

## 해양환경 모니터링을 위한 순환 신경망 기반의 돌고래 클릭 신호 분류 알고리즘 개발

정서제<sup>1</sup> · 정우근<sup>2</sup> · 신성렬<sup>2</sup> · 김동현<sup>3</sup> · 김재수<sup>4</sup> · 변기훈<sup>5</sup> · 이다운<sup>6\*</sup>

<sup>1</sup>한국해양대학교 해양과학기술융합학과 박사과정

<sup>2</sup>한국해양대학교 에너지자원공학과 교수

<sup>3</sup>한국해양대학교 수중운동체특화연구센터 학연연구교수

<sup>4</sup>한국해양대학교 해양공학과 교수

<sup>5</sup>한국해양대학교 해양과학기술융합학과 교수

<sup>6</sup>한국해양대학교 에너지자원공학과 박사후연구원

### Development of Dolphin Click Signal Classification Algorithm Based on Recurrent Neural Network for Marine Environment Monitoring

Seoje Jeong<sup>1</sup>, Wookeen Chung<sup>2</sup>, Sungryul Shin<sup>2</sup>, Donghyeon Kim<sup>3</sup>, Jeasoo Kim<sup>4</sup>, Gihoon Byun<sup>5</sup>, and Dawoon Lee<sup>6\*</sup>

<sup>1</sup>Ph.D. student, Department of Convergence Study on the Ocean Science and Technology, Ocean Science and Technology (OST) School, Korea Maritime & Ocean University, Busan, Korea

<sup>2</sup>Professor, Department of Energy and Resources Engineering, Korea Maritime & Ocean University, Busan, Korea

<sup>3</sup>Research Professor, Underwater Vehicle Research Center, Korea Maritime & Ocean University, Busan, Korea

<sup>4</sup>Professor, Department of Ocean Engineering, Korea Maritime & Ocean University, Busan, Korea

<sup>5</sup>Professor, Department of Convergence Study on the Ocean Science and Technology, Ocean Science and Technology (OST) School, Korea Maritime & Ocean University, Busan, Korea

<sup>6</sup>Postdoctoral researcher, Department of Energy and Resources Engineering, Korea Maritime & Ocean University, Busan, Korea

#### 요약

본 연구에서는 해양 모니터링 중에 기록된 돌고래 클릭 신호를 분류하기 위해 순환 신경망(RNN)을 적용하는 방법을 검토했다. 클릭 신호 분류의 정확도를 높이기 위해 단일 시계열 자료를 분수 푸리에 변환을 사용하여 분수 영역으로 변환하여 특징을 확장했으며, 분류를 위한 최적의 네트워크를 결정하기 위해 세 가지 순환 신경망 모델(LSTM, GRU, BiLSTM)을 비교 분석하였다. 순환 신경망 모델의 입력 자료로써 이용된 분수 영역 자료의 경우, 분수 푸리에 변환 시 회전 각도에 따라 다른 특성을 가지므로, 각 네트워크 모델에 따라 우수한 성능을 가지는 회전 각도 범위를 분석했다. 이때 네트워크 성능 분석을 위해 정확도, 정밀도, 재현율, F1-점수와 같은 성능 평가 지표를 도입했다. 수치실험 결과, 세 가지 네트워크 모두 높은 성능을 보였으며, BiLSTM 네트워크가 LSTM, GRU에 비해 뛰어난 학습 결과를 제공했다. 마지막으로, 현장 자료 적용 가능성 측면에서 BiLSTM 네트워크가 다른 네트워크에 비해 낮은 오답지 결과를 제공하였다.

#### 주요어

분류, 돌고래 클릭 신호, 순환 신경망, 모니터링, 분수 푸리에 변환

#### ABSTRACT

In this study, a recurrent neural network (RNN) was employed as a methodological approach to classify dolphin click signals derived from ocean monitoring data. To improve the accuracy of click signal classification, the single time series data were transformed into fractional domains using fractional Fourier transform to expand its features. Transformed data were used as input for three RNN models: long short-term memory (LSTM), gated recurrent unit (GRU), and bidirectional LSTM (BiLSTM), which were compared to determine the optimal network for the classification of signals. Because the fractional Fourier transform displayed different

Received: 3 July 2023; Revised: 10 August 2023; Accepted: 10 August 2023

\*Corresponding author

E-mail: dwlee@gkmou.ac.kr

Address: Postdoctoral Researcher, Department of Energy and Resources Engineering, Korea Maritime & Ocean University, Busan, Korea

©2023, Korean Society of Earth and Exploration Geophysicists

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

characteristics depending on the chosen angle parameter, the optimal angle range for each RNN was first determined. To evaluate network performance, metrics such as accuracy, precision, recall, and F1-score were employed. Numerical experiments demonstrated that all three networks performed well, however, the BiLSTM network outperformed LSTM and GRU in terms of learning results. Furthermore, the BiLSTM network provided lower misclassification than the other networks and was deemed the most practically applicable to field data.

## KEYWORDS

classification, dolphin click signals, recurrent neural network, monitoring, fractional Fourier transform

## 서 론

해양환경에서 탐사 혹은 모니터링을 통해 수집된 자료에서는 인공송신원에 의한 신호뿐만 아니라 해양에 서식하는 다양한 해양생물의 신호가 기록된다. 해양생물에 의해 기록된 신호는 해양 생태계를 분석하기 위한 중요한 도구로써 해양환경 분석 자료 제공 및 해상풍력단지 부지선정 등의 다양한 분야에 적용할 수 있다(Jin *et al.*, 2022). 또한, 탄성과 탐사를 통해 이용되는 인공송신원의 주파수 대역은 일부 해양생물의 주파수 대역에 포함되어 해양생물에 영향을 미치며, 기록된 해양생물의 신호 주기를 분석함으로써 해양생물에 미치는 영향을 분석할 수 있다(Kavanagh *et al.*, 2019). 국내 해역에서는 해양생물 중 딱총새우에 의한 스냅 신호(snap signal)와 돌고래에 의한 클릭 신호(click signal)가 주로 관측되며, 클릭 신호는 충격파(impulse) 형태의 신호로써 나타난다(Wenz, 1962; Radford *et al.*, 2008). 모니터링 자료에서 기록된 충격파 형태의 클릭 신호를 해양 생태계 분석을 위한 자료로써 활용하기 위해서는 클릭 신호의 특성에 대한 정확한 이해와 분류가 요구된다. 그러나 모니터링 자료에는 클릭 신호뿐만 아니라 바람, 파도, 난류에 의해 기록되는 주변 잡음(ambient noise)이 함께 기록되므로, 이를 분류하기 위한 알고리즘이 요구된다. 기존 신호 분류 알고리즘의 경우 자료처리자의 주관적인 판단이 들어가므로 부정확한 자료처리 결과를 얻게 될 위험이 있다(Sugimatsu *et al.*, 2015). 이러한 한계를 극복하고자 일관성 있는 기준을 갖는 자동 분류 시스템이 연구되었으며, 최근 수중음향학 분야에서 기계학습을 도입하여 우수한 성능을 갖는 신호 분류 모델을 제안되었다(Carusio *et al.*, 2020; Jin *et al.*, 2022; Ditria *et al.*, 2022). Jiang *et al.* (2019)는 고래 신호를 탐지 및 분류하기 위해 합성곱 신경망(convolution neural network, CNN)을 훈련시켰고, Caruso *et al.* (2020)은 생물학적 신호 탐지를 위해 전달 손실(transmission loss, TL) 모델과 서포트 벡터 머신(support vector machine, SVM)을 사용하였으며, Thomas *et al.* (2019), Jin *et al.* (2022), 그리고 Kim *et al.* (2022b)은 합성곱 신경망에 입력 자료로써 단시간 푸리에 변환(short-time Fourier transform, STFT)으로 변환된 스펙트로그램을 활용하여 해양생물의 신호를 성공적으로 검출하였다. 그러나 서포트 벡터 머신과 합성곱 신경망의 경우, 자료의 시간적 순서를 고려하지 않으므로 실시간으로 기록되는 정보들의 순서 관계를

처리할 수 없다는 문제점이 있다(Schmidhuber, 2015).

