

ỨNG DỤNG HỌC MÁY TRONG PHÁT HIỆN SỚM TỖN THƯƠNG PHỔI DỰA TRÊN IRIDOLOGY

Phạm Thị Hồng Ngân ⁽¹⁾, Nguyễn Chí Cường ^(2,*), Hồ Đắc Hưng ⁽³⁾

(1) Viện Công nghệ Số, Trường Đại học Thủ Dầu Một
Email: 2224802010434@student.tdmu.edu.vn

(2) Viện Công nghệ Số, Trường Đại học Thủ Dầu Một
Email: 2224802010781@student.tdmu.edu.vn

(3) Viện Công nghệ Số, Trường Đại học Thủ Dầu Một
Email: hunghd@tdmu.edu.vn

(*) Tác giả liên hệ

Ngày nhận bài: 12/4/2026; Chấp nhận đăng: 22/6/2026

DOI: 10.37550/tdmu.EJS/V2026.03.761

Tóm tắt

Phát hiện sớm nguy cơ tổn thương phổi có ý nghĩa quan trọng trong hỗ trợ sàng lọc và can thiệp kịp thời, trong khi các phương pháp như X-quang hoặc CT đòi hỏi thiết bị chuyên dụng và chi phí cao. Nghiên cứu này đề xuất một hướng tiếp cận hỗ trợ dựa trên học sâu nhằm phân loại nguy cơ tổn thương phổi từ ảnh móng mắt theo lý thuyết iridology. Quy trình nghiên cứu gồm tiền xử lý ảnh, trích xuất vùng quan tâm của móng mắt, huấn luyện và so sánh bốn mô hình MobileNetV2, MobileNetV3, Vision Transformer và TinyViT trên cùng bộ dữ liệu. Kết quả thực nghiệm cho thấy TinyViT đạt hiệu năng tốt nhất với Accuracy 91,30%, Recall 95,80%, F1-score 94,48% và thời gian phản hồi 103 ms, cho thấy khả năng cân bằng giữa độ chính xác và chi phí tính toán. Nghiên cứu cũng định hướng tích hợp mô hình vào ứng dụng di động để hỗ trợ sàng lọc ban đầu trong cộng đồng. Hệ thống đề xuất không thay thế các phương pháp chẩn đoán y khoa chuẩn, mà đóng vai trò công cụ hỗ trợ cảnh báo sớm với chi phí thấp và khả năng triển khai linh hoạt.

Từ khóa: Iridology, học máy, MobileNetV2, MobileNetV3, phát hiện tổn thương phổi, TinyViT, Vision Transformer

Abstract

MACHINE LEARNING APPLICATION FOR EARLY DETECTION OF LUNG DAMAGE BASED ON IRIDOLOGY

Early screening of lung damage risk is important for timely health intervention, whereas conventional methods such as X-ray and CT require specialized equipment and relatively high costs. This study proposes a supportive deep learning approach for classifying lung damage risk from iris images based on iridology. The research workflow includes image preprocessing, iris region-of-interest extraction, and the training and comparison of four models, namely MobileNetV2, MobileNetV3, Vision Transformer, and TinyViT, under consistent experimental conditions. The results show that TinyViT achieved the best performance, with an accuracy of 91.30%, recall of 95.80%, F1-score of 94.48%, and response time of 103 ms, indicating a favorable balance between classification performance and computational cost. The study also suggests integrating the selected model into a mobile application for preliminary community-based screening. The proposed system is not intended to replace standard medical diagnostic methods, but to serve as a

low-cost and flexible early warning support tool.

Keywords: *Iridology, machine learning, lung lesion detection, MobileNetV2, MobileNetV3, TinyViT, Vision Transformer*

1. Đặt vấn đề

Phát hiện sớm tổn thương phổi có vai trò quan trọng trong việc nâng cao hiệu quả điều trị và giảm tỷ lệ tử vong do các bệnh lý hô hấp mạn tính cũng như ung thư phổi. Các nghiên cứu đã chỉ ra rằng việc sàng lọc sớm giúp cải thiện đáng kể tiên lượng và cơ hội điều trị cho người bệnh (Aberle và nnk., 2011; Oudkerk và nnk., 2021). Các phương pháp sàng lọc phổ biến hiện nay như chụp X-quang và CT liều thấp cho độ tin cậy cao, tuy nhiên đòi hỏi thiết bị chuyên dụng, chi phí lớn và khó triển khai trên diện rộng, đặc biệt tại các khu vực hạn chế về nguồn lực (National Lung Screening Trial Research Team, 2011). Vì vậy, nhu cầu về các phương pháp sàng lọc không xâm lấn, chi phí thấp và có khả năng ứng dụng cộng đồng đang ngày càng được quan tâm, đặc biệt với sự hỗ trợ của các công nghệ trí tuệ nhân tạo trong y tế (Topol, 2019; Lee và nnk., 2017).

