Asian Journal of Innovation and Policy (2025) 14.3:380-401
DOI: https://doi.org/10.7545/ajip.2025.14.3.380

Multiclass Classification of Coast in 3D LiDAR Point Clouds for Shoreline Erosion Monitoring

Sannareth Sou [*], Surasak Boonkla [**], Hongly Va [***], Jessada Karnjana [****]

Abstract Coastal erosion poses a serious threat to ecosystems, infrastructure, and coastal livelihoods. Traditional monitoring using satellite or aerial imagery often struggles to distinguish between forest, beach, and sea regions, particularly in complex coastal environments. This study aims to enhance shoreline monitoring accuracy by introducing a LiDAR-based multiclass classification framework that effectively differentiates these terrain types. The proposed method integrates geometric and color-based features extracted from UAV-based LIDAR data collected in Khlung, Chanthaburi, Thailand. Six features were derived from segmented point cloud grids: elevation variability (v1), directional slope variability (v2), and the mean RGB values (Rmean, Gmean, Bmean, RGBmean). Four machine learning models, Random Forest, Decision Tree, Support Vector Machine, and Logistic Regression, were evaluated using cross-validation on 1500 samples. Results indicate that the Random Forest classifier achieved the highest accuracy of 99.78%, outperforming other models. The proposed approach effectively separates forest, beach, and sea regions, reducing misclassification errors and improving the reliability of coastal terrain classification. This study demonstrates the potential of combining LiDAR point cloud geometry and radiometric features for robust multiclass classification. The findings support more precise and automated coastal erosion analysis, offering a foundation for future integration with deep learning and hybrid image point cloud systems.

Keywords Coastal erosion, 3D point clouds, multiclass classification, machine learning, Random Forest, LiDAR.

Submitted, December 27, 2025; Accepted, March 6, 2026
* Graduate student, CADT, Cambodia Academy of Digital Technology, Phnom Penh, Cambodia; sannareth.sou@student.cadt.edu.kh
** Researcher, CADT, Cambodia Academy of Digital Technology, Phnom Penh, Cambodia; hongly.va@cadt.edu.kh
*** Researcher, NECTEC, National Science and Technology Development Agency, Pathum Thani, Thailand; surasak.boonkla@nectec.or.th
**** Corresponding, Researcher, NECTEC, National Science and Technology Development Agency, Pathum Thani, Thailand; jessada.karnjana@nectec.or.th

I. Introduction

Coastal erosion is a significant issue that is getting worse and is a threat to ecosystems, infrastructure, and people all over the world. It leads to the loss of natural ecosystems, erosion of coastal vegetation, property damage, and severe socio-economic impact affecting local populations (Cui et al., 2021; El-Mahdy et al., 2022; Leatherman, 2018; Lin et al., 2019). Climate change, rising sea levels, and unpredictable weather conditions are worsening situation by accelerating an erosion of shorelines from land. As a result, the need for accurate, reliable, and consistent monitoring systems has become more critical compared to recently (Lv et al., 2024; Neumann et al., 2015; Depountis et al., 2023). Such techniques are required for effective coastal management, establishing risk-reduction strategies, and promoting long-term growth in vulnerable coastal areas.

LiDAR-derived point clouds provide detailed spatial information for detecting changes along the coast, but they often make mistakes when classifying features. It is particularly relevant in complex areas like sea, forest, and muddy beach areas, where changes in canopy and surface make interpretation more difficult. These three terrain types—forest, beach, and sea—are specifically targeted in this study because they are the primary sources of misclassification in coastal LiDAR data and contribute the largest errors in shoreline extraction. It is also hard to tell the difference between terrain classes when you only have elevation or intensity data because beach sand and shallow sea areas have similar spectral features. Such issues lead to wrong conclusions about how fast and how much the coast is changing.

For addressing this problem, the study presents a multiclass classification method for LiDAR-derived point cloud data that categorizes coastal areas into three groups: forest, beach, and sea. The suggested solution combines geometric and color-based elements to make surface categorization more accurate and reliable in difficult coastal areas. Specifically, geometric feature elevation variability (v1) and directional variability of surface slopes (v2) capture structural and morphological characteristics of terrain, reflecting differences in surface roughness and topographic variation. Adding additional features, such as color information from LiDAR RGB characteristics, to geometric variability measurements can improve classification accuracy. By using this technique, it can help us to understand coastal topography better and make it easier to build strong classification models that can tell the difference between forest, beach, and sea areas. It also makes it possible to maintain monitoring of coastal erosion. Meanwhile, radiometric is the science of measuring electromagnetic radiation; its features derived from LiDAR RGB attributes (Rmean, Gmean, Bmean, and RGBmean) provide spectral information that complements geometric description, allowing for more effective differentiation between visually similar terrain types. This mix of structural and spectral information lets the model

accurately separate areas that are overlapping or unclear, like muddy beaches and shallow waters. This makes coastal surface classification more reliable overall.

The integration greatly enhances the model's ability to assess complex coastal environments, especially in regions where geometric and spectral characteristics tend to overlap, such as muddy beaches, vegetated dunes, or shallow coastal waters. By incorporating both geometric and radiometric features, the model achieves a more comprehensive understanding of surface variation. As a result, it can correctly find and separate forest, beach, and sea regions before performing coastal change analysis. This preprocessing stage eliminates errors in the classification of sea, forest, and beach, which makes it easier to draw accurate lines on the coastline, improves coastal monitoring, and provides a more solid basis for long-term shoreline management and environmental assessment. This research is conducted within an international collaboration under the ASEAN IVO framework, with the goal of delivering critical and actionable insights to inform policy development and address challenges posed by coastal erosion.

