

Khảo sát thuật toán lan truyền ngược với thời gian phân rã thích nghi cho mạng nơron đột biến trong nhận dạng hệ số lực nâng thiết bị bay

Trương Đăng Khoa^{1*}, Nguyễn Văn Tuấn¹, Phạm Trung Dũng¹, Nguyễn Văn Hoa²

¹Viện Tên lửa và Kỹ thuật điều khiển, Học viện Kỹ thuật Quân sự, 236 Hoàng Quốc Việt, Bắc Từ Liêm, Hà Nội, Việt Nam.

²Khoa Vũ khí dưới nước, Học viện Hải Quân, 30 Trần Phú, Vĩnh Nguyên, Nha Trang, Khánh Hòa.

*Email: tuanguyen.mtak44@gmail.com

Nhận bài: 18/01/2024; Hoàn thiện: 10/3/2024; Chấp nhận đăng: 14/3/2024; Xuất bản: 01/4/2024.

DOI: <https://doi.org/10.54939/1859-1043.j.mst.CAPITI.2024.69-74>

TÓM TẮT

Bài báo này tiến hành khảo sát thuật toán lan truyền ngược với thời gian phân rã thích nghi cho mạng nơron đột biến. Từ kết quả khảo sát, tác giả đã xác định được khoảng giá trị phù hợp của thời gian phân rã và tốc độ học để nâng cao hiệu quả luyện mạng. Hiệu quả của thuật toán với giá trị các tham số được chọn sau khảo sát, cho thấy tốc độ hội tụ của mạng được cải thiện so với thuật toán ban đầu với các tham số khi chưa khảo sát thông qua bài toán nhận dạng tham số khí động thiết bị bay.

Từ khóa: Mạng nơron đột biến; Thời gian phân rã; Nhận dạng hệ thống.

1. GIỚI THIỆU

Mạng nơron nhân tạo (ANN) được xây dựng dựa trên những đặc điểm cơ bản của bộ não, nhưng chúng khác so với mạng nơron sinh học (BNN) về cơ chế hoạt động và cấu trúc. Do đó, hiệu suất của ANN vẫn thấp hơn rất nhiều so với BNN. Đây chính là động lực để thúc đẩy một thể hệ mạng mới có các cơ chế hoạt động và hiệu suất tiệm cận BNN, nó được gọi là mạng nơron đột biến (SNN) [1]. SNN là một bước đột phá trong việc khắc phục các hạn chế của ANN. Do sự tương đồng về cấu trúc và cơ chế tính toán với BNN, SNN có thể xử lý các tín hiệu rời rạc (các đột biến) như trong sinh học và có khả năng tương thích cao với mã thời gian [2].

Để phục vụ cho việc huấn luyện SNN, nhiều thuật toán được xây dựng dựa trên các ý tưởng khác nhau ra đời. Đầu tiên, các thuật toán lấy cảm hứng từ quy tắc lan truyền ngược sai số truyền thống như thuật toán Spikeprop [3] và các biến thể của nó. Tiếp theo là các thuật toán dựa trên ý tưởng về một cửa sổ học tập, được phát triển từ quy tắc Hebbian như ReSuMe [4] và các biến thể của nó. Ngoài ra, còn có một số thuật toán học không giám sát như STDP [5] và các biến thể của nó. Tổng quát, các thuật toán này đều sử dụng quy tắc cập nhật trọng số khớp nối trong quá trình huấn luyện mạng để tối thiểu hóa hàm sai số.

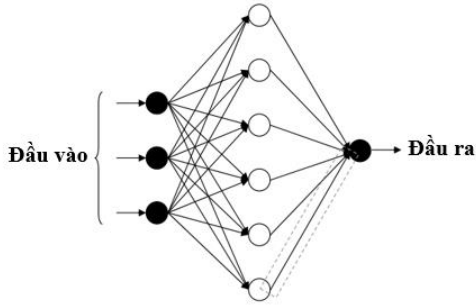
Thuật toán đề xuất [6] với giá trị phù hợp của các tham số xác định sau khảo sát đã cải thiện độ chính xác và tốc độ hội tụ của SNN trong nhận dạng hệ số lực nâng theo mô hình phi tuyến so với thuật toán có các tham số được chọn trước khảo sát.