본 연구에서는 해양에서 실시간으로 취득되는 시계열 자료의 시퀀스 특징을 고려하여 돌고래의 클릭 신호를 분류하기 위해 순환 신경망을 이용하고자 하였다. 순환 신경망은 텍스트, 음성 단어와 같은 시퀀스 특징을 가지는 자료의 시간적 순서를 고려하여 입력된 자료의 반복되는 특징들을 스스로 학습하는 인공신경망이다(Graves *et al.*, 2013). 순환 신경망 알고리즘은 대표적으로 LSTM (long short-term memory) (Hochreiter and Schmidhuber, 1997), GRU (gated recurrent unit) (Dey and Salem, 2017; Cho *et al.*, 2014), BiLSTM (bidirectional long short term memory) (Schuster and Paliwal, 1997) 네트워크가 있으며, 각 네트워크는 상황과 자료에 따라 서로 다른 성능을 가지므로 최적의 네트워크 선정을 위해서는 학습을 통한 직접적인 비교가 필요하다(Jozefowicz *et al.*, 2015; Kim and Kim, 2022). 따라서, 본 연구에서는 돌고래 클릭 신호 분류를 위해 동일한 학습 자료를 이용하여 세 가지 네트워크를 학습하였으며, 결과를 비교 분석하여 최적의 네트워크를 선정하고자 하였다.

모니터링을 통해 기록된 시계열 자료에서 돌고래의 클릭 신호를 정확하게 분류하기 위해서는 시계열 자료에서 클릭 신호의 적합한 특징(feature) 선정이 필요하다. 하지만, 단일 시계열 자료만을 이용하여 목적하는 특징을 추출하는 것은 어렵기 때문에, 다양한 영역으로 시계열 자료를 변환함으로써 원활히 특징을 추출하기 위한 연구가 수행되고 있다(Saxena and Singh, 2005; Tian, 2021). 돌고래 클릭 신호의 경우 주변 잡음과 다른 충격파 형태의 신호를 가지고 있어, 넓은 영역으로 주파수 성분이 분포하는 특징을 가지고 있다. 이러한 돌고래 클릭 신호와 주변 잡음과의 주파수 특성 차이를 이용하여 돌고래 클릭 신호를 분류하기 위해, 본 연구에서는 시계열 자료의 시간-주파수 분포 특성을 파악할 수 있는 분수 푸리에 변환(fractional Fourier transform, FrFT)(Almeida, 1994)을 이용하고자 하였다.

분수 푸리에 변환의 경우 시계열 자료의 특성 분석에 활용되는 기법으로, 회전 각도에 따른 시간-주파수 분포 특성을 분석할 수 있으며 신호 처리에 있어서 기존의 푸리에 변환을 이용한 방법보다 우수한 결과를 보인다고 알려져 있다(Namias, 1980; Ozaktas and Kutay, 2001; Zhai, 2014). 변환을 위해 이용되는 회전 각도의 경우 0°는 시계열 특성을, 90°는 주파수

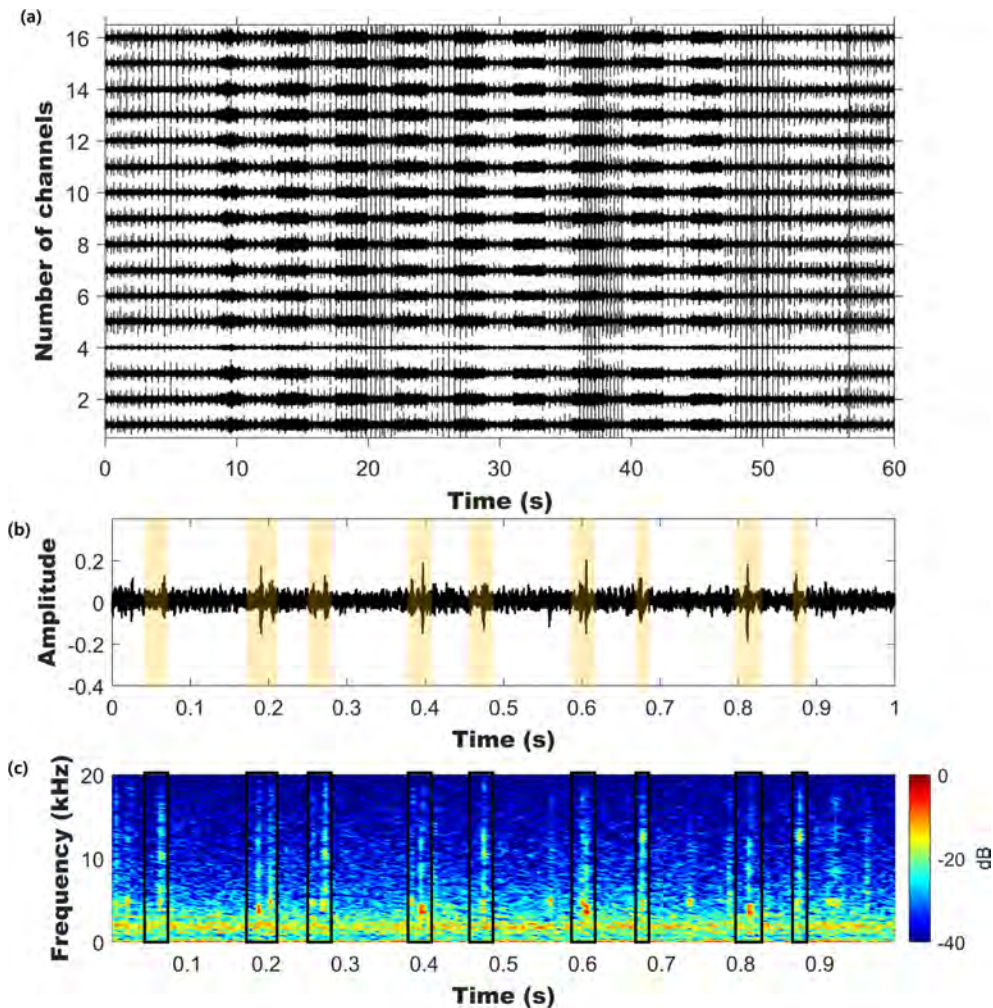
특성을 나타낸다. 시간-주파수 분포 특성을 입력 자료로써 사용하여 돌고래 클릭 신호를 분류하기 위해, 입력 자료 변환에 이용되는 회전 각도의 최대 범위를 9°에서부터 90°까지 고려하였다.

따라서, 본 연구에서는 해양환경 모니터링 자료에서 기록된 돌고래 클릭 신호를 분류하기 위해 순환 신경망 기반 자동 분류 알고리즘을 개발하고자 하였다. 또한, 개발된 알고리즘의 세 가지 네트워크에서 우수한 성능을 보이는 회전 각도 범위를 선정하고, 돌고래 클릭 신호 분류를 위한 최적의 네트워크를 선정하고자 하였다. 이때, 최적의 네트워크 및 회전 각도 범위 선정을 위해, 정확도(accuracy), 정밀도(precision), 재현율(recall), F1-점수(F1-score)와 같은 분류 문제의 성능 평가 지표를 이용하였다(Read *et al.*, 2011; Cui *et al.*, 2022). 마지막으로, 선정된 최적의 네트워크 및 회전 각도 범위에 따라 학습된 네트워크를 현장 자료에 적용함으로써 개발된 알고리즘의 현장 자료 적용 가능성을 확인하였다.

## 이 론

### 클릭 신호 특징 및 음향 특성 분석

본 연구에서 사용된 음향 신호 자료는 2018년 포항 동쪽 해역에서 수행된 해상실험 자료이다(Kim *et al.*, 2022a). 음향 신호는 16개 수신기가 2.8 m 간격으로 구성된 42 m 길이의 수직 선 배열 센서를 통해 심도 179 m부터 221 m까지 기록되었다. 본 연구에서 사용된 음향 신호 자료는 최대 수심 1,550 m 해역에서 기록되었으므로, 저서 생활을 하는 딱총새우의 신호 영향을 고려하지 않았으며, 기록된 자료의 클릭 신호는 돌고래 클릭 신호로써 가정하였다. 이때, 일반적으로 돌고래가 발생시키는 클릭 신호는 수중에서 주변 잡음과 서로 다른 스펙트럼 특성을 가지는 것으로 알려져 있다(Au and Herzing, 2003; Frasier *et al.*, 2017). 또한, 동해 해역에서 취득되는 클릭 신호는 함께 기록되는 주변 잡음과 비교했을 때 매우 짧은 충격과 형태의 반복적인 신호로, 비교적 높은 주파수를 가지고 있다(Choi *et al.*, 2019; Kim *et al.*, 2022a). 본 연구에서는 순



**Fig. 1.** (a) Collection of 16 acoustic dolphin click signals from ocean monitoring data, (b) Enlarged view of the first acoustic signal within 0 ~ 1 s range, and (c) Frequency spectrum of signal displayed in (b).

