

Sử dụng biến đổi Wavelet và các thuật toán học máy để phân loại các trạng thái biểu cảm khuôn mặt thông qua tín hiệu điện não đồ

Võ Hoàng Thủy Tiên^{1,2}, Nguyễn Thị Như Quỳnh^{1,2}, Nguyễn Thanh Phước^{1,2}, Huỳnh Văn Tuấn^{1,2,*}



Use your smartphone to scan this QR code and download this article

TÓM TẮT

Các thuật toán máy học được sử dụng để giải quyết các bài toán phân loại là để đánh giá hiệu suất của các giải thuật mang tính thời sự nhất. Mục đích là cố gắng cải thiện kết quả của một phương pháp đã được sử dụng trong một nghiên cứu tương tự như phương pháp được sử dụng trong nghiên cứu này. Bài báo trình bày việc xây dựng bộ dữ liệu đa trạng thái liên quan đến các hoạt động của mắt và biểu hiện của khuôn mặt. Tín hiệu được ghi và lưu trữ bởi sự kết nối của thiết bị di động ít kênh. Dữ liệu được tối ưu bằng các kỹ thuật chuẩn hóa z-score, max-min. Kỹ thuật kiểm chứng chéo ứng dụng phương pháp phân tầng lặp lại được sử dụng để phân chia dữ liệu thành các đoạn huấn luyện/kiểm tra. Các đặc trưng của tín hiệu điện não gồm các dải tần delta, theta, alpha và beta được phân tách bằng phương pháp biến đổi Wavelet họ Daubechies. Các tính năng trên miền thời gian và tần số được trích xuất tính toán năng lượng tổng cộng, năng lượng thành phần chi tiết, năng lượng thành phần xấp xỉ, năng lượng tương đối. Ba thuật toán, máy véc-tơ hỗ trợ, k-láng giềng gần nhất và thuật toán tổng hợp, được sử dụng để phát triển thành 17 mô hình phân loại, nhằm tối ưu hóa hiệu suất phân loại của các thuật toán máy học. Các mô hình này được sử dụng qua đó các tham số được khảo sát và tối ưu hóa, nhằm có thể đề xuất được một mô hình phân loại tốt nhất cho bộ dữ liệu Data-021. Mô hình Subspace ensemble được đề nghị vì hiệu suất mô hình đạt 87,7%.

Từ khóa: biểu hiện khuôn mặt, điện não đồ, máy học, phân loại, xử lý tín hiệu

¹Khoa Vật lý – Vật lý Kỹ thuật, Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM, Việt Nam

²Đại học Quốc gia Thành phố Hồ Chí Minh, Việt Nam

Liên hệ

Huỳnh Văn Tuấn, Khoa Vật lý – Vật lý Kỹ thuật, Trường Đại học Khoa học Tự nhiên, ĐHQG- HCM, Việt Nam

Đại học Quốc gia Thành phố Hồ Chí Minh, Việt Nam

Email: hvtuan@hcmus.edu.vn

Lịch sử

- Ngày nhận: 18-12-2021
- Ngày chấp nhận: 30-5-2022
- Ngày đăng: 30-6-2022

DOI: 10.32508/stdjns.v6i2.1157



Bản quyền

© ĐHQG Tp.HCM. Đây là bài báo công bố mở được phát hành theo các điều khoản của the Creative Commons Attribution 4.0 International license.



GIỚI THIỆU

Thông tin trạng thái của não được ghi lại bằng các tín hiệu điện não đồ sinh lý (EEG), được sử dụng rộng rãi để nghiên cứu các hoạt động khác nhau của não. Một ý nghĩa điển hình trong việc hỗ trợ chẩn đoán các loại bệnh như động kinh¹, các loại bệnh liên quan đến thần kinh vận động như bệnh parkinson², hay bệnh Alzheimer³. Bên cạnh đó, phân loại các trạng thái của tín hiệu điện não đồ vẫn là bài toán được các nhà nghiên cứu quan tâm vì tính nền tảng và khả năng ứng dụng cao. Cụ thể như các bài toán phân loại dữ liệu ghi nhận các trạng thái cảm xúc⁴, nhận dạng và mô phỏng lại các trạng thái cảm xúc⁵, nhận dạng nét mặt dựa trên điện não đồ⁶, nhận dạng các hành vi khi đang lái xe bằng tín hiệu EEG⁷, các trạng thái tín hiệu tương tự động tay, chân, hay điều khiển cầm, nắm vật^{8,9}. Phân loại tín hiệu EEG về tư duy, suy nghĩ¹⁰, hay tính bảo mật xác thực của sóng não¹¹. Dữ liệu về sóng não là đa dạng và không giới hạn lĩnh vực, đây vẫn là hộp đen rộng lớn đầy bí ẩn để khám phá. Một trong những phương thức hỗ trợ tối ưu để giải đáp các bài toán phân loại, dự đoán tín hiệu điện não đồ chính là các thuật toán máy học. Gần đây, các phương pháp tổng hợp để phân loại tín hiệu EEG đã thu hút sự chú ý ngày càng tăng trong

giới học thuật. Sun và cộng sự đã đánh giá hiệu suất của ba phương pháp tổng hợp phổ biến, đó là bagging, boosting và random subspace. Họ báo cáo rằng khả năng của các phương thức tổng hợp phụ thuộc vào các bộ phân loại cơ sở, đặc biệt là các cài đặt và tham số được sử dụng cho từng bộ phân loại riêng lẻ¹². Dehuri và cộng sự¹³ đã trình bày tập hợp các phương pháp mạng thần kinh chức năng cơ sở xuyên tâm (RBFNs) để xác định các cơn co giật động kinh. Phương pháp này dựa trên phương pháp bagging và sử dụng các RBFN tiến hóa khác biệt (DE) làm bộ phân loại cơ sở. Điện não đồ được phân tách với sự biến đổi Wavelet thành các dải con khác nhau và một số thông tin thống kê được trích xuất từ các hệ số Wavelet để cung cấp làm đầu vào cho một tập hợp các DE-RBFN. Kết quả phân loại xác nhận rằng nhóm DE-RBFN được đề xuất có tiềm năng lớn hơn để xác định các rối loạn động kinh. Nhận biết cảm xúc từ các tín hiệu điện não đồ bằng cách sử dụng thuật toán phân tích dạng kinh nghiệm (Empirical Mode Decomposition, EMD) được Degirmenci và các tác giả trình bày. Họ sử dụng EMD trong giai đoạn xử lý bởi những ưu điểm như phân tích được những tín hiệu không tuyến tính và không cố định. Nhóm tác giả sử dụng một số thuật toán học máy để phân loại

Trích dẫn bài báo này: Tiên V H T, Quỳnh N T N, Phước N T, Tuấn H V. **Sử dụng biến đổi Wavelet và các thuật toán học máy để phân loại các trạng thái biểu cảm khuôn mặt thông qua tín hiệu điện não đồ.** *Sci. Tech. Dev. J. - Nat. Sci.*; 6(2):2116-2130.

tín hiệu như máy véc-tơ hỗ trợ (Support Vector Machine, SVM), tích biệt thức tuyến tính (Linear Discriminant Analysis, LDA) và Naïve Bayes. Thuật toán máy véc-tơ hỗ trợ cho kết quả tốt nhất với độ chính xác, độ nhạy và độ đặc hiệu ứng với các giá trị lần lượt là 87%, 86% và 97%¹⁴. Nhóm tác giả sử dụng tín hiệu EEG để phục vụ cho việc phân loại bệnh động kinh. Bài báo đưa ra một mô hình nhằm nâng cao độ chính xác trong phân loại các tín hiệu nhiễu trong khi vẫn giữ được lượng thông tin phức tạp. Phương pháp biến đổi Wavelet rời rạc được sử dụng để trích xuất tính năng và đưa vào khảo sát với các thuật toán phân loại như máy véc-tơ hỗ trợ (Support Vector Machine), mạng thần kinh nhân tạo (Artificial Neural Network), Naïve Bayes cùng với bộ phân loại tổ hợp tín hiệu nhận biết nhiễu (NSC) này kết hợp bốn mô hình phân loại dựa trên hiệu suất riêng lẻ của chúng. NSC cho các kết quả phân loại tốt nhất với các tín hiệu có tỷ lệ tín hiệu trên nhiễu (SNR) với 1dB, 5dB, 10dB lần lượt là 0,80 0,84 và 0,88. Đặc biệt độ chính xác với tín hiệu “sạch” lên đến 0,9 và cao hơn so với các thuật toán phân loại khác¹⁵. Yu Chen và cộng sự¹⁶ sử dụng các tính năng ở miền thời gian, miền thời gian-tần số và các tính năng phi tuyến do chúng có những đặc điểm mang tính toàn diện và thích hợp. Ngoài ra, phương pháp LDA còn được sử dụng để lựa chọn tính năng nhằm cải thiện kết quả phân loại. Thuật toán Ensemble với mô hình Adaboost được sử dụng và đạt được độ chính xác trung bình ở chiều dominance là 88,70%. Bộ phân loại có thể tập trung tốt hơn vào các mẫu phân loại sai, nhờ đó cải thiện được khả năng tổng quát hóa, tránh tình trạng quá mức (overfitting) và cải thiện hiệu suất phân loại cảm xúc. Zhuang và các cộng sự đã sử dụng bộ dữ liệu DEAP và xử lý tín hiệu điện não bởi thuật toán EMD. Ưu điểm của EMD là lợi thế sử dụng những thông tin dao động hơn những phương pháp khác. Ngoài ra, khi so sánh với phép biến đổi Wavelet thì EMD còn có thể phân tích tín hiệu tự động và bỏ qua việc lựa chọn cửa sổ. Tín hiệu EEG sẽ được phân tích thành những hàm bản chất (Intrinsic Mode Function, IMF). Các thông tin của IMF được chọn làm các tính năng như sự khác nhau về thời gian, sự khác nhau về pha và năng lượng được chuẩn hóa. Thành phần IMF1 cho kết quả tốt nhất, độ chính xác ở chiều valence và arousal lần lượt là 70,41% và 72,10%¹⁷.

