

Convolutional neural network 기법을 이용한 턱수염물범 신호 판별

Classification of bearded seals signal based on convolutional neural network

김지섭,¹ 윤영글,¹ 한동균,² 나형술,² 최지웅^{3,4†}

(Ji Seop Kim,¹ Young Geul Yoon,¹ Dong-Gyun Han,² Hyoung Sul La,² and Jee Woong Choi^{3,4†})

¹한양대학교 해양융합학과, ²한국해양과학기술원 극지연구소, ³한양대학교ERICA 해양융합공학과,
⁴한양대학교ERICA 국방정보공학과

(Received January 18, 2022; revised February 25, 2022; accepted March 4, 2022)

초 록: 수동 음향 관측을 통해 수집된 방대한 양의 데이터에서 해양포유류의 소리를 탐지하고 식별하기 위해 합성곱 신경망(Convolutional Neural Network, CNN)을 활용한 연구가 많이 수행되고 있다. 본 연구는 2017년 8월부터 2018년 8월까지 동시베리아 해에서 수집된 수중음향 스펙트럼 이미지를 기반으로 CNN을 활용하여 턱수염물범 소리의 분류 자동화 가능성을 확인해 보았다. 학습 데이터로서 다른 소음이 거의 포함되지 않은 뚜렷한 턱수염물범 소리를 사용하였을 때, 암기로 인한 과적합이 발생하였다. 일부 데이터를 소음이 포함된 데이터로 교체하여 학습시켜 수집된 전체 데이터로 평가한 결과 정확도(0.9743), 정밀도(0.9783), 재현율(0.9520)으로 모델이 이전보다 일반화되어 과적합이 방지되는 것을 확인하였다. 본 연구를 통해 물범신호 분류는 학습 데이터에 소음이 포함되었을 때 성능이 증가하는 것으로 나타났다.

핵심용어: 수동 음향 관측, 턱수염물범, 딥러닝, 합성곱 신경망, 식별

ABSTRACT: Several studies using Convolutional Neural Network (CNN) have been conducted to detect and classify the sounds of marine mammals in underwater acoustic data collected through passive acoustic monitoring. In this study, the possibility of automatic classification of bearded seal sounds was confirmed using a CNN model based on the underwater acoustic spectrogram images collected from August 2017 to August 2018 in East Siberian Sea. When only the clear seal sound was used as training dataset, overfitting due to memorization was occurred. By evaluating the entire training data by replacing some training data with data containing noise, it was confirmed that overfitting was prevented as the model was generalized more than before with accuracy (0.9743), precision (0.9783), recall (0.9520). As a result, the performance of the classification model for bearded seals signal has improved when the noise was included in the training data.

Keywords: Passive acoustic monitoring, Bearded seal, Deep learning, Convolution neural network, Classification

PACS numbers: 43.60.Lq, 43.30.Sf

I. 서 론

최근 머신러닝 기법이 크게 발전함에 따라 다양한 분야에서 이 기법을 적용하기 위해 많은 연구들이

수행되고 있다.^[1] 음향 신호 처리 및 음성 인식 분야에서 기계학습을 활용하여 다양한 연구가 수행되고 있으며, 음향 신호에 대한 과학적 및 공학적 이해를 바탕으로 기존 데이터에 기계학습 기법을 결합한

†Corresponding author: Jee Woong Choi (choijw@hanyang.ac.kr)

Department of Marine Sciences and Convergence Engineering, Hanyang University, 55 Hanyangdaehak-ro, Sangnok-gu, Ansan, Gyeonggi Province 15588, Republic of Korea

(Tel: 82-31-400-5531, Fax: 82-31-400-5457)



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.

