

BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ KỸ THUẬT  
THÀNH PHỐ HỒ CHÍ MINH

NGUYỄN THANH TÂM

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

LUẬN ÁN TIẾN SĨ  
NGÀNH: KỸ THUẬT ĐIỆN TỬ

Tp. Hồ Chí Minh, tháng 1 năm 2026

BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ KỸ THUẬT  
THÀNH PHỐ HỒ CHÍ MINH

NGUYỄN THANH TÂM

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

LUẬN ÁN TIẾN SĨ  
NGÀNH: KỸ THUẬT ĐIỆN TỬ

Người hướng dẫn khoa học 1: PGS.TS. Nguyễn Thanh Hải

Phản biện 1:

Phản biện 2:

Phản biện 3:

Tp. Hồ Chí Minh, tháng 1 năm 2026

BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT  
THÀNH PHỐ HỒ CHÍ MINH

Số: 3459/QĐ-ĐHSPKT

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM  
Độc lập – Tự do – Hạnh phúc

TP. Hồ Chí Minh, ngày 10 tháng 12 năm 2024

**QUYẾT ĐỊNH**  
**Về việc đổi tên luận án tiến sĩ**

**HIỆU TRƯỞNG TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HỒ CHÍ MINH**

Căn cứ Luật Giáo dục đại học ngày 18/6/2012 và Luật sửa đổi, bổ sung một số điều của Luật Giáo dục đại học ngày 19/11/2018;

Căn cứ Nghị định 99/2019/NĐ-CP ngày 30/12/2019 của Chính phủ Quy định chi tiết và hướng dẫn thi hành một số điều của Luật sửa đổi, bổ sung một số điều của Luật giáo dục đại học;

Căn cứ Quyết định số 465/QĐ-BGDĐT ngày 31 tháng 01 năm 2024 của Bộ Giáo dục và Đào tạo về việc giao quyền tự chủ tài chính giai đoạn 2023 - 2025 cho các đơn vị sự nghiệp công lập trực thuộc;

Căn cứ Nghị quyết số 11/NQ-HĐT ngày 08/01/2021 của Hội đồng trường ban hành Quy chế tổ chức hoạt động của Trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh; Nghị quyết số 118/NQ-HĐT ngày 27/01/2023, Nghị quyết số 136/NQ-HĐT ngày 14/10/2023 và Nghị quyết số 182/NQ-HĐT ngày 28/09/2024 của Hội đồng trường về việc sửa đổi, bổ sung Quy chế tổ chức hoạt động của Trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh;

Căn cứ Nghị quyết số 188/NQ-HĐT ngày 23/10/2024 của Hội đồng trường về việc tiếp tục giao Quyền Hiệu trưởng Trường Đại học Sư phạm Kỹ thuật Thành phố Hồ Chí Minh;

Căn cứ Thông tư số 08/2017/TT-BGDĐT ngày 04/4/2017 của Bộ Giáo dục và Đào tạo về việc Ban hành Quy chế tuyển sinh và đào tạo trình độ tiến sĩ;

Căn cứ Quyết định số 1557/QĐ-ĐHSPKT ngày 30/8/2017 của Trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh về việc ban hành Quy định đào tạo tiến sĩ của trường đại học Sư phạm Kỹ thuật TP.HCM;

Theo đề nghị của Nghiên cứu sinh, Khoa quản ngành và Trường phòng Đào tạo.

**QUYẾT ĐỊNH:**

**Điều 1.** Đổi tên luận án tiến sĩ cho:

Nghiên cứu sinh : **Nguyễn Thanh Tâm**

Ngành : Kỹ thuật điện tử

Khoá: 2020 – 2023

Tên luận án mới : **Phát triển hệ thống hỗ trợ chẩn đoán bệnh ung thư vú ứng dụng trí tuệ nhân tạo**

Người HD thứ nhất (HD chính): **PGS.TS. Nguyễn Thanh Hải**

Người HD thứ hai : **không**

Thời gian thực hiện : **01/6/2020 đến 31/5/2023**

3771

**Điều 2.** Giao cho Phòng Đào tạo quản lý, thực hiện theo đúng Quy chế đào tạo trình độ tiến sĩ của Bộ Giáo dục & Đào tạo và Nhà trường đã ban hành.

**Điều 3.** Trưởng các đơn vị: phòng Đào tạo, khoa/viện quản ngành, phòng KHTC và các Ông (Bà) có tên ở Điều 1 chịu trách nhiệm thi hành quyết định này. /.

**Nơi nhận:**

- Như điều 3;
- BGH (để chỉ đạo);
- Lưu: VT, ĐT (3b).

**Q. HIỆU TRƯỞNG**



**PGS.TS. Lê Hiếu Giang**

## LỜI CAM ĐOAN

Tôi cam đoan đây là công trình nghiên cứu của tôi.

Các số liệu, kết quả nêu trong Luận án là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác

*Tp. Hồ Chí Minh, ngày 26 tháng 1 năm 2026*

Nguyễn Thanh Tâm

## LỜI CẢM ƠN

Trước tiên, tôi xin gửi lời cảm ơn chân thành và sâu sắc đến Thầy **PGS. TS. Nguyễn Thanh Hải**, người Thầy luôn rất nhiệt tình và tận tâm hướng dẫn tôi trong thời gian thực hiện chuyên đề. Hơn nữa, trong suốt quá trình thực hiện từ lúc lập đề cương cho đến khi thực hiện luận án, Thầy luôn có những góp ý và định hướng giúp tôi đạt được những kết quả tốt nhất.

Tôi cũng đồng gửi lời cảm ơn đến quý Thầy Cô ở Trường Đại Học Sư Phạm Kỹ Thuật Tp. HCM nói chung cũng như các Thầy Cô ở khoa Điện – Điện tử nói riêng đã truyền đạt cho tôi những nền tảng kiến thức quý báu, giúp tôi có thể hoàn thành được chuyên đề.

Cuối cùng, tôi cũng xin được gửi lời biết ơn sâu sắc đến gia đình tôi, những người đã luôn là chỗ dựa tinh thần, là nguồn động viên vô cùng to lớn trong những lúc khó khăn, giúp tôi có thể an tâm thực hiện công việc học tập và nghiêm cứu của mình trong suốt thời gian thực hiện chuyên đề.

Xin chân thành cảm ơn!

*Tp. Hồ Chí Minh, tháng 1 năm 2026*

NCS

NGUYỄN THANH TÂM

## TÓM TẮT

Luận án nghiên cứu và phát triển một hệ thống hỗ trợ chẩn đoán Ung Thư Vú (UTV) dựa trên Trí Tuệ Nhân Tạo (AI), nhằm nâng cao độ chính xác chẩn đoán và giảm tải cho bác sĩ trong bối cảnh số ca mắc UTV tại Việt Nam ngày càng gia tăng, trong khi nguồn lực y tế còn hạn chế. Luận án tập trung vào ba hướng nghiên cứu chính: thu thập – xử lý dữ liệu hình ảnh X-quang vú, phân loại UTV bằng các mô hình học sâu, và xây dựng hệ thống chẩn đoán từ xa dựa trên AI.

Trong phần thu thập và tiền xử lý ảnh, luận án sử dụng nhiều bộ dữ liệu lớn như MIAS, CBIS-DDSM, VinDr-Mammo và ảnh thu trực tiếp từ Bệnh viện Ung Bướu TP.HCM. Các kỹ thuật xử lý ảnh nâng cao như cân bằng histogram, điều chỉnh gamma, colormap và trích xuất vùng ROI (vùng tổn thương) được áp dụng để cải thiện chất lượng ảnh đầu vào. Kết quả cho thấy các bước tiền xử lý đóng vai trò quan trọng, giúp tăng độ chính xác phân loại thêm 5–8%.

Trong phần phân loại ảnh, luận án thử nghiệm nhiều kiến trúc học sâu gồm VGG16, ResNet, DenseNet và EfficientNet. Ảnh sau tiền xử lý được đưa vào mô hình để giải quyết các bài toán: phát hiện có/không UTV, phân loại các dạng tổn thương (u, vôi hóa, bất đối xứng...), và đánh giá mức độ nguy cơ theo hệ thống BIRADS. Mô hình tinh chỉnh cho kết quả nổi bật: khoảng 95% độ chính xác cho phân loại tổn thương và khoảng 92.6% cho phân loại BIRADS, cao hơn rõ rệt so với nhiều phương pháp truyền thống.

Điểm nổi bật của luận án nằm ở quy trình hoàn chỉnh từ dữ liệu đến mô hình, bao gồm: thu thập – chuẩn hóa ảnh – trích xuất ROI – huấn luyện mô hình – đánh giá – triển khai. Ngoài ra, luận án còn phát triển hệ thống chẩn đoán từ xa, cho phép bác sĩ tải ảnh DICOM/JPEG lên máy chủ, hệ thống AI sẽ tự động phân loại và trả kết quả. Hệ thống tích hợp PACS, giao diện web và module AI, giúp hỗ trợ bác sĩ tuyến dưới, giảm tải bệnh viện lớn và tăng khả năng tiếp cận y tế vùng xa.

Cuối cùng, luận án khẳng định tính khả thi của việc ứng dụng AI trong chẩn đoán UTV tại Việt Nam, đồng thời đề xuất hướng phát triển: mở rộng dữ liệu từ nhiều bệnh viện, cải tiến mô hình đa ảnh (CC/MLO), tích hợp attention để tăng độ nhạy phát hiện vi vôi hóa, và triển khai thử nghiệm lâm sàng thực tế.

## ABSTRACT

This dissertation focuses on developing an AI-based system for classifying breast lesion classes to support breast cancer diagnosis. This can improve diagnostic accuracy and reduce the workload for radiologists in Vietnam, where breast cancer cases are rapidly increasing every day. Moreover, the dissertation covers three main research directions: data acquisition and preprocessing of mammograms; breast cancer classification using deep learning models; and the development of a remote AI-based diagnostic system.

In the data acquisition and preprocessing phase, large datasets such as MIAS, CBIS-DDSM, VinDr-Mammo, as well as real mammograms collected from Ho Chi Minh City Oncology Hospital have been used. Moreover, advanced image-processing techniques such as histogram equalization, gamma correction, colormap enhancement, and ROI (Region of Interest) extraction are applied for improving the quality of input images. Experimental results show that preprocessing plays a important role, particularly the classification accuracy achieved about 5% to 8%.

In the classification phase, multiple deep learning architectures including VGG16, ResNet, DenseNet, and EfficientNet are considered and developed. In particular, images after preprocessing and enhancing are fed into these models to solve tasks such as detecting the presence of breast cancer; classifying different lesion types (mass, calcification, asymmetry, etc.); and assessing cancer risk according to the BIRADS system. Fine-tuned models achieve notable performance: around 95% accuracy for lesion classification and around 92.6% accuracy for BIRADS classification, significantly higher than many traditional approaches.

A major contribution of this dissertation is establishing a complete end-to-end pipeline from data collection, standardization, ROI extraction, model training, evaluation, to deployment. Furthermore, the dissertation has developed a remote AI-based diagnostic platform that allows doctors to upload DICOM/JPEG mammograms to a server, where the AI system automatically analyzes and returns diagnostic results. The system integrates PACS, a web interface, and AI processing modules, providing support for lower-tier hospitals, reducing workload for central hospitals, and improving healthcare access in remote regions.

The dissertation concludes by confirming the feasibility of applying AI to breast cancer diagnosis in Vietnam and suggests future directions: expanding datasets from more hospitals, improving multi-view (CC/MLO) processing, integrating attention mechanisms to enhance microcalcification detection, conducting clinical tests to validate real-world performance.

---

---

**MỤC LỤC**

Trang tựa	TRANG
LỜI CAM ĐOAN.....	iii
LỜI CẢM ƠN.....	i
TÓM TẮT.....	i
ABSTRACT.....	i
MỤC LỤC.....	ii
DANH SÁCH CÁC CHỮ VIẾT TẮT.....	v
DANH SÁCH CÁC BẢNG.....	vii
DANH SÁCH CÁC HÌNH.....	ix
<b>Chương 1 GIỚI THIỆU TỔNG QUAN.....</b>	<b>1</b>
1.1. TÍNH CẤP THIẾT CỦA ĐỀ TÀI.....	1
1.2. MỤC TIÊU NGHIÊN CỨU CỦA LUẬN ÁN.....	9
1.2.1 Mục tiêu nghiên cứu.....	9
1.2.2 Mục tiêu cụ thể.....	9
1.3. ĐỐI TƯỢNG VÀ PHẠM VI NGHIÊN CỨU.....	9
1.3.1 Đối tượng nghiên cứu.....	9
1.3.2 Phạm vi nghiên cứu.....	10
1.4. CÁC TIẾP CẬN VÀ PHƯƠNG PHÁP NGHIÊN CỨU.....	10
1.4.1 Cách thức tiếp cận.....	10
1.4.2 Phương pháp nghiên cứu.....	10
1.5. ĐÓNG GÓP MỚI VỀ KHOA HỌC VÀ Ý NGHĨA THỰC TIỄN CỦA LUẬN ÁN 12	
1.6. CẤU TRÚC CỦA LUẬN ÁN.....	13
<b>Chương 2 CƠ SỞ LÝ THUYẾT.....</b>	<b>16</b>
2.1. TỔNG QUAN VỀ BỆNH UTV.....	16
2.1.1 Bệnh UTV.....	16
2.1.2 Chẩn đoán bệnh UTV và ứng dụng TTNT trong chẩn đoán.....	19
2.2. CÁC TẬP DỮ LIỆU HÌNH ẢNH VỀ BỆNH UTV.....	21
2.2.1 Bộ dữ liệu chụp nhũ ảnh nguồn mở.....	21
2.2.2 Bộ dữ liệu siêu âm công cộng.....	26
2.3. PHÂN ĐOẠN ẢNH VÀ XLA TRONG ẢNH UTV.....	27
2.3.1 Phân đoạn ảnh UTV.....	27
2.3.2 Cải thiện chất lượng hình ảnh trong chẩn đoán UTV.....	31

## MỤC LỤC

---

---

2.4.	PHÂN LOẠI ẢNH UTV SỬ DỤNG MÁY HỌC .....	34
2.4.1	Phát hiện và phân loại khối u vú .....	35
2.4.2	Đánh giá nguy cơ UTV .....	39
2.5.	CHẨN ĐOÁN UTV TỪ XA DỰA TRÊN TTNT .....	42
2.5.1	Mạng CNN và ứng dụng trong chẩn đoán hình ảnh .....	42
2.5.2	Tình hình nghiên cứu.....	46
Chương 3	THU THẬP VÀ TIỀN XLA X-QUANG VÚ .....	48
3.1.	THU THẬP DỮ LIỆU ẢNH X-QUANG VÚ .....	48
3.1.1	Các tập dữ liệu ảnh X-quang vú sử dụng trong phạm vi nghiên cứu .....	48
3.1.2	Thu thập ảnh X-quang vú tại BV Ung bướu .....	52
3.1.3	Gán nhãn ảnh X-quang vú .....	55
3.2.	TĂNG CƯỜNG HÌNH ẢNH .....	58
3.2.1	Điều chỉnh độ sáng và tương phản ảnh UTV.....	59
3.2.2	Hiệu chỉnh Gamma.....	61
3.2.3	Cân bằng biểu đồ ảnh UTV.....	62
3.3.	TÔ MÀU CHO ẢNH X-QUANG VÚ .....	64
3.3.1	Chuyển sang ảnh màu sử dụng Colormap.....	64
3.3.2	Hiệu quả của việc tô màu ảnh X-quang vú .....	66
Chương 4	TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV .....	69
4.1.	TRÍCH ĐẶC TRUNG ẢNH UTV .....	69
4.1.1	Vai trò của đặc trưng ảnh UTV .....	69
4.1.2	Trích đặc trưng GLCM.....	70
4.2.	TRÍCH XUẤT VÙNG ROI CHO PHÂN LOẠI UTV .....	73
4.2.1	Phân đoạn ảnh UTV để trích xuất vùng ROI.....	73
4.2.2	Tô màu cho vùng ROI .....	78
4.3.	ỨNG DỤNG PHÂN LOẠI UTV DỰA TRÊN ĐẶC TRƯNG ẢNH.....	80
4.3.1	Phân loại ảnh UTV dựa vào đặc trưng độ tương phản .....	80
4.3.2	Phân loại ảnh UTV dựa trên đặc trưng GLCM và mô hình SVM .....	83
4.3.3	Phân loại ảnh UTV dựa trên vùng ROI.....	86
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	
	93	
5.1.	MẠNG HỌC SÂU TRONG PHÂN LOẠI BỆNH UTV .....	94
5.1.1	Vai trò mạng học sâu cho phân loại UTV .....	95

## MỤC LỤC

---

---

5.1.2	Mô hình mạng VGG16 .....	98
5.1.3	Mô hình mạng EfficientNet .....	102
5.1.4	Mô hình mạng Resnet-152.....	104
5.1.5	Mô hình mạng Resnet-201.....	105
5.2.	<b>PHÂN LOẠI ẢNH UTV THEO TỔN THƯƠNG DÙNG MẠNG HỌC SÂU</b> 106	
5.2.1	Bài toán phân loại có ung thư - không có ung thư.....	106
5.2.2	Bài toán phân loại các loại tổn thương dựa trên tập ảnh X-quang vú MIAS 110	
5.2.3	Bài toán phân loại các loại tổn thương dựa trên tập ảnh X-quang vú VinDr- Mammo118	
5.2.4	Kết quả phân loại bệnh UTV theo BIRADS .....	125
5.2.5	Hợp nhất bộ dữ liệu để phân loại tổn thương với EfficientNet kết hợp với Multihead Attention.....	135
5.3.	<b>HỆ THỐNG HỖ TRỢ CHẨN ĐOÁN TỪ XA DÙNG MẠNG HỌC SÂU.</b>	147
5.3.1	Các yêu cầu thiết kế.....	148
5.3.2	THIẾT KẾ HỆ THỐNG.....	149
Chương 6	<b>KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN.....</b>	160
6.1.	<b>KẾT LUẬN.....</b>	160
6.2.	<b>HƯỚNG PHÁT TRIỂN .....</b>	162
	<b>TÀI LIỆU THAM KHẢO.....</b>	163
	<b>CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ.....</b>	177

## DANH SÁCH CÁC CHỮ VIẾT TẮT

### DANH SÁCH CÁC CHỮ VIẾT TẮT

AI	Artificial Intelligence	TTNT
ANN	Artificial Neural Network	Mạng nơ-ron nhân tạo
API	Application Programming Interface	Giao diện lập trình ứng dụng
BCDR	Breast Cancer Digital Repository	Kho lưu trữ số về UTV
BIRADS	Breast Imaging Reporting and Data System	Hệ thống báo cáo và dữ liệu hình ảnh vú
BRCA1	BReast CAncer gene 1	Gen UTV 1
BRCA2	BReast CAncer gene 2	Gen UTV 2
BS	-	Bác sĩ
BUS	Breast Ultrasound	Siêu âm vú
BUSI	Breast Ultrasound Image	Hình ảnh siêu âm vú
BV	-	Bệnh viện
CAD	Computer Aided Diagnosis	Chẩn đoán hỗ trợ máy tính
CBIS-DDSM	Curated Breast Imaging Subset of DDSM	Phân nhóm hình ảnh vú được tuyển chọn của DDSM
CC	Craniocaudal	Từ đầu đến chân
CLAHE	Contrast-Limited Adaptive Histogram Equalization	Cân bằng Histogram thích ứng giới hạn độ tương phản
CNN	Convolutional Neural Network	Mạng nơ-ron tích chập
CROI	Color Region of Interest	Vùng màu quan tâm
CSDL	-	Cơ sở dữ liệu
CT	Computed Tomography	Chụp cắt lớp vi tính
CWT	Continuous Wavelet Transform	Biến đổi Wavelet liên tục
DBT	Digital Breast Tomosynthesis	Chụp cắt lớp vú kỹ thuật số
DDSM	Digital Database for Screening Mammography	CSDL số để sàng lọc chụp nhũ ảnh
DICOM	Digital Imaging and Communications in Medicine	Hình ảnh kỹ thuật số và truyền thông trong y học
DL	-	Dữ liệu
DWT	Discrete Wavelet Transform	Biến đổi Wavelet rời rạc
FADHECAL	Fuzzy Anisotropic Diffusion Histogram Equalization Contrast Adaptive Limited	Biểu đồ khuếch tán dị hướng mờ Cân bằng Tương phản Thích ứng Giới hạn

## DANH SÁCH CÁC CHỮ VIẾT TẮT

FC	Fully Connected Layer	Lớp được kết nối đầy đủ
GLCM	Gray Level Cooccurrence matrix	Ma trận đồng hiện mức xám
GROI	Gray Region of Interest	Vùng xám quan tâm
JPEG	Joint Photographic Experts Group	Nhóm chuyên gia nhiếp ảnh chung
KHDL	-	Khoa học dữ liệu
kNN	k Nearest Neighbors	k Hàng xóm gần nhất
LAPIMO	Laboratório de Análise e Processamento de Imagens Médicas e Odontológicas	Phòng thí nghiệm phân tích và xử lý hình ảnh Y khoa và Nha khoa
LDA	Linear Discriminant Analysis	Phân tích phân biệt tuyến tính
MIAS	Mammographic Image Analysis Society	Hội phân tích hình ảnh chụp nhũ ảnh
MLO	Mediolateral Oblique	Chéo giữa bên
MLP	Multi Layer Perceptron	Perceptron nhiều lớp
MRI	Magnetic Resonance Imaging	Chụp cộng hưởng từ
NC	-	Nghiên cứu
NN	Neural Networks	Mạng lưới nơ-ron
PCA	Principal Component Analysis	Phân tích thành phần chính
PC	Personal Computer	Máy tính cá nhân
PL	-	Phân loại
ReLU	Rectified Linear Unit	Đơn vị tuyến tính chỉnh lưu
ROI	Region of Interest	Khu vực quan tâm
SNR	Signal to Noise Ratio	Tỷ lệ tín hiệu trên nhiễu
SVM	Support Vector Machine	Máy hỗ trợ vector
TTNT	-	Trí tuệ nhân tạo
UTV	-	Ung thư vú
VGG	Visual Geometry Group	Nhóm hình học trực quan
VICTRE	Virtual Imaging Clinical Trial for Regulatory Evaluation	Thử nghiệm lâm sàng hình ảnh ảo để đánh giá theo quy định
WT	Wavelet Transform	Biến đổi Wavelet
XLA	-	Xử lý ảnh
XQV	-	X-quang vú

**DANH SÁCH CÁC BẢNG**

<b>BẢNG</b>	<b>TRANG</b>
Bảng 1.1 Các yếu tố làm tăng nguy cơ UTV [6] .....	1
Bảng 2.1 Tổng quan về các nghiên cứu AI trong chẩn đoán UTV.....	36
Bảng 3.1 Thông tin bộ dữ liệu MIAS.....	49
Bảng 3.2 Thông tin bộ dữ liệu CBIS-DDSM .....	51
Bảng 3.2 Thông Tin Bộ Dữ Liệu Vindr-Mammo .....	52
Bảng 3.3 Biểu diễn thông tin chi tiết của hình ảnh sau khi thu thập .....	54
Bảng 3.5 So sánh độ chính xác phân loại giữa ảnh được tô màu và ảnh xám .....	66
Bảng 4.1. Trích xuất đặc trưng của các trường hợp ung thư và bình thường .....	82
Bảng 4.2. Biểu diễn tập dữ liệu cho huấn luyện và kiểm thử của 2 loại ảnh UTV ..	83
Bảng 4.3. So sánh độ chính xác phân loại giữa ảnh được tô màu và ảnh xám .....	91
Bảng 5.1 Tóm tắt các phương pháp tiền xử lý và học sâu cho bài toán phân loại UTV .....	97
Bảng 5.2 Chi tiết kiến trúc của mạng EfficientNet-B0 .....	102
Bảng 5.3 Chi tiết bộ dữ liệu Mini-DDSM.....	107
Bảng 5.4 Chi tiết số lượng Train và Vaid.....	107
Bảng 5.5 Độ chính xác huấn luyện và kiểm tra của 2 mô hình .....	109
Bảng 5.6 Thông tin của tập dữ liệu CBIS-DDSM. ....	112
Bảng 5.6 Thông tin của tập dữ liệu MIAS. ....	112
Bảng 5.8. So sánh giữa các bộ ảnh khác nhau (có/không XLA).....	116
Bảng 5.9. So sánh với các kết quả nghiên cứu khác .....	117
Bảng 5.10. Tập dữ liệu VinDr-Mammo với bốn loại bệnh tương ứng .....	119
Bảng 5.11. Tập dữ liệu VinDr-Mammo với bốn loại bệnh tương ứng .....	124

## DANH SÁCH CÁC BẢNG

---

---

Bảng 5.12. Thống kê đánh giá mức độ tổn thương theo thang đo BIRADS của tập dữ liệu gốc .....	127
Bảng 5.13. Số lượng ảnh sau khi tăng cường của tập CBIS-DDSM.....	128
Bảng 5.14 Kết quả so sánh khi có áp dụng/ không có áp dụng các bước XLA .....	132
Bảng 5.15. Thống kê và đánh giá mức độ tổn thương của hình ảnh với BIRADS	134
Bảng 5.16 Các thông số yêu cầu hệ thống.....	149

## DANH SÁCH CÁC HÌNH

HÌNH	TRANG
Hình 1.1 Mô hình dự kiến hệ thống hỗ trợ chẩn đoán từ xa sử dụng TTNT. ....	8
Hình 2.1 Năm loại ung thư phổ biến nhất đối với phụ nữ, số liệu năm 2022 [1].....	17
Hình 2.2 Tỷ lệ mắc bệnh và tỷ lệ tử vong trên 100.000 người của 15 bệnh ung thư hàng đầu tại Việt Nam, số liệu năm 2022 [1].....	18
Hình 2.3 Một vài ảnh từ tập dữ liệu MIAS.....	22
Hình 2.4 Sơ đồ chuẩn bị dữ liệu của tập dữ liệu CBIS-DDSM.....	23
Hình 2.5 Mẫu dữ liệu của INBreast .....	23
Hình 2.6 Mẫu hình ảnh từ BCDR .....	24
Hình 2.7 Mẫu hình ảnh từ tập dữ liệu LAPIMO.....	25
Hình 2.8 Mẫu hình ảnh từ tập dữ liệu OPTIMAM .....	26
Hình 2.9 Sơ đồ phân đoạn khối u sử dụng đường biên mờ.....	29
Hình 2.10 Sơ đồ hệ thống tăng cường chất lượng hình ảnh đề xuất bởi [95] .....	33
Hình 2.11 Sơ đồ hệ thống chẩn đoán hỗ trợ máy tính (CAD) được đề xuất dựa trên học sâu để phát hiện, phân đoạn và phân loại khối u UTV từ ảnh X-quang kỹ thuật số đầu vào .....	35
Hình 2.12 Sự khác biệt giữa phương pháp máy học truyền thống (bên trái) và phương pháp học sâu (bên phải).....	43
Hình 3.1 Một vài ảnh từ tập dữ liệu MIAS.....	50
Hình 3.2 Sơ đồ chuẩn bị dữ liệu của tập dữ liệu CBIS-DDSM [66] .....	50
Hình 3.3. Một số ảnh của tập dữ liệu CBIS-DDSM .....	51
Hình 3.4. Một bộ ảnh mẫu của tập dữ liệu VinDr-Mammo: (a) Left CC; (b) Right CC; (c) Left MLO; (d) Right MLO.....	52

## DANH SÁCH CÁC HÌNH

---

Hình 3.5 Sơ đồ thu thập dữ liệu tại bệnh viện Ung Bướu TP.HCM.....	53
Hình 3.6 Một số ảnh từ tập dữ liệu thu tại bệnh viện Ung Bướu TP.HCM.....	54
Hình 3.7 Ví dụ một kết luận chẩn đoán của BS và ảnh X-quang vú tương ứng. ....	57
Hình 3.8 Sơ đồ khối của quá trình đánh giá ung thư .....	59
Hình 3.9 Biểu diễn hình ảnh gốc và hình ảnh nâng cao.....	61
Hình 3.10 Kết quả của việc thực hiện hiệu chỉnh gamma.....	62
Hình 3.11 Cân bằng biểu đồ sử dụng CLAHE .....	64
Hình 3.12 Kết quả colormap đối với 5 loại ảnh với các mức BIRADS (1,2,3,4,5) .	66
Hình 4.1 Thống kê các đặc trưng của ảnh có ung thư và không có ung thư.....	71
Hình 4.2 Quy trình trích xuất vùng ROI lớn nhất.....	74
Hình 4.3. 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.....	77
Hình 4.4. Mô tả kết quả chuẩn hóa và tăng cường hình ảnh ROI lớn nhất bằng cách sử dụng cân bằng lược đồ: (a1-a3) Ảnh gốc; (b1-b3) Hình ảnh được chuẩn hóa; (c) Ảnh có cân bằng lược đồ.....	78
Hình 4.5. Kết quả to 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.....	80
Hình 4.6. Lược đồ xám của ảnh bình thường: (a) Ảnh vú bình thường; (b) Lược đồ xám của ảnh .....	81
Hình 4.7. Lược đồ xám của ảnh UTV: (a) Ảnh vUTV; (b) Lược đồ xám của ảnh ..	81
Hình 4.8. 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.....	84

## DANH SÁCH CÁC HÌNH

---

---

Hình 4.9. Các đặc trưng của ảnh bình thường và ung thư: (a) Contrast; (b) Energy; (c) Homogeneity; (c) Mean .....	85
Hình 4.10. Ma trận nhầm lẫn để đánh giá 50 hình ảnh thử nghiệm mỗi loại.....	86
Hình 4.11. Sơ đồ tổng thể của phương pháp phân loại ảnh UTV dựa trên vùng ROI .....	87
Hình 4.12. 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. ....	87
Hình 4.13. Minh họa quá trình XLA: (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 GROI sau hậu xử lý.....	88
Hình 4.14. Minh họa năm loại ảnh với GROI và nền đen: (a) bình thường; (b) bất đối xứng cục bộ; (c) biến dạng cấu trúc; (d) vôi hóa; (e) khối u.....	89
Hình 4.15. Minh họa năm loại tổn thương khác nhau của ảnh với CROI: (a) Không phát hiện; (b) Bất đối xứng cục bộ; (c) Biến dạng cấu trúc; (d) Vôi hóa; (e) Khối u.....	89
Hình 4.16. Ma trận nhầm lẫn để đánh giá hiệu suất phân loại đối với các tập ảnh xám. ....	90
Hình 4.17. Kết quả Ma trận nhầm lẫn để đánh giá hiệu suất phân loại đối với các tập ảnh có tô màu.....	91
Hình 5.1. Kiến trúc của mạng VGG16 đã tinh chỉnh.....	98
Hình 5.2. Lớp FC cuối cùng đã được sửa đổi cho đầu ra thứ 4.....	100
Hình 5.3. Cấu trúc của mạng EfficientNet-B7 được tinh chỉnh .....	103
Hình 5.4. Cấu Trúc Của Resnet .....	104
Hình 5.5 Cấu trúc của mạng DenseNet-201 .....	106
Hình 5.6. Thay đổi lớp FC = 3 .....	108

## DANH SÁCH CÁC HÌNH

---

Hình 5.7 Hàm mất mát huấn luyện và kiểm tra tương ứng của mô hình ResNet-152 và DenseNet-201.....	109
Hình 5.8 Ma trận nhầm lẫn của kết quả bài toán phân loại 3 ngõ ra: bình thường, lành tính, và ác tính.....	110
Hình 5.9 Sơ đồ khối phân loại và quy trình: (1) hai tập dữ liệu 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.....	111
Hình 5.10. Mô tả quy trình đào tạo với hơn 350 epoch cho Adam với tốc độ học tập 0,0001: (a) đường cong độ chính xác để đào tạo và xác nhận; (b) đường cong tổn thất cho việc đào tạo và xác nhận .....	114
Hình 5.11. Mô tả hiệu quả phân loại của hai loại tập dữ liệu với Vô hình ảnh lạnh tính (BC), Khối lành tính (BM), Vô hình ảnh ác tính (MC), Khối lượng ác tính (MM). (a) Bộ dữ liệu CBIS-DDSM; (b) Bộ dữ liệu MIAS.....	115
Hình 5.12. Đường cong tỷ lệ chính xác của huấn luyện và kiểm tra sử dụng mô hình trên.....	116
Hình 5.13. Đề xuất hệ thống phân loại các loại tổn thương vú dựa vào tập ảnh X-quang vú VinDr-Mammo .....	120
Hình 5.14. Minh họa ảnh X-quang vú của bệnh nhân ở hai tư thế khác nhau và được chụp theo 2 hướng.....	121
Hình 5.15. Kết quả ảnh X-quang vú sau XLA và đưa về kích thước 512x512.....	122
Hình 5.16. Kết quả trình bày đồ thị hàm mất mát và độ chính xác của hệ thống mạng học sâu đã được đề xuất. ....	123
Hình 5.17. Ma trận nhầm lẫn với số lượng ảnh trong mỗi lớp được tăng cường lên 10.000.....	124
Hình 5.18. Sơ đồ khối quá trình xử lý của hệ thống phân loại ảnh X-quang vú theo mức độ tổn thương theo thang BIRADS.....	126

## DANH SÁCH CÁC HÌNH

---

---

Hình 5.19. Bộ 4 ảnh trong một kết quả khám X-quang vú của tập ảnh VinDr-Mammo .....	127
Hình 5.20. Ví dụ lật ảnh .....	129
Hình 5.21. Ví dụ xoay ảnh.....	130
Hình 5.22. Các Đường cong độ chính xác và mất mát của mô hình .....	131
Hình 5.23. Ma trận nhầm lẫn .....	132
Hình 5.24 Kiến trúc của một hệ thống phân loại mới.....	136
Hình 5.25 Biểu diễn việc tạo ra một hình ảnh ngược được tô màu. ....	137
Hình 5.26 Kiến trúc kết hợp EfficientNet-B7 để tạo ra vector đặc trưng của MLO và CC.....	139
Hình 5.27 Cấu trúc của mạng Multihead Attention.....	141
Hình 5.28 Biểu diễn xử lý hình ảnh: (a) Hình ảnh gốc; (b) Mặt nạ đa ROI; (c) Mặt nạ ROI lớn nhất; (d) ROI lớn nhất RGB (e) ROI được tô màu và (1) Bình thường; (2) Khối lượng; (3) Vôի hóa; (4) Bất đối xứng cục bộ.....	143
Hình 5.29 Biểu diễn ma trận nhầm lẫn cho bốn loại tổn thương vú không có Multihead Attention. ....	144
Hình 5.30 Biểu diễn ma trận nhầm lẫn cho bốn loại tổn thương vú với Multihead Attention .....	144
Hình 5.31 Minh họa hệ thống y tế từ xa để phân loại hình ảnh UTV sử dụng mô hình AI.....	150
Hình 5.32. Mô hình hệ thống máy chủ lưu trữ hình ảnh (PACS).....	151
Hình 5.33. Sơ đồ các giao diện phần mềm chính của hệ thống.....	152
Hình 5.34. Sơ đồ kết nối Backend và Bộ xử lý AI .....	153
Hình 5.35. Mô hình AI trong hệ thống.....	154
Hình 5.36. Mô hình giao thức DICOM.....	155

## Chương 1

# GIỚI THIỆU TỔNG QUAN

### 1.1. TÍNH CẤP THIẾT CỦA ĐỀ TÀI

UTV là bệnh dẫn tới tử vong cao nhất đối với phụ nữ trên toàn thế giới, với số lượng ca bệnh ngày càng gia tăng, cả trong nhóm người trẻ tuổi. Theo những thống kê về bệnh ung thư gân đây [1], tại Việt Nam mỗi năm có khoảng 22.000 trường hợp được phát hiện mắc UTV và có hơn 9.000 trường hợp tử vong, chiếm 25,8% tổng số trường hợp bệnh ung thư ở nữ giới. Hơn nữa, theo số liệu của BV Bạch Mai tại Việt Nam, gần 70% bệnh nhân UTV được chẩn đoán và phát hiện ở giai đoạn muộn (giai đoạn 3-4), giảm đáng kể khả năng chữa trị thành công. Nguyên nhân chính có thể là do thiếu hệ thống sàng lọc quy mô lớn và nguồn lực chuyên môn cao còn hạn chế. Hơn nữa, BV Bạch Mai tiếp nhận hơn 5.000 bệnh nhân mỗi ngày, gây áp lực lớn cho các BS chẩn đoán hình ảnh [2].

Nguyên nhân gây ra bệnh UTV rất đa dạng và khó xác định, gồm những yếu tố như tuổi tác, di truyền, hormone, lối sống và có thể cả yếu tố môi trường sống. Những yếu tố khác ví dụ như tuổi sinh sản và cho con bú cũng có tác động tới tỷ lệ nguy cơ UTV [3]-[8]. Về mặt y khoa, UTV có thể có nguyên nhân từ các tế bào phát triển bất thường trong mô tuyến vú. Những tế bào này có khả năng xâm lấn vào các mô lân cận hoặc di căn tới các cơ quan khác. UTV được phân chia thành nhiều dạng khác nhau, trong đó UTV loại biểu mô xâm lấn là phổ biến và nguy hiểm nhất. May mắn thay, nếu được phát hiện và chẩn đoán từ giai đoạn sớm, bệnh UTV có khả năng được chữa trị hiệu quả và có tỷ lệ sống cao [9]-[11].

**Bảng 1.1 Những yếu tố làm tăng nguy cơ UTV [6]**

Nhóm yếu tố	Yếu tố có khả năng làm tăng nguy cơ UTV
Nội tiết tố và sinh sản	Tuổi của kỳ kinh nguyệt đầu tiên sớm Tuổi của kỳ kinh nguyệt cuối cùng muộn Lần đầu tiên có thai ở độ tuổi muộn (sau 30 tuổi) Không có thai

## Chương 1 GIỚI THIỆU TỔNG QUAN

	Tình trạng sau mãn kinh Sử dụng thuốc tránh thai Các liệu pháp hormone thay thế
Yếu tố sinh lý và tình trạng sức khỏe	Lớn tuổi (nguy cơ tăng từ 35 tuổi) Tiền sử trong gia đình có trường hợp bệnh UTV UTV, buồng trứng và nội mạc tử cung quá khứ Xuất hiện những thay đổi lành tính ở vú, dẫn đến tăng sản không điển hình Tiếp xúc bức xạ ion, ví dụ như liệu pháp điều trị ung thư hạch Hodgkin Tăng trưởng nhanh ở tuổi thiếu niên và tăng trưởng cao ở tuổi trưởng thành Nhiễm vi-rút gây ung thư (ví dụ: Epstein-Barr)
Dinh dưỡng	Chế độ ăn kiêng kiểu phương Tây Ăn quá nhiều chất béo (mỡ động vật) Tiêu thụ nhiều thịt chiên và thịt đỏ Lượng sắt cao Sự phát triển thừa cân/béo phì sau mãn kinh Tiêu thụ ít rau và trái cây tươi Lượng phytoestrogen hấp thụ thấp
Những yếu tố khác có liên quan đến lối sống	Thường xuyên uống rượu Hoạt động thể chất kém Làm việc ban đêm

Theo thống kê, trên toàn thế giới có hàng triệu phụ nữ mắc bệnh UTV hàng năm [3]. Điều này cho thấy sự quan trọng trong vấn đề nâng cao nhận thức về căn bệnh này, cũng như tầm soát và điều trị sớm. Ở Việt Nam, hàng năm hơn 22.000 ca mắc bệnh với hơn 9.000 trường hợp tử vong do bệnh UTV [1]. Trong khi, trên toàn cầu, con số ca tử vong năm 2022 là 670.000 [12]. Tại Mỹ, trong năm 2024 có 313.510 ca UTV được phát hiện trong tổng số 2.001.140 ca ung thư và trong số ca UTV có 42.780 ca tiên lượng tử vong [3]. Rõ ràng UTV cho thấy rất nguy hiểm, do vậy việc chẩn đoán được thực hiện sớm sẽ cho khả năng trị khỏi là rất cao. Cụ thể, nếu chẩn đoán và phát hiện UTV ở giai đoạn đầu sẽ đạt tỷ lệ chữa khỏi lên tới 80%, còn giai đoạn 2 thì chiếm khoảng 60% khả năng chữa khỏi [11].

Tỷ lệ UTV ở phụ nữ đang có chiều hướng gia tăng [1], khiến việc phát hiện sớm và xác định loại ung thư để điều trị trở nên hết sức quan trọng. Chẩn đoán sớm giúp quá trình điều trị được thực hiện hiệu quả và dễ dàng hơn, nâng cao tỷ lệ chữa khỏi.

## Chương 1 GIỚI THIỆU TỔNG QUAN

---

Bên cạnh việc chú ý đến các biểu hiện trong cuộc sống, tầm soát ung thư thường xuyên là yếu tố rất quan trọng để phát hiện sớm bệnh. Tuy nhiên, do nhiều lý do, nhiều người không thực hiện tầm soát định kỳ 1-2 lần mỗi năm. Hơn nữa, việc tầm soát ban đầu có thể không phát hiện ra ung thư do hạn chế về thiết bị, công nghệ hoặc các nguyên nhân khác. Khi bệnh có biểu hiện rõ ràng để BS chẩn đoán, thường đã ở giai đoạn muộn, gây khó khăn cho điều trị và giảm khả năng chữa khỏi. Do đó, ứng dụng công nghệ, đặc biệt là TTNT, đang đóng vai trò trọng yếu trong việc hỗ trợ BS phát hiện sớm ung thư, đồng thời giúp đưa ra các khuyến nghị y khoa để điều trị kịp thời. Hiện nay, các phương pháp chẩn đoán hình ảnh thường được áp dụng trong khám sàng lọc UTV gồm chụp X-quang vú, chụp cắt lớp, siêu âm và chụp cộng hưởng từ.

Hiện nay, X-quang vú vẫn đang là phương pháp khám bệnh để tầm soát UTV hiệu quả và phổ biến nhất. Để hỗ trợ BS chẩn đoán dựa trên hình ảnh X-quang, các phương pháp XLA UTV đóng vai trò quan trọng [13]. Cụ thể, các phương pháp lọc ảnh, tăng cường ảnh giúp làm rõ hình ảnh, và phương pháp phân đoạn dò biên giúp tách biệt các vật thể nghi là khối u để phân tích đánh giá.

Trong vài năm gần đây, TTNT đã được ứng dụng vào các hệ thống chẩn đoán từ xa được coi là một bước đột phá trong các ứng dụng y tế [14]. Đặc biệt, các nghiên cứu gần đây từ Đức, Hàn Quốc, Thụy Điển và một số quốc gia khác đã chứng minh AI không chỉ gia tăng độ chính xác trong phát hiện sớm UTV lên đến 20% mà còn giảm 44% tải công việc cho BS [15]. Các BV lớn của Việt Nam, bao gồm BV Bạch Mai và BV Đại học Y Thái Bình, đã triển khai thành công các công cụ AI như Cadai-BTM để diễn giải hình ảnh siêu âm. Các công cụ này không chỉ giúp cho quá trình chẩn đoán tự động mà còn hỗ trợ BS chẩn đoán ảnh X-quang hoặc siêu âm vú tốt hơn [2], [16].

Các phương pháp chẩn đoán truyền thống chủ yếu dựa trên hình ảnh chụp nhũ ảnh hoặc siêu âm [17], và điều này đòi hỏi BS phải có kỹ năng cao mới có thể phát hiện các tổn thương nhỏ như vi vôi hóa hoặc khối u có đường kính dưới 1 cm [11]. Tuy nhiên, sự chủ quan của BS trong việc đọc hình ảnh và làm việc quá tải đôi khi

## Chương 1 GIỚI THIỆU TỔNG QUAN

---

có thể dẫn đến sai sót trong chẩn đoán hình ảnh UTV. Đặc biệt, một nghiên cứu từ Đại học Lund ở Thụy Điển chỉ ra rằng khoảng 15% đến 30% các trường hợp UTV đã bị bỏ sót trong quá trình sàng lọc thường quy [18]. Điều này rõ ràng là rất đáng tiếc và nguy hiểm cho nhiều bệnh nhân.

Khối lượng công việc và áp lực gia tăng hàng ngày đối với các BS tại BV có thể dẫn đến việc chẩn đoán không chính xác. Do đó, một hệ thống chẩn đoán từ xa tích hợp AI là một giải pháp để gia tăng độ chính xác và cũng giảm bớt áp lực về công việc cho các BS. Điều này có nghĩa là việc tích hợp AI vào hệ thống y tế từ xa đang định hình lại cách tiếp cận UTV trên toàn cầu. Các nghiên cứu gần đây đã chỉ ra rằng AI không chỉ cải thiện độ chính xác của chẩn đoán mà còn cung cấp kết quả chẩn đoán lên đến 94,5% so với 88% của chẩn đoán truyền thống [19]. Hơn nữa, việc sử dụng hệ thống y tế từ xa dựa trên AI giúp giảm số lần khám bệnh khoảng 40% so với khám trực tiếp. Tại Việt Nam, nơi tỷ lệ phát hiện muộn UTV lên tới 70% [1], hệ thống y tế từ xa này hứa hẹn sẽ thu hẹp khoảng cách tiếp cận các dịch vụ chăm sóc sức khỏe chất lượng.

Ngoài ra, AI còn được sử dụng để hỗ trợ các BS trong việc đưa ra quyết định chẩn đoán và điều trị. Nghiên cứu gần đây đã chỉ ra rằng mạng thần kinh tích chập (CNN) có khả năng phù hợp với khả năng phát hiện của một BS X quang có kinh nghiệm [20]. Các hệ thống AI có thể phân tích dữ liệu bệnh nhân từ nhiều nguồn khác nhau, bao gồm hình ảnh y tế, hồ sơ bệnh án, và kết quả xét nghiệm, để đưa ra những gợi ý và dự đoán chính xác.

### ❖ CÁC KẾT QUẢ ĐÃ ĐƯỢC CÔNG BỐ TRONG VÀ NGOÀI NƯỚC

#### *a. Trong nước*

Một số nhà nghiên cứu đã dựa trên ảnh X-quang và sử dụng các phương pháp XLA nhằm làm rõ những khối u trong ảnh vú. Với sự phát triển của khoa học dữ liệu và công nghệ AI được áp dụng khá phổ biến trong nhiều lĩnh vực dựa vào dữ liệu lớn, thì việc phân loại và nhận biết ảnh UTV sử dụng công nghệ AI là rất khả thi.

## Chương 1 GIỚI THIỆU TỔNG QUAN

---

Nhiều nghiên cứu đã được thực hiện tại Việt Nam ứng dụng AI trong chẩn đoán UTV cho thấy những kết quả khả quan. Có nhiều nghiên cứu tiêu biểu sử dụng AI để phân loại ảnh X-quang vú và phát hiện UTV với độ chính xác cao. Ví dụ, nghiên cứu được thực hiện tại Đại học Y Dược TP.HCM [5] đạt độ chính xác 87,5% trong phân loại UTV ác tính và lành tính, trong khi nghiên cứu của Viện Công nghệ Thông tin - ĐHQG-HN sử dụng mạng nơ-ron tích chập đạt độ chính xác lên tới 92,1% trong việc phân loại 4 loại tổn thương vú [6].

