

Hệ thống ứng dụng trí tuệ nhân tạo (AI) chẩn đoán hội chứng ngưng thở khi ngủ do tắc nghẽn (OSA) sử dụng tín hiệu điện tâm đồ ECG và độ bão hoà oxy SpO2

Nông Nguyễn Tường Vy*, Bùi Minh Triết*, Nguyễn Hải Đăng**

** Trường Phổ thông Năng khiếu, ĐHQG-HCM*

*** Phòng thí nghiệm Công nghệ Phần mềm, Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM*

1. TÓM TẮT

Hội chứng ngưng thở khi ngủ do tắc nghẽn (OSA) là một bệnh lý phổ biến có thể dẫn đến hàng loạt các tình trạng gây hại cho sức khoẻ bệnh nhân. Kỹ thuật đa ký giấc ngủ là một tiêu chuẩn vàng để chẩn đoán hội chứng này, tuy nhiên chi phí sử dụng rất đắt đỏ, tốn nhiều thời gian và cần có chuyên gia về giấc ngủ để phân tích thủ công. Bài nghiên cứu này đề xuất một phương pháp chẩn đoán OSA dựa trên tín hiệu điện tâm đồ (ECG) và độ bão hoà oxy trong máu (SpO2) bằng mô hình học sâu Convolutional Neural Network (CNN) kết hợp với kỹ thuật rút trích đặc trưng Mel Frequency Cepstral Coefficient (MFCC). Kết quả cho thấy với mô hình sử dụng ECG có độ chính xác ổn định 83% trên các lần thử nghiệm và mô hình sử dụng SpO2 có độ chính xác 93.4%. Khi kết hợp cả hai tín hiệu cho ra độ chính xác đến 96.2%. Kết quả trên đã nhấn mạnh tiềm năng và khả năng ứng dụng của mô hình mà chúng tôi như một công cụ chẩn đoán OSA hiệu quả, chính xác và tiện lợi.

2. ĐẶT VẤN ĐỀ

Hội chứng ngưng thở khi ngủ do tắc nghẽn (OSA) là một bệnh lý phổ biến, ảnh hưởng từ 3% đến 7% dân số toàn cầu. Tại Việt Nam, Bộ Y tế cho biết tỷ lệ người mắc OSA đang gia tăng, ước tính khoảng 5-10% dân số. Theo Viện Y học Giấc ngủ Hoa Kỳ, OSA ở người trưởng thành được xác định khi biên độ hô hấp qua miệng và mũi giảm đến 90% trong ít nhất 10 giây. Sự tắc nghẽn luồng khí lặp đi lặp lại này dẫn đến gián đoạn và giảm chất lượng giấc ngủ. Các triệu chứng thường gặp bao gồm ngáy to, thở không đều và buồn ngủ ban ngày. OSA không chỉ ảnh hưởng đến cuộc sống hàng ngày mà còn gây ra các vấn đề nghiêm trọng như suy giảm nhận thức, mệt mỏi, trầm cảm, và tăng nguy cơ đột quỵ, huyết áp cao, loạn nhịp tim, hoặc nhồi máu cơ tim.

Đa ký giấc ngủ (Polysomnography) là một phương pháp đo đạc các thông số của cơ thể khi ngủ để chẩn đoán và đánh giá mức độ nghiêm trọng của hội chứng ngưng thở khi ngủ và các rối loạn giấc ngủ khác [7]. Đây là tiêu chuẩn vàng trong chẩn đoán OSA, mặc dù có độ chính xác cao, nhưng đòi hỏi kỹ thuật đo phức tạp và chỉ được thực hiện trong môi trường bệnh viện. Phương pháp này không phù hợp cho việc sử dụng tại nhà hàng ngày do bệnh nhân phải ngủ ít nhất một đêm tại bệnh viện. Điều này làm mất nhiều thời gian, chi phí cao, và không thoải mái, nhất là với người Việt Nam, khi các triệu chứng chưa rõ ràng. Hơn nữa, bệnh nhân thường khó ngủ tự nhiên khi phải gắn nhiều điện cực trên cơ thể, dẫn đến giấc ngủ bị gián đoạn và làm sai lệch kết quả đo.

Ngoài ra, bệnh nhân thường sẽ gặp khó khăn trong việc tuân thủ khi phải đeo cảm biến áp suất để đo luồng khí qua miệng và mũi, do cảm giác khó chịu khi ngủ khiến họ thường xuyên tháo ống thông mũi. Điều này làm cho dữ liệu luồng khí từ mũi trở nên không ổn định và dễ bị sai lệch. Vì vậy, chúng tôi chọn sử dụng tín hiệu ECG để phát hiện các sự kiện OSA, nhằm đảm bảo tính ổn định và chính xác hơn trong quá trình chẩn đoán. Cũng như tâm lý “ngại” đến bệnh viện của người Việt Nam khi chưa có dấu hiệu rõ ràng [21].

Nhiều phương pháp hiện tại sử dụng mạng nơ-ron hồi tiếp (RNN) do tính liên tục của dữ liệu. Tuy nhiên, mục tiêu của mô hình là phân tích một đoạn tín hiệu dữ liệu ngắn (dài 1 phút) để dự đoán sự hiện diện của bệnh, cho phép dự đoán ngay cả trên các đoạn tín hiệu ngắn mà không cần yêu cầu bệnh nhân phải ngủ dài 7-8 tiếng. Do đó, việc ưu tiên trích xuất mẫu là rất quan trọng. Vì thế chúng tôi đề xuất một mô hình học máy (Machine Learning) để chẩn đoán hội chứng ngưng thở khi ngủ do tắc nghẽn, sử dụng thuật toán nơ-ron tích chập Convolutional Neural Network (CNN) và Mel Frequency Cepstral Coefficient (MFCC) để trích xuất các tín hiệu ECG và tín hiệu SpO2 để phân tích. Bài nghiên cứu này sử dụng dữ liệu từ 35 bệnh nhân có tín hiệu ECG dạng số với tốc độ đo là 100 mẫu mỗi giây, mỗi phút và có đều có nhãn là bệnh hoặc không bệnh. Cùng với dữ liệu của 8 bệnh nhân có chỉ số phụ là SpO2 có cùng tốc độ đo và được dán nhãn cho mỗi phút đo.

Sự đóng góp

- Thực nghiệm trên tập dữ liệu mở Apnea-ECG Database của Physio.Net, một cơ sở dữ liệu lớn bao gồm cả tín hiệu ECG và SpO2 để kiểm chứng tính hiệu quả của mô hình đề xuất. Sử dụng tập dữ liệu mở khiến tính khả thi của nghiên cứu trở nên cao hơn vì dữ liệu đã được xác thực bởi cộng đồng nghiên cứu.
- Đề xuất mô hình chẩn đoán bằng tín hiệu ECG và SpO2 dựa trên các đặc trưng của phương pháp CNN và MFCC. Với CNN là một kỹ thuật để xử lý tín hiệu thời gian, có khả năng trích xuất đặc trưng từ các tín hiệu sinh học hiệu quả. Còn MFCC trong nghiên cứu này giúp trích xuất đặc trưng tần số của dữ liệu, cải thiện độ chính xác của mô hình chẩn đoán. Khi kết hợp cả hai phương pháp, mô hình cải thiện hiệu suất so sánh chẩn đoán.
- Mục tiêu cuối mà nghiên cứu này đề ra là đưa mô hình lên thiết bị đo lường di động giúp việc chẩn đoán được thực hiện trực tiếp lên thiết bị một cách tự động và chính xác hơn.

