

어텐션 기반 게이트 순환 유닛을 이용한 수동소나 신호분류

Passive sonar signal classification using attention based gated recurrent unit

이기배,¹ 고건혁,¹ 이종현[†]

(Kibae Lee,¹ Guhn Hyeok Ko,¹ and Chong Hyun Lee^{1†})

¹제주대학교 해양시스템공학과

(Received May 3, 2023; revised June 14, 2023; accepted June 23, 2023)

초 록: 수동소나의 표적신호는 수초 내 세기의 변화를 갖는 협대역 고조파 특성과 로이드 거울 효과에 의한 장시간 주파수 변이 특성을 나타낸다. 본 논문에서는 지역 및 전역적 시계열 특징을 학습하는 게이트 순환 유닛 기반의 신호분류 알고리즘을 제안한다. 제안하는 알고리즘은 게이트 순환 유닛을 이용한 다층 네트워크를 구성하고 확장된 연결을 통해 지역 및 전역적 시계열 특징들을 추출한다. 이후 어텐션 메커니즘을 학습하여 시계열 특징들을 가중하고 수동소나 신호를 분류한다. 공개된 수중 음향 데이터를 이용한 실험에서 제안된 네트워크는 96.50%의 우수한 분류 정확도를 보였다. 이러한 결과는 기존의 잔차 연결된 게이트 순환 유닛 네트워크와 비교하여 4.17% 높은 분류 정확도를 갖는다.
핵심용어: 지역 및 전역적 시계열 특징, 어텐션 메커니즘, 확장된 연결, 게이트 순환 유닛

ABSTRACT: Target signal of passive sonar shows narrow band harmonic characteristic with a variation in intensity within a few seconds and long term frequency variation due to the Lloyd's mirror effect. We propose a signal classification algorithm based on Gated Recurrent Unit (GRU) that learns local and global time series features. The algorithm proposed implements a multi layer network using GRU and extracts local and global time series features via dilated connections. We learn attention mechanism to weight time series features and classify passive sonar signals. In experiments using public underwater acoustic data, the proposed network showed superior classification accuracy of 96.50%. This result is 4.17% higher classification accuracy compared to existing skip connected GRU network.

Keywords: Local and global time series features, Attention mechanism, Dilated connection, Gated Recurrent Unit (GRU)

PACS numbers: 43.30.Wi, 43.60.Bf, 43.60.Lq

1. 서 론

항해중인 선박 및 항만의 안전을 위해 수동소나 기반의 감시체계가 운용되고 있다. 수중 상황 인지를 위해 수동소나의 주 표적인 선박 신호의 분류가 우선적으로 요구된다. 선박 신호는 기계 장치에 의한 기계소음과 프로펠러에 의한 공동소음을 동반한다.^[1,2] 기존 연구에서는 기계소음과 공동소음 분석을 위해 LOw Frequency Analysis and Recording(LOFAR)와

DEMOdulation of Noise(DEMON) 처리를 수행하고 협대역의 주파수선을 추적하여 선박신호를 식별한다.^[3,4] 최근에는 영상처리 분야에서 주로 사용되는 심층 학습 알고리즘을 사용하여 괄목할만한 성과를 도출하였다.^[5-11] 심층 학습을 통한 수동소나 신호의 분류에는 단시간 푸리에 변환으로 생성된 시간-주파수 영역의 특징이 주로 사용되며, Mel Frequency Cepstral Coefficients(MFCC), Mel Filter Bank(MFB) 에너지 등의 특징이 사용되기도 한다. 수동소나 신호에서 추

[†]Corresponding author: Chong Hyun Lee (chonglee@jejunu.ac.kr)

Department of Ocean System Engineering, Jeju National University, 102, Jeju Daehakro, Jeju, Jeju-do 63243, Republic of Korea
(Tel: 82-64-754-3481, Fax: 82-64-754-3480)



Copyright©2023 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.

출된 시간-주파수 영역 특징은 영상화되어 합성곱 신경망(Convolutional Neural Network, CNN)을 기반으로 분류된다. 합성곱 신경망은 영상화된 특징에서 시간-주파수 영역의 상관관계를 학습하지만, 순차적으로 입력되는 시계열 데이터의 특성을 고려하지 않는다.

순환 신경망(Recurrent Neural Network, RNN)은 시계열 데이터의 특징을 학습할 수 있는 대표적인 심층 학습 모델이다.^[12] 특히 게이트 순환 유닛(Gated Recurrent Unit, GRU)은 기존 순환 신경망 모델들의 성능을 유지하면서 구조를 단순화하여 주목받고 있다.^[13] 이러한 이유로 게이트 순환 유닛을 이용한 네트워크가 시계열 데이터 처리를 위해 주로 사용된다.^[14,15] 일반적인 게이트 순환 유닛 네트워크는 인접한 시퀀스 데이터 간 상관관계를 학습하므로 전역적 시계열 특징 추출에 한계를 갖는다. 이러한 문제를 개선하고자 잔차 연결된 게이트 순환 유닛 네트워크(skip connected GRU network)가 제안되었다.^[16] 잔차 연결된 게이트 순환 유닛 네트워크는 게이트 순환 유닛 간의 잔차 연결을 추가하여 지역 및 전역적 시계열 특징을 효율적으로 학습한다. 하지만 지역 및 전역적 시계열 특징 정보를 목적에 따라 가중할 수 없어서 데이터 특성에 따른 복잡한 설계가 요구된다.

본 논문에서는 지역 및 전역적 시계열 특징을 학습하는 게이트 순환 유닛 기반의 신호분류 알고리즘을 제안한다. 제안하는 알고리즘은 게이트 순환 유닛을 이용한 다층 네트워크를 구성하고 점진적으로 확장된 연결(dilated connection)을 통해 지역 및 전역적 시계열 특징을 추출한다. 이후 어텐션 메커니즘(attention mechanism)을 학습하여 시계열 특징들을 가중하고 신호를 분류한다. 수동소나 신호 분류에서는 단시간 푸리에 변환으로 시간-주파수 영역 특징을 생성하고 시간에 따른 특징벡터를 순차적으로 입력한다. 본 논문에서는 공개된 수중 음향 데이터를 이용한 실험을 통해 제안된 분류 알고리즘의 성능을 기존 방법들과 비교하여 평가한다.

II. 수동소나 표적신호

수동소나의 주 표적인 선박신호는 엔진 및 기타

장비에 의한 기계소음, 프로펠러의 공동현상에 의한 공동소음, 유체와의 마찰에 의한 유체소음을 포함한다.^[1,2] 선박신호의 주된 특성인 기계소음은 Eq. (1)과 같이 혼재된 고조파 협대역 신호로 정의된다.

$$s(t) = \sum_{i=1}^N \sum_{j=1}^M a_{ij}(t) \sin(j2\pi f_i t + \phi_{ij}), \quad (1)$$

여기서 N 은 기계소음원의 수, M 은 기계소음원의 고조파 개수를 의미한다. 또한 $a_{ij}(t)$ 는 시간에 따른 신호 세기, ϕ_{ij} 는 신호의 초기 위상을 나타내며, f_i 는 i 번째 신호의 기본 주파수를 의미한다. 기계소음 중에서 선박 엔진음의 f_i 는 엔진의 분당 회전수 r 에 따라 $0.16r$ 으로 결정된다. 이러한 선박의 기계소음은 일정한 협대역 주파수에서 수초 내 세기가 변동되는 특성을 갖는다.

