A Convolutional Neural Networks Based Coral Reef Annotation and Localization

Rohit Raj Gunti and Abebe Rorissa

A Convolutional Neural Networks based Coral Reef Annotation and Localization

Rohit R. Gunti [1] and Abebe Rorissa [1,2]

[1] *University at Albany, Albany, UAlbany, USA*
[2] *University of Tennessee, Knoxville, UAlbany, USA*

Abstract
The purpose of this study was to examine the effectiveness and flexibility of our image retrieval system by participating in the ImageCLEFcoral 2021 challenge. The system was developed for object detection and identification in any dataset. Initially, multiple trials were conducted to train the system with patterns of thirteen substrates and searching the relationships between them. Since the key for better performance of machine learning systems is repeated inputs, we used datasets with three distinct groups, each group having a different characterization of substrates. For submissions to the ImageCLEFcoral challenge, we tested the system with the provided test dataset where it was able to find the patterns and relationships between the substrates in a massive amount of data that was also too complex. We sought to extract high-level image features and use deep learning and the CNN-RNN neural network. We obtained an acceptable range of accuracies for each characterization of substrates, although the average accuracy was 70 percent.

Keywords
Image classification; coral reef; convolutional neural networks; annotation and localization

1. Introduction

Conventional classification techniques employ spectral responses from coral types based on single-pixel values without capturing spatial information [1]. Several attempts, such as the use of fractal, spatial autocorrelation, and spatial co-occurrence matrices, have been made to improve the spectral analysis of remotely sensed image data. Especially when dealing with complex spatial features and many different coral substrates like fire coral and branching coral that share a similar spatial pattern. In recent years, wavelet transform has been investigated and applied in image processing due to its innovative mathematical framework for multiscale time-frequency signal analysis. The literature on image processing and analysis with wavelet transform generally focus on image compression, image fusion, image watermarking, face-image-detection, image noise removal, identification of tumorous regions and microcalcification clusters, image segmentation, etc. However, image classification, a type of categorization of image data using spectral, spatial, and temporal information that assigns the pixels of a continuous raster image to discrete categories concerning wavelet transforms, has not been explored enough. Wavelet transform opens the possibility for capturing image features at different scales and are more discriminative than spectral features.

In this working note, a novel frequency-based classification framework and algorithm is proposed using an overcomplete decomposition procedure at multiple scales to examine if the proposed method can effectively identify detailed coral reef substrates in Marine Technology Research Unit dataset [2] [3]. Because overcomplete wavelet can generate spatial arrangements of objects and features at any

scale level (infinite-scale analysis), a frequency-based multiscale classification algorithm using an overcomplete wavelet proposed by [4] and superior to the dyadic-based wavelet classifier approach was utilized for the current study. Furthermore, because the approach has the flexibility to use any window size, we believe that it can be applied at any level and at either high spatial or medium-resolution structure-from-motion photogrammetry (action cameras attached to drones). The newly developed algorithm is hereafter referred to as Wave-CLASS [4].

This study's classification decision rule computes one distance for each class among all bands; hence, five distances are computed for the classification decision. The training sample with the shortest distance to the unknown feature vector wins the new texture. In other words, the training class that is closest to the local window is assigned to the center of that window.

Mallet [5] initiated the multiresolution analysis theory using the orthonormal wavelet basis. Wavelet decompositions can be categorized into dyadic and overcomplete approaches. The idea of employing overcomplete wavelet decomposition is motivated by the fact that it can provide translational invariant features. The Overcomplete decomposition approach employed in this study omits the down sampling procedure and produces four-texture information with the same dimension of the original image at an infinite scale. This advantage is crucial for classification. A higher decomposition level can be expected to improve accuracy. A set of feature vectors is used to identify detailed coral reef substrates.

An entire image, which consists of both unknown blocks (blocks that need to be classified) and sample blocks (training samples), undergoes a multiscale overcomplete wavelet transform. Several training samples need to be selected for each coral reef substrate class. A spatial measure can be used as a feature to represent unknown blocks and sample blocks. The distance between the feature vectors leads to supervised classification, such as the Euclidean distance classifier.

The overall accuracy, producer's accuracy, user's accuracy, and kappa coefficient were generated using the error matrix. A minimum of 50 sample points for each benthic substrate is generally suggested for any image classification accuracy assessment. Thus, a total of 525 samples was generated using a stratified random sampling approach with a minimum per size sample size of 35 points, resulting in an average of 75 points per class (a total of seven classes). Table 3 shows the overall accuracy, producer's accuracy, user's accuracy, and kappa coefficient produced by the Wave-CLASS algorithm.

