DOA Estimation in MIMO Radars via Deep Learning

1st Kerem Maden
Radar, Electronic Warfare and Intelligence Systems Division
ASELSAN
Ankara, Türkiye
kmaden@aselsan.com.tr

2nd Işın Erer
Electrical and Electronics faculty
Istanbul Technical University
Istanbul, Türkiye
ierer@itu.edu.tr

***Abstract*—The Direction of Arrival (DOA) estimation is an active research area in array signal processing. Conventional DOA estimation methods require high computational complexity for the multiple-input and multiple-output (MIMO) radars which require the use virtual data vector. In addition, while most conventional methods perform well in high signal-to-noise ratio (SNR) environments, the results in low SNR conditions are not satisfactory. To address these problems, this paper introduces an architecture composed of denoising convolutional autoencoders (DCAE) and convolutional neural networks (CNN) named as DCAE-CNN architecture. The DCAE is used to restore the data prior to DOA estimation, and CNN is employed to estimate the angle of arrival by mapping the restored data to the corresponding angles. Compared to the conventional MUSIC algorithm, experimental results of the proposed DCAE-CNN scheme demonstrate more satisfactory performance in terms of accuracy in low SNR circumstances and reduce the computation time considerably which makes it's use possible for in real-time applications.**

***Index Terms*—direction of arrival (DOA), multiple-input multiple-output (MIMO) radar, convolutional neural network (CNN), denoising convolutional neural networks (DCAE)**

I. Introduction

The direction of arrival (DOA) is the estimation process of the angle of arrival of targets by a sensor array. DOA has many applications in a wide range of technologies e.g., radar, acoustics, and wireless systems [1]. Both multiple signal classification (MUSIC) [2] and the estimation of signal parameters via rotational invariance techniques (ESPRIT) [3] are the most successful conventional methods proposed for the DOA estimation task. MUSIC and ESPRIT are sub-space-based methods that require many samples to create a covariance matrix that is correlated with the target and the estimation accuracy suffers heavily under heavy noise.

Precise estimation of DOA can be achieved with conventional methods by employing large antenna arrays or increasing the number of snapshots which also increases computational and hardware complexity. Multiple-input multiple-output (MIMO) radars can create virtual aperture and this virtual aperture increases the number of antenna elements in a sensor array without increasing hardware complexity. Thus, MIMO radars increase the angular resolution broadly by using fewer antenna elements for transmitting and receiving than conventional active radars [4] [5]. However, the extension of the virtual aperture results in a covariance matrix larger than the covariance matrix created from the received signal. Both state-of-the-art MUSIC and ESPRIT algorithms suffer from the high complexity of the eigendecomposition process. As the antenna array gets larger, the covariance matrix grows exponentially and the complexity of the sub-space-based algorithms also increases exponentially. Consequently, The increment in complexity makes the implementation of these methods hard for real-time applications.

Artificial neural networks(ANN) are used widely for DOA estimation with different antenna arrays and network architectures. CNN is a specialized version of ANN that has convolutional layers for feature mapping and pooling layers for downsampling. In [6], [7] are used CNN and [8] used ANN for DOA estimation task.

In this paper, a DOA estimation method based on DCAE-CNN architecture for MIMO radar is proposed. The method aims to enhance the angular resolution while reducing the processing time. This architecture utilizes the virtual covariance matrix as an image and solves the DOA estimation task as a classification problem. To decrease the complexity as well as the processing time, spatial smoothing [9] is applied to the covariance matrix of the virtual array to diminish it's size. DOA estimation methods that use covariance matrix performs unsatisfactorily in poor SNR conditions since the covariance matrix is heavily corrupted. To improve robustness and accuracy, the DCAE [10] is employed to predict the ideal (noise-free) covariance matrix from corrupted one. Then, the predicted covariance matrix is used as an input to the CNN estimator to solve the DOA estimation problem as a classification problem. The CNN is trained to learn the mapping between the predicted covariance matrix and the angle of arrival. The main reason CNN has been chosen over conventional methods is to avoid the computationally intensive eigendecomposition process. After the training and validation process, CNN models are fast and are widely used in real-time image and video classification systems [11]. The processing time and accuracy of the proposed network are compared to the MUSIC algorithm. A short review of a MIMO radar and virtual array vector are provided in section 2. The proposed DCAE-CNN model with data pre-processing is introduced in

978-1-6654-7273-9/22/$31.00 ©2022 IEEE

section 3. Experimental results are given in section 4 and section 5 concludes the paper.

