

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP.HCM

PHÁT TRIỂN HỆ THỐNG HỖ TRỢ CHẨN ĐOÁN BỆNH UNG THƯ
VỤ ỨNG DỤNG TRÍ TUỆ NHÂN TẠO

Chuyên ngành: Kỹ thuật Điện tử
Mã số chuyên ngành: 9520203

TÓM TẮT LUẬN ÁN TIẾN SĨ

TP. HỒ CHÍ MINH – NĂM 2025

Công trình được hoàn thành tại **Trường Đại học Sư phạm Kỹ thuật Tp.HCM**

Người hướng dẫn khoa học 1: PGS.TS. -----

Phản biện 1:

Phản biện 2:

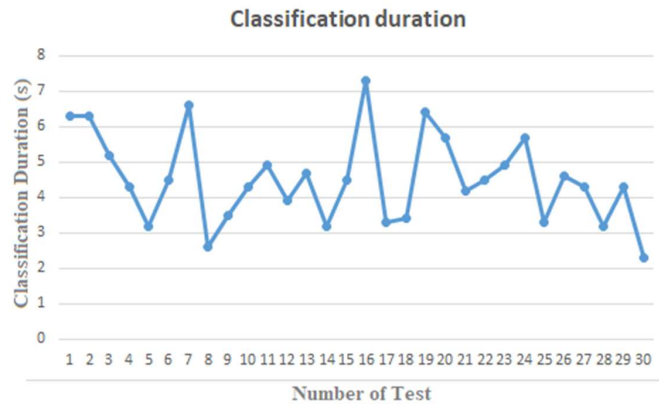
Phản biện 3:

Luận án đã được bảo vệ trước Hội đồng đánh giá luận án Cấp Cơ sở học tại
Trường Đại học Sư phạm Kỹ thuật Tp.HCM
vào ngày 01 tháng 11 năm 2025

DANH MỤC CÔNG TRÌNH ĐÃ CÔNG BỐ

1. T.-T. Nguyen, T.-H. Nguyen, B.-V. Ngo, and T.-N. Nguyen, “Largest roi segmentation for breast cancer classification using a VGG16 deep learning network,” Advances in Electrical and Electronic Engineering, 2024, DOI: 10.15598/aeec.v22i4.240303.
2. Thanh-Tam Nguyen, Thanh-Hai Nguyen, “Segmentation and Color ROI Extraction from Breast Imaging Datasets for Cancer Classification”, Engineering, Technology and Applied Science Research, 2025, accepted paper.
3. Thanh-Nghia Nguyen, Thanh-Tam Nguyen, Thanh-Hai Nguyen, B.V. Ngo, “A Robust Approach for Breast Cancer Classification from DICOM Images”, Engineering, Technology and Applied Science Research, 04 Jun 2025, Vol. 15, Issue 3, pages 23499 – 23505, doi:10.48084/etasr.10931.
4. Thanh-Tam Nguyen, Nguyen Thanh Hai, Tin-Trung Nguyen, “Design of A Telemedicine System for Classification of Breast Cancer Images”, Journal of Technical Education Science, 2025, accepted paper.
5. T.-T. Nguyen, T.-H. Nguyen, B.-V. Ngo, and D.-D. Vo, “Breast Image Segmentation for evaluation of Cancer Disease,” in 2020 5th International Conference on Green Technology and Sustainable Development (GTSD), 2020, pp. 344–348, doi: 10.1109/GTSD50082.2020.9303133.
6. T.-T. Nguyen, T.-H. Nguyen, and B.-V. Ngo, “A GLCM Algorithm for Optimal Features of Mammographic Images for Detection of Breast Cancer,” in 2021 International Conference on System Science and Engineering (ICSSE), 2021, pp. 295–299, doi: 10.1109/ICSSE52999.2021.9538426.

Hình 7 thể hiện kết quả thử nghiệm hệ thống y tế từ xa. Cụ thể, biểu đồ cho thấy tốc độ đọc/ghi của hệ thống luôn vượt ngưỡng cho phép là 200Mbps và tốc độ trung bình là 284Mbps.



Hình 8. Kết quả đo thời gian phân loại.

Hình 8 cho thấy thời gian phân loại trung bình của hệ thống là 4,5 giây (nhạy nhất 2,3 giây, chậm nhất 7,3 giây) sau 30 lần thử. Hệ thống EfficientNet-B7 huấn luyện trên dữ liệu DICOM từ Bệnh viện Ung bướu TP.HCM đạt độ chính xác 97% trên tập huấn luyện và 89,58% trên tập thử nghiệm, đáp ứng yêu cầu triển khai thực tế với tốc độ truyền 284 Mbps.

CHƯƠNG 6 KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Luận án tập trung vào các thuật toán tăng cường và tiền xử lý ảnh X-quang vú nhằm nâng cao chất lượng hình ảnh, từ đó cải thiện độ chính xác phân loại giữa vùng tổn thương và không tổn thương. Dữ liệu nghiên cứu gồm hơn 4.000 ảnh từ Bệnh viện Ung Bướu TP.HCM, đã được xử lý và gán nhãn theo chuẩn. Các phương pháp trích xuất đặc trưng, phân đoạn ROI và tô màu ảnh (CROI) giúp cải thiện 6–8% độ chính xác, đạt gần 95% trong phân loại tổn thương và 92,6% trong phân loại nguy cơ ung thư theo BIRADS. Luận án cũng xây dựng các mô hình học sâu kết hợp CNN tiên tiến, đồng thời phát triển hệ thống hỗ trợ chẩn đoán từ xa cho bác sĩ. Kết quả thử nghiệm trên nhiều tập dữ liệu cho thấy hiệu quả rõ rệt, được công bố trên các tạp chí khoa học (Q2, Q3), góp phần vào phát hiện sớm và hỗ trợ chẩn đoán ung thư vú trong y tế từ xa.

CHƯƠNG 1 GIỚI THIỆU TỔNG QUAN

1.1 TÍNH CẤP THIẾT CỦA ĐỀ TÀI

Ung thư vú là nguyên nhân tử vong hàng đầu ở phụ nữ, với số ca mắc ngày càng tăng, kể cả ở người trẻ tại Việt Nam. Mỗi năm có khoảng 22.000 ca mới và hơn 9.000 ca tử vong, trong đó gần 70% phát hiện ở giai đoạn muộn, chủ yếu do thiếu hệ thống sàng lọc và nguồn lực y tế hạn chế. Bệnh viện Bạch Mai tiếp nhận hơn 5.000 bệnh nhân/ngày, tạo áp lực lớn cho bác sĩ, dễ dẫn đến chẩn đoán sai.

Giải pháp là tích hợp AI vào hệ thống chẩn đoán từ xa, giúp nâng độ chính xác lên tới 94,5% (so với 88% truyền thống) và giảm khoảng 40% số lần khám trực tiếp. Tại Việt Nam, nơi tỷ lệ phát hiện muộn vẫn rất cao, công nghệ này hứa hẹn cải thiện khả năng tiếp cận và chất lượng chăm sóc sức khỏe.

1.2 MỤC TIÊU NGHIÊN CỨU CỦA LUẬN ÁN

1.2.1 Mục tiêu nghiên cứu

Phát triển một hệ thống chẩn đoán bệnh UTV dựa trên trí tuệ nhân tạo được huấn luyện thông qua các tập ảnh X-quang lớn khác nhau đã được xử lý để chuẩn hóa để hỗ trợ bác sĩ trong chẩn đoán bệnh UTV.

1.2.2 Mục tiêu cụ thể

Xuất phát từ mục tiêu nghiên cứu, đề tài sẽ được phân thành 3 mục tiêu cụ thể để tiến tới mục tiêu đã đặt ra:

Mục tiêu 1: Thu thập nhiều tập dữ liệu lớn hình ảnh X-quang vú khác nhau (từ khoảng năm ngàn đến gần chục ngàn ảnh) đã được gán nhãn. Trên các tập ảnh đó, phát triển các thuật toán xử lý ảnh trước khi trích xuất các đặc trưng ảnh cho dò tìm bất thường của ảnh X-quang vú.

Mục tiêu 2: Xây dựng một thuật toán sử dụng trí tuệ nhân tạo để có thể dự đoán tình trạng bệnh UTV của bệnh nhân (có bệnh hay không có bệnh) dựa trên các tập dữ liệu ảnh đầu vào thu được.

Mục tiêu 3: Triển khai một mô hình hệ thống chẩn đoán từ xa dùng AI cho phép bác sĩ có thể tải lên các hình ảnh X-quang vú cho dự đoán tình trạng bệnh UTV từ xa thông qua điện thoại hoặc PC, từ đó cho hỗ trợ bác sĩ đưa ra kết quả chẩn đoán. Thông qua hệ thống này chúng ta có thể thu thập dữ liệu và kết quả về bệnh UTV của bệnh nhân tại Việt Nam.

