**A Hybrid Feature Representation: A Novel Method for Coral Classification**

Md Zaheeda Parveen,

Research Scholar,

Department of Computer Science and Engineering,

Annamalai University, Chidambaram, India

Dr. L.R. Sudha,

Associate Professor,

Department of Computer Science and Engineering,

Annamalai University, Chidambaram, India

Dr. R. Santhoshkumar,

Associate Professor,

Department of Computer Science and Engineering,

St. Martin’s Engineering College, Secunderabad, India

**Abstract:**

For marine life, coral reefs are an absolute must. Reports have surfaced recently indicating that coral reefs are becoming less healthy and abundant. The quantity of data accessible for analysis on coral reefs has been greatly expanded by underwater imaging systems including towed diver sleds and autonomous underwater vehicles (AUVs). The picture clarity of coral reefs varies from one class to another, and there are complicated boundaries between different classes. Classifying corals is thus a difficult task. An overview of deep learning and machine learning applications to underwater image analysis, with a focus on coral species categorization, is provided in this chapter. When it comes to computer vision tasks like picture classification, object detection, and scene interpretation, deep learning methods have produced state-of-the-art outcomes. It is challenging to approach marine ecosystems from a computer vision standpoint due to the complexity of the scenes. The use of automated technology in ocean health monitoring has the potential to alleviate the laborious process of human annotation by making it easier to find and identify marine organisms. Using deep learning to efficiently classify coral species is an ambitious goal in and of itself.

Keywords: Machine Learning, Deep Learning, marine images, corals, classification, convolutional neural networks

**1 Introduction**

In marine environments, coral reefs play an essential role. Many marine species find refuge in their nutrient-rich habitat. They provide benthic organisms with nitrogen and other necessary nutrients in abundance. Along with their crucial role in nutrient recycling, they shield coasts from the destructive power of waves and storms at sea. Because many fish and other species congregate near coral reefs, these ecosystems play an important role in supporting the expanding fishing industry. The tourist business also reaps the benefits of coral reefs in shallow water, like Australia's Great Barrier Reef.

The coral population is declining, according to marine biologists. A 2011 study found that 19% of coral reefs had already died off, and another 75% are under danger of extinction [1]. Coral reefs have been enormously affected, for better or worse, by the rise in global warming, urbanization, human population, and the heavy reliance on the sea for transportation, resource extraction, recreational activities like boating, and industrial commerce and activity [2]. Corals discolor and die when water temperatures rise too high [3, 4]. The marine biodiversity of our globe has been drastically reduced as a result of this [5]. Consistent monitoring of marine ecosystems is necessary to lessen the detrimental effects of these activities on the water. Optical imaging while submerged comes to the rescue in this situation. Most management systems today include long-term monitoring of wide regions, remote sensing, and tracking of marine animals and their habitats. Therefore, a research priority is the automatic annotation of acquired maritime data, since it is currently a front-runner in management applications [6]. Standard standards can be established to analyze and reduce the negative consequences on seawater environmental sustainability as underwater optical imaging techniques continue to advance. Storage and automatic analysis of such data is also necessary due to the exponential growth of digital camera and video usage.





**Figure 1: (a) Fringing reef (b) The Great Barrier Reef (c) Lighthouse Atoll Reef (d) Coral Bleaching**

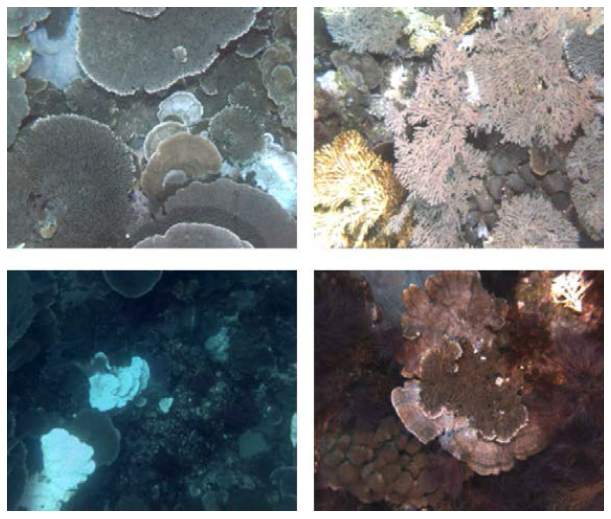
Researchers in the marine field have access to a mountain of unannotated photos of coral reefs. In Australia, coral reefs are captured in millions of photos annually by monitoring systems such as the Integrated Marine Observing System (IMOS). Yet, marine specialists only examine a tiny fraction of these images—usually less than 5%. In addition, manual annotation is a time-consuming and labor-intensive process that requires a lot of human effort. In order to continuously monitor marine ecosystems without input from human specialists, automated monitoring technologies are essential. In light of the above, it is possible to accomplish visually understandable outcomes by automated annotation of underwater photographs. A worldwide hazard may be lessened as a consequence of the suggested outcomes. The level of interest in this area of study is indicative of the research's potential utility. Our goal in writing this article is to investigate the potential uses of deep learning for automated coral reef image annotation and to discuss some of the difficulties inherent in automated marine data processing.

