Convolutional Neural Network to Retrieve Water Depth in Marine Shallow Water Area From Remote Sensing Images

Bo Ai, Zhen Wen, Zhenhua Wang, Ruifu Wang, Dianpeng Su, Chengming Li, and Fanlin Yang

***Abstract*—The local connection characteristics of convolutional neural network (CNN) are linked with the local spatial correlation of image pixels for water depth retrieval in this article. The method has greater advantages and higher precision than traditional retrieval methods. Traditional remote sensing empirical models require manual extraction of retrieval factors and the process is complex. This article proposes a model based on CNN, which uses different remote sensing images in four spectral bands, red, green, blue, and near-infrared, to retrieve the water depth. In general, CNN is mostly used for image recognition and classification tasks, which can make full use of the local spatial correlation between pixels. The method in this article exploits this feature of CNN for water depth retrieval, taking into consideration the nonlinear relationship between the radiance value and water depth value from adjacent and central pixels. In this article, remote sensing image data, measured water depth data, and lidar sounding data are used as input data to build the model. Then, the retrieval error is analyzed and the parameters are adjusted to further optimize the model. Quantitative analysis and experimental results show that the accuracy of the CNN model in shallow sea areas retrieval is improved by more than 50%. The mean absolute error can reach within 0.8 m. Finally, the model is shown to be highly portable and capable of retrieving water depth data with resolution equal to the spatial resolution of the remote sensing image using only a small amount of input water depth data.**

***Index Terms*—Convolutional neural network (CNN), deep learning, shallow water, remote sensing images, water depth retrieval, spatial correlation.**

I. Introduction

OCEAN water depth measurement has always been one of the most important tasks in ocean mapping. The measurement of water depth in shallow sea areas is of great significance for marine transportation, coastal management, and marine environmental monitoring. At present, the methods for obtaining water depth in shallow ocean waters are mainly divided into two categories: field observation and remote sensing image retrieval. The traditional method is shipborne sonar measurement. This method can be highly precise, but it is greatly affected by environmental factors such as weather and sea conditions, it costs a lot of money, and ships cannot enter sensitive areas. Lidar bathymetry is estimated to have good accuracy for river water depth [1], [2]. In recent years, airborne lidar sounding has developed rapidly for shallow waters [3]. This method can be very precise, with an error of less than 0.2 m, but the cost of equipment is very high, and it is best suited for small-scale observations [4], [5]. Using models to retrieve water depth from high-resolution remote-sensing images has become a possibility with the development of remote sensing technology, and much exploratory work on remote sensing image-based water depth retrieval models has been carried out at home and abroad. This work can be roughly divided into three categories: physical models, empirical models, and mathematical models. Chen *et al.* proposed a new physics-based dual-band that uses the blue and green bands of high spatial resolution multispectral images to estimate shallow water depth [6]. This algorithm can retrieve the water depth in the absence of real data, which is a huge breakthrough, but the extraction of the retrieval factor is complex and the retrieval accuracy is not ideal. Kerr *et al.* proposed an algorithm for optically extracting water depth from multispectral images of coral reef landscapes which is effective, but not very portable [7]. A bathymetry estimation approach that combines a physical wave model with a statistical method based on Gaussian process regression learned was derived [8]. Local water depth was successfully estimated through wave dispersion relationship [9]. It is difficult for physical models to comprehensively estimate the factors and weights that affect water depth [10]–[12]. Stumpf *et al.* proposed an empirical model that uses high-resolution satellite images of different seabed types to determine water depth with low precision [13]. Icebergs were used as depth sounders, but the uncertainties associated with the methods were high [14]. Although this method has been adopted for the purpose of navigational charting [15], they are difficult to propose accurate and universal water depth retrieval models from a scientific perspective [16], [17]. Bian *et al.* proposed a through-water photogrammetric approach by using feature extraction and image geometry from multitemporal Sentinel-1

Manuscript received March 5, 2020; revised April 19, 2020; accepted April 30, 2020. Date of publication May 28, 2020; date of current version June 16, 2020. This work was supported in part by The Key Program of the National Natural Science Foundation of China under Grant 41930535, in part by the SDUST (Shandong University of Science and Technology) Research Fund under Grant 2019TDJH103. *(Corresponding author: Fanlin Yang.)*

Bo Ai, Zhen Wen, Ruifu Wang, Dianpeng Su, and Fanlin Yang are with the College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266000, China (e-mail: aibo@sdust.edu.cn; wwzhen01@163.com; wrf@sdust.edu.cn; sudianpeng@126.com; flyang@126.com).

Zhenhua Wang is with the Duke University, Durham, NC 27706 USA (e-mail: zhenhua.wang@duke.edu).

Chengming Li is with the Chinese Academy of Surveying & Mapping, Beijing 100000, China (e-mail: cmli@casm.ac.cn).

This article has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the authors.

Digital Object Identifier 10.1109/JSTARS.2020.2993731

SAR data [18]. It is difficult to achieve high precision because the low correlation between SAR images and water depth. The regression kriging (RK) approach was used to combine the optimal spatial interpolation of kriging with the high-resolution auxiliary information of multispectral imagery for a detailed bathymetric mapping [19]. A multi-Gauss function was used to retrieve water depth using laser radar waveform data [20]. A quasi-analytical algorithm approach was proposed for estimating the water depth around Weizhou Island [21]. Mathematical models are difficult to achieve good portability due to strict calculations. In addition, many companies also have commercial-oriented water depth products such as EOMAP. The cost of data is high, and the accuracy of the data needs to be verified with the measured data. Some teams also made 3-D water depth models for specific areas, all without considering the portability of the model [22]–[24].

At present, most research is based on physical models, empirical models, and mathematical models. These traditional water depth retrieval methods do not consider the spatial correlation between neighboring pixels and central pixels in remote sensing images [25], and it is often necessary to manually extract the retrieval factor. This tends to limit the accuracy of these models and complicate the process of building them. In summary, there are still many problems in current water depth retrieval models: they tend to only discuss a single region or a single type of image, have poor portability, lack practical application, and have low precision. With the development of machine learning, new algorithms have been applied to calculate water depth. Wang *et al.* achieves high precision using a spatial distribution support vector machine to perform water depth retrieval from optical images, but the portability of this algorithm has not been verified [26]. Random forest machine learning and multitemporal satellite images were used to create a generalized depth estimation model, but the accuracy was not ideal [27]. As one of the key research objects in deep learning, convolutional neural network (CNN) has been successfully applied to remote sensing images. At present, CNN is mainly used for remote sensing image classification and target detection. CNN performed remote sensing image scene classification, and achieved good results, such as land use and land cover classification [28]–[30]. CNN was also used to detect targets from remote sensing images, such as ships, clouds, and cyanobacteria, which detected targets effectively [31]–[33].