Iridology - phương pháp đánh giá tình trạng cơ quan nội tạng thông qua đặc điểm móng mắt gần đây được xem là một nguồn dữ liệu hình ảnh tiềm năng cho các hệ thống trí tuệ nhân tạo. Khi kết hợp với các mô hình học sâu, nhiều nghiên cứu đã cho thấy khả năng khai thác đặc trưng móng mắt để hỗ trợ sàng lọc nguy cơ bệnh lý. Tiêu biểu, nghiên cứu ReLiris: Lung damage pre-screening via iridology đã đề xuất mô hình học sâu lai giữa CNN (ResNet) và BiLSTM kèm cơ chế self-attention nhằm nhận diện dấu hiệu tổn thương phổi từ ảnh móng mắt, đạt hiệu năng cao trên tập dữ liệu hơn 10.000 ảnh (Huynh và nnk., 2025). Công trình này chứng minh tính khả thi của kiến trúc CNN-LSTM trong việc học đặc trưng không gian và quan hệ ngữ cảnh, tuy nhiên chủ yếu tập trung vào một cấu trúc mô hình và chưa hướng mạnh đến tối ưu hóa cho thiết bị di động.

Bên cạnh đó, nghiên cứu Iris-based lung cancer pre-scanning for mobile platforms đã bước đầu đưa mô hình học sâu phân tích ảnh móng mắt lên nền tảng di động, cho thấy tiềm năng của hướng tiếp cận sàng lọc không xâm lấn và có tính cơ động cao trong thực tế (Ho-Dac và nnk., 2025). Tuy nhiên, nghiên cứu này chủ yếu sử dụng một kiến trúc đơn và chưa thực hiện đánh giá so sánh hệ thống giữa nhiều dòng mô hình học sâu hiện đại.

Xuất phát từ các hạn chế trên, nghiên cứu này đề xuất một hướng tiếp cận có tính hệ thống hơn bằng cách so sánh và đánh giá đa kiến trúc học sâu hiện đại gồm MobileNetV2, MobileNetV3, TinyViT và Vision Transformer (ViT) cho bài toán phát hiện sớm tổn thương phổi dựa trên ảnh móng mắt. Khác với các công trình trước thường tập trung vào một mô hình hoặc một nhóm kiến trúc, nghiên cứu này thực hiện thực nghiệm đối chuẩn trên cùng tập dữ liệu, cùng quy trình tiền xử lý và cùng cấu hình huấn luyện nhằm đảm bảo tính công bằng khi so sánh hiệu năng. Đồng thời, việc đưa các kiến trúc nhẹ tối ưu cho thiết bị biên như MobileNetV3 và TinyViT vào đánh giá giúp làm rõ khả năng cân bằng giữa độ chính xác và chi phí tính toán.

Ngoài ra, nghiên cứu còn tích hợp mô hình có hiệu năng tốt nhất vào ứng dụng di động phục vụ sàng lọc ban đầu, cho phép thu nhận và phân tích ảnh móng mắt trực tiếp trên thiết bị cầm tay. Cách tiếp cận này góp phần thu hẹp khoảng cách giữa nghiên cứu mô hình và khả năng triển khai thực tế, hướng tới một giải pháp sàng lọc hỗ trợ nhanh, chi phí thấp và không xâm lấn trong cộng đồng.

2. Tổng quan tình hình nghiên cứu

2.1. Iridology và ứng dụng sàng lọc bệnh lý

Iridology là phương pháp đánh giá tình trạng sức khỏe dựa trên các đặc điểm hình thái và màu sắc của mống mắt. Theo lý thuyết này, mỗi khu vực trên mống mắt được cho là có liên hệ với một cơ quan nội tạng nhất định trong cơ thể. Đối với hệ hô hấp, một số vùng nhất định trên mống mắt được cho là phản ánh tình trạng phổi (Sharma và nnk., 2018). Trong các hệ thống học máy, ảnh mống mắt thường được tiền xử lý để xác định và cắt vùng quan tâm (Region of Interest - ROI), loại bỏ các yếu tố gây nhiễu như mi mắt, lông mi và vùng củng mạc. Việc chuẩn hóa ROI giúp giảm nhiễu và tăng tính nhất quán của dữ liệu đầu vào cho mô hình học sâu (Ahmed & Yaseen, 2025).

2.2. Học sâu trong phân tích ảnh y sinh

Học sâu (Deep Learning) là một nhánh của học máy sử dụng mạng nơ-ron nhiều tầng để tự động học đặc trưng từ dữ liệu (LeCun và nnk., 2015). Trong bài toán phân tích ảnh y sinh, các mô hình học sâu có khả năng trích xuất đặc trưng không gian phức tạp mà không cần thiết kế thủ công đặc trưng.

Đối với bài toán phát hiện tổn thương phổi dựa trên ảnh mống mắt, học sâu đóng vai trò:

- Tự động học đặc trưng hình thái vi mô
- Phân loại nguy cơ dựa trên embedding học được
- Giảm phụ thuộc vào chuyên gia iridology thủ công

2.3. Vision Transformer (ViT)

Vision Transformer là kiến trúc dựa trên cơ chế self-attention, cho phép mô hình học quan hệ toàn cục giữa các vùng ảnh. Thay vì sử dụng phép tích chập cục bộ như CNN, ViT chia ảnh thành các patch và xử lý chúng như một chuỗi token tương tự mô hình Transformer trong xử lý ngôn ngữ tự nhiên. Cách tiếp cận này giúp mô hình khai thác hiệu quả thông tin toàn cục của ảnh và cải thiện khả năng học đặc trưng (Yang và nnk., 2022).

