Convolutional Neural Network-Based DOA Estimation Using Non-uniform Linear Array for Multipath Channels

Van-Sang Doan[1], Thien Huynh-The[1], Van-Phuc Hoang[2],
and Dong-Seong Kim[1](✉)

[1] Kumoh National Institute of Technology, Gumi, South Korea
{vansang.doan,thienht,dskim}@kumoh.ac.kr
[2] Le Quy Don Technical University, Hanoi, Vietnam
phuchv@lqdtu.edu.vn

Abstract. In this paper, a novel convolutional neural network (CNN) was designed for DOA estimation, which could deploy in radio-electronics systems for improving the accuracy and operation efficiency. The proposed model was evaluated with different hyper-parameter configurations for optimization, and then a suitable model was compared with other existing models to demonstrate its preeminence. Regarding dataset generation, our work considered the influence of both Gaussian noise and multipath channels to DOA estimation accuracy. According to the analysis, in frame of this study, the model with 5 conv-blocks, 48 filters, and a filter size of 1×7 achieved the best performance in terms of accuracy (75.27% at +5 dB SNR) and prediction time (10.1 ms) that notably outperformed two other state-of-the-art CNN model-based DOA estimation techniques.

Keywords: Convolution neural network · DOA estimation · Multipath channels · Antenna array

1 Introduction

Many years, direction of arrival (DOA) estimation has been an active research topic for application in various areas, including radar, sonar, and communication. In military radar and sonar systems, DOA estimation facilitates determining the location of targets to help for surveillance, tracking and control in the ground, air, and water areas of nations. In communications, the quality of wireless connections is greatly improved if DOA estimation is employed to target the user. In many contexts, the DOA estimation plays a role as a spatial filter. In addition, the arrival angle information of a signal source in the electronic warfare

---

Supported by ICTCRC, Kumoh National Institute of Technology, South Korea.

N.-S. Vo and V.-P. Hoang (Eds.): INISCOM 2020, LNICST 334, pp. 45–56, 2020.
https://doi.org/10.1007/978-3-030-63083-6_4

operation is very useful to explore further data about the enemy situation that enable us to build a plan or decide a proper activity.

In almost applications, an antenna array with multiple elements is often used for steering a beam pattern toward the defined direction. As well-known methods, multiple signal classification (MUSIC) [1] and estimation of signal parameters via rotation invariance (ESPRIT) exploit the Eigen-decomposition of the covariance matrix of received signals to determine the arrival angles [2]. In decades, researchers attempt to solve the DOA estimation for coherent signals. Therefore, plenty of array pre-processing techniques are proposed, such as forward spatial smoothing (FSS) [3], forward/backward spatial smoothing (FBSS) [4], joint spatial-temporal method [5], and Toeplitz approximation [6]. They all were deeply analyzed and evaluated.

Recently, the machine learning technique is introduced as an effective approach to solving the DOA estimation problem, for example, support vector machine (SVM) is used to estimate DOA of multiple plane waves in [7–9]. As part of machine learning, deep learning is a well-known universal approximation theorem that can learn features deeply by designing neural networks with multiple hidden layers [10]. Accordingly, a deep neural network (DNN) proposed in [11] was designed for multitask autoencoder and a sequence of parallel multiple-layer classifiers. As a result, the DNN achieved a DOA estimation performance with higher accuracy than SVM and MUSIC techniques. Despite enhancing DOA estimation accuracy, the multilayer perception consumes a high computational complexity, expensive architecture, and time delay. Besides, convolutional neural network (CNN) allow learning feature automatically without expert knowledge in a specific domain. In 2017, *Adavanne et al.* demon-strated a stacked convolutional recurrent neural network, namely DOAnet, for esti-mating DOA in both azimuth and elevation angles [12]. In that work, the input data of DOAnet was signal frames in time domain. The study evaluated the DOA estimation possibility of DOAnet on anechoic, matched and unmatched reverberation dataset, and indicated that the approach performed better than MUSIC in most scenarios. In 2019, *Chakrabarty et al.* proposed a CNN-based method for DOA estima-tion of multiple speakers, here, the short-time Fourier transform (STFT) of received signals was used as a pre-processing to generate the input feature map, which is di-rectly fed into the CNN [13]. In another approach, the input feature map was columns of a covariance matrix, which played a role as a spatial spectrum [14]. However, the method did not consider multipath propagation conditions.

