

Đánh giá hiệu quả của ứng dụng AI trong phân tích kết quả kháng sinh đồ phương pháp Kirby-Bauer so với phương pháp đo truyền thống

Nguyễn Hoàng Bách*, Ung Thị Thủy, Thái Khắc Trường,

Đinh Thị Hải, Trần Thị Quỳnh Tâm

Bộ môn Vi sinh, Trường Đại học Y - Dược, Đại học Huế

Tóm tắt

Đặt vấn đề: Phương pháp Kirby-Bauer là kỹ thuật phổ biến để đánh giá độ nhạy kháng sinh của vi khuẩn dựa trên đường kính vùng ức chế. Tuy nhiên, việc đo thủ công tiêu tốn nhiều thời gian. Ứng dụng trí tuệ nhân tạo (AI) cho phép tự động hoá quy trình phân tích hình ảnh, giảm thiểu yếu tố chủ quan và nâng cao hiệu suất. **Đối tượng và phương pháp nghiên cứu:** Nghiên cứu được thực hiện trên 608 ảnh kháng sinh đồ của năm loài vi khuẩn thường gặp: *Escherichia coli*, *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae* và *Acinetobacter baumannii*. Mỗi đĩa có 4-6 kháng sinh, ảnh chụp bằng điện thoại thông minh trong điều kiện ánh sáng tiêu chuẩn trong hộp chụp ảnh DDr box. Dữ liệu thu được được phân tích bởi mô hình AI nhận diện nhãn kháng sinh và đo đường kính vùng ức chế, so sánh với kết quả đo thủ công bằng thước kỹ thuật. **Kết quả:** AI đạt độ chính xác nhận diện nhãn kháng sinh lần đầu ~97%, tăng lên ~99,6% sau khi tinh chỉnh (fine-tuning). Độ chính xác trong đo đường kính vùng ức chế nhìn chung khá cao, nhưng vẫn ghi nhận sai lệch ở một số trường hợp chồng lấn vòng vô khuẩn hoặc biên giới không rõ. Thời gian xử lý trung bình của AI thấp hơn đáng kể so với đo thủ công. Cơ chế kết hợp “Human-In-The-Loop” (HITL) giúp phát hiện sớm và giảm thiểu sai số. **Kết luận:** Việc ứng dụng AI trong phân tích kháng sinh đồ Kirby-Bauer cho thấy tiềm năng nâng cao hiệu quả chẩn đoán, tiết kiệm thời gian và giảm sai sót so với phương pháp thủ công. Kết hợp HITL và tinh chỉnh mô hình là giải pháp quan trọng nhằm duy trì tính chính xác, đồng thời mở rộng khả năng triển khai trong các đơn vị xét nghiệm.

Từ khóa: trí tuệ nhân tạo, Human in the loop, tinh chỉnh, Kirby-Bauer.

Evaluation of the effectiveness of AI application in analyzing Kirby-Bauer method's antibiotic susceptibility test results compared to traditional measurement methods

Nguyen Hoang Bach*, Ung Thi Thuy, Thai Khac Truong,

Dinh Thi Hai, Tran Thi Quynh Tam

Department of Microbiology, Hue University of Medicine and Pharmacy, Hue University

Abstract

Background: The Kirby-Bauer method is a widely used technique to evaluate bacterial antibiotic susceptibility based on the diameter of the inhibition zone. However, manual measurement is time-consuming. The application of artificial intelligence (AI) allows for the automation of the image analysis process, minimizing subjective factors and enhancing efficiency. **Material and Methods:** The study was conducted on 608 antibiotic susceptibility images of five common bacterial species: *Escherichia coli*, *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, and *Acinetobacter baumannii*. Each plate contained 4–6 antibiotics, and the images were captured using a smartphone under standard lighting conditions in a DDr box. The collected data were analyzed by an AI model that identifies antibiotic labels and measures the inhibition zone diameter, with the results compared to manual measurements using a technical ruler. **Results:** The AI achieved an initial accuracy of approximately 97% in recognizing antibiotic labels, which increased to approximately 99.6% after fine-tuning. The accuracy in measuring the inhibition zone diameter was generally high, although some discrepancies were noted in cases with overlapping zones or unclear boundaries. The average processing time of the AI was significantly lower than that of manual measurement. The integration of a “Human-In-The-Loop” (HITL) mechanism helped in early detection and minimization of errors. **Conclusion:** The application of AI in Kirby-Bauer antibiotic susceptibility analysis

*Tác giả liên hệ: Nguyễn Hoàng Bách. Email: nhbach@huemed-univ.edu.vn

Ngày nhận bài: 28/1/2025; Ngày đồng ý đăng: 24/2/2025; Ngày xuất bản: 25/3/2025

DOI: 10.34071/jmp.2025.1.24

demonstrates the potential to improve diagnostic efficiency, save time, and reduce errors compared to manual methods. Integrating HITL and fine-tuning the model are key solutions to maintain accuracy while expanding deployment in testing laboratories.

Keywords: Artificial intelligent, Human-in-the-loop, fine-tune, Kirby-Bauer.

1. ĐẶT VẤN ĐỀ

Phương pháp Kirby-Bauer là một kỹ thuật phổ biến trong vi sinh lâm sàng để đánh giá độ nhạy cảm của vi khuẩn với kháng sinh. Đọc kết quả của phương pháp này dựa vào việc đo đường kính vùng ức chế vi khuẩn xung quanh đĩa kháng sinh trên môi trường thạch, từ đó xác định mức độ nhạy cảm của vi khuẩn đối với từng loại kháng sinh, hỗ trợ bác sĩ lâm sàng lựa chọn kháng sinh hiệu quả nhất trong điều trị [1]. Hiện nay đã có nhiều hệ thống tự động phân tích kết quả nhạy cảm kháng sinh như Vitek 2 (bioMérieux) và BD Phoenix (BD Diagnostic Systems) giúp cải thiện tốc độ và độ chính xác của xét nghiệm, nhưng phương pháp Kirby-Bauer vẫn là lựa chọn phổ biến tại các phòng xét nghiệm có nguồn lực hạn chế, đặc biệt ở các quốc gia đang phát triển [2]. Điều này chủ yếu do chi phí đầu tư ban đầu và yêu cầu bảo trì của các hệ thống tự động thường cao, trong khi phương pháp Kirby-Bauer có chi phí thấp, dễ thực hiện và không đòi hỏi trang thiết bị phức tạp, giúp nhiều phòng xét nghiệm có thể tiếp cận và triển khai dễ dàng. Mặc dù đơn giản và hiệu quả, phương pháp này đòi hỏi thời gian và sự chính xác trong việc đọc kết quả, đặc biệt khi số lượng mẫu lớn. Phương pháp đo thủ công trong thử nghiệm kháng sinh đồ Kirby-Bauer có nhiều hạn chế, bao gồm sai số do con người, độ lặp lại kém, tốn thời gian và khó đo chính xác các vùng ức chế phức tạp. Các kỹ thuật kháng sinh đồ (KSD) tự động hoá có thể khắc phục được các hạn chế này, đặc biệt là ứng dụng AI, để nâng cao độ chính xác và hiệu quả trong phân tích kháng sinh đồ [3,4].

