



Received: 2023/08/25  
Revised: 2023/09/06  
Accepted: 2023/09/26  
Published: 2023/09/30

**\*Corresponding Author:**

**Dooyoung Kim**

Dept. of Artificial Intelligence, Republic of Korea  
Naval Academy  
Jungwon-ro, Jinhae-gu, Changwon-si,  
Gyungsangnam-do, 51704, Republic of Korea  
Tel: +82-55-907-5246  
E-mail: dykim07@navy.ac.kr

# 선박소음의 자동 식별을 위한 원샷 학습기법

## Use of One-shot Learning for Automatic Target Recognition of Surface Ship Noise

김두영\*

해군소령/해군사관학교 인공지능학과 부교수

Dooyoung Kim\*

LCDR, ROK Navy/Associate Professor, Dept. of Artificial Intelligence,  
Republic of Korea Naval Academy

**Abstract**

본 논문은 수중에 방사되는 선박 소음을 자동으로 식별하기 위한 원샷 학습기법을 제시한다. 제안한 학습기법은 기존의 지도학습 기반의 딥 러닝 기법이 가지는 학습데이터의 의존성을 극복하여, 단 하나의 신호 표본만으로 동일 선박을 식별할 수 있는 모델을 공개 데이터 기반의 실험을 통해 성능을 검증한 결과 후보군이 다섯 척인 상황에서 72.5%의 정확도로 대상 선박의 선명을 예측함을 확인하였다.

This paper presents a one-shot learning technique for automatically identifying ship noise emitted underwater. The proposed learning technique overcomes the dependence of the learning data of the existing supervised learning-based deep learning technique and a model that can identify the same ship with only a single signal sample. As a result of verifying the performance through open data-based experiments, it was confirmed that the clarity of the target ship was predicted with an accuracy of 72.5% in the situation of five candidates.

**Keywords**

자동 표적 인지(Automatic Target Recognition),  
메타 학습(Meta Learning),  
원샷 학습(One-shot Learning),  
선박 소음(Ship Noise)

**Acknowledgement**

이 논문은 2023년 해군사관학교 해양연구소  
학술연구과제 연구비 지원으로 수행된 연구임.

### 1. 서론

선박은 고유의 방사 소음 특성을 가진다[1]. 이는 선박에 장착된 엔진, 프로펠러, 추진축 등 추진체계와 조수기, 발전기 등의 보기류에서 발생하는 다양한 소음이 중첩되어 만들어진다. 이러한 방사 소음은 수중에서 활동하는 잠수함이 대상을 식별할 수 있는 유일한 수단이다[1,2]. 따라서 각국의 해군은 군함뿐만 아니라 다양한 선박의 방사 소음을 항시 수집하고 있다.

이러한 방사 소음의 특성을 이용하여 선박의 종류를 자동으로 식별(automatic target recognition, ATR)하려는 다양한 연구가 진행되었다. 복잡한 해양 환경을 극복하기 위해 빔 형성 기법, 표적 위치 추정, 다중경로 상 음파 제거 등의 다양한 신호처리 기법을 적용한 연구가 진행되어 높은 성능을 보여주었다[3-5].

한편, 인공지능(artificial intelligence, AI)과 기계 학습(machine learning) 기술의 발전은 선박 식별 기술의 성능 향상에 많은 이바지를 하고 있다. 전통적인 기계 학습을 이용한 음향 식별 알고리즘[6]에서 더욱 발전한 딥러닝 기반의 전방 신경망(feed-forward neural network, FNN), 합성곱 신경망(convolutional neural network, CNN) 등 다양한 모델이 제시되었으며[7-10], 이를 학습하기 위한 다수의 학습데이터 집합(data set)이 공개되었다[11,12].

그러나, 최근 발표된 딥러닝 기반 모델 다수가 뛰어난 성능을 보여주고 있음에도 불구하고, 지도학습(supervised learning) 또는 준지도학습(semi-supervised learning)[13] 기반의 기존 모델은 현실 세계에 바로 적용되기 어렵다[14,15]. 이는 모델의 학습을 위해 모든 분

류 대상에 대한 대량의 학습데이터 집합이 필요하고, 학습 이후 실제 운용 시 학습데이터 집합에 빠진 새로운 선박의 신호는 식별할 수 없기 때문이다[15]. 따라서 수중 방사 소음 수집이 쉽지 않은 현재의 환경에서 딥러닝 기반의 식별 알고리즘이 현장에 적용되기 위해서는 이러한 제약사항이 반드시 해결되어야 한다.