Tuy nhiên, việc ứng dụng AI trong chẩn đoán UTV tại Việt Nam vẫn còn nhiều hạn chế. Một trong những thách thức lớn nhất là thiếu dữ liệu huấn luyện mô hình AI. Việc thu thập dữ liệu y tế, đặc biệt là dữ liệu hình ảnh y tế, tại Việt Nam còn nhiều khó khăn do các vấn đề về bảo mật thông tin và đạo đức nghiên cứu.

Bên cạnh đó, hạn chế về nguồn nhân lực chuyên môn cao trong lĩnh vực AI và y tế cũng là một rào cản lớn [7]. Việc ứng dụng AI vào thực tế cũng cần sự phối hợp chặt chẽ giữa các nhà khoa học, BS và các cơ quan quản lý để đảm bảo hiệu quả và tính ứng dụng cao. Mặc dù còn nhiều thách thức, ứng dụng AI chẩn đoán UTV tại Việt Nam vẫn có rất nhiều tiềm năng phát triển to lớn. AI có thể góp phần quan trọng trong nâng cao hiệu quả chẩn đoán, giúp phát hiện sớm UTV, từ đó cải thiện tỷ lệ sống cho bệnh nhân.

### ***b. Ngoài nước***

Trong những năm gần đây, có rất nhiều nghiên cứu về TTNT ứng dụng trong y tế, đặc biệt là việc ứng dụng công nghệ AI để nhận dạng và phân loại các bệnh lý như tim [8], da [9], ung thư não [10], UTV [11], v.v. Trong các nghiên cứu này, AI được sử dụng để nhận dạng UTV và đánh giá mức độ ung thư dựa trên hình ảnh X-quang vú. Bên cạnh đó, việc sử dụng AI để phân loại và nhận dạng ung thư, các loại ung thư để điều trị sớm dựa trên hình ảnh chụp bởi các công nghệ khác nhau như CT scanner, MRI [12] và X-quang [13] đóng vai trò rất quan trọng.

Đối với vấn đề nhận dạng và phân loại UTV, các BS giàu kinh nghiệm sẽ cung cấp các tập dữ liệu lớn hình ảnh X-quang tuyến vú cùng với thông tin liên quan đến

## Chương 1 GIỚI THIỆU TỔNG QUAN

---

bệnh UTV để xây dựng hệ thống học máy. Đặc biệt, dữ liệu báo cáo X-quang vú theo Hệ thống dữ liệu và xếp hạng hình ảnh tuyến vú BIRADS với 7 mức độ từ BIRADS 0 đến BIRADS 6 được sử dụng cùng với các hệ thống học máy để đánh giá. Nhờ việc đánh giá này, các xét nghiệm tiếp theo được sử dụng để hỗ trợ BS đưa ra kết luận liên quan đến UTV [14].

XLA đóng một vai trò to lớn trong việc chẩn đoán UTV bằng hình ảnh, cả trong vai trò giúp làm rõ và nổi bật những vùng đáng lưu ý trên ảnh để giúp BS có thể chẩn đoán tốt hơn, cũng như để tạo ra các ảnh ngõ vào cho các hệ thống phân loại ảnh UTV dựa trên các bộ phân loại truyền thống cũng như các mạng CNN. Các phương pháp XLA thường được sử dụng là các bộ lọc nhiễu, tăng cường hình ảnh và điều chỉnh hoặc cân bằng mức xám [21].

Việc ứng dụng các mạng Nơ-ron hoặc các bộ phân loại khác để hỗ trợ việc chẩn đoán bệnh UTV đã được triển khai khá rộng rãi trong nhiều nghiên cứu gần đây [22]. Ảnh X-quang vú sau xử lý có thể được trích xuất các đặc trưng, sau đó đưa vào các bộ phân loại truyền thống để giúp xác định có hoặc không có tổn thương, và lành tính hay ác tính. Các bộ phân loại như SVM, kNN, có thể được sử dụng cho việc này. Tuy nhiên phương pháp này thường cho độ chính xác không cao quá 90% và gặp nhiều khó khăn trong việc chọn các đặc trưng thích hợp cho việc phân loại.

Hiện nay, có nhiều nghiên cứu sử dụng mạng CNN để chẩn đoán các tổn thương vú, trong đó việc phát hiện UTV rất quan trọng [23]. Ngoài ra, CNN được ứng dụng để phân loại khối u lành tính và ác tính trên hình ảnh kỹ thuật số với độ chính xác lên đến 98% [24]. Một mạng CNN khác, được gọi là mạng nơ-ron chập cải tiến để phân loại UTV (CNNI-BCC), được xây dựng vào năm 2019 và được sử dụng để phân loại 3 trường hợp tổn thương như lành tính, ác tính và bình thường với độ chính xác lần lượt là 89,5%, 90,5% và 90,7% [25].

Trong những năm gần đây, nhiều mô hình học sâu để phân tích hình ảnh đã được phát triển, có thể kể đến YOLO, ResNet-50 và các mô hình khác đã được áp dụng để nhận dạng các tổn thương khác nhau. Một nghiên cứu cho thấy YOLOv6 kết hợp với

## Chương 1 GIỚI THIỆU TỔNG QUAN

---

Federated Learning (FedL) sử dụng các tập dữ liệu BreakHis và BUSI đã đạt được độ chính xác 98% trong phân loại mô lành tính và ác tính [26]. Đặc biệt, việc học liên kết này để đào tạo mô hình trên nhiều máy chủ mà không chia sẻ dữ liệu bệnh nhân do vấn đề bảo mật thông tin. Trong khi đó, MammoScreen với giải pháp AI được FDA chấp thuận đã áp dụng phương pháp tiếp cận từ trên xuống để đánh giá toàn bộ ảnh chụp nhũ ảnh và phương pháp tiếp cận từ dưới lên để phân tích từng vùng nghi ngờ, giúp giảm 44% khối lượng công việc của BS [15].

Các hệ thống chăm sóc sức khỏe gần đây đã được phát triển để hỗ trợ BS trong quá trình chẩn đoán và điều trị y tế [27], [28] và AI là một trong những công nghệ được phát triển trong một số hệ thống này. Google Health đang phát triển hệ thống Med-PaLM 2, một Mô hình Ngôn ngữ Lớn (LLM) dành riêng cho y tế để phân tích kết quả xét nghiệm, hình ảnh và sau đó đề xuất các phác đồ điều trị. Tại Ấn Độ, một thử nghiệm hệ thống IBM Watson đã đạt được sự đồng thuận 90% với quyết định của một hội đồng chuyên gia về điều trị UTV giai đoạn đầu [29]. Các hệ thống này sử dụng học tăng cường để liên tục cập nhật từ dữ liệu lâm sàng mới nhằm tối ưu hóa các khuyến nghị điều trị cá nhân hóa. Cụ thể, hệ thống hỗ trợ AI của Google Health đã giảm 5,7% kết quả dương tính giả và 9,4% kết quả âm tính giả so với các phương pháp truyền thống. Một nghiên cứu khác cho thấy, với việc phân tích 44.755 hình ảnh siêu âm và AUROC 0,976, hệ thống đã phát hiện các tổn thương có kích thước rất nhỏ từ 0,5mm đến 2mm mà mắt người dễ bỏ sót [30].

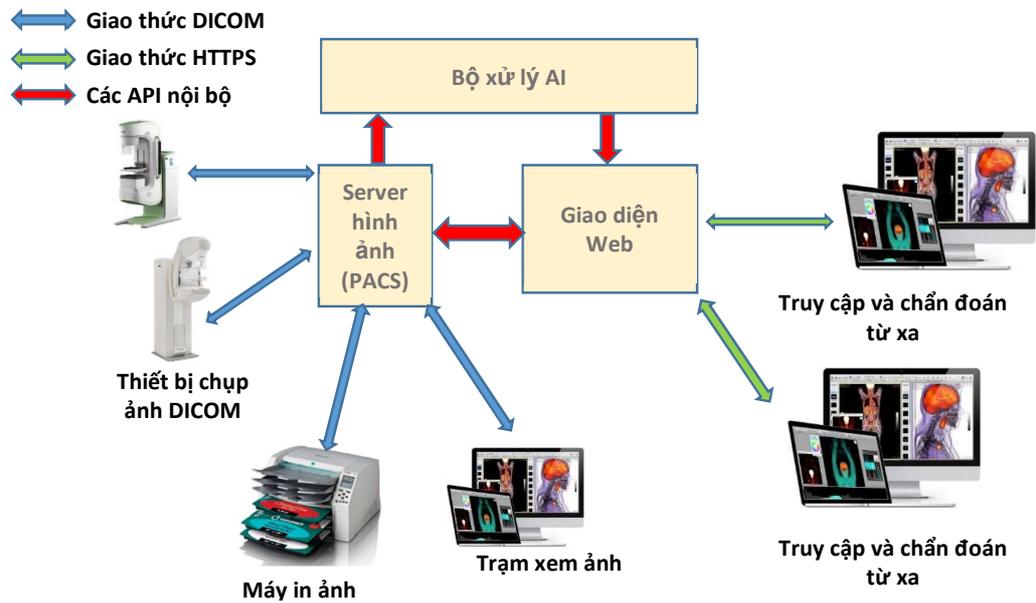
### ❖ GIỚI THIỆU CHUNG VỀ ĐỀ TÀI

Hình ảnh X-quang của các bệnh liên quan đến tổn thương vú thường có chất lượng không đồng nhất và chứa nhiều nhiễu loạn không cần thiết. Do đó, trong hầu hết các nghiên cứu, các thuật toán tiền xử lý hình ảnh được áp dụng để cải thiện chất lượng hình ảnh đồng nhất và nâng cao hiệu quả phân loại. Tuy nhiên, việc tiền xử lý hình ảnh trước khi áp dụng để đào tạo các mạng học sâu tiên tiến như VGG16 [31], EfficientNet [32] hoặc các mạng khác có thể giúp gia tăng hiệu quả của việc phân loại. Trong nghiên cứu này, các tập ảnh UTV từ các nguồn mở và tập ảnh từ BV ở Việt Nam sẽ được sử dụng để xây dựng một hệ thống chẩn đoán bệnh UTV sử dụng

## Chương 1 GIỚI THIỆU TỔNG QUAN

---

mạng học sâu. Từ tập ảnh ban đầu, các bước tiền xử lý như tăng cường dữ liệu, lọc nhiễu ảnh, tăng cường hình ảnh và cân bằng mức xám sẽ được áp dụng. Việc phân đoạn hình ảnh để tìm vùng ROI cũng như trích xuất đặc trưng ảnh cũng sẽ được nghiên cứu trước khi đưa vào các mạng học sâu để đào tạo và phân loại các tổn thương ví.



**Hình 1.1** Mô hình dự kiến hệ thống hỗ trợ chẩn đoán từ xa sử dụng TTNT.

Hình 1.1 mô tả sơ đồ nguyên tắc của một hệ thống chẩn đoán từ xa hỗ trợ bằng AI mà đề tài muốn xây dựng. Một hệ thống như vậy bao gồm thành phần trung tâm là một Server lưu trữ dữ liệu, giao diện web và bộ xử lý các thuật toán AI.

Việc xây dựng một hệ thống chẩn đoán từ xa dựa vào các kết quả đạt được từ việc huấn luyện mạng học sâu cũng sẽ được xem xét. Một hệ thống như vậy sẽ cho phép các BS từ các BV ở xa tải lên hình ảnh dưới dạng JPEG hoặc DICOM lên một server thông qua mạng Internet để được phân loại và cho ra các kết quả đánh giá về loại thương tổn để giúp các BS có thể chẩn đoán một cách hiệu quả hơn. Các BS chuyên gia từ các BV tuyến trên cũng có thể thông qua đó để hỗ trợ một cách hiệu quả cho các BV ở vùng xa, hệ thống như vậy cũng có thể ứng dụng vào việc đào tạo và huấn luyện.

## 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 TTNT được huấn luyện thông qua các tập ảnh X-quang lớn khác nhau đã được xử lý để chuẩn hóa để hỗ trợ BS 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 XLA 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 TTNT 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 BS 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 sỹ đưa ra kết quả chẩn đoán.

## 1.3. ĐỐI TƯỢNG VÀ PHẠM VI NGHIÊN CỨU

### 1.3.1 Đối tượng nghiên cứu

Để thực hiện đề tài nghiên cứu, những đối tượng nghiên cứu có thể được sử dụng như sau:

- Tập lớn hình ảnh X-quang vú từ nhiều nguồn và bao gồm nhiều loại tổn thương khác nhau của các bệnh nhân sử dụng trong chẩn đoán bệnh UTV.

## Chương 1 GIỚI THIỆU TỔNG QUAN

---

- Những thông tin đa dạng của dữ liệu khác nhau của nhiều bệnh nhân như di truyền, tiền sử bệnh lý và những thông tin khác có liên quan đến bệnh UTV.
- Các thuật toán trong XLA và các mô hình mạng AI để phân loại ảnh UTV.
- Hệ thống hỗ trợ chẩn đoán từ xa bệnh UTV sử dụng AI

### 1.3.2 Phạm vi nghiên cứu

- Nghiên cứu phát hiện các dấu hiệu đặc trưng của của ảnh X-quang vú để có thể phân loại thành các nhóm có ung thư và không có ung thư và có thể đưa ra các dự đoán về loại tổn thương của UTV.
- Thử nghiệm trên nhóm mẫu với các ca bệnh tại các BV Việt Nam.

## 1.4. CÁC TIẾP CẬN VÀ PHƯƠNG PHÁP NGHIÊN CỨU

### 1.4.1 Cách thức tiếp cận

Tìm hiểu tổng quan các nghiên cứu trên Thế giới và tại Việt Nam, thu thập tư liệu, dữ liệu cần thiết, lập quy trình mô phỏng, xây dựng thuật toán, thí nghiệm, điều tra, phân tích, đối chiếu và so sánh các khối dữ liệu.

### 1.4.2 Phương pháp nghiên cứu

Dựa trên những nghiên cứu hiện nay, đề tài kiến nghị một số phương pháp có tính khả thi như sau:

- Phương pháp nghiên cứu định tính, định lượng, toán thống kê
- Phương pháp thu thập dữ liệu từ các nguồn đáng tin cậy, bảo đảm tính chính xác và tính hợp pháp của dữ liệu.
- Phương pháp quan sát khoa học, cụ thể quan sát quy trình mô phỏng, dữ liệu và các kết quả thử nghiệm.
- Phương pháp thực nghiệm khoa học như thử nghiệm nhiều phương pháp XLA và phân đoạn khác nhau để tìm ra giải pháp cho kết quả tốt nhất.
- Phương pháp phân tích và tổng kết kinh nghiệm, Đây là cách nghiên cứu lại những gì đã thực hiện trong quá khứ để rút ra các bài học quý báu.

## Chương 1 GIỚI THIỆU TỔNG QUAN

---

- Phương pháp phân tích và tổng hợp lý thuyết tập trung vào việc nghiên cứu các tài liệu và lý luận hiện có, phân tích chúng thành từng bộ phận để tìm hiểu sâu sắc về đối tượng và liên kết từng mặt.

- Phương pháp nhận biết, phân loại và hệ thống hóa lý thuyết: giúp sắp xếp và tổ chức kiến thức khoa học một cách có hệ thống. Mục tiêu là phân loại các tài liệu theo từng chủ đề, đặc điểm chung hoặc hướng phát triển, sau đó hệ thống hóa chúng thành một mô hình lý thuyết mạch lạc.

- Phương pháp giả thuyết: dự đoán về quy luật của đối tượng sau đó tìm cách chứng minh dự đoán đó là chính xác, ví dụ như dự đoán các đặc trưng ảnh sẽ đóng vai trò quan trọng trong việc giúp phân loại ảnh UTV, từ đó thử nghiệm và mô phỏng để kiểm chứng.

- Phương pháp thống kê đánh giá các kết quả thu được.

Giải pháp dự kiến:

- Giải pháp cho mục tiêu 1: sử dụng các nguồn dữ liệu ảnh X-quang vú mở và hợp pháp, đáng tin cậy, đã được sử dụng trong nghiên cứu khoa học, kết hợp với các nguồn dữ liệu ảnh thu được hợp pháp từ các BV để tạo ra một CSDL các tập ảnh X-quang vú với nhiều loại thương tổn và có số lượng ảnh đủ lớn phục vụ nghiên cứu. Thông qua quá trình chọn lọc để loại bỏ các ảnh không phù hợp và quá trình XLA để tạo ra tập dữ liệu ảnh đồng đều về chất lượng và kích thước.

- Giải pháp cho mục tiêu 2: nghiên cứu các đặc trưng của nhũ ảnh, các phương pháp trích đặc trưng và áp dụng các thuật toán phân loại như SVM hoặc kNN để thử nghiệm phân loại ảnh UTV để có thể hiểu rõ vai trò của các đặc trưng ảnh cho phân biệt các loại tổn thương của ảnh X-quang vú. Trên cơ sở những hiểu biết đó xây dựng mô hình mạng học sâu và huấn luyện mô hình này với tập ảnh đã thu thập và được chuẩn hóa để đạt được độ chính xác cần thiết trong phân loại ảnh UTV.

- Giải pháp cho mục tiêu 3: Hệ thống phân loại ảnh UTV từ xa sử dụng AI với mô hình mạng từ kết quả thu được ở trên cho phép các BS dễ dàng truy cập và tải lên

các hình ảnh sử dụng giao thức DICOM hoặc ảnh JPEG và được hệ thống phân loại tự động, từ đó giúp BS trong việc chẩn đoán bệnh UTV.

### **1.5. ĐÓNG GÓP MỚI VỀ KHOA HỌC VÀ Ý NGHĨA THỰC TIỄN CỦA LUẬN ÁN**

Luận án đã đóng góp một cách toàn diện vào lĩnh vực ứng dụng trí tuệ nhân tạo trong chẩn đoán bệnh ung thư vú thông qua việc xây dựng một quy trình xử lý và phân tích ảnh X-quang vú hoàn chỉnh từ dữ liệu đến mô hình phân loại. Trước hết, luận án đã phát triển một quy trình tiền xử lý ảnh có hệ thống, bao gồm các bước tăng cường chất lượng ảnh, điều chỉnh sáng – tương phản, hiệu chỉnh gamma, cân bằng lược đồ và chuyển đổi ảnh xám sang ảnh màu bằng colormap. Quy trình này giúp chuẩn hóa ảnh từ nhiều nguồn dữ liệu khác nhau, tạo ra một bộ dữ liệu đồng nhất và đạt chất lượng cao, đóng vai trò nền tảng quan trọng để tối ưu hóa hiệu quả của các mô hình học sâu trong những bước nghiên cứu tiếp theo.

Một đóng góp đáng kể khác của luận án là việc đề xuất phương pháp trích xuất vùng quan tâm (ROI) hiệu quả, kết hợp giữa phân đoạn dựa trên ngưỡng Otsu, loại bỏ nhiễu và chuẩn hóa kích thước. Đặc biệt, luận án chứng minh rằng việc tô màu ROI (CROI) giúp mô hình học sâu phân biệt các tổn thương rõ ràng hơn, làm tăng độ chính xác từ 5–8% so với ROI xám. Đây là hướng tiếp cận mới mẻ và có ý nghĩa thực tiễn cao đối với các mô hình AI trong y tế, vốn thường gặp khó khăn khi xử lý ảnh X-quang có độ tương phản thấp và cấu trúc phức tạp.

Ngoài việc xử lý dữ liệu, luận án còn thực hiện phân tích sâu các đặc trưng hình ảnh, bao gồm các đặc trưng texture từ GLCM và histogram mức xám. Từ đó, tác giả xây dựng và đánh giá các mô hình học máy như SVM và kNN nhằm xác định vai trò của từng loại đặc trưng trong phân loại ung thư vú. Kết quả cho thấy rằng việc kết hợp đặc trưng texture và vùng ROI giúp cải thiện rõ rệt hiệu suất phân loại, tạo tiền đề để tinh chỉnh các mạng học sâu ở những chương sau.

Đóng góp nổi bật nhất của luận án nằm ở việc tinh chỉnh và phát triển các mô hình học sâu như VGG16, ResNet-152/201, DenseNet-201 và EfficientNet-B0/B7.

## Chương 1 GIỚI THIỆU TỔNG QUAN

---

Nhờ kết hợp chặt chẽ giữa tiền xử lý ảnh, trích ROI và huấn luyện sâu, các mô hình này đạt độ chính xác cao, khoảng 95% trong phân loại tổn thương và 92.6% trong phân loại mức độ BIRADS. Luận án cũng đề xuất kiến trúc phân loại đa ảnh (multi-view), kết hợp đặc trưng từ hai tư thế chụp (CC và MLO) và sử dụng cơ chế Multi-head Attention để mô hình tập trung vào vùng tổn thương quan trọng. Cách tiếp cận này mô phỏng quy trình đọc ảnh thực tế của bác sĩ X-quang, từ đó nâng cao tính hiệu quả và độ tin cậy của mô hình.

Không chỉ dừng lại ở mô hình thuật toán, luận án còn triển khai một hệ thống chẩn đoán từ xa sử dụng AI, bao gồm PACS, backend xử lý AI, API nội bộ và giao diện web cho bác sĩ. Hệ thống hỗ trợ tải ảnh DICOM/JPEG từ xa, tự động phân loại và trả kết quả dựa trên mô hình đã huấn luyện. Đây là một đóng góp có ý nghĩa thực tiễn rất lớn, đặc biệt đối với các bệnh viện tuyến dưới hoặc các khu vực thiếu bác sĩ chuyên khoa, góp phần thu hẹp khoảng cách y tế và nâng cao khả năng phát hiện sớm ung thư vú.

Cuối cùng, luận án có giá trị đặc biệt khi sử dụng dữ liệu thực từ Bệnh viện Ung Bướu TP.HCM để xây dựng và đánh giá mô hình. Việc phát triển mô hình AI phù hợp với dữ liệu Việt Nam giúp tăng khả năng ứng dụng thực tế và làm chủ công nghệ trong nước, thay vì phụ thuộc vào dữ liệu quốc tế. Điều này không chỉ nâng cao giá trị khoa học mà còn góp phần thúc đẩy phát triển công nghệ AI trong y tế Việt Nam theo hướng bền vững và hiệu quả.

### 1.6. CẤU TRÚC CỦA LUẬN ÁN

Với những kết quả nghiên cứu trong và ngoài nước được tóm tắt như ở trên, cấu trúc dự kiến của luận án bao gồm 6 chương, được mô tả chi tiết như sau:

Chương 1 - Giới thiệu tổng quan: nêu rõ tính cấp thiết của việc ứng dụng AI trong chẩn đoán ung thư vú (UTV), khi tỷ lệ mắc và tử vong ngày càng cao, đặc biệt ở Việt Nam với ~22.000 ca mắc mới mỗi năm. Các phương pháp chẩn đoán truyền thống còn hạn chế do sự phụ thuộc vào chuyên môn bác sĩ và áp lực công việc lớn. Nghiên cứu trong nước và quốc tế cho thấy AI có tiềm năng nâng cao độ chính xác, giảm tải

## Chương 1 GIỚI THIỆU TỔNG QUAN

---

cho bác sĩ và hỗ trợ phát hiện sớm. Chương cũng trình bày mục tiêu nghiên cứu, đối tượng - phạm vi, phương pháp tiếp cận, và đóng góp mới, gồm: xây dựng hệ thống hỗ trợ chẩn đoán UTV từ xa dựa trên mạng học sâu và quy trình XLA chuẩn hóa dữ liệu.

Chương 2 - Cơ sở lý thuyết: Chương này tổng hợp kiến thức nền tảng về UTV, các loại tổn thương và vai trò chẩn đoán hình ảnh. Trình bày các tập dữ liệu nhũ ảnh nổi tiếng như MIAS, CBIS-DDSM, INBreast, OPTIMAM... cùng đặc thù của từng bộ. Chương cũng hệ thống các phương pháp phân đoạn, tăng cường ảnh, phát hiện ROI, và tổng quan các thuật toán học máy/học sâu (SVM, CNN, VGG, ResNet...) áp dụng trong phân loại UTV. Ngoài ra, các mô hình chẩn đoán từ xa sử dụng AI cũng được giới thiệu, làm nền cho hệ thống đề xuất.

Chương 3 - Thu thập và tiền xử lý ảnh X-quang vú: Tập trung mô tả việc thu thập dữ liệu từ MIAS, CBIS-DDSM, VinDr-Mammo và đặc biệt là ảnh X-quang vú thu từ Bệnh viện Ung Bướu TP.HCM. Một quy trình chuẩn hóa dữ liệu được xây dựng gồm: tăng cường ảnh (điều chỉnh sáng - tương phản, gamma, CLAHE), loại nhiễu, chuyển đổi màu qua colormap. Các bước này giúp ảnh trở nên đồng nhất và làm nổi bật các dấu hiệu tổn thương, tạo tiền đề cho việc phân loại tốt hơn trong các chương sau.

Chương 4 - Trích đặc trưng và xử lý vùng ROI: Chương 4 phân tích vai trò của đặc trưng ảnh trong phân loại UTV. Sử dụng GLCM để trích đặc trưng kết cấu (contrast, homogeneity, energy...). Tiếp theo, chương trình bày phương pháp phân đoạn bằng ngưỡng Otsu và trích xuất vùng tổn thương ROI (GROI, CROI). Quy trình chuẩn hóa ROI, tô màu ROI và loại bỏ thành phần dư thừa được mô tả chi tiết. Chương cũng đánh giá thử nghiệm phân loại dựa trên đặc trưng và SVM, cho thấy mức cải thiện khi dùng ROI so với ảnh gốc.

Chương 5 - Phân loại UTV bằng mạng học sâu & hệ thống chẩn đoán từ xa: Đây là chương trọng tâm, áp dụng các mô hình học sâu gồm VGG16, EfficientNet, ResNet-152/201, DenseNet-201 để phân loại ảnh X-quang vú theo nhiều dạng bài

## Chương 1 GIỚI THIỆU TỔNG QUAN

---

toán: có/không ung thư, phân loại 4 loại tổn thương, và phân loại mức độ BIRADS. Các mô hình được tinh chỉnh (fine-tuning), kết hợp tăng cường dữ liệu và xử lý ROI, đạt độ chính xác lên đến 95% cho phân loại tổn thương, 92.6% cho phân loại theo BIRADS.

Trong chương này còn bao gồm nội dung xây dựng hệ thống chẩn đoán từ xa gồm PACS, backend AI, giao diện web, tích hợp xử lý DICOM. Bác sĩ có thể tải ảnh từ xa để hệ thống tự động phân loại.

Chương 6 - Kết luận và hướng phát triển: Luận án kết luận rằng việc kết hợp tiền xử lý XLA, trích ROI và mạng học sâu giúp cải thiện hiệu suất phân loại UTV một cách rõ rệt. Đồng thời khẳng định tính khả thi của hệ thống chẩn đoán từ xa dùng AI trong thực tiễn. Hướng phát triển gồm: mở rộng dữ liệu từ nhiều bệnh viện, cải thiện mô hình đa đầu vào (multi-view), ứng dụng các kỹ thuật attention mới và triển khai thử nghiệm lâm sàng trên diện rộng.

## Chương 2

# CƠ SỞ LÝ THUYẾT

Với sự phát triển không ngừng của các kỹ thuật TTNT, triển vọng ứng dụng của TTNT đã mở rộng sang các lĩnh vực y tế khác nhau, trong đó có các lĩnh vực như chẩn đoán trong ống nghiệm, phục hồi chức năng thông minh, hình ảnh y tế và tiên lượng bệnh. UTV là căn bệnh ác tính phổ biến, ảnh hưởng nghiêm trọng đến sức khỏe thể chất và tinh thần của phụ nữ. Sàng lọc UTV sớm—thông qua chụp nhũ ảnh, siêu âm hoặc chụp cộng hưởng từ (MRI)—có thể cải thiện đáng kể tiên lượng cho bệnh nhân UTV. Các ứng dụng AI đã cho thấy hiệu suất tuyệt vời trong các nhiệm vụ nhận dạng hình ảnh khác nhau và việc sử dụng chúng trong sàng lọc UTV đã được khám phá trong nhiều nghiên cứu khác nhau trong thời gian gần đây. Phần cơ sở lý thuyết sẽ giới thiệu các kỹ thuật AI có liên quan và ứng dụng của chúng trong lĩnh vực chẩn đoán hình ảnh y tế X-quang vú, cụ thể là trong việc xác định, phân chia và phân loại các tổn thương; đánh giá nguy cơ UTV; và cải thiện chất lượng hình ảnh. Tập trung vào hình ảnh y tế cho bệnh UTV, chương này cũng xem xét những thách thức và triển vọng liên quan đến AI.

### 2.1. TỔNG QUAN VỀ BỆNH UTV

#### 2.1.1 Bệnh UTV

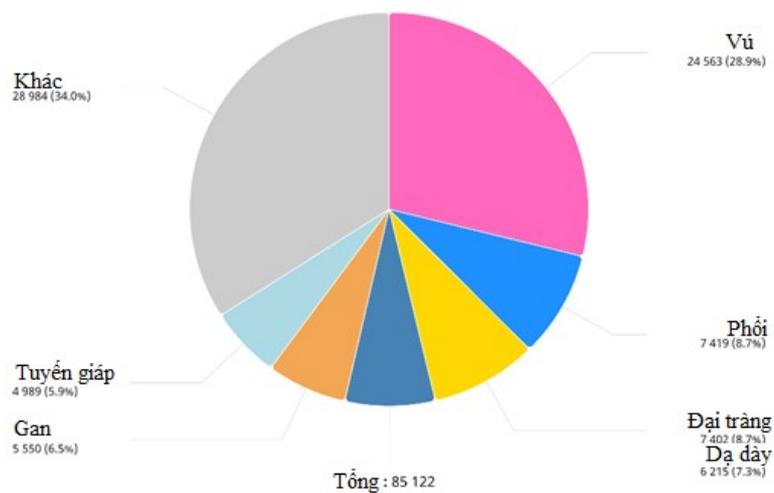
UTV là một bệnh lý ác tính phát triển từ các tế bào vú. Bệnh thường gặp nhất ở phụ nữ, nhưng nam giới cũng có thể mắc. Giai đoạn UTV dựa trên kích thước của khối u, mức độ lan rộng của ung thư sang các hạch bạch huyết và các bộ phận khác của cơ thể. Giai đoạn UTV từ 0 đến IV, với giai đoạn 0 là sớm nhất và giai đoạn IV là tiến triển nhất [6]. Nguyên nhân gây UTV vẫn chưa được xác định rõ [33]. Tuy nhiên, một số yếu tố chính làm gia tăng nguy cơ đã được xác định, bao gồm: giới tính, trong đó nữ giới có nguy cơ mắc UTV cao hơn nam giới; tuổi tác, nguy cơ mắc UTV tăng theo độ tuổi; lịch sử gia đình, UTV sẽ làm tăng nguy cơ mắc bệnh liên quan đến người thân; di truyền: có thể một số đột biến gen, chẳng hạn như BRCA1

## Chương 2 CƠ SỞ LÝ THUYẾT

---

và BRCA2 [34] có khả năng làm tăng nguy cơ bệnh UTV và ung thư buồng trứng; lối sống: có thể liên quan đến một số vấn đề như béo phì, uống rượu bia, ít hoạt động thể chất và hút thuốc lá.

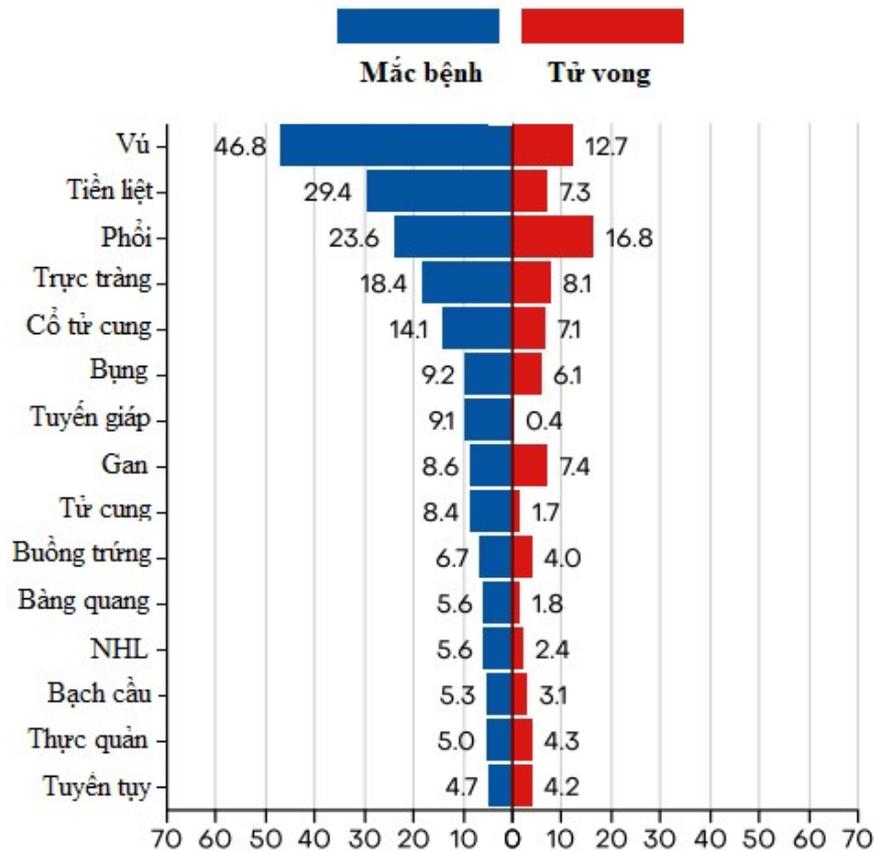
UTV là loại ung thư được chẩn đoán phổ biến nhất ở 154 quốc gia và là nguyên nhân chính gây tử vong liên quan đến ung thư ở hơn 100 quốc gia [35]. Ở Việt Nam, năm 2024 có khoảng 10.008 phụ nữ chết vì UTV trong số 24.563 người mắc bệnh [1].



**Hình 2.1 Năm loại ung thư phổ biến nhất đối với phụ nữ, số liệu năm 2022**

[1]

Hình 2.1 cho thấy bệnh UTV chiếm tỷ lệ cao nhất trong các bệnh ung thư cho phụ nữ với 24.56%, cao gấp 3 lần so với bệnh ung thư đứng tiếp theo là ung thư phổi. Trong khi Hình 2.2 biểu diễn tỷ lệ mắc bệnh và tỷ lệ tử vong của 10 loại bệnh ung thư hàng đầu tại Việt Nam, trong đó UTV xếp đầu tiên, tiếp theo là ung thư phổi và bệnh ung thư gan, nhưng tỷ lệ gây ra tử vong do UTV thấp hơn so với hai loại ung thư kia.



**Hình 2.2 Tỷ lệ mắc bệnh và tỷ lệ tử vong trên 100.000 người của 15 bệnh ung thư hàng đầu tại Việt Nam, số liệu năm 2022 [1]**

UTV có thể được phân loại là ung thư biểu mô lành tính, bình thường, tại chỗ hoặc ung thư biểu mô xâm lấn [36]. Một khối u lành tính sẽ gây ra những thay đổi nhỏ về giải phẫu vú; tuy nhiên, không có chất gây hại nên không thể coi là khối u nguy hiểm [37]. Mặt khác, ung thư biểu mô tại chỗ chỉ ảnh hưởng đến các tiểu thùy ống vú và không lan sang các mô khác [38]. Chẩn đoán UTV thường bao gồm các phương pháp: khám vú lâm sàng, chụp X-quang vú, siêu âm vú, và sinh thiết vú [37], [39]. Trong đó X-quang vú là một kỹ thuật chẩn đoán hình ảnh sử dụng tia X năng lượng thấp để tạo ra hình ảnh chi tiết của mô vú. Đây là một công cụ quan trọng trong việc phát hiện sớm UTV, giúp tăng khả năng điều trị thành công và cải thiện tỷ lệ sống sót cho bệnh nhân [40]. Việc chụp X-quang vú đóng vai trò thiết yếu trong chẩn

đoán UTV vì có nhiều ưu điểm so với các phương pháp khác: Chụp X-quang vú có thể phát hiện UTV ở giai đoạn rất sớm, khi khối u còn nhỏ và chưa lan rộng sang các bộ phận khác của cơ thể. Việc phát hiện sớm UTV giúp tăng khả năng điều trị thành công với các phương pháp ít xâm lấn hơn, chẳng hạn như phẫu thuật cắt bỏ khối u vú và xạ trị. Nhờ khả năng phát hiện sớm, chụp X-quang vú góp phần cải thiện đáng kể tỷ lệ sống sót cho bệnh nhân UTV. Theo thống kê, phụ nữ được chẩn đoán UTV ở giai đoạn 0 (giai đoạn sớm nhất) có tỷ lệ sống sót sau 5 năm lên đến 99%, cao hơn nhiều so với những người được chẩn đoán ở giai đoạn muộn hơn. Bên cạnh đó, chụp X-quang vú là một kỹ thuật an toàn và không xâm lấn, phù hợp với hầu hết phụ nữ từ 40 tuổi trở lên [41].

### 2.1.2 Chẩn đoán bệnh UTV và ứng dụng TTNT trong chẩn đoán

Đối với các Bác sĩ để có thể tìm ra bệnh ung thư sớm nhất và chẩn đoán đúng bệnh có thể thì họ thường tập trung vào các phương pháp áp dụng trong chương trình tầm soát UTV bao gồm khám lâm sàng tuyến vú và chẩn đoán hình ảnh bệnh lý UTV. Cụ thể, các phương pháp chẩn đoán hình ảnh phổ biến có thể được áp dụng cho chẩn đoán UTV gồm chụp nhũ ảnh, siêu âm, chụp cắt lớp và chụp cộng hưởng từ [41].

Các tiến bộ đáng kể trong các chức năng tính toán liên quan đến sự gia tăng của dữ liệu lớn trong 5 thập kỷ qua đã thúc đẩy ứng dụng AI vào các lĩnh vực mới [42]. Ví dụ: AI có sẵn để nhận dạng khuôn mặt và giọng nói, cùng với các công nghệ mới khác. Việc sử dụng các phương pháp AI trong hình ảnh y tế ngày càng thu hút được sự quan tâm nghiên cứu. Đã đạt được tiến bộ đáng kể trong việc áp dụng thuật toán AI, đặc biệt là thuật toán học sâu (DL), vào các nhiệm vụ nhận dạng hình ảnh. Sự sẵn có của nhiều phương pháp khác nhau trong lĩnh vực phân tích hình ảnh y tế, từ mạng thần kinh tích chập (CNN) đến bộ mã hóa tự động đa dạng, đã góp phần vào sự phát triển nhanh chóng trong hình ảnh y tế [43]. Nhiều loại ung thư khác nhau đã có thể được chẩn đoán bằng cách sử dụng công cụ AI và máy học (ML) [44]-[46].

Các BS X quang đọc và phân tích hình ảnh vú và sử dụng chúng để chẩn đoán. Tuy nhiên, khối lượng công việc lớn và thời gian làm việc dài có thể gây mệt mỏi,

## Chương 2 CƠ SỞ LÝ THUYẾT

---

dẫn đến đánh giá sai và chẩn đoán sai hoặc chẩn đoán sai. Sử dụng AI, hoặc trong trường hợp này là chẩn đoán y tế với sự hỗ trợ của máy tính (CAD), cho mục đích này có thể giảm bớt các lỗi tiềm ẩn của con người. Trong hệ thống CAD, thuật toán phù hợp sẽ hoàn tất quá trình xử lý và phân tích hình ảnh [47], [48]. Những đột phá gần đây với DL trong AI, đặc biệt là CNN, đã nâng cao đáng kể lĩnh vực hình ảnh y tế [49], [50]. AI đề cập đến khả năng của các ứng dụng hoặc máy móc có thể tái tạo con người (hoặc chức năng não người) để học hỏi và giải quyết vấn đề [51]. Công nghệ AI đã đạt được tiến bộ đáng kể kể từ đó, đặc biệt là trong thập kỷ qua. Khi AI trở thành một phần quan trọng trong lĩnh vực khoa học máy tính, đã có những nỗ lực liên tục để tạo ra các loại máy thông minh mới có thể bắt chước bộ não con người, mở rộng sang các ứng dụng đa dạng như nhận dạng hình ảnh, khai thác dữ liệu, hệ thống chuyên gia, xử lý ngôn ngữ tự nhiên, ngôn ngữ. nhận dạng, nhận dạng mẫu và robot. Trong lĩnh vực y tế, AI có thể được sử dụng để dự đoán bệnh sớm, sàng lọc bệnh, hỗ trợ quyết định lâm sàng, quản lý sức khỏe, quản lý BV, hình ảnh y tế và phân tích hồ sơ bệnh án hoặc tài liệu. Ngoài việc hỗ trợ BS đưa ra chẩn đoán chính xác, AI còn có thể được sử dụng để phân tích hình ảnh và thông tin y tế nhằm sàng lọc và dự đoán bệnh. Tập trung vào ứng dụng AI trong chụp ảnh vú, Al-antari et al. báo cáo độ chính xác cao cho một hệ thống CAD tích hợp hoàn chỉnh (>92%) trong việc phát hiện, phân đoạn và phân loại các khối nhìn thấy trên ảnh chụp quang tuyến vú [52]. Dựa trên 2654 bài kiểm tra và bài đọc của 101 BS X quang, Rodriguez-Ruiz et al. báo cáo đã giảm đáng kể khối lượng công việc của BS X quang (17%) khi sử dụng hệ thống AI được đào tạo với tính năng chọn trước tự động (với điểm AI là 2 làm ngưỡng khi chấm điểm khả năng mắc bệnh ung thư từ 1 đến 10) [53].

Trong những năm gần đây, đã có rất nhiều nghiên cứu về TTNT trong lĩnh vực y tế, đặc biệt là ứng dụng công nghệ AI trong việc nhận dạng và phân loại các bệnh như tim [54][55], da người [56], ung thư não [57], UTV [58] và các bệnh khác. Một phương pháp áp dụng trên mạng CNN cho tập ảnh MRI của CSDL Brain Tumor Segmentation Challenge 2013 (BRATS 2013). Trong các nghiên cứu này, AI đã được ứng dụng để nhận dạng UTV và đánh giá mức độ ung thư dựa trên hình ảnh

chụp nhũ ảnh. Ngoài ra, việc sử dụng AI để phân loại và nhận dạng ung thư, việc điều trị sớm ung thư dựa trên hình ảnh chụp từ các công nghệ khác nhau như máy quét CT, MRI [59], X-quang [60] là rất quan trọng. Trong bài toán xác định và phân loại UTV, các bộ ảnh chụp nhũ ảnh lớn nhất chứa thông tin liên quan đến tổn thương vú được cung cấp bởi các BS giàu kinh nghiệm để xây dựng hệ thống học máy. Cụ thể, các báo cáo dữ liệu chụp nhũ ảnh như BIRADS với 7 điểm từ BIRADS-0 đến BIRADS-6 được sử dụng với các hệ thống học máy để đánh giá. Với đánh giá này, việc kiểm tra có thể được sử dụng để giúp các BS kết luận chính xác hơn liên quan đến UTV [61].

Hiện nay, có nhiều nghiên cứu sử dụng Mạng nơ-ron tích chập (CNN) để phân loại tổn thương vú, trong đó việc phát hiện UTV là rất quan trọng [62]. Ngoài ra, CNN còn được ứng dụng để phân loại khối u lành tính hoặc ác tính trên ảnh kỹ thuật số với độ chính xác lên đến 91%. Một CNN khác, được gọi là mạng nơ-ron tích chập cải tiến để phân loại ung thư vú (CNNI-BCC), được xây dựng vào năm 2019 và được ứng dụng để phân loại 3 trường hợp tổn thương là lành tính, ác tính và bình thường với độ chính xác lần lượt là 89,5%, 90,5% và 90,7% [25].

### **2.2. CÁC TẬP DỮ LIỆU HÌNH ẢNH VỀ BỆNH UTV**

Phần này trình bày một số bộ dữ liệu công khai phổ biến nhất trong cả hình ảnh chụp quang tuyến vú và siêu âm. Các nhà nghiên cứu thường xuyên sử dụng các kỹ thuật học máy để phân tích và rút ra những hiểu biết sâu sắc từ các bộ dữ liệu này, nhằm tạo điều kiện thuận lợi cho những tiến bộ trong lĩnh vực hình ảnh y tế. Những bộ dữ liệu này góp phần chung vào sự tiến bộ của kỹ thuật máy học trong chụp nhũ ảnh, thúc đẩy sự đổi mới trong phát hiện và chẩn đoán UTV.

#### **2.2.1 Bộ dữ liệu chụp nhũ ảnh nguồn mở**

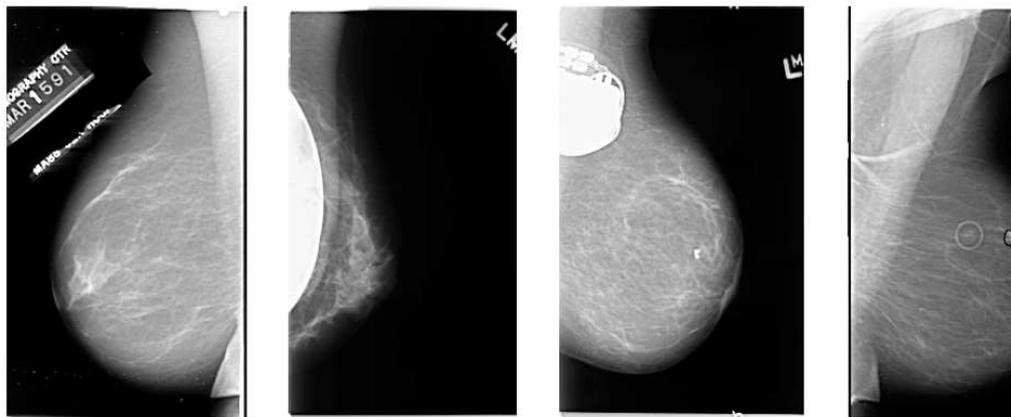
Trong lĩnh vực nghiên cứu chụp nhũ ảnh, một số bộ dữ liệu có sẵn công khai đóng vai trò là nguồn tài nguyên quan trọng để thúc đẩy các kỹ thuật học máy. Các bộ dữ liệu này cung cấp thông tin có chú thích và hình ảnh đa dạng, cho phép các nhà

## Chương 2 CƠ SỞ LÝ THUYẾT

---

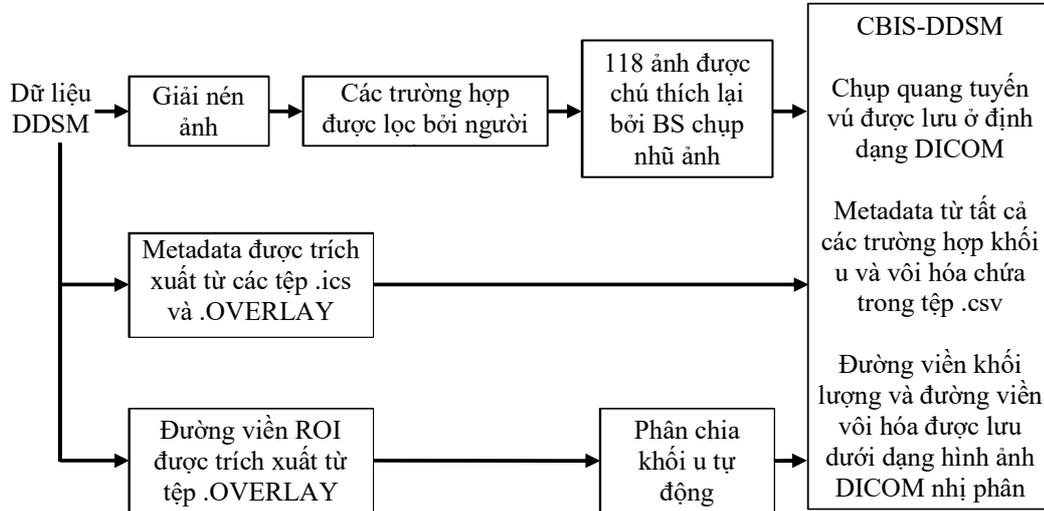
ngiên cứu phát triển và đánh giá thuật toán một cách hiệu quả. Ở đây, chúng tôi nêu bật một số bộ dữ liệu chụp nhũ ảnh công cộng nổi bật:

MIAS (Hiệp hội phân tích hình ảnh nhũ ảnh) [63]: Tập dữ liệu này bao gồm hơn 300 ảnh chụp X quang tuyến vú được chú thích với thông tin về loại mô nền, những bất thường xuất hiện ở vú và mức độ nghiêm trọng của chúng. Ngoài ra, các tổn thương trong chụp quang tuyến vú được đánh dấu bằng tọa độ X và Y, cùng với nhãn cho các bất thường khác nhau [64]. Hình 2.3 trình bày mẫu hình ảnh từ tập MIAS.



