MVCNN: A Deep Learning-Based Ocean–Land Waveform Classification Network for Single-Wavelength LiDAR Bathymetry

Gang Liang, Xinglei Zhao, *Member, IEEE*, Jianhu Zhao, and Fengnian Zhou

***Abstract*—Ocean–land waveform classification (OLWC) is crucial in airborne LiDAR bathymetry (ALB) data processing and can be used for ocean–land discrimination and waterline extraction. However, the accuracy of OLWC for single-wavelength ALB systems is low given the nature of the green laser waveform in complex environments. Thus, in this article, a deep learning-based OLWC method called the multichannel voting convolutional neural network (MVCNN) is proposed based on the comprehensive utilization of multichannel green laser waveforms. First, multiple green laser waveforms collected in deep and shallow channels are input into a multichannel input module. Second, a one-dimensional (1-D) convolutional neural network (CNN) structure is proposed to handle each green channel waveform. Finally, a multichannel voting module is introduced to perform majority voting on the predicted categories derived by each 1-D CNN model and output the final waveform category (i.e., ocean or land waveforms). The proposed MVCNN is evaluated using the raw green laser waveforms collected by Optech coastal zone mapping and imaging LiDAR (CZMIL). Results show that the overall accuracy, kappa coefficient, and standard deviation of the overall accuracy for the OLWC utilizing green laser waveforms based on MVCNN can reach 99.41%, 0.9800, and 0.03%, respectively. Results further show that the classification accuracy of the MVCNN is improved gradually with the increase in the number of laser channels. The multichannel voting module can select the correct waveform category from the deep and shallow channels. The proposed MVCNN is highly accurate and robust, and it is slightly affected by aquaculture rafts and the merging effect of green laser waveform in very shallow waters. Thus, the use of MVCNN in OLWC for single-wavelength ALB systems is recommended. In addition, this article explores the relationships between green deep and shallow channel waveforms based on the analysis of CZMIL waveform data.**

***Index Terms*—Airborne lidar bathymetry, deep learning, multichannel green laser waveforms, multichannel voting, ocean–land waveform classification.**

Manuscript received 27 June 2022; revised 15 September 2022, 14 October 2022, and 12 November 2022; accepted 8 December 2022. Date of publication 14 December 2022; date of current version 23 December 2022. This work was supported by the National Natural Science Foundation of China under Grant 41906166. *(Corresponding author: Xinglei Zhao.)*

Gang Liang and Xinglei Zhao are with the College of Information Science and Engineering, Shandong Agricultural University, Tai'an 271018, China (e-mail: 2020110670@sdau.edu.cn; xingleizhao@126.com).

Jianhu Zhao is with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China (e-mail: jhzhao@sgg.whu.edu.cn).

Fengnian Zhou is with the Survey Bureau of Hydrology and Water Resources of Yangtze Estuary, Shanghai 200136, China (e-mail: cjkfnzhou@126.com).

Digital Object Identifier 10.1109/JSTARS.2022.3229062

I. Introduction

AIRBORNE LiDAR bathymetry (ALB) uses a scanning, pulsed, and green laser beam for shallow water measurements, which has the advantages of high efficiency and flexibility [1], [2], [3], [4], [5], [6], [7], [8], [9]. Apart from depth measurements, ALB has been applied to ocean–land discrimination, suspended sediment concentration monitoring, habitat mapping, ocean wave pattern analysis, and seabed classification [9], [10], [11], [12]. Presently, many advanced and commercial ALB systems exist, such as Optech coastal zone mapping and imaging LiDAR (CZMIL), Optech Aquarius, Optech SHOALS series, AHAB HawkEye III, Fugro LADS-MK3, USGS experimental advanced airborne research LiDAR (EAARL), and RIEGL VQ-880-G [13], [14]. Single-wavelength ALB systems, such as Fugro LADS-MK3, use a single green laser that penetrates typical waters with the least attenuation to detect the water surface and bottom [2]. Dual-wavelength ALB systems, such as Optech CZMIL, emit a green laser to detect the water bottom and an additional infrared (IR) laser to accurately detect the water surface and overcome the water surface uncertainty problem of green lasers [2], [5].

A typical ALB system is equipped with a laser sensor, an inertial navigation device, a global navigation satellite system receiver, and an integrated digital camera if necessary [15]. Photomultiplier tubes or avalanche photodiodes in laser sensors are used to detect return pulses, which are then digitized and recorded per nanosecond as return laser full waveforms [16]. ALB systems can integrate ocean and land measurements and receive ocean and land return waveforms. Ocean–land waveform classification (OLWC) is vital for ALB systems for the following reasons.

1) OLWC is crucial in ALB data processing. The components of ocean and land return waveforms differ from one another. On the one hand, a typical green laser waveform of land comprises multiple returns reflected from vegetation and land surfaces. On the other hand, a typical green laser waveform of oceans with clear water comprises three returns, namely, water surface, volume backscatter, and water bottom returns [1], [14]. Waveform processing purposes, techniques, and methods for land and ocean waveforms also differ from one another. Thus, ocean and land waveforms must be properly classified.

2) OLWC can be used for ocean–land discrimination and waterline extraction by integrating the waveform categories (ocean or land) derived by OLWC and the corresponding positions of laser irradiation points [14].

OLWC was performed for the single- and dual-wavelength ALB systems.

1) Dual-Wavelength ALB Systems: IR and green lasers are used in dual-wavelength ALB systems; thus, IR and green laser waveforms can be used for OLWC. OLWC utilizing IR laser waveforms achieves high accuracy due to the significant difference in IR laser waveform features of ocean and land [9], [13]. For example, the amplitudes of IR laser waveforms of land are typically larger than those of ocean [9], [13], [14]. An IR laser waveform saturation method was proposed [1] and verified as an effective OLWC method for the Optech SHOALS series [12]. With the development of ALB systems, full waveforms with a wide amplitude range can be digitized and recorded [16]; thus, the IR laser waveform of land is no longer saturated. The dual-clustering method was proposed to solve the limitation of the IR laser waveform saturation method by performing K-means clustering for IR laser waveform amplitudes and density-based spatial clustering of applications with noise (DBSCAN) for laser points, achieving an overall accuracy of nearly 100% [9].

Furthermore, features derived from green laser waveforms can help improve the overall accuracy of OLWC [13], [17], [18], [19], [20]. Waveform features, such as amplitude, full width at half maxima (FWHM), area, and amplitude ratio of IR and green lasers, can be extracted and combined to perform OLWC by using support vector machines (SVMs) [13], [18], [19]. A one-dimensional (1-D) convolutional neural network (CNN) was used to perform OLWC for dual-wavelength ALB systems by automatically extracting waveform features using convolutional layers [20].

2) Single-Wavelength ALB Systems: Ideally, OLWC can be realized using only green laser waveforms [21]. High-accuracy OLWC utilizing only green laser waveforms without additional information, such as IR laser waveforms, is extremely important, especially for single-wavelength ALB systems [14]. Unfortunately, this is not an easy goal to achieve given the nature of green laser waveforms in complex environments [14], [21]. Water surface, volume backscatter, and water bottom returns in the green laser waveform of very shallow waters cannot be discriminated and merged into a single return [22]. This merging effect causes the green laser waveforms of very shallow waters to be similar to those of land [14]. Meanwhile, land topography also causes some green laser waveforms of land to be similar to the green laser waveforms of oceans [14]. These special waveforms are misclassified. Currently, OLWC utilizing green laser waveform features via fuzzy c-means clustering (FCM) [14] or SVM [13] can achieve an overall accuracy of less than 94%, which is significantly lower than that using IR laser waveforms.

In summary, OLWC for dual-wavelength ALB systems that depend on IR and green laser waveforms is reliable and accurate. However, the accuracy of OLWC for single-wavelength ALB systems that depends only on green laser waveforms is low and must be improved. With the development of ALB devices, ALB can collect multichannel waveforms by segmenting the field of view (FOV) of receiver telescopes. Different channel waveforms vary in their responses to the ocean and land environment. The accuracy of OLWC can be improved if multichannel green waveforms are comprehensively used. Thus, this article proposes a deep learning-based OLWC method called the multichannel voting CNN (MVCNN), which comprehensively utilizes multichannel green laser waveforms by performing majority voting, to improve the accuracy of OLWC for single-wavelength ALB systems.

II. Data and Methods

A. Study Area and Data

The ALB data used in this article were collected using Optech CZMIL around Qinshan Island of Lianyungang City, Jiangsu Province, China. The detailed descriptions of the experimental area can be found in [9]. The nominal parameters of the CZMIL system for laser pulse repetition rate, scanner frequency, scan angle, flight altitude, flight velocity, strip width, maximum measurable depth, and minimum measurable depth are 10 kHz, 27 Hz, 20°, 400 m, 140 kts, 294 m, $4.2/K_d$ (bottom power reflectivity $> 15\%$), and 0.15 m, respectively [23]. The CZMIL system adopts a collinear and synchronous method to emit green and IR laser pulses with a circular beam scanning pattern [24], [25]. The CZMIL optical system uses a rotating Fresnel prism to emit a laser beam and multiple channels to collect return laser waveforms [16]. CZMIL is a dual-wavelength ALB system, but it can be treated as a single-wavelength ALB system when only the green laser waveforms collected in its deep and shallow channels are used for OLWC. The IR laser waveforms of the CZMIL system can be used as a highly accurate external reference for OLWC utilizing green laser waveforms.

Fig. 1(a) shows the location of the study area and the track line of three strips, i.e., strips C14–C16, with the arrows representing the flight directions. Each strip contains ocean and land waveforms. The C14 strip is selected as the training set to train the model, the C15 strip is selected as the validation set to adjust the network structure and parameters, and the C16 strip is used as the test set to evaluate the model. The CZMIL system can collect nine laser waveforms by using multiple channels for each laser shot, and the effective signal length of each waveform is 320 bins. There are 304322, 286328, and 294654 laser shots in the training, validation, and test sets, respectively. Typical multichannel green laser waveforms of the ocean and land collected by the CZMIL system are shown in Figs. 2 and 3, respectively. The positions of the laser irradiation points (i.e., longitude, latitude, and elevation) of the corresponding waveforms can be calculated by merging the propagation distances of the laser pulses, the scanning angles of the laser beams, the position and orientation data of the sensor, and the configuration parameters for system installation [15]. The locations of the laser irradiation points of the corresponding typical ocean (blue triangle) and land (red triangle) waveforms are shown in Fig. 1(a). The color shaded-relief image obtained using the laser points is shown in Fig. 1(b).

Traditional machine learning (ML) requires time-consuming waveform analysis and manual feature extraction, and the fixed

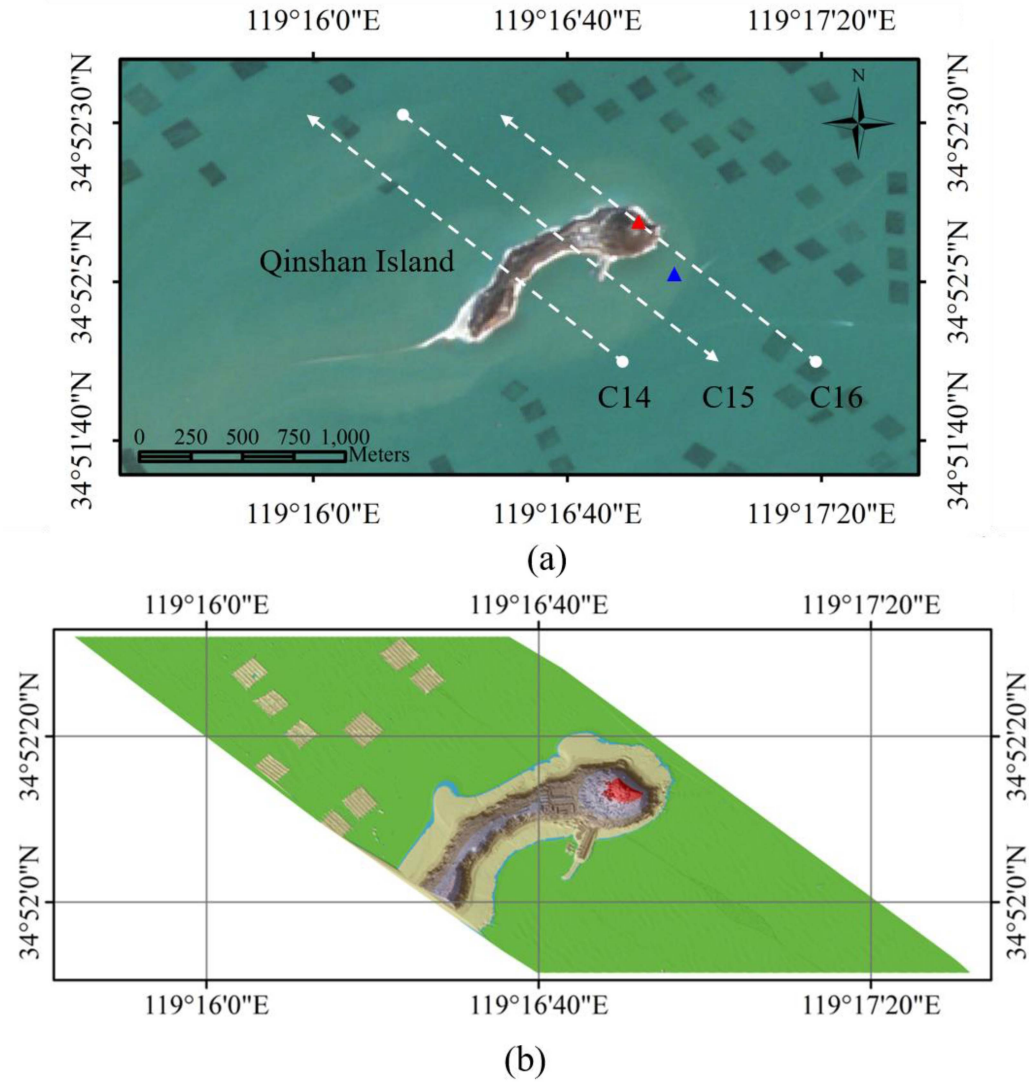

Fig. 1. Study area. (a) Location of the study area. (b) Color shaded-relief image.

