

DOA-CNN: Mô hình tự động hiệu chỉnh sai số hệ thống, nâng cao độ chính xác ước lượng hướng đến của tín hiệu vô tuyến

Nguyễn Duy Thái^{1*}, Hoàng Văn Phúc², Lê Thanh Hải¹

¹Viện Điện tử, Viện Khoa học và Công nghệ quân sự;

²Viện Tích hợp hệ thống, Học viện Kỹ thuật quân sự.

*Email: ndthai03@gmail.com

Nhận bài: 23/5/2023; Hoàn thiện: 03/8/2023; Chấp nhận đăng: 08/8/2023; Xuất bản: 25/8/2023.

DOI: <https://doi.org/10.54939/1859-1043.j.mst.89.2023.43-51>

TÓM TẮT

Bài báo trình bày kết quả nghiên cứu đề xuất mạng nơ-ron tích chập học sâu, đặt tên là DOA-CNN, cho bài toán ước lượng hướng đến (DOA: Direction of Arrival) của tín hiệu vô tuyến sử dụng mảng ăng ten tuyến tính đồng đều (ULA: Uniform Linear Array) nhằm nâng cao độ chính xác ước lượng trong các trường hợp xảy ra lỗi hệ thống phổ biến, như: sai lệch vị trí các phần tử của mảng ăng ten; sai số biên độ và pha gây ra do sai lệch đường truyền tuyến thu. Mô hình DOA-CNN được thiết kế với các lớp nhân tích lũy nhằm tăng cường đặc trưng đại diện cho dữ liệu thông qua các tầng tích chập, từ đó có thể thực hiện bài toán phân lớp góc DOA. Kết quả đánh giá, so sánh độ chính xác ước lượng của mô hình DOA-CNN với các phương pháp CBF, Capon, MUSIC, Root-MUSIC, ESPRIT và một số mô hình học máy khác cho thấy: nếu cấu hình của mảng ULA và máy thu là lý tưởng thì Root-MUSIC và ESPRIT cho độ chính xác tốt nhất, do các phương pháp này tính trực tiếp góc DOA trong khi những phương pháp còn lại ước lượng DOA thông qua phổ góc nên độ chính xác phụ thuộc vào độ phân giải phổ; Tuy nhiên, nếu xuất hiện sai số vị trí các phần tử mảng ULA và sai lệch đường truyền trong tuyến thu thì mô hình DOA-CNN đề xuất cho độ chính xác vượt trội so với các phương pháp truyền thống và tốc độ xử lý nhanh hơn một số mô hình học máy khác.

Từ khóa: Ước lượng hướng đến; Mạng nơ-ron tích chập; Sai số vị trí; Sai số biên độ và pha.

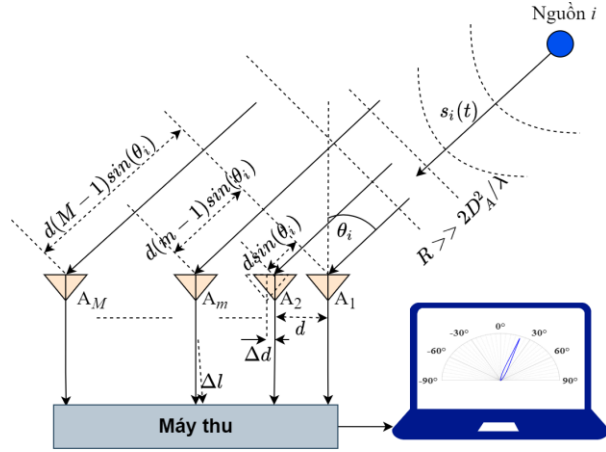
1. ĐẶT VẤN ĐỀ

Ước lượng hướng đến của tín hiệu vô tuyến là một trong những kỹ thuật xử lý mảng quan trọng, được áp dụng trong nhiều hệ thống từ thông tin vô tuyến đến trinh sát, tác chiến điện tử. Có rất nhiều phương pháp từ truyền thống đến hiện đại được áp dụng để ước lượng hướng đến của tín hiệu vô tuyến, có thể chia thành các nhóm như sau: (i) phương pháp so sánh tham số tín hiệu (so sánh biên độ, so sánh pha, hiệu thời gian tới) [1]; (ii) phương pháp quét búp sóng (tạo búp sóng thông thường (CBF: Conventional Beamforming), đáp ứng không biến dạng phương sai cực tiểu (MVDR: Minimum Variance Distortionless Response)) [2]; phương pháp phân tích không gian con tín hiệu (phân loại đa tín hiệu (MUSIC: Multiple Signal Classification) [3] và ước lượng tham số tín hiệu bằng kỹ thuật bất biến xoay (ESPRIT: Estimation of Signal Parameters via Rotational Invariance Techniques)) [4]. Ngoài ra, một số phương pháp mở rộng như: Root-MUSIC [5], Smoothing MUSIC [6], được đề xuất nhằm nâng cao độ chính xác ước lượng DOA trong trường hợp các nguồn tín hiệu vô tuyến tương quan. Điểm chung của tất cả các phương pháp nêu trên là dựa vào mô hình mảng ăng ten giả định từ trước, vì vậy chúng rất nhạy cảm với các sai số hệ thống của mảng ăng ten.

