ia-PNCC: Noise Processing Method for Underwater Target Recognition Convolutional Neural Network

Nianbin Wang[1], Ming He[1,2], Jianguo Sun[1,*], Hongbin Wang[1], Lianke Zhou[1], Ci Chu[1] and Lei Chen[3]

Abstract: Underwater target recognition is a key technology for underwater acoustic countermeasure. How to classify and recognize underwater targets according to the noise information of underwater targets has been a hot topic in the field of underwater acoustic signals. In this paper, the deep learning model is applied to underwater target recognition. Improved anti-noise Power-Normalized Cepstral Coefficients (ia-PNCC) is proposed, based on PNCC applied to underwater noises. Multitaper and normalized Gammatone filter banks are applied to improve the anti-noise capacity. The method is combined with a convolutional neural network in order to recognize the underwater target. Experiment results show that the acoustic feature presented by ia-PNCC has lower noise and are well-suited to underwater target recognition using a convolutional neural network. Compared with the combination of convolutional neural network with single acoustic feature, such as MFCC (Mel-scale Frequency Cepstral Coefficients) or LPCC (Linear Prediction Cepstral Coefficients), the combination of the ia-PNCC with a convolutional neural network offers better accuracy for underwater target recognition.

1 Introduction

With the development of marine resources and the implications of national security, underwater target recognition technology is becoming more widely used. It is a key area in the study of target recognition technology and is a vital issue in the field of acoustic signal processing. Scholars at across the globe have studied the topic from many aspects proposed solutions that analyze and resolve problems of underwater target recognition from different perspectives. However, based on current requirements, the main challenge in underwater target recognition is one-sidedness of feature representation, resulting from multiple feature representation methods. The combination of time and frequency signals of underwater target noise, and the formation of a feature extraction method based on both is the main focus of this paper.

[1] College of Computer Science and Technology, Harbin Engineering University, Harbin, 150001, China.

[2] College of Computer and Information Engineering, Heilongjiang University of Science and Technology Harbin,150022, China.

[3] College of Engineering and Computing, Georgia Southern University, Georgia, 30458, USA.

[*] Corresponding Author: Jianguo Sun. Email: sunjianguo@hrbeu.edu.cn.

Great importance has been attached to underwater target recognition technology by the academic and application sectors since earlier 1950s. Scholars in the USA have studied it since 1960s and Feigenbaum et al. in Stanford university developed the underwater prediction expert system and the improved SIAP by extracting features of narrow-band signal with signal recognition spectrum and related algorithms, in addition, related submarines in the USA and Britain have been equipped with the recognition system [Purton, Kourousis, Clothier et al. (2014)]. After that, scholars in Japan developed the SK-8 underwater target warning system based on FFT system, which compares the target signal with the existed spectrum to judge the target type [Xiao, Cai and Liao (2006)]. A.J. Bonner et al. in Canada developed the expert analysis system called INTERSENSOR based on the vessel radiated noise signal [Araghi, Khaloozade and Arvan (2009)]. Wu et al. [Wu, Jing, Chen et al. (1998); Wu, Li and Chen (1999)] have combined the energy spectrum of vessel noise with traditional statistical theory to recognize target with clustering. Yang Desen proposed the three-factor theory and judgement for line spectrum [Gu and Yang (2004); Li and Yang (2007)]. Han Shuping proposed a target spectrum of Spatio-temporal joint to differentiate target spectrum and self-noise spectrum effectively [Shu and Ping (2009)]. Yang Chunying applied the multiresolution decomposition algorithm and wavelet transform theory in the extraction of underwater target recognition power spectrum, in which the variable scale features of wavelet transform are used to obtain the better frequency resolution compared with the traditional short time Fourier transform method, improving the feature extraction accuracy greatly [Xi, Zou, Yang et al. (2011)].

With the rapid development of recognition technology, such technologies as neural network and deep learning have performed well in classification in competitions for their great nonlinear representation capacity. The SD-Section company in Britain has used two supervised learning methods and neural network structure to develop underwater target analysis system. Similarly, scholars at home have followed the study well. Wu et al. [Wu, Li and Chen (1999)] have extracted features like line spectrum and other characteristics of radiated noise, and established the template matching base as well as made use of fuzzy statistical theory and neural network to predict the classification of underwater vessels, what's more, they have made a good prediction performance. Wang et al. [Wang (2007)] have studied several joint extraction technology in spectrum features and improved the stochastic adaptive genetic algorithm and trained it with neural network, thus a good classification model was obtained.

