

# 수중 표적 분류를 위한 합성곱 신경망의 전처리 성능 비교

## Preprocessing performance of convolutional neural networks according to characteristic of underwater targets

박경민,<sup>1,2</sup> 김두영<sup>3†</sup>

(Kyung-Min Park<sup>1,2</sup> and Dooyoung Kim<sup>3†</sup>)

<sup>1</sup>한국해양대학교 해양공학과, <sup>2</sup>대한민국 해군 전력지원체계사업단, <sup>3</sup>대한민국 해군사관학교 사이버과학과  
(Received July 21, 2022; revised September 21, 2022; accepted October 5, 2022)

**초 록:** 본 논문은 합성곱 신경망 기반 수중 표적 분류기의 성능 향상을 위한 최적의 전처리 기법을 제시한다. 실제 선박 수중신호를 수집한 데이터 세트의 주파수 분석을 통해 강한 저주파 신호로 인한 특성 표현의 문제점을 확인하였다. 이를 해결하기 위해 다양한 스펙트로그램 기법과 특성 스케일링 기법을 조합한 전처리 기법들을 구현하였다. 최적의 전처리 기법을 확인하기 위해 실제 데이터를 기반으로 합성곱 신경망을 훈련하는 실험을 수행하였다. 실험 결과, 로그 멜 스펙트로그램과 표준화 및 로버스트 정규화 스케일링 기법의 조합이 높은 인식 성능과 빠른 학습 속도를 보임을 확인하였다.

**핵심용어:** 수중 표적, 합성곱 신경망, 전처리, 스케일러, 이상치

**ABSTRACT:** We present a preprocessing method for an underwater target detection model based on a convolutional neural network. The acoustic characteristics of the ship show ambiguous expression due to the strong signal power of the low frequency. To solve this problem, we combine feature preprocessing methods with various feature scaling methods and spectrogram methods. Define a simple convolutional neural network model and train it to measure preprocessing performance. Through experiment, we found that the combination of log Mel-spectrogram and standardization and robust scaling methods gave the best classification performance.

**Keywords:** Underwater target, Convolution neural network, Preprocessing, Scaler, Outlier

**PACS numbers:** 43.60.Bf, 43.60.Lq

### 1. 서 론

선박의 수중 방사 소음을 이용하여 표적을 탐지하고 식별하는 문제는 소나 시스템의 주요한 부분이다.<sup>[1]</sup> 하지만, 해양의 높은 배경 소음과 음파의 다중 경로 전달 등으로 인한 낮은 신호 대 잡음 비는 성능 향상을 어렵게 하고 있다.<sup>[2]</sup> 이를 해결하기 위해 다중경로상 음파 전달 영향을 제거하거나, 빔 형성 기법 적용, 표적 위치 추정 및 추적 등 여러 가지 신호 처리 방법이 적용되었다.<sup>[3-6]</sup>

한편, 기계 학습을 비롯한 심층 신경망(Deep Neural Network, DNN), 합성곱 신경망(Convolution Neural Network, CNN)<sup>[7]</sup> 등 인공지능 알고리즘의 급속한 발전에 따라 이를 수중 표적 탐지 및 분류에 활용하는 연구가 진행되고 있다. 예를 들어, 의사 결정 트리(Decision Tree), 서포트 벡터 머신(Support vector machine, SVM) 및 K-근접 이웃(K-nearest neighbor, KNN) 등 전통적인 기계 학습 알고리즘을 수중신호에 적용하거나,<sup>[8]</sup> Feed-forward neural network(FNN), Random Forest(RF) 등을 이용하여 수중 표적의 위치 추정을

†Corresponding author: Dooyoung Kim (dykim07@outlook.com)

Department of Cyber Science, Republic of Korea Naval Academy, 1 Jungwon-ro, Jinhae-gu, Changwon-si Gyeongsangnam-do 51704, Republic of Korea

(Tel: 82-55-907-5246)



Copyright©2022 The Acoustical Society of Korea. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

연구한 사례도 있다.<sup>[9]</sup>

또한, 인공지능 알고리즘을 적용하기 위해 실해역에서 데이터를 확보하여 공개한 연구도 진행되었다. “ShipsEar”라고 명명된 수중 표적 데이터 세트는 스페인 대서양 연안에서 11개 선박 종류의 수중 방사 음원이 수집된 자료로, 총 6,189s 길이의 90개 파일로 구성되어 있다.<sup>[10]</sup> 가장 최근에 공개된 수중 표적 데이터 세트인 “DeepShip”은 캐나다 태평양 북서 해안 중 통항량이 많은 지점에 음원 수집 센서를 설치하여 데이터를 획득하였으며, 총 265개 선박에 대한 약 47시간 길이의 음원 데이터를 보유하고 있다.<sup>[11]</sup> 수중음향 데이터의 공개에 따라 이를 이용한 인공지능 알고리즘 적용에 관한 연구도 계속되고 있다.