Trong những năm gần đây, trí tuệ nhân tạo (Artificial Intelligence -AI) đã được ứng dụng rộng rãi trong y học, đặc biệt là trong phân tích hình ảnh và dữ liệu. Các thuật toán học sâu (deep learning) và máy học (machine learning) đã chứng minh tiềm năng trong việc tự động hóa quy trình phân tích kết quả kháng sinh đồ, giúp giảm thiểu sai sót do con người và tăng tốc độ xử lý [5,6]. AI có thể phân tích hình ảnh vùng ức chế vi khuẩn một cách chính xác, đồng thời dự đoán kết quả dựa trên các mô hình được huấn luyện từ dữ liệu trong dataset. Mặc dù tiềm năng lớn, việc ứng dụng AI trong phân tích kháng sinh đồ vẫn gặp nhiều thách thức. Đầu tiên, chất lượng dữ liệu huấn luyện là yếu tố quyết định

hiệu quả của mô hình AI. Dữ liệu không đồng nhất hoặc thiếu chính xác có thể dẫn đến kết quả sai lệch. Thứ hai, việc tích hợp AI vào quy trình lâm sàng đòi hỏi sự đầu tư về công nghệ và đào tạo nhân lực. Cuối cùng, các vấn đề về đạo đức và pháp lý, chẳng hạn như trách nhiệm khi xảy ra sai sót, cũng cần được xem xét [7].

Trong các báo cáo trước đây, nhóm nghiên cứu đã đề xuất các mô hình máy học cũng như xây dựng công cụ phân tích hình ảnh kết quả kháng sinh đồ phương pháp Kirby-Bauer tại địa chỉ <https://amr.ai.vn> [TKTK mới của nhóm]. Các nghiên cứu ban đầu nhằm đánh giá một số tiêu chí trên dữ liệu mẫu thử nghiệm. Để đánh giá hiệu quả của ứng dụng AI trong phân tích kháng sinh đồ, các nghiên cứu thường áp dụng nhiều phương pháp toàn diện. Trước hết, kết quả do AI đưa ra được so sánh với kết quả từ phương pháp truyền thống do kỹ thuật viên thực hiện thông qua các chỉ số như độ chính xác (accuracy), độ nhạy (sensitivity) và độ đặc hiệu (specificity). Ngoài ra, thời gian xử lý của AI cũng được đánh giá để xác định khả năng tăng tốc độ phân tích so với phương pháp thủ công. Bên cạnh đó, khả năng tích hợp của AI vào hệ thống hiện có, chẳng hạn như kết nối với hệ thống quản lý dữ liệu phòng xét nghiệm (LIS), cũng là yếu tố quan trọng. Mức độ chấp nhận của người dùng, bao gồm kỹ thuật viên đánh giá thông qua khảo sát hoặc phỏng vấn để hiểu rõ tính dễ sử dụng và độ tin cậy của AI. Các nghiên cứu cũng cần kiểm tra hiệu suất của AI trên các bộ dữ liệu đa dạng, bao gồm nhiều loại vi khuẩn, điều kiện nuôi cấy và kháng sinh khác nhau, để đảm bảo tính ứng dụng rộng rãi. Cuối cùng, việc đánh giá chi phí triển khai so với lợi ích mang lại, cũng như độ ổn định và khả năng mở rộng của AI trong môi trường thực tế, là những yếu tố không thể bỏ qua. Những phương pháp này giúp đảm bảo AI không chỉ chính xác mà còn thực tế và hữu ích trong phân tích kháng sinh đồ [8].

Nghiên cứu này nhằm đánh giá hiệu quả của ứng dụng AI trong phân tích kết quả kháng sinh đồ bằng phương pháp Kirby-Bauer so với phương pháp đo truyền thống. Cụ thể, nghiên cứu sẽ tập trung vào so sánh độ chính xác và thời gian xử lý giữa hai phương pháp, phân tích các yếu tố ảnh hưởng đến hiệu quả của AI trong phân tích kháng sinh đồ.

2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng nghiên cứu

Nghiên cứu sử dụng kết quả thử nghiệm kháng sinh đồ kỹ thuật đĩa thạch khuếch tán của các bệnh nhân xét nghiệm tại khoa Vi sinh, Bệnh viện Trường Đại học Y-Dược Huế. Dữ liệu gồm 608 ảnh kết quả kháng sinh đồ của 5 loài vi khuẩn gây bệnh phổ biến, cụ thể 270 ảnh kháng sinh đồ trên *Escherichia coli*, 156 ảnh kháng sinh đồ trên *Staphylococcus aureus*, 98 ảnh kháng sinh đồ *Pseudomonas aeruginosa*, 19 ảnh kháng sinh đồ *Klebsiella pneumoniae* và 18 ảnh kháng sinh đồ *Acinetobacter baumannii*.

Các bộ kháng sinh thử nghiệm trên các chủng vi khuẩn được sử dụng theo hướng dẫn Bộ Y tế và theo quy trình thao tác chuẩn (Standard Operating Procedure-SOP) của Khoa Vi sinh, Bệnh viện Trường Đại học Y-Dược Huế. Số lượng kháng sinh đặt trên một đĩa thạch kích thước 90mm là 4-6 khoanh giấy kháng sinh. Hình ảnh thu nhận bằng camera trước của điện thoại thông minh với đĩa kháng sinh đồ đặt trong thiết bị DDr Box. Hình ảnh có độ phân giải tối thiểu 2560×1440 px, được chụp trong điều kiện ánh sáng LED trắng được cung cấp bởi thiết bị DDr Box (Hình 1).



Hình 1. Thiết bị DDr Box sử dụng để tạo môi trường chụp ảnh bằng điện thoại di động kết quả kháng sinh đồ phương pháp Kirby-Bauer sử dụng trong nghiên cứu.

Dữ liệu kết quả kháng sinh đồ đối chứng để đánh giá độ chính xác được xuất ra từ phần mềm quản lý phòng xét nghiệm (LIS) của Khoa Vi Sinh, Bệnh viện Trường Đại học Y-Dược Huế.

2.2. Phương pháp nghiên cứu

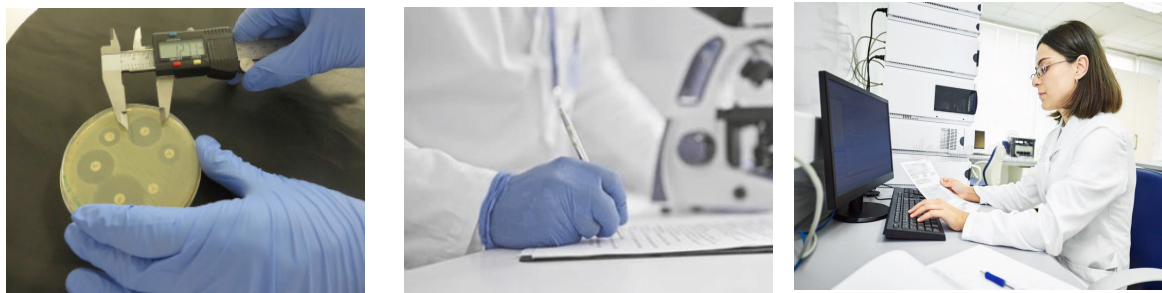
Thiết kế nghiên cứu: Thiết kế nghiên cứu mô tả cắt ngang.

