A Robust Two-Dimensional DOA Estimation Approach Based on Convolutional Attention Network

Rui Xiao[1], Ming-Yi You[2], Rui Zhou[3], Qingjiang Shi[1]
[1]School of Software Engineering, Tongji University, Shanghai 200092, China
[2]Science and Technology on Communication Information Security Control Laboratory, Jiaxing 314033, China
[3]Shenzhen Research Institute of Big Data, Shenzhen 518172, China
xiaorui1999@tongji.edu.cn, youmingyi@126.com, rui.zhou@sribd.cn, qing.j.shi@gmail.com

2023 International Joint Conference on Neural Networks (IJCNN) | 978-1-6654-8867-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/IJCNN54540.2023.10191726

***Abstract*—The direction of arrival (DOA) estimation of the signal is an important task in radio signal positioning. Various methods have been investigated to cope with the DOA task. However, since the imperfect interference factors are often present in practical antenna arrays, the performance of DOA estimation is often significantly degraded. Besides, few methods deal with the DOA estimation for signals of multiple frequencies. In this paper, we consider the problem of two-dimensional DOA estimation in the presence of imperfect factors, and propose a novel approach where the convolutional attention network is used for DOA estimation. The frequency information is introduced as a token added to the network, which improves the network robustness while taking into account the case of the signal of multiple frequencies. Besides, we extend the mean square error (MSE) to the design of a new loss function for training to improve the accuracy of the model. The advantages of the proposed DOA estimation scheme are demonstrated through numerical experiments.**

***Index Terms*—DOA estimation, convolutional neural network, attention mechanism, deep learning, imperfect effect**

I. Introduction

Direction-of-arrival (DOA) estimation of signals in the radio direction finding problem has been studied for a long time [1]–[5]. It is a fundamental problem in the fields of radar sensing, wireless communication [6], [7], and has a wide range of applications in military, public security, aviation, navigation, surface transportation, disaster relief, etc. The core of DOA estimation is the mapping relationship of the parameters hidden in the received signal with the DOA [8], [9]. These parameters can be interpreted as the responses which are very relevant to the DOA, and they can be measured from the received signal. The phase difference [10], amplitude difference [11], Doppler effect [12], array subspace [13], and beam formed signal power [14] can be the responses.

A number of traditional algorithms have been investigated for DOA estimation. Two representative algorithms are the multiple-signal classification (MUSIC) algorithm [15] and the estimation of signal parameters via rotational invariance technique (ESPRIT) method [16]. They have been the basis for several other methods. For example, a tree-structured frequency-space-frequency (FSF) MUSIC-based algorithm is proposed in [17], the parameter estimation and filtering processes are combined to joint estimate the DOA and frequency. Aiming to the direction of departure (DOD) and DOA estimation problem of multiple-input multiple-output (MIMO) radio, the work [18] presented a reduced-dimensional MUSIC (RD-MUSIC) algorithm with a lower computational cost. The ESPRIT method and Root-MUSIC method were used in [19] to estimate the DOD and DOA of MIMO radio, respectively. A novel compressed MUSIC (C-MUSIC) spatial spectrum was studied in [20] involving a limited spectral search as compared to the MUSIC which requires a much larger search space. To perform DOA estimation of MIMO radar with imperfect waveforms, the RD-MUSIC algorithm with the noiseless cross-covariance matrix was proposed in [21]. The work [22] proposed a parallel model for DOA estimation of colocated MIMO radar with imperfect waveforms. However, these algorithms have a limited ability to fit interference factors in imperfect condition, and the optimization of the received signal measurement model becomes complicated when multiple factors are considered.

Benefit from the excellent feature extraction and efficient modeling capability of neural networks, many works use deep learning techniques to overcome the effects of interference factors. In [23], an end-to-end deep neural network (DNN) was proposed for channel estimation and DOA estimation of a massive MIMO system. In [24], a DOA estimation method based on convolutional neural network (CNN) is given in a low signal-to-noise ratio (SNR). In [25], an iterative sparse signal recovery algorithm was unfolded as a deep network to estimate DOA in the presence of array imperfections. In [26], multiple deep CNNs were presented, in which the MUSIC spectrum of the received signal can be obtained from the CNN of the corresponding angular subregion. In [27], the Long Short-Term Memory (LSTM) network was used to estimate the DOA, where the network input is the correlation vector of signals while the network output is the one-hot encoding corresponding to the DOA. Besides, the works [28]–[30] have also been proposed for complex receiving signal systems with the mutual coupling effect, inconsistent phases and gains, positions perturbations, etc. However, most of the deep learning based methods nowadays are for the one-dimensional case, while the two-dimensional direction finding is more common in practice and has been widely used in satellite or airborne payloads [9]. Besides, these methods can not handle the case of the signal of multiple frequencies. Constructing a network for each frequency is time-consuming and irrational, which would simply ignore the inner relationship between the samples for different frequencies. Therefore, it is desired to consider the inner relationship and design a generic network framework to satisfy the different frequency cases.