수중 표적 신호와 같이 시간-음압의 1차원 데이터를 CNN과 같은 인공지능 알고리즘에 적용하기 위해서는 적합한 전처리 과정이 필요하다. 수중음향 분야에서 널리 사용되는 Low-frequency analysis(LOFAR)로 전처리를 하여 CNN에 적용한 연구<sup>[12]</sup>도 있지만, 음향을 입력으로 사용하는 많은 경우에 심리음향 기반의 멜 스펙트로그램(Mel spectrogram)이나 Mel-Frequency Cepstral Coefficient(MFCC) 등으로 전처리한다. 수중 표적 분류 연구에서도 멜스펙트로그램과 MFCC으로 전처리하여 CNN 알고리즘에 적용하였으며,<sup>[13,14]</sup> 같은 인공지능 알고리즘에 멜 스펙트로그램, MFCC, Cepstrum, Gammatone, Wavelet 및 CQT(Constant-Q transform)의 전처리에 따른 분류 성능을 비교한 연구도 진행되었다.<sup>[11]</sup>

하지만, 위의 연구 등에서는 수중 표적 신호에 대해 다양한 변환 방식으로 인공지능 알고리즘의 분류 성능을 확인하였지만, 데이터 스케일링 프로세스가 전체 모델 성능에 미치는 영향을 고려한 연구는 미흡하였다. 한편, 최근 의학 분야에서는 데이터 스케일링 변화에 따른 인공지능 알고리즘 분류 성능을 비교한 연구가 수행되었으며,<sup>[15]</sup> 음향 분야에서도 환경음 데이터 세트에 대해 멜 스펙트로그램과 MFCC에 각각 다른 스케일러를 적용하여 CNN의 분류 성능을 확인한 연구 결과가 제시되었다.<sup>[16]</sup>

따라서, 본 논문에서는 CNN을 이용한 수중 표적 분류에서 다양한 전처리 기법 적용에 따른 성능 변화를 확인하고자 한다. 수중 표적 분류를 위해 데이

터 세트의 음향 신호를 전처리 과정에서 널리 사용되는 멜 스펙트로그램과 로그 멜 스펙트로그램으로 변환하고, 각각에 대해 4가지의 스케일러를 적용하여 전처리하였을 때 CNN 알고리즘의 분류 성능과 학습 성능을 비교하였다.

논문의 구성은 다음과 같다. 먼저 2장에서는 논문에 사용된 “DeepShip” 수중 표적 데이터에 대해 탐색적 데이터 분석(Exploratory Data Analysis, EDA)을 수행하였다. 3장에서는 수중 표적 분류에 적합한 전처리 기법의 조합을 제안하고 4장에서 표준적인 CNN 모델에 대한 학습 실험을 통해 각 전처리 기법의 성능을 확인하였다.

## II. 수중음향 신호의 변환 및 분석

### 2.1 수중 표적 데이터 형태

논문에서는 2021년에 공개된 “DeepShip” 수중 표적 데이터를 사용하였다.<sup>[11]</sup> 이 데이터는 태평양 북서부 해역 통항량이 많은 경로의 수심 약 140 m에 음향 수집기를 설치하고, 이를 기준으로 반경 2 km 이내로 통과하는 선박의 수중 소음을 녹음한 데이터로 약 47 h 분량, 총 265대의 선박 데이터로 구성되어 있다. 정확한 선박의 구분을 위해 선박 자동식별시스템(Automatic Identification System, AIS)을 사용하였으며, 이를 통해 수집된 수중 표적 데이터는 Tug, Passenger ship, Tank, Cargo의 네 가지 클래스로 라벨링 되었다. 클래스별 수집된 수중 표적 데이터의 신호 길이와 데이터 수는 Table 1과 같다.

수집된 시간 영역의 음향 신호는 푸리에 변환을 통해서 시간-주파수 영역의 신호로 나타내고, 이후 멜 스펙트로그램 및 로그 멜 스펙트로그램으로 각각 변환하여 비교하였다.

먼저 멜 스펙트로그램은 심리음향 및 음성 신호

Table 1. Size of the DeepShip<sup>[11]</sup> dataset.

