
Article:
Tian, Y., Liu, S., Liu, W. orcid.org/0000-0003-2968-2888 et al. (2 more authors) (2022) Vehicle positioning with deep learning-based direction-of-arrival estimation of incoherently distributed sources. IEEE Internet of Things, 9 (20). pp. 20083-20095. ISSN 2327-4662

https://doi.org/10.1109/jiot.2022.3171820

Vehicle Positioning with Deep Learning-Based Direction-of-Arrival Estimation of Incoherently Distributed Sources

Ye Tian, *Member, IEEE*, Shuai Liu, *Member, IEEE*, Wei Liu, *Senior Member, IEEE*, Hua Chen, *Member, IEEE*, and Zhiyan Dong

***Abstract*—In this paper, a novel vehicle positioning system architecture based on direction-of-arrival (DOA) estimation of incoherently distributed (ID) sources is proposed employing massive multiple-input multiple-output (MIMO) arrays. Such an architecture with the associated signal model is more consistent with the actual array application and multipath transmission scenarios. First, an end-to-end two-dimensional (2-D) DOA estimation of ID sources utilizing a dual one-dimensional (1-D) convolutional neural network (D1D-CNN) under the deep learning (DL) framework is performed, where the normalized covariance matrix data is used for both offline training and online estimation. Then, the received SNR information is exploited to select a set of DOA estimates provided by multiple collaborative BSs for positioning. Moreover, transfer learning and an attention mechanism are employed to promote its generalization ability and achieve robustness against array perturbations. Simulation results are provided to show that the proposed method outperforms the state-of-the-art methods in terms of computational complexity, positioning accuracy and robustness against array perturbations.**

***Index Terms*—Vehicle positioning, 2-D DOA estimation, incoherently distributed (ID) sources, deep learning (DL), transfer learning, Internet of Vehicles (IoV).**

I. Introduction

HIGH-performance vehicle positioning technology for Internet of Vehicles (IoV) plays a crucial role for safe driving [1]. In most IoV application scenarios, vehicle positioning based on the global positioning system (GPS) is the most common and basic solution. However, the current commercial GPS may fail to work in some covered areas, such as tunnels and underground car park. Under such a circumstance, exploiting cooperative positioning methods to improve vehicle positioning accuracy, continuity and stability has become a key development direction of IoV.

Cooperative positioning methods usually depend on various types of sensors, such as radar, lidar and camera, to achieve vehicle positioning [2], which perform well in terms of latency and reliability. According to the positioning mechanism, it can be roughly divided into four categories. The first is based on the received signal strength (RSS) [3] or differential RSS (DRSS) [4], which can provide a simple way to achieve vehicle localization. However, its performance is not guaranteed since it depends on the prior information of fading characteristics, which is not easily available in practice. The second and the third categories are established on the time of arrival (TOA) [5] and the time difference of arrival (TDOA) [6], respectively. These two can provide good performance, provided that the synchronization of clocks among all nodes/sensors are perfect. Obviously, this is also difficult to achieve and subject to existing technologies and device level. In contrast, the fourth category, i.e., the direction of arrival (DOA) based methods, provide a more robust and efficient way for vehicle positioning, since their positioning performance only relies on the accuracy of DOA estimation, and it is easier to obtain good DOA estimates in comparison with the RSS/DRSS, TOA and TDOA based methods [7].

DOA estimation has a long research history, and a plethora of effective methods have been proposed in the past decades. According to different types of wireless channels, these DOA estimation methods can be roughly classified into three categories. The first one is built on the ideal channel without considering multipath propagation, and the point source model is employed [8]. Under this model, three interesting vehicle positioning architectures that respectively applying perturbed multiple signal classification (MUSIC) [9], sparse Bayesian learning (SBL) [10] and deep SBL network [11] have been investigated recently. However, such a channel and signal model cannot perfectly match the actual IoV and wireless communication environment, since the effect of multipath propagation does exist and cannot be ignored.

Unlike the first category, the remaining two are established on the multipath channel and a scattered/distributed source model, which can characterize the propagation environment of a realistic wireless communication process better [12]. Specifically, the second category is built on a slowly time-varying channel and coherently distributed (CD) source model, whereas the third one is established on rapidly time-varying

Manuscript received September 17, 2021. This work was supported in part by the National Natural Science Foundation of China under Grants 62001256, 61601398, and in part by the Scientific and Technological Achievement Cultivation Project of Guangxi Normal University under Grant RZ200000314. (*Corresponding Authors: Shuai Liu and Hua Chen*).

Y. Tian and H. Chen are with the Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China (e-mail: tianfield@126.com; dkchenhua0714@hotmail.com).

S. Liu is with the School of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, China (e-mail: liushuai@ysu.edu.cn).

W. Liu is with the Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield S1 3JD, U.K. (e-mail: w.liu@sheffield.ac.uk).

Z. Dong is with the Academy for Engineering and Technology, Fudan University, Shanghai 200433, China (e-mail: dongzhiyan@fudan.edu.cn).

channel and incoherently distributed (ID) source model. For DOA estimation of CD sources, many methods have been proposed, such as the modified MUSIC-like method [13], the ESPRIT [14] or unitary ESPRIT methods [15], the parallel factor (PARAFAC) analysis based method [16], and the sparse signal reconstruction (SSR) based method [17], etc. However, unfortunately these methods are not good choices for vehicle positioning in actual IoV system either, since IoV is constructed based on cellular mobile communication systems (CMCS) and it is well known that the rapidly time-varying channels are more representative of the realistic circumstances of the CMCS. Consequently, vehicle positioning based on the ID source model is more appropriate for practical applications. However, in comparison with point sources and CD sources, the DOA estimation problem of ID sources is more complicated since the signal component may span the whole observation space. Nevertheless, several DOA estimation methods for ID sources have been introduced, such as the covariance matching estimation technique (COMET) [18], the maximum likelihood (ML) method [19], the Capon method [20] and the dispersed signal parametric estimation (DISPARE) method [21]. However, it is well known that these methods are more suitable for conventional/small-scale arrays, since they require multi-dimensional search or iterative optimization. Therefore, it may be extremely difficult if not impossible for them to well balance the accuracy and real-time requirements of vehicle positioning in an actual IoV system.

In the past ten years, with rapid development of 5G and 6G wireless communication technologies, massive multiple-input multiple-output (MIMO) arrays or large-scale arrays have been widely deployed at BS. Its high degrees of freedom and large array aperture provide an easier way to improve the DOA estimation accuracy without extra hardware. In view of this, several DOA estimation methods exploiting massive MIMO arrays were successively developed. Based on the point source model, the quaternion non-circular MUSIC (QNC-MUSIC) method [22] and the higher-order propagator method (HOPM) [23] are presented; while based on the ID source model, the estimating signal parameters via rotational invariance technique (ESPRIT) [24] and the Beamspace-based approach [25] are introduced for two-dimensional (2-D) DOA estimation utilizing a massive uniform rectangle array (URA). Since these two methods are established on the multipath propagation environment, they are relatively consistent with the application scenarios of vehicles in the IoV system. However, two challenges still remain to further meet the positioning accuracy and the real-time response of the system. Firstly, these two methods mentioned above are established on an approximate ID source model and an ideal array manifold, which cannot accurately model the actual transmission and reception environment. Therefore, the DOA estimation and subsequent vehicle positioning accuracy may not be satisfactory, since various array uncertainties (such as mutual coupling and gain-phase errors) always exist in practice. Secondly, although these two methods can yield closed-form solutions, they still involves the high-dimensional matrix inversion and eigenvalue decomposition (EVD) operations. As indicated in [26], these operations incur prohibitively high computational cost for massive MIMO systems.

In order to tackle these challenges and at the same time provide a new and effective strategy for vehicle positioning in actual IoV scenarios, a deep learning (DL)-based approach is proposed in this paper, whose key idea is to obtain an efficient 2-D DOA estimation of ID sources utilizing a massive MIMO array, and then achieve vehicle positioning with preferred DOA information provided by multiple BSs. The main contributions of this paper are listed as follows:

1) As opposed to existing methods for DOA estimation of ID sources in massive MIMO systems, a DL-based approach that employs an unbiased ID system model is proposed, which can achieve 2-D DOA estimation without performing spectral search, parameter matching, high-dimensional matrix inversion and EVD, and thus computationally very attractive.
2) A dual 1-D convolutional neural network (D1D-CNN) is constructed to realize an end-to-end performance, and the real and imaginary parts of normalized array covariance matrix are exploited to train their corresponding sub-networks. Through this operation as well as information fusion, robustness in both DOA estimation and subsequent vehicle positioning is achieved.
3) Considering the influence of array perturbations present in antenna array systems, it is proposed to employ transfer learning and attention mechanism to promote the generalization ability of the network trained with a small amount of array perturbed data. As a result, a satisfactory 2-D DOA estimation performance in the presence of such perturbations is obtained.
4) Instead of using DOA estimates provided by multiple collaborative BSs directly, the received signal-to-noise ratio (SNR) information is exploited to select the best set of DOA estimates, and use them to achieve robust vehicle positioning.

The remainder of this paper is organized as follows. In Section II, the vehicle positioning system architecture is provided for 2-D DOA estimation of ID sources with a massive MIMO URA. In Section III, the proposed D1D-CNN based approach is presented in detail. In Section IV, the problem of how to exploit the received SNR information to select the set of DOA estimates to obtain robust vehicle positioning result is addressed. Simulation results are given in Section V, and conclusions are drawn in Section VI.