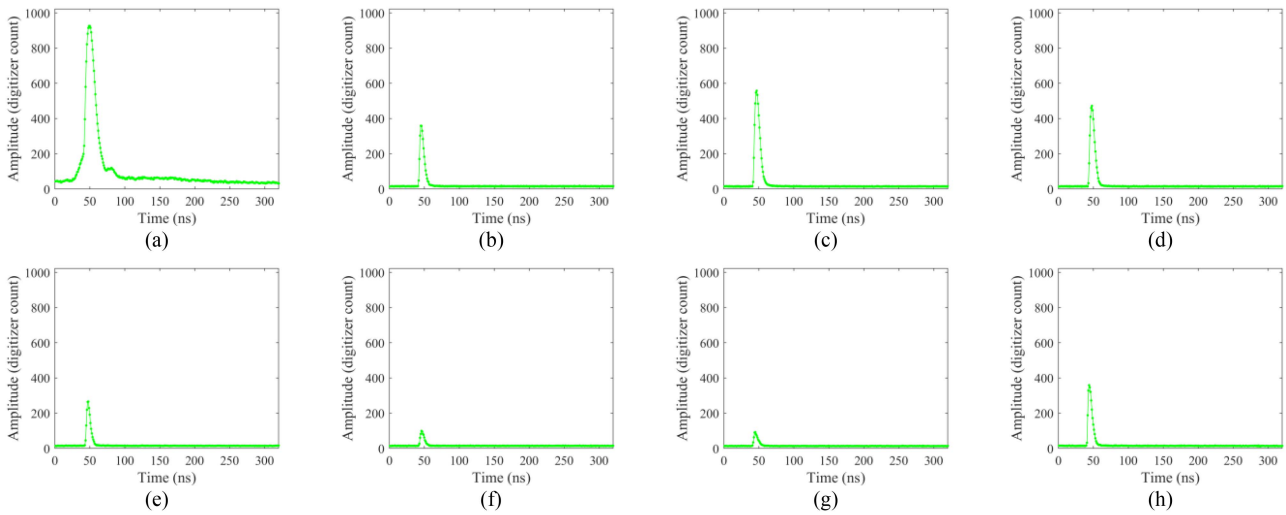

Fig. 2. Typical green laser waveforms of the ocean collected by (a) deep channel and (b)–(h) shallow 0–6 channels.

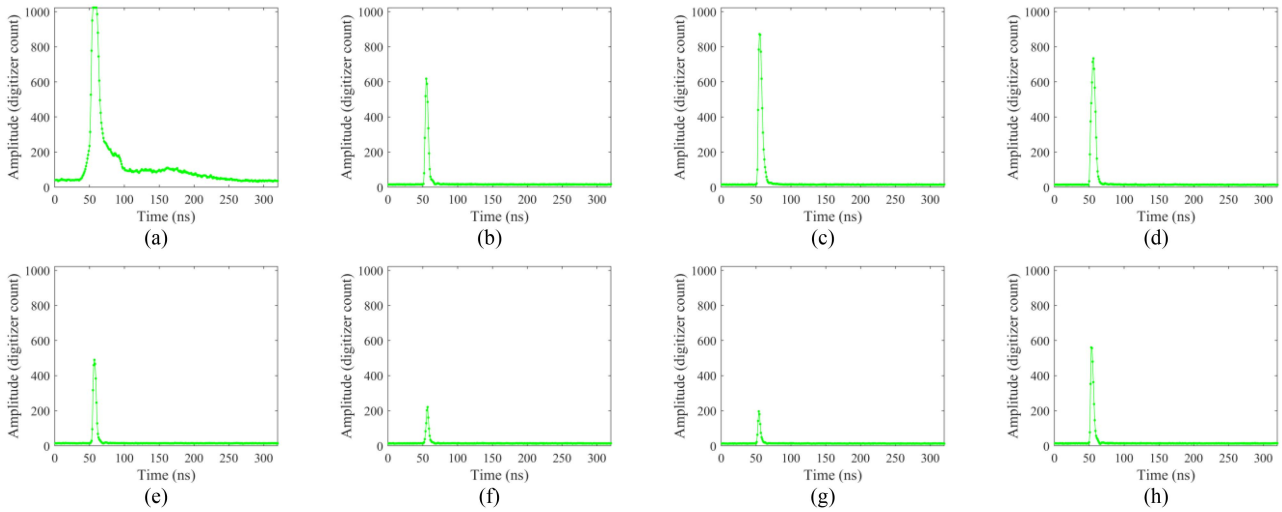

Fig. 3. Typical green laser waveforms of land collected by (a) deep channel and (b)–(h) shallow 0–6 channels.

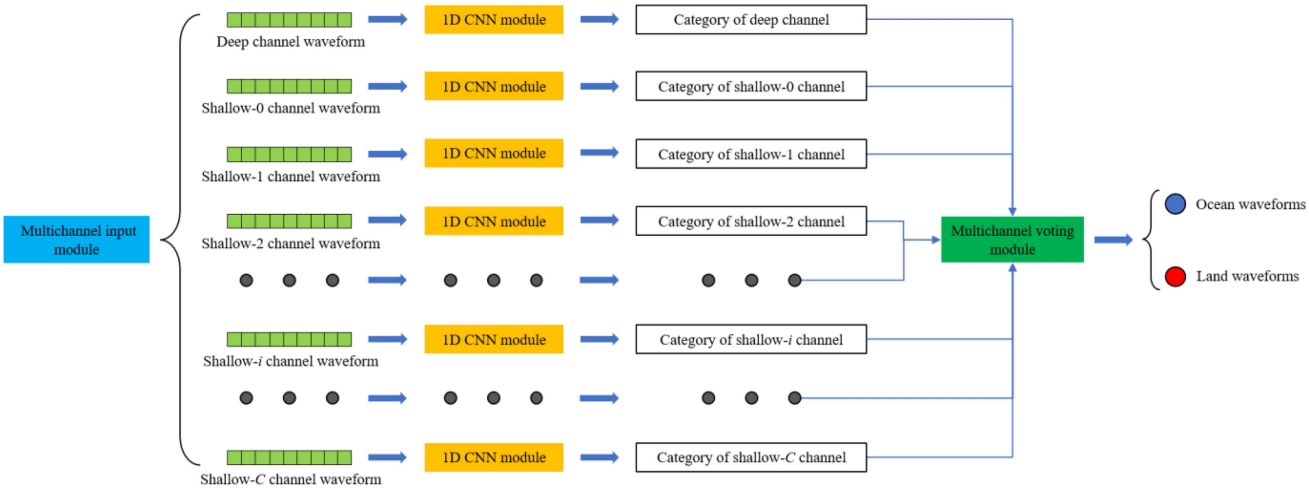

Fig. 4. Overview of the proposed model, MVCNN, for OLWC utilizing single-wavelength ALB systems.

and handcrafted features may not be optimal. As the size of the data increases, the disadvantages of ML gradually become significant. Deep learning (DL) can automatically extract features from numerous raw laser waveforms to maximize the classification accuracy without the need for prior knowledge, making it ideal for big ALB data processing. CNN is a well-known deep learning architecture characterized by a sequence of convolutional and pooling layers, followed by one or more fully connected layers. As shown in Fig. 4, our proposed MVCNN comprises one multichannel input module, multiple 1-D CNN modules, and one multichannel voting module.

B. Multichannel Input Module

A traditional green laser pulse from the laser emitter of the ALB system interacts with the environment to produce a return pulse that is detected by a photomultiplier tube or avalanche photodiode in the laser sensor. Then, the return pulse is digitized and recorded per nanosecond as a return laser full waveform. The green laser waveform can be expressed as

$$A = \{a_{t1}, a_{t2}, \ldots, a_{ti}, \ldots, a_{tn}\} \tag{1}$$

where a_{ti} is the amplitude (digitizer count) of the ith nanosecond of the green laser waveform A, and n is the waveform length. The green laser waveform is considered a 1-D time series that can be input into a 1-D CNN.

Multiple green laser channels are proposed to meet the challenge of seamless transitions between measurements in bathymetry and topography modes [16]. ALB systems have started to adopt multiple green laser channels by segmenting the FOV of receiver telescopes [14]. Fig. 5 shows a nonscaled illustration of the distributions of the multichannel FOVs of Optech CZMIL, with the X-axis representing the flight direction. Optech CZMIL presents one IR channel with 6 mrad, one green deep channel with 6–40 mrad FOV, and seven green shallow channels with 1.9 mrad FOV. The MVCNN can use multichannel green

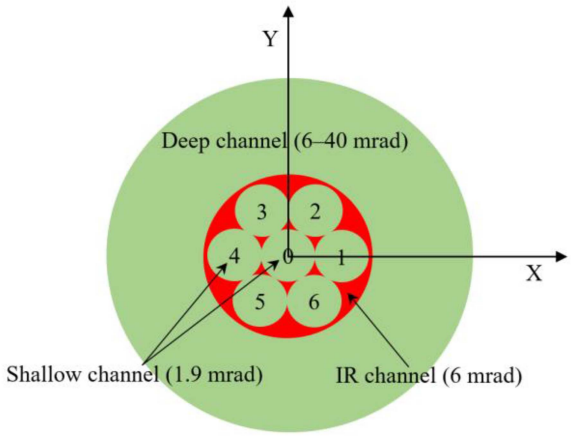

Fig. 5. Distributions of the multichannel FOVs of optech CZMIL (nonscaled).

laser waveforms by setting a multichannel input module in front of the 1-D CNN module for separately inputting multichannel green laser waveforms into the corresponding 1D CNN. The CZMIL system used in this article has eight green laser channels; thus, the number of channels of the multichannel input module is set to 8.

C. One-Dimensional CNN Module

Fig. 6 shows the structure of the 1-D CNN module proposed in this article. The 1-D CNN module mainly comprises four feature extraction modules, one attention mechanism (AM) module, and one classification module.

1) Feature Extraction Module: The 1-D convolutional [26] layer uses a 1-D filter of the preset width to locally convolve the input data in the order of step size and then outputs the corresponding features. The convolution is calculated as follows [27]:

$$z_{ik}^{l} = w_{k}^{l\mathrm{T}} x_{i}^{l} + b_{k}^{l} \tag{2}$$

where x_i^l is the input patch centered at location i of the lth layer, w_k^l and b_k^l are the weight matrix and bias term of the kth filter of the lth layer, respectively, and z_k^l is the kth feature map of the lth layer. The width of the convolution kernels of the four feature extraction modules is set to 3, and the number is set to 32, 32, 64, and 64. The number of convolution kernels per layer is equal to the number of feature maps in that convolutional layer due to the parameter sharing mechanism. Therefore, the shallow convolutional layer uses fewer convolution kernels to extract low-level features, and the deep convolutional layer uses more convolution kernels to extract enriched higher-level features to reduce the number of parameters and prevent overfitting.

Batch normalization (BN) can dramatically accelerate the training of deep networks and address the gradients that explode or vanish, and it also acts as a regularizer [28]. Therefore, a BN layer is set after the convolutional layer to accelerate the convergence and improve the robustness of the network. The nonlinear representation capability of the network is enhanced by setting a rectified linear unit [29], [30] layer after the BN layer.

Pooling transforms the joint feature representation into a new and more usable representation to preserve important information while discarding irrelevant details [31]. For CNNs, pooling reduces the input size of the next layer by subsampling the feature maps, thus lowering the computational burden and achieving shift invariance [27], [32]. Max pooling selects the largest element in each pooling region to better preserve the data texture features; thus, a max pooling layer is set after the second feature extraction module. Max pooling is calculated as follows [33]:

$$p_{ik} = \max_{(o,p)\in R_{ik}} \left(y_{koq}\right) \tag{3}$$

where y_{kop} is the element at (o, p) within the pooling region R_{ik}, which represents a local neighborhood around position i of the kth feature map, and p_{ik} is the related output of the pooling operator. The size and step size of the pooling kernel are set to 2.

Global max pooling [34] selects the largest element in each feature map to significantly reduce the parameters of the fully connected layers in the AM and classification modules, thereby saving computational resource and time while effectively preventing overfitting. Therefore, a global max pooling layer is set between the fourth feature extraction module and the AM module. The global max pooling is calculated as

$$g_k = \max\left(y_k\right) \tag{4}$$

where y_k is the activation value of the kth feature map and g_k is the related output of the pooling operation.