Class	Duration (s)	Number of ships
Cargo	38,400	69
Passenger ship	44,520	46
Tug	40,620	17
Tanker	45,900	133

처리에서 널리 사용되는 기법이다. 멜 스펙트로그램은 Mel scale에서 나타나는 특성을 고려한 것으로, Eq. (1)과 같이 선형의 주파수 스케일을 Mel scale로 나타낼 수 있다. 먼저 계산된 스펙트로그램 값은 Mel scale에서 밴드의 수만큼 생성된 Triangular filter를 통과하여 멜 스펙트로그램 결과로 변환된다.

$$Mel(f) = 2595 \log\left(1 + \frac{f}{700}\right). \quad (1)$$

로그 멜 스펙트로그램 신호 변환은 Eq. (2)와 같이 멜 스펙트로그램의 제곱  $S$ 에 로그를 취해 데시벨 값으로 변환한다. 이때, 기준값인  $ref$ 는 멜 스펙트로그램의 최댓값이며, 차원의 크기 변화는 없다.<sup>[16]</sup>

$$S_{log} = 10 \log_{10}\left(\frac{S}{ref}\right). \quad (2)$$

세부적인 신호 변환과정은 다음과 같다. 먼저, 주어진 음향 데이터를 16 kHz로 재 표본화한 후 이를 3 초 단위로 분할 하였다. 분할된 각각의 신호에 대해

Table 2. Number of data per classes.

Class	Train	Validation	Test
Cargo	8189	2081	2531
Passenger ship	9869	2397	3144
Tug	8618	2181	2696
Tanker	9527	2359	2876

FFT를 수행하였다. 이때, 윈도우 단위는 265 ms를 사용하였으며, 프레임은 64 ms씩 이동하여 프레임별로 75%가 중첩되도록 하였다. 이를 통해 총 47개의 프레임을 구성하였다. 최종적으로 멜 밴드는 40개로 하여 40×47 크기의 멜 스펙트로그램을 생성하였다. 생성된 데이터를 무작위로 혼합하고, 혼합된 데이터를 6:2:2 비율로 나누어 각각 학습, 검증, 테스트 데이터로 사용하였다. Table 2는 클래스별 데이터의 수를 나타낸다.

## 2.2 데이터 특성 분석

Fig. 1(a)와 (b)는 같은 수중 표적 음원에 대해 각각 멜 스펙트로그램 기법과 로그 멜 스펙트로그램 기법을 통해 추출한 특징을 시각화한 결과이다. 그림에서 멜 스펙트로그램의 이미지는 로그 멜 스펙트로그램의 이미지와 비교해 클래스별 구별이 명확하지 않은 모습을 보인다. 특히 강한 저주파 영역의 신호로 인해 상대적으로 고주파 영역의 특징들이 약해지는 특성이 나타난다. 이러한 현상은 선박 고유의 기계적 소음과 프로펠러 소음 등이 에너지의 대부분이 1 kHz 이내의 저주파 대역에 집중되어 있기 때문이다.<sup>[17]</sup>

하지만, 이러한 경향은 선박의 종류 및 속도에 따라 각각 다른 형태로 나타난다. 따라서 해당 저주파 수 대역을 제거하는 것은 모델이 활용할 수 있는 공간적, 지역적 특징이 소실되는 것을 의미한다. 따라서, 분류 성능 향상을 위해서는 단순히 저주파 대역

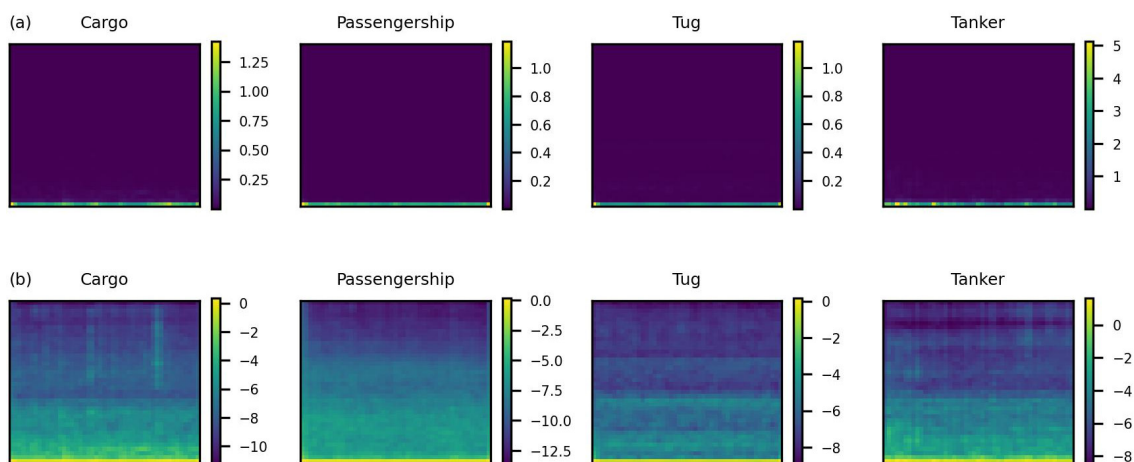


Fig. 1. (Color available online) Preprocessing results (a) Mel-spectrogram (b) log Mel-spectrogram.

