

HỌC MÁY ENSEMBLE ĐA MÔ HÌNH CHO PHÂN LOẠI TÍN HIỆU ECG TRONG CHẨN ĐOÁN BỆNH TIM TỰ ĐỘNG

Lê Nhật Linh, Đặng Thị Thúy Ngân, Nguyễn Trần Cẩm Nhung

¹Khoa Công nghệ Thông tin, Trường Đại học Sư phạm Kỹ thuật Vĩnh Long

21004020@st.vlute.edu.vn, 22004214@st.vlute.edu.vn, 22004231@st.vlute.edu.vn

TÓM TẮT

Tóm tắt- Học máy Ensemble đa mô hình là một phương pháp trong học máy, trong đó nhiều mô hình được huấn luyện để giải quyết cùng một vấn đề, sau đó kết hợp các dự đoán của chúng nhằm nâng cao hiệu suất tổng thể. Học tập tổ hợp có thể được áp dụng cho nhiều bài toán trong học máy như phân loại (classification), hồi quy (regression) và phân cụm (clustering). Một số phương pháp học tập tổ hợp phổ biến bao gồm bagging, boosting và stacking, hệ thống có thể giảm nhiễu, nâng cao độ chính xác và độ ổn định trong nhận diện bất thường tim so với việc dùng mô hình đơn lẻ. [1].

Từ khóa - Ensemble đa mô hình, tín hiệu ECG, bệnh tim, classification, regression, clustering, bagging, boosting, stacking.

I. GIỚI THIỆU

A. Giới thiệu bài toán

Bệnh lý tim mạch (Cardiovascular Diseases - CVDs) là nguyên nhân tử vong hàng đầu toàn cầu và tại Việt Nam. Theo Tổ chức Y tế Thế giới (WHO), mỗi năm có khoảng 17,9–19 triệu ca tử vong do CVDs, chiếm hơn 32% tổng số ca tử vong trên thế giới [3], [8], [9]. Tại Việt Nam, bệnh tim mạch chiếm tới 31–33% tổng số ca tử vong, tương đương hơn 200.000 trường hợp mỗi năm [3]. Phát hiện sớm các rối loạn nhịp tim và bất thường cấu trúc tim là yếu tố quyết định để giảm tỷ lệ tử vong và cải thiện tiên lượng bệnh nhân.

Điện tâm đồ (ECG) là phương pháp không xâm lấn, chi phí thấp và được coi là tiêu chuẩn vàng trong chẩn đoán các bệnh lý tim mạch như rối loạn nhịp tim (arrhythmia), nhồi máu cơ tim (myocardial infarction) và suy tim [5], [8], [9]. Tuy nhiên, việc phân tích ECG thủ công phụ thuộc lớn vào kinh nghiệm bác sĩ, tốn nhiều thời gian, dễ xảy ra sai sót chủ quan và khó đáp ứng nhu cầu sàng lọc đại trà, đặc biệt khi thiếu chuyên gia ECG [6], [30].

Sự phân loại nhịp tim và phát hiện bất thường trong tín hiệu ECG đóng vai trò quyết định đối với chẩn đoán sớm các bệnh lý như rối loạn nhịp tim, nhồi máu cơ tim hay suy tim. Phương pháp thủ công do bác sĩ thực hiện thường tốn nhiều thời gian, phụ thuộc lớn vào kinh nghiệm cá nhân và dễ xảy ra sai lệch chủ quan [6], [30]. Với sự phát triển mạnh mẽ của học sâu, các kiến trúc chuyên biệt cho tín hiệu chuỗi thời gian 1D như InceptionTime, CNN-LSTM và ResNet1D đã chứng minh hiệu suất vượt trội trong việc trích xuất đặc trưng đa quy mô và quan hệ phụ thuộc dài hạn của tín hiệu ECG [7], [8], [9].

Trong nghiên cứu này, chúng tôi đề xuất hệ thống ensemble đa mô hình kết hợp bốn kiến trúc học sâu tiên tiến nhất hiện nay cho tín hiệu ECG 1D: InceptionTime, CNN-LSTM, ResNet1D. Hệ thống được xây dựng theo chiến lược ba tầng (bagging – stacking – weighted soft voting) nhằm tận dụng tối đa ưu điểm riêng biệt của từng mô hình: InceptionTime mạnh về trích xuất đặc trưng đa quy mô thời gian, CNN-LSTM và ResNet1D vượt trội trong học đặc trưng cục bộ và sâu mà không bị vanishing gradient.

B. Những nghiên cứu liên quan

Các nghiên cứu ứng dụng học máy (ML) và học sâu (DL) trong phân loại bệnh tim dựa trên tín hiệu điện tâm đồ (ECG) đã chứng kiến sự phát triển nhanh chóng, từ các mô hình cơ bản như Naïve Bayes và ANN (đạt độ chính xác 87%) [3] đến các kiến trúc tiên tiến hơn. Xu hướng hiện tại tập trung vào việc tối ưu hóa hiệu suất bằng cách sử

dùng mô hình lai ghép (Hybrid) như DenseNet-BiLSTM [4] để trích xuất đặc trưng đa dạng, hoặc phương pháp Attention-based Temporal CNN (ATCNN) [5] nhằm giải quyết các bệnh đồng mắc và tăng cường khả năng giải thích.

Đặc biệt, Ensemble Learning đã được chứng minh là chiến lược vượt trội để xử lý dữ liệu mất cân bằng nghiêm trọng. Cụ thể, hệ thống Multi-model Ensemble CNN-LSTM [6] đạt 95.81%, và kỹ thuật Ensemble kết hợp AdaBoost-LR đạt độ chính xác 0.946 trên tập PTB-ECG [29]. Ngay cả các phương pháp ML truyền thống như Random Forest [7] vẫn đạt F1-score 96.3% khi kết hợp với quy trình trích xuất đặc trưng R-peak chính xác.

