Convolutional Neural Networks Predict Fish Abundance from Underlying Coral Reef Texture

G. C. Young^{1,2,3*}, V. Balntas³, V. A. Prisacariu³

Last edited May 25, 2018

¹Department of Zoology, University of Oxford, Oxford, OX1 3PS, UK

²Department of Engineering Science, University of Oxford, Oxford, OX1 3PJ

³Operation Wallacea, Old Bolingbroke, Spilsby, Lincolnshire, PE23 4EX, UK

*Corresponding author email <grace@robots.ox.ac.uk>

1 Abstract

2 Coral reefs are among the most biodiverse ecosystems on Earth in large part
3 owing to their unique three-dimensional (3D) structure, which provides niches for
4 a variety of species. Metrics of structural complexity have been shown to correlate
5 with the abundance and diversity of fish and other marine organisms, but they
6 are imperfect representations of a surface that can oversimplify key structural ele-
7 ments and bias discoveries. Moreover, they require researchers to make relatively
8 uninformed guesses about the features and spatial scales relevant to species of
9 interest. This paper introduces a machine-learning method for automating infer-
10 ences about fish abundance from reef 3D models. It demonstrates the capacity
11 of a convolutional neural network (ConvNet) to learn ecological patterns that
12 are extremely subtle, if not invisible, to the human eye. It is the first time in
13 the literature that no *a priori* assumptions are made about the bathymetry–fish
14 relationship.

15
16 *Keywords:* neural networks, structural complexity, habitat complexity, habitat
17 heterogeneity, fish, coral reef, machine learning, artificial intelligence, ecology

1 Introduction

Like forests, coral reefs have high structural complexity enabling them to maintain some of the greatest levels of biodiversity on Earth [1]. Reefs’ unique three-dimensional (3D) structure provides organisms refuges from predators [2], places to spawn [3], surface areas for grazing [4], and niches with optimal light [5], sound [6], or hydrodynamic conditions [7]. Hard corals are the principal architect, but even dead corals or concrete blocks can harbour reef communities [8, 9].

Structural complexity can be measured with a variety of metrics [10], the majority of which significantly correlate with indicators of ecosystem health on land and sea [11, 12]. The most common metric in marine ecology is rugosity, the ratio of a chain’s draped length to its fully extended length [13]. It is slowly being replaced with more sophisticated metrics from 3D models, however [14]. 3D models of reef can be created from sonar or LiDAR data [15], areal drone footage if the water is clear enough [16], or structure-from-motion (SfM) photogrammetry [17]. The relationship between fish and structure is of particular interest to not only marine ecologists, but also fisheries managers and conservationists trying to optimise yield or resilience [2, 18]. Terrain descriptors extracted from 3D models have been shown to correlate with fish abundance [19–23], as have more traditional structural complexity metrics [19, 24, 25].

While the basic ecological rationale for these correlations is clear, nuances in the relationship between structure and function are not so clear, and metrics leave much unanswered. They can oversimplify key structural elements [26], bias discoveries [27], and require researchers to make relatively uninformed guesses about which structural features and spatial scales will be relevant to species of interest [10, 19, 28]. Ideally the 3D model is not summarized into a metric, and instead is fed directly into the data analysis pipeline.

Machine-learning (ML) algorithms are a promising approach for this task because they can learn patterns from raw data without *a priori* information and are

not constrained by the assumptions of parametric statistics. Boosted regression trees (BRTs) from ML have successfully modelled bathymetry–fish relationships from variables derived from 3D models at spatial scales 2–300 m [29–31]. The BRTs can reveal variables’ relative importance and are robust to extraneous variables, missing values, and interactions, but they still require a researcher to make decisions about which variables to extract from the data; *e.g.*, Costa et al. [31] choose six terrain variables (depth, depth standard deviation, curvature, distance to shelf edge, rugosity, and slope of slope) from their 3D models and measured each at five scales (2, 25, 50, 100, and 300 m). Convolutional neural networks (ConvNets), also known as deep networks, do not require any pre-existing assumptions about which descriptions of the data determine relationships. They can learn patterns (if present) from raw data through layers of linear and non-linear operations. The result is that the network can discover relationships more intricate than what even hundreds of multi-scale variables could predict. Already ConvNets have proven to be a powerful tool in marine ecology for automating image-identification tasks that otherwise require human experts, including annotating coral [32] and identifying polyp activity [33].