을 제거하는 것이 아니라, 이에 적합한 전처리 기법을 적용하는 것이 필요하다.

### III. 특징 스케일링 적용

#### 3.1 특징 스케일링의 필요성

이미지 분류를 위한 딥러닝 모델에서 CNN은 입력 벡터의 내재한 특성을 자동으로 추출하기 위해 사용된다. 이때 입력 벡터의 원소별 값의 차이가 크면 CNN 내부의 각 필터의 학습을 저해하여 분류 성능이 저하된다.<sup>[7]</sup>

또한, 입력 벡터의 각 원소값의 절댓값이 큰 경우, 이에 연결되는 가중치 값의 이동 범위도 함께 커지게 된다. 따라서 최적화 단계에서 최적해에 수렴하는데 걸리는 시간이 증가한다.<sup>[7]</sup> 따라서 입력 벡터의 각 원소가 가지는 절댓값을 적절하게 조절하는 것은 딥러닝 모델의 성능 향상을 위한 중요한 작업이다.<sup>[7]</sup>

특징 스케일링(Feature scaling) 또는 정규화(Normalization)는 딥러닝에 사용되는 데이터의 크기를 조절하여 학습 성능을 향상하는 기법이다.<sup>[7]</sup> 앞서 데이터 분석에서 확인한 바와 같이, 수중에서의 선박 소음 데이터는 특정 고주파 영역의 절댓값이 상대적으로 크기 때문에, 식별 가능한 패턴을 보이는 고주파 영역의 중요성이 감소하는 경향이 있다. 따라서, 딥러닝 계열의 수중 표적 분류 기법을 사용할 때 적절한 특징 스케일링 기법의 적용은 성능 향상을 위해 필요하다.

#### 3.2 스케일링 기법

본 논문에서 사용한 스케일링 기법은 다음과 같다.<sup>[7]</sup> 학습데이터  $\mathbf{X}$ 는 2장에서 언급한 바와 같이 데이터의 신호변환 결과로 총 40-by-47 행렬이다.  $\mathbf{X}$ 의 열벡터인  $\mathbf{x}_j = [x_{1,j}, x_{2,j}, \dots, x_{40,j}]^T$ 는  $j$ 번째 시간프레임에서 나타난 주파수별 멜스펙트로그램 성분의 집합이다, 이에 대해 각각의 스케일링 기법을 적용하였으며, 적용된 결과는  $\hat{\mathbf{x}}_j$ 로 나타내었다.

##### 3.2.1 Standardization

일반적으로 널리 사용되는 스케일링 기법으로 학

습데이터  $\mathbf{X}$ 의 각  $\mathbf{x}_j$  열이 평균은 0, 표준편차는 1을 갖는 표준정규분포로 재조정된다.

각각의 성분에 대해 평균값을 제거하고 표준편차로 나누어줌으로써 데이터의 성분이 일정 범위 안에 존재한다. 변환 과정에서 최댓값과 최솟값을 제한하지 않아 데이터 분포에 영향을 주는 등 이상치에 민감하나, 이러한 특징으로 특정 범위를 벗어난 결과를 쉽게 확인하고 제거할 수도 있는 기법이다.

$$\hat{\mathbf{x}}_j = \frac{\mathbf{x}_j - \bar{\mathbf{x}}_j}{\sigma_j}, \forall j, \quad (3)$$

여기서  $\bar{\mathbf{x}}_j$ 는  $\mathbf{x}_j$  성분의 평균값,  $\sigma_j$ 는  $\mathbf{x}_j$  성분의 표준편차다.

##### 3.2.2 Min-max

학습데이터  $\mathbf{X}$ 의 각  $\mathbf{x}_j$ 에 대해 최댓값과 최솟값을 기준으로 하여 0과 1 사이의 내에 연속 값으로 사상하는 전처리 기법이다. 일반적으로 이상치에 민감하다는 단점이 있지만, 각 특징의 범위가 모두 0에서 1로 동등하게 분포한다는 특성이 있다.