본 논문에서는 이러한 문제를 해결하기 위해 메타 학습(meta learning)[16]의 일종인 원샷 학습(one-shot learning)[14] 기반의 선박 소음 자동 식별 기법을 제안한다. 기존 지도학습 기반의 알고리즘이 주어진 소음 신호를 기반으로 각 선박 집합을 구분하기 위한 “경계”를 학습하는 반면에, 제안한 기법은 서로 다른 입력 신호를 구분하는 “방법”을 학습한다. 이를 통해 새로운 선박의 소음 신호가 단 하나만 존재하는 상황(one-shot)에서 해당 선박을 식별할 수 있다. 이와 더불어 기존 연구의 목표는 “여객선”, “예인선” 등 선박의 종류를 식별하는 것이었으나, 제안한 방법은 선박의 “선명”을 예측하는 것을 목표로 하였다. 제안한 모델의 성능검증을 위해 공개된 선박 방사 소음데이터를 이용하여 실험한 결과 대상 선박 후보군이 다섯 척인 상황에서 약 72.5%의 정확도로 동일 선박을 식별하였다.

본 논문의 구성은 다음과 같다. 먼저 2장에서 본 논문이 제안하는 선박 소음 식별에 특화된 원샷 학습 기법을 서술한다. 3장에서는 성능검증을 위한 데이터 집합 및 실험환경 구성과 실험결과를, 4장에서는 논문의 결론과 향후 연구 방향을 제시한다.

## 2. 원샷 학습 기법

Fig. 1은 제안하는 기법의 전체적인 구성을 도식화한 그림이다. 제안하는 기법은 크게 학습 단계(training stage)와 예측 단계(prediction stage)로 구성된다.

학습 단계에서는 두 개의 선박 소음 신호  $x_1, x_2$ 를 입력받아 두 신호 간의 유사도  $y \in \{0, 1\}$ 를 추정하는 신경망 네트워크  $f(x_1, x_2)$ 를 학습시킨다. 이때,  $y$ 가 1에 가까울수록 두 신호의 유사도가 높음을 의미한다.

예측 단계에서는 식별 대상 선박의 소음 신호인 질의 신호(query) 신호와 후보 선박의 신호 집합인 지지 집합(support set) 내 각 신호 간의 유사도를  $f$ 로 추정하여 대상 선박을 식별한다. 이때 학습 단계 사용하는 데이터 집합(data set)은 메타-학습 집합(meta-training set), 예측 단계에서 지지 집합을 구성하기 위해 사용

하는 데이터 집합은 메타-테스트 집합(meta-test set)으로 정의하며, 두 데이터 집합 간에 중복되는 원소는 없다[16].

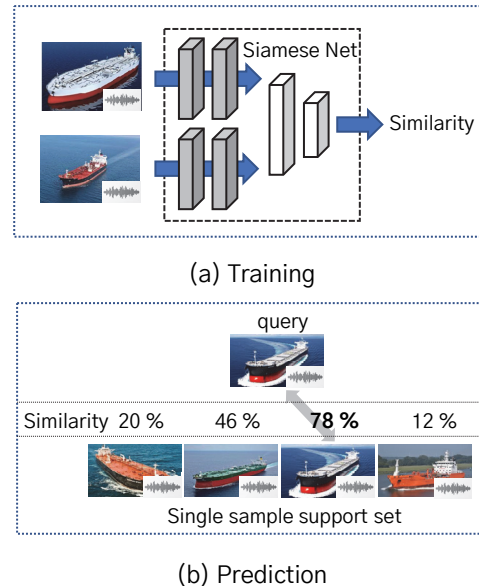


Fig. 1. Concept of the proposed method

### 2.1 신경망 네트워크(Neural network) 구조

본 논문에서는 대표적인 원샷 러닝 기법의 하나인 합성곱 삼 신경망 네트워크(convolutional Siamese neural network)[15] 구조를 기반으로 수중 소음 분석에 적합하도록 재구성하였다. 전체 네트워크는 크게 3개의 CNN layer로 구성된 특징 임베딩 네트워크(feature embedding network; FEN)와 3개의 FCNN(fully connected neural network)으로 구성된 유사도 추정 네트워크(similarity estimation network; SEN)를 가지고 있다.