3. TỔNG QUAN NGHIÊN CỨU

Nhiều nhà nghiên cứu đã đề xuất nhiều phương pháp thay thế nhằm khắc phục các nhược điểm của kỹ thuật đa ký giấc ngủ (PSG) và phát hiện hội chứng ngưng thở khi ngủ do tắc nghẽn (OSA) bằng cách sử dụng ít tín hiệu và đơn giản hơn như tín hiệu điện tâm đồ (ECG), lưu lượng luồng khí (AF), sự thay đổi nhịp tim (HRV) và độ bão hòa oxy trong máu (SpO2). Một số nhà nghiên cứu cũng đã kết hợp các tín hiệu này để phát hiện OSA một cách hiệu quả hơn.

Trong một nghiên cứu của Ying-Chieh Hsu et al [1], các nhà khoa học đã nghiên cứu về việc áp dụng mô hình học máy và học sâu để phân loại tiếng ngáy dựa trên các mức độ hẹp của đường hô hấp trên ở bệnh nhân mắc OSA, và tìm thấy rằng các mô hình như SVM, BiLSTM, và ResNet-50 có hiệu quả cao trong việc phân loại âm thanh ngáy. Đã từng có những nghiên cứu về vai trò của mạng nơ-ron tích chập trong việc tự động trích xuất dữ liệu, như trong nghiên cứu của Paul et al [2], họ tìm thấy rằng thuật toán single CNN có khả năng phát hiện các giai đoạn ngưng thở từ tín hiệu ECG và độ bão hòa oxy (SpO2). Các nghiên cứu trước đã tiếp cận vấn đề này từ một góc độ khác, như nghiên cứu của Về tín hiệu SpO2. Trong một nghiên cứu của Hnin Thiri et al [3], các nhà khoa học đã phát triển mô hình CNN để phân tích tín hiệu bão hòa oxy (SpO2) từ cảm biến thông minh và tìm thấy rằng CNN có hiệu suất cao khi thử nghiệm trên dữ liệu bệnh nhân thực tế. Nghiên cứu này cũng so sánh CNN với các phương pháp học máy truyền thống như LDA, SVM và ANN để đánh giá hiệu quả phân loại tín hiệu. Tuy nhiên nghiên cứu trên cũng đã chỉ ra rằng tín hiệu SpO2 một chiều có thể hạn chế những thông tin cần thiết để phát hiện ra các sự kiện ngưng thở. Trong bài nghiên cứu của A.Leino et al [8] đã đề xuất một phương pháp ước tính mức độ nghiêm trọng của

hội chứng ngưng thở khi ngủ ở bệnh nhân đột quỵ cấp tính dựa trên tín hiệu SpO2 bằng mô hình CNN. Nghiên cứu đã đưa ra một phương pháp phát hiện tự động chứng ngưng thở ở các bệnh nhân mắc bệnh máu não cấp tính.

Trong một nghiên cứu của Fanlin Shen et al [4], các nhà khoa học đã nghiên cứu về việc sử dụng phương pháp học sâu để phân loại tiếng ngáy liên quan đến các sự kiện ngưng thở và tiếng ngáy không liên quan, và tìm thấy rằng việc trích xuất đặc trưng MFCC kết hợp với các mô hình CNN và LSTM mang lại độ chính xác cao, trong đó LSTM đạt độ chính xác cao nhất, là 0.87. Nhưng hạn chế của bài viết này là thiếu khả năng tổng quát, nghiên cứu không đưa ra các chi tiết về việc thử nghiệm trên các tập dữ liệu đa dạng. Muqing Deng et al [5] đã từng có những nghiên cứu về vai trò của thuật toán MFCC kết hợp với CRNN trong phân loại âm thanh tim và tìm thấy rằng việc sử dụng CNN để trích xuất đặc trưng cục bộ và RNN để nắm bắt các phụ thuộc dài hạn mang lại hiệu quả cao. Trong nghiên cứu của Y. Jiang [6] đã phát triển một thuật toán phát hiện tiếng ngáy tự động để phân loại âm thanh ngáy và không ngáy bằng phương pháp quang phổ và CNN. Để phát hiện các giai đoạn ngáy, bài viết sử dụng MFCC để trích xuất các đoạn âm thanh đã được phân đoạn thủ công. Sau đó các âm thanh này được phân loại thành tiếng ngáy và không phải tiếng ngáy bằng cách sử dụng mô hình HMM. ElMoaqet [9] cũng đã đưa ra một phương pháp mới dựa trên RNN sâu để phát hiện các sự kiện ngưng thở, sử dụng tín hiệu hô hấp từ một kênh duy nhất. Kết hợp với LSTM và BiLSTM, mô hình RNN sâu có khả năng xử lý tốt các phụ thuộc thời gian trong chuỗi dữ liệu. Tuy nhiên, nghiên cứu này cho thấy một hạn chế của RNN là việc học các mối quan hệ thời gian dài vẫn là thách thức và tiêu tốn dung lượng. Haijun Zhang et al [10] đã chứng minh rằng RNN, một phiên bản mở rộng của mạng nơ-ron truyền thẳng, có khả năng xử lý hiệu quả các chuỗi dữ liệu có độ dài thay đổi và dữ liệu thời gian, giúp mô hình hiểu được các mối liên hệ giữa các điểm trong chuỗi. Tuy nhiên, nghiên cứu cũng đã tìm thấy kết quả trái ngược khi chỉ ra rằng việc học các phụ thuộc thời gian dài vẫn gặp nhiều khó khăn do gradient bị suy giảm. Ngoài ra, RNN đã thể hiện hiệu suất vượt trội trong các ứng dụng như nhận dạng giọng nói và xử lý ngôn ngữ tự nhiên, nơi dữ liệu chuỗi và các phụ thuộc dài hạn đóng vai trò quan trọng [11].

Ngoài ra cũng đã có các bài nghiên cứu áp dụng những thuật toán hay phương pháp khác để chẩn đoán bệnh. Xie et al [12] nghiên cứu cách phân loại tiếng ngáy để chẩn đoán OSA sử dụng dữ liệu được ghi âm tại bệnh viện bằng cách kết hợp cả hai thuật toán Convolutional Neural Network (CNN) và Recurrent Neural Network (RNN). Để trích xuất tiếng ho, bài nghiên cứu đã sử dụng Mel-Frequency Cepstral Coefficient (MFCC), một phương pháp thường được sử dụng để nhận dạng giọng nói và nhận dạng người nói [13].

Cho dù các bài nghiên cứu trên có độ chính xác cao nhưng nhiều nghiên cứu vẫn chưa được kiểm định trong môi trường lâm sàng, chưa được áp dụng thực tế mà chỉ có mô hình. Có nhiều mô hình yêu cầu người dùng phải thực hiện nhiều bước thủ công để ghi lại quá trình ngáy hay để đo các tín hiệu dẫn đến trải nghiệm không thoải mái và gây khó khăn cho người dùng, đặc biệt là người cao tuổi. Còn ở bác sĩ, nếu hệ thống không cung cấp đủ dữ liệu và phân tích trong thời gian thực tế mà chỉ lấy những dữ liệu có sẵn sẽ làm nghiên cứu mất đi tín cập nhật dữ liệu trong thời gian thực.

