

ỨNG DỤNG MÔ HÌNH TRÍ TUỆ NHÂN TẠO QUẢN LÝ BỆNH NHÂN RỐI LOẠN GIỌNG

Lê Minh Đạt^{1,✉}, Phạm Thị Bích Đào¹, Nguyễn Thị Hằng¹
Đỗ Tiến Lộc¹, Nguyễn Thị Anh Đào¹, Nguyễn Quý Đôn¹
Nguyễn Thị Xuân Hoà¹, Nguyễn Diệu My¹, Nguyễn Mạnh Hưng¹
Phan Xuân Nam²

¹Bệnh viện Đại học Y Hà Nội

²Bệnh viện Đa khoa tỉnh Quảng Trị

Nghiên cứu thực hiện tại Bệnh viện Đại học Y Hà Nội từ năm 2024 đến năm 2025 nhằm đánh giá hiệu quả mô hình trí tuệ nhân tạo (AI-VoiceCare) quản lý bệnh nhân rối loạn giọng. 312 bệnh nhân được chẩn đoán rối loạn giọng. Mô hình AI sử dụng mạng nơ-ron tích chập (Convolutional Neural Network - CNN) phân tích phổ Mel, tích hợp hệ thống AI-VoiceCare quản lý hồ sơ điện tử ngoại trú. Kết quả thấy mô hình đạt độ chính xác 92,1%, độ nhạy 0,90, đặc hiệu 0,93 trong phân loại giọng bệnh lý; thời gian xử lý hồ sơ giảm 61,1%, sai sót ghi chép giảm 60,9%. Sau sáu tháng triển khai, tỷ lệ bệnh nhân tái khám đúng hẹn tăng từ 48,7% lên 76,3%, 68,5% bệnh nhân ghi nhận cải thiện rõ rệt về giọng sau trị liệu có hỗ trợ AI. 84,3% bác sĩ đánh giá hệ thống “hữu ích” hoặc “rất hữu ích”. Ứng dụng AI-VoiceCare giúp nâng cao hiệu quả chẩn đoán, quản lý và phục hồi giọng, hướng tới mô hình chăm sóc cá thể hóa cho người bệnh.

Từ khóa: Rối loạn giọng, trí tuệ nhân tạo, AI-VoiceCare, phân tích giọng nói.

I. ĐẶT VẤN ĐỀ

Rối loạn giọng là một trong những nguyên nhân phổ biến khiến bệnh nhân tìm đến chuyên khoa Tai Mũi Họng, với tỷ lệ mắc ước tính từ 3 - 9% dân số toàn cầu.^{1,2} Mặc dù, không đe dọa tính mạng, rối loạn giọng gây ảnh hưởng nghiêm trọng tới giao tiếp, công việc và chất lượng cuộc sống.³ Các triệu chứng khàn tiếng, hụt hơi hoặc cảm giác vướng họng thường bắt đầu nhẹ và tiến triển chậm, khiến nhiều bệnh nhân chủ quan, chỉ đến khám khi tình trạng đã nặng hoặc dai dẳng.⁴

Ở giai đoạn đầu, rối loạn giọng thường thuộc dạng rối loạn chức năng (Functional Disorders - FD), đây là nhóm rối loạn giọng nói

không bắt nguồn từ tổn thương cấu trúc giải phẫu của dây thanh hay thanh quản, mà xuất phát từ sự sử dụng giọng không đúng cách, thói quen phát âm sai, hoặc điều hòa cơ - thần kinh bất thường trong quá trình tạo giọng. Đây là nhóm rối loạn có thể đáp ứng tốt với can thiệp sớm như liệu pháp phát âm và thay đổi hành vi.⁵ Tuy nhiên, nếu để kéo dài, tổn thương dây thanh có thể tiến triển thành thực thể như polyp, hạt, xơ hoặc viêm mạn tính, đòi hỏi điều trị phối hợp (nội khoa, phẫu thuật và phục hồi âm).^{6,7} Quản lý những bệnh nhân này từ khi phát hiện sớm đến theo dõi lâu dài gặp rất nhiều khó khăn như tỷ lệ tuân thủ tái khám của bệnh nhân thấp, khó giám sát tiến triển lâm sàng và hiệu quả điều trị, đặc biệt khi người bệnh cảm thấy triệu chứng nhẹ gần như không ảnh hưởng đến sinh hoạt hằng ngày.^{8,9} Trong bối cảnh công nghệ thông tin - điện tử y tế - phát triển nhanh chóng, hệ thống lưu trữ