Nội dung nghiên cứu:

Nghiên cứu thu thập dữ liệu để đánh giá các tiêu chí độ chính xác, độ đặc hiệu, hiệu suất từ lúc

bắt đầu đọc kết quả đến phiên giải kết quả trên 2 phương thức: sử dụng AI và phương pháp truyền thống đo bằng thước kỹ thuật calliper.

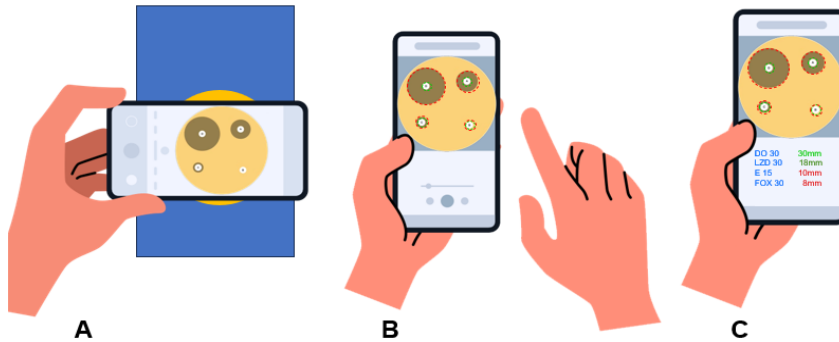
Kỹ thuật viên đo bằng thước calliper: Tiến hành đo đường kính vùng ức chế (inhibition zone), nhận biết nhãn kháng sinh, ghi kết quả đo đường kính (mm) và nhập dữ liệu vào hệ thống LIS (Laboratory information system). Thời gian thực hiện được tính từ lúc bắt đầu đo đến lúc ra kết luận nhạy cảm (S), đề kháng (R) và trung gian (I) của mỗi loại kháng sinh (Hình 2).



Hình 2. Quy trình đọc và phân tích kết quả kháng sinh đồ truyền thống.

A: đọc nhãn kháng sinh và đo đường kính vùng ức chế bằng thước calliper. B: ghi kết quả đo từng loại nhãn kháng sinh vào phiếu tiến trình. C: nhập đường kính vùng ức chế vào đúng tên kháng sinh trên hệ thống LIS để có kết luận nhạy cảm (S), đề kháng (R), trung gian (I).

Kỹ thuật viên sử dụng ứng dụng tích hợp AI hỗ trợ phân tích kết quả: chụp ảnh, phân tích kết quả tự động và có giám sát, ra kết quả cuối cùng. Dữ liệu phiên giải kháng sinh sử dụng API của WHONET 2024. Quy trình thao tác hoàn toàn trên điện thoại thông minh được mô tả tại hình 3.



Hình 3. Quy trình chụp ảnh và phân tích kết quả kháng sinh đồ bằng ứng dụng AI.

A: chụp ảnh bằng điện thoại thông minh với đĩa kháng sinh đồ đặt trong DDr Box. B: Nhập thông tin xét nghiệm gồm loài vi khuẩn, mã bệnh nhân, bệnh phẩm.

C: phân tích và phiên giải kết quả gồm tên kháng sinh, kết luận nhạy cảm (S), đề kháng (R), trung gian (I).

Dữ liệu nhãn kháng sinh, đường kính vùng vô khuẩn và kết luận sẽ được thu thập để đánh giá độ chính xác, độ nhạy và độ chính xác khi dự đoán dương tính của phương pháp có sử dụng AI. Thời gian thực hiện trên từng kháng sinh, trên từng đĩa kết quả và trên một lô kết quả sẽ được ghi nhận và lưu trữ dữ liệu để tính hiệu suất của của từng phương pháp.

Accuracy (Độ chính xác)

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Trong đó:

TP (True Positive): Dự đoán đúng trường hợp thực sự dương (positive).

TN (True Negative): Dự đoán đúng trường hợp thực sự âm (negative).

FP (False Positive): Dự đoán sai trường hợp thực sự âm.

FN (False Negative): Dự đoán sai trường hợp thực sự dương.

Sensitivity (Độ nhạy)

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

Độ nhạy thể hiện khả năng phát hiện đúng các trường hợp dương tính (positive) trong tất cả các trường hợp dương tính có thật.

Precision (Độ chính xác khi dự đoán dương)

$$\text{Precision} = \frac{TP}{TP+FP}$$

Precision thể hiện trong số các trường hợp được mô hình dự đoán là dương tính, có bao nhiêu trường hợp thực sự dương tính.

F1-score: là trung bình điều hòa giữa Precision và Sensitivity, được khuyến nghị sử dụng khi cần cân bằng cả hai chỉ số và/hoặc khi có **sự mất cân đối giữa** số lượng dữ liệu dương và âm

$$F1 = 2 \times \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}}$$

3. KẾT QUẢ

3.1. Chức năng Human-in-the-loop và tinh chỉnh (fine-tuning), tái huấn luyện dữ liệu

Trong khi phát triển ứng dụng AI, chúng tôi đã kết hợp chức năng Human-in-the-loop (HITL) cùng các kỹ thuật tinh chỉnh (fine-tuning) và tái huấn luyện dữ liệu đã trở thành yếu tố then chốt để nâng cao hiệu quả và độ tin cậy của hệ thống. HITL đảm bảo sự tham gia trực tiếp của kỹ thuật viên trong quá trình đào tạo, giám sát và điều chỉnh mô hình, từ đó tối ưu hóa kết quả thông qua phản hồi chuyên sâu. Trong khi đó, fine-tuning cho phép tùy chỉnh mô hình tiền huấn luyện trên tập dữ liệu cụ thể, tái huấn luyện lại giúp cập nhật dữ liệu mới hoặc sửa lỗi hệ thống. Sự kết hợp này không chỉ cải thiện độ chính xác mà còn thích ứng linh hoạt với các tình huống thực tế, đặc biệt trong lĩnh vực đòi hỏi tính an toàn cao như y tế. Hình 3 mô tả các chức năng giám sát dữ liệu tự động nhận diện và đề xuất dữ liệu thay thế cũng như lưu lại dữ liệu vào tập dataset để tái huấn luyện.

Trong nghiên cứu chúng tôi chỉ sử dụng ban đầu 42 loại nhãn kháng sinh của các công ty có thương hiệu được sử dụng ở nhiều phòng xét nghiệm. Để đảm bảo hệ thống tự học thông qua dữ liệu của

