



OPEN

DATA DESCRIPTOR

Long-term High Resolution Image Dataset of Antarctic Coastal Benthic Fauna

Simone Marini^{1,2,5} , Federico Bonofiglio^{1,5}, Lorenzo Paolo Corgnati^{1,5}, Andrea Bordone^{3,5}, Stefano Schiaparelli^{4,5} & Andrea Peirano^{3,5}

Antarctica is a remote place, the continent is covered by ice and its surrounding coastal areas are frozen for the majority of the year. Due to its peculiarity the observation of the underwater organisms is particularly difficult, complicated by logistic factors. We present a long-term dataset consisting of 755 images acquired by using a non-invasive, autonomous imaging device and encompassing both the Antarctic daylight and dark periods, including the corresponding transition phases. All images have the same field of view showing the benthic fauna and part of the water column above, including fishes present in the monitored period. All the images are manually annotated after a visual inspection performed by expert biologists. The extended monitoring period and the annotated images make the dataset a valuable benchmark suitable for studying the dynamics of the long-term Antarctic underwater fauna as well as for developing and testing algorithms for automated image analysis focused on the recognition and classification of the Antarctic organisms and the automated analysis of their long-term dynamics.

Background & Summary

Antarctica is one of the most remote and inaccessible places on the Earth. The continent is constantly covered by ice and its surrounding coastal areas are frozen for the vast majority of the year¹. Due to its peculiarity, the Antarctic continent and the corresponding marine environment are extremely sensible to both climate changes and anthropic impacts^{2,3}. For this reason the need of long-term monitoring actions is strongly supported by the scientific community²⁻⁶ and by the international committees steering the Antarctic research⁷⁻⁹.

Many studies collect and analyze Antarctic long-term datasets dealing with benthic fauna, fishery and plankton communities^{3,10-12}, while other studies investigate ice-sheet dynamics, iceberg tracking and meteorological data¹³⁻¹⁸.

Long-term time series are particularly important as they disclose relevant information on the temporal variation of the environment and of the living communities. This is true especially for the benthic fauna, where systematic observations repeated in time provide information useful for a detailed understanding and predictions of their dynamics. Furthermore, structure and composition of the benthic fauna are used as an effective tool to identify the impacts of environmental factors as well as the impact of human activities^{5,6,19-21}.

Underwater video-monitoring methodologies are a consolidated approach for the ecosystem investigation²²⁻²⁴. Their action is non-destructive and the acquired data provide a highly informative content about the observed organisms^{19,25,26}. Despite its effectiveness underwater video monitoring in the Antarctica regions is rarely practiced. Examples of short-term video monitoring activities are the photogrammetry approach proposed in^{6,20}, where seabed footage was used for producing 3D models of the sea floor for identifying and counting the lying organisms. Other examples rely on time-lapse cameras deployed for few days in coastal areas, for observing benthic organisms^{27,28}.

Within this context, this article describes the image dataset acquired within the ICE-LAPSE project²⁹ through a long-term autonomous video monitoring activity³⁰ continuously performed from January to November 2017 near the Italian Antarctic station Mario Zucchelli (Terra Nova Bay, Ross Sea).

¹National Research Council of Italy (CNR), Institute of Marine Sciences, La Spezia, 19132, Italy. ²Stazione Zoologica Anton Dohrn, Naples, 80121, Italy. ³ENEA-Marine Environment Research Centre, La Spezia, 19132, Italy. ⁴DISTAV, Università di Genova, Genova, 16132, Italy. ⁵These authors contributed equally: Simone Marini, Federico Bonofiglio, Lorenzo Paolo Corgnati, Andrea Bordone, Stefano Schiaparelli, Andrea Peirano. ✉e-mail: simone.marini@sp.ismar.cnr.it

The image dataset was acquired by using a stand-alone imaging device, specifically conceived for autonomous deployment extended in time^{30–33}, positioned on the seabed of a coastal area at 20 m depth. The dataset consists of 755 images acquired every nine hours, encompassing both the Antarctic daylight and dark periods, including the corresponding transition phases. The imaging device was deployed onboard a fixed lander, all the images have the same field of view and show a sponge *Mycale (Oxymycale) acerata* Kirkpatrick, 1907 and its surrounding area including the corresponding fauna. The acquired dataset was visually inspected by expert biologists and the organisms present in each image were manually tagged for a total of 23881 individuals belonging to 12 different taxa.

The monitoring period extended in time and the tagged images make the dataset a valuable benchmark suitable for studying the dynamics of the Antarctic benthic fauna as well as for developing and testing algorithms for automated image analysis focused on the recognition and the classification of the Antarctic organisms and the automated analysis of their long-term dynamics.

Methods

Study area. The underwater images were acquired in Tethys Bay, a coastal area near the Italian Antarctic station Mario Zucchelli (74° 41.410' S, 164° 06.233' E), as shown in Fig. 1a,b.

Tethys Bay is an inlet 1600m wide and 3000 m long, oriented to Northwest. It is surrounded by glaciers and granite cliffs and its maximum depth is about 280 m in the central part. The sea pack in the bay persists on average from the end of February to the beginning of January^{6,34,35}. The monitored site, known as “Zecca”, is located in front of a granite cliff as shown in Fig. 1c. The seabed at the “Zecca” site is characterized by a moderate slope composed of pebbles and scattered with large stone blocks as shown in Fig. 1d. The site, was selected because it is suitable for long term experiments as the seabed is protected from icebergs and wave storms.

