remote sensing

MDPI

Article

Few-Shot Classification of Shallow-Water Seabed Sediment and Benthic Cover by Fusing Airborne LiDAR Bathymetry and Multispectral Imagery

Shuohao Chen [1], Xueshan Song [2], Jinfeng Mao [3], Yu Huang [3], Anxiu Yang [1,4,*], Rui Shan [5], Han Gao [1] and Dianpeng Su [1,4,6]

[1] College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China; 202311030701@sdust.edu.cn (S.C.); g_han@sdust.edu.cn (H.G.); sudianpeng@sdust.edu.cn (D.S.)
[2] China Water Resources Beifang Investigation, Design and Research Co., Ltd., Tianjin 300222, China; song_xs@bidr.com.cn
[3] Hydrological Bureau of the Yangtze River Water Resources Commission Hydrological and Water Resources Survey Bureau of the Lower Yangtze River, Nanjing 210011, China; mjf@sdust.edu.cn (J.M.); 202182020019@sdust.edu.cn (Y.H.)
[4] State Key Laboratory of Disaster Prevention and Ecology Protection in Open-Pit Coal Mines, Shandong University of Science and Technology, Qingdao 266590, China
[5] Qingdao Institute of Marine Geology, China Geological Survey, Qingdao 266071, China; shanrui@mail.cgs.gov.cn
[6] Key Laboratory of Ocean Geomatics, Ministry of Natural Resources, Qingdao 266590, China
* Correspondence: skd996737@sdust.edu.cn

Highlights

What are the main findings?

- A robust cross-modal alignment between airborne LiDAR bathymetry and multispectral imagery is achieved using SIFT-PROSAC and perspective transformation, ensuring high geometric consistency.
- The proposed FCA-Relief-F feature selection and GAT-PN model enable accurate classification of five sediment and benthic-cover types under few-shot conditions, significantly improving discriminative performance.

What are the implications of the main findings?

- The study demonstrates that effective feature fusion and selection can substantially reduce redundancy and enhance cross-modal representation for shallow-water seabed mapping.
- The proposed framework provides a practical solution for sediment and benthic-cover classification in data-scarce scenarios, with strong potential for coastal applications.

Abstract

The accurate classification of seabed sediment and benthic covers in shallow-water environments remains a key challenge for marine activities and oceanographic research. However, coastal areas of shallow waters are influenced by complex dynamic environments, making it difficult to obtain authentic sediment and benthic-cover samples. Therefore, to address the problem of few-shot classification of seabed sediment and benthic covers, a few-shot classification algorithm of seabed sediment and benthic covers based on the fusion model of airborne LiDAR bathymetry (ALB) and multispectral images is proposed in this article. Based on the extracted features, a scale-invariant feature transform-progressive sample consensus (SIFT-PROSAC) algorithm and perspective transform model were constructed to achieve feature fusion. Then, multi-modal feature selection is realized using a formal

Check for updates

Academic Editor: Francesco Serafino

Received: 20 April 2026
Revised: 13 June 2026
Accepted: 25 June 2026
Published: 1 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

concept analysis-Relief-F (FCA-Relief-F) algorithm. Finally, a graph attention network-prototype network (GAT-PN) model was established to classify five types of sediment and benthic cover (coral reef, stone, sand, vegetation, and coastal zone). To validate the effectiveness of the proposed method, experimental data from actual measurements at Ganquan Island in the Xisha Islands of China were used. Compared to other classical classifiers, the GAT-PN algorithm achieves a higher classification accuracy, with an overall accuracy (OA) and Kappa coefficient of 97.50% and 0.97, respectively. The findings of this study provide effective technical support for marine engineering and related fields.

Keywords: airborne lidar bathymetry; multispectral image; few-shot; feature fusion; sediment and benthic-cover classification

1. Introduction

The seabed sediment and benthic-cover category is an important marine environmental parameter that serves as a foundation for seabed scientific research. With the development of marine pasture construction, marine environment monitoring, and seabed resource surveys, the need for seabed sediment and benthic-cover classification has become increasingly pressing [1]. Verification of the truth data is required for the classification of the seabed. Due to various collected restrictions on sampling in the field, sample data with actual labels are difficult to obtain [2]. In addition, the application of few-shot classification in traditional deep learning classifiers often leads to overfitting, which is an important factor that restricts the development of seabed sediment and benthic-cover classification. Therefore, the exploration of a reasonable and effective classification method for few-shot classification of seabed sediment and benthic covers is of significant value for marine environmental protection activities and marine sciences.

Oceanographic surveys are typically based on acoustic technologies such as the Multi-Beam Echo Sounder (MBES), which is used for depth detection purposes [3]. Although these technologies are relatively mature, there are still some limitations, such as low efficiency in shallow waters and the risk of running aground [4]. Airborne lidar bathymetry (ALB) technology is characterized by high measurement accuracy, density of measurement points, and strong mobility, making it especially suitable for rapid water depth detection in shallow water [5]. In addition, the blue and green bands of multispectral remote sensing have strong penetration capabilities for clear water. Therefore, multispectral remote sensing has been widely applied to oceanic measurements [6,7]. Because of the lack of spectral and texture information, although ALB technology can directly obtain a wealth of original waveform and seabed terrain information, it has some deficiencies in seabed sediment and benthic-cover classification [8]. Multispectral images contain rich spectral information that can accurately describe the spectral and textural features of objects. However, the phenomenon of the same subject with different spectra or different subjects with the same spectra usually occurs in the context of seabed sediment and benthic-cover classification [9]. Therefore, integrating multi-source data that reflect seabed sediment and benthic-cover properties can provide complementary feature information, which is of significant research value for shallow-water sediment and benthic-cover classification. Owing to measurement errors, large discrepancies are introduced when only the position information is used for fusion. Registration before fusion can improve the accuracy of the fusion results. Currently, from the perspective of registration primitives, common LiDAR and optical image registration can be divided into grayscale primitives and feature primitives. Umeda et al. proposed a grayscale-based original method, which is simple to operate but requires a large amount of computation and is sensitive to noise and lighting changes [10]. A feature-based method

is used to achieve registration by identifying the corresponding geometric information, such as contour lines and intersection points. Among them, the Scale-Invariant Feature Transform (SIFT) algorithm is widely applied owing to its rotation and scale invariance as well as its robustness to brightness variations and noise [11]. However, a situation in which feature point pairs do not match still exists. Based on the sift-random sample consensus (RANSAC) algorithm, Kaur et al. removed incorrect matching point pairs and realized accurate registration. However, the RANSAC method is confronted with a trade-off between efficiency and optimal solutions in parameter selection [12]. Currently, most studies on the registration of ALB point cloud and satellite images have concentrated on land areas and are notably sparse in shallow water. Accordingly, a coarse and fine registration method for shallow water ALB point clouds and satellite image data must be identified with the aim of optimizing the accuracy of the registrations and the precision of the fused data.

Owing to the strong correlation between multispectral bands, high redundancy, strong correlation, and low signal-to-noise ratio were introduced in the extracted feature vectors. The removal of redundant interference features can be achieved through feature optimization, which effectively reduces the dimensionality of high-dimensional data and improves the accuracy of classification [13]. Le et al. used the Relief-F algorithm for feature selection, where features are arranged in descending order according to their contribution rate to achieve optimization [14]. However, correlations between features still exist after dimensionality reduction, and redundant feature information cannot be removed. Su et al. used the Principal Component Analysis (PCA) algorithm to reduce the ALB waveform and terrain features to irrelevant principal component vectors [15]. However, if feature preprocessing is not conducted prior to PCA dimensionality reduction, the computational efficiency would be diminished. Consequently, finding a suitable feature selection method is of great importance for the classification of seabed sediment and benthic covers. There are three types of few-shot classification methods: model fine-tuning- based, data enhancement-based, and transfer learning-based. The method based on transfer learning is a relatively cutting-edge method at present, which can be subdivided into three types of methods: metric-based learning, meta-learning, and graph-based neural networks [16]. The metric learning-based method classifies unknown samples by calculating their distance. However, its effectiveness is limited when only a few samples are available [17]. The meta-learning method is designed to use tasks as training units, allowing the model to acquire learning capabilities. However, improvements in the few-shot classification accuracy are limited by the high complexity of this method and the lack of consideration of the relationships between samples [18]. It has been proven that few-shot learning algorithms must make full use of the relationship between the support set and query set [19–21]. Feature aggregation from neighbors is iteratively performed through a message-passing mechanism in graph neural networks [22], allowing complex interactions between data instances to be expressed, making them a hot topic for few-shot learning. Therefore, combining the advantages of the above algorithms may improve the accuracy of the few-shot classification method.

To address the issues inherent to the few-shot classification of seabed sediment and benthic covers, a fusion algorithm combining ALB with a multispectral image and a classifier for seabed sediment and benthic covers was proposed. Multi-source features can be fully explored and accurately fused by combining the SIFT-Progressive Sample Consensus (PROSAC) algorithm and the perspective transform model. Then, the FCA-Relief-F algorithm was applied to analyze and explore the correlation and contribution rate between multi-source characteristics and seabed sediment and benthic-cover categories. Eventually, the graph attention network-prototype network (GAT-PN) shallow water few-shot sediment and benthic-cover classification model was constructed to divide the

sediment and benthic covers into five categories, and the accuracy of the classification results was evaluated and analyzed.

