

Broadband signal source localization using convolutional neural networks combined with coherent signal subspace methods

Nguyen Thanh Thien¹, Huynh Huy Cuong², Ta Quang Hien³,
Nguyen Manh Hung^{1*}, Do Doanh Dien⁴

¹Institute of Defense Equipment, Academy of Military Science and Technology, 17 Hoang Sam, Nghia Do, Hanoi, Vietnam;

²Institute for Tropical Technology, Academy of Military Science and Technology, 57A Truong Quoc Dung, Phu Nhuan, Ho Chi Minh City, Vietnam;

³International University, Vietnam National University HCMC, Dong Hoa, Ho Chi Minh City, Vietnam;

⁴Institute of Information Technology and Electronics, Academy of Military Science and Technology, 17 Hoang Sam, Nghia Do, Hanoi, Vietnam.

*Corresponding author: hungnmpcu@gmail.com

Received 31 Jul. 2025; Revised 23 Sep. 2025; Accepted 10 Oct. 2025; Published 30 Oct. 2025.

DOI: <https://doi.org/10.54939/1859-1043.j.mst.IITE.2025.115-122>

ABSTRACT

In array signal processing, the Direction Of Arrival (DOA) of the wideband radio frequency signal source plays a key role in fields such as sonar, radar, communications, and electronic warfare. Root Multiple Signal Classification (Root-MUSIC) is one of the good algorithms in the subspace method. This algorithm is based on some assumptions, including the initial angular, but their performance is strongly degraded when these assumptions do not hold. The article presents a novel approach that uses Deep Neural Network (DNN) in combination with Root-MUSIC algorithm for improvement in DOA estimation. Observation samples are divided into distinguishable subspaces. The input autocorrelation matrix is trained by DNN to analyze real values. An estimator based on Super-resolution subspace improving the accuracy of direction-of-arrival estimation is presented and analyzed in the paper. Our simulation results show that wideband signal sources can be calculated in low SNR and limited snapshot conditions.

Keywords: Array Signal Processing; DOA Estimation; Deep neural network; Subspace method; Root-MUSIC.

1. INTRODUCTION

In practice, high-accuracy Direction of Arrival (DOA) estimation plays an important role in a wide range of real-world applications in radar, sonar and wireless communications. Accurately determining the direction of broadband signal sources is challenging due to the influence of noise, array configuration, and other system-related uncertainties. Due to the complexity of real-world environments, conventional methods are generally insufficient to effectively address these errors and maintain reliable estimation performance. Among the various approaches to DOA estimation, SubSpace-based methods - such as MUSIC [1], ESPRIT [2], and their broadband extensions have gained widespread popularity due to their high resolution and solid theoretical foundation. However, these methods often suffer performance degradation in challenging conditions, such as low SNR, coherent sources, limited snapshots, and broadband environments. The challenge becomes even more pronounced in broadband scenarios, where the frequency-dependent characteristics of the signals require specialized processing techniques. A common approach is to perform narrowband decomposition followed by coherent subspace combination [3]. While incoherent methods are simple to implement, they generally exhibit poor performance with coherent or correlated sources. Coherent Signal Subspace Methods (CSSMs) address this by aligning narrowband subspaces before combining them, but they often rely heavily on accurate model assumptions and can be sensitive to estimation errors in practical settings.

Recently, deep learning networks have achieved significant breakthroughs and have been

rapidly advancing in various domains, including signal processing. The proposed approach leverages a combination of Deep Neural Network (DNN) and classical subspace-based methods to enhance the accuracy of broadband DOA estimation. In this approach, DNN is designed as a deep mapping module capable of extracting features from broadband signals and transforming them into an equivalent narrowband representation in the form of a covariance matrix. This transformation enables the effective application of conventional subspace algorithms.

To overcome the limitations of classical methods, data-driven approaches have recently emerged as powerful alternatives. In particular, Convolutional Neural Networks (CNNs) have demonstrated strong capabilities in learning spatial and spectral patterns from array data, enabling robust DOA estimation even in adverse conditions [4-6]. Several recent works have explored CNN-based models trained directly on raw snapshots or covariance matrices for narrowband DOA estimation. However, extending these approaches to broadband, coherent signal settings remains challenging. Based on the research [7] and motivated by the complementary strengths of deep learning and subspace signal processing, this paper proposes a novel hybrid framework that integrates CNNs with coherent signal subspace methods for broadband source localization. Specifically, we introduce one type of SubspaceNet, a model that learns to construct an equivalent covariance matrix from autocorrelation-based input features, enabling classical subspace methods such as Root-MUSIC to operate effectively in broadband and coherent conditions.