환 신경망 알고리즘을 학습할 때, 기존 16,384 Hz의 고르기 비율(sampling rate)로 취득된 자료를 동일한 입력 크기에서 보다 자세한 시퀀스 특성을 제공할 수 있는 40,000 Hz로 다시 고르기(resampling)하였다. Fig. 1a는 해양에서 취득된 16개의 음향 신호를 나타내며, 1b는 16개의 음향 신호 중 첫 번째 음향 신호에서 1 s까지의 구간을 확대한 것이다. 1b의 강조된 구간을 보면, 클릭 신호가 주변 잡음과 비교했을 때 진폭이 약 2 배 이상 차이가 나타나는 것을 확인할 수 있다. 측정된 음향 신호의 주파수 특성을 분석하기 위해 단시간 푸리에 스펙트럼 분석을 수행한 결과, Fig. 1c와 같이 주파수 스펙트럼에서도 낮은 주파수 대역을 갖는 주변 잡음에 비해 상대적으로 높은 주파수 대역까지 클릭 신호가 나타남을 확인하였다.

### 분수 푸리에 변환

본 연구에서 시간-주파수 영역에서 일정한 패턴을 나타내는 클릭 신호의 특징을 활용하기 위해 분수 푸리에 변환을 통해 시계열 자료를 2차원 입력 자료로 변환하였다. 분수 푸리에 변환은 식 (1), (2)와 같이 계산된다(Almeida, 1994).

$$F_{\alpha}(\xi) = \int_{-\infty}^{+\infty} K_{\alpha}(\xi, t) f(t) dt, \quad (1)$$

$$K_{\alpha}(\xi, t) = \begin{cases} \sqrt{\frac{1 - j \cot(\alpha)}{2\pi}} e^{\frac{j \cot(\alpha) \xi^2}{2} - (j \csc(\alpha) \xi t + \frac{j \cot(\alpha) t^2}{2})} & \text{if } \alpha \text{ is not a multiple of } \pi \\ \delta(t - \xi) & \text{if } \alpha \text{ is a multiple of } \pi \\ \delta(t + \xi) & \text{if } \alpha + \pi \text{ is a multiple of } 2\pi \end{cases} \quad (2)$$

if  $\alpha$  is not a multiple of  $\pi$

if  $\alpha$  is a multiple of  $\pi$

if  $\alpha + \pi$  is a multiple of  $2\pi$

여기서  $\alpha$ 는 회전 각도,  $\xi$ 는 주파수 변수,  $K_{\alpha}$ 은 변환을 위한 커널,  $j$ 는 허수 단위,  $f(t)$ 는 변환하고자 하는 시간영역의 신호, 그리고  $F_{\alpha}(\xi)$ 는 회전 각도에 따라 변환된 신호를 나타낸다. 분수 푸리에 변환은 시계열 자료를  $\alpha$ 의 값에 따라 시간-

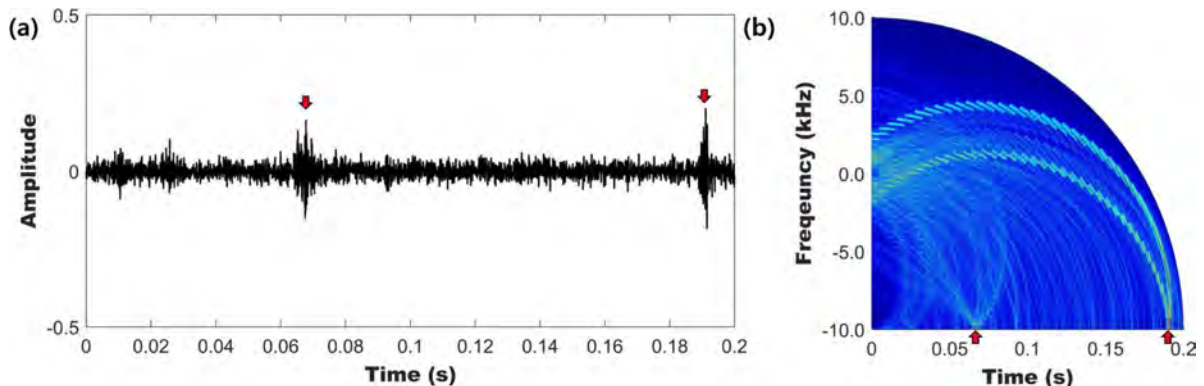
주파수 영역 모두에서 신호를 변환할 수 있으며, 회전 각도에 따른 시간-주파수 특성을 분석할 수 있다. Fig. 2a의 자료는 Fig. 1a의 자료에서 14번째 자료를 임의로 선정하여 0 ~ 0.2 s 구간을 추출한 것이다. Fig. 2a의 자료에 분수 푸리에 변환을 수행한 결과, Fig. 2b와 같이 저주파수 특징을 가지는 주변 잡음의 경우 회전 각도에 따라 저주파수 대역으로 신호가 집중되고, 클릭 신호의 경우 모든 주파수 대역에서 신호가 퍼지는 것을 확인할 수 있다. 특히, 주변 잡음과 진폭 차이가 약 2 배 차이 나는 0.06 s, 0.19 s 구간에서 주변 잡음과 비교하여 클릭 신호의 세기가 강하게 나타나는 패턴이 생성된 것을 확인할 수 있다. 이와 같이, 클릭 신호와 주변 잡음의 경우 서로 다른 패턴을 가지는 것을 확인할 수 있으며, 본 연구에서는 이러한 특성을 가진 회전 각도에 따른 시간-주파수 특성 자료를 학습을 위한 입력 자료로써 활용하고자 하였다.

### 수치 실험

본 연구에서 제안하는 순환 신경망 기반의 클릭 신호 분류 알고리즘의 흐름도는 Fig. 3에 도시하였다. 크게 학습 자료 생성, 학습 네트워크 구축 및 학습 결과, 분류 모델 성능 평가로 구성되었으며, 돌고래 클릭 신호를 분류하는 최적의 네트워크 선정을 위해 각각의 LSTM, GRU, BiLSTM의 학습 환경을 동일하게 설정하였다(Table 1). 또한, 학습을 위해 동일한 하이퍼 파라미터(hyperparameter)를 Table 2와 같이 설정하였으며, 앞서 생성된 총 20,000개의 자료 중 85%에 해당하는 자료를 학

**Table 1.** Parameters of the experimental environment.

Parameters of the experimental environment	
CPU	Intel(R) Core™ i9-10980XE CPU @ 3.00GHz
GPU	NVIDIA GeForce RTX 3090 *2
RAM	96 GB
Operating system	Win 10_64bit
MATLAB version	MATLAB R2022a
Lab environment	Deep Learning Toolbox



**Fig. 2.** (a) Time-series data with click signals and (b) Data after fractional Fourier transform.

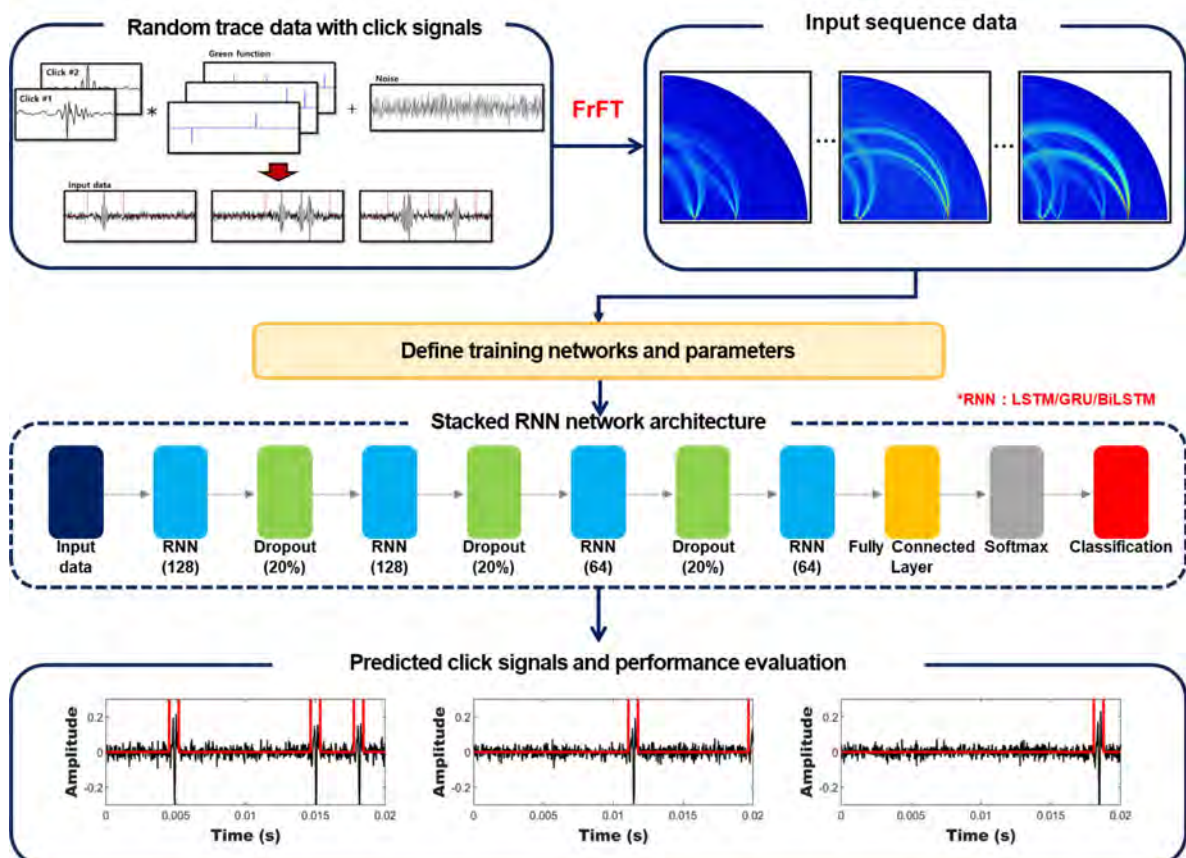
**Table 2.** Hyperparameter and training settings for the proposed network models.