Mô hình mà nhóm nghiên cứu đề xuất hỗ trợ chẩn đoán, giảm tải cho các bác sĩ và hỗ trợ cho con người nhận biết được nguy cơ mắc bệnh, tránh trường hợp bệnh xuất hiện triệu chứng xấu gây hại đến sức khỏe bản thân người bệnh. Đối với người sử dụng, họ sẽ có thể sử dụng thiết bị tại nhà để chẩn đoán bệnh có thể giảm được tâm lý né tránh khi phải đi đến bệnh viện và đồng thời có thể phát hiện bệnh sớm mà không cần phải đợi đến khi nào có triệu chứng nặng mới đi thì đã trễ để có thể điều trị. Còn đối với bác sĩ, việc

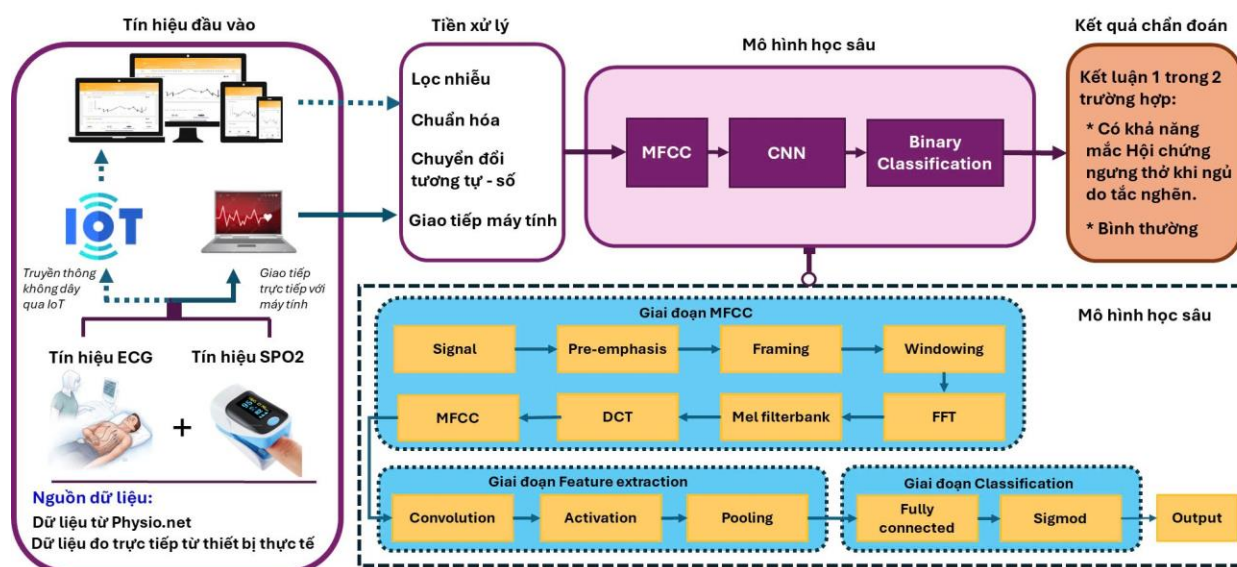
thiết kế ra mô hình và thiết bị này đã góp phần lọc bớt đi các bệnh nhân không có bệnh cũng như giảm tải cho bệnh viện.

4. DATASET

Dự án của chúng tôi thực hiện các thí nghiệm trên tập dữ liệu Apnea-ECG Database - [PhysioNet](#). [14] Tập dữ liệu bao gồm 70 bản ghi chỉ số đo ECG của các bệnh nhân khi ngủ, được chia thành một tập gồm 35 bản ghi chỉ có dán nhãn. Các bản ghi có độ dài khác nhau, từ dưới 7 giờ một chút đến gần 10 giờ mỗi bản. Mỗi bản ghi bao gồm tín hiệu ECG được số hóa liên tục, một bộ chú thích (có bệnh hay không bệnh) về OSA (do các chuyên gia con người lấy ra trên cơ sở nhịp thở được ghi đồng thời và các tín hiệu liên quan). Trong đó có 8 bệnh nhân có các chỉ số phụ, bao gồm có SpO2 - chỉ số mà nghiên cứu sử dụng cho mô hình phụ. Các bản ghi chứa các ECG được số hóa (16 bit trên mỗi mẫu, byte đầu tiên có ý nghĩa nhỏ nhất trong mỗi cặp, 100 mẫu mỗi giây, thông thường là 200 đơn vị A/D trên mỗi milivolt).

5. PHƯƠNG PHÁP

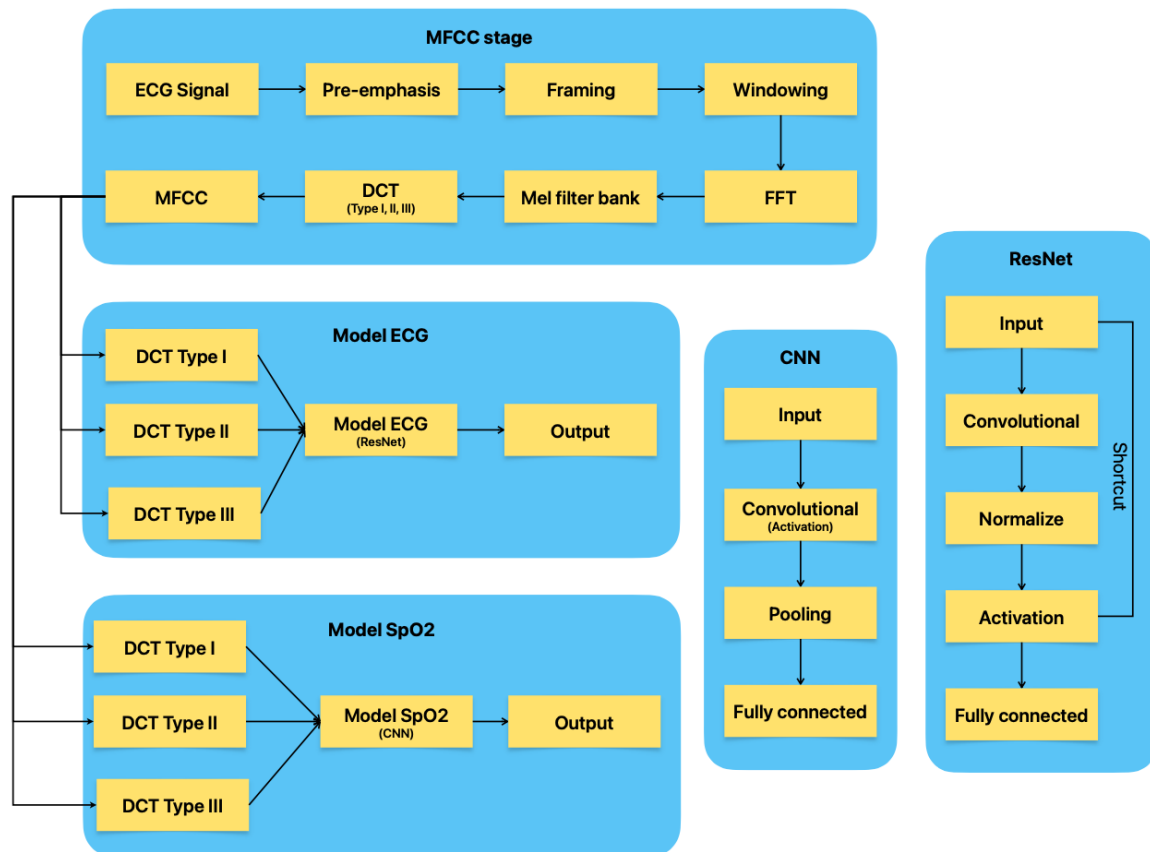
Chúng tôi đề xuất áp dụng hai mô hình riêng biệt cho hai loại tín hiệu ECG và SpO2:



Hình 1. Sơ đồ tổng mô hình chẩn đoán hội chứng ngưng thở khi ngủ do tắc nghẽn.