**1.1 Annotation of Coral Reefs: Methods and Challenges**

In terms of biodiversity and monetary value, coral reefs rank among the world's most significant ecosystems. Fringing, barrier, and atoll reefs are the three most common kinds. Reefs that form borders around islands and coastlines are typically located closer to the shore and are named "fringing" (for example, reefs off the coast of Eilat, Israel). For example, the Great Barrier Reef in Australia is a barrier reef that lies offshore from the shore. A body of water known as a lagoon forms as a result, separating the reef from the land. Circular or oval atoll reefs, like Belize's Lighthouse Atoll Reefs, are located deep beneath the surface of the ocean.

In order to learn more about the effects of coral bleaching on reefs over time, long-term monitoring is essential. This research is carried out on both a national and international level. In a coral reef database, you might find:

* A site survey, which may include details such as the exact spot, water temperature and turbidity, and depth
* Species inventory of corals (including varieties such as hard, soft, bleached, and dead corals)
* Analysis of the substrate (non-coral organisms such as sponges, sand, rock, and debris)



**Figure 2: Seabed under different illumination conditions and color distortion**

Many subspecies of corals exist. There are over 600 subspecies in the Great Barrier Reef of Australia. Their size, shape, and color variation reflects their diversity as a species. Hard corals and soft corals are the two most common types of corals.

Soft corals are pliable and sometimes confused for plants since they do not have a skeleton, in contrast to hard corals that are composed of limestone. The presence of hard corals is a good sign that a coral reef is healthy. The majority of coral reef population estimates rely on their percentage cover. Coral bleaching and death are consequences of harmful climatic effects such pollution and rising sea floor temperatures.

**1.2 Challenges**

We now have a fantastic chance to examine the enormous and intricate marine environment thanks to sea floor investigation and imagery. While collecting data from the ocean floor is essential for our knowledge of these complex ecosystems, the practical challenges of operating underwater frequently make this task more difficult than it needs to be. There is a pressing need to advance imaging techniques due to the proliferation of high-tech underwater cameras and the growing fascination with exploring underwater worlds. The majority of seabed observations, archaeological digs, marine geological studies, biological surveys, and biodiversity surveys rely on optical imaging [1, 10 and 11]. Modern oceanography relies heavily on the use of ROVs and AUVs, or autonomous underwater vehicles, to acquire digital photographs of the ocean floor [12]. An effective automated annotation method, as opposed to depending on human labeling, is necessary to transform the marine data into analytically valuable information.

There are unique obstacles that every theoretical system must overcome when operating in a real-world underwater setting. Similarly, with a big enough training set, algorithms developed for object classification in the actual world can handle organized textures and objects, but simply don't cut it when it comes to underwater environments. Blurring, scattering, sun flicker, and color attenuation are just a few of the numerous difficulties that need fixing for improved classification accuracy. Consequently, underwater picture classification via automatic annotation is a complex and demanding subject.Automatic species classification using underwater digital photography is a very difficult undertaking. Shape, color, texture, size, illumination, view angle, camera distance, and light conditions are some of the ways in which underwater classes vary in training datasets. The most common difficulties are:

* The acquired images exhibit a high degree of intra-class and inter-site variety.
* Class boundaries that are both complex and confusing in space.
* Annotation by hand differs amongst specialists.
* Differences in the cameras' field of vision, spatial and spectral resolution limitations, and picture quality.
* Objects of interest being either partly or totally obscured.
* Alterations to the marine seabed's structure that take place gradually over prolonged time periods.
* Lighting artifacts caused by wave refraction and optical characteristics that vary with depth
* Water turbidity that varies, color distortions, and insufficient lighting.

**2 Literature Survey**

**2.1 Automatic Coral Classification using Machine Learning**

***Coral Classification with Hand-crafted Features***

The main ways in which corals can be classified are by their color and texture. Therefore, the extraction of characteristics based on texture and color for the purpose of image representation has been the subject of much research. Due of the lack of clarity around class borders and the arbitrary shapes of corals, features that encode shape information are not well suited. The majority of the time, people like traits that are based on both color and texture. When it comes to generic coral datasets, there are no guaranteed feature combinations that will work. It is common practice to choose features according on the characteristics that distinguish corals from non-corals in a particular dataset. Here we will review a few of the most important papers on the topic of coral image categorization using custom features.

Color was measured by Normalized Chromaticity Coordinate (NCC), and texture was measured by Local Binary Pattern (LBP), according to Marcos [13]. The five types of corals were identified using a three-layer feed-forward back propagation neural network: living, dead, corals with algae, abiotic, and algae. It was postulated that whereas LBP remains stable regardless of changes in brightness, NCC color characteristics remain constant regardless of lighting conditions. On the other hand, complicated underwater photos defy classification using NCC and LBP characteristics. Afterwards, this technique was put to the test using just 300 photos on all three coral classes. When comparing the results of using LBP with hue information to those of using LBP with NCC, the former group demonstrated superior performance.

For color descriptors, Stokes and Deane [14] employed normalized color histograms. For texture, they utilized a feature vector based on the discrete cosine transform (DCT). There were 18 separate classes in their training set of 3000 photographs. The "probability density weighted mean distance (PDWMD)" method is a new suggestion for categorization. You can get this system up and running quickly. The feature extraction process does, however, include manually setting the weights of the texture and color features. And DCT descriptors aren't exactly the gold standard when it comes to texture descriptions.