2. Materials and Methods

The workflow of this study consists of three sequential stages. First, cross-modal multi-source features are extracted from ALB and multispectral data, including spectral, texture, waveform, and terrain features. Second, the extracted features are fused and optimized through feature selection to remove redundancy and retain the most informative feature subset. Finally, the optimized fusion features are used as input to the proposed few-shot seabed sediment and benthic-cover classification model. The overall framework of the proposed method is illustrated in Figure 1.

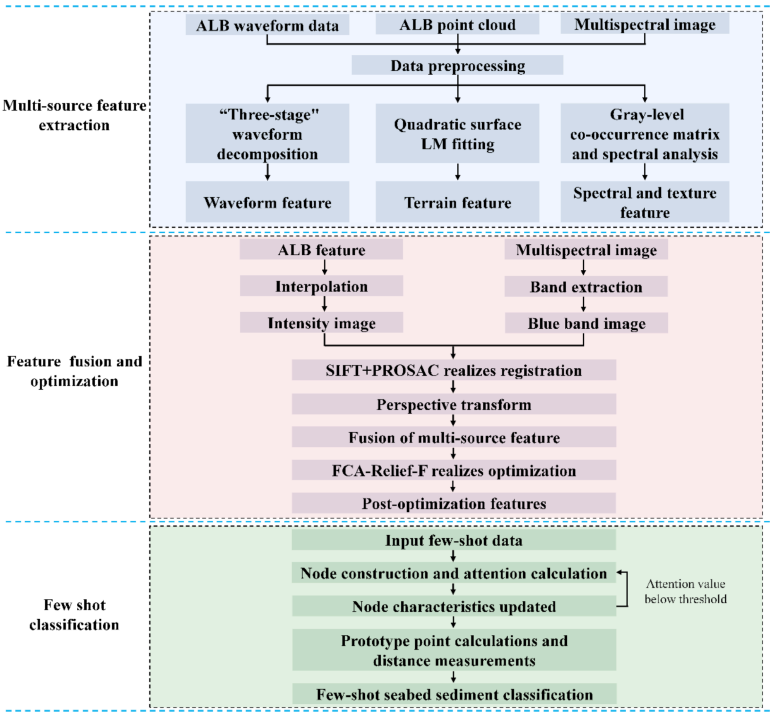


Figure 1. Algorithm flowchart.

2.1. Data Preprocessing and Feature Extraction

The purpose of ALB data pre-processing is to provide more complete and accurate detailed information on each component of the waveform echo signal. Waveform data preprocessing includes two main parts: waveform denoising and waveform fitting [23]. The wavelet adaptive threshold denoising method was used to denoise the ALB waveform, and the water surface, water body, and bottom waveforms were decomposed based on a hierarchical heterogeneous algorithm [24,25]. Specifically, the "three-stage" waveform decomposition method involves fitting the water surface echo with a Gaussian function, water body echo with a double exponential function, and bottom echo with a Weibull function. Furthermore, the processes of radiation calibration and atmospheric correction are essential for eliminating sensor-induced errors, reducing the impact of atmospheric molecules and aerosols on target objects, obtaining true reflectivity, and obtaining more accurate multispectral images.

After data preprocessing, based on the "three-stage" waveform decomposition method mentioned previously, multivariate waveform characteristics of ALB were extracted, including amplitude, wave width, half-width wave width, waveform area, skewness, kurtosis, backscattering cross-sectional coefficient, residual amplitude, and residual wave width.

In addition, a quadratic surface LM fitting model was constructed to extract multivariate terrain features from ALB, such as the slope, Gaussian curvature, roughness, water depth standard deviation, elevation entropy, undulation, mean deviation, coefficient of variation, and concavity coefficient. In addition, the texture and spectral features of the four multispectral bands were extracted using a gray-level co-occurrence matrix and spectral analysis methods. The extracted texture features included mean, variance, homogeneity, contrast, dissimilarity, information entropy, second moment, and correlation. The extracted spectral features included pixel values of the blue, green, red, and near-infrared bands, and band-ratio features derived from pairwise combinations of these bands (e.g., blue/green, blue/red, green/red, and near-infrared-related ratios), as well as the normalized difference vegetation index (NDVI) and normalized difference water index (NDWI).

2.2. Muti-Source Features Registration and Fusion Model Construction

To fully explore the characteristic information of the seabed and build a detailed data foundation, a high-efficiency and high-precision SIFT-PROSAC algorithm combined with the perspective transformation method was used to realize the registration and fusion of ALB and multispectral data. The specific steps are as follows:

(1) Scale space construction and feature point detection. Image locations were sought at all scales, with feature points unresponsive to scale and rotation identified using Gaussian derivative functions. The location and scale of the feature points are then determined through the application of a fitted model [26].

(2) Feature point main orientation estimation and descriptor creation. The main direction is determined by calculating the gradient direction of each extreme point in the image. Gradient calculation and direction assignment were performed on the pixels around the extreme points to generate the feature descriptor of each extreme point. The gradient direction and modulus were calculated using the following equations [27]:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (1)$$

$$\theta(x,y) = \tan^{-1}(L(x,y+1) - L(x,y-1))/(L(x+1,y) - L(x-1,y)) \quad (2)$$

where $L(x,y)$ denotes the feature point, m denotes the gradient modulus, and θ denotes the gradient direction.

(3) Erroneous feature point-pair elimination. For the matching points obtained after the SIFT transformation, all matching points were sorted in descending order according to the Hamming distance similarity. The first n groups of high-quality matching point pairs were selected, and m groups of matching point pairs were chosen to form a sample set. Then, the basic matrix F is obtained according to Equation (3) [28,29]. The inliers are defined as the number of matching point pairs that satisfy the matrix F. When the number of inliers exceeds the set maximum threshold or the number of inliers after two consecutive samplings does not increase, the above cycle is stopped, and the corresponding inliers are output.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = F \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3)$$

where $(x, y, 1)$ and $(x', y', 1)$ are the homogeneous coordinates of a pair of correctly matched points $a(x, y)$ and $a'(x', y')$, respectively.

(4) Perspective transformation matrix calculation and multi-source feature fusion. To address the geometric discrepancies between ALB intensity range images and multi-

spectral blue-band range images caused by different imaging geometries and viewing perspectives, the perspective transformation model was adopted. Compared with affine transformation, the perspective model can better handle non-linear projective distortions. Therefore, the transformation matrix for pixel coordinate conversion between the two datasets was calculated based on the perspective transformation model. Using this matrix, ALB features were then aligned and fused with multispectral features. The perspective transformation model is expressed as follows [30]:

$$\begin{bmatrix} q' \\ w' \\ 1 \end{bmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{bmatrix} q \\ w \\ 1 \end{bmatrix} \tag{4}$$

where q and ω are the coordinate points in the original multispectral image, (q', ω') are the corresponding two-dimensional coordinates after the perspective transformation, and $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ is the perspective transformation parameter matrix, which can be divided into four parts. a_{11}, a_{12} and a_{13} are the parameters that control the scaling, rotation, and translation in the horizontal direction; a_{21}, a_{22} and a_{23} are the same in the vertical direction; a_{31}, a_{32} and a_{33} are the parameters of the perspective projection; a_{31} and a_{32} determine the nonlinear scaling of the q and ω coordinates, and a_{33} is the scale factor.

2.3. Cross-Modal Feature Optimization Model Construction

To address the problem of feature redundancy in the few-shot classification of seabed sediment and benthic covers, a feature optimization model based on the FCA-Relief-F algorithm was proposed. The dimensions of high-dimensional data can be effectively reduced through this model, thereby improving classification accuracy. The specific contents of this paper are as follows:

First, based on the FCA algorithm, the correlation coefficient between two features was calculated [31]. The calculation equations are shown in Equations (5) and (6):

$$S_{ou} = \frac{1}{K-1} \sum_{L=1}^{K} (x_o^L - \frac{1}{K} \sum_{L=1}^{K} x_o^L)(x_u^L - \frac{1}{K} \sum_{L=1}^{K} x_u^L) \tag{5}$$

$$r_{ou} = \frac{S_{ou}}{\sqrt{S_{oo} \cdot S_{uu}}} \tag{6}$$

In Equation (5), S_{ou} represents the covariance between feature o and feature u, k represents the total number of samples, x_o^L represents the o-th feature quantity of the L-th sample in the feature set, and x_u^L represents the u-th feature quantity of the L-th sample in the feature set. In Equation (6), r_{ou} is the linear correlation coefficient between the o-th and u-th features, while S_{oo} and S_{uu} are the variances of features o and u, respectively.

The correlation threshold δ ($0 \leq \delta \leq 1$) was then set. When the correlation threshold between two features satisfies $r_{ij} > \delta$, the features with a large correlation are eliminated, and the first-dimension reduction of the features is achieved. Subsequently, based on the Relief-F algorithm, the k-nearest neighbor sample set H is identified as the set of the same type as feature x_i. The quantitative expression of the difference between feature x_i and each feature in H in feature set A is as follows:

$$diff_hit(A, x, H) = \sum_{j=1}^{k} \frac{|x_i - h_j|}{\max(X) - \min(X)} \tag{7}$$

where $j = 1, 2, 3, \ldots k$, and h_j is the feature in H. Similarly, the k-nearest neighbor sample sets $M(c)$ which are different from x_i need to be identified. The calculation equation is shown in (8).

$$diff_miss[A, x_i, M(c)] = \sum_{c \neq class(x_i)} \frac{P(c)}{1 - P[class(x_i)]} Q \tag{8}$$

where $Q = \sum_{i=1}^{k} \frac{|x_i - m_{cj}|}{\max(X) - \min(X)}$, $P(c)$ is the ratio of the number of target c samples to the total number of samples. Based on the above two differences, the weight of each feature was updated, and features with high weights were screened for optimization.