FEN은 스펙트로그램(spectrogram) 이미지 형태의 선박 소음 신호를 입력으로 받아 이를 특징 벡터(feature vector)로 임베딩한다. 이때 가중치 공유(weight sharing)를 적용하여 하나의 네트워크로 두 소음 신호의 특성 벡터를 각각 추출한다. 추출된 각 신호의 특징 벡터는 SEN을 통과하고 그 결과값으로 두 신호의 유사도를 추정한다. SEN의 마지막 계층에 시그모이드(Sigmoid)를 활성화 함수(activation function)로 적용하여, 네트워크의 출력값  $y$ 는 0에서 1 사이의 실숫값을 가지도록 하였다. 전체 모델 구조와 세부 매개변수는 Fig. 2 및 Table 1과 같다.

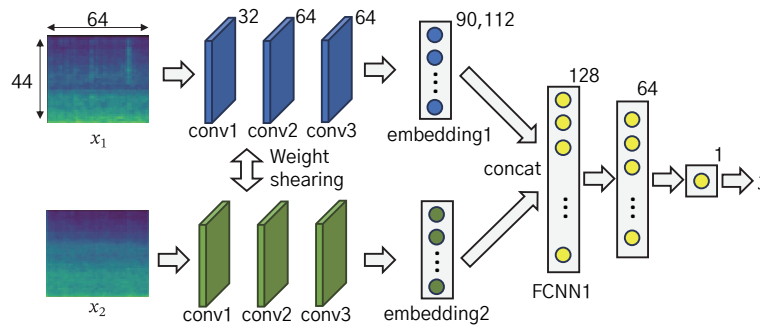


Fig. 2. Neural network structure of the proposed method

Table 1. Feature extraction parameters for librosa

Sampling rate	Segment size	Window size	Hop time	Number of bins
16 kHz	3.0s	0.25s	0.0625s	64

## 2.2 학습단계

메타-학습 집합에서 학습을 위한 미니 배치를 추출하는 과정은 다음과 같다. 학습데이터는 두 개의 입력 신호  $x_1, x_2$ 와 유사도 출력값  $y \in \{0, 1\}$ 로 구성된다. 미니 배치(mini batch)의 배치 크기(batch size)를  $n_b$ , 배치 내 각 원소의 번호를  $i_b$ 라고 할 때, 메타 학습데이터에서 총  $n_b$ 만큼의 선박을 무작위로 추출한다. 이때,  $i_b$ 가 홀수이면 동종 신호,  $i_b$ 가 짝수이면 이종 신호로 정의한다. 동종 신호는 두 입력신호 값이 같은 선박에서 추출된 것으로 메타-학습 집합 내 동일 선박 데이터 중 서로 다른 임의의 두 신호를 선택하여 학습데이터를 구성한다. 이종 신호는 두 입력신호가 다른 선박에서 추출된 것을 말하며, 메타-학습 집합 내에서 서로 다른 두 선박의 데이터에서 각각 하나의 신호를 선택하여 학습데이터를 구성한다. 이때, 동종 신호의 출력값은  $y = 1$ , 이종 신호는  $y = 0$ 를 부여하여 유사도를 표현한다.

학습을 위한 비용 함수로는 식 (1)과 같은 이산 크로스 엔트로피(binary cross entropy, BCE) 함수를 사용하였다.

$$(y_i)[- \log(y_p)] + (1 - y_i)[- \log(1 - y_p)] \quad (1)$$

여기서,  $y_i$ 는 실제 유사도 값,  $y_p$ 는 예측 유사도 값을 의미한다.

## 2.3 예측 단계

앞서 설명한 바와 같이, 예측 단계는 식별 대상 선박의 방사 소음 신호인 질의(query) 신호와 가장 유사한 신호를 후보군인 지지 집합(support set) 안에서 찾아내는 것을 목적으로 한다. 이때 지지 집합의 원소 수, 즉 후보군의 수를  $n$ -way라 하며, 지지 집합 내 각 후보군의 데이터 수를  $n$ -shot이라 한다[16]. 일반적으로  $n$ -way가 클수록 후보군이 많아지므로 난도가 상승하고,  $n$ -shot 값이 클수록 예측을 위한 기준 데이터가 많으므로 난도가 하락한다. 본 논문에서는 원샷(one-shot) 학습을 하기에 지지 집합의 각 클래스는 하나의 방사 소음 신호만을 가지고 있다.

