IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 29 August 2023, accepted 15 November 2023, date of publication 8 December 2023,
date of current version 25 January 2024.

Digital Object Identifier 10.1109/ACCESS.2023.3341156

RESEARCH ARTICLE

ReefCoreSeg: A Clustering-Based Framework for Multi-Source Data Fusion for Segmentation of Reef Drill Cores

RATNEEL DEO[1,2,3], JODY M. WEBSTER[1,3], TRISTAN SALLES[1,3], AND ROHITASH CHANDRA[2,3,4], (Senior Member, IEEE)

[1]Geocoastal Research Group, School of Geosciences, The University of Sydney, Sydney, NSW 2050, Australia
[2]Transitional Artificial Intelligence Research Group, School of Mathematics and Statistics, UNSW Sydney, Sydney, NSW 2052, Australia
[3]ARC-ITTC Data Analytics for Resources and Environments, Sydney, NSW 2751, Australia
[4]UNSW Data Science Hub, UNSW Sydney, Sydney, NSW 2052, Australia

Corresponding author: Ratneel Deo (ratneel.deo@sydney.edu.au)

This work was supported in part by the Australian Research Council under Grant DP1094001, in part by the Australian and New Zealand International Ocean Discovery Program (ANZIC), and in part by the Australian Research Council under Grant IC190100031.

ABSTRACT Coral reefs are among the most biologically diverse and economically valuable ecosystems on Earth, but they are threatened by climate change. Understanding how reefs developed over geological timescales can provide important information about past environmental changes and their impacts on reef systems. Significant effort and capital have been invested in drilling and analyzing reef cores. Recognizing coral and sediment patterns visually from fossil reefs is a laborious task that demands domain expertise. In this paper, we present a machine learning-based framework that utilizes clustering and classification methods to fuse multiple sources of data for the segmentation and annotation of reef cores. The framework produces an annotated image of a reef core with six lithologies identified; massive corals, encrusted corals, coralline algae, microbialite, sand, and silt. We utilize reef cores recovered from Expedition 325 of the International Ocean Discovery Program (IODP) to the Great Barrier Reef. We use *reef core image data* and *physical properties data* to segment reef cores. We evaluate the framework using selected clustering and classification models. The results show that *Gaussian mixture models* can provide accurate segmentation of reef core image data, with a clear visual distinction between two major classes: massive corals and stromatolitic microbialites. Furthermore, we find that the *random forest classifier* provides the best annotations for the segmented reef core image data with an accuracy of 96%.

INDEX TERMS Clustering, segmentation, multi-source data, classification, reef core analysis, Gaussian mixture models.

I. INTRODUCTION

Coral reefs are biosystems and significantly contribute to the social and economic sustenance of coastal communities [1], [2], [3]. Coastal regions and reefs are the source of many livelihoods and support a diverse range of marine life [4], [5], [6]. The ecology and environment associated with past reef development are inferred from the fossil records, with the prevalence or absence of specific reef lithologies and coral assemblages corresponding to particular environmental conditions [7]. Studies in reef development are limited to the amount of surface data and drill cores available [8]. Drill cores from these reefs provide great insights into the dynamics of coral growth. Marshall and Davies [9] used six drill cores from One Tree Island in the Southern Great Barrier Reef (GBR) to model vertical accretion of the reef post-Holocene. Sanborn et al. [10] used 12 cores drilled along three transects from One Tree Reef across various geomorphic and hydrodynamic environments to document three distinct stages of Holocene reef development. The authors conducted

The associate editor coordinating the review of this manuscript and approving it for publication was Julien Le Kernec.

a manual classification of paleoecological assemblages that included coral, coralline algae, and associated biota. They also used extensive radiometric dating to identify three distinct phases of reef development. Advances in reef drilling and underwater imaging have seen unprecedented growth in data availability and quality, but given limited resources for manual analyses, novel and automated approaches are required [8], [11], [12], [13], [14].

Computer vision and machine learning methods have been prominent for studying environmental change and restoration [15], [16], and effective for coastal reef studies via remote sensing [17]. Satellite image-based classification models have shown good performance in estimating coral cover and mapping of various coral reef systems around the world [18]. Deep learning methods such as convolutional neural networks (CNNs) reported good performance in the classification of algae, hard corals, and soft corals [19]. Gonzalez-Rivero et al. [19] used CNNs for automated point annotation of benthic images from the *XL Catlin Seaview Survey*.[1] Lyons et al. [20] developed a framework that mapped coral reef habitats at the individual reefs to vast ocean reef levels. They produced 10 maps with up to 78% accuracy in classifying reef habitats in the GBR and other reefs in the South West Pacific. Kennedy et al. [21] enhanced the framework and developed a reef classification tool called *Reef Cover* to map coral reef habitat by fusing remote sensing data (surface reflectance, bathymetry, and wave modeling) with field data (benthic imagery, surface morphology, and geological data). Their model has been successfully used for supporting management and conservation efforts in the Cairns Management region and the Mariana Islands. Mogstad et al. [22] used the support vector machine for classifying coralline algae and other invertebrates from underwater hyperspectral images. Hence, there is potential for machine learning models to study and automate reef drill-core data processing.

Lithology and lithofacies identification is an important part of analyzing fossil reef cores. This has traditionally been done by analyzing cores and drilling cuttings, as well as interpreting well-log data through visualizations [10], [23]. There have been successful attempts to automate lithology and lithofacies prediction using computer vision and supervised machine learning models. Alzubaidi et al. [24] used residual networks for predicting four classes of rocks in drill cores. Galdamesh et al. [25] used region-based CNNs, for the segmentation of rock instances from 230 images. Thomas et al. [26] used a nearest neighbor classifier to identify and predict sand, shale, and carbonate cement from drill core photographs. Dawson et al. [27] used CNNs with transfer learning to classify carbonate rocks from drill core images on varying dataset sizes. They explored the limits of machine learning models on classifying datasets smaller than 100,000 points. Baraboshkin et al. [28] compared five deep learning models for classification of rock lithologies using 20,000 images of drill cores from oil and gas fields, where the GoogLeNet model reported impressive performance accuracy. Insua et al. [29] created a labelled six-class dataset of all the reef cores obtained from International Ocean Discovery Program (IODP) Expedition 325 and developed multiclass classification of lithologies using *physical properties data*. In summary, the processing of drill-core data works well with supervised machine learning models for classification; however, generating the ground truth labels for these drill cores is a highly laborious and time-consuming process. Hence, there is a need to develop unsupervised and hybrid machine learning frameworks that can facilitate efficient drill core analysis and classification.

Clustering is an unsupervised machine learning method used for grouping samples in a given dataset based on a similarity index [30]. We can model the distribution of multi-class data with methods such as k-means clustering [31] making them suited for unlabeled data [32]. Clustering has been used for image segmentation that partitions an image into spatially adjoining and homogeneous regions (segments) [33], [34], [35]. Segmentation methods rely on three key criteria When clustering pixels into groups, which include the homogeneity within each segment, differentiation from neighboring segments, and shape consistency [36]. There are some key applications of segmentation based on clustering in the area of coral reef analyses. Song et al. [37] showed good performance on semantic segmentation of single-channel images of modern corals. Chiriyath and Instrella [38] devised a robust technique for segmenting coral reefs utilizing airborne fluid lensing information by implementing a naive Bayes maximum a posteriori estimation scheme. Wang et al. [39] developed a color gradient method to segment the pores in the coral reef borehole images. Steinberg et al. [40] used the Bayesian non-parametric Dirichlet process for clustering large quantities of seafloor imagery in the O'Hara Marine Protected Area in Tasmania, Australia.

There is an abundance of reef core data available from the GBR and reefs globally for automating reef core analysis. However, we still lack a specific framework that incorporates unsupervised machine learning for the analysis of reef core data. In this paper, we present a novel framework that utilizes different clustering methods to fuse multiple sources (formats) of data for the segmentation and annotation of reef drill cores. The framework produces a classification scheme for labelling the reef drill core, with segments annotated based on the type of lithologies present in the drill core. We utilize reef cores selected from the Hydrographers Passage in the Central GBR. In our analysis, we combine image data and physical properties data such as bulk density (gamma-ray attenuation), porosity, and electrical resistivity. Then, we evaluate the potential of three key clustering methods including k-means clustering [31], agglomerative hierarchical clustering [41] and Gaussian mixture models [42] on segmenting the data. The primary source of data is based on the image data that embeds

[1] https://www.catlinseaviewsurvey.com/

TABLE 1. Physical properties data description showing the range of values for each of physical properties. There are 3852 measurements taken for each property across all the cores from Exp 325.

	Bulk Density g/cm^3	Porosity $\%vol$	Resistivity $Ohmm$
mean	1.994231	0.436058	2.470501
std	0.229167	0.133041	3.510888
min	1.017700	0.123800	0.331800
max	2.536500	1.004800	28.860000

information from the color and texture. However, to further distinguish the classes, we also use data from physical properties measurements taken by a multi-sensor core logger on the same core. We follow this by evaluating four selected classification methods for annotating the segmented image that includes support vector machines, multilayer perceptron, random forests, and k-nearest neighbors.

We structure the remainder of the paper in the following manner. Section II provides an overview of the data, and Section II-C presents our clustering framework. Section III presents the results, Section IV presents the discussion and finally, Section V provides the conclusions.

II. METHODOLOGY

A. STUDY AREA

We utilize reef drill-cores from the IODP Exp 325.[2] The coring (drilling) was done on board the research vessel, the Greatship Maya. The vessel was equipped with a dynamic positioning system which was used to drill cores offshore at the reef shelf edge. The drill site is located east of Townsville, as shown in Figure 1. The team drilled 34 boreholes across 17 sites, varying in depth from 42.27 to 211.70 meters below sea level. The drill holes ranged from 1.40 to 46 meters in length, starting at the sea floor. Most of the cores had a recovery rate over 26%, however, there were core recovery rates between 1.4 to 40%.

B. DATASET DESCRIPTION

We use two types of data from Expedition 325, core physical properties data, and core image data. In our study, we have only selected physical properties data that were measured using a *multi-sensor core logger* (MSCL). These data include measurements of bulk density, porosity, and resistivity. The MSCL used an array of sensors to measure the physical properties of individual drill core sections onboard the drilling vessel. Table 1, shows the different values for each property, and it was worth noting that a total of 3852 measurements were taken across all the cores. Insua et al. [29] used this data for all the cores from the expedition and did a visual analysis of the core samples into six different lithologies shown in Figure 3a, which created a labelled six-class classification dataset. Figure 3b shows the distribution of the six lithologies in the dataset.

[2]website: http://publications.iodp.org/proceedings/325/325title.htm

It is evident that there is a high imbalance in the data, as the distribution of data across different classes is not equal [43]. In other words, the amount of sand present in the dataset overwhelms all the other classes. The number of samples of corals, microbialites, and algae is much smaller and this is problematic for machine learning models since are biased towards the majority class (which in our case is sand). Consequently, these models can also ignore the minority classes as there is fewer data samples to learn which presents the challenge as the minority classes (coral, microbialite, and aglae) are of more significance in our dataset since we are interested in classifying all the components of the drill cores instead of just labelling sand.

