

ỨNG DỤNG PHƯƠNG PHÁP HỌC MÁY XÂY DỰNG MÔ HÌNH NHẬN DẠNG RUNG NHĨ

Nguyễn Trọng Tuyển^{1*}; Đào Ngọc Diệp¹

TÓM TẮT

Rung nhĩ là một trong những bệnh lý nguy hiểm hàng đầu có thể dẫn đến tử vong do đột quỵ ở các bệnh nhân mắc bệnh. Hiện nay, phương pháp chẩn đoán các bệnh lý tim mạch nói chung, trong đó có rung nhĩ, chủ yếu sử dụng máy đo điện tim Holter để nhận dạng các bất thường trong tín hiệu điện tim. Tuy nhiên, ở bệnh nhân rung nhĩ, các chu kỳ rối loạn nhịp tim có thể diễn ra ở mọi thời điểm, không thể dự đoán trước được, thậm chí có thể xảy ra vào thời điểm mà thiết bị đo Holter không ghi lại được. Xây dựng thuật toán nhận dạng và phát hiện rung nhĩ nhằm ứng dụng cho các thiết bị đo chẩn đoán rung nhĩ theo thời gian thực đóng vai trò hết sức cần thiết. Trong bài báo này, chúng tôi xây dựng mô hình nhận dạng rung nhĩ sử dụng phương pháp học máy dựa trên các bộ cơ sở dữ liệu đã được kiểm chứng. Độ chính xác trung bình khi sử dụng các mô hình học máy đạt được là 91.95%. Kết quả trên cho thấy, mô hình có thể áp dụng trong theo dõi và chẩn đoán rung nhĩ, nâng cao hiệu quả điều trị các bệnh lý tim mạch và đào tạo kỹ sư kỹ thuật y sinh.

Từ khoá: rung nhĩ, học máy.

ABSTRACT

Atrial fibrillation is one of the leading life-threatening conditions that can lead to stroke in affected patients. Currently, the diagnosis of cardiovascular diseases in general, including atrial fibrillation, primarily relies on Holter electrocardiogram devices to identify abnormalities in the electrocardiogram signal. However, in patients with atrial fibrillation, arrhythmic cycles can occur at any time, unpredictably, and may even occur during periods when the Holter device is not recording. Building an atrial fibrillation recognition and detection algorithm for real-time diagnostic devices is crucial. In this paper, we develop an atrial fibrillation recognition model using machine learning methods based on validated datasets. The average accuracy achieved using machine learning models is 91.95%. The results suggest that the model can be applied in atrial fibrillation monitoring and diagnosis, enhancing the effectiveness of cardiovascular disease treatment and training biomedical engineering engineers.

Keywords: Atrial fibrillation, machine learning

¹ Trường ĐH Cửu Long

* Người chịu trách nhiệm bài viết: Nguyễn Trọng Tuyển (Email: nguyentuyen1988@gmail.com).

1. Mở đầu

Rung nhĩ là một trong những rối loạn nhịp tim nguy hiểm hàng đầu, có thể dẫn đến tử vong ở các bệnh nhân mắc bệnh. Theo thống kê, khoảng 5% số trường hợp đột quỵ hàng năm xuất phát do rung nhĩ và tỷ lệ mắc bệnh cũng tăng dần theo độ tuổi. Nguyên nhân gây bệnh do xuất hiện các vùng hoạt động điện hỗn loạn trên tâm nhĩ, dẫn tới việc co bóp nhanh quá mức và không đều ở tâm nhĩ. Điều này dẫn đến xuất hiện cục máu đông ở tim, di chuyển theo mạch máu lên não và các cơ quan gây đột quỵ hoặc tắc nghẽn mạch máu cấp ở các cơ quan trong cơ thể. Triệu chứng của bệnh lý rung nhĩ bao gồm đánh trống ngực, giảm khả năng gắng sức, mệt mỏi, khó thở và thoáng ngất. Các triệu chứng này có thể dễ gây nhầm lẫn với các bệnh lý khác [1]. Do đó, việc chẩn đoán bệnh lý đóng vai trò hết sức quan trọng, giúp bệnh nhân được điều trị kịp thời.

Hiện nay, chẩn đoán các bệnh lý rối loạn nhịp tim được thực hiện bằng phép đo tín hiệu điện tim, sử dụng máy đo điện tim Holter. Bệnh nhân sau thăm khám có thể được bác sĩ yêu cầu ghi đo tín hiệu điện tim trong vài ngày. Tuy nhiên, đối với bệnh lý rung nhĩ, chu kỳ rối loạn nhịp tim có thể diễn ra ở bất kỳ thời điểm nào trong một khoảng thời gian theo dõi, chu kỳ này có thể diễn ra trong một khoảng thời gian ngắn, không thể dự đoán trước được. Do đó, trong nhiều trường hợp, bệnh nhân mắc rối loạn nhịp tim do rung nhĩ không thể phát hiện bằng thiết bị đo Holter, đặc biệt, trong trường hợp bệnh lý rung nhĩ kịch phát.

