Deep-Learning-Based Marine Aquaculture Zone Extractions From Dual-Polarimetric SAR Imagery

Wantai Chen, *Student Member, IEEE*, and Xiaofeng Li, *Fellow, IEEE*

***Abstract*—Efficient monitoring of marine aquaculture zones (MAZs) is crucial for facilitating coastal resource management. To achieve this, we developed a specialized deep convolutional neural network tailored for extracting MAZs from synthetic aperture radar (SAR) imagery, integrating prior analytical knowledge of MAZ imaging features. A total of 47 Sentinel-1 dual-polarized (VV and VH) SAR images spanning 2016–2023 in China's Subei Sandbanks along the Yellow Sea coast were collected due to appropriate tidal level and acquired time. We first comprehensively analyzed of normalized radar cross section (NRCS) values for MAZs under varying tidal levels and aquaculture facility structures. Rising tide-induced submergence resulted in a significant mean NRCS reduction of 7.01 dB (VV) and 4.54 dB (VH), causing MAZ signals to resemble seawater. In addition, during low tide, volume scattering from the net screen on the aquaculture rafts increased VH-polarized image recognizability, with a smaller NRCS overlap (64%) between MAZs and tidal flats compared to VV-polarized images. Hence, VH-polarized images taken during low tide with intact aquaculture facilities were selected for dataset construction due to their reliability in characterizing MAZs. Building upon the classical U-Net framework, we introduced four modifications informed by our imaging characteristics analysis to enhance the model's performance. Testing experiments demonstrated an impressive F1-score of 94.77%, highlighting the effectiveness of incorporating prior knowledge into refining deep learning models. Applying the model to SAR images from 2016 to 2023 revealed concentrated MAZs in the relatively flat southeastern Subei Sandbanks, with a noticeable scale decline post-2021 resulting in a 67.65% reduction over the years.**

***Index Terms*—Deep convolution neural networks (DCNNs), imaging characteristics, marine aquaculture zones (MAZs), Sentinel-1, synthetic aperture radar (SAR).**

Manuscript received 13 January 2024; revised 16 February 2024; accepted 29 March 2024. Date of publication 3 April 2024; date of current version 16 April 2024. This work was supported in part by the National Key Research and Development Program of China under Grant 2021YFB3901304, in part by the National Natural Science Foundation of China under Grant 42221005 and Grant 42076200, in part by the National Natural Science Foundation of China under Grant U2006211, in part by the Strategic Priority Research Program of the Chinese Academy of Sciences under Grant XDB42000000, in part by the Natural Science Foundation of Shandong Province under Grant ZR2020MD083, and in part by the Key Research and Development Program of Shandong province under Grant 2022CXPT020. *(Corresponding author: Xiaofeng Li.)*

Wantai Chen is with the Key Laboratory of Ocean Circulation and Waves, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China, and also with the College of Marine Sciences, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: chenwantai@qdio.ac.cn).

Xiaofeng Li is with the Key Laboratory of Ocean Circulation and Waves, Institute of Oceanology, and the Center for Ocean Mega- Science, Chinese Academy of Sciences, Qingdao 266071, China (e-mail: xiaofeng.li@ieee.org).

Digital Object Identifier 10.1109/JSTARS.2024.3384511

I. Introduction

MARINE aquaculture is crucial and steadily expanding worldwide [1]. In 2020, marine aquaculture production reached 33.6 million tons, serving as a substantial source of sustenance and nutrition for the global population [2], [3]. Nonetheless, aquaculture processes generate residual nutrients and significant secretion volumes, making marine aquaculture zones (MAZs) hotbeds of environmental problems [4]. As an illustration, the recurrent emergence of the green tide phenomenon in the Yellow Sea, which previously inflicted substantial economic consequences during the 2008 Olympics, was attributed to the rapid expansion of MAZs in Subei Sandbanks to the south of the affected area [5]. Therefore, people have been effectively managing marine aquaculture resources and reasonably planning to reduce environmental impact and facilitate sustainable development [6].

Effective monitoring of MAZs is the key to managing marine aquaculture resources. Traditional monitoring methods relied on site-specific on-the-spot investigations, which were laborious and expensive due to weather and sea-state limitations [7]. Conversely, remote sensing technology is extensively employed for this task, benefitting from its rapid response, large spatial coverage, and cost-efficiency. Optical and synthetic aperture radar (SAR) data are basic data sources for monitoring tidal flat targets. Optical sensor images are rich in spectral features operating during daylight hours but are frequently contaminated by cloudy or foggy weather, especially in the intertidal zones. On the contrary, SAR is an active radar highly suitable for monitoring MAZs. Its all-weather and all-day observation capability is supplementary with the optical sensor in the detection tasks [8]. Scientists have employed SAR imagery to observe diverse categories of MAZs, including oyster farming zones [9], [10], mussel beds [11], [12], bivalve beds [13], and full-floating raft aquaculture zones [14]. Variations in scattering mechanisms arise among distinct MAZs due to disparities in aquaculture structures and cultivated crops. For example, Geng et al. [15] found the scattering mechanism of the semifloating raft aquaculture zones is dominated by double-bounce and volume scattering at low tides. However, few studies focused on analyzing polarimetric characteristics that vary with the raft structures and ever-changing sea conditions, which may result in an ineffective observation of the complete MAZ signals.

Traditional MAZ extraction methods have been proposed mainly based on the following three approaches: threshold method [16], [17], objected-based method [14], [18], [19], and classic machine learning method [20], [21]. Although the first

two approaches are straightforward and easy to use, challenges emerge when employing them for large-scale remote sensing monitoring tasks. In contrast, classical machine learning methods exhibit greater adaptability. For instance, Cheng et al. [20] efficiently combined SAR data from the Sentinel-1 mission and optical data from the Sentinel-2 mission to extract MAZs in two different sea areas with a modified random forest method. Nonetheless, these methods necessitated intricate calibration of empirical parameters, such as input features, potentially resulting in increased labor requirements or the introduction of human-induced errors.

Deep learning has been comprehensively applied to spaceborne SAR image processing with its high efficiency in recognition and detection tasks [22], [23], [24], [25]. In recent years, several deep convolution neural network (DCNN) methods have been proposed for the MAZs extraction tasks [26], [27], [28], [29], [30]. For instance, Wang et al. [31] incorporated dilated and offset convolution to mitigate the negative effect of speckle noise on MAZs extraction and Wang et al. [32] also presented a MAZs extraction method named incremental double unsupervised deep learning, reducing the labeled data requirement during model training. However, the design of DCNN models often overlooks the impact of imaging conditions on target features. This oversight implies that the described DCNN methods may lack robustness when monitoring MAZs, whose imaging features are consistently influenced by both the raft structures and the dynamic sea state.

To bolster the robustness of DCNN in extracting MAZs from SAR images, we devise a scheme to incorporate excavated SAR knowledge into the DCNN. The main contributions of this article are as follows.

1) Utilizing Sentinel-1 dual-polarized images, we statistically analyzed the imaging characteristics of MAZs in the Subei Sandbanks along the Yellow Sea coast of China under varying tidal conditions and raft structures. This analysis yielded suitable imaging conditions and polarization channels for MAZ extraction.
2) A specially tailored deep learning-based network (MAZ-Net) was proposed for automatic MAZ extraction from SAR imagery, embodying input optimization and model architectural enhancement according to the polarized and raft-structural characteristics.
3) MAZ-Net was employed to extract MAZs from SAR images spanning 2016 to 2023, which facilitated the reveal of changes in the distribution and scale of MAZs over recent years.

The rest of this article is organized as follows. Section II provides details about the study area and data. In Section III, we delve into analyzing the SAR imaging characteristics of the MAZs. The structure of the MAZ-Net and its performance evaluation are described in Section IV. Section V demonstrates the practical applications of the trained model. Finally, Section VI concludes this article.

II. Study Area and Objects

Subei Sandbanks is located along the southeastern Yellow Sea coast (see Fig. 1). It constitutes the largest intertidal zone

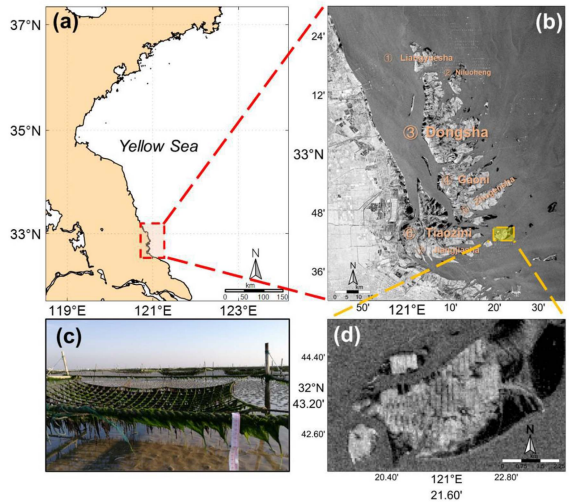

Fig. 1. (a) Overview of the study area. (b) VV-polarized Sentinel-1B SAR image of the study area at low tidal level, acquired at greenwich mean time (GMT) 09:54 on March 19, 2019. Seven main tidal flats: Liangyuesha, Niluoheng, Dongsha, Gaoni, Zhugensha, Tiaozini, and Jiangjiasha. (c) Field survey photo of the MAZs in the study area, acquired on March 15, 2017. (d) Detailed image of the MAZs in (b).

in China, encompassing seven diverging tidal flats characterized by varying sizes and shapes [33]. Notably, this region features distinctive manufactured facilities, such as MAZs and wind turbines in the tidal flats. The study area experiences complex hydrodynamic conditions under the influence of semidiurnal tides dominated by a progressive Poincaré wave from the East China Sea and an amphidromic system in the Yellow Sea [34]. These sandbanks experience exposure and submergence fluctuations as the tide rises and falls [35]. These favorable tidal conditions render the Subei Sandbanks a primary production area for Porphyra Yezensis, with the semifloating rafts as the only aquaculture facilities [36].

As shown in Fig. 1(c), the semifloating raft we focused on has a unique structure consisting of bamboo poles, ropes, and nylon net screens for supporting, fixation, and attaching the Porphyra Yezensis [37]. These rafts float on the sea surface and rise with seawater during tide flooding but stand on the tidal flats upheld by vertical bamboo poles at low tidal levels [38], [39], making the net screens exposed to the air for a certain duration every day. In addition, partial rafts are submerged when the tide is high due to the limited length of the ropes. In each aquaculture cycle, the semifloating raft structures have two states: they remain intact throughout the growing period (November to April), while during the preparation (September to November) and recovery (April to May) periods, only the nets are removed because of the facility's preparatory and recycling procedures [20], [38]. MAZs are shown as organized blocks composed of light and dark stripes in the SAR imagery at low tidal levels [see Fig. 1(d)], with the bright stripes corresponding to arranged rafts and the narrow dark bands to the intermediate tidal flats. It is worth noting that changes in tidal level or in the structure of semifloating rafts due to the aquaculture cycles can affect the imaging characteristics of MAZs in SAR images, which will be discussed in Section III.