Core 33-A-16R from Exp 325 was used for this study which comes from hole 33-A of transect HYD-01C given in Figure 1. Core 33-A was drilled 32.9 meters below the sea floor with a radius of 0.066 meters. The drilling process resulted in a collection of 13.41 meters of cores, which were obtained in 23 runs. Figure 2 shows the core 33-A-16R which means that this core section comes from the 16th drill run of hole 33-A. This section of the core was selected for our study since Insua et al. [29] studied the classification of six carbonates lithologies in this core using supervised machine learning. Additionally, this core had one of the highest rates of coral recovery at 40.9%, thus making it an ideal candidate for the annotation of multiple carbonate types. Figure 2 also shows five out of the six lithologies present in the dataset.

C. FRAMEWORK

We present the ReefCoreSeg framework that fuses clustering and classification techniques on different formats of data to automatically annotate carbonates present in the cores as shown in Figure 4. The framework is divided into three independent modules; 1.) Clustering (segmentation of the cores is achieved using this module), 2.) Classification, and 3.) Annotation.

In the framework, we first execute the clustering module using the image data and the physical properties data. We restrict the image data to drill core 33-A-16R as a benchmark for the study. Firstly, we take the image and downscale its resolution to ensure efficient clustering. We reduced the dimensions by a factor of 0.95 as this worked best with the clustering modules in our initial experimental runs. This was selected over the flattening of the image colors into grayscale as the colors are a major distinguishing feature in the corals. In the clustering module labelled as 1 (Figure 4), we also add the physical properties data for core 33-A-16R and create a stacking of the two formats of data. Next, we present the stacked image and physical properties data to the three clustering methods that are discussed in detail in the next section. We enable each of the models to segment the data and produce a cluster label for each pixel of the image data. In order to identify the most optimal clustering method, we employ a comprehensive evaluation process that takes into account several factors. Firstly, we use the Silhouette score, which measures the distance between

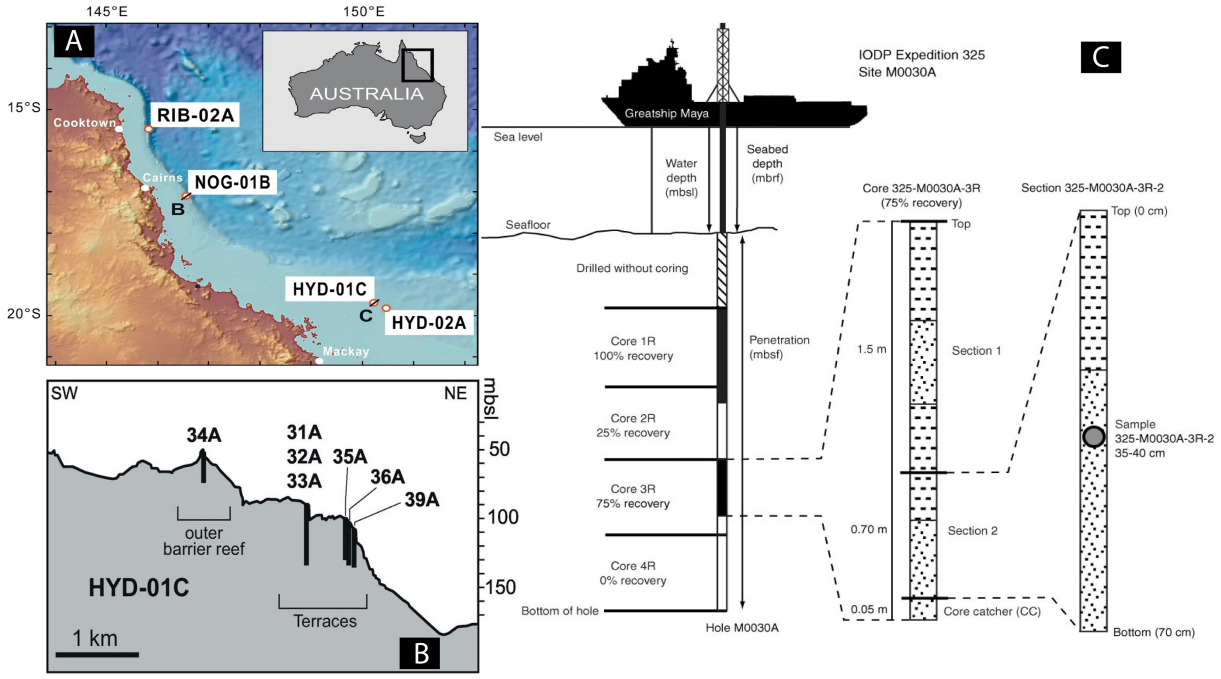


FIGURE 1. A. IODP drilling sites for Exp 325 located along the NE Australian shelf. B. Shelf edge bathymetric profile and location of drilling holes at the Hydrographer's Passage transect HYD-01C of which we use Core 33-A. C. Drill rig used to extract the reef cores offshore from onboard the ship and the IODP core classification scheme.

clusters and determines the quality of the clusters based on how well each observation fits within its assigned cluster. We also utilize the Calinski Harabasz score, which takes into account the size and dispersion of the clusters. Additionally, we take a human visual analysis approach, where experts in the field manually inspect and evaluate the clusters to ensure that they make sense from a business perspective.

Furthermore, we also perform a control experiment, where we evaluate the clustering performance of models that solely use image data. This allows us to compare and contrast the results of the control group with those of the other groups, which in turn enables us to determine the effectiveness of the clustering methods that we are evaluating. By taking these multiple factors into consideration, we can confidently select the best clustering method that will yield the most accurate and reliable results.

Figure 4 shows the classification module, where we only use the physical properties dataset. We apply standard data pre-processing which includes min-max normalization of the features, and one hot encoding of the carbonate class labels. Our dataset was found to be highly imbalanced, and initial experiments showed that oversampling did not improve the classification of minority classes. To address this, we used the neighborhood cleaning rule to implement undersampling and obtained a balanced dataset. We chose to undersample the data since the majority class was sand. We give less importance to the classification of sand since we are more interested in improving the prediction of the minority classes, such as encrusted corals. Initial experiments showed better performance via undersampling. In order to finish the pre-processing, we separated the data into distinct training and testing sets. We assigned 33% of the data for testing and the remaining portion for training. In this study, we did not use a validation set, since we are comparing the classification performance to Insua et al. [29] and we are not implementing early stopping of training via the validation set. Then, as shown in Figure 4, we pass the balanced data to four machine learning models for classification (classifiers), namely the support vector machines, random forests, simple neural networks (multilayer perceptron), and k-nearest neighbors. We note that support vector machines and random forests have been used in a previous study [29] and other models and have also been used in related literature [27], [44]. We provided each of the classifiers with the same training and testing data to ensure fair comparison and tuned their hyperparameters empirically.

The *classification* module in the ReefCoreSeg framework (Figure 4) accounts for the missing labels in the drill-core data. In order to obtain data on the physical properties of larger rock samples within the cores from Exp 325, we relied on the results of a previous study. In this study, Insua et al. [29] meticulously analyzed and labeled each piece of material present in the drill core section. They then mapped this information onto the physical properties data obtained from the MSCL. The resulting dataset is shown in Figure 3b.. This dataset was sufficient for classification; however, it was incomplete for annotating the core section in our study (core 33-A-16R). There were 75 observations

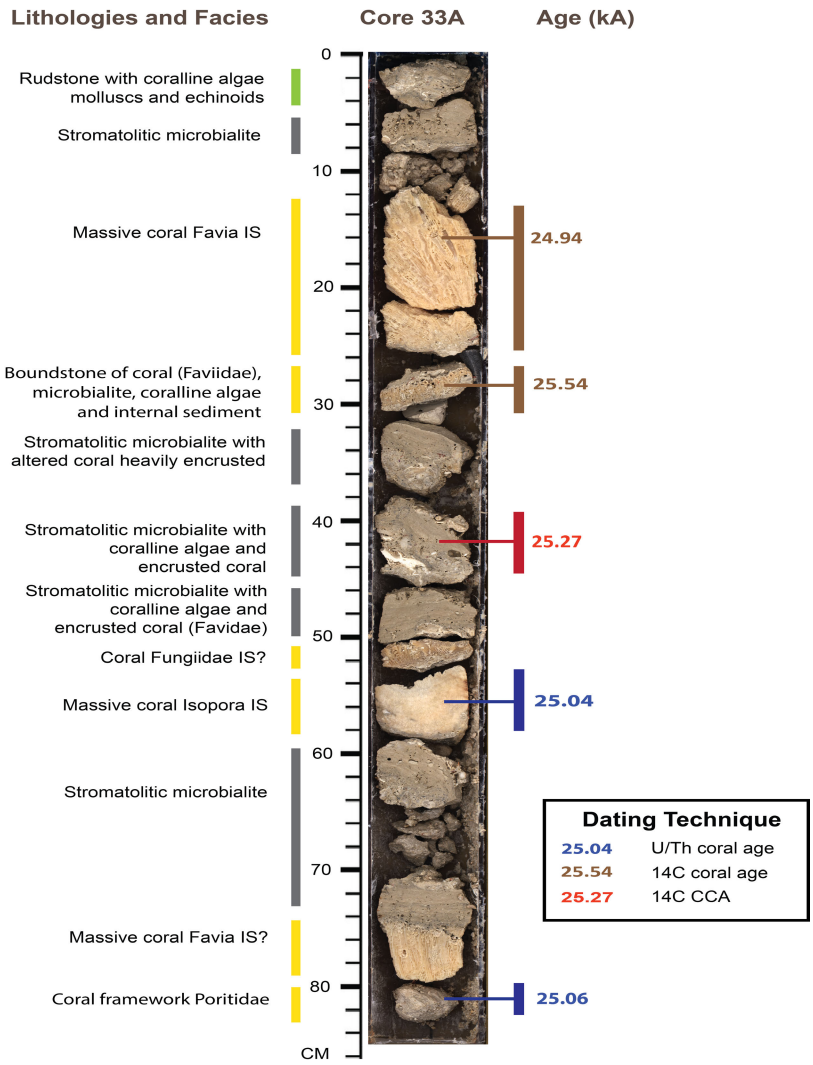


FIGURE 2. Core 33-A-16R used in this study with all the lithologies and facies identified on the left. The legend on the right gives the age of the core sections in the last thousand years (ka) as well as the dating technique used to estimate the age. Uranium/Thorium (U/Th) and radiocarbon (14C-AMS) were used in dating the core. The figure is adapted from [29].

of physical properties taken from core 33-A-16R at 10-centimeter intervals using the MSCL; however, only 6 of these readings had a mapping to the class of carbonate present at that interval. This led to having labels for only 8 percent of the core section in our study (core 33-A-16R). We used the classification methods in the ReefCoreSeg framework to predict the labels for the remainder of the data (90 percent), which we then used to annotate the clustered image.

Finally, we execute the annotation of clustered images as the final phase of the ReefCoreSeg framework as shown by the *automatic annotation* module of Figure 4. We use the clustered image from the clustering module as the base for annotation. We then extract the physical properties data for stacking in the clustering module and then use the classification module to get the predicted labels. Furthermore, we stack the clustered image alongside the classifications to get the annotations of the reef drill-core data. We finally validate the quality of the annotation qualitatively with consultation from a domain reef scientist (Jody Webster from the University of Sydney) and present the final annotated core image.