The feature vector utilized by Pizarro [15] for scale-invariant feature Transform (SIFT) and Hue-histograms was derived from the NCC histogram and bag of words (BoW). A BoW is built using a subset of the training samples, and the test picture is subsequently described using this vocabulary. They voted for the best matches to perform the categorization. There are a total of eight classes, and their technique assigns one to each image. When it came time to teach and acquire new words, 453 images were utilized.When it comes to pixel annotations, this annotation approach isn't very good and often misses important features. This study does not deal with classification at the sub-image level. Underwater environments are complicated, making texture features defined by BoW on SIFT features inefficient.

Using a Maximum Response (MR) filter bank and text on maps for feature extraction at several scales, Beijbom [8] created the Moorea Labelled Coral (MLC) dataset, which includes four non-coral and five coral classifications. To create a texture dictionary, we utilized k-means clustering and a subset of the training photos. Furthermore, they demonstrated that L\*a\*b color space pre-processing yields better results than RGB. Their classification process relied on a Support Vector Machine (SVM) equipped with a Radial Basis Function (RBF) kernel.

Coral images from three different years were automatically annotated to yield coral maps across the reef sites.

For optimal classification accuracy across various benthic datasets, the authors of [16] examined a mix of manually constructed features and numerous classifiers. Their set of descriptors includes the following: opponent angle and hue channel color histograms, Gabor feature, Completed Local Binary Patterns (CLBP), and grey level co-occurrence matrix (GLCM). The feature vectors employed in this study were all resistant to color distortion and poor illumination, and they were also scale invariant. The chosen classifiers were neural networks, support vector machines (SVMs), k-nearest neighbours (KNNs), and probability density weighted mean distance (PDWMD). For optimal performance across all six test datasets, we also tried out various feature-classifier combinations. Unfortunately, this study does not deal with problems like identifying overlapping classes or how to determine the best scale for patch extraction.

***Coral Classification with Learned Features***

A strong class of machine learning algorithms, deep neural networks are built by stacking smaller architectures' neural network layers along their depth and width. Recent years have seen deep networks show off their representation learning and discriminative skills across a variety of applications. Deep learning is being explored by ML researchers for potential uses in a wide variety of new fields. The classification of scenes in maritime environments is one such emerging field. When trained, deep networks necessitate copious amounts of data with annotations. Deep neural networks can sort through millions of tagged photos with the help of effective training techniques.The trained network has additional applications in learning effective picture representations for related benthic datasets. We provide a short overview of deep learning and its state-of-the-art architectures in the next part before we talk about its uses in coral categorization.

**2.2 Automatic Coral Classification using Deep Learning**

The extraction of discriminative features or picture representations from the input data is crucial for any image or video processing task to function well, including classification, object recognition, and scene understating. In computer vision, domain-specific, manually-crafted picture representations have seen heavy use for many years. In recent years, representation learning—a method for learning features using machine learning algorithms—has outperformed more conventional, manually-created representations. Deep learning methods use more complex and extensive versions of traditional neural networks.

Extraction of high-level abstractions from raw data is possible with neural networks that have several hidden layers [17]. A key competency of many cutting-edge computer vision systems is the capacity to derive abstract meaning from complex images. While neural networks (NNs) enjoyed widespread use in the '90s, support vector machines (SVMs) rose to prominence in the '00s thanks to their superior performance. Following the groundbreaking work in [18], deep neural networks gained immense popularity in the field of computer vision.

***Convolutional Neural Networks***

One further significant class of neural networks that can learn picture representations and be used to many computer vision issues is convolutional neural networks (CNNs) [18]. In example, deep convolutional neural networks (CNNs) learn both linear and non-linear operations in parallel across several layers. Parameters of these layers are learned over multiple iterations in order to solve a certain task. Feature extraction from video and image data using CNN based approaches has grown in popularity in the past several years.

Convolutional neural networks (CNNs) use a series of interconnected layers called convolutional and pooling layers. The main reasons why convolutional neural networks (CNNs) are resistant to scaling, shifting, and distortions in the input data are sparse connectivity, parameter sharing, sub-sampling, and local reception fields. Reducing the number of connections between the input and output layers, or "sparse connectivity," is accomplished by making the kernel size lower than the input picture. To make the most of scarcity, you can take use of redundancy inside and between channels. Also, there are fewer operations and less memory needs to store the weights for the output computation. A weight element in a non-convolutional neural network is multiplied by the input just once and then discarded. On the other hand, a convolutional layer convolves the input image with each element of the kernel matrix many times.Stacks of filters with predetermined sizes convolve with the layer's input to form the convolutional layers. In comparison to dense matrix multiplication, convolutional layer parameter sharing is more efficient in terms of computation and memory usage. By sharing parameters, the convolutional layers can become equivariant to linear translations, meaning that changing the input will cause the output to change by the same amount. Convolutional layers, on the other hand, do not maintain their shape when subjected to rotational or scalar distortions.

By convolutionally feeding the pooling layer's output into the following convolutional layer, we can raise the CNN's depth. Improved results have been observed using CNNs that use deeper architectures and have a lower filter size (3x3). Figure 1.4 shows the VGGnet [19] as an example. In comparison to state-of-the-art arrangements, they achieved a notable improvement by increasing the depth to 16–19 hidden layers. They took first and second place, respectively, in the localization and classification tracks of the 2014 ImageNet Challenge.