To address these deficiencies, [9] proposed a DNN-based two-dimensional direction finding approach, in which the whole DOA estimation process is divided into two stages: classification and regression. The frequency information is used to partition multi-group networks, and the estimated DOA is obtained by the the corresponding regression network according to the classification result. However, to enjoy higher accuracy by increasing the number of frequency and angle partitions, more classification and regression networks are required. It will unavoidably enlarge the network structure and increase the training overhead. In addition, the frequency of the signal is an estimated value, and the deviation in frequency may lead to incorrect selection of sub-networks, which has a significant impact on the accuracy of the network.

Inspired by the DNN-based approach [9], we focus on developing a convolutional attention network based approach. It is used for the two-dimensional DOA estimation problem of receiving antenna arrays with interference factors. Unlike [9], we do not use frequency partition sub-networks to achieve DOA estimation for multiple frequencies of signal, so the network architecture is significantly simpler. In [9], the frequency is copied as a long vector and added into the input, which increases the complexity and uncertainty of the input data. In our approach, this is avoided by a frequency token. In summary, the contribution of this paper are as follows:

- We propose a DOA estimation approach based on convolutional attention network.
- We design a new input form, where the frequency information is added to the network as a token and improves the robustness.
- We propose a loss function for the network training by extending the mean square error (MSE) loss function.
- Extensive experiments are conducted to determine the optimal network architecture and to verify the effectiveness of our proposed method.

II. Signal Model

For a single far-field signal, we can use the following expression to represent its form at time t

$$s(t) = ue^{j(\omega t+\varphi)}, \tag{1}$$

where $\omega = 2\pi f = 2\pi c/\lambda$ in which f is the signal frequency, c is the speed of propagation of electromagnetic waves, and λ is the wavelength. u and φ are the amplitude and phase of the signal, respectively. For convenience, we assume that the

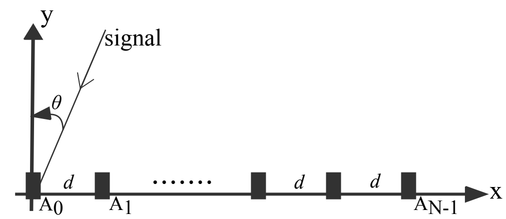

Fig. 1. Planar coordinate system of one-dimensional direction finding.

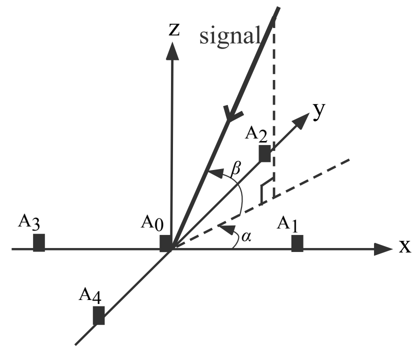

Fig. 2. Spatial coordinate system of two-dimensional direction finding.

amplitude and phase of the same far-field signal do not change with time t. Then the far-field signal with delay τ is

$$\begin{aligned} s(t-\tau) &= ue^{j(\omega(t-\tau)+\varphi)} \\ &= ue^{j(\omega t+\varphi)}e^{-j\omega\tau} = s(t)e^{-j\omega\tau}. \end{aligned} \tag{2}$$

According to the above mathematical modeling of the original far-field signal, it is easy to know that the signal received by the i-th antenna is

$$x_i(t) = s(t-\tau_i) + n_i(t) = e^{-j\omega\tau_i}s(t) + n_i(t), \tag{3}$$

where $i = 0, 1, \cdots, N-1$ and N is the number of elements in the antenna array, τ_i denotes the delay of the original signal imping onto the i-th antenna, which is relative to the reference antenna (i.e., the origin of the coordinate system), and $n_i(t)$ is the noise at time t which is usually considered to be Gaussian white noise. By expanding (3) into the vector form, we can obtain the received signal mathematical model of the entire antenna array

$$\boldsymbol{x}(t) = \boldsymbol{a}s(t) + \boldsymbol{n}(t), \tag{4}$$

where $\boldsymbol{x}(t) = [x_0(t), \cdots, x_{N-1}(t)]^T \in \mathbb{C}^N$ and $\boldsymbol{n}(t) = [n_0(t), \cdots, n_{N-1}(t)]^T \in \mathbb{C}^N$. The steering vector is

$$\boldsymbol{a} = [1, e^{-j\omega\tau_1}, \cdots, e^{-j\omega\tau_{N-1}}]^T. \tag{5}$$

In the one-dimensional direction finding problem, the DOA estimation of signal is performed in a planar coordinate system. As shown in Fig. 1, a uniform linear array (ULA) with N antenna elements ($\mathrm{A}_0 \sim \mathrm{A}_{\mathrm{N}-1}$) is distributed on the x-axis with their spacing being d, θ is the DOA of the signal. With the antenna at the origin of the coordinate system as the reference, the wave path difference of the signal imping onto each antenna is $d_i \sin\theta$, where d_i is the length of the baseline between the i-th and the reference antenna. Therefore, the delay of the i-th antenna is

$$\tau_i = \frac{d_i \sin\theta}{c}. \tag{6}$$

It is not difficult to find that the wave path difference is actually the projection length of the baseline vector on the signal direction vector. The spatial coordinate system of two-dimensional direction finding is shown in Fig. 2, with the DOA angle $\boldsymbol{\theta} = [\alpha, \beta]^T$, where α and β are the azimuth and elevation angles, respectively. The unit-length signal direction vector can be defined as

$$\boldsymbol{r} = [\cos\alpha\cos\beta, \sin\alpha\cos\beta, \sin\beta]^T. \tag{7}$$