Hyperparameters	Value
Signal length of the input	50.0 ms
Sampling interval	0.025 ms
Input feature dimension	41
Number of epochs	10
Data shuffle	Per every epoch
Mini batch size	128
Optimizer	Adam
Learning rate	1e-4
L2 regularization	0.001
Training data set	Value
Training data count	17,000 (85%)
Validation data count	2,000 (10%)
Test data count	1,000 (5%)
Total data count	20,000 (100%)

습에 사용하고, 10%는 자료를 검증에 사용하고, 5%는 테스트 자료로써 사용하였다.

### 학습 자료 생성

학습에 사용된 자료의 경우, 클릭 신호의 정확한 음향 특성을 파악하기 위해 주변 잡음보다 충분히 높은 신호대잡음비 (signal to noise ratio, SNR)를 갖는 신호 구간을 선정하여 클릭 신호로써 추출하고 무작위로 생성된 그린 함수와의 합성곱으로 생성하였다. 먼저, 입력 자료 생성을 위해 Fig. 1a의 16개 음향 신호 중 총 2개의 신호(1번째, 16번째 신호)를 임의로 선정하였다(Fig. 4a). 이때 Fig. 1b에서 확인된 클릭 신호 길이를 고려하여 각각의 클릭 신호에서 주변 잡음에 비해 우세한 세기를 가지는 꼭지(peak) 신호를 추출한 후 정규화하였다(Figs. 4b, 4c). 현장 자료에서 관측되는 클릭 신호의 세기를 고려하여 0.3에서 0.5의 진폭을 갖는 임의의 정수형 난수를 생성하고, 이를 정규화 된 클릭 신호와 합성곱하여 클릭 신호가 포함된 음향 신호를 생성하였다. 마지막으로, 생성된 자료에 현장 자료 적용성을 높이기 위해 최소 진폭 -0.1에서 최대 진폭 0.1을 갖는 무작위 잡음을 추가하였다(Fig. 5). Fig. 6은 학습에 사용된 입력 자료이며, Fig. 6a의 단일 시계열 자료를 분수 푸리에 변환함으로써 Fig. 6b와 같은 시계열 자료의 회전 각도에 따른 시간-주파수 특성을 확인할 수 있는 결과를 얻을 수 있다. Fig. 6a에서 강조된 구간을 보면, 주변 잡음과 상대적으로 큰 진폭 차이를 갖는 클릭 신호가 생성됨을 알 수 있다. 또한 Fig. 6b의 인공합성자료를 분수 푸리에 변환한 자료에도 Fig.

**Fig. 3.** Flowchart of the proposed network model.

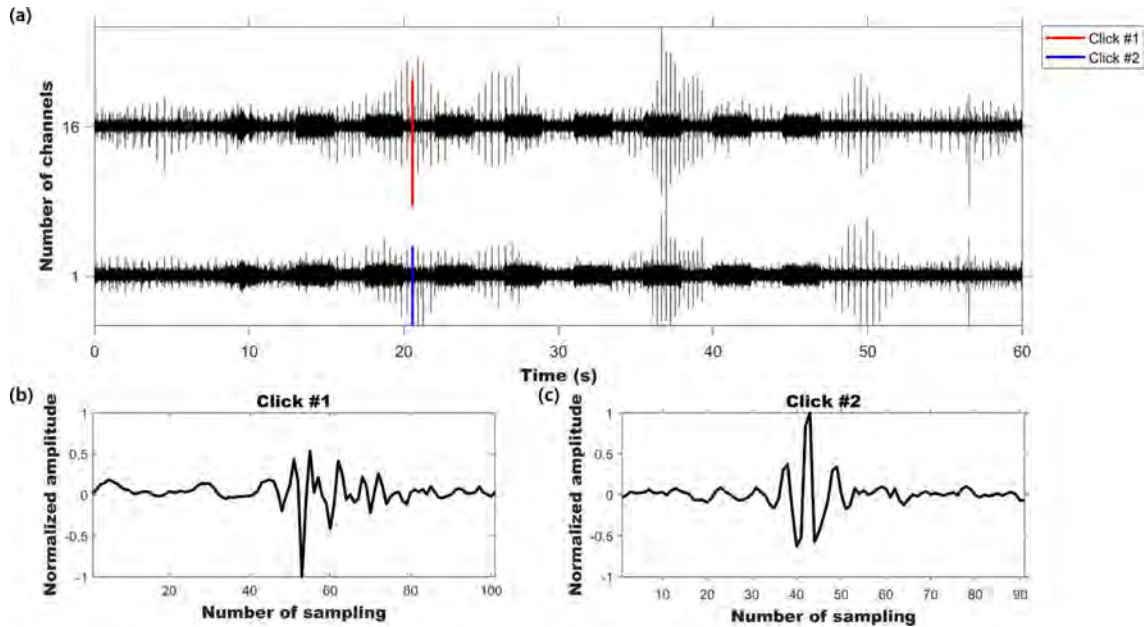


Fig. 4. (a) Time-series data with click signals, (b) click #1, and (c) click #2.

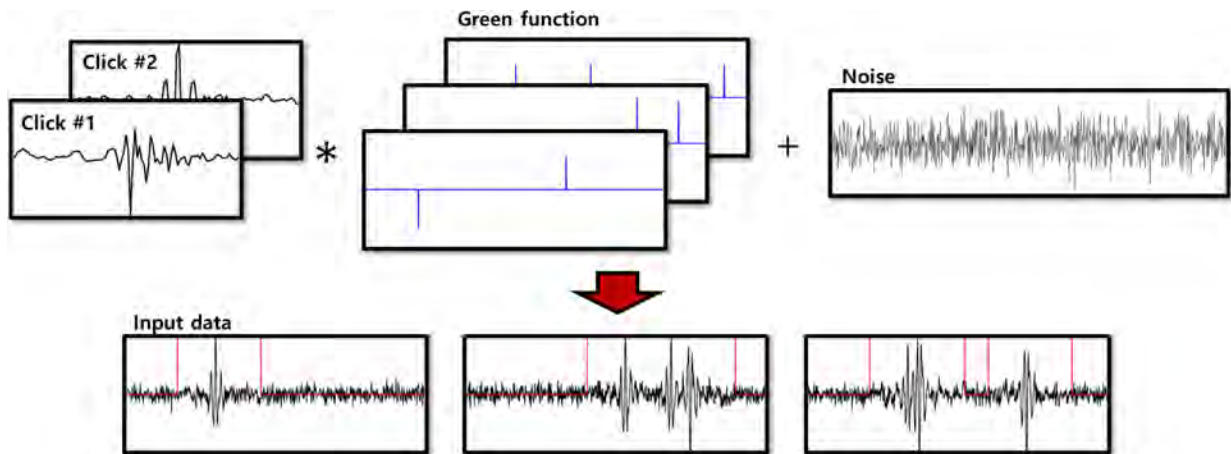


Fig. 5. Schematic representation of synthetic data generation.

