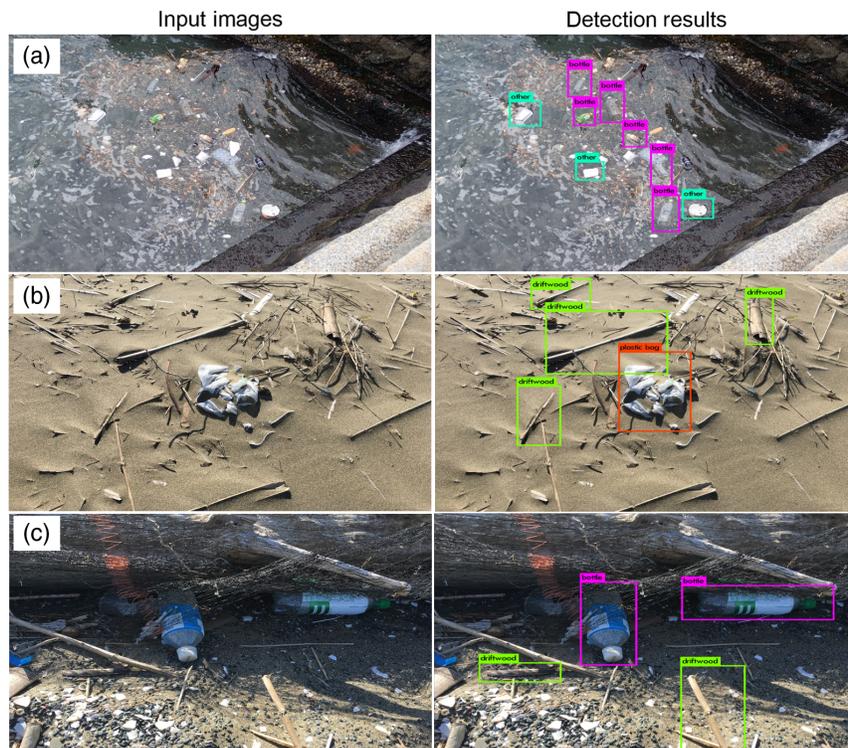
Fig. 4 Example detection results of debris. (a) Floating on ocean surface and (b), (c) drifting ashore. Plastic bottles, plastic bag, drift wood, and other debris are denoted in pink, orange, light green, and light blue bounding boxes, respectively. Video taken manually was input.

(a)   (b)
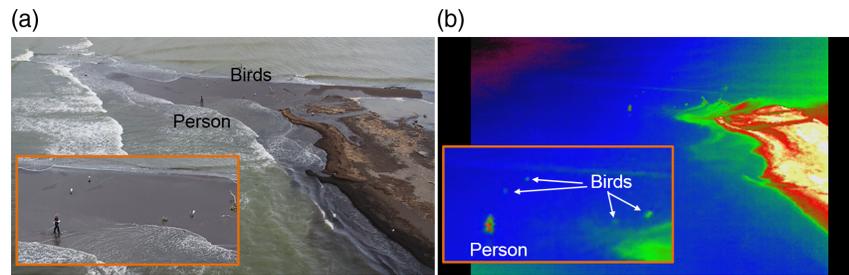
Fig. 5 150-m bird's-eye view. Person and four birds on edge of water observed using (a) visible light camera and (b) IR camera. Insets are enlarged views.

We found that aerial marine monitoring using UAVs that fly at lower altitudes, i.e., less than 30 m, provides sufficient image quality for capturing small targets. In our experiment, the visual light camera on a drone successfully captured images of a 30-cm-long toy fish floating in the water. From higher altitudes, it may be difficult to identify objects due to various water-surface conditions or color and size of targets. Figure 5 shows images taken using both visible light and IR cameras from a higher altitude, i.e., 150 m. In the image taken with the visible light camera [Fig. 5(a)], we can identify something black appearing to be a person walking on the edge of the water. In the enlarged view, we can find four birds as well, though it is difficult to recognize them in the original image. In the IR image [Fig. 5(b)], however, the person and birds are represented by bright colors different from the background. This indicates that not only visible light cameras but also IR cameras are useful for aerial marine monitoring. We believe that multi- and hyperspectral cameras can also be applied for various marine-monitoring tasks.