Suppose the position coordinate of the reference antenna is the origin $[0,0,0]^T$ and the position of the i-th antenna is at coordinate $[x_i, y_i, z_i]^T$, then the baseline vector is $\boldsymbol{d}_i = [x_i, y_i, z_i]^T$ and the delay τ_i can be calculated by

$$\tau_i = \frac{\boldsymbol{d}_i^T \cdot \boldsymbol{r}}{c}. \tag{8}$$

Based on (8), the phase difference of the i-th antenna with respect to the reference antenna is

$$\phi_i = \omega\tau_i = 2\pi f \tau_i = 2\pi \frac{f}{c} \boldsymbol{d}_i^T \cdot \boldsymbol{r}. \tag{9}$$

There are usually various imperfect factors that usually exist in practical arrays, such as mutual coupling and receiving channel inconsistency [9]. All of these imperfect effects can bias the response model of (9), which will be reflected in the steering vector $\boldsymbol{a}$. Therefore, the problem we need to solve is to estimate the DOA with the help of received signal $\boldsymbol{x}(t) = \boldsymbol{a}(\boldsymbol{\theta}, e)s(t) + \boldsymbol{n}(t)$ from the antenna array under the influence of various imperfect factors, where $\boldsymbol{a}(\boldsymbol{\theta}, e)$ denotes the steering vector with the imperfect effect e when the DOA angle is $\boldsymbol{\theta}$. In this paper, the mutual coupling effect and inconsistent gains/phases effect are considered

$$\boldsymbol{x}(t) = \boldsymbol{B}\boldsymbol{G}\boldsymbol{a}(\boldsymbol{\theta})s(t) + \boldsymbol{n}(t), \tag{10}$$

where $\boldsymbol{B} \in \mathbb{C}^{N\times N}$ is a symmetric matrix whose diagonal elements are all one and denote the mutual coupling effect. The diagonal matrix $\boldsymbol{G} \in \mathbb{C}^{N\times N}$ represents the inconsistent gains/phases effect with the diagonal element $G_i = g_i e^{j\psi_i}$, where g_i and ψ_i are inconsistent gain and phase of the i-th antenna with respect to the reference antenna, respectively.

III. Proposed Approach

A. The Network Architecture

In this paper, we use the convolutional attention network to realize DOA estimation, for convenience we named it CANN, and its architecture is shown in Fig. 3. The network consists of several convolutional modules (gray part) connected with each other to form the main structure. The attention mechanism module (blue part) is followed by these convolutional modules. The regression module (green part) is composed of two fully connected layers and a nonlinear layer located between them.

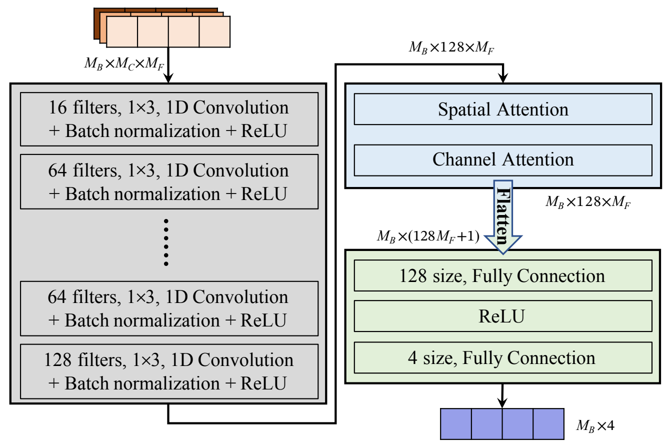

Fig. 3. The architecture of the convolutional attention network.

Among them, each convolutional module is composed of a one-dimensional convolutional layer, a batch normalization layer and a nonlinear layer. Moreover, the attention mechanism module is divided into a spatial attention module and a channel attention module.

In convolutional modules, the first convolutional module is represented as an input layer whose number of filters should not be set too large for the case of oversampling of the feature map, so 16 filters are used in this module. The parameters of the convolutional layers of the other modules between the first and the last convolutional module are set the same. For the considered case of a 5 antenna receiving array, we heuristically set the number of filters for these convolutional layers to be 4 times that of the first convolutional module, i.e. 64 (because there is a reference antenna). And the number of filters of the last convolutional module is 128. For all of them, considering the length of the input data feature the convolutional kernel size is 1×3, and the padding and stride are set to keep the feature length of convolution output consistent with the input. Moreover, the ReLU function is used as a nonlinear layer.

