

## RESEARCH ARTICLE

# A low-cost, long-term underwater camera trap network coupled with deep residual learning image analysis

Stephanie M. Bilodeau<sup>1,2\*</sup>, Austin W. H. Schwartz<sup>1,2</sup>, Binfeng Xu<sup>3</sup>, V. Paúl Pauca<sup>3</sup>, Miles R. Silman<sup>1,2</sup>

**1** Department of Biology, Wake Forest University, Winston-Salem, NC, United States of America, **2** Center for Energy, Environment, and Sustainability, Wake Forest University, Winston-Salem, NC, United States of America, **3** Department of Computer Science, Wake Forest University, Winston-Salem, NC, United States of America

\* Current address: Georgia Institute of Technology, School of Biological Sciences, Atlanta, GA, United States of America

\* [sbilodeau@gatech.edu](mailto:sbilodeau@gatech.edu)



## OPEN ACCESS

**Citation:** Bilodeau SM, Schwartz AWH, Xu B, Paúl Pauca V, Silman MR (2022) A low-cost, long-term underwater camera trap network coupled with deep residual learning image analysis. PLoS ONE 17(2): e0263377. <https://doi.org/10.1371/journal.pone.0263377>

**Editor:** Fraser Andrew Januchowski-Hartley, Swansea University, UNITED KINGDOM

**Received:** March 31, 2021

**Accepted:** January 18, 2022

**Published:** February 2, 2022

**Copyright:** © 2022 Bilodeau et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** This image set is available on Harvard Dataverse (<https://doi.org/10.7910/DVN/5IOMOW>). Code from this project can be accessed on GitHub (<https://github.com/silmanmr/PONE-D-21-10609>).

**Funding:** This work was supported by grant #2018-007 to MRS and SMB from the Wake Forest University Center for Energy, Environment, and Sustainability (<https://cees.wfu.edu>), a Wake Forest University Richter Scholarship (<https://ureca.wfu.edu>), Sullivan Scholarship (<https://biology.wfu.edu>).

## Abstract

Understanding long-term trends in marine ecosystems requires accurate and repeatable counts of fishes and other aquatic organisms on spatial and temporal scales that are difficult or impossible to achieve with diver-based surveys. Long-term, spatially distributed cameras, like those used in terrestrial camera trapping, have not been successfully applied in marine systems due to limitations of the aquatic environment. Here, we develop methodology for a system of low-cost, long-term camera traps (**Dispersed Environment Aquatic Cameras**), deployable over large spatial scales in remote marine environments. We use machine learning to classify the large volume of images collected by the cameras. We present a case study of these combined techniques' use by addressing fish movement and feeding behavior related to halos, a well-documented benthic pattern in shallow tropical reefs. Cameras proved able to function continuously underwater at deployed depths (up to 7 m, with later versions deployed to 40 m) with no maintenance or monitoring for over five months and collected a total of over 100,000 images in time-lapse mode (by 15 minutes) during daylight hours. Our ResNet-50-based deep learning model achieved 92.5% overall accuracy in sorting images with and without fishes, and diver surveys revealed that the camera images accurately represented local fish communities. The cameras and machine learning classification represent the first successful method for broad-scale underwater camera trap deployment, and our case study demonstrates the cameras' potential for addressing questions of marine animal behavior, distributions, and large-scale spatial patterns.

## Introduction

Understanding species' habitat use and biodiversity is central to ecology and conservation of marine and terrestrial ecosystems, though these data are often some of the most difficult to

edu), and Undergraduate Research Fellowship (<https://ureca.wfu.edu/>) to AWHS, and a Wake Forest University Vecellio Grant (<https://biology.wfu.edu/>) to SMB. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

gather. Terrestrial camera trapping is a growing field and a technique widely applied in global biodiversity monitoring [1]. Camera traps are remotely activated cameras that usually trigger based on the difference between background radiation and a warm-bodied animal passing through the sensor's field. They have been used primarily to study mammals and birds, although the field is now expanding to include some ectotherms [2].

Camera traps are used to study species richness [2] and the distribution, abundance, habitat use, and behavior of wildlife around the world, with many studies surveying more than one species at a time [3]. Most camera traps are small, relatively inexpensive, and often deployed in groups or networks over a wide area for months at a time. They are typically less invasive and more reliable than comparable observation techniques [4].

Analogous techniques for monitoring underwater species face several operational challenges, chiefly the attenuation of infrared radiation in water, which renders a standard commercial camera trap unlikely to trigger underwater, except at very close range [5], and the rigors of operating in the marine environment where water intrusion and algal and faunal fouling are persistent issues, especially in nearshore systems [6]. The fact that most fishes are ectotherms further complicates use of the traditional heat-triggered infrared sensor technology used in most terrestrial camera traps. Far-red illumination invisible to most fishes provides one potential alternative to infrared [7], although far-red light still attenuates over short distances underwater. Given the ability of sound to propagate well underwater, acoustic techniques provide another possible alternative to infrared sensing [5]. Acoustic cameras [8,9] have been used in the past to image sharks and other fishes in low-light, turbid environments, replacing light-based imaging entirely [10] for species that are morphologically distinct enough to be differentiated in this manner [11]. This suggests that specific acoustic cues could also be used to trigger conventional optical cameras [12], although this would have limited applications outside of underwater camera trapping.

Due to the power requirements of active triggering (far-red light, sonar) and recording methods, most current underwater fish monitoring and measurement techniques are either limited by short battery life and operate on the scale of hours, as with baited remote underwater video (BRUV) and similar short-term recording devices [7,13–17], or require a tethered, external power source [18,19]. Because of this, the spatial extent, number of cameras, and duration of monitoring for marine systems are vastly smaller in scope than terrestrial efforts (compare the work of Williams et al. [7] and Siddiqui et al. [20] to the TEAM [21] or Snapshot Serengeti [22] datasets). Even long-term underwater monitoring with an external power supply may be limited to recording images or video during daylight hours [18] due to the difficulties of avoiding reflected particulate matter in underwater images taken at night with direct illumination. Without a side-mounted flash or otherwise indirect lighting, both white light and infrared images may be obscured by the illumination of biotic and abiotic particulates suspended in the water column, although optical backscatter filters and other multispectral image processing techniques can improve image quality in low-light and/or turbid environments [23,24]. Continuous illumination may also attract fishes and other marine organisms, depending on the color of the light [25]. Thus, long-term underwater observations are limited with regard to both power and nighttime illumination, and available solutions (e.g., batteries, specialty lighting rigs) are expensive. Both monetary costs and logistics can limit deployment of underwater cameras in remote locations without access to external power or consistent upkeep.

A related challenge faced by both terrestrial and marine camera traps is the time cost related to processing and analyzing large photosets obtained from multiple cameras over the course of months or years [22]. Applications of computer vision are growing in the field of camera trapping. They have the potential to reduce time cost [2] and have already been successfully applied to the extensive Snapshot Serengeti camera trap dataset [22] as well as images and video frames

from cabled observatories or short-term marine cameras [18–20,26–28], long-term GUARD1 cameras on marine floats [29,30], and autonomous underwater vehicles [31].

Here we present a simple design for an affordable, long-running, autonomous underwater camera based on an existing commercially-available terrestrial model. We outline the deployment of our Dispersed Environment Aquatic Cameras (DEACs) across a 270 km<sup>2</sup> tropical reefscape, our testing, and our subsequent analysis of the 100,000+ images obtained by implementing a deep convolutional neural network (CNN) technique for image classification. To demonstrate both the efficacy of our design and the value of long-term unmanned underwater observations to marine ecology research, we present a case study in fish feeding behavior as it relates to a well-documented benthic pattern at Lighthouse Reef Atoll, Belize of sand halos around coral heads [32,33].

Halos consist of a bare sand or lightly vegetated border surrounding a coral patch or similar underwater structure that separates the reef from surrounding dense vegetation (i.e., seagrass or algae). The heightened grazing observed inside these halos could be due to a landscape of fear [34], where herbivores are afraid to venture past a threshold distance from the reef due to predation risk [35]. However, this threshold of fish density could also be due simply to the natural dispersion of grazers and other fishes as they venture farther from the reef, which serves as a central aggregating structure for many species [36–38]. We proposed that if a strong landscape of fear is in effect, herbivorous species observed and photographed in the halo will never venture out into the surrounding seagrass, except perhaps when traveling in schools. However, a simple drop-off in fish density with distance would still result in the occasional grazing reef fish being seen by our cameras, and if reef fishes are food limited rather than predator limited, they should forage widely in the seagrass, which is a preferred food [39]. Therefore, regular detection of reef herbivores out in the seagrass over multiple months could be taken as evidence refuting the landscape of fear at Lighthouse Reef.

## Methods

### Design

**Camera selection.** Cuddeback Silver Series scouting cameras, model 1231 (Cuddeback, Green Bay, WI, USA), were chosen for their compact size, their relatively high 20 megapixel (MP) image resolution, and their time-lapse function, which allows images to be taken at pre-programmed time intervals and certain light levels without requiring the use of additional video or motion-triggered image settings. Since infrared motion triggering does not function well in the underwater environment, the presence of a time-lapse option was critical and is increasingly available in terrestrial trail cams. The cameras were programmed to take one 20 MP image every 15 minutes whenever ambient light levels were high enough to allow for color photography without flash, using the “Day” setting. Nighttime images and video and all infrared sensor-triggered images and video were disabled to conserve both power and memory space because initial field tests showed that few or no additional usable images were captured using the infrared, low-light “Night” setting.

All cameras deployed throughout the study period were synchronized to record photos on the hour and every 15 minutes following to ensure that images from different cameras and sites were captured at the same time of day and under the same local conditions. The 15-minute interval was chosen to balance the need for regular observations of reef and seagrass communities that may include transient fish species and the constraints of storing and processing thousands of high-resolution images collected by multiple cameras over a months-long deployment. The ability of this interval to adequately capture the community composition and species present at a given reef was validated using in-person diver surveys (described below).

Memory cards were 32 GB SanDisk (Western Digital Corporation, Milpitas, CA, USA) or Kingston (Kingston Technology Corporation, Fountain Valley, CA, USA) microSD cards with adapters, capable of holding over 25,000 20 MP color images each. Each camera required eight Energizer Ultimate Lithium AA batteries (Energizer Holdings, Inc., St. Louis, MO, USA), which provided enough power for over five months of continuous function under the settings described here.

The total cost of each camera, including the batteries and SD card, came to just \$125 per unit with the housing (discussed below). The use of pre-built commercial trail cameras, which are designed for energy efficiency over long deployments, significantly reduced both material and energy costs, relative to constructing a similar camera from scratch using components like a Raspberry Pi (Raspberry Pi Foundation, Cambridgeshire, UK) and GoPro (GoPro, Inc., San Mateo, CA, USA), Canon (Canon, Inc., Ota City, Tokyo, Japan), or Sony (Sony Corporation, Minato City, Tokyo, Japan) cameras, as used in previous underwater camera applications [7,15,16,20,26].