최근 연구들은 지속적으로 발전하고 있다.^[1] 수중 음향 분야에서도 다양한 기계학습 기법을 활용한 연구가 시도되고 있으며 수중 음원 및 표적의 탐지/분류 및 위치 파악과 같은 기존의 원격 감시와 해양환경 정보 추출을 위한 음향 신호처리 분야 등 다양하게 적용되고 있다.^[2-5] 특히 시간과 주파수 측면에서 소리를 시각화하여 파악하는 스펙트로그램을 입력 데이터로 하여, 합성곱 신경망(Convolutional Neural Network, CNN) 기법을 통해 신호들을 분류하는 사례들이 점차 늘어나고 있다.^[6-8]

합성곱 신경망 기법이란 기계학습에 속한 딥러닝 기법의 하나로써 합성곱 연산을 통한 필터링 기법을 인공신경망에 적용해 이미지를 더욱 효과적으로 처리하기 위해 처음 제안되었다.^[9] 이미지들을 데이터로 입력 받아 시간이나 위치에 관계없이 동일하게 반복되는 특징들을 스스로 학습하여, 이미지의 분류나 객체 식별에서 뛰어난 성능을 보인다. 이 기법은 여러 연구자들에 의해 단순화 되고 발전하면서, 음성/오디오 인식, 객체 인식, 동영상 및 이미지의 검색/분류, 음악의 분류 등 여러 분야에서 대표적인 분류 및 인식 알고리즘으로써 사용되고 있다.^[10-12]

수동 음향 관측(Passive Acoustic Monitoring, PAM) 방법을 활용하는 해양포유류 연구에서도 해양포유류 소리를 탐지하고 식별하기 위해 합성곱 신경망이 많이 활용되고 있다.^[12,13] 수동 음향 관측은 해양 포유류를 연구하기 위한 효율적인 기술로 장기모니터링을 통해 대량의 수중음향 데이터를 지속적으로 획득할 수 있다.^[12,14] 이 방법을 통해 수집된 대량의 데이터를 분석해야하므로, 녹음된 음향신호에서 해양포유류의 소리를 탐지하고 분석하기 위한 자동화 시스템에 대한 연구의 필요성이 증가하고 있다.^[15,16] 일반적으로 다양한 해양포유류 신호의 특성을 추출하고 추출된 음향특성을 기반으로 분류하지만, 기존의 고정된 기능 기반의 분류 방법은 몇 가지 단점들로 인해 성능이 떨어진다.^[12] 탐지 또는 분류 성능은 특징 추출에 크게 좌우되는 반면, 수동으로 특징 추출을 설계하려면 신호 처리 기술을 갖춘 전문가가 필요하며 많은 시간이 소요된다.^[12,17] 합성곱 신경망은 스펙트로그램 내에서 발성의 유무를 간단히 결정할 수 있으며, 시간과 주파수 측면에서 음향신호를 정

확하게 탐지할 수 있어 종을 분류할 수 있다는 장점을 유지한다.^[15]

본 연구에서는 북극에서 측정한 수중 음향 데이터를 이용하여 턱수염물범을 대상으로 합성곱 신경망 기법을 사용하여 발성 분석의 자동화 가능성을 확인해 보았다. 또한 1년간의 수동 음향 관측 측정 데이터로 턱수염물범의 시간별 등장 유무를 통해 계절적 변동성 특성을 확인하여, 아직 많은 연구가 되지 않은 동시베리아 해에서 턱수염물범의 생태학적 특성을 유추할 수 있는 기초자료를 확인하였다.

II. CNN 모델의 이미지 학습 및 평가 방법

이미지는 통계적인 특성이 시간에 따라 변하지 않으며, 한 픽셀과 관련있는 픽셀들은 그 픽셀 주변에 제한되어 있다.^[18] 합성곱 신경망은 이러한 이미지의 특성이 변하지 않는다는 가정하에 이미지들을 데이터로 입력 받아 시간이나 위치에 관계없이 동일하게 반복되는 특징들을 스스로 학습해 이미지의 분류나 객체 인식 등 여러 분야에서 뛰어난 성능을 보여주고 있다.^[19]

합성곱 신경망이 이미지 특징 추출에 뛰어난 성능을 나타내는 이유는 합성곱 연산의 특징과 깊은 관련이 있다. 합성곱 연산을 통해 이미지의 특정 부분을 이미지와 필터의 내적 연산에 의해 큰 값을 더욱 크게, 작은 값을 더욱 작게 하여 특정 부분에 존재하는 이미지의 패턴들과 유사한 형태로 학습하기 때문이다. 또한 필터는 이미지 위를 이동하며 필터의 파라미터를 공유하며 연산을 진행하는데 이로 인해 위치를 고려하지 않고 이미지의 패턴을 추출해 학습할 수 있어, 위치를 고려하여 특징을 학습하는 방식보다 더 적은 연산량으로 이미지를 다룰 수 있다.