II. Literature Review

Coastal erosion monitoring has typically relied on satellite imagery, aerial photography, and other remote sensing techniques to observe changes along the coastline and identify areas in which the shoreline is likely to weaken. These methods are good for making general observations, but they usually do not have the spatial resolution and elevation accuracy needed for a more detailed study of terrain. On the other hand, light detection and ranging (LiDAR) has grown into a strong alternative that can collect accurate latitude, longitude, and elevation values for dense point clouds in three dimensions (Kramer et al., 2021). This technique facilitates visualization of coastal topography and provides researchers with extensive information regarding both elevation and surface shape. Moreover, LiDAR is capable of detecting subtle variations in elevation that are difficult to observe with other remote sensing techniques, as it gathers millions of data points along coastlines. Using these precise measurements, you can identify and examine trends of erosion, sediment accumulation, and other variations in the land's topography that occur over time. The high-resolution three-dimensional maps that were created give a robust foundation for analysis and modeling of coastal processes, which provide valuable insights into the dynamics of coastal change, the impact of seas on terrain, and processes by which the land's form is altered over time. This knowledge is also very important for long-term coastal management because it helps scientists and

policymakers decide how to stop erosion, build infrastructure, and protect habitats based on facts. LiDAR technology helps protect coastal areas in the long term by making tracking of changing shorelines more accurate, consistent, and time-based (Bui & Pham, 2022; Cifuentes et al., 2014; Topouzelis et al., 2017).

For instance, Insung et al. (2024) utilized LiDAR technology to monitor beach environments by capturing coastal regions in a 3D point cloud format, enabling detailed terrain analysis. Their research presented a classification tree-based methodology that proficiently differentiates between sandy and rocky terrains utilizing two statistical attributes: altitude variance and gradient orientation. This technique accurately identified various types of terrain by analyzing an alteration in the shapes of coastal surfaces. The experimental results demonstrated that this approach achieved a classification accuracy exceeding 99.69%, highlighting the strong potential of feature-based LiDAR analysis for accurate coastal surface differentiation and monitoring applications. Obu et al. (2017) employed repeated LiDAR elevation measurements to analyze elevation and volume changes as well as coastline movement along the Yukon Coastal Plain and Herschel Island. By comparing results across different geomorphic units in both hinterland and coastal zones, this study revealed notable spatial variability in erosion dynamics. Results indicated that low-lying coasts (with elevations up to 10 meters) had uniform and consistent erosion, while higher backshore areas had more complicated and varied erosion patterns. This work demonstrated the capability of multi-temporal LiDAR analysis to capture detailed morphological changes and enhance understanding of coastal erosion processes across diverse terrain types.

Moreover, S. Quan utilized vessel-based LiDAR technology to collect high-resolution coastal topography data around Monterey Bay, California, enabling precise measurement of shoreline dynamics. The analysis revealed average erosion rates of 1.8 meters versus 0.1 meters in the southern bay and 0.5 meters versus 0.0 meters in the northern bay, indicating distinct spatial variations in coastal change intensity. The research showed that LiDAR systems mounted on boats are very good at finding small-scale erosion patterns and shoreline retreat. This gives researchers very accurate information about how coastal morphology changes over time and where it happens (Quan et al., 2012). Furthermore, S. Hui proposed an automatic coastline detection method combining multi-scale segmentation and multi-level inheritance classification to improve accuracy in complex coastal environments. The approach successfully extracted various coastal types from GF-2 satellite pictures of Jiaozhou Bay, including harbor-wharf, silt, pond, rocky, and sandy beaches. Compared with traditional and OBRGIE methods, it achieved higher precision and reliability, demonstrating strong potential for supporting refined coastline management (Hui et al., 2022).

Yastikli and Cetin (2016) introduced two automated point-based classification techniques designed to work directly on raw airborne LiDAR point clouds, thereby avoiding the traditional gridding process to maintain the original point density and elevation precision. The proposed methods use hierarchical rule-based classification applied at the individual point level. The first approach combines both spatial (height-related) and echo-based attributes such as return number and intensity, while the second approach relies solely on spatial characteristics. These methods were thoroughly tested using a high-density full waveform LiDAR dataset (16 points/m²) collected by a Riegl LSM-Q680i sensor over a topographically complex urban–forested region in Zekeriyaköy, Istanbul. The experimental results indicate that both methods achieve highly accurate and statistically consistent classifications of ground, building, and vegetation categories, surpassing conventional grid-based methods by preventing interpolation errors and preserving fine-scale geometric detail (Yastikli & Cetin, 2016).

Although these studies demonstrate significant progress in LiDAR-based coastal analysis, an important gap remains: existing works rarely address the challenge of simultaneously differentiating forest, beach, and sea surfaces within a unified classification framework. Most prior research focuses on either terrain separation, such as sandy vs. rocky, coastline extraction, or structural object classification, such as buildings and vegetation. However, in many real coastal settings, forest canopy, muddy beach surfaces, and shallow sea areas display overlapping geometric or spectral properties. This results in frequent misclassification when using elevation-only or intensity-based LiDAR features. No existing study integrates geometric variability with radiometric LiDAR RGB attributes to improve multi-class separation across these three critical terrain categories. Therefore, a robust method is needed to accurately classify forest, beach, and sea regions as a preprocessing step for reliable shoreline delineation and coastal erosion monitoring. This study aims to address that gap.

III. Methodology

The proposed approach focuses on classifying coastal regions from LiDAR point cloud data into three distinct surface categories: forest, beach, and sea. The overall framework is composed of two primary stages: feature extraction and classification (see Figure 1). The initial step is to break up the raw point cloud data into smaller, properly sized grid tiles so that surface analysis can be done in specific areas. For each grid, a set of geometric and radiometric features is calculated. Its characteristics illustrate the surface's height variations and light reflection. These features collectively represent the spatial and spectral

properties of coastal terrains. In the second stage, the extracted features are used to train a supervised machine learning model for multiclass classification, allowing each grid cell to be categorized according to its corresponding surface type. That technique makes it possible to identify the difference between forest, beach, and sea areas with high precision. It is an important step toward better coastal change detection and erosion analysis.

