VERİ BİLİMİ DERGİSİ
www.dergipark.gov.tr/veri

Ensemble NASNet Deep Feature Generator Based Underwater Acoustic Classification Model

Orhan YAMAN[1*],Turker TUNCER[2]

[1,2]*Firat University, Digital Forensics Engineering, Elazig, Turkey*

Abstract

Underwater acoustics is one of the important and complex research areas for advanced signal processing. Underwater acoustics is used in many fields, including defense technologies in the literature. In this study, a deep learning and machine learning-based method has been developed and underwater direction determination has been made. A new dataset was collected to determine direction underwater. First, a microphone is placed underwater. Propeller sounds of Remote Controlled Underwater Vehicle (ROV) moving in x, y, and z directions underwater were collected. Four classes were obtained with the sounds taken from the x, y, z directions, and the quiet environment. The resulting sounds were converted into images. Features were extracted from these images with deep learning methods. In this study, NASNetLarge and NASNetMobile deep learning models were preferred. The features extracted from these two models are combined. The chi2 algorithm was used to select the most weighted features among these features. Support Vector Machine (SVM) algorithm is used to classify selected features. With the proposed method, an accuracy of 77% and above was calculated.

Keywords: *Underwater Sound Classification, Underwater Direction Detection, SVM, NASNetLarge, NASNetMobile.*

Article Info

Received 30/06/2021
Accepted 05/08/2021

[1*]Corresponding author: orhanyaman@firat.edu.tr

[**] Part of this work was presented orally at the I. International Conference on Interdisciplinary Applications of Artificial Intelligence (ICIDAAI) 2021.

1 Introduction

Today, most countries make high-budget investments in defense technology. Many applications based on artificial intelligence have been developed for land, air, and naval defense technologies. Underwater communication, underwater imaging, and underwater acoustic studies are carried out in the naval defense system. Vehicles are moving underwater or on the water, the surface can be detected by underwater acoustic methods [1]. Underwater acoustic studies are especially important for border security. In sound-based methods, underwater microphones are used to listen underwater. By analyzing the obtained sound data, underwater vehicles, or ships on the water surface can be easily detected. There are many object detection studies using underwater acoustic data in the literature [2,3]. Acoustic based object detection methods have also been developed in many areas, mostly underwater mine searching [4]. Underwater, the image is often blurry, making it difficult to view images. For this reason, acoustic-based methods are more preferred. Also, the direction can be determined by using underwater acoustic data [5,6].

Jiang et al. [7] developed a CNN model for acoustic target recognition. He achieved an accuracy of 64.17%, 99.17%, and 59.17% in the CNN model he developed using three different data sets. Neves et al. [8] proposed a deep learning-based method for detecting underwater structures and detecting shipwrecks. It used RBoxNet, and YOLOv2 + RBoxNet models. They used the RBoxNet and YOLOv2 + RBoxNet models and calculated 90.3% and 77.5% accuracy with this method, respectively. Reis et al. [9] perform acoustic signature and boat detection using underwater acoustic data. It uses frequency amplitude variation. Choi et al. [10] proposed a multi-target localization method for underwater acoustics. It used signal processing methods. Sierra et al. [11] proposed a fuzzy logic-based method for classifying small ships moving on the water surface. Many studies in the literature detect objects using underwater acoustics [12–14].

Our motivation in this study is to collect an underwater acoustic dataset and suggest a deep learning-based hybrid method. The acoustic dataset obtained from underwater consists of four classes. This sound data was normalized and transformed into an image. Feature extraction was made with deep learning from images. NASNetLarge and NASNetMobile models are used for deep learning. Features derived from these two models are combined. Then, the most significant features were selected by the Chi2 method. SVM algorithm was used for classification after the feature selection process. A hybrid model has been developed in the proposed method by combining both deep learning and machine learning methods. The innovations of the proposed method are collecting underwater acoustic data and contributing to the literature. Also, a hybrid model has been proposed by combining NASNetLarge and NASNetMobile methods with SVM.

2 Materials

In this study, GLADIUS MINI [15] model underwater robot was used to create Underwater Direction Dataset (UDD). This robot has 100 meters of cable. It can transfer images in real-time. It can also be controlled by phone, tablet, or computer. EKEN brand sound and video recording device was used to capture underwater acoustic data. The underwater robot and voice recorder used for UDD are given in Figure 1.

