GIScience & Remote Sensing

ISSN: 1548-1603 (Print) 1943-7226 (Online) Journal homepage: www.tandfonline.com/journals/tgrs20

CRUHI-Mamba: a Mamba-based underwater hyperspectral image classification framework for coral reef benthic mapping

Feifei Zhang , Yabin Hu , Yi Ma , Guangbo Ren , Wenshuo Zhu , Shou Feng & Zhongwei Li

To cite this article: Feifei Zhang , Yabin Hu , Yi Ma , Guangbo Ren , Wenshuo Zhu , Shou Feng & Zhongwei Li (2026) CRUHI-Mamba: a Mamba-based underwater hyperspectral image classification framework for coral reef benthic mapping, GIScience & Remote Sensing, 63:1, 2679361, DOI: 10.1080/15481603.2026.2679361

To link to this article: https://doi.org/10.1080/15481603.2026.2679361

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

View supplementary material

Published online: 08 Jun 2026.

Submit your article to this journal

Article views: 674

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tgrs20

GISCIENCE & REMOTE SENSING
2026, VOL. 63, NO. 1, 2679361
https://doi.org/10.1080/15481603.2026.2679361

Taylor & Francis
by informa

RESEARCH ARTICLE

OPEN ACCESS Check for updates

CRUHI-Mamba: a Mamba-based underwater hyperspectral image classification framework for coral reef benthic mapping

Feifei Zhang[a,b,c,d], Yabin Hu[a,c,e], Yi Ma[a,c,d], Guangbo Ren[a,c], Wenshuo Zhu[a,c,f], Shou Feng[g] and Zhongwei Li[b]

[a]Lab of Marine Physics and Remote Sensing, First Institute of Oceanography, Ministry of Natural Resources, Qingdao, People's Republic of China; [b]College of Oceanography and Space Informatics, China University of Petroleum, Qingdao, People's Republic of China; [c]Technology Innovation Center for Ocean Telemetry, Ministry of Natural Resources, Qingdao, People's Republic of China; [d]National Engineering Laboratory for Integrated Aero-Space-Ground-Ocean Big Data Application Technology, Northwestern Polytechnical University, Xi'an, People's Republic of China; [e]Key Laboratory of Marine Environment Detection Technology and Application, Ministry of Natural Resources, Guangzhou, People's Republic of China; [f]College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, People's Republic of China; [g]College of Information and Communication Engineering, Harbin Engineering University, Harbin, People's Republic of China

ABSTRACT

Scientific management of coral reefs is of significant ecological and economic importance. Underwater hyperspectral imagery (UHI) enables fine-scale benthic classification and mapping, providing effective support for coral reef monitoring and conservation. However, the high-dimensional nature of UHI and the pronounced spatial heterogeneity of coral reef benthic types, coupled with limited training samples, hinder the effective extraction of discriminative spatial–spectral features for fine-grained benthic classification. Moreover, the imbalance in spatial distribution and information density across benthic types biases feature learning toward prominent regions, thereby compromising the model's discrimination of small-scale or weak-response types. To address these limitations, a coral reef underwater hyperspectral image classification model based on the Mamba architecture, termed CRUHI-Mamba, is proposed. First, a spatial–spectral Mamba backbone (SSMB) is constructed, which, through the synergistic design of dual pooling layers and spatial–spectral dual-branch Mamba blocks, enables the collaborative modeling and feature extraction of global spatial dependencies and fine-grained spectral differences, even with limited training samples. Next, to mitigate classification bias caused by the uneven information density among benthic types, an adaptive feature–decision fusion module (AFDFM) is designed to dynamically balance spatial–spectral feature information, prevent information loss, enhance feature representations, and correct classification bias, ultimately generating robust classification maps. Classification experiments on coral reef UHI data collected at three different water depths demonstrate that CRUHI-Mamba outperforms state-of-the-art methods in quantitative accuracy, computational complexity, and efficiency, achieving an overall accuracy exceeding 99.22% with an inference time of less than 10 seconds. CRUHI-Mamba therefore provides an effective methodological support for fine-scale benthic habitat mapping in coral reefs.

ARTICLE HISTORY
Received 26 January 2026
Accepted 21 May 2026

KEYWORDS
Underwater hyperspectral image (UHI); Mamba; classification; coral reef benthic

1. Introduction

The coral reef ecosystem is among the most productive and biodiverse on Earth (Hughes et al. 2020), not only providing critical habitat for one quarter of all marine life but also benefiting about one billion people directly or indirectly (Sing Wong, Vrontos, and Taylor 2022). In April 2025, the International Coral Reef Initiative (ICRI) reported that 84% of the world's coral reefs were impacted by the most intense global coral bleaching event ever recorded (ICRI and I.C.R.I 2025). With the global decline in coral reef health, there is an urgent need for rapid, efficient, and accurate methods for the mapping and classification of seafloor composition and benthic types (Chen et al. 2024; Zhang et al. 2024; Zuo et al. 2025). In this context, image-based remote sensing,

CONTACT Yi Ma mayimail@fio.org.cn

Supplemental data for this article can be accessed online at https://doi.org/10.1080/15481603.2026.2679361.

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

particularly underwater remote sensing, has become a reliable and indispensable tool for coral reef monitoring (Chen, Cheng, and Zhang 2025; Ji et al. 2024; Li et al. 2024).

Underwater hyperspectral images (UHI) of coral reefs contain high-resolution spatial and fine-grained spectral information, providing abundant discriminative cues for precise benthic classification at the species and genus levels (Rashid and Chennu 2020; Zeng et al. 2022). Notably, the blue–green spectral band (400–600 nm), characterised by minimal water absorption, forms the primary band for underwater optical information. High-fidelity spectral feature extraction from this band is therefore indispensable for accurate benthic classification. In contrast, conventional RGB imagery captures only three broad spectral bands, which limits its ability to distinguish fine benthic types and typically results in coarse-grained classification (Kurinchi-Vendhan, Grey, and Cole 2023; Zhang, Gruen, and Li 2022). Consequently, UHI has attracted growing research attention in coral reef monitoring (Mills et al. 2023; Schürholz and Chennu 2023), with existing studies demonstrating its effectiveness in estimating benthic community composition (Chennu et al. 2017). However, most work has focused on data acquisition rather than fully exploiting UHI characteristics or developing advanced classification methods. Current approaches are often restricted to coarse benthic categories and rely on traditional classifiers such as support vector machine (SVM) and CNN-based models, which typically require extensive labelled data, leading to high annotation costs and limited utilisation efficiency of UHI data.

It is worth noting that the underwater environment of coral reefs is highly complex. Compared with terrestrial hyperspectral imagery, benthic types in coral reef UHIs exhibit greater spatial heterogeneity and more intricate spectral–spatial coupling (Mills et al. 2024; Zhu, Du, and Zhang 2021, 2024). Such high spatial heterogeneity requires modelling long-range spatial dependencies over large areas, i.e. capturing spatial interactions—such as adjacency and interleaving patterns—among benthic types across extended distances to effectively distinguish them. However, these dependencies are difficult to capture using classification models that rely solely on local receptive fields. In addition, in coral reef UHIs, some benthic types (e.g. sediment, turf algae) occupy large continuous areas with a large number of pixels, whereas others (e.g. rare coral species or small benthic types) are confined to small patches or have intrinsically small sizes, resulting in low information density (Li et al. 2024). This imbalance in information density biases feature learning toward dominant regions, thereby degrading the model's ability to accurately discriminate small-scale or sparsely distributed benthic types. Consequently, a key challenge in coral reef UHI classification lies in effectively modelling long-range dependencies and adaptively modelling critical information.

Recently, the selective state-space model Mamba has been introduced into hyperspectral remote sensing image classification, providing a new paradigm to address the aforementioned challenges (Gu and Dao 2023; Li et al. 2024; Sheng et al. 2025). Unlike CNN-based models that rely on local convolutional operations, Mamba captures long-range dependencies at the sequence level through state-space modelling. Employing selective gating mechanisms, it enables effective integration of global contextual information in underwater hyperspectral scenarios characterised by strong spatial heterogeneity and uneven information distribution (Pan et al. 2025; Zhao et al. 2025). Furthermore, Mamba supports efficient modelling of long sequences with linear computational complexity, alleviating the high computational overhead associated with Transformer-based models in high-resolution underwater applications. Thus, Mamba offers a theoretically grounded foundation for developing efficient and robust coral reef UHI classification frameworks.

Given the demand for rapid, efficient, and accurate acquisition of information on coral reef benthic communities and community structure, as well as the challenges posed by the high spatial heterogeneity, uneven distribution of information density, and substantial requirement for labelled samples in coral reef UHI, we propose a Mamba-based coral reef underwater hyperspectral image classification model (CRUHI-Mamba) for benthic classification. CRUHI-Mamba fully leverages the long-range modelling capabilities and the advantage of linear computational complexity of the Mamba architecture, enabling precise and efficient classification of coral reef UHI even when only a limited number of labelled samples are available. The main contributions of this study are summarised as follows.

1. CRUHI-Mamba uses less than 2% of labelled samples to model long-range interactions across the entire UHI at the pixel level, achieving high-quality benthic classification results with complete and clear boundaries. This provides an efficient and practical method for large-scale coral reef benthic mapping.

2. To effectively extract discriminative features from coral reef UHI with high spatial heterogeneity, we propose a Spatial–Spectral Mamba Backbone (SSMB). The SSMB decouples the spatial and spectral dimensions, separately modelling global spatial dependencies and long-range spectral correlations, thereby enabling efficient extraction of discriminative spatial–spectral features.
3. To mitigate classification bias and fusion difficulties arising from imbalanced information density across benthic types, we propose an Adaptive Feature–Decision Fusion Module (AFDFM). By decoupling feature-level and decision-level fusion, AFDFM enhances feature representation, corrects classification biases, and improves model robustness, leading to more accurate benthic classification.

2. Related works

2.1. Classification for UHI of coral reefs

Early studies on coral reef UHI benthic classification primarily relied on the Spectral Angle Mapper (Foglini et al. 2019) and the Minimum Distance Classifier (Teague et al. 2023) in the commercial software ENVI to classify broad benthic categories such as hard corals, soft corals, and sand. These studies demonstrated the practicality and applicability of UHI for coral reef benthic monitoring. Using hyperspectral transect data from the HyperDiver system, Chennu et al. (2017) applied Multi-layer Perceptron and SVM algorithms to classify 11 benthic cover classes. Furthermore, Rashid and Chennu (2020) utilised HyperDiver in Curaçao, the Caribbean Sea, to collect a large volume of coral reef UHI data. They applied a patch-based spectral–spatial residual network (SSRN) (Zhong et al. 2018) to perform fine-scale classification of 43 benthic categories, achieving overall accuracies of 76% (Schürholz and Chennu 2023). In addition, using UHI data of Guam coral reefs collected with the beta version of HyperDiver, Mills et al. (2024) applied the SSRN model to classify 10 benthic community types, achieving an average classification accuracy exceeding 88%.

These studies primarily focused on testing UHI systems. In terms of classification approaches, most existing methods rely on commercial software or CNN-based models. However, pixel-based models such as SVMs fail to capture spatial context, while CNNs suffer from localised receptive fields and high computational costs, reducing accuracy and efficiency. Furthermore, these methods require large labelled datasets, which significantly hinder the efficient use of UHI data. These limitations motivate the proposed Spatial–Spectral Mamba Backbone (SSMB), which is detailed in Section 3.

2.2. Mamba and its application to hyperspectral image classification

Mamba is built upon State Space Modelling (SSM). To apply continuous SSMs to discrete input sequences, SSMs are discretized using the zero-order hold method (Pechlivanidou and Karampetakis 2022). By adopting the HiPPO matrix as the state transition matrix to capture long-range dependencies (Gu et al. 2020), the Structured State Space Sequence model (S4) is constructed (Gu, Goel, and R'E 2021). Building on S4, Mamba introduces an input-dependent selectivity mechanism, which dynamically adjusts internal parameters based on the input, forming the Selective State Space model (S6). Combined with hardware-aware parallel scanning and memory layout optimisation, Mamba achieves high inference efficiency with linear computational complexity (Wang et al. 2025).

Mamba has been increasingly applied to hyperspectral image classification (He et al. 2024; Ming et al. 2025; Wang et al. 2024). Existing studies mainly explore its spatial–spectral feature extraction through various scanning strategies or dual-branch spatial–spectral architectures (Liu et al. 2024; Sheng et al. 2025; Wang et al. 2025). However, these methods typically rely on hyperspectral patches as input, limiting the modelling of global spatial–spectral context. Nevertheless, Mamba-based methods have shown competitive performance even with simple scanning strategies (Zhu et al. 2024). Currently, few studies have explored pixel-level extraction of global spatial–spectral features directly from entire hyperspectral images (Li et al. 2024; Pan et al. 2025; Zhu et al. 2023).

However, the above Mamba-based models exhibit complex structures, and the datasets used are typically small in spatial scale, mainly derived from urban and agricultural scenes. These models require further optimisation and validation in terms of structure and generalisation. Additionally, coral reef benthic

types have imbalanced information density, and fusing only spectral and spatial features at the feature level often results in the loss of modality-specific discriminative information, compromising classification accuracy. For example, macroalgae and sediment may appear spectrally similar but differ in spatial texture, and feature-only fusion can suppress such spatial cues, causing confusion. These benthic types also display strong spatial heterogeneity and complex distributions, making them difficult to distinguish with local spatial information alone. To address this, we use the entire UHI as input and employ a specifically designed SSMB for pixel-level global feature extraction. The proposed AFDFM adaptively fuses spatial–spectral features and dynamically weights predictions from each branch at the decision level, enhancing accuracy while maintaining linear computational complexity.

3. Methodology

This section provides a detailed description of the proposed CRUHI-Mamba. In Section 3.1, we present the overall framework of CRUHI-Mamba. In Sections 3.2 to 3.4, we describe the motivation and design ideas behind each innovation in detail, including the feature extraction SSMB, the AFDFM, and the Chunked Linear Weighted Averaging (CLWA) strategy for training and inference.

3.1. Overall framework

Figure 1 illustrates the proposed CRUHI-Mamba framework, a pure Mamba-based architecture comprising three components: a pixel-based embedding layer, the SSMB, and the AFDFM.