Then, we use the convolutional block attention module (CBAM) [31] to compute the attention weights in the feature dimension and the channel dimension successively. The parameters of the convolutional layer in the spatial attention module are similar to those of the convolutional module, where the convolution kernel size is 1×3 and the padding is set to keep the size of output. While in the channel attention module, the length of the feature dimension is replaced by the number of channels due to the global maximum and global average pooling operations, so the shared multi-layer perceptron (MLP) is replaced by a one-dimension convolutional layer with a kernel size of 1×1.

Finally, as the regression module used by the network for DOA estimation, the output of the two fully connected layers are 128 and 4, which correspond to the number of filters in the last convolutional module and the number of sine and cosine values of the DOA angle. And the nonlinear layer is ReLU.

It is noted that after the flatten operation we add a token ξ to the data and then input it into the regression module. We use $\mathcal{I} \in \mathbb{R}^{M_B \times M_C \times M_F}$ to denote the input data, then the whole network computation process can be expressed as

$$\mathcal{O} = F_6 \left(\xi, F_5 \left(F_4 \left(F_3 \left(F_2 \left(\cdots F_2 \left(F_1 \left(\mathcal{I}\right)\right) \cdots\right)\right)\right)\right)\right), \quad (11)$$

where M_B, M_C, M_F denote the batch size, the number of channel and the feature length of input data, respectively. F_1, F_2, F_3 denote the convolutional modules, F_4 and F_5 refer to the spatial and channel attention module, then F_6 corresponds to the regression module. We use $\mathcal{O} = [\sin\alpha, \cos\alpha, \sin\beta, \cos\beta]^T \in \mathbb{R}^4$ as the output of this network, which is the same as in [9] and has been proven to be closed to optimal.

B. Model Input

The input form of the current deep learning based DOA estimation methods can be broadly divided into two categories. On one hand, the raw sampled signal data from the antenna array could be used directly as the input, and on the other hand, the covariance matrix of the received signal could be used. There is a significant shortcoming for the former, i.e., the number of sampling points can have an impact on the model. Specifically, a small number of samples means limited information contained and would cause a inaccurate estimation result, while a large number makes the model input complicated and increases the training cost. In contrast, the covariance matrix contains information about the sampled data while reducing the input dimension, which achieves a balanced effect.

In order to reduce the network input size we only use the upper diagonal elements of the covariance matrix (12) and construct a vector form (13), which contains information about the interrelationship among the individual antennas

$$\boldsymbol{C} = \frac{1}{M} \sum_{i=0}^{M-1} \boldsymbol{x}(t_i) \boldsymbol{x}^H(t_i), \quad (12)$$

$$\boldsymbol{C}_{up} = [C_{1,2}, C_{1,3}, C_{2,3}, \cdots, C_{1,N}, \cdots, C_{N-1,N}]^T, \quad (13)$$

where M is the number of sample points, and $C_{i,j}$ denotes the element of covariance matrix $\boldsymbol{C}$ in row i column j. Therefore, the input form of our proposed network model is as follows

$$\mathcal{I} = [real\{\boldsymbol{C}_{up}\}, imag\{\boldsymbol{C}_{up}\}]^T \in \mathbb{R}^{2 \times N(N-1)/2}, \quad (14)$$

where $real\{\cdot\}$ and $imag\{\cdot\}$ denote the real and imaginary parts of a complex value, respectively.

Furthermore, as mentioned in III-A, in addition to the input at the beginning of the network, a token is added as an input to the regression module after the flatten operation. In the proposed method, the token is set to the frequency of the received signal $\xi = f$ in (11), which is, in practice, a feature estimated by a certain method. The frequency information used before the regression module is due to two considerations. First, if feed in at the beginning of the network, it is usually required to repeat the frequency to align with the dimension of $\boldsymbol{C}_{up}$ (e.g., [9]). However, as an estimation feature, the repetition increases the complexity and uncertainty of the input information. Second, the various convolution operations of the network play the role of feature extraction, and the frequency is also a feature estimated from the received signal, so it is reasonable to connect it with the output of these convolution operations for input into the regression module.

C. Loss Function

Based on the network model output $\mathcal{O}$, an intuitive loss function is MSE. The process of DOA estimation with a neural network is solving a regression problem for which the MSE loss function is very common. In intelligent direction finding methods based on neural network model [6], most of the models are trained with the MSE loss function. In this proposed network, the commonly used MSE loss function is extended to conform to the output of the network, which is named ExMSE

$$\text{ExMSE} = \mu \text{MSE} + \nu \text{Ex}, \quad (15a)$$

$$\text{MSE} = \frac{1}{M_B} \sum_{i=0}^{M_B-1} \left(\hat{\mathcal{O}}_i - \mathcal{O}_i\right)^T \left(\hat{\mathcal{O}}_i - \mathcal{O}_i\right), \quad (15b)$$

$$\begin{aligned} \text{Ex} = \frac{1}{M_B} \sum_{i=0}^{M_B-1} [\left(1 - (\hat{\sin}\,\alpha_i)^2 - (\hat{\cos}\,\alpha_i)^2\right)^2 + \\ \left(1 - (\hat{\sin}\,\beta_i)^2 - (\hat{\cos}\,\beta_i)^2\right)^2], \end{aligned} \quad (15c)$$

where $\hat{\mathcal{O}} = \left[\hat{\sin}\,\alpha, \hat{\cos}\,\alpha, \hat{\sin}\,\beta, \hat{\cos}\,\beta\right]^T$ denotes the prediction result of the network. Both μ and ν are weighting coefficients that can be found by cross-validating to obtain better performance. For simplicity, we set $\mu = \nu = 1$ in our experiments.