2) AM Module: The feature extraction module extracts various features from the green laser waveforms. However, differences exist in the features of green laser waveforms collected at different times and regions in complex ocean–land environments. Some features can be used to classify ocean–land waveforms, whereas others can interfere with the classification and reduce the generalization ability and robustness of the model. Therefore, an AM must be introduced to adaptively weight the feature maps enabling the model to focus more on features with important information while suppressing the invalid features. The AM module comprises a fully connected layer and a sigmoid layer. The fully connected layer takes all neurons in the previous layer and connects them to every single neuron of the current layer to generate global semantic information [27]. The number of neurons in the fully connected layer of the AM module is the same as the number of feature maps, which is set to 64. Let the activation value sequence of each feature map after the global max pooling operation be $G = [g_1, g_2, \ldots, g_K]^{\mathrm{T}}$, and let K be the number of feature maps. First, the feature sequence is input to the fully connected layer and activated by sigmoid to obtain the attention weight S of each feature map. This scheme can be expressed as

$$S = \delta\left(WG + b\right) = \left[s_1, s_2, \ldots, s_K\right]^{\mathrm{T}} \tag{5}$$

where W and b are the weight matrix and bias vector of the fully connected layer, respectively, and $\delta(\cdot)$ is the sigmoid function.

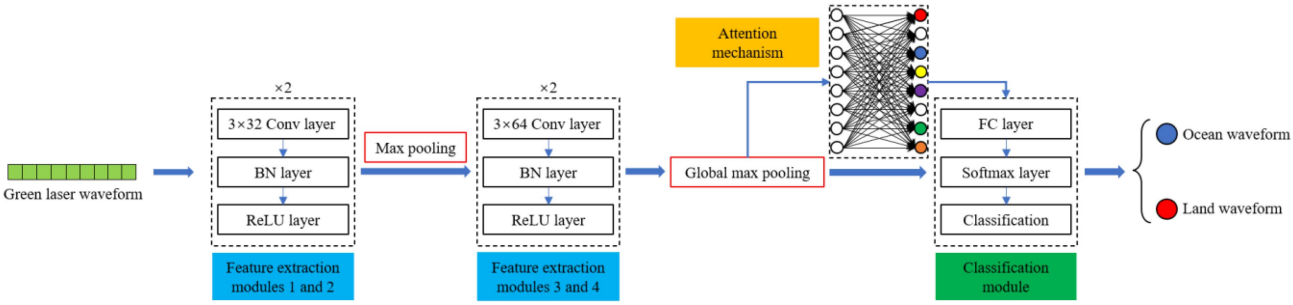

Fig. 6. Structure of the proposed 1-D CNN module.

The sigmoid function is expressed as follows [28]:

$$\delta(x) = \frac{1}{1+e^{-x}} \tag{6}$$

where x is the output of the fully connected layer. After sigmoid activation, the attention weights take values distributed between $(0, 1)$. Then, the attention weights S and the primitive feature sequence G are operated as element products to obtain the feature sequence G' after the AM screening, which is expressed as

$$G' = SG. \tag{7}$$

3) Classification Module: The feature extraction module performs deep feature extraction of the green laser waveforms and adaptively weights the obtained features by using the AM module to filter out the useful classification feature information. The feature sequences filtered by AM are input to the classification module for the classification task. The classification module comprises a fully connected layer and a softmax layer. In the classification module, the number of fully connected layer neurons is the same as the number of waveform classes, which is set to 2. The last layer of the CNN is an output layer. For the classification task of CNN, the softmax operator is commonly used as the last layer [35]. The softmax function is defined as follows [36]:

$$y_i = \frac{e^{x_i}}{\sum_{j=1}^{2} e^{x_j}} \tag{8}$$

where x_i is the output of the ith neuron of the fully connected layer, $i = \{1, 2\}$, and y_i is the activation value (i.e., predicted scores) corresponding to the probability of the input data of the CNN affiliated with category i.

Training a CNN is a global optimization problem. The optimum parameters (e.g., weight matrices and bias vectors) can be determined by minimizing the loss function. The CNN model in this article is trained using the following binary cross-entropy loss function [37]:

$$J = -\frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{2} t_{mi}\log(y_{mi}) \tag{9}$$

where M is the number of training samples, and t_{mi} is the indicator that the mth sample belongs to the ith category. t_{mi} is 1 if the sample belongs to the ith category; otherwise, t_{mi} is 0. y_{mi} is the output for the training sample m for category i of the softmax function, and J is the value of the loss function.

D. Multichannel Voting Module

Green laser waveforms of very shallow waters misclassified as land waveforms and green laser waveforms of special land topography misclassified as ocean waveforms both reduce the overall accuracy of OLWC for single-wavelength ALB systems [14]. The complex environment distinctly affects the OLWC of different channel waveforms. The responses of lasers in different channels to the environment are different. The characteristics of green laser return waveforms collected by multiple channels with different FOVs for the same laser shot are dissimilar. OLWC utilizing green laser waveforms collected in the deep channel with a large FOV is highly influenced by the merging effect [14]. Green laser waveforms collected in shallow channels with a small FOV can classify more very shallow waters than those collected in deep channels. However, green shallow channel waveforms are largely influenced by noise because of the small FOVs, which reduce the overall accuracy of the OLWC utilizing green shallow channel waveforms [14]. For the green laser waveforms of land with multiple returns, the laser energy is distributed to multiple returns, thereby reducing the amplitude of a single return and further resulting in the misclassification of such special waveforms. Therefore, the overall accuracy is low when green laser waveforms collected in deep channels or in shallow channels are used separately. Thus, the multichannel voting module is proposed to take advantage of the green laser waveforms collected in deep and shallow channels with different responses to the environment, hence reducing the influence of special waveforms on OLWC.

The multichannel voting module votes on the predicted categories of each channel 1-D CNN module according to the majority voting rules and selects the category with more than half of the votes as the final category φ. The number of ocean and land category votes may be equal when the number of channels is even. In this case, the predicted category of the deep channel is the final category because the deep channel waveform performs better than the shallow channel waveform for OLWC [14]. This process can be described as follows:

$$\varphi = \begin{cases} \text{ocean waveform} & O > L \\ \text{land waveform} & O < L \\ \text{category of deep channel} & O = L \end{cases} \tag{10}$$

where O and L are multiple 1-D CNN module vote numbers for the waveform of the laser shot belonging to ocean and

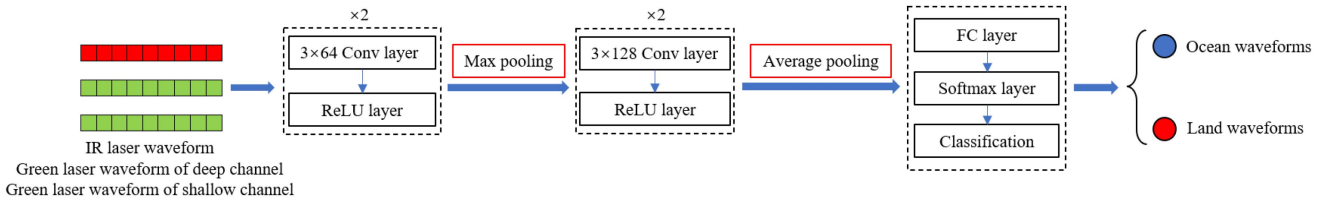

Fig. 7. Structure of Hu's 1-D CNN.

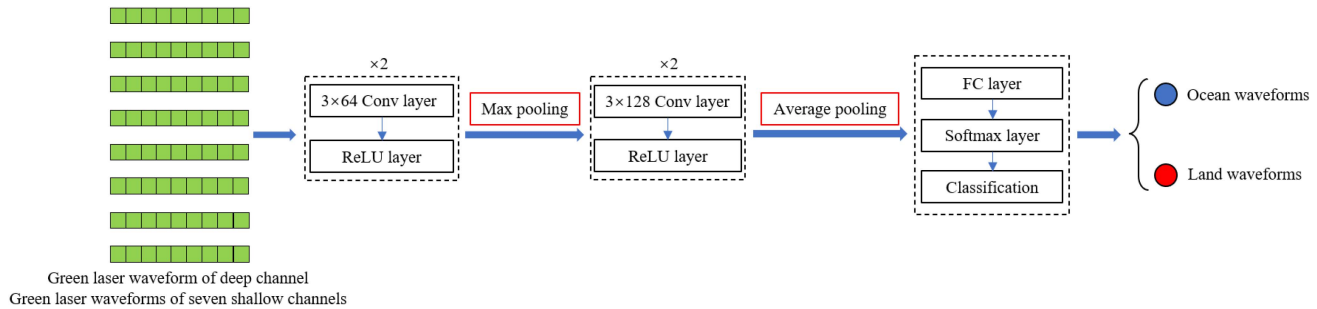

Fig. 8. Structure of the modified Hu's 1-D CNN.

land categories, respectively. The multichannel voting module ensures that the classification results of each laser shot are jointly decided by the waveforms of multiple channels, thus avoiding extreme instances and improving the accuracy and robustness of the OLWC.

E. Comparison Methods

1) ML Methods: The OLWC of the single-wavelength ALB is commonly performed by traditional ML methods, including unsupervised clustering algorithms, such as FCM, or supervised classification algorithms, such as SVM. The shape of the green laser waveform of water differs from that of land. Thus, ocean–land waveforms can be classified using green laser waveform amplitudes through FCM [14]. Other features, such as the FWHM and area of the green laser waveform, can be used to train the SVM for OLWC [13].

2) DL Methods: Hu et al. [20] constructed a 1-D CNN (hereafter referred to as Hu's 1-D CNN) for OLWC utilizing dual-wavelength ALB systems. Fig. 7 shows the structure of Hu's 1-D CNN. This 1-D CNN includes four convolutional layers, two pooling layers, and one fully connected layer. Unlike MVCNN, this method inputs three channels of waveform data into CNN simultaneously for OLWC and has only one end-to-end 1-D CNN model. Our proposed method inputs each channel waveform data into one CNN individually, performs OLWC utilizing the multichannel voting module, and has multiple end-to-end 1-D CNN models. Hu's 1-D CNN is oriented to dual-wavelength ALB systems and cannot be directly applied to single-wavelength ALB systems. Thus, as shown in Fig. 8, Hu's 1-D CNN is modified (hereafter referred to as modified Hu's 1-D CNN) by removing the IR laser waveform from the input layer and increasing the number of green laser channels to make it suitable for single-wavelength ALB systems.

F. Evaluation Methods

1) Reference Method: The dual-clustering OLWC utilizing IR laser waveforms is robust and highly accurate [9]. The dual-clustering method has two levels. The first level removes outliers in the IR laser waveform amplitudes via K-means clustering, while the second level removes geographic outliers by using DBSCAN to correct the misclassified IR laser waveforms derived by the first level [9]. This method is used as an external reference method to evaluate different OLWC methods utilizing green laser waveforms.

2) Evaluation Metrics: Precision (P), recall (R), F_1 score (F_1), overall accuracy (OA), and kappa coefficient (Kappa) are calculated using the reference and prediction categories as a way of quantifying the classification performance of different models.

The DL algorithm has a certain randomness, such as the random initialization of model parameters. If the model is trained several times using the same training data, then the model parameters obtained from the model training are different, resulting in the different accuracies of the model for the test data. The standard deviation of OA (SDOA) of the DL models obtained from each training for the test data is proposed to evaluate the robustness of the model. SDOA is calculated as

$$\mathrm{SDOA} = \sqrt{\frac{\sum_{i=1}^{N}\left(\mathrm{OA}_i - \frac{1}{N}\sum_{j=1}^{N}\mathrm{OA}_j\right)^2}{N-1}} \tag{11}$$

where N is the number of training and OA_i is the overall accuracy of the ith training model. The smaller the SDOA, the more robust the model.

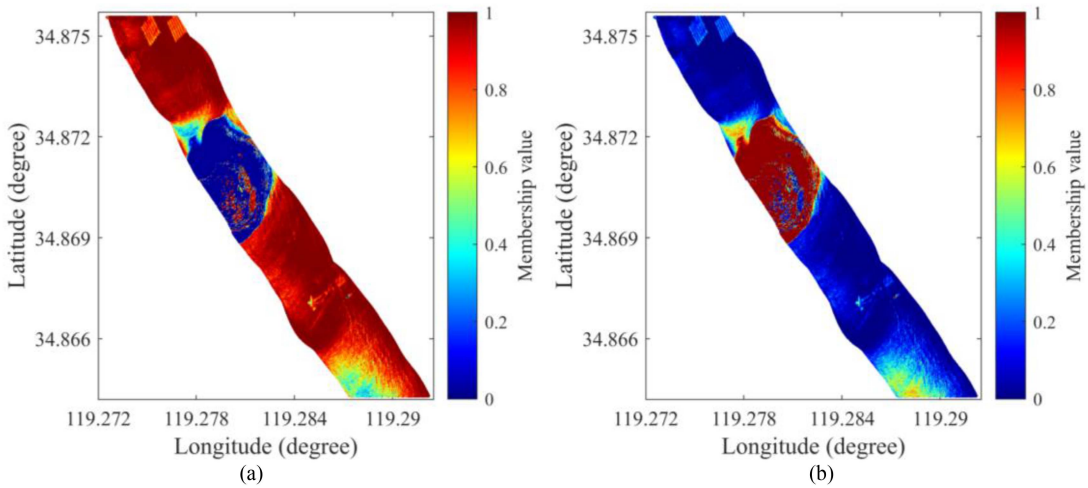

Fig. 9. Membership values belonging to (a) ocean and (b) land categories derived by the FCM.