The monitored area is rich of sponges *Mycale (Oxymycale) acerata* Kirkpatrick, 1907, sea urchins *Sterechinus neumayeri* (Meissner, 1900) and sea stars, mainly *Odontaster validus* Koehler, 1906 and *Diplasterias brucei* (Koehler, 1907) present all over the year, while fishes and several benthic organisms show different behaviors depending on the Antarctic season³⁰.

The average duration of daytime at the Mario Zucchelli Station was estimated and discussed in³⁶, resulting that from 3rd November to 7th February the sun is above the horizon for 24 hours per day continuously, while from 3rd May to 10th August natural light is completely absent. The natural light dynamics in Antarctica, are one of the parameter driving the faunal dynamics^{37,38} and the diffused light reaching the seabed through the icepack is an important factor conditioning the benthic fauna. The accumulation of snow on the sea pack or the shadow projected by cliffs surrounding the studied area can locally change the underwater diffused light dynamics and in the period in which the sun is constantly present above the horizon, the monitored site shows a daily light/shadow cycle. As discussed in^{28,30}, the shadow effect is caused by the position of the cliff with respect to the monitoring site and such a light variation has been captured by the acquired images.

Image acquisition device. The image dataset was acquired by an imaging device based on the European patent EP 2863257 B1³¹. The imaging device was installed on a metal frame deployed through a whole practiced in the sea pack (Fig. 1b) and secured by divers with metal pins and ballasts on the seabed at 20 m depth, distant about one meter from a sponge *M. acerata* (Fig. 1d). Though the technology described in the patent can execute on board image processing algorithms and perform data communication³⁹, the device deployed for acquiring the presented dataset was designed only for acquiring and storing underwater images, without any communication and image processing features.

The imaging device was based on a commercial Canon EOS 600D reflex camera whose behavior was controlled by a firmware based on the Magic Lantern toolkit⁴⁰ and installed on the camera for the automated acquisition of the image dataset. The customized firmware was aimed at the automatic shooting of the images and, during the deployment activities performed by divers, at redirecting a video stream to a service display used for the correct positioning of the camera (see Fig. 2).

The optical group used for acquiring the images was a CANON Zoom Lens EF-S 10–18 mm (1:4.5–5.6 IS STM ϕ 67 mm), set with focal length equal to 12 mm. The lighting system consist of four high performance LEDs, (3.4 V, 1 W, 350 mA, 75 lm, cold white color) embedded in epoxy resin and wired to the device underwater case, as shown in Fig. 2a. The four LEDs were split into two groups (two LEDs each group) each one displaced about 40 cm from the device and oriented toward the sponge in front of the lander as shown in Fig. 1d. In order to position the imaging device correctly in front of the sponge, a service display was installed inside the transparent underwater case (Fig. 2) aimed at providing a live stream of the subject framed by the camera. The overall behavior of the imaging device was defined by a control hardware component managed by a diver operator through a magnetic-switch that allowed to switch between the live view stream and two operative modes for the image acquisition³⁰.

The imaging device was powered by a main battery pack consisting of nine primary batteries Tadiran SL-2700 Series (3.6 V, 35.0 Ah, size DD) coupled with a secondary rechargeable battery pack consisting of four Samsung 18650 batteries (3.6 V, 3450mAh, max discharge current 8 A). The main battery pack was used for providing the energy needed for the long term monitoring action, while the secondary battery pack was used for supplying the impulse of power needed for switching on the camera and the lighting system.

More details on the imaging device used for acquiring the Antarctic image dataset are provided in³⁰, while other monitoring activities using a similar technology can be found in^{32,33,41}.



Fig. 1 Overview of the monitoring activity near Mario Zucchelli Station, Terra Nova Bay (Ross Sea) - $74^{\circ} 41.410' S, 164^{\circ} 06.233' E$ (a) and (b); the deployment site on the frozen sea surface (c) and at the seabed (d).

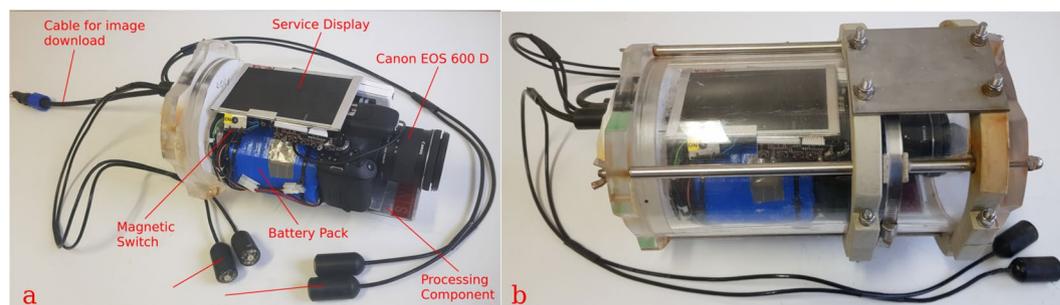


Fig. 2 The autonomous and stand-alone imaging device used for the image acquisition and its main hardware components (a); The imaging device inside its transparent underwater case (b).