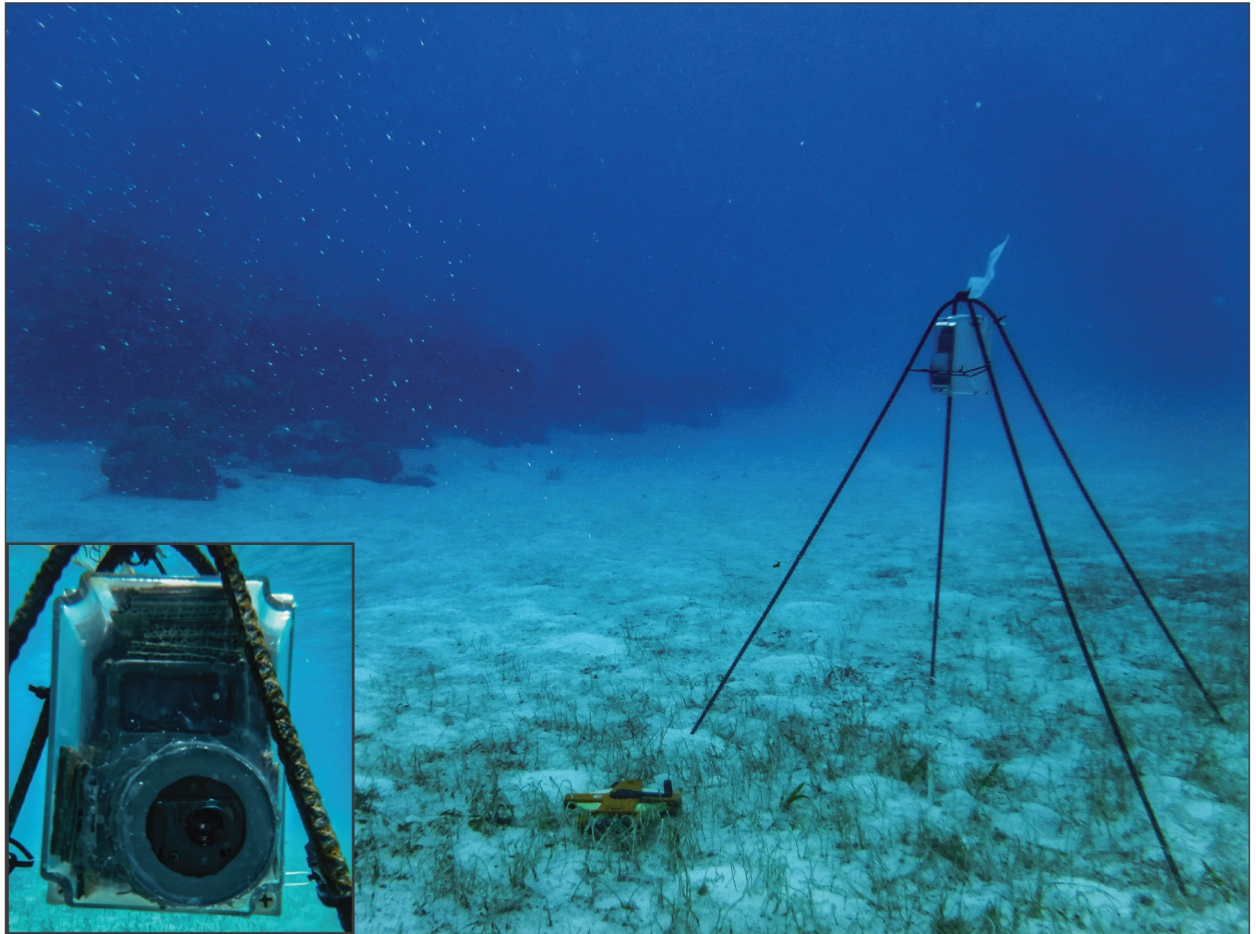
**Housing construction.** Two housings of different materials were tested in the field, both based on commercially available junction box enclosures. Housing 1, based on item DS-AT-1217-1 available from “Saipwell” (Saip Electric Group Co., Ltd, Wenzhou, China), has thinner walls (2.4–3.9 mm) made of an unspecified plastic. The top secures with specially-shaped plastic screws and it contains an O-ring-like insert made of foam. Housing 2, based on model ML-47F\* 1508 from Polycase, Inc. (Avon, OH, USA), is a thick-walled (3.5–4.0 mm) design made of polycarbonate resin secured with stainless steel screws and a silicon rubber gasket. Both housings had a 2-inch diameter hole drilled in the faceplate and a 3-inch disk of 1/8 inch thick acrylic epoxied to the opening with MarineWeld (J-B Weld Company, Atlanta, GA, USA) to act as a window. Since both housings used identical acrylic windows and contained the same cameras, images collected using each design were indistinguishable; the chief difference between the housings was their pressure tolerance and leakage at depth. Each housing cost approximately \$30 for all components (included in the total price given above). Ablative anti-fouling boat paint was applied to the housing exterior, excluding the back and acrylic window, of a subset of cameras with Housing 2 to reduce biofouling.

Cameras were programmed, armed, and packed inside their housings with cardboard spacers. Housings were sealed with Star brite marine silicone sealant (Star brite, Fort Lauderdale, FL, USA) around the seam of the enclosure lid. Silicone sealant was also used to reinforce the edges of the epoxy seal around the lens window, both inside and outside.

**Field installation.** Cameras were secured to four-legged bent rebar stands with plastic cable ties threaded through mounting holes pre-built into the housings (Fig 1). Due to their buoyancy, camera housings were placed underneath the crossed rebar forming the top of each stand and secured laterally to the four legs, which were sunk into the sediment to keep the stands upright. Lightweight plastic or polystyrene buoys were tethered to a small subset of cameras and stands located in particularly shallow water (2 m or less) to prevent collision by boats.

## Deployment

**Study system.** Lighthouse Reef Atoll off the coast of Belize is primarily a shallow (1–8 m depth) lagoon environment, dominated by scattered patch reefs interspersed with a mosaic of seagrass, macroalgae, and sand [40]. Although relatively isolated from the mainland, the atoll is subject to heavy local fishing pressure for conch, lobster, and certain fish species. A small no-take marine protected area (MPA) surrounds the island of Half Moon Caye in the southeastern corner of the atoll. Because of its shallow benthos, wide variety of benthic cover types, and spatial variation in protected status, Lighthouse Reef provides an ideal location to test our DEACs



**Fig 1. DEACs deployed on four-legged rebar stands in the field.** Plastic cable ties attached to the camera housing and to the legs were used to secure buoyant cameras in housings below the crossed rebar at the top of each stand. Bottom left inset: A detailed front view of a camera in Housing 1 underwater. Reprinted from [39] under a CC BY license with permission.

<https://doi.org/10.1371/journal.pone.0263377.g001>

under a variety of conditions. The shallow marine grazing system of Lighthouse Reef and similar Caribbean locations is in many ways analogous to terrestrial grazing systems like the African savanna [41], where camera trap networks have been effectively deployed for years [22].

To structure our testing and demonstrate the utility of the DEACs for addressing ecological questions at large spatial scales, we organized our deployments around detection of herbivorous fishes grazing in coral reef halos at Lighthouse Reef. Grazers at Lighthouse Reef are mostly large parrotfishes (Scaridae) and surgeonfishes (Acanthuridae). These diurnal herbivores readily consume seagrass when it is provided to them within the halo surrounding a reef [39]. We conducted nighttime observations and grazing assays (see [39]) to ascertain that nocturnal grazers like urchins were not particularly active at our sites. We deployed our DEACs in a paired design that allowed us to monitor the reef and halo as well as the surrounding seagrass or macroalgae at multiple sites to determine whether these herbivorous reef fishes regularly venture into the surrounding seagrass to feed and are simply more abundant in the halo due to its proximity to the reef or if they are constrained to the reef and adjacent halo by predation (a landscape of fear).

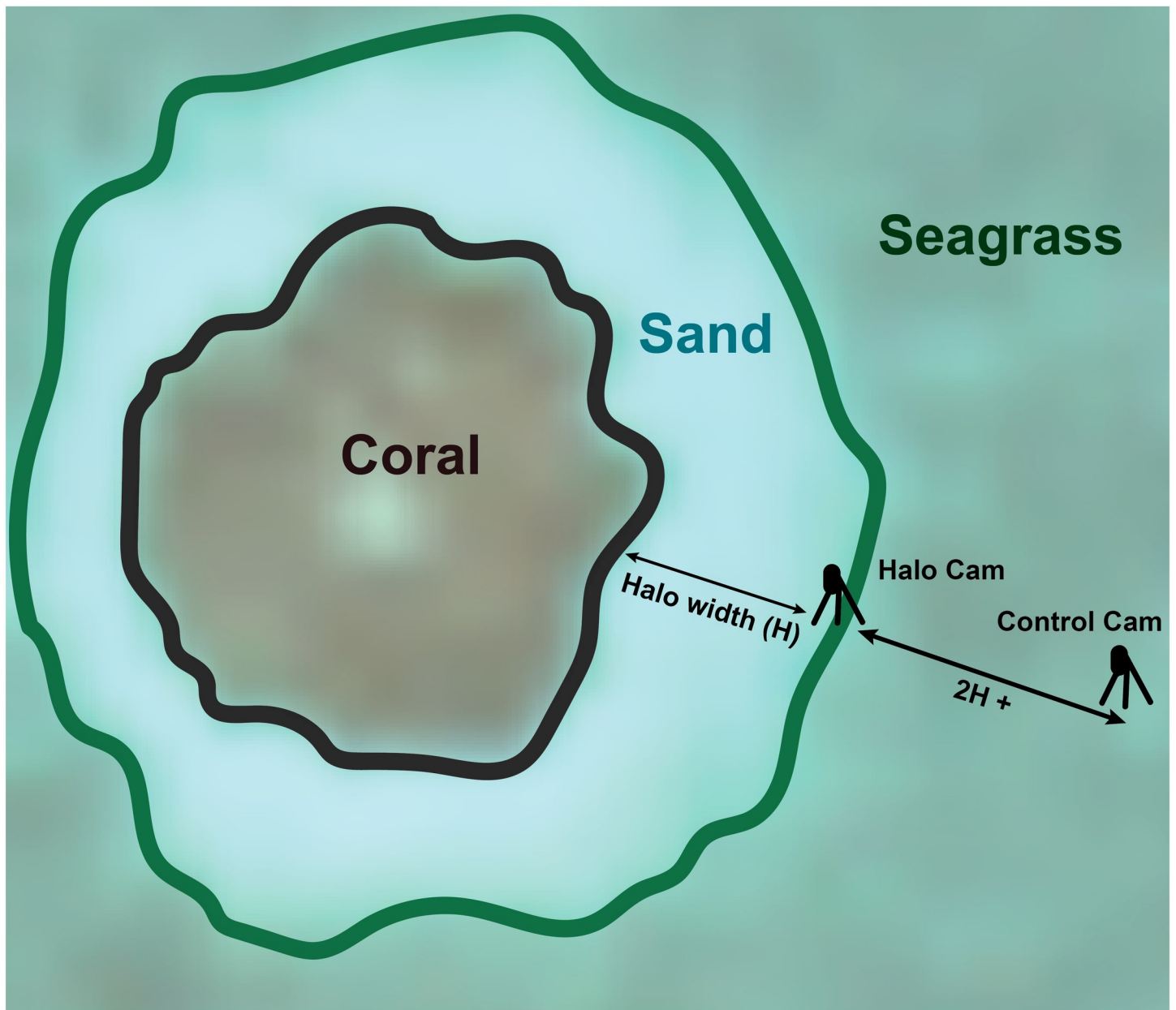
**Spatial arrangement.** Cameras were deployed in pairs at 21 patch reef sites within Lighthouse Reef Atoll. Deployments occurred asynchronously between March 2018 and March

2019 (S1 Fig), and a total of 28 individual cameras were used across all deployments, with some cameras deployed multiple times. The majority of sites were occupied by cameras between June and August 2018, with the longest deployments lasting over 5 months (August 7, 2018 to January 12, 2019). DEAC sites were distributed evenly inside and outside of the Half Moon Caye Natural Monument MPA in the southeastern corner of the atoll. Seven sites (3 inside the MPA, 4 outside) featured predominantly algal bottom cover; the rest were in areas surrounded by seagrass, primarily *Thalassia testudinum*. Patch reef sites for camera deployment were chosen via random point placement using satellite imagery and a depth map of the atoll (courtesy of the Carnegie Airborne Observatory), which allowed sites to be evenly stratified across depths from 2–7 m.

