

NHẬN DẠNG TÍN HIỆU RA ĐA SỬ DỤNG KỸ THUẬT MẠNG NƠ-RON HỌC SÂU

LPI RADAR SIGNAL RECOGNITION USING DEEP-LEARNING NEURAL NETWORK

Nguyễn Văn Linh, Đoàn Văn Sáng*, Trần Công Tráng, Trần Văn Cường

Học viện Hải quân¹

*Tác giả liên hệ: doansang.g1@gmail.com

(Nhận bài: 01/7/2021; Chấp nhận đăng: 07/9/2021)

Tóm tắt - Hiện nay, các thiết bị ra đa sử dụng kỹ thuật điều chế tín hiệu phức tạp nhằm giảm xác suất bị thu chặn (LPI). Trong khi đó, dạng điều chế tín hiệu ra đa là một trong những thông tin quan trọng trong trinh sát điện tử, cho phép định danh nguồn phát xạ. Do đó, một mô hình mạng nơ-ron học sâu tích chập (CNN) sẽ được đề xuất trong bài báo này, nhằm nâng cao khả năng nhận dạng tín hiệu ra đa LPI. Cụ thể, mô hình CNN đề xuất được khảo sát với số kênh và kích thước bộ lọc khác nhau. Kết quả khảo sát cho thấy, tham số càng cao thì độ chính xác nhận dạng càng tăng, tuy nhiên, thời gian thực thi càng chậm. Vì vậy, cần lựa chọn mạng có độ lớn phù hợp để đạt được độ chính xác cần thiết với thời gian thực thi cho phép. Ngoài ra, các kỹ thuật tiền xử lý cũng đóng vai trò quan trọng trong việc nâng cao độ chính xác nhận dạng của mạng CNN. Do đó, hai kỹ thuật STFT và WVD được khảo sát. Kết quả cho thấy, mặc dù WVD cho độ chính xác nhận dạng cao hơn nhưng thời gian xử lý chậm hơn STFT.

Từ khóa - Trinh sát điện tử; mạng nơ-ron tích chập; tín hiệu ra đa; kỹ thuật tiền xử lý; điều chế tín hiệu

1. Giới thiệu

Trong tác chiến điện tử (EW: Electronic Warfare), các thiết bị ra đa chủ động thường sử dụng các giải pháp quản lý công suất, thời gian, dải phổ và không gian phát sóng để tránh hoặc làm giảm xác suất bị thu chặn. Ngoài ra, các kỹ thuật điều chế tín hiệu ra đa cũng giúp cho xác suất thu chặn giảm đi đáng kể. Những tín hiệu điều chế này được gọi là tín hiệu thu chặn xác suất thấp (LPI: Low Probability of Intercept). Chính vì vậy, các hệ thống trinh sát điện tử sẽ gặp phải nhiều thách thức trong việc nhận dạng các tín hiệu LPI. Thật vậy, trong hoạt động trinh sát điện tử, nhận dạng được các tín hiệu LPI của nguồn phát mà đối phương sử dụng sẽ góp phần định danh nguồn phát, từ đó có thể nắm bắt được tình hình về lực lượng và trang bị của đối phương. Bài toán nhận dạng tín hiệu LPI càng trở nên khó khăn khi hoạt động trong điều kiện có rất nhiều nguồn bức xạ điện từ (trạm phát thanh, truyền hình, trạm thu phát sóng viễn thông, thiết bị wifi, các loại ra đa, máy thông tin,...) khiến cho mật độ truyền sóng trong môi trường trở nên dày đặc.

Trước đây, để nhận biết được kiểu loại tín hiệu, các thiết bị trinh sát điện tử cần phải thu chặn, xử lý và trích chọn một cách hiệu quả các đặc trưng tín hiệu để tạo dữ liệu đầu vào cho các thuật toán nhận dạng chuyên dụng. Theo đó, các kỹ thuật phân tích phổ thời gian tần số được sử dụng rộng rãi để trích rút các đặc trưng mật độ công suất tín hiệu trong miền thời gian và tần số. Trong đó, đáng chú ý là các kỹ thuật xử lý biến đổi Fourier thời gian ngắn (STFT: Short

Abstract - Currently, radar equipment uses Low Probability Intercepted (LPI) signals. Meanwhile, modulated radar signal is one of the important information in electronic reconnaissance, allowing the identification of the emission source. In order to improve the recognition of LPI radar signals, convolutional deep learning neural networks (CNN) are proposed in this paper. Specifically, the proposed CNN model is surveyed with different channel numbers and filter sizes. Survey results show that, the higher the parameter, the higher the identification accuracy; however, the slower the execution time. Therefore, it is necessary to select a network of a suitable size to achieve the required accuracy with the allowed execution time. In addition, preprocessing techniques also play an important role in enhancing the identity accuracy of the CNN network. Therefore, two techniques of STFT and WVD were explored. Although WVD offers higher recognition accuracy, the results show that it has a slower processing time than STFT.

Key words - Electronic reconnaissance; Convolutional Neural Network (CNN); radar signal; pretreatment techniques; modulation of signals

Time Fourier Transform) và Xử lý phân bố Wigner-Ville (WVD: Wigner-Ville Distribution).

Gần đây, mạng nơ-ron tích chập (CNN) đã và đang được sử dụng rộng rãi trong nhiều lĩnh vực như: Nhận dạng hình ảnh; Nhận dạng giọng nói; Nhận dạng chữ viết tay,... Mạng CNN có tiềm năng to lớn để có thể được ứng dụng vào trong lĩnh vực quân sự mà cụ thể là nhận dạng tín hiệu ra đa LPI cho nhiệm vụ trinh sát điện tử.