D. CLUSTERING MODULE

We provide details about the three clustering methods in the ReefCoreSeg framework. As noted earlier, clustering methods divide data into groups based on the distance or similarity between the observations in the dataset. The distance metric is used to capture the topology of the data sample for identifying clusters of specific shapes [31].

1) K-MEANS CLUSTERING

K-means clustering is a highly effective algorithm because it can handle a vast number of features [31] and has been applied to several domains that include image

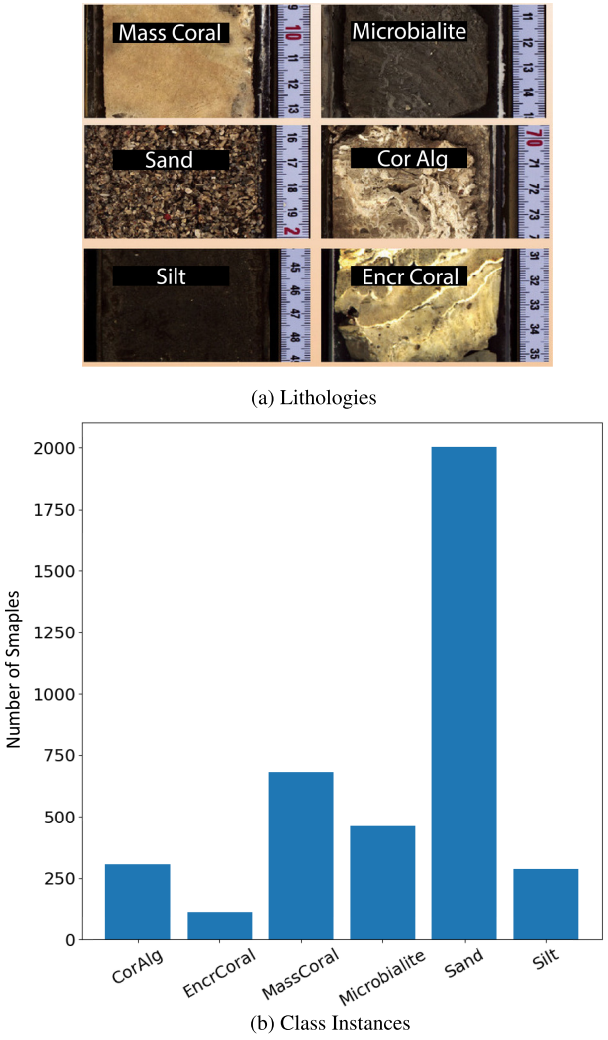


FIGURE 3. The classification dataset generated by visual inspection of the individual core samples, adapted from Insua et al. [29]. (a) Images of the six lithologies found in the drill cores. The lithologies can be distinguished by color and texture. (b) The number of instances of each lithology present in the dataset with high imbalance.

segmentation [45]. K-means clustering uses k clusters (centroids) and assigns every data point to the nearest centroid. The centroid shifts towards the mean of all the assigned data points, and this cycle continues until a specific level of error is achieved. One significant constraint of the algorithm is that it needs the number of centroids, i.e. k to be predetermined. Furthermore, k-means clustering is highly reliant on the initial placement of centroids in the data space. To address this, multiple random assignments of centroid location can be utilized and the clusters obtained can be averaged out [46]. The number of clusters needs to be informed by expert advice or determined empirically. In the study, we use the *evidence lower bound* (ELBO) [47] method for determining the number of clusters or centroids in the given data. We use the distortion measured via the average of the squared Euclidean distance as the primary comparison metric for the clusters. Figure 5 displays the ELBO results for the target drill-core data. It is evident from the plot that the optimal number of clusters is 4, represented by $k = 4$, i.e. "elbow" of the plot that refers to the region after which the distortion starts decreasing linearly.

2) AGGLOMERATIVE HIERARCHICAL CLUSTERING

Agglomerative hierarchical clustering (AHC) [41] builds a cluster hierarchy by recursively merging pairs of clusters of sample data using linkage distance. AHC is known as a "bottom-up" approach, where each observation begins with its own cluster. As one moves up the hierarchy (level) of AHC, the clusters are merged in pairs. Hence, it is important to identify which elements should be merged in a cluster. Usually, it is customary to select the two elements that are closest to each other based on the chosen distance measurement. The approach allows having up to n clusters, where n is one less than the number of data points. A *dendrogram* is a tree-like structure that is produced as the result of AHC where the child nodes of the dendrogram represent each data point and the clusters are shown as parents. We can cut across the tree at any level to have the desired number of clusters. By default, the dendrogram features all the available clusters; however, we can specify how many parents (clusters) to retain, which will affect the number of clusters formed. However, for comparisons between the different clustering methods in the module shown in Figure 4, we also provide the number of parents (clusters) to AHC. The limitation of AHC is that it is very inefficient and has issues with scaling to large datasets [48], [49], [50].

3) GAUSSIAN MIXTURE MODEL

The Gaussian mixture model(GMM) [42] features a mixture of Gaussian distributions to map out the clusters in the data. The ability to use multiple Gaussian allows the segmentation to allow some level of uncertainty at the class boundaries in the prediction. Maximum a posteriori estimation [51] can map out the rigid boundaries of the classes in the data. GMM has been applied to a wide range of clustering and image segmentation problems that include reef habitat mapping. We need to implement a grid search using the Bayesian information criterion (BIC) plot to determine the optimal number of clusters in the data. This will help identify the minimum number of classes with low compactness, which can be used for further hyperparameter tuning in the GMM. BIC penalizes models that have a high number of clusters to prevent over-fitting. The most suitable number of k is determined by selecting the smallest cluster with a low change in gradient. This helps to minimize the cost of running the GMM model. Figure 5b shows the BIC analysis of the GMM model over different numbers of clusters k. We select the optimal k similarly to the ELBO method in k-means and there is an extra hyperparameter in GMM that requires manual tuning - the initial state of the Gaussian. In our trial

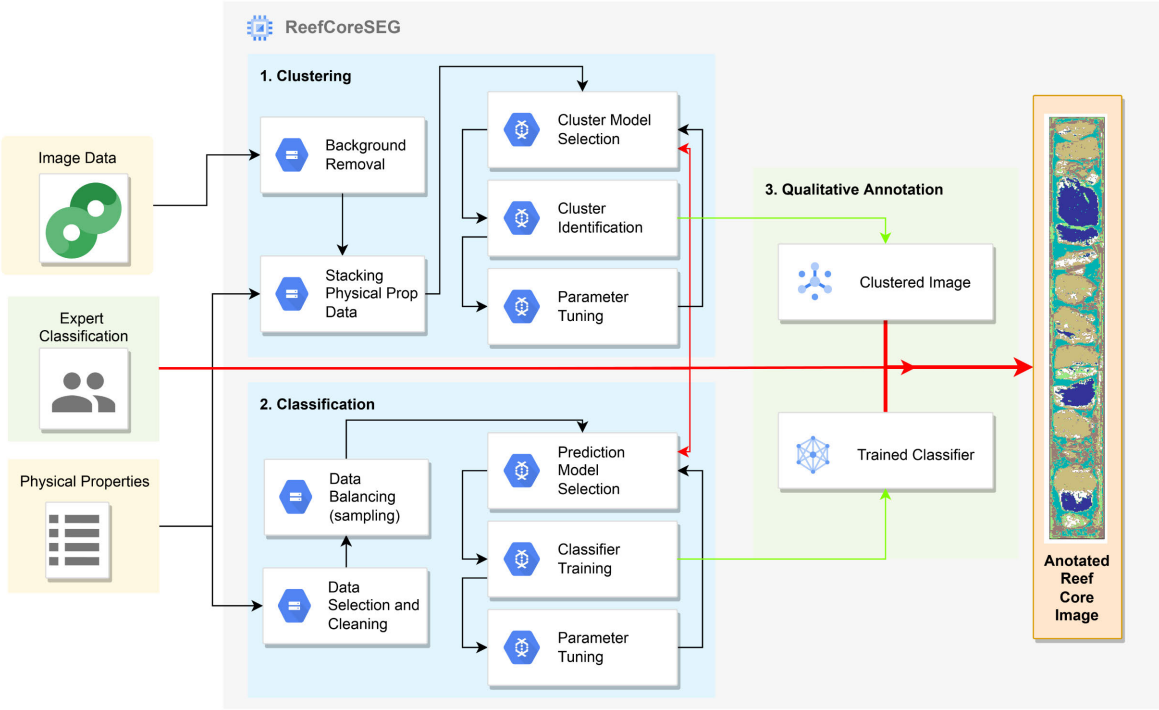


FIGURE 4. ReefCoreSeg framework showing three independent modules; 1.) Clustering, 2.) Classification, and 3.) Annotation.

experiments, we found that setting the initial state to 10 is optimal.

E. CLUSTERING EVALUATION METRICS

We need a common metric to compare the clustering methods although each clustering method features a metric for the quality of clusters. Firstly, we look at the numerical stats such as the Silhouette score [52] and the Calinski score [53] which measure the compactness and separability between clusters. The silhouette score measures an object's proximity to its assigned cluster compared to neighboring clusters and ranges from -1 to 1. The higher the silhouette score, the better the alignment between the object and its designated cluster, and the more dissimilar it is from adjacent clusters. Scores close to 1 indicate highly effective clustering outcomes, while scores nearing -1 suggest that the object should be reassigned to a different cluster. The Calinki score (Calinski-Harabasz index) considers both inter-cluster separation and intra-cluster cohesion to derive the ratio of between-cluster variance to within-cluster variance. It incorporates the number of clusters in the numerical formulation for a nuanced quantitative measure of clustering quality. Elevated Calinski scores indicate well-defined and distinct clusters. We also compare the models based on visual analyses by plotting the clustering results as images using unique colour mapping for individual clusters. We then do a side-by-side comparison of the plots to identify the tightness and separation of the clustered segments. The clearer and more distinct each cluster is, the better the performance.

F. CLASSIFICATION MODULE

We compare four supervised machine learning models for the classification module in the ReefCoreSeg framework. We extend the idea of using a classification model proposed by Insua et al. [29]; however, we use only three input features from Table 1 and ignore the magnetic susceptibility feature used in the original experiments. We implement this since there are not enough measurements for the magnetic susceptibility in core 33-A-16R; therefore, we cannot use it for our annotation module.

The support vector classifier (SVC) [54], [55] is an extension of the *perceptron* model that distinguishes between two classes of data. The SVC creates a hyperplane to separate the different classes and to extend classification to multiple classes, we can use either the one-versus-one or the one-versus-all approach. The classification of new instances is mapped to either each class or a collection of the other classes treated as one class. We use the one-versus-all approach as it is computationally more efficient than its counterpart. The success of SVC is based on the choice of kernel and we use the most simplistic kernel, i.e. the linear kernel to prioritize efficiency. We determine the best combination of the other hyperparameters, including regularization, by a grid search.

The random forest classifier (RFC) [56] is an ensemble-based machine learning model for classification that features a group of *decision trees*. Each decision tree uses a subset of the original data containing features and labels to develop a set of rules for making predictions where the predictions get combined either by averaging or voting depending on the

TABLE 2. Comparison of the three clustering methods with 1.) image data, and 2.) image data combined with physical properties data. The physical properties data did not have any significant contribution towards improving the clustering performance. We see that K-means provided the highest silhouette and Calinski score, while GMM underperformed. Therefore, we conducted a visual analysis to verify the results.