2b와 같이 클릭 신호 구간에서 신호의 세기가 강하게 나타나는 패턴이 나타나는 것을 확인하였다. 분수 푸리에 변환의 회전 각도 범위에 따라 기록된 시간-주파수 특성 자료가 나타내는 특성이 상이하기 때문에, 학습을 위한 입력 자료로써 이용되는 회전 각도 범위에 따른 결과 분석이 필요하다. 따라서, 본 연구에서는 네트워크별 최적의 회전 각도 범위를 선정하기 위해, 41개의 동일한 특징 수를 유지하면서 사용되는 각도의 최대 범위를 9°에서부터 90°까지 9° 간격으로, 총 10개의 회전 각도 범위를 입력 자료로써 활용하였다.

**학습 네트워크 구축 및 학습 결과**

본 연구는 순환 신경망을 기반으로 네트워크를 학습하였으며, 네트워크 구조의 경우 반복적으로 순환 신경망을 활용하는

심층 순환 신경망 구조(stacked-RNN)로 설계하였다. 설계된 네트워크는 Fig. 3에서 제시된 순환 신경망 네트워크 구조와 같이 입력층(input layer), 시퀀스층(LSTM, GRU, BiLSTM layer), 완전연결층(fully connected layer), 드롭아웃층(dropout layer), 소프트맥스층(softmax layer), 분류층(classification layer)으로 구성하였다. 해당 네트워크의 입력층에서는 분수 푸리에 변환의 회전 각도 범위에 따라 변환된 분수 푸리에 영역의 자료를 시퀀스 입력 자료로써 이용하였으며, 분류층의 경우 시간에 따른 분류 결과를 출력하도록 네트워크를 구성하였다. 그리고, 각각의 시퀀스층 뒤에 드롭아웃층을 추가하여 과적합을 방지하고자 하였다.

세 가지 네트워크에 대한 심층 순환 신경망 학습 결과, 회전 각도에 따른 검증 정확도와 검증 손실은 Fig. 7과 같다. Figs.

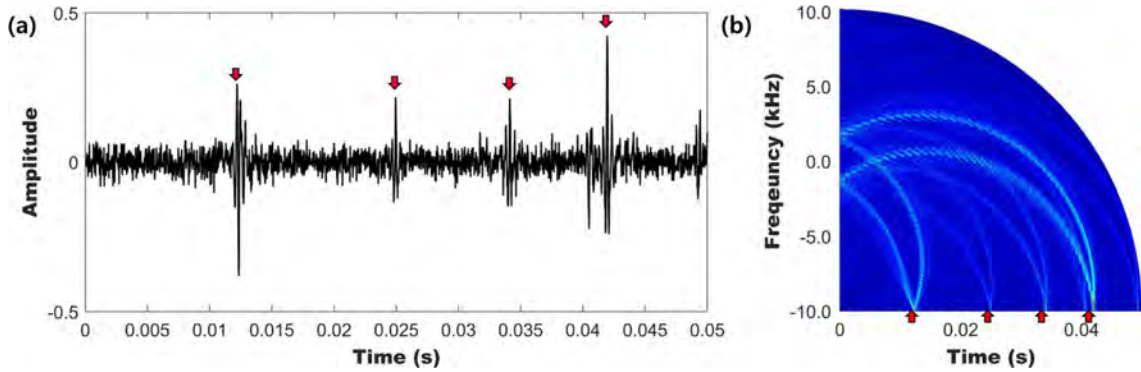


Fig. 6. Input data: (a) Synthetic data with click signals before FrFT, and (b) Data after FrFT.

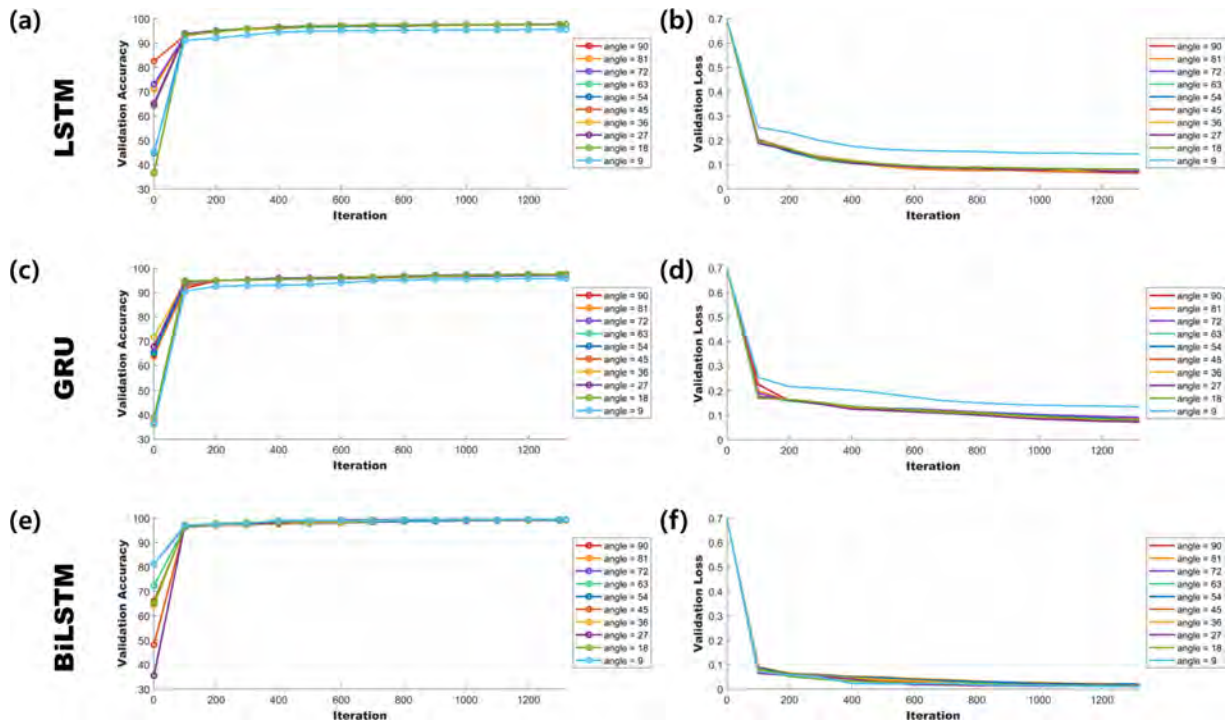


Fig. 7. Validation accuracy (a, c, e) and loss curve (b, d, f): (a, b) LSTM, (c, d) GRU, and (e, f) BiLSTM.

Table 3. Validation loss and validation accuracy by angle for LSTM, GRU, and BiLSTM.

Range of Angle (interval)	Validation Accuracy (%)			Validation Loss		
	LSTM	GRU	BiLSTM	LSTM	GRU	BiLSTM
0°~90° (2.250°)	97.686	97.265	99.395	0.0708	0.0824	0.0155
0°~81° (2.025°)	97.675	97.241	99.180	0.0724	0.0863	0.0204
0°~72° (1.800°)	97.763	97.087	99.486	<b>0.0667</b>	0.0906	0.0140
0°~63° (1.575°)	97.677	97.336	99.496	0.0705	0.0788	0.0129
0°~54° (1.350°)	97.739	97.349	99.301	0.0691	0.0784	0.0174
0°~45° (1.125°)	97.737	97.343	99.403	0.0672	0.0789	0.0147
0°~36° (0.900°)	97.661	<b>97.623</b>	99.510	0.0713	<b>0.0719</b>	0.0124
0°~27° (0.675°)	<b>97.756</b>	97.524	99.586	0.0726	<b>0.0719</b>	0.0110
0°~18° (0.450°)	97.618	97.514	99.555	0.0803	0.0805	0.0119
0°~9° (0.225°)	95.528	95.839	<b>99.623</b>	0.1449	0.1339	<b>0.0102</b>

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

Fig. 8. Confusion matrix used to evaluate performance of the three classification models.