Gần đây, trong luận án tiến sĩ của mình [1] Grancharova đã đề xuất mạng CNN (có ba lớp tích chập, mỗi lớp có 32 kênh với kích thước bộ lọc 3x3) để phân loại 9 dạng điều chế tín hiệu vô tuyến. Mặc dù, mạng CNN đó có thể cải thiện độ chính xác nhận dạng trung bình lên đến 86% cho tỉ số tín / tạp (SNR) trong khoảng (-10, +10) dB, cao hơn so với một mạng CNN khác (với độ chính xác 73%) được đề xuất trong [2], nhưng mô hình CNN trong [1] vẫn còn khá đơn giản và sử dụng dữ liệu đầu vào ở dạng biên độ-thời gian. Bên cạnh đó, mô hình LSTM (Long Short Term Memory) cũng được nghiên cứu, khảo sát và so sánh với mô hình CNN. Kết quả nghiên cứu trong [3] cho thấy khả năng phân loại của LSTM có tính ổn định hơn so với CNN khi thay đổi môi trường truyền sóng. Mặc dù vậy, khi xét trong một điều kiện nhất định, mô hình CNN có khả năng nhận dạng tín hiệu tốt hơn LSTM. Do đó, kết hợp giữa CNN và LSTM có thể tăng tính ổn định của thuật toán nhận dạng khi có tác động của nhiễu mà vẫn đảm bảo độ chính xác nhận dạng cao [4].

¹ Naval Academy (Nguyen Van Linh, Van-Sang Doan, Tran Cong Trang, Tran Van Cuong)

Một số nghiên cứu cho thấy, các kỹ thuật tiền xử lý để tạo dữ liệu đầu vào cho mô hình nhận dạng tác động rất lớn đến hiệu năng nhận dạng (độ chính xác và tốc độ xử lý). Trong đó, các kỹ thuật tiền xử lý điển hình như STFT [5] biến đổi Wavelet [6], WVD [7], và Choi-Williams Distribution [8] được ứng dụng rộng rãi. Tuy nhiên, mỗi kỹ thuật xử lý có những ưu và nhược điểm riêng và phù hợp với các mô hình nhận dạng khác nhau. Chính vì vậy, trong nghiên cứu này nhóm tác giả thực hiện khảo sát tính hiệu quả của mạng CNN khi sử dụng hai kỹ thuật tiền xử lý STFT và WVD. Cụ thể, nhóm tác giả đề xuất một mô hình mạng CNN mới có khả năng ứng dụng trong nhận dạng các tín hiệu LPI (bao gồm LFM, Barker và Rect). Đây là ba loại điều chế tín hiệu ra đa phổ biến nhất hiện nay. Mô hình CNN mà nhóm tác giả đề xuất được huấn luyện và đánh giá trên tập dữ liệu được giả lập bằng chương trình Matlab. Theo đó, hiệu quả nhận dạng tín hiệu LPI của mô hình đề xuất được đánh giá bằng cách thay đổi các tham số của mạng. Kết quả thử nghiệm cho thấy, mô hình CNN càng phức tạp thì độ chính xác nhận dạng càng cao, tuy nhiên thời gian xử lý lâu hơn. Ngoài ra, hiệu quả của hai kỹ thuật tiền xử lý (STFT và WVD) để tạo tập dữ liệu cũng được so sánh với nhau khi áp dụng mạng CNN của nhóm tác giả. Mặc dù, kỹ thuật WVD cho độ chính xác cao hơn STFT nhưng tốc độ xử lý chậm hơn.

2. Các tín hiệu ra đa LPI và kỹ thuật tiền xử lý

2.1. Tín hiệu LPI

Trong bài báo này, nhóm tác giả xem xét ba dạng sóng ra đa điển hình, bao gồm: tín hiệu điều chế nội xung đơn tần (Rect), tín hiệu điều tần tuyến tính (LFM) và tín hiệu điều pha mã Barker (Barker). Các tín hiệu này được giả lập để phục vụ cho việc huấn luyện mạng CNN và kiểm nghiệm. Như đã biết, một tín hiệu ra đa ở dạng số thực có thể được mô tả như sau:

$$s(t) = A(t) \cos(2\pi f(t)t + \varphi(t)) \quad (1)$$

Trong đó, $A(t)$ là giá trị biên độ - một hàm phụ thuộc thời gian mô tả đường bao tín hiệu điều chế, $f(t)$ là tần số sóng mang phụ thuộc thời gian và $\varphi(t)$ là pha của tín hiệu.

2.1.1. Tín hiệu Rect

Đây là tín hiệu phổ biến nhất được sử dụng trong phần lớn các hệ thống ra đa, đặc biệt là các hệ thống sử dụng máy phát dùng Magnetron. Tín hiệu $s(t)$ trở thành tín hiệu điều chế nội xung đơn tần khi:

$$A(t) = \begin{cases} A_0 & \text{khi } t \in [0, \tau_x] \\ 0 & \text{khi } t \notin [0, \tau_x] \end{cases}, \quad f(t) = f_0, \quad \varphi(t) = \varphi_0 \quad (2)$$

Trong đó, τ_x là độ rộng xung; A_0 là biên độ (không đổi); f_0 là tần số sóng mang (không đổi); φ_0 là pha ban đầu của tín hiệu (không đổi).