It is intuitive to add this extension term Ex to the loss function for network models with trigonometric values as output. Each of the sine and cosine values are estimated independently. We want to ensure that they are conform this mathematical constraint, so this extension term need to be added to the loss function as a penalty. The extension term can improve the accuracy of the network, which will be verified in subsequent experiments.

IV. Experiments

A. Dataset

In this paper, we consider the DOA estimation problem in the two-dimensional direction finding. A surface array containing five antennas is used to receive the signal. The frequency range of the signal is $f \in [300, 350]$MHz with step length 5MHz, the wave speed is $c = 3 \times 10^8$ m/s, and $\boldsymbol{n}(t)$ is 0-mean Gaussian white noise. The mutual coupling effect and the inconsistency effect in (10) are the same as [9]. It is noted that we set the antenna elements at the same position as [9], but we use the baseline between the i-th and the reference antenna, where the $\boldsymbol{d}_0 = [0, 0, 0]^T$ is the baseline vector of the reference antenna.

Based on the above settings, we generate the signal data in ideal and complex receiving systems according to (4) and (10). If conditions permit, we recommend collecting samples in a microwave unreflected chamber or other controlled environments. In the simulation of the training samples, the azimuth and elevation angles of the signals are $\alpha \in [0°, 358°]$ and $\beta \in [30°, 88°]$ with step length $2°$, respectively, and a perturbation within $(0°, 2°)$ is added to each angle. The signal of each set of f, α and β is repeatedly sampled 5 times with a sampling frequency of 1000, and is randomly discarded with 40% possibility. The validation samples are similar to training, but the step length is changed to $10°$ in α, repeated 10 times and without randomly discarding. The test samples are all in complex receiving system, in which α is in the range $[2°, 180°]$ with step length $5°$, β is in the range $[30°, 88°]$ with step length $2°$, and repeat sampling for 5 times without randomly discarding. It is noted that the SNR of the training and validation set is 10 and is in the range of $[0, 30]$ with step length 5 in the test set.

B. Evaluation Metrics

To measure the performance of the proposed DOA estimation method, we calculate the direction vector of the signal with the help of the trigonometric values of the network output. The angle between the estimated direction vector and the real direction vector of the signal is used as the evaluation metric [9], that is

$$\hat{\boldsymbol{r}} = \left[\hat{\cos\alpha}\hat{\cos\beta}, \hat{\sin\alpha}\hat{\cos\beta}, \hat{\sin\beta}\right]^T, \tag{16a}$$

$$\text{error} = \arccos\frac{\hat{\boldsymbol{r}}^T\boldsymbol{r}}{\|\hat{\boldsymbol{r}}\|\|\boldsymbol{r}\|}, \tag{16b}$$

where $\hat{\boldsymbol{r}}$ is the estimated direction vector with bias, so $\|\hat{\boldsymbol{r}}\| \approx 1$, while $\boldsymbol{r}$ is a unit direction vector and $\|\boldsymbol{r}\| = 1$.

A very straightforward idea in the DOA estimation problem is to use the difference between the estimated angle and the real angle as the evaluation metric. However, in the method where the trigonometric values are used as the output, this consideration can lead to bias in the evaluation metric. Obviously, $\sin 1°$ and $\sin 359°$ are close enough that the error is not very large when the evaluation metric is calculated using the direction vector. But converting them to angle values, the error between $1°$ and $359°$ will be unacceptable. Therefore, it is a reasonable and effective idea to use (16b) as the evaluation metric.

C. Experimental Results

In this section, we will demonstrate the DOA estimation performance of the proposed approach on signals containing interference factors through a series of experiments.

First, in the CANN, an important hyperparameter must be considered to achieve better DOA estimation. In section III-A we mentioned that there are several identical convolutional modules between the first and the last convolutional module of CANN, and different numbers of these middle convolutional modules can have different effects on the feature extraction ability of the network. In Fig. 4, we show the DOA estimation performance with different numbers of middle convolutional modules.

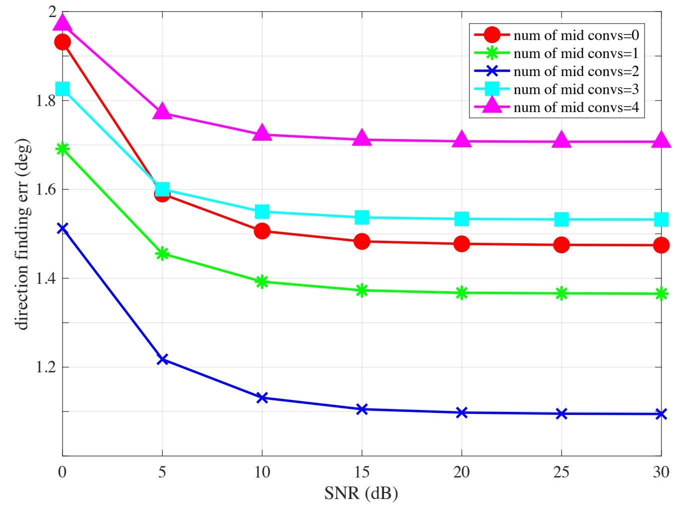

Fig. 4. The DOA estimation performance with different numbers of middle convolutional modules.