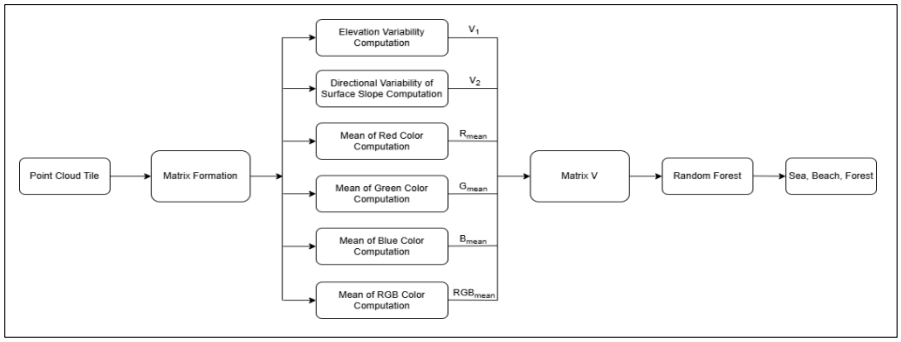

Figure 1. Proposed Framework

1. Feature Extraction

Let P be an n × 6 matrix that represents the LiDAR point cloud, incorporating both spatial and radiometric information for each point in the dataset. Each row of this matrix corresponds to a single LiDAR point, containing six attributes:

$$\boldsymbol{P} = \begin{bmatrix} x_1 & y_1 & z_1 & R_1 & G_1 & B_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & z_n & R_n & G_n & B_n \end{bmatrix} \quad (1)$$

where three spatial coordinates (x_i, y_i, z_i), which describe an exact three-dimensional location of a point in space, and three radiometric values (R_i, G_i, B_i), which represent the intensity of the red, green, and blue channels captured by the LiDAR sensor.

Spatial coordinates (x_i, y_i, z_i) provide detailed geometric information about the terrain surface, such as elevation and shape, while radiometric values (R_i, G_i, B_i) capture surface reflectance characteristics related to color and material properties. Together, these six parameters form a comprehensive representation of LiDAR data, combining both physical structure and visual attributes of the observed coastal landscape.

The spatial patch $\boldsymbol{P}$ is converted into a square grid matrix $\boldsymbol{Q}$ of size $d \times d$, in which each element $q_{j,k}$ represents the mean elevation of points falling within that grid cell:

$$\boldsymbol{Q} = \begin{bmatrix} q_{1,1} & q_{1,2} & \cdots & q_{1,d} \\ \vdots & \vdots & \ddots & \vdots \\ q_{d,1} & q_{d,2} & \cdots & q_{d,d} \end{bmatrix} \quad (2)$$

Each element $q_{j,k}$ in **Q** corresponds to the mean elevation of all LiDAR points contained within the $(j, k)^{\text{th}}$ grid cell.

$$q_{j,k} = \frac{1}{N_{j,k}} \sum_{i \in \Omega_{j,k}} z_i \quad (3)$$

$\Omega_{j,k}$ denotes the set of points in the (j, k)th grid cell, $N_{j,k}$ is the number of points within that cell, and z_i is the elevation value of a point i.

2. Geometric Feature

The geographical attributes of each surface section are defined by two geometric properties based on the grid matrix Q. The first feature, elevation variability, measures the extent of vertical change in altitude values within a grid. It shows the general roughness and topographical irregularity. The second feature, directional variability of surface slopes, looks at changes in slope directions measured in angles. This variable shows how the slopes of the land change from one cell to the next. Both of these elements collaboratively characterize both vertical and directional aspects of surface morphology. By using these geometric markers, the model can tell the difference between flat, smooth surfaces like sea, forest tops with highly irregular patterns, and beaches with somewhat uneven textures. Therefore, the combined use of elevation and slope-based variability provides a thorough and meaningful description of the terrain's structure, thereby enhancing the precision of tasks related to coastal landscape classification.

2.1 Elevation Variability

The elevation variability measures the distribution of surface height across the grid:

$$v_1 = \frac{1}{d^2 - 1}\sum_{j=1}^{d}\sum_{k=1}^{d}(q_{j,k} - \overline{q})^2 \quad (4)$$

where $\overline{q}$ is the mean elevation over the entire grid;

$$\overline{q} = \frac{1}{d^2}\sum_{l=1}^{d}\sum_{m=1}^{d} q_{l,m} \quad (5)$$

2.2 Directional Variability

The directional variability captures variation in slope orientation. The gradient direction $\theta_{j,k}$ is computed from elevation matrix $\boldsymbol{Q}$, and variability is expressed as:

$$v_2 = ln\left(\frac{1}{d^2 - 1}\sum_{j=1}^{d}\sum_{k=1}^{d}(\theta_{j,k} - \overline{\theta}_Q)^2\right) \quad (6)$$

where $\overline{\theta}_Q$ is the mean slope orientation;

$$\overline{\theta}_Q = \frac{1}{d^2}\sum_{l=1}^{d}\sum_{m=1}^{d}\theta_{l,m} \quad (7)$$

3. Radiometric Features

Additional radiometric features are added to the classification framework to enhance the ability to distinguish between distinct types of terrain that have similar geometric properties but differing spectral appearances, such as muddy beaches and shallow sea surfaces. RGB attributes associated with LiDAR point cloud data enable these radiometric properties. The mean intensity values of the red (R_{mean}), green (G_{mean}), and blue (B_{mean}) channels are determined by calculating an average intensity of each point within a grid cell. Subsequently, we calculate the mean of these three channel averages to obtain an overall mean reflectance value (RGB_{mean}), providing a comprehensive assessment of the surface's luminance and color balance.

