

Research Article

Multisource DOA Estimation in Impulsive Noise Environments Using Convolutional Neural Networks

Dong Chen ^{1,2} and Young Hoon Joo ²

¹School of Electronic and Information Engineering, Jiujiang University, Jiujiang 332005, China

²School of IT Information and Control Engineering, Kunsan National University, Gunsan 54150, Republic of Korea

Correspondence should be addressed to Young Hoon Joo; yhjoo@kunsan.ac.kr

Received 13 October 2021; Accepted 14 February 2022; Published 4 March 2022

Academic Editor: Francisco Falcone

Copyright © 2022 Dong Chen and Young Hoon Joo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This work proposes an effective high-resolution multisource direction-of-arrival (DOA) estimation method in impulsive noise scenarios based on convolutional neural networks (CNNs). First of all, the array observation matrix is preprocessed and fed into a denoising network to suppress outliers and filter out impulsive noise. Secondly, the denoising network output is fed into a model order selection network to estimate the model order. Next, according to the estimation, the denoising network output is fed into a DOA subnetwork corresponding to the model order in a DOA network to estimate the DOA of each signal. Comprehensive simulations demonstrate that, in the presence of impulsive noise, the proposed method is effective and superior in accuracy and computation speed for multisource DOA estimation. Therefore, it is concluded that CNN can be well generalized for DOA estimation.

1. Introduction

In recent years, direction-of-arrival (DOA) estimation has been extensively applied in many fields, such as radar, sonar, electronic monitoring, and mobile communication [1–3]. It is the critical technology of array signal processing, and the purpose is to estimate the transmitter positions of the signals received by arrays. The representative conventional DOA estimation algorithms are the estimation of signal parameters via rotational invariance techniques (ESPRIT) [4, 5] and multiple signal classification (MUSIC) [6, 7]. ESPRIT uses two subarrays with translation invariance to realize DOA estimation. Although the calculation is less than that of MUSIC, its accuracy is not enough. MUSIC obtains the spatial spectrum by constructing orthogonal signal subspace and noise subspace, and the DOA can be estimated via spectral peak search. The smaller the search step, the higher the estimation accuracy, while the computational burden may increase. In addition, both the ESPRIT and MUSIC algorithms are developed in Gaussian noise. However, many noises and signals in applications are impulsive, such as

underwater acoustic signals, radar clutter, and artificial interference. In the presence of impulsive noise, the performance of algorithms represented by ESPRIT and MUSIC would be significantly degraded because of the non-convergence of array output second-order moments. To improve the robustness, FLOM [8], PFLOM [9], CRCO [10], COBU [11], and other algorithms, which effectively achieve DOA estimation in impulsive noise, have been developed successively. However, the high computational complexity may make these algorithms difficult to be applied in real-time.

Many DOA estimation methods based on neural networks have been developed in recent years to reduce the computational burden. References [12–15] using a convolutional neural network (CNN), references [16, 17] using a support vector regression (SVR), reference [18] using a residual network, reference [19] using a fully connected neural network (FNN), reference [20] using a long short-term memory network, and references [21, 22] using a radial basis function (RBF) achieve high accuracy DOA estimation. However, they can only be used in single-source scenarios,

which may be extremely limited in practical applications. Therefore, reference [23] using CNN, references [24, 25] using SVR, and references [26, 27] using RBF achieve multisource DOA estimation. Although high estimation accuracy is obtained, these networks are only suitable for fixed model order scenarios and may not adapt to a variable source number. To improve the robustness to model orders, reference [28] divides the region of interest into several subregions, and each subregion corresponds to an RBF. A multilayer perceptron is used to judge the presence of a single source in each subregion, and RBFs perform position estimations of subregions with a source. Likewise, reference [29] using a circularly fully convolutional network divides the region of interest into several subregions. In the first stage, a coarse DOA is estimated, and then a refined result is obtained in the second stage. Although references [28, 29] can adapt to a variable model order, multiple sources are not permitted to be presented simultaneously in a subregion, which will affect the resolution of DOA estimation. The above neural network-based methods have effectively overcome computational complexity and improved DOA estimation accuracy. Yet, it may be difficult for these neural network-based methods to meet both high-resolution and variable model order requirements. Furthermore, these methods do not consider impulsive noise scenarios, in which neural networks may fail to be fed into practical input features, resulting in performance degradation of neural networks.

Motivated by the above investigation, we propose a high-resolution DOA estimation method with a variable model order using CNNs in impulsive noise environments. A DOA estimation model is developed, which comprises four modules: the preprocessing, denoising network, model order selection (MOS) network, and DOA network. The preprocessing processes array observation matrices into input features for the denoising network. The MOS network input is from the denoising network, and the output is the model order estimation. The DOA network comprises several DOA subnetworks corresponding to different model orders. The subnetworks, which adopt the transfer learning strategy, are trained sequentially from the model order of 1. The estimation of the MOS network decides to feed the denoising network output into the DOA subnetwork corresponding to the model order, which outputs DOA estimation. A series of simulation experiments verify the effectiveness and superiority of the DOA estimation model.

The main contributions of the study are as follows: (1) a high-resolution DOA estimation method with a variable model order using CNNs is achieved; (2) the robustness of DOA estimation is improved in impulsive noise scenarios; (3) for DOA estimation, it is proved that employing transfer learning can reduce training data.

The remainder of this paper is organized as follows: Section 2 briefly introduces the impulsive noise and CNN, and the problem of interest is defined; Section 3 develops the novel DOA estimation model; the experimental results are presented in Section 4, and the study is summarized in Section 5.

Table 1 lists the main notations used in this study. Other notations follow the conventional expression unless otherwise specified.

TABLE 1: Main notations.

Notations	Descriptions
$ \cdot $	Modulus of a complex number
$\ \cdot\ $	Infinite norm of a vector
$\text{Diag}(\cdot)$	Diagonal matrix operator
$\text{Re}(\cdot)$	Real part operator
$\text{Im}(\cdot)$	Imaginary part operator
$E(\cdot)$	Expectation operator
$[A1: A2: A3]$	A set from A1 to A3 (A2 step)
$\ \cdot\ _1$	L1 norm of a matrix
$\ \cdot\ _F$	Frobenius norm of a matrix

2. Preliminary and Problem Formulation

2.1. Impulsive Noise. The a -stable distribution, a highly flexible tool, can model the impulsive noise. It is normally defined via the characteristic function containing four variables: location parameter a ($-\infty < a < +\infty$), symmetric parameter β ($-1 \leq \beta \leq 1$), dispersion coefficient γ ($\gamma > 0$), and characteristic exponent α ($0 < \alpha \leq 2$) [30]. a reflects the deviation of the probability density function (PDF) on the x -axis ($a = 0$ in the study). β indicates the PDF inclination to the left or right about a ($\beta = 0$ in the study). γ denotes the dispersion of samples ($\gamma = 1$ in the study). α signifies the thickness of the a -stable distribution tails. Figure 1 displays the PDF of impulsive noise with different α . The smaller the α , the stronger the impulses, and the more impulsive the outliers.

