IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 10, 2019, accepted June 26, 2019, date of publication July 8, 2019, date of current version August 7, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2927366

Sediment Classification of Small-Size Seabed Acoustic Images Using Convolutional Neural Networks

XIAOWEN LUO[1], XIAOMING QIN[1], ZIYIN WU[1,2], FANLIN YANG[3], MINGWEI WANG[3], AND JIHONG SHANG[1]

[1]Key Laboratory of Submarine Geosciences, State Oceanic Administration and Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou 310012, China
[2]School of Oceanography, Shanghai Jiao Tong University, Shanghai 200240, China
[3]College of Geomatics, Shandong University of Science and Technology, Qingdao 266590, China

Corresponding authors: Xiaowen Luo (cdslxw@163.com) and Ziyin Wu (zywu@vip.163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 41830540 and Grant 41676037, in part by the Scientific Research Fund of the Second Institute of Oceanography, MNR, under Grant JZ1902, and in part by the Project of State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, under Grant SOEDZZ1802.

ABSTRACT Seabed acoustic images are image data mosaics derived from seafloor acoustic backscattering intensity data, which is related to the type of sediment covering the seabed. Therefore, submarine sediment classification can be realized using seabed acoustic images, and has been studied extensively. Recently, deep learning has also rapidly advanced; in particular, deep convolutional neural networks (CNNs) are now being used to achieve remarkable results in the field of image processing—showing that they are well-suited for image classification tasks. Previous studies have used GoogleNet to classify large-scale side-scan sonar data, with some sediments being well-classified. However, deep learning is data-driven and, theoretically, the greater the depth, the stronger is the learning ability of the feature. It is worth noting that the dataset used for sediment classification can sometimes be small. Hitherto, no related research has analyzed the feasibility and applicability of a CNN classifier for a small-sized seabed acoustic image dataset, so we adopted two different CNN classifier models to conduct the classification experiment in this study. As the results show, the CNN classifier can be applied to the classification of sediments based on a small-sized seabed acoustic image dataset, and the classification performance of shallow CNN was found to be better than that of the deep CNN on existing side-scan sonar data. In particular, the accuracy obtained from the results of several sediment classification experiments using a shallow CNN classifier ranged between 93.4% (Sand Wave) and 87.54% (Reef).

INDEX TERMS Deep convolutional neural network, seabed acoustic image, seabed sediment classification.

I. INTRODUCTION

The characteristics of marine sediment types are important environmental information used in the fields of Marine Geological Surveying, Marine Engineering Construction, and Seabed Mineral Resources Development. The traditional means of sediment surveying by direct sampling cannot be used on a large scale, as it is a cumbersome, complex, and time-consuming process. An acoustic surveying method for seabed sediment classification is more feasible [1]. The principal method of acoustic sediment classification uses acoustic information derived from acoustic reflection and scattering data, with a defined number of collected samples or visual observations, combined with the appropriate classification algorithms to determine the sediment type. This method is more efficient than traditional surveys and has a wider application range, which is of great significance for Marine Geological Surveying and Marine Engineering Construction.

Numerous studies have been conducted on sediment classification of seabed acoustic data. Lucieer and Lucieer [2] used Fuzzy C-means (FCM) and fuzzy maximum likelihood estimation (FMLE) to classify sediments based on the Seascapes datasets. Samsudin and Hansn [3] used K-means clustering to classify sediments, and Ahmed *et al.* [4] used K-means clustering, quality threshold clustering (QTC), and spectral clustering (SC) to perform sediment classification based

The associate editor coordinating the review of this manuscript and approving it for publication was Alessandro Pozzebon.

on multi-beam echosounder data. Snellen *et al.* [5] used Bayesian techniques and K-means clustering to classify sediments based on multi-beam echosounder data. Ma *et al.* [6] applied the ISODATA clustering algorithm to the classification of sediments. Atallah *et al.* [7] used wavelet analysis to process side-scan sonar data and employed the multilayer perceptron (MLP) and K-nearest neighbors (KNN) algorithms to classify the sediments. Preston [8] used simulated annealing (SA) algorithms for sediment classification based on multi-beam data. Wang *et al.* [9] compared the performances of different support vector machine (SVM) models on sediment classification tasks based on multi-beam and LiDAR data. Stewart *et al.* [10] combined side-scan sonar imagery with Back Propagation neural networks (BPNNs) to classify sediments.

Seafloor backscattering intensity is affected by the type of sediment covering the seafloor; thus, seabed acoustic images generated using this method can be utilized to classify seabed sediment type [11]. A deep convolution neural network (CNN) is a deep learning technique that learns features and performs well in image recognition, unmanned driving, and other fields. CNNs have many applications in remote sensing, such as using CNN combined with remote sensing image data for geospatial object detection [12], ship detection [13], and identification of maize leaf diseases [14]. In addition, a CNN can also be applied for remote sensing image scene classification [15]–[19], remote sensing image registration [20], and road extraction from a high spatial resolution remote sensing image [21]. Related work in the field of remote sensing images undoubtedly shows the advantages of using CNN in image processing, and CNNs are also applied in the processing of seabed image data [22]–[25]. Deep CNN (specifically, GoogleNet) has been used to classify large scale side-scan sonar acoustic image samples with segmentation sizes of 320 × 320, with good classification results achieved for some sediment types and a sand classification accuracy of 83% [25].