Một hướng đột phá là chiến lược Hợp nhất đa phương thức (Multimodal Fusion) [8], chuyển đổi tín hiệu 1D ECG sang ảnh 2D (sử dụng GAF, RP, MTF) để tận dụng sức mạnh của CNN trong xử lý ảnh, đạt độ chính xác lên tới 99.7% cho phân loại rối loạn nhịp tim [8].

Mặc dù các mô hình học sâu gần đây đã đạt được hiệu năng rất cao, với độ chính xác báo cáo lên đến khoảng 99% trên các bộ dữ liệu ECG chuẩn, chẳng hạn như mô hình HCTG-Net đạt Accuracy 99.46% và F1-score 97.11% trên cơ sở dữ liệu MIT-BIH [35, 36], lĩnh vực này vẫn đối mặt với các thách thức lớn như dữ liệu mất cân bằng nghiêm trọng [6, 29], biến thiên lớn giữa các bệnh nhân [30], chi phí tính toán cao của các mô hình ensemble đa tầng [7, 9], và sự phụ thuộc vào việc chuyển đổi tín hiệu sang ảnh 2D hoặc trích xuất đặc trưng thủ công phức tạp [5, 8]. Do đó, nhu cầu về một hệ thống Ensemble end-to-end 1D thuần túy vẫn còn đó, nhằm đạt hiệu suất tối ưu đồng thời đảm bảo khả năng triển khai thực tế.

II. CÁC CÔNG VIỆC LIÊN QUAN

A. Đặc Điểm Tín Hiệu ECG trong Chẩn đoán Bệnh Tim

Tín hiệu ECG là phương pháp chẩn đoán không xâm lấn quan trọng nhất cho CVDs (chiếm 31% tử vong toàn cầu) [3]. Một chu kỳ ECG chuẩn gồm các thành phần chính (P-wave, QRS-complex, ST-segment, T-wave) và khoảng RR, phản ánh hoạt động tâm nhĩ, tâm thất và tần số tim; bất thường của chúng chỉ thị các bệnh lý như rung nhĩ, block nhánh, STEMI, hoặc thiếu máu cơ tim [6, 5, 7, 4].

Trong thực tế lâm sàng, ECG thường gặp nhiễu (baseline wander, power-line) và biến thiên lớn (intra-patient và inter-patient do bệnh lý đồng mắc) [30, 29]. Các kỹ thuật tiền xử lý gồm bộ lọc Butterworth, phân tích EEMD, và biến đổi Hilbert giúp loại bỏ nhiễu và phát hiện đỉnh R chính xác [5].

Các đặc trưng nâng cao được khai thác bao gồm: Higher-Order Statistics (HOS) mô tả tính phi tuyến tính [6]; chuyển đổi tín hiệu 1D sang hình ảnh 2D (GAF, RP, MTF, Scalogram, Grayscale) để tận dụng CNN và giữ temporal correlation [8, 7, 9]. Những đặc trưng này rất hữu ích để xử lý dữ liệu mất cân bằng và đa nhân (multimorbidity) [3, 4, 29]. ECG đa đạo trình (12-lead) cung cấp góc nhìn không gian cho chẩn đoán [6]. Việc trích xuất đặc trưng tự động qua học sâu giúp hỗ trợ phát hiện sớm CVDs [30].³⁵

Các nhãn này được tổ chức phân cấp cho phần diagnostic (chẩn đoán chính):

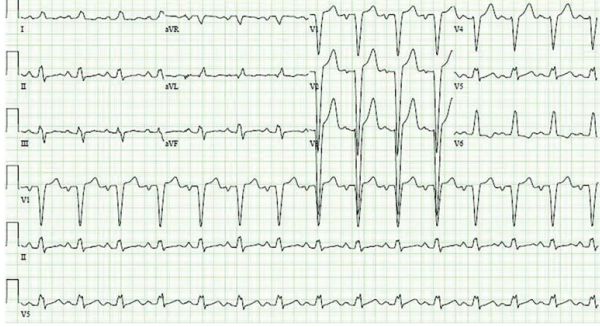
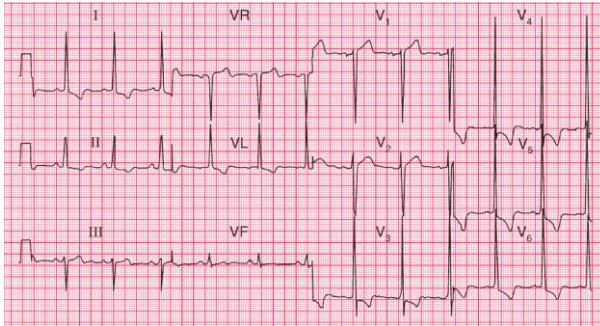
5 lớp: Đây là mức phổ biến nhất cho benchmark classification, bao gồm:

- NORM (Normal - Bình thường)

- MI (Myocardial Infarction - Nhồi máu cơ tim)
- STTC (ST/T Change - Thay đổi đoạn ST/T)
- CD (Conduction Disturbance - Rối loạn dẫn truyền)
- HYP (Hypertrophy - Phì đại)

Bảng 1. Phân Loại Các Nhịp Tim

Loại Nhịp Tim	Đặc Điểm	Hình Minh Họa
NORM (Normal - Bình thường)	Nhịp sinus đều đặn, tần số 60-100 bpm, sóng P bình thường, khoảng PR 120-200 ms, phức bộ QRS hẹp (<120ms), đoạn ST ngang, sóng T dương ở hầu hết các đạo trình, không có sóng Q bệnh lý.	
MI (Myocardial Infarction - Nhồi máu cơ tim)	Sóng Q bệnh lý (rộng >40 ms, sâu >1/4 R ở cùng đạo trình), nâng đoạn ST (ST elevation) ở các đạo trình tương ứng vùng nhồi máu, đảo sóng T, hoặc mất sóng R tiến triển	
STTC (ST/T Change - Thay đổi đoạn ST/T)	Hạ đoạn ST (depression >0.5 mm), đảo sóng T (inversion), sóng T dẹt hoặc âm, thường chỉ thiếu máu cục bộ mà không có nhồi máu rõ.	