Tác giả liên hệ: Lê Minh Đạt

Bệnh viện Đại học Y Hà Nội

Email: leminhdatt96@gmail.com

Ngày nhận: 23/10/2025

Ngày được chấp nhận: 15/12/2025

hồ sơ điện tử và kết nối dữ liệu y tế cung cấp nền tảng kỹ thuật để ứng dụng các mô hình AI vào chăm sóc bệnh nhân.^{10,11} Các nghiên cứu quốc tế cho thấy mô hình học sâu như mạng nơ-ron tích chập (Convolutional Neural Network - CNN) phân tích các đặc trưng âm học (Mel-spectrogram) có thể phân biệt giọng bình thường và giọng bệnh lý với độ chính xác cao (Area Under the Curve > 0,9).^{8,9,12} Các ứng dụng đa phương thức - kết hợp âm học, hình ảnh nội soi và hồ sơ lâm sàng - cũng đã chứng minh tiềm năng quản lý toàn diện bệnh lý thanh quản.^{13,14} Tại Việt Nam, mặc dù cơ sở công nghệ thông tin trong lĩnh vực y tế đã cải thiện, việc áp dụng AI cho quản lý rối loạn giọng vẫn còn hạn chế. Nghiên cứu này nhằm đánh giá hiệu quả bước đầu của mô hình AI-VoiceCare, tích hợp xử lý giọng nói và hồ sơ điện tử, trong thực hành lâm sàng quản lý bệnh nhân rối loạn giọng, hướng tới tiếp cận công nghệ tiên tiến để nâng cao chất lượng chăm sóc.

II. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP

1. Đối tượng

Đối tượng nghiên cứu gồm 312 bệnh nhân được chẩn đoán rối loạn giọng đến khám và điều trị tại các khoa khám bệnh - Bệnh viện Đại học Y Hà Nội trong thời gian từ tháng 01/2024 đến 09/2025.

Tiêu chuẩn lựa chọn

Bệnh nhân ≥ 18 tuổi, được chẩn đoán rối loạn giọng chức năng hoặc thực thể, có bản ghi giọng nói, hình ảnh nội soi thanh quản và hồ sơ lâm sàng đầy đủ.

Tiêu chuẩn loại trừ

Bệnh nhân có tổn thương thần kinh trung ương, phẫu thuật thanh quản trước đó, khiếm thính nặng hoặc không hợp tác trong quá trình ghi âm, nội soi hoặc trị liệu giọng. Các bệnh nhân bị viêm đường hô hấp cấp (viêm họng, viêm thanh quản cấp), ho cấp kéo dài, COPD

mất bù... vì có thể làm thay đổi giọng tạm thời.

2. Phương pháp

Thiết kế nghiên cứu

Can thiệp trước-sau (pre-post implementation) có ứng dụng công nghệ trí tuệ nhân tạo (AI).

Thời gian và địa điểm nghiên cứu

Thực hiện tại Bệnh viện Đại học Y Hà Nội từ 01/2024-09/2025.

Cỡ mẫu và chọn mẫu

Chọn toàn bộ 312 bệnh nhân đáp ứng tiêu chí trong thời gian nghiên cứu.

Bệnh nhân đã được chẩn đoán rối loạn giọng dựa trên:

- Triệu chứng cơ năng (khàn, hụt hơi, mệt khi nói...).
- Khám Tai Mũi Họng.
- Nội soi thanh quản (có/không tổn thương thực thể).
- Thang đánh giá giọng (nếu có dùng GRBAS/VHI-10/VHI... thì ghi rõ; nếu chưa dùng thì phải nêu “đánh giá bởi bác sĩ chuyên khoa theo thực hành thường quy” và đưa vào hạn chế).