***Representation Learning***

A promising research field has emerged: learning discriminative image representations from data. An effective visual representation learns the picture characteristics and so captures the data's past distributions. Typically, these characteristics are hierarchical, with lower-level features serving as the foundation and higher-level features serving as the caps. As a result, picture representations learn to define abstract concepts using terms from less abstract ones. The characteristics of a well-learned representation are simplicity (often depending on a linear relationship), sparsity, and temporal and spatial coherence. An additional critical component of representation learning is the network's depth. Deep networks learn representations at deeper layers that encode data properties at a high level.

Convolutional neural network (CNN) image representations have demonstrated state-of-the-art performance in several picture classification tasks [20], after being trained on big datasets like ImageNet [18] and fine-tuned on domain specific datasets. Among the many computer vision tasks that have benefited greatly from these learnt characteristics are instance retrieval, object detection, fine grained recognition, image categorization, and object detection. The majority of studies favor activating the first fully linked layer of convolutional neural networks (CNNs). On the other hand, similar results have been obtained with the activations of intermediate convolutional layers. In a 'local feature' configuration, subarrays of convolutional layer activations are utilized as region descriptors in [21]. In order to create the final feature vector, the local features that were collected from two successive convolutional layers are combined.The method, which is referred to as "cross-convolutional layer pooling," greatly enhanced performance on scene classification tests [21].

Why are these CNN features so effective in so many different contexts? These deep networks' intrinsic nature remains a mystery, despite their exceptional performance. In [22], a method for visualizing the relationship between the input image and the output of the different layers of the convolutional neural network (CNN) architecture was suggested (in [18]). Upon analyzing the outputs of various convolutional layers, the following findings were reached: The second layer handles edge and corner detection, the third handles sophisticated invariances like mesh patterns and texture, the fourth is more class-specific, and the fifth handles object detection in its whole, regardless of posture variations.

The computer vision community is seeing a rise in the use of visualization techniques that clarify learnt deep representations and other computer vision picture representations. Prior to the creation of these techniques, deep feature extraction used CNN-based picture representations as black boxes. In a recent publication, a novel approach to visualization was detailed in [23]. The foundation of this approach are pre-images with significant picture representations that appear natural. These pre-images are created by inversely transforming the learned representation. "Natural pre-images" describes these kinds of pictures. To compare the efficacy of traditional hand-crafted representations with CNN representations, three image visualizations were employed: inversion, activation maximization, and caricaturization. The results showed that representations such as HOG are more accurately invertible than CNN features. Nevertheless, many convolutional neural network (CNN) layers store pertinent input visual data, including variations in stance and lighting. Deep convolutional neural network (CNN) layers maintained both object-specific and global variance information. In addition, completely linked layers were able to pick up on significant differences in the object arrangements. Lines, edges, curves, and components, as well as local variations, appeared to be preserved by intermediate convolutional layers. A significant stride towards comprehending general deep features was made by these findings.

In [24], multi-scale order-less pooling (MOP-CNN) was introduced to further strengthen the invariance of deep features without lowering their discriminative power. Vector of Locally Aggregated Descriptors (VLAD) pooling is used to combine CNN features at different scales locally. Then, by merging these local feature vectors, a general descriptor for supervised or unsupervised identification tasks, picture classification, or scene interpretation is produced—the final feature vector. Compared to global CNN activations, MOP-CNN features routinely outperform them. Additionally, they have done away with the requirement to train prediction layers for a specific domain together.An effective CNN version that has lately gained a lot of popularity for object detection tasks is regions with CNN characteristics, or R-CNN [25].

(1) R-CNN uses deep CNNs in conjunction with region proposals to localize and segment objects, and (2) for smaller training datasets, domain-specific fine-tuning follows supervised pre-training. In the instance of object detection tasks, our approach produced a notable performance boost.

For example, VGGnet requires an input picture size of 224 by 224 in order to be used with a pre-trained traditional deep network. For pictures of any size, the identification accuracy could suffer due to this "artificial" criterion. The authors of [26] proposed a new pooling approach they dubbed "spatial pyramid pooling (SPP)" to get over this CNN restriction. Using SPP-nets, you can get a feature vector of a set length regardless of the size of the input images.

Additionally, SPP strengthened the network's resistance to scale deformations and posture fluctuations. SPP-net utilized full-image representations and accomplished state-of-the-art classification results without fine-tuning. Because feature maps are created from all of the images at once, convolutional features don't have to be computed again.

Several computer vision challenges have been successfully tackled using deep learning methods. However, there is still a lot of space for improvement in the methodologies and these activities are still difficult. When it comes to computer vision, the history of deep learning applications shows that deeper networks perform better [18].

**2.3 Going Deeper with Neural Nets**

In [27] and [28], researchers laid out the reasons why deep learning is superior to other shallow networks in great detail. The difficulty of a deep network-encoded function was addressed in [27] by using the number of different linear sections as a crucial parameter. The capacity to encode fragments of input data was shown to be exponentially increasing for every particular deep network layer. Although the deeper layers computed more complex functions, they still exhibited an inherent stiffness due to the replication of the buried layers. Due to their inherent rigidity, deep networks outperform shallow models when it comes to generalizing unknown input data.

Computing the expressiveness of a model using computational geometry for piecewise linear functions was introduced in a new way in [28]. When comparing networks with the same amount of computing units, deep and narrow rectifier MLPs produced more linear regions than shallow networks.

