

# ỨNG DỤNG BIẾN ĐỔI WAVELET CHO MỤC ĐÍCH NÂNG CAO CHẤT LƯỢNG PHÂN LOẠI DỮ LIỆU DẠNG CHUỖI THỜI GIAN

USING WAVELET TRANSFORM TO IMPROVE QUALITY CLASSIFICATION FOR TIME-SERIES DATA SEQUENCE

Đinh Thị Thùy Dương<sup>1,\*</sup>

## TÓM TẮT

Bài báo đề xuất giải pháp sử dụng phương pháp biến đổi sóng con (wavelet) để trích xuất các đặc trưng từ một chuỗi thời gian, kết quả đầu ra của giai đoạn tiền xử lý được đưa vào một mô hình mạng nơ-ron để phân loại và dự đoán xu hướng tương lai gần của chuỗi dữ liệu. Kết quả của đề xuất được thử nghiệm trên tập dữ liệu thực tế về hành vi hoạt động của con người cho thấy sự cải thiện về độ chính xác đạt 94%, cao hơn so với các kết quả khi sử dụng các phương pháp cũ.

**Từ khóa:** Chuỗi thời gian, biến đổi sóng con, trích xuất đặc trưng, học máy.

## ABSTRACT

This paper proposes a solution using wavelet transform to extract features from a time-series, the outputs of the pre-processing is input of a neural network in order to classify and predict near future trends of the data. The approach is based on the CWT and DWT of time-series. The result which is tested on real datasets HAR (Human Activity Recognition), shows the improvements in accuracy, reaching 94%. It is an improvement compared to previously reported results for previous systems.

**Keywords:** Time-series, wavelet transform, machine learning, deep learning.

<sup>1</sup>Viện Khoa học và Công nghệ Quân sự

\*Email: duongmta@gmail.com

Ngày nhận bài: 03/3/2022

Ngày nhận bài sửa sau phản biện: 18/4/2022

Ngày chấp nhận đăng: 25/4/2022

## 1. GIỚI THIỆU

Cùng với cuộc sống đang dần hiện đại, con người đang đối mặt với những bài toán phức tạp về khoa học dữ liệu như dự đoán thời tiết, phân tích thị trường chứng khoán... đi kèm với dữ liệu thay đổi theo thời gian. Ta gọi những dữ liệu đó là chuỗi thời gian (time series-TS) hoặc tín hiệu. Cho dù sử dụng để phân loại đối tượng hay thống kê hồi quy, việc trích xuất đặc trưng của đối tượng là một bước quan trọng sẽ ảnh hưởng đến độ chính xác của các dự đoán. Trong tập dữ liệu chuỗi thời gian, giá trị dự đoán ( $y$ ) là một hàm phụ thuộc thời gian có thể mô tả bất kỳ điều gì, như giá trị của tiền hoặc một cổ phiếu cụ thể theo thời gian ... Một tín hiệu là một định nghĩa tổng quát hơn, giá trị ( $y$ ) không chỉ là một hàm của thời gian, mà còn là một hàm phụ thuộc tọa độ không gian ( $y = y(x, y)$ ), khoảng cách ( $y = y(r)$ )... Về bản chất, hầu hết mọi thứ đều có thể được hiểu là