Các nghiên cứu về các phương pháp ước lượng DOA nêu trên đều cho thấy sai số ước lượng phụ thuộc vào các yếu tố như: khâu độ mảng ăng ten, số lượng phần tử trong mảng, khoảng cách giữa các phần tử, số lượng mẫu tín hiệu và góc đến của tín hiệu. Tuy nhiên, các nghiên cứu chủ yếu giả định mảng ăng ten cũng như đường truyền trong tuyến thu là lý tưởng (chính xác hóa vị trí các phần tử, các phần tử là đẳng hướng và giống nhau, đường truyền từ các phần tử đến máy thu giống nhau), ít đề cập sự ảnh hưởng của các yếu tố hay xảy ra trong thực tế như sai lệch vị trí

đặt các phần tử và sai khác đường truyền trong tuyến thu tới độ chính xác ước lượng. Có một số nghiên cứu khi xem xét vấn đề này đã đề xuất thực hiện đo đặc hiệu chỉnh nhằm khắc phục các sai số ước lượng, tuy nhiên việc đo đặc hiệu chỉnh không giải quyết triệt để vấn đề sai pha và biên độ (tập hợp bởi nhiều nguyên nhân khác nhau) [7]. Gần đây, việc áp dụng mạng nơ-ron học sâu cho bài toán ước lượng DOA tín hiệu vô tuyến đang nhận được sự quan tâm rộng rãi từ các nhà khoa học [8]. Do không phụ thuộc nhiều vào mô hình mảng ăng ten dựng trước, có thể cập nhật trọng số theo dữ liệu đo thực tế và tự động điều chỉnh sai số hệ thống, mạng nơ-ron học sâu hứa hẹn là phương pháp đầy tiềm năng cho việc nâng cao độ chính xác ước lượng DOA [9].

Trong bài báo này, mô hình mạng nơ-ron tích chập DOA-CNN được đề xuất để loại bỏ sai số ước lượng DOA cho mảng ăng ten ULA do ảnh hưởng bởi các yếu tố: sai lệch vị trí đặt phần tử, sai lệch đường truyền từ các phần tử đến máy thu. Với dữ liệu thu giả định tồn tại các sai số nêu trên, kết quả mô phỏng cho thấy, mô hình mạng đề xuất cải thiện đáng kể độ chính xác ước lượng DOA so với các phương pháp truyền thống như CBF, Capon, MUSIC, Root-MUSIC và ESPRIT, đảm bảo tốc độ xử lý nhanh so với các mô hình học máy khác.



Hình 1. Mô hình thu tín hiệu mảng ăng ten ULA.

2. MÔ HÌNH THU TÍN HIỆU

Để phân tích, đánh giá, so sánh sai số ước lượng DOA, một mô hình thu tín hiệu của mảng ăng ten ULA với M phần tử được xây dựng như minh họa trong hình 1.

Giả thiết có P nguồn tín hiệu truyền tới mảng ăng ten trong mặt phẳng phương vị ở các góc lần lượt là $\theta_1, \theta_2, \dots, \theta_p$. Khi đó, tín hiệu tại đầu ra của mảng ULA là:

$$\mathbf{x}(t) = \mathbf{A}(\theta) \cdot \mathbf{s}(t) + \mathbf{n}(t), \quad (1)$$

trong đó, $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_p(t)]^T$ là véc-tơ tín hiệu của P nguồn, $\mathbf{n}(t) = [n_1(t), n_2(t), \dots, n_M(t)]^T$ là véc-tơ của nhiễu tạp tương ứng với M kênh thu, $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T$ là véc-tơ tín hiệu đầu ra của mảng ULA, $\mathbf{A}(\theta)$ là ma trận định hướng của mảng, được biểu diễn như sau [2]:

$$\mathbf{A}(\theta) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-j\frac{2\pi d_1 \sin \theta_1}{\lambda}} & e^{-j\frac{2\pi d_1 \sin \theta_2}{\lambda}} & \dots & e^{-j\frac{2\pi d_1 \sin \theta_p}{\lambda}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j\frac{2\pi d_{M-1} \sin \theta_1}{\lambda}} & e^{-j\frac{2\pi d_{M-1} \sin \theta_2}{\lambda}} & \dots & e^{-j\frac{2\pi d_{M-1} \sin \theta_p}{\lambda}} \end{bmatrix}. \quad (2)$$

Nếu tồn tại các sai lệch thiết đặt vị trí phần tử ăng ten (minh họa Δd trong hình 1) và sai số đường truyền (minh họa Δl trong hình 1) thì sẽ phát sinh các sai pha và biên độ trong ma trận định hướng. Chọn phần tử 1 là tham chiếu, sai số vị trí đặt của các phần tử còn lại sẽ gây ra sai pha đối với nguồn thứ p như sau:

$$\Delta\varphi_{d_m, \theta_p} = \frac{2\pi\Delta d_m \sin \theta_p}{\lambda}, \quad (3)$$

trong đó, Δd_m là sai số vị trí đặt của phần tử ăng ten thứ m, θ_p là góc đến của nguồn thứ p. Tín hiệu truyền từ các phần tử đến máy thu đi qua các mạch đường truyền, mạch lọc và khuếch đại,

do đó, sẽ có sự lệch pha và suy hao biên độ nhất định trước khi được lấy mẫu chuyển sang dữ liệu số. Những sai lệch về pha và biên độ của kênh thu thứ $m-1$ có thể biểu diễn bởi $\Delta\varphi_{m-1}$ và α_{m-1} . Như vậy, ma trận định hướng $\mathbf{A}(\theta)$ trong phương trình (2) có thể được viết lại như sau:

$$\mathbf{A}(\theta, \varepsilon) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \alpha_1 e^{-j\left(\frac{2\pi(d_1+\Delta d_1)\sin\theta_1}{\lambda} + \Delta\varphi_1\right)} & \alpha_1 e^{-j\left(\frac{2\pi(d_1+\Delta d_1)\sin\theta_2}{\lambda} + \Delta\varphi_1\right)} & \dots & \alpha_1 e^{-j\left(\frac{2\pi(d_1+\Delta d_1)\sin\theta_p}{\lambda} + \Delta\varphi_1\right)} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{M-1} e^{-j\left(\frac{2\pi(d_{M-1}+\Delta d_{M-1})\sin\theta_1}{\lambda} + \Delta\varphi_{M-1}\right)} & \alpha_{M-1} e^{-j\left(\frac{2\pi(d_{M-1}+\Delta d_{M-1})\sin\theta_2}{\lambda} + \Delta\varphi_{M-1}\right)} & \dots & \alpha_{M-1} e^{-j\left(\frac{2\pi(d_{M-1}+\Delta d_{M-1})\sin\theta_p}{\lambda} + \Delta\varphi_{M-1}\right)} \end{bmatrix} \quad (4)$$

trong đó, $\varepsilon = \alpha e^{-j\left(\frac{2\pi\Delta d \sin\theta}{\lambda} + \Delta\varphi\right)}$ đại diện cho lỗi pha và biên độ trong ma trận định hướng. Phương trình (1) được viết lại như sau:

$$\mathbf{x}(t, \varepsilon) = \mathbf{A}(\theta, \varepsilon) \cdot \mathbf{s}(t) + \mathbf{n}(t). \quad (5)$$