2.2. Convolutional Neural Network. Compared with RBF, SVR, and FNN, the weight sharing strategy of CNN effectively improves the problem of the enormous input feature dimension of neural networks. Usually, CNN comprises multiple convolutional layers, pooling layers, and fully connected layers [31]. As an essential part of CNN, the convolutional layer comprises convolution filters. The feature image, as the input of the next layer, can be obtained by convolution of the input image and convolution filters.

The convolution process is displayed in Figure 2. The input image size is 8×8 . Convolve with a 3×3 convolution filter with a stride of 1, and obtain an 8×8 feature image. For ensuring the equal size of the input and feature images, the periphery of the input image is filled with zeros. The convolution filter is sliding on the input image, and the element $a_{i,j}$ in row i and column j of the feature image can be calculated by

$$a_{i,j} = F \left(v + \sum_{I=i}^{i+2} \sum_{J=j}^{j+2} \omega_{I-i+1, J-j+1} a'_{I,J} \right), \quad (1)$$

where F indicates a nonlinear activation function. $\omega_{I-i+1, J-j+1}$ denotes the weight in row $I-i+1$ and column $J-j+1$ of the convolution filter, and v denotes the bias. $a'_{I,J}$ signifies the element in row I and column J of the padded input image. Typically, a convolution layer has multiple convolution filters. The number of feature images equals that of convolution filters. The channel number of a convolution filter equals that of the input image.

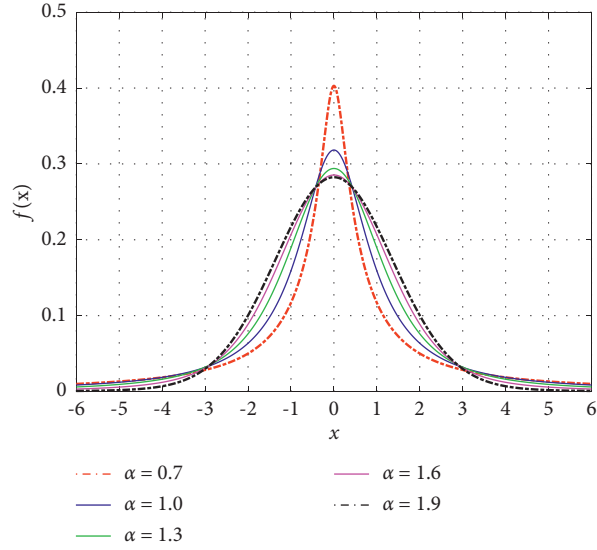


FIGURE 1: PDF of impulsive noise.

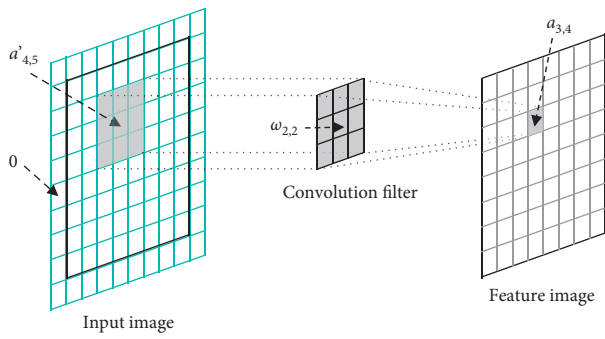


FIGURE 2: Convolution process diagram.

2.3. Problem Formulation. Figure 3 depicts p far-field stable electromagnetic signals impinge on a uniform linear array (ULA). The array element number is M , and the distance

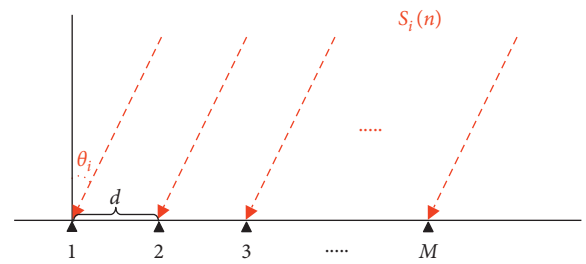


FIGURE 3: ULA model.

between two elements is d . The incident angle of the i -th signal $s_i(n)$ is θ_i ($0^\circ \leq \theta_i \leq 180^\circ$). The element denoted by 1 is selected as the phase reference point, and the $M \times p$ steering matrix can be expressed as [32]

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-j2\pi(df/c)\sin \theta_1} & e^{-j2\pi(df/c)\sin \theta_2} & \dots & e^{-j2\pi(df/c)\sin \theta_p} \\ \vdots & \vdots & \vdots & \vdots \\ e^{-j(M-1)2\pi(df/c)\sin \theta_1} & \dots & \dots & e^{-j(M-1)2\pi(df/c)\sin \theta_p} \end{bmatrix}, \quad (2)$$

where c denotes the propagation speed and f indicates the known carrier frequency. The $p \times N$ signal matrix can be expressed as

$$\mathbf{s} = \begin{bmatrix} s_1(1) & s_1(2) & \dots & s_1(N) \\ s_2(1) & s_2(2) & \dots & s_2(N) \\ \vdots & \vdots & \vdots & \vdots \\ s_p(1) & s_p(2) & \dots & s_p(N) \end{bmatrix}, \quad (3)$$

where N signifies the snapshot number. The $M \times N$ observation matrix is expressed as

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{e}_i \quad (4)$$

where \mathbf{e}_i signifies the $M \times N$ impulsive noise matrix and subscript i indicates impulse.

Obviously, θ_1 to θ_p can be obtained from \mathbf{x} , while the impulsive noise may disturb \mathbf{x} . The problem addressed in the study is establishing the mapping between \mathbf{x} and θ_1 to θ_p via the developed DOA estimation model.

3. DOA Estimation Model

In this section, we proposed the DOA estimation model, of which the input is x and the output is p DOA estimations, as shown in Figure 4. The model contains four modules: the preprocessing, denoising network, MOS network, and DOA network. x is processed by the preprocessing module and fed into the denoising network, of which the output is fed into the MOS and DOA networks. The DOA network comprises a set of DOAp ($1 \leq p \leq P$) subnetworks corresponding to different model orders. The estimation of the MOS network decides to feed the denoising network output into the DOAp subnetwork corresponding to the model order of p . For the convenience of training, the denoising network, MOS network, and each DOAp subnetwork are trained independently.