This article makes the novel proposal that CNN be used to retrieve water depth data in shallow marine areas. Two characteristics of CNN make it suitable for the problem of retrieving water depth from remote sensing images. First, the CNN's local join feature makes CNN suitable for processing raster data with a spatial grid structure. It can be seen from the geographic interpolation algorithm that the depths of unknown points can be regarded as weighted averages of nearby known depths. Therefore, this article explores the impact of adjacent pixels on water depth retrieval through CNN. In addition, in the traditional remote sensing water depth retrieval research, people often use the relationships between water depth and radiance values in remote sensing images to create a retrieval factor which relates these two values. The convolutional layer of CNN in the model can be used to extract these data feature values. Therefore, the complicated process of manually extracting the retrieval factor can be avoided [34]. The resulting model has high precision and high portability.

II. Methods

A. Overview

The CNN-based water depth retrieval method proposed in this article first preprocesses the image and water depth data, so that the remote sensing image and the measured water depth match spatially, then divides the image into subimages centered on the measured points. Next, it randomly samples the training and test data and inputs data into the CNN for training. During training, the CNN adjusts its parameters and builds the water-depth retrieval model. The model takes subimages as input and outputs water depth values. The model is validated using data from the area around Robert Island to verify the scalability of the model. The Robert Island data are first simplified, then the model is used to retrieve depth values for the shallow waters areas, and finally the results are compared with high-resolution data obtained by airborne lidar sounding. The goal is to obtain water depth measurement data for an entire region using only a small amount of preexisting water depth data. The water depth data retrieved by the model can achieve the same resolution as the remote sensing image used as model input. A detailed flow chart for this process is shown in Fig. 1.

B. Data Preparation

1) Research Area: The research area utilized in this article is the region surrounding North Island and Robert Island in the Xisha Islands, as shown in Fig. 2. These two islands are under the jurisdiction of Sansha City, Hainan Province. They are in the central part of South China Sea, southeast of Hainan Province. They have a tropical maritime monsoon climate. Robert Island covers an area of 0.3 square kilometers and North Island covers an area of 0.4 square kilometers. These islands are surrounded by shallow seas with a water depth of less than 50 m, which are suitable for an application of the shallow water depth retrieval method in this article.

2) Data: The remote sensing image data comes from three satellites: Resources Satellite Three (ZY3), Gaofen-1 (GF1), and Worldview-2 (WV2) [35]. All three images use four bands: three visible light bands (450–525 nm) and one near-infrared band (700–2500 nm) [36]. ZY3 is the first independent civilian high-resolution stereo mapping satellite in China, which can stably acquire full-color stereo images, multispectral images, and auxiliary data. This article uses the multispectral images (two scenes) from ZY3 with a 5.8-m resolution. The GF1 satellite is the first satellite of China's high-resolution Earth observation system. It incorporates breakthroughs in the key technologies of high spatial resolution, multispectral imaging, and wide-coverage combined with optical remote sensing. This letter uses multispectral image data (one scene) from GF1 with 8 m resolution. The WV2 is a commercial satellite that provides its users around the world with high-performance imaging products that meet their needs. This letter uses multispectral image data (one scene) from WV2 with 1.8 m resolution.

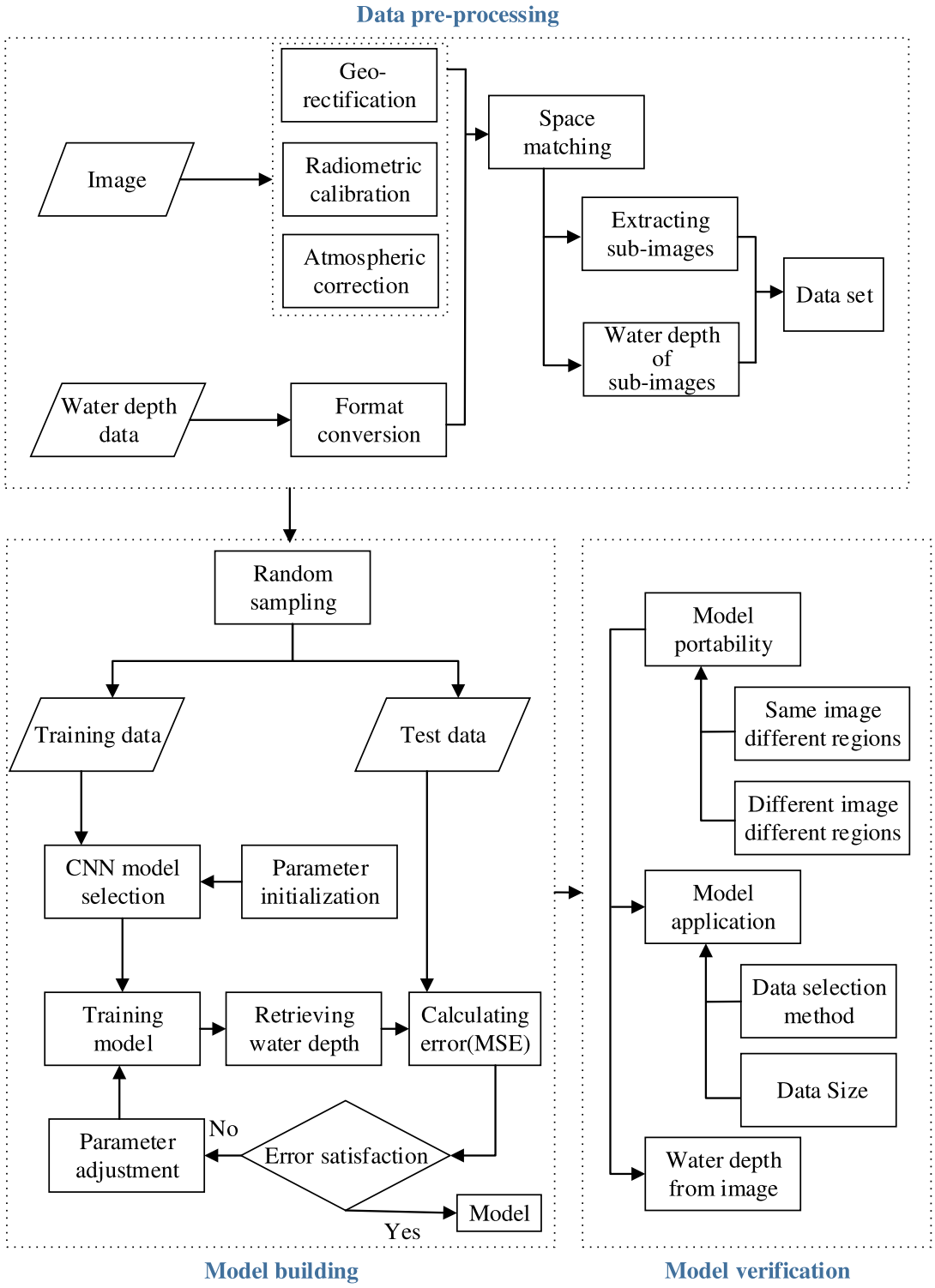

Fig. 1. Water depth retrieval flowchart.

The experimental data also contains two types of water depth data. The first is the land and underwater topographic survey map of the North Island, measured in April 2013, which uses a measuring scale of 1:2000 and the China geodetic coordinate system 2000. The second is airborne lidar sounding data with an error of less than 0.2 m, including measurements taken in both 2016 and 2019.