해수면에 의해 생성된 가상음원의 간섭에 의한 로이드 거울 효과(Lloyd's mirror effect)는 시간-주파수 영역에서 U자 패턴을 형성화한다.^[17,18] 로이드 거울 효과로 인한 간섭 패턴은 Eq. (2)와 같이 발생한다.

$$p(t) = 4\sin^2\left(\frac{kh_{sou}h_{hyd}}{d(t)}\right), \quad (2)$$

여기서 k 는 음원의 파수, h_{sou} 는 음원의 수심, h_{hyd} 는 하이드로폰의 수심, $d(t)$ 는 시간에 따른 음원과 하이드로폰 간의 수평거리를 의미한다. 이러한 로이드 거울 효과는 수심 초 간 주파수 변이를 유발한다.

III. 신호분류 알고리즘

3.1 게이트 순환 유닛

게이트 순환 유닛은 순환 신경망의 일종으로 장단기 메모리(Long Short Term Memory, LSTM)를 경량화한 모델이다.^[13] 본 논문에서는 시간에 따른 특징벡터의 처리에 기반하여 게이트 순환 유닛을 소개한다. 게이트 순환 유닛은 Fig. 1과 같이 초기화 게이트, 갱신 게이트 후보군 계산, 은닉층 계산 과정을 수행한다. 초기화 게이트는 시그모이드 함수 $\sigma(\cdot)$ 를 이용하여 과거 정보를 초기화하며, τ 시간 이전 시퀀

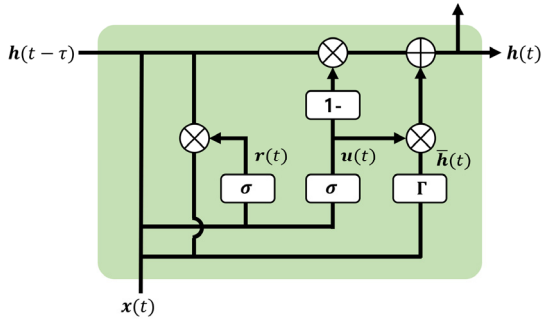


Fig. 1. (Color available online) Architecture of GRU.

스의 은닉층 출력 $h(t-\tau)$ 과 현재 시퀀스의 정보 $x(t)$ 에 가중치 W_r 과 U_r 을 적용하여 Eq. (3)과 같이 수행된다.

$$r(t) = \sigma(W_r h(t-\tau) + U_r x(t)). \quad (3)$$

일반적으로 τ 는 현재 시점에서 직전 시퀀스의 시간차로 정의된다. 갱신 게이트는 직전 시퀀스와 현재 시퀀스의 가중 정도 $u(t)$ 를 Eq. (4)와 같이 결정한다.

$$u(t) = \sigma(W_u h(t-\tau) + U_u x(t)), \quad (4)$$

여기서 W_u 와 U_u 는 $h(t-\tau)$ 과 $x(t)$ 에 대한 갱신 게이트 내 가중치를 의미한다. 후보군 계산은 $r(t)$ 를 $h(t-\tau)$ 에 곱하고 가중치 W , U 를 적용하여 Eq. (5)와 같이 수행된다.

$$\bar{h}(t) = \Gamma(W h(t-\tau) \odot r(t) + U x(t)), \quad (5)$$

여기서 $\Gamma(\cdot)$ 는 탄젠트 하이퍼볼릭 함수, \odot 요소 별 곱셈 연산을 의미한다. 끝으로 $u(t)$ 와 $\bar{h}(t)$ 를 결합하여 현재 시퀀스의 은닉층 출력 $h(t)$ 을 Eq. (6)과 같이 계산한다.

$$h(t) = (1-u(t)) \odot h(t-\tau) + u(t) \odot \bar{h}(t). \quad (6)$$

3.2 게이트 순환 유닛을 이용한 신호분류

본 절에서는 게이트 순환 유닛을 이용한 수동소나의 신호분류 방법들을 소개하고, 기존 방법들의 한계를 개선하고자 어텐션 기반의 신호분류 알고리즘

을 제안한다. Fig. 2는 게이트 순환 유닛을 이용한 신호 분류 방법들을 도시화한다. 게이트 순환 유닛은 시계열 특징의 수만큼 정렬되어 순차적으로 특징을 입력받는다.^[13] 이러한 게이트 순환 유닛 네트워크를 Fig. 2(a)와 같이 적층하여 심층 학습을 수행하고, 최종 층에서 종료 시점의 은닉층 출력을 사용하여 신호를 분류할 수 있다. 적층된 게이트 순환 유닛 네트워크(stacked GRU network)는 인접 시간의 특징 간 상관관계를 학습하므로 전역적 시계열 특징 추출에 한계를 갖는다. 이러한 한계를 극복하고자 Fig. 2(b)와 같이 잔차 연결된 게이트 순환 유닛 네트워크가 사용될 수 있다. 잔차 연결된 게이트 순환 유닛 네트워크는 게이트 순환 유닛 간의 잔차 연결을 추가하여 지역 및 전역적 시계열 특징을 효율적으로 학습한다.^[16] 현재 시점의 게이트 순환 유닛은 $h(t-\tau)$ 과 τ_s 이전 시점의 은닉층 출력 $h(t-\tau_s)$ 을 Eq. (7)과 같이 병합하여 현재 시점의 은닉 상태 $h_s(t)$ 로 입력받는다.

$$h_s(t) = h(t-\tau) + h(t-\tau_s), \tau < \tau_s. \quad (7)$$

따라서 Eqs. (3) ~ (6)의 $h(t-\tau)$ 을 $h_s(t)$ 로 변경하여 특징벡터를 처리한다. 잔차 연결된 게이트 순환 유닛 네트워크는 지역 및 전역적 시계열 특징 정보를 목적에 따라 가중할 수 없어서 데이터 특성에 따른 복잡한 설계가 요구된다.

