Miller Scott Donald (Orcid ID: 0000-0001-5092-6037)
Schmitt Russell J (Orcid ID: 0000-0002-9674-4244)
Rassweiler Andrew (Orcid ID: 0000-0002-8760-3888)

Title: Using machine learning to achieve simultaneous, georeferenced surveys of fish and benthic communities on shallow coral reefs

Authors:

Scott D. Miller[1,*], Alexandra K. Dubel[1], Thomas C. Adam[3], Dana T. Cook[2,3], Sally J. Holbrook[2,3], Russell J. Schmitt[2,3], Andrew Rassweiler[1]

Institutional affiliations:

[1]Department of Biological Science, Florida State University, 319 Stadium Drive, Tallahassee, FL, 32306, USA

[2]Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, CA, 93106, USA

[3]Marine Science Institute, University of California, Santa Barbara, CA, 93106, USA

***Corresponding author:** sdmiller@bio.fsu.edu

Running head: Simultaneous fish and benthic surveys

Key Words: computer vision, coral reef fishes, CoralNet, machine learning, photoquadrats, subtidal surveys, underwater visual census

Abstract

Surveying coastal systems to estimate distribution and abundance of fish and benthic organisms is labor-intensive, often resulting in spatially limited data that are difficult to scale up to an entire reef or island. We developed a method that leverages the automation of a machine learning platform, CoralNet, to efficiently and cost-effectively allow a single observer to simultaneously generate georeferenced data on abundances of fish and benthic taxa over large areas in shallow coastal environments. Briefly, a researcher conducts a fish survey while snorkeling on the surface and towing a float equipped with a handheld GPS and a downward-facing GoPro, passively taking ~10 photographs per meter of benthos. Photographs and surveys are later georeferenced and photographs are automatically annotated by CoralNet. We found that this method provides similar biomass and density values for common fishes as traditional scuba-based fish counts on fixed transects, with the advantage of covering a larger area. Our CoralNet validation determined that while photographs automatically annotated by CoralNet are less accurate than photographs annotated by humans at the level of a single image, the automated approach provides comparable or better estimations of the percent cover of the benthic substrates at the level of a minute of survey (~ 50 m^2 of reef) due to the volume of photographs that can be automatically annotated, providing greater spatial coverage of the site. This method can be used in a variety of shallow systems and is particularly advantageous when spatially-explicit data or surveys of large spatial extents are necessary.

Introduction

The detection and interpretation of ecological patterns and processes depend on the scale observed (Levin 1992). Mismatches in scale between observations and processes can lead to erroneous interpretations of the drivers of patterns in nature, making it imperative to consider multiple scales in study designs (Sandel and Smith 2009). Unfortunately, the scale at which observations of nature are made is often dictated by logistical constraints rather than by the underlying biology or specific research questions (Lindenmayer and Likens 2010; Estes et al. 2018). The appreciation of scale and spatial configuration in marine reef systems has galvanized the emerging field of seascape ecology, the subtidal equivalent of landscape ecology (Pittman et al. 2011), leading to novel insights into marine ecology and management (Boström et al. 2011). Due to the focus on spatial processes, spatially explicit data are critical to this burgeoning framework (Pittman 2017). With anthropogenic impacts affecting marine systems at a range of local (e.g., overharvesting and nutrient pollution) and global (e.g., climate change) scales (Bellwood et al. 2004; Zaneveld et al. 2016), new methodologies are needed for study on multiple scales and to integrate spatially explicit data from diverse sources (González-Rivero et al. 2020; Holbrook et al. 2022). Ideally, new methods would be able to efficiently and accurately survey biological communities at fine taxonomic and spatial resolutions with the ability to scale up to cover larger spatial extents when necessary.

A multitude of methodologies to survey subtidal communities have been developed that span the continuum from spatially-localized with high resolution to broader scales coming at the cost of coarser taxonomic resolution. At the extremes, *in situ* subtidal surveys can identify individual species with high accuracy in small areas, while remote sensing techniques can encompass entire study regions but may only be able to distinguish a few classes of organisms averaged over relatively large areas (i.e., multiple m^2) due to resolution limits (Hedley et al. 2017). Subtidal surveys can be both time- and labor-intensive, leading to estimates of community structure based on a limited number of observations or area sampled. For example, fish surveys typically employ either a belt transect or a stationary point count method (Caldwell et al. 2016). These sample fixed areas on the reef and require extensive replication to infer the density of patchily distributed and mobile fishes at larger scales (Colvocoresses and Acosta 2007), although GPS-enabled timed swims and towed-diver methods have been employed to allow greater area to be covered (Richards et al. 2011; Beck et al. 2014; Rizzari et al. 2014). Methods for surveying cover of benthic taxa are similarly varied, though most utilize *in situ* point-count methods that are time-intensive in the field, or photoquadrats, which are easier to collect in the field but are time-intensive to annotate in the laboratory (Beijbom et al. 2015). As with fish surveys, these methods ordinarily capture only small areas of the targeted ecosystem.

Although there will always be tradeoffs in survey methods along the axes of efficiency, spatial coverage, and taxonomic resolution, advances in technology may reduce some of these tradeoffs by increasing the spatial coverage while maintaining or even improving efficiency and taxonomic resolution (González-Rivero et al. 2014; Griffin et al. 2017). Recent advances in computer vision algorithms provide the ability to automatically detect a number of marine organisms from photographs and videos (e.g., Beijbom et al. 2015; Villon et al. 2018). In benthic marine systems, these algorithms have been leveraged to help alleviate the time-consuming work of manually annotating benthic photographs (Williams et al. 2019; González-Rivero et al. 2020). Different methods have been employed by researchers, but the CoralNet platform (Beijbom et al. 2015) is open source and commonly used in studies of marine ecosystems including coral reefs (Williams et al. 2019), temperate reefs (Griffin et al. 2017), and biofouling communities (Gormley et al. 2018).

The CoralNet platform (coralnet.ucsd.edu) has been the subject of several evaluations in the recent literature (Beijbom et al. 2015; Williams et al. 2019). These evaluations have revealed that CoralNet's automatic annotations provide comparable estimates of benthic cover to human experts in a variety of systems (Williams et al. 2019). Although these results are promising, prior assessments tended to focus on how well the automatic classifier performs compared to human annotations done on the same photographs, essentially testing how well CoralNet performs on data collected from traditional study designs where the intent is for experts to annotate most of the images. Images typically take more time to annotate than to collect in the field, so studies and sample sizes are often designed around this rate-limiting step of expert annotation, termed the "manual-annotation bottleneck" (Beijbom et al. 2015). While reducing this analytical bottleneck is certainly an advantage of the machine learning approach, we argue that the power of automatic annotation has yet to be fully realized. Automating the annotation process can be used in tandem with the automated collection of large numbers of photographs to greatly improve the efficiency in the field and expand the spatial coverage of benthic survey data with limited loss of taxonomic resolution. In this study, we build upon previous work by using CoralNet to automatically annotate hundreds of thousands of photographs. This is many more than could practically be scored by a human, resulting in greater spatial coverage of the site, and potentially a more accurate description of the ecological community.

To address the need for efficient, cost-effective methods to characterize fish and benthic communities across large areas, we describe novel methodology that leverages advances in machine learning to enable collection of paired, georeferenced survey data on both fish and benthic communities in shallow marine environments. Briefly, a snorkeler swimming on the surface visually surveys conspicuous, mobile fishes while towing a float equipped with a GPS and a downward facing camera that captures images of the benthos. The benthic photographs are later processed using CoralNet to automatically annotate hundreds of thousands of images. Although GPS-enabled fish surveys and automatic annotation of benthic photographs have both been utilized in the marine ecology literature (Beck et al. 2014; Beijbom et al. 2015), our approach combines these two data collection techniques into a single methodology conducted simultaneously by a single observer in the field, which enables questions to be addressed at larger spatial extents while maintaining fine spatial resolution. Here we describe the materials used to build the float, field methodology for the paired fish and benthic surveys, the methods to synchronize the different data sources and georeference the surveys, as well as the process used to validate the accuracy of cover estimates of benthic substrates from CoralNet. During our CoralNet validation, we specifically examine how well different numbers of manually- and automatically-annotated images estimate benthic cover based on a manually-annotated reference set of images. Additionally, we compare results from the described GPS fish counts with those of a traditional, scuba-based fixed transect method to demonstrate that despite different methodology, the two methods provide similar estimates of the biomass and density of herbivorous fish functional groups and key carnivorous families.

Materials and Procedures

Study Site

We collected data to assess and validate this method in June and July 2018 in the shallow lagoons (< 4 m) around the island of Moorea, French Polynesia (17.54°S, 149.83°W). Moorea is a high volcanic island located in the central South Pacific with a perimeter of ~ 50 km and an offshore barrier reef that encloses a shallow lagoon. The lagoon is characterized by a contiguous reef tract on the back reef just shoreward of the reef crest that transitions into scattered coral bommies interspersed with sand and coral rubble in the mid-lagoon. Alongshore is a shallow fringing reef that in places is separated from the mid-lagoon by a deeper channel. Moorea is the location of the National Science Foundation-funded Moorea Coral Reef Long Term Ecological Research (MCR LTER) site that has been collecting time series data on fish and benthic communities at six sites and two habitats nested within site (back reef and fringing reef) within the lagoon since 2004. These time series indicate that some areas on the north and northeast sides of the island have been losing coral and gaining macroalgae recently, while coral cover in other areas of the island has remained relatively constant (Schmitt et al. 2019, 2021; Carpenter 2020; Adam et al. 2021). Thus, the lagoons currently consist of a patchy landscape with different benthic ecosystem states, ranging from predominantly sand and other low relief habitats to locations with greater amounts of hard substrate covered with varying degrees of coral and macroalgae.