II. Problem Formulation

Consider a MIMO radar system that consists of M uniform linear array (ULA) elements on both transmiter and receiver. Both the transmit and receive elements are in the same array (close to each other) and it is assumed that the received signal meets the narrowband condition so the target is perceived in the same direction by the receive elements. For the MIMO radar case, transmit elements are assumed to satisfy the orthogonality condition

$$\int \phi_m(t)\phi_m^H(t)\,dt = I_m \tag{1}$$

where $\phi(t)_m$ is the m^{th} element of the waveform vector $\boldsymbol{\phi(t)} = [\phi_1(t), \cdots, \phi_M(t)]^T$.

The received vector by the ULA can be expressed as

$$\boldsymbol{x(t)} = \boldsymbol{x_s(t)} + \boldsymbol{n(t)} \tag{2}$$

where the $\boldsymbol{x(t)} = [x_1(t), \cdots, x_m(t)]$ is the received signal vector, $\boldsymbol{x_s(t)} = [x_{s1}(t), \cdots, x_{sm}(t)]$ is the source signal vector and $\boldsymbol{n(t)} = [n_1(t), \cdots, n_m(t)]$ is the zero-mean uncorrelated noise vector.

Under the simplification of the point target, source signal vector can be described by the equation

$$\boldsymbol{x_s(t)} = \beta(\boldsymbol{a^T}\theta_s\ \boldsymbol{\phi(t)})\boldsymbol{b}(\theta_s) \tag{3}$$

where β is the complex reflection coefficient that is unique to the source signal, $\boldsymbol{a}(\theta_s)$ is uplink and $\boldsymbol{b}(\theta_s)$ is downlink steering vectors whereas θ_s is the source direction and the $\boldsymbol{\phi(t)}$ is the waveform vector.
Transmitted m^{th} waveform for each receive element can be recovered by matched filtering which is multiplying the received signal by the waveform vector

$$y_m(t) = \int x_m(t)\phi_m^*(t)\,dt \tag{4}$$

where y_m is the m^{th} element of the virtual data vector $\boldsymbol{y} = [y_1(t), \cdots, y_m(t)]$. After the matched filtering, the waveform vector is eliminated so $M \times N$ data vector where the N is the number of snapshots $\boldsymbol{y}$ can be simplified as follows

$$\boldsymbol{y} = \beta\boldsymbol{a}(\theta_s) \otimes \boldsymbol{b}(\theta_s) \tag{5}$$

As stated in [12] [13], using virtual data vector $\boldsymbol{y}$ provides higher performance for the angular resolution and detection for the MIMO radar case.

III. Proposed Method

The proposed DCAE-CNN architecture is shown in Fig. 2. The architecture includes a preprocessing stage where the data size is reduced, following by a pre-trained DCAE used to reduce the noise. Lastly, a CNN is employed as an estimator.

Preprocessing step: The sample covariance matrix $\boldsymbol{R_{yy}}$ of the virtual data vector $\boldsymbol{y}$ is given as

$$\boldsymbol{R_{yy}} = E\{\boldsymbol{yy^H}\} \tag{6}$$

where $\boldsymbol{R_{yy}}$ is a $MM \times MM$ complex matrix and each element of $\boldsymbol{R_{yy}}$ can be represented as real and imaginary components. The angle of arrival information mostly resides in the sample covariance matrix. Even though most of the information lies in the phase difference between different elements of $\boldsymbol{R_{yy}}$, to extract every bit of information, both real and imaginary components will be utilized to find DOA. The real and imaginary representation of $\boldsymbol{R_{yy}}$ is given as

In the pre-processing stage spatial smoothing is introduced to reduce the size of the sample covariance matrix. Since $\boldsymbol{R_{yy}}$ contains M^2 individual sub-array matrices with each size $M \times M$ as in (7) they can be smoothed as followed in equation (8)

$$\boldsymbol{R} = \boldsymbol{R_{yy}} = \begin{pmatrix} R_{11} & \cdots & R_{1N} \\ \vdots & \vdots & \ddots \\ R_{M1} & \cdots & R_{MN} \end{pmatrix} \tag{7}$$

$$\overline{\boldsymbol{R}} = \sum_{m=0}^{M}\sum_{n=0}^{N} \boldsymbol{R}_{mn} \tag{8}$$

where $\overline{\boldsymbol{R}}$ is $M \times M$ smoothed covariance matrix that is normalized and partitioned into real and imaginary components to feed in DCAE network. The normalization type is min-max normalization with the formulation as follows

$$\overline{\boldsymbol{R}}_{scaled} = \frac{\overline{\boldsymbol{R}} - \overline{\boldsymbol{R}}_{min}}{\overline{\boldsymbol{R}}_{max} - \overline{\boldsymbol{R}}_{min}}. \tag{9}$$