7a, 7c, 7e는 회전 각도 범위에 따른 LSTM, GRU, BiLSTM의 검증 정확도이며, 세 네트워크 모두 반복 횟수가 증가함에 따라 안정적으로 정확도가 향상되는 것을 볼 수 있다. 또한 Figs. 7b, 7d, 7f를 보면, 세 학습 네트워크 모두 회전 각도 범위별로 반복 횟수가 증가함에 따라 검증 손실이 감소하는 추세를 보였다. 학습 후 최종적으로 마지막 반복 횟수에서 얻은 검증 손실 및 검증 정확도는 Table 3와 같다. LSTM과 GRU의 경우, 각각 최대 검증 정확도가 97.756%, 97.623%, 최소 검증 손실이 0.0667, 0.0719로 나타났다. BiLSTM의 경우, 최대 검증 정확도가 99.623%, 최소 검증 손실이 0.0102로 LSTM과 GRU

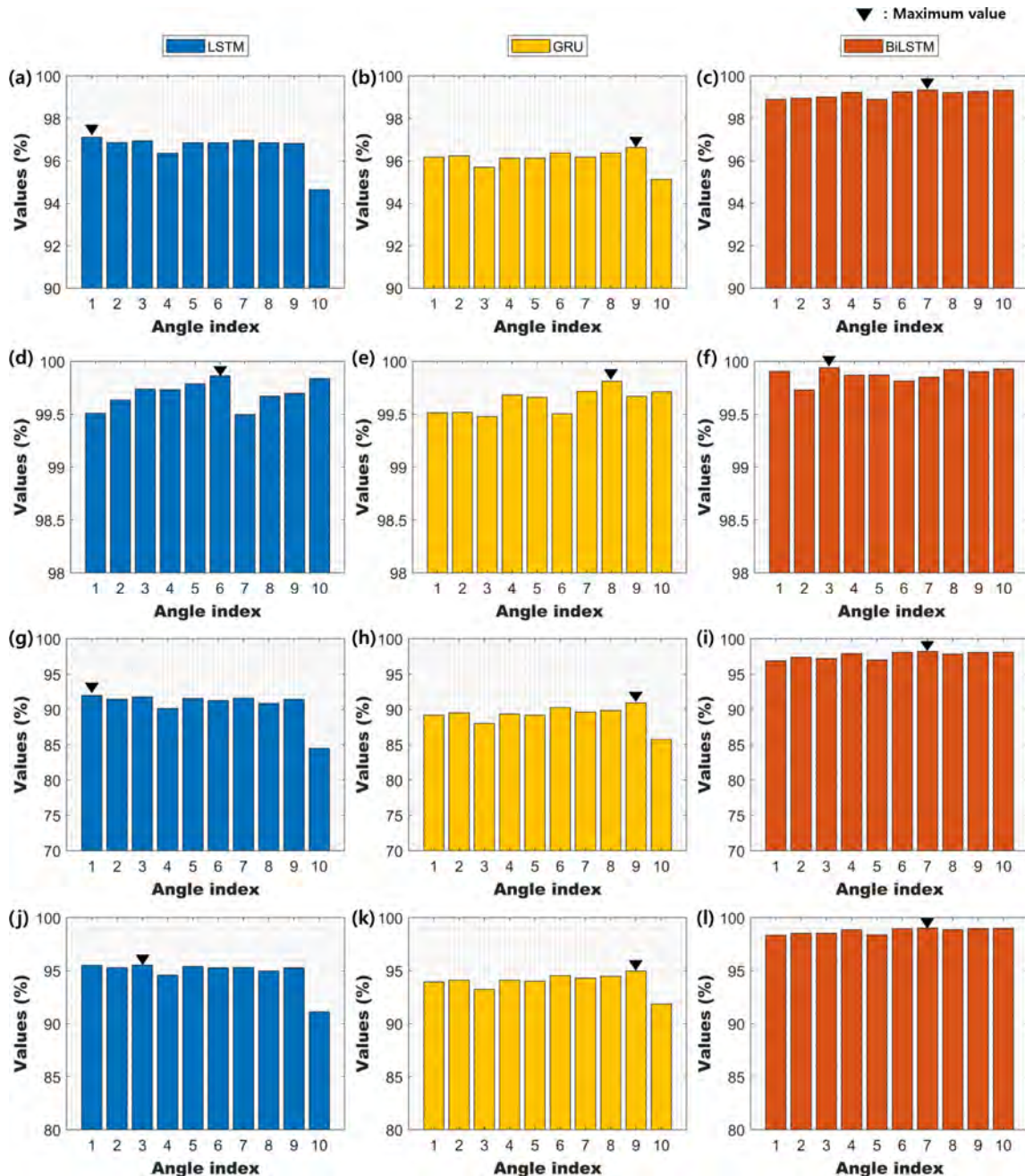


Fig. 9. Performance evaluation of the three classification models: (a-c) Accuracy, (d-f) Precision, (g-i) Recall, and (j-l) F1-score.



보다 높은 성능을 보였다. 즉, 세 가지 네트워크 모두 높은 검증 정확도를 보였으며, 이는 신경망이 최소한의 오류로 클릭 신호의 오류 유형을 성공적으로 식별함을 의미한다.

### 분류 모델 성능 평가

제안된 기법 성능 비교 및 최적의 네트워크 선정을 위해 성능 평가 지표를 사용하여 해당 모델의 정량적 평가를 실시하였다. 분류 모델의 성능 평가 지표는 정확도, 정밀도, 재현율, 그리고 F1-점수이다. 정밀도는 분류 모델을 참으로 분류한 것 중 실제로 참인 비율을, 회수율은 실제 참인 것 중 모델이 참이라고 예측한 비율을, F1-점수는 정밀도와 재현율의 조화평균을, 그리고 정확도는 전체 자료 중 올바르게 분류된 비율을 측정한다(Read *et al.*, 2011; Cui *et al.*, 2022). 분류 성능의 평가는 Fig. 8의 오차 행렬(confusion matrix)을 기반으로 구성되며, 성능 평가 지표는 식 (3) ~ (6)을 통해 계산할 수 있다.

$$Accuracy (\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100, \quad (3)$$

$$Precision (\%) = \frac{TP}{TP + FP} \times 100, \quad (4)$$

$$Recall (\%) = \frac{TP}{TP + FN} \times 100, \quad (5)$$

$$F1 - score (\%) = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100. \quad (6)$$

여기서 TP는 실제 참을 참이라고 예측한 경우, FP는 실제 거짓을 참이라고 예측한 경우, FN은 실제 참을 거짓이라고 예측한 경우, 그리고 TN은 실제 거짓을 거짓이라고 예측한 경우이다. 이때, 성능 평가 지표는 100에 가까울 때 우수한 성능을 가진다고 평가된다.

Fig. 9는 각각의 성능 평가 지표(정확도, 정밀도, 재현율, F1-점수) 항목에서 회전 각도 범위별 LSTM, GRU, BiLSTM 네트워크를 비교하여 그래프로 도시한 결과이다. Figs. 9a-c는 정