In this paper, the auditory perceptual feature is introduced for underwater target recognition, and the improved anti-noise power-normalized cepstral coefficients method (ia-PNCC) is proposed, based on the PNCC auditory feature with multitaper and normalized Gammatone filter banks. It has been combined with a convolutional neural network to interpret data from both the time and frequency domains, and applied to underwater target recognition. The extracted frequency domain features of the target noise found by the ia-PNCC are represented as a vector of the convolutional neural network. In this work, we first show the optimized convolutional neural network model and parameters, then present the auditory perception of the sonar using attributes of the convolutional neural network. We showed that the underwater target noise vector representation based on the ia-PNCC was able to improve underwater target recognition compared with previous methods, including the MFCC and LPCC.

This paper demonstrates an improvement of the PNCC, to make it applicable to underwater target noise feature extraction and representation. It also shows how to combine underwater target noise feature representation with convolutional neural networks. The paper is organized as follows: the second part introduces the related concepts and application technology, the third part reports the problem and improvement to the PNCC method, and the fourth part reports experimental verification.

2 Related work

2.1 Underwater target noise processing

Underwater target recognition covers several disciplines, including transducer technology, testing technologies and digital signal processing. With the development of damping and noise reduction technology, the noise levels radiated by vessels is becoming lower and lower. Thus, it is more difficult to obtain samples of target-radiated noise. How to extract the underwater target-radiated noise signal and effective feature parameters in low signal-to-noise conditions is a key issue in identifying target types.

The features of underwater target-radiated noise are physical forms, and the detection via sonar has relatively objective standards. Tucker has verified the feasibility of analyzing water signal with auditory perception method in 2001, since then, with the development of study on auditory perception in recent decade, people have attached great importance to the role of auditory perceptual feature in underwater target recognition. Liu et al. [Liu, Sun and Yang (2008)] have proposed a feature extraction method based on human hearing model targeting to the defects of MFCC in recognition accuracy and robustness [Qu and Li (2007)]. Yang Yixin verified the better recognition performance of MFCC under non-noise disturbance condition, however, the recognition performance under noise disturbance condition will decrease a lot [Liu, Sun and Yang (2008); Yang, Yang and Wang (2016)]. Wang et al. [Wang, Zuo, Huang et al. (2011)] shows that the LPC cepstrum is a good method to study the time domain feature of the signals, and the signals are able to be separated into excitation component and vocal channel component. Literature [O'Shaughnessy (2002)] introduces it into underwater target recognition, achieving a better experiment performance. However, with both MFCC and LPCC there are limitations resulting from by their own characteristics. LPCC assumes that there is a linear predictive structure in the signal, which is suitable for vocal sounds with periodic features. However, for an underwater target passing through water at high speed, most of the radiated noise is turbulent. In fact, the LPCC method imposes an incorrect structure on turbulent noise. Similarly, there are problems with the MFCC-as the adjacent frame feature is extracted independently, which ignores internal correlations within signals [Wang, Li, Yang et al. (2016)]. Although the relation can be compensated by overlapping adjacent frames, there is no reasonable overlapping parameter for the real-world applications.

Compared with the MFCC and LPCC, the PNCC shows some anti-noise capacity [Lu, Zhang and Hu (2004)]. In this paper, we use auditory perception principles to describe the target noise. According to the obtained sound noise power spectrum, the ia-PNCC is extracted and applied to the underwater target recognition convolutional neural network. This process performed well under real-world conditions.

2.2 Convolutional neural network

Deep learning is a nonlinear information processing technology, which is able to achieve multiple tiers of information representation for the extraction and transformation of supervised or unsupervised features. Current acoustic target recognition method depend on previous experimental and expert knowledge, however deep learning algorithms are able to learn and find the classification information of noise automatically and dynamically, and adaptively construct a decision classification system. In underwater target recognition, most of the applications attached great importance to the artificial neural network.