Despite achieving the remarkable performance of DOA estimation based on deep learning techniques, the above-mentioned researches have concerned the assessment of uniform linear array (ULA) but not non-uniform linear array (NLA); therefore, it still remains room for exploration. In addition, no signal dataset with the multipath effect is synthesized in those publications for learning CNN models, and this, therefore, motivates us to generate a new dataset for performance evaluation, in which diverse factors affecting the incoming signals are considered exhaustively. Accordingly, this paper presented a CNN model, namely DOA-ConvNet, which was designed according to a combination of dense

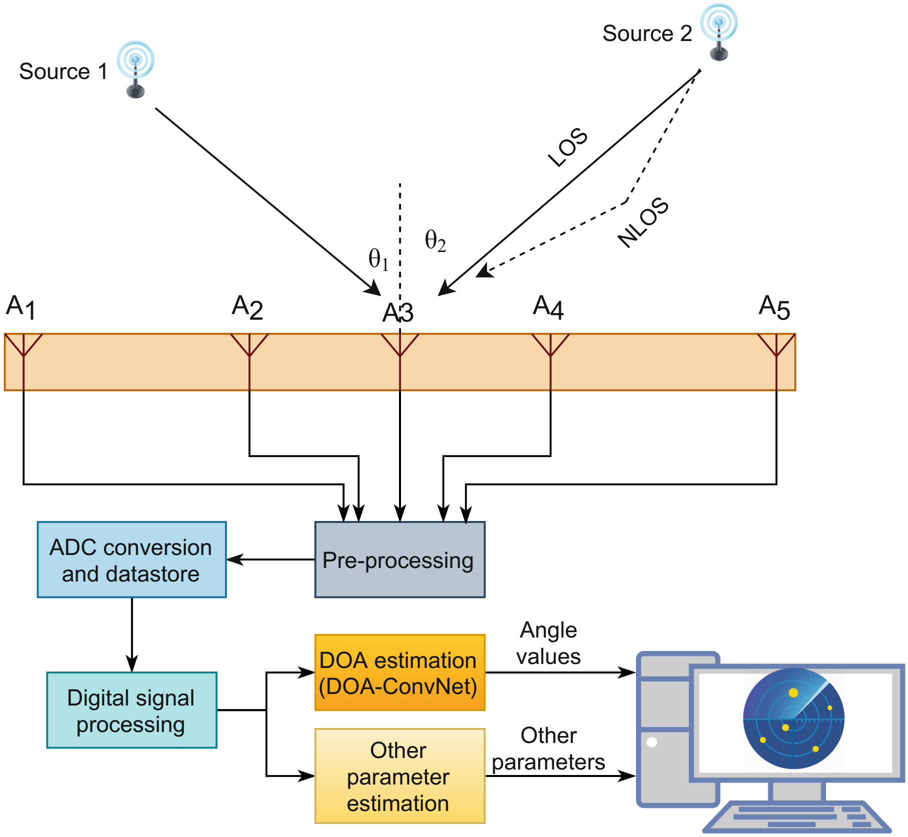


Fig. 1. Direction finding system with 5-element NLA for DOA estimation using DOA-ConvNet model.

and residual network architecture. The proposed model was evaluated to achieve the optimal performance in terms of accuracy and time-consuming.

2 Signal Model and Dataset Generation

2.1 Signal Model of Antenna Array

As shown in Fig. 1, a signal model of a linear antenna array is illustrated. In principle, an incoming signal is captured by the antenna array. Then, the received signals of elements are pre-processed at pre-processing block and converted to digital form at ADC conversion and datastore block. In the digital signal processing block, the signals are re-constructed and prepared for DOA-ConvNet. The output of DOA-ConvNet is a DOA value of the received signal. Along with other signal parameters, the estimated DOA is indicated in a display for surveillance and tracking.