3.1. Preprocessing. The preprocessing input is x , and the output is the input feature of the denoising network. In the presence of impulsive noise, elements of x may be extreme outliers at some moments. First of all, x is normalized with the infinite norm to suppress impulsive outliers and facilitate the training of the denoising network. The definition of the infinite norm is

$$\|\mathbf{x}(n)\|_{\infty} = \max_{1 \leq i \leq M} |x_i(n)|, \quad (5)$$

where $x_i(n)$ indicates the n -th snapshot of the i -th array element. After normalizing x with the infinity norm, x expressed by (4) should be modified as

$$\mathbf{x}_{\infty} = \mathbf{x} \Lambda_{\infty}, \quad (6)$$

where $\Lambda_{\infty} = \text{diag}(1/\|\mathbf{x}(1)\|_{\infty}, 1/\|\mathbf{x}(2)\|_{\infty}, \dots, 1/\|\mathbf{x}(N)\|_{\infty})$. Secondly, \mathbf{x}_{∞} is used to construct the covariance matrix \mathbf{R}_{∞} to reduce the dimension of neural network input features. Considering that \mathbf{R}_{∞} is a Hermitian matrix and CNN requires input features to be real numbers, we construct the following $M \times M$ input feature \mathbf{F}_d for the denoising network:

$$\mathbf{F}_d = \begin{bmatrix} R_{\infty 11} & \text{Im}(R_{\infty 12}) & \cdots & \text{Im}(R_{\infty 1M}) \\ \text{Re}(R_{\infty 12}) & R_{\infty 22} & \cdots & \text{Im}(R_{\infty 2M}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Re}(R_{\infty M1}) & \text{Re}(R_{\infty M2}) & \cdots & R_{\infty MM} \end{bmatrix}, \quad (7)$$

where $R_{\infty jk}$ represents the element in row j and column k of \mathbf{R}_{∞} .

3.2. Denoising Network. The function of the denoising network is to eliminate the effect of noise and signal powers from \mathbf{F}_d on the MOS network and DOA network. Therefore, we must construct appropriate labels for \mathbf{F}_d to train the denoising network.

In the absence of impulsive noise, the observation matrix is expressed by \mathbf{x}_d , and (4) can be modified as

$$\mathbf{x}_d = \mathbf{A}\mathbf{s}, \quad (8)$$

\mathbf{x}_d is used to construct the covariance matrix \mathbf{R}_d , and the $M \times M$ label \mathbf{L}_d constructed with \mathbf{R}_d can be expressed as

$$\mathbf{L}_d = \begin{bmatrix} 0 & \text{Im}(\overline{R_{d12}}) & \cdots & \text{Im}(\overline{R_{d1M}}) \\ \text{Re}(\overline{R_{d21}}) & 0 & \cdots & \text{Im}(\overline{R_{d2M}}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Re}(\overline{R_{dM1}}) & \text{Re}(\overline{R_{dM2}}) & \cdots & 0 \end{bmatrix}, \quad (9)$$

where $\overline{R_{djk}} = R_{djk}/N_{\mathbf{R}_d}$. R_{djk} indicates the element in row j and column k of \mathbf{R}_d , and $N_{\mathbf{R}_d} = \|\mathbf{R}_{d11}\mathbf{R}_{d22} \cdots \mathbf{R}_{dMM}\|_1/M$, namely, the estimation of the signal power sum. And the principal diagonal elements of \mathbf{R}_d are replaced by zeros because they fail to contain angle information. Obviously, \mathbf{L}_d does not contain impulsive noise, and $N_{\mathbf{R}_d}$ normalizes the elements of \mathbf{R}_d to eliminate the effect of signal powers.

The cost function of the denoising network employs the mean squared error (MSE), which is expressed by MSEd as

$$\text{MSE}_d(\boldsymbol{\omega}, \mathbf{v}) = \frac{1}{M^2} E\left(\|\mathbf{L}_d - \widehat{\mathbf{L}}_d\|_F^2\right), \quad (10)$$

where $\widehat{\mathbf{L}}_d$ signifies the denoising network response. $\boldsymbol{\omega}$ and \mathbf{v} denote the denoising network weights and biases, respectively.

3.3. MOS Network. The MOS network input is $\widehat{\mathbf{L}}_d$ from the denoising network. The $P \times 1$ label \mathbf{L}_{mos} using one-hot [33] represents the model order, and P denotes the preset maximum model order. When the model order is p , the p -th element of \mathbf{L}_{mos} is 1, and other elements are zeros.

Because of the independent MOS network training, the input feature is designed separately for the MOS network. Considering that the previous denoising network may not wholly filter out noise, Gaussian noise is added to the MOS network training set to improve the robustness of the DOA estimation model. In this case, the observation matrix is expressed by \mathbf{x}_{mos} , and (4) is modified as

$$\mathbf{x}_{\text{mos}} = \mathbf{A}\mathbf{s} + \mathbf{e}_g, \quad (11)$$

where \mathbf{e}_g signifies the $M \times N$ Gaussian noise matrix and subscript g indicates Gauss. The covariance matrix \mathbf{R}_{mos} is constructed by \mathbf{x}_{mos} , and the noise power σ^2 can be estimated by eigenvalue decomposition. The input feature is expressed by \mathbf{F}_{mos} , and (9) can be modified as

$$\mathbf{F}_{\text{mos}} = \begin{bmatrix} 0 & \text{Im}(\overline{R_{\text{mos}12}}) & \cdots & \text{Im}(\overline{R_{\text{mos}1M}}) \\ \text{Re}(\overline{R_{\text{mos}21}}) & 0 & \cdots & \text{Im}(\overline{R_{\text{mos}2M}}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Re}(\overline{R_{\text{mos}M1}}) & \text{Re}(\overline{R_{\text{mos}M2}}) & \cdots & 0 \end{bmatrix}, \quad (12)$$

where $\overline{R_{\text{mos}jk}} = R_{\text{mos}jk}/N_{\mathbf{R}_{\text{mos}}}$. $R_{\text{mos}jk}$ signifies the element in row j and column k of \mathbf{R}_{mos} , and $N_{\mathbf{R}_{\text{mos}}} = (\|\mathbf{R}_{\text{mos}11}\mathbf{R}_{\text{mos}22} \cdots \mathbf{R}_{\text{mos}MM}\|_1 - \sigma^2)/M$, namely, the estimation of the signal power sum. Compared with \mathbf{L}_d , elements in \mathbf{F}_{mos} are disturbed by Gaussian noise. Training the MOS network with \mathbf{F}_{mos} can improve the network robustness.

The cost function of the MOS network employs cross-entropy, which is expressed as

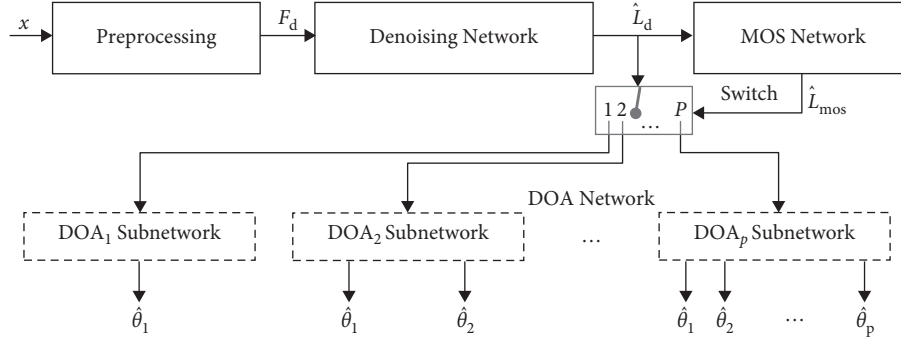


FIGURE 4: Block diagram of the DOA estimation model.

$$C(\boldsymbol{\omega}, \mathbf{v}) = -E \left(\sum_{i=1}^P \mathbf{L}_{\text{mos}i} \log \hat{\mathbf{L}}_{\text{mos}i} \right), \quad (13)$$

where $\hat{\mathbf{L}}_{\text{mos}i}$ and $\mathbf{L}_{\text{mos}i}$ represent the i -th element of the response and ground-truth label, respectively. $\boldsymbol{\omega}$ and \mathbf{v} denote the MOS network weights and biases, respectively.