2.1.2. Tín hiệu LFM

Tín hiệu LFM được sử dụng trong các đài ra đa hiện đại nhằm nâng cao cự ly hoạt động và khả năng phân biệt mục tiêu, đồng thời giảm công suất phát bằng việc tăng độ rộng xung phát và nén xung tại máy thu. Tín hiệu LFM thường được tạo ra bởi các máy phát sử dụng công nghệ bán dẫn. Tín hiệu $s(t)$ trở thành tín hiệu điều tần tuyến tính nội xung khi:

$$s(t) = A(t) \cos(2\pi f_0 t + \frac{\eta}{2} t^2 + \varphi_0);$$

$$A(t) = \begin{cases} A_0, & t \in [0, \tau_x]; \\ 0, & t \notin [0, \tau_x], \end{cases} \quad (3)$$

$$\eta = \frac{2\pi \Delta f}{\Delta t} \quad (\text{rad} / s^2).$$

Trong đó, τ_x là độ rộng xung; A_0 là biên độ; f_0 là tần số gốc sóng mang; η là hệ số nén của tín hiệu điều tần; Δf là lượng dịch tần trong thời gian Δt ; φ_0 là pha ban đầu của tín hiệu.

2.1.3. Tín hiệu điều chế pha mã Barker

Tín hiệu điều pha (hay tín hiệu mã pha) là tổ hợp n xung vô tuyến liên tiếp nhau, có độ rộng τ_0 và tần số f_0 bằng nhau với một số hữu hạn p các giá trị pha ban đầu $\varphi_{0(q)}$ ($q = 0, 1, \dots, p-1$) tạo thành một xung tín hiệu có độ rộng $\tau_x = n \cdot \tau_0$ tương đối lớn. Giá trị pha ban đầu của các xung con thành phần có thể phân bố trong đoạn $[0, 2\pi]$ theo quy luật nhất định [9]. Trong trường hợp $p = 2$, hai giá trị pha ban đầu có thể là $\varphi_{0(0)} = 0$ và $\varphi_{0(1)} = \pi$. Tín hiệu $s(t)$ trở thành tín hiệu điều pha mã Barker khi:

$$A(t) = \begin{cases} A_0 & \text{khi } t \in [0, \tau_x] \\ 0 & \text{khi } t \notin [0, \tau_x] \end{cases},$$

$$f(t) = f_0, \quad (4)$$

$$\varphi(t) = \text{Barker}\{0; \pi\},$$

Trong đó, τ_x là độ rộng xung; A_0 là biên độ (không đổi); f_0 là tần số sóng mang (không đổi); $\varphi(t)$ là pha ban đầu của mỗi xung con tín hiệu (theo quy luật của mã Barker).

Bảng 1. Bộ mã Barker đã biết

n	Mã Barker
2	+1 -1 +1 +1
3	+1 +1 -1
4	+1 +1 -1 +1 +1 +1 +1 -1
5	+1 +1 +1 -1 +1
7	+1 +1 +1 -1 -1 +1 -1
11	+1 +1 +1 -1 -1 -1 +1 -1 -1 -1
13	+1 +1 +1 +1 +1 -1 -1 +1 +1 -1 +1 -1 +1

2.2. Tiền xử lý tín hiệu

Hình dạng phổ tần số-thời gian của tín hiệu là một dấu hiệu quan trọng để phân biệt các dạng điều chế tín hiệu LPI. Đã có nhiều phương pháp xử lý tín hiệu trong miền thời gian và miền tần số được nghiên cứu và phát triển. Trong nghiên cứu này, nhóm tác giả sử dụng hai kỹ thuật xử lý phổ biến, bao gồm STFT và WVD để tạo hình ảnh phổ tần số-thời gian. Tập dữ liệu được tạo ra bởi hai kỹ thuật này sẽ được sử dụng để huấn luyện và đánh giá các mạng CNN.

2.2.1. Xử lý STFT

Biến đổi STFT là một trong những phép phân tích tín hiệu mà kết quả thu được là ảnh phổ mật độ công suất của tín hiệu theo thời gian và tần số. Theo đó, trước khi biến đổi Fourier, tín hiệu phải được chia thành các đoạn ngắn hơn bằng cách nhân với một hàm cửa sổ (window function). Một số hàm cửa sổ điển hình được sử dụng như:

Rectangular window, Gauss window, Cassion. Như vậy, độ mượt của phổ công suất phụ thuộc vào độ rộng hàm cửa sổ và tỉ lệ chông lẫn giữa các cửa sổ khi trượt qua tín hiệu. Phép biến đổi STFT có thể được biểu diễn tổng quát bằng biểu thức toán học sau:

$$S(\tau, \omega) = \text{STFT}\{s(t)\} = \int_{-\infty}^{\infty} s(t)w(t-\tau)e^{-i\omega t} dt \quad (5)$$

Trong đó, $w(-)$ là hàm cửa sổ.

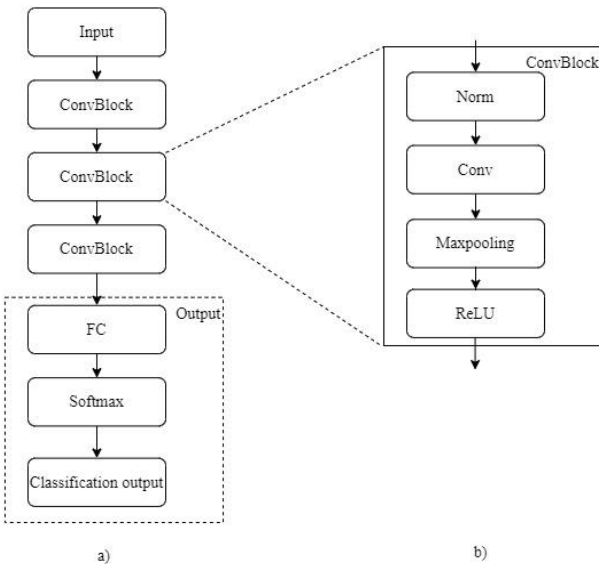
2.2.2. Xử lý phân bố Wigner-Ville

So với xử lý STFT, xử lý phân bố Wigner-Ville cung cấp độ phân giải về thời gian và tần số cao nhất có thể. Về mặt toán học, phép tính WVD được mô tả như sau:

$$W_s(t, \omega) = \int_{-\infty}^{\infty} s\left(t + \frac{\tau}{2}\right) s^*\left(t - \frac{\tau}{2}\right) e^{-j\tau\omega} d\tau. \quad (6)$$

