

DEEP LEARNING FOR BREAST CANCER RECOGNITION

ỨNG DỤNG KỸ THUẬT HỌC SÂU CHO BÀI TOÁN PHÁT HIỆN UNG THƯ VÚ

Vũ Hữu Đạo¹, Hoàng Văn Thành²

¹Học viện Công nghệ Bưu chính Viễn thông

²Trường Đại học Quảng Bình

ABSTRACT: Breast cancer is one of the most prevalent cancers among women and has a high mortality rate if not detected and treated in a timely manner. This study presents research findings on the application of artificial intelligence in breast cancer diagnosis, aiming to assist physicians in early disease detection. We employ advanced machine learning models to analyze medical images and evaluate their performance based on accuracy, sensitivity, and specificity metrics. The findings highlight the potential of artificial intelligence in enhancing diagnostic efficiency, improving treatment quality, and minimizing risks for patients.

Keywords: Breast cancer, deep learning techniques, CNN.

TÓM TẮT: Ung thư vú là một trong những bệnh ung thư phổ biến nhất ở phụ nữ và có tỷ lệ tử vong cao nếu không được phát hiện và điều trị kịp thời. Bài báo này trình bày kết quả nghiên cứu về ứng dụng trí tuệ nhân tạo trong chẩn đoán ung thư vú, nhằm hỗ trợ các bác sĩ trong quá trình phát hiện bệnh sớm. Chúng tôi sử dụng các mô hình học máy tiên tiến để phân tích hình ảnh y khoa và đánh giá hiệu suất của chúng dựa trên các chỉ số chính xác, độ nhạy và độ đặc hiệu. Kết quả nghiên cứu cho thấy tiềm năng ứng dụng của trí tuệ nhân tạo trong cải thiện hiệu quả chẩn đoán ung thư vú, góp phần nâng cao chất lượng điều trị và giảm thiểu rủi ro cho bệnh nhân.

Từ khóa: Ung thư vú, kỹ thuật học sâu, CNN.

1. GIỚI THIỆU

Ung thư vú là một trong những loại ung thư phổ biến và nguy hiểm nhất đối với phụ nữ trên toàn thế giới. Theo báo cáo mới nhất vào năm 2022 của Tổ chức Nghiên cứu Quốc tế về Ung thư (IARC) [6], hơn 2.3 triệu phụ nữ được chẩn đoán mắc bệnh ung thư vú trong năm 2018. Cùng với đó, số ca tử vong do căn bệnh này lên đến 670.000 trường hợp, cho thấy mức độ nghiêm trọng của vấn đề. Dự báo đến năm 2050, số ca mắc ung thư vú mới và chết vì ung thư vú có thể tăng lên lần lượt là 38% và 68%.

Tại các quốc gia đang phát triển, bao gồm cả Việt Nam, việc nhận thức về các triệu chứng của ung thư vú còn hạn chế. Nhiều bệnh nhân chỉ phát hiện bệnh khi đã

ở giai đoạn muộn, làm gia tăng nguy cơ tử vong. Hơn nữa, sự thiếu hụt các chuyên gia y tế và thiết bị chẩn đoán hiện đại, đặc biệt ở vùng sâu, vùng xa, cũng là một rào cản lớn trong việc phát hiện sớm và điều trị kịp thời căn bệnh này. Do đó, việc áp dụng công nghệ thông tin và trí tuệ nhân tạo (AI) trong hỗ trợ chẩn đoán ung thư vú đang trở thành một hướng đi đầy tiềm năng, giúp cải thiện độ chính xác và tốc độ chẩn đoán, qua đó nâng cao hiệu quả điều trị và giảm tỷ lệ tử vong.

Dữ liệu chụp X-quang tuyến vú (mammography) là một trong những phương pháp hình ảnh y học quan trọng giúp phát hiện sớm ung thư vú. Trong những năm gần đây, nhiều nghiên cứu đã

chứng minh rằng các thuật toán AI có khả năng phân tích hình ảnh X-quang vú để hỗ trợ sàng lọc ung thư. Nhiều quốc gia như Hoa Kỳ, Phần Lan, Hà Lan, Pháp, Hàn Quốc và Trung Quốc đã triển khai nghiên cứu về lĩnh vực này, góp phần cải thiện độ chính xác trong chẩn đoán [8].

Một trong những nghiên cứu tiêu biểu là nghiên cứu của McKinney và cộng sự (2020), được thực hiện tại Anh và Hoa Kỳ và công bố trên tạp chí Nature. Nhóm nghiên cứu đã thiết kế một mô hình AI để phân tích hình ảnh X-quang vú của gần 29.000 phụ nữ [9]. Kết quả cho thấy thuật toán AI có độ chính xác cao hơn so với sáu bác sĩ X-quang độc lập. Khi so sánh với tiêu chuẩn chẩn đoán “chuẩn vàng” (gold standard) dựa trên ảnh scan giải phẫu bệnh của Hoa Kỳ, mô hình AI giúp giảm 9,4% trường hợp âm tính giả và 5,7% trường hợp dương tính giả. Ở Anh, hệ thống giúp giảm 2,7% âm tính giả và 1,2% dương tính giả. Dữ liệu nghiên cứu được thu thập từ các trung tâm y tế uy tín như Bệnh viện St George, Trung tâm Vú Jarvis và Bệnh viện Addenbrooke [9].