It is worth noting that different application scenarios put forward different requirements for the application of CNN's structure. However, in some sediment classification tasks, no large seabed acoustic image dataset is available. There is no relevant research to analyze CNN classifiers when using a small-sized seabed acoustic image dataset, so we have no way of knowing the feasibility and applicability of this method, which is the core of our research. In theory, the depth of a CNN determines its feature learning ability, but its performance varies significantly with different datasets in practical applications. The selection of the CNN classifier structure should be based on the characteristics of the dataset used, and repeated modifications and debugging should be conducted to obtain the best classifier model. A deeper structure is not better for every task. Therefore, in our experiment, we constructed two CNN classifiers to classify small size seabed acoustic image dataset, with one being a shallow CNN modified based on LeNet-5 and the other being a deep CNN modified based on AlexNet. As the results show, CNN can be applied to the sediment classification task for a small-sized seabed acoustic image dataset, and the performance of the two CNN classifier is difference in this case. It can be seen from the experiment that in such application scenarios (dataset size is small and single sample size is small), the shallow CNN performs significantly better. Although deep CNN has superior feature learning ability and requires few training epochs to achieve the objective, it is not applicable in this case. In this study, we utilized existing data and obtained good sediment classification results with shallow CNN, with an accuracy better than 87.5% for seabed sediment classification, compared to 66.75% when using deep CNN.

II. PRINCIPLE

The principle underlying the use of a CNN to achieve sediment classification from seabed acoustic images is based on two factors: (1) seabed acoustic images reflect seabed sediment type information, and (2) a CNN facilitates effective feature extraction and learning processes.

A. SEABED ACOUSTIC IMAGES AND THEIR CONNECTION TO SEDIMENT TYPES

In previous studies, the correlation between seabed backscattering intensity and sediment type was investigated, and several general theories—such as the relationship between the acoustic characteristics and physical properties of sediments [26]–[28], the relationship between the geoacoustic properties and input parameters necessary for sound wave propagation theory of seabed sediments [29]–[31], and the relationship between submarine surface roughness and high frequency acoustic scattering [32]—were developed. Therefore, classifying seabed sediments based on acoustic images is theoretically feasible.

After processing the multi-beam echo intensity data and calculating the latitude and longitude of the beam sampling points, the positions of the backscattering intensity values within the geographical coordinate framework can be determined. After meshing, striping, and mosaicing, a backscattering intensity graph (i.e. a seabed acoustic image) is obtained.

As shown in Fig. 1, seafloor backscattering intensity is defined as the ratio of the intensity of the sound scattered

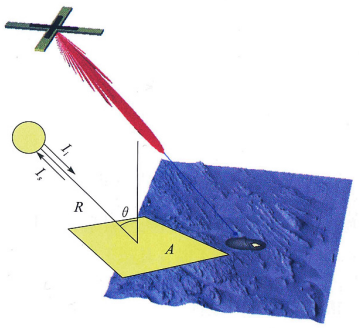


FIGURE 1. Schematic diagram of the definition of seafloor backscattering intensity.

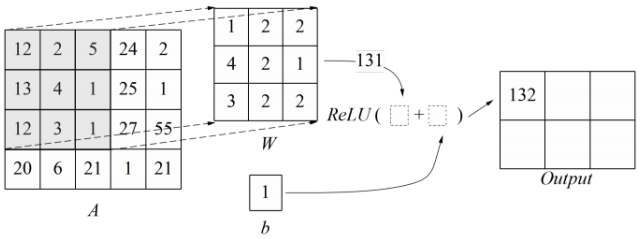


FIGURE 2. Principle of single convolution kernel calculation, where ***A*** is the input data, ***W*** is the convolution kernel, ***b*** is the offset, ***Output*** is the output data, and the shaded area is the field of view. Here, the traversal mode is 1 pixel in the horizontal direction and 1 pixel in the vertical direction; thus, it can be inferred that the output size is 2 × 3.

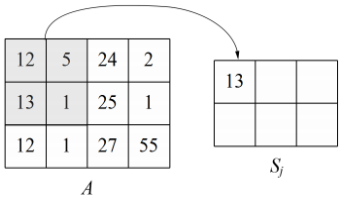


FIGURE 3. Calculation principle of the pooling layer: ***A*** is input data, ***Sj*** is output data, the shaded area is the pooled area, and the pooled area traversal mode is the same as in Fig 1. The size of ***Sj*** can be increased to 2 × 3.

per unit area or volume to the intensity of the incident plane wave (at a reference distance of 1 m) in decibels (dB). The intensity of the sound scattered is measured in the far field and reduced to the unit of distance. Given an acoustic intensity emitted by a sound source of I_i, propagation distance R, submarine incident angle θ, backscattering coefficient $s(\theta)$, and instantaneous acoustic region area A, the accepted acoustic intensity I_S can be expressed as:

$$I_s = I_i \frac{As(\theta)}{R^2} \tag{1}$$

The resulting seafloor backscatter coefficients are then calculated from far field measurements:

$$s(\theta) = \frac{I_S R^2}{I_i A} \tag{2}$$

Expressed in decibels (dB), the seabed backscattering intensity is:

$$BS_B(\theta) = 10 \log_{10}[s(\theta)] \tag{3}$$

B. PRINCIPLES OF A CNN

In [33], the core idea of using CNNs to extract features from the input data through the convolution kernel and the pooling layer is shown. After the convolution layer and pooling layer are processed, a fully connected layer can be terminated for classification calculation.

A typical CNN classifier comprises of convolutional layers, pooled layers, and fully connected layers. CNNs also include feedforward propagation and back propagation. Considering the relative simplicity of the fully connected layer, this study focused on the principle of feature extraction (namely, the convolutional and pooling layers) by formula.