The original UHI is denoted as $I \in R^{H\times W\times C}$, where H and W denote the spatial dimensions (height and width), and C denotes the spectral dimension. To eliminate redundant spectral information caused by linear correlation, Principal Component Analysis (PCA) is applied to the UHI to obtain $X \in R^{H\times W\times V}$. V represents the reduced spectral dimension, while the spatial dimensions H and W remain unchanged. The data cube X serves as the input to the CRUHI-Mamba framework.

The ***Pixel embedding layer*** takes as input the cube $X \in R^{H\times W\times V}$ and performs nonlinear mapping and compression of the spectral dimension at the pixel level via 1×1 convolutions, thus generating a spectral embedding for each spatial location. Unlike traditional convolutions that aggregate spatial information

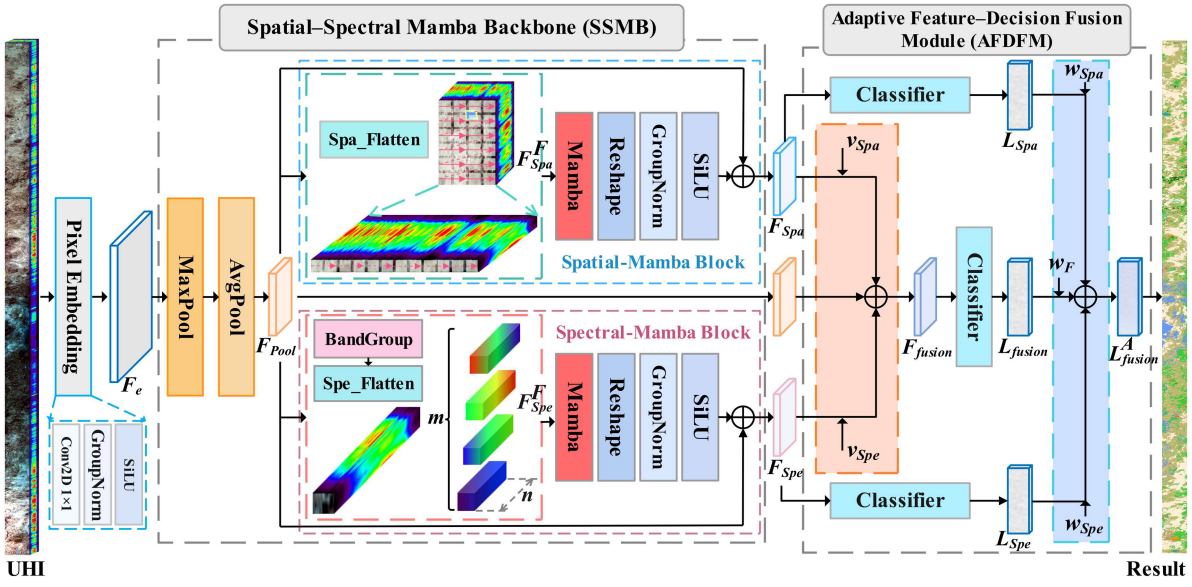

Figure 1. Overview of the proposed CRUHI-Mamba framework, consisting of three main components: a Pixel Embedding layer, a Spatial–Spectral Mamba Backbone (SSMB), and an Adaptive Feature–Decision Fusion Module (AFDFM). The model takes the entire coral reef underwater hyperspectral image (UHI) as input and processes it through these components to generate the benthic classification map.

early, this design avoids premature mixing of spatial neighbourhoods, preserving fine-grained spatial structure in coral reef UHI. 1×1 convolutions model correlations among spectral bands and align feature dimensions, all without expanding the spatial receptive field. Consequently, they produce a unified and compact input representation, which is well-suited for the subsequent SSMB modules that rely on long-sequence modelling. This process can be formulated as follows:

$$F_e = \mathrm{SiLU}(\mathrm{GN}(\mathrm{Conv}_{2D}(X))) \tag{1}$$

where $F_e \in R^{H \times W \times D}$ denotes the output embedding feature map, and D is the embedding dimension. Conv_{2D} denotes the 1×1 convolutional layer. GroupNorm (GN) is a group normalisation layer that divides channels into G groups and normalises each group independently. SiLU is a smooth non-linear activation function.

The ***SSMB*** efficiently extracts discriminative spatial–spectral features with linear computational complexity. It consists of a dual-pooling layer that enhances local saliency and models global statistics, preserving key structural information while generating compact representations. Subsequently, dual-branch Mamba blocks—spatial (SSMB_{spa}) and spectral (SSMB_{spe})—perform joint modelling of global spatial dependencies and fine-grained spectral differences. The feature extraction process is represented as follows:

$$F_j = \mathrm{SSMB}(F_e) \tag{2}$$

where SSMB denotes the feature extraction backbone, and $F_j \in R^{(H/4) \times (W/4) \times D}$ represents the output feature maps. Here, the subscript $j \in \{Pool, Spa, Spe\}$ refers to the output features from the dual-pooling layer, the spatial Mamba block, and the spectral Mamba block, respectively.

The ***AFDFM*** is designed to adaptively fuse spatial and spectral features at the feature and decision level, thereby alleviating the imbalance in representational strength and category-discriminative preference across feature branches. The fused logits output L^A_{fusion} provides an integrated decision basis for coral reef benthic type classification. At a high level, the fusion process can be summarised as:

$$L^A_{fusion} = \mathrm{AFDFM}(\{F_j\}) \tag{3}$$

where $\{F_j\}$ denotes the feature set output by the SSMB. The fused output logits are given by $L^A_{fusion} = [L_1, L_2, \ldots, L_k, \ldots, L_K] \in R^{(H/4) \times (W/4) \times K}$. Here $L_k \in R^{(H/4) \times (W/4)}$ represents the logit map for the kth class at each spatial location, and K denotes the total number of classes. The detailed formulations of feature-level and decision-level fusion within AFDFM are presented in Section 3.3.

Finally, bilinear interpolation is applied to resize L^A_{fusion} to match the size of the input image. Then, the argmax function is used to determine the class of each pixel, yielding the final coral reef benthic type classification result $Y \in R^{H \times W}$, represented as follows:

$$Y = \mathrm{argmax}\,(\mathrm{Resize}(L^A_{fusion})) \tag{4}$$

3.2. Spatial–spectral Mamba backbone

The SSMB contains a dual-pooling layer and a dual-branch spatial-spectral Mamba block. First, the dual pooling layer is used to incrementally recode features, rather than simply stacking down-sampled layers, reducing spatial resolution and computational complexity. Sequentially, max pooling is applied first, followed by average pooling. Max pooling highlights local salient features, such as edges and textures, while average pooling performs global statistical modelling to enhance consistency and suppress noise. This process balances local saliency and global structure in a compact representation, avoiding reliance on multiple pooling operations for spatial compression. The resulting feature F_{Pool} is then fed into the dual-branch spatial-spectral Mamba block.

Given that UHI suffers from spectral distortion and geometric deformation due to water absorption, scattering, and push-broom acquisition, our dual-branch spatial–spectral Mamba feature extraction module decouples the processing of spatial and spectral dimensions to prevent interference between distorted features. Specifically, the spatial branch captures long-range spatial dependencies among benthic types over large areas, directly addressing the challenge of extreme spatial heterogeneity and imbalanced

information density. The spectral branch models fine-grained spectral variations while being robust to spectral distortion by decoupling spectral processing from spatial features. This architecture collaboratively models locally salient features and global spatial–spectral information while maintaining linear computational complexity. The detailed structure is illustrated in Figure 1. The forward process can be expressed as follows:

$$F_{Pool} = \mathrm{MaxPool}_{2D}(\mathrm{AvgPool}_{2D}(F_e)) \tag{5}$$

$$F_{Spa} = \mathrm{SSMB}_{\mathrm{Spa}}(F_{Pool})$$

$$F_{Spe} = \mathrm{SSMB}_{\mathrm{Spe}}(F_{Pool}) \tag{6}$$

where $F_{Pool} \in R^{(H/4)\times(W/4)\times D}$ denotes the output features of the dual pooling layer, $\mathrm{SSMB}_{\mathrm{Spa}}$ and $\mathrm{SSMB}_{\mathrm{Spe}}$ are the spatial and spectral feature extraction Mamba blocks, F_{Spa} and $F_{Spe} \in R^{(H/4)\times(W/4)\times D}$ are the extracted spatial and spectral features, respectively.

3.2.1. $SSMB_{Spa}$ block

To address the complex distribution of coral reef benthic types with high spatial heterogeneity, we utilise the entire F_{Pool} as input and employ Mamba modules as the fundamental units to construct a spatial feature extractor, thereby establishing global pixel-level spatial dependencies with linear computational complexity. The $\mathrm{SSMB}_{\mathrm{Spa}}$ block is shown in Figure 1. The forward process can be formulated as follows:

$$F_{Spa}^{F} = \mathrm{Flatten}(F_{Pool})$$

$$F_{Spa}^{M} = \mathrm{Mamba}(F_{Spa}^{F})$$

$$F_{Spa} = \mathrm{SiLU}(\mathrm{GN}(\mathrm{Reshape}(F_{Spa}^{M}))) + F_{Pool} \tag{7}$$

where $F_{Spa}^{F} \in R^{(H/4)\times(W/4)\times D}$ denotes the flattened spatial output. Zhu et al. (2024) demonstrated that the Mamba-based classification method can achieve satisfactory performance using a simple flattening strategy. In this paper, we adopt the row-optimised flattening strategy, as illustrated in Figure 1, which preserves local spatial continuity along the row direction. This ensures that spatially adjacent pixels remain adjacent in the resulting sequence, a property that benefits the sequential modelling of Mamba. Compared with column or block flattening, row flattening introduces a more regular and consistent traversal order, facilitating efficient spatial dependency modelling with lower structural complexity. $F_{Spa}^{M} \in R^{(H/4)\times(W/4)\times D}$ denotes the spatial features extracted by the standard Mamba architecture (Gu and Dao 2023). The output is sequentially processed through a reshaping layer, a GN layer, and a SiLU activation function, with a residual connection added to enhance representational capacity. Finally, the spatial features F_{Spa} are obtained.

3.2.2. $SSMB_{Spe}$ block

To fully exploit the spectral information in the blue–green wavelength range of coral reef UHI data, we extract fine-grained and discriminative spectral features by dividing the spectral bands into m groups. This grouping strategy reduces the impact of local noise and enhances spectral coherence, as illustrated in Figure 1. Long-range spectral correlations among different spectral groups are modelled using the Mamba architecture. The spectral features are updated based on learned relationships among spectral groups. The output spectral feature F_{Spe} of the $\mathrm{SSMB}_{\mathrm{Spe}}$ block is computed as follows:

$$F_{Spe}^{B} = \mathrm{BandGroup}_{m,n}(F_{Pool})$$

$$F_{Spe}^{F} = \mathrm{Flatten}(F_{Spe}^{B})$$

$$F_{Spe}^{M} = \mathrm{Mamba}(F_{Spe}^{F})$$

$$F_{Spe} = \mathrm{SiLU}(\mathrm{GN}(\mathrm{Reshape}(F_{Spe}^{M}))) + F_{Pool} \quad (8)$$

where $F_{Spe}^{B} \in R^{(H/4)\times(W/4)\times m\times n}$ denotes the output of the spectral grouping layer, $F_{Spe}^{F} \in R^{((H/4)\times(W/4))\times m\times n}$ denotes the flattened spectral group features, and $F_{Spe}^{M} \in R^{((H/4)\times(W/4))\times m\times n}$ corresponds to the spectral feature representations extracted by Mamba. The spectral grouping layer (BandGroup) partitions the input channels into m groups, where $n = D/m$ is the number of channels per group. The flattening operation arranges, for each spatial location, its m group features into a sequence of length m while collapsing the spatial dimensions, producing F_{Spe}^{F}. Mamba processes the sequence dimension (length m) to capture cross-group spectral dependencies and outputs F_{Spe}^{M}. The feature F_{Spe}^{M} is then reshaped back to the spatial dimensions, resulting in a feature of shape $(H/4) \times (W/4) \times$ m $\times$ n. Then, as in the $\mathrm{SSMB_{Spa}}$ block, GN, SiLU activation, and residual concatenation are applied sequentially. Finally, the spectral output $\mathrm{F_{Spe}}$ of the $\mathrm{SSMB_{Spe}}$ block is produced.

