IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 12, 2022, accepted May 22, 2022, date of publication May 23, 2022, date of current version May 27, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3177594

Deep Learning-Based Near-Field Source Localization Without a Priori Knowledge of the Number of Sources

HOJUN LEE, YONGCHEOL KIM, SEUNGHWAN SEOL, AND JAEHAK CHUNG
Department of Electronics Engineering, Inha University, Incheon 22212, South Korea
Corresponding author: Jaehak Chung (jchung@inha.ac.kr)

This work was supported by the Inha University Research Program.

ABSTRACT In this paper, we propose a high resolution grid-based deep learning source localization that precisely estimates the locations of near-field sources without a priori knowledge of the number of sources. The proposed method consists of a principal component analysis network (PCAnet) and a spatial spectrum network (Sp2net). The proposed PCAnet calculates the noise spaces of the received signals by convolutional layers without a priori knowledge or the estimation of the number of sources and has the lower complexity than eigenvalue decomposition (EVD). The proposed Sp2net calculates the spatial spectrum with a linear layer from the output of the PCAnet and classifies dense location grids with a convolutional neural network (CNN). From the spatial spectrum, this paper also proposes an activation function to enlarge the values at the grid points where the near-field sources exist, which are differentiable for all input values. Then, the direction of arrivals (DOAs) and the ranges of the near-field sources are estimated with high resolution. Computer simulations demonstrated that the proposed method had better DOA and range estimation performances than those of the conventional methods.

INDEX TERMS Array signal processing, direction-of-arrival estimation, machine learning, signal detection.

I. INTRODUCTION

Source localizations (SLs) play an important role in the research area of array signal processing such as radar, sonar, and wireless communications [1]–[6]. Most SLs focused on estimating the direction-of-arrivals (DOAs) for the far-field sources with a plane wave model. When the sources are located in the near-field, the plane wave assumption is no longer appropriate, and the wavefronts of the near-field sources are considered as a spherical form [7], [8]. Thus, the near-field SLs (NFSLs) needed to estimate the DOAs and the ranges of the sources.

For the NFSLs, the two-dimensional multiple signal classification (2D-MUSIC) and lots of MUSIC-based methods have been studied [7]–[11]. The conventional MUSIC-based methods had large computational complexities because eigenvector decomposition (EVD) is utilized to decompose the received signals into the signal and noise spaces. In [12], the reduced-dimension MUSIC (RD-MUSIC) was proposed to reduce the computational complexity. However, their estimation performances decreased when the estimated number of sources is incorrect.

To solve the drawback of the MUSIC-based methods, the NFSLs without the knowledge of the number of sources were proposed [13], [14], [16]. In [13], the discrete fractional Fourier transform (DFrFT) based SL was proposed. The DFrFT was designed for one snapshot, which was unsuitable for low signal-to-noise power ratio (SNR) scenarios. In [14], the low complexity localization algorithm (LCLA) was developed. The LCLA estimated the DOAs and the ranges by DFT and orthogonal matching pursuit (OMP), respectively. Since DFT has a frequency bias by the DFT size, the frequency bias caused the DOA estimation errors, which also led to the range estimation errors. In [15], [16], the fourth-order cumulant (FOC) based methods were proposed. The FOC-based methods had the error propagation between the DOA and the range estimations by the dependence of the DOA and the range estimation processes.

Various deep learning (DL) based researches have been conducted in many applications such as image processing,

The associate editor coordinating the review of this manuscript and approving it for publication was Bilal Alatas.

signal detection, and estimation theory. DL-based SLs also have been studied [17]–[26]. The DLSLs were implemented by a deep neural network (DNN) or a convolutional neural network (CNN) with preprocessed data [17]–[26]. The DLSLs were classified as grid-free and grid-based methods.

The grid-free method calculates the coordinates of the source as a specific value by the regression of the DLSL. Since the grid-free method has no constraint of the resolution, the locations of the sources are precisely estimated. In the grid-free method, however, the maximum number of sources needs to be determined in advance. If the number of actual sources exceeds the predetermined number, the SL results may be incorrect [25], [26].

The grid-based scheme assigns labels to all range- and azimuth-grids on the target space and estimates the source locations by the classification. The resolution of the estimation is determined according to the grid interval, but all sources are detected without the predetermined number of sources. To obtain the high resolution, however, the grid-based method needs so many classes to map all range- and azimuth-grids, and the learning process with the conventional DNN and CNN structures also requires many features and training data to classify the classes [27]. Thus, the high resolution grid-based DLSL may not be implemented by the conventional learning structures. In the conventional grid-based methods, the DOAs of the far-field sources were estimated with a large grid interval of 5° [18], [22] or 10° [23], [24] to keep the acceptable number of classes.

Based on the literature reviews, the problems of the conventional NFSLs are as follows: (a) The MUSIC-based methods require a priori knowledge or the perfect estimation of the number of sources; (b) The NFSLs without the knowledge of the number of sources perform worse in a low SNR; (c) The grid-free DLSLs need the predetermined maximum number of sources; (d) The grid-based DLSLs are difficult to achieve the high resolution.

In this paper, therefore, we propose a novel high resolution grid-based DLSL structure that learns and estimates both the DOAs and the ranges of the near-field sources without a priori knowledge of the number of sources. Since the proposed network works on the grid-based, the maximum number of the near-field sources is not required. To improve the learning and estimation performances for many classes, the proposed network consists of a principal component analysis network (PCAnet) that estimates the noise space and a spatial spectrum network (Sp2net) that precisely estimates the DOAs and the ranges of the near-field sources. The proposed PCAnet solves the problems of the conventional MUSIC-based methods by calculating the noise spaces without a priori knowledge or the estimation of the number of sources and with lower computational complexity than EVD. The proposed Sp2net calculates the spatial spectrum from the noise spaces with a linear layer. An activation function in the Sp2net is also proposed to be differentiable for all inputs and to enlarge the values of the null spaces from the spatial spectrum. Then, the proposed Sp2net classifies the source location grids with a CNN and estimates both the DOAs and the ranges of the near-field sources with high resolution. Computer simulations demonstrated that the SL estimation performance of the proposed method was better than that of the conventional methods.

The contributions of this paper are fourfold:

- We propose the PCAnet that calculates the noise spaces with the lower computational complexity than EVD and without a priori knowledge or the estimation of the number of sources.
- We propose the Sp2net that precisely estimates both the DOAs and the ranges of the near-field sources with high resolution.
- We propose the custom activation function that is differentiable for all input values and enlarges the values of the null spaces from the spatial spectrum.
- The proposed method improved the SL estimation performance without a priori knowledge of the number of sources by using DL.

II. SIGNAL MODEL

Assume that L far-field or near-field sources independently radiate narrowband signals. This paper focuses on the NFSL in mixed near- and far-field signal scenarios. The signals are statistically mutually independent and uncorrelated with noise. The signals are captured by a uniform linear array (ULA) consisting of N sensors with spacing d. The configuration of the ULA with the l-th signal is displayed in Fig. 1.

