Deep Learning-Based Robust Imaging Exploiting 2-D Array Compressive Measurement

Md. Saidur R. Pavel, Yimin D. Zhang

Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19122, USA

***Abstract*— This paper proposes a neural network-based robust imaging method from compressive measurements exploiting a two-dimensional (2-D) array. The practical implementation of a 2-D array becomes much more complicated as the number of antennas increases due to the requirement to allocate a different radio frequency front-end circuit to each antenna. An effective solution to this problem is to compress the received signal prior to digitization at the array. In this paper, we use the maximization of the mutual information between compressed measurement and the signal locations to determine the optimal compressive measurement. A neural network-based strategy for localizing sources using these compressed measurements is then proposed. We treat neural network training as a 2-D multilabel classification problem and design an appropriate loss function to train the network. Compared to the conventional approach, the proposed neural network-based approach provides more robust performance as it does not rely on any prior knowledge of received signals and the antenna configuration.**

keywords: 2-D array, source localization, DOA estimation, compressive measurement, convolutional neural network.

I. Introduction

Direction-of-arrival (DOA) estimation, which determines the spatial spectrum of the impinging electromagnetic or acoustic waves, is one of the fundamental research areas in array signal processing. It finds importance in various applications, including wireless communication, radar, sonar, and astronomical observations [1–3]. A number of methods have been developed for DOA estimation purpose. Among them subspace-based methods [4, 5] is popularly used. Despite having numerous advantages, the subspace-based method suffers from several issues. The subspace-based methods require singular value decomposition of the covariance matrix, which is quite time-consuming in real-time source detection. Additionally, the subspace-based method relies on some prior knowledge of the array structure and signal forms. Any deviation from these prior knowledge may prevent the subspace-based method from accurately estimating DOA [33, 34]. Sparsity-inducing methods [6–13] are also widely employed to solve the DOA estimation problems. Although sparsity inducing methods offer many advantages such as immunity to the lack of source number information [35–37], robustness to noise, and limited number of snapshots [38], the complexity involved solving ℓ_p norm optimization problem can pose a challenge for real-time DOA estimation.

With the emergence of deep learning in a variety of applications, including image, speech, and array signal processing [14–17], these techniques are also widely employed for DOA estimation and source localization [21–27]. Neural network-based approach for DOA estimation is computationally more efficient compared to the convectional methods because once the network is trained it does not require performing any complicated matrix operation or solving any optimization problem. Furthermore, neural network-based method does not rely on any array configuration assumptions, making it more robust to potential imperfections such as imperfect calibration or knowledge in sensor gain, phase, position, and inter-sensor mutual coupling [28–30].

Two-dimensional DOA estimation simultaneously obtains the azimuth and elevation angles of a source and can more precisely describe the spatial characteristics of an incident signal. Therefore, 2-D DOA estimation often requires attention for source localization in real situations. The practical implementation of a 2-D array becomes complicated if the number of antennas increase. Assigning a separate radio frequency front-end circuit and high resolution analog to digital converter (ADC) to each antenna is impractical considering the hardware cost and the power consumption. To overcome this issue, we introduce a compressive sampling matrix as discussed in [18–20] to project the high dimensional signal into a low dimensional manifold. The compressive sampling matrix is optimized by maximizing the mutual information between the compressed measurements and the source locations.

Deep learning-based approaches are becoming popular for 2-D DOA estimation. The authors in [31] consider estimating the azimuth and elevation angle as a regression problem. They trained multiple convolutional neural networks (CNNs) to estimate the source location independently, and the final result was obtained by averaging the estimates of the independently trained networks. The paper in [32] employed two CNNs for estimating the azimuth and elevation angle information separately and a third for pairing them. However, as they consider 5 sources, they only require $5! = 120$ possible pairing combinations. On the other hand, we consider 16 sources in this paper, and it is unrealistic to consider $16!$ output nodes in a neural network. Considering these issues, we design a convolutional neural network (CNN) that simultaneously provides the azimuth and the elevation angles. To achieve this, we treat our location finding problem as a 2-D multilabel classification problem, where we train the network using 2-D labels containing information about azimuth and elevation angles.

