

PREPRINT

Posted on 13/11/2020

DOI: <https://doi.org/10.3897/arphapreprints.e60597>

A system for automated analysis of subsea movies using citizen science and machine learning

 Victor Anton, Jannes Germishuys,  Matthias Obst

Not peer-reviewed, not copy-edited manuscript.

A system for automated analysis of subsea movies using citizen science and machine learning

Victor Anton[‡], Jannes Germishuys[§], Matthias Obst[¶]

[‡] Wildlife.ai, New Plymouth, New Zealand

[§] Combine AB, Gothenburg, Sweden

| Department of Marine Sciences, Göteborg University, Gothenburg, Sweden

[¶] SeAnalytics AB, Gothenburg, Sweden

Corresponding author: Victor Anton (victor@wildlife.ai), Matthias Obst (matthias.obst@marine.gu.se)

Abstract

Background

This paper describes a data system to analyse large amounts of subsea movie data for marine ecological research. The system consists of three distinct modules for data management and archiving, citizen science, and machine learning in a high performance computation environment. It allows scientists to upload underwater footage to a customised citizen science website hosted by Zooniverse, where volunteers from the public classify the footage. Classifications with high agreement among citizen scientists are then used to train machine learning algorithms. An application programming interface allows researchers to test the algorithms and track biological objects in new footage. We tested our system using recordings from remotely operated vehicles (ROVs) in a Marine Protected Area, the Kosterhavet National Park in Sweden. Results indicate a strong decline of cold-water corals in the park over a period of 15 years, showing that our system allows to effectively extract valuable occurrence and abundance data for key ecological species from underwater footage. We argue that the combination of citizen science tools, machine learning, and high performance computers are key to successfully analyse large amounts of image data in the future, suggesting that these services should be consolidated and interlinked by national and international research infrastructures.

New information

Novel information system to analyse marine underwater footage.

Keywords

marine biodiversity, autonomous underwater vehicles, remotely operated vehicles, artificial intelligence, big data, image analysis, participatory science, Essential Biodiversity Variables, research infrastructure

Introduction

Biological observation techniques in the marine environment need to improve radically to serve our understanding of marine ecosystems under long-term global change and under the influence of multiple stressors ([Benedetti-Cecchi et al. 2018](#)). Today biologists increasingly have access to autonomously operated technologies for data collection, offering the opportunity to generate enormous volumes of data. This is especially the case for high-definition optical imagery recorded by ROV's (remotely operated vehicles), AUVs (autonomous underwater vehicles), drop-cameras, video plankton recorders, and drones (Bean et al. 2017, Danovaro et al. 2016). Although such image-based observations may revolutionize the fields of marine biology and biodiversity monitoring, these methods also impose completely new demands for data management and processing on researchers. Hence, such in-situ monitoring systems need to be coupled to data services that allow for swift exploration, processing, and long-term storage (Guidi et al. 2020). Some of these services already exist including the EcoTaxa application for analysis of large amounts of plankton imagery (M et al. 2017), the CoralNet tool that allows researchers to analyse coral images (Williams et al. 2019), and [FathomNet](#) a system offering machine learning algorithms and training data to analyse deep sea footage. Although these platforms have pioneered the daily use of image analysis tools in marine science, they can not provide all functions needed by the fast growing community of users. Archiving functions for example often fall under national responsibility and can not be provided only by a single global system. Also, training datasets need to be specifically developed for the region of interest and for the scientific question at hand. For this reason, it is important to establish local image analysis services, which can interoperate with larger and global platforms in the future.

Here we present a system for managing, processing, and analysing large amounts of subsea movie data for marine ecological research. The system allows scientists to upload underwater movies to a customised citizen science website hosted by Zooniverse, where volunteers from the public classify the footage. Classifications with high agreement among citizen scientists are used to train machine learning algorithms. These algorithms can be accessed through an application programming interface (API) allowing users to extract species observations from new footage and test the performance of the algorithms under different confidence and overlapping thresholds.