Cùng với sự phát triển của hạ tầng công nghệ thông tin, các hệ thống y tế từ xa ngày nay càng được sử dụng phổ biến hơn. Các thiết bị đo điện tim, chẩn đoán bệnh lý theo thời gian thực đang ngày càng được đưa vào sử dụng giúp bệnh nhân chẩn đoán rối loạn nhịp tim

tại nhà. Điều này giúp bệnh nhân được chẩn đoán, giảm thiểu khả năng phát triển bệnh lý và được điều trị một cách kịp thời. Một số thiết bị đo cảm tay hiện nay đang được phát triển bởi một số hãng như Apple Watch, Samsung Watch, AliveCor, Zenicor, Omron giúp bệnh nhân nhanh chóng bệnh lý rung nhĩ giúp thu đo tín hiệu điện tim và đưa ra chẩn đoán rối loạn nhịp tim trong thời gian ngắn [2]. Tuy nhiên, đa phần các thiết bị này, vẫn chưa đem lại hiệu quả cao do ảnh hưởng bởi nhiễu, thời gian đo tín hiệu ngắn, nên chưa thể phát hiện được các bất thường trong mọi thời điểm.

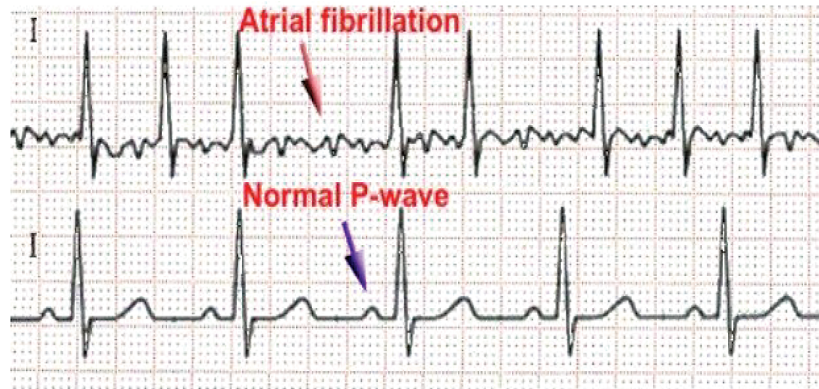
Do đó, việc phát triển thuật toán nhận dạng rối loạn nhịp tim do rung nhĩ là điều cần thiết. Thuật toán có thể được ứng dụng trong xử lý tín hiệu trong bản ghi tín hiệu điện tim của thiết bị đo điện tim Holter, giúp bác sĩ có thể rút ngắn thời gian chẩn đoán. Đồng thời, việc áp dụng thuật toán cũng có thể thực hiện bằng cách nhúng mô hình nhận dạng tới các thiết bị thu đo điện tim thời gian thực giúp bệnh nhân và bác sĩ có thể theo dõi và phát hiện kịp thời bệnh lý. Từ đó, bệnh nhân có thể được điều trị với phác đồ thích hợp. Trong nghiên cứu này, chúng tôi sử dụng thuật toán xử lý tín hiệu điện tim, dựa trên các đặc trưng của bệnh lý rung nhĩ. Sau đó, các đặc trưng được sử dụng như dữ liệu đầu vào để xây dựng mô hình nhận dạng ứng dụng các phương pháp học máy khác nhau.

2. Xây dựng mô hình nhận dạng rung nhĩ dựa trên các đặc trưng tín hiệu ECG

Để chẩn đoán rung nhĩ bằng tín hiệu điện tim ECG, bác sĩ thường dựa trên các đặc trưng lâm sàng của tín hiệu điện tim ở trong chu kỳ rối loạn nhịp tim. Khi đó, trên tín hiệu điện tim có thể nhận thấy các đặc điểm đặc trưng như sau. Thứ nhất, tính không đồng đều của nhịp co bóp tâm thất, hay sự không đồng đều giữa

các đoạn RR trong khoảng tín hiệu theo dõi. Thứ hai, đó là sự biến mất của sóng P trên tín hiệu điện tim và thay vào đó là sự xuất hiện

của các sóng nhĩ f có biên độ và tần số khác nhau, với biểu hiện rõ ràng nhất là trên các đạo trình II, III, aVR, V1, V2 (Hình 1).