Sites deeper than 4 m were initially avoided due to leakage of Housing 1 past this depth, although depths of up to 7 m (maximum depth required in the study) were successfully achieved with Housing 2, which was used for initial deployments in March 2018 and all deployments from August 2018 onward. Cameras were deployed in sets of two, one camera located at the edge of the sandy halo surrounding a patch reef and the other located at least twice the halo's width away from the edge in the surrounding seagrass or algal benthic cover (Fig 2). In the case of particularly narrow halos, "control" cameras were placed a minimum of 15 m from the halo edge. This allowed control cameras a view of the same bottom cover (seagrass or macroalgae) as that adjacent to the halo but placed them well beyond the fish density thresholds observed by Layman et al. [38], while also accounting for the possibility that larger halos could represent reefs with larger or farther-ranging fish populations. "Halo" cameras had a relatively wide field of view, as is typical of terrestrial trail cameras, which included both the halo in front of them and the patch reef beyond. Both cameras were pointed toward the reef, although the edge of the halo was beyond the range of view for the grass ("control") camera at most sites. Camera orientation with regard to compass direction was varied between sites during deployment, although both cameras at any given location were oriented in the same direction.

**Monitoring and secondary deployments.** Camera deployments occurred in stages, with the first cameras deployed in March 2018 and the last cameras collected in March 2019 (S1 Fig). The longest-running DEACs remained continuously active underwater for five months (see above) and were still operating at retrieval. Initial deployment of cameras in March 2018 included one camera at the edge of a halo and a second camera pointed at the first in order to assess whether the presence of the camera and stand had any effect on the presence or behavior of fishes. No obvious confounding effects of camera presence were noted in this initial test (neither herbivores nor large predators aggregated around the camera structure, the same species were photographed near the camera and in other parts of the halo, etc.), so doubling of cameras in this manner was discontinued for future deployments. For all subsequent deployments (March 2018 onward), the two cameras at each site were positioned to monitor different benthic environments and unable to see each other. Diver observations while cleaning and monitoring cameras in the field over the duration of this study supported our initial conclusion that reef fishes largely ignore cameras, except to occasionally eat algae off the structures, as they do with any other hard substrate in the halo. During the summer of 2018 (June to August), cameras were consistently checked every 1–2 weeks for leakage, algal overgrowth, or stand displacement. Cameras deployed at longer intervals from March to June 2018, August 2018 to January 2019, and January to March 2019 were unmonitored during these periods in order to test long-term underwater function and determine the effects of biofouling on housings and image quality in the absence of cleaning or regular adjustments.

**Diver observations.** In-water observations were conducted by a team of 2–3 divers at each camera site and at several additional locations in order to validate the cameras' ability to detect



**Fig 2. Camera arrangement at a patch reef site.** One camera was placed on the edge of the halo facing in toward the reef (“Halo Cam”), while the second (“Control Cam”) was placed twice the halo’s width away in the surrounding seagrass or algal cover, in line with the first camera. Modified from [39] with permission.

<https://doi.org/10.1371/journal.pone.0263377.g002>

fishes and other animals. Observations consisted of all divers sitting directly behind each camera for 15 minutes and recording the presence and abundance of all fishes and other animals observed on the reef, in the halo, or in the grass to the genus or species level, when possible. Divers faced forward toward the reef and recorded all fishes observed within the halo or on the reef itself, as this was the primary field of view of the camera. In the seagrass or macroalgae, divers again faced toward the reef and recorded only fishes that swam in front of them and the camera. Divers remained stationary for the entire observation period at each camera. Fish species and counts were determined by a consensus of all the divers present at each observation.

Fish counts were recorded by divers as 1 individual, 2–3 fish, or 5 or more (coded as 1, 3, and 5), which could under-represent species for which there were many individuals at a given site. In practice, the majority of fishes observed by divers were solo or in small groups.

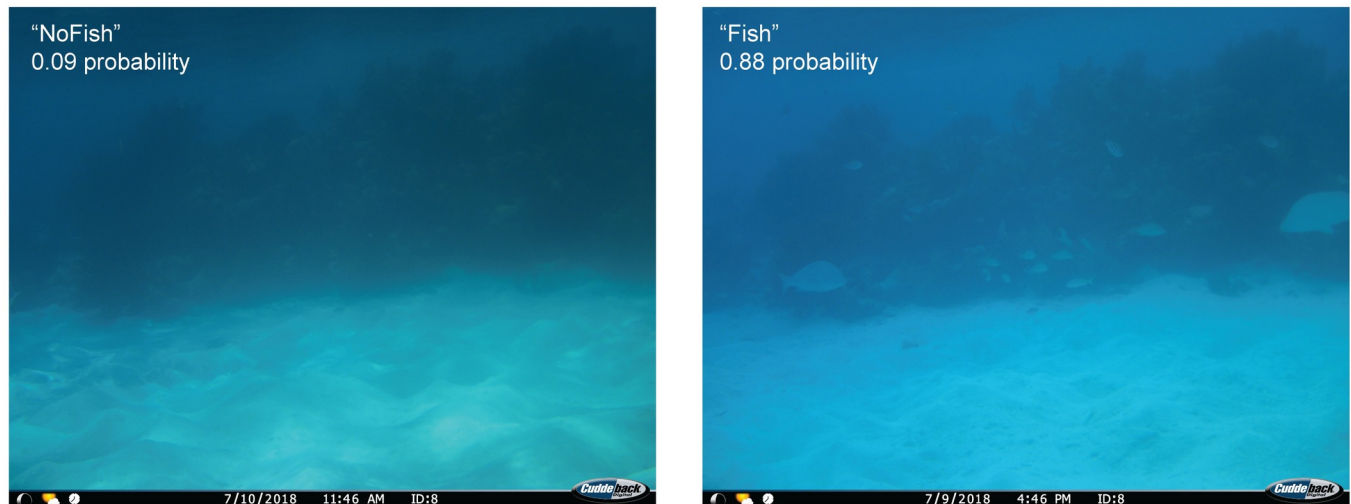
## Data analysis

**Community composition.** A non-metric multidimensional scaling (NMDS) analysis was run on species communities inside and outside of the halo using data from diver observations at 20 camera sites and image analyses from 18 of these cameras. For each camera, 3 images taken before, during, and after diver observation (15-minute intervals) were analyzed, for a total of 54 images. Only images from immediately before, after, and during diver observations were used for comparison in order to control for natural variation in fish communities over time, since the main goal of the comparison was to assess the ability of the cameras to capture known community composition at a given location and time. All fishes in this subset of images were identified to the genus or species level by the same divers who conducted the in-water observations. Species recorded in images from all three time points (before, during, and after divers) were pooled for each site to determine presence or absence. The NMDS used a Sørensen distance matrix calculated through the binary version of the “bray” method in *vegan*’s *vegdist* function [42] and was based on species presence or absence at each site, not abundance, to reduce the chances of “double counting” the same individual fish in successive images and to limit effects of the binned counts recorded by the divers. Fish counts across different habitat types and methods (divers and cameras) were also compared with a two-way PERMANOVA, using the *adonis* function in *vegan* [42]. To determine whether divers had any effect on fish presence, counts from images captured before, during, and after diver observations were compared using a nonparametric Friedman rank sum test. All analyses were conducted using R statistical software version 4.0.3 [43] with packages *qdap* [44], *reshape* [45], *rstatix* [46], and *vegan* [42].

**Image sorting.** Images from the larger dataset encompassing all cameras were initially named and sorted by location, time, and date using the *camtrapR* [47] package in R [43]. A subset of over 13,000 images were sorted by trained undergraduate student volunteers into categories containing at least one visible fish (“Fish”) and without any visible fishes (“NoFish”). To ensure consistency, all volunteers were initially trained on the same subset of ~300 images drawn from multiple different cameras and their accuracy assessed by the research team before they were assigned a larger subset of images to sort individually. Due to the nature of timed rather than motion or heat triggered photos, many images did not contain fishes.

**Model choice.** In order to streamline future analyses of fish photos and reduce the need to manually classify over 80,000 images, we built and trained a CNN based on ImageNet pre-trained ResNet-50 [48], a deep residual network widely used in image classification. Deep CNNs (i.e., those with many layers) are important tools for visual recognition tasks but may be hard to train because of vanishing gradients and degradation. The residual network (“ResNet”) approach solves this by explicitly allowing stacked layers to fit a residual mapping instead of directly fitting a desired underlying mapping, which enables neural network architectures to go deeper [48]. In our application, ResNet allowed us to learn Fish/NoFish representations from a noisy and relatively small dataset. We chose ResNet as our base structure because we wanted conservative results in classifying fish pictures. ResNet is powerful at preventing overfitting, making it less likely to omit pictures containing fishes. The reproducible code was implemented in Pytorch [48–50] to identify images with animals (Fig 3). It took roughly twelve hours to train the ResNet model for 50 epochs on a custom-built desktop PC with a single Nvidia RTX 2080 GPU, Intel i9-9900k CPU, and 32 GB of RAM.





**Fig 3.** Empty (left) and fish-filled (right) images captured at the same camera site within 24 hours. The empty or “NoFish” photo was assigned only a 0.09 probability of containing a fish by our ResNet-50 model, whereas the model predicted a 0.88 probability of the second photo containing at least one fish, hence its “Fish” designation. Probabilities like these were used to sort images into “Fish” and “NoFish” categories to streamline further analysis. Reprinted from [39] under a CC BY license with permission.