The data used in this article is divided into two parts: model construction data and model verification data. Model construction data are used to build the model and adjust the parameters. Model validation data are used to test the robustness and application of the model. The model construction data includes the images of North Island from ZY3 and WV2, as well as the measured water depth data from the 2013 survey. The model

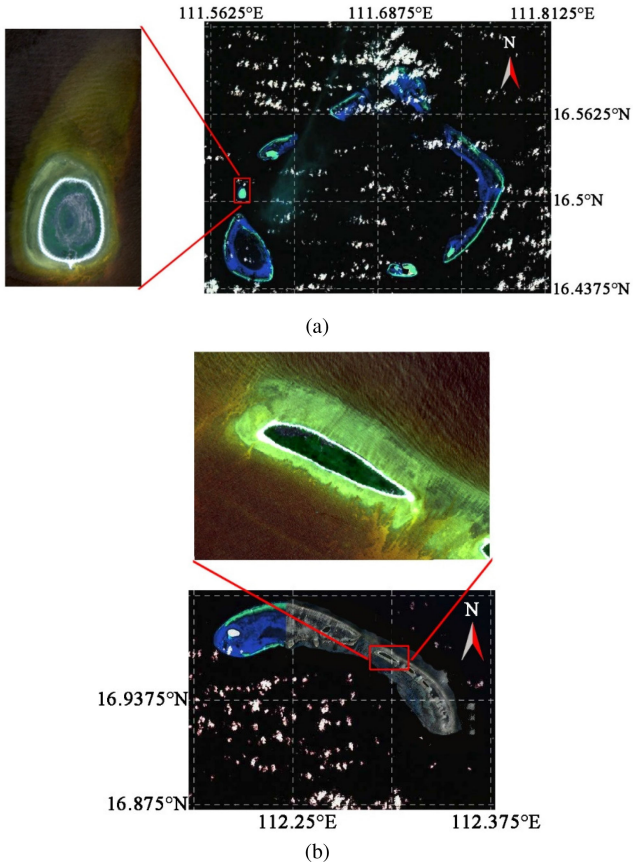

Fig. 2. Study area. (a) Robert Island. (b) North Island.

TABLE I
DATASET INFORMATION

Data	Date	Area	Application
ZY3 (Scene 1)	Apr. 26, 2013	North Island	Construction
ZY3 (Scene 2)	Feb. 28, 2019	Robert Island	Verification
WV2	Feb. 10, 2010	North Island	Construction
GF1	May. 17, 2015	Robert Island	Verification
Survey Data	Apr. 2013	North Island	Construction
Lidar Data 1	May. 2016	Robert Island	Verification
Lidar Data 2	May. 2019	Robert Island	Verification

verification data includes the images of Robert Island from ZY3 and GF1, as well as the airborne lidar sounding water depth data. Table I shows the information of all datasets in detail, including date, area, and application.

3) Data Preprocessing: The water depth data are preprocessed. First, the airborne lidar sounding data (.las) and the measured water depth (.dwg) are converted into shapefile. Then, coordinate correction is performed on the data points. Finally, the remote sensing image is preprocessed, geometrically corrected, radiometrically corrected, and atmospherically corrected, which results in the water depth data matching the remote sensing image in space. The pixel values of the remote sensing image reflect the radiance value, so there is spatial correlation between the pixels of the remote sensing image. Meanwhile, the local

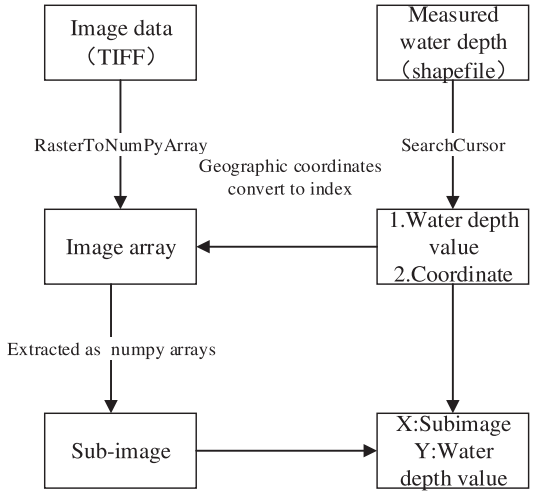

Fig. 3. Subimage extraction process.

connection characteristics of CNN make CNN suitable for processing the relationship between the local pixels. This article uses CNN to study the influence of adjacent pixels on water depth retrieval, a method inspired by the geographic interpolation algorithm. Before inputting images to CNN, the whole remote sensing image should first be divided into subimages with a certain width centered on the measured water depth points. In the process of extraction, the RasterToNumPyArray tool and the SearchCursor tool provided by ArcPy are used to spatially superimpose the image and the measured water depth point, then a buffer of CellSize is constructed at the corresponding point position for buffer analysis. Finally, the size of the CellSize subimage block is extracted [37]. Fig. 3 is a flow chart of this algorithm.

The geographic coordinates of the remote sensing image need to be converted into an index and stored when extracting subimages. The conversion formulas are as follows:

$$\mathrm{Index}_{\mathrm{col}} = \frac{X - X_{\min}}{\mathrm{CellSize}_X} \tag{1}$$

$$\mathrm{Index}_{\mathrm{row}} = \frac{Y - Y_{\max}}{-\mathrm{CellSize}_Y}. \tag{2}$$

In the formula, $\mathrm{Index}_{\mathrm{col}}$ represents the index of the converted column, $\mathrm{Index}_{\mathrm{row}}$ represents the index of the converted row, X represents the latitude of the current point, $X_{\min}$ represents the latitude of the lower-left corner of the image, Y represents the longitude of the current point, $Y_{\max}$ represents the longitude of the upper right corner of the image, $\mathrm{CellSize}_X$ represents the width of a cell in the remote sensing image, and $\mathrm{CellSize}_Y$ represents the length of a cell.

C. Model

1) Model Basis: This article designs a CNN model for water depth retrieval with the objective of improved accuracy. Since the original intention of CNN is to tackle classification tasks such as image recognition, this article needs to adjust the network

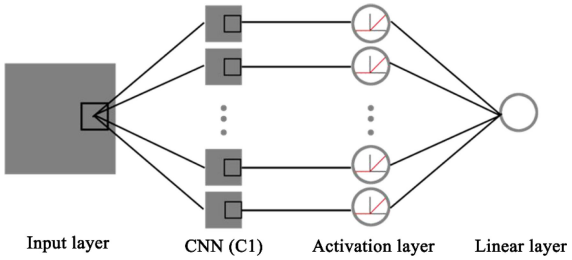

Fig. 4. CNN model structure.