3.3. Adaptive feature–decision fusion module

To exploit the complementary strengths of spatial and spectral features in benthic type recognition and alleviate the imbalance in representational capacity and class-discriminative preference across feature branches caused by uneven information density, we design the AFDFM module (Figure 1). At the feature level, spatial features F_{Spa} and spectral features F_{Spe} are adaptively fused within a shared representation space to enhance complementary information, while residual connections are introduced to prevent the suppression of key discriminative features during fusion, yielding a stable and more discriminative fused feature F_{fusion}. This adaptive fusion also helps mitigate the negative effects of spectral distortion and geometric deformation, as the combined feature can compensate for the unreliability of each individual branch. This stage focuses on representation-level integration but cannot fully eliminate decision-level bias across different feature branches. Therefore, at the decision level, three parallel classifiers (Figure 2a) are applied to F_{Spa}, F_{fusion}, and F_{Spe} to generate independent classification logits, which are further combined through an adaptive decision fusion mechanism that dynamically estimates the relative importance of each branch for different categories, enabling robust classification under class imbalance and pronounced differences in feature discriminability. By decoupling feature-level and decision-level fusion, AFDFM jointly

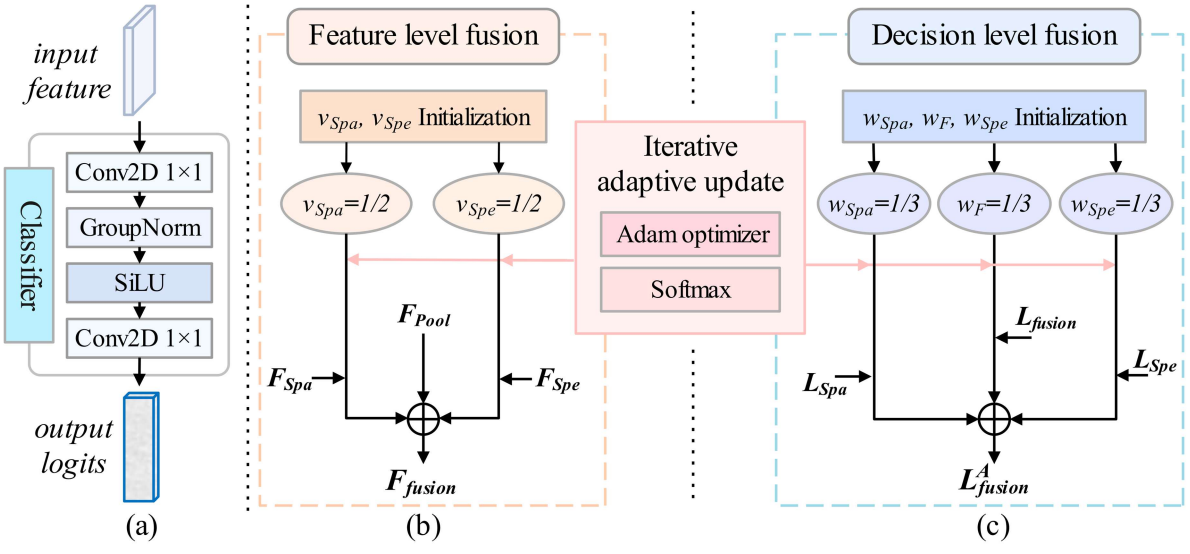

Figure 2. Partial module structure. (a) Classifier structure. (b) Feature-level fusion weight initialisation and update; the learnable scalar weights are denoted as v_{Spa} and v_{Spe}. (c) Decision-level fusion weight initialisation and update; the learnable scalar weights are denoted as w_{Spa}, w_F, and w_{Spe}.

enhances feature representation and corrects classification bias within a unified framework, leading to more robust spatial–spectral classification. The fusion process is represented as follows:

$$L_{Spa} = \text{Classifier}(F_{Spa})$$

$$L_{Spe} = \text{Classifier}(F_{Spe}) \tag{9}$$

$$F_{fusion} = v_{Spa} F_{Spa} + v_{Spe} F_{Spe} + F_{Pool} \tag{10}$$

$$L_{fusion} = \text{Classifier}(F_{fusion}) \tag{11}$$

$$L^{A}_{fusion} = w_{Spa} L_{Spa} + w_F L_{fusion} + w_{Spe} L_{Spe} \tag{12}$$

where $\text{Classifier}(\cdot)$ is the classifier. $v_{Spa} and v_{Spe}$ represent the learnable scalar fusion weights of spatial and spectral features, respectively. Both are initialised to 0.5 and are adaptively updated during model training, as illustrated in Figure 2b. L_{Spa}, L_{Spe}, $and L_{fusion} \in R^{(H/4)\times(W/4)\times K}$ are the classification predictions corresponding to the spatial, spectral, and spatial–spectral fused features, respectively, obtained from the classifier. w_{Spa}, w_{Spe}, $and w_F$ are the corresponding learnable scalar weights for decision fusion, each initialised to one-third. These weights are adaptively updated using the Adam optimiser at each iteration to yield the final fusion weights (Figure 2c). L^{A}_{fusion} denotes the final fused classification prediction generated by the AFDFM.

3.4. CRUHI-Mamba training and inference

CRUHI-Mamba is a pixel-based classification method that takes the entire hyperspectral image as input. Owing to the large data volume of coral reef UHI data and hardware limitations, direct training on full images is impractical. To balance high training/inference efficiency with high classification accuracy, we leverage the characteristics of UHI data to perform chunked training and inference along the H dimension of the input $X \in R^{H\times W\times V}$. The training loss $\mathcal{L}$ is calculated as follows:

$$s_i = i \cdot P, \quad i \in [1, N]$$

$$e_i = \begin{cases} (i+1)P + O, & 1 \le i \le N-1 \\ H, & i = N \end{cases} \tag{13}$$

$$\mathcal{L} = \frac{1}{N} \sum_{i=0}^{N-1} \mathcal{L}_{CE} \left(\text{CRUHI} - \text{Mamba} \left(X_{[s_i:e_i,W,V]} \right), Y_{[s_i:e_i,W,V]} \right) \tag{14}$$

where s_i and e_i denote the start and end positions of the i-th chunk along the H dimension, $i = (1,2,\ldots,N)$, N is the total number of chunks, $P = H/N$ is the average height of each chunk, and O is the number of overlapping rows. $X_{[\cdot]}$ and $Y_{[\cdot]}$ represent the training data and corresponding labels for each chunk, respectively. $\mathcal{L}_{CE}$ is the cross-entropy loss function.

In the validation and testing phase, we adopt the chunked linear weighted averaging (CLWA) strategy to reduce the stitching artifacts between neighbouring chunks and improve the consistency of boundary predictions. The weighting function $w_i(x)$ and the stitching function $\hat{Y}(x)$ are calculated using the following equation:

$$w_i(x) = \begin{cases} 1 - \frac{0.5(x - s_i)}{e_i - s_i}, & i = 1,\ x \in [s_1, e_1] \\ 0.5 + \frac{0.5(x - s_i)}{(P+O)/2}, & 2 \le i \le N-1,\ x \in \left[s_i,\ s_i + \frac{P+O}{2} \right) \\ 1 - \frac{0.5(x - (s_i + (P+O)/2))}{(P+O)/2}, & 2 \le i \le N-1,\ x \in \left[s_i + \frac{P+O}{2},\ e_i \right] \\ 0.5 + \frac{0.5(x - s_i)}{H - s_i}, & i = N,\ x \in [s_N, e_N] \end{cases} \tag{15}$$

$$\hat{Y}(x) = \frac{\sum_{i=0}^{N-1}(w_i(x)\cdot \text{CRUHI} - \text{Mamba}(X_{[s_i:e_i,\ W,V]}))}{\sum_{i=0}^{N-1}(w_i(x))} \tag{16}$$

where $w_i(x)$ for the first chunk decreases linearly from 1 to 0.5 between the starting position s_1 to the ending position e_1. Each middle chunk ($2 \le i \le N-1$) is equally divided into two parts: the weight increases from 0.5 to 1 in the first half and decreases from 1 to 0.5 in the second half. The weight of the last chunk is set in the opposite manner to the first chunk, with its weight increasing linearly from 0.5 to 1, as shown in Figure 3. The chunk-by-chunk output of CRUHI-Mamba is multiplied by their corresponding weights w_i and normalised. Finally, the spatially continuous and accurate prediction result of the entire coral reef UHI, denoted as $\hat{Y}$ is obtained. The CLWA strategy not only improves the model efficiency but also effectively reduces the stitching artifacts and ensures the continuity and consistency of the prediction results.

4. Experiments setup

4.1. Description of datasets

The benthic classification experiments were conducted on three coral reef UHIs, indexed as 005 (depth 6 m), 006 (depth 3 m), and 019 (depth 9 m) (Figure 4), which were acquired by a diver-operated hyperspectral imaging system (HyperDiver) at different water depths at the Carmabi site on the island of Curaçao in the Caribbean Sea (Rashid and Chennu 2020). During the surveys, the seafloor depth remained approximately constant, and only ambient natural light was used. These UHIs contain diverse benthic types and are suitable for evaluating the classification performance of the proposed method. The benthic types and sample statistics are summarised in Table 1, and the benthic classification scheme follows that of (Schürholz and Chennu 2023). The ground truth labels were annotated by marine biology experts as regions of interest (ROIs) on natural-colour renderings of the hyperspectral images, assisted by synchronised underwater video. Each ROI was assigned the deepest possible taxonomic/substrate label following the World Register of Marine Species (WoRMS) and the CATAMI hierarchical schemes. A deliberately biased annotation strategy was adopted to maximise class coverage with minimal effort, leading to severe class imbalance across the images (Table 1). The type "Material" in Table 1 refers to any non-biotic, anthropogenic, or unidentified objects found in the hyperspectral transect surveys besides float, reel, refboard, and tape (e.g. plastic fragments, discarded fishing gear, metal pieces).

Therefore, for each UHI, we select labelled samples via class-balanced random sampling, with a fixed number of samples drawn per class (Table 2 for the full dataset and experimental details). In preprocessing, infra-red bands carrying no useful signal are discarded; only the 350 bands within the 400–750 nm range are retained for benthic classification. No radiometric calibration, reflectance conversion, interpolation, or spectral smoothing was applied. The original sensor digital number (DN) data, without physical unit conversion, are employed for coral reef benthic classification (Figure 5). Additionally, the UHI data were rectified to approximately square pixels without geographic reference. Each UHI is about 50 m long and 1 m

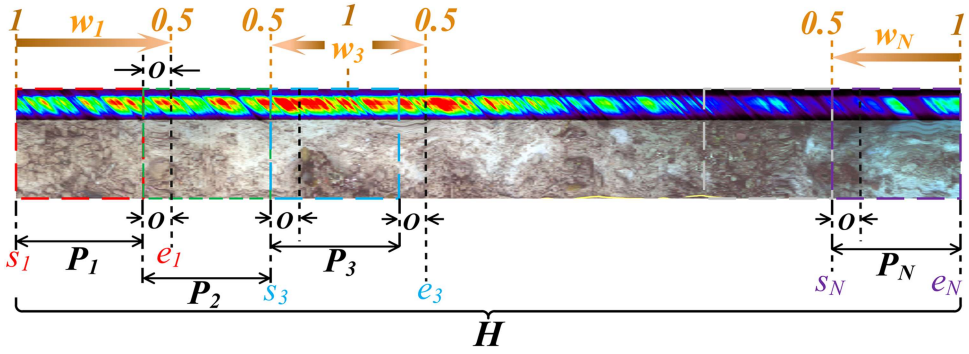

Figure 3. Schematic representation of the chunked linear weighted averaging (CLWA) strategy used for the training and inference process. CLWA divides the underwater hyperspectral image (UHI) into overlapping chunks, applies linear weighting to the outputs of each chunk, and stitches them to produce a seamless full-image classification result.

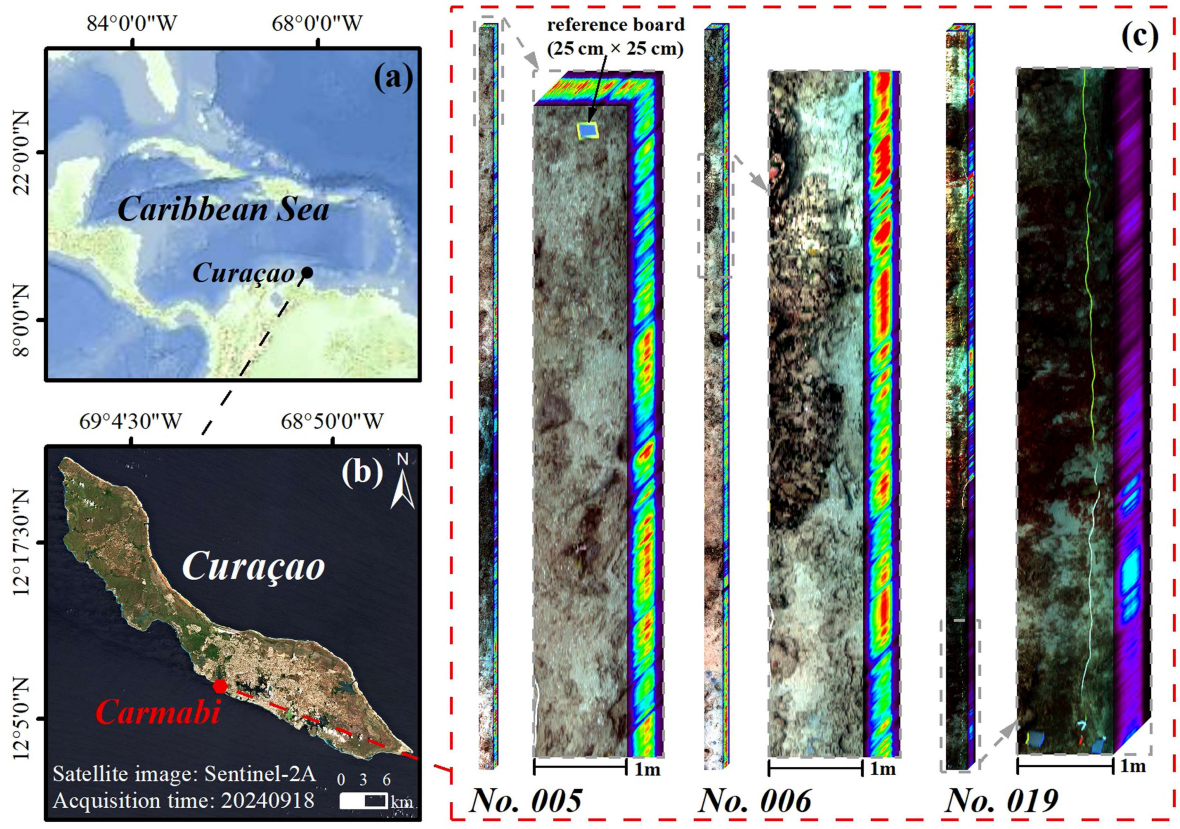

Figure 4. Standard RGB synthetic true-colour map of hyperspectral data and acquisition locations, (a) location of Curaçao in the Caribbean Sea, (b) location of the Carmabi survey site, and (c) three underwater hyperspectral images (UHIs) and local magnification. Each full UHI is approximately 50 m long and 1 m wide, with a cross-track ground sampling distance (GSD) of ~1.6 mm.

wide, with a fixed width of 640 pixels, giving a cross-track ground sampling distance (GSD) of ~1.6 mm (i.e. each pixel corresponds to ~1.6 mm on the seafloor). The along-track GSD varies slightly with swimming speed. A 25 cm × 25 cm grey plastic reference board is present at one or both ends of each UHI for scale validation. Acquisition at near constant altitude and compensation for cross-currents further reduce geometric distortions (Rashid and Chennu 2020).

4.2. Implementation details

The CRUHI-Mamba model was implemented in PyTorch and executed on identical hardware: Hygon C86-3G CPU@3.00 GHz, NVIDIA GeForce RTX4090D GPU, and 128 GB RAM. For each experiment, 135 labelled pixels were used for training, 30 for validation, with the remaining samples reserved for testing (Table 2). The batch size was fixed at 1, as CRUHI-Mamba processes the entire coral reef UHI and predicts the full image simultaneously. Training was conducted for 200 epochs using Adam optimiser with a learning rate of 0.0001. In the $SSMB_{Spe}$ block, the spectral grouping number m was set to 8, and the GN groups G to 16. The Mamba-SSM version used was 1.2.0, with an embedding dimension $D = 64$. The state dimension (d_state), local convolution width (d_conv), and expansion factor (expand) for $SSMB_{Spe}$ were set to 16, 4, and 1, respectively, and for $SSMB_{Spa}$ to 16, 4, and 4. All other Mamba block parameters followed default settings.