2.4. Few-Shot Sediment and Benthic-Cover Classification Model Construction

The classic classifier model has problems with low robustness and generalization for the few-shot classification of shallow seabed sediment and benthic covers. To solve these problems, a GAT-PN classification model was proposed. The specific contents of this paper are as follows:

First, the optimal features selected by the FCA-Relief-F method are used to construct a graph structure. Each sediment and benthic-cover sample is regarded as a graph node, and its feature vector is composed of the selected features. To model the relationships among samples, a k-nearest neighbor (KNN) algorithm is employed to establish graph edges in the feature space. Specifically, each node is connected to its k-nearest neighboring nodes according to the Euclidean distance between feature vectors, where k is set to 3–5. Each node is represented as a feature vector, and the attention value of each node is calculated [32]. The attention value can be expressed as in Equation (9):

$$e_{ij} = (\vec{a}^{T}[W\vec{h_i} \Big|\Big| W\vec{h_j}]) \tag{9}$$

where e_{ij} is the influence coefficient of node i on node j, $h = \{h_1, h_2, \ldots, h_N\}$ is the input feature, $\vec{a}^{T}$ is the weight vector, T is the transpose function, and W is the linear-transformation weight matrix. The ratio of the attention value to the sum of the attention values of the neighboring nodes is defined as the attention score a_{ij}, which can be expressed as:

$$a_{ij} = \frac{\exp(\text{LeakyReLu}(e_{ij}))}{\sum_{k \in N_i} \exp(\text{LeakyReLu}(e_{ik}))} \tag{10}$$

where k is a neighbor node, N_i is the number of neighbor nodes of node i, and LeakyReLu is the activation function. Eventually, the new feature value is obtained, which can be expressed as follows:

$$\vec{h_i'} = \sigma(\sum_{j \in N_i} a_{ij} W \vec{h}_j) \tag{11}$$

where $\vec{h_i'}$ denotes the updated feature vector of node i, σ denotes the activation function, and W denotes the linear transformation weight matrix.

Based on the prototype network method, support set samples were used to map the features to a feature space. In the feature space, support set samples of the same type were averaged to obtain the prototype center of each class. The query set was then mapped to the feature space. The distance between the sample and each prototype point was measured for each query set sample. The Euclidean distance equation is employed to measure distances,

as shown in (12) [33]. The negative distance metric is then fed into the softmax layer to predict the label.

$$D(f(x_i), proto_k) = \sqrt{\sum_{j=1}^{r} (f(x_i)_j - proto_{k_j})^2}, x_i \in Q \quad (12)$$

where $proto_k$ is the prototype point generated for each class, k is the label of the sample in this class, f is the feature mapping process of the model, x_i is the sample, y_i is the sample label, Q is the query set, and r is the dimension of the feature vector. The workflow of the GAT-PN model is shown in Figure 2.

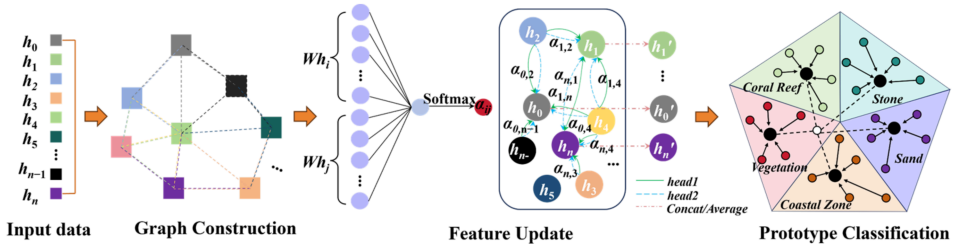


Figure 2. The workflow of the GAT-PN model.

In this study, the GAT-PN model was implemented using the PyTorch framework (version 2.0.1) and trained with the Adam optimizer. The maximum number of training epochs was set to 150, and early stopping with a patience of 20 epochs was employed to prevent overfitting. The learning rate was searched within $\{1 \times 10^{-4}, 5 \times 10^{-4}, 1 \times 10^{-3}\}$, the hidden dimension within {8, 16, 32}, the dropout rate within {0.2, 0.4}, and the number of attention heads was set to 2.

3. Results

3.1. Study Data

The data used in this study were collected from Ganquan Island, which is located northwest of the South China Sea. The location of Ganquan Island is shown in Figure 3a. Ganquan Island is a typical area with coral reefs, and the marine environment in this area is relatively pristine, which makes it suitable for classification of seabed sediment and benthic covers in shallow water.

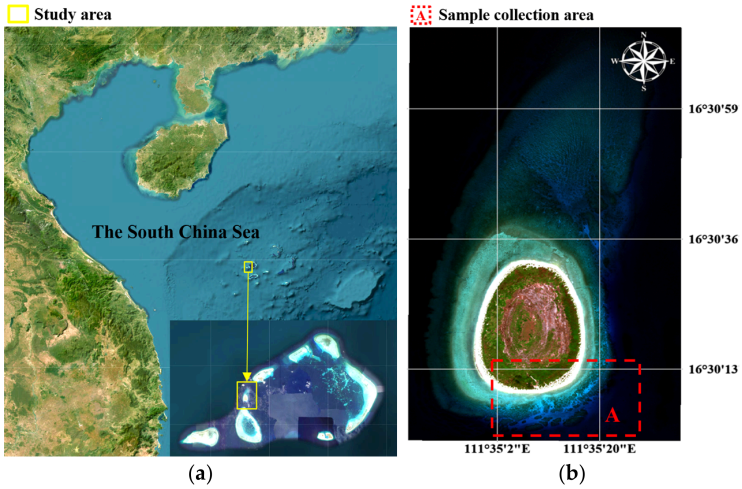


Figure 3. Study area. (**a**) Location of the Ganquan Islands in the South China Sea. (**b**) Remote sensing image of Ganquan Island.

Multispectral remote sensing data were collected from the GeoEye-1 satellite in February 2013, and the multispectral image is shown in Figure 3b. The altitude angle of the satellite was 28.2°. The multispectral image provides information in four bands: blue, green, red, and near-infrared. The wavelengths were 450–510 nm, 510–580 nm, 655–690 nm, 780–920 nm, respectively. The spatial resolution of each band is 2 m. The image is in the Universal Transverse Mercator (UTM) projection and World Geodetic System 1984 (WGS84).

The ALB data were collected using the Optech Aquarius system in January 2013, which employed a 532 nm and 70 kHz green laser. The aircraft was flying at an altitude of 300 m with a minimum laser scanning angle of 15°, laser emission angle of 1 mrad, and pulse width of 8.3 ns. After data preprocessing, a total of 1.8×10^7 seabed points were obtained, and the point cloud density was approximately 4 pts/m^2. To select training and validation samples, high-definition digital photos with a resolution of 5 cm were collected simultaneously. The sample collection area is shown in Area A of Figure 3b.

3.2. Result of Feature Extraction, Registration and Fusion

(1) Waveform and terrain features extraction from the ALB data

According to the waveform and terrain feature extraction method, 9-dimensional waveform and 9-dimensional terrain feature variables were extracted from the Ganquan Island ALB data. The waveform feature extraction results are shown in Figure 4, and the terrain feature extraction results are shown in Figure 5.

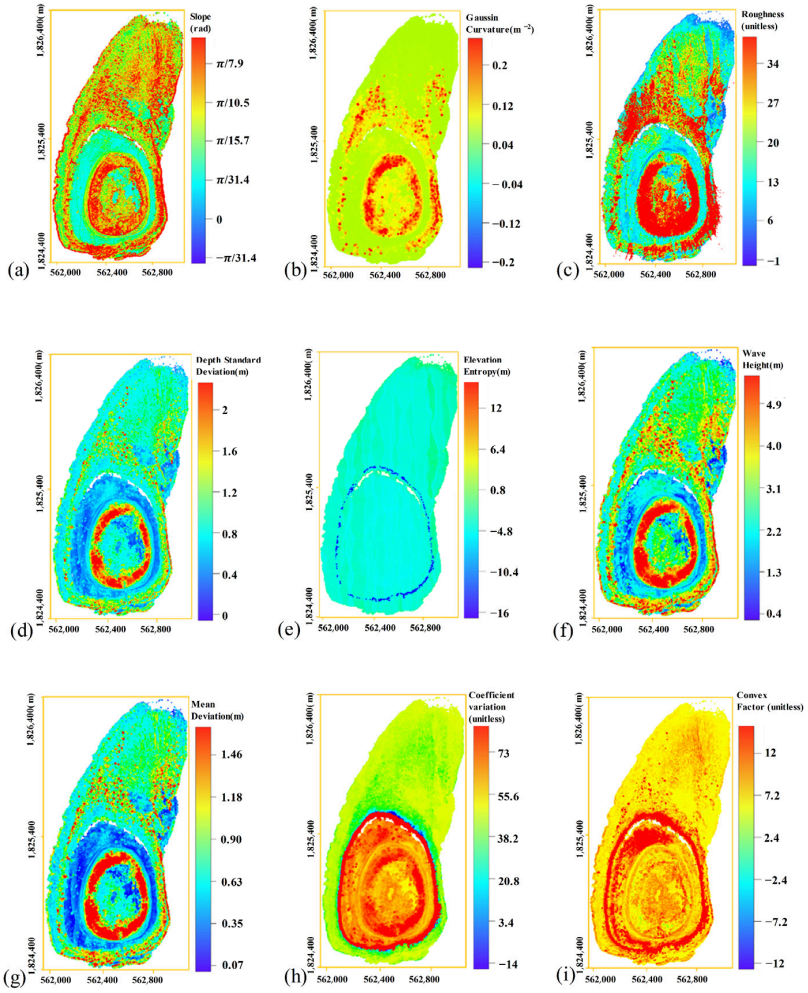


Figure 4. Terrain feature extraction results. (**a**) Slope. (**b**) Gaussian curvature. (**c**) Roughness. (**d**) Depth standard deviation. (**e**) Elevation entropy. (**f**) Wave height. (**g**) Mean deviation. (**h**) Coefficient variation. (**i**) Convex factor.