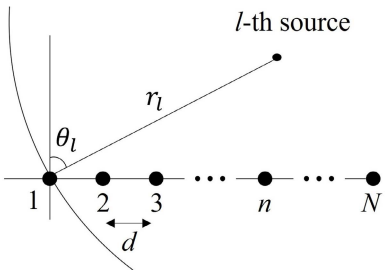


FIGURE 1. Uniform linear array configuration.

Let the DOA and the range of the arbitrary l-th source be θ_l and r_l, respectively, and the radiated narrowband signal at the k-th snapshot be $s_l(k)$. The received signal ($x_n(k)$) of the n-th sensor is given as,

$$x_n(k) = \sum_{l=1}^{L} s_l(k)\, g_{n,l} e^{j\tau_n(\theta_l, r_l)} + w_n(k), \qquad (1)$$
$$n = 1, \ldots, N,\ k = 0, \ldots, K-1,$$

where K and $w_n(k)$ denote the number of snapshots and the additive Gaussian noise of the n-th sensor, respectively. $g_{n,l}$ denotes the l-th source's magnitude attenuation for the n-th sensor with respect to the reference sensor [28]–[31]. $\tau_n(\theta_l, r_l)$ denotes the relative propagation delay of the n-th sensor to the reference sensor, which is set to the first sensor

in Fig. 1, and is expressed as [4], [26],

$$\tau_n(\theta_l, r_l) = \begin{cases} \frac{2\pi}{\lambda}\left(\sqrt{r_l^2 + (nd)^2 - 2r_l nd\sin(\theta_l)} - r_l\right), & \text{for near-field,} \\ -2\pi(n-1)\frac{d}{\lambda}sin(\theta_l), & \text{for far-field,} \end{cases} \tag{2}$$

where λ denotes wavelength. The near-field sources are located in the near-field and the Fresnel region, i.e., $r_l \in (0, 2D^2/\lambda)$ where D is the array aperture [8], [26], and the far-field sources are located further than $2D^2/\lambda$.

In a matrix form for N sensors, Eq. (1) can be written as,

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{w}(k), \tag{3}$$

where

$$\mathbf{x}(k) = [x_1(k) \cdots x_n(k) \cdots x_N(k)]^{\mathrm{T}}, \tag{4}$$
$$\mathbf{s}(k) = [s_1(k) \cdots s_l(k) \cdots s_L(k)]^{\mathrm{T}}, \tag{5}$$
$$\mathbf{w}(k) = [w_1(k) \cdots w_n(k) \cdots w_N(k)]^{\mathrm{T}}, \tag{6}$$
$$\mathbf{A} = [\boldsymbol{\alpha}(\theta_1, r_1) \cdots \boldsymbol{\alpha}(\theta_l, r_l) \cdots \boldsymbol{\alpha}(\theta_L, r_L)], \tag{7}$$
$$\boldsymbol{\alpha}(\theta_l, r_l) = \left[g_{1,l}e^{j\tau_1(\theta_l, r_l)} \cdots g_{N,l}e^{j\tau_N(\theta_l, r_l)}\right]^{\mathrm{T}}. \tag{8}$$

III. PROPOSED METHOD

The proposed method learns and estimates the DOAs and ranges of the near-field sources with the input of the covariance matrix of the received signals. For the high resolution SL, the proposed method utilizes the orthogonality between the signal and noise spaces, which is the same idea as the conventional MUSIC. The proposed method consists of the PCAnet and the Sp2net in Fig. 2. The proposed PCAnet calculates the noise spaces from the covariance matrix of the received signals with the low complexity and without a priori knowledge or the estimation of the number of sources. The proposed Sp2net precisely estimates the DOAs and ranges of the near-field sources with the CNN and the proposed activation function.

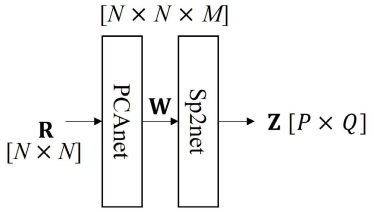


FIGURE 2. Block diagram of the proposed network.

In Fig. 2, $\mathbf{R}$ denotes the covariance matrix of the received signals ($\mathbf{x}(k)$) in Eq. (3), and is given as,

$$\mathbf{R} = \mathrm{E}\left\{\mathbf{x}(k)\ \mathbf{x}^{\mathrm{H}}(k)\right\} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^{\mathrm{H}} + \sigma_w^2 I, \tag{9}$$

where $\mathbf{R}_{ss} = \mathrm{E}\left\{\mathbf{s}(k)\ \mathbf{s}^{\mathrm{H}}(k)\right\}$ and σ_w^2 denotes the variance of the noise. $\mathbf{W}$ denotes the noise spaces, and M denotes the number of channels of the noise spaces. $\mathbf{Z}$ denotes the estimated DOAs and ranges of the near-field sources. P and Q denote the numbers of azimuth- and range-classes, respectively. The number of all classes for the source location is $P \times Q$.

The following subsections describe the proposed networks.

A. PRINCIPAL COMPONENT ANALYSIS NETWORK

In the conventional MUSIC, the EVD is utilized to decompose the signal and noise spaces. However, the EVD not only has a high complexity but also needs knowledge of the number of sources. To solve the problems, this paper proposes the PCAnet that decomposes the signal and noise spaces by the convolutional layers (Convs) that utilize the Gram-Schmidt process (GS) idea. The PCAnet divides $\mathbf{R}$ into multiple blocks to reduce the complexity and orthogonalizes the divided blocks in parallel by the proposed GS neural network (GNN). The PCAnet calculates $\mathbf{W}$ by a depthwise Conv (DepthConv) without a priori knowledge or the estimation of the number of sources. The block diagram of the proposed PCAnet is represented in Fig. 3.

To reduce the complexity, the proposed PCAnet divides $\mathbf{R}$ into D block covariance matrices with $[N \times B]$ in Fig. 3, where B denotes the column block size and $D = N/B$. The divided $\mathbf{R}$ is expressed as,

$$\mathbf{R} = \begin{bmatrix}\mathbf{R}_1 & \cdots & \mathbf{R}_d & \cdots & \mathbf{R}_D\end{bmatrix}, \tag{10}$$

where

$$\mathbf{R}_d = \begin{bmatrix}\mathbf{v}_{d,1} \cdots \mathbf{v}_{d,b} \cdots \mathbf{v}_{d,B}\end{bmatrix}, \tag{11}$$
$$\mathbf{v}_{d,b} = \begin{bmatrix}v_{d,b}^1 \cdots v_{d,b}^n \cdots v_{d,b}^N\end{bmatrix}^{\mathrm{T}}. \tag{12}$$

The detailed GNN process is described as follows: For the d-th block, the proposed GNN computes the block eigenvectors $\left(\mathbf{Q}_d \in \mathbb{R}^{N\times B}\right)$ from the block covariance matrix $\left(\mathbf{R}_d \in \mathbb{R}^{N\times B}\right)$. $\mathbf{Q}_d$ is obtained by removing dependent components between column vectors of $\mathbf{R}_d$ like the conventional GS.