3.4. DOA Network. The DOA network comprises a set of DOAp subnetworks ($1 \leq p \leq P$) corresponding to different model orders. The MOS network output activates the corresponding DOAp subnetwork. The input feature of each DOAp subnetwork is the denoising network output, and the response is the DOA estimation of p sources.

Theoretically, a single large-scale network can be trained to achieve multisource DOA estimation with a variable model order. However, this method is challenging to implement because of the limitation of computing power and the difficulty in obtaining a large number of training data to ensure the DOA estimation accuracy. Therefore, we adopt the strategy of training each DOAp subnetwork independently. Even so, to ensure the DOAp subnetwork estimation accuracy, the amount of training data required by conventional training methods also increases significantly with p . Given that DOA estimation of p sources and $p+1$ sources are two different but related tasks, we adopt parameter-transfer learning [34] to train DOAp subnetworks sequentially from $p=1$ to P , and training sets are generated separately according to (12). First of all, we construct a single source training set to train the DOA1 subnetwork. Secondly, based on the DOAp subnetwork architecture, convolution filters, convolution layers, or lengths of fully connected layers are increased, and then one output layer port is added. Finally, the DOAp+1 subnetwork architecture is developed. The parameters of the trained DOAp subnetwork architecture are used as initialization parameters for corresponding parts of the DOAp+1 subnetwork architecture, and other newly added parameters of the DOAp+1 subnetwork adopt conventional initialization methods. Thus, the training data required by the DOAp+1 network can be significantly reduced.

The cost function of each DOAp subnetwork adopts the MSE, which is expressed by MSEp as

$$\text{MSE}_p(\boldsymbol{\omega}, \mathbf{v}) = \frac{1}{P} E \left(\sum_{i=1}^P (\theta_i - \hat{\theta}_i)^2 \right) + \frac{\lambda}{2} E \left(\sum_{l=1}^L \|\boldsymbol{\omega}_l\|_F^2 \right), \quad (14)$$

where θ_i and $\hat{\theta}_i$ denote the ground-truth label and DOA estimation of the i -th signal, respectively. L signifies the layer number of the subnetwork, and λ is the regularization parameter. $\boldsymbol{\omega}$ and \mathbf{v} denote the DOAp subnetwork network weights and biases, respectively.

4. Simulation Results

Firstly, the denoising network performance is presented. Secondly, Section 4.2 shows the MOS network accuracy and compares it with a modified Akaike information criterion (AIC) method and a modified minimum description length (MDL) method. Finally, Section 4.3 demonstrates the DOA estimation model performance, and the MSE and computation speed are compared with those of CRCO and COBU.

Unless otherwise specified, the simulation conditions are as follows: (1) the ULA element number of M is 8, and the maximum model order P is set to 3; (2) the signal carrier frequency is 600 MHz, and the ULA element spacing is set to 0.48 times the wavelength; (3) the signal amplitude is randomly sampled to improve the DOA estimation model robustness for the signal amplitude [35], and the snapshot number is set to 2,000. The network initialization settings are as follows: (1) initialize the network weights with lecun_normal [36] and the biases with zeros; (2) the convolution filter is 3×3 , and the stride is 1; (3) the padding mode is "same"; (4) the mini-batch [37] is 512; (5) Adam [38] is adopted in the backpropagation. Also, in Gaussian noise scenarios, the signal-to-noise ratio (SNR) is defined as $\text{SNR} = 10 \lg(\sigma_s^2 / \sigma^2)$, where σ_s^2 denotes the signal power. In non-Gaussian noise scenarios, the ratio of the signal to noise dispersion, as generalized SNR, is defined as $\text{GSNR} = 10 \lg(\sigma_s^2 / \gamma)$ [11].

4.1. Performance of the Denoising Network. GSNR is set to 0 dB and 20 dB, respectively, and 30,000 samples are generated in $p \in [1:1:3]$, $\alpha \in [0.1:0.1:2.0]$, and $\theta \in [0^\circ:1^\circ:180^\circ]$. 60,000 samples constitute the data set, which is divided into the training set (90%) and validation set (10%). GSNR is set to 0 dB to enhance the network generalization ability in the

scenario of low GSNR, while 20 dB is set to prevent the network from overfitting. The learning rate gradually decreases with the epoch. Table 2 illustrates the denoising network architecture.

The relative error (RE) is defined as the evaluation criterion, which can be defined as

$$RE = E\left(\left\|\mathbf{L}_d - \hat{\mathbf{L}}_d\right\|_1 / \left\|\mathbf{L}_d\right\|_1\right). \quad (15)$$

Figure 5(a) depicts the variation of RE with GSNR of the denoising network in different characteristic exponents after preprocessing test data. $GSNR \in [-5 \text{ dB}: 1 \text{ dB}: 20 \text{ dB}]$, and 500 test samples are generated in $p \in [1: 1: 3]$ and $\theta \in [0^\circ: 1^\circ: 180^\circ]$ for each GSNR. Figure 5(b) depicts the variation of RE with the snapshot number in different characteristic exponents with GSNR of 10 dB. $N \in [100: 100: 2000]$, and 500 test samples are generated in $p \in [1: 1: 3]$ and $\theta \in [0^\circ: 1^\circ: 180^\circ]$ for each N. Figure 5 illustrates that increasing the GSNR, snapshot number, or characteristic exponent can improve the denoising network performance.

4.2. Performance of the MOS Network. The MOS network and denoising network are trained independently. Considering that the denoising network RE is challenging to be 0%, Gaussian noise is added during the MOS network training to enhance the model generalization ability. The SNR is set as -15 dB and 20 dB , respectively, and 20,000 samples are generated in $p \in [1: 1: 3]$ and $\theta \in [0^\circ: 1^\circ: 180^\circ]$. 40,000 samples constitute the data set, which is divided into the training set (85%) and validation set (15%). The learning rate gradually decreases with the epoch. Table 3 illustrates the MOS network architecture.

Accuracy N_C/N_T is defined as the evaluation criterion. NT is the total amount of test data, and NC is the amount of correctly estimated data. The test data is the same as that in Section 4.1, and the denoising network output is fed into the MOS network. Figure 6 demonstrates the relationship between the MOS network accuracy and SNR or snapshot number. When $\alpha < 0.4$, the MOS network performance is significantly degraded. The main reasons are as follows: (1) the amplitude of outliers is extreme, resulting in significant attenuation to sources after normalization with the infinite norm; (2) the angle step of random sampling is set to 0.01° , which may cause the angular distance of simultaneous incident sources to be very close, resulting in the inability of the MOS network to distinguish. However, when $\alpha > 0.4$, the MOS network accuracy approaches 1 and hardly changes with the SNR or snapshot number, indicating that the MOS network has high-resolution and better robustness.