합성곱 신경망은 여러 역할을 가지는 층들을 예를 들어 Input layer, Convolution layer, Activation layer, Pooling layer, Fully Connected layer, Dropout layer, Classification layer 등으로 구성된다(Fig. 1). Input layer는 이미지를 입력받는 층으로써 연구자가 설정한 너비와 높이의 이미지를 입력 받는다. Convolution layer는 입력받은 이미지를 필터와의 합성곱 연산을 통해

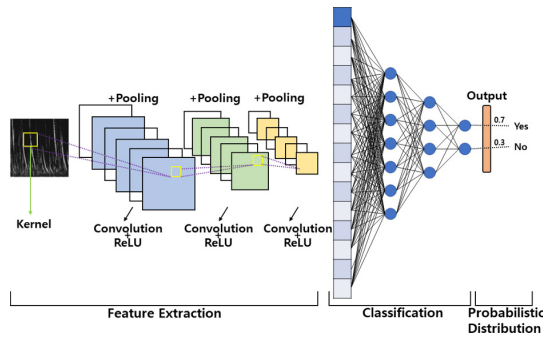


Fig. 1. (Color available online) Typical convolutional neural network architecture.

Table 1. Configuration of CNN model used in this study.

| Order | Type | Size |
|-------|-------------|---------------------------|
| 1 | Input | $100 \times 100 \times 1$ |
| 2 | Convolution | $96 \times 96 \times 12$ |
| 3 | ReLU | $96 \times 96 \times 12$ |
| 4 | Avg pooling | $48 \times 48 \times 12$ |
| 5 | Convolution | $44 \times 44 \times 16$ |
| 6 | ReLU | $44 \times 44 \times 16$ |
| 7 | Avg pooling | $22 \times 22 \times 16$ |
| 8 | Convolution | $20 \times 20 \times 20$ |
| 9 | ReLU | $20 \times 20 \times 20$ |
| 10 | Avg pooling | $10 \times 10 \times 20$ |
| 11 | Dropout | $10 \times 10 \times 20$ |
| 12 | FC | 150 |
| 13 | ReLU | 2 |
| 14 | FC | 2 |
| 15 | Softmax | Object, Background |

특징 값들을 추출하고, 추출된 특징 값들로 특징 맵을 생성한다. Activation layer는 활성화 함수를 사용하여 Convolution layer에서 받은 각각의 특징 값들을 활성화, 혹은 비활성화 시키는 층이다. Pooling layer는 인접한 픽셀들 중 가장 특징적인 값만 뽑아내 파라미터들을 줄여 연산량을 감소시키고 학습 데이터와의 과적합을 막는 역할을 한다. Dropout layer는 모델의 학습 중 Fully Connected layer와 연결된 노드들의 일부 연결을 끊어 모델이 과적합 되는 것을 방지하는 층이다.^[20] Fully Connected layer는 앞의 과정을 거치면서 최종적으로 추출된 특징 맵들을 입력으로 받아 역전파 알고리즘을 통해 가중치와 편향값을 계속해서 업데이트 한다. Classification layer는 이전 층

| | | Predicted | |
|--------|-------|---------------------|---------------------|
| | | True | False |
| Actual | True | True Positive (TP) | False Negative (FN) |
| | False | False Positive (FP) | True Negative (TN) |

Fig. 2. (Color available online) Confusion matrix to evaluate performance.

에서 계산된 값들을 사용하여, 처음 입력으로 받았던 이미지가 분류하고자 하는 클래스와 일치하는 정도를 확률로써 나타내게 된다.

