

NGHIÊN CỨU PHƯƠNG PHÁP LOẠI BỎ NHIỀU DO CHUYỂN ĐỘNG TRONG TÍN HIỆU PPG SỬ DỤNG AUTOENCODER

RESEARCH ON MONTION ARTIFACTS REMOVAOL IN PPG SIGNAL USING AUTOENCODER

Lưu Thanh Tùng, Trần Quang Đức Hòa, Trần Vũ Anh Duy

Bộ môn Kỹ thuật Máy Xây dựng và Nâng chuyển, Khoa Cơ Khí, Trường Đại học Bách Khoa,
Đại học Quốc gia Thành phố Hồ Chí Minh

Email: tthuu@hcmut.edu.vn

TÓM TẮT

Công nghệ đo quang thể tích (Photoplethysmography - PPG) sử dụng ánh sáng để phát hiện các thay đổi nhỏ trong thể tích máu ngoại vi, dựa trên khả năng hấp thụ ánh sáng khác nhau của các thành phần mô tuân theo định luật Beer-Lambert. Mặc dù có nhiều ứng dụng tiềm năng trong y tế, tín hiệu PPG rất nhạy cảm với nhiễu do chuyển động (Motion Artifacts - MA) của người dùng, gây khó khăn cho việc thu thập và phân tích dữ liệu. Nghiên cứu này đề xuất một phương pháp ứng dụng trí tuệ nhân tạo (Deep Learning), cụ thể là mô hình mạng nơ-ron Autoencoder, để tự động phát hiện và hỗ trợ loại bỏ các đoạn tín hiệu PPG bị ảnh hưởng bởi MA. Thuật toán được đánh giá trên bộ dữ liệu thực tế thu thập tại Bệnh viện Thống Nhất, đạt độ chính xác 95,2% trong việc phân loại đoạn tín hiệu sạch và nhiễu, cho thấy tiềm năng ứng dụng trong các thiết bị theo dõi sức khỏe sử dụng cảm biến PPG.

Từ khóa: PPG; Cảm biến; Autoencoder; Photoplethysmography.

ABSTRACT

Photoplethysmography (PPG) is a technology employing light to measure subtle variations in peripheral blood volume, based on the Beer-Lambert law regarding differential light absorption by tissue components. Despite its potential medical applications, PPG signals are highly susceptible to motion artifacts (MA) caused by user movement, challenging accurate data acquisition and analysis. This study proposes an approach using deep learning (DL), specifically an Autoencoder neural network model, to automatically detect and aid in the removal of MA-corrupted PPG signal segments. Evaluated on a real-world dataset collected from Thong Nhat Hospital, the algorithm achieved 95.2% accuracy in classifying clean versus noisy signal segments, demonstrating its potential applicability in PPG-based health monitoring devices.

Keywords: PPG; Sensor; Autoencoder; Photoplethysmography.



1. GIỚI THIỆU

Công nghệ đo quang thể tích (Photoplethysmography – PPG) hoạt động dựa trên định luật Beer-Lambert, cho phép đo lường sự thay đổi thể tích máu bằng cách chiếu ánh sáng đơn sắc vào mô và đo cường độ ánh sáng phản xạ hoặc truyền qua [1]. Ban đầu, PPG được ứng dụng chủ yếu trong các thiết bị đo nhịp tim và độ bão hòa oxy trong máu (SpO₂) như Pulse Oximeter.

Với sự phát triển của khoa học máy tính và trí tuệ nhân tạo (AI), việc phân tích các đặc trưng bậc một và bậc hai của tín hiệu PPG đã mở rộng ứng dụng sang các lĩnh vực phức tạp hơn như phát hiện rối loạn nhịp tim, ước tính nồng độ glucose máu, và nhiều thông số sinh lý khác [2].

Tuy nhiên, một trong những thách thức lớn nhất của công nghệ PPG là sự nhạy cảm với nhiễu, đặc biệt là nhiễu do chuyển động của người dùng (Motion Artifacts - MA) [3]. Các chuyển động này làm thay đổi vị trí giữa cảm biến và da, gây nhiễu tín hiệu. Việc sử dụng các đoạn tín hiệu chứa MA có thể dẫn đến sai số lớn trong tính toán hoặc thậm chí gây chẩn đoán sai, đặc biệt khi sử dụng các mô hình AI. Mặc dù còn các nguồn nhiễu khác như: sai vị trí đo, do biên độ tín hiệu hoặc do hiện tượng co thắt tâm thất sớm gây gián đoạn nhịp tim nhưng nguyên nhân chính và thường gặp nhất là do MA.