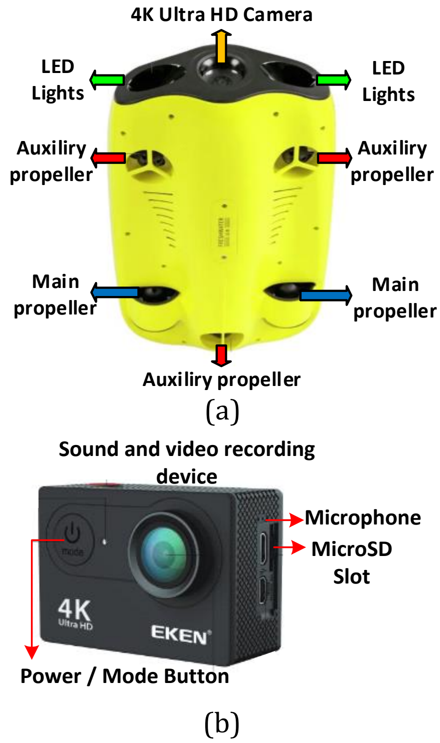

Fig. 1. Devices used while collecting underwater direction dataset (a) Underwater robot (b) Sound and video recording device

Acoustic data for UDD has been obtained using sound and video recorders. QuickTime Video File (MOV), bit rate 1411 kbps, channels 2 (stereo), Sample rate 44.100 kHz are used for sound recording. While collecting the UDD dataset, an experimental setup was created, as shown in Figure 2.

As shown in Figure 2, the voice recorder is fixed underwater at a depth of five meters. Sound data were collected underwater for five minutes from the quiet environment without the use of an ROV. Later, the GLADIUS MINI [15] robot moved on the x-axis for 5 minutes and recorded it. This process was repeated for the y and z axes for five minutes. At the end of these steps, a total of twenty minutes UDD was created for four different situations. The features of the underwater direction sound dataset collected are shown in Table 1.

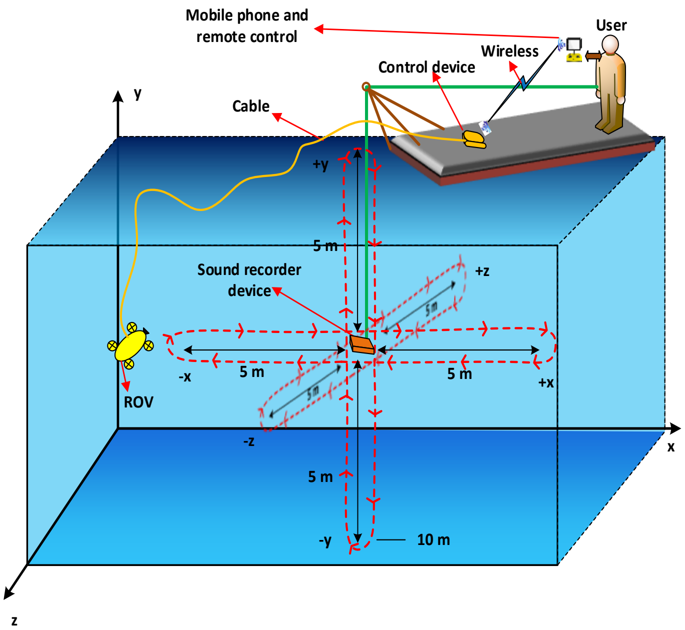

Fig. 2. Experimental setup for direction underwater sounds acquisition

Table 1. Class Information of the Collected Underwater Direction Dataset

Class number	**Class definition**	**Time (min)**	**Number of Samples**
Class 1	No movement	5	300
Class 2	x-axis movement	5	300
Class 3	y-axis movement	5	300
Class 4	z-axis movement	5	300

In the direction dataset, we obtained 300x44100 data by taking 300 seconds of sound for each class. This case total of 1200x44100 data for four classes in the direction dataset. In this study, sample sound signals taken from microphones for the direction dataset have shown in Figure 3.