Quy trình nghiên cứu:

- Thu thập hồ sơ bệnh án, bản ghi giọng nói (phát âm /a/, /i/, câu đọc chuẩn) từ người không có triệu chứng, nội soi thanh quản bình thường bình thường và bác sĩ đánh giá giọng bình thường. Dữ liệu từ các bản ghi đạt chất lượng.
- Mô hình mạng nơ-ron tích chập (CNN) được huấn luyện để phân tích phổ Mel và phân loại giọng bình thường - bệnh lý. Mô hình sử dụng mạng nơ-ron tích chập (CNN) với các lớp:
 - + lớp tích chập (convolution).
 - + lớp gộp (pooling).
 - + lớp kết nối đầy đủ (fully connected).

Dữ liệu được chia thành các tập huấn luyện, kiểm định và kiểm tra theo tỷ lệ phù hợp. Quá

trình huấn luyện sử dụng:

- + hàm mất mát cho phân loại nhị phân.
- + thuật toán tối ưu chuẩn.

+ dừng sớm (early stopping) để tránh quá khớp.

+ Nhãn giọng bình thường/gọng bệnh lý được gán dựa trên chẩn đoán lâm sàng và nội soi thanh quản bởi bác sĩ chuyên khoa Tai Mũi Họng.

- Hệ thống AI-VoiceCare tích hợp dữ liệu lâm sàng, âm học, nội soi và hồ sơ bệnh án điện tử nhằm hỗ trợ quản lý, theo dõi tiến triển và tái khám của bệnh nhân.

Quy trình tích hợp AI-VoiceCare vào thực hành ngoại trú theo các bước:

- Bệnh nhân được ghi âm giọng nói trong quá trình khám.

- Dữ liệu giọng được hệ thống tự động xử lý, trích xuất phổ Mel và phân tích bằng mô hình CNN.

- Kết quả phân loại giọng (bình thường/bệnh lý) được lưu vào hồ sơ bệnh án điện tử.

Thông tin này hỗ trợ bác sĩ trong đánh giá, theo dõi và quản lý bệnh nhân, đồng thời phục vụ thống kê và nhắc tái khám.

- Thời gian xử lý hồ sơ được định nghĩa là khoảng thời gian từ khi bắt đầu mở hồ sơ bệnh án ngoại trú đến khi hoàn tất nhập liệu và lưu trữ đầy đủ thông tin. Thời gian sẽ được ghi nhận trực tiếp trong quá trình làm việc thường quy. Giá trị được tính trung bình trên các hồ sơ của từng giai đoạn (trước AI và sau AI). Kết

quả được biểu diễn dưới dạng trung bình $\pm$ độ lệch chuẩn.

- Sai sót ghi chép hồ sơ bao gồm: thiếu các trường thông tin bắt buộc, nhập sai hoặc không nhất quán dữ liệu lâm sàng, nhầm lẫn phân loại chẩn đoán hoặc mã hóa.

- Sai sót được phát hiện thông qua rà soát hồ sơ định kỳ sau khám. Việc rà soát này được thực hiện bởi nhóm nghiên cứu và/hoặc bộ phận quản lý hồ sơ của bệnh viện theo quy trình thường quy.

- Cải thiện giọng được đánh giá sau quá trình theo dõi và điều trị có hỗ trợ AI-VoiceCare. Việc đánh giá do bác sĩ chuyên khoa Tai Mũi Họng và/hoặc chuyên viên âm ngữ trị liệu thực hiện trong quá trình tái khám. Với các tiêu chí cải thiện đánh giá như chất lượng giọng nói cải thiện rõ rệt so với thời điểm ban đầu, triệu chứng cơ năng giảm, giọng nói được bác sĩ đánh giá là tốt hơn so với trước điều trị.

Xử lý số liệu

Sử dụng phần mềm SPSS 26.0; thống kê mô tả, so sánh tỷ lệ bằng kiểm định Chi-square và so sánh trung bình bằng T-test, mức ý nghĩa $p < 0,05$.

3. Đạo đức nghiên cứu

Nghiên cứu tuân thủ các quy định về đạo đức trong nghiên cứu y sinh học, đảm bảo quyền riêng tư và bảo mật thông tin người bệnh. Tất cả bệnh nhân được giải thích mục tiêu nghiên cứu và ký cam kết đồng ý tham gia.