Notations: Upper-case (lower-case) boldface letters denote matrices (vectors). $\mathbf{I}_M$ represents the $M \times M$ identity matrix, $[\cdot]_{m,n}$ the (m,n)th element of a matrix, and $[\cdot]_m$ the mth element of a vector. Superscripts $(\cdot)^*, (\cdot)^T$ and $(\cdot)^H$ represent the conjugate, transpose and conjugate transpose operators, respectively. $\|\cdot\|$, $\mathbb{E}\{\cdot\}$ and $\mathrm{Tr}(\cdot)$ stand for the Frobenius norm, statistical expectation and trace of a matrix, respectively.

II. Positioning System and Data Model

The architecture of the employed vehicle positioning system is shown in Fig. 1(a), which mainly has three parts: 1) vehicle terminal equipped with a wireless signal transmitter, 2)

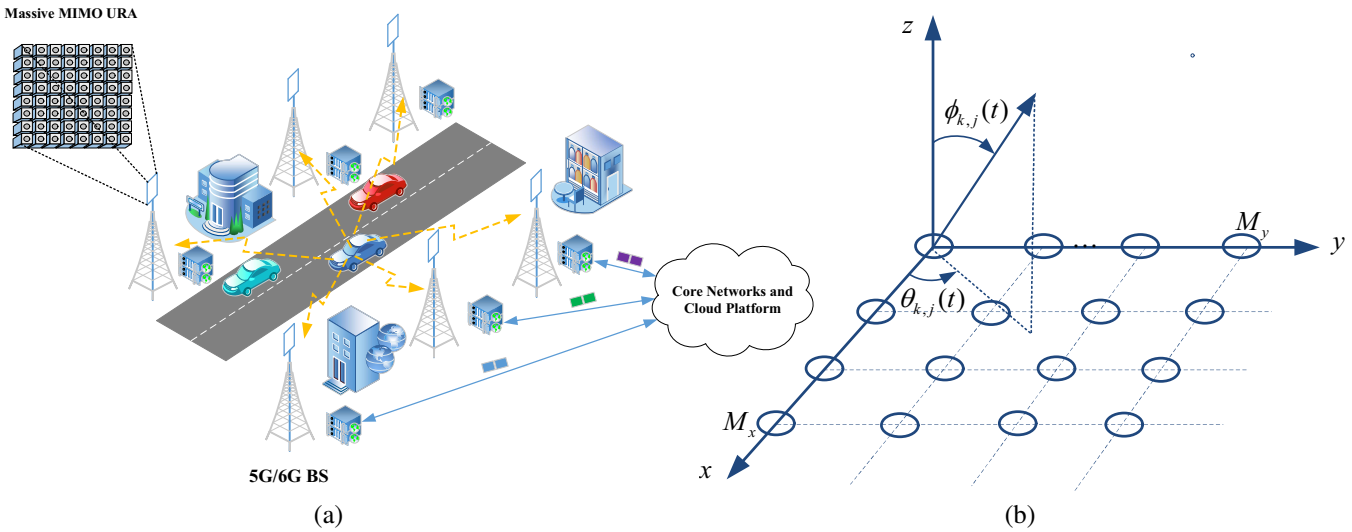


Fig. 1. (a) Illustration of the vehicle positioning system utilizing multiple collaborative BSs. (b) Massive MIMO URA geometry adopted for 2-D DOA estimation of ID sources.

fifth/sixth generation (5G/6G) BS equipped with a mssive MIMO URA, (3) the core network for information transmission and the cloud platform for providing computing services. In its operation, the vehicle first transmits a wireless positioning signal via its transmitter, the BS receives this signal and obtains the required DOA information, and then, the DOA estimates provided by multiple BSs are uploaded to the cloud platform via the core network; finally, high-accuracy vehicle positioning is achieved by triangulation using a selected set of DOA estimates.

The key to this architecture lies in robust and efficient DOA estimation. Consider a massive MIMO URA as shown in Fig. 1(b) with $M = M_x \times M_y$ antennas, where M_x and M_y are the number of antennas in the x-direction and y-direction, respectively. Taking into account the effect of rapidly time-varying channels and multipath propagation, the received signal at the antenna array can be expressed as [24]

$$\mathbf{x}(t) = \sum_{k=1}^{K} s_k(t) \sum_{l=1}^{L_k} \gamma_{k,l}(t) \mathbf{a}(\theta_{k,l}(t), \phi_{k,l}(t)) + \mathbf{n}(t), \quad (1)$$

where K denotes the number of vehicles, $s_k(t)$ the complex-valued signal transmitted by the kth vehicle, $t = 1, 2, \ldots, T$ with T denoting the number of samples; L_k is the number of multipaths of the kth vehicle, and $\gamma_{k,l}$ the complex-valued gain of the lth path from the kth vehicle; $\mathbf{a}(\theta_{k,l}(t), \phi_{k,l}(t)) \in \mathbb{C}^{M\times 1}$ is the response of the array with $\theta_{k,l}(t)$ and $\phi_{k,l}(t)$ representing the real-valued azimuth and elevation DOAs of the lth path from the kth vehicle, respectively, with $0 \le \theta_{k,l}(t) < \pi$ and $0 \le \phi_{k,l}(t) < \pi/2$; $\mathbf{n}(t) \in \mathbb{C}^{M\times 1}$ is the complex-valued additive Gaussian white noise.

Taking the origin as the phase reference point, the mth element of $\mathbf{a}(\theta_{k,l}(t), \phi_{k,l}(t))$ is given by [24]

$$\begin{aligned}[\mathbf{a}(\theta_{k,l}(t), \phi_{k,l}(t))]_m = \exp(ju\sin(\phi_{k,l}(t))[(m_x - 1) \\ \times \cos(\theta_{k,l}(t)) + (m_y - 1)\sin(\theta_{k,l}(t))]), \\ m = (m_y - 1)M_x + m_x, m_x = 1, 2, \ldots, M_x, \\ m_y = 1, 2, \ldots, M_y, \end{aligned} \quad (2)$$

where $u = 2\pi d/\lambda$ with d and λ respectively being the distance of two adjacent antennas and the wavelength of the carrier. According to the scattering characteristics [27], the azimuth and elevation DOAs of the lth ray at time instant t can be expressed as

$$\theta_{k,l}(t) = \theta_k + \tilde{\theta}_{k,l}(t), \quad (3)$$

$$\phi_{k,l}(t) = \phi_k + \tilde{\phi}_{k,l}(t), \quad (4)$$

where θ_k and ϕ_k are the nominal azimuth and elevation DOAs for the kth vehicle, respectively, and $\tilde{\theta}_{k,l}(t)$ and $\tilde{\phi}_{k,l}(t)$ are their corresponding random angular deviations with zero mean and standard deviations σ_{θ_k} and σ_{ϕ_k}.

For simplicity and also for comparison in later simulations, the same assumptions as in [24] and [25] are adopted here. Subsequently, the array covariance matrix $\mathbf{R}$ is given by

$$\mathbf{R} = \mathbb{E}\{\mathbf{x}(t)\mathbf{x}^H(t)\} = \mathbf{R}_s + \sigma_n^2 \mathbf{I}_M. \quad (5)$$

It should be emphasized here that detailed expressions of $\mathbf{R}$ and $\mathbf{R}_s$ are not given, since this paper attempts to establish a generalized approach for DOA estimation of ID sources without knowing the specific expressions of $\mathbf{R}$ and $\mathbf{R}_s$. Nevertheless, $\mathbf{R}$ and $\mathbf{R}_s$ contain all the required information about the vehicle, including 2-D nominal DOA, angular spreads, signal power emitted by the vehicle, etc. In what follows, we will demonstrate how to obtain a satisfactory 2-D DOA estimation performance efficiently with $\mathbf{R}$ and further achieve an improved vehicle positioning result.

III. The Proposed Method

ESPRIT [24] and Beamspace [25] are two representative approaches for 2-D DOA estimation of ID sources with massive MIMO arrays, which respectively employ approximate covariance matrix $\mathbf{R}_E = \mathbf{A}\mathbf{\Lambda}_s\mathbf{A}^H + \sigma_n^2\mathbf{I}_M$ and dimension-reduced covariance matrix $\mathbf{R}_B = \mathbf{W}\mathbf{A}\mathbf{\Lambda}_s\mathbf{A}^H\mathbf{W}^H + \sigma_n^2\mathbf{I}_{PM_y}$ for DOA estimation. These two approaches can yield good performance, provided that the angular spreads are sufficiently small and there are no array perturbation or model errors.