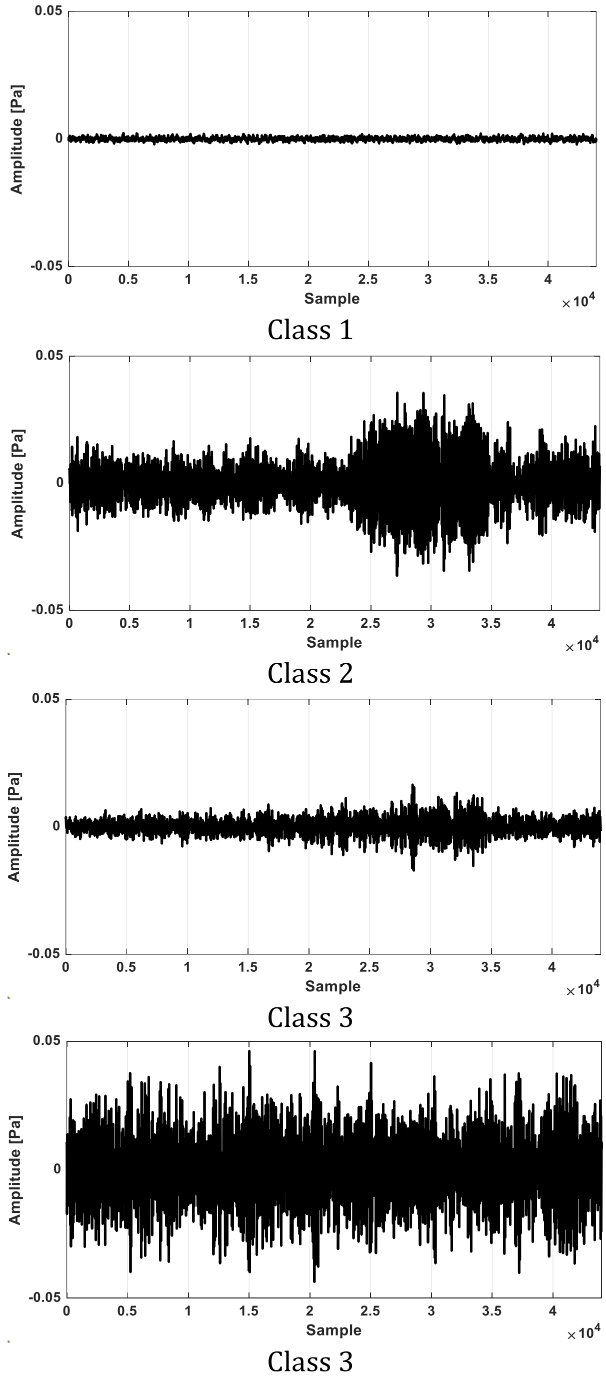

Fig. 3. Sound samples collected for direction determination

3 Proposed Method

The block diagram of the method suggested in this study is shown in Figure 4.

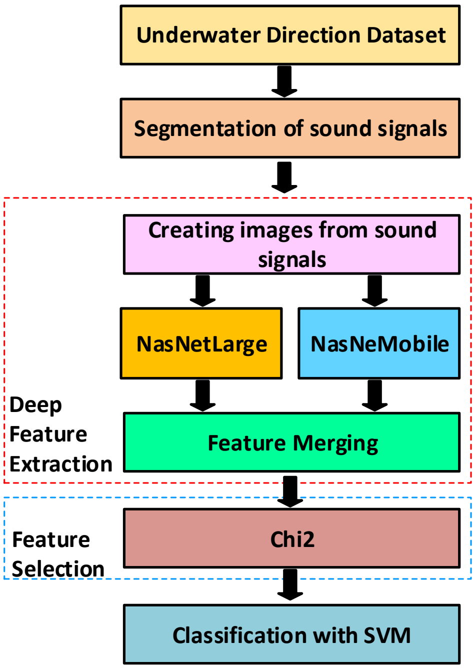

Fig. 4. Block diagram of the proposed method

In this study, a four-class UDD was created using an underwater robot and a sound recorder. A total of 300 seconds of voice recording was taken for each class. The recorder records 44100 samples every 1 second. There are 1200 sound data for four classes in total. These data consist of a 1200x44100 matrix. The sound data obtained are divided into one-second segments. The sound signal received in one second is 1x44100. This sound signal has been normalized between 0-255 and converted into 200x220 sized images. These images have been used with NASNetLarge and NASNetMobile models and feature extracted. For this reason, 200x220 sized images are arranged in 331x331 size. Thus, the size of the images is suitable for NASNetLarge and NASNetMobile models. Sample images obtained for four classes in this study are shown in Figure 5.