Denoising Convolutional Autoencoder step: The DCAE is used the recover theoretical covariance matrix from the corrupted version. In the encoding stage, input x which may be corrupted by noise is compressed to a low dimensional matrix to extract principal components by encoder $f_\theta(x)$ where W is the weight matrix and b is the offset vector of the encoder. In decoding stage, hidden representation (encoded) y is mapped back by decoder $g_{\theta'}(y)$ to x' where W' is the weight matrix and b' is the offset vector of the decoder. x' is not an exact reconstruction of x but "recovered version" of x.

$$y = f_\theta(x) = s(Wx + b) \tag{10}$$

$$x' = g_{\theta'}(y) = s(W'y + b') \tag{11}$$

Convolutional Neural Network Estimator step: The CNN estimator has three groups of convolutional and pooling layers followed by dense layers and a softmax layer. As the activation function, rectified linear activation function (ReLU) is used for both convolutional and fully connected layers. ReLU is used because other activation functions lead to slow convergence and computationally ReLU and derivative of ReLu are faster than sigmoid or tanh. Adam optimizer to minimize the loss

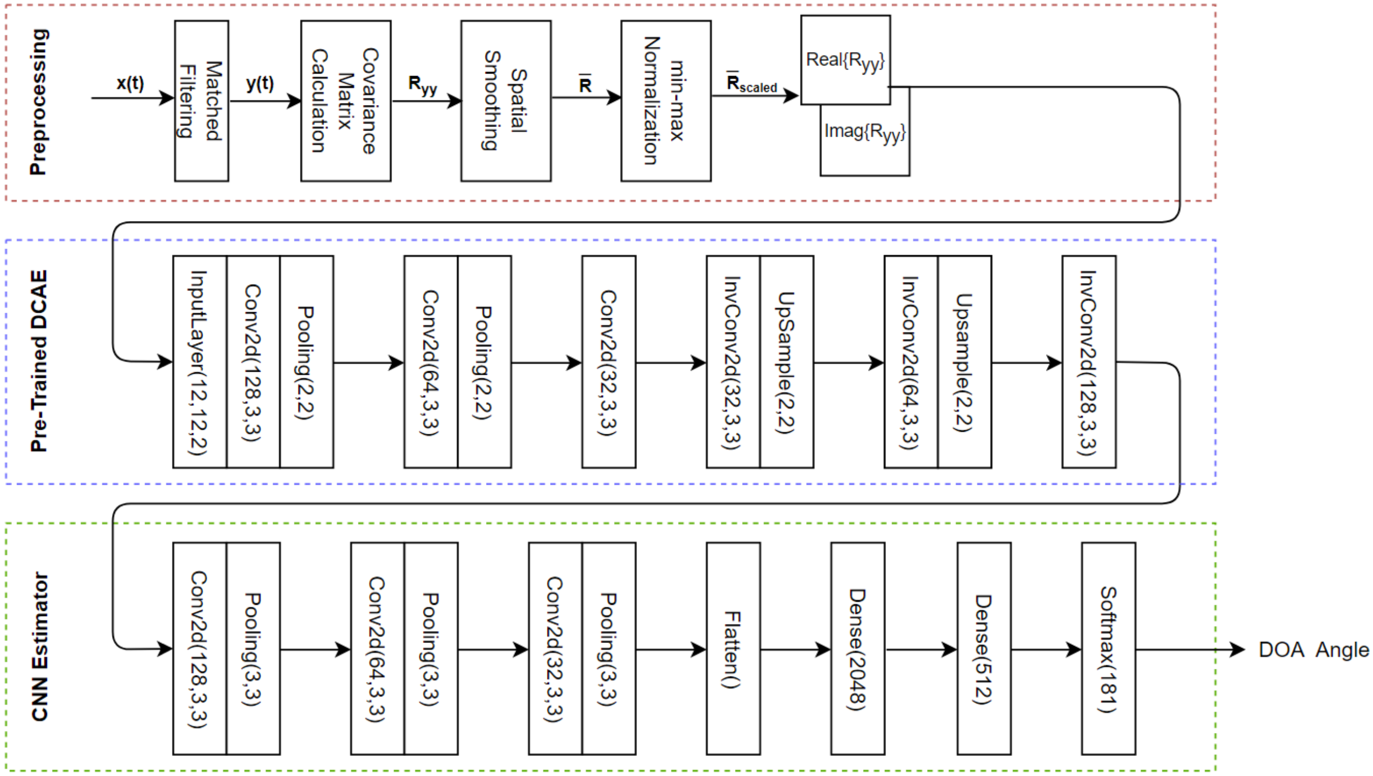

Fig. 1. The Proposed DCAE-CNN Architecture.

function, sparse categorical cross entropy is used. In contrast to the gradient descent-based algorithms which maintain a single learning rate for all weights and do not change during training, the Adam optimizer is an adaptive method. It can update the learning rate. Thus, it can converge to global minimum more efficiently.

IV. Experimental Results