<https://doi.org/10.1371/journal.pone.0263377.g003>

**Training and validation.** Our model was trained on a sample of 10,727 images sorted by our team of trained volunteers and then run on 75,470 additional images to sort them into “Fish” and “NoFish” categories, based on the model’s calculated probability of each image containing at least one fish. To augment the training set, we preprocessed the fish images with spatial transformations, cropping and lightening variation, and used focal loss [51] as an objective function to address the problem of imbalanced labels. The model’s accuracy was validated by comparing the categories produced by sorting above and below 0.5 probability with the actual “Fish” and “NoFish” designations assigned to 2,702 images hand-sorted by trained volunteers.

**Image analysis.** Images with a “Fish” probability of 0.5 or above were used to determine the relative presence of fishes at camera sites inside and outside of the Half Moon Caye MPA, in different habitat types (reef/halo vs. seagrass/algae), and at different times of day. Detection is an imperfect proxy for actual fish presence [52], so these comparisons could represent differences in occupancy, changes in behavior, or other unknown factors. This analysis of fish detections was conducted based on the relative proportions of the total images classified that did or did not contain fishes inside and outside of the MPA and at patch reef and seagrass/algae locations, compared using the `prop.test` function in R [43]. The null hypothesis was that both groups in each test should have equal proportions of images containing fishes. Only predictions from images taken within the first month of deployment were used in these analyses to account for the varying degradation in image quality (and resulting increased potential for misclassification) due to biofouling at different sites. Trends in fish detection probability across different times and locations were visualized using `ggplot2` [53].

## Results

### Camera performance in the field

**Cameras and housings.** Housing 1 proved vulnerable to flooding at depths exceeding 4 m and prone to leaking even at shallower depths. This appears to be due primarily to the thin nature of Housing 1’s walls and lid, which deformed substantially under pressure, breaking the epoxy seal with the lens and allowing water to enter. Housing 2 proved far more robust to

pressure and had minimal leakage even at a maximum deployed depth of 7 m over multiple months. Further testing with this housing will be necessary to determine its maximum functional depth, but preliminary tests with a revised housing design have reached depths of 40 m.

The cameras themselves proved resilient to flooding and were typically still armed and taking photos when extracted from partially-flooded housings. However, many cameras recovered from flooded housings were unable to be redeployed due to lasting damage to their internal systems. Batteries were sufficient to power the cameras past the five-month extraction date of our longest deployment, with all non-flooded cameras still displaying the starting battery status of “OK” (not “LOW” or “DEAD”). Maximum water temperature measured at any of our deployment sites, which could affect battery life, was 30° C (86° F). Cuddeback advertises that their cameras can operate up to 12 months continuously with efficient battery usage. Memory cards were more than sufficient to store images on the schedule that they were collected during this period (approximately 8.5 GB of images over 5 months), suggesting that image frequency could be doubled or tripled in future deployments, assuming that processing of these additional images is sufficiently streamlined with the use of computer vision or similar techniques.

**Biofouling.** The most significant impediment to long-term camera function at our sites was biofouling, the growth of marine organisms over the camera housing and stands. Image quality declined rapidly due to biofouling, making fish identification impractical in as little as one month depending on site conditions (S2 Fig), with a median value of two months endurance, although imagery from some cameras remained usable up to four or five months after deployment. This biofouling decreased both fish detections in images and the ability of our model to accurately classify images taken by biofouled cameras (S3 and S4 Figs). Algal grazing by fishes was an important factor reducing fouling on cameras and stands placed near patch reefs, while cameras placed in seagrass or algae beds away from patch reefs were overgrown with algae over the same deployment period that halo cameras remained relatively unobstructed (Fig 4). The addition of antifouling boat paint to the camera housings before the August 2018 deployment appeared to be only a minor deterrent to organisms growing on the



**Fig 4. Comparison of biofouling and algal overgrowth after a month based on location.** Structures were located at the edge of the halo, adjacent to seagrass (left), and in the middle of the surrounding seagrass (right). Similar results seen across all camera sites suggest that reef-based herbivores help to control algal growth both within and at the edges of the halo but do not graze heavily, if at all, on algae and related marine organisms in the surrounding algal or seagrass beds. Reprinted from [39] under a CC BY license with permission.

<https://doi.org/10.1371/journal.pone.0263377.g004>

housing in general and did not prevent the camera lens window from being almost completely obscured 2–3 months after deployment. In addition to biofouling, camera function was also limited by reduced visibility at certain sites, which varied based on local turbidity and light availability, and by depth.

### Image classification and accuracy

Of the 130,621 images collected, 88,899 were deemed usable (68%) based on a visual examination, despite some level of biofouling in many of these. Images were included in the model if the researchers could still visually distinguish the shapes of fishes, the reef, and other underwater features despite the biofouling. The ability to identify individual species was impaired in the most heavily fouled images we included, but this was not a primary goal of our particular model. Since the determination regarding image quality was made by humans, it is possible that the threshold at which we considered an image unusable was not the optimal cut-off point for the processes used by the model to evaluate image content. Our ResNet-50 model had an overall accuracy of 92.5% when classifying these images into “Fish” or “NoFish” categories, with higher accuracy in identifying empty or “NoFish” photos, likely due to the increased number of these in the training dataset (Table 1). The training and validation sets of manually classified images together made up approximately 15% of the total usable images obtained from the cameras.

Accuracy is determined by comparing the computer’s prediction for each image with the actual label of “Fish” (image contains at least one fish or similar animal) or “NoFish” (image is empty) assigned to the photo by a team of trained volunteers. The model was trained on 10,727 images classified by volunteers and then validated with an additional 2,702 human-sorted images, the results of which are shown in this table. Prediction accuracy is higher for “NoFish” (empty) images, likely due to the higher proportion of this type of image in both the training and validation datasets.

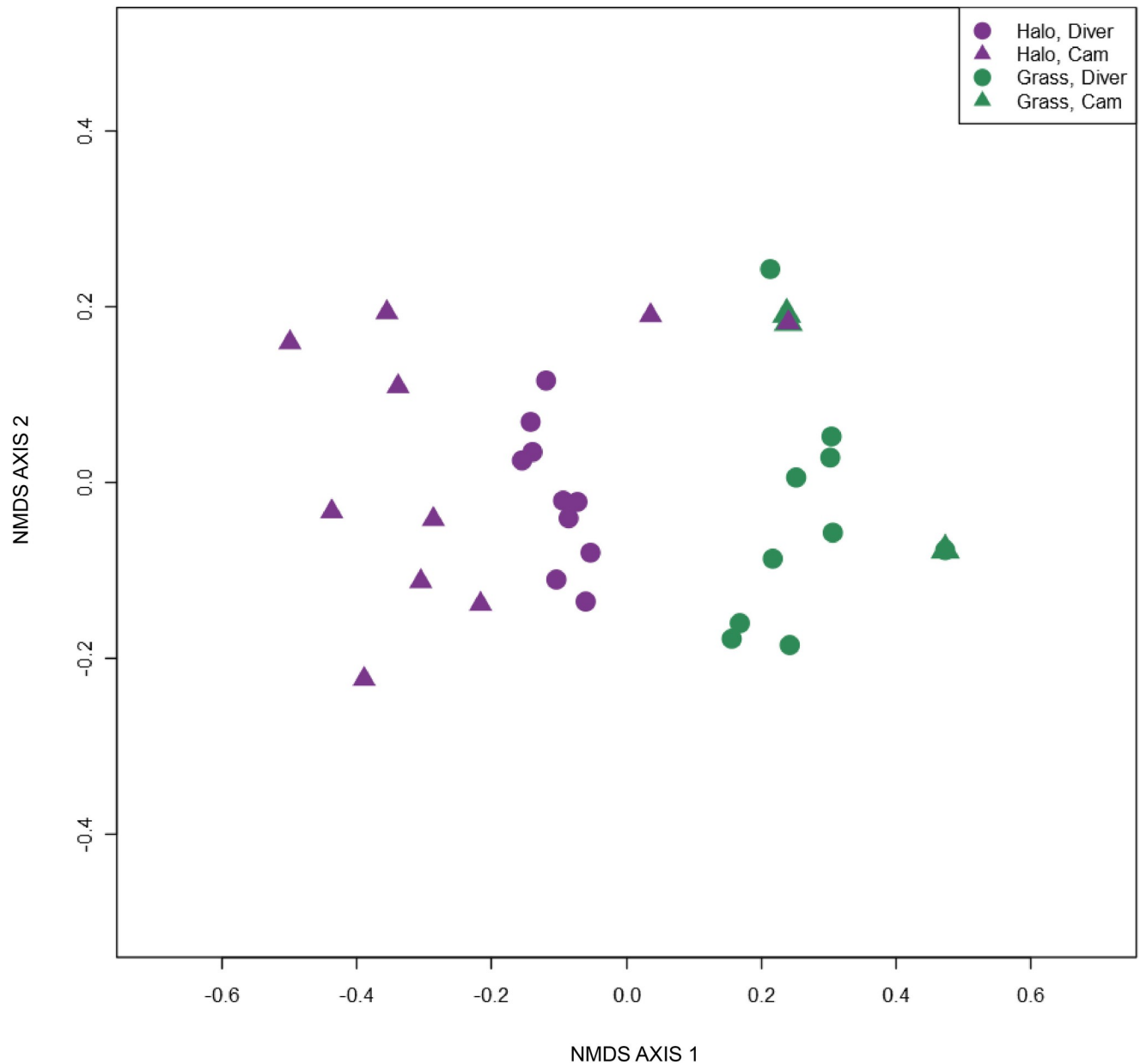
### Community composition

**Comparison with diver observations.** Large compositional differences were identified between the halo/reef and seagrass/algae fish communities, which were reflected in both the diver species observations and those from camera images (NMDS; Fig 5), although each 45-minute set of three images consistently contained fewer fishes than were observed by divers during a 15-minute period at the same site. An ANOSIM conducted on the NMDS output showed a significant difference between the halo/reef and seagrass/algae communities ( $R = 0.68$ ,  $p < 0.001$ ), which was reflected in both the camera and diver species observations. The PERMANOVA revealed significant differences between diver and camera methods ( $F = 3.92$ ,  $df = 1$ ,  $p = 0.002$ ) and between habitat types ( $F = 7.15$ ,  $df = 1$ ,  $p < 0.001$ ), as well as an interaction between the two ( $F = 2.31$ ,  $df = 1$ ,  $p = 0.028$ ). Overall, there was no significant effect of diver presence (before, during, or after) on fish counts from the cameras (Friedman rank sum test  $p = 0.15$ , Kendall’s  $W = 0.09$ ).

Table 1. Accuracy of model predictions.