2. Tasks performed

For our project, we initially conducted satellite image classification where we trained a MATLAB based system high spatial resolution images as our datasets. During the second phase of the project, we trained the system using high spatial resolution images along with medium resolution ones to increase system flexibility. Additionally, we used some images from the Internet to train and identify individual substrates.

Initially, the training and testing was done using a training dataset to analyze system performance and calculate the standardized metrics like precision, recall, Intersection over union, dice coefficient, and coefficient kappa. After training using single substrate online images, subsequent trainings used images that have a smaller number of substrates and some with more substrates than the previous images. To improve our results and the accuracy of the system in detecting some of the substrates, we started the training and testing from scratch by removing third-party images and adding images that have fewer number of substrates. The first training phase consisted of eight (8) trials covering a total of 13 substrates, 120 images, and 405 annotations (see Table 1).

Table 1

Images used for the trials in the first training phase

Trials	1st trial	2nd trial	3rd trial	4th trial	5th trial	6th trial	7th trial	8th trial
Images trained	12	15	18	13	11	18	18	15

Because our first phase training involved saving the annotations in array format, removing and retraining was important in analyzing the training patterns. During the 2nd trial, with input from the 1st

trial, we trained 96 images and 515 annotations that showed better test results. However, for the actual submission of our results to ImageCLEFcoral (AIcrowd), we divided the training dataset and test dataset (downloaded from the challenge's resources) into four groups

1. Group 1: trained and tested images from dominica-cabrits folder
2. Group 2: trained and tested images from spermonde-keke folder
3. Group 3: trained and tested images from Seychelles-BL
4. Group 4: trained and tested images from PK-20180729-02

The system was at its highest accuracy when classifying test images separately compared to the classification that was done during trial rounds.

3. Methods

3.1. Approaches Used

A six-step method was used to process and conduct the classification of an image:
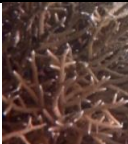
1. Image Acquisition: Capturing the input image from the source file using the "uigetfile" and "imread" functions. However, if the image has not been acquired satisfactorily, the intended tasks may not be achievable, even with some form of image enhancement.
2. Image Resizing or Scaling: A process that involves a trade-off between efficiency, smoothness, and sharpness.
3. Fuzzy C Means Clustering: If the Marine Technology Research Unit dataset input image bandwidth and cell value are between 0.5 to 0.9 and 0.1 to 0.5, respectively, then that class is identified as hard coral mushroom. The same goes for branching, soft coral, sponge and every other annotated class.
4. Image Segmentation: The process of partitioning an image into multiple segments (set of pixels or superpixels). The goal is to represent the image as something more meaningful and to assign a label to every pixel in an image such that pixels with the same label share specific characteristics.
5. Feature Extraction: A convolutional layer condenses the input by extracting features of interest from it and produces feature maps in response to different feature detectors.
6. Classification: This is the top layer of a network that collects the final convoluted feature and returns a column vector where each row points towards a class.

3.2. Test dataset

The dataset for the ImageCLEFcoral 2021 task consists of images from coral reefs collected through the Marine Technology Research Unit at the University of Essex from around the world. While the images vary in terms of their quality, with some that are blurry and others having poor color balance, they come with annotations of 13 types of substrates ranging from hard coral branching to soft-coral – Gorgonian to Sponge – Barrel. With no oceanographers on our team to identify the substrates while training the images, we used every possible online resource for substrate detection. Table 3 presents three types of sample images for each substrate.

Table 2
Table title

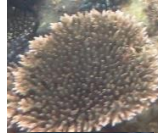| Substrates	Type 1	Type 2	Type 3
Hard Coral – Branching			

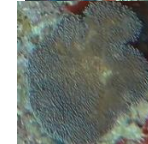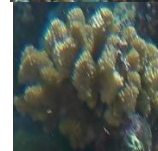Hard Coral – Submassive
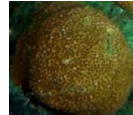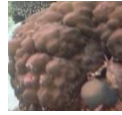
Hard Coral – Boulder
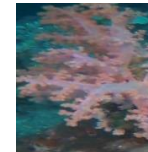
Hard Coral – Encrusting
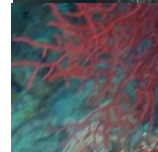
Hard Coral – Table

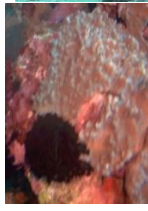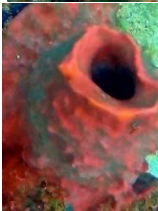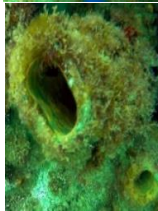Hard Coral – Foliose

Hard Coral – Mushroom
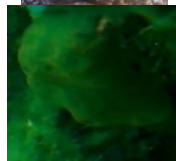
Soft Coral

Soft Coral – Gorgonian

Sponge

Sponge – Barrel

Fire Coral – Millepora

Algae - Macro or Leaves

4. Resources Employed

We used a Windows 10 PC with Intel Processor and 8 GB RAM. For training and classification purposes, we used MATLAB for Windows. For errors and calculations we used the Anaconda distribution of Python IDE 3.7.