2. PROBLEM

2.1. Theoretical background

We consider an ideal uniform linear array (ULA) consisting of m ($m = 1, 2, \dots, M$) elements and an inter-element spacing d ($d < \lambda / 2$). It is assumed that there are k ($k = 1, 2, \dots, P$) far-field Radio Frequency signal sources impinging from different directions. From [8], the DOA estimation using root-MUSIC algorithm is obtained by solving the equation:

$$\theta_k = \arcsin \left[\frac{\lambda}{2\pi d} \arg(z_k) \right] \quad (1)$$

Where θ_k is the angle of the signal and $z_k = e^{j \frac{2\pi d \sin \theta_k}{\lambda}}$.

Various subspace methods and the root-MUSIC algorithm in particular can achieve DOA estimation with a resolution that is not constrained by the physical aperture of the array. In many real-world scenarios, one must contend with broadband signals, coherent sources, calibration errors, a limited snapshot budget, and low SNR. Our aim is to develop a subspace technique that remains effective even when these ideal conditions for classical subspace algorithms are violated. In this paper, we focus on reconstructing the directly angles of arrival signal from the array's output of observation matrix. Although the number of signal sources is unknown, we restrict our attention to scenarios in which the number of signal sources is smaller than the number of elements of the antenna array. Similarly, the number of snapshots N is not assumed known during system design and may be limited to a minimal value, even as low as a single snapshot.

To mitigate the algorithm's reliance on ideal operating conditions, one may construct a supervised dataset D in which the data comprise J pairs of array observations and their corresponding DOAs, denoted as:

$$D = \left\{ \left(x(t)^{(j)}, \theta^{(j)} \right) \right\}_{j=1}^J \quad (2)$$

2.2. Method

The proposed method, tailored for a ULA, leverages the representational and reconstruction

capabilities of a deep neural network to transform a broadband input signal into an equivalent narrowband form expressed as an estimated covariance matrix in order to bolster subspace-based DOA recovery. We refer to this framework as one type of SubspaceNet. The key challenge is to extract the covariance-matrix features of the broadband observation and map them into an equivalent narrowband covariance estimate; this learned mapping can then be combined with any classical subspace algorithm without compromising its functionality.

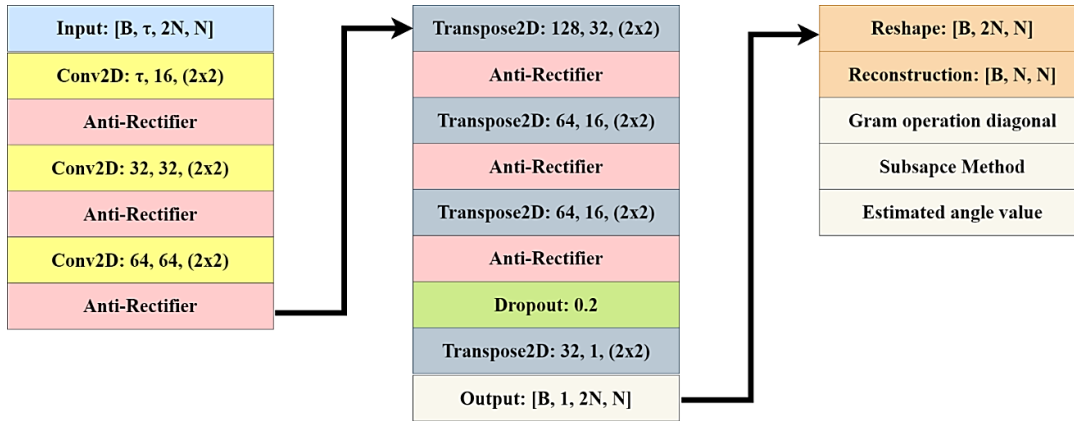


Figure 1. Structure of the SubspaceNet model.

Importantly, SubspaceNet does not rely on supervised labels for its DNN outputs. Instead, the quality of the network’s estimated covariance matrix is judged by the usefulness of the resulting DOA estimates. Concretely, the training loss is defined by comparing the DOA values obtained from the network’s covariance estimate against those produced by a chosen DOA estimator, as in equation (2). The research train is conducted through a first-order optimization of the DOA estimates. Hence, the DOA-estimation module must be differentiable - i.e., it must admit back-propagation of gradients from the estimated DOAs back to the covariance matrix.