A convolutional neural network (CNN) is an idea originated from the biologists' deep study on animal auditory perception principles after the neuron model was proposed in the end of 1940s. In the 1990s, Lecun et al. [Lecun and Bengio (1998)] constructed the LeNet-5 model based on the study of handwriting recognition, which adopted the alternately connected convolution tier and pooling tier as well as the full connection tier classification to set the foundation for modern convolutional neural network. In recent years, studies on structure of convolutional neural network enjoy great passions and some network structures with great performance are proposed. Since 2012, the AlexNet [Krizhevsky, Sutskever and Hinton (2012)] proposed by Krizhevsky et al. [Krizhevsky, Sutskever and Hinton (2012)] has won in the diagram classification competition in the large diagram data base ImageNet, surpassing the second place with 11%, for which, incessant convolutional neural network models are proposed after the convolutional neural network becomes the academic focus. And the models include the Visual Geometry Group (VGG) in Oxford University, GoogLeNet [Szegedy, Liu, Jia et al. (2015)] of Google and ResNet [Kaiming, Xiangyu, Shaoqing et al. (2015)] of Microsoft and so on, which have all broke the records of AlexNet in ImageNet.

In this paper, the CNN is applied to underwater target recognition, and the network model and parameters are adjusted and set by experiment and analysis, making full use of CNN to describe the auditory perception capacity in target recognition. The experimental results show that in underwater target recognition, the application of the CNN improves the recognition rate by 10-15% in the accuracy rate compared to a BP neural network.

3 Noise feature extraction of underwater target based on anti-noise PNCC

The PNCCis proposed by Kim et al. [Kim and Stern (2016)] and has a structure very similar to the MFCC and PLP. Compared with MFCC and GFCC, PNCC is able to remove the influence of background noise effectively with long time frame power analysis without losing recognition performance and calculation complexity. However, in terms of underwater target recognition, some processing methods of PNCC would eliminate noise of equipment and target channel. Therefore, improvements are made from two aspects in this paper to maintain target channel noise information when eliminating background noise. And the accuracy and robustness of recognition system are promoted greatly. The ia-PNCC processing is showed in Fig. 1.
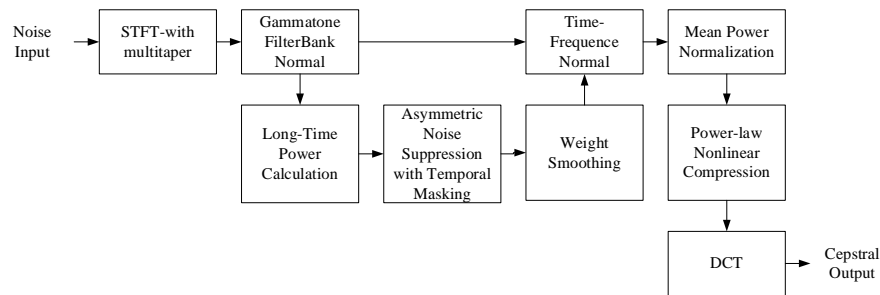
Figure 1: The extraction process of ia-PNCC

The PNCC is composed of three parts, comprising preprocessing, environment compensation, and post-processing. The initial processing stage is made up of pre-emphasis, a short-time Fourier transform, magnitude squaring, and Gammatone frequency integration. In the environment compensation stage, the long-time power calculation is completed, and asymmetric and temporal masking is carried out, as well as weight smoothing. Normalization of the time-frequency domain and power are also finished in this stage. In the post-processing stage, the initial PNCC is similar to the process of the MFCC process, in which the previous data results are processed in a nonlinear way and transformed inversely. Generally, it is obtained by the discrete cosine transform (DCT), and the value of it is the cepestrum result. However, different from the MFCC, power function is used here to finish nonlinear processing. Compared with the initial PNCC, the ia-PNCC proposed in this paper targets to improving the underwater target noise. And in the preprocessing stage, the preprocessing is removed and the original signal is maintained, also, the multitaper is used to extract the frequency signal; and in the environment compensation stage, normalized Gammatone filter banks is used to replace the traditional Gammatone filter banks; and in the post processing stage, CMU is removed to simplify the operation to improve the processing speed.