Tại máy thu, tín hiệu tương tự nhận được từ các phần tử ăng ten sẽ được chuyển sang dạng số để xử lý. Gọi T_s là chu kỳ lấy mẫu của bộ biến đổi ADC, đầu ra của nó được biểu diễn như sau:

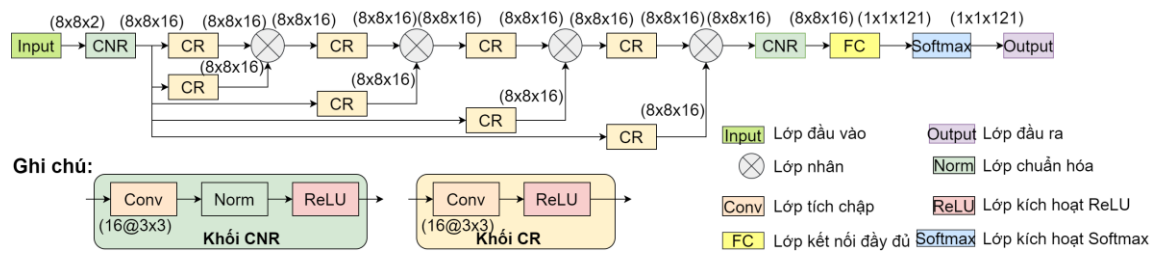
$$\mathbf{x}(nT_s, \varepsilon) = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_M \end{bmatrix} = \begin{bmatrix} x_1(T_s) & x_1(2T_s) & \dots & x_1(NT_s) \\ x_2(T_s) & x_2(2T_s) & \dots & x_2(NT_s) \\ \vdots & \vdots & \ddots & \vdots \\ x_M(T_s) & x_M(2T_s) & \dots & x_M(NT_s) \end{bmatrix}, \quad (6)$$

ở đây, N là số lượng mẫu trong một cửa sổ tín hiệu. Tín hiệu thu \mathbf{x} là ma trận số có kích thước $M \times N$. Ma trận hiệp phương sai \mathbf{R}_x của tín hiệu thu được tính theo công thức sau [2]:

$$\mathbf{R}_x = \mathbf{x}\mathbf{x}^H = \mathbf{A}\mathbf{s}\mathbf{s}^H\mathbf{A}^H + \mathbf{R}_n = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \sigma_n^2\mathbf{I} \quad (7)$$

trong đó, σ_n^2 đại diện cho công suất trung bình của nhiễu, \mathbf{I} là ma trận đơn vị, \mathbf{R}_s là ma trận hiệp phương sai của tín hiệu nguồn. Như vậy, ma trận \mathbf{R}_x có kích thước $M \times M \times 2$ sẽ là dữ liệu đầu vào của mô hình DOA-CNN đề xuất và các thuật toán ước lượng khác, trong đó, 2 thể hiện cho việc bao gồm cả phần thực và phần ảo của ma trận \mathbf{R}_x .

3. MÔ HÌNH DOA-CNN



Hình 2. Sơ đồ mô hình DOA-CNN đề xuất.

Mô hình mạng nơ-ron tích chập giải quyết các vấn đề đặt ra bên trên của bài toán ước lượng DOA tín hiệu vô tuyến được chúng tôi đề xuất và đặt tên là DOA-CNN. Cấu trúc tổng quát của DOA-CNN được thể hiện như trên hình 2. Có thể thấy, mô hình được cấu thành từ nhiều lớp, nhiều khối để trích xuất đặc trưng dữ liệu. Đầu ra mỗi lớp (hoặc khối) có kích thước $w \times h \times c$ thể hiện kích thước ma trận trực trung, ví dụ: lớp đầu vào (ma trận hiệp phương sai) có kích thước $(8 \times 8 \times 2)$, đầu ra của các khối CNR (Convolution-Normalization-ReLU), CR (Convolution-ReLU) và lớp nhân có kích thước $(8 \times 8 \times 16)$. Khối CNR gồm ba lớp nối tiếp: tích chập (Conv), chuẩn hóa (Norm) và kích hoạt (ReLU). Khối CR gồm hai lớp Conv và ReLU. Các khối CR kết

hợp với nhau thông qua lớp nhân nhằm tăng cường tính đại diện các đặc trưng dữ liệu theo phép nhân. Đặt đầu ra khối CNR thứ nhất là z_1 thì đầu vào của khối CNR thứ hai sẽ được biểu diễn như sau:

$$z_2 = F_{CR} \{F_{CR} \{F_{CR} \{F_{CR} \{z_1\} \otimes F_{CR} \{z_1\}\} \otimes F_{CR} \{z_1\}\} \otimes F_{CR} \{z_1\}\} \otimes F_{CR} \{z_1\}, \quad (8)$$

trong đó, $F_{CR} \{z_1\}$ thể hiện phép toán truyền thẳng của khối CR với đặc trưng đầu vào là z_1 ; ký hiệu \otimes thể hiện phép nhân từng phần tử của hai ma trận.

5	6	1	5	4	5	8	8
8	3	4	8	8	5	3	8
5	5	4	1	6	5	8	1
7	3	5	1	5	3	1	6
6	1	2	7	4	5	2	6
4	7	1	1	8	3	7	4
2	7	2	8	7	4	5	4
1	8	8	7	5	5	4	4

 \otimes

2	1	2
2	3	1
1	2	1

 $=$

41	52	45	56	61	62	74	46
61	66	60	80	78	84	79	68
57	73	51	68	71	73	69	38
50	62	52	49	68	57	64	39
55	48	52	62	59	68	55	51
55	52	56	66	82	75	68	47
52	72	77	84	86	73	70	44
30	59	65	64	58	48	43	29

Hình 3. Minh họa phép tích chập trong lớp tích chập.