2.4. Convolutional Neural Network (CNN)

CNN là kiến trúc phổ biến trong xử lý ảnh, sử dụng các lớp tích chập (convolution) để trích xuất đặc trưng cục bộ như cạnh, kết cấu và hoa văn. Các đặc trưng này được học theo từng tầng, từ mức thấp đến mức cao, giúp mô hình nhận diện hiệu quả các cấu trúc trong ảnh (Chua & Roska, 1993).

Trong nghiên cứu này, MobileNetV2, MobileNetV3 đại diện cho nhóm CNN tối ưu hóa cho thiết bị di động. Kiến trúc này sử dụng:

- Depthwise separable convolution
- Cơ chế attention nhẹ
- Thiết kế tối ưu cho tài nguyên thấp

2.5. TinyViT và mô hình nhẹ cho thiết bị biên

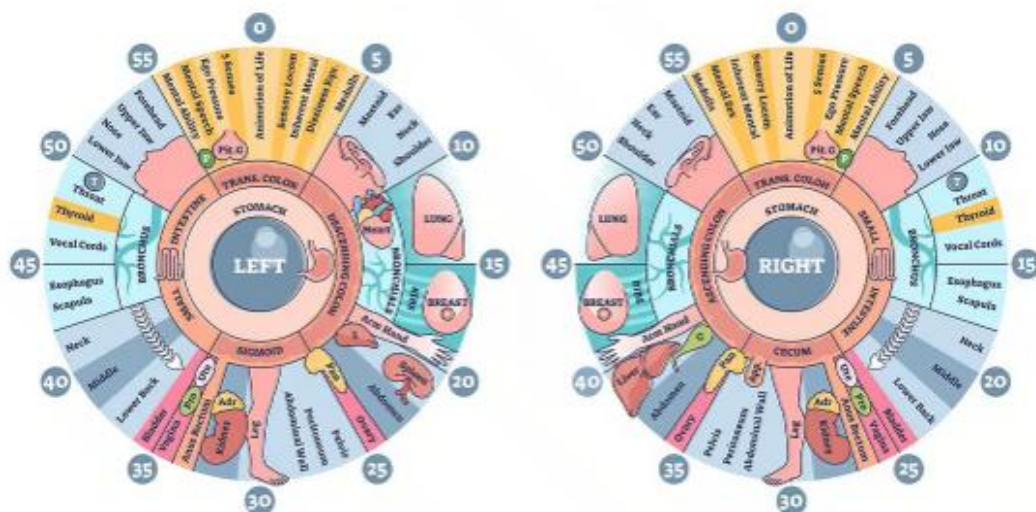
TinyViT là phiên bản tối ưu hóa của Vision Transformer nhằm giảm số tham số và chi phí tính toán nhưng vẫn giữ được khả năng học quan hệ toàn cục. Mô hình này phù hợp cho triển khai trên thiết bị di động hoặc hệ thống tài nguyên hạn chế nhờ kiến trúc nhẹ và hiệu quả (Wu và nnk., 2022).

2.6. Sàng lọc không xâm lấn

Sàng lọc không xâm lấn là phương pháp đánh giá nguy cơ bệnh lý mà không cần can thiệp vào cơ thể như chụp CT hoặc sinh thiết. Phân tích ảnh mống mắt là một hướng tiếp cận không xâm lấn, có tiềm năng ứng dụng trong sàng lọc ban đầu.

3. Phương pháp nghiên cứu

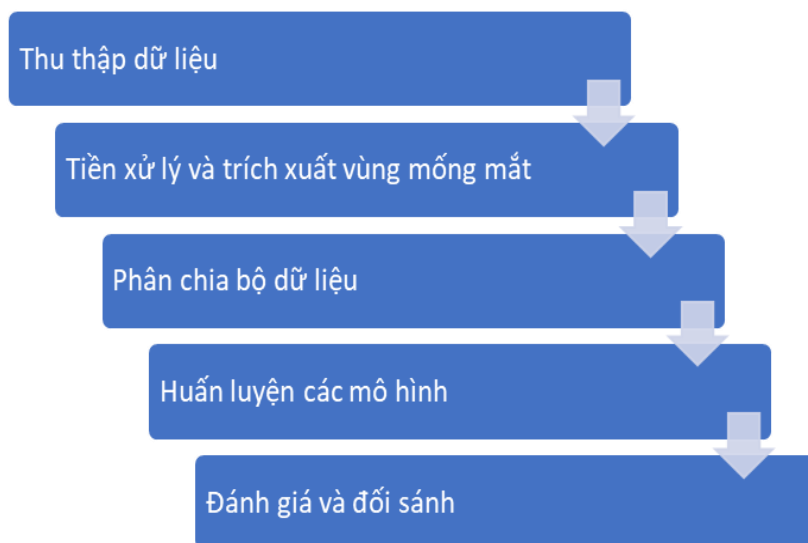
Nghiên cứu này sử dụng một bộ hình ảnh mống mắt cụ thể để kiểm tra các dấu hiệu tổn thương phổi. Để tránh sai lệch, tập dữ liệu được cân bằng. Một vùng quan tâm (ROI) bên trong mống mắt được phát hiện để thu thập dữ liệu liên quan đến phổi dựa trên lý thuyết mống mắt học. Biểu đồ mống mắt học cho thấy vùng này, nằm giữa rìa đồng tử và rìa ngoài của mống mắt, đại diện cho vùng kết nối với phổi (Yogeshwaran và nnk, 2023) (Hình 1). Một hệ thống nhận dạng mống mắt tinh vi đã được sử dụng để tự động xác định và trích xuất ROI. Phương pháp này mang lại phân đoạn chính xác bằng cách xác định tâm và bán kính của cả đồng tử và mống mắt. Vùng 2h–4h (hoặc 8h–10h) được trích xuất từ dữ liệu này bằng cách sử dụng mặt nạ hình quạt, đảm bảo rằng nghiên cứu chỉ tập trung vào phần mống mắt liên quan đến phổi.