True Image Designation	Correct Predictions	Incorrect Predictions	Percent Accuracy (%)
<i>Fish</i>	444	165	72.9
<i>NoFish</i>	2057	36	98.3
<b>All images</b>	<b>2501</b>	<b>201</b>	<b>92.5</b>

<https://doi.org/10.1371/journal.pone.0263377.t001>

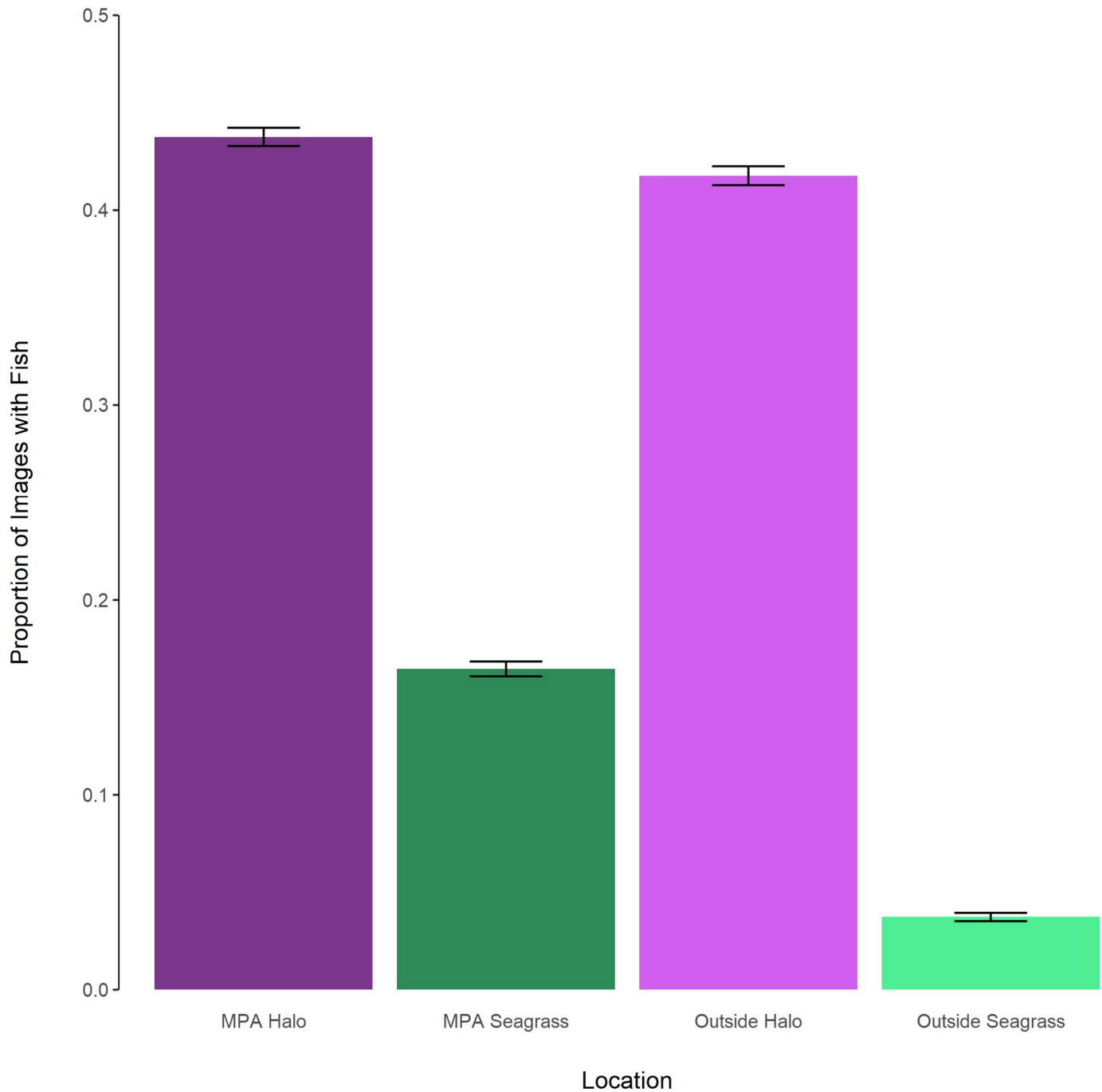


**Fig 5. Non-metric multidimensional scaling (NMDS) based on presence/absence data.** Points represent 20 individual site observations by divers or 3 pooled images from 18 given sites and are colored purple for halo/reef cameras and dark green for seagrass or algae cameras. ANOSIM of this NMDS output found a significant difference between halo/reef (purple) and seagrass/algae (green) communities ( $R = 0.68$ ,  $p < 0.001$ ). Dots represent diver observations, while triangles represent camera images. Both diver and camera observations reflect the same clear difference between fish communities on the reef and those in the seagrass. See S5 Fig for a stressplot of this NMDS.

<https://doi.org/10.1371/journal.pone.0263377.g005>

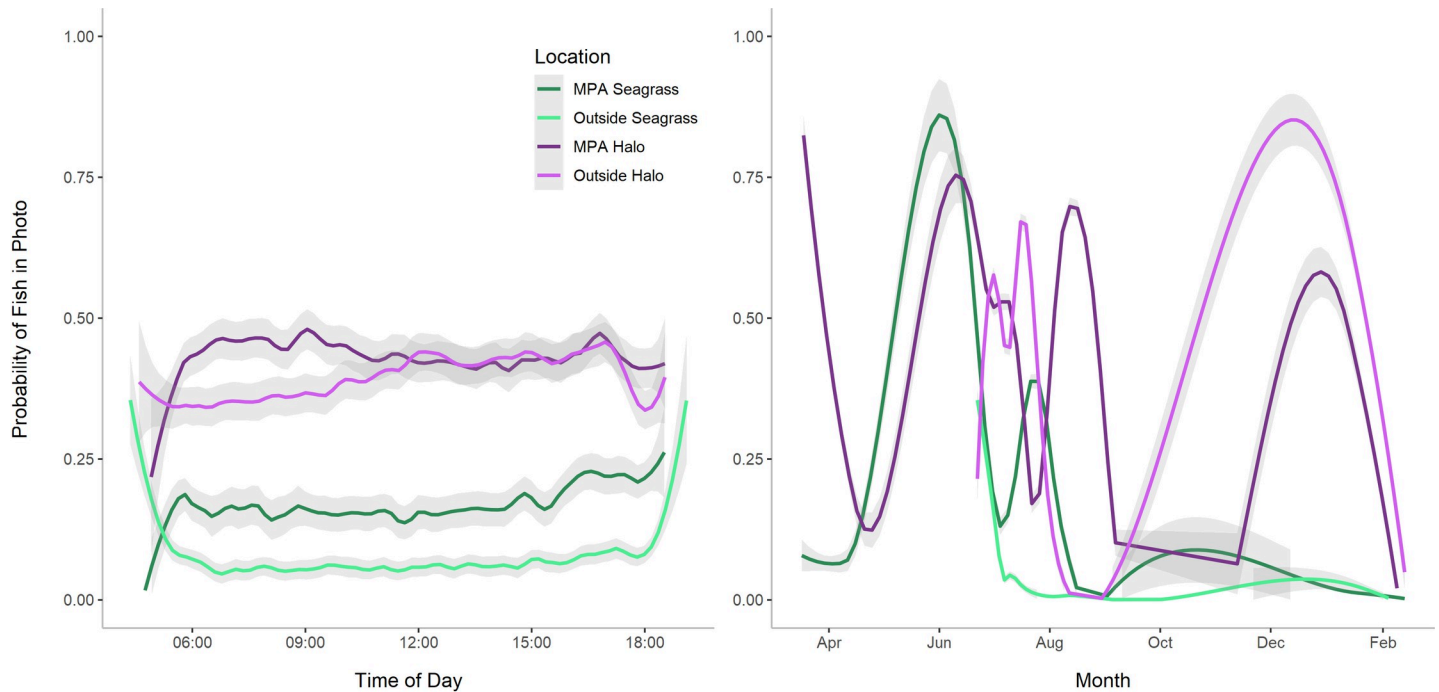
**Fish distributions.** The fish detection counts based on camera data revealed that the proportion of images with fishes was 32% higher in halos compared to seagrass or algae camera locations ( $\chi^2 = 4892.2$ ,  $df = 1$ ,  $p < 0.001$ ). Fish counts were 38% higher in algae environments than in seagrass ( $\chi^2 = 5298.3$ ,  $df = 1$ ,  $p < 0.001$ ), which could represent the actual presence of animals in these habitats or might reflect differences in the ability of humans and human-trained computer vision to distinguish small fishes in the denser seagrass cover. Protected status also had a smaller but significant effect, with a 6% increase in the proportion of photos

with fishes within the MPA, compared to outside the MPA ( $\chi^2 = 174.09$ ,  $df = 1$ ,  $p < 0.001$ ). Analysis of camera locations relative to both benthic cover (i.e., reef/halo or algae/seagrass) and protection status (inside or outside MPA) revealed that fish detection is always higher in halos than in seagrass or algae communities at Lighthouse Reef, and the number of fishes in each community type is relatively higher inside the MPA (Figs 6 and 7). Fishes were



**Fig 6. Proportion of images containing fishes at halo and grass/algae sites inside and outside of the Half Moon Caye MPA.** Halo sites are represented in shades of purple, seagrass/algae sites in shades of green. Each bar is a proportion based on the  $n$  for that treatment. Error bars represent  $\pm 1$  standard deviation from the binomial distribution, scaled by the respective  $n$  for each treatment.

<https://doi.org/10.1371/journal.pone.0263377.g006>



**Fig 7. Time series showing smoothed data from 40 different cameras deployed asynchronously between March 2018 and March 2019.** (A) Fish probabilities for individual images in halo/reef and seagrass/algae habitats by time of day. Probabilities ranged from 0 to 1 for this dataset, but smoothing visually compresses the highest and lowest probability values on the graph. (B) Fish probabilities for images by date, showing variation over the course of one year at Lighthouse Reef Atoll. Trend lines for all groups were generated using the `geom_smooth` function with method “loess” in `ggplot2` [53]. Light gray shading represents 95% confidence intervals.

<https://doi.org/10.1371/journal.pone.0263377.g007>

consistently more abundant around reefs and halos than in surrounding vegetation. However, there may also be a seasonal effect, with fewer fishes detected in seagrass environments during the fall and winter (October through February), while fish detections in halos increased during this time (Fig 7).

A preliminary analysis of diver observations from 20 camera locations as well as an additional 60 images from those sites did not contain a single instance of an herbivorous reef fish (surgeonfish and all adult parrotfish except for *Sparisoma radians*) outside of a halo. Predatory fishes, including barracuda (*Sphyraena barracuda*) and multiple species of jacks (*Caranx* spp.), were seen in both the reef/halo and seagrass/algae environments at different sites. Snappers (*Ocyurus chrysurus*, *Lutjanus* spp.) were seen primarily on the reef.

## Discussion

### Camera performance

**Deployment duration and function.** Underwater camera traps proved to be energy-efficient, durable, and capable of producing large volumes of quality images representative of the fish communities at their locations. The DEAC camera trap design is a reliable, cost-effective, and easy-to-implement solution allowing the expansion of terrestrial camera trapping techniques to shallow marine environments. The set-up can easily be used to vastly expand the capability of BRUVs and associated techniques [13–15] and can also provide marine observations over periods of months. In addition to being long-term and more scalable, un-baited camera traps such as these may be more accurate than BRUVs in their estimation of population and community composition [54] and allow the use of random encounter models for occupancy that are widely used with terrestrial camera traps [52].