Method	k	Image Data		Image + Physical Properties	
		Silhouette	Calinski	Silhouette	Calinski
K-means	3	0.578	135278.576	0.568	134506.071
	4	0.571	172865.850	0.564	170318.221
	5	0.532	185517.386	0.526	183267.867
	6	0.486	191894.409	0.479	188536.297
	7	0.471	198750.707	0.464	195674.078
	8	0.451	204668.864	0.444	200515.483
	9	0.437	207444.994	0.425	202218.540
	10	0.422	206365.794	0.425	204680.450
AHC	3	0.557	117498.730	0.583	129266.675
	4	0.530	142985.726	0.539	129949.629
	5	0.508	170455.586	0.510	176768.472
	6	0.491	170848.071	0.466	166020.464
	7	0.455	174758.601	0.430	166239.892
	8	0.435	184487.133	0.411	171012.290
	9	0.402	188474.925	0.411	177458.553
	10	0.387	181522.784	0.379	184676.437
GMM	3	0.289	37710.312	0.279	39057.007
	4	0.333	68554.137	0.330	69245.508
	5	0.310	71918.885	0.305	72996.008
	6	0.222	63897.055	0.264	63848.365
	7	0.235	70671.966	0.208	70371.997
	8	0.136	41691.842	0.207	57259.417
	9	0.159	43757.392	0.126	45453.273
	10	0.128	36392.288	0.070	44688.634

type of problem. Hence, RFC combines multiple uncorrelated models (decision trees) that perform much better as a group than they do alone. In the case of classification, each randomly selected tree gives a classification and the majority ''vote'' is selected as the class label. RFC are well suited for multi-class classification and are resistant to noise [57] and also demonstrated effective for class-imbalance datasets and have been prominent in machine learning competitions [29].

The multilayer perceptron (MLP) is a machine learning model inspired by biological learning systems for classification and regression problems. MLP consists of a series of layers featuring nodes (neurons) with interconnections (synapses) that are trained based on data and a training algorithm; hence, also referred to as neural networks. Although MLP models have been used in a wide range of applications, we review selected applications to coral reefs and geoscientific minerology. Awalludin et al. [58] created a model for the classification of coral reef components using color and texture features from drill core imagery. Zhong et.al [59] performed a classification of tectonic settings using data from bulk basaltic rocks and basaltic volcanic glasses. MLP forms the basis of deep learning models that are prominent for various applications in health, Earth sciences, and multimedia applications [60], [61], [62], [63]. In our implementation, we train our MLP with the *adaptive moment estimation* (ADAM) optimizer [64] and determined model hyperparameters using grid search.

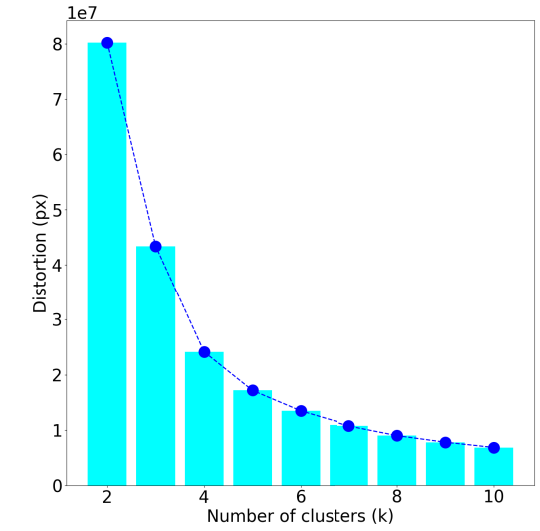
K-nearest neighbors (KNN) [65] is a supervised learning method that relies on sample proximity to classify the grouping of a data point. We select KNN for this specific reason as our data has a high correlation to the distance of measurement as physical properties on a single coral at 10-centimeter intervals yield the same lithology. KNNs use a majority vote approach to assign class labels. We use Euclidean distance as a measure of proximity in the KNN approach. KNNs have been successfully applied to a wide range of Earth science and reef studies such as classifying seismic-volcanic signals, classifying corals in underwater coral reef images, and reef benthic habitat mapping with object-based image analysis [44], [66], [67].

We compare the respective classification models using the classification accuracy given in Equation 1 as well as the F1 score given in Equation 4 for each class label. We also use confusion matrices [68] to show the classification per-class as the data features a high-class imbalance that are prominently used in the multi-class classification literature [69], [70], [71]. We select the best model using the *reciever operating characteristic* (ROC) curves and the *area under the curve* (AUC) score. We also report the *macro-averaged* AUC score of all the individual classes, which is also consistent with the literature.

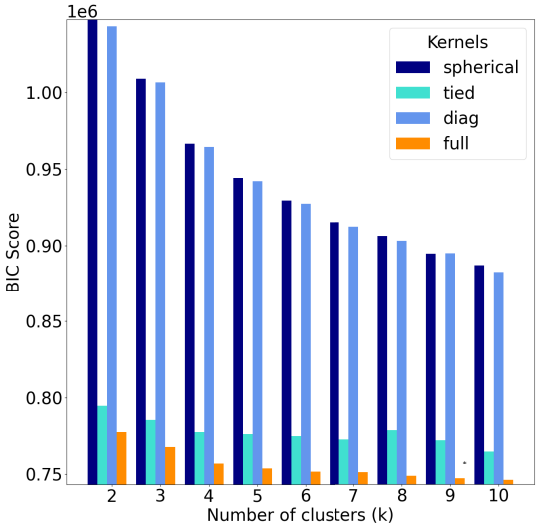
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$



(a) ELBO



(b) BIC

FIGURE 5. Grid search for the optimal number of segments in clustering. [a] ELBO plot for *distortions* over a possible *number of clusters k*. [b] We use the Bayesian information criterion (BIC) to search for the optimal number of components in GMM with four different kernels (spherical, tied, diagonal, and full).

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \tag{4}$$

where the *true negative* (TN) reflects the number of negative examples correctly classified, and *true positive* (TP) indicates the number of positive examples correctly classified. The *false positive* (FP) represents the number of actual negative examples classified as positive and *false negative* (FN) represents the number of positive examples classified as negative.

G. QUALITATIVE ANNOTATION MODULE

The final stage of our framework creates an annotated image of drill core data by selecting and combining the best models from the clustering and classification modules. This is the stage of the framework where we use a human expert for visual analysis of the clustered images to select the model that has the highest degree of separation within the clusters. The selected model results are then stacked against the classification of the best-performing classifier, which is only fed the selected core data and the annotations are then drawn visually. We note that several papers in the literature have used similar expert-based qualitative annotations [10], [72], [73], [74]. However, these papers do not have any labelled physical properties data for supervised machine learning. Their datasets are very small with respect to the number of labelled images of each class of coral; therefore, we can not directly use them. It would also require a lot of time to manually annotate any physical properties data associated with these datasets. We only use core 33-A-16R to test the performance of the framework, which could then be applied to the other datasets for validation.

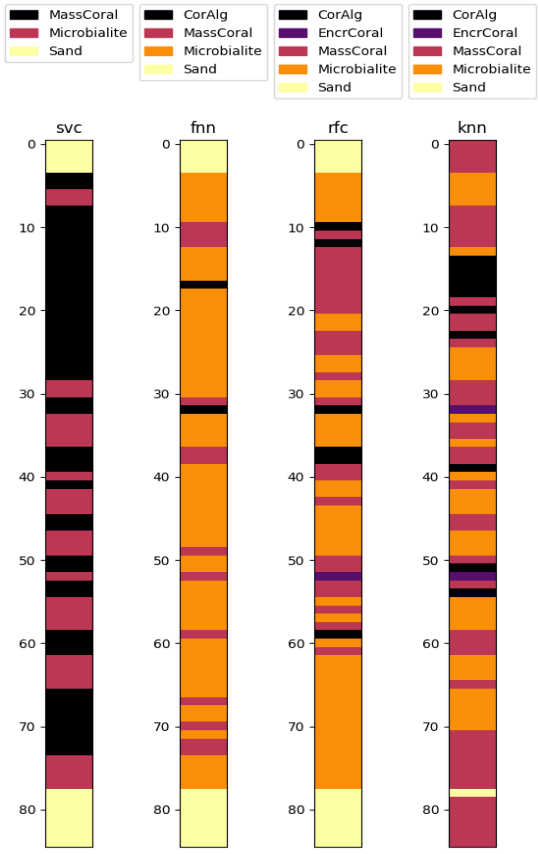


FIGURE 6. Classification performance for the respective models on Core 33-A-16R. Each model can identify some of the classes present in the data; however, KNN and RFC identify the most number of classes. We observe that KNN is over-predicting the presence of Coralline Algae, and RFC has the best prediction compared to the expert-labelled image by visual inspection.

III. RESULTS

A. CLUSTERING MODULE

We first evaluate the performance of the clustering methods (k-means, HAC, and GMM) used in the study. We present

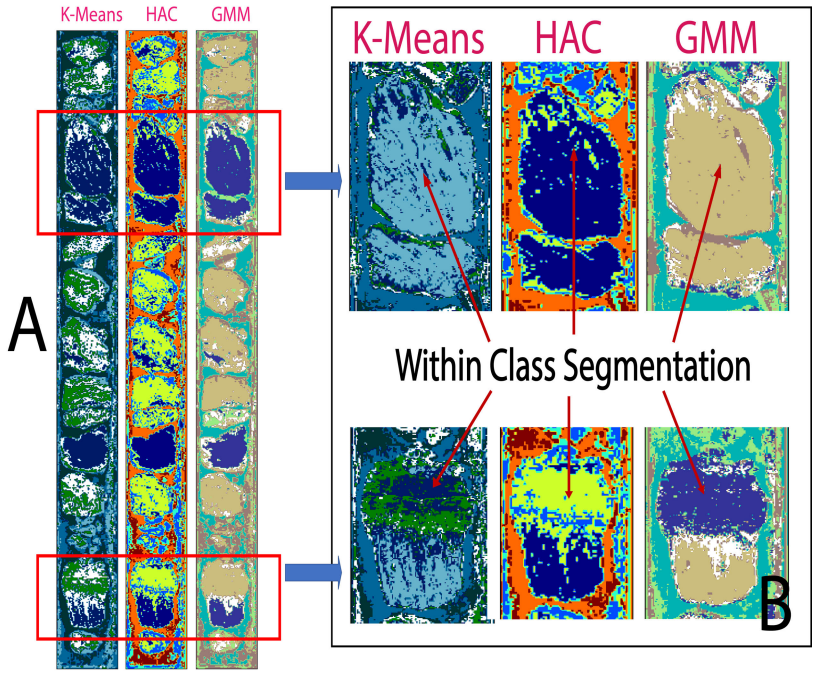


FIGURE 7. Clustering performance with $k = 6$ for k-means, HAC, and GMM. A. Clustering on the core section, showing little distinction between the models. B. Within-class segmentation performance on massive corals, showing that GMM has the best performance.

the results for the standalone image data and the physical properties data combined with the image data. Table 2 shows results for a grid search for the optimal number of clusters within the data. The silhouette and Calinski scores are given to demonstrate the performance of each of these models. K-means had the highest scores for both the Shilouette and Calinksi scores. There is no significant change in performance for K-means and AHC after adding physical properties information as shown in Table 2. The only exception is GMM, where we see a slight increase in performance at larger values of k. We also observe that GMM has a significantly lower silhouette score with both types of data. We later provide a visual analysis (Figure 7) of the clustered cores to further analyze the performance of the three clustering methods, since in image-based clustering a higher Shilouette and Calinski score do not necessarily imply better clustering performance.