III. Experimental Results and Analysis

A. Experiments of ML Methods

1) FCM Experiment: The waveform amplitude is extracted from the test set of the deep channel waveforms to conduct OLWC utilizing FCM. Fig. 9(a) and (b) shows the membership values of each laser shot belonging to the ocean and land categories generated by the FCM, respectively. The cluster centroids for ocean and land waveform amplitudes are 925.57 and 1008.50 digitizer counts, respectively. Fig. 10 presents the laser point positions (i.e., longitude and latitude of the laser irradiation points) of the corresponding ocean (blue) and land (red) waveforms of the test set classified by the FCM. The yellow curve represents the waterline obtained by the reference method. A typical misclassified waveform collected by the deep channel is attached to the right of each position figure. As shown in the figure, the special green laser waveform of land is wider than the typical waveform of land and has multiple peaks, thus reducing the amplitude of each peak, further leading to misclassification. The misclassified waveforms of the ocean are mainly induced by the merging effect. In very shallow waters, the water surface, volume backscatter, and water bottom returns of the green laser waveform merge into a single return, causing a large amplitude of the ocean's green laser waveform and a misclassification [14].

2) SVM Experiment: The amplitude, FWHM, and area of the deep channel raw waveforms are calculated. Then, the features of the training set of the deep channel waveforms are input into the SVM for training. Finally, the trained SVM is used for the OLWC test set. Fig. 11 shows the laser point positions of the corresponding ocean (blue) and land (red) waveforms of the test set classified by the SVM. The number of misclassified waveforms obtained by the SVM is less compared with that obtained by the FCM. The SVM can correctly classify many land waveforms affected by topography, but it cannot correctly classify ocean waveforms affected by the merging effect.

B. Experiments of DL Methods

1) Network Model Training Environment and Parameters: Network models are built, trained, and tested using the MATLAB deep learning framework installed on a Windows 10 operating system with an Intel Core i5-7300HQ, 32 GB random access memory (RAM), and an NVIDIA GeForce GTX 1050 Ti (4 GB). The training parameters include a learning rate of 0.001, a batch size of 1024, several batches of 891 (three epochs), a gradient descent optimizer of Adam, and the training set shuffled before each epoch round.

2) Modified Hu's 1-D CNN Experiment: The training sets of eight-channel raw green laser waveforms are simultaneously input into the modified Hu's 1-D CNN. Fig. 12 shows the trends of accuracy and loss function values of the model during the training process. The model converges after 891 batches, and the accuracy and loss are close to 100% and 0, respectively. The modified Hu's 1-D CNN model is trained ten times under the same experimental conditions, and the ten groups of modified Hu's 1-D CNN model are obtained and used to classify the validation and test sets. Fig. 13 exhibits the laser point positions of the corresponding ocean (blue) and land (red) waveforms of the test set classified by the modified Hu's 1-D CNN, with the OA closest to the averaged OAs of the ten groups. The modified Hu's 1-D CNN can correctly classify most of the land waveforms affected by topography. However, the modified Hu's 1-D CNN misclassifies many ocean waveforms because of the presence of aquaculture rafts and the merging effect in very shallow waters. The merging effect causes the green laser waveforms of very shallow waters to be similar to those of land [14]. Multiple peaks tend to appear when green laser irradiates the rafts and water, resulting in similarity with land waveforms with multiple returns, further leading to misclassification [9].

3) MVCNN Experiment: The multichannel input module is used to input the training set of the eight-channel raw green laser waveforms to the eight 1-D CNNs for training. Fig. 14 shows the training process of the proposed 1-D CNN. The accuracy

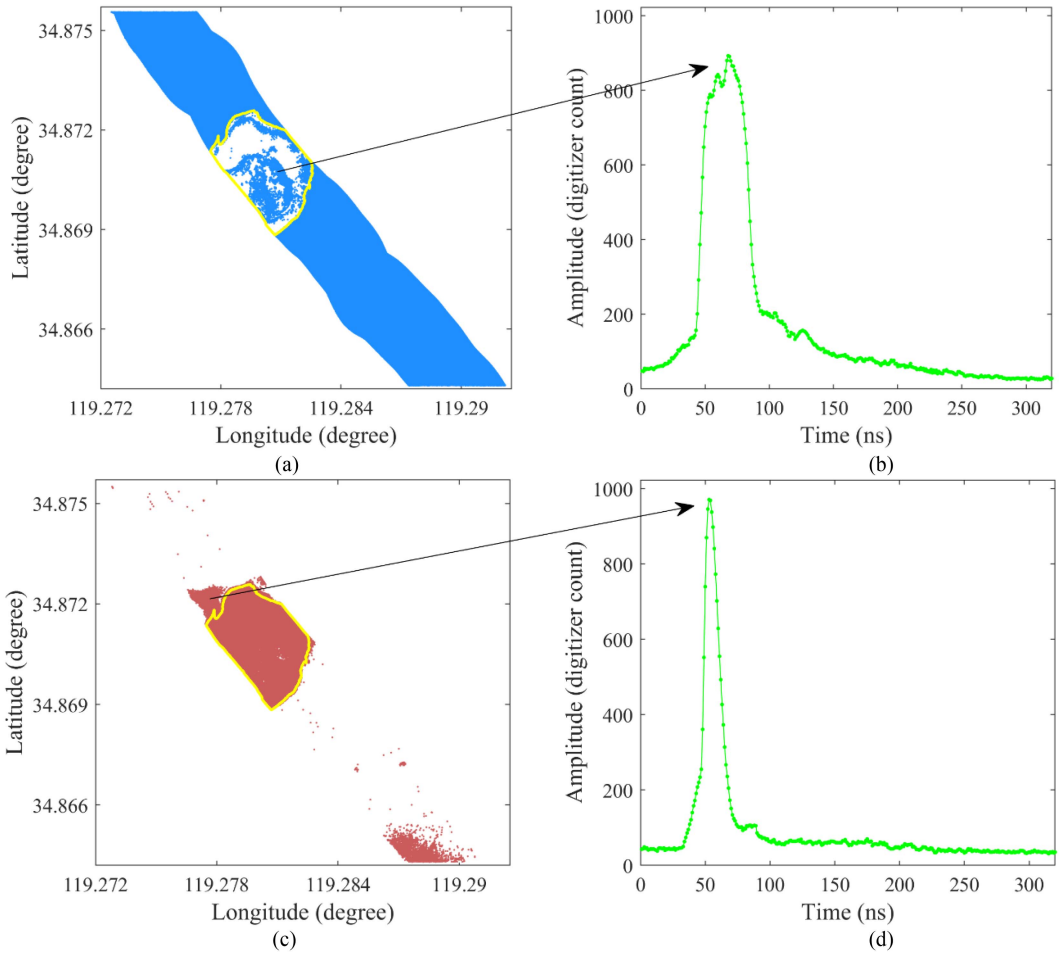

Fig. 10. Classification results obtained from the FCM for the test data. Positions of laser irradiation points of the corresponding (a) ocean and (c) land waveforms classified by the FCM, (b) a typical land waveform of the deep channel misclassified as an ocean waveform, and (d) a typical ocean waveform of the deep channel misclassified as a land waveform.

and loss of the model converge after 891 batches. Then, the categories of the validation and test sets are obtained using the multichannel voting module. The MVCNN is trained ten times under the same experimental conditions, and the ten groups of MVCNNs are obtained and used to classify the validation and test sets. Fig. 15 shows the laser irradiation point positions of the corresponding ocean (blue) and land (red) waveforms of the test set classified by the MVCNN, with the OA closest to the averaged OAs of the ten groups. The MVCNN has the highest boundary consistency with the reference waterline and the fewest misclassified waveforms compared with the other three methods.

TABLE I
ACCURACY EVALUATION RESULTS OF THE FCM, SVM, MODIFIED HU'S 1-D CNN, AND MVCNN

Metrics		FCM	SVM	Modified Hu's 1D CNN	MVCNN
Ocean	P (%)	96.97	98.53	**99.71**	99.58
	R (%)	97.61	99.33	98.52	**99.70**
	F_1 (%)	97.29	98.92	99.11	**99.64**
Land	P (%)	88.90	96.83	93.77	**98.62**
	R (%)	86.26	93.28	**98.71**	98.11
	F_1 (%)	87.56	95.03	96.15	**98.36**
OA (%)		95.55	98.23	98.56	**99.41**
Kappa (×100)		84.85	93.95	95.27	**98.00**
SDOA (%)		—	—	0.51	**0.03**

C. Results of the Accuracy Evaluation

The test set is not involved in the training adjustment of an OLWC model, and the generalization ability of the model can be tested. Table I shows the accuracy evaluation results of the FCM, SVM, modified Hu's 1-D CNN, and MVCNN on the test set, in which the optimal metrics are bolded, and all metrics for the CNN models are averaged except for SDOA. Fig. 16 shows the confusion matrix of all methods. The number of green laser waveforms of the test set used by SVM is less than that of the other three methods because the FWHM features cannot be extracted from 202 laser waveforms. Table I and Fig. 16 demonstrate that the MVCNN has obvious advantages over the other methods in terms of accuracy and robustness, as shown by the better F_1, OA, Kappa, and SDOA than those of the other methods.

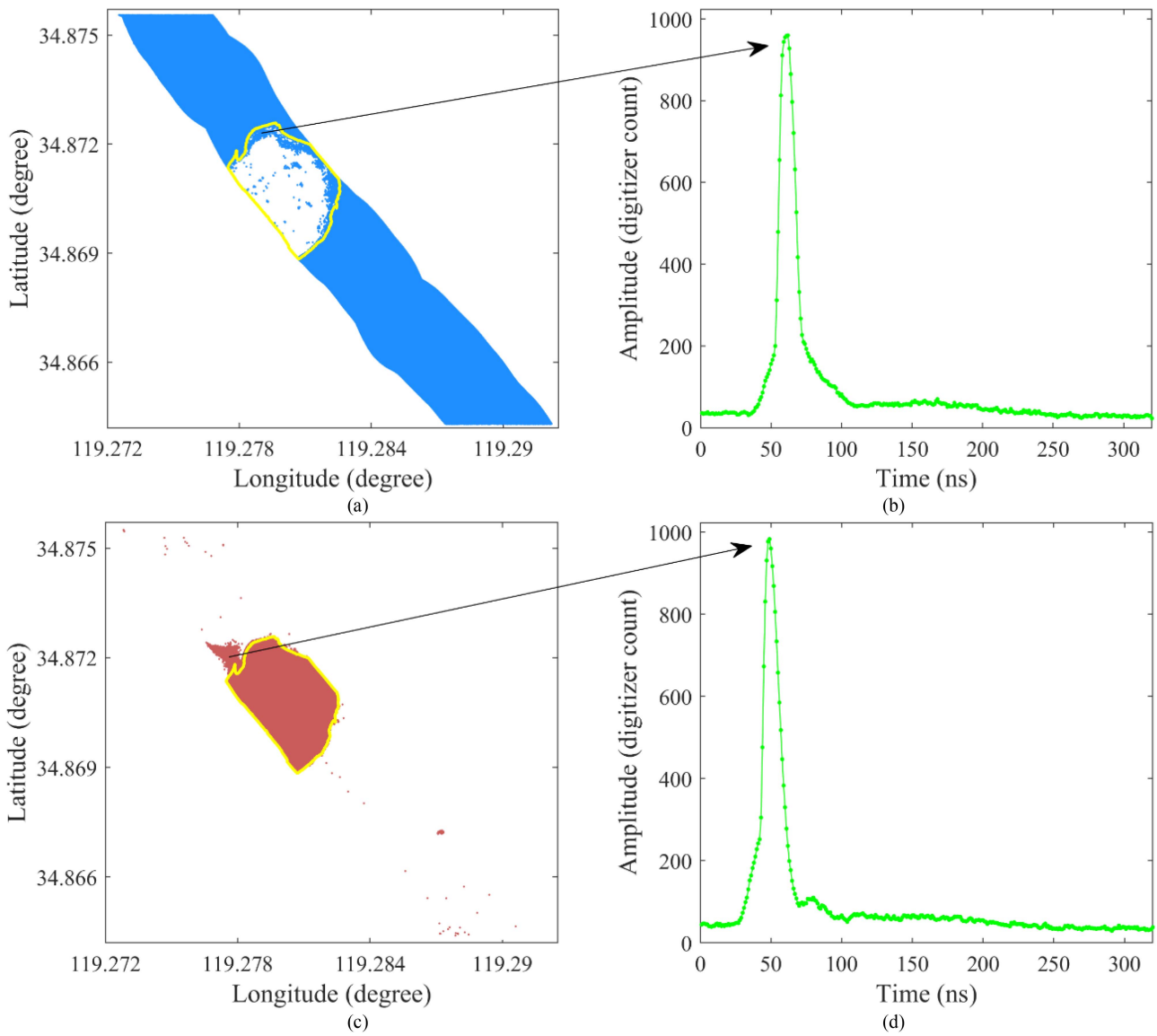

Fig. 11. Classification results obtained from the SVM for the test data. Positions of laser irradiation points of the corresponding (a) ocean and (c) land waveforms classified by the SVM, (b) a typical land waveform of the deep channel misclassified as an ocean waveform, and (d) a typical ocean waveform of the deep channel misclassified as a land waveform.

