

PHÂN LOẠI TÌNH TRẠNG THAI NHI TỪ DỮ LIỆU CARDIOTOCOGRAM DỰA TRÊN MÔ HÌNH ALEXNET-SVM

CLASSIFICATION OF FETAL STATUS BASED ON CARDIOTOCOGRAM DATA USING ALEXNET-SVM MODEL

Trần Thị Huế^{1,*}

DOI: <http://doi.org/10.57001/huih5804.2025.058>

TÓM TẮT

Bài báo này đề xuất một phương pháp phân loại tình trạng thai nhi dựa trên dữ liệu cardiotocogram (CTG) bằng cách sử dụng mô hình kết hợp AlexNet và SVM (AlexNet-SVM). Dữ liệu CTG bao gồm các thông số liên quan đến nhịp tim thai nhi (FHR) và cơn co tử cung (UC) của các bà mẹ mang thai, được phân loại thành ba nhóm: bình thường, nghi ngờ và bệnh lý. Nhằm tăng độ chính xác trong việc phân loại dữ liệu phức tạp này, tác giả đã sử dụng mạng nơ-ron tích chập (CNN) AlexNet để trích xuất đặc trưng và máy vector hỗ trợ (SVM) để thực hiện việc phân loại. Kết quả cho thấy mô hình AlexNet-SVM đạt độ chính xác 98,02%. Mô hình này có tiềm năng ứng dụng trong việc hỗ trợ bác sĩ phát hiện sớm các vấn đề sức khỏe thai nhi, qua đó giảm thiểu rủi ro và nâng cao sức khỏe cho cả mẹ và bé.

Từ khóa: Cardiotocogram; AlexNet; SVM.

ABSTRACT

This paper proposes a method for classifying fetal status based on cardiotocogram (CTG) data using a hybrid model combining AlexNet and SVM (AlexNet-SVM). The CTG data includes parameters related to fetal heart rate (FHR) and uterine contractions (UC) of pregnant mothers, classified into three categories: normal, suspicious, and pathological. To improve accuracy in classifying this complex data, the author employed the AlexNet convolutional neural network (CNN) for feature extraction and support vector machine (SVM) for classification. The results show that the AlexNet-SVM model achieved an accuracy of 98.02%. This model has potential applications in assisting doctors with early detection of fetal health issues, thereby reducing risks and improving health outcomes for both mother and child.

Keywords: Cardiotocogram; AlexNet; SVM.

¹Viện Tên lửa và Kỹ thuật điều khiển, Học viện Kỹ thuật Quân sự

*Email: tranhue3cx@gmail.com

Ngày nhận bài: 10/10/2024

Ngày nhận bài sửa sau phản biện: 23/01/2025

Ngày chấp nhận đăng: 28/3/2025

1. GIỚI THIỆU

Thời gian mang thai kéo dài khoảng 9 tháng là giai đoạn quan trọng, trong đó có nhiều biến chứng có thể

phát sinh, dẫn đến nguy cơ tử vong ở trẻ sơ sinh. Các biến chứng này có thể xuất phát từ nhiều yếu tố như tình trạng sức khỏe của người mẹ, nhiễm trùng hoặc các vấn đề trong sự phát triển của thai nhi. Các biến chứng khi mang thai bao gồm tăng huyết áp, tiểu đường thai kỳ, nhiễm trùng, tiền sản giật, sẩy thai và sảy thai, sinh non và thai chết lưu. Ngoài ra, cũng có thể buồn nôn, nôn mửa và thiếu máu nghiêm trọng do thiếu sắt [1]. Trước những rủi ro lớn đối với cả mẹ và bé, việc giảm thiểu các tai biến tiềm ẩn là vô cùng quan trọng. Sử dụng các công cụ theo dõi và chẩn đoán hiệu quả có thể giúp phát hiện sớm và can thiệp kịp thời, đảm bảo an toàn và sức khỏe cho cả mẹ và trẻ sơ sinh.