Figure 7 shows the clustering performance of the models on the core image. We only show the results for $k = 6$ as they are a good proxy for all other cases. Each cluster of images has the performance of the three algorithms side-by-side. The images include the clustering performance of k-means, HAC, and GMM, respectively. We specifically show the performance for 6 clusters as we are working with physical properties data that were annotated into 6 different lithologies. Figure 7–A shows that GMM is doing a much better job at clustering the larger corals. This is evident from looking at Figure 7–B which shows a close-up of the clustering of a massive coral section within the core. The top two images are k-means and HAC results, and we can see that these two models are trying to take the larger corals and cluster within them by separating all the smaller holes within the coral. We observe that the GMM is better at identifying the boundaries of the massive coral segment and treating it as a single object with some minor imperfections.

B. CLASSIFICATION MODULE

The next present the classification performance on the physical properties data. Table 4 shows the average classification accuracy of the respective methods for the test dataset. We observe that RFC gives the best overall accuracy followed by MLP and KNN and SVC has the worst performance. We further review the per-class classification performance in Table 3 and observe that RFC is the best for individual class classification. KNN is also close to the RFC; however, it struggled at classifying encrusted corals. Furthermore, we observe that the respective methods have difficulty in classifying the encrusted corals and coralline algae. The confusion matrix given in Figure 8 shows a breakdown of the misclassification. Moreover, we can observe in Figure 8d that RFC has the lowest misclassification rate for most of the classes, and SVC has the worst performance. SVC could not predict any of the coralline algae or the encrusted corals properly. MLP and KNN have similar performances as shown in Figure 8c and 8b; therefore, we use the ROC curves and the AUC score to select the best classifier. Figure 9 shows the RFC with a score of 0.95 beating out all the other classifiers.

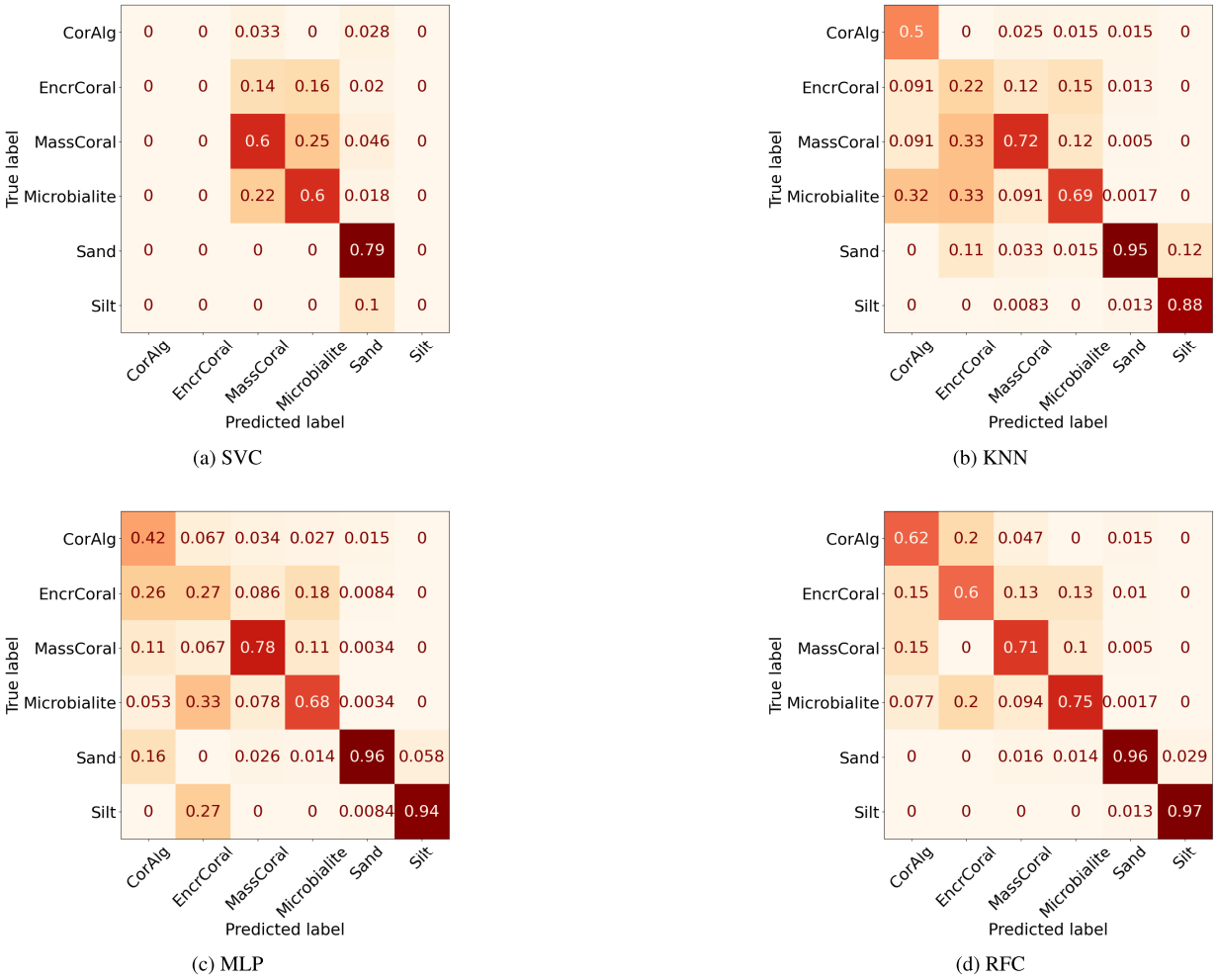


FIGURE 8. **Confusion matrix for a typical experimental run by the respective classifiers (SVC, KNN, MLP, RFC), where RFC has the best classification across all the classes.**

C. EXPERT-BASED QUALITATIVE EVALUATION

We present the final set of results in Figure 6 which shows the predictions of each of the trained classifiers on predicting the classes of data present in core 33-A-16R. This result is then used for the qualitative annotation for further evaluation by our domain expert, Jody Webster from the University of Sydney, who is also a coauthor in this study. The qualitative analysis is done by comparing model performance by sight and the expert-labelled image given in Insua et al. [29] is used as the benchmark for comparison in our study. As expected, SVC had the worst performance as it only managed to find three classes. The MLP in Figure 6 is only marginally better than the SVC since it found four of the five classes present in the data. The KNN and RFC did well and found all five classes of data. However, RFC gives the best classification performance, both in terms of accuracy score, and expert-based visual analysis as shown in Figure 10. This is also consistent with the performance measured by the AUC score, where RFC has the highest score.

Figure 10 shows the output of the ReefCoreSeg framework highlighting the classification of the core segments with GMM and RFC. The legend shows that the framework can clearly segment the image into six distinct segments; however, the classifier can only find five classes of data. We notice the misclassification of the coralline algae and sand; however, the model did exceptionally well at identifying the massive corals as well as the stromatolitic microbialites. A small section of encrusted coral has been annotated correctly.

IV. DISCUSSION

In general, we find that ReefCoreSeg successfully automates the annotation of coral core images into six lithologies by combining unsupervised (clustering) and supervised learning (classification). We find that GMM clustering gives the best performance out of all the clustering methods tested, since it managed to properly identify the boundaries of the different classes present in the data. The visual analysis done by coral reef geologists consists of distinguishing distinct boundaries between corals. This is a specifically challenging task due to the muted colors and textures within the core image. The hard boundaries between larger pieces

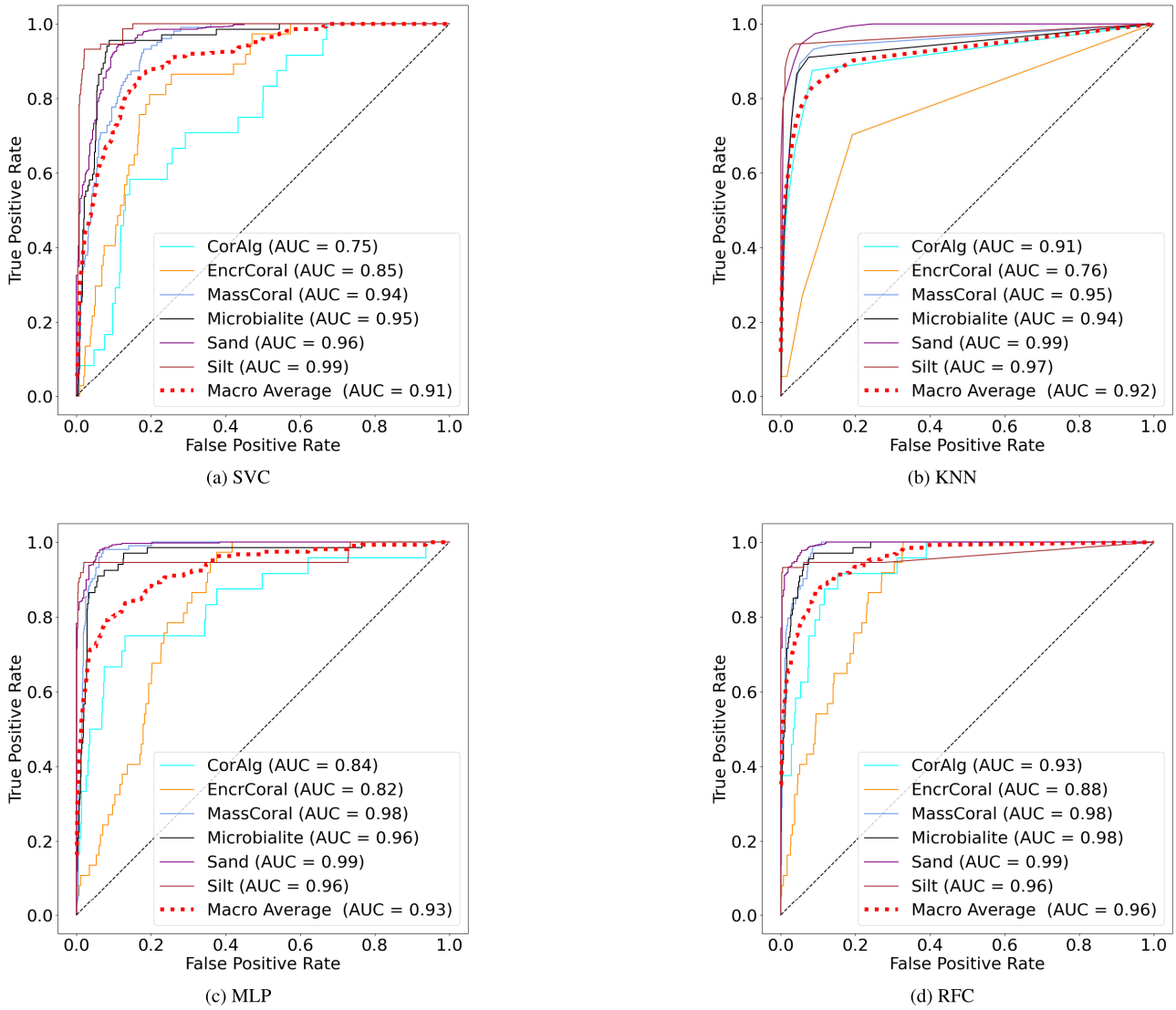


FIGURE 9. ROC curve for the respective classifiers (SVC, KNN, MLP, RFC). The macro-averaged AUC (Macro AUC) score shows that the RFC has the best overall classification performance.

of corals, microbialites, and other classes have been well identified by all the clustering methods; however, the soft boundaries within a lithology (such as the textures within a massive coral) or more complex structures (such as encrusted corals) have been difficult to segment. GMM-based clustering identified the major segments of the drill-core data and was able to distinguish between skeletal textures within a single coral fragment better than the other clustering techniques. We also find that stacking physical properties data with the image allowed a small improvement in the clustering performance compared to only using image data. The similarities within the corals are much better captured by the multi-sensor core logger, as the physical properties were approximately the same for a single piece of the core segment. This could have contributed to the slight improvement in performance.