Xu hướng xác định một mô hình tốt nhất đã phổ biến từ lâu, cho dù mô hình đó dựa trên máy học hay thống kê thì độ chính xác cho một mô hình cũng phần nào hỗ trợ mạnh mẽ trong các phát triển ứng dụng phân loại EEG. Bên cạnh các mô hình ý nghĩa sâu sắc trong lĩnh vực y tế, đáp ứng sự phát triển và nhu cầu sử dụng của đại đa số con người cũng ngày càng tăng. Các ứng

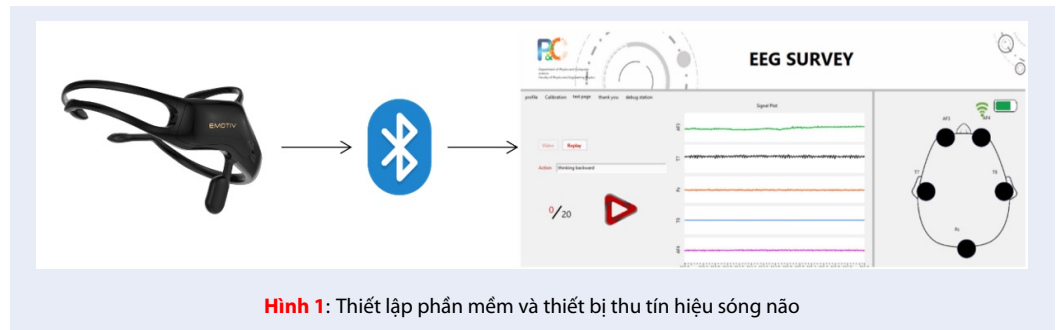
dụng tối ưu hóa cuộc sống con người được phát triển và sử dụng càng ngày càng phổ biến. Để đóng góp vào nguồn dữ liệu nghiên cứu, chúng tôi tiến hành thí nghiệm thu thập dữ liệu, khảo sát các kỹ thuật chuẩn hóa và phân tách dữ liệu huấn luyện, kiểm tra. Các thuật toán máy học được sử dụng để phân loại các trạng thái dữ liệu thu được.

DỮ LIỆU VÀ PHƯƠNG PHÁP

Dữ liệu phân loại

Bộ dữ liệu *data-021* là sản phẩm của Khoa Vật lý-Vật lý Kỹ thuật, Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM. Dữ liệu ghi lại các hoạt động thay đổi điện thế ở bề mặt vỏ não bằng kỹ thuật điện não đồ. Các tín hiệu sóng não thu được ở dạng tín hiệu số liên tục theo thời gian. Nghiên cứu sử dụng thiết bị EMOTIV Insight với 5 kênh tín hiệu lần lượt là AF3, AF4, T7, T8 và Pz. Thiết bị được kết nối không dây đến phần mềm thu dữ liệu EEG SURVEY. Đây là phần mềm được nhóm nghiên cứu xây dựng trên môi trường lập trình LabVIEW. Công cụ ghi tín hiệu cho phép người sử dụng nhập các thông tin khảo sát được từ người tham gia thí nghiệm, giao diện ứng dụng cho phép quan sát các đối tượng như nhân/ trạng thái tín hiệu, số lượng mẫu đã thu, đồ thị sóng của năm kênh tín hiệu, chất lượng tiếp xúc của các kênh tín hiệu, thời lượng pin của thiết bị cũng như chất lượng kết nối không dây giữa máy tính và thiết bị thu tín hiệu (Hình 1).

Thiết kế thí nghiệm gồm một máy tính đã cài đặt phần mềm EEG SURVEY, một thiết bị ghi tín hiệu là EMOTIV Insight, một màn hình hiển thị các đoạn băng ghi hình ảnh mô tả từng trạng thái để người tham gia thí nghiệm thực hiện theo các mô tả có trong đoạn băng. Kỹ thuật viên làm việc với máy tính và thiết bị thu. Người tham gia thí nghiệm được yêu cầu giữ sức khỏe tốt, ngủ đủ giấc vào đêm trước thí nghiệm, không sử dụng chất kích thích, không có tiền sử các bệnh mãn tính liên quan đến thần kinh. Nơi diễn ra thí nghiệm là không gian kín, hạn chế tối đa các loại tiếng ồn từ môi trường xung quanh cũng như các tác nhân gây ảnh hưởng đến sự tập trung của người tham gia thí nghiệm, tuy nhiên tất cả chỉ nằm ở mức tương đối cho phép. Người tham gia được yêu cầu ngồi thoải mái, thả lỏng, ổn định tinh thần trong suốt quá trình ghi tín hiệu. Bộ dữ liệu *data-021* gồm có 10 người tham gia thí nghiệm, mỗi người sẽ được thu bảy trạng thái gồm nhắm mắt, mở mắt, liếc mắt sang trái, liếc mắt sang phải, nhướn mày, cười mỉm và trạng thái bình thường (lần lượt tương ứng với các nhãn *close eye*, *open eye*, *eye left*, *eye right*, *eye brown*, *smile* và *normal* như mô tả Hình 2). Mỗi nhãn được



Hình 1: Thiết lập phần mềm và thiết bị thu tín hiệu sóng não

lặp lại 20 lần, mỗi lần kéo dài 8 giây. Tần số lấy mẫu là 128 Hz. Mô tả cụ thể của từng nhân/ trạng thái như sau:

- Trạng thái bình thường (*normal*): Đối tượng ngồi thoải mái trên ghế, mở mắt, bình tĩnh và tránh bất kỳ công việc suy nghĩ.
- Nhắm mắt (*close eye*): Từ trạng thái bình thường, đối tượng đang mở mắt sau khi ghi tín hiệu từ 2–4 giây đối tượng được yêu cầu nhắm mắt nhân nhả, tránh dao động và cơ mắt co mạnh trong phần còn lại của bản ghi.
- Mở mắt (*open eye*): Từ trạng thái bình thường, đối tượng đang nhắm mắt sau khi ghi tín hiệu từ 2–4 giây đối tượng được yêu cầu mở mắt cho đến hết 8 giây.
- Liếc mắt sang trái (*eye left*): Từ trạng thái bình thường, 2–4 giây sau khi bắt đầu ghi, đối tượng được yêu cầu liếc mắt sang trái và giữ nguyên trạng thái cho đến khi quá trình ghi dừng.
- Liếc mắt sang phải (*eye right*): Từ trạng thái bình thường, 2–4 giây sau khi bắt đầu ghi, đối tượng được yêu cầu liếc mắt sang phải và tiếp tục cho đến khi quá trình ghi dừng lại.
- Cười mỉm (*smile*): Từ trạng thái bình thường, 2–4 giây sau khi bắt đầu ghi, đối tượng được yêu cầu cười mỉm nhẹ nhàng (không hé môi, hở răng) và tiếp tục giữ nguyên trạng thái cho đến khi quá trình ghi dừng lại.
- Nhướn mày (*eye brown*): Từ trạng thái bình thường, 2–4 giây sau khi bắt đầu ghi hình, đối tượng được yêu cầu nhướn mày như thể đang ngạc nhiên.

Sau khi hoàn tất mỗi bản ghi được lưu trữ lại như cấu trúc thư mục ở Hình 2 với định dạng .txt. Như mô tả Hình 3, thông tin trong mỗi tệp dữ liệu gồm có giá trị của năm kênh tín hiệu tương ứng với các cột IED–AF3, IED–T7, IED–O1, IED–T8 và IED–AF4, các thông số thời gian và giá trị theo hai trục tọa độ của con quay hồi chuyển. Trong nghiên cứu này tập trung sử dụng giá trị năm kênh tín hiệu điện não để phân loại bẫy nhân.

Chuẩn hóa dữ liệu

Mỗi thuật toán có những giả định về dữ liệu khác nhau nên dữ liệu cần được chuẩn hóa trước khi phân loại. Có hai kỹ thuật chuẩn hóa dữ liệu là bình thường hóa dữ liệu (*normalization*) bằng phương pháp chuẩn hóa max-min và chuẩn hóa dữ liệu (*standardization*) bằng phương pháp chuẩn hóa z-score^{18,19}. Bình thường hóa max-min là một kỹ thuật đơn giản trong đó kỹ thuật này có thể khớp dữ liệu một cách cụ thể trong một ranh giới xác định trước với một ranh giới xác định trước. Đây là sự điều chỉnh tỷ lệ sao cho dữ liệu nằm trong khoảng [0; 1] hoặc cũng có thể [-1; 1] bằng cách áp dụng công thức (1). Trong đó, x là giá trị dữ liệu chưa chuẩn hóa, x_{min} là giá trị dữ liệu nhỏ nhất, x_{max} là giá trị dữ liệu lớn nhất, x_{new} là dữ liệu sau khi chuẩn hóa²⁰.

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

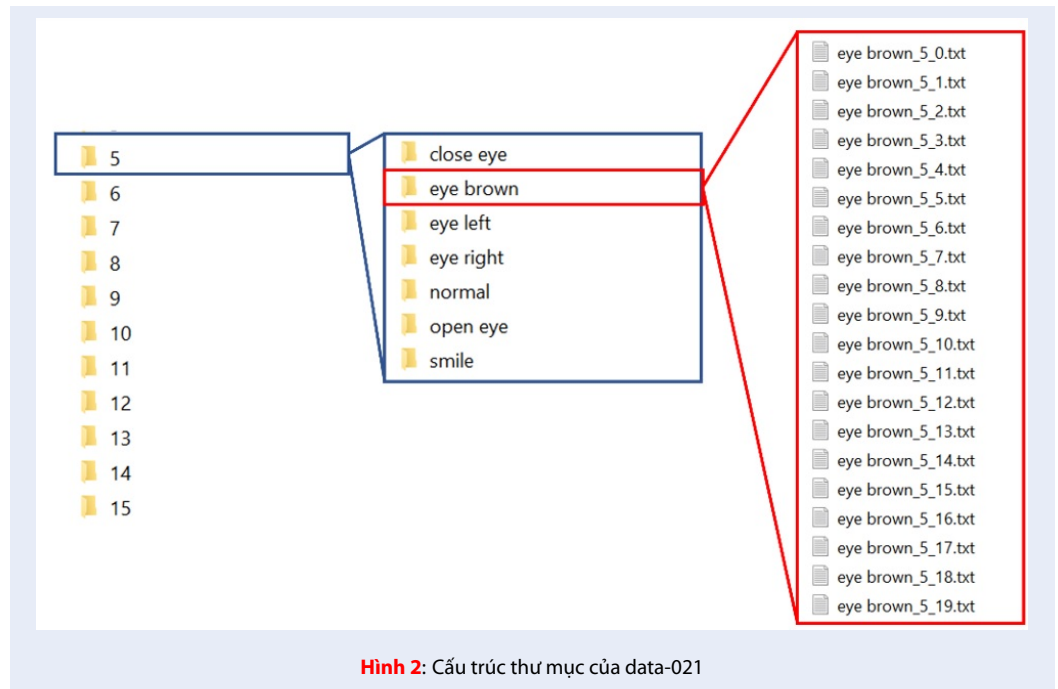
Kỹ thuật chuẩn hóa dữ liệu được sử dụng phổ biến nhất là z-score được tính bằng giá trị trung bình cộng (μ) và độ lệch chuẩn (σ) của dữ liệu đã cho dựa vào công thức (2). Đây là kỹ thuật tỷ lệ dữ liệu giúp cho giá trị của mỗi đặc trưng hay quan sát có giá trị trung bình bằng 0 bằng cách trừ đi μ ở tử số và phương sai của phân phối bằng 1 khi chia cho σ ở mẫu số²¹.