Bên cạnh nghiên cứu của McKinney, các phương pháp học sâu, đặc biệt là mạng nơ-ron tích chập (CNN), cũng được ứng dụng rộng rãi trong chẩn đoán ung thư vú. Dhungel và cộng sự (2017) đã đề xuất mô hình mạng nơ-ron dư sâu (ResNet) để phân loại ảnh ung thư vú thành hai nhóm: ác tính và lành tính. Mô hình sử dụng sáu ảnh đầu vào từ các tư thế chụp khác nhau và đạt độ chính xác cao khi thử nghiệm trên tập dữ liệu INBreast [10].

Levy và Jain (2016) cũng đã sử dụng mạng CNN kết hợp với học chuyển để phân loại u vú trên tập dữ liệu DDSM. Kết quả cho thấy độ nhạy (recall) đạt 0,934, cao hơn

so với mức 0,745 - 0,923 của các bác sĩ chẩn đoán hình ảnh [7]. Tương tự, Li Shen (2019) đã đề xuất một thuật toán AI phân tích toàn bộ ảnh nhũ ảnh, đạt AUC = 0,88 trên tập dữ liệu DDSM và AUC = 0,91 trên tập INBreast.

Ngoài ra, Xi và cộng sự (2018) đã nghiên cứu phương pháp phân loại và khoanh vùng canxi cũng như khối u trong ảnh chụp X-quang vú. Nhóm nghiên cứu sử dụng mô hình VGGNet và đạt độ chính xác 92,53% trong phân loại vùng tổn thương [15]. Một nghiên cứu khác của Đại học New York (2019) đã phát triển hệ thống AI để phân loại các lớp BI-RADS trong ảnh X-quang vú. Hệ thống này áp dụng mạng nơ-ron tích chập sâu trên 886.000 ảnh X-quang và đạt kết quả macAUC = 0,688, tương đương với các bác sĩ chuyên khoa (macAUC = 0,704) [4].

Trong những năm gần đây, nhiều nghiên cứu đã khám phá các phương pháp học chuyển giao trong chẩn đoán ung thư bằng hình ảnh y tế. Sandhya Aneja và cộng sự (2022) đã tiến hành so sánh toàn diện 14 mô hình CNN được huấn luyện trước trên hình ảnh mô bệnh học, trong đó DenseNet161 đạt hiệu suất cao nhất với AUC là 0.9924 và F1 score là 0.95 trong trích xuất đặc trưng. Kết quả của họ cho thấy mặc dù DenseNet161 vượt trội về độ chính xác, nhưng ResNet101 lại có độ nhạy cao hơn. Đáng chú ý, họ quan sát thấy rằng việc tinh chỉnh mô hình mang lại ít lợi ích về hiệu suất, cho thấy rằng phương pháp trích xuất đặc trưng đã nắm bắt hiệu quả các mẫu liên quan [3].

Dựa trên các ứng dụng của học chuyển giao, Tuan Linh Le và cộng sự (2023) đã nghiên cứu tối ưu hóa ResNet-34 cho phân loại X-quang ung thư vú theo tiêu chuẩn BI-

RADS. Phương pháp của họ bao gồm tiền huấn luyện trên 10.418 hình ảnh từ cơ sở dữ liệu DDSM, sau đó tinh chỉnh với 6.248 hình ảnh từ cơ sở dữ liệu HMU của Việt Nam. Mô hình đạt macAUC là 0.766, vượt qua hiệu suất của ResNet-50 (macAUC = 0,754) khi chỉ sử dụng dữ liệu HMU. Hệ thống cho thấy hiệu suất xuất sắc trong việc xác định các trường hợp BI-RADS 1 (bình thường), nhưng gặp khó khăn hơn trong phân loại BI-RADS 23 (lành tính)[14].

Năm 2024, Shuaa S. Alharbi và cộng sự đã đề xuất ECSAnet, một mô hình cải tiến dựa trên kiến trúc EfficientNetV2, tích hợp CBAM và các lớp FC để nâng cao khả năng học đặc trưng trong phân tích hình ảnh mô bệnh học. Phương pháp của họ áp dụng chuẩn hóa màu Reinhard và kỹ thuật tăng cường dữ liệu nhằm giải quyết vấn đề mất cân bằng tập dữ liệu và cải thiện khả năng tổng quát hóa. Các thí nghiệm trên tập dữ liệu BreakHis cho thấy hiệu suất vượt trội, đạt độ chính xác 94.2%, vượt qua các kiến trúc đã được thiết lập như AlexNet, DenseNet121 và ResNet50. Kết quả của họ nhấn mạnh tác động đáng kể của các thành phần CBAM, các lớp FC và chuẩn hóa màu đối với độ chính xác của mô hình.