Nhiều phương pháp đã được nghiên cứu để xử lý MA, bao gồm các kỹ thuật dựa trên chỉ số chất lượng tín hiệu (Signal Quality Index - SQI) hoặc so sánh với tín hiệu mẫu [4]. Tuy nhiên, chưa có phương pháp nào được coi là tối ưu cho mọi ứng dụng. Trong bối cảnh đó, nghiên cứu này đề xuất một hướng tiếp cận mới

sử dụng mô hình Autoencoder, một dạng mạng nơ-ron sâu, để nhận diện và hỗ trợ loại bỏ các đoạn tín hiệu PPG bị ảnh hưởng bởi MA.

2. CƠ SỞ LÝ THUYẾT

2.1. Tổng quan về nhiễu do chuyển động (Motion Artifact Noise)

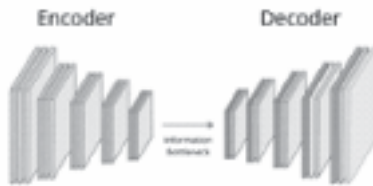
Nhiễu do chuyển động (MA) là sự biến đổi ảo trong tín hiệu PPG, phát sinh khi có sự chuyển động tương đối giữa cảm biến và bề mặt da của đối tượng đo. Các hoạt động như đi bộ, cử động tay, hoặc thậm chí gõ ngón tay có thể gây ra MA. Đặc điểm của MA là các dao động thường có biên độ lớn, không đều (có thể tuần hoàn hoặc không tuần hoàn), và có thể che lấp hoàn toàn sóng mạch đập (pulse waveform) vốn có của tín hiệu PPG sạch.

Một thách thức cốt lõi trong việc loại bỏ MA là phổ tần số của nó thường chồng lấp lên phổ tần số của các tín hiệu sinh lý quan trọng như nhịp tim (thường 0,5 – 3,5 Hz) và nhịp thở (thường 0,13 – 0,67 Hz) [5]. Sự chồng lấp này làm cho các bộ lọc tần số tuyến tính đơn giản (như lọc thông dải) trở nên kém hiệu quả. Do đó, cần có các phương pháp xử lý tín hiệu tiên tiến hơn, có khả năng phân biệt và tách biệt tín hiệu sinh lý khỏi nhiễu MA ngay cả khi chúng cùng tồn tại trong một dải tần số.

2.2. Tổng quan về Autoencoder

Autoencoder là một mạng nơ-ron đặc biệt được huấn luyện để cố gắng làm cho dữ liệu đầu ra giống với dữ liệu đầu vào nhất có thể [6]. Kiến trúc của một mạng autoencoder bao gồm: Bộ mã hóa (Encoder), không gian tiềm ẩn (Latent space) hay còn gọi là nút thắt (bottleneck), và một bộ giải mã (decoder). Nguyên lý hoạt động của mạng nơ-ron này

sẽ như sau: Khi một dữ liệu đầu vào mạng autoencoder, nó sẽ bị mã hóa trở thành một biểu diễn tiềm ẩn có số chiều thấp hơn, sau đó từ nút thắt cổ chai ấy nó sẽ được giải mã trở thành dữ liệu đầu ra.



Hình 1. Cấu trúc Autoencoder

Bộ mã hóa có nhiệm vụ mã hóa dữ liệu đầu vào, giảm số chiều của dữ liệu bằng cách nén chúng lại, giữ lại các đặc trưng quan trọng của dữ liệu. Dữ liệu nén sẽ được biểu diễn ở vùng không gian tiềm ẩn, nơi có số chiều nhỏ nhất. Từ đó, bộ giải mã sẽ cố gắng tái tạo lại dữ liệu đầu vào ban đầu một cách chính xác nhất có thể. Autoencoder thường được huấn luyện bằng cách sử dụng học không giám sát (unsupervised learning) với mục tiêu là giảm thiểu sai số giữa đầu vào và đầu ra nhất có thể, thường được đo bằng một hàm mất mát như sai số toàn phương trung bình (Mean square error - MSE) [7].