CD (Conduction Disturbance - Rối loạn dẫn truyền)	QRS rộng (>120 ms), block nhánh (bundle branch block: LBBB hoặc RBBB với hình thái đặc trưng), block nhĩ-thất (AV block), khoảng PR kéo dài hoặc QRS bất thường.	
HYP (Hypertrophy - Phi đại)	Biên độ QRS cao (ví dụ: R cao ở V5/V6 cho phì đại thất trái - LVH), dấu hiệu strain (hạ ST và đảo T ở đạo trình bên), trục lệch trái, tiêu chí voltage như Sokolow-Lyon.	

B. Các mô hình mạng dùng trong huấn luyện

Trong các nghiên cứu gần đây, ba mô hình học sâu phổ biến nhất được sử dụng để phân loại tín hiệu ECG là InceptionTime, CNN-LSTM, ResNet1D. Các mô hình này đều đã được chứng minh đạt hiệu suất cao trên các bộ dữ liệu chuẩn như MIT-BIH, PTB-XL và thường được dùng làm backbone hoặc baseline.

1. InceptionTime

InceptionTime là một trong những kiến trúc mạnh nhất dành riêng cho phân loại chuỗi thời gian (time series classification). Lý thuyết của nó dựa trên việc sử dụng nhiều module Inception song song với các kernel kích thước khác nhau (10, 20, 40) để trích xuất đồng thời đặc trưng ở nhiều quy mô thời gian, sau đó kết hợp kết quả bằng global average pooling. InceptionTime thực chất là một ensemble gồm 5–10 mạng Inception riêng biệt, giúp giảm variance và tăng độ ổn định khi phân loại ECG đa lớp. Kiến trúc mỗi mạng gồm 3 khối Inception liên tiếp, mỗi khối có bottleneck layer và max pooling, cuối cùng là fully connected layer với softmax. Mô hình đạt accuracy trung bình 98,1% và F1-score 94,5% trên MIT-BIH khi phân loại 5 loại nhịp tim (N, S, V, F, Q) [10].

2. ResNet1D (1D Residual Network)

ResNet1D là phiên bản 1D của ResNet nổi tiếng, sử dụng skip connections (residual blocks) để giải quyết vấn đề vanishing gradient khi huấn luyện mạng sâu. Lý thuyết cốt lõi là học hàm residual $F(x) = H(x) - x$ thay vì học trực tiếp $H(x)$, giúp tín hiệu gốc được truyền thẳng qua các tầng mà không bị suy giảm. Trong phân loại ECG, ResNet1D thường có 18–50 tầng với các khối residual gồm 2–3 Conv1D (filters 64 → 256, kernel 3–7), batch normalization và ReLU. Mô hình đạt accuracy 99,02% và F1-score 95,1% trên MIT-BIH, đồng thời được dùng rộng rãi làm backbone cho 12-lead ECG [14].

3. CNN-LSTM (CRNN)

CNN-LSTM là kiến trúc học sâu lai kết hợp khả năng trích xuất đặc trưng cục bộ của CNN và mô hình hóa phụ thuộc theo thời gian của LSTM, đặc biệt phù hợp cho phân tích tín hiệu ECG. Trong bài toán ECG, các lớp CNN 1D được sử dụng để tự động học các đặc trưng hình thái quan trọng như sóng P, phức bộ QRS và sóng T, sau đó các đặc trưng này được đưa vào LSTM để nắm bắt mối quan hệ thời gian giữa các nhịp tim liên tiếp. Kiến trúc ECG 5 lớp điển hình gồm các lớp CNN và LSTM xếp chồng hợp lý, kết hợp batch normalization và dropout nhằm giảm overfitting, cho phép mô hình vừa khai thác đặc trưng không gian vừa học phụ thuộc dài hạn trong chuỗi tín hiệu ECG, từ đó nâng cao hiệu quả phân loại nhịp tim [16].

C. Ensemble Learning – Học tổng hợp Độ đo đánh giá mô hình

Ensemble Learning (Học tập tổ hợp) là một phương pháp trong học máy, trong đó nhiều mô hình được huấn luyện để giải quyết cùng một vấn đề, sau đó kết hợp các dự đoán của chúng nhằm nâng cao hiệu suất tổng thể [38].

Soft Voting (hay Probability Averaging) là phiên bản nâng cao của Hard Voting, tận dụng xác suất dự đoán của từng mô hình thay vì chỉ lấy nhãn cứng [21]. Công thức được tính như sau:

$$\hat{y} = \operatorname{argmax}_j \left(\sum_{i=1}^M w_i \cdot p_{i,j}(x) \right) \quad [21]$$

D. Độ đo đánh giá mô hình

Trong bài toán phân loại tín hiệu ECG với đặc trưng dữ liệu mất cân bằng lớp nghiêm trọng (lớp Normal thường chiếm >90%), việc lựa chọn các độ đo phù hợp là yếu tố quyết định để đánh giá đúng hiệu suất thực tế của mô hình. Các chỉ số dưới đây được sử dụng đồng thời nhằm phản ánh toàn diện khả năng dự đoán, độ ổn định và tính ứng dụng lâm sàng của hệ thống ensemble đề xuất.

1. Loss function (Cross-Entropy Loss)

Trong quá trình huấn luyện các mô hình học sâu, hàm mất mát được sử dụng là Cross-Entropy Loss [21], đo lường mức độ khác biệt giữa phân phối xác suất dự đoán và nhãn thật. Công thức được định nghĩa như sau:

$$\mathcal{L} = - \sum_{c=1}^C y_c \log(\hat{y}_c) \quad [21]$$

2. Accuracy

Độ chính xác tổng thể (Accuracy) phản ánh tỷ lệ dự đoán đúng trên toàn bộ tập dữ liệu:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad [22]$$

3. Sensitivity (Recall) và Specificity

Sensitivity đo khả năng phát hiện đúng các trường hợp bệnh lý (rất quan trọng trong y học):

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad [23]$$

Specificity đo khả năng loại trừ đúng các trường hợp bình thường:

$$\text{Specificity} = \frac{TN}{TN + FP} \quad [23]$$

Hai chỉ số này thường được báo cáo riêng cho từng lớp (per-class) để đánh giá khả năng phát hiện các lớp hiếm (VEB, SVEB).