In the proposed GNN, each column vector of $\mathbf{R}_d$ is sequentially orthogonalized. The first column vector of $\mathbf{R}_d$ is normalized and set to the first block eigenvector $\left(\mathbf{Q}_d^1 = \mathbf{q}_{d,1} \in \mathbb{R}^{N\times 1}\right)$, i.e., $\mathbf{q}_{d,1} = \mathbf{v}_{d,1}/\|\mathbf{v}_{d,1}\|$, where $\|\cdot\|$ denotes a norm operator. For the b-th column vector ($\mathbf{v}_{d,b} \in \mathbb{R}^{N\times 1}$, where $b = 2, 3, \ldots, B$) of $\mathbf{R}_d$, the orthonormal basis $\left(\mathbf{q}_{d,b} \in \mathbb{R}^{N\times 1}\right)$ is calculated by subtracting the dependent components of $\mathbf{v}_{d,b}$ for $\mathbf{Q}_d^{b-1} = \begin{bmatrix}\mathbf{q}_{d,1}\ \mathbf{q}_{d,2} \cdots \mathbf{q}_{d,b-1}\end{bmatrix}$ from $\mathbf{v}_{d,b}$.

To calculate the dependent components, the conventional GS utilizes the inner product, but this paper replaces the inner product with the convolutional layer (Conv) because the Conv attains the features from the inputs [17], [18], [26], [32]–[36]. However, since the general Conv sums the convolution results for each channel, this paper utilizes the 1-dimensional DepthConv (DepthConv1D) to calculate the

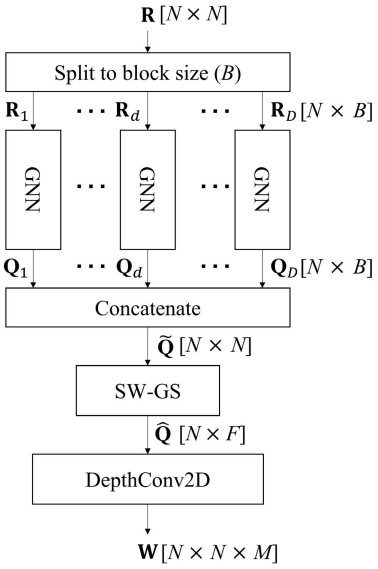


FIGURE 3. Block diagram of the PCAnet.

dependent components of $\mathbf{v}_{d,b}$ for $\mathbf{Q}_d^{b-1}$ with learnable weights. For computing the dependent components of $\mathbf{v}_{d,b}$ for each column vector of $\mathbf{Q}_d^{b-1}$, the DepthConv1D is designed with the kernel size of N, the filter size of one, and the stride of one. The input of the DepthConv1D is set to the result of the Hadamard product for the complex conjugate of $\mathbf{Q}_d^{b-1}$ and $\mathbf{V}_{d,b} \in \mathbb{R}^{N\times(b-1)}$, where $\mathbf{V}_{d,b}$ is $(b-1)$ replications of $\mathbf{v}_{d,b}$. The DepthConv1D executes a weighted sum to the column vector of $\left\{\mathbf{Q}_d^{b-1}\right\}^* \odot \mathbf{V}_{d,b}$ with the learnable weights and attains the dependent components $\left(\mathbf{c}_{d,b} \in \mathbb{R}^{1\times(b-1)}\right)$, where $\odot$ and $*$ denote the Hadamard product operator and complex conjugate, respectively. Assume $f_{b,d}^G(\cdot)$ is the DepthConv1D for the b-th state in the d-th block. $\mathbf{c}_{d,b}$ is expressed as,

$$\mathbf{c}_{d,b} = f_{b,d}^G\left(\left\{\mathbf{Q}_d^{b-1}\right\}^* \odot \mathbf{V}_{d,b}\right). \tag{13}$$

Since $\mathbf{c}_{d,b}$ in Eq. (13) have positive or negative values, a leaky rectified linear unit (LeakyReLU) is used as an activation function [35]. Note that the DepthConv1D includes the activation function.

Then, the block orthonormal basis $\left(\mathbf{q}_{d,b} \in \mathbb{R}^{N\times 1}\right)$ of $\mathbf{v}_{d,b}$ is calculated by removing $\mathbf{c}_{d,b}$ from $\mathbf{v}_{d,b}$, and is as,

$$\mathbf{q}_{d,b} = \widetilde{\mathbf{q}}_{d,b}/\|\widetilde{\mathbf{q}}_{d,b}\|, \tag{14}$$

where

$$\widetilde{\mathbf{q}}_{d,b} = \mathbf{v}_{d,b} - \mathbf{Q}_d^{b-1}\mathbf{c}_{d,b}{}^{\mathrm{T}}. \tag{15}$$

The block eigenvectors $\left(\mathbf{Q}_d^b \in \mathbb{R}^{N\times b}\right)$ for $\mathbf{v}_{d,1}, \mathbf{v}_{d,2}, \ldots, \mathbf{v}_{d,b}$ are computed by concatenating $\mathbf{Q}_d^{b-1}$ and $\mathbf{q}_{d,b}$, and is given as,

$$\mathbf{Q}_d^b = \left[\mathbf{Q}_d^{b-1} \mid \mathbf{q}_{d,b}\right]. \tag{16}$$

The d-th block eigenvectors $\left(\mathbf{Q}_d = \mathbf{Q}_d^B \in \mathbb{R}^{N\times B}\right)$ are obtained from Eq. (13) to Eq. (16) for $b = 2, 3, \ldots, B$, and the block eigenvectors for all blocks are attained.

Let all block eigenvectors be concatenated, which is referred to as $\widetilde{\mathbf{Q}} \in \mathbb{R}^{N\times N}$, and is expressed as,

$$\widetilde{\mathbf{Q}} = \left[\mathbf{Q}_1 \cdots \mathbf{Q}_d \cdots \mathbf{Q}_D\right]. \tag{17}$$

Since $\widetilde{\mathbf{Q}}$ in (17) is calculated by the divided block covariance matrices, the orthogonality among the block eigenvectors may not be satisfied. Thus, the dependent components among the blocks need to be removed.

To obtain approximated eigenvectors ($\widehat{\mathbf{Q}}$) from $\widetilde{\mathbf{Q}}$, this paper proposes a sliding window-GS (SW-GS). The proposed SW-GS divides $\widetilde{\mathbf{Q}}$ into submatrices by the size of the sliding window and orthogonalizes $\widetilde{\mathbf{Q}}$ by removing dependent components among the submatrices. When the sliding interval and the window size are set to $s\,(\leq B)$ and $[N \times B]$, respectively, $\widetilde{\mathbf{Q}}$ is divided into E submatrices, where $E = (N-B)/s + 1$. The e-th submatrix ($\widetilde{\mathbf{Q}}_e$, where $e \in [1, E]$) consists of the column vectors from the $\{(e-1)\,s+1\}$-th to the $\{(e-1)\,s+B\}$-th of $\widetilde{\mathbf{Q}}$.