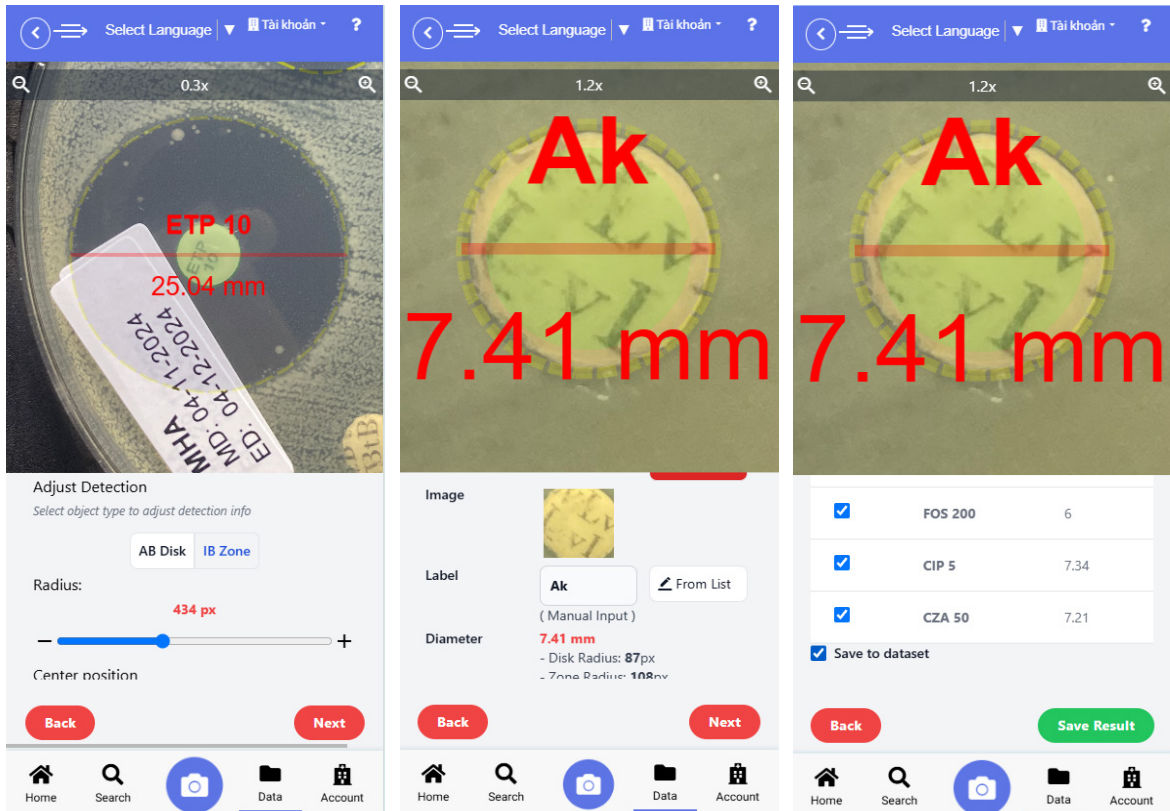
nhiều phòng xét nghiệm thì chức năng bổ sung dữ liệu đã xây dựng và thao tác dễ dàng trên điện thoại di động hoặc máy tính để bàn. Chức năng này gồm các bước chính sau (hình 4):

- Bước 1: hệ thống tự động đề xuất số lượng khoan giấy kháng sinh, nhãn kháng sinh, vùng ức chế và đường kính vùng ức chế.

- Bước 2: hệ thống HITL đề xuất chỉnh sửa thông số:
+ Nhãn kháng sinh: KTV có thể nhập dữ liệu mới hoặc lấy từ cơ sở dữ liệu có sẵn.

+ Đường kính vùng ức chế: thay đổi kích thước đường kính (mm) dựa trên công cụ tăng giảm đường kính (fine-tuning).

- Bước 3: lưu dữ liệu và lưu dữ liệu chỉnh sửa vào dataset.



Hình 4. Các thao tác giúp KTV tham gia điều chỉnh kết quả tự động nhận diện của AI, lưu trữ dữ liệu phục vụ tái huấn luyện mô hình.

3.2. Khả năng xác định đường kính vùng ức chế của AI so với kỹ thuật viên

Mô hình AI được sử dụng để xác định vùng ức chế và khoan giấy kháng sinh. Từ khoan giấy kháng sinh thì vùng ức chế sẽ là hình tròn đồng tâm với khoan giấy kháng sinh. Mô hình tính toán nội suy đường kính vùng ức chế dựa vào hằng số là đường kính khoan giấy kháng sinh là 6mm. Kết quả sử dụng AI trong dữ liệu nghiên cứu được thể hiện ở Bảng 1.

Bảng 1. Kết quả xác định đường kính vùng ức chế bằng AI so với kỹ thuật viên.

Vi khuẩn	Ảnh kết quả KSD	Kết quả tương đồng KTV*	Vùng ức chế	Đường kính tương đồng (<±1 mm)	Khác biệt (> ±1mm)	Khác biệt thay đổi kết luận**
<i>E. coli</i>	235	214 (91,1%)	1479	1237 (83,6%)	242 (16,6%)	53 (3,6%)
<i>S. aureus</i>	126	115 (91,3%)	559	464 (83,0%)	95 (17,0%)	22 (3,9%)
<i>P. aeruginosa</i>	98	91 (92,9%)	507	470 (92,7%)	37 (7,3%)	15 (3,0%)

<i>K. pneumoniae</i>	19	14 (73,7%)	136	82 (60,3%)	54 (39,7%)	17 (12,5%)
<i>A.baumannii</i>	16	14 (87,5%)	105	83 (79,0%)	22 (21,0%)	12 (11,4%)
Tổng cộng	494		2786	2336	450	119

* Số ảnh không có tất cả đường kính vùng vô khuẩn khác biệt ± 1 mm

** Vùng ức chế có đường kính sai lệch khác biệt về kết luận: nhạy cảm, đề kháng, trung gian

Bảng 1 so sánh kết quả kháng sinh đồ (KSD) giữa hệ thống AI và đánh giá của kỹ thuật viên dựa trên đường kính vùng ức chế. Các chỉ số được trình bày bao gồm số lượng mẫu theo từng loại vi khuẩn, tỷ lệ “kết quả tương đồng” trong việc phân loại nhạy cảm/đề kháng, số liệu định lượng về vùng ức chế (bao gồm số mẫu có đường kính tương đồng trong khoảng ± 1 mm và số mẫu có sự khác biệt vượt quá ± 1 mm) cũng như tỷ lệ các trường hợp sai lệch đủ lớn để thay đổi kết luận. Nhìn chung, tỷ lệ tương đồng của kết quả KSD giữa AI so với kỹ thuật viên ở mức khá cao đối với hầu hết các vi khuẩn (trên 90% đối với một số chủng như *E. coli*, *S. aureus* và *P. aeruginosa*), trong khi đối với *K. pneumoniae* tỷ lệ

này giảm xuống chỉ khoảng 73,7%. Ngoài ra, sự nhất quán về số liệu đo đường kính vùng ức chế cũng dao động, với tỷ lệ mẫu có chênh lệch ≤ 1 mm khá cao ở một số chủng (đến 92,7% ở *P. aeruginosa*) nhưng thấp hơn đáng kể ở *K. pneumoniae* (60,3%), dẫn đến tỷ lệ sai lệch thay đổi kết luận cao nhất ở chủng này (12,5%). Những số liệu này chỉ ra rằng trong khi hệ thống AI có khả năng tái hiện kết quả của kỹ thuật viên ở đa số trường hợp, thì vẫn tồn tại một số hạn chế khi áp dụng cho một số loại vi khuẩn nhất định.