$$x_{new} = \frac{x - \mu}{\sigma} \quad (2)$$

Trong nghiên cứu này, dữ liệu được chuẩn hóa bằng các kỹ thuật max-min, z-score trước khi tiến hành các bước trích xuất đặc trưng của dữ liệu. Nghiên cứu cũng trình bày một số kết quả khảo sát được khi dữ liệu chưa được chuẩn hóa để có những nhận định khách quan về tầm quan trọng của chuẩn hóa dữ liệu.

Kiểm chứng chéo dữ liệu

Phân tách dữ liệu thành các tập huấn luyện và kiểm tra là một đặc trưng của các mô hình học có giám sát của các thuật toán máy học^{22,23}. Khi cho các tham số vào một hàm dự đoán và kiểm tra trên cùng một tập dữ



Hình 2: Cấu trúc thư mục của data-021

Giá trị lấy mẫu		Ref.	Giá trị kênh tín hiệu AF3, T7, O1, T8, AF4					GYRO		Thời gian	Giá trị đánh dấu			Mark
IED_COUNTER	IED_INTERPOLATED		IED_RAW_CQ	IED_AF3	IED_T7	IED_O1	IED_T8	IED_AF4	IED_GYROX	IED_GYROZ	IED_TIMESTAMP	IED_MARKER	IED_SYNC_SIGNAL	
116	0	450	4203.077	4157.436	4121.026	4133.846	4165.641	8190	8212	2351.389	0	0	0	0
117	0	450	4209.231	4151.282	4103.59	4139.487	4156.923	8192	8214	2351.397	0	0	0	0
118	0	450	4182.051	4127.692	4066.667	4116.41	4140	8189	8216	2351.404	0	0	0	0
119	0	16197	4173.333	4131.795	4068.718	4110.769	4146.667	8195	8213	2351.412	0	0	0	0
120	0	450	4178.462	4131.795	4072.821	4104.103	4153.333	8197	8208	2351.42	0	0	0	0

Hình 3: Cấu trúc của một tệp dữ liệu

liệu, đây là sai lầm vì mô hình chỉ gặp các nhân mà mô hình vừa kiểm tra nên kết quả thu được tuyệt đối đúng và kết quả này không có ý nghĩa. Đây chính là vấn đề được gọi là quá mức (*overfitting*)²⁴. Kiểm chứng chéo là giải pháp cho vấn đề quá mức sử dụng một phương pháp thống kê đánh giá và so sánh các thuật toán học tập bằng cách chia dữ liệu thành hai phân đoạn: một phân đoạn dùng để học hoặc đào tạo mô hình và phân đoạn còn lại dùng để xác thực mô hình. Bên cạnh đó, kiểm chứng chéo là kỹ thuật lấy mẫu để đánh giá mô hình học máy trong trường hợp dữ liệu không được đối ngẫu²⁵. Dữ liệu được chia thành k đoạn (*k-fold*) và lặp lại k lần. Trong đó, k đại diện cho số nhóm dữ liệu được chia như mô tả Hình 4. Giá trị của k được chọn sao cho mỗi tập dữ liệu huấn luyện/ kiểm tra chứa số lượng mẫu dữ liệu đủ lớn để đại diện về mặt thống kê cho tập dữ liệu rộng hơn. Kiểm chứng chéo có bốn

biến thể gồm:

- *Train/ Test split*: chỉ tạo ra duy nhất một tập huấn luyện và một tập kiểm tra để đánh giá mô hình. Đây là trường hợp đặc biệt của CV khi $k=2$ ²⁶.
- *Leave-one-out CV*: đây là trường hợp k bằng với kích thước của tập dữ liệu, tức tất cả đối tượng trong dữ liệu đều có cơ hội được đưa ra khỏi tập dữ liệu²⁷.
- *Stratified*: khi dữ liệu được chia thành *k-fold*, mỗi phần sẽ chứa cùng một tỷ lệ phân loại nhất định²⁸. Tương ứng với dữ liệu trong nghiên cứu, data-021 có kích thước 1400 mẫu, được chia thành 10-fold, mỗi fold chứa 220 mẫu bao gồm đầy đủ bảy nhân.
- *Repeated k-fold* kiểm chứng chéo lần lượt được lặp lại *n* lần, trong đó quan trọng là mẫu dữ liệu

được xáo trộn trước mỗi lần lặp lại, dẫn đến sự phân chia mẫu khác nhau²⁹.

Trong nghiên cứu này, các biến thể *Stratified* và *Repeated* được kết hợp với nhau nhằm tạo ra các *k-fold* với kích thước và số lượng mẫu tương ứng mỗi nhân là như nhau ở mỗi đoạn (*fold*), quá trình kiểm chứng được lặp lại 10 lần và dữ liệu được xáo trộn một cách ngẫu nhiên. Phương pháp này được gọi là kiểm chứng chéo *k-fold Stratified Repeated* hay gọi tắt là *k-fold SRCV*. Bước đầu sử dụng kỹ thuật 2-fold SRCV để khảo sát tất cả các mô hình phân loại nhằm lựa chọn mô hình tối ưu trước khi khảo sát tối ưu hóa các tham số.

Trích xuất đặc trưng

Nghiên cứu sử dụng phương pháp biến đổi Wavelet rời rạc (DWT) họ db4 để phân tách dữ liệu thành năm mức³⁰. Tín hiệu đầu vào có tần số lấy mẫu là 128 Hz, thu được các giá trị tần số ở các mức sau mỗi lần phân tách tương ứng với nhịp sóng não đặc trưng (Bảng 1).

Bảng 1: Mối liên hệ giữa các hệ số Wavelet với các nhịp EEG

Hệ số Wavelet	Tần số (Hz)	Nhịp EEG
D_1	> 63	Nhiều
D_2	32 – 63	γ
D_3	16 – 32	β
D_4	8 – 16	α
D_5	4 – 8	θ
A_5	0 – 4	δ

Điện não đồ được phân tách thành năm dãy sóng đặc trưng là delta, theta, alpha, beta và gamma³¹. Tuy nhiên, gamma là nhịp sóng chỉ có thể thu được nếu sử dụng kỹ thuật lấy mẫu xâm lấn (tức các phương pháp cấy ghép điện não tiếp xúc với não hoặc sâu trong tế bào não). Nghiên cứu này ghi tín hiệu bằng kỹ thuật không xâm lấn (sử dụng điện cực khô tiếp xúc với da đầu) nên các tần số tương ứng gamma và lớn hơn sẽ được xem là nhiễu và loại bỏ. Các nhịp sóng còn lại được sử dụng tương ứng với các hệ số chi tiết Wavelet là D_5, D_4, D_3 và hệ số xấp xỉ là A_5 ³².

$$E_{D_i} = \sum_{j=1}^N |D_{ij}|^2, \quad i = 3, 4, 5 \tag{3}$$

$$E_{A_5} = \sum_{j=1}^N |A_{5j}|^2 \tag{4}$$

$$E_{total} = \sum_{i=3}^5 E_{D_i} + E_{A_5} \tag{5}$$

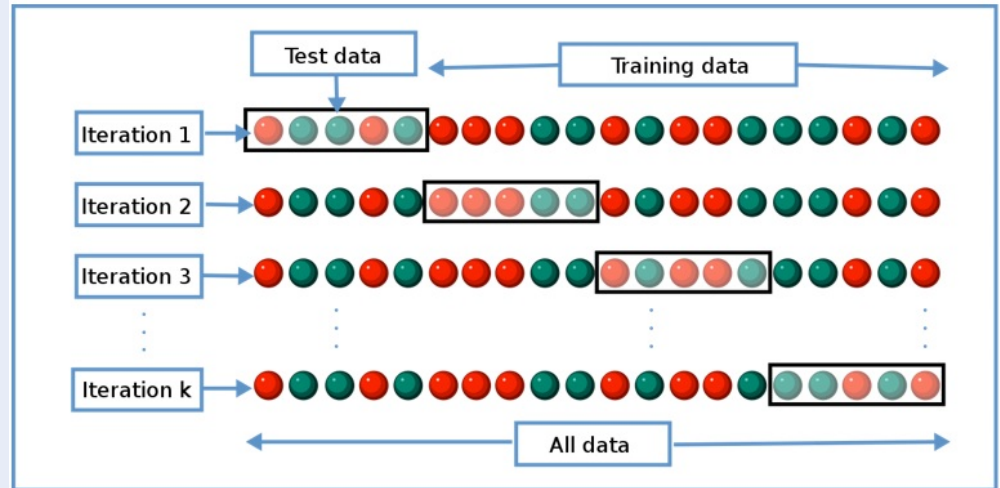
Từ các thành phần đặc trưng trên, nghiên cứu đề xuất một dạng ma trận đặc trưng làm đầu vào cho các

thuật toán máy học. Mỗi kênh tín hiệu được phân tách thành bốn dãy nhịp sóng đặc trưng của EEG, trong đó ba nhịp theta, alpha, beta tương ứng hệ số chi tiết D_5, D_4, D_3 , nhịp delta tương ứng thành phần hệ số xấp xỉ A_5 . Năng lượng thành phần xấp xỉ của năm kênh tín hiệu, $E_A (5 \times N)$ được tính theo công thức (4), trong đó N là kích thước dữ liệu. Năng lượng thành phần chi tiết của năm kênh tín hiệu $E_{D_{3-5}} (15 \times N)$ tính theo công thức (3), năng lượng tổng cộng $E_{total} (1 \times N)$ tính theo công thức (5). Do đó, một ma trận gồm 21 tính năng được xây dựng để cải thiện hiệu quả phân loại, trong đó N tương ứng với 1400 mẫu dữ liệu.