The first approximated eigenvectors ($\widehat{\mathbf{Q}}_1 = \widehat{\mathbf{q}}_1 \in \mathbb{R}^{N\times B}$) are set to the first submatrix ($\widetilde{\mathbf{Q}}_1$), i.e., the same as $\mathbf{Q}_1$ in Eq. (17). The e-th submatrix ($\widetilde{\mathbf{Q}}_e$) is orthogonalized by removing the dependent components for $\widehat{\mathbf{q}}_1, \widehat{\mathbf{q}}_2, \ldots, \widehat{\mathbf{q}}_{e-1}$, which is referred to as $\widehat{\mathbf{q}}_e \in \mathbb{R}^{N\times B}$. Let $\left[\widehat{\mathbf{q}}_1\ \widehat{\mathbf{q}}_2 \cdots \widehat{\mathbf{q}}_{e-1}\right]$ be the $(e-1)$-th approximated eigenvectors ($\widehat{\mathbf{Q}}_{e-1} \in \mathbb{R}^{N\times(e-1)B}$). Then, $\widehat{\mathbf{q}}_e$ is as,

$$\widehat{\mathbf{q}}_e = \widetilde{\mathbf{Q}}_e - \widehat{\mathbf{Q}}_{e-1}\left[\widehat{\mathbf{Q}}_{e-1}^{\mathrm{H}}\widetilde{\mathbf{Q}}_e\right]. \tag{18}$$

The e-th approximated eigenvectors ($\widehat{\mathbf{Q}}_e \in \mathbb{R}^{N\times eB}$) are obtained by concatenating $\widehat{\mathbf{Q}}_{e-1}$ and $\widehat{\mathbf{q}}_e$, and is expressed as,

$$\widehat{\mathbf{Q}}_e = \left[\widehat{\mathbf{Q}}_{e-1} \mid \widehat{\mathbf{q}}_e\right]. \tag{19}$$

The approximated eigenvectors ($\widehat{\mathbf{Q}} = \widehat{\mathbf{Q}}_E \in \mathbb{R}^{N\times F}$, where $F = E \times B$) are calculated from Eq. (18) to Eq. (19) for $e = 2, 3, \ldots, E$.

The estimation of the noise spaces from $\widehat{\mathbf{Q}}$ in the conventional method is important to maximize the SL performance. In the proposed PCAnet, the noise spaces are estimated by assigning different learnable weights to each column vector of $\widehat{\mathbf{Q}}$ without estimating the number of sources and by learning the weights. After the learning process, the learnable weights of the signal and the noise spaces are converged to zero and arbitrary positive values, respectively. The noise spaces can be found from $\widehat{\mathbf{Q}}$.

To assign and learn the weights, a 2-D DepthConv (DepthConv2D) with the kernel size of (1,1), the filter size of M, and the stride of (1,1) is used. The kernels of the DepthConv2D are constrained to nonnegative values without biases to prevent the phase variation of the noise space. The activation function in the DepthConv2D is not utilized to preserve the orthogonality between the signal and noise spaces. To apply to the different learnable weights for each column vector of $\widehat{\mathbf{Q}} \in \mathbb{R}^{N\times F}$, the column vectors are channelized to a matrix of $[N \times 1 \times F]$. Since the filter size of the DepthConv2D

is M, the outputs for F channels increase by M times and the diversity is attained by M filters.

The output of the DepthConv2D is reshaped to $\mathbf{U} \in \mathbb{R}^{N\times F\times M}$ for M channels, i.e., $\mathbf{U}_1, \ldots, \mathbf{U}_m, \ldots, \mathbf{U}_M$. The noise spaces ($\mathbf{W} \in \mathbb{R}^{N\times N\times M}$) with M channels are obtained from each channel of $\mathbf{U}$, and the noise space ($\mathbf{W}_m \in \mathbb{R}^{N\times N}$) of the m-th channel is calculated as,

$$\mathbf{W}_m = \mathbf{U}_m \mathbf{U}_m^{\mathrm{H}}, \quad \text{for } m = 1, 2, \ldots, M. \tag{20}$$

Therefore, the proposed PCAnet calculates M noise spaces without a priori knowledge or the estimation of the number of sources by Eq. (20).

The following subsection describes the proposed Sp2net, which estimates the DOAs and the ranges of the near-field sources by using the estimated M noise spaces.

B. SPATIAL SPECTRUM NETWORK

The proposed Sp2net estimates the DOAs and the ranges of the near-field sources with high resolution by utilizing the linear layer and the CNN. The target space is gridded by two dimensions consisting of the azimuths and the ranges. The virtual steering vectors are generated by Eqs. (2) and (8) for the azimuth- and range-grid points. At the linear layer, the product of $\mathbf{W}$ and the virtual steering vectors is inputted, and the spatial spectrum is outputted. To obtain the high resolution spatial spectrum, the proposed Sp2net utilizes the orthogonality between the signal and noise spaces like the conventional MUSIC. Thus, this paper also proposes an activation function that is differentiable and enlarges the values of the null spaces from the spatial spectrum. Then, a CNN estimates the DOAs and the ranges of the near-field sources with the high resolution from the spatial spectrum by classifying the dense grids of the source locations. The proposed Sp2net is depicted in Fig. 4.

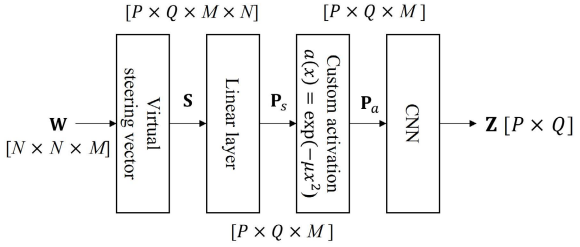


FIGURE 4. Block diagram of Sp2net.

In Fig. 4, P, Q, M, and N denote the numbers of azimuth classes, range classes, noise space channels, and sensors, respectively. Assume that $p \in [1, P]$, $q \in [1, Q]$, $n \in [1, N]$, and $m \in [1, M]$. $\mathbf{S} \in \mathbb{R}^{P\times Q\times M\times N}$ is the result of multiplying $\mathbf{W} \in \mathbb{R}^{N\times N\times M}$ by the virtual steering vector ($\alpha(p, q, n)$), and the element of $\mathbf{S}$ is calculated as,

$$S(p, q, m, n) = \alpha(p, q, n) \times \sum_{j=1}^{N} \left\{\alpha^*(p, q, j) \times W(j, n, m)\right\}, \tag{21}$$