The SubspaceNet architecture is engineered to transform the broadband input observation X into an estimated covariance matrix R , irrespective of the number of snapshots T . The model operates in three sequential stages: Feature Extraction - extracting from the broadband input the representations needed to form an equivalent covariance matrix for the DOA estimation task; Autoencoder Mapping - processing those features through an encoder-decoder network to reconstruct a preliminary narrowband-equivalent covariance estimate; Post-Processing - refining the DNN’s output to ensure Hermitian symmetry and positive semi-definiteness, as figure 1.

- **Feature Extraction:** The model is designed to compute an equivalent covariance matrix inspired by the Test of Orthogonality of Projected Subspaces (TOPS) technique. Under the hypothesis that, when steered to the correct angle, the signals at different frequencies lie in congruent subspaces, the goal is to convert the DOA features of a broadband signal into those of an equivalent narrowband signal. This is achieved by estimating covariance matrices as linear combinations of multivariate power spectral densities, using the Fourier transform of the autocorrelation function of the array outputs $x(t)$ as input features. The autocorrelation $\hat{R}_x[\tau]$ of the m -th array element’s output is computed as:

$$\hat{R}_x[\tau] = \frac{1}{T - \tau} \sum_{t=1}^{T-\tau} x(t)x^H(t + \tau) \quad (3)$$

Where $\tau > 0$ is the delay of the received signal $\tau \in \{0, \dots, \tau_{\max}\}$. These autocorrelation values across delays and elements form the feature set from which the model learns to reconstruct a narrowband-equivalent covariance matrix. The extracted features are then fed into the DNN as

follows: each resulting feature tensor has dimensions $N \times N \times \tau_{\max}$ independent of the number of snapshots T . Every tensor produced by the feature-extraction stage has a fixed size, enabling the network to learn covariance-relevant representations without variation in input dimensionality. This design both supplies the DNN with the essential information needed to reconstruct the covariance matrix and guarantees that the model's complexity remains unaffected by the snapshot count.

- **Auto-encoder and Auto-decoder:** DNN architecture inspired by denoising autoencoders is utilized to handle the covariance matrix. The Auto-encoder is a deep learning architecture used to extract relevant features from signal and noise data, and subsequently map them into equivalent representations. The encoder part of the DNN consists of three convolutional layers with progressively increasing filter sizes. Its objective is to map the input into a lower-dimensional latent feature space while preserving only the most informative components for subsequent processing. The decoder comprises transposed convolutional layers with decreasing filter sizes, aiming to reconstruct the hidden features that were compressed during encoding. The final output is post-processed into an equivalent covariance matrix.

- **Post-processing of the output:** the output of the autoencoder is a real-valued matrix $2N \times N$ because this results from the separation of the real and imaginary components during the input processing stage. To convert the output into a surrogate covariance matrix, the real and imaginary parts at the output are first combined to form a complex matrix K of size $N \times N$. Subsequently, to construct the surrogate covariance matrix, matrix K is transformed into a Hermitian positive semi-definite (PSD) covariance estimate using the following formula:

$$\hat{R} = KK^H + \varepsilon I_N \quad (4)$$

Where $\varepsilon > 0$ is a fixed architectural hyperparameter. The trainable parameters of the architecture are the weights of the autoencoder.

2.3. Training procedure using ADAM

The Root-MUSIC algorithm is selected as the subspace-based method for estimating DOAs from the covariance matrix output of SubspaceNet. The DOA estimates are obtained from the roots of an implicit polynomial function. This property enables end-to-end optimization of the DNN parameters in SubspaceNet using first-order methods such as ADAM. The training procedure of SubspaceNet using ADAM is described as follows:

- Training Parameter Settings: The following hyperparameters are used during training: Delay τ_{\max} ; Batch size $B=256$; Learning rate $\mu = 10^{-6}$; Number of training epochs $e_{\max} = 80$.

- Input Training Dataset: $D=10^5$ samples.

- Weight Initialization for the CNN Model.