1) FEEDFORWARD PROPAGATION OF A CNN

The feedforward algorithm of a CNN includes convolution operations, pooling operations, and fully connected layers for classification calculations. The role of the convolution layers is to extract features from data, which can be calculated as follows:

$$C_i = conv2d\left(A, W^i, ' Padding'\right) + b^i \tag{4}$$

where $conv2d\ (\cdot)$ is a formula expression of convolving a two-dimensional image, and C_i is the mapping of the convolution kernel W^i onto the data A. The field of view and the moving step size of the kernel W^i must be set in advance, b^i is the offset parameter, and '*Padding* ' is the fill method used when traversing to the edge of the image. Then the C_i is transformed into a nonlinear function by an excitation function:

$$output_i = \sigma\ (C_i) \tag{5}$$

where $\sigma(\cdot)$ is the excitation function. In general, the excitation function in the CNNs uses the Rectified Linear Unit (*ReLU*) function. If we assume the input data are given as x, the function is thus:

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

The output of the convolution layers can be obtained in accordance with Equations (5) and (6):

$$output_i = Relu\ (C_i) \tag{7}$$

The convolution kernel acts as a filter to extract features from the input data. The process of extracting these features is called feature mapping.

Pooling operations are much simpler than convolution operations. Pooling methods include maximum pooling and mean pooling, and so on. In CNNs, maximum pooling is commonly used [33], with the objective of outputting the maximum value in the pooling window. Traversal of the pooling window is very similar to that of the convolution kernel. The supposed $maxpool\ (\cdot)$ is the public expression of the maximum pooling, A is the input data, and S_j is the pooled map. The equation is as follows:

$$S_j = maxpool\left(A, ' Padding'\right) \tag{8}$$

The pooling layer can reduce the data size without losing the features expressed by the data, thereby increasing the speed of the operation while ensuring feature acquisition.

After a series of convolution and pooling operations, some of the feature parameters can be extracted from the input data. These feature parameters are transformed into one dimension and input into the full connection layer, which is essentially a BPNN classifier. The specific principle behind this method is beyond the scope of this paper (see [34] for a detailed introduction).

2) ERROR BACKPROPAGATION OF A CNN

In general, a CNN is trained based on an error backpropagation algorithm. The principle of backpropagation error can be understood as the optimization and adjustment of each parameter using the gradient descent method, according to the output mean square error (MSE). This makes the real output approximate to the ideal output and continuously improves the performance of the classifier. The process of error backpropagation can be divided into two parts: error propagation and parameter update.

a: ERROR PROPAGATION

Error propagation in the convolutional layer and the pooled layer are not the same. When the sensitivity (or gradient) of the current layer (layer l) is δ^l, the S^{l-1} is the feature map of the prior layer, and the error back propagation in the pooling layer is expressed as:

$$\delta^{l-1} = \frac{\partial J(W,b)}{\partial S^{l-1}} = upsample\left(\delta^l\right) \tag{9}$$

where $J(W, b)$ is the MSE of the training process, δ^{l-1} is the sensitivity of the layer before the current layer. Here, the *upsample* $(\cdot)$ process completes the logic of pooled error amplification and error redistribution; that is, it establishes the corresponding relationship between the sensitivity and the feature map before pooling.

In the convolution layer the error propagation is more complicated. Because the input data have been convolved, the convolution operation should be taken into account during the error propagation process. This relationship is as follows:

$$\delta^{l-1} = \frac{\partial J(W,b)}{\partial C^{l-1}} = \frac{\partial J(W,b)}{\partial C^l}\frac{\partial C^l}{\partial C^{l-1}} = \delta^l \frac{\partial C^l}{\partial C^{l-1}} \tag{10}$$

In (10), C^l is the feature map of the current layer, and C^{l-1} is the feature map before the current layer. According to the derivation, this can be written in another form:

$$\delta^{l-1} = \delta^l * rot180\left(W^l\right) \odot \sigma'\left(C^{l-1}\right) \tag{11}$$

where $rot180\,(\cdot)$ is the operation of flipping the matrix, $\sigma(\cdot)$ is the active function and $\sigma'(\cdot)$ is the derivative calculation of the excitation function. This process redistributes the current sensitivity to obtain the sensitivity δ^{l-1} before the convolution operation.

b: PARAMETER UPDATE

Parameter updating occurs in the convolution operation. The convolution kernel parameter and the offset are updated using the gradient descent method, according to the current gradient (sensitivity), which reduces the output mean square error (MSE). The CNN classifier then has the ideal output. The current layer sensitivity is denoted as δ^l, the layer convolution kernel parameter W^l, and the offset b^l are updated. First, we calculate the gradient value:

$$\begin{aligned}\frac{\partial J(W,b)}{\partial W^l} &= \frac{\partial J(W,b)}{\partial C^l}\frac{\partial C^l}{\partial W^l} = a^{l-1}\delta^l \\ \frac{\partial J(W,b)}{\partial b^l} &= \sum_{u,v}\left(\delta^l\right)_{u,v}\end{aligned} \tag{12}$$

According to the obtained gradient value, W^l and b^l are updated on the basis of the given learning rate η; and the relationship can be expressed as follows:

$$\begin{aligned}W^l_{new} &= W^l - \eta\frac{\partial J(W,b)}{\partial W^l} \\ b^l_{new} &= b^l - \eta\frac{\partial J(W,b)}{\partial b^l}\end{aligned} \tag{13}$$

The combination of feedforward propagation and backpropagation of CNN training on the input training samples is such that the output of the classifier is continuously optimized until a terminal condition is reached, resulting in a classifier with good performance.