기존의 게이트 순환 유닛을 이용한 분류 방법들의 한계를 개선하고자 Fig. 2(c)와 같이 어텐션 기반 확장된 게이트 순환 유닛 네트워크(attention dilated GRU network)를 제안한다. 첫 번째 층에서는 시계열 입력 특징의 수와 동일하게 게이트 순환 유닛을 배치하고 순차적으로 특징을 입력받아 처리한다. 이후 층이 적층됨에 따라 일정 시간 간격을 두고 게이트 순환 유닛을 배치하여 점진적으로 확장된 연결을 구현한다. 제안하는 네트워크의 게이트 순환 유닛은 Eqs. (8) ~ (11)과 같이 동작한다.

$$r(t) = \sigma(W_r h(t-2^{(l-1)}\tau) + U_r x(t)). \quad (8)$$

$$u(t) = \sigma(W_u h(t-2^{(l-1)}\tau) + U_u x(t)). \quad (9)$$

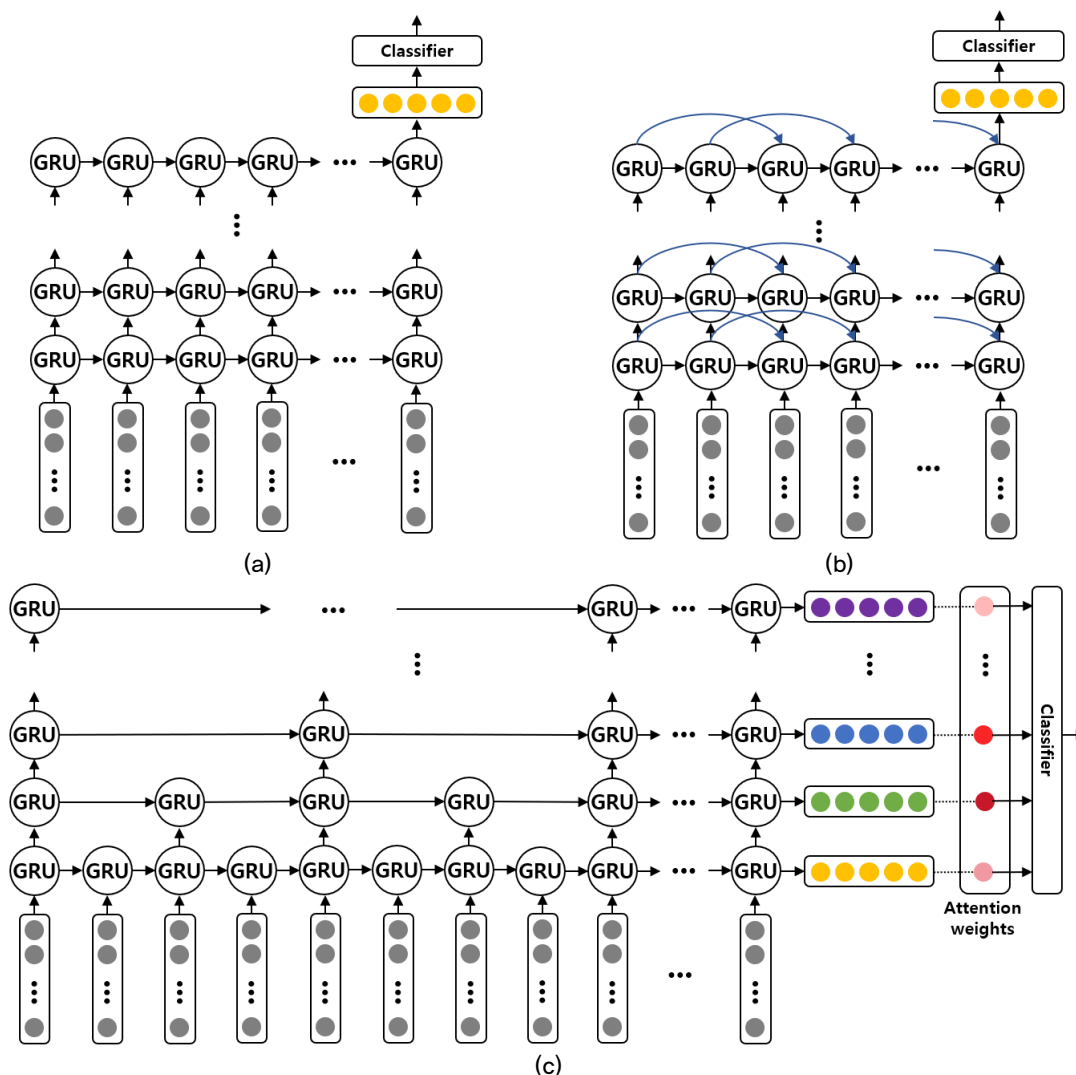


Fig. 2. (Color available online) Classification models based on (a) stacked GRU network, (b) skip connected GRU network and (c) attention dilated GRU network.

$$\bar{\mathbf{h}}(t) = I(W\mathbf{h}(t-2^{(l-1)}\tau) \odot r(t) + U\mathbf{x}(t)). \quad (10)$$

$$\mathbf{h}(t) = (1 - \mathbf{u}(t)) \odot \mathbf{h}(t-2^{(l-1)}\tau) + \mathbf{u}(t) \odot \bar{\mathbf{h}}(t), \quad (11)$$

여기서 $l \in \{1, 2, \dots, L\}$ 은 L 개 층으로 적층된 네트워크의 층 순번을 의미하며, τ 는 Eq. (12)와 같이 l 에 따라 결정된다.

$$\tau = \begin{cases} \tau_0, & l = 1 \\ \tau_d, & l > 1 \end{cases} \quad (12)$$

여기서 τ_0 는 현재 시점과 직전 시퀀스 간 시간 차이

며, τ_d 는 두 번째 층부터 확장 연결을 위한 기준 시간 간격으로 $\tau_d > \tau_0$ 의 조건을 만족한다. 따라서 제안하는 네트워크는 다양한 스케일의 시계열 특징을 추출하고자, 두 번째 층부터 층수에 비례하여 가장 작은 정수배인 2배수로 확장된 연결을 갖는다. 제안하는 네트워크는 각 층에서 종료 시점 유닛의 은닉층 출력 $\mathbf{h}_{end}^{(l)}$ 을 특징벡터로 추출한다. 이후 어텐션 메커니즘에 기반하여 학습된 가중치 $w^{(l)}$ 을 Eq. (13)과 같이 $\mathbf{h}_{end}^{(l)}$ 에 적용하고 통합된 특징벡터 \mathbf{u} 를 획득한다.

$$\mathbf{u} = \sum_{l=1}^L w^{(l)} \mathbf{h}_{end}^{(l)}. \quad (13)$$

이후 분류기는 \mathbf{u} 를 입력받아 신호의 분류 결과를 출력한다. 제안하는 네트워크는 Eq. (14)와 같이 분류기의 예측 결과 \hat{y} 와 실제 값 y 를 이용한 교차 엔트로피 손실을 통해 종단간 학습(end to end learning)을 수행한다.

$$\mathcal{L}(y, \hat{y}) = -y \log(\hat{y}). \quad (14)$$

제안하는 네트워크는 서로 다른 스케일의 시계열 특징들에 어텐션 기반 가중치를 적용하여 분류를 위한 효율적인 학습을 수행한다.

IV. 실험 및 결과

4.1 데이터 구성

공개된 수중 음향 데이터인 ShipEar database^[19]의 다양한 선종에 따른 음향 신호를 이용하여 제안된 알고리즘의 성능을 평가한다. 본 논문에서는 ShipEar database의 11종 선박신호 중에서 학습 데이터를 충분히 확보할 수 있는 어선, 모터보트 등 6종의 선박신호를 사용한다. Fig. 3은 실험에 사용되는 선박 신호들의 대표적인 스펙트로그램을 보여준다. 앞서 2장에서 기술한 바와 같이 선박신호는 기계적 동작에 의한 고조파 협대역 신호로 나타나며, 로이드 거울 효과에 의한 주파수 변이를 수반한다. 고조파 협대역 신호는 시간에 따라 일정한 고조파 주파수에서 수초 내 세기의 변화를 갖는다. 또한 로이드 거울 효과에 의한 주파수 변이는 수십 초 동안 지속된다. Fig. 4는 모터보트 신호에서 관찰되는 고조파 협대역 성분과 로이드 거울 효과에 의한 주파수 변이 현상을 보여준다.