CHƯƠNG 2 CƠ SỞ LÝ THUYẾT

2.1 CÁC TẬP DỮ LIỆU HÌNH ẢNH VỀ BỆNH UNG THU VÚ

MIAS [63]: Hơn 300 ảnh X-quang tuyến vú có chú thích về loại mô nền, bất thường và mức độ nghiêm trọng, kèm tọa độ X, Y của vùng tổn thương [64].

DDSM [65]: Gồm hơn 2.600 nghiên cứu nhũ ảnh quét từ phim. Tập con CBIS-DDSM [66] được chú thích chi tiết về loại, cấp và giai đoạn khối u.

INBreast [67]: Cung cấp hơn 410 ảnh nhũ ảnh kỹ thuật số chất lượng cao, hỗ trợ nghiên cứu UTV.

BCDR [68]: Tập trung vào nhũ ảnh kỹ thuật số toàn trường (FFDM), khuyến khích đóng góp từ cộng đồng nghiên cứu [69].

BancoWeb LAPIMO [70]: Gồm hơn 1.400 ảnh TIFF từ 320 bệnh nhân, thuộc dự án BancoWeb LAPIMO.

VICTRE [71]: Bộ dữ liệu mô phỏng với 2.986 đối tượng đại diện cho nhiều kích thước và mật độ vú.

OPTIMAM (OMI-DB) [71]: CSDL nhũ ảnh được chú thích đầy đủ, tập trung cho mục đích nghiên cứu.

2.2 PHÂN ĐOẠN ẢNH VÀ XỬ LÝ ẢNH TRONG ẢNH UTV

Phân đoạn ảnh

Ảnh X-quang vú thường không đồng đều và chứa nhiều nhiễu, nên các nghiên cứu thường áp dụng thuật toán tiền xử lý để cải thiện chất lượng và tăng hiệu quả phân loại với các mạng học sâu như VGG16, EfficientNet. Ở giai đoạn này, các bước phổ biến gồm lọc nhiễu, loại bỏ thành phần không quan trọng (cơ ngực, nhân...) và tách nền bằng phương pháp như ngưỡng Otsu. Sau khi tiền xử lý, ảnh được phân đoạn để xác định vùng nghi ngờ (ROI) bằng nhiều kỹ thuật: dựa trên ngưỡng, vùng, đặc trưng, cạnh, hoặc học sâu. Các phương pháp này giúp loại bỏ nền, giữ lại khối u và nâng cao độ nhạy trong phát hiện. Trong đó, ngưỡng Otsu là kỹ thuật phổ biến trong ảnh y tế, cho kết quả phân đoạn lên đến 77,43%, và được áp dụng trong nghiên cứu này để trích xuất ROI từ ảnh nhũ ảnh có bốn loại tổn thương.

Cải thiện chất lượng hình ảnh trong chẩn đoán UTV

Độ chính xác chẩn đoán phụ thuộc nhiều vào chất lượng nhũ ảnh. Hình ảnh rõ nét giúp AI phát hiện các tổn thương vi thể hiệu quả hơn. Để cải thiện chất lượng và độ tương phản, nhiều kỹ thuật tăng cường ảnh đã được nghiên cứu, gồm: phương pháp truyền thống, tăng cường theo vùng, dựa trên đặc trưng (như biến

Bảng 4. Biểu diễn kết quả phân loại bằng cách sử dụng ma trận nhầm lẫn.

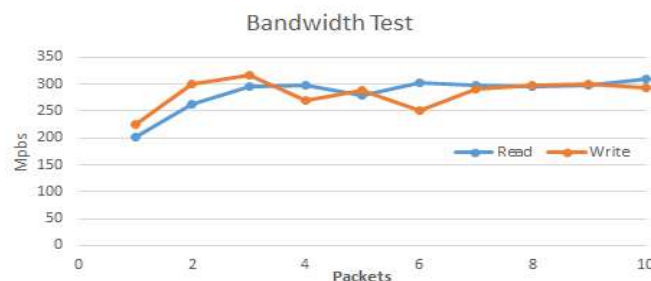
Categories	Number of Images	True Positive	False Positive	
Normal	160	144	7 to Benign	9 to Malignant
Benign	160	141	8 to Normal	11 to Malignant
Malignant	160	145	9 to Normal	6 to Benign

Phân tích kết quả huấn luyện cho thấy mô hình AI với EfficientNet-B7 đạt độ chính xác huấn luyện cao, đặc biệt hội tụ ở mức khoảng 97%, cho thấy khả năng học tốt trên dữ liệu huấn luyện. Tuy nhiên, độ chính xác trên tập kiểm tra chỉ đạt khoảng 89,58% và đây có thể là một vấn đề nhỏ liên quan đến việc xây dựng dữ liệu mới. Sự khác biệt giữa độ chính xác huấn luyện và kiểm tra có thể là do tập dữ liệu được sử dụng là dữ liệu thực, có sự khác biệt lớn về chất lượng và được thu thập từ các thiết bị khác nhau. Độ chính xác và điểm F1 được thể hiện trong Bảng 5.

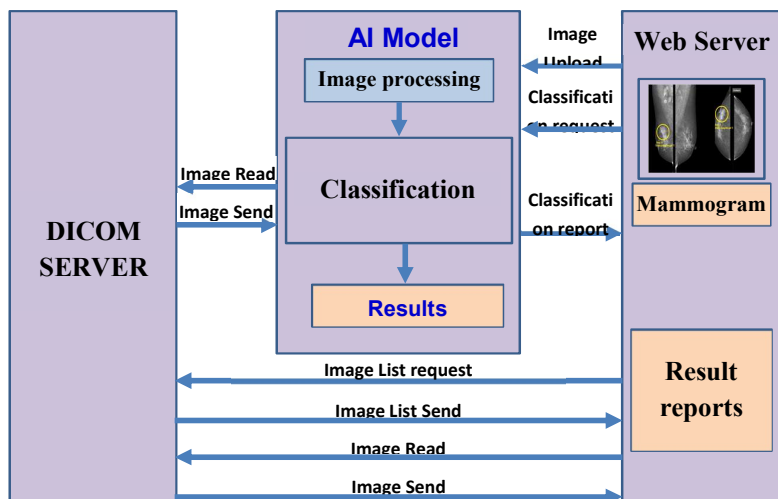
Table 5. Representation of the precision and F1-score.

Categories	Precision	recall	f1-score
Normal	0.8944	0.9000	0.8972
Benign	0.9156	0.8812	0.8981
Malignant	0.8788	0.9062	0.8923

Kết quả của mô hình Telemedicine



Hình 7. Kết quả đo băng thông hệ thống.

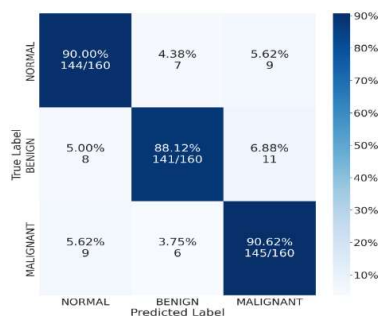


Hình 3. Mô tả các giao thức nội bộ để trao đổi thông tin giữa ba thành phần chính trong lõi hệ thống.

Kết quả

Kết quả phân loại ung thư vú

Độ chính xác phân loại được đánh giá thông qua ma trận nhầm lẫn, trong đó các tập ảnh được thu thập từ Bệnh viện Ung bướu Thành phố Hồ Chí Minh và tập dữ liệu kiểm tra hoàn toàn độc lập với tập dữ liệu huấn luyện, với tỷ lệ số lượng là 8:2. Về cơ bản, có 3 loại tập ảnh là Lành tính, Ác tính và Bình thường. Do đó, kết quả huấn luyện của mô hình EfficientNet-B7 được đánh giá bằng các hàm Mật mát và Độ chính xác.



Hình 6. Biểu diễn ma trận nhầm lẫn cho ba loại tổn thương như lành tính, ác tính và bình thường.

đổi wavelet), và tăng cường mờ (fuzzy entropy). Các kỹ thuật này hỗ trợ làm rõ vi vôi hóa, khối u, đường viền và chi tiết ảnh, song có thể làm tăng nhiễu.

Các nghiên cứu gần đây chứng minh hiệu quả: Shenbagavalli đạt độ chính xác 93,45% với biến đổi shearlet; Teare sử dụng CLAHE kết hợp CNN và rừng ngẫu nhiên, cho độ nhạy 0,91 và độ đặc hiệu 0,80. Ngoài ra, kỹ thuật FADHECAL giúp giảm nhiễu mà vẫn giữ chi tiết hình ảnh tốt hơn các phương pháp khác. Gần đây, các phương pháp tăng cường dựa trên tập mờ nâng cao và thuật toán tối ưu hóa cũng cho kết quả khả quan. Nhìn chung, việc cải thiện chất lượng ảnh là bước quan trọng để hỗ trợ hệ thống AI và bác sĩ X-quang chẩn đoán chính xác hơn.