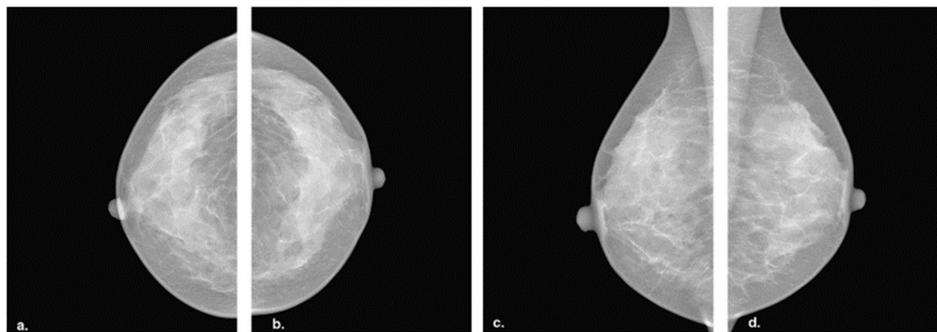
**Hình 2.3 Một vài ảnh từ tập dữ liệu MIAS**

DDSM (CSDL kỹ thuật số về chụp nhũ ảnh sàng lọc) [65]: Với hơn 2600 nghiên cứu về chụp nhũ ảnh bằng phim được quét, DDSM cung cấp một kho lưu trữ toàn diện. Một tập hợp con của DDSM, được gọi là CBIS-DDSM [66], cung cấp các hình ảnh được chú thích rõ ràng với thông tin bệnh lý chi tiết như loại khối u vú, cấp độ khối u và giai đoạn. Hình 2.4 cho thấy quy trình sàng lọc của tập CBIS-DDSM.



**Hình 2.4 Sơ đồ chuẩn bị dữ liệu của tập dữ liệu CBIS-DDSM**

CSDL INBreast [67]: Bộ dữ liệu INbreast là nguồn tài nguyên có giá trị cho các nhà nghiên cứu về UTV, cung cấp một bộ sưu tập lớn các hình ảnh chụp quang tuyến vú kỹ thuật số chất lượng cao, với hơn 410 ảnh X-quang vú đã được sàng lọc. Bộ dữ liệu này cung cấp nhiều trường hợp khác nhau, bao gồm cả các ca lành tính và ác tính. Thông tin đi kèm bao gồm thông tin về các loại bất thường và dữ liệu đường viền khối u.



(a) Ảnh CC - Right; (b) Ảnh CC - Left;  
(c) Ảnh MLO - Right; (d) Ảnh MLO - Left.

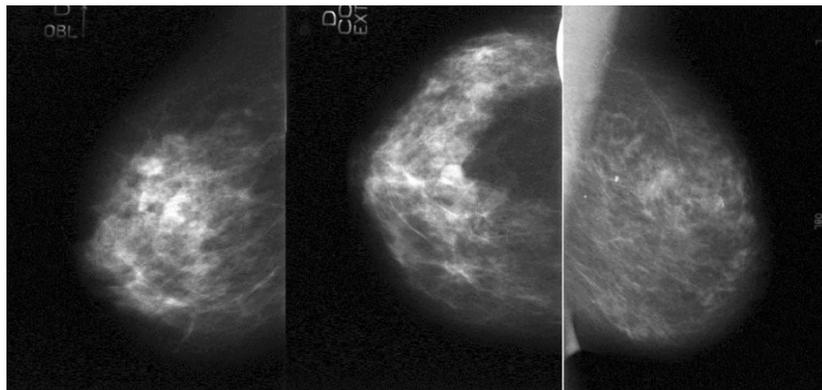
**Hình 2.5 Mẫu dữ liệu của INBreast**

## Chương 2 CƠ SỞ LÝ THUYẾT

---

Hình 2.5 giới thiệu một số ảnh của tập INBreast. Đây là tập ảnh chụp từ máy chụp X-Quang vú kỹ thuật số chứ không phải là ảnh số hóa từ phim nên có chất lượng tốt hơn hẳn tập MIAS và DDSM.

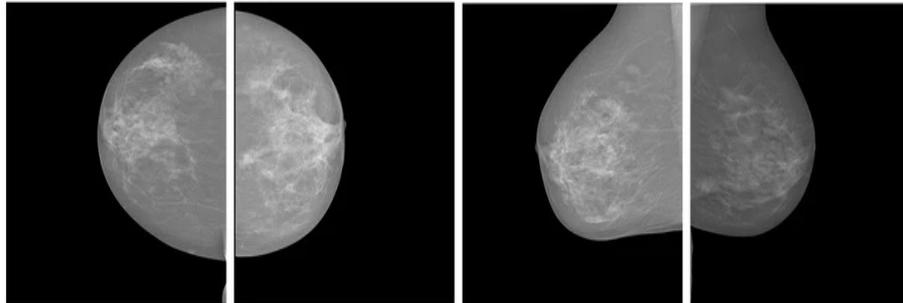
BCDR (Kho lưu trữ kỹ thuật số về UTV) [68]: Bộ dữ liệu này tập trung vào chụp quang tuyến vú kỹ thuật số toàn trường (FFDM) và khuyến khích sự đóng góp từ các nhà nghiên cứu [69]. Các chi tiết cụ thể của bộ dữ liệu khác nhau phụ thuộc vào phiên bản hoặc bản cập nhật, nhưng nhìn chung nó chứa một lượng lớn hình ảnh, bao gồm cả các trường hợp lành tính và ác tính. Nó cung cấp hình ảnh ở cả chế độ chụp Trên - dưới (chụp từ trên xuống dưới theo hướng đầu-chân) và Chéo - giữa bên (chụp từ trên xuống dưới, chệch 45-60 độ theo hướng trong-ngoài).



**Hình 2.6 Mẫu hình ảnh từ BCDR**

Tập BCDR có nguồn gốc ảnh khá đa dạng nên chất lượng ảnh không đồng nhất, như thể hiện trên Hình 2.6, tập này bao gồm cả ảnh phim được số hóa và ảnh chụp bằng thiết bị X-quang số.

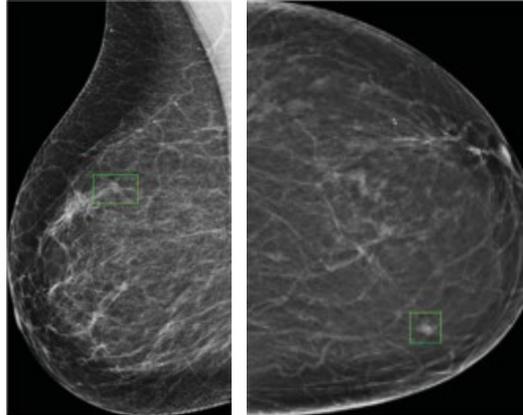
BancoWeb LAPIMO [70]: Bao gồm hơn 1400 ảnh chụp X quang tuyến vú, bộ dữ liệu này cung cấp hình ảnh ở định dạng TIFF được thu thập từ 320 đối tượng. Bộ dữ liệu này liên quan đến dự án BancoWeb LAPIMO, một nỗ lực hợp tác có sự tham gia của các nhà nghiên cứu từ nhiều tổ chức khác nhau nhằm thúc đẩy nghiên cứu về UTV bằng cách sử dụng phương pháp chụp nhũ ảnh kỹ thuật số. Nó bao gồm nhiều loại tổn thương được phân loại là lành tính, ác tính hoặc khỏe mạnh.



**Hình 2.7** Mẫu hình ảnh từ tập dữ liệu LAPIMO

Tập dữ liệu thử nghiệm VICTRE (Virtual Imaging Clinical Trial for Regulatory Evaluation) [71]: Không giống như những tập dữ liệu khác, tập dữ liệu này hoàn toàn tổng hợp, mô phỏng 2986 đối tượng với kích thước và mật độ vú đại diện. Tập dữ liệu này được tạo ra từ dự án VICTRE, một bộ công cụ cho phép tạo hình ảnh X-quang vú mô phỏng (in-silico) và cung cấp các công cụ phần mềm nguồn mở để phân tích và chèn tổn thương [72].

Bộ dữ liệu OPTIMAM [71]: CSDL hình ảnh chụp nhũ ảnh OPTIMAM (OMI-DB) được tạo ra để cung cấp một tập dữ liệu tập trung, được chú thích đầy đủ cho mục đích nghiên cứu. Lý do ban đầu để tạo ra CSDL này là dành cho các dự án OPTIMAM (2008-2013) và OPTIMAM2 (2013-2018) do Cancer Research United Kingdom tài trợ [72], nhằm đánh giá cách các yếu tố khác nhau ảnh hưởng đến việc phát hiện UTV trên phim chụp nhũ ảnh. Các hình ảnh được lấy từ các trung tâm sàng lọc tại Vương quốc Anh và kết hợp với dữ liệu được thu thập một cách có hệ thống. Tại Vương quốc Anh, Chương trình sàng lọc vú của Dịch vụ Y tế Quốc gia (NHSBSP) mời phụ nữ tham gia sàng lọc vú 3 năm một lần trong độ tuổi từ 50 đến 70. OPTIMAM chứa hình ảnh DBT 3D có chú thích và cung cấp gói Python nguồn mở để dễ dàng tích hợp vào các hệ thống nghiên cứu, tạo điều kiện xử lý liền mạch. Hình 2.8 trình bày các ảnh mẫu từ tập OPTIMAM.



**Hình 2.8 Mẫu hình ảnh từ tập dữ liệu OPTIMAM**

Trên đây liệt kê các tập dữ liệu về ảnh UTV phổ biến, năm phát hành, quốc gia và đặc điểm chính của chúng, cũng như số lượng nghiên cứu (gần đúng) sử dụng các tập dữ liệu đó. Từ các số liệu này, chúng ta có thể thấy được mức độ phổ biến của các tập dữ liệu nguồn mở trong nghiên cứu cũng như có thể lựa chọn tập dữ liệu phù hợp với yêu cầu nghiên cứu.

Việc lựa chọn tập dữ liệu phụ thuộc vào nhiều yếu tố. Các tập dữ liệu lớn hơn thường cung cấp khả năng đào tạo tốt hơn cho các mô hình học máy, trong khi chất lượng hình ảnh có thể ảnh hưởng đến độ chính xác của phân tích. Sự đa dạng trong tập dữ liệu, đại diện cho các nhóm tuổi, dân tộc và mật độ vú khác nhau, là rất quan trọng để cải thiện khả năng khái quát hóa mô hình. Các chú thích chính xác chỉ ra vị trí tổn thương là cần thiết cho việc học có giám sát. Các điều khoản cấp phép của tập dữ liệu và các vấn đề về y đức cũng phải được xem xét kỹ lưỡng để đảm bảo tập dữ liệu phù hợp với mục tiêu nghiên cứu.

Bằng cách tận dụng các tập dữ liệu chụp nhũ ảnh, các nhà nghiên cứu có thể phát triển các phương pháp để phát hiện, phân loại và tiên lượng UTV. Các thuật toán học máy, bao gồm các kỹ thuật học sâu, đã cho thấy kết quả đầy hứa hẹn trong việc xác định các bất thường nhỏ trong chụp nhũ ảnh mà các BS X quang có thể bỏ sót.

### **2.2.2 Bộ dữ liệu siêu âm công cộng**

Những bộ dữ liệu này là nguồn tài nguyên quan trọng thường được các nhà nghiên cứu sử dụng trong học máy. Mỗi tập dữ liệu cung cấp một bộ sưu tập hình ảnh siêu âm vú riêng biệt, phục vụ cho các nhu cầu nghiên cứu khác nhau.

BUS [73]: Được lấy từ Trung tâm Chẩn đoán UDIAT của Tập đoàn Parc Tauli ở Sabadell, Tây Ban Nha, BUS bao gồm 163 hình ảnh siêu âm vú. Trong số này, 109 trường hợp mô tả tình trạng lành tính, trong khi 54 trường hợp biểu hiện các đặc điểm ác tính.

BUSI [74]: Được tập hợp từ BV Baheya để phát hiện sớm và điều trị ung thư phụ nữ ở Cairo, Ai Cập, BUSI có bộ sưu tập phong phú hơn. Bộ dữ liệu này bao gồm các hình ảnh siêu âm thu được từ 600 bệnh nhân nữ trong độ tuổi từ 25 đến 75. Nó bao gồm 437 hình ảnh lành tính, 210 hình ảnh ác tính và 133 hình ảnh mô tả tình trạng vú bình thường, tổng cộng 730 hình ảnh siêu âm.

BUSIS [75]: Xuất phát từ BV liên kết thứ hai của Đại học Y Cáp Nhĩ Tân, BV liên kết của Đại học Thanh Đảo và BV thứ hai của Đại học Y Hà Bắc, BUSIS bao gồm 562 hình ảnh mô tả các đối tượng nữ từ 26 đến 78 tuổi. Đáng chú ý, những bộ dữ liệu này có thể chứa nhiều hình ảnh đại diện cho cùng một bệnh nhân.

Về nhãn hình ảnh, cả BUS và BUSI đều cung cấp nhãn hình dạng tổn thương cùng với các phân loại để phân biệt tình trạng lành tính và ác tính. Ngược lại, BUSIS chỉ cung cấp nhãn hình dạng tổn thương mà không cung cấp thêm chi tiết phân loại.

### 2.3. PHÂN ĐOẠN ẢNH VÀ XLA TRONG ẢNH UTV

#### 2.3.1 Phân đoạn ảnh UTV

Ảnh X-quang của các bệnh liên quan đến tổn thương vú thường có chất lượng không đồng đều và cũng chứa nhiều hiện tượng nhiễu không cần thiết. Do đó, trong hầu hết các nghiên cứu, các thuật toán tiền XLA được áp dụng để cải thiện đồng đều chất lượng ảnh và nâng cao hiệu quả phân loại. Do đó, việc tiền XLA trước khi áp dụng để huấn luyện các mạng học sâu tiên tiến như VGG16 [31], EfficientNet [32] hoặc các mạng khác có thể làm tăng hiệu quả phân loại.

Ở giai đoạn tiền xử lý trong các phương pháp máy học truyền thống, đa phần các nghiên cứu tập trung vào việc lọc nhiễu và lọc các đối tượng không quan trọng như cơ ngực, nhãn ... nhằm tạo điều kiện cho việc định vị vùng bất thường hoặc tổn thương một cách chính xác và giảm thiểu thời gian xử lý ở các giai đoạn sau [76]. Việc loại bỏ hình nền hoặc các thành phần không mong muốn trên ảnh X-quang vú có thể được thực hiện với các thuật toán tách vùng đơn giản như sử dụng ngưỡng Otsu [77].

Phương pháp Otsu lựa chọn giá trị ngưỡng  $k^*$  tối ưu để phân đoạn hình ảnh sao cho sự chênh lệch mức xám giữa 2 đoạn là lớn nhất.

$$\sigma_b^2(k^*) = \max_{1 \leq k \leq L-1} \sigma_b^2(k) \quad (2.1)$$

Trong đó  $k^*$  là ngưỡng Otsu tối ưu,  $k$  là ngưỡng mức xám nằm trong khoảng  $[1, L-1]$ ,  $\sigma_b^2$  là giá trị chênh lệch giữa 2 vùng  $C_1, C_2$  được phân chia theo ngưỡng  $k$ .

$$\sigma_b^2(k) = P_1(m_1 - m_G)^2 + P_2(m_2 - m_G)^2$$

Trong đó  $m_G$  là giá trị trung bình của ảnh

$$m_G = \sum_{i=0}^{L-1} ip_i \quad (2.2)$$

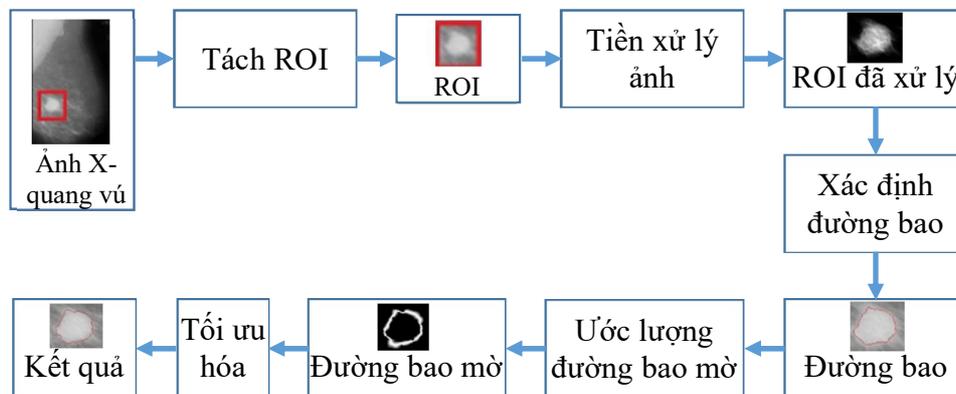
Và  $m_k$  là giá trị trung bình đến ngưỡng  $k$

$$m_k = \sum_{i=0}^k ip_i \quad (2.3)$$

Khi đã có nhũ ảnh với chất lượng tốt, các nhà nghiên cứu tiếp tục tập trung vào việc phân chia nhũ ảnh thành một số vùng có thể có tổn thương không chồng lấp bằng các kỹ thuật phân đoạn. Nhìn chung, các kỹ thuật phân đoạn có thể chia làm 4 nhóm: sử dụng ngưỡng, dựa trên vùng, dựa trên đặc trưng và dựa trên cạnh. Mỗi kỹ thuật sẽ có những ưu nhược điểm riêng, tuy nhiên ở giai đoạn này, để tăng tỷ lệ phát hiện khối u, độ nhạy và tỷ lệ dương tính sai thường được kỳ vọng có giá trị cao [30]. Có nhiều phương pháp phân đoạn ảnh khác nhau để tìm đối tượng, được gọi là Vùng quan tâm (ROI), chẳng hạn như phương pháp dựa trên vùng, phương pháp dựa trên ranh giới, phương pháp dựa trên atlas, phương pháp dựa trên mô hình và học sâu trong ảnh y

tế [78], [79]. Các phương pháp này có thể được áp dụng để loại bỏ nền, nhiễu và chỉ giữ lại đối tượng cần đánh giá, được gọi là khối u, trong ảnh ung thư [80], [81]. Phương pháp Otsu là một trong những phương pháp phổ biến nhất và được sử dụng rộng rãi trong ảnh y tế, chẳng hạn như ảnh X-quang, ảnh chụp CT và nhiều ảnh y tế khác [82], [83]. Đặc biệt, phương pháp Otsu được áp dụng để xác định giá trị ngưỡng kết hợp với đơn vị Hounsfield để chuyển đổi ảnh chụp CT thành ảnh nhị phân. Với phương pháp này, kết quả phân đoạn cao nhất là 77,43%. Trong bài báo này, ngưỡng Otsu tối ưu được sử dụng để trích xuất ROI trong ảnh xám có nền đen trong tập ảnh chụp nhũ ảnh có bốn loại tổn thương.

Sau khi thu được lượng lớn các vùng khả nghi (ROIs), các nhà nghiên cứu bắt đầu trích xuất các đặc trưng cần thiết cho việc phân loại. Thông thường, đối với mỗi vùng ROI, các đặc tính của nó như kích cỡ, hình dạng của khối u, tính đồng nhất của các đường biên, mật độ mô thường được xem xét đến [84]. Hình 2.9 mô tả sơ đồ thuật toán phân đoạn sử dụng đường bao mờ.



**Hình 2.9** Sơ đồ phân đoạn khối u sử dụng đường biên mờ

Ở bước phân đoạn cụ thể hơn, các thuật toán tập trung vào việc xác định chính xác vùng của khối u trên hình ảnh X-quang vú. Hiệu quả chẩn đoán trực tiếp phụ thuộc vào tính chính xác của việc xác định kích thước của khối u. Nhiều thuật toán máy học khác nhau đã được sử dụng để phân chia khối u một cách tự động. Một nghiên cứu đã ghi nhận tỷ lệ dương tính thực sự trung bình cao (91,12%) và độ chính

xác (88,08%) đối với RoI được trích xuất từ bộ dữ liệu mini-MIAS [85]. Độ tương phản thấp của hình ảnh chụp X quang tuyến vú, hình dạng khối không đều, lẻ có gai và cường độ điểm ảnh khác nhau góp phần tạo nên sự phức tạp trong việc phân chia tổng thể các khối trong chụp X quang tuyến vú. Trong một nghiên cứu khác, chức năng thiết lập mức độ tiến hóa để phân chia vú và các vùng khối đáng ngờ đã được khám phá bằng cách sử dụng phương pháp sắp xếp chức năng cơ sở xuyên tâm không có lưới và các vùng khối đáng ngờ được phân loại thành bất thường và bình thường bằng cách sử dụng bộ phân loại SVM, đạt được độ nhạy cao (97,12%) và độ đặc hiệu (92,43%) trên DDSM [86]. Phân đoạn chính xác các tổn thương vú đảm bảo phân loại và chẩn đoán bệnh chính xác [87]. Các thuật toán phân đoạn hình ảnh tự động như vậy chứng tỏ tiềm năng đầy hứa hẹn của DL trong các hệ thống y tế chính xác.

Một phương pháp phân đoạn mới cho chụp quang tuyến vú khối u được trình bày trong [88], phương pháp này trích xuất cả các vùng có gai và khối lõi. Có một mô hình tuyến tính chung trong việc sắp xếp các pixel trong các vùng được mô phỏng, được phản chiếu trong các vùng lõi lớn nơi các pixel cũng tuyến tính. Phương pháp đề xuất trích xuất các vùng này dựa trên sự khác biệt giữa các pixel liền kề. Trong vùng lõi khối và vùng có hình dạng, các pixel dư thừa có thể bị xóa bằng ba ngưỡng; các khối u được phân đoạn sau đó được hình thành bằng cách hợp nhất các vùng này. Phương pháp này trình bày hệ số Dice trung bình là 0,9309 trên tập dữ liệu mini-MIAS và hệ số Jaccard trung bình là 0,9557 và 0,9132 trên DDSM. Theo kết quả, khi so sánh với các kỹ thuật khác, kỹ thuật phân đoạn được đề xuất có thể trích xuất chính xác các phân đoạn khối u.

Các mô hình mạng CNN ngày càng đóng vai trò mạnh mẽ hơn trong bài toán phân đoạn khối u trên ảnh X-quang vú. Salama và Aly [89] đã thu thập hình ảnh từ DDSM, bộ dữ liệu MIAS mini và CBIS-DDSM. Họ đã sử dụng nhiều mô hình khác nhau để phân đoạn và phân loại hình ảnh là lành tính hay ác tính, bao gồm DenseNet121, InceptionV3, VGG16, ResNet50 và MobileNetV2. Khi sử dụng InceptionV3 với tính năng tăng cường dữ liệu, độ chính xác tốt nhất là 88,87%. Li và cộng sự. [90] đã đề xuất một phương pháp hoàn toàn tự động kết hợp các U-Net được

kết nối dày đặc với các công chú ý (AG) để phân chia khối lượng vú trên hình ảnh chụp quang tuyến vú. Nó bao gồm một bộ mã hóa và một bộ giải mã: Mạng tích chập mã hóa thông tin và U-Net được tích hợp với AG sẽ giải mã thông tin đó. DDSM, CSDL sàng lọc công khai được ủy quyền, đã được sử dụng để kiểm tra phương pháp được đề xuất. F1 score, nghĩa là sự giao nhau giữa độ kết hợp, độ nhạy, độ đặc hiệu và độ chính xác tổng thể, được sử dụng để đánh giá tính hiệu quả của phương pháp. So với U-Net, U-Net, DenseNet và các kỹ thuật phân đoạn tiên tiến, U-Net dày đặc được tích hợp với AG vượt trội hơn các phương pháp khác và đạt được kết quả phân đoạn tốt hơn, với độ chính xác 78,38% và điểm F1 là 78,24%. Ngoài ra, mạng có độ lệch chuẩn thấp hơn, cho thấy nó có khả năng khái quát hóa cao hơn.

Các kết quả của nghiên cứu được báo cáo trong [91] đã chứng minh sự phân chia UTV có độ chính xác cao bằng cách kết hợp các CNN vùng mặt nạ (mặt nạ R-CNN) với các CNN nhóm (G-CNN). Những cách tiếp cận này tối đa hóa tỷ lệ trọng số và khả năng biểu đạt của mô hình trong khi vẫn duy trì tính bất biến quay. Bộ dữ liệu INbreast và mini-MIAS đã được sử dụng để kiểm tra mô hình. So sánh kết quả với kết quả của các kiến trúc thông thường, mô hình đạt độ chính xác 99,01%, hệ số Dice là 86,63%, chỉ số Jaccard là 87,76%, độ nhạy 99,24% và độ đặc hiệu 98,55%.

### 2.3.2 Cải thiện chất lượng hình ảnh trong chẩn đoán UTV

Độ chính xác của chẩn đoán phụ thuộc rất nhiều vào chất lượng hình ảnh. Chất lượng hình ảnh tốt có nghĩa là hình ảnh rõ ràng, giúp cải thiện đáng kể tỷ lệ chẩn đoán và độ chính xác của mô hình AI khi phát hiện và chẩn đoán các tổn thương vi thể trong chụp X-quang tuyến vú. Các thuật toán máy tính khác nhau đã được phát triển để nâng cao chất lượng hình ảnh. Chất lượng hình ảnh cao hơn cung cấp nhiều thông tin hơn về pha, tính định hướng và tính bất biến dịch chuyển của dữ liệu.

Để cải thiện chất lượng của nhũ ảnh và tăng độ tương phản, những kỹ thuật tăng cường ảnh được nghiên cứu áp dụng. Thường những kỹ thuật này tập trung chính vào việc cải thiện chất lượng nhũ ảnh ở những vùng có độ tương phản thấp, vì ở những vùng này các bất thường nhỏ thường bị che giấu trong các mô xung quanh, dễ dẫn

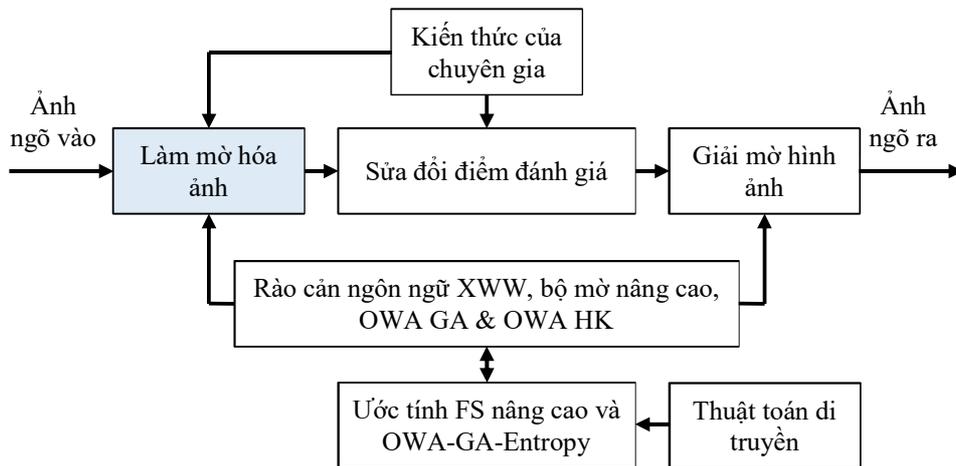
đến chẩn đoán sai. Những kỹ thuật tăng cường ảnh thường được chia ra làm 4 loại [92]: các kỹ thuật truyền thống, các kỹ thuật tăng cường theo vùng, các kỹ thuật tăng cường dựa trên đặc trưng, và các kỹ thuật tăng cường mờ. Các kỹ thuật tăng cường truyền thống thường có thể đáp ứng cho việc tăng cường dù là từng phần hay là toàn bộ nhũ ảnh. Tuy nhiên, ngoài việc tăng cường độ tương phản, các kỹ thuật này cũng có xu hướng tăng cường nhiễu. Do đó, để tăng cường chất lượng các chi tiết ảnh mà không kèm theo bất kỳ ảnh hưởng tiêu cực nào, các phương pháp tăng cường theo vùng được phát triển. Các kỹ thuật này khá phù hợp với việc tăng cường các vùng vi vôi hoá ở nhũ ảnh khi mô dày đặc. Ngoài ra, để tăng cường nhũ ảnh với cả vi vôi hoá và mass, những kỹ thuật tăng cường dựa trên đặc trưng cũng có thể được sử dụng. Ở những kỹ thuật này, thông thường nhũ ảnh sẽ được biến đổi sang vùng tần số với các bộ biến đổi đa mức như wavelets, sau đó lọc đi những vùng tần số thấp. Nhũ ảnh sẽ được xây dựng lại chỉ với những vùng có tần số cao chứa đựng những khu vực có khả năng bị tổn thương. Với nguyên lý tối đa hoá entropy mờ (maximum fuzzy entropy), các kỹ thuật tăng cường mờ cũng có thể được áp dụng trên nhũ ảnh để tăng cường độ tương phản và triệt tiêu nhiễu. Các kỹ thuật này khá hiệu quả trong việc tăng cường các đường viền khối và làm rõ hơn các chi tiết trong nhũ ảnh. Tuy nhiên gần đây, sự phát triển của công nghệ kỹ thuật số với dải động cao (high dynamic range) [84], độ tương phản và tỷ lệ tín hiệu trên nhiễu được cải thiện một cách trực tiếp làm cho chất lượng nhũ ảnh ngày càng được tăng cao đồng thời cũng thu hẹp không gian nghiên cứu cho các kỹ thuật tăng cường nhũ ảnh [92].

Trong các nghiên cứu gần đây, nhiều thuật toán đã được áp dụng trong việc cải thiện chất lượng hình ảnh X-quang vú. Nghiên cứu của Shenbagavalli cho thấy các trường hợp lành tính và ác tính trong DDSM đã được phân loại với độ chính xác lên tới 93,45% bằng phương pháp nâng cao hình ảnh biến đổi shearlet, cho thấy hiệu quả của nó trong việc nâng cao chất lượng hình ảnh chụp quang tuyến vú [47]. Nhóm nghiên cứu của Teare đã sử dụng một dạng mới của phương pháp tăng cường màu giả để tối ưu hóa các đặc điểm của ảnh chụp X quang tuyến vú thông qua cân bằng biểu đồ thích ứng giới hạn độ tương phản (CLAHE) và họ sử dụng CNN sâu kép ở

## Chương 2 CƠ SỞ LÝ THUYẾT

các tỷ lệ khác nhau để phân loại hình ảnh và các mảng phái sinh, cùng với một khung ngẫu nhiên mạng lưới cổng. Bằng cách này, họ đạt được độ nhạy 0,91 và độ đặc hiệu 0,80 [93]. Do tầm quan trọng của chất lượng hình ảnh đối với chẩn đoán chính xác, việc đánh giá và cải thiện chất lượng hình ảnh một cách nghiêm ngặt là điều cần thiết cho các phân tích và chẩn đoán tiếp theo của hệ thống ANN và BS X quang.

Để giảm nhiễu của hình ảnh chụp quang tuyến vú trong khi vẫn giữ được độ tương phản và độ sáng, Suradi [94] đã phát triển kỹ thuật giới hạn thích ứng tương phản thích ứng biểu đồ khuếch tán dị hướng mờ (FADHECAL). Ngoài kỹ thuật FADHECAL, hệ thống suy luận cắt bớt mờ (FCIS) được áp dụng trong quá trình nâng cao, tự động chọn giới hạn cắt từ các tùy chọn có sẵn. Bộ dữ liệu DDSM và mini-MIAS đã được sử dụng để truy cập hình ảnh chụp quang tuyến vú. Kết quả cho thấy kỹ thuật FAD-HECAL cho kết quả tốt hơn các phương pháp nâng cao được lựa chọn khác. Kỹ thuật FADHECAL có thể được sử dụng để cải thiện hình ảnh chụp quang tuyến vú nhằm phát hiện các tổn thương UTV chính xác hơn với mức độ nhiễu giảm trong khi vẫn giữ được chi tiết hình ảnh.



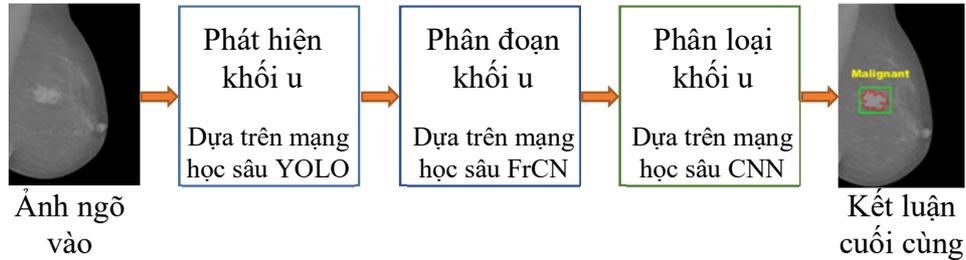
**Hình 2.10** Sơ đồ hệ thống tăng cường chất lượng hình ảnh đề xuất bởi [95]

Một số cách tiếp cận được trình bày trong [95] để tăng cường độ tương phản của ảnh chụp nhũ ảnh, bao gồm các phương pháp cổ điển (hàng rào ngôn ngữ và hàm tăng cường mờ), tập mờ nâng cao (tập mờ trực quan, Pythagore và Fermatean) và tối

ưu hóa thuật toán di truyền. Tập mờ nâng cao cung cấp đánh giá chính xác hơn về tính không chắc chắn của hàm thành viên. Vì lý do này, phương pháp trực quan là hiệu quả nhất, nhưng hầu hết các kỹ thuật khác cũng có hiệu quả, tùy thuộc vào vấn đề. So với các phương pháp thông thường, các phương pháp ngôn ngữ có thể cung cấp một cách dễ quản lý hơn để trải rộng biểu đồ, tiết lộ nhiều giá trị cực đoan hơn. Có thể thu được hình ảnh cuối cùng chất lượng cao bằng cách sử dụng toán tử lấy trung bình trọng số (OWA) theo thứ tự kết hợp với hình ảnh chụp nhũ ảnh nâng cao.

### 2.4. PHÂN LOẠI ẢNH UTV SỬ DỤNG MÁY HỌC

Các phương pháp sàng lọc UTV thường dựa trên việc sử dụng chụp nhũ ảnh [96]. Phương pháp sàng lọc không xâm lấn này dễ thực hiện và mang lại hình ảnh có độ phân giải và độ tin cậy cao. Nó có thể được áp dụng cho hầu hết đối tượng khác nhau về tuổi tác hoặc thể trạng. Các hình ảnh cũng dễ dàng được lưu trữ để có thể so sánh tiến trình phát triển của bệnh. Các khối vú không thể phát hiện được bằng khám lâm sàng, các tổn thương lành tính và khối u vú ác tính có thể được phát hiện chính xác bằng chụp X quang tuyến vú. Chụp nhũ ảnh bao gồm các hệ thống chụp nhũ ảnh kỹ thuật số toàn trường (DM) và mang lại các định dạng hình ảnh để xử lý (dữ liệu hình ảnh thô) và để trình bày (các phiên bản sau xử lý của dữ liệu thô) [97]. Tuy nhiên, ảnh x-quang vú với nhiều vùng khác nhau rất khó để chẩn đoán bởi BS và đòi hỏi chuyên môn cao cũng như kinh nghiệm. Việc sử dụng TTNT để hỗ trợ cho quá trình phân loại nhũ ảnh đã được nghiên cứu rất sớm cùng với sự phát triển của công nghệ TTNT đã giúp cải thiện đáng kể quá trình này. Các nghiên cứu về việc sử dụng AI để phát hiện, phân loại và phân chia khối vú, đánh giá nguy cơ UTV sẽ được thảo luận trong phần này. Sơ đồ của hệ thống này được trình bày trong Hình 2.11.



**Hình 2.11 Sơ đồ hệ thống chẩn đoán hỗ trợ máy tính (CAD) được đề xuất dựa trên học sâu để phát hiện, phân đoạn và phân loại khối u UTV từ ảnh X-quang kỹ thuật số đầu vào**

#### 2.4.1 Phát hiện và phân loại khối u vú

Một trong những triệu chứng phổ biến nhất của UTV là sự xuất hiện của khối u. Việc phát hiện và chẩn đoán những bất thường như vậy có thể là một thách thức, đặc biệt là ở những bộ ngực dày đặc, do hình dạng, kích thước và đường viền khác nhau của chúng. Đây là lý do tại sao việc phát hiện khối vú là một bước quan trọng trong chẩn đoán UTV. Một số nghiên cứu đã khuyến nghị tối ưu hóa tìm kiếm theo đám đông dựa trên phương pháp phân cụm mờ trực quan với sức hút vùng lân cận (CrSA-IFCM-NA). Hiệu quả của nó trong việc tách các khối trong hình ảnh chụp quang tuyến vú và kết quả tốt trên các chỉ số hợp lệ của cụm, cho thấy sự phân chia vùng chính xác, đã được chứng minh [98]. Các nghiên cứu khác khuyến nghị sử dụng hệ thống hỗ trợ chẩn đoán tích hợp hoàn toàn bao gồm phương pháp DL khu vực YOLO (You Only Look One), mô hình mạng sâu mới của mạng tích chập có độ phân giải đầy đủ (FrCN) và CNN sâu để phát hiện, phân đoạn và phân loại khối lượng trong chụp quang tuyến vú, tương ứng. Sử dụng bộ dữ liệu INbreast (2022), các nghiên cứu này báo cáo độ chính xác phát hiện đáng chú ý là 98,96%, cho thấy tiềm năng hệ thống này hỗ trợ các BS X quang chẩn đoán hiệu quả và chính xác những khối này [99], [100].

Ertoşun và cộng sự [101] đã trình bày một phương pháp tìm kiếm trực quan dựa trên CNN được thiết kế để xác định vị trí các khối trong ảnh chụp X quang tuyến vú.

## Chương 2 CƠ SỞ LÝ THUYẾT

Mô hình của họ bao gồm hai mô-đun phụ chính: một máy dò dị thường, sau đó là một bộ định vị khối lượng. Ban đầu, máy phát hiện dị thường sẽ phân biệt liệu ảnh chụp X quang tuyến vú có chứa các khối hay không, sau đó chuyển các hình ảnh có chứa khối vào bộ định vị để định vị chính xác. Các mô-đun này, tận dụng các lớp CNN phân cấp, trải qua quá trình đào tạo trên tập dữ liệu vượt quá 2500 hình ảnh có nguồn gốc từ DDSM. Al-Masni và cộng sự. [102] đã chọn mạng sâu bạn chỉ nhìn một lần (YOLO) do Redmon và cộng sự giới thiệu. [103] để phát hiện và phân loại khối lượng đồng thời. YOLO sử dụng phương pháp học từ đầu đến cuối, sử dụng các lớp tích chập liên tiếp để phân chia hình ảnh thành các vùng phụ, sau đó đặt các hộp giới hạn xung quanh các đối tượng quan trọng và gán nhãn lớp. Sau khi xác thực chéo năm lần, các tác giả đã báo cáo điểm số về độ nhạy (100%) và độ đặc hiệu (94%) ấn tượng trên tập dữ liệu INBreast.

**Bảng 2.1 Tổng quan về các nghiên cứu AI trong chẩn đoán UTV**

Nghiên cứu	Phương pháp	Ứng dụng	Kích thước tập dữ liệu	Độ chính xác
Alsheh Ali 2019 [97]	DoG và HoG	Phát hiện chuỗi vi vôi hóa	373 ca	-
Ertosun 2015 [101]	CNN	Máy nhận dạng và định vị	2420 ca	85%
Al-masni 2018 [102]	YoLo	Phát hiện	600 ca	99.7%
Ribli 2018 [24]	Fast-RCNN Deep	Phát hiện	DDSM với 2620 ca, SU-D với 847 ca, INbreast với 115 ca	-
Zhu W. 2019 [104]	Deep multi-instance networks	Phân loại	410 ca	90%
Bouzar-Benlabiod 2023 [105]	CBR	Phân loại	2620 ca	91.34%
Gao 2018 [106]	CNN	Phân loại	INbreast 89 ca, MCA 49 ca	90%
Elkorany 2023 [107]	CNN features + MSVM	Phân loại	416 ca	90%
Yu 2020 [108]	Deep fusion learning	Phân loại	208 ca	89.06%
Hmida 2018 [85]	Fuzzy contours	Phân đoạn	57 ca	88.08%
Kashyap 2017 [86]	Mesh-free + SVM	Phân đoạn	322 ca	94.77%
Hussain 2022 [87]	Dense U-Net + AGs	Phân đoạn	D-A 186 ca, D-B 163 ca	78.38%

## Chương 2 CƠ SỞ LÝ THUYẾT

Pezeshki 2022 [88]	Mask-RCNNs + GCNNs	Phân đoạn	MIAS 58 ca, DDSM 200 ca	99.01%
Salama 2021 [89]	CNN	Phân đoạn	885 ca	91%
Li 2019 [91]	U-Net với attention gates (AGs)	Phân đoạn	400 ca	78.38%
Sani 2024 [90]	Mask-RCNN với GCNN	Phân đoạn	MIAS với 322 ca và Inbreast với 115 ca	99.1%
Teare 2017 [93]	CLAHE and CNN	Tăng cường hình ảnh	DDSM 6000 ca, ZMDS 1739 ca	85.5%
Saifullah Harith Suradi 2022 [94]	FADHECAL and FCIS	Tăng cường hình ảnh	DDSM 2620 ca, MIAS 322 ca	-
Dounis 2023 [95]	LH and FEF	Tăng cường hình ảnh	97 ca	-

Ribli và cộng sự. [24] đã triển khai phương pháp R-CNN nhanh hơn, đạt được AUC đáng khen ngợi là 0,95 trên tập dữ liệu INBreast. Đáng chú ý, chúng góp phần vào khả năng tái tạo bằng cách cung cấp nguồn mở cho việc triển khai trên GitHub, một phương pháp hiếm thấy trong lĩnh vực này. Để nâng cao độ tin cậy của mô hình, các tác giả đào tạo dữ liệu DDSM và đánh giá trên INBreast, thể hiện khả năng khái quát hóa miền. Platania và cộng sự. [104] đã sử dụng các trọng số CNN được huấn luyện trước để khởi tạo các trọng số của bộ phân loại nhị phân theo cách bán giám sát. Cách tiếp cận của nghiên cứu này sử dụng một hệ thống hai mô-đun, trong đó CNN lấy cảm hứng từ YOLO sẽ phát hiện các vùng quan tâm (ROI). Sau đó, trọng số của máy dò ban đầu được chuyển sang bộ phân loại FFDM, bộ phân loại này trải qua quá trình huấn luyện trên toàn bộ hình ảnh chụp quang tuyến vú. Sau thử nghiệm trên DDSM, các tác giả báo cáo điểm AUC khoảng 92,3% và độ chính xác là 93,5%.

Trong một nghiên cứu [105], các tác giả đã phát triển một phương pháp mới sử dụng TTNT để phát hiện UTV từ ảnh chụp quang tuyến vú. Mục tiêu của họ là tạo ra một hệ thống không chỉ chính xác mà còn dễ giải thích. Điểm nổi bật của phương pháp này là sự kết hợp giữa hai kỹ thuật: U-Net và Case-Based Reasoning (CBR). Trong đó U-Net là một kiến trúc học sâu được dùng để phân đoạn ảnh, giúp xác định các vùng quan tâm (ROI) một cách hiệu quả. Điều này giúp trích xuất các đặc trưng ảnh quan trọng từ ảnh chụp quang tuyến vú. Case-Based Reasoning (CBR) là một phương pháp phân loại dựa trên việc so sánh các trường hợp mới với những trường

## Chương 2 CƠ SỞ LÝ THUYẾT

---

hợp đã có. CBR giúp đưa ra chẩn đoán cuối cùng một cách dễ hiểu, bởi vì nó dựa trên những trường hợp thực tế đã được phân tích trước đó. Nhờ việc kết hợp này, các nhà nghiên cứu đã tạo ra một hệ thống vừa đạt được độ chính xác cao từ U-Net, vừa đảm bảo tính giải thích từ CBR. Kết quả cho thấy, phương pháp này vượt trội hơn nhiều so với các phương pháp học máy (Machine Learning - ML) và học sâu (DL) khác. Cụ thể, trên CSDL CBIS-DDSM, độ chính xác đạt được là 86,71% khi sử dụng CBR và 91,34% khi kết hợp cả hai kỹ thuật.

Một nghiên cứu khác của Gao và cộng sự [106] cũng đề xuất một mô hình CNN (mạng nơ-ron tích chập) nông-sâu để phân loại khối u là lành tính hay ác tính. Mô hình này sử dụng CNN nông để kết hợp các hình ảnh năng lượng thấp, sau đó dùng CNN sâu để trích xuất các đặc trưng riêng biệt. Phương pháp này đã đạt được độ chính xác lên tới 90%.

Một kỹ thuật lai để chẩn đoán UTV đã được đề xuất và thử nghiệm trong một nghiên cứu [107]. Các tác giả đã phát triển và thử nghiệm một phương pháp lai để chẩn đoán UTV. Phương pháp này kết hợp các mô hình học sâu (DL) với thuật toán thống kê và học máy (ML) để đạt được độ chính xác cao. Họ đã sử dụng ba mô hình DL CNN phổ biến là Inception-V3, ResNet50, và AlexNet để trích xuất các đặc trưng từ ảnh chụp quang tuyến vú. Các đặc trưng này sau đó được tinh lọc bằng thuật toán lựa chọn đặc trưng phương sai (TV). TV giúp xác định và giữ lại các đặc trưng quan trọng, có độ biến động cao, đồng thời loại bỏ những đặc trưng không cần thiết. Sau khi các đặc trưng đã được TV lựa chọn, chúng được kết hợp lại và chuyển đến máy véc-tơ hỗ trợ đa lớp (MSVM) để phân loại. MSVM sẽ xác định xem khối u là lành tính hay ác tính. Để đánh giá hiệu quả, phương pháp này đã được kiểm tra trên CSDL mini-MIAS. Kết quả cho thấy: Với 70% dữ liệu dùng để huấn luyện, độ chính xác đạt 97,81%; Với 80% dữ liệu dùng để huấn luyện, độ chính xác đạt 98%; Với 90% dữ liệu dùng để huấn luyện, độ chính xác đạt 99%. Kết quả này cho thấy phương pháp lai này đạt được độ chính xác phân loại cao hơn nhiều so với các nghiên cứu trước đó. Một phân tích bổ sung đã được thực hiện để làm nổi bật các thông số chính của mô hình.