Hệ thống thu thập và hiển thị dữ liệu y tế gồm hai phương thức truyền dữ liệu từ thiết bị trực tiếp đến hệ thống lưu trữ và phân tích thông tin. Bằng phương thức sử dụng giao tiếp truyền thông không dây, dữ liệu truyền từ các thiết bị IoT được lưu trữ lên các dịch vụ đám mây (cloud). Dữ liệu sau khi thu thập sẽ được tổng hợp và lưu trữ thành các tệp, đưa vào mô hình xử lý dữ liệu chuyên sâu để hỗ trợ chẩn đoán hoặc đưa cho bác sĩ để giúp điều trị bệnh nhân. Tín hiệu đầu vào bao gồm 2 tín hiệu ECG và SpO2 thu được từ thiết bị, các tín hiệu này được tiền xử lý như lọc nhiễu, chuẩn hoá và chuyển đổi tương tự - số để phù hợp với quá trình xử lý trên máy tính. Đối với tín hiệu ECG, sau khi qua quá trình số hoá sẽ trải qua quá trình xuất trích đặc trưng Mel Frequency Cepstral Coefficient MFCC. Các tín hiệu ECG sẽ được nhấn mạnh ở các tần số cao, sau đó sẽ được chia khung để phân tích (Framing, Windowing) và áp dụng các phép biến đổi và nhấn mạnh các tần số quan trọng (FFT, Mel-filterbank, DCT). Đầu ra của MFCC sẽ được đưa vào CNN: Mỗi bộ lọc nhỏ sẽ tích chập để phân tích các mẫu hình đặc trưng và được gộp lại để bảo toàn các tín hiệu

mẫu chốt. Đối với tín hiệu SpO2, sau được tiền xử lý lọc nhiễu, chuẩn hoá và số hoá, tín hiệu sẽ được đưa trực tiếp vào CNN thì không cần trải qua quá trình rút trích đặc trưng. Sau khi đã phân tích cả hai tín hiệu đầu vào, mô hình sẽ đưa ra chẩn đoán cho bệnh nhân là có bệnh hoặc không có bệnh. Cả hai tín hiệu, ECG và SpO2, sau khi được xử lý tiền xử lý và chuyển đổi, đều được đưa vào cho mô hình học sâu để giúp tăng cường khả năng phân tích và chẩn đoán hội chứng ngưng thở khi ngủ do tắc nghẽn (OSA).



Hình 2. Sơ đồ khối chi tiết mô hình học sâu

Với tín hiệu ECG (Mô hình 1): Đầu tiên, phép biến đổi MFCC được áp dụng để trích xuất các biên độ phổ công suất từ sóng ECG. Tín hiệu ECG đặc trưng bởi các đoạn sóng lên xuống tương tự như sóng âm thanh. Việc sử dụng MFCC không chỉ giúp giảm số lượng đầu vào cho mô hình, giúp cho mô hình chạy nhanh hơn và tiết kiệm được bộ nhớ (CPU) mà còn loại bỏ các tín hiệu dư thừa. Thêm vào đó, đạo hàm của MFCC cũng được thêm vào để tạo ra các giá trị đặc trưng duy nhất cho mỗi đoạn tín hiệu. Sau khi áp dụng MFCC, tín hiệu được chuyển thành dạng ma trận và đầu ra của MFCC được xem như một hình ảnh, từ đó áp dụng cấu trúc CNN (Convolutional Neural Network) để trích xuất các “mẫu hình” đặc trưng nhằm phân tích và đưa ra kết quả cuối cùng.

5. 1 Mel-frequency Cepstral Coefficients (MFCC)

MFCC là một phương pháp phổ biến để trích xuất các đặc trưng âm thanh, đặc biệt trong nhận dạng giọng nói và phân tích tín hiệu âm thanh. MFCC hoạt động bằng cách chuyển đổi tín hiệu âm thanh từ miền thời gian sang miền tần số và trích xuất các đặc trưng dựa trên thang Mel, một thang tần số dựa trên cách con người cảm nhận âm thanh. [14, 15]

Dựa vào đó, tín hiệu ECG được đưa qua bộ lọc High-pass để khuếch đại các tần số cao. Điều này giúp cân bằng phổ âm và làm nổi bật các thành phần tần số cao, vốn thường bị yếu đi so với tần số thấp. Ta có phương trình $y[n] = x[n] - \alpha \cdot x[n-1]$, với α thường bằng 0.95. Tín hiệu ECG thường không ổn định trong thời gian dài, vì vậy tín hiệu được chia thành các khung nhỏ (frames) với thời lượng thường từ 20-40ms. Mỗi khung chứa một tập hợp dữ liệu tín hiệu được coi là gần như ổn định. Sau khi chia tín hiệu thành các khung, mỗi khung được nhân với một hàm cửa sổ (Windowing function) để giảm thiểu các lỗi ở biên khung. Hàm Hamming là hàm cửa sổ thường được sử dụng, được định nghĩa là:

$$w[n] = 0.54 - 0.46 \cdot \cos(2\pi nN - 1)$$

Mục đích của việc này là tránh hiện tượng spectral leakage, tức là làm giảm sự bất liên tục ở rìa các khung. Mỗi khung sau khi nhân cửa sổ sẽ được biến đổi Fourier để chuyển tín hiệu từ miền thời gian sang miền tần số. FFT Fourier Transform (Biến đổi Fourier nhanh - FFT) tạo ra một phổ biên độ (amplitude spectrum), giúp xác định tần số của các thành phần trong tín hiệu ECG. Thang tần Mel dựa trên cách con người cảm nhận ECG, với độ nhạy cao hơn ở các tần số thấp và giảm dần ở tần số cao. Vì thế, phổ tần số được chuyển sang thang Mel. Tín hiệu sẽ được đưa qua một loạt các bộ lọc tam giác phân bố đều trên thang Mel. Những bộ lọc này tập trung ở các tần số mà con người nhạy cảm hơn. Công thức chuyển đổi từ tần số thông thường fff sang thang Mel:

$$m = 2595 \log_{10}(1 + f/700)$$

Kết quả của bước này là một tập hợp các Mel spectrum, biểu diễn độ mạnh của tín hiệu trên thang Mel. Áp dụng hàm logarit lên kết quả của Mel filter bank để chuyển từ miền tuyến tính sang miền logarit. Điều này giúp mô hình hóa cách con người cảm nhận cường độ ECG, vì thính giác của con người phản ứng theo thang logarit đối với âm lượng. DCT Discrete Cosine Transform (Biến đổi Cosin rời rạc - DCT) được áp dụng lên Mel spectrum để nén các thông tin lại và tạo ra các hệ số MFCC. Các hệ số đầu tiên đại diện cho thông tin quan trọng nhất (giống như các tần số thấp trong FFT). DCT loại bỏ sự tương quan giữa các giá trị và giữ lại các đặc trưng cốt lõi của tín hiệu. Các giá trị MFCC thu được là các hệ số Cepstral đầu tiên, mô tả hình dạng của phổ tín hiệu và thường từ 12 đến 13 hệ số đầu tiên được giữ lại. Các hệ số này là đại diện nén gọn của phổ tín hiệu và chứa thông tin về các nét đặc trưng của tín hiệu. MFCC được sử dụng rộng rãi trong nhận diện giọng nói, âm nhạc và nhiều ứng dụng liên quan đến ECG vì nó giúp trích xuất các đặc trưng ECG mà hệ thống có thể sử dụng để phân loại hoặc nhận diện. Tương tự ECG, tín hiệu ECG cũng có dạng sóng và cũng đặc trưng bởi cái tần số nên rất thích hợp để dùng MFCC để nhận diện các “mẫu hình” đặc trưng của bệnh.