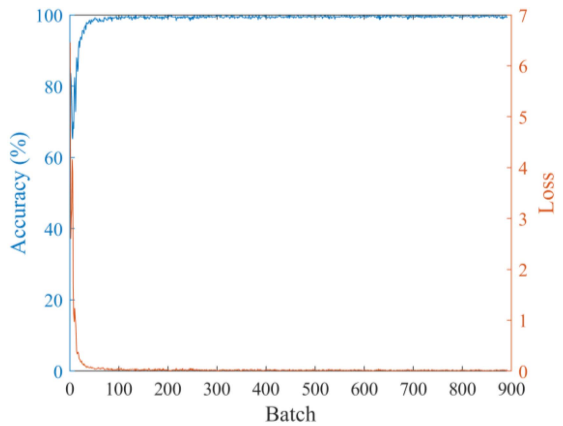

Fig. 12. Trends of accuracy and loss in training the modified Hu's 1-D CNN.

Figs. 10, 11, 13, and 15 show the inability of the manually extracted features to accurately classify the ocean waveforms affected by the merging effect and the land waveforms affected by topography. By contrast, the features automatically extracted by the DL methods can well identify the special land waveforms from ocean waveforms. As shown in Fig. 13(c), many ocean waveforms are misclassified as land waveforms by the modified Hu's 1-D CNN, taking the waterline obtained by dual-clustering of IR laser waveforms as the reference. These misclassified ocean waveforms are concentrated in the areas of very shallow waters and aquaculture rafts, indicating that the modified Hu's 1-D CNN is significantly affected by the aquaculture rafts and the merging effect of the green laser waveform in very shallow waters. As shown in Fig. 15(c), the MVCNN is slightly affected by the merging effect and the aquaculture rafts.

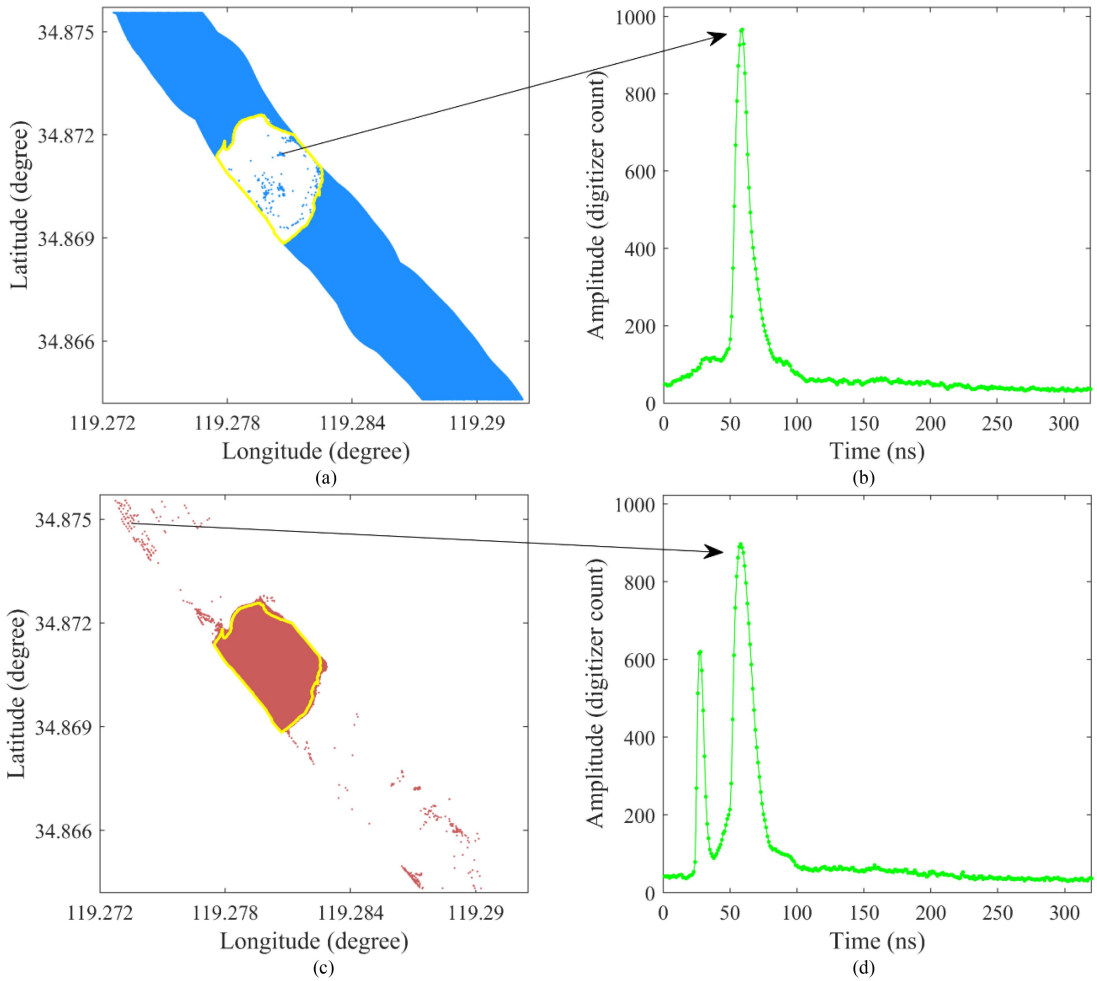

Fig. 13. Classification results obtained from the modified Hu's 1-D CNN for the test data. Positions of laser irradiation points of the corresponding (a) ocean and (c) land waveforms classified by the modified Hu's 1-D CNN, (b) a typical land waveform of deep channel misclassified as an ocean waveform, and (d) a typical ocean waveform of deep channel misclassified as a land waveform.

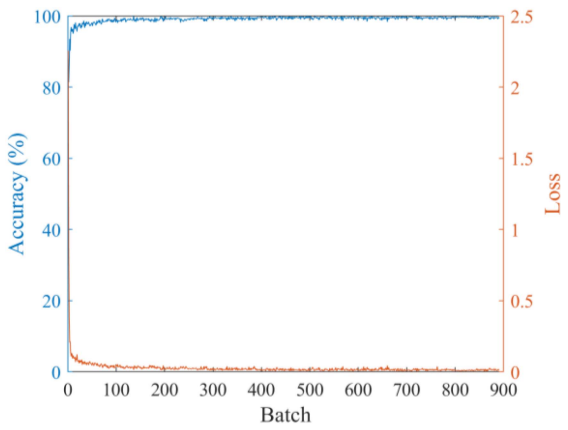

Fig. 14. Trends of accuracy and loss in training the proposed 1-D CNN.

The box plot can reflect the robustness of the OLWC model. The narrower the box, the smaller the OA floating range and the more robust the model. Fig. 17 shows the box plot of the modified Hu's 1-D CNN (blue) and the MVCNN (red). The classification accuracy of each model on the validation set (Groups 1 and 2) is high and fluctuates slightly, indicating that each model fits well on the validation set. On the test set (Groups 3 and 4), the MVCNN remains robust, but the modified Hu's 1-D CNN fluctuates immensely. Although the accuracy of the validation set meets expectations, the result of the test set has great uncertainty, indicating the weak generalization ability of the model. By contrast, the MVCNN has strong generalization ability and robustness, which is important for model design, training, and practical applications.

D. Comprehensive Comparison of OLWC Methods for Practical Application

This section provides a comprehensive comparison of the FCM, SVM, modified Hu's 1-D CNN, and MVCNN for practical applications. The results are shown in Table II. FCM does not need sample data, and its classification accuracy is the lowest. The feature extraction in the DL methods is automatic, whereas that of the FCM and SVM methods is manual. MVCNN does

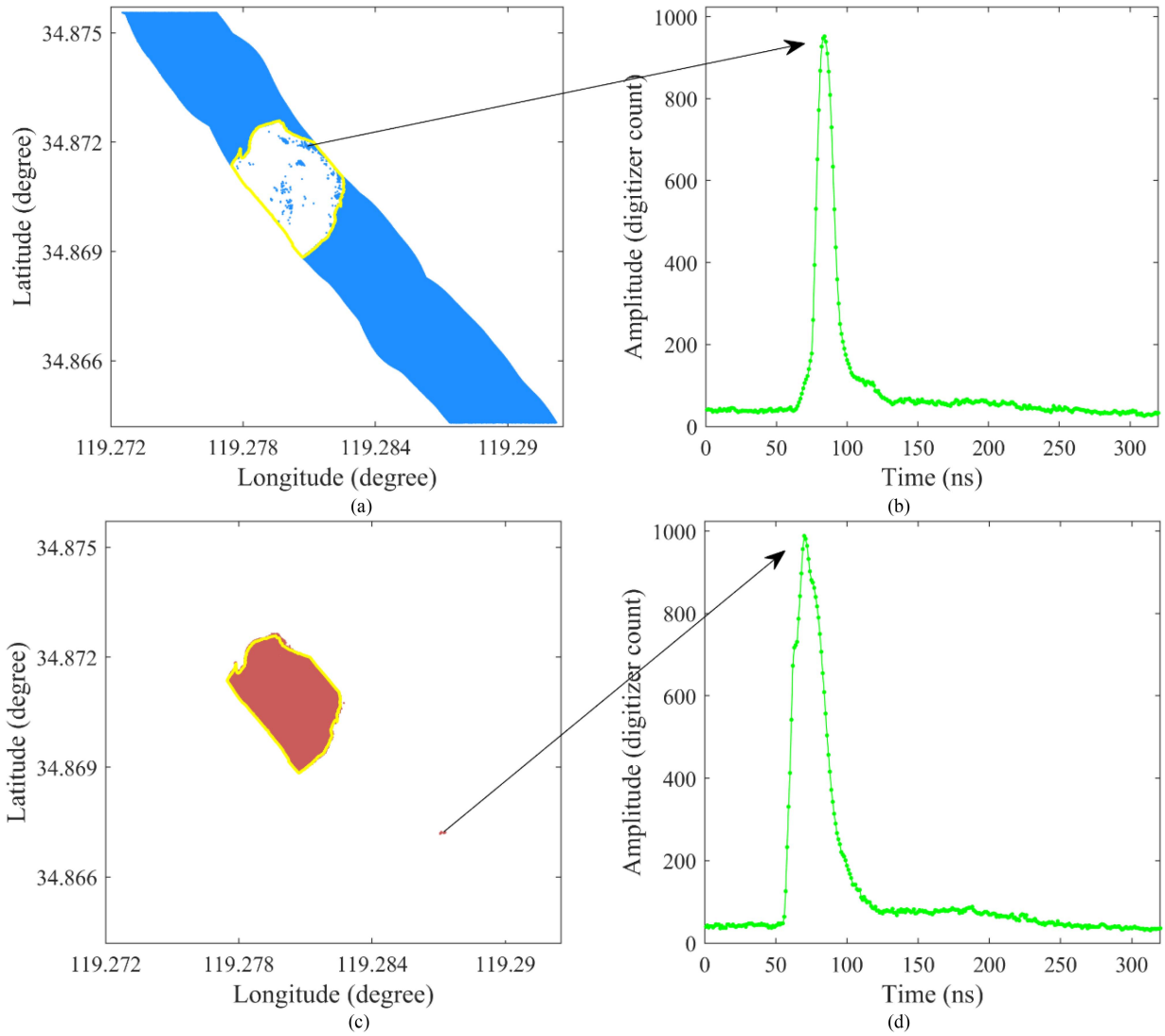

Fig. 15. Classification results obtained from the MVCNN for the test data. Positions of laser irradiation points of the corresponding (a) ocean and (c) land waveforms classified by the MVCNN, (b) a typical land waveform of the deep channel misclassified as an ocean waveform, and (d) a typical ocean waveform of the deep channel misclassified as a land waveform.

not need to manually extract waveform features and is easy to use. MVCNN can achieve an OA of 99.41% in the experimental area and has the highest OA among the existing OLWC methods for single-wavelength ALB systems. In the experimental area, MVCNN can correctly identify 2500–11366 more ocean/land laser waveforms than the other methods, which is beneficial for subsequent data processing, such as depth extraction from ocean waveforms and waterline identification by using ocean/land laser points.

The RAMs required for the normal operation of the methods during training and classification and the time required for classifying test sets by using the model are listed in Table II. The RAM consumption of FCM, which requires 0.7 GB, mainly lies in the feature extraction process. SVM requires 1 GB RAM to train the model. CNN consumes minimal RAM for training and classification but requires a sizable RAM to convert the raw data into a specific format. The input of the modified Hu's 1-D CNN for training takes the form of big data (320 bins $\times$ 304332 laser pulses $\times$ eight channels), which consume 6.3 GB of RAM. The input of each 1-D CNN of MVCNN takes the form of small data (320 bins $\times$ 304332 laser pulses $\times$ one channel), which require only 1.4 GB of RAM. The modified Hu's 1-D CNN must simultaneously process eight-channel waveforms and therefore requires a large amount of RAM. Conducting modified Hu's 1-D CNN method on a large test set with large pulse numbers is difficult due to the hardware constraints. MVCNN can process single-channel waveforms and, therefore, requires relatively lower RAM (roughly 22% of modified Hu's 1-D CNN) and can be applied to a large test set. The classification time of the ML methods is less than that of the DL methods. Although MVCNN requires the maximum time among the four methods, it can classify approximately 3466 pulse waveforms per second, which can meet the efficiency requirements of OLWC.