III. KẾT QUẢ

Bảng 1. Đặc điểm chung của nhóm bệnh nhân (n = 312)

Đặc điểm	Số lượng (n)	Tỷ lệ (%)
Nam	118	37,8
Nữ	194	62,2
Tuổi trung bình (năm)	44,6 $\pm$ 12,3	-

Đặc điểm	Số lượng (n)	Tỷ lệ (%)
Nghề nghiệp sử dụng giọng thường xuyên (giáo viên, MC, ca sĩ...)	102	32,7
Thời gian khàn tiếng trung bình (tuần)	10,2 ± 7,5	-

Phần lớn bệnh nhân là nữ trong độ tuổi lao động, gần một phần ba thuộc nhóm nghề nghiệp cần sử dụng giọng nhiều như giáo viên, giảng viên, phát thanh viên, MC, ca sĩ, diễn viên, nhân viên tổng đài, bán hàng, hướng dẫn viên... Thời gian khàn tiếng trung bình kéo dài,

phản ánh tình trạng đến Thời gian khàn tiếng trung bình kéo dài cho thấy xu hướng bệnh nhân đến khám sau một thời gian xuất hiện triệu chứng, tuy nhiên cần nghiên cứu thêm để khẳng định tình trạng đi khám muộn.

Bảng 2. Phân loại rối loạn giọng theo lâm sàng và mô hình AI-VoiceCare

Phân loại	Lâm sàng (n, %)	AI-VoiceCare (n, %)
Rối loạn chức năng	146 (46,8%)	148 (47,4%)
Rối loạn thực thể (u hạt, polyp, hạt xơ...)	114 (36,5%)	112 (35,9%)
Rối loạn phối hợp (chức năng + thực thể nhẹ)	52 (16,7%)	52 (16,7%)

Phân loại giữa AI và lâm sàng có sự tương đồng cao, chênh lệch < 1%, cho thấy mô hình AI-VoiceCare tái tạo tốt đánh giá của bác sĩ chuyên khoa.

Bảng 3. Hiệu năng mô hình trí tuệ nhân tạo (AI) trong phân loại giọng bệnh lý

Chỉ số đánh giá	Giá trị
Độ chính xác tổng thể	92,1%
Độ nhạy (Sensitivity)	0,90
Độ đặc hiệu (Specificity)	0,93
AUC (Area Under Curve)	0,94
Sai số trung bình tuyệt đối (MAE)	0,08

Mô hình trí tuệ nhân tạo đạt hiệu năng cao trong phân loại giọng bệnh lý với độ chính xác tổng thể 92,1%, độ nhạy 0,90 và độ đặc hiệu 0,93. Giá trị AUC đạt 0,94 cho thấy khả năng phân biệt xuất sắc giữa giọng bình thường và giọng bệnh lý, tương đương với các nghiên cứu quốc tế sử dụng học sâu trong nhận diện rối

loạn giọng. Các chỉ số này cho thấy mô hình có độ tin cậy cao và tiềm năng ứng dụng trong thực hành lâm sàng Sai số trung bình tuyệt đối (MAE) được tính trên giá trị xác suất dự đoán (0-1) của mô hình so với nhãn thực (0/1), phản ánh mức độ sai lệch trung bình của xác suất dự đoán. Tuy nhiên, trong bối cảnh phân loại nhị phân y sinh, MAE chỉ mang tính tham khảo và các chỉ số chính được sử dụng để đánh giá mô hình là độ nhạy, độ đặc hiệu và AUC. Các chỉ số này cho thấy mô hình có độ tin cậy cao và tiềm năng ứng dụng trong thực hành lâm sàng..

Sau khi được tích hợp vào hệ thống bệnh viện, Chúng tôi đã tiến hành đo Thời gian xử lý hồ sơ từ khi bắt đầu mở hồ sơ bệnh án ngoại trú đến khi hoàn tất nhập liệu và lưu trữ đầy đủ thông tin đồng thời sẽ rà soát hồ sơ định kỳ sau khám thực hiện bởi nhóm nghiên cứu và/hoặc bộ phận quản lý hồ sơ của bệnh viện theo quy trình thường quy. Số liệu này sẽ được dùng để đánh giá hiệu quả của ứng dụng AI-VoiceCare trong việc quản lý hồ sơ bệnh nhân.