Thuật toán máy học

Thuật toán tổng hợp – ensemble method

Bagging, *boosting* và *random subspace* là ba kỹ thuật phổ biến nhất của thuật toán tổng hợp. *Bagging* là một kỹ thuật sử dụng phương pháp bỏ phiếu theo đa số. Kỹ thuật quy tắc biểu quyết đa số (majority voting) thu thập các phiếu bầu của tất cả các bộ phân loại và điều tra tên lớp mà hầu hết được các bộ phân loại báo cáo. Sau đó, mô hình chọn lớp được báo cáo nhiều nhất như một quyết định cuối cùng³³. *Boosting* là một phương pháp đưa ra kết quả học tập cuối cùng bằng cách dựa vào các dự đoán kết quả của những lần học trước. Đầu tiên tạo ra mô hình phân loại yếu, các mô hình cải tiến sau kế thừa và các điểm bị phân loại sai từ các mô hình trước được đánh trọng số lớn hơn trung bình và tiếp tục lặp lại, kế thừa, cải tiến mô hình cho đến khi tạo được mô hình được xem là học giỏi³⁴. Như vậy, các mô hình mới bị ảnh hưởng bởi hiệu suất của những mô hình trước. Kỹ thuật *random subspace* được sử dụng cho biến phân hồi phân loại được gọi là phân loại và phân hồi liên tục được gọi là hồi quy. Đây là một phương pháp kết hợp các mô hình tương đối gần đây. Máy học được huấn luyện trên không gian con được chọn ngẫu nhiên của không gian đầu vào ban đầu (tức là tập huấn luyện được lấy mẫu trong không gian đặc trưng). Kết quả đầu ra của các mô hình sau đó được kết hợp với nhau, thường là dùng bình chọn đa số để đưa ra quyết định phân loại cuối cùng³⁵. Nghiên cứu trình bày năm mô hình điển hình của thuật toán tổng hợp. Mô hình Bagged Trees sử dụng kỹ thuật *bagging* với kiểu ra quyết định là bình chọn theo số đồng. Mô hình Boosted Trees, RUSboosted trees sử dụng kỹ thuật *boosting* cùng với bình chọn kết quả theo số đồng. Mô hình Random Subspace, Subspace Ensemble sử dụng kỹ thuật *random subspace*.



Hình 4: Kỹ thuật xác thực chéo k-lần

Bảng 2: Mô tả các mô hình phân loại của thuật toán SVM

Models	Kernel function	Constraint	Kernel scale	Multiclass method
Linear SVM	linear	1	1	one vs one
Quadratic SVM	quadratic	1	1	one vs one
Cubic SVM	cubic	1	1	one vs one
Fine Gaussian SVM	gaussian	1	1.3	one vs one
Medium Gaussian SVM	gaussian	1	5.2	one vs one
Coarse Gaussian SVM	gaussian	1	21	one vs one

Máy véc-tơ hỗ trợ - SVM

Động lực chính đằng sau SVM là giải quyết trực tiếp mục tiêu tổng quát hóa tốt bằng cách đồng thời tối đa hóa hiệu suất của máy trong khi giảm thiểu độ phức tạp của mô hình đã học. Đối với dữ liệu *data-021* có bảy nhãn, đây là bài toán phân loại đa lớp. Cách thức để giải quyết bài toán là giảm vấn đề phân loại đa lớp thành một tập hợp các bài toán con phân loại nhị phân, với mỗi lần học SVM sẽ giải quyết lần lượt các bài toán phân loại hai lớp. Có hai phương pháp chuyển hóa phân loại đa lớp thành hai lớp, được gọi là kỹ thuật một – một (*one vs one*) và một – nhiều (*one vs all*). Khi một nhãn được xem là tích cực và tất cả các nhãn còn lại là tiêu cực, số bài toán con mà mô hình phải học là k , trong đó k là số nhãn, được gọi là phương pháp một – nhiều (*one vs all*)³⁶. Như vậy, *data-021* có bảy nhãn thì các mô hình phân loại phải học và giải quyết bảy bài toán con. Khi một nhãn là tích cực, nhãn khác là tiêu cực và phần còn lại bị bỏ qua. Thiết kế này loại bỏ tất cả sự kết hợp của các bài

tập cặp lớp. Số lần SVM phải học và phân loại các bài toán con là $(k(k-1))/2$. Tương ứng dữ liệu *data-021* là 21 lần học và phân loại của SVM. Hàm hạt nhân (*kernel*) giúp chuyển vấn đề không phân tách được thành phân tách được, tức chuyển đổi bài toán đa lớp sang nhị phân. Quy trình chuyển đổi của hàm hạt nhân khá phức tạp, yêu cầu là cần tìm ra quy trình tách dữ liệu dựa trên các nhãn hoặc các kết quả đầu ra trước đó đã xác định được³⁷. Các hàm hạt nhân được sử dụng là *gaussian*, *cubic*, *quadratic*, *linear*^{38,39}. Khi dữ liệu không phân loại được gọi là phân loại sai, có hệ số C gọi là hệ số ràng buộc (*constraint*)⁴⁰. Hệ số C là tham số kiểm soát sự cân bằng giữa hai điều kiện là khoảng cách siêu mặt phẳng đến điểm dữ liệu phải lớn nhất nhưng tỷ lệ huấn luyện sai của tập huấn luyện đạt nhỏ nhất. Khi tăng C làm tăng trọng số của các phân loại sai giúp phân loại được cải thiện và chặt chẽ hơn. Trong thuật toán SVM có sáu mô hình phân loại điển hình là Linear SVM, Quadratic SVM, Cubic SVM, Fine Gaussian SVM, Medium Gaussian SVM, Coarse Gaussian SVM được mô tả ở Bảng 2.

K-láng giếng gần nhất – k-NN

kNN đây là một thuật toán lười học hay được gọi là thuật toán học dựa trên phiên bản hay học dựa trên trí nhớ. Thuật toán sử dụng kỹ thuật bình chọn số đồng để quyết định nhãn của bộ dữ liệu huấn luyện, các bộ phân loại kNN thường có độ chính xác dự đoán tốt ở kích thước thấp⁴¹. Một dữ liệu không được gán nhãn sau đó được phân loại theo nhãn của các dữ liệu xung quanh và được gán nhãn theo điểm dữ liệu gần nhất. Các biến thể của lược đồ này bao gồm bộ phân loại kNN, sử dụng phiếu bầu của kNN được gán nhãn gần nhất và bộ phân loại chọn lọc, lưu trữ và sử dụng các ví dụ được gán nhãn một cách chọn lọc⁴². Nghiên cứu này sử dụng bộ phân loại kNN với các nhãn được quy định bằng kỹ thuật sử dụng số phiếu bầu nhiều nhất. Các mô hình phân loại kNN được sử dụng gồm các mô hình kết hợp công thức tính khoảng cách Euclidean: Fine kNN, Medium kNN, Coarse kNN; Cosine kNN; Cubic kNN; Weighted kNN. Bên cạnh đó, các kỹ thuật đánh trọng số cũng giúp cải thiện hiệu suất mô hình hơn. Trong k điểm gần nhất, tất cả các điểm dữ liệu được gán nhãn theo nhãn có số lượng điểm dữ liệu nhiều nhất trong k điểm. Tuy nhiên, khi đánh giá như vậy thì vai trò của k điểm là như nhau. Trong k điểm, có điểm gần hơn đáng tin cậy hơn, điểm ở xa kém tin cậy hơn. Để cải thiện vấn đề, các điểm dữ liệu được đánh trọng số với các công thức $w=1/d^2$ (trọng số bằng nghịch đảo bình phương khoảng cách); $w=1/d$ (trọng số bằng nghịch đảo khoảng cách). Nhờ vào việc đánh trọng số, đặc tính của từng điểm dữ liệu đã được phân định rõ ràng. Nghiên cứu này sử dụng phương thức đánh trọng số bằng nghịch đảo bình phương khoảng cách, đây là phương thức thấy rõ nhất khi điểm càng tin cậy thì trọng số càng lớn.

KẾT QUẢ VÀ THẢO LUẬN

Đầu tiên, dữ liệu được chuẩn hóa bằng hai kỹ thuật *max-min*, *z-score* và không chuẩn hóa dữ liệu. Sau đó, DWT – db4 phân tách 5 mức được sử dụng để trích xuất 21 tính năng và khởi tạo ma trận 1400×21 được gọi là ma trận tính năng làm đầu vào cho 17 mô hình phân loại. Nghiên cứu sử dụng kỹ thuật 2-fold SRCV để kiểm chứng chéo dữ liệu. Độ chính xác và thời gian xử lý của các mô hình được trình bày chi tiết ở Bảng 3. Hiệu suất của các kỹ thuật kết hợp với các mô hình được đánh giá thông qua độ chính xác và thời gian xử lý dữ liệu. Dựa vào kết quả khảo sát, bình thường hóa dữ liệu bằng kỹ thuật *max-min* cho kết quả tốt hơn so với chuẩn hóa dữ liệu *z-score* và không chuẩn hóa dữ liệu. Chuẩn hóa dữ liệu *z-score* là bước quan trọng đối với các tín hiệu không cùng đơn vị, vì các biến được đo lường ở các tỷ lệ khác nhau không đóng góp như

nhau vào phân tích và cuối cùng có thể tạo ra một rào cản. Tuy nhiên, đối với dữ liệu *data-021* là kiểu tín hiệu số liên tục theo thời gian (với đơn vị biên độ là μV). Vì vậy, chuẩn hóa *z-score* không thật sự nổi bật bằng bình thường hóa *max-min* mặc dù so với không chuẩn hóa dữ liệu thì kết quả phân loại được cải thiện hơn.