Notations: We use lower-case (upper-case) bold characters to describe vectors (matrices). In particular, $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ respectively denote the transpose and conjugate transpose of a matrix or vector. $\mathrm{diag}(\cdot)$ denotes a diagonal matrix with the elements of a vector constituting the diagonal entries. $\mathrm{triu}(\cdot)$ denotes the upper triangular elements of a matrix. $\mathrm{vec}(\cdot)$ denotes vectorizing of a matrix. $\mathbb{E}(\cdot)$ denotes the expectation operation. $\circ$ is the Hadamard product operator. $\jmath = \sqrt{-1}$ denotes the unit imaginary number. $\boldsymbol{I}_M$ stands for the $M \times M$ identity matrix. In addition, $\mathcal{R}(\cdot)$ and $\mathcal{I}(\cdot)$ denote the real part and imaginary part, respectively.

II. Signal Model

We consider a rectangular (2-D) array as depicted in Fig. 1 consisting of N_1 and N_2 antennas arranged in an uniform linear fashion in each direction. The total number of physical antennas are then $N_t = N_1 \times N_2$. Consider, D uncorrelated signals are assumed to impinge on the array with azimuth angles $\boldsymbol{\phi} = [\phi_1, \phi_2, \cdots, \phi_D]^{\mathrm{T}}$ and with elevation angle $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_D]^{\mathrm{T}}$. Then the baseband received signal vector at time t is modeled as

$$\begin{aligned} \boldsymbol{x}(t) &= \sum_{d=1}^{D} \boldsymbol{a}(\phi_d, \theta_d) s_d(t) + \boldsymbol{n}(t) \\ &= \bar{\boldsymbol{A}}(\boldsymbol{\phi}, \boldsymbol{\theta}) \boldsymbol{s}(t) + \boldsymbol{n}(t), \end{aligned} \tag{1}$$

where $\bar{\boldsymbol{A}}(\boldsymbol{\phi}, \boldsymbol{\theta}) = [\boldsymbol{a}(\phi_1, \theta_1), \boldsymbol{a}(\phi_2, \theta_2), \cdots, a(\phi_D, \theta_D)] \in \mathbb{C}^{N_t \times D}$ is the presumed array manifold matrix whose column $\boldsymbol{a}(\phi_d, \theta_d) \in \mathbb{C}^{N_t}$ represents the steering vector of the dth user with an azimuth angle of ϕ_d and an elevation angle of θ_d, $\boldsymbol{s}(t) = [s_1(t), s_2(t), \cdots, s_D(t)]^{\mathrm{T}}$ represents the signal waveform vector, and $\boldsymbol{n}(t) \sim \mathcal{CN}(\boldsymbol{0}, \sigma_n^2 \boldsymbol{I}_D)$ represents the zero mean additive white Gaussian noise (AWGN) vector.

When an array sensor has calibration error, we describe the gain and phase errors of the nth sensor as $g_n = \alpha_n e^{\jmath \beta_n}$ for $n = 1, \cdots, N_t$, and denote $\boldsymbol{g} = [\alpha_1 e^{\jmath \beta_1}, \cdots, \alpha_{N_t} e^{\jmath \beta_{N_t}}]^{\mathrm{T}}$. Then, the actual array manifold $\boldsymbol{A}(\boldsymbol{\phi}, \boldsymbol{\theta})$ becomes

$$\boldsymbol{A}(\boldsymbol{\phi}, \boldsymbol{\theta}) = \mathrm{diag}(\boldsymbol{g}) \bar{\boldsymbol{A}}(\boldsymbol{\phi}, \boldsymbol{\theta}). \tag{2}$$

We then introduce a compressive sampling matrix $\boldsymbol{\Gamma} = [\boldsymbol{\gamma}_1^{\mathrm{T}}, \boldsymbol{\gamma}_2^{\mathrm{T}}, \cdots, \boldsymbol{\gamma}_M^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{M \times N_t}$ with $M \ll N_t$. $\boldsymbol{\gamma}_m \in \mathbb{C}^{1 \times N_t}$ denoting the measurement kernel for $m = 1 \cdots, M$. $\boldsymbol{\Gamma}$ is designed to be row-orthonormal, i.e., $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{H}} = \boldsymbol{I}_M$, to keep the noise power unchanged after applying the compression. Projection of the high-dimensional signal vector $\boldsymbol{x}(t)$ into the mth measurement kernel $\boldsymbol{\gamma}_m$ yields the mth compressive measurement $y_m(t) = \boldsymbol{\gamma}_m \boldsymbol{x}(t)$. Stacking the M compressed measurement vector yields $\boldsymbol{y}(t) = [y_1(t), y_2(t), \cdots, y_M(t)]^{\mathrm{T}} \in \mathbb{C}^M$, which is as

$$\boldsymbol{y}(t) = \boldsymbol{\Gamma} \boldsymbol{x}(t) = \boldsymbol{\Gamma} \boldsymbol{A}(\boldsymbol{\phi}, \boldsymbol{\theta}) \boldsymbol{s}(t) + \boldsymbol{\Gamma} \mathbf{n}(t), \tag{3}$$

where $\boldsymbol{\Gamma} \boldsymbol{A}(\boldsymbol{\phi}, \boldsymbol{\theta}) \in \mathbb{C}^{M \times D}$ represents the compressed array manifold with significantly reduced dimension.