Yu, X và cộng sự. [108] đã nghiên cứu các mô hình phân biệt giữa các loại bình thường và khối u dựa trên phương pháp học tổng hợp sâu. Khung phân loại hình ảnh chụp nhũ ảnh bằng phương pháp học tổng hợp sâu bao gồm hai bước. Các mô hình phản ứng tổng hợp sâu được đề xuất trước tiên được huấn luyện trên các bản vá RoI được chọn ngẫu nhiên từ tất cả các RoI trong tập dữ liệu gốc sau khi chúng được lấy. Các tác giả đã phát triển một mô hình phản ứng tổng hợp sâu để phân loại các mảng RoI trong các mô khối u và mô bình thường. Một mô hình khác tích hợp các tính năng sâu đa kênh bằng cách sử dụng tích chập một-một để khám phá mối liên kết giữa các kênh của cùng một khối. Các bản vá dự đoán đa số phiếu bầu tạo nên một RoI và dự đoán cuối cùng. Mô hình 1 đạt được tỷ lệ thu hồi là 0,913, tỷ lệ chính xác là 0,8077 và tỷ lệ chính xác tổng thể là 0,8906, trong khi Mô hình 2 đạt được độ chính xác tổng thể là 0,875, tỷ lệ thu hồi là 0,9565 và tỷ lệ chính xác là 0,7586 đối với dữ liệu loại khối u.

### 2.4.2 Đánh giá nguy cơ UTV

Các phương pháp học sâu ngày càng trở nên thông dụng và gặt hái được các thành công nhất định cho bài toán phát hiện UTV sử dụng nhũ ảnh. Đơn cử như nghiên cứu của Mohamed và đồng nghiệp [109] về việc phân loại nhũ ảnh dựa trên mật độ vú sử dụng mạng nơ-ron tích chập (CNN). Tuy nhiên nhóm nghiên cứu chỉ tập trung vào phân loại hai loại mật độ khó nhất là: mật độ mô vú dày đặc rải rác và mật độ mô vú không đồng nhất trong 4 loại về mật độ của hệ thống BIRADS. Ở một số nghiên cứu khác, các chiến lược khác nhau được áp dụng như: (1) sử dụng các patch nhỏ trích xuất từ toàn bộ nhũ ảnh để huấn luyện [94], (2) thu thập và dán nhãn hơn 20.000 nhũ ảnh [110] để huấn luyện, hoặc (3) sử dụng residual CNN [111] để phân loại 4 loại mật độ này. Bên cạnh đó, để giảm thiểu số lượng dữ liệu dán nhãn cần thiết, nhóm nghiên cứu của Kallenberg [112] đã đề xuất sử dụng kỹ thuật học không giám sát để phân loại mật độ mô vú và chỉ số nguy hiểm (ung thư hoặc bình thường). Trong kỹ thuật này, sparse autoencoder được sử dụng để học các đặc trưng hữu ích.

Ngoài các hệ thống phân loại mật độ mô vù, các nhà nghiên cứu cũng xây dựng thành công các hệ thống phát hiện và phân loại UTV sử dụng học sâu. Các đặc trưng có kết cấu và có ngữ nghĩa sâu trong các mạng học sâu cho phép các bộ phân loại phân biệt giữa các tổn thương bình thường và bất thường với nhiều hình dạng, kích cỡ và hướng khác nhau. Ví dụ như nhóm nghiên cứu của Levy và Jian [113] đã áp dụng cả học chuyển giao và tăng cường dữ liệu để phát triển hệ thống phân loại các vùng có nguy cơ ung thư thành hai loại lành tính và ác tính. Gần đây, để cải thiện hiệu suất phân loại, nhóm nghiên cứu của Samala [62] tiếp tục đề xuất việc học chuyển giao đa tầng (multistage transfer learning) để lựa chọn sự kết hợp các đặc trưng học sâu một cách tối ưu cho việc phân loại. Các phương án lai kết hợp giữa học sâu và các phương pháp truyền thống cũng được áp dụng và đạt được những kết quả khả quan cho việc phân loại UTV [114], [115].

Ngoài số lượng dữ liệu hạn chế, vấn đề mất cân bằng giữa các loại dữ liệu cũng là một vấn đề nan giải trong các phương pháp học sâu. Việc mất cân bằng giữa các loại dữ liệu dễ gây ra hiện tượng thiên vị trong các mô hình học sâu, nghĩa là kết quả dự đoán có xu hướng tập trung vào một hoặc vài loại có số lượng nhiều trong tập dữ liệu huấn luyện. Để giải quyết vấn đề này, Wu và đồng nghiệp [116] đề xuất tạo ra các ảnh tổng hợp bằng mạng cGAN. Kết quả cho thấy, với việc tăng cường các dữ liệu tổng hợp, kết quả phân loại được cải thiện một cách hiệu quả. Cũng để giải quyết vấn đề trên, nhóm nghiên cứu của Shams [117] đề xuất một phương pháp học sâu đa nhiệm, phương pháp này cũng cho thấy khá hiệu quả trên các tập dữ liệu hạn chế. Một nghiên cứu gần đây nhất, ngày 1 tháng 1 năm 2020, của Google được công bố ở tạp chí Nature [118] được xem như là một nghiên cứu bước ngoặt chứng minh tiềm năng của AI trong việc phát hiện UTV sử dụng ảnh. Nhóm nghiên cứu cho thấy hệ thống TTNT của họ có thể giúp xác định bệnh UTV với độ chính xác bằng với kết quả của chuyên gia. Không những thế, hệ thống của họ còn chiến thắng các chuyên gia trong việc giảm thiểu các trường hợp dương tính sai (positive false) và âm tính sai (negative false).

## Chương 2 CƠ SỞ LÝ THUYẾT

---

Do tỷ lệ mắc và tỷ lệ tử vong cao của bệnh UTV, sức khỏe thể chất và cả tinh thần của các bệnh nhân UTV bị ảnh hưởng xấu. Có nhiều yếu tố nguy cơ, chẳng hạn như tuổi tác, tiền sử gia đình, các yếu tố sinh sản (ví dụ: có kinh sớm, mãn kinh muộn, mang thai lần đầu ở độ tuổi muộn, số lần sinh thấp), estrogen (ví dụ: estrogen nội sinh hoặc ngoại sinh) và lối sống (ví dụ: hút thuốc, uống quá nhiều rượu, ăn nhiều chất béo) [119]. Nhận thức và hiểu rõ hơn về nguy cơ UTV có thể thúc đẩy việc phát hiện và phòng ngừa sớm. Nhiều nghiên cứu đã khám phá rộng rãi việc sử dụng AI trong dự đoán nguy cơ UTV. Ví dụ, trong một đánh giá có hệ thống về các thuật toán ML để dự đoán nguy cơ UTV được báo cáo từ tháng 1 năm 2000 đến tháng 5 năm 2018, Nindrea et al. nhận thấy rằng thuật toán SVM có độ chính xác cao nhất so với ANN, cây quyết định, k-láng giềng gần nhất và thuật toán Naive Bayes [120]. Một số nghiên cứu khác đã chứng minh rằng sự kết hợp giữa ANN và chẩn đoán tế bào học có thể được sử dụng để đánh giá nguy cơ UTV bằng cách phân tích và tìm hiểu kết quả chụp X-quang tuyến vú, các yếu tố nguy cơ và phát hiện lâm sàng, có thể giúp BS đưa ra ước tính sáng suốt về nguy cơ ác tính và cải thiện kết quả điều trị. giá trị tiên đoán dương (PPV) của quyết định thực hiện sinh thiết [121].

Một số nghiên cứu [122]-[126] đã sử dụng các nhóm sàng lọc cắt ngang lớn đại diện cho nhóm sàng lọc chung để đào tạo các mô hình DL. Những nghiên cứu này sử dụng hình ảnh chụp quang tuyến vú bình thường thu được ít nhất một năm trước khi chẩn đoán UTV hoặc theo dõi âm tính (tức là BIRADS 1 và 2). Thiết kế của những nghiên cứu này phản ánh chặt chẽ hơn nhiệm vụ đánh giá nguy cơ UTV, vì chúng nhằm mục đích xác định các nhóm phụ nữ nguy cơ cao trước khi thực sự phát triển thành ung thư. Điều quan trọng là sử dụng các trường hợp và đối chứng ở cùng độ tuổi trong một nghiên cứu như vậy và báo cáo các biện pháp đánh giá được điều chỉnh theo độ tuổi. Nếu không, ước tính hiệu suất dự đoán rủi ro có thể bị thổi phồng. Các vùng bên dưới đường cong đặc tính vận hành máy thu (AUC) của các mô hình nằm trong khoảng từ 0,60 tới 0,84 và thường hoạt động tốt hơn các mô hình nguy cơ UTV hiện đại [95], [119], [122], [124]. Hà và cộng sự. [126] cho thấy điểm rủi ro DL dựa trên chụp nhũ ảnh kỹ thuật số toàn trường (FFDM) có tính dự đoán cao hơn so với

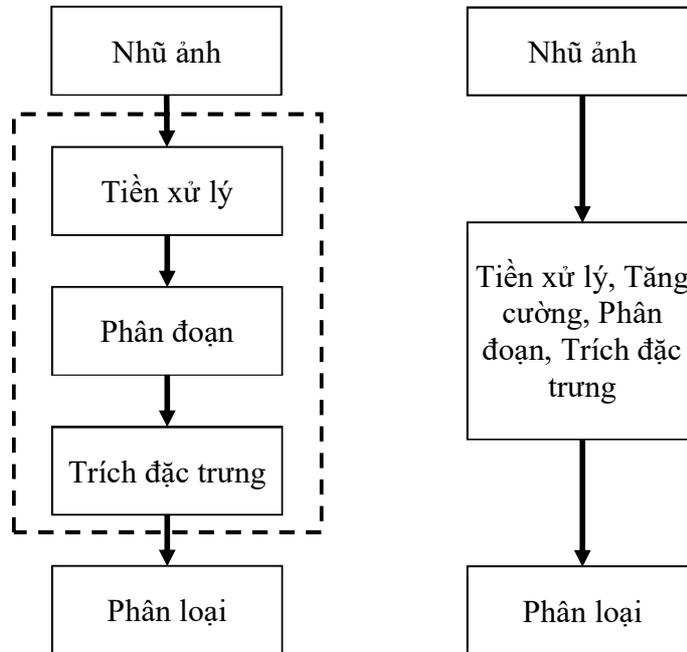
mật độ vú của hệ thống dữ liệu báo cáo và hình ảnh vú (BIRADS) (tỷ lệ chênh lệch 4,4 so với 1,7). Điểm rủi ro DL dựa trên FFDM vượt trội hơn các phép đo mật độ vú tự động theo Dembrower et al. [122]. Cuối cùng, Yala và cộng sự. nhận thấy rằng mô hình DL lai bao gồm cả chụp quang tuyến vú toàn trường và các yếu tố rủi ro truyền thống ghi lại độ chính xác cao hơn mô hình Tyrer-Cusick, đại diện cho tiêu chuẩn lâm sàng hiện tại (AUC lần lượt là 0,68 so với 0,62) [124]. Những nghiên cứu kể trên cung cấp các bằng chứng sơ bộ rằng các mô hình DL dựa trên FFDM hứa hẹn đưa ra dự đoán chính xác hơn về nguy cơ UTV so với các mô hình dựa trên mật độ và các mô hình dựa trên dịch tễ học hiện có. Với những điều trên, độ chính xác của kỹ thuật AI trong dự đoán nguy cơ UTV là rõ ràng, điều này chắc chắn có thể giúp các BS đưa ra các biện pháp can thiệp thích hợp để giảm thiểu nguy cơ UTV.

### 2.5. CHẨN ĐOÁN UTV TỪ XA DỰA TRÊN TTNT

#### 2.5.1 Mạng CNN và ứng dụng trong chẩn đoán hình ảnh

Một số hệ thống hỗ trợ chẩn đoán UTV từ xa dựa trên nhũ ảnh sử dụng các phương pháp XLA [127] hoặc các mạng TTNT [128] để hỗ trợ việc chẩn đoán cũng được phát triển trong các nghiên cứu gần đây. Tuy nhiên chưa thật sự có một hệ thống nào có thể cung cấp sự chẩn đoán thay thế được các BS với độ chính xác và tin cậy cao.

Trên thế giới, các hệ thống AI chẩn đoán bệnh UTV sử dụng nhũ ảnh được quan tâm và phát triển từ rất lâu. Có hai nhóm phương pháp cơ bản được sử dụng đó là: nhóm phương pháp máy học truyền thống và nhóm phương pháp học sâu [15]. Sự khác biệt giữa hai nhóm phương pháp này được thể hiện trên Hình 2.12. Đối với các phương pháp máy học truyền thống, để phân loại nhũ ảnh, rất nhiều bước phải được quan tâm thiết kế như: tiền xử lý, phân đoạn, trích xuất đặc trưng và phân loại. Đặc biệt các đặc trưng được sử dụng trong phương pháp này thường là các đặc trưng thiết kế thủ công theo ý kiến các chuyên gia. Còn trong các phương pháp học sâu, tất cả các bước này được tập hợp trong một cấu trúc mạng duy nhất, và đặc biệt là các đặc trưng sẽ được huấn luyện và trích xuất một cách tự động từ dữ liệu huấn luyện.



**Hình 2.12 Sự khác biệt giữa phương pháp máy học truyền thống (bên trái) và phương pháp học sâu (bên phải)**

Ở giai đoạn tiền xử lý trong các phương pháp máy học truyền thống, đa phần các nghiên cứu tập trung vào việc lọc nhiễu và lọc các đối tượng không quan trọng như cơ ngực, nhãn ... nhằm tạo điều kiện cho việc định vị vùng bất thường hoặc tổn thương một cách chính xác và giảm thiểu thời gian xử lý ở các giai đoạn sau. Sau đó, để cải thiện chất lượng của nhũ ảnh và tăng độ tương phản, những kỹ thuật tăng cường ảnh được nghiên cứu áp dụng. Thường những kỹ thuật này tập trung chính vào việc cải thiện chất lượng nhũ ảnh ở những vùng có độ tương phản thấp, vì ở những vùng này các bất thường nhỏ thường bị che giấu trong các mô xung quanh, dễ dẫn đến chẩn đoán sai.

Cấu trúc mạng CNN gồm các lớp sau: Convolutional layers: lớp tích chập, lớp hiệu chỉnh tuyến tính (ReLU), lớp Pooling, lớp kết nối đầy đủ

**Lớp tích chập:**

Lớp này sử dụng bộ lọc kích thước nhỏ và lần lượt quét qua toàn bộ ảnh. Mỗi lần dịch chuyển theo một giá trị bước trượt. Tại mỗi bước, bộ lọc sẽ tiến hành tích chập với các điểm ảnh tại vùng ảnh đó. Ta cũng có thể dùng nhiều bộ lọc khác nhau lên ảnh đầu vào để tìm nhiều đặc trưng của ảnh.

$$g(x, y) = \sum_{s=0}^{m-1} \sum_{t=0}^{n-1} w(s, t) f(x+s, y+t) \quad (2.4)$$

Với  $f(x, y)$  ảnh vào,  $w(s, t)$  bộ lọc và  $g(x, y)$  ảnh ngõ ra.

Ngõ vào với kích cỡ cố định  $W_1 * H_1 * D_1$  (rộng x cao x sâu), số lượng bộ lọc K, kích cỡ bộ lọc F, bước trượt (stride) S, số lượng padding P. Cho ra kích thước ngõ ra, với

$$W_2 = \frac{W_1 - F + 2P}{S} + 1, H_2 = \frac{H_1 - F + 2P}{S} + 1, D_2 = K \quad (2.5)$$

Và  $F \times F \times D_1$  là trọng số trên mỗi bộ lọc. Vậy tổng cộng ta có  $(F \times F \times D_1) \times K$  trọng số và K bias. Ngõ ra của lát thứ d (kích cỡ  $W_2 * H_2$ ) là kết quả của quá trình tích chập bộ lọc thứ d trên ảnh ngõ vào với stride S và offset bởi bias thứ d.

### **Lớp hiệu chỉnh tuyến tính (ReLU):**

ReLU là một hàm phi tuyến tính, chức năng của lớp này là chuyển toàn bộ giá trị âm từ kết quả của lớp tích chập thành giá trị 0 mà vẫn giữ sự tin cậy toán học của mạng. Nghĩa là  $\text{Output} = \max(0, \text{Input})$ . Đầu ra của lớp ReLU có kích thước giống với đầu vào, chỉ là tất cả các giá trị âm trong đó được loại bỏ

### **Lớp Pooling:**

Sau khi qua lớp Pooling, ảnh sẽ được giảm kích thước nhưng vẫn giữ được những đặc trưng quan trọng của ảnh. Khác với lớp tích chập, lớp này chỉ tiến hành lấy mẫu (subsampling) thay vì tích chập.

Hình thức phổ biến nhất của một lớp pooling với các bộ lọc có kích thước  $2 \times 2$  được áp dụng với stride 2 sẽ giảm 2 lần mỗi lát cắt sâu trong đầu vào dọc theo cả chiều rộng và chiều cao. Mỗi thao tác MAX trong trường hợp này sẽ lấy tối đa trong

4 số (vùng 2x2 nhỏ trong một số lát cắt sâu). Kích thước chiều sâu vẫn không thay đổi. Ngõ vào với kích cỡ cố định  $W_1 \times H_1 \times D_1$ , kích cỡ bộ lọc  $F$ , bước trượt (stride)  $S$ . Cho kích thước ngõ ra  $W_2 \times H_2 \times D_2$ , với

$$W_2 = \frac{W_1 - F}{S} + 1, \quad H_2 = \frac{H_1 - F}{S} + 1, \quad D_2 = D_1 \quad (2.6)$$

Ví dụ: Khối lượng đầu vào có kích thước 224x224x64 được chập với cửa sổ kích thước 2x2, stride 2 sẽ cho đầu ra có kích thước 112x112x64

### **Lớp kết nối đầy đủ (Fully connected layer)**

Lớp được kết nối đầy đủ là một mạng nơ-ron nhiều tầng sử dụng chức năng kích hoạt softmax trong lớp đầu ra. Kết nối đầy đủ có nghĩa là mọi nơ-ron ở lớp trước được kết nối với mọi nơ-ron ở lớp kế tiếp.

### **Một số thuật toán tối ưu dùng trong huấn luyện mạng nơ-ron**

#### ***Giảm độ dốc:***

Với thuật toán giảm độ dốc (Gradient Descent) [59] ta dự đoán một điểm khởi tạo  $\theta = \theta_0$ , trong đó  $\theta$  là tập hợp các tham số của mô hình. Cập nhật đến khi đạt được kết quả chấp nhận được:

$$\theta = \theta - \eta \nabla_{\theta} J(\theta) \quad (2.7)$$

Với  $\nabla_{\theta} J(\theta)$  là đạo hàm của hàm mất mát tại  $\theta$ ,  $\eta$  là tốc độ học.

Vì ta cần tính toán độ dốc cho toàn bộ tập dữ liệu để thực hiện chỉ một bản cập nhật, việc giảm độ dốc có thể rất chậm và không thể truy cập được đối với các bộ dữ liệu không phù hợp với bộ nhớ. Hiện thuật toán này có rất nhiều biến thể, điển hình là giảm độ dốc ngẫu nhiên (SGD) đã khắc phục được nhược điểm trên.

#### ***Giảm độ dốc ngẫu nhiên***

Giảm độ dốc ngẫu nhiên (SGD) thực hiện cập nhật tham số cho mỗi ví dụ đào tạo  $x^{(i)}$  và nhãn  $y^{(i)}$ :

$$\theta = \theta - \eta \nabla_{\theta} J(\theta; x^{(i)}; y^{(i)}) \quad (2.8)$$

Trong thuật toán này, tại một thời điểm ta chỉ tính đạo hàm của hàm mất mát dựa trên chỉ một điểm dữ liệu  $x^{(i)}$  rồi cập nhật  $\theta$  dựa trên đạo hàm này và tiếp tục thực hiện với từng điểm trên toàn bộ dữ liệu.

### *Độ dốc thích ứng*

Độ dốc thích ứng (Adagrad) [60] là một thuật toán để tối ưu hóa dựa trên độ dốc, nó điều chỉnh tốc độ học tập cho các tham số, thực hiện các cập nhật nhỏ hơn (tức là tốc độ học tập thấp) cho các tham số liên quan đến các tính năng thường xuyên xảy ra và cập nhật lớn hơn (nghĩa là học cao tỷ lệ) cho các tham số liên quan đến các tính năng không thường xuyên. Nó sử dụng tỷ lệ học tập khác nhau cho mỗi tham số  $\theta_i$  tại mỗi thời điểm bước  $t$ . Để cho ngắn gọn, ta sử dụng  $g_t$  để biểu thị độ dốc tại thời điểm bước  $t$  và  $g_{t,i}$  là đạo hàm riêng của hàm mục tiêu có liên quan đến tham số  $\theta_i$  tại thời điểm bước  $t$ .

$$g_{t,i} = \nabla_{\theta} J(\theta_{t,i}) \quad (2.9)$$

Bản cập nhật giảm độ dốc ngẫu nhiên cho mọi tham số  $\theta_i$  tại mỗi thời điểm bước  $t$  là:

$$\theta_{t+1,i} = \theta_{t,i} - \eta \cdot g_{t,i} \quad (2.10)$$

Độ dốc thích ứng đổi tỷ lệ học chung  $\eta$  tại mỗi thời điểm bước  $t$  cho mọi tham số  $\theta_i$  dựa trên các độ dốc trước đây đã được tính cho  $\theta_i$ :

$$\theta_{t+1,i} = \theta_{t,i} - \frac{\eta}{\sqrt{G_{t,ii} + \varepsilon}} \cdot g_{t,i} \quad (2.11)$$

$G_t$  là một ma trận đường chéo, trong đó mỗi phần tử đường chéo  $i$ ,  $i$  là tổng bình phương của độ dốc liên quan đến tham số  $\theta_i$  tại thời điểm bước  $t$ ,  $\varepsilon$  là một hằng số nhỏ bổ sung vào để mẫu tránh bằng 0 (thông thường bằng  $1e^{-8}$ ).

### 2.5.2 Tình hình nghiên cứu

## Chương 2 CƠ SỞ LÝ THUYẾT

---

Trong nghiên cứu năm 2021, Yiqiu Shen và các cộng sự đã đề xuất một hệ thống hỗ trợ từ xa dựa trên TTNT để nhằm làm giảm chẩn đoán sai trong các chẩn đoán bệnh UTV bằng hình ảnh siêu âm. Sử dụng tập dữ liệu NYU Breast Ultrasound, nghiên cứu này đã huấn luyện một mô hình mạng học sâu sử dụng AUROC (Area Under The Receiver Operating Characteristics) và AUPRC (Area Under The Precision-Recall Curve), để đạt độ chính xác cao nhất lên đến 95% [30].

Các kỹ thuật AI, đặc biệt là DL, ngày càng được sử dụng nhiều trong hình ảnh y tế nhờ tiềm năng đầy hứa hẹn và hiệu suất vượt trội trong việc phân tích hình ảnh y tế. Kỹ thuật AI có thể mang lại tốc độ tính toán nhanh với khả năng lặp lại tốt và không gây mệt mỏi, đảm bảo BS có được thông tin khách quan và chính xác cao, từ đó giúp giảm thiểu khối lượng công việc cũng như chẩn đoán sai hoặc chẩn đoán sai [129]. Việc sử dụng hệ thống CAD để sàng lọc UTV đã được khám phá trong nhiều nghiên cứu. Các hệ thống này có thể xác định và phân chia các tổn thương vú một cách đáng tin cậy, trích xuất và phân loại các đặc điểm, ước tính bệnh vú và nguy cơ UTV, đồng thời đánh giá hiệu quả điều trị và tiên lượng bất kể phương pháp hình ảnh y tế nào được sử dụng (ví dụ: chụp nhũ ảnh, siêu âm, MRI hoặc các loại hình ảnh khác) [130]-[134]. Các hệ thống CAD này rất hứa hẹn và mang lại nhiều lợi ích khác nhau trong việc hỗ trợ BS, tối ưu hóa phân bổ nguồn lực và cải thiện độ chính xác.

Chương này đã tiến hành đánh giá toàn diện về những tiến bộ gần đây trong kỹ thuật học máy và học sâu để phát hiện và phân loại UTV. Bằng cách khám phá các phương pháp khác nhau được sử dụng trên các loại hình ảnh y tế khác nhau, nghiên cứu này nhằm mục đích cung cấp sự hiểu biết thấu đáo về các phương pháp hiện tại trong việc xác định UTV. Đặc biệt nhấn mạnh vào các phương pháp học máy và học sâu truyền thống, nêu bật tầm quan trọng của chúng trong lĩnh vực này.

### Chương 3

## THU THẬP VÀ TIỀN XLA X-QUANG VÚ

Chương này tập trung vào việc mô tả các tập dữ liệu được sử dụng trong nghiên cứu, bao gồm cả tập dữ liệu được thu thập tại BV Ung Bướu TP.HCM, cũng như quy trình thu thập dữ liệu. Các vấn đề liên quan đến tiền XLA X-quang vú cũng được thảo luận trong chương này, bao gồm tăng cường hình ảnh và hiệu chỉnh, cũng như các bước tiền xử lý khác.

Các kết quả nghiên cứu trong chương này đã được công bố trong 1 bài báo Hội nghị Quốc tế và 1 bài báo tạp chí trong nước.

[CT1] 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.

[CT6] 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.

### 3.1. THU THẬP DỮ LIỆU ẢNH X-QUANG VÚ

#### 3.1.1 Các tập dữ liệu ảnh X-quang vú sử dụng trong phạm vi nghiên cứu

Ảnh X-quang vú là một công cụ quan trọng trong chẩn đoán và phát hiện sớm UTV, góp phần cải thiện đáng kể kết quả điều trị UTV. Để thúc đẩy nghiên cứu trong lĩnh vực này và phát triển các công cụ chẩn đoán và tiên lượng chính xác hơn, các nhà nghiên cứu dựa vào các tập dữ liệu chụp ảnh X-quang vú quy mô lớn. Các tập dữ liệu này cung cấp nguồn dữ liệu phong phú, bao gồm hình ảnh mô vú bình thường, lành tính và ác tính, cho phép phát triển và đánh giá các thuật toán học máy, hệ thống phát hiện hỗ trợ máy tính và các phương pháp tiếp cận sáng tạo khác.

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

---

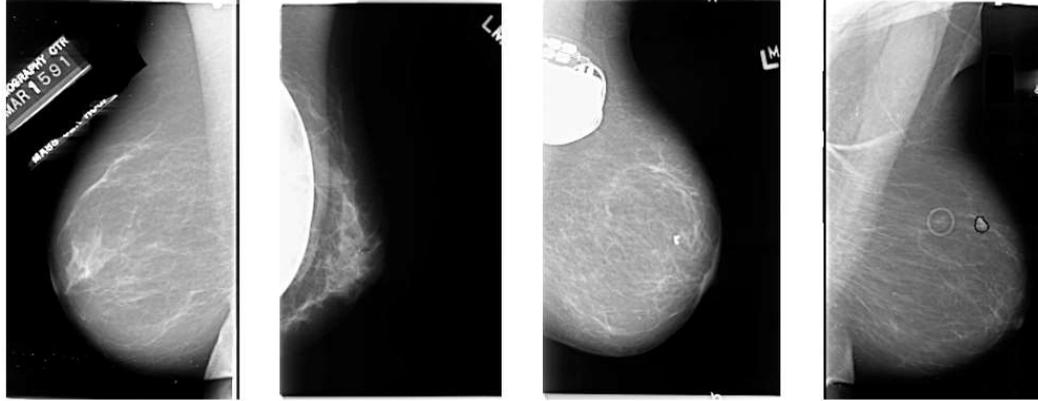
Có nhiều tập dữ liệu chụp ảnh X-quang vú khác nhau, mỗi tập dữ liệu có các đặc điểm và mục đích riêng. Các tập dữ liệu có thể truy cập công khai, chẳng hạn như CSDL kỹ thuật số để sàng lọc chụp ảnh X-quang vú DDSM và tập con của nó là CBIS-DDSM, cung cấp một bộ sưu tập lớn các ảnh chụp ảnh X-quang vú với nhiều mức chú thích khác nhau. Các tập dữ liệu này được sử dụng rộng rãi để đào tạo và thử nghiệm các mô hình học máy, cho phép các nhà nghiên cứu khám phá các phương pháp tiếp cận khác nhau để phát hiện và phân loại UTV. Các tập dữ liệu thương mại, như OPTIMAM và CSDL chụp ảnh X-quang vú Trung Quốc CMMD, cung cấp hình ảnh chất lượng cao hơn và chú thích toàn diện hơn. Các tập dữ liệu này rất phù hợp cho các dự án nghiên cứu cụ thể hoặc hợp tác với các nhà cung cấp dịch vụ chăm sóc sức khỏe. Một số các tập dữ liệu thường được sử dụng phổ biến trong các nghiên cứu về UTV gần đây có thể kể dưới đây:

Tập dữ liệu ảnh UTV MIAS [63]: Tập dữ liệu này gồm hơn 300 ảnh chụp X-quang tuyến vú được chú thích với thông tin về loại mô nền, những bất thường xuất hiện ở vú và mức độ nghiêm trọng của chúng. Ngoài ra, các tổn thương trong chụp quang tuyến vú được đánh dấu bằng tọa độ X và Y, cùng với nhãn cho các bất thường khác nhau [64]. Các đặc điểm chính của tập dữ liệu MIAS được thống kê trong Bảng 3.1.

**Bảng 3.1 Thông tin bộ dữ liệu MIAS**

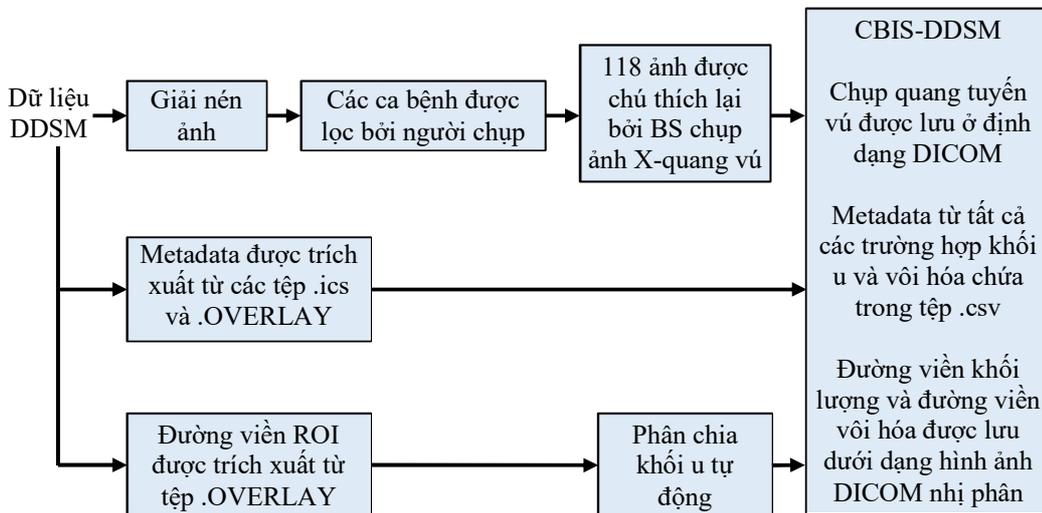
Nội dung	Giá trị	Đơn vị
Tổng số ảnh	322	ảnh
Kích thước ảnh	224 × 224	Pixel
Mã màu	RGB	-
Khối u lành tính	37	ảnh
Vôi hóa dạng lành tính	12	ảnh
Vôi hóa dạng ác tính	13	ảnh
Khối u ác tính	19	ảnh

Ảnh MIAS là ảnh scan từ phim X-quang nên có chất lượng không cao, thể hiện trong Hình 3.1. Các ảnh còn giữ lại các nội dung ghi chú và các thành phần không mong muốn trên ảnh.



Hình 3.1 Một vài ảnh từ tập dữ liệu MIAS

DDSM (CSDL kỹ thuật số về chụp ảnh X-quang vú sàng lọc) [65]: Với hơn 2600 nghiên cứu về chụp ảnh X-quang vú bằng phim được quét, DDSM cung cấp một kho lưu trữ toàn diện. Một tập hợp con của DDSM, được gọi là CBIS-DDSM [66], cung cấp các hình ảnh được chú thích rõ ràng với thông tin bệnh lý chi tiết như loại khối u vú, cấp độ khối u và giai đoạn.



Hình 3.2 Sơ đồ chuẩn bị dữ liệu của tập dữ liệu CBIS-DDSM [66]

CBIS-DDSM là phiên bản cập nhật của DDSM cung cấp dữ liệu để truy cập và phân đoạn ROI được cải thiện [23]. Tập DDSM ban đầu gồm 2.620 ảnh X-quang vú

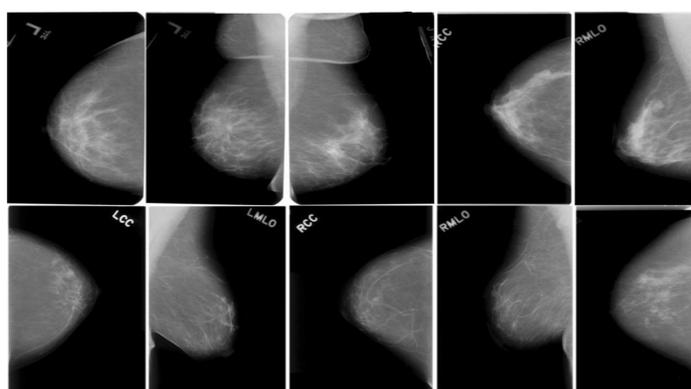
### Chương 3 THU THẬP VÀ TIỀN XỬ LÝ ẢNH X-QUANG VÚ

được chuyển sang ảnh số bằng cách scan phim x-quang, trong khi CBIS-DDSM bao gồm 1.644 hình ảnh của bốn loại tổn thương vú (Khối u lành tính; Khối u ác tính; Vôùi hóa dạng lành tính; Vôùi hóa dạng ác tính). Việc lọc ảnh từ DDSM sang CBIS-DDSM được thực hiện theo quy trình thể hiện trong Hình 3.2. Số lượng ảnh trong mỗi loại và thông tin của chúng được thể hiện trong Bảng 3.2.

**Bảng 3.2 Thông tin bộ dữ liệu CBIS-DDSM**

Nội dung	Giá trị	Đơn vị
Tổng số ảnh	1644	ảnh
Kích thước ảnh	224 × 224	điểm ảnh
Mã màu	RGB	-
Khối u lành tính	753	ảnh
Vôi hóa dạng lành tính	414	ảnh
Vôi hóa dạng ác tính	339	ảnh
Khối u ác tính	419	ảnh

Một số hình ảnh từ tập CBIS-DDSM được trình bày trong Hình 3.3. Đây cũng là hình ảnh scan từ phim X-quang, tuy nhiên đã qua chọn lọc và xử lý nên có chất lượng tốt hơn tập MIAS.

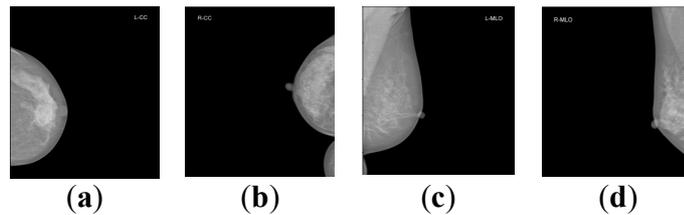


**Hình 3.3. Một số ảnh của tập dữ liệu CBIS-DDSM**

Bộ dữ liệu VinDr-Mammo [135]: trong đó các ảnh chụp nữ ảnh được thu thập từ năm 2018 đến năm 2020. Đặc biệt, các bộ ảnh này được lưu trữ tại Hệ thống PACS tại BV Đại học Y Hà Nội (ĐHYHN) và BV 108 (H108). Ngoài ra, các bộ ảnh này có 5.000 lần chụp nữ ảnh, tương ứng với 20.000 ảnh chụp nữ ảnh, vì mỗi lần chụp tạo

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

ra 4 ảnh vú bao gồm: Right Cranial-Caudal (CC), Left CC, Right Mediolateral-Oblique (MLO), Left MLO như mô tả trong Hình 3.4. Điều quan trọng là thông tin cá nhân của bệnh nhân trên ảnh luôn được bảo mật và các ảnh này chỉ lưu giữ các đối tượng liên quan đến tổn thương. Do đó, 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. Số lượng ảnh trong bộ dữ liệu VinDr-Mammo với các loại tổn thương khác nhau là không cân bằng, vì vậy chúng tôi chỉ chọn 4 loại tổn thương và một loại không có trong nghiên cứu này như mô tả trong Bảng 1.



**Hình 3.4. Một bộ ảnh mẫu của tập dữ liệu VinDr-Mammo: (a) Left CC; (b) Right CC; (c) Left MLO; (d) Right MLO**

**Bảng 3.3 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

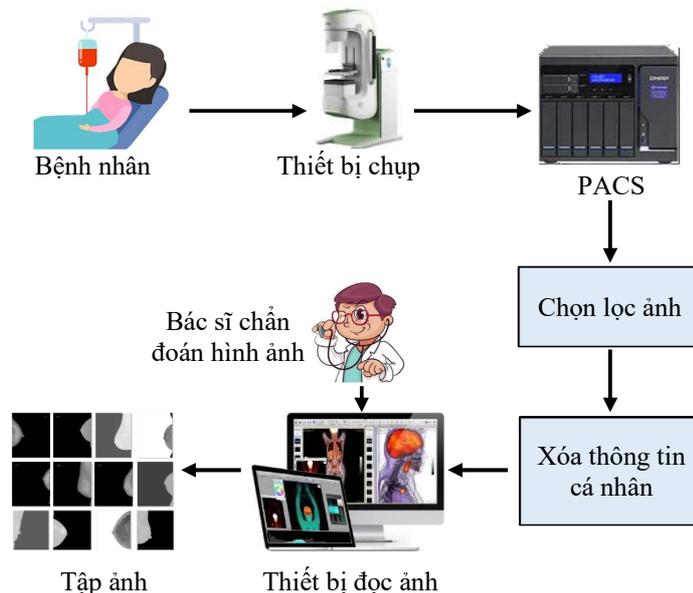
Số lượng ảnh của mỗi loại tổn thương trong tập ảnh này không đồng đều, trong đó ảnh không có tổn thương có số lượng nhiều nhất. Trong số các loại tổn thương thì tổn thương loại khối u chiếm số lượng lớn nhất, các tổn thương biến dạng cấu trúc có số lượng ít nhất.

#### 3.1.2 Thu thập ảnh X-quang vú tại BV Ung bướu

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

Để phục vụ cho luận án, nghiên cứu sinh đã đệ trình kế hoạch nghiên cứu tại hội đồng đạo đức trong nghiên cứu y sinh tại bệnh viện Ung Bướu TP.HCM và được Hội đồng chấp thuận tại Quyết định số 74/BVUB-HĐĐĐ ngày 20/1/2025 để thực hiện nghiên cứu và thu thập hình ảnh X-quang vú tại khoa Chẩn đoán hình ảnh của BV (thông tin về giấy chấp nhận của Hội đồng Đạo đức tại Phụ lục).

Trong nghiên cứu, các hình ảnh chụp ảnh X-quang vú DICOM thu thập từ BV 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, ảnh ảnh X-quang vú DICOM được thu thập từ thiết bị lưu trữ hình ảnh PACS của BV và được ẩn danh để bảo vệ quyền riêng tư của bệnh nhân. Tiếp theo, tập dữ liệu được một số BS chẩn đoán hình ảnh chú thích thông qua công cụ gắn nhãn trên máy tính của BV. Cuối cùng, hình ảnh đã được chú 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

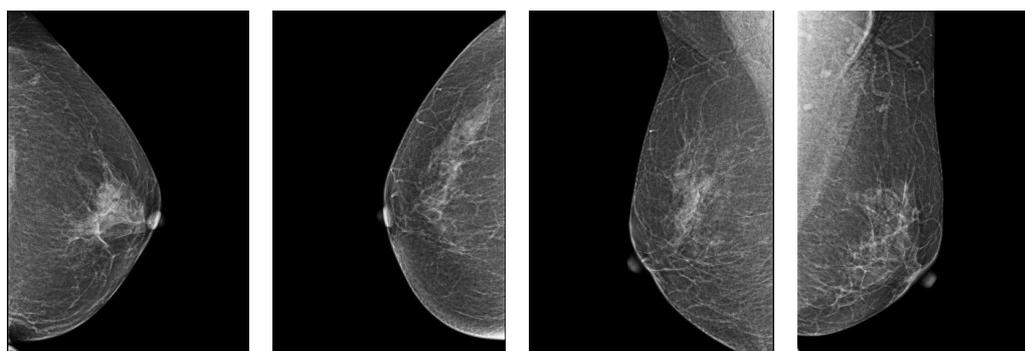
Trong bước này, hơn 4000 ảnh ảnh X-quang vú ở định dạng DICOM từ Hệ thống PACS của hệ thống DICOM của BV Ung Bướu TP.HCM, với các lần chụp trong

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

---

khoảng từ đầu năm 2018 đến cuối năm 2023 được thu thập. Hình ảnh được thu thập trên các thiết bị chủ yếu từ nhà cung cấp là SIEMENS và GE HEALTHCARE. Để đảm bảo các quy định về quyền riêng tư, 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. Bên cạnh siêu dữ liệu DICOM, thông tin liên quan có thể xuất hiện trong hình ảnh, chẳng hạn như góc chụp và góc nhìn của hình ảnh và đôi khi là tên của bệnh nhân.

Một số hình ảnh đã thu thập từ BV Ung Bướu TP.HCM được trình bày ở Hình 3.6, tất cả đều là ảnh số thu thập từ hệ thống DICOM và đã loại bỏ thông tin cá nhân.



**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

Đọc ảnh X-quang vú: Kết quả đọc được đánh giá các tổn thương bao gồm khối lượng, vôi hóa, bất đối xứng, biến dạng cấu trúc và các đặc điểm liên quan khác. Mỗi phát hiện được đánh dấu bằng một hộp giới hạn để xác định vị trí phát hiện bất thường.

### 3.1.3 Gán nhãn ảnh X-quang vú

Đối với tập dữ liệu tại BV ung bướu, việc gán nhãn được thực hiện bởi các BS chẩn đoán hình ảnh của BV Ung Bướu TP.HCM. Sau khi hình được chụp bởi kỹ thuật viên, BS chẩn đoán hình ảnh sẽ tiến hành phân tích ảnh X-quang một cách hệ thống, thường theo một trình tự nhất định để không bỏ sót bất kỳ chi tiết nào, và gán nhãn theo trình tự như sau:

- BS đánh giá tổng thể chất lượng của ảnh, đảm bảo ảnh chụp đúng tiêu chuẩn kỹ thuật (độ tương phản, độ sáng, tư thế chụp). Nếu ảnh chụp không đủ tiêu chuẩn để tiến hành chẩn đoán, BS sẽ yêu cầu chụp lại.

- Kiểm tra đối xứng: BS sẽ so sánh hai bên vú để tìm kiếm sự khác biệt bất thường về mật độ, cấu trúc. Bước này cho phép BS đưa ra các đánh giá về các loại tổn thương như bất đối xứng, biến dạng cấu trúc.

- Tìm kiếm các khối u (Masses): BS sẽ xác định xem có khối u hay không và phân tích các chi tiết của khối u, bao gồm: Hình dạng (khối u có hình tròn, bầu dục hay không đều); Bờ (Bờ của khối u là một đặc điểm rất quan trọng. Bờ rõ, đều thường là dấu hiệu của u lành tính, trong khi bờ gai, tua, hoặc không rõ ràng có thể có nguy cơ u ác tính); Mật độ (Mật độ của khối u so với mô vú xung quanh)

- Phát hiện vi vôi hóa (Microcalcifications): Đây là những đốm vôi hóa rất nhỏ, khó nhìn thấy bằng mắt thường, cần phóng to ảnh để tìm kiếm và phân tích. Việc phân tích các chi tiết của vôi hóa thường bao gồm: Hình dạng và kích thước (Các đốm vôi hóa có hình dạng tròn, bầu dục hay không đều); Sự phân bố (Các đốm vôi hóa nằm rải rác, tập trung thành cụm hay theo đường ống tuyến sữa. Sự phân bố theo cụm hoặc theo đường ống có thể là dấu hiệu của ung thư).

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

---

- Tìm kiếm các dấu hiệu khác: các dấu hiệu gián tiếp của ung thư như co kéo da, núm vú, dày da hay thay đổi kiến trúc mô vú có thể bổ sung thêm các chứng cứ cho kết luận của BS.

Bên cạnh việc phân tích hình ảnh, BS cũng sẽ xem xét một số thông tin quan trọng của bệnh nhân, bao gồm tiền sử bệnh; yếu tố nguy cơ; triệu chứng lâm sàng; kết quả chụp trước đây (trong trường hợp bệnh nhân đã từng chụp X-quang vú trước đó). Sau khi phân tích kỹ lưỡng, BS sẽ sử dụng hệ thống BIRADS để đưa ra kết luận và đề xuất hướng xử lý tiếp theo. Đây là hệ thống phân loại chuẩn quốc tế, giúp thống nhất cách báo cáo kết quả và quản lý bệnh nhân, gồm có các cấp độ sau:

- BIRADS 0: Không có đủ thông tin để chẩn đoán.
- BIRADS 1: Âm tính, không có dấu hiệu bất thường.
- BIRADS 2: Lành tính.
- BIRADS 3: Khả năng lành tính cao, cần theo dõi định kỳ.
- BIRADS 4: Nghi ngờ ác tính. Cần sinh thiết để chẩn đoán xác định. Trong đó 4a là khả năng thấp, 4b là khả năng trung bình, và 4c là khả năng cao.
- BIRADS 5: Khả năng ác tính rất cao. Cần sinh thiết.
- BIRADS 6: Đã xác định là ác tính qua sinh thiết.