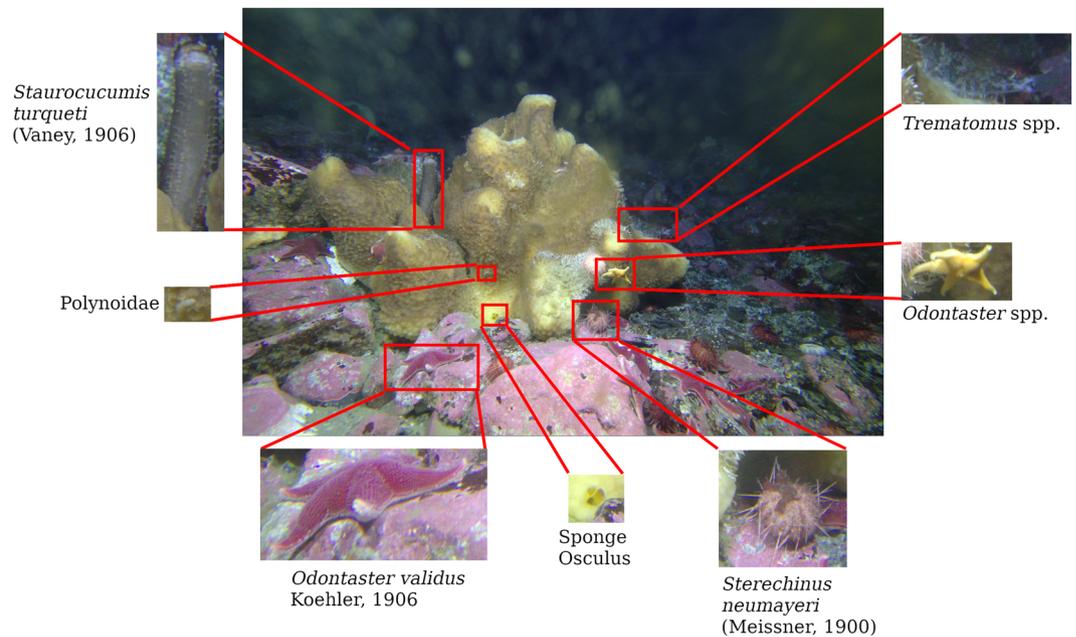


Fig. 3 Example of underwater image acquired by the imaging device. Detailed examples of the present faunal organisms are highlighted.

Data Records

The dataset consists of 755 high resolution color images acquired every nine hours from 24 January 2017 to 15 November 2017. Each image has a resolution of 5207×3469 pixels and was saved in the Canon proprietary raw format CR2⁴². The collected images were visually inspected and the framed organisms were counted, tagged and labeled. The Fig. 3 shows an example of acquired image highlighting some organisms captured by the imaging device.

The image dataset encompasses different environmental conditions including short-term and seasonal light dynamics, icepack coverage and a short period with occurrence of bio-fouling on the camera porthole. The images shown in Fig. 4 provide several examples of the environmental conditions occurring along the acquisition period of the image dataset. Each row contains five images acquired consecutively, while different rows correspond to images sequences acquired in different periods and spanning the whole image dataset.

The images in Fig. 4a–c exemplify the acquisitions between the end of February and the beginning of May. Each of the three sequences alternate bright images, where the natural ambient light is diffused through the water column (and the icepack when present), and darker images corresponding to the dark short-term periods produced by the light/shadow cycle caused by the cliff in proximity of the monitored site^{28,30}. During the shadow periods the natural diffused light is attenuated and the foreground objects (e.g. the sponge) are illuminated by the lighting system of the imaging device. The water's greenish hue in Fig. 4a, acquired in February, is caused by a huge phytoplankton bloom occurring after the icepack melting. In the sequence shown in Fig. 4b, corresponding to march, the icepack is present again and the strong reduction of diffused light might be probably caused by a temporary cloudy sky or by the accumulation of snow on the icepack surface. The images in Fig. 4b also show the presence of bio-fouling on the porthole of the camera, following the phytoplankton bloom. Such bio-fouling is suddenly reduced from the sequence in Fig. 4c, corresponding to April, until the end of the acquisition period (Fig. 4j), corresponding to November.

Examples of images acquired during the Antarctic night are shown in Fig. 4d, corresponding to June. In this case the bio-fouling is not present anymore, the diffused natural light is completely absent and all the images in the sequence are illuminated by the lighting system of the imaging device. After the end of the Antarctic night (end of August) the natural light increases with some discontinuities, Fig. 4e–g, until the end of September where a significant decrease happens as shown in Fig. 4h–j, yet probably caused by a cloudy sky or by the accumulation of snow on the sea pack. In all the images from Fig. 4b to Fig. 4j it is possible to note the presence of the frozen sea surface, during the daylight periods.

After the acquisition of the first 109 images (out of 755 images), the imaging device moved slightly downward and rightward, producing an offset of a few pixels (30×108 pixels) in the position of the sponge with respect to the borders of the image.

The image dataset is accessible through the Zenodo repository⁴³ and consists of the records presented in Table 1.