예측 단계의 각 과정은 다음과 같다. 먼저 질의 선박과 해당 선박의 질의 신호  $x^q$ 를 기준으로 후보군 선박과 해당 선박의 신호  $x_s^i \in \{1, 2, \dots, m\}$ 로 구성된 지지 집합을 생성한다. 이때  $n$ 은  $n$ -way의 값이다. 이후, 앞서 훈련 단계에서 학습이 완료된  $f(x^q, x_s^i)$ 를 이용하여, 질의 신호와 각 후보 신호의 유사도  $y^i$ 를 추정한다. 마지막으로, 추정된  $y^i$ 값이 가장 큰 신호에 해당하는 선박을 질의 신호의 선박과 같은 선박으로 추정한다.

## 3. 실험 설계 및 결과

### 3.1 데이터 집합 및 실험 환경 구성

실험 및 검증을 위해 “DeepShip” 선박 소음데이터를 사용하였다[12]. 이 데이터는 고정형 음탐기를 선박 통행량이 많은 태평양 북서부의 수심 140 m에 배치하여 통행하는 선박의 수중 방사 소음을 네 종류의 선박 클래스로 구분하여 수집한 데이터이다. 본 논문에서는 그중 가장 많은 선박의 데이터를 가진 “Tanker” 클

래스의 데이터를 실험에 사용하였으며, 해당 클래스는 총 133척 45,900초 분량의 선박 수중 방사 소음데이터를 가지고 있다.

선박 방사 소음데이터의 전처리를 위하여 음향 특징 추출 기법 중 가장 널리 쓰이는 기법인 CQT transform[17,18]을 사용하였다. 특징 추출의 세부 과정은 선행연구[2]에서 검증한 방식을 본 논문의 목적에 적합하도록 변형하여 적용하였다. Python의 대표적인 음향 분석 라이브러리인 librosa[19]를 사용하였으며, 특징 추출을 위해 사용한 매개변수는 Table 2와 같다.

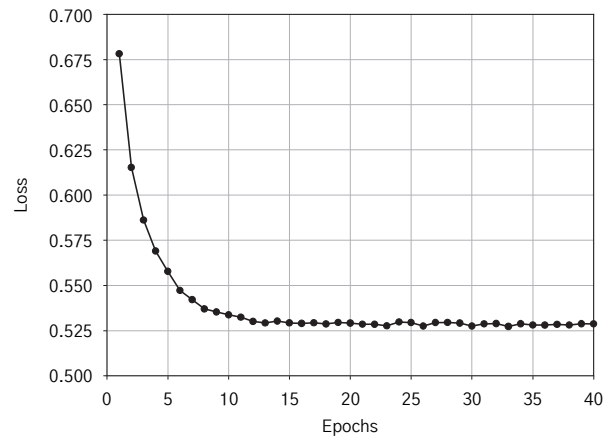
추출 과정에서 선박별 음향 표본 신호의 합이 200초 미만인 선박의 표본은 제외하였다. 이를 통해 만들어진 표본 중 각 선박의 모든 표본 길이의 합이 4,000초 이상이면 메타-학습 집합으로, 미만이면 메타-테스트 집합으로 배정하였다. 최종적으로 구축된 실험 데이터 집합은 총 123척의 선박 데이터로 그중 65척은 메타-학습 집합, 58척은 메타-테스트 집합으로 구분하였으며, 각 데이터의 크기는 637,008개, 64,840개이다.

실험 수행을 위해 NVIDIA사의 DGX-A100 서버를 이용하였고, 딥러닝 모델 구현 과정에서 Python 3.10과 Pytorch 2.0을 사용하였다. 학습 시 배치 크기는 64, 에포크는 40이었으며 최적화 알고리즘은 Ada-delta를, 알고리즘의 학습률은 0.01, gamma는 0.7을 사용하였다. 모든 초매개변수(hyper-parameter)는 반복실험을 통해 최적값을 도출하였다.