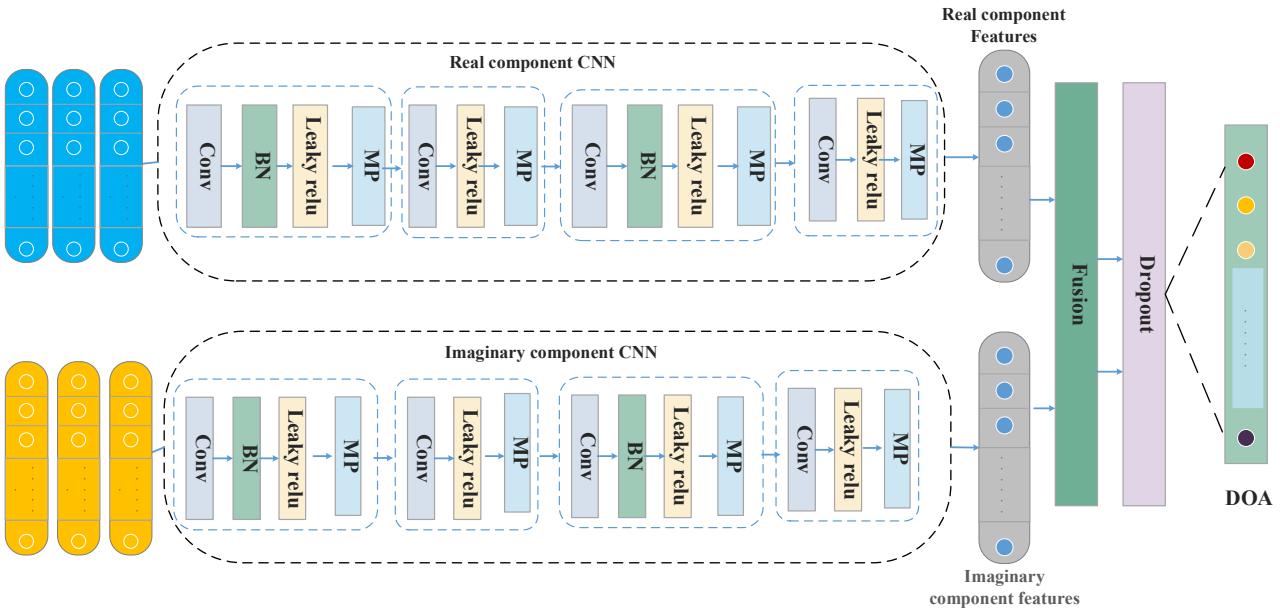


Fig. 2. Dual one-dimensional (1-D) convolutional neural network (D1D-CNN) structure for DOA estimation.

However, such conditions may not be satisfied in practice, and as a result, their performance will degrade. In the following, a D1D-CNN based approach is presented to tackle these issues.

A. Basic Configuration of D1D-CNN

CNN is a special type of deep learning model. By using local connections and shared weights, it can reduce the number of neural network parameters. CNN usually consists of convolution, pooling, activation function, and batch normalization layers. A deep CNN architecture is built by stacking several convolution layers. Given the capability of high level feature learning, a dual one-dimensional (1-D) convolutional neural network (D1D-CNN) as shown in Fig. 2 is constructed, which consists of an input module, a real component CNN feature learning branch, an imaginary component CNN feature learning branch, and a real-imaginary component feature fusion module. Each feature learning branch uses four convolution (Conv) layers, two batch normalization (BN) layers, four activation function layers and four max pooling (MP) layers. Note that such a configuration is decided through extensive simulations to reach a good balance between accuracy and complexity.

Different from the existing DL-based methods studied in [28] and [29], the array covariance elements are used as the input of the D1D-CNN, which have a fixed dimension in comparison with the time-domain data, and provide a particularly efficient representation of the data when a sufficient number of samples are available. In the process of DOA estimation, the covariance matrix elements are first normalized and transformed into one-dimensional vector, which is then divided into real and imaginary parts. The real part contains the noise variance components, while the imaginary part does not. Subsequently, discriminative real and imaginary features are extracted and fused. Finally, the fused features are fed into the fully connected layers to achieve 2-D DOA estimation. The output of D1D-CNN is the real-valued azimuth and elevation DOAs $\alpha(\theta, \phi)$. Let $\mathbf{x}_r$ and $\mathbf{x}_i$ indicate the input real and imaginary components of $\mathbf{R}$, respectively. Then the output of the D1D-CNN is given by

$$\alpha(\theta, \phi) = f(\mathbf{x}_r, \mathbf{x}_i, w) = f^{(n-1)}(f^{(n-2)}(\dots f^{1}(\mathbf{x}_r, \mathbf{x}_i))), \tag{6}$$

where $f(\cdot)$ represents the output of the model, n and w are the number of layers and the weights of D1D-CNN, respectively.

In particular, when the input value of rectified linear unit (ReLU) is negative, the output as well as its first derivative is always zero, which directly yields that the neuron cannot update the parameters any more. To avoid this problem, the Leaky ReLU is used to keep a small gradient value when the input is less than zero [30]. The function of the Leaky ReLU is given by

$$\text{LeakyReLU}(x) = \max(0, x) + \gamma \min(0, x), \tag{7}$$

where γ is a very small constant. The network adopts the BN layer to normalize the data and MP layer to reduce the number of parameters and simultaneously facilitate convergence of the model. For the fully connected layers, we apply a dropout technique to randomly drop units from the neural network during training, which can effectively prevent the overfitting issue. For the fully connected output layer, the Sigmoid function is selected as the activation function of the processing layer to realize classification and obtain azimuth and elevation angle estimation.

B. Real and Imaginary Component Features Fusion

Let $\mathbf{R}_r$ and $\mathbf{R}_i$ denote the learned features for the real and imaginary component CNNs, respectively. In order to balance these two features efficiently and effectively, the summation and maximization strategy is applied. The summation fusion

Algorithm 1: D1D-CNN Training

Input: $\{\mathbf{R}, \theta, \phi\}$, training epochs
Output: All trainable parameters in each layer
1: Initialize all weights
2: **While** epoch<epochs **do**
Stage 1:
Train the real and imaginary component CNNs; sample minibatch of $\{\mathbf{R}, \theta, \phi\}$ from training set, where the parameter matrices $\mathbf{W}_r$, $\mathbf{W}_i$ associated with real CNN and imaginary CNN are all learned by minimizing the following loss function:
$L_{oss} = \frac{1}{KN}\sum_{i=1}^{N}\sum_{k=1}^{K}\left\{\left\|\theta_{k,i} - \hat{\theta}_{k,i}\right\|^2 + \left\|\phi_{k,i} - \hat{\phi}_{k,i}\right\|^2\right\}$.
3: **End while**
4: **While** epoch<epochs **do**
Stage 2:
Merge the separated trained models, freeze the weights of real and imaginary CNNs and fine-tune fully connected layers, where the parameter matrix $\mathbf{W}_f$ is learned by minimizing the loss function:
$L_{oss} = \frac{1}{KN}\sum_{i=1}^{N}\sum_{k=1}^{K}\left\{\left\|\theta_{k,i} - \hat{\theta}_{k,i}\right\|^2 + \left\|\phi_{k,i} - \hat{\phi}_{k,i}\right\|^2\right\}$.
5: **End while**

aims at performing an element-wise sum of the two feature representations, i.e.,

$$F_{\text{sum}} = \mathbf{R}_r + \mathbf{R}_i, \tag{8}$$

while the maximization fusion computes the element-wise maximum value, i.e.,

$$F_{\max} = \max\left(\mathbf{R}_r + \mathbf{R}_i\right). \tag{9}$$

The whole fusion layer is formulated as

$$F_{\text{fusion}} = \lambda_1 F_{\text{sum}} + \lambda_2 F_{\max}, \tag{10}$$

where λ_1 and λ_2 are in the range [0,1] with $\lambda_1 + \lambda_2 = 1$.

C. Network Training

After establishing the DOA estimation framework, an offline learning scheme is adopted to train the D1D-CNN. The parameters of the model are randomly initialized and trained by an error back propagation algorithm. It is difficult to optimize the parameters in the D1D-CNN branch. In the experiment, the real and imaginary component CNNs are first trained using real and imaginary data individually. Then, a fine-tune strategy is employed, which loads the weights from pre-trained two branches to greatly reduce the time of numerical calculation. As described in the D1D-CNN model, two different feature learning components are first trained on real and imaginary data sets with a large learning rate. Then, the two pre-trained models extract the corresponding features from real and imaginary training data pairs. When the two component branches are merged, the fusion features will be fine-tuned with a smaller learning rate.

For obtaining accurate estimates of θ_k and ϕ_k, the mean square error (MSE) based loss function is used, which can be formulated as

$$L_{oss} = \frac{1}{KN}\sum_{i=1}^{N}\sum_{k=1}^{K}\left\{\left\|\theta_{k,i} - \hat{\theta}_{k,i}\right\|^2 + \left\|\phi_{k,i} - \hat{\phi}_{k,i}\right\|^2\right\}, \tag{11}$$

where N is the number of samples in second-order statistics domain, $\hat{\theta}_{k,i}$ and $\hat{\phi}_{k,i}$ are the estimated values in the training process. The Adaptive moment estimation (Adam) optimizer with a batch size of 100 is employed [31].

Once the network is trained, it can be used directly to estimate θ and ϕ. In the procedure, online deployment is conducted by feeding the raw pair data into the model without requiring iterations and parameter pairing. The stages of D1D-CNN training are illustrated in Algorithm 1.

D. Enhanced D1D-CNN for DOA Estimation with Array Perturbations

To the best of our knowledge, all the existing DOA estimation methods for ID sources are based on the ideal array manifold without considering array perturbations, such as mutual coupling, gain-phase errors and array position uncertainties. It has been demonstrated that such array perturbations can degrade the estimator's performance substantially [32]. Although there have been some array calibration methods available, such as the eigenstructure-based method [33], the active calibration method [34], and the covariance difference based method [35], most of them are based on the point source model, uniform linear array and subspace theory, and it is very difficult to extend them to deal with the scenario considered in this work. Fortunately, based on the characteristics of DL, this issue can be tackled by enhancing the constructed D1D-CNN in the following two aspects:

i) Transfer learning (TL) is used to promote the generalization ability of already trained D1D-CNN with a small amount of perturbed array data, which aims to align the features across array perturbations and reduce the distribution divergence. Instead of using the fine-tuning strategy mentioned above, a more suitable and effective TL approach is exploited here. Let D_i and D_p represent the feature of data with ideal array manifold and that with array perturbations, respectively. Then the learning tasks are represented as $L_i = \{\alpha_i, P(\alpha_i|\mathbf{r}_i)\}$ and $L_p = \{\alpha_p, P(\alpha_p|\mathbf{r}_p)\}$, respectively, where α denotes the DOA information, and $\mathbf{r}$ the normalized array covariance matrix vector. For DOA estimation, it is important to match the conditional probability distributions $P(\alpha_i|\mathbf{r}_i)$ and $P(\alpha_p|\mathbf{r}_p)$ across array perturbations, since the conditional distributions reflect the difference information. Thus the TL can be regarded as learning a feature space to align different distributions. Based on the statistics, the posterior probability distribution $P(\alpha|\mathbf{r})$ can be replaced by the conditional distribution $P(\mathbf{r}|\alpha)$, since the posterior distribution is quite complicated. The objective function to measure the divergences of the conditional distributions across data with array perturbations can be defined as below

$$J = \min\sum\left\|\frac{1}{N_i^{\alpha}}\sum_{x_i\in D_i^{\alpha}}F(x_i) - \frac{1}{N_p^{\alpha}}\sum_{x_p\in D_p^{\alpha}}F(x_p)\right\|, \tag{12}$$

where $F(\cdot)$ stands for the CNN mapping function to project data into the common feature subspace, N_i^{α} and N_p^{α} denote the same DOA number of the source domain and the target domain, respectively. The function is optimized to obtain the mapping function, and thus the conditional distributions of features across ideal and perturbation arrays are aligned.