Với n là số lượng điểm dữ liệu, y_i là giá trị thực tế của điểm dữ liệu thứ i và $\hat{y}_i$ là giá trị dự đoán của điểm dữ liệu thứ i , ta có phương trình của MSE:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

Trong quá trình huấn luyện, các trọng số của autoencoder được tối ưu hóa thông qua thuật toán gradient descent để giảm thiểu sai số tái tạo.

3. PHƯƠNG PHÁP NGHIÊN CỨU

3.1. Dữ liệu sử dụng trong nghiên cứu

Có hai bộ dữ liệu được sử dụng trong luận văn gồm bộ dữ liệu mở MIMICIII [8] và bộ dữ liệu thu thập từ các bệnh nhân ở phòng khám tim mạch của bệnh viện Thống Nhất (Số 1 Lý Thường Kiệt, phường 7, quận Tân Bình, TP. Hồ Chí Minh). Bộ dữ liệu MIMICIII được sử dụng để phát triển thuật toán cả xử lý tín hiệu và phát hiện rối loạn nhịp tim dựa trên tín hiệu PPG, trong khi đó bộ dữ liệu từ Bệnh viện Thống Nhất sẽ được dùng để kiểm tra kết quả của thuật toán.

Chúng tôi sử dụng phương pháp 5-fold cross validation, nghĩa là chia dữ liệu thành 5 phần bằng nhau và tiến hành 5 lần huấn luyện, mỗi lần huấn luyện là một fold, trong đó chọn bốn phần để làm dữ liệu huấn luyện mới trong vòng 80 epoch, phần còn lại sẽ là tập thử nghiệm và đảm bảo rằng tập thử nghiệm là khác nhau giữa các fold.

3.2. Thuật toán tiền xử lý tín hiệu PPG

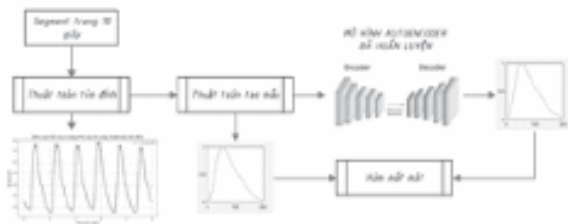
Quy trình tiền xử lý tín hiệu PPG thô và tạo mẫu đầu vào cho mô hình Autoencoder bao gồm các bước sau:

- Đầu tiên, lọc nhiễu bằng bộ lọc thông dải (band – pass filter) Butterworth bậc 4 với tần số từ 0,5 Hz đến 3 Hz với mục đích là loại bỏ các thành phần nhiễu tần số thấp (baseline drift) và nhiễu tần số cao, giữ lại các thông tin quan trọng.

- Chia đoạn tín hiệu thành các đoạn với độ dài 10 giây vì nó chứa đủ thông tin về nhiều chu kỳ tim [9].



- Xác định đỉnh của tín hiệu PPG trong mỗi đoạn segment 10 giây, sau đó sử dụng thuật toán tạo mẫu để chuẩn bị dữ liệu đầu vào cho mô hình Autoencoder đã huấn luyện. Các mẫu được coi là “sạch” nếu giá trị hàm mất mát tái tạo “reconstruction loss” nhỏ hơn ngưỡng cho phép. Những đoạn tín hiệu có ít nhất 95,2% mẫu đạt chuẩn được giữ lại. Quy trình này được minh họa chi tiết ở hình dưới.



Hình 2. Quy trình nghiên cứu

3.3. Thuật toán tìm đỉnh và tạo mẫu của tín hiệu PPG

Thuật toán tìm đỉnh tham khảo từ [10], sử dụng ngưỡng động MA_{beat} – chu kỳ nhịp tim trung bình và a là hằng số bù để lọc nhiễu và xác định đỉnh là điểm có giá trị lớn nhất trong mỗi khối tín hiệu hợp lệ. Phương pháp này tự động thích nghi với biến đổi tín hiệu, giảm sai số do nhiễu và trôi đường cơ sở (baseline drift), đảm bảo phát hiện đỉnh chính xác ngay cả trong điều kiện nhiễu.

$$THR_1 = MA_{beat} + a \quad (2)$$

Đường bao dưới (lower envelope) của tín hiệu PPG được xác định bằng nội suy spline. Tín hiệu sau đó được tách thành các nhịp, tái mẫu thành 100 điểm, đưa về mức không (zero-baselined) và chuẩn hóa.