선박신호 분류 모델의 학습 및 평가에 사용되는 데이터는 90% 오버랩하여 30s 길이로 생성되었다. 각 선종 별 학습 데이터는 무작위로 80개가 선택되고, 평가 데이터는 잔여 데이터 중 무작위로 20개가 선택된다. Table 1은 학습 및 평가에 사용되는 선박신호의 데이터 구성을 보여준다. 분류 모델의 학습에는 30s 길이의 6종 선박신호가 총 480개, 모델의 평가에는 120개가 사용된다. 실험에 사용되는 30s 길이 데이터는 8789 Hz 표본화율을 가지며, 단시간 푸리에 변환

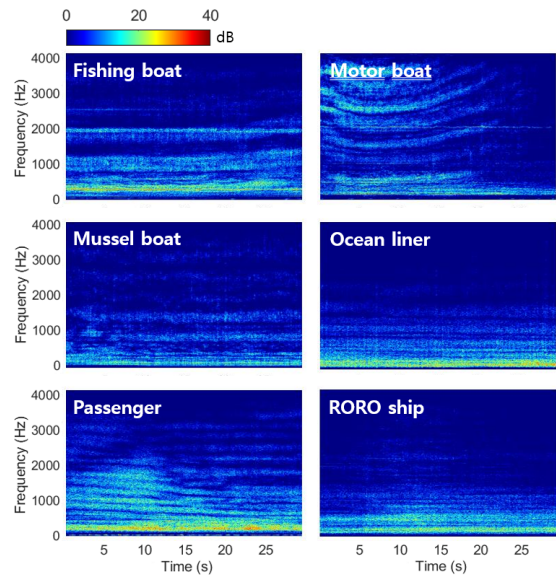


Fig. 3. (Color available online) Spectrograms according to ship type.

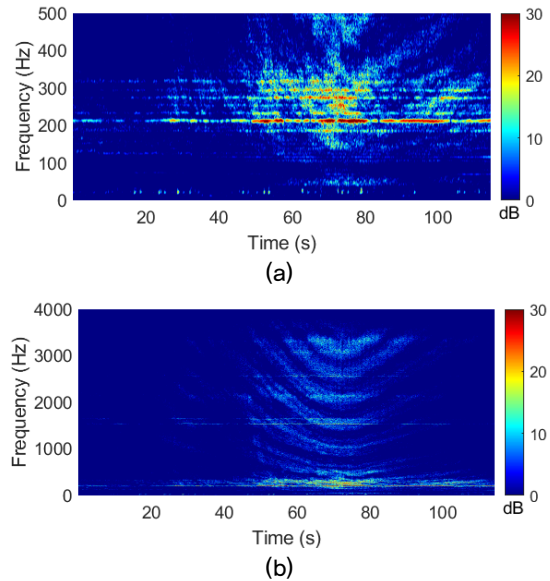


Fig. 4. (Color available online) (a) Narrowband harmonics and (b) frequency shift characteristics of motor boat signal.

을 통해 스펙트로그램으로 변환된다. 스펙트로그램은 0.5s 단위 시간 윈도우에 88.5% 오버랩이 적용된 총 512번의 단시간 푸리에 변환에 의해 생성된다. 변환 과정에서 단시간 푸리에 변환을 통해 표현되는 주파수 간격은 8.5 Hz로 설정된다. 이와 같이 생성된 512 × 512 크기의 스펙트로그램은 최소-최대 스케일링 정규화 과정을 거쳐 분류 모델에 입력된다. Table 2는

Table 1. Underwater acoustic dataset for experiments.

| Data type | Train data | Test data |
|----------------------------|------------|------------|
| Fishing boat | 80 | 20 |
| Motor boat | 80 | 20 |
| Mussel boat | 80 | 20 |
| Ocean liner | 80 | 20 |
| Passenger | 80 | 20 |
| Roll-On/Roll-Off RORO ship | 80 | 20 |
| Total | 480 | 120 |

Table 2. Parameters and settings for data processing.

| Audio data | |
|----------------------|-----------------|
| Sampling rate | 8789 Hz |
| Audio length | 30 s |
| Spectrogram | |
| Window length | 0.5 s |
| Overlap | 88.5 % |
| Frequency interval | 8.5 Hz |
| Normalization method | Min-max scaling |
| Input size | 512 × 512 |

데이터 처리를 위한 설정들을 보여준다.

4.2 분류 모델 구현

본 절에서는 선박신호 분류를 위해 합성곱 신경망, 적층된 게이트 순환 유닛 네트워크, 잔차 연결된 게이트 순환 유닛 네트워크, 어텐션 기반 확장된 게이트 순환 유닛 네트워크를 구현하고 성능을 비교한다. 합성곱 신경망은 64, 32, 16, 16, 16개의 합성곱 필터, 3, 3, 2, 2, 2의 커널 크기로 구성되는 5층의 은닉층을 보유한다. 5번째 층을 제외한 은닉 층에서는 (3, 3), (3, 3), (2, 2), (2, 2)의 최대값 추출을 수행한다. 각각의 은닉 층은 합성곱 연산 이후에 배치 정규화와 Rectified Linear Unit(ReLU) 활성화 함수에 대한 수행을 포함한다. 5번째 은닉 층에서는 과적합 방지를 위해 0.3의 드롭아웃을 수행한다. 합성곱 층에 의해 추출된 특징은 2304개의 입력 노드와 6개의 출력 노드를 갖는 완전 연결층에 입력되고 소프트맥스(Softmax) 연산을 통해 분류 결과를 출력한다. Table A1은 구현된 합성곱 신경망의 구조를 보여준다.

적층된 게이트 순환 유닛 네트워크는 5층의 게이트 순환 유닛 네트워크와 분류를 위한 완전 연결층

Table 3. Attention dilated GRU network architecture.

| |
|---|
| Input: 512 × 512 × 1 Shape |
| Layer 1: 512 × 256 GRU sequence output shape, 1 × 256 GRU output shape |
| Layer 2: 52 × 256 GRU sequence output shape, 1 × 256 GRU output shape 10 skip connection interval |
| Layer 3: 26 × 256 GRU sequence output shape, 1 × 256 GRU output shape 20 skip connection interval |
| Layer 4: 13 × 256 GRU sequence output shape, 1 × 256 GRU output shape 40 skip connection interval |
| Layer 5: 6 × 256 GRU sequence output shape, 1 × 256 GRU output shape 80 skip connection interval |
| Layer 6: 5 Attention weights, Fully connected layer (256 input, 6 output), Softmax |

으로 구성된다. 각 층의 네트워크는 512개의 게이트 순환 유닛으로 구성되며, 각 유닛은 512, 256, 128, 64 길이의 벡터를 출력한다. 5번째 층의 종료 시점 유닛의 벡터를 64개 입력 노드와 6개 출력 노드를 갖는 완전 연결층에 입력하고 소프트맥스 연산을 통해 분류 결과를 출력한다. Table A2는 적층된 게이트 순환 유닛 네트워크 구조를 보여준다. Table A2의 네트워크는 현 시점에서 바로 이전 시점의 은닉층 출력을 입력받는다. 따라서 Table 2와 같이 생성된 스펙트로그램을 시퀀스로 입력받는 경우 τ 가 0.05 s인 시계열 특징 정보를 학습한다. 이와 유사하게 본 논문에서는 τ 가 4 s인 특징을 입력받아 처리할 수 있는 네트워크를 Table A3와 같이 구현한다.

