Classification of sea and land waveforms based on deep learning for airborne laser bathymetry

Hu Shanjiang[1,2], He Yan[1], Tao Bangyi[3], Yu Jiayong[4], Chen Weibiao[1]

(1. Key Laboratory of Space Laser Communication and Detection Technology, Shanghai Institute of Fine Mechanics and Optics, Chinese Academy of Sciences, Shanghai 201800, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China; 3. Second Institute of Oceanography, MNR, Hangzhou 310012, China; 4. Shandong University of Science and Technology, Qingdao 266590, China)

Abstract: Classification of sea and land returns in airborne lidar was essential for the research of coastal zones and their changing nature. A method for classification using deep learning on the original airborne lidar echo was proposed. A fully connected neural network, and a one-dimensional convolutional neural network (CNN), were used on a training dataset and test datasets from in-situ measurements, and a classification accuracy of 99.6% was obtained. The model was utilized on the datasets from different areas, a classification accuracy of 95.6% was achieved and the processing speed was increased by about 52% compared to support vector machine (SVM) method. The results denote that the deep learning method is very effective for classification of airborne lidar echo waveforms with high precision and speed. It may present further use as a candidate method for classifying species on the sea floor with airborne laser bathymetry.

Key words: bathymetry; lidar; classification; deep learning

CLC number: TN958.98 **Document code:** A **DOI:** 10.3788/IRLA201948.1113004

基于深度学习的机载激光海洋测深海陆波形分类

胡善江[1,2],贺 岩[1],陶邦一[3],俞家勇[4],陈卫标[1]

(1. 中国科学院上海光学精密机械研究所 空间激光传输与探测技术重点实验室,上海 201800;
2. 中国科学院大学,北京 100049; 3. 自然资源部第二海洋研究所,浙江 杭州 310012;
4. 山东科技大学,山东 青岛 266590)

摘 要: 机载激光雷达的海陆波形分类对于沿海地区及其变化性质的研究至关重要。提出了一种在原始的机载激光雷达回波上使用深度学习进行分类的方法。构建全连接神经网络和一维卷积神经网络(CNN),在一个测量海域的数据集上进行训练和测试,最优模型获得了 99.6%的分类精度。该最优模型对来自不同测量海域的数据进行分类,分类精度达到了 95.6%,相比支持向量机方法,处理速度提高了约 52%。结果表明:深度学习方法对机载激光雷达回波波形的分类具有较高的精度和速度,它可以进一步作为通过机载激光测深技术对海底种类进行分类的候选方法。

关键词: 海洋测深; 激光雷达; 分类; 深度学习

收稿日期:2019-03-13; **修订日期**:2019-05-10

基金项目:国家重点研发计划(2016YFC1400902);国家重大科学仪器设备开发专项(2013YQ120343)

作者简介:胡善江(1981-),男,高级工程师,博士生,主要从事激光遥感和仪器开发设计等方面的研究。Email:sjhu@siom.ac.cn

导师简介:陈卫标(1969-),男,研究员,博士,主要从事激光遥感和空间全固态激光器技术研究,重点开展激光遥感、遥测技术在航天、大气和海洋中的应用基础方面的研究。Email:wbchen@mail.shcnc.ac.cn

0 Introduction

Coastal zone topography is of fundamental importance to marine science and economic development. Airborne lidar bathymetry is a powerful tool for examining coastal zones, islands, reefs, and their surrounding seabed topography. Major commercial lidar products include CZMIL [1] from Optech and HawkEye [2] from Leica, and these have already been used for survey applications. The authors′ lab at the Shanghai Institute of Optics and Fine Mechanics (SIOM), Chinese Academy of Sciences, has also been developing airborne laser bathymetry lidar systems, called LADM-I, LADM-II, and Mapper5000.

The return echoes of airborne lidar bathymetry are digitized into full waveforms from the sea surface to sea floor. To determine the sea surface and sea floor respectively, both infrared and green laser are used. The waveform of an airborne ocean lidar return signal is complicated and variable, due to the random rough sea surface, different types of sea water, and mixing with a shallow sea floor. Moreover, the lidar returns from the infrared laser are sometimes lost. In addition, in the case of coastal zones, it is hard to record the lidar echo from either the land or the sea surface. Therefore, an automatic classification method for airborne laser bathymetry is preferred to achieve a high-performance three-dimensional Digital Elevation Model (DEM) of a coastal zone.