Một trong những công cụ quan trọng nhất để phân tích sức khỏe của thai nhi trong bụng mẹ là thực hiện Cardiotocogram (CTG). CTG là một kỹ thuật phổ biến được sử dụng để liên tục theo dõi và ghi lại nhịp tim thai nhi (FHR) và các cơn co tử cung (UC) trong suốt thai kỳ, giúp đánh giá tình trạng sức khỏe của thai nhi và chẩn đoán sớm nguy cơ gặp phải các vấn đề thai kỳ [2]. Phương pháp này cho phép phát hiện sớm tình trạng thiếu oxy ở thai nhi, từ đó có thể can thiệp kịp thời trước khi xảy ra các biến chứng nghiêm trọng như ngạt nặng hoặc tử vong [3]. Theo Liên đoàn Phụ khoa và Sản khoa Quốc tế (FIGO), trong quá trình co tử cung, các dấu hiệu quan trọng như nhịp tim thai nhi, sự biến thiên của nhịp, khả năng đáp ứng và giảm tốc độ là các chỉ số chính cho thấy thai nhi có đang phát triển bình thường hay không [4]. CTG được thực hiện đơn giản, chi phí thấp, tiết kiệm thời gian, đồng thời cung cấp các kết quả với độ chính xác cao nên được các chuyên gia y tế lựa chọn nhằm phát hiện sớm tình trạng thai nhi, giảm biến chứng và giảm tỷ lệ tử vong cho mẹ và bé.

Trí tuệ nhân tạo (AI) đang trở thành một trong những công nghệ đột phá trong lĩnh vực chăm sóc sức khỏe, mang lại những giải pháp tiên tiến trong việc hỗ trợ chẩn

đoán và điều trị [5, 6]. Một trong những ứng dụng quan trọng của AI trong y tế là sử dụng các mô hình học sâu (deep learning) để phân tích dữ liệu y tế, từ đó hỗ trợ bác sĩ đưa ra các quyết định kịp thời và chính xác. Trong những năm gần đây, các phương pháp học máy đã được ứng dụng rộng rãi trong việc phân loại dữ liệu y tế, đặc biệt là dữ liệu CTG, giúp hỗ trợ bác sĩ trong quá trình chẩn đoán tình trạng sức khỏe thai nhi. Các mô hình như Random Forest [7], LS-SVM [8] và DenseNet [9] đã được áp dụng thành công trong các nghiên cứu trước đây với độ chính xác lần lượt đạt 83,30%, 80,64% và 93,67%. Tuy nhiên, mặc dù đạt được những kết quả đáng kể, các mô hình này vẫn chưa thể tối ưu hóa hoàn toàn độ chính xác và khả năng phân loại, nhất là trong các trường hợp dữ liệu phức tạp và không đồng nhất.

Trong những năm gần đây, mô hình AlexNet-SVM đã được ứng dụng rộng rãi trong các lĩnh vực khác như tài chính, an ninh mạng, xử lý ngôn ngữ tự nhiên, và thị giác máy tính. Trong phân tích tài chính, mô hình này hỗ trợ phân loại dữ liệu phức tạp và dự đoán xu hướng thị trường nhờ khả năng trích xuất đặc trưng mạnh mẽ của AlexNet và độ chính xác cao của SVM [10]. Trong an ninh mạng, AlexNet-SVM được sử dụng để phát hiện xâm nhập và phân loại các mối đe dọa, trong đó AlexNet đảm nhận việc trích xuất đặc trưng từ dữ liệu mạng và SVM đảm bảo khả năng phân loại chính xác các mối nguy [11]. Trong xử lý ngôn ngữ tự nhiên, mô hình này hỗ trợ phân loại văn bản, phát hiện thư rác, và phân tích cảm xúc, tận dụng khả năng học sâu của AlexNet và tính phân loại hiệu quả của SVM [12].

Những ứng dụng này cho thấy sự linh hoạt và hiệu quả của mô hình AlexNet-SVM trong nhiều lĩnh vực, đặc biệt trong xử lý dữ liệu phức tạp. Do đó, nghiên cứu này đề xuất sử dụng AlexNet-SVM cho việc phân loại tình trạng thai nhi từ dữ liệu CTG, với kết quả ban đầu đạt được cho thấy tiềm năng lớn trong chẩn đoán y tế.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Dữ liệu và tiền xử lý

Dữ liệu được sử dụng trong nghiên cứu này là bộ dữ liệu Cardiotocogram (CTG) từ Kho Dữ liệu Học Máy UCI [13]. Bộ dữ liệu bao gồm 2126 mẫu, được ghi nhận từ các bà mẹ mang thai, với 21 đặc trưng liên quan đến nhịp tim thai nhi (FHR) và cơn co tử cung (UC). Dữ liệu được phân loại thành ba nhóm: bình thường (Normal), nghi ngờ (Suspect), và bệnh lý (Pathological) dựa trên phân tích của các chuyên gia y tế. Các đặc trưng được mô tả ở bảng 1.