All the images in the archives Raw_Images, JPG_Images, 640×427 and ICE-LAPSE_tags have a filename starting with the string IMG_<id>, where <id> is a number varying between 609 and 1367 that unambiguously identifies each image file. The two archives Raw_Images and JPG_Images contain the full resolution

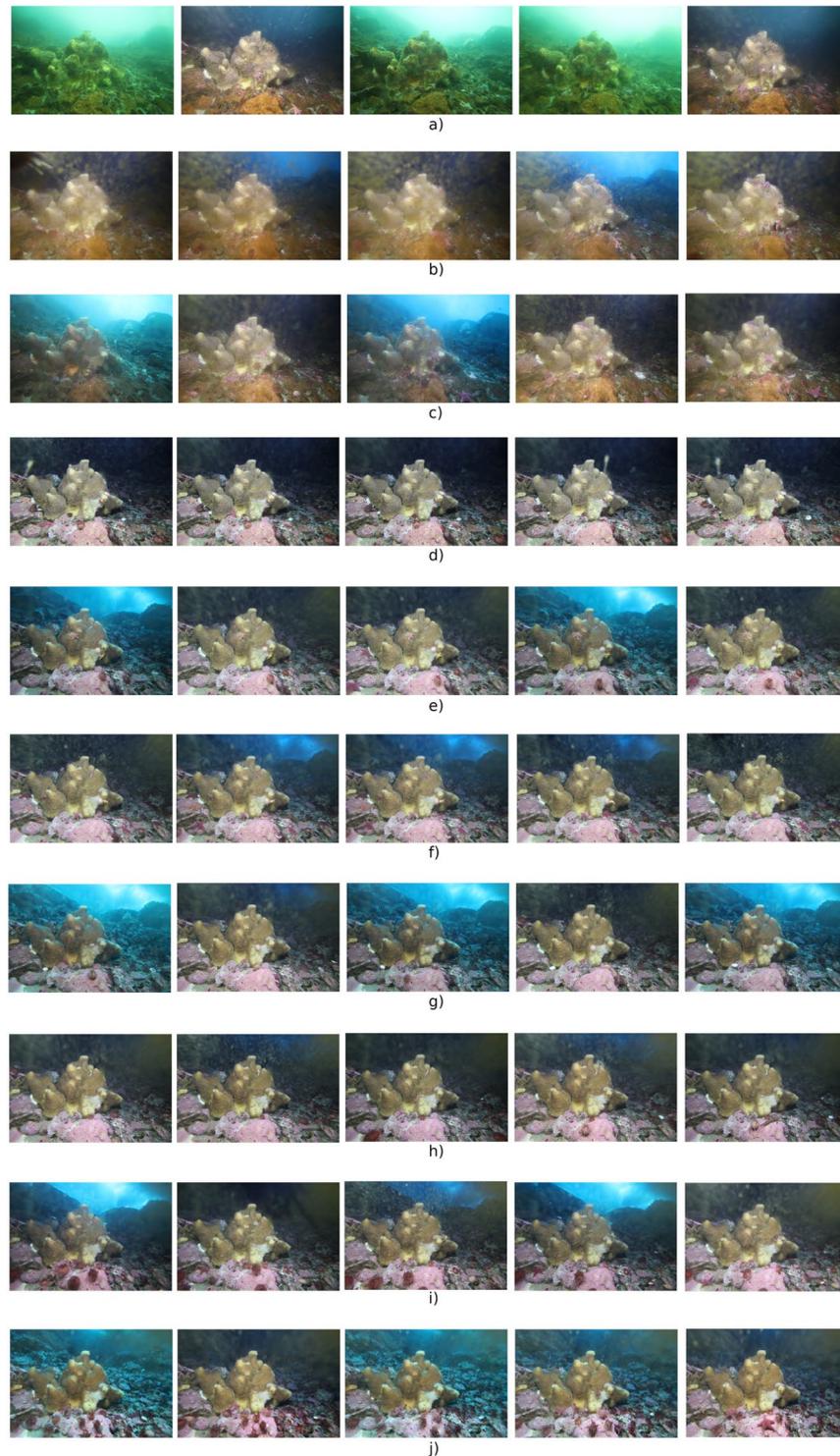


Fig. 4 A set of examples sampled from the image dataset. Each row consists of images acquired consecutively, while different rows correspond to different periods within the monitoring activity.

images in raw and JPEG formats, respectively. The ICE-LAPSE_tags archive provides information on the organisms classification performed by expert biologists. It contains two folders and three files:

- TaggedImages
- TextualTags
- Classes.txt
- SpeciesColorLabel.jpg
- SpeciesDistribution.csv

Filename	Content	Size
Raw_Images.tgz	Raw images with 5207 × 3469 pixels resolution in CR2 format	18.6 GB
JPG_Images.tgz	Images with 5207 × 3469 pixels resolution in JPEG format	2.0 GB
640 × 427.tgz	Images with 640 × 427 pixels resolution in JPEG format	72.3 MB
ICE-LAPSE_tags.tgz	Information on the organisms' tags and species	2.7 GB
iFDO_Dataset.tgz	Images to be used within the iFDO standard ⁴⁵	2.0 GB
Deployment in Tethys bay - Mario Zucchelli Station - 2015_GUARD1-PNRA-2015_iFDO.yaml	iFDO metadata ⁴⁵	359 KB

Table 1. The image dataset archives.