The covariance matrix of the compressed measurement vector is then computed as

$$\begin{aligned} \boldsymbol{R} &= \mathbb{E}[\boldsymbol{y}(t) \boldsymbol{y}^{\mathrm{H}}(t)] = \boldsymbol{A} \boldsymbol{S} \boldsymbol{A}^{\mathrm{H}} + \sigma_{\mathrm{n}}^2 \boldsymbol{I}_M \\ &= \sum_{d=1}^{D} \sigma_d^2 \boldsymbol{a}(\phi_d, \theta_d) \boldsymbol{a}^{\mathrm{H}}(\phi_d, \theta_d) + \sigma_{\mathrm{n}}^2 \boldsymbol{I}_M, \end{aligned} \tag{4}$$

where $\boldsymbol{S} = \mathbb{E}[\boldsymbol{s}(t) \boldsymbol{s}^{\mathrm{H}}(t)] = \mathrm{diag}([\sigma_1^2, \cdots, \sigma_{\mathrm{D}}^2])$ is the source covariance matrix with σ_d^2 denoting the power of the dth source.

III. Optimization of compressive sampling matrix

In this section, we extend the the optimization procedure of the compressive sampling matrix discussed in [18–20] into 2-D scenario.

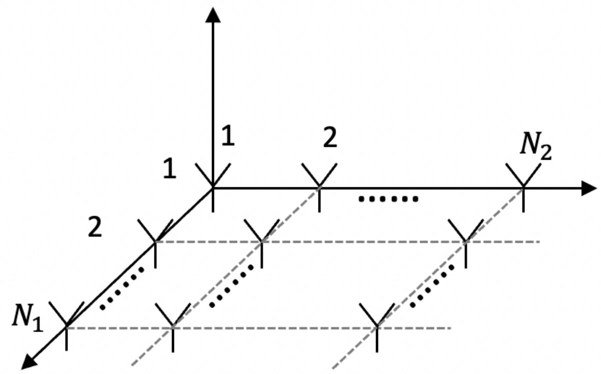

Fig. 1: Antenna Configuration

A. Probabilistic Signal Model

Consider the azimuth ϕ and the elevation θ is a random variable with a known joint probability density function (PDF) $f(\phi, \theta)$. Then, according to the law of total probability, the PDF of the compressed measurement vector can be expressed as,

$$f(\boldsymbol{y}) = \mathbb{E}_{\phi,\theta}\{\boldsymbol{y}|\phi, \theta\} = \int_{\phi} \int_{\theta} f(\boldsymbol{y}|\phi, \theta) f(\phi, \theta) d\phi d\theta \tag{5}$$

We then discretize the marginal distributions $f(\phi)$ and $f(\theta)$ into U and V angular bins with an equal width of $\Delta\bar{\phi}$ and $\Delta\bar{\theta}$ respectively. The the PDF of the compressed measurement vector can be approximated as,

$$f(\boldsymbol{y}) \approx \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} p_{u,v} f(\boldsymbol{y}|\bar{\phi}_u, \bar{\theta}_v), \tag{6}$$

where $p_{u,v} = f(\bar{\phi}_u, \bar{\theta}_v) \Delta\bar{\phi} \Delta\bar{\theta}$ denotes the probability of the (u, v)th azimuth and elevation bin pair with $\sum_{u \in \mathcal{U}, v \in \mathcal{V}} p_u p_v = 1$. $f(\boldsymbol{y}|\bar{\phi}_u, \bar{\theta}_u)$ denotes the conditional distribution of the measurement vector $\boldsymbol{y}$ given a specific azimuth, elevation pair $(\bar{\phi}_u, \bar{\theta}_v)$. Here, $\mathcal{U} = \{1, \cdots, U\}$ and $\mathcal{V} = \{1, \cdots, V\}$ represents the index set of the angular bin.

