Using Deep Learning to Monitor Coral Reef Health

Rithika Narayan

Personal Section

A few years ago, my best friend returned from her vacation to the island of Cozumel off the coast of Mexico with a GoPro full of images she had taken while paragliding, hiking, and relaxing on the beach. Most interesting to me, however, were the photos she had taken while scuba diving in the coral reefs that surround the island. As we scrolled through dozens of pictures of corals, I realized that my friend, like millions of other tourists, had in her possession a treasure trove of information about the health of the coral reefs of Cozumel at the time she visited. However, the issue was how to collect and analyze all that information efficiently.

I first became interested in coral reefs in my childhood. My father and I spent the evenings watching nature programs on channels such as National Geographic; I remember watching hours of footage of lions hunting zebras, penguins swimming in the waters of the Antarctic, and hundreds of species making their homes in the reefs of the world. I was awed by the natural world and I continue to be, but as we are all familiar, the incredible ecosystems we have the joy of seeing captured on our screens are being devastated by the effects of climate change. Coral reefs, specifically, face the increasingly severe threat of bleaching: warming ocean waters stress corals and cause them to expel the microscopic algae that live within them and that provide them food. Nearly half of coral reefs have been severely damaged by bleaching and other climate change-related diseases. It is vital that we rally to preserve them for the future of our oceans.

My childhood fascination with reefs, my friend’s scuba diving adventures, and the grim news articles I scrolled through about how yet another section of the Great Barrier Reef had been destroyed inspired me to undertake research on how to contribute to reef recovery efforts. I had long felt a sense of hopelessness about the destruction of the environment; this project provided me with an opportunity to be an active participant in ecological efforts that I had watched from the sidelines. It also gave me the chance to explore a technology that had fascinated me for years: machine learning (ML).

ML is a field of computer science concerning the study of algorithms that learn from exposure to previous data in order to improve at autonomously completing a task. Chances are you’ve already interacted with ML algorithms in your everyday life: they drive the show recommendations that Netflix gives you, virtual assistants like Siri and Alexa, and facial recognition software. I developed my interest in this field as I was thinking about my future in college and beyond; in fact, this project was a principal factor behind my decision to study computer science in college. I decided to dive head first into exploring ML by applying it environmental monitoring. With the guidance of my mentor Mr. Anthony Pellicano, the ML specialist at Angion Biomedica Corp. in New York, I used a convolutional neural network to analyze underwater images of coral reefs in order to detect the presence of healthy and bleached corals within these images. In order to do so, I took online courses on ML to better acquaint myself with the mathematics behind the technology I was harnessing. Although the calculus and statistics discussed were above my level, the mere mention of them piqued my interest. I can’t wait to get to the point in my academic journey when I am able to fully understand and appreciate the mathematics behind ML.

As far as advice to aspiring high school researchers goes, I urge everyone to explore their curiosities fully. I have seen incredible projects built from a single, seemingly insignificant question; the idea for my own came from spending time with my friend. Research also doesn’t require access to a state-of-the-art lab; there are endless contributions you can make to science and mathematics in your backyard, on your laptop, or with a pencil and a piece of paper.

Research

INTRODUCTION

Coral reefs are among the most biodiverse marine ecosystems, housing millions of organisms across thousands of species within just 1% of the ocean floor [1]. In fact, the largest living structure is a coral reef itself, the Great Barrier Reef [2]. In addition to providing living habitats, coral reefs protect coastal areas from tidal waves and erosion. Healthy reefs absorb 97% of a wave’s energy, buffering shorelines from currents, waves, and storms, thereby mitigating loss of life and property damage [1]. In addition to providing tremendous ecological benefits, coral reefs support the millions of people that rely on them for food and to draw tourists (over 70 million trips made annually), making these fragile and beautiful organisms a powerful engine of coastal and marine tourism [3]. In one estimate, the net annual benefit of the world's coral reefs is approximately $30 billion in the forms of tourism, recreation, coastal protection, fisheries, and biodiversity [4].

Despite their importance, corals and coral reefs are in crisis. Environmental changes including oceanic warming and acidification could result in issues such as bleaching. This phenomenon is associated with the expulsion of in-dwelling algae by coral polyps, causing them to lose more than just their bright coloration [5]. Coral polyps have symbiotic relationships with these algae, which provide a majority of the coral's energy and are vital for the survival of the individual coral and the reef as a whole. While bleached corals might still survive, starvation eventually sets in [5]. The leading cause of coral bleaching is increasing water temperatures, with mass bleaching events occurring across hundreds of miles or more of ocean [6-9]. In 2016, heat stress encompassed more than half the coral reefs globally and other mass bleaching events have eliminated swathes of healthy coral in the Great Barrier Reef, the western Indian Ocean, and the Seychelles [6-8].