Float Construction and Field Methods

To simultaneously survey fish and benthic assemblages over large areas in shallow subtidal habitats, we constructed a float that was towed behind a snorkeler conducting a fish count. The float captured photos of the benthos and obtained a GPS track for each survey. The float was constructed by adding PVC crossbars (~ 70 cm long) to the bottom of an inflatable, ~ 60 cm diameter dive safety "inner-tube" float (JC Scuba, Inc., Warwick, Rhode Island, USA) to provide a stable attachment point for a GoPro Hero 5 Black camera (San Mateo, California, USA), which was oriented directly downward (**Fig. 1a**). We also attached a small 1.4 kg dive weight to the PVC to stabilize the crossbars. Inside the float, we placed a Garmin GPSMap 78 handheld GPS (Olathe, Kansas, USA) contained inside a waterproof case, which was set to record a track of georeferenced locations at 5 s intervals. Prior to leaving the dock, the GPS was turned on, the GPS track was cleared, and the unit was placed into the waterproof case, remaining there until returning to the dock upon which it was removed from the case and the track was saved. We chose this GPS unit because it is a newer model of a unit that was successfully used to georeference fish surveys in prior work (Beck et al. 2014), and GoPro cameras were chosen due to their low cost, ruggedness in the field, and wide implementation in a variety of underwater research projects involving photography, including benthic surveys (Koester et al. 2020) and photogrammetry (Nocerino et al. 2020). The float was towed by a snorkeler via a 3 m long rope (the length is necessary to prevent turbulence from fin kicks impacting the quality of images; **Fig. 1b**). The snorkeler wore a digital, waterproof wristwatch (Timex Ironman Classic 30; Middlebury, Connecticut, USA) whose time was synchronized to the internal time of the GPS every 48 hrs to account for slight drift in the watch time away from that of the GPS. On the boat prior to beginning a transect, the GoPro was set to time lapse mode (Linear FOV, 0.5 s interval), and the observer began the time lapse, photographing the watch face (**Fig. S1**) for approximately 10 s before attaching the GoPro to the float, attempting to ensure that the watch face was illuminated by sunlight to make identifying time in the photographs easier; because photos were taken at fixed increments, each subsequent photograph could later be associated with an exact time and thus with coordinates from the GPS track. Each watch was set to produce a looping alarm at 1-min intervals to allow finer-scale binning of fish counts (these 1-min long bins are hereafter referred to as “transect segments”), with the alarm signaling the beginning of a new transect segment. To conduct the fish survey, the snorkeler swam for up to 30 min at a speed of roughly 10 m per min, writing fish observations in a new row on a datasheet based on the current transect segment. Surveys were conducted between the hours of 10:00 and 16:00 on days of no rain and low cloud cover to limit differences in fish behavior among surveys. A consistent swim speed within transect segments is necessary to limit bias in the resulting data; swimming slowly over one area for half of a segment and swimming quickly in the other will increase the percent cover estimates from the slowly-swum area relative to the quickly-swum area. Although this can be corrected by sub-sampling images taken a set distance apart from one another after the photographs are georeferenced, maintaining a similar swim speed in the field provides the easiest solution to this problem. Different swim speeds between transect segments should not suffer this bias due to the fish surveys being standardized to the area of each transect segment and the benthic surveys being standardized to the same number of points between segments. While the observer counted fish, the GoPro was passively taking thousands of images of the benthos, and these photographs are clear enough to identify key reef substrates and benthic organisms (**Fig. 1c,d**). At the end of each survey the snorkeler swam an additional 30 s forward to capture photographs of the benthos in which fish were counted in the final minute of the survey.

The surveys allow coverage of a large area within relatively short time periods. For example, in 2018, our island-wide campaign in Moorea surveyed approximately 225,000 m^2 of reef in only ~ 67 person-hours of sampling (**Fig. 2a**). We surveyed different lagoon habitats, including the mid-lagoon, fringing reef, along the shallow edges of reef passes, and just behind the reef crest. (**Fig. 2b**). The snorkeler counted and visually estimated the total length (TL) to the nearest cm of each fish ≥ 10 cm from a fixed species list (57 taxa, including mobile herbivores and important fishes in the local fishery; **Table S1**; Rassweiler et al. 2020) in a 5 m wide swath. We did not include small individuals or certain cryptic taxa (e.g., Synodontidae, the lizardfishes) in this species list due to low confidence in our ability to accurately estimate them from the surface while swimming at our speed of 10 m per minute. There are tradeoffs in the amount of area covered and the taxa surveyed; by focusing on larger, more mobile individuals and taxa, more area can be covered, while including smaller, more cryptic individuals would force a slower swimming speed to properly estimate their abundances. The approach described here is appropriate for research questions focused on larger fish. Binning these counts by minute leads to 30, 1-min-long transect segments, which can be used to obtain fine-scale spatial information (~ 10 x 5 m surveyed per transect segment) on the location of counted fishes (**Fig. 2c**). This technique generated fish counts and fish sizes, paired with 90 benthic photographs taken during every minute of fish survey.

Georeferencing Survey Data

Data from the fish and benthic surveys were georeferenced using the recorded times for the GPS. GPX (GPS Exchange Format) tracks from the GPS were accessed using the “PlotKML” package in R (Hengl et al. 2015), and the times associated with each point of the track were converted to local time. We assigned a time to every photograph in a transect by extrapolating times from a photograph of the watch face taken at the beginning of the transect, calculated based on a constant interval between each photograph in the time lapse. We ran tests to identify the true GoPro timelapse interval and to validate our method by recording a digital watch until the GoPro battery expired and identifying the average number of photographs taken during this time period.

These tests revealed that the GoPro's nominal 0.5 s interval actually took a photo every 2/3 of a second, and by assuming this value as the interval, we found that by the end of a time lapse trial (trial length range: 76.7 – 85.5 minutes), the mismatch between the time we associated with the photographs and the true time on the watch face was less than 1 s in all trials (n = 3). This behavior was consistent across multiple GoPro cameras. The timing of the fish transect segments were defined by the known fixed 1-min intervals. However, because the GPS and camera were towed behind the observer, and fish were counted just ahead of the observer, we performed a spatial correction to match the GPS coordinates more closely with the locations the fish surveys were conducted on the reef. For the fish data, we took the beginning and end coordinates for each transect segment recorded by the GPS and calculated a directional bearing using the "geosphere" package in R (Hijmans et al. 2017) and used this to project the coordinates from the GPS 4 m forward (3 m for the float to observer and 1 m for the distance from the observer's waist to their eyes) along this bearing. The two points at the beginning and end were used to calculate a rectangular polygon 5 m wide to identify the observation window (**Fig. 2c**). To match the benthic photographs with this shift, photos associated with each fish count minute were lagged 22 seconds behind the fish counting, because this on average equates to a 4 m distance forward (i.e., the first image associated with a fish count minute beginning at 12:00:00 would be the image taken at 12:00:22, and the last associate image would be taken at 12:01:21). Although we could have matched the benthic photographs and the fish surveys more directly by associating photographs with a transect segment based on whether their coordinates were located within the observation polygon of each minute, this would have led to different numbers of photographs assigned to each transect segment. As our calculated locations on the reef are estimates within a few meters due to the error of our handheld GPS, we decided that a consistent number of replicate images within a minute was more important than slightly more accurate spatial matching with the fish survey observation window. However, depending on specific research questions (such as when precise location data are imperative), it is possible to aggregate photographs based on location rather than time by selecting photographs that fall within the fish survey polygons.

Sample R code and associated documentation to perform these procedures can be found in the Supplemental Information.

CoralNet Training and Benthic Data Analysis

Our 2018 sampling campaign in the lagoons of Moorea generated 330,660 photographs of the benthos across 3966 transect segments nested within 152 fish survey transects (each up to 30 min in duration). To begin analysis of the benthic photographs, two photographs from every transect segment were uploaded to CoralNet 1.0 (coralnet.ucsd.edu; Chen et al. 2021), an online image repository and machine learning automatic classifier tool (Beijbom et al. 2015; Williams et al. 2019), for a total of 7932 images. Trained researchers scored 30 stratified random points per photograph; points were not placed within 15 % of the edges of photographs, as objects near the edge can be out of focus due to the fisheye lens of the GoPro camera. Using a different camera may allow points to be accurately annotated up to the edges of the photograph, but as this was done across all photographs, we do not expect it to introduce any bias in this case. We chose 30 points as an acceptable tradeoff between time required to annotate an image and the precision of percent cover estimates for each image. Scoring only 30 points per image increased the number of images we could process; utilizing more images rather than more points per image has been found to improve statistical power in previous studies (Houk and Van Woesik 2006; Perkins et al. 2016). Each point was classified by the observer into one of 36 benthic substrate categories, including several categories representing points in which the substrate could not be identified (e.g., motion blur, bubble interference, too deep, etc.), but ultimately only 17 of the 36 were common enough to be included in our validation procedure (**Table 1**). The 19 of our original categories that were not included in the validation procedure only comprised 1.3% of the total number of points used to train the CoralNet classifier, so although these are a high number of categories, they represent rare occurrences on the reef. These categories were selected because they represent the dominant benthic taxa and morphotypes present in the lagoon of Moorea, comprising ~95% of the coral cover in available MCR LTER lagoon data in 2018 (Edmunds 2022), but researchers in higher-diversity systems will likely require more categories, which may negatively impact the ability of the CoralNet system to distinguish among taxa.

We used the two annotated images per transect segment to train CoralNet's automatic classifier. This training triggers automatically without additional user input once 20 images are annotated by a human and continually attempts to retrain to greater accuracy as more images are annotated. During training the classifier uses the images annotated by a human to learn how to identify substrates automatically, and once this training is completed it begins making guesses on points from newly uploaded images. Our classifier achieved 84% average accuracy at the point level across all substrates based on CoralNet's internal metrics, and this final classifier was used to automatically annotate photographs for the validation procedure described in the Assessment.