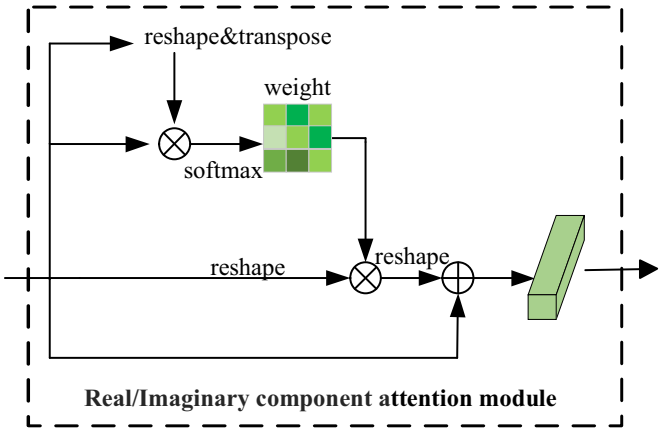


Fig. 3. The adopted plug and play attention module.

ii) Attention mechanism is exploited to increase the robustness of the D1D-CNN against array perturbations. Specifically, we design plug and play attention module to add soft weights on feature maps from the feature learning module, as shown in Fig. 3. After processing by the real and imaginary component CNNs, the resulting real and imaginary features have redundant information. The real component attention aims to refine the features of real component features, which model the relationships between each channel of the feature map to learn a real component weight, and then multiply it to the real component features. Given real data, 1-D CNN will produce a feature map $F \in \mathbb{R}^{D\times C}$, where D and C denote the dimension of the feature map. By performing a matrix multiplication operation, the real component feature weight map $\mathbf{X}$ is obtained through a softmax operation:

$$x_{ji} = \frac{\exp(F_i \cdot F_j)}{\sum_{i=1}^{C} \exp(F_i \cdot F_j)}, \tag{13}$$

where x_{ji} is used to calculate the ith channel's impact on the jth channel. After that, F and the real component feature weight map $\mathbf{X}$ are multiplied together. Here, an element-wise sum operation with F is performed to obtain the final real component feature map F^{real}, i.e.,

$$F_j^{real} = \sum_{i=1}^{C} (x_{ji} F_i) + F_j. \tag{14}$$

The imaginary component attention module is designed in the same way. Finally, we obtain the final imaginary component feature map F^{imag}.

IV. Vehicle Positioning Based on DOA Estimation

Once the 2-D nominal DOAs have been estimated by multiple collaborative BSs, the location of the vehicle can then be obtained. Let $(P_{x,\bar{m}}, P_{y,\bar{m}}, P_{z,\bar{m}})$ and (X_k, Y_k, Z_k) denote the known position of the $\bar{m}$th BS and the unknown position of the kth vehicle, respectively. Then the following relationships hold,

$$\tan\theta_{k,\bar{m}} = \frac{P_{y,\bar{m}} - Y_k}{P_{x,\bar{m}} - X_k},\ \tan\phi_{k,\bar{m}} = \frac{P_{y,\bar{m}} - Y_k}{P_{z,\bar{m}} - Z_k}. \tag{15}$$

According to the localization method of triangulation and the widely used average operation, the coordinates of the kth vehicle can be calculated by equations at the bottom of this page, where $\bar{M}$ denotes the number of available BSs, $\bar{m}, \tilde{m} \in [1, \bar{M}]$, $(\theta_{k,\bar{m}}, \phi_{k,\bar{m}})$ and $(\theta_{k,\tilde{m}}, \phi_{k,\tilde{m}})$ are 2-D nominal DOA estimates provided by the $\bar{m}$th BS and the $\tilde{m}$th BS, respectively.

However, due to influence of the vehicle location and wireless transmission environment in the actual IoV system, not all DOA information provided by the BSs is reliable. Under this situation, exploiting some auxiliary/prior information to select the best set of DOA estimates for vehicle positioning holds the key for a final satisfactory result. In this paper, the received signal-to-noise ratio (SNR) at the BS is used to select DOA estimates, since it is well known that the DOA estimation accuracy increases with the received SNR.

For simplicity, a *priori* knowledge of the noise variance σ_n^2 is assumed. For the case with unknown σ_n^2, one can estimate it first according to [36]. Since the 2-D nominal DOAs have been estimated, $\mathbf{\Lambda}_s$ under the approximated model can be reconstructed as

$$\hat{\mathbf{\Lambda}}_s = \hat{\mathbf{A}}^{\dagger}(\mathbf{R} - \hat{\sigma}_n^2 \mathbf{I}_M)(\hat{\mathbf{A}}^H)^{\dagger}, \tag{16}$$

where $\hat{\mathbf{A}}$ is the estimate of $\mathbf{A}$. Consequently, the received SNR corresponding the kth vehicle is given by

$$\mathrm{SNR}_k = 10 \times \log_{10}\left(\hat{\mathbf{\Lambda}}_s(k,k)/\sigma_n^2\right), k \in [1, K]. \tag{17}$$

The proposed vehicle positioning method is summarized in Algorithm 2.

Several Remarks:

$$X_k = \frac{2}{\bar{M}(\bar{M}-1)} \sum_{\bar{m}=1}^{\bar{M}} \sum_{\substack{\tilde{m}=1 \\ \tilde{m}\neq\bar{m}}}^{\bar{M}} \frac{P_{y,\bar{m}} - P_{y,\tilde{m}} + P_{x,\tilde{m}} \tan\theta_{k,\tilde{m}} - P_{x,\bar{m}} \tan\theta_{k,\bar{m}}}{\tan\theta_{k,\tilde{m}} - \tan\theta_{k,\bar{m}}},$$

$$Y_k = \frac{2}{\bar{M}(\bar{M}-1)} \sum_{\bar{m}=1}^{\bar{M}} \sum_{\substack{\tilde{m}=1 \\ \tilde{m}\neq\bar{m}}}^{\bar{M}} \frac{(P_{x,\bar{m}} - P_{x,\tilde{m}}) \tan\theta_{k,\tilde{m}} \tan\theta_{k,\bar{m}} + P_{y,\tilde{m}} \tan\theta_{k,\bar{m}} - P_{y,\bar{m}} \tan\theta_{k,\tilde{m}}}{\tan\theta_{k,\bar{m}} - \tan\theta_{k,\tilde{m}}},$$

$$Z_k = \frac{2}{\bar{M}(\bar{M}-1)} \sum_{\bar{m}=1}^{\bar{M}} \sum_{\substack{\tilde{m}=1 \\ \tilde{m}\neq\bar{m}}}^{\bar{M}} \frac{P_{y,\bar{m}} - P_{y,\tilde{m}} + P_{z,\tilde{m}} \tan\phi_{k,\tilde{m}} - P_{z,\bar{m}} \tan\phi_{k,\bar{m}}}{\tan\phi_{k,\tilde{m}} - \tan\phi_{k,\bar{m}}}.$$

Algorithm 2: The Proposed Vehicle Positioning Method

Input: D1D-CNN, covariance matrix $\mathbf{R}$, K, $\bar{M}$
Output: $(X_k, Y_k, Z_k), k = 1, \ldots, K$
1: Load the D1D-CNN which has been trained thoroughly.
2: Utilize $\mathbf{R}$ to update the output $(\theta_{k,\bar{m}}, \phi_{k,\bar{m}}), k \in [1, K], \bar{m} \in [1, \bar{M}]$.
3: Perform EVD on $\mathbf{R}$ to obtain eigenvalues and eigenvectors.
4: Estimate the array perturbation based on the orthogonality between the signal and noise subspaces.
5: Calculate the received SNR according to (17) after compensating for this array perturbation.
6: Select the DOA estimates provided by BSs with a large received SNR.
7: Obtain the vehicle location with selected DOA estimates.
8: **return:** $(X_k, Y_k, Z_k), k = 1, \ldots, K$.