Bài báo được tổ chức như sau: Phần 1 giới thiệu về SNN và một số thuật toán luyện mạng. Phần 2 thảo luận ngắn gọn về mô hình SNN và thuật toán đề xuất. Phần 3 giới thiệu mô hình chuyển động của máy bay trong kênh độ cao. Phần 4 khảo sát và xác định khoảng giá trị phù hợp cho thời gian phân rã và tốc độ học. Phần 5 đưa ra một số kết luận và hướng phát triển tiếp theo.

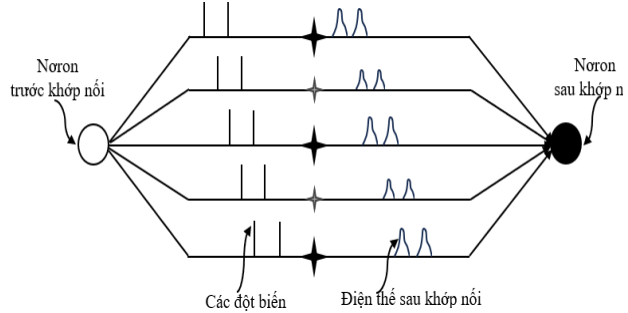
2. MÔ HÌNH MẠNG NƠRON ĐỘT BIẾN VÀ THUẬT TOÁN

2.1. Cấu trúc SNN

Trước khi đi vào nghiên cứu thuật toán, một kiến trúc mạng nơron đột biến [6] được xem xét như hình 1. Sự liên kết giữa các nơron được minh họa như hình 2.



Hình 1. Kiến trúc mạng SNN.



Hình 2. Liên kết giữa noron trước và sau khớp nối.

2.2. Mô hình phản ứng đột biến của noron SNN

Trong cấu trúc SNN, các noron được xây dựng với mô hình phản ứng đột biến (SRM) như trong [6].

$$u_j(t) = \sum_{i=1}^{F_{m+1}} \sum_{k=1}^S \sum_{g=1}^{G_i} w_{ij}^k \varepsilon(t - t_i^{(g)} - d^k) + \psi(t - t_j^{(f)}) \quad (1)$$

trong đó, các thành phần trong (1) được giải thích chi tiết trong [6].

2.3. Thuật toán cập nhật thích nghi đối với thời gian phân rã

Tương tự các thuật toán lan truyền ngược sai số truyền thống, thuật toán này cũng thực hiện tối thiểu hóa hàm sai số cho các đột biến đầu ra bằng việc cập nhật biến thời gian phân rã τ_d cho các hàm phản ứng đột biến được thực hiện giữa các lớp của mạng SNN như sau:

$$\Delta \tau_d = -\lambda \nabla E \quad (2)$$

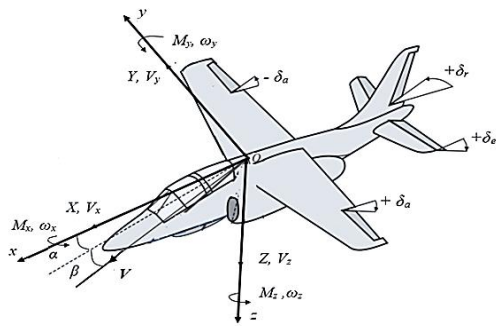
trong đó, λ là tốc độ học.

Việc tính toán lượng cập nhật $\Delta \tau_d$ được trình bày chi tiết trong [6].