where

$$\alpha(p, q, n) = e^{j\tau_n(\theta_p, r_q)}. \tag{22}$$

In Eq. (21), the fourth dimension of $\mathbf{S}$ is the results of the sensors. If the results in $\mathbf{S}$ are summed with proper weights, the estimation accuracy and the resolution of the SL are improved [36], [37]. However, obtaining the proper weights is difficult. Thus, the proposed Sp2net learns the weights by using the linear layer for the fourth dimension of $\mathbf{S}$. The linear layer ($f^L(\cdot)$) consists of one node for the weighted sum and calculates the spatial spectrum ($\mathbf{P}_s \in \mathbb{R}^{P\times Q\times M}$) for the azimuth- and range-grid points, which is as,

$$\mathbf{P}_s = f^L(\mathbf{S}). \tag{23}$$

In Eq. (23), the elements of $\mathbf{P}_s$ at the grid points where the near-field sources exist have small values because of the orthogonality between the signal and the noise spaces. The conventional MUSIC spectrum is obtained by taking the reciprocal of $\mathbf{P}_s$. However, if the activation function of the linear layer is set to $1/x$, the learning process may not be carried out since $1/x$ is not differentiable at $x = 0$. Thus, this paper proposes the custom activation function ($a(x)$) which is differentiable for all x and is given as,

$$a(x) = e^{-\mu x^2}, \tag{24}$$

where μ denotes a sharpening factor. $\mathbf{P}_a \in \mathbb{R}^{P\times Q\times M}$ is calculated by $a(\mathbf{P}_s)$.

The third dimension of $\mathbf{P}_a$ consists of the M noise space channels. The M channels are combined by the CNN to attain one final spatial spectrum. The CNN is composed of J Conv2Ds and linear layers in Fig. 5.

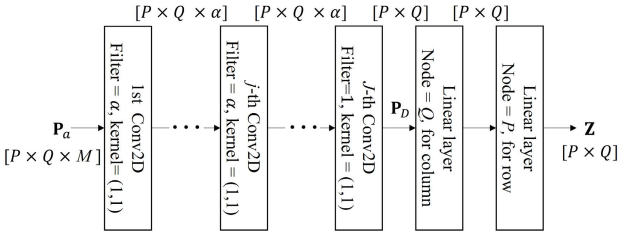


FIGURE 5. CNN structure in Sp2net.

The kernel size and the stride of the Conv2Ds are (1,1). The filter size is one for the last Conv2D, and the remainder is α. The activation functions of all Conv2Ds are the ReLU. The output ($\mathbf{P}_D \in \mathbb{R}^{P\times Q}$) of the Conv2D is obtained by combining all channels.

For increasing the number of features of the proposed network, the linear layers are applied to the rows and the columns of $\mathbf{P}_D$. Note that the rows and the columns of $\mathbf{P}_D$ are mapped to the azimuth classes (P) and the range classes (Q), respectively. The columns of $\mathbf{P}_D$ pass through the linear layer with Q nodes and the ReLU, and the rows of the output of the linear layer pass through the next linear layer with P nodes. To obtain the final output ($\mathbf{Z} \in \mathbb{R}^{P\times Q}$) of the proposed

network, a softmax is used as the activation function for the classification. The output of the Sp2net ($\mathbf{Z}$) is the estimation result for the DOAs and the ranges of the near-field sources. The labels of $\mathbf{Z}$ for the azimuth- and range-grid points are given as,

$$\mathbf{Z} = \begin{bmatrix} z(\theta_1, r_1) & \cdots & z(\theta_1, r_q) & \cdots & z(\theta_1, r_Q) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ z(\theta_p, r_1) & \cdots & z(\theta_p, r_q) & \cdots & z(\theta_p, r_Q) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ z(\theta_P, r_1) & \cdots & z(\theta_P, r_q) & \cdots & z(\theta_P, r_Q) \end{bmatrix}, \tag{25}$$

Therefore, the proposed PCAnet has lower complexity than the conventional EVD and calculates the noise spaces without a priori knowledge of the number of sources. The proposed Sp2net classifies many DOA and range classes by using the CNN and the proposed activation function and estimates both the DOAs and the ranges of the near-field sources with the high resolution.

IV. COMPUTER SIMULATIONS

In this section, the SL estimation performance and the computational complexity of the proposed method are analyzed by the computer simulations, which have been widely utilized for the evaluation of the conventional DLSL methods [26], [38], [39]. For the comparisons with the proposed method, the RD-MUSIC [12], the LCLA [14], the FOC-based root propagator (FOC-RP) [15], and the Cramer-Rao bound (CRB) [3], [40], [41] were tested.

A. TRAINING

For training and evaluating the proposed network, a fixed ULA passive sonar was assumed. Since the distances between the near-field source and each array element were different, $g_{n,l}$ in Eq. (1) was set according to the underwater transmission losses (TL) with Thorp's model [42], [43], whose dB scale model is given as,

$$TL = k_{TL} 10\log_{10} l_{TL} + \frac{l_{TL}}{10^3} 10\log_{10} a(f), \tag{26}$$

where k_{TL}, l_{TL} and f denote a spreading factor that describes the propagation geometry, a distance (m), and frequency (kHz), respectively. $a(f)$ denotes the absorption coefficient, which is expressed as,

$$10\log_{10} a(f) = 0.11\frac{f^2}{1+f^2} + \frac{44f^2}{4100+f^2} + 2.75\times 10^{-4} f^2 + 0.003. \tag{27}$$

For generating the training data set, k_{TL} was set to 1.5 for the practical spreading model.

Dataset configurations are demonstrated in Table 1. The number of the fixed ULA elements was set to $N = 50$ to cover a wide underwater area. To avoid overfitting problems, the training data set was generated with random parameters and the random initial phases of the near- and far-field signals. The number of sources was uniformly randomized from one to five for the near-field sources and from zero to five for the far-field sources. The grids for the ranges were set from 7.5 m ($\lambda/4$) to 600 m (20λ) with a grid interval of 7.5 m ($\lambda/4$). The numbers of DOA and range classes (P and Q in Fig. 4) were 180 and 80, respectively. Near-field sources randomly lay on the grid points with the uniform distribution, and the DOAs of the far-field sources were randomly set. The near-field signal strengths were set to be stronger than those of the far-field sources from 0 dB to 10 dB. The SNRs of the near-field signals were set randomly from -20 dB to 20 dB with a uniform distribution. The covariance matrix of the received signals with 200 snapshots was utilized for training as the input of the proposed network.

TABLE 1. Dataset configurations.

Parameters	Values
The number of the near-field sources	$1 \sim 5$
The number of the far-field sources	$0 \sim 5$
Grids for the ranges	7.5 m $\sim$ 600 m
Grid interval of the range	7.5 m
Grids for the azimuths	$-90^\circ \sim 89^\circ$
Grid interval of the azimuth	1°
SNR of the near-field signal	-20 dB $\sim$ 20 dB
SNR of the far-field signal	-30 dB $\sim$ 20 dB
The number of snapshots	200
The number of the training set	360,000
The number of the validation set	120,000

The output of the proposed network was a one-hot encoded label in Eq. (25) that assigns '1' to the grids where the near-field sources are located and allocates '0' to other grids. To classify the one-hot encoded label, the loss function of the proposed network used categorical cross-entropy [44]. The parameters of the proposed network are shown in Table 2.

