Journal of Engineering and Sustainable Development
Vol. 28, No. 06, November 2024
https://jeasd.uomustansiriyah.edu.iq/index.php/jeasd
ISSN 2520-0917
Research Article
https://doi.org/10.31272/jeasd.28.6.12

Recognition of Underwater Acoustic Radar Signals Based on Multiresolution and Dense Convolutional Neural Network

Taqwa Oday Fahad[1], Abbas Hussien Miry[2]*, Ammar Al-Gizi[3], Mohammed Hussein Miry[4], Ahmed Talib Razzooqee[5]

[1]Biomedical Engineering Department, University of Technology, Baghdad, Iraq
[2,3,5]Electrical Engineering Department, College of Engineering, Mustansiriyah University, Baghdad, Iraq
[4]Communication Engineering Department, University of Technology, Baghdad, Iraq

[1]https://orcid.org/0000-0001-7537-9682
[2]https://orcid.org/0000-0002-7456-287X
[3]https://orcid.org/0000-0002-4200-665X
[4]https://orcid.org/0000-0002-6921-5152
[5]https://orcid.org/0009-0007-3802-2669
*Email: abbasmiry83@uomustansiriyah.edu.iq

Article Info

Received	06/02/2024
Revised	17/10/2024
Accepted	19/10/2024

Abstract

Recognizing underwater objects based on radiated noise information is one of the most crucial issues in underwater acoustics. Underwater acoustic target signals are altered by elements such as the undersea environment and the ship's operational circumstances; hence, generalizing the recognition model is crucial. Most conventional Machine Learning (ML) algorithms often encounter difficulties when dealing with the costly recognition model for massive data analysis. However, Convolutional Neural Networks (CNNs) can automatically extract features for precise categorization. DenseNet is a powerful CNN network, but it has a data duplication problem, so in this paper, an approach using multi-resolution with a dense CNN model for underwater acoustic radar signal detection is proposed to overcome the DensNet problem. At first, the wavelet decomposition with different levels is applied to the input signal to represent the suitable data. The decomposed signals are inputs to the dense CNN. Our detection approach beats other CNN models and achieves an overall accuracy of 99.5% at 0 dB SNR based on experimental findings evaluated on a real-world passive sonar data set.

Keywords: Convolution Neural Network; Deep Learning; Underwater Acoustic Signal Detection; Multiresolution

1. Introduction

After the end of the Cold War, the Western nations' navy shifted from being based in the deep sea to being based offshore. The port region's safety is crucial to the economies, politics, and defenses of coastal states and localities. As a result, several nations have given attack defense a lot of attention. The port region's technological capabilities. Because of this, there has been much interest in underwater surveillance systems for commercial management, military blockade, and safety defense in recent years. For both military and civil applications, underwater surveillance systems must be able to detect and identify entering and leaving boats as well as underwater vehicles [1]-[]2]. The army began using underwater acoustic communication systems (UWACS) extensively in the late 19th century [3-[4]. Underwater environments are home to various relevant auditory signals that vessel propellers, marine mammals, and other ambient organisms emit—autographs as a token of appreciation. As a result, passive radar systems use sound to detect and categorize nearby objects for critical applications like navigation and surveillance [5].

Target radiation noise is utilized in underwater acoustic target passive detection technology, which uses a radar system to detect and identify the kind of target. Generally, the hull constructions and mechanical vibration properties of various ship targets and propeller construction vary. These causes cause disparities in radiated noise. Hydrophones record a complex and hazy image of the ship's radiated noise because of the gap in operating conditions on board and the interference from time- and space-varying underwater acoustic channels and ocean noise. The challenges associated with underwater audio signal detection increase with complexity and fuzziness. As a result, enhancing a radar system's underwater audio signal detection capabilities might be challenging [6]. The artificial intelligence approach makes it possible to model complicated data and is appropriate for designing algorithms in intricate settings. Several researchers have used artificial intelligence to detect and recognize underwater audio targets [6]. It may be broadly classified into two categories: deep learning and standard machine learning. Classifier design and feature extraction are traditional machine-learning techniques [6]. Using conventional techniques, researchers extract a variety of data from the ship's radiated noise signal, including signal structure features [7], Time-frequency analysis and frequency characteristics [8]–[]9], and auditory perception aspects [10–[11]. Next, the retrieved characteristics are fed into classic machine learning classifiers, including the one based on a basic neural network or the approach based on statistical analysis. While classic machine learning methods can complete specific tasks, the complexity of the undersea environment, the variety of ship target operating circumstances, the extraction of features from excessive artificial involvement, and the simplicity of classifier design all restrict recognition accuracy. Researchers used deep learning for underwater acoustic signal recognition to address these issues.