본 논문에서는 5층의 잔차 연결된 게이트 순환 유닛 네트워크를 구현한다. 잔차 연결된 게이트 순환 유닛 네트워크는 τ 는 0.05 s인 직전 시퀀스 연결을 가지며, 2에서 5번째 층에서 τ_s 가 각각 0.5 s, 1 s, 2 s, 4 s인 잔차 연결을 수행한다. Table A4는 잔차 연결된 게이트 순환 유닛 네트워크 구조를 보여준다.

다른 게이트 순환 유닛 네트워크들과 동일하게 5층의 어텐션 기반 확장된 게이트 순환 유닛 네트워크를 구현한다. 모든 층의 게이트 순환 유닛들은 256 길이의 벡터를 출력하며, τ_0 와 τ_d 는 각각 0.05 s와 0.5 s로 설정된다. 따라서 2에서 5번째 층에서 0.5 s, 1 s, 2 s,

4s 간격의 2배수되는 점진적 확장 연결을 수행한다. 잔차 연결과 다르게 확장 연결은 인접한 유닛과의 연결을 완전히 차단한다. 또한 어텐션 기반 확장된 게이트 순환 유닛 네트워크는 각 층의 종료 시점 유닛의 벡터들을 출력하고 어텐션 메커니즘에 의한 가중치를 적용한다. Table 3는 어텐션 기반 게이트 순환 유닛 네트워크 구조를 보여준다. 본 논문에서 구현하는 모든 분류 모델들은 32 배치 크기 설정과 0.01 학습률의 Adam optimizer를 사용하여 학습되며, 모든 모델들의 손실 값이 수렴할 수 있도록 에폭(Epoch)은 500으로 설정된다. 여기서 배치 크기는 'NVIDIA GeForce GTX 1070' Graphical Processing Unit(GPU)의 구동이 가능도록 설정되고, 학습률은 학습 과정에서 손실 값이 수렴하도록 결정된다.

4.3 분류 결과

4.2절에서 구현된 분류 모델들의 성능을 비교한다. 또한 확장된 게이트 순환 유닛 네트워크에 어텐션 메커니즘의 적용 유무에 따른 성능을 함께 비교한다. 어텐션 메커니즘이 적용되지 않은 경우는 Eq. (13)에서 $w^{(l)}$ 은 모두 1로 결정됨을 의미한다. 본 논문에서는 이러한 네트워크를 균일 확장된 게이트 순환 유닛 네트워크(Uniform dilated GRU network)로 정의한다.

Table 4는 분류 모델에 따른 선종 별 10회 평균된 분류 정확도와 학습 소요 시간을 보여준다. Table 4에서 적층된 게이트 순환 유닛 네트워크의 $\tau=0.05$ 와 $\tau=4$ 는 각각 Tables A2와 A3와 같이 구현된 네트워크의 성능을 의미한다. Table 4의 결과에서 $\tau=0.05$ 인

적층된 게이트 순환 유닛 네트워크는 영상의 상관관계를 학습하는 합성곱 신경망과 비교하여 2.00% 우수한 평균 분류 정확도를 나타낸다. 반면에 $\tau=4$ 인 적층된 게이트 순환 유닛 네트워크는 이산적인 시계열 특징 입력으로 인해 76.16%의 낮은 평균 정확도를 보인다. 잔차 연결된 게이트 순환 유닛 네트워크는 유닛 간 다양한 시간 간격을 학습하여 92.33%의 우수한 평균 분류 정확도를 획득하며, 이러한 결과는 합성곱 네트워크 및 적층된 게이트 순환 유닛 네트워크보다 우수한 성능을 나타낸다.

제안하는 어텐션 기반 확장된 게이트 순환 유닛 네트워크는 96.50%의 가장 우수한 평균 분류 정확도를 보여준다. 이러한 결과는 균일 확장된 게이트 순환 유닛 네트워크와 비교하여 2.84% 향상된 성능을 나타낸다. 따라서 제안하는 네트워크는 매 층별로 서로 다른 시간 간격에 의한 특징벡터들을 추출하고 어텐션 메커니즘을 통해 효율적으로 통합하여 우수한 분류 성능을 획득한다. 추가적으로 제안하는 네트워크는 층수에 비례한 게이트 순환 유닛의 감소로 7.40 min의 적은 학습 시간을 소요한다. 이러한 네트워크의 학습 파라미터 감소는 적은 데이터의 학습에 효율적일 수 있다. Fig. 5는 분류 네트워크 별 대표적인 혼동 행렬을 보여준다. Table 4과 Fig. 5에서 게이트 순환 유닛을 이용한 네트워크들은 합성곱 신경망과 비교하여 여객선과 로로선에 대한 분류 성능을 개선한다. 특히, 제안하는 어텐션 기반 게이트 순환 유닛 네트워크는 여객선과 로로선에 대한 분류 성능을 크게 향상시킨다.

2 kHz 이하의 저주파 대역 스펙트로그램 특징을

Table 4. Accuracy in % and training time for classification models evaluated with ShipEar dataset.

| Ship type | CNN | Stacked GRU | | Skip connected GRU | Uniform dilated GRU | Attention dilated GRU |
|----------------------------|--------------|--------------|--------------|--------------------|---------------------|-----------------------|
| | | $\tau=0.05$ | $\tau=4$ | | | |
| Fishing boat | 97.50 | 92.00 | 82.50 | 95.50 | 96.00 | 98.00 |
| Motor boat | 90.00 | 91.50 | 74.00 | 95.00 | 96.50 | 97.50 |
| Mussel boat | 88.00 | 95.50 | 72.00 | 96.00 | 97.50 | 98.00 |
| Ocean liner | 97.50 | 96.00 | 80.00 | 95.50 | 98.00 | 98.50 |
| Passenger | 70.00 | 75.00 | 72.50 | 80.00 | 77.00 | 90.00 |
| RORO ship | 80.00 | 85.00 | 76.00 | 92.00 | 97.00 | 97.00 |
| Total accuracy | 87.16 | 89.16 | 76.16 | 92.33 | 93.66 | 96.50 |
| Training time (min) | 21.06 | 28.64 | 3.46 | 29.16 | 7.31 | 7.40 |