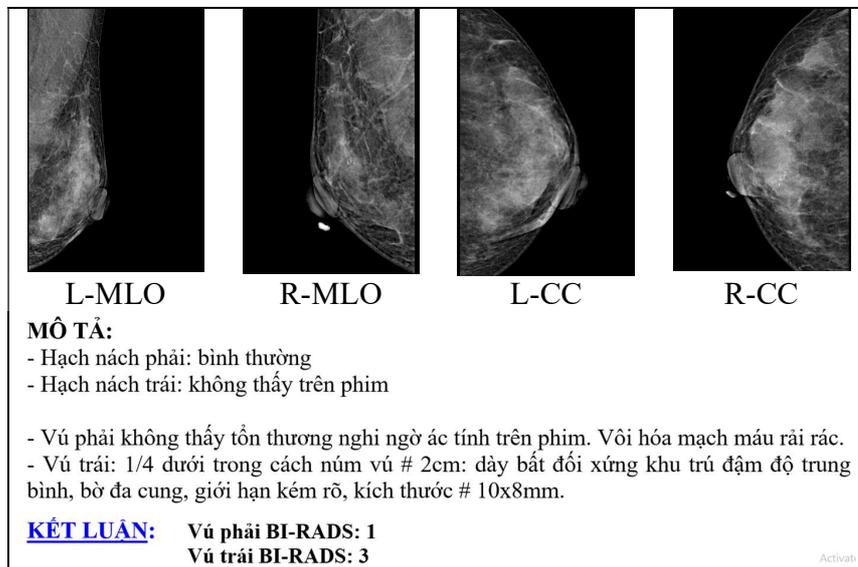
Giữa phân loại nguy cơ ung thư theo thang đo BIRADS và các loại tổn thương của ảnh X-quang vú có mối liên hệ chặt chẽ và được đánh giá bởi các bác sĩ chuyên khoa.

**Bảng 3.5 Các loại tổn thương trên nhũ ảnh**

Thành phần mô vú	A. Hoàn toàn là mô mỡ B. Rải rác các vùng mô tuyến – mô sợi C. Mật độ không đồng nhất, có thể che khuất các khối bất thường D. Mật độ rất cao, làm giảm độ nhạy phát hiện	
Khối	Hình dạng	hình bầu dục – tròn – không đều
	Bờ	rõ nét – mờ – thùy nhỏ – không rõ – tua gai
	Mật độ	mỡ – thấp – bằng – cao

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

Bất đối xứng	bất đối xứng – toàn bộ – khu trú – đang phát triển (mới xuất hiện / tăng dần)		
Biến dạng cấu trúc	mô tuyến vú bị biến dạng mà không thấy khối rõ ràng		
Vi vôi hóa	Hình thái	Thường lành tính	
		Khả nghi	<ol style="list-style-type: none"> <li>1. vô định hình</li> <li>2. thô (coarse)</li> <li>3. không đồng nhất</li> <li>4. nhỏ dạng đa hình (fine pleomorphic)</li> <li>5. nhỏ dạng nhánh hoặc tuyến tính mảnh (fine linear / fine linear branching)</li> </ol>
	Phân bố	lan tỏa – vùng – tụ nhóm – tuyến tính – phân đoạn	
Các đặc điểm kèm theo	kéo tụ da - tụ núm vú - dày da - dày bì – dày nhu mô - hạch nách - biến dạng kiến trúc - vi vôi hóa		



Hình 3.7 Ví dụ một kết luận chẩn đoán của BS và ảnh X-quang vú tương ứng.

Đối với tập dữ liệu thu thập từ BV Ung Bướu, chỉ những hình ảnh đã được chẩn đoán mới được đưa vào tập dữ liệu. Trong đó 100% các ảnh đều có kết luận chẩn đoán của bác sỹ chẩn đoán hình ảnh kết luận về:

- Phân loại tổn thương.
- Phân loại nguy cơ ung thư (theo thang đo BIRADS)

Đối với các trường hợp có phân loại BIRADS từ 4 trở lên, có kèm theo kết quả sinh thiết (bằng các kỹ thuật chọc hút bằng kim nhỏ (FNA) và sinh thiết kim lõi (CNB)). Tất cả các thông tin trên cho thấy việc gán nhãn cho tập dữ liệu này có cơ sở chuyên môn vững chắc và đã được xác nhận. Các dữ liệu này là căn cứ để thực hiện huấn luyện mô hình.

#### 3.2. TĂNG CƯỜNG HÌNH ẢNH

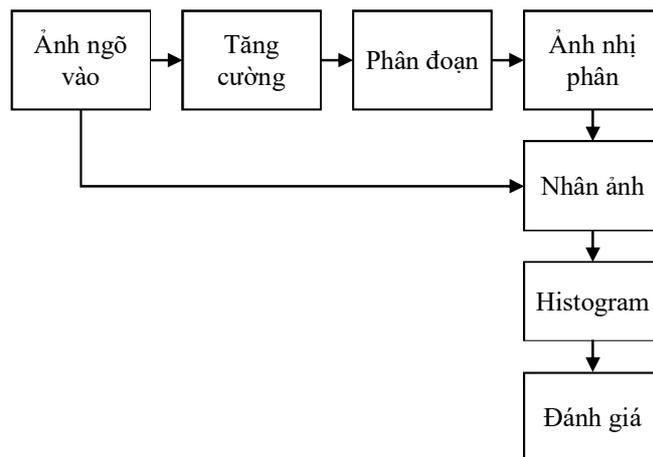
Tăng cường ảnh là công đoạn điều chỉnh hình ảnh kỹ thuật số để kết quả phù hợp hơn cho việc hiển thị hoặc phân tích hình ảnh sâu hơn. Ví dụ: bạn có thể loại bỏ nhiễu, làm sắc nét hoặc làm sáng hình ảnh, giúp xác định các đặc điểm chính dễ dàng hơn.

Đối với ảnh UTV, các kỹ thuật tăng cường hình ảnh được áp dụng được chia ra làm 4 loại [92]: các kỹ thuật truyền thống, các kỹ thuật tăng cường theo vùng (region-based), các kỹ thuật tăng cường dựa trên đặc trưng ảnh (feature based), và các kỹ thuật tăng cường khác. Các kỹ thuật tăng cường truyền thống thường có thể đáp ứng cho việc tăng cường dù là từng phần hay là toàn bộ ảnh X-quang vú. Tuy nhiên, ngoài việc điều chỉnh độ tương phản, các kỹ thuật này cũng có xu hướng làm gia tăng nhiễu. Do đó, để tăng cường chất lượng các chi tiết của hình ảnh mà không kèm theo bất kì ảnh hưởng tiêu cực nào, các phương pháp tăng cường theo vùng được phát triển. Các kỹ thuật này khá phù hợp với việc tăng cường các vùng vi vôi hoá ở ảnh X-quang vú khi mô dày đặc. Ngoài ra, đối với ảnh X-quang vú với cả vi vôi hoá và khối u, những kỹ thuật tăng cường hình ảnh dựa trên đặc trưng cũng có thể được áp dụng. Ở những kỹ thuật này, thông thường ảnh X-quang vú sẽ được biến đổi sang vùng tần số với các bộ biến đổi đa mức như wavelets, sau đó lọc đi những vùng tần số thấp. Ảnh X-quang vú sẽ được xây dựng lại chỉ với những vùng có tần số cao chứa đựng những

khu vực có khả năng bị tổn thương. Với nguyên lý tối đa hoá entropy mờ, các kỹ thuật tăng cường mờ cũng có thể được áp dụng trên ảnh X-quang vú để tăng cường độ tương phản và triệt tiêu nhiễu. Các kỹ thuật này khá hiệu quả trong việc tăng cường các đường viền khối và làm rõ hơn các chi tiết trong ảnh X-quang vú. Tuy nhiên gần đây, sự phát triển của công nghệ kỹ thuật số với dải động cao, độ tương phản và tỷ lệ tín hiệu trên nhiễu được cải thiện một cách trực tiếp làm cho chất lượng ảnh X-quang vú ngày càng được tăng cao đồng thời cũng thu hẹp không gian nghiên cứu cho các kỹ thuật tăng cường ảnh.

### 3.2.1 Điều chỉnh độ sáng và tương phản ảnh UTV

#### *Thuật toán tăng cường ảnh*



**Hình 3.8** Sơ đồ khối của quá trình đánh giá ung thư

Trong phần này, nguồn tập dữ liệu được thu thập từ Mammographic Image Analysis Society (MIAS). Hình ảnh ban đầu thường khó chẩn đoán chính xác tình trạng hình ảnh vú bị ung thư hay bình thường. Do đó, hình ảnh có thể được nâng cao trước khi phân đoạn để phát hiện ROI và sau đó thống kê để đánh giá tình trạng bệnh như trong Hình 3.8 [77].

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 [136]. Trong nghiên cứu, các giá

trị nhỏ nhất, lớn nhất và trung bình của các pixel của một hình ảnh mức xám được tính toán và sau đó hai ngưỡng  $Th_1$ ,  $Th_2$  được tính như sau:

$$Th_1 = (mean + min)/2 \quad (3.1)$$

$$Th_2 = (mean + max)/2 \quad (3.2)$$

Từ hai giá trị ngưỡng, hình ảnh mức xám được chia thành bốn nhóm để xác định bốn ngưỡng tương ứng như sau:

$$min \leq P_0 \leq Th_1$$

$$Th_1 \leq P_1 \leq mean$$

$$mean \leq P_2 \leq Th_2$$

$$Th_2 \leq P_3 \leq max$$

trong đó,

$P_0, P_1, P_2, P_3$  là các ngưỡng mức xám mới

min biểu thị giá trị mức xám nhỏ nhất trong hình ảnh

max là giá trị mức xám lớn nhất

nghĩa là mô tả giá trị trung bình mức xám

Mức xám của pixel trong mỗi nhóm được điều chỉnh và sau đó tất cả pixel trong hình ảnh được tính toán để tương ứng với mức xám bằng cách áp dụng các phương trình dưới đây:

$$P_{N0} = 2 \times \left( \frac{P_0 - min}{mean - min} \right)^2 \quad (3.3)$$

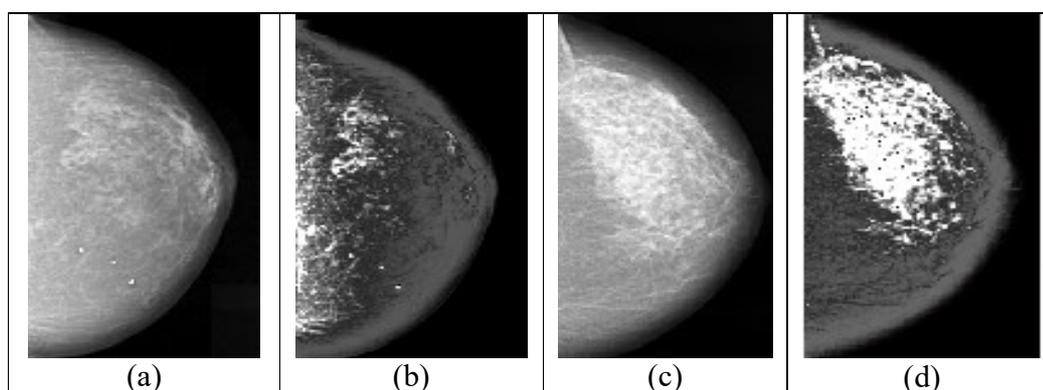
$$P_{N1} = 1 - 2 \times \left( \frac{P_1 - mean}{mean - min} \right)^2 \quad (3.4)$$

$$P_{N2} = 1 - 2 \times \left( \frac{P_2 - mean}{max - min} \right)^2 \quad (3.5)$$

$$P_{N3} = 2 \times \left( \frac{P_3 - mean}{max - mean} \right)^2 \quad (3.6)$$

trong đó  $P_{N0}$ ,  $P_{N1}$ ,  $P_{N2}$ ,  $P_{N3}$  là bốn ngưỡng mới được tính toán căn cứ vào các ngưỡng trước đó và các giá trị tối thiểu, tối đa và trung bình trong hình ảnh vú.

Tất cả các hình ảnh vú được nâng cao và mức độ tương phản được điều chỉnh. Rõ ràng là hình ảnh vú sau khi nâng cao để phân đoạn hình ảnh tốt hơn để tìm kiếm vùng ROI.



(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**

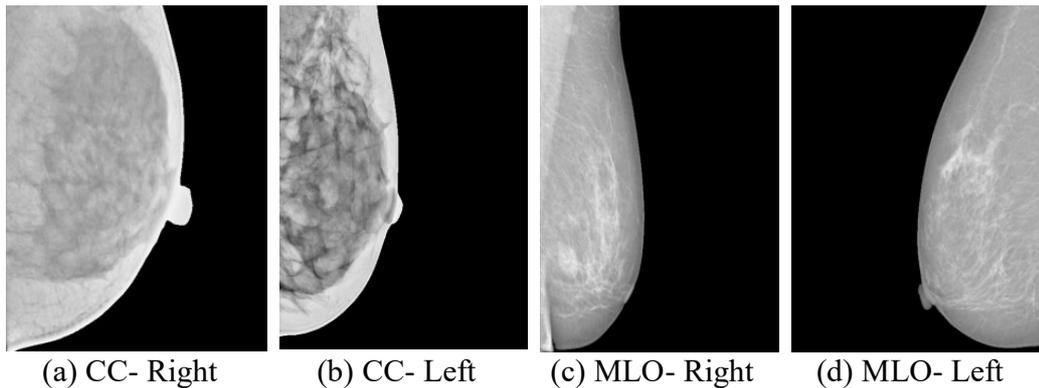
#### 3.2.2 Hiệu chỉnh Gamma

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. Bằng cách thay đổi giá trị gamma để thích nghi với độ contrast chung, có thể đạt được độ sáng tối ưu để phát hiện vật thể chính xác hơn trên hình ảnh gốc [137]. Khi giá trị gamma tăng, các pixel tối của hình ảnh sẽ trở nên tối hơn và các pixel sáng sẽ trở nên sáng hơn. Hiệu ứng này không giống nhau đối với tất cả các pixel vì nó cân bằng tông màu sáng và tối tổng thể. Vì mục đích của quá trình tiền xử lý là làm nổi bật tổn thương vú trên nền tối nên hiệu chỉnh gamma được áp dụng để phân bố các vùng sáng và tối tốt hơn. 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

### 3.2.3 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].

Các phương trình và công thức toán học của CLAHE được tác giả nghiên cứu này giải thích ngắn gọn [139]. Cụ thể, các điểm ảnh trong ảnh tương phản được phân bố lại đồng đều để ảnh có độ sáng tốt hơn. Do đó, thuật toán CLAHE được áp dụng cho hình ảnh đầu vào  $f(x, y)$ , được chia thành các ô có kích thước  $n \times m$  để xác định số

điểm ảnh trung bình đối với mỗi mức xám. Tiếp theo, giá trị giới hạn clip  $N_{CL}$  được tính bằng biểu thức sau:

$$N_{CL} = CL_{nor} \times \frac{n \times m}{N_g} \quad (3.7)$$

Trong đó,  $CL_{nor}$  biểu thị giới hạn clip được chuẩn hóa ( $0 \leq CL_{nor} \leq 1$ ),  $N_g$  là số mức xám.

Do đó, chiều cao biểu đồ của mỗi ô  $h_i$  được tính toán bằng cách áp dụng biểu thức sau:

$$h_i = \begin{cases} N_{CL}, & \text{if } N_i \geq N_{CL} \\ N_i, & \text{else} \end{cases}, i = 0, 1, 2, 3, \dots, L - 1 \quad (3.8)$$

trong đó,  $N_i$  là biểu đồ Histogram của ô thứ  $i$ ,  $L$  là số mức độ xám trong một hình ảnh.

Vì vậy, số lượng điểm ảnh bị cắt có thể được xác định như sau:

$$N_c = n \times m - \sum_{i=0}^{L-1} h_i \quad (3.9)$$

Số lượng điểm ảnh được phân phối lại có thể được tính với biểu thức sau:

$$N_r = \frac{n \times m - \sum_{i=0}^{L-1} h_i}{L-1} \quad (3.10)$$

Cuối cùng, một hình ảnh mới được tạo ra từ các kết quả trên bằng cách chuyển đổi cường độ ánh sáng của từng điểm ảnh tương ứng với biểu đồ tích lũy của vùng ngữ cảnh  $p_c(x, y)$  bằng cách sử dụng công thức sau:

$$p_c(x, y) = \frac{L-1}{n \times m} \sum_j^K n_j \quad (3.11)$$

trong đó, chuẩn hóa của biểu đồ xén histogram  $n_j$  là:

$$n_j = \begin{cases} N_{CL}, & \text{if } N_j + N_r \geq N_{CL} \\ N_j + N_r, & \text{else} \end{cases}, i = 0, 1, 2, 3, \dots, L - 1 \quad (3.12)$$

Từ các biểu thức trên, thuật toán của CLAHE được thể hiện như sau:

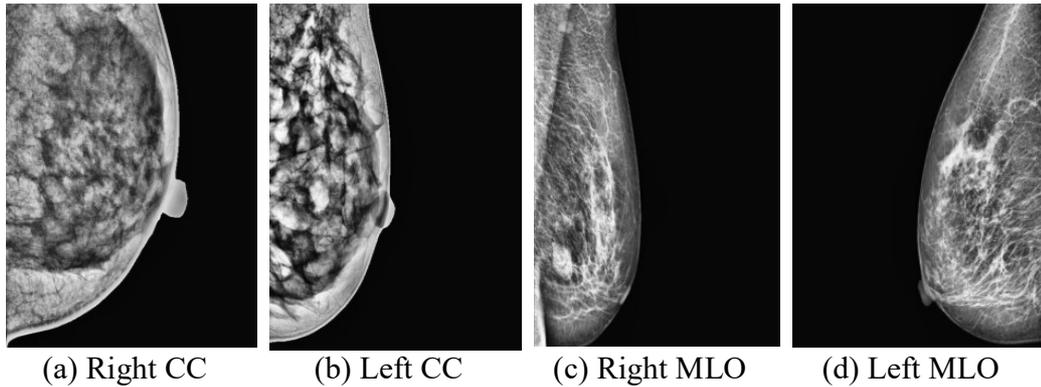
**Thuật toán 3. Cân bằng Histogram thích ứng giới hạn tương phản**

---

```
1  START
2  READ input image  $f(x,y)$ 
   Split  $f(x,y)$  into  $n \times m$  tiles
3  Calculate clip limit  $N_{CL}$ 
4  Apply clip limit to the histogram
5  Calculate the number of pixels to be redistributed
6  Calculate the new value for each pixel based on
   the cumulative histogram
7  OUTPUT  $h(x,y)$ 
8  END
```

---

Kết quả của bước này được thể hiện ở Hình 3.11. 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**

### 3.3. TÔ MÀU CHO ẢNH X-QUANG VÚ

#### 3.3.1 Chuyển sang ảnh màu sử dụng Colormap

Ảnh X-quang vú là ảnh xám có 1 kênh màu, trong khi đó các mạng học sâu như EfficientNet đều được huấn luyện trước với CSDL ảnh phổ biến là tập ảnh màu ImageNet và phân loại hiệu quả trên các ảnh màu có 3 kênh màu. Vì vậy, trong nghiên cứu này, để sử dụng hiệu quả khả năng của mô hình học sâu và việc phân loại chính xác hơn, chúng tôi sử dụng colormap để tô màu cho ảnh xám dựa trên ảnh X-quang gốc. Phương pháp này giúp làm nổi bật khối u, mô dày đặc và các cấu trúc xung quanh nhờ sử dụng các dải màu khác nhau. Có nhiều bộ lọc màu colormap khác nhau trong thư viện Matplotlib, chúng tôi đã thử nghiệm các colormap khác nhau để tìm ra

bộ lọc phù hợp nhất cho tập dữ liệu của mình. ‘JET’ là bộ lọc colormap hoạt động tốt nhất theo các thử nghiệm này.

Colormap là một kỹ thuật nhằm trực quan hóa dữ liệu trong Matplotlib. Cho phép thể hiện các biến thể dữ liệu một cách trực quan. Đây là bảng phân tích các bản đồ màu trong Matplotlib. Các loại colormap phổ biến là:

- Tuần tự (Sequential): Chuyển màu từ một màu (thường là giá trị thấp) sang màu khác (thường là giá trị cao), phù hợp với dữ liệu liên tục như nhiệt độ hoặc độ cao (ví dụ các bộ lọc: viridis, plasma).
- Phân kỳ (Diverging): Chuyển đổi từ màu trung điểm sang hai màu đối lập ở hai bên, lý tưởng cho dữ liệu có giá trị trung tính ở giữa và độ lệch dương hoặc âm (ví dụ: RdBu, bwr).
- Chu kỳ (Cyclic): Lặp lại các chuỗi màu, hữu ích để biểu diễn dữ liệu định kỳ (ví dụ: hsv, coolwarm).
- Định tính (Qualitative): Màu sắc riêng biệt để biểu thị dữ liệu phân loại (ví dụ: tab10, Set2).

Trong đó JET là một trong những bộ lọc có quang phổ chuyển từ xanh dương sang đỏ. Thuật toán của phương pháp tô màu sử dụng colormap này được thể hiện dưới đây.

---

**Thuật toán 4. Chuyển sang ảnh màu sử dụng Colormap**

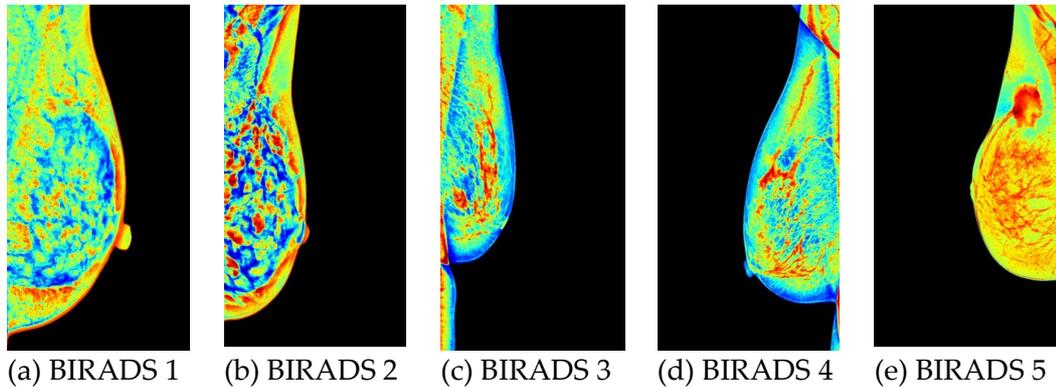
---

```
1 START
2   READ source image  $I_{IN}(x,y)$  where  $x$  and  $y$  denotes the
   pixel coordinates
3   Segment input image into  $N$  regions  $S_i, i=1, \dots, N$  based
   on grayscale thresholds
4   FOR  $i = 1$  to  $N$ 
5     Colorize  $S_i$  by the corresponding color  $C_i$ 
6   END FOR
7   OUTPUT  $I_{OUT}(x,y)$ 
8 END
```

---

Hình 3.12 cho thấy kết quả colormap đối với 5 loại ảnh với các mức BIRADS khác nhau làm nổi bật các dấu hiệu của vùng tổn thương trên ảnh UTV. Các ảnh

BIRADS 4,5 cho thấy các vùng màu rõ ràng, khác biệt với các vùng mô bình thường, làm nổi bật các điểm khác biệt của các trường hợp có nguy cơ UTV cao.



Hình 3.12 Kết quả colormap đối với 5 loại ảnh với các mức BIRADS (1,2,3,4,5)

### 3.3.2 Hiệu quả của việc tô màu ảnh X-quang vú

Bảng 3.5 cho thấy kết quả cho thấy các phương pháp tiền XLA được đề xuất làm tăng đáng kể độ chính xác tổng thể của quá trình phân loại các tập ảnh tổn thương vú. Cụ thể, tập ảnh bằng GROI tạo ra độ chính xác lớn hơn, đạt 82,3% so sánh 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 Effi-cientNet-B7 rất phù hợp với tập dữ liệu hình ảnh có CROI.

Bảng 3.6 So sánh độ chính xác phân loại giữa ảnh được tô màu và ảnh xám

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

Thí nghiệm được tiến hành sử dụng tập dữ liệu VinDrMammo với mạng học sâu EfficientNetB7, các chi tiết của thí nghiệm và mạng học sâu được trình bày ở mục 5.2.4 trong chương 5.

## 3.4. Xử lý sự mất cân bằng dữ liệu và vai trò của tăng cường dữ liệu

### 3.4.1 Đặc điểm mất cân bằng dữ liệu trong bài toán phân loại ảnh X-quang vú

Trong các tập dữ liệu ảnh X-quang vú được sử dụng trong luận án (MIAS, CBIS-DDSM, VinDr-Mammo và dữ liệu thu thập tại bệnh viện), hiện tượng mất cân bằng dữ liệu (class imbalance) là vấn đề phổ biến và mang tính bản chất. Cụ thể, số lượng ảnh bình thường (normal) thường chiếm tỷ lệ lớn hơn đáng kể so với các ảnh có tổn thương. Trong nhóm ảnh có tổn thương, các lớp như vi vôi hóa ác tính, bất đối xứng cục bộ hoặc BIRADS 4–5 thường có số lượng mẫu rất hạn chế so với các lớp phổ biến hơn như khối u lành tính hoặc BIRADS 1–2. Ở bài toán phân loại đa lớp và phân loại theo BIRADS, sự chênh lệch số lượng mẫu giữa các lớp có thể lên đến 1:5 hoặc 1:10.

Nếu không có chiến lược xử lý phù hợp, mô hình học sâu có xu hướng thiên lệch về các lớp chiếm ưu thế, dẫn đến độ chính xác tổng thể (accuracy) cao nhưng độ nhạy (recall) đối với các lớp quan trọng về mặt lâm sàng lại thấp, đặc biệt là các tổn thương ác tính.

### **3.4.2 Chiến lược tăng cường dữ liệu (Data Augmentation) được áp dụng**

Luận án áp dụng tăng cường dữ liệu có kiểm soát (controlled augmentation) như một giải pháp chính để giảm thiểu ảnh hưởng của mất cân bằng dữ liệu, thay vì chỉ dựa vào việc thu thập thêm dữ liệu – vốn rất khó khăn trong bối cảnh y tế.

#### **3.4.2.1. Các kỹ thuật tăng cường hình học (Geometric augmentation)**

Các phép biến đổi hình học được áp dụng bao gồm: Lật ảnh theo phương ngang/dọc (horizontal/vertical flip), Xoay ảnh với góc nhỏ ( $\pm 5^\circ$  đến  $\pm 15^\circ$ ), Tịnh tiến và co giãn nhẹ (shift, zoom). Các phép biến đổi này không làm thay đổi bản chất hình thái tổn thương, đồng thời phản ánh sự đa dạng thực tế của tư thế chụp và vị trí mô vú giữa các bệnh nhân.

#### **3.4.2.2. Tăng cường cường độ và tương phản (Intensity-based augmentation)**

Đối với ảnh X-quang vú, luận án đặc biệt chú trọng các phép điều chỉnh độ sáng – tương phản, hiệu chỉnh gamma, thay đổi nhẹ phân bố mức xám trước khi áp dụng colormap. Nhóm phép biến đổi này giúp mô hình tăng khả năng thích nghi với sự

khác biệt về thiết bị chụp, liều tia và điều kiện thu nhận ảnh giữa các bệnh viện và các bộ dữ liệu khác nhau.

#### **3.4.2.3. Chiến lược tăng cường có chọn lọc theo lớp (Class-aware augmentation)**

Khác với tăng cường dữ liệu đồng đều cho mọi lớp, luận án áp dụng chiến lược tăng cường mạnh hơn cho các lớp thiểu số (ví dụ: vi vôi hóa ác tính, BIRADS 4–5) và hạn chế tăng cường quá mức đối với các lớp chiếm ưu thế (normal, BIRADS 1–2). Cách tiếp cận này giúp giảm chênh lệch số lượng mẫu giữa các lớp trong tập huấn luyện, đồng thời tránh hiện tượng mô hình bị “ghi nhớ” các mẫu lặp lại của lớp thiểu số (overfitting do oversampling).

Chiến lược xử lý mất cân bằng dữ liệu được đề xuất trong luận án cho thấy tăng cường dữ liệu không chỉ là bước hỗ trợ kỹ thuật, mà đóng vai trò trung tâm trong việc nâng cao độ nhạy phát hiện các tổn thương quan trọng về mặt lâm sàng; giúp mô hình học sâu tổng quát hóa tốt hơn khi triển khai trên dữ liệu thực tế tại bệnh viện; phù hợp với bối cảnh y tế Việt Nam, nơi dữ liệu ác tính thường khan hiếm và khó thu thập. Cách tiếp cận này góp phần đảm bảo rằng hệ thống AI không chỉ đạt hiệu năng cao về mặt thống kê, mà còn đáp ứng yêu cầu an toàn và tin cậy trong hỗ trợ chẩn đoán y khoa.

Các kết quả đạt được trong chương 3 tập trung vào các thuật toán tiền xử lý ảnh, cần thiết và đã chứng minh được tầm quan trọng trong việc nâng cao hiệu suất phân loại của các mạng học sâu. Việc thu thập dữ liệu hình ảnh UTV cũng đã được thực hiện với 2 nguồn là các tập dữ liệu nguồn mở và dữ liệu từ BV Ung Bướu TP.HCM. Các kết quả này là cơ sở quan trọng cho việc thực hiện các bước tiếp theo trong nghiên cứu.

## Chương 4

# TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

Chương này tập trung vào việc trích xuất đặc trưng hình ảnh và xử lý vùng ROI trong phân loại UTV. Nội dung chính bao gồm: (i) trình bày vai trò quan trọng của đặc trưng ảnh trong việc hỗ trợ chẩn đoán, giảm tính chủ quan và làm đầu vào cho các mô hình mạng máy học, học sâu; (ii) ứng dụng các phương pháp trích đặc trưng như GLCM để mô tả cấu trúc mô vú và hỗ trợ phân biệt khối u lành tính, ác tính; (iii) đề xuất các kỹ thuật phân đoạn ảnh nhằm loại bỏ thành phần không mong muốn, xác định và tăng cường vùng ROI; (iv) phát triển phương pháp tô màu ROI (CROI) nhằm tận dụng hiệu quả mô hình EfficientNet-B7 vốn được huấn luyện trên ảnh màu; và (v) tiến hành phân loại dựa trên đặc trưng ảnh và ROI để đánh giá tính hiệu quả của các mô hình.

Các kết quả nghiên cứu trong chương này đã được công bố trong 1 bài báo Hội nghị Quốc tế:

[CT2] 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.

### 4.1. TRÍCH ĐẶC TRUNG ẢNH UTV

#### 4.1.1 Vai trò của đặc trưng ảnh UTV

Đặc trưng ảnh UTV giữ vai trò quan trọng trong việc hỗ trợ chẩn đoán và phân loại khối u. Các đặc trưng này bao gồm hình dạng, biên dạng, cấu trúc kết cấu, mức xám và các thông số thống kê từ ảnh. Chúng cung cấp thông tin có hệ thống, giúp mô tả rõ ràng hơn các đặc điểm vốn khó nhận diện chỉ bằng quan sát trực tiếp. Một vai

## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

---

trò nổi bật của đặc trưng ảnh là làm cơ sở cho việc phân loại khối u lành tính và ác tính. Thông qua việc trích xuất và phân tích đặc trưng, các mô hình máy tính có thể phát hiện ra sự khác biệt nhỏ giữa các loại khối u, từ đó hỗ trợ BS đưa ra quyết định chính xác hơn. Điều này đặc biệt quan trọng trong bối cảnh phát hiện sớm UTV nhằm tăng hiệu quả điều trị.

Đặc trưng ảnh còn giúp giảm tính chủ quan trong quá trình chẩn đoán. Việc đọc và phân tích ảnh UTV thường phụ thuộc nhiều vào kinh nghiệm của BS, dẫn đến sự khác biệt trong đánh giá. Khi áp dụng hệ thống trích xuất và chuẩn hóa đặc trưng, việc phân tích trở nên khách quan hơn, tăng tính nhất quán và độ tin cậy trong chẩn đoán. Ngoài ra, các đặc trưng được trích xuất chính là nền tảng đầu vào cho các thuật toán học máy và học sâu. Chất lượng của các đặc trưng quyết định trực tiếp đến hiệu quả của mô hình phân loại, chẳng hạn như SVM, Random Forest, CNN hoặc LNN. Việc lựa chọn và trích xuất đặc trưng phù hợp sẽ giúp mô hình học tốt hơn, cải thiện độ chính xác và khả năng khái quát hóa.

Cuối cùng, đặc trưng ảnh đóng vai trò quan trọng trong nghiên cứu và so sánh các phương pháp chẩn đoán. Nhờ có bộ đặc trưng chuẩn hóa, các nhà nghiên cứu có thể dễ dàng kiểm nghiệm, đánh giá và đối chiếu hiệu quả của nhiều thuật toán khác nhau, qua đó thúc đẩy sự phát triển của các hệ thống hỗ trợ chẩn đoán UTV.

### 4.1.2 Trích đặc trưng GLCM

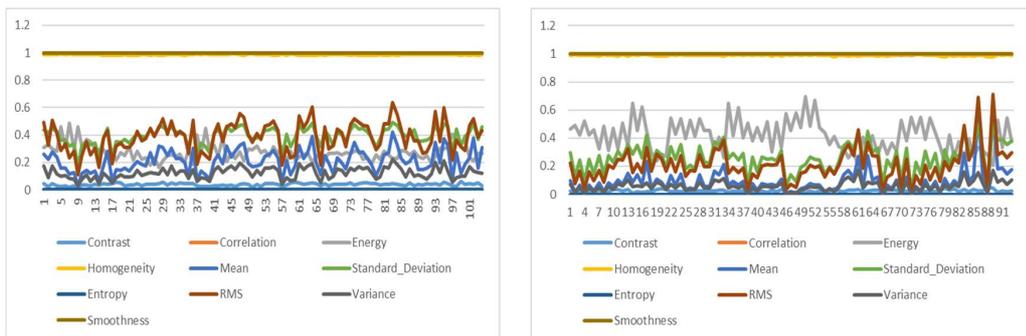
Phương pháp phân tích ma trận đồng mức xám được sử dụng nhiều trong trích đặc trưng ảnh [140]. Đặc trưng GLCM được ứng dụng rộng rãi trong lĩnh vực phân tích hình ảnh y sinh, trong đó có ảnh UTV. Từ GLCM, có thể tính toán được 10 đặc trưng phổ biến, bao gồm:

- *Energy (Năng lượng)* – đo mức độ đồng nhất của ảnh.
- *Entropy (Độ hỗn loạn)* – phản ánh sự ngẫu nhiên trong phân bố mức xám.
- *Contrast (Độ tương phản)* – biểu thị sự khác biệt về cường độ xám giữa các điểm ảnh.

## Chương 4 TRÍCH ĐẶC TRƯNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

- *Homogeneity (Tính đồng nhất)* – thể hiện sự phân bố đều của mức xám.
- *Correlation (Trương quan)* – đo mối quan hệ tuyến tính giữa các điểm ảnh.
- *Variance (Phương sai)* – phản ánh độ phân tán của mức xám.
- *Mean (Giá trị trung bình)* – cho biết mức xám trung bình của vùng quan sát.
- *Standard Deviation (Độ lệch chuẩn)* – thể hiện mức độ dao động quanh giá trị trung bình.
- *Maximum Probability (Xác suất cực đại)* – biểu diễn cặp giá trị mức xám xuất hiện nhiều nhất.
- *Inverse Difference Moment (IDM – Moment khác biệt nghịch đảo)* – đo sự tương đồng cục bộ trong ảnh.

Các đặc trưng này có vai trò quan trọng trong việc mô tả cấu trúc kết cấu mô vú. Ví dụ, khối u ác tính thường có kết cấu phức tạp, biên mờ và phân bố mức xám không đồng nhất, do đó các chỉ số như entropy hay contrast thường cao hơn so với khối u lành tính. Ngược lại, energy và homogeneity thường thấp hơn ở khối u ác tính. Thông qua việc phân tích 10 đặc trưng này, hệ thống có thể xây dựng bộ tiêu chí khách quan để phân loại khối u. Khi đưa vào các mô hình học máy hoặc học sâu, các đặc trưng GLCM giúp tăng cường khả năng phát hiện sớm UTV, đồng thời giảm thiểu sự phụ thuộc vào đánh giá chủ quan của BS.



(a) 10 đặc trưng của ảnh không có ung thư

(b) 10 đặc trưng của ảnh có ung thư

Hình 4.1 Thống kê các đặc trưng của ảnh có ung thư và không có ung thư

## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

---

Với 10 đặc trưng thu được bằng GLCM, kết quả khảo sát bằng đồ thị được trình bày ở hình 4.1. Kết quả thống kê cho thấy 4 đặc trưng có sự thay đổi lớn nhất giữa ảnh có và không có ung thư là Độ tương phản, Năng lượng, Tính đồng nhất và Giá trị trung bình, trong đó ý nghĩa và công thức của nó được mô tả dưới đây. 4 đặc trưng phù hợp với tập dữ liệu ảnh UTV sẽ được sử dụng trong bộ phân loại SVM [141].

**Độ tương phản** thể hiện ma trận có sự phân bố các phần tử, trong đó mỗi phần tử nằm xa đường chéo chính và giá trị độ tương phản của nó càng lớn. Giá trị tương phản cho chúng ta biết lượng đa dạng cường độ xám có trong hình ảnh UTV mẫu và nó được tính bằng công thức sau:

$$Con = \sum_i^{L-1} \sum_j^{L-1} (i-j)^2 P(i,j) \quad (4.1)$$

trong đó  $|i-j|$  là sự khác biệt về thang độ xám giữa các điểm ảnh liền kề,  $P(i, j)$  là phần tử  $(i, j)$  của GLCM đối xứng chuẩn hóa, được gọi là xác suất phân phối của các mức thang độ xám khác nhau giữa các điểm ảnh liền kề.  $L$  là số mức xám trong hình ảnh UTV.

**Năng lượng** có thể được xác định bằng cách đo thang mức xám của hình ảnh trong ảnh UTV phản ánh sự phân bố đồng đều trọng lượng và kết cấu của nó. Giá trị năng lượng  $Egy$  được biểu thị như sau:

$$Egy = \sum_i^{L-1} \sum_j^{L-1} P(i,j)^2 \quad (4.2)$$

**Tính đồng nhất** có thể mô tả sự đồng nhất về cấu trúc của các thang màu xám trong hình ảnh UTV và giá trị đồng nhất  $Hom$  được tính như sau:

$$Hom = \sum_i^{L-1} \sum_j^{L-1} \frac{1}{1+(i-j)^2} P(i,j) \quad (4.3)$$

## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

---

**Giá trị trung bình** cho biết mức xám trung bình của ảnh  $G(x, y)$  có kích thước  $m \times n$ . Ngoài ra, cường độ xám của các điểm ảnh được chuẩn hóa trong khoảng  $[0, 1]$  trước khi tính giá trị trung bình  $Me$  của ảnh  $G(x, y)$  như sau:

$$Me = \frac{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} G(x, y)}{(2^{16} - 1) \times m \times n} \quad (4.4)$$

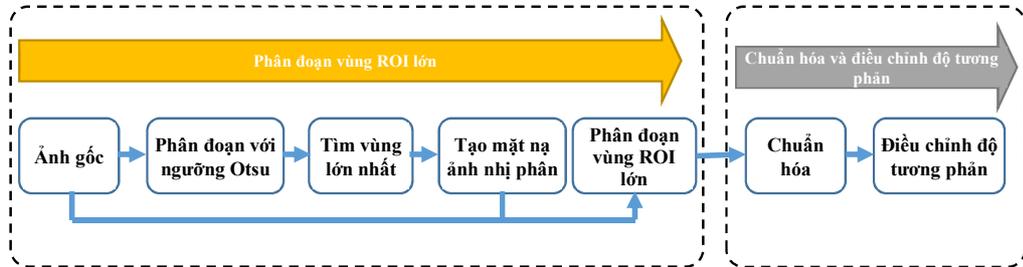
### 4.2. TRÍCH XUẤT VÙNG ROI CHO PHÂN LOẠI UTV

#### 4.2.1 Phân đoạn ảnh UTV để trích xuất vùng ROI

Trong ảnh UTV, thông thường có thể chứa các thành phần không mong muốn, không có đóng góp trong việc nhận dạng ảnh và có thể làm giảm độ chính xác phân loại cũng như tăng thời gian xử lý. Đặc biệt, các dòng văn bản hoặc ghi chú trên ảnh hoặc các vùng sáng ở rìa ảnh có thể ảnh hưởng đến quá trình huấn luyện và phân loại. Các vùng này cần được loại bỏ trước khi đưa vào xử lý tiếp theo. Phương pháp Otsu được sử dụng để loại bỏ các thành phần không cần thiết, chỉ giữ lại phần vật thể có chứa những tổn thương cho việc đánh giá và phân loại UTV.

Để sử dụng trong các hệ thống phân loại ảnh UTV, tất cả các hình ảnh cần được xử lý để xác định vùng ROI lớn nhất tập trung vào vùng tổn thương ví trước khi huấn luyện mô hình. Với việc xử lý hình ảnh để tạo ra vùng ROI lớn nhất, độ chính xác phân loại sẽ cao và cũng giảm thời gian tính toán của mô hình. Đặc biệt, thuật toán Otsu được áp dụng cho ảnh gốc để xác định ngưỡng phân đoạn ảnh nhằm thu được các đối tượng trong ảnh [77]. Hơn nữa, vùng ROI lớn nhất với phần tổn thương ở vù được xác định dựa trên mật độ điểm ảnh lớn nhất và các vùng ROI nhỏ hơn sẽ bị loại bỏ [142]. Hình ảnh có vùng ROI lớn nhất này là hình ảnh nhị phân được nhân với hình ảnh gốc để tạo ra hình ảnh gốc chỉ có vùng ROI lớn nhất. Cuối cùng, hình ảnh ROI lớn nhất được chuẩn hóa và điều chỉnh để tạo ra hình ảnh ROI mong muốn như trong hình 4.2.

Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV



Hình 4.2 Quy trình trích xuất vùng ROI lớn nhất

Phân đoạn ảnh là quá trình chia ảnh thành các vùng có cùng đặc điểm. Đặc điểm cơ bản của việc phân đoạn là độ chói đối với ảnh đơn sắc và thành phần màu đối với ảnh màu. Ngoài ra, cạnh và kết cấu là những đặc điểm quan trọng để phân đoạn. Trong nghiên cứu này, ảnh được phân đoạn dựa trên ngưỡng bằng phương pháp Otsu [77]. Giả sử ngưỡng  $k = 1, 2, \dots, (L-1)$  được chọn để chia số mức xám  $L$  thành hai vùng  $C_0$  và  $C_1$  và phương pháp ngưỡng Otsu cho phép tìm  $k$  dựa trên phương sai lớn nhất giữa hai vùng  $C_0$  và  $C_1$  bằng biểu thức sau:

$$\sigma_B^2 = \omega_0 \omega_1 (\mu_0 - \mu_1)^2 \quad (4.5)$$

Trong đó trọng số  $\omega_0$  và trị trung bình mức xám  $\mu_0$  của miền  $C_0$ , trọng số  $\omega_1$  và trung bình mức xám  $\mu_1$  của miền  $C_1$  được tính như sau:

$$\omega_0 = \sum_{q=0}^{k-1} p_q(r_q) \quad (4.6)$$

$$\omega_1 = \sum_{q=k}^{L-1} p_q(r_q) \quad (4.7)$$

$$\mu_0 = \sum_{q=0}^{k-1} qp_q(r_q) / \omega_0 \quad (4.8)$$

$$\mu_1 = \sum_{q=k}^{L-1} qp_q(r_q) / \omega_1 \quad (4.9)$$

## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

---

trong đó  $r_q$  là giá trị mức xám thứ  $q$ .

Do đó, mật độ xác suất  $p_q$  của mức xám được tính như sau:

$$p_q(r_q) = \frac{n_q}{M \times N} \quad (4.10)$$

trong đó  $n_q$  là số điểm thuộc mức xám thứ  $q$  và  $M \times N$  là tổng số điểm ảnh của một ảnh xám.

Phương pháp phân đoạn Otsu sẽ cho nhiều ngưỡng khác nhau, tuy nhiên để chọn ngưỡng chính xác người ta dựa vào giá trị phương sai của hai vùng ảnh như trong biểu thức (4.5). Cụ thể, phương sai lớn nhất là giá trị được coi là đạt được ngưỡng tốt nhất để phân đoạn. Tuy nhiên, ngưỡng gần với ngưỡng khuyến nghị có thể được chọn để phân đoạn tốt nhất và phù hợp với dữ liệu điển hình.

Để loại bỏ hiệu quả các thành phần không mong muốn và chỉ thu được vùng ROI lớn nhất, thuật toán tìm vùng lớn nhất dựa trên điểm ảnh sẽ được áp dụng [142]. Đặc biệt, hình ảnh được phân đoạn màu xám sẽ được chuyển đổi thành hình ảnh nhị phân để xác định vùng ROI lớn nhất dựa trên các điểm ảnh màu trắng, đồng thời loại bỏ các vùng nhỏ hơn. Hơn nữa, để có được hình ảnh gốc chỉ có vùng ROI lớn nhất, hình ảnh nhị phân có vùng ROI lớn nhất sẽ được nhân với hình ảnh gốc.

Để xác định ROI lớn nhất, ngưỡng Otsu được áp dụng để tạo ra các đối tượng có cùng mức xám điểm ảnh và sau đó sẽ tìm thấy đối tượng lớn nhất, được gọi là vùng ROI lớn nhất. Cụ thể, thuật toán tìm vùng ROI lớn nhất sẽ được xác định như sau:

$$\chi_i = \sum_{x=0}^{h_i} \sum_{y=0}^{w_i} O_i(x, y) \quad (4.11)$$

trong đó  $\chi_i$  là tổng số điểm ảnh màu trắng,  $O_i$  là vùng ROI thứ  $i$  có kích thước  $h_i \times w_i$  ( $i = 1, 2, \dots, n$ ).

Sau khi hoàn thành phân đoạn vùng ROI lớn nhất, hình ảnh tiếp tục được chuẩn hóa và tăng cường sử dụng kỹ thuật cân bằng lược đồ. Việc tăng cường sẽ làm cho tất cả các hình ảnh đồng bộ về đặc trưng, độ sáng và việc phân loại sẽ chính xác hơn.

## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

---

Các hình ảnh sau khi tăng cường sẽ được đưa vào các mô hình mạng học sâu để huấn luyện và phân loại.

Để chuẩn hóa hình ảnh, phương pháp với các giá trị mức xám tối thiểu và tối đa được áp dụng. Giá trị mức xám mới của điểm ảnh được chuẩn hóa bằng phương trình sau:

$$x_{new} = x_{old} \frac{x_{old} - x_{min}}{x_{max} - x_{min}} \quad (4.12)$$

Với  $x_{new}$  là mức xám mới,  $x_{old}$  là mức xám cũ của điểm ảnh,  $x_{max}$  là mức xám tối đa,  $x_{min}$  là mức xám tối thiểu.