본 연구에서는 MATLAB(Mathworks, USA)의 Deep Learning Toolbox와 Parallel Computing Toolbox를 사용하였으며, CNN 모델의 구성은 Table 1과 같다. 네트워크 학습 시 활성화 함수는 ReLU로 사용하였으며, Epoch는 130으로 설정하였다. Mini batch size는 64로 설정하였으며 최적화함수는 Adam을 사용하였다. 입력 이미지는 흑백 이미지를 사용하였으며, 입력 크기를 리샘플링하여 $100 \times 100 \times 1$ 로 축소하여 연산량을 줄였다. 입력 이미지의 크기가 작아 발생할 수 있는 과적합 가능성을 줄이기 위해 층을 깊지 않게 구성하였으며 Fully Connected layer의 앞에 Dropout layer를 추가하였다.^[21,22]

본 모델의 검증은 머신러닝 분야에서 사용하는 컨퓨전 행렬(Fig. 2)을 적용하여 정확도, 정밀도, 재현율로 평가하였다. 정확도는 모든 데이터 중 모델이 정답을 맞춘 비율을 의미한다. 정밀도는 모델이 True라고 판단한 것 중 데이터가 True인 것의 비율이다. 재현율은 데이터가 True인 것 중 모델이 True라고 답한 것의 비율이다.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}. \quad (1)$$

$$Precision = \frac{TP}{TP + FP}. \quad (2)$$

$$Recall = \frac{TP}{TP + FN}. \quad (3)$$

정밀도는 모델의 관점에서, 재현율은 데이터의 관점에서 모델을 평가하며, 둘 중 하나만 높다고 좋은 모델이 아니라 둘 모두 높았을 때 좋은 모델이라 할 수 있다.

III. 데이터 획득 및 학습자료 구성

음향자료는 북극 동시베리아 해에서 2017년 8월 21일부터 2018년 8월 13일까지 약 1년 동안 저장되었다. 수심이 약 62 m로 비교적 얇은 동시베리아 대륙붕에서 52 m 수심에 음향이탈기와 함께 채류되었다(74° 37.332'N, 174° 56.382'E). 사용된 자가기록식 수중청음기(Aural-M2, Multi-Electronique Inc., Canada)는 32,768 Hz 샘플링주기로 설정하였으며, 장비의 저장 공간과 배터리 용량의 제한으로 매시간 10 min 동안의 음향자료를 저장하였다. 수중청음기의 수신민감도는 텍수염물범의 발성주파수 내에서는 선형이다.^[23]

텍수염물범의 주요 발성은 트릴(trill), 상승(ascent), 소인(sweep), 신음(moan)의 4가지로 분류되며, 번식기 동안 수컷 수염 바다표범이 내는 트릴이 주요한 발성이다.^[24,25] 트릴은 주파수가 급격히 감소하였다가 점차적으로 천천히 감소하는 긴 down-sweep 발성이며(Fig. 3),^[25] 합성곱 신경망 패턴 인식으로 텍수염물범 유무를 판단할 수 있는 특성을 가진다.

본 연구에서는 10 min 동안의 음향 자료로 그린 스펙트럼 이미지들을 입력으로 사용하여 모델을 학습시켰다(Fig. 4). 노이즈(Noise)는 스펙트럼 이미지에서 텍수염물범 신호가 존재하지 않는 데이터들이며, 시그널(Signal)은 약간의 저주파 소음은 존재하나 텍수염물범 신호가 뚜렷하게 구분되는 이미지들이며, 시그널 + 노이즈는 스펙트럼 이미지에서 텍수염물범 신호를 육안으로는 식별되나 다른 소음들로 인해 신호가 겹쳐진 데이터들을 의미한다.