Project description

Title: Analysis of spatio-temporal trends in deep water habitat builders (Use case)

Study area description: We piloted our data system to extract data on spatial distribution and relative abundance of habitat-building species from deep-water recordings in a Marine Protected Area, the [Kosterhavets National Park](#) in Sweden. The park contains a highly diverse and unique marine ecosystem that has been under active protection and management since 2009. The seafloor in the deeper waters of the park has direct oceanographic connections to the open Atlantic ocean and hence contains much of the bottom-dwelling fauna, which is otherwise only found in deep oceanic waters (Lavaleye et al. 2009). This fauna includes large habitat-building species (Costello et al. 2005) such as sponges (e.g. *Geodia baretii*, *Phakellia ventilabrum*) and cold-water corals (e.g. *Lophelia pertusa*) as well as other large species which can be easily identified from camera footage (e.g. starfish *Porania pulvillus*, *Crossaster papposus*, *Echinus esculentus*).

Design description: Our data system is divided into three main modules: data management, citizen science, and machine learning with high performance computing (Fig. 1).

- **Module 1: Data management (Anton et al. 2019)**

In the data management module researchers store and process the data in ways that maximises efficiency, convenience, and opportunities for sharing and collaboration. To store and access the raw data we use long term and short term storage servers. The long term storage server, or cold storage, archives large amounts of files that don't need to be accessed frequently. These include recordings from Remotely Operated Vehicles (ROVs) managed by the Tjärnö Marine Biological Laboratory, Sweden. The movies (mp4 and mov formats) are on average 1-2 hours long and are systematically collected from all expeditions since the late 1990s (Fig. 1).

The short term storage server, or hot storage, stores a small proportion of files that are frequently used. Here, we transferred 60 movies from the cold storage to the project-specific short term storage server (Fig. 1). The number of movies we selected was a compromise between selecting a representative sample and efficiently using the limited storage of our server. This "hot server" was Linux-based and hosted by [Chalmers University of Technology](#), Gothenburg. The specifications of this High Performance Computing server consisted of a GTX2080Ti GPU with 2 x 8 core Intel(R) Core(TM) i9-9900 CPU @ 3.10GHz (total 16 cores) and 2GB DDR4 RAM.

We created a SQLite database to link all information related to the movies and the classifications provided by both citizen scientists and machine learning algorithms (Fig. 1). The database has seven interconnected tables (Fig. 2). The "movies", "sites", and "species" tables have project-specific information from the underwater movie metadata, as well as the species choices available for citizen scientists to annotate the clips, retrieved

from Zooniverse. The “agg_annotations_frame” and “agg_annotations_clip” tables contain information related to the annotations provided by citizen scientists. The “subjects” table has information related to the clips and frames uploaded to the Koster Seafloor Observatory. The “model_annotations” table holds information related to the annotations provided by the machine learning algorithms. The database followed the [Darwin Core](#) (DwC) standards to maximise the sharing, use and reuse of open-access biodiversity data.

- **Module 2: Citizen science (Anton et al. 2019)**

In the citizen science module researchers and citizen scientists work together to efficiently and accurately annotate raw data. To identify the species recorded in our footage we have created a citizen science website, the [Koster Seafloor Observatory](#). The Koster Seafloor Observatory is hosted in Zooniverse, the largest citizen science platform in the world. The website contains rich supporting material (e.g. background, tutorials, field guides) and features two workflows that help citizen scientists to correctly classify biological objects in video (workflow 1) and locate these in still images (workflow 2).

Workflow 1 (species identification):

Citizen scientists are presented with 10s clips of underwater footage and need to select at least one of the 27 available choices (Fig. 3). The choices include species of scientific importance, animals grouped at different taxonomic levels (e.g. “gastropods” or “fish”), as well as a few miscellaneous options (“Nothing here”, “Human objects”). If citizen scientists select a species or animal, they also need to specify the number of individuals of the taxon selected and the time (in seconds) when any of the individuals fully appears on the screen. Each clip is annotated by at least three different citizen scientists. If the annotations provided by the different citizen scientists match, the clip is “retired” from the website, meaning the clip is not displayed to new users anymore. If the annotations did not match, the clip is annotated by more citizen scientists (up to nine citizen scientists per clip).

Workflow 2 (object location):

Citizen scientists are presented with a still image of a species of interest. To annotate the image, citizen scientists need to draw rectangles around the individuals of the species of interest (Fig. 4). If citizen scientists can not identify any individual of the species of interest in the frame they will not draw any rectangle. Each still image is annotated by five different citizen scientists before it is “retired” from the website.