This paper introduces a method for learning ecological patterns that are extremely subtle, if not invisible, to the human eye. Our aim is not just to automate inference, but also to let artificial intelligence reveal ecological patterns. We demonstrate the capacity of a neural network to learn fish aggregation patterns from just 85 3D models of the reef. It is the first time in the literature that no *a priori* assumptions are made about the bathymetry–fish relationship.

2 Methods

2.1 Study Area and Data Collection

Data were collected from the south-west coast of the Caribbean island of Utila, Honduras (16.09°N, -86.91°W; GPS coordinates in WGS84 format) under a re-

search permit from the Instituto de Conservación Forestal (#ICF-DE-MP-080-2016). SCUBA divers 3D modelled 85 2 x 2 m patches (“quadrats”) of reef 5 ± 2 m deep using underwater photogrammetry (following Young et al. [34]). Additionally, divers recorded the number of fish over each quadrat during a 10-minute fish survey, including 9 minutes spent observing at a distance of ~ 1 m and 1 minute spent searching up-close for cryptobenthic fish (following Gratwicke and Speight [24]).

For analysis, fish data were binned into 2–8 classes of roughly equal proportions (Fig. 1). We started with the simplest division, low or high amounts of fish, and increased the granularity of classes to test how well the network could cope. We stopped at 8 classes because divisions started to lose ecological meaning (*e.g.*, 12–14 versus 14–16 fish). We formulated our problem as a classification rather than regression problem because ConvNets are generally more robust to classification problems [35].

Figure 1: Fish abundance data. (a) Distribution of fish data ($N = 85$, minimum = 4 fish, maximum = 44 fish) with Gaussian smoothing kernel density curve. (b) Division of data into classes of roughly equal proportions.

2.2 Conversion of 3D Models into 2D Images

We converted the 3D models into 2D images because ConvNets are presently better at dealing with images than 3D objects. To do so, a grid of points was projected downward onto the 3D model along the Z-axis (Fig. 2(a)–(b)). The grid was 189 x 189 cm rather than the theoretical maximum of 200 x 200 cm,

the size of the quadrat, so that it could consistently avoid the PVC-pipe that demarcated the edges of the quadrat. Grid points were spaced 1.50 cm apart in X and Y (Fig. 2(b)) because the original 3D models had root mean squared spatial accuracy to 1.48 cm in X and Y [34]. X and Y positions corresponded to pixel indexes in the 2D image, while Z positions corresponded to pixel values (Fig. 2(c)). A model's minimum Z-position was assigned a pixel value of 0, and an integer increment in pixel value represented a 1 cm increase in Z-position. Z-positions did not exceed 255 cm, meaning pixel values did not exceed 255 (Fig. 2(c)). The Python script that performed these steps worked with Rhinoceros 3D software ("Rhino"; Robert McNeel & Associates, Seattle, WA, USA) and is available open-source.¹

Figure 2: Process for converting a 3D model into an image. The X-Y plane is parallel with the ocean surface and the Z-axis points upwards. (a) Points (in red) are projected along the Z-axis onto the uppermost points of the reef 3D model (in grey). (b) Points are extracted from the 3D model. Points cover a 189 x 189 cm grid and each is spaced 1.5 cm from its nearest neighbour. (c) The grid of points is represented as a 126 x 126 pixel greyscale image. Pixel values 0–255 represent Z-positions, each integer representing 1 cm.

The difference between 3D models that hosted high or low fish abundances is extremely subtle to the human eye, and even more so in the image representations (Fig. 3).