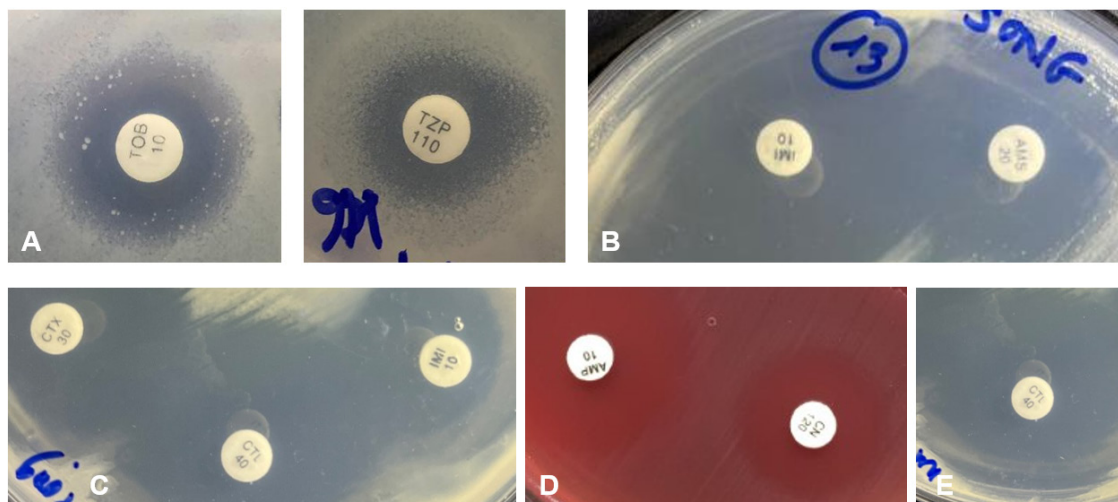
Dữ liệu hiệu chỉnh của nghiên cứu đã phân loại các nhóm nguyên nhân dẫn đến sai lệch đường kính vùng ức chế so với KTV đo bằng tay. Các yếu tố ảnh hưởng và số liệu được thể hiện ở bảng 2.

Bảng 2. Các nguyên nhân gây sai lệch kết quả đo bằng AI.

Các nguyên nhân AI đo sai đường kính (>1mm)	Số lượng vùng vô khuẩn của từng loại vi khuẩn sai lệch so với KTV				
	<i>E. coli</i>	<i>S. aureus</i>	<i>P. aeruginosa</i>	<i>K. pneumoniae</i>	<i>A. baumannii</i>
Hình ảnh vùng ức chế chồng nhau	152	20	29	36	8
Có vi khuẩn mọc trong vòng ức chế	32	0	0	3	5
Có hình ảnh vòng ức chế mờ	38	34	3	5	0
Vùng ức chế có vùng biên giới hạn không rõ	0	22	0	3	6
Hình ảnh bị bóng mờ	20	19	5	7	3
Tổng cộng	242	95	37	54	22

Từ số liệu Bảng 2 cho thấy nguyên nhân và tần suất gặp các sai số trong việc tính toán đường kính vùng vô khuẩn khá đa dạng. Trong quá trình đánh giá độ chính xác của mô hình AI khi xác định đường kính vùng ức chế, tổng cộng 450 trường hợp sai số đã được ghi nhận ở năm loài vi khuẩn (*E. coli*, *S. aureus*, *P. aeruginosa*, *K. pneumoniae* và *A. baumannii*). Dữ liệu cho thấy năm nhóm nguyên nhân chính gây ra sai số, bao gồm: (i) hình ảnh vùng ức chế chồng nhau, (ii) có vi khuẩn mọc trong vùng ức chế, (iii) vùng ức

chế mờ (fuzzy ring), (iv) vùng biên giới hạn không rõ và (v) hình ảnh bị bóng mờ. Trong số này, vùng ức chế chồng chéo nhau chiếm tỷ lệ cao nhất (245/450 trường hợp, tương đương 54,4%), tiếp đến là vòng ức chế mờ (80/450, chiếm 17,8%) và hình ảnh bị bóng mờ (54/450, chiếm 12,0%). Các lỗi còn lại, như vi khuẩn mọc trong vòng ức chế (40/450, 8,9%) và biên giới hạn không rõ (31/450, 6,9%), tuy chiếm tỷ lệ thấp hơn nhưng cũng có thể ảnh hưởng đáng kể đến quá trình nhận dạng và đo đạc của mô hình.



Hình 5. Các hình ảnh vùng ức chế mô hình AI nhận định chưa chính xác.

A: có vi khuẩn mọc trong vòng ức chế. B: vùng ức chế chồng nhau. C: khoan giấy kháng sinh có ảnh bóng mờ. D: không xác định rõ vùng ức chế. E: vùng biên vùng ức chế không xác định rõ.

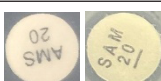
Xét theo từng loài vi khuẩn, *E. coli* có số trường hợp sai số cao nhất (242), phản ánh cả tần suất thử nghiệm lớn lẫn khả năng phát triển lan rộng, dễ gây hiện tượng vòng ức chế chồng lấn. Tương tự, *S. aureus* (95) cũng thường gặp sai sót, đặc biệt là ở hai nhóm vòng mờ và vùng biên giới không rõ. Trong khi đó, *P. aeruginosa* (37) và *A. baumannii* (22) ghi nhận số mẫu sai sót ít hơn, song vẫn có các lỗi đặc trưng liên quan đến vòng ức chế mờ hoặc bóng mờ do điều kiện chụp ảnh. *K. pneumoniae* (54) đáng chú ý vì có tỷ lệ vòng ức chế chồng chéo tương đối cao, gây khó khăn cho mô hình AI khi xác định ranh giới riêng biệt của mỗi đĩa kháng sinh.

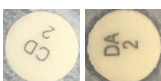
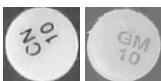
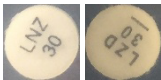
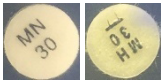
Nguyên nhân chính của các sai sót này có thể xuất phát từ nhiều yếu tố. Thứ nhất, việc bố trí đĩa kháng sinh quá gần nhau trên môi trường nuôi cấy hoặc tốc độ mọc nhanh của vi khuẩn dẫn đến hiện tượng vòng ức chế chồng lấn, gây nhiễu cho thuật toán xác định đường viền. Thứ hai, chất lượng môi trường thạch, độ ổn định của quá trình chuẩn bị mẫu, cũng như đặc điểm tăng trưởng của vi khuẩn có thể dẫn đến vòng ức chế mờ hoặc có khuẩn mọc rải rác bên trong vòng. Thứ ba, điều kiện chụp ảnh không tối ưu (ánh sáng không đồng đều, lóa sáng, hoặc hiện tượng bóng mờ) làm mất độ tương phản, khiến AI không thể tách bạch ranh giới vùng ức chế (hình 5).