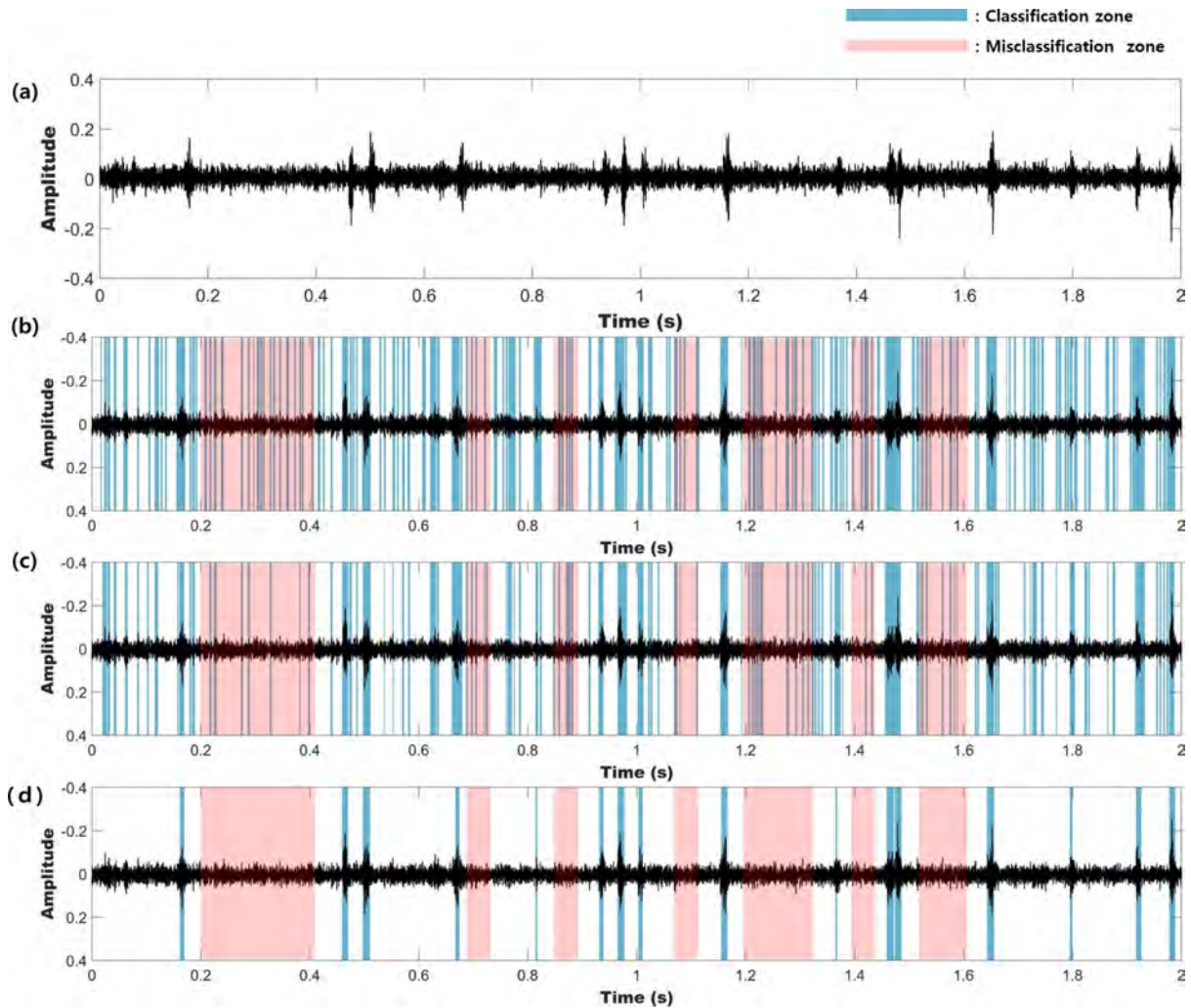
확도, Figs. 9d-f는 정밀도, Figs. 9g-i는 재현율, Figs. 9j-l은 F1-점수를 의미한다. 회전 각도 범위별 성능 평가 결과는 총 1,000개의 자료 중 평균값을 사용하였으며, 탐지확률을 99%로 설정하여 99% 미만이면 0(주변 잡음), 99% 이상이면 1(클릭 신호)로 분류하였다. 이때 그래프의 x축은 1일 때 0° ~ 90°, 10일 때 0° ~ 9°를 의미하며, 9° 간격으로 최대 회전 각도가 감소하도록 회전 각도 범위를 총 10개로 나누었다. 세 네트워크를 비교한 결과, 성능 평가 결과에서 BiLSTM이 모든 항목에서 최대값이 98% 이상의 결과를 보이며 가장 우세하게 나타났다. 수치적으로 비교하기 위해 테스트 자료에 대한 탐지 결과를 바탕으로 산출된 성능 평가 결과를 표로 나타내었다(Table 4). 성능 평가 지표를 종합적으로 고려하여 학습 네트워크별 최적의 회전 각도 범위를 평가한 결과, LSTM의 경우 네 가지 지표 중 정확도, 재현율에서 0° ~ 90°가 가장 높은 결과가 나왔으며, GRU는 정확도, 재현율, F1-점수에서 0° ~ 18°가 가장 높은 결과가 나왔다. 마지막으로 BiLSTM은 0° ~ 36°에서 네 가지 지표 중 정확도, 재현율, F1-점수가 우수한 결과를 보였다. 성능 평가를 통해 얻어진 결과를 통해, 각 순환 신경망에서의 최적의 회전 각도 범위를 선정하였으며, LSTM은 0° ~ 90°, GRU는 0° ~ 18°, BiLSTM은 0° ~ 36°로 지정하였다.

### 현장 자료 적용성 평가

최종적으로 클릭 신호 분류에 가장 적합한 순환 신경망 네트워크 선정을 위해, 현장 자료에 성능 평가에서 선정된 네트워크별 최적의 회전 각도 범위를 적용하여 테스트를 진행하였다. 이때 현장 자료는 Fig. 1a에서 해양에서 취득된 16개의 음향 신호 중 16번째 음향 신호를 임의로 선정하였고, 0 ~ 2 s 구간에서의 클릭 신호를 탐지했다. Figs. 10b-d는 Fig. 10a의 테스트 자료에 훈련된 네트워크를 적용하여 탐지확률 99%로 클릭 신호를 분류한 결과이다. 이때, 파랑색은 클릭 신호 분류 결과, 빨간색은 오탐지 구간으로 각각 색상으로 분석하였다. 분류모델 성능 평가에 의해 도출된 네트워크별 최적의 회전

**Table 4.** Performance metrics (accuracy, precision, recall, F1-score) by angle for LSTM, GRU, and BiLSTM by angle.

Maximum Angle(°)	Accuracy			Precision			Recall			F1-score		
	LSTM	GRU	BiLSTM	LSTM	GRU	BiLSTM	LSTM	GRU	BiLSTM	LSTM	GRU	BiLSTM
90	<b>97.112</b>	96.166	98.885	99.508	99.512	99.905	<b>91.957</b>	89.178	96.858	95.497	93.934	98.350
81	96.861	96.232	98.948	99.636	99.518	99.731	91.445	89.501	97.344	95.285	94.112	98.516
72	96.938	95.697	99.006	99.737	99.478	<b>99.940</b>	91.775	88.008	97.191	<b>95.513</b>	93.226	98.539
63	96.354	96.117	99.222	99.732	99.683	99.869	90.116	89.356	97.890	94.577	94.112	98.855
54	96.851	96.119	98.892	99.788	99.661	99.871	91.539	89.184	96.987	95.406	94.010	98.399
45	96.849	96.373	99.250	<b>99.864</b>	99.505	99.814	91.256	90.252	98.092	95.278	94.532	98.940
36	96.965	96.170	<b>99.339</b>	99.497	99.715	99.851	91.588	89.615	<b>98.200</b>	95.307	94.291	<b>99.014</b>
27	96.853	96.368	99.214	99.671	<b>99.812</b>	99.923	90.844	89.847	97.838	94.970	94.460	98.866
18	96.818	<b>96.627</b>	99.275	99.697	99.670	99.903	91.410	<b>90.897</b>	98.070	95.273	<b>94.970</b>	98.975
9	94.637	95.116	99.326	99.839	99.712	99.929	84.478	85.752	98.104	91.118	91.847	99.005



**Fig. 10.** (a) Field data, and test results of: (b) LSTM test result, (c) GRU test result, and (d) BiLSTM. Blue zones represent section that were accurately classified the proposed algorithms, whereas red zones highlight misclassified sections.

각도 범위에서의 탐지 결과를 살펴보면, Figs. 10b, 10c의 LSTM과 GRU는 클릭 신호를 모두 탐지하나, 주변 잡음을 클릭 신호로 분류하는 오탐지가 다수 나타났다. Fig. 10d의 BiLSTM 테스트 결과, 주변 잡음과 큰 차이를 보이는 클릭 신호뿐만 아니라, 주변 잡음과 유사한 세기를 갖는 클릭 신호도 분류할 수 있었으며, 세 가지 네트워크 중 오탐지가 가장 적게 나타났다.

### 결론

본 연구에서는 해양환경 모니터링을 위한 돌고래 클릭 신호 분류 알고리즘 개발을 목적으로 순환 신경망 기반의 분류 기법을 제안하였다. 클릭 신호와 주변 잡음을 효과적으로 분류하기 위해 회전 각도 범위별 분수 푸리에 변환을 통해 입력 자료를 변환하여 특징을 추출하였으며, 최적의 네트워크 선정을 위해 LSTM, GRU, BiLSTM 총 세 가지 네트워크를 학습하였

다. 인공합성자료를 이용한 학습 결과, 세 네트워크 모두 반복 횟수가 증가함에 따라 검증 손실이 0.0667, 0.0719, 0.0102로 0에 가깝게 감소하였다. 또한, 최대 검증 정확도의 경우, LSTM과 GRU는 각각 97.756%, 97.623%이며, BiLSTM은 99.623%의 최대 정확도를 보였다. 정량적인 평가를 위해 성능 평가 지표를 도입하여 평가한 결과, 정확도, 정밀도, 재현율, F1-점수를 종합적으로 고려했을 때 LSTM는 회전 각도 범위 0° ~ 90°, GRU는 회전 각도 범위 0° ~ 18°에서 높은 평가를 받았으며, BiLSTM은 회전 각도 범위 0° ~ 36°에서 다른 네트워크에 비해 우수한 성능을 보였다. 마지막으로 네트워크별 최적의 회전 각도 범위에서 학습 네트워크를 현장 자료에 적용한 결과, BiLSTM 네트워크를 사용했을 때 주변 잡음과 클릭 신호 간의 차이와 무관하게 오탐지가 가장 적게 나타났다.

본 연구는 동해 해역에서 취득된 음향 신호에서 돌고래 클릭 신호를 추출하여 학습에 이용하였다. 향후 본 연구에서 제안하는 알고리즘에 통해 인근 해역에서 나타나는 돌고래 클릭

신호뿐만 아니라 다른 해역에서 서식하는 해양 포유류의 음향 신호를 학습하여 다양한 특성을 보이는 해양 포유류의 음향 신호의 다중 분류에 적용할 수 있을 것으로 기대된다. 또한, 추후 탄성파 탐사 자료에서 기록되는 클릭 신호를 탐지하여 억제 혹은 제거 연구에 적용 가능할 것으로 사료된다.