Consider a signal impinging from the angular bin pair (u, v) with a nominal azimuth, elevation pair $(\bar{\phi}_u, \bar{\theta}_v)$, the compressed measurement vector $\boldsymbol{y}$ can be expressed as

$$\boldsymbol{y}|_{\phi = \bar{\phi}_u, \theta = \bar{\theta}_v} = \boldsymbol{\Gamma}(\boldsymbol{a}(\bar{\phi}_u, \bar{\theta}_v) s(t) + \boldsymbol{n}(t) \tag{7}$$

with conditional PDF

$$f(\boldsymbol{y}|\bar{\phi}_u, \bar{\theta}_v) = \frac{1}{\pi^M |\boldsymbol{C}_{\boldsymbol{yy}|\bar{\phi}_u, \bar{\theta}_v}|} e^{-\boldsymbol{y}^{\mathrm{H}} \boldsymbol{C}^{-1}_{\boldsymbol{yy}|\bar{\phi}_u, \bar{\theta}_v} \boldsymbol{y}}, \tag{8}$$

where $\boldsymbol{C}_{\boldsymbol{yy}|\bar{\phi}_u, \bar{\theta}_v} = \boldsymbol{\Gamma}(\sigma_s^2 \boldsymbol{a}(\bar{\phi}_u, \bar{\theta}_v) \boldsymbol{a}^{\mathrm{H}}(\bar{\phi}_u, \bar{\theta}_v) + \sigma_n^2 \boldsymbol{I}) \boldsymbol{\Gamma}^{\mathrm{H}}$ is the covariance matrix of the compressed measurement for a specific azimuth, elevation pair $(\bar{\phi}_u, \bar{\theta}_v)$.

B. Optimization of the compressive sampling matrix

The compressive sampling matrix $\boldsymbol{\Gamma}$ is optimized by maximizing the mutual information between the compressed measurement vector $\boldsymbol{y}$ and the source location (ϕ, θ). The gradient of the mutual information $I(\boldsymbol{y}; \phi, \theta)$ with respect to the compressive sampling matrix $\boldsymbol{\Gamma}$ can be expressed as

$$\nabla_{\boldsymbol{\Gamma}} I(\boldsymbol{y}; \phi, \theta) = \nabla_{\boldsymbol{\Gamma}} h(\boldsymbol{y}) - \nabla_{\boldsymbol{\Gamma}} h(\boldsymbol{y}|\phi, \theta), \tag{9}$$

where $\nabla_{\boldsymbol{\Gamma}}$ denotes the gradient operator with respect to $\boldsymbol{\Gamma}$, $h(\boldsymbol{y}) = -\mathbb{E}_y\{\log[f(\boldsymbol{y})]\}$ is the differential entropy of $\boldsymbol{y}$, and $h(\boldsymbol{y}|\phi,\theta) = \mathbb{E}_{\boldsymbol{y},\phi,\theta}\{\log[f(\boldsymbol{y}|\phi,\theta)]\}$ is the conditional differential entropy of $\boldsymbol{y}$ given the source location (ϕ,θ). An extension of the mutual information gradient derived in [18, 19] into 2-D case

$$\nabla_{\boldsymbol{\Gamma}} I(\boldsymbol{y};\phi,\theta) \approx$$

$$\frac{\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}} \frac{p_{u,v}}{\left|\frac{\boldsymbol{C}_{\boldsymbol{yy}|\bar{\phi}_u,\bar{\theta}_v}}{\sigma_n^2}\right|}\left[\frac{\boldsymbol{C}_{\boldsymbol{yy}|\bar{\phi}_u,\bar{\theta}_v}}{\sigma_n^2}\right]^{-1}\boldsymbol{\Gamma}\left(\frac{\sigma_s^2}{\sigma_n^2}\boldsymbol{a}(\bar{\phi}_u,\bar{\theta}_v)\boldsymbol{a}^{\mathrm{H}}(\bar{\phi}_u,\bar{\theta}_v)+\boldsymbol{I}\right)}{\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}} p_{u,v}\left|\frac{\boldsymbol{C}_{\boldsymbol{yy}|\bar{\phi}_u,\bar{\theta}_v}}{\sigma_n^2}\right|^{-1}}$$

$$-\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\left[\frac{\boldsymbol{C}_{\boldsymbol{yy}|\bar{\phi}_u,\bar{\theta}_v}}{\sigma_n^2}\right]^{-1}\boldsymbol{\Gamma}\left(\frac{\sigma_s^2}{\sigma_n^2}\boldsymbol{a}(\bar{\phi}_u,\bar{\theta}_v)\boldsymbol{a}^{\mathrm{H}}(\bar{\phi}_u,\bar{\theta}_v)+\boldsymbol{I}\right), \tag{10}$$

where σ_s^2/σ_n^2 deontes the estimated signal to noise ration (SNR) of the input signal.