3. MÔ HÌNH ĐỘNG HỌC MÁY BAY TRONG KÊNH ĐỘ CAO

3.1. Mô hình chuyển động của máy bay trong kênh độ cao

Chuyển động theo kênh độ cao của máy bay khi này được mô tả bằng hệ phương trình phi tuyến như trong [7]:



Hình 3. Hệ tọa độ liên kết của máy bay và các ký hiệu.

$$\begin{cases} \dot{V} = \frac{P}{m} \cos \alpha - \frac{q_V S}{m} C_D - g \sin(\vartheta - \alpha) \\ \dot{\alpha} = \omega_y - \frac{P}{mV} \sin \alpha - \frac{q_V S}{m} C_L \cos(\vartheta - \alpha) \\ \dot{\vartheta} = \omega_y \\ \dot{\omega}_y = \frac{q_V S b_A}{I_y} m_y \end{cases} \quad (3)$$

Trong hình 3 biểu diễn hệ tọa độ liên kết của máy bay Oxyz và các ký hiệu được giải thích chi tiết như trong [8].

Các hệ số lực nâng, hệ số lực cản và hệ số mômen trong (3) được xác định trong hệ tọa độ tốc độ, khi chuyển qua hệ tọa độ liên kết, các hệ số này được xác định như sau [7]:

$$\begin{cases} C_x = \frac{1}{q_v S} (m a_x - P); C_z = \frac{m a_z}{q_v S}; m_y = \frac{1}{q_v S b_A} I_y \dot{\omega}_y \\ C_L = C_x \sin \alpha - C_z \cos \alpha; C_D = -C_x \cos \alpha - C_z \sin \alpha \end{cases} \quad (4)$$

Trong thực tế, với chuyển động của máy bay trong giai đoạn cất cánh, hạ cánh hoặc chiếm lĩnh độ cao, thì mô hình hệ số lực nâng trong giai đoạn này được xác định [7]:

$$C_L = C_{L0} + C_{L1} \Delta \alpha + C_{L2} (\Delta \alpha)^2 + C_{L3} \frac{\omega_y b_A}{2V_0} + C_{L4} \left(\frac{\omega_y b_A}{2V_0} \right)^2 + C_{L5} \left(\Delta \alpha \frac{\omega_y b_A}{2V_0} \right) \dots \quad (5)$$

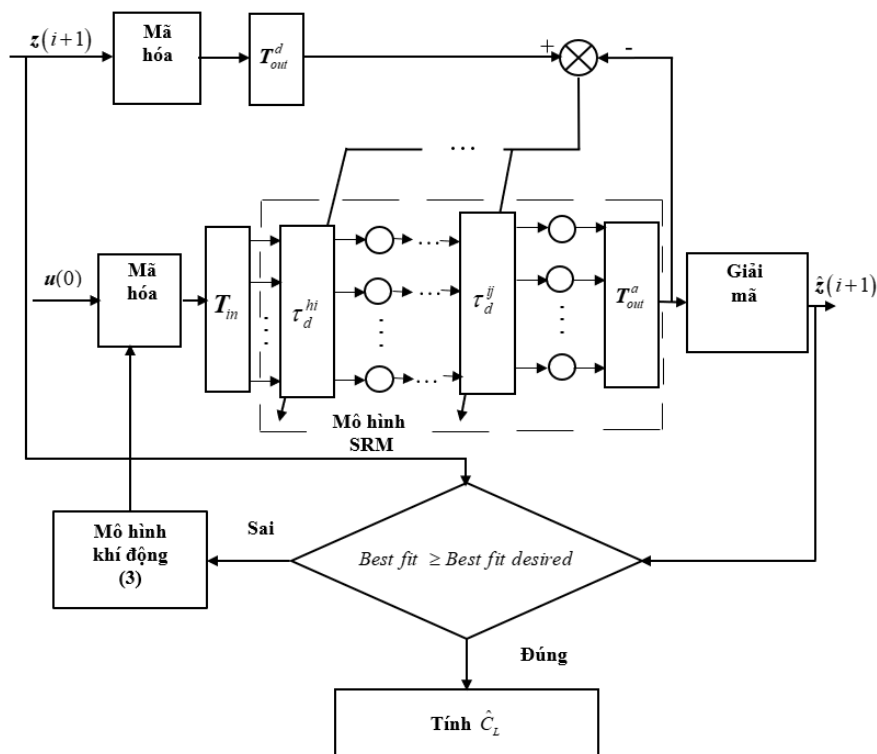
ở đây: C_{L0} - Hệ số lực nâng khi $\alpha = \alpha_0, V = V_0, \omega_y = 0$; $\Delta \alpha$ - Gia số góc tấn công so với α_0 ; α_0 - Góc tấn công cân bằng, đảm bảo cho máy bay bay với tốc độ V_0 không đổi và độ cao ổn định; $C_{L1}, C_{L2}, C_{L3}, C_{L4}, C_{L5}$ - Dẫn xuất khí động (aerodynamic derivatives) của lực nâng tương ứng với các thành phần tạo nên lực nâng.