2.3 Neural Network Architecture

The 85 images were randomly sorted into a training set containing 60 images and a test set containing 25 images. Three image transformations augmented the

¹www.github.com/gracecalvertyoung/Rhino-Python-Scripts-for-3D-to-2.5D-Conversion

Figure 3: Six randomly selected images from (a) the class with the lowest fish abundance (4–8 fish), and (b) the class with the highest fish abundance (33–44 fish).

training dataset: (1) random cropping to 124 x 124 pixels, (2) center cropping to 124 x 124 pixels, and (3) random horizontal flips. These are standard practice in deep learning to increase models' robustness [36, 37]. The image reductions do not affect the validity of the associated fish data, as fish would have moved beyond the 3 x 189 cm cm cropped slivers of the quadrat during surveys. Moreover, this data augmentation is a common practice that improves model accuracy for small training sets [38].

We started with standard network architecture including AlexNet [37] and VGG-16 [39], but these were vastly over parametrised for our limited data set. We therefore implemented our own shallow neural network.

The network contained two layers (Fig. 4). Layer 1 composed of a 2D convolution with 8 channels and a 6x6 pixel kernel. Layer 2 also composed of a 2D convolution, but with 16 channels and a 4x4 pixel kernel. In both layers, the 2D convolution had 2 pixel padding and preceded a rectified linear unit (ReLU) and 2D max pooling with a kernel size of 2 pixels. The network trained using Cross-entropy loss [38] and optimized using the Adam algorithm [40]. The learning

rate was held constant at 0.001. An epoch is an iteration through the network of data that the network has already seen. For each epoch, a random mini-batch of 50 images from the training data fed through the network. The ConvNet was implemented in PyTorch [41].²

Figure 4: Convolutional neural network (ConvNet) architecture.

We also tested adding a third layer and batch normalization to the network, but they over-fit too quickly to the training set so the more lightweight network was best. We tried dividing the data 50/25/10 into training, validation, and test sets so that we could optimise hyper-parameters such as kernel size, but this only worsened over-fitting.

2.4 Assessment

The accuracy of the ConvNet ($A_{ConvNet}$) was the number of correctly classed images from the training set divided by the total number of images in the training set:

$$A_{ConvNet} = \frac{\text{\#Correctly classified images}}{\text{\#Total test images}} \quad (1)$$

It was compared to the accuracy of a “dumb classifier” (A_{wg}) that classified the test set by making random, weighted guesses based on the proportion of each class in the training set:

²www.github.com/gracecalvertyoung/ConvNet-for-Ecology

$$A_{wg} = \frac{1}{\# \text{Total test images}} \sum_{i=1}^{\# \text{ classes}} \left(\frac{\# \text{Training images from class } i}{\# \text{Total training images}} \right)^2 \quad (2)$$

For each number of classes, the data was re-shuffled three times into the training and test sets. The ConvNet was then evaluated once on each shuffling, so $A_{ConvNet}$ was computed three times for each number of class. This helped ensure that results were not determined by a “lucky” division of the data where patterns in the test set closely matched the training set. No parameters changed between evaluations.

2.5 Data & Code Availability

Data and code available on an open GitHub repository upon paper acceptance.

3 Results & Discussion

The ConvNet predicted the correct class of fish with a maximum accuracy of 68% (Table 1), and $A_{ConvNet}$ was consistently greater than A_{wg} for 2–7 classes (Fig. 5 shows $A_{ConvNet} - A_{wg}$). The discrepancy at 8 classes (Fig. 5) likely results from there being insufficient data to generalize trends, as there were only 5–10 images in the training set for each classification at the 8 class granularity. Alternatively, it could be an ecological threshold; *e.g.*, the difference between reef texture for two bins in the 8 class granularity is truly random. In this problem the upper threshold for attainable accuracy is likely considerably lower than 100% because numerous stochastic factors affect fish abundance [42].