4.3. Comparison methods and evaluation metrics

The classification accuracy and computational cost of CRUHI-Mamba were compared with eight hyperspectral image classification methods: Random Forest (RF); the CNN-based 3-DCNN (Ben Hamida et al. 2018), FullyContNet (Wang, Du, and Zhang 2022), and SSRN (Zhong et al. 2018); the Transformer-based

Table 1. Benthic types and sample statistics of UHI of coral reefs, "-" indicates type absence.

	No. 005 UHI		No. 006 UHI		No. 019 UHI	
Type	Class	No. samples	Class	No. samples	Class	No. samples
Agaricia	–	–	1	3397	1	673
Aplysina cauliformis	1	3438	–	–	–	–
Diploria labyrinthiformis	2	9703	2	7804	2	496
Diploria strigosa	3	19706	3	11227	3	5292
Meandrina meandrites	4	27324	4	9759	4	2726
Montastraea cavernosa	–	–	–	5	609	
Porites astreoides	5	6650	5	3760	–	–
Siderastrea siderea	6	9210	6	4785	6	3816
Millepora	7	4235	7	4885	–	–
Zoanthid	–	–	–	–	7	2067
Agelas conifera	–	–	8	456	8	12626
Aiolochroia crassa	–	–	–	–	9	2240
Desmapsamma anchorata	8	2539	–	–	–	–
Ircinia campana	9	4915	–	–	–	
Neofibularia nolitangere	10	59950	–	–	10	6789
Corallinales	–	–	9	3543	11	1797
Dictyota	11	15863	–	–	12	25656
Lobophora variegata	–	–	–	–	13	1720
Halimeda opuntia	12	9290	10	30517	14	1972
Turf algae	13	60658	11	74988	15	17612
Cyanobacterial mat	–	–	–	–	16	46902
Sediment	14	91004	12	80463	17	25674
Float	–	–	–	–	18	1033
Material	–	–	–	–	19	10548
Reel	–	–	–	–	20	2781
Refboard	15	6291	13	6560	21	10497
Tape	16	863	14	1972	22	6128
Total	16	331639	14	244116	22	189654

Table 2. Detailed information of experimental data for UHI classification of coral reefs.

No. UHI	Spatial size (pixels)	Approx. ground size (m)	Depth (m)	No. Class	No. of samples per class		Ratio of train and validation to total sample
					Train	Validation	
005	35235 × 640	50 × 1	6	16	135	30	0.79%
006	27345 × 640		3	14			0.95%
019	22198 × 640		9	22			1.91%

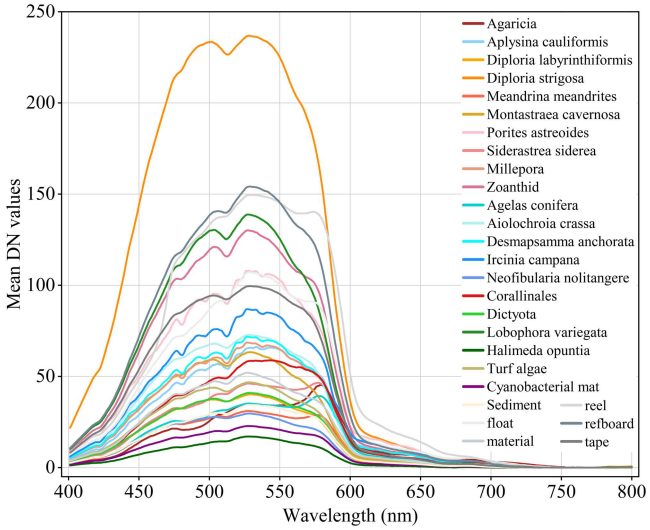

Figure 5. Mean digital number (DN) spectral curves of the 27 benthic types across the three underwater hyperspectral images (UHI No. 005, 006, and 019).

Spectralformer (Hong et al. 2022), SSFTT (Sun et al. 2022), and GSC-ViT (Zhao et al. 2024) and the Mamba-based MambaHSI (Li et al. 2024). SSRN is a leading deep learning model for UHI classification. All models were trained and tested under the same conditions, with hyperparameters set according to their respective original papers.

Model performance was evaluated using three standard hyperspectral classification metrics: overall accuracy (OA), average accuracy (AA), and the Kappa coefficient. Time complexity was assessed by training and inference time, and space complexity by the floating-point operations (FLOPs) and the number of trainable parameters. Each experiment was repeated five times with different random seeds to reduce the effects of sample selection and parameter initialisation. Higher mean values and lower standard deviations (STD) indicate better performance.

5. Results

5.1. Experimental comparison with competitive approaches

Based on the quantitative results in Tables 3 and 4, it is evident that the proposed CRUHI-Mamba performs the best among all methods. It achieves the highest OA, AA, and Kappa scores across all three UHI datasets of coral reefs, with OA exceeding 99%, which demonstrates the superior classification capability of the proposed model. On the 005 hyperspectral image, the highest classification accuracy for the 12 benthic types is achieved using only 0.79% of the labelled samples, with an OA of 99.56%. CRUHI-Mamba outperforms RF by more than 8.9% in OA, primarily due to RF's limitations in using hand-crafted features and its inability to capture spatial feature information. The OA of the proposed model is on average more than 3.3% and 5.9% higher than those of the CNN-based and Transformer-based models, respectively (Table 3). These models struggle to establish long-term dependencies due to their limited receptive fields and the underutilisation of detailed spectral interactions and global spatial information of coral reef benthic types. These limitations result in lower classification accuracy.

For the 006 image (Table 4), using only 0.95% of labelled samples for training and validation, the proposed model achieves an accuracy of $99.23 \pm 0.14\%$. Compared with the second-best method, MambaHSI, CRUHI-Mamba improves OA, AA, and Kappa by approximately 1.18%, 0.89%, and 1.53%, respectively. On the 019 image, trained with 1.91% of labelled samples, CRUHI-Mamba again achieves the best performance, with OA improvements of 8.98% and 5.81% over CNN- and Transformer-based models, and corresponding Kappa gains of 10.18% and 6.60% (Table 4). Notably, CRUHI-Mamba and MambaHSI achieve OA values exceeding 98% across all three images (Tables 3 and 4), whereas the accuracy of other methods varies significantly, indicating the superior and more stable performance of the Mamba-based models. Although CRUHI-Mamba achieves only 0.29%–1.14% higher OA than MambaHSI, the latter has a more complex structure and lower feature extraction efficiency. Moreover, in training efficiency, inference speed, and parameter scale, MambaHSI performs significantly worse than CRUHI-Mamba, as quantitatively shown in Section 5.3.

5.2. Qualitative assessment of benthic classification maps

The qualitative visual results presented in Figures 6–8 indicate that RF-generated classification maps are affected by noticeable salt-and-pepper noise. The classification maps produced by the CNN-based FullyContNet show the poorest performance, with extensive misclassified regions and incomplete identification of benthic types. The classification maps generated by Transformer-based models exhibit fuzzy boundaries and considerable noise in localised regions. The Spectralformer model exhibits significant misclassifications among coral types, such as *Meandrina meandrites*, *Porites astreoides*, and *Millepora*, in UHI 005. MambaHSI cannot process the entire hyperspectral image with fewer segments during training and inference due to its complex architecture and high computational demands, leading to more visible stitching artifacts in the classification map that compromise its spatial continuity. In the enlarged local regions (Figure 6–8), the proposed CRUHI-Mamba exhibits the fewest misclassifications and aligns most closely with the ground truth, as highlighted by the bright blue boxes. The resulting high-quality coral reef benthic classification maps show smooth and clear boundaries with minimal noise, demonstrating the effectiveness and superiority of the proposed method. In summary, compared with all competing methods, the proposed CRUHI-Mamba achieves the best benthic classification performance on coral reef UHI in terms of both accuracy and visual quality.

Table 3. Classification accuracies of different methods on UHI No. 005, presented in the "Accuracy% ± Standard deviations (STD)%" format. **Black** bold is the best.

Class	Conventional	CNN-based methods			Transformer-based methods			Mamba-based methods	
	RF	SSRN	3DCNN	FullyContNet	Spectralformer	SSFTT	GSC-ViT	MambaHSI	CRUHI-Mamba
1	96.8 ± 0.79	99.07 ± 1.13	99.58 ± 0.40	**100**	97.48 ± 1.28	99.82 ± 0.25	99.99 ± 0.02	**100**	**100**
2	80.46 ± 1.74	97.28 ± 1.51	92.27 ± 1.35	99.41 ± 0.26	82.81 ± 1.80	97.43 ± 1.21	98.41 ± 0.81	99.56 ± 0.07	**99.86 ± 0.17**
3	90.78 ± 1.06	94.33 ± 6.24	96.47 ± 1.29	99.72 ± 0.63	87.73 ± 1.85	99.13 ± 0.76	**99.73 ± 0.20**	99.54 ± 0.11	99.72 ± 0.30
4	89.19 ± 1.41	95.34 ± 2.25	96.92 ± 0.37	96.23 ± 0.61	80.82 ± 2.95	98.40 ± 0.53	98.83 ± 0.70	99.02 ± 0.66	**99.23 ± 0.26**
5	88.76 ± 1.16	97.94 ± 0.36	96.16 ± 1.17	87.94 ± 5.27	81.11 ± 1.72	98.11 ± 1.10	97.24 ± 3.89	99.16 ± 0.06	**99.33 ± 0.54**
6	99.2 ± 0.21	99.49 ± 0.28	99.62 ± 0.17	98.98 ± 0.44	98.37 ± 0.46	99.81 ± 0.11	99.69 ± 0.43	99.97 ± 0.05	**99.99 ± 0.02**
7	84.6 ± 1.28	94.81 ± 2.90	94.45 ± 0.60	98.53 ± 0.85	82.38 ± 3.29	97.39 ± 1.33	**98.69 ± 1.28**	98.57 ± 0.33	98.29 ± 0.44
8	98.58 ± 1.03	99.95 ± 0.08	99.88 ± 0.22	99.75 ± 0.35	99.65 ± 0.29	99.79 ± 0.18	99.97 ± 0.07	**100**	99.70 ± 0.37
9	98.3 ± 0.46	**100**	99.97 ± 0.03	**100**	97.46 ± 1.20	99.93 ± 0.14	99.88 ± 0.21	**100**	**100**
10	90.02 ± 0.42	98.74 ± 1.36	97.20 ± 1.01	99.96 ± 0.25	74.19 ± 3.07	97.39 ± 0.42	98.13 ± 1.11	**100**	99.99 ± 0.02
11	82.76 ± 2.27	88.28 ± 3.55	91.27 ± 1.33	85.28 ± 7.08	58.88 ± 3.49	94.39 ± 1.00	92.28 ± 6.50	98.77 ± 0.74	**99.35 ± 0.59**
12	86.79 ± 1.38	96.10 ± 1.44	94.08 ± 1.50	98.6 ± 0.67	84.78 ± 1.81	96.98 ± 1.23	94.40 ± 6.52	99.28 ± 0.92	**99.69 ± 0.14**
13	81.59 ± 2.61	93.38 ± 4.34	92.98 ± 1.50	86.82 ± 3.63	86.15 ± 5.22	93.86 ± 1.63	95.70 ± 1.35	98.38 ± 0.54	**99.35 ± 0.32**
14	98.15 ± 0.27	96.71 ± 2.34	98.36 ± 0.22	95.71 ± 1.15	98.04 ± 1.35	97.76 ± 0.63	97.71 ± 1.20	99.29 ± 0.51	**99.42 ± 0.22**
15	**100**	**100**	**100**	99.27 ± 0.27	**100**	**100**	**100**	**100**	**100**
16	99.78 ± 0.21	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
OA(%)	90.6 ± 0.38	96.05 ± 0.65	96.32 ± 0.51	96.41 ± 0.88	86.33 ± 1.53	97.09 ± 0.36	97.47 ± 0.33	99.27 ± 0.07	**99.56 ± 0.05**
AA(%)	91.61 ± 0.22	96.96 ± 0.53	96.83 ± 0.45	97.35 ± 0.59	88.12 ± 0.74	98.14 ± 0.09	98.17 ± 0.70	99.47 ± 0.13	**99.62 ± 0.06**
Kappa(%)	88.85 ± 0.45	95.30 ± 0.77	95.63 ± 0.60	95.74 ± 1.05	83.85 ± 1.82	96.54 ± 0.43	97.00 ± 0.39	99.13 ± 0.08	**99.48 ± 0.06**

Table 4. Overall classification performance of different methods on UHI No. 006 and No. 019, presented in the "Accuracy% ± Standard deviations (STD)%" format. **Black** bold is the best. Full per-class accuracies are provided in Supplementary Tables S1 and S2.