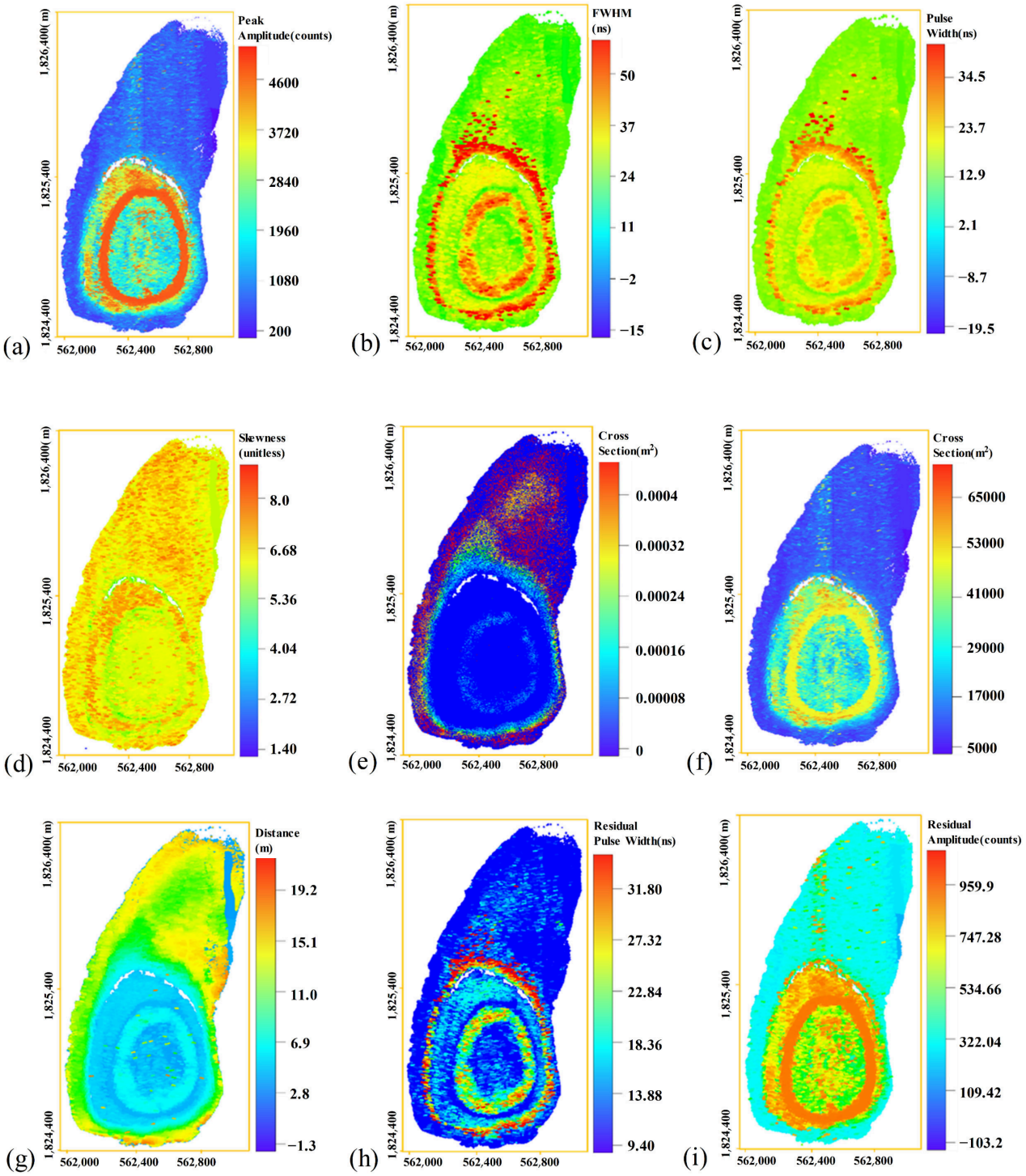


Figure 5. Waveform feature extraction results. (**a**) Peak amplitude. (**b**) FWHM. (**c**) Pulse width. (**d**) Skewness. (**e**) Cross Section. (**f**) Area. (**g**) Distance. (**h**) Residual pulse width. (**i**) Residual amplitude.

(2) Texture and spectral features extraction from multispectral image

According to the texture and spectral feature extraction methods, feature extraction is performed on the four bands of the multispectral image. The extraction results of some texture features (blue-band texture features) are shown in Figure 6, and the spectral feature extraction results are shown in Figure 7.

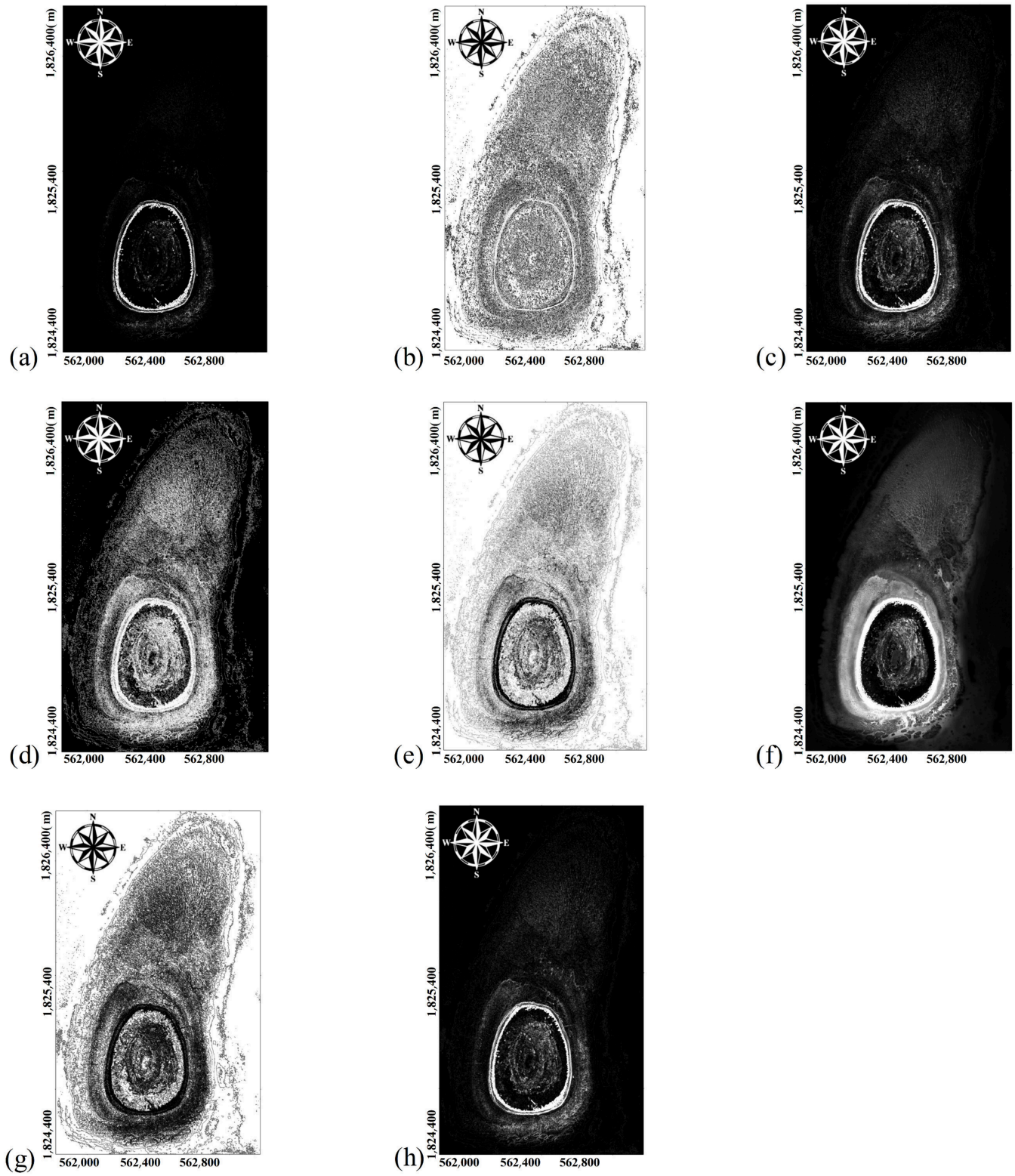


Figure 6. Texture feature extraction results. (**a**) Contrast. (**b**) Correlation. (**c**) Dissimilarity. (**d**) Information entropy. (**e**) Homogeneity. (**f**) Mean. (**g**) Second moment. (**h**) Variance.