- In practical applications, if the difference of the received SNR among different BSs is small, a simple average operation is used to realize vehicle positioning. Otherwise, if the received SNR is very different, only $\bar{M} \geq 2$ DOAs provided by the BSs with a large received SNR are chosen. This strategy can yield an improved positioning accuracy and robustness on vehicle positioning.
- When there are array perturbations, one can exploit some approach to estimate it first with the estimated 2-D DOAs, such as the approach in [35], and then obtain the received SNR after compensating for such array perturbations.
- Estimation of the received SNR (17) is based on the approximated model, which only holds when the angular spreads are sufficiently small. In case of large angular spreads, (17) is invalid, and there have been no reports for estimating the received SNR under this circumstance. Through our analysis, it may be feasible to accomplish this through a DL-based approach, but further research is needed to solve the problem.
- The proposed method can achieve 2-D DOA estimation with a computational complexity of $O(M^2T + M^2K)$. As a comparison, the complexities of ESPRIT [24] and Beamspace [25] are $(O(M^2T + M^3 + M^2K))$ and $(O((PM_x)^3 + (PM_y)^2N + PM_xN + PM_yK^2))$, respectively. Fig. 4 provides an intuitive comparison on complexities in logarithm scale for different algorithms with $M_x = M_y = \sqrt{M}$, $P = M_x - 2, K = 2, N = 200$, from which we can clearly see that the computational complexity of the proposed one is lower than those of the other two methods when $M > 100$, making it well suited for massive MIMO systems.

V. Simulation Results

In this section, numerical simulations are carried out to demonstrate the effectiveness of the proposed vehicle positioning method. Firstly, the 2-D nominal DOA estimation performance of the proposed D1D-CNN based method is evaluated and compared with that of the ESPRIT [24], the Beamspace [25] based approaches and the Cramér-Rao bound (CRB) [24]. Secondly, the vehicle positioning performance of the proposed algorithm with 2-D DOA estimates is shown. Each BS in the positioning system is equipped with a massive URA, which consists of $M = M_x \times M_y$ antennas, and the distance between adjacent sensors is $d = \lambda/2$. All the wireless

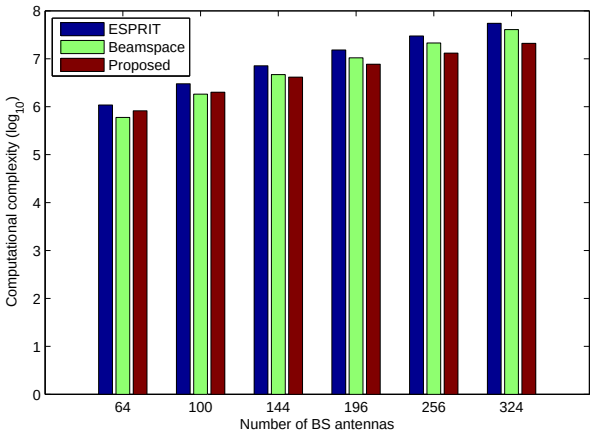


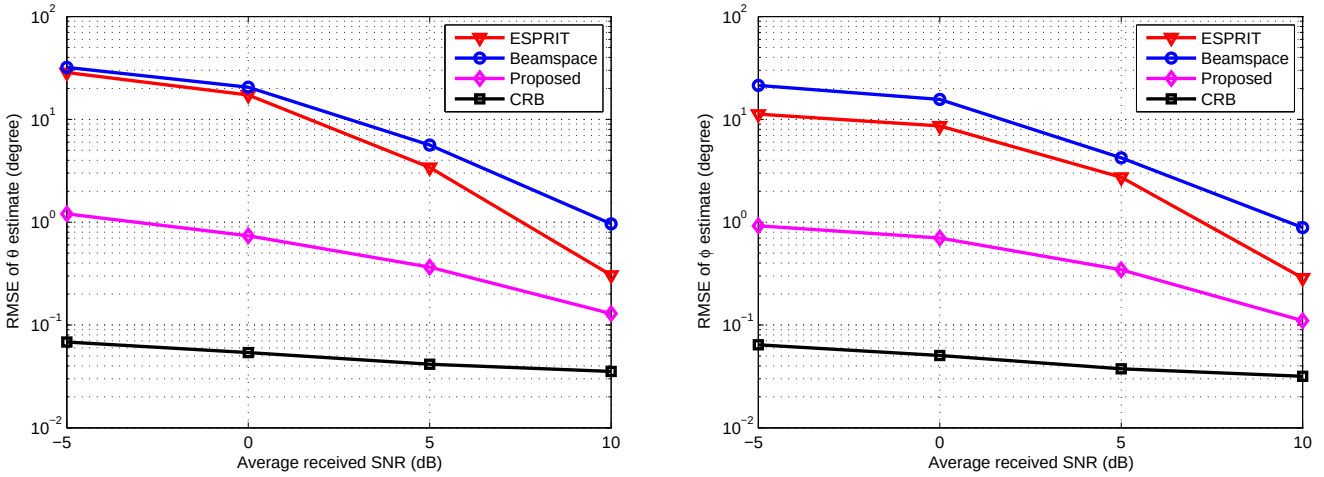
Fig. 4. Computational complexities versus number of BS antennas.

positioning signals transmitted by the vehicle are BPSK-modulated. Without loss of generality, a Long Term Evolution (LTE) uplink system is considered here, which operates at 2 GHz, with a channel bandwidth of 2.5 MHz, and the sampling rate is 3.84MHz. The training samples of the D1D-CNN are fromed by the direction information $\{\theta_{k,i}, \phi_{k,i}\}$, where $\theta_{k,i}$ and $\phi_{k,i}$ are randomly distributed in the range $[0, \pi)$ and $[0, \pi/2)$, respectively. In addition, the training set comprises 50000 examples in the simulation, while a validation set with 4000 samples for each SNR.

A. Performance of DOA Estimation without Perturbations

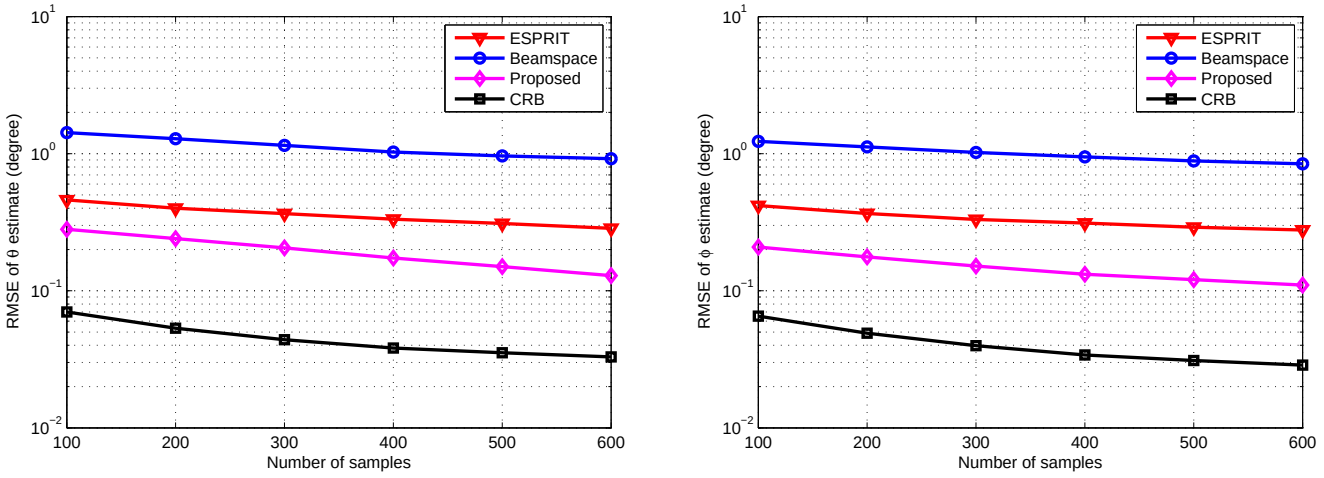
1) RMSE versus the Average Received SNR: In the first simulation, the performance of the proposed D1D-CNN based method versus the average received SNR is tested with an ideal array manifold, whose root mean square error (RMSE) curves are shown in Fig. 5 with two sources located at $\{\theta_1 = 10^\circ, \phi_1 = 30^\circ\}$ and $\{\theta_2 = 50^\circ, \phi_2 = 40^\circ\}$. The number of multipaths is $L_k = 50$, the ray gain variance, the azimuth angular spreads and the elevation angular spreads are $\sigma^2_{\gamma_k} = 1, \sigma_{\theta_k} = 1^\circ, \sigma_{\phi_k} = 1^\circ$, respectively, $k = 1, 2$. $M_x = M_y = 10$, the number of samples is fixed at 500, and SNR varies from -5 dB to 10 dB. From the result, it can be seen that the proposed method yields a better result than the ESPRIT and the Beamspace based algorithms. However, there is a clear gap between the estimation performance of the proposed method and the CRB, which indicates that there is still much room for improvement in the adopted network model and training set, which can be a topic of our further research.

2) RMSE versus the Number of Samples: In the second simulation, we evaluate the impact of the number of samples on DOA estimation performance of different algorithms, whose RMSE curves are illustrated in Fig. 6. SNR=10 dB and the number of samples is varied from 100 to 600, and the other conditions are the same as in the first simulation. It can be seen from Fig. 6 that the performance of all algorithms improves monotonically with the number of samples. In particular, the proposed method still performs better than the other two algorithms in the whole sample region.



(a) (b)

Fig. 5. RMSEs of 2-D nominal DOA estimation versus the average received SNR without considering array perturbations, $M = 100, N = 500$. (a) Estimation of θ. (b) Estimation of ϕ.



(a) (b)

Fig. 6. RMSEs of 2-D nominal DOA estimation versus the number of samples without considering array perturbations, $M = 100$, average received SNR=10 dB. (a) Estimation of θ. (b) Estimation of ϕ.