These color-based descriptions function as crucial indications of surface reflectance and texture, effectively augmenting geometric parameters such as elevation and slope variability. By integrating these spectral characteristics, the

model gains improved capability to distinguish between terrain classes that may appear similar in elevation but exhibit distinct reflectance signatures. For instance, bright sandy beaches usually have higher mean RGB values than darker sea surfaces or shaded forest areas. Mathematical formulation for these features is presented as follows:

$$R_{mean} = \frac{1}{n}\sum_{i=1}^{n} R_i \quad (8)$$

$$G_{mean} = \frac{1}{n}\sum_{i=1}^{n} G_i \quad (9)$$

$$B_{mean} = \frac{1}{n}\sum_{i=1}^{n} B_i \quad (10)$$

and the overall mean intensity value is calculated as the average of the three-color components:

$$RGB_{mean} = \frac{R_{mean} + G_{mean} + B_{mean}}{3} \quad (11)$$

This composite feature provides an accurate estimate of surface reflectance by providing the overall brightness and color properties of each grid cell. We can combine spectral data with geometric features to improve the model's ability to distinguish between terrain types that have similar structural properties but different visual aspects. In particular, this feature improves discrimination between bright sandy beaches, darker ocean surfaces, and vegetated forest regions. As a result, it strengthens the classification framework by providing complementary radiometric information, leading to more reliable and accurate separation of land cover types in complex coastal environments.

4. Feature Vector Formation

The six extracted features are combined to construct the final feature vector:

$$v = [v_1, v_2, R_{mean}, G_{mean}, B_{mean}, RGB_{mean}]^T \quad (12)$$

Its feature vector combines both geometric descriptors (v_1, v_2) and radiometric attributes (R_{mean}, G_{mean}, B_{mean}, and RGB_{mean}) to give a complete and unified picture of each grid tile in point cloud dataset. Geometric features capture variations in surface elevation and slope orientation, effectively describing structural complexity and roughness of terrain. By contrast, radiometric features

characterize spectral reflectance properties derived from LiDAR RGB values, which convey essential information about surface color and brightness. Combining these two related feature types, the feature vector covers both the spatial structure and the spectral identity of each coastal surface segment. The integration significantly improves the model's ability to distinguish among three terrain categories—forest, beach, and sea—even in regions where visual or topographic similarities make classification challenging. Therefore, this multidimensional feature representation forms a robust foundation for accurate multiclass classification in complex coastal environments.

IV. Experiments and Results

1. Dataset description

A three-dimensional (3D) point cloud dataset was collected using an unmanned aerial vehicle (UAV) equipped with a DJI Zenmuse L2 LiDAR sensor over the coastal region of Khlung, Chanthaburi, Thailand. Data collection was conducted between October 2024 and September 2025, including multiple flights under various seasonal and environmental conditions, which were used to record temporal variations across the study area. Additionally, flights were performed at altitudes ranging from 1 to 300 meters to ensure extensive and precise geographic coverage of both coastal and inland areas. LiDAR system recorded precise x, y, and z coordinates along with return intensity and RGB color attributes, facilitating a thorough characterization of both terrain geometry and surface reflectance. All point cloud data were stored in LAS format (*.las), preserving the original 3D spatial and radiometric information. The dataset provides a solid basis for robust terrain modeling, feature extraction, and multiclass surface classification in coastal conditions.

2. Data Preprocessing

Point cloud data were spatially divided into uniform tiles measuring 5×5 meters (see Figure 2). The grid resolution was determined considering the spatial variability of LiDAR point density across the study area, which ranges from approximately 100 to 3000 points/m². Lower densities typically occur in sea regions, while higher densities are observed in vegetated areas. Using tiling results in systematic segmentation of the study area, allowing localized analysis of terrain characteristics. Each tile was subsequently subdivided into smaller sub-grids with resolutions ranging from 2×2 to 35×35 cells, facilitating multi-scale feature extraction and enabling the evaluation of how spatial granularity

influences classification accuracy. For each sub-grid, six descriptive features were calculated to represent both geometric and radiometric surface properties: elevation variability (v_1), which quantifies height fluctuations; directional variability of surface slopes (v_2), which characterizes changes in terrain orientation; and three-color intensity averages (R_{mean}, G_{mean}, and B_{mean}) derived from LiDAR's RGB attributes. The overall mean intensity (RGB_{mean}) was computed to capture the composite reflectance value. Together, these features provide a thorough representation of each grid's structural and spectral characteristics for classification.

RGB values were obtained from UAV-based LiDAR imagery captured under stable daylight conditions during a single flight mission, minimizing variations caused by illumination differences. Since LiDAR RGB attributes are typically stored as 16-bit integer values (ranging from 0 to 65,535), using their raw mean values can lead to disproportionately large feature magnitudes that negatively affect model performance and numerical stability. To get around this problem, RGB channels were rescaled to the standard 8-bit range (0–255). This made them more consistent with regular optical images and made it easier to understand color-based features. This normalizing step allows radiometric features to increase proportionally while training the model. Subsequently, all extracted geometric and radiometric features were standardized using Min–Max Scaler before being used as inputs for multiclass classification models.

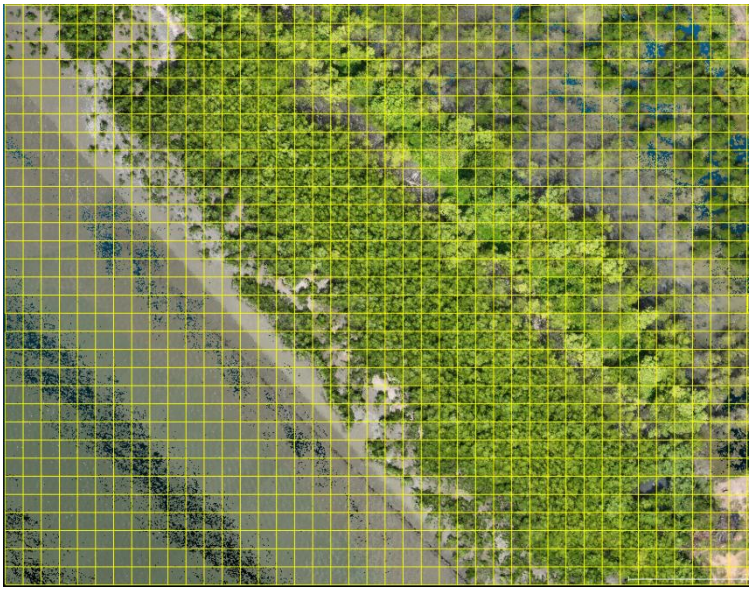

Figure 2. Uniform tiles of point cloud data (5x5 meters)