TABLE I
DETAILS OF 47 SAR IMAGES USED IN THIS STUDY

	No.	Satellite	Acquisition Day (yyyy/MM/dd)	Acquisition Time (hh: mm)	Tidal Level (m)	Structure of aquaculture facilities	wind speed (m/s)
Training Set	1*	Sentinel-1A	January 17, 2018	09:55	−2.40	Rafts with nets	2.52
	2*	Sentinel-1A	January 29, 2018	09:55	−1.72	Rafts with nets	5.37
	3*	Sentinel-1A	March 30, 2018	09:55	−2.54	Rafts with nets	2.99
	4	Sentinel-1A	August 9, 2018	09:55	−1.56	None	2.31
	5*	Sentinel-1A	December 7, 2018	09:55	−2.50	Rafts with nets	8.81
	6*	**Sentinel-1A**	**February 5, 2019**	**09:55**	**−2.39**	**Rafts with nets**	**5.31**
	7*	**Sentinel-1A**	**February 17, 2019**	**09:55**	**−1.73**	**Rafts with nets**	**2.87**
	8*	**Sentinel-1A**	**March 7, 2019**	**09:54**	**−2.37**	**Rafts with nets**	**5.16**
	9*	Sentinel-1A	April 6, 2019	09:55	−2.30	Rafts with nets	6.46
	10	Sentinel-1A	June 5, 2019	09:55	−1.58	None	8.29
	11	Sentinel-1A	June 17, 2019	09:55	−1.92	None	6.24
	12	Sentinel-1B	July 17, 2019	09:54	−1.65	None	6.41
	13*	Sentinel-1B	January 13, 2020	09:54	−1.89	Rafts with nets	1.31
	14*	Sentinel-1B	January 25, 2020	09:54	−2.53	Rafts with nets	7.25
	15*	Sentinel-1B	February 6, 2020	09:54	−1.63	Rafts with nets	5.88
	16*	Sentinel-1A	March 7, 2020	09:55	−2.05	Rafts with nets	3.91
	17*	Sentinel-1B	March 25, 2020	09:54	−2.19	Rafts with nets	7.98
	18	Sentinel-1A	June 23, 2020	09:55	−1.35	None	7.43
	19	Sentinel-1A	July 5, 2020	09:55	−2.14	None	5.26
	20*	Sentinel-1A	January 1	09:55	−1.92	Rafts with nets	4.77
	21*	Sentinel-1A	January 13, 2021	09:55	−2.71	Rafts with nets	2.35
	22*	Sentinel-1B	January 31, 2021	09:54	−1.83	Rafts with nets	3.65
	23*	Sentinel-1B	February 12, 2021	09:54	−2.61	Rafts with nets	3.71
	24*	Sentinel-1B	February 24, 2021	09:54	−1.64	Rafts with nets	7.51
	25	Sentinel-1A	May 25, 2021	09:55	−2.64	None	4.14
	26	Sentinel-1B	June 12, 2021	09:54	−1.39	None	5.09
	27	Sentinel-1A	July 24, 2021	09:55	−1.99	None	8.41
Testing Set	28*	Sentinel-1A	March 18, 2018	09:55	−2.54	Rafts with nets	2.96
	29*	**Sentinel-1B**	**March 19, 2019**	**09:54**	**−2.32**	**Rafts with nets**	**8.35**
	30*	Sentinel-1A	February 24, 2020	09:55	−2.39	Rafts with nets	6.08
	31*	Sentinel-1A	March 14, 2021	09:55	−2.39	Rafts with nets	6.41
Application	32	Sentinel-1A	December 23, 2015	09:55	−2.04	Rafts with nets	4.59
	33	Sentinel-1A	February 9, 2016	09:55	−2.41	Rafts with nets	7.05
	34	**Sentinel-1A**	**February 27, 2017**	**09:55**	**−2.74**	**Rafts with nets**	**4.43**
	35	Sentinel-1A	March 11, 2017	09:55	−2.86	Rafts with nets	6.63
	36	Sentinel-1B	December 21, 2021	09:55	−2.07	Rafts with nets	2.88
	37	Sentinel-1A	February 1, 2022	09:55	−2.82	Rafts with nets	5.63
	38	Sentinel-1A	February 13, 2022	09:55	−1.51	Rafts with nets	3.12
	39	Sentinel-1A	February 20, 2023	09:55	−3.25	Rafts with nets	8.02
Others	**40**	**Sentinel-1A**	**May 10, 2017**	**09:55**	**−2.24**	**Rafts without nets**	**6.82**
	41	**Sentinel-1A**	**July 28, 2018**	**09:55**	**−1.66**	**None**	**2.92**
	42	**Sentinel-1A**	**January 12, 2019**	**09:55**	**1.42**	**Rafts with nets**	**4.6**
	43	**Sentinel-1A**	**January 24, 2019**	**09:55**	**−0.54**	**Rafts with nets**	**3.48**
	44	**Sentinel-1B**	**February 23, 2019**	**09:54**	**0.24**	**Rafts with nets**	**7.78**
	45	**Sentinel-1A**	**March 1, 2019**	**09:55**	**0.31**	**Rafts with nets**	**2.95**
	46	**Sentinel-1A**	**March 13, 2019**	**09:55**	**1.50**	**Rafts with nets**	**3.77**
	47	**Sentinel-1A**	**March 25, 2019**	**09:55**	**−0.03**	**Rafts with nets**	**4.42**

The images with intact MAZs at low tidal levels for model training and testing are marked by asterisks (*); the images without MAZs at low tidal levels for model training are marked with shadow; the images for SAR imaging analysis are bolded.

III. SAR IMAGERY COLLECTION AND ANALYSIS

Knowledge of target imaging characteristics facilitates data screening and modeling, which in turn improves monitoring accuracy. In this section, we first display the data utilized in this article and then statistically address the imaging discrepancies under different polarization, focusing on two pivotal influencing factors: tidal level and structure of semifloating rafts.

A. Data Collection

The European Space Agency's Sentinel-1 provides C-band SAR images with ground range detected high-resolution mode, proven to be effective in tidal flats monitoring tasks [40]. The imaging mode is interferometric wide (IW) with two polarizations, including VH polarization (transmitted and received in vertical polarization and received in horizontal polarization) and VV polarization (both transmitted and received in vertical polarization). We collected 47 dual- (VV and VH-) polarized SAR images (see Table I) with 10 m resolution from 2015 to 2023 under diverse sea conditions. The sentinel application platform software was used to preprocess the SAR images, including cropping and splicing, radiometric correction (in dB), speckle filtering, and geo-correction. Tidal levels were calibrated from the TPXO ocean tide model [41], whose outputs showed a good performance in tidal phase prediction but presented a systematic underestimation of amplitude [42]. Wind speed data were the hourly data at the 10 m level from the Fifth

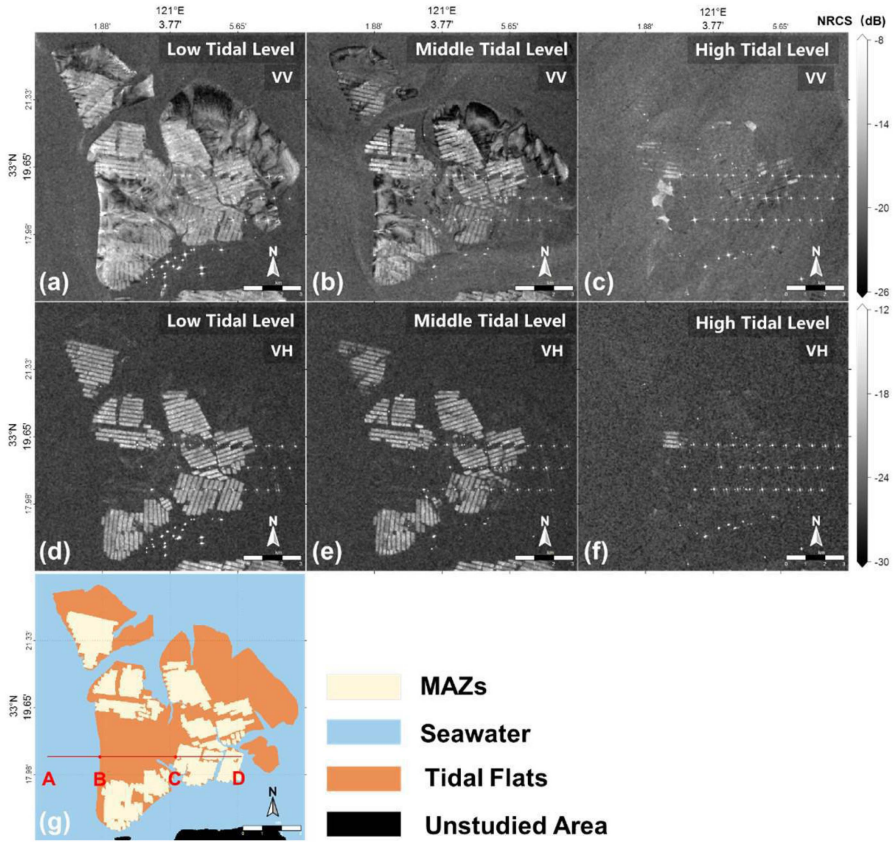

Fig. 2. VV-polarized SAR images acquired on (a) February 5, 2019, (b) January 24, 2019, and (c) March 1, 2019. Their corresponding VH-polarized images (d)–(f). (g) Classification map of Liangyuesha and its surroundings during the growing period in 2019 [not showing the wind turbines and ships appeared as bright points in (a)–(f)], with the red line A-D denoting the cross-section shown in Fig. 4.

Generation of ECMWF Atmospheric Reanalysis of the Global Climate (ERA5) (https://cds.climate.copernicus.eu/). Environmental factors, such as tidal level and wind speed can significantly influence the imaging of tidal flat targets in SAR images. Hence, tide level and wind speed are gathered to aid in selecting SAR images.