We used a four-stage image processing framework to upload clips and still images to the Koster Seafloor Observatory and download the annotations provided by citizen scientists (Fig. 5).

Stage 1: Generate and upload clips (Fig. 5, circle a). In this stage we split the +1 hour long movies into 10s clips. After the clips were created, we randomly selected 5,702 clips from the original 60 movies and uploaded them to the workflow 1 of the Koster Seafloor Observatory.

Stage 2: Process clip annotations (Fig. 5, circle b). We retrieved the annotations provided by citizen scientists in workflow 1 and aggregated them on a per-clip basis. To aggregate the workflow 1 annotations, we grouped the annotations each clip received and retained only those choices that were selected by at least 80% of the citizen scientists that annotated the clip (Table 1). In our study, there were 194 clips for which cold-water coral was identified at least by 80% of the citizen scientists. We also averaged the answers from citizen scientists to the question "When is the first time the species appears fully in the video?".

Stage 3: Generate and upload frames (Fig. 5, circle c). We extracted up to three frames per clip from the 194 clips containing cold-water corals, and extracted one frame per second after the first time the species fully appeared in the clip. After extracting 533 frames, we then uploaded them to the workflow 2 of the Koster Seafloor Observatory.

Stage 4: Process frame annotations (Fig. 5, circle d). We retrieved the workflow 2 annotations provided by citizen scientists and aggregated them on a per-frame basis. To aggregate the workflow 2 annotations, we retained the area of overlapping between those rectangles drawn by 80% of the citizen scientists who annotated the frame. Once we aggregated the annotations, we formatted them appropriately to train YOLOv3 algorithms Redmon and Farhadi 2018 (Table 2)

- **Module 3: Machine learning and High Performance Computing (Germishuys et al. 2019)**

In the machine learning and High Performance Computing module researchers train, test and expose state-of-the-art machine learning models. The aggregated citizen scientist annotations are used to train object detection models in tracking and identifying the species of interest. In our case study we used 544 user-annotated ground-truth frames obtained from workflow 2 (Suppl. material 1) to train an algorithm to identify deep water corals (*Lophelia pertusa*). We augmented this data by using a frame tracker which filled subsequent movie frames with bounding boxes with the highest probability of containing the object of interest. This typically increased the amount of data by a factor of 10. The frames were then pre-processed to remove background distortion (as far as possible), since colours often lose intensity underwater, mainly due to poor visibility. Three datasets were then created, one for training the model, another for validation (which is used to tune the model hyperparameters) and finally a testing set. Once the data were prepared, the model training was done until satisfactory metrics were achieved on evaluation measures including F1, Recall, Precision and mAP@0.5 (Table 3). The trained model could then be used to detect species of interest in new footage, whether it be from a webcam, video or a still image.

We made the trained model available through an application programming interface (API), where it can be used by researchers to run predictions of the species of interest in new recordings (Fig. 1). To this end we used FastAPI (Ramírez 2020) as it provides the speed, scalability and reliability required to have multiple users making use of the service at the

same time. The API was also supplied with a user-friendly front-end, using the Streamlit (Teixeira 2020) framework, allowing a broader audience of scientific users (i.e. ecologists, ROV and AUV-pilots, students) to access the service through a web application. The interface allows researchers to browse through already-classified footage, or upload their own footage as either images or video. Once the media has been uploaded/selected, users are able to manipulate hyperparameter thresholds (IOU threshold, confidence threshold) and interactively see the impact on the model output. The API is described by Germishuys et al. (2019).

The last component of this module is a data visualization toolkit that enables researchers to explore and visualise the ecological data extracted from the outputs of the machine learning model (Fig. 1). In our case study, we tested the ability of the model to estimate the relative abundance of cold-water corals over time. To this end, we created 20 standardized recordings (i.e. same length) from a revisited area in the Kosterhavets National Park during the years 2000-2015. These movies were analysed with the trained model to obtain a timeline for the relative abundance of coral heads in the area before and after the national park was established (Fig. 6).