Để thực hiện được phép toán nhân trong mô hình cần phải thiết kế các lớp sao cho hai đầu vào của lớp nhân có cùng kích thước. Nếu đầu vào của khối CNR là z_2 thì đầu ra của nó có thể được biểu diễn như sau:

$$z_3 = \max\{0, \text{Norm}\{\text{Conv}\{z_2\}\}\}, \quad (9)$$

trong đó, $\max\{0, z\}$ là biểu thức của hàm ReLU trong lớp kích hoạt với z là biến đầu vào của hàm. $\text{Norm}\{-\}$ thể hiện hàm chuẩn hóa trong lớp chuẩn hóa, được tính như sau [10]:

$$\hat{z} = \text{Norm}\{z\} = \frac{z - \mu_B}{\sqrt{\sigma_B^2 - \gamma}}, \quad (10)$$

ở đây, μ_B và σ_B^2 là giá trị trung bình và phương sai của cụm dữ liệu gồm B đầu vào, γ là một hằng số rất nhỏ được thêm vào nhằm tránh trường hợp xảy ra với phương sai $\sigma_B^2 = 0$. $\text{Conv}\{z_2\}$ thể hiện phép tính tích chập của đặc trưng đầu vào z_2 với bộ lọc kích thước $k \times k$. Trong trường hợp bài toán đang xem xét: kích thước bộ lọc được lựa chọn bằng 3×3 , bước trượt là $(1, 1)$, với bản đồ đặc trưng có kích thước là 8×8 , phép tích chập được minh họa như trong hình 3. Có thể thấy, phép tích chập 2 chiều thực hiện phép toán nhân từng phần tử của bộ lọc với từng phân vùng (cùng kích thước) của ma trận đặc trưng đầu vào, sau đó tính tổng gộp để được một giá trị tương ứng cho một phần tử của ma trận đầu ra. Trong trường hợp này, các hàng và cột ngoại biên của ma trận vào được thêm với giá trị bằng 0 để ma trận kết quả có cùng kích thước với ma trận vào. Lớp kết nối đầy đủ (FC: Fully connected) thực hiện duỗi thẳng ma trận đặc trưng đầu vào của nó, sau đó thực hiện phép toán như đối với mạng nơ-ron nhân tạo tiêu chuẩn [10]. Trong nghiên cứu này, lớp FC được chỉ định với 121 nơ-ron tương ứng với 121 phân lớp góc DOA từ -60° đến 60° , bước cách 1° . Lớp softmax được sử dụng để tính toán xác suất cho các phân lớp. Softmax là một hàm kích hoạt được sử dụng chủ yếu cho lớp cuối cùng của bài toán phân loại đa lớp. Nó có đặc tính biến một véc-tơ với K phần tử có giá trị thực thành một véc-tơ gồm K phần tử mà tổng giá trị các phần tử bằng 1. Hàm Softmax được biểu diễn như sau [10]:

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_{i=1}^K e^{z_i}} \quad (11)$$

Cuối cùng, mô hình sẽ quyết định góc ước lượng dựa vào xác suất tương ứng với mỗi góc.

4. ĐÁNH GIÁ MÔ HÌNH DOA-CNN

4.1. Mô tả dữ liệu

Trước khi mô hình DOA-CNN được sử dụng để thực thi quá trình ước lượng hướng đến, nó

phải được huấn luyện với tập dữ liệu được tạo ra với các tham số giả định liệt kê trong bảng 1. Cụ thể, một tập dữ liệu $D = \{\mathbf{R} | \theta\}$ được tạo ra với đầu vào là ma trận hiệp phương sai \mathbf{R} , nhãn đầu ra là các góc θ tương ứng. Bằng giải pháp lựa chọn ngẫu nhiên phân bố đều các góc trong giới hạn -60° đến $+60^\circ$, bước cách 1° và các mức nhiễu với SNR trong giới hạn -20 dB đến 10 dB, bước cách 1 dB, có thể tạo ra tổng cộng 100000 cặp $\{\mathbf{R} | \theta\}$, trong đó, 80000 cặp (80%) được dùng cho huấn luyện và 20000 cặp (20%) dùng để kiểm chứng mô hình.

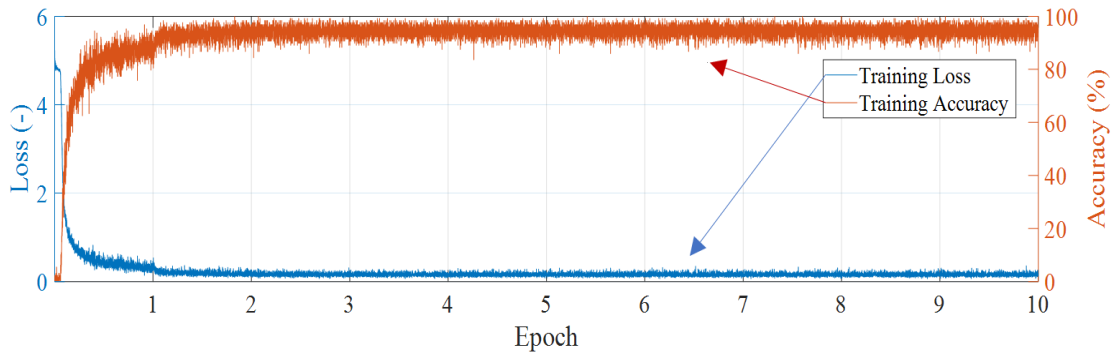
Bảng 1. Tham số giả định tạo tập dữ liệu huấn luyện và kiểm chứng mô hình DOA-CNN.