Công thức thuật toán đưa về không: Với y là tín hiệu đã được đưa về không, y_0 là tín hiệu gốc, t là thời gian (đối với tín hiệu liên tục)

và chỉ số (đối với tín hiệu đã được rời rạc hóa), y_1 là giá trị tín hiệu tại đỉnh bắt đầu.

$$y = y_0 - k.t - y_1 \quad (3)$$

Công thức chuẩn hóa: Trong đó y là tín hiệu cần được chuẩn hóa, $\hat{y}$ là tín hiệu đã được chuẩn hóa, $\max\{y\}$ là giá trị lớn nhất của tín hiệu, $\min\{y\}$ là giá trị nhỏ nhất của tín hiệu.

$$\hat{y} = \frac{y - \min\{y\}}{\max\{y\} - \min\{y\}} \quad (4)$$

3.4. Mô hình autoencoder về phát hiện đoạn tín hiệu bị nhiễu do cử động

Mô hình autoencoder có thể được sử dụng để phát hiện sự bất thường của dữ liệu vì mô hình này có khả năng biểu diễn dữ liệu trên một “không gian” khác được gọi là không gian tiềm ẩn (latent space), với số chiều ít hơn và có tính phân cụm cao hơn. Trong không gian này, các dữ liệu có cùng tính chất sẽ có phân phối gần nhau, hay có thể hiểu là nằm chung một cụm. Dựa trên tính chất này, việc xác định tín hiệu đạt chuẩn được tiến hành bằng cách huấn luyện mô hình với dữ liệu đạt chuẩn. Sau khi hoàn tất quá trình huấn luyện, mô hình sẽ phát hiện các dữ liệu không đạt chuẩn dựa trên giá trị hàm mất mát giữa tín hiệu tái tạo và tín hiệu đầu vào. Những dữ liệu đạt chuẩn sẽ có giá trị hàm mất mát thấp và gần bằng với giá trị hàm mất mát trên tập huấn luyện, còn các dữ liệu không đạt chuẩn sẽ có giá trị hàm mất mát lớn.

Tuy nhiên, câu hỏi về việc giá trị hàm mất mát lớn bao nhiêu là đủ để phân loại tín hiệu chưa được đề cập cụ thể trong các nghiên cứu trước đây và thường phụ thuộc vào từng trường hợp cụ thể. Để giải quyết vấn đề này, chúng tôi chia tập huấn luyện thành hai phần: phần để huấn luyện và phần để kiểm tra. Sau

khi hoàn tất việc huấn luyện, giá trị ngưỡng được xác định bằng cách so sánh giới hạn bao của phân phối giá trị hàm mất mát của tập huấn luyện và tập kiểm tra. Nếu hai phân phối này tương đồng nhau, giá trị bao trung bình của cả hai sẽ được dùng làm giá trị ngưỡng để đánh giá tín hiệu đạt chuẩn và không đạt chuẩn.

Mô hình autoencoder sử dụng để xác định nhiễu do chuyển động sử dụng hàm mất mát MSE (1) để huấn luyện và có cấu trúc như sau:

STT	Lớp		Kích thước
1	Encoder	Đầu vào	100
		Hàm ReLU	
2		Lớp 1	64
		Hàm ReLU	
3		Lớp 2	32
		Hàm ReLU	
5	Lớp ẩn		16
7	Decoder	Lớp 5	32
		Hàm ReLU	
8		Lớp 6	64
		Hàm ReLU	
9		Đầu ra	100
		Hàm Sigmoid	

4. KẾT QUẢ VÀ THẢO LUẬN

Giá trị hàm mất mát trong quá trình huấn luyện được thể hiện như ở dưới đây:



Hình 3. Giá trị hàm mất mát

Trong quá trình huấn luyện, giá trị hàm mất mát giảm dần và đạt mức ổn định ở các epoch sau. Giá trị hàm mất mát cuối cùng khá nhỏ, cho thấy tín hiệu đầu vào và tín hiệu tái tạo ở đầu ra của mô hình có độ tương đồng cao. Điều này chứng tỏ mô hình autoencoder đã trích xuất tốt các đặc trưng của tín hiệu PPG sạch.