3.1 Removing pre-emphasis in the noise signal

Pre-emphasis is a method to augment the signal in the overall detected signal at the sending end. The signal may be lost while travelling through the transmission medium, so the damaged signal must be processed to augment the signal wave at the receiving end. In fact, pre-emphasis suppresses the low frequencies to flatten the signal spectrum for the channel parameter estimation later. All of these works are effective for the recognition and identification of sound. The power in low frequency part is more while less in high frequency part. However, the power spectral density of the output of the frequency detector is proportional to the square of the frequency, which means that the signal-to-noise ratio is larger for low frequencies.

In underwater target recognition, processing of the noise signal not only requires extracting the phoneme characteristics of targets, but also needs to focus on the channel information of the sounding equipment. Previous work has shown that the equipment channel information changes rather slowly compared to other phoneme parts and most of

the information is in the low-frequency part. Given this, pre-emphasis in the original noise input is likely to suppress the channel information in the low-frequency part. In addition, the classifier after feature extraction is a convolutional neural network, which has excellent nonlinear processing capacity. The training data in the convolutional neural network should be of high quality; therefore, in order to avoid damaging the underwater target noise feature in the initial stage, the pre-emphasis of the PNCC was removed in this paper first to reduce the artificial disturbance in the original signal. The effect of removing pre-emphasis is showed in Fig. 2.
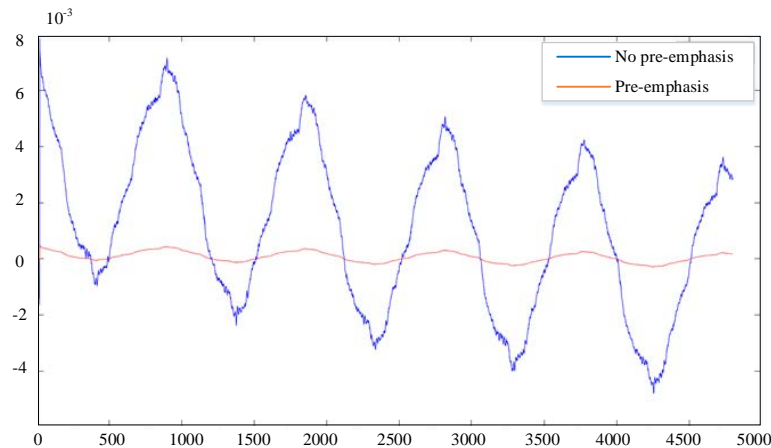

Figure 2: The frequency spectrum obtained by removing pre-emphasis and normal process

3.2 Replacing the hamming window with multitaper window

Fourier change is able to analyze the components in frequency spectrum in the noise signal, however, the real noise is the non-robust information that changes with time. And there is limitation of Fourier transformation in processing non-robust signals. In order to solve the problem, the short time Fourier transformation is chose to extract features. For the variability in the underwater targets, there are too much target feature information in the underwater noise signal, which is higher than the number needed for single recognition tasks. When adopting the single time-frequency window function to extract information, not only the extracted feature information is limited, but the extraction will be unsmooth and the spectrum will be distorted, therefore, Kinnunen et al. combined multitaper with MFCC, which improved the recognition accuracy characterized by MFCC and PLP. In this paper, the multitaper function was used to extract the features of underwater target noise signal. With the multitaper function, most of the frequency spectrum information would be maintained and a smooth result would be provided. Therefore, the multitaper function is adopted in this paper to extract the features of underwater target noise signal.

A given underwater target noise segment can be represented as a vector $x = [x(1), x(2), \ldots, x(N)]$, and the short time Fourier transformation can be represented in Eq. (1).

$$S(f) = \left| \sum_{t=1}^{N} w(t)x(t)e^{\frac{-i2\pi tf}{N}} \right|^2 \tag{1}$$

Where $w(t)$ is the window function, which is typically a Hamming function-a single trigonometric function window formed by a tapered sine or cosine function. When adopting a single time-frequency window function, the extracted information is limited and the processing of signal would not be smooth enough. Thus, in this paper, the multitaper was used to realize weighted mean for frequency information to obtain the final spectrum estimation.