Training deep neural networks with more depth and width (more weight parameters to change) is computationally expensive. Larger datasets, deeper models, faster hardware, and, most importantly, new techniques for optimizing and efficiently training deeper networks have all contributed to CNNs' phenomenal success over the past five years.

Partitioning and abstraction are the two primary features of traditional convolutional neural networks (CNNs). Using tiny filters at the network's outset and gradually increasing their size as we go deeper improves partitioning. A typical convolutional neural network (CNN) uses a non-linear activation function in conjunction with a linear classifier to generate an abstraction from the input patch. These generalizations lack specificity. A unique structure known as "Network in network (NIN)" was suggested in [29] to make these abstractions stronger.

Multilayer perceptrons and micro neural networks are both initialized at the same time. The rectified linear activation layer is usually followed by an extra 1×1 convolutional layer in this micro network. There are two benefits to using 1x1 convolutions (small filters): first, they increase the network's width and second, they decrease the computational cost by reducing the input vector's dimension. Compared to using a linear classifier followed by a non-linear activation function, this combination improves the approximation of any given function.So the term "network in network" comes from the process of convolving these smaller networks with the input picture inside a bigger network. Repeatedly stacking these structures produces deep NINs. Additionally, global average pooling layers, which are less likely to overfit, replace fully connected layers. The CIFAR-10, CIFAR-100, and SVHN datasets showcased the top-tier performance of deep NINs.

So far, we have proven that the computational difficulty of training a neural network is exactly proportional to its depth. Once we reach a certain depth when adding layers after layers, the accuracy of a deep network becomes saturated. An increase in depth may not necessarily translate to better performance, but it can if the training computations are optimized well. Remainder networks (ResNets) are one method that was described in [34]. Multiple residual blocks, each of which is a miniature CNN, make up a residual network. This collection of leftover blocks isn't merely stacked; each block also features a shortcut connection to the outputs of the blocks below it. Figure 1.5 shows a residual block in action. Decreased network complexity is a result of these shortcut connections. In comparison to a 19-layer VGGnet's 19.6 billion multiply-add operations, a 34-layer ResNet's total number of operations is 3.6 billion. As a result, ResNets do not experience saturation in training accuracy and are easier to train. Afterwards, a study utilizing a 1001-layer deep ResNet produced better results on CIFAR-10.

**3 Hybrid and Ensemble Deep Learning for Coral Classification**

The picture clarity of coral reefs varies from one class to another, and there are complicated boundaries between different classes. How well a classification algorithm performs is directly correlated to how well the features collected from images can discriminate. Given the difficulties mentioned in Section 2, there are a number of restrictions on hand-crafted features. Typically, features that are hand-crafted only encode a single or double attribute, such shape, texture, or color. It is not an easy process to create a new hand-crafted feature representation that takes into account all the difficulties associated with maritime photos. Using pre-trained deep networks on huge image datasets as a source for off-the-shelf CNN features is much more practical. The discriminatory power of CNN features has been demonstrated when applied to a different domain [20]. An intriguing research challenge (as seen below) is how to enhance classification performance by combining CNN features with domain specific hand-crafted features.

***Hybrid and Quantized Features***

In order to classify actions in movies, the concept of integrating CNN with manually created features was applied . To classify actions in movies, the majority of submissions to the THUMOS challenge have used a combination of CNN-based features and features that were hand-crafted. Fisher vectors and Vector of Locally Aggregated Descriptors (VLAD) are commonly used to encode hand-crafted features before they are combined with CNN features.

Using a convolutional neural network (CNN) feature cascaded with manually-crafted morphology and texture data, Wang et al.were able to detect mitosis. A CNN classifier, a hand-crafted feature classifier, and a third classifier for test samples misclassified by the first two have all been trained. For applications involving massive datasets, this method is computationally too expensive. The combination of convolutional neural networks (CNNs) and manually created features yields promising outcomes for RGB-D object detection, as demonstrated by Jin et al.. They integrated the CNN features with those derived from spatial pyramid matching using Locality-constrained Linear Coding (LLC).

Since benthic datasets have pixel annotations rather than bounding box annotations, CNN features cannot be utilized directly for coral image classification. For the coral reef categorization issue, deep features have only lately been investigated. To classify coral reefs, an approach was suggested in that makes use of both VGGnet-extracted generic deep features and custom-crafted features, capitalizing on the complimentary strengths of the two representation types. There wasn't enough data to train a convolutional neural network (CNN) using random initializations.Thus, in order to make the picture representations scale-invariant, features based on convolutional neural networks (CNNs) that had already been trained were extracted from patches that were centered at labelled pixels at various scales. A local variation of support vector propagation (SPP) was then applied. Additionally, features based on color and texture were hand-crafted using the same patches and utilized to augment the CNN characteristics. In order to cut memory needs in half, researchers looked at a 2-bit feature representation technique that was efficient with memory.

On the MLC benchmark dataset for corals, the suggested method outperformed the state-of-the-art methods in terms of classification accuracy.The most effective features were those that combined learned and hand-crafted elements.Additionally, it is inferred that the class imbalance problem was more effectively tackled by the CNN features and the hybrid features. When coral patches are extracted at a single scale, the most numerous class takes center stage, while the less common classes are overshadowed. The feature vectors that emerge from max-pooling patches extracted at different scales provide greater weight to the less abundant classes. The outcomes of the experiments proved this. Because of this, the classifier is better able to deal with the problem of class imbalance.