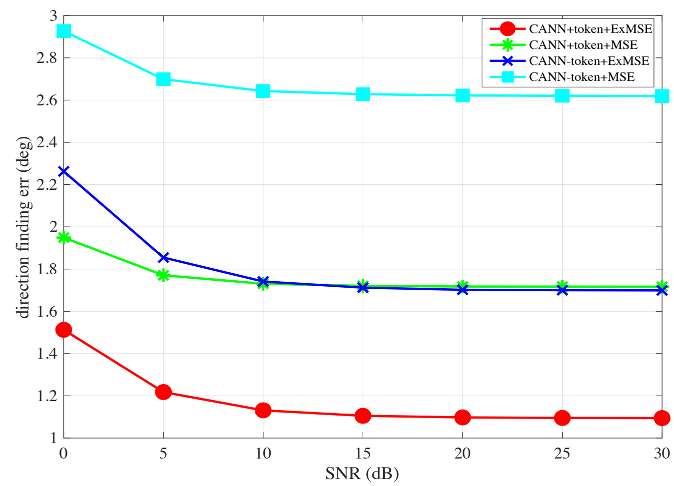

Fig. 5. The DOA estimation performance with different influencing factors.

As shown in this figure, the direction finding error decreases and then increases with the increase of the number of intermediate convolution modules. When the number is small, the network structure is relatively simple and has limited ability to fit the data. The network will be over-fitted and the complexity becomes larger when the number increases. Obviously, when the number of the middle convolutional modules is 2, a better trade-off between DOA estimation performance and network complexity is achieved. So added with the first and last one, we have a total of 4 convolutional modules for CANN.

Then, we compare the influence of the loss function ExMSE and the token in the regression module on the network for DOA estimation. The results are shown in Fig 5, where the symbols $+$*token* and $-$*token* means to use and not to use the token in the regression module, respectively, and the $+$*ExMSE* and $+$*MSE* indicate which loss function is used to train the network model.

In Fig 5, it is easy to find that the CANN, which combines the additional token and trained with the ExMSE loss function, has a better DOA estimation performance. Oppositely, the network without token and trained with MSE has worse performance. Looking only at the error curves of the networks trained with different loss functions, it can be found that the trend of these curves of the same loss function is similar, and the ExMSE-trained network is more sensitive to the change of SNR in the low SNR case. Consistent with what we stated in III-C, the ExMSE has a better direction finding accuracy than the MSE loss function. When using ExMSE to train the network, the Ex term in (15c) makes the output trigonometric function values more mathematically correct, and improves the fitting ability of the network to achieve higher accuracy. Overall, the networks with the token have better DOA estimation performance.

Next, we compare our method with other methods, i.e., the MUSIC method [15] and the DNN-based method [9]. Both two methods are suitable for the two-dimensional DOA estimation problem. We re-implemented these two methods and evaluate them on the previously mentioned dataset. The comparison results of them and our proposed method are arranged in Table I.

As shown in the table, our proposed method outperforms the baseline methods for DOA estimation results for signal data with different SNR. Although the DNN-based method is slightly worse than our proposed method in direction finding error, the $2.4^{\circ} \sim 2.5^{\circ}$ of error is acceptable. However, the MUSIC algorithm cannot model the interference factors in the input signal data as better as the network model, which leads to worse performance, and to achieve higher accuracy, the grid needs to be more refined and the computational overhead is significantly increased.

The number of network modules of the DNN-based method and ours is shown in Table II, where N_f and N_z denote the number of classification and regression parts, N_m denotes the number of middle convolutional modules. In the DNN-based method, the entire network is divided into a number of classification and regression networks based on partitioning. And a classification part contains 2 convolutional modules and 2 fully connected layers, while a regression network contains 4 convolutional modules and 2 fully connected layers. Because of the existence of partition, the original single network structure is divided into multiple parts which undoubtedly makes the network structure huger and requires more training as well as deployment costs. From the practical application, we need to consider the resource issue, and the complex network may not be suitable for use in miniaturized devices with limited resources. So, both in terms of DOA estimation accuracy and complexity of the model, the method we proposed is better than the DNN-based method.

Finally, we add certain perturbations to the frequency in the test data, which is a simulation of the frequency estimation in practice, to test the robustness of the network when adding frequency as the token to the regression module. The experimental results are shown in Table III, where the symbol $\pm$*1MHz* means the perturbation is between $-$*1MHz* and $+$*1MHz*. Obviously, the variation of the direction finding error is small for different degrees of frequency perturbation, which is a good demonstration of the robustness of our proposed approach. On the other hand, the frequency partition in the DNN-based method to determine the sub-network may lead to bias in the presence of perturbations, while in our approach this phenomenon can be avoided by the frequency token. In a sense, the token increases the resistance of the network to perturbation.