Theo tiêu chuẩn của Viện Sức khỏe Trẻ em và Phát triển Con người Quốc gia, các biến có nguy cơ chính được

sử dụng để xác định tình trạng của thai nhi dựa trên mô tả của FHR là nhịp tim cơ sở, độ biến thiên cơ bản, số lần tăng tốc mỗi giây, số lần giảm tốc sớm, muộn và thay đổi mỗi giây, số lần giảm tốc kéo dài mỗi giây và các cơn co tử cung có thể được nhìn thấy từ giai điệu cơ bản của tử cung, tần suất, thời gian và cường độ của cơn co.

Bảng 1. Mô tả các thuộc tính

Thuộc tính	Mô tả
Baseline Value (BV)	Giá trị nhịp tim cơ bản của thai nhi (Fetal Heart Rate - FHR)
Accelerations (AC)	Số lần tăng tốc của nhịp tim thai nhi (tăng FHR trên 15bpm kéo dài ≥ 15 giây)
Fetal Movement (FM)	Chuyển động của thai nhi
Uterine Contractions (UC)	Cơn co tử cung
Light Decelerations (LD)	Giảm nhịp tim nhẹ của thai nhi (giảm FHR nhẹ)
Severe Decelerations (SD)	Giảm nhịp tim nghiêm trọng của thai nhi (giảm FHR > 30 bpm kéo dài > 2 phút)
Prolonged Decelerations (PD)	Prolonged Decelerations - Giảm nhịp tim kéo dài của thai nhi (giảm FHR kéo dài ≥ 2 phút nhưng < 10 phút)
Abnormal Short-Term Variability (ASTV)	Biến thiên ngắn hạn bất thường của nhịp tim thai nhi (dao động nhịp tim ≤ 5 bpm)
Mean Value of Short-Term Variability (MSTV)	Giá trị trung bình của độ biến thiên ngắn hạn của FHR (dao động tức thời giữa các nhịp tim kế tiếp)
Percentage of Time with Abnormal Long-Term Variability (ALTV)	Tỷ lệ thời gian với biến thiên dài hạn bất thường của FHR
Mean Value of Long-Term Variability (MLTV)	Giá trị trung bình của biến thiên dài hạn của FHR (dao động lớn hơn và chậm hơn trong khoảng thời gian dài hơn)
Histogram Width (HW)	Độ rộng của biểu đồ tần suất nhịp tim (FHR)
Histogram Max (Hmax)	Giá trị nhịp tim tối đa trong biểu đồ FHR
Histogram Min (Hmin)	Giá trị nhịp tim tối thiểu trong biểu đồ FHR
Number of Histogram Peaks (NP)	Số lượng đỉnh trong biểu đồ FHR
Number of Histogram Zeroes (NZ)	Số lần FHR bằng 0 trong biểu đồ
Histogram Mode (HMo)	Giá trị mode (giá trị xuất hiện nhiều nhất) của FHR
Histogram Mean (HMe)	Giá trị trung bình của nhịp tim trong biểu đồ FHR
Histogram Median (HMed)	Giá trị trung vị của nhịp tim trong biểu đồ FHR
Histogram Variance (HV)	Độ biến thiên của FHR trong biểu đồ
Histogram Tendency (HT)	Xu hướng của biểu đồ FHR

Bộ dữ liệu gồm có 1655 mẫu bình thường (77,8%), 295 mẫu nghi ngờ (13,9%) và 176 mẫu bệnh lý (8,3%). Do sự mất cân bằng của tập dữ liệu, lớp nghi ngờ và bệnh lý có số lượng mẫu ít hơn nhiều so với lớp bình thường. Vì vậy, kỹ thuật tăng cường dữ liệu (oversampling) được áp dụng cho các lớp có số lượng mẫu ít hơn để cân bằng dữ liệu [14]. Kỹ thuật này sử dụng thuật toán k-láng giềng gần nhất (k-nearest neighbors) để tổng hợp thêm các mẫu mới. Over sampling là nhóm các phương pháp giúp giải quyết hiện tượng mất cân bằng mẫu bằng cách gia tăng kích thước mẫu thuộc nhóm thiểu số. Các mô hình được đào tạo trên dữ liệu đã được cân bằng thường có hiệu suất cao hơn đáng kể so với các mô hình được đào tạo trên dữ liệu mất cân bằng.