CNN can recognize distinct items in visual input, give each object a learnable weight, and then distinguish one object from another. Another advantage of CNN is that it requires much less pre-processing than other classification methods.

This work studies one of the advanced deep learning types (DensNet). Because DensNet duplicates the data several times, an input data preprocessing stage is added to overcome this problem.

This paper uses the wavelet transform as a preprocessing stage and the DenseNet neural network as a classifier. Section 2 presents related works that recognize underwater signals. Section 3 describes the fundamentals of DenseNet. Section 4 presents the dataset used in this work. Section 6 describes the details of the DenseNet model. Section 6 shows the results with discussions, and finally, section 6 presents the work's conclusion.

2. Related Works

Underwater Acoustic (UA) signal detection and recognition has long been achieved by skillfully manipulating handmade aspects, including temporal and spectral parameters; nevertheless, their efficiency significantly impacts the sonar system's ultimate performance. Many feature extraction algorithms have been investigated to address this problem by capturing the acoustic properties of propellers, which may be divided into three groups: time [12], frequency [13], and time-frequency combination domains [14].

Time-frequency analysis methods are extensively employed for classifying UA signals because they are better appropriate for nonstationary signals and are inspired by human auditory perception. Bark-wavelet analysis and the Hilbert-Huang transform were investigated.

Zeng and Wang [15] proposed UA signal frequency decomposition and signal reconstruction based on instantaneous frequency and amplitude . To model up to 16 UA targets with the gamma tone coefficient. Zhang et al. [16] looked at several standard classifiers, including decision tree (DT), support vector machine (SVM), and k-nearest neighbor (KNN).

Several machine learning (ML) recognition methods have been extensively used to increase accuracy

in the past ten years.

Yuan et al. [17] investigated an innovative supervised feature separation method to optimize the deep features extracted by the one-dimensional convolutional auto-encoder-decoder model to enhance the precision of underwater acoustic target categorization on the data limit. Despite a satneeds to display moretrategy displays many shortcomings, such as the Fourier transform's high complexity for data transformation and its delicate performance. In the presence of additive noise.

Deep learning (DL), which has recently made significant progress in many different academic fields [18], has performed very well. For instance, Wang et al. [19] used a modified empirical mode decomposition (MEMD) and deep learning to train the GFCC feature. The GMM layer of the network reduces redundant features and increases the recognition rate. Another study [20] looks into a multimodal deep learning (DL) approach for ship recognition using ship-radiated sound. This approach involves simultaneously extracting deep features from the visual and auditory modalities and combining them to a moderate degree, perhaps to increase the accuracy and dependability of sonar systems. However, in addition to the requirement of multimodal synchronization, this technique incurs high computational costs [21]. A dense convolutional neural network was utilized for automatically training representative features, eliminating the need for expert knowledge in feature extraction and domain translation [5]. The CNN-based classifier outperforms other current CNN and ML models in an accuracy competition, reaching a recognition accuracy at SNR 0 dB equal to 98.85%.

The general formulation of the problem statement is as follows: It requires accurate identification of the signal type, which results from precise feature extraction and classification.

The objectives of the paper are:

1. It preprocesses the input into the network to get good data representation, improving the classification accuracy.
2. The use of a dense network of CNN types improves classification accuracy.
3. Compared to related works, the above two factors result in the best signal categorization accuracy.

3. Densely Connected Convolutional Networks

To address the issue of gradient vanishing caused by overly deep learning, the Densely Connected Convolutional Networks is offered, which decreases the parameters due to adding bypass multiplexing, making sure the neural network's layers are deep enough to extract enough features. Fig.1 depicts the neural network topology of DenseNet. As shown in Fig. 1, the input of each layer is derived from the output of all preceding levels.

From the input data, each neural network layer will extract features that will become more pronounced as the layer depth increases. Through feature reuse and bypass setup, DenseNet significantly decreases the number of network parameters and, to some extent, resolves the gradient vanishing issue [22].

4. Dataset Description

Doan [5] created the dataset used in this work. A passive sonar system records a data set at a sampling rate of 22050 Hz, consisting of 11 underwater acoustic signals (representing 11 classes for recognition) and a single noisy signal for performance testing.

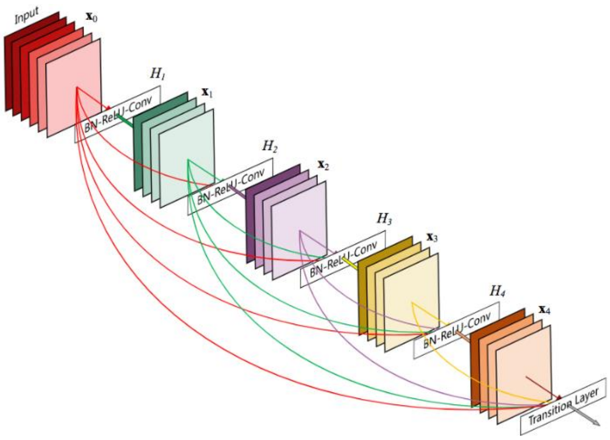

Figure 1. DenseNet neural network

A sonar specialist with years of experience utilizing acoustic-based sonar systems for target detection and reidentification labels each signal.