Assessment

Comparison of GPS-Enabled Snorkel Fish Surveys to Fixed-Area SCUBA Surveys

To compare the biomass and density values of the fish counts with more commonly used methods, we utilized the MCR LTER time series data on the lagoon fish community, which are obtained from scuba-based, fixed-area permanent transects (n = 4, 50 x 5 m transects per site) at six back reef and six fringing reef locations located around the island (Brooks 2019). We compared data on fish abundance and biomass obtained using our method with data collected in the same year (2018) by the MCR LTER project. To make the comparisons as similar as possible, we filtered both datasets to include only back reef habitats and the LTER dataset to only include individuals greater than 10 cm TL of the species we counted. Because we used a truncated species list in our method, we limited comparisons to functional groups of herbivorous fishes (browsers, scrapers, excavators, grazers, and detritivores) that are of ecological interest and often aggregated in this way (Han et al. 2016; Schmitt et al. 2021). Herbivores were selected because as a group they constituted the majority of biomass observed in the back reef habitat in 2018 when sharks are excluded (~58% of biomass; Brooks 2019). Due to their tendency to form large, highly mobile and spatially patchy schools, we excluded the grazer, *Acanthurus triostegus,* from both datasets to avoid a single school influencing the comparisons (Han et al. 2016). We then selected transect segments from our 2018 sampling (n = 483 total) that were within 500 m of the center of the LTER fixed transects (n = 5 back reef sites, because we did not have transects located within 500 m of one of the LTER sites), which helped minimize potentially confounding effects driven by spatial heterogeneity of the reef environment. Values for average biomass of each functional group from the datasets at each location were calculated by converting observed lengths to biomass using length-weight relationships retrieved from online and published sources (Kulbicki et al. 2005; Brooks & Adam 2019). We plotted the mean biomass and density of each functional group in both methods, and we calculated Pearson's correlations using all functional groups for both metrics. Additionally, we show how the mean biomass and density (calculated as mean at the site level, n = 5) of each functional group compares between the methods. We ran linear models with biomass and density as the response variables and method and functional group as interacting explanatory variables and performed an ANOVA to determine the effect of method, functional group, and their interaction on both metrics. To test the applicability of the method to other taxa, we used the same approach to evaluate four non-herbivore families in which all species were counted in both our method and the LTER surveys (Mullidae, Serranidae, Carangidae, and Lutjanidae). These species are all mobile and ambush predators.

CoralNet Validation

Because we captured many more images than can be realistically hand-annotated by even the best resourced groups of researchers, we wanted to examine whether the automated classifier could replace human annotators in estimating the percent cover at the scale of a transect segment (roughly 50 m^2 of reef). The accuracy of estimates at this scale are a function both of point-by-point accuracy in identification of substrates, and of the number of images that are scored, which raises a potential tradeoff between the automated and the human approach. The human is equally or more accurate on a point-by-point basis, but the manual-annotation bottleneck limits the number of images processed resulting in low spatial coverage of the site, while the algorithm is often less accurate at the point level but can score enough images to nearly capture the entire benthic makeup of the transect, potentially providing a better representation of the ‘true’ percent cover of the site (e.g., Griffin et al. 2017).

To compare the ability of humans and the algorithm to estimate the benthic composition of the reef, we randomly selected 50 transect segments, 40 selected at random from unique full transects conducted on the back reef and 10 selected from each of 10 fringing reef transects. We limited our analyses to transects from fringing and back reef habitats because of the relatively consistent depths (< 3 m). We then allocated 21 of the 90 photos from each transect segment to be annotated by humans and the remaining 69 photos to be automatically annotated by CoralNet. The manually-annotated photos were evenly spaced across the minute and were scored by the same group of trained human annotators that trained the classifier. The other 69 photos were automatically annotated using our previously trained classifier.

We randomly selected 12 of the manually-annotated images from each transect segment to calculate the percent cover of each substrate at the site. This was used as a reference set for comparison with other sets of manually-annotated and automatically-annotated images from the same transect segment to assess the increase in accuracy that could be achieved by annotating larger numbers of images. We developed a resampling procedure to compare the effectiveness of using different numbers of manually-annotated images versus automatically-annotated images at estimating this reference percent cover within each transect segment. First, we randomly selected one of the 9 remaining images scored by a human for all 50 transect segments (i.e., an image not in the reference set), and calculated the mean percent cover of the substrate of interest for each transect segment. We then generated a regression using the mean percent cover in each transect segment (in this example based on a single hand-annotated image) to predict the percent cover calculated from the set of twelve reference images (**Fig. S2**). If the sample estimate perfectly predicted the percent cover from the reference set in all transect segments, the r^2, slope, and intercept of this regression would be 1, 1, and 0, respectively, thus the closer the observed results are to these values, the more accurately the sample predicted the reference set. For each substrate of interest, we repeated this process 1000 times for different numbers of manually-annotated images (n = 1 to 8) and automatically-annotated images (n = 1, 2, 4, 8, 16, 32, 64) using different sample images in each replicate (**Figs. S3-S5)**. To ensure that observed results were not driven by the details of the randomly chosen reference set, this process was repeated an additional 100 times with different randomly chosen reference sets, resulting in n = 100,000 regression results for each substrate for each number of manually-annotated and automatically-annotated images. Importantly, this process uses separate images to calculate the percent cover in the reference set and the sample set in each iteration, which allows us to infer how well different numbers of images from each classification method perform at estimating percent cover in an unknown reference set, and thus how accurately they can describe variation in ecosystem state at this spatial scale. We performed this analysis on each class of substrates and organisms identified, as well as on aggregated classes that are more likely to be used in broader-scale ecological analyses (**Table 1**). It is relevant to note that these aggregated classes were combined post-annotation from the same CoralNet source (i.e., we did not train new classifiers on aggregated label sets) and training a new classifier directly on the combined categories may have increased the performance (Williams et al. 2019).

For each regression generated in the resampling procedure, the r^2, slope, and intercept were determined, and for visualization, the mean r^2 and 95 % quantiles of these were plotted for each substrate across each number of manually- and automatically-annotated images. As a summary, we plotted just the mean r^2 values for 8 manually-annotated images, 8 automatically-annotated images, and 64 automatically-annotated images for each substrate to visualize the mean effects for key values of our analyses.

As one of the main benefits of using the algorithm to automatically annotate images is the reduction in time spent processing images, we wanted to compare how many person-hours it would take to annotate one hundred 30-min long transects using our method versus the manual annotation approach. To maintain consistency with our evaluation, we performed these calculations using 8 manually-annotated and 64 automatically-annotated images per transect segment. We assumed that it would take a human two minutes on average to annotate an image in our dataset (a speed equivalent to that of an experienced annotator). We then assumed that image processing takes approximately 20 minutes per 5000 images in administrative time (uploading images, downloading data, etc.) based on records from our usage of the CoralNet platform. For the CoralNet approach, we assumed that it took 3485 images to train the classifier (as it initially did in our project) and that these training images are not used in the final dataset, thus the time spent training the classifier is a time "cost" and represents an intercept in a model of time spent per transect. Due to additional effort required to upload and process the higher volume of images, we assumed that it takes approximately 30 minutes per 5000 images in administrative time. Although CoralNet does not instantly annotate the images, we did not consider this waiting time in the calculations, because this does not represent human time required to generate the data, although such waiting is relevant for project timelines. These two functions were plotted and we calculated the intercept where the two are equivalent in efficiency. R code and a more detailed walkthrough of this validation procedure are included in the Supplemental Information.

All statistical analyses and maps were created using R 3.6.0 (R Core Team 2019).

Results

Comparison of GPS-Enabled Snorkel Fish Surveys to Fixed-Area SCUBA Surveys

Despite significant differences in methodology between our GPS-enabled snorkel surveys and fixed-area scuba transects conducted in 2018 by the MCR LTER project, the two methods provided similar results with respect to the biomass and density of the 5 major functional groups of herbivorous fish (**Fig. 3**). When comparing our snorkeling transects situated within 500 m of each of the LTER back reef sites, both the biomass ($r = 0.64$, $p < 0.001$; **Fig. 3a**) and density ($r = 0.54$, $p < 0.01$; **Fig. 3b**) were positively and significantly correlated based on Pearson correlations. There was some tendency for the GPS-enabled counts to yield a greater biomass of grazers and scrapers, and for the fixed-area scuba counts to feature a greater density of excavators (**Fig. 3c,d**). However, the ANOVA test showed that while average biomass and density of fish functional groups differed ($p < 0.005$ for both metrics), there was no significant effect of sampling method on either metric, nor a significant interaction between method and functional group. Results for the four carnivorous families were similar to our analysis of herbivorous groups, with both biomass ($r = 0.75$, $p < 0.001$; **Fig. S6a**) and density ($r = 0.78$, $p < 0.001$; **Fig. S6b**) being positively correlated between the two methods. Our snorkel method tended to find a greater biomass of Mullidae (**Fig. S6c**), while the fixed-area scuba counts found greater density of Serranidae (**Fig. S6d**). However, the ANOVA results showed the same pattern as for herbivores; the biomass and density of the carnivore families differed ($p < 0.005$), but there were no significant effects between sampling methods or the interaction of sampling method and family. Overall, differences between methods were minor, particularly given the differences in the area covered by the two types of surveys.

CoralNet Validation

For all substrates, the accuracy in estimating the reference image set increased as we used more images to estimate benthic cover within a transect segment; the mean r^2 value of comparisons with the reference dataset increased and approached 1, and the variance of those r^2 values decreased (**Figs. 4a, Figs. S7-27; Table 2**). The automatically-annotated images tended to perform similarly to the manually-annotated images across the range of values where both manually- and automatically-annotated images were tested (n = 1 to 8 images) (**Figs. 4a**, **Figs. S7-27**). However, when 16 or more computer-annotated images were used, performance continued to increase for most substrates. For brevity, only the results for a single substrate (live coral) are depicted in **Fig. 4**, but other substrates show a qualitatively similar pattern and are provided in the Supplemental Information (**Figs. S7-27**). In addition to providing greater accuracy for many substrates, using CoralNet to annotate 64 images per minute represents significant time savings over annotating 8 images per minute by a human when many transects were surveyed (**Fig. 4b**). The intersection of these lines occurs at approximately 447 transect segments, representing just under 15 full-length (30-min) transects; below this sampling effort manual annotation by humans would be quicker, but beyond this level the CoralNet approach provides significant time savings, increasing dramatically as the number of transects becomes larger (**Fig. 4b**).

When 8 manually- and automatically-annotated images were compared, the automatically-annotated images outperformed (i.e., had a higher mean r^2 than) the manually-annotated images in 13/17 (76.5%) of substrates. When 64 images per minute were automatically annotated, CoralNet's ability to estimate the percent cover of the reference image set was boosted and it outperformed the 8 manually-annotated images in 16/17 (94.1%) substrates (**Fig. 5a; Table 2**). The lone substrate that the manually-annotated images outperformed the automatically-annotated images was non-branching *Acropora.* Non-branching *Acropora* is rare on the reef, so this substrate tends to be misidentified as other corals by the algorithm, particularly branching *Acropora* and *Pocillopora*, likely due to a low amount of training data. Substrates common on the reef (and thus in the training data) were generally estimated better (i.e., had a higher mean r^2) than rare substrates, and this was true for both manually- and automatically-annotated images (**Fig. 5a**). When substrates were aggregated into four major groups (Soft Substrate, Hard Substrate, Macroalgae, Coral), the overall patterns remained the same (**Fig. 5b; Table 3**). All four aggregated substrate categories had a mean r^2 value of greater than 0.85 when scored by 64 automatic images and outperformed estimates resulting from 8 manually-annotated images (**Fig. 5b**).