Mặc dù AI cho thấy tiềm năng lớn trong việc hỗ trợ phát hiện ung thư vú, nhưng vẫn còn nhiều thách thức cần vượt qua. Thứ nhất, mô hình AI cần lượng dữ

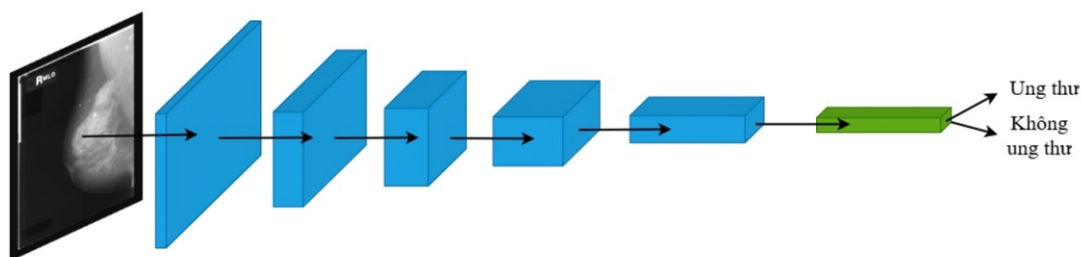
liệu lớn để huấn luyện, trong khi dữ liệu y tế thường gặp vấn đề về bảo mật và quyền riêng tư. Thứ hai, AI không thể thay thế hoàn toàn bác sĩ X-quang mà chỉ đóng vai trò hỗ trợ. Theo nghiên cứu của McKinney và cộng sự, AI có thể giúp giảm tải khối lượng công việc và tăng tốc độ chẩn đoán, nhưng vẫn cần bác sĩ giám sát để đảm bảo độ chính xác[9].

Tuy nhiên, triển vọng của AI trong y học là rất lớn. AI có thể được ứng dụng trong nhiều giai đoạn của quá trình chẩn đoán, từ sàng lọc bệnh nhân có nguy cơ cao đến hỗ trợ bác sĩ đưa ra quyết định điều trị. Với sự phát triển không ngừng của các mô hình học sâu và dữ liệu y tế, AI có thể giúp nâng cao chất lượng chăm sóc sức khỏe, đặc biệt tại các quốc gia đang phát triển, nơi mà nguồn nhân lực y tế còn hạn chế.

2. ỨNG DỤNG MẠNG NƠ-RON TÍCH CHẬP CHO BÀI TOÁN PHÁT HIỆN UNG THƯ VÚ

Mạng nơ-ron tích chập (Convolutional Neural Networks - CNN) là một trong những mô hình học sâu phổ biến và mạnh mẽ nhất trong xử lý dữ liệu hình ảnh và thị giác máy tính. CNN được thiết kế để tự động học và trích xuất các đặc trưng quan trọng từ hình ảnh mà không cần đến quá nhiều tiền xử lý thủ công.

Một mạng nơ-ron tích chập thường bao gồm các lớp chính sau:



Hình 1. Các thành phần của mạng tích chập cho bài toán phát hiện ung thư vú

- Lớp tích chập (Convolutional Layer): Đây là thành phần cốt lõi của CNN, giúp phát hiện các đặc trưng quan trọng trong dữ liệu đầu vào. Lớp này sử dụng các bộ lọc (kernels) để quét qua ảnh, tạo ra các bản đồ đặc trưng (feature maps) phản ánh các mẫu quan trọng như đường biên, góc cạnh hoặc kết cấu.

- Lớp phi tuyến (Activation Layer): Sau khi thực hiện phép tích chập, CNN sử dụng các hàm kích hoạt như ReLU (Rectified Linear Unit) để tăng tính phi tuyến và giúp mạng học được các đặc trưng phức tạp hơn.

- Lớp lấy mẫu (Pooling Layer): Lớp này có nhiệm vụ giảm kích thước của bản đồ đặc trưng, giảm số lượng tham số và tính toán, đồng thời tăng tính bất biến đối với các biến đổi của hình ảnh như tịnh tiến hoặc xoay. Phương pháp phổ biến nhất là Max Pooling, giúp lấy ra giá trị lớn nhất trong mỗi vùng nhỏ của bản đồ đặc trưng.

- Lớp kết nối đầy đủ (Fully Connected Layer - FC): Sau khi trích xuất đặc trưng, các lớp FC giúp đưa ra dự đoán cuối cùng bằng cách kết nối toàn bộ các đặc trưng lại với nhau.

- Lớp đầu ra (Output Layer): Tùy vào bài toán, lớp này có thể sử dụng Softmax (đối với phân loại nhiều lớp) hoặc Sigmoid (đối với bài toán nhị phân) để đưa ra xác suất dự đoán.

Khi áp dụng mạng nơ-ron tích chập để giải quyết bài toán ung thư vú, ta sẽ điều chỉnh lớp đầu ra sao cho còn 1 hoặc 2 đầu ra. Nếu ta chỉ sử dụng 1 đầu ra, thì so sánh giá trị của đầu ra với một ngưỡng để xác định ảnh đầu vào có phải là ảnh ung thư vú hay không. Nếu lớn hơn giá trị ngưỡng thì ảnh đầu vào được xác định là ung thư vú, nếu không thì là không phải. Giá trị ngưỡng

thường được lấy là 0.5. Nếu ta sử dụng 2 đầu ra thì 1 đầu ra sẽ chỉ xác suất ảnh đầu vào là ung thư vú và còn lại là xác suất không ung thư vú. Ta sẽ xác định ảnh đầu vào là ung thư vú hay không dựa vào đầu vào có giá trị xác suất lớn hơn.