2.3 PHÂN LOẠI ẢNH UTV SỬ DỤNG MÁY HỌC

Phát hiện và phân loại khối u vú

Khối u là triệu chứng phổ biến nhất của ung thư vú, song việc phát hiện trên nhũ ảnh, đặc biệt ở mô vú dày đặc, khá khó khăn. Nhiều nghiên cứu đã ứng dụng AI để nâng cao độ chính xác. Các phương pháp tiêu biểu gồm: phân cụm mờ trực quan (CrSA-IFCM-NA), CNN nhiều tầng, YOLO, Fast-RCNN, U-Net, Mask-RCNN, và các mô hình lai kết hợp học sâu với ML.

Kết quả báo cáo rất khả quan: độ chính xác phát hiện có thể đạt 98,96% với YOLO và FrCN trên INBreast, độ nhạy 100% và độ đặc hiệu 94% với YOLO, hoặc AUC 0,95 với Fast-RCNN. Ngoài ra, nhiều nghiên cứu cũng tập trung vào phân loại lành tính/ác tính bằng CNN, học sâu lai, và deep fusion learning, với độ chính xác dao động từ 89% đến trên 99%. Các phương pháp mới còn chú trọng khả năng giải thích kết quả, như kết hợp U-Net và Case-Based Reasoning.

Đánh giá nguy cơ UTV

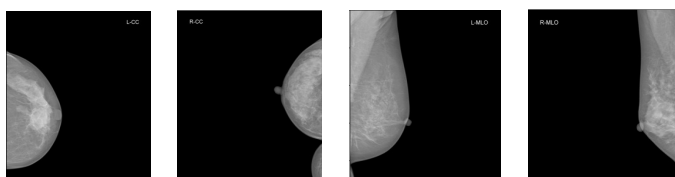
Một số nghiên cứu tập trung vào phân loại mật độ mô vú (theo BIRADS) bằng CNN, residual CNN hoặc học không giám sát với autoencoder. Ngoài ra, các mô hình học sâu còn được dùng để phân loại khối u lành/ác tính, với các chiến lược như học chuyển giao, tăng cường dữ liệu, hay mô hình lai kết hợp DL và ML, cho độ chính xác cao. Vấn đề dữ liệu hạn chế và mất cân bằng được khắc phục bằng cách tạo dữ liệu tổng hợp (cGAN) hoặc học đa nhiệm. Nghiên cứu của Google (2020, Nature) cho thấy AI có thể chẩn đoán UTV từ nhũ ảnh với độ chính xác ngang hoặc vượt chuyên gia, đồng thời giảm dương tính và âm tính sai.

Ngoài hình ảnh, nhiều mô hình ML (SVM, ANN, CBR...) cũng được ứng dụng để dự đoán nguy cơ dựa trên yếu tố lâm sàng và dịch tễ. Đặc biệt, các nghiên cứu sử dụng nhũ ảnh thường quy để ước lượng nguy cơ trong nhóm sàng lọc cộng đồng cho thấy AI dựa trên FFDM dự đoán tốt hơn mật độ vú BIRADS và các mô hình dịch tễ học truyền thống (AUC 0,60–0,84).

CHƯƠNG 3 THU THẬP VÀ TIỀN XỬ LÝ ẢNH X-QUANG VÚ

3.1 BỘ DỮ LIỆU VINDR-MAMMO

Bộ dữ liệu VinDr-Mammo [135]: trong đó các ảnh chụp nhũ ảnh được thu thập từ năm 2018 đến năm 2020. Đặc biệt, các bộ ảnh này được lưu trữ tại Hệ thống PACS của Bệnh viện Đại học Y Hà Nội (ĐHYHN) và Bệnh viện 108 (H108). Tất cả 20.000 ảnh chụp nhũ ảnh trong bộ dữ liệu VinDr-Mammo được phân loại thành năm nhóm ảnh chụp nhũ ảnh và 11 nhiễu loạn vú (khối u; canxi hóa đáng ngờ; bất đối xứng; bất đối xứng khu trú; bất đối xứng toàn bộ; biến dạng cấu trúc; dày da; co rút da; co rút núm vú; và hạch bạch huyết đáng ngờ), trong đó bệnh nhân có độ tuổi từ 20 đến 86 và tập trung nhiều nhất là 40-45 tuổi.



Hình 3.1. Một bộ ảnh mẫu của tập dữ liệu VinDr-Mammo: (a) Left CC; (b) Right CC; (c) Left MLO; (d) Right MLO

Bảng 3.1 Thông Tin Bộ Dữ Liệu Vindr-Mammo

Loại tổn thương	Số lượng ảnh
Khối u	1,226
Vôi hóa	543
Bất đối xứng tiêu điểm	269
Biến dạng cấu trúc	119
Bình thường	2,551

Số lượng ảnh trong bộ dữ liệu VinDr-Mammo với các loại tổn thương khác nhau là không cân bằng.

3.2 THU THẬP ẢNH X-QUANG VÚ TẠI BỆNH VIỆN UNG BƯỞU

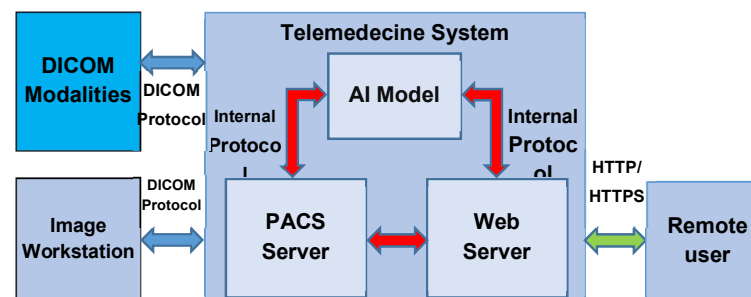
Trong nghiên cứu này, hình ảnh chụp ảnh X-quang vú DICOM thu thập từ Bệnh viện Ung bướu Việt Nam được sử dụng. Quá trình thu thập hình ảnh được thực hiện theo một quy trình được thiết kế chặt chẽ để đảm bảo tính hợp lệ của tập dữ liệu như thể hiện trong Hình 3.5. Đầu tiên, hình ảnh chụp ảnh X-quang vú DICOM được thu thập từ thiết bị lưu trữ hình ảnh PACS của bệnh viện và được

5.3 HỆ THỐNG HỖ TRỢ CHẨN ĐOÁN TỪ XA DÙNG MẠNG HỌC SÂU

Hệ thống hoạt động trên nền tảng web, giúp bác sĩ dễ dàng truy cập bằng nhiều thiết bị, đồng thời đảm bảo bảo mật và tính tương thích. Các yêu cầu chính gồm: (1) truy cập mọi lúc, mọi nơi qua internet; (2) hỗ trợ DICOM và JPEG; (3) giao tiếp qua giao diện web; (4) cung cấp kết quả phân loại loại tổn thương và mức độ nguy cơ.

2.1. Kiến trúc phần cứng của hệ thống y tế từ xa

Hình dưới minh họa một hệ thống được đề xuất để phân loại hình ảnh ung thư vú từ xa. Hệ thống này bao gồm các thành phần cốt lõi, các thiết bị ngoại vi tương tác và các kết nối giữa chúng. Cụ thể, hệ thống bao gồm ba thành phần chính: 1. Máy chủ Hệ thống Lưu trữ và Truyền thông Hình ảnh (PACS) cho hình ảnh; 2. Máy chủ web để kết nối dữ liệu với mô hình AI và PACS; 3. Mô hình AI để phân loại hình ảnh.



Hình 1. Minh họa hệ thống y tế từ xa dùng để phân loại hình ảnh ung thư vú bằng mô hình AI.

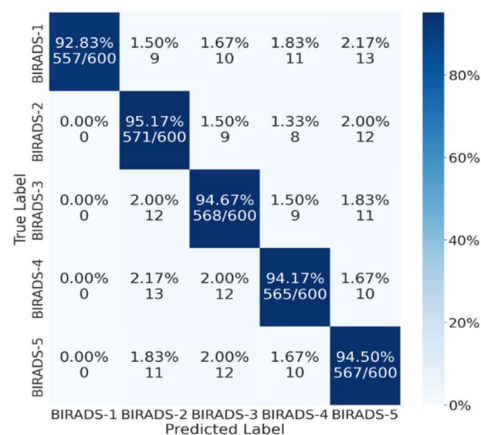
Mô tả các giao thức nội bộ của lõi hệ thống

Mô hình AI sử dụng các giao thức nội bộ để trao đổi thông tin giữa các thành phần và giao tiếp với thiết bị ngoại vi qua các chuẩn phổ biến như DICOM (máy X-quang, thiết bị hiển thị) và HTTP/HTTPS (thiết bị đầu cuối). Hệ thống kết nối với PACS để lấy và xử lý hình ảnh trước khi phân loại, đồng thời máy chủ web quản lý dữ liệu, hiển thị và gửi yêu cầu chẩn đoán đến mô hình AI để trả kết quả cho người dùng. Các tham số hệ thống (kích thước ảnh, dung lượng lưu trữ...) được tối ưu dựa trên dữ liệu và yêu cầu triển khai nhằm đảm bảo hiệu suất cao.