3. Mô hình CNN cho bài toán nhận dạng tín hiệu LPI

CNN là một trong những mô hình của mạng nơ-ron học sâu trong các hệ thống thông minh, có khả năng tự động trích chọn đặc trưng của dữ liệu với độ chính xác cao. Nó đã được áp dụng vào đời sống thực tiễn với những ứng dụng điển hình như nhận dạng khuôn mặt, chuyển giọng nói thành văn bản, chẩn đoán hình ảnh y học, ... Trong lĩnh vực trình sát điện tử, mô hình CNN có tiềm năng ứng dụng vào nhận dạng tín hiệu ra đa LPI. Mạng nơ-ron học sâu mà nhóm tác giả đề xuất được trình bày trong Hình 1, gồm có: Lớp đầu vào, ba khối tự động trích xuất đặc tính (ConvBlock) và khối đầu ra.



Hình 1. Mô hình CNN đề xuất cho bài toán nhận dạng tín hiệu ra đa LPI. a) Mô hình tổng quát; b) mô hình của ConvBlock

Tham số lớp đầu vào của mạng CNN được chỉ định theo kích thước của ảnh phổ tín hiệu. Cụ thể trong nghiên cứu này, khi sử dụng dữ liệu phổ STFT thì kích thước của lớp đầu vào là 128×17 , và kích thước này sẽ là 200×200 khi sử dụng dữ liệu phổ WVD. Các đặc tính hữu ích để phân biệt các dạng tín hiệu LPI trong ảnh phổ đầu vào sẽ được tự động trích chọn khi đi qua các khối Convblock. Chi tiết các lớp trong Convblock được thể hiện trong Hình 1b), ở đó,

các lớp chuẩn hóa (Norm), tích chập (Conv), gộp (Maxpool) và hàm kích hoạt (ReLU) được kết nối liên tiếp với nhau.

Lớp Norm được dùng để chuẩn hóa dữ liệu đầu vào trước khi thực hiện phép tính tích chập tại lớp Conv. Như đã đề cập trong [10], lớp Norm giúp đẩy nhanh quá trình huấn luyện nhờ việc đưa các giá trị đặc tính về dạng chuẩn hóa. Giả thiết đầu vào của Norm là x_i thì giá trị đầu ra \hat{x}_i nhận được theo công thức sau:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (7)$$

Trong đó, $\mu_B = \frac{1}{B} \sum_{i=1}^B x_i$ và $\sigma_B^2 = \frac{1}{B} \sum_{i=1}^B (x_i - \mu_B)^2$ lần lượt là giá trị trung bình và phương sai của một nhóm gồm B ảnh phổ đầu vào, ϵ là hằng số đủ nhỏ để ổn định đầu ra trong trường hợp phương sai rất nhỏ. Trong nghiên cứu này nhóm tác giả chọn $\epsilon = 10^{-5}$.

Đặc tính phổ thời gian-tần số sau khi được chuẩn hóa sẽ được đưa vào lớp Conv có K kênh song song, mỗi kênh là một bộ lọc có kích thước $k \times k$. Lớp Conv có chức năng tăng cường các đặc tính hữu ích đồng thời làm suy yếu các đặc tính không cần thiết và nhiễu. Phép tính tích chập với tín hiệu đầu vào là véc-tơ a và véc-tơ trọng số w của bộ lọc được biểu diễn như sau:

$$y_n = \sum_{i=0}^{k-1} a_{n+i} w_i \quad (8)$$

Với n thỏa mãn $0 \leq n+i < N, \forall i = 0, 1, \dots, k-1$.

Theo sau lớp Conv là lớp Maxpool. Lớp này được sử dụng để chọn lọc các đặc tính mạnh, loại bỏ các đặc tính yếu, đồng thời làm giảm kích thước của dữ liệu cho lớp tiếp theo. Các lớp gộp (Pool) có thể có nhiều loại khác nhau: Max pooling lấy phần tử lớn nhất từ ma trận đối tượng; Average Pooling lấy tổng trung bình hoặc Sum Pooling lấy tổng tất cả các phần tử trong map. Trong mô hình CNN đề xuất, lớp Maxpool được sử dụng.

Hàm kích hoạt (Activation function) là một hàm phi tuyến được áp dụng vào đầu ra của các lớp Maxpooling, như trong Hình 1b. Nhiều nghiên cứu chỉ ra rằng, hàm ReLU (Rectified Linear Unit) cho tốc độ tính toán nhanh, đồng thời khắc phục được một số nhược điểm của các hàm khác như Sigmoid hoặc Tanh. Do đó, hàm ReLU được lựa chọn áp dụng trong mô hình mạng của nhóm tác giả. Theo đó hàm ReLU được mô tả như sau:

$$f(x) = \max(0, x). \quad (9)$$

Khối đầu ra (Output) bao gồm các lớp: FC (Fully Connected layer), Softmax, và Classification.