Bảng 4. Hiệu quả quản lý hồ sơ bệnh nhân trước và sau khi ứng dụng AI-VoiceCare

Chỉ số	Trước AI (n = 156)	Sau AI (n = 156)	Thay đổi (%)	p-value
Thời gian xử lý hồ sơ (phút)	18,2 ± 4,1	7,1 ± 2,3	-61,1	< 0,001
Sai sót ghi chép hồ sơ (%)	12,3	4,8	-60,9	0,004
Mức độ hài lòng của bác sĩ (%)	32,7	84,3	+51,6	0,001

Các chỉ số trong Bảng 4 phản ánh hiệu quả quản lý hồ sơ bệnh nhân ngoại trú, bao gồm thời gian xử lý, tỷ lệ sai sót ghi chép và mức độ hài lòng của bác sĩ. Tỷ lệ sai sót giảm hơn một nửa cho thấy AI-VoiceCare không chỉ rút ngắn thời gian làm việc mà còn nâng cao độ chính xác và tính nhất quán của hồ sơ bệnh án. Đây là yếu tố quan trọng trong quản lý bệnh mạn tính và theo dõi lâu dài, nơi chất lượng dữ liệu đóng vai trò then chốt cho quyết định lâm sàng.

Ngoài ra việc giảm hơn 60% thời gian xử lý và từ số liệu thu thập được từ bảng đánh giá hệ thống AI-VoiceCare theo các mức độ từ không hữu ích đến rất hữu ích thì 84,3% số bác sĩ đều đánh giá hệ thống ở mức “hữu ích” hoặc “rất hữu ích”. với sự gia tăng đáng kể mức độ hài lòng của bác sĩ sau khi được hỏi thu thập bằng bảng hỏi., cho thấy AI-VoiceCare mang lại lợi ích rõ rệt trong thực hành lâm sàng.

Bảng 5. Hiệu quả theo dõi và phục hồi giọng ở bệnh nhân sau triển khai AI-VoiceCare

Chỉ số theo dõi	Trước triển khai	Sau 6 tháng triển khai	Thay đổi (%)
Tỷ lệ bệnh nhân tái khám đúng hẹn (%)	48,7	76,3	+27,6
Tỷ lệ bệnh nhân cải thiện rõ rệt giọng (%)	-	68,5	-
Bác sĩ đánh giá hệ thống “hữu ích” hoặc “rất hữu ích” (%)	-	84,3	-

Sau 6 tháng ứng dụng được triển khai ở viện, tỷ lệ tái khám đúng hẹn tăng mạnh. Dựa vào các đánh giá chất lượng giọng nói của bệnh nhân cải thiện rõ rệt so với thời điểm ban đầu, triệu chứng cơ năng giảm, và giọng nói được bác sĩ đánh giá là tốt hơn so với trước điều trị. Hơn hai phần ba bệnh nhân có cải thiện rõ rệt về giọng, và đa số bác sĩ ghi nhận AI-VoiceCare hỗ trợ tích cực trong phục hồi giọng và quản lý theo dõi lâu dài.

IV. BÀN LUẬN

Nghiên cứu này đánh giá hiệu quả bước đầu của hệ thống AI-VoiceCare trong quản lý bệnh

nhân rối loạn giọng tại Bệnh viện Đại học Y Hà Nội, với ba khía cạnh chính: hiệu năng phân loại giọng bệnh lý của mô hình trí tuệ nhân tạo, hiệu quả cải thiện quy trình quản lý hồ sơ ngoại trú, và tác động đến theo dõi - phục hồi giọng của bệnh nhân.

Kết quả cho thấy mô hình mạng nơ-ron tích chập (CNN) phân tích phổ Mel đạt độ chính xác 92,1%, độ nhạy 0,90, độ đặc hiệu 0,93 và AUC 0,94 trong phân loại giọng bệnh lý. Các chỉ số này phản ánh khả năng phân biệt tốt giữa giọng bình thường và giọng bệnh lý, đồng thời tương đương với nhiều nghiên cứu quốc tế ứng dụng học sâu trong nhận diện rối loạn giọng.