Khi kiểm tra mô hình trên tập dữ liệu độc lập từ Bệnh viện Thống Nhất, kết quả cho thấy khả năng phân loại hiệu quả giữa các đoạn tín hiệu sạch và nhiễu. Các đoạn tín hiệu được xác định là sạch (đạt chuẩn) được đánh dấu bằng màu xanh. Đường mức màu cam thể hiện việc chọn và loại bỏ tín hiệu, với mức cao tương ứng với việc chọn giữ lại tín hiệu và mức thấp tương ứng với việc loại bỏ tín hiệu.



Hình 4. Kết quả thuật toán trên dữ liệu của Bệnh viện Thống Nhất

Kết quả cho thấy thuật toán có khả năng phát hiện chính xác các đoạn tín hiệu bị nhiễu do chuyển động, với độ chính xác đạt 95,2% trong việc phân loại giữa đoạn tín hiệu sạch và nhiễu. Tuy nhiên, vẫn còn một số trường hợp thuật toán nhận diện sai, đặc biệt là khi nhiễu có biên độ thấp hoặc có đặc tính gần giống với tín hiệu sinh lý.

5. KẾT LUẬN

Trong nghiên cứu này, chúng tôi đã ứng dụng mô hình autoencoder để phát hiện và loại bỏ nhiễu do chuyển động trong tín hiệu PPG. Kết quả thực nghiệm cho thấy phương pháp đề xuất đạt độ chính xác 95,2% trong việc phân loại giữa tín hiệu sạch và nhiễu. Mặc dù còn



một số trường hợp nhận diện sai, những hạn chế này có thể được cải thiện trong tương lai thông qua việc tăng cường dữ liệu huấn luyện. Kết quả nghiên cứu khẳng định tiềm năng ứng dụng của autoencoder trong việc xử lý nhiễu do chuyển động cho các thiết bị theo dõi sức khỏe sử dụng cảm biến PPG.

Lời cảm ơn:

Nghiên cứu này được tài trợ bởi Trường Đại học Bách Khoa, Đại học Quốc gia Thành phố Hồ Chí Minh trong khuôn khổ đề tài mã số SVOISP-2024 – CK-29. Chúng tôi xin cảm ơn Trường Đại học Bách Khoa, ĐHQG – HCM đã hỗ trợ cho nghiên cứu này. ❖

Ngày nhận bài: **28/3/2025**

Ngày phản biện: **14/4/2025**

Tài liệu tham khảo:

- [1]. Allen, J. (2007), “*Photoplethysmography and its application in clinical physiological measurement*”. *Physiological Measurement*, 28(3), R1-R39.
- [2]. Shelley, K. H. (2007), “*Photoplethysmography: Beyond the calculation of arterial oxygen saturation and heart rate*”. *Anesthesia & Analgesia*, 105(6 Suppl), S31-S36.
- [3]. Foo, J. Y. A., & Wilson, S. J. (2006), “*A computational system to optimise noise rejection in photoplethysmography signals during motion or poor perfusion states*”. *Medical & Biological Engineering & Computing*, 44(1-2), 140-145.
- [4]. Zhang, Z., Pi, Z., & Liu, B. (2015), “*TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise*”. *IEEE Transactions on Biomedical Engineering*, 62(2), 522-531.
- [5]. Zong, W., Heldt, T., Moody, G. B., & Mark, R. G. (2003), “*An open-source algorithm to detect onset of arterial blood pressure pulses*”. *Computers in Cardiology*, 30, 259-262.
- [6]. Goodfellow, I., Bengio, Y., & Courville, A. (2016), “*Deep Learning*”. MIT Press.
- [7]. IBM. (n.d.). Autoencoder. IBM Think Topics.
- [8]. Johnson, A. E. W., Pollard, T. J., Shen, L., Lehman, L. H., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Celi, L. A., & Mark, R. G. (2016), “*MIMIC-III, a freely accessible critical care database*”. *Scientific Data*, 3, 160035.
- [9]. Elgendi, M. (2012), “*On the analysis of fingertip photoplethysmogram signals*”. *Current Cardiology Reviews*, 8(1), 14-25.
- [10]. Elgendi, M. (2012), “*On the analysis of fingertip photoplethysmogram signals*”. *Current Cardiology Reviews*, 8(1), 14-25.