Đối với thuật toán tổng hợp, việc sử dụng bình chọn theo số đồng để đưa ra nhận định cuối là một phương pháp rất phổ biến. Tuy nhiên, khi đánh giá như vậy thì vai trò của tất cả kết quả bỏ phiếu đều là như nhau, các kết quả đáng tin cậy hơn được đánh giá ngang với các kết quả không đáng tin cậy. Vì nhược điểm này, các mô hình sử dụng bình chọn theo số đồng sẽ cho kết quả không tốt bằng mô hình sử dụng kỹ thuật khác, điển hình là kết hợp kỹ thuật quyết định kNN. Các mô hình như Boosted trees, RUSboosted trees, Bagged trees có độ chính xác lần lượt 51,5%, 45,5% và 74,6%. Ngoài ra, Random subspace là phần mở rộng ý tưởng của *bagging* và được phát triển như là một đối thủ cạnh tranh với *bagging*. Random subspace đã thể hiện rõ lợi thế khi kết quả thu được là cao nhất với 81,2%, hiệu quả hơn so với Bagged trees, mặc dù đây cũng là một trong những mô hình tốt nhất của thuật toán tổng hợp.

Đối với thuật toán SVM, được sử dụng để ánh xạ không gian đầu vào của thuật toán vào không gian tính năng có chiều cao. Khi hàm hạt nhân là *linear*, các dữ liệu được xem là tuyến tính, dữ liệu trong nghiên cứu là tín hiệu biến thiên theo thời gian nên dẫn đến chồng lấp các thông tin và tính năng, hiệu suất phân loại cũng giảm đi, kết quả đạt được 59,6%. Hàm hạt nhân *Gaussian* sử dụng các đường cong thông thường xung quanh các điểm dữ liệu và tính tổng các điểm dữ liệu sao cho ranh giới quyết định có thể được xác định bởi một loại điều kiện để phân tách được các đặc trưng của từng nhãn. Chính vì vậy, kết quả khả quan hơn so với hàm *linear* với độ chính xác là 78,1% cho mô hình Fine Gaussian SVM. Fine Gaussian SVM là mô hình có tính chất của một đa thức bậc cao. Tích vô hướng giữa hai hàm đặc trưng được tính bởi sự kết hợp giữa hai véc-tơ mà không cần tìm biểu diễn của dữ liệu trước khi ánh xạ sang chiều không gian cao. Điều này giúp tiết kiệm chi phí tính toán khi biết trước được định dạng của hàm hạt nhân. Chính vì vậy, mô hình cubic SVM cho kết quả phân loại 81,7% và thời gian xử lý dữ liệu khá tốt so các mô hình còn lại.

Đối với kNN, khoảng cách Euclidean áp dụng định lý Pytago để tính khoảng cách trong không gian hai chiều. Đây là công thức rất phổ biến, dễ thực hiện và đạt kết quả tốt trong nhiều trường hợp. Khoảng cách này chỉ hiệu quả với dữ liệu có chiều không gian thấp

Bảng 3: Độ chính xác và thời gian xử lý của 17 mô hình phân loại của các thuật toán tổng hợp, thuật toán SVM và thuật toán kNN

Mô hình phân loại	max-min		zscore		Không chuẩn hóa	
	ACC (%)	Thời gian (giây)	ACC (%)	Thời gian (giây)	ACC (%)	Thời gian (giây)
Boosted trees	55,1	12,11	54,6	15,84	54,9	12,85
Bagged trees	74,6	6,14	74,8	8,62	73,6	6,20
Subspace discriminant	55,4	9,07	55,1	11,62	54,6	9,57
Subspace Ensemble	81,2	15,98	80,2	18,99	81,4	16,03
RUSboosted trees	45,5	21,44	44,1	31,15	45,1	21,89
Linear SVM	59,6	3,87	60,0	3,44	60,3	3,36
Quadratic SVM	67,5	7,53	66,7	6,99	68,0	6,05
Cubic SVM	81,7	5,63	81,9	5,59	81,9	5,10
Fine Gaussian SVM	78,1	4,60	79,4	4,94	78,4	4,42
Medium Gaussian SVM	74,9	8,94	75,4	7,89	76,6	6,89
Coarse Gaussian SVM	54,1	10,86	53,9	8,79	53,7	7,65
Fine kNN	80,1	1,61	77,5	1,98	78,2	1,92
Medium kNN	76,0	0,75	74,7	0,84	75,9	0,81
Coarse kNN	49,8	2,11	49,8	2,80	50,5	2,44
Cosine kNN	70,1	2,00	70,2	2,60	71,0	2,31
Cubic kNN	73,4	8,39	72,6	8,85	73,2	8,36
Weighted kNN	79,3	1,82	77,9	2,07	78,6	2,03

và dễ bị ảnh hưởng bởi các tính năng. Vì vậy, cần phải có bước chuẩn hóa dữ liệu trước khi tính toán. Mô hình Fine kNN kết hợp chuẩn hóa dữ liệu *max-min* đã chứng minh được những nhận định trên bằng độ chính xác tốt nhất 80,1%. Medium kNN và Coarse kNN cùng chung công thức khoảng cách nhưng việc lấy số điểm k lân cận quá lớn đã làm tỷ lệ chống lẫn giữa các trạng thái bị nâng cao dẫn đến hiệu suất phân loại của hai mô hình không được tối ưu.

Sau khi đã xác định được mô hình tốt nhất của từng thuật toán tương ứng, nghiên cứu sẽ khảo sát các kết quả liên quan đến kỹ thuật kiểm chứng chéo. Cần phải lựa chọn tham số k phù hợp để thỏa đủ hai điều kiện là *Stratified* và *Repeated*. Số lượng mẫu giữa các k-fold phải bằng nhau, mỗi k-fold phải chứa bảy nhãn và số lượng nhãn trong mỗi k-fold phải bằng nhau (thỏa *Stratified*). Dữ liệu data-021 có 1400 mẫu, gồm bảy nhãn, mỗi nhãn có 220 mẫu. Vậy k phải là số mà cả 1400 và 220 đều chia hết cho k. Đối với máy học, khi tăng số lượng mẫu huấn luyện thì hiệu suất mô hình sẽ được cải thiện và đi kèm là thời gian huấn luyện cũng tăng. Các giá trị k có thể nhận là [5, 10, 20, 25,

50, 100], mỗi giá trị k tương ứng sẽ được lặp lại n=10 lần (thỏa *Repeated*). Kết quả của mô hình Subspace Ensemble, Cubic SVM, Fine kNN lần lượt được thể hiện ở các Bảng 4 và Bảng 5. Các kết quả sẽ được đánh giá dựa trên độ chính xác và độ lệch chuẩn, vì tính chất lặp lại xáo trộn ngẫu nhiên các mẫu dữ liệu nên việc đánh giá mô hình dựa trên độ lệch chuẩn sẽ giúp dễ dàng nhận biết mức độ ổn định của mô hình. Như mô tả ở Bảng 4 và Bảng 5, Subspace Ensemble với 50-fold SRCV cho hiệu suất phân loại tốt nhất với độ chính xác trung bình là 86,8% và độ lệch chuẩn StD 0,06%. Mô hình Fine kNN đạt kết quả 84,46% và StD 0,04% với kỹ thuật kiểm chứng chéo là 100-fold SRCV. Đối với Cubic SVM, phương pháp phân loại một – một với số lần học và phân loại ít hơn xấp xỉ một nửa đã giảm thiểu chi phí tính toán cho mô hình, thời gian huấn luyện và kiểm tra nhanh hơn. Bên cạnh đó, số lượng nhãn phân loại khá nhiều (bảy nhãn) nhưng độ chính xác phân loại 85,42% khi áp dụng phương pháp một – một. Độ lệch chuẩn ở 100-fold ổn định hơn so với 25-fold và 50-fold. Tuy nhiên, khi phân đoạn dữ liệu càng nhiều, số lượng mẫu huấn

Bảng 4: Kết quả khảo sát SRCV của mô hình Subspace Ensemble và Fine k-NN

k-fold	Subspace Ensemble		Fine k-NN	
	ACC (%)	StD (%)	ACC (%)	StD (%)
5	85,8	0,19	83,46	0,12
10	86,2	0,12	84,29	0,09
25	86,7	0,11	84,21	0,05
50	86,8	0,06	84,45	0,05
100	86,7	0,07	84,46	0,04

Bảng 5: Kết quả khảo sát SRCV của mô hình phân loại Cubic SVM

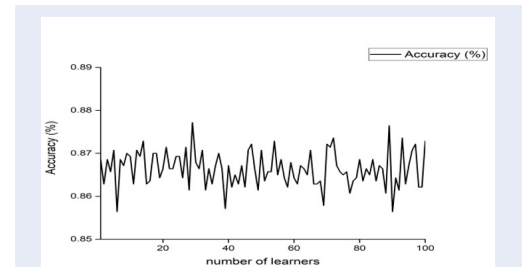
k-fold	One - one		One - all	
	ACC (%)	StD (%)	ACC (%)	StD (%)
5	84,12	0,12	83,68	0,15
10	84,74	0,16	84,46	0,10
25	85,42	0,07	84,48	0,09
50	85,54	0,07	84,90	0,11
100	85,54	0,06	85,04	0,07

luyện càng tăng giúp cải thiện hiệu suất của mô hình phân loại trở nên tốt hơn, độ ổn định của mô hình được thể hiện qua phần trăm độ lệch chuẩn rất thấp nhưng thời gian huấn luyện mô hình cũng tăng theo do chi phí tính toán bị thay đổi. Vì vậy, mô hình mang đầy đủ tính chất hiệu suất cao, ổn định và hoạt động nhanh khi áp dụng 25-fold Cubic SVM.