B. *Effect of Tidal Level on SAR Imaging Characteristics*

In order to compare MAZs' imaging characteristics at different tidal levels while maintaining the structural integrity of the rafts, we collected all images from January to March 2019 (10 images in total, shown in Table I as bolded ones). Liangyuesha, located in the northernmost of the Subei Sandbanks, was selected as the study area due to the regularly distributed instances of three target classes [shown in Fig. 2(g)]: MAZs, tidal flats, and seawater. In addition, to facilitate the description of tidal conditions, we categorized tidal levels into the following three classes: low tidal levels (−1.31 m or less), middle ones (−1.31 m to −0.15 m), and high ones (−0.15 m or more). In both polarized (VV and VH) channels of the ten images, MAZs or tidal flats exhibit similar completeness and clarity under the same tidal class, while they may differ for different tidal classes, as depicted in Fig. 2(a)–(f). To clearly see the imaging difference, the variation of the overall normalized radar cross section (NRCS) for each of the target classes based on the 10 images was counted in Fig. 3. Then, we systematically organized the imaging characteristics of each of the target classes under various tidal classes in Table II.

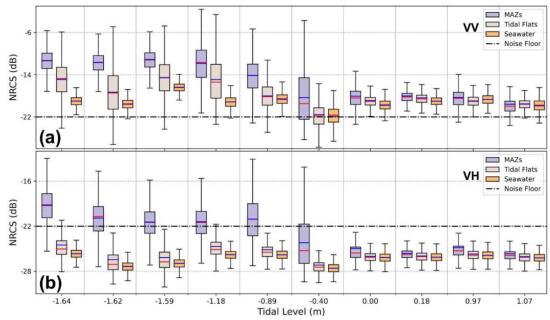

Fig. 3. Box plots of three categories of targets (MAZs, tidal flats, and seawater) in VV (a) and VH (b) channels with the blue solid lines representing the mean NRCS values, the red ones representing the median values, and the dash-dotted lines representing the noise floor (outliers are omitted in these plots). We replace the image date with the corresponding tide level and arrange them in the horizontal coordinate in order of magnitude to highlight the effect of tide level on the NRCS.

Table II illustrates the sensitivity of MAZs' imaging features to tidal variations in both polarizations. MAZs appear as organized blocks characterized by alternating wide bright and narrow

TABLE II
IMAGING CHARACTERISTICS OF EACH OF THE TARGET CLASSES UNDER VARIOUS TIDAL CLASSES

		Low Tide	Middle Tide	High Tide
VV Channel	MAZs	Indistinct organized blocks; relatively large median and small range of NRCS values	Partial disappeared and block edge blurred; small median and large range of the NRCS values	Almost completely disappeared; similar distribution of the NRCS values to seawater
	Tidal Flats	Irregular brightness; relatively large range of median and NRCS values	Partial disappeared	Almost completely disappeared; similar distribution of the NRCS values to seawater
	Seawater	Very low brightness; low and stable NRCS values with narrow range		
VH Channel	MAZs	Pronounced organized blocks; unique high NRCS values, large range of NRCS values	Partial disappearance;block edge blurred; the NRCS median near noise floor	Almost completely disappeared; the NRCS values below the noise floor
	Tidal Flats	Low brightness and blurry; the NRCS almost below the noise floor		
	Seawater	Below the noise floor		

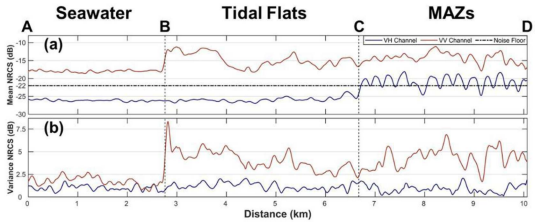

Fig. 4. (a) Mean and (b) variance of NRCS along the cross-section A-D of SAR images acquired at low tidal levels with the dash-dotted line representing the noise floor.

dark stripes at the low tidal level [see Fig. 2(a) and (d)], while seawater generally displays a more subdued appearance with low NRCS. As the tide rises, the MAZs become progressively submerged, reducing visibility, recognition [see Fig. 2(a)–(f)] and the value of NRCS (as shown in Fig. 3). At high tidal level, the mean NRCS values of MAZs in the VV and VH channels are −18.50 dB and −25.29 dB, respectively, which are 7.01 dB and 4.54 dB lower than that at low tide, indicating the challenge of clearly recognizing the signals of MAZs. On the other hand, at low tidal levels, tidal flats and MAZs show indistinguishable larger values in the VV channel [see Fig. 3(a)]. Interestingly, in their corresponding VH-polarized images [see Fig. 3(b)], tidal flats exhibit weaker backscattering intensity, even lower than the sensors' noise floor of the IW imaging mode [43]. In contrast, MAZs display consistently elevated NRCS during low tidal levels in the VH channel [see Fig. 2(d)]. At low tidal levels, the overlap between the NRCS value ranges of MAZs and tidal flats is 99% in the VV channel and 64% in the VH channel, which suggests that MAZs are more distinguishable from the surrounding tidal flats and seawater in the VH channel.

We further calculated the mean and variance of NRCS along the cross-section A-D (see Fig. 2 for location) of four images acquired at low tidal levels from these 10 SAR images. As depicted by the solid lines in Fig. 4(a), tidal flats show similar stronger backscattering intensity as MAZs in VV-polarized images but appear much weaker and smoother than seawater in VH-polarized images. Conversely, MAZs show similar stronger signals in both channels but are more distinguishable in the VH-polarized images due to the weak backscattering from the tidal flats and seawater. In addition, VH-polarized MAZ signature manifests as regular funnel-shaped fluctuations due to the narrow gaps between the rafts, which also contributes to their recognizability in VH-polarized images.

On the other side, variance in NRCS reveals the susceptibility of the polarization characteristics of the three classes to sea conditions [44]. As illustrated in Fig. 4(b), MAZs and tidal flats display obvious variations in the VV channel, whereas the values for all targets remain low in the VH channel. It implies that VV-polarized images demonstrate significantly higher sensitivity to environmental changes than cross-polarized ones, particularly within the land parts of these images. Therefore, despite the substantial variations in imaging among the three target classes in the study area due to the ever-shifting sea conditions, the VH-polarized images obtained during low tidal levels can reliably portray the characteristics of the MAZs, which alleviates the challenges in monitoring MAZs arising from the intricate topography and tidal conditions in the Subei Sandbanks.

C. *Effect of Raft Structure on SAR Imaging Characteristics*

Due to the aquaculture process, the structure of the semifloating rafts can be artificially altered, leading to variations in the SAR imaging of MAZs. As main parts of rafts, the net screens and bamboo poles are successively removed from April to May each year. For instance, Zhang et al. [45] pointed out that all the net screens within the Zhugensha MAZs were taken down by mid-April 2017, but the recovery of bamboo poles and ropes did not commence until May 11 of the same year. Therefore, SAR images in two polarized channels acquired on February 27 (intact rafts structure), May 10 (rafts without nets and mainly bamboo poles), 2017, and July 28, 2018 (bamboo poles were removed and devoid of rafts), under similar tidal levels and wind conditions, were selected to explore the effect of raft structures on the imaging characteristics of MAZs. According to the corresponding Fig. 5(a)–(f), the imaging characteristics of MAZs with various raft structures in both channels are summarized in Table III for clarity.

According to Table III, the MAZs in the VH-polarized and VV-polarized images exhibit strong sensitivity to the removal of net screens and bamboo poles, respectively. As for the VV channel, the disappearance of brightness and shape of the MAZs [comparing Fig. 5(b) and (c)] caused by bamboo poles removal

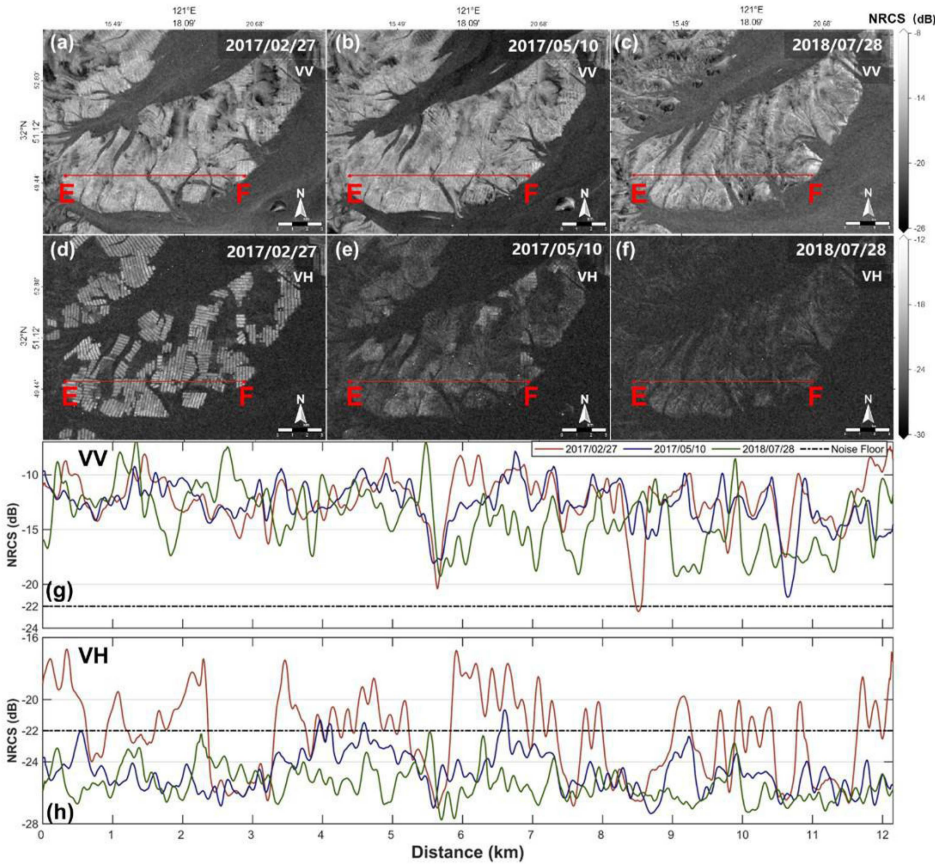

Fig. 5. VV-polarized SAR images acquired on (a) February 27, 2017, (b) May 10, 2017, and (c) July 28, 2018; (d)–(f) Their corresponding VH-polarized images; NRCS along the cross-section EF in (g) VV and (h) VH channel, respectively, with the dash-dotted line representing the noise floor.