Currently, there are several methods of classification[3]. The first one is based on the saturation of the peak intensity of lidar returns owing to the different reflectance of the land and sea surface [4]. While the saturation is variable because of the varying status of the sea surface, the returns can also be saturated at different aircraft altitudes or with different sea waters. The second method is distinguished by another channel of Raman scattering, or the polarization of sea water[5]. This method is effective and simple, yet it nonetheless attaches a higher complexity and cost to the lidar system. The third method is classification of signal waveforms using the SVM method of machine learning[6], which functions by constructing a vector model of features by extracting some characteristic parameters from the echo signal waveforms. Good results were demonstrated with this method, although a great amount of work is involved in the extraction of the characteristic parameters and calculation from huge waveforms; thus, it is time consuming. The fourth method is to classify the terrains by known analytic models of the terrains[7], which is suitable for the case where the object characteristics change little. Therefore, a higher-efficiency, simpler, and faster classification method is desired.

As a branch of machine learning algorithms, deep learning is based on the establishment and simulation of deep nonlinear neural networks for human brain analysis and learning. Deep learning systems can learn a deep nonlinear network structure, characterize input data, implement complex function approximation, and demonstrate a powerful ability to learn the essential features of datasets from a small sample set. In recent years, deep learning algorithms have been rapidly developing, and have become highly effective for object recognition. The method of deep learning is widely used in the fields of language recognition and image recognition. Recently, deep learning has been applied to the field of lidar and used with lidar data in a multitude of areas, such as the modeling of buildings[8], the classification of land cover or objects[9-10], the automatic landing of aircraft[11], and automatic driving target recognition[12-15]. These applications are based on the post-processed point cloud data from the lidars.

In this paper, the deep learning method is implemented to classify land and sea from the original echo signal waveforms from airborne laser bathymetry. An in-situ measurement dataset is used for training and testing. The mapping of the coastal zone is

demonstrated using lidar data processed with the deep learning method.

1 Materials and methods

1.1 Datasets of airborne laser bathymetry

The datasets are acquired by the Mapper5000－S airborne laser bathymetry system developed by SIOM. The lidar system has lasers of two wavelengths, 1 064 nm and 532 nm, with a repetition rate of 1 000 Hz. It can function in conjunction with either a helicopter or a fixed-wing aircraft, at an altitude of 300－1 000 m. From an altitude of 300 m, the swath width is 210 m, with a grid of 2.5 m×2.5 m, and the horizontal and depth accuracy is 0.3 m and 0.2 m, respectively. There are three receiver channels on the Mapper 5000－S, including a sea surface channel, a shallow water channel, and a deep water channel. The avalanche photo diode (APD) is used to detect the echo signals of 1 064 nm. Both photomultiplier tubes (PMT) are applied to detect shallow and deep water with different fields-of-view.

The training dataset and testing datasets are extracted from the flight data of Mapper5000－S, collected at Wuzhizhou Island, Sanya, China, on December 25, 2015. The flight paths are shown in Fig.1(a); the data from flight path 0 is selected as the training dataset, and flight paths 1, 2, and 3 are selected as testing datasets and are named as testing dataset No.1, No.2, and No.3, respectively. Through satellite imagery, hydrological data, manual assisted surveys, and GPS positioning data, all the echo data are labeled as "Land" or "Sea".

The flight area for the applied demonstration is selected as shown in Fig.1(b), and denoted *Y*. This area is about 42 km from the training and testing area, denoted *X*. Although the lidar data was recorded on the same day as the training and testing data, the terrain and sea water quality are quite different from the training and testing area. The sea water around *Y* is clearer than that in the vicinity of *X*.