특징 벡터의 각 원소에 대하여 최대 최소 범위를 계산하는 방법은 다음과 같다.

$$\hat{\mathbf{x}}_j = \frac{\mathbf{x}_j - \min(\mathbf{x}_j)}{\max(\mathbf{x}_j) - \min(\mathbf{x}_j)}, \forall j. \quad (4)$$

##### 3.2.3 Max-abs

학습데이터  $\mathbf{X}$ 의 각  $\mathbf{x}_j$  열에 대해 절댓값 0과 1 사이의 내에 위치하도록 한다. 즉, -1부터 1 사이로 재조정하며, 데이터를 이동 및 중앙화하지 않으므로 희소성이 손상되지 않는다. 기존 데이터가 양의 값으로만 구성된 경우 Min-max 스케일링과 유사하게 동작하는 특징이 있으며, 역시 이상치(Outlier)에 민감하다.

$$\hat{\mathbf{x}}_j = \frac{\mathbf{x}_j}{\max(\mathbf{x}_j)}, \forall j. \quad (5)$$

### 3.2.4 Robust scaler

데이터에 이상치가 많은 경우, 데이터의 평균 및 표준 편차를 사용한 스케일링 기법은 알고리즘 학습에 있어서 제대로 동작하지 않을 수 있다. 이러한 단점을 개선하기 위해 Robust scaler는 Standard scaler의 평균값과 표준 편차 대신 중앙값(Median)과 4분위 간 범위(InterQuartile Range, IQR)를 사용하여 상대적으로 더 넓게 분포시키며, 이상치의 영향을 최소화한 방법이다.

$$\hat{x}_j = \frac{x_j - \text{median}(x_j)}{Q3(x_j) - Q1(x_j)}, \forall j, \quad (6)$$

여기서  $Q3(x_j)$ 는 해당 벡터의 75위,  $Q1(x_j)$ 는 25위이며, median에 해당하는 값을 0으로 위치시키고, IQR ( $Q3-Q1$ ) 차이만큼을 기준으로 정규화를 진행한다.

## IV. 식별 모델 및 실험 결과

### 4.1 학습 대상 모델과 실험 구성

본 논문의 목적은 수중 표적 분류를 위한 최적의 전처리 기법의 조합을 실험으로 확인하는 것이다. 이를 위해 성능향상기법을 적용하지 않은 단순한 구조의 CNN 기반의 모델을 실험 대상으로 구현하였다.

모델의 전체적인 구조는 Table 3과 같다. 먼저 음향 특징 추출을 위해 총 3개의 2차원 CNN 계층을 구성하였고, 각 필터는 kernel size 3 by 3, zero-padding 을 사용하였으며, Rectified Linear Unit(ReLU) 활성화함수 (Activation function)<sup>[18]</sup>를 적용하였다. 이후 출력값을 평활화한 후 3개의 전 연결 인공 신경망(Fully Connected Neural Network, FCNN)을 적용하였다. 각 계층의 활성화함수는 ReLU를 사용하였다. 모든 계층의 초기값은 Groat<sup>[19]</sup> 알고리즘으로 초기화하였다.

구현 언어는 Python 3.8 을 사용하였으며, 음향 신호 처리 및 특징 추출에는 Librosa,<sup>[20]</sup> 전처리에는 Scikit-learn<sup>[21]</sup>과 Numpy<sup>[22]</sup>를 사용하였으며, 딥러닝 모델 구현은 Tensorflow 2.6<sup>[23]</sup>를 사용하였다. 학습을 위해 NVIDIA DGX-A100 서버를 사용하였다. 모델학습 시 배치 사이즈 256으로 총 40 Epoch를 학습하였으며, 최적화 알고리즘은 Adam optimizer<sup>[24]</sup>를 학습률 0.001

Table 3. The architecture of CNN model.

Layer	Parameter
CNN 1	input channel : 1, output channel : 32, kernel size = (3,3), stride = (1,1), no padding, activation = relu
CNN 2	input channel : 32, output channel : 16, kernel size = (3,3), stride = (1,1), no padding, activation = relu
FCNN1	hidden unit = 128, activation = relu
FCNN2	hidden unit = 64, activation = relu
FCNN3	hidden unit = 4, activation = softmax

로 설정하여 사용하였다. 네트워크의 매개 변수 수와 기타 초매개 변수(Hyper parameter)는 학습 및 검증 데이터를 이용한 실험을 통해 최적값을 결정하였다.

실험은 수중 표적 데이터에 대해 2.2절과 같이 멜 스펙트로그램과 로그 멜 스펙트로그램으로 특징을 추출하고, 각각에 대해 스케일링을 적용하지 않은 경우(None)와 Standard, Min-max, Max-abs, robust의 5가지의 방법을 적용하여 모두 10개의 경우에 대해 실험을 진행하였다. 특징 추출, 전처리, 모델 훈련에는 학습데이터와 검증데이터만을 사용하였으며, 성능 검증 시에는 테스트 데이터를 사용하였다.