Hundreds to thousands more training datum should improve network accuracy. In the coming years, it is indeed expected that there will be vastly more data available, as techniques for creating 3D reef models become evermore time and cost-effective and are adopted by researchers and citizen scientists alike [43]. The underwater 3D models could be generated from any method (*e.g.*, photogram-

Table 1: Accuracy of random weighted guess (A_{wg}) compared to the accuracy of the ConvNet ($A_{ConvNet}$). Averages are from three random shufflings of the data into the training and test sets. Results are reported from the 100-epoch mark, after which the network tended to overfit to the training set.

	Number of Classes						
	2	3	4	5	6	7	8
Average A_{wg}	47%	30%	23%	19%	15%	13%	11%
Average $A_{ConvNet}$	63%	40%	36%	28%	25%	17%	8%
Top A_{wg}	50%	34%	25%	20%	16%	15%	12%
Top $A_{ConvNet}$	68%	44%	44%	32%	28%	20%	12%

Figure 5: Increase in accuracy of the convolutional neural network (ConvNet) compared to the accuracy of a “dumb classifier” that makes a random guess about each image in the test set based on the distribution of data in training set. For each number of classes, the ConvNet was tested on three random sortings of the data into the training and test sets.

metry, sonar, LiDAR) and still be applied to the ConvNet framework as long as they are in a standardized format. Additional descriptors could also augment the data, such as site information (if different between inputs) or scaling factors (if quadrats are different sizes); these could be incorporated following what Huang et al. [44] developed for adding text descriptors into their image classification problem. The caveat is that any extra parameters require even more training data [45]. Including the colours of points may also improve accuracy, especially if colour reliably corresponded to algae, coral cover, or other ecological features that help predict fish abundance. To do so, colour fidelity should be consistent across the data, which can be accomplished through modifications to the photogrammetry method [46].

Another way to improve accuracy may be to train the network on 3D rep-

176 representations rather than images. A 3D model would contain features that are
177 collapsed in the image representation, such as the canopy structures and crevices
178 that provide creatures refuge. The irregular format of point clouds traditionally
179 poses challenges for deep learning networks, but this hurdle can be overcome by
180 voxel transformations or with more advanced networks [47, 48]. Again, however,
181 training more network parameters requires more data. Additionally, benchmark-
182 ing the neural network’s accuracy against traditional metrics of reef structural
183 complexity, such as rugosity, rather than weighted guesses may be more revealing.

184 Once it achieves considerable accuracy, this network could be used to validate
185 artificial reef designs aiming to maximize fish abundance, predict hotspots that
186 should be conservation priorities, or conduct a range of other science, conserva-
187 tion, and management tasks. The framework is versatile and could also learn
188 other ecological patterns, from land or sea, that are not obvious to the human
189 eye.

190 4 Conclusion

191 This paper has demonstrated that a neural network has the capacity to learn
192 fish abundance patterns from raw representations of the underlying coral reef
193 topography. The results are encouraging considering the small amount of samples
194 used to train the network. The ConvNet achieved an accuracy of 68% from just
195 85 images, and this accuracy should improve with more data. Once extended with
196 hundreds more data points, this method of learning patterns could reveal insights
197 into how fish use structure and be a tool for making automated predictions of
198 where fish will aggregate. Additionally, the versatile framework could learn other
199 complex ecological patterns, from land or sea.

200 **5 Acknowledgements**

201 The Marshall Aid Commemoration Commission funded author GCY and Op-
202 eration Wallacea funded and facilitated data collection.

203 **6 Author Contributions**

204 Conceptualization: GCY, VAP; Data curation: GCY; Formal analysis: GCY;
205 Funding acquisition: GCY, VAP.; Methodology: GCY, VB, VAP; Project admin-
206 istration: VAP; Resources: VAP; Software: GCY, VB; Supervision: VAP; Vali-
207 dation: GCY, VB; Visualization: GCY; Writing – original draft: GCY; Writing
208 – review & editing: GCY, VAP.