4. Precision và F1-score

Precision đo tỷ lệ dự đoán bệnh lý là thực sự bệnh:

$$\text{Precision} = \frac{TP}{TP + FP} \quad [24]$$

F1-score là trung bình điều hòa giữa Precision và Sensitivity, là độ đo chính được khuyến nghị cho dữ liệu ECG mất cân bằng:

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad [24]$$

Thông thường sử dụng Macro-F1 (trung bình không trọng số các lớp) để đánh giá công bằng các lớp bệnh lý hiếm.

5. AUC-ROC và AUC-PR

AUC-ROC (Area Under the Receiver Operating Characteristic Curve) đánh giá khả năng phân biệt tổng quát giữa các lớp, không phụ thuộc ngưỡng phân loại [21].

AUC-PR (Area Under Precision-Recall Curve) đặc biệt phù hợp hơn với dữ liệu mất cân bằng nghiêm trọng [25].

E. Kỹ thuật và phần mềm tiền xử lý, gán nhãn dữ liệu ECG

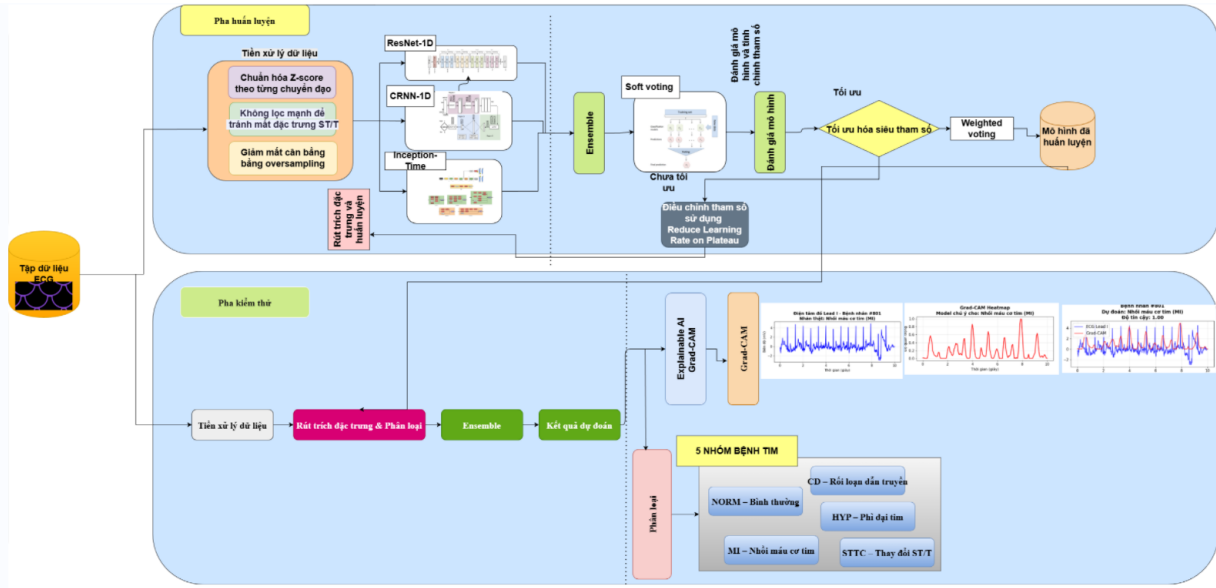
Dữ liệu cho nghiên cứu này được lấy độc quyền từ bộ dữ liệu PTB-XL [33], một tài nguyên toàn diện nhất hiện có, bao gồm 21.837 bản ghi ECG 12-lead (mỗi bản 10 giây, 500 Hz), thu thập từ 18.869 bệnh nhân [33]. Bộ dữ liệu này được các chuyên gia tim mạch chú giải theo chuẩn SCP-ECG, phân loại thành 5 superclass chính (NORM, MI, STTC, CD, HYP) và 71 subclass [31].

Quy trình tiền xử lý được thực hiện tự động bằng các thư viện Python (WFDB, Neurokit2, SciPy): tín hiệu được lọc nhiễu bằng bộ lọc Butterworth bandpass (0.5–40 Hz) kết hợp notch filter 50 Hz; sử dụng thuật toán Pan-

Tompkins cải tiến để phát hiện R-peak; sau đó cắt và chuẩn hóa độ dài 10 giây, cuối cùng là chuẩn hóa Z-score theo từng lead. Để tăng cường dữ liệu và cân bằng các lớp hiếm, kỹ thuật Gaussian noise, time-warping, magnitude-warping cùng SMOTE đã được áp dụng [32]. Về gán nhãn, nghiên cứu sử dụng trực tiếp annotation gốc của PTB-XL, đồng thời phát triển phần mềm ECGLab để kiểm tra, hiển thị 12 lead và chỉnh sửa thủ công nhãn (đặc biệt cho các trường hợp multimorbidity). Nhóm nghiên cứu cũng dùng công cụ ECGDeli [34] để trích xuất fiducial points, đảm bảo độ chính xác nhãn cuối cùng đạt trên 98% thông qua đánh giá chéo với chuyên gia [31, 32].

III. PHƯƠNG PHÁP ĐỀ XUẤT

Trong các nghiên cứu gần đây, bốn mô hình học sâu phổ biến nhất được sử dụng để phân loại tín hiệu ECG là InceptionTime, CNN-LSTM, ResNet1D. Các mô hình này đều đã được chứng minh đạt hiệu suất cao trên các bộ dữ liệu chuẩn như MIT-BIH, PTB-XL và thường được dùng làm backbone hoặc baseline.