**Coral Dataset**

Bleaching corals is throwing aquatic ecosystems out of whack, since they are essential to many aquatic organisms. Thus, aquatic life is spared from calamity when bleached corals are detected early. You can see both healthy and bleached corals in this collection of pictures. There is a 227 × 227 pixel input image.



(b) Bleached Coral

1. Healthy Coral

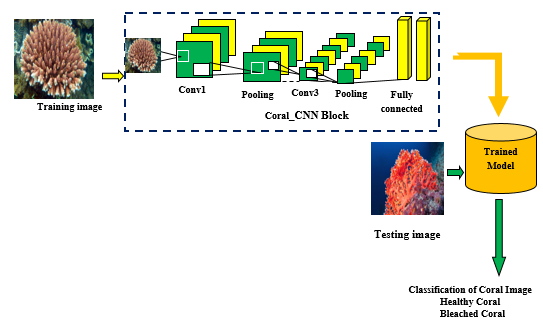
**Figure 3: Healthy Coral and Bleached Coral**

**Preprocessing**

Sharpening: The pictures aren't all that great, and a lot of them are fuzzy or out of focus. To get around this and make a better dataset (C), we boosted the contrast of the whole photos and brought attention to the details. In order to do this, the weighted average of each pixel's 3 × 3 neighborhood was used to replace its original value. Here is the filter displayed in the matrix:

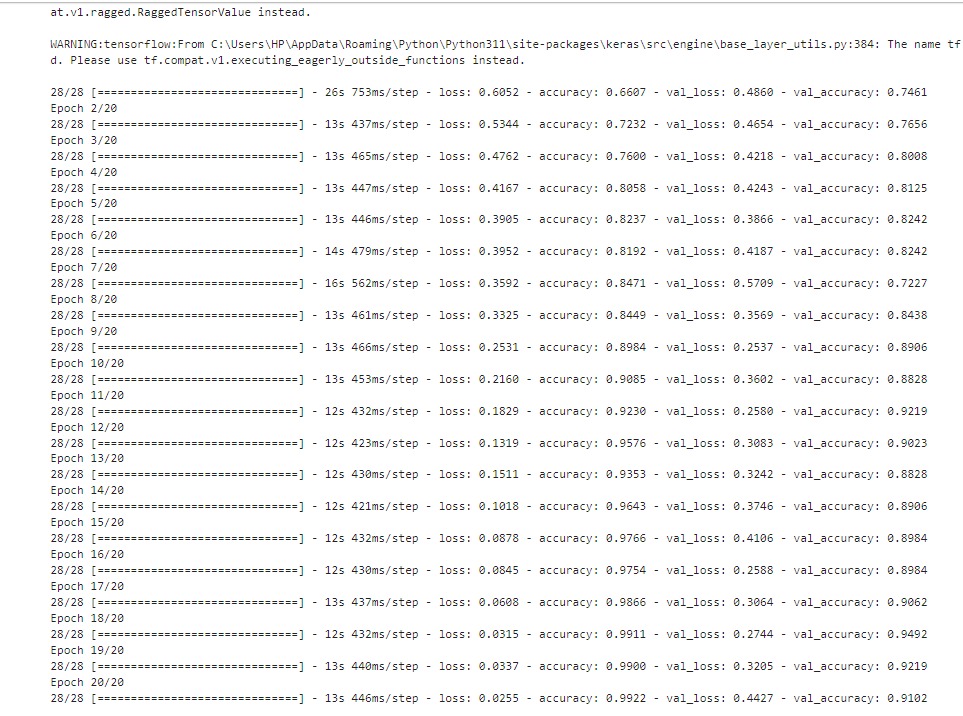


**Proposed Coral CNN Model**



**Figure 4: Architecture for Proposed Model**

Since deep learning algorithms, such as convolutional neural networks (CNNs), are structurally comparable to the human brain, they find widespread application in computer vision. Conventional neural networks (CNNs) have many applications, including image recognition and classification. Ranks images using a convolutional neural network (CNN) technique. A convolutional neural network (CNN) model typically has four layers: an input layer, multiple hidden layers, and an output layer (as seen in the previous figure). Layers for feature extraction (convolutional), feature reduction (pooling), flattening (fully linked), and prediction (soft-max).



**Figure 5: Summary for the Proposed Model**

**4 Result and Discussion**

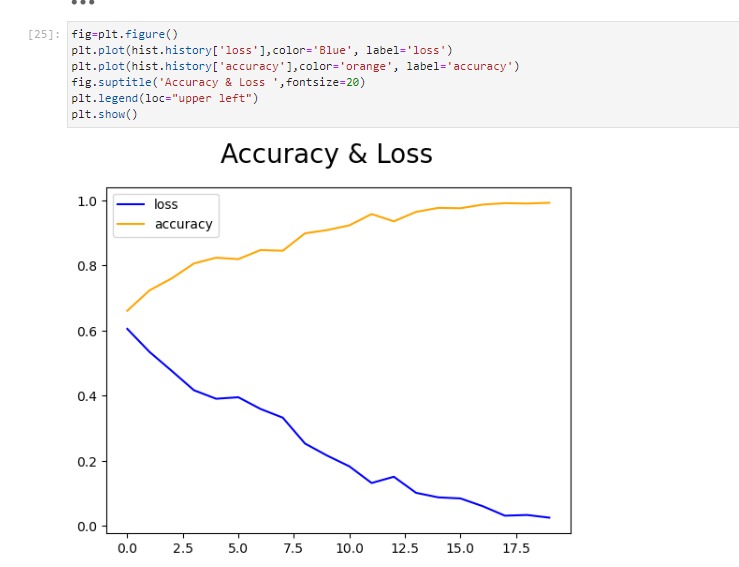
Performance is determined by using f1-score, recall, accuracy, and precision.