To test recognition models, the signals were subjected to the noise with a 2 dB step size and an SNR (standard deviation of the received signal power divided by the noise power) ranging from −20 to 10 dB. Each signal is successively split into 1000 observation frames as a preprocessing step before data modification; each frame contains 4096 amplitude samples. A total of 192000 signal frames are gathered for 12 UA signals tested at various SNRs. The data set is then split into 70% for training and 30% for testing at random.

5. CNN Model

The proposed method has two parts. The first part applies a wavelet transform to the input data to get an efficient data set resonation for the CNN.

The second part uses DenseNet, which is used in [5]. Fig.2 shows the flowchart for the proposed method. The model classifies 11 underwater acoustic signals and a single blank signal. As part of the data preparation procedure, the continuous audio signal is divided into many frames in the time domain, each having a length of 4096 samples.

At the start of the network, a batch normalization layer follows the input layer to help with the optimization process throughout the training phase. As shown in Fig.3 with Table 1, CNN is constructed by multiple convolutional blocks stacked, or "cTonv-blocks," each comprising activation, max-pooling, and convolutional layers.

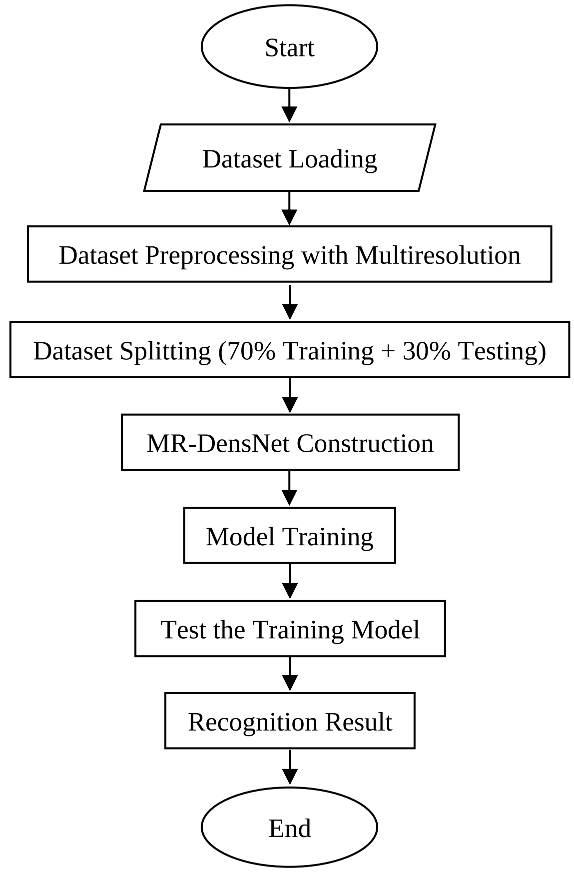

Figure 2. Flow chart of the proposed method

To create 32 feature maps, the convolutional layer is equipped with 32 one-dimension kernels measuring 7. This allows for a precise design. Next, a max-pooling layer is set up with a pool length of 3, and spatial pooling is used to downsample the output feature map y by removing any weak features that may exist. The dimension size of the output is reduced by defining the stride of (1, 2), which lowers the computing volume for several subsequent layers. In the network design, the activation layer comes after the max-pooling layer and is usually crucial in CNNs.

Due to rapid convergence, the (eLU) function is typically considered in several renowned CNN designs. Nevertheless, it indicates a barrier where information vanishes for input values smaller than zero, which the (eLU) function appears to be able to surmount to improve network training efficiency.

In UATC-DenseNet, there are many skip connections and a backbone flow. Three convolutional blocks stack the flow to extract deep features. Meanwhile, skip connections are carefully examined to enable the gradient flow farther into the network.