Hình 1. Iridology chart

Nguồn: Yogeshwaran và nnk, 2023

Chúng tôi đã sử dụng một tập dữ liệu gồm 3488 hình ảnh thuộc hai lớp: âm tính và dương tính. Các hình ảnh được thu thập và dán nhãn bởi các bác sĩ và chuyên gia dựa trên các phương pháp kiểm tra đã được phê duyệt. Chúng tôi đã chia tập dữ liệu thành các tập huấn luyện, xác nhận và kiểm tra với tỷ lệ lần lượt là 70%, 15% và 15%. Quy trình thực nghiệm cụ thể được mô hình hóa trong hình 2.



Hình 2. Quy trình thực nghiệm

Nguồn: Tác giả, 2026

Do đặc thù của bài toán liên quan đến lĩnh vực y sinh học, hiệu năng của các mô hình được đánh giá thông qua nhiều chỉ số phân loại phổ biến bao gồm: Accuracy, Precision, Recall (Sensitivity), F1-score và thời gian phản hồi. Trong các hệ thống hỗ trợ phát hiện bệnh sớm, Recall và Precision đóng vai trò đặc biệt quan trọng, vì chúng phản ánh khả năng của mô hình trong việc phát hiện đúng các trường hợp bệnh lý đồng thời hạn chế các cảnh báo sai.

- Recall (Sensitivity): thể hiện tỷ lệ các trường hợp dương tính thực sự được mô hình phát hiện đúng. Đây là chỉ số quan trọng trong các hệ thống sàng lọc bệnh.
- Precision: biểu thị tỷ lệ các dự đoán dương tính của mô hình thực sự là đúng.
- F1-score: là trung bình điều hòa của Precision và Recall, phản ánh sự cân bằng giữa hai chỉ số này.

4. Kết quả và thảo luận

Bảng 1 trình bày kết quả thực nghiệm của bốn mô hình học sâu gồm MobileNetV2, MobileNetV3, Vision Transformer (ViT) và TinyViT trong bài toán phát hiện sớm tổn thương võng mạc từ ảnh võng mạc dựa trên lý thuyết Iridology.

Bảng 1. So sánh hiệu năng của các mô hình học sâu

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Response time (ms)
MobileNetV2	57.07	93.24	48.25	63.60	185
MobileNetV3	77.17	97.20	72.72	83.20	206
ViT	78.26	90.55	80.42	85.19	521
TinyViT	91.30	93.20	95.80	94.48	103

Nguồn: Tác giả, 2026

Mô hình TinyViT đạt hiệu năng tốt nhất, với Accuracy đạt 91.30%, Recall đạt 95.80% và F1-score đạt 94.48%, đồng thời có thời gian phản hồi thấp nhất (103 ms).

Điều này cho thấy TinyViT có khả năng phát hiện chính xác các dấu hiệu tổn thương phổi trong ảnh mỏng mắt và có thể phù hợp cho các hệ thống hỗ trợ chẩn đoán trong thời gian thực.

Mô hình Vision Transformer (ViT) đạt Accuracy 78.26% và F1-score 85.19%. Mặc dù khả năng phát hiện bệnh khá tốt với Recall đạt 80.42%, nhưng thời gian phản hồi của mô hình khá cao (521 ms). Điều này có thể giải thích bởi kiến trúc Transformer có chi phí tính toán lớn hơn so với các mạng CNN truyền thống.