As can be seen in Figure 5, a total of 1200 images were created, 300 for each class. Feature extraction was performed on these images using NASNetLarge and NASNetMobile models. NASNet was developed by the Google Brain team for the CIFAR-10 dataset [16,17]. The NASNet model is advantageous due to its small size and low complexity. NASNet model uses the reinforcement learning method. In this study, feature extraction has been made using both NASNetLarge and NASNetMobile models. 1200x1000 feature extraction has been made with each model. 1200x2000 features were obtained by combining the features obtained from these models. Chi2 method was used to select the most significant features from these features. The Chi2 method is a widely used feature selection method [18]. The mathematical equivalent of the Chi2 algorithm is given in Equation 1.

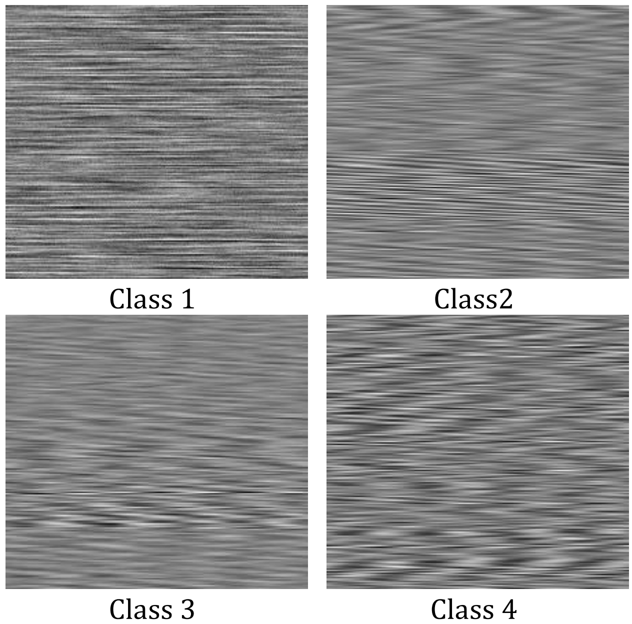

Fig. 5. Images created from the underwater direction dataset

$$x_c^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{1}$$

In the Chi2 method, the observed O_i and expected E_i values are handled, and the degrees of freedom c of the values are determined. After the deep feature extraction, the Chi2 method was applied to the 1200x2000 feature matrix, and the 1200x944 feature was selected. The most weighted features obtained were classified using the SVM algorithm. SVM algorithm is a machine learning method widely used in the literature [18–23]. In this study, results are obtained by using Linear, Cubic, and Quadratic models of the SVM algorithm. The parameters of SVM algorithms used in the proposed method are given in Table 2.

Table 2. Parameters of SVM Algorithms Used in the Proposed Method

	Linear SVM	**Quadratic SVM**	**Cubic SVM**
Kernel function	Linear	Quadratic	Cubic
Box constraint level	5	1	1
Kernel scale mode	Auto	Auto	Auto
Kernel scale	1	1	1
Multiclass method	One-vs-one	One-vs-one	One-vs-one

4 Experimental Results

To apply the method suggested in this study, a computer with an i7-9700 CPU 3.00 GHz, 32GB RAM, and 64 operating systems were used. The proposed method was developed in the MATLAB 2020a program. MATLAB Classification Learner Toolbox was used for the classification process. Confusion matrices were calculated using 10-fold cross-validation. In the proposed method, 100 iterations were run to get results with SVM classification methods. Confusion Matrixes obtained for Linear SVM, Cubic SVM, and Quadratic SVM are given in Figure 6.

The results given in Figure 6 were examined, and it was shown that the best result among the three classifiers was calculated in Class 1. Accuracy, Precision, Recall, Geometric mean, and F1-Score results calculated for Linear SVM, Cubic SVM, and Quadratic SVM are given in Table 3.