The multitaper weighted mean spectrum estimation for frequency information can be represented in Eq. (2).

$$S(f) = \sum_{i=1}^{M} \phi(i) \left| \sum_{t=1}^{N} w_k(t)x(t)e^{\frac{-i2\pi tf}{N}} \right|^2 \tag{2}$$

Where $\phi(i)$ is the weighted coefficient of added subwindow, and $w_k(t)$ is the multitaper window.

The sine taper is adopted in this paper. In addition, the number of window function is an important parameter, the value of which would affect the calculation efficiency, and the number of multitaper is set as 3 in this paper.

3.3 Normalization of the Gammatone filter banks

In studies on auditory perception, it was found that the analysis and processing of acoustic signals in the basal membrane of the human cochlea are equal to the filtering and frequency decomposition processes. The Gammatone filter group is very close to the auditory properties of the human ear and the impulse response of the Gammatone in time–frequency can be represented in Eq. (3).

$$h(t) = ct^{n-1}e^{-2\pi bt}\cos(2\pi f_0 t + \phi), t > 0 \tag{3}$$

where c is the constant to adjust proportion; n is the filter order, which is set as 4; b is the declining speed, which is usually a positive number; f0 is the center frequency, and; ϕ is the phase, which can be omitted for human ears' insensitivity to the phases.

There is an inhibitory action in the Gammatone with respect to additive background noise and white Gaussian noise. However, processing the Gammatone increases the high-frequency part of the noise, so as to decrease the low-frequency signal-to-noise ratio, which is related to speech recognition applications. Compared with the background noise of slow changes, the speech needs to be strengthened. However, this approach is not suitable for underwater target recognition. The poor underwater acoustic environment and multiple recognition of object is the main reasons for the noise itself to be hidden within background noise and be undetectable. Therefore, a simple improvement of the high-frequency part in the target-radiated noise would not be effective. In this paper, the value of the channels in the Gammatone filter group is normalized and the traditional Gammatone filter for high-frequency enhancement of the noise signal is removed, so as to promote the ratio of low-frequency waves in the processing results. The normalized

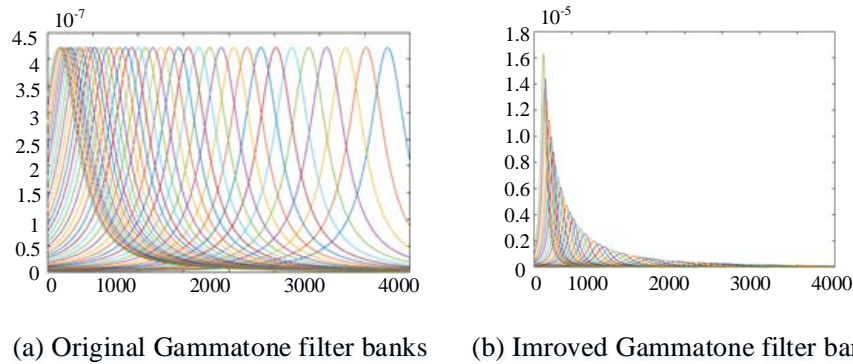Gammatone filter group in this paper is shown in Fig. 3.

(a) Original Gammatone filter banks   (b) Imroved Gammatone filter banks

Figure 3: Comparison between the original Gammatone filter banks (a) and the improved Gammatone filter banks (b)

4 Experiment and analysis

4.1 Data set

We used two kinds of experimental data. One was underwater data recorded by different sounding bodies with 18-channel hydrophones in an anechoic tank, for which the experimental structure is shown in Fig. 4. The experimental targets were divided into three types. The data sizes are shown in Tab. 1.
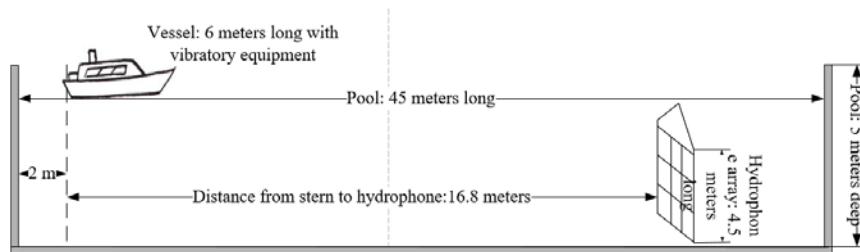

Figure 4: Experimental environment of the muffling pool and data collecting plan

Table 1: Experiment data size and distribution

Name	Total	Train	Test
Target 1	4800	3600	1200
Target 2	10000	7200	2800
Target 3	3470	2450	1020

It can be seen in Tab. 1 that a fine classification has been made for the three kinds of targets in the experiment. The total number of samples is 18270, and 13250 of them are used as training set, which accounts for 72.5% in the total sample. 5020 are used as testing set, accounting for 27.5% in the total sample.

