Deep Learning Inspired Acoustic-based Underwater Object Detection and Localization

Rubina Akter[†], Ahmad Zainudin*, and Dong-Seong Kim[§]
[†]ICT Convergence Research Center, Kumoh National Institute of Technology, Gumi 39177, South Korea
*Department of Electronic Engineering, Kumoh National Institute of Technology, Gumi 39177, Korea
[§]Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi 39177, Korea
(rubinaakter2836[†], zai*, dskim[§])@kumoh.ac.kr

***Abstract*—Detection and tracking of underwater objects have a wide range of uses. Due to the intrinsic instability of acoustic communications, Underwater Acoustic Sensor Networks (UASNs) face challenges in detecting and localizing underwater objects. To solve these problems, novel architectures and protocols are required for the efficient drive of UASNs. The most fundamental UASNs duties are object detection and tracking, which are necessary for target detection, node tracking, and data tagging, and can be utilized to enhance medium access performance and protocols for networks. Therefore, this study presents Convolution Neural Network (CNN) inspired underwater object detection and localization using acoustic signals. In this paper, the MobileNetV2 network is used to classify underwater mines and the Direction of Arrival (DoA) technique is applied to estimate the location of the object.**

***Index Terms*—Convolution neural network, direction of arrival, mine classification, target tracking, Underwater acoustic sensor networks.**

I. Introduction

There are various civil and military uses for the detection, localization [1], and tracking of underwater objects, including ocean environmental monitoring, offshore oil leak detection, underwater rescue, and many more. Hence, this is essential to develop an efficient future internet of underwater things (IoUT). The underwater nodes can be used to maintain item tracking in sizable water expanses since they organize into arrays. Passive localization and tracking systems can be used by employing direction of arrival (DoA) [2] if the targets emit sound. Typically, static position-aware nodes work as anchors for the localization and tracking of objects. Recently, using autonomous underwater vehicles (AUVs) as mobile anchors has gained popularity since the AUV-based architecture offers better and more flexible coverage for a significantly reduced price.

Typically, static position-aware nodes work as anchors for the localization and tracking of objects. Recently, using Autonomous Underwater Vehicles (AUVs) as mobile anchors has gained popularity since the AUV-based architecture offers better and more flexible coverage for a significantly reduced price [3]. However, passive sonar systems are used to detect and track the surrounding objects in Underwater Acoustic System (UAS). But the sonar systems face challenges in working at a low signal-to-noise ratio (SNR) and multipath propagation. As a result, a Machine Learning (ML)-based approach has been developed to solve the sonar-based systems' limitations. Importantly, recently, Convolution Neural Networks (CNN) [4] have been studied to analyze the propeller acoustic signals and extract the features for underwater mining detection and localization. However, most of the existing literature adopts CNN for either detection or localization purposes, instead of serving both purposes. But an efficient UAS requires a framework for underwater object detection and localization to ensure the security measures of the under-water surveillance system. Motivated by the aforementioned issue, this paper proposed a multitasking scheme for underwater acoustic object detection and tracking using MobileNetV2.

II. Underwater Object Detection and Localization Using MobileNetV2

This study uses two acoustic datasets for mine classification and localization, respectively. The object detection dataset comprises eleven types of acoustic signals gathered from underwater objects. And for the object localization, we created a synthetic dataset labeled as multiclass to estimate DOA. To create the dataset, a uniform linear array (ULA) was configured with five microphones, where the enter distance of each sensor was 10 cm. $-60°$ to $60°$ angles were considered with an angle gap of $1°$. These two data-sets were fed into the MobileNetV2 network, which is one of the light wight pretrained networks in CNN. The detail architecture of the MobileNetV2 is shown in Fig. 1. The input acoustic signals are processed with three key processing blocks of the MobileNet, named the Stem, Stack, and the ResidualStack. The stem is configured with one 3×3 convolution layer, three 1×1 convolution layers, and two 3×3 Depthwise convolution layers. This block extracts the common features of the acoustic signals. In this network, ten ResidualStacks are incorporated, where each block is comprised of two 1×1 convolution layers, one 3×3 Depthwise convolution layer, and a residual connectivity. Moreover, MobileNetV2 is comprised of five stack connections. Three 1×1 convolution layers, and one 3×3 Depthwise convolution layer are sequentially stacked in this block. The ResidualStack and Stack extract intermediate features of the acoustic signals that significantly contribute to the acoustic object classification and angle estimation of the corresponding objects.