TABLE III
IMAGING CHARACTERISTICS OF MAZS AT VARIOUS SEMIFLOATING RAFT STRUCTURES

	Intact	Only without nets	Devoid
VV Channel	Indistinctorganized blocks; high brightness	Similar to when the rafts are intact	Texture disappeared; shapeless
VH Channel	Pronounced regular blocks; high brightness	Faint brightness; Almost shapeless	Disappeared

implies that the bamboo poles act as a source of double-bounce scattering, which dominates the MAZ features in the VV channel. As for the VH channel, the MAZs in Fig. 5(e) appear significantly darker than in Fig. 5(d) but exhibit homologous brightness to tidal flats in Fig. 5(f), where no MAZs are present. It is more evident in Fig. 5(h): the curves corresponding to rafts without nets and regions lacking rafts are nearly indistinguishable below the noise floor, whereas the curve representing the intact structure reveals valuable fluctuation features precisely at the MAZs' location. It illustrates that the net screens on rafts play a crucial role in generating the depolarization effect observed in MAZs, thus influencing the imaging characteristics in the VH channel.

In conclusion, in contrast to the copolarization, features of the MAZs (primarily attributed to the raft net screens) in the VH channel are more distinct and less influenced by the sea conditions than in the VV channel, implying that VH imagery acquired at low tidal levels is well suited for efficiently monitoring the MAZs. Besides, MAZs exhibit organized blocks characterized by alternating wide bright and narrow dark stripes at low tidal levels, inspiring us to build extraction models based on textural features.

IV. DEEP LEARNING FRAMEWORK DEVELOPMENT

A. Data Preparation

Through the imaging characteristics analysis, we posit that SAR images acquired at low tidal levels with intact raft structures offer an accurate representation of MAZ signals. Therefore, we collected 22 Sentinel-1 SAR images during the growing periods of 2018 to 2021 (marked with "*" in Table I). One image was randomly selected from each growing period to constitute the testing set, and the remaining formed the training set (shown in the training set/testing set section in Table I). In addition, 9

images (marked with shadow in Table I) without MAZs were supplemented into the training set to enhance the representation of background (such as tidal flats and seawater) features within the deep learning model. Finally, the training and testing set consisted of 27 and 4 SAR images, respectively. Preprocessed SAR images were downsampled to a 50 m spatial resolution to expedite the training process.

Accurate ground truth is vital for model training. We manually extracted MAZs in each SAR image based on visual interpretation. The procedure involves overlaying a transparent layer onto the VH-polarized SAR image and then accurately marking the MAZs using a stylus. The final ground truth maps were binary images with 255 and 0 labeling the MAZs and other areas, respectively. All the SAR images and their corresponding ground truth were cropped into 256×256 pixels subimages for model training. Therefore, the final training set contains 1944 sample pairs (90% for training and 10% for validation) and other 288 pairs for testing. During model training, we use online data augmentation [35] to enrich the datasets.

B. Model Structure

The U-Net framework [46] is renowned for delivering effective segmentation results in medical images even with limited datasets and stands as one of the efficient and frequently employed base models in remote sensing classification tasks [24], [35], [47], [48], [49], [50], [51], [52]. However, U-Net may encounter challenges when segmenting satellite data, particularly when dealing with SAR images characterized by intricate imaging mechanisms and features [53], [54]. The MAZ-Net (see Fig. 6) was proposed guided by the analysis of backscattering and texture features within SAR images, leveraging four modifications to improve the performance compared to the original U-Net:

1) GLCM-Based Dual-Branch Encoder: The imaging characteristics of MAZs vary dramatically due to changing environmental factors, posing a challenge to the task of automatic extraction. We first introduced three gray level co-occurrence matrix (GLCM) texture maps as additional channels to enrich the input features. Texture features derived from the GLCM have demonstrated practical utility in SAR image classification tasks [47], [55], [56]. The GLCM provides the following nine texture features: contrast, dissimilarity, homogeneity, angular second moment, energy, maximum probability, entropy, mean, variance, and correlation. However, the high correlation among some of these features can lead to redundancy and increased computational demands in modeling. To address this, it is beneficial to select features with low correlation that distinctly represent the target. Initially, we extracted these nine GLCM texture features from a randomly selected SAR image and calculated their correlation matrix to evaluate interfeature relationships. Following a detailed analysis focused on their noise reduction capabilities and differentiation between MAZs and other elements, angular second moment, contrast, and mean were identified as optimal for modeling due to their low intercorrelation and effective representation of MAZ characteristics.

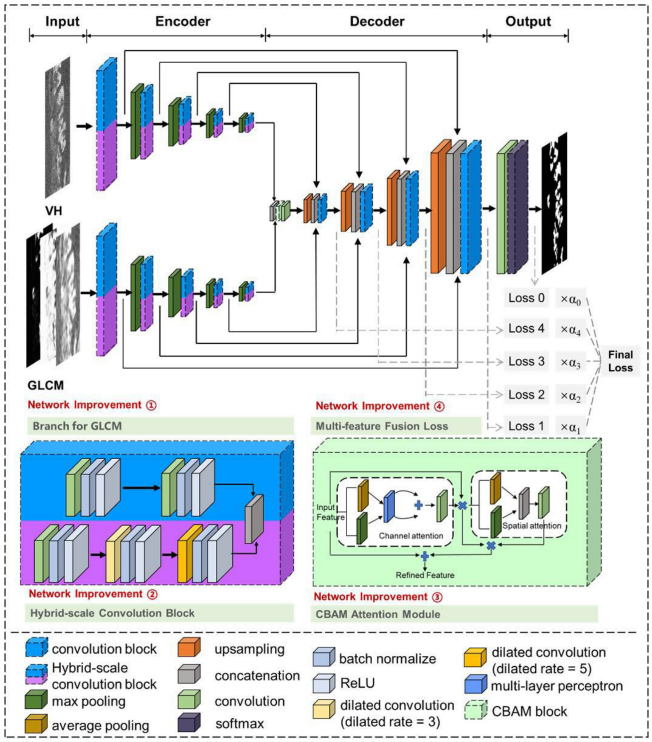

Fig. 6. Architecture of the proposed MAZ-Net with four highlighted modifications: GLCM-based dual-branch encoder, hybrid-scale convolution block, CBAM embedding, and multifeature fusion loss.

In practice, the three supplementary GLCM texture maps are not simply integrated into the same input channel as the SAR image. The MAZ-Net incorporated a dual-branch encoder structure resembling two symmetrical flowcharts. These flowcharts were connected to VH channel SAR images and GLCM texture maps, shown in Fig. 6 encoder section. In each branch, input features are processed with their respective weights and parameters during encoding. Subsequently, the high-level semantic features are combined by concatenation and fused using a convolutional layer. Compared to the single-branch encoder, the dual-branch structure extracts brightness and texture features independently and efficiently, thereby augmenting the model's perception of target features by considering visual-spatial and statistical perspectives [57].

2) Hybrid-Scale Convolution Block: Convolution achieves feature extraction within the receptive fields by linear operations on matrices. Nevertheless, the width of the bright strip in MAZs exceeds 3 pixels, which implies that the 3 × 3 traditional convolution kernels may struggle to capture sufficient contextual information due to their limited receptive field, potentially leading to misclassification of small and bright objects (such as ships and offshore wind turbines) as MAZs. Dilated convolution [58] broadens the receptive field to acquire more comprehensive spatial information, effectively mitigating the adverse effects of speckle noise on SAR images [31]. Moreover, Wang et al. [31] introduced a convolutional structure with a progressively increasing dilated rate to address the issue of information omission resulting from a large receptive field.

Therefore, we proposed a hybrid-scale convolution block, as the module made up of blue and purple modules spliced together at the top and bottom, as shown in Fig. 6. The module incorporates a traditional 3×3 convolution module (the blue module in Fig. 6) in parallel with a progressively expanding dilated rate (1, 3, 5) convolution module (the purple module in Fig. 6). On the one hand, the normal convolution module prevents the ignoration of the small-scale MAZs and other fine-grained information, such as brightness. On the other hand, the dilated convolution module emphasizes large-scale information, such as the regularly spaced alternating dark and bright stripes. These blocks are implanted at all convolutions of the encoder, providing the model with more comprehensive multiscale features.

3) Convolutional Block Attention Module (CBAM): DCNNs often learn redundant features due to translation invariance, overlapping convolution kernels, and interchannel correlations. These networks inherently exhibit translation invariance to recognize features in an image without being sensitive to their exact spatial locations, resulting in learning multiple similar feature representations. Convolution kernels typically overlap and are applied at different positions in the input image, which results in spatially adjacent regions in feature maps containing highly overlapping and redundant information. Multiple convolutional channels are responsible for detecting different features in the images but exhibit correlations to share similar information. While this redundancy can enhance the robustness of networks, it might also raise complexity and computational requirements without delivering significant performance improvements for complicated tasks.

The attention mechanism enables CNNs to concentrate on pertinent image features while suppressing responses to irrelevant information, and it has found extensive application in image segmentation tasks [59]. Woo et al. [60] introduced an innovative CBAM to enhance the filtering of effective features. This versatile module can effectively optimize weights and biases by considering dependencies in convolutional features from both spatial and channel perspectives. Furthermore, the deepest layers of CNNs often contain the most complex and abstract features. As a result, we embedded CBAM after aggregating the highest level information to harness the richest semantic features from where the two encoder branches intersect, thereby increasing the model's sensitivity to MAZs.

4) Multifeature Fusion Loss: Conventional loss functions like cross entropy (CE) adjust parameters solely based on the difference between the final output and the ground truth, potentially leading to a lack of detailed feature learning at different scales or levels of abstraction. Therefore, we proposed a loss function capable of learning hierarchical representations from multiscale feature maps at various levels [61]. This loss function, named multifeature fusion loss, aggregates the output from each stage of the decoder (shown in the Fig. 6 decoder section) and assigns weights accordingly, expressed as follows:

$$L = \alpha_0 \, L_{\mathrm{CE0}} + \alpha_1 \, L_{\mathrm{CE1}} + \cdots + \alpha_4 \, L_{\mathrm{CE4}} \tag{1}$$

where L_{CE0}, L_{CE1}, ..., L_{CE4} represent the CE losses calculated from predictions with ground truth in each decoder stage, as illustrated along the grey dashed line in Fig. 6. The α_0, α_1, ..., α_4 denote the corresponding weights, which were selected as 0.8, 0.05, 0.05, 0.05, and 0.05 during the training process. It is designed to guide the model in learning a representation that effectively balances and combines information from multiple sources (SAR image with GLCM texture maps).