A total of 1,500 sample sizes were manually selected from the processed LiDAR dataset to provide a balanced and precise representation of the varied coastal environment in the study area. Such samples included 500 sea tiles (class 0), 500 forest tiles (class 1), and 500 beach tiles (class 2), each with unique surface features described by elevation, texture, and spectral reflectance. Additionally, a manual selection process was developed to cover a wide variety of environmental and topographic factors, ensuring that the dataset accurately represents both coastal and shoreline variations. To maintain experimental consistency, reproducibility, and avoid spatial autocorrelation between training and testing, the dataset was randomly shuffled using a fixed random seed before splitting. Samples were then divided into 70% for training and 30% for testing, enabling the models to generalize from diverse data while being evaluated on unseen instances. Moreover, a five-fold cross-validation procedure was applied to assess model reliability and robustness, minimizing overfitting and confirming consistent performance across multiple spatial partitions of LiDAR data.

3. Model Training and Evaluation

Multiple classical machine learning models were trained using six extracted geometric and radiometric features to perform a multiclass classification task.

The models included Decision Tree, Support Vector Machine (SVM), Logistic Regression, and Random Forest (Celik & Gazioglu, 2022). Each was selected for the proven effectiveness in handling structured data and feature-based learning. The training process included fine-tuning the model's parameters to get the best separation between three surface classes: forest, beach, and sea. To ensure a comprehensive assessment of model capability, performance was evaluated using four key metrics: precision, recall, F1-score, and overall accuracy (Yacouby & Axman, 2020). These evaluation measures provided a balanced view of classification reliability, allowing analysis of both the model's discriminative power and its consistency across classes.

Random Forest demonstrated the highest overall performance among all evaluated classifiers, achieving an outstanding accuracy of 99.78% at 30 × 30 grid resolution, as seen in Table 1; the Decision Tree classifier also did well, with an accuracy of 99.56%. This shows that tree-based models are very good at finding non-linear relationships between geometric and radiometric features that were extracted. In comparison, logistic regression achieved 98.67%, while support vector machine (SVM) recorded a slightly lower accuracy of 94.00%, indicating that linear models are less capable of handling complex feature interactions present in the dataset. These results clearly show that ensemble-based algorithms, especially Random Forest, are strong, flexible, and able to generalize well for multi-class point cloud classification tasks. By effectively integrating multiple decision trees, the Random Forest model reduces overfitting and enhances classification stability, making it a reliable tool for accurately differentiating forest, beach, and sea regions in challenging coastal environments.

Moreover, we conducted comparative experiments using multiple grid resolutions, ranging from 2×2 to 35×35, to evaluate the influence of spatial granularity on the performance of the Random Forest classifier (see Table 2). Results show that the classifier maintained consistently high accuracy across all tested resolutions, with precision, recall, and F1-score values above 97% in every case. At small grid resolutions of 2×2, a slight decrease in accuracy was observed because of the model's increased sensitivity to elevation noise and reduced feature stability caused by small grid sizes. Also, as the grid size increased, the performance of the classifier improved consistently, reaching its highest accuracy of 99.78% at 30×30 resolution. This improvement illustrates that larger grid cells help mitigate noise and capture more stable geometric patterns while preserving essential topographic characteristics. As a result, the model achieved superior differentiation between forest, beach, and sea regions, confirming the effectiveness of multi-scale feature extraction for coastal LiDAR classification.

Additionally, we conducted a comparative analysis of model performance using two different feature configurations: one with three features (v1, v2, and RGBmean) and another with six features (v1, v2, Rmean, Gmean, Bmean, and

RGBmean). The results clearly indicate that incorporating additional radiometric features substantially improves classification accuracy across all models. With only three features, Random Forest achieved 99.11% accuracy, already outperforming other classifiers. However, when extended to six features, it reached an exceptional 99.78% accuracy, precision, recall, and F1-score, demonstrating the superior capability of ensemble learning in capturing complex relationships between geometric and spectral attributes. Similarly, the Decision Tree classifier also benefited significantly from the inclusion of radiometric information, improving from 98.44% to 99.56% accuracy. Logistic Regression and SVM, although less flexible in handling nonlinear data, showed notable improvements as well, increasing from 94.89% to 98.67% and 78.89% to 94.00%, respectively. These results confirm that combining geometric variability (v1, v2) with detailed color-based features enhances class separability, particularly between visually similar regions such as beaches and shallow seas. Adding spectral information from RGB channels to geometric descriptors helps models better understand small differences in surfaces that topography alone can't show. Overall, the results validate that using both geometric and radiometric features yields a more comprehensive feature space, significantly boosting classification performance and ensuring greater reliability in coastal terrain differentiation (see Table 3).

Table 1. Model comparison at the best grid size (30x30).

Model	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
Logistic Regression	98.72	98.67	98.67	98.67
SVM	94.63	94.00	93.97	94.00
Decision Tree	99.56	99.56	99.56	99.56
Random Forest	**99.78**	**99.78**	**99.78**	**99.78**

Table 2. Random Forest performance under varying grid size from 2x2 to 35x35.

Grid size	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
2x2	97.28	97.11	97.11	97.11
5x5	98.48	98.44	98.44	98.44
11x11	98.03	98.00	98.00	98.00
15x15	98.48	98.44	98.44	98.44
19x19	98.67	98.67	98.67	98.67
20x20	98.67	98.67	98.67	98.67
27x27	99.56	99.56	99.56	99.56
30x30	**99.78**	**99.78**	**99.78**	**99.78**

Table 3. Comparison of each model with 3 features vs 6 features at the best grid size.