Discussion

We described a novel methodology for collecting paired, georeferenced surveys of fish and benthic communities simultaneously by a single observer in a cost-effective, time-efficient manner. We also showed how the resulting data could be efficiently and accurately processed to determine benthic community composition. After an initial investment in time (to train the CoralNet classifier) and cost (to construct the camera float – roughly $700 USD per float), this method allows a small group of researchers to significantly scale up surveys in shallow subtidal systems with minimal marginal cost beyond the time spent in the field. The method is specifically designed to use materials that are lightweight, easy to assemble, and cost-effective, and it accomplishes this with little sacrifice to accuracy. However, if greater accuracy (such as increased taxonomic resolution or more precise locational data) is required, the framework presented here can easily be modified to achieve those goals. For example, use of a differential

GPS rather than the handheld GPS we employed could potentially yield cm-scale georeferencing accuracy, and upgraded camera systems could obtain higher quality images, improving the classification potential of CoralNet. Even more sophisticated elaborations can be imagined, such as direct circuitry linking the GPS and the camera, eliminating the need to assign a location to each photograph by extrapolating time based on the time lapse interval. Thus, our already-promising assessment of the method can be thought of as a conservative estimate of the potential for this general approach.

Our methodology is an extension of prior work investigating the statistical properties of point count analyses, which found that increasing the number of images provides an improvement in accuracy at estimating benthic cover, particularly in heterogeneous environments (Perkins et al. 2016). While using computer vision algorithms to scale up surveys is not a new idea in the ecological literature (e.g., González-Rivero et al. 2016; Griffin et al. 2017), our study provides an explicit test of the tradeoffs between using different numbers of manually- and automatically scored images to assess unknown benthic cover at ecologically-relevant scales. Our findings suggest that leveraging computer vision algorithms can do more than simply alleviate the “manual-annotation bottleneck” from photoquadrat ecological surveys by simultaneously increasing the scale and accuracy at which ecological data can be collected. Because the accuracy improves with most categories as more photographs are annotated, and more photographs are efficient to both collect and analyze in a computer vision framework, our method scales up to cover greater area than traditional surveys with little-to-no tradeoff in efficiency in terms of cost or accuracy. Therefore, our validation procedure is more than just an assessment of the CoralNet system. We show that our method unlocks new data collection options, both the paired fish and benthic surveys demonstrated here and potentially through the use of different methods specific to benthic data, such as autonomous camera systems (e.g., Raber and Schill 2019).

Our validation procedure indicates that the CoralNet system performs well when estimating cover of the benthos compared to a set of reference images taken from the same transect segment. However, it is worth noting that these reference images were collected using the same surface-borne camera, thus the estimates of benthic cover derived from the reference set are subject to any limitations of the photos themselves, including low resolution at greater depths and the inability to discern cryptic benthic classes. Further validation could be performed by using *in situ* validation of benthic substrates or through comparing these surface-borne photographs to traditional SCUBA-based photoquadrats. Because of the cost of sampling benthic habitats, many research projects have focused narrowly on a single habitat type (e.g., structurally similar hard bottom habitats). The efficiency of our method permits sampling a wider array of habitats with the same effort, potentially giving a more realistic description of the overall seascape. For example, we record a much higher (and likely more representative) percentage of sand and coral rubble in our transects compared to the LTER surveys (52.9% vs. 31.1%) which were intentionally stratified to be in areas dominated by hard-bottom habitat (Carpenter 2020). Sand cover is likely an important variable to consider when describing fish assemblages because many key fish groups respond to physical structure (Holbrook et al. 2002a,b; Schmitt et al. 2007). Researchers in areas with greater resources and access to airborne drones, LiDAR surveys, or high-quality remote sensing products may be able to use those data sources to estimate reef structure and sand across greater extents (e.g., Casella et al. 2017; Collin et al. 2018), but our method provides these data at the same scales as the associated fish surveys with relatively low additional time and monetary investment. Further, many remote sensing techniques classify individual pixels to a single class (and may have difficulty distinguishing between certain reef taxa, such as coral and brown algae), so our continuous percent cover estimates provide useful information that these methods in their current state cannot.

Although our methods were developed and validated in a shallow coral reef system, the methodology described is modular and the different pieces can be isolated and used for other purposes. For example, the georeferenced fish survey methods could be used for any shallow-water system where visual estimates are used, and our method of aggregating photographic data is even more broadly applicable. One application could be the development of autonomous surface vehicles that take similar surface-borne photographs to map other shallow subtidal systems (Raber and Schill 2019), or it could be used to generate ground-truthed field data to improve satellite-derived habitat classifications (Roelfsema and Phinn 2010). We chose to aggregate at the level of a minute of swimming, but finer or broader aggregations can be made, depending on the level of precision versus spatial coverage required.

Considerations and Limitations

Utilizing this method unlocks the ability to address ecological questions at larger spatial scales than traditional site-based methods, but this comes with several considerations, tradeoffs, and limitations. The end product is a dataset of spatially-explicit transect segments. While the collection of segments can cover an extensive stretch of reef, each segment represents a fairly small area, describing benthic cover and fish abundance therin. Considering the area of these transects segments is small and there are up to 30 per transect, they can be aggregated into coarser units of replication to address ecological questions related to scale (e.g., a multiscale approach) and can be layered with other spatially-explicit data sources, such as bathymetry, to explore distributions of organisms or as predictor variables in other analyses (Pittman & Brown 2011). One important concern moving forward would be that of autocorrelation; although the lowest unit of replication are individual transect segments, they are adjacent along a transect and thus do not represent statistically independent observations (Legendre 1993). Whether this is merely an issue that needs to be accounted for or it is used to elucidate additional properties of the system depends on the research question (Legendre 1993; Lennon 2000; Diniz-Filho et al. 2003), but in any case, it should be considered when designing a study that implements the method described in this article.

In addition to the implementation concerns above, there are more direct limits to its application in the field. The most obvious limitation is that of depth; benthic substrates at depths below ~ 2-3m become more difficult to discern from the surface, even in the clear waters of our coral reef system. It is for this reason that we limited our analysis to fringing and back reef habitats. We had collected some data from deeper, steeply sloping reef passes and fringing reefs but excluded these due to low confidence in benthic characterization. However, for many coastal marine systems, such as shallow lagoon habitats and seagrass beds, the shallower depths are relevant. For example, in our surveys, ~89% of total transect segments (i.e., before removing those that were too deep) had a mean depth shallower than 2 m, with ~98% being shallower than 3 m, and LiDAR bathymetry of Moorea shows that ~81% of the reef within the lagoons is shallower than 3 m. In our data, transect segments between 2-3 m had similar total percent cover for most substrates to transects shallower than 2 m, suggesting that the identification methods remained reliable down to 3 m (**Table S2**). It is possible that estimating percent cover deeper than 3 m may be possible depending on local taxa and water clarity, but we do not have the ability to examine these depths with our current data and therefore offer a conservative recommendation of utilizing this approach in areas shallower than 3 m.

Even in shallower water, the ability to fully sample both fish and benthic communities is limited due to our faster swim speed compared scuba-based survey methods and due to the location of the researcher at the surface. We chose to swim at a target speed of 10 m per minute because this quicker, more mobile approach allows better quantification of rare and large fish species which were of primary concern for our research questions and tend to be underrepresented in traditional surveys (Richards et al. 2011). This speed could be slowed to focus on surveying smaller individuals or different taxa. But even with a slower swim speed, it seems unlikely that proper quantification of cryptic species or smaller individuals will be possible given the position of the observer and camera at the surface, due to cryptic fishes being obscured within the reef matrix. These limitations could be overcome in some situations if a scuba diver working at depth pulled a GPS float deployed at the surface.

Considering that the surveys are conducted from the surface, conditions on the surface can influence the quality of data collected. Notably, strong wind or currents (which are predominantly wave-driven in our system) can push the float around so that it is not directly behind the researcher, leading to mismatches of a few meters in the georeferencing of the surveys. If conditions are too rough, the researcher must focus too much on maintaining direction and speed, which can impact their ability to properly survey fishes. This caused us to avoid collecting data when winds were > ~18 knots and when swells were > 2.5 m on the reef crest behind which we were surveying (with occasional cancellations at values weaker than those based on local conditions). Turbidity in the water column could also influence the quality of data collected. We found that the auto white balance of our GoPro cameras was sufficient to capture quality photographs of the benthos as long as visibility was high enough to conduct a fish survey, although higher turbidity would likely reduce the depths this method is applicable even further. Turbidity was not an issue for us most of the time and was only an issue nearshore, but we did have to cancel a survey due to visibility concerns (horizontal visibility < ~5 m) when we attempted to survey the fringing reef after a rainstorm.

Additionally, although our results indicate that the CoralNet system performs well on our surface-borne photographs, we caution against moving forward with automatic annotation of entire datasets without first undergoing a validation procedure in each system. Our results demonstrate that the benthic photographs perform well in Moorea, but this is a relatively low diversity system and higher diversity areas may require additional modifications to properly survey. These modifications may include a slower swimming speed to properly enumerate a greater diversity of fishes, or it may require more training of the benthic classifier to properly identify increased diversity on the benthos than was necessary in Moorea's lagoons. A properly trained source can perform well under similar conditions it was trained (i.e., based on our source's performance across the different habitats of Moorea, we have confidence that it would do well in the lagoons of neighboring Tahiti), but it would likely require extensive training or starting fresh if applied to a location with different taxa (e.g., the Great Barrier Reef). Thus the applicability to new locations depends on the local taxa and conditions. In our experience, the CoralNet system can be retrained and expanded to include additional benthic categories if the community changes within a system. For example, in subsequent work with the CoralNet system after a bleaching event in 2019 we were able to add bleached corals to a new CoralNet automatic classifier by uploading additional photographs with bleached corals present and the associated manual annotation files to add these categories to the CoralNet source.