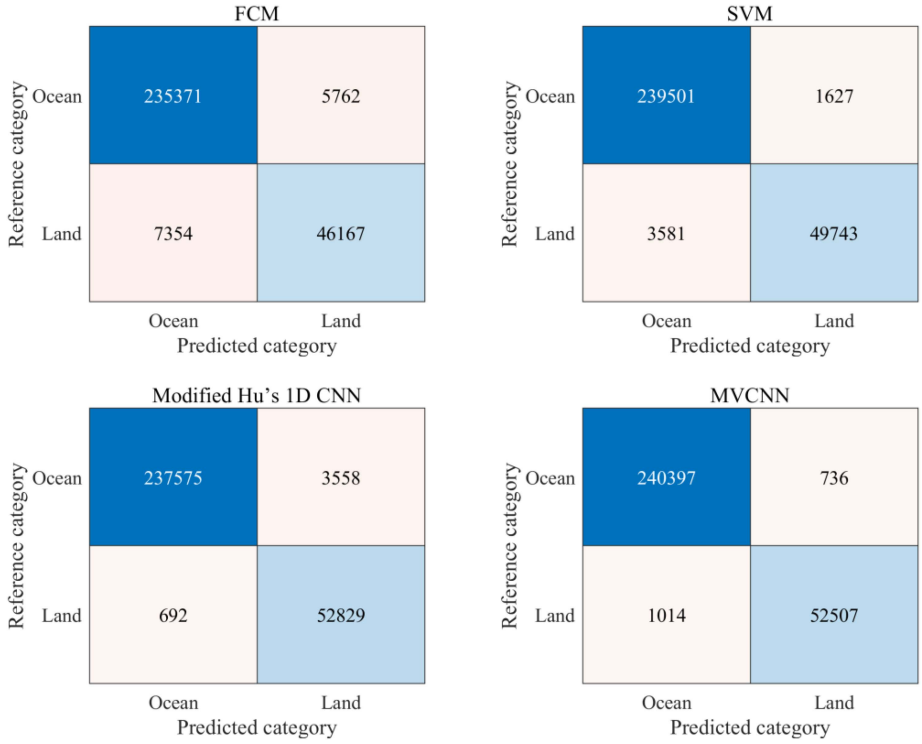

Fig. 16. Confusion matrix of the test data based on the prediction of each method.

TABLE II
COMPREHENSIVE COMPARISON OF OLWC USING FCM, SVM, MODIFIED HU'S 1-D CNN, AND MVCNN FOR PRACTICAL APPLICATION

Method	Sample data	Features	Key parameters	OA	RAM	Time	Reference
FCM	Not required	Waveforms amplitudes of green deep channel	Fuzzy index and number of cluster centroids are set to 2.	95.55%	0.7 GB	2 s	Zhao et al, 2021
SVM	Required	Waveform amplitudes, FWHMs, and areas of green deep channel	Gaussian kernel function, kernel parameter σ is set to 1.7, and penalty coefficient is set to 1.	98.23%	1.0 GB	13 s	Zhao et al, 2019
Modified Hu's 1D CNN	Required	Automatically extracted features	One 1D CNN with four convolutional layers. The width and number of convolution kernels of the first and second convolutional layers are 3 and 64, respectively. The width and number of convolution kernels of the third and fourth convolutional layers are 3 and 128, respectively.	98.56%	6.3 GB	21 s	Hu et al, 2019
MVCNN	Required	Automatically extracted features	One multichannel input module, multiple 1D CNNs, and one multichannel voting module. Each 1D CNN has four convolutional layers and one AM module. The width and number of convolution kernels of the first and second convolutional layers are 3 and 32, respectively. The width and number of convolution kernels of the third and fourth convolutional layers are 3 and 64, respectively.	99.41%	1.4 GB	85 s	—

IV. DISCUSSION

The advantages of using MVCNN for green laser OLWC have been demonstrated in Section III. In this section, the effectiveness of the three submodules of MVCNN, i.e., multichannel input module, 1-D CNN module, and multichannel voting module, are separately discussed to demonstrate their contributions to MVCNN. In addition, the relationships of different channel waveforms are analyzed, and the reasons for the amplitude difference between the different channels are given.

A. Multichannel Input Module Evaluation

The proposed 1-D CNN trained by the raw laser waveforms of different channels are used to classify the test set. Then,

TABLE III
PERFORMANCE METRICS OF THE PROPOSED 1-D CNN FOR THE TEST SET COLLECTED BY EIGHT CHANNELS

Metrics		Deep	Shallow-0	Shallow-1	Shallow-2	Shallow-3	Shallow-4	Shallow-5	Shallow-6
Ocean	P (%)	**99.25**	96.79	98.75	98.64	98.85	97.88	97.80	98.45
	R (%)	99.60	99.70	**99.75**	99.72	99.58	99.48	99.48	99.65
	F_1 (%)	**99.43**	98.23	99.25	99.18	99.21	98.67	98.63	99.05
Land	P (%)	98.19	98.46	**98.85**	98.70	98.05	97.50	97.46	98.34
	R (%)	**96.60**	85.12	94.31	93.80	94.77	90.24	89.89	92.94
	F_1 (%)	**97.39**	91.30	96.52	96.17	96.37	93.67	93.50	95.56
OA (%)		**99.06**	97.05	98.77	98.65	98.71	97.80	97.74	98.43
Kappa (×100)		**96.81**	89.53	95.77	95.35	95.58	92.35	92.13	94.61
SDOA (%)		**0.14**	0.20	0.16	0.39	0.29	0.55	0.34	0.22

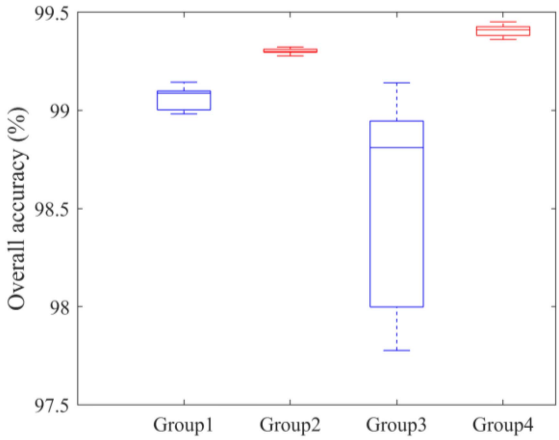

Fig. 17. Box plot of the modified Hu's 1-D CNN (blue) and the MVCNN (red). (Groups 1 and 2 are the validation set, and groups 3 and 4 are the test set).

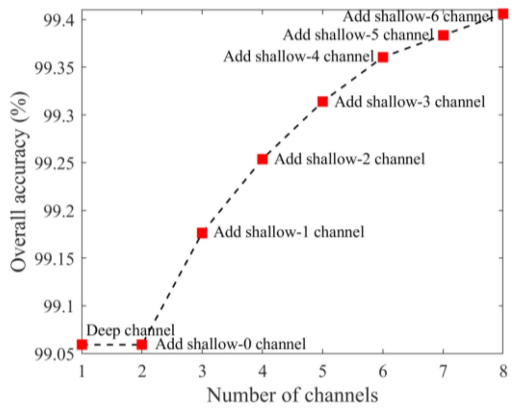

Fig. 18. Trends of the overall accuracy with the increase in the number of channels.

the performance metrics of each channel model are derived to analyze the effect of the multichannel input module for the MVCNN and the OLWC ability of different channel waveforms. As shown in Table III, the OLWC ability of the green laser waveforms collected in the different green laser channels is dissimilar. The model trained by the deep channel waveforms has higher accuracy and better robustness than that trained by the shallow channel waveforms. The advantage of using the deep channel waveform is especially obvious in the performance of classifying land waveforms. In addition, the OLWC ability of green laser waveforms collected in the different shallow channels is dissimilar. The central shallow channel (shallow-0 channel) is significantly poorer than the other shallow channel, especially for classifying land waveforms. Subsequently, deep channel and shallow 0–6 channel waveforms are added gradually to the MVCNN, and the results are shown in Fig. 18. The classification accuracy is improved gradually with the increase in the number of channels, indicating that the multichannel input module plays a crucial role in the accuracy improvement of the MVCNN.

TABLE IV
PERFORMANCE METRICS OF THE 1-D CNN STRUCTURE OF MODIFIED HU'S 1-D CNN AND MVCNN

Metrics		Modified Hu's 1D CNN	MVCNN
Ocean	P (%)	**99.49**	99.25
	R (%)	98.82	**99.60**
	F_1 (%)	99.15	**99.43**
Land	P (%)	94.93	**98.19**
	R (%)	**97.71**	96.60
	F_1 (%)	96.26	**97.39**
OA (%)		98.62	**99.06**
Kappa (×100)		95.42	**96.81**
SDOA (%)		0.45	**0.14**

B. *One-Dimensional CNN Module Evaluation*

The MVCNN with only deep channel waveforms and the modified Hu's 1-D CNN with only deep channel waveforms are used for OLWC to compare the performance of the different 1-D CNN structures. The results are shown in Table IV. The 1-D CNN structure of MVCNN is better than that of the modified Hu's 1-D CNN. Moreover, the OLWC accuracy improvement by using MVCNN can be attributed to not only the multichannel input and voting modules but also the improvement of the 1-D CNN structure.

Then, the classification capability of the proposed 1-D CNN is further investigated by performing an ablation experiment and feature visualization.

1) AM Module Evaluation: Aiming to investigate whether the AM can improve the classification ability of the proposed 1-D CNN, an ablation experiment is designed to compare and analyze the OLWC ability of the CNN model with and without the AM for deep channel waveforms. Table V shows the performance metrics of the 1-D CNN without and with AM for the test set. The model with AM has higher OLWC accuracy

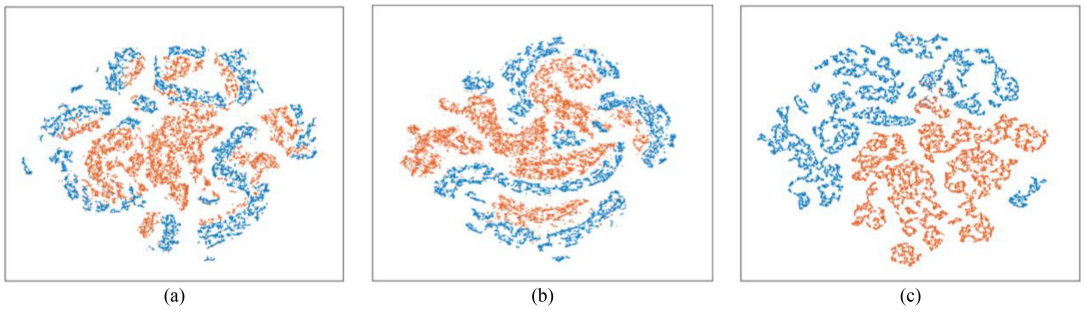

Fig. 19. Feature visualization using the proposed network. (a) Raw green laser waveforms. (b) Fourth convolutional layer. (c) Final fully connected layer. The blue and orange colors are waveform features representing ocean and land, respectively.

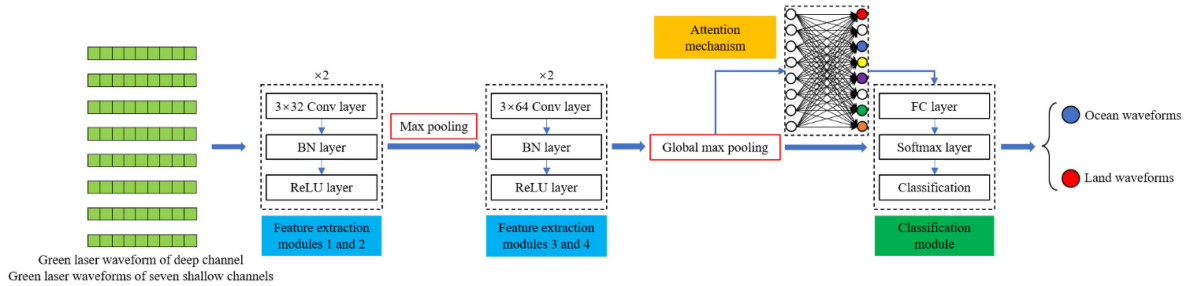

Fig. 20. Structure of the proposed 1-D CNN with eight-channel waveforms but without the multichannel voting module.

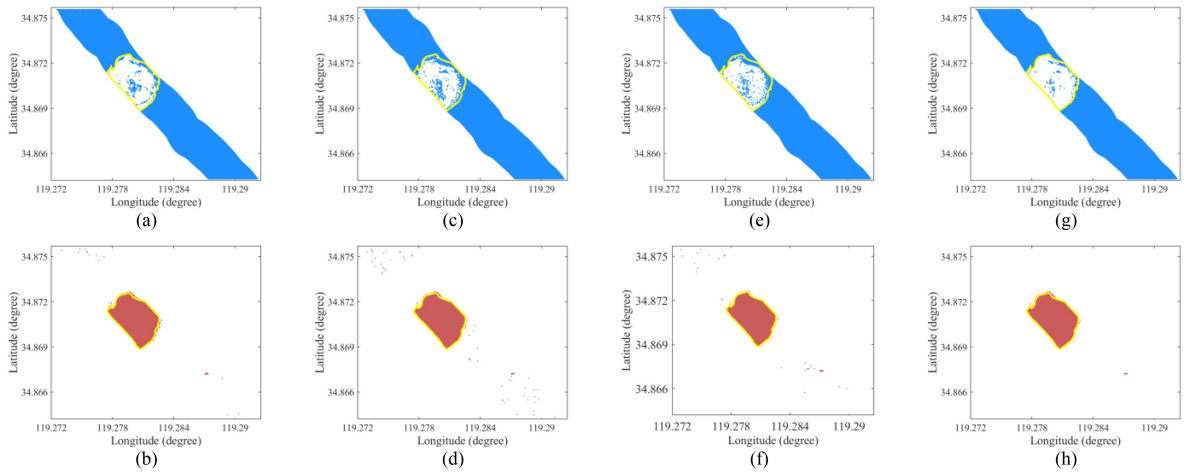

Fig. 21. Positions of laser irradiation points of the corresponding ocean and land waveforms classified by the 1-D CNN with deep channel waveforms (a) and (b), 1-D CNN with shallow-1 channel waveforms (c) and (d), 1-D CNN with eight-channel waveforms (e) and (f), and MVCNN (g) and (h), respectively.