Accordingly, assume that a signal travels from one transmitter to a non-uniform linear array (NLA) with M elements in K paths of the multipath propagation. The light-of-sight (LOS) path is in the directions of θ_1, and other non-light-of-sight (NLOS) ones are in the direction of θ_2, θ_3, ..., θ_K. Then, the output signals of NLA expressed as a matrix $\mathbf{x}$ in time domain are written as:

$$\mathbf{x} = \mathbf{A} \cdot \mathbf{s} + \mathbf{n} \tag{1}$$

where $\mathbf{x} = [x_1(t), x_2(t), ..., x_M(t)]^T$ is output signal vector of the array, $x_m(t)$ is signal value at time t of m^{th} antenna element; $\mathbf{s} = [s_1(t), s_2(t), ..., s_K(t)]^T$ presents incoming signal vector, $s_k(t)$ is signal value at time t of k^{th} signal from direction of θ_k; $\mathbf{n} = [n_1(t), n_2(t), ..., n_M(t)]^T$ stands for white noise vector, whose elements are independent; $\mathbf{A} = [\mathbf{a}_1(\theta_1), \mathbf{a}_2(\theta_2), ..., \mathbf{a}_K(\theta_K)]$ is defined as a steering matrix of the array, $\mathbf{a}_k(\theta_k)$ is a steering vector for angle θ_k as following:

$$\mathbf{a}_k(\theta_k) = \left[1, e^{-j\frac{2\pi d_1 \sin\theta_k}{\lambda_k}}, ..., e^{-j\frac{2\pi d_{M-1} \sin\theta_k}{\lambda_k}}\right]^T \tag{2}$$

The multipath propagation is a phenomenon that strongly affects the DOA estimation performance of a direction finding system. The angle bias occurs by the non-useful correlation of non-line-of-sight (NLOS) signals. The bias problem sometimes is solved by calibration for stationary propagation environment, but not for mobile devices that are deployed on the battlefield. As a result, the antenna array output will contain LOS signal, their NLOS components, and white noise. By acquisition of the array output signals with multiple samples, the direction-finding system can estimate the arrival angle of strong incoming signals. After converting output signals to digital, the matrix $\mathbf{x}$ has a size of $M \times N$, with M is number of antenna elements, and N is number of signal samples. In the case of I/Q (In-phase/Quadratic-phase) receiver, the matrix $\mathbf{x}$ is assigned with the size of $M \times N \times 2$. The matrix $\mathbf{x}$ now becomes input data for the proposed DOA-ConvNet.

2.2 Dataset Generation

In this study, an NLA with configuration of $\mathbf{d} = \{0, d_1, d_2, d_3, d_4\} = \{0, 3, 5, 7, 10\}\lambda/2$ is considered for evaluating the performance of DOA-ConvNet for DOA estimation task. The NLA configuration has a symmetric property that provides a high accuracy of DOA estimation of multiple coherent signals, as demonstrated in [18]. Regard to the data structure, the signal frame for each acquisition is of size 256 samples. By using the array signal model as presented in the previous subsection, a dataset with consideration of multiple noise levels and randomized multipath propagation is generated for training and validating the proposed network. Accordingly, the dataset is produced according to parameters summarized in Table 1. Herein, the input feature map is defined by the size of matrix $\mathbf{x}$, concretely is of size $5 \times 256 \times 2$. All signal feature maps in the

Table 1. Summary of signal dataset generation parameters.

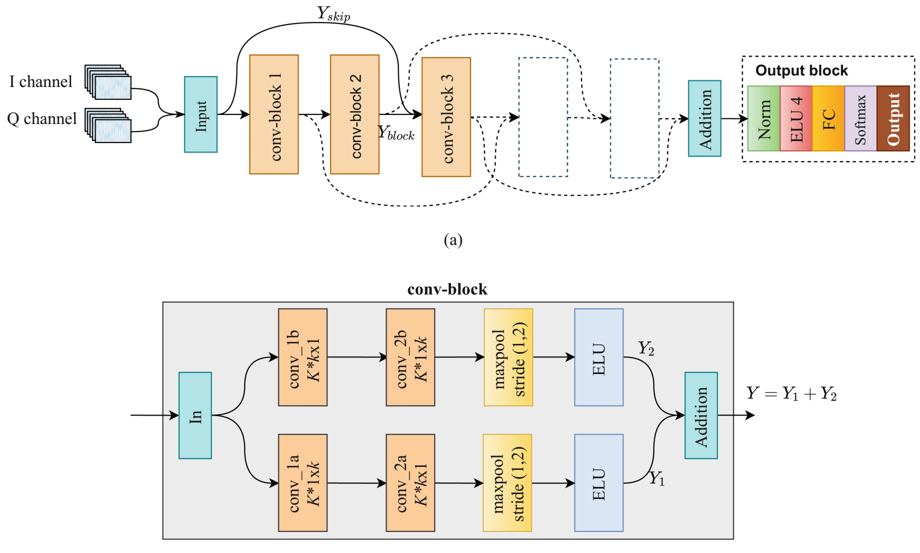
Parameter	Value
Array	NLA, $d = [0, 3, 5, 7, 10]\lambda/2$
Carrier Frequency (LOS)	35 MHz
Sampling frequency	350 MHz
Signal strength	1 V
Number of NLOS signals	6
Multipath delay (NLOS)	$\mathbf{U}\{[1, 1000]\}$ ns
Maximum Doppler shift	10 Hz
Multipath attenuation (NLOS)	$\mathbf{U}\{[-100, -10]\}$ dB
SNR	-20 dB to $+25$ dB, step 5 dB
Angle of arrival	$-89°$ to $+89°$, step $1°$
Number of angle classes	179 class names
Data size of one signal window	$5 \times 256 \times 2$