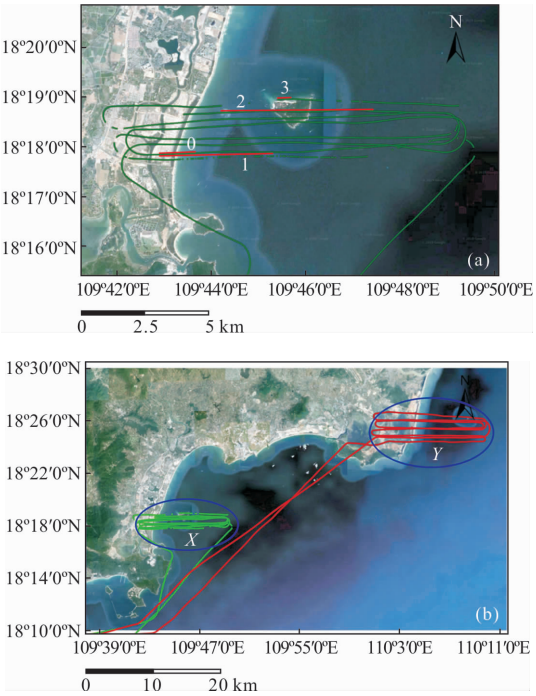


Fig.1 Flight area and flight paths for (a) the basis of the training and test datasets and (b) the training area and application area

Because of the variation in terrain and sea water quality, the echo signal waveforms in the different areas are distinct. Sample echo signal waveforms from one measurement point of the lidar are shown in Fig.2. Figs.2 (a) and (b) show echo signals from shallow water and land in the training and testing area, and Figs.2(c) and (d) show echo signals from the shallow water and land in the application area. It can be seen that the surface reflecting echoes from land and sea have particular characteristics, and the characteristics of the echoes at each measurement point are different in the datasets.

The echo signal from each measurement point consists of three waveforms on three channels; an echo waveform has 5 500 digitized sample points, and each digitized sample point consists of 2 bytes, thus the total data length of one echo signal is 3×5 500×2 bytes. The training dataset contains 23 411 echo signals; test datasets No.1, No.2, and No.3 contain 82 651, 11 165, and 18 716, respectively. We implemented 2 851 448 echo signals for the applied demonstration. As shown in Fig.2, only a limited amount of data from each returned waveform is taken as the effective

signal sampling for training. The length of the effective signal is 100 digitized sampling points, and its center point is approximately the peak of the return echo. The amount of storing and calculating is thereby reduced greatly, so the deep training can be run on either the GPU or CPU of a personal computer.