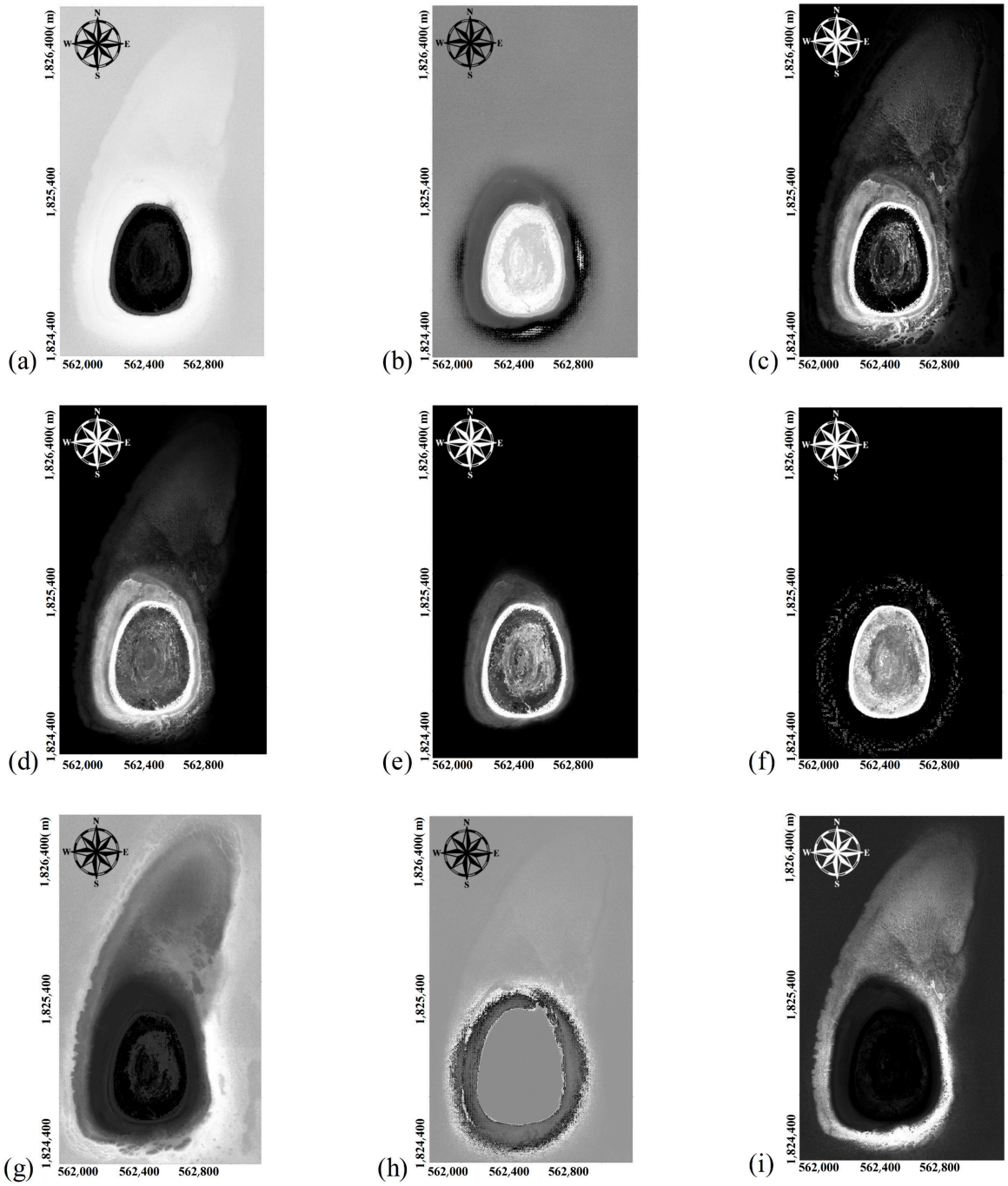


Figure 7. Spectral feature extraction results. (**a**) NDWI. (**b**) NDVI. (**c**) Blue band. (**d**) Green band. (**e**) Red band. (**f**) near-infrared band. (**g**) Ratio of blue band to green band. (**h**) Ratio of blue band to near-infrared band. (**i**) Ratio of green band to red band.

(3) Results of ALB data and multispectral data registration and fusion

As shown in Figure 8a, the ALB intensity image (left image) is used as the reference image. Based on the SIFT algorithm, the matching points of the ALB intensity image and multispectral blue band image were found to achieve a coarse alignment. The pixel grid of the two range images was 1 m × 1 m. It is evident that many incorrectly matched point pairs exist, as indicated by the non-parallel lines. The PROSAC algorithm was then used to achieve precise alignment between the two images, as shown in Figure 8b. The incorrect matching point pairs were eliminated, and 24 pairs of correct matching point pairs were retained in the two images, with a Root Mean Square Error (RMSE) of

0.263 m, where RMSE is used to quantify the geometric registration accuracy between the ALB intensity image and the multispectral image. Finally, according to the pixel coordinates and XY coordinates of the ALB intensity image and the multispectral blue band image calculated, the transformation matrix of the two images was calculated through the perspective transformation model to achieve the fusion of the two datasets. The inverse transformation matrix was obtained using $\begin{pmatrix} 1.000e^{+00} & 1.930e^{-04} & -7.613e^{+01} \\ -6.306e^{-03} & 1.100e^{+00} & -1.374e^{+02} \\ -3.794e^{-07} & 1.054e^{-06} & 9.999e^{-01} \end{pmatrix}$. The ALB point cloud data *xyz* are colored based on multispectral RGB in the fused data, as shown in Figure 8c. The registration results of the SIFT-PROSAC and SIFT-RANSAC algorithms are presented in Table 1. According to the above experiments, the PROSAC algorithm can remove incorrect matching points and obtain the best matching point pairs compared with the RANSAC algorithm. The reprojection deviation and RMSE of the PROSAC algorithm were relatively low, which are 0.205 m and 0.263 m, respectively. Therefore, the PROSAC algorithm obtains better registration and fusion results.

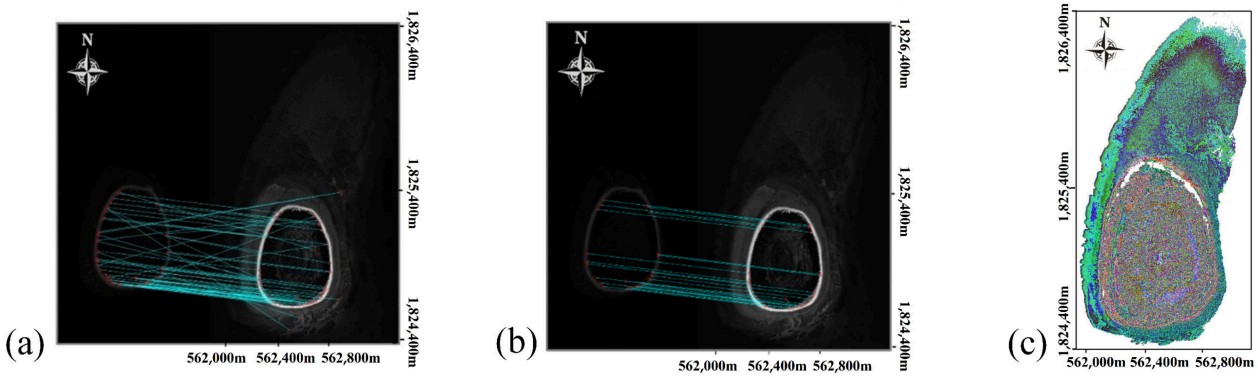


Figure 8. Feature registration and fusion result. (**a**) SIFT coarse registration matching result. (**b**) PROSAC precise registration matching result. (**c**) Fusion point cloud image after RGB coloring.

Table 1. Comparison analysis of registration results. (Bold values indicate the best registration performance among all methods.)

Registration Method	Reprojection Deviation/m	RMSE/m
SIFT-RANSAC	0.341	0.419
SIFT-PROSAC (proposed)	**0.205**	**0.263**

3.3. Result of Classification and Accuracy Assessment

As shown in Figure 9, the experimental area was divided into five sediment and benthic covers: coral reefs, vegetation, coastal zone, sand, and stone. The vegetation category is mainly distributed on the island surface, while the coastal zone forms a narrow belt surrounding the island shoreline. Sand is the dominant seabed type in the shallow-water area around the island and occupies a large proportion of the study area. Coral reefs and stone are primarily distributed in the reef-flat and reef-edge regions, exhibiting more heterogeneous spatial patterns than the sand areas. These sediment and benthic-cover categories present distinct spatial distributions and environmental characteristics, providing a basis for their discrimination using fused ALB and multispectral features.