3.2. Cấu trúc mô hình SNN kênh độ cao

Cấu trúc SNN kênh độ cao theo mô hình SRM được sử dụng để khảo sát thuật toán đề xuất cho mạng nơron đột biến trong nhận dạng hệ số lực nâng máy bay được thể hiện trong hình 4.

Véc tơ biến đầu vào quá trình nhận dạng $u(0) = [\alpha(0), \mathcal{G}(0), \omega_y(0), V(0), C_L(0), C_D(0)]^T$ tại thời điểm ban đầu $t = 0$ và véc tơ đầu ra là $z(i+1)$ đối với các tham số chuyển động của máy bay: $z(i+1) = [\alpha(i+1), \mathcal{G}(i+1), \omega_y(i+1), V(i+1)]^T$ được mã hóa thành chuỗi thời điểm đột biến đưa vào SNN.

Công thức mã hóa các biến đầu vào thành thời điểm đột biến và giải mã các chuỗi đột biến đầu ra thành tín hiệu tương tự được trình bày tương tự như trong [9].



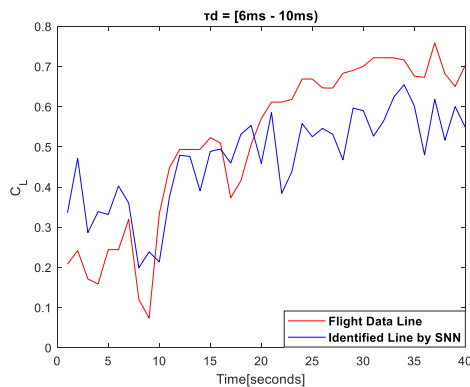
Hình 4. Cấu trúc SNN phục vụ nhận dạng hệ số lực nâng.

4. KHẢO SÁT VÀ XÁC ĐỊNH KHOẢNG GIÁ TRỊ PHÙ HỢP CHO THỜI GIAN PHÂN RÃ VÀ TỐC ĐỘ HỌC

4.1. Khảo sát ảnh hưởng của thời gian phân rã τ_d đến hiệu quả của thuật toán

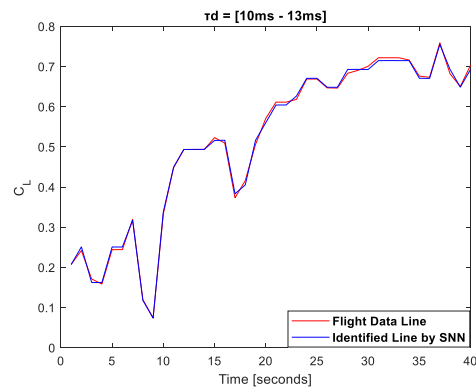
Tác giả sẽ sử dụng thuật toán đề xuất và tập dữ liệu thiết bị bay được ghi lại từ các chuyến bay thực tế để nhận dạng hệ số lực nâng của thiết bị bay trong giai đoạn cất cánh. SNN được xây dựng với cấu trúc 5-15-1 với các tham số được thiết lập như sau: $\tau_u = 5$ ms, $\tau_d \in [6$ ms ÷ 30 ms]; trọng số khớp nối cố định $w = 0.5$; tốc độ học $\lambda = 0.5$; số lượng epochs cho phép là 100.