một tín hiệu miễn là nó mang thông tin bên trong chính nó. Ví dụ: Một hình ảnh có thể được coi là một tín hiệu chứa thông tin về độ sáng của ba màu (RGB) trên hai chiều không gian hoặc tín hiệu sonar cung cấp thông tin về tín hiệu dạng âm như một hàm của thời gian và không gian ba chiều. Ta có thể sử dụng biến đổi Fourier nhanh (Fast Fourier transform-FFT) để biến đổi tín hiệu từ miền thời gian sang miền tần số của nó [1]. Các đỉnh trong phổ tần số cho biết tần số xuất hiện nhiều nhất trong tín hiệu. Đỉnh càng lớn và sắc nét thì tần số của tín hiệu càng phổ biến. Sau đó, các giá trị về tần số và biên độ của các đỉnh trong phổ tần số được sử dụng làm đầu vào cho các mô hình Học máy như Random Forest, Gradient Boosting hoặc mạng nơ-ron để phân loại và đưa ra kết luận mang tính thống kê [2]. Một số mô hình kết hợp như vậy đã cho kết quả phân loại trên 80% với các bộ dữ liệu nổi tiếng được công khai và công nhận trên Internet như bộ dữ liệu về ảnh đối tượng dân sự ImageNet [3], âm thanh đường phố URBAN [4]... Tuy nhiên, phương pháp sử dụng biến đổi Fourier này sẽ hoạt động rất tốt khi phổ tần số là tĩnh. Nghĩa là, các tần số tồn tại trong tín hiệu không phụ thuộc vào thời gian; ví dụ, nếu một tín hiệu có tần số 10kHz thì tần số này phải có mặt ở bất kỳ vị trí nào trong tín hiệu. Tín hiệu càng không tĩnh thì kết quả sẽ bị sai lệch rất nhiều. Tuy vậy, hầu hết các tín hiệu chúng ta thấy trong cuộc sống thực đều có bản chất không cố định, do đó việc ứng dụng những mô hình cũ vào các bài toán thực tế gặp rất nhiều hạn chế. Dựa trên những phân tích đó, tác giả đề xuất một mô hình kết hợp biến đổi Wavelet (Wavelet transform-WT) với các thuật toán học máy để tăng tính chính xác, khắc phục các nhược điểm của các mô hình cũ, có khả năng ứng dụng trong thực tế đời sống con người. Phần tiếp theo sẽ trình bày cơ sở lý thuyết về mặt toán học và phân tích tính cải tiến, kết quả mô phỏng mô hình và kết luận.

## 2. CƠ SỞ LÝ THUYẾT

### 2.1. Biến đổi Fourier và biến đổi Wavelet

Biến đổi Fourier hoạt động bằng cách nhân một tín hiệu với một chuỗi các sóng hình sin với các tần số khác nhau, từ đó ta xác định được tần số nào có trong tín hiệu. Nếu tích số giữa tín hiệu của chúng ta và sóng hình sin của một tần số nhất định dẫn đến một biên độ lớn, điều này có nghĩa là có nhiều sự chồng chéo giữa hai tín hiệu và tín hiệu của chúng ta chứa tần số cụ thể này.

Biến đổi Fourier về bản chất có thể được biểu diễn theo công thức:

$$F\{s(t)\}(\omega) = \langle b(\omega, t), s(t) \rangle = \int_{-\infty}^{\infty} \exp(-i\omega\tau) s(\tau) d\tau \tag{1}$$

Với  $t$  là thời gian và  $\omega$  là tần số đơn. Tuy nhiên, biến đổi Fourier có độ phân giải cao trong miền tần số nhưng độ phân giải bằng không trong miền thời gian. Điều này có nghĩa là nó có thể cho chúng ta biết chính xác tần số nào xuất hiện trong một tín hiệu, nhưng không biết thời điểm nào mà các tần số này đã xuất hiện trong tập dữ liệu. Do đó, ta không thể biết tần số phát hiện được xảy ra trong thời gian nào của dữ liệu và cũng không thể phân biệt cụ thể vậy tín hiệu đó liệu có thuộc về nguồn âm cần tìm hay không.

Để cố gắng khắc phục vấn đề này, các nhà khoa học đã đưa ra phép biến đổi Fourier trong thời gian ngắn (Short time Fourier transform-STFT). Phép biến đổi STFT thêm chiều thời gian vào các tham số của hàm cơ sở bằng cách nhân hàm mũ phức vô hạn với một cửa sổ.

$$b(\omega, t_0)(t) := \omega(t - t_0) \exp(i\omega t) \tag{2}$$

trong đó,  $\omega(t)$  là các hàm cửa sổ và  $(\omega, t_0)$  là tọa độ thời gian-tần số của hàm cơ sở. Ta có công thức tổng quát cho STFT như sau:

$$S\{s(t)\}(\omega, t_0) = \int_{-\infty}^{\infty} \omega(\tau) * \exp(-i\omega\tau) s(\tau) d\tau \tag{3}$$