시그널 데이터만을 사용하여 모델을 학습시켰을 때 과적합이 발생하는 경향이 보였다. 과적합은 모델의 가중치 벡터들이 학습 데이터에 거의 완벽하게 맞춰져 학습 데이터에서는 잘 작동하지만 실제 데이터에서는 오차가 증가하는 현상을 의미한다.^[26] 따라서 모델을 더 일반화하여 과적합을 방지하기 위해

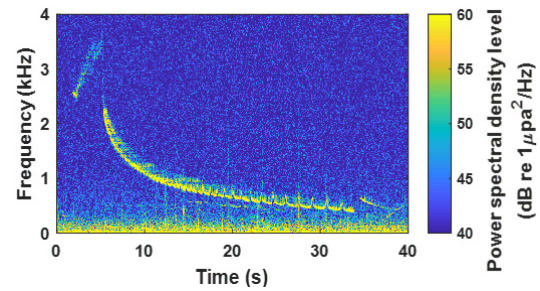


Fig. 3. (Color available online) An example of spectrogram of a bearded seal trill.

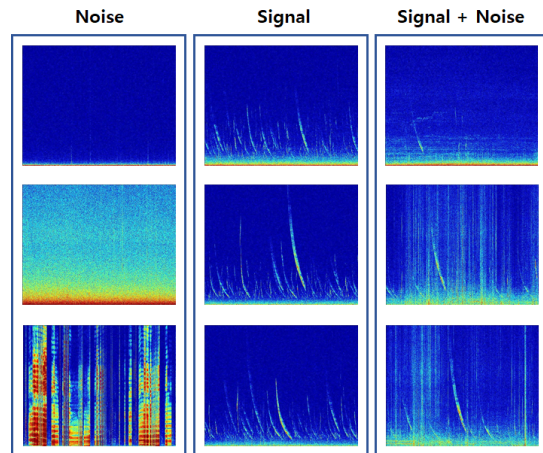


Fig. 4. (Color available online) Example images of training data of ambient noise, clear seal sound, seal sound with other noise.

시그널 데이터의 일부를 시그널 + 노이즈 데이터로 교환하여 학습 데이터를 구성하였다.^[27]

따라서 모델은 2가지로 구분하였으며, 시그널 데이터 958개, 노이즈 데이터 831개로 학습을 진행한 모델과 시그널 + 노이즈 데이터 316개와 시그널 데이터 646개, 노이즈 데이터 831개로 학습을 진행한 모델로 구성되었다. 각각의 데이터를 7:3의 비율로 나누어 학습 데이터셋, 교차 검증 데이터셋을 구성하여 교차 검증을 실시했다. 또한 1년간의 수중 소음 스펙트럼 이미지를 데이터 모델에 입력하여 결과를 도출하고, 육안으로 판별한 결과와 검증하였다.

IV. 결과 및 토의

동시베리아 해에서 획득한 8,550개의 음향신호 스펙트럼 이미지를 사용하여 합성곱 신경망 기법으로

Table 2. CNN results - validation data set.

| Input data (number) | Accuracy | Precision | Recall |
|--------------------------------------|----------|-----------|--------|
| Signal (958) | 0.9925 | 0.9862 | 1 |
| Signal (646) Signal + Noise (316) | 0.9310 | 0.9596 | 0.9094 |

Table 3. CNN results - 1 year data set collected in east siberian sea.

| Input data (number) | Accuracy | Precision | Recall |
|--------------------------------------|----------|-----------|--------|
| Signal | 0.9476 | 0.9896 | 0.8677 |
| Signal (646) Signal + Noise (316) | 0.9743 | 0.9783 | 0.9520 |

턱수염물범의 발성을 판별하였다.

시그널 데이터로 구성된 데이터셋으로 학습한 모델에 대한 교차검증 성능 지표는 정확도 0.9925, 정밀도 0.9862, 재현율 1로 나타났다(Table 2). 그러나 동시베리아 해에서 수집한 1년치 데이터에 적용한 결과 정확도 0.9478, 정밀도 0.9896, 재현율 0.8677로 재현율이 교차검증 결과에 비해 매우 낮게 나타났다(Table 3). 재현율이 매우 낮게 나타나 데이터셋 내에 존재하는 데이터는 잘 예측하지만 그 이외의 데이터는 잘 예측하지 못하는 모습을 보여 데이터셋에 과적합된 것을 확인할 수 있었다.