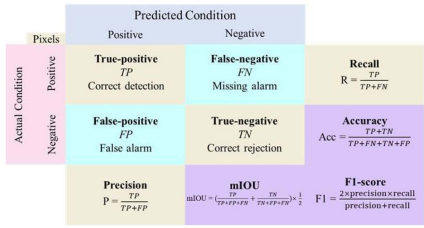

Fig. 7. Evaluation index and calculation formula used in this study.

C. Model Training and Performance

The MAZ-Net was trained and tested under the PyTorch framework. The NVIDIA Tesla A100 graphics processing unit was utilized to accelerate the training and testing of the model. The batch size was set to 8, and the optimizer was Adam [62], with an initial learning rate of 0.0001. The number of epochs was set to 2500 for each experiment. The proposed model has a total of approximately 51.3 million parameters and requires around 28 h for training on our device.

Several commonly used metrics, such as precision, recall, F1-score, accuracy, and mean intersection over union (mIOU), are employed to evaluate the performance of models. The evaluation metrics are shown in Fig. 7. True positive (TP) and false positive (FP) represent the number of pixels correctly and incorrectly predicted as MAZs, while true negative and false negative indicate the number correctly and incorrectly predicted as non-MAZs. High Recall means the model misses a small number of MAZs, whereas high precision suggests the model extracts a small amount of FP. The high mIOU, F1-score, and accuracy indicate high comprehensive performance.

The extraction of MAZs presents a unique and challenging semantic segmentation task due to their complex characteristics in SAR images. Three classes of segmentation methods for MAZs were compared in this study. The first class is the traditional automatic threshold methods, including OTSU [63], a relatively simple but fast image segmentation method. The second is the classic supervised learning segmentation method, Random Forest [64], widely used in SAR-based segmentation tasks [65]. The third is the DCNN-based segmentation method represented by the SegNet [66]. Besides, some recent effective semantic segmentation models are also added to the experiment for comparison with our proposed model MAZ-Net (No. 10 in Table IV), including HRNetV2 [67], UPerNet [68], and Nested U-Net [69] (No. 7 to 9 in Table IV).

Table IV shows the evaluation metric values for each experiment. The OTSU method achieves the highest recall but the lowest precision, suggesting that in large-scale remote sensing image segmentation tasks, it is challenging for automatic

TABLE IV
EVALUATION INDICATORS OF DIFFERENT DATASETS AND METHODS

NO.	Polarization	Model	F1-score	Precise	Recall	Accuracy	mIOU
1	VH	OTSU	79.40%	69.41%	94.47%	98.50%	82.38%
2	VH	Random Forest	82.85%	87.13%	79.34%	99.09%	86.47%
3	VH	SegNet	90.65%	96.80%	85.37%	99.46%	91.22%
4	VH	U-Net	92.72%	95.62%	90.05%	99.57%	93.02%
5	VV	U-Net	84.30%	90.72%	79.23%	99.12%	86.10%
6	VV+VH	U-Net	92.42%	95.48%	89.64%	99.56%	92.76%
7	VH	HRNetV2	86.53%	95.77%	79.23%	99.26%	87.84%
8	VH	UPerNet	94.02%	94.70%	93.35%	99.66%	94.19%
9	VH	Nested U-Net	92.68%	95.63%	89.96%	99.58%	92.98%
10	VH	MAZ-Net	94.77%	96.55%	93.07%	99.70%	94.88%

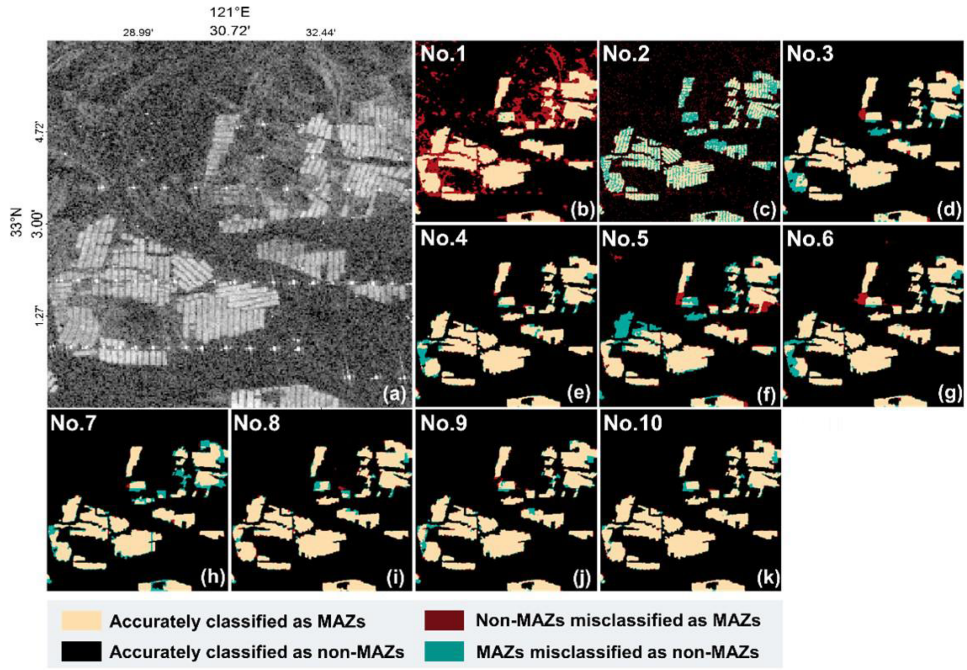

Fig. 8. VH-polarized subimage of the central Tiaozini tidal flats acquired on March 19, 2019, and its corresponding extracted results from different segmentation models.

thresholding methods to offer a threshold range suitable for all small-scale targets. The random forest performs poorly in all metrics because it does not adequately consider the spatial correlation among pixels within the MAZs. SegNet achieves the highest precision with low recall, suggesting poor generalization ability due to its simple structure. Compared to SegNet, the U-Net with the VH channel as input exhibits a higher F1-score and mIOU, implying that it is a more suitable choice as a baseline for this task.

Interestingly, the U-Net with VH-polarized images as input (No. 4 in Table IV) outperforms the VV ones (No. 5 and 6 in Table IV), attributed to the fact that described in Section III: the tidal flats show similar scattering intensities as MAZs in the VV channel, diminishing the distinctiveness of MAZ characteristics. This observation may suggest that, even with disparate ground truth maps, highly similar features can considerably mislead the model during the learning process, ultimately reducing the accuracy of the deep learning model. In contrast, the VH channel's more pronounced backscatter coefficient features provide clearer guidance to the model, resulting in superior performance in this task. UPerNet and Nested U-Net exhibit satisfactory performance in segmenting VH polarization channel inputs (No. 8 and 9 in Table IV). Nevertheless, despite the UPerNet achieving an F1-score of 94.02%, it still falls short of the performance of our model MAZ-Net, indicating the superior overall performance of MAZ-Net designed explicitly for MAZ extraction in this task.

The slice [see Fig. 8(a)] of the partial Subei Sandbnaks poses a significant challenge for MAZ extraction due to their complex topography and dense MAZ distribution. It is, therefore, selected to illustrate the extraction effect of the methods further. Fig. 8(b) and (c) shows the widespread green and red blocks, illustrating traditional extraction methods' difficulty in showing good robustness under complex imaging conditions. Conversely, all the DCNN-based models can accurately extract most MAZs. In addition, a comparison between Fig. 8(e) and

TABLE V
EVALUATION INDICATORS OF ABLATION EXPERIMENTS

NO.	Experiment	F1-score	Precise	Recall	Accuracy	mIOU
1	Exp. 1	94.43%	95.93%	92.98%	99.68%	94.57%
2	Exp. 2	94.14%	95.52%	92.81%	99.67%	94.30%
3	Exp. 3	93.46%	95.46%	91.55%	99.63%	93.67%
4	Exp. 4	92.72%	95.62%	90.05%	99.57%	93.02%
5	Control group (MAZ-Net)	94.77%	96.55%	93.07%	99.70%	94.88%

(f) reveals significantly fewer misclassifications in the former, indicating that the VH-polarized data is more suitable for MAZ extraction than VV-polarized ones. Fig. 8(h) shows widespread green blocks with few red blocks, indicating that HRNetV2 achieves high accuracy at the expense of low recall, as shown in experiment No. 7 in Table IV. While UPerNet, Nested U-Net, and our MAZ-Net demonstrate better segmentation results [see Fig. 8(i)–(k)], they also exhibit more misjudgment and misclassification at the edges of the MAZs than the interior. This result is influenced by tidal fluctuations, with unstable backscattering intensity at the edges of the MAZs. Based on the percentage of cream-colored blocks in Fig. 8(k), it is evident that our proposed MAZ-Net demonstrates the highest segmentation performance and robustness, possibly due to its emphasis on multi-scale information, enhancing the model's ability to capture minute details.

Furthermore, as shown in Table V, we set up a series of experiments to assess the effectiveness of each model optimization [55]. We first set the MAZ-Net as the control group. Then, we substituted the multifeature fusion loss of MAZ-Net with CE loss as Exp. 1. Then, we conducted Exp. 2 by eliminating CBAM. Following this, we replaced the Hybrid-scale convolution block with the convolution module as Exp. 3. Lastly, Exp. 4 involved the removal of the GLCM-informed dual-branch encoder in demonstrating the effectiveness of incorporating GLCM. According to the result shown in Table V, with each optimization incrementally removed from the MAZ-Net, the model's overall performance was gradually decreased, as shown in the change of F1-score from 94.77% to 92.72%, indicating these modifications effectively enhance the complexity of the model. Notably, eliminating the GLCM-informed dual-branch encoder led to the most substantial decrease in performance metrics (see Table V), highlighting significant enhancement from the GLCM-informed dual-branch encoder in the MAZ detection performance.

V. MODEL APPLICATIONS

A. Annual Spatial Distribution From 2016 to 2023

Understanding the annual spatial distribution of MAZs in the Subei Sandbanks is essential for effective planning and management. Therefore, our model was employed to comprehensively extract all MAZs within the Subei Sandbanks region, facilitating an exploration of the Porphyra aquaculture activity over the past eight years in this section. To ensure complete coverage, we selected as many high-quality SAR images as possible from each growing period these years as model inputs (see Table I). Then, we extracted MAZs from each VH-polarized SAR image and superimposed the outputs for each aquaculture cycle. As the annual spatial distribution shown in Fig. 9(a)–(h), the MAZs exhibited a roughly J-shaped pattern along the Subei Sandbanks, with a predominant concentration in the southeastern part of these tidal flats. From 2016 to 2019, these aquaculture zones were installed on each sandbank, with Dongsha having the largest extent of MAZs and Niluosha having the smallest. However, there has been a substantial reduction in the MAZ densities on all sandbanks since 2021, with Dongsha Sandbanks experiencing a particularly pronounced decline that nearly disappeared in the past three years. The phenomenon is more evident in the difference maps [see Fig. 9(i)–(p)]. As the blue zones depicted in Fig. 9(i)–(pc), MAZs remained relatively stable in the first four years, but there has been an overall trend of decreasing the MAZs' scales since 2021. The rapid reduction is attributed to the government's strategic planning for marine aquaculture resources, such as the "Planning of Tidal Flats in Aquaculture Waters in Jiangsu Province (2020–2030)" (http://nynct.jiangsu.gov.cn/), which aims to promote harmonizing environmental protection and coastal cultivation.