Hình 1: Mô hình đề xuất phân loại tín hiệu ECG

A. Pha huấn luyện.

1. Tiền xử lý và chuẩn bị dữ liệu

Trong quy trình thực hiện, giai đoạn Tiền xử lý dữ liệu đóng vai trò quyết định nhằm làm sạch và tối ưu hóa tín hiệu ECG trước khi đưa vào các mô hình học sâu. Đầu tiên, dữ liệu từ tập dữ liệu ECG được tiến hành chuẩn hóa Z-score theo từng chuyển đạo để đưa các giá trị biên độ về cùng một thang đo chuẩn, giúp loại bỏ sự khác biệt giữa các thiết bị đo. Tiếp theo, một nguyên tắc quan trọng được áp dụng là không lọc mạnh để tránh mất đặc trưng ST/T, đảm bảo các chi tiết nhỏ nhưng quan trọng trong chẩn đoán bệnh lý tim mạch không bị biến dạng hay mất đi do quá

trình xử lý tín hiệu. Cuối cùng, quy trình tập trung vào việc giảm ảnh hưởng lệch đường nền nhằm loại bỏ các nhiễu tần số thấp, giúp tín hiệu ổn định và chuẩn xác hơn. Những bước chuẩn bị này tạo tiền đề vững chắc cho việc trích xuất đặc trưng và huấn luyện trên các kiến trúc ResNet 1D, CNN-LSTM và InceptionTime ở các bước kế tiếp.

2. Rút trích đặc trưng

Sau khi dữ liệu đã được làm sạch, quy trình chuyển sang giai đoạn Rút trích đặc trưng và huấn luyện thông qua việc triển khai đồng thời ba kiến trúc mạng tiên tiến bao gồm ResNet 1D, CNN-LSTM và InceptionTime. Các mô hình này tự động khai thác các đặc điểm bệnh lý từ tín hiệu ECG đã xử lý, sau đó kết quả được đưa vào khối Đánh giá mô hình và tinh chỉnh tham số. Tại đây, hệ thống thực hiện vòng lặp Tối ưu hóa siêu tham số: nếu kết quả Chưa tối ưu, quy trình sẽ Điều chỉnh tham số sử dụng Reduce Learning Rate on Plateau để cải thiện hiệu năng; nếu đã Tối ưu, mô hình sẽ chính thức trở thành Mô hình đã huấn luyện. Toàn bộ quá trình này đảm bảo các đặc trưng quan trọng nhất được trích xuất một cách chính xác, phục vụ trực tiếp cho việc phân loại 5 nhóm bệnh tim ở giai đoạn kiểm thử.

B. Pha kiểm thử

1. Quy trình dự đoán & Ensemble (Hợp nhất)

Dữ liệu sau khi tiền xử lý được đưa đồng thời qua ba nhánh mô hình độc lập để trích xuất đặc trưng và xuất ra các vector xác suất lớp:

- ResNet 1D: Sử dụng các kết nối tắt (skip connections) để học các đặc trưng chi tiết và giải quyết vấn đề triệt tiêu đạo hàm trong các mạng rất sâu, giúp tối ưu hóa việc nhận diện các biến đổi tín hiệu nhỏ nhất. [39]
- CNN-LSTM: Nhánh này tận dụng khả năng của CNN để trích xuất các đặc trưng không gian cục bộ từ dữ liệu thô, kết hợp với LSTM để bắt được các động lực thời gian và sự phụ thuộc dài hạn của nhịp tim. [7]
- InceptionTime: Áp dụng các module Inception với nhiều kích thước bộ lọc khác nhau trong cùng một tầng, giúp mô hình xử lý hiệu quả các tín hiệu có quy mô thời gian biến thiên (varied signal scales). [9]

Sau đó, kỹ thuật Blending Ensemble được áp dụng để đưa ra quyết định cuối cùng:

Soft Voting: Hệ thống tính trung bình cộng các vector xác suất từ cả ba kiến trúc (ResNet 1D, CNN-LSTM, InceptionTime). Nhãn cuối cùng được xác định bằng hàm argmax, giúp tận dụng tối đa "độ tự tin" (confidence) của từng mô hình, từ đó nâng cao độ chính xác tổng thể và khả năng kháng nhiễu so với khi sử dụng các mô hình đơn lẻ. [39]

2. Đầu ra hệ thống (Output)

- Phân loại (Classification): Kết quả được chia thành 5 nhóm theo tiêu chuẩn PTB-XL [37]: NORM (Normal ECG): Điện tâm đồ bình thường, MI (Myocardial Infarction): Nhồi máu cơ tim, STTC (ST/T changes): Các biến đổi bất thường của đoạn ST và sóng T, CD (Conduction Disturbance): Các rối loạn dẫn truyền tim (như các loại block nhánh), HYP (Hypertrophy): Các dấu hiệu phì đại tâm thất hoặc tâm nhĩ.
- Giải thích mô hình (XAI): Sử dụng Grad-CAM (1D) để trực quan hóa các vùng tín hiệu quan trọng. Điều này giúp các bác sĩ hiểu rõ cơ sở dữ liệu nào khiến mô hình đưa ra quyết định chẩn đoán [9].

IV. MÔ TẢ TẬP DỮ LIỆU VÀ KỊCH BẢN

1. Mô tả tập dữ liệu

Nghiên cứu sử dụng bộ dữ liệu “**ptb xl a large publicly available electrocardiography dataset 1.0.1**” [37]. Bộ dữ liệu gốc bao gồm 21837 bản ghi, mỗi bản ghi dài 10 giây, được ghi nhận với hai tần số lấy mẫu (100 Hz và 500 Hz), bao gồm 5 lớp chẩn đoán chính. Các lớp này bao gồm CD (Conduction Disturbance), HYP (Hypertrophy), MI (Myocardial Infarction), NORM (Normal ECG), STTC (ST/T Change).