MIMO radar transmitter and receiver models are created in MATLAB 2020b. The number of array elements M is 12 with element spacing d equals 0.5 for both transmit and receive arrays. The target angle range is between -90 to 90 with a total of 181 direction angles and each angle is sampled 5000 times for -10, -5, 0, 5, and 10 SNR values. In total, 905000 (R, θ) pairs are generated for each SNR. These generated pairs were split into three groups for the training of the CNN Estimator as 60% for training, 15% for validation, and 25% for test datasets. The DCAE network is pre-trained by an entirely different data set with the same amount of data.

DCAE-CNN architecture is built on Python using Tensorflow. For the training, the Adam optimizer learning rate is set to 0.001 and the batch size is fixed to 32. The training, testing and time measurements are done on a PC with 32 GB RAM, AMD Ryzen(TM) 5 3600X CPU, and AMD Radeon(TM) RX 5700 XT GPU. For the training of the designed CNN, the PlaidML tensor compiler is used since the Tensorflow framework does not have native support for Radeon(TM) architecture.

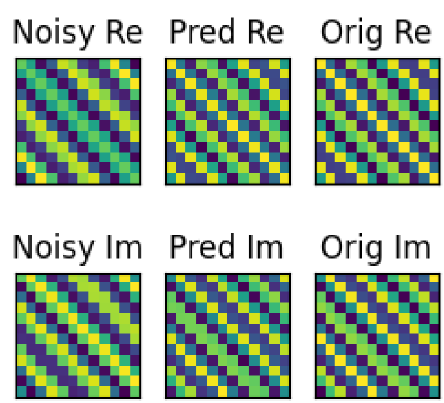

Fig. 2. DCAE Reconstruction Result of 53° at SNR = -10 dB.

An example of the reconstruction results of the DCAE is shown in Fig. 2. The *Noisy Re* and *Noisy Im* refers to real and imaginary parts of covariance matrix created under SNR = -10 dB environment for 53°, whereas *Pred Re* and *Pred Im* are the reconstruction results of DCAE. Lastly, the *Orig Re* and *Orig Im* refer to the theoretical covariance matrix for 53° and it can be seen that the "noisy smoothed covariance matrix" is reconstructed as similar to possible to the original (theoretical) one. The training process of CNN for SNR = 0 dB is shown in Fig. 3 and it can be observed that CNN trained at SNR = 0 dB environment converges pretty fast and accurately.

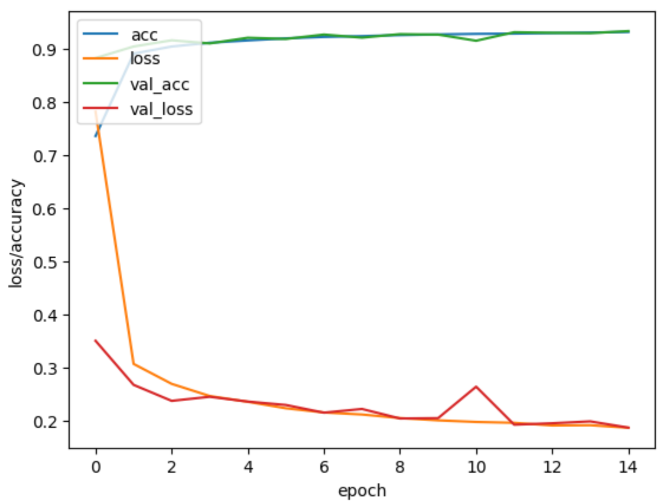

Fig. 3. CNN Training Process for SNR = 0 dB.