Để đảm bảo hiệu quả trong việc huấn luyện mô hình, chuẩn hóa Z-score [15] được áp dụng cho các đặc trưng trong bộ dữ liệu. Đây là một phương pháp phổ biến để chuẩn hóa các đặc trưng có thang đo khác nhau, đặc biệt khi các đặc trưng có giá trị dao động lớn. Chuẩn hóa Z-score giúp đưa tất cả các đặc trưng về cùng một thang đo với trung bình bằng 0 và độ lệch chuẩn bằng 1, thông qua công thức (1):

$$X_{\text{std}} = \frac{x - \mu}{\sigma} \quad (1)$$

Trong đó:

x: Giá trị của đặc trưng cần chuẩn hóa.

μ : Giá trị trung bình của đặc trưng.

σ : Độ lệch chuẩn của đặc trưng.

Sau khi chuẩn hóa, dữ liệu được chuyển đổi thành định dạng ảnh để phù hợp với mô hình AlexNet. Cụ thể, mỗi mẫu dữ liệu bao gồm 21 đặc trưng được sắp xếp thành dạng ma trận hoặc lưới, giúp biểu diễn dưới dạng ảnh. Các đặc trưng này sau đó được chuẩn hóa và chuyển đổi thành ảnh có kích thước $227 \times 227 \times 3$, phù hợp với đầu vào của mô hình AlexNet. Dữ liệu dạng số từ 21 đặc trưng được mô phỏng như một ảnh màu với ba kênh RGB. Quá trình này đảm bảo rằng dữ liệu đầu vào phù hợp với các lớp tích chập (convolutional layers) của AlexNet.

2.2. Kiến trúc AlexNet

Mô hình mạng AlexNet được Alex Krizhevsky giới thiệu trong nghiên cứu ImageNet Classification with Deep Convolutional Networks. Mô hình học sâu này đã giành chiến thắng trong cuộc thi ILSVRC (ImageNet Large-Scale Visual Recognition Challenge) năm 2012 [16]. AlexNet là một mạng nơ-ron tích chập (Convolutional Neural Network - CNN) sâu, bao gồm năm lớp tích chập (convolutional layers), ba lớp kết nối đầy đủ (fully connected layers) và các lớp kích hoạt phi tuyến ReLU

(Rectified Linear Unit) và Softmax. AlexNet ban đầu được phát triển cho nhiệm vụ phân loại hình ảnh, nhưng trong nghiên cứu này, nó được điều chỉnh để làm việc với dữ liệu tín hiệu y tế. Các lớp tích chập sẽ trích xuất các đặc trưng từ dữ liệu đầu vào và các lớp kết nối đầy đủ sẽ học cách phân loại dựa trên các đặc trưng này.

Mạng học sâu AlexNet sử dụng các lớp kích hoạt phi tuyến - ReLU cho một đơn vị huấn luyện thần kinh [17]. Theo đó, phương pháp ReLU có thời gian huấn luyện nhanh hơn phương pháp cổ điển sử dụng hàm $f(x) = \tanh(x)$ hoặc $f(x) = (1 + e^{-x})^{-1}$ trong mạng tích chập 4 lớp. Đây là ưu điểm rất lớn của mạng AlexNet khi huấn luyện các tập dữ liệu có số lượng lớn ảnh. Phương trình hàm ReLU được miêu tả như sau:

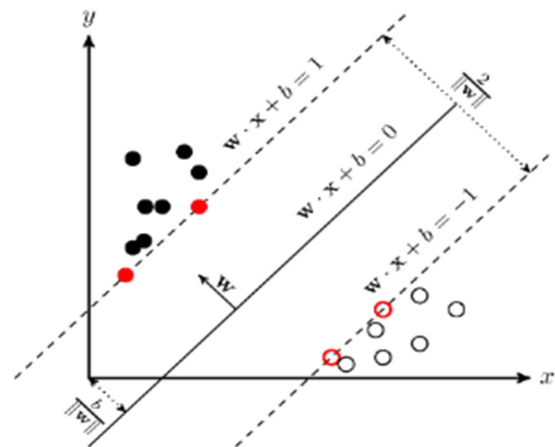
$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (2)$$

Với x là giá trị ngõ vào của hàm số.

Dựa trên phương trình (2), ngõ ra cho giá trị 0 nếu ngõ vào là âm, và ngõ ra x nếu ngõ vào lớn hơn 0. Từ đó, ta có thể miêu tả lại ReLU dưới dạng $f(x) = \max(0, x)$.

Phương trình này cho thấy hàm ReLU dễ lập trình và khó bị bão hòa hơn các hàm cổ điển.