TABLE V
PERFORMANCE METRICS OF THE 1-D CNN WITHOUT AND WITH AM FOR THE TEST SET

Metrics		Without AM	With AM
Ocean	P (%)	98.83	**99.25**
	R (%)	**99.69**	99.60
	F_1 (%)	99.26	**99.43**
Land	P (%)	**98.55**	98.19
	R (%)	94.68	**96.60**
	F_1 (%)	96.56	**97.39**
OA (%)		98.78	**99.06**
Kappa (×100)		95.82	**96.81**
SDOA (%)		0.27	**0.14**

and stronger robustness than the model without AM. The AM improves the OLWC ability of the 1-D CNN for green laser waveforms.

2) Convolutional and Fully Connected Layer Evaluation: The effects of the convolutional and fully connected layers of the proposed 1-D CNN are demonstrated by randomly selecting 10 000 ocean and land deep channel waveforms of the test set and then inputting them to the model trained by the deep channel waveforms. The high-level semantic features are reduced to a 2-D space with a t-distributed stochastic neighbor embedding algorithm [38]. The waveform features from the raw green laser waveforms, fourth convolutional layer, and final fully connected

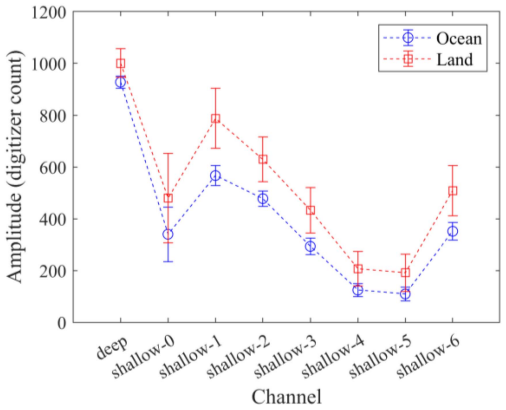

Fig. 22. Averaged amplitude of multichannel green laser waveforms.

TABLE VI
PERFORMANCE METRICS OF THE CONTRAST MODEL AND MVCNN

Metrics		1D CNN with eight-channel waveforms	MVCNN
Ocean	P (%)	98.80	**99.58**
	R (%)	**99.78**	99.70
	F_1 (%)	99.29	**99.64**
Land	P (%)	**99.00**	98.62
	R (%)	94.55	**98.11**
	F1 (%)	96.70	**98.36**
OA (%)		98.83	**99.41**
Kappa (×100)		95.99	**98.00**
SDOA (%)		0.34	**0.03**

layer are shown in Fig. 19(a)–(c). The blue and orange parts are the waveform features representing ocean and land, respectively. As shown in Fig. 19, the features of the raw green laser ocean and land waveforms are scattered and intertwined. After four convolution operations, the waveform features of ocean and land remain linearly indistinguishable, but each of them is gathered and no longer appear intertwined. Finally, the ocean and land waveform features are completely separated by the global integration of the fully connected layers. Overall, when the 1-D CNN entails four convolutional layers and one fully connected layer, it exhibits a strong nonlinear representation of the green laser waveform and can achieve the task of OLWC.

C. Multichannel Voting Module Evaluation

Fig. 20 shows the removal of the multichannel voting module from the MVCNN to form a new model (i.e., the proposed 1-D CNN with eight-channel waveforms). In this manner, the effects of the multichannel voting module of the MVCNN can be evaluated. The eight-channel waveforms (i.e., one deep channel and seven shallow channels) are simultaneously input into the 1-D CNN for training, and the trained model is used to classify the test set. The performance metrics of the 1-D CNN with eight-channel waveforms but without a multichannel voting module and the MVCNN are shown in Table VI. As shown in Tables III and VI, the overall accuracies of the OLWC utilizing 1-D CNN with deep channel waveforms, 1-D CNN with shallow-1 channel waveforms, 1-D CNN with eight-channel waveforms, and MVCNN are 99.06%, 98.77%, 98.83%, and 99.41%, respectively.

Fig. 21 presents the positions of the laser irradiation points of the corresponding ocean (blue) and land (red) waveforms classified by the 1-D CNN with deep channel waveforms, 1-D CNN with shallow-1 channel waveforms, 1-D CNN with eight-channel waveforms, and MVCNN. As shown in Fig. 21(a)–(d), the proposed 1-D CNN with the deep channel waveform or shallow-1 channel can correctly classify most of the ocean waveforms, even for the deep channel waveforms affected by the merging effect, but it misclassified numerous land waveforms affected by topography. Fig. 21(e) and (f) reveals the tendency of the 1-D CNN with eight-channel waveforms but without a multichannel voting module to superimpose false categories derived by deep and shallow channels, leading to more confusing classification results. By contrast, Fig. 21(g) and (h) shows the ability of the multichannel voting module of MVCNN to pick out the correct waveform category from the eight channels, thus improving the classification accuracy.

In summary, the multichannel voting module can take advantage of the green laser waveforms collected in the deep and shallow channels with different responses to the environment, which is a determinant of the MVCNN for achieving a highly accurate and robust OLWC.

D. Analysis of the Relationships of Multichannel Green Waveforms

Figs. 2 and 3 shows the multichannel waveforms of the two laser pulses. The amplitudes of the green laser waveforms collected by each channel are significantly different. The amplitudes of the multichannel waveforms of 294654 laser pulses in strip C16 are averaged separately according to the ocean and land categories as a means of analyzing the multichannel laser waveforms. The results are shown in Fig. 22. The waveform amplitudes of the different channels significantly differ from one another in terms of both ocean and land waveforms, and the multichannel waveform amplitude curves also show varying patterns.

The factors affecting the intensity of the return waveform are mainly the environmental and equipment parameters. Land has a strong reflection effect on the green laser; meanwhile, the green laser penetrates the ocean, and energy is absorbed by the ocean. Hence, the amplitudes of the ocean waveforms of each channel are smaller than those of the land waveforms. As the return waveforms of different channels are reflected from the same target, the environmental parameters (i.e., wave pattern and turbidity) have a slight influence on the amplitude differences between the different channels. The equipment parameters (i.e., FOV, aircraft attitude, beam scanning pattern, and transmitted laser pulse) are the main influencing factors of the amplitude difference of multichannel waveforms. The larger the FOV of the channel, the higher the energy of the received laser pulse and the greater the return waveform intensity. Thus, the deep channel (6–40 mrad) has a greater waveform intensity than the shallow channel (1.9 mrad).

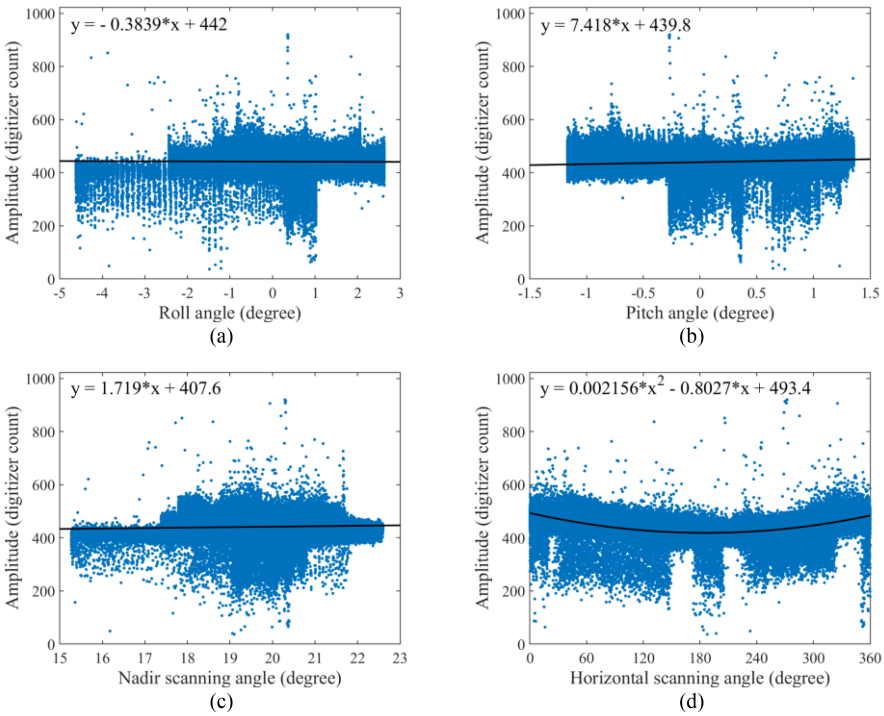

Fig. 23. Relationship between amplitude difference and (a) roll angle, (b) pitch angle, (c) nadir scanning angle, and (d) horizontal scanning angle.

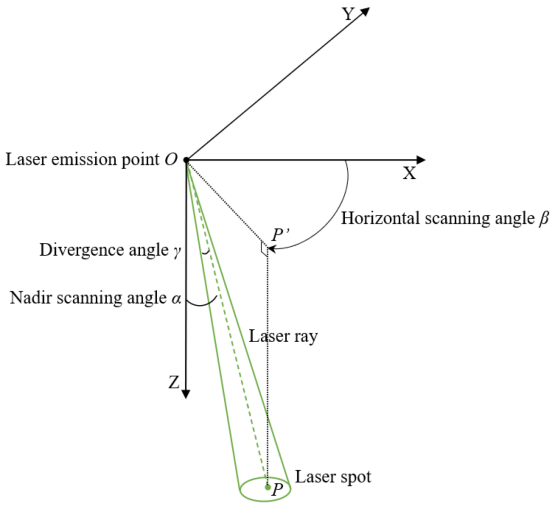

Fig. 24. Schematic diagram of the nadir and horizontal beam scanning angles.

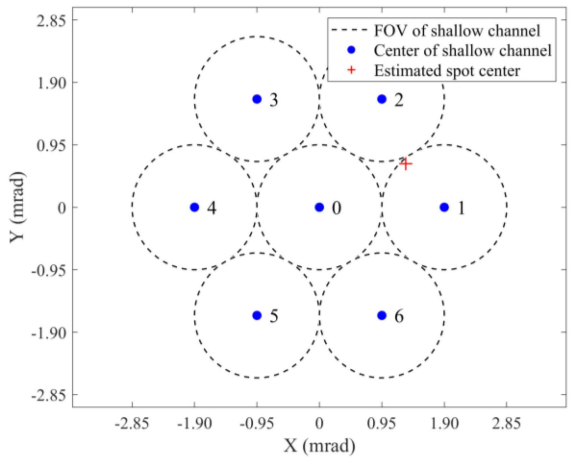

Fig. 25. Estimated spot center in shallow channel FOVs.

Notably, FOV is not the reason for the amplitude difference of shallow channels because each shallow channel has the same FOV. Fig. 23 shows the scatter plots and fitted lines of the amplitude difference between the shallow-1 channel and shallow-4 channel of the ocean waveforms varying with roll angle, pitch angle, nadir scanning angle, and horizontal scanning angle. Fig. 24 shows a schematic diagram of the nadir and horizontal beam scanning angles, with the *X*-axis representing the flight direction and the *Z*-axis representing the vertical direction. As shown in Fig. 23, the roll angle, pitch angle, and nadir scanning angle have a slight effect on the amplitude difference, and the horizontal scanning angle has a quadratic function relationship with the amplitude difference. As previously shown in Fig. 22, the amplitude difference between the shallow-1 channel and shallow-4 channel is greater than 400 digitizer counts, whereas the effect of the horizontal scanning angle on the amplitude difference is less than 100 digitizer counts. Hence, the horizontal scanning angle is not the main factor of the return waveform intensity difference of the shallow channels.