dataset are labeled corresponding to angles from $-89°$ to $+89°$ with the step of $1°$, so there is a total of 179 label names that correspond to 179 output classes of the proposed network. In detail, $10,000$ signal feature maps are generated for each angle with various SNRs from -15 dB to $+15$ dB with the step of 5 dB. Each SNR has $179,000$ signal feature maps. Overall, the dataset has $1,790,000$ observation maps, in which 80% is used for training and 20% is for validation.

3 CNN-Based DOA Estimation Model

An ELINT receiver always operates with two channels I and Q, which provide full amplitude and phase information of received signals. Therefore, the input data size should be of $M \times N \times 2$, where M is number of antenna elements, N is number of samples per signal window, and number 2 presents two signal channels, I and Q. According to the problem in this work, the input data size is assigned by $5 \times 256 \times 2$. Inspired by the ResNet architecture [15], our deep neural network model (namely DOA-ConvNet) is designed with a primary flow (main path) and skip-connections (residual paths), as shown in Fig. 2a and listed in Table 2. Accordingly, the primary flow is constructed by a series of conv-blocks connected consecutively from input to output layers. Each conv-block, as shown in Fig. 2b, consists of two branches, whose outputs are synthesized by an addition layer. Each branch contains two consecutive one-dimensional (1D) convolution layers, one max-pool layer, and one ELU activation function layer. In 1D-convolution layer, the convolution operation is executed between an input data and a kernel so that its formula can express as follows:

$$\mathbf{y}(i) = \sum_{m} \mathbf{c}(m) \cdot \mathbf{x}(i - m) \tag{3}$$



(a)

(b)

Fig. 2. DOA-ConvNet model structure: (a) *overall*; (b) *conv-block.*

where $\mathbf{y}$ is an output matrix, $\mathbf{x}$ is an input matrix, $\mathbf{c}$ is a kernel matrix, and m, i, are element indices.

It can be observed that two 1D-convolution layers with filter sizes of $1 \times k$ and $k \times 1$ are employed instead of one convolution layer with a filter size of $k \times k$. This way helps to reduce the number of learnable parameters that produce the model with a light-weight structure [16,17]. In the second branch, two convolution layers with two mentioned kernel sizes are exchanged their positions. The second convolution layer is followed by a maxpool layer whose hyper-parameters are designed by a pool-size of 1×3 and stride of $(1, 2)$. The role of maxpool layer is to perform down-sampling to reduce the size of feature maps and extract the robust features before going to the activation layer. The activation function plays a crucial role in CNNs as a brain neuron, which activates the useful features and impairs the redundant ones. In our model, we use an exponential linear unit (ELU) function, a smooth activation function, which provides better signal propagation through the network than ReLU one. The ELU function is expressed as follows:

$$y = \begin{cases} x & \text{if } x \geq 0 \\ \alpha(e^x - 1) & \text{if } x < 0 \end{cases} \tag{4}$$

where y denotes the output, x stands for input value and α is a scalar number between 0.1 and 0.3. In this work, it is chosen by 0.2.

4 Experiment Result and Discussion

Our model is intended with changeable hyper-parameters, including number of conv-blocks, number of filters in a convolution layer, and size of filters. Different configurations of the model will be evaluated, and then a suitable one will be chosen for competing with other existing models in terms of accuracy and time-consuming. For each configuration, the model must be trained with the synthesized dataset, which is described in Sect. 2 before the experiment and assessment.