Các lớp được kết nối đầy đủ (FC) trong mạng nơ-ron là những lớp mà tất cả các đầu vào từ một lớp được kết nối với mọi đơn vị kích hoạt của lớp tiếp theo. Trong hầu hết các mô hình học máy phổ biến, một vài lớp cuối cùng là các lớp được kết nối đầy đủ sẽ biến đổi dữ liệu được trích xuất bởi các lớp trước đó để tạo thành đầu ra cuối cùng. Đây là lớp tiêu tốn nhiều thời gian thứ hai sau Lớp Convolution.

Hàm Softmax là một hàm biến một vectơ gồm K giá trị thực thành một vectơ gồm K giá trị thực có tổng bằng 1. Các giá trị đầu vào có thể tùy ý, hàm Softmax biến chúng thành các giá trị giữa 0 và 1, để chúng có thể được hiểu là xác suất. Nếu một trong các đầu vào nhỏ hoặc âm, thì Softmax sẽ biến nó thành một xác suất nhỏ và nếu một đầu vào lớn, thì nó sẽ biến nó thành một xác suất lớn, nhưng nó sẽ luôn nằm trong khoảng từ 0 đến 1. Softmax rất hữu ích vì nó chuyển đổi điểm số thành phân phối xác suất chuẩn hóa, có thể được hiển thị cho người dùng hoặc được sử dụng làm đầu vào cho các hệ thống khác. Vì lý do này, người ta thường thêm một hàm Softmax làm lớp cuối cùng của mạng nơ-ron.

Lớp Classification thực hiện ra quyết định dạng sóng tín hiệu ra đa LPI tương ứng với dữ liệu phổ đầu vào. Việc ra quyết định phụ thuộc vào xác suất của mỗi dạng sóng được tính bởi hàm Softmax trước đó. Giả sử p_i là xác suất của các phân lớp tín hiệu, khi đó kết quả dự đoán được xác định như sau:

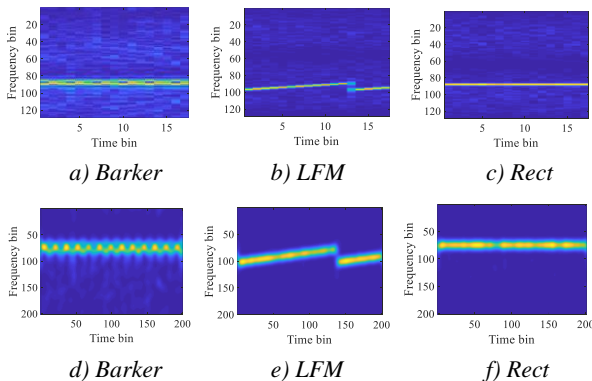
$$Waveform_{predicted} = \arg \{ \max(p_i) \} \quad (10)$$

4. Kết quả kiểm nghiệm và thảo luận

4.1. Tập dữ liệu và phần cứng huấn luyện

Dữ liệu cho huấn luyện mạng nơ-ron đóng vai trò quan trọng, quyết định độ chính xác nhận dạng tín hiệu điều chế. Nhóm tác giả mô phỏng 3 dạng tín hiệu ra đa LPI cơ bản là tín hiệu điều chế nội xung đơn tần; Tín hiệu điều tần tuyến tính nội xung; Và tín hiệu điều pha mã Barker bằng Matlab. Mỗi tín hiệu được thêm nhiễu tạp trắng dạng phân bố Gauss có tỉ số SNR ngẫu nhiên trong dải từ -20 dB đến 10 dB với bước cách là 1 dB.

Các tín hiệu tạo ra ở trên được đưa vào xử lý trong miền thời gian - tần số bởi 2 kỹ thuật tiền xử lý STFT và WVD. Đầu ra của các bộ tiền xử lý là các ảnh phổ thời gian - tần số của tín hiệu, như được thể hiện trong Hình 2. Trong đó, Hình 2a,b,c lần lượt là ảnh phổ của các tín hiệu Barker, LFM và Rect khi xử lý bằng STFT với SNR = 10 dB. Tương tự, Hình 2d,e,f là ảnh phổ khi xử lý bằng WVD.



Hình 2. Hình ảnh phổ tín hiệu với SNR = 10 dB sử dụng các kỹ thuật tiền xử lý STFT (a, b, c) và WVD (d, e, f)

Mỗi kỹ thuật tiền xử lý tạo ra 30 000 file ảnh phổ thời gian-tần số được gán cho 3 nhãn đầu ra là Rect; LFM và Barker. Mỗi ảnh có kích thước 128×17 và 200×200 lần lượt tương ứng với các kỹ thuật tiền xử lý STFT và WVD.

Tập kiểm tra là các ảnh phổ thời gian - tần số được tạo

ra độc lập so với tập huấn luyện với các giá trị SNR từ -20 dB đến 10 dB với bước cách 1 dB. Như vậy, tập kiểm tra có tổng số 186 000 ảnh phổ cho mỗi kỹ thuật tiền xử lý (2 000 ảnh phổ / kiểu điều chế / SNR).

Các tham số chính của tín hiệu mô phỏng được thể hiện trong Bảng 2.

Quá trình huấn luyện và kiểm tra hiệu năng của mạng CNN đề xuất được thực hiện trên GPU của máy tính có cấu hình như sau: CPU Intel Core 9300H CPU @ 2.40GHz, RAM 16 GB, GPU GTX 1660ti 6GB.