To assist with the hurdle in performing site-specific validations, we provide R code and a straightforward, general process for conducting this validation. Our validation procedure requires some additional time investment, with 21 photographs per minute used in the validation procedure to be manually-annotated (n = 1050 images for the 50 transect segments in this study), but this is small compared to the number of photographs needed to initially train the classifier and the long-term efficiency achieved once validation is completed (**Fig. 4b**). Not only does independent validation allow estimates of accuracy and bias uniquely tailored to each system, but it also ensures that some human expertise is involved in the process. Although here we argue for increased automation in the annotation of benthic imagery, humans should remain involved throughout the process in some capacity (Portelli et al. 2020). Without proper validation and/or spot-checking of data by a trained human observer, the procedure could generate erroneous data, either due to human error during processing steps or poor performance by the computer vision system.

Summary

Here we present a novel methodology that allows a single user to simultaneously conduct georeferenced surveys of fish and benthic communities with minimal material and time investment. We demonstrate that both components of the method provide accurate estimations of the fish abundance/biomass and benthic cover of conspicuous categories in a shallow coral reef system. The spatially-explicit nature of the data allows straightforward layering with other spatial data sources, and the binning of transects into minute-long segments enables fine-scale modeling that can be aggregated at different scales depending on research needs. Further, aspects of this methodology could be combined with existing subtidal survey approaches (e.g., González-Rivero et al. 2014) to expand the depths accessed. We believe our method fills a distinct and useful gap within the pantheon of sampling techniques, falling between fixed-area surveys and remote sensing along the continuum of tradeoffs in area sampled and taxonomic resolution. While the method we describe cannot fully replace traditional monitoring protocols that are able to survey entire benthic and fish communities, particularly if smaller or cryptic taxa are the focus of interest, it could be widely-implemented in research programs where the increased spatial coverage offered or the ability to obtain georeferenced surveys outweigh the need for fine-scale taxonomic information.

Data Availability Statement

The version of data and code referenced in this study are archived and openly available in Zenodo at http://doi.org/10.5281/zenodo.7920817. The associated code may be updated in the future, and the latest version can be found in the author's Github at https://github.com/SMillerTime57/GPS_fish_benthic_surveys.

Author Contribution Statement

SM, TA, SH, RS, and AR conceived the method. SM, TA, and DC collected field data. SM, AK, and AR led photographic analysis. SM wrote code and led the writing of the manuscript. All authors contributed critical feedback throughout the development of the method and the validation procedures, assisted with the writing of drafts, and approved the final manuscript.

Acknowledgements

We gratefully acknowledge the team at CoralNet for the creation of the tool and their assistance throughout this project. We thank A. Turner, L. Turner, A. Siceloff, P. Bengston, and J. Baker for their assistance in annotating images, K. Kopecky for assistance in the field, and S. Lester for helpful discussions while developing the procedures. This work was funded by NSF BCS 1714704 and OCE 1637396, and the CoralNet platform was funded through various grants from NOAA and NSF. We thank the Moorea Coral Reef LTER site and the University of California Gump Research Station for logistic support, particularly A.J. Brooks for helpful discussions regarding the LTER fish surveys. Research was completed under permits issued by the Territorial Government of French Polynesia (Délégation à la Recherche) and the Haut-commissariat de la République en Polynésie Française (DTRT) (Protocole d'Accueil 2015-2018); we thank them for their continued support. The authors disclose no conflicts of interest.

References

Adam, T.C., D.E. Burkepile, S.J. Holbrook, R.C. Carpenter, J. Claudet, C. Loiseau, L. Thiault, A.J. Brooks, L. Washburn, and R.J. Schmitt. 2021. Landscape-scale patterns of nutrient enrichment in a coral reef ecosystem: implications for coral to algae phase shifts. Ecol. Appl. 31(1): e02227. doi:10.1002/eap.2227.

Beck, H.J., D.A. Feary, W.F. Figueira, and D.J. Booth. 2014. Assessing range shifts of tropical reef fishes: a comparison of belt transect and roaming underwater visual census methods. Bull. Mar. Sci. 90: 705-721. doi:10.5343/bms.2013.1055

Beijbom, O., P.J. Edmunds, C. Roelfsema, and others. 2015. Towards automated annotation of benthic survey images: variability of human experts and operational modes of automation. PLoS ONE 10(7): e0130312.

Bellwood, D.R., T.P. Hughes, C. Folke, and M. Nyström. 2004. Confronting the coral reef crisis. Nature 429: 827-733.

Boström, C., S.J. Pittman, C. Simenstad, and R.T. Kneib. 2011. Seascape ecology of coastal biogenic habitats: advances, gaps, and challenges. Mar. Ecol. Prog. Ser. 427: 191-217.

Brooks, A.J. 2019. MCR LTER: Coral Reef: Long-term population and community dynamics: Fishes, ongoing since 2005. knb-lter-mcr.6.58. doi:10.6073/pasta/a667eed481d9743c69c4209f6479acb4

Brooks, A.J., and T.C. Adam. 2019. MCR LTER: Reference: Fish taxonomy, trophic groups and morphometry. knb-lter-mcr.6001.6. doi:10.6073/pasta/f6feeebdbe44f3865ce4cd233a744b83

Caldwell, Z.R., B.J. Zgliczynski, G.J. Williams, and S.A. Sandin. 2016. Reef fish survey techniques: assessing the potential for standardizing methodologies. PLoS ONE 11(4): e0153066.

Carpenter, R., and Moorea Coral Reef LTER. 2020. MCR LTER: Coral Reef: Long-term population and community dynamics: Benthic algae and other community components, ongoing since 2005. knb-lter-mcr.8.32. doi:10.6073/pasta/0bf200e9e0f099de69826f57b18ff3da

Casella, E., A. Collin, D. Harris, S. Ferse, S. Bejarano, V. Parravicini, J.L. Hench, and A. Rovere. 2017. Mapping coral reefs using consumer-grade drones and structure from motion photogrammetry techniques. Coral Reefs 36: 269-275.

Chen, Q., O. Beijbom, S. Chan, J. Bouwmeester, and D. Kriegman. 2021. A new deep learning engine for CoralNet. International Conference on Computer Vision (ICCV) Workshops 2021: 3693-3702.

Collin, A., C. Ramambason, Y. Pastol, and others. 2018. Very high resolution mapping of coral reef state using airborne bathymetric LiDAR surface-intensity and drone imagery. Int. J. Remote Sensing 39: 5676-5688.

Colvocoresses, J., and A. Acosta. 2007. A large-scale field comparison of strip transect and stationary point count methods for conducting length-based underwater visual surveys of reef fish populations. Fish. Res. 85: 130-141.

Diniz-Filho, J.A.F., L.M. Bini, and B.A. Hawkins. 2003. Spatial autocorrelation and red herrings in geographical ecology. Glob. Ecol. Biogeogr. 12: 53-64.

Edmunds, P., and Moorea Coral Reef LTER. 2022. MCR LTER: Coral Reef: Long-term Population and Community Dynamics: Corals, ongoing since 2005. knb-lter-mcr.4.39 doi:10.6073/pasta/913b1532594fcaf02de7671ed940b4c5

Estes, L., P.R. Elsen, T. Treuer, L. Ahmen, K. Caylor, J. Chang, J.J. Chol, and E.C. Ellis. 2018. The spatial and temporal domains of modern ecology. Nat. Ecol. Evol. 2: 819-826.

González-Rivero, M., P. Bongaerts, O. Beijbom, and others. 2014. The Catlin Seaview Survey – kilometre-scale seascape assessment, and monitoring of coral reef ecosystems. Aquat. Conserv. 24: 184-198.

González-Rivero, M., O. Beijbom, A. Rodriguez-Ramirez, and others. 2016. Scaling up ecological measurements of coral reefs using semi-automated field image collection and analysis. Remote Sens. 8: 30. doi:10.3390/rs8010030

González-Rivero, M., O. Beijbom, A. Rodriguez-Ramirez, and others. 2020. Monitoring of coral reefs using artificial intelligence: a feasible and cost-effective approach. Remote Sens. 12: 489.

Gormley, K., F. McLellan, C. McCabe, C. Hinton, J. Ferris, D.I. Kline, and B.E. Scott. 2018. Automated image analysis of offshore infrastructure marine biofouling. J. Mar. Sci. Eng. 6(1), 2. doi: 10.3390/jmse6010002

Griffin, K.J., L.H. Hedge, M. González-Rivero, O.I. Hoegh-Guldberg, and E.L. Johnston. 2017. An evaluation of semi-automated methods for collecting ecosystem-level data in temperate marine systems. Ecol. Evol. 7: 4640-4650.

Han, X., T.C. Adam, R.J. Schmitt, A.J. Brooks, and S.J. Holbrook. 2016. Response of herbivore functional groups to sequential perturbations in Moorea, French Polynesia. Coral Reefs 35: 999-1009.

Hedley, J.D., C.M. Roelfsema, I. Chollett, and others. 2017. Remote sensing of coral reefs for monitoring and management: a review. Remote Sens. 8: 118. doi: 10.3390/rs8020118

Hengl, T., P. Roudier, D. Beaudette, and E. Pebesma. 2015. plotKML: Scientific visualization of spatio-temporal data. J. Stat. Softw. 63(5): 1-25.

Hijmans, R.J., E. Williams, and C. Vennes. 2017. Geosphere: spherical trigonometry. R package version 1.5-7. URL: https://CRAN.R-project.org/package=geosphere

Holbrook, S.J., A.J. Brooks, and R.J. Schmitt. 2002a. Predictability of fish assemblages on coral patch reefs. Mar. Freshwater Res. 53: 181-188.

Holbrook, S.J., A.J. Brooks, and R.J. Schmitt. 2002b. Variation in structural attributes of patch-forming corals and patterns of abundance of associated fishes. Mar. Freshwater Res. 53: 1045-1053.

Holbrook, S.J., J. Wencélius, A.K. Dubel, and others. 2022. Spatial co-variation in nutrient enrichment and fishing of herbivores in an oceanic coral reef ecosystem. Ecol. Appl. 32: e2515. doi:10.1002/eap.2515

Houk, P., and R. Van Woesik. 2006. Coral reef benthic video surveys facilitate long-term monitoring in the commonwealth of the Northern Mariana Islands: toward an optimal sampling strategy. Pac. Sci. 60(2): 177-189.