### 4.2 전처리 기법별 분류 성능

모델의 성능평가 지표로 Precision, Recall, F1-Score 를 사용하였으며, 각각은 Eqs. (7)~(9)와 같다.<sup>[7]</sup>

$$\text{Precision} = \frac{TP}{TP+FP}. \quad (7)$$

$$\text{Recall} = \frac{TP}{TP+FN}. \quad (8)$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (9)$$

여기서 TP는 True Positive, TN은 True Negative, FP는 False Positive, 그리고 FN은 False Negative를 나타낸다.

검증 시 테스트 데이터만을 사용하였으며, 네 가지 클래스 전체에 대한 성능을 검증하였다. 실험 결과는 Table 4와 같으며 스케일러의 적용과 관계없이 로그 멜 스펙트로그램으로 특징을 추출한 경우가 멜

Table 4. Accuracy comparison for scalers and spectrograms.

Methods	Mel-spectrogram					Log mel-spectrogram				
	None	Standard	Min-max	Max-abs	Robust	None	Standard	Min-max	Max-abs	Robust
Precision	0.8743	0.8312	0.8482	0.8499	0.8966	0.8833	0.9220	0.8921	0.8893	<b>0.9242</b>
Recall	0.8735	0.8320	0.8511	0.8411	0.8963	0.8821	0.9219	0.8913	0.8890	<b>0.9240</b>
F1-score	0.8734	0.8309	0.8489	0.8417	0.8963	0.8826	0.9219	0.8923	0.8891	<b>0.9241</b>

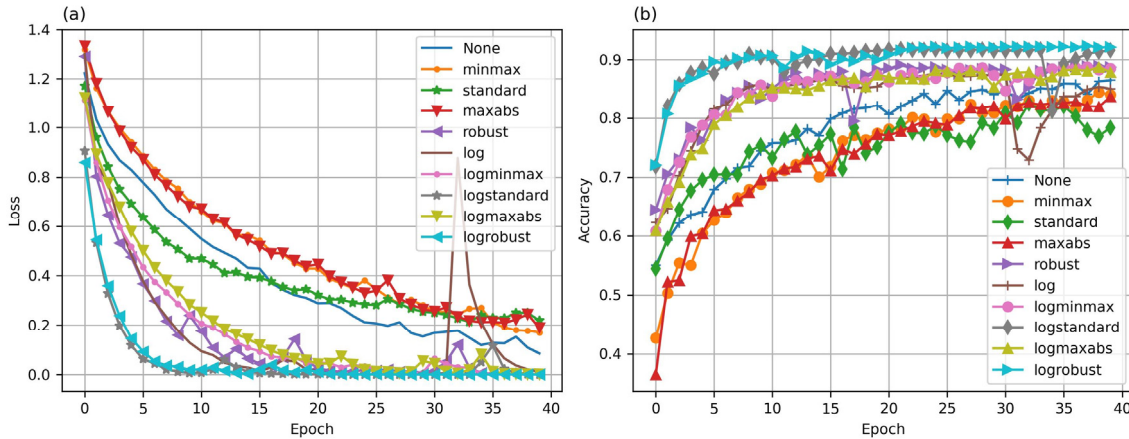


Fig. 2. (Color available online) (a) Training loss curve for each scaler (b) Validation accuracy curve for each scaler.

스펙트로그램을 적용했을 때보다 높은 성능을 나타냈다.

한편, 스케일러가 단순히 좋은 결과를 보장하지 않았다. 멜 스펙트로그램으로 특징을 추출한 경우 Robust 스케일러를 제외한 다른 3가지의 스케일러는 그렇지 않은 경우보다 낮은 성능을 보였다. 반면, 로그 멜 스펙트로그램으로 적용했을 때, 스케일러를 적용한 모든 경우에서 성능이 4% 이상 향상되었고, 특히 Standard 스케일러는 9%까지 성능이 향상되는 결과를 얻을 수 있었다.

이러한 현상은 수중 표적 데이터의 특성과 연관이 있다. Fig. 2에서 나타난 바와 같이 수중 표적 데이터 중 소수를 차지하는 저주파 신호는 상대적으로 매우 큰 값으로 존재한다. 이는 CNN 알고리즘의 학습 시 이상치(Outlier)의 역할을 하는 것으로 볼 수 있다. 이상치가 적용된 평균 및 분산을 사용하는 Standard 스케일러, 이상치로 인해 다른 값들이 상대적으로 매우 좁은 범위로 압축되는 Min-max 및 Max-abs 스케일러의 성능 결과를 통해서 알 수 있다. 4분위 값을 사용하여 이상치의 영향을 적게 받은 Robust 스케일러의 성능 결과도 실험 결과의 분석을 뒷받침한다.