Bảng 5.1. Số lượng ảnh sau khi tăng cường của tập ảnh

Category	Original	Augmentation/Choice	Division of image sets	
			Train n	Test
BIRADS-1	13406	3000	2400	600
BIRADS-2	4676	3000	2400	600
BIRADS-3	930	3000	2400	600
BIRADS-4	762	3000	2400	600
BIRADS-5	226	3000	2400	600
Total	20000	15000	12000	3000

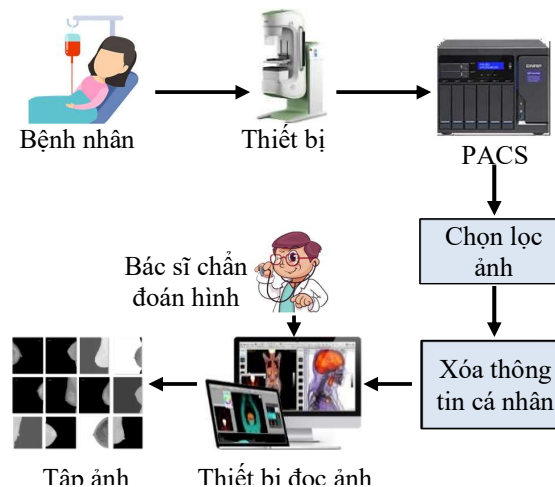
b) Kết quả phân loại hình ảnh tổn thương vú theo BIRADS



Hình 5.5. Ma trận nhầm lẫn

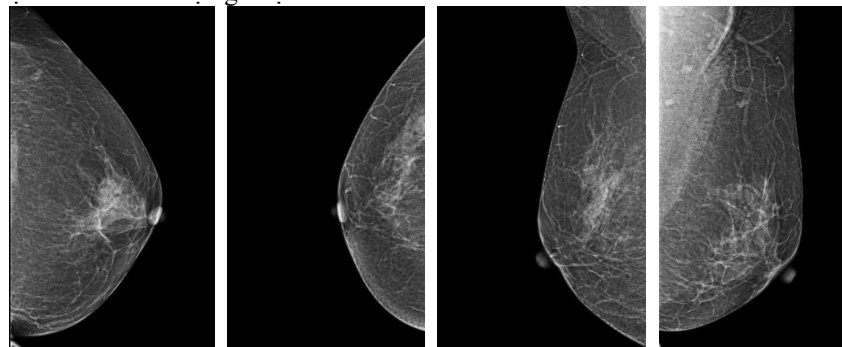
Kết quả cho thấy tiền xử lý ảnh (XLA) giúp cải thiện rõ rệt độ chính xác phân loại ảnh UTV theo thang BIRADS. Đặc biệt, phân đoạn ảnh và dùng Colormap tăng độ chính xác khoảng 8%, trong khi hiệu chỉnh Gamma chỉ cải thiện ~1% và cân bằng Histogram gần như không đáng kể. Mô hình EfficientNet phù hợp hơn với ảnh màu. Trong nghiên cứu, hệ thống đạt độ chính xác trung bình 92,6% và F1-score 92,33% trên bộ VinDr-Mammo, vượt trội so với LightGBM trước đây chỉ đạt 67%.

ẩn danh để bảo vệ quyền riêng tư của bệnh nhân. Tiếp theo, tập dữ liệu được một số bác sĩ chẩn đoán hình ảnh chủ thích thông qua công cụ gắn nhãn trên máy tính của bệnh viện. Cuối cùng, hình ảnh đã được chủ thích được lưu trữ dưới dạng tập dữ liệu.



Hình 3.5 Sơ đồ thu thập dữ liệu tại bệnh viện Ung Bướu TP.HCM

Trong bước này, hơn 4000 hình ảnh chụp ảnh X-quang vú ở định dạng DICOM từ Hệ thống PACS của hệ thống DICOM của Bệnh viện Ung Bướu TP.HCM, với các lần chụp trong khoảng thời gian từ năm 2018 đến năm 2023 được thu thập. Để đảm bảo quyền riêng tư của bệnh nhân được bảo vệ, thông tin có thể nhận dạng được của bệnh nhân trong các thẻ DICOM sẽ được xóa hoàn toàn. Chỉ thông tin cần thiết được sử dụng xử lý hình ảnh DICOM và thông tin độ tuổi của bệnh nhân mới được giữ lại.



Hình 3.6 Một số ảnh từ tập dữ liệu thu tại bệnh viện Ung Bướu TP.HCM

Thông số của hình ảnh thu thập được tại BV Ung Bướu được mô tả trong bảng dưới đây.

Bảng 3.4 Biểu diễn thông tin chi tiết của hình ảnh sau khi thu thập

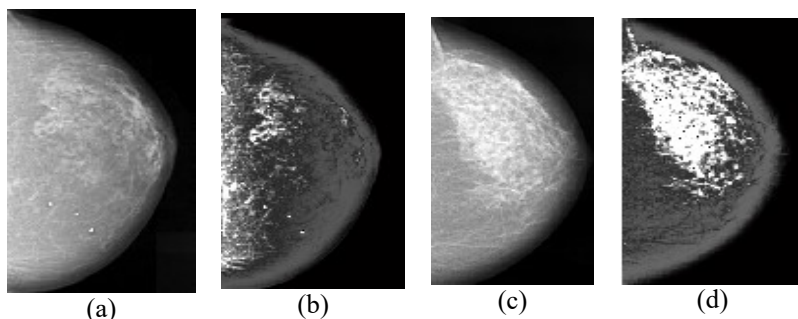
STT	Thông số	Giá trị
1	Số lượng ảnh	4034
2	Kích thước ảnh	1024×1024
3	Định dạng	DICOM
4	Kích thước file	8.8MB

Đối với tập dữ liệu tại bệnh viện ung bướu, việc gán nhãn được thực hiện bởi các bác sĩ chẩn đoán hình ảnh của bệnh viện ung bướu. Sau khi hình được chụp bởi kỹ thuật viên, bác sĩ chẩn đoán hình ảnh sẽ tiến hành phân tích ảnh X-quang một cách có hệ thống, thường theo một trình tự nhất định để không bỏ sót bất kỳ chi tiết nào.

3.3 TĂNG CƯỜNG HÌNH ẢNH

Điều chỉnh độ sáng và tương phản ảnh UTV

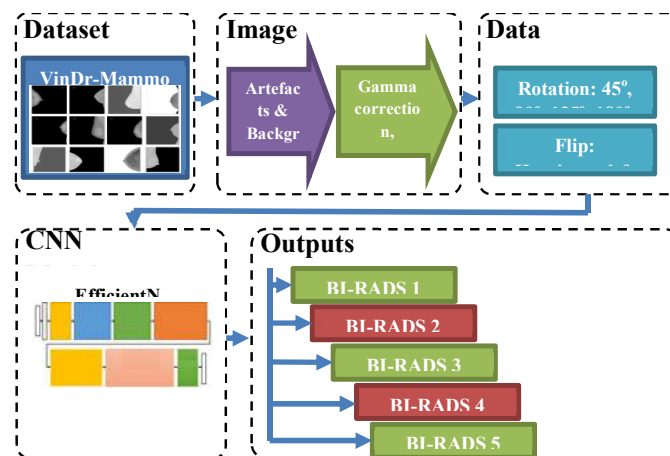
Sau khi thu thập tập dữ liệu từ MIAS, các ảnh đầu vào được nâng cao bằng cách sử dụng logic mờ và thuật toán phân phối khả năng.



(a) hình ảnh ban đầu của trường hợp bình thường; (b) hình ảnh nâng cao của trường hợp bình thường; (c) hình ảnh ban đầu của trường hợp ung thư; (d) hình ảnh nâng cao về trường hợp ung thư.