Công thức (3) cho ta thông tin về thời gian và tần số. Từ sự lựa chọn các cửa sổ ta thu được cân bằng giữa thời gian và tần số. Kết quả của phép biến đổi này cũng có thể được coi là một bộ lọc thông dải có biến đổi fourier là cửa sổ  $\omega(t)$  là đáp ứng tần số, nhưng được dịch chuyển sang tần số trung tâm  $\omega$ . Do đó, tất cả các bộ lọc có cùng băng thông. Trong cách tiếp cận này, tín hiệu gốc được chia thành nhiều phần có độ dài bằng nhau (có thể có hoặc không có chồng lên nhau) bằng cách sử dụng cửa sổ trượt trước khi áp dụng biến đổi Fourier. Tuy nhiên cách tiếp cận này gặp phải các giới hạn về lý thuyết của biến đổi Fourier được gọi là nguyên lý bất định [5].

Nếu kích thước của cửa sổ càng nhỏ, ta sẽ biết nhiều hơn về thời điểm tần số đã xảy ra trong tín hiệu, nhưng ít thông tin hơn về tần số của chính nó và ngược lại, làm cho kích thước của cửa sổ càng lớn thì ta sẽ biết nhiều hơn về giá trị tần số và ít hơn về thời gian.

Một cách tiếp cận tốt hơn để phân tích tín hiệu với phổ tần số không tĩnh là biến đổi Wavelet [6], WT có độ phân giải cao trong cả miền tần số và miền thời gian. WT không chỉ cho chúng ta biết tần số nào hiện diện trong một tín hiệu mà còn cho biết những tần số này đã xảy ra tại thời điểm nào. Điều này được thực hiện bằng cách làm việc với các cửa sổ có kích thước khác nhau.

**2.2. Biến đổi Wavelet liên tục và rời rạc**

WT sử dụng một loạt các hàm được gọi là wavelet, mỗi hàm có một tỷ lệ khác nhau. Về mặt toán học, Biến đổi Wavelet liên tục (Continuous Wavelet Transform-CWT) được mô tả bằng phương trình sau:

$$X_w(a, b) = \frac{1}{|a|^{1/2}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt \tag{4}$$

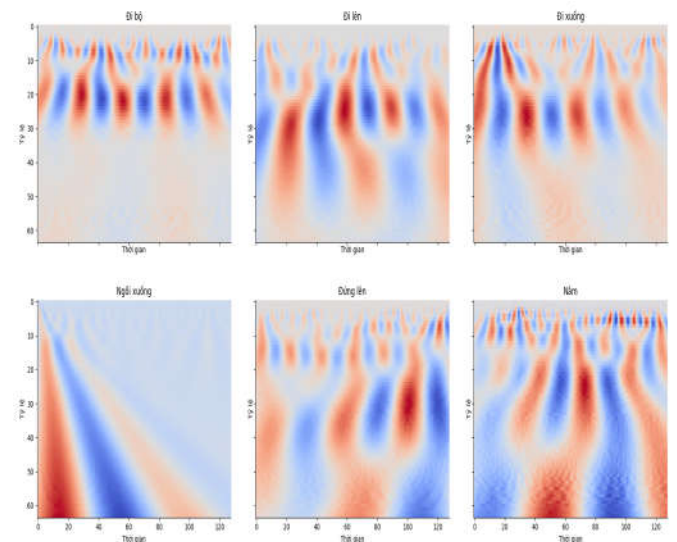
trong đó,  $\psi$  là wavelet mẹ liên tục được chia tỷ lệ theo hệ số  $a$  và dịch theo hệ số  $b$ . Các giá trị của các yếu tố tỷ lệ và dịch là liên tục, có nghĩa là có thể có vô số wavelet. Ta có thể chia tỷ lệ wavelet mẹ với hệ số 1,3, hoặc 1,31, và 1,311...