The compressive sampling matrix is then iteratively updated in a gradient ascent manner according to

$$\boldsymbol{\Gamma} \leftarrow \boldsymbol{\Gamma} + \alpha\nabla_{\boldsymbol{\Gamma}} I(\boldsymbol{y};\phi,\theta), \tag{11}$$

where $\alpha > 0$ is the step size.

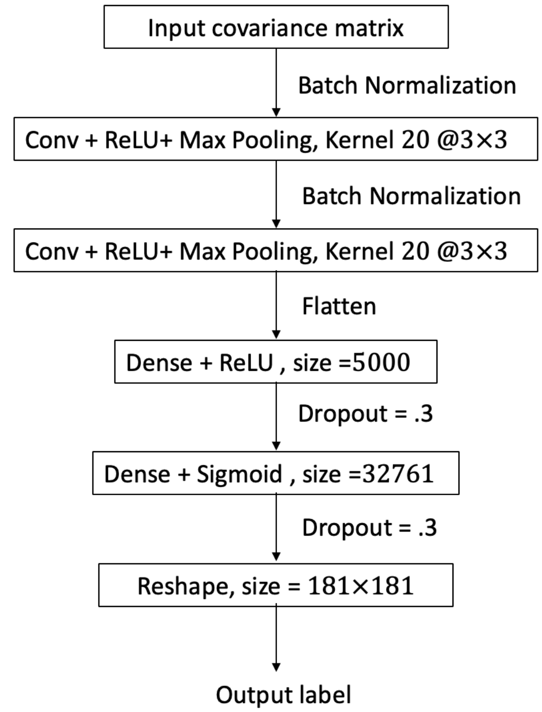

Fig. 2: Proposed Convolutional Neural Network

IV. Proposed Convolutional Neural Network

We consider a convolutional neural network as depicted in Fig. 2 for source localization. As the input to the network, we employ the covariance matrices computed from the compressed measurement vectors, and as its output, we wish to determine the azimuth and the elevation angle.

In the proposed CNN model, two convolutional layers are utilized. In each layer we consider C learnable filters with a size of $F \times F$. In each layer of the CNN, the convolutional operation between the input to this layer and the kernels produce C feature maps. The ReLU activation is used in conjunction with each convolutional layer to introduce nonlinearity. We use batchnormalization to the convolutional layers to normalize the inputs feeding to the layers. The output from a convolutional layer is followed by a max-pooling layer. The max-pooling layer divides the feature map into several non-overlapping regions, and maps the largest values from each region to its output feature map. The max-pooling operation reduce the input dimension hence lower the number of trainable parameters, add transnational invariancy and avoid trivial solutions. After the max pooling two fully connected layers are Incorporated. Since the number of nodes in these fully connected layers is very large to accommodate the dimension of the 2-D layer at the output, we adopt dropout regularization technique to prevent potential overfitting.

For a $\mathcal{L}$ layer neural network, denote $\boldsymbol{W}^{\ell}$ and b^{ℓ} as the wights and biases of the ℓth hidden layer with $\ell \in \{1,2,\cdots,\mathcal{L}\}$. $\mathcal{F}^{[\ell]}$ is a nonlinear activation function applied on the ℓth hidden layer. Then the output from the ℓth layer, $\boldsymbol{\mathcal{A}}^{[\ell]}$, can be expressed as

$$[\boldsymbol{\mathcal{A}}]^{\ell} = \mathcal{F}^{[\ell]}\left(\boldsymbol{W}^{[l]}\boldsymbol{\mathcal{A}}^{[l-1]} + b^{[l]}\right). \tag{12}$$

We consider the source localization as a (2-D) multilabel classification problem. Therefore, as the label of the network we construct a matrix of ones and zeros. Each element of the matrix represents a particular azimuth and elevation angle pair, and the objective is to make a binary decision whether a signal present in a particular pair or not. The sigmoid activation function implemented at the output ensures the output nodes of the network has values between 0 and 1. To make this kind of binary decisions, the most common loss function is the binary cross-entropy loss function as expressed in Eq. 14

$$\min_{\boldsymbol{W},\boldsymbol{b}} -\frac{1}{J}\sum_{j=1}^{J}\left[\boldsymbol{\mathcal{Y}}_j^{[i]}\log\hat{\boldsymbol{\mathcal{Y}}}_j^{[i]} + \left(1-\boldsymbol{\mathcal{Y}}_j^{[i]}\right)\log\left(1-\hat{\boldsymbol{\mathcal{Y}}}_j^{[i]}\right)\right], \tag{13}$$

where J is the number of training samples in each batch, and $\hat{\boldsymbol{\mathcal{Y}}}_j^{[i]}$ and $\boldsymbol{\mathcal{Y}}_j^{[i]}$ are, respectively, the predicted output and the actual label of the jth sample at the ith batch of the training data.