III. CLASSIFICATION EXPERIMENT

The data used in this study are from side-scan sonar images obtained in the Pearl River Estuary as part of the "China Offshore and Ocean Comprehensive Survey and Evaluation Special." A DF1000/560D digital dual-frequency side-scan sonar system, differential GPS positioning system, SDH-13D depth sounder, CAP-6600 shallow profiler system, professional processing software C-View 1.52, and Sonar Wiz Map were used. The micro-geomorphology of the Pearl River Estuary is diverse, with various types of sediment cover. The side-scan sonar's operational frequency is 500 KHz, and the A/D resolution of the side-scan sonar system is 12 bits/sample. The width and height of a pixel in the seabed acoustic images correspond to 0.2 m each.

The seabed acoustic imagery was divided into multiple segmented small-scale images, which were used for classification. The image data used in this experiment are shown in Fig. 4, comprising three main sediment types: reef, mud, and sand wave. The type of each sediment was determined by direct sampling analysis. The acoustic images of each sediment type have peculiar features, which are the basis used for the classifier, in theory. In this study, the processing method used for the original image data was the division of the original picture into a plurality of moments, as shown in Fig. 5.

Data partitioning utilized the hold-out method. The number of classified reef, mud, and sand wave images were 234, 217, and 228, respectively. From each sediment type, 20% of the images were randomly selected as test samples: 46, 43, and 45 images of reef, mud, and sand wave, respectively. In order to eliminate any influence of sample partitioning on the performance of the classifier, the hold-out method was used to randomly generate 10 datasets to verify the classifier.

Considering that there is no existing research that analyzed the application of a CNN classifier on a small-sized seabed acoustic image dataset, we used two types of CNN in this

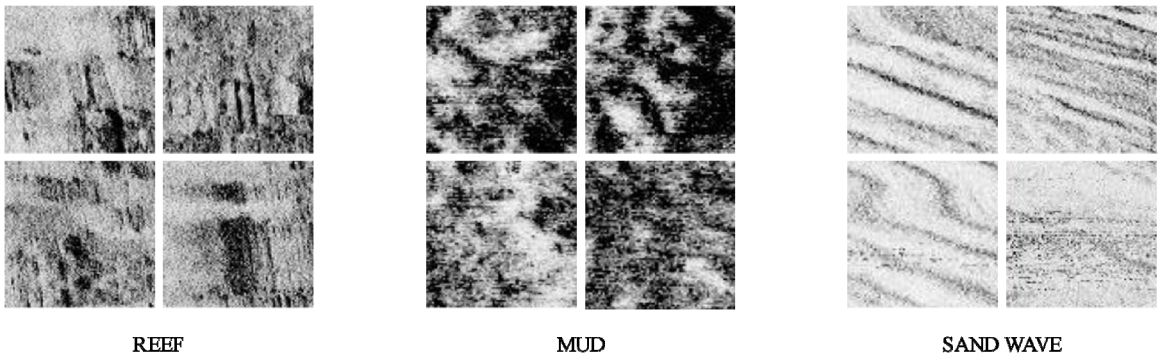


FIGURE 4. Seabed acoustic images of the three sediment types in the dataset.

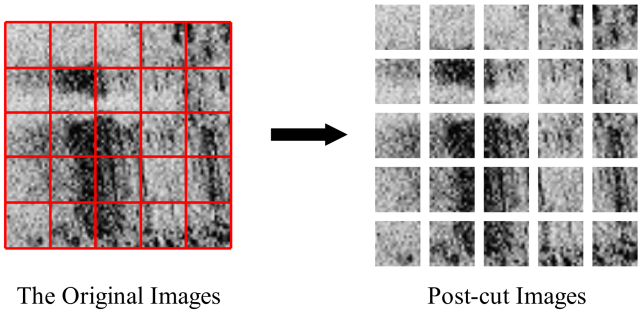


FIGURE 5. Example of the image cutting process.

TABLE 1. Number of kernels (or neurons) in each layer.

Deep CNN		Shallow CNN	
C1	64	C1	32
C3	128	C3	64
C5	256	FC1	1024
C6	256		
C7	128		
FC1	4096		
FC2	4096		

experiment: deep CNN based on AlexNet and shallow CNN based on LeNet-5. As shown in Fig. 6, the structure of the deep CNN is relatively complicated, based on modifications of AlexNet. These mainly involved removing the LRN layer and remotely modifying the structure to fit small-size image data. In contrast, as shown in Fig. 7, the structure of the shallow CNN is much simpler and is based on modifications of LeNet-5. The complexity and depth of the two CNN classifiers are quite different. The deep CNN consists of five convolutional layers, three pooling layers, and two fully connected layers, whereas the shallow CNN contains only two convolutional layers, two pooling layers, and one fully connected layer.

The convolution kernel size of the deep CNN was 5×5 pixels and the pooled area was 3×3 pixels. The traversal mode of the convolution kernel was 1 pixel in the horizontal direction and 1 pixel in the vertical direction (abbreviated as (1,1)), and the traversal mode of the pooled region was (2, 2). The convolution kernel size in the shallow CNN was 3×3 and the pooled area was 2×2. The convolution kernel traversal mode was (1,1) and the pooled area traversal mode was (2,2). Both types of CNN classifiers used the same padding model at the edges of the image. The number of convolutional kernels (or neurons) in each layer is shown in Table 1.