To the best of our knowledge, no MOS algorithm has been published in the presence of impulsive noise. To validate that the MOS network is superior to conventional algebraic algorithms, we replace the covariance matrix in AIC and MDL [39] with R_{∞} in Section 3.1 and obtain the modified AIC and MDL. Figure 7 shows that the MOS network outperforms the modified AIC and MDL, and the test data is the same as that with $\alpha = 1.3$ in Figure 5. In addition, the processing time of the modified AIC and MDL is about 3 times that of the MOS network.

TABLE 2: Denoising network architecture.

Layer and activation	Output shape	Parameter
L0: Input	8×8	0
L1: Convolution + Relu	$64 \times 8 \times 8$	640
L2: Convolution + Relu	$32 \times 8 \times 8$	18464
L3: Convolution + Relu	$32 \times 8 \times 8$	9248
L4: Convolution + Relu	$32 \times 8 \times 8$	9248
L5: Convolution + Relu	$16 \times 8 \times 8$	4624
L6: Convolution	8×8	145

4.3. Performance of the DOA Estimation Model

4.3.1. DOA Network Architecture. Each subnetwork of the DOA network corresponds to a different model order and is trained independently. Firstly, the data set is constructed for the DOA1 subnetwork. The SNR is set as -15 dB and 20 dB , respectively, and 4,500 samples are generated in $\theta \in [0^\circ: 1^\circ: 180^\circ]$ with p of 1.9,000 samples constituting the data set, which is divided into the training set (80%) and validation set (20%). The learning rate gradually decreases with the epoch. Table 4 illustrates the DOA1 subnetwork architecture.

Secondly, The SNR is set as -15 dB and 20 dB , respectively, and 20,000 samples are generated in $\theta \in [0^\circ: 1^\circ: 180^\circ]$ with p of 2.40,000 samples constitute the data set, which is divided into the training set (90%) and validation set (10%). The learning rate gradually decreases with the epoch, and early stopping is adopted to prevent overfitting. Table 5 illustrates the DOA2 subnetwork architecture obtained by expanding the corresponding layer of the DOA1 subnetwork.

Then, the SNR is set as -15 dB and 20 dB , respectively, and 80,000 samples are generated in $\theta \in [0^\circ: 1^\circ: 180^\circ]$ with p of 3.160,000 samples constituting the data set, which is divided into the training set (85%) and validation set (15%). The learning rate gradually decreases with the epoch. Early stopping and regularization are adopted to avoid overfitting. Table 6 illustrates the DOA3 subnetwork architecture. L1 to L4, L7, and L8 of the DOA3 subnetwork are expanded from L1 to L4, L6, and L7 of the DOA2 subnetwork in turn, and L5 is a new layer.

4.3.2. Performance of the Model. The MSE is defined as the evaluation criterion, which can be expressed as

$$MSE = E\left(\left[\frac{1}{p} \sum_{i=1}^p (\theta_i - \hat{\theta}_i)^2\right]\right), \quad (16)$$

where p represents a variable model order. After the test data is preprocessed, the MOS network decides to feed the denoising network output into the corresponding subnetwork in the DOA network. Figure 8(a) shows the variation of the MSE with GSNR of the DOA estimation model in different characteristic exponents. $GSNR \in [-5 \text{ dB}: 1 \text{ dB}: 20 \text{ dB}]$, and 500 test samples are generated in $p \in [1: 1: 3]$ and $\theta \in [0^\circ: 2^\circ: 180^\circ]$ for each GSNR. Figure 8(b) shows the variation of the MSE with snapshot number in different characteristic exponents with GSNR of 10 dB. $N \in [100: 100:$

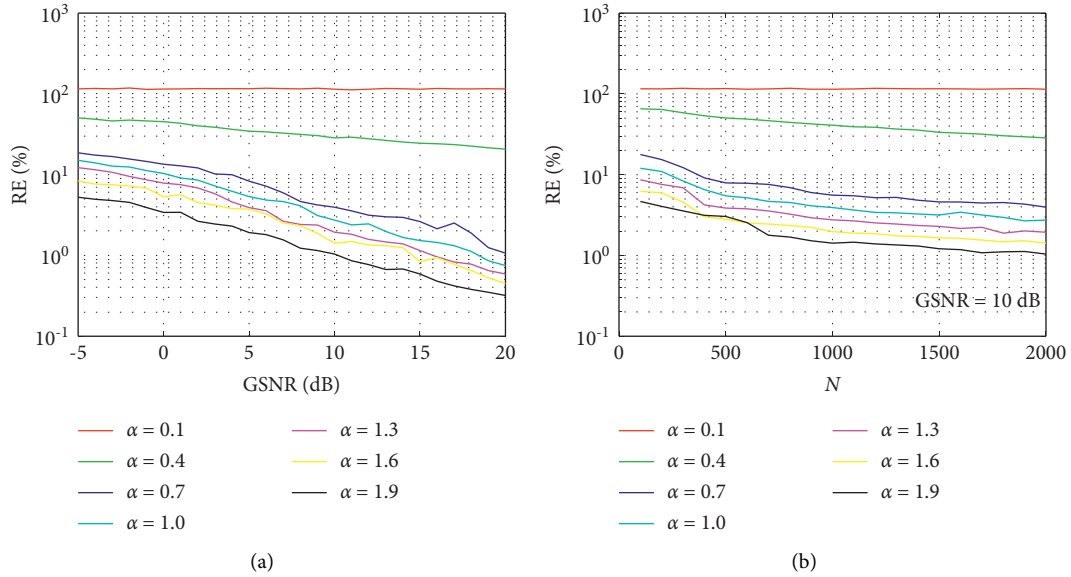


FIGURE 5: Performance of the denoising network. (a) RE versus GSNR. (b) RE versus snapshot number.

TABLE 3: MOS network architecture.

Layer and activation	Output shape	Parameter
L0: Input	8×8	0
L1: Convolution + Relu	64×8×8	640
L2: Convolution + Relu	32×8×8	18464
L3: Convolution + Relu	16×8×8	4624
L4: Convolution + Relu	8×8×8	1160
L5: Flatten	512	0
L6: Dense + Relu	16	8208
L7: Dense + Softmax	3	51

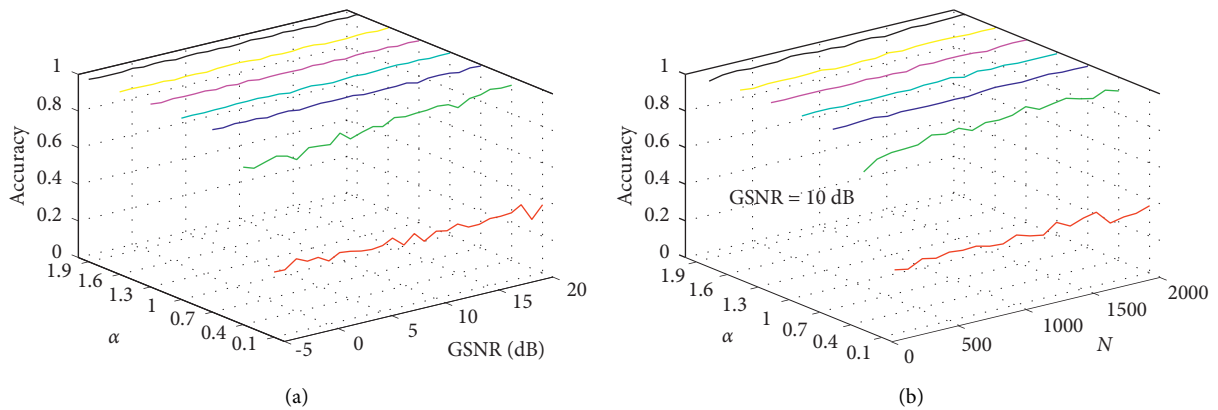


FIGURE 6: Performance of the MOS network. (a) Accuracy versus GSNR. (b) Accuracy versus snapshot number.