Discussion

The system described here has been tested in a scientific case study, estimating abundance levels of the habitat builder *Lophelia pertusa* in a Marine Protected Area across a period of 15 years (Fig. 6). Results show a steep decline of this key ecological species in the national park during the investigated period. We also found no indication of recovery of corals as a consequence of the establishment of the national park in 2009. These results suggest that physical protection alone may not lead to recovery of the coral stocks, and that external climatic pressures, changes in water quality, and oceanographic connectivity may strongly impact the coral populations in the park.

Our scientific case study exemplifies how the presented system can be used to extract ecological data on abundance and distribution for many benthic species from underwater recordings. Underwater recordings are routinely collected by research institutes, which may allow for a concerted analysis of such data over broad spatial and temporal scales in the future. Such analyses may calculate data products for biological state variables on regional or even global level, so-called Essential Biodiversity Variables or EBVs (Pereira et al. 2013, Hardisty et al. 2019). A recent study by Kissling et al. (2018) suggests that image-based sensor networks are promising candidates for EBVs. Our case study provides empirical support that these data products can indeed be derived from image-based sensors, and importantly from marine environments which are particularly difficult to access and survey. The case study also shows that hincasts of species occurrence and abundance are possible if archived video material is available.

In order to scale up analysis of underwater imagery in the future to extract ecological data for larger regions, longer time periods, and more species several technical bottlenecks have to be addressed. Currently, underwater recordings are locally archived and often can not be discovered. We suggest making underwater recordings discoverable by publishing

metadata in national and/or European archives and data portals (e.g. [European Marine Data Archive](#), [EMODnet portal](#)). Another important technical bottleneck is the disconnection between many essential data services that need to interact to successfully analyse image data. We suggest that seamless links should be developed especially between citizen science platforms (for training of machine learning models) and high-performance computation services (for extracting ecological data from large amounts of imagery). In this development national and European research infrastructures should take a leading role.

Funding: The project was funded by Ocean Data Factory, an expert network supported by grants from Sweden's innovation agency (grant agreement no. 2019-02256), the Swedish Agency for Marine and Water Management (grant agreement no. 956-19), and the Swedish Research Council (through Swedish LifeWatch grant agreement no. 829-2009-6278). The presented work was furthermore supported by the [NeIC](#) program DeepDive and the Horizon 2020 project ENVRIplus (grant agreement no. 654182).

Web location (URIs)

Homepage: <https://www.zooniverse.org/projects/victorav/the-koster-seafloor-observatory/about/results>

Usage licence

Usage licence: Creative Commons Public Domain Waiver (CC-Zero)

IP rights notes: The system is open for use in research, as well as public and academic education for analysis of community composition in marine ecosystems.

Acknowledgements

We thank the data providers who allowed us using movie material. These include ROV-pilots at Tjärnö Marine Laboratory (especially Tomas Lundälv, Lisbeth Jonsson, and Roger A. Johansson) as well as the data curators at University of Gothenburg (Lars Ove Loo, Per Bergström) and Chalmers Technical University (Ola Benderius). We acknowledge the tremendous help from taxonomic experts Thomas Dahlgren, Kennet Lundin, and Björn Källström who actively curated the citizen science platform as well as the ROV pilots who offered their material for use (Tomas Lundälv, Lisbeth Jonsson, and Roger A. Johansson), while Emil Burman helped with the translation of the site. Finally we are grateful for support by the Center for Sea and Society and the Gothenburg Global Biodiversity Center.

Author contributions

VA, MO, and JG conceived and designed the study. VA and MO set up and continue to maintain the Zooniverse site. VA and JG wrote the code for both Github projects (data

processing workflow and model). MO worked with the public contributions to the Zooniverse site. All three authors contributed equally to the writing and revision of the manuscript.