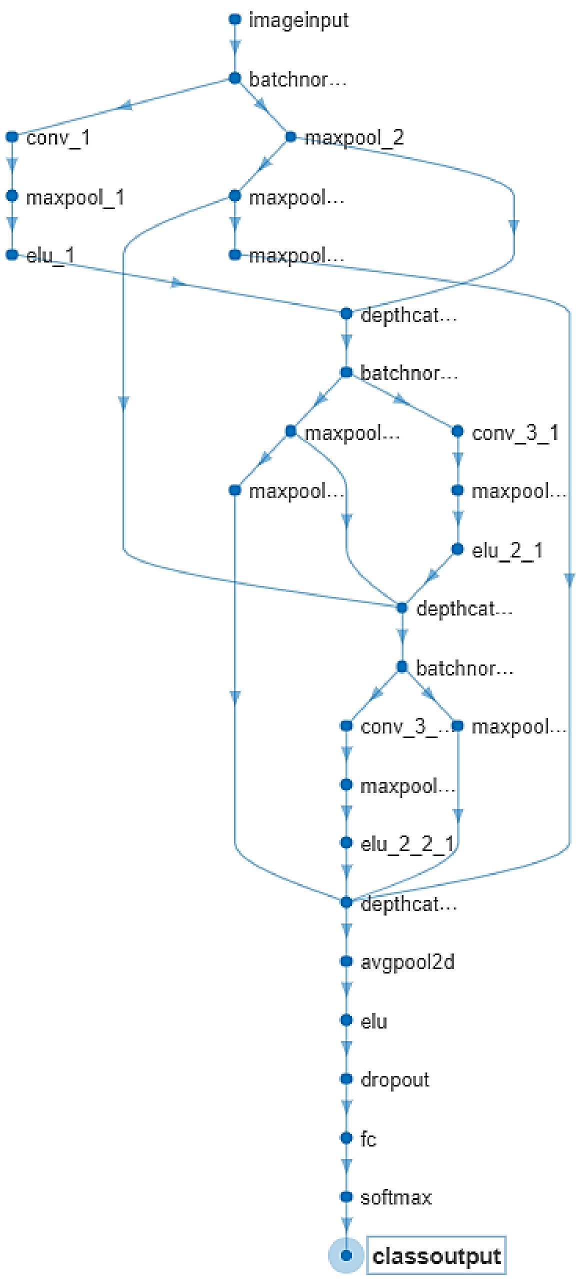

Figure 3. Structure of the CNN

In contrast to certain traditional CNNs, the skip connection protects the network from the vanishing gradient issue while maximizing the use of feature maps collected from several previous conv-blocks. The three most common skip connection strategies are addition, sidewise concatenation, and depth concatenation.

5. Discussions and results

The paper's technique categorizes 12 different types of signals using a dense CNN network with several filters. The simulation is Matlab 2022a. Table 2 shows the CNN Training Parameters.

Table 1. Layers of the CNN

	Name	Type	Activations
1	imageinput	Image Input	1(S) x 4096(S) x 1(C) x 1(B)
2	batchnorm_1	Batch Normalization	1(S) x 4096(S) x 1(C) x 1(B)
3	conv_1	Convolution	1(S) x 4096(S) x 32(C) x 1(B)
4	maxpool_1	Max Pooling	1(S) x 2048(S) x 32(C) x 1(B)
5	maxpool_2	Max Pooling	1(S) x 2048(S) x 1(C) x 1(B)
6	maxpool_5_1	Max Pooling	1(S) x 1024(S) x 1(C) x 1(B)
7	elu_1	ELU	1(S) x 2048(S) x 32(C) x 1(B)
8	Depthcat_1	Depth concatenation	1(S) x 2048(S) x 33(C) x 1(B)
9	batchnomi_3_1	Batch Normalization	1(S) x 2048(S) x 33(C) x 1(B)
10	maxpool_4_1_1	Max Pooling	1(S) x 1024(S) x 33(C) x 1(B)
11	Conv_3_1	Convolution	1(S) x 2048(S) x 32(C) x 1(B)
12	maxpool_3_1_1	Max Pooling	1(S) x 1024(S) x 32(C) x 1(B)
13	elu_2_1	ELU	1(S) x 1024(S) x 32(C) x 1(B)
14	maxpool_4_1_2_1	Max Pooling	1(S) x 512(S) x 33(C) x 1(B)
15	depthcat_2_1	Depth concatenation	1(S) x 1024(S) x 66(C) x 1(B)
16	batchnorm_3_2_1	Batch Normalization	1(S) x 1024(S) x 66(C) x 1(B)
17	maxpool_4_2_1	Max Pooling	1(S) x 512(S) x 66(C) x 1(B)
18	maxpool_5_2_1	Max Pooling	1(S) x 512(S) x 1(C) x 1(B)
19	conv_3_2_1	Convolution	1(S) x 1024(S) x 32(C) x 1(B)
20	maxpool_3_2_1_1	Max Pooling	1(S) x 512(S) x 32(C) x 1(B)
21	elu_2_2_1	ELU	1(S) x 512(S) x 32(C) x 1(B)
22	deptheat_2_2_2_2	Depth concatenation	1(S) x 512(S) x 132(C) x 1(B)
23	avgpool2d	Average Pooling	1(S) x 64(S) x 132(C) x 1(B)
24	elu	ELU	1(S) x 64(S) x 132(C) x 1(B)
25	dropout	Dropout	1(S) x 64(S) x 132(C) x 1(B)
26	fc	Fully Connected	1(S) x 1(S) x 12(C) x 1(B)
27	softmax	Softmax	1(S) x 1(S) x 12(C) x 1(B)
28	classoutput	Classification Output	1(S) x 1(S) x 12(C) x 1(B)