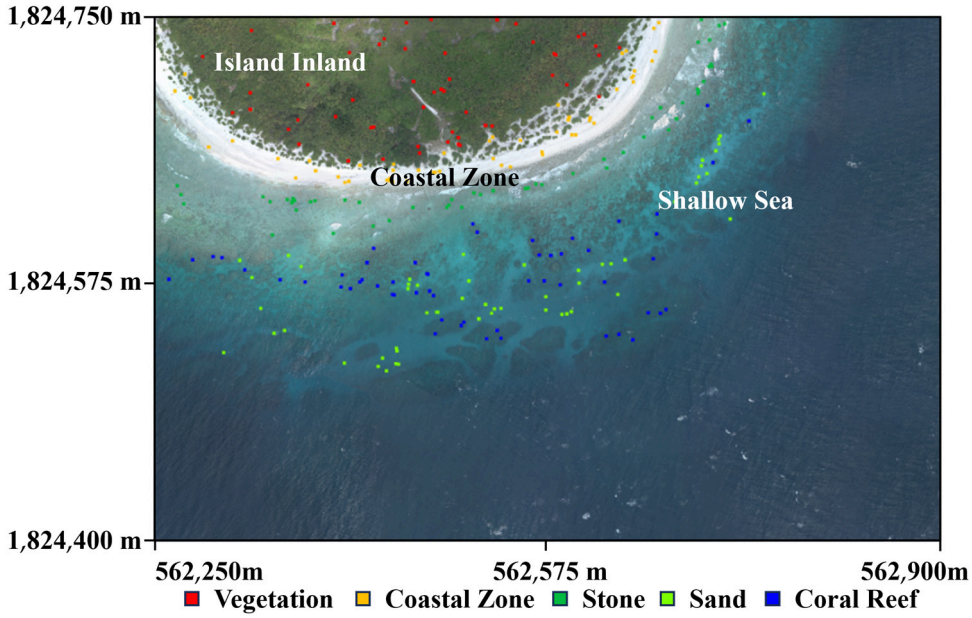


Figure 9. Sample selection distribution results.

The sediment and benthic-cover samples used for training and validation were labeled based on the aerial photographs acquired by the Optech aerial digital camera (Teledyne Optech, Vaughan, ON, Canada) and the underwater videos collected using a GoPro video recorder (GoPro, Inc., San Mateo, CA, USA) during the field survey. The aerial photographs, with a spatial resolution of 5 cm, were used to assist in the selection and delineation of sediment and benthic-cover samples, while the underwater videos served as reference data for sediment and benthic-cover identification. Using the GPS-recorded survey locations, the underwater observations were associated with the corresponding aerial images. For each category, 53 representative samples were identified based on the reference data, yielding a total of 265 labeled samples. Samples located in transitional or overlapping areas between different sediment and benthic-cover types were excluded to ensure the reliability of the reference dataset.

In the classification process, 70% of the labeled samples (185 samples in total) were randomly selected as the support set for prototype construction, while the remaining 30% (80 samples in total) were used as the query set for few-shot classification and performance evaluation. The results were obtained from a single random support-query split. Based on the fusion features of ALB and multispectral images, traditional classifiers such as support vector machine, BP neural network, random forest, prototype network, graph neural network, GraphSAGE, graph attention network, and GAT-PN proposed in this study were used to realize few-shot sediment and benthic-cover classification in shallow water. The overall accuracy (OA) and Kappa coefficient were used as evaluation metrics to assess the classification performance. OA represents the proportion of correctly classified samples, while the Kappa coefficient evaluates the agreement between predicted and reference classes while accounting for agreement that may occur by chance. A Kappa value close to 1 indicates very strong classification agreement. The results are presented in Table 2. The confusion matrix of the proposed GAT-PN model, obtained on the query set, is shown in Figure 10.

Table 2. Comparison and analysis of classification accuracy results. (Bold values indicate the best classification performance among all methods.)

Method	Overall Accuracy	Kappa
Support Vector Machines	43.75%	0.326
Backpropagation neural network	71.25%	0.637
Random Forest	88.78%	0.861
Prototypical Network	86.90%	0.836
Graph Convolutional Networks	87.50%	0.840
Graph Sample and Aggregate (GraphSAGE)	92.50%	0.904
GAT	96.25%	0.953
GAT-PN (proposed)	**97.50%**	**0.969**

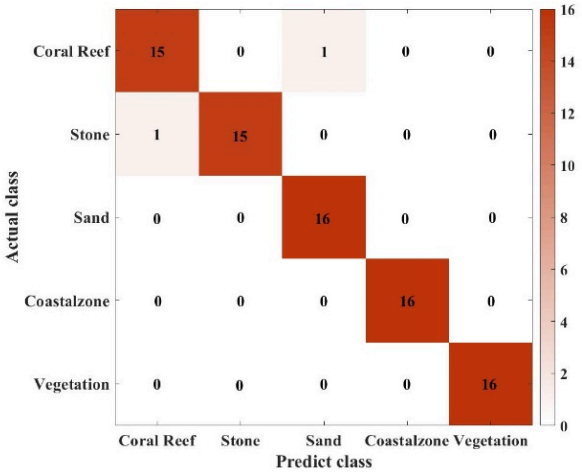


Figure 10. Confusion matrix of test results.

According to the experimental results in Table 2 and Figure 10, it can be seen that among the 8 classification methods, the classification accuracy of algorithms such as support vector machines, backpropagation neural networks, and random forests is low. Therefore, the classic machine learning algorithm is not suitable for few-shot classification. Compared with traditional few-shot classifiers, such as the prototype network, graph convolutional network, GraphSAGE, and graph attention network, the GAT-PN classifier demonstrates strong generalization ability and achieves the highest classification accuracy, with an overall accuracy of 97.50% and a kappa coefficient of 0.969. The prediction results for Ganquan Island based on the model trained with GAT-PN are shown in Figure 11a. A high-definition digital image is shown in Figure 11b.

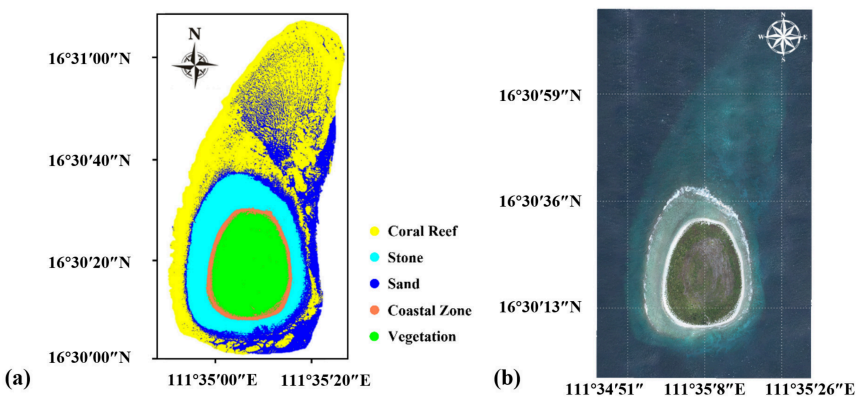


Figure 11. Overall prediction of Ganquan Island and high-definition digital photo. (**a**) Prediction result of Ganquan Island based on GAT-PN. (**b**) High-definition digital photo of Ganquan Island.

4. Discussion

The proposed feature selection strategy consists of two stages. First, highly correlated features were removed through correlation analysis to reduce feature redundancy. Subsequently, feature contribution analysis was performed to identify the most informative features for sediment and benthic-cover classification. The effectiveness of the proposed feature selection strategy, as well as the overall performance and limitations of the proposed framework, are discussed in the following subsections.

4.1. Correlation Analysis of Features

To reduce feature redundancy before Relief-F-based feature evaluation, FCA was applied to analyze the correlations among terrain, waveform, spectral, and texture features. The heat maps of the correlation coefficients of terrain features, waveform features, spectral features, and texture features are shown in Figure 12a,b, 13 and 14, respectively. A correlation coefficient threshold of 0.9 was adopted to identify highly correlated features. Features with correlation coefficients greater than 0.9 were considered redundant and removed from the feature subset.

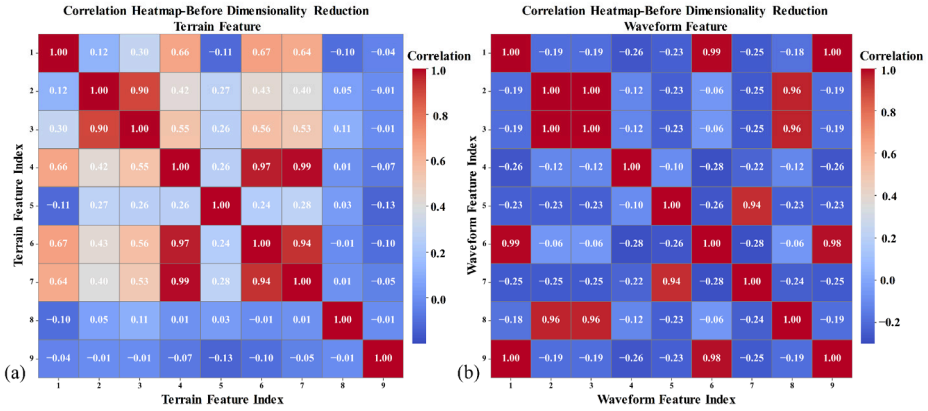


Figure 12. The heat maps of correlation coefficients of ALB features. (**a**) Correlation heat map of terrain features; (**b**) Correlation heat map of waveform features.