3) RMSE versus the Number of BS Antennas: In the third simulation, the performance is examined with different number of BS antennas. The simulation conditions are also the same as in the first simulation, except that SNR is fixed at 10 dB, and the number of BS antennas M varies from 36 to 196, with $\sqrt{M} = M_x = M_y$. The result is shown in Fig. 7, from which we can observe that the RMSEs of the proposed method decrease rapidly as M increases and are also lower than the other two compared algorithms. On the other hand, it can also be clearly seen that the performance of the other two algorithms is influenced by M significantly more, and their performance is far worse when $M < 100$, which is consistent with the simulation results in [24] and [25]. As explained in [24], the signal subspace $\hat{\mathbf{E}}_s$ and $\mathbf{A}$ might not be in the same subspace in cases of small M, which directly yields that their related performance is not satisfactory. Since the proposed solution is not established under the subspace framework, it avoids this mismatch problem effectively.

4) RMSE versus the Number of Multipaths: In the fourth simulation, the performance of the proposed method with different number of multipaths is investigated. The simulation parameters are the same as those in the third simulation, except that M is fixed at 100, and L_k is varied from 20 to 60 in steps of 10. As can be seen in Fig. 8, the RMSE of the proposed method is almost invariant with the number of multipaths, which effectively validates the robustness of the proposed method in a multipath scenario. Note that the number of multipaths is normally time-varying, especially in an urban environment, and therefore, this robust performance will play an important role for safe and reliable autonomous driving.

B. Performance of DOA Estimation with Mutual Coupling

1) RMSE versus the Average Received SNR: In the fifth simulation, the performance of the proposed method in the

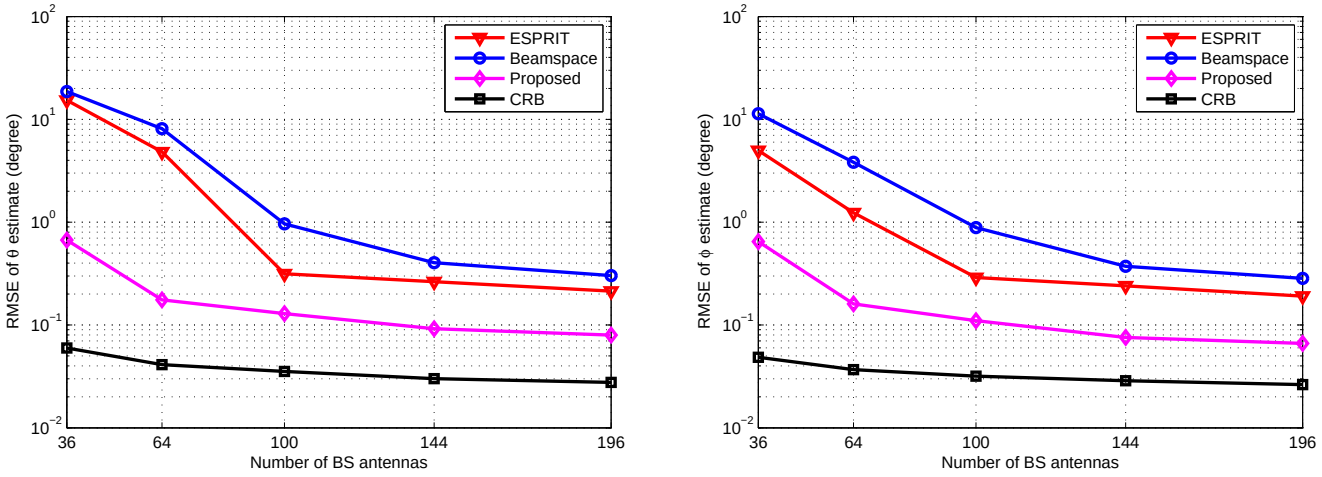


Fig. 7. RMSEs of 2-D nominal DOA estimation versus the number of BS antennas without considering array perturbations, $N = 500$, average received SNR=10 dB. (a) Estimation of θ. (b) Estimation of ϕ.

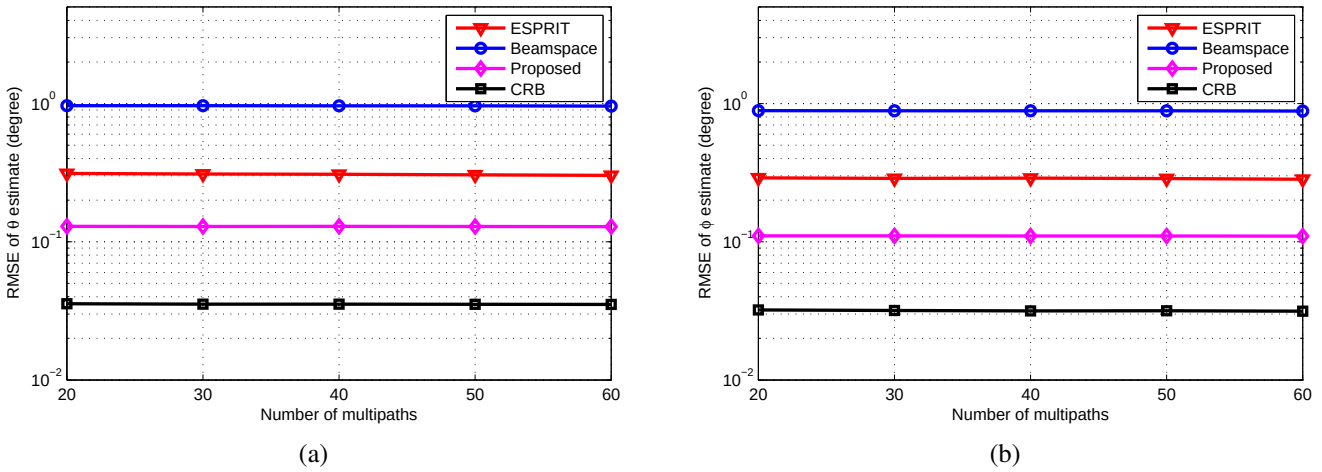


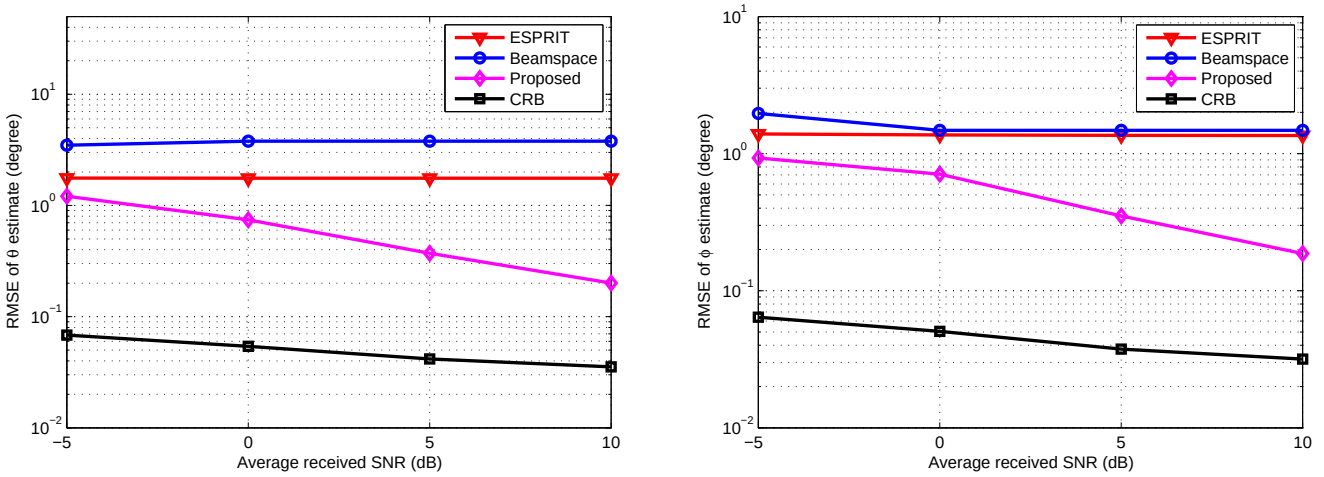
Fig. 8. RMSEs of 2-D nominal DOA estimation versus the number of multipaths without considering array perturbations, $M = 100$, $N = 500$, average received SNR=10 dB. (a) Estimation of θ. (b) Estimation of ϕ.

presence of array perturbations is studied, where the array mutual coupling is taken into account as an example. The mutual coupling coefficients between two adjacent sensors with distance $\lambda/2$ and λ are $c_1 = 0.7e^{-j\pi/4}$ and $c_2 = 0.3e^{j\pi/10}$, respectively, while the mutual coupling coefficient between other antennas are assumed to be relatively small. TL and attention mechanism are utilized to enhance the trained D1D-CNN. Except for the mutual coupling configuration, the remaining simulation conditions are the same as in the first simulation. It should be noted that the ESPRIT and Beamspace based methods are established on an ideal array model without considering mutual coupling. Due to the model mismatch problem, it can be seen from Fig. 9 that their performance drops sharply, and barely improves as the SNR increases, in the presence of mutual coupling. In contrast, the proposed D1D-CNN based method can still reach a satisfactory DOA estimation result, which effectively validates the robustness of the proposed solution. Note that mutual coupling always exists and cannot be ignored or perfectly calibrated in practical applications of massive MIMO arrays, which implies that the proposed solution would be a better choice in practice.

2) RMSE versus the Number of Samples: In the sixth simulation, the average received SNR is set to 10 dB, and the number of samples is varied from 100 to 600. The simulation result is illustrated in Fig. 10, and again it shows that the proposed D1D-CNN based method can achieve lower RMSE than the other two, and exhibits great robustness against array perturbations.