3. MÔ HÌNH THỰC NGHIỆM

3.1. Quá trình xử lý đầu vào và Backbone

Hình ảnh đầu vào là ảnh chụp X-quang vú (mammogram) của bệnh nhân. Sau khi được tiền xử lý, hình ảnh này được đưa vào mô hình để trích xuất đặc trưng. Backbone đóng vai trò là bộ trích xuất đặc trưng từ ảnh đầu vào. Nó có thể là một mạng nơ-ron tích chập (CNN) phổ biến như ResNet, EfficientNet hoặc một mô hình phù hợp khác. Backbone sẽ tạo ra một tập hợp các đặc trưng quan trọng, giúp mô hình hiểu rõ hơn về cấu trúc của ảnh và các dấu hiệu có thể liên quan đến ung thư.

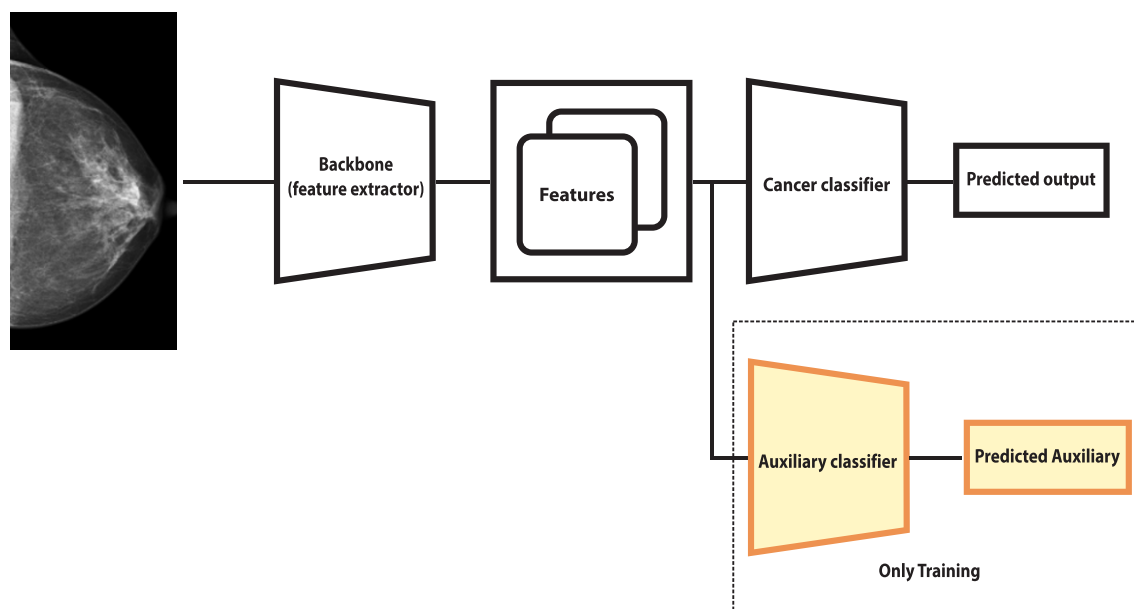
3.2. Phân loại ung thư với Cancer Classifier

Sau khi trích xuất đặc trưng từ backbone, chúng được đưa vào một thành phần gọi là *Cancer Classifier*. Thành phần này có nhiệm vụ thực hiện phân loại chính, đưa ra dự đoán liệu ảnh chụp có dấu hiệu ung thư hay không. Để làm điều này, mô hình sử dụng *Equalization Loss (EQL)*, giúp cân bằng trọng số trong quá trình học và tối ưu hóa việc phân loại giữa các nhóm dữ liệu.

3.3. Auxiliary Classifier - Bộ phân loại phụ trợ

Trong quá trình huấn luyện, một Bộ phân loại phụ trợ được thêm vào với trọng số 0.1. Bộ phân loại này giúp mô hình học tốt hơn bằng cách bổ sung thêm thông tin phụ trợ, cụ thể là:

- Cung cấp thông tin bổ sung về đặc trưng của ảnh: Bộ phân loại phụ trợ có thể



Hình 2. Mô hình thực nghiệm

học và cung cấp thêm thông tin về độ sáng, độ tương phản, kết cấu của tổn thương, giúp mô hình có góc nhìn đa chiều hơn.

- Giảm nhiễu và tăng độ ổn định: Nó hoạt động như một bộ lọc, giúp mô hình tập trung vào các đặc trưng quan trọng hơn thay vì bị ảnh hưởng bởi các yếu tố không liên quan.

- Cải thiện khả năng học đặc trưng: Các thông tin từ Bộ phân loại phụ trợ giúp mô hình nhận diện và học được các mẫu dữ liệu phức tạp hơn.

- Tăng độ chính xác: Việc bổ sung một bộ phân loại phụ trợ giúp mô hình phân biệt rõ hơn giữa các tổn thương lành tính và ác tính, đặc biệt là các trường hợp khó chẩn đoán.

Cách hoạt động của mô hình trong quá trình huấn luyện và dự đoán

- Trong quá trình huấn luyện, Auxiliary Classifier giúp cải thiện chất lượng mô hình bằng cách đưa ra các dự đoán bổ sung.

- Khi mô hình đã được huấn luyện hoàn chỉnh, Auxiliary Classifier chỉ được sử dụng trong giai đoạn huấn luyện và không tham gia vào quá trình dự đoán thực tế.