209 **7 Competing Interests**

210 The authors declare there are no competing interests.

211 **References**

- 212 [1] Andrew Scott Hoey, Shaun K Wilson, Morgan S Pratchett, Andrew S
213 Hoey, and Shaun K Wilson. Reef degradation and the loss of critical
214 ecosystem goods and services provided by coral reef fishes. *Current Opin-*
215 *ion in Environmental Sustainability*, 7:37–43, 2014. ISSN 1877-3435. doi:
216 10.1016/j.cosust.2013.11.022.
- 217 [2] Alice Rogers, Julia L. Blanchard, and Peter J. Mumby. Vulnerability of
218 coral reef fisheries to a loss of structural complexity. *Current Biology*, 24(9):
219 1000–1005, 2014. ISSN 09609822. doi: 10.1016/j.cub.2014.03.026.
- 220 [3] D. J. Coker, N. A J Graham, and M. S. Pratchett. Interactive effects of live
221 coral and structural complexity on the recruitment of reef fishes. *Coral Reefs*,
222 31(4):919–927, 2012. ISSN 07224028. doi: 10.1007/s00338-012-0920-1.

- 223 [4] Adriana Vergés, Mathew A. Vanderklift, Christopher Doropoulos, and
224 Glenn A. Hyndes. Spatial patterns in herbivory on a coral reef are influ-
225 enced by structural complexity but not by algal traits. *PLoS ONE*, 6(2),
226 2011. ISSN 19326203. doi: 10.1371/journal.pone.0017115.
- 227 [5] David O Obura. Resilience and climate change: Lessons from coral reefs
228 and bleaching in the Western Indian Ocean. *Estuarine, Coastal and Shelf
229 Science*, 63(3):353–372, 2005. doi: 10.1016/j.ecss.2004.11.010.
- 230 [6] C. A. Radford, J. A. Stanley, S. D. Simpson, and A. G. Jeffs. Juvenile coral
231 reef fish use sound to locate habitats. *Coral Reefs*, 30(2):295–305, 2011. ISSN
232 07224028. doi: 10.1007/s00338-010-0710-6.
- 233 [7] J. L. Johansen, D. R. Bellwood, and C. J. Fulton. Coral reef fishes exploit
234 flow refuges in high-flow habitats. *Marine Ecology Progress Series*, 360:219–
235 226, 2008. ISSN 01718630. doi: 10.3354/meps07482.
- 236 [8] B Gratwicke and Mr Speight. Effects of Habitat Complexity on Caribbean
237 Marine Fish Assemblages. *Marine Ecology Progress Series*, 292:301–310,
238 2005. ISSN 0171-8630. doi: 10.3354/meps292301.
- 239 [9] Shaun K. Wilson, Nicholas A J Graham, Morgan S. Pratchett, Geoffrey P.
240 Jones, and Nicholas V C Polunin. Multiple disturbances and the global
241 degradation of coral reefs: Are reef fishes at risk or resilient? *Global Change
242 Biology*, 12(11):2220–2234, 2006. ISSN 13541013. doi: 10.1111/j.1365-2486.
243 2006.01252.x.
- 244 [10] Margaret F. J. Wilson, Brian O’Connell, Colin Brown, Janine C. Guinan, and
245 Anthony J. Grehan. Multiscale Terrain Analysis of Multibeam Bathymetry
246 Data for Habitat Mapping on the Continental Slope. *Marine Geodesy*, 30
247 (1-2):3–35, 2007. ISSN 0149-0419. doi: 10.1080/01490410701295962.
- 248 [11] J. Tews, U. Brose, V. Grimm, K. Tielbörger, M. C. Wichmann, M. Schwa-
249 ger, and F. Jeltsch. Animal species diversity driven by habitat heterogene-