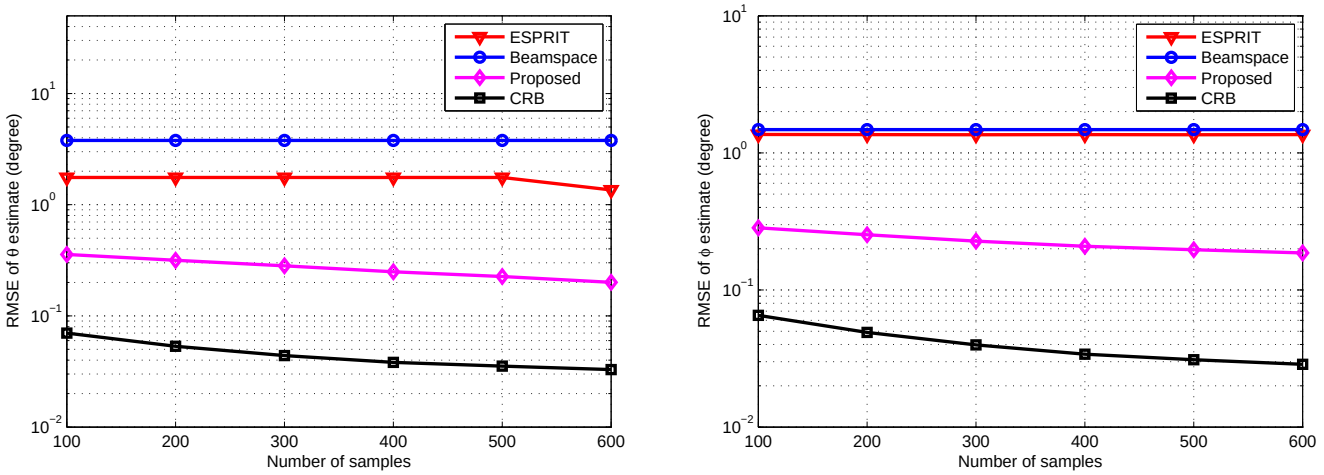
C. Performance of Vehicle Positioning Utilizing Three Collaborative BSs without Array Perturbation

In this part, the vehicle positioning performance is studied, where three collaborative BSs with their location coordinates $\#1(200\text{m}, 0\text{m}, 50\text{m})$, $\#2(0\text{m}, -20\text{m}, 50\text{m})$, $\#3(100\text{m}, 100\text{m}, 50\text{m})$ are considered. The coordinate of the vehicle is $(-30\text{m}, -40\text{m}, 0\text{m})$, and its position is assumed to be fixed within a time interval of 600 samples. Notice that the



(a) (b)

Fig. 9. RMSEs of 2-D nominal DOA estimation versus the average received SNR with array mutual coupling, $M = 100, N = 500$. (a) Estimation of θ. (b) Estimation of ϕ.



(a) (b)

Fig. 10. RMSEs of 2-D nominal DOA estimation versus the number of samples with array mutual coupling, $M = 100$, average received SNR=10 dB. (a) Estimation of θ. (b) Estimation of ϕ.

delay caused by 600 samples is about $600/3.84M \approx 1.56 \times 10^{-4}$s, and the speed of vehicle is usually limited to less than 100 m/s, and therefore, the change of nominal DOA caused by 600 samples is $100 \times 1.56 \times 10^{-4}/10^3/\pi \times 180 < 0.001°$, which means that the assumption about a fixed vehicle position is reasonable. Two different scenarios are considered. *Scenario 1*: The received SNR at each BS is the same and equals 10 dB, which is consistent with the assumptions in [9]-[11]. *Scenario 2*: The received SNRs at different BSs are different. For simplicity, the received SNR obtained by the BS closest to the vehicle is assumed to be 10 dB, and the received SNR obtained by the lth BS ($l = 1, 2, 3$) is $[10-2D_d]$ dB, where D_d (unit: km) denotes the distance difference between the vehicle to the nearest BS and the lth BS. Note that it is difficult to guarantee the attenuation of actual wireless signals arriving at different BSs to be the same, and hence *Scenario 2* is more representative of the actual situation. The positioning results are shown in Tables I and II, respectively, and it can be seen that the proposed method performs much better than the compared algorithms, and its corresponding positioning accuracy can reach sub-meter level, which provides reliable information for assisted driving. Particularly, when the received SNRs of different BSs are different, the proposed one can still maintain a satisfactory performance, whereas the performance of the compared algorithms gets much worse.

D. Performance of Vehicle Positioning Utilizing Three BSs with Mutual Coupling

In the last simulation, vehicle positioning performance of different algorithms in the presence of array mutual coupling is further examined. The array mutual coupling at all BSs is set to $c_1 = 0.7e^{-j\pi/4}$, and $c_2 = 0.3e^{j\pi/10}$. The other simulation conditions are the same as in the previous simulation. It can be seen from Tables III and IV that the proposed method

TABLE I
VEHICLE POSITIONING RESULTS FOR THE SAME RECEIVED SNR WITHOUT ARRAY PERTURBATIONS: SCENARIO 1

Number of Samples	Absolute error (m)		
	ESPRIT [24]	Beamspace [25]	Proposed
200	1.09	3.43	0.79
300	0.99	3.14	0.70
400	0.93	2.86	0.63
500	0.86	2.67	0.59
600	0.82	2.54	0.55

TABLE II
VEHICLE POSITIONING RESULTS FOR DIFFERENT RECEIVED SNRS WITHOUT ARRAY PERTURBATIONS: SCENARIO 2

Number of Samples	Absolute error (m)		
	ESPRIT [24]	Beamspace [25]	Proposed
200	1.91	5.46	0.94
300	1.75	5.03	0.83
400	1.63	4.60	0.74
500	1.50	4.27	0.69
600	1.40	3.99	0.61

TABLE III
VEHICLE POSITIONING RESULTS FOR THE SAME RECEIVED SNR WITH MUTUAL COUPLING: SCENARIO 1

Number of Samples	Absolute error (m)		
	ESPRIT [24]	Beamspace [25]	Proposed
200	4.44	7.52	0.81
300	4.44	7.51	0.73
400	4.41	7.50	0.66
500	4.41	7.50	0.62
600	3.98	7.49	0.57

TABLE IV
VEHICLE POSITIONING RESULTS FOR DIFFERENT RECEIVED SNRS WITH MUTUAL COUPLING: SCENARIO 2

Number of Samples	Absolute error (m)		
	ESPRIT [24]	Beamspace [25]	Proposed
200	4.53	7.62	0.97
300	4.52	7.59	0.86
400	4.49	7.57	0.77
500	4.49	7.56	0.72
600	4.10	7.52	0.65

can provide a satisfactory positioning result, which can better meet the requirements of vehicles in IoV environment with a positioning accuracy reaching the sub-meter level, whereas the error of the compared algorithms is larger than four meters.

VI. CONCLUSION

In this paper, a novel vehicle positioning method based on efficient 2-D DOA estimation of ID sources employing massive MIMO arrays has been proposed, which is more suitable for the actual IoV or vehicle application environments. A D1D-CNN was constructed first under the DL framework for 2-D DOA estimation, and it was enhanced further by adding transfer learning and an attention mechanism. As a result, it can avoid the process of parameters matching and simultaneously provide an improved estimation result, especially in the presence of array perturbations. Based on the collaborative BSs and the proposed DOA estimation method, a way to select the set of DOA estimation results for effective vehicle positioning using information of the received SNR was then developed. As demonstrated by extensive computer simulations, the proposed vehicle positioning method can achieve increased accuracy and robustness against array perturbations in comparison with two other state-of-the-art algorithms.

REFERENCES

[1] A. Sarker, H. Shen, M. Rahman, *et al*., "A review of sensing and communication, human factors, and controller aspects for information-aware connected and automated vehicles," *IEEE Trans. Intell. Transp. Syst*., vol. 21, no. 1, pp. 7-29, Jan. 2020.

[2] T. Kim, J. Lee, and T. Park, "Fusing Lidar, radar, and camera using extended kalman filter for estimating the forward position of vehicles," in *Proc. 2019 IEEE Conf. Cybernetics and Intell. Syst. (CIS) and 2019 IEEE Conf. Robotics, Antomation and Mech. (RAM)*, Bangkok, Thailand, Nov. 2019, pp. 374-379.

[3] S. Tomic, M. Beko, and R. Dinis, "RSS-based localization in wireless sensor networks using convex relaxation: noncooperative and cooperative schemes," *IEEE Trans. Veh. Technol*., vol. 64, no. 5, pp. 2037-2050, May 2015.

[4] Y. Hu and G. Leus, "Robust differential received signal strength-based localization," *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3261-3276, Jun. 2017.

[5] G. Wang, S. Cai, Y. Li, and M. Jin, "Second-order cone relaxation for TOA-based source localization with unknown start transmission time," *IEEE Trans. Veh. Technol*., vol. 63, no. 6, pp. 2973-2977, Jul. 2014.

[6] Y. Zou and H. Liu, "TDOA localization with unknown signal propagation speed and sensor position errors," *IEEE Commun. Lett*., vol. 24, no. 5, pp. 1024-1027, May 2020.

[7] G. Han, L. Wan, L. Shu, and N. Feng, "Two novel DOA estimation approaches for real-time assistant calibration systems in future vehicle industrial," *IEEE Syst. J*., vol. 11, no. 3, pp. 1361-1372, Sept. 2017.

[8] Z. Zheng and S. Mu, "Two-dimensional DOA estimation using two parallel nested arrays," *IEEE Commun. Lett.*, vol. 24, no. 3, pp. 568-571, Mar. 2020.