References

- Anton V, Germishuys J, Obst M (2019) Koster Data Management Repository. 0.1. Github. URL: https://github.com/ocean-data-factory-sweden/koster_data_management
- Bean TP, Greenwood N, Beckett R, Biermann L, Bignell JP, Brant JL, Copp GH, Devlin MJ, Dye S, Feist SW, Fernand L, Foden D, Hyder K, Jenkins CM, van der Kooij J, Kroger S, Kupschus S, Leech C, Leonard KS, Lynam CP, Lyons BP, Maes T, Nicolaus EEM, Malcolm SJ, McIlwaine P, Merchant ND, Paltriguera L, Pearce DJ, Pitois SG, Stebbing PD, Townhill B, Ware S, Williams O, Righton D, et al. (2017) A review of the tools used for marine monitoring in the UK: Combining historic and contemporary methods with modeling and socioeconomics to fulfill legislative needs and scientific ambitions. *Frontiers in Marine Science* 4 <https://doi.org/10.3389/fmars.2017.00263>
- Benedetti-Cecchi L, Crowe T, Boehme L, Boero F, Christensen A, Grémare A, Hernandez F, Kromkamp JC, Nogueira García E, Petihakis G, Robidart J, Sousa Pinto I, Zingone A (2018) Strengthening Europe's capability in biological ocean observations. European Marine Board Future Science Brief 3, Ostend, Belgium. [ISBN 9789492043559]
- Costello MJ, McCrea M, Freiwald A, Lundalv T, Jonsson L, Bett BJ, van Weering TCE, de Haas H, Roberts JM, Allen D (2005) Role of cold-water *Lophelia pertusa* coral reefs as fish habitat in the NE Atlantic. Springer, 34 pp. [ISBN 3-540-24136-1] https://doi.org/10.1007/3-540-27673-4_41
- Danovaro R, Carugati L, Berzano M, Cahill AE, Carvalho S, Chenuil A, Corinaldesi C, Cristina S, David R, Dell'Anno A, Dzhenbekova N, Garces E, Gasol JM, Goela P, Feral JP, Ferrera I, Forster RM, Kurekin AA, Rastelli E, Marinova V, Miller PI, Moncheva S, Newton A, Pearman JK, Pitois SG, Rene A, Rodriguez-Ezpeleta N, Saggiomo V, Simis SGH, Stefanova K, Wilson C, Lo Martire M, Greco S, Cochrane SKJ, Mangoni O, Borja A (2016) Implementing and innovating marine monitoring approaches for assessing marine environmental status. *Frontiers in Marine Science* 3: 21. <https://doi.org/10.3389/fmars.2016.00213>
- Germishuys J, Anton V, Obst M (2019) Koster YOLO object detection model. 0.1. Github. URL: https://github.com/ocean-data-factory-sweden/koster_ml
- Guidi L, Fernández-Guerra A, Bakker D, Canchaya C, Curry E, Foglini F, Irission JO, Malde K, Marshall CT, Obst M, Ribeiro RP, Tjiputra J (2020) Big data in marine science. Future Science Brief 6. European Marine Board, Ostend, Belgium. [ISBN 9789492043931] <https://doi.org/10.5281/zenodo.3755793>
- Hardisty AR, Belbin L, Hobern D, McGeoch MA, Pirzl R, Williams KJ, Kissling WD (2019) Research infrastructure challenges in preparing essential biodiversity variables data products for alien invasive species. *Environmental Research Letters* 14 (2). <https://doi.org/10.1088/1748-9326/aaf5db>
- Kissling WD, Ahumada JA, Bowser A, Fernandez M, Fernandez N, Garcia EA, Guralnick RP, Isaac NJB, Kelling S, Los W, McRae L, Mihoub JB, Obst M, Santamaria M, Skidmore AK, Williams KJ, Agosti D, Amariles D, Arvanitidis C, Bastin L, De Leo F,