Độ nhạy cao cho thấy mô hình có khả năng phát hiện hiệu quả các trường hợp giọng bệnh lý, góp phần giảm nguy cơ bỏ sót bệnh nhân, đặc biệt trong giai đoạn rối loạn chức năng - giai đoạn có khả năng hồi phục tốt nếu được can thiệp sớm. Bên cạnh đó, độ đặc hiệu cao giúp hạn chế chẩn đoán dương tính giả, tránh làm gia tăng số bệnh nhân cần theo dõi hoặc can thiệp không cần thiết. Giá trị AUC vượt ngưỡng 0,9 cho thấy mô hình đạt mức phân biệt “xuất sắc”, khẳng định tiềm năng ứng dụng trong hỗ trợ sàng lọc và theo dõi lâm sàng.

Tuy nhiên, cần nhấn mạnh rằng mô hình AI trong nghiên cứu này được sử dụng như công cụ hỗ trợ quyết định, không thay thế đánh giá lâm sàng của bác sĩ. Việc kết hợp kết quả phân tích giọng với thăm khám và nội soi thanh quản vẫn là yếu tố then chốt trong chẩn đoán chính xác rối loạn giọng.

Một điểm nổi bật của nghiên cứu là việc đánh giá tác động của AI-VoiceCare lên quy trình quản lý hồ sơ bệnh nhân ngoại trú. Kết quả cho thấy thời gian xử lý hồ sơ giảm hơn 60% và tỷ lệ sai sót ghi chép giảm tương ứng sau khi triển khai hệ thống. Điều này cho thấy AI-VoiceCare không chỉ đóng vai trò phân tích giọng nói, mà còn góp phần chuẩn hóa và tự động hóa quy trình hành chính trong thực hành lâm sàng.

Trong bối cảnh các phòng khám Tai Mũi Họng thường có lưu lượng bệnh nhân lớn, việc rút ngắn thời gian xử lý hồ sơ giúp giảm tải cho bác sĩ và nhân viên y tế, đồng thời cải thiện hiệu suất tiếp nhận và chăm sóc người bệnh. Việc giảm sai sót ghi chép cũng có ý nghĩa quan trọng đối với chất lượng dữ liệu bệnh án, đặc biệt trong theo dõi bệnh lý chức năng và mạn tính, nơi dữ liệu không đầy đủ hoặc không nhất quán có thể ảnh hưởng đến quyết định điều trị và đánh giá hiệu quả can thiệp.

Mức độ hài lòng của bác sĩ tăng rõ rệt sau khi triển khai AI-VoiceCare phản ánh tính khả

thi và mức độ chấp nhận của hệ thống trong thực hành thường quy. Đây là yếu tố mang tính quyết định đối với khả năng duy trì và mở rộng ứng dụng công nghệ AI trong môi trường lâm sàng, bởi một hệ thống dù có hiệu năng kỹ thuật cao nhưng không được người sử dụng chấp nhận sẽ khó có giá trị thực tiễn lâu dài.

Sau 6 tháng triển khai, tỷ lệ bệnh nhân tái khám đúng hẹn tăng đáng kể, và hơn hai phần ba bệnh nhân ghi nhận cải thiện rõ rệt về chất lượng giọng. Kết quả này gợi ý rằng AI-VoiceCare có thể đóng vai trò tích cực trong việc tăng cường tuân thủ theo dõi và điều trị của bệnh nhân rối loạn giọng.

Rối loạn giọng, đặc biệt là rối loạn chức năng, thường không gây triệu chứng nặng ngay từ đầu, dẫn đến tình trạng bệnh nhân chủ quan và bỏ tái khám. Việc tích hợp hệ thống AI hỗ trợ theo dõi, lưu trữ và phản hồi tiến triển giọng nói giúp bệnh nhân nhận thức rõ hơn về tình trạng của mình, từ đó tăng động lực tham gia điều trị và tái khám đúng hẹn. Tuy nhiên, do giai đoạn trước triển khai AI chưa có quy trình chuẩn hóa ghi nhận cải thiện giọng, nghiên cứu không thể so sánh trực tiếp mức độ cải thiện giọng giữa hai giai đoạn, đây là một hạn chế cần được lưu ý.