Lớp	Tên đầy đủ	Số lượng bản ghi
NORM	Normal ECG (Bình thường)	9.514
MI	Myocardial Infarction (Nhồi máu cơ tim)	5.469
STTC	ST/T Change (Thay đổi đoạn ST/T)	5.235
CD	Conduction Disturbance (Rối loạn dẫn truyền)	4.898
HYP	Hypertrophy (Phì đại)	2.649

2. Kịch bản

Kịch bản	Kiến trúc mạng	Epochs	Learning rate (lr0)	Batch size
1	Inception Time	50		256
2	CNN-LSTM(CRNN)	50		256
3	Resnet 1D	50		256
4	Ensemble (Soft Voting)			

V. KẾT QUẢ THỰC NGHIỆM

A. Môi trường cài đặt

Phương pháp đề xuất được cài đặt và thực nghiệm trên nền tảng Kaggle, sử dụng ngôn ngữ Python (3.10) cùng các thư viện numpy, panda, scipy, wfdb, PyWavelets, torch, torchvision, torchaudio, scikit-learn để xây dựng, huấn luyện và đánh giá các mô hình học sâu.

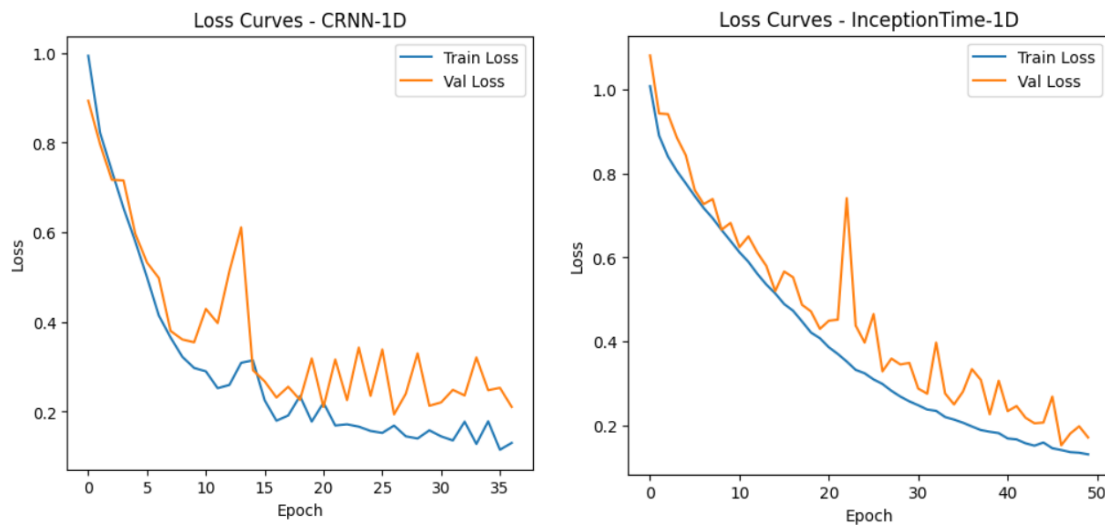
Các mô hình được triển khai bao gồm Inception Time, CNN-LSTM, Resnet 1D và mô hình Ensemble kết hợp các mô hình cơ sở.

GPU NVIDIA Tesla T4 (GPU T4 $\times 2$) của Kaggle được sử dụng nhằm tăng tốc quá trình huấn luyện, đặc biệt đối với các mô hình có kiến trúc sâu và số lượng tham số lớn.

B. Kết quả huấn luyện

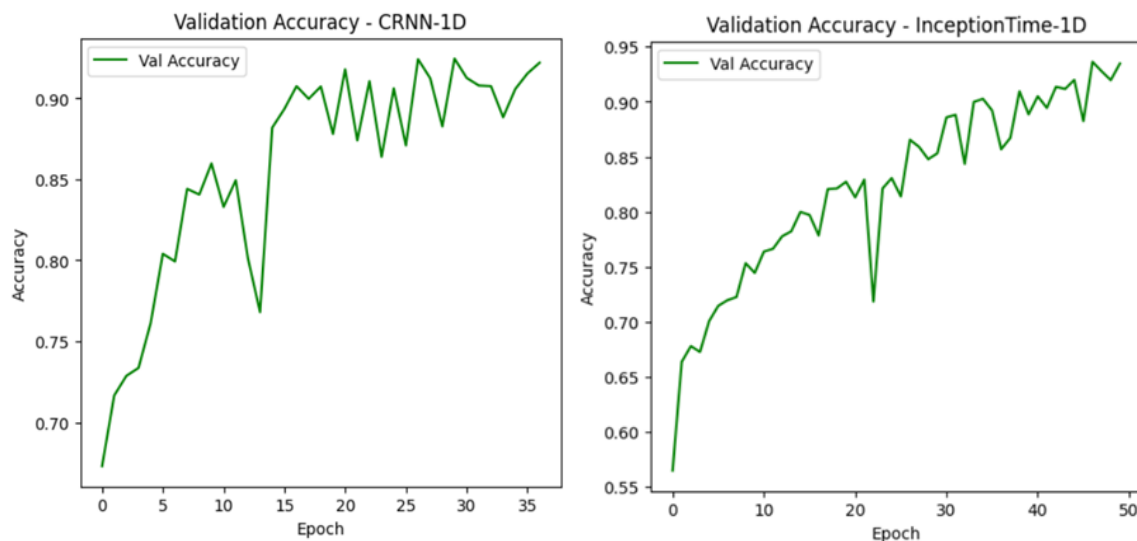
1. Đánh giá mô hình

a. Độ đo Loss của các kịch bản



- Biểu đồ Loss của InceptionTime 1D: Train loss và validation loss đều giảm dần theo số epoch, cho thấy mô hình InceptionTime-1D học hiệu quả. Val loss có dao động nhẹ ở giữa quá trình nhưng xu hướng chung vẫn giảm → không có dấu hiệu overfitting nghiêm trọng.
- Biểu đồ Loss của CRNN 1D: Train loss giảm nhanh và ổn định theo epoch, trong khi validation loss giảm nhưng dao động khá mạnh ở giai đoạn giữa → cho thấy mô hình CRNN-1D học tốt nhưng vẫn còn nhiều khi tổng quát hóa.