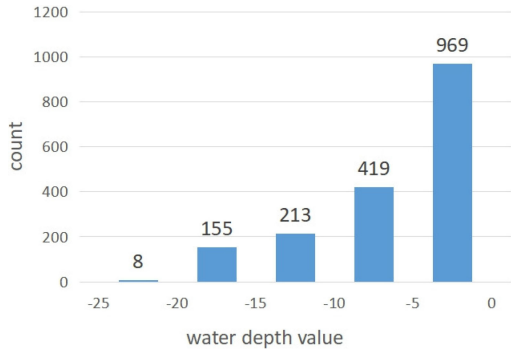

Fig. 5. Measured data distribution.

model structure so that the CNN model can be applied to the regression task of water depth retrieval [38].

This article considers other classic CNN, such as AlexNet, VGG, and ResNet. AlexNet successfully applied tricks such as ReLU, Dropout and LRN in CNN for the first time, which applies the basic principles of CNN to deeper and wider network [39]. In the VGG structure, the convolution kernel focuses on expanding the number of channels, and pooling focuses on narrowing the width and height, making the model architecture deeper, and wider and the slow down the increase of the calculation [40]. ResNet has designed a residual structure using skip connection, which makes the network reach a deep level and improves the performance [41]. These classic algorithms have a common feature: exploring deeper and wider neural networks to process larger dimensional images. However, the test shows that the deeper networks are not suitable for the retrival method. Because the dataset cannot reach a larger dimension in terms of structure and data volume, and is not suitable for deeper networks.

There is no theoretical support for the setting of the number of convolutional layers currently. Through the experiment, the multilayer convolutional layer has no obvious effect on the retrieval results, and the time of model construction is longer. So this article, uses CNN with only one convolutional layer to perform the retrieval work. The structure of the CNN model is shown in Fig. 4.

The convolutional layer is the core structure of CNN, and it has the characteristics of a local receptive field. The local receptive field means that each neuron of convolution layer is only connected to the input domain in a small part of the area [42]. This structure enables CNN to better grasp the spatial correlation between remote sensing image pixels when processing remote sensing data, which improves the accuracy of water depth retrieval [43]. The expression for the convolution operation is as follows [44]:

$$\begin{aligned} Y(i,j) &= (I*K)(i,j) \\ &= \sum_m \sum_n I(i+m, j+n) K(m,n) \end{aligned} \tag{3}$$

where I can be regarded as a 2-D image as input, K represents a 2-D kernel function, i and j are variables, and m and n are constants. In the CNN model, the variation of the parameters of the front layer network will cause the distribution of the input values for each layer to change, which may lead to the failure of the model training. In response to this problem, this article selects the BatchNormalization layer to normalize the results of each hidden layer and then input them to the next layer [45]. The formula for the standardization process is as follows:

$$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \tag{4}$$

$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv \mathrm{BN}_{\gamma,\beta}(x_i). \tag{5}$$

In this formula, $\mu_B = \frac{1}{m}\sum_{i=1}^{m} x_i$ is the mean of the data, $\sigma_B^2 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_B)^2$ is the data variance, $\hat{x}_i$ is the standardized remote sensing pixel data, y_i is the water depth data after scaling and moving, and γ and β are parameters calculated in the learning process.

The purpose of the activation function is to add nonlinear capability to the CNN model, which improves the ability of the CNN model to fit the nonlinear relationship between water depth and image pixel values. This article uses a ReLU activation function with sparse characteristics. ReLU can increase the nonlinear fitting ability, save computing resources, and reduce gradient disappearance, and over-fitting [46]. The formula of ReLU is as follows:

$$y = \max(0, x). \tag{6}$$

In a typical CNN, a tanh activation layer or softmax activation layer is output, because the purpose of the classification task is to output the probability that the image belongs to a certain category or directly output the category of the image [47]. In the water depth retrieval task, however, the output is the water depth value calculated by the model. The water depth value is a number with a range of R, so the tanh or softmax activation layer cannot be selected. In this article, the linear activation layer is adopted as the output layer at the end of the model, which makes the value range of the output value satisfy the requirements of water depth retrieval [48].

2) Model Creation: Models are constructed using North Island image data and measured water depth data. Fig. 5 is a statistical diagram of the water depth distribution of the study area, showing that the measured water depth data used in this article are mostly distributed in the shallow sea area, with only a small amount of water more than 20-m deep. In this area, the maximum water depth is 20.2 m, the minimum water depth is 0.1 m, and the average water depth is 7.6 m.

Data partitioning is one of the important prerequisites for successful CNN training. The uniform distribution of data can

TABLE II
NEAR PIXEL SELECTION EXPERIMENT

Number of adjacent pixels	MSE	
	ZY3	WV2
9*9	0.889	0.731
8*8	0.865	0.649
7*7	0.849	0.663
6*6	0.864	0.721
5*5	0.975	0.790
4*4	1.047	0.842

reduce the influence of outliers in the data on the model, thus improving the precision of model training. Therefore, this article randomize the data to divide the two sets of data into two separate partitions such that the test data and training data have uniform distributions of water depths. Training data, verification data, and test data basically follow a 6:2:2 ratio. This article divides the 1721 sets of remote image sensing data from WV2 into 1100 sets of training data, 300 sets of verification data, and 321 sets of test data. The 427 sets of extracted remote sensing image data from ZY3 are divided into 320 sets of training data, 50 sets of verification data, and 57 sets of test data.

The randomly sampled training data and verification data, along with the ZY3 data, are input into the CNN model, and the corresponding training errors and test errors of WV2 and ZY3 are calculated, respectively. The parameters are optimized according to two types of errors.

The error function used in the CNN model training is mean square error (MSE), and the formula is as follows:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(X_i - X)^2. \tag{7}$$

In the above formula, X is the measured water depth value; X_i is the water depth value obtained from the model retrieval; N is the total amount of data.

A gradient-based optimization algorithm called adamax is used for training. The optimization parameters are set as follows: the learning rate is set to 0.02, $\beta 1$ is set to 0.9, $\beta 2$ is set to 0.999, ε for numerical stability is set to $1e$-8, and the attenuation rate is set to 0.0 [49].

The curves of the training error and the verification error are output after the training is completed. Fig. 6(a) shows the error curve of ZY3 data. The two types of errors reach saturation after about 200 training cycles, and then the error fluctuates within a certain small range. Fig. 6(b) shows the error curve of WV2 data. The resulting shows that the two types of errors reach saturation after about 150 training cycles.

3) Parameter Adjustment: The topography of the shallow sea area is often continuously changing. Therefore, considering the terrain of the adjacent area in the experiment will have a positive impact on the retrieval of water depth. In this section, the subimage widths selected in this article are 3, 4, 5, 6, 7, and 8 pixels, and preset experiments are established to compare the corresponding retrieval precisions between the different sizes of pixel.

As can be seen from Table II, the number of suitable neighboring pixels is approximately distributed around $6 * 6$–$8 * 8$.