Model	Precision (%)		Recall (%)		F1 score (%)		Accuracy (%)	
	3 features	6 features	3 features	6 features	3 features	6 features	3 features	6 features
Logistic Regression	95.01	**98.72**	94.89	**98.67**	94.91	**98.67**	94.89	**98.67**
SVM	81.96	**94.63**	78.89	**94.00**	78.58	**93.97**	78.89	**94.00**
Decision Tree	98.45	**99.56**	98.44	**99.56**	98.44	**99.56**	98.44	**99.56**
Random Forest	99.11	**99.78**	99.11	**99.78**	99.11	**99.78**	99.11	**99.78**

*3 features: v_1, v_2, RGB_{mean}
*6 features: v_1, v_2, R_{mean}, G_{mean}, B_{mean}, RBG_{mean}

Pair plot (see Figure 3) shows the relationships among six extracted features: v_1, v_2, R_{mean}, G_{mean}, $Bmean$, and RGB_{mean} for three distinct coastal classes: sea (blue), forest (green), and beach (orange). The visualization clearly separates these classes across multiple feature dimensions. Sea class has a lot of variation in elevation ($v_1 \approx 20$) and not much variation in slope ($v_2 \approx -2$). The brightness is very low ($RGB_{mean} \approx 60$), which means water surfaces are darker. Forest class exhibits moderate elevation values and peaks at high green intensity ($G_{mean} \approx 150$) with low blue reflectance ($B_{mean} \approx 70$), reflecting dense vegetation. Beach class, by contrast, shows the highest overall brightness ($RGB_{mean} \approx 190$) with strong red and blue components, typical of sandy surfaces. Diagonal histograms show that there is very little overlap between classes, and the scatterplots show clear clusters, especially when v_1 is combined with radiometric features, and G_{mean} is compared to B_{mean}. These observations confirm that the combined use of geometric and color-based features yields clear linear separability among three terrain types, enabling even simple classifiers such as decision trees or linear SVMs to achieve accuracy exceeding 99%.

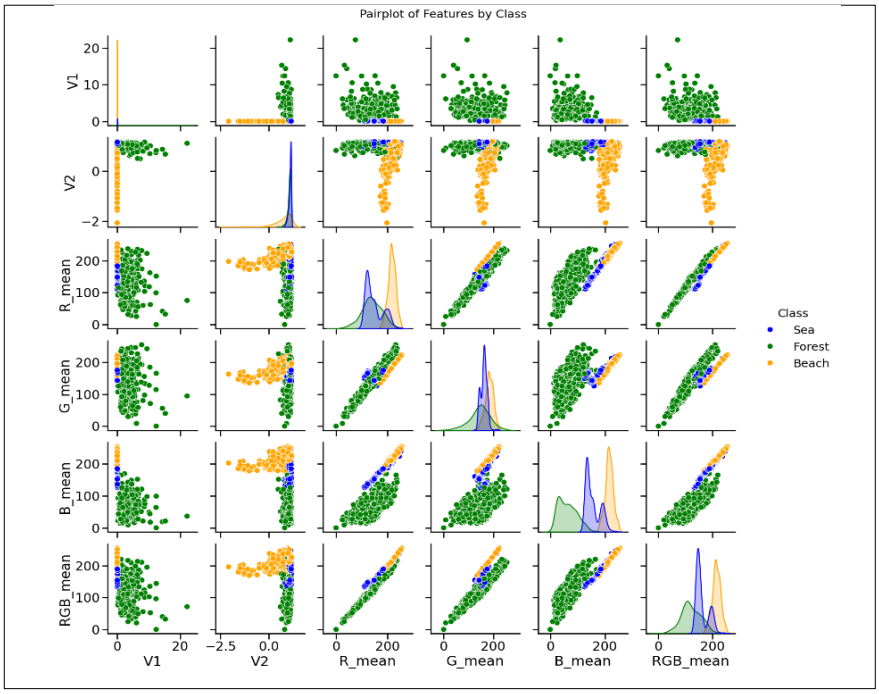

Figure 3. Pairwise distribution of features (v_1, v_2, R_{mean}, G_{mean}, B_{mean}, RGB_{mean}) by class.

Last but not least, we also conducted a comparison among four classical machine learning classifiers—Decision Tree, Logistic Regression, Support Vector Machine (SVM), and Random Forest—to evaluate their effectiveness in multiclass coastal surface classification. Each model was tested using 5-fold cross-validation, a robust statistical validation technique designed to ensure reliable and unbiased performance evaluation across different data partitions (Pal & Patel, 2020). By combining the results from several training and testing cycles, this approach mitigates overfitting and provides a more comprehensive understanding of the model's performance. Through this systematic comparison, an investigation was conducted to find the best technique for properly differentiating forests, beaches, and sea regions using geometric and radiometric properties taken from LiDAR point cloud data.

Results show that the Random Forest model had the highest and most stable accuracy of 98.9%. The decision Tree model came in second with 98.8%, and the Logistic Regression and SVM models came in third with 97.9% and 93.9%, respectively (see Figure 4). This demonstrates that ensemble-based methods, such as Random Forest, are more robust and generalize better for this dataset.

Moreover, the bar chart also confirms the trend observed in this feature pair plot, showing that three land-cover classes—sea, forest, and beach—are almost perfectly separable. Geometric features like v_1 that clearly separate water surfaces and radiometric features (RGB channels) that clearly separate forested and sandy areas made it possible for even simple models to get almost perfect classification accuracy (over 99%). This highlights the high discriminative power of the proposed features and overall simplicity of the classification task when using the combined geometric and color-based feature set.