Based on the annual extraction results, a further statistical analysis of the spatial distribution frequency of MAZs in the study area from 2016 to 2023 is depicted in Fig. 10(a). One can see that the typical deployment area for MAZs is primarily located in the southern part of the near-shore sandbanks and the eastern part of the offshore sandbanks. The positioning of these aquaculture zones farther from the coastline minimizes exposure to land-based pollution, consequently enhancing the quality and overall health of Porphyra cultivation [38]. We also compared the probability density function (PDF) distribution of the digital elevation model (DEM, data from [35]) of the MAZ-located tidal flats with the PDF of the overall DEM of the study area. As shown in Fig. 10(b), MAZs were situated on relatively flat tidal flats due to the frequent tidal flow that is conducive to the growth of Porphyra here [39].

B. Area Estimation of the Marine Aquaculture Rafts

The statistical data on the spatial extent of marine aquaculture raft installations can be a predictive tool for estimating aquaculture production and contribute valuable information to the fishery statistical yearbook [7]. Nevertheless, these semifloating rafts in MAZs have a narrow width of only three meters [37], rendering them challenging to visually distinguish using Sentinel-1 SAR imagery due to the spatial resolution limitation. For instance, we collected three types of images to scrutinize the variations in raft characteristics across various imaging modes and

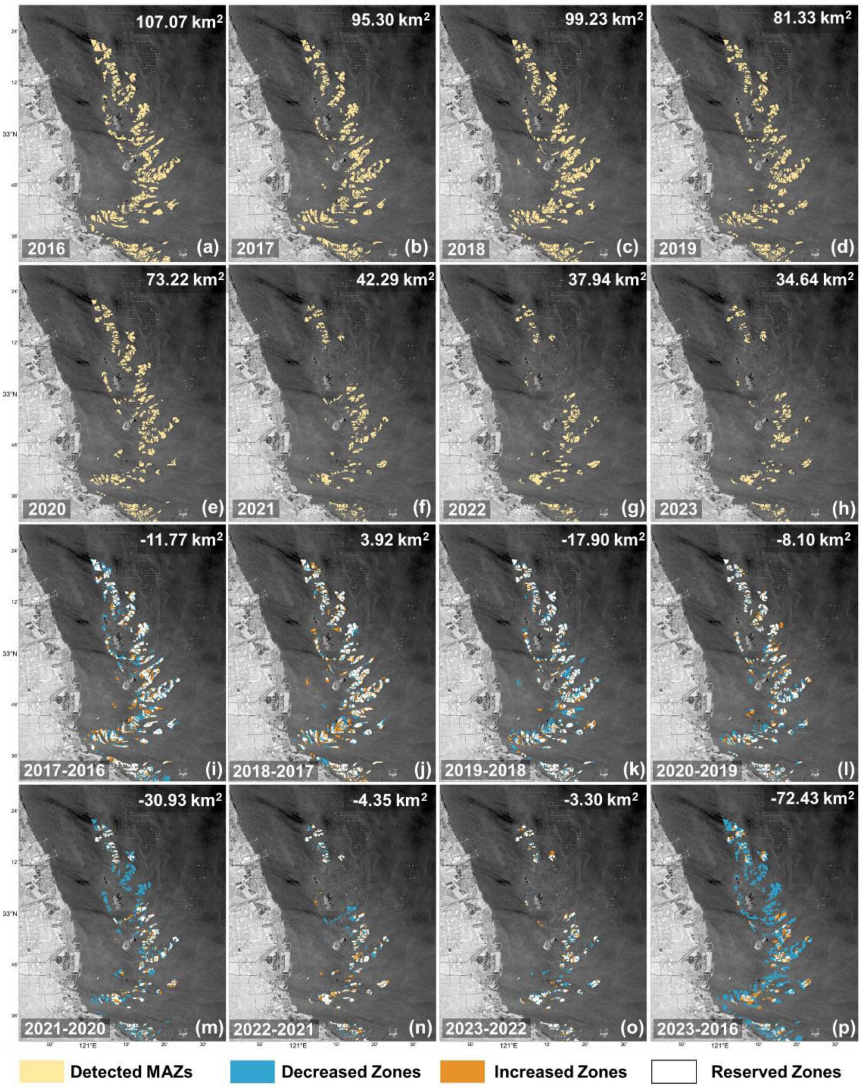

Fig. 9. (a)–(h) Extraction maps for MAZs in Subei Sandbanks from 2016 to 2023, generated by the MAZ-Net; (i)–(o) Annual spatial distribution difference maps of consecutive years; (p) Comparison between the years 2023 and 2016; The area or area changes of rafts are shown in the upper-left corner; The background SAR image acquired on March 13, 2019.

spatial resolutions. Fig. 11(a) shows an optical image acquired on March 7, 2019, by the Pleiades with a fine spatial resolution of 0.5 m. Fig. 11(b) and (c) displays the quasi-contemporaneous Sentinel-2 true-color and Sentinel-1 VH-polarized SAR images at the same spatial resolution of 10 m. The dark lines in Fig. 11(a) delineate the distribution of marine aquaculture rafts in this sea area. However, these lines coalesce into elongated blocks in Fig. 11(b) and (c), owing to their reduced spatial resolution. As the compiled extracted raft areas depicted in Fig. 11(d), one can see that the width of these dark elongated blocks is determined by the length of the lines shown in Fig. 11(a), while their length consists of all the lines' widths and the spaces between them. Interestingly, despite two images capturing the same objects at identical spatial resolutions, the bright stripes in the SAR image generally encompass larger areas than the dark elongated blocks in the optical image. This amplification effect could be attributed to edge blurring caused by speckle noise in the SAR images and the overlapping of backscattering signals emanating from the surrounding tidal flats with those originating from the MAZs.

Furthermore, we calculated the extent of raft coverage in the Pleiades and Sentinel-1 SAR images to establish a scaling factor to estimate raft areas from the MAZ-Net's output. This estimation can be expressed as follows:

$$\text{Area}_{\text{rafts}} = 0.27 \times \text{Area}_{\text{MAZs}} \tag{2}$$

where the $\text{Area}_{\text{rafts}}$ is the estimated area of the rafts in the Subei Sandbanks, and the $\text{Area}_{\text{MAZs}}$ is the MAZs' area statistically derived from extracting results from the proposed method.

The preceding analysis delineates the impact of resolution and imaging mechanisms on visually interpreting semifloating rafts and MAZs, which establishes a theoretical foundation for estimating raft areas through model extraction results. To monitor the interannual production changes in the Subei Sandbanks, we calculated the area from 2016 to 2023 using (2). As shown in

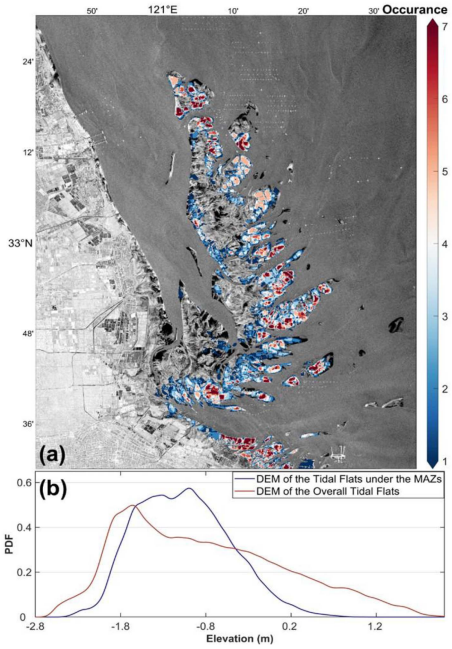

Fig. 10. (a) Occurrence of MAZ distribution in Subei Sandbank from 2016 to 2023, overlaid on the SAR image acquired on March 19, 2019. (b) PDF distribution statistics of MAZ-located tidal flats' DEM (blue line) and the overall DEM of Subei Sandbanks (red line).

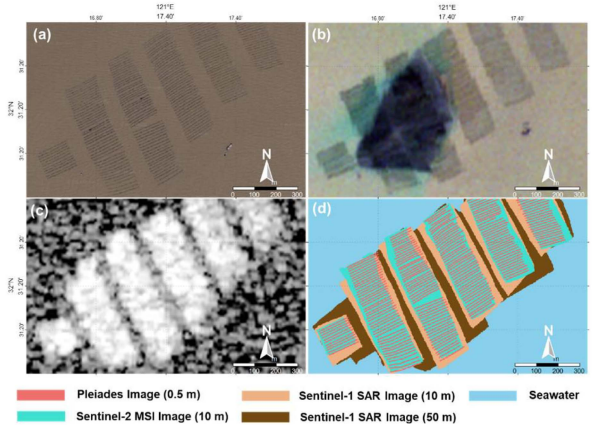

Fig. 11. (a) Pleiades true-color image, (b) Sentinel-2 false-color image, and (c) Sentinel-1 VH-polarized image of the MAZs in the Subei Sandbanks acquired on March 7, 2019, January 24, 2019, and February 5, 2019, respectively; (d) Extracted rafts from (a), (b), and (c), as well as the training label with a resolution of 50 m in the same zone.

Fig. 12, the area of the aquaculture rafts in the Subei Sandbanks was estimated to be approximately 100 km2 during 2016–2018, experiencing a slight decline in 2019–2020. However, starting in 2020, there was a significant and sustained reduction in the raft area, stabilizing at less than 50 km2. By 2023, the raft area had dwindled to 34.64 km2, decreasing to 32.35% of its 2016 extent. Surprisingly, the yellow sea green tide in 2021 is more extensive than in 2020 [55], suggesting that the outbreak of the green tide in the Yellow Sea might be regulated by multiple factors [70],

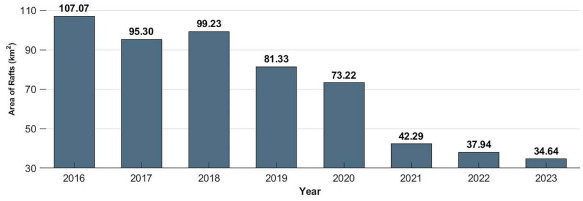

Fig. 12. Estimation of the annual area of marine aquaculture rafts from 2016 to 2023.

rather than being solely dominated by Porphyra aquaculture in the Subei Sandbanks.