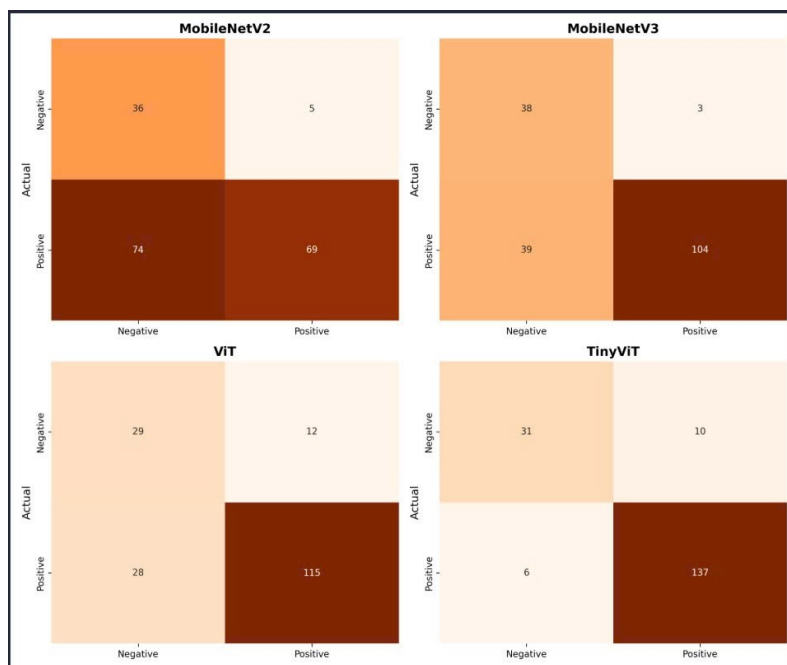
Trong khi đó, MobileNetV3 đạt Accuracy 77.17% và F1-score 83.20%. Điểm đáng chú ý là mô hình này có Precision rất cao (97.20%), cho thấy khi mô hình dự đoán một trường hợp là dương tính thì khả năng đúng là rất cao. Tuy nhiên, Recall chỉ đạt 72.72%, nghĩa là vẫn còn một số trường hợp bệnh chưa được phát hiện.

Ngược lại, MobileNetV2 có hiệu năng thấp nhất, với Accuracy chỉ đạt 57.07% và Recall đạt 48.25%. Mặc dù Precision vẫn ở mức cao (93.24%), nhưng Recall thấp cho thấy mô hình bỏ sót nhiều trường hợp có tổn thương phổi, điều này làm giảm tính hữu ích của mô hình trong các ứng dụng sàng lọc bệnh sớm.

Nhìn chung, kết quả cho thấy các kiến trúc Transformer cải tiến như TinyViT có khả năng học đặc trưng tốt hơn trên dữ liệu ảnh mỏng mắt so với các mô hình CNN nhẹ như MobileNetV2 và MobileNetV3.

4.1. Phân tích ma trận nhầm lẫn (Confusion Matrix)

Hình 3 trình bày ma trận nhầm lẫn của bốn mô hình nhằm phân tích chi tiết khả năng phân loại giữa hai lớp có tổn thương phổi (Positive) và không có tổn thương phổi (Negative).



Hình 3. Kết quả thống kê các mô hình MobileNetV2, MobileNetV3, ViT, và TinyViT

Nguồn: Tác giả, 2026

Đối với MobileNetV2, mô hình dự đoán đúng 36 mẫu âm tính và 69 mẫu dương tính. Tuy nhiên, số lượng False Negative (74 trường hợp) khá cao, cho thấy mô hình bỏ

sót nhiều bệnh nhân có tổn thương phổi. Điều này giải thích vì sao Recall của mô hình thấp và hiệu năng tổng thể chưa cao.

Mô hình MobileNetV3 cho kết quả tốt hơn với 38 mẫu âm tính và 104 mẫu dương tính được dự đoán chính xác. Tuy nhiên, vẫn còn 39 trường hợp False Negative, cho thấy mô hình vẫn chưa phát hiện được toàn bộ các trường hợp bệnh.

Đối với Vision Transformer (ViT), mô hình cho thấy khả năng phát hiện bệnh tốt hơn với 115 trường hợp True Positive, đồng thời số lượng False Negative giảm xuống còn 28. Điều này cho thấy ViT có khả năng học được các đặc trưng toàn cục trong ảnh mờ mắt, giúp cải thiện khả năng phát hiện tổn thương phổi.

Mô hình TinyViT đạt kết quả tốt nhất trong các ma trận nhầm lẫn. Cụ thể, mô hình dự đoán đúng 137 trường hợp dương tính và 31 trường hợp âm tính, trong khi số lượng False Negative chỉ còn 6 trường hợp. Điều này cho thấy TinyViT có khả năng phát hiện bệnh rất cao và hạn chế tối đa việc bỏ sót bệnh nhân.

Trong các ứng dụng y học, giảm số lượng False Negative là yếu tố đặc biệt quan trọng, vì việc bỏ sót bệnh nhân có thể dẫn đến chậm trễ trong quá trình điều trị. Do đó, kết quả từ ma trận nhầm lẫn cho thấy TinyViT là mô hình phù hợp nhất cho bài toán phát hiện sớm tổn thương phổi dựa trên Iridology.

4.2. Thảo luận

Từ các kết quả thực nghiệm có thể thấy rằng các mô hình dựa trên Transformer có xu hướng đạt hiệu năng tốt hơn trong bài toán phân tích ảnh mờ mắt. Điều này có thể được giải thích bởi khả năng của cơ chế self-attention, cho phép mô hình học được các mối quan hệ không gian dài hạn trong ảnh.