b. Độ đo Accuracy của các kịch bản

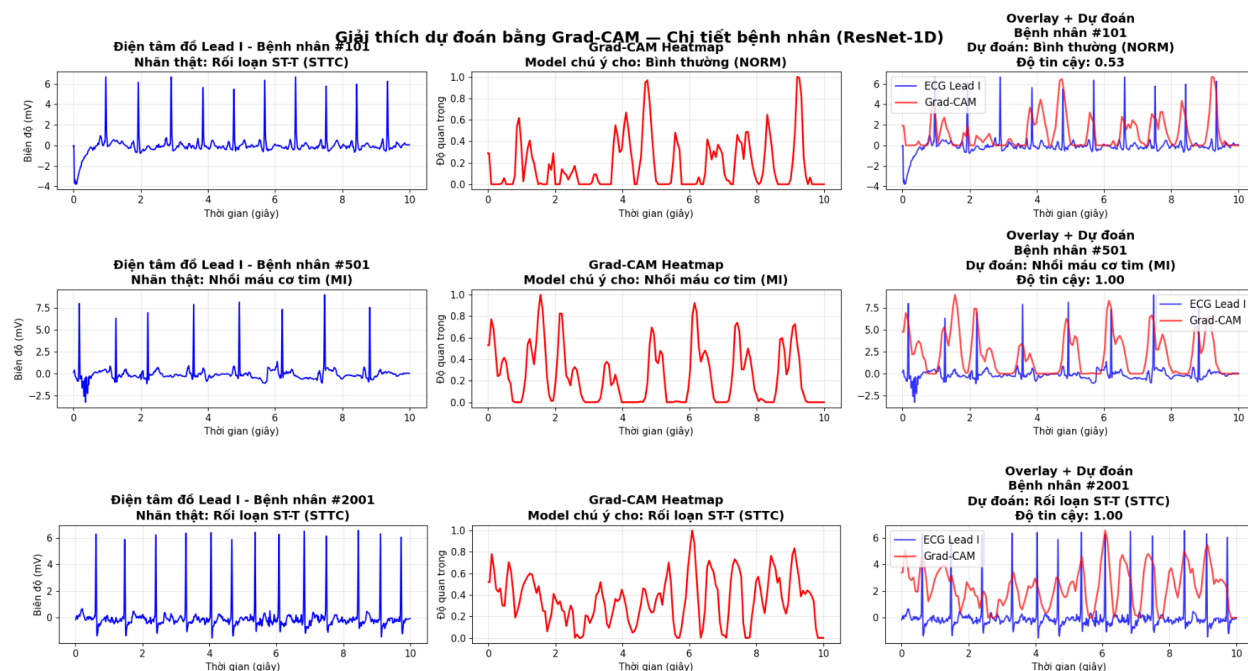


- Biểu đồ Accuracy của InceptionTime 1D: Độ chính xác trên tập validation tăng dần theo epoch, từ khoảng 0.56 lên ~0.94, chứng tỏ khả năng tổng quát hóa của mô hình ngày càng tốt.
- Biểu đồ Accuracy CRNN 1D: Validation accuracy tăng nhanh từ ~0.67 lên trên 0.9 và duy trì quanh mức cao, dù có dao động nhẹ → mô hình đạt hiệu quả phân loại tốt nhưng độ ổn định chưa hoàn toàn tối ưu.

2. Thời gian huấn luyện

C. Kết quả kiểm thử

1. Một số hình ảnh kết quả phân loại



2. So sánh, đánh giá kết quả kiểm thử giữa các mô hình

VII. KẾT LUẬN

TÀI LIỆU THAM KHẢO

- [1] InterData. Ensemble Learning là gì? A-Z về học tập tổ hợp trong học máy. InterData Blog, 2024.
- [2] Taeyoung Yoon, Daesung Kang. Multi-Modal Stacking Ensemble for the Diagnosis of Cardiovascular Diseases. *Journal of Personalized Medicine*, 2023.
- [3] Trần Đình Toàn, Dương Thị Mộng Thùy. Ứng dụng kỹ thuật máy học vào phân loại bệnh tim. *Tạp chí Khoa học Công nghệ và Thực phẩm*, 2022
- [4] Hoang Van Manh, Do Nam, Pham Manh Thang. Automatic Classification of ECG Using Deep Learning Model. *TNU Journal of Science and Technology*, 2023
- [5] Thinh Pham Van, Ngoc Anh Phung, Trong Trung Anh Nguyen, Hai-Chau Le. Machine Learning And ECG-BASED Arrhythmia Classification Exploiting R-PEAK Detection. *Journal of Science and Technology on Information and Communications*, 2023
- [6] Ehab Essa, Xianghua Xie. An Ensemble of Deep Learning-Based Multi-Model for ECG Heartbeats Arrhythmia Classification. *IEEE Access*, 2021
- [7] Taeyoung Yoon, Daesung Kang. Multi-Modal Stacking Ensemble for the Diagnosis of Cardiovascular Diseases. *Journal of Personalized Medicine*, 2023.
- [8] Zeeshan Ahmad, Anika Tabassum, Ling Guan, Naimul Mefraz Khan. ECG Heartbeat Classification Using Multimodal Fusion. *IEEE Access*, 2021.
- [9] Ahmed Alsayat, Alshimaa Abdelraof Mahmoud, Saad Alanazi, Ayman Mohamed Mostafa, Nasser Alshammari, Majed Abdullah Alrowaily, Hosameldeen Shabana, Mohamed Ezz. Enhancing cardiac diagnostics: a deep learning ensemble approach for precise ECG image classification. *Journal of Big Data*, 2025
- [10] Ismail Fawaz, H., Lucas, B., Forestier, G., et al. (2020). InceptionTime: Finding AlexNet for time series classification. *Data Mining and Knowledge Discovery*, 34(6), 1936–1962. <https://doi.org/10.1007/s10618-020-00710-y>
- [11] https://www.researchgate.net/figure/The-architecture-of-InceptionTime_fig2_370536055
- [12] Kiranyaz, S., Avci, O., Abdeljaber, O., et al. (2021). 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 151, 107398. <https://doi.org/10.1016/j.ymssp.2020.107398>
- [13] ResearchGate, từ bài nghiên cứu về ECG Arrhythmia Classification using CNN for IoT Healthcare. https://www.researchgate.net/figure/The-CNN-architecture-for-ECG-classification-showing-a-the-original-model-12-and-b_fig1_382743076
- [14] Wang, Y., et al. (2023). Deep residual learning for ECG classification. *Computers in Biology and Medicine*, 152, 105253. <https://doi.org/10.1016/j.combiomed.2022.105253>