5. Results

A total of 216 images in the trial round and 485 images with 2611 annotations in the actual run round were tested. Several of them were successfully classified while some were partially classified. The first step is the processing of the input image by resizing it. Figure 1 shows the input image during the trial round (left) and the resized image (right).
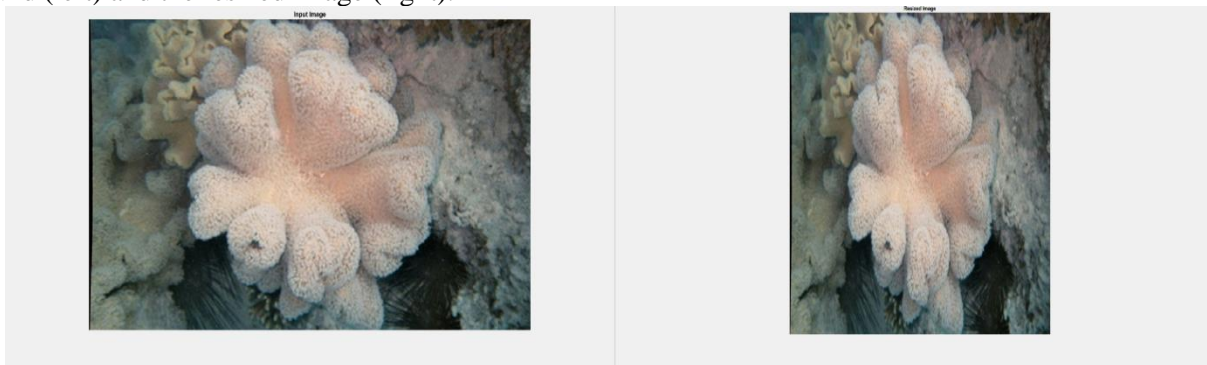
Figure 1: Input image for the trial round and the resized version of the image

To determine the exact part of size of the substrates in an image, we drew the region boundaries by forming the mask and applying the Sobel edge detector to a labelled image. In short, we applied thresholding to narrow in the resulting image. To find the precise boundaries of the regions, we converted the resulting RGB image to a binary format and measured the data points at the corners of the substrate. In the images (Figure 2), each data point was measured from the binary image covering the exact corners of the substrates. Noting the values of X axes and Y axes, a bounding box was drawn covering the area of the substrate.
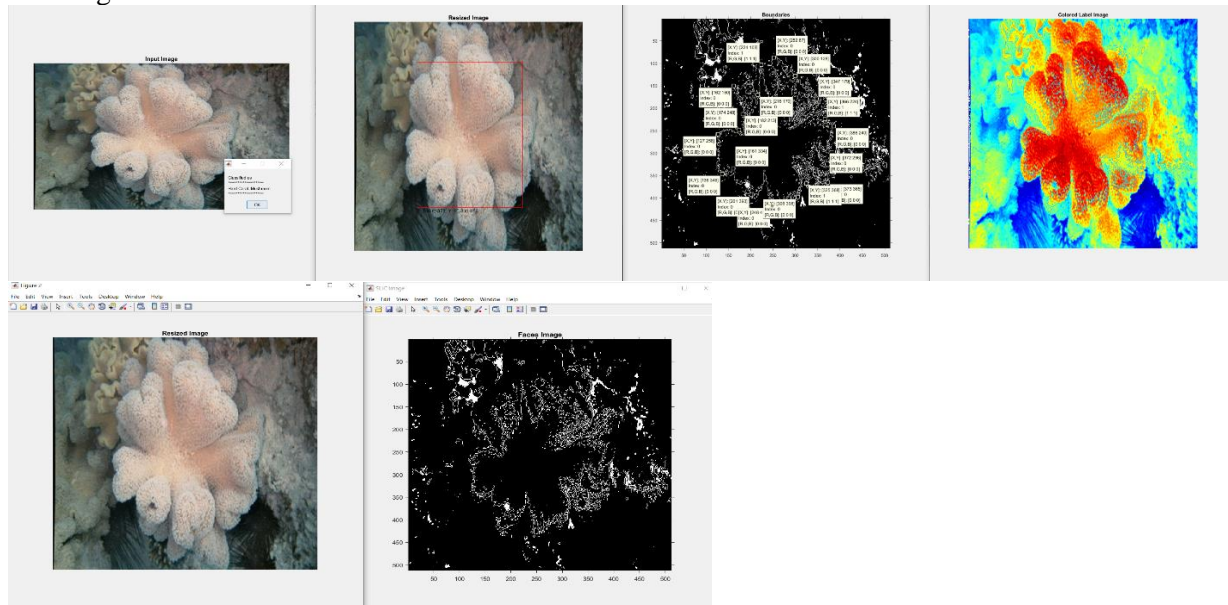
Figure 2: Input image for the trial round and the resized version of the image