In addition, data were collected by underwater sound collecting equipment in real waters, in order to verify the feature extraction capacity of the auditory perception plan. The target vessel used only sound monitoring equipment and moved at various speeds along different trajectories in accordance with the experiment requirements. The experiment environment and measurement conditions are shown in Fig. 5.
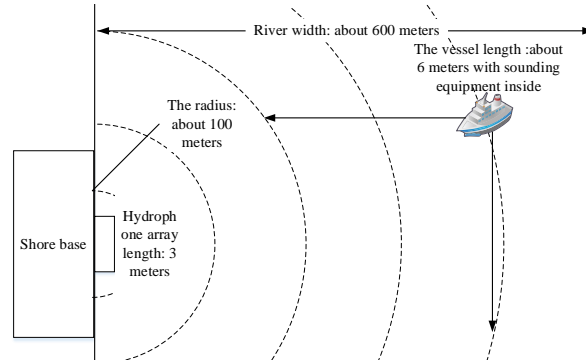

Figure 5: Real underwater experimental environment and data collecting plan

The vessel moves with different sounding bodies and at different speeds, and the experiment defines them as different targets, and the specific data size is showed in Tab. 2.

Table 2: Experiment data size and distribution for testing in river

Data content		Audio parameter				Data number
Power	Distance	Channel	Digitalizing bit	Sampling frequency	Sampling number	
8 horsepower	160 m/260 m/350 m	3	2	48000	4800	8192
90% power	50 m/260 m/350 m	3	2	48000	4800	31744
90% power	260 m	3	2	48000	4800	24576
8 horsepower	350 m with salvage ship	3	2	48000	4800	28672
90% power	350 m with salvage ship	3	2	48000	4800	30720

It can be seen from Tab. 2 that a fine classification has been made for the five types of targets in the experiment. The total number of samples was 100,000, and 74,496 of them are used as the training set, accounting for 75% of the total with 24,832 used as testing set, accounting for 25% of the total.

4.2 Experimental results and analysis

In this paper, a convolutional neural network was used as the underwater target recognition classifier, with the nonlinear features of the convolutional neural network being used to represent the sonar's perception capacity. The multiple channel underwater target noise was spliced in accordance with the arrangement of the multichannel hydrophones and the time-sequence relation between each segment of noise data. This data formed the input dataset for the deep learning model. The spliced data are shown in Fig. 6.
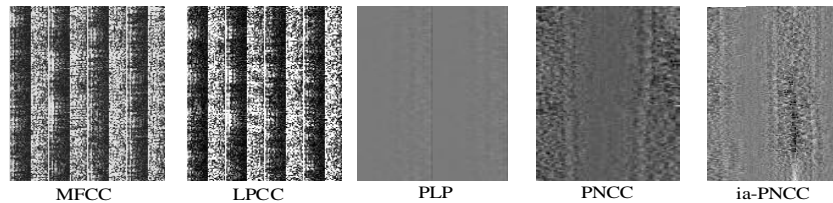