Hình 3.9 Biểu diễn hình ảnh gốc và hình ảnh nâng cao

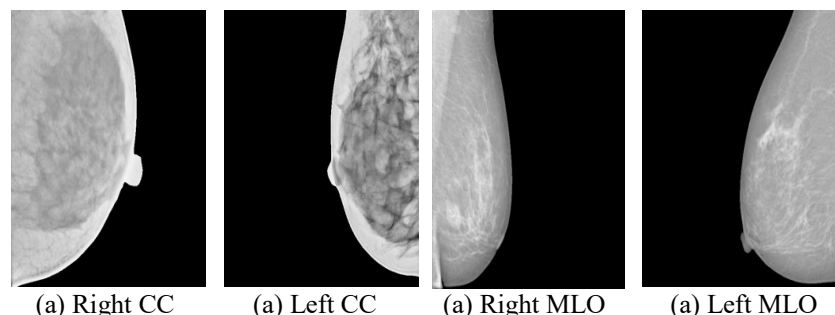
Hiệu chỉnh gamma:



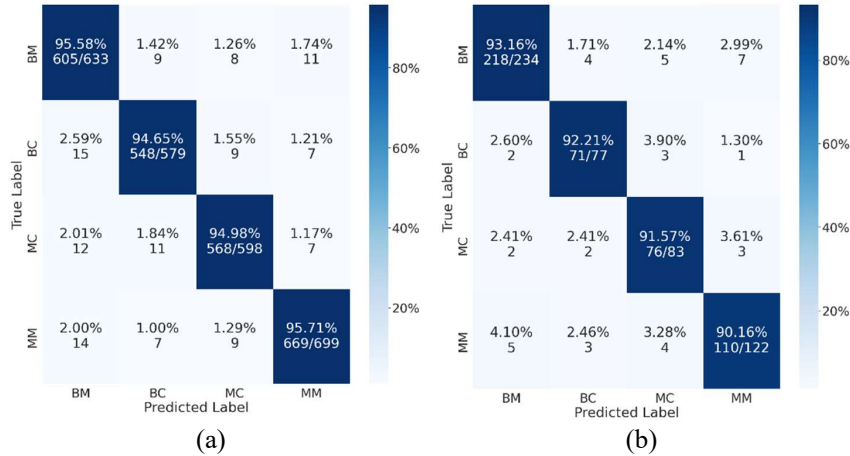
Hình 5.3. Sơ đồ khối quá trình xử lý của hệ thống phân loại ảnh X-quang vú theo mức độ tổn thương theo thang BIRADS

Trước khi đưa vào CNN, ảnh trong tập DL được tiền xử lý bằng cách loại bỏ thông tin không cần thiết và tăng cường chất lượng. Tập dữ liệu sau đó được mở rộng bằng các phép xoay, lật và thay đổi độ sáng để cân bằng số lượng ảnh giữa các mức tổn thương. Các ảnh được huấn luyện trên EfficientNet (khởi tạo từ ImageNet), tinh chỉnh các tầng cuối và thêm lớp ngõ ra phù hợp. Kết quả phân loại được so sánh để chọn mô hình hiệu quả nhất cho phân loại ảnh X-quang vú theo thang BIRADS [171].

a) Tập DL



(a) Right CC (a) Left CC (a) Right MLO (a) Left MLO
Hình 5.4. Bộ 4 ảnh trong một kết quả khám X-quang vú của tập ảnh VinDr-Mammo



Hình 5.2. Mô tả hiệu quả phân loại của hai loại tập DL với Vô hình lành tính (BC), Khối lành tính (BM), Vô hình ác tính (MC), Khối lượng ác tính (MM). (a) Bộ DL CBIS-DDSM; (b) Bộ DL MIAS.

Việc trích xuất ROI lớn và nâng cao chất lượng hình ảnh đã cải thiện hiệu suất phân loại của mô hình đề xuất, đặc biệt, mô hình đề xuất trong bài viết này đạt độ chính xác 95%. Để có hiệu suất cao, các bộ hình ảnh được trích xuất với ROI lớn để ghi lại nhiều đặc điểm của tổn thương vú, sau đó được chuẩn hóa và nâng cao. Để đánh giá hiệu quả của mô hình đề xuất, bài viết trình bày ba trường hợp tập DL như trong Bảng 5.8. Trong đó, tập ảnh được xử lý theo đề xuất có ROI lớn nhất cho hiệu suất cao nhất là 95%, trong khi tập ảnh được xử lý theo mô hình đề xuất, không xử lý thì độ chính xác nhỏ hơn 82,1% so với chỉ phân đoạn Otsu. Kết quả này chứng tỏ mô hình đề xuất có thể phân loại 4 loại tổn thương vú giúp xác định UTV qua chụp X quang tuyến vú dễ dàng hơn với độ chính xác cao nhất.

5.2 KẾT QUẢ PHÂN LOẠI BỆNH UTV THEO BIRADS

Nghiên cứu này nhằm phân loại ảnh nhũ ảnh theo thang BIRADS bằng mô hình EfficientNet kết hợp các phương pháp XLA, với mục tiêu tìm ra mô hình xử lý tối ưu để hỗ trợ chẩn đoán sớm ung thư vú. Thí nghiệm được thực hiện trên bộ dữ liệu VinDr-Mammo [24] thu thập từ các bệnh viện tại Việt Nam.

Hiệu chỉnh gamma, nhằm điều chỉnh độ sáng của hình ảnh để làm nổi bật sự chuyển đổi từ tối sang sáng và ngược lại bằng cách sử dụng các phép biến đổi phi tuyến. Thuật toán hiệu chỉnh gamma được thực hiện như sau.

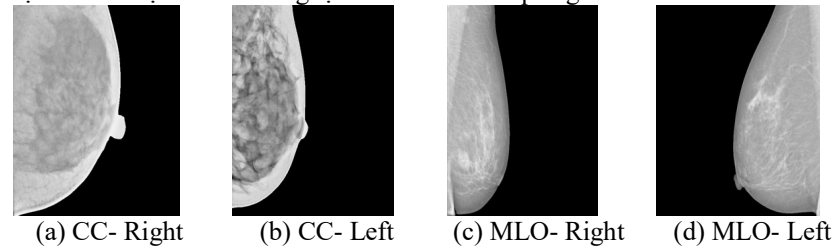
Thuật toán 2. Hiệu chỉnh Gamma thích ứng

```

1  START
2      READ original image  $I_{IN}(x,y)$  where  $x$  and  $y$ 
    denotes coordinates of the pixel
3      Select Gamma initial value  $\gamma_0$ 
4      WHILE NOT CorrectGamma( $I_{OUT}(x,y)$ )
5          Apply Gamma correction  $I_{OUT}(x,y) =$ 
             $I_{IN}(x,y)^\gamma$ 
6      END WHILE
7      OUTPUT  $I_{OUT}(x,y)$ 
8  END

```

Hình 3.10 bên dưới cho thấy kết quả của việc thực hiện hiệu chỉnh gamma trên ảnh đã được loại bỏ nền và các thành phần không mong muốn. Thuật toán Adaptive Gamma correction giúp đồng nhất chất lượng ảnh và loại bỏ sự khác biệt do thiết bị thu ảnh mang lại trên các ảnh X-quang vú.

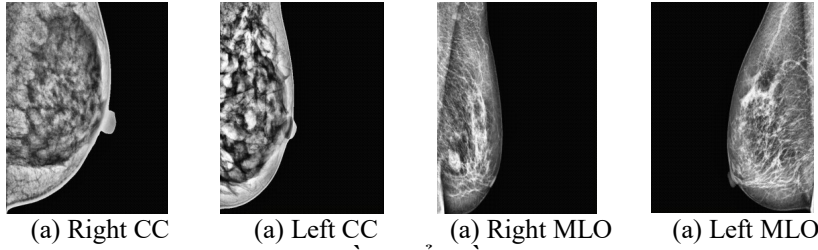


Hình 3.10 Kết quả của việc thực hiện hiệu chỉnh gamma

Cân bằng biểu đồ ảnh UTV

Sau khi hiệu chỉnh gamma, thuật toán cân bằng độ tương phản CLAHE được sử dụng để đạt được độ tương phản tối ưu trên hình ảnh. CLAHE là một biến thể cải tiến của thuật toán AHE (Adaptive Histogram Equalization) [138]. CLAHE thường được dùng cho hình ảnh y tế nhằm cải thiện chất lượng của các cấu trúc phức tạp [139].

Kết quả của bước này được thể hiện ở **Error! Reference source not found..** Trong đó có thể thấy rõ sự khác biệt khi so sánh giữa hình ảnh trước và sau khi điều chỉnh độ sáng.

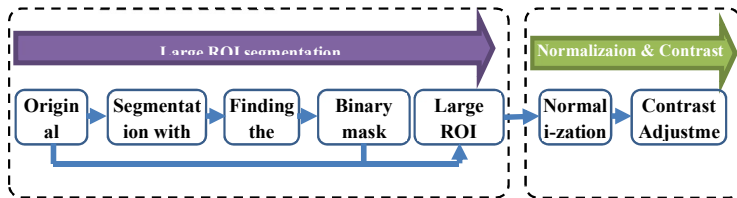


Hình 3.11 Cân bằng biểu đồ sử dụng CLAHE

CHƯƠNG 4 TRÍCH ĐẶC TRƯNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UNG THƯ VÚ