Koester, A., V. Migani, N. Bunbury, A. Ford, C. Sanchez, and C. Wild. 2020. Early trajectories of benthic coral reef communities following the 2015/16 coral bleaching event at remote Aldabra Atoll, Seychelles. Sci. Rep. 10: 17034

Kulbicki, M., G. Nicolas, and A. Marion. 2005. A general approach to length-weight relationships for New Caledonian lagoon fishes. Cybium 29: 235-252.

Lennon, J.J. 2000. Red-shifts and red herrings in geographical ecology. Ecography. 23: 101-113.

Levin, S.A. 1992. The problem of pattern and scale in ecology. Ecology 73(6): 1943-1967.

Lindenmayer, D.B., and G.E. Likens. 2010. The science and application of ecological monitoring. Biol. Conserv. 143: 1317-1328.

Nocerino, E., F. Menna, A. Gruen, and others. 2020. Coral reef monitoring by scuba divers using underwater photogrammetry and geodetic surveying. Remote Sens. 12(18): 3036.

Perkins N.R., S.D. Foster, N.A. Hill, and N.S. Barrett. 2016. Image subsampling and point scoring approaches for large-scale marine benthic monitoring programs. Estuar. Coast. Shelf Sci. 176: 36-46.

Pittman, S.J., and K.A. Brown. Multi-scale approach for predicting fish species distributions across coral reef seascapes. PLoS ONE 6(5): e20583.

Pittman, S.J., R.T. Kneib, and C.A. Simenstad. 2011. Practicing coastal seascape ecology. Mar. Ecol. Prog. Ser. 427: 187-190.

Pittman, S.J. 2017. Introducing seascape ecology, p. c3-25. *In* S.J. Pittman [Ed.], Seascape Ecology. Wiley.

Portelli, R.A. 2020. Don't throw the baby out with the bathwater: reappreciating the dynamic relationship between humans, machines, and landscape images. Landscape Ecol. 35: 815-822.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/

Raber, G.T., and S.R. Schill. 2019. Reef rover: a low cost small autonomous unmanned surface vehicle (USV) for mapping and monitoring coral reefs. Drones 3(2): 38.

Rassweiler, A., M. Lauer, S.E. Lester, and others. 2020. Perceptions and responses of Pacific Island Fishers to changing coral reefs. Ambio 49: 130-143. doi.org/10.1007/s13280-019-01154-5

Richards, B.L., I.D. Williams, M.O. Nadon, and B.J. Zgliczynski. 2011. A towed-diver survey method for mesoscale fishery-independent assessment of large-bodied reef fishes. Bull. Mar. Sci. 87(1): 55-74.

Rizzari, J.R., A.J. Frisch, and S.R. Connolly. 2014. How robust are estimates of coral reef shark depletion? Biol. Conserv. 176: 39-47.

Roelfsema, C., and S. Phinn. 2010. Integrating field data with high spatial resolution multispectral satellite imagery for calibration and validation of coral reef benthic community maps. J. Appl. Remote Sens. 4: 043527.

Sandel, B. and A.B. Smith. 2009. Scale as a lurking factor: incorporating scale-dependence in experimental ecology. Oikos 118: 1284-1291.

Schmitt, R.J. and S.J. Holbrook. 2007. The scale and cause of spatial heterogeneity in the strength of temporal density dependence. Ecology 88: 1241-1249.

Schmitt, R.J., S.J. Holbrook, S.L. Davis, A.J. Brooks, and T.C. Adam. 2019. Experimental support for alternate attractors on coral reefs. Proc. Natl. Acad. Sci. U.S.A. 116: 4372-4381.

Schmitt, R.J., S.J. Holbrook, A.J. Brooks, and T.C. Adam. 2021. Evaluating the precariousness of coral recovery when coral and macroalgae are alternative basins of attraction. Limnol. Oceanogr. 67: S285-S297. doi: 10.1002/lno.11929

Villon, S., D. Mouillot, M. Chaumont, E.S. Darling, G. Subsol, T. Claverie, and S. Villéger. 2018. A deep learning method for accurate and fast identification of coral reef fishes in underwater images. Ecol. Inform. 48: 238-244.

Williams, I.D., C.S. Couch, O. Beijbom, T.A. Oliver, B. Vargas-Angel, B.D. Schumacher, and R.E. Brainard. 2019. Leveraging automated image analysis tools to transform our capacity to assess status and trends of coral reefs. Front. Mar. Sci. 6: 222.

Zaneveld, J.R., D.E. Burkepile, A.A. Shantz, and others. 2016. Overfishing and nutrient pollution interact with temperature to disrupt coral reefs down to microbial scales. Nat. Commun. 7: 11833

Table 1. List of individual substrates used in the analysis, along with each substrate's aggregated group and a description. Only substrates identified by both humans and CoralNet in our validation procedure are shown here.

Individual substrate	Aggregated substrate	Description

Sand	Soft substrate	Sandy, fine substrate
Rubble	Soft substrate	Coral rubble (broken, detached dead coral)
Hard substrate	Hard substrate	Hard substrate (dead coral/rock, including crustose coralline algae, and low turf-algae)
Pavement	Hard substrate	Pavement (low relief hard substrate)
Turf	Hard substrate	Longer turfing algae (e.g., damselfish garden)
Obscured	Obscured	Unable to see substrate due to it being too dark, blurry, or otherwise covered (e.g., by a fish or a bubble)
Dictyota	Algae	Any algae in the genus *Dictyota*
Halimeda	Algae	Any algae in the genus *Halimeda*
Padina	Algae	Any algae in the genus *Padina*
Sargassum	Algae	Any algae in the genus *Sargassum*
Turbinaria	Algae	Any algae in the genus *Turbinaria*
Acropora - branching	Coral	Coral in genus *Acropora* showing branching morphology
Acropora	Coral	Corals in genus *Acropora* showing other growth forms (e.g., tabular)
Massive *Porites*	Coral	Corals in genus *Porites* showing massive growth forms (e.g., *Porites lobata*)
Montipora	Coral	Corals in genus *Montipora* showing any growth form
Pocillopora	Coral	Corals in genus *Pocillopora* (except for those identified as *P. damicornis*)
Porites rus	Coral	Corals of the species *Porites rus*

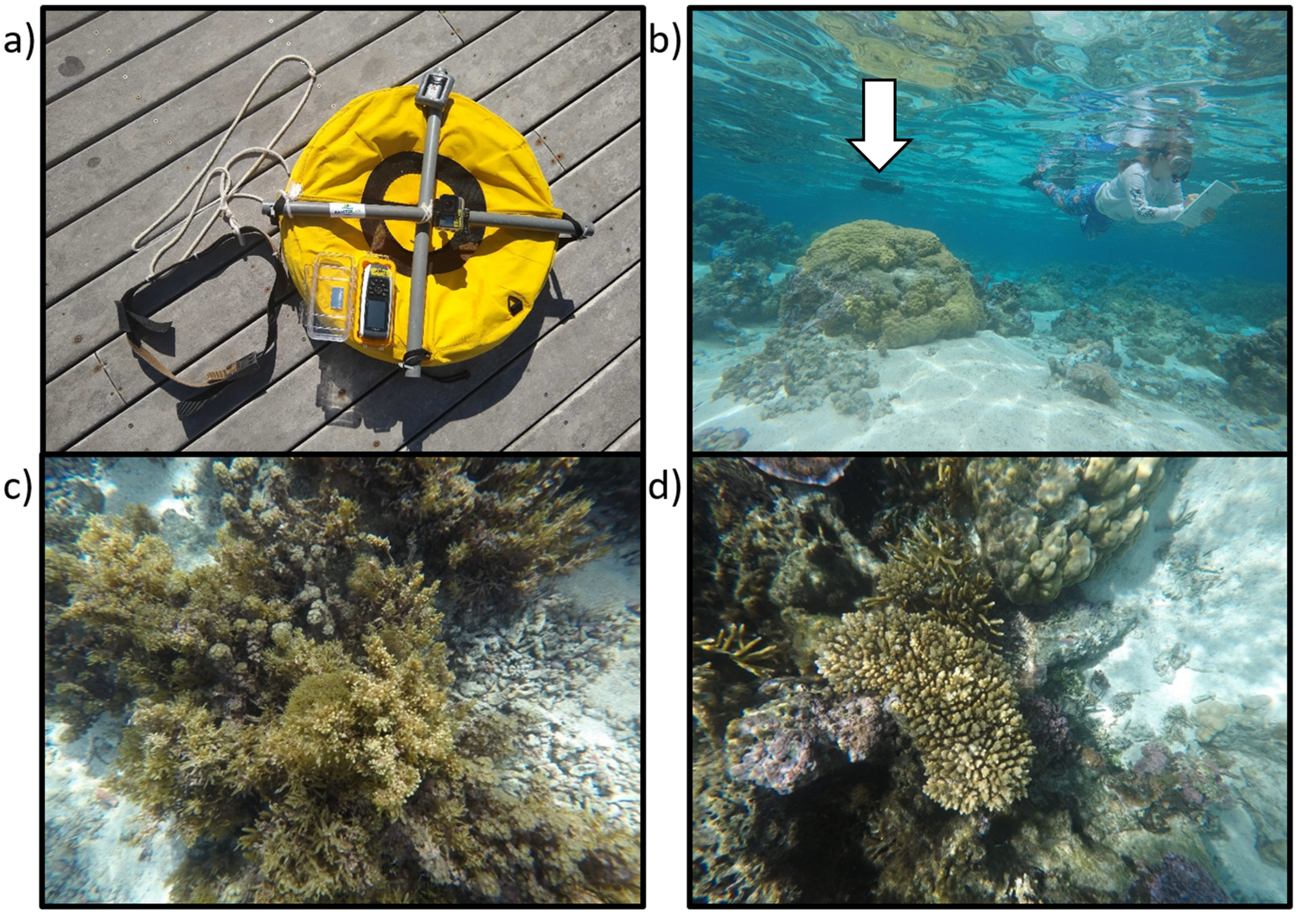
Table 2. Compilation of mean r^2, slope, and intercept from all replicates of each raw substrate scored by 8 CoralNet images, 8 human images, 64 CoralNet images.