Tham số	Giá trị	Tham số	Giá trị
Tần số trung tâm	1090 MHz	Mảng ăng ten	ULA
Băng thông	10 MHz	Khoảng cách các phần tử	$d = \lambda/2$
Số lượng mẫu	1024	Số lượng phần tử	8
Góc đến	$U\{-60^\circ, 60^\circ, 1^\circ\}$	Mức nhiễu	$U\{-20, 10, 1\}$ dB

Chú thích: $U\{a, b, c\}$ thể hiện phép toán tạo dữ liệu ngẫu nhiên phân bố đều trong giới hạn $[a, b]$ với bước cách c .

4.2. Huấn luyện mô hình DOA-CNN

Mô hình DOA-CNN đề xuất được xây dựng, huấn luyện và đánh giá bằng phần mềm Matlab. Các tham số huấn luyện được thiết lập như sau: tốc độ học khởi tạo là $0,01$ và sẽ giảm 10 lần sau 2 lượt (epoch) huấn luyện; số lượt huấn luyện là 10 ; trình tối ưu SGDM với hệ số moment = $0,9$; hàm mất mát Cross-Entropy được sử dụng. Mô hình DOA-CNN được huấn luyện và kiểm tra trên nền tảng phần cứng GPU của máy Laptop với cấu hình CPU Intel Core i5 9300H, RAM 16 GB và GPU GTX 1660ti. Hình 4 thể hiện sự thay đổi giá trị mất mát và độ chính xác phân loại góc trong quá trình huấn luyện mô hình DOA-CNN. Có thể nhận thấy, sau hai lượt huấn luyện các trọng số của mô hình bắt đầu ổn định. Sau 10 lượt huấn luyện thì mô hình cho độ chính xác phân loại góc đạt $94,5\%$.



Hình 4. Sự thay đổi giá trị mất mát và độ chính xác phân loại góc trong quá trình huấn luyện mô hình DOA-CNN.

4.3. Đánh giá độ chính xác ước lượng góc của mô hình DOA-CNN

Mô hình DOA-CNN sau huấn luyện sẽ được đánh giá, so sánh độ chính xác ước lượng góc với các phương pháp truyền thống: CBF, Capon, MUSIC, Root-MUSIC, ESPRIT và các phương pháp học máy như cây quyết định (DT: Decision Tree) [11], GNB (Gaussian Naive Bayes) [12], KNN (K-Nearest Neighbor) [13], MLP (Multi-Layer Perceptron) [8] và CNN [14] thông qua việc cùng thực hiện ước lượng góc DOA của 20000 mẫu dữ liệu với các giá trị SNR khác nhau, thay đổi từ -20 dB đến 10 dB, bước cách 1 dB. Lưu ý, dữ liệu dùng cho đánh giá không được dùng để huấn luyện mô hình DOA-CNN.

Tham số đánh giá là sai số ước lượng quân phương, được tính theo công thức sau:

$$RMSE_{DOA} = \sqrt{\frac{1}{QP} \sum_{i=1}^Q \sum_{j=1}^P (\tilde{\theta}_{i,j} - \theta_{i,j})^2} \quad (12)$$

Trong đó, $\tilde{\theta}$ và θ lần lượt là góc ước lượng và góc đúng; P là số nguồn tín hiệu cho một lần ước lượng, với $P \leq 4$; Q là số lượng kiểm tra.

Trong nghiên cứu này, các đánh giá sẽ được thực hiện lần lượt với 4 trường hợp khác nhau của mảng ULA và đường truyền: a) Thiết đặt chính xác vị trí các phần tử ăng ten, không có sai lệch pha và biên độ do đường truyền; b) Có sai lệch thiết đặt vị trí các phần tử ăng ten, không có sai lệch pha và biên độ do đường truyền; c) Thiết đặt chính xác vị trí các phần tử ăng ten, có sai lệch pha và biên độ do đường truyền; d) Có sai lệch thiết đặt vị trí các phần tử ăng ten và sai lệch pha và biên độ do đường truyền. Dưới đây là các kết quả nhận được và bình luận.

a. Thiết đặt chính xác vị trí các phần tử, không có sai lệch pha và biên độ do đường truyền

Kết quả thể hiện trên hình 5a cho thấy, mô hình DOA-CNN và ba phương pháp CBF, Capon và MUSIC cho độ chính xác ước lượng DOA tương đương nhau và kém hơn hai phương pháp Root-MUSIC và ESPRIT. Điều này là do DOA-CNN và các phương pháp CBF, Capon và MUSIC ước lượng DOA dựa trên phổ góc với độ phân giải là 1° , nếu góc đến là số lẻ thập phân (ví dụ $10,45^\circ$) thì các phương pháp này sẽ đưa ra góc ước lượng là một số nguyên (ví dụ 10°). Đây là một hạn chế của các phương pháp ước lượng dựa trên phổ góc. Trong khi đó, Root-MUSIC và ESPRIT có khả năng cho ra kết quả tường minh thông qua tính toán trực tiếp nên có thể đạt được độ phân giải cao hơn, do đó, hai phương pháp này cho độ chính xác ước lượng tốt hơn nếu được đảm bảo tỉ số SNR cao.