- ity/diversity: the importance of keystone structures. *Journal of Biogeography*, 31(1):79–92, 2004. ISSN 03050270. doi: 10.1046/j.0305-0270.2003.00994.x.
- [12] N. A. J. Graham and K. L. Nash. The Importance of Structural Complexity in Coral Reef Ecosystems. *Coral Reefs*, 32:315–326, 2013. ISSN 07224028. doi: 10.1007/s00338-012-0984-y.
- [13] E Luckhurst and K Luckhurst. Analysis of the Influence of Substrate Variables on Coral Reef Fish Communities. *Marine Biology*, 323(49):317–323, 1978.
- [14] Curt D. Storlazzi, Peter Dartnell, Gerald A. Hatcher, and Ann E. Gibbs. End of the chain? Rugosity and fine-scale bathymetry from existing underwater digital imagery using structure-from-motion (SfM) technology. *Coral Reefs*, 2016. ISSN 0722-4028. doi: 10.1007/s00338-016-1462-8.
- [15] B. M. Costa, T. A. Battista, and S. J. Pittman. Comparative evaluation of airborne LiDAR and ship-based multibeam SoNAR bathymetry and intensity for mapping coral reef ecosystems. *Remote Sensing of Environment*, 113(5): 1082–1100, 2009. ISSN 00344257. doi: 10.1016/j.rse.2009.01.015.
- [16] Elisa Casella, Antoine Collin, Daniel Harris, Sebastian Ferse, Sonia Bejarano, Valeriano Parravicini, James L. Hench, and Alessio Rovere. Mapping coral reefs using consumer-grade drones and structure from motion photogrammetry techniques. *Coral Reefs*, 36(1), 2016. ISSN 07224028. doi: 10.1007/s00338-016-1522-0.
- [17] Jhr Burns, D Delparte, Rd Gates, and M Takabayashi. Integrating Structure-from-Motion Photogrammetry with Geospatial Software as a Novel Technique for Quantifying 3D Ecological Characteristics of Coral Reefs. *PeerJ*, 3 (e1077), 2015. ISSN 2167-8359. doi: 10.7717/peerj.1077.
- [18] Shaun K. Wilson, Russ C. Babcock, Rebecca Fisher, Thomas H. Holmes,

- 277 James A.Y. Moore, and Damian P. Thomson. Relative and combined ef-
278 fects of habitat and fishing on reef fish communities across a limited fishing
279 gradient at Ningaloo. *Marine Environmental Research*, 81:1–11, 2012. ISSN
280 01411136. doi: 10.1016/j.marenvres.2012.08.002.
- 281 [19] Brian K. Walker, Lance K. B. Jordan, and Richard E. Spieler. Relationship of
282 Reef Fish Assemblages and Topographic Complexity on Southeastern Florida
283 Coral Reef Habitats. *Journal of Coastal Research*, 10053:39–48, 2009. ISSN
284 0749-0208. doi: 10.2112/SI53-005.1.
- 285 [20] Simon J. Pittman and Kerry A. Brown. Multi-Scale Approach for Predicting
286 Fish Species Distributions across Coral Reef Seascapes. *PLoS ONE*, 6(5):
287 e20583, 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0020583.
- 288 [21] Sonia Bejarano, Peter J. Mumby, and Ian Sotheran. Predicting struc-
289 tural complexity of reefs and fish abundance using acoustic remote sensing
290 (RoxAnn). *Marine Biology*, 158(3):489–504, 2011. ISSN 00253162. doi:
291 10.1007/s00227-010-1575-5.
- 292 [22] Esteban A. Agudo-Adriani, Jose Cappelletto, Francoise Cavada-Blanco, and
293 Aldo Croquer. Colony geometry and structural complexity of the endangered
294 species *Acropora cervicornis* partly explains the structure of their associated
295 fish assemblage. *PeerJ*, 4(e1861), 2016. ISSN 2167-8359. doi: 10.7717/peerj.
296 1861.
- 297 [23] M. González-Rivero, A. R. Harborne, A. Herrera-Reveles, Y.-M. Bozec,
298 A. Rogers, A. Friedman, A. Ganase, and O. Hoegh-Guldberg. Linking
299 fishes to multiple metrics of coral reef structural complexity using three-
300 dimensional technology. *Scientific Reports*, 7(1):13965, 2017. ISSN 2045-
301 2322. doi: 10.1038/s41598-017-14272-5.
- 302 [24] Brian Gratwicke and Martin R. Speight. The Relationship Between Fish
303 Species Richness, Abundance and Habitat Complexity in a Range of Shallow