Trong dữ liệu iridology, các dấu hiệu liên quan đến tổn thương phổi thường xuất hiện dưới dạng những vùng kết cấu nhỏ và phân bố không đồng đều trên móng mắt. Do đó, các mô hình có khả năng khai thác thông tin toàn cục như Transformer sẽ có lợi thế hơn so với các mạng CNN truyền thống vốn chủ yếu tập trung vào các đặc trưng cục bộ.

Bên cạnh đó, kiến trúc TinyViT được thiết kế nhằm tối ưu hóa hiệu quả tính toán, giúp giảm độ phức tạp của cơ chế attention trong khi vẫn duy trì khả năng biểu diễn đặc trưng mạnh mẽ. Điều này giúp mô hình đạt được sự cân bằng tốt giữa độ chính xác và tốc độ xử lý, làm cho TinyViT trở thành lựa chọn phù hợp cho các hệ thống hỗ trợ chẩn đoán tự động.

Nghiên cứu này chưa triển khai khía cạnh Explainable AI. Trong các nghiên cứu tiếp theo, khía cạnh Explainable AI sẽ được triển khai để tăng khả năng tin cậy cũng như khả năng kiểm chứng của hướng tiếp cận.

5. Thách thức và hướng nghiên cứu mở

Một trong những thách thức quan trọng của nghiên cứu này liên quan đến cơ sở khoa học của iridology. Mặc dù phương pháp này được sử dụng trong một số hệ thống sàng lọc thay thế, tính chính xác và độ tin cậy lâm sàng của iridology vẫn còn gây tranh luận. Ernst đã chỉ ra rằng các bằng chứng thực nghiệm ủng hộ iridology còn hạn chế và chưa đủ mạnh để sử dụng như một công cụ chẩn đoán y khoa chính thức (Ernst, 2000). Do đó, trong bối cảnh nghiên cứu này, phương pháp phân tích ảnh mờ mắt nên được định vị như một công cụ hỗ trợ sàng lọc ban đầu (pre-screening), thay vì thay thế các kỹ

thuật chẩn đoán tiêu chuẩn như X-quang hoặc CT. Thêm vào đó, vào thời điểm công bố của Ernst, các trang thiết bị chụp ảnh mỏng mắt và các kỹ thuật xử lý tính toán còn hạn chế. Đến nay, với cách tiếp cận computational iridology dựa trên các thiết bị chụp ảnh mỏng mắt tiên tiến và các kỹ thuật xử lý tính toán hiệu năng cao đã cho kết quả tốt trên nhiều nghiên cứu (Ahmed & Yaseen, 2025).

Thách thức thứ hai liên quan đến dữ liệu và khả năng tổng quát hóa của mô hình học sâu. Trong lĩnh vực hình ảnh y khoa, hiệu năng của mô hình phụ thuộc lớn vào kích thước, chất lượng và tính đa dạng của dữ liệu huấn luyện. Theo tổng quan về deep learning trong y học, các hệ thống AI cần tập dữ liệu lớn và được gán nhãn chính xác để đạt độ tin cậy cao (Lee và nnk., 2017). Trong trường hợp ảnh mỏng mắt, sự khác biệt về thiết bị chụp, điều kiện ánh sáng và đặc điểm sinh học cá nhân có thể gây sai lệch phân phối dữ liệu, dẫn đến hiện tượng overfitting hoặc giảm hiệu năng khi triển khai thực tế.

Ngoài ra, vấn đề khả năng giải thích của mô hình cũng là một thách thức đáng lưu ý. Các mô hình học sâu thường hoạt động như “hộp đen”, khó lý giải cơ sở đưa ra quyết định. Trong môi trường y khoa, tính minh bạch là yêu cầu quan trọng nhằm đảm bảo độ tin cậy và khả năng chấp nhận lâm sàng (Lee và nnk., 2017). Nếu không có các kỹ thuật hỗ trợ như trực quan hóa attention hoặc Grad-CAM, rất khó xác định liệu mô hình đang học các đặc trưng liên quan đến bệnh lý hay chỉ khai thác các mẫu nhiễu trong dữ liệu.

Mặc dù còn tồn tại nhiều thách thức, tiềm năng của phương pháp không xâm lấn trong chẩn đoán sớm bệnh phổi vẫn được ghi nhận. Nghiên cứu của Rehman và cộng sự đã cho thấy các thuật toán học máy có thể hỗ trợ chẩn đoán sớm bệnh phổi tắc nghẽn thông qua dữ liệu không xâm lấn, góp phần mở rộng khả năng sàng lọc chi phí thấp (Rehman và nnk., 2022). Kết quả này củng cố quan điểm rằng các tín hiệu sinh học không xâm lấn, khi được xử lý bằng học máy, có thể đóng vai trò bổ trợ trong hệ thống phát hiện sớm bệnh lý.