- Cancer Classifier sẽ đưa ra dự đoán cuối cùng về tình trạng ung thư dựa trên những đặc trưng đã học được từ backbone và quá trình tối ưu hóa với EQL Loss.

3.4. Các mô hình thực nghiệm

Để giải quyết bài toán phát hiện ung thư vú, chúng tôi đã nghiên cứu và tìm hiểu về các mô hình học máy và nhận thấy một số mô hình phù hợp cho vai trò backbone để giải quyết bài toán phát hiện ung thư vú. Các mô hình phù hợp được lựa chọn để thử nghiệm bao gồm:

MobileNetV3Large/Small [2]: MobileNetV3 bao gồm hai biến thể, Lớn và Nhỏ, với sự đánh đổi khác nhau giữa độ chính xác và độ trễ. Nó kết hợp nhiều cải tiến kiến trúc khác nhau, bao gồm cải thiện tính phi tuyến tính, khối bộp và kích thích và kích hoạt h-swish. Các mẫu MobileNetV3 được

thiết kế để đạt hiệu suất cao nhất trên các thiết bị di động và thiết bị nhúng.

ResNet50 [5]: ResNet50 là một phần của dòng mô hình ResNet (Mạng dư) đã giới thiệu các kết nối dư để giảm bớt vấn đề độ dốc biến mất. ResNet50 có 50 lớp và bao gồm các khối còn lại với các kết nối bỏ qua. Các mô hình ResNet được biết đến với kiến trúc sâu và đã được áp dụng rộng rãi trong các ứng dụng thị giác máy tính khác nhau.

ConvNeXtV2-nano [13]: Đây là phiên bản nhẹ nhất trong họ ConvNeXtV2, được thiết kế để cân bằng giữa hiệu suất và kích thước mô hình. Mô hình vẫn giữ nguyên kiến trúc cốt lõi của ConvNeXt nhưng với số lượng tham số ít hơn, khoảng 15,6 triệu. ConvNeXtV2 sử dụng kiến trúc hoàn toàn dựa trên convolution (fully convolutional architecture) với các khối fully convolutional block, giúp tăng tốc độ xử lý so với các phiên bản lớn hơn. Mô hình cũng tích hợp Global Response Normalization (GRN) để cải thiện khả năng học các mẫu toàn cục (global patterns). Ngoài ra, ConvNeXtV2 còn được tối ưu hóa với các kết nối tắt (skip connections) và cải tiến kiến trúc nhằm giảm độ phức tạp tính toán.

ResNeXt50 [11]: ResNeXt50 là một

biến thể của ResNet50, kết hợp ý tưởng từ mạng Inception. Mô hình này sử dụng “cardinality” (số lượng nhóm song song) thay vì chỉ đơn thuần tăng độ rộng hoặc độ sâu của mạng. Kiến trúc gồm 50 tầng với các khối đặc trưng ResNeXt. Tổng số tham số khoảng 25 triệu (25M). Thông thường, cardinality được đặt là 32. Kích thước đầu vào tiêu chuẩn là 224×224 pixel, với độ rộng bottleneck là 4.

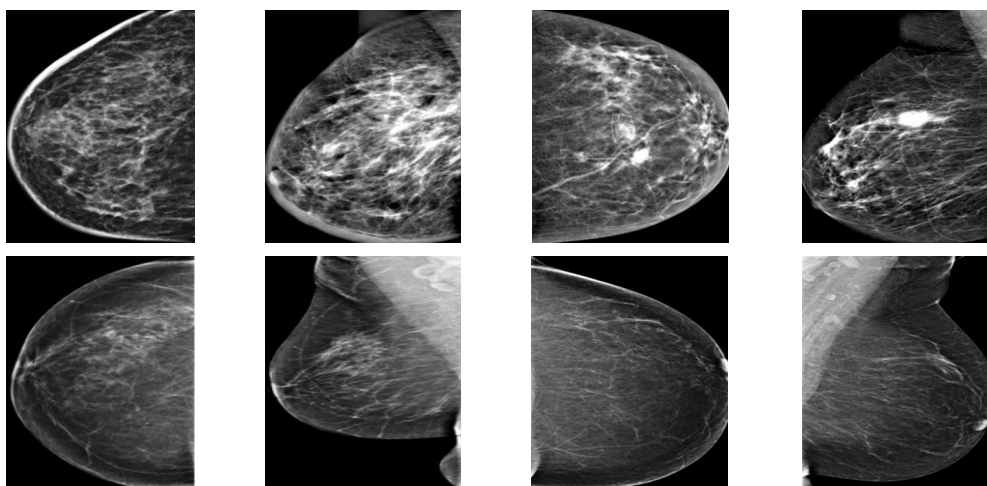
MobileViT-small [10]: Kết hợp điểm mạnh của CNN và Transformer, MobileViT sử dụng các khối MobileViT đặc trưng trong một kiến trúc nhẹ, phù hợp với các thiết bị di động và edge devices. Kích thước đầu vào tiêu chuẩn là 256×256 pixel. Thay vì chia ảnh thành các phần (patches) rời rạc, mô hình áp dụng phương pháp cửa sổ trượt. MobileViT tích hợp xử lý cục bộ (CNN) và cơ chế toàn cục (Transformer), giúp cải thiện hiệu quả trong việc xử lý thông tin không gian. Mô hình có số lượng tham số ít hơn đáng kể so với các kiến trúc tương tự. Ngoài ra, MobileViT tối ưu hóa bộ nhớ thông qua kỹ thuật đặt kiểm tra gradient và cải tiến các lớp Mobile Convolution nhằm nâng cao hiệu suất hoạt động.