- Egloff W, Elith J, Hobern D, Martin D, Pereira HM, Pesole G, Peterseil J, Saarenmaa H, Schigel D, Schmeller DS, Segata N, Turak E, Uhler PF, Wee B, Hardisty AR (2018) Building Essential Biodiversity Variables (EBVs) of species distribution and abundance at a global scale. *Biological Reviews* 93 (1): 600-625. <https://doi.org/10.1111/brv.12359>
- Lavaleye M, Duineveld G, Lundalv T, White M, Guihen D, Kiriakoulakis K, Wolff GA (2009) Cold-water corals in the Tilser reef. Preliminary observations on the dynamic reef environment. *Oceanography* 22 (1): 76-84. <https://doi.org/10.5670/oceanog.2009.08>
 - M P, S C, J-O I (2017) EcoTaxa, a tool for the taxonomic classification of images. <http://ecotaxa.obs-vlfr.fr>.
 - Pereira HM, Ferrier S, Walters M, Geller GN, Jongman RHG, Scholes RJ, Bruford MW, Brummitt N, Butchart SHM, Cardoso AC, Coops NC, Dulloo E, Faith DP, Freyhof J, Gregory RD, Heip C, Hoft R, Hurr G, Jetz W, Karp DS, McGeoch MA, Obura D, Onoda Y, Pettorelli N, Reyers B, Sayre R, Scharlemann JPW, Stuart SN, Turak E, Walpole M, Wegmann M (2013) Essential Biodiversity Variables. *Science* 339 (6117): 277-278. <https://doi.org/10.1126/science.1229931>
 - Ramírez S, et al. (2020) FastAPI framework. 0.61.1. Github. URL: <https://github.com/tiangolo/fastapi>
 - Redmon J, Farhadi A (2018) YOLOv3: An Incremental Improvement. arXiv URL: <https://arxiv.org/abs/1804.02767>
 - Teixeira T, et al. (2020) Streamlit. 0.69.0. Github. URL: <https://github.com/streamlit/streamlit>
 - Williams ID, Couch CS, Beijbom O, Oliver TA, Vargas-Angel B, Schumacher BD, Brainard RE (2019) Leveraging automated image analysis tools to transform our capacity to assess status and trends of coral reefs. *Frontiers in Marine Science* 6 <https://doi.org/10.3389/fmars.2019.00222>

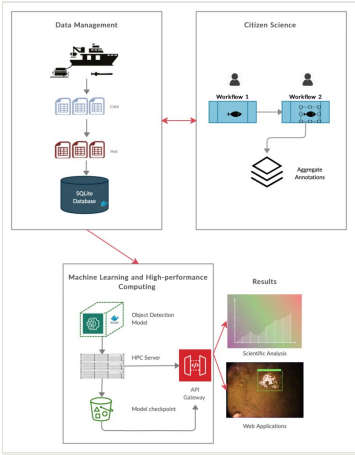


Figure 1.
High-level overview of Koster Seafloor Observatory process.

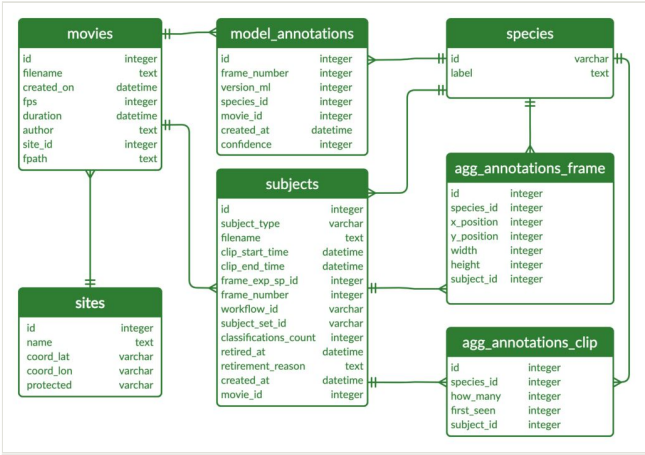


Figure 2.
Entity relationship diagram of the SQLite database used by the Koster Seafloor Observatory.

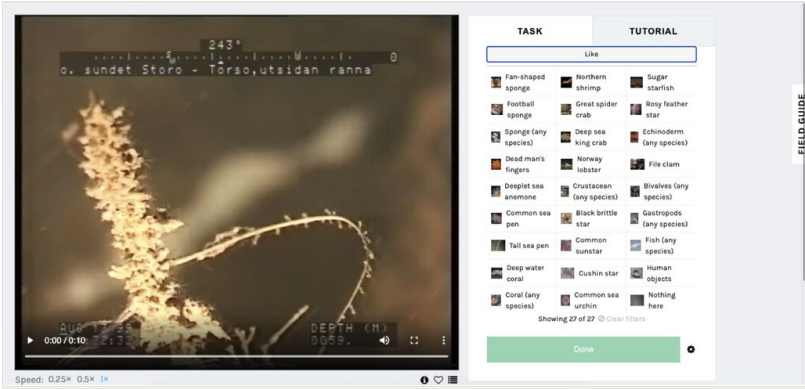


Figure 3. Screenshot of the Zooniverse annotation interface. On the left, display of the clips. On the right, species choices available.