- In the epochs, implementation:

- Randomly split the training dataset D into B mini-batches.
- With $b = 1, 2, \dots, B$ implementation
 - Computing the covariance matrix $\hat{R}(X_b^{(j)}; \psi)$ with $(X_b^{(j)}, \theta_b^{(j)}) \in D_b$
 - Estimation $\hat{\theta}(\hat{R}(X_b^{(j)}; \psi))$ by *Root-MUSIC*
 - Calculating $L_{D_b}(\psi)$ by the equation:

$$L_{D_b}(\psi) = \frac{1}{|D|} \sum_{(X^{(j)}, \theta^{(j)}) \in D} l(X^{(j)}, \theta^{(j)}; \psi) \quad (5)$$

- Updating the Weight by $\psi \leftarrow \psi - \mu \nabla_{\psi} L_{D_b}(\psi)$

3. RESULTS AND DISCUSSION

3.1. Signal model

In conducting model evaluation, the simulation of a Uniform Linear Array (ULA) with M sensors receiving broadband signals from different directions, P is considered. The broadband signal model in the simulation can be designed in the frequency domain, at a frequency ω given by [9]:

$$\chi(\omega) = A(\omega, \theta)S(\omega) + \nu(\omega) \quad (6)$$

Where $\chi(\omega)$ and $S(\omega)$ are the Fourier transform of $x(t)$ and $s(t)$, $A(\omega, \theta)$ is the driven vector, $\nu(\omega)$ is Gaussian noise.

The research simulates broadband signals in the time domain using an Orthogonal Frequency-Division Multiplexing (OFDM) signal with L subcarrier frequencies, with the satisfied broadband condition. To evaluate the performance of the SubspaceNet model in wideband signal DOA estimation, we employ the RMSPE (Root Mean Squared Periodic Error) metric. As previously mentioned regarding the periodic nature of DOA values, RMSPE addresses this characteristic by computing the angular error as the arc-length distance on the unit circle. This allows for a more accurate and objective assessment of DOA estimation models. The RMSPE is defined by the following formula:

$$RMSPE = \sqrt{E \left[\left((\hat{\theta} - \theta) \bmod 2\pi \right)^2 \right]} \quad (7)$$

Where are $\hat{\theta}$ and θ estimated and true arrival angles of the evaluated signals. The performance of the SubspaceNet model combined with the root-MUSIC algorithm is compared with the conventional approach, in which the wideband signals are divided into frequency bins using the FFT. The root-MUSIC algorithm is then applied to each frequency bin as if it were a narrowband signal, and the final result is obtained by aggregating the estimations across all bins.

3.2. Wideband and coherent OFDM signals

In this section, simulations and evaluations are conducted with a number of incoming signal sources $P = 3$ and sources are mutually incoherent.

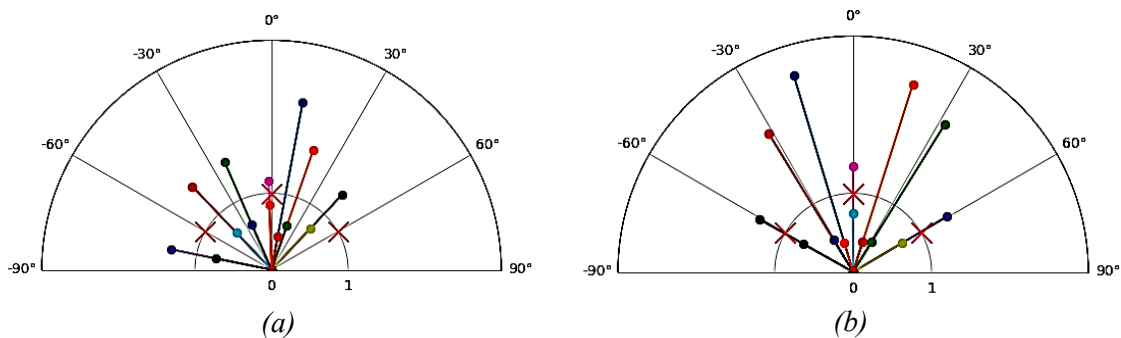


Figure 2. Simulation results of Root-MUSIC (a) and SubspaceNet+Root-MUSIC (b) methods.