VI. Conclusion

Deep learning has demonstrated superior performance to classical techniques in extracting valuable information across various applications over the past few years [23]. Compared to conventional photographs, applying deep-learning-based methods to remote sensing images poses greater challenges due to their unique characteristics, including geographic information, multiple spectral/polarized bands, and variations in spatial resolution. Both optimizing the input and enhancing the model's architecture are integral components of improving the performance of a deep learning model.

On the one hand, unselected input data introduce irrelevant information to the model [71]. For instance, high tidal level or copolarized SAR data in our task injects noise and redundancy into the extraction model because the MAZ features become less distinct or ambiguous under such conditions. The improvement achieved by using cross-polarized and additional texture information implies that incorporating domain knowledge into an input can guide data processing and feature engineering decisions to aid the model's comprehension of this task. On the other hand, additional inputs impose requirements on the model structure to ensure effective handling of the extra information because the original model may not have been well-suited to adapt to the diversity of the data it encounters [72]. However, an excessively simplistic model may lack the capacity to capture the intricate patterns and features in the data, whereas an overly complex model can result in overfitting and the inefficient use of resources. It suggests that effective model complexity is necessary to enable the model to learn higher level feature representations, facilitating better fitting to the training data and enhancing generalization ability [73].

However, as mentioned earlier, the availability of SAR imagery is influenced by tidal fluctuations, restricting the frequency of monitoring MAZs. It makes sense to increase the monitoring frequency, especially during the recovery of the aquaculture facility (April to May), which can significantly contribute to the early warning of green tide outbreak [37]. In the future, optical imageries, such as Sentinel-2 and Landsat satellite images, and additional SAR imageries, like TerraSAR-X and GaoFen-3 images, will be attempted to be incorporated into the deep learning model in order to enhance the number of effective observations of MAZs during critical periods. Besides, the MAZ extraction

results of MAZ-Net indicate a higher likelihood of misjudgment and misclassification along the edges of MAZs. This drawback may stem from overlapping backscatter signals or fluctuations in seawater, resulting in unstable backscattering intensity at the MAZ edges. Hence, future studies will prioritize improving the model's sensitivity to MAZ edges to reduce inaccuracies in MAZ edge extraction.

In this study, we first analyzed the polarization characteristics of semifloating raft aquaculture zones in the Subei Sandbanks utilizing Sentinel-1 dual-polarized SAR imagery. Due to the volume scattering from the net screens of rafts, the MAZs' features in the VH channel exhibit superior recognizability and are less susceptible to the influence of sea conditions than copolarization data, especially evident during low tidal levels. Therefore, prior domain knowledge was incorporated to suggest a deep-learning-based extraction model, named MAZ-Net, involving input optimization and model architectural enhancement. The input optimization entailed two strategies: 1) constructing the dataset using cross-polarized images acquired at a low tidal level and with the intact raft structure to reduce the misleading information; 2) introducing relevant GLCM texture maps as the additional input to augment the feature injection. The model structure enhancement based on the classic U-Net framework encompasses three key facets:

1) designing the hybrid-scale convolution block to capture MAZ features across various scales;
2) embedding the CBAM module to enhance the model's focus on pertinent channels and regions;
3) using the multifeature Fusion loss function to guide the model in maintaining an effective equilibrium among multi-source and multiscale information.

The performance evaluation (F1-score 94.77%) indicated that the prior knowledge-driven optimizations significantly improved the automatic extraction of MAZs in Subei Sandbanks from SAR imagery under varying imaging conditions.

Moreover, we utilized the proposed model to map the spatial distributions of MAZs and estimate the areas occupied by marine aquaculture rafts in Subei Sandbanks spanning 2016 to 2023. The outputs unveiled a substantial decline in the expanse of MAZs within the Subei Sandbanks after 2020, aligning with the implementation of stringent governmental regulations on marine aquaculture resources. However, it is intriguing to note that the Yellow Sea green tide has not exhibited a decrease in recent years despite the reduction in MAZs within the Subei Sandbanks, highlighting that the outbreak of green tide in the Yellow Sea is a complicated ecological phenomenon likely influenced by multiple contributing factors, not solely reliant on the Porphyra farming within the Subei Sandbanks.

ACKNOWLEDGMENT

The authors would like to thank the European Space Agency and NASA for providing the Sentinel-1 images (https://scihub.copernicus.eu and https://search.asf.alaska.edu/) and the Oregon State University for providing TPXO tide model and TMD toolbox (https://www.tpxo.net/). They also would like to thank the European Centre for Medium-Range Weather Forecasts for providing the ERA5 10 m level sea winds data (https://cds.climate.copernicus.eu/), the Airbus Defence and Space for the Pleiades 0.5-m image (https://data.oneatlas.airbus.com/), and Dr. Qing-Chun Zhang for providing the field survey photograph used in this research.

REFERENCES

[1] H. E. Froehlich, R. R. Gentry, and B. S. Halpern, "Global change in marine aquaculture production potential under climate change," *Nature Ecol. Evol.*, vol. 2, no. 11, pp. 1745–1750, 2018.
[2] R. R. Gentry et al., "Mapping the global potential for marine aquaculture," *Nature Ecol. Evol.*, vol. 1, no. 9, pp. 1317–1324, 2017.
[3] FAO, *The State of World Fisheries and Aquaculture*, Food Agriculture Org., Rome, Italy, 2022.
[4] V. Smetacek and A. Zingone, "Green and golden seaweed tides on the rise," *Nature*, vol. 504, no. 7478, pp. 84–88, 2013.
[5] Z. Wang, J. Xiao, S. Fan, Y. Li, X. Liu, and D. Liu, "Who made the world's largest green tide in China?—An integrated study on the initiation and early development of the green tide in yellow sea," *Limnology Oceanogr.*, vol. 60, no. 4, pp. 1105–1117, 2015.
[6] D. Peng, Y. Zhu, and J. Chu, "Strengthen management of offshore aquaculture," *Science*, vol. 381, no. 6661, pp. 955–955, 2023.
[7] Y. Liu et al., "Satellite-based monitoring and statistics for raft and cage aquaculture in China's offshore waters," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 91, 2020, Art. no. 102118.
[8] H. Zhang et al., "Optical and SAR image dense registration using a robust deep optical flow framework," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 1269–1294, 2023.
[9] S.-K. Lee, S.-H. Hong, S.-W. Kim, Y. Yamaguchi, and J.-S. Won, "Polarimetric features of oyster farm observed by AIRSAR and JERS-1," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2728–2735, Oct. 2006.
[10] B.-H. Choe, D.-j. Kim, J.-H. Hwang, Y. Oh, and W. M. Moon, "Detection of oyster habitat in tidal flats using multi-frequency polarimetric SAR data," *Estuarine, Coastal Shelf Sci.*, vol. 97, pp. 28–37, 2012.
[11] M. Gade and S. Melchionna, "Joint use of multiple synthetic aperture radar imagery for the detection of bivalve beds and morphological changes on intertidal flats," *Estuarine, Coastal Shelf Sci.*, vol. 171, pp. 1–10, 2016.
[12] S. Schmitz, E. Wortmeyer, A. Thiele, H. Dirks, and A. Wurpts, "Detection of mussel beds using airborne polarimetric SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2021, pp. 7501–7504.
[13] W. Wang and M. Gade, "A PolSAR classification algorithm and its verification for intertidal flats in the German wadden SEA," in *Proc. IGARSS IEEE Int. Geosci. Remote Sens. Symp.*, 2022, pp. 5156–5159.
[14] J. Fan, J. Zhao, W. An, and Y. Hu, "Marine floating raft aquaculture detection of GF-3 PolSAR images based on collective multikernel fuzzy clustering," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 8, pp. 2741–2754, Aug. 2019.
[15] J. Geng, J. Fan, and H. Wang, "Weighted fusion-based representation classifiers for marine floating raft detection of SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 3, pp. 444–448, Mar. 2017.
[16] B. Cheng, Y. Liu, X. Liu, G. Wang, and X. Ma, "Research on extraction method of coastal aquaculture areas on high resolution remote sensing image based on multi-features fusion," *Remote Sens. Technol. Application*, vol. 33, no. 2, pp. 296–304, 2018.
[17] Z. Wang, X. Yang, Y. Liu, and C. Lu, "Extraction of coastal raft cultivation area with heterogeneous water background by thresholding object-based visually salient NDVI from high spatial resolution imagery," *Remote Sens. Lett.*, vol. 9, no. 9, pp. 839–846, 2018.
[18] B. Ai, P. Wang, Z. Yang, Y. Tian, and D. Liu, "Spatiotemporal dynamics analysis of aquaculture zones and its impact on green tide disaster in Haizhou Bay, China," *Mar. Environ. Res.*, vol. 183, 2023, Art. no. 105825.
[19] M. Wang, Q. Cui, J. Wang, D. Ming, and G. Lv, "Raft cultivation area extraction from high resolution remote sensing imagery by fusing multi-scale region-line primitive association features," *ISPRS J. Photogrammetry Remote Sens.*, vol. 123, pp. 104–113, 2017.
[20] J. Cheng, N. Jia, R. Chen, X. Guo, J. Ge, and F. Zhou, "High-resolution mapping of seaweed aquaculture along the Jiangsu coast of China using Google Earth Engine (2016–2022)," *Remote Sens.*, vol. 14, no. 24, 2022, Art. no. 6202.

[21] T. Hou, W. Sun, C. Chen, G. Yang, X. Meng, and J. Peng, "Marine floating raft aquaculture extraction of hyperspectral remote sensing images based decision tree algorithm," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 111, 2022, Art. no. 102846.

[22] B. Liu, X. Li, and G. Zheng, "Coastal inundation mapping from bitemporal and dual-polarization SAR imagery based on deep convolutional neural networks," *J. Geophysical Res.: Oceans*, vol. 124, no. 12, pp. 9101–9113, 2019.

[23] X. Li et al., "Deep-learning-based information mining from ocean remote-sensing imagery," *Nat. Sci. Rev.*, vol. 7, no. 10, pp. 1584–1605, 2020.