Because objects had to be detected in every image, we used the label2rgb function to determine colors of each object based on the number of objects in the label matrix of the resized images. Additionally, we used optional perimeters like 'spring' setting the background pixels to the color cyan and randomizing color assignment to the labels. Through these color filters, we were able to average all colormap values within the range of 0.5 to 1 and average all the values to get the highest confidence value. After training the system using half of the images from the training dataset, we then tested the system on the other half. The coordinates of the tested images were noted in a text editor to validate the results.

Figure 3: Transforming the resized image into a colored label image

For the annotation and localization task, we separately trained 472 images and tested 485 images with 2611 annotations. The output of the tests was submitted to the AIcrowd system in the following format: The height and width of the boundaries is measured through the bounding box by subtracting the lowest x datapoint (Xmin) from the highest x datapoint (Ymin). We noted all the values X min, Y min, height, and width in the given format. A scale of 1 unit per 50 px was used for both the x and y axes.

[image_ID];[substrate1]
[[confidence1,1]:][width1,1]x[height1,1]+[xmin1,1]+[ymin1,1],[[confidence1,2]:][width1,2]x[height1,2]+[xmin1,2]

Each RGB image was converted to a binary format to locate the points around the boundaries. A data tip pointer tool was used to draw the polygons and all the data tip information was exported to a MATLAB workspace. Figure 4 presents the results for the annotations and localizations task.

Figure 4: Results of the annotations and localizations task

Some of the coding was done on the Anaconda distribution of Python IDE 3.7 to remove all the errors such as wrong format, unnecessary spaces, commas, semicolons, and colons. This was done to increase axes limit from 513 to 2000 and to check the different results for score validity. Figure 5 presents the increased axes results for the annotation and localization task.

Figure 5: Increased axes results for the annotation and localization task

Table 3 presents results of the classification tests conducted on the training dataset in terms of Precision, Recall, Jaccard coefficient, Dice coefficient, Kappa coefficient, and Hue ratio for every substrate. Our tests also achieved an average sensitivity, average specificity, and overall accuracy of 16.3, 21.1, and 17.3, respectively.

Table 3

Results of tests for the training dataset

Substrates	Precision	Recall	Jaccard coefficient	Dice coefficient	Kappa coefficient	Hue
Hard coral - Branching	0.117927	0.17712	0.12914	0.12866	0.12814	0.7
Hard Coral - Submassive	0.117782	0.17727	0.12914	0.12866	0.12847	0.6
Hard Coral - Boulder	0.118244	0.17687	0.12914	0.12866	0.12815	0.8
Hard Coral - encrusting	0.115589	0.17755	0.12914	0.12865	0.12847	0.5
Hard Coral - table	0.116410	0.17725	0.12914	0.12866	0.12847	0.6
Hard Coral - Foliose	0.117296	0.17703	0.12914	1.2866	0.12847	0.7
Hard Coral - Mushroom	0.114476	0.178439	0.12914	0.12865	0.12814	0.8
Soft Coral	0.115979	0.17808	0.12914	0.12865	0.12847	0.9
Gorgorian	0.118785	0.17700	0.12914	0.12866	0.12847	0.8
Sponge	0.116738	0.17759	0.12914	0.12866	0.12847	0.8
Barrel Sponge	0.117568	0.17726	0.12914	0.12866	0.12847	0.4
Fire Coral	0.117440	0.17760	0.12914	0.12866	0.12847	0.7
Algae	0.118422	0.17632	0.12914	0.12866	0.12847	0.2

Overall, our results from the training dataset show that the producer's and user's accuracies for all the categories are about equal for all image subsets, which is one of the critical criteria for systematic substrate mapping. Because both the training and ImageCLEFcoral data consist of coral substrate

images from four different locations and some images were downloaded from the Internet, they were classified into three categories: (1) substrates that are uniquely grown or long distanced from other substrates in the Marine Technology Research Unit dataset; (2) coral reef substrates that grow on together or short distanced from other substrates; and (3) single individual substrates that are downloaded online.

