

신경망 기반 분포 외 검출기를 활용한 수중 음향 환경의 이상 음원 탐지 방법

김 나 연*, 전 찬 준*, 김 홍 국[○]

Novelty Detection in Underwater Acoustic Environments Using Out-of-Distribution Detector for Neural Networks

Nayeon Kim*, Chanjun Chun*,
Hong Kook Kim[○]

요 약

본 연구에서는 수중 음향 환경에서 학습되지 않은 미지의 음향 신호를 효과적으로 탐지하기 위한 ODIN 기반의 이상 음원 탐지 프레임워크를 제안한다. 구체적으로, 사전 학습된 분류기의 소프트맥스 출력에 온도 스케일링과 입력 교란을 적용하여 훈련된 클래스와 미지의 클래스간의 출력 차이를 유도하고, 최대 소프트맥스 확률을 이상 점수로 활용하여 이상 탐지를 수행한다.

키워드 : 이상 음원 탐지, 신경망 기반 분포 외 검출기, 온도 스케일링, 입력 교란

Key Words : Novelty detection, Out-of-Distribution Detector for Neural Networks, Temperature scaling, Input perturbation

ABSTRACT

In this paper, we propose an ODIN-based novelty detection framework to effectively identify unknown

acoustic signals in underwater environments. Specifically, temperature scaling and input perturbation are applied to the softmax output of a pre-trained classifier to induce differences between known and unknown samples, and the calibrated maximum softmax probability is used as a novelty score to perform novelty detection.

I. 서 론

수중 음향 환경은 날씨, 조류, 수온, 계절 변화 등으로 인해 매우 동적이고 예측이 어렵기 때문에 정밀한 감지와 분석이 어렵다^[1]. 특히 실시간 감지 시스템에서는 기존에 학습되지 않은 새로운 음향 신호가 출현할 가능성이 높고, 이는 위협 요소나 이상 상황을 나타낼 수 있다. 이러한 환경에서 널리 사용되는 수동 소나 시스템은 하이드로폰을 통해 음향을 수집하며, 자체 신호를 방사하지 않아 보안성이 높다는 장점이 있다. 이와 같은 수동 소나 데이터를 기반으로 다양한 표적 인식 기법이 제안되어 왔으며, 최근에는 CNN 및 트랜스포머 기반의 딥러닝 모델이 높은 성능을 보여주고 있다^[2,3].

그러나 위 연구는 학습된 클래스만을 대상으로 하며, 학습되지 않은 새로운 신호가 입력되었을 때 기존 클래스 중 하나로 잘못 분류하는 문제가 있다. 이를 해결하기 위해 이상 탐지 기법(novelty detection)에 대한 연구가 진행되고 있으며, 계층적 LSTM 오토인코더나 stacked autoencoder를 활용한 방법이 제안되었다^[4,5]. 하지만 클래스 수 증가에 따른 복잡도 증가, 임계값 설정의 민감도 등 여러 한계를 지닌다. 또 다른 방법으로는 MSP(Maximum Softmax Probabilities) 기법이 있으나, 이는 과신(over-confidence)문제로 인해 분포 외 데이터를 정확히 탐지하는 데 한계가 있다^[6].

본 연구에서는 이러한 한계를 극복하기 위해 ODIN(Out-of-Distribution Detector for Neural Networks) 기법을 수중 음향 환경에 적용한 이상 탐지 방법을 제안한다. ODIN은 추론 시 소프트맥스 출력에 온도 스케일링을 적용하고, 입력에 미세한 교란을 추가

※ 이 논문은 2023년도 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원을 받아 수행된 연구임(No. 21-107-B00-008-02(KRIT-CT-23-009-02), 원해 수중감시용 소형 가변심도 소나시스템 기술).

• First Author : (ORCID:0009-0008-6991-5539) I Graduate School, Gwangju Institute of Science and Technology (GIST), nayeunk1117@gm.gist.ac.kr, 학생회원

○ Corresponding Author : (ORCID:0000-0002-0105-6693) School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology (GIST), hongkook@gist.ac.kr, 종신회원

* Dept. of Computer Engineering, Chosun University, cjchun@chosun.ac.kr, 종신회원

논문번호 : 202507-170-A-LU, Received July 23, 2025; Revised August 7, 2025; Accepted August 7, 2025

하여 분포 내/외 샘플 간의 출력 차이를 유도한다^[7]. 이후 소프트맥스 최대 출력값에 임계값을 적용함으로써 이상 신호를 효과적으로 탐지할 수 있으며, 이를 통해 수중 환경에서의 미지 신호에 대한 탐지 성능을 향상시키고자 한다.