Table 2. Summary of signal dataset generation parameters.

Block	Layer	Output	Description
Input	Input	$5 \times 256 \times 2$	256 Sample/frame
			5 antenna element
			2 signal channels I/Q
p^{th} conv-blocks ($p = 1, 2, \ldots, P$), P is number of conv-blocks	*Branch 1:*		
	- Conv	$5 \times 256/2^{p-1} \times K$	K filters, $1 \times k$, stride $(1, 1)$
	- Conv	$5 \times 256/2^{p-1} \times K$	K filters, $k \times 1$, stride $(1, 1)$
	- Maxpool	$5 \times 256/2^{p} \times K$	Poolsize 1×3, stride $(1, 2)$
	- Activation	$5 \times 256/2^{p} \times K$	ELU: exponential linear unit
	Branch 2:		
	- Conv	$5 \times 256/2^{p-1} \times K$	K filters, $k \times 1$, stride $(1, 1)$
	- Conv	$5 \times 256/2^{p-1} \times K$	K filters, $1 \times k$, stride $(1, 1)$
	- Maxpool	$5 \times 256/2^{p} \times K$	Poolsize 1×3, stride $(1, 2)$
	- Activation	$5 \times 256/2^{p} \times K$	ELU: exponential linear unit
	- *Addition*	$5 \times 256/2^{p} \times K$	Branch 1 + branch 2 + skip-connection (except $n = 1$)
p^{th} Skipconnection ($p = 1, 2, \ldots, P - 1$)	- Conv	$5 \times 256/2^{p+1} \times K$	K filters, 1×1, stride $(1, 1)$
	- Maxpool	$5 \times 256/2^{p-1} \times K$	Poolsize 1×6, stride $(1, 4)$
Output block	Norm	$5 \times 256/2^{P} \times K$	Batch normalization
	Activation	$5 \times 256/2^{P} \times K$	ELU
	FC	$1 \times 1 \times 179$	179 classes
	Softmax	$1 \times 1 \times 179$	Softmax function
	Classification		Prediction execution

4.1 DOA-ConvNet with Different Filter Size

As mentioned in the previous section, the pair of consecutive convolution layers in each conv-block branch of DOA-ConvNet are assigned with K filters, and filter sizes of $1 \times k$ and $k \times 1$, respectively. For evaluating the impact of filter size on the performance, the network is configured with following hyper-parameters: 5 conv-blocks, 64 filters per convolution layer, and various filter sizes with $k = [3, 5, 7, 9]$.

The experimental result presented in Fig. 3 shows the comparison of angle classification accuracy depending on various SNR levels and different convolution filter sizes. It shows that in spite of giving the best accuracy at low SNR levels (from -20 dB to 0 dB), the model with the filter size of 1×9 has the worst

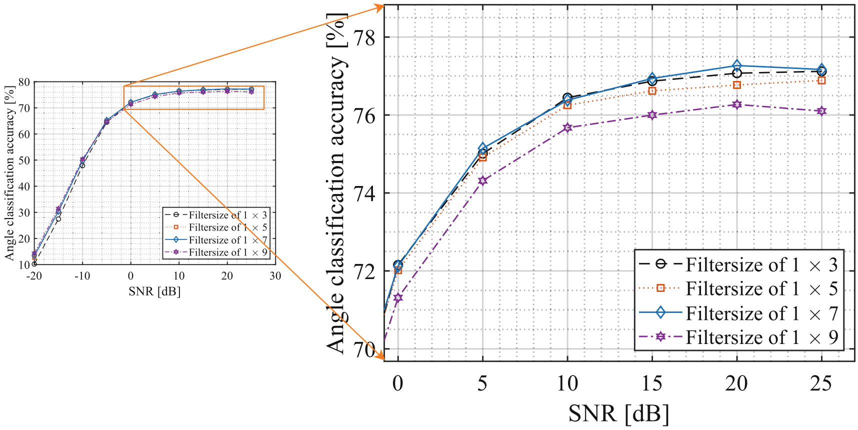


Fig. 3. Angle classification performance of DOA-ConvNet with different filter sizes depending on various SNRs.

performance at high SNRs ($> 0\,$dB). Whereas, the model with the filter size of 1×3 seem lowest correct angle classification rate for SNR < 0 dB but it outperforms other ones at higher SNRs excepting 1×7. By observation, the model with the filter size of 1×7 obtains very good performance at all SNR values, especially it achieves the best accuracy with SNR $> 0\,$dB.