Hình 1. Tín hiệu điện tim trong trường hợp rung nhĩ và bình thường

Dựa trên đặc điểm lâm sàng của bệnh lý như trên, chúng tôi xây dựng bộ cơ sở dữ liệu các đặc trưng bằng phương pháp xử lý tín hiệu số tín hiệu điện tim. Đối với mỗi đoạn tín hiệu với độ dài 10 giây, sơ đồ thuật toán sau dùng để tính toán các đặc trưng tín hiệu điện tim được thể hiện như trên hình 2.

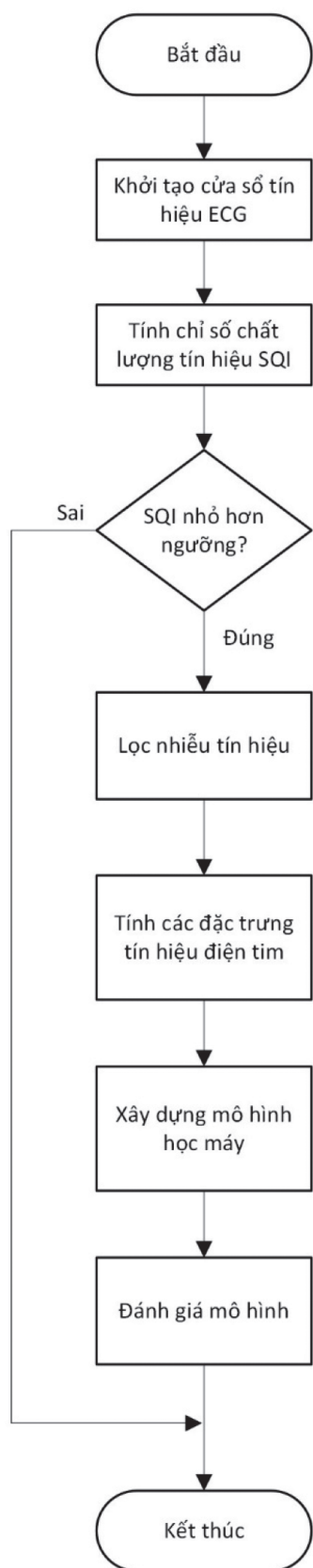
Tín hiệu điện tim trong quá trình đo thường bị ảnh hưởng bởi nhiễu từ các nguồn nhiễu khác nhau. Một trong số các nguồn nhiễu ảnh hưởng đáng kể đến quá trình đo tín hiệu điện tim là nhiễu chuyển động của người đo, gây trôi đường cơ sở của tín hiệu điện tim. Ngoài ra là các nhiễu tần số cao cũng ảnh hưởng đáng kể đến phân tích tín hiệu điện tim. Do đó, trong các hệ thống thu nhận và phân tích, xử lý tín hiệu điện tim, đầu tiên, cần tiến hành lọc các loại nhiễu kể trên. Trong nghiên cứu này, cần đánh giá mức độ nhiễu thông qua chỉ số chất lượng tín hiệu. Khi đó, nếu chỉ số chất lượng tín hiệu dưới ngưỡng đặt trước nào đó, các bộ lọc nhiễu trôi đường cơ sở và nhiễu lọc tần số cao được áp dụng. Sau khi lọc nhiễu, thực hiện tính các đặc trưng của tín hiệu điện tim ECG, được đưa vào như dữ liệu đầu vào của mô hình học máy.

Các đặc trưng được tính toán trong nghiên cứu này, được xác định dựa trên các đặc trưng lâm sàng bệnh lý của tín hiệu điện tim trong trường hợp rối loạn nhịp tim do rung nhĩ, bao gồm: các đặc trưng của biến thiên nhịp tim (HRV) và đặc trưng biến thiên đoạn TQ của tín hiệu điện tim. Để đưa ra các đặc trưng kể trên, trong một số trường hợp, cần loại bỏ các phức bộ dị thường trên tín hiệu điện tim, đáng chú ý có thể kể đến là các phức bộ ngoại tâm thu thất và ngoại tâm thu nhĩ thường xuất hiện ở bệnh nhân mắc rung nhĩ.

Các đặc trưng biến thiên nhịp tim bao gồm: độ rộng HRV dRR , giá trị trung bình RR_{mean} và độ phân tán HRV $disRR$ được tính theo công thức sau:

$$\begin{cases} dRR = RR_{max} - RR_{min}; \\ RR_{mean} = \sum RR / N_B; \\ disRR = \sqrt{\sum (RR - RR_{mean})^2 / (N_B - 1)}, \end{cases}$$

Trong đó, RR_{max} , RR_{min} , tương ứng là các giá trị lớn nhất và nhỏ nhất của chuỗi các khoảng RR tín hiệu điện tim trên cửa sổ đang theo dõi, N_B là số chu kỳ nhịp tim sau khi loại bỏ các phức bộ dị thường.