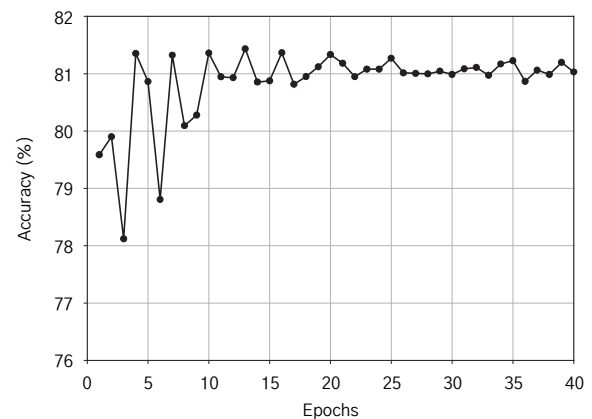
### 3.2 학습 결과

Fig. 3는 각 학습 에포크 별로 나타낸 학습 결과이다. Fig. 3(a)는 학습 과정에서의 loss 값을 보여준다. 학습이 진행될수록 loss 값이 점차 감소하여 15 - 20 에포크를 전후로 하여 학습이 안정화되는 것을 볼 수 있다. Fig. 3(b)의 검증 정확도(validation accuracy)는  $n$ -way의 수를 2로 하고, 동종 신호와 이종 신호의 비율을 1:1로 구성한 데이터를 사용하여 실험하였다. 이때 모델의 출력값이 0.5 이상일 경우 동종 신호, 출력값이 0.5 미만일 때 이종 신호로 간주하여 얼마나 정확하게 구분하였는지 여부를 측정하였다. 해당 그래프는 일반적인 원샷 학습 기법에 나타나는 검증 데이터의 학습 그래프 패턴과 유사한 형태를 보인다. 5 - 20 에포크까지는 검증값의 변동성이 크나, 25 에포크 이후는 검증값이 안정화된 모습을 보여준다. 즉 위 실험 결과를 통

해 제안한 기법이 안정적으로 학습되었음을 확인할 수 있다.



(a) Train loss



(b) Validation accuracy

Fig. 3. Learning curves

### 3.3 테스트 결과

테스트를 위한 질의 및 지지 집합의 배치(batch) 생성 과정은 다음과 같다. 먼저 메타-테스트 집합에서 차례로 질의 신호를 추출한다. 다음으로, 해당 신호를 추출한 선박에서 임의의 다른 신호를 추가로 추출하여 지지 집합에 추가한다. 이후, 해당 선박을 제외한 나머지 선박 중에서  $n$ -way - 1 만큼의 선박을 임의로 추출하고, 추출된 각 선박에서 하나의 신호를 임의로 추출하여 지지 집합을 구성하였다.

Fig. 4는 학습이 완료된 유사도 측정 함수  $f$ 를 이용하여  $n$ -ways 수를 변경하면서 측정한 정확도를 표현한 것이다. 여기서 정확도는 전체 메타-테스트 집합 내 모든 질의 신호를 대상으로 각 지지 집합 내에서 동일 선박의 신호를 정확하게 추정된 비율을 의미한다.

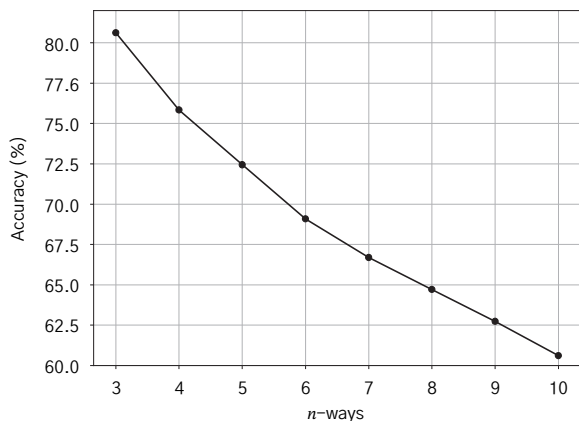


Fig. 4. Recognition accuracy per  $n$ -ways

Fig. 4에서 볼 수 있듯이 지지 집합의 수인  $n$ -ways, 즉 분류 후보군이 증가할수록 정확도는 감소한다. 후보군의 수가 세 척일 때 80.1% 이상의 정확도를, 후보군이 다섯 척일 때 약 72.5%의 정확도를 보여주었다. 특히 후보군이 열 척인 악조건에서도 60%의 정확도로 동일 선박을 식별하였다.

#### 4. 결론

본 논문에서는 원샷 학습을 이용한 선박 소음 자동 식별 기법을 제안하였다. 기존의 지도학습 기반 소음 식별 알고리즘의 경우 학습에 다수의 데이터가 필요하고, 학습되지 않은 신호에 대해서는 분류할 수 없는 단점이 있었으나, 본 연구에서는 메타-학습 기법의 일종인 원샷 학습 기법을 이용하여 단 하나의 신호 표본만으로 학습 없이 선박 소음 식별이 가능하게 하였다. 특히 선박의 종류 식별에 집중된 기존 연구와 다르게 선명 수준의 분류 방법을 제안하였으며, 5척의 후보군을 기준으로 약 72.5%의 분류 성능을 달성하였다.