In this problem, the 2-D label contains very few numbers of ones compared to the entire 2-D grid. Therefore, the loss function defined by Eq. 14 can be very small by just making every prediction to 0. This prevents the network from learning. To combat this issue, we modified the standard binary cross-entropy loss function by employing appropriate wights f_1 and f_0 with $f_1 \gg f_0$ to emphasize predicting $1'$s.

$$\min_{\boldsymbol{W},\boldsymbol{b}} -\frac{1}{J}\sum_{j=1}^{J}\left[f_1\boldsymbol{\mathcal{Y}}_j^{[i]}\log\hat{\boldsymbol{\mathcal{Y}}}_j^{[i]} + f_0\left(1-\boldsymbol{\mathcal{Y}}_j^{[i]}\right)\log\left(1-\hat{\boldsymbol{\mathcal{Y}}}_j^{[i]}\right)\right], \tag{14}$$

V. Experimental results

We consider a rectangular 2-D array, as shown in Fig. 1, located in the XY plane. $N_1 = 10$ and $N_2 = 10$ antennas are arranged in a uniform linear fashion with half-wavelength

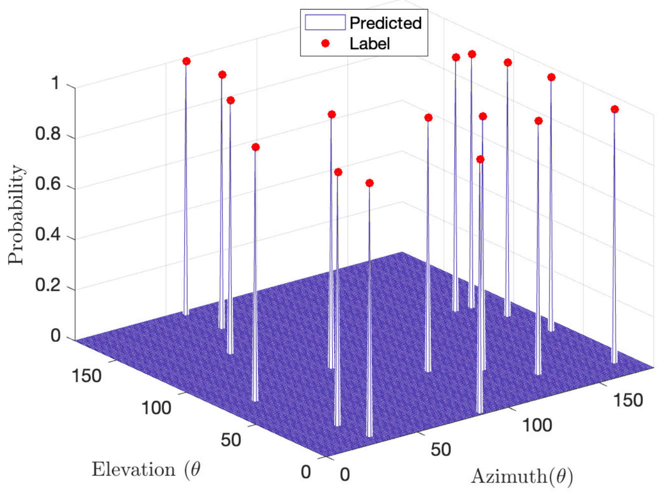

Fig. 3: Source localization performance

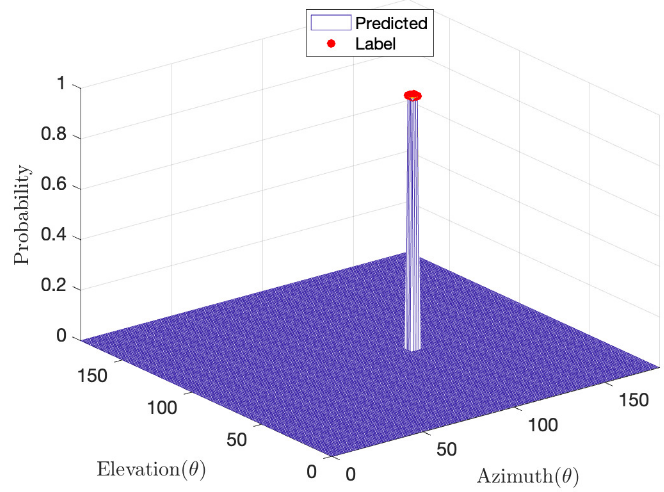

Fig. 4: Source localization performance for distributed source

separation on each axis, resulting $N_t = 100$ physical antennas. 16 uncorrelated far-field sources are assumed to impinge on the array with an input SNR of 20 dB, and the number of snapshots is $T = 100$. A compression ratio $N_t/M = 5$ is chosen to get the dimension of the compressed measurement vector $M = 20$. We generate our training dataset by considering the signals impinging from the range of azimuth angles $[0°, 180°]$ and elevations angle $[0°, 180°]$. The spatial domain is discretized with $1°$ interval rendering 181×181 direction grids. The antenna gains are independently generated from a uniform distribution between 0.5 and 1.5, whereas the phase errors are independently generated from a uniform distribution between $-6°$ and $6°$. We randomly sample 10000 observations of 16 sources from a uniform distribution to generate the training dataset.