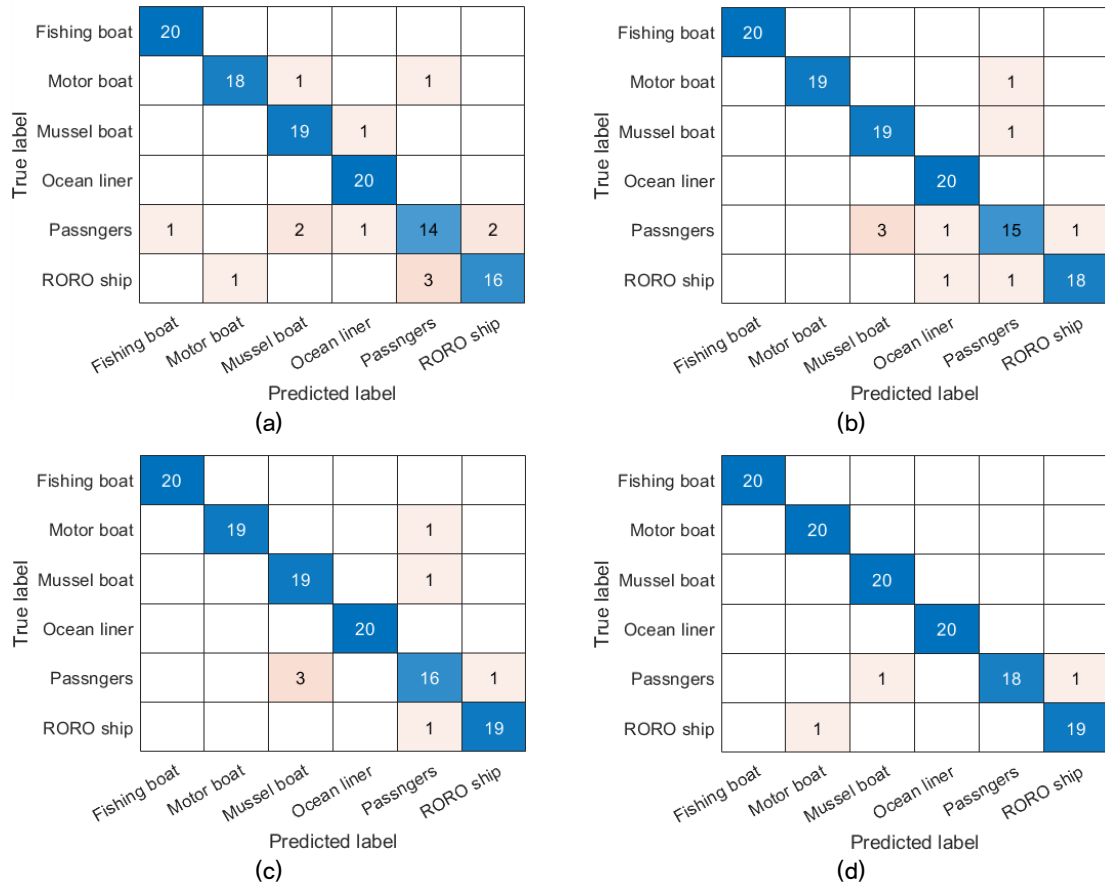


Fig. 5. (Color available online) Confusion matrices. (a) CNN, (b) skip connected GRU network, (c) uniform dilated GRU network, (d) attention dilated GRU network.

Table 5. Accuracy in % for classification models evaluated using low frequency features.

| Ship type | CNN | stacked GRU | | skip connected GRU | uniform dilated GRU | attention dilated GRU |
|----------------------------|--------------|---------------|--------------|--------------------|---------------------|-----------------------|
| | | $\tau = 0.05$ | $\tau = 4$ | | | |
| Fishing boat | 97.00 | 92.00 | 81.00 | 95.00 | 95.50 | 96.00 |
| Motor boat | 90.00 | 91.00 | 70.00 | 93.00 | 95.50 | 96.50 |
| Mussel boat | 84.50 | 94.50 | 71.50 | 95.00 | 96.00 | 97.00 |
| Ocean liner | 95.50 | 95.00 | 78.50 | 94.50 | 95.00 | 96.00 |
| Passenger | 70.00 | 74.50 | 72.00 | 75.50 | 75.00 | 76.00 |
| RORO ship | 75.50 | 85.00 | 75.00 | 87.50 | 90.00 | 90.50 |
| Total accuracy | 85.41 | 88.66 | 74.66 | 90.08 | 91.16 | 92.00 |
| Training time (min) | 21.10 | 28.46 | 3.20 | 29.35 | 7.32 | 7.39 |

사용하여 로이드 거울 효과가 최소화된 상황의 모델 별 분류 성능 및 학습 소요 시간을 Table 5에서 보여 준다. Table 5의 선종 별 분류 정확도는 10회 평균된 값으로 기술된다. 평가에 사용된 저주파 스펙트로그램은 3.9 Hz의 주파수 간격을 제외하고 Table 2와 동일하게 생성되며, 구현된 모델들은 512 × 512 크기의

스펙트로그램에 대한 분류를 수행한다. 로이드 거울 효과에 의한 전역적 시계열 특징이 감소됨에 따라 모든 모델들이 Table 4의 결과와 비교하여 저하된 평균 분류 정확도를 보여준다. 특히 Table 4에서 상대적으로 낮은 분류 정확도를 보이는 여객선과 로로선에 대한 분류 성능이 Table 5에서 더욱 크게 저하된다.

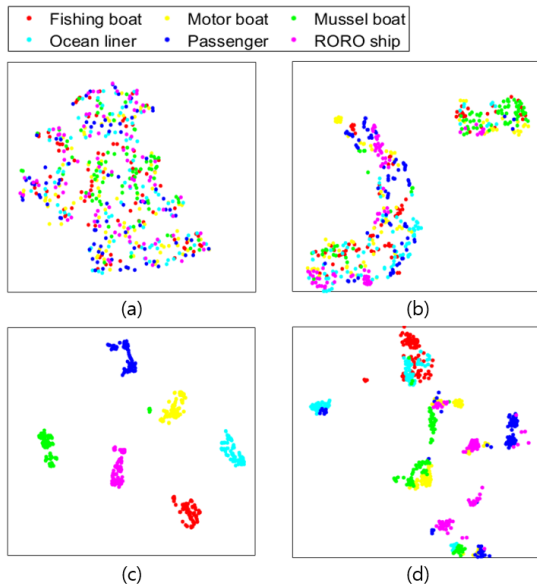


Fig. 6. (Color available online) t-SNE plots of feature vectors extracted from the (a) 1st layer, (b) 3rd layer, (c) 5th layer and (d) 7th layer.

Tables 4와 5에서 제안하는 어텐션 기반 확장된 게이트 순환 유닛 네트워크의 평균 분류 정확도가 4.5% 저하되며, 균일 확장된 게이트 순환 유닛 네트워크와 비교하여 단지 1.16% 높은 평균 정확도를 갖는다. Table 5에서 스펙트로그램의 전역적 시계열 특징이 감소됨에도 불구하고 제안하는 네트워크는 다중 스케일의 시계열 특징들을 효율적으로 통합하여 가장 우수한 92.00%의 평균 분류 정확도를 보여준다. 추가적으로 학습 소요 시간은 입력되는 스펙트로그램의 크기와 모델 별 파라미터의 변화가 없으므로 크게 변동되지 않는다.