5.2 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) là một loại mạng nơ-ron nhân tạo phổ biến trong xử lý ảnh và nhận diện mẫu. CNN được thiết kế để học các đặc trưng từ dữ liệu hình ảnh thông qua các lớp tích chập (convolutional layers) [16]. Dưới đây là cách hoạt động cơ bản của một CNN:

Dữ liệu đầu vào là một hình ảnh, được biểu diễn dưới dạng ma trận của các pixel. Lớp tích chập áp dụng các bộ lọc (filters) nhỏ lên hình ảnh đầu vào. Kernel (bộ lọc) là một ma trận nhỏ có kích thước như 3x3 hoặc 5x5, trượt qua toàn bộ hình ảnh và tính toán tích chập giữa kernel và vùng tương ứng của ảnh. Kết quả này tạo ra một ma trận đặc trưng gọi là feature map. Tính năng trích xuất: Bộ lọc này học các đặc trưng khác nhau như cạnh, đường cong, hoặc cấu trúc phức tạp hơn qua các lớp chồng nhau. Sau mỗi lớp tích chập, một hàm kích hoạt thường được áp dụng, phổ biến nhất là ReLU (Rectified Linear Unit). ReLU giữ lại các giá trị dương và loại bỏ các giá trị âm, giúp mô hình học các đặc trưng phi tuyến tính. Lớp gộp giảm kích thước không gian của feature map nhưng vẫn giữ lại các đặc trưng quan trọng nhất. Max Pooling là phương pháp phổ biến nhất, lấy giá trị lớn nhất trong một vùng cụ thể (ví dụ: 2x2) để giảm độ phân giải mà

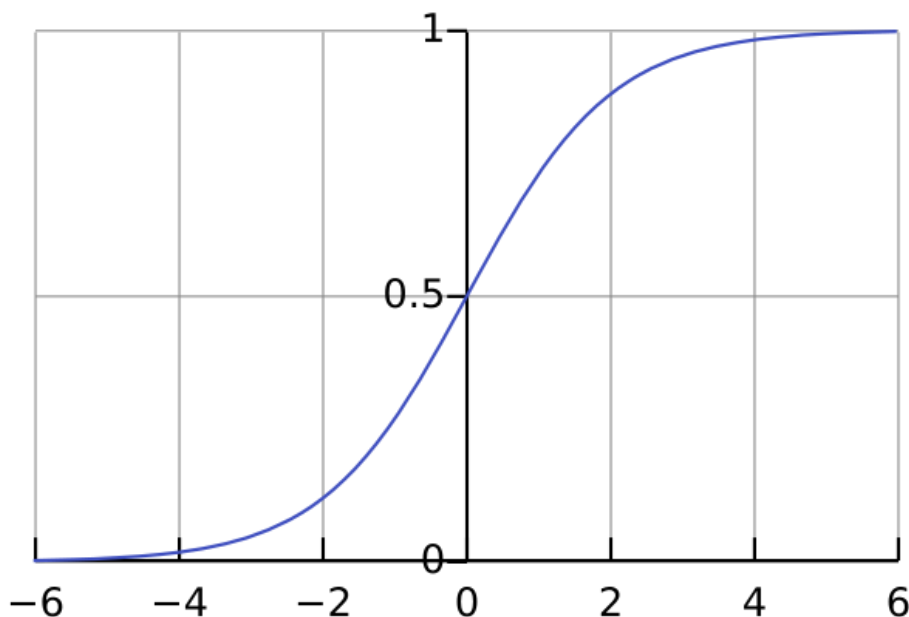
không làm mất thông tin quan trọng. Sau một số lớp tích chập và gộp, ta có được các feature maps có kích thước nhỏ gọn hơn, nhưng chứa thông tin chi tiết về ảnh. Các đặc trưng này được chuyển thành một vector và đi qua lớp fully connected. Ở đây, mọi nơ-ron đều kết nối với nhau như trong mạng nơ-ron truyền thống (Fully Connected Neural Network). Lớp cuối cùng sẽ sử dụng các đặc trưng học được để dự đoán. Ở lớp cuối cùng sử dụng hàm kích hoạt sigmoid để đảm bảo đầu ra nằm trong khoảng 0.0 đến 1.0. CNN đặc biệt hiệu quả trong các bài toán thị giác máy tính như phân loại ảnh, phát hiện đối tượng, và nhận diện khuôn mặt nhờ khả năng học tự động các đặc trưng từ dữ liệu thô.

Với tín hiệu SpO2 (Mô hình 2): Đối với tín hiệu SpO2 trong mô hình 2, do có hai lớp phân loại (có bệnh và không bệnh) có các “mẫu hình” khác nhau một cách rõ rệt, khiến cho việc sử dụng các phương pháp trích xuất đặc trưng phức tạp trở nên không cần thiết. Điều này khiến nghiên cứu tập trung vào việc trực tiếp áp dụng CNN để tự động học các đặc trưng từ dữ liệu đầu vào. CNN có khả năng tự động phát hiện các “mẫu hình” đặc trưng trong dữ liệu mà không yêu cầu sự can thiệp trực tiếp vào quá trình trích xuất. Khi áp dụng vào tín hiệu SpO2, từng đoạn tín hiệu sẽ được xử lý dưới dạng ma trận đầu vào và đi qua các lớp tích chập (convolutional layers), trong đó các bộ lọc (filters) sẽ học cách phát hiện các mẫu hình đặc trưng của tín hiệu tương ứng với hai lớp “có bệnh” và “không bệnh”.

5.3 Các thuật toán tối ưu hóa mô hình

Với cả 2 mô hình, hàm kích hoạt cho đầu ra đều sử dụng hàm sigmoid:

$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$



5.3.1 Hàm định lượng sai số

Sau đó sử dụng công thức Binary Cross Entropy để tính toán sai số của mô hình [17]:

$$- \frac{1}{N} \sum_{i=1}^N [y_i \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i))]$$

Sau lần mỗi dữ liệu được đưa vào mô hình để học, chúng tôi áp dụng thuật toán ADAM (Adaptive Momentum Estimation) để cập nhật các “cân nặng” của mô hình [18, 19]

5.3.2 Phương pháp cập nhật tham số

5.3.2.1 Quy luật cập nhật:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2$$
$$m_{t+1} = m_t + m_t^2 \quad v_{t+1} = v_t + v_t^2$$

m_t : là ước tính sai lệch thời điểm đầu tiên (trung bình của độ dốc),

v_t : là ước tính sai lệch mô men thứ hai (phương sai của độ dốc).