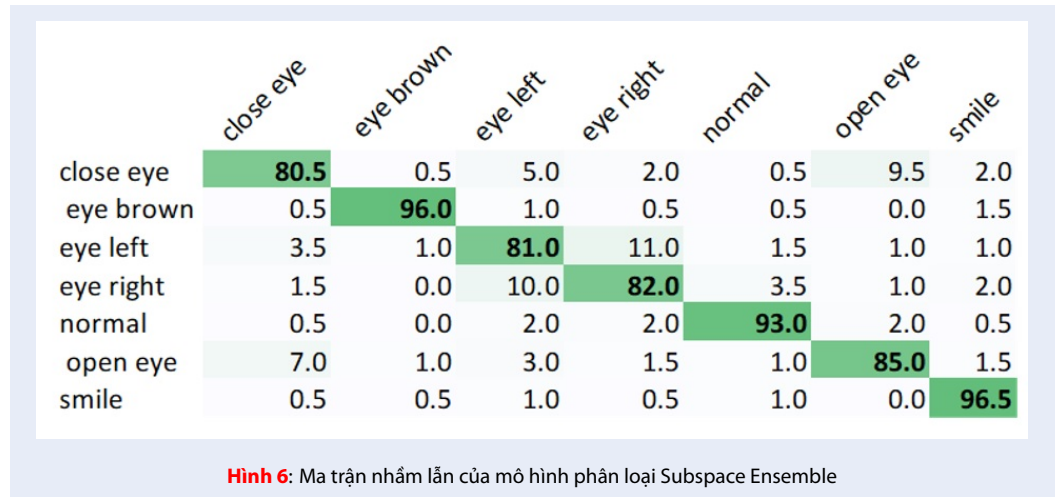
Subspace Ensemble

Thuật toán tổng hợp sử dụng mô hình phân loại Subspace Ensemble 50-fold SRCV độ chính xác 86,8% là kết quả ghi nhận được sau các khảo sát. Tham số có thể tối ưu hóa thêm cho mô hình Subspace Ensemble chính là số lần học. Hình 5 mô tả sự biến đổi của độ chính xác theo số lần học của mô hình. Nghiên cứu khảo sát tham số này từ 1-100 với bước nhảy là 1. Ở lần học thứ 29, mô hình ghi nhận được kết quả phân loại 87,7%. Như vậy, sau 29 lần học mô hình đạt hiệu suất tối ưu và kết quả phân loại từng nhãn được thể hiện ở ma trận nhầm lẫn Hình 6. Tất cả các nhãn đều đạt kết quả hơn 80%, tỷ lệ nhầm lẫn giữa hai nhãn *open eye* và *close eye* từ 7% đến 9%. Hai nhãn có tỷ lệ phân loại sai lên đến hơn 10% là *eye left* và *eye right*. Các nhãn liên quan đến biểu hiện gương mặt, khá độc lập và đặc trưng về mặt hình ảnh nên kết quả phân loại rất tốt lần lượt 96% và 96,5% tương ứng với *eye brown* và *smile*. Rajdeep và cộng sự⁴³ đã sử dụng kỹ thuật *Adaboost* thuộc phương pháp *boosting* của thuật toán tổng hợp để phân loại tín hiệu hành động điều khiển

động cơ với bốn nhãn. Trích xuất đặc trưng tín hiệu bằng cách tính năng lượng dải và năng lượng *entropy*. Các tác giả các kỹ thuật trích xuất đặc trưng chẳng hạn như biến đổi Wavelet rời rạc (DWT) hoặc căn bậc hai dựa trên Wavelet (RMS) và năng lượng-entropy (EngEnt), mật độ phổ công suất, công suất dải (Bp) và các thông số tự động phục hồi thích ứng (AAR). Phân loại bằng kỹ thuật *Adaboost* đạt 83,57% khi sử dụng tính năng năng lượng-entropy. Chúng tôi đã cải thiện hơn khoảng 3% so với các kết quả của Rajdeep nhờ vào việc tập trung vào các thành phần chi tiết của biến đổi Wavelet và kết hợp ba tính năng là năng lượng dải, năng lượng tổng cộng và năng lượng thành phần chi tiết.



Hình 5: Độ chính xác của mô hình Subspace Ensemble và số lần học



Cubic SVM

Đối với mô hình phân loại là Cubic SVM sử dụng hàm hạt nhân là cubic (hay được gọi tắt là poly) nên tham số tối ưu hóa cho mô hình này chính là ràng buộc C . Hình 7 mô tả sự phụ thuộc của kết quả phân loại với hệ số ràng buộc. Tham số C được khảo sát từ 10–200 với bước nhảy là 10. Hiệu quả của mô hình ghi nhận được khi tham số $C=60$ độ chính xác là 86,3%. Khi tăng ràng buộc, thời gian xử lý dữ liệu của mô hình càng tăng. Tuy nhiên, C càng lớn thì biên của mặt phẳng phân loại càng nhỏ cho phép sai lệch càng bé, tỷ lệ phân loại sai giảm, chính điều này làm cho kết quả phân loại của mô hình được cải thiện hơn. Bên cạnh đó, Hình 8 cũng cho thấy kết quả phân loại của các nhãn và tỷ lệ nhầm lẫn, phân loại sai giữa các nhãn với nhau. Các nhãn về biểu hiện khuôn mặt cho kết quả phân loại tốt nhất 92,5% cho nhãn *eye brown* và 90,5% cho nhãn *smile*. Hai nhãn *eye left* và *eye right* có cải thiện tỷ lệ nhầm lẫn hơn, chỉ còn 6% đến 10%. Đối với SVM, Rajdeep và cộng sự⁴³ cũng thu được kết quả 76.7% khi sử dụng hàm hạt nhân là RBF (hàm gaussian) khi sử dụng tính năng tránh xuất đặc trưng là công suất dải. Ngoài ra, Chatterjee và cộng sự⁴⁴ cũng sử dụng các tính năng tương tự Rajdeep là năng lượng dải và năng lượng entropy với kết quả lần lượt là 81,43% và 85% để phân loại bốn nhãn dữ liệu (tượng tượng cũ động tay, chân trái, phải). Khi trích xuất bằng năng lượng entropy thể hiện rõ đặc tính của dữ liệu nên kết quả được cải thiện hơn khi trích xuất bằng năng lượng dải. Bên cạnh đó, Isa và các tác giả⁴⁵ cũng sử dụng SVM để phân loại nhãn dữ liệu khi sử dụng FFT để trích xuất đặc trưng của tín hiệu trên miền tần số. Dữ liệu được kiểm chứng chéo 10-fold và kết quả là 78,61% cho phân loại hai nhóm dữ liệu chuyển động tay (gồm hai nhãn tay trái và tay phải) và chân

(gồm chân trái và chân phải). Tuy nhiên, số lượng nhãn phân loại đã được trình bày ở đây nhiều hơn với nghiên cứu của Chatterjee, Isa và kết quả mô hình Cubic SVM SRCV là 86,3% cho thấy mô hình được khảo sát và tối ưu hóa các tham số tốt hơn giúp độ chính xác được nâng cao.



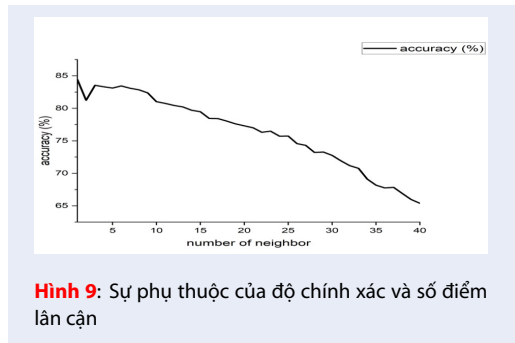
Fine k – Nearest Neighbor

Mô hình kNN giảm tính linh hoạt khi bắt đầu cài đặt thông số k , tức khảo sát số điểm lân cận. Như Hình 9, k được khảo sát trong khoảng từ 1 – 40 điểm, độ chính xác chỉ tốt khi số điểm lân cận ít, k lớn dần thì tỷ lệ phân loại nhầm lẫn càng lớn dẫn đến độ chính xác giảm liên tục. Khi lựa chọn số điểm xung quanh càng lớn, xác suất chông lẩn khi phân loại giữa các nhãn dữ liệu với nhau sẽ tăng. Vì vậy, dựa vào khảo sát ở Hình 9 tham số k tốt nhất khi bằng 1 tối ưu hóa được độ chính xác lên đến 84,4%. Isa và các tác giả⁴⁵ cũng sử dụng kNN để phân loại các tín hiệu phân loại các hành động (bốn nhãn). Áp dụng kiểm chứng chéo 10-fold để huấn luyện và kiểm tra dữ liệu, mô hình phân loại sử dụng 15 điểm lân cận và tính khoảng cách theo

	close eye	eye brown	eye left	eye right	normal	open eye	smile
close eye	85.5	0.0	2.0	3.5	0.0	8.0	1.0
eye brown	0.0	92.5	1.0	1.5	1.5	1.0	2.5
eye left	3.0	0.0	80.5	10.5	3.0	1.5	1.5
eye right	2.0	1.5	6.5	81.0	4.5	2.5	2.0
normal	1.0	0.5	2.0	4.5	88.0	2.5	1.5
open eye	9.0	0.5	2.5	1.5	4.0	82.0	0.5
smile	0.5	0.5	2.5	3.5	2.0	0.5	90.5

Hình 8: Ma trận nhầm lẫn của mô hình phân loại Cubic SVM

công thức Euclidean thu được kết quả 70,8%. Mô hình của chúng tôi đa dạng về số lượng nhãn, mặc dù kích thước dữ liệu chưa lớn nhưng vẫn được cải thiện bằng kỹ thuật SRCV đã giúp cho mô hình có kết quả trội hơn. Độ nhạy giữa các trạng thái liên quan về mắt vẫn cao, có thể do chất lượng tập trung của người tham gia thí nghiệm cũng ảnh hưởng đến chất lượng của dữ liệu. Vì môi trường diễn ra thí nghiệm chỉ nằm ở mức tương đối chứ không hoàn toàn lý tưởng. Tuy nhiên, các trạng thái về biểu hiện khuôn mặt ghi nhận kết quả 94,5% và 94,9% (Hình 10) lần lượt cho các trạng thái *eye brown* và *smile*. Điều này mở ra cơ hội tiến đến các ứng dụng thực tế sử dụng các trạng thái liên quan đến biểu hiện khuôn mặt để điều khiển hay các ứng dụng liên quan trong lĩnh vực giao tiếp người – máy. Bên cạnh thuật toán SVM, Isa và các tác giả⁴⁵ cũng trình bày các kết quả từ thuật toán k-NN. Kết quả của việc phân loại dữ liệu EEG bằng cách áp dụng trình phân loại k-NN với các giá trị khác nhau của k và số liệu khoảng cách được đã được Isa trình bày. Độ chính xác 70,8% với số điểm lân cận $k=15$ và công thức tính khoảng cách là Minkowski. Ngoài ra, Hindarto và các tác giả⁴⁶ cũng đưa ra kết quả cải thiện hơn so với Isa khoảng 6% (độ chính xác 76%) chỉ với số điểm lân cận bằng 3, để phân loại hoạt động của não để điều khiển con trỏ trên màn hình máy tính. Như vậy, mô hình của chúng tôi (Euclidean 1-NN 84,4%) chỉ với số điểm $k=1$ và dùng hàm tính khoảng cách Euclidean đã cải thiện độ phức tạp của mô hình tốt hơn so với khoảng cách Minkowski của Isa và hạn chế sự chồng chập dữ liệu dẫn đến sự phân loại sai được giảm thiểu tốt hơn từ 8% đến 14% so với kết quả của Hindarto và Isa. Những kết quả đối sánh được trình bày như Bảng 6.