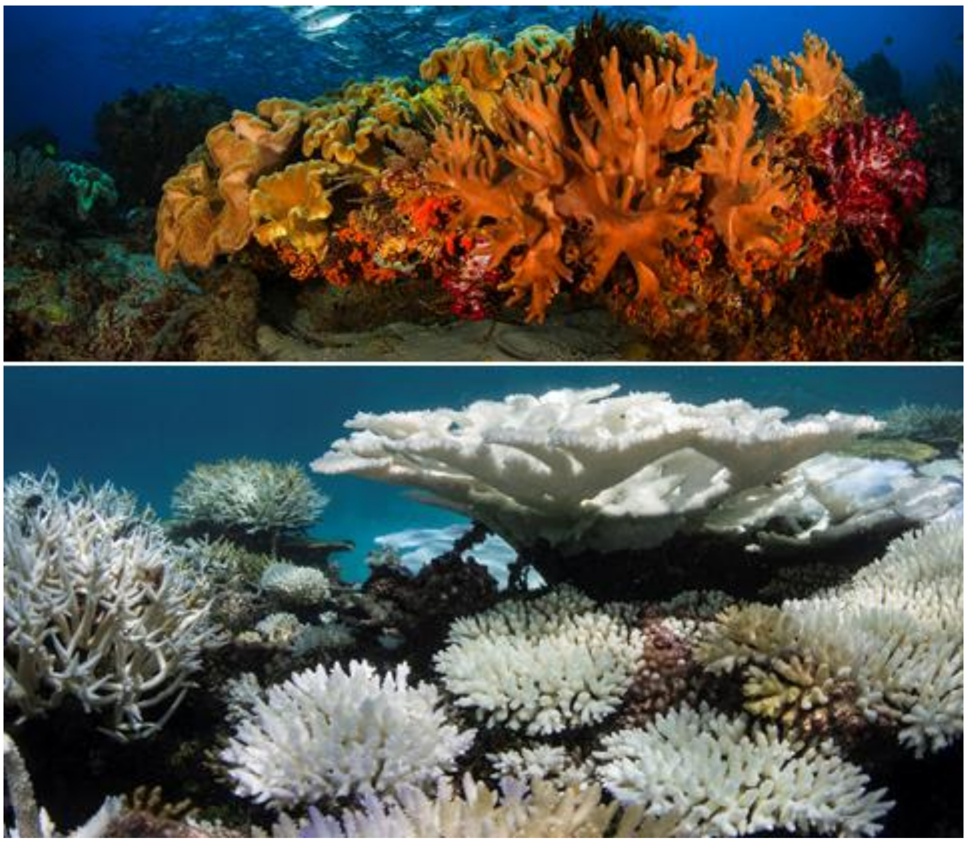

Figure 1. A healthy reef, characterized by dense and diverse coral cover, in the top image is compared with a reef that has undergone a bleaching event, shown in the bottom image. Image Credit: [12].

In addition to warming waters, the impact from land-based sources of pollution—including coastal development, deforestation, agricultural runoff, and oil and chemical spills—can impede coral growth and reproduction, disrupt overall ecological function, and cause disease and mortality in sensitive species [6-9]. Under the United States Endangered Species Act, 22 coral species are currently listed as threatened, and 3 are listed as endangered [10]. With more than 40% of the world's live coral cover lost in the last three decades, the number of people adversely impacted - 500 million - is 35% more than the entire population of the United States [11].

In order to address these problems, global efforts including the XL Catlin Seaview Survey, the Allen Coral Atlas, and CORALNET are using underwater and satellite-based images of the coral benthic cover, i.e. top view of the coral canopy, to both create a coral map and assess coral health in terms of the percentage of live coral [12-14]. These analyses are useful to track the health of a reef over time; for example, seeing a decrease in live coral cover while noticing an increase in algal cover could indicate that corals are losing the battle for space, sunlight, and oxygen to algae [15]. Traditionally, these analyses have been done by human annotators, but this is both an expensive and laborious process. As described in "Automated annotation of coral reef survey images", of the millions of images of reefs taken each year, only 1% is annotated by humans, which leaves a wealth of information untapped [16]. To bridge this gap, machine learning (ML) approaches have recently been unveiled that analyze benthic cover – including types of corals in a reef, other reef invertebrates, and sand – and achieve the same levels of accuracy as traditional methods but more efficiently [12-14].