The classification performance of the four methods shows great promise in terms of classification accuracy; however, the performance of the models on Core 33-A-16R shows the importance of visual analysis and expert intervention in the ReefCoreSeg framework. Although the respective models have high classification accuracy, only the RFC properly identified the various components within the core. RFC had the best overall classification in terms of the ROC curves as shown in Figure 9. The performance in classifying the mass coral class was better on the given drill core, since it was the second most abundant class within our dataset. RFC has some drawbacks in terms of the misclassification of sand and coralline algae as shown in Figure 10 with green labels. We attribute this to the high class imbalance within the data. Balancing the data improves the performance, giving a better score than the previous classification approach [29]; however,

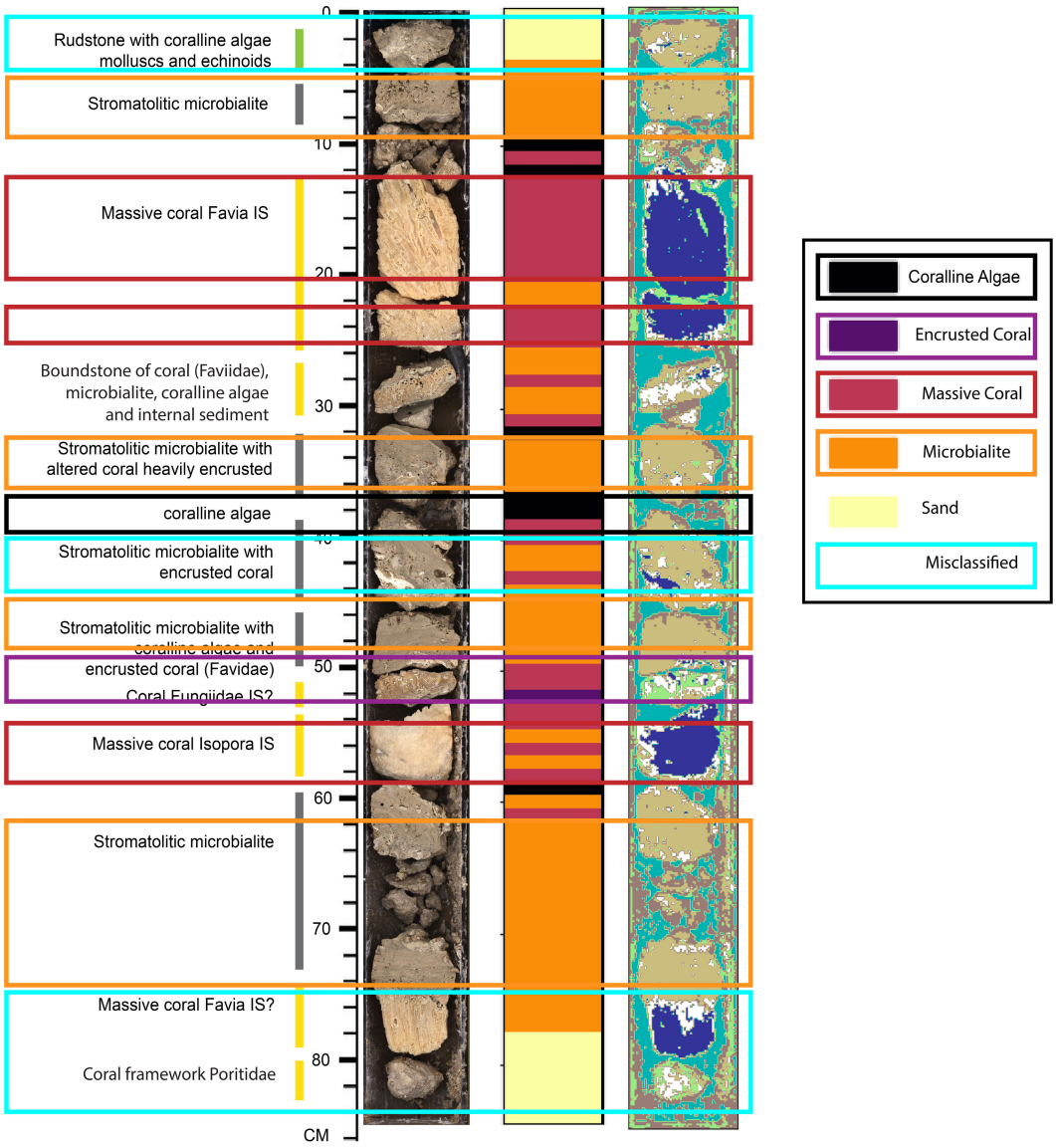


FIGURE 10. **Annotated image of Core 33-A-16R with GMM-based clustering and classification by RFC.**

the example of core 33-A-16R shows that there is scope for improvement in classifying the minority classes.

We note that multi-class classification on an imbalanced dataset is a challenging task. We found that the random forest model managed to give a good predictive performance on the data as compared to the other tested models. However, the minority classes are still underrepresented, and the majority class remains overrepresented. This shows the importance of expert intervention in the ReefCoreSeg module as we want to select the best classifier for the annotation via expert visual analysis. This hybrid framework shows promising annotation performance and sets up a clear platform for further enhancements. Further work can focus on uncertainty quantification in the annotation framework. This can be done with methods such as Bayesian neural networks for model based uncertainty to the classification module. There is also scope for connecting the framework with one-dimensional reef evolution models such as pyReef-core and its counterpart known as Bayesreef that employed Bayesian inference for model parameter estimation representing environmental conditions that affect reef growth over thousands of years. The class imbalance within the framework could also be addressed with more robust adversarial learning sampling techniques and their combination with resemble models. Furthermore, we also need to understand the decision-making process; since we have mostly used black-box models, it is not clear how the decision has been made. This can be addressed by extracting if-then rules from the random forest model and from novel methods in the field of explainable artificial intelligence (XAI).

TABLE 3. Classification performance for the physical properties data where RFC has the best overall accuracy as well as F1 scores for per-class classification. We note that SVC has not been able to predict the coralline algae, encrusted corals, and silt.

		Precision	Recall	F1-score	Accuracy
SVC	CorAlg	0.00	0.00	0.00	0.76
	EncrCoral	0.00	0.00	0.00	
	MassCoral	0.60	0.53	0.57	
	Microbialite	0.60	0.51	0.55	
	Sand	0.79	1.00	0.88	
	Silt	0.00	0.00	0.00	
KNN	CorAlg	0.50	0.46	0.48	0.88
	EncrCoral	0.22	0.05	0.09	
	MassCoral	0.72	0.84	0.78	
	Microbialite	0.69	0.67	0.68	
	Sand	0.95	0.97	0.96	
	Silt	0.88	0.88	0.88	
MLP	CorAlg	0.42	0.33	0.37	0.89
	EncrCoral	0.27	0.11	0.15	
	MassCoral	0.78	0.87	0.82	
	Microbialite	0.68	0.75	0.71	
	Sand	0.96	0.98	0.97	
	Silt	0.94	0.88	0.91	
RFC	CorAlg	0.62	0.33	0.43	0.90
	EncrCoral	0.60	0.08	0.14	
	MassCoral	0.71	0.88	0.79	
	Microbialite	0.75	0.78	0.76	
	Sand	0.96	0.99	0.97	
	Silt	0.97	0.89	0.93	

TABLE 4. Average classification accuracy over 10 experimental runs with different initial weights and biases. '-' indicates the method not being used.

Classifier	Balanced data testing performance	[29] testing performance
SVC	0.7562	0.7448
KNN	0.8758	-
MLP	0.8882	-
RFC	0.8984	0.7896

Furthermore, with reef drill cores, we are presented with a limited data problem as we have limited expert annotated core images for testing the robustness of the framework. The annotations that do exist for such datasets, cannot be directly used by machine learning as they are stored in formats that need to be manually transformed into labelled training datasets. Finally, there is scope to use the ReefCoreSeg on other drill core data such as ice cores from glaciers and ice sheets, rock cores from mineral exploration, and oil exploration that have been studied and annotated to improve and validate the framework.

V. CONCLUSION

Fossil reef drill-core represent valuable archives that can be used to reconstruct paleoenvironmental changes over thousands of years, yet their description and classification remain challenging. Hence, we presented a framework that utilized reef-core data with a combination of unsupervised and supervised learning to identify and annotate coral assemblages given as lithologies. We found that there are a lot of muted colors and textures in the core image. This coupled with soft transition boundaries within corals made it a very challenging task for clustering techniques to distinguish the different features of lithologies present in the drill-core data. The GMM-based clustering distinguished the boundaries well and identified larger corals whilst avoiding the problem of segmenting within one coral lithology. This was further verified by qualitative evaluation of the results. We also found that adding physical properties data to image data slightly impacts clustering performance. Hence, stacking more data types, such as hyperspectral and computed tomography scan (CT scan) data could further enhance the clustering performance.

ACKNOWLEDGMENT

The authors express gratitude toward the crew, technical support and science party of IODP Expedition 325. They also like to thank Tania Lado Insua for providing the labelled training data used for training the classification module within this project.

CODE AND DATA

The code and data for our framework is available: https://github.com/rvdeo/reefcoreseg

REFERENCES

[1] D. W. Kinsey and D. Hopley, "The significance of coral reefs as global carbon sinks—Response to greenhouse," *Global Planet. Change*, vol. 3, no. 4, pp. 363–377, Mar. 1991. [Online]. Available: https://www.sciencedirect.com/science/article/pii/003101829190172N

[2] C. J. Crossland, B. G. Hatcher, and S. V. Smith, "Role of coral reefs in global ocean production," *Coral Reefs*, vol. 10, no. 2, pp. 55–64, Jun. 1991, doi: 10.1007/bf00571824.

[3] H. S. Cesar and P. J. van Beukering, "Economic valuation of the coral reefs of Hawai'i," *Pacific Sci.*, vol. 58, no. 2, pp. 231–242, 2004.

[4] J. Cinner, "Coral reef livelihoods," *Current Opinion Environ. Sustainability*, vol. 7, pp. 65–71, Apr. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877343513001875

[5] R. A. Morais and D. R. Bellwood, "Principles for estimating fish productivity on coral reefs," *Coral Reefs*, vol. 39, no. 5, pp. 1221–1231, Oct. 2020, doi: 10.1007/s00338-020-01969-9.