4.2 DOA-ConvNet with Different Number of Filters

In Subsect. 4.1, the filter size with $k = 7$ is selected as a suitable configuration of the convolution dimension; thus, the second experiment is performed with different numbers of filters in a convolution layer while 5 conv-blocks are still used in this evaluation.

As a result, Fig. 4 shows the different performances of the model due to various filter numbers. Logically, more filters result in more neurons in the model; thus, more representative information should be extracted for better prediction performance. However, we observe that the model with a greater number of filters does not gain higher angle classification accuracy that the number of filters should be satisfied with dataset properties, that only be discovered by experiments. In particular, in this framework the model with 48 filters estimates DOA with the highest accuracy, whereas the models with 64 and 96 filters obtain lower performance.

4.3 DOA-ConvNet with Different Conv-Blocks

"*How many convolution layers are optimal for a neural network model-based DOA estimation*" is always a difficult question for answering. Indeed, how big is the network, it depends on the dataset properties and the number of output classes. Therefore, a classifier must be fitted with a particular dataset. For the

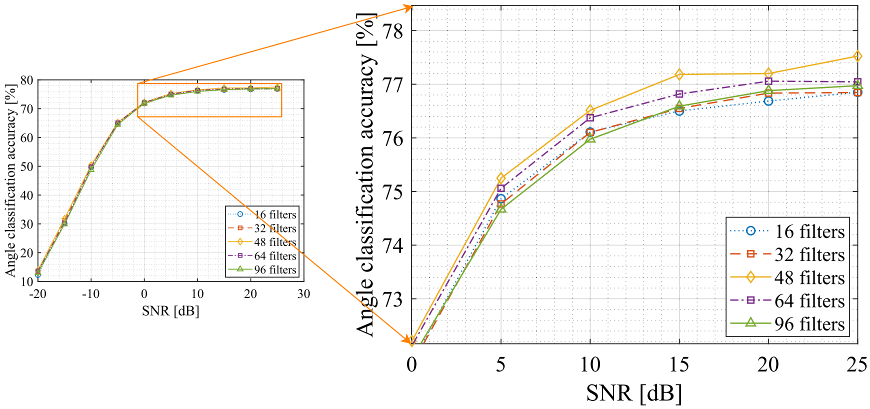


Fig. 4. Angle classification performance of DOA-ConvNet with different numbers of filters depending on various SNRs.

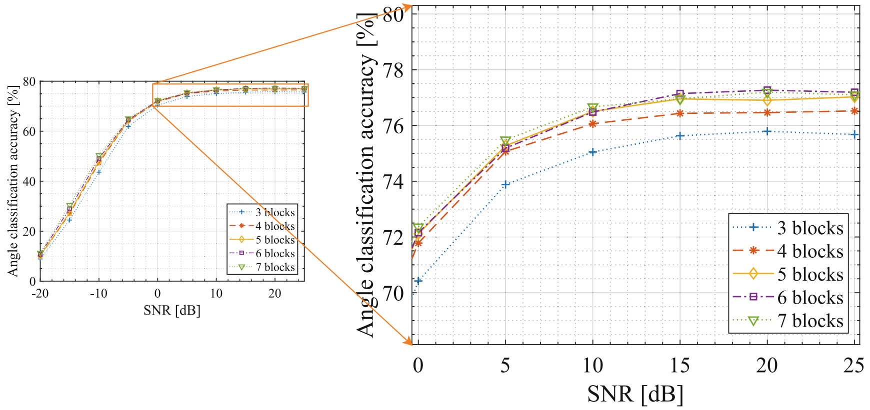


Fig. 5. Angle classification performance of DOA-ConvNet with different numbers of conv-blocks depending on various SNRs.

DOA estimation problem, DOA-ConvNet al.so has to be optimized with a suitable number of convolution layers, which must obtain good accuracy and fast execution. Herein, the number of convolution layers is represented by the number of conv-blocks.

By experiment, the accuracy performance depending on SNRs and different numbers of conv-blocks is plotted in Fig. 5, which shows that DOA-ConvNet with 5, 6, and 7 conv-blocks achieves the almost same classification efficiency and significantly better than that with 3 and 4 conv-blocks. Moreover, the execution time performance also is taken into account for assessment. This result is listed in Table 3, where the numerical results enable us to suggest that the model with 5 conv-blocks is chosen to satisfy the accuracy, structural, and time cost requirements.