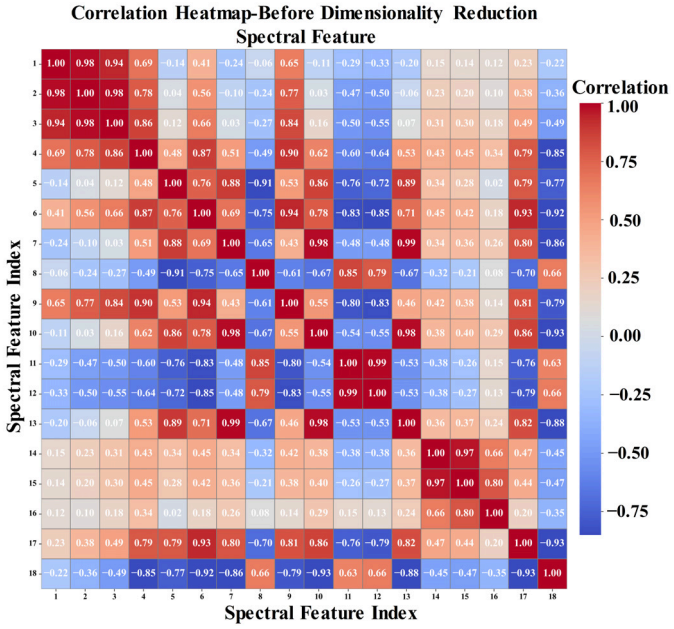


Figure 13. The heat maps of correlation coefficients of spectral features.

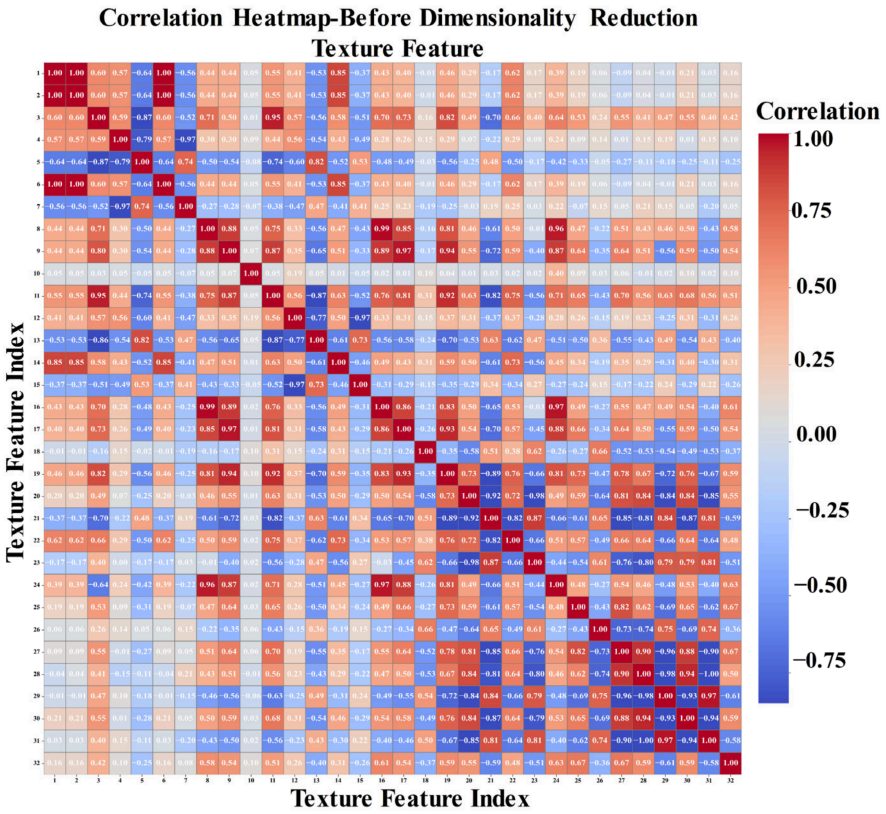


Figure 14. The heat maps of correlation coefficients of texture features.

As shown in Figure 12a, terrain feature 4 (depth standard deviation), terrain feature 6 (wave height), and terrain feature 7 (mean deviation) exhibit strong correlations with each other. These features characterize similar seabed topographic variations and therefore contain overlapping information. Consequently, terrain features 6 and 7 were removed during feature selection. Similarly, Figure 12b shows strong correlations among waveform feature 2 (FWHM), waveform feature 3 (pulse width), and waveform feature 8 (residual pulse width), as well as among waveform feature 1 (peak amplitude), waveform feature 6 (area), and waveform feature 9 (residual amplitude). In addition, waveform feature 5 (cross section) is highly correlated with waveform feature 7 (distance). These results indicate that several waveform descriptors are derived from similar waveform characteristics and therefore provide redundant information. Consequently, redundant waveform features were removed.

As shown in Figure 13, strong correlations are also observed among several spectral features, including the original spectral bands, band-ratio features, and spectral indices. Since these features are derived from related spectral information, a certain degree of redundancy is expected. Likewise, Figure 14 indicates that correlations exist among several texture features representing similar statistical texture properties, such as contrast, correlation, dissimilarity, entropy, mean, and variance. Therefore, some spectral and texture features contain overlapping information and can be removed without significantly reducing the discriminative capability of the feature set.

Overall, the correlation analysis demonstrates that considerable redundancy exists among the original feature set. Removing highly correlated features effectively reduces feature redundancy while preserving the primary discriminative information required for sediment and benthic-cover classification. As a result, 30 features were removed after correlation analysis.

4.2. Feature Contribution Analysis

After FCA-based feature reduction, the contribution rates of the remaining features were evaluated using the Relief-F algorithm. The feature-weight distribution is shown in Figure 15.

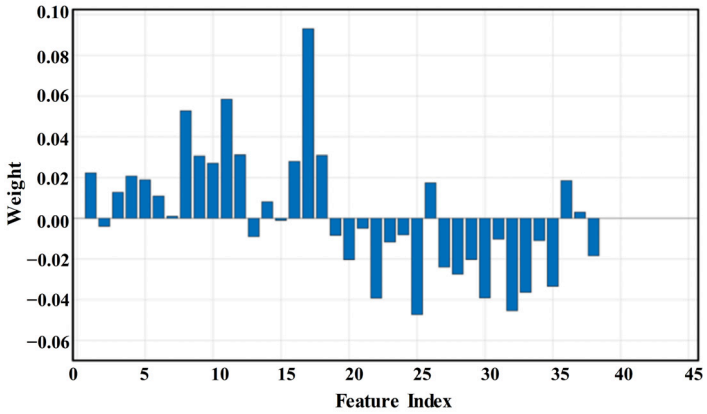


Figure 15. Feature weight distribution diagram.

The feature-weight distribution obtained by Relief-F is shown in Figure 15. The contribution rates of the 38 features retained after FCA vary considerably, with several features exhibiting negative weights. These negative weights indicate that the corresponding features provide limited discriminative information for sediment and benthic-cover classification. Therefore, features with negative contribution rates were removed during the second stage of feature selection, reducing the feature dimensionality from 38 to 15. Together with the FCA stage, which removed highly correlated features, the proposed feature selection strategy reduced the original feature set from 68 to 15 dimensions, indicating that only a limited number of features contain the majority of the discriminative information required for classification. The substantial reduction in feature dimensionality decreases the amount of information processed during the subsequent classification stage, thereby reducing computational complexity.

4.3. Limitations and Future Work

Although the proposed feature selection strategy effectively reduced feature redundancy and retained informative features, several factors may still affect the discriminative capability of the selected features. First, variations in water depth may influence the spectral response of seabed sediment and benthic covers and reduce the separability of different sediment and benthic-cover classes. Second, misclassification is more likely to occur in transitional areas where multiple sediment and benthic-cover types coexist, leading to mixed spectral and structural characteristics. In addition, the current evaluation was conducted using a single random support-query split, and the influence of different sample partitioning strategies on classification performance was not investigated. Furthermore, the proposed method was evaluated using data from a single study area, and its applicability to other coastal environments requires further investigation. Future studies should incorporate more diverse datasets, repeated random trials, and different environmental conditions to further evaluate the robustness and transferability of the selected features and the proposed method.

5. Conclusions

To address the challenges of few-shot sediment and benthic-cover classification in shallow water areas, a classification method based on the fusion of ALB and multispectral images was proposed. The proposed algorithm was applied to the measured data, and the following conclusions were drawn from the accuracy evaluation and analysis of the classification results:

(1) Based on ALB and multispectral features, the constructed SIFT-PROSAC and perspective transformation models were used for feature registration and fusion. The experimental results demonstrate that the algorithm can obtain higher registration accuracy, leading to the generation of more precise fusion features.

(2) To select an effective classifier, after feature optimization, traditional classifiers, such as support vector machines, BP neural networks, random forests, prototype networks, graph convolutional networks, GraphSAGE, graph attention networks, and the GAT-PN classifier proposed in this article, were used to extract five sediment and benthic covers. The experimental results demonstrate that the GAT-PN classifier has the best classification effect, with an overall accuracy and kappa coefficients of 97.50% and 0.969, respectively. The results can provide effective technical support for the classification of seabed sediment and benthic covers in marine engineering and other fields.

Author Contributions: S.C.: conceptualization, methodology, supervision, writing—original-draft, writing—review and editing. X.S.: software, data curation, visualization. J.M.: data curation, validation. Y.H.: conceptualization, methodology, investigation, validation. A.Y.: methodology, validation, visualization, writing—original-draft, writing—review and editing, funding acquisition. R.S.: software, data curation, visualization. H.G.: conceptualization, methodology, validation. D.S.: methodology, supervision, writing—review and editing, funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Key R&D Program of Shandong Province (Competitive Innovation Platform) under Grant 2025CXPT205, National Natural Science Foundation of China under Grant 42406186 and 52001189, Open Innovative Fund of Key Laboratory of Ecological Prewarning and Protection of Bohai Sea, Ministry of Natural Resources under Grant 2023107, Natural Science Foundation of Qingdao under Grant 23-2-1-66-zyyd-jch and 24-8-4-zrjj-2-jch, Key Technology Research and Industrialization Demonstration Project of Qingdao under Grant 23-1-3-hygg-1-hy, Project of Shandong Province Higher Educational Youth Innovation Science and Technology Program under Grant 2023KJ088, Natural Science Foundation of Shandong Province under Grant ZR2023QD050, and Open Innovative Fund of Key Laboratory of Submarine Geosciences, Ministry of Natural Resources under Grant KLSG2306.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to legal restrictions imposed by the Surveying and Mapping Law of China.