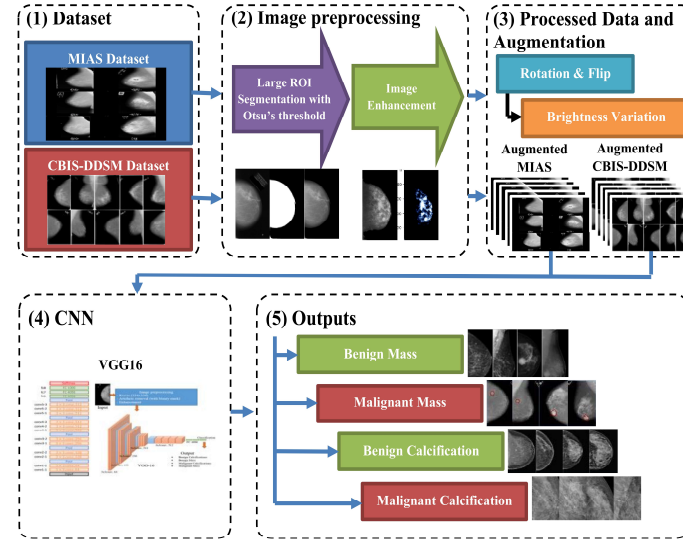
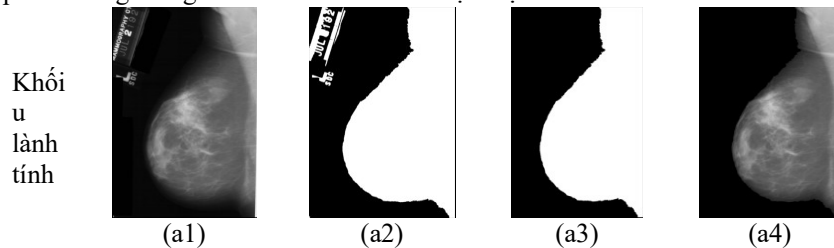
4.1 PHÂN ĐOẠN ẢNH UTV ĐỂ TRÍCH XUẤT VÙNG ROI

Trong hệ thống phân loại ảnh UTV, tất cả hình ảnh cần được xử lý để trích xuất ROI lớn nhất, tập trung vào vùng tổn thương nhằm tăng độ chính xác và giảm thời gian tính toán. Thuật toán Otsu được dùng để phân đoạn ảnh, sau đó chỉ giữ lại ROI lớn nhất dựa trên mật độ điểm ảnh, loại bỏ các ROI nhỏ hơn. ROI này được nhân với ảnh gốc để thu được vùng tổn thương, sau đó chuẩn hóa và điều chỉnh như minh họa ở Hình 4.2.



Hình 4.1 Quy trình trích xuất vùng ROI lớn nhất

Hình 4.3 cho thấy kết quả của việc tách vùng ROI lớn nhất, trong đó các thành phần không mong muốn và nền ảnh đã được loại bỏ.



Hình 5.1 Sơ đồ khối phân loại và quy trình: (1) hai tập DL với bốn bộ ảnh tổn thương vú; (2) Hình ảnh được phân đoạn để tìm ROI lớn nhất và được tăng cường; (3) Hình ảnh được tăng cường từ các hình ảnh ROI lớn nhất; (4) và (5) VGG16 với bốn đầu ra để phân loại tổn thương vú.

Phần này đánh giá độ chính xác của VGG16 trên ba bộ ảnh đầu vào: không phân đoạn Otsu, chỉ phân đoạn Otsu, và ROI lớn nhất từ phân đoạn Otsu. Dữ liệu nhũ ảnh lấy từ CBIS-DDSM và MIAS, trong đó CBIS-DDSM (phiên bản nâng cấp của DDSM) gồm 1.644 ảnh thuộc bốn loại tổn thương vú. Tất cả ảnh được chuẩn hóa về kích thước 224×224 pixel và mã màu RGB.

Kết quả phân loại các loại tổn thương tập ảnh MIAS

Trên bộ DL CBIS-DDSM, mô hình đạt độ chính xác trung bình 95,23%, cao nhất với MM (95,71%) và BM (95,58%). Trên bộ DL MIAS, độ chính xác trung bình là 91,78%, nổi bật ở BM (93,16%) và BC (92,21%). Sai lệch chủ yếu xảy ra giữa BM và MM, nhưng tỷ lệ thấp (<5%). Kết quả này khẳng định hiệu quả và tiềm năng của mô hình.

Trong Bảng 4.3, kết quả mô phỏng cho thấy các phương pháp tiền xử lý ảnh được đề xuất làm tăng đáng kể độ chính xác tổng thể của việc phân loại các tập ảnh tổn thương vú. Cụ thể, tập ảnh bằng GROI tạo ra độ chính xác cao hơn 82,3% so với tập ảnh gốc chỉ đạt 74,3%. Trong khi đó, tập ảnh bằng CROI đạt độ chính xác tăng dần khoảng 92,6% điểm F1. Điều này có nghĩa là mô hình EfficientNet-B7 rất phù hợp với tập dữ liệu hình ảnh có CROI.

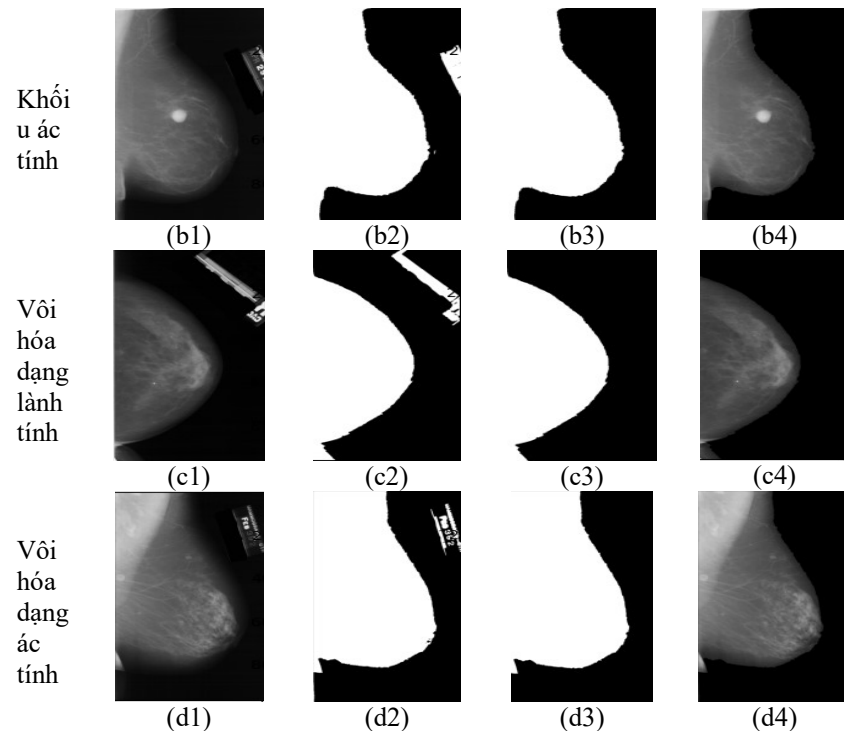
Bảng 4.3. So sánh độ chính xác phân loại giữa ảnh được tô màu và ảnh xám

Loại ảnh	Độ chính xác huấn luyện (%)	Độ chính xác kiểm tra (%)	F1 score (%)
Ảnh xám	87.6	84.3	84.2
Ảnh tô màu	94.2	92.6	92.6

CHƯƠNG 5 PHÂN LOẠI ẢNH UTV SỬ DỤNG MẠNG HỌC SÂU VÀ HỆ THỐNG HỖ TRỢ CHẨN ĐOÁN UTV TỪ XA SỬ DỤNG MẠNG HỌC SÂU

5.1 BÀI TOÁN PHÂN LOẠI CÁC LOẠI TỔN THƯƠNG DỰA TRÊN TẬP ẢNH X-QUANG VÚ MIAS

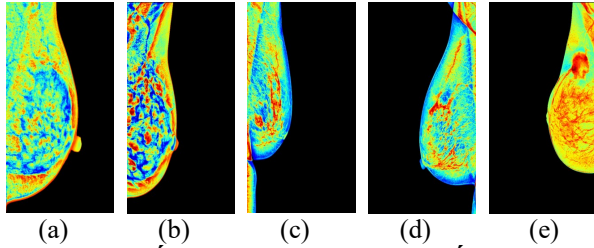
Dữ liệu được lấy từ CBIS-DDSM và MIAS, sau đó tăng cường bằng xoay, lật và thay đổi độ sáng với bốn loại tổn thương (u lạnh, u ác, vôi hóa lạnh, vôi hóa ác). Các ảnh được xử lý bằng phân đoạn Otsu để chọn ROI lớn nhất rồi đưa vào mạng VGG16, tinh chỉnh hai lớp cuối và thêm lớp đầu ra. Hiệu quả phân loại được đánh giá qua ma trận nhầm lẫn [142].



Hình 4.2. Mô tả trích xuất ROI dựa trên đường viền lớn nhất và loại bỏ các thành phần không mong muốn: (a1-a4) Hình ảnh gốc của bốn tổn thương vú; (b1-b4) Hình ảnh được phân đoạn bằng phương pháp ngưỡng Otsu (c1-c4) Hình ảnh nhị phân có ROI lớn nhất đã loại các thành phần không mong muốn; (d1-d4) Ảnh gốc có ROI lớn nhất không có thành phần không mong muốn

4.2 TÔ MÀU CHO VÙNG ROI

Ảnh UTV là ảnh xám một kênh, trong khi EfficientNet được huấn luyện trên ảnh màu ba kênh. Để tận dụng mô hình, nghiên cứu phát triển phương pháp tô màu ROI (CROI), làm nổi bật khối u hoặc vôi hóa. Thuật toán gồm hai bước: Nhận diện vùng xám bất thường và So sánh sự khác biệt với nền bằng chỉ số mức xám và các phép biến đổi hình thái (TopHat, BlackHat). Kết quả chuyển thành ICROI có màu, giúp tăng hiệu quả phân loại. Hình 4.5 minh họa 5 ảnh UTV với các mức BIRADS khác nhau, trong đó BIRADS 4 và 5 thể hiện rõ các vùng màu, nổi bật dấu hiệu tổn thương nguy cơ cao.