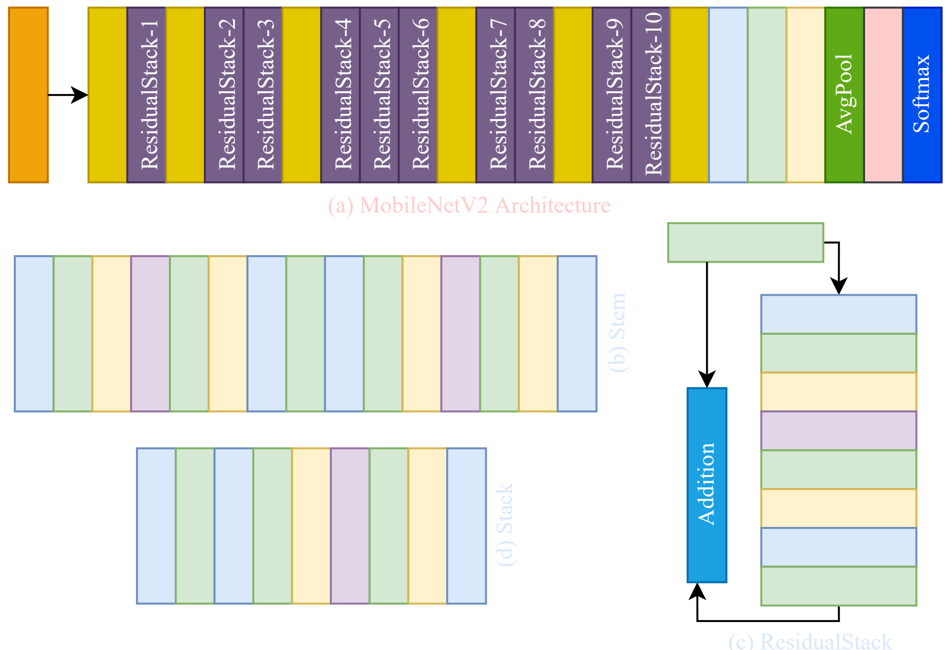

Fig. 1. Network configuration of MobilNetV2.

Confusion Matrix

True Class \ Predicted Class	Noi	T01	T02	T03	T04	T05	T06	T07	T08	T09	T10	T11		
Noi	3629		45	1	80	36	25	11	77				93.0%	7.0%
T01	487		48	3		3				3	57	180		100.0%
T02	293		3217	36	27	66	108	83	146	17	115	71	77.0%	23.0%
T03	4			3570	95	113	4	133	25		148	46	86.3%	13.7%
T04	159		7	116	3512	96	40	86	81		8	26	85.0%	15.0%
T05	95		33	481	179	3057	21	89	111	1	42	25	73.9%	26.1%
T06	56		78	12	80	21	3708	39	77		2	68	89.5%	10.5%
T07	3		10	315	50	29	10	3479	45		149	44	84.2%	15.8%
T08	116		91	50	161	102	64	148	3338		6	14	81.6%	18.4%
T09	84		21	4	2		1	17	7	892	22		85.0%	15.0%
T10			9	163	3	7		178	1	2	3698	77	89.4%	10.6%
T11			16	59	38	3	146	54	10	6	112	3690	89.3%	10.7%

Fig. 2. Confusion matrix of the 11-types of acoustic signal classification..

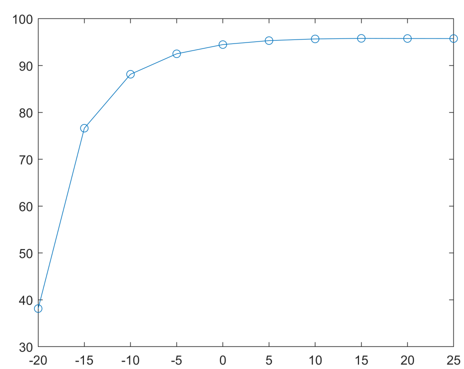

Fig. 3. DoA estimation accuracy of MobileNetV2.

III. Experimental Results

Fig. 2. shows the confusion matrix of the acoustic signal classification using MobileNetV2. In this experiment, the MobileNet acquired 83.6% classification accuracy. The detection ratio drops due to the difficulty of identifying target T01 objects, as shown in the confusion matrix. Fig. 3. shows the angle classification accuracy of the MobileNetV2 at various SNR levels. The simulation results show that the angle classification increases in terms of the increment of the SNR levels. At $5dB$ SNR, MobileNet achieves 94.2% DoA estimation accuracy.

IV. Conclusion

This paper presents acoustic-based underwater acoustic signal classification and DOA estimation for object tracking using MobileNetV2. MobileNetV2 shows efficiencies in both object detection and localization. However, due to the deep design strategies of MobileNetV2, this network requires high computational expense. In the future, we will propose lightweight CNN for both acoustic-based object detection and localization for UAS.

Acknowledgment

This research was supported by the MSIT (Ministry of Science and ICT), and Priority Research Centers Program, Korea, under the Grand Information Technology Research Center support program (IITP-2022-2020-0-01612) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation) and the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2018R1A6A1A03024003), respectively.

REFERENCES

[1] R. Akter, V.-S. Doan, T. Huynh-The, and D.-S. Kim, "Rfdoa-net: An efficient convnet for rf-based doa estimation in uav surveillance systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 11, pp. 12 209–12 214, 2021.

[2] R. Akter, M. Golam, V.-S. Doan, J.-M. Lee, and D.-S. Kim, "IoMT-Net: Blockchain Integrated Unauthorized UAV Localization Using Lightweight Convolution Neural Network for Internet of Military Things," *IEEE Internet of Things Journal*, 2022.

[3] M. Erol-Kantarci, H. T. Mouftah, and S. Oktug, "A survey of architectures and localization techniques for underwater acoustic sensor networks," *IEEE communications surveys & tutorials*, vol. 13, no. 3, pp. 487–502, 2011.

[4] R. Akter, V.-S. Doan, J.-M. Lee, and D.-S. Kim, "CNN-SSDI: Convolution neural network inspired surveillance system for UAVs detection and identification," *Computer Networks*, vol. 201, p. 108519, 2021.