Bảng 2. Các tham số của tín hiệu mô phỏng

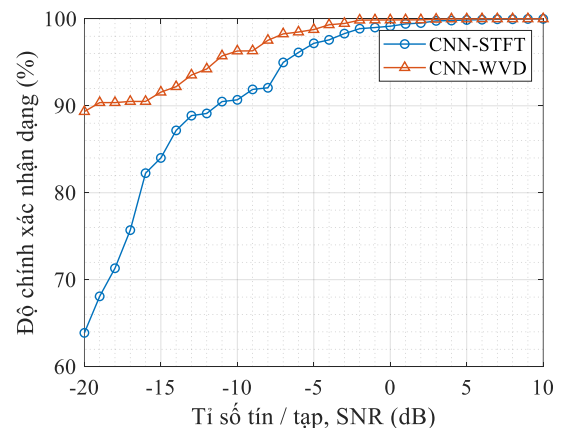
Tín hiệu	Các tham số chính	Ghi chú
Rect	$F_s = 100$ MHz $F_c = U\{F_s/6, F_s/5\}$ $\tau = U\{0.5, 2\}$ μ s	F_s : tần số lấy mẫu F_c : tần số sóng mang τ : Độ rộng xung điều chế
LFM	$F_s = 100$ MHz $F_c = U\{F_s/6, F_s/5\}$ $\tau = U\{0.5, 2\}$ μ s $B = U\{F_s/20, F_s/16\}$	$U\{a, b\}$: phân bố đều trong khoảng $[a, b]$ B : Độ rộng phổ
Barker	$F_s = 100$ MHz $F_c = U\{F_s/6, F_s/5\}$ $\tau_0 = U\{0.1, 0.5\}$ μ s $B = U\{F_s/20, F_s/16\}$ $N = \{3, 4, 5, 7, 11\}$	τ_0 : Độ rộng 1 bit trong mã Barker N : Số bit của mã Barker

4.2. Đánh giá hiệu năng nhận dạng của mạng CNN đề xuất

Ở phần này, mạng CNN đề xuất được đánh giá về độ chính xác cũng như thời gian xử lý khi áp dụng các kỹ thuật tiền xử lý STFT và WVD. Tiếp theo, độ chính xác nhận dạng của mạng CNN được khảo sát bằng việc thay đổi một số tham số như số lượng bộ lọc trong lớp Conv và kích thước bộ lọc.

4.2.1. So sánh độ chính xác nhận dạng khi thay đổi kỹ thuật tiền xử lý

Trong thử nghiệm này, mạng CNN với cấu Hình 3 khối ConvBlocks, mỗi khối có 8 kênh (= 8 bộ lọc) và kích thước bộ lọc của mỗi kênh được chỉ định là 3×3 , được lựa chọn để khảo sát thực hiện nhận dạng tín hiệu ra đa LPI khi kết hợp với hai kỹ thuật tiền xử lý STFT và WVD.



Hình 3. So sánh độ chính xác nhận dạng

Quá trình huấn luyện mạng CNN cho mỗi kỹ thuật tiền xử lý được thiết lập như nhau, cụ thể là: Tốc độ huấn luyện

khởi tạo là 0,01, kích thước gói huấn luyện cho mỗi vòng lặp là 64. Quá trình huấn luyện được thực hiện trong 20 epochs. Sau khi huấn luyện, hai mạng CNN (lần lượt ký hiệu là CNN-STFT và CNN-WVD) được kiểm chứng với tập dữ liệu kiểm tra.

Kết quả kiểm tra được thể hiện trong Hình 3 cho thấy, mạng CNN-WVD cho độ chính xác nhận dạng cao hơn đáng kể so với mạng CNN-STFT khi SNR < 0 dB. Khi SNR > 0 dB thì sự chênh lệch về độ chính xác nhận dạng của hai mạng là không nhiều, bởi vì lúc này cường độ của tín hiệu cao hơn nhiều khi biểu diễn trong miền thời gian. Khi phân tích phổ, cả hai kỹ thuật tiền xử lý STFT và WVD đều cho đặc trưng rõ nét về phổ mật độ công suất. Mặc dù vậy, mô hình CNN-WVD vẫn cho độ chính xác nhìn hơn một chút so với CNN-STFT vì độ phân giải ảnh phổ của WVD cao hơn so với STFT.

Bảng 3. So sánh hiệu năng của CNN-STFT với CNN-WVD

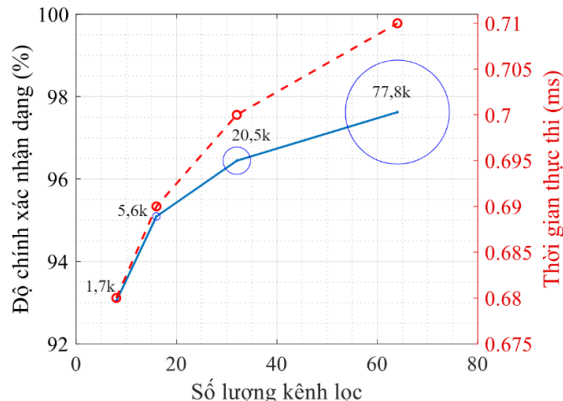
Mô hình mạng CNN	Độ chính xác nhận dạng trung bình (%)	Thời gian thực hiện (ms)
CNN-STFT	93,03	0,68
CNN-WVD	96,84	1,2

Tuy nhiên, khi so sánh thời gian thực hiện một lần dự đoán, mạng CNN-STFT cho thấy khả năng xử lý nhanh hơn nhiều so với mạng CNN-WVD, như số liệu thể hiện trong Bảng 3. Có sự khác biệt về thời gian xử lý là do sự khác nhau về kích thước đầu vào của mỗi mạng. Cụ thể, mạng CNN-STFT có kích thước dữ liệu đầu vào là 128×17 nhỏ hơn nhiều (khoảng 18 lần) so với kích thước đầu vào 200×200 của mạng CNN-WVD. Hơn nữa, tiền xử lý STFT cũng được thực hiện nhanh hơn so với WVD. Nhưng đổi lại, mạng CNN-WVD cho độ chính xác nhận dạng trung bình cao hơn so với mạng STFT khoảng 3,81%.