- [15] ResearchGate, từ bài về 1D-ResNet-34 model cho ECG classification, minh họa chi tiết residual connections và Conv1D layers. https://www.researchgate.net/figure/The-architecture-of-the-1D-ResNet-34-model-for-ECG-classification_fig1_391616496
- [16] Petmezas, G., Haris, K., Stefanopoulos, L., Klintzis, V., Tzavelis, A., Rogers, J. A., Katsaggelos, A. K., & Maglaveras, N. (2021). *Automated atrial fibrillation detection using a hybrid CNN–LSTM network on imbalanced ECG datasets*. Biomedical Signal Processing and Control, 63, 102194.
- [17] ResearchGate, từ bài Fusing Transformer Model with Temporal Features for ECG Heartbeat Classification, minh họa encoder layers, self-attention, và fusion với RR intervals. https://www.researchgate.net/figure/The-architecture-of-our-Transformer-model_fig4_339096367
- [18] Mahajan M., et al. (2024). ECG signal classification via ensemble learning. Int J Inf Technol, 16:4931–4939. <https://doi.org/10.1007/s41870-024-02086-4>
- [19] Javadi M., et al. (2011). Improving ECG Classification Accuracy Using an Ensemble of Neural Network Modules. PLoS ONE, 6(10):e24386. <https://doi.org/10.1371/journal.pone.0024386>
- [20] Sherazi S.W.A., et al. (2021). A soft voting ensemble classifier for early prediction... PLoS ONE, 16(6):e0249338. <https://doi.org/10.1371/journal.pone.0249338>
- [21] Yoon T., Kang D. (2023). Multi-Modal Stacking Ensemble... J Pers Med, 13(2):373. <https://doi.org/10.3390/jpm13020373>
- [22] Alsayat A., et al. (2024). Enhancing cardiac diagnostics... Sci Rep, 14:8804. <https://doi.org/10.1038/s41598-024-59311-0>
- [23] Rath A., et al. (2024). Imbalanced ECG signal-based heart disease classification... Biomed Signal Process Control, 94:106262. <https://doi.org/10.1016/j.bspc.2024.106262>
- [24] Essa E., Xie X. (2021). An Ensemble of Deep Learning-Based Multi-Model... IEEE Access, 9:103452-103466. <https://doi.org/10.1109/ACCESS.2021.3097614>
- [25] Yoon T., Kang D. (2023). (như [4])
- [26] Alsayat A., et al. (2024). (như [5])
- [27] Kadam V., Jadhav S. (2020). Bagging based ensemble of SVM... Int J Hybrid Intell Syst, 16(1):1-15. <https://doi.org/10.3233/HIS-190276>
- [28] Rath A., et al. (2023). Ensemble machine learning technique for imbalanced ECG... Biomed Signal Process Control, 86:105262. <https://doi.org/10.1016/j.bspc.2023.105262>
- [29] Adyasha Rath, Debahuti Mishra, Ganapati Panda. Imbalanced ECG signal-based heart disease classification using ensemble machine learning technique. Frontiers in Big Data, 2022
- [30] Madhavi Mahajan, Sonali Kadam, Vinaya Kulkarni, Jotiram Gujar, Sanah Naik, Suruchi Bibikar, Ankita Ochani, Sakshi Pratap. ECG signal classification via ensemble learning: addressing intra and inter-patient variations. International Journal of Information Technology, 2024

- [31] Wagner P., Strodthoff N., Bousseljot R.-D., et al. (2020). PTB-XL, a large publicly available electrocardiography dataset. *Scientific Data*, 7:154. <https://doi.org/10.1038/s41597-020-0495-6>
- [32] Strodthoff N., Wagner P., Schaeffter T., Samek W. (2020). Deep Learning for ECG Analysis: Benchmarks and Insights from PTB-XL. *IEEE Journal of Biomedical and Health Informatics*, 25(5), 1519–1528. <https://doi.org/10.1109/JBHI.2020.3022989>
- [33] Makowski D., et al. (2023). NeuroKit2: A Python toolbox for neurophysiological signal processing. *Journal of Open Source Software*, 8(82):5231. <https://doi.org/10.21105/joss.05231>
- [34] Pilia N., et al. (2021). ECGDeli – An open source ECG delineation toolbox for MATLAB. *SoftwareX*, 13:100639. <https://doi.org/10.1016/j.softx.2021.100639>
- [35] Ni Xiong , Zibo Wei , Xuehua Wang , Yan Wang and Zhaohui Wang .HCTG-Net: A Hybrid CNN–Transformer Network with Gated Fusion for Automatic ECG Arrhythmia Diagnosis, 2025.
- [36] Mahmud T., Barua A., Islam D., Hossain M. S., Chakma R., Barua K. Ensemble Deep Learning Approach for ECG-Based Cardiac Disease Detection: Signal and Image Analysis. *Proceedings of the 2023 International Conference on ...*, IEEE, 2023.
- [37] Wagner, P., Strothoff, N., Bousseljot, R., Samek, W., & Schaeffter, T. *PTB-XL, a large publicly available electrocardiography dataset* (version 1.0.3). 2022
- [38] Interdata. (n.d.). *Ensemble Learning là gì? A-Z về học tập tổ hợp trong học máy*.2025
- [39] Zheng, Z., Chen, Z., Hu, F., Zhu, J., Tang, Q., & Liang, Y. (2020). *An automatic diagnosis of arrhythmias using a combination of CNN and LSTM technology*. *Electronics*, 9(1), 121.