Bộ dữ liệu các chuyến bay thực tế của máy bay SU30MK2 trong [8]. Kết quả khảo sát hiệu quả của thuật toán trong mục 2.3 với các khoảng thời gian phân rã khác nhau.



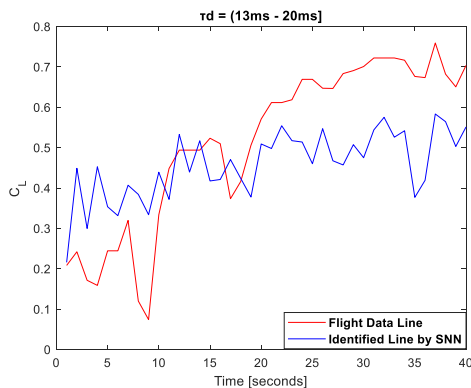
Hình 5. Trường hợp 1:

$\tau_u = 5$ ms; $\tau_d = 9,23$ ms.



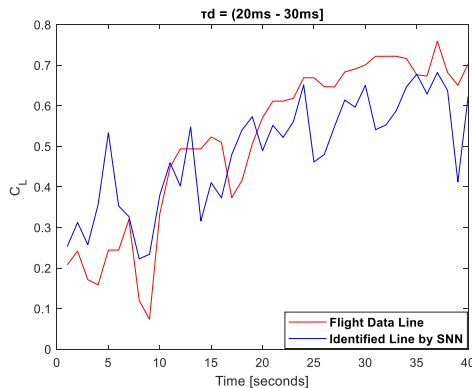
Hình 6. Trường hợp 2:

$\tau_u = 5$ ms; $\tau_d = 12,12$ ms.



Hình 7. Trường hợp 3:

$\tau_u = 5$ ms; $\tau_d = 17,26$ ms.



Hình 8. Trường hợp 4:

$\tau_u = 5$ ms; $\tau_d = 27,63$ ms.

- Trong khoảng 6 ms $\leq \tau_d < 10$ ms có thể thấy rằng, khoảng 10 s đầu tiên thì hệ số lực nâng nhận được theo nhận dạng có giá trị lớn hơn; khoảng 10 s tiếp theo, hệ số lực nâng nhận được tính toán theo các tham số khác trong bộ dữ liệu chuyến bay, tuy nhiên, lại có sự khác biệt lớn trong thời gian còn lại (hình 5 tương ứng $\tau_d = 9,23$ ms). Điều này cho thấy rằng, khoảng giá trị của thời gian phân rã 6 ms $\leq \tau_d < 10$ ms là chưa phù hợp.

- Trong khoảng 10 ms $\leq \tau_d \leq 13$ ms. Độ chính xác của hệ số lực nâng được nhận dạng bởi mạng SNN so với hệ số lực nâng tính qua dữ liệu nằm trong dải 95,7% ÷ 97,6%, đặc biệt với

$\tau_d = 12,12$ ms, thời gian phân rã cho độ chính xác đầu ra của mạng đạt 97.6% sau 13 epochs luyện mạng (hình 6).

- Trong khoảng $13 \text{ ms} < \tau_d \leq 20 \text{ ms}$ và $20 \text{ ms} < \tau_d \leq 30 \text{ ms}$, có thể thấy rằng, kết quả gần tương tự như trường hợp đầu tiên (hình 7 và 8 tương ứng $\tau_d = 17,26$ ms và $\tau_d = 27,63$ ms).