Provided empirical data, ML makes new predictions or recognizes new patterns. The field is very broad and includes many types of algorithms with various advantages and disadvantages for a given task, such as the support vector machine (SVM) used as the architecture for CORALNET or convolutional neural networks (CNNs), a type of deep learning architecture most often used in image classification which generates functions that map the features in an input image to a classification output [16, 17]. ML has been making headway in medicine, where it has already been successfully utilized in order to diagnose skin cancers, diabetic retinopathy, lymph node metastasis, pneumonia, emphysema, and many more pathologies [18]. Indeed, most diagnoses can trigger intervention and this is my goal for use of ML in environmental, and more specifically coral reef, monitoring. However, assessment of the benthic cover alone is insufficient to fully diagnose the extent of reef disease, as it only provides information about the quantity of coral but not the health condition of that coral. Satellite imagery from above water might show healthy coral tentacle tops whereas bleaching might be from bottom-up. Furthermore, satellite imagery lacks the clarity and depth of detail needed to diagnose the small lesions in corals which may only be partially bleached.

While there is tremendous cause for concern, there is much more need for action. Historical evidence indicates that reefs are able to recover, even from mass bleaching or predation events [19]. However, their ability to recover is dependent on factors such as herbivory, algal cover, and coral recruitment [19]. There are already initiatives in place, such as the Mars Symbioscience which places complex structures in open spaces in depleted reefs to serve as recruitment locations for new corals, that have successfully begun rebuilding our marine ecosystems [20].

My overarching hypothesis is that coupling ML with crowd-sourced images can be used to analyze existing and new databases to identify coral disease. Hundreds of thousands of photographs routinely obtained by professional and amateur scuba divers, tourists, and other deep-sea forays represent a warehouse of untapped information that can be used to preserve coral health across the globe. This approach, which encourages citizen science and empowers stakeholders like you and I, couples ML with crowd-sourced images to complement existing efforts to identify the extent of worldwide coral bleaching. Analyzing images uploaded by users across the globe for evidence of bleaching by traditional methods would be a Herculean task. By contrast, ML could be deployed to detect and identify bleaching in corals by first training the machines on an annotated set of images. As described below, a CNN has now been trained on images of corals with varying degrees of health in order to create a model that can identify disease.

The overarching deliverable is an ML-based platform that will be used to analyze incoming crowd-sourced images of corals along with the location of these corals. Heightened public awareness can translate to increased data and intensified conservation efforts. The data collected could be integrated into the National Oceanographic and Atmospheric Administration's (NOAA's) Deep Sea Coral Data Portal (DSCDP), which houses images of and information about coral found around the world [21]. This technology can also be used in citizen science projects such as the Great Reef Census, which asks volunteers to upload 10 photos of their diving site [22]. This coral health database can be rapidly used to identify trends in the health of different reefs and anticipate outbreaks. Based on these trends, resources for treatment of bleaching can be allocated, protection zones can be set up, and local awareness can be raised to lower risk factors.

METHODS

I selected an ML algorithm after a literature review which indicated that a CNN has a high Mathew's Correlation Coefficient (MCC), a performance parameter that assesses classifier algorithms, compared to other algorithms when applied to coral images [17, 23]. The MCC is calculated on a scale between -1 and 1, with algorithms that score near -1 making predictions that oppose the true values, algorithms that score closer to a 0 making their decisions near randomly, and algorithms that score closer to 1 being in near agreement with the true values of what they are classifying. I chose to use the Mask Region (R) CNN algorithm, an open source CNN, specifically because of its successful use in the analysis of nuclei in microscope images, detection of sports fields in satellite images, and other object detection and segmentation projects [24]. An R-CNN was preferable to a CNN because it not only can classify objects in an image, but detect where they are located in the image and output a “mask” over the region of interest, thus providing more detailed analysis of the image [24]. I then modified the Mask R-CNN to best suit the qualities of my dataset by adding and changing the names of the classes (the health conditions of the coral) as well as making the algorithm compatible with my data annotation platform of choice, Labelbox [25].