Bảng 1. Thông số của các tập dữ liệu

Tập dữ liệu			Số bệnh nhân	Số mẫu	Số mẫu dương tính	Phần trăm mẫu dương tính
BMCD			82	328	22	6,71%
CDD-CESM	326	1.003	331	33,00%		
CMMD	1.775	5.202	2.632	50,60%		
MiniDDSM	1.952	7.808	1.480	18,95%		

Bảng 2. Bảng kết quả thực nghiệm với nhiều mô hình khác nhau

Mô hình	Độ chính xác (Acc.)	Điểm F1 (F1 score)	Độ tin cậy (Precision)	Độ nhạy (Recall)	Điểm F1-xác suất (pF1 Score)	Điểm F1-beta (F1- β Score)
mobilenet_v3-large	84,53%	84,28%	78,21%	70,05%	67,12%	73,91%
convnextv2-nano	81,15%	80,43%	75,57%	58,73%	57,66%	66,10%
resnet50	83,58%	83,43%	75,36%	70,59%	66,16%	72,89%
resnext50	84,33%	83,93%	79,72%	66,93%	67,84%	72,77%
mobilevit-small	83,75%	83,32%	78,64%	65,95%	65,11%	71,74%



Hình 3. Minh họa kết quả dự đoán sai của Mô hình ResNet50. Những ảnh ở hàng trên là những ảnh nhân thực tế là không ung thư nhưng mô hình lại dự đoán là ung thư. Những ảnh ở hàng dưới là những ảnh có nhãn là ung thư nhưng mô hình lại dự đoán là không ung thư

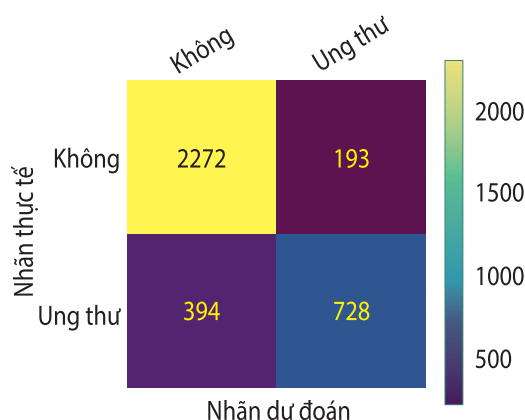
4. KẾT QUẢ THỰC NGHIỆM

4.1. Cơ sở dữ liệu thực nghiệm

Trong phần thực nghiệm này, chúng tôi sử dụng 4 tập dữ liệu sau đây: BMCD, CDD-CESM, CMMD, MiniDDSM để tiến hành huấn luyện và đánh giá hiệu năng mô hình. Thông tin của 4 bộ dữ liệu được trình bày ở Bảng 1. Chúng tôi đã tổng hợp 4 tập dữ liệu này thành 1 tập dữ liệu tổng hợp. Sau đó tiến hành phân chia tập dữ liệu huấn luyện và tập dữ liệu đánh giá theo tỉ lệ 80% và 20%. Chúng tôi làm điều này để đánh giá mức độ hiệu quả của các mô hình được thực nghiệm khi được huấn luyện và kiểm thử với các kiểu ảnh X-quang tuyến vú khác nhau từ nhiều loại cơ sở dữ liệu. Điều này gần với thực tế khi ảnh chụp X-quang từ các bệnh viện có thể khác nhau.

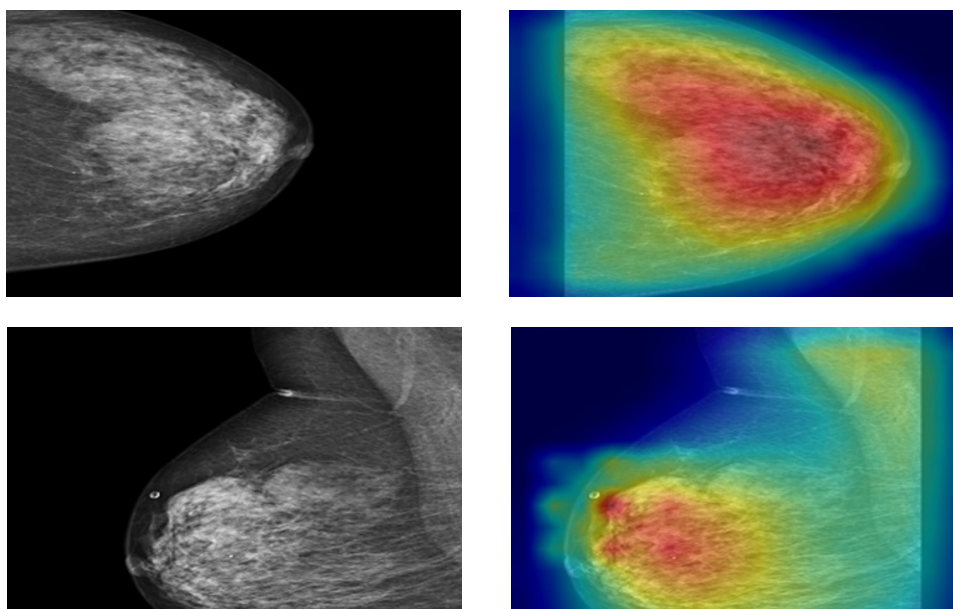
4.2. Kết quả thực nghiệm

Chúng tôi đã thực hiện chạy dữ liệu đã phân bổ với các mô hình được lựa chọn đó là MobilenetV3 Large, ConvnextV2 Nano, Resnet50, Resnext50, MobileViT Small.