TABLE 2. Network parameters.

Parameters	Values
Block size of the PCAnet (B)	10
Sliding interval of the SW-GS (s)	8
β of LeakyReLU	0.2
The number of noise space channels (M)	10
μ of Eq. (24)	128
α of Fig. 5	4
J of Fig. 5	3
The number of DOA classes (P)	180
The number of range classes (Q)	80

B. EVALUATION OF PERFORMANCES

The SL estimation performance of the proposed method was analyzed according to various SNRs and the number of snapshots. The estimation performance was measured by the root mean square error (RMSE) from 100,000 Monte-Carlo trials with the test data set, which was independent of the training data set. The near-field sources were located at $(46^\circ, 0.75\lambda)$, $(14^\circ, 9.5\lambda)$, and $(-34^\circ, 19\lambda)$. The far-field sources were located at $(5^\circ, \infty)$ and $(55^\circ, \infty)$. The example of the SL result of the proposed method is displayed in Fig. 6. In Fig. 6, the SNRs of the near-field sources were set to 0 dB, and the SNRs of the far-field sources were set to -5 dB.

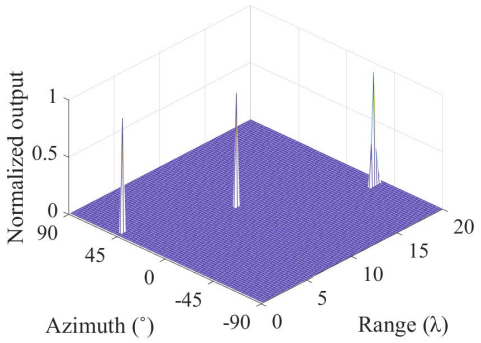


FIGURE 6. Source localization result of the proposed method.

The proposed method precisely estimated the DOAs and the ranges of all near-field sources.

Figure 7 demonstrates the DOA estimation performances of the tested schemes. A solid block line denotes the CRB. A blue triangle, green circle, magenta square, and red diamond markers denote the RD-MUSIC, FOC-RP, LCLA, and proposed method, respectively. Figures 7(a) and 7(b) show the RMSEs according to various SNRs of the near-field sources and the number of snapshots, respectively. The locations of the near- and far-field sources were the same as in Fig. 6. In Fig. 7(a), the number of snapshots was set to 200. Since the FOC-RP and the LCLA dependently estimated two parameters (DOAs and ranges), the estimation errors were propagated to each parameter in low SNR regions and the RMSEs did not monotonically decrease. The proposed method also had large RMSEs in lower SNRs of -8 dB. However, the proposed method had the smallest RMSEs in all SNR regions. At the SNR $=$ 10 dB, the RMSE of the proposed method had 22, 15, and five times better than that of the RD-MUSIC, the LCLA, and the FOC-RP, respectively. In Fig. 7(b), the SNR was set to -5 dB. The RMSEs of the proposed method were the smallest among the tested methods for all snapshot scenarios. When the number of snapshots was 500, the proposed method had six, four and three times better RMSEs than the RD-MUSIC, the FOC-RP, and the LCLA, respectively.

Figure 8 shows the range estimation performances of the tested schemes. The legends were the same as in Fig. 7. Figures 8(a) and 8(b) display the RMSEs according to various SNRs and the number of snapshots, respectively. In Fig. 8(a), the number of snapshots was set to 200. The proposed method had the smallest RMSEs among the tested schemes. At the SNR $=$ -10 dB, the RMSE of the proposed method had 80, 24, and four times better than that of the LCLA, the FOC-RP, and the RD-MUSIC, respectively. In Fig. 8(b), the SNR was set to -5 dB, and the numbers of snapshots were from 100 to 2000. The RMSEs of the proposed method also were the smallest among the tested methods for all snapshot scenarios. When the number of snapshots was 500, the proposed method had six, four, and three times better RMSEs than the RD-MUSIC, the FOC-RP, and the LCLA, respectively.

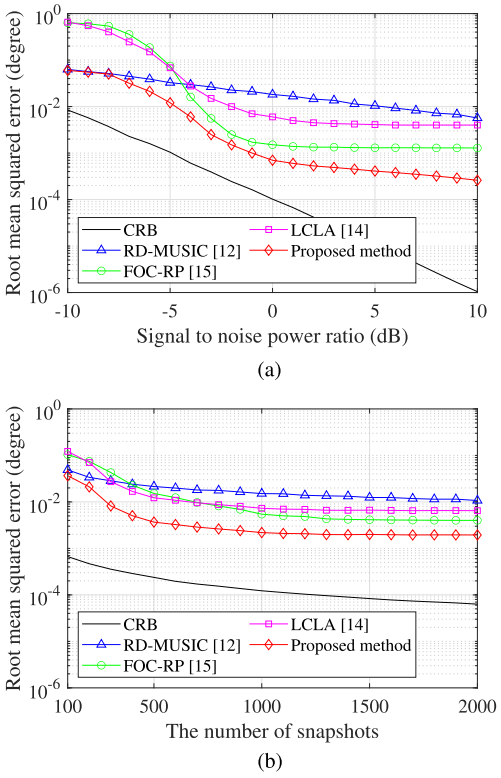


FIGURE 7. DOA estimation performances of (a) RMSE with varying SNR, (b) RMSE with varying the number of snapshots.

The computer simulations demonstrated that the proposed high resolution grid-based DLSL with the proposed activation function had better SL estimation performance than the conventional methods.

C. COMPLEXITY COMPARISONS

In this subsection, the computational complexities of the tested methods were compared by using the theoretical analysis and the average operation time. Table 3 demonstrates the theoretical complexities of the tested methods. Table 4 and Fig. 9 exhibit the average operation time and the complexities based on Table 3, respectively, according to 200 and 500 snapshots. Other parameters were the same as in Tables 1 and 2. For the fair comparisons of the average operation time, the AMD Ryzen 7 3700X CPU was utilized, and the GPU was not utilized.

In Table 4, the LCLA had the smallest average operation time among the compared methods, but the LCLA had larger RMSEs than the proposed method in Figs. 7 and 8. For 500 snapshots, the average operation time of the LCLA was

TABLE 3. Comparison of computational complexity.