For $N_t = 100$ and $T = 100$, the dimension of the array received signal is 100×100. After the compressive sampling matrix $\mathbf{\Gamma}$ is optimized, the compressed measurement vector is computed with a dimension of 20×100. The covariance matrix is estimated from T samples of the compressed measurement as $\hat{\boldsymbol{R}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{y}(t)\boldsymbol{y}^H(t)$. The dimension of this covariance matrix is 20×20. We then separate the real and imaginary parts and stacked them together to get a real-valued matrix with a dimension of 20×40.

This real-valued covariance matrix feed to the CNN network. The CNN is trained to learn the nonlinear relationship between the input covariance matrix and the source location. We use Two convolutional layers followed by ReLU and Max-pooling and one fully connected layer to construct the network. The sigmoid activation function is used at the output. Since we have 181×181 directional grid, we construct a matrix of dimension 181×181 containing true signal location as the label of the network. In this 2-D label, 1 is assigned for the true azimuth-elevation pair and 0 is assigned for the rest of the elements. Since we have only 16 sources, only 16 elements will have a value of 1 out of 181×181 elements. Therefore, the training loss is mostly regulated by 0's, and the network has a poor performance to predict the 1's correctly. To combat this issue, we overweight the loss corresponding to predicting 1's so that any miss-classification of 1 will be highly penalized.

Fig. 3 illustrates the performance of the proposed CNN model to an example test data. This example considers a test scenario of point sources with azimuth angles [168°,162°,27°,172°,44°,58°,99°,170°,68°,60°,119°,135°,30°, 17°,154°,86°] and elevation angles [89°,107°,27°,63°,126°, 176°, 56°,15°,85°,153°,43°,107°,8°, 73°,108°,2°]

Fig. 4 depicts the performance of the proposed model in the case of a distributed source. In both cases the proposed model successfully locate every sources.

.

VI. Conclusion

In this paper, we consider source localization in a 2-D platform. For practical implementation of a 2-D array with a large number of antennas, we optimize a compressive sampling matrix based on the maximization of the mutual information between the compressed measurements and the source location. Using the obtained compressed measurements, our proposed convolutional neural network model determined the azimuth and elevation angle of the source location. The proposed method offers improved and robust performance in the presence of imperfections.

VII. References

[1] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Simon & Schuster, 1992.

[2] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. Wiley, 2002.

[3] T. E. Tuncer and B. Friedlander (Eds.), *Classical and Modern Direction-of-Arrival Estimation.* Elsevier, 2009.

[4] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.

[5] R. Roy and T. Kailath, "ESPRIT–Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, July 1989.

[6] D. Malioutov, M. Çetin, and A. S. Willsky, “A sparse signal reconstruction perspective for source localization with sensor arrays,” *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[7] Z.-M. Liu, Z.-T. Huang, and Y.-Y. Zhou, “An efficient maximum likelihood method for direction-of-arrival estimation via sparse Bayesian learning,” *IEEE Trans. Wireless Commun.,* vol. 11, no. 10, pp. 1–11, Oct. 2012.

[8] Z. Yang, L. Xie, and C. Zhang, “Off-grid direction of arrival estimation using sparse Bayesian inference,” *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, Jan. 2013.

[9] Y. D. Zhang, M. G. Amin, and B. Himed, “Sparsity-based DOA estimation using co-prime arrays,” in *Proc. IEEE Int. Conf. Acousti., Speech, Signal Process. (ICASSP)*, Vancouver, Canada, May 2013.

[10] S. Qin, Y. D. Zhang, and M. G. Amin, “Generalized coprime array configurations for direction-of-arrival estimation,” *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1377–1390, March 2015.

[11] C. Zhou, Y. Gu, Y. D. Zhang, Z. Shi, T. Jin, and X. Wu, “Compressive sensing based coprime array direction-of-arrival estimation,” *IET Commun.*, vol. 11, no. 11, pp. 1719–1724, Aug. 2017.

[12] Z.-M. Liu and F.-C. Guo, “Azimuth and elevation estimation with rotating long-baseline interferometers,” *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2405-2419, May 2015.

[13] M. Guo, Y. D. Zhang, and T. Chen, “DOA estimation using compressed sparse array,” *IEEE Trans. Signal Process.*, vol. 66, no. 15, pp. 4133–4146, Aug. 2018.

[14] G. Hinton *et al.*, “Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups,” *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.

[15] L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu, “Road crack detection with deep convolution neural network,” in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, AZ, Sept. 2016.

[16] P. Kuang, T. Ma, Z. Chen, and F. Li, “Image super-resolution with densely connected convolutional networks,” *Appl. Intell.*, vol.49, pp.125–136, 2019.