- 304 Tropical Marine Habitats. *Journal of Fish Biology*, 66(3):650–667, 2005.
305 ISSN 00221112. doi: 10.1111/j.1095-8649.2005.00629.x.
- 306 [25] Alastair R. Harborne, Peter J. Mumby, and Renata Ferrari. The effectiveness
307 of different meso-scale rugosity metrics for predicting intra-habitat variation
308 in coral-reef fish assemblages. *Environmental Biology of Fishes*, 94(2):431–
309 442, 2012. ISSN 03781909. doi: 10.1007/s10641-011-9956-2.
- 310 [26] Christopher H R Goatley and David R Bellwood. The Roles of Dimensional-
311 ity, Canopies and Complexity in Ecosystem Monitoring. *PLoS ONE*, 6(11),
312 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0027307.
- 313 [27] Jesús Aguirre-Gutiérrez, Luísa G. Carvalheiro, Chiara Polce, E. Emiel van
314 Loon, Niels Raes, Menno Reemer, and Jacobus C. Biesmeijer. Fit-for-
315 Purpose: Species Distribution Model Performance Depends on Evaluation
316 Criteria - Dutch Hoverflies as a Case Study. *PLoS ONE*, 8(5), 2013. ISSN
317 19326203. doi: 10.1371/journal.pone.0063708.
- 318 [28] Kirsty L. Nash, Nicholas A.J. Graham, Shaun K. Wilson, and David R.
319 Bellwood. Cross-scale Habitat Structure Drives Fish Body Size Distributions
320 on Coral Reefs. *Ecosystems*, 16(3):478–490, 2013. ISSN 14329840. doi:
321 10.1007/s10021-012-9625-0.
- 322 [29] Simon J. Pittman, Bryan M. Costa, and Tim A. Battista. Using Lidar
323 Bathymetry and Boosted Regression Trees to Predict the Diversity and
324 Abundance of Fish and Corals. *Journal of Coastal Research*, 10053:27–38,
325 2009. ISSN 0749-0208. doi: 10.2112/SI53-004.1.
- 326 [30] Anders Knudby, Ellsworth LeDrew, and Alexander Brenning. Predic-
327 tive mapping of reef fish species richness, diversity and biomass in Zanz-
328 ibar using IKONOS imagery and machine-learning techniques. *Remote*
329 *Sensing of Environment*, 114(6):1230–1241, 2010. ISSN 00344257. doi:
330 10.1016/j.rse.2010.01.007.