반면, 로그 멜 스펙트로그램의 CNN 학습 결과 성능이 향상되었다. 로그 변환을 통해 편향된 값의 멜 스펙트로그램의 편차를 조절할 수 있다. 스케일러를 적용하지 않은 경우에도 성능은 향상되었지만, 이상치의 영향을 받은 스케일러로 로그 멜 스펙트로그램을 입력 데이터로 적용했을 때 성능이 급격히 향상되었으며, 특히 Standard 스케일러는 이상치의 영향을 적게 받는 Robust 스케일러의 성능과 유사하게 높은 성능 수치를 확인하였다.

### 4.3 전처리 기법별 학습 성능

실험에 사용된 CNN 모델에 대해 Fig. 3과 같이 스케일러 적용 형태별로 학습 곡선을 나타내었다. 일반적으로 학습 곡선은 사용된 CNN 모델이 학습을 반복하면서 데이터에 과대 적합 또는 과소 적합 등 모델의 일반화 성능을 추정하기 위해 사용된다.

Fig. 3(a)는 수중 표적 신호의 훈련 세트를 대상으로 한 스케일러 종류별 손실 함숫값이다. 전반적으로 모든 경우에 대해 손실이 감소하면서 모델의 학습이 진행되었다. 특히, 로그 멜 스펙트로그램의 Standard 스케일러 및 Robust 스케일러는 빠르게 학습하고 수렴

하였다. 상대적으로 다른 방법들은 학습 속도가 늦었으며, 멜 스펙트로그램의 Robust 스케일러와 스케일러 없는 로그 멜 스펙트로그램에서는 일시적으로 손실 값이 증가하는 현상이 나타났다.

Fig. 2(b)는 검증 세트를 대상으로 스케일러별 정확도 함수 곡선을 나타낸다. Fig. 2(a)와 유사하게 로그 멜 스펙트로그램의 Standard 스케일러 및 Robust 스케일러가 정확도 면에서도 빠르게 학습된 것을 알 수 있다.

## V. 결 론

본 논문에서는 딥러닝 기반의 선박 수중 표적 분류기의 성능 향상을 위한 최적의 전처리 기법을 제안하였다. 먼저 실제 선박 수중 소음의 시각화를 통해 강한 저주파 대역 신호와 중요한 정보를 포함하고 있으나 상대적으로 신호가 작은 고주파 영역의 특성을 관찰함으로써 전처리 기법 적용의 필요성을 확인하였다.

입력 데이터의 특성 전처리를 위해 2종의 특징 추출 알고리즘과 5종의 특징 스케일링 기법을 조합하여 10종의 전처리 기법 조합을 제작하였다. 실제 수중 선박 신호 데이터를 제작한 기법으로 전처리한 후 CNN 모델을 학습함으로써, 각 전처리 기법 조합의 성능을 확인하였다. 실험 결과 제작한 전처리 기법을 실제 데이터를 이용하여 CNN 모델을 학습시킴으로써 로그 멜 스펙트로그램과 Robust 스케일러 조합이 가장 높은 성능을 보임을 확인하였다.

이번 연구에서는 특정해역에서 확보된 수중 표적 신호에 대해 휴리스틱 기반의 특징 추출 및 스케일링 기법을 적용한 전처리 기법의 성능을 확인하였다. 스케일러 구현 시 해양 환경이 아닌 표적 선박의 구동 특성에 초점을 두었다. 따라서 선박 분류 문제에 한정한다면, 다른 해역에서 수집된 데이터라도 같은 성능 향상이 있으리라 판단한다. 향후 비지도 학습(Unsupervised learning)과 자기 지도 학습(Self-supervised learning) 기반의 자동 특징 추출 기법을 적용한 연구와 서로 다른 해역에서 확보된 수중 표적 신호에 대한 CNN 분류 성능연구로 확장하고자 한다.

## 감사의 글

이 논문은 2022년 해군사관학교 해양연구소 학술연구과제 연구비 지원으로 수행된 연구임.