The energy distribution in the laser spot, which is nonuniform because of the laser beam divergence, has a great impact on the return waveform intensity of each shallow channel. The shallow channel FOV near the spot center receives more energy and

has a greater return waveform intensity than that positioned far away from the spot center. Figs. 5 and 22 have previously demonstrated that the return waveform intensity of the shallow-1 channel is the highest among all shallow channels. Furthermore, the closer the channel to the shallow-1 channel, the higher the waveform intensity of this shallow channel. Hence, the actual spot center is not likely in the FOV center of the shallow-0 channel but in the FOV of the shallow-1 channel. Subsequently, the actual spot center is estimated by defining the coordinate system (see Fig. 25) and setting the coordinates of the FOV center of the shallow-0 channel to $(0, 0)$. Assuming that the laser energy is linearly inverse to the distance from the spot center, the following mathematical equation can be obtained:

$$E = \frac{k}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \tag{12}$$

where k is the scale factor, (x_0, y_0) represents the coordinates of the spot center, (x, y) represents the coordinates of the FOV center of the shallow channel, and E is the waveform amplitude of the shallow channel. The FOV center coordinates and the averaged waveform amplitudes of shallow 0–2 channels are substituted into (12). The three unknown variables in (12) can be obtained as $x_0 = 1.31$, $y_0 = 0.66$, and $k = 500.96$. The location of the estimated spot center is shown in Fig. 25, which is shifted by approximately 1.47 mrad from the FOV center of the shallow-0 channel. In summary, FOV is the main factor affecting the amplitude difference between deep and shallow channels, and spot center offset is the main influencing factor for the amplitude difference between the shallow channels.

V. Conclusion

In this article, an MVCNN comprising one multichannel input module, multiple 1-D CNN modules, and one multichannel voting module is proposed to perform OLWC utilizing multichannel green laser waveforms for single-wavelength ALB systems. First, multiple green laser waveforms collected in the deep and shallow channels are separately input into a multichannel input module. Second, the proposed 1-D CNN structure, which comprises four feature extraction modules, one attention mechanism module, and one classification module, is used to handle each green channel waveform. Finally, the proposed multichannel voting module is used to perform majority voting on the predicted categories derived by each 1-D CNN model and output the final waveform categories.

The effectiveness of the proposed MVCNN is evaluated using a dataset collected by Optech CZMIL in a study area located in the coastal waters of Lianyungang City, China. Then, the proposed MVCNN is compared with three well-known state-of-the-art methods for OLWC utilizing single-wavelength ALB systems. The results indicate that MVCNN can achieve an OA, Kappa, and SDOA of 99.41%, 0.9800, and 0.03%, respectively, which are better than those of the other methods. MVCNN is slightly affected by aquaculture rafts and the merging effect of the green laser waveform in very shallow waters. An in-depth evaluation of the different submodules of the MVCNN is conducted using various comparison, quantitative, and visual experiments.

1) The multichannel input module plays a crucial role in improving the accuracy of the MVCNN. Experimentally, the OLWC ability of the green laser waveforms collected in different green laser channels is dissimilar, and the classification accuracy of the MVCNN is improved gradually with the increase in the number of channels.
2) Compared with the previous 1-D CNN structure, the proposed 1-D CNN structure is more suitable for OLWC utilizing green laser. This configuration is an essential factor for the excellent performance of the MVCNN.
3) The multichannel voting module is the critical structure of the MVCNN, as it can take advantage of multichannel green laser waveforms by selecting the correct waveform category from deep and shallow channels with different responses to the environment.

The proposed MVCNN outperforms all well-known methods and is recommended for OLWC utilizing single-wavelength ALB systems.

Acknowledgment

The authors would like to thank the Survey Bureau of Hydrology and Water Resources of Yangtze Estuary for the data used in this article.

References

[1] G. C. Guenther, *Airborne Laser Hydrography: System Design and Performance Factors*. Rockville, MD, USA: NOAA, 1985, pp. 1–385.
[2] G. C. Guenther, A. G. Cunningham, P. E. LaRocque, and D. J. Reid, "Meeting the accuracy challenge in airborne LiDAR bathymetry," in *Proc. 20th EARSeL Symp., Workshop Lidar Remote Sens. Land Sea*, 2000, pp. 1–27.
[3] R. C. Hilldale and D. Raff, "Assessing the ability of airborne LiDAR to map river bathymetry," *Earth Surf. Process. Landforms*, vol. 33, no. 5, pp. 773–783, Apr. 2008, doi: 10.1002/esp.1575.
[4] P. J. Kinzel, C. J. Legleiter, and J. M. Nelson, "Mapping river bathymetry with a small footprint green LiDAR: Applications and challenges," *J. Amer. Water Resour. Assoc.*, vol. 49, no. 1, pp. 183–204, Feb. 2013, doi: 10.1111/jawr.12008.
[5] G. Mandlburger, M. Pfennigbauer, and N. Pfeifer, "Analyzing near water surface penetration in laser bathymetry—A case study at the River Pielach," in *Proc. ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, 2013, pp. 175–180, doi: 10.5194/isprsannals-II-5-W2-175-2013.
[6] K. Saylam, J. R. Hupp, A. R. Averett, W. F. Gutelius, and B. W. Gelhar, "Airborne LiDAR bathymetry: Assessing quality assurance and quality control methods with Leica Chiroptera examples," *Int. J. Remote Sens.*, vol. 39, no. 8, pp. 2518–2542, Jan. 2018, doi: 10.1080/01431161.2018.1430916.
[7] F. Yang, D. Su, Y. Ma, C. Feng, A. Yang, and M. Wang, "Refraction correction of airborne LiDAR bathymetry based on sea surface profile and ray tracing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6141–6149, Nov. 2017, doi: 10.1109/TGRS.2017.2721442.
[8] J. Zhao, X. Zhao, H. Zhang, and F. Zhou, "Shallow water measurements using a single green laser corrected by building a near water surface penetration model," *Remote Sens.*, vol. 9, no. 5, Apr. 2017, Art. no. 426, doi: 10.3390/rs9050426.
[9] G. Liang, X. Zhao, J. Zhao, and F. Zhou, "Feature selection and mislabeled waveform correction for water–land discrimination using airborne infrared laser," *Remote Sens.*, vol. 13, no. 18, Sep. 2021, Art. no. 3628, doi: 10.3390/rs13183628.
[10] A. Collin, B. Long, and P. Archambault, "Salt-marsh characterization, zonation assessment and mapping through a dual-wavelength LiDAR," *Remote Sens. Environ.*, vol. 114, no. 3, pp. 520–530, Mar. 2010, doi: 10.1016/j.rse.2009.10.011.

[11] A. Collin, P. Archambault, and B. Long, "Predicting species diversity of benthic communities within turbid nearshore using full-waveform bathymetric LiDAR and machine learners," *PLoS One*, vol. 6, no. 6, Jun. 2011, Art. no. e21265, doi: 10.1371/journal.pone.0021265.
[12] A. Collin, B. Long, and P. Archambault, "Merging land-marine realms: Spatial patterns of seamless coastal habitats using a multispectral LiDAR," *Remote Sens. Environ.*, vol. 123, pp. 390–399, Aug. 2012, doi: 10.1016/j.rse.2012.03.015.
[13] X. Zhao, X. Wang, J. Zhao, and F. Zhou, "Water–land classification using three-dimensional point cloud data of airborne LiDAR bathymetry based on elevation threshold intervals," *J. Appl. Remote Sens.*, vol. 13, no. 3, Aug. 2019, Art. no. 034511, doi: 10.1117/1.JRS.13.034511.
[14] X. Zhao, X. Wang, J. Zhao, and F. Zhou, "An improved water–land discriminator using laser waveform amplitude and point cloud elevation of airborne LIDAR," *J. Coastal Res.*, vol. 37, no. 6, pp. 1158–1172, Aug. 2021, doi: 10.2112/JCOASTRES-D-20-00185.1.
[15] X. Zhao, G. Liang, Y. Liang, J. Zhao, and F. Zhou, "Background noise reduction for airborne bathymetric full waveforms by creating trend models using optech CZMIL in the Yellow Sea of China," *Appl. Opt.*, vol. 59, no. 35, pp. 11019–11026, Dec. 2020, doi: 10.1364/AO.402973.
[16] E. Fuchs and G. Tuell, "Conceptual design of the CZMIL data acquisition system (DAS): Integrating a new bathymetric lidar with a commercial spectrometer and metric camera for coastal mapping applications," in *Proc. SPIE*, vol. 7695, 2010, Art. no. 76950U, doi: 10.1117/12.851919.
[17] B. Cao, S. Zhu, Z. Qiu, and B. Cao, "Water–land classification for LiDAR bathymetric data based on echo waveform characteristics," *Hydrographic Surv. Charting*, vol. 38, no. 3, pp. 12–16, May 2018, doi: 10.3969/j.issn.1671-3044.2018.03.004.
[18] T. Huang et al., "Classification of sea and land waveform based on multichannel ocean LiDAR," *Chin. J. Lasers*, vol. 44, no. 6, Jun. 2017, Art. no. 0610002, doi: 10.3788/CJL201744.0610002.
[19] Z. Qiu and B. Cao, "Water–land classification method for airborne LiDAR bathymetric data," in *Proc. China Inertial Technol. Soc. High–End Front. Special Topics Acad. Conf.*, 2017, pp. 325–335.
[20] S. Hu, Y. He, B. Tao, J. Yu, and W. Chen, "Classification of sea and land waveforms based on deep learning for airborne laser bathymetry," *Infrared Laser Eng.*, vol. 48, no. 11, Nov. 2019, Art. no. 1113004, doi: 10.3788/IRLA201948.1113004.
[21] S. Pe'eri, L. V. Morgan, W. D. Philpot, and A. A. Armstrong, "Land–water interface resolved from airborne LIDAR bathymetry (ALB) waveforms," *J. Coastal Res.*, vol. 62, pp. 75–85, Mar. 2011, doi: 10.2112/SI_62_8.
[22] S. Pe'eri and W. Philpot, "Increasing the existence of very shallow-water LIDAR measurements using the red-channel waveforms," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1217–1223, May 2007, doi: 10.1109/TGRS.2007.894584.
[23] E. Fuchs and A. Mathur, "Utilizing circular scanning in the CZMIL system," in *Proc. SPIE*, vol. 7695, 2010, Art. no. 76950W, doi: 10.1117/12.851936.
[24] A. Payment, V. Feygels, and E. Fuchs, "Proposed LiDAR receiver architecture for the CZMIL system," in *Proc. SPIE*, vol. 7695, 2010, Art. no. 76950Y, doi: 10.1117/12.851971.
[25] J. W. Pierce, E. Fuchs, S. Nelson, V. Feygels, and G. Tuell, "Development of a novel laser system for the CZMIL lidar," in *Proc. SPIE*, vol. 7695, 2010, Art. no. 76950V, doi: 10.1117/12.851933.
[26] S. Kiranyaz, T. Ince, R. Hamila, and M. Gabbouj, "Convolutional neural networks for patient-specific ECG classification," in *Proc. EMBC*, 2015, pp. 2608–2611, doi: 10.1109/EMBC.2015.7318926.
[27] J. Gu et al., "Recent advances in convolutional neural networks," *Pattern Recognit.*, vol. 77, pp. 354–377, May 2018, doi: 10.1016/j.patcog.2017.10.013.
[28] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, vol. 37, pp. 448–456.
[29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, Jun. 2017, doi: 10.1145/3065386.
[30] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. Int. Conf. Mach. Learn.*, 2010, pp. 807–814.
[31] Y. L. Boureau, J. Ponce, and Y. LeCun, "A theoretical analysis of feature pooling in visual recognition," in *Proc. Int. Conf. Mach. Learn.*, 2010, pp. 111–118.
[32] M. Vogt, "An overview of deep learning techniques," *At-Automatisierungstechnik*, vol. 66, no. 9, pp. 690–703, Sep. 2018, doi: 10.1515/auto-2018-0076.
[33] D. Yu, H. Wang, P. Chen, and Z. Wei, "Mixed pooling for convolutional neural networks," in *Proc. RSKT*, 2014, vol. 8818, pp. 364–375, doi: 10.1007/978-3-319-11740-9_34.
[34] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Is object localization for free?—Weakly-supervised learning with convolutional neural networks," in *Proc. Comput. Vis. Pattern Recognit. Conf.*, 2015, pp. 685–694, doi: 10.1109/CVPR.2015.7298668.
[35] O. Russakovsky et al., "Imagenet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Apr. 2015, doi: 10.1007/s11263-015-0816-y.
[36] K. Banerjee et al., "Exploring alternatives to softmax function," Nov. 2020, doi: 10.48550/arXiv.2011.11538.
[37] Y. Ho and S. Wookey, "The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling," *IEEE Access*, vol. 8, pp. 4806–4813, 2020, doi: 10.1109/ACCESS.2019.2962617.
[38] V. D. M. Laurens and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, Nov. 2008.

Gang Liang received the B.S. degree in remote sensing science and technology in 2020 from Shandong Agricultural University, Shandong, China, where he is currently working toward the M.S. degree in photogrammetry and remote sensing.

His current research interest includes airborne LiDAR bathymetry.

Xinglei Zhao (Member, IEEE) received the Ph.D. degree in geodesy and geomatics from Wuhan University, Hubei, China, in 2018.

He is currently an Associate Professor with Shandong Agricultural University, Shandong, China. His current research interests include ocean remote sensing and airborne oceanic LiDAR.

Jianhu Zhao received the Ph.D. degree in geodesy and geomatics from Wuhan University, Hubei, China, in 2002.

He is currently a Professor with the School of Geodesy and Geomatics, Wuhan University. His research interests include underwater topography, seafloor sediment classification, marine hydrology, and underwater navigation.

Fengnian Zhou received the M.S. degree in surveying and mapping engineering from Wuhan University, Hubei, China, in 2006.

He is currently a Professorate Senior Engineer with the Survey Bureau of Hydrology and Water Resources of Yangtze Estuary, Shanghai, China. He is also the Director with the Survey Bureau of Hydrology and Water Resources of Yangtze Estuary. His research interests include hydrology and water resources.