Nghiên cứu này có một số hạn chế. Thứ nhất, đây là nghiên cứu đơn trung tâm với cỡ mẫu tương đối hạn chế, do đó khả năng khái quát hóa kết quả còn hạn chế. Thứ hai, việc so sánh trước - sau triển khai AI sử dụng nhóm chứng lịch sử, không phải thử nghiệm ngẫu nhiên, nên có nguy cơ sai số chọn mẫu theo thời gian. Thứ ba, việc đánh giá cải thiện giọng chưa sử dụng đầy đủ các thang đo chuẩn hóa quốc tế, mà chủ yếu dựa trên đánh giá lâm sàng trong thực hành thường quy. Cuối cùng, nghiên cứu chưa thực hiện phân tích chi phí - hiệu quả, nên chưa thể kết luận về lợi ích kinh tế của hệ thống AI-VoiceCare.

Mặc dù còn hạn chế, kết quả nghiên cứu cho thấy AI-VoiceCare có tính khả thi và hiệu quả bước đầu trong quản lý bệnh nhân rối loạn giọng. Các nghiên cứu tiếp theo cần được thiết kế theo hướng đa trung tâm, cỡ mẫu lớn hơn, sử dụng thang đo giọng chuẩn hóa và bổ sung phân tích chi phí - hiệu quả, nhằm đánh giá toàn diện hơn giá trị lâm sàng và kinh tế của hệ thống này.

V. KẾT LUẬN

Nghiên cứu cho thấy hệ thống AI-VoiceCare có tính khả thi trong hỗ trợ quản lý bệnh nhân rối loạn giọng tại phòng khám ngoại trú. Mô hình trí tuệ nhân tạo đạt hiệu năng phân loại giọng bệnh lý ở mức cao, cho thấy tiềm năng ứng dụng trong hỗ trợ sàng lọc và theo dõi lâm sàng. Việc triển khai AI-VoiceCare gắn với cải thiện hiệu quả quản lý hồ sơ bệnh nhân, thể hiện qua việc rút ngắn thời gian xử lý và giảm sai sót ghi chép, đồng thời được bác sĩ lâm sàng chấp nhận ở mức cao. Tỷ lệ tái khám đúng hẹn và cải thiện giọng ở giai đoạn triển khai AI cho thấy xu hướng tích cực trong theo dõi và điều trị bệnh nhân. Do những hạn chế về thiết kế nghiên cứu và phạm vi dữ liệu, các kết quả này cần được kiểm chứng thêm bằng các nghiên cứu có thiết kế chặt chẽ hơn trước khi khẳng định giá trị ứng dụng rộng rãi.

LỜI CẢM ƠN

Nhóm nghiên cứu xin chân thành cảm ơn các bác sĩ, kỹ thuật viên và điều dưỡng của các Khoa Khám bệnh đã phối hợp chặt chẽ trong quá trình thu thập mẫu giọng, thực hiện nội soi và theo dõi lâm sàng bệnh nhân. Nhóm nghiên cứu cũng gửi lời cảm ơn sâu sắc đến toàn thể bệnh nhân đã tham gia nghiên cứu vì sự hợp tác, tin tưởng và kiên trì trong quá trình điều trị và theo dõi. Xin trân trọng cảm ơn Phòng Công nghệ thông tin đã hỗ trợ kỹ thuật trong xây dựng và vận hành hệ thống AI-VoiceCare,

cũng như Phòng Kế hoạch Tổng hợp đã hỗ trợ quản lý, thống kê và điều phối dữ liệu trong suốt quá trình nghiên cứu.