Có thể thấy rằng, với giá trị của thời gian phân rã nằm trong dải $\tau_d \in [10 \text{ ms} \div 13 \text{ ms}]$, thuật toán đề xuất có hiệu quả tốt nhất với bộ dữ liệu để nhận dạng hệ số lực nâng, khi mà chỉ sau khoảng 13 epochs mạng đã đạt được chỉ số Best fit mong muốn.

4.2. Khảo sát ảnh hưởng của tốc độ học và thời gian phân rã đến tốc độ hội tụ của mạng

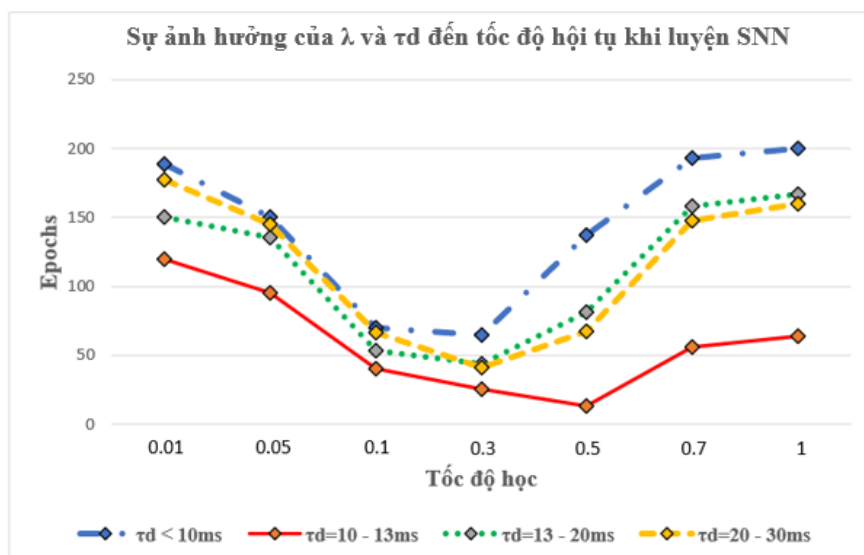
Trong phần này, thực hiện thay đổi tốc độ học λ trong (2) trong khoảng $\lambda = [0.01 \div 1]$ với các khoảng giá trị của thời gian phân rã $\tau_d \in [6 \text{ ms} \div 30 \text{ ms}]$; cấu trúc mạng SNN, các tham số khác của SNN tương tự như phần 4.1.

Từ hình 9 có thể thấy rằng, với các khoảng thời gian phân rã khác nhau khi tốc độ học λ nhỏ thì tốc độ hội tụ của mạng SNN với bộ dữ liệu TBB cũng chậm hơn. Nếu tiếp tục tăng tốc độ học thì tốc độ hội tụ của mạng cũng nhanh hơn. Tuy nhiên, khi tăng tốc độ học tới một giá trị nhất định thì tốc độ hội tụ của mạng cũng bắt đầu suy giảm, cụ thể như sau:

- Thời gian phân rã được thiết lập trong khoảng $10 \text{ ms} \leq \tau_d \leq 13 \text{ ms}$, mạng có tốc độ hội tụ tốt nhất với epochs lần lượt nhỏ nhất là 13 và lớn nhất là 87, tại $\lambda = 0.01$ và $\lambda = 0.5$. Trong khi đó, tốc độ hội tụ của mạng chậm nhất tương ứng số epochs nhỏ nhất là 75 và lớn nhất 100, tại $\lambda = 0.3$ và $\lambda = 1$ cho $6 \text{ ms} \leq \tau_d < 10 \text{ ms}$.

- Thời gian phân rã được thiết lập trong hai khoảng $13 \text{ ms} < \tau_d \leq 20 \text{ ms}$ và $20 \text{ ms} < \tau_d \leq 30 \text{ ms}$, tốc độ hội tụ của mạng có xu hướng tăng và giảm tương đồng nhau với các giá trị của tốc độ học, tốc độ hội tụ của mạng nhanh nhất khi epochs nhỏ nhất là 45 tại tốc độ học $\lambda = 0.3$ cho cả hai trường hợp.