Figure 4.
Example of a frame containing deep water coral displayed to the citizen scientists (left) and the same frame with annotated rectangles provided by a citizen scientist (right).

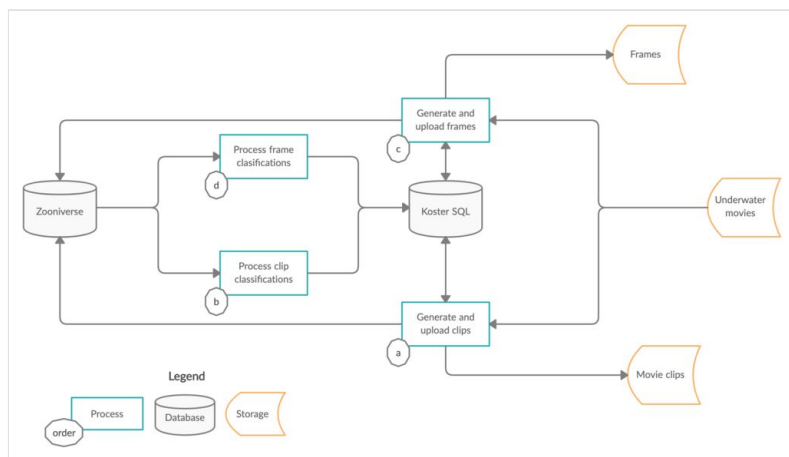


Figure 5.

Four-stage image processing framework used to identify species of interest.

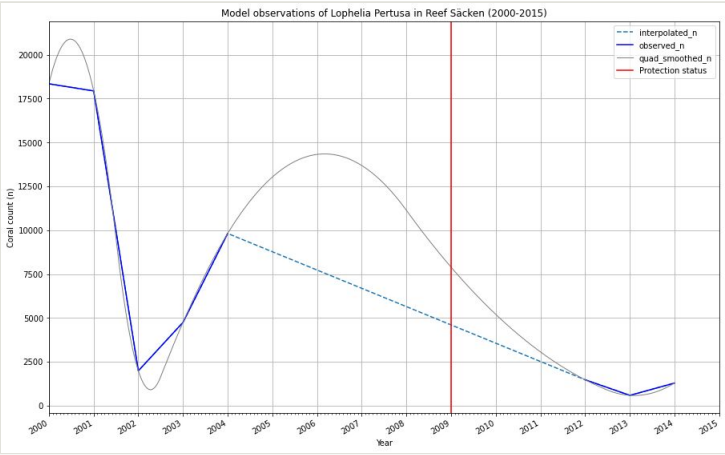


Figure 6. Model observations of *Lophelia pertusa* in the reef area Säcken between 2000-2015, where n denotes the number of instances of *Lophelia Pertusa* (as a count). Observed counts are shown by the thick blue line, along with linearly interpolated values between 2004 and 2012 (shown as a dashed blue line), and quadratic smoothed fitted to the data (shown in gray). Protection status refers to the date at which Kosterhavets National Park was granted protection status and serves as a reference point for monitoring purposes.

Table 1.
Clip Annotation Summary for cold water corals (*Lophelia pertusa*).

Number of classified clips	Median number of users per clip	Median agreement proportion per clip	Number of clips with agreement proportion exceeding 0.8
1945	9	0.44	194

Table 2.
Frame Annotation Summary (*Lophelia pertusa*).

Number of classified frames	Median number of classifications per frame	Number of frames with agreement proportion exceeding 0.8
533	5	409

Table 3.
Classification metrics: Lophelia Pertusa.

Precision	Recall	mAP@0.5	F1 Score
0.979	0.962	0.962	0.970

Supplementary material

Suppl. material 1: *Lophelia pertusa* Dataset

Authors: Victor Anton, Jannes Germishuys, Matthias Obst

Data type: images

Brief description: Instances of *Lophelia pertusa* used to train Koster YOLO machine learning model

[Download file](#) (33.00 MB)