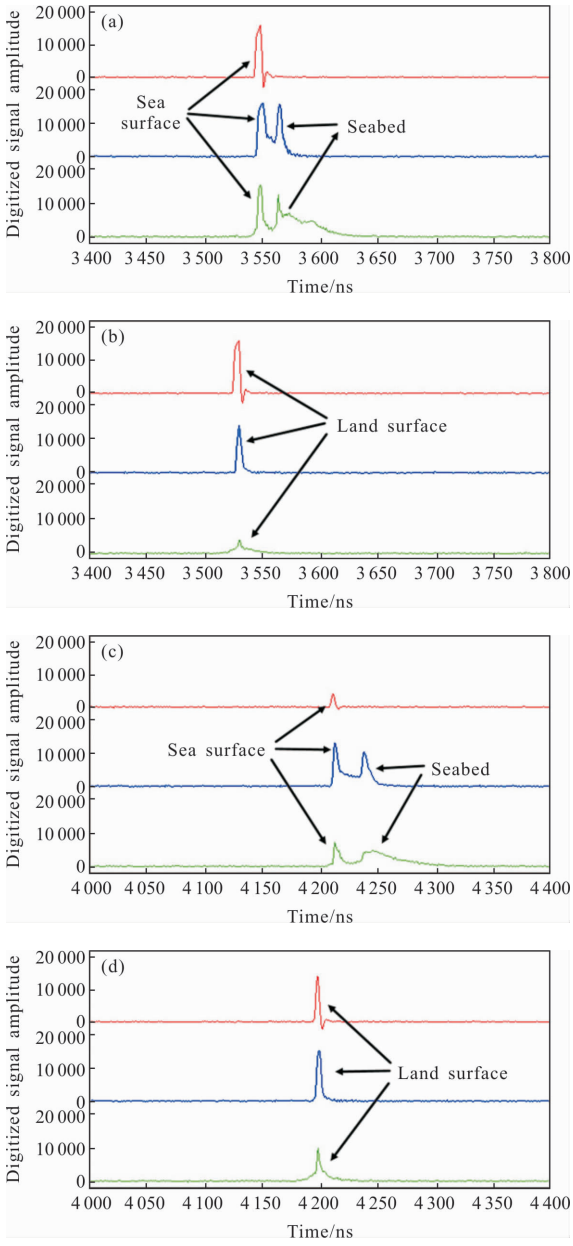


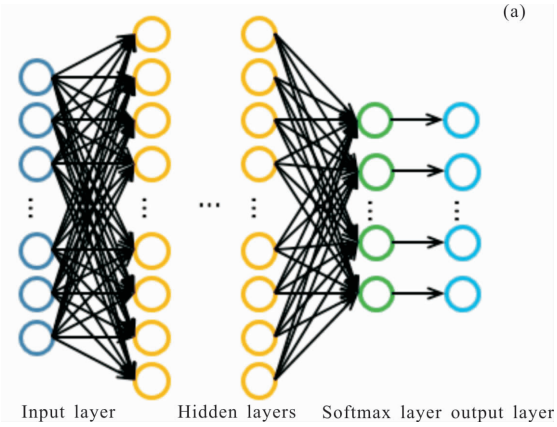
Fig.2 Sample echo signal waveforms in the training and test area, and application area: (a) echo signal from shallow water in the training and test area, (b) echo signal from land in the training and test area, (c) echo signal from shallow water in the application area, and (d) echo signal from land in the application area

1.2 Methodology

Deep learning is used with Google´s deep learning framework Tensorflow, which is an open source software library for high performance numerical calculations, and provides powerful support for machine learning and deep learning. Because of its flexible architecture, users can easily deploy computing work across multiple platforms (CPU, GPU, TPU) and devices(desktops, server clusters, mobile devices, edge devices, etc.)(https://www.tensorflow.org).

Two network models of a multi-layer fully connected neural network and a one-dimensional convolutional neural network (1D CNN) are chosen, which are commonly used in deep learning for waveform classification. In a multi-layer fully connected neural network, all neurons in the adjacent two-layer neural network are connected, whereas in a CNN, the neurons are partially connected between layers. Their structures are shown in Fig.3.

In Fig.3 (a), the first layer is an input layer with 3×100 neurons, the second layer is a hidden layer with 1 024 neurons, the third layer is a hidden layer with 128 neurons, and the fourth layer is a softmax layer with two neurons-this returns an array of two probability scores that sum to 1. Each neuron contains a score that indicates the probability that the current signal waveform belongs to one of the two classes. The dropout layers between hidden layers are generally not counted in the total number of layers and are not specifically included in the diagram. This network has 439 682 parameters to be trained.



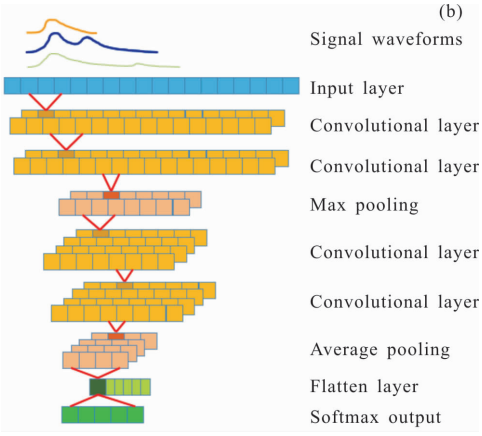


Fig.3 Network models structures of: (a) a multi-layer fully connected neural network and (b) a one-dimensional convolutional neural network

The structure of a CNN generally uses the classical, widely used LeNet-5 [16] structure or its variant, but this structure is generally applicable to the field of image processing and is a two-dimensional CNN. The data sequence changes with time, and needs to adopt a one-dimensional CNN [17-18]. We designed a one-dimensional CNN structure as shown in Fig.3 (b). The first layer is an input layer with 3×100 neurons; the second and third layers are convolutional layers with 64 filters and a 3×1 kernel size; the fourth layer is a max pooling layer that filters out half of the features by subsampling; the fifth and sixth layers are convolutional layers with 128 filters and a 3×1 kernel size; the seventh layer is an average pooling layer; the eighth layer is a flatten layer with 256 neurons; and the last layer is a softmax layer with two neurons. This 1D CNN has 120 514 parameters to be trained.