Công thức chỉnh sửa tham số thiên vị:

$$m_t = 1 - m_{t-1}$$
$$v_t = v_{t-1} - 2t$$

5.3.2.2 Công thức cuối cùng:

$$\theta = \theta - \eta m_t$$

t : Cân nặng tại bước thứ t

Tốc độ học (từ 0.0 \Rightarrow 1.0), model sử dụng tốc độ học 0.001

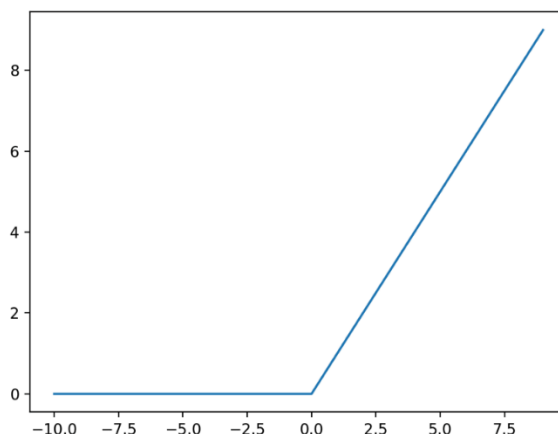
Epsilon, thường rất nhỏ, được dùng để tránh trường hợp chia cho 0

5.4 ResNet (Residual Network)

5.4.1 Residual Block

Input: Tầng đầu vào có thể là một vector hoặc hình ảnh, tùy thuộc vào kích thước của mô hình. Layer 1 (Convolution): Một lớp tích chập (Convolutional Layer) được áp dụng lên đầu vào với một số bộ lọc, kích thước kernel (thường là 3x3), và stride (thường là 1). Batch Normalization: Batch Normalization giúp chuẩn hóa các đặc trưng, ổn định quá trình huấn luyện. Activation (ReLU): Hàm kích hoạt ReLU được áp dụng để giới hạn giá trị của output trong khoảng $[0, \infty)$, sử dụng hàm ReLU là vì nó là một hàm phi tuyến tính, giúp tăng độ phức tạp cho mô hình.

$$\text{Hàm ReLU } f(x) = \max(0, x)$$



Layer 2 (Convolution): Một lớp tích chập khác được áp dụng (thường là 3x3 kernel). Batch Normalization: Batch Normalization giúp duy trì ổn định gradient khi truyền ngược. Skip Connection (kết nối tắt): Kết quả của tầng convolution sẽ được cộng với đầu vào ban đầu trước khi qua hàm kích hoạt ReLU, tạo ra "residual" giúp thông tin không bị mất khi truyền qua các tầng. Output: Kết quả từ tầng cộng.

5.4.2 Kiến trúc tổng quát của ResNet

ResNet gồm nhiều lớp residual block xếp chồng lên nhau. Cấu trúc cơ bản của một ResNet như sau: Input Layer: Lớp tích chập ban đầu với kernel lớn (thường là 7×7) và stride 2 để giảm kích thước đầu vào. Pooling Layer: Thường là lớp Max Pooling với kích thước kernel 3×3 , stride 2. Residual Blocks: Chuỗi các residual blocks với số lượng thay đổi dựa trên độ sâu của ResNet (ví dụ, ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet-152). Fully Connected Layer: Sau khi qua toàn bộ residual blocks, các đặc trưng được nén lại và qua một tầng fully connected để xuất kết quả cuối cùng.

5.4.3 Đặc điểm quan trọng của ResNet

Skip Connection: Truyền thẳng đầu vào qua các tầng giúp giảm thiểu việc mất mát thông tin, tránh hiện tượng "vanishing gradient". Bottleneck Block: Ở các phiên bản sâu hơn, các phép tích chập 1×1 giảm số chiều của vector đặc trưng trước khi qua các lớp tiếp theo, giảm chi phí tính toán.

5.5 Quá trình thí nghiệm

Thí nghiệm được thực hiện trên máy chủ huấn luyện với cấu hình gồm CPU Intel i7-1068NG7, xung nhịp 2.30GHz, GPU Intel Iris Plus Graphics, Ram 16GB. Khởi đầu với 3 lớp CNN, mỗi lớp có $2^3 = 4$ filters và kích thước bộ lọc 3×3 , từ từ tăng dần lên 2^4 , 2^5 , 2^6 để tăng dần chiều sâu cho mô hình. Epochs cố định là 5 và learning rate cố định là 0.001 với mỗi lần kiểm tra chéo.

Thời gian huấn luyện được trình bày trong bảng 1.

Bảng 1. Thời gian huấn luyện mô hình

Huấn luyện mô hình ECG	19 Phút
Huấn luyện mô hình SpO2	2 Phút
Test mô hình ECG	1 Giờ 40 Phút
Test mô hình SpO2	11 Phút

6. Kết quả

Việc sử dụng thêm Bidirectional LSTM vào mô hình không phù hợp với mục tiêu dự đoán trên đoạn tín hiệu ngắn lẫn dài mà còn tăng độ phức tạp cho mô hình. Điều này dẫn đến việc kéo dài thời gian học do mỗi bước cần nhiều thời gian tính toán hơn, đồng thời tiêu thụ nhiều tài nguyên CPU và bộ nhớ. Việc tăng độ phức tạp không chỉ không cải thiện hiệu suất, lãng phí tài nguyên mà còn làm giảm hiệu quả cho tổng thể mô hình. Nên chúng tôi quyết định sẽ không áp dụng BiLSTM.

6.1. Kết quả thống kê của mô hình ECG sử dụng CNN

Mô hình có hiệu năng ổn định, dao động từ 86- 89%, độ chính xác trên mỗi class gần như tương đương. Tuy vậy có thể thấy tại vào bộ test, mô hình sẽ có sự chênh lệch độ chính xác giữa hai lớp nhưng độ chính xác trung bình không đổi (precision và recall biến thiên, accuracy không đổi) dẫn đến việc mô hình đôi lúc sẽ có sự thiên vị cho một lớp nhất định và sẽ đánh đổi độ chính xác của lớp này cho lớp kia.

Bảng 2. Độ chính xác của các lớp sử dụng ECG

		Precision	Recall	F1-Score	Accuracy	Support cases
Fold 1	No OSA	94%	77%	85%	86%	382
	OSA	82%	95%	88%		409
Fold 2	No OSA	90%	89%	89%	89%	421
	OSA	87%	89%	88%		370
Fold 3	No OSA	88%	87%	88%	87%	406
	OSA	87%	87%	87%		385
Fold 4	No OSA	92%	77%	84%	86%	391
	OSA	81%	94%	87%		400

6.2. Kết quả thống kê của mô hình SpO2 sử dụng CNN

Mô hình biểu diễn tốt trên cả 2 lớp với độ chính xác xấp xỉ 96% , tuy vậy trong một vài lần thí nghiệm mô hình thể hiện không tốt lắm với các trường hợp có bệnh OSA. Tổng quan, đây là một mô hình tiềm năng nhưng tuy nhiên dữ liệu còn ít và không mang tính tổng quát.