따라서 일반화 성능을 높이기 위해 일부 시그널 데이터를 데이터셋에서 제외하고 시그널 + 노이즈 데이터로 교체하였다. 시그널 데이터와 시그널 + 노이즈 데이터를 합친 데이터셋으로 학습시킨 모델의 교차검증 성능 지표 결과는 정확도 0.9310, 정밀도 0.9596, 재현율 0.9094로 나타나, 시그널로 학습한 모델의 교차검증 성능 지표보다 낮게 나타났다(Table 2). 그러나 1년치 데이터에 적용한 결과 정확도 0.9743, 정밀도는 0.9783, 재현율이 0.9520로 모든 성능 지표가 교차검증 성능 보다 증가하여, 모델이 데이터셋에 의해 과적합되지 않고 턱수염물범 소리의 일반적인 특징을 잘 학습한 것으로 생각된다(Tables 2, 3).

각 모델의 성능을 1년치 데이터로 비교한 결과, 시그널 데이터로 학습한 결과보다 시그널 + 노이즈 데이터를 데이터셋에 포함하였을 때 일부 향상된 성능이 나타났다(Table 3). 분석에 사용된 동시베리아 해의 수중 소음은 해빙, 에어건 등의 다양한 원인으로 발생한 수중소음이 포함되어 있는 환경이므로,^[23] 모

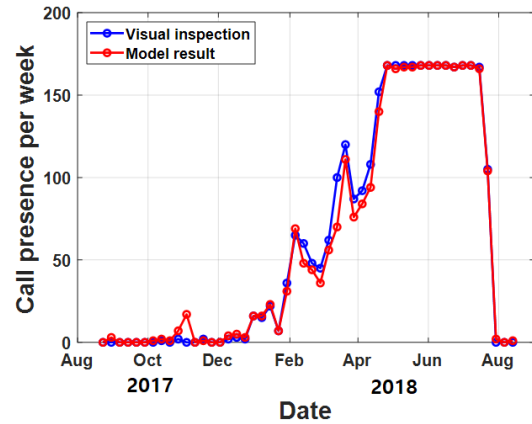


Fig. 5. (Color available online) Call presence result of bearded seals from August 2017 to August 2018.

델의 성능을 향상시키기 위해서는 일반화가 중요한 역할을 담당한 것으로 생각된다.

마지막으로 합성곱 신경망 기법에 의해서 판별된 턱수염물범 소리의 신뢰성을 확인하기 위하여 모델에 의해 도출된 결과와 육안으로 판별한 결과를 기간별로 비교하였다(Fig. 5). x축은 일주일 단위로 표기하였으며, 1시간을 대표하는 10분간 턱수염물범의 유무판별 결과를 일주일로 누적하여 최소 0 최대 168의 범위에서 구분된다. 일부 시기에서 오탐지 또는 미탐지 되는 경우가 발생하였으나, 물범 소리의 존재 여부에 대한 경향은 비교적 잘 일치하는 것으로 나타났다.

본 연구 결과로는 물범의 정량적인 개체수는 알 수 없지만 계절에 따른 측정 지역으로의 출현 여부는 확인할 수 있었다. 8월과 9월에는 소리는 녹음되지 않았으며, 4월에서 7월사이 이들의 음향 활동은 매우 활발하게 나타났다. 동시베리아 해와 지역적으로 멀지 않은 축치 해에서 측정된 턱수염물범 소리는 7,8월에는 존재하지 않았으며 3월에서 6월사이에 소리 존재율이 최대로 나타났으므로,^[28] 한 달 정도의 분포 시기 차이가 있는 것으로 생각된다.

본 연구에서는 수중소음 데이터에서 합성곱 신경망을 이용하여 턱수염물범의 음향 신호 분류를 시도하였으며, 합성곱 신경망의 자동적으로 특징을 학습하는 기능으로 인해 이미지 기반 분류 작업에서 우수한 성능을 보였다. 또한 다양한 소음이 발생하는 해양환경에서 해양포유류 소리 분류의 정확성을 높이

기 위해서는 소음을 포함시켜 학습 시키는 것이 효과적으로 활용될 수 있다는 것을 확인하였다. 본 논문은 수중소음의 스펙트럼 이미지에서 합성곱 신경망 기법을 활용하여 해양포유류 소리를 추출해 낼 수 있음을 확인한 기초적인 연구이다. 추후 분류할 해양 포유류 종의 수를 늘리고 각 신호에 대한 연구가 수행되어 실시간 자동 탐지 및 분류도 가능해 진다면 수동 음향 관측이 더 효율적으로 활용될 것으로 생각한다.