The directions of arrival (DOAs) of the sources are $-60^\circ, 0^\circ$ and 60° , respectively. The array consists of $M=8$ elements, and the received signal has an $SNR=10$ dB with 200 snapshots. The estimated polynomial root spectrum of the root-MUSIC algorithm is illustrated as the figure 2. The proposed method, which combines SubspaceNet with the root-MUSIC algorithm, clearly yields more accurate and reliable results compared to the conventional approach that estimates DOAs independently on individual frequency bins and then aggregates the results. This is intuitive, as the

steering vector varies with frequency and the phase incoherence between sources prevents the independent bin-wise processing from fully capturing the frequency-domain relationships across the entire signal spectrum. This limitation leads to higher estimation errors, especially under high noise conditions or system imperfections. In contrast, SubspaceNet combined with root-MUSIC enables a more holistic treatment of wideband signals. The deep learning model is capable of learning subspace features from the entire frequency spectrum without relying on coherence among sources.

3.3. Evaluation of the impact of the number of Snapshots

The number of snapshots is a critical factor that directly affects the performance of DOA estimation algorithms. As the number of snapshots increases, the estimation of the signal covariance matrix becomes more accurate. However, in practical scenarios where only a small number of snapshots are available, conventional algorithms tend to perform poorly. In contrast, deep learning-based models such as SubspaceNet can maintain robust performance due to their ability to learn from prior data and their independence from snapshot quantity during training.

Therefore, under conditions of reduced snapshot availability, the proposed model demonstrates higher stability. The deep neural network is capable of interpolating and learning appropriate representations even with incomplete input information, thereby preserving estimation accuracy. In this experiment, the signal parameters and simulation conditions are kept the same as in section 3.2, with the only difference being that the number of snapshots is halved to 100. The estimation results of both methods under this significantly reduced snapshot scenario are presented in the table 1 below. Simulation results demonstrate that the proposed approach yields marginally improved performance compared to [7].

Table 1. RMSPE results in small snapshot condition.

Algorithm	RMSPE [°]
Root-MUSIC	28.3456
SubspaceNet+root-MUSIC	1.6731

3.4. Evaluation of the impact of SNR

In the root-MUSIC algorithm, the estimated covariance matrix becomes increasingly dominated by noise, which distorts the subspace structure and leads to large estimation errors at the low SNR value. However, deep learning-based models such as SubspaceNet exhibit improved robustness to low-SNR conditions due to their ability to learn denoised and generalized representations from training data. Figure 3 shows that in the low SNR (below 5 dB), the RMSPE of SubspaceNet increases more slowly compared to classical algorithms, indicating higher stability and noise resilience. In the high SNR (above 10 dB), all methods tend to converge to their optimal performance. It means that SubspaceNet still achieves lower estimation errors thanks to being trained directly with a loss function defined on DOA error. The experiments are conducted under the same signal parameters and simulation in various SNR levels in the range from 5 dB to 10 dB.

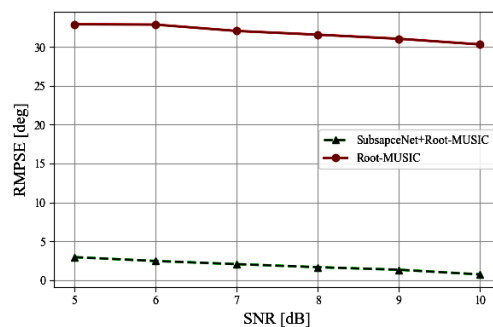


Figure 3. Impact of SNR on estimation error.

3.5. Evaluation of the impact of the number of array elements

The number of elements in the antenna array is a critical factor that determines angular resolution and the ability to separate signal sources in DOA estimation problems. As the number of elements increases, the signal subspace gains more degrees of freedom, enabling classical subspace algorithms to better distinguish between sources with closely spaced angles. In addition, the peaks in the estimated spatial spectrum become sharper, enhancing the accuracy and stability of DOA determination.

The experiment is conducted under the same signal parameters and simulation conditions as in section 3.2, with the number of array elements set to $M = 4, 6$ and 8 . Table 2 shows that the performance of conventional methods typically degrades rapidly when the number of elements is reduced to a minimal configuration (e.g., equal to the number of sources). In contrast, the proposed SubspaceNet model demonstrates adaptability to flexible array configurations, even when the number of array elements is limited. Through training on diverse data, the deep learning network can effectively extract the remaining spatial features, maintaining a lower RMSPE compared to traditional algorithms in scenarios with a small number of sensors. Moreover, in cases with larger array sizes, SubspaceNet continues to leverage structural advantages and converges toward optimal performance.

Table 2. RMSPE results with the number of array elements set to $M = 4, 6$ and 8 .