Bảng 3. Độ chính xác của các lớp sử dụng SpO2

		Precision	Recall	F1-Score	Accuracy	Support cases
Fold 1	No OSA	94%	96%	95%	95%	188
	OSA	96%	93%	95%		179
Fold 2	No OSA	96%	92%	94%	94%	174
	OSA	93%	96%	95%		193
Fold 3	No OSA	95%	97%	96%	96%	173
	OSA	97%	95%	96%		194
Fold 4	No OSA	95%	97%	96%	96%	173
	OSA	97%	96%	96%		194

Chú thích:

Precision: $\frac{\text{Tổng trường hợp chẩn đoán đúng}}{\text{Tổng trường hợp thực sự có}} \times 100$

Recall: $\frac{\text{Tổng trường hợp chẩn đoán đúng}}{\text{Tổng trường hợp được chẩn đoán}} \times 100$

Support cases: Số trường hợp của lớp đó

$$\text{F1-Score: } 2 * (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

6.3 Kết quả khi kết hợp hai mô hình bằng công thức weighted sum

Đây là kết quả trên 8 người có chỉ số phụ SpO2, mô hình ECG đã được học trước từ 27 người trước đó. Phương pháp kiểm tra áp dụng k-fold cross validation.

Công thức weighted sum có dạng: $W_e * Pred_e + W_s + Pred_s$

- **Pred_e**: kết quả khi sử dụng ECG
- **Pred_s**: kết quả khi sử dụng SpO2
- **W_e**: trọng số của kết quả ECG
- **W_s**: trọng số của kết quả SpO2

Fold 1

Train set: Apnea cases [1]: 1286 - Normal cases [0]: 1747

Test set: Apnea cases [1]: 180 - Normal cases [0]: 254

Bảng 4. Độ chính xác của mô hình kết hợp ECG và SpO2_Fold1

W_ECG	W_SpO2	Accuracy
0.0	1.0	95.85%
0.2	0.8	96.08%
0.4	0.6	96.54%
0.5	0.5	96.94%
0.6	0.4	97.09%
0.8	0.2	97.19%
1.0	0.0	97.07%

Fold 2

Train set: Apnea cases [1]: 1287 - Normal cases [0]: 1746

Test set: Apnea cases [1]: 179 - Normal cases [0]: 255

Bảng 5. Độ chính xác của mô hình kết hợp ECG và SpO2_Fold2

W_ECG	W_SpO2	Accuracy
0.0	1.0	97.11%
0.2	0.8	97.20%
0.4	0.6	97.35%
0.5	0.5	97.48%
0.6	0.4	97.54%
0.8	0.2	97.51%
1.0	0.0	97.46%

Fold 3

Train set: Apnea cases [1]: 1288 - Normal cases [0]: 1745

Test set: Apnea cases [1]: 178 - Normal cases [0]: 256

Bảng 6. Độ chính xác của mô hình kết hợp ECG và SpO2_Fold3

W_ECG	W_SpO2	Accuracy
0.0	1.0	97.32%
0.2	0.8	97.23%
0.4	0.6	97.22%
0.5	0.5	97.22%
0.6	0.4	97.19%
0.8	0.2	97.17%
1.0	0.0	97.14%

Fold 4

Train set: Apnea cases [1]: 1296 - Normal cases [0]: 1765

Test set: Apnea cases [1]: 197 - Normal cases [0]: 236

Bảng 7. Độ chính xác của mô hình kết hợp ECG và SpO2_Fold4

W_ECG	W_SpO2	Accuracy
0.0	1.0	97.14%
0.2	0.8	97.15%
0.4	0.6	97.16%
0.5	0.5	97.15%
0.6	0.4	97.03%
0.8	0.2	96.73%
1.0	0.0	96.36%

Độ chính xác tương đối khoảng 97% cho hầu hết mọi lần test. Độ chính xác cao nhất khi nghiêng về ECG trong 2 folds đầu và SpO2 trong 2 folds cuối nhiều hơn, tuy vậy sự chênh lệch không cao $\approx 0.1\%$. Điều này cho thấy hai chuẩn đoán của hai mô hình làm tăng đáng kể độ chính xác so với khi hai mô hình hoạt động riêng lẻ, đồng thời tín hiệu ECG và tín hiệu SpO2 có giá trị chẩn đoán cao: ECG tăng từ 87% - 97% và SpO2 tăng từ 96% -> 97%. Do đó chúng tôi quyết định kết quả ra được của mỗi mô hình sẽ có cùng giá trị chẩn đoán theo tỉ lệ 50:50.

7. Thảo luận và kết luận

Nghiên cứu của chúng tôi hướng đến việc phát triển một mô hình lai tạo mới từ việc áp dụng phương pháp trích xuất đặc trưng của âm thanh MFCC và một loại mạng nơ-ron nhân tạo phổ biến trong xử lý ảnh và nhận diện mẫu CNN để phân tích các tín hiệu điện tâm đồ ECG và độ bão hoà oxy SpO2 và chẩn đoán bệnh cho bệnh nhân. Qua nghiên cứu, chúng tôi đã chứng minh được tiềm năng của mô hình chẩn đoán OSA. Với mô hình sử dụng tín hiệu ECG cho thấy độ chính xác dao động từ 81% - 83%, trong khi mô hình sử dụng tín hiệu SpO2 có độ chính xác từ 95% - 96%. Đặc biệt thay khi kết hợp 2 mô hình chỉ ra độ chính xác cao lên đến 94% - 95%. Những kết quả này khẳng định rằng khả năng ứng dụng của mô hình chúng tôi đề xuất như một công cụ hỗ trợ chẩn đoán OSA chính xác, hiệu quả và tiết kiệm chi phí hơn so với phương pháp PSG.

Có thể thấy cả hai mô hình hoạt động ổn định trên cả hai lớp, các độ chính xác trên mỗi thang đo không chênh lệch nhau quá nhiều chứng tỏ sự hài hòa của chuẩn đoán của mô hình với trường hợp có bệnh lẫn không bệnh. Khi kết hợp cả hai mô hình và áp dụng weighted sum thì độ chính xác được đẩy lên đáng kể 96% - 97% dù cho khi hai mô hình nếu hoạt động riêng lẻ sẽ cho ra độ chính xác không cao. Độ chính xác giữa các lần kiểm tra chéo chênh lệch nhau rất nhỏ chỉ ra rằng cả hai loại tín hiệu không có độ thiên vị đáng kể. Một lần nữa khẳng định tính đúng đắn khi áp dụng kết hợp hai mô hình và giá trị chẩn đoán của hai tín hiệu ECG và SpO2. Tuy vậy, dữ liệu dùng để kiểm tra còn tương đối ít, không mang tính khái quát cao nên không đánh giá đúng hoàn toàn tính chính xác của cả hai mô hình. Ngoài ra, với tín hiệu ECG có phần tương đối phức tạp và sự khác biệt của hai lớp là không đáng kể nên mô hình tuy hoạt động ổn định nhưng có độ chính xác của mô hình không cao lắm 85%-87%. Có thể thấy lý do chính yếu là việc mô hình chưa thực sự “sâu” và có đủ độ phức tạp để thực thi trên các tín hiệu phức tạp vì máy chủ huấn luyện không đáp ứng nổi. Trong tương lai chúng tôi có thể đổi sang một máy chủ mạnh hơn và sử dụng các mô hình học sâu có cấu trúc phức tạp cao như ResNet 101, ResNet 152 để có thể nhận diện sự khác biệt dù là nhỏ nhất giữa 2 lớp, khi đó càng tăng độ chính xác khi kết hợp hai mô hình.