			CNN-based methods			Transformer-based methods			Mamba-based methods	
Method		Conventional RF	SSRN	3DCNN	FullyContNet	Spectralformer	SSFTT	GSC-ViT	MambaHSI	CRUHI-Mamba
UHI	OA(%)	89.05 ± 0.7	95.02 ± 0.97	94.60 ± 1.55	95.23 ± 1.4	96.06 ± 1.26	94.07 ± 0.96	95.87 ± 0.41	98.05 ± 0.85	**99.23 ± 0.14**
No.	AA(%)	95.71 ± 0.23	98.29 ± 0.31	98.15 ± 0.37	91.52 ± 3.42	98.73 ± 0.44	98.40 ± 0.20	98.92 ± 0.12	98.89 ± 1.2	**99.78 ± 0.04**
006	Kappa(%)	85.93 ± 0.89	93.58 ± 1.23	93.04 ± 1.99	93.87 ± 1.79	94.92 ± 1.62	92.34 ± 1.23	94.66 ± 0.53	97.47 ± 1.1	**99.0 ± 0.18**
UHI	OA(%)	86.21 ± 0.45	87.27 ± 1.98	92.79 ± 0.59	91.5 ± 1.03	95.98 ± 0.66	94.63 ± 0.30	90.46 ± 3.04	98.82 ± 0.25	**99.5 ± 0.15**
No.	AA(%)	92.73 ± 0.13	93.62 ± 1.32	97.26 ± 0.19	94.91 ± 1.49	98.42 ± 0.34	97.92 ± 0.10	95.74 ± 2.08	99.56 ± 0.12	**99.84 ± 0.04**
019	Kappa(%)	84.39 ± 0.51	85.57 ± 2.20	91.80 ± 0.66	90.36 ± 1.16	95.42 ± 0.75	93.88 ± 0.34	89.17 ± 3.40	98.65 ± 0.29	**99.43 ± 0.17**

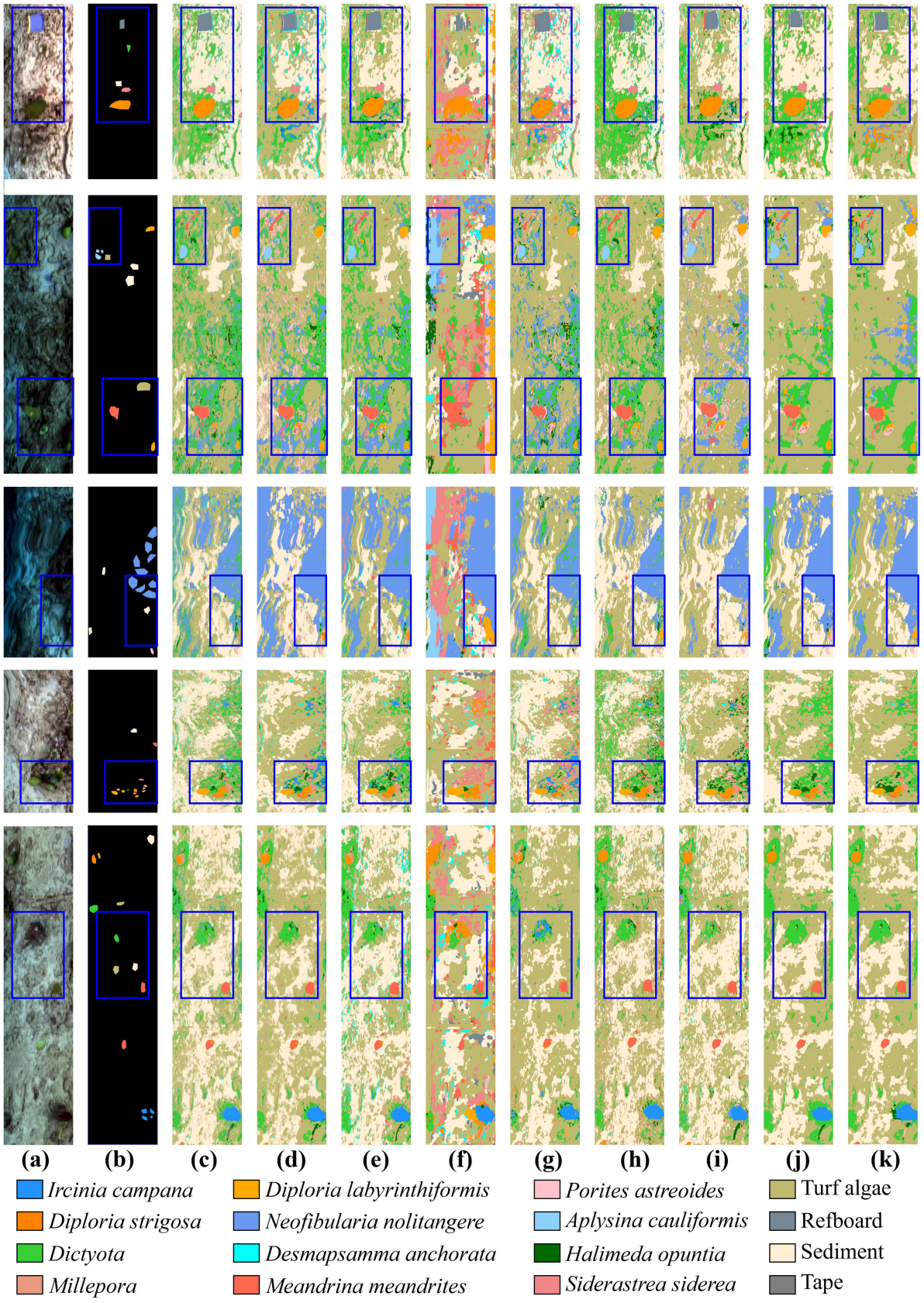

Figure 6. Enlarged local regions of benthic classification maps for UHI No. 005. (a) Standard RGB synthetic natural image, (b) Ground truth, (c) RF, (d) SSRN, (e) 3DCNN, (f) FullyContNet, (g) Spectralformer, (h) SSFTT, (i) GSC-ViT, (j) MambaHSI, (k) Proposed CRUHI-Mamba. Bright blue boxes highlight areas where the proposed method outperforms other models. Full classification maps are provided in Supplementary Figure S1.

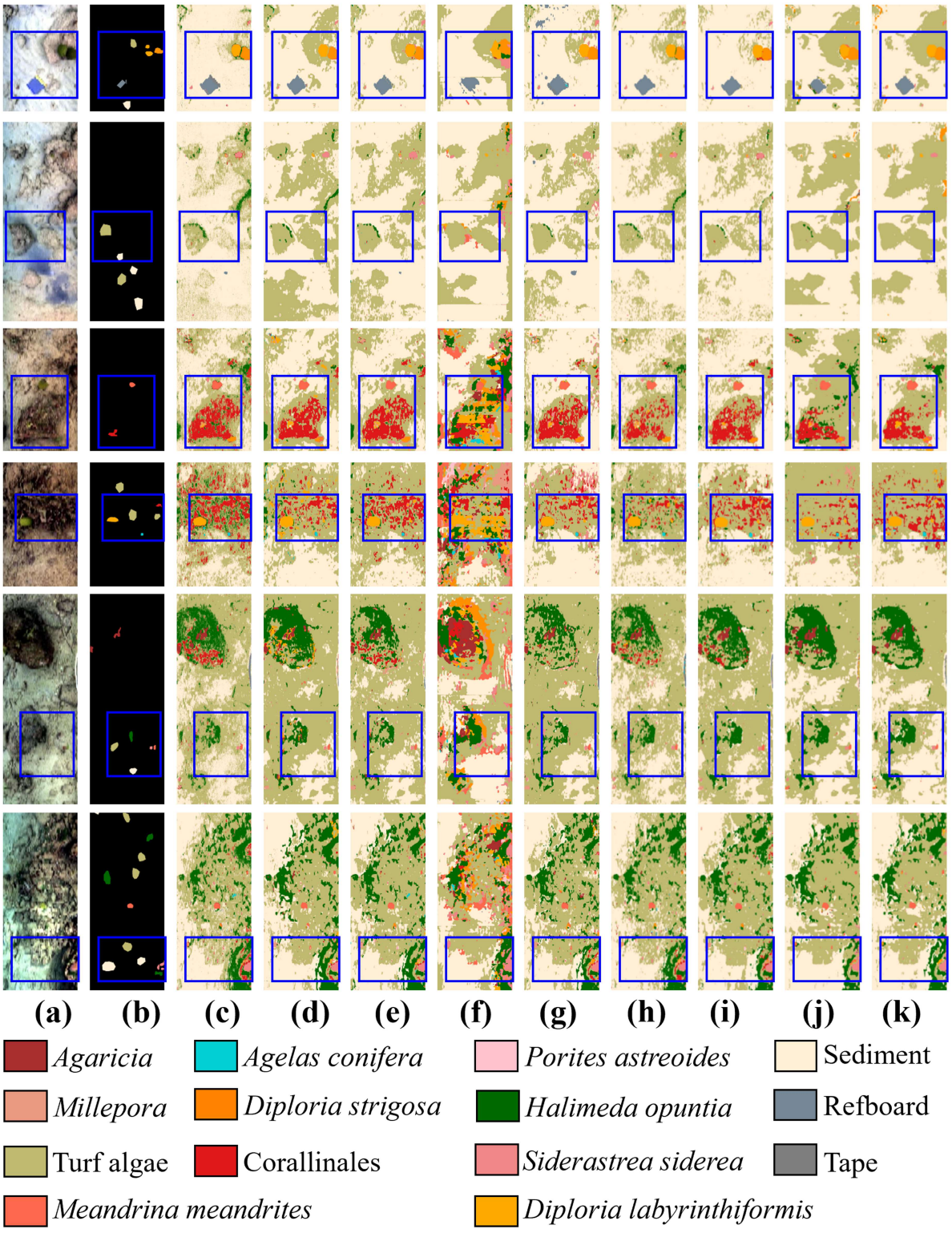

Figure 7. Enlarged local regions of benthic classification maps for UHI No. 006. (a) Standard RGB synthetic natural image, (b) Ground truth, (c) RF, (d) SSRN, (e) 3DCNN, (f) FullyContNet, (g) Spectralformer, (h) SSFTT, (i) GSC-ViT, (j) MambaHSI, (k) Proposed CRUHI-Mamba. Bright blue boxes highlight areas where the proposed method outperforms other models. Full classification maps are provided in Supplementary Figure S2.

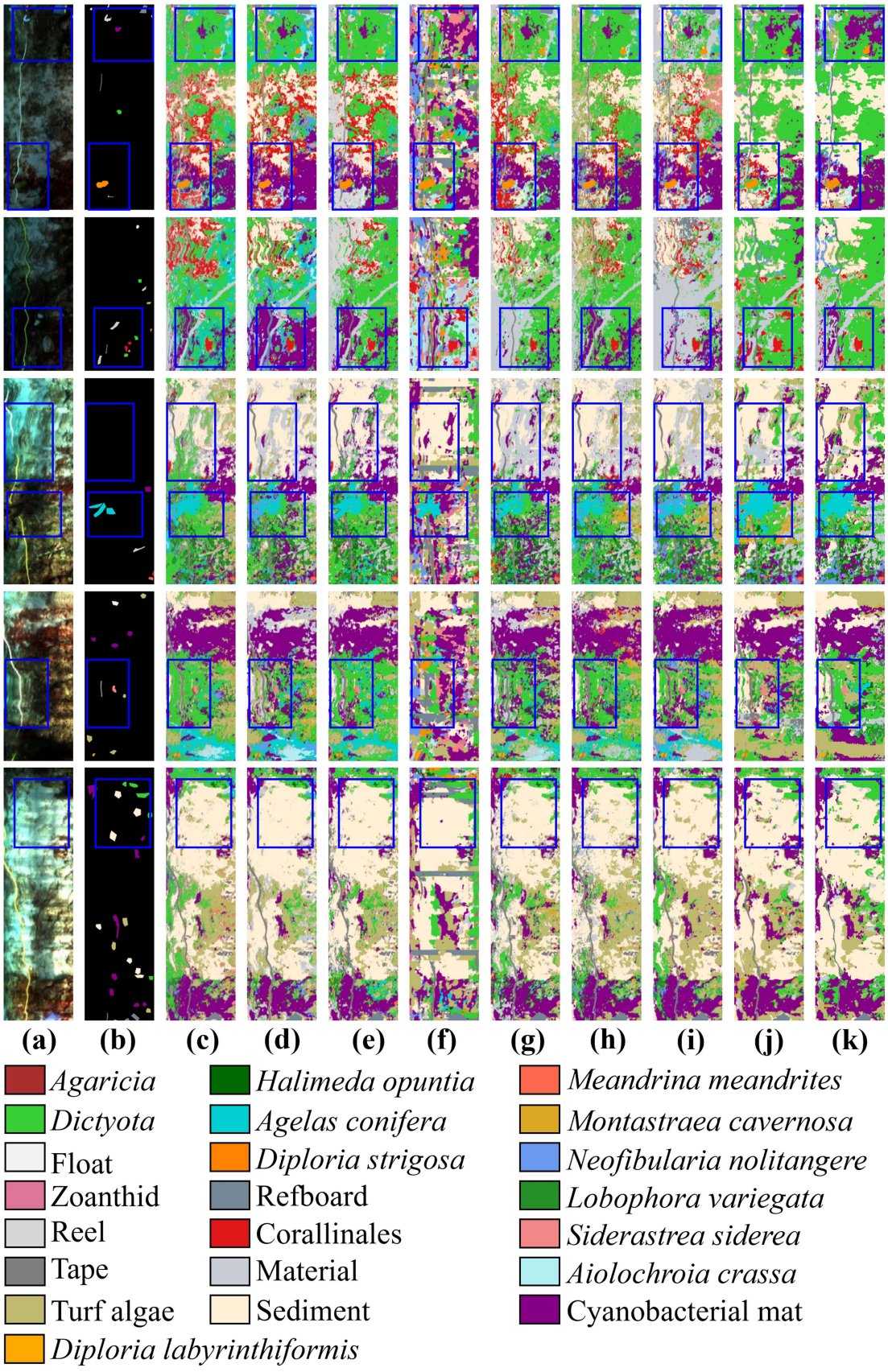

Figure 8. Enlarged local regions of benthic classification maps for UHI No. 019. (a) Standard RGB synthetic natural image, (b) Ground truth, (c) RF, (d) SSRN, (e) 3DCNN, (f) FullyContNet, (g) Spectralformer, (h) SSFTT, (i) GSC-ViT, (j) MambaHSI, (k) Proposed CRUHI-Mamba. Bright blue boxes highlight areas where the proposed method outperforms other models. Full classification maps are provided in Supplementary Figure S3.

5.3. Computational complexity and efficiency comparison

The advantages of CRUHI-Mamba extend beyond classification accuracy. We also evaluate the model's parameters, runtime, and FLOPs. In terms of parameter count and OA, CRUHI-Mamba achieves the best results, reducing parameters by an average of 64%, 49%, and 15% compared with the CNN-based, Transformer-based, and Mamba-based models, respectively. This aligns with our design objective of achieving high classification accuracy with minimal parameters (Figure 9). Furthermore, CRUHI-Mamba achieves the shortest inference time for full coral reef UHI—approximately half that of FullyContNet and MambaHSI (also image-level models, Table 5). Notably, its inference time is 1-3 orders of magnitude shorter than other compared models. A comprehensive runtime analysis confirms that CRUHI-Mamba remains highly competitive in classification efficiency.