4.2.2. So sánh độ chính xác nhận dạng khi thay đổi số lượng kênh

Từ kết quả kiểm tra ở Phần 4.2.1, phổ đặc trưng xử lý bằng STFT được lựa chọn để làm dữ liệu đầu vào cho thử nghiệm này. Cụ thể, số lượng kênh của lớp Conv thay đổi lần lượt là $K = 8, 16, 32$ và 64; Kích thước bộ lọc của mỗi kênh là 3×3 . Quá trình huấn luyện và kiểm tra thực hiện tương tự như trong Phần 4.2.1. Hình 4 thể hiện sự phụ thuộc của độ chính xác nhận dạng tín hiệu ra đa LPI, thời gian thực thi và độ lớn (số lượng tham số) của mạng CNN vào số lượng kênh trong lớp Conv. Cụ thể, biểu đồ nét liền thể hiện độ chính xác nhận dạng; Biểu đồ nét đứt thể hiện thời gian thực thi; Và kích thước đường tròn biểu thị số lượng tham số của mạng tương ứng. Hình 4 cho thấy, khi tăng số lượng kênh của lớp tích chập thì số lượng tham số tính toán của các mạng CNN tăng rất nhiều từ $1,7 \times 10^3$ đến $77,8 \times 10^3$ khi số kênh thay đổi từ 8 đến 64, nhưng đổi lại độ chính xác nhận dạng tăng từ 93,03% lên 97,62% và giá phải trả là thời gian thực thi bị làm chậm từ 0,68 ms lên 0,71 ms. Ngoài ra, Hình 4 cũng cho thấy, độ chính xác nhận dạng tín hiệu ra đa được cải thiện đáng kể ($\approx 3,4\%$) khi thay đổi số kênh từ 8 lên 32. Trong khi đó, mạng CNN 64 kênh có độ chính xác cao hơn mạng CNN 32 kênh chỉ $\approx 1,2\%$ nhưng kích thước mạng tăng $57,3 \times 10^3$ tham số. Vì vậy, mạng CNN 32 kênh nên được lựa chọn để đạt được hiệu năng tốt nhất trong trường hợp nghiên cứu này khi xét đến

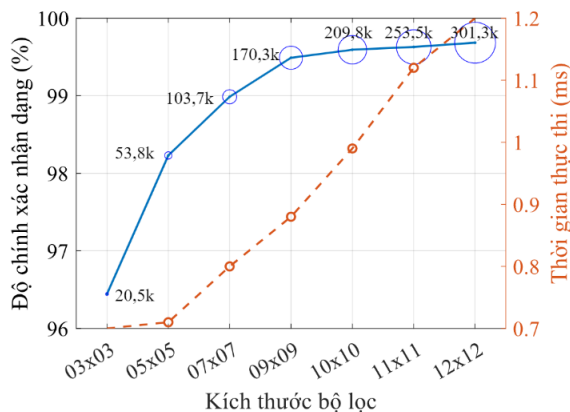
sự cân bằng giữa độ chính xác nhận dạng, tham số mạng và thời gian thực thi.



Hình 4. Độ chính xác nhận dạng và thời gian tính toán khi thay đổi số lượng kênh của mạng CNN

4.2.3. So sánh độ chính xác nhận dạng khi thay đổi kích thước bộ lọc

Từ kết quả kiểm tra ở Phần 4.2.2, số lượng kênh của lớp Conv được lựa chọn là 32 cho khảo sát này, trong khi đó, kích thước của bộ lọc được thay đổi lần lượt là $3 \times 3, 5 \times 5, 7 \times 7, 9 \times 9, 10 \times 10, 11 \times 11$, và 12×12 . Quá trình huấn luyện và kiểm tra thực hiện tương tự như các phần trên. Kết quả hiệu năng nhận dạng, kích thước mạng và thời gian thực thi được thể hiện trên Hình 5. Có thể thấy, các tham số của mạng CNN tăng từ $20,5 \times 10^3$ lên $301,3 \times 10^3$ khi thay đổi kích thước bộ lọc từ 3×3 đến 12×12 . Thời gian tính toán tăng gần như tuyến tính từ 0,7 ms lên 1,2 ms.