In this experiment, the 10 datasets outlined previously were utilized. In order to minimize the impact of randomly generated initialization values and to evaluate the real performance of the classifier as best as possible, repeated tests (set as 10 times) were conducted on each dataset and the mean value was taken as the result of the classification on this dataset. In addition, in order to adapt to the subsamples in the CNN, the original 16×16 pixel sample data were expanded to 32×32 pixels. Considering that the number of samples is small, all training samples are divided into one batch during the training process to identify the global optimum in the batch gradient descent (BGD) process. Finally, in order to avoid random initialization of the neural network parameters, the average of the results of 10 datasets was taken as the final result. The respective final performances of the deep CNN and shallow CNN classifiers are shown in Table 2.

TABLE 2. Performance of the two CNN classifiers.

Sediment	Deep CNN		Shallow CNN	
	Training Samples	Testing Samples	Training Samples	Testing Samples
Reef	100%	66.75%	100%	87.54%
Mud	100%	73.8%	100%	90.56%
Sand Wave	100%	79.58%	100%	93.43%

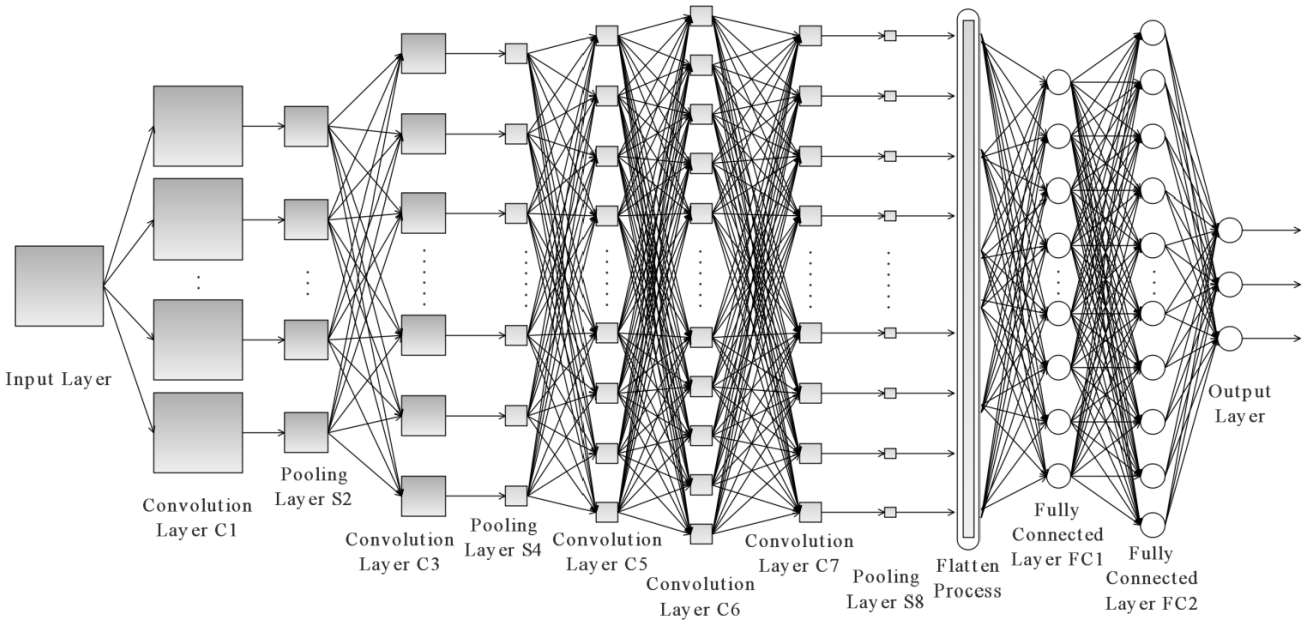


FIGURE 6. Structure of the deep CNN (based on AlexNet).

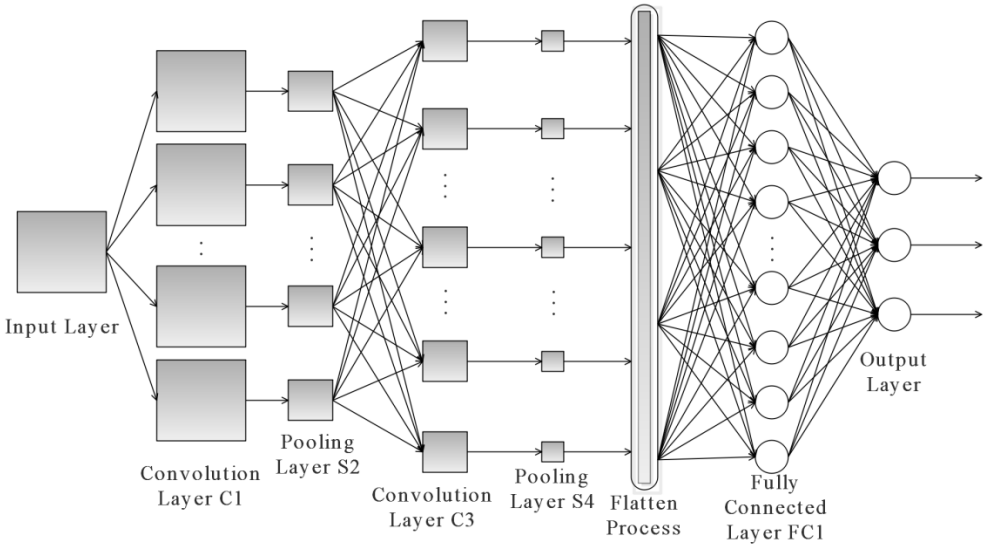


FIGURE 7. Structure of the shallow CNN (based on LeNet-5).