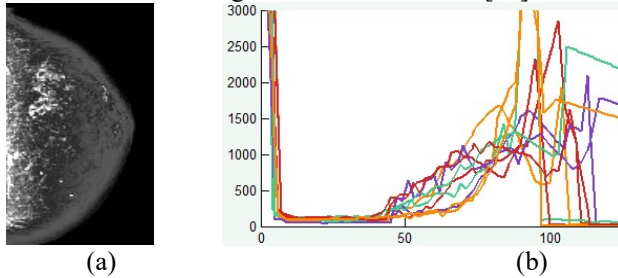


Hình 4.3. Kết quả tô màu vùng ROI đối với 5 ảnh với các mức BIRADS: (a) BIRADS 1; (b) BIRADS 2; (c) BIRADS 3; (d) BIRADS 4; (e) BIRADS 5

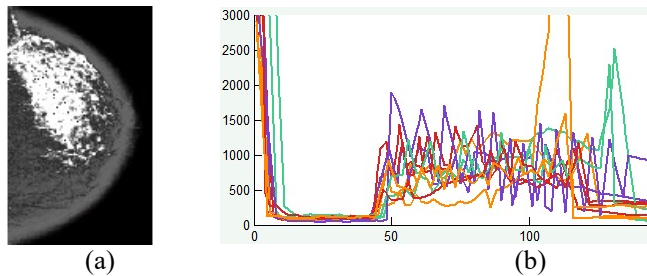
4.3 ỨNG DỤNG PHÂN LOẠI UTV DỰA TRÊN ĐẶC TRƯNG ẢNH

Phân loại ảnh UTV dựa vào đặc trưng độ tương phản

Tất cả ảnh UTV được tính toán để xây dựng các lược đồ mức xám. Kết quả cho thấy mật độ mức xám giữa hai loại ảnh UTV (ung thư và bình thường) có sự khác biệt, như thể hiện trong Hình 4.6 và Hình 4.7 [77].

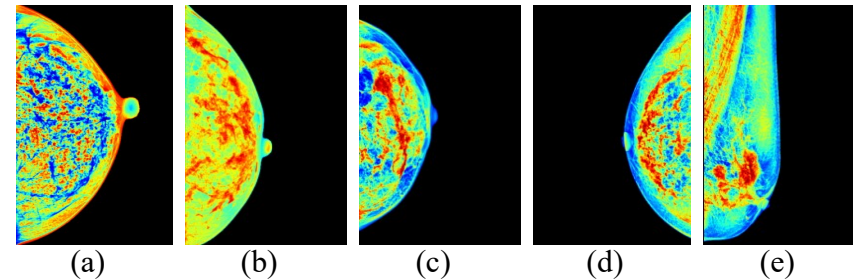


Hình 4.4. Lược đồ xám của ảnh bình thường: (a) Ảnh vú bình thường; (b) Lược đồ xám của ảnh



Hình 4.5. Lược đồ xám của ảnh UTV: (a) Ảnh vUTV; (b) Lược đồ xám của ảnh

Sau khi tạo ra các ảnh với GROI, các ảnh này được chuyển đổi thành ảnh CROI bằng thuật toán tô màu. Hình 4.15 minh họa kết quả của năm loại ảnh màu với các loại tổn thương khác nhau. Có thể thấy rằng các vùng màu trên CROI khác biệt rõ ràng so với các vùng mô bình thường. Điều này có nghĩa là màu sắc trong CROI đã làm nổi bật một số khu vực có khả năng nguy cơ cao mắc UTV. Do đó, các ảnh CROI này có thể mang lại độ chính xác phân loại cao khi sử dụng với mô hình EfficientNet-B7.



Hình 4.12. Minh họa bốn loại tổn thương khác nhau của ảnh với CROI: (a) Không phát hiện; (b) Khối u; (c) Vô hóa; (d) Khối u & Vô hóa.

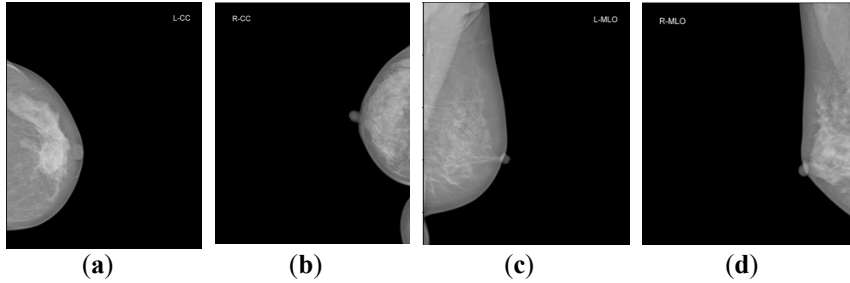
Hình 4.15 cho thấy CROI làm nổi bật tổn thương bằng màu sắc (đỏ cho khối u, xanh và đỏ cho vô hóa), giúp phân biệt rõ với ảnh không tổn thương. Khi so sánh ma trận nhầm lẫn (Hình 4.16 và 4.17), mô hình EfficientNet-B7 đạt độ chính xác cao hơn với ảnh ROI tô màu (92,6%) so với ROI xám (84,3%), chứng minh hiệu quả của phương pháp tô màu.

	Normal	Arch Disor	Arch Disor	Arch Disor	Normal
True Label	85.00% 510/600	4.50% 27	3.67% 22	2.83% 17	4.00% 24
Arch Disor	2.50% 15	87.67% 526/600	3.17% 19	4.67% 28	2.00% 12
Arch Disor	3.67% 22	3.50% 21	84.50% 507/600	3.17% 19	5.17% 31
Calibration	3.83% 23	6.50% 39	6.17% 37	80.17% 481/600	3.33% 20
Mass	5.17% 31	3.67% 22	3.33% 20	3.83% 23	84.00% 504/600
Predicted Label	Normal	Arch Disor	Arch Disor	Arch Disor	Normal

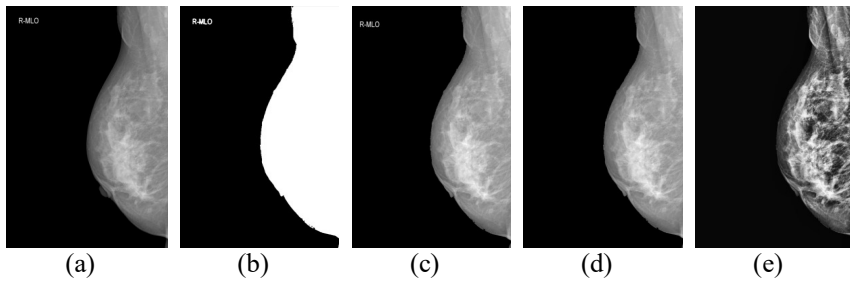
Hình 4.13. Ma trận nhầm lẫn để đánh giá hiệu suất phân loại đối với các tập ảnh xám

	Normal	Arch Disor	Arch Disor	Arch Disor	Normal
True Label	92.83% 557/600	1.50% 9	1.67% 10	1.83% 11	2.17% 13
Arch Disor	2.50% 15	92.67% 556/600	1.50% 9	1.33% 8	2.00% 12
Arch Disor	2.00% 12	1.83% 11	92.83% 557/600	1.50% 9	1.83% 11
Calibration	2.17% 13	2.00% 12	1.83% 11	92.83% 557/600	1.67% 10
Mass	1.83% 11	2.00% 12	1.67% 10	2.17% 13	92.33% 554/600
Predicted Label	Normal	Arch Disor	Arch Disor	Arch Disor	Normal

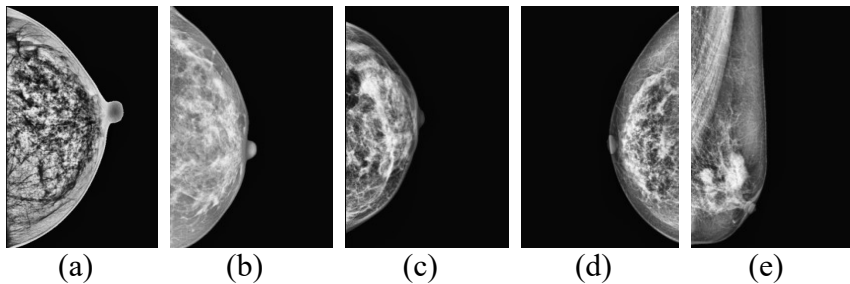
Hình 4.14. Kết quả Ma trận nhầm lẫn để đánh giá hiệu suất phân loại đối với các tập ảnh có tô màu



Hình 4.9. Bộ 4 ảnh chụp tuyến vú mẫu cho một lần chụp trong bộ dữ liệu VinDr-Mammo: (a) CC bên trái; (b) CC bên phải; (c) MLO bên trái; (d) MLO bên phải.