Scheme	Complexity
Proposed method	$O((B+E+K)N^2+PQMN$ $+PQ^2+QP^2)$
RD-MUSIC	$O\left(\left[(N+L)(3N+1)/4+(N+1)^2/8\right]\right.$ $\left.\times P(N+1)+N^3+KN^2\right)$
LALC	$O\left(KN^2+2N\log_2 N+NLQ\right)$
FOC-RP	$O\left((9K+2)N^2+5N^3\right)$

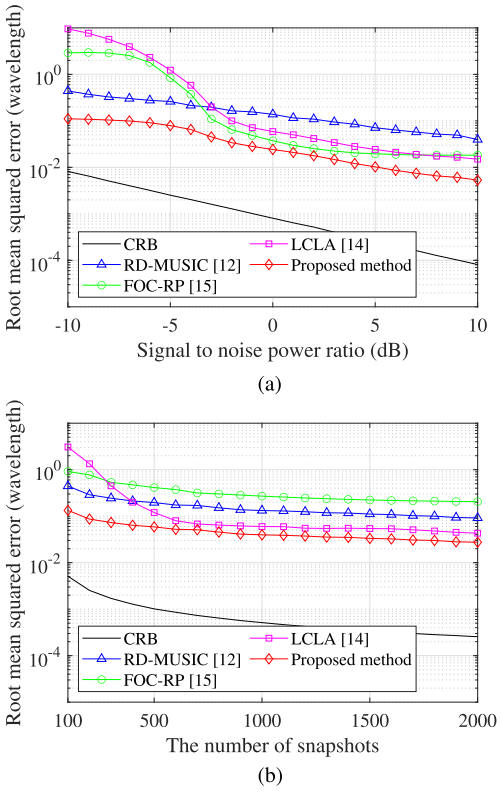


FIGURE 8. Range estimation performances of (a) RMSE with varying SNR, (b) RMSE with varying the number of snapshots.

TABLE 4. Comparison of average operation time.

Scheme	Snapshots	Average operation time (msec)
Proposed method	200	9.5
	500	10.1
RD-MUSIC	200	20.2
	500	18.5
LCLA	200	4.1
	500	3.5
FOC-RP	200	7.8
	500	25.5

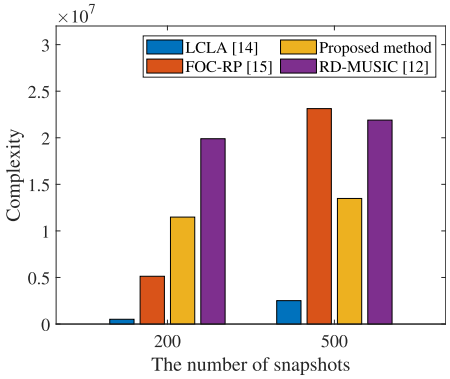


FIGURE 9. Comparison of computational complexities according to the number of snapshots.

2.5 times faster than the proposed method, but the proposed method had 80 times better RMSE performance than the LCLA in Fig. 8(a).

In Fig. 9, the complexity of the FOC-RP was lower than that of the proposed method at 200 snapshots but larger than that of the proposed method at 500 snapshots. These results were well matched with Table 4. Note that the proposed method had better RMSEs than the FOC-RP for all snapshots.

Thus, the proposed method demonstrated the lower complexity and RMSE than other conventional schemes.

V. CONCLUSION

This paper proposed the high resolution grid-based DLSL that precisely estimates the locations of near-field sources by classifying many source location classes. The proposed PCAnet estimated the noise space without a priori knowledge of the number of sources. The proposed Sp2net improved the SL estimation performance by classifying dense grids with the proposed activation function. The computer simulation showed that the estimation performances for the DOAs and the ranges of the proposed method were better than those of the conventional methods.

REFERENCES

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.

[2] L. Kumar and R. M. Hegde, "Near-field acoustic source localization and beamforming in spherical harmonics domain," *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3351–3361, Jul. 2016.

[3] E. Grosicki, K. Abed-Meraim, and Y. Hua, "A weighted linear prediction method for near-field source localization," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3651–3660, Oct. 2005.

[4] H. Lee, J. Ahn, Y. Kim, and J. Chung, "Direction-of-arrival estimation of far-field sources under near-field interferences in passive sonar array," *IEEE Access*, vol. 9, pp. 28413–28420, 2021.

[5] E. Fisher and B. Rafaely, "Near-field spherical microphone array processing with radial filtering," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 2, pp. 256–265, Feb. 2011.

[6] H.-J. Lee, J.-M. Ahn, J.-P. Seo, J.-K. Ahn, S.-I. Kim, and J.-H. Chung, "Left right discrimination performance improvement for the line array sonar system," *J. Acoust. Soc. Korea*, vol. 36, no. 1, pp. 49–56, Jan. 2017.

[7] G. Liu and X. Sun, "Spatial differencing method for mixed far-field and near-field sources localization," *IEEE Signal Process. Lett.*, vol. 21, no. 11, pp. 1331–1335, Nov. 2014.

[8] J. Liang and D. Liu, "Passive localization of mixed near-field and far-field sources using two-stage MUSIC algorithm," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 108–120, Jan. 2010.

[9] X. Li, S. Yan, X. Ma, and C. Hou, "Spherical harmonics MUSIC versus conventional MUSIC," *Appl. Acoust.*, vol. 72, no. 9, pp. 646–652, Sep. 2011.

[10] Y. D. Huang and M. Barkat, "Near-field multiple source localization by passive sensor array," *IEEE Trans. Antennas Propag.*, vol. 39, no. 7, pp. 968–975, Jul. 1991.

[11] J. He, M. N. S. Swamy, and M. O. Ahmad, "Efficient application of MUSIC algorithm under the coexistence of far-field and near-field sources," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 2066–2070, Apr. 2012.

[12] X. Zhang, W. Chen, W. Zheng, Z. Xia, and Y. Wang, "Localization of near-field sources: A reduced-dimension MUSIC algorithm," *IEEE Commun. Lett.*, vol. 22, no. 7, pp. 1422–1425, Jul. 2018.

[13] Z. Huang, W. Wang, F. Dong, and D. Wang, "A one-snapshot localization algorithm for mixed far-field and near-field sources," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1010–1014, May 2020.

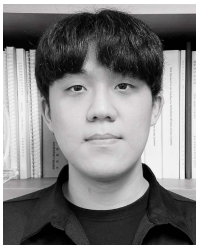
[14] Z. Huang, B. Xue, W. Wang, F. Dong, and D. Wang, "A low complexity localization algorithm for mixed far-field and near-field sources," *IEEE Commun. Lett.*, vol. 25, no. 12, pp. 3838–3842, Dec. 2021.

[15] J. Li, Y. Wang, Z. Ren, X. Gu, M. Yin, and Z. Wu, "DOA and range estimation using a uniform linear antenna array without a priori knowledge of the source number," *IEEE Trans. Antennas Propag.*, vol. 69, no. 5, pp. 2929–2939, May 2021.