Với biến đổi Wavelet rời rạc (Discrete Wavelet Transform - DWT), sự khác biệt chính là DWT sử dụng các giá trị rời rạc cho tỷ lệ và hệ số dịch. Hệ số tỉ lệ tăng theo lũy thừa của hai, do đó  $a = 1, 2, 4, \dots$  và hệ số tịnh tiến làm tăng các giá trị nguyên ( $b = 1, 2, 3, \dots$ ). DWT chỉ rời rạc trong phạm vi tỷ lệ và miền dịch, không phải trong miền thời gian. Để có thể làm việc với các tín hiệu kỹ thuật số và tín hiệu rời rạc, chúng ta cũng cần phân biệt các biến đổi wavelet của mình trong miền thời gian. Các dạng biến đổi wavelet này được gọi là Biến đổi Wavelet theo thời gian rời rạc và Biến đổi Wavelet liên tục theo thời gian rời rạc.

**3. KẾT QUẢ MÔ PHỎNG VÀ THẢO LUẬN**

Để so sánh kết quả giữa các phương pháp đã phân tích ở phần 2, thực hiện kiểm tra lần lượt các phương pháp trên bộ dữ liệu thực tế về hành vi hoạt động của con người HAR [7]. HAR chứa các dữ liệu đo của những người khác nhau được thu từ một gia tốc kế và một con quay hồi chuyển, khi đang thực hiện tổng cộng 06 hoạt động: đi bộ chậm, đứng lên, ngồi xuống, nằm, lên cầu thang và xuống cầu thang. Có 7.532 mẫu trong tập huấn luyện và 2.947 mẫu trong tập kiểm thử với độ chèn 50% giữa mỗi mẫu.

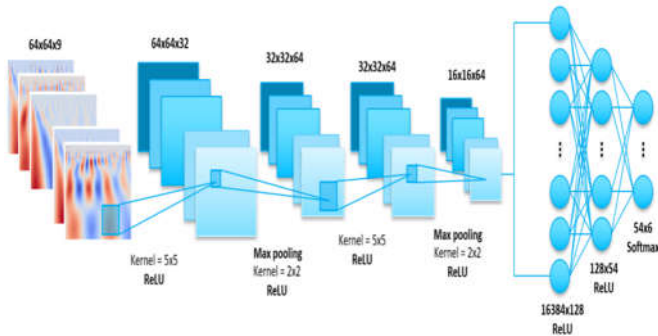
**3.1. Biến đổi Wavelet liên tục và mạng nơ-ron tích chập cho bộ dữ liệu HAR**



Hình 1. Khai triển Scalogram của một mẫu trong bộ dữ liệu HAR

WT tín hiệu một chiều dẫn đến biểu đồ tỷ lệ (scalogram) hai chiều chứa nhiều thông tin hơn FT. Nó không chỉ cho

chúng ta thấy chu kỳ của các dao động lớn nhất là bao nhiêu mà còn cho biết thời điểm có các dao động này và khi nào thì không. Một biểu đồ tỷ lệ không chỉ được dùng để phân tích về hành vi của một hệ thống, mà còn được sử dụng để phân biệt các loại tín hiệu khác nhau do một nguồn tạo ra với nhau. Biểu đồ scalogram của một mẫu trong bộ dữ liệu sau khi biến đổi wavelet có dạng như trong hình 1.