Từ những phân tích trên, một số hướng nghiên cứu mở có thể được đề xuất. Thứ nhất, cần thực hiện các nghiên cứu đối chứng lâm sàng nhằm so sánh dự đoán từ ảnh mỏng mắt với kết quả chẩn đoán dựa trên phương pháp chuẩn như CT hoặc X-quang, nhằm đánh giá độ tin cậy của hệ thống trong môi trường y tế thực tế (Kelly và nnk., 2019). Thứ hai, mở rộng quy mô và đa dạng nguồn dữ liệu sẽ giúp cải thiện tính tổng quát hóa của mô hình (Wang và nnk., 2020). Thứ ba, tích hợp các phương pháp Explainable AI để tăng tính minh bạch và khả năng giải thích của hệ thống (Ribeiro và nnk., 2016; Lundberg & Lee, 2017). Cuối cùng, hướng tiếp cận đa phương thức kết hợp ảnh mỏng mắt với dữ liệu lâm sàng hoặc sinh lý khác có thể nâng cao độ chính xác và giảm rủi ro phụ thuộc vào một nguồn dữ liệu duy nhất (Baltrušaitis và nnk., 2019).

Nhìn chung, mặc dù còn nhiều hạn chế về cơ sở khoa học và dữ liệu, việc ứng dụng học sâu vào các phương pháp sàng lọc không xâm lấn vẫn mở ra tiềm năng phát triển các công cụ hỗ trợ phát hiện sớm tổn thương phổi, đặc biệt trong bối cảnh cần các giải pháp chi phí thấp và dễ tiếp cận. Nhiều nghiên cứu đã chỉ ra rằng các mô hình học sâu có thể hỗ trợ hiệu quả trong phát hiện sớm bệnh lý thông qua phân tích dữ liệu hình ảnh và tín hiệu sinh học mà không cần can thiệp xâm lấn (LeCun và nnk., 2015; Lee và nnk., 2017). Ngoài ra, các hệ thống AI nhẹ triển khai trên thiết bị di động cũng đang được quan tâm nhằm mở rộng khả năng tiếp cận trong cộng đồng, đặc biệt tại các khu vực thiếu điều kiện y tế chuyên sâu (Wu và nnk., 2022). Bên cạnh đó, các phương pháp sàng lọc không xâm lấn kết hợp học máy đã cho thấy tiềm năng trong phát hiện sớm các

bệnh hô hấp, giúp giảm chi phí và hỗ trợ quyết định lâm sàng ban đầu (Rehman và nnk., 2022). Những kết quả này cho thấy việc tiếp tục nghiên cứu và cải tiến các mô hình học sâu cho mục tiêu sàng lọc sớm là hướng đi khả thi và có ý nghĩa thực tiễn trong tương lai.

6. Kết luận

Nghiên cứu này đã đề xuất và đánh giá một hướng tiếp cận ứng dụng học sâu trong phát hiện sớm tổn thương phổi dựa trên ảnh móng mắt theo lý thuyết iridology. Bốn kiến trúc hiện đại gồm MobileNetV2, MobileNetV3, TinyViT và Vision Transformer (ViT) đã được triển khai và so sánh trên cùng tập dữ liệu với quy trình tiền xử lý thống nhất nhằm đảm bảo tính khách quan trong đánh giá. Kết quả thực nghiệm cho thấy các mô hình học sâu có khả năng khai thác đặc trưng từ ảnh móng mắt và đạt hiệu năng phân loại khả quan trong bài toán sàng lọc ban đầu.

Bên cạnh việc so sánh đa kiến trúc, nghiên cứu còn hướng tới khả năng triển khai thực tế thông qua tích hợp mô hình tối ưu lên nền tảng di động. Điều này góp phần chứng minh tiềm năng xây dựng một công cụ hỗ trợ sàng lọc không xâm lấn, chi phí thấp và có tính cơ động cao. So với các nghiên cứu trước chủ yếu tập trung vào một kiến trúc đơn lẻ hoặc chưa chú trọng tối ưu hóa cho thiết bị di động, nghiên cứu này cung cấp một góc nhìn hệ thống hơn về sự đánh đổi giữa độ chính xác và chi phí tính toán.

Tuy nhiên, nghiên cứu vẫn tồn tại những hạn chế nhất định. Cơ sở khoa học của iridology vẫn còn gây tranh luận trong cộng đồng y học (Ernst, 2000), do đó hệ thống đề xuất không nhằm thay thế các phương pháp chẩn đoán chuẩn như X-quang hoặc CT mà chỉ đóng vai trò công cụ hỗ trợ sàng lọc ban đầu. Ngoài ra, kích thước và tính đa dạng của tập dữ liệu còn hạn chế, có thể ảnh hưởng đến khả năng tổng quát hóa của mô hình khi triển khai trong môi trường thực tế (Lee và nnk., 2017).