As some previous studies have applied Faster R-CNN for underwater fish detection (e.g., Refs. 35 and 36), we implemented Faster R-CNN for sea-life and debris detection to compare performance with that of YOLO v3 for our datasets. We found that the YOLO v3 was outstanding in terms of accuracy and processing speed. Our Faster R-CNN models achieved only 40.0% and 41.2% of mAP for sea-life and debris detection, respectively, when IoU was set to 0.5. The values were 28.8% and 21.0 when IoU was 0.75 showing much lower performances compare to those of YOLO v3 (see Table 2). We found that Faster R-CNN was not good at detecting small targets [e.g., schools of fishes shown in Fig. 3(b)] that YOLO v3 could successfully detect.

To develop an environmental monitoring system consisting of autonomous robots scattered throughout a natural environment, such robots should have an object-recognition function to determine their next action depending on the detection results. When we ran YOLO v3 on a single board computer designed for developing such robots, Raspberry pi 3 Model B+ (CPU Broadcom BCM2837B0, Cortex-A53(ARMv8), and 1-GB memory), it took a very long processing time, which was not realistic. This is because the processing ability of the board is too weak for a very deep network with 53 convolution layers. Therefore, we applied tiny models with shallower network structure, e.g., tiny-YOLO,[42] to run on the board with a visual processing unit (VPU) called Intel® Movidius™ Neural Compute Stick. In our preliminary experiment, we could detect objects almost in real-time using this architecture. We plan to evaluate the performance quantitatively for different hardware settings (multi VPUs, for example) and various deep network structures.

We believe that using the swarm control technique of drones and underwater vehicles (e.g., Refs. 43 and 44) with various types of cameras and combining satellite observation will enable the development of a high-spatiotemporal marine-monitoring system. For example, sending multiple drones to a specific area, where a satellite discovered abnormal trends, to collect precise local data will provide us with richer marine information that can be used not only for disaster prevention but also for industrial fields such as smart fisheries. To realize such marine monitoring system, we need to develop an algorithm to optimally control multiple robots based on the object-detection results. We believe reinforcement learning is powerful under various natural environments, where robots have to move under incomplete observable conditions, to run the robots with cooperation to execute a task.

6 Conclusion

We explored methods for monitoring marine environments by using autonomous robots including commercial UAVs and AUVs. We found that by using the deep-learning object-detection algorithm YOLO v3, underwater sea life and debris floating on the ocean surface can be detected with mAP of 69.6% and 77.2%, respectively. Our results indicate the fundamental feasibility of detecting underwater sea life and small objects, such as debris, using the state-of-the-art deep-learning-based object detection algorithm that could be implemented on the robots scattered in the nature to autonomously monitor the environment. For the next step, we plan to increase the number of classes that the model can classify to observe various ecosystems and types of marine debris. We also plan to develop a prototype robot with the object detection function controlled optimally by reinforcement algorithm for the environmental monitoring. We believe that this study will facilitate the development of a marine-monitoring system consisting of scattered autonomous robots with deep-network-based object detector.

