An estimation method of underwater target spectrum line parameters based on convolutional neural network

Zhaorui Luo12*, Shuanping Du12, Yuechao Chen12

1Hangzhou Applied Acoustics Research Institute, Hangzhou, Zhejiang, 310012, China
2Science and Technology on Sonar Laboratory, Hangzhou, Zhejiang, 310012, China
*Correspondence author's email address: lizr426@alumni.sjtu.edu.cn

Abstract: The spectrum line parameters of underwater target radiated noise are important basis for underwater target recognition. In this paper, the convolution neural network (CNN) is applied to estimate the spectrum line parameters. The low frequency analysis record (LOFAR) spectrum of radiated noise is used as input data, and a suitable convolution neural network is constructed to estimate the spectrum line parameters. The processing results of simulation data show that the estimation sensitivity of convolution neural network reaches 91.6% under the condition of low SNR. It proves the effectiveness of CNN in estimating spectrum line parameters of underwater targets.

1. Introduction
The radiated noise of ships is rich in spectrum line. Low frequency analysis record (LOFAR) originally means low frequency analysis and record, but in the field of underwater acoustic engineering, the low frequency line spectrum of ship radiated noise is often called LOFAR spectrum. LOFAR spectrum is an important characteristic of ship radiated noise. It is widely used for sonar target detection and recognition[1]. However, in the face of the increasing detection range of sonar and the decreasing radiated noise of underwater targets, the traditional signal processing methods are difficult to realize accurate estimation of spectrum line parameters because of relying on prior models and rules. This becomes the main factor restricting the development of underwater target recognition.

In recent years, due to the improvement of parallel computing capability and the emergence of GPU acceleration technology, convolutional neural network (CNN) becomes a research hotspot in the field of artificial intelligence. It is clearly proved that, given sufficient computing resources and training data, CNN can learn complex nonlinear mappings. And different from traditional methods, CNN can automatically extract features from data for recognition and analysis. Based on the above advantages, CNN is rapidly applied to computer vision, image processing and other fields, achieving gratifying results[2]. In the field of underwater acoustic target recognition, some scholars preliminarily carried out the application research of CNN[3]. But some of them did not fully utilize the traditional underwater acoustic target recognition knowledge[4][5].

In this paper, CNN is combined with the existing target recognition technology. LOFAR spectrogram is proposed as the input feature. And a CNN matched with LOFAR spectrogram is constructed to estimate the spectrum line parameters of underwater targets.
2. Materials and methods

2.1. Convolutional neural network
At present, CNN is one of the representatives of deep learning network. It is very suitable for processing image data. It makes remarkable achievements in image detection, image segmentation, image classification and other fields. CNN is essentially a deep neural network with convolution structure. Convolution structure has three key operations: local receptive field, weight sharing and pooling layer. These three operations effectively reduce the number of network parameters to reduce the memory occupied by deep network. It becomes possible to process a large number of complex high-dimensional data.

CNN generally consists of input layer, convolution layer, pooling layer, full connection layer and output layer in turn. The input layer inputs each pixel into the network. The convolution layer is composed of multiple convolution kernels, which are used to extract features. The pooling layer reduces the dimension of convolution results. The full connection layer connects the feature mapping and the output layer. The number of nodes in the output layer corresponds to the number of classification categories, and the value of each node represents the probability of this type.

2.2. Preprocessing of acoustic target noise
In order to reduce the computation of deep learning algorithm and improve the efficiency of parameter estimation, underwater acoustic data are preprocessed. Finally, The normalized LOFAR spectrogram is taken as the input data of the deep learning model.

LOFAR analysis is an important analysis method of signal feature extraction, applying to the extraction of signal spectrum line in additive noise. The additive noise is

\[
x(t) = s(t) + n(t)
\]

As in equation (1), \(s(t)\) is the original signal and \(n(t)\) is the noise.

![Figure 1 The process of LOFAR.](image)

The process of LOFAR is shown in figure 1, and the specific steps are as follows:

- The sampling sequence of the original signal is divided into several continuous frames, each frame has \(N\) sampling points.
- Normalized and centralized processing is carried out for each frame signal sample \(L(n)\), as in equation (2) and equation (3). The purpose of normalized processing is to make the amplitude (or variance) of the received signal uniform in time scale. And the purpose of centralized processing is to make the average value of samples zero.

\[
u(n) = \frac{L(n)}{\max_{1 \leq i \leq N}[L(i)]}
\]

\[
x(n) = u(n) - \frac{1}{N} \sum_{i=1}^{N} u(i)
\]

- The LOFAR spectrum is obtained by performing short-time Fourier transform on the signal \(x(n)\), as in equation (4).

\[
X(\omega) = \text{FFT}[x(n)]
\]

- The LOFAR spectra are arranged according to the time dimension to obtain LOFAR spectrogram.
2.3. Construction and training of deep learning model

In order to accurately extract the low-frequency spectrum line parameters of the target under the condition of low SNR, it is necessary to rely on the powerful computing and nonlinear fitting capabilities of the depth CNN. However, with the increase of network depth, the network degradation will appear. This problem is solved by ResNet network. So this paper refers to this network to construct a deep CNN, which matches LOFAR spectrogram. The specific method is as follows.