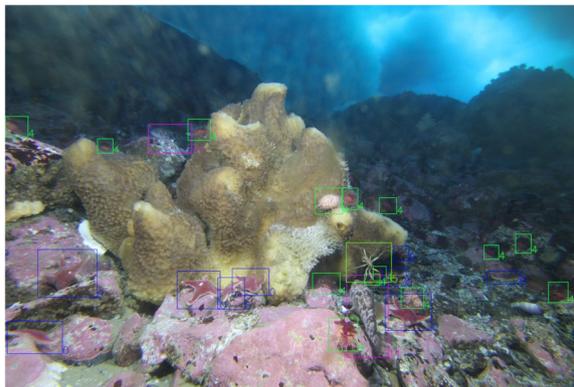


Fig. 5 An example of tagged image, where the labelled (colored) boxes represent the following taxa: 0 *Odontaster validus* Koehler, 1906; 4 *Sterechinus neumayeri* (Meissner, 1900); 5 *Ammothea* sp.; 11 *Trematomus* spp.

The TaggedImages folder contains the dataset images where each organism identified in the visual inspection is highlighted by a colored bounding-box (tag) accompanied by a number. The color and the number associated to each tag represent the organism's species as listed in the files Classes.txt or SpeciesColorLabel.jpg. An example of a tagged image is shown in Fig. 5.

The TextualTags folder contains one text file for each image of the dataset. Each row in the file represents a tag in the corresponding image. The image tagging has been performed with the labelImage⁴⁴ graphical tool and each row in the file has the following format:

$$\langle \text{label_id} \rangle \langle x \rangle \langle y \rangle \langle w \rangle \langle h \rangle ,$$

where $\langle \text{label_id} \rangle$ is a number identifying the species of the tagged organisms, as listed in the file Classes.txt, $\langle x \rangle$, $\langle y \rangle$, $\langle w \rangle$ and $\langle h \rangle$ are the coordinates of the centre of the tagging box and the corresponding width and height, respectively. The file SpeciesDistribution.csv is an abundance matrix per species where the first column contains the time-stamp of each image, the second column contains the image indices and the remaining thirteen columns contain the number of individuals occurring in the images for each species.

The archive iFDO_Dataset.tgz contains the image dataset formatted according to the image FAIR Digital Object (iFDO) standard⁴⁵, whose metadata are encoded in the file "Deployment in Tethys bay - Mario Zucchelli Station - 2015_GUARD1-PNRA-2015_iFDO.yaml". Each image of the iFDO dataset is characterized by a universally unique identifier (uuid) and the corresponding SHA-256 image encoding. The metadata encoded in the yaml file contain general information about the whole dataset (e.g. the dataset uuid, dataset abstract, dataset authors, type of illumination, type of instrument used) and specific metadata for each image (e.g. depth, latitude, longitude, altitude, uuid, the SHA-256 hashing code).

Finally, each dataset image, either in raw or JPEG format, contains the corresponding EXIF data, encoding several image acquisition information like for example time exposure, iris aperture, focal length and ISO, together with other low level metadata.

Technical Validation

All the images in the dataset were visually inspected and all the animals in the images were annotated using the labelImg software tool⁴⁴.

In order to avoid any ambiguity in the classification of organisms, the dataset was independently inspected by two expert biologists. The inspection activity produced the list of taxa shown in Fig. 6, that was used by a third operator to annotate the images. In the event of lack of agreement between the two experts, the genus or the family was used to classify an individual.

The visual inspection resulted in to a total of 23881 organisms visually classified and tagged, as summarized in Table 2. The fauna classification shown in the table is based on the scientific literature describing the fauna

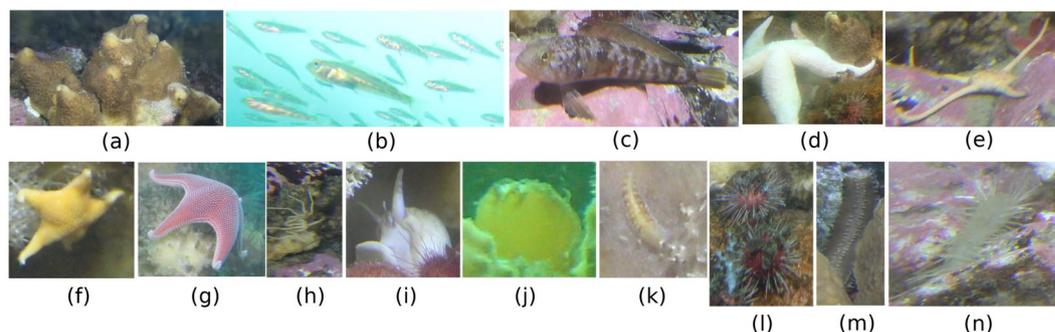


Fig. 6 Examples of species present in the image dataset as reported in³⁰. **(a)** *Mycale (Oxymycale) acerata* Kirkpatrick, 1907; **(b)** *Trematomus newnesi* Boulenger, 1902; **(c)** *Trematomus* sp.; **(d)** *Diplasterias brucei* (Koehler, 1907); **(e)** Ophiuridae; **(f)** *Odontaster* sp.; **(g)** *Odontaster validus* Koehler, 1906; **(h)** *Ammothea* sp.; **(i)** *Neobuccinum eatoni* (E. A. Smith, 1875); **(j)** Unclassified; **(k)** Polynoidae; **(l)** *Sterechinus neumayeri* (Meissner, 1900); **(m)** *Staurocucumis turqueti* (Vaney, 1906); **(n)** *Flabegraviera mundata* (Gravier, 1906).