Camera placement and orientation is an important consideration when deploying DEACs in the field, and the ideal placement may vary with geographic location, season, target time of day, water quality, and shading from local structures like patch reefs or docks. The cameras have no observable effect on the behavior of marine animals and may therefore be used to document species that typically evade divers. Even cameras that are “cleaned” by herbivorous fishes in the halo do not appear to attract any more attention than natural patches of algae or coral rubble. They are also capable of remaining underwater at depth for long periods and recording high volumes of photo or video observations—our longest running cameras were deployed for 5 months, captured over 7,500 images each, and still had battery life to extend to a year. The exceptional duration and the ability to capture vast numbers of images, and to automatically recognize photos with fishes (see below) gives the capability to not only provide a continuous record of rare species behaviors or visits by transient species that divers would miss or only observe by chance, but also, for the first time, to record changes in behavior, relative abundance, and migrations across seasons and lunar phases. DEACs are flexible in their ability to record both photo and video at a variety of time intervals suitable for addressing different ecological and behavioral questions.

**Comparison to existing methods.** DEACs offer significant advantages over existing underwater camera options for remote, shallow marine environments, as well as surveys by divers. While the longer time interval chosen for our extended duration field tests did not capture as many individual fishes in each image as divers recorded during the same 15-minute period corresponding to a single photo, analysis of synchronous diver and camera observations at different sites demonstrated the ability of both methods to capture differences in community composition between distinct benthic environments (Fig 5). These compositional differences also support the placement of the halo and control cameras at each site, suggesting that the cameras were separated enough to capture these different adjacent community types while remaining close enough to control for local site conditions. The only group of animals observed by divers that were not well captured by DEAC images were schools of roving juvenile fishes and small parrotfishes that camouflage well within seagrass and algal environments and are best detected through movement, which may have contributed to the interaction between habitat type and observation method detected by the PERMANOVA. The inability of the camera images and corresponding CNN to identify these fishes may be due to a combination of more limited image quality in these turbid environments and the inability of human sorters (on whose data the CNN was trained) to distinguish these small fishes later in a static image. Since these species are neither threatening predators nor significant grazers on seagrass or foliose macroalgae [55,56], their detection or inclusion in fish community composition was relatively unimportant for addressing the question used here to assess DEAC utility. However, for applications where detection of these species or similarly small or cryptic organisms is important, use of short video clips in place of still images (an option readily enabled with the DEAC design) could increase the ability of both humans and machines to detect these animals through motion. Because of their long deployment time, the DEACs captured differences inside and outside of the MPA that were not revealed by the diver observations, as well as changes based on time of day and season (Fig 7), which are hard to quantify with a limited number of diver surveys.

The current duration of camera operation is over an order of magnitude longer than the operational duration achieved by similar non-tethered underwater cameras [7,20], even with biofouling seriously impairing image quality after 1–2 months (S2 Fig). To the best of our knowledge, these are the first such underwater camera traps to be affordable (<\$200 per unit), power-efficient (therefore deployable over the span of months) and also self-contained, without the challenges imposed by a surface-tethered external power source, which typically precludes deployment in remote regions and at multiple sites over large areas.

The most closely related marine camera application to the DEACs is GUARD1 [29,57,58], an autonomous camera trap also operating on a time-lapse and for similarly long durations (48,298 images over 5.5 months [58]), which uses on-board image recognition and classification to detect gelatinous zooplankton in open waters. While GUARD1 is also a “low-cost” solution, the price of just the commercial camera used in the GUARD1 system is more than the entire DEAC set-up, including housing, memory, and batteries. This makes the DEAC design more suitable for deploying networks of many individual cameras over a large area and for study designs involving pairs or groups of cameras in close proximity. DEACs are inexpensive and easy to assemble, allowing many of them to be quickly deployed to address a variety of ecological questions, while GUARD1 units are individually more costly, complicated to build, and currently engineered to address a very specific monitoring challenge. DEACs and GUARD1 are designed for different marine environments and applications, with GUARD1 optimized to detect specific zooplankton [29] and potentially also fishes [30] in open water, while DEACs are suitable for detecting fishes and other marine animals, changes in habitat or other ecosystem features, and general monitoring in shallow, nearshore benthic environments, extending the kinds of landscape-level data taken with terrestrial camera traps into the marine realm.

**Future improvements.** The greatest immediate limitation to the DEAC design is the inability to prevent or seriously reduce biofouling without periodic manual cleaning of the cameras, although the magnitude of this challenge may be location-specific. Biofouling is a common problem with unsupervised underwater monitoring equipment [59], with multiple solutions proposed to control it [6], including local chlorination [59,60], copper sheeting and mesh, UV radiation [61], and wipers [62]. Ongoing tests of camera and housing design are incorporating these methods to reduce biofouling by marine organisms and extend the functional life of the lens window to better reflect the power and storage capabilities of the camera. Preliminary tests of a similar camera design in Hawai'i suggest that the extreme biofouling observed at Lighthouse Reef is not representative of all such reef environments and may be related to especially high productivity and suspended organic matter in the water column at this location. Simple monthly cleaning of the lens with a saltwater aquarist's sponge is another remedy if revisitation is possible.

While both the batteries and memory cards we employed were sufficient for the deployment duration and image capture frequency tested, the inability to utilize memory cards with greater than 32 GB of storage in Cuddeback cameras is problematic with a more frequent time lapse interval or in the case of short videos collected in place of single images. Therefore, optimizing and expanding both power and data storage capacity of these cameras via commercial means (cameras with larger SD storage capability) or noncommercial modifications is another potentially valuable direction for future research. For example, SD Ultra Capacity (SDUC) storage media currently affords 2 TB of storage, vastly expanding the capability of underwater camera traps for either smaller time intervals between images or multi-year deployments. Indeed, large storage capacity may obviate the need for camera triggering, making time-lapse or video with on-board object recognition (as used in GUARD1 [29]) a superior alternative, giving the ability to capture even rare or transient species while effectively using the millions of potential images generated. Design improvements currently in progress focus on reducing biofouling, extending the depth range of the camera housings to over 50 m, and implementing an external flash or other nighttime illumination.

## Image analysis

The success of our initial image sorting using the CNN ResNet-50 illustrates the power of machine learning and computer vision techniques to drastically reduce time and cost when



dealing with large image sets in marine environments. Our ability to use this trained model to pre-sort images with high accuracy before attempting further analysis via manual or autonomous machine-learning based methods also reduces the cost of the most disadvantageous aspect of our timed image capture method: the number of frames with no objects of interest to the current study. Now that datasets obtained from this and similar shallow marine environments can be easily sorted to exclude non-target images, future underwater camera trap projects using a similar time-lapse method will be able to quickly remove the majority of empty frames, while retaining the ability to measure frequency of detection events and variations in fish presence by time of day, season, or other environmental variables by comparison of occupied and empty images. Expansion of the deep neural network model to focus on identification of individual species, functional groups, and/or the sizes of different individuals, as has been done in similar image analyses [16,18,20,22], will further streamline the analysis of this and related large image datasets.

Our community composition analysis also demonstrates that the 15-minute photo interval of our cameras was sufficient to capture the same community differences between benthic environments that were observed by divers in the water. This suggests that while a single image does not capture every fish active in the area during the 15-minute period it represents, the collection of images from any given site are representative of the community at that location and are likely to accurately reflect changes in species behavior, abundance, or diversity at the site over a range of time scales (e.g., daily vs. seasonal changes), perhaps more accurately than a limited series of diver observations would. Camera networks are an emerging method for monitoring fish activity and other behaviors on diel and seasonal timescales [63,64] and may reveal important information about temporal resource partitioning, predator avoidance, and other interspecific dynamics in the community [65]. Further study of camera captures vs. diver observations regarding species known to be wary of divers, either those using traditional open-circuit SCUBA or closed-circuit rebreathers [66], is an important line of future investigation to understand true reef fish occupancy and abundance.

Our analysis of the number of images with fishes collected at different groups of cameras revealed a difference undetected by the fish counts from our diver observations, showing that fishes were detected by cameras more frequently (and are therefore likely to have higher occupancy) inside of the MPA. The inability of traditional diver surveys alone to detect this difference at all reinforces the value of spatially-distributed, long-term datasets like the images collected from our cameras. The observed differences in fish detection between halo and grass/algae control sites also reinforce diver observations of both fewer fishes and a different fish community in the algae or seagrass beds away from the reef and support the results of the community composition analysis (Fig 5).

### Case study and future applications

Our observations of fish species presence across different environments are consistent with the landscape of fear explanation for heightened herbivory inside coral reef halos. We obtained a large volume of usable images from sites with a variety of depths and benthic cover types, subjected to different fishing pressures inside and outside of a local MPA, and monitored over the course of different seasons. This allows us to reasonably conclude that our observations are likely representative of fish presence and behavior at patch reef sites within Lighthouse Reef Atoll as a whole. The complete lack of grazing reef fishes observed outside of the halo region by cameras located in surrounding seagrass within 30 m or less of the halo supports the idea that a predator modification of prey fish behavior imposes real constraints on fish movement and that heightened grazing pressure adjacent to reef structures is not simply the result of

fishes randomly dispersing with distance from the reef. If the latter were true, the pattern would be a simple exponential decrease in fishes as distance from the patch reef increased, which would likely be reflected in a lower (but nonzero) number of grazing fishes detected at seagrass and algae camera locations, compared to those at the halo edge.

An alternative hypothesis is that the lack of grazing fishes appearing in images taken outside of the halo is due to unequal detection of fishes between the two environments [67], possibly the result of faster or more furtive movements outside the relative safety of the halo. However, this explanation is refuted by two pieces of evidence: First, diver observations also showed no herbivorous reef fishes at grass or algal camera sites; second, other fish species that were observed by divers to move quickly through the seagrass or algal environment (e.g., bar jacks, *Caranx ruber*) appear in both halo and control camera images, indicating that cameras placed in the grass are very capable of photographing fishes moving through that environment. It is therefore more likely that the complete lack of detection of reef-based grazers outside of the halo, even though their food is in much higher supply there, is due to their total or near-total absence from this environment because of the combined lack of shelter and exposure to predators [35].

The most common large predators seen by both divers and cameras were barracuda (*S. barracuda*). This is based only on the 54 images used for the NMDS and anecdotal observations made in the course of sorting and processing additional images, since our CNN did not include detection of individual fishes to the species level. One of us (MR Silman) observed a barracuda preying on a large adult parrotfish, which smaller predatory fishes are unlikely to target. This demonstrates that barracuda present a real threat to even the largest herbivorous fishes in this system. Barracuda were observed in both reef and seagrass/algae environments, but the reef provides more structure (and likely protection [68,69]) to potential prey than the open seagrass beds. This could lead foraging fishes to limit their movement to within a certain radius of the reef, which is the most commonly-cited explanation for reef halos and is consistent with our observations.