II. ODIN 기반 이상 음원 탐지

본 연구에서는 수중 음향 환경에서 학습에 포함되지 않은 신호를 효과적으로 탐지하기 위해 ODIN 기반의 이상 음원 탐지 프레임워크를 제안한다. 제안하는 방법은 그림 1의 (1) Attention ResNet(AResNet) 분류 모델을 학습하는 과정과^[8], (2) 온도 스케일링과 입력 교란을 포함하는 ODIN 기법을 통해 얻은 보정된 소프트맥스 출력값 중 최대값을 이상 점수로 사용하여 이상 여부를 판단하는 추론 과정으로 구성된다. 이상 점수가 설정된 임계값 이상이면 입력 신호는 학습된 클래스 중 하나로 간주되며, 분류기의 예측 결과를 따른다. 반대로 이상 점수가 임계값 미만이면 입력은 이상 음원으로 탐지된다.

추론 과정에서 사용되는 온도 스케일링 계수 T 와 입력 교란 크기 ϵ 은 각각 $T = \{1, 2, 5, 10, 50, 100, 500, 1000\}$, $\epsilon = \{0, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}\}$ 의 값에 대해 실험을 수행하였다.

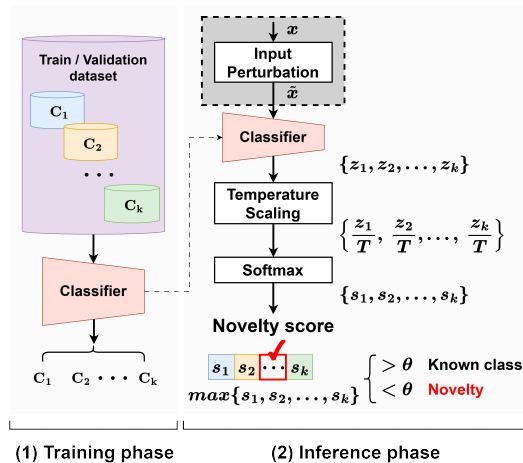


그림 1. 제안하는 ODIN 기반 이상 탐지 방법의 개요도
Fig. 1. Scheme of the proposed ODIN based novelty detection method

III. 실험 결과 및 분석

본 논문의 검증을 위해 4종류로 분류된 선박 운항시

의 소음 데이터를 포함하는 DeepShip 공개 데이터셋을 사용하였다^[9]. 학습/검증/테스트를 7/2/1 비율로 나누었으며, mel-frequency cepstral coefficient(MFCC)를 입력 피쳐로 사용하였다.

이상 탐지 모델의 성능 평가를 위해 AUROC, FPR@95TPR, AUPR-In, AUPR-Out을 사용하였다. AUROC는 이상 점수를 기준으로 다양한 임계값에서의 TPR(True Positive Rate)과 FPR(False Positive Rate)을 계산해 ROC 곡선 아래 면적을 측정한 값으로, 전반적인 분류 성능을 나타낸다. FPR@95TPR은 TPR이 95%일 때의 FPR 값을 의미하며, 미지 클래스를 학습된 클래스와 혼동하는 비율을 나타낸다. AUPR-In과 AUPR-Out은 각각 학습된 클래스와 미지 클래스를 양성으로 간주하여 계산한 precision-recall 곡선의 면적으로, 정상 및 이상 클래스에 대한 탐지 성능을 정량화한 지표이다.

표 1은 제안한 ODIN과 MSP의 이상 탐지 성능을 각 클래스에 대해 각각 미지 클래스로 설정하여 비교한 결과이다. FPR@95TPR이 가장 낮았던 $T = 10, \epsilon = 10^{-5}$ 에서, ODIN은 모든 지표에 대해 클래스별 샘플 수를 고려한 평균 성능값이 MSP 대비 우수한 결과를 보였다. 특히 AUPR-In의 성능저하 없이 AUPR-Out에서 2.75%의 성능 향상이 나타나, 제안한 방법이 정상 클래스 분류 성능을 유지하면서도 미지 클래스 탐지 능력을 효과적으로 향상시켰음을 알 수 있다.