I obtained a total of 178 images of individual corals and coral reefs with the aforementioned health conditions (healthy and bleached), which corresponded to a total of 1530 individual instances of these health conditions, in order to train and test the algorithm (Table 1). I then annotated these images in Labelbox using the segmentation polygon tool to outline masks of the shape of the corals as well as disease lesions (Figure 2). I ensured the accuracy of the annotation of these images by collecting them from reliable research sources such as the XL Catlin Seaview Survey and published ecological papers, among others [12, 13]. These masks from the annotations are what the algorithm learned from and emulated.

	Healthy	Bleached	Total
Training	553	830	1383
Validation	91	56	147

Table 1. Distribution of individual instances of coral health conditions across training and validation data sets.

Figure 2. Example of image data annotated (right) in Labelbox and used for model-training. Healthy corals were outlined in green and bleached corals were annotated in purple using the polygonal selection tool. The annotations are stored in the JSON file for the model to learn from. Image Credit: [12], Researcher.

I split the dataset into two groups, one for training and one for validation. The setup for both groups was identical in terms of classes and annotation type, but the data were split between the two, approximately 80-20 training to validation, which the Pareto Principle indicates is an ideal split for midsize data sets [26]. Training data is the data that the model learns from, and validation data is the data that it tests itself on between rounds of training. The corals in the images could be classified as healthy or bleached, and the background of the image (along with any other organisms present) was left unannotated and was treated as a third class. The platform generated a JavaScript Object Notation (JSON) format that contains image paths and the annotations for each image. JSONs are readable by humans and computers alike, making them a viable method for providing data to a model [27]. This JSON was used by the Mask R-CNN algorithm to train on and validate itself using the images.

I started an Amazon Web Services (AWS) EC2 instance in which I developed the model using the images from the dataset, the JSON files containing the image annotations, and the modified algorithm code. The instance was equipped with Tensorflow graphical processing unit (GPU) capabilities, which increases computing power and quickens the training process [28]. Over time, the algorithm generated logs as it trained that allow it to build on what was previously learned with each training session rather than relearning each session, the basis of the neural network. In order to overcome the small size of my datasets (178 images and 1530 total instances of healthy and bleached corals), I implemented image augmentation. This is a technique used to artificially inflate the size of the dataset by performing transformations — such as flipping or rotation —of annotated images such that these images act as new images for the model to learn from. Image augmentation prevents overfitting, a phenomenon in which a machine learning model trains for too long on too few images and as a result is unable to extrapolate the trends from the training images to completely unfamiliar images. Without using image augmentation, it would be possible for my model to overfit and only recognize bleached and healthy corals in the 178 images it was trained and validated on and not be able to perform well on images it has not seen before.

I monitored the progress of the algorithm using Tensorboard, a feature of Tensorflow, which generates loss graphs and accuracy graphs for different aspects of the model's performance including its class accuracy (how well the model is able to identify healthy corals as healthy and bleached corals as bleached) and segmentation box accuracy (how well the model can outline the corals of each health condition). Loss graphs ideally approach but never reach zero and the accuracy graphs ideally approach but never reach one. Training was halted after the training and validation losses converged, or approached the same value. The convergence of training and validation losses demonstrates that the model has not been overfitted and that it is able to perform well on the validation images, as well as the training images which it learned from. If validation loss were significantly higher than training loss, this would indicate that the model was overfitted.

I repeated this process for two more versions of the algorithm, in a procedure known as hyperparameter tuning. Hyperparameters are specific settings assigned to the model that help control the way it learns and by adjusting the hyperparameters, a researcher can influence their algorithm to train more efficiently and achieve better results regarding loss and accuracy. Table 2 displays selected hyperparameters for each of the three versions of the algorithm I trained.

	Version A	Version B	Version C
Steps per Epoch	1000	1000	1000
Validation Steps	50	50	50
Learning Rate	0.001	0.01	0.001
Weight Decay	0.0001	0.001	0.001
Learning Momentum	0.9	0.9	0.5

Table 2. Selected hyperparameters from training. Hyperparameters are values used by a researcher to control the training process and can be tuned during training to achieve the best results.