2.3. Support Vector Machine (SVM)



Hình 1. Siêu phẳng (w,b) tối ưu phân chia thành 2 class [19]

SVM là một giải thuật máy học dựa trên lý thuyết học thống kê do Vapnik và Chervonenkis xây dựng [18]. Bài toán cơ bản của SVM là bài toán phân loại hai lớp: Cho trước r điểm trong không gian n chiều (mỗi điểm thuộc vào một lớp kí hiệu là +1 hoặc -1), mục đích của giải thuật SVM là tìm một siêu phẳng (hyperplane) phân hoạch tối ưu cho phép chia các điểm này thành hai phần sao cho các điểm cùng một lớp nằm về một phía với siêu phẳng này.

Sau quá trình luyện tập với tập dữ liệu đầu vào, SVM sẽ đưa ra các tham số của phương trình siêu phẳng ranh giới giữa hai lớp, hai tham số này là véc tơ w có p chiều và tham số b .

Phương trình của siêu phẳng là: $w \cdot x + b = 0$

Khi đó hàm số biểu diễn dấu của biểu thức: $f(x) = \text{sign}(wx + b)$.

$f(x)$ là hàm số có khả năng phân chia hoàn toàn dữ liệu vào một trong hai lớp.

Sau quá trình luyện tập, các véc tơ mẫu thử có thể được phân vào một trong hai lớp bằng hàm số $f(x)$:

- $f(x) = \text{sign}(wx + b) > 0$: xếp vào lớp (+), $y = 1$
- $f(x) = \text{sign}(wx + b) < 0$: xếp vào lớp (-), $y = -1$

2.4. Mô hình đề xuất AlexNet-SVM

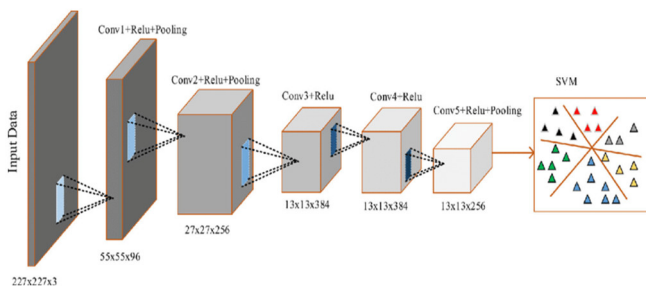
Trong nghiên cứu này, tác giả đề xuất sử dụng mô hình AlexNet-SVM, một sự kết hợp giữa mạng nơ-ron tích chập AlexNet để trích xuất đặc trưng và SVM để phân loại. Đây là cách tiếp cận nhằm tận dụng ưu thế của cả hai kỹ thuật:

Trích xuất đặc trưng bằng AlexNet:

- Dữ liệu đầu vào là các ảnh kích thước $227 \times 227 \times 3$, được đưa qua 5 lớp tích chập của AlexNet.
- Sau mỗi lớp tích chập, phép max-pooling được sử dụng để giảm kích thước không gian.

Thay thế lớp Fully Connected bằng SVM:

- Thay các lớp fully connected ở phần cuối của mô hình AlexNet bằng Support Vector Machine (SVM). SVM có khả năng phân loại tốt hơn, đặc biệt khi đối mặt với dữ liệu y tế có sự phức tạp.
- SVM giúp tìm ra siêu phẳng tối ưu để phân tách các lớp trong không gian đặc trưng, giảm thiểu hiện tượng quá khớp (overfitting) so với các lớp fully connected thông thường.



Hình 2. Mô hình đề xuất AlexNet-SVM

Mô hình AlexNet-SVM trong nghiên cứu này nổi bật với việc áp dụng kỹ thuật học chuyển tiếp. Thay vì huấn luyện toàn bộ mạng từ đầu, nghiên cứu tận dụng các tham số đã được huấn luyện, giúp tiết kiệm thời gian và

cải thiện đáng kể hiệu suất phân loại trên dữ liệu cardiocotogram (CTG). Cách tiếp cận này không chỉ tăng độ chính xác mà còn giúp mô hình thích nghi tốt hơn với dữ liệu y tế phức tạp và không đồng nhất. Đây là điểm cải tiến chính, làm nổi bật sự khác biệt và tính tối ưu của mô hình được đề xuất.

Học chuyển tiếp [20] là một kỹ thuật học máy cho phép sử dụng lại các tham số đã được huấn luyện từ một mô hình trước đó. Trong nghiên cứu này, tác giả sử dụng mô hình AlexNet đã được huấn luyện, sau đó tinh chỉnh nó cho nhiệm vụ phân loại dữ liệu CTG.