2000], and 500 test samples are generated in $p \in [1 : 1 : 3]$ and $\theta \in [0^\circ : 2^\circ : 180^\circ]$ for each N . Figure 8 suggests that increasing the GSNR, snapshot number, or characteristic exponent can improve the model performance. Still, DOA estimation is a challenging task in small characteristic exponent scenarios. Therefore, Figure 8 does not display the DOA estimation model performance when $\alpha < 0.7$.

To highlight the superiority of the DOA estimation model, the model is compared with COBU and CRCO, which have been proved to be superior to classic FLOM and PFLM in references [10, 11]. The search steps of COBU and CRCO are set as 0.01° . Both the weight factor and kernel size of COBU are set as 1. The scale factor and parameter μ of CRCO are set as 1.4 and 0.5, respectively. Other simulation

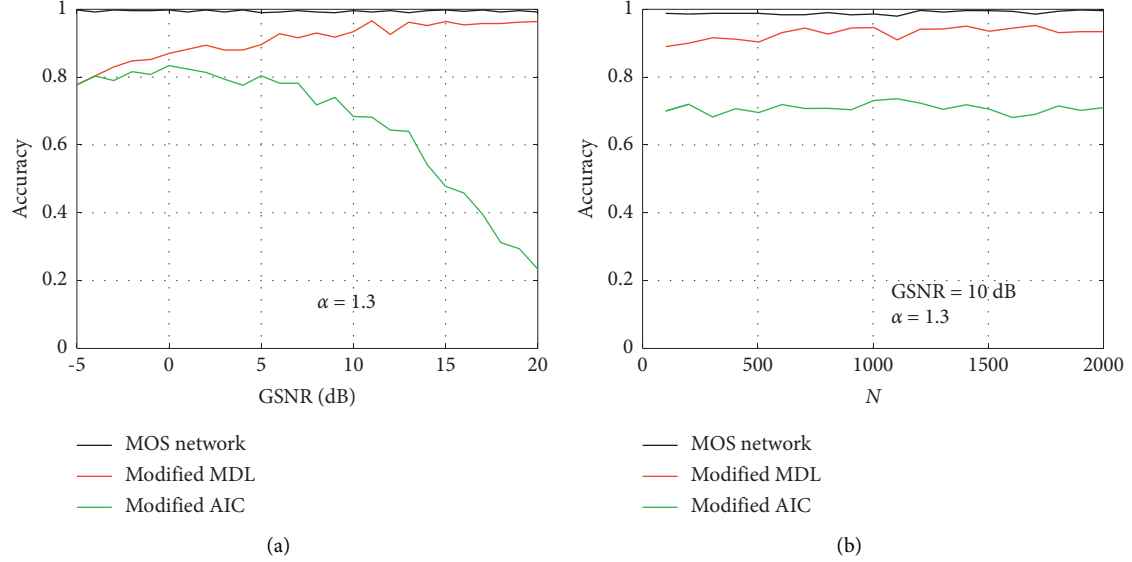


FIGURE 7: Performance comparison of the MOS network, modified AIC, and modified MDL. (a) Accuracy versus GSNR. (b) Accuracy versus snapshot number.

TABLE 4: DOA1 subnetwork architecture.

Layer and activation		Output shape	Parameter
L0:	Input	8×8	0
L1:	Convolution + Relu	$32 \times 8 \times 8$	320
L2:	Convolution + Relu	$32 \times 8 \times 8$	9248
L3:	Convolution + Relu	$16 \times 8 \times 8$	4624
L4:	Convolution + Relu	$16 \times 8 \times 8$	2320
L5:	Flatten	1024	0
L6:	Dense + Relu	8	820
L7:	Dense + Relu	1	9

TABLE 5: DOA2 subnetwork architecture.

Layer and activation		Output shape	Parameter
L0:	Input	8×8	0
L1:	Convolution + Relu	$64 \times 8 \times 8$	640
L2:	Convolution + Relu	$64 \times 8 \times 8$	36928
L3:	Convolution + Relu	$32 \times 8 \times 8$	18464
L4:	Convolution + Relu	$16 \times 8 \times 8$	4624
L5:	Flatten	1024	0
L6:	Dense + Relu	10	10250
L7:	Dense + Relu	2	22

TABLE 6: DOA3 subnetwork architecture.

Layer and activation		Output shape	Parameter
L0:	Input	8×8	0
L1:	Convolution + Relu	$80 \times 8 \times 8$	800
L2:	Convolution + Relu	$80 \times 8 \times 8$	57680
L3:	Convolution + Relu	$48 \times 8 \times 8$	34608
L4:	Convolution + Relu	$48 \times 8 \times 8$	20784
L5:	Convolution + Relu	$16 \times 8 \times 8$	6928
L6:	Flatten	1024	0
L7:	Dense + Relu	12	12300
L8:	Dense + Relu	3	39

conditions are the same as those of the DOA estimation model. Assume that three signals impinge the ULA with the incident angles of 50.01° , 100.08° , and 120.05° , respectively. Based on 500 Monte Carlo runs, Figure 9 demonstrates the variation of the MSE with GSNR or snapshot number of each method with α of 1.3. $\text{GSNR} \in [-5 \text{ dB}; 1 \text{ dB}; 20 \text{ dB}]$ and $N \in [100:100:2000]$. Although COBU and CRCO are high-resolution algorithms based on spatial spectrum search, the estimation is still a discrete value limited by search step, and appropriate parameters should be set according to the characteristic exponent and GSNR. However, the DOA estimation model needs no prior knowledge about the

characteristic exponent and GSNR, and its output is a continuous value. Therefore, the DOA estimation model is superior to COBU and CRCO.