After hyperparameter tuning was completed, I chose the version of the algorithm that performed best on the validation datasets (see Results) and tested that model on a set of thirty images of healthy and bleached corals. Tests sets are used in addition to validation sets as another measure against overfitting; by choosing the model that performed best on the validation set, the researcher is inserting bias into the model by optimizing it to the validation set. The model is not optimized to the test set, so assessing its performance on these images provides a better understanding of its true abilities when applied to new images. The size of the test set is thirty images because thirty images is approximately 20% of the size of the training set, and this is a standard convention in the ML field. In order to choose an operating point (the confidence level at which a model will choose to display its prediction of the health condition of the coral rather than discard it as a poor guess), I constructed a precision-recall curve. These curves are constructed by choosing operating points at intervals for values between 0 and 1 (I chose intervals of 0.1) and recording the amount of true positives, true negatives, false positives, and false negatives generated by the model at that operating point. They are used to determine the balance between the amount of false positives generated and the amount of true positives generated when the model is given a certain confidence threshold and trained on a dataset which may not be balanced in terms of class representation. Depending on the application of the model, the developer may choose to prioritize decreasing the amount of false positives or may prioritize increasing the number of true positives. From this curve, I chose the operating point which led to the optimal balance between precision (the number of true positives out of total positives) and recall (the number of true positives out of the sum of true positives and false negatives). This curve is discussed in the Results section.

RESULTS

The models were trained on the dataset of 1530 instances for 283 epochs with heavy image augmentation to prevent overfitting. Each epoch represents one cycle of training and one pass over a subset of the training dataset. All training and validation losses were minimized to near zero – segmentation loss approached 0.136, binding-box loss approached 0.013, and class loss approached .003 for the model which was determined to have the most optimal hyperparameters (Figure 3). Binding box and segmentation loss are measures of the model's performance in identifying the coral in the image and being able to outline them. Class loss is a measure of the model's ability to predict whether a particular coral is healthy or bleached.

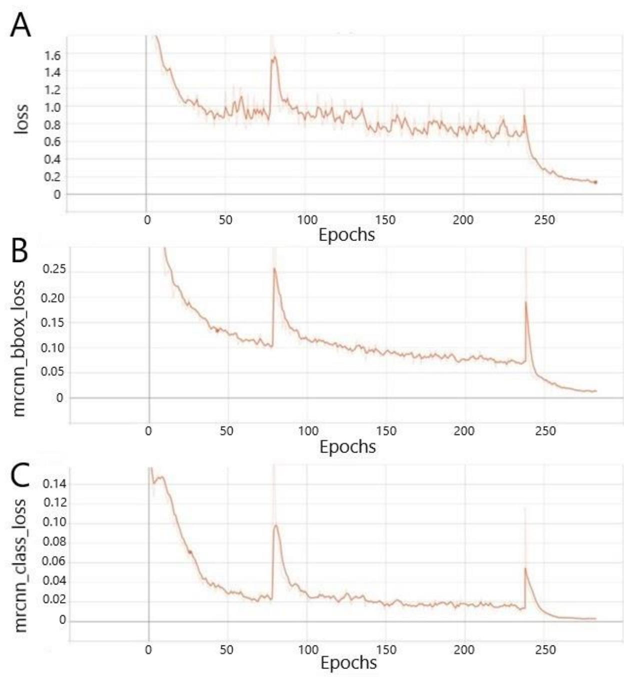

Figure 3. Loss is defined as the difference between the results of the ground-truth annotation training data and the model's prediction attempt. Pictured here: segmentation loss (A), binding-box loss (B), and class loss (C). Segmentation and binding-box identifies individual coral features while class loss predicts which class those features belong to: healthy or bleached coral. Image Credit: Researcher.

Based on the validation losses, the model whose hyperparameters are outlined in Table 2 under Version A was selected as the most optimal model, because its overall validation loss was the lowest of the three versions of the model that were trained. This optimal model was tested on 30 test images of healthy and bleached corals. A sample image from the test set and an analysis done of it by the model are shown in Figure 4. The model was run on the test set at operating points spaced at intervals of 0.1 between 0.0 and 1.0 in order to determine the optimal operating point. This operating point was determined to be 0.8 because it balanced the rate of true positives (represented by recall) and false positives (represented by precision) most appropriately for the application to detecting bleaching in coral reefs. Detecting true instances of bleaching is more critical to protecting reef health than minimizing false positives, therefore the operating point of 0.8, which resulted in a recall of 0.9 and a precision of 0.7, was appropriate. At this operating point, the model achieved 85% accuracy in distinguishing between healthy and bleached corals in the test set. Precision, recall, and accuracy combined provide a more complete picture of the model's performance; the operating point is also flexible and can be changed for implementation of the model.