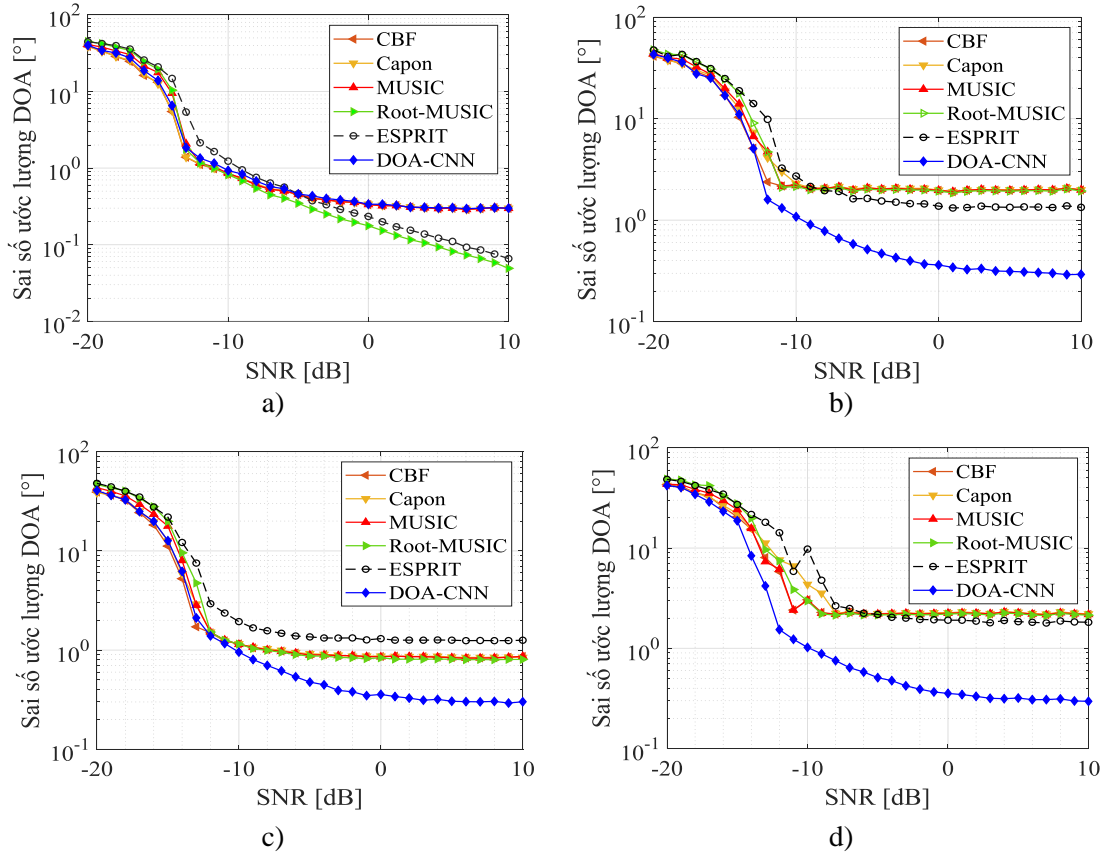
b. Có sai lệch thiết đặt vị trí các phần tử, không có sai lệch pha và biên độ do đường truyền

Sai lệch vị trí các phần tử gây ra sai số hiệu pha giữa các tín hiệu nhận được tại các phần tử của mảng. Nếu áp dụng các phương pháp ước lượng dựa theo cấu trúc giả định của mảng (không có sai lệch vị trí các phần tử) thì độ chính xác ước lượng sẽ giảm. Với mô hình DOA-CNN đề xuất, do được học tập từ dữ liệu thu bởi các phần tử của mảng để ước lượng DOA nên nó có thể tự động hiệu chỉnh các sai số, duy trì độ chính xác ước lượng. Trong mô phỏng này, vị trí các phần tử được giả định ngẫu nhiên là $[0; 1,37; 2,35; 2,83, 4,15; 4,85; 6,23; 6,76] \lambda/2$, các vị trí này sai lệch so với các vị trí đúng $[0; 1; 2; 3; 4; 5; 6; 7] \lambda/2$. Kết quả so sánh sai số ước lượng DOA giữa các phương pháp trong trường hợp này được thể hiện trên hình 5b. Có thể thấy, mô hình DOA-CNN mang lại hiệu quả rõ rệt so với các phương pháp khác. Cụ thể: DOA-CNN cho độ chính xác với sai số nhỏ hơn 1° khi SNR > 10 dB và tiệm cận $0,3^\circ$ khi SNR > 0 dB. Trong khi đó, hai phương pháp Root-MUSIC và ESPRIT cho sai số ước lượng lớn, luôn lớn hơn 1° với mọi tỉ số SNR. Điều này là do các phương pháp truyền thống không thể thay đổi tham số thuật toán để thích ứng với vị trí thực tế của các phần tử mảng ăng ten.

c. Thiết đặt chính xác vị trí các phần tử ăng ten, có sai lệch pha và biên độ do đường truyền

Đường truyền tín hiệu từ các phần tử của mảng ăng ten đến đầu vào của bộ biến đổi ADC cũng gây ra những sai lệch về biên độ và pha, đặc biệt khi tín hiệu phải đi qua nhiều mạch trung gian như bộ lọc, khuếch đại, bộ trộn tần số,... Mặc dù có thể hiệu chỉnh, bù sai lệch pha và biên độ thông qua đo đạc (hiệu chỉnh tuần tự từng phần tử và đường truyền), nhưng phương pháp này tiêu tốn thời gian và khó kiểm soát. Giải pháp sử dụng mô hình DOA-CNN đề xuất có thể tự động hiệu chỉnh những sai số về biên độ và pha thông qua việc học tập từ dữ liệu thu. Trong mô phỏng này, các sai lệch về biên độ và pha được giả định ngẫu nhiên trong giới hạn 10%. Cụ thể, suy hao biên độ $\alpha = [1; 1,21; 1,09; 1,1; 1,12; 1,22; 1,2; 1,05]$, sai số pha $\Delta\phi = [-5^\circ; 16^\circ; 17^\circ; 8^\circ; 4^\circ; 1^\circ; 35^\circ; 13^\circ]$. Kết quả so sánh sai số ước lượng giữa các phương pháp được thể hiện trên hình 5c. Có thể nhận thấy, mô hình DOA-CNN tiếp tục cho độ chính xác ước lượng tốt hơn các phương pháp còn lại vì nó có khả năng tự động hiệu chỉnh các sai lệch về pha cũng như biên độ

gây ra bởi đường truyền. Phương pháp ESPRIT cho độ chính xác ước lượng kém nhất. Capon, CBF, MUSIC và Root-MUSIC cho độ chính xác tốt hơn ESPRIT nhưng sai số ước lượng cũng không nhỏ hơn $0,9^\circ$. Mô hình DOA-CNN cho độ chính xác tốt nhất với sai số ước lượng nhỏ hơn $0,9^\circ$ khi SNR > -10 dB và đạt nhỏ hơn $0,3^\circ$ độ khi SNR > 2 dB.