Table 3의 어텐션 기반 게이트 순환 유닛 네트워크를 7층으로 확장하고 시계열 특징의 스케일에 따른 변별력을 평가한다. 여기서 확장된 6, 7번째 네트워크는 8s와 16s의 확장된 연결을 갖는다. Figs. 6과 7은 층 별로 추출된 특징벡터의 t-SNE plot과 학습된 어텐션 가중치를 보여준다. Figs. 6과 7에서 특징벡터의 변별력에 따라 어텐션 가중치가 효율적으로 학습됨을 확인할 수 있다. 특히 5번째 층에서 추출된 특징벡터가 우수한 변별력을 보유하고 동시에 높은 어텐션 가중치로 분류에 사용된다. Fig. 6(c)에서 5번째 층의 특징벡터는 홍합선(Mussel boat)의 분류에서 상대적으로 저하된 변별력을 갖는다. 이를 개선하고자 Fig.

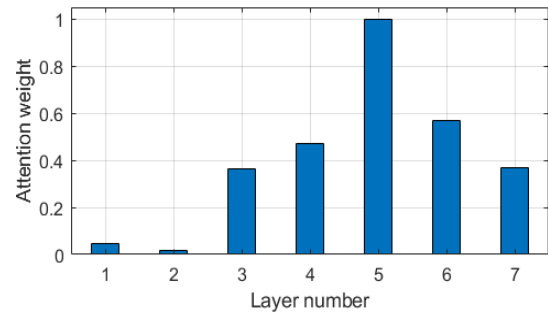


Fig. 7. (Color available online) Attention weights learned according to feature vectors extracted from each layer.

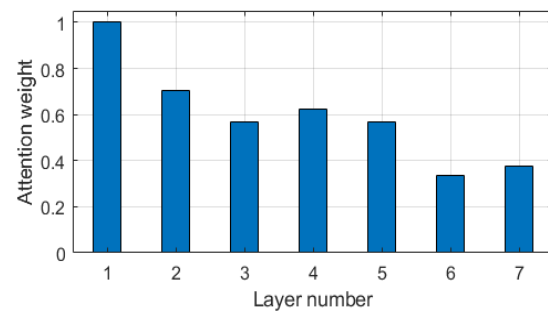


Fig. 8. (Color available online) Attention weights learned using low frequency features.

6(b)와 같이 홍합선에 대해 우수한 변별력을 갖는 3번째 층의 특징벡터에 상대적으로 높은 어텐션 가중치가 설정된다.

Fig. 8은 2 kHz 이하의 저주파 대역 스펙트로그램을 사용하여 학습된 어텐션 가중치를 보여준다. 저주파 대역 신호는 수초 내 세기가 변동되는 특징을 가지며, 어텐션 가중치는 0.05 s 간격의 상관관계를 학습하는 첫 번째 층의 특징벡터에 가장 높게 학습된다. 이로써 제안하는 네트워크의 어텐션 메커니즘이 다중 스케일의 시계열 특징벡터들을 변별력에 따라 효율적으로 가중하는 것을 검증한다.

1에서 7층까지의 어텐션 가중치가 적용된 특징벡터를 순차적으로 누적 합하여 Fig. 9와 같이 분류 성능을 평가한다. Fig. 9의 분류 정확도는 10회 평균하여 계산된다. Fig. 9에서 특징벡터가 누적됨에 따라 분류 성능이 향상되고 5번째부터 96.50%로 수렴되는 것을 확인할 수 있다. Fig. 10은 어텐션 가중치가 높은 순서로 특징벡터들을 누적 합하여 평가된 분류 성능을 보여준다. 따라서 Fig. 7에서 결정된 어텐션

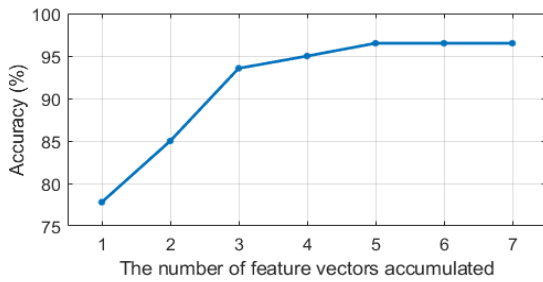


Fig. 9. (Color available online) Average accuracies for cumulative sum of feature vectors according to layer order.

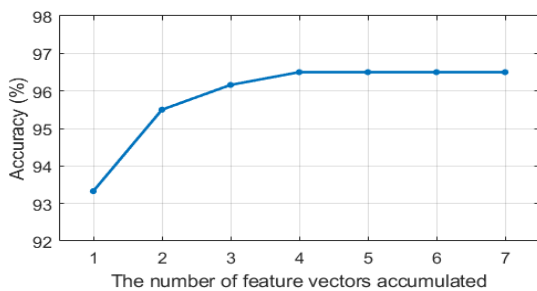


Fig. 10. (Color available online) Average accuracies for cumulative sum of feature vectors according to order of attention weights.

가중치에 따라 5, 6, 4, 7, 3, 1, 2층의 순서로 누적 합을 수행하고, 특징벡터의 누적 합에 따른 분류 정확도는 10회 평균하여 도출된다. Fig. 10에서 5번째 단일 특징벡터를 사용할 경우 93.33%의 평균 분류 정확도를 가지며, 특징벡터를 누적함에 따라 96.50%의 향상된 성능을 얻을 수 있다. 따라서 제안하는 네트워크는 어텐션 메커니즘을 이용하여 변별력이 낮은 특징벡터의 사용으로 인한 성능 저하를 방지한다.

V. 결론

수동소나의 표적신호는 기계소음에 의한 협대역 고조파 특성과 로이드 거울 효과에 의한 시간에 따른 주파수 변이 특성을 갖는다. 협대역 고조파 신호는 단시간 내 세기의 변화를 가지며, 로이드 거울 효과는 장시간에 걸친 주파수 변이를 야기한다. 본 논문에서는 수동소나 신호분류를 위해 다양한 스케일의 시계열 특징을 학습하는 어텐션 기반 게이트 순환 유닛 네트워크를 제안하였다. 제안하는 분류 모델은 효율적인 시계열 특징의 학습이 가능한 게이트

순환 유닛을 사용하여 다층 구조의 네트워크를 구성하고, 층에 따라 점진적으로 확장된 연결을 통해 다중 스케일의 시계열 특징벡터들을 추출한다. 추출된 특징벡터들은 어텐션 메커니즘 학습을 통해 가중되어 신호분류에 사용된다. 수중 음향 데이터를 이용한 실험에서 제안된 모델은 기존 모델들과 비교하여 96.50%의 가장 우수한 평균 분류 정확도를 보였다.

제안된 분류 알고리즘은 수동소나를 이용한 감시 체계에 응용되어 군사적 목적에 크게 기여할 수 있을 것으로 기대된다. 향후 다양한 수중 음향 환경과 표적 신호에 대한 알고리즘의 실험적 검증이 수행되어야 한다.