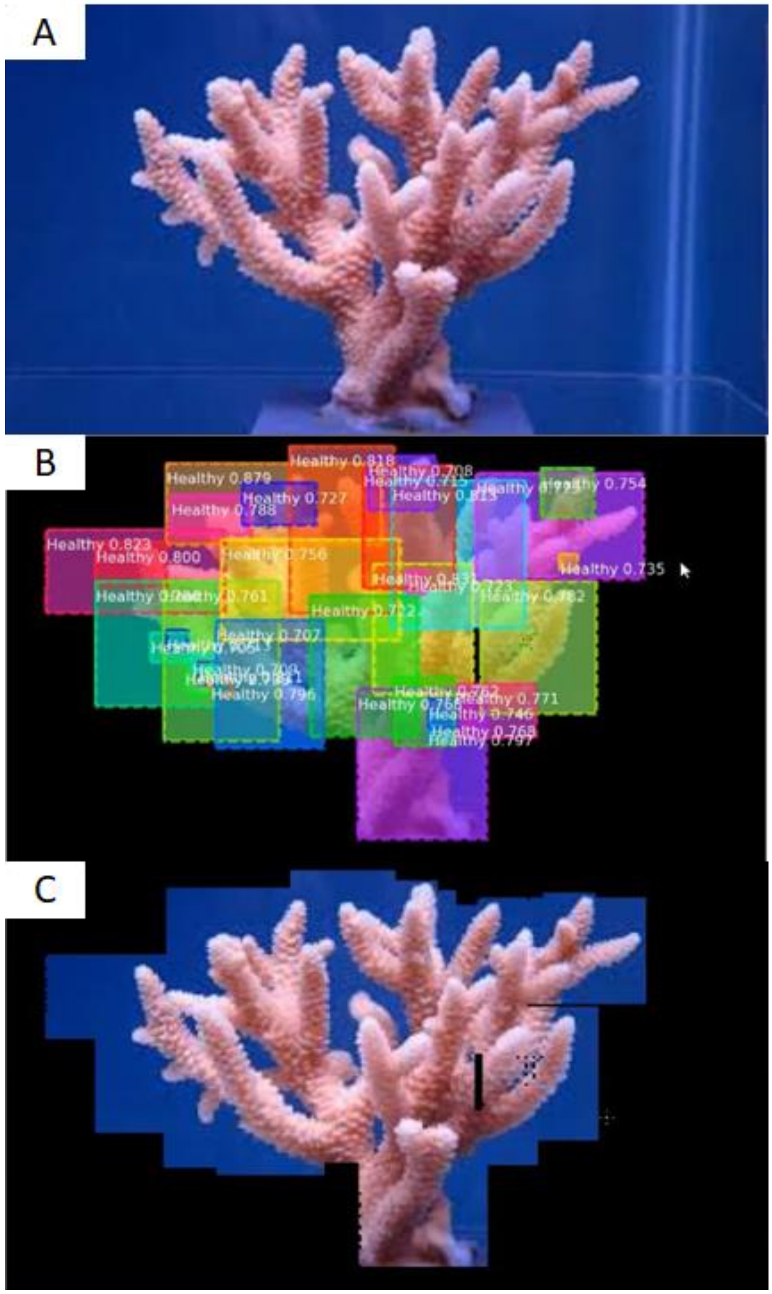

Figure 4. Model detection of healthy coral. An image of healthy coral (A) is analyzed with the model which makes predictions (B) breaking the image down into segmented regions that the computer recognizes as healthy coral (visualized in C). Image Credit: Researcher.

DISCUSSION

This research represents an advancement in the application of ML to environmental monitoring, specifically in the area of coral reef monitoring. The standard thus far has been benthic cover analysis, aided by algorithms such as SVM [16]. While these analyses are an appropriate first step in measuring reef health, they do not provide information beyond the percent of live coral coral. In order to make informed decisions to protect coral, researchers and policymakers need in-depth knowledge of the state of reefs. My algorithm provides this type of information because it is an instance segmentation algorithm; not only does the algorithm detect the presence of bleached or healthy corals in an image, it can also identify how many corals of these types are present and where they are located in the image frame.

The main limitation of this project was the small dataset. Only high-resolution, captioned, publicly-available images were used in the training and validation datasets. This limited the amount of data available for the model to learn from; the time required to annotate the datasets (see Methods) further limited the amount of data because labeling the regions of interest is a time-consuming process. These issues may have resulted in overfitting, despite the image augmentation methods used to combat overfitting. The model could be improved by increasing the size of the dataset and ensuring diversity in the types of coral represented in the images.