## References

1. X. Xiao, W. Wang, Q. Ren, P. Gerstoft, and L. Ma, "Underwater acoustic target recognition using attention-based deep neural network," *JASA Express Lett.* **1**, 106001 (2021).
2. D. Wang, L. Zhang, C. Bao, S. Ma, and Y. Wang, "Passive Ship Localization in a Shallow Water Using Pre-Trained Deep Learning Networks," *Proc. the 23<sup>rd</sup> ICA*, 1956-1962 (2019).
3. J. S. Abel and K. Lashkari, "Track parameter estimation from multipath delay information," *J. Ocean Eng.* **12**, 207-221 (1987).
4. B. Friedlander, "Accuracy of source localization using multipath delays," *IEEE Trans. Aerospace and Electronic Systems*, **24**, 346-359 (1988).
5. M. Johnson, L. Freitag, and M. Stojanovic, "Improved doppler tracking and correction for underwater acoustic communications," *Proc. IEEE ICASSP.* **1**, 575-578 (1997).
6. N. Owsley and G. Swope, "Time delay estimation in a sensor array," *IEEE Trans. ASSP.* **29**, 519-523 (1981).
7. Y. Bengio, I. Goodfellow, and A. Courville, *Deep Learning* (MIT press, Cambridge, 2017), pp. 1-272.
8. 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*, 42-47 (2018).
9. H. Niu, E. Reeves, and P. Gerstoft, "Source localization in an ocean waveguide using supervised machine learning," *J. Acoust. Soc. Am.* **142**, 1176-1188 (2017).
10. D. Santos-Dominquez, S. Torres-Guijarro, A. Cardenal-Lopez, and A. Pena-Gimenez, "ShipsEar: An underwater vessel noise database," *Appl. Acoust.* **113**, 64-69 (2016).
11. M. Irfan, Z. Jiangbin, S. Ali, M. Iqbal, Z. Masood, and U. Hamid, "DeepShip: An underwater acoustic benchmark dataset and a separable convolution based auto-encoder for classification," *Expert Syst. Appl.* **183**, 115270 (2021).
12. X. Yin, X. Sun, P. Liu, L. Wang, and R. Tang "Underwater acoustic target classification based on LOFAR spectrum and convolutional neural network," *Proc. 2<sup>nd</sup> Int. Conf. AIAM*, 59-63 (2020).
13. H. Yang, J. Li, S. Shen, and G. Xu, "A deep convolu-

- tional neural network inspired by auditory perception for underwater acoustic target recognition,” *Sensors* **19**, 1104 (2019).
14. Z. Wei, Y. Ju, and M. Song, “A method of underwater acoustic signal classification based on deep neural network,” *Proc. 5<sup>th</sup> ICISCE*, IEEE Computer Society, 46-50 (2018).
  15. M. M. Ahsan, M. A. P. Mahmud, P. K. Saha, K. D. Gupta, and Z. Siddique, “Effect of data scaling methods on machine learning algorithms and model performance,” *Technologies*, **9**, 52 (2021).
  16. W. E. Oh, “Comparison of environmental sound classification performance of convolutional neural networks according to audio preprocessing methods” (in Korean), *J. Acoust. Soc. Kr.* **39**, 143-149 (2020).
  17. G. M. Wenz, “Acoustic ambient noise in the ocean : spectra and sources,” *J. Acoust. Soc. Am.* **34**, 1936-1956 (1962).
  18. V. Nair and G. E. Hinton, “Rectified linear units improve restricted Boltzmann machines,” *Proc. 27<sup>th</sup> ICML*, 807-814 (2010).
  19. X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” *Proc. 13<sup>th</sup> Int. Conf. AISTATS*, 249-256 (2010).
  20. B. McFee, C. Raffel, D. Liang, D. Ellis, M. Mcvcar, E. Battenberg, and O. Nieto, “Librosa: Audio and music signal analysis in python,” *Proc. 14<sup>th</sup> Python Sci. Conf.* 18-24 (2015).
  21. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, and D. Cournapeau, “Scikit-learn: Machine learning in Python,” *J. Mach. Learn Res.* **12**, 2825-2830 (2011).
  22. C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H van Kerkwijk, M. Brett, A. Haldane, J. Fernández Del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard., T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E Oliphant, “Array programming with NumPy,” *Nature*, **585**, 357-362 (2020).
  23. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, Cr. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, “Tensorflow: Large-scale machine learning on heterogeneous distributed systems,” *arXiv preprint arXiv:1603.04467* (2016).
  24. D. P. Kingma and J. Ba, “Adam: a method for stochastic optimization,” *Proc. 3<sup>th</sup> ICLR*, 1-13 (2015).

## 저자 약력

### ▶ 박 경 민 (Kyung-Min Park)



2003년 2월 : 대한민국 해군사관학교 전자공학과 학사  
 2012년 2월 : 서울대학교 전기·컴퓨터공학부 석사  
 2017년 9월 ~ 현재 : 한국해양대학교 해양공학과 박사과정

### ▶ 김 두 영 (Dooyoung Kim)



2007년 2월 : 대한민국 해군사관학교 전산학과 학사  
 2012년 2월 : 서울대학교 전기·컴퓨터공학부 석사  
 2019년 2월 : KAIST 전산학부 박사  
 2019년 2월 ~ 현재 : 해군사관학교 인공지능학과 조교수