- [31] Bryan Costa, J. Christopher Taylor, Laura Kracker, Tim Battista, and Simon Pittman. Mapping reef fish and the seascape: Using acoustics and spatial modeling to guide coastal management. *PLoS ONE*, 9(1), 2014. ISSN 19326203. doi: 10.1371/journal.pone.0085555.
- [32] A. Mahmood, M. Bennamoun, S. An, F. Sohel, F. Boussaid, R. Hovey, G. Kendrick, and R. B. Fisher. Automatic annotation of coral reefs using deep learning. *OCEANS 2016 MTS/IEEE Monterey, OCE 2016*, 2016. doi: 10.1109/OCEANS.2016.7761105.
- [33] Jonas Osterloff, Ingunn Nilssen, Johanna Järnegren, Pål Buhl-Mortensen, and Tim W. Nattkemper. Polyp activity estimation and monitoring for cold water corals with a deep learning approach. In *International Conference on Pattern Recognition 2nd Workshop on Computer Vision for Analysis of Underwater Imagery*, 2016. ISBN 9781509058709. doi: 10.1109/CVAUI.2016.14.
- [34] G. C. Young, S. Dey, A. D. Rogers, and D. Exton. Cost and time-effective method for multi-scale measures of rugosity, fractal dimension, and vector dispersion from coral reef 3D models. *Plos One*, 12(4):e0175341, 2017. ISSN 1932-6203. doi: 10.1371/journal.pone.0175341.
- [35] Vasileios Belagiannis, Christian Rupprecht, Gustavo Carneiro, and Nassir Navab. Robust optimization for deep regression. *Proceedings of the IEEE International Conference on Computer Vision*, 2015 Inter:2830–2838, 2015. ISSN 15505499. doi: 10.1109/ICCV.2015.324.
- [36] Chiyuan Zhang, Samy Bengio, Moritz Hardt, Benjamin Recht, and Oriol Vinyals. Understanding deep learning requires rethinking generalization. *arXiv preprint arXiv:1611.03530*, 2017. ISSN 10414347. doi: 10.1109/TKDE.2015.2507132.
- [37] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. ImageNet Clas-

- sification with Deep Convolutional Neural Networks. *Advances In Neural Information Processing Systems*, 2012. ISSN 10495258. doi: <http://dx.doi.org/10.1016/j.protcy.2014.09.007>.
- [38] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, Cambridge, MA, 2016. ISBN 978-0262035613.
- [39] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *3rd International Conference on Learning Representations*, 2015. ISBN 0950-5849. doi: 10.1016/j.infsof.2008.09.005.
- [40] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. In *3rd International Conference on Learning Representations (ICLR2015)*, 2015. ISBN 9781450300728. doi: <http://doi.acm.org.ezproxy.lib.ucf.edu/10.1145/1830483.1830503>.
- [41] Gregory Chanan Adam Paszke, Sam Gross, Soumith Chintala. PyTorch, 2017.
- [42] Steven P Newman, Erik H Meesters, Charlie S Dryden, Stacey M Williams, Cristina Sanchez, Peter J Mumby, and Nicholas V C Polunin. Reef Flattening Effects on Total Richness and Species Responses in the Caribbean. *Journal of Animal Ecology*, 84(6):1678–1689, 2015. doi: 10.1111/1365-2656.12429.
- [43] Vincent Raoult, Peter A. David, Sally F. Dupont, Ciaran P. Mathewson, Samuel J. O’Neill, Nicholas N. Powell, and Jane E. Williamson. GoPros as an underwater photogrammetry tool for citizen science. *PeerJ*, 4:e1960, 2016. ISSN 2167-8359. doi: 10.7717/peerj.1960.
- [44] Feiran Huang, Xiaoming Zhang, Zhoujun Li, Tao Mei, Yueying He, and Zhonghua Zhao. Learning Social Image Embedding with Deep Multimodal Attention Networks. In *Thematic Workshops, October 23–27, 2017, Moun-*

- tain View, CA, USA, 2017. ISBN 9781450354165. doi: 10.1145/3126686.
3126720.
- [45] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York City, second edition, 2009.
- [46] M. Bryson, M. Johnson-Roberson, O. Pizarro, and S. B. Williams. Colour-consistent structure-from-motion models using underwater imagery. In *Robotics: Science and Systems VIII*, 2013.
- [47] Gernot Riegler, Ali Osman Ulusoy, and Andreas Geiger. OctNet: Learning Deep 3D Representations at High Resolutions. *arXiv preprint*, 2016. ISSN 1063-6919. doi: 10.1109/CVPR.2017.701.
- [48] Yizhak Ben-Shabat, Michael Lindenbaum, and Anath Fischer. 3D Point Cloud Classification and Segmentation using 3D Modified Fisher Vector Representation for Convolutional Neural Networks. *arXiv preprint*, 2017.