Acknowledgments

References

1. United Nations, "SDGs: sustainable development knowledge platform," https://sustainabledevelopment.un.org/sdgs.
2. E. Honkavaara et al., "Processing and assessment of spectrometric, stereoscopic imagery collected using a lightweight UAV spectral camera for precision agriculture," *Remote Sens.* **5**(10), 5006–5039 (2013).
3. J. Watanabe et al., "Power line-tree conflict detection and 3D mapping using aerial images taken from UAV," *Proc. SPIE* **10643**, 106430U (2018).
4. E. Galceran and M Carreras, "Efficient seabed coverage path planning for ASVs and AUVs," in *IEEE/RSJ Int. Conf. Intell. Robots Syst.*, pp. 88–93 (2012).
5. S. B. Williams et al., "AUV benthic habitat mapping in south eastern Tasmania," in *Field and Service Robotics*, S. Bruno and K. Oussama, Eds., pp. 275–284, Springer, Berlin, Heidelberg (2010).
6. C. Doya et al., "Seasonal monitoring of deep-sea megabenthos in Barkley Canyon cold seep by internet operated vehicle (IOV)," *PLoS One* **12**(5), e0176917 (2017).
7. J. Redmon and A. Farhadi, "YOLOv3: an incremental improvement," arXiv:1804.02767 (2018).
8. C. R. McClain, G. C. Feldman, and S. B. Hooker, "An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series," *Deep Sea Res. Part II: Top. Stud. Oceanogr.* **51**(1–3), 5–42 (2004).
9. W. E. Esaias et al., "An overview of MODIS capabilities for ocean science observations," *IEEE Trans. Geosci. Remote Sens.* **36**(4), 1250–1265 (1998).
10. B. A. Block et al., "A new satellite technology for tracking the movements of Atlantic bluefin tuna," *Proc. Natl. Acad. Sci. U. S. A.* **95**(16), 9384–9389 (1998).
11. T. Sagawa et al., "Mapping seagrass beds using IKONOS satellite image and side scan sonar measurements: a Japanese case study," *Int. J. Remote Sens.* **29**(1), 281–291 (2008).
12. P. T. Fretwell, I. J. Staniland, and J. Forcada, "Whales from space: counting southern right whales by satellite," *PLoS One* **9**(2), e88655 (2014).
13. V. Klemas, "Remote sensing techniques for studying coastal ecosystems: an overview," *J. Coastal Res.* **27**(1), 2–17 (2010).
14. F. Shkurti et al., "Multi-domain monitoring of marine environments using a heterogeneous robot team," in *IEEE/RSJ Int. Conf. Intell. Rob. Syst.*, pp. 1747–1753 (2012).
15. Y. Zhang et al., "Using an autonomous underwater vehicle to track a coastal upwelling front," *IEEE J. Oceanic Eng.* **37**(3), 338–347 (2012).

16. R. N. Smith and V. T. Huynh, "Controlling buoyancy-driven profiling floats for applications in ocean observation," *IEEE J. Oceanic Eng.* **39**(3), 571–586 (2014).

17. T. Alam et al., "A data-driven deployment approach for persistent monitoring in aquatic environments," in *IEEE Int. Conf. Rob. Comput.*, pp. 147–154 (2018).

18. H. Lu et al., "CONet: a cognitive ocean network," arXiv:1901.06253 (2019).

19. Argo Project, "Welcome on ARGO.NET, the International Argo Program homepage," http://www.argo.net/.

20. A. F. Thompson et al., "Satellites to seafloor: toward fully autonomous ocean sampling," *Oceanography* **30**(2), 160–168 (2017).

21. XPRIZE, "Ocean discovery XPRIZE," https://oceandiscovery.xprize.org/prizes/ocean-discovery.

22. C. M. Rochman et al., "Policy: classify plastic waste as hazardous," *Nature* **494**, 169–171 (2013).

23. A. L. Lusher et al., "Microplastics in arctic polar waters: the first reported values of particles in surface and sub-surface samples," *Sci. Rep.* **5**, 14947 (2015).

24. M. L. Pedrotti et al., "Changes in the floating plastic pollution of the Mediterranean Sea in relation to the distance to land," *PLoS One* **11**(8), e0161581 (2016).

25. S. D. Smith and R. J. Edgar, "Documenting the density of subtidal marine debris across multiple marine and coastal habitats," *PLoS One* **9**(4), e94593 (2014).