This study demonstrates the value of long-term, spatially-distributed underwater camera trap observations for addressing a subtle difference in fish communities and benthic pattern generation. DEACs can be easily deployed alone or in large spatial arrays for short or extended time periods, and they are capable of recording periodic still images for long-term studies or short videos for detailed behavioral observations. Most importantly, these cameras are highly energy-efficient and require little-to-no maintenance while deployed, making them ideal for remote locations or extended observations that surface-anchored systems or commercial underwater cameras with limited battery life are not suitable for. Multiple networks of camera traps like TEAM [21] and Snapshot Serengeti [22] have been successfully deployed over large regions in terrestrial environments, and DEACs offer the option to now expand such long-term, spatially-extensive monitoring efforts to the marine realm. The use of a CNN makes processing the volume of images collected over such a long-term study a practical option, and continuing advancements in machine learning and computer vision are likely to enable further processing of similar large visual datasets in the future.

## Supporting information

**S1 Fig. Deployment periods for individual cameras, typically deployed in pairs at a given site.** Treatment designations represent if the camera was deployed in the Half Moon Caye marine protected area (MPA) or outside (OUT), in algae (a) or seagrass (g), the site number or other site designation within that treatment, and camera location at the halo edge (H) or farther away (C). The final number present for only certain treatments indicates which of

multiple deployments at the same location this period refers to.  
(JPEG)

**S2 Fig. Decline in fish detections due to biofouling over time for the 13 longest deployments.** Deployment time in days is shown on the x-axis, while the probability of a given image containing at least one fish is shown on the y-axis, as calculated by the ResNet-50 model. Fish detection probabilities decrease over the first month, at which point the pattern becomes less clear, likely the result of inaccurate image classifications at certain sites in response to biofouling in the images. The red line represents the smoothed results of all sites combined, using the “loess” function in R package ggplot2 (Wickham 2009).  
(PNG)

**S3 Fig. Fitted lines for fish detection probabilities over long-term reef-adjacent halo deployments.** Lines were fitted with the “loess” local polynomial regression fitting function with span = 1 from R package ggplot2 (Wickham 2009). The dark line represents the smoothed results of all sites shown on this graph. Not only do fish detections decline consistently at most sites due to biofouling of the camera lens over the first month, model accuracy is clearly impaired past this point at certain sites, leading to a false increase in fish detection probability for some cameras (e.g., note the erroneously high probabilities around 80–100 days for MPATr12H, shown in blue).  
(JPEG)

**S4 Fig. Fitted lines for fish detection probabilities over long-term seagrass or algae meadow deployments.** Lines were fitted with the “loess” local polynomial regression fitting function with span = 1 from R package ggplot2 (Wickham 2009). The dark line represents the smoothed results of all sites shown on this graph. Note that fish detection probabilities are lower at grass sites, relative to reef/halo sites, and the y-axis is scaled appropriately. After the first month, fish detection probabilities erroneously rise at a couple sites, likely reflecting reduced accuracy in the ResNet-50 model’s image classifications as a result of biofouling.  
(JPEG)

**S5 Fig. Stressplot for the NMDS shown in Fig 5 of the main text.** Created using the stressplot() function in R package vegan (Oksanen et al., 2018) [42].  
(PNG)

## Acknowledgments

We thank the Belize Audubon Society for hosting us at Half Moon Caye and assisting with data collection, especially Eli Romero, Dominique Lizama, and Shane Young. This work was conducted under Belize Fisheries Permit 00021–18. We thank Greg Asner and the Carnegie Airborne Observatory/ASU Center for Global Discovery and Conservation Science for use of their depth models. We also thank Elvis Solis for logistics help and counsel, Connor Walsh and John Gorelick for help in the field and in the lab, and all undergraduates who contributed to classifying and analyzing the data, especially Tatianna Stroud and Joseph Chen.

## Author Contributions

**Conceptualization:** Stephanie M. Bilodeau, Miles R. Silman.

**Data curation:** Stephanie M. Bilodeau, Binfeng Xu.

**Formal analysis:** Stephanie M. Bilodeau, Binfeng Xu.

**Funding acquisition:** Stephanie M. Bilodeau, Austin W. H. Schwartz, Miles R. Silman.

**Investigation:** Stephanie M. Bilodeau, Austin W. H. Schwartz, Binfeng Xu, V. Paúl Pauca, Miles R. Silman.

**Methodology:** Stephanie M. Bilodeau, Binfeng Xu, V. Paúl Pauca, Miles R. Silman.

**Resources:** Miles R. Silman.

**Software:** Binfeng Xu, V. Paúl Pauca.

**Supervision:** V. Paúl Pauca, Miles R. Silman.

**Validation:** Stephanie M. Bilodeau, Austin W. H. Schwartz, Binfeng Xu.

**Visualization:** Stephanie M. Bilodeau.

**Writing – original draft:** Stephanie M. Bilodeau, Austin W. H. Schwartz, Binfeng Xu, V. Paúl Pauca, Miles R. Silman.

**Writing – review & editing:** Stephanie M. Bilodeau, Austin W. H. Schwartz, Binfeng Xu, V. Paúl Pauca, Miles R. Silman.

## References

1. Steenweg R, Hebblewhite M, Kays R, Ahumada J, Fisher JT, Burton C, et al. Scaling-up camera traps: Monitoring the planet's biodiversity with networks of remote sensors. *Front Ecol Environ*. 2017; 15(1):26–34.
2. Rowcliffe JM. Key frontiers in camera trapping research. *Remote Sens Ecol Conserv*. 2017; 3(3):107–8.
3. Burton AC, Neilson E, Moreira D, Ladle A, Steenweg R, Fisher JT, et al. REVIEW: Wildlife camera trapping: a review and recommendations for linking surveys to ecological processes. *Methods Ecol Evol*. 2015;675–85.
4. Cutler TL, Swann DE. Using remote photography in wildlife ecology: a review. *Wildl Soc Bull* 1973–2006. 1999; 27(3):571–81.
5. Giles JW, Bankman IN. Underwater optical communications systems. Part 2: basic design considerations. In: MILCOM 2005–2005 IEEE Military Communications Conference. 2005. p. 1700–1705 Vol. 3.
6. Del-Rio J, Noguera M, Toma DM, Martínez E, Artero-Delgado C, Bghiel I, et al. Obsea: a decadal balance for a cabled observatory deployment. *IEEE Access*. 2020; 8:33163–77.
7. Williams K, De Robertis A, Berkowitz Z, Rooper C, Towler R. An underwater stereo-camera trap. *Methods Oceanogr*. 2014 Dec 1; 11:1–12.
8. Martignac F, Daroux A, Bagliniere J-L, Ombredane D, Guillard J. The use of acoustic cameras in shallow waters: new hydroacoustic tools for monitoring migratory fish population. A review of DIDSON technology. *Fish Fish*. 2015; 16(3):486–510.
9. Jones RE, Griffin RA, Unsworth RKF. Adaptive Resolution Imaging Sonar (ARIS) as a tool for marine fish identification. *Fish Res*. 2021 Nov 1; 243:106092.
10. McCauley DJ, DeSalles PA, Young HS, Gardner JPA, Micheli F. Use of high-resolution acoustic cameras to study reef shark behavioral ecology. *J Exp Mar Biol Ecol*. 2016 Sep; 482:128–33.
11. Aguzzi J, Chatzievangelou D, Marini S, Fanelli E, Danovaro R, Flögel S, et al. New high-tech flexible networks for the monitoring of deep-sea ecosystems. *Environ Sci Technol*. 2019 Jun 18; 53(12):6616–31. <https://doi.org/10.1021/acs.est.9b00409> PMID: 31074981
12. Wilby A, Kastner R, Hostler A, Slattery E. Design of a low-cost and extensible acoustically-triggered camera system for marine population monitoring. In: OCEANS 2016 MTS/IEEE Monterey [Internet]. Monterey, CA, USA: IEEE; 2016 [cited 2021 Dec 10]. p. 1–9. Available from: <http://ieeexplore.ieee.org/document/7761320/>.
13. Cappel M, Speare P, De'ath G. Comparison of baited remote underwater video stations (BRUVS) and prawn (shrimp) trawls for assessments of fish biodiversity in inter-reefal areas of the Great Barrier Reef Marine Park. *J Exp Mar Biol Ecol*. 2004 May 12; 302(2):123–52.
14. Colton MA, Swearer SE. A comparison of two survey methods: differences between underwater visual census and baited remote underwater video. *Mar Ecol Prog Ser*. 2010 Feb 11; 400:19–36.