Table 2. CNN Training Parameters

Parameter	Value
MaxEpochs	20
batchSize	64
InitialLearnRate	0.001
LearnRateDropPeriod	5
LearnRateDropFactor	0.1
L2Regularization	0.0001

This work is done on a Laptop Core i7-9750H CPU and NVIDIA GeForce RTX 2060 GPU.

In the first experiment, the impact of the wavelet level on the overall accuracy is thoroughly investigated, in which the wavelet level decomposition configured varies in the set of {3, 4, 5, 6, 7}.

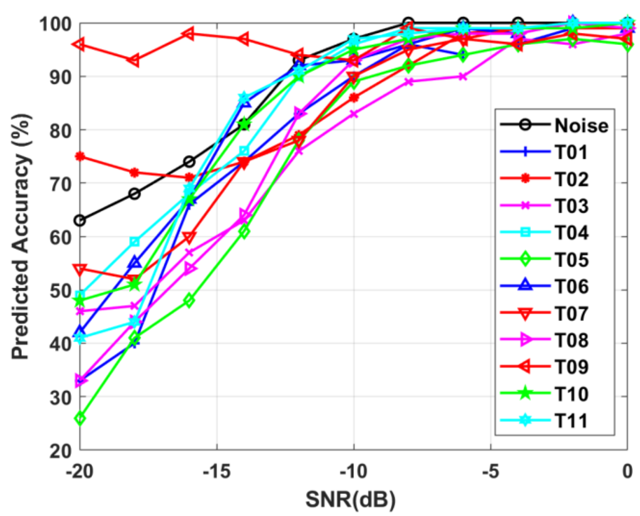

Figure 4. Detection accuracy with wavelet level 3

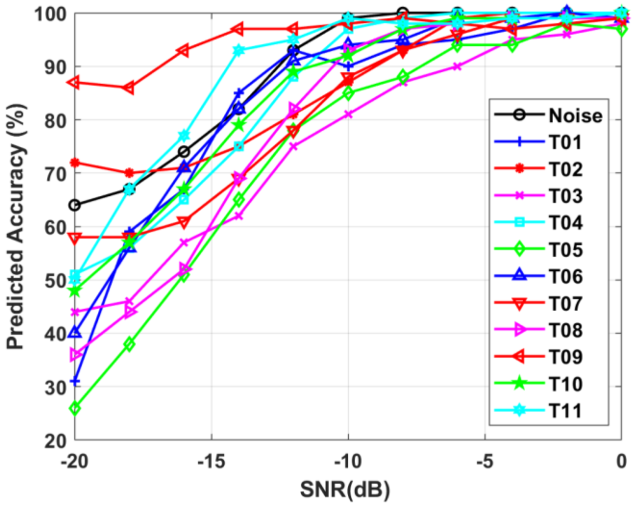

Figure 5. Detection accuracy with wavelet level 4

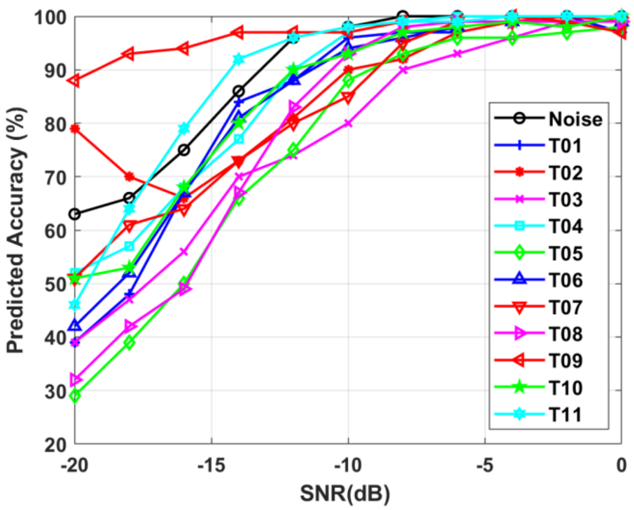

Figure 6. Detection accuracy with wavelet level 5

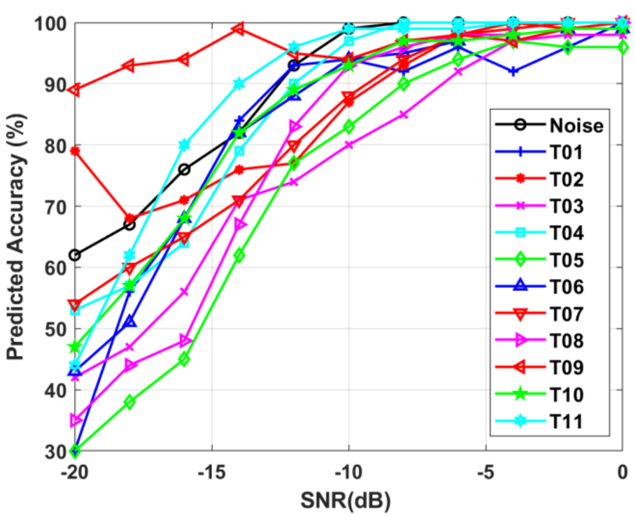

Figure 7. Detection accuracy with wavelet level 6

According to the result in Fig (4-8), different wavelet levels to feed data to CNN allow for a unique analysis of UA signals. Multiple feature representations capture temporal correlations, which are then integrated via the skip-connection process in MR-DenseNet. Fig. 9 shows the performance comparison of all wavelet levels. The performance of the wavelets levels is similar. Wavelet level 5 is better than another wavelet level.