[9] H. Wang, L. Wan, M. Dong, *et al*., "Assistant vehicle localization based on three collaborative base stations via SBL-based robust DOA estimation," *IEEE Internet of Things J.*, vol. 6, no. 3, pp. 5766-5777, Jun. 2019.

[10] F. Wen, J. Wang, J. Shi and G. Gui, "Auxiliary vehicle positioning based on robust DOA estimation with unknown mutual coupling," *IEEE Internet of Things J.*, vol. 7, no. 6, pp. 5521-5532, Jun. 2020.

[11] L. Wan, Y. Sun, L. Sun, *et al*., "Deep learning based autonomous vehicle super resolution DOA estimation for safety driving", *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4301-4315, Jul. 2021.

[12] M. Ghogho, O. Besson and A. Swami, "Estimation of directions of arrival of multiple scattered sources," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2467-2480, Nov. 2001.

[13] L. Wan, G. Han, J. Jiang, *et al*., "DOA estimation for coherently distributed sources considering circular and noncircular signals in massive MIMO systems," *IEEE Syst. J.*, vol. 11, no. 1, pp. 41-49, Mar. 2017.

[14] S. Shahbazpanahi, S. Valaee and M. H. Bastani, "Distributed source localization using ESPRIT algorithm," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2169-2178, Oct. 2001.

[15] Y. Zhou, Z. Fei, S. Yang, *et al*., "Joint angle estimation and signal reconstruction for coherently distributed sources in massive MIMO systems based on 2-D Unitary ESPRIT," *IEEE Access*, vol. 5, pp. 9632-9646, Jun. 2017.

[16] Q. Cheng, X. Zhang and R. Cao, "Fast parallel factor decomposition technique for coherently distributed source localization," *J. Syst. Eng. Electron*., vol. 29, no. 4, pp. 667-675, Aug. 2018.

[17] Y. Tian, H. Yue and X. Rong, "Multi-parameters estimation of coherently distributed sources under coexistence of circular and noncircular signals," *IEEE Commun. Lett.*, vol. 24, no. 6, pp. 1254-1257, Jun. 2020.

[18] H. Boujeâa, "Extension of COMET algorithm to multiple diffuse source localization in azimuth and elevation," *Eur. Trans. Telecommun.*, vol. 16, no. 6, pp. 557-566, Nov./Dec. 2005.

[19] B. T. Sieskul, "An asymptotic maximum likelihood for joint estimation of nominal angles and angular spreads of multiple spatially distributed sources," *IEEE Trans. Veh. Technol*., vol. 59, no. 3, pp. 1534-1538, Mar. 2010.
[20] J. Lee, J. Joung, and J. D. Kim, "A method for the direction-of-arrival estimation of incoherently distributed sources," *IEEE Trans. Veh. Technol*., vol. 57, no. 5, pp. 2885-2893, Sep. 2008.
[21] Y. Meng, P. Stoica and K. Wong, "Estimation of the directions of arrival of spatially dispersed signals in array processing," *Inst. Electron. Eng. Proc. Radar, Sonar Navigat*., vol. 143, no. 1, pp. 1-9, Feb. 1996.
[22] L. Wan, K. Liu, Y. -C. Liang and T. Zhu, "DOA and polarization estimation for non-circular signals in 3-D millimeter wave polarized massive MIMO systems," *IEEE Trans. Wireless Commun*., vol. 20, no. 5, pp. 3152-3167, May 2021.
[23] T. Ahmed, X. Zhang and W. U. Hassan, "A higher-order propagator method for 2D-DOA estimation in massive MIMO systems," *IEEE Commun. Lett*., vol. 24, no. 3, pp. 543-547, Mar. 2020.
[24] A. Hu, T. Lv, H. Cao, *et al.*, "An ESPRIT-based approach for 2-D localization of incoherenltly distributed sources in massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 996-1011, Oct. 2014.
[25] Z. Zheng, W. Wang, H. Meng, *et al.*, "Efficient beanspace-based algorithm for two-dimensional DOA estimation of incoherently distributed sources in massive MIMO systems," *IEEE Trans. Veh. Technol*., vol. 67, no. 12, pp. 11776-11789, Dec. 2018.
[26] L. Cheng, Y.-C. Wu, J. Zhang, and L. Liu, "Subspace identification for DOA estimation in massive/full-dimension MIMO systems: Bad data mitigation and automatic source enumeration," *IEEE Trans. Signal Process*., vol. 63, no. 22, pp. 5897-5909, Nov. 2015.
[27] R. Cao, F. Gao, and X. Zhang, "An angular parameter estimation method for incoherently distributed sources via generalized shift invariance," *IEEE Trans. Signal Process*., vol. 64, no. 17, pp. 4493-4503, Sep. 2016.
[28] H. Huang, J. Yang, H. Huang, *et al*., "Deep learning for super-Resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Trans. Veh. Technol*., vol. 67, no. 9, pp. 8549-8560, Sep. 2018.
[29] D. Hu, Y. Zhang, L. He and J. Wu, "Low-complexity deep-learning-based DOA estimation for hybrid massive MIMO systems with uniform circular arrays," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 83-86, Jan. 2020.
[30] A. L. Mass, A. Y. Hannun and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. 30th Int. Conf. Machine Learning*, Atlanta, Georgia, USA, 2013, pp. 1-6.
[31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980*, 2014.
[32] Y. Tian and H. Xu, "Calibration nested arrays for underdetermined DOA estimation using fourth-order cumulant," *IEEE Commun. Lett.*, vol. 24, no. 9, pp. 1949-1952, Sep. 2020.
[33] L. Qin, C. Li, Y. Du and B. Li, "DOA estimation and mutual coupling calibration algorithm for array in plasma environment," *IEEE Trans. Plasma Sci*., vol. 48, no. 6, pp. 2075-2082, Jun. 2020.
[34] S. Liu, Z. Zhang and Y. Guo, "2-D DOA estimation with imperfect L-shaped array using active calibration," *IEEE Commun. Lett.*, vol. 25, no. 4, pp. 1178-1182, Apr. 2021.
[35] Y. Tian and X. Rong, "Generalized cablibration of model errors for uniform linear arrays," *Meas. Sci. Technol*., vol. 31, pp. 115101, 2020.
[36] Y. Bresler, "Maximum likelihood estimation of a linearly structured covariance with application to antenna array processing, in *Proc. 4th Annual ASSP Workshop Spectrum Estimation Model.*, Minneapolis, MN, USA, Aug. 1988, pp. 172-175.

Ye Tian (Member, IEEE) received the B.S. and Ph.D. degrees from the College of Communication Engineering, Jilin University, Changchun, China, in 2009 and 2014, respectively. He won a Huawei scholarship in 2013 and was selected as a young top talent by the Hebei Provincial Department of Education in 2016. He is currently an Associate Professor in Faculty of Information Science and Engineering, Ningbo University. He has published more than 30 international peer-reviewed journal/conference papers and more than 10 patents. His research interests include array signal processing, autonomous vehicle positioning, massive MIMO as well as large-dimensional random matrix theory.

Shuai Liu (Member, IEEE) received the M.Eng. degree from the School of Resources and Information, China University of Petroleum, Beijing, China, in 2008, and Ph.D. degree from the Institute of Remote Sensing and Geographic Information System, Peking University, China, in 2012, respectively. He is currently a Lecturer with School of Information Science and Engineering, Yanshan University. His research interests include Intelligent remote sensing information processing, array signal Processing and machine learning.

Wei Liu (Senior Member, IEEE) received his BSc and LLB. degrees from Peking University, China, in 1996 and 1997, respectively, MPhil from the University of Hong Kong in 2001, and PhD from the School of Electronics and Computer Science, University of Southampton, UK, in 2003. He then worked as a postdoc first at Southampton and later at the Department of Electrical and Electronic Engineering, Imperial College London. Since September 2005, he has been with the Department of Electronic and Electrical Engineering, University of Sheffield, UK, first as a Lecturer and then a Senior Lecturer. He has published about 350 journal and conference papers, five book chapters, and two research monographs titled "Wideband Beamforming: Concepts and Techniques" (John Wiley, March 2010) and "Low-Cost Smart Antennas" (by Wiley-IEEE, March 2019), respectively. His research interests cover a wide range of topics in signal processing, with a focus on sensor array signal processing and its various applications, such as robotics and autonomous systems, human computer interface, radar, sonar, satellite navigation, and wireless communications.

He is a member of the Digital Signal Processing Technical Committee of the IEEE Circuits and Systems Society (Secretary from May 2020) and the Sensor Array and Multichannel Signal Processing Technical Committee of the IEEE Signal Processing Society (Chair from Jan 2021). He was an Associate Editor for IEEE Trans. on Signal Processing (2015-2019) and IEEE Access (2016-2021), and is currently an editorial board member of the journal Frontiers of Information Technology and Electronic Engineering and the Journal of The Franklin Institute.

Hua Chen (Member, IEEE) received the M.Eng. and Ph.D. degrees in Information and Communication Engineering from Tianjin University, Tianjin, China, in 2013 and 2017, respectively. He is currently an Associate Professor in Faculty of Information Science and Engineering, Ningbo University, China. His research interests include array signal processing and MIMO radar.

Dr. Chen is currently an Associated Editor for Circuits, Systems, and Signal Processing.

Zhiyan Dong received the B.Eng. degree in Communication Engineering and Ph.D. degree in Mechanical and Aerospace Engineering from Jilin University, Changchun, China, in 2011 and 2016, respectively. He is currently an Associate Professor in Faculty of Institute of AI and Robotics, Fudan University. His research interests include wireless communications, intelligent control and perception, and swarm for unmanned vehicle.