15. Brooks EJ, Sloman KA, Sims DW, Danylchuk AJ. Validating the use of baited remote underwater video surveys for assessing the diversity, distribution and abundance of sharks in the Bahamas. *Endanger Species Res.* 2011 Mar 17; 13(3):231–43.
16. Boussarie G, Teichert N, Lagarde R, Ponton D. BichiCAM, an underwater automated video tracking system for the study of migratory dynamics of benthic diadromous species in streams. *River Res Appl.* 2016 Jul 1; 32(6):1392–401.
17. Sheehan EV, Bridger D, Nancollas SJ, Pittman SJ. PelagiCam: A novel underwater imaging system with computer vision for semi-automated monitoring of mobile marine fauna at offshore structures. *Environ Monit Assess.* 2020;192. <https://doi.org/10.1007/s10661-020-8155-z> PMID: 32080771
18. Boom BJ, He J, Palazzo S, Huang PX, Beyan C, Chou H-M, et al. A research tool for long-term and continuous analysis of fish assemblage in coral-reefs using underwater camera footage. *Ecol Inform.* 2014 Sep 1; 23:83–97.
19. Marini S, Fanelli E, Sbragaglia V, Azzurro E, Fernandez JDR, Aguzzi J. Tracking fish abundance by underwater image recognition. *Sci Rep.* 2018 Sep 13; 8(1):13748. <https://doi.org/10.1038/s41598-018-32089-8> PMID: 30213999
20. Siddiqui SA, Salman A, Malik MI, Shafait F, Mian A, Shortis MR, et al. Automatic fish species classification in underwater videos: exploiting pre-trained deep neural network models to compensate for limited labelled data. *ICES J Mar Sci.* 2018 Jan 1; 75(1):374–89.
21. Beaudrot L, Ahumada JA, O'Brien T, Alvarez-Loayza P, Boekee K, Campos-Arceiz A, et al. Standardized assessment of biodiversity trends in tropical forest protected areas: The end is not in sight. *PLOS Biol.* 2016 Jan 19; 14(1):e1002357. <https://doi.org/10.1371/journal.pbio.1002357> PMID: 26785119
22. Norouzzadeh MS, Nguyen A, Kosmala M, Swanson A, Palmer MS, Packer C, et al. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proc Natl Acad Sci.* 2018 Jun 19; 115(25):E5716–25. <https://doi.org/10.1073/pnas.1719367115> PMID: 29871948
23. Mortazavi H, Oakley JP, Barkat B. Mitigating the effect of optical back-scatter in multispectral underwater imaging. In: 2012 IEEE International Conference on Imaging Systems and Techniques Proceedings. 2012. p. 633–6.
24. Bicknell AW, Godley BJ, Sheehan EV, Votier SC, Witt MJ. Camera technology for monitoring marine biodiversity and human impact. *Front Ecol Environ.* 2016; 14(8):424–32.
25. Ko D, Gu B, Kim J. Analysis of the luring characteristics of phototactic fishes under LED illumination in water. 2018; 13(1):4.
26. Villon S, Mouillot D, Chaumont M, Darling ES, Subsol G, Claverie T, et al. A deep learning method for accurate and fast identification of coral reef fishes in underwater images. *Ecol Inform.* 2018 Nov; 48:238–44.
27. Ditria EM, Lopez-Marcano S, Sievers M, Jinks EL, Brown CJ, Connolly RM. Automating the analysis of fish abundance using object detection: Optimizing animal ecology with deep learning. *Front Mar Sci.* 2020;7. <https://doi.org/10.3389/fmars.2020.00429>
28. Zuazo A, Grinyó J, López-Vázquez V, Rodríguez E, Costa C, Ortenzi L, et al. An automated pipeline for image processing and data treatment to track activity rhythms of *Paragorgia arborea* in relation to hydrographic conditions. *Sensors.* 2020 Jan; 20(21):6281. <https://doi.org/10.3390/s20216281> PMID: 33158174
29. Corgnati L, Marini S, Mazzei L, Ottaviani E, Aliani S, Conversi A, et al. Looking inside the ocean: toward an autonomous imaging system for monitoring gelatinous zooplankton. *Sensors.* 2016 Dec; 16(12):2124. <https://doi.org/10.3390/s16122124> PMID: 27983638
30. Marini S, Corgnati L, Mantovani C, Bastianini M, Ottaviani E, Fanelli E, et al. Automated estimate of fish abundance through the autonomous imaging device GUARD1. *Measurement.* 2018 Oct; 126:72–5.
31. Zurowietz M, Langenkämper D, Hosking B, Ruhl HA, Nattkemper TW. MAIA—A machine learning assisted image annotation method for environmental monitoring and exploration. *PLOS ONE.* 2018 Nov 16; 13(11):e0207498. <https://doi.org/10.1371/journal.pone.0207498> PMID: 30444917
32. Hay ME. Patterns of fish and urchin grazing on Caribbean coral reefs: Are previous results typical? *Ecology.* 1984 Apr 1; 65(2):446–54.
33. Bilodeau SM, Layman CA, Silman MR. Benthic pattern formation in shallow tropical reefs: does grazing explain grazing halos? *Landsc Ecol.* 2021; 36:1605–20.
34. Hammerschlag N, Broderick AC, Coker JW, Coyne MS, Dodd M, Frick MG, et al. Evaluating the landscape of fear between apex predatory sharks and mobile sea turtles across a large dynamic seascape. *Ecology.* 2015; 96(8):2117–26. <https://doi.org/10.1890/14-2113.1> PMID: 26405737
35. Madin EMP, Madin JS, Booth DJ. Landscape of fear visible from space. *Sci Rep.* 2011 Dec;1. <https://doi.org/10.1038/srep00001> PMID: 22355520

36. Sale PF, Douglas WA. Temporal variability in the community structure of fish on coral patch reefs and the relation of community structure to reef structure. *Ecology*. 1984; 65(2):409–22.
37. Bohnsack JA. Are high densities of fishes at artificial reefs the result of habitat limitation or behavioral preference? *Bull Mar Sci*. 1989; 44(2):631–45.
38. Layman CA, Allgeier JE, Yeager LA, Stoner EW. Thresholds of ecosystem response to nutrient enrichment from fish aggregations. *Ecology*. 2013 Feb 1; 94(2):530–6. <https://doi.org/10.1890/12-0705.1> PMID: 23691671
39. Bilodeau SM. Ecological process in pattern generation in tropical coral-seagrass reefs. M.Sc. Thesis, Wake Forest University. 2019. Available from: <http://hdl.handle.net/10339/93912>.
40. Stoddart DR. Three Caribbean atolls: Turneffe Islands, Lighthouse Reef, and Glover's Reef, British Honduras. *Atoll Res Bull*. 1962; 87:1–151.
41. Burkepile DE. Comparing aquatic and terrestrial grazing ecosystems: is the grass really greener? *Oikos*. 2013; 122(2):306–12.
42. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGinn D, et al. *vegan: Community ecology package* [Internet]. 2018 [cited 2021 Jan 12]. Available from: <https://CRAN.R-project.org/package=vegan>.
43. R Core Team. *R: A language and environment for statistical computing* [software]. Vienna, Austria: R Foundation for Statistical Computing; 2020 [cited 2021 Jan 12]. Available from: <https://www.R-project.org/>.
44. Rinker TW. *qdap: Quantitative discourse analysis package* [Internet]. Buffalo, New York, New York, USA; 2019 [cited 2021 Jan 12]. Available from: <http://github.com/trinker/qdap>. <https://doi.org/10.5281/zenodo.11124>
45. Wickham H. Reshaping data with the reshape package. *J Stat Softw*. 2007 Nov 13; 21(1):1–20.
46. Kassambara A. *rstatix: Pipe-Friendly Framework for Basic Statistical Tests* [Internet]. 2020 [cited 2021 Jan 12]. Available from: <https://CRAN.R-project.org/package=rstatix>.
47. Niedballa J, Sollmann R, Courtiol A, Wilting A. *camtrapR: an R package for efficient camera trap data management*. *Methods Ecol Evol*. 2016; 7(12):1457–62.
48. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. Las Vegas, NV, USA: IEEE; 2016 [cited 2019 Apr 14]. p. 770–8. Available from: <http://ieeexplore.ieee.org/document/7780459/>.
49. Paszke A, Gross S, Chintala S, Chanan G, Yang E, DeVito Z, et al. Automatic differentiation in PyTorch. 2017 Oct 28 [cited 2019 Apr 14]; Available from: <https://openreview.net/forum?id=BJJsrnfCZ>.
50. Howard J, others. *fastai* [Internet]. GitHub; 2018. Available from: <https://github.com/fastai/fastai>.
51. Lin T-Y, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. In: The IEEE International Conference on Computer Vision (ICCV) [Internet]. 2017 [cited 2019 Jun 16]. p. 2980–8. Available from: <http://arxiv.org/abs/1708.02002>.
52. Rowcliffe JM, Carbone C, Kays R, Kranstauber B, Jansen PA. Density estimation using camera trap surveys: the random encounter model. In: Meek P, Fleming P, editors. *Camera Trapping: Wildlife Management and Research*. Melbourne, Australia: CSIRO Publishing; 2014. p. 317–24.
53. Wickham H. *ggplot2: Elegant Graphics for Data Analysis* [Internet]. Cham: Springer International Publishing; 2016 [cited 2021 Sep 9]. (Use R!). Available from: <http://link.springer.com/10.1007/978-3-319-24277-4>.
54. McCoy JC, Ditchkoff SS, Steury TD. Bias associated with baited camera sites for assessing population characteristics of deer. *J Wildl Manag*. 2011; 75(2):472–7.
55. Randall JE. Food habits of reef fishes of the West Indies. 1967.
56. Overholtzer K, Motta P. Comparative resource use by juvenile parrotfishes in the Florida Keys. *Mar Ecol Prog Ser*. 1999; 177:177–87.
57. Corgnati L, Mazzei L, Marini S, Aliani S, Conversi A, Griffa A, et al. Automated gelatinous zooplankton acquisition and recognition. In: 2014 ICPR Workshop on Computer Vision for Analysis of Underwater Imagery [Internet]. Stockholm, Sweden: IEEE; 2014 [cited 2021 Sep 13]. p. 1–8. Available from: <http://ieeexplore.ieee.org/document/6961262/>.
58. Marini S, Corgnati L, Mazzei L, Ottaviano E, Isoppo B, Aliani S, et al. GUARD1: An autonomous system for gelatinous zooplankton image-based recognition. In: OCEANS 2015—Genova [Internet]. Genova, Italy: IEEE; 2015 [cited 2021 Sep 13]. p. 1–7. Available from: <http://ieeexplore.ieee.org/document/7271704/>.
59. Delauney L, Compère C. An Example: Biofouling Protection for Marine Environmental Sensors by Local Chlorination. In: Flemming H-C, Murthy PS, Venkatesan R, Cooksey K, editors. *Marine and Industrial Biofouling* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009 [cited 2018 Sep

- 14]. p. 119–34. (Springer Series on Biofilms). Available from: [https://doi.org/10.1007/978-3-540-69796-1\\_6](https://doi.org/10.1007/978-3-540-69796-1_6).
60. Xue Y, Zhao J, Qiu R, Zheng J, Lin C, Ma B, et al. In situ glass antifouling using Pt nanoparticle coating for periodic electrolysis of seawater. *Appl Surf Sci*. 2015 Dec 1; 357:60–8.
  61. Patil JS, Kimoto H, Kimoto T, Saino T. Ultraviolet radiation (UV-C): a potential tool for the control of biofouling on marine optical instruments. *Biofouling*. 2007 Jan 1; 23(4):215–30. <https://doi.org/10.1080/08927010701275598> PMID: 17653932
  62. Chabanet P, Loiseau N, Join J-L, Ponton D. VideoSolo, an autonomous video system for high-frequency monitoring of aquatic biota, applied to coral reef fishes in the Glorioso Islands (SWIO). *J Exp Mar Biol Ecol*. 2012 Nov 1; 430–431:10–6.
  63. Aguzzi J, Doya C, Tecchio S, De Leo FC, Azzurro E, Costa C, et al. Coastal observatories for monitoring of fish behaviour and their responses to environmental changes. *Rev Fish Biol Fish*. 2015 Sep 1; 25(3):463–83.
  64. Aguzzi J, Chatzievangelou D, Company JB, Thomsen L, Marini S, Bonofiglio F, et al. The potential of video imagery from worldwide cabled observatory networks to provide information supporting fish-stock and biodiversity assessment. *ICES J Mar Sci*. 2020 Dec 1; 77(7–8):2396–410.
  65. Kronfeld-Schor N, Dayan T. Partitioning of time as an ecological resource. *Annu Rev Ecol Evol Syst*. 2003; 34(1):153–81.
  66. Lindfield SJ, Harvey ES, McIlwain JL, Halford AR. Silent fish surveys: Bubble-free diving highlights inaccuracies associated with SCUBA-based surveys in heavily fished areas. *Methods Ecol Evol*. 2014; 5(10):1061–9.
  67. Thompson SK, Seber GAF. Detectability in conventional and adaptive sampling. *Biometrics*. 1994; 50(3):712–24. PMID: 7981397
  68. Kerry JT, Bellwood DR. The effect of coral morphology on shelter selection by coral reef fishes. *Coral Reefs*. 2012 Jun; 31(2):415–24.
  69. Ménard A, Turgeon K, Roche DG, Binning SA, Kramer DL. Shelters and their use by fishes on fringing coral reefs. *PLOS ONE*. 2012 Jun 20; 7(6):e38450. <https://doi.org/10.1371/journal.pone.0038450> PMID: 22745664