An ML approach does have weaknesses in that some false discovery rate is inherent with ML methods. However, this feature can be minimized through more training on additional data [29]. Secondly, images taken at greater depths might not have the same resolution and quality as images obtained at shallower depths, resulting in false positive findings. This issue could be tackled through the use of algorithms such as the sea-thru algorithm that reconstructs color in underwater images in the preprocessing of input images [30]. Nevertheless, this approach is an important and stakeholder-empowering first step to address a global crisis because diagnosis can trigger intervention.

CONCLUSION

The research presented herein couples crowd-sourced data with ML to provide a deliverable that complements existing strategies to diagnose the extent of coral bleaching. It has the potential to spur citizen science and empower stakeholders as active participants and trigger corrective measures to halt or slow coral bleaching.

Almost on a seasonal basis, we are experiencing irreversible ecological tipping points. Human footprints, including pollution and global warming, are starting to leave permanent marks on our ecosystem, including the calcium carbonate skeletons of bleached and starved corals. Eventually, these coral skeletons will erode, causing the reef structure to collapse. The mass bleaching events in recent years have been triggered by global warming and, if current trends continue, corals are expected to become increasingly rare on reef systems.

Bleaching is reversible if the stressors are removed quickly [31]. There is a global ongoing effort to harness satellite imagery to map benthic cover as a measure of reef health [12-14]. However, additional resources must be taken advantage of. Fortunately, there is a large crowd-sourced databank in the form of videos and photographs of corals and coral reefs across the globe. These close-range photographs hold valuable information on the health of the corals. The research described herein utilizes a ML-based approach to diagnose coral bleaching based on these photographs. ML has precedent in diagnosis: the technology has revolutionized medicine, now routinely outperforming and outpacing human experts in making diagnoses at a fraction of the cost and saving lives across the world [17].

The ML algorithm is currently able to identify healthy and bleached corals at an 85% accuracy. ML approaches come with notable benefits including: efficiency allowing for larger sample analyses, the presence of a “trained expert” wherever there is a computer rather than relying on thinly spread marine biologists, the detection of patterns and correlations which would require a mass workforce working in unison for large amounts of time to replicate, and learning based on expert evaluations.

This algorithm will be made available to NOAA and private companies such as National Geographic so that it can be used freely. It can also be incorporated into an application for use by the general public in order to encourage more tourists to obtain photographs of corals, run them through the algorithm, and upload the results to coral health databases such as NOAA’s DSCDP [21]. With the successful implementation of widespread reef monitoring and analysis will come the ability to efficiently respond to bleaching events. Although there is no method of treating bleached corals, these vulnerable corals can be protected from further harm and given the opportunity to recover by halting trawling, industrial runoff, and mining in the area after bleaching events [19].

REFERENCES

1. “Corals,” *NOAA's National Ocean Service* , 01-Jun-2013. [Online]. Available: https://oceanservice.noaa.gov/education/tutorial_corals/welcome.html. [Accessed: 06-Jul-2019].

2. “What is the Great Barrier Reef?,” *National Oceanic and Atmospheric Administration*, 01-Jun-2013. [Online]. https://oceanservice.noaa.gov/facts/gbrlargeststructure.html. [Accessed: 11-Jul-2019].

3. M. Spalding, L. Burke, S. A. Wood, J. Ashpole, J. Hutchison, and P. Z. Ermgassen, “Mapping the global value and distribution of coral reef tourism,” *Marine Policy*, vol. 82, pp. 104–113, Aug. 2017.

4. “Value of Corals,” *Coral Reef Systems*. [Online]. Available https://sioweb.ucsd.edu/projects/coralreefsystems/about-coral-reefs/value-of-corals/. [Accessed: 10-Jul-2019].

5. M. S. Pratchett, D. Mccowan, J. A. Maynard, and S. F. Heron, “Changes in Bleaching Susceptibility among Corals Subject to Ocean Warming and Recurrent Bleaching in Moorea, French Polynesia,” *PLOS One*, vol. 8, no. 7, Jul. 2013.

6. T. P. Hughes et al., “Global warming and recurrent mass bleaching of corals,” *Nature*, vol. 543, no. 7645, pp. 373–377, Mar. 2017.

7. T. P. Hughes et al, “Spatial and temporal patterns of mass bleaching of corals in the Anthropocene,” *Science*, vol. 359, no. 6371, pp. 80–83, Jan. 2018.

8. M. Le Nohaïc et al, “Marine heatwave causes unprecedented regional mass bleaching of thermally resistant corals in northwestern Australia,” *Scientific Reports*, vol. 7, no. 1, Nov. 2017.

9. G. S. Aeby et al, “Patterns of Coral Disease across the Hawaiian Archipelago: Relating Disease to

Environment," *PLOS One*, vol. 6, no. 5, May 2011.