3.3. Độ chính xác nhận diện nhãn kháng sinh

Trong nghiên cứu chúng tôi đã sử dụng các kết quả kháng sinh đồ từ 5 chủng vi khuẩn. Với phương thức thực hiện định danh và kháng sinh đồ song song áp dụng trong một số trường hợp, chúng tôi đã sử dụng 42 loại kháng sinh khác nhau từ các 2 công ty lớn và phổ biến Liofilchem (Teramo, Italy), Oxoid (ThermoFisher Scientific, Hoa Kỳ). Ngoài ra chúng tôi cũng thử nghiệm các kháng sinh từ công ty Nam Khoa (Tp Hồ Chí Minh, Việt Nam). Mỗi loại kháng sinh có từ 1 đến 2 nhãn thương mại khác nhau. Trong đó, các kháng sinh thuộc công ty Nam Khoa (Tp. Hồ Chí Minh, Việt Nam) được sản xuất khác tiêu chuẩn của các công ty khác trên thế giới. Dữ liệu cho thấy có 2711/2786 (97,3%) các nhãn kháng sinh đã được xác định thành công. Trong đó, 11 (0,39%) nhãn không phát hiện do nhãn đĩa petri hoặc các nguyên nhân khác che khuất nhãn khoan giấy kháng sinh. 64 (2,3%) số nhãn không thể đọc lần đầu sử dụng do tập dữ liệu huấn luyện. Đây là các kháng sinh thuộc công ty Nam Khoa với các nhãn gồm Ak, Ac, Am, Bt, Cm, Cz, Pt, Tb. Công cụ gắn nhãn khi sử dụng giúp các KTV tự gắn nhãn cho các kháng sinh mới, lưu trữ vào bộ dữ liệu và tái huấn luyện mô hình AI. Sau khi phân tích lại mô hình dữ liệu mới thì toàn bộ 64 nhãn kháng sinh đã nhận diện đúng (Bảng 3).

Bảng 3. Dữ liệu sai lệch khi xác định nhãn kháng sinh bằng mô hình AI.

STT	Nhãn kháng sinh thương mại	Ảnh khoanh giấy kháng sinh	Mã code WHONET	Tên kháng sinh	Số nhãn nhận diện sai lệch chưa tái huấn luyện
1	AK 30, Ak		AMK	Amikacin	Ak: 9
2	AMC 30, Ac		AMC	Amoxicillin/ Clavulanic acid	Ac: 6
3	AMP 10, Am		AMP	Ampicillin	Am: 13
4	AMS 20, SAM 20		SAM	Ampicillin/ Sulbactam	0
5	ATM 30		ATM	Aztreonam	0
6	KZ 30		CZO	Cefazolin	0
7	FEP 30, Cm		FEP	Cefepime	Cm: 6
8	CTX 30		CTX	Cefotaxime	CTX 30: 2
9	CTC 40, CTL 40		CTC	Cefotaxime/ Clavulanic acid	CTL 40: 2
10	FOX 30		FOX	Cefoxitin	0
11	CPT 30		CPT	Ceftaroline	CPT 30: 2
12	CAZ 30, Cz		CAZ	Ceftazidime	Cz: 3
13	CZA 50		CZA	Ceftazidime/ Avibactam	CZA 50: 4
14	CAL 40		CAL	Ceftazidime/ Ceftazidime/ Clavulanic acid	CAL 40: 2
15	CRO 30		CRO	Ceftriaxone	0
16	CXM 30		CXM	Cefuroxime	CMX 30: 1

17	C 30		CHL	Chloramphenicol	0
18	CIP 5		CIP	Ciprofloxacin	0
19	CLR 15		CLR	Clarithromycin	0
20	CD 2, DA 2		CLI	Clindamycin	0
21	DO 30, DXT 30		DOX	Doxycycline	0
22	ETP 10		ETP	Ertapenem	0
23	E 15		ERY	Erythromycin	0
24	FOS 200		FOS	Fosfomycin	0
25	CN 10, GM 10		GEN	Gentamicin	0
26	CN120		GEH	Gentamicin-High	0
27	IMI 10, IPM 10		IPM	Imipenem	0
28	LEV 5, Lv		LVX	Levofloxacin	Lv: 8
29	LNZ 30, LZD 30		LNZ	Linezolid	0
30	MEM 10, MRP 10		MEM	Meropenem	0
31	MN 30, MH 30		MNO	Minocycline	0
32	MXF 5		MXF	Moxifloxacin	0
33	F 300		NIT	Nitrofurantoin	0

34	NOR 10		NOR	Norfloxacin	0
35	OX 1		OXA	Oxacillin	0
36	P 10		PEN	Penicillin G	0
37	TZP 110, Pt	 	TZP	Piperacillin/ Tazobactam	Pt: 12
38	TE 30		TCY	Tetracycline	0
39	TTC 85, TIM 85	 	TCC	Ticarcillin/ Clavulanic acid	TIM 85: 3
40	TOB 10, Tb	 	TOB	Tobramycin	0
41	SXT 25, Bt	 	SXT	Trimethoprim/ Sulfamethoxazole	Bt: 7
42	VA 30		VAN	Vancomycin	0

Kết quả nhận diện nhãn kháng sinh từ hình ảnh đĩa kháng sinh đồ được phân tích dựa trên bốn chỉ số chính: (1) Độ chính xác tổng thể (Accuracy), (2) Độ nhạy (Sensitivity), (3) Độ đặc hiệu (Specificity), và (4) F1-score (để đánh giá cân bằng giữa độ nhạy và độ chính xác). Bảng 4 tóm tắt các chỉ số thống kê về khả năng nhận diện nhãn kháng sinh của mô hình AI so với nhãn thực tế.

- Tổng số nhãn kháng sinh cần nhận diện: 2786 nhãn
- AI nhận diện đúng nhãn lần đầu: 2711 nhãn
- Sai nhãn (lần đầu): 75 nhãn
- AI nhận diện cùng HITL (Human In The Loop): 2786 nhãn (tức 100% đúng sau khi có KTV hiệu chỉnh)
- AI nhận diện khi đã áp dụng Fine-Tuning: 2775 nhãn đúng, 11 nhãn sai

Bảng 4. Chỉ số đánh giá hiệu năng nhận diện nhãn kháng sinh so với kỹ thuật viên

Chỉ số	AI (chưa sử dụng Fine-Tuning)	AI + HITL	AI + Fine-Tuning
True-positive	2711	2786	2775
False-negative	0	0	0
False-positive	75	0	11
True-negative	0	0	0
Accuracy	~97,3%	100,00%	~99,6%
Sensitivity	100,00%	100,00%	100,0%
Precision	~97,3%	100,00%	~99,6%
F1-score	~98,6	100,00%	~99,8%

Do trong bài toán này chỉ có hai kết quả “đúng” hoặc “sai”, hoàn toàn không xuất hiện trường hợp âm tính giả (FN = 0) cũng như âm tính thực sự (TN

= 0), nên chỉ số “specificity” không xác định. Kết quả thống kê cho thấy AI (chưa Fine-Tuning) đạt độ chính xác (accuracy) khoảng 97,3% (tương đương

2711/2786 nhãn được nhận diện đúng), tức vẫn còn 75 nhãn bị sai. Nguyên nhân chủ yếu xuất phát từ việc nhãn bị che khuất nên hình ảnh khoan giấy trong các trường hợp này AI khó phân biệt.

Khi kết hợp AI với HITL (có kỹ thuật viên kiểm tra và chỉnh sửa khi cần), toàn bộ 2786 nhãn đều được nhận diện chính xác (accuracy = 100%). Đây là phương pháp thường được áp dụng trong giai đoạn vận hành, giúp khắc phục hoàn toàn các lỗi còn lại; điều này đặc biệt quan trọng khi yêu cầu tính chính xác tuyệt đối cho các ứng dụng trong y tế.

Sau khi áp dụng tinh chỉnh (Fine-Tuning), AI tự động nâng tỷ lệ nhận diện đúng lên đến gần 99,6% (tương đương 2775/2786 nhãn), chỉ còn 11 nhãn sai.