TABLE I
COMPARISON RESULTS WITH MUSIC AND DNN-BASED METHODS

Method	Direction Finding Error In Different SNR (deg)						
	0dB	5dB	10dB	15dB	20dB	25dB	30dB
MUSIC [15]	27.572	27.609	27.621	27.603	27.605	27.598	27.602
DNN-Based [9]	2.496	2.423	2.411	2.407	2.406	2.407	2.406
Ours	**1.513**	**1.218**	**1.131**	**1.105**	**1.098**	**1.095**	**1.094**

TABLE II
COMPARISON RESULTS WITH DNN-BASED METHODS IN TERMS OF THE NUMBER OF DIFFERENT NETWORK MODULES

Method	Conv	FC	CBAM
DNN-Based [9]	$2N_f + 4N_f N_z$	$2N_f + 2N_f N_z$	-
Ours	$2 + N_m$	2	1

TABLE III
COMPARISON RESULTS WITH DIFFERENT FREQUENCY PERTURBATION

Frequency Perturbation	Direction Finding Error In Different SNR (deg)						
	0dB	5dB	10dB	15dB	20dB	25dB	30dB
$\pm$0MHz	1.513	1.218	1.131	1.105	1.098	1.095	1.094
$\pm$1MHz	1.512	1.185	1.273	1.279	1.095	1.167	1.082
$\pm$2MHz	1.719	1.179	1.002	1.123	0.949	1.016	1.313
$\pm$3MHz	1.541	1.404	0.952	1.006	1.397	1.091	1.011

V. CONCLUSION

In this paper, we consider the problem of two-dimensional DOA estimation in imperfect condition. A mathematically model of the received signal in the antenna array with the mutual coupling effect and the inconsistent effect has been formulated. Then a convolutional neural network with attention mechanism has been used to address this problem. The ExMSE loss function has been used for the network training. Moreover, the input data of the network has been redesigned, in which the frequency information is passed to the regression module as a token. While avoiding the complexity and uncertainty of the input data brought by the copy operation, it increases the robustness of the network to frequency perturbation. The simulation results has shown the advantages of our proposed two-dimensional DOA estimation scheme.

REFERENCES

[1] Dhabale, Ashmika. ”Direction of arrival (DOA) estimation using array signal processing,” University of California, Riverside, 2018.

[2] Ching-Sung Shieh and Chin-Teng Lin, ”Direction of arrival estimation based on phase differences using neural fuzzy network,” in IEEE Transactions on Antennas and Propagation, vol. 48, no. 7, pp. 1115-1124, July 2000

[3] L. Lizzi, G. Oliveri, P. Rocca, and A. Massa, “Estimation of the directions-of-arrival of correlated signals by means of a SVM-based multi-resolution approach,” in Proceedings of the 2010 IEEE Antennas and Propagation Society International Symposium, IEEE, Toronto, Canada, July 2010.

[4] Zhang Ke, Ma Peng and Zhang Jian-yun, ”DOA estimation algorithm based on FFT in switch antenna array,” Proceedings of 2011 IEEE CIE International Conference on Radar, 2011.

[5] 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 Proceedings of the ICASSP, pp. 2814–2818, IEEE, Brisbane, Australia, April 2015.

[6] J. A. Zhang et al., ”An Overview of Signal Processing Techniques for Joint Communication and Radar Sensing,” in IEEE Journal of Selected Topics in Signal Processing, vol. 15, no. 6, pp. 1295-1315, Nov. 2021.

[7] C. Xu, B. Clerckx, S. Chen, Y. Mao and J. Zhang, ”Rate-Splitting Multiple Access for Multi-Antenna Joint Radar and Communications,” in IEEE Journal of Selected Topics in Signal Processing, vol. 15, no. 6, pp. 1332-1347, Nov. 2021.

[8] M.-Y. You, A.-N. Lu, Y.-X. Ye, K. Huang and B. Jiang, ”A review on machine learning based radio direction finding,” Mathematical Problems in Engineering, vol. 2020, Feb. 2020.

[9] M. -Y. You, ”A Unified Two-Dimensional Direction Finding Approach for Sensor Arrays Based on Deep Neural Networks With Inhomogeneous Angle and Frequency Partition,” in IEEE Sensors Journal, vol. 22, no. 7, pp. 6840-6850, 1 April1, 2022.

[10] Z. -M. Liu and F. -C. Guo, ”Azimuth and Elevation Estimation With Rotating Long-Baseline Interferometers,” in IEEE Transactions on Signal Processing, vol. 63, no. 9, pp. 2405-2419, May1, 2015.

[11] Garrido Lopez, M. A. Al-Tarifi, G. Lasser and D. S. Filipovic, ”Wideband Antenna Systems for Millimeter-Wave Amplitude-Only Direction Finding,” in IEEE Transactions on Antennas and Propagation, vol. 66, no. 6, pp. 3122-3129, June 2018.