Hình 9: Sự phụ thuộc của độ chính xác và số điểm lân cận

KẾT LUẬN

Nghiên cứu và phân tích các tín hiệu điện não bằng các thuật toán phân loại được phát triển ngày càng tăng trong thế giới khoa học. Nghiên cứu này cung cấp một bộ dữ liệu gồm bảy nhãn chứa các tín hiệu khác nhau liên quan đến hành vi mắt và biểu hiện khuôn mặt. Một ứng dụng ghi, lưu và xuất dữ liệu được đề xuất giúp tối giản các bước chuyển đổi dữ liệu EEG hơn so với ứng dụng hỗ trợ kèm theo có sẵn của thiết bị ghi tín hiệu. Các tính năng được xây dựng thành dạng ma trận hai chiều giúp cải thiện làm việc phân loại được nhiều mẫu dữ liệu hơn so với đầu vào là các véc-tơ một chiều. Đề xuất kỹ thuật kiểm chứng dữ liệu mới là SRCV giải quyết được vấn đề quá mức mà các bộ dữ liệu có kích thước nhỏ dễ mắc sai lầm. Khảo sát đa dạng các mô hình có được trong ba thuật toán phân loại máy véc-tơ hỗ trợ, k-NN và thuật toán tổng hợp. Nghiên cứu này đề xuất mô hình Subspace Ensemble 50-fold SRCV với kết quả phân loại 87,7%, mô hình Cubic SVM 25-SRCV với kết quả 86,3% và mô hình có hiệu suất phân loại là 84,4% Euclidean 1-NN. Bên cạnh đó, các thách thức vẫn còn như dữ liệu

	close eye	eye brown	eye left	eye right	normal	open eye	smile
close eye	76.6	0.5	7.5	2.0	1.0	11.0	1.5
eye brown	0.1	94.5	1.6	1.5	0.5	0.0	2.1
eye left	4.5	1.0	72.6	17.4	2.5	1.0	1.0
eye right	2.5	0.0	13.5	78.1	2.5	1.5	2.0
normal	0.5	0.0	2.0	2.5	92.0	1.6	1.4
open eye	7.5	1.0	4.1	2.5	1.0	82.5	1.5
smile	0.5	0.5	0.6	0.5	2.1	1.0	94.9

Hình 10: Ma trận nhầm lẫn của mô hình phân loại Fine k-NN

Bảng 6: So sánh kết quả với các công bố trước

Tác giả/ tài liệu	Phương pháp	Kết quả
⁴⁴ Chatterjee và cộng sự	Năng lượng dài + SVM	81,43%
	Năng lượng entropy Wavelet + SVM	85,00%
⁴³ Rajdeep và cộng sự	Năng lượng dài + SVM	76,70%
	Năng lượng dài + Adaboost	83,57%
⁴⁵ Isa và các tác giả	15-NN	70,80%
	SVM	78,61%
⁴⁶ Hindarto và các tác giả	3-NN	76,00%
Kết quả trình bày trong bài báo	Subspace Ensemble SRCV	87,70%
	Cubic SVM SRCV	86,30%
	Euclidean 1-NN	84,40%

cần được chất lượng hóa và nâng cao kích thước, số lượng mẫu. Hướng đến phương pháp khảo sát tối ưu hóa tất cả tham số và đề xuất mô hình tự động. Mô hình phân tích, xử lý và phân loại định tính theo thời gian thực cũng là một bài toán hướng đến giải quyết. Trong tương lai, nhiều thuật toán sẽ được khảo sát kỹ lưỡng để phân loại điện não đồ.

LỜI CẢM ƠN

Nghiên cứu được tài trợ bởi Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM trong khuôn khổ Đề tài mã số T2021-01.

DANH MỤC CÁC TỪ VIẾT TẮT

AAR Adaptive Autoregressive Thích nghi hồi quy tự động
 CV Cross-Validation Xác thực chéo
 DE Differential Evolution Tiến hóa khác biệt

EMD Empirical Mode Decomposition Phân tích dạng kinh nghiệm

k-NN k-Nearest Neighbor K láng giềng gần nhất

LDA Linear Discriminant Analysis Tích biệt thức tuyến tính

NSC Noise-aware Signal Combination Tổ hợp tín hiệu nhận biết nhiễu

RBFNs Radial Basis Function Neural Networks Mạng thần kinh chức năng cơ sở xuyên tâm

RBF Radial Basis Function Hàm cơ sở xuyên tâm

RMS Root mean square Sai số toàn phương trung bình

SVM Support Vector Machine Máy véc-tơ hỗ trợ

SRCV Stratified Repeated Cross-Validation Xác thực chéo phân tầng lặp lại

SNR Signal to Noise Ratio Tỷ lệ tín hiệu trên nhiễu

XUNG ĐỘT LỢI ÍCH

Các tác giả đồng ý không có bất kì xung đột lợi ích nào liên quan đến các kết quả công bố.

ĐÓNG GÓP CỦA CÁC TÁC GIẢ

Tác giả Võ Hoàng Thủy Tiên viết chương trình, viết và tổng hợp bản thảo.

Tác giả Nguyễn Thị Như Quỳnh tham gia đo đạc và xử lý tín hiệu.

Tác giả Nguyễn Thanh Phước tham gia đo đạc tín hiệu.

Tác giả Huỳnh Văn Tuấn tham gia viết và chỉnh sửa bản thảo.