Hình 4.10. Minh họa quá trình xử lý ảnh: (a) Ảnh gốc; (b) Ảnh sau khi áp dụng mặt nạ với ngưỡng tối ưu; (c) Ảnh đã phân đoạn; (d) Ảnh sau các phép toán hình thái học và làm sạch; (e) Ảnh với ROI sau hậu xử lý.



Hình 4.11. Minh họa năm loại ảnh với ROI và nền đen: (a) bình thường; (b) bất đối xứng khu trú; (c) biến dạng cấu trúc; (d) vôi hóa; (e) khối u.

Hình 4.6 và Hình 4.7 cho thấy rằng trong các trường hợp ung thư, các điểm ảnh tương ứng với khoảng mức xám từ khoảng 50 đến 130 trong lược đồ thay đổi đột ngột từ thấp lên cao, trong khi ở các trường hợp bình thường, các điểm ảnh trong cùng khoảng mức xám lại tăng chậm. Do đó, về cơ bản có thể đánh giá sự khác biệt giữa hai loại ảnh vú. Tuy nhiên, các giá trị tương phản và số lượng điểm ảnh tương ứng cung cấp thêm nhiều thông tin hơn cho việc đánh giá tình trạng bệnh vú.

Bảng 4.1. Trích xuất đặc trưng của các trường hợp ung thư và bình thường

Trường hợp	Ung thư		Bình thường	
	Độ tương phản	Số điểm ảnh	Độ tương phản	Số điểm ảnh
1	0.1927	558	0.0758	319
2	0.1478	417	0.0711	331
3	0.1582	484	0.0702	207
4	0.1419	402	0.0269	343
5	0.1501	601	0.0458	391
6	0.0817	513	0.0181	212
7	0.2799	527	0.0584	388
8	0.2483	560	0.0855	326
9	0.1185	442	0.0746	98
10	0.1455	420	0.0295	77

Bảng 4.1 cho thấy các giá trị tương phản tương ứng với số lượng điểm ảnh trung bình của ảnh ung thư và ảnh bình thường trong khoảng mức xám từ 0,2 đến 0,3 trên lược đồ. Rõ ràng, giá trị tương phản trong ảnh bình thường nằm trong khoảng từ 0,0181 đến 0,0855, trong khi ở ảnh ung thư là từ 0,0817 đến 0,2799. Ngoài ra, số lượng điểm ảnh tương ứng với độ tương phản ảnh của 10 ảnh cho mỗi loại cũng khác nhau. Cụ thể, số lượng điểm ảnh trong ảnh ung thư nằm trong khoảng từ 402 đến 601, trong khi ở ảnh bình thường là từ 77 đến 391.

Phân loại ảnh UTV dựa trên đặc trưng GLCM và mô hình SVM

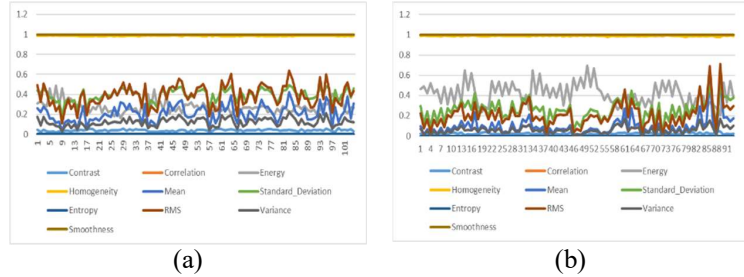
Dữ liệu hình ảnh được khai thác trong nghiên cứu này được lấy từ bộ dữ liệu Mini-DDSM [144]. Các ảnh X-quang tuyến vú (mammogram) được sử dụng có định dạng ảnh xám 16-bit, bao gồm 9.684 ảnh thuộc 3 loại: lành tính (Benign), ung thư (Cancer) và bình thường (Normal). Nghiên cứu này sử dụng một phần của bộ dữ liệu Mini-DDSM để xử lý, trích xuất đặc trưng và phân loại UTV.

Bảng 4.2. Biểu diễn tập dữ liệu cho huấn luyện và kiểm thử của 2 loại ảnh

Tập dữ liệu	Ung thư	Bình thường
-------------	---------	-------------

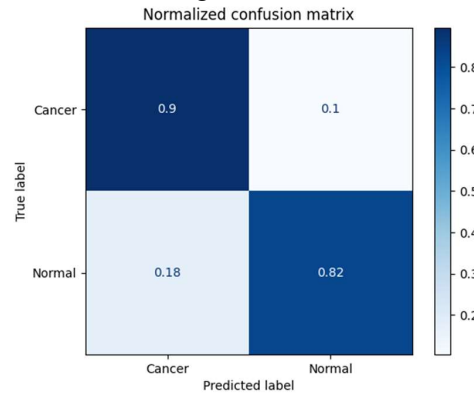
Huấn luyện	100	100
Kiểm tra	50	50

Hình ảnh UTV được phân tích để trích xuất đặc trưng bằng GLCM và điều này cho phép phát hiện ung thư tiềm ẩn ở vú, tốt nhất cho phân loại SVM.



Hình 4.6. Thống kê đặc điểm của hình ảnh vú sử dụng GLCM: (a) Đặc trưng của hình ảnh bình thường; (b) Đặc trưng của hình ảnh ung thư

Trong bước kiểm tra sau khi phân loại SVM cho dữ liệu được kiểm tra, việc tính toán để tìm ra độ chính xác và độ nhạy của hệ thống được thực hiện bằng cách áp dụng ma trận nhầm lẫn như trong hình 4.10.



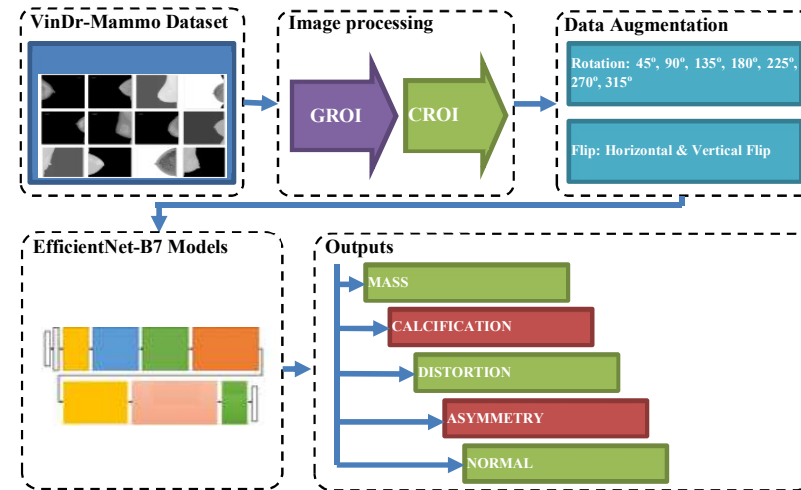
Hình 4.7. Ma trận nhầm lẫn

Với bộ dữ liệu hình ảnh ung thư gồm 50 hình ảnh, 45 hình ảnh đã được phân loại đúng với độ chính xác 90%. Ngoài ra, với 50 ảnh thông thường, hệ thống đã phân loại để cho ra kết quả chính xác là 41 ảnh, chiếm 82%. Như vậy có thể thấy, độ

chính xác trung bình cho cả hai loại là 86%. Các kết quả này cho thấy ý nghĩa của các đặc trưng tìm được trong việc phân loại ảnh UTV [145].

Phân loại ảnh UTV dựa trên vùng ROI

Tập ảnh UTV được chia thành năm loại tổn thương: khối u (mass), vôi hóa (calcification), biến dạng cấu trúc (architectural distortion), bất đối xứng (asymmetry) và bình thường (normal). Để thực hiện phân loại các tập ảnh này bằng mô hình EfficientNet-B7, dữ liệu đầu vào sẽ được xử lý qua hai bước: trích xuất GROI trong các tập ảnh; chuyển đổi CROI sang CROI như minh họa trong Hình 4.11 [143]. Bên cạnh đó, các tập dữ liệu sẽ được cân bằng bằng cách tăng số lượng hình ảnh, đồng thời ngẫu nhiên giảm bớt các tập có kích thước quá lớn.



Hình 4.8. Sơ đồ tổng thể của phương pháp phân loại ảnh UTV dựa trên vùng ROI

Trong nghiên cứu này, bộ dữ liệu VinDr-Mammo [135] được sử dụng. Bộ dữ liệu này bao gồm 5.000 lần chụp tuyến vú, tương ứng với 20.000 ảnh, mỗi lần chụp tạo ra 4 ảnh vú: Cranial-Caudal (CC) bên phải, CC bên trái, Mediolateral-Oblique (MLO) bên phải, và MLO bên trái, như minh họa trong Hình 4.12.