The standard deep learning training process is one of cyclical training, evaluation, and parameter tuning. The process involves training the model with the training dataset, testing the model with the testing data sets, analyzing the results, then looping the above process and visually evaluating the training effect through the Tensorboard module in Tensorflow.

We experiment with different gradient descent optimizers, learning rates, activation functions, and loss functions to improve the learning effect. Training epochs, the number of layers, and number of neurons in the network model, for comparison accuracy and performance are adjusted by each other. After more than 50 loops of training and testing with varying configurations of models, an optimal model configuration is chosen. Subsequently, the optimal model configuration is used to train 200 models on the training dataset, and the best model is chosen based on the accuracy of the evaluation on the test datasets. Finally, the best model is used to predict the classification result on the applied region dataset, and the prediction accuracy through satellite imagery, hydrological data, manual assisted surveys, and GPS positioning data is also calculated. The comparison of the calculation performance with a CPU and GPU is also conducted.

2 Results

Because the deep learning algorithm has a certain randomness, the accuracy of the model for each episode of training is also random. The accuracy of the best model of the 200 trained models on each test set is shown in Tab.1.

Tab.1 Accuracy on test datasets

Model structure	Accuracy on test set No.1	Accuracy on test set No.2	Accuracy on test set No.3
Fully connected	98.78%	97.90%	98.02%
1D CNN	99.65%	99.33%	98.64%

The results indicate that the accuracy of the 1D CNN structure network is greater than the accuracy of the fully connected structure network, although the 1D CNN structure network has a smaller number of parameters. The quantity weighted average accuracy of 1D CNN is 99.45%, and that of the fully connected structure network is 98.56%.

In the training of a neural network, one epoch refers to one pass of the full training set; one training epoch contains many training batches. The trends for

accuracy and loss are shown in Fig.4. Figs.4 (a) and (c) indicate that in batch training, the accuracy and loss of each batch tend to converge. Figs.4(b) and (d) indicate that after ten epochs, the accuracy and loss level off at favorably high and low values, respectively.

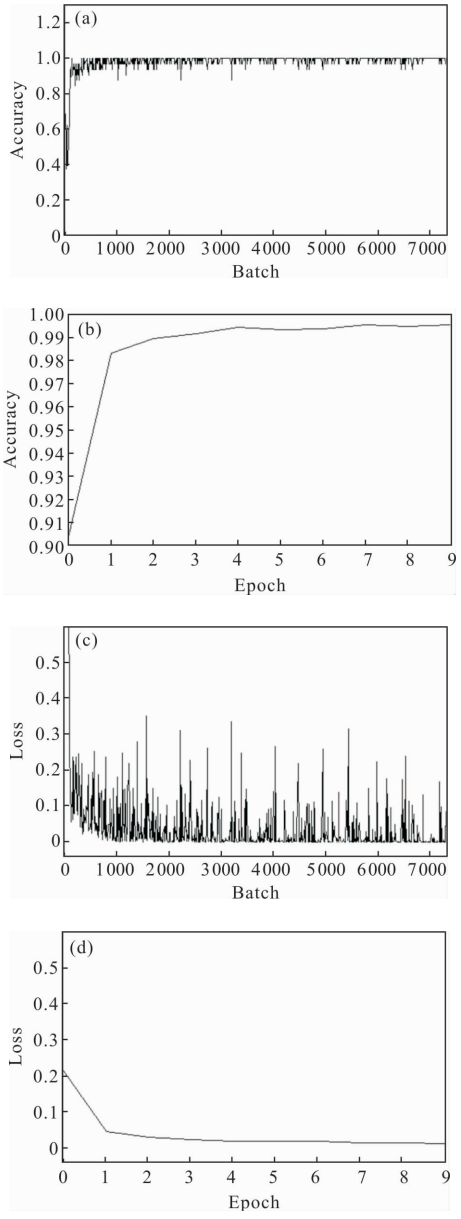


Fig.4 Trends of accuracy and loss in training: (a) accuracy vs batch; (b) accuracy vs epoch; (c) loss vs batch; and (d) loss vs epoch