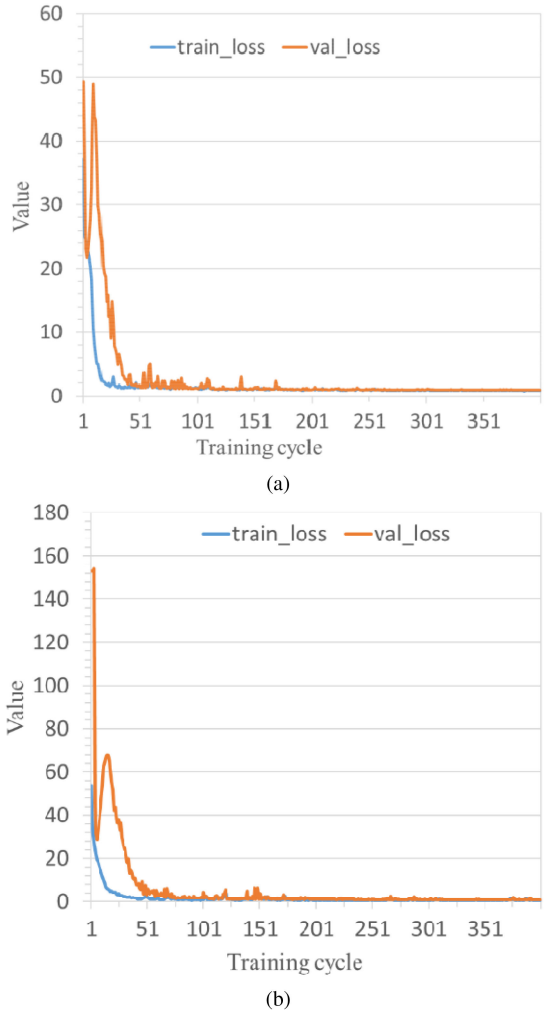

Fig. 6. Error curve of training. (a) ZY3 data. (b) WV2 data.

The best result of ZY3 appears in $7 * 7$, and the best result of WV2 appears in $8 * 8$. Since the resolution of the WV2 image used in this article is slightly higher than the resolution of the ZY3, WV2 includes more pixels than ZY3 for the same area, So, more adjacent pixels can be selected when using a higher resolution image.

There is no accurate method for determining the number of neurons for water depth retrieval currently. Therefore, the prediction experiment is conducted to establish the water depth retrieval model several times with varying numbers of neurons: 9, 12, 15, 18, 21, and 24 [50]. The experimental results are shown in Table III. When using the ZY3 images, best results occur when the number of neurons is 21. When using WV2 images, best results occur when the number of neurons is 18.

III. MODEL RESULTS

Two trained models are used for two sets of test data (WV2, ZY3). The measured water depth value and retrieved water depth value are mapped to observe the result of the retrieval model. The abscissa is the serial number of the water depth point, and the ordinate is the water depth value. The retrieval accuracy map

TABLE III
NEURON SELECTION EXPERIMENT

Number of neurons	MSE	
	ZY3	WV2
24	1.003	0.810
21	0.849	0.652
18	0.876	0.595
15	0.955	0.798
12	1.003	0.790
9	1.125	0.797

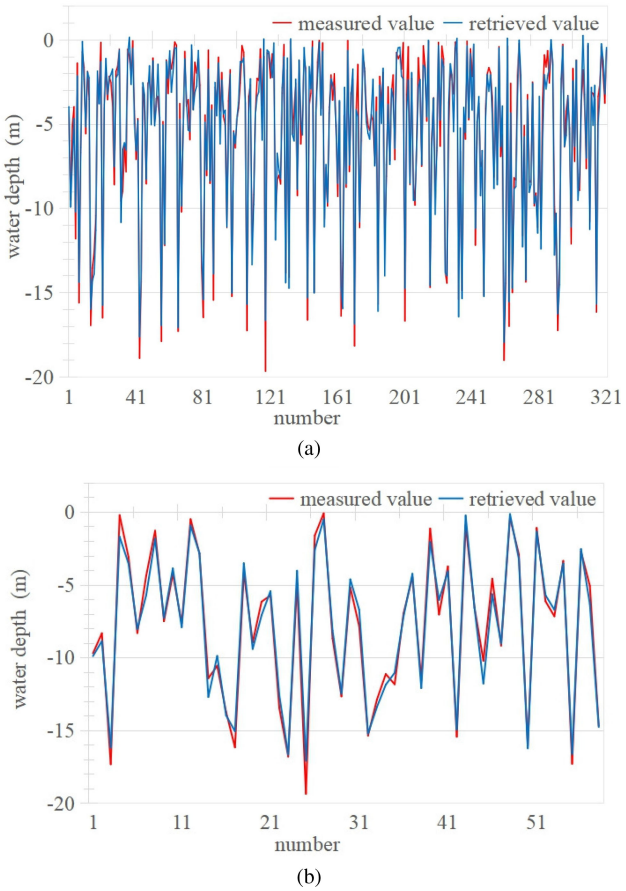

Fig. 7. Retrieval accuracy. (a) ZY3. (b) WV2.

is shown in Fig. 7. The part where the blue line and the red line do not coincide is the error.

As can be seen from the above figure, the model has higher retrieval accuracy in the shallow water area. Visible light and near-infrared cannot pass through the deeper water layer due to the occlusion effect of sediment and marine plants in deeper areas, so the true radiance value of deeper water depths cannot be obtained, which reduces retrieval accuracy. The retrieval errors of different water depth ranges are further summarized in Table IV. The statistical results show that the accuracy of water depth retrieval is significantly reduced when the depth is lower than 15 m; the accuracy of retrieval is better when the depth is greater than –10 m. The mean absolute error (MAE) of ZY3 image can reach 0.668; the MAE of WV2 image can reach 0.447.

TABLE IV
ERROR STATISTICS

Water depth（m）	MAE（m）	
	ZY3	WV2
(-5, 0)	0.688	0.521
(-10, -5)	0.668	0.447
(-15, -10)	0.785	0.673
(-∞, -15)	0.780	0.711

TABLE V
ERROR STATISTICS IN DIFFERENT MODEL

Precision index	linear regression	single-layer neural network	CNN
Mean absolute deviation	3.540（m）	0.729 (m）	0.578 (m）
Mean relative deviation	53.04%	39.56%	19.72%

In this article, Numpy is applied to establish a linear regression model [51]; Keras is applied to establish a single-layer neural network model and a CNN model [52]; and then the above three models are applied to perform retrieval experiments on WV2 data. A precision comparison analysis for these three models is shown in the Table V.

The results show that the CNN has higher water depth retrieval accuracy than the statistical linear regression model and the single-layer neural network without considering adjacent pixels.

IV. DISCUSSION

The data used in the model application test section is the ZY3 imagery, the GF1 satellite imagery, and the airborne lidar sounding data from the Robert Island area. First, the portability of the model is verified, demonstrating that the model can be used in different regions and with different types of images. Then, an application analysis of the model is carried out, and the model is applied to the actual work, which verifies the feasibility of the model in high-resolution water depth retrieval.