Furthermore, the processing time of the three methods is compared, and Table 7 displays the results from 500 Monte Carlo runs. The calculations are performed on a computer equipped with 16 GB RAM and Intel Core i7-9700K CPU. The processing time of the proposed model is mainly spent on preprocessing. If the snapshot number is slightly reduced, Figure 8(b) reveals that the DOA estimation model performance will not be significantly degraded, while the

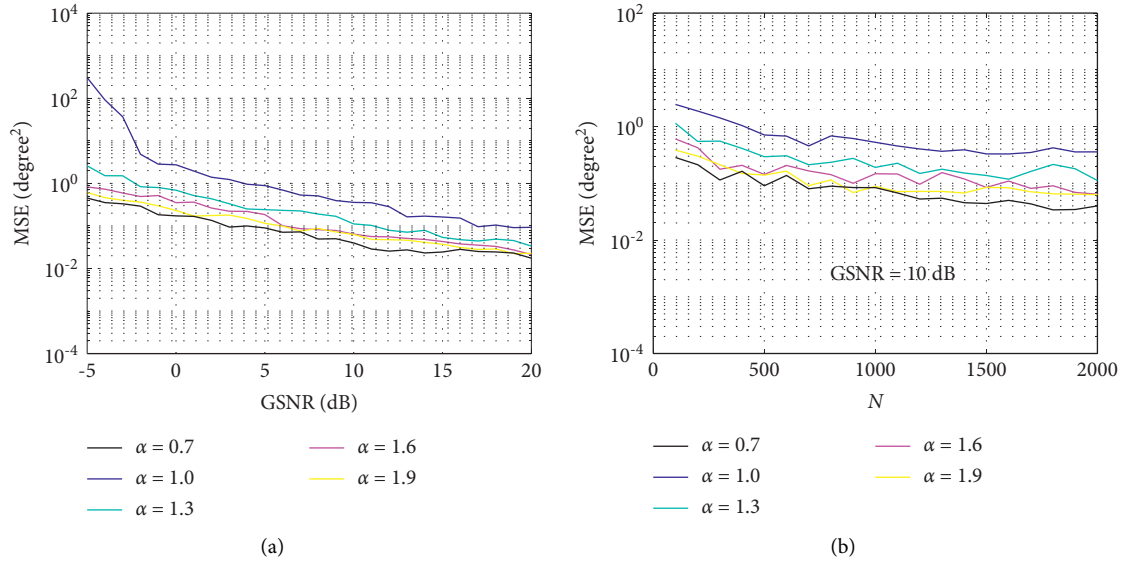


FIGURE 8: Performance of the DOA Estimation Model. (a) MSE versus GSNR. (b) MSE versus snapshot number.

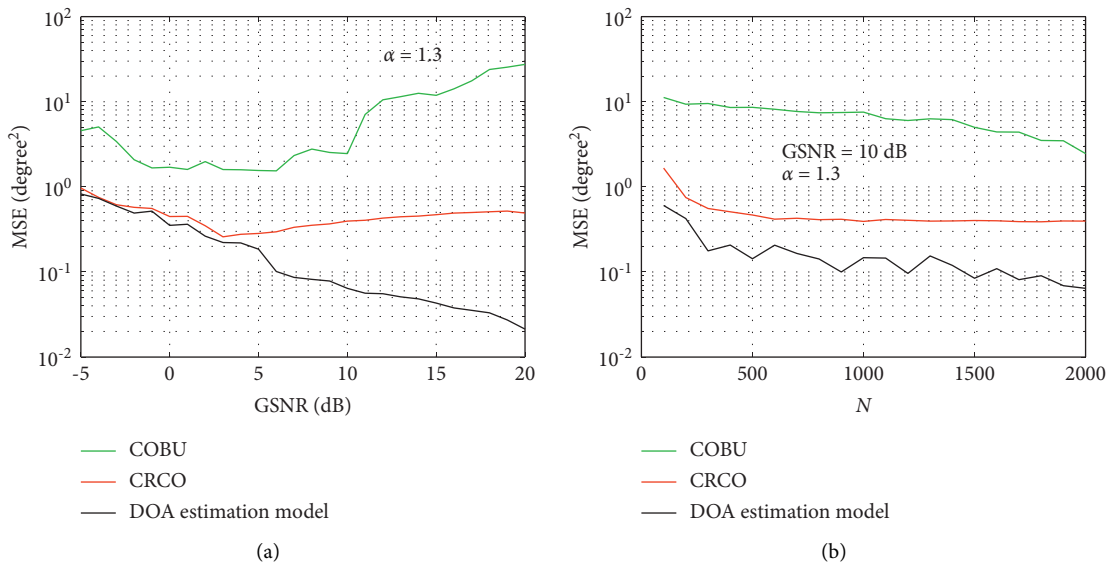


FIGURE 9: Performance comparison of the DOA estimation model, COBU, and CRCO. (a) MSE versus GSNR. (b) MSE versus snapshot number.

TABLE 7: Processing time of the methods.

Methods	Phases	Times (s)	Total time (s)
COBU	Construct spatial spectrum	58.2012	58.8681
	Spectral peak search	0.6669	
CRCO	Construct spatial spectrum	61.6728	62.3397
	Spectral peak search	0.6669	
Proposed method	Preprocessing	12.1932	12.2767
	Denoising network	0.0268	
	MOS network	0.0215	
	DOA network	0.0352	

processing time of preprocessing will be dramatically reduced. However, the processing time of COBU and CRCO will not be reduced considerably.

5. Conclusions

This study presented the DOA estimation model in impulsive noise environments. The model consists of the preprocessing, denoising network, MOS network, and DOA network. For the convenience of training, each network is trained independently. The preprocessing processes array observation matrices into appropriate input features for the denoising network. The DOA network consists of several DOA subnetworks corresponding to different model orders. The estimation of the MOS network decides to feed the denoising network output into the corresponding DOA subnetwork, which outputs DOA estimation.

The experiments reached the following conclusions: (1) in the presence of impulsive noise, the preprocessing and denoising network effectively suppress outliers and filter out impulsive noise; (2) the MOS network can estimate the model order with high accuracy; (3) the proposed DOA estimation model is effective and superior in accuracy and computation speed.

The problems that still need to be solved are as follows: (1) to develop DOA estimation in low characteristic exponent scenarios; (2) to extend the dimension of DOA estimation.

Data Availability

The data used to support the findings are available from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) funded by the Korea Government (MSIT) (NRF-2016R1A6A1A03013567 and NRF-2021R1A2B5B01001484).