MFCC   LPCC   PLP   PNCC   ia-PNCC

Figure 6: Feature map of noise collected by multiple channel hydrophone

4.2.1 Parameter determination of convolutional neural network

For this paper, five-layer convolutional neural networks were adopted. The underwater target recognition experiments were initially conducted with different convolutional kernels. The experimental results are shown in Fig. 7.
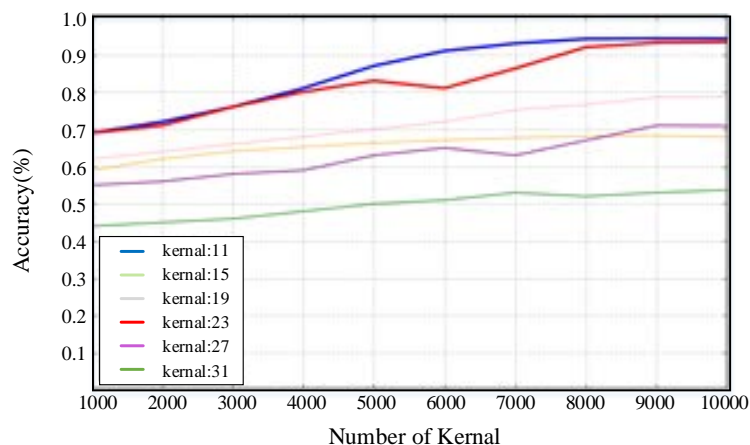

Figure 7: Comparison diagram of the effect of different Kernel size on test results

In Fig. 7, underwater target recognition accuracy of different convolutional kernels is presented. The abscissa is the number of iterations, the ordinate is the accuracy rate of different convolutional kernels acting on underwater target recognition, and the thickness of the lines reflects the corresponding accuracy. It can be seen from the figure that as the number of iterations increases, the model accuracy also increases, but tends to be flat in the middle and late stages of the iteration. When the sizes of the convolution kernel are 11×11 and 23×23, the recognition rates are higher, being 94.33% and 93.42%, respectively. For other sizes of convolutional kernel, the accuracy rates are not affected

by the size of the convolutional kernel. The accuracy was 78.7% when the convolutional kernel size was 19; 70.9% when the size was 27; 67.9% when the size was 15; and 53.7% when the size was 31. In addition, by comparing the curves when the convolutional sizes are 11 and 23, it is seen that accuracy rate curves tend to be flat in the middle and late stages of the iteration, indicating that accuracy does not change following the previous iterations. The former curve flattens before than the latter, indicating that it would enter into flat stage first. Furthermore, the Area under the curve (AUC) of the former is clearly higher than the latter, indicating that when the convolutional kernel size is 11, the performance is optimal.

4.2.2 Results from different processing methods

Experimental results of different underwater target noise processing methods with the convolutional neural network classifier of same parameters are showed in Tab. 3. It can be found that the recognition effect of the data preprocessed with PLP in the anechoic tank is best, and the effect of data preprocessed with ia-PNCC in real condition is best. Also, it can be found from the experimental results that the recognition rate of data obtained in anechoic tank with different noise perception feature extraction methods in the same convolutional structure classifiers are almost the same.

Table 3: Comparison of experimental results of different noise processing methods

Processing methods	Recognition rate in anechoic tank	Recognition rate in real condition
MFCC+CNN	89.1%	67.3%
LPCC+CNN	97.2%	76.58%
PLP+CNN	98.9%	89.32%
PNCC+CNN	93.24%	86.2%
ia-PNCC+CNN	96.47%	90.6%

The performances of the MFCC feature processing methods were quite different in the anechoic tank and under real-world conditions. This proved that the performance of the MFCC is better in a pure sound environment, but weaker in a noisy environment. Since the feature extraction of the LPCC is based on a linear assumption, its performance is generally better under the real conditions. The experimental results show that under the controlled conditions, the ia-PNCC feature processing and PLP perform similarly, and the performance of the ia-PNCC does not decline under real conditions. Furthermore, by decreasing the SNR (signal-to-noise ratio), the performance of the ia-PNCC is better than that of the PNCC, verifying the effectiveness and robustness of the method.

5 Conclusion

The improved PNCC auditory perceptual feature was adopted for underwater target noise feature extraction in this paper, and the underwater target noise signal perception attribute was quantified. Additionally, a convolutional neural network was used as the classifier to describe the auditory perception capacity of sonars for underwater target recognition.

Compared with the previous feature extraction methods, the improved PNCC auditory perceptual feature was able to offer improved feature extraction in noisy conditions, and the combination with the convolutional neural network was able to improve underwater target recognition accuracy. This was verified by the experimental results, in which the auditory perceptual feature offers advantages of simple calculation, large amounts of information, and wide application, etc. Combining the auditory perceptual features with such deep learning models as a convolutional neural network allows improvement of the underwater target recognition rate, providing a new direction for studies on underwater target recognition.