본 논문의 연구는 사전 지정된 후보군에 대해서만 식별하고 있으나, 향후 추가 연구를 통해 실제 잠수함의 운용 시나리오를 기반으로 한 온라인 식별 및 추적 알고리즘으로 연구를 확장하여 실용성을 높일 예정이다.

#### 참고문헌

- [1] Xu Xiao, Wenbo Wang, Qunyan Ren, Peter Gerstoff, and Li Ma, "Underwater Acoustic Target Recognition Using Attention-based Deep Neural Network," *J. Acoust. Soc. Am. Express Letters* 1, 106001 (2021).
- [2] Kyung-Min Park, and Dooyoung Kim. "Preprocessing

Performance of Convolutional Neural Networks According to Characteristic of Underwater Targets," (in Korean), *J. Acoust. Soc. Kr*, Vol. 41, pp. 629–636 (2022).

- [3] B. Friedlander, "Accuracy of Source Localization Using Multipath Delays," *IEEE Trans. Aerospace and Electronic Systems*, Vol. 24, pp. 346–359 (1988).
- [4] M. Johnson, L. Freitag, and M. Stojanovic, "Improved Doppler Tracking and Correction for Underwater Acoustic communications," *Proc. IEEE ICASSP*, 1, pp. 575–578 (1997).
- [5] N. Owsley and G. Swope, "Time Delay Estimation in a Sensor Array," *IEEE Trans. ASSP*, Vol. 29, pp. 519–523 (1981).
- [6] W. Zhang, Y. Wu, D. Wang, Y. Wang, Y. Wang, and L. Zhang, "Underwater Target Feature Extraction and Classification Based on Gammatone Filter and Machine Learning," *Proc. ICWAPR*, pp. 42–47 (2018).
- [7] X. Yin, X. Sun, P. Liu, L. Wang, and R. Tang "Underwater Acoustic Target Classification Based on LOFAR Spectrum and Convolutional Neural Network," *Proc. 2nd Int. Conf. AIAM*, pp. 59–63 (2020).
- [8] H. Yang, J. Li, S. Shen, and G. Xu, "A Deep Convolutional Neural Network Inspired by Auditory Perception for Underwater Acoustic Target Recognition," *Sensors*, Vol. 19, 1104 (2019).
- [9] Z. Wei, Y. Ju, and M. Song, "A Method of Underwater Acoustic Signal Classification Based on Deep Neural Network," *Proc. 5th ICISCE*, IEEE Computer Society, pp. 46–50 (2018).
- [10] Yang, S. et al. "A Lightweight Network Model Based on an Attention Mechanism for Ship-radiated Noise Classification," *Journal of Marine Science and Engineering*, Vol. 11, No. 2, 432. (2023).
- [11] D. Santos-Dominquez, S. Torres-Guijarro, A. CardenalLopez, and A. Pena-Gimenez, "ShipsEar: An Underwater Vessel Noise Database," *Appl. Acoust.* Vol. 113, pp. 64–69 (2016).
- [12] 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. Appl.* Vol. 183, 115270 (2021).
- [13] Bengio, Yoshua, Ian Goodfellow, and Aaron Courville. *Deep learning*. Vol. 1. Cambridge, MA, USA: MIT Press, 2017.
- [14] Fei-Fei, L., Fergus, R., Perona, P., "One-shot Learning of Object Categories," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 4, pp. 594–611 (2006).
- [15] Koch, G. et al, "Siamese Neural Networks for One-shot Image Recognition," In: *ICML Deep Learning Workshop*, Vol. 2 (2015).
- [16] Hospedales, T., Antoniou, A., Micaelli, P., and Storkey, A., "Meta-learning in Neural Networks: A Survey," *IEEE Transactions on Pattern Analysis and Machine*

- Intelligence, Vol. 44, No. 9, pp. 5149–5169, (2021).
- [17] Schörkhuber, C. and Klapuri, A, "Constant-q Transform Toolbox for Music Processing," Proc. 7th Sound and Music Computing Conf. (2010).
- [18] Brown, J.C, "Calculation of a Constant q Spectral Transform," The Journal of the Acoustical Society of America, Vol. 89, No. 1, pp. 425–434 (1991).
- [19]. B. McFee, C. Raffel, D. Liang, D. Ellis, M. Mcvcar, E. Battenberg, and O. Nieto, "Librosa: Audio and Music Signal Analysis in Python," Proc. 14th Python Sci. Conf. pp. 18–24 (2015).