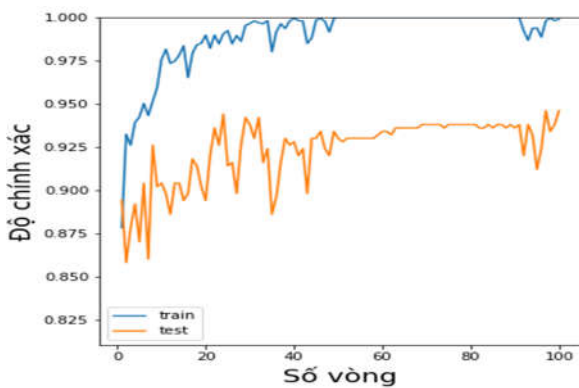


Hình 2. Kiến trúc mô hình CNN-LeNet

Ta có thể tự động hóa quá trình này bằng cách xây dựng một Mạng nơ ron phù hợp, để phát hiện các lớp mà mỗi biểu đồ tỷ lệ thuộc về và phân loại chúng. Bài báo sử dụng mạng Nơ ron tích chập kiến trúc LeNet với hai cải tiến như trong hình 2.

- Sử dụng max-pooling thay cho average-pooling, vì max-pooling cho hiệu suất tốt hơn khi trích xuất các tính năng thay đổi đột ngột và khi dữ liệu đầu vào có mật độ điểm ảnh cao (chín kênh đầu vào).
- Hàm kích hoạt ReLU thay cho hàm Tanh để khắc phục nhược điểm suy giảm gradient, tăng thời gian huấn luyện và đạt được hiệu suất cao hơn.

Hình 3 biểu diễn kết quả phân loại của mô hình đạt: 94,59%



Hình 3. Kết quả huấn luyện và kiểm thử trên tập dữ liệu HAR

### 3.2. Biến đổi Wavelet rời rạc và một số mô hình học máy

DWT được sử dụng để chia tín hiệu thành các dải tần phụ khác nhau. Nếu các loại tín hiệu khác nhau thể hiện các đặc tính tần số khác nhau, thì sự khác biệt về hành vi này phải được thể hiện ở một trong các dải tần con. Vì vậy, nếu ta tổng quát hóa các đặc trưng từ mỗi dải phụ và sử dụng làm đầu vào cho mô hình học máy và đào tạo nó bằng cách

sử dụng các tính năng này, bộ phân loại sẽ có thể phân biệt giữa các loại tín hiệu khác nhau.

Kết quả thử nghiệm DWT và một số phương pháp học máy tại bảng 1.

Bảng 1. Kết quả thử nghiệm HAR trên một số phương pháp học máy với DWT

	Độ chính xác trên tập huấn luyện	Độ chính xác trên tập thử nghiệm	Thời gian huấn luyện mô hình (s)
<b>Gradient Boosting</b>	1,000000	0,948422	828,721307
<b>Random Forest</b>	1,000000	0,922633	11,424119
<b>Logistic Regression</b>	0,959875	0,912793	7,944989

### 3.3. So sánh kết quả và phân tích

Sử dụng biến đổi FFT và các phép phân loại tương tự như mục 3.2 ta thu được kết quả phân loại trên bộ dữ liệu HAR đạt cao nhất 89% tại bảng 2.

Bảng 2. Kết quả thử nghiệm HAR trên một số phương pháp học máy với FFT

	Độ chính xác trên tập huấn luyện	Độ chính xác trên tập thử nghiệm	Thời gian huấn luyện mô hình (s)
<b>Gradient Boosting</b>	0,990751	0,899559	87,668457
<b>Random Forest</b>	0,998096	0,857482	0,524725
<b>Logistic Regression</b>	0,882617	0,811673	0,811673