TÀI LIỆU THAM KHẢO

1. Srinivasan V, Eswaran C, Sriraam NJToTiB. Approximate entropy-based epileptic EEG detection using artificial neural networks. *IEEE Transactions on information Technology in Biomedicine*;11(3):288-295;PMID: 17521078. Available from: <https://doi.org/10.1109/TITB.2006.884369>.
2. Stam C, Jelles B, Achtereekte H, Rombouts S, Slaets J, Kenen RJE., Investigation of EEG non-linearity in dementia and Parkinson's disease. *Electroencephalography and clinical neurophysiology*. 1995; 95(5):309-317; Available from: [https://doi.org/10.1016/0013-4694\(95\)00147-Q](https://doi.org/10.1016/0013-4694(95)00147-Q).
3. Pritchard WS, Duke DW, Coburn KL, Moore NC, Tucker KA, Jann MW. EEG-based, neural-net predictive classification of Alzheimer's disease versus control subjects is augmented by non-linear EEG measures. *Electroencephalography and clinical Neurophysiology*;1994;91(2):118-130; Available from: [https://doi.org/10.1016/0013-4694\(94\)90033-7](https://doi.org/10.1016/0013-4694(94)90033-7).
4. Wang X-W, Nie D, Lu B-LJN. Emotional state classification from EEG data using machine learning approach. *Neurocomputing*;2014;129:94-106; Available from: <https://doi.org/10.1016/j.neucom.2013.06.046>.
5. Liu Y, Sourina O, Nguyen MK, editors. Real-time EEG-based human emotion recognition and visualization. 2010 international conference on cyberworlds; 2010: IEEE; Available from: <https://doi.org/10.1109/CW.2010.37>.
6. Raheel A, Majid M, Anwar SM, editors. Facial expression recognition based on electroencephalography. 2019 2nd international conference on computing, mathematics and engineering technologies (iCoMET); 2019: IEEE; Available from: <https://doi.org/10.1109/ICOMET.2019.8673408>.
7. Yang L, Ma R, Zhang HM, Guan W, Jiang SJAA, Prevention. Driving behavior recognition using EEG data from a simulated car-following experiment. *Accident Analysis & Prevention*;2018;116:30-40;PMID: 29174606. Available from: <https://doi.org/10.1016/j.aap.2017.11.010>.
8. Guger C, Schlögl A, Neuper C, Walterspacher D, Strein T, Pfurtscheller GJIToNS, Rapid prototyping of an EEG-based brain-computer interface (BCI). *IEEE Transactions on Neural Systems and Rehabilitation Engineering*;2001;9(1):49-58;PMID: 11482363. Available from: <https://doi.org/10.1109/7333.918276>.
9. Penalzoa CI, Nishio SJSR. BMI control of a third arm for multi-tasking. *Science Robotics*;2018;3(20);PMID: 33141729. Available from: <https://doi.org/10.1126/scirobotics.aat1228>.
10. Stevens Jr CE, Zabelina DLJN. Classifying creativity: Applying machine learning techniques to divergent thinking EEG data. *NeuroImage*. 2020;219:116990;PMID: 32474083. Available from: <https://doi.org/10.1016/j.neuroimage.2020.116990>.
11. Chuang J, Nguyen H, Wang C, Johnson B, editors. I think, therefore I am: Usability and security of authentication using brainwaves. International conference on financial cryptography and data security; 2013: Springer; Available from: https://doi.org/10.1007/978-3-642-41320-9_1.
12. Sun S, Zhang C, Zhang DJPRL. An experimental evaluation of ensemble methods for EEG signal classification. *Pattern Recognition Letters*;2007;28(15):2157-63; Available from: <https://doi.org/10.1016/j.patrec.2007.06.018>.
13. Dehuri S, Jagadev AK, Cho S-BJPCS. Epileptic seizure identification from electroencephalography signal using DE-RBFNs ensemble. *Procedia Computer Science*;2013;23:84-95; Available from: <https://doi.org/10.1016/j.procs.2013.10.012>.
14. Degirmenci M, Ozdemir MA, Sadighzadeh R, Akan A, editors. Emotion Recognition from EEG Signals by Using Empirical Mode Decomposition. 2018 Medical Technologies National Congress (TIPTEKNO); 2018: IEEE;PMID: 29389513. Available from: <https://doi.org/10.1109/TIPTEKNO.2018.8597061>.
15. Abualsaud K, Mahmuddin M, Saleh M, Mohamed AJTJSWJ. Ensemble classifier for epileptic seizure detection for imperfect EEG data. *The Scientific World Journal*;2015;2015;PMID: 25759863. Available from: <https://doi.org/10.1155/2015/945689>.
16. Chen Y, Chang R, Guo JJMPIE. Emotion Recognition of EEG Signals Based on the Ensemble Learning Method: AdaBoost. *Mathematical Problems in Engineering*;2021;2021; Available from: <https://doi.org/10.1155/2021/8896062>.
17. Zhuang N, Zeng Y, Tong L, Zhang C, Zhang H, Yan BJBrI. Emotion recognition from EEG signals using multidimensional information in EMD domain. *BioMed research international*;2017;2017;PMID: 28900626. Available from: <https://doi.org/10.1155/2017/8317357>.
18. Patro S, Sahu KKJapa. Normalization: A preprocessing stage. *arXiv preprint arXiv*;2015; Available from: <https://doi.org/10.17148/IARJSET.2015.2305>.
19. Vanzant ES, Cochran RC, Titgemeyer ECJJoas. Standardization of in situ techniques for ruminant feedstuff evaluation. *Journal of animal science*;1998;76(10):2717-29;PMID: 9814915. Available from: <https://doi.org/10.2527/1998.76102717x>.
20. Liu ZJPES. A method of SVM with normalization in intrusion detection. *Procedia Environmental Sciences* 2011;11:256-62; Available from: <https://doi.org/10.1016/j.proenv.2011.12.040>.
21. Saranya C, Manikandan GJJJoE, Technology. A study on normalization techniques for privacy preserving data mining. *International Journal of Engineering and Technology (IJET)*. 2013;5(3):2701-4;.
22. Crisci C, Ghattas B, Perera GJEM. A review of supervised machine learning algorithms and their applications to ecological data. *Ecological Modelling*;2012;240:113-22; Available from: <https://doi.org/10.1016/j.ecolmodel.2012.03.001>.
23. Vabalas A, Gowen E, Poliakoff E, Casson AJJPo. Machine learning algorithm validation with a limited sample size. *PLoS one*;2019;14(11):e0224365;PMID: 31697686. Available from: <https://doi.org/10.1371/journal.pone.0224365>.
24. Dietterich TJAcS. Overfitting and undercomputing in machine learning. *ACM computing surveys (CSUR)*. 1995;27(3):326-7; Available from: <https://doi.org/10.1145/212094.212114>.
25. Refaeilzadeh P, Tang L, Liu HJEOds. Cross-validation. *Encyclopedia of database systems*;2009;5:532-8; Available from: https://doi.org/10.1007/978-0-387-39940-9_565.
26. Reitermanova Z, editor Data splitting. WDS; 2010;.
27. Wong T-TJPR. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*;2015;48(9):2839-46; Available from: <https://doi.org/10.1016/j.patcog.2015.03.009>.
28. Zeng X, Martinez TRJJoE, Intelligence TA. Distribution-balanced stratified cross-validation for accuracy estimation. *Journal of Experimental & Theoretical Artificial Intelligence*;2000;12(1):1-12; Available from: <https://doi.org/10.1080/095281300146272>.
29. Kim J-HJCS, analysis d. Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. *Computational statistics & data analysis*;2009;53(11):3735-45; Available from: <https://doi.org/10.1016/j.csda.2009.04.009>.

30. Jawerth B, Sweldens WJSr. An overview of wavelet based multiresolution analyses. *SIAM review*.1994;36(3):377-412;Available from: <https://doi.org/10.1137/1036095>.
31. Buzsaki G. *Rhythms of the Brain*: Oxford university press; 2006;.
32. Portilla J, Simoncelli EPJlJocv. A parametric texture model based on joint statistics of complex wavelet coefficients. *International journal of computer vision*.2000;40(1):49-70;Available from: <https://doi.org/10.1023/A:1026553619983>.
33. Ahangi A, Karamnejad M, Mohammadi N, Ebrahimpour R, Bagheri NJNC, Applications. Multiple classifier system for EEG signal classification with application to brain-computer interfaces. *Neural Computing and Applications*.2013;23(5):1319-27;Available from: <https://doi.org/10.1007/s00521-012-1074-3>.
34. Zhang C, Ma Y. *Ensemble machine learning: methods and applications*: Springer; 2012;Available from: <https://doi.org/10.1007/978-1-4419-9326-7>.
35. Dong X, Yu Z, Cao W, Shi Y, Ma QJFoCS. A survey on ensemble learning. *Frontiers of Computer Science*.2020;14(2):241-58;Available from: <https://doi.org/10.1007/s11704-019-8208-z>.
36. Hsu C-W, Lin C-JJltoNN. A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks*.2002;13(2):415-25;PMID: 18244442. Available from: <https://doi.org/10.1109/72.991427>.
37. Patle A, Chouhan DS, editors. SVM kernel functions for classification. 2013 International Conference on Advances in Technology and Engineering (ICATE); 2013: IEEE;Available from: <https://doi.org/10.1109/ICAdTE.2013.6524743>.
38. Han S, Qubo C, Meng H, editors. Parameter selection in SVM with RBF kernel function. *World Automation Congress 2012*; 2012: IEEE;.
39. Hussain M, Wajid SK, Elzaart A, Berbar M, editors. A comparison of SVM kernel functions for breast cancer detection. 2011 eighth international conference computer graphics, imaging and visualization; 2011: IEEE;Available from: <https://doi.org/10.1109/CGIV.2011.31>.
40. Reddy S, Reddy KT, Kumari VV, Varma KVJJoCS, Technologies I. An SVM based approach to breast cancer classification using RBF and polynomial kernel functions with varying arguments. *International Journal of Computer Science and Information Technologies*.2014;5(4):5901-4;.
41. Raymer ML, Punch WF, Goodman ED, Kuhn LA, Jain AKJl-toec. Dimensionality reduction using genetic algorithms. *IEEE transactions on evolutionary computation*. 2000;4(2):164-71;Available from: <https://doi.org/10.1109/4235.850656>.
42. Lindenbaum M, Markovitch S, Rusakov DJMI. Selective sampling for nearest neighbor classifiers. *Machine learning*.2004;54(2):125-52;Available from: <https://doi.org/10.1023/B:MACH.0000011805.60520.fe>.
43. Chatterjee R, Datta A, Sanyal DK. Ensemble learning approach to motor imagery EEG signal classification. *Machine Learning in Bio-Signal Analysis and Diagnostic Imaging*: Elsevier; 2019. p. 183-208;Available from: <https://doi.org/10.1016/B978-0-12-816086-2.00008-4>.
44. Chatterjee R, Bandyopadhyay T, Sanyal DK, Guha D, editors. Comparative analysis of feature extraction techniques in motor imagery EEG signal classification. *Proceedings of First International Conference on Smart System, Innovations and Computing*; 2018: Springer;Available from: https://doi.org/10.1007/978-981-10-5828-8_8.
45. Isa NEzM, Amir A, Ilyas MZ, Razalli MS, editors. The performance analysis of K-nearest neighbors (K-NN) algorithm for motor imagery classification based on EEG signal. *MATEC web of conferences*; 2017: EDP Sciences;.
46. Hindarto H, Muntasa A, Efiyanti A, editors. Identification of ElectroEncephaloGraph signals using sampling technique and K-nearest neighbor. *Journal of Physics: Conference Series*; 2019: IOP Publishing;Available from: <https://doi.org/10.1088/1742-6596/1381/1/012009>.

Using wavelet transform for features extraction and machine learning algorithms to classify the facial expression by eeg signals

Hoang-Thuy-Tien Vo^{1,2}, Thi-Nhu-Quynh Nguyen^{1,2}, Phuoc Thanh Nguyen^{1,2}, Tuan Van Huynh^{1,2,*}



Use your smartphone to scan this QR code and download this article

ABSTRACT

The goal was to evaluate the performance of the state-of-the-art algorithms. A secondary goal was to try to improve upon the result of a method that was used in a study similar to the one used in this work. This paper presented the building of multi-state datasets relating to eye behaviors and facial expressions. Signals were recorded and stored by the connection of a channel-less mobile device. Z-score, max-min normalization techniques were used to optimize data. The cross-validation technique divided the data into training/testing segments. The features of the electrical brain signals (delta, theta, alpha and beta band) were analyzed by the Daubechies wavelet transform method. The extracted time and frequency domain features calculate total energy, detailed component energy, approximate component energy, relative energy. Three algorithms, support vector machine, k-nearest neighbor, and ensemble algorithm, were used to develop into 17 models to optimize the classification efficiency of the machine learning algorithms. Parameters of these models were surveyed and optimized to propose a best classification one for the Data-021 dataset. The Subspace ensemble model was proposed because its model efficiency was more than 87,7%.

Key words: classification, electroencephalogram, facial expression, machine learning, signal processing

¹Faculty of Physics and Engineering Physics, University of Science, Ho Chi Minh City, VietNam

²Vietnam National University, Ho Chi Minh City, Vietnam

Correspondence

Tuan Van Huynh, Faculty of Physics and Engineering Physics, University of Science, Ho Chi Minh City, VietNam

Vietnam National University, Ho Chi Minh City, Vietnam

Email: hvtuan@hcmus.edu.vn

History

- Received: 18-12-2021
- Accepted: 30-5-2022
- Published: 30-6-2022

DOI : 10.32508/stdjns.v6i2.1157



Copyright

© VNUHCM Press. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.



Cite this article : Vo H, Nguyen T, Nguyen P T, Huynh T V. Using wavelet transform for features extraction and machine learning algorithms to classify the facial expression by eeg signals. *Sci. Tech. Dev. J. - Nat. Sci.*; 2022, 6(2):2116-2130.