A. *Model Portability*

The established model is applied to the model test data, and the adjusted model parameters are retained to perform the test retrieval of the water depth. First, the model is used to construct water depth data for the Robert Island area using the ZY3 images, and the results are compared to the airborne lidar sounding data of the same area. This verification tests the model using images of the same type as the training images in a different geographical region. The model retrieves 8869 sets of water depth data. The average relative error of the water depth data calculated by the model calculation and the airborne lidar sounding data is 0.9485. The two types of data are plotted as scatter plots as shown in Fig. 8(a), in which the *x*-axis is water depth data generated by model, the *y*-axis is the airborne lidar sounding data, and the

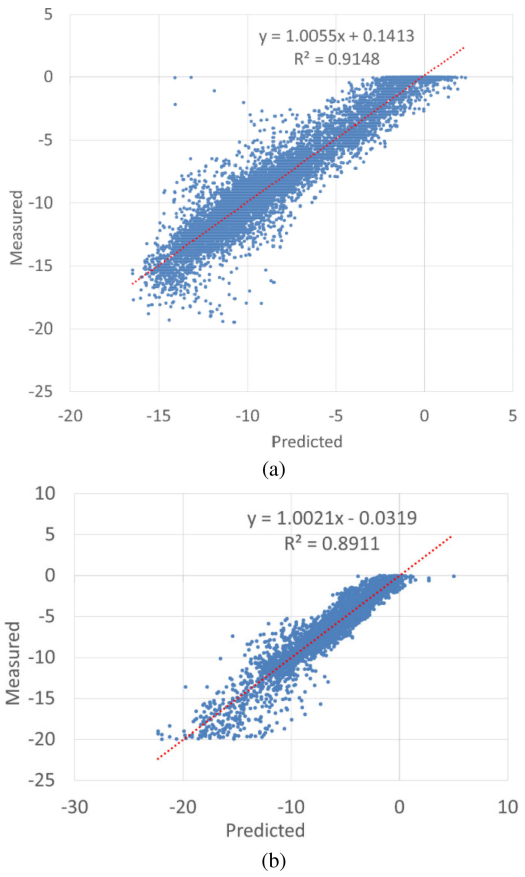

Fig. 8. Correlation coefficient between predicted and measured values. (a) Predicted values from ZY3. (b) Predicted values from GF1.

TABLE VI
DIFFERENT DATA SELECTION METHOD ERRORS

Image source	Number of water depth data	Number of retrieval pixel	Data selection method	MAE (m)
GF1	1462	24367	Random	1.143
GF1	1462	24367	Rectangle	1.244
GF1	731	24367	Random	1.231
GF1	731	24367	Rectangle	1.455

TABLE VII
INFLUENCE OF DATA VOLUME ON THE ERROR OF DIFFERENT REMOTE SENSING IMAGES

Image source	Number of water depth data	Number of retrieval pixel	Proportion	MAE (m)
ZY3	2607	43457	6%	1.405
ZY3	1304	43457	3%	1.540
ZY3	869	43457	2%	1.731
GF1	1462	24367	6%	1.143
GF1	731	24367	3%	1.231
GF1	487	24367	2%	1.268

TABLE VIII
EFFECT OF WATER DEPTH ON ERROR

Water depth range	Number of water depth data	MAE(m)
(-5, 0)	9685	0.965
(-10, -5)	10715	1.327
(-15, -10)	3733	2.088
(-∞, -15)	234	3.765

correlation coefficient of the two types of data reaches 0.9148. Next, the high-resolution satellite imagery and airborne lidar sounding data in Robert Island area are used for verification in the same way, testing the model with both a region and type of image data which are different from those used in model construction. The model retrieves 4874 sets of water depth data and validates both types of data in the same way. The average relative error is 0.9141 and the correlation coefficient is 0.8911, as shown in Fig. 8(b). Through verifying the model against these two sets of data, it can be seen that neither the type of remote sensing images used nor geographical region affect the model training test, and the robustness of the model is thereby illustrated. By observing the two sets of test data maps, the figure shows that the closer to the *X*-axis and *Y*-axis, the greater the water depth value, the more scattered the data, the farther from the fitted straight line. It can be seen that as the degree of data aggregation and accuracy of water depth retrieval decrease as water depth increases.

B. Application Test

This model can be applied to the field of remote sensing image retrieval. Water depth data with the same resolution of high-resolution remote sensing image are retrieved using only a small amount of water depth measurement data and images, which reduces the burden of the actual measurement task. Two remote sensing images and sounding data are used for high-resolution retrieval of water depth. The accuracy of the two selection methods, random selection and strip selection, is tested, and two sets of control experiments are set. Table VI shows the test results. Experiments show that the randomly selected points retrieve the entire region with higher precision because the randomly selected points are more able to summarize the characteristics of the entire region. The fewer the points, the more obvious the advantages of random sampling. The model provides an idea for measurement. Higher accuracy can be obtained with fewer points if point measurements are sampled randomly. Strip-like domain airborne sounding data can also produce high accuracy. Table VI shows the randomly selected data retrieval test parameters.

It can be seen from the table that a small number of points can adequately reflect the high-resolution water depth data. It is possible to obtain better accuracy by using data wherein a higher percentage of retrieval pixels have corresponding water depth data, as shown in Table VII. An example is used to analyze the retrieval accuracy of each water depth range. These 24 367 sets of water depth data retrieved from the GF1 image 3% data (the fifth line in the table) are divided into four parts: –20 to –15 m, –15 to –10 m, –10 to –5m, and –5 to 0 m. MAE is calculated separately for each part. The results are shown in Table VIII.

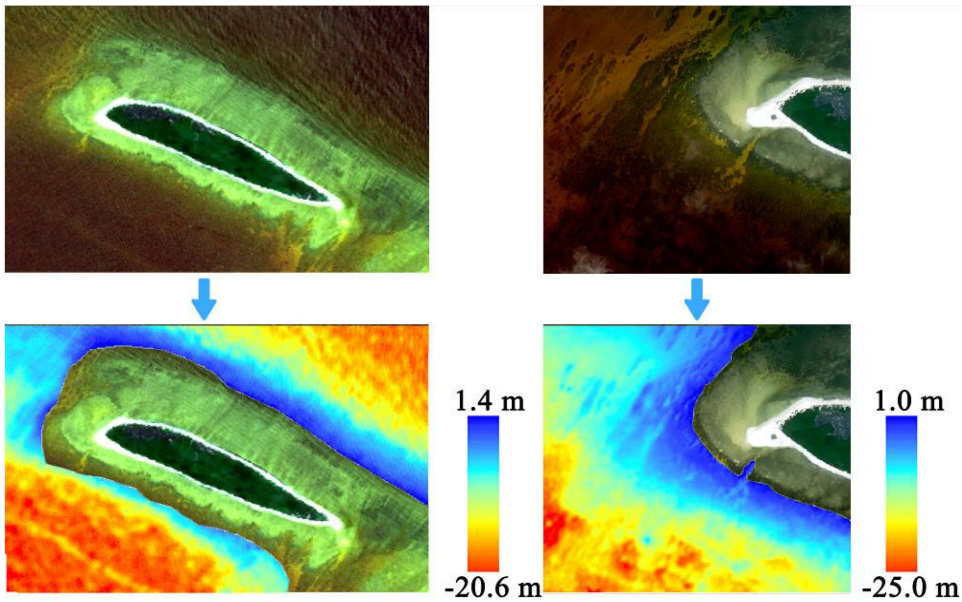

Fig. 9. Retrieval result. (a) Water Depth from ZY3 image. (b) Water Depth from WV2 image.