[6] R. Costanza, R. de Groot, P. Sutton, S. van der Ploeg, S. J. Anderson, I. Kubiszewski, S. Farber, and R. K. Turner, "Changes in the global value of ecosystem services," *Global Environ. Change*, vol. 26, pp. 152–158, May 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959378014000685

[7] D. K. Hubbard, "Reef drilling," in *Encyclopedia of Modern Coral Reefs: Structure, Form and Process*, D. Hopley, Ed. Dordrecht, The Netherlands: Springer, 2011, pp. 856–869, doi: 10.1007/978-90-481-2639-2_54.

[8] T. Salles, X. Ding, J. M. Webster, A. Vila-Concejo, G. Brocard, and J. Pall, "A unified framework for modelling sediment fate from source to sink and its interactions with reef systems over geological times," *Sci. Rep.*, vol. 8, no. 1, p. 5252, Mar. 2018. [Online]. Available: https://www.nature.com/articles/s41598-018-23519-8

[9] J. F. Marshall and P. J. Davies, "Internal structure and Holocene evolution of one tree reef, southern great barrier reef," *Coral Reefs*, vol. 1, no. 1, pp. 21–28, Jun. 1982, doi: 10.1007/bf00286536.

[10] K. L. Sanborn, J. M. Webster, G. E. Webb, J. C. Braga, M. Humblet, L. Nothdurft, M. A. Patterson, B. Dechnik, S. Warner, T. Graham, R. J. Murphy, Y. Yokoyama, S. P. Obrochta, J.-X. Zhao, and M. Salas-Saavedra, "A new model of Holocene reef initiation and growth in response to sea-level rise on the southern great barrier reef," *Sedimentary Geol.*, vol. 397, Mar. 2020, Art. no. 105556. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0037073819302088

[11] A. Lim, A. J. Wheeler, and L. Conti, "Cold-water coral habitat mapping: Trends and developments in acquisition and processing methods," *Geosciences*, vol. 11, no. 1, p. 9, Dec. 2020. [Online]. Available: https://www.mdpi.com/2076-3263/11/1/9

[12] D. M. Thompson, "Environmental records from coral skeletons: A decade of novel insights and innovation," *WIREs Climate Change*, vol. 13, no. 1, p. e745, Jan. 2022. [Online]. Available: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcc.745

[13] F. Foglini, L. Angeletti, V. Bracchi, G. Chimienti, V. Grande, I. M. Hansen, A. N. Meroni, F. Marchese, A. Mercorella, M. Prampolini, M. Taviani, A. Vertino, F. Badalamenti, C. Corselli, I. Erdal, E. Martorelli, and A. Savini, "Underwater hyperspectral imaging for seafloor and benthic habitat mapping," in *Proc. IEEE Int. Workshop Metrology Sea, Learn. Measure Sea Health Parameters (MetroSea)*, Oct. 2018, pp. 201–205.

[14] Y. Yokoyama, J. M. Webster, C. Cotterill, J. C. Braga, L. Jovane, H. Mills, S. Morgan, and A. Suzuki, "IODP expedition 325: Great barrier reefs reveals past sea-level, climate and environmental changes since the last ice age," *Sci. Drilling*, vol. 12, pp. 32–45, Sep. 2011, doi: 10.2204/iodp.sd.12.04.2011.

[15] A. Fu, M. S. Hosseini, and K. N. Plataniotis, "Reconsidering CO_2 emissions from computer vision," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2021, pp. 2311–2317. [Online]. Available: https://openaccess.thecvf.com/content/CVPR2021W/RCV/html/Fu_Reconsidering_CO2_Emissions_From_Computer_Vision_CVPRW_2021_paper.html

[16] P. Ramos-Giraldo, C. Reberg-Horton, A. M. Locke, S. Mirsky, and E. Lobaton, "Drought stress detection using low-cost computer vision systems and machine learning techniques," *IT Prof.*, vol. 22, no. 3, pp. 27–29, May 2020.

[17] S. J. Purkis, "Remote sensing tropical coral reefs: The view from above," *Annu. Rev. Marine Sci.*, vol. 10, pp. 149–168, Jan. 2018, doi: 10.1146/annurev-marine-121916-063249.

[18] J. Li, D. E. Knapp, N. S. Fabina, E. V. Kennedy, K. Larsen, M. B. Lyons, N. J. Murray, S. R. Phinn, C. M. Roelfsema, and G. P. Asner, "A global coral reef probability map generated using convolutional neural networks," *Coral Reefs*, vol. 39, no. 6, pp. 1805–1815, Dec. 2020, doi: 10.1007/s00338-020-02005-6.

[19] M. González-Rivero, O. Beijbom, A. Rodriguez-Ramirez, D. E. P. Bryant, A. Ganase, Y. Gonzalez-Marrero, A. Herrera-Reveles, E. V. Kennedy, C. J. S. Kim, S. Lopez-Marcano, K. Markey, B. P. Neal, K. Osborne, C. Reyes-Nivia, E. M. Sampayo, K. Stolberg, A. Taylor, J. Vercelloni, M. Wyatt, and O. Hoegh-Guldberg, "Monitoring of coral reefs using artificial intelligence: A feasible and cost-effective approach," *Remote Sens.*, vol. 12, no. 3, p. 489, Feb. 2020.

[20] M. B. Lyons, C. M. Roelfsema, E. V. Kennedy, E. M. Kovacs, R. Borrego-Acevedo, K. Markey, M. Roe, D. M. Yuwono, D. L. Harris, S. R. Phinn, G. P. Asner, J. Li, D. E. Knapp, N. S. Fabina, K. Larsen, D. Traganos, and N. J. Murray, "Mapping the world's coral reefs using a global multiscale Earth observation framework," *Remote Sens. Ecol. Conservation*, vol. 6, no. 4, pp. 557–568, Dec. 2020.

[21] E. V. Kennedy, C. M. Roelfsema, M. B. Lyons, E. M. Kovacs, R. Borrego-Acevedo, M. Roe, S. R. Phinn, K. Larsen, N. J. Murray, D. Yuwono, J. Wolff, and P. Tudman, "Reef cover, a coral reef classification for global habitat mapping from remote sensing," *Sci. Data*, vol. 8, no. 1, p. 196, Aug. 2021, doi: 10.1038/s41597-021-00958-z.

[22] A. A. Mogstad, G. Johnsen, and M. Ludvigsen, "Shallow-water habitat mapping using underwater hyperspectral imaging from an unmanned surface vehicle: A pilot study," *Remote Sens.*, vol. 11, no. 6, p. 685, Mar. 2019. [Online]. Available: https://www.mdpi.com/2072-4292/11/6/685

[23] J. M. Webster and P. J. Davies, "Coral variation in two deep drill cores: Significance for the pleistocene development of the great barrier reef," *Sedimentary Geol.*, vol. 159, nos. 1–2, pp. 61–80, Jun. 2003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0037073803000952

[24] F. Alzubaidi, P. Mostaghimi, P. Swietojanski, S. R. Clark, and R. T. Armstrong, "Automated lithology classification from drill core images using convolutional neural networks," *J. Petroleum Sci. Eng.*, vol. 197, Feb. 2021, Art. no. 107933.

[25] F. J. Galdames, C. A. Perez, P. A. Estévez, and M. Adams, "Rock lithological instance classification by hyperspectral images using dimensionality reduction and deep learning," *Chemometric Intell. Lab. Syst.*, vol. 224, May 2022, Art. no. 104538. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169743922000491

[26] A. Thomas, M. Rider, A. Curtis, and A. MacArthur, "Automated lithology extraction from core photographs," *1st Break*, vol. 29, no. 6, Jun. 2011. [Online]. Available: https://www.earthdoc.org/content/journals/0.3997/1365-2397.29.6.51281

[27] H. L. Dawson, O. Dubrule, and C. M. John, "Impact of dataset size and convolutional neural network architecture on transfer learning for carbonate rock classification," *Comput. Geosci.*, vol. 171, Feb. 2023, Art. no. 105284. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0098300422002333

[28] E. E. Baraboshkin, L. S. Ismailova, D. M. Orlov, E. A. Zhukovskaya, G. A. Kalmykov, O. V. Khotylev, E. Y. Baraboshkin, and D. A. Koroteev, "Deep convolutions for in-depth automated rock typing," *Comput. Geosci.*, vol. 135, Feb. 2020, Art. no. 104330. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0098300419304686

[29] T. L. Insua, L. Hamel, K. Moran, L. M. Anderson, and J. M. Webster, "Advanced classification of carbonate sediments based on physical properties," *Sedimentology*, vol. 62, no. 2, pp. 590–606, Feb. 2015, doi: 10.1111/sed.12168.

[30] A. W. F. Edwards and L. L. Cavalli-Sforza, "A method for cluster analysis," *Biometrics*, vol. 21, no. 2, pp. 362–375, Jun. 1965.

[31] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm," *J. Roy. Stat. Society. Ser. C, Appl. Statist.*, vol. 28, no. 1, p. 100, 1979.

[32] S. Wazarkar and B. N. Keshavamurthy, "A survey on image data analysis through clustering techniques for real world applications," *J. Vis. Commun. Image Represent.*, vol. 55, pp. 596–626, Aug. 2018.

[33] I. Kotaridis and M. Lazaridou, "Remote sensing image segmentation advances: A meta-analysis," *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 309–322, Mar. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271621000265

[34] J. Yuan, D. Wang, and R. Li, "Remote sensing image segmentation by combining spectral and texture features," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 16–24, Jan. 2014.

[35] M. D. Hossain and D. Chen, "Segmentation for object-based image analysis (OBIA): A review of algorithms and challenges from remote sensing perspective," *ISPRS J. Photogramm. Remote Sens.*, vol. 150, pp. 115–134, Apr. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271619300425

[36] S. Nussbaum, *Object-Based Image Analysis and Treaty Verification : New Approaches in Remote Sensing—Applied to Nuclear Facilities in Iran*. Dordrecht, The Netherlands: Springer, 2008. [Online]. Available: http://repository.vnu.edu.vn/handle/VNU_123/27576

[37] H. Song, S. R. Mehdi, Y. Zhang, Y. Shentu, Q. Wan, W. Wang, K. Raza, and H. Huang, "Development of coral investigation system based on semantic segmentation of single-channel images," *Sensors*, vol. 21, no. 5, p. 1848, Mar. 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/5/1848

[38] V. Chirayath and R. Instrella, "Fluid lensing and machine learning for centimeter-resolution airborne assessment of coral reefs in American Samoa," *Remote Sens. Environ.*, vol. 235, Dec. 2019, Art. no. 111475. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425719304948

[39] J. Wang, W. Chen, Y. Wang, and J. Zou, "Coral reef pore recognition and feature extraction based on borehole image," *Mar. Georesources Geotechnol.*, vol. 40, no. 2, pp. 159–170, Feb. 2022. [Online]. Available: https://www.tandfonline.com/doi/pdf/10.1080/1064119X.2021.1874576

[40] D. Steinberg, A. Friedman, O. Pizarro, and S. B. Williams, "A Bayesian nonparametric approach to clustering data from underwater robotic surveys," in *Proc. 15th Int. Symp. Robot. Res.* Princeton, NJ, USA: CiteSeer, 2011, pp. 1–16.