Việc sử dụng học chuyển tiếp giúp tiết kiệm thời gian huấn luyện, cải thiện độ chính xác và cho phép mô hình nhanh chóng thích ứng với dữ liệu y tế.

Sau khi đặc trưng được trích xuất qua các lớp tích chập, và sau khi phân loại bằng SVM, kết quả cuối cùng sẽ được chuyển đổi thành phân phối xác suất qua lớp SoftMax. Mô hình sẽ đưa ra phân loại giữa các nhãn: bình thường, nghi ngờ và bệnh lý.

AlexNet-SVM kết hợp sức mạnh của CNN trong việc học đặc trưng từ dữ liệu hình ảnh và khả năng phân loại mạnh mẽ của SVM. Sử dụng học chuyển tiếp giúp tăng hiệu quả và độ chính xác, giảm thời gian huấn luyện mà vẫn đảm bảo hiệu quả cao.

3. KẾT QUẢ THỰC NGHIỆM

3.1. Phương pháp đánh giá

Đánh giá hiệu quả của các thuật toán dựa trên 4 độ đo sau đây: Accuracy, Precision, Recall và F1-score.

Cụ thể để đánh giá một bộ phân loại hai lớp được gọi là dương và âm:

- True Positive (TP): số lượng phần tử dương được phân loại dương;
- Âm tính giả (FN): số phần tử dương được phân loại là âm;
- Âm thật (TN): số phần tử âm được phân loại là âm;
- Dương tính giả (FP): số phần tử âm được phân loại là dương.

Các chỉ số dùng để đánh giá mô hình học máy [21]:

Accuracy: Độ chính xác phân loại biểu thị tỷ lệ của các mẫu được phân loại chính xác trên tổng số mẫu và được mô tả ở phương trình (3).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \cdot 100\% \quad (3)$$

Precision: Giá trị dự đoán dương tính, precision biểu thị tỷ lệ số mẫu dương tính được phân loại chính xác trên

tổng số mẫu dự đoán dương tính và được mô tả ở phương trình (4).

$$\text{Precision} = \frac{TP}{TP + FP} \cdot 100\% \quad (4)$$

Recall: Tỷ lệ dương tính thật hay độ nhạy, recall biểu thị tỷ lệ giữa mẫu dương tính với tổng số mẫu dương tính và số mẫu âm tính giả, được mô tả ở phương trình (5).

$$\text{Recall} = \frac{TP}{TP + FN} \cdot 100\% \quad (5)$$

F1-score: F1-score được định nghĩa là giá trị trung bình hài hòa giữa Precision và Recall, được mô tả ở phương trình (6). F1-score có phạm vi từ 0 đến 1, giá trị cao của nó thể hiện hiệu suất phân loại cao.

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \cdot 100\% \quad (6)$$

3.2. Kết quả

Dữ liệu được tách ngẫu nhiên thành hai phần để phát triển các thuật toán (phần dữ liệu đào tạo và phần dữ liệu kiểm tra). Phần dữ liệu đào tạo chiếm 70% tổng dữ liệu và phần dữ liệu kiểm tra chiếm 30% dữ liệu còn lại.

Mô hình sử dụng kiến trúc AlexNet với sự kết hợp của SVM nhằm cải thiện khả năng phân loại và tăng tốc độ hội tụ. Cấu hình siêu tham số của mô hình được tối ưu hóa thông qua quá trình thử nghiệm trên tập dữ liệu huấn luyện và kiểm tra, với mục tiêu đạt hiệu suất tốt nhất về độ chính xác và tốc độ xử lý.

Cụ thể, mô hình sử dụng 5 lớp tích chập với các bộ lọc có kích thước khác nhau, được thiết kế để trích xuất các đặc trưng từ dữ liệu một cách hiệu quả. Mô hình cũng sử dụng cách thức tổng hợp dữ liệu max pooling để giảm số lượng thông tin đầu ra từ các lớp tích chập, giúp giảm thiểu độ phức tạp tính toán mà vẫn giữ được các đặc trưng quan trọng. Thay vì sử dụng các lớp kết nối đầy đủ (fully connected layers) thông thường, mô hình này thay thế bằng SVM ở lớp cuối cùng, giúp mô hình tránh bị quá khớp (overfitting) và cải thiện độ chính xác trong phân loại. Quá trình huấn luyện sử dụng thuật toán tối ưu Adam với hàm mất mát cross-entropy, đảm bảo quá trình cập nhật trọng số diễn ra hiệu quả.