[16] J. Xie, H. Tao, X. Rao, and J. Su, "Passive localization of mixed far-field and near-field sources without estimating the number of sources," *Sensors*, vol. 15, no. 2, pp. 3834–3853, Feb. 2015.
[17] W. Liu, Y. Yang, M. Xu, L. Lü, Z. Liu, and Y. Shi, "Source localization in the deep ocean using a convolutional neural network," *J. Acoust. Soc. Amer.*, vol. 147, no. 4, pp. EL314–EL319, Apr. 2020.
[18] S. Chakrabarty and E. A. P. Habets, "Multi-speaker DOA estimation using deep convolutional networks trained with noise signals," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 1, pp. 8–21, Mar. 2019.
[19] J. Cong, X. Wang, M. Huang, and L. Wan, "Robust DOA estimation method for MIMO radar via deep neural networks," *IEEE Sensors J.*, vol. 21, no. 6, pp. 7498–7507, Mar. 2021.
[20] Z.-M. Liu, C. Zhang, and P. S. Yu, "Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 7315–7327, Dec. 2018.
[21] X. Xiao, S. Zhao, X. Zhong, D. L. Jones, E. S. Chng, and H. Li, "A learning-based approach to direction of arrival estimation in noisy and reverberant environments," in *Proc. IEEE ICASSP*, South Brisbane, QLD, Australia, Apr. 2015, pp. 2814–2818.
[22] S. Chakrabarty and E. A. P. Habets, "Broadband DOA estimation using convolutional neural networks trained with noise signals," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust. (WASPAA)*, New Paltz, NY, USA, Oct. 2017, pp. 136–140.
[23] Q. Li, X. Zhang, and H. Li, "Online direction of arrival estimation based on deep learning," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Calgary, AB, Canada, Apr. 2018, pp. 2616–2620.
[24] S. Adavanne, A. Politis, and T. Virtanen, "Direction of arrival estimation for multiple sound sources using convolutional recurrent neural network," in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Rome, Italy, Sep. 2018, pp. 1462–1466.
[25] S. Y. Lee, J. Chang, and S. Lee, "Deep learning-based method for multiple sound source localization with high resolution and accuracy," *Mech. Syst. Signal Process.*, vol. 161, no. 1, Dec. 2021, Art. no. 107959.
[26] W. Liu, J. Xin, W. Zuo, J. Li, N. Zheng, and A. Sano, "Deep learning based localization of near-field sources with exact spherical wavefront model," in *Proc. 27th Eur. Signal Process. Conf. (EUSIPCO)*, A Coruna, Spain, Sep. 2019, pp. 1–5.
[27] A. Kamilaris and F. X. Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agricult.*, vol. 147, no. 1, pp. 70–90, Apr. 2018.
[28] J. He, L. Li, and T. Shu, "Near-field parameter estimation for polarized source using spatial amplitude ratio," *IEEE Commun. Lett.*, vol. 24, no. 9, pp. 1961–1965, Sep. 2020.
[29] J. He, L. Li, and T. Shu, "Localization of near-field sources for exact source-sensor spatial geometry," *IEEE Signal Process. Lett.*, vol. 27, pp. 1040–1044, 2020.
[30] J. He, L. Li, T. Shu, and T.-K. Truong, "Mixed near-field and far-field source localization based on exact spatial propagation geometry," *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3540–3551, Apr. 2021.
[31] T. Shu, J. He, and L. Li, "Near-field passive localization and gain–phase compensation with partly calibrated arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 1, pp. 712–719, Feb. 2022.
[32] J. Ahn, H. Lee, Y. Kim, W. Kim, and J. Chung, "Machine learning based biomimetic underwater covert acoustic communication method using dolphin whistle contours," *Sensors*, vol. 20, no. 21, p. 6166, Oct. 2020.
[33] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.
[34] Y. Kim, H. Lee, J. Ahn, and J. Chung, "Selection of CDMA and OFDM using machine learning in underwater wireless networks," *ICT Exp.*, vol. 5, no. 4, pp. 215–218, Dec. 2019.
[35] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. ICML*, Atlanta, GA, USA, 2013, pp. 1–6.
[36] S. A. S. Karam, D. O'Loughlin, B. L. Oliveira, M. O'Halloran, and B. M. Asl, "Weighted delay-and-sum beamformer for breast cancer detection using microwave imaging," *Measurement*, vol. 177, no. 1, Jun. 2021, Art. no. 109283.
[37] S. S. Ali, C. Liu, J. Liu, M. Jin, S.-J. Yoo, and J. M. Kim, "A blind weighted MUSIC based detection for spatial spectrum sensing," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Jeju, South Korea, Oct. 2017, pp. 480–485.
[38] Y. Yao, H. Lei, and W. He, "A-CRNN-based method for coherent DOA estimation with unknown source number," *Sensors*, vol. 20, no. 8, p. 2296, Apr. 2020.
[39] X. Su, P. Hu, Z. Liu, T. Liu, B. Peng, and X. Li, "Mixed near-field and far-field source localization based on convolution neural networks via symmetric nested array," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 7908–7920, Aug. 2021.
[40] P. Stoica, E. G. Larsson, and A. B. Gershman, "The stochastic CRB for array processing: A textbook derivation," *IEEE Signal Process. Lett.*, vol. 8, no. 5, pp. 148–150, May 2001.
[41] Y. Begriche, M. Thameri, and K. Abed-Meraim, "Exact Cramér–Rao bound for near field source localization," in *Proc. 11th Int. Conf. Inf. Sci., Signal Process. Their Appl. (ISSPA)*, Montreal, QC, Canada, Jul. 2012, pp. 718–721.
[42] J. M. Jornet, M. Stojanovic, and M. Zorzi, "On joint frequency and power allocation in a cross-layer protocol for underwater acoustic networks," *IEEE J. Ocean. Eng.*, vol. 35, no. 4, pp. 936–947, Oct. 2010.
[43] M. Stojanovic, "On the relationship between capacity and distance in an underwater acoustic communication channel," in *Proc. 1st ACM Int. workshop Underwater Netw. (WUWNet)*, Los Angeles, CA, USA, 2006, pp. 34–43.
[44] Y. Ho and S. Wookey, "The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling," *IEEE Access*, vol. 8, pp. 4806–4813, 2020.

HOJUN LEE was born in Pohang, South Korea, in 1990. He received the B.S. and M.S. degrees in electronics engineering from Inha University, Incheon, South Korea, in 2015 and 2018, respectively, where he is currently pursuing the Ph.D. degree in electronics engineering. His research interests include sonar systems, array signal processing, underwater acoustic communications, and machine learning.

YONGCHEOL KIM was born in Jeju, South Korea, in 1995. He received the B.S. degree in electronics engineering from Inha University, Incheon, in 2018, where he is currently pursuing the Ph.D. degree in electronics engineering. His research interests include underwater acoustic communication systems, signal processing, estimation theory, and machine learning.



SEUNGHWAN SEOL was born in Seoul, South Korea, in 1995. He received the B.S. degree in electronics engineering from Inha University, Incheon, South Korea, in 2020, where he is currently pursuing the Ph.D. degree in electronics engineering. His research interests include underwater acoustic communications, wireless communications, signal processing, and machine learning.

JAEHAK CHUNG received the B.S. and M.S. degrees in electronics engineering from Yonsei University, Seoul, South Korea, in 1988 and 1990, respectively, and the Ph.D. degree from the University of Texas at Austin, in 2000.

He was a member of technical staff at the Samsung Advanced Institute of Technology (SAIT), Kyungkido, South Korea, from 2001 to 2005. He is currently a Professor with Inha University, Incheon, South Korea. His research interests include 6G, underwater acoustic communications, and machine learning.