Species	Number of Tags
<i>Sterechinus neumayeri</i> (Meissner, 1900)	16605
<i>Odontaster validus</i> Koehler, 1906	5965
<i>Trematomus newnesi</i> Boulenger, 1902	401
<i>Trematomus</i> spp.	267
Polynoidae	143
<i>Odontaster</i> spp.	139
<i>Diplasterias brucei</i> (Koehler, 1907)	137
Ophiuridae	79
<i>Ammothea</i> sp.	70
<i>Neobuccinum eatoni</i> (E. A. Smith, 1875)	23
Unclassified	21
<i>Staurocucumis turqueti</i> (Vaney, 1906)	17
<i>Flabegraviera mundata</i> (Gravier, 1906)	14
Total tags	23881

Table 2. Distribution of image tags per species, as reported in³⁰.

present in the area of Mario Zucchelli Station. This classification is supported by the continuous samplings and taxonomic analysis performed in the last twenty years^{3,46–48}.

Figure 6 shows some examples of the 14 taxa present in the acquired images.

Several analysis on the image dataset were presented and discussed in³⁰. The manual tags were used to investigate the long- and mid- term trends of the organism abundance, as well as the single species distribution along the monitored period. The abundance trends capture the overall correlation between the seasonal trend of the natural light and the behavior of the benthic fauna, together with short-term fluctuation of the individual counts suggesting more complex faunal behavior driven by environmental parameters independent by the underwater diffused light. The single species temporal distribution shows that different species have different behaviors, depending on the seasonal characteristics, as reported in³⁰.

Also automated image analyses have been performed on the proposed image dataset. A computer vision based approach was used for estimating the underwater diffused light from the acquired images. The obtained results are coherent with the average daylight duration at Mario Zucchelli Station discussed in³⁶ and highlight the light/shadow cycle reported in²⁸. Other automated analyses reported in³⁰ deal with the movement estimation of the organisms in proximity of the framed sponge, the analysis of the temporal dynamics of the sponge's oscula and the use of the YOLO Convolutional Neural Network^{49,50} for automatically detecting sea stars and sea urchins.

Code availability

The image dataset was tagged by using the labelImage⁴⁴ graphical tool.

Received: 20 July 2022; Accepted: 23 November 2022;