Hình 5. So sánh độ chính xác ước lượng DOA của mô hình DOA-CNN với các phương pháp khác trong các trường hợp: a) Thiết đặt chính xác vị trí các phần tử, không có sai lệch pha và biên độ đường truyền; b) Có sai lệch thiết đặt vị trí các phần tử, không có sai lệch pha và biên độ đường truyền; c) Thiết đặt chính xác vị trí các phần tử, có sai lệch pha và biên độ đường truyền; d) Có sai lệch thiết đặt vị trí các phần tử và sai lệch pha và biên độ đường truyền.

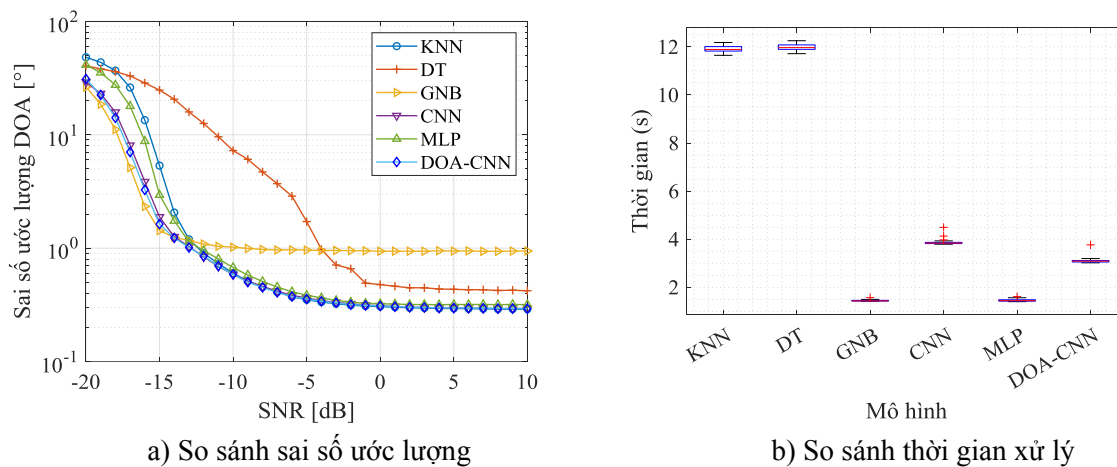
d. Có sai lệch thiết đặt vị trí các phần tử ăng ten và sai lệch pha và biên độ do đường truyền

Trong trường hợp này, sai số tổng hợp từ sai lệch thiết đặt vị trí các phần tử ăng ten và sai lệch pha và biên độ do đường truyền được xem xét đánh giá. Các sai lệch được giả định giống như hai trường hợp b, c. Kết quả so sánh sai số ước lượng DOA của các phương pháp thể hiện trên hình 5d. Dễ dàng nhận ra sai số của mô hình DOA-CNN thấp hơn rất nhiều sai số của các phương pháp còn lại. Điều này tiếp tục khẳng định hiệu quả của mô hình DOA-CNN đề xuất trong việc khắc phục sai số hệ thống do lỗi thiết đặt mảng ăng ten và đường truyền gây ra.

Quan sát hầu hết các kịch bản mô phỏng, khi SNR ≤ -10 dB các tín hiệu hữu ích quá yếu so với nhiễu tạp; do đó, dưới sự tác động mạnh của nhiễu, tất cả các phương pháp đều cho sai số lớn. Chỉ khi SNR > -10 dB thì các phương pháp mới bắt đầu đạt được mức sai số ước lượng DOA ổn định. Đặc biệt, trong các trường hợp có sai số mảng ăng ten, giải pháp sử dụng DOA-CNN cho thấy rõ hiệu quả so với các phương pháp truyền thống. DOA-CNN có thể đạt và duy trì các mức sai số ước lượng nhỏ khi SNR > -10 dB nhờ khả năng học tập từ dữ liệu để tự động điều chỉnh các sai số hệ thống.

e. So sánh mô hình DOA-CNN với các mô hình học máy, học sâu khác

Bên cạnh việc so sánh với các phương pháp truyền thống như trên, mô hình DOA-CNN còn được xem xét so sánh với một số mô hình học máy như KNN, DT, GNB, MLP và CNN. Phép so sánh được thực hiện trong trường hợp sai số tổng hợp (có sai lệch vị trí các phần tử ăng ten và sai lệch pha và biên độ do đường truyền), kết quả được thể hiện trên hình 6. Có thể thấy, các phương pháp học máy và học sâu đều có khả năng thích ứng tốt với các sai số hệ thống, từ đó cho kết quả ước lượng tương đối ổn định. Cụ thể, các mô hình DOA-CNN, CNN và KNN cho khả năng ước lượng DOA tương đương nhau và tốt hơn đáng kể so với thuật toán DT và GNB (hình 6a). Hình 6b cho thấy, mô hình DOA-CNN có tốc độ xử lý xấp xỉ 3 s cho 20000 lần ước lượng, nhanh hơn so với CNN (≈ 4 s) và KNN (≈ 12 s). Thuật toán GNB và MLP cho tốc độ xử lý nhanh nhất ($\approx 1,5$ s) nhưng độ chính xác thấp hơn DOA-CNN, đặc biệt là GNB cho sai số khoảng 1° khi SNR > -10 dB.



Hình 6. So sánh sai số ước lượng và thời gian xử lý của mô hình DOA-CNN với các mô hình học máy, học sâu khác.