Conflicts of Interest: Author Xueshan Song was employed by China Water Resources Beifang Investigation, Design and Research Co., Ltd. Authors Jinfeng Mao and Yu Huang were employed by the Hydrological Bureau of the Yangtze River Water Resources Commission Hydrological and Water Resources Survey Bureau of the Lower Yangtze River. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ALB	Airborne LiDAR Bathymetry
SIFT	Scale-Invariant Feature Transform
PROSAC	Progressive Sample Consensus
FCA	Formal Concept Analysis
GAT	Graph Attention Network
PN	Prototype Network
OA	Overall Accuracy
MBES	Multi-Beam Echo Sounder
RANSAC	Random Sample Consensus
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
UTM	Universal Transverse Mercator
RMSE	Root Mean Square Error
GraphSAGE	Graph Sample and Aggregate
PCA	Principal Component Analysis
KNN	K-nearest neighbor

References

1. Anokye, M.; Cui, X.; Yang, F.; Wang, P.; Sun, Y.; Ma, H.; Amoako, E.O. Optimizing multi-classifier fusion for seabed sediment classification using machine learning. *Int. J. Digit. Earth* **2024**, *17*, 2295988. [CrossRef]
2. Zhao, Y.; Zhu, K.; Zhao, T.; Zheng, L.; Deng, X. Small-sample seabed sediment and benthic-cover classification based on deep learning. *Remote Sens.* **2023**, *15*, 2178. [CrossRef]
3. Foglini, F.; Rovere, M.; Tonielli, R.; Castellan, G.; Prampolini, M.; Budillon, F.; Cuffaro, M.; Di Martino, G.; Grande, V.; Innangi, S.; et al. A new multi-grid bathymetric dataset of the Gulf of Naples (Italy) from complementary multi-beam echo sounders. *Earth Syst. Sci. Data* **2025**, *17*, 181–203. [CrossRef]
4. Guenther, G.C. *Airborne Laser Hydrography System Design and Performance Factors*; NOAA Professional Paper Series; Nat. Ocean Service 1; National Oceanic and Atmospheric Administration, U.S. Department of Commerce: Washington, DC, USA, 1985.
5. Szafarczyk, A.; Toś, C. The use of green laser in LiDAR bathymetry: State of the art and recent advancements. *Sensors* **2022**, *23*, 292. [CrossRef] [PubMed]
6. Amani, M.; Moghimi, A.; Mirmazloumi, S.M.; Ranjgar, B.; Ghorbanian, A.; Ojaghi, S.; Ebrahimy, H.; Naboureh, A.; Nazari, M.E.; Mahdavi, S.; et al. Ocean remote sensing techniques and applications: A review (part I). *Water* **2022**, *14*, 3400. [CrossRef]
7. Amani, M.; Mehravar, S.; Asiyabi, R.M.; Moghimi, A.; Ghorbanian, A.; Ahmadi, S.A.; Ebrahimy, H.; Moghaddam, S.H.A.; Naboureh, A.; Ranjgar, B.; et al. Ocean remote sensing techniques and applications: A review (part II). *Water* **2022**, *14*, 3401. [CrossRef]
8. Pan, Z.; Glennie, C.L.; Fernandez-Diaz, J.C.; Legleiter, C.J.; Overstreet, B. Fusion of LiDAR orthowaveforms and hyperspectral imagery for shallow river bathymetry and turbidity estimation. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4165–4177. [CrossRef]
9. He, X.; Tang, C.; Liu, X.; Zhang, W.; Sun, K.; Xu, J. Object detection in hyperspectral image via unified spectral–spatial feature aggregation. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5521213. [CrossRef]
10. Umeda, K.; Godin, G.; Rioux, M. Registration of range and color images using gradient constraints and range intensity images. In Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, UK, 23–26 August 2004. [CrossRef]
11. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]
12. Cavalli, L.; Barath, D.; Pollefeys, M.; Larsson, V. Consensus-adaptive ransac. *arXiv* **2023**, arXiv:2307.14030. [CrossRef]
13. Zhang, L.; Wang, Y.; Chen, H. Optimal feature selection from non-stationary signals for early anomaly detection using adaptive sparse regression. *Mech. Syst. Signal Process.* **2024**, *208*, 111045. [CrossRef]
14. Le, T.T.; Simmons, W.K.; Misaki, M.; Bodurka, J.; White, B.C.; Savitz, J.; McKinney, B.A. Differential privacy-based evaporative cooling feature selection and classification with relief-F and random forests. *Bioinformatics* **2017**, *33*, 2906–2913. [CrossRef] [PubMed]
15. Su, D.; Yang, F.; Ma, Y.; Zhang, K.; Huang, J.; Wang, M. Classification of coral reefs in the South China Sea by combining airborne LiDAR bathymetry bottom waveforms and bathymetric features. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 815–828. [CrossRef]
16. Zhang, J.; Liu, L.; Silvén, O.; Pietikäinen, M.; Hu, D. Few-shot class-incremental learning for classification and object detection: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2025**, *47*, 2924–2945. [CrossRef] [PubMed]

17. Liu, Y.; Shi, D.; Lin, H. Few-shot learning with representative global prototype. *Neural Netw.* **2024**, *180*, 106600. [CrossRef] [PubMed]
18. Sun, Q.; Liu, Y.; Chua, T.S.; Schiele, B. Meta-transfer learning for few-shot learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019. [CrossRef]
19. Zhang, H.; Wang, Y.; Liu, X. Task-adaptive meta-metric learning for few-shot image classification. *Pattern Recognit.* **2024**, *146*, 110025. [CrossRef]
20. Zhang, J.; Huang, J.; Jin, S.; Lu, S. Vision-language models for vision tasks: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2024**, *46*, 5625–5644. [CrossRef] [PubMed]
21. Wang, S.; Wu, J.; Zhang, L. Adaptive global-local prototypical networks for few-shot representation learning. *Neural Netw.* **2024**, *171*, 215–228. [CrossRef] [PubMed]
22. Corso, G.; Stark, H.; Jegelka, S.; Jaakkola, T.; Barzilay, R. Graph neural networks. *Nat. Rev. Methods Prim.* **2024**, *4*, 17. [CrossRef]
23. Li, W.; Du, L.; Meng, X.; Liu, J.; Li, Y.; Zhang, X.; Wan, D. Adaptive denoising for airborne lidar bathymetric full waveforms using emd-based multiresolution analysis. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 6500705. [CrossRef]
24. Chao, Q.; Dianpeng, S.; Xiankun, W.; Mingwei, W.; Bo, S.; Fanlin, Y. Fitting algorithm for airborne laser bathymetric waveforms based on layered heterogeneous model. *Infrared Laser Eng.* **2019**, *48*, 0206004. [CrossRef]
25. Chao, Q.; Fengnian, Z.; Jingwen, W.; Dianpeng, S.; Xiankun, W.; Fanlin, Y. Extraction method for diffuse attenuation coefficient based on airborne LiDAR bathymetric water column waveform. *Acta Oceanol. Sin.* **2021**, *43*, 147–154. [CrossRef]
26. Guastella, M.; Pisciotta, A.; Martorana, R.; D'aLessandro, A. Automated Detection of Recent Mud Extrusions Using UAV Imagery and Deep Learning: A Comparative Analysis of Traditional and CNN-Based Approaches. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2025**, *48*, 169–179. [CrossRef]
27. Rahman, M.D.; Matta, G.R. PySIFT: GPU-Resident Deterministic SIFT for Deep Learning Vision Pipelines. *arXiv* **2026**, arXiv:2605.17869. [CrossRef]
28. Luo, Y.; Wang, X.; Liao, Y.; Fu, Q.; Shu, C.; Wu, Y.; He, Y. A review of homography estimation: Advances and challenges. *Electronics* **2023**, *12*, 4977. [CrossRef]
29. Wei, X.; Sen, G.; Chengzhi, C.; Zhenwei, J.; Changhai, R. Real-time correction of image drift in scanning electron microscope. *Opto-Electron. Eng.* **2018**, *45*, 180198–180201. [CrossRef]
30. Zhenyu, S.; Lihui, W.; Liqiang, L. Three-dimensional reconstruction algorithm based on inverse perspective transformation. In Proceedings of the 2020 International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), Fuzhou, China, 12–14 June 2020. [CrossRef]
31. Sun, X.; Guo, S.; Liu, S.; Guo, J.; Du, B. A Correlation-Redundancy Guided Evolutionary Algorithm and Its Application to High-Dimensional Feature Selection in Classification. *Neural Process. Lett.* **2024**, *56*, 86. [CrossRef]
32. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. *arXiv* **2017**, arXiv:1710.10903.
33. Zhang, S.; Li, Z.; Zhang, K.; Lu, Y.; Deng, Y.; Tang, L.; Jiang, X.; Ma, J. Deep learning reforms image matching: A survey and outlook. *arXiv* **2025**, arXiv:2506.04619. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.