7.1 Hướng phát triển

Bên cạnh những ưu điểm nổi bật mang lại hiệu quả cao trong việc chẩn đoán, hệ thống vẫn còn tồn tại một số giới hạn. Những giới hạn này đang được chúng tôi nghiên cứu và hoàn thiện dần trong thời gian sắp tới. Để phát triển một hệ thống sử dụng trí tuệ nhân tạo chính xác và linh hoạt, chúng tôi sẽ cải thiện độ chính xác của mô hình bằng cách tối ưu hoá thuật toán và nâng cấp cảm biến để đảm bảo các chẩn đoán đạt độ tin cậy cao hơn. Tiếp theo là tích hợp các chức năng cảnh báo, báo động và gửi thông báo đến người dùng trong các tình huống khẩn cấp để kịp thời xử lý. Đồng thời, chúng tôi sẽ thiết kế ra thiết bị cầm tay nhỏ gọn để tăng tính linh hoạt và tiện dụng. Thiết bị sẽ được kết nối với thiết bị di động để cung cấp chẩn đoán tức thì và cá nhân hoá thông tin cũng như cấp tài khoản riêng cho từng người dùng, đảm bảo tính phù hợp và hiệu quả trong việc chăm sóc sức khỏe hoặc giám sát cá nhân.

8. REFERENCE

- [1] *Real-Time Obstructive Sleep Apnea Detection from Raw ECG and SpO2 Signal Using Convolutional Neural Network*. (2024). <https://arxiv.org/pdf/2408.16030>
- [2] Paul, T., Hassan, O., Islam, S. K., & Mosa, Abu. S. M. (2024). *Real-time obstructive sleep apnea detection from raw ECG and SPO2 signal using convolutional neural network*. AMIA Jt Summits Transl Sci Proc. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11141842/>
- [3] Chaw, H. T., Kamolphiwong, T., Kamolphiwong, S., Tawaranurak, K., & Wongtanawijit, R. (2023). ZleepNet: A deep convolutional neural network model for predicting sleep apnea using SPO2 Signal. *Applied Computational Intelligence and Soft Computing*, 2023, 1–12.

<https://doi.org/10.1155/2023/8888004>

[4] Shen, F., Cheng, S., Li, Z., Yue, K., Li, W., & Dai, L. (2020). Detection of Snore from OSAHS Patients Based on Deep Learning. *Journal of Healthcare Engineering*, 2020, 1–10.

<https://doi.org/10.1155/2020/8864863>

[5] Deng, M., Meng, T., Cao, J., Wang, S., Zhang, J., & Fan, H. (2020). Heart sound classification based on improved MFCC features and convolutional recurrent neural networks. National Natural Science Foundation of China. <https://doi.org/10.1016/j.neunet.2020.06.015>

[6] Jiang, Y., Peng, J., & Zhang, X. (2020). Automatic snoring sounds detection from sleep sounds based on deep learning. *Physical and Engineering Sciences in Medicine*, 43(2), 679–689.

<https://doi.org/10.1007/s13246-020-00876-1>

[7] P. K. B. V. Đ. H. Y. (2022, April 14). ĐO ĐA KÝ GIÁC NGỦ (Polysomnography - PSG). *Bệnh Viện Đại Học Y Dược Tp.HCM*. <https://umcclinic.com.vn/dich-vu/tam-soat-chuc-nang-ho-hap/do-da-ky-giac-ngu-polysomnography-psg>

[8] Leino, A. et al (2021). *Neural network analysis of nocturnal SpO2 signal enables easy screening of sleep apnea in patients with acute cerebrovascular disease*. <https://doi.org/10.1016/j.sleep.2020.12.032>

[9] ElMoaqet, H., Eid, M., Glos, M., Ryalat, M., & Penzel, T. (2020). Deep Recurrent Neural Networks for Automatic Detection of Sleep Apnea from Single Channel Respiration Signals. *Sensors*, 20(18), 5037.

<https://doi.org/10.3390/s20185037>

[10] *Understanding subtitles by Character-Level Sequence-to-Sequence learning*. (2017, April 1). IEEE Journals & Magazine | IEEE Xplore. <https://ieeexplore.ieee.org/abstract/document/7547277>

[11] Sak, H., Senior, A., & Beaufays, F. (2014). *LONG SHORT-TERM MEMORY BASED RECURRENT NEURAL NETWORK ARCHITECTURES FOR LARGE VOCABULARY SPEECH RECOGNITION*.

<https://arxiv.org/pdf/1402.1128>

[12] Xie, J., Aubert, X., Dijik, J. v., Arsenali, B., Fonseca, P., & Overeem, S. (2021). *Audio-based snore detection using deep neural networks*. ScienceDirect. <https://doi.org/10.1016/j.cmpb.2020.105917>

[13] Ittichaichareon, C., Suksri, S., & Yingthawornsuk, T. (2012). *Speech recognition using MFCC*. International Conference on Computer Graphics, Simulation and Modeling (ICGSM'2012).

[14] Wikipedia contributors. (2024, July 9). *Mel-frequency cepstrum*. Wikipedia.

https://en.wikipedia.org/wiki/Mel-frequency_cepstrum

[15] Ilyamich. (2022, November 12). *MFCC implementation and tutorial*. Kaggle.

<https://www.kaggle.com/code/ilyamich/mfcc-implementation-and-tutorial>

[16] Ibm. (2024, October 2). *Convolutional Neural Networks*. IBM,

<https://www.ibm.com/topics/convolutional-neural-networks>

[17] Shipra_Saxena. (2024, October 10). *Binary Cross Entropy/Log loss for binary classification*.

Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2021/03/binary-cross-entropy-log-loss-for-binary-classification/>

[18] DeepLearningAI. (2017, August 25). *Adam Optimization Algorithm (C2W2L08)* [Video]. YouTube.

https://www.youtube.com/watch?v=JXQT_vxqwIs

[19] Tanner, G. (n.d.). *Adaptive Moment Estimation (ADAM)*. Machine Learning Explained. <https://ml-explained.com/blog/adam-explained>

[20] *Apnea-ECG Database V1.0.0*. (2000, February 10). <https://physionet.org/content/apnea-ecg/1.0.0/>

[21] Huyền V. (2024, June 11). Nguy kích vì tâm lý “ngại” đi khám bệnh. *Báo Giáo Dục Và Thời Đại Online*. <https://giaoducthoidai.vn/nguy-kich-vi-tam-ly-ngai-di-kham-benh-post687099.html>

9. Lời cảm ơn

Chân thành cảm ơn Chương trình Khoa học Kỹ thuật Liên ngành trường Phổ thông Năng khiếu – ĐHQG-HCM, Ban Giám hiệu trường Phổ thông Năng khiếu – ĐHQG-HCM, Thầy Nguyễn Hải Đăng, và Cha mẹ đã luôn hướng dẫn, đồng hành và động viên trong suốt quá trình thực hiện dự án.

10. Đóng góp của tác giả

Bùi Minh Triết: Phát triển phần mềm thuật toán, phát triển phương pháp nghiên cứu, phát triển phần cứng, xử lý số liệu.

Nông Nguyễn Tường Vy: Đóng góp vào ý tưởng chính của bài nghiên cứu, phát triển phương pháp nghiên cứu, tìm kiếm thông tin cho bài nghiên cứu, phát triển phần cứng, viết bản thảo ban đầu và hoàn thiện báo cáo.