[12] Y. -Y. Shih, A. -C. Pang and P. -C. Hsiu, ”A Doppler Effect Based Framework for Wi-Fi Signal Tracking in Search and Rescue Operations,” in IEEE Transactions on Vehicular Technology, vol. 67, no. 5, pp. 3924-3936, May 2018.

[13] E. D. D. Claudio, R. Parisi and G. Jacovitti, ”Space Time MUSIC: Consistent Signal Subspace Estimation for Wideband Sensor Arrays,” in IEEE Transactions on Signal Processing, vol. 66, no. 10, pp. 2685-2699, 15 May15, 2018.

[14] M. Hawkes and A. Nehorai, ”Acoustic vector-sensor beamforming and Capon direction estimation,” in IEEE Transactions on Signal Processing, vol. 46, no. 9, pp. 2291-2304, Sept. 1998.

[15] R. Schmidt, ”Multiple emitter location and signal parameter estimation,” in IEEE Transactions on Antennas and Propagation, vol. 34, no. 3, pp. 276-280, March 1986.

[16] A. Paulraj, R. Roy and T. Kailath, ”A subspace rotation approach to signal parameter estimation,” in Proceedings of the IEEE, vol. 74, no. 7, pp. 1044-1046, July 1986.

[17] J. -D. Lin, W. -H. Fang, Y. -Y. Wang and J. -T. Chen, ”FSF MUSIC for Joint DOA and Frequency Estimation and Its Performance Analysis,” in IEEE Transactions on Signal Processing, vol. 54, no. 12, pp. 4529-4542, Dec. 2006.

[18] X. Zhang, L. Xu, L. Xu and D. Xu, ”Direction of Departure (DOD) and Direction of Arrival (DOA) Estimation in MIMO Radar with Reduced-Dimension MUSIC,” in IEEE Communications Letters, vol. 14, no. 12, pp. 1161-1163, December 2010.

[19] M. L. Bencheikh and Y. Wang, ”Joint DOD-DOA estimation using combined ESPRIT-MUSIC approach in MIMO radar,” Electronics Letters, vol. 46, no. 15, pp. 1081–1083, 2010.

[20] F. Yan, M. Jin and X. Qiao, ”Low-Complexity DOA Estimation Based on Compressed MUSIC and Its Performance Analysis,” in IEEE Transactions on Signal Processing, vol. 61, no. 8, pp. 1915-1930, April15, 2013.

[21] F. Wen, ”Computationally Efficient DOA Estimation Algorithm for MIMO Radar With Imperfect Waveforms,” in IEEE Communications Letters, vol. 23, no. 6, pp. 1037-1040, June 2019.

[22] N. -J. Ruan, F. -q. Wen, L. Ai and K. Xie, ”A PARAFAC Decomposition Algorithm for DOA Estimation in Colocated MIMO Radar With Imperfect Waveforms,” in IEEE Access, vol. 7, pp. 14680-14688, 2019.

[23] 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,” in IEEE Transactions on Vehicular Technology, vol. 67, no. 9, pp. 8549-8560, Sept. 2018.

[24] G. K. Papageorgiou, M. Sellathurai and Y. C. Eldar, ”Deep Networks for Direction-of-Arrival Estimation in Low SNR,” in IEEE Transactions on Signal Processing, vol. 69, pp. 3714-3729, 2021.

[25] L. Wu, Z. Liu, and J. Liao, “DOA estimation using an unfolded deep network in the presence of array imperfections,” in 2022 7th International Conference on Signal and Image Processing (ICSIP), 2022, pp.182–187.

[26] A. M. Elbir, ”DeepMUSIC: Multiple Signal Classification via Deep Learning,” in IEEE Sensors Letters, vol. 4, no. 4, pp. 1-4, April 2020.

[27] T. Han-Trong, N. Ngo Duc, H. Tran Van, H. Pham-Viet, and A. Bardhan, ”Direction of Arrival Estimation for Coherent Signals’ Method Based on LSTM Neural Network,” in Applied Computational Intelligence and Soft Computing, vol. 2022, Jan. 2022.

[28] S. Vigneshwaran, N. Sundararajan and P. Saratchandran, ”Direction of Arrival (DoA) Estimation Under Array Sensor Failures Using a Minimal Resource Allocation Neural Network,” in IEEE Transactions on Antennas and Propagation, vol. 55, no. 2, pp. 334-343, Feb. 2007.

[29] R. Wang, B. Wen and W. Huang, ”A Support Vector Regression-Based Method for Target Direction of Arrival Estimation From HF Radar Data,” in IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 5, pp. 674-678, May 2018.

[30] Z. -M. Liu, C. Zhang and P. S. Yu, ”Direction-of-Arrival Estimation Based on Deep Neural Networks With Robustness to Array Imperfections,” in IEEE Transactions on Antennas and Propagation, vol. 66, no. 12, pp. 7315-7327, Dec. 2018.

[31] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon, ”Cbam: Convolutional block attention module,” in Proceedings of the European Conference on Computer Vision (ECCV), pp. 3-19, 2018.