Hình 2. Sơ đồ thuật toán xử lý tín hiệu điện tim trong nhận dạng rung nhĩ

Để phát hiện sự xuất hiện của sóng nhĩ f trên tín hiệu ECG, chúng tôi thực hiện đồng bộ hóa các khoảng TQ (khoảng TQ là đoạn tín hiệu tính từ điểm kết thúc của sóng T và điểm bắt đầu của sóng Q trên tín hiệu điện tim) của tín hiệu điện tim trong cửa sổ theo dõi. Ở nghiên cứu này, sử dụng phương pháp nội suy theo độ dài lớn nhất các khoảng TQ. Để đánh giá biến thiên khoảng TQ trong hai trường hợp tín hiệu điện tim với bệnh lý rung nhĩ và tín hiệu điện tim bình thường, nghiên cứu đề xuất đánh giá khoảng cách Euclidean giữa các khoảng TQ của cửa sổ theo dõi với khoảng TQ trung bình. Đầu tiên, cần tính các chỉ số khoảng TQ trung bình và biến thiên khoảng TQ.

$$\begin{cases} \overline{TQ} = \frac{\sum_{i=1}^{N_B} TQ_i}{N_B}; \\ De = \frac{\|TQ_j - \overline{TQ}\|}{\|TQ_j\|}. \end{cases}$$

Tỉ số De thường được sử dụng để đánh giá độ lệch của đại lượng vật lý so với giá trị trung bình của chúng trong một tập hợp. Khi giá trị De tăng lên thể hiện độ lệch của khoảng TQ đang xét với giá trị khoảng TQ trung bình, nghĩa là tăng độ biến thiên của khoảng TQ trong cửa sổ tín hiệu điện tim đang theo dõi. Để đưa ra đặc trưng cho chu kỳ rung nhĩ, tham số sau được tính.

$$VarDe = \frac{N_{De \geq D_f}}{N_B},$$

Trong đó, $N_{De \geq D_f}$ là số các giá trị De vượt quá giá trị ngưỡng D_f , N_B là số chu kỳ nhịp tim sau khi loại bỏ các phức bộ dị thường.

Phương pháp tính đặc trưng trên có một số ưu điểm như độ phức tạp tính toán thấp (low-cost computing), giảm ảnh hưởng của nhiễu

đến quá trình đánh giá độ biến thiên khoảng TQ dựa trên việc tính giá trị trung bình khoảng TQ và chuẩn hóa giá trị khoảng theo dõi.

Để xây dựng mô hình nhận dạng rung nhĩ, chúng tôi sử dụng bộ cơ sở dữ liệu được kiểm chứng là bộ cơ sở dữ liệu MIT BIH Atrial Fibrillation Database gồm 25 bản ghi tín hiệu ECG của các bệnh nhân mắc rung nhĩ. Mỗi bản ghi có thời gian ghi tín hiệu trong vòng 10 giờ và chứa 02 tín hiệu ECG, được lấy mẫu với tần số 250Hz và độ phân giải 12 bit. Mỗi bản ghi đều được đánh dấu điểm bắt đầu chu kỳ rối loạn nhịp tim do rung nhĩ (AF) và chu kỳ tín hiệu điện tim bình thường (N).

Trong nghiên cứu này, sử dụng 4150 đoạn tín hiệu có dán nhãn AFIB và 3790 đoạn tín hiệu được dán nhãn N. Các đoạn tín hiệu này có độ dài 10 giây. Chúng tôi xây dựng mạng neuron tuần tự 08 lớp, bao gồm 01 lớp đầu vào, 06 lớp ẩn và 01 lớp đầu ra. Cụ thể, lớp đầu vào gồm 3 neuron tương ứng với 3 mẫu đầu bao gồm dRR, dispRR, VarDe và sử dụng hàm ReLU là hàm kích hoạt, do các giá trị đầu vào đều nằm trong khoảng 0 và 1 với số lượng mẫu tương đối lớn. Tiếp theo xây dựng mạng neuron gồm 6 lớp ẩn, mỗi lớp ẩn có 11 đơn vị. Hàm kích hoạt cho mỗi lớp ẩn là hàm ReLU giúp đẩy nhanh tốc độ xử lý. Cuối cùng lớp đầu ra có 1 lớp ẩn tương ứng dự đoán dương tính hoặc âm tính với bệnh lý rung nhĩ. Hàm sigmoid được sử dụng là hàm kích hoạt của lớp đầu ra.