On UHI 005, MambaHSI exceeds the GPU memory capacity, making a fair comparison of training time infeasible. On the smaller datasets UHI 006 and UHI 019, where both methods fit in GPU memory, MambaHSI still takes over seven times longer to train than CRUHI-Mamba. FLOPs are measured using a synthetic input tensor with the same spectral dimensions as the experimental data (i.e. $1 \times 100 \times 256 \times 256$). Results show that the proposed model achieves the lowest FLOPs, using only one-third of the computations required by the baseline model MambaHSI. These results demonstrate that CRUHI-Mamba not only delivers superior classification performance but also offers a lightweight, efficient structure.

6. Discussion

6.1. Ablation study

We conduct an ablation study to evaluate the contribution of each module, including the $SSMB_{Spa}$ block, the $SSMB_{Spe}$ block, the SSMB, and the AFDFM. Here, $SSMB_{Spa}$ and $SSMB_{Spe}$ refer to the standalone spatial and spectral blocks that exclude the dual-pooling (AvgPool & MaxPool) operation. Based on the ablation results (Table 6), (1) CRUHI-Mamba achieves the highest accuracy across all settings, with all modules

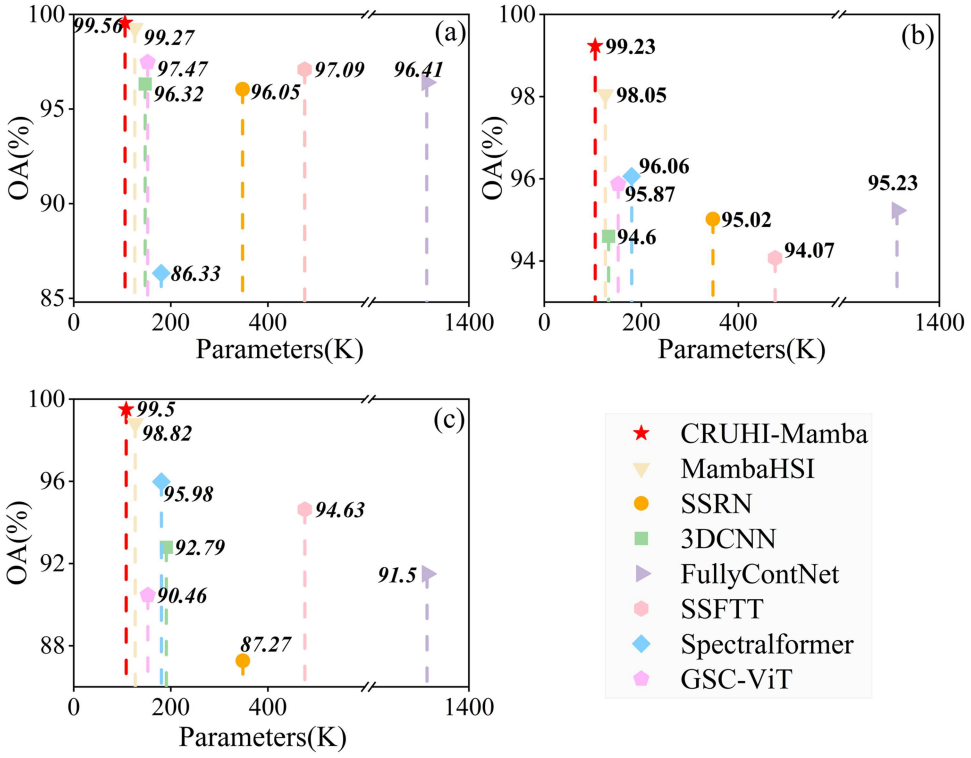

Figure 9. Comparison of the number of trainable parameters (in thousands, K) across different models. (a) Underwater hyperspectral image (UHI) No. 005, (b) UHI No. 006, (c) UHI No. 019.

Table 5. Model Performance Comparison. T_{tr}: training time (min); T_{infer}: inference time for the entire hyperspectral image (min); Floating-point operations (FLOPs): computed on a random tensor of size [1, 100, 256, 256], measured in gigaflops (G).

Data		No. 005 UHI		No. 006 UHI		No. 019 UHI		[1, 100, 256, 256]
Metrics		T_{tr} (min)	T_{infer} (min)	T_{tr} (min)	T_{infer} (min)	T_{tr} (min)	T_{infer} (min)	FLOPs (G)
Conventional	RF	0.01	1.66	0.01	1.19	0.01	1.25	-
CNN-based Methods	SSRN	2.58	20.2	2.33	15.56	3.44	12.72	5.06
	3DCNN	1.49	11.46	1.32	8.91	1.84	7	1.51
	FullyContNet	98.93	0.2	54	0.15	45.12	0.12	48.99
Transformer-based Methods	Spectralformer	8.48	125.55	8.15	77.63	10.7	55.49	5.91
	SSFTT	3.43	16.39	2.28	12.02	3.13	12.098	60.89
	GSC-ViT	2.95	27.22	2.49	22.73	3.79	17.26	9.94
Mamba-based Methods	MambaHSI	2107.06	0.34	152.66	0.26	132.01	0.22	4.61
	CRUHI-Mamba	22.22	**0.14**	19.76	**0.13**	16.85	**0.08**	**1.4**

contributing to performance gains and yielding OA improvements exceeding 5% on all three datasets. (2) $SSMB_{Spa}$ consistently outperforms $SSMB_{Spe}$, suggesting that spatial context and morphological structure are more discriminative than spectral information for coral reef benthic classification, as spectral features are susceptible to water-column absorption and illumination variations. (3) Directly combining the spatial and spectral features extracted by SSMB improves classification accuracy while reducing boundary artifacts and isolated misclassifications observed in spatial-only and spectral-only results (Figure 10 $SSMB_{Spa}$ and $SSMB_{Spe}$ columns). Although the SSMB + (Spaeq\o\ac(○,+)Spe) results exhibit clear boundaries and coherent regions, misclassified areas remain, particularly for low–information-density classes such as coral. (4) With the introduction of AFDFM, classification accuracy is further improved, and previously misclassified regions are correctly recovered (Figure 10, blue boxes). In particular, rare classes, including different corals, Refboards, and Tapes, are classified more accurately, consistent with the visible-light image. This demonstrates that AFDFM effectively alleviates classification bias and fusion issues caused by uneven information density through adaptive multi-level feature fusion, leading to more accurate and robust classification.

6.2. Effect of data dimensionality reduction on classification performance

We investigate the impact of raw DN data and PCA-reduced representations of coral reef UHIs on the classification performance of the proposed CRUHI-Mamba. Experimental results (Figure 11) show that OA exceeds 98.9% across all experiments, indicating strong overall performance. Furthermore, nearly all PCA-reduced datasets outperform raw DN data. This improvement is not solely due to noise suppression or redundancy removal, but more likely to PCA-enhanced feature representation. By reducing band dimensionality while emphasising high-energy and discriminative features, PCA improves the retention of key information in the state space, thereby alleviating memory decay in selective state space models during long-sequence modelling in CRUHI-Mamba and improving both classification accuracy and training efficiency. Experiments with varying numbers of PCA components further indicate that performance is sensitive to PCA dimensionality, peaking at 100 components for UHI 005 and 019 and at 50 components for UHI 006. Overall, the model achieves more stable and superior performance when the number of principal components lies between 50 and 100.

6.3. Effect of the number of training samples on classification results

Based on the PCA-reduced data from UHI No. 005, we analyse the impact of training sample size on benthic classification performance, where each training sample corresponds to a single pixel. As the number of training samples per class increases from 10 to 210, with 30 fixed validation samples, the following conclusions are drawn (Figure 12a): (1) Initially, OA, AA, and Kappa improve rapidly with increasing training samples. (2) The highest OA of 99.56% is achieved at 135 training samples, corresponding to a labelled data proportion of only 0.79%. Beyond this point, the OA remains stable at around 99.5%, indicating that CRUHI-Mamba maintains excellent performance with less than 1% of the data being labelled, demonstrating

Table 6. Accuracy of ablation experiments on the three datasets, presented in the "Accuracy% ± Standard deviations (STD)%" format. **Black** bold is the best.

					No. 005 UHI			No. 006 UHI			No. 019 UHI		
SSMB_spa	SSMB_spe	SSMB	Spaeq\o\ac(○,+)Spe	AFDFM	OA	AA	Kappa	OA	AA	Kappa	OA	AA	Kappa
✓					98.44 ± 0.1	98.81 ± 0.14	98.15 ± 0.12	98.04 ± 0.23	99.22 ± 0.1	97.47 ± 0.29	98.15 ± 0.36	99.23 ± 0.15	97.89 ± 0.41
	✓				95.04 ± 0.48	95.57 ± 0.4	94.1 ± 0.57	95.2 ± 0.69	97.97 ± 0.27	93.81 ± 0.88	94.29 ± 0.49	97.13 ± 0.23	93.5 ± 0.55
		✓	✓		99.23 ± 0.21	99.4 ± 0.1	99.08 ± 0.25	99.13 ± 0.33	99.73 ± 0.09	98.87 ± 0.42	99.22 ± 0.24	99.73 ± 0.1	99.11 ± 0.27
		✓		✓	**99.56** ± 0.05	**99.62** ± 0.06	**99.48** ± 0.06	**99.23** ± 0.14	**99.78** ± 0.04	**99.0** ± 0.18	**99.5** ± 0.15	**99.84** ± 0.04	**99.43** ± 0.17

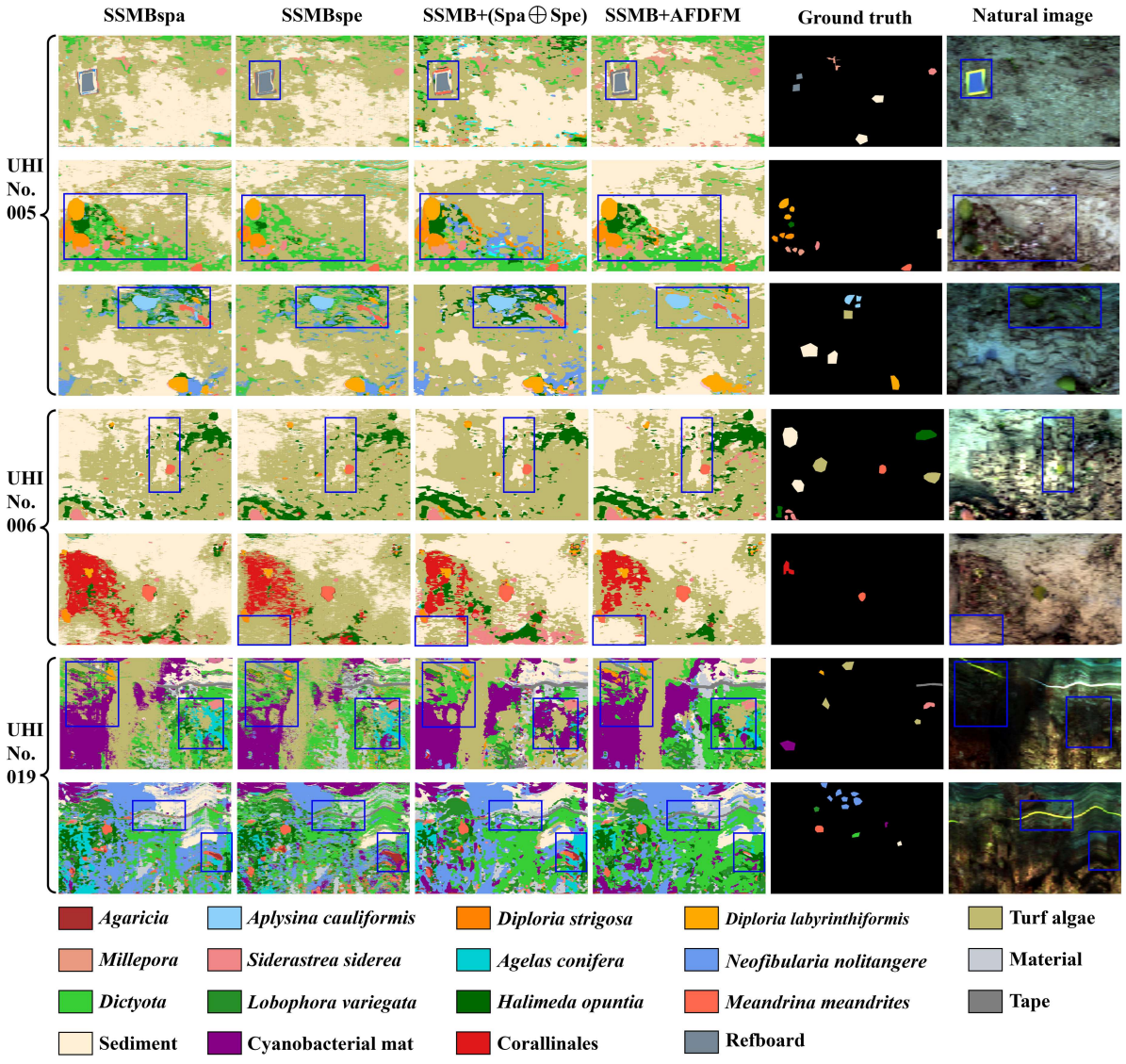

Figure 10. Partial classification results from the ablation study. The natural images are standard RGB synthetic composites of the corresponding UHIs. The source UHI for each cropped local classification result is labelled on the left with braces.

efficient data utilisation. (3) The model's training time remains stable at 22–27 minutes (Figure 12a), regardless of training size, indicating strong scalability in practical applications.