Các siêu tham số chi tiết của mô hình được tóm tắt trong bảng 2.

Hiệu suất của mô hình được đánh giá bằng các chỉ số chính bao gồm độ chính xác (Accuracy), độ nhạy (Recall), độ chính xác dự đoán (Precision) và điểm F1 (F1-score). Những chỉ số này đánh giá mức độ hiệu quả của mô hình trong việc phân loại trạng thái sức khỏe thai nhi, và được trình bày trong bảng 3.

Bảng 2. Các giá trị siêu tham số

Siêu tham số	Giá trị
Số lớp tích chập	5
Số bộ lọc (lớp 1)	96 bộ lọc, kích thước 11x11
Số bộ lọc (lớp 2)	256 bộ lọc, kích thước 5x5
Số bộ lọc (lớp 3-5)	384 bộ lọc, kích thước 3x3
Loại pooling	Max pooling (stride = 2, kích thước = 3x3)
Số lớp kết nối đầy đủ	Thay bằng SVM
Hàm kích hoạt (lớp tích chập)	ReLU
Hàm kích hoạt (lớp đầu ra)	SoftMax
Tốc độ học	Ban đầu: 0,0001
Bộ tối ưu	Adam
Hàm mất mát	Cross-entropy
Kích thước batch	10
Số epoch	10

Bảng 3. Kết quả của mô hình AlexNet-SVM

Đơn vị tính: (%)

Mô hình	Accuracy	Recall	Precision	F1-score
AlexNet-SVM	98,02	98,61	97,99	98,29

Các kết quả trên cho thấy mô hình AlexNet-SVM đạt được hiệu suất cao với độ chính xác tổng thể là 98,02%, điểm F1 là 98,29%, cho thấy mô hình có khả năng phân loại chính xác và ổn định. Với độ nhạy cao 98,61% và độ chính xác dự đoán 97,99%, mô hình đã thể hiện khả năng dự đoán đúng các trạng thái sức khỏe thai nhi một cách hiệu quả.

Các kết quả của mô hình AlexNet-SVM được so sánh với các mô hình trước đây như Random Forest, LS-SVM, AlexNet và DenseNet. Kết quả cụ thể được thể hiện trong bảng 4.

Bảng 4. Kết quả của mô hình AlexNet-SVM và các mô hình nghiên cứu khác

Đơn vị tính: (%)

Mô hình	Accuracy	Recall	Precision	F1-score	Nguồn
AlexNet-SVM	98,02	98,61	97,99	98,29	Nghiên cứu đề xuất
Random Forest	83,30	90,29	76,32	82,36	[7]
LS-SVM	80,64	81,28	80,00	80,64	[8]
DenseNet	93,67	84,45	82,36	90,52	[9]

Kết quả cho thấy mô hình AlexNet-SVM đạt hiệu suất vượt trội so với các mô hình khác, với độ chính xác 98,02%, cao hơn so với các mô hình như Random Forest (83,30%), LS-SVM (80,64%) và DenseNet (93,67%). Điểm F1 của AlexNet-SVM cũng cao nhất với 98,29%, cho thấy khả năng cân bằng tốt giữa độ nhạy và độ chính xác dự đoán.

4. KẾT LUẬN

Trong nghiên cứu này, tác giả đã đề xuất mô hình AlexNet-SVM kết hợp giữa mạng nơ-ron tích chập (CNN) và SVM để nâng cao khả năng phân loại trạng thái sức khỏe thai nhi. Kết quả cho thấy mô hình AlexNet-SVM đạt độ chính xác 98,02%, cao hơn so với các mô hình trước đây. Điều này cho thấy tiềm năng lớn của việc kết hợp các phương pháp học sâu với thuật toán phân loại truyền thống như SVM trong việc phân tích và dự đoán dữ liệu y tế phức tạp. Sự cải tiến chính của mô hình là sự kết hợp giữa mạng nơ-ron tích chập (AlexNet) và SVM ở lớp đầu ra, giúp tăng cường khả năng phân loại và giảm thiểu hiện tượng quá khớp. Mô hình này còn có ưu điểm về tốc độ hội tụ nhanh hơn và hiệu quả tính toán cao hơn so với các mô hình mạng nơ-ron thông thường.