10. R. M. Gregg, "Listing of Coral Reef Species under the U.S. Endangered Species Act," *Climate Adaption Knowledge Exchange*, 30-Jun-2010. [Online]. Available: https://www.cakex.org/case-studies/listing-coral-reef-species-under-us-endangered-species-act. [Accessed: 14-Jul-2019].

11. S. Wismer, S. B. Tebbett, R. P. Streit, and D. R. Bellwood, "Young fishes persist despite coral loss on the Great Barrier Reef," *Communications Biology*, vol. 2, no. 1, Dec. 2019.

12. *XL Catlin Seaview Survey*, 2011. [Online]. Available: http://catlinseaviewsurvey.com/. [Accessed: 18-Jul-2019].

13. *Allen Coral Atlas*. [Online]. Available: https://allencoralatlas.org/. [Accessed: 18-Jul-2019].

14. "CoralNet BETA," *CoralNet*. [Online]. Available: https://coralnet.ucsd.edu/. [Accessed: 16-Jul-2019].

15. S. Clements, "The Faces and Functions of Algae on the Reef," *Scripps Institution of Oceanography*, 27-Feb-2015. [Online]. Available: https://sioweb.ucsd.edu/labs/coralreefecology/the-faces-and-functions-of-algae-on-the-reef/. [Accessed: 19-Jul-2019].

16. O. Beijbom, P. J. Edmunds, D. I. Kline, B. G. Mitchell, and D. Kriegman, "Automated annotation of coral reef survey images," *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012.

17. P. Raghav, "Understanding of Convolutional Neural Network (CNN) - Deep Learning," *Medium*, 04-Mar-2018. [Online]. Available: https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148. [Accessed: 24-Jul-2019].

18. P. Rajpurkar et al, “Deep learning for chest radiograph diagnosis: A retrospective comparison of
the CheXNeXt algorithm to practicing radiologists,” *PLOS Medicine*, vol. 15, no. 11, Nov. 2018.

19. O. Hoegh-Guldberg, E. S. Poloczanska, W. Skirving, and S. Dove, “Coral Reef Ecosystems under
Climate Change and Ocean Acidification,” *Frontiers in Marine Science*, vol. 4, May 2017.

20. “Taking Responsibility to Rebuild Coral Reefs,” *Mars*. [Online]. Available: https://www.mars.com/news-and-stories/articles/coral-reef-rehabilitation. [Accessed: 01-Aug-2019].

21. “Welcome to the NOAA Deep Sea Coral Data Portal,” *NOAA Deep-Sea Coral Data Portal*. [Online]. Available: https://deepseacoraldata.noaa.gov/. [Accessed: 26-Jul-2019].

22. The Great Reef Census. [Online]. Available: https://greatreefcensus.org/. [Accessed: 29-Jul-2019].

23. “Matthews correlation coefficient,” Wikipedia, 04-Jul-2017. [Online]. Available: https://en.wikipedia.org/wiki/Matthews_correlation_coefficient. [Accessed: 02-Aug-2019].

24. Abdulla, W (2017) Mask R-CNN for Object Detection and Segmentation [Source code]. https://github.com/matterport/Mask_RCNN

25. Labelbox. [Online]. Available: https://labelbox.com/. [Accessed: 02-Aug-2019].

26. “Pareto principle,” Wikipedia, 10-Jun-2005. [Online]. Available: https://en.wikipedia.org/wiki/Pareto_principle. [Accessed: 09-Aug-2019].

27. “Introducing JSON,” JSON. [Online]. Available: https://www.json.org/json-en.html. [Accessed: 29-Aug-2019].

28. “Use a GPU,” TensorFlow. [Online]. Available: https://www.tensorflow.org/guide/gpu. [Accessed: 20-Aug-2019].

29. Li, Junning, and Z. Jane Wang. “Controlling the False Discovery Rate of the Association/Causality Structure Learned with the PC Algorithm.” *The Journal of Machine Learning Research*, vol. 10, June 2009, pp. 475–514.

30. Akkaynak, Derya. “Sea-Thru.” *Derya Akkaynak*, https://www.deryaakkaynak.com/sea-thru.

31. Normile, Dennis. “There’s Only One Way to Save the Great Barrier Reef, Scientists Conclude.” *Science*, 15 Mar. 2017, https://www.sciencemag.org/news/2017/03/there-s-only-one-way-save-great-barrier-reef-scientists-conclude.