Kết quả này cho thấy tiềm năng cải thiện đáng kể của mô hình khi được huấn luyện thêm với dữ liệu đặc thù, đồng thời giảm bớt đáng kể công sức can thiệp thủ công của kỹ thuật viên. Tuy nhiên, trong một số trường hợp nhãn in quá mờ hoặc chất lượng kém, AI vẫn có thể nhận diện nhầm; do đó, cơ chế kiểm tra hoặc xác nhận từ con người vẫn được khuyến nghị cho các tình huống quan trọng, nhằm đảm bảo tính chính xác tối đa.

3.4. Đánh giá hiệu suất và thời gian phân tích kết quả kháng sinh đồ

Thời gian thực hiện phân tích kết quả kháng sinh đồ bằng cả 2 phương pháp được đo đạc và ghi nhận kết quả.

Bảng 5. So sánh thời gian hoàn thành phân tích kết quả của 2 phương pháp.

Tiêu chí	Đo thủ công bằng thước	Ứng dụng AI + HITL
Thời gian trung bình đo đĩa 5 kháng sinh	60 - 75 giây	20 giây
Thời gian trung bình đo 10 đĩa (phút)	12 - 15 phút	3 - 4 phút
Sai số ước tính (mm)	±0,5 - 1,0 mm	±0,5 - 1,0 mm

Kết quả cho thấy, ứng dụng AI trong phân tích đĩa kháng sinh đồ giúp rút ngắn đáng kể thời gian so với đo thủ công. Thay vì mất vài phút, AI chỉ cần vài giây đến dưới một phút để xử lý mỗi mẫu và tự động lưu kết quả. Điều này nâng cao năng suất, giảm sai sót do con người và giúp kỹ thuật viên tập trung vào các công đoạn quan trọng khác, cải thiện hiệu quả và chất lượng xét nghiệm.

4. BÀN LUẬN

Kết quả nghiên cứu cho thấy việc ứng dụng trí tuệ nhân tạo (AI) trong phân tích kháng sinh đồ phương pháp Kirby-Bauer không chỉ đáp ứng yêu cầu về độ chính xác mà còn rút ngắn đáng kể thời gian thực hiện xét nghiệm. Mặc dù tỷ lệ sai số đo đường kính vùng ức chế ở *E. coli*, *S. aureus* và *P. aeruginosa* tương đối thấp, song vẫn ghi nhận tỷ lệ sai lệch đáng kể khi áp dụng cho *K. pneumoniae* và *A. baumannii* (Bảng 1). Điều này cho thấy sự cần thiết của việc hiệu chỉnh mô hình AI hoặc áp dụng cơ chế can thiệp thủ công (human-in-the-loop, HITL) để kịp thời phát hiện và sửa sai các trường hợp ngoại lệ. Mô hình AI đạt độ chính xác nhận diện nhãn kháng sinh gần 97,31% ở lần chạy đầu, nhưng khi kết hợp với HITL đã khắc phục toàn bộ lỗi (100%) và khi tinh chỉnh (fine-tuning) mô hình, tỷ lệ nhận diện đúng tự động nâng lên 99,60%. Kết quả này khẳng định tầm quan trọng của hai chiến lược: (i) sử dụng HITL để kiểm soát chất lượng và điều chỉnh kịp thời những lỗi phức tạp và (ii) áp dụng quy trình huấn luyện bổ sung (fine-tuning) với dữ liệu đặc thù, giúp giảm đáng kể sự phụ thuộc vào can thiệp thủ công [9,10]. Khái niệm “Human-in-the-Loop” (HITL) đề cập đến cách tiếp cận mà con người được tích hợp vào vòng lặp ra quyết định của AI. Thay vì để hệ thống AI hoạt động hoàn toàn tự động, bác sĩ và nhân viên y tế

tham gia vào các bước quan trọng: giám sát đầu ra của AI, xác nhận hoặc hiệu chỉnh kết quả, và cung cấp phản hồi để cải thiện hệ thống. Mục tiêu của HITL là kết hợp thế mạnh của AI (xử lý nhanh, phân tích mẫu phức tạp) với kiến thức và kinh nghiệm của chuyên gia. Trong lĩnh vực y tế, điều này rất có ý nghĩa nhằm đảm bảo an toàn và chất lượng [11]. Sử dụng HITL không chỉ giám sát mà còn đóng góp vào việc cải tiến AI. Thông qua phản hồi liên tục (ví dụ: bác sĩ đánh giá kết quả dự đoán của AI là đúng hay sai), hệ thống có thể được huấn luyện bổ sung để học từ những sai lầm. Một số hệ thống cho phép bác sĩ điều chỉnh tham số hoặc ngưỡng cảnh báo cho phù hợp với thực tiễn. Về mặt dữ liệu, sự tham gia của con người cũng quan trọng trong gán nhãn dữ liệu huấn luyện: hàng triệu ảnh y khoa cần được chuyên gia chú thích chính xác để dạy cho AI, đây chính là quá trình human-in-the-loop ở giai đoạn phát triển.

Mô hình AI phân tích kháng sinh đồ có thể mất chính xác khi áp dụng tại các phòng xét nghiệm khác nhau do dịch chuyển miền (domain shift) – sự khác biệt về thiết bị xét nghiệm, điều kiện nuôi cấy, và mô hình kháng thuốc theo khu vực. Ví dụ, hình ảnh đĩa kháng sinh đồ có thể khác biệt về độ tương phản, kích thước vùng ức chế, ảnh hưởng đến khả năng nhận diện của AI. Do đó, cần tinh chỉnh (fine-tuning) mô hình bằng cách huấn luyện thêm trên dữ liệu nội

bộ của từng phòng xét nghiệm với tốc độ học thấp. Quá trình này giúp điều chỉnh trọng số để AI nhận diện chính xác hơn vùng ức chế trên đĩa thạch trong điều kiện thực tế [12, 13].

Sự kết hợp giữa AI và HITL trong phân tích kháng sinh đồ cũng đồng thời giải quyết bài toán về tính linh hoạt và khả năng mở rộng. Với những trường hợp nhận kháng sinh không phổ biến hoặc ảnh chụp thiếu chất lượng, kỹ thuật viên (KTV) có thể bổ sung dữ liệu và huấn luyện lại (fine-tuning) để cải thiện mô hình. Cách tiếp cận này đặc biệt hữu ích cho các phòng xét nghiệm quy mô nhỏ hoặc chưa có đủ điều kiện triển khai hoàn toàn hệ thống tự động. Thay vì đòi hỏi đầu tư lớn về trang thiết bị, quy trình này tập trung vào nâng cao chất lượng dữ liệu và cải tiến giải thuật, từ đó gia tăng hiệu quả chẩn đoán và tối ưu chi phí. Hơn nữa, việc rút ngắn thời gian phân tích xuống chỉ còn khoảng 20 giây cho mỗi đĩa kháng sinh đồ (thay vì 60 - 75 giây khi đo thủ công) giúp tiết kiệm nguồn nhân lực đáng kể, tạo điều kiện để KTV tập trung vào các hoạt động khác như bảo đảm chất lượng mẫu, giải thích lâm sàng, hoặc phối hợp với bác sĩ nhằm nâng cao hiệu quả sử dụng kháng sinh.