Hướng phát triển tiếp theo của báo cáo là mở rộng việc thu thập thêm dữ liệu CTG từ nhiều khu vực địa lý và ở các nhóm tuổi khác nhau để tạo ra một bộ dữ liệu phong phú hơn, giúp mô hình có khả năng tổng quát tốt hơn. Bên cạnh đó, áp dụng thêm nhiều phương pháp học máy khác để nâng cao độ chính xác của mô hình. Ngoài ra, sẽ tập trung vào việc tích hợp mô hình vào các hệ thống hỗ trợ chẩn đoán tự động, giúp bác sĩ theo dõi và phân loại tình trạng sức khỏe thai nhi hiệu quả hơn, qua đó giảm thiểu rủi ro trong thai kỳ và nâng cao sức khỏe cho cả mẹ và bé.

TÀI LIỆU THAM KHẢO

- [1]. <https://www.nichd.nih.gov/health/topics/pregnancy/conditioninfo/Pages/complications.aspx>, truy cập ngày 25/9/2024
- [2]. Signorini MG, Pini N, Malovini A, Bellazzi R, Magenes G, "Integrating machine learning techniques and physiology-based heart rate features for antepartum fetal monitoring," *Computer Methods and Programs in Biomedicine*, 185, 105015, 2020.
- [3]. I. Ingemarsson, "Fetal monitoring during labor," *Neonatology*, 95(4), 342-346, 2009.
- [4]. Ayres-De-Campos D., C. Y. Spong, E. Chandraran, "FIGO GUIDELINES FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography," *FIGO Intrapartum Fetal Monitoring Expert Consensus Panel Int J Gynaecol Obs*, 131, 13-24, 2015.
- [5]. K. H. Miao, J. H. Miao, G. J. Miao, "Diagnosing coronary heart disease using ensemble machine learning," *International Journal of Advanced Computer Science and Applications*, 7(10), 30-39, 2016.
- [6]. K. H. Miao, G. J. Miao, "Mammographic diagnosis for breast cancer biopsy predictions using neural network classification model and receiver operating characteristic (ROC) curve evaluation," *Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Bioinformatics*, 3(9), 1-10, 2013.

- [7]. Peterek T., P. Gajdoš, P. Dohnálek, J. Krohová, "Human fetus health classification on cardiotocographic data using random forests," *Advances in Intelligent Systems and Computing*, 317, 323-330, 2014.
- [8]. Yilmaz H., M. Ekinci, F. Gürgen, "Classification of fetal state from cardiotocography data using LS-SVM," *IEEE Transactions on Information Technology in Biomedicine*, 17(1), 27-32, 2013.
- [9]. Huang G., Z. Liu, L. Van Der Maaten, K.Q. Weinberger, "Densely connected convolutional networks," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 9(4), 2261-2269, 2020.
- [10]. Miao K. H., J. H. Miao, G. J. Miao, "Diagnosing coronary heart disease using ensemble machine learning," *International Journal of Advanced Computer Science and Applications*, 7(10), 30-39, 2016.
- [11]. Krizhevsky A., Sutskever I., Hinton G. E., "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, 25, 1097-1105, 2012.
- [12]. Jain A. K., Duin R. P. W., Mao J., "Statistical pattern recognition: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 4-37, 2000.
- [13]. <http://archive.ics.uci.edu/ml/datasets/Cardiotocography>, truy cập lần cuối ngày 25/9/2024
- [14]. Douzas G., Bacao F., "Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE," *Information Sciences*, 465, 1-20, 2019.
- [15]. Jain A.K., Duin R.P.W., Mao J., "Statistical Pattern Recognition: A Review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 4-37, 2000.
- [16]. Krizhevsky A., Sutskever I., Hinton G. E., "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, 25, 1097-1105, 2012.
- [17]. Nair V., Hinton G. E., "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th International Conference on Machine Learning (ICML)*, Haifa, Israel, 807-814, 2010.
- [18]. Vapnik V. N., *The nature of statistical learning theory*. Springer-Verlag, New York, USA, 211-248, 1995.
- [19]. Đệ T. C., Khang P. N., "Phân loại văn bản với máy học vector hỗ trợ và cây quyết định," *Tạp chí Khoa học Đại học Cần Thơ*, (21a), 52-63, 2012.
- [20]. Pan S. J., Yang Q., "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345-1359, 2010.
- [21]. A. K. Dwivedi, "Performance evaluation of different machine learning techniques for prediction of heart disease," *Neural Computing and Applications*, 29(10), 685-693, 2018.

AUTHOR INFORMATION

Tran Thi Hue

Institute of Missiles and Control Engineering, Military Technical Academy, Vietnam