References

- [1] X. Li and B. Wang, "Thinned virtual array for cramer rao bound optimization in MIMO radar," *International journal of Antennas and propagation*, vol. 2021, Article ID 1408498, 2021.
- [2] Q. Liu, J. Xu, Z. Ding, and H. C. So, "Target localization with jammer removal using frequency diverse array," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11685–11696, 2020.
- [3] Q. Yang, Y. Zhang, C. Tang, and J. Lian, "A combined antijamming and antispoofing algorithm for GPS arrays," *International journal of Antennas and propagation*, vol. 2019, Article ID 8012569, 2019.
- [4] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, & Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.
- [5] R. Wang, Y. Wang, Y. Li, W. Cao, and Y. Yan, "Geometric Algebra-Based ESPRIT Algorithm for DOA Estimation," *Sensors*, vol. 21, no. 17, 2021.
- [6] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.
- [7] W. K. Park, "Application of MUSIC algorithm in real-world microwave imaging of unknown anomalies from scattering matrix," *Mechanical Systems and Signal Processing*, vol. 153, 2021.
- [8] T. H. Liu and J. M. Mendel, "A subspace-based direction finding algorithm using fractional lower order statistics," *IEEE Transactions on Signal Processing*, vol. 49, no. 8, pp. 1605–1613, 2001.
- [9] H. Belkacemi and S. Marcos, "Robust subspace-based algorithms for joint angle/Doppler estimation in non-Gaussian clutter," *Signal Processing*, vol. 87, no. 7, pp. 1547–1558, 2007.
- [10] J. Zhang, T. Qiu, A. Song, and H. Tang, "A novel correntropy based DOA estimation algorithm in impulsive noise environments," *Signal Processing*, vol. 104, pp. 346–357, 2014.
- [11] Q. Tian, T. Qiu, J. Ma, J. Li, and R. Li, "A simplified DOA estimation method based on correntropy in the presence of impulsive noise," *IEEE Access*, vol. 6, pp. 67010–67025, 2018.
- [12] L. Zhang, C. Shi, J. Niu, Y. Ji, and Q. M. J. Wu, "DOA estimation for HFSWR target based on PSO-elm," *IEEE geoscience and remote sensing letters*, vol. 19, pp. 1–5, 2021.
- [13] W. Zhu, M. Zhang, C. Wu, and L. Zeng, "Broadband direction of arrival estimation based on convolutional neural network," *IEICE transactions on communications*, vol. E103, no. 3, pp. 148–154, 2019.
- [14] X. Lyu and J. Wang, "Direction of arrival estimation in passive radar based on deep neural network," *IET Signal Processing*, vol. 15, pp. 612–621, 2021.
- [15] Y. Liu, H. Chen, and B. Wang, "DOA estimation based on CNN for underwater acoustic array," *Applied Acoustics*, vol. 172, 2021.
- [16] L. L. Wu and Z. T. Huang, "Coherent SVR learning for wideband direction-of-arrival estimation," *IEEE Signal Processing Letters*, vol. 26, no. 4, pp. 642–646, 2019.
- [17] K. Huang, M. Y. You, Y. X. Ye, B. Jiang, and A. N. Lu, "Direction of arrival based on the multioutput least squares support vector regression model mathematical problems in engineering, vol." vol. 2020, Article ID 8601376, 2020.
- [18] Y. Cao, T. Lv, Z. Lin, P. Huang, and F. Lin, "Complex ResNet aided DoA estimation for near-field MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11139–11151, 2020.
- [19] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8549–8560, 2018.
- [20] H. Xiang, B. Chen, M. Yang, S. Xu, and Z. Li, "Improved direction-of-arrival estimation method based on LSTM neural networks with robustness to array imperfections," *Applied Intelligence*, vol. 51, no. 7, pp. 4420–4433, 2021.
- [21] M. Stoilkovic, Z. Stankovic, and B. Milovanovic, "A cascade-connected neural model for improved 2D DOA estimation of an EM signal," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 29, no. 2, pp. 343–353, 2016.
- [22] M. P. Enriconi, F. C. C. de Castro, C. Müller, and M. C. F. de Castro, "Phase transmittance RBF neural network

- beamforming for static and dynamic channels,” *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 2, pp. 243–247, 2020.
- [23] Y. Yuan, S. Wu, Y. Ma, L. Huang, and N. Yuan, “KR product and sparse prior based CNN estimator for 2-D DOA estimation,” *AEU - International Journal of Electronics and Communications*, vol. 137, 2021.
- [24] F. Y. Sun, Y. B. Tian, G. B. Hu, and Q. Y. Shen, “DOA estimation based on support vector machine ensemble,” *International Journal of Numerical Modelling: electronic Networks, Devices and Fields*, vol. 32, no. 5, 2019.
- [25] Z.-T. Huang, L.-L. Wu, and Z.-M. Liu, “Toward wide-frequency-range direction finding with support vector regression,” *IEEE Communications Letters*, vol. 23, no. 6, pp. 1029–1032, 2019.
- [26] X. Chen, D. Wang, Z. P. Liu, and Y. Wu, “A fast direct position determination for multiple sources based on radial basis function neural network,” in *Proceedings of the 2018 10th international conference on communication software and networks*, Chengdu, China, July 2018.
- [27] M. Agatonović, Z. Stanković, N. Dončov, B. Milovanović, and I. Milovanović, “Neural network model for 2D DOA estimation of two coherent sources,” *International Journal of Reasoning-based Intelligent Systems*, vol. 7, pp. 62–69, 2015.
- [28] X. Chen, D. Wang, J. Yin, and Y. Wu, “A Direct Position-Determination Approach for Multiple Sources Based on Neural Network Computation,” *Sensors*, vol. 18, no. 6, 2018.
- [29] W. Zhang, Y. Huang, J. Tong, M. Bao, and X. Li, “Off-Grid DOA Estimation Based on Circularly Fully Convolutional Networks (CFCN) Using Space-Frequency Pseudo-Spectrum,” *Sensors*, vol. 21, no. 8, 2021.
- [30] Q. Liu, Y. Gu, and H. C. So, “DOA estimation in impulsive noise via low-rank matrix approximation and weakly convex optimization,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 6, pp. 3603–3616, 2019.
- [31] X. Y. Zhang, Y. B. Tian, and X. Zheng, “Antenna optimization design based on deep Gaussian process model,” *International journal of Antennas and propagation*, vol. 2020, Article ID 2154928, 2020.
- [32] F. G. Yan, S. Liu, J. Wang, and M. Jin, “Two-Step root-MUSIC for direction of arrival estimation without EVD/SVD computation,” *International journal of Antennas and propagation*, vol. 2018, Article ID 9695326, 2018.
- [33] Y. Yousfi and J. Fridrich, “An intriguing struggle of CNNs in JPEG steganalysis and the OneHot solution,” *IEEE Signal Processing Letters*, vol. 27, pp. 830–834, 2020.
- [34] F. Zhuang, Z. Qi, K. Duan et al., “A comprehensive survey on transfer learning,” *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2021.
- [35] D. Chen and Y. H. Joo, “A Novel Approach to 3D-DOA Estimation of Stationary EM Signals Using Convolutional Neural Networks,” *Sensors*, vol. 20, no. 10, 2020.
- [36] H. Li, M. Krček, and G. Perin, “A comparison of weight initializers in deep learning-based side-channel analysis,” in *Proceedings of the Applied Cryptography and Network Security*, Rome Italy, October 2020.
- [37] C. Peng, T. Xiao, Z. Li et al., “MegDet: a large mini-batch object detector,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, Salt Lake City, USA, June 2018.
- [38] D. P. Kingma and J. L. Ba, “Adam: a method for stochastic optimization,” in *Proceedings of the 3rd international conference for learning representations*, San Diego, USA, May 2015.
- [39] J. Shi, W. Zhang, L. Tan, L. Mao, and W. Shi, “Under-determined DOA estimation for wideband signals via focused atomic norm minimization,” *Entropy*, vol. 22, no. 3, 2020.