[41] F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: An overview," *WIREs Data Mining Knowl. Discovery*, vol. 2, no. 1, pp. 86–97, Jan. 2012. [Online]. Available: https://onlinelibrary.wiley.com/doi/pdf/10.1002/widm.53

[42] G. J. McLachlan and K. E. Basford, *Mixture Models: Inference and Applications to Clustering*, vol. 38. New York, NY, USA: Marcel Dekker, 1988.

[43] F. Thabtah, S. Hammoud, F. Kamalov, and A. Gonsalves, "Data imbalance in classification: Experimental evaluation," *Inf. Sci.*, vol. 513, pp. 429–441, Mar. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025519310497

[44] N. A. B. Mary and D. Dharma, "Coral reef image classification employing improved LDP for feature extraction," *J. Vis. Commun. Image Represent.*, vol. 49, pp. 225–242, Nov. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1047320317301827

[45] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using k-means clustering algorithm and subtractive clustering algorithm," *Proc. Comput. Sci.*, vol. 54, pp. 764–771, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050915014143

[46] M. E. Celebi, H. A. Kingravi, and P. A. Vela, "A comparative study of efficient initialization methods for the k-means clustering algorithm," *Expert Syst. Appl.*, vol. 40, no. 1, pp. 200–210, Jan. 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417412008767

[47] D. C. Hambrick, "Taxonomic approaches to studying strategy: Some conceptual and methodological issues," *J. Manage.*, vol. 10, no. 1, pp. 27–41, Apr. 1984, doi: 10.1177/014920638401000104.

[48] E. M. Rasmussen and P. Willett, "Efficiency of hierarchic agglomerative clustering using the ICL distributed array processor," *J. Documentation*, vol. 45, no. 1, pp. 1–24, Jan. 1989, doi: 10.1108/eb026836.

[49] P. Govender and V. Sivakumar, "Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019)," *Atmos. Pollut. Res.*, vol. 11, no. 1, pp. 40–56, Jan. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1309104219304556

[50] A. Rajabi, M. Eskandari, M. J. Ghadi, L. Li, J. Zhang, and P. Siano, "A comparative study of clustering techniques for electrical load pattern segmentation," *Renew. Sustain. Energy Rev.*, vol. 120, Mar. 2020, Art. no. 109628. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032119308354

[51] J.-L. Gauvain and C.-H. Lee, "Maximum a posteriori estimation for multivariate Gaussian mixture observations of Markov chains," *IEEE Trans. Speech Audio Process.*, vol. 2, no. 2, pp. 291–298, Apr. 1994.

[52] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, pp. 53–65, Nov. 1987. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0377042787901257

[53] T. Calinski and J. Harabasz, "A dendrite method for cluster analysis," *Commun. Statist. Theory Methods*, vol. 3, no. 1, pp. 1–27, Jan. 1974. [Online]. Available: https://www.tandfonline.com/doi/pdf/10.1080/03610927408827101

[54] J. Weston and C. Watkins, "Multi-class support vector machines," CiteSeer, Princeton, NJ, USA, Tech. Rep., 1998.

[55] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/bf00994018.

[56] T. K. Ho, "Random decision forests," in *Proc. 3rd Int. Conf. Document Anal. Recognit.*, vol. 1, Aug. 1995, pp. 278–282.

[57] Y. Mishina, R. Murata, Y. Yamauchi, T. Yamashita, and H. Fujiyoshi, "Boosted random forest," *IEICE Trans. Inf. Syst.*, vol. 98, no. 9, pp. 1630–1636, 2015. [Online]. Available: https://search.ieice.org/bin/summary.php?id=e98-d_9_1630&category=D&year=2015&lang=E&abst=

[58] E. A. Awalludin, M. S. Hitam, W. N. J. H. W. Yussof, and Z. Bachok, "Classification of coral reef components using color and texture features," *J. Telecommun., Electron. Comput. Eng.*, vol. 9, nos. 3–4, pp. 109–113, Oct. 2017. [Online]. Available: https://jtec.utem.edu.my/jtec/article/view/2927

[59] R. Zhong, Y. Deng, and C. Yu, "Multi-layer perceptron-based tectonic discrimination of basaltic rocks and an application on the Paleoproterozoic Xiong'er volcanic province in the North China Craton," *Comput. Geosci.*, vol. 149, Apr. 2021, Art. no. 104717. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0098300421000303

[60] D. Ravì, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G.-Z. Yang, "Deep learning for health informatics," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 1, pp. 4–21, Jan. 2017.

[61] A. King, S. M. Bhandarkar, and B. M. Hopkinson, "A comparison of deep learning methods for semantic segmentation of coral reef survey images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 1394–1402.

[62] G. Camps-Valls, D. Tuia, X. X. Zhu, and M. Reichstein, *Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences*. Hoboken, NJ, USA: Wiley, 2021.

[63] M. Zampoglou, F. Markatopoulou, G. Mercier, D. Touska, E. Apostolidis, S. Papadopoulos, R. Cozien, I. Patras, V. Mezaris, and I. Kompatsiaris, "Detecting tampered videos with multimedia forensics and deep learning," in *MultiMedia Modeling* (Lecture Notes in Computer Science), I. Kompatsiaris, B. Huet, V. Mezaris, C. Gurrin, W.-H. Cheng, and S. Vrochidis, Eds. Cham, Switzerland: Springer, 2019, pp. 374–386.

[64] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[65] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 1, pp. 21–27, Jan. 1967.

[66] M. Bicego, A. Rossetto, M. Olivieri, J. M. Londoño-Bonilla, and M. Orozco-Alzate, "Advanced KNN approaches for explainable seismic-volcanic signal classification," *Math. Geosci.*, vol. 55, no. 1, pp. 59–80, Jan. 2023, doi: 10.1007/s11004-022-10026-w.

[67] N. Wahidin, V. P. Siregar, B. Nababan, I. Jaya, and S. Wouthuyzen, "Object-based image analysis for coral reef benthic habitat mapping with several classification algorithms," *Proc. Environ. Sci.*, vol. 24, pp. 222–227, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1878029615000973

[68] M. Grandini, E. Bagli, and G. Visani, "Metrics for multi-class classification: An overview," 2020, *arXiv:2008.05756*.

[69] R. Yacouby and D. Axman, "Probabilistic extension of precision, recall, and F1 score for more thorough evaluation of classification models," in *Proc. 1st Workshop Eval. Comparison NLP Syst.* Stroudsburg, PA, USA: Association for Computational Linguistics, Nov. 2020, pp. 79–91, doi: https://aclanthology.org/2020.eval4nlp-1.9.

[70] M. Raza, F. K. Hussain, O. K. Hussain, M. Zhao, and Z. U. Rehman, "A comparative analysis of machine learning models for quality pillar assessment of SaaS services by multi-class text classification of users' reviews," *Future Gener. Comput. Syst.*, vol. 101, pp. 341–371, Dec. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X19300196

[71] D. M. Farid, L. Zhang, C. M. Rahman, M. A. Hossain, and R. Strachan, "Hybrid decision tree and naïve Bayes classifiers for multi-class classification tasks," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1937–1946, Mar. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417413007100

[72] B. Dechnik, J. M. Webster, P. J. Davies, J.-C. Braga, and P. J. Reimer, "Holocene 'turn-on' and evolution of the Southern Great Barrier Reef: Revisiting reef cores from the Capricorn Bunker Group," *Mar. Geol.*, vol. 363, pp. 174–190, May 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0025322715000596

[73] M. Humblet and J. M. Webster, "Coral community changes in the great barrier reef in response to major environmental changes over glacial-interglacial timescales," *Palaeogeography, Palaeoclimatol., Palaeoecol.*, vol. 472, pp. 216–235, Apr. 2017, doi: 10.1016/j.palaeo.2017.02.003.

[74] E. Abbey, J. M. Webster, J. C. Braga, K. Sugihara, C. Wallace, Y. Iryu, D. Potts, T. Done, G. Camoin, and C. Seard, "Variation in deglacial coralgal assemblages and their paleoenvironmental significance: IODP expedition 310, 'Tahiti sea level,'" *Global Planet. Change*, vol. 76, nos. 1–2, pp. 1–15, Mar. 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0921818110002523

RATNEEL DEO received the Master of Science degree in computing science from the University of the South Pacific, in 2018. He is currently pursuing the Ph.D. degree with The University of Sydney. His master's thesis entitled time series forecasting problems using deep learning with application to cyclone intensity and path prediction. His current research interests include optimization and machine learning for geophysical modeling and large-scale climatic modeling methods specifically targeted to coral reef analysis.

JODY M. WEBSTER received the Ph.D. degree from The University of Sydney, in 2000.

He spent several years in USA, carrying out his postdoctoral research on fossil coral reefs and carbonate platforms. He is currently an Expert of carbonate sedimentology and stratigraphy and interpretation of marine geology and geophysical data and the Co-Coordinator of the Geocoastal Research Group, School of Geosciences, The University of Sydney. He was the Co-Chief Scientist of the International Ocean Discovery Program (IODP) Expedition 325 Great Barrier Reef Environmental Changes and will also lead the upcoming IODP Expedition 389 (Hawaiian Drowned Reefs). His research interests include

understanding coral-reef and carbonate platform evolution and their implications for addressing fundamental problems in climate change and tectonics. His interests also involve linking observational and numerical modeling data to accurately show how reef and carbonate platform evolution in the Indo-Pacific are controlled by changes in sea level, subsidence, and growth rates.

TRISTAN SALLES received the Bachelor of Engineering degree from Ecole Centrale, France, in 2003, the Bachelor of Science degree (Hons.) in physical oceanography from the University of Aix Marseille, and the Ph.D. degree in geophysics from the University of Bordeaux, in 2006. He joined CSIRO, as a Postdoctoral Fellow, for two years, in 2007, to work on the wealth from ocean flagship before becoming a Research Scientist for six years with the Australian Resources Research Centre. He moved to The University of Sydney, in 2015, and since then he has been involved on a diverse range of projects related to Earth surface processes. He is currently involved with the Basin Genesis Hub Industrial Transformation Training Centre in collaboration with the University of Melbourne, ANU, Curtin, and Caltech, and partners from Geoscience Australia, OilSearch, Chevron, and Equinor. He is also the Chief Investigator in a linkage grant led by The University of Western Australia on the evolution of rift basins, including Monash University, Geological Survey of Western Australia, CSIRO, and BHP Billiton, Independence Group, and Anglo American. His research interests include sediment transport and sedimentary systems, geodynamic and landscape evolution, carbonate platforms, and ocean dynamics. His main activities consist in the design and implementation of open-source numerical codes that improve our understanding of the complex interactions between sedimentary systems, climatic forcing, and the physical processes that erode, transport, and deposit sediments. He has extensive experience in industry collaboration and translating academic research output into industrial applications.

ROHITASH CHANDRA (Senior Member, IEEE) is currently a Senior Lecturer of data science with the School of Mathematics and Statistics, UNSW Sydney. He leads a program of research encircling methodologies and applications of artificial intelligence, particularly in the areas of Bayesian deep learning, climate extremes, geoscientific models, remote sensing and mineral exploration, and large language models for media and philosophy of religion. His research has been funded by the Australian Research Council (ARC) and the National Health and Medical Research Council (NHMRC). Since 2021, he has been appearing in Stanford's list of top 2% scientists in the world. He is an Associate Editor of *Geoscientific Model Development*, *Neurocomputing*, and IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS (2021–2022). He is originally from Fiji (Nausori) and has a Girmit Indian heritage.