Trong tương lai, cần thực hiện các nghiên cứu đối chứng lâm sàng để đánh giá mức độ tương quan giữa dự đoán từ ảnh móng mắt và kết quả chẩn đoán y khoa tiêu chuẩn, từ đó xác định độ tin cậy khi triển khai trong môi trường thực tế (Kelly và nnk., 2019). Đồng thời, việc mở rộng dữ liệu đa trung tâm với quy mô lớn và đa dạng hơn được kỳ vọng sẽ cải thiện khả năng tổng quát hóa của mô hình học sâu (Wang và nnk., 2020). Bên cạnh đó, tích hợp các phương pháp Explainable AI có thể giúp tăng tính minh bạch, hỗ trợ bác sĩ hiểu rõ hơn quá trình ra quyết định của hệ thống và nâng cao mức độ tin cậy trong ứng dụng lâm sàng (Ribeiro và nnk., 2016; Lundberg & Lee, 2017). Ngoài ra, hướng tiếp cận đa phương thức kết hợp ảnh móng mắt với dữ liệu lâm sàng, tín hiệu sinh lý hoặc thông tin bệnh sử được cho là có tiềm năng nâng cao độ chính xác và giảm phụ thuộc vào một nguồn dữ liệu duy nhất (Baltrušaitis và nnk., 2019). Nếu được xác thực đầy đủ, hướng tiếp cận này có thể góp phần phát triển các công cụ sàng lọc sớm bệnh lý hô hấp dựa trên trí tuệ nhân tạo, đặc biệt trong bối cảnh cần các giải pháp tiếp cận cộng đồng, tối ưu chi phí và hỗ trợ phát hiện sớm tại tuyến cơ sở.

Lời cảm ơn

Nghiên cứu này được tài trợ bởi Trường Đại học Thủ Dầu Một trong đề tài mã số DTSV.26.2-007

TÀI LIỆU THAM KHẢO

- [1] Aberle, D. R., Adams, A. M., Berg, C. D., Black, W. C., Clapp, J. D., Fagerstrom, R. M., ... & National Lung Screening Trial Research Team. (2011). Reduced lung-cancer mortality with low-dose computed tomographic screening. *New England Journal of Medicine*, 365(5), 395–409.
- [2] Ahmed, A. A., & Yaseen, M. T. (2025). Recent advances in iridology based disease detection: A comprehensive review. *International Journal of Scientific Engineering Research (IJSER)*. <https://doi.org/10.58564/IJSER.4.1.2025.291>
- [3] Baltrušaitis, T., Ahuja, C., & Morency, L.-P. (2019). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2), 423–443.
- [4] Chua, L. O., & Roska, T. (1993). The CNN paradigm. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 40(3), 147–156. <https://doi.org/10.1109/81.222795>
- [5] Ernst, E. (2000). Iridology: Not useful and potentially harmful. *Archives of Ophthalmology*, 118(1), 120–121. <https://doi.org/10.1001/archophth.118.1.120>
- [6] Ho-Dac, H. (2025). Dual loss and Vision Transformer integration for iris-based early lung damage detection.
- [7] Ho-Dac, H., Le, T. A., & Huynh, T. T. (2025). Iris-based lung cancer pre-scanning for mobile platforms.
- [8] Huynh, T. T., Huynh, D. T., Vu, H. P., & Nguyen, N. (2025). ReLiris: Lung damage pre-screening via iridology.
- [9] Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *Nature Medicine*, 25(1), 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
- [10] Lee, J.-G., Jun, S., Cho, Y.-W., Lee, H., Kim, G. B., Seo, J. B., & Kim, N. (2017). Deep learning in medical imaging: General overview. *Korean Journal of Radiology*, 18(4), 570–584. <https://doi.org/10.3348/kjr.2017.18.4.570>
- [11] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444. <https://doi.org/10.1038/nature14539>
- [12] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems (NeurIPS)*.
- [13] Oudkerk, M., Devaraj, A., Vliegthart, R., Henzler, T., Prokop, M., Heussel, C. P., ... & Sverzellati, N. (2021). European position statement on lung cancer screening. *The Lancet Oncology*, 22(12), e558–e566.
- [14] Rehman, M. U., Driss, M., Khakimov, A., & Khalid, S. (2022). Non-invasive early diagnosis of obstructive lung diseases leveraging machine learning algorithms. *Computers, Materials & Continua*, 72(3), 5681–5697.
- [15] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD*, 1135–1144.
- [16] Sharma, A., Kumar, R., & Singh, M. (2018). Pre-diagnostic tool to predict obstructive lung diseases using iris recognition system. In *Smart Innovations in Communication and Computational Sciences* (pp. 71–79). Springer.
- [17] Topol, E. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56.

- [18] Wang, S., Summers, R. M., & Yao, J. (2020). Machine learning and radiology. *Medical Image Analysis*, 63, 101693. <https://doi.org/10.1016/j.media.2020.101693>
- [19] Wu, K., Zhang, J., Peng, H., Liu, M., Xiao, B., Fu, J., & Yuan, L. (2022). TinyViT: Fast pretraining distillation for small Vision Transformers. In *Computer Vision – ECCV 2022* (pp. 68–85). Springer.
- [20] Yang, C., Wang, Y., Zhang, J., Zhang, H., Wei, Z., Lin, Z., & Yuille, A. (2022). Lite Vision Transformer with enhanced self-attention. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 11998–12008).
- [21] Yogeshwaran, R., Vishal, S., et al.: Disease detection based on iris recognition. In: 2023 International Conference on Energy, Materials and Communication Engineering (ICEMCE). pp. 1–5. IEEE (2023)