감사의 글

이 논문은 2021년도 정부(산업통상자원부)의 재원으로 한국에너지기술평가원의 지원(20203030020080, 해상풍력 단지 해양공간 환경 영향 분석 및 데이터베이스 구축)과 2022년도 정부(해양수산부)의 재원으로 해양수산과학기술진흥원의 지원(1525011760, 북극해 온난화-해양생태계 변화 감시 및 미래전망 연구)을 받아 수행된 연구임.

References

1. S. Lee, M. Kim, and Y. Jeong, "A study on the waveform-based end-to-end deep convolutional neural network for weakly supervised sound event detection" (in Korean), *J. Acoust. Soc. Kr.* **39**, 24-31 (2020).
2. H. Yang, K. Lee, Y. Choo, and K. Kim, "Underwater Acoustic Research Trends with Machine Learning General Background," *J. Ocean Eng. Technol.* **34**, 147-154 (2020).
3. Y. Wang and H. Peng, "Underwater acoustic source localization using generalized regression neural network," *J. Acoust. Soc. Am.* **143**, 2321-2331 (2018).
4. C. Frederick, S. Villar, and Z. H. Michalopoulou, "Seabed classification using physics-based modeling and machine learning," *J. Acoust. Soc. Am.* **148**, 859-872 (2020).
5. S. Kim, S.-K. Jung, D. Kang, M. Kim, and S. Cho, "Application of the artificial intelligence for automatic detection of shipping noise in shallow-water" (in Korean), *J. Acoust. Soc. Kr.* **39**, 279-285 (2020).
6. S. Liu, M. Liu, M. Wang, T. Ma, and X. Qing, "Classification of cetacean whistles based on convolutional neural network," *Proc. IEEE 10th ICWCSP.* 1-5 (2018).
7. W. Luo, W. Yang, and Y. Zhang, "Convolutional neural network for detecting odontocete echolocation clicks," *J. Acoust. Soc. Am.* **145**, EL7-12 (2019).
8. D. Wang, L. Zhang, Z. Lu, and K. Xu, "Large-scale whale call classification using deep convolutional neural network architectures," *Proc. IEEE ICSPCC.* 1-5 (2018).
9. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, **1**, 541-551 (1989).
10. S. Behnke, *Hierarchical Neural Networks for Image Interpretation* (Springer, New York, 2003), pp. 44-46.
11. A. Maccagno, A. Mastropietro, U. Mazziotta, M. Scarpiniti, Y.-C. Lee, and A. Uncini, "A CNN approach for audio classification in construction sites," in *Handbook of Progresses in Artificial Intelligence and Neural Systems*, edited by A. Esposito, M. Faundez-Zanuy, F. C. Morabito, and E. Pasero (Springer, Singapore, 2021).
12. T. Lu, B. Han, and F. Yu, "Detection and classification of marine mammal sounds using AlexNet with transfer learning," *Ecological Informatics*, **62**, 101277 (2021).
13. Y. Chen, W. Wang, Y. Liang, D. Zhou, C. Dong, and J. Li, "Real-time detection and classification for targeted marine mammals," *Proc. 2021 OES China Ocean Acoustics*, 1027-1031 (2021).
14. S. M. Van Parijs, C. W. Clark, R. S. Sousa-Lima, S. E. Parks, S. Rankin, D. Risch, and I. C. Van Opzeeland, "Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales," *Marine Ecology Progress Series*, **395**, 21-36 (2009).
15. M. Thomas, B. Martin, K. Kowarski, B. Gaudet, and S. Matwin, "An end-to-end approach for true detection of low frequency marine mammal vocalizations," *J. Acoust. Soc. Am.* **146**, 2959 (2019).
16. M. Thomas, B. Martin, K. Kowarski, B. Gaudet, and S. Matwin, "Marine mammal species classification using convolutional neural networks and a novel acoustic representation," *Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, 290-305 (2019).
17. C. Li, R.V. Sanchez, G. Zutita, M. Cerrada, D. Cabrera, and R. E. Vasquez, "Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis," *Neurocomputing*, **168**, 119-127 (2015).
18. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Proc. Advances in neural information processing systems*, **25**, 1097-1105 (2012).

19. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, **86**, 2278-2324 (1998).
20. G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv preprint arXiv:1207.0580* (2012).
21. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE CVPR*, 770-778 (2016).
22. A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Proc. Advances in Neural Information Processing Systems*, 1097-1105 (2012).
23. D.-G. Han, J. Joo, W. Son, K. H. Cho, J. W. Choi, E. J. Yang, J.-H. Kim, S.-H. Kang, and H. S. La, "Effects of geophony and anthrophony on the underwater acoustic environment in the East Siberian Sea, arctic ocean," *Geophysical Research Letters*, **48**, e2021GL093097 (2021).
24. D. Risch, C. W. Clark, P. J. Corkeron, K. M. Elepfandt, K. M. Kovacs, C. Lydersen, I. Stirling, and S. M. van Parijs, "Vocalizations of male bearded seals, *Erignathus barbatus*: Classification and geographical variation," *Animal Behavior*, **73**, 747-762 (2007).
25. I. Charrier, N. Mathevon, and T. Aubin, "Bearded seal males perceive geographic variation in their trills," *Behav. Ecol. Sociobiol.* **67**, 1679-1689 (2013).
26. G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv Tech. Rep.*, 2012.
27. J. Sietsma and R. Dow "Creating artificial neural networks that generalize," *Neural Networks*, **4**, 67-79 (1991).
28. M. Jimbo, D. Mizuguchi, H. Shirakawa, K. Tsujii, A. Fujiwara, K. Miyashita, and Y. Mitani, "Seasonal variations in the call presence of bearded seals in relation to sea ice in the southern Chukchi Sea," *Polar Biol.* **42**, 1953-1958 (2019).

저자 약력

▶ 김 지 섭 (Ji Seop Kim)



2021년 2월 : 한양대학교 해양환경공학과
공학사
2021년 3월 ~ 현재 : 한양대학교 해양환경
과학과 석사과정

▶ 윤 영 글 (Young Geul Yoon)



2012년 2월 : 한양대학교 해양환경과학과
학사
2014년 2월 : 한양대학교 해양환경과학과
석사
2014년 3월 ~ 현재 : 한양대학교 해양융합
과학과 박사과정

▶ 한 동 균 (Dong-Gyun Han)



2012년 2월 : 한양대학교 해양환경과학과
학사
2014년 2월 : 한양대학교 해양환경과학과
석사
2020년 2월 : 한양대학교 해양환경과학과
박사
2020년 6월 : 한양대학교 미래해양연구센
터 박사후연구원
2020년 6월 ~ 현재 : 한국해양과학기술원
부설 극지연구소 연수연구원

▶ 나 형 술 (Hyoung Sul La)



2001년 2월 : 한양대학교 지구해양과학과
학사
2005년 2월 : 한양대학교 지구해양과학과
석사
2009년 8월 : 한양대학교 해양환경과학과
박사
2015년 4월 : 극지연구소 연수연구원
2017년 3월 ~ 현재 : 과학기술연합대학교
부교수
2020년 3월 ~ 현재 : 극지연구소 책임연구원

▶ 최 지 웅 (Jee Wonng Choi)



1996년 2월 : 한양대학교 지구해양과학과
학사
1998년 2월 : 한양대학교 지구해양과학과
석사
2002년 8월 : 한양대학교 지구해양과학과
박사
2002년 11월 : Applied Physics Lab., Univ. of
Washington, USA, Research Associate
2007년 9월 ~ 현재 : 한양대학교 해양융합
공학과/국방정보공학과 교수
2020년 5월 ~ 현재 : 미국음향학회 석학
(Fellow) 회원