The learning rate, gradient descent optimizer, activation function, loss function, and data preprocessing have little effect on the accuracy and only affect the speed of learning. However, the number of epochs has a great influence on the accuracy; too few epochs will cause under-fitting, whereas excessive epochs will cause over-fitting. When training the model, it is important to adjust the number of epochs of the training.

We used the model to classify the application set. The results of the classification of one flight strip of the application area are shown in Fig.5. In Fig.5(a), on the left end of the strip is a seafood farm with a residential area in the middle, and a coastline and sea on the right. On the high-resolution overall map, we found that there is a small number of classification errors on the left and in the middle, but there is no obvious classification error on the right. We believe the reason to be that the training set does not contain similar target areas, and the model does not learn the corresponding special diagnoses. In Fig.5(b), it is obvious that the rate of correct classification in the ocean and on land is very high, though there are mistakes at the land and sea junction. Accuracy of the classification in the area marked in Fig.5(a) is 95.61%.

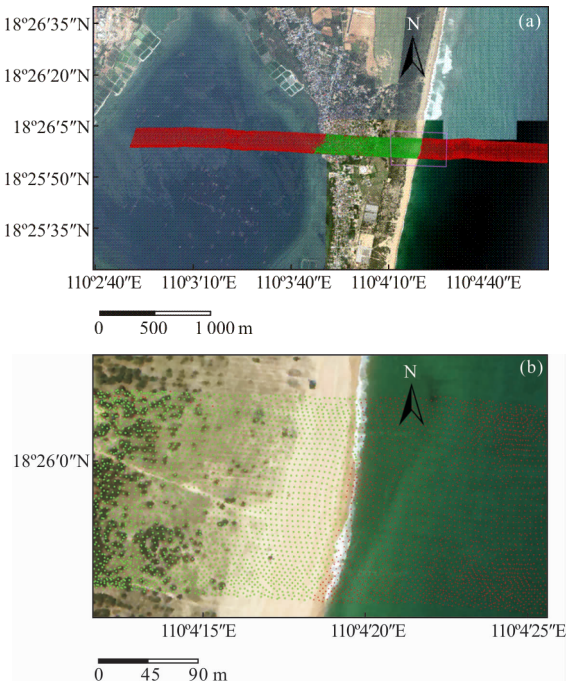


Fig.5 Mapped results of the classification of one flight strip of the application area: (a) the whole flight strip and (b) partially enlarged and subsampled image of the coastline

On an equivalent computer with Intel Core i7-8 700 K and NVIDIA GTX 1050, the training of 200 models on a GPU and CPU took approximately 20.9 hours and 42.9 hours. Using the trained model to classify 2 851 448 signals of the application dataset takes about 120 s, it indicates that the processing speed is high compared to the hours spent measuring these data.

The classification accuracy of the SVM method on the test set is 99.03%[6]. Including the extraction of effective signal operations and classification operations from raw data, the deep learning method takes about 279 s to process 122 GB of application area data, and the SVM method takes about 582 s to processes the same data. The processing speed is increased by about 52% because the deep learning method does not require complex feature extraction operations.

Point cloud based classification can only be performed after point cloud generation, while point cloud generation generally requires time equivalent to measurement time, and does not include classification time.

3 Discussion

The traditional classification of lidar returns is based on point cloud data, which does not contain any raw signal information, and the classification must be executed after the flight test and cloud data generation. From the above results, we find that the deep learning method based on the original echo data can realize real-time processing, and the target classification can be performed during the flight process. This method offers very attractive efficiency improvements.

The original waveform may contain more information that we have not yet exploited, and deep learning can automatically extract features that we may not have discovered, so it may not only be useful for classification of land and sea. In future work, we will classify the training dataset and the test dataset with further target area types, such as sand, stone, shallow water, and deep water, and expect that an improved target classification may be achieved.