## 감사의 글

이 논문은 2023년도 해양수산부 재원으로 해양수산과학기술진흥원의 지원을 받아 수행된 연구임(RS-2023-00259633).

## References

- Almeida, L. B., 1994, The fractional Fourier transform and time-frequency representations, *IEEE Trans. Signal Process.*, **42(11)**, 3084-3091. doi: 10.1109/78.330368
- Au, W. W., and Herzing, D. L., 2003, Echolocation signals of wild Atlantic spotted dolphin (*Stenella frontalis*), *J. Acoust. Soc. Am.*, **113(1)**, 598-604. <https://doi.org/10.1121/1.1518980>
- Caruso, F., Dong, L., Lin, M., Liu, M., Gong, Z., Xu, W., Alonge, G., and Li, S., 2020, Monitoring of a Nearshore Small Dolphin Species Using Passive Acoustic Platforms and Supervised Machine Learning Techniques, *Front. Mar. Sci.*, **7**, 267. <https://doi.org/10.3389/fmars.2020.00267>
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y., 2014, Learning phrase representations using RNN encoder-decoder for statistical machine translation, *arXiv preprint arXiv:1406.1078*. <https://doi.org/10.48550/arXiv.1406.1078>
- Choi, K. H., Yoon, Y. G., Kim, S., Kim, H., Cho, J. W., Bae, H. S., and Park, K., 2019, Analysis of echolocation click signals of Indo-Pacific bottlenose dolphin (*Tursiops aduncus*) in Jeju Island, *J. Acoust. Soc. Korea*, **38(1)**, 56-65 (in Korean with English abstract). <https://doi.org/10.7776/ASK.2019.38.1.056>
- Cui, Q., Yang, B., Liu, B., Li, Y., and Ning, J., 2022, Tea Category Identification Using Wavelet Signal Reconstruction of Hyperspectral Imagery and Machine Learning, *Agriculture*, **12(8)**, 1085. <https://doi.org/10.3390/agriculture12081085>
- Dey, R., and Salem, F. M., 2017, Gate-variants of gated recurrent unit (GRU) neural networks, *In 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 1597-1600. <https://doi.org/10.48550/arXiv.1701.05923>
- Ditria, E. M., Buelow, C. A., Gonzalez-Rivero, M., and Connolly, R. M., 2022, Artificial intelligence and automated monitoring for assisting conservation of marine ecosystems: A perspective, *Front. Mar. Sci.*, **9**, 918104. <https://doi.org/10.3389/fmars.2022.918104>
- Frasier, K. E., Roch, M. A., Soldevilla, M. S., Wiggins, S. M., Garrison, L. P., and Hildebrand, J. A., 2017, Automated classification of dolphin echolocation click types from the Gulf of Mexico, *PLoS Comput. Biol.*, **13(12)**, e1005823. <https://doi.org/10.1371/journal.pcbi.1005823>
- Graves, A., Mohamed, A. R., and Hinton, G., 2013, Speech recognition with deep recurrent neural networks, *In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 6645-6649. doi: 10.1109/ICASSP.2013.6638947
- Hochreiter, S., and Schmidhuber, J., 1997, Long short-term memory, *Neural Comput.*, **9(8)**, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Jiang, J. J., Bu, L. R., Duan, F. J., Wang, X. Q., Liu, W., Sun, Z. B., and Li, C. Y., 2019, Whistle detection and classification for whales based on convolutional neural networks, *Applied Acoustics*, **150**, 169-178. <https://doi.org/10.1016/j.apacoust.2019.02.007>
- Jin, C., Kim, M., Jang, S., and Paeng, D. G., 2022, Semantic segmentation-based whistle extraction of indo-pacific bottlenose dolphin residing at the coast of jeju island, *Ecol. Indic.*, **137**, 108792. doi: 10.1016/j.ecolind.2022.108792
- Jozefowicz, R., Zaremba, W., and Sutskever, I., 2015, An empirical exploration of recurrent network architectures, *In International conference on machine learning*, PMLR, 2342-2350. <https://proceedings.mlr.press/v37/jozefowicz15.html>
- Kavanagh, A. S., Nykänen, M., Hunt, W., Richardson, N., and Jessopp, M. J., 2019, Seismic surveys reduce cetacean sightings across a large marine ecosystem, *Scientific Reports*, **9(1)**, 19164. <https://doi.org/10.1038/s41598-019-55500-4>
- Kim, D., Kim, J. S., and Song, J., 2022a, Cancellation of dolphin sonar clicks in a communication signal based on adaptive time reversal processing, *JASA Express Lett.*, **2(5)**, 056001. <https://doi.org/10.1121/10.0010375>
- Kim, J. H., and Kim, J. Y., 2022, Comparative analysis of performance of BI-LSTM and GRU algorithm for predicting the number of Covid-19 confirmed cases, *Journal of the Korea Institute of Information and Communication Engineering*, **26(2)**, 187-192 (in Korean with English abstract). <https://doi.org/10.6109/jkiice.2022.26.2.187>
- Kim, J. S., Yoon, Y. G., Han, D. G., La, H. S., and Choi, J. W., 2022b, Classification of bearded seals signal based on convolutional neural network, *J. Acoust. Soc. Korea*, **41(2)**, 235-241 (in Korean with English abstract). <https://doi.org/10.7776/ASK.2022.41.2.235>
- Namias, V., 1980, The fractional order Fourier transform and its application to quantum mechanics, *IMA J. Appl. Math.*, **25(3)**, 241-265. <https://doi.org/10.1093/imamat/25.3.241>
- Ozaktas, H. M., and Kutay, M. A., 2001, The Fractional Fourier Transform, *In 2001 European Control Conference (ECC)*. IEEE, 1477-1483. doi: 10.23919/ECC.2001.7076127
- Radford, C. A., Jeffs, A. G., Tindle, C. T., and Montgomery, J. C., 2008, Temporal patterns in ambient noise of biological origin from a shallow water temperate reef, *Oecologia*, **156(4)**, 921-929. doi: 10.1007/s00442-008-1041-y
- Read, J., Pfahringer, B., Holmes, G., and Frank, E., 2011, Classifier chains for multi-label classification, *Mach. Learn.*, **85**, 333-359. <https://doi.org/10.1007/s10994-011-5256-5>
- Saxena, R., and Singh, K., 2005, Fractional Fourier transform: A

- novel tool for signal processing, *J. Indian Inst. Sci.*, **85(1)**, 11-26. [https://www.researchgate.net/publication/228341636\\_Fractional\\_Fourier\\_transform\\_A\\_novel\\_tool\\_for\\_signal\\_processing](https://www.researchgate.net/publication/228341636_Fractional_Fourier_transform_A_novel_tool_for_signal_processing)
- Schmidhuber, J., 2015, Deep learning in neural networks: An overview, *Neural Networks*, **61**, 85-117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- Schuster, M., and Paliwal, K. K., 1997, Bidirectional recurrent neural networks, *IEEE Trans. Signal Process*, **45(11)**, 2673-2681. <https://ieeexplore.ieee.org/document/650093>
- Sugimatsu, H., Kojima, J., Ura, T., Bahl, R., and Tomuro, S., 2015, Development of an automatic discrimination method of the bio-sonar clicks of Irrawaddy dolphin (*Orcaella brevirostris*) from various types of boat noises, *OCEANS 2015 – MTS/IEEE Washington*, 1-9. doi: 10.23919/OCEANS.2015.7404489
- Thomas, M., Martin, B., Kowarski, K., Gaudet, B., and Matwin, S., 2019, An end-to-end approach for true detection of low frequency marine mammal vocalizations, *J. Acoust. Soc. Am.*, **146(4)**, 2959. <https://doi.org/10.1121/1.5137278>
- Tian, L., 2021, Seismic spectral decomposition using short-time fractional Fourier transform spectrograms, *J. Appl. Geophys.*, **192**, 104400. <https://doi.org/10.1016/j.jappgeo.2021.104400>
- Wenz, G. M., 1962, Acoustic ambient noise in the ocean: Spectra and sources, *J. Acoust. Soc. Am.*, **34(12)**, 1936-1956. <https://doi.org/10.1121/1.1909155>
- Zhai, M. Y., 2014, Seismic data denoising based on the fractional Fourier transformation, *J. Appl. Geophys.*, **109**, 62-70. <https://doi.org/10.1016/j.jappgeo.2014.07.012>