5. KẾT LUẬN

Bài báo đã trình bày kết quả nghiên cứu, đề xuất mô hình DOA-CNN tự động hiệu chỉnh sai số hệ thống, nâng cao độ chính xác ước lượng hướng đến của tín hiệu vô tuyến cho mảng ăng ten ULA. Mô hình DOA-CNN được thiết kế với cấu trúc phép nhân tích lũy của nhiều tầng trích chọn nhằm tạo ra tính phân biệt dữ liệu trong quá trình huấn luyện và thực thi. Độ chính xác ước lượng của mô hình được đánh giá, so sánh với các phương pháp truyền thống như: CBF, Capon, MUSIC, Root-MUSIC, ESPRIT; và các mô hình học máy học sâu gồm KNN, DT, GNB, CNN và MLP. Kết quả cho thấy: trong trường hợp lý tưởng, mảng ăng ten có vị trí các phần tử là chính xác tuyệt đối và không có sai lệch biên độ, pha do đường truyền thì phương pháp Root-MUSIC và ESPRIT cho độ chính xác ước lượng hướng đến cao hơn mô hình đề xuất và các phương pháp CBF, Capon, MUSIC. Tuy nhiên, trong thực tế với các sai số hệ thống do sai lệch thiết đặt vị trí các phần tử của mảng ăng ten và sai lệch biên độ, pha do đường truyền luôn tồn tại, mô hình DOA-CNN đề xuất đã chứng minh khả năng duy trì độ chính xác vượt trội so với các phương pháp còn lại. Khi so sánh với các mô hình học máy khác, mô hình DOA-CNN cũng cho thấy lợi thế về độ chính xác ước lượng DOA và thời gian xử lý. Hướng nghiên cứu tiếp theo: (1) xem xét bài toán ước lượng nhiều nguồn tín hiệu tương quan và không tương quan; (2) thực nghiệm đo đạc, kiểm chứng và đưa mô hình vào ứng dụng trong thực tiễn.

Lời cảm ơn: Bài báo này được tài trợ trong khuôn khổ đề tài Nghị định thư với Cộng hòa Séc mã số NĐT/CZ/22/12.

TÀI LIỆU THAM KHẢO

- [1]. R. Wiley, "ELINT: The Interception and Analysis of Radar Signals", Artech House, (2006).
- [2]. S. Chandran, "Advances in Direction-of-Arrival Estimation", Artech House, (2006).
- [3]. R. Schmidt, "Multiple emitter location and signal parameter estimation," in IEEE Transactions on Antennas and Propagation, vol. 34, no. 3, pp. 276-280, (1986).
- [4]. Y. -M. Ning, S. Ma, F. -Y. Meng and Q. Wu, "DOA Estimation Based on ESPRIT Algorithm Method for Frequency Scanning LWA," in IEEE Communications Letters, vol. 24, no. 7, pp. 1441-1445, (2020).
- [5]. B. D. Rao and K. V. S. Hari, "Performance analysis of Root-MUSIC," in IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 37, no. 12, pp. 1939-1949, (1989).
- [6]. V.-S. Doan and D.-S. Kim, "DOA estimation of multiple non-coherent and coherent signals using element transposition of covariance matrix," ICT Express, vol. 6, no. 2, pp. 67-75, (2020).
- [7]. M. Zhang and Z.-D. Zhu, "DOA estimation with sensor gain, phase and position perturbations," Proceedings of the IEEE 1993 National Aerospace and Electronics Conference-NAECON 1993, Dayton, OH, USA, pp. 67-69 vol.1, (1993).
- [8]. L. Wu, Z.-M. Liu, and Z.-T. Huang, "Deep convolution network for direction of arrival estimation with sparse prior", IEEE Signal Processing Letters, vol. 26, no. 11, pp. 1688-1692, (2019).
- [9]. S. Ge, K. Li, and S. N. B. M. Rum, "Deep Learning Approach in DOA Estimation: A Systematic Literature Review," Mobile Information Systems, vol. 2021, pp. 1-14, (2021).
- [10]. V. A. Devi and M. Naved, "Dive in Deep Learning," in Machine Learning in Signal Processing, Chapman and Hall/CRC, pp. 97-126, (2021).
- [11]. A. R. Carballeira, F. A. P. de Figueiredo, and J. M. C. Brito, "Simultaneous Estimation of Azimuth and Elevation Angles Using a Decision Tree-Based Method," (2023).
- [12]. L. Jian, X. Wang, J. Shi, and X. Lan, "Robust Sparse Bayesian Learning Scheme for DOA Estimation with Non-Circular Sources," Mathematics, vol. 10, no. 6, p. 923, (2022).
- [13]. Y. Liu, H. Chen, and B. Wang, "DOA Estimation of Underwater Acoustic Signals Based on PCA-kNN Algorithm," in 2020 International Conference on Computer Information and Big Data Applications (CIBDA), (2020).
- [14]. F. Zhao, G. Hu, C. Zhan, and Y. Zhang, "DOA Estimation Method Based on Improved Deep Convolutional Neural Network," Sensors, vol. 22, no. 4, p. 1305, (2022).

ABSTRACT

DOA-CNN: An automatic system error calibration model for enhancing accuracy in direction of arrival estimation of radio frequency signals

This paper presents a research proposal on a deep convolutional neural network for the problem of direction of arrival estimation of radio frequency signals (called DOA-CNN). The DOA-CNN model is designed with multiplication layers to enhance strong features of the data through convolutional stacks enabling the DOA classification accuracy. The evaluation considers several factors affecting the accuracy of DOA estimation for uniform linear array (ULA), including antenna element position errors, and amplitude and phase errors caused by transmission path deviations in the receiver. The analysis and comparison of DOA-CNN with CBF, Capon, MUSIC, Root-MUSIC, and ESPRIT methods and other machine learning methods show that considering the ideal configuration of the ULA array and the receiver, the Root-MUSIC and ESPRIT methods achieve the best accuracy since they can directly compute the DOA, while the other methods estimate the DOA via angular spectrum, leading to accuracy dependent on the spectral resolution. However, considering ULA errors and transmission path deviations in the receiver, the proposed DOA-CNN model outperforms in terms of accuracy compared to traditional methods and processes faster than some other machine learning models.

Keywords: DOA estimation; Convolutional neural network; Position errors; Amplitude and phase errors.