Sau khi huấn luyện mô hình, chúng tôi sử dụng hai tham số để đánh giá mô hình là độ chính xác (Acc, Accuracy) và F1-Score. Kết quả đánh giá mô hình mạng neuron cho giá trị các tham số Acc và F1-Score lần lượt là 91.91% và 91.29%. Ngoài ra, trong nghiên cứu còn đánh giá mô hình dựa trên kiểm chứng

chéo, thu được các tham số Acc và F1-Score sử dụng kiểm chứng chéo là 91.90% và 91.40%.

Để so sánh với các mô hình học máy khác, ở đây, chúng tôi sử dụng các mô hình cây quyết định và k láng giềng gần nhất để so sánh. Kết quả đối chiếu với mô hình mạng neuron được đưa ra dưới bảng sau.

Bảng 1. Kết quả huấn luyện sử dụng các mô hình học máy khác nhau

Mô hình	Độ chính xác (Acc)	F1-Score
Mạng Neuron	91.91%	91.29%
Cây quyết định	91.86%	91.24%
K-láng giềng	92.09%	91.5%

Các mô hình học máy xây dựng trong nghiên cứu này có độ chính xác tương đối tốt so với các mô hình nhận dạng khác như mô hình của Babaeizadeh [3] với độ chính xác khoảng 93%, Lake và Moorman [4] với độ chính xác khoảng 92.5%.

3. Phát triển hướng nghiên cứu

Hướng nghiên cứu tiếp theo của nhóm tác giả là xây dựng thêm các đặc trưng tín hiệu ECG trong miền tần số, nhằm nhận dạng và phân loại chu kỳ rung nhĩ trên nền nhiễu và các chu kỳ rối loạn nhịp tim do cuồng nhĩ và tín hiệu điện tim bình thường. Ngoài ra, các tác giả cũng xem xét, biến thiên nhịp tim như một quá trình hỗn loạn, hay một hệ động lực phi tuyến. Từ đó, có thể bổ sung thêm các đặc trưng của hệ phi tuyến như phân tích fractal hoặc entropy hệ thống, nhằm nâng cao độ chính xác của thuật toán nhận dạng.

Ngoài ra, các tác giả cũng định hướng xây dựng mô hình học sâu nhằm phát hiện rối loạn nhịp tim do rung nhĩ trong thời gian cực ngắn (xấp xỉ 01 chu kỳ co bóp). Đây là các chu kỳ có khả năng phát triển thành các

rối loạn nhịp tim rất lớn và khó phát hiện. Từ đó, đưa ra được các dự đoán chính xác và hiệu quả hơn.

4. Kết luận

Trong nghiên cứu này, các tác giả đề xuất mô hình nhận dạng rối loạn nhịp tim do rung nhĩ dựa trên các đặc trưng của tín hiệu điện tim ở bệnh nhân. Mô hình được xây dựng cho kết quả tương đối tốt, thích hợp để tích hợp vào các hệ thống nhận dạng rối loạn nhịp tim do rung nhĩ theo thời gian thực. Điều này cho phép bệnh nhân nhanh chóng phát hiện được chu kỳ rung nhĩ dựa trên các thiết bị đo điện tim tại nhà. Từ đó nhanh chóng liên hệ với bác sĩ, nhằm kịp thời thăm khám, đưa ra phác đồ điều trị đối với bệnh nhân dựa trên tình trạng của bệnh nhân. Mô hình cũng rất hữu ích trong quá trình đào tạo những chuyên viên kỹ thuật của ngành kỹ thuật y sinh hiện nay.

TÀI LIỆU THAM KHẢO

- [1] January, Craig T., et al. "2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society." *Journal of the American College of Cardiology* 64.21 (2014): e1-e76.
- [2] Phong, Phan Đình, et al. "Bước đầu đánh giá khả năng nhận biết rung nhĩ bởi trí tuệ nhân tạo Kardia Alivecor Sử dụng thiết bị ghi điện tâm đồ lưu động." *Tạp chí Tim mạch học Việt Nam* 90 (2019): 170-175.
- [3] Babaeizadeh, Saeed, et al. "Improvements in atrial fibrillation detection for real-time monitoring." *Journal of electrocardiology* 42.6 (2009): 522-526.
- [4] Lake, Douglas E., and J. Randall Moorman. "Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices." *American Journal of Physiology-Heart and Circulatory Physiology* 300.1 (2011): H319-H325.