True Class \ Predicted Class	1	2	3	4
1	**299**	**0**	**0**	**1**
2	**0**	**214**	**55**	**31**
3	**0**	**65**	**203**	**32**
4	**0**	**38**	**46**	**216**

Linear SVM

True Class \ Predicted Class	1	2	3	4
1	**298**	**0**	**1**	**1**
2	**0**	**207**	**65**	**28**
3	**0**	**65**	**199**	**36**
4	**0**	**32**	**43**	**225**

Cubic SVM

True Class \ Predicted Class	1	2	3	4
1	**297**	**2**	**0**	**1**
2	**0**	**214**	**62**	**24**
3	**0**	**70**	**199**	**31**
4	**0**	**31**	**48**	**221**

Quadratic SVM

Fig. 6. Confusion Matrix obtained for Linear SVM, Cubic SVM and Quadratic SVM

Table 3. Accuracy, Precision, Recall, Geometric Mean and F1-Score (%) Results of the Used Direction Underwater Acoustic Classifiers

Classification Statistics		**Accuracy**	**Precision**	**Recall**	**Geometric mean**	**F1-Score**
Linear	**Max**	**77.66**	77.85	**77.66**	**76.71**	**77.76**
	Min	74.41	75.90	75.75	74.58	75.82
	Mean	75.95	77.46	77.27	76.27	77.36
	Std	0.56	0.33	0.33	0.36	0.33
Cubic	**Max**	77.41	77.57	77.41	76.41	77.49
	Min	73.66	75.37	75.16	73.98	75.27
	Mean	75.19	77.32	77.15	76.12	77.23
	Std	0.63	0.47	0.50	0.54	0.49
Quadratic	**Max**	77.58	**77.92**	77.58	76.64	77.75
	Min	74.33	74.57	74.33	73.02	74.45
	Mean	76.10	76.33	76.10	74.98	76.22
	Std	0.55	0.54	0.55	0.61	0.54

When Table 3 is examined, the best results were obtained with Linear SVM. The accuracy was 77.66% with Linear SVM, 77.41% with Cubic SVM, and 77.58% with Quadratic SVM.

5 Conclusion and Discussions

Underwater acoustic systems, methods such as depth detection, object detection, and object tracking have been developed. Underwater acoustic methods are significant for defense technology. In this study, a deep learning and machine learning-based hybrid method are proposed for underwater direction detection. First, UDD was collected by using an underwater robot. Feature extraction is made by doing deep feature extraction on this dataset. The most weighted features were selected from the obtained features and classified with SVM. In the proposed method, the best accuracy was calculated as 77.66%. In this study, 10-fold cross-validation was used while classifying. In the proposed method, Fold by Fold results for Linear SVM, Cubic SVM, and Quadratic SVM are calculated and presented in Figure 7.

As seen in Figure 7, the results of Fold-1, Fold-2, Fold-3, ..., Fold-10 have been calculated for classification algorithms. The highest result was seen in Fold-1, and the lowest result was seen in Fold-5. Since 10-fold cross-validation is used in the proposed method, the best result was calculated as 77.66%. Also, the recommended method Class by Class results are shown in Figure 8.