6.4. Experimental parameters analysis

6.4.1. The number of chunks N and the embedding dimension D

Using PCA-reduced data from UHI No. 005, we evaluate how the UHI chunk number N (applied in training and inference) and embedding dimension D impact classification performance. As N increases, classification accuracy improves while training time decreases, and both metrics stabilise at $N = 30$ (Figure 12b), demonstrating the effectiveness of the chunked input design during training and the CLWA-based output stitching during inference. However, larger N introduces more unlabelled chunks and spatial boundaries, leading to residual stitching artifacts even with CLWA. Therefore, $N = 30$ is selected as the default setting. Increasing D yields rapid performance gains initially, but accuracy improvements become marginal beyond $D = 64$, while training time increases sharply and exhibits instability (Figure 12c). When $D \geq 128$, accuracy

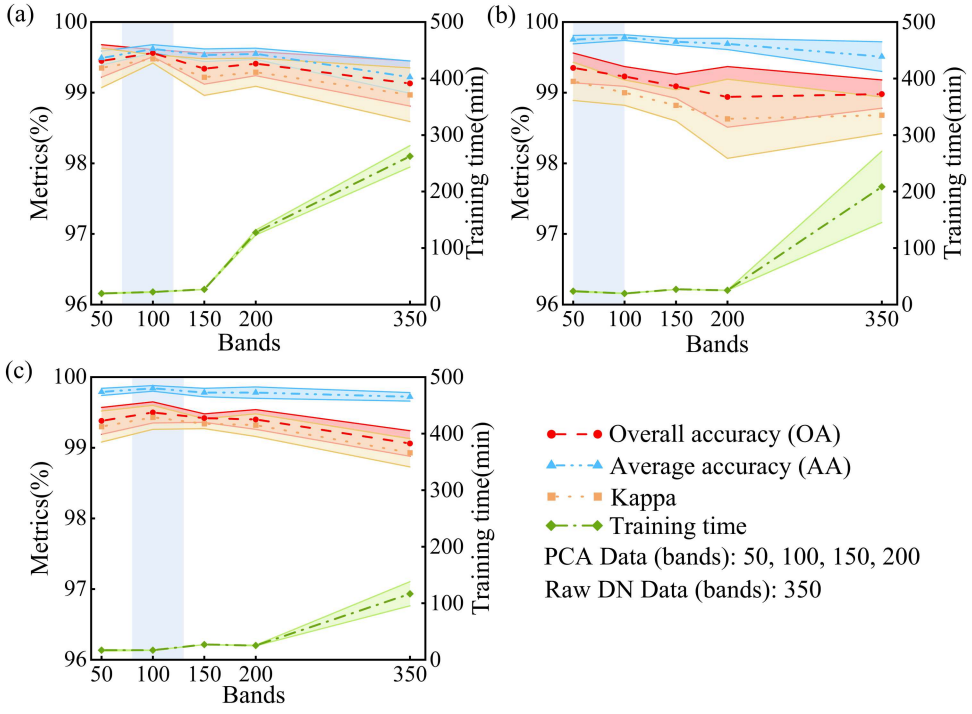

Figure 11. Classification performance of raw digital number (DN) data and PCA-reduced data with different numbers of principal components. Results are presented as Accuracy% ± Standard deviations (STD)% over five runs. (a) Underwater hyperspectral image (UHI) No. 005, (b) UHI No. 006, (c) UHI No. 019. The optimal range of principal components is highlighted.

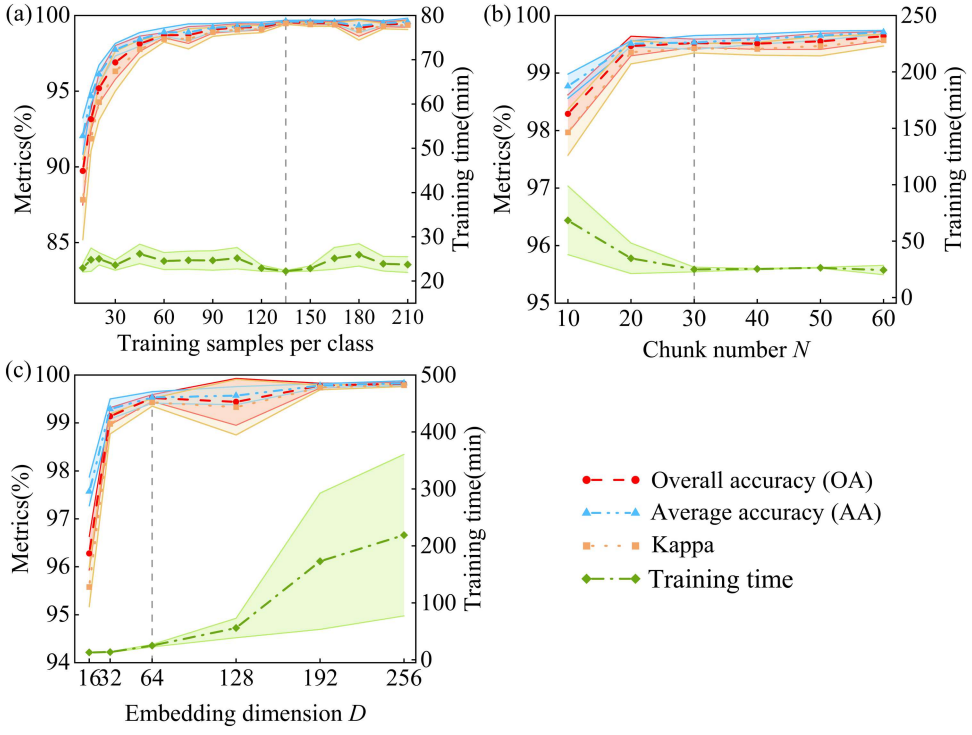

Figure 12. Effect of different parameters on classification performance based on underwater hyperspectral image (UHI) No. 005. (a) Number of training samples. (b) Number of training/inference chunks *N*. (c) Embedding dimension *D*. Results are presented as Accuracy% ± Standard deviations (STD)% over five runs. The vertical dashed lines indicate the optimal parameter values balancing accuracy and training time.

shows little improvement over $D = 64$, but training time increases by two to eight times. Consequently, $D = 64$ is adopted as the default embedding dimension.

6.4.2. $SSMB_{spe}$ groups m and GN groups G

The effects of the number of $SSMB_{spe}$ groupings m and the number of GN groups G on classification accuracy and training time are shown in Figure 13. These parameters are determined in an empirical, data-driven manner. When the number of $SSMB_{spe}$ groups m is fixed, OA gradually increases as the number of GN groups G increases, and eventually stabilises. In contrast, the training time does not exhibit a clear trend. Conversely, when G is held constant, OA remains relatively stable as m increases, while training time gradually increases. Therefore, considering both computational efficiency and classification accuracy, $m = 8$ and $G = 16$ are adopted as the default parameters, achieving high classification performance with an acceptable computational cost.

6.5. Applicability and robustness analysis

From the perspective of underwater environmental conditions and data characteristics, although the experimental data in this study were collected from relatively clear waters, water depth can be regarded as an integrated proxy for multiple environmental factors, leading to noticeable variations in illumination conditions and benthic class distributions across depths. As demonstrated in Section 5, CRUHI-Mamba exhibits consistent and stable performance at depths of 3 m, 6 m, and 9 m, indicating its robustness and

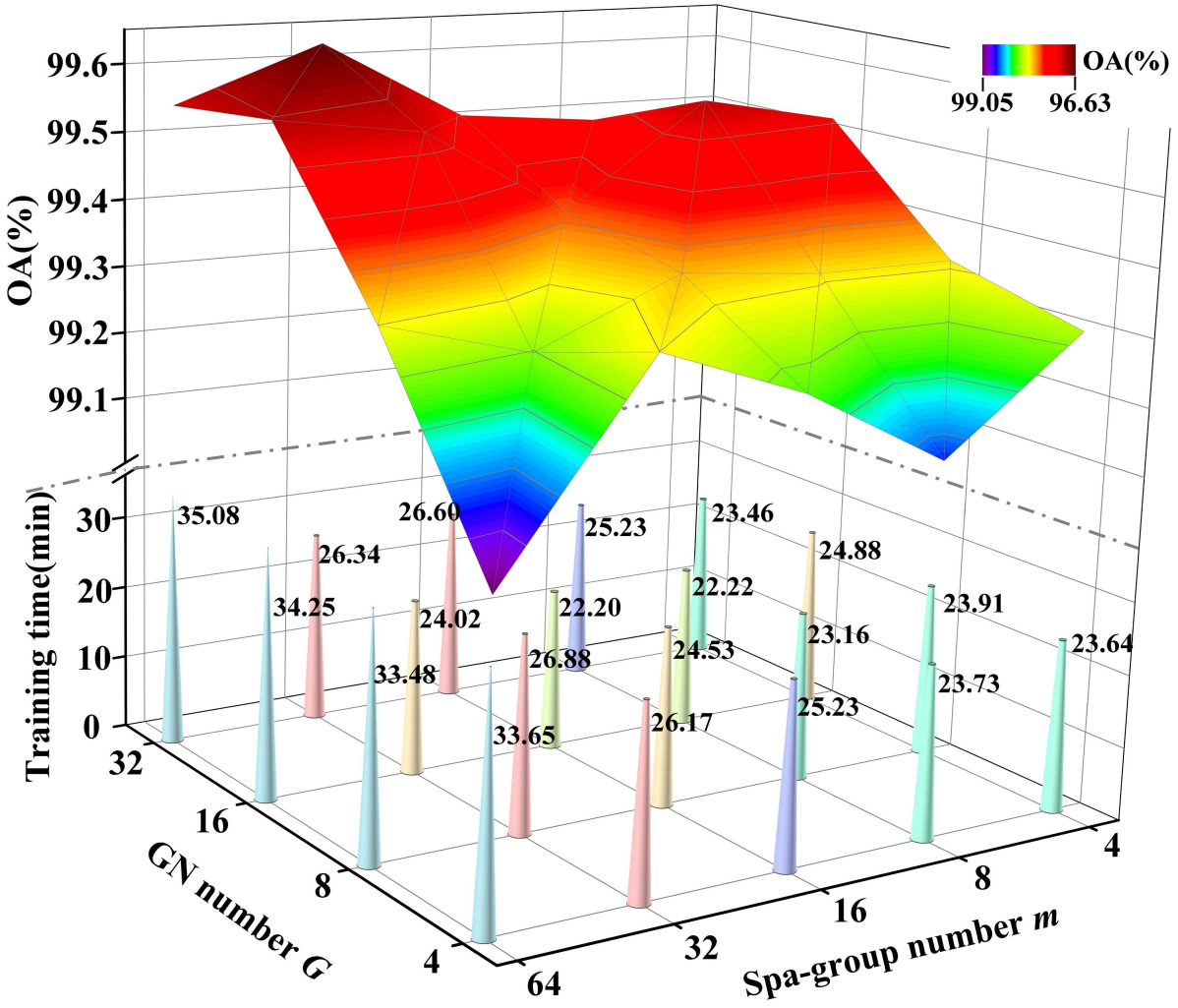

Figure 13. Effect of the number of $SSMB_{spe}$ groups m and GroupNorm (GN) groups G on classification performance (based on UHI No. 005). Classification accuracy and training time are separated by dotted lines, and the figure reports the average of the results from five runs.

generalisation capability under varying environmental conditions, particularly depth-related changes. However, this study does not explicitly evaluate the model under lateral illumination differences induced by terrain (e.g. slopes) within the same survey area, nor does it assess its generalisability across different marine regions. Future work will incorporate physics-based underwater illumination correction methods to address such variability and further validate the model across diverse environments.

In addition, a systematic comparison between raw DN values and reflectance data will be conducted to investigate their impacts on coral reef benthic classification using UHI. Moreover, the lack of publicly available coral reef hyperspectral datasets acquired in turbid, highly dynamic, or complex underwater terrains limits the evaluation of the proposed method in such conditions. To address these challenges, future research will focus on two directions: (1) acquiring large-scale UHI datasets spanning a wide range of turbidity, terrain complexity, and water depths; and (2) improving the classification framework through spectral correction, data augmentation, and integration of depth information to mitigate spectral attenuation, scattering effects, and contrast degradation. It is worth noting that UHIs often exhibit more severe spectral distortion compared with land-based data, due to water absorption, scattering, and variable illumination. In addition, because of the push-broom acquisition mode, many features will be deformed, potentially impacting algorithm performance. In our dataset, ground-truth labels were annotated in regions with relatively minor distortion and deformation, minimising the impact on quantitative evaluation.

From an architectural perspective, the proposed method decouples spatial and spectral features via the SSMB and adaptively fuses them at both feature and decision levels. This design helps correct local classification biases caused by spectral degradation, making the model more suitable for practical underwater scenarios where distortion or deformation may be present. Although we adopt the CLWA strategy, the potential artifacts that might still exist at the stitching edges remain an issue. Future work should further improve the algorithm to better address these artifacts. It should be noted that we did not enforce strict spatial separation between training, validation, and test pixels. Due to spatial continuity within the same ROI, this may introduce spatial autocorrelation and lead to over-optimistic generalisation estimates. In future work, we will adopt spatial cross-validation to more rigorously evaluate the model's effectiveness.

7. Conclusion

This article proposes CRUHI-Mamba, a Mamba-based method for coral reef UHI benthic classification that achieves high accuracy, high efficiency, and low annotation cost. By leveraging long-sequence modelling with linear computational complexity, CRUHI-Mamba captures global spatial dependencies and spectral correlations across the entire UHI, while enabling optimised fusion of spatial and spectral information for coral reef benthic types. The designed SSMB enables the model to capture discriminative global spatial and fine-grained spectral features of benthic types from highly heterogeneous UHI, even under limited training samples. Furthermore, the AFDFM alleviates classification bias and fusion issues caused by imbalanced information density among benthic types, achieving adaptive balanced fusion of spatial–spectral information at feature and decision levels to produce precise classification results. Based on extensive experimental results obtained from UHI data of coral reefs collected at various water depths, we derive the following key conclusions: (1) CRUHI-Mamba outperforms state-of-the-art methods in both classification accuracy and result visualisation, demonstrating high precision and strong generalisability. (2) With less than 2% of labelled samples, it achieves competitive performance with fewer parameters and the shortest inference time among compared methods. (3) The high-precision benthic habitat maps generated by CRUHI-Mamba provide critical data support for studies on species diversity, community composition, and coral reef health assessment. Future research will explore extending CRUHI-Mamba to a broader range of coral reef UHI applications by incorporating multi-modal data and investigating lightweight model designs for efficient large-scale benthic classification and habitat mapping.

Acknowledgements

We gratefully acknowledge the Microsensor Group (Max Planck Institute for Marine Microbiology), the Data Science and Technology Unit (Leibniz Centre for Tropical Marine Research), and PANGAEA: Data Publisher for Earth and

Environmental Science for providing the underwater hyperspectral coral reef data utilised in this study. We also thank the ESA Copernicus data space ecosystem for providing Sentinel-2A data.