- Đối với tất cả các trường hợp, mạng có tốc độ hội tụ nhanh khi tốc độ học $\lambda = [0.1 \div 0.5]$, ngược lại tốc độ hội tụ sẽ chậm lại thậm chí mạng không hội tụ.



Hình 9. Ảnh hưởng của tốc độ học và thời gian phân rã đến tốc độ hội tụ khi luyện mạng.

5. KẾT LUẬN

Bài báo đã thực hiện khảo sát và xác định được khoảng tốc độ học và khoảng thời gian phân rã tương đối phù hợp cho thuật toán đề xuất, nâng cao độ chính xác và tốc độ hội tụ của SNN trong bài toán nhận dạng hệ số lực nâng từ các bộ dữ liệu nhận được của máy bay trong giai đoạn cất cánh. Do quan hệ phi tuyến phức tạp giữa hai tham số này đối với sai số đầu ra trong mỗi epoch luyện mạng, việc chứng minh toán học để tường minh sự lựa chọn này chưa được thực hiện. Các nghiên cứu tiếp theo của nhóm tác giả sẽ tiếp tục nghiên cứu vấn đề này.

TÀI LIỆU THAM KHẢO

- [1]. Kashu Yamazaki, Viet Khoa Vo Ho, Darshan Bulsara, Ngan Le, “Spiking neural networks and their applications: A review,” *Brain Sci*, PMID: 35884670, 12(7): 863, (2022).
- [2]. Kasabov, N.K., “Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence,” Springer: Berlin/Heidelberg, Germany, (2019).
- [3]. Bohte, S.M., Kok, J.N., La Poutre, H, “Error-backpropagation in temporally encoded networks of spiking neurons,” *Neurocomputing* 48(1–4), 17–37, (2002).
- [4]. Ponulak F, Kasinski A, “Supervised learning in spiking neural networks with ReSuMe: sequence learning, classification, and spike shifting,” *Neural Computation*. 22: 467–510, (2010).
- [5]. H. Sun, X. Cui, Y. Guo and A. Ding, “Simplified Spike-timing Dependent Plasticity Learning Rule of Spiking Neural Networks for Unsupervised Clustering,” *The IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, Chongqing, China, pp. 26-30, (2019).
- [6]. Van Tuan Nguyen, Dang Khoa Truong, Trung Dung Pham, “A new supervised learning algorithm with the adaptive decay time for the spike neural network,” *The 12th International Conference on Control, Automation and Information Sciences, Vietnam*, (2023).
- [7]. E. A. M. Vladislav Klein, “Aircraft System Identification Theory and Practice,” American Institute of Aeronautics and Astronautics, Inc., (2006).
- [8]. N, Q, Vinh., Duc Thanh, N., Minh Dac, H., Dang Khoa, T., “Identify aerodynamic derivatives of the airplane attitude channel using a spiking neural network,” *International Journal of Aviation, Aeronautics, and Aerospace*, vol. 7, no. 3, (2020).
- [9]. Nguyen Van Tuan, Truong Dang Khoa, Pham Trung Dung, Dinh Huu Tai, “A spike trains encoding and decoding solution for the spiking neural networks,” *Journal of Military Science and Technology*, Vol 91, pp. 28-34, (2023).

ABSTRACT

Survey on error back propagation algorithm with the adaptive decay time for spike neural network in identifying the lift coefficient of an aircraft

This paper investigates the backpropagation algorithm with the adaptive decay time for the spike neural network. From the survey results, the author has determined the appropriate value range of decay time and learning rate to improve network training efficiency. The efficiency of the algorithm with the parameter values selected after the survey shows that the convergence speed of the network is improved compared to the original algorithm through the problem of identifying aircraft aerodynamic parameters.

Keywords: Spike neural network; The decay time; System identification.