[24] L. Gao, X. Li, F. Kong, R. Yu, Y. Guo, and Y. Ren, "AlgaeNet: A deep-learning framework to detect floating green algae from optical and SAR imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 2782–2796, Mar. 2022.

[25] Y. Ren, X. Li, and H. Xu, "A deep learning model to extract ship size from Sentinel-1 SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, Mar. 2022.

[26] Y. Fu et al., "A new satellite-derived dataset for marine aquaculture areas in China's coastal region," *Earth Syst. Sci. Data*, vol. 13, no. 5, pp. 1829–1842, 2021.

[27] J. Fan, J. Zhou, X. Wang, and J. Wang, "A self-supervised transformer with feature fusion for SAR image semantic segmentation in marine aquaculture monitoring," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–15, Mar. 2023.

[28] J. Liu, Y. Lu, X. Guo, and W. Ke, "A deep learning method for offshore raft aquaculture extraction based on medium-resolution remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 6296–6309, Jul. 2023.

[29] B. Cui, Y. Zhao, M. Yang, L. Huang, and Y. Lu, "Reverse attention dual-stream network for extracting laver aquaculture areas from GF-1 remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 5271–5283, Jun. 2023.

[30] Y. Han et al., "Dynamic mapping of inland freshwater aquaculture areas in Jianghan plain, China," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 4349–4361, Apr. 2023.

[31] J. Wang, J. Fan, and J. Wang, "MDOAU-Net: A lightweight and robust deep learning model for SAR image segmentation in aquaculture raft monitoring," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, Jan. 2022.

[32] X. Wang, J. Zhou, and J. Fan, "IDUDL: Incremental double unsupervised deep learning model for marine aquaculture SAR images segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–12, Aug. 2022.

[33] Y. Wang, Y. Liu, S. Jin, C. Sun, and X. Wei, "Evolution of the topography of tidal flats and sandbanks along the Jiangsu coast from 1973 to 2016 observed from satellites," *ISPRS J. Photogrammetry Remote Sens.*, vol. 150, pp. 27–43, 2019.

[34] W.-F. Ni, Y.-P. Wang, X.-Q. Zou, J.-C. Zhang, and J.-H. Gao, "Sediment dynamics in an offshore tidal channel in the southern yellow sea," *Int. J. Sediment Res.*, vol. 29, no. 2, pp. 246–259, 2014.

[35] S. Zhang, Q. Xu, H. Wang, Y. Kang, and X. Li, "Automatic waterline extraction and topographic mapping of tidal flats from sar images based on deep learning," *Geophysical Res. Lett.*, vol. 49, no. 2, 2022, Art. no. e2021GL096007.

[36] G. Deng, H. Wu, P. Guo, and M. Li, "Evolution and development trend of marine raft cultivation model in China," *Chin. Fish. Econ.*, vol. 31, pp. 164–169, 2013.

[37] Q. Xing et al., "Monitoring seaweed aquaculture in the yellow sea with multiple sensors for managing the disaster of macroalgal blooms," *Remote Sens. Environ.*, vol. 231, 2019, Art. no. 111279.

[38] Q. Qi, "Seedling and culture technology of porphyra yezensis striata," *Special Econ. Animals Plants*, vol. 15, pp. 44–45, 2012.

[39] J. Tang, "Technical instructions for porphyra yezensis culture," *J. Aquaculture*, vol. 31, pp. 39–41, 2010.

[40] C. Zhao, C.-Z. Qin, and J. Teng, "Mapping large-area tidal flats without the dependence on tidal elevations: A case study of southern China," *ISPRS J. Photogrammetry Remote Sens.*, vol. 159, pp. 256–270, 2020.

[41] G. D. Egbert and S. Y. Erofeeva, "Efficient inverse modeling of barotropic ocean tides," *J. Atmospheric Ocean. Technol.*, vol. 19, no. 2, pp. 183–204, 2002.

[42] S. Zhang, Q. Xu, Q. Zheng, and X. Li, "Mechanisms of SAR imaging of shallow water topography of the Subei bank," *Remote Sens.*, vol. 9, no. 11, 2017, Art. no. 1203.

[43] K. Millard, S. Darling, N. Pelletier, and S. Schultz, "Seasonally-decomposed Sentinel-1 backscatter time-series are useful indicators of peatland wildfire vulnerability," *Remote Sens. Environ.*, vol. 283, 2022, Art. no. 113329.

[44] M. Gade, W. Wang, and L. Kemme, "On the imaging of exposed intertidal flats by single-and dual-co-polarization synthetic aperture radar," *Remote Sens. Environ.*, vol. 205, pp. 315–328, 2018.

[45] Q. Zhang et al., "Green algae detached from aquaculture rafts into seawater resulted in green tide occurrence in the yellow sea," *Oceanologia et Limnologia Sinica*, vol. 49, pp. 1014–1020, 2018.

[46] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervention*, 2015, pp. 234–241.

[47] Y. Wang et al., "Subseasonal prediction of regional Antarctic sea ice by a deep learning model," *Geophysical Res. Lett.*, vol. 50, no. 17, 2023, Art. no. e2023GL104347.

[48] Y. Ren, X. Li, and W. Zhang, "A data-driven deep learning model for weekly sea ice concentration prediction of the Pan-Arctic during the melting season," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–19, May 2022.

[49] Y. Ren, X. Li, X. Yang, and H. Xu, "Development of a dual-attention U-Net model for sea ice and open water classification on SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, Feb. 2022.

[50] Y. Liu, Q. Zheng, and X. Li, "Characteristics of global ocean abnormal mesoscale eddies derived from the fusion of sea surface height and temperature data by deep learning," *Geophysical Res. Lett.*, vol. 48, no. 17, 2021, Art. no. e2021GL094772.

[51] H. Wang and X. Li, "Deepblue: Advanced CNN applications for ocean remote sensing," *IEEE Geosci. Remote Sens. Mag.*, vol. 12, no. 1, pp. 138–161, Mar. 2024, doi: 10.1109/MGRS.2023.3343623.

[52] A. Akhtarmanesh, D. Abbasi-Moghadam, A. Sharifi, M. H. Yadkouri, A. Tariq, and L. Lu, "Road extraction from satellite images using attention-assisted UNet," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 1126–1136, Nov. 2024.

[53] Z. Huang, M. Datcu, Z. Pan, and B. Lei, "Deep SAR-Net: Learning objects from signals," *ISPRS J. Photogrammetry Remote Sens.*, vol. 161, pp. 179–193, 2020.

[54] D. Patil, K. Patil, R. Nale, and S. Chaudhari, "Semantic segmentation of satellite images using modified U-Net," in *Proc. IEEE Region 10 Symp.*, 2022, pp. 1–6.

[55] Y. Guo, L. Gao, and X. Li, "A deep learning model for green algae detection on SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, Oct. 2022.

[56] S. Mu, X. Li, and H. Wang, "The fusion of physical, textural, and morphological information in SAR imagery for hurricane wind speed retrieval based on deep learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–13, May. 2022.

[57] Y. Zhou, K. Chen, and X. Li, "Dual-branch neural network for sea fog detection in geostationary ocean color imager," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–17, Aug. 2022.

[58] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2015, *arXiv:1511.07122*.

[59] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 11531–11539.

[60] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 3–19.

[61] H. Huang et al., "Unet 3+: A full-scale connected Unet for medical image segmentation," in *Proc. ICASSP IEEE Int. Conf. Acoust., Speech Signal Process.*, 2020, pp. 1055–1059.

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

[63] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.

[64] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, 2001.

[65] F. Mohammadimanesh, B. Salehi, M. Mahdianpari, E. Gill, and M. Molinier, "A new fully convolutional neural network for semantic segmentation of polarimetric SAR imagery in complex land cover ecosystem," *ISPRS J. Photogrammetry Remote Sens.*, vol. 151, pp. 223–236, 2019.

[66] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[67] K. Sun et al., "High-resolution representations for labeling pixels and regions," Apr. 2019, *arXiv:1904.04514*.

[68] T. Xiao, Y. Liu, B. Zhou, Y. Jiang, and J. Sun, "Unified perceptual parsing for scene understanding," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 418–434.

[69] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Proc. 4th Int. Workshop, Deep Learn. Med. Image Anal. 8th Int. Workshop Multimodal Learn. Clin. Decis. Support*, 2018, pp. 3–11.
[70] L. Zheng et al., "What causes the great green tide disaster in the south yellow sea of China in 2021?," *Ecological Indicators*, vol. 140, 2022, Art. no. 108988.
[71] Z. Huang, X. Yao, and J. Han, "Progress and perspective on physically explainable deep learning for synthetic aperture radar image interpretation," *J. Radar*, vol. 11, no. 1, pp. 107–125, 2021.
[72] X. Li, Y. Zhou, and F. Wang, "Advanced information mining from ocean remote sensing imagery with deep learning," *J. Remote Sens.*, vol. 2022, 2022, Art. no. 9849645.
[73] Z. Huang, X. Yao, Y. Liu, C. O. Dumitru, M. Datcu, and J. Han, "Physically explainable CNN for SAR image classification," *ISPRS J. Photogrammetry Remote Sens.*, vol. 190, pp. 25–37, 2022.

Wantai Chen (Student Member, IEEE) received the B.S. degree in oceanography from the Ocean University of China, Qingdao, China, in 2022. He is currently working toward the M.S. degree in physical oceanography with the Institute of Oceanology, Chinese Academy of Sciences, Qingdao, and the University of Chinese Academy of Sciences, Beijing, China.

His research interests are in coastal zone monitoring with synthetic aperture radar (SAR) and artificial intelligence oceanography.

Xiaofeng Li (Fellow, IEEE) received the bachelor's degree in optical engineering from Zhejiang University, Hangzhou, China, in 1985, the master's degree in physical oceanography from the First Institute of Oceanography, Qingdao, China, in 1992, and the Ph.D. degree in physical oceanography from North Carolina State University, Raleigh, NC, USA, in 1997.

From 1997 to 2019, he was with the National Environmental Satellite, Data, and Information Service (NESDIS), National Ocean and Atmospheric Administration (NOAA), Washington, DC, USA, where he was involved in developing many operational satellite ocean remote sensing products. He is with the Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China. His research interests include synthetic aperture radar (SAR) applications in oceanography and marine meteorology, the application of artificial intelligence in oceanographic studies, big data analytics, and satellite image processing.

Dr. Li holds editorial positions as an Associate Editor for both IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the *International Journal of Remote Sensing*, as an Editor for *JGR: Machine Learning and Computation*, and as the Executive Editor-in-Chief for the *Journal of Remote Sensing*.