Sameer Khan, Gordon Doore, Drew Hinton
Professor Tahiya Chowdhury
JP-297 D
January 09, 2023

Annotation and Datasheet for "Coral Reef" Data-set

Getting the dataset of "Coral Reef," which is taken on Maldives, French Polynesia, and other Pacific Atolls, we chose to label the data with Dashdoodler. DashDoodler is a web application built for image segmentation with minimal supervision which helps images to be labeled by the pattern followed by AI. Thus, first, let's install dashdoodler. Instructions for installing dash doodler:

- Go to https://github.com/Doodleverse/dash_doodler, scroll to "Installation"
- Run git command to download dash doodler
- Install anaconda at https://docs.anaconda.com/anaconda/install/index.html
- Create conda environment and install dependencies as instructed
- Find dash_doodler in your files, should be in your user folder

Now, as we are done installing the dash-doodler in our computer using terminal, we will move the reef images to the assets folder in the dash doodler folder so that we can get our selected images labeled and saved. Before that, one thing we need to do is to change classes.txt to match following format:

- vegetation
- reef flat
- Ocean

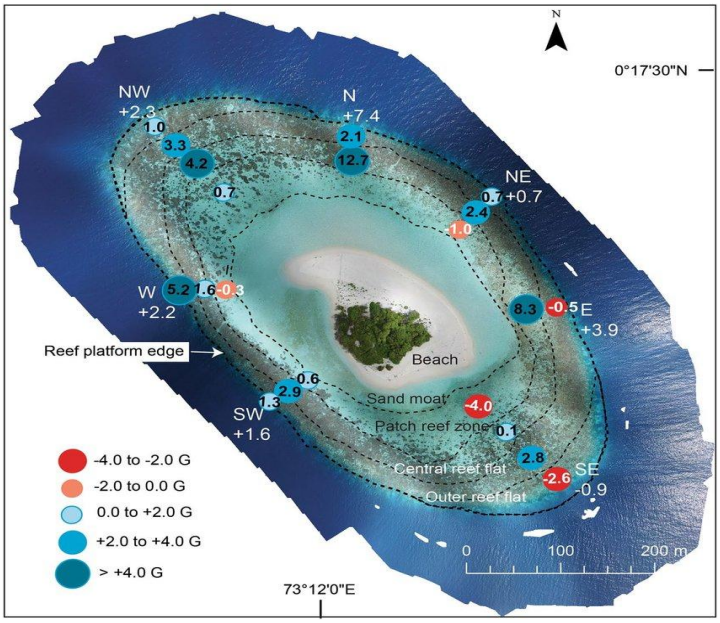
Next, once DashDoodler is installed, we open the conda terminal and enter the USER/dash_doodler/ directory and then run "*python doodler.py*" in the terminal. If there is an error saying module doodler_engine is not installed, just install the package listed and try again. Additional issues may be caused from improper installation. Maybe we can check if we are in the anaconda environment. For that, we need (dashdoodler) instead of (base) before the current working directory is listed. If it says (base) we are not in our conda environment, which we just created, then we just need to type:

conda activate dashdoodler

When doodler.py is run, we should be given a URL to access with our browser. For note, we need to remember that it can be run in Edge, Chrome, and Firefox. Next, we set the user ID to something that distinguishes us and select an image to annotate. While selecting an image, we need to ensure that the image is of a small size, ideally much smaller than 3000 x 3000 pixels to avoid memory leaks.

Next, we switched to the "imagery and controls" tab in order to begin annotating. After that, we selected a class and color liberally into the feature: green areas are "vegetation," white/sandy areas are "reef flat," and anything water is "ocean." Once this was done, we checked the compute/show segmentation box.

Now, referring to the image below, we can learn the basics of Atoll feature classification.



Massive corals maintain a positive carbonate budget of a Maldivian upper reef platform despite major bleaching event - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Average-net-carbonate-budget-rates-G-in-different-eco-geomorphologic-zones-around_fig5_332621102 [accessed 9 Jan, 2023]

Figure: Basics of Atoll feature Classification for Coral Reef

Labeled reef flat should include outer reef flat, central reef flat, and patch reef zone. Obvious vegetation should be labeled for vegetation, but what looks like sand should be labeled for reef. This is the format of the data which was already labeled. Finally, everything else will be ocean.

In terms of pre-labeling image preparation, one should perform canny edge detection and overlay this into a given image to be labeled. This can be implemented by using cv.Canny() on an image, then "overlaying" by mapping white pixels into the image to be labeled. This makes it easier for dashdoodler to identify the different classes.

Other image preparation methods did not work effectively. In the future, data augmentation techniques may be used to increase the size of the dataset. This is to be determined in the future.