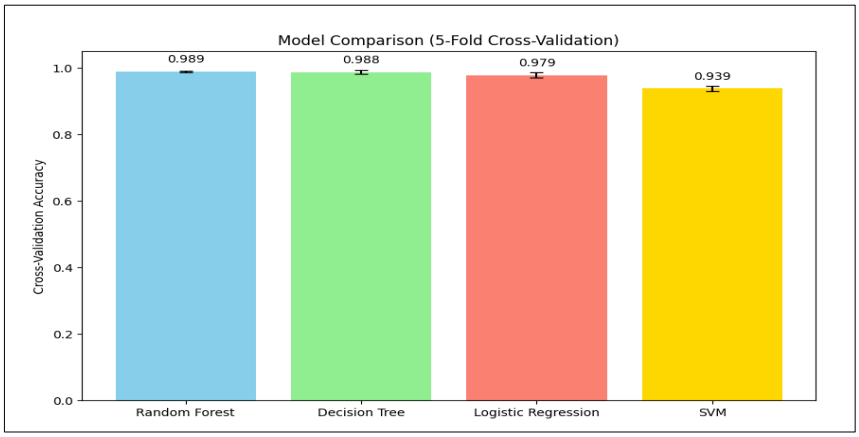

Figure 4. Model comparison using 5-fold cross-validation

The confusion matrix measures the classification performance of the Random Forest model across three classes: sea, forest, and beach (Grandini et al., 2020). The model is almost perfect, correctly classifying almost every sample with only one mistake. It accurately predicted 149 sea, 150 forest, and 150 beach samples out of 450 total test samples (150 per class), with an overall accuracy of 99.78% (see Figure 5). Only error occurred when one sea sample was incorrectly labeled as beach, likely due to the presence of a bright water pixel resembling sand in its RGB intensity values. This transitional zone presents inherent ambiguity in LiDAR-based classification and represents a limitation of the current feature set. Moreover, an almost perfect diagonal pattern in the confusion matrix illustrates the high discriminative ability of these retrieved features and the exceptional reliability of the Random Forest classifier. Results are consistent with previous analyses, like cross-validation results and pair plots, which indicate that the combined geometric and radiometric features allow for highly successful separation of sea, forest, and beach regions with minimum confusion or overlap.

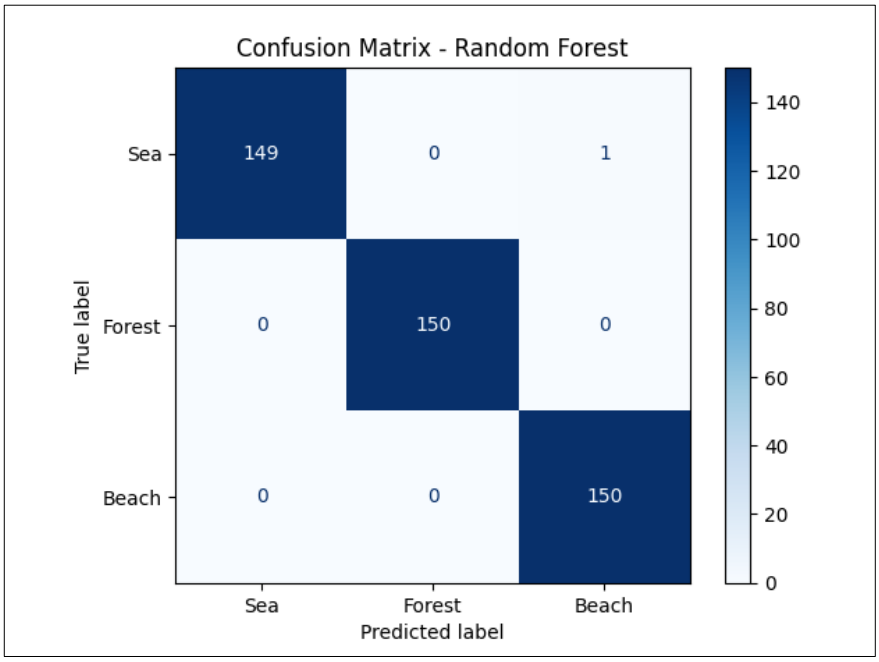

Figure 5. Confusion Matrix of Random Forest

V. Discussion

The results obtained in this study align closely with the outcomes of pair plot visualization and cross-validation analyses, both of which demonstrate that integration of geometric and radiometric features enables a clear and accurate separation between sea, forest, and beach regions. This strong separability confirms the discriminative power of extracted features and validates the effectiveness of the proposed feature engineering strategy. The experimental results reveal that the proposed method successfully distinguishes among three terrain types in LiDAR-derived point cloud data, substantially reducing overall misclassification rates. Among the tested models, the Random Forest classifier achieved the best overall performance, outperforming the Decision Tree, Support Vector Machine (SVM), and Logistic Regression models, with perfect accuracy attained at a 30×30 grid resolution. This superior performance highlights the robustness of Random Forest algorithm, which benefits from ensemble learning and random feature selection, making it particularly effective for complex coastal datasets characterized by high variability in elevation and reflectance values.

The steady improvement in model accuracy observed across larger grid resolutions reinforces the importance of selecting an optimal spatial scale for feature computation. Smaller grids are likely to capture excessive noise and fine-scale irregularities, resulting in unpredictability in pattern structures. In contrast, larger grids smooth out such variations while preserving essential surface characteristics, thereby enhancing the reliability of classification. This trend demonstrates that the Random Forest model is highly capable of handling multi-scale spatial data, adapting effectively to both micro- and macro-level terrain differences.

Furthermore, the study emphasizes the value of combining geometric variability features (v_1, v_2) with radiometric descriptors (R_{mean}, G_{mean}, B_{mean}, and RGB_{mean}) for improving class separability, particularly between spectrally similar areas such as beach and shallow sea zones. The fusion of geometric and color-based information provides a more holistic representation of coastal surfaces, allowing the model to leverage both shape and reflectance cues for classification. This hybrid approach captures complex non-linear relationships within data, enhancing the model's discriminative capability and resilience to environmental variations.