The proposed network outperforms popular networks like CNN-ELM [23], ResNet18 [24], and SqueezeNet [25] in detection accuracy—the obtained results in Fig.9.

MR-DenseNet maximizes data representational features to attain greater accuracy, with gains ranging from 1%-9%.

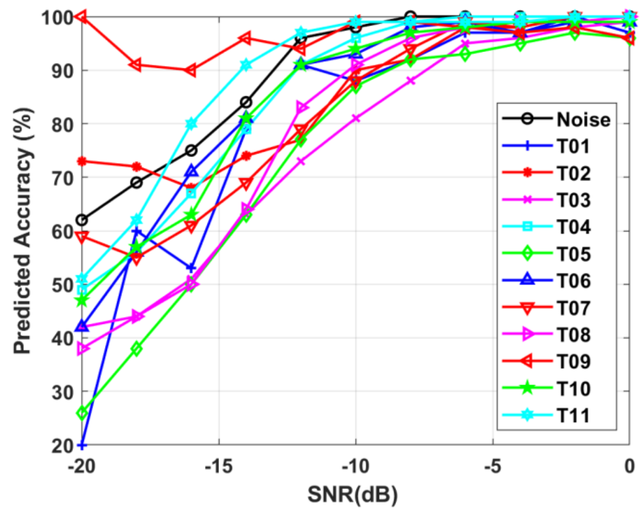

Figure 8. Detection accuracy with wavelet level 7

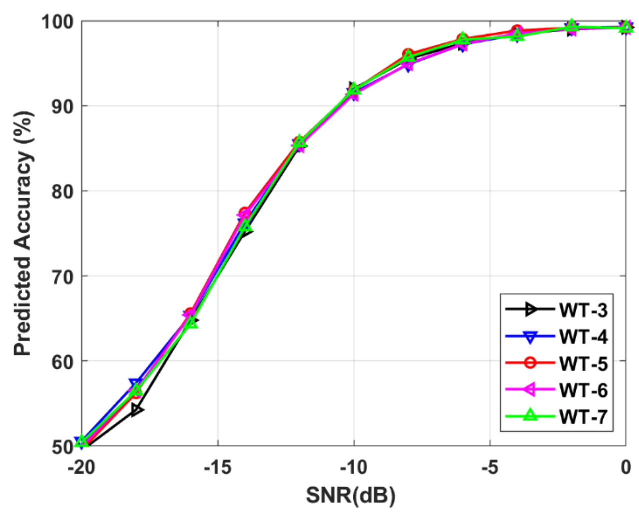

Figure 9. Detection accuracy with all wavelet levels

As shown in Fig. 10 at SNR = -20 dB, the MR-DenseNet reaches an accuracy of 50%, UATC-DenseNet almost reaches 40%, and the other networks are less than 30%. This shows the importance of modifying the input data set. The proposed network performs significantly better at low SNRs. When the wavelet transform is added, the five-level decomposition is combined with the Db2 basis function. The gathered underwater target signals are preprocessed using the wavelet decomposition technique, and the high-frequency underwater target noise signal coefficient is isolated. Still, the original underwater target signal remains useful after processing. The underwater target signal following wavelet transform decomposition serves as the input for CNN. This producer improves DensNet performance, as seen in Fig. 9.