RMSPE [$^{\circ}$]	$M = 4$	$M = 6$	$M = 8$
Root-MUSIC	29.1416	28.1682	28.6462
SubspaceNet+root-MUSIC	4.4769	1.1076	0.7240

4. CONCLUSIONS

In this article, the DOA simulation results in this research proved that wideband signal sources can be analyzed and compared in low SNR and limited snapshot conditions. As it can be seen from the simulation, estimation performance using convolutional neural networks combined with coherent signal subspace is one of the good solutions against the challenges of low SNR, coherent sources, limited snapshots, and broadband environments. The contributions of this research include the following:

- A hybrid architecture that combines CNN-based feature learning with coherent signal subspace methods is proposed for broadband DOA estimation.
- Moreover, the article demonstrates that the proposed method achieves superior performance compared to both traditional CSSMs and purely learning-based approaches, particularly under low SNR and limited snapshot scenarios.

REFERENCES

- [1]. P. R. O. Schmidt, "Multiple emitter location and signal parameter estimation," IEEE transactions on antennas and propagation, Vol. 34, No. 3, pp. 276–280, (1986).
- [2]. R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters via rotational invariance techniques," IEEE transactions on acoustics, speech, and signal processing, Vol. 37, No. 7, pp. 984–995, (1989).
- [3]. M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," IEEE transactions on acoustics, speech, and signal processing, Vol. 36, No. 4, pp. 477–489, (1988).
- [4]. Y. Jiang, X. Zhang, and W. Liu, "Deep neural networks for direction-of-arrival estimation with massive MIMO systems," in proceedings of the IEEE global communications conference (GLOBECOM), pp. 1–6, (2017).
- [5]. L. Liu, C. Huang, and Y. Shi, "Convolutional neural network-based DOA estimation using multiple snapshots," in proceedings of the IEEE international conference on acoustics, speech and signal processing (ICASSP), pp. 3286–3290, (2018).

- [6]. D. Kim and J. C. Ye, "Deep learning-based direction-of-arrival estimation with circular microphone arrays," *IEEE signal processing letters*, Vol. 26, No. 8, pp. 1177–1181, (2019).
- [7]. D. H. Shmuel, J. P. Merkofer, G. Revach, R. J. G. van Sloun, and N. Shlezinger, "SubspaceNet: Deep learning-aided subspace methods for DoA estimation," *arXiv preprint*, arXiv:2306.02271, doi:10.48550/arXiv.2306.02271, (2024).
- [8]. L. M. K., J. H. M., and Y.-S. Yoon, "TOPS: new DOA estimator for wideband signal," *IEEE transactions on signal processing*, Vol. 54, No. 6, pp. 1977–1989, (2006).
- [9]. K. Forsythe, "Utilizing waveform features for adaptive beamforming and direction finding with narrowband signals," *Lincoln laboratory journal*, Vol. 10, No. 2, pp. 35–50, (1997).

TÓM TẮT

Định hướng nguồn tín hiệu dải rộng sử dụng mạng nơ-ron tích chập kết hợp phương pháp không gian con tín hiệu đồng kết hợp

Trong xử lý tín hiệu mảng, ước lượng hướng đến của nguồn tín hiệu tần số vô tuyến băng rộng đóng vai trò quan trọng trong các lĩnh vực như sonar, radar, thông tin liên lạc, tác chiến điện tử. Root-MUSIC là một trong những thuật toán tốt trong các phương pháp không gian con. Thuật toán này dựa trên một số giả định bao gồm góc ban đầu nhưng hiệu quả ước lượng của thuật toán bị giảm mạnh khi những giả định này không đúng. Bài báo trình bày một phương pháp mới sử dụng mạng nơ-ron sâu (DNN) kết hợp với thuật toán Root-MUSIC để cải thiện ước lượng hướng đến của nguồn tín hiệu. Các mẫu quan sát được chia thành các không gian con có thể phân biệt được. Ma trận tự tương quan đầu vào được DNN đào tạo để phân tích các giá trị thực. Bộ ước lượng dựa trên không gian con siêu phân giải giúp cải thiện độ chính xác được trình bày và phân tích trong bài báo. Kết quả mô phỏng của chúng tôi cho thấy các nguồn tín hiệu băng rộng có thể được tính toán trong điều kiện SNR thấp và các mẫu quan sát ngắn.

Từ khoá: Xử lý tín hiệu mảng; Ước lượng hướng tín hiệu đến; Mạng thần kinh học sâu; Phương pháp không gian con; Root-MUSIC.