Hình 4. Ma trận nhầm lẫn của mô hình ResNet50

Sau quá trình huấn luyện, kết quả thu được như được thể hiện trong Bảng 2. Kết quả trên đánh giá năm mô hình này dựa trên các chỉ số Độ chính xác (Accuracy), Điểm F1 (F1 Score), Độ tin cậy (Precision), Độ nhạy (Recall), Điểm F1-xác suất (pF1 Score), và điểm F1-beta (F1 β Score). Mỗi chỉ số phản ánh một khía cạnh quan trọng của hiệu suất mô hình trong bài toán phân loại ung thư vú từ ảnh chụp X-quang.



Hình 5. Minh họa hình ảnh đầu vào và GradCAM tương ứng cho các kết quả phát hiện ung thư vú của mô hình ResNet50

Về Độ chính xác, MobileNetV3-Large (84,53%) đạt cao nhất, theo sau là ResNeXt50 (84,33%) và MobileViT-Small (83,75%), cho thấy hiệu suất phân loại tổng thể tốt. Ngược lại, ConvNeXtV2-Nano có Accuracy thấp nhất (81,15%), làm giảm tính tin cậy của mô hình trong thực tế.

Độ tin cậy cao giúp giảm dương tính giả, trong khi Recall cao đảm bảo không bỏ sót ca bệnh. ResNeXt50 có Precision cao nhất (79,72%), giúp giảm cảnh báo sai. Trong khi đó, ResNet50 có Recall cao nhất (70,59%), cho thấy khả năng phát hiện bệnh tốt hơn. ConvNeXtV2-Nano có Recall thấp nhất (58,73%), đồng nghĩa với việc có thể bỏ sót nhiều ca bệnh nguy hiểm.

MobileNetV3-Large đạt điểm F1 cao nhất (0,8428), theo sau là ResNeXt50 (0,8393) và ResNet50 (0,8343). Điều này cho thấy ba mô hình này có sự cân bằng tốt giữa Độ tin cậy và Độ nhạy, phù hợp với bài toán phân loại ung thư. Ngược lại, ConvNeXtV2-Nano có điểm F1 thấp nhất (0,8043), xác nhận hiệu suất kém hơn.

ResNeXt50 có điểm F1-xác suất cao nhất (0,6784), tiếp theo là MobileNetV3-Large (0,6712), cho thấy độ chính xác cao khi phát hiện ca bệnh dương tính. MobileNetV3-Large có F1 β Score cao nhất (0,7391), theo sau là ResNet50 (0,7289), phản ánh khả năng phát hiện ung thư tốt mà không quá ưu tiên Độ tin cậy. ConvNeXtV2-Nano có chỉ số thấp nhất ở cả hai tiêu chí, tiếp tục khẳng định hiệu suất kém nhất trong nhóm.

Như vậy, MobileNetV3-Large và ResNeXt50 có hiệu suất tổng thể tốt nhất, cân bằng giữa độ chính xác và khả năng phát hiện bệnh. ResNet50 có Độ nhạy cao, phù hợp nếu ưu tiên phát hiện bệnh sớm. MobileViT-Small đạt hiệu suất khá nhưng

không vượt trội. ConvNeXtV2-Nano có hiệu suất kém nhất và không phù hợp cho bài toán này. Tóm lại, nếu cần mô hình chính xác và đáng tin cậy, MobileNetV3-Large hoặc ResNeXt50 là lựa chọn tốt nhất. Nếu ưu tiên Độ nhạy, ResNet50 sẽ phù hợp hơn.

Hình 3 minh họa kết quả dự đoán sai của Mô hình ResNet50. Ta có thể thấy rằng, chủ yếu những trường hợp đánh giá sai của mô hình là do ảnh đầu vào do được tập hợp từ nhiều cơ sở dữ liệu khác nhau, dẫn đến dễ khiến mô hình nhầm lẫn.

Hình 4 thể hiện ma trận nhầm lẫn của mô hình ResNet50. Ta có thể dễ dàng thấy rằng phần lớn những dự đoán nhầm lẫn của mô hình là ở những ảnh có nhân thực tế là ung thư nhưng mô hình lại dự đoán là không ung thư.

Hình 5 minh họa hình ảnh đầu vào và GradCAM tương ứng cho kết quả phát hiện ung thư vú của mô hình ResNet50. Trong đó, vùng màu đỏ thể hiện vùng chú ý cao của mô hình khi thực hiện phân loại đối với ảnh đầu vào. Mức độ chú ý sẽ giảm dần đối với vùng màu vàng. Và vùng màu xanh là vùng ít được chú ý nhất. Ta có thể dễ dàng thấy rằng mô hình ResNet50 đã chú ý vào vùng có xuất hiện khối u trong ảnh đầu vào.