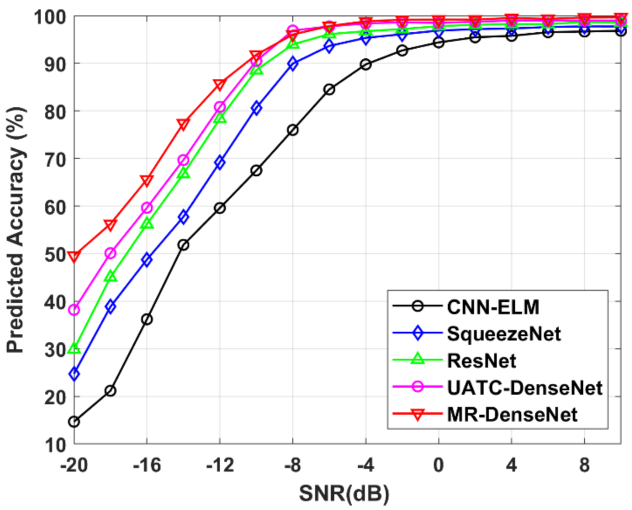

Figure 10. Compares BL, VGG, RN, and proposed CNN.

6. Conclusion:

Our work revealed the assessment of MR-DenseNet with different wavelet levels for UA target detection and recognition at varying SNRs and compared it to other current works. MR-DenseNet recognizes the acoustic radar signals more correctly due to its efficient data preprocessing with deeper network design. MR-DenseNet outperforms different CNNs on 12-target detection. With such accuracy, MR-DenseNet's future is promising, and it can be used in underwater radar systems to categorize acoustic objects effectively. The additional process for the dataset is the limitation of the proposed method. This work validates the practical applicability using an actual passive sonar dataset evaluation. This exemplifies the power of this methodology and its possible use in actual underwater situations. It can be used for another application that requires recognition.

Acknowledgments

The authors would like to express their appreciation to Mustansiriyah University (www.uomustansiriyah.edu.iq), Baghdad-Iraq, for its assistance and academic advice in writing this work.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Author Contributions Statement

Authors Mohammed Hussein and Abbas Hussien proposed the research problem .

Authors Mohammed Hussein and Ali Hussien developed the theory and carried out the calculations .

Authors Mohammed Hussien and Ammar Al-Gizi verified the analytical methods, investigated algorithm performance comparisons, and managed the results of this article .

All authors analyzed the findings and participated in the final article.

References

[1] Yuan, F., Ran, A., and Yue, W., (2012). *Modeling and characteristic analysis of underwater acoustic signal of the accelerating propeller*. SCIENCE CHINA Information Sciences Vol. 55 No. 2: 270–280. http://doi.org/10.1007/s11432-011-4285-9.

[2] Asada A, Maeda F, Kuramoto K, et al., (2007). *Advanced surveillance technology for underwater security sonar systems*. Proc IEEE OCEANS'07, Aberdeen, UK, pp:1–5. https://doi.org/10.1109/OCEANSE.2007.4302220.

[3] Murad, T. E., Al-Aboosi, Y.Y.,(2023). *Bit error performance enhancement for underwater acoustic noise channel by using channel coding.* Journal of Engineering and Sustainable Development, Vol. 27, No. 5. pp: 659- 670. https://doi.org/10.31272/jeasd.27.5.8.

[4] Mahmood, M. H., and Al-Aboosi, Y.Y. (2023). *Effects of multipath propagation channel in Tigris river.* Journal of Engineering and Sustainable Development, Vol. 27, No. 2, pp: 659- 670. https://doi.org/10.31272/jeasd.27.2.9.

[5] Doan, V.S., Thien, H.T., and Kim, D.S., (2022). *Underwater Acoustic Target Classification Based on Dense Convolutional Neural Network*, IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5. http://doi.org/10.1109/LGRS.2020.3029584.

[6] Li, D., Liu, F., Shen, T. and et al., (2022). *Generalizable Underwater Acoustic Target Recognition Using Feature Extraction Module of Neural Network.*. Applied Sciences, Vol. 12, No. 21, pp. 1-17. https://doi.org/10.3390/app122110804.

[7] Jiang, J., Wu, Z., Lu, J.A. and et al. (2020). *Interpretable features for underwater acoustic target recognition*. Measurement Vol.173. https://doi.org/10.1016/j.measurement.2020.108586.

[8] Liu, J., He, Y., Liu, Z., et al. (2014) *Underwater Target Recognition Based on Line Spectrum and Support Vector Machine*. International Conference on Mechatronics, Control and Electronic Engineering (MCE-14), Shenyang, China. https://doi.org/10.2991/mce-14.2014.17.