Hình 5. Độ chính xác nhận dạng và thời gian tính toán khi thay đổi kích thước bộ lọc

Độ chính xác nhận dạng được cải thiện đáng kể khi thay đổi kích thước bộ lọc từ 3×3 đến 9×9 (tăng $\approx 3\%$), nhưng tốc độ tăng chậm lại đối với các kích thước $10 \times 10, 11 \times 11$ và 12×12 . Có thể thấy, khi tăng kích thước bộ lọc từ 3×3 lên 5×5 thì độ lớn của mạng tăng thêm $33,3 \times 10^3$ tham số, kết quả là có thể cải thiện độ chính xác nhận dạng thêm 1,8% và thời gian thực thi cho một lần dự đoán gần như nhau ($\approx 0,7$ ms). Trong khi đó, khi tăng kích thước từ 9×9 lên 12×12 làm cho kích thước mạng CNN tăng 131×10^3 tham số, tuy nhiên độ chính xác nhận dạng chỉ tăng có 0,19% mà thời gian thực thi dự đoán lại chậm đi đáng kể ($\approx 1,2$ ms). Như vậy, mạng CNN với 32 bộ lọc kích thước 5×5 có thể được lựa chọn như cấu hình tối ưu nhất trong bài toán đang xem xét ở nghiên cứu này khi tính đến sự cân

bảng giữa độ chính xác nhận dạng, độ lớn của mạng và thời gian thực thi. Hình 6 thể hiện ma trận so sánh khi thực hiện phân loại 3 dạng tín hiệu radar LPI khi tỉ số SNR = -5 dB. Có thể thấy rằng, tín hiệu ra đa dạng xung điều chế một sóng mang (Rect) được mạng CNN dự đoán chính xác nhất do dạng tín hiệu này đơn giản nhất. Hai dạng tín hiệu mã Barker và LFM cho tỉ lệ nhận dạng đúng thấp hơn do tính phức tạp biến đổi pha và tần số của chúng. Trong đó, nhận dạng tín hiệu Barker chính xác hơn tín hiệu LFM. Hơn nữa, tỉ lệ nhận dạng nhầm của Barker và LFM chủ yếu rơi vào tín hiệu Rect, lần lượt là 1,6% và 2,8%. Điều này đúng với thực tế nhận dạng thủ công khi dựa vào kinh nghiệm của người vận hành hệ thống, rằng phân biệt tín hiệu Barker và LFM khó hơn so với tín hiệu Rect và thường nhầm sang tín hiệu Rect nếu nhiễu mạnh.

Ma trận so sánh

Nhận đúng	Barker	97.6%	0.8%	1.6%
	LFM		97.3%	2.8%
	Rect		0.1%	100.0%
		Barker	LFM	Rect

Nhận dự đoán

Hình 6. Ma trận so sánh khả năng nhận dạng đối với từng dạng sóng khi SNR = -5 dB

5. Kết luận

Qua các kết quả khảo sát, để đạt được độ chính xác cao mà số lượng các tham số tính toán và thời gian xử lý tối ưu, mô hình mạng CNN với 32 kênh, kích thước bộ lọc 5×5 được đề xuất để nhận dạng các tín hiệu ra đa LPI. Kỹ thuật tiền xử lý STFT được áp dụng để trích rút đặc trưng phổ tín hiệu đầu vào phục vụ nhiệm vụ nhận dạng trong điều kiện

truyền lan sóng điện từ tương đối phức tạp (SNR = -10dB) vì thời gian xử lý nhanh hơn mà vẫn bảo đảm được độ chính xác nhận dạng cần thiết. Như vậy, nghiên cứu này đã đưa ra được minh chứng về tính khả thi khi áp dụng mạng CNN cho bài toán nhận dạng tín hiệu radar LPI. Do đó, nhóm tác giả định hướng sẽ mở rộng nghiên cứu sử dụng mạng CNN bài toán nhận dạng tín hiệu ra đa LPI với nhiều kiểu loại đa dạng hơn trong các điều kiện thực tế khác nhau. Hơn nữa, một mô hình tối ưu sẽ được kiểm chứng với tín hiệu thực nghiệm và có thể đưa vào ứng dụng trong hệ thống thực tế.

TÀI LIỆU THAM KHẢO

- [1] M. Grancharova, "Representation Learning for Modulation Recognition of LPI Radar Signals Through Clustering", Dissertation, 2020.
- [2] L. J. Wong, W. C. Headley, S. Andrews, R. M. Gerdes and A. J. Michaels, "Clustering Learned CNN Features from Raw I/Q Data for Emitter Identification", *MILCOM 2018 - 2018 IEEE Military Communications Conference (MILCOM)*, Los Angeles, CA, USA, 2018, pp. 26-33.
- [3] Z. Ma, Z. Huang, A. Lin, G. Huang, "LPI Radar Waveform Recognition Based on Features from Multiple Images", *Sensors*, 2020 (2), pp. 526.
- [4] G. Ruan and Ya. Wang and Sh. L. Wang and Yu. Zheng and Q. Guo and S. N. Shulga, "Automatic recognition of radar signal types based on CNN-LSTM", *Telecommunications and Radio Engineering*, Vol. 79, No. 4, pp. 305-321, 2020.
- [5] A. Gupta, and A. A. Bazil Rai. "Feature Extraction of Intra-Pulse Modulated LPI Waveforms Using STFT", *2019 4th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT)*, 2019, pp. 742-746.
- [6] E. Świercz, "Automatic Classification of LFM Signals for Radar Emitter Recognition Using Wavelet Decomposition and LVQ Classifier", *Acta Physica Polonica A*. Vol. 119, pp. 488-494, 2011.
- [7] T. O. Gulum, P. E. Pace and R. Cristi, "Extraction of polyphase radar modulation parameters using a wigner-ville distribution - radon transform", *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, NV, USA, 2008, pp. 1505-1508
- [8] M. Zhang, L. Liu, and M. Diao, "LPI Radar Waveform Recognition Based on Time-Frequency Distribution", *Sensors*, Vol. 16, No. 10, pp. 1682, 2016.
- [9] Nguyễn Thanh Hùng, *Nguyên lý ra đa*, Học viện Hải quân, 2011.
- [10] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift", 2015-02-11, *arXiv:1502.03167*.