	8 CoralNet			**8 Human**			**64 CoralNet**		
Substrate	r^2	Slope	Intercept	r^2	Slope	Intercept	r^2	Slope	Intercept
Sand	0.928	0.919	1.702	0.920	0.951	1.562	0.965	0.955	0.525
Rubble	0.806	0.917	1.902	0.799	0.881	1.949	0.891	1.012	0.376
Hard substrate	0.781	0.806	2.639	0.791	0.874	2.778	0.862	0.889	0.661
Pavement	0.892	0.993	0.512	0.897	0.945	0.342	0.939	1.038	0.275
Turf	0.754	1.761	0.116	0.703	0.954	0.081	0.876	1.752	0.073
Obscured	0.666	0.825	0.949	0.678	0.801	1.238	0.802	0.987	-0.075
Dictyota	0.674	0.868	0.206	0.602	0.778	0.166	0.822	1.039	0.109
Halimeda	0.496	0.626	0.109	0.327	0.533	0.104	0.687	0.822	0.068
Padina	0.946	0.886	0.039	0.928	1.022	0.022	0.976	0.888	0.025
Sargassum	0.920	1.080	0.032	0.910	0.937	0.109	0.978	1.118	-0.066
Turbinaria	0.855	0.930	0.728	0.821	0.905	0.724	0.929	1.005	0.193

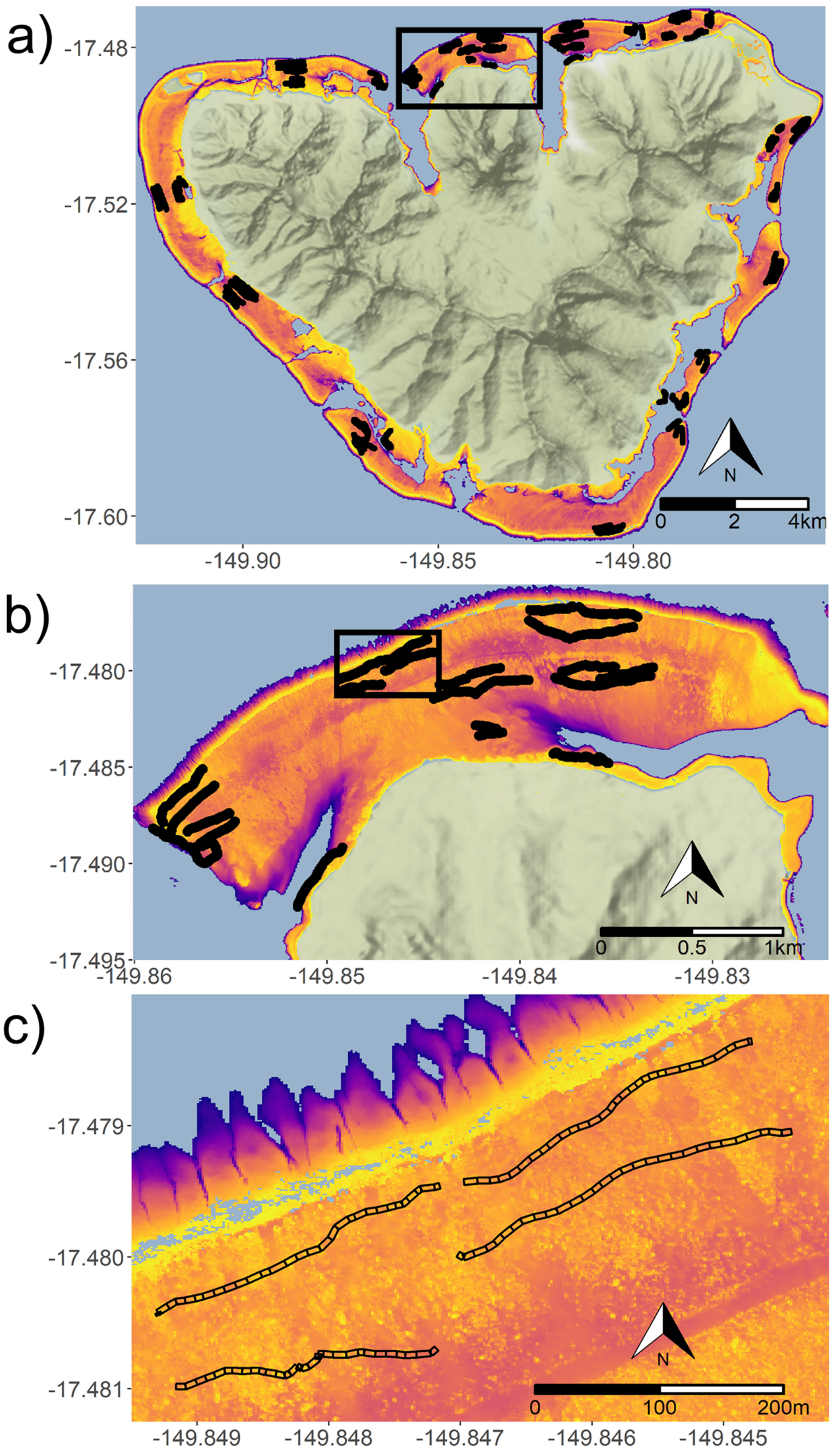
Acropora - branching	0.869	0.890	0.103	0.800	0.921	0.153	0.964	0.935	-0.012
Acropora	0.124	1.975	0.092	0.401	0.348	0.061	0.138	5.616	0.047
Massive *Porites*	0.738	0.774	0.553	0.671	0.790	0.662	0.878	0.914	0.081
Montipora	0.932	1.118	0.143	0.920	0.956	0.072	0.972	1.144	0.091
Pocillopora	0.304	0.434	0.206	0.160	0.367	0.242	0.565	0.817	0.052
Porites rus	0.540	0.515	0.089	0.491	0.843	0.112	0.712	0.603	0.020

Table 3. Compilation of mean r^2, slope, and intercept from all replicates of each aggregated substrate scored by 8 CoralNet images, 8 human images, 64 CoralNet images.

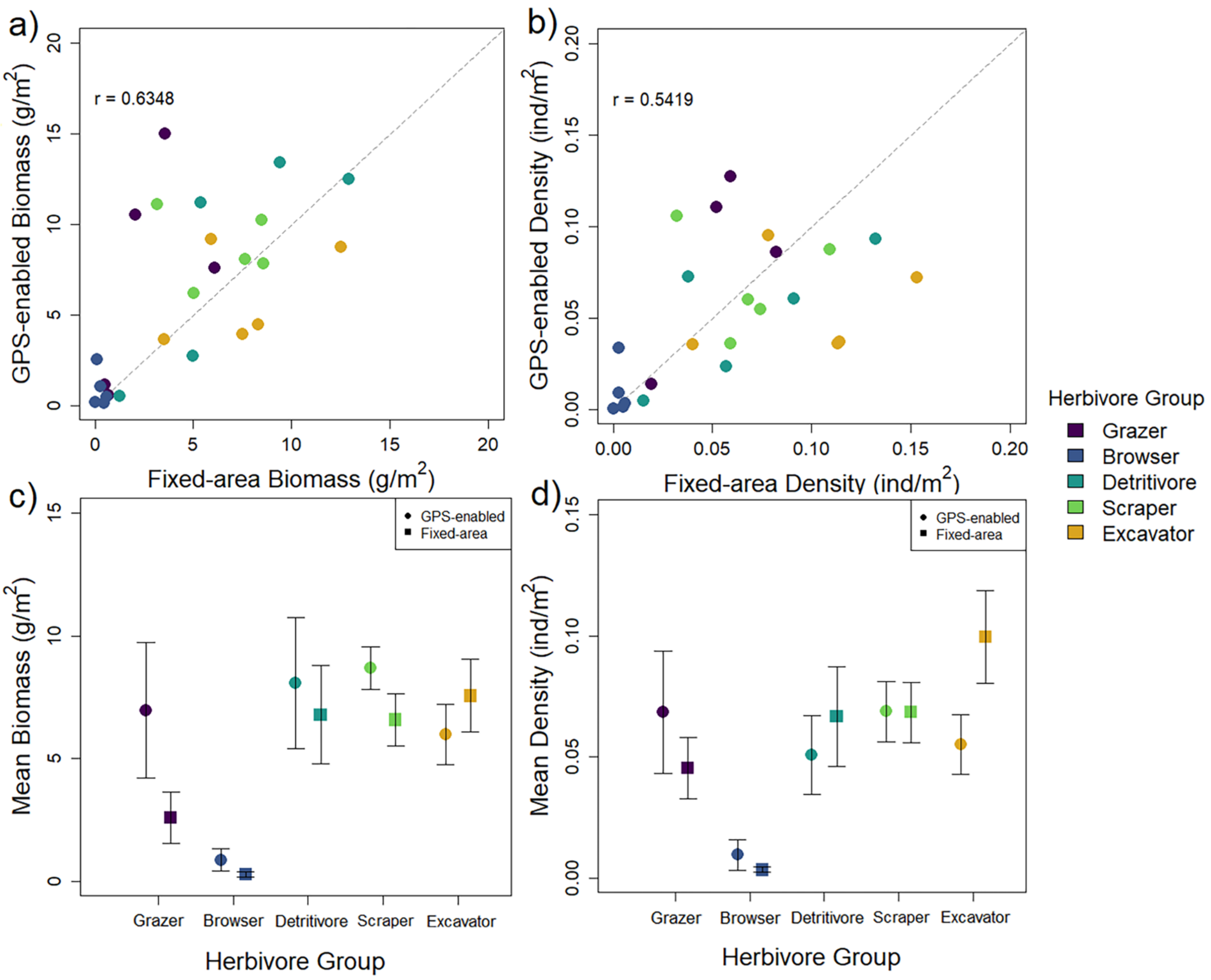
	8 CoralNet			**8 Human**			**64 CoralNet**		
Substrate	r^2	Slope	Intercept	r^2	Slope	Intercept	r^2	Slope	Intercept
Live Coral	0.821	0.909	0.774	0.794	0.893	0.847	0.926	1.017	0.012
Macroalgae	0.895	0.961	1.091	0.865	0.902	0.949	0.951	1.017	0.521
Hard Substrate	0.790	0.804	4.457	0.803	0.871	3.592	0.864	0.878	2.309
Soft Substrate	0.875	0.869	5.888	0.834	0.891	5.148	0.940	0.932	2.802