Để tăng cường hình ảnh, phương pháp cân bằng lược đồ được sử dụng. Đặc biệt, phương pháp này sẽ tính toán lại các giá trị mức xám dựa trên mật độ xác suất của từng mức xám, trong đó thể hiện mối liên hệ giữa mật độ xác suất của mức xám tiếp theo và mật độ xác suất của mức xám trước đó thông qua phép tính tổng. Do đó, mỗi mức xám được tính toán để cân bằng dựa trên lược đồ trong ảnh được mô tả theo phương trình sau:

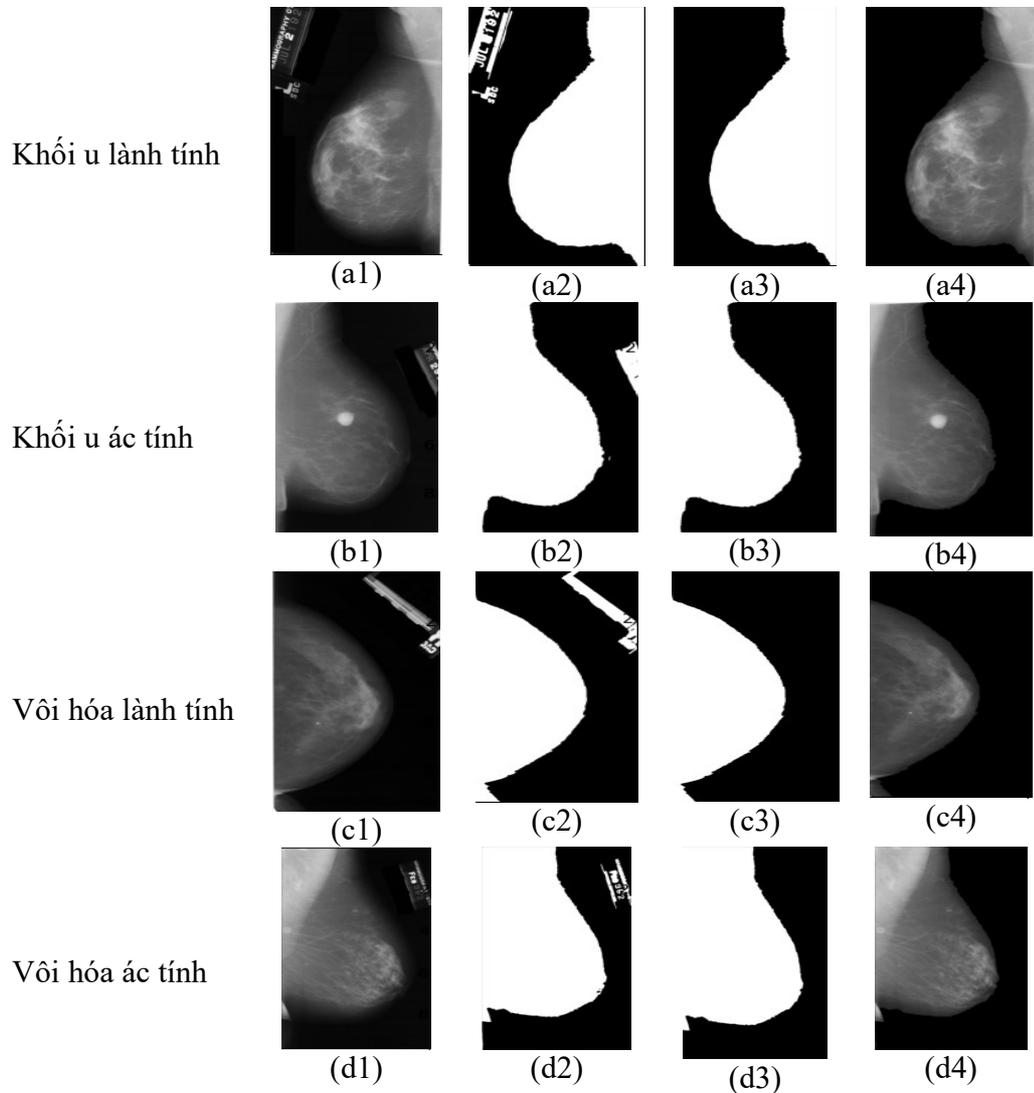
$$s_q = (L - 1) \sum_{q=0}^{L-1} p_q(r_q) \quad (4.13)$$

trong đó  $L$  là số mức xám có  $q=0,1,2,\dots,L-1$ .

Giả sử rằng  $r_q$  là giá trị mức xám trong ảnh gốc và  $s_q$  là mức xám trong ảnh được cân bằng và  $p_q$  là mật độ xác suất của các mức xám của ảnh gốc. Hơn nữa,  $n_q$  là số lượng điểm ảnh ở mức xám  $q=0,1,2,\dots,L-1$  và  $M \times N$  là số lượng điểm ảnh theo chiều ngang và chiều dọc trong một ảnh.

Với hình ảnh tổn thương vú, việc xác định chính xác đó có phải là UTV hay không là rất cần thiết vì điều này có thể giúp BS có thêm thông tin để đưa ra quyết định về bệnh UTV. Trong nghiên cứu, các tập dữ liệu hình ảnh về các loại tổn thương ở vú được phân đoạn để trích xuất các vùng ROI lớn nhất với nhiều điểm đặc trưng nhằm phân loại tổn thương tốt hơn. 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.

Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

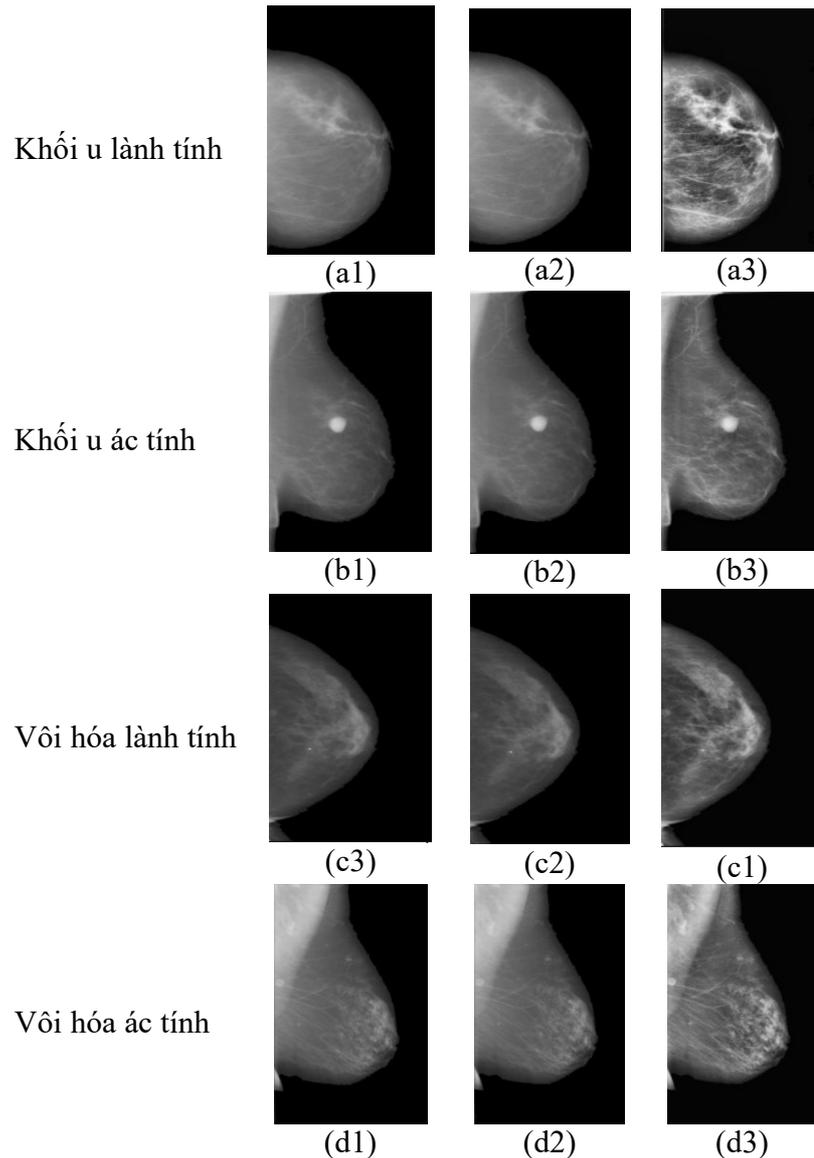


**Hình 4.3. Mô tả trích xuất ROI dựa trên đường viền lớn nhất và loại bỏ các thành phần không mong muốn: (a1-a4) Hình ảnh gốc của bốn tổn thương vú; (b1-b4) Hình ảnh được phân đoạn bằng phương pháp ngưỡng Otsu (c1-c4) Hình ảnh nhị phân có ROI lớn nhất đã loại các thành phần không mong muốn; (d1-d4) Ảnh gốc có ROI lớn nhất không có thành phần không mong muốn**

Phân đoạn với các giá trị ngưỡng khác nhau có thể sinh ra các ảnh được phân đoạn có ROI khác nhau. Vì vậy, kết quả thực hiện phân đoạn cho thấy việc điều chỉnh độ sáng hình ảnh trước khi phân đoạn đã cải thiện đáng kể các vùng được phân đoạn bằng thuật toán Otsu. Hình 4.4 cho thấy hình ảnh ROI được chuẩn hóa và điều chỉnh

#### Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

độ tương phản bằng thuật toán cân bằng lược đồ. Cụ thể, hình ảnh sau khi tăng cường sẽ rõ ràng hơn với các đặc trưng tập trung vào vùng ROI lớn nhất và điều này có thể giúp hệ thống phân loại khối u đạt hiệu suất tốt hơn.



**Hình 4.4. Mô tả kết quả chuẩn hóa và tăng cường hình ảnh ROI lớn nhất bằng cách sử dụng cân bằng lược đồ: (a1-a3) Ảnh gốc; (b1-b3) Hình ảnh được chuẩn hóa; (c) Ảnh có cân bằng lược đồ**

Việc trích xuất vùng ROI trong nghiên cứu này chỉ thực hiện với vùng ROI lớn, bao gồm gần như toàn bộ vùng hình ảnh của vú, chỉ loại bỏ các thành phần nằm trên

## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

---

nền ảnh có thể ảnh hưởng đến hiệu suất huấn luyện và phân loại. Do đó quá trình này có thể hạn chế tối đa việc bị mất thông tin do khu vực tổn thương bị cắt bỏ. Để đảm bảo kiểm chứng cho việc không mất thông tin, trong quá trình điều chỉnh thuật toán trích xuất vùng ROI, các ảnh kết quả sau khi trích xuất ROI được rà soát lại ngẫu nhiên nhằm phát hiện các khả năng các vùng thông tin quan trọng được cắt bỏ và điều chỉnh vùng ROI cho phù hợp. Kết quả của việc phân loại thử nghiệm với các trường hợp có và không có tách vùng ROI cũng cho thấy hiệu quả của quá trình này và việc tách ROI không làm mất các thông tin quan trọng.

Về lý thuyết, vẫn có khả năng xảy ra trường hợp vùng tổn thương nhỏ và nằm sát biên của vùng ảnh vù, dẫn đến có thể bị ảnh hưởng bởi việc trích xuất ROI. Tuy nhiên trong các tập ảnh đã sử dụng trong nghiên cứu, chưa thấy xuất hiện trường hợp này.

### 4.2.2 Tô màu cho vùng ROI

Ảnh UTV là ảnh xám có 1 kênh màu, trong khi đó, hầu hết các mô hình học sâu như EfficientNet đều được huấn luyện trước trên tập ảnh ImageNet là tập ảnh màu và hoạt động tốt trên các ảnh màu có 3 kênh màu. Do đó, trong nghiên cứu này, để sử dụng hiệu quả khả năng của mô hình học sâu và việc phân loại chính xác hơn, các vùng ROI đều được tô màu, tạo thành vùng CROI. Để tạo CROI, chúng tôi phát triển một thuật toán trực quan hóa các đặc trưng cần được nhận diện, chẳng hạn như vùng khối u trên ảnh nhũ ảnh, bằng cách áp dụng một dải màu được chọn để làm nổi bật các vùng tổn thương [143].

Để tạo ảnh CROI, thuật toán tô màu được xây dựng với hai nhiệm vụ: Nhận diện và So sánh (IaC). Cụ thể, nhiệm vụ thứ nhất (Nhận diện) là xác định các vùng có mức xám bất thường trong vùng ROI trên ảnh xám (GROI), được đánh dấu như dấu hiệu của khối u hoặc các loại tổn thương khác. Nhiệm vụ thứ hai (So sánh) là tìm sự khác biệt giữa vùng của một nhóm điểm nhỏ so với nền xung quanh, qua đó cũng thể hiện các dấu hiệu tổn thương hoặc vô hóa. Như vậy, với nhiệm vụ thứ nhất, một chỉ số mức xám được tính toán và sau đó chuyển đổi thành các mức màu để bổ sung vào

## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

ảnh ROI. Trong khi đó, các vùng nghi ngờ như vôi hóa có thể được phát hiện nhờ việc áp dụng các thuật toán biến đổi hình thái học là TopHat và BlackHat. Cuối cùng, thuật toán IaC có thể được tính toán để tạo ra ảnh CROI (ký hiệu  $I_{CROI}$ ), trong đó các vùng nghi ngờ được làm nổi bật và tô màu theo công thức dưới đây:

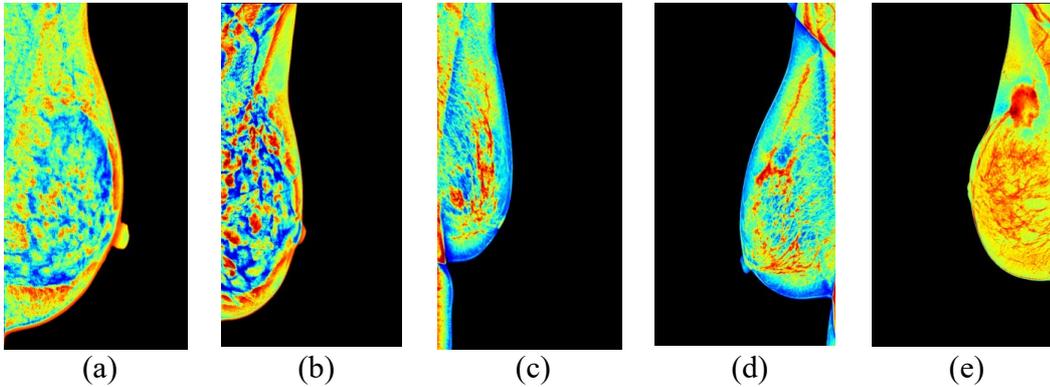
$$I_{CROI} = (G(x, y) + I_{TH}) - I_{BH} - Z[[SE_Z^{\wedge} \cap G(x, y)] \subseteq G(x, y) \quad (4.14)$$

trong đó,  $G(x, y)$  là ảnh đầu vào,  $SE$  là phân tử cấu trúc của ảnh,  $I_{TH}$  và  $I_{BH}$  được xác định như sau:

$$I_{TH} = G(x, y) - (G(x, y) \ominus SE) \oplus SE \quad (4.15)$$

$$I_{BH} = (G(x, y) \oplus SE) \ominus SE - G(x, y) \quad (4.16)$$

Hình 4.5 cho thấy kết quả tô màu đối với 5 ảnh UTV với các mức BIRADS khác nhau làm nổi bật các dấu hiệu nhận dạng tổn thương trên ảnh UTV. Các ảnh BIRADS 4,5 cho thấy các vùng màu rõ ràng, khác biệt với các vùng mô bình thường, làm nổi bật các điểm khác biệt của các trường hợp có nguy cơ UTV cao.



**Hình 4.5. 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**

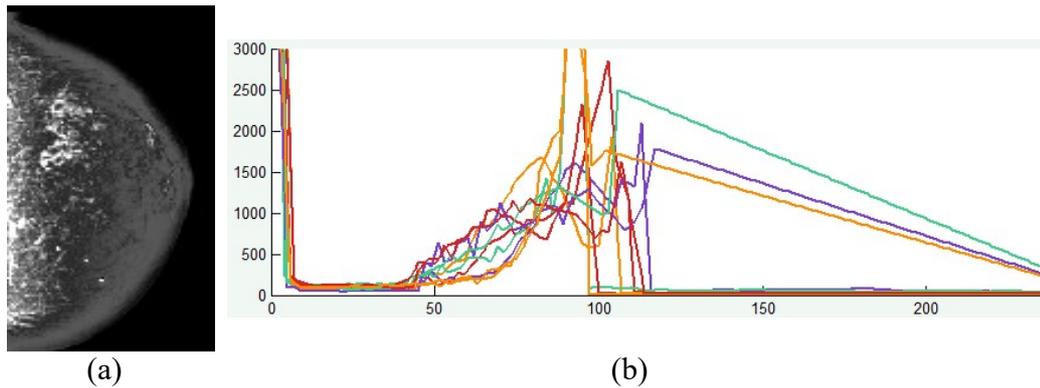
### 4.3. ỨNG DỤNG PHÂN LOẠI UTV DỰA TRÊN ĐẶC TRƯNG ẢNH

#### 4.3.1 Phân loại ảnh UTV dựa vào đặc trưng độ tương phản

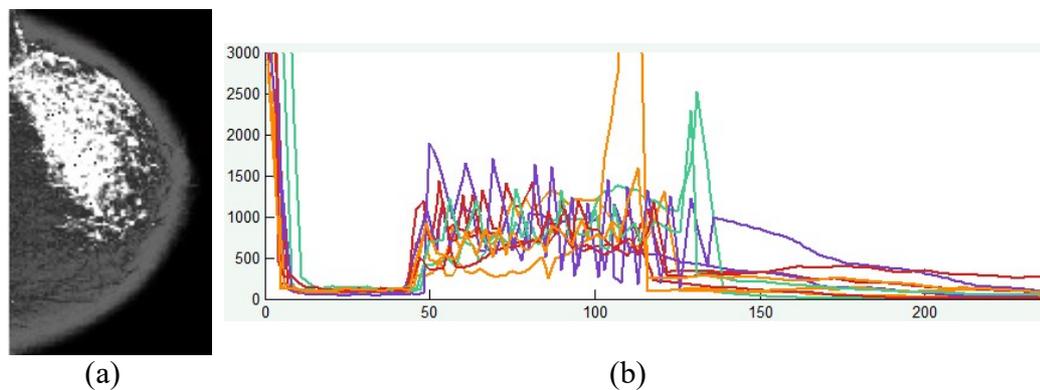
Tất cả ảnh UTV được tính toán để xây dựng các lược đồ mức xám. Kết quả cho thấy mật độ mức xám giữa hai loại ảnh UTV (ung thư và bình thường) có sự khác

## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

biệt, như thể hiện trong Hình 4.6 và Hình 4.7 [77]. Cụ thể, mật độ điểm ảnh của ảnh ung thư có sự thay đổi đột ngột trong khoảng mức xám từ 0,18 đến 0,20, trong khi đó ở ảnh bình thường, mật độ thay đổi chậm hơn trong cùng khoảng này.



**Hình 4.6. Lược đồ xám của ảnh bình thường: (a) Ảnh vú bình thường; (b) Lược đồ xám của ảnh**



**Hình 4.7. Lược đồ xám của ảnh UTV: (a) Ảnh vUTV; (b) Lược đồ xám của ảnh**

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.

Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

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

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

Trong thực tế, các ảnh UTV được phân đoạn vùng ROI bằng phương pháp phân cụm mờ Fuzzy C-Means và sau đó các vùng ROI trên ảnh được tính toán để xác định vùng bình thường và bất thường [21]. Kết quả cho thấy độ chính xác phân loại trong các nghiên cứu này đạt khoảng 92%. Trong luận án này, các kỹ thuật XLA được áp dụng, lược đồ của ảnh UTV được biểu diễn và sau đó thống kê mật độ điểm ảnh trên cả ảnh bình thường và bất thường được thực hiện làm cơ sở cho việc đánh giá. Các giá trị tương phản và số lượng điểm ảnh tương ứng cũng được xác định để phục vụ việc đánh giá tình trạng bệnh. Với thí nghiệm gồm 20 ảnh UTV, trong đó có 10 ảnh cho mỗi loại, có thể đánh giá được khả năng nhận diện tình trạng bệnh dựa trên ảnh sau xử lý. Đặc biệt, với số liệu thống kê ở bảng 4.1 có thể đánh giá được tình trạng tổn thương của ảnh UTV. Đây có thể là một nguồn thông tin hữu ích cho BS trong chẩn đoán bệnh.

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

Để phát hiện UTV, thuật toán SVM được áp dụng, trong đó 4 loại đặc trưng tối ưu bao gồm Độ tương phản, Tương quan, Năng lượng, Entropy được sử dụng để làm đầu vào cho quá trình huấn luyện và kiểm tra. Hệ thống phân loại này cho phép có thể phát hiện UTV. Từ các đặc trưng được trích xuất bằng GLCM, siêu phẳng tuyến tính trong thuật toán SVM được xác định để phân chia các tập dữ liệu đặc trưng để huấn luyện và phân loại nhằm phát hiện UTV. Bộ phân loại SVM được dùng sử dụng nhân tuyến tính, với hệ số nhân Lagrange.

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. Định dạng ảnh ở dạng Portable Network Graphics (PNG), bao gồm 9.684 ảnh thuộc 3 loại: lành tính (Benign), ung thư (Cancer) và bình thường (Normal). Cụ thể, nghiên cứu này sử dụng một phần của bộ dữ liệu Mini-DDSM để xử lý, trích xuất đặc trưng và phân loại UTV, như được mô tả trong Bảng 4.2.

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

Tập dữ liệu	Ung thư	Bình thường
Huấn luyện	100	100
Kiểm tra	50	50

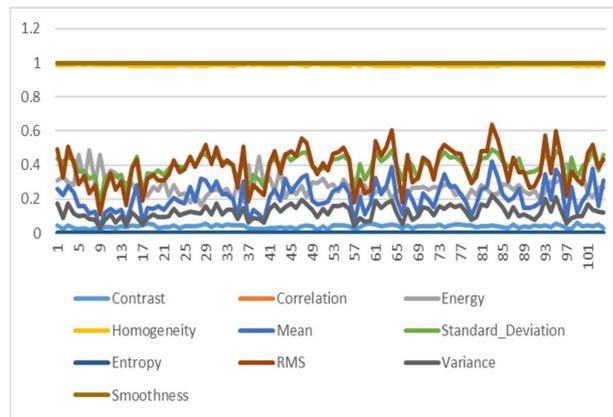
Để phân loại UTV bằng thuật toán SVM [145], hệ số Lagrange là cần thiết. Trong thuật toán này, các hệ số Lagrange, với  $i = 1, 2, \dots, m$  tương ứng với các ràng buộc bất đẳng thức và Lagrangian  $L(v, w, \alpha)$  của chúng liên quan đến vectơ pháp tuyến  $w$  và vectơ đặc trưng  $v$  và được biểu diễn như sau:

$$f(v, w, b) = w^T \cdot v + b = \sum_i^p w_i v_i + b \quad (4.17)$$

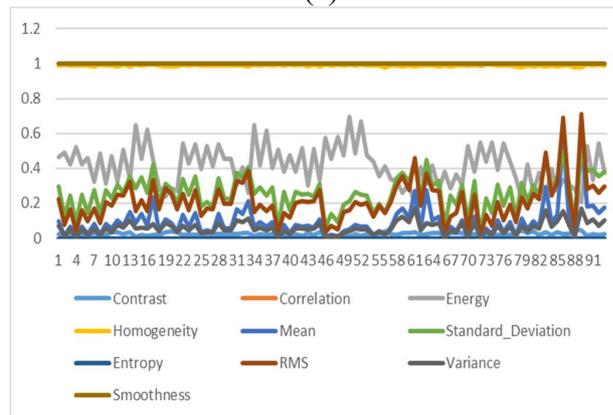
Trong thuật toán này, mỗi mẫu huấn luyện tương ứng với  $v_i$  và các Lagrange. Vì vậy, sau khi huấn luyện sẽ trở thành vector hỗ trợ nằm trên một trong hai siêu mặt phẳng.

## Chương 4 TRÍCH ĐẶC TRƯNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

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. Do đó, các đặc trưng GLCM được sử dụng để huấn luyện và phân loại bằng phương pháp SVM. Quy trình trích xuất đặc trưng dựa trên số liệu thống kê để tìm nhóm đặc trưng tốt nhất cho phân loại SVM.



(a)



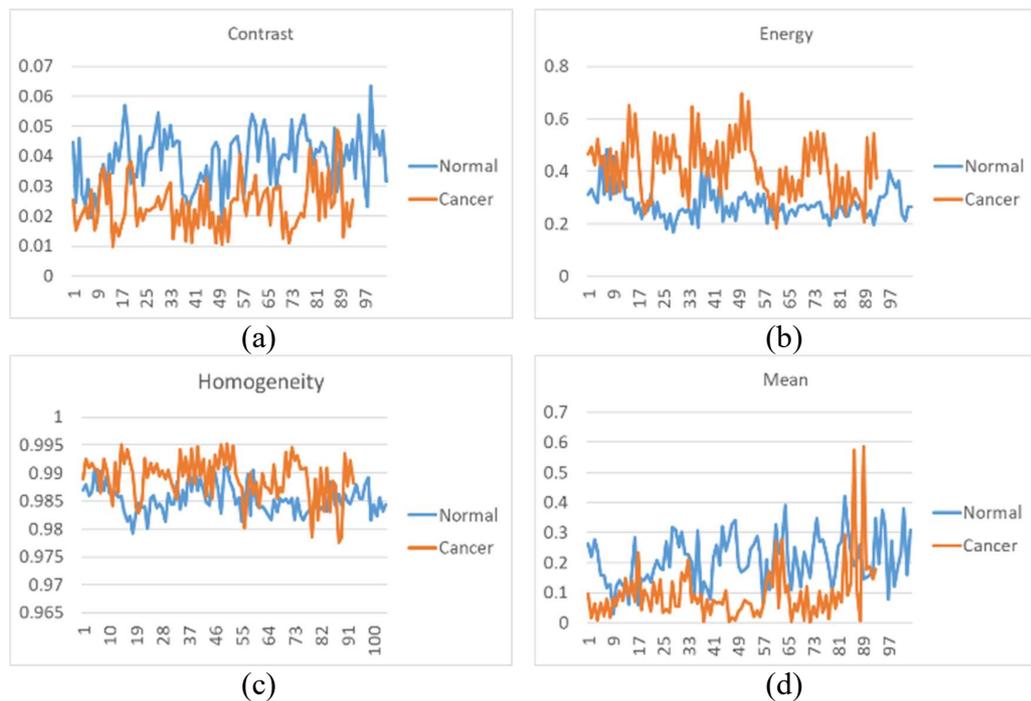
(b)

**Hình 4.8. 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**

Từ hình 4.8 có thể thấy rằng từ thống kê kết quả trích xuất đặc trưng GLCM cho 10 loại trong 300 hình ảnh UTV, 4 đặc trưng tối ưu được tìm thấy trong 10 loại đặc trưng, bao gồm Năng lượng, Tương phản, Đồng nhất và Trung bình như mô tả trong hình 4.9. Ngoài ra, các đặc trưng này cho thấy mô tả tốt nhất về sự khác biệt giữa vùng tổn thương và vùng không tổn thương trong hình ảnh UTV. Do đó, 4 đặc trưng tối ưu là đầu vào của bộ phân loại SVM để phân loại các trường hợp bình thường và

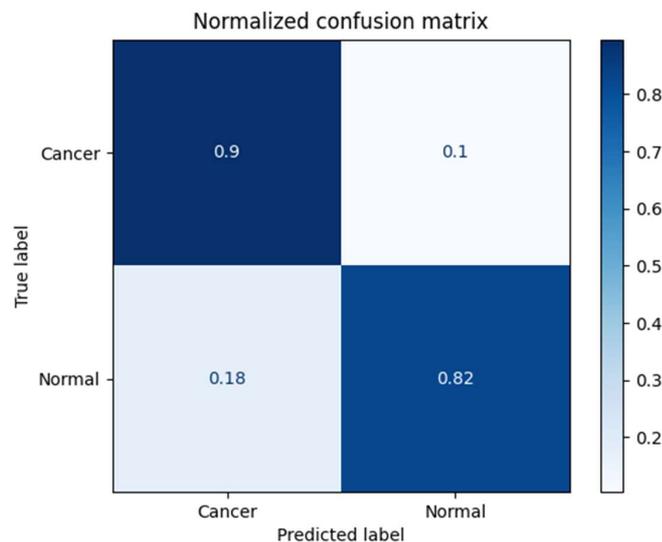
## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

ung thư. Ở bước tiếp theo, các đặc trưng kể trên sẽ sử dụng để phân loại với bộ phân loại SVM để đánh giá khả năng phân loại khối u dựa vào các đặc trưng vừa tìm được. Kết quả sẽ cho chúng ta biết tính khả thi của việc phân loại ảnh UTV dựa vào các bộ phân loại truyền thống như SVM.



**Hình 4.9. Các đặc trưng của ảnh bình thường và ung thư: (a) Contrast; (b) Energy; (c) Homogeneity; (c) Mean**

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.



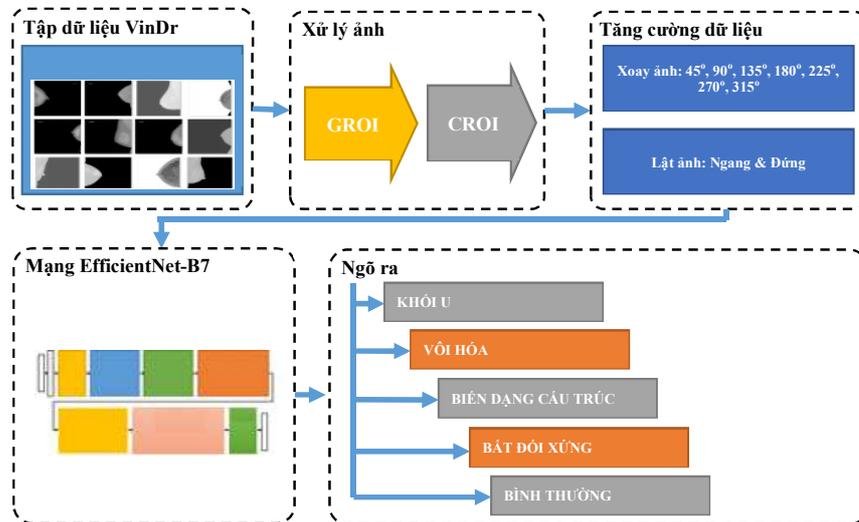
**Hình 4.10. Ma trận nhầm lẫn để đánh giá 50 hình ảnh thử nghiệm mỗi loại**

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].

### 4.3.3 Phân loại ảnh UTV dựa trên vùng ROI

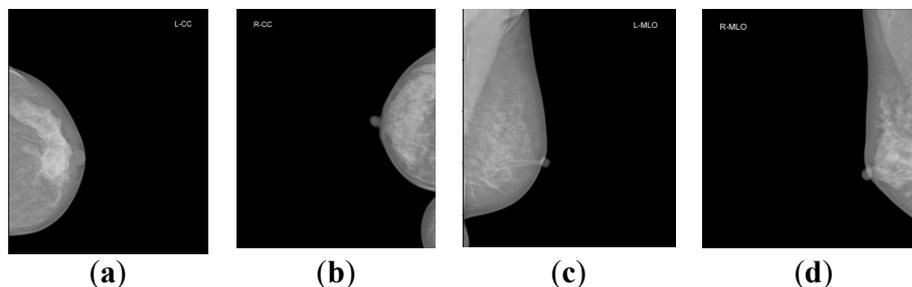
Tập ảnh UTV được chia thành năm loại tổn thương: khối u (mass), vôi hóa (calcification), biến dạng cấu trúc (architectural distortion), bất đối xứng (asymmetry) và bình thường (normal). Để thực hiện phân loại các tập ảnh này bằng mô hình EfficientNet-B7, dữ liệu đầu vào sẽ được xử lý qua hai bước: trích 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.

## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV



**Hình 4.11. 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, trong đó các ảnh thu thập được là ảnh X-quang tuyến vú (mammogram) được chụp trong giai đoạn từ năm 2018 đến 2020. Cụ thể, các tập ảnh này được lưu trữ trong hệ thống Lưu trữ và Truyền thông Hình ảnh (PACS) của BV Đại học Y Hà Nội (HMH) và BV Quân đội 108 (H108). Ngoài ra, bộ dữ liệu này bao gồm 5.000 lần chụp tuyến vú, tương ứng với 20.000 ảnh X-quang tuyến vú, vì mỗi lần chụp tạo ra 4 ảnh vú: CC (Cranial-Caudal) bên phải, CC bên trái, MLO (Mediolateral-Oblique) bên phải, và MLO bên trái, như minh họa trong Hình 4.12.

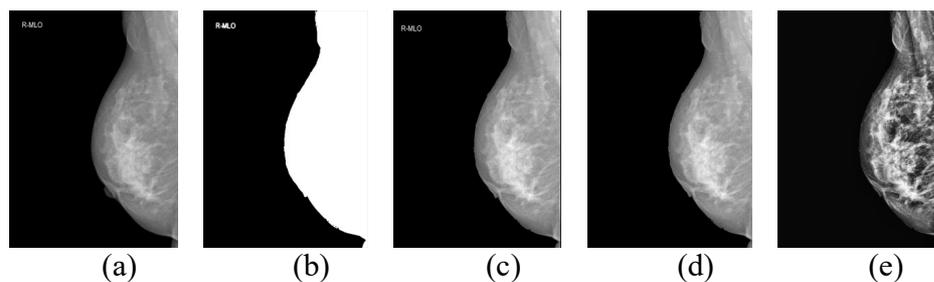


**Hình 4.12. Bộ 4 ảnh chụp tuyến vú mẫu cho một lần chụp trong bộ dữ liệu VinDr-Mammo: (a) CC bên trái; (b) CC bên phải; (c) MLO bên trái; (d) MLO bên phải.**

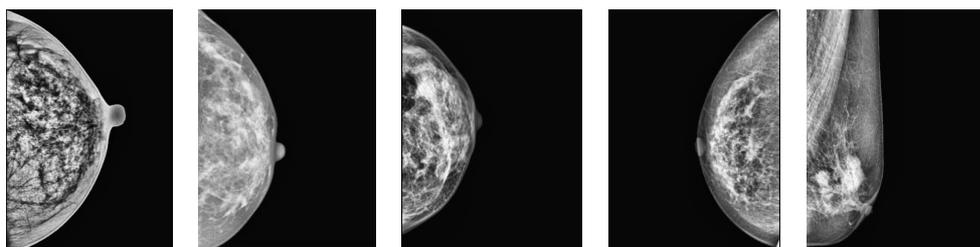
## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

Điểm quan trọng là thông tin cá nhân của bệnh nhân trên ảnh luôn được bảo mật, và các ảnh này chỉ giữ lại các đối tượng liên quan đến tổn thương. Do đó, toàn bộ 20.000 ảnh mammogram của bộ dữ liệu VinDr-Mammo được phân loại thành năm nhóm ảnh UTV và 11 loại bất thường tuyến vú (bao gồm: khối u, vôi hóa nghi ngờ, bất đối xứng, bất đối xứng khu trú, bất đối xứng toàn cục, biến dạng cấu trúc, dày da, co rút da, co rút núm vú, và hạch bạch huyết nghi ngờ). Các bệnh nhân trong bộ dữ liệu có độ tuổi từ 20 đến 86, trong đó tập trung nhiều nhất ở nhóm 40–45 tuổi.

Một ảnh tổn thương tuyến vú gốc thường chứa nhiều thông tin về bệnh nhân và các thành phần khác, điều này có thể ảnh hưởng đến quá trình phân loại. Do đó, các bước phân đoạn ảnh được thực hiện nhằm loại bỏ những thành phần này và chỉ giữ lại GROI cùng nền đen để phục vụ cho việc phân loại như hình 4.13. Kết quả phân đoạn bằng phương pháp ngưỡng Otsu cho thấy thông tin về nền ảnh và các nhiễu đã được loại bỏ, như minh họa trong Hình 4.14. Cụ thể, năm ảnh chỉ còn lại phần nền đen và các vùng GROI. Hơn nữa, các ảnh chỉ chứa nền đen và GROI sẽ giúp giảm thời gian huấn luyện khi đào tạo mạng học sâu, đồng thời cải thiện khả năng phân loại chính xác.



**Hình 4.13. Minh họa quá trình XLA: (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 GROI sau hậu xử lý.**

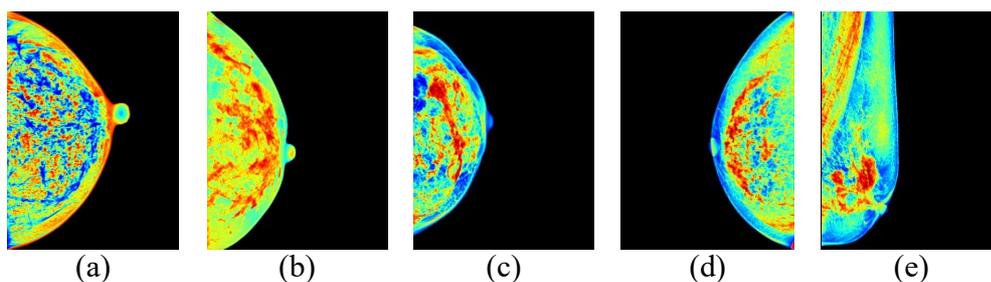


#### Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

(a) (b) (c) (d) (e)

**Hình 4.14. Minh họa năm loại ảnh với GROU và nền đen: (a) bình thường; (b) bất đối xứng cục bộ; (c) biến dạng cấu trúc; (d) vôi hóa; (e) khối u.**

Sau khi tạo ra các ảnh với GROU, 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, trong đó các đặc trưng bên trong CROI được làm nổi bật. 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.



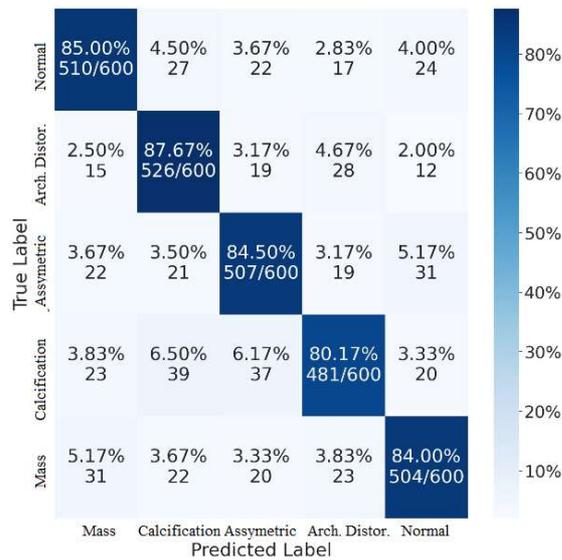
**Hình 4.15. Minh họa năm loại tổn thương khác nhau của ảnh với CROI: (a) Không phát hiện; (b) Bất đối xứng cục bộ; (c) Biến dạng cấu trúc; (d) Vôi hóa; (e) Khối u**

Trong Hình 4.15, các ảnh với CROI thể hiện dấu hiệu tổn thương (màu đỏ) như vôi hóa hoặc các loại tổn thương khác với đặc điểm màu sắc khác biệt so với các trường hợp không có tổn thương. Bên cạnh đó, ảnh được tô màu cho thấy sự khác biệt rõ rệt: trong một số trường hợp, vùng màu đỏ có thể là khối u, còn vôi hóa hiển thị màu xanh kèm theo các đốm đỏ. Tóm lại, một ảnh có chứa các vùng màu này được xem là có tổn thương, trong khi ảnh không có tổn thương sẽ không xuất hiện đặc trưng màu sắc rõ ràng. Rõ ràng, hiệu quả phân loại của các ảnh CROI như vậy sẽ cao khi sử dụng các mạng học sâu.

Hình 4.16 và Hình 4.17 thể hiện kết quả phân loại khi sử dụng ma trận nhầm lẫn cho hai trường hợp: ảnh vùng ROI xám và ảnh vùng ROI đã tô màu với mô hình EfficientNet-B7 [143]. Cụ thể, các giá trị hàng biểu thị nhãn thực tế của các tập

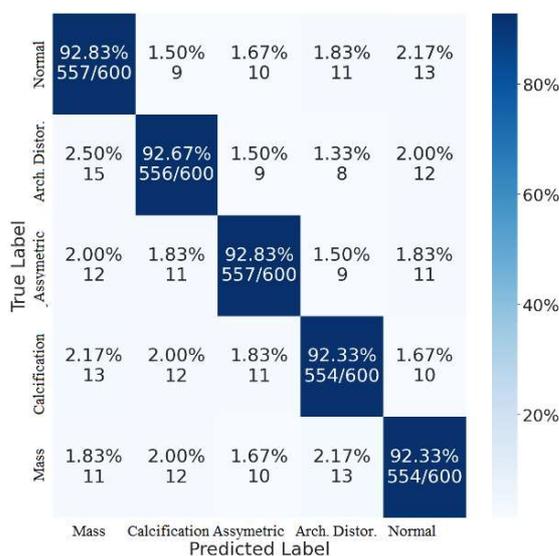
## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

ảnh (khối lượng; vôi hóa; biến dạng kiến trúc; bất đối xứng; và bình thường), các giá trị cột biểu thị nhãn dự đoán của năm loại tập ảnh này, và các đường chéo biểu thị giá trị True Positive (TP). Kết quả là việc phân loại ảnh bằng ảnh ROI đã tô màu, như thể hiện trong Hình 4.17, cho độ chính xác vượt trội so với ảnh ROI xám. Cụ thể, độ chính xác phân loại trung bình là 92,6% so với 84,3%. Điều này cho thấy hiệu quả của việc tô màu ảnh.



**Hình 4.16. Ma trận nhầm lẫn để đánh giá hiệu suất phân loại đối với các tập ảnh xám.**

## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV



**Hình 4.17. Kết quả Ma trận nhầm lẫn để đánh giá hiệu suất phân loại đối với các tập ảnh có tô màu**

Trong Bảng 4.3, kết quả huấn luyện phân loại với mô hình trên cho thấy các phương pháp tiền XLA 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

### ❖ Kết luận

Chương này đã trình bày một quy trình toàn diện cho việc trích xuất đặc trưng hình ảnh và xử lý vùng ROI trong phân loại UTV. Các kết quả nghiên cứu cho thấy rằng đặc trưng thống kê từ ma trận đồng xuất hiện mức xám (GLCM) có khả năng mô tả rõ ràng sự khác biệt về cấu trúc giữa khối u lành tính và ác tính, qua đó nâng

## Chương 4 TRÍCH ĐẶC TRUNG HÌNH ẢNH VÀ XỬ LÝ VÙNG ROI CHO PHÂN LOẠI UTV

---

cao hiệu quả của các mô hình phân loại truyền thống [77] [145]. Đồng thời, việc áp dụng kỹ thuật phân đoạn dựa trên ngưỡng Otsu kết hợp với xác định vùng ROI lớn nhất đã chứng minh tính hữu ích trong việc loại bỏ các thành phần không mong muốn, giữ lại vùng tổn thương quan trọng và rút ngắn thời gian huấn luyện mô hình [142]. Đặc biệt, phương pháp chuyển đổi từ ROI xám sang ROI màu (CROI) đã làm nổi bật các dấu hiệu bất thường trong ảnh nhũ ảnh, đồng thời tận dụng hiệu quả sức mạnh của các mô hình học sâu vốn được huấn luyện trên tập dữ liệu ảnh màu [143]. Thực nghiệm với EfficientNet-B7 khẳng định rằng cách tiếp cận này cải thiện đáng kể độ chính xác phân loại, vượt trội hơn so với việc chỉ sử dụng ROI xám. Những kết quả đạt được đã chứng minh vai trò then chốt của tiền xử lý và lựa chọn đặc trưng trong hệ thống hỗ trợ chẩn đoán, đồng thời khẳng định CROI là một hướng tiếp cận khả thi và hiệu quả nhằm nâng cao độ chính xác của các mô hình phân loại UTV. Các bước tiền xử lý trong chương này được sử dụng trong quy trình phân loại trước khi đưa vào mạng học sâu ở chương 5.

## 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

Trong chương này, các mạng học sâu được sử dụng để phân loại các ảnh X-quang vú được thu thập từ các tập dữ liệu được chọn lọc trong số các nguồn dữ liệu đã được đề cập tới trong chương 3, sử dụng các kỹ thuật tiền xử lý ảnh đã được đề cập đến cũng trong chương này để xử lý hình ảnh trước khi đưa vào huấn luyện trên các mô hình mạng học sâu. Nhiều mạng học sâu khác nhau đã được sử dụng, kết hợp với một số tập ảnh khác nhau và các cách phân loại ngõ ra khác nhau để khám phá các cách thức mà mạng học sâu có thể được sử dụng để hỗ trợ chẩn đoán bệnh UTV cũng như đánh giá và so sánh hiệu quả giữa các phương pháp này. Từ đó có thể chọn ra các kết quả tốt nhất để sử dụng trong việc triển khai thử nghiệm một hệ thống hỗ trợ chẩn đoán từ xa cho bệnh UTV.

Với mục đích như vậy, nghiên cứu đã sử dụng các mạng học sâu mới nhất với nhiều bài toán phân loại khác nhau (số lượng ngõ ra khác nhau, tập dữ liệu khác nhau) bao gồm phân loại ảnh có ung thư hay không có ung thư (3 ngõ ra gồm bình thường, khối u lành tính, khối u ác tính); phân loại để xác định loại tổn thương (4 ngõ ra bao gồm: tổn thương dạng vi vôi hóa lành tính, vi vôi hóa ác tính, khối u lành tính, khối u ác tính); và phân loại theo nguy cơ ung thư đánh giá bằng thang đo BIRADS. Các tập ảnh khác nhau được sử dụng, được lựa chọn phù hợp với ngõ ra mong muốn và các đặc điểm phù hợp. Kết quả nghiên cứu trong chương này đã cho thấy sự phù hợp của hệ thống được đề xuất cho bài toán phân loại các loại tổn thương vú dựa vào tập ảnh X-quang vú thu thập được.

Các kết quả nghiên cứu trong chương này đã được công bố trong 3 bài báo Tạp chí Quốc tế và một bài báo đã nộp, đang chờ phản biện:

## 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

---

- [CT3] 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/aeee.v22i4.240303.
- [CT4] 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, DOI: 10.48084/etasr.12067.
- [CT5] 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.
- [CT7] Thanh-Tam Nguyen, Thanh-Hai Nguyen, Thanh-Nghia Nguyen, Ba-Viet Ngo, “Dataset Fusion for Breast Lesion Classification with a Combined EfficientNet-B7 and Attention Network”, *IEEE Access*, submitted.

### 5.1. MẠNG HỌC SÂU TRONG PHÂN LOẠI BỆNH UTV

Các mạng nơ ron tích chập (CNN) được sử dụng phổ biến trong các bài toán nhận dạng và phân loại ảnh vì chúng có nhiều ưu điểm so với các kỹ thuật khác. Cấu trúc mạng CNN thường bao gồm các lớp tích chập, lớp tuyến tính chỉnh lưu (ReLU), lớp lớp tổng hợp (Pooling), và lớp kết nối đầy đủ. Các mạng CNN thường có từ 5 đến 25 lớp trong mô hình đầy đủ. Các mạng CNN phổ biến thường được sử dụng với độ chính xác cao cho nhận dạng hình ảnh gồm có VGG16, ResNet, DenseNet, EfficientNet. Các mạng này đều được huấn luyện và kiểm tra trên tập dữ liệu ImageNet và cho độ chính xác nhận dạng rất cao. Trong phần này, vai trò của mạng học sâu cho ứng dụng phân loại ảnh UTV sẽ được trình bày. Hơn nữa, cấu trúc và đặc điểm cơ bản của các loại mạng học sâu hiện đại sẽ được trình bày để làm cơ sở thiết kế mạng học sâu phù hợp cho phân loại bệnh UTV từ tập ảnh UTV thu thập và

## 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

---

xử lý được ở các chương trước.