As shown in Table 2, the performance of both classifiers on the training samples were 100%, but the performance of the shallow CNN on the test samples was much better than that of the deep CNN. Considering that the accuracy of the two classifiers on the training samples is 100%, they are not comparable, and to more directly represent the difference between the two classifiers, their performances on the test samples are intuitively compared in Fig 8. The classification accuracy of the shallow CNN is a minimum of 87.5% (Reef) and a maximum of 93.43% (Sand Wave). The classification accuracy of the deep CNN is a minimum of 66.75% (Reef) and a maximum of 79.58% (Sand Wave).

In addition, there are significant differences between the calculation speeds of the two classifiers, as shown in Table 3. The program was written in Tensorflow1.12.0 using Python and executed on a GPU with GTX780M. In order to reflect the calculation speed, the average value was taken as the speed after running 10 times. The calculation speed of the shallow CNN was much higher than that of the deep CNN, both for the training data and for the testing data. It is worth noting that the training epoch of the shallow CNN was 500, whereas that of the deep CNN was 300, which further reflects the difference in the speed of the two classifiers, indicating that the computational power resources consumed by the deep

TABLE 3. Runtime of the two CNN classifiers.

Classifier \ Data	Training Data	Testing Data
Shallow CNN	176.8 (s)	0.23291 (s)
Deep CNN	327.6 (s)	1.07436 (s)

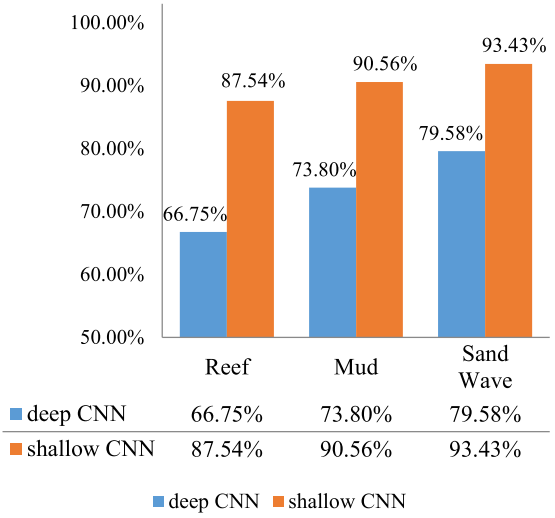


FIGURE 8. Classification results of the two CNN classifiers on the testing samples.

CNN far exceeded those consumed by the shallow CNN. The runtime of the deep CNN on the testing samples is approximately five times that of the shallow CNN, which indicate that the structural complexity and the parameter amount of the deep CNN far exceed those of the shallow CNN, which directly reflects the greatly increased computational power required by the deep CNN over the shallow CNN.

Although the deep CNN has obvious disadvantages compared to the shallow CNN in terms of classification performance for testing samples and computing speed, in the training process, the deep CNN is much faster than the shallow CNN for the epochs required to converge to a classification accuracy of 100%, as shown in Fig. 9. The difference in convergence speed can also indicate that the deep CNN is better than the shallow CNN in terms of learning ability, because theoretically, the greater the depth, the stronger the expression ability of neural networks.

However, it also seems possible for the deep CNN to learn too many features during the learning process, and some of these features seems to be non-universal, which leads to the weak generalization ability of the deep CNN (or overfitting); this may explain its poor performance on testing samples. Summarizing the above and comparing the actual performances of the two classifiers, the reasons for the weak performance of the deep CNN compared to the

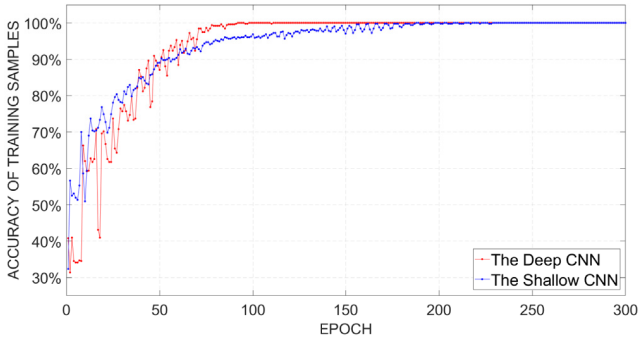


FIGURE 9. The accuracy of the two CNN classifiers on training samples.

shallow CNN may be as follows: (1) the stronger learning ability of the deep CNN leads to overfitting problems and a weaker generalization ability (even after adoption of the dropout method to eliminate overfitting as much as possible); (2) overly small final output feature map sizes resulted in poor representation; (3) insufficient number of training samples.

It is clear that the CNNs can be applied to sediment classification based on a small-sized seabed acoustic image dataset, and the classification effect of the shallow CNN classifier was superior to that of the deep CNN classifier on the small seabed acoustic image dataset. However, this is only in the case of a small dataset with few easily distinguishable types. If the dataset is expanded or if the sample content diversity and complexity are increased, it is difficult to determine whether the deep CNN or the shallow CNN are performing better, which is also the subject of our subsequent research. In addition, the performance of deep learning is closely related to its structure and training tricks; thus, we will test more CNN structures and training methods in the follow-up work, to better apply deep learning to the processing of seabed acoustic images.