TABLE IX
APPLICATION ANALYSIS

Image	Total	>0	Proportion(>0)
WV2	175453	55	<0.1%
ZY3	43559	126	<0.3%

Table VII further illustrates that MAE increases as water depth increases.

C. Model Application

The amount of water depth measurement data in the North Island data used in model construction is small. The ZY3 images and WV2 images are input as retrieval data into the water depth retrieval model, and the water depth retrieval effect map (see Fig. 9) is obtained after the retrieval. Map (a) shows the water depth retrieval results from the ZY3 image, and map (b) shows the water depth retrieval result from the WV2 image.

According to all the above results, this model also has the following disadvantages. First, the result contains some outliers (the water depth value of some points is greater than zero) for many reasons, such as the individual singular points in the pixels of the remote sensing image, the influence of water vapor between the sea and land, and the bottom of the shallow water. Table IX shows the statistics of outliers, which records the total number of pixels, the number and proportion of outliers. From the table, the result contains some only a small number of outliers. Second, the portability of the model is limited because the amount of data is limited; better results can be achieved if the model samples are increased.

V. CONCLUSION

This article considers the spatial correlation between remote sensing image pixels and the local connectivity of CNN. This represents the first time that CNN has been applied to the task of retrieving water depth data for shallow marine regions. The data are applied to the model by preprocessing the remote sensing image data and water depth data. The model is initially established to achieve a certain precision through the basic activation function, optimization algorithm, and standardization processing of CNN. It can be concluded that the water depth retrieval model is a complex nonlinear model, and the mapping relationship between water depth values and radiance values in remote sensing images is related to the spatial correlation in the remote sensing images. At the same time, there are many abnormal points in the measured water depth values, which makes the distribution function of the data deviate from the distribution function of the true water depth.

The model is further optimized by adjusting the number of neurons and adjacent pixels, so that the model achieves higher precision. The spatial resolution of the remote sensing image should be considered when selecting the number of neighbors. More neighboring pixels can be selected in the experiment for images with high spatial resolution. The MSE test results of the two scenes of images are 0.849 and 0.595, respectively. The images are ZY3 (Scene1) and WV2, which are shown in Table I. A graphical analysis of the model results can lead to some conclusions: the retrieval effect is good in shallow water areas, while in deep waters, the retrieval accuracy will be adversely affected because the light cannot penetrate deeper waters. The proposed CNN model has higher precision compared to the linear model and the single-layer neural network which doesn't consider adjacent pixels.

The model proposed in this article has a higher application value. The model is verified by using different types of remote sensing images. The results show that the two sets of different data retrieval accuracy MSE can reach 0.9485, 0.9141, respectively, and the water depth and water depth measurement after retrieval have a high correlation. It proves that the model is highly portable, which is superior to the traditional method. Finally, the model is used in practical applications to verify the practical value of the model. High-resolution water depth data are obtained with good accuracy through a small amount of water depth data and remote sensing image data, which is of great significance for water depth measurement in shallow waters of the ocean.

REFERENCES

[1] R. C. Hilldale and D. Raff, "Assessing the ability of airborne LiDAR to map river bathymetry," *Earth Surf. Processes Landforms*, vol. 33, no. 5, pp. 773–783, 2008.

[2] T. Daniele *et al.*, "Mapping river bathymetries: Evaluating topobathymetric LiDAR survey," *Earth Surf. Processes Landforms*, vol. 44, no. 2, pp. 403–678, 2018.

[3] R. Schwarz, G. Mandlburger, M. Pfennigbauer, and N. Pfeifer, "Design and evaluation of a full-wave surface and bottom-detection algorithm for LiDAR bathymetry of very shallow waters," *ISPRS J. Photogram. Remote Sens.*, vol. 150, pp. 1–10, 2019.

[4] P. Westfeld, H. G. Maas, K. Richter, and R. Weiß, "Analysis and correction of ocean wave pattern induced systematic coordinate errors in airborne LiDAR bathymetry," *ISPRS J. Photogram. Remote Sens.*, vol. 128, pp. 314–325, 2017.

[5] 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. 5, no. 11, pp. 6141–6149, Nov. 2017.

[6] B. Chen, Y. Yang, D. Xu, and E. Huang, "A dual band algorithm for shallow water depth retrieval from high spatial resolution imagery with no ground truth," *ISPRS J. Photogram. Remote Sens.*, vol. 151, pp. 1–13 2019.

[7] J. M. Kerr, and S. Purkis, "An algorithm for optically-deriving water depth from multispectral imagery in coral reef landscapes in the absence of ground-truth data," *Remote Sens. Environ.*, vol. 210, pp. 307–324, 2018.

[8] C. Danilo, and F. Melgani, "High-coverage satellite-based coastal bathymetry through a fusion of physical and learning methods," *Remote Sens.*, vol. 11, no. 4, 2019.

[9] S. Gonzalo, C. Daniel, L. Pau, O. Alejandro, and R. Francesca, "UBathy: A new approach for bathymetric inversion from video imagery," *Remote Sens.*, vol. 11, no. 23, 2019.

[10] J. Hedley, C. Roelfsema, and S. R. Phinn, "Efficient radiative transfer model inversion for remote sensing applications," *Remote Sens. Environ.*. vol. 113, no. 11, pp. 2527–2532, 2009.

[11] A. G. Dekker *et al.*, "Intercomparison of shallow water bathymetry, hydro-optics, and benthos mapping techniques in Australian and Caribbean coastal environments," *Limnol. Oceanogr., Methods*, vol. 9, pp. 396–425, 2011.

[12] M. Nazeer, J. E. Nichol, and Y. K. Yung, "Evaluation of atmospheric correction models and Landsat surface reflectance product in an urban coastal environment," *Int. J. Remote Sens.*, vol. 35 no. 16, pp. 6271–6291. 2014.

[13] R. P. Stumpf, K. Holderied, and M. Sinclair, "Determination of water depth with high-resolution satellite imagery over variable bottom types," *Limnol. Oceanogr.*, vol. 48 (1-part-2), pp. 547–556. 2003.

[14] J. Scheick, E. M. Enderlin, E. E. Miller, and G. Hamilton, "First-order estimates of coastal Bathymetry in Ilulissat and Naajarsuit Fjords, Greenland, from remotely sensed iceberg observations," *Remote Sens.*, vol. 11, no. 8, 2019.

[15] M. Hodúl, S. Bird, A. Knudby, and R. Chénier, "Satellite derived photogrammetric bathymetry," *ISPRS J. Photogram. Remote Sens.*, vol. 142, pp. 268–277, 2018.