Author contributions

CRediT: **Feifei Zhang:** Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft; **Yabin Hu:** Supervision, Writing – review & editing; **Yi Ma:** Conceptualization, Methodology, Supervision, Writing – review & editing; **Guangbo Ren:** Funding acquisition, Supervision; **Wenshuo Zhu:** Methodology, Visualization; **Shou Feng:** Supervision; **Zhongwei Li:** Supervision, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Key Research and Development Plan under Grant (No. 2022YFC3105100).

Data availability statement

The underwater hyperspectral benthic classification dataset for coral reefs was preprocessed prior to use, including hyperspectral band clipping and class reindexing of the ground truth labels. The preprocessed dataset supporting the findings of this study is available at https://doi.org/10.6084/m9.figshare.31130086 (Zhang et al. 2026). The original, unprocessed coral reef hyperspectral dataset is publicly available and can be accessed at https://doi.org/10.1594/PANGAEA.911300 (Chennu et al. 2020). The source code for the proposed method is available at https://github.com/Feifei2025-z/CRUHI-Mamba.

References

Ben Hamida, A., A. Benoit, P. Lambert, and C. Ben Amar. 2018. "3-D Deep Learning Approach for Remote Sensing Image Classification." *IEEE Transactions on Geoscience and Remote Sensing* 56: 4420–4434. https://doi.org/10.1109/TGRS.2018.2818945.

Chen, H., L. Cheng, and K. Zhang. 2025. "Bathymetry-Guided Multi-Source Remote Sensing Image Domain Adaptive Coral Reef Benthic Habitat Classification." *GIScience & Remote Sensing* 62: 2471193. https://doi.org/10.1080/15481603.2025.2471193.

Chen, H., J. Cheng, X. Ruan, J. Li, L. Ye, S. Chu, L. Cheng, and K. Zhang. 2024. "Satellite Remote Sensing and Bathymetry Co-Driven Deep Neural Network for Coral Reef Shallow Water Benthic Habitat Classification." *International Journal of Applied Earth Observation and Geoinformation* 132: 104054. https://doi.org/10.1016/j.jag.2024.104054.

Chennu, A., Rashid, A. R., Haan, J. den, and Beer, D. de 2020. Taxonomically annotated underwater hyperspectral and color images of coral reef transects from Curaçao. https://doi.org/10.1594/PANGAEA.911300

Chennu, A., P. F Rber, G. De Ath, D. De Beer, and K. E. Fabricius. 2017. "A Diver-Operated Hyperspectral Imaging and Topographic Surveying System for Automated Mapping of Benthic Habitats." *Scientific Reports* 7: 7122. https://doi.org/10.1038/s41598-017-07337-y.

Foglini, F., V. Grande, F. Marchese, V. A. Bracchi, M. Prampolini, L. Angeletti, G. Castellan, et al. 2019. "Application of Hyperspectral Imaging to Underwater Habitat Mapping, Southern Adriatic Sea." *Sensors (Basel, Switzerland)* 19: 2261. https://doi.org/10.3390/s19102261.

Gu, A., and Dao, T. 2023. Mamba: Linear-Time Sequence Modeling with Selective State Spaces. arXiv. https://doi.org/10.48550/arxiv.2312.00752

Gu, A., Goel, K., and R'E, C. 2021. Efficiently Modeling Long Sequences with Structured State Spaces. ArXiv abs/2111.00396.

Gu, A., Dao, T., Ermon, S., Rudra, A., and Ré, C. 2020. HiPPO: Recurrent Memory with Optimal Polynomial Projections. CoRR abs/2008.07669.

He, Y., B. Tu, B. Liu, J. Li, and A. Plaza. 2024. "3DSS-Mamba: 3D-Spectral-Spatial Mamba for Hyperspectral Image Classification." *IEEE Transactions on Geoscience and Remote Sensing* 62: 1–16. https://doi.org/10.1109/TGRS.2024.3472091.

Hong, D., Z. Han, J. Yao, L. Gao, B. Zhang, A. Plaza, and J. Chanussot. 2022. "SpectralFormer: Rethinking Hyperspectral Image Classification with Transformers." *IEEE Transactions on Geoscience and Remote Sensing* 60: 1–15. https://doi.org/10.1109/TGRS.2021.3130716.

Hughes, D. J., R. Alderdice, C. Cooney, M. Kühl, M. Pernice, C. R. Voolstra, and D. J. Suggett. 2020. "Coral Reef Survival under Accelerating Ocean Deoxygenation." *Nature Climate Change* 10: 296–307. https://doi.org/10.1038/s41558-020-0737-9.

ICRI, I.C.R.I. 2025. 84% of the world's coral reefs impacted in the most intense global coral bleaching event ever [WWW Document]. (accessed 4.23.25). URL https://icriforum.org/4gbe-2025/

Ji, X., B. Yang, Z. Wei, M. Wang, Q. Tang, W. Xu, Y. Wang, J. Zhang, and L. Zhang. 2024. "Benthic Habitat Sediments Mapping in Coral Reef Area Using Amalgamation of Multi-Source and Multi-Modal Remote Sensing Data." *Remote Sensing of Environment* 304: 114032. https://doi.org/10.1016/j.rse.2024.114032.

Kurinchi-Vendhan, R., Gray, D., and Cole, E. 2023. BenthIQ: a Transformer-Based Benthic Classification Model for Coral Restoration. ArXiv abs/2311.13661.

Li, M., H. Zhang, A. Gruen, and D. Li. 2024. "A Survey on Underwater Coral Image Segmentation Based on Deep Learning." *Geo-spatial Information Science 1–25* 28: 472–496. https://doi.org/10.1080/10095020.2024.2343323.

Li, Y., Y. Luo, L. Zhang, Z. Wang, and B. Du. 2024. "MambaHSI: Spatial–Spectral Mamba for Hyperspectral Image Classification." *IEEE Transactions on Geoscience and Remote Sensing* 62: 1–16. https://doi.org/10.1109/TGRS.2024.3430985.

Liu, Q., J. Yue, Y. Fang, S. Xia, and L. Fang. 2024. "HyperMamba: A Spectral-Spatial Adaptive Mamba for Hyperspectral Image Classification." *IEEE Transactions on Geoscience and Remote Sensing* 62: 1–14. https://doi.org/10.1109/TGRS.2024.3482473.

Mills, M. S., M. Ungermann, G. Rigot, J. den Haan, J. X. Leon, and T. Schils. 2023. "Assessment of the Utility of Underwater Hyperspectral Imaging for Surveying and Monitoring Coral Reef Ecosystems." *Scientific Reports* 13: 21103. https://doi.org/10.1038/s41598-023-48263-6.

Mills, M. S., M. Ungermann, G. Rigot, J. den Haan, J. X. Leon, and T. Schils. 2024. "Coral Reefs in Transition: Temporal Photoquadrat Analyses and Validation of Underwater Hyperspectral Imaging for Resource-Efficient Monitoring in Guam." *PLoS One* 19: e0299523. https://doi.org/10.1371/journal.pone.0299523.

Ming, R., N. Chen, J. Peng, W. Sun, and Z. Ye. 2025. "Semantic Tokenization-Based Mamba for Hyperspectral Image Classification." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 18: 4227–4241. https://doi.org/10.1109/JSTARS.2025.3528122.

Pan, Z., C. Li, A. Plaza, J. Chanussot, and D. Hong. 2025. "Hyperspectral Image Classification with Mamba." *IEEE Transactions on Geoscience and Remote Sensing* 63: 1–14. https://doi.org/10.1109/TGRS.2024.3521411.

Pechlivanidou, G., and N. Karampetakis. 2022. "Zero-Order Hold Discretization of General State Space Systems with Input Delay." *IMA journal of mathematical control and information* 39: 708–730. https://doi.org/10.1093/imamci/dnac005.

Rashid, A. R., and A. Chennu. 2020. "A Trillion Coral Reef Colors: Deeply Annotated Underwater Hyperspectral Images for Automated Classification and Habitat Mapping." *Data* 5: 19. https://doi.org/10.3390/data5010019.

Schürholz, D., and A. Chennu. 2023. "Digitizing the Coral Reef: Machine Learning of Underwater Spectral Images Enables Dense Taxonomic Mapping of Benthic Habitats." *Methods in Ecology and Evolution* 14: 596–613. https://doi.org/10.1111/2041-210X.14029.

Sheng, J., J. Zhou, J. Wang, P. Ye, and J. Fan. 2025. "DualMamba: A Lightweight Spectral–Spatial Mamba-Convolution Network for Hyperspectral Image Classification." *IEEE Transactions on Geoscience and Remote Sensing* 63: 1–15. https://doi.org/10.1109/TGRS.2024.3516817.

Sing Wong, A., S. Vrontos, and M. L. Taylor. 2022. "An Assessment of People Living By Coral Reefs over Space and Time." *Global Change Biology* 28: 7139–7153. https://doi.org/10.1111/gcb.16391.

Sun, L., G. Zhao, Y. Zheng, and Z. Wu. 2022. "Spectral–Spatial Feature Tokenization Transformer for Hyperspectral Image Classification." *IEEE Transactions on Geoscience and Remote Sensing* 60: 1–14. https://doi.org/10.1109/TGRS.2022.3144158.

Teague, J., J. C. C. Day, M. J. Allen, T. B. Scott, E. J. Hochberg, and D. Megson-Smith. 2023. "A Demonstration of the Capability of Low-Cost Hyperspectral Imaging for the Characterisation of Coral Reefs." *Oceans* 4: 286–300. https://doi.org/10.3390/oceans4030020.

Wang, D., B. Du, and L. Zhang. 2022. "Fully Contextual Network for Hyperspectral Scene Parsing." *IEEE Transactions on Geoscience and Remote Sensing* 60: 1–16. https://doi.org/10.1109/TGRS.2021.3050491.

Wang, Z., N. Xu, Z. You, and S. Zhang. 2025. "DiffMamba: Semantic Diffusion Guided Feature Modeling Network for Semantic Segmentation of Remote Sensing Images." *GIScience & Remote Sensing* 62: 2484829. https://doi.org/10.1080/15481603.2025.2484829.

Wang, G., X. Zhang, Z. Peng, T. Zhang, and L. Jiao. 2025. "S2 Mamba: A Spatial-Spectral State Space Model for Hyperspectral Image Classification." *IEEE Transactions on Geoscience and Remote Sensing* 63: 1–1. https://doi.org/10.1109/TGRS.2025.3530993.

Wang, C., J. Huang, M. Lv, H. Du, Y. Wu, and R. Qin. 2024. "A Local Enhanced Mamba Network for Hyperspectral Image Classification." *International Journal of Applied Earth Observation and Geoinformation* 133: 104092. https://doi.org/10.1016/j.jag.2024.104092.

Zeng, K., Z. Xu, Y. Yang, Y. Liu, H. Zhao, Y. Zhang, B. Xie, W. Zhou, C. Li, and W. Cao. 2022. "In Situ Hyperspectral Characteristics and the Discriminative Ability of Remote Sensing to Coral Species in the South China Sea." *GIScience & Remote Sensing* 59: 272–294. https://doi.org/10.1080/15481603.2022.2026641.

Zhang, H., A. Gruen, and M. Li. 2022. "Deep Learning for Semantic Segmentation of Coral Images in Underwater Photogrammetry." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences V-2–2022* V-2-2022: 343–350. https://doi.org/10.5194/isprs-annals-V-2-2022-343-2022.

Zhang, F., Y. Hu, Y. Ma, G. Ren, Z. Li, and Y. Bao. 2024. "GCU-Net: Remote Sensing Classification Method for Coral Reef Geomorphology Integrating Geospatial Cognition." *IEEE Geoscience and Remote Sensing Letters* 21: 1–5. https://doi.org/10.1109/LGRS.2024.3406681.

Zhang, F., Hu, Y., Ma, Y., Ren, G., Zhu, W., Feng, S., and Li, Z. 2026. Coral Reef Underwater Hyperspectral Dataset. https://doi.org/10.6084/m9.figshare.31130086.v2

Zhao, Z., X. Xu, S. Li, and A. Plaza. 2024. "Hyperspectral Image Classification Using Groupwise Separable Convolutional Vision Transformer Network." *IEEE Transactions on Geoscience and Remote Sensing* 62: 1–17. https://doi.org/10.1109/TGRS.2024.3377610.

Zhao, M., C. Zhang, P. Yue, C. Cai, and F. Ye. 2025. "MDA-RSM: Multi-Directional Adaptive Remote Sensing Mamba for Building Extraction." *GIScience & Remote Sensing* 62: 2568776. https://doi.org/10.1080/15481603.2025.2568776.

Zhong, Z., J. Li, Z. Luo, and M. Chapman. 2018. "Spectral-Spatial Residual Network for Hyperspectral Image Classification: A 3-D Deep Learning Framework." *IEEE transactions on geoscience and remote sensing* 56: 847–858. https://doi.org/10.1109/TGRS.2017.2755542.

Zhu, D., B. Du, and L. Zhang. 2021. "Two-Stream Convolutional Networks for Hyperspectral Target Detection." *IEEE Transactions on Geoscience and Remote Sensing* 59: 6907–6921. https://doi.org/10.1109/TGRS.2020.3031902.

Zhu, D., B. Du, and L. Zhang. 2024. "Learning Single Spectral Abundance for Hyperspectral Subpixel Target Detection." *IEEE Transactions on Neural Networks and Learning Systems* 35: 10134–10144. https://doi.org/10.1109/TNNLS.2023.3239061.

Zhu, D., B. Du, Y. Dong, and L. Zhang. 2023. "Target Detection with Spatial-Spectral Adaptive Sample Generation and Deep Metric Learning for Hyperspectral Imagery." *IEEE Transactions on Multimedia* 25: 6538–6550. https://doi.org/10.1109/TMM.2022.3210389.

Zhu, Q., Y. Fang, Y. Cai, C. Chen, and L. Fan. 2024. "Rethinking Scanning Strategies with Vision Mamba in Semantic Segmentation of Remote Sensing Imagery: An Experimental Study." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 17: 18223–18234. https://doi.org/10.1109/JSTARS.2024.3472296.

Zuo, X., J. Teng, F. Su, Z. Duan, and K. Yu. 2025. "Multimodel Combination Bathymetry Inversion Approach Based on Geomorphic Segmentation in Coral Reef Habitats Using ICESat-2 and Multispectral Satellite Images." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 18: 3267–3280. https://doi.org/10.1109/JSTARS.2024.3523296.