TABLE I
COMPARISON OF DCAE-CNN ARCHITECTURE AND MUSIC ALGORITHM.

	MUSIC		DCAE-CNN	
SNR [dB]	Accuracy	Time [ms]	Accuracy	Time [ms]
10	**96.77%**	2.18	96.69%	0.372
5	94.93%	2.178	**95.38%**	0.361
0	91.45%	2.101	**93.17%**	0.359
-5	83.45%	2.139	**87.4%**	0.369
-10	60.66%	2.052	***80.09%***	0.366

From Tab. 1. the performance comparison of the MUSIC algorithm and DCAE-CNN can be observed. Overall, the proposed method have shown better performance than the MUSIC algorithm in low SNR circumstances. However, in high SNR environments, the results are very similar. The processing time of a single DOA estimation is calculated by taking the average of 1000 single estimations. On average, the MUSIC algorithm takes 2.1 milliseconds to complete. Meanwhile, the proposed DCAE-CNN takes 0.37 milliseconds. As a result of elapsed time comparison, DCAE-CNN is approximately 6 times faster than the traditional MUSIC algorithm.

V. CONCLUSION

A new architecture based on DCAE and CNN has been proposed for DOA estimation of MIMO radars. The DCAE is used the repair the noise corrupted covariance matrix, and the CNN was employed to learn the relationship between the covariance matrix and the angle of arrival. As results indicate, the proposed DCAE-CNN architecture performs more satisfactorily than the traditional MUSIC algorithm in low SNR conditions. Moreover, the proposed architecture reduces the processing time significantly making it suitable for real-time applications. However, creating the proper neural network architecture, training the networks with considerable amount of data, and optimizing the hyper-parameters for specific antenna arrays cost some time. Nevertheless, the simulation results demonstrate that the optimized networks perform better than conventional algorithms in DOA estimation task.

REFERENCES

[1] Zhezhen Xu, Hui Li, and Kunde Yang. A modified differential beamforming and its application for doa estimation of low frequency underwater signal. *IEEE Sensors Journal*, 20(16):8890–8902, 2020.

[2] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, 1986.

[3] R. Roy and T. Kailath. Esprit-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(7):984–995, 1989.

[4] Jian Li and Petre Stoica. Mimo radar with colocated antennas. *IEEE Signal Processing Magazine*, 24(5):106–114, 2007.

[5] D.W. Bliss and K.W. Forsythe. Multiple-input multiple-output (mimo) radar and imaging: degrees of freedom and resolution. In *The Thirty-Seventh Asilomar Conference on Signals, Systems Computers, 2003*, volume 1, pages 54–59 Vol.1, 2003.

[6] Yuji Liu, Huixiu Chen, and Biao Wang. Doa estimation based on cnn for underwater acoustic array. *Applied Acoustics*, 172:107594, 01 2021.

[7] Wenli Zhu and Min Zhang. A deep learning architecture for broadband doa estimation. In *2019 IEEE 19th International Conference on Communication Technology (ICCT)*, pages 244–247, 2019.

[8] Aya Mostafa Ahmed, Udaya Sampath K. P. Miriya Thanthrige, Aly El Gamal, and Aydin Sezgin. Deep learning for doa estimation in mimo radar systems via emulation of large antenna arrays. *IEEE Communications Letters*, 25(5):1559–1563, 2021.

[9] Tie-Jun Shan, M. Wax, and T. Kailath. On spatial smoothing for direction-of-arrival estimation of coherent signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(4):806–811, 1985.

[10] Pascal Vincent, Hugo Larochelle, Y. Bengio, and Pierre-Antoine Manzagol. Extracting and composing robust features with denoising autoencoders. pages 1096–1103, 01 2008.

[11] Mahbub Hussain, Jordan J. Bird, and Diego R. Faria. A study on cnn transfer learning for image classification. In Ahmad Lotfi, Hamid Bouchachia, Alexander Gegov, Caroline Langensiepen, and Martin McGinnity, editors, *Advances in Computational Intelligence Systems*, Cham, 2019. Springer International Publishing.

[12] Aboulnasr Hassanien and Sergiy A. Vorobyov. Phased-mimo radar: A tradeoff between phased-array and mimo radars. *IEEE Transactions on Signal Processing*, 58(6):3137–3151, 2010.

[13] Li Jian. *Mimo Radar Signal Processing*. Wiley, 2009.