[9] Jahromi, M.S., Bagheri, V., Rostami, H. et al. (2019). *Feature Extraction in Fractional Fourier Domain for Classification of Passive Sonar Signals*. Journal of Signal Processing Systems, Vol.91, pp: 511–520. https://doi.org/10.1007/s11265-018-1347-x.

[10] Zhang, L., Wu, D., Han, X. et al. (2016). Feature Extraction of Underwater Target Signal Using Mel Frequency Cepstrum Coefficients Based on Acoustic Vector Sensor. Journal of Sensors, Vol. 2016. https://doi.org/10.1155/2016/7864213.

[11] Mohankumar, K., Supriya, M.H. and Pillai, P.S., (2015). *Maximum power point tracking techniques for photovoltaic systems: a comparative study.* International Symposium on Underwater Technology, http://doi.org/10.1109/UT.2015.7108321.

[12] Pinheiro, B. C., Moreno, U. F., Sousa, J. T. B., and et al. (2017). *Kernel-function-based models for acoustic localization of underwater vehicles*. IEEE Journal of Oceanic Engineering, Vol. 42, No. 3, pp. 603–618. http://doi.org/10.1109/JOE.2016.2578218.

[13] Rahmati, M. and Pompili, D. (2017). *UNISeC: Inspection, separation, and classification of underwater acoustic noise point sources*. IEEE Journal of Oceanic Engineering, Vol. 43, No. 3, pp. 777–791. http://doi.org/10.1109/JOE.2017.2731061.

[14] Sharma G., Umapathy, K., and Krishnan, S. (2020). *Trends in audio signal feature extraction methods*. Applied Acoustics, Vol. 158.https://doi.org/10.1016/j.apacoust.2019.107020.

[15] Zeng, X.Y. and Wang, S.G., (2013). Bark-*wavelet analysis and Hilbert Huang transform for underwater target recognition*, Defence Technology, Vol. 9, No. 2, pp. 115–120. https://doi.org/10.1016/j.dt.2012.09.001.

[16] Zhang, W., Wu, Y., Wang, D., and et al. (2018). Underwater *target feature extraction and classification based on gammatone filter and machine learning*. International Conference on Wavelet Analysis and Pattern Recognition, pp. 42–47. https://doi.org/10.1109/ICWAPR.2018.8521356.

[17] Yuan, X., Ke, F., and Cheng,E. (2018). *Underwater acoustic target recognition based on supervised feature-separation algorithm*. Sensors, Vol. 18, No.: 4318. https://doi.org/10.3390/s18124318.

[18] Shah, A. H. ,Miry ,A.A. and Salman,T.M. ,(2023) . *Automatic Modulation Classification Using Deep Learning Polar Feature.* Journal of Engineering and Sustainable Development, Vol. 27, No. 4. pp: 477- 486. https://doi.org/10.31272/jeasd.27.4.5.

[19] Wang, X., Liu, A., Zhang, Y. and et al. (2019). Underwater *acoustic target recognition: A combination of multi-dimensional fusion features and modified deep neural network*. Remote Sensing 11, No. 16: 1888 https://doi.org/10.3390/rs11161888.

[20] Yuan F, Ke X, Cheng E. (2019) *Joint Representation and Recognition for Ship-Radiated Noise Based on Multimodal Deep Learning*. Journal of Marine Science and Engineering. Vol.7, No.11:380. https://doi.org/10.3390/jmse7110380.

[21] Cao, X., Zhang, X., Togneri, R., and et al. (2019). Underwater *target classification at greater depths using deep neural network with joint multipledomain feature*. IET Radar, Sonar Navigation., Vol. 13, No. 3, pp. 484–491. https://doi.org/10.1049/iet-rsn.2018.5279.

[22] Xie, X., Yang, G., Jiang, M. Q. and et al. (2021). *A Kind of Wireless Modulation Recognition Method Based on DenseNet and BLSTM*. IEEE Access, Vol. 9, pp. 125706-125713. https://doi.org/10.1109/ACCESS.2021.3111406.

[23] Hu, G., Wang, K., Peng, Y. and et al. (2018). *Deep Learning Methods for Underwater Target Feature Extraction and Recognition.* Computational Intelligence and Neuroscience, Vol.2018, pp:1-10. https://doi.org/10.1155/2018/1214301.

[24] Kaiming H., Xiangyu Z., Shaoqing R., et al. (2016). *Deep Residual Learning for Image Recognition.* IEEE Conference on Computer Vision and Pattern Recognition. pp:770-778. https://doi.org/10.48550/arXiv.1512.03385.

[25] Forrest N., Song H., Matthew W., et al. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. arXiv:1602.07360. http://arxiv.org/abs/1602.07360.