### 5.1.1 Vai trò mạng học sâu cho phân loại UTV

Khi áp dụng các mạng học sâu, việc sở hữu một tập dữ liệu ảnh đủ lớn và cân bằng là điều tối quan trọng. Do đó, một trong những phương pháp phổ biến nhằm tăng cường dữ liệu là sử dụng nhiều phép biến đổi khác nhau. Cụ thể, các kỹ thuật tăng cường dữ liệu có thể bao gồm biến đổi hình học, tăng cường ở mức điểm ảnh, giả lập màu, xoá ngẫu nhiên và các bộ lọc hạt nhân [146]. Trong nghiên cứu này, chúng tôi áp dụng các kỹ thuật xoay và lật ảnh để mở rộng quy mô tập dữ liệu. Nhờ quá trình nâng cao và cân bằng tập ảnh, bộ phân loại EfficientNet-B7 có khả năng đạt được độ chính xác phân loại cao.

Mạng học sâu được sử dụng rộng rãi trong ứng dụng y học, phổ biến nhất là trong chẩn đoán UTV dựa vào ảnh X-quang vú [147]–[149]. Việc ứng dụng mạng học sâu mang lại nhiều lợi ích và tiềm năng trong việc gia tăng hiệu quả và độ chính xác của quy trình chẩn đoán. Trong các nghiên cứu gần đây, việc áp dụng mạng CNN cho chẩn đoán UTV chủ yếu tập trung vào các ứng dụng phân loại ảnh X-quang vú [59], [150], [151]. Cụ thể, một nghiên cứu đã dùng mạng AlexNet để phân loại ảnh nhũ ảnh bình thường, UTV xâm lấn (Invasive Breast Cancer – IBC) và ung thư biểu mô ống tại chỗ (Ductal Carcinoma in Situ – DCIS), đạt độ chính xác 83,1% với tập dữ liệu DDSM [152]. Mạng GoogLeNet được sử dụng để phân loại ảnh X-quang vú với các loại: bình thường, IBC và DCIS đạt độ chính xác 96,37% trên một tập dữ liệu thu thập riêng [153].

Với sự phát triển mạnh mẽ của các mạng học sâu, kết quả nhận dạng và phân loại đã đạt độ chính xác cao. Một nghiên cứu đã áp dụng mạng VGG16 để phát hiện khối u trên ảnh X-quang vú, đạt độ nhạy 95,4% và độ đặc hiệu 98,3% trên tập dữ liệu CBIS [154]. Trong nghiên cứu của Li [155], mạng ResNet50 được dùng để phát hiện khối u trên ảnh nhũ ảnh, thu được độ nhạy 96,2% và độ đặc hiệu 99,1% trên tập dữ liệu INbreast. Bên cạnh đó, phân loại theo mức độ ác tính sử dụng mạng DenseNet cho phân loại UTV dựa trên năm mức độ của tập ảnh nhũ ảnh, cho độ chính xác

## 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

---

92,5% trên tập dữ liệu Breast Cancer Histopathology Image Database [156]. Trong nghiên cứu [61], mạng AlexNet được sử dụng để phân loại UTV theo năm mức độ của tập ảnh nhũ ảnh, đạt độ chính xác 83,4% trên các tập ảnh này.

Phiên bản đầu tiên của mô hình EfficientNet là EfficientNet-B0, sau đó mô hình này được phát triển thành EfficientNet-B1 đến EfficientNet-B7. Các kết quả cho thấy EfficientNet-B7 đạt độ chính xác cao hơn so với nhiều mô hình học sâu khác. Ngoài ra, EfficientNet-B7 thường cho độ chính xác tốt trên tập dữ liệu ImageNet, trong khi kích thước mạng chỉ nhỏ hơn 8,4 lần và tốc độ nhanh hơn 6,1 lần so với các mạng học sâu hiện tại [32]. Trong các nghiên cứu gần đây, các mô hình EfficientNet đã được kết hợp với nhiều thuật toán khác trong nhiều lĩnh vực phân loại. Cụ thể, Lung-EfficientNet được đề xuất cho bài toán phân loại ung thư phổi dựa trên tập ảnh CT-scan, đạt độ chính xác 99,10% [157]. Một nghiên cứu khác áp dụng các mô hình EfficientNet-U-Net để phân đoạn khối u thận trên ảnh CT-scan, đạt độ chính xác cao trong phân đoạn [158], [159]. Trong bài báo này, chúng tôi thực hiện tinh chỉnh phù hợp để cải thiện độ chính xác phân loại của EfficientNet-B7 cho nhiệm vụ phân loại ảnh có CROI trên tập dữ liệu VinDr-Mammo và thu được kết quả có độ chính xác cao.

Với sự phát triển của mạng học sâu, các kỹ thuật học sâu đã được ứng dụng hiệu quả trong phân loại UTV, cho phép chẩn đoán sớm và qua đó nâng cao tỷ lệ sống của bệnh nhân [9]–[11]. Nhiều nghiên cứu gần đây đã áp dụng các phương pháp học sâu để phân loại UTV dựa trên nhiều cách tiếp cận ảnh khác nhau. Dheeb Albashish và cộng sự [160] đề xuất mô hình mạng VGG16 kết hợp transfer learning để trích xuất đặc trưng từ tập ảnh mô bệnh học BreakHis, đạt độ chính xác 89%. S. Shamy và cộng sự [161] giới thiệu mô hình dựa trên k-means GMM và CNN, sử dụng vùng quan tâm (ROI) để trích xuất đặc trưng, đạt độ chính xác 95,8%. Một nghiên cứu khác của V. Sansya Vijayam [162] tập trung vào thuật toán Lloyd cho phân cụm kết hợp CNN để phân loại, đạt độ chính xác 96%. Các nghiên cứu này cho thấy hiệu quả của việc ứng dụng học sâu trong phân loại UTV.

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

**Bảng 5.1 Tóm tắt các phương pháp tiền xử lý và học sâu cho bài toán phân loại UTV**

Tác giả	Tập dữ liệu (Ngõ ra)	Phương pháp tiền xử lý	Mô hình mạng học sâu	Kết quả
Nan Wu, và các cộng sự. [163]	The NYU breast cancer screening dataset V1.0 with over 1,000,000 images (Benign and Malignant)	Image normalization, data augmentation.	Custom ResNet-based CNN	AUC = 0.895%
Arslan Khalid, và các cộng sự. [164]	The assessment measures EDA dataset 3,002 mammographic images (Benign and Malignant)	Removal of low-variance features, univariate feature selection, recursive feature elimination.	CNNI-BCC	Mang lại hiệu quả cao
Zede Zhu, và các cộng sự. [59]	Breast Ultrasound Images, MIAS, Mini-DDSM,	Wiener filtering, total variation filtering, image segmentation.	ViT, U-KAN	ViT accuracy: 99.3%, U-KAN: 93.3% accuracy
Z. Wang, và các cộng sự. [165]	3,002 digital mammography images (Benign and Malignant)	CNN deep feature extraction, feature fusion.	CNN-ELM hybrid	Cải thiện hiệu suất phân loại

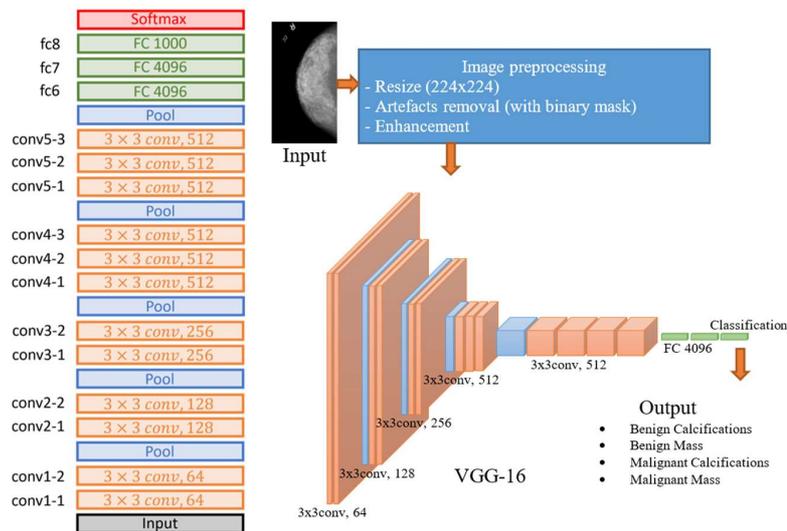
Bảng 5.1 trình bày sự so sánh một số công trình nghiên cứu tập trung vào phân loại UTV bằng các phương pháp máy học và mạng học sâu. Mỗi nghiên cứu áp dụng các phương pháp tiền xử lý khác nhau như chuẩn hóa ảnh, lựa chọn đặc trưng, lọc và tăng cường độ tương phản nhằm cải thiện chất lượng hình ảnh trước khi phân loại. Các phương pháp phân loại trải dài từ các mô hình ML truyền thống [164] đến những mô hình học sâu tiên tiến (ví dụ: CNN, Vision Transformer, U-Net, Extreme Learning Machine) [163][59] [165]. Nguồn dữ liệu bao gồm ảnh nhũ ảnh từ các chương trình tầm soát quy mô lớn (ví dụ, hơn 200.000 ảnh nhũ ảnh trong một nghiên cứu) [163] cũng như nhiều bộ dữ liệu công khai (BreastDM, MIAS, BUSI, BreakHis, DDSM) [59]. Kết quả phân loại chủ yếu tập trung phân biệt khối u lành và ác tính, một số nghiên cứu mở rộng tới phân loại theo phân nhóm [59]. Các kết quả cho thấy các mô hình học sâu nhìn chung vượt trội so với phương pháp ML truyền thống, trong đó các

## 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

mô hình dựa trên CNN và Transformer đạt độ chính xác cao nhất (tới 99,3% trong một số trường hợp). Tuy nhiên, khả năng khái quát và hiệu quả tính toán vẫn còn là thách thức, nhấn mạnh nhu cầu phát triển thêm các hệ thống học sâu nhằm nâng cao độ chính xác phân loại UTV. Nghiên cứu so sánh này làm nổi bật những tiến bộ nhanh chóng trong phân loại UTV dựa trên AI và tiềm năng của học sâu trong cải thiện độ chính xác chẩn đoán bệnh UTV cũng như hỗ trợ BS trong thực hành lâm sàng.

### 5.1.2 Mô hình mạng VGG16

VGG16 ban đầu có kích thước độ sâu là 23 và đúng như tên gọi, nó có 16 lớp. Nó bao gồm năm khối và mỗi khối chứa hai hoặc ba lớp tích chập theo sau là lớp Maxpooling. Lớp tích chập thực hiện thao tác tuyến tính bằng cách nhân một mảng dữ liệu đầu vào với mảng trọng số hai chiều, được gọi là bộ lọc hoặc hạt nhân. Hình 5.1 thể hiện cấu trúc của mạng VGG16, trong đó bao gồm các lớp Conv 3x3 sắp xếp theo từng nhóm, xen kẽ với các lớp Pool, hàm Softmax là hàm kích hoạt ngõ ra, đi kèm với các lớp FC.



**Hình 5.1. Kiến trúc của mạng VGG16 đã tinh chỉnh**

Theo mặc định, lớp đầu vào của kiến trúc yêu cầu kích thước của hình ảnh, là hình ảnh RGB, là  $224 \times 224 \times 3$ , kích thước đầu vào của chúng tôi cho lớp chập đầu tiên là  $224 \times 224 \times 3$ . Trong ngữ cảnh này, khối đầu tiên chứa hai lớp tích chập với

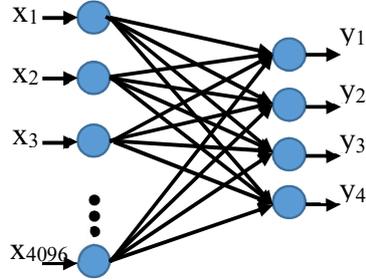
## Chương 5 PHÂN LOẠI ẢNH UTV SỬ DỤNG MẠNG HỌC SÂU VÀ HỆ THỐNG HỖ TRỢ CHẨN ĐOÁN UTV TỪ XA SỬ DỤNG MẠNG HỌC SÂU

---

64 kênh có kích thước hạt nhân  $3 \times 3$  và cùng một phần đệm, theo sau là lớp Maxpooling  $2 \times 2$  có bước tiến  $2 \times 2$ . Tương tự, khối thứ hai chứa hai lớp tích chập gồm 128 kênh, với kích thước hạt nhân  $3 \times 3$ , theo sau là lớp Maxpooling có bước  $2 \times 2$  như trong khối một. Ba khối cuối cùng chứa ba lớp chập theo sau là lớp Maxpooling. Kích thước kênh của ba lớp chập trong khối 3, 4 và 5 lần lượt là 256, 512 và 512, tất cả đều có cùng kích thước hạt nhân là  $3 \times 3$ . Hình ảnh đầu vào ban đầu được thu nhỏ xuống một nửa kích thước trong mỗi lớp Maxpooling. Sau chông các lớp chập và Maxpooling, bản đồ đặc trưng xuất ra từ lớp Maxpooling cuối cùng có kích thước  $7 \times 7 \times 512$ . Một lớp làm phẳng đã được thêm vào để tạo vectơ đặc trưng  $1 \times 25.088$ . Các lớp tích chập trong mô hình VGG16 sử dụng hàm mất mát ReLU, bằng 0 với  $x < 0$  và bằng  $x$  với mọi  $x \geq 0$  và được mô tả như sau:

$$ReLU(x) = \max\{x, 0\} = \begin{cases} x, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases} \quad (5.1)$$

Trong bài viết này, việc xây dựng mô hình VGG16 về cơ bản có hai giai đoạn: đào tạo và tinh chỉnh. Trong quá trình đào tạo, hai lớp cuối cùng trong tổng số 16 lớp trong mô hình VGG16 được tinh chỉnh, trọng số ở các lớp còn lại được giữ nguyên từ nguyên mẫu được đào tạo trên ImageNet. Ở lớp đầu ra của mô hình, các lớp dày đặc và lớp FC của mô hình mạng được cập nhật các tính năng mới của hình ảnh trong quá trình tinh chỉnh. Hơn nữa, lớp FC cuối cùng sẽ được sửa đổi để phân loại bốn loại hình ảnh tổn thương vú. Cụ thể, để đạt được mục đích phân loại 4 lớp ảnh UTV là Khối u lành tính, khối u ác tính, Vôùi hóa dạng lành tính và vôùi hóa dạng ác tính, lớp FC cuối cùng được điều chỉnh thành 4 đầu ra. Cấu trúc của lớp FC cuối cùng có thể hiện ở Hình 5.2.



**Hình 5.2. Lớp FC cuối cùng đã được sửa đổi cho đầu ra thứ 4.**

Do đó, trọng số được đào tạo trước của các lớp được chuyển giao bằng cách sử dụng phương pháp học chuyển giao và sau đó mô hình được đào tạo lại và tinh chỉnh với các tập dữ liệu hình ảnh mới. Mô hình này sử dụng hàm kích hoạt 'Softmax' để tính toán xác suất cho mỗi lớp và tạo ra kết quả dự đoán. Do đó, Softmax được thể hiện bằng phương trình sau:

$$z_j = \frac{e^{x_j}}{\sum_{i=1}^n e^{x_i}}; \quad j = 1, 2, \dots, n \quad (5.2)$$

Để đào tạo mô hình trong nghiên cứu này, số lượng epoch tối đa được đặt thành 350 và kích thước loạt là 16. Thuật toán tối ưu hóa Adam được áp dụng với tốc độ học là 0,001. Ngoài ra, thuật toán Adam là phiên bản nâng cao của Stochastic Gradient Descent, là sự kết hợp của thuật toán Momentum và RMSProp và được mô tả bằng phương trình toán học sau:

$$w_{t+1} = w_t - \alpha m_t \quad (5.3)$$

Trong đó:

$$m_t = \beta m_{t-1} + (1 - \beta) \left[ \frac{\delta L}{\delta w_t} \right] \quad (5.4)$$

Với  $w_{t+1}$  là trọng số tại thời điểm  $t+1$ .  $w_t$  là trọng số tại thời điểm  $t$ .  $\alpha$  là tốc độ học.  $m_t$  là tổng hợp các gradient tại thời điểm  $t$ .  $\delta L$  là đạo hàm của Hàm mất mát.  $\delta w_t$  là đạo hàm của trọng số tại thời điểm  $t$ . Và  $\beta$  là tham số trung bình động.

Hiệu chỉnh tham số mô hình học sâu:

Mô hình VGG16 trong luận án được khởi tạo từ trọng số tiền huấn luyện trên

## 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

---

ImageNet, sau đó được hiệu chỉnh có chọn lọc (selective fine-tuning) nhằm thích nghi với đặc thù của ảnh X-quang vú, vốn khác biệt đáng kể so với ảnh tự nhiên.

Quá trình hiệu chỉnh tập trung vào ba nhóm chính: Cấu trúc tầng cuối (classification head); Chiến lược đóng/mở băng (freeze/unfreeze) các tầng tích chập; Các siêu tham số huấn luyện (learning rate, optimizer, batch size, epoch, regularization). Mục tiêu là tăng khả năng phân biệt tổn thương, đồng thời giảm overfitting trong bối cảnh dữ liệu y tế có quy mô vừa và mất cân bằng.

Cụ thể đối với mô hình VGG16 trên, các thông số và cấu trúc được hiệu chỉnh bao gồm:

(a) Hiệu chỉnh tầng Fully Connected (FC): so với kiến trúc gốc của VGG16, luận án đã loại bỏ hoàn toàn 3 tầng FC gốc (4096–4096–1000), thay thế bằng Global Average Pooling (GAP), 1–2 tầng Dense với số nút giảm (256 hoặc 512), Dropout (tỷ lệ 0.3–0.5), Tầng Dense cuối với số nút tương ứng số lớp (2, 3, 4 hoặc 5).

Việc này nhằm giảm mạnh số lượng tham số huấn luyện, hạn chế overfitting trên dữ liệu X-quang và giữ lại các đặc trưng không gian quan trọng từ tầng tích chập cuối.

(b) Chiến lược freeze/unfreeze tầng tích chập: bằng cách chia quá trình huấn luyện thành 2 giai đoạn: Giai đoạn 1: đóng băng toàn bộ các tầng tích chập, chỉ huấn luyện head phân loại; Giai đoạn 2: mở băng 1–2 block tích chập cuối (Conv5\_x) để fine-tune. Việc này nhằm giữ ổn định các đặc trưng cơ bản (edge, texture), đồng thời cho phép mạng thích nghi với đặc trưng mô vú và vi vôi hóa ở tầng sâu.

(c) Siêu tham số huấn luyện: Optimizer: Adam; Learning rate:  $1e-4 \rightarrow 1e-5$  (giảm dần); Batch size: 16–32; Epoch: 100–350 (early stopping); Loss: categorical cross-entropy.

Trong nghiên cứu này, VGG16 được sử dụng để phân loại bốn loại tổn thương vú (Khối u lành tính; Khối u ác tính; Vôi hóa dạng lành tính; Vôi hóa dạng ác tính) của các bộ ảnh X-quang vú có tổn thương. Cấu hình máy tính được sử dụng để đào tạo mạng VGG16 bao gồm bộ xử lý Intel Core i9-9980HK, GPU NVidia Tesla P4,

## 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

Bộ nhớ RAM DDR4 32GB và SSD 1TB để lưu trữ. Với cấu hình này, mạng hoạt động tương đối tốt và đưa ra kết quả phân loại chính xác hơn. Các kết quả của mô hình này sẽ được trình bày ở phần sau.

### 5.1.3 Mô hình mạng EfficientNet

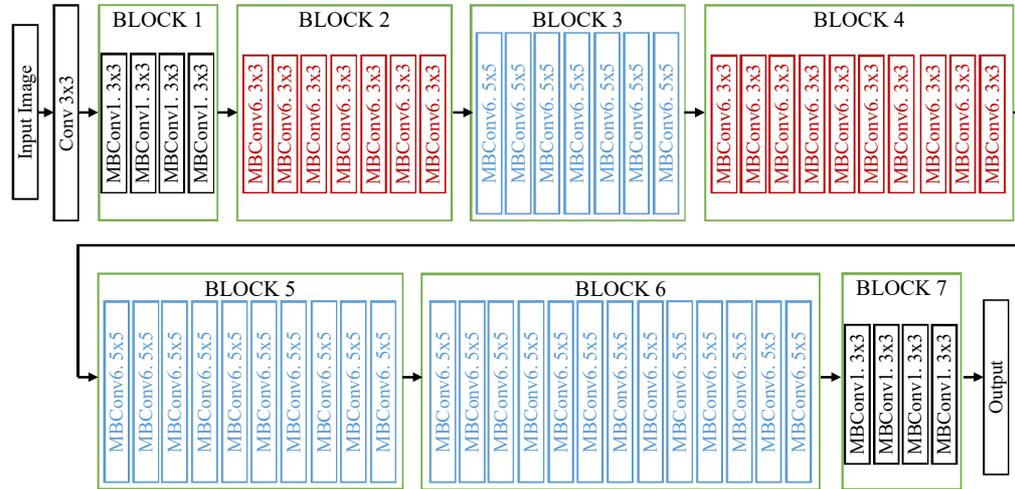
Năm 2019 nhóm nghiên cứu của Google đã giới thiệu mạng học sâu mới của họ - EfficientNet [32]. Mô hình CNN này không chỉ nhằm tăng độ chính xác mà còn hướng tới việc cải thiện hiệu suất của mô hình bằng cách giảm số tham số và khối lượng tính toán (FLOPS - Floating Point Operations Per Second) so với các mô hình khác.

**Bảng 5.2 Chi tiết kiến trúc của mạng EfficientNet-B0**

Tầng $i$	Toán tử $\hat{F}_i$	Độ phân giải $\hat{H}_i \times \hat{W}_i$	Số kênh $\hat{C}_i$	Số lớp $\hat{L}_i$
1	Conv 3x3	224×224	32	1
2	MBCConv1 k3x3	112×112	16	1
3	MBCConv6 k3x3	112×112	24	2
4	MBCConv6 k5x5	56×56	40	2
5	MBCConv6 k3x3	28×28	80	3
6	MBCConv6 k5x5	14×14	112	3
7	MBCConv6 k5x5	14×14	192	4
8	MBCConv6 k3x3	7×7	320	1
9	Conv 1x1 & Pooling & FC	7×7	1280	1

Các tác giả sử dụng phương pháp tìm kiếm kiến trúc thần kinh để xây dựng kiến trúc mạng, phiên bản đầu tiên của mạng EfficientNet là EfficientNet-B0 với mô tả chi tiết như ở Bảng 5.2. Từ mạng EfficientNet-B0, bằng cách tăng số lớp và mở rộng, sẽ thu được các mạng tiếp theo từ EfficientNet-B1 đến EfficientNetB7. Kiến trúc của mô hình mạng EfficientNet-B7 được trình bày ở Hình 5.3 bên dưới:

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



**Hình 5.3. Cấu trúc của mạng EfficientNet-B7 được tinh chỉnh**

Các mô hình mạng EfficientNet đạt được độ chính xác tốt nhất trên các tập dữ liệu CIFAR-100 (91,7%), Flowers (98,8%). Trong nghiên cứu này, việc tinh chỉnh thích hợp được thực hiện nhằm nâng cao hiệu suất phân loại của mạng EfficientNet trên các tập dữ liệu hình ảnh UTV đã đề cập ở trên.

Hiệu chỉnh mô hình EfficientNet-B7: các thông số và cấu trúc của mô hình được hiệu chỉnh như sau:

(a) Giữ nguyên backbone – hiệu chỉnh head phân loại: Backbone EfficientNet-B7 được giữ nguyên cấu trúc MBConv + SE, thay thế tầng phân loại gốc bằng: Global Average Pooling, Dense (512 hoặc 1024), Dropout (0.4–0.5), Dense output theo số lớp. Việc hiệu chỉnh này nhằm tận dụng tối đa khả năng trích xuất đặc trưng đa tỉ lệ của EfficientNet, giảm nguy cơ overfitting khi số mẫu huấn luyện không quá lớn.

(b) Fine-tuning có chọn lọc các block sâu: giai đoạn đầu freeze toàn bộ backbone. Giai đoạn sau: unfreeze các block MBConv cuối (high-level features). Việc này cho phép mô hình điều chỉnh các đặc trưng ngữ nghĩa cao, tăng khả năng phát hiện các dấu hiệu tinh vi như Biên khối u mờ, Cụm vi vôi hóa, Bất đối xứng cấu trúc.

(c) Tham số huấn luyện: Optimizer: Adam; Learning rate:  $1e-4 \rightarrow 1e-5$  (cosine decay / step decay); Batch size: 8–16 (do kích thước mô hình lớn); Epoch: 150–300;

## 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

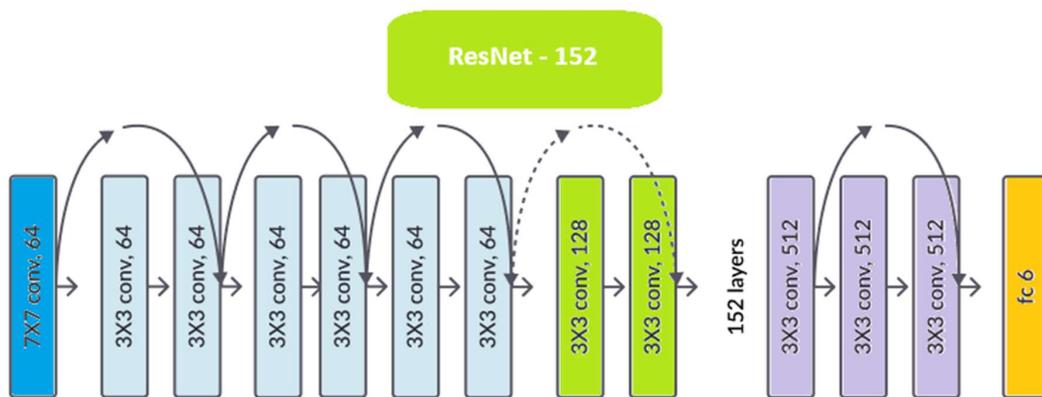
Loss: categorical cross-entropy

So với EfficientNet-B7 không fine-tune hoặc fine-tune toàn bộ mạng ngay từ đầu, chiến lược hiệu chỉnh cho thấy tốc độ hội tụ nhanh hơn, loss giảm ổn định, độ chính xác và F1-score cao hơn, đặc biệt ở các lớp thiếu số, phân loại đa lớp và BIRADS.

### 5.1.4 Mô hình mạng Resnet-152

Resnet-152 thuộc mạng Resnet (Residual Network), là một mạng kiến trúc CNN được thiết kế với hàng trăm lớp và có khả năng khắc phục hiện tượng Vanishing Gradient dẫn tới quá trình học tập không tốt [166]. Resnet-152 là mạng thần kinh nơ-ron tích chập 152 lớp (150 lớp tích chập, 1 lớp Maxpool và 1 lớp Average Pool) và hoàn toàn phát triển dựa trên kiến trúc Resnet ban đầu là Resnet-34 được bổ sung thêm nhiều lớp tích chập vào CNN hơn.

Điều này giúp cho Resnet-152 có khả năng học tập và chiết suất các đặc trưng của hình ảnh đầu vào tốt hơn các mạng Resnet khác và nhờ vào đó độ chính xác của model pretrained Resnet-152 khá cao.



**Hình 5.4. Cấu Trúc Của Resnet**

Kiến trúc Resnet-152 bao gồm các thành phần sau:

- Conv1: Kích thước 7x7 Kernel Convolution với stride là 2
- Conv2\_x: là lớp tổng hợp (Pooling layer) với stride là 2 và 9 lớp filter (3x3,64).

## 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

---

- Conv3\_x: Có 24 lớp filter (3x3,128).
- Conv4\_x: có 108 lớp filter (3x3,256).
- Conv5\_x: có 9 lớp filter (3x3,512).
- 1 lớp Average Pool

Resnet-152 là một kiến trúc dày đặc, hơn 152 lớp layer cho phép mô hình có thể học được các chi tiết đặc trưng của nhữ ảnh (Input đầu vào). Nhờ vào đó mà mô hình có khả năng tiếp cận và học hỏi rất nhanh trong việc chẩn đoán bệnh UTV với độ chính xác cao.

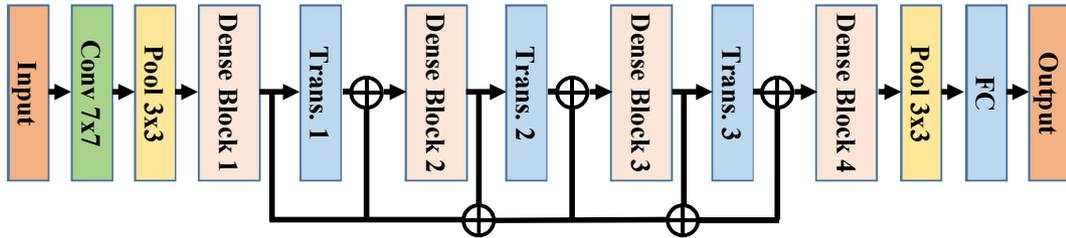
Với việc sử dụng lại mô hình đã được huấn luyện trước trên tập dữ liệu ImageNet cho phép rút ngắn thời gian huấn luyện, mô hình nhanh chóng học được các đặc trưng chính của hình ảnh. Sau đó việc tinh chỉnh trên tập dữ liệu nghiên cứu giúp mô hình cải thiện hiệu suất phân loại và nâng cao độ chính xác.

### 5.1.5 Mô hình mạng Resnet-201

Densenet-201 là một kiến trúc mạng nơ-ron tích chập (CNN) được phát triển để thực hiện các nhiệm vụ nhận diện hình ảnh và phân loại. Điểm đặc biệt của DenseNet là cách mà nó kết nối các lớp nơ-ron trong quá trình học, trong đó các ngõ ra của các lớp trước kết nối trực tiếp vào đầu vào của các lớp sau thay vì thực hiện phép cộng như ResNet.

Mô hình này giúp tăng cường độ sâu của mạng và chia sẻ thông tin hiệu quả, giúp cải thiện hiệu suất với số lượng tham số ít hơn so với một số kiến trúc khác. Densenet-201 được biết đến với độ chính xác cao trên bộ dữ liệu ImageNet, cho thấy hiệu quả của mô hình trong việc thực hiện các nhiệm vụ nhận diện và phân loại ảnh [24].

## 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



Hình 5.5 Cấu trúc của mạng DenseNet-201

Đối với yêu cầu đề tài, model pretrained DenseNet-201 rất phù hợp các dạng đề tài phân loại và nhận diện ảnh mà độ chính xác lại cao nên hầu hết DenseNet-201 được sử dụng phổ biến trong các đề tài phân loại.

### 5.2. PHÂN LOẠI ẢNH UTV THEO TỔN THƯƠNG DÙNG MẠNG HỌC SÂU

#### 5.2.1 Bài toán phân loại có ung thư - không có ung thư

Trong nghiên cứu đầu tiên, một bài toán với ngõ ra tương đối đơn giản được chọn để đánh giá sự hiệu quả của các phương pháp xử lý trong nghiên cứu này. Các ngõ ra được chọn cho bài toán phân loại có ung thư là: bình thường, lành tính (có tổn thương nhưng lành tính, không phải là ung thư, và ác tính (tức là ung thư). Chúng tôi sử dụng phương pháp transfer learning với các mạng DenseNet và ResNet trong bài toán phân loại này.

##### a) Tập dữ liệu

Việc lựa chọn tập dữ liệu cho nghiên cứu phụ thuộc rất nhiều vào ngõ ra của bài toán: tập dữ liệu được cho cần phải có thông tin gán nhãn phù hợp với ngõ ra đã chọn, nếu không chúng ta sẽ không có đủ dữ liệu cần thiết để huấn luyện mô hình. Trong trường hợp này, tập DDSM cho thấy chúng có đủ 3 phân loại đã lựa chọn như mô tả ở bảng 5.3. Một tập con của DDSM đã được chọn lọc là Mini-DDSM được sử dụng do sự sẵn sàng của các hình ảnh và việc đã loại bỏ các hình ảnh kém chất lượng. Tập Mini-DDSM đã được gán nhãn sẵn với 3 phân loại phù hợp là bình thường, lành tính và ác tính. Số lượng của mỗi phân loại thể hiện ở bảng dưới đây:

## 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

**Bảng 5.3 Chi tiết bộ dữ liệu Mini-DDSM**

Nhãn	Số Lượng	Tỷ lệ
Lành tính (Benign)	2685	34,4%
Ác tính (Cancer)	2716	34,7%
Bình thường (Normal)	2408	30,9%

Bảng số liệu cho thấy số lượng của các nhãn được chia tỷ lệ chênh lệch không quá nhiều, cụ thể nhãn lành tính “Benign” có số lượng 2685 mẫu chiếm 34,4% và nhãn ác tính “Cancer” có số lượng 2716 mẫu chiếm 34,7%, cuối cùng là nhãn bình thường “Normal” có số lượng 2408 chiếm 30,9%. Điều này có nghĩa là số lượng mẫu chênh lệch không quá nhiều và không ảnh hưởng nhiều đến quá trình huấn luyện mô hình. Tiếp theo tập dữ liệu này được chia thành 2 tập con là tập huấn luyện và tập xác thực để huấn luyện đánh giá hiệu suất mô hình như bảng 5.4. Việc chia dữ liệu này được thực hiện cụ thể như sau:

**Bảng 5.4 Chi tiết số lượng Train và Vaid**

Label	Số Lượng	Tỷ lệ
Train	6247	79,9%
Vaid	1562	20,1%

Trong bảng trên, tỷ lệ chia hai tập lệnh Train và Vaid có sự chênh lệch rất lớn, số lượng mẫu trong tập Train gấp 4 lần tập Vaid. Đây là chủ đích của nhóm nhằm mục đích muốn mô hình có thể học được nhiều đặc tính của nhữ ảnh trong tập Train để mô hình có hiệu suất và độ chính xác cao.

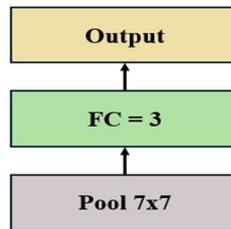
### **b) Tinh chỉnh mô hình**

Trước khi huấn luyện sử dụng kỹ thuật học chuyển giao, các trọng số huấn luyện của các mô hình đã được huấn luyện trước vẫn được sử dụng, vì hầu hết các trọng số huấn luyện này đã được xác định và cung cấp cho người dùng để cho ra kết quả tốt nhất. Tuy nhiên, số lượng đầu ra của mỗi model pretrained hoàn toàn khác nhau và khác với đầu ra của mô hình chẩn đoán bệnh UTV nên vì thế ta sẽ thay đổi đầu ra của các model pretrained sao cho phù hợp với mô hình.

## 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

---

Mục tiêu đặt ra của đề tài là huấn luyện mô hình chẩn đoán 3 phân loại là benign, cancer, normal nên vì thế ta thay đổi đầu ra của các model pretrained trong lớp Fc và Classifier sao cho phù hợp và model pretrained được sử dụng cho bài toán này là Resnet-152 và DenseNet-210. Để đạt được mục tiêu của bài toán đặt ra, ta thay đổi lớp FC của model pretrained bằng 3 để thỏa yêu cầu bài toán là phân loại 3 labels: Normal, Benign, Cancer như trình bày ở hình 5.6.



**Hình 5.6. Thay đổi lớp FC = 3**

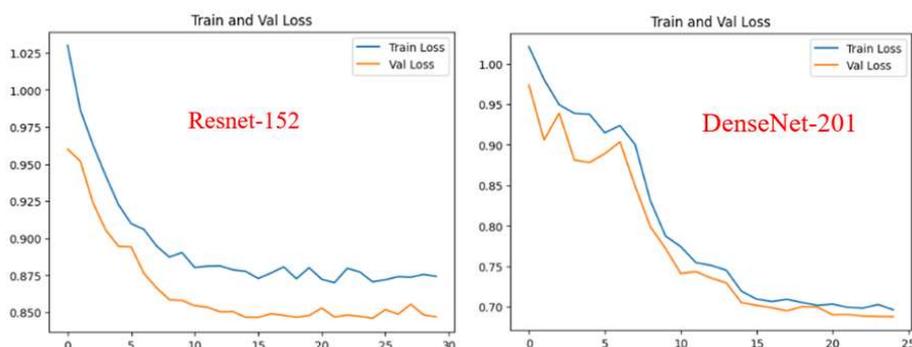
Trong mô hình mạng này, hàm mất mát CrossEntropyLoss được sử dụng cho các lớp. CrossEntropyLoss là một hàm mất mát phổ biến được sử dụng trong huấn luyện mô hình phân loại hình ảnh. Nó được dùng để tính toán mất mát cross-entropy giữa các mức logits đầu ra và nhãn tương ứng, và là một phần của quá trình tối ưu hóa mô hình để đạt được dự đoán phân loại chính xác. Công thức tính của hàm được cho dưới đây:

$$CE = -\log L(\{y^{(n)}\}, \{\hat{y}^{(n)}\}) = \sum_n [-\sum_i y_i \log \hat{y}_i^{(n)}] = \sum_n H(y^{(n)}, \hat{y}^{(n)}) \quad (5.5)$$

Hàm CrossEntropyLoss được dùng để đánh giá sự tương đồng giữa các mức logic ngõ ra của hình ảnh và các nhãn tương ứng để xác định phân loại.

### c) Kết quả phân loại

## 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



**Hình 5.7 Hàm mất mát huấn luyện và kiểm tra tương ứng của mô hình ResNet-152 và DenseNet-201**

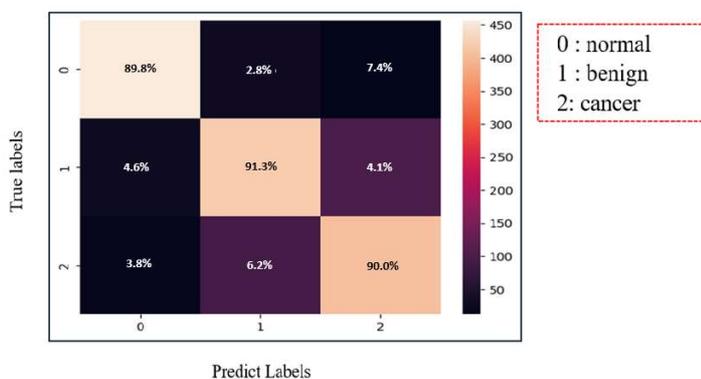
Hình 5.7 cho thấy kết quả sau nhiều lần thử nghiệm, và độ chính xác (Accuracy) cao nhất mà mô hình huấn luyện chẩn đoán bệnh UTV đạt được là 92,01%. Kết quả này không cao như nhóm kỳ vọng mô hình có thể đạt được, vì thế nhóm quyết định thử nghiệm và tra cứu để so sánh kết quả của mô hình với các model pretrained được áp dụng ở những mô hình khác.

**Bảng 5.5 Độ chính xác huấn luyện và kiểm tra của 2 mô hình**

Mô hình	Độ chính xác huấn luyện	Độ chính xác kiểm tra
ResNet-152	91,21%	89,7%
DenseNet-201	92,01%	90,1%

Bảng 5.5 cho thấy cả 2 mô hình đều đạt độ chính xác rất tốt trên tập ảnh Mini-DDSM. Trong đó DenseNet-201 cho độ chính xác cao hơn một chút. Kết quả độ chính xác cũng giảm khoảng 2% khi kiểm tra so với kết quả khi huấn luyện. Điều này là bình thường đối với các mô hình mạng học sâu.

## 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



**Hình 5.8** Ma trận nhầm lẫn của kết quả bài toán phân loại 3 ngõ ra: bình thường, lành tính, và ác tính

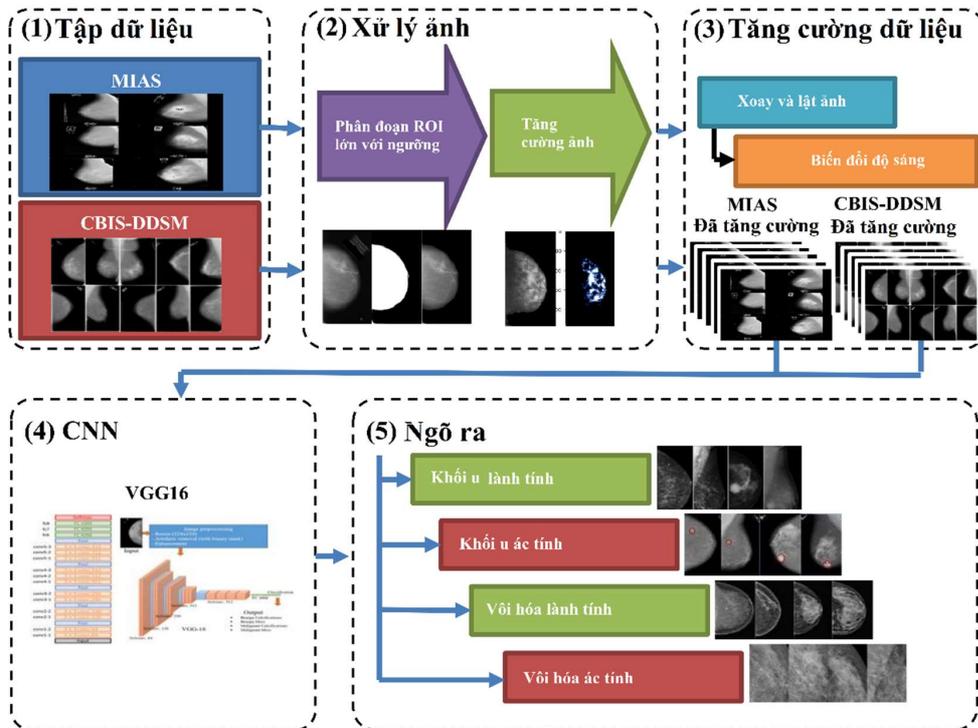
Hình 5.8 trên biểu diễn ma trận nhầm lẫn của mạng DeneseNet-201 trong bài toán phân loại, các ô trên đường chéo đạt giá trị cao xấp xỉ 90% cho thấy độ chính xác của mô hình, trong khi các ô còn lại có giá trị từ 2.8% đến 7.4%.

### 5.2.2 Bài toán phân loại các loại tổn thương dựa trên tập ảnh X-quang vú MIAS

#### a) Tập dữ liệu

Trong phần này, phương pháp của một quy trình phân loại tổn thương vú được đề xuất như mô tả trong Hình 5.9. Cụ thể, các tập dữ liệu được lấy từ CBIS-DDSM và MIAS và các tập dữ liệu hình ảnh tăng cường sử dụng xoay, lật và thay đổi độ sáng, trong đó bốn loại hình ảnh tổn thương (Khối u lành tính, Khối u ác tính, Vôi hóa dạng lành tính, Vôi hóa dạng ác tính) được sử dụng. Tất cả các tập dữ liệu hình ảnh được xử lý trước khi phân đoạn Otsu để chọn các đặc điểm chứa ROI lớn nhất trong mỗi hình ảnh tổn thương. Do đó, các tập dữ liệu ROI lớn nhất được đưa vào mạng học sâu VGG16. Mạng này được cấu hình để tinh chỉnh hai lớp cuối cùng và cũng thêm các lớp đầu ra thích hợp. Cuối cùng, ma trận nhầm lẫn được dùng để đánh giá hiệu quả phân loại của phương pháp được đề xuất [142].

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



**Hình 5.9 Sơ đồ khối phân loại và quy trình: (1) hai tập dữ liệu 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.**

Ngoài ra, phân này thực hiện ba trường hợp tập ảnh đầu vào vào đầu vào của mạng VGG16 để đánh giá độ chính xác giữa chúng. Đặc biệt, các bộ ảnh được sử dụng cho ba trường hợp như bộ ảnh không có phân đoạn Otsu, bộ ảnh chỉ có phân đoạn Otsu và bộ ảnh có ROI lớn nhất sử dụng phân đoạn Otsu.

Nhũ ảnh được sử dụng trong nghiên cứu này từ hai nguồn hình ảnh: CBIS-DDSM và MIAS. Hơn nữa, CBIS-DDSM là phiên bản cập nhật của DDSM cung cấp dữ liệu để truy cập và phân đoạn ROI được cải thiện. DDSM ban đầu là CSDL gồm 2.620 nghiên cứu chụp nhũ ảnh phim được quét, trong khi CBIS-DDSM có 1.644 hình ảnh của bốn loại tổn thương vú (Khối u lành tính; Khối u ác tính; Vôi hóa lành tính; Vôi hóa ác tính) và số hình ảnh trong mỗi loại cùng thông tin của chúng được hiển thị trong Bảng 5.6. Theo đó, kích thước ảnh đã được đồng nhất là 224x224 pixel và mã

## 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

màu là RGB.

**Bảng 5.6 Thông tin của tập dữ liệu CBIS-DDSM.**

Đặc tính	Giá trị
Tổng số ảnh	1644
Kích thước ảnh	224 × 224
Mã hóa màu	RGB
Số ảnh u lành tính	753
Số ảnh vôi hóa lành tính	414
Số ảnh vôi hóa ác tính	339
Số ảnh u ác tính	419

Tương tự như vậy, tập dữ liệu Chụp nhũ ảnh MIAS bao gồm 1.128 hình ảnh của 362 phụ nữ và chỉ có 322 hình ảnh của bốn loại tổn thương vú cùng thông tin và số hình ảnh của chúng được mô tả trong Bảng 5.7. Các thông số về kích thước ảnh và mã màu cũng tương tự.

**Bảng 5.7 Thông tin của tập dữ liệu MIAS.**

Đặc tính	Giá trị
Số lượng ảnh	322
Kích thước ảnh	224 × 224
Mã hóa	RGB
Số lượng ảnh u lành tính	37
Số lượng ảnh vôi hóa lành tính	12
Số lượng ảnh vôi hóa ác tính	13
Số lượng ảnh u ác tính	19

### b) Huấn luyện mô hình

Để đánh giá hiệu quả của các mô hình, cùng một cấu hình máy tính được sử dụng trong quá trình huấn luyện các mạng trên. Cấu hình máy tính sử dụng trong huấn

## 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

---

luyện mạng VGG16 bao gồm bộ xử lý Intel Core i9-9980HK, NVidia GPU Tesla P4, Bộ nhớ Ram DDR4 32GB và SSD 1TB để lưu trữ. Với cấu hình này, mạng hoạt động tương đối tốt và cho kết quả phân loại chính xác hơn.

Như đã trình bày ở trên, trong tất cả các trường hợp mô hình mạng được sử dụng, các trọng số đã được huấn luyện trước bằng tập dữ liệu ImageNet được sử dụng cho mô hình mạng ban đầu, trước khi các lớp điều chỉnh được thêm vào và quá trình học chuyển đổi được thực hiện. Trong quá trình huấn luyện của học chuyển đổi, trọng số của các lớp cuối cùng trong mô hình gốc được cho phép thay đổi để thực hiện việc tinh chỉnh, trong khi trọng số của các lớp trước đó được giữ cố định (với các giá trị từ các trọng số đã được huấn luyện trên tập dữ liệu ImageNet).

Bên cạnh đó, do các mô hình mạng được sử dụng đều được huấn luyện trên tập ImageNet là tập ảnh màu gồm 3 kênh màu, và có ngõ vào phù hợp với ảnh màu, nên các ảnh xám của tập dữ liệu ảnh X-quang vù cũng được chuyển thành ảnh 3 kênh màu tương ứng (được đề cập trong chương 3) trước khi đưa vào huấn luyện.