TÀI LIỆU THAM KHẢO

1. Roy N, Merrill RM, Gray SD, Smith EM. Voice disorders in the general population. *Laryngoscope*. 2005; 115(11): 1988-1995. doi:10.1097/01.mlg.0000179174.32345.41.
2. Verdolini K, Ramig LO. Review: occupational risks for voice problems. *Logoped Phoniatr Vocol*. 2001; 26(1): 37-46. doi:10.1080/140154301300109150.
3. Behlau M, Zambon F, Guerrieri AC, Roy N. Epidemiology of voice disorders. *J Voice*. 2017; 31(1): 127.e1-127.e7. doi:10.1016/j.jvoice.2015.08.017.
4. Zhang Y, Jiang J. Acoustic analysis in voice pathology detection. *IEEE Trans Biomed Eng*. 2020; 67(8): 2197-2205. doi:10.1109/TBME.2020.2969782.
5. Phạm TTB, Võ TH, Nguyễn DM. Ứng dụng trí tuệ nhân tạo trong chẩn đoán bệnh lý giọng nói. *Tạp chí Tai Mũi Họng Việt Nam*. 2023; 18(3): 45-52.
6. Alhindi T, Elshafeey A, Alotaibi M, et al. CNN-based detection of voice disorders. *J Voice*. 2022; 36(3): 412-420. doi:10.1016/j.jvoice.2021.03.010.
7. Paja W, Stylianou Y. Automatic detection of pathological voices. *IEEE Access*. 2021; 9: 10456-10467. doi:10.1109/ACCESS.2021.3051234.
8. Chen L, Zhang Y, Yang X, et al. Deep learning in voice pathology. *Speech Commun*. 2021; 128: 14-22. doi:10.1016/j.specom.2021.02.004.
9. Umapathy K, Krishnan S, Rao A, et al. Feature extraction for pathological voice detection. *Biomed Signal Process Control*. 2020;

59:101904. doi:10.1016/j.bspc.2020.101904.

10. Chien CY, Lee YC, Lin MH, et al. AI in laryngeal image diagnosis. *Otolaryngol Head Neck Surg.* 2022; 166(4): 703-710. doi:10.1177/01945998211066723.

11. Arora R, Sharma D, et al. Voice pathology detection using transfer learning. *Comput Speech Lang.* 2023; 80: 101470. doi:10.1016/j.csl.2022.101470.

12. Hu T, Wang X, Li Q, et al. Multimodal AI for laryngeal diseases. *IEEE Trans Neural Netw Learn Syst.* 2024; 35(2): 543-556. doi:10.1109/TNNLS.2023.3298715.

13. Wang S, Zhao Y, Lin H, et al. AI-powered screening of voice disorders. *Expert Syst Appl.* 2023; 224:119990. doi:10.1016/j.eswa.2023.119990.

14. International Organization for Standardization. ISO/IEC 19794-13:2021. Information Technology-Biometric Data Interchange Formats-Part 13: Voice Data. Geneva: ISO; 2021.

15. Patel RR, Huber JE, Kreiman J, et al. Clinical implications of AI in speech pathology. *Int J Speech Lang Pathol.* 2023; 25(3): 251-260. doi:10.1080/17549507.2022.2142658.

Summary

APPLICATION OF ARTIFICIAL INTELLIGENCE MODEL IN MANAGING PATIENTS WITH VOICE DISORDERS

The study was conducted at Hanoi Medical University Hospital from 2024 to 2025 to evaluate the effectiveness of the artificial intelligence (AI) model *AI-VoiceCare* in managing patients with voice disorders. A total of 312 patients diagnosed with voice disorders were enrolled. The AI model utilized a convolutional neural network (CNN) to analyze Mel-spectrogram features and integrated the *AI-VoiceCare* system for outpatient electronic medical record management. The model achieved an accuracy of 92.1%, sensitivity of 0.90, and specificity of 0.93 in classifying pathological voices. Record-processing time was reduced by 61.1%, and documentation errors decreased by 60.9%. After six months of implementation, the proportion of patients returning for scheduled follow-up visits increased from 48.7% to 76.3%, and 68.5% of patients reported significant improvement in voice quality after AI-assisted therapy. In total, 84.3% of physicians rated the system as “useful” or “very useful.” The application of *AI-VoiceCare* enhances diagnostic accuracy, patient management, and voice rehabilitation, contributing to the development of a personalized care model for individuals with voice disorders.

Keywords: Voice disorder, artificial intelligence, *AI-VoiceCare*, voice analysis.