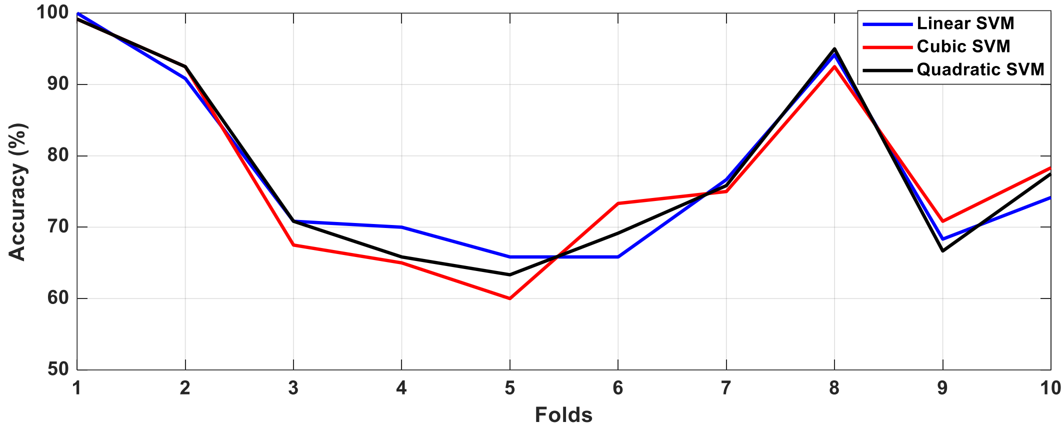

Fig. 7. Plot of best accuracy (%) obtained for different folds using UDD

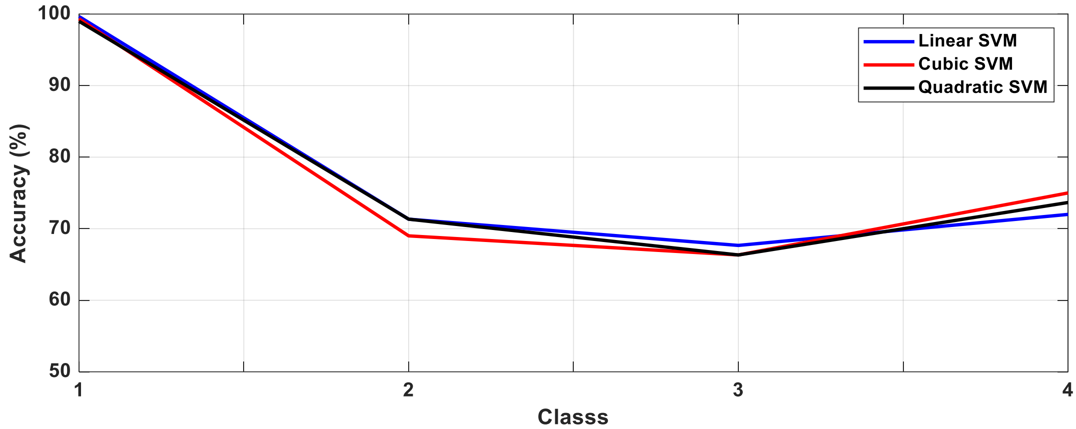

Fig. 8. Classification accuracies (%) obtained for various classes using our proposed method with UDD

As seen in Figure 8, the best results were obtained in Class 1, while the lowest results were seen in Class 3. Class by Class results support the accuracy of the confusion matrix obtained in the study.

Acknowledgment

This work is supported by Firat University Research Fund, Turkey Project Number: MMY.20.01.

References

[1] Choi J, Park J, Lee Y, Jung J, Choi HT. Robust directional angle estimation of underwater acoustic sources using a marine vehicle. Sensors (Switzerland) 2018;18. https://doi.org/10.3390/s18003062.

[2] Koseoglu M, Bereketli A, Yazgi I, Yeni B. Probabilistic broadcast for dense AUV networks. Ocean 2016 MTS/IEEE Monterey, OCE 2016 2016. https://doi.org/10.1109/OCEANS.2016.7761118.

[3] Cho H, Gu J, Joe H, Asada A, Yu SC. Acoustic beam profile-based rapid underwater object detection for an imaging sonar. J Mar Sci Technol 2015;20:180–97. https://doi.org/10.1007/s00773-014-0294-x.

[4] Lee H, Jung HK, Cho SH, Kim Y, Rim H, Lee SK. Real-Time Localization for Underwater Moving Object Using Precalculated DC Electric Field Template. IEEE Trans Geosci Remote Sens 2018;56:5813–23. https://doi.org/10.1109/TGRS.2018.2826556.

[5] Nie D, Sun Z, Qiao G, Liu S, Yin Y. Kite-type passive acoustic detection system for underwater small targets. 2014 Ocean - St John's, Ocean 2014 2015. https://doi.org/10.1109/OCEANS.2014.7003207.

[6] Isbitiren G, Akan OB. Three-dimensional underwater target tracking with acoustic sensor networks. IEEE Trans Veh Technol 2011;60:3897–906. https://doi.org/10.1109/TVT.2011.2163538.

[7] Jiang J, Wu Z, Lu J, Huang M, Xiao Z. Interpretable features for underwater acoustic target recognition. Meas J Int Meas Confed 2020:108586. https://doi.org/10.1016/j.measurement.2020.108586.