The producer's and user's accuracies of the classification of all the selected classes in the unique substrates (#1), grouped substrates (#2), and individual substrates (#3) achieved above 75% accuracy, except the user's accuracy of uniquely grown substrates (#1) (75%) and grouped substrates (#2) (71%). The user accuracy of 75% indicates that only 75% of the animals identified as a specific substrate within the uniquely grown class (#1) produced more accurate classification results, although the producer's accuracy of the same category reached 88%, which can be considered adequate. In other words, remote sensing analysis of the algorithm shows that 88% of the time, a coral reef that was identified as a substrate, say mushroom, grown with no other different coral substrate nearby was uniquely grown, whereas a user of the output map would argue that only 75% of the time the classification correctly identifies "mushroom" as actually being a mushroom. In contrast, individual class substrates from the Internet achieved high producer's (94%) and user's (99%) accuracies. We anticipated that encrusting class that are grown alone and in a group with other substrates, could potentially lower the overall accuracy. The relatively low user's accuracy could be due to the spectral similarity among substrates like a fully grown hard coral branching and partially grown fire corals having similar tree like structures, and coexistence of hard coral branching on the hard coral table.

Another critical factor is that almost all encrustings in these subsets are grouped. These encrustings are generally much smaller than those grown uniquely, where most are large stone like structures and purple in color. Most of the hard coral encrusting that were grouped with other substrates are bright (red, blue, and so forth), surrounded by dark impervious surfaces (algae). Whereas, uniquely grown substrates have many shapes and colors that are mixed or interlocked with diverse materials or complex background objects and features (e.g., stones, soil, grass, plants). The pool category results achieved 100% for both producer's and user's accuracy in both the individual and uniquely grown subsets because the spectral signatures of adherence to hard rocky surfaces are highly distinguishable from those of other coral substrate features.

The relatively low classification accuracy for rock surfaces in the grouped substrates is likely due to its smaller size, which inevitably causes mixed pixels in the training samples. The next highest producer's accuracy is 98% for the individual subset, 99% for the uniquely grown subset, and 94% for the grouped subset.

6. Future work

Our findings indicate that it will be possible to improve the accuracy of the Wave-CLASS classifier we utilized and build a supervised/unsupervised system. We believe that the algorithm and classifier can also be used in other domains and for applications in Medicine, geography, astronomy and so on. In medicine, it could help detect specific diseases and their characteristics. In geography, it could be used to study land-use for cultivating and farming. In astronomy, one could use it to detect land that is destroyed by natural disasters. The system is flexible enough that by using the wavelet algorithm to classify substrates, we could use other algorithms such as random forest in conjunction with the Wave-CLASS classifier.

7. References

[1] B. Kartikeyan, B. Gopalakrishna, M. H. Kalubarme & K. L. Majumder (1994) Contextual techniques for classification of high and low resolution remote sensing data, International Journal of Remote Sensing, 15:5, 1037-1051, DOI: 10.1080/01431169408954132.

[2] Jon Chamberlain, Antonio Campello, Wright Jessica, Clift Louis, Clark Adrian, and Alba García Seco de Herrera, (2019). Overview of ImageCLEFcoral 2019 task. In: CLEF2019

Working Notes. Vol. 2380. CEUR Workshop Proceedings. Lugano, Switzerland: CEUR-WS.org, 2019.

[3] Jon Chamberlain, Alba García Seco de Herrera, Antonio Campello, Adrian Clark, Thomas A. Oliver and Hassan Moustahfid. Overview of the ImageCLEFcoral 20201 Task: Coral Reef Image Annotation of a 3D environment, in Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the 12th International Conference of the CLEF Association (CLEF 2021), Bucharest, Romania, Springer Lecture Notes in Computer Science LNCS, September 21-24, 2021.

[4] Soe W. Myint, Tong Zhu, and Baojuan Zheng, "A Novel Image Classification Algorithm Using Overcomplete Wavelet Transforms," in IEEE Geoscience and Remote Sensing Letters, vol. 12, no. 6, pp. 1232-1236, June 2015, doi: 10.1109/LGRS.2015.2390133.

[5] Stephane Georges Mallat, "A Theory of Multiresolution Signal Decomposition: The Wavelet Representation," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 11, pp. 674-693, 1989.