Table 3. Performance of DOA-CovNet with different numbers of conv-blocks at SNR = +5 dB.

Number of Conv-blocks	Learnable parameters	Classification accuracy	Time-consuming
3	1.9M	73.88%	7.3 ms
4	1.1M	75.07%	8.8 ms
5	0.7M	75.27%	10.1 ms
6	0.5M	75.19%	11.8 ms
7	0.4M	75.48%	13.6 ms

4.4 Comparison of DOA-ConvNet with Other Existing CNN Models

As a result of the aforementioned analysis, DOA-ConvNet with 5 conv-blocks,48 filters, and a filter size of $k = 7$ is our designated CNN model to compete with two other existing ones (namely DOA-CNN-1 and DOA-CNN-2) for DOA estimation task. DOA-CNN-1 consists of 4 consecutive pairs of Conv+ReLU layers, following by three fully connected layers with output sizes of 512 for two first ones and 179 classes for the last one, and finalized by softmax layer and output decision layer [13]. DOA-CNN-2 is constructed by four consecutive pairs of Conv+ReLU layers also, but its input data is a spatial spectrum, which is pre-processed by beamforming technique [14]. Both mentioned CNN models were

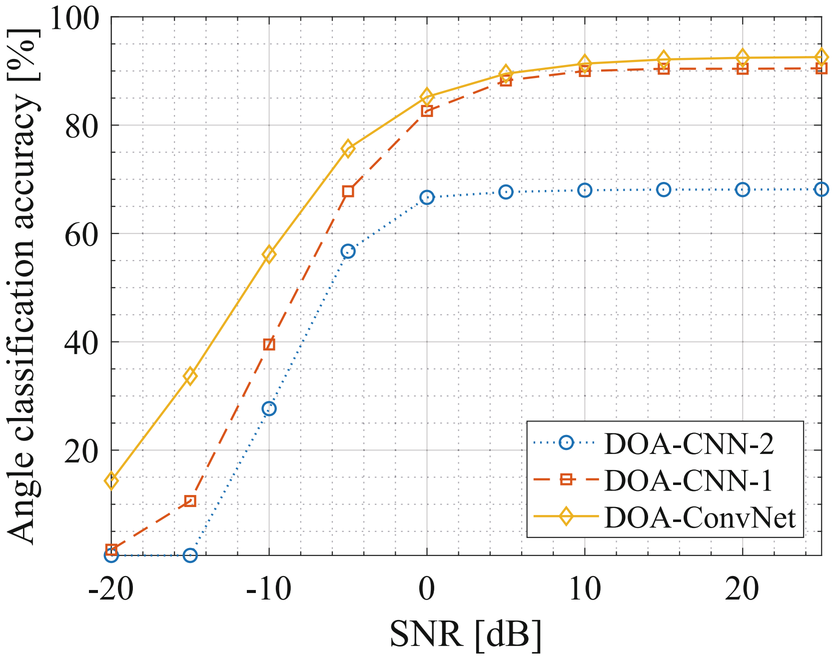


Fig. 6. Comparison of DOA-ConvNet with DOA-CNN-1 and DOA-CNN-2.

demonstrated remarkable outcomes in their own case studies. In this comparison, we applied those two CNN models for our signal dataset design, based on 5-element NLA along with Gaussian noise and multipath channels, to compare with DOA-ConvNet, while other setting up parameters remain as original. The comparison result is plotted in Fig. 6.

Regarding the structural volume, DOA-CNN-1, and DOA-CNN-2 have respectively $8,662,451$ and $960,488$ learnable parameters, whereas DOA-ConvNet has the number of parameters of 646,619 only. However, in the accuracy comparison, DOA-ConvNet gains the highest performance at all SNR values, the second place for DOA-CNN-2, and the worst for DOA-CNN-1. As we mentioned above, DOA-ConvNet is designed for re-usage of former feature maps, which can improve the robust features, therefore, it achieves higher accuracy.