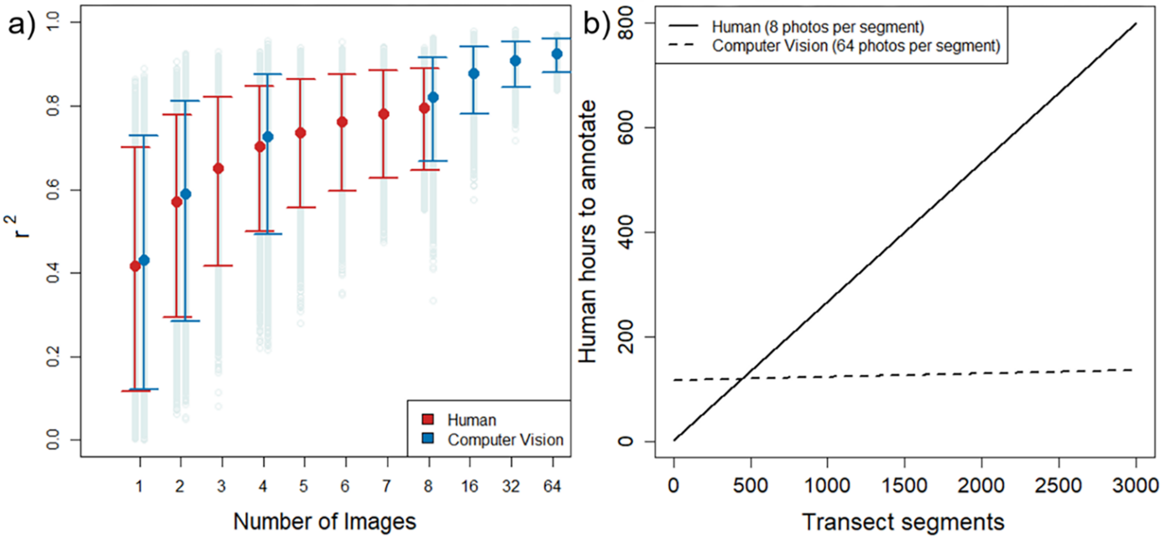
LOM3_10557_figure1_tiff_final.tif



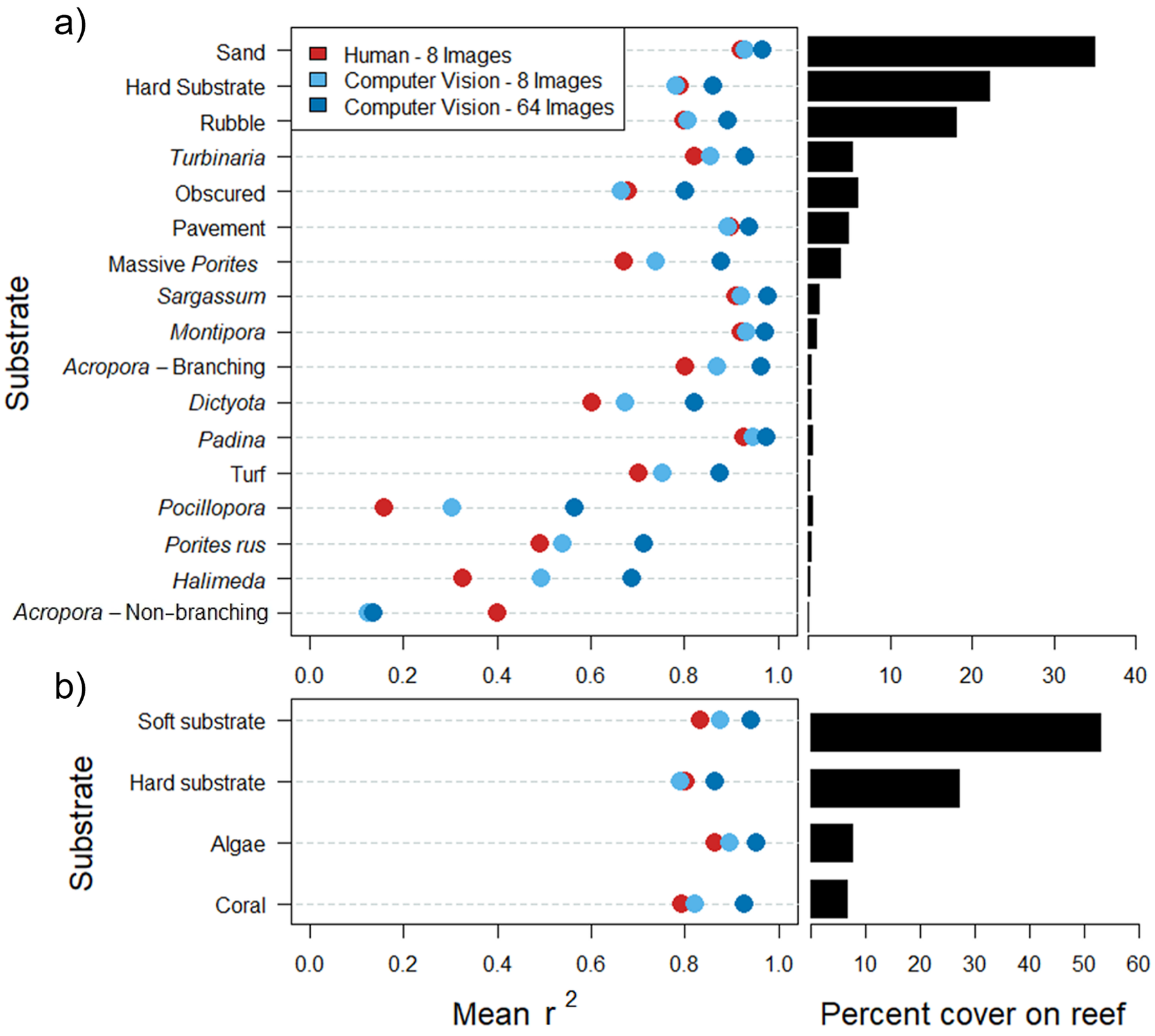
LOM3_10557_figure2_tiff_final.tif



LOM3_10557_figure3_tiff_final.tif



LOM3_10557_figure4_tiff_final.tif



LOM3_10557_figure5_tiff_final.tif

Table 1. List of individual substrates used in the analysis, along with each substrate's aggregated group and a description. Only substrates identified by both humans and CoralNet in our validation procedure are shown here.

Individual substrate	Aggregated substrate	Description
Sand	Soft substrate	Sandy, fine substrate
Rubble	Soft substrate	Coral rubble (broken, detached dead coral)
Hard substrate	Hard substrate	Hard substrate (dead coral/rock, including crustose coralline algae, and low turf-algae)
Pavement	Hard substrate	Pavement (low relief hard substrate)
Turf	Hard substrate	Longer turfing algae (e.g., damselfish garden)
Obscured	Obscured	Unable to see substrate due to it being too dark, blurry, or otherwise covered (e.g., by a fish or a bubble)
Dictyota	Algae	Any algae in the genus *Dictyota*
Halimeda	Algae	Any algae in the genus *Halimeda*
Padina	Algae	Any algae in the genus *Padina*
Sargassum	Algae	Any algae in the genus *Sargassum*
Turbinaria	Algae	Any algae in the genus *Turbinaria*
Acropora - branching	Coral	Coral in genus *Acropora* showing branching morphology
Acropora	Coral	Corals in genus *Acropora* showing other growth forms (e.g., tabular)
Massive *Porites*	Coral	Corals in genus *Porites* showing massive growth forms (e.g., *Porites lobata*)
Montipora	Coral	Corals in genus *Montipora* showing any growth form
Pocillopora	Coral	Corals in genus *Pocillopora* (except for those identified as *P. damicornis*)
Porites rus	Coral	Corals of the species *Porites rus*

Table 2. Compilation of mean r^2, slope, and intercept from all replicates of each raw substrate scored by 8 CoralNet images, 8 human images, 64 CoralNet images.

	8 CoralNet			**8 Human**			**64 CoralNet**		
Substrate	r^2	Slope	Intercept	r^2	Slope	Intercept	r^2	Slope	Intercept
Sand	0.928	0.919	1.702	0.920	0.951	1.562	0.965	0.955	0.525
Rubble	0.806	0.917	1.902	0.799	0.881	1.949	0.891	1.012	0.376
Hard substrate	0.781	0.806	2.639	0.791	0.874	2.778	0.862	0.889	0.661
Pavement	0.892	0.993	0.512	0.897	0.945	0.342	0.939	1.038	0.275
Turf	0.754	1.761	0.116	0.703	0.954	0.081	0.876	1.752	0.073
Obscured	0.666	0.825	0.949	0.678	0.801	1.238	0.802	0.987	-0.075
Dictyota	0.674	0.868	0.206	0.602	0.778	0.166	0.822	1.039	0.109
Halimeda	0.496	0.626	0.109	0.327	0.533	0.104	0.687	0.822	0.068
Padina	0.946	0.886	0.039	0.928	1.022	0.022	0.976	0.888	0.025
Sargassum	0.920	1.080	0.032	0.910	0.937	0.109	0.978	1.118	-0.066
Turbinaria	0.855	0.930	0.728	0.821	0.905	0.724	0.929	1.005	0.193
Acropora - branching	0.869	0.890	0.103	0.800	0.921	0.153	0.964	0.935	-0.012
Acropora	0.124	1.975	0.092	0.401	0.348	0.061	0.138	5.616	0.047
Massive *Porites*	0.738	0.774	0.553	0.671	0.790	0.662	0.878	0.914	0.081
Montipora	0.932	1.118	0.143	0.920	0.956	0.072	0.972	1.144	0.091
Pocillopora	0.304	0.434	0.206	0.160	0.367	0.242	0.565	0.817	0.052
Porites rus	0.540	0.515	0.089	0.491	0.843	0.112	0.712	0.603	0.020

Table 3. Compilation of mean r^2, slope, and intercept from all replicates of each aggregated substrate scored by 8 CoralNet images, 8 human images, 64 CoralNet images.

	8 CoralNet			**8 Human**			**64 CoralNet**		
Substrate	r^2	Slope	Intercept	r^2	Slope	Intercept	r^2	Slope	Intercept
Live Coral	0.821	0.909	0.774	0.794	0.893	0.847	0.926	1.017	0.012
Macroalgae	0.895	0.961	1.091	0.865	0.902	0.949	0.951	1.017	0.521
Hard Substrate	0.790	0.804	4.457	0.803	0.871	3.592	0.864	0.878	2.309
Soft Substrate	0.875	0.869	5.888	0.834	0.891	5.148	0.940	0.932	2.802