[17] H. Xian, B. Chen, T. Yang, D. Liu, “Phase enhancement model based on supervised convolutional neural network for coherent DOA estimation,” *Appl. Intell.*, vol. 50, pp. 2411–2422, 2020.

[18] Y. Gu, Y. D. Zhang, and N. A. Goodman, “Optimized compressive sensing-based direction-of-arrival estimation in massive MIMO,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, New Orleans, LA, March 2017, pp. 3181–3185.

[19] Y. Gu and Y. D. Zhang, “Compressive sampling optimization for user signal parameter estimation in massive MIMO systems,” *Digital Signal Process.*, vol. 94, pp. 105–113, Oct. 2019.

[20] Y. D. Zhang, “Iterative learning for optimized compressive measurements in massive MIMO systems,” in *Proc. IEEE Radar Conf.*, New York, NY, March 2022.

[21] R. Takeda & K. Komatani, “Discriminative multiple sound source localization based on deep neural networks using independent location model,” in Proc. IEEE Spoken Lang. Technol. Workshop (SLT), Dec. 2016, pp. 603–609.

[22] X. Xiao, S. Zhao, X. Zhong, D. L. Jones, E. S. Chng, & H. Li, “A learning-based approach to direction of arrival estimation in noisy and reverberant environments,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Apr. 2015, pp. 2814–2818.

[23] F. Vesperini, P. Vecchiotti, E. Principi, S. Squartini, and F. Piazza, “A neural network based algorithm for speaker localization in a multiroom environment,” in Proc. IEEE 26th Int. Workshop Mach. Learn. Signal Process. (MLSP), Sep. 2016, pp. 1–6.

[24] S. Chakrabarty, & E. A. Habets (2017, October). Broadband DOA estimation using convolutional neural networks trained with noise signals. In 2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) (pp. 136-140). IEEE.

[25] S. Adavanne, A. Politis, & T. Virtanen (2018, September). Direction of arrival estimation for multiple sound sources using convolutional recurrent neural network. In 2018 26th European Signal Processing Conference (EUSIPCO) (pp. 1462-1466). IEEE.

[26] A. Rawat, R. N. Yadav, and S. Shrivastava, “Neural network applications in smart antenna arrays: A review,” *AEU-Int. J. Electron. Commun.*, vol. 66, no. 11, pp. 903–912, Nov. 2012.

[27] L. Wu, Z. Liu, and Z. Huang, “Deep convolution network for direction of arrival estimation with sparse prior,” *IEEE Signal Process. Lett.*, vol. 26, no. 11, pp. 1688–1692, Nov. 2019.

[28] S. R. Pavel, W. T. S. Chowdhury, Y. D. Zhang, D. Shen, and G. Chen, “Machine learning-based direction-of-arrival estimation exploiting distributed sparse arrays,” in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 2021.

[29] Z. Liu, C. Zhang, and P. S. Yu, “Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections,” *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 7315–7327, Dec. 2018.

[30] 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,” *Appl. Intell.*, vol. 51, pp. 4420–4433, 2021.

[31] W. Zhu, M. Zhang, P. Li, & C. Wu. (2020). Two-dimensional DOA estimation via deep ensemble learning. IEEE Access, 8, 124544-124552.

[32] Y. Yuan, S. Wu, Y. Ma, L. Huang, & N. Yuan. (2021). KR product and sparse prior based CNN estimator for 2-D DOA estimation. AEU-International Journal of Electronics and Communications, 137, 153780.

[33] M. Viberg, & A. L. Swindlehurst, (1994). Analysis of the combined effects of finite samples and model errors on array processing performance. IEEE Transactions on Signal Processing, 42(11), 3073-3083.

[34] Z. M. Liu, Z. T. Huang, & Y. Y. Zhou, (2009). Bias analysis of MUSIC in the presence of mutual coupling. IET Signal Processing, 3(1), 74-84.

[35] D. Malioutov, M. Cetin, & A. S. Willsky. (2005). A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE transactions on signal processing, 53(8), 3010-3022.

[36] I. F. Gorodnitsky, & B. D. Rao. (1997). Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm. IEEE Transactions on signal processing, 45(3), 600-616.

[37] H. Zhu, & G. B. Giannakis. (2010). Exploiting sparse user activity in multiuser detection. IEEE Transactions on Communications, 59(2), 454-465.

[38] D. Malioutov, M. Cetin, & A. S. Willsky. (2005). A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE transactions on signal processing, 53(8), 3010-3022.