Published online: 03 December 2022

References

- Rogers, A. *et al.* Antarctic futures: An assessment of climate-driven changes in ecosystem structure, function, and service provisioning in the southern ocean. *Annual Review of Marine Science* **12**, 87–120, <https://doi.org/10.1146/annurev-marine-010419-011028> (2020).
- Tin, T. *et al.* Impacts of local human activities on the antarctic environment. *Antarctic Science* **21**, 3–33, <https://doi.org/10.1017/S0954102009001722> (2009).
- Pineda-Metz, S. E. A., Gerdes, D. & Richter, C. Benthic fauna declined on a whitening antarctic continental shelf. *Nature Communications* **11**, 2226, <https://doi.org/10.1038/s41467-020-16093-z> (2020).
- Convey, P. Antarctic terrestrial biodiversity in a changing world. *Polar Biology* **34**, 1629, <https://doi.org/10.1007/s00300-011-1068-0> (2011).
- Kang, Y. H. *et al.* Composition and structure of the marine benthic community in terra nova bay, antarctica: Responses of the benthic assemblage to disturbances. *PLOS ONE* **14**, 1–16, <https://doi.org/10.1371/journal.pone.0225551> (2019).
- Piazza, P. *et al.* Underwater photogrammetry in antarctica: long-term observations in benthic ecosystems and legacy data rescue. *Polar Biology* **42**, 1061–1079, <https://doi.org/10.1007/s00300-019-02480-w> (2019).
- SOOS. Southern Ocean Observing System - Report on the 2017 Ross Sea Working Group Meeting. <http://www.soos.aq>. [Online; accessed 2022/11/15] (2017).
- SCAR. Scientific Committee on Antarctic Research. <https://www.scar.org>. [Online; accessed 2022/11/15] (2021).
- ANTOS. Antarctic near-shore and terrestrial observing system. <https://www.scar.org/science/antos/home>. [Online; accessed 2022/11/15] (2015).
- Dayton, P. K. *et al.* Benthic responses to an antarctic regime shift: food particle size and recruitment biology. *Ecological Applications* **29**, e01823, <https://doi.org/10.1002/eap.1823> (2019).
- Watters, G. M., Hinke, J. T. & Reiss, C. S. Long-term observations from antarctica demonstrate that mismatched scales of fisheries management and predator-prey interaction lead to erroneous conclusions about precaution. *Scientific Reports* **10**, 2314, <https://doi.org/10.1038/s41598-020-59223-9> (2020).
- Bolinesi, F. *et al.* Spatial-related community structure and dynamics in phytoplankton of the ross sea, antarctica. *Frontiers in Marine Science* **7**, <https://doi.org/10.3389/fmars.2020.574963> (2020).
- Stenni, B. *et al.* Three-year monitoring of stable isotopes of precipitation at concordia station, east antarctica. *The Cryosphere* **10**, 2415–2428, <https://doi.org/10.5194/tc-10-2415-2016> (2016).
- Ramesh, K. & Soni, V. Perspectives of antarctic weather monitoring and research efforts. *Polar Science* **18**, 183–188, <https://doi.org/10.1016/j.polar.2018.04.005> (2018). Recent Advances in Climate Science of Polar Region (to commemorate the contributions of Late Dr. S.Z. Qasim, a pioneering doyen of the Indian Polar programme).
- Shepherd, A. *et al.* Mass balance of the antarctic ice sheet from 1992 to 2017. *Nature* **558**, 219–222, <https://doi.org/10.1038/s41586-018-0179-y> (2018).
- Budge, J. S. & Long, D. G. A comprehensive database for antarctic iceberg tracking using scatterometer data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **11**, 434–442, <https://doi.org/10.1109/JSTARS.2017.2784186> (2018).
- Rignot, E. *et al.* Four decades of antarctic ice sheet mass balance from 1979–2017. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 1095–1103, <https://doi.org/10.1073/pnas.1812883116> (2019).
- Barbat, M. M., Rackow, T., Wesche, C., Hellmer, H. H. & Mata, M. M. Automated iceberg tracking with a machine learning approach applied to sar imagery: A weddell sea case study. *ISPRS Journal of Photogrammetry and Remote Sensing* **172**, 189–206, <https://doi.org/10.1016/j.isprsjprs.2020.12.006> (2021).
- Aguzzi, J. *et al.* New high-tech flexible networks for the monitoring of deep-sea ecosystems. *Environmental Science & Technology* **53**, 6616–6631, <https://doi.org/10.1021/acs.est.9b00409> (2019).
- Piazza, P., Gattone, S., Guzzi, A. & Schiaparelli, S. Towards a robust baseline for long-term monitoring of antarctic coastal benthos. *Hydrobiologia* **847**, 1753–1771, <https://doi.org/10.1007/s10750-020-04177-2> (2020).
- Rountree, R. *et al.* Towards an optimal design for ecosystem-level ocean observatories. *Oceanography and Marine Biology* **58**, 79–105, <https://doi.org/10.1201/9780429351495-2> (2020).
- Katsanevakis, S. *et al.* Monitoring marine populations and communities: Methods dealing with imperfect detectability. *Aquatic Biology* **16**, 31–52, <https://doi.org/10.3354/ab00426> (2012).
- Zampoukas, N. *et al.* Technical guidance on monitoring for the marine strategy framework directive. *Tech. Rep., European Commission, Report EUR 26499* (2014).
- Bicknell, A. W., Godley, B. J., Sheehan, E. V., Votier, S. C. & Witt, M. J. Camera technology for monitoring marine biodiversity and human impact. *Frontiers in Ecology and the Environment* **14**, 424–432, <https://doi.org/10.1002/fee.1322> (2016).
- European Marine Board. Working Group on Big Data in Marine Science. <https://www.marineboard.eu/publications/big-data-marine-science>. [Online; accessed 2022/11/15] (2020).
- Zurovietz, M. & Nattkemper, T. W. Current trends and future directions of large scale image and video annotation: Observations from four years of biigle 2.0. *Frontiers in Marine Science* **8**, <https://doi.org/10.3389/fmars.2021.760036> (2021).
- Kim, S. L., Thurber, A., Hammerstrom, K. & Conlan, K. Seastar response to organic enrichment in an oligotrophic polar habitat. *Journal of Experimental Marine Biology and Ecology* **346**, 66–75, <https://doi.org/10.1016/j.jembe.2007.03.004> (2007).
- Peirano, A., Bordone, A., Marini, S., Piazza, P. & Schiaparelli, S. A simple time-lapse apparatus for monitoring macrozoobenthos activity in antarctica. *Antarctic Science* **28**, 473–474, <https://doi.org/10.1017/S0954102016000377> (2016).
- Peirano, A., Marini, S., Bordone, A. & Schiaparelli, S. ICE-LAPSE: Analysis of antarctic benthos dynamics by using non-destructive monitoring devices and permanent stations, pnra 2013/az1.16, funded by the italian national antarctic program (2015–2016).
- Marini, S. *et al.* Long-term automated visual monitoring of antarctic benthic fauna. *Methods in Ecology and Evolution* **13**, 1746–1764, <https://doi.org/10.1111/2041-210X.13898> (2022).
- Marini, S. *et al.* EP2863257 (A1) - Underwater images acquisition and processing system. <https://data.epo.org/gpi/EP2863257B1>. [Online; accessed 2022/11/15] (2013).
- Corgnati, L. *et al.* Looking inside the ocean: Toward an autonomous imaging system for monitoring gelatinous zooplankton. *Sensors* **16**, 2124, <https://doi.org/10.3390/s16122124> (2016).
- Marini, S. *et al.* Automated estimate of fish abundance through the autonomous imaging device guard1. *Measurement* **126**, 72–75, <https://doi.org/10.1016/j.measurement.2018.05.035> (2018).
- Pensieri, S. *et al.* Environmental acoustic noise observations in tethys bay (terra nova bay, ross sea, antarctica). In *2014 Oceans - St. John's*, 1–6, <https://doi.org/10.1109/OCEANS.2014.7003196> (2014).
- Jung, J. *et al.* Multibeam bathymetry and distribution of clay minerals on surface sediments of a small bay in terra nova bay, antarctica. *Minerals* **11**, <https://doi.org/10.3390/min11010072> (2021).
- Balog, I. *et al.* Estimation of direct normal irradiance at antarctica for concentrated solar technology. *Applied System Innovation* **2**, <https://doi.org/10.3390/asi2030021> (2019).
- Caputi, S. S. *et al.* Seasonal food web dynamics in the antarctic benthos of tethys bay (ross sea): Implications for biodiversity persistence under different seasonal sea-ice coverage. *Frontiers in Marine Science* **7**, 1046, <https://doi.org/10.3389/fmars.2020.594454> (2020).
- van Leeuwe, M. A. *et al.* Annual patterns in phytoplankton phenology in antarctic coastal waters explained by environmental drivers. *Limnology and Oceanography* **65**, 1651–1668, <https://doi.org/10.1002/lno.11477> (2020).