IV. CONCLUSION

This study verifies that a CNN classifier can be applied to a small-sized seabed acoustic image dataset, and through a comparison of the experimental results, it is discovered that CNN with shallow depth seems to be more effective than deep CNN in this case. More specifically, the classification accuracy and calculation speed of a deep CNN classifier and a shallow CNN classifier were compared using an existing small-size dataset and the shallow CNN classifier is found to be more appropriate for such datasets. Theoretically, a deep CNN has better feature learning ability; however, for the small-sized seabed acoustic image dataset in this experiment, the shallow CNN outperformed the deep CNN in terms of classification accuracy and speed.

Deep learning is an excellent learning algorithm with a wide range of applications and impressive results for a variety of tasks, such as classification, object detection, and natural language processing. Through this experiment, we conclude that CNNs are feasible for a small-sized seabed acoustic

image dataset and that shallow CNN is more feasible in this application scenario. Considering that deep learning theory and its applications are still developing rapidly, we will continue to attempt to apply more models and training skills to the sediment classification task in the following work. Moreover, we will continue to expand the dataset and enrich the application scenarios for experiments to consolidate its theoretical feasibility. Although deep learning has problems such as randomness, dataset requirements, and being computationally intensive, it is undeniable that it provides excellent results in this sediment classification experiment. In addition to sediment classification, it can also be applied to other tasks in marine science, such as submarine target detection and surface remote sensing image processing. Therefore, deep learning still has potential applications in the field of marine science.

REFERENCES

[1] Q. Tang, B.-H. Liu, Y.-Q. Chen, X.-H. Zhou, and J.-S. Ding, "Application of LVQ neural network combined with the genetic algorithm in acoustic seafloor classification," (in Chinese), *Chin. J. Geophys.*, vol. 50, no. 1, pp. 291–298, Jan. 2007.
[2] V. Lucieer and A. Lucieer, "Fuzzy clustering for seafloor classification," *Mar. Geol.*, vol. 264, nos. 3–4, pp. 230–241, Aug. 2009.
[3] S. A. Samsudin and R. C. Hasan, "Assessment of multibeam backscatter texture analysis for seafloor sediment classification," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 42, pp. 177–183, Oct. 2017.
[4] K. I. Ahmed, H. Caughey, P. Hung, U. Demšar, S. Mcloone, A. S. Fotheringham, X. Monteys, R. O'Toole, and P. Harris, "Classification and mapping of seabed type from deep water multibeam echosounder (MBES) data," presented at the 6th Int. Conf. Geograph. Inf. Sci., Zürich, Switzerland, 2010.
[5] M. Snellen, T. C. Gaida, L. Koop, E. Alevizos, and D. G. Simons, "Performance of multibeam echosounder backscatter-based classification for monitoring sediment distributions using multitemporal large-scale ocean data sets," *IEEE J. Ocean. Eng.*, vol. 44, no. 1, pp. 142–155, Jan. 2019.
[6] F. Ma, E. Dongchen, J. Zhao, and X. Zhou, "Seabed classification based on ISODATA algorithm," *J. Geomatics*, vol. 33, no. 6, pp. 43–45, 2008.
[7] L. Atallah, P. J. P. Smith, and C. R. Bates, "Wavelet analysis of bathymetric sidescan sonar data for the classification of seafloor sediments in Hopvågen bay—Norway," *Mar. Geophys. Res.*, vol. 23, nos. 5–6, pp. 431–442, 2002.
[8] J. Preston, "Automated acoustic seabed classification of multibeam images of stanton banks," *Appl. Acoust.*, vol. 70, no. 10, pp. 1277–1287, Oct. 2009.
[9] M. Wang, Z. Wu, F. Yang, Y. Ma, X. H. Wang, and D. Zhao, "Multifeature extraction and seafloor classification combining LiDAR and MBES data around Yuanzhi Island in the South China Sea," *Sensors*, vol. 18, no. 11, p. 3828, Nov. 2018.
[10] W. K. Stewart, M. Jiang, and M. Marra, "A neural network approach to classification of sidescan sonar imagery from a midocean ridge area," *IEEE J. Ocean. Eng.*, vol. 19, no. 2, pp. 214–224, Apr. 1994.
[11] Z. Wu, F. Yang, and X. Luo *High Resolution Submarine Geomorphology*. Beijing, China: Science Press, 2017.
[12] X. Han, Y. Zhong, and L. Zhang, "An efficient and robust integrated geospatial object detection framework for high spatial resolution remote sensing imagery," *Remote Sens.*, vol. 9, no. 7, p. 666, 2017.
[13] R. Zhang, J. Yao, K. Zhang, C. Feng, and J. Zhang, "S-CNN based ship detection from high-resolution remote sensing images," *Int. Arch. Photogram., Remote Sens. Spatial Inf. Sci.*, vols. 41–B7, pp. 423–430, Jul. 2016.
[14] X. Zhang, Y. Qiao, F. Meng, C. Fan, and M. Zhang, "Identification of maize leaf diseases using improved deep convolutional neural networks," *IEEE Access*, vol. 6, pp. 30370–30377, 2018.
[15] F. Hu, G.-S. Xia, J. Hu, and L. Zhang, "Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery," *Remote Sens.*, vol. 7, no. 11, pp. 14680–14707, 2015.
[16] W. Zhou, S. Newsam, C. Li, and Z. Shao, "Learning low dimensional convolutional neural networks for high-resolution remote sensing image retrieval," *Remote Sens.*, vol. 9, no. 5, p. 489, 2017.
[17] N. Liu, L. Wan, Y. Zhang, T. Zhou, H. Huo, and T. Fang, "Exploiting convolutional neural networks with deeply local description for remote sensing image classification," *IEEE Access*, vol. 6, pp. 11215–11228, 2018.
[18] W. Li, H. Liu, Y. Wang, Z. Li, Y. Jia, and G. Gui, "Deep learning-based classification methods for remote sensing images in urban built-up areas," *IEEE Access*, vol. 7, pp. 36274–36284, 2019.
[19] G. Cheng, C. Yang, X. Yao, L. Guo, and J. Han, "When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2811–2821, May 2018.
[20] Z. Yang, T. Dan, and Y. Yang, "Multi-temporal remote sensing image registration using deep convolutional features," *IEEE Access*, vol. 6, pp. 38544–38555, 2018.
[21] Z. Hong, D. Ming, K. Zhou, Y. Guo, and T. Lu, "Road extraction from a high spatial resolution remote sensing image based on richer convolutional features," *IEEE Access*, vol. 6, pp. 46988–47000, 2018.
[22] T. Rimavicius and A. Gelzinis, "A comparison of the deep learning methods for solving seafloor image classification task," *Inform. Softw. Technol.*, vol. 756, pp. 442–453, 2017.
[23] A. Diegues, J. Pinto, P. Ribeiro, R. Frias, and D. C. Alegre, "Automatic habitat mapping using convolutional neural networks," in *Proc. IEEE/OES Auton. Underwater Vehicle Workshop (AUV)*, Nov. 2018, pp. 1–6.
[24] P. Feldens, A. Darr, A. Feldens, and F. Tauber, "Detection of boulders in side scan sonar mosaics by a neural network," *Geosciences*, vol. 9, no. 4, p. 159, 2019.
[25] T. Berthold, A. Leichter, B. Rosenhahn, V. Berkhahn, and J. Valerius, "Seabed sediment classification of side-scan sonar data using convolutional neural networks," presented at the IEEE Symp. Ser. Comput. Intell., Honolulu, HI, USA, 2017.
[26] C. M. Mckinney and C. D. Anderson, "Measurements of backscattering of sound from the ocean bottom," *J. Acoust. Soc. Amer.*, vol. 36, no. 1, pp. 158–163, 1964.
[27] L. Hampton, *Physical of Sound in Marine Sediments*. New York, NY, USA: Plenum Press, 1974.
[28] G. Shumway, "Sound speed and absorption studies of marine sediments by a resonance method," *Geophysics*, vol. 25, no. 2, pp. 451–467, 1960.
[29] D. M. McCann, "Measurement of the acoustic properties of marine sediments," *Acta Acustica United Acustica*, vol. 26, no. 2, pp. 55–66, 1972.
[30] E. L. Hamilton and R. T. Bachman, "Sound velocity and related properties of marine sediments," *J. Acoust. Soc. Amer.*, vol. 72, no. 6, pp. 1891–1904, 1982.
[31] R. T. Bachman, "Estimating velocity ratio in marine sediment," *J. Acoust. Soc. Amer.*, vol. 72, no. 5, pp. 1891–1904, 1989.
[32] D. R. Jackson, D. P. Winebrenner, and A. Ishimaru, "Application of the Composite Roughness Model to High-Frequency Bottom Backscattering," *J. Acoust. Soc. Amer.*, vol. 86, no. 5, pp. 2029 –2032, 1986.
[33] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*. Cambridge, U.K.: MIT Press, 2016.
[34] J. Li, J.-H. Cheng, J.-Y. Shi, and F. Huang, "Brief introduction of back propagation (BP) neural network algorithm and its improvement," in *Advances in Computer Science and Information Engineering*, vol. 2. Berlin, Germany: Springer, 2012.