[16] H. Su, H. Liu, L. Wang, A. M. Filippi, W. D. Heyman, and R. A. Beck, "Geographically adaptive inversion model for improvising bathymetric retrieval from satellite multispectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52 no. 1, pp. 465–476, 2014.

[17] P. Vinayaraj, V. Raghavan, and S. Masumoto, "Satellite-derived bathymetry using adaptive geographically weighted regression model," *Marine Geodesy*, vol. 9 no. 6, pp. 458–478, 2016.

[18] X. Bian, Y. Shao, S. Wang, W. Tian, X. Wang, and C. Zhang, "Shallow water depth retrieval from multitemporal sentinel-1 SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sen.*, vol. 11, no. 9, pp. 2991–3000, Sep. 2018.

[19] H. Su, H. Liu, and Q. Wu, "Prediction of water depth from multispectral satellite imagery—the regression kriging alternative," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2511–2515, Dec. 2015.

[20] Y. Ma, J. Zhang, Z. Zhang, J. Y. Zhang, "Bathymetry retrieval method of LiDAR waveform based on multi-Gaussian functions," *J. Coastal Res.*, vol. 90, pp. 324–331, 2019.

[21] H. Rongyong, Y. Kefu, W. Yinghui, W. Jikun, M. Lin, and W. Wenhuan,. "Bathymetry of the coral reefs of Weizhou island based on multispectral satellite images," *Remote Sens.*, vol. 9, no. 7, 2017.

[22] I. Y. Filina, D. D. Blankenship, M. Thoma, V. V. Lukin, V. N. Masolov, and M. K. Sen, "New 3D bathymetry and sediment distribution in lake Vostok: Implication for pre-glacial origin and numerical modeling of the internal processes within the lake," *Earth Planet. Sci. Lett.*, vol. 276, no. 1–2, pp. 1–114, 2008.

[23] C. Li *et al.*, "Multi-band remote sensing based retrieval model and 3D analysis of water depth in hulun lake, China," *Math. Comput. Model.*, vol. 58, no. 3–4, pp. 765–775, 2013.

[24] T. Uchida and S. Fukuoka, "Quasi-3D two-phase model for dam-break flow over movable bed based on a non-hydrostatic depth-integrated model with a dynamic rough wall law," *Adv. Water Resour.*, vol. 129, pp. 311–327, 2019.

[25] A. Tremeau and P. Colantoni, "Regions adjacency graph applied to color image segmentation," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 735–744, Apr. 2000.

[26] L. Wang, H. Liu, H. Su, and J. Wang, "Bathymetry retrieval from optical images with spatially distributed support vector machines," *Geosci. Remote Sens.*, vol. 56, no. 3, pp. 323–337, 2018.

[27] T. Sagawa, Y. Yamashita, T. Okumura, and T. Yamanokuchi, "Satellite derived bathymetry using machine learning and multi-temporal satellite images," *Remote Sens.*, vol. 11, no. 10, 2019.

[28] Z. Ce *et al.*, "An object-based convolutional neural network (OCNN) for urban land use classification," *Remote Sens. Environ.*, vol. 216, pp. 57–70, 2018.

[29] C. Zhang *et al.*, "Joint deep learning for land cover and land use classification," *Remote Sens. Environ.*, vol. 221, pp. 173–187, 2019.

[30] W. Zhang, P. Tang, and L. Zhao, "Remote sensing image scene classification using CNN-CapsNet," *Remote Sens.*, vol. 11, no. 5, 2019.

[31] D. Chai, S. Newsam, H. K. Zhang, Y. Qiu, and J. Huang, "Cloud and cloud shadow detection in Landsat imagery based on deep convolutional neural networks," *Remote Sens. Environ.*, vol. 225, pp. 307–316, 2019.

[32] J. Pyo *et al.*, "A convolutional neural network regression for quantifying cyanobacteria using hyperspectral imagery," *Remote Sens. Environ.*, vol. 233, 2019.

[33] S. Zhang, R. Wu, K. Xu, J. Wang, and W. Sun, "R-CNN-based ship detection from high resolution remote sensing imagery," *Remote Sens.*, vol. 11, no. 6, 2019.

[34] D. Cunefare, L. Fang, R. F. Cooper, A. Dubra, J. Carroll, and S. Farsiu, "Open source software for automatic detection of cone photoreceptors in adaptive optics ophthalmoscopy using convolutional neural networks," *Scientific Rep.*, vol. 7, 2017.

[35] M. J. Willis, B. G. Herried, M. G. Bevis, and R. E. Bell, "Recharge of a subglacial lake by surface meltwater in northeast Greenland," *Nature*, vol. 518, no. 7538, pp. 223–227, 2015.

[36] G. Cheng, P. Zhou, and J. Han, "Learning rotation-invariant convolutional neural networks for object detection in vhr optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7405–7415, Dec. 2016.

[37] F. Hedefalk, P. Svensson, and L. Harrie, "Spatiotemporal historical datasets at micro-level for geocoded individuals in five Swedish parishes," *Scientific Data*, vol. 4, pp. 1813–1914, 2017.

[38] H. C. Shin *et al.*, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, 2016.

[39] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[40] K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *Comput. Sci.*, 2014.
[41] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. 2016 IEEE Conf. Comput. Vision Pattern Recognit.*, Seattle, WA, 2016, pp. 770–778.
[42] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2015.
[43] J. Nagi, F. Ducatelle, G.A.D. Caro, D. Ciresan, and L. M. Gambardella, "Max-pooling convolutional neural networks for vision-based hand gesture recognition," in *Proc. IEEE Int. Conf. Signal Image Process. Appl.*, Kuala Lumpur, Malaysia, 2011, pp. 342–347.
[44] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2017.
[45] S. Ioffe, and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Int. Conf. Mach. Learn.*, Lille, France, 2015.
[46] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," 2017, [Online]. Available: https://arxiv.org/abs/1710.05941
[47] M. Gandhi, K. Lee, Y. Pan, and E. Theodorou, "Propagating uncertainty through the tanh function with application to reservoir computing," 2018. [Online]. Available: https://arxiv.org/abs/1806.09431
[48] A. W. Kabani, and M. R. El-Sakka, "Object detection and localization using deep convolutional networks with softmax activation and multi-class log loss," in *Proc. Int. Conf. Image Anal. Recognit*. Niagara Falls, Canada, 2016.
[49] D. P. Kingma, and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations*, San Diego, CA, USA, 2015.
[50] F. Chen, G. Chen, Q. He, G. He, and X. Xu, "On CNN universal perceptron," in *Proc. Cellular Nanoscale Netw. Appl., 12th Int. Workshop*, Berkeley, CA, USA, 2010.
[51] B. Ai, Z. Wen, Y. Jiang, S. Gao, and G. Lv, "Sea surface temperature retrieval model for infrared remote sensing images based on deep neural network," *Infrared Phys. Technol.*, vol. 99, pp. 231–239, 2019.
[52] S. Wang *et al.*, "Target analysis by integration of transcriptome and chip-seq data with BETA," *Nature Protocols*, vol. 8, no. 1, pp. 2502–2515, 2013.