Ta nhận thấy kết quả mô phỏng trong bảng 1 và 2 có độ chính xác khác nhau tương đối nhiều, điều này chứng minh cho phân tích lý thuyết là hợp lý vì hàm wavelet có đặc tính về mặt thời gian, do đó ta có thể nhận tín hiệu của mình với các hàm wavelet tại các thời điểm khác nhau của chuỗi dữ liệu. Bắt đầu từ phần đầu của tín hiệu và từ từ di chuyển wavelet về phía cuối của tín hiệu. Như trong phần 3.1 ta sử dụng phép tích chập của mạng CNN, ta có thể tăng tỷ lệ lớn hơn và lặp lại quá trình. Mỗi lần như vậy, mô hình sẽ trích xuất được nhiều đặc trưng hơn, và cho kết quả phân loại chính xác hơn. Trong trường hợp biến đổi wavelet rời rạc, ta áp dụng một filter-bank trên tín hiệu tổng thể cùng một lúc. Bằng cách này, ta có được một mẫu có độ phân giải từ đơn giản đến phức tạp trên miền thời gian - tần số. Sự khác biệt chính giữa WT và STFT đó là: (1) STFT là đồng nhất nhưng CWT thì không. (2) STFT xử lý trên các đoạn dữ liệu đã phân đoạn, nhưng CWT xử lý tín hiệu một cách tổng thể. (3) STFT liên quan đến các phép biến đổi Fourier nhưng CWT yêu cầu một bộ lọc trực giao. Sử dụng biến đổi wavelet sẽ là một cách tiếp cận mới để phân tích tín hiệu động thay vì Biến đổi Fourier.

Trong trường hợp biến đổi wavelet liên tục, một yếu tố quan trọng phải nghiên cứu đó là cách đưa các ma trận hệ số CWT vào CNN. Cách tiếp cận hiện nay là đặt các hệ số 2D (hình ảnh) của chín tín hiệu chồng lên nhau như ba kênh đỏ, lục, lam (RGB) của một hình ảnh màu. Theo đó, tất cả các đặc trưng phụ thuộc giữa các dữ liệu thu từ các cảm biến khác nhau có thể được tính toán đồng thời. Tuy nhiên, nếu kết nối liên tiếp các hệ số CWT của chín tín hiệu khác

nhau thành một mảng, ta sẽ quan sát được những thay đổi đột ngột giữa các ma trận. Điều này có thể dẫn đến mô hình CNN tập trung vào các điểm nổi hay các ranh giới của sự thay đổi thay vì các đặc tính quan trọng của mỗi tín hiệu. Theo đó, ta sẽ cần một mạng CNN có độ phức tạp và sâu hơn để khắc phục loại nhiễu này.

#### 4. KẾT LUẬN

Bài báo trình bày một kết quả so sánh độ chính xác giữa các phương pháp xử lý tín hiệu số bằng biến đổi Fourier, biến đổi Fourier nhanh và biến đổi Wavelet. Bằng việc thử nghiệm trên bộ dữ liệu thống kê thực tế hoạt động của người (HAR), bài báo đã phân tích và so sánh các ưu nhược điểm của từng phương pháp và mô hình đề xuất sử dụng biến đổi wavelet kết hợp các phương pháp học máy cho kết quả phân loại đạt 94% cải thiện đáng kể so với các phương pháp truyền thống khác. Kết quả nghiên cứu này cho thấy tiềm năng ứng dụng của công nghệ học máy kết hợp các phương pháp tiền xử lý cho kết quả cao, có khả năng ứng dụng thực tế.

---

#### TÀI LIỆU THAM KHẢO

- [1]. N. Q. Trung, 2016. *Xu lý tín hiệu và lọc số*. Science and Technics Publishing House, Hanoi.
- [2]. Soh Ping Jack, et al, 2016. *Advances in Machine Learning and Signal Processing*. Proceedings of MALSIP 2015. Vol. 387, Springer.
- [3]. Deng Jia, et al, 2009. *Imagenet: A large-scale hierarchical image database*. IEEE conference on computer vision and pattern recognition.
- [4]. Lezhenin Iurii, et al, 2019. *Urban sound classification using long short-term memory neural network*. Federated Conference on Computer Science and Information Systems (FedCSIS), IEEE.
- [5]. P. D. Tuan, 2013. *The Fourier transform and its application to solving a differential and integral equation*. PhD thesis, VNU University of Science.
- [6]. Fujieda Shin, et al, 2018. *Wavelet convolutional neural networks*. arXiv preprint arXiv:1805.08620.
- [7]. <https://archive.ics.uci.edu/ml/datasets/human+activity+recognition+using+smartphones>.

---

#### AUTHOR INFORMATION

**Dinh Thi Thuy Duong**

Academy of Military Science and Technology