VI. Conclusion

This study introduces a multiclass classification framework aimed at improving the interpretation and analysis of LiDAR point cloud data for coastal erosion monitoring by effectively distinguishing forest, beach, and sea regions. By integrating geometric and radiometric features, the proposed method improves the precision of shoreline detection compared to unclassified point clouds. This facilitates the reliable calculation of shoreline change and coastal erosion rates over time. Experimental results confirm that the Random Forest model delivers the best overall performance, achieving an outstanding accuracy of 99.78%, surpassing other classical machine learning models such as Decision Tree, Logistic Regression, and SVM. These results demonstrate the method's reliability and precision in coastal surface classification, supporting more accurate shoreline delineation and erosion analysis. The proposed framework also lays a strong foundation for future research, such as adding deep learning architectures and combining data from many sources to make coastal monitoring scalable, automated, and real-time for long-term shoreline management.

Acknowledgment

The ASEAN IVO project (https://www.nict.go.jp/en/asean_ivo/index.html), titled "Coastal Erosion Monitoring Platform Based on Wireless Sensor Networks and 3D Point Clouds from Airborne LiDAR," was involved in the production of the contents of this work and was financially supported by NICT (http://www.nict.go.jp/en/index.html). This study is financially supported by the Faculty of Informatics, Burapha University.

References

Bui, L.T., & Pham, H.T. (2022). Coastal erosion trend analysis using a combination of remote sensing and hydrodynamic models: Case study of ca mau cape, mekong delta. Remote Sensing Applications: Society and Environment, 26, 100734. https://doi.org/https://doi.org/10.1016/j. rsase.2022.100734

Celik, O.I., & Gazioglu, C. (2022). Coast type-based accuracy assessment for coastline extraction from satellite image with machine learning classifiers. The Egyptian Journal of Remote Sensing and Space Science, 25 (1), 289–299. https://doi.org/https://doi.org/10.1016/j.ejrs.2022. 01.010

Cifuentes, R., Van der Zande, D., Salas, C., Farifteh, J., & Coppin, P. (2014). Correction of erroneous lidar measurements in artificial forest canopy experimental setups. Forests, 5 (7), 1565–1583.

Cui, B., Jing, W., Huang, L., Li, Z., & Lu, Y. (2021). Sanet: A sea–land segmentation network via adaptive multiscale feature learning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 116–126. https://doi.org/10.1109/JSTARS.2020.3040176

Depountis, N., Apostolopoulos, D., Boumpoulis, V., Christodoulou, D., Dimas, A., Fakiris, E., Leftheriotis, G., Menegatos, A., Nikolakopoulos, K., Papatheodorou, G., et al. (2023). Coastal erosion identification and monitoring in the patras gulf (greece) using multi-discipline approaches. Journal of Marine Science and Engineering, 11 (3), 654.

El-Mahdy, M.E.-S., Saber, A., Moursy, F.E., Sharaky, A., & Saleh, N. (2022). Coastal erosion risk assessment and applied mitigation measures at ezbet elborg village, egyptian delta. Ain Shams Engineering Journal, 13 (3), 101621.

Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for multi-class classification: An overview. arXiv preprint arXiv:2008.05756.

Hui, S., Mengliang, G., Yuliang, G., Mingming, X., Shanwei, L., Yasir, M., Jianyong, C., & Jianhua, W. (2022). Coastline extraction based on multi-scale segmentation and multi-level inheritance classification. Frontiers in Marine Science, 9, 1031417.

Insung, P., Supasri, P., Limsuwat, P., Thitathanapat, S., Takhom, A., Galajit, K., Boonkla, S., & Karnjana, J. (2024). Utilizing a classification tree for rocky terrain detection in 3d point clouds for coastal erosion studies. 2024 SICE Festival with Annual Conference (SICE FES), 852–857.

Kramer, H., M¨ucher, S., & van der Hagen, H. (2021). Hotspot vegetation structure and terrain monitoring of dutch coastal dunes with lidar and optical camera's mounted on drones. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 739–742.

Leatherman, S.P. (2018). Coastal erosion and the United States national flood insurance program. Ocean & coastal management, 156, 35–42.

Lin, Y.-C., Cheng, Y.-T., Zhou, T., Ravi, R., Hasheminasab, S.M., Flatt, J. E., Troy, C., & Habib, A. (2019). Evaluation of uav lidar for mapping coastal environments. Remote Sensing, 11 (24). https://doi.org/10.3390/rs11242893

Lv, Q., Wang, Q., Song, X., Ge, B., Guan, H., Lu, T., & Tao, Z. (2024). Research on coastline extraction and dynamic change from remote sensing images based on deep learning. Frontiers in Environmental Science, 12, 1443512.

Neumann, B., Vafeidis, A.T., Zimmermann, J., & Nicholls, R. J. (2015). Future coastal population growth and exposure to sea-level rise and coastal flooding-a global assessment. PloS one, 10 (3), e0118571.

Obu, J., Lantuit, H., Grosse, G., G¨unther, F., Sachs, T., Helm, V., & Fritz, M. (2017). Coastal erosion and mass wasting along the canadian beaufort sea based on annual airborne lidar elevation data [Permafrost and periglacial research from coasts to mountains]. Geomorphology, 293, 331–346. https://doi.org/ https://doi.org/10. 1016/j.geomorph.2016.02.014

Pal, K., & Patel, B.V. (2020). Data classification with k-fold cross validation and holdout accuracy estimation methods with 5 different machine learning techniques. 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 83–87. https://doi.org/10.1109/ICCMC48092.2020. ICCMC-00016

Quan, S., Kvitek, R.G., Smith, D.P., & Griggs, G.B. (2012). Using vessel-based lidar to quantify coastal erosion during el ni˜no and inter-el ni˜no periods in monterey bay, california. Journal of Coastal Research, 29 (3), 555–565. https://doi.org/10. 2112/JCOASTRES-D-12-00005.1

Topouzelis, K., Papakonstantinou, A., & Doukari, M. (2017). Coastline change detection using unmanned aerial vehicles and image processing technique. Fresen. Environ. Bull, 26, 5564–5571.

Yacouby, R., & Axman, D. (2020). Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models. Proceedings of the first workshop on evaluation and comparison of NLP systems, 79–91.

Yastikli, N., & Cetin, Z. (2016). Classification of lidar data with point based classification methods. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 41, 441–445.