[8] Neves G, Ruiz M, Fontinele J, Oliveira L. Rotated object detection with forward-looking sonar in underwater applications. Expert Syst Appl 2020;140:112870. https://doi.org/10.1016/j.eswa.2019.112870.

[9] Reis CDG, Padovese LR, de Oliveira MCF. Automatic detection of vessel signatures in audio recordings with spectral amplitude variation signature. Methods Ecol Evol 2019;10:1501–16. https://doi.org/10.1111/2041-210X.13245.

[10] Choi J, Choi HT. Multi-target localization of underwater acoustic sources based on probabilistic estimation of direction angle. MTS/IEEE Ocean 2015 - Genova Discov Sustain Ocean Energy a New World 2015. https://doi.org/10.1109/OCEANS-Genova.2015.7271437.

[11] Sierra E, Contreras J. Classification of small boats using fuzzy classifier. Annu Conf North Am Fuzzy Inf Process Soc - NAFIPS 2015;2015-Septe:0–4. https://doi.org/10.1109/NAFIPS-WConSC.2015.7284174.

[12] Fischell EM, Viquez O, Schmidt H. Passive acoustic tracking for behavior mode classification between surface and underwater vehicles. IEEE Int Conf Intell Robot Syst 2018:2383–8. https://doi.org/10.1109/IROS.2018.8593981.

[13] Santos-Domínguez D, Torres-Guijarro S, Cardenal-López A, Pena-Gimenez A. ShipsEar: An underwater vessel noise database. Appl Acoust 2016;113:64–9. https://doi.org/10.1016/j.apacoust.2016.06.008.

[14] Sutin A, Bunin B, Sedunov A, Sedunov N, Fillinger L, Tsionskiy M, et al. Stevens passive acoustic system for underwater surveillance. 2010 Int Waterside Secur Conf WSS 2010 2010. https://doi.org/10.1109/WSSC.2010.5730286.

[15] CHASING | GLADIUS MINI - 4K Underwater Drone with Camera - Chasing Innovation n.d. https://www.chasing.com/gladius-mini.html (accessed November 27, 2020).

[16] Saxen F, Werner P, Handrich S, Othman E, Dinges L, Al-Hamadi A. Face attribute detection with mobilenetv2 and nasnet-mobile. Int Symp Image Signal Process Anal ISPA 2019;2019-Septe:176–80. https://doi.org/10.1109/ISPA.2019.8868585.

[17] Zoph B, Vasudevan V, Shlens J, Le Q V. Learning Transferable Architectures for Scalable Image Recognition Barret. Proc IEEE Conf Comput Vis Pattern Recognit 2018:8697–710.

[18] Baygin N, Baygin M, Karakose M. A SVM-PSO Classifier for Robot Motion in Environment with Obstacles. 2019 Int. Conf. Artif. Intell. Data Process. Symp. IDAP 2019, 2019. https://doi.org/10.1109/IDAP.2019.8875921.

[19] Tuncer T, Ertam F. Neighborhood component analysis and reliefF based survival recognition methods for Hepatocellular carcinoma. Phys A Stat Mech Its Appl 2020;540:123143. https://doi.org/10.1016/j.physa.2019.123143.

[20] Tsanas A, Little MA, McSharry PE, Spielman J, Ramig LO. Novel speech signal processing algorithms for high-accuracy classification of Parkinsons disease. IEEE Trans Biomed Eng 2012;59:1264–71. https://doi.org/10.1109/TBME.2012.2183367.

[21] Singh M, Shaik AG. Faulty bearing detection, classification and location in a three-phase induction motor based on Stockwell transform and support vector machine. Meas J Int Meas Confed 2019;131:524–33. https://doi.org/10.1016/j.measurement.2018.09.013.

[22] Yaman O. An automated faults classification method based on binary pattern and neighborhood component analysis using induction motor. Meas J Int Meas Confed 2021;168:108323. https://doi.org/10.1016/j.measurement.2020.108323.

[23] Lahmiri S. Parkinson's disease detection based on dysphonia measurements. Phys A Stat Mech Its Appl 2017;471:98–105. https://doi.org/10.1016/j.physa.2016.12.009.