References

Araghi, L. F.; Khaloozade, H.; Arvan, M. R. (2009): Ship identification using Probabilistic Neural Networks (PNN). *International Multi Conference of Engineers & Computer Scientists*, vol. 2175.

Qu, B. D.; Li, J. B. (2007): Speech recognition based on two-dimensional PMCC robust feature parameter. *Journal of Computer Applications*, vol. 27, no. 10, pp. 2547-2548.

Gulin, O.; Yang, D. (2004): On the certain semi-analytical models of low-frequency acoustic fields in terms of scalar-vector description. *Chinese Journal of Acoustics*, vol. 25, no. 1, pp. 58-70.

Guoqing, W.; Jing, L. I.; Chen, Y.; Yuan, Y.; Chen, Y. (1998): Ship radiated-noise recognition (I) the overall framework, analysis and extraction of line-spectrum. *Acta Acustica*, vol. 23, pp. 394-400.

Kaiming, H.; Xiangyu, Z.; Shaoqing, R.; Jian, S. (2015): Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778.

Kim, C.; Stern, R. M. (2016): *Power-Normalized Cepstral Coefficients (PNCC) for Robust Speech Recognition*. IEEE Press, USA.

Krizhevsky, A.; Sutskever, I.; Hinton, G. E. (2012): ImageNet classification with deep convolutional neural networks. *International Conference on Neural Information Processing Systems*, pp. 1097-1105.

Lecun, Y.; Bengio, Y. (1998): *Convolutional Networks for Images, Speech, and Time-Series*. MIT Press, USA.

Li, S.; Yang, D. (2007): DEMON feature extraction of acoustic vector signal based on 3/2-D spectrum. *IEEE Conference on Industrial Electronics and Applications*, pp. 2239-2243.

Liu, G.; Sun, C.; Yang, Y. (2008): Target feature extraction for passive sonar based on two cepstrums. *International Conference on Bioinformatics and Biomedical Engineering*. pp. 539-542.

Lu, Z. B.; Zhang, X. H.; Hu, H. B. (2004): Auditory feature extraction of noise radiated from an underwater target. *Systems Engineering & Electronics*, vol. 26, no. 12, pp. 1801-1803.

O'Shaughnessy, D. (2002): Linear predictive coding. *IEEE Potentials*, vol. 7, no. 1, pp. 29-32.

Purton, L.; Kourousis, K. I.; Clothier, R.; Massey, K. (2014): Mutual recognition of national military airworthiness authorities. *International Journal of Aeronautical & Space Sciences*, vol. 15, no. 1, pp. 54.

Shu, X.; Ping, H. S. (2009): Improvement of DOA estimation using wavelet denoising. *IEEE International Conference on Information Science and Engineering*. pp. 587-590.

Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S. et al. (2015): Going deeper with convolutions. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-9.

Wang, F. (2007): Research on technique of passive sonar target recognition based on data fusion. *China Water Transport*, vol. 5, no. 4, pp. 123-125.

Wang, J.; Zuo, Y.; Huang, Y.; Yang, B.; Pan, S. (2011): Arc sound recogniting penetration state using LPCC features. *Lecture Notes in Electrical Engineering*, vol. 88, pp. 229-233.

Wang, W.; Li, S.; Yang, J.; Liu, Z.; Zhou, W. (2016): Feature extraction of underwater target in auditory sensation area based on MFCC. *Ocean Acoustics*, pp. 1-6.

Wu, G.; Li, J.; Chen, Y. (1999): Ship radiated-noise recognition (Ⅱ)-stability and uniqueness of line spectrum. *Acta Acustica*, vol. 24, no. 1, pp. 6-10.

Xiao, H. G.; Cai, C. Z.; Liao, K. J. (2006): Recognition of military vehicles by using acoustic and seismic signals. *Systems Engineering-Theory & Practice*, vol. 26, no. 4, pp. 108-113.

Xi, H. S.; Zou, K.; Yang, C. Y.; Zhou, P. (2011): Sea clutter covariance matrix estimation and its impact on signal detection performance. *Journal of Systems Engineering and Electronics*, vol. 33, no. 10, pp. 2174-2178.

Yang, L.; Yang, Y.; Wang, Y. (2016): Sparse spatial spectral estimation in directional noise environment. *Journal of the Acoustical Society of America*, vol. 140, no. 3.