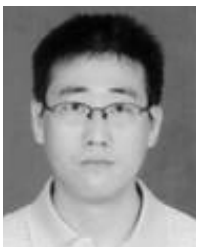
XIAOWEN LUO received the Ph.D. degree in geodesy and geomatics from the Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Hubei, China, in 2007. He is currently an Associate Professor with the Second Institute of Oceanography, Ministry of Natural Resources. His primary research interests include seabed survey and information systems.

XIAOMING QIN received the B.E. degree from the College of Geomatics, Shandong University of Science and Technology, Qingdao, China, in 2018. He is currently with the Key Laboratory of Submarine Geosciences, Second Institute of Oceanography, Ministry of Natural Resources. His current research interests include the earth exploration and information technology.

ZIYIN WU received the Ph.D. degree in marine geology from Zhejiang University, Zhejiang, China, in 2008. He is currently a Professor with the Second Institute of Oceanography, Ministry of Natural Resources and School of Oceanography, Shanghai Jiao Tong University. His primary research interests include seabed survey and information systems.

FANLIN YANG received the Ph.D. degree in geodesy and geomatics from Wuhan University, Hubei, China, in 2003. He is currently a Professor with the College of Geomatics, Shandong University of Science and Technology, Qingdao, China. His research areas focus on bathymetry survey and marine navigation.

MINGWEI WANG received the B.S. and M.S. degrees from the Shandong University of Science and Technology, Qingdao, China, in 2014 and 2017, respectively, where he is currently pursuing the Ph.D. degree. His research interests include bathymetry survey and seafloor classification.



JIHONG SHANG received the Ph.D. degree in marine geology from the Institute of Oceanology, Chinese Academy of Sciences, China, in 2008. He is currently an Associate Research with the Second Institute of Oceanography, Ministry of Natural Resources, China. His research interests include submarine topographic survey and morphotectonics study.