부 록(Appendix)

Table A1. Convolutional network architecture.

| |
|--|
| Input: $512 \times 512 \times 1$ Shape |
| Layer 1: 64 Filters, 3×3 Convolution, 1×1 Stride, ReLU activation function, 3×3 Max pooling, Batch normalization, |
| Layer 2: 32 Filters, 3×3 Convolution, 1×1 Stride, ReLU activation function, 3×3 Max pooling, Batch normalization, |
| Layer 3: 16 Filters, 2×2 Convolution, 1×1 Stride, ReLU activation function, 2×2 Max pooling, Batch normalization, |
| Layer 4: 16 Filters, 2×2 Convolution, 1×1 Stride, ReLU activation function, 2×2 Max pooling, Batch normalization, |
| Layer 5: 16 Filters, 2×2 Convolution, 1×1 Stride, ReLU activation function, Batch normalization, |
| Layer 6: Fully connected layer (2304 input, 6 output), Softmax |

Table A2. Stacked GRU network architecture with 0.05 s interval.

| |
|--|
| Input: $512 \times 512 \times 1$ Shape |
| Layer 1: 512×512 GRU sequence output shape |
| Layer 2: 512×256 GRU sequence output shape |
| Layer 3: 512×128 GRU sequence output shape |
| Layer 4: 512×64 GRU sequence output shape |
| Layer 5: 512×64 GRU sequence output shape, 1×64 GRU output shape |
| Layer 6: Fully connected layer (64 input, 6 output), Softmax |

Table A3. Stacked GRU network architecture with 4 s interval.

| |
|---|
| Input: $6 \times 512 \times 1$ Shape |
| Layer 1: 6×512 GRU sequence output shape |
| Layer 2: 6×256 GRU sequence output shape |
| Layer 3: 6×128 GRU sequence output shape |
| Layer 4: 6×64 GRU sequence output shape |
| Layer 5: 6×64 GRU sequence output shape, 1×64 GRU output shape |
| Layer 6: Fully connected layer (64 input, 6 output), Softmax |

Table A4. Skip connected GRU network architecture.

| |
|---|
| Input: $512 \times 512 \times 1$ Shape |
| Layer 1: 512×512 GRU sequence output shape, |
| Layer 2: 512×256 GRU sequence output shape, 10 skip connection interval |
| Layer 3: 512×128 GRU sequence output shape, 20 skip connection interval |
| Layer 4: 512×64 GRU sequence output shape, 40 skip connection interval |
| Layer 5: 512×64 GRU sequence output shape, 1×64 GRU output shape, 80 skip connection interval |
| Layer 6: Fully connected layer (64 input, 6 output), Softmax |

감사의 글

이 논문은 2023년도 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원을 받아 수행된 연구임(20-106-B00-003).

References

1. Y. C. Jung, B. U. Kim, S. K. An, W. J. Seong, and K. H. Lee, "An algorithm for submarine passive sonar simulator" (in Korean), *J. Acoust. Soc. Kr.* **32**, 472-483 (2013).
2. K. Lee, J. Kim, M. J. Cheong, Y. W. Ryu, J. Bae, and C. H. Lee, "Separation and denoising algorithm of passive sonar target signals in frequency domain" (in Korean), *J. IEIE*, **55**, 88-96 (2018).
3. J. K. Ahn, H. D. Cho, D. Shin, T. Kwon, and G. T. Kim, "LOFAR/DEMON grams compression method for passive sonar" (in Korean), *J. Acoust. Soc. Kr.* **39**, 38-46 (2020).
4. S. E. Lee, S. B. Hwang, and D. Y. Noh, "A study on the algorithm for underwater target automatic classification using the passive sonar" (in Korean), *J. KIMS Technol.* **3**, 76-84 (2000).
5. B. Mishachandar and S. Vairamuthu, "Diverse ocean noise classification using deep learning," *Appl. Acoust.* **181**, 108141 (2021).
6. G. Song, X. Guo, W. Wang, Q. Ren, J. Li, and L. Ma, "A machine learning-based underwater noise classification method," *Appl. Acoust.* **184**, 108333 (2021).
7. C. Satheesh, S. Kamel, A. Mujeeb, and M. H. Supriya, "Passive sonar target classification using deep generative β -VAE," *IEEE Signal Process. Lett.* **28**, 808-812 (2021).
8. N. H. Bach, L. H. Vu, and V. D. Nguyen, "Classification of surface vehicle propeller cavitation noise using spectrogram processing in combination with convolution neural network," *Sensors*, **21**, 3353 (2021).
9. F. Liu, T. Shen, Z. Luo, D. Zhao, and S. Guo, "Underwater target recognition using convolutional recurrent neural networks with 3-D mel-spectrogram and data augmentation," *Appl. Acoust.* **178**, 107989 (2021).
10. 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).
11. K. M. Park and D. Kim, "Preprocessing performance of convolutional neural networks according to characteristic of underwater targets" (in Korean), *J. Acoust. Soc. Kr.* **41**, 629-636 (2022).
12. A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D: Nonlinear Phenomena*, **404**, 132306 (2020).
13. J. Chung, C. Gulcehre, K. H. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *Proc. NIPS Workshop on Deep Learning*, 1-9 (2014).
14. X. Zhang, Y. Sun, K. Jiang, C. Li, L. Jiao, and H. Zhou, "Spatial sequential recurrent neural network for hyperspectral image classification," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **11**, 4141-4155 (2018).
15. S. Singh, S. K. Pandey, U. Pawar, and R. R. Janghel, "Classification of ECG arrhythmia using recurrent neural networks," *Procedia Computer Science*, **132**, 1290-1297 (2018).
16. S. Chang, Y. Zhang, W. Han, M. Yu, X. Guo, W. Tan, X. Cui, M. Witbrock, M. Hasegawa-Johnson, and T. S. Huang, "Dilated recurrent neural networks," *Proc. 31st Conf. NIPS*, 1-13 (2017).
17. S. H. Kang, "A study on the Lloyd's mirror effect on the underwater radiated noise for the underwater vehicle" (in Korean), *J. Acoust. Soc. Kr.* **40**, 314-319

(2021).

18. L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics* (John Wiley & Sons, New Jersey, 1999), Chap. 15.
19. D. S. Domingues, S. T. Guijarro, A. C. Lopez, and A. P. Gimenez, "ShipEar: An underwater vessel noise database," *Appl. Acoust.* **113**, 64-69 (2016).

저자 약력

▶ 이 기 배 (Kibae Lee)



2016년: 제주대학교 해양시스템공학과 학사 졸업
 2017년: 제주대학교 해양시스템공학과 석사 졸업
 2017년~2021년: 경원산업(주) 기업부설연구소 연구원
 2021년~현재: 제주대학교 지구해양융합학부 해양시스템공학과 박사과정

▶ 고 건 혁 (Guhn Hyeok Ko)



2023년: 제주대학교 해양시스템공학과 학사 졸업
 2023년~현재: 제주대학교 지구해양융합학부 해양시스템공학과 석사과정

▶ 이 종 현 (Chong Hyun Lee)



1985년: 한양대학교 전자공학과 학사 졸업
 1987년: Michigan Technological University 석사 졸업
 2002년: 한국과학기술원(KAIST) 전기 및 전자공학과 박사 졸업
 1990년~1995년: 한국전자통신연구원 선임연구원
 2000년~2002년: (주)KM Telecom 연구소장
 2003년~2006년: 서경대학교 전자공학과 전임강사
 2017년~2018년: Georgia Institute of Technology 방문교수
 2006년~현재: 제주대학교 해양시스템공학과 교수