Firstly, considering the difficulty of the task, ResNet152, which has large number of layers and powerful performance, is selected as a reference. The convolution layer of the model is set to 152 layers. Secondly, because the image size of LOFAR spectrogram is smaller than the original image size. Ensuring the feature dimension in the network, the step size of some largest pooling layers is modified to 1. Finally, in order to adapt to the spectrum line estimation task of LOFAR spectrogram, the number of nodes in the fully connected layer is changed to be equal to the number of frequency points of LOFAR spectrogram. The "sigmoid" function is used as the activation function, because it can independently predict whether each frequency point has a spectrum line or not.

During the training, the Batch size is 50, an Epoch has 100 steps, and the learning rate is 1e-3 by using the "Adam" optimizer. Take the binary cross entropy loss function as the model loss function.

3. Results and discussion

The model was trained using simulation data set. In simulation LOFAR spectrogram, the number of spectrum line is 0 to 20 and the SNR is -15dB to 0dB. The frequency range of LOFAR spectrogram is 400Hz. So LOFAR spectrogram has 400 frequency points. The time scale of LOFAR spectrogram has 150 frames. LOFAR spectrogram corresponds to a vector label (1*400), with element 1 representing a spectrum line and element 0 representing no spectrum line.

![Figure 2](image)

Figure 2 Part (a) is training loss value. Part (b) is validating loss value.

Figure 2 shows the variation of loss during model training. It can be seen that with the increase of training times, the loss of training set and validating set tend to be stable. The best model is obtained.
Figure 3 A simulation results of spectrum line parameters estimation. Part (a) is a simulation LOFAR spectrogram. Part (b) is the true label of spectrum line. Part (c) is the predicted result.

Figure 3 is a simulation result of spectrum line parameters estimation. It shows that the depth CNN can accurately predict the frequency of spectrum line in LOFAR spectrogram, even though two spectrum lines are very close together.

In order to test the performance of the model, the confusion matrix is obtained by testing 5000 LOFAR spectrogram of labeled data, as in table 1.

The accuracy, precision and sensitivity of the model are obtained from the confusion matrix, as in table 2. The performance of the model is balanced.

| Predicted Positive spectrum line | Positive spectrum line | Negative spectrum line |
|---------------------------------|------------------------|------------------------|
| 50573                           | 5169                   |
| 5985                            | 1938252                |

Table 2 A table of accuracy, precision and sensitivity.

| Accuracy  | Precision | Sensitivity |
|-----------|-----------|-------------|
| 99.5%     | 91.5%     | 91.6%       |

Under the conditions of different SNR, the sensitivity shows in table 3. With the decreasing of SNR, the sensitivity decreases, which means the performance of the model decreases. But, in the condition of low SNR, the model still has good performance.

| Range of SNR (dB) | Sensitivity |
|-------------------|-------------|
| [-15 , -10]       | 88.3%       |
| (-10 , -5]        | 91.3%       |
| (-5 , 0 ]         | 92.8%       |

4. Conclusion

In this paper, an estimation method of underwater target spectrum line parameters based on CNN is proposed. This method combines CNN with traditional underwater acoustic target recognition
knowledge, realizing parameters estimation of spectrum line under the condition of low SNR. It also provides a basis for further underwater target recognition. However, the effectiveness of this method is only proved by simulation data. It needs to be further verified by real data.

Reference
[1] Fang S, Du S, Luo X, et al. (2019) Development of underwater acoustic target feature analysis and recognition technology. Bulletin of Chinese Academy of Sciences, 34(3): 297-305.
[2] LeCun Y, Bengio Y, Hinton G. (2015) Deep learning. Nature, 521: 436-444.
[3] Chen Y, Fu J. (2019) Underwater acoustic target recognition method based on convolution neural network. Technical Acoustics, 38(2): 424-426.
[4] Cao X, Zhang X, Yu Y. (2017) Deep learning based recognition of underwater target. In: IEEE International Conference on Digital Signal Processing. Kuala Lumpur. pp. 89-93.
[5] Choi J, Choo Y, Lee K. (2019) Acoustic classification of surface and underwater vessels in the ocean using supervised machine learning. Sensors, 19(16): 1-15.
[6] He K, Zhang X, Ren S, et al. (2016) Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas. pp. 770-778.