Prediction Accuracy: The correctly predicted diseases are shown as

The number of accurate positive predictions, or "precision," is a measure of the accuracy of the model.

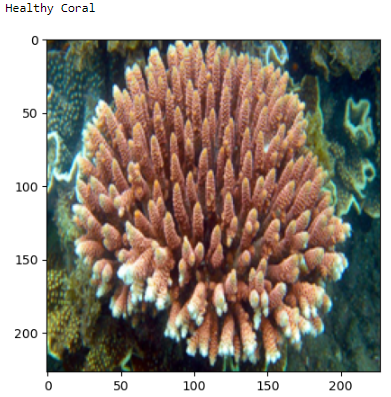
A classifier's recall is the percentage of true positive predictions made by the model relative to the total number of true positives in the data.

The F1-score takes precision and recall into account.

****

**Figure 6: Accuracy and Loss for Proposed Model**

In the graphic above, the blue line represents the suggested model's loss and the yellow line represents its accuracy. In this case, the y-axis can take values between zero and one. The numbers 0–20 make up the x-axis.

****

**Figure 7: Classified Healthy Coral**

Figure 7 displays the input image's categorized coral, whereas Figure 6 displays the proposed model's accuracy and loss.

**5 Future Prospects**

An objective metric for detecting, discriminating, and identifying species, as well as their behavior and morphology, can be provided by deep learning solutions applied to ecological investigations.

Some other future prospects of this research are:

* Create deep learning algorithms that can automatically categorize vast amounts of marine data, not just corals.
* To evaluate several deep learning approaches side by side in order to establish a firm foundation for effective evaluation of marine ecosystems.
* To eliminate the need for human labelers by creating an automated annotation system that can handle various datasets.
* Using financially viable monitoring programs, we aim to learn how marine ecosystems withstand human-caused changes like climate change, pollution, resource exploitation, and coastal expansion.
* In order to: Determine patterns in population dynamics and examine interspecies connections in the ocean.

**5 Conclusion**

In this chapter of the book, we introduced a simple model called Coral CNN that makes use of the latest architectures in deep neural networks and the evolution of deep learning. We covered the basics of ocean bottom investigation and the difficulties of gathering and analyzing marine data. Our next presentation was a synopsis of the research on methods for marine picture classification. By reviewing the most current research from our lab and others, we dug deeper into the possible uses of deep learning for benthic picture classification. In addition, we went over several potential future research areas in deep learning and underwater scene interpretation, and we presented a Coral CNN Model. Through writing this chapter, we hope to inspire other computer vision and marine society researchers to work together on similar, long-term projects.

.