5 Conclusion

This paper has demonstrated a novel CNN model-based DOA estimation method (namely DOA-ConvNet). The proposed model was evaluated based on different structural hyper-parameters by experiments to gain an optimal one with the good trade-off in terms of accuracy, structural, and time cost. In addition, a signal dataset was synthesized with consideration of both Gaussian noise and multipath channels. A non-uniform linear array with the configuration of $\mathbf{d} = [0, 3, 5, 7, 10]\lambda/2$, along with the optimal DOA-ConvNet, is proposed to outperform other existing models for the same DOA estimation task. In the future work, we intend to develop the network for different array geometries such as circular, rectangle planar or spherical arrays, which can estimate signal DOA in both azimuth and elevation planes.

Acknowledgment. This work was supported under the framework of international cooperation program managed by National Research Foundation of Korea (NRF-2019K2A9A1A09081533) and Priority Research Centers Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (NRF-2018R1A6A1A03024003).

References

1. Schmidt, R.: Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propag. **34**, 276–280 (1986)
2. Roy, R., Kailath, T.: ESPRIT-estimation of signal parameters via rotational invariance techniques. IEEE Trans. Acoust. Speech Signal Process. **37**, 984–995 (1989)
3. Reddy, K.M., Reddy, V.U.: Analysis of spatial smoothing with uniform circular arrays. IEEE Trans. Signal Process. **47**, 1726–1730 (1999)
4. Pillai, S.U., Kwon, B.H.: Forward/backward spatial smoothing techniques for coherent signal identification. IEEE Trans. Acoust. Speech Signal Process. **37**, 8–15 (1989)
5. Lin, J.-D., Fang, W.-H., Wu, M.-L.: Joint spatial-temporal channel parameter estimation using tree-structured MUSIC. In: IEEE 55th Vehicular Technology Conference. VTC Spring 2002 (Cat. No.02CH37367) (2002)

6. Wang, Q., Chen, H., Zhao, G., Chen, B., Wang, P.: An improved direction finding algorithm based on Toeplitz approximation. Sensors **13**, 746–757 (2013)
7. Pastorino, M., Randazzo, A.: A smart antenna system for direction of arrival estimation based on a support vector regression. IEEE Trans. Antennas Propag. **53**, 2161–2168 (2005)
8. Gonnouni, A.E., Martinez-Ramon, M., Rojo-Alvarez, J.L., Camps-Valls, G., Figueiras-Vidal, A.R., Christodoulou, C.G.: A support vector machine MUSIC algorithm. IEEE Trans. Antennas Propag. **60**(10), 4901–4910 (2012)
9. Sun, F.-Y., Tian, Y.-B., Hu, G.-B., Shen, Q.-Y.: DOA estimation based on support vector machine ensemble. Int. J. Numer. Model. Electron. Netw. Dev. Fields, e2614 (2019)
10. Huang, H., Yang, J., Huang, H., Song, Y., Gui, G.: Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system. IEEE Trans. Veh. Technol. **67**(9), 8549–8560 (2018)
11. Liu, Z.-M., Zhang, C., Yu, P.S.: Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections. IEEE Trans. Antennas Propag. **66**(12), 7315–7327 (2018)
12. Adavanne, S., Politis, A., Virtanen, T.: Direction of arrival estimation for multiple sound sources using convolutional recurrent neural network, CoRR, vol. abs/1710.10059 (2017)
13. Chakrabarty, S., Habets, E.A.P.: Multi-speaker DOA estimation using deep convolutional networks trained with noise signals. IEEE J. Sel. Top. Signal Process. **13**(1), 8–21 (2019)
14. Wu, L., Liu, Z.-M., Huang, Z.-T.: Deep convolution network for direction of arrival estimation with sparse prior. IEEE Signal Process. Lett. **26**(11), 1688–1692 (2019)
15. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778 (2016)
16. Huynh-The, T., Hua, C.-H., Pham, Q.-V., Kim, D.-S.: MCNet: an efficient CNN architecture for robust automatic modulation classification. IEEE Commun. Lett. **24**(4), 811–815 (2020)
17. Huynh-The, T., Hua, C., Kim, J., Kim, S., Kim, D.: Exploiting a low-cost CNN with skip connection for robust automatic modulation classification. In: 2020 IEEE Wireless Communications and Networking Conference (WCNC), pp. 1–6, Seoul, Korea (South) (2020)
18. Doan, V.-S., Kim, D.-S.: DOA estimation of multiple non-coherent and coherent signals using element transposition of covariance matrix. ICT Express **6**(2), 67–75 (2020)