5. KẾT LUẬN

Bài báo này đã trình bày việc ứng dụng kỹ thuật học sâu trong bài toán phát hiện ung thư vú dựa trên hình ảnh chụp X-quang tuyến vú. Thông qua việc thử nghiệm nhiều mô hình khác nhau, kết quả cho thấy các mô hình như MobileNetV3-Large và ResNeXt50 đạt hiệu suất cao nhất, với sự cân bằng giữa độ chính xác, khả năng phát hiện bệnh (Recall) và độ tin cậy của dự đoán (Precision). Trong khi đó, ResNet50 cũng là một lựa chọn phù hợp nếu ưu tiên khả năng

phát hiện sớm các ca bệnh dương tính.

Mặc dù các phương pháp học sâu mang lại tiềm năng lớn trong việc cải thiện độ chính xác và tốc độ chẩn đoán ung thư vú, vẫn còn một số thách thức cần giải quyết. Đầu tiên, chất lượng và số lượng dữ liệu huấn luyện đóng vai trò quan trọng, do đó việc mở rộng và chuẩn hóa dữ liệu là cần thiết để nâng cao hiệu suất mô hình. Thứ hai, mô hình AI không thể thay thế hoàn toàn bác sĩ, mà chỉ đóng vai trò hỗ trợ trong quá trình chẩn đoán. Việc kết hợp AI với chuyên môn của bác sĩ có thể giúp nâng cao

hiệu quả chẩn đoán và tối ưu hóa quy trình điều trị.

Trong tương lai, nghiên cứu có thể mở rộng theo hướng tối ưu hóa mô hình để phù hợp hơn với các thiết bị di động và hệ thống lâm sàng thực tế, giúp ứng dụng AI vào chẩn đoán ung thư vú một cách rộng rãi và hiệu quả hơn. Đồng thời, việc tích hợp các phương pháp xử lý ảnh tiên tiến và học tăng cường có thể giúp cải thiện khả năng nhận diện các tổn thương ở giai đoạn sớm, góp phần giảm tỷ lệ tử vong do ung thư vú.

TÀI LIỆU THAM KHẢO

- [1] Aldakhil, L.A., Alhasson, H.F., Alharbi, S.S (2024), “Attention-based deep learning approach for breast cancer histopathological image multi-classification”. *Diagnostics*.
- [2] Andrew Howard, Mark Sandler. (2019), “Searching for MobileNetV3”, *ICCV* 2019.
- [3] Aneja, S., Aneja, và cộng sự (2022), “Transfer learning for cancer diagnosis in histopathological images”. *IAES International Journal of Artificial Intelligence* 11, 129–136.
- [4] Geras K.J., Wolfson S., Shen Y. (2018), “High-Resolution Breast Cancer Screening with Multi-View Deep Convolutional Neural Networks”, *arXiv:170307047*
- [5] Kaiming He, Xiangyu Zhang và cộng sự. (2019), “Deep Residual Learning for Image Recognition”, *ICCV* 2019
- [6] Kim, J., Harper, A., McCormack, V. et al. Global patterns and trends in breast cancer incidence and mortality across 185 countries. *Nat Med* (2025).
- [7] Lévy D. và Jain A. (2016), “Breast Mass Classification from Mammograms using Deep Convolutional Neural Networks”, *arXiv:161200542*.
- [8] M. Schneider, M. Yaffe (2000), “Better detection: improving our chances, Digital Mammography”, 5th International Workshop on Digital Mammography.
- [9] McKinney S.M., Sieniek M., Godbole V. và cộng sự. (2020) “International evaluation of an AI system for breast cancer screening”, *Nature*, 577(7788), Trang 89–94.
- [10] Moreira I.C., Amaral I., Domingues I. (2012), “INbreast: toward a full-field digital mammographic database”, *Academic Radiology*, 19(2), Trang 236–248.
- [11] Sachin Mehta, Mohammad Rastegari. (2021), “MobileViT: Light-weight, General-purpose, and Mobile-friendly Vision Transformer”, *ICLR'22*
- [12] Saining Xie, Ross Girshick, và cộng sự. (2017), “Aggregated Residual Transformations for Deep Neural Networks”, *CVPR* 2017
- [13] Sanghyun Woo, Shoubhik Debnath, và

- cộng sự, (2023), “ConvNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders”, arXiv:2301.00808
- [14] Tuan Linh Le, My Hanh Bui, và cộng sự (2024), “Transfer learning for deep neural networks-based classification of breast cancer x-ray images”. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization 12.
- [15] Xi P., Shu C., và Goubran R. (2018), “Abnormality Detection in Mammography using Deep Convolutional Neural Networks”, arXiv:1803.01906.

Lời cảm ơn: Nghiên cứu này được tài trợ bởi đề tài cấp cơ sở Trường Đại học Quảng Bình, mã số đề tài: CS.13.2025

Liên hệ:

TS. Hoàng Văn Thành

Khoa Công nghệ thông tin, Trường Đại học Quảng Bình

Địa chỉ: 18 Nguyễn Văn Linh, Đồng Hới, Quảng Bình

Email: thanhhvnbq@gmail.com

Ngày nhận bài: 12/02/2025

Ngày gửi phản biện: 12/02/2025

Ngày duyệt đăng: 26/02/2025