References

1. P. Fairweather, "Statistical power & design requirements for environmental monitoring," Australian Journal of Marine and Freshwater Research, vol. 42, pp. 555-567, 1991.
2. A. Underwood, "Experiments in ecology," in Cambridge University Press, Cambridge, 1997.
3. E. Graham, M. Dethier, S. Cohen and L. Tear, "Visual versus random-point percent cover estimations: objective is not always better.," Marine Ecology Progress Series, vol. 96, p. 93–100, 1993.
4. A. Wojcicka and Z. Wrobel , "The panoramic visualization of metallic materials in macro- and microstructure of surface analysis using Microsoft Image Composite Editor (ICE)," pp. 358-368, 2012.
5. C. N. Rooper and M. Zimmermann, "A bottom-up methodology for integrating underwater video and acoustic mapping for seafloor substrate classification," Continental Shelf Research, vol. 27 (7), pp. 947–957, Apr 2007.
6. R. Hummel, "Image enhancement by histogram transformtion , computer graphics and image processing," vol. 6, pp. 184-195, 1997.
7. C. E. Woodcock, "The factor of scale in remote sensing," Remote Sensing of Environment, vol. 21, no. 3, pp. 311-332, April 1987.
8. "What is spectral signature in remote sensing?," 2013. https://articles.extension.org/pages/40071/what-is-a-spectral-signature-in-remote-sensing (last accessed: Nov 3, 2017)
9. P. J. Mumby, C. D. Clark, E. P. Green and A. J. Edwards, "Benefits of water column correction and contextual editing for mapping coral reefs," international Journal of Remote sensing, vol. 19 (1), 1998.
10. H. Lu, Y. Li, Y. Zhang, M. Chen, S. Serikawa and H. Kim, "Underwater optical image processing: a comprehensive review," Proc. CVPR, 2017.
11. R. Coles, L. McKenzie, M. Rashed, J. Mellors, H. Taylor, K. Dew, S. McKenna, T. Sankey, A. Carter and A. Grech, "Status and trend of seagrass habitats in the Great Barrier Reef world heritage area," Report to the Marine and Tropical Sciences Research Facility, Reef and Rainforest Research Centre on behalf of the Australian Government’s Marine and Tropical Sciences Research Facility, Cairns, Australia, 2007.
12. F. T. Shorta, B. Polidoro, S. R. Livingstone, K. E. Carpenter, S. Bandeira, J. S. Bujang, H. P. Calumpong and T. J. Carruthers, "Extinction risk assessment of the world's seagrass species," Biological Conservation, vol. 144, pp. 1961-1971, 2011.
13. D. I. Walker and R. I. T. Prince, "Distribution and biography of seagrass species on the northwest coast of Australia," Aquatic Botany, vol. 29(1), pp.19-32, 1987.
14. M. Vanderklift, D. Bearham, M. Haywood, H. Lozano-Montes, R. McCallum, J. McLaughlin, K. McMahon, N. Mortimer and P. Lavery, "Natural dynamics: understanding natural dynamics of seagrasses in north-western Australia," WAMSI Dredging Science Node Report Perth, Oct 2016.
15. M. I. SAUNDERS et al., "Coastal retreat and improved water quality mitigate losses of seagrass from sea level rise," global change Biology, vol. 19, p. 2569–2583, 2013.
16. R. Pu and S. Bell, "A protocol for improving mapping and assessing of seagrass abundance along the West Central Coast of Florida using Landsat TM and EO-1 ALI/Hyperion image," ISPRS Journal of Photogrammetry and Remote Sensing, pp. 116-129, Jul 2013.
17. C.-F. Chen, V.-K. Lau, N.-B. Chang, N.-T. Son, P.-H.-S. Tong and S.-H. Chiang, "Multi-temporal change detection of seagrass beds using integrated Landsat TM/ETM+/OLI imageries in Cam Ranh Bay, Vietnam," Ecological Informatics, pp. 43-54, Jul 2016.
18. Marouf M, Abu-Naser SS. Fine-tuning MobileNetV2 for Sea Animal Classification.
19. Sharan S, Kininmonth S, Mehta UV. Automated cnn based coral reef classification using image augmentation and deep learning. International Journal of Engineering Intelligent Systems. 2021 Jul 1;29(4):253-61.
20. Firdous RJ, Sabena S. Collaborative CNN with Multiple Tuning for Automated Coral Reef Classification. InInternational Conference on Computer, Communication, and Signal Processing 2023 Jan 4 (pp. 81-93). Cham: Springer Nature Switzerland.
21. King A, M Bhandarkar S, Hopkinson BM. Deep learning for semantic segmentation of coral reef images using multi-view information. InProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops 2019 (pp. 1-10).
22. Wulff J. Assessing and monitoring coral reef sponges: why and how?. Bulletin of Marine Science. 2001 Sep 1;69(2):831-46.
23. Arif Mahmood, [Mohammed Bennamoun](https://research-repository.uwa.edu.au/en/persons/mohammed-bennamoun), Senjian An, Ferdous Sohel, [Farid Boussaid](https://research-repository.uwa.edu.au/en/persons/farid-boussaid), [Renae Hovey](https://research-repository.uwa.edu.au/en/persons/renae-hovey), [Gary Kendrick](https://research-repository.uwa.edu.au/en/persons/gary-kendrick), Robert Fisher.Coral classification with hybrid feature representations.
24. Forest Rohwer, Merry Youle, and Derek Vosten, Coral reefs in the microbial seas, vol. 1, Plaid Press United States, 2010.
25. Oscar Pizarro, Paul Rigby, Matthew Johnson-Roberson, Stefan B Williams, and Jamie Colquhoun, “Towards image-based marine habitat classification,” in OCEANS 2008. IEEE, 2008, pp. 1–7
26. Ma Shiela Angeli Marcos, Maricor Soriano, and Caesar Saloma, “Classification of coral reef images from underwater video using neural networks,” Optics express, vol. 13, no. 22, pp. 8766–8771, 2005.
27. Ross Girshick, Jeff Donahue, Trevor Darrell, and Jagannath Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on. IEEE, 2014, pp. 580–587.
28. David G Lowe, “Object recognition from local scale-invariant features,” in Computer vision, 1999. The proceedings of the seventh IEEE international conference on. Ieee, 1999, vol. 2, pp. 1150–1157
29. Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell, “Decaf: A deep convolutional activation feature for generic visual recognition,” arXiv preprint arXiv:1310.1531, 2013.
30. MR Patterson and NJ Relles, “Autonomous underwater vehicles resurvey bonaire: a new tool for coral reef management,” in Proceedings of the 11th International Coral Reef Symposium, 2008, pp. 539–543.
31. M Dale Stokes and Grant B Deane, “Automated processing of coral reef benthic images,” Limnology and Oceanography: Methods, vol. 7, no. 2, pp. 157–168, 2009.
32. B Boser Le Cun, John S Denker, D Henderson, Richard E Howard, W Hubbard, and Lawrence D Jackel, “Handwritten digit recognition with a back-propagation network,” in Advances in neural information processing systems. Citeseer, 1990.
33. Oscar Beijbom, Peter J Edmunds, David Kline, B Greg Mitchell, David Kriegman, et al., “Automated annotation of coral reef survey images,” in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012, pp. 1170–1177.
34. Ali S Razavian, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson, “Cnn features off-the-shelf: an astounding baseline for recognition,” in Computer Vision and Pattern Recognition Workshops (CVPRW), 2014 IEEE Conference on. IEEE, 2014, pp. 512–519.