표 1. 이상 신호 탐지 성능 비교(MSP/ODIN)
Table 1. Novelty detection performance(MSP/ODIN)

Novelty label	AUROC ↑	FPR@95 TPR ↓	AUPR-In ↑	AUPR-Out ↑
Cargo	76.61	73.77	91.47	47.57
	78.94	66.23	92.35	54.62
Passengership	71.36	81.80	86.58	47.15
	73.61	77.47	88.01	51.63
Tanker	73.36	76.94	91.62	38.51
	73.49	79.35	91.74	37.66
Tug	79.00	72.03	86.59	63.98
	79.35	72.93	86.77	64.55
Average	75.42	75.92	88.52	51.39
	76.66	73.88	89.18	54.14

IV. 결 론

본 연구에서는 수중 음향 환경에서 학습되지 않은 미지의 음향 신호를 효과적으로 탐지하기 위해 ODIN

기반의 이상 음원 탐지 프레임워크를 제안하였다. 제안한 방법이 모든 지표에서 기존 방법 대비 우수한 성능을 나타내었다. 이러한 성능 향상은 온도 스케일링을 통해 소프트맥스 출력의 과신을 완화하고, 입력 교란을 통해 분포 내/외 샘플 간의 출력 차이가 효과적으로 보정된 결과로 볼 수 있다. 이는 제안한 방법이 정상 클래스의 분류 성능을 유지하면서도 미지의 수중 음향 신호를 보다 안정적으로 탐지할 수 있는 가능성을 보여주며, 제안한 방법의 실용성과 적용 가능성을 입증하였다.

References

- [1] Z. Lei, X. Lei, N. Wang, and Q. Zhang, "Present status and challenges of underwater acoustic target recognition technology: A review," *Front. Phys.*, vol. 10, p. 1044890, Oct. 2022.
(<https://doi.org/10.3389/fphy.2022.1044890>)
- [2] B.-K. Kim, W. Chang, H. Kim, S. Kim, D. Kang, M. Kim, D. Kang, and S. Cho, "CNN-based shipping noise detection using short-time underwater acoustics signal," *J. Inst. Electron. Inf. Eng. (IEIE)*, vol. 60, no. 3, pp. 61-68, Dec. 2023.
(<https://doi.org/10.5573/jieie.2023.60.3.61>)
- [3] S. Feng and X. Zhu, "A transformer-based deep learning network for underwater acoustic target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1-5, 2022.
(<https://doi.org/10.1109/LGRS.2022.3201396>)
- [4] E. S. Honorato, J. B. O. e S. Filho, and V. H. S. Muniz, "A hierarchical ensemble of LSTM-based autoencoders for novelty detection in passive sonar systems," in *Proc. 2021 IEEE Latin Am. Conf. Comput. Intell. (LA-CCI)*, pp. 1-6, Temuco, Chile, November 2021.
(<https://doi.org/10.1109/LA-CCI48322.2021.9769821>)
- [5] V. dos Santos Mello, N. N. de Moura, and J. M. de Seixas, "Novelty detection in passive sonar systems using stacked autoencoders," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, pp. 1-7, Rio de Janeiro, Brazil, Jul 2018.
(<https://doi.org/10.1109/IJCNN.2018.8489559>)
- [6] D. Hendrycks and K. Gimpel, "A baseline for detecting misclassified and out-of-distribution examples in neural networks," in *Proc. Int. Conf. Learn. Represent. 2017 (ICLR)*, pp. 1-10, Toulon, France, Apr. 2017.
(<https://doi.org/10.48550/arXiv.1610.02136>)
- [7] S. Liang, Y. Li, and R. Srikant, "Enhancing the reliability of out-of-distribution image detection in neural networks," in *Proc. Int. Conf. Learn. Represent. 2018 (ICLR)*, pp. 1-13, Vancouver, Canada, May 2018.
(<https://doi.org/10.48550/arXiv.1706.02690>)
- [8] J. Li, B. Wang, X. Cui, S. Li, and J. Liu, "Underwater acoustic target recognition based on attention residual network," *Entropy*, vol. 24, no. 11, pp. 1-16, Aug. 2023.
(<https://doi.org/10.3390/e24111657>)
- [9] M. Irfan, Z. Jiangbin, S. Ali, M. Iqbal, Z. Masood, and U. Hamid, "DeepShip: An underwater acoustic benchmark dataset and a separable convolution based autoencoder for classification," *Expert Syst. with Appl.*, vol. 183, no. C, 115270, Nov. 2021.
(<https://doi.org/10.1016/j.eswa.2021.115270>)