26. B. Munier and L. I. Bendell, "Macro and micro plastics sorb and desorb metals and act as a point source of trace metals to coastal ecosystems," *PLoS One* **13**(2), e0191759 (2018).

27. L. G. A. Barboza et al., "Microplastics cause neurotoxicity, oxidative damage and energy-related changes and interact with the bioaccumulation of mercury in the European seabass, Dicentrarchus labrax (Linnaeus, 1758)," *Aquat. Toxicol.* **195**, 49–57 (2018).

28. J. R. Jambeck et al., "Plastic waste inputs from land into the ocean," *Science* **347**, 768–771 (2015).

29. The Ocean Cleanup, "Ocean cleanup," https://www.theoceancleanup.com/.

30. L. Lebreton et al., "Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic," *Sci. Rep.* **8**(1), 4666 (2018).

31. Y. LeCun et al., "Backpropagation applied to handwritten zip code recognition," *Neural Comput.* **1**(4), 541–551 (1989).

32. S. Ren et al., "Faster R-CNN: towards real-time object detection with region proposal networks," in *Adv. Neural Inf. Process. Syst. (NIPS 2015)*, pp. 91–99 (2015).

33. J. Redmon et al., "You only look once: unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 779–788 (2016).

34. W. Liu et al., "SSD: single shot multibox detector," *Lect. Notes Comput. Sci.* **9905**, 21–37 (2016).

35. X. Li et al., "Fast accurate fish detection and recognition of underwater images with Fast R-CNN," in *OCEANS'15 MTS/IEEE*, Washington, pp. 1–5 (2015).

36. R. Mandal et al., "Assessing fish abundance from underwater video using deep neural networks," in *2018 Int. Joint Conf. Neural Networks (IJCNN)*, IEEE, pp. 1–6 (2018).

37. W. Xu and S. Matzner, "Underwater fish detection using deep learning for water power applications," arXiv:1811.01494 (2018).

38. M. Sung, S. C. Yu, and Y. Girdhar, "Vision based real-time fish detection using convolutional neural network," in *IEEE OCEANS 2017*, pp. 1–6 (2017).

39. Google, "Open images dataset V4," https://storage.googleapis.com/openimages/web/index.html.

40. "Darknet open source neural networks in C," https://pjreddie.com/darknet/.

41. T.-Y. Lin et al., "Microsoft coco: common objects in context," *Lect. Notes Comput. Sci.* **8693**, 740–755 (2014).

42. J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 7263–7271 (2017).

43. V. Roberge, M. Tarbouchi, and G. Labonté, "Comparison of parallel genetic algorithm and particle swarm optimization for real-time UAV path planning," *IEEE Trans. Ind. Inf.* **9**(1), 132–141 (2013).

44. T. Schmickl et al., "CoCoRo: the self-aware underwater swarm," in *IEEE Conf. Self-Adapt. and Self-Organ. Syst. Workshops*, pp. 120–126 (2011).

Jun-ichiro Watanabe is a senior researcher at the Center for Exploratory Research, Research and Development Group, Hitachi, Ltd. He received his BS and MS degrees in physics in 1996 and 1998 from Tohoku University, Japan. He received his PhD in computer science from Tsukuba University in 2009. His research interests include computer-human interaction, artificial intelligence, and remote sensing technologies.

Yang Shao is a researcher at the Center for Exploratory Research, Research & Development Group, Hitachi, Ltd. He received his BS degree in mathematics and physics from Tsinghua University, Beijing, in 2007 and his PhD in numerical analysis from the University of Tokyo, Japan, in 2013. His research interests include the development of reinforcement learning technology and artificial intelligence.

Naoto Miura is a senior researcher at the Center for Exploratory Research, Research & Development Group, Hitachi, Ltd. He received his BS degree in information and communication engineering in 1998, MS degree in information engineering in 2000, and his PhD in information science and technology from the University of Tokyo, Japan, in 2013. His research interests include the development of biometric authentication technology, computer vision, pattern recognition, and artificial intelligence.