Một số thách thức vẫn tồn tại. Thứ nhất, các loại vi khuẩn có vòng ức chế chồng nhau hoặc hình ảnh mờ dễ gây nhầm lẫn cho AI, đòi hỏi quy trình chuẩn hoá việc bố trí đĩa kháng sinh trên môi trường thạch và kiểm soát chất lượng ảnh chụp. Thứ hai, trong trường hợp vi khuẩn mọc lấn vào vùng ức chế, mô hình AI thường thiếu dữ liệu đặc trưng để phân biệt ranh giới, dẫn đến sai lệch về đường kính đủ lớn để thay đổi kết luận (Bảng 2). Thứ ba, xây dựng bộ dữ liệu đa dạng để AI huấn luyện là quá trình tốn kém và đòi hỏi nhiều thời gian. Do đó, ngay cả khi áp dụng fine-tuning, sự tham gia chủ động của KTV qua mô hình "AI + HITL" vẫn rất quan trọng để duy trì mức độ tin cậy, đồng thời bảo đảm mô hình có khả năng cập nhật liên tục trong bối cảnh vi khuẩn và kháng sinh đa dạng [14]. Nhằm khắc phục và giảm thiểu các sai số này, cần tối ưu đồng thời cả yếu tố kỹ thuật phòng xét nghiệm và phương pháp xử lý ảnh. Về phía phòng

xét nghiệm, có thể bố trí đĩa kháng sinh cách nhau hợp lý, bảo đảm ánh sáng và góc chụp nhất quán, cũng như kiểm soát chất lượng môi trường nuôi cấy. Về phía AI, cần xây dựng và huấn luyện mô hình với bộ dữ liệu đa dạng hơn, phản ánh đầy đủ các trường hợp thực tế (vòng chồng chéo, vòng mờ, bóng mờ, v.v.), đồng thời tích hợp cơ chế nhận diện hoặc cảnh báo chất lượng ảnh trước khi đo đường kính vùng ức chế. Những cải tiến này sẽ nâng cao độ chính xác, đồng thời giảm thiểu sai lệch trong chẩn đoán và lựa chọn liệu pháp kháng sinh.

Nhìn chung, AI kết hợp với HITL và fine-tuning đã cho thấy tiềm năng lớn trong việc tự động hoá và nâng cao hiệu quả phân tích kháng sinh đồ Kirby-Bauer. Dù vẫn còn những hạn chế về chất lượng hình ảnh, tính phức tạp của từng loài vi khuẩn và chi phí xây dựng dữ liệu, kết quả nghiên cứu này gợi mở hướng tiếp cận mới, giúp tối ưu hoá quy trình xét nghiệm vi sinh lâm sàng. Điều quan trọng là duy trì chiến lược phối hợp "con người-máy" nhằm không ngừng cải tiến độ chính xác, tốc độ và khả năng ứng dụng rộng rãi của hệ thống AI trong bối cảnh đề kháng kháng sinh ngày càng trở nên phức tạp.

5. KẾT LUẬN

Nghiên cứu cho thấy ứng dụng trí tuệ nhân tạo (AI) trong phân tích kháng sinh đồ Kirby-Bauer có tiềm năng nâng cao độ chính xác và rút ngắn thời gian thực hiện so với phương pháp thủ công. Kết hợp cơ chế "Human-In-The-Loop" (HITL) và quy trình tinh chỉnh (fine-tuning) giúp giảm đáng kể sai sót, đặc biệt ở các trường hợp vòng ức chế phức tạp. Giải pháp này mở ra hướng ứng dụng rộng rãi trong lâm sàng, góp phần kiểm soát tốt hơn tình trạng kháng kháng sinh.

LỜI CẢM ƠN

Nhóm nghiên cứu xin gửi lời cảm ơn đến Khoa Vi Sinh, Bệnh viện Trường Đại học Y-Dược Huế đã tham gia hỗ trợ nghiên cứu này. Đây là một phần kết quả của đề tài KHCN cấp Đại học Huế mã số DDH2024-04-208.

TÀI LIỆU THAM KHẢO

1. Bauer AW, Kirby WM, Sherris JC, Turck M. Antibiotic susceptibility testing by a standardized single disk method. *Am J Clin Pathol* 1966;45:493-6.
2. Höring S, Massarani AS, Löffler B, Rödel J. Rapid antibiotic susceptibility testing in blood culture diagnostics performed by direct inoculation using the VITEK®-2 and BD Phoenix™ platforms. *European Journal of Clinical*

- Microbiology & Infectious Diseases* 2019;38:471-8. <https://doi.org/10.1007/s10096-018-03445-3>.
3. Khan ZA, Siddiqui MF, Park S. Current and Emerging Methods of Antibiotic Susceptibility Testing. *Diagnostics* 2019;9:49. <https://doi.org/10.3390/diagnostics9020049>.
4. Balouiri M, Sadiki M, Ibensouda SK. Methods for in vitro evaluating antimicrobial activity: A review. *J*

Pharm Anal 2016;6:71–9. <https://doi.org/10.1016/j.jpha.2015.11.005>.

5. Alonso CA, Domínguez C, Heras J, Mata E, Pascual V, Torres C, et al. Antibioqramj: A tool for analysing images from disk diffusion tests. *Comput Methods Programs Biomed* 2017;143:159–69. <https://doi.org/10.1016/j.cmpb.2017.03.010>.

6. Pascucci M, Royer G, Adamek J, Asmar M Al, Aristizabal D, Blanche L, et al. AI-based mobile application to fight antibiotic resistance. *Nat Commun* 2021;12:1173. <https://doi.org/10.1038/s41467-021-21187-3>.

7. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;25:44–56. <https://doi.org/10.1038/s41591-018-0300-7>.

8. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017;42:60–88. <https://doi.org/10.1016/j.media.2017.07.005>.

9. Mosqueira-Rey E, Hernández-Pereira E, Alonso-Ríos D, Bobes-Bascarán J, Fernández-Leal Á. Human-in-the-loop machine learning: a state of the art. *Artif Intell Rev* 2023;56:3005–54. [https://doi.org/10.1007/s10462-022-](https://doi.org/10.1007/s10462-022-10246-w)

10246-w.

10. Holzinger A. Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Inform* 2016;3:119–31. <https://doi.org/10.1007/s40708-016-0042-6>.

11. Sezgin E. Artificial intelligence in healthcare: Complementing, not replacing, doctors and healthcare providers. *Digit Health* 2023;9. <https://doi.org/10.1177/20552076231186520>.

12. Sahiner B, Chen W, Samala RK, Petrick N. Data drift in medical machine learning: implications and potential remedies. *Br J Radiol* 2023;96. <https://doi.org/10.1259/bjr.20220878>.

13. Kim MJ, Kim SH, Kim SM, Nam JH, Hwang YB, Lim YJ. The Advent of Domain Adaptation into Artificial Intelligence for Gastrointestinal Endoscopy and Medical Imaging. *Diagnostics* 2023;13:3023. <https://doi.org/10.3390/diagnostics13193023>.

14. Feng L, Feng F, Yang Y, Tan M, Wang L. Model adaptive parameter fine-tuning Based on contribution measure for image classification. *Neurocomputing* 2025;632:129634. <https://doi.org/10.1016/j.neucom.2025.129634>.