39. OEngineering. OEngineering s.r.l. - GUARD-1, Underwater Autonomous Smart Camera. <https://www.oengineering.eu/GUARD-1/>. [Online; accessed 2022/11/15] (2021).
40. Magic Lantern. <https://magiclantern.fm>. [Online; accessed 2022/11/15] (2021).
41. Marini, S. *et al.* Guard1: An autonomous system for gelatinous zooplankton image-based recognition. In *OCEANS 2015 - Genova*, 1–7, <https://doi.org/10.1109/OCEANS-Genova.2015.7271704> (2015).
42. CR2. The Canon RAW (CRW) File Format. https://exiftool.org/canon_raw.html. [Online; accessed 2022/11/15] (2022).
43. Marini, S. *et al.* ICE-LAPSE image dataset. *Zenodo* <https://doi.org/10.5281/zenodo.6418163> (2022).
44. LabelImg. A graphical image annotation tool. <https://github.com/tzutalin/labelImg>. [Online; accessed 2022/11/15] (2021).
45. Schoening, T. *et al.* Making marine image data fair. *Scientific Data* **9**, 414, <https://doi.org/10.1038/s41597-022-01491-3> (2022).
46. Cattaneo-Vietti, R., Chiantore, M., Schiaparelli, S. & Albertelli, G. Shallow- and deep-water mollusc distribution at terra nova bay (ross sea, antarctica). *Polar Biology* **23**, 173–182, <https://doi.org/10.1007/s003000050024> (2000).
47. Cattaneo-Vietti, R. *et al.* *Spatial and Vertical Distribution of Benthic Littoral Communities in Terra Nova Bay*, 503–514 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2000).
48. Cummings, V. J. *et al.* Linking ross sea coastal benthic communities to environmental conditions: Documenting baselines in a spatially variable and changing world. *Frontiers in Marine Science* **5**, 232, <https://doi.org/10.3389/fmars.2018.00232> (2018).
49. Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788, <https://doi.org/10.1109/CVPR.2016.91> (2016).
50. YOLO V5. <https://github.com/ultralytics/yolov5>. [Online; accessed 2022/11/15] (2022).

Acknowledgements

The Research leading to these results has received funding from the project “ICE-LAPSE: Analysis of Antarctic benthos dynamics by using non-destructive monitoring devices and permanent stations”, PNRA 2013/AZ1.16, funded by the Italian National Antarctic Program. We are indebted to the Comando Subacquei ed Incursori (COMSUBIN) of the Italian Navy for help and assistance during the dives and to S. Cocito and C. Lombardi for their support during the test of the imaging device. We are also grateful to M. La Mesa for the determination of fish species in the captured images. This paper is also an Italian contribution to the CCAMLR CONSERVATION MEASURE 91-05 (2016) for the Ross Sea region Marine Protected Area, specifically, addressing the priorities of Annex 91-05/C, and a contribution to the SCAR-ANTOS Expert Group (<https://www.scar.org/science/antos/home/>). This research activity was also partially funded by the ENDURUNS project [Horizon 2020; Grant Agreement H2020-MG-2018-2019-2020 n.824348] and the JERICO-S3 project [Horizon 2020; Grant Agreement no. 871153].

Author contributions

A.P., S.M. and S.S. conceived the pilot study; S.M. and L.P.C. conceived the imaging device; A.P., S.M. and A.B. prepared and tested the imaging device for the Antarctic long term monitoring; A.P. and S.M. participated to the XXXI Antarctic expedition where the imaging device was deployed; A.P., S.S. and F.B. visually inspected the acquired images and determined the species; all the authors wrote and reviewed the manuscript.

Competing interests

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

Additional information

Correspondence and requests for materials should be addressed to S.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022