4 Conclusions

The results prove that the method of deep learning is very effective, with high precision and speed for sea-land classification of lidar echo signal waveforms. Compared with the traditional target classification based on the final DEM point cloud data, classification based on the original echo waveforms requires less computation and simpler implementation, and it may be applied in data-processing for mapping, and the classification of other target area types, with appropriate training. Moreover, it may also be a good candidate method for classifying species on the sea floor with airborne laser bathymetry.

References:

[1] Tuell G, Barbor K, Wozencraft J. Overview of the coastal zone mapping and imaging lidar (CZMIL): A new multisensor airborne mapping system for the US army corps of engineers [C]//Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVI. International Society for Optics and Photonics, 2010, 7695: 76950R.

[2] Baltsavias E P. Airborne laser scanning: existing systems and firms and other resources [J]. ***ISPRS Journal of Photogrammetry and Remote Sensing***, 1999, 54(2-3): 164-198.

[3] Pe′eri S, Morgan L V, Philpot W D, et al. Land-water interface resolved from airborne LiDAR bathymetry (ALB) waveforms [J]. ***Journal of Coastal Research***, 2011,62: 75-85.

[4] Collin A, Long B, Archambault P. Merging land-marine realms: Spatial patterns of seamless coastal habitats using a multispectral LiDAR [J]. ***Remote Sensing of Environment***, 2012, 123: 390-399.

[5] Allouis T, Bailly J S, Pastol Y, et al. Comparison of LiDAR waveform processing methods for very shallow water bathymetry using Raman, near-infrared and green signals[J]. ***Earth Surface Processes and Landforms,*** 2010, 35(6): 640-650.

[6] Huang Tiancheng, Tao Bangyi, Mao Zhihua, et al.

Classification of sea and land waveform based on multi-channel ocean lidar[J]. ***Chinese Journal of Lasers***, 2017, 44(6): 0610002. (in Chinese)

[7] Ma Yue, Zhang Wenhao, Zhang Zhiyu, et al. Sea and sea-ice waveform classification for the laser altimeter based on semi-analytic model [J]. ***Infrared and Laser Engineering***, 2018, 47(5): 0506005. (in Chinese)

[8] Nahhas F H, Shafri H Z M, Sameen M I, et al. Deep learning approach for building detection using lidar-orthophoto fusion[J]. ***Journal of Sensors***, 2018:7212307.

[9] Hu X, Yuan Y. Deep-learning-based classification for DTM extraction from ALS point cloud [J]. ***Remote Sensing***, 2016, 8(9): 730.

[10] Arief H, Strand G H, Tveite H, et al. Land cover segmentation of airborne LiDAR data using stochastic atrous network[J]. ***Remote Sensing***, 2018, 10(6): 973.

[11] Maturana D, Scherer S. 3d convolutional neural networks for landing zone detection from lidar [C]//2015 IEEE International Conference on Robotics and Automation (ICRA), 2015: 3471-3478.

[12] Velas M, Spanel M, Hradis M, et al. Cnn for very fast ground segmentation in velodyne lidar data [C]//2018 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC), 2018: 97-103.

[13] Matti D, Ekenel H K, Thiran J P. Combining lidar space clustering and convolutional neural networks for pedestrian detection [C]//2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2017: 1-6.

[14] Dewan A, Oliveira G L, Burgard W. Deep semantic classification for 3d lidar data [C]//2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017: 3544-3549.

[15] Wang A, He X, Ghamisi P, et al. Lidar data classification using morphological profiles and convolutional neural networks[J]. ***IEEE Geoscience and Remote Sensing Letters***, 2018, 15(5): 774-778.

[16] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [J]. ***Proceedings of the IEEE***, 1998, 86(11): 2278-2324.

[17] Dai W, Dai C, Qu S, et al. Very deep convolutional neural networks for raw waveforms [C]//2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017: 421-425.

[18] Zhao Ming, Chen Shi, Yuen Dave. Waveform classification and seismic recognition by convolution neural network [J]. ***Chinese Journal of Geophysics***, 2019, 62(1): 374-382. (in Chinese)