### c) Kết quả phân loại các loại tổn dùng tập ảnh MIAS

Hình 5.10 hiển thị các đường cong huấn luyện cho mô hình phân loại, bao gồm đường cong tổn thất và đường cong độ chính xác. Rõ ràng, đường cong huấn luyện hội tụ suôn sẻ từ giai đoạn đầu tiên đến giai đoạn cuối cùng mà không xảy ra va chạm. Hơn nữa, khoảng cách giữa đường cong độ chính xác thực (màu đỏ) và đường cong huấn luyện (màu xanh) cho thấy chúng gần như không bị trang bị quá mức trong quá trình huấn luyện. Tương tự như đường cong huấn luyện, đường cong tổn thất thể hiện trong Hình 5.10 hội tụ đều đặn. Với những kết quả của đường cong huấn luyện và mất mát mà không khớp quá mức hoặc không phù hợp, chúng ta có thể thấy rằng mô hình đề xuất để phân loại bốn tổn thương vù là hiệu quả. Ngoài ra, đường cong Accuracy rate over epochs cho thấy độ chính xác tăng lên khi số lần lặp tăng lên, trong đó đường màu xanh dành cho tập huấn luyện, đường màu đỏ dành cho tập kiểm tra.

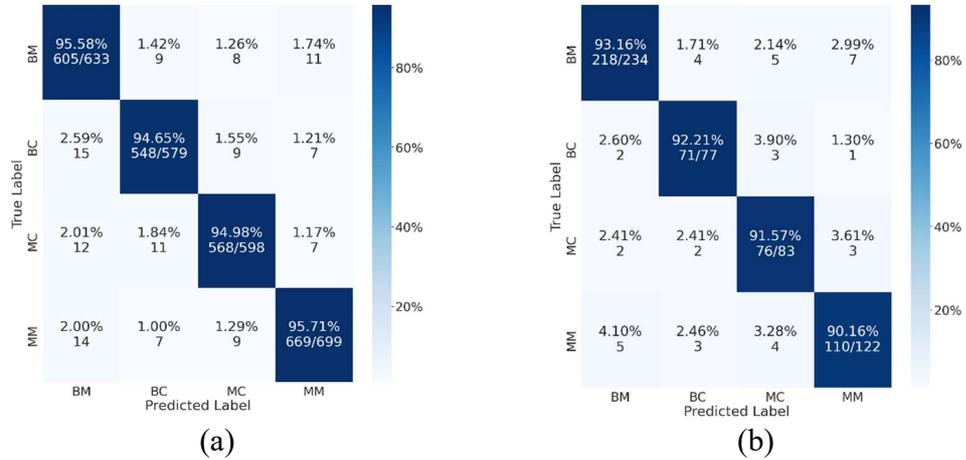
## Chương 5 PHÂN LOẠI ẢNH UTV SỬ DỤNG MẠNG HỌC SÂU VÀ HỆ THỐNG HỖ TRỢ CHẨN ĐOÁN UTV TỪ XA SỬ DỤNG MẠNG HỌC SÂU



**Hình 5.10. Mô tả quy trình đào tạo với hơn 350 epoch cho Adam với tốc độ học tập 0,0001: (a) đường cong độ chính xác để đào tạo và xác nhận; (b) đường cong tổn thất cho việc đào tạo và xác nhận**

Hình 5.10a và Hình 5.10b thể hiện kết quả sử dụng ma trận nhầm lẫn cho mô hình VGG16 với độ chính xác phân loại cao nhất. Cụ thể, các giá trị hàng biểu thị nhãn thực tế của các bộ ảnh, BC, BM, MC, MM, các giá trị cột mô tả nhãn dự đoán của bốn loại bộ ảnh này và các đường chéo biểu thị các giá trị True Positive (TP). Ngoài ra, ma trận nhầm lẫn cho thấy rằng không có sai lệch nào tồn tại đối với bất kỳ lớp nào và chúng cũng dự đoán tất cả các lớp tương tự. Cụ thể, với bộ dữ liệu CBIS-DDSM, mô hình có thể phân loại tổn thương MM và BM với độ chính xác cao nhất lần lượt là 95,71% và 95,58%. Trong khi đó, với bộ dữ liệu MIAS, tổn thương BM và BC cho độ chính xác phân loại cao nhất lần lượt là 93,16% và 92,21%. Độ chính xác trung bình trên bộ dữ liệu CBIS-DDSM là 95,23%, cao hơn độ chính xác đạt được khi phân loại trên bộ dữ liệu MIAS là 91,78%. Đáng chú ý trong Hình 8, tỷ lệ phân loại sai từ BM đến MM đều cao đối với cả hai tập dữ liệu thử nghiệm, lần lượt là 1,74% và 2,99%. Ngoài ra, MM cũng bị phân loại sai thành BM với tỷ lệ 2,00% và 4,10% trên bộ dữ liệu CBIS-DDSM và MIAS. Những kết quả phân loại này chứng tỏ tính hiệu quả và tiềm năng của mô hình đề xuất.

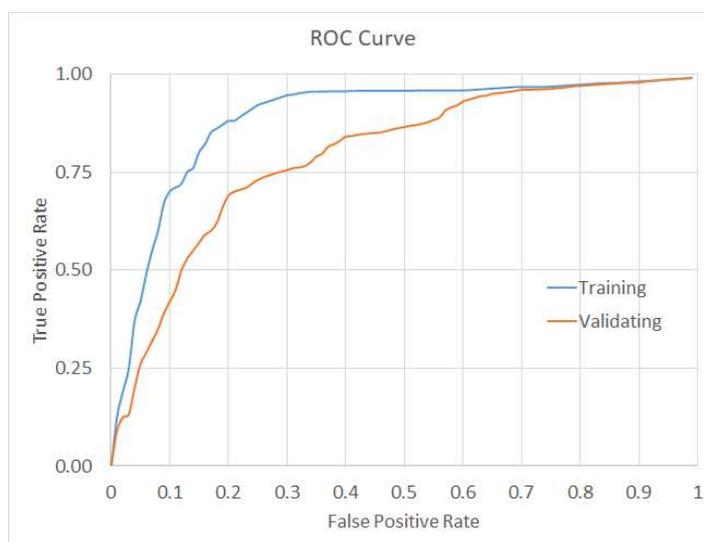
Chương 5 PHÂN LOẠI ẢNH UTV SỬ DỤNG MẠNG HỌC SÂU VÀ HỆ THỐNG  
HỖ TRỢ CHẨN ĐOÁN UTV TỪ XA SỬ DỤNG MẠNG HỌC SÂU



**Hình 5.11. Mô tả hiệu quả phân loại của hai loại tập dữ liệu với Vô hình lạnh tính (BC), Khối lạnh tính (BM), Vô hình ác tính (MC), Khối lượng ác tính (MM). (a) Bộ dữ liệu CBIS-DDSM; (b) Bộ dữ liệu MIAS.**

Trong Hình 5.11, các đường cong độ chính xác để kiểm tra và xác nhận cho thấy mô hình đạt được độ chính xác cao với tỷ lệ dương tính giả thấp đối với các tập dữ liệu. Trong đó, đường cong màu vàng là kết quả của tập huấn luyện và đường cong màu xanh là kết quả của tập kiểm tra. Do đó, phương pháp xử lý vận hành nhận được (ROC) và diện tích dưới đường cong (AUC) được lấy từ đường cong ROC cho thấy mô hình đề xuất cùng với phương pháp đề xuất có hiệu quả để phân biệt các loại khác nhau. Đặc biệt, giá trị AUC gần bằng 1 cho thấy mô hình này có thể phát hiện hầu hết các lớp, trong đó đường cong ROC gần như chạm vào đỉnh của trục y, có nghĩa là tỷ lệ dương tính giả gần bằng 0. Trong khi tỷ lệ dương tính thực là gần đến 1, điều này chứng tỏ hiệu quả của mô hình là rất cao.

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



**Hình 5.12. Đường cong tỷ lệ chính xác của huấn luyện và kiểm tra sử dụng mô hình trên**

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 dữ liệu 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.

**Bảng 5.8. So sánh giữa các bộ ảnh khác nhau (có/không XLA)**

Phương pháp	Độ chính xác TB
Không có tiền XLA	82.1%
Phân đoạn với Otsu	89.2%
Phân đoạn Otsu với xác định vùng ROI lớn nhất	95%

Bảng 5.9 cho thấy mô hình đề xuất đạt được độ chính xác cao nhất so với các công trình trước đó sử dụng một số phương pháp, bộ dữ liệu hoặc mô hình tương tự.

## 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

Hơn một nửa số nghiên cứu trước đây trong Bảng 5 đã sử dụng VGG16 và các bộ dữ liệu tương tự so với nghiên cứu trong bài viết này. Trong [167], phương pháp học chuyển giao với mô hình MVGG được đề xuất để phân loại ảnh dựa trên mô hình VGG16 và kết quả là 94,3% với 15 kỹ nguyên sử dụng CSDL DDSM.

Trong một nghiên cứu của Shallu và Mehra [168] sử dụng VGG16 đã được huấn luyện trước và tinh chỉnh, hiệu suất của nó đạt độ chính xác cao nhất là 92,60%. Với VGG16 và VGG19 sử dụng cùng một tập dữ liệu, Hameed [169] đã có kết quả với độ chính xác thu được là 94,71%. Tương tự, trong nghiên cứu [156] đã áp dụng một số mô hình được huấn luyện trước và độ chính xác cao nhất là 92,78% với VGGNet. Mạng VGG16 được tinh chỉnh của chúng tôi với lớp FC đã đạt được độ chính xác 95% trong phân loại tổn thương vú bằng cách sử dụng bộ dữ liệu CBIS-DDSM sau khi tối ưu hóa Adam. Nhìn chung, với số liệu thống kê trên, các mô hình sử dụng VGG16 đều đạt được kết quả phân loại với độ chính xác rất cao.

Với các nghiên cứu sử dụng các mạng khác, Nasir Khan [170] đã phát triển hệ thống MVFF CADx và bộ ảnh CBIS-DDSM và MIAS để phân loại. Với mô hình CADx này, kết quả phân loại đạt độ chính xác tốt nhất là 93,73%. Một nghiên cứu là Al-antari [52] đã phát triển một hệ thống dựa trên CAD với INbreast hướng tới CSDL chụp nhũ ảnh kỹ thuật số toàn trường và độ chính xác phân loại đạt được 95,64%. Trong khi Zahra Jafari [23] sử dụng nhiều mạng CNN khác nhau để trích xuất đặc trưng bằng các bộ phân loại như Random Forest, SVM, KNN, NNN, độ chính xác cao nhất là 96%. Tuy nhiên, hiệu suất của các mô hình này có thể bị ảnh hưởng bởi độ phân giải điểm ảnh thấp, mô xung quanh, kích thước nốt nhỏ hoặc vị trí khó chụp ảnh vú.

**Bảng 5.9. So sánh với các kết quả nghiên cứu khác**

Nghiên cứu	Mạng CNN	Tập dữ liệu	Accuracy
Khamparia et al. (2021) [167]	Hybrid MVGG16 ImageNet	DDSM Gồm 2620 ảnh.	94.3%
Shallu and Mehra,	VGG16 + LR VGG19 + LR	Dữ liệu từ BV Shanxi Medical University với 2042 ảnh số.	92.60% 90.40%

## 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

(2018) [168]	ResNet50 + LR		79.40%
	Full trained		
Hameed et al. (2020) [169]	VGG16 + VGG19	Dữ liệu từ Đại học Colsanitas Colombia với 544 ảnh UTV.	93.53%
	Fine-tuned VGG16 + VGG19		95.29%
	AlexNet		92.70%
	VGGNet		92.78%
Li et al. (2019) [156]	GoogLeNet	Dữ liệu từ BV Shanxi Medical University với 2042 ảnh số.	93.54%
	DenseNet		93.87%
	DenseNet-II		94.55%
Nasir Khan et al. (2019) [170]	MVFF CADx	CBIS-DDSM: 3568 ảnh và MIAS: 322 ảnh	93.73%
Al-antari et al. (2018) [52]	Fully Integrated CAD	Tập dữ liệu INbreast Toward với 410 ảnh	95.64%
Zahra Jafari and Ebrahim Karami (2023) [23]	Concat. Model with the NN classifier	Tập dữ liệu RSNA với 54,713 ảnh	96%
Nghiên cứu này	Fine-tuned VGG16	CBIS-DDSM, MIAS	94%
			95%

Rõ ràng hai bộ ảnh (CBIS-DDSM và MIAS) rất phổ biến cho các nghiên cứu liên quan đến tổn thương vú. Mặc dù các thuật toán phân loại khác nhau nhưng độ chính xác rất cao, khoảng 92%. Trong nghiên cứu của chúng tôi sử dụng hai bộ ảnh với 4 loại tổn thương ở vú, trong đó có bộ ảnh tăng cường. Ngoài ra, các bộ ảnh này được xử lý để trích xuất ROI lớn nhất, dùng để áp dụng cho VGG16 tinh chỉnh và đạt kết quả có độ chính xác từ 94% đến 95%. Điều này có nghĩa là tính hiệu quả của thuật toán đề xuất được chứng minh.

### 5.2.3 Bài toán phân loại các loại tổn thương dựa trên tập ảnh X-quang vú VinDr-Mammo

#### a) Tập dữ liệu

Tập dữ liệu ảnh UTV được sử dụng trong nghiên cứu này là VinDr-Mammo [17], bao gồm các ảnh UTV ở định dạng DICOM, được mô tả trong Bảng 5.10. Ảnh DICOM không chỉ cung cấp hình ảnh y tế chất lượng cao mà còn lưu trữ các thông tin quan trọng của bệnh nhân và các tham số kỹ thuật của thiết bị chụp, chẳng hạn

## 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

như tên bệnh nhân, tuổi, ngày chụp, loại máy chụp và các thông số kỹ thuật cụ thể. Hơn nữa, mỗi liên kết trực tiếp giữa hình ảnh y tế và thông tin bệnh nhân giúp nâng cao hiệu quả và độ chính xác trong chẩn đoán cũng như điều trị. Mỗi bệnh nhân được chụp ở nhiều tư thế khác nhau: L-CC, R-CC, L-MLO, và R-MLO.

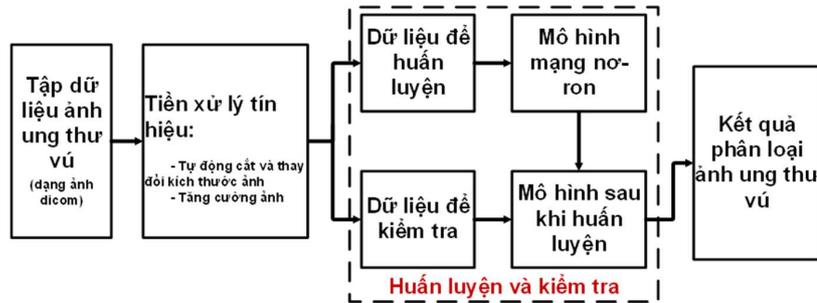
**Bảng 5.10. Tập dữ liệu VinDr-Mammo với bốn loại bệnh tương ứng**

Số thứ tự (Ngõ ra)	Nhãn	Theo BIRADS	Tổng
1	Khối u (Mass)	3, 4 and 5	1113
2	Vôi hóa (Calcification)	3, 4 and 5	212
3	Thay đổi cấu trúc (Architectural and Asymmetry)	3, 4 and 5	390
4	Bình thường (Normal)	1 and 2	18232

### b) Đề xuất hệ thống phân loại 4 loại tổn thương từ tập dữ liệu VinDr-Mammo

Hệ thống đề xuất cho phân loại UTV, như minh họa trong Hình 5.13, bao gồm các khối chức năng sau: dữ liệu ảnh UTV, tiền XLA, mạng học sâu cho phân loại và kết quả phân loại. Cụ thể, các ảnh UTV là tập dữ liệu DICOM của các ca UTV, sẽ được tiền xử lý tự động bằng cách cắt và thay đổi kích thước ảnh DICOM để thu được ảnh đầu vào mong muốn. Ngoài ra, kỹ thuật tăng cường dữ liệu được áp dụng nhằm tăng số lượng ảnh trong bốn loại dữ liệu khác nhau, qua đó cân bằng số lượng ảnh giữa các loại này. Các khối huấn luyện và kiểm thử được thiết kế để thực hiện huấn luyện và phân loại các loại UTV khác nhau. Cuối cùng, khối kết quả phân loại cung cấp đầu ra phân loại từ hệ thống đề xuất.

## 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



**Hình 5.13. Đề xuất hệ thống phân loại các loại tổn thương vú dựa vào tập ảnh X-quang vú VinDr-Mammo**

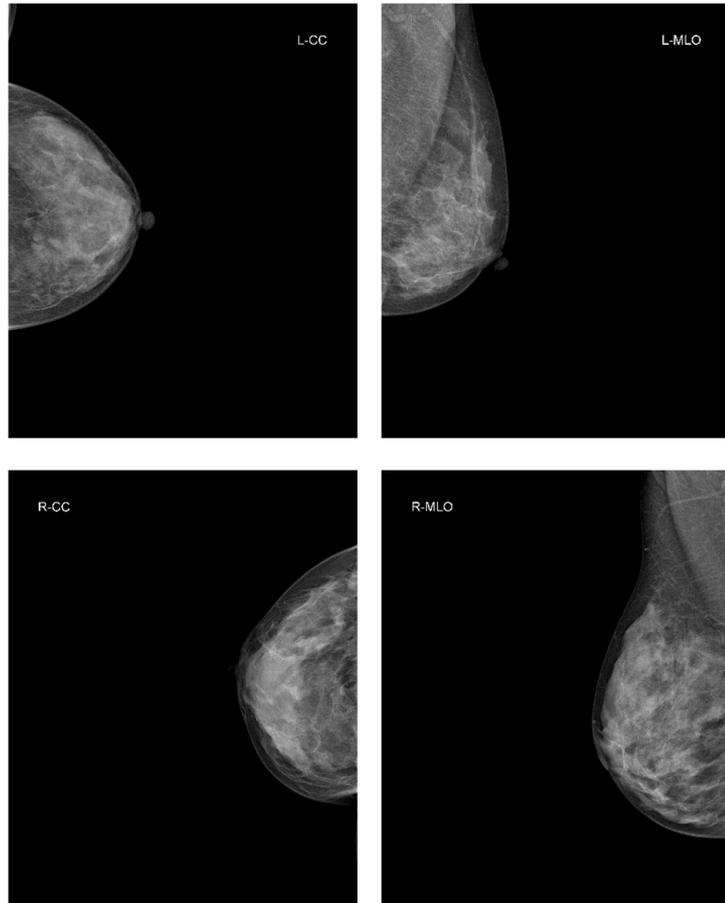
Hệ thống phân loại UTV được đề xuất, như minh họa trong Hình 3.1, được cấu trúc thành bốn khối chức năng chính: (i) khối dữ liệu hình ảnh, (ii) khối tiền xử lý, (iii) khối huấn luyện và kiểm tra, và (iv) khối kết quả phân loại. Cụ thể, khối dữ liệu bao gồm 20.247 ảnh DICOM thu thập từ các ca bệnh UTV, được sử dụng trong quá trình huấn luyện và cả đánh giá mô hình. Trong khối tiền xử lý, hình ảnh được xử lý tự động thông qua phân ngưỡng Otsu kết hợp với kỹ thuật tìm biên Contour nhằm trích xuất chính xác vùng nhũ ảnh. Sau đó, ảnh được chuẩn hóa kích thước về  $512 \times 512$  pixel mà vẫn giữ nguyên tỷ lệ gốc. Đồng thời, các phương pháp tăng cường dữ liệu (data augmentation) được áp dụng để mở rộng tập dữ liệu và cân bằng số lượng mẫu giữa các nhóm. Khối huấn luyện và kiểm tra được thiết kế dựa trên kiến trúc ResNet-50, trong đó mô hình được huấn luyện trên tập dữ liệu huấn luyện và kiểm định bằng tập dữ liệu kiểm tra nhằm đánh giá khả năng phân loại. Cuối cùng, khối kết quả phân loại cung cấp đầu ra của hệ thống, thể hiện khả năng nhận dạng và phân loại các loại UTV từ mô hình đã được huấn luyện.

### c) Kết quả phân loại tổn thương dùng tập ảnh VinDr-Mammo

Mỗi ca bệnh được chụp X-quang tuyến vú ở bốn tư thế khác nhau, bao gồm: L-CC (trái – thẳng trên dưới), L-MLO (trái – chệch trong ngoài), R-CC (phải – thẳng trên dưới) và R-MLO (phải – chệch trong ngoài). Hình 5.14 minh họa một trường hợp bệnh nhân UTV với đầy đủ bốn góc chụp: (a) L-CC, (b) L-MLO, (c) R-CC và (d) R-MLO.

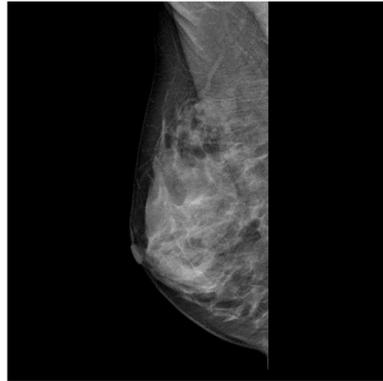
Chương 5 PHÂN LOẠI ẢNH UTV SỬ DỤNG MẠNG HỌC SÂU VÀ HỆ THỐNG  
HỖ TRỢ CHẨN ĐOÁN UTV TỪ XA SỬ DỤNG MẠNG HỌC SÂU

---



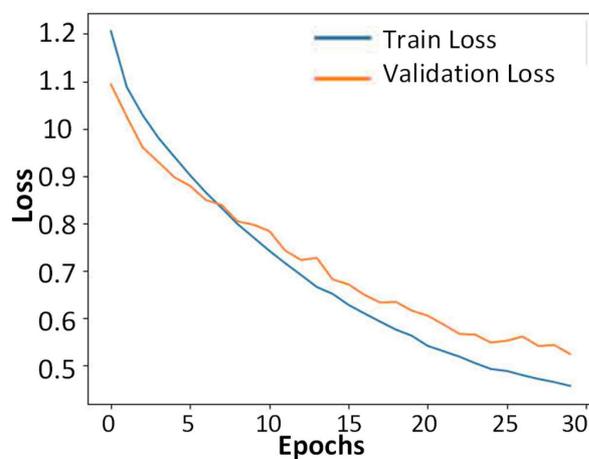
**Hình 5.14. Minh họa ảnh X-quang vú của bệnh nhân ở hai tư thế khác nhau và được chụp theo 2 hướng.**

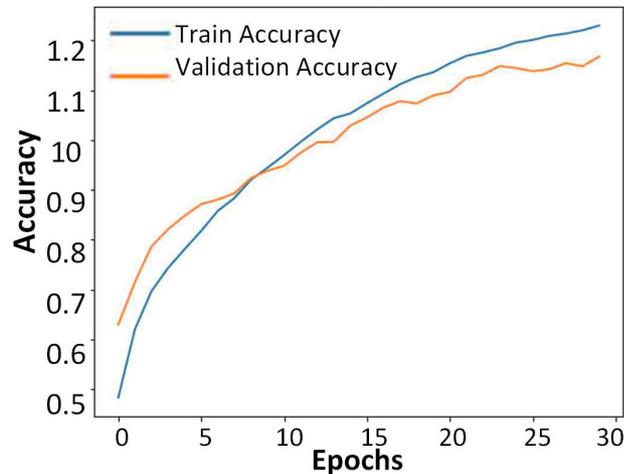
Tập dữ liệu kể trên gồm các hình ảnh X-quang vú với kích thước khác nhau do được thu nhận từ nhiều thiết bị chụp. Nếu chỉ áp dụng phương pháp cắt và thay đổi kích thước thông thường, một số ảnh có thể xuất hiện viền trắng sau khi chuẩn hóa, làm giảm độ chính xác trong việc thể hiện vùng vú. Phương pháp cắt đề xuất khắc phục hạn chế này bằng cách tự động nhận diện và trích xuất chính xác vùng vú, loại bỏ các phần nền không cần thiết. Sau đó, ảnh được đưa vào khung nền đen có kích thước chuẩn  $512 \times 512$  pixel và căn chỉnh ở vị trí trung tâm để đồng bộ kích thước giữa các mẫu. Kết quả xử lý được minh họa trong Hình 5.15.



**Hình 5.15. Kết quả ảnh X-quang vú sau XLA và đưa về kích thước 512x512.**

Trong bộ dữ liệu đã trình bày, số lượng ảnh giữa các lớp có sự chênh lệch lớn, điển hình là lớp 2 chỉ có 212 ảnh trong khi lớp 4 lên tới 18.232 ảnh. Sự mất cân đối này ảnh hưởng tới hiệu suất phân loại, do đó kỹ thuật tăng cường dữ liệu cần được áp dụng. Quá trình tăng cường được thực hiện thông qua các phép biến đổi hình học như lật ngang, lật dọc và xoay trong khoảng  $\pm 50^\circ$ . Đối với lớp 4, một tập ảnh được chọn ngẫu nhiên để điều chỉnh số lượng sao cho cân bằng với các lớp còn lại. Nghiên cứu tiến hành bốn kịch bản thí nghiệm: (i) mỗi lớp gồm 2.000 ảnh, (ii) mỗi lớp gồm 5.000 ảnh, (iii) mỗi lớp gồm 10.000 ảnh, và (iv) mỗi lớp 10.000 ảnh, riêng lớp 4 giữ nguyên 18.232 ảnh. Toàn bộ tập dữ liệu sau xử lý được chia theo tỷ lệ 80% cho huấn luyện, 10% cho xác thực và 10% cho kiểm tra mô hình.

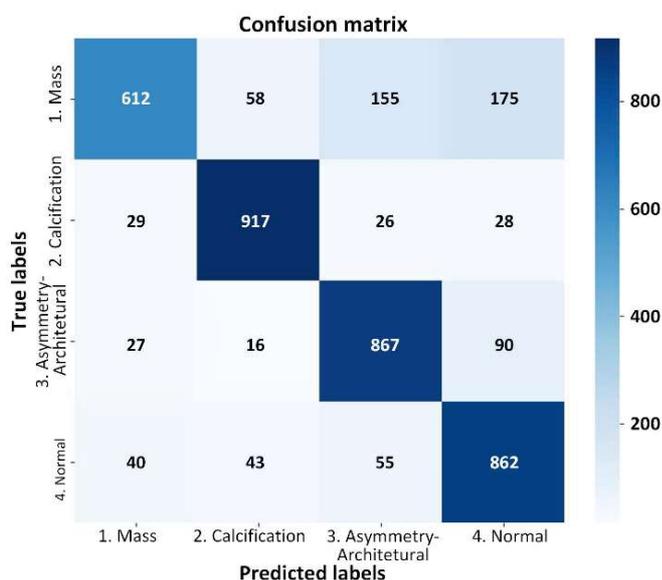




**Hình 5.16. Kết quả trình bày đồ thị hàm mất mát và độ chính xác của hệ thống mạng học sâu đã được đề xuất.**

Trong nghiên cứu này, mô hình ResNet50 được triển khai để phân loại ảnh UTV. Tất cả các ảnh được chuẩn hóa về kích thước  $512 \times 512$  pixel và huấn luyện trong 30 vòng lặp (epochs). Kết quả huấn luyện với mỗi lớp gồm 10.000 ảnh, được trình bày trong Hình 5.16. Kết quả cho thấy giá trị hàm mất mát giảm dần theo số vòng lặp, đồng thời độ chính xác được cải thiện rõ rệt khi số epoch tăng. Bên cạnh đó, Hình 5.17 thể hiện ma trận nhầm lẫn, cho thấy khả năng phân loại ảnh thuộc nhóm vôi hóa đạt hiệu suất cao, trong khi nhóm khối u có độ chính xác thấp hơn.

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



**Hình 5.17. Ma trận nhầm lẫn với số lượng ảnh trong mỗi lớp được tăng cường lên 10.000.**

Bảng 5.11 trình bày kết quả so sánh hiệu suất phân loại ở bốn kịch bản thí nghiệm. Trong Trường hợp 1, khi mỗi lớp được tăng cường lên 2.000 ảnh, hiệu suất phân loại đạt mức thấp nhất với các chỉ số: Độ chính xác (ACC) 72,92%, Độ chính xác riêng (PRE) 75,64%, Độ nhạy (REC) 72,92% và F1-score (F1S) 70,66%. Khi số lượng ảnh tăng lên 5.000 (Trường hợp 2) và 10.000 (Trường hợp 3), các chỉ số đều được cải thiện, trong đó ACC lần lượt đạt 77,83% và 81,45%. Tuy nhiên, ở Trường hợp 4, khi ba lớp còn lại được tăng cường lên 10.000 ảnh nhưng lớp Bình thường (Normal) vẫn giữ nguyên 18.232 ảnh, hiệu suất phân loại có xu hướng giảm do sự mất cân bằng của dữ liệu. Cụ thể, kết quả đạt ACC = 81,65%, PRE = 84,14%, REC = 78,41% và F1S = 79,95%. Điều này cho thấy sự không đồng đều về số lượng mẫu giữa các lớp dẫn đến chênh lệch trong khả năng phân loại giữa các nhóm.

**Bảng 5.11. Tập dữ liệu VinDr-Mammo với bốn loại bệnh tương ứng**

	Trường hợp 1	Trường hợp 2	Trường hợp 3	Trường hợp 4
ACC	72.92	77.83	81.45	81.65

## 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

PRE	75.64	77.80	82.09	84.14
REC	72.92	77.84	81.45	78.41
F1S	70.66	77.82	81.07	79.95

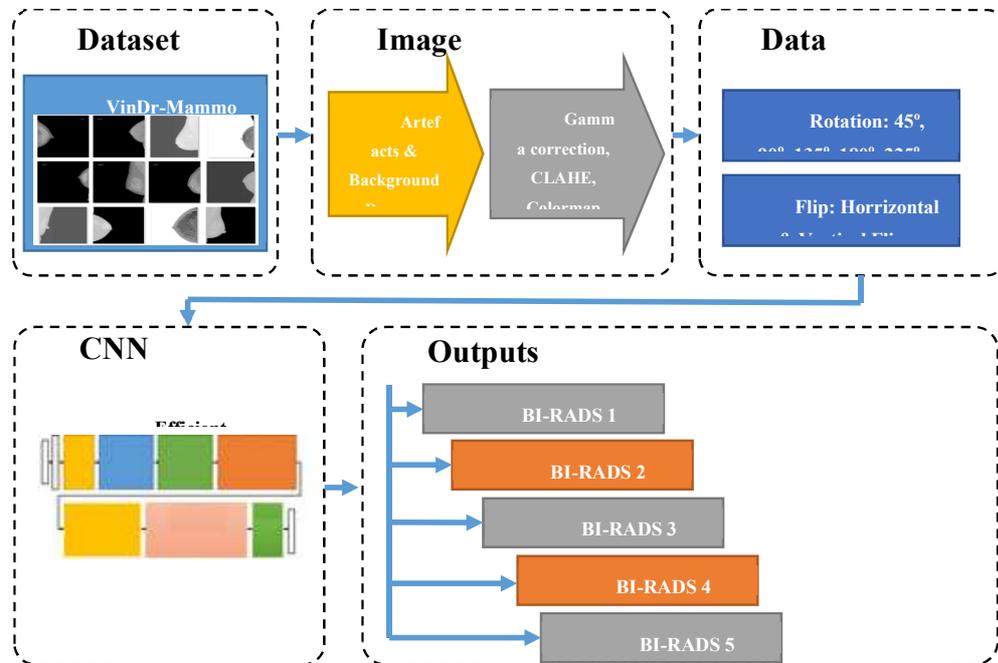
Trong nghiên cứu này, phương pháp đề xuất kết hợp bước tiền XLA DICOM với mô hình học sâu ResNet nhằm phân loại các dạng phụ của UTV trên bộ dữ liệu VinDr-Mammo. Kết quả thực nghiệm cho thấy việc tăng số lượng ảnh mỗi lớp lên 10.000 giúp nâng cao hiệu suất phân loại. Tuy nhiên, khi lớp Bình thường vẫn duy trì ở mức 18.232 ảnh, sự mất cân bằng dữ liệu khiến hiệu quả phân loại giữa các lớp chưa đạt mức tối ưu. Vì vậy, các nghiên cứu tiếp theo hướng tới việc mở rộng quy mô dữ liệu và cải tiến các mô hình học sâu tiên tiến hơn nhằm tiếp tục cải thiện hiệu quả phân loại ảnh UTV.

Bộ dữ liệu VinDr-Mammo mới được thu thập và công bố tại Việt Nam nên hiện còn ít nghiên cứu khai thác. Do đó, nghiên cứu này có thể được xem như một nền tảng ban đầu, góp phần thúc đẩy các công trình tiếp theo trong việc phát triển thuật toán và nâng cao hiệu quả phân loại UTV.

### 5.2.4 Kết quả phân loại bệnh UTV theo BIRADS

Mục đích của nghiên cứu này là phân loại ảnh chụp nhũ ảnh theo mức độ tổn thương trên thang BIRADS bằng một trong những mô hình CNN tiên tiến nhất, cụ thể là mô hình EfficientNet được kết hợp với một số phương pháp XLA phù hợp. Phương pháp đề xuất này là tìm ra một mô hình xử lý hiệu quả nhất trong việc phân loại mức độ tổn thương dựa trên ảnh chụp nhũ ảnh như Hình 5.18. Đây là nhiệm vụ quan trọng góp phần chẩn đoán chính xác UTV sớm hơn để điều trị. Ngoài ra, nghiên cứu được tiến hành trên một bộ ảnh chụp nhũ ảnh tương đối mới, được gọi là bộ ảnh VinDr-Mammo [24], được thu thập từ các BV tại Việt Nam.

Chương 5 PHÂN LOẠI ẢNH UT V SỬ DỤNG MẠNG HỌC SÂU VÀ HỆ THỐNG HỖ TRỢ CHẨN ĐOÁN UT V TỪ XA SỬ DỤNG MẠNG HỌC SÂU



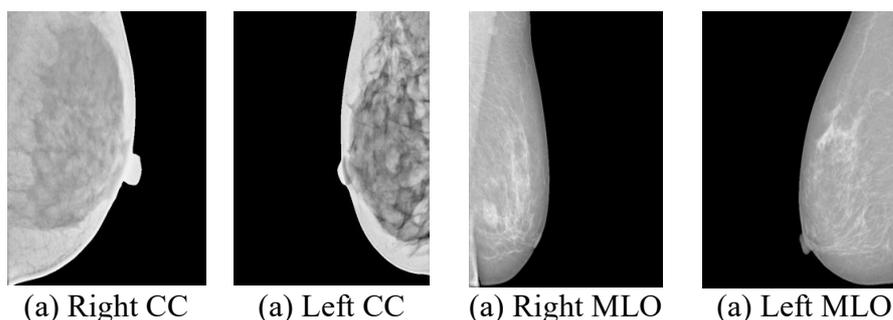
**Hình 5.18. Sơ đồ khối quá trình xử lý của hệ thống phân loại ảnh X-quang vú theo mức độ tổn thương theo thang BIRADS**

Để thực hiện được việc phân loại các ảnh trong tập dữ liệu theo mức độ tổn thương, trước khi đưa vào mô hình CNN, các ảnh trong tập dữ liệu đầu vào sẽ được tiền xử lý qua hai bước: loại bỏ các thành phần không cần thiết trên ảnh gốc (ví dụ ngày tháng chụp, thông tin ảnh, ...) và tăng cường hình ảnh. Sau đó tập dữ liệu sẽ được tăng cường bằng cách áp dụng 2 phương pháp là hình học (xoay và lật ảnh), và biến thiên độ sáng ảnh. Bước này sẽ giúp các tập ảnh gia tăng đến số lượng cần thiết cho việc huấn luyện và tạo ra sự cân bằng giữa số lượng ảnh của mỗi loại tổn thương, nhằm tránh việc mất cân bằng trong quá trình huấn luyện. Sau đó, các tập dữ liệu sẽ được đưa vào mạng học sâu EfficientNet với các trọng số ban đầu được thiết lập từ kết quả huấn luyện sẵn với ImageNet. Các mô hình mạng học sâu này được tinh chỉnh các tầng cuối và thêm vào các lớp ngõ ra phù hợp. Kết quả phân loại ảnh sẽ được so sánh để tìm ra được mô hình có hiệu quả tốt nhất để giải bài toán phân loại ảnh X-quang vú theo mức độ tổn thương theo thang BIRADS [171].

**a) Tập dữ liệu**

## 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

Tập dữ liệu VinDr-Mammo [135] được thu thập từ các kết quả kiểm tra chụp X-quang vú được thực hiện từ năm 2018 đến năm 2020 được lưu trữ trong Hệ thống Lưu trữ và Truyền thông Hình ảnh (PACS) của BV Đại học Y Hà Nội (HMHU) và BV 108 (H108), 5.000 kết quả kiểm tra chụp X-quang tuyến vú, tương ứng với 20.000 hình ảnh (mỗi kết quả khám gồm 4 ảnh: Right CC, Left CC, Right MLO, Left MLO), được lấy mẫu ngẫu nhiên, sau đó được xác nhận lại. Các thông số hình ảnh được sử dụng để xử lý, tuổi của bệnh nhân và kiểu thiết bị hình ảnh được giữ lại, trong khi tất cả thông tin bệnh nhân khác bị xóa để bảo vệ quyền riêng tư của bệnh nhân [135]. Toàn bộ 20.000 ảnh X-quang vú của tập dữ liệu VinDr-Mammo được phân thành 5 loại theo thang đo BIRADS. Với độ tuổi của các bệnh nhân nằm trong khoảng từ 20 đến 86 tuổi, trong đó tập trung nhiều nhất ở độ tuổi từ 40-45 tuổi.



**Hình 5.19. Bộ 4 ảnh trong một kết quả khám X-quang vú của tập ảnh VinDr-Mammo**

Số lượng ảnh tương ứng với các mức độ tổn thương theo thang đo BIRADS của tập dữ liệu VinDr-Mammo không đồng đều, thể hiện ở số lượng ảnh lớn nhất tập trung ở mức BIRADS 1 và giảm dần ở các mức BIRADS cao hơn, điều này dẫn tới sự mất cân bằng đáng kể trong số lượng của mỗi loại ảnh cho bài toán phân loại. Các đặc điểm tổn thương cụ thể của mỗi trường hợp bệnh nhân cũng được xác định, số lượng các loại đặc điểm tổn thương cụ thể được liệt kê trong Bảng 5.12.

**Bảng 5.12. Thống kê đánh giá tổn thương theo thang BIRADS của tập dữ liệu gốc**

	BIRADS 1	BIRADS 2	BIRADS 3	BIRADS 4	BIRADS 5	TỔNG
Training	5.362	1.871	372	305	90	8.000

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

	(67,03%)	(23,39%)	(04,65%)	(03,81%)	(01,12%)	
Testing	1.341	467	93	76	23	<b>2.000</b>
	(67,05%)	(23,35%)	(04,65%)	(03,80%)	(01,15%)	
<b>Tổng</b>	<b>6.703</b>	<b>2.338</b>	<b>465</b>	<b>381</b>	<b>113</b>	<b>10.000</b>
	(67,03%)	(23,38%)	(04,65%)	(03,81%)	(01,13%)	

**Ghi chú:** số lượng trong bảng trên là số lượng ca bệnh (tương ứng với 1 bên vú, tổng là 10.000) mỗi bên vú sẽ được chụp 2 ảnh là MLO và CC, do đó tổng số ảnh là 20.000, và số bệnh nhân là 5.000.

Do số lượng ảnh mỗi loại không đều nhau (có thể ảnh hưởng đến việc huấn luyện) nên cần tiến hành bỏ bớt đối với nhóm BIRADS-1 (nhiều nhất) và tăng cường dữ liệu các nhóm còn lại để có sự cân bằng.

Phương pháp tăng cường dữ liệu:

- BIRADS 1, BIRADS 2: chọn ngẫu nhiên 3.000 ảnh số ảnh ban đầu
- BIRADS 3,4,5: tăng cường ảnh với các phương pháp lật, xoay, dịch trái, phải để có 3000 ảnh.

**Bảng 5.13. Số lượng ảnh sau khi tăng cường của tập VinDr-Mammo**

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

**Lật ảnh:**

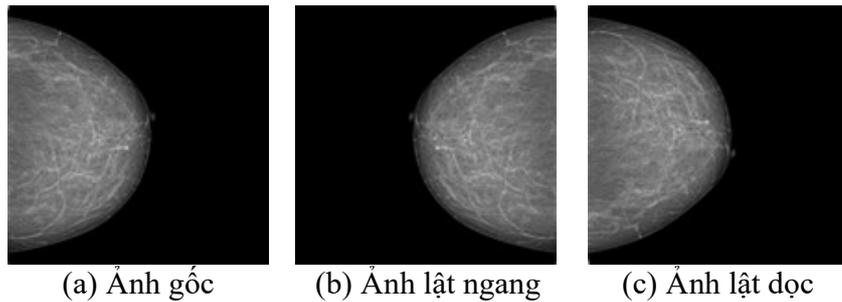
Các hình ảnh có thể được lật theo chiều ngang hoặc chiều dọc để tạo ra các ảnh mirror của ảnh ban đầu. Với  $p(x, y)$  là ảnh ban đầu, ở đây  $x$  và  $y$  biểu diễn cho tọa độ của pixel, công thức lật ảnh được cho bên dưới.

Cho  $x \in [0, w]$ ;  $y \in [0, h]$ :

$$\text{Phép lật ngang: } p_{out}(x, y) = p_{in}(w - x, y) \quad (5.6)$$

$$\text{Phép lật dọc: } p_{out}(x, y) = p_{in}(x, h - y) \quad (5.7)$$

Ở đây:  $p_{in}(x, y)$  là điểm ảnh tại tọa độ  $(x, y)$  từ ảnh đầu vào  
 $p_{out}(x, y)$  là điểm ảnh của ảnh đầu ra  
 $w, h$  là chiều rộng và chiều cao của ảnh tính theo pixel



**Hình 5.20. Ví dụ lật ảnh**

#### **Xoay ảnh:**

Việc xoay hình ảnh có thể được thực hiện với một góc xoay bất kỳ, trong nghiên cứu này, các góc xoay  $45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ$  được sử dụng để xoay ảnh ban đầu quanh tâm của hình (center point) nhằm tạo ra nhiều dữ liệu hơn cho việc huấn luyện. Mỗi điểm ảnh  $p(x, y)$  ban đầu được chuyển sang vị trí mới  $p(x_r, y_r)$ , trong đó  $x_r, y_r$  là tọa độ mới được tính theo công thức:

$$x_r = r \cos(\alpha + \theta) \quad (5.8)$$

$$y_r = r \sin(\alpha + \theta) \quad (5.9)$$

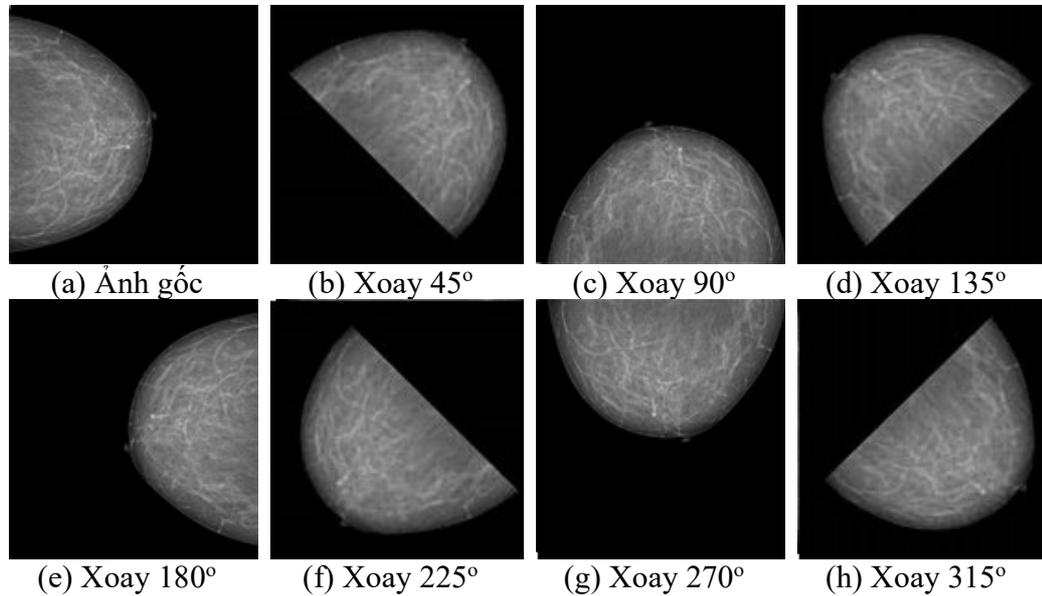
ở đây:  $r = \sqrt{x^2 + y^2}$ ,  $x, y$  là tọa độ của điểm ảnh gốc

$$\alpha = \arctan\left(\frac{y}{x}\right) \quad (5.10)$$

$\theta$  là góc lật ( $\theta = 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ$ )

Kết quả của việc xoay ảnh được minh họa ở hình dưới đây:

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

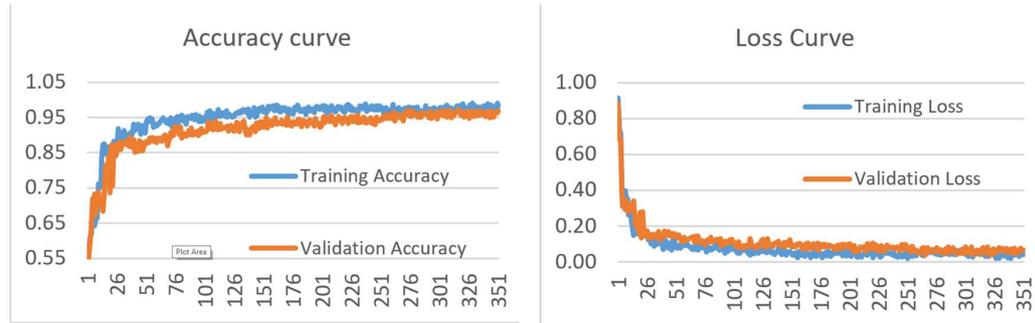


**Hình 5.21. Ví dụ xoay ảnh**

Trong nghiên cứu này, việc tinh chỉnh được thực hiện một cách thích hợp nhằm cải thiện hiệu suất trong kết quả phân loại của EfficientNet trên các tập dữ liệu ảnh UTV đã đề cập ở trên. Việc tinh chỉnh mô hình bao gồm huấn luyện lại 2 lớp cuối cùng của EfficientNet và thêm vào các lớp đầu ra, bao gồm 1 lớp Flatten, tiếp theo là 1 lớp Dense với 1024 node kết hợp với hàm activation là relu, một lớp Dropout với hệ số 0.5 và cuối cùng là lớp Dense với 4 node và hàm activation softmax để phù hợp với số lượng lớp ngõ ra. Các lớp này sẽ được kết hợp với mô hình mạng EfficientNet đã đề cập ở trên và huấn luyện cùng với các tập dataset đã được chuẩn bị cho nghiên cứu này. Kết quả của bước tiền XLA ở trên sẽ được đưa vào mạng học sâu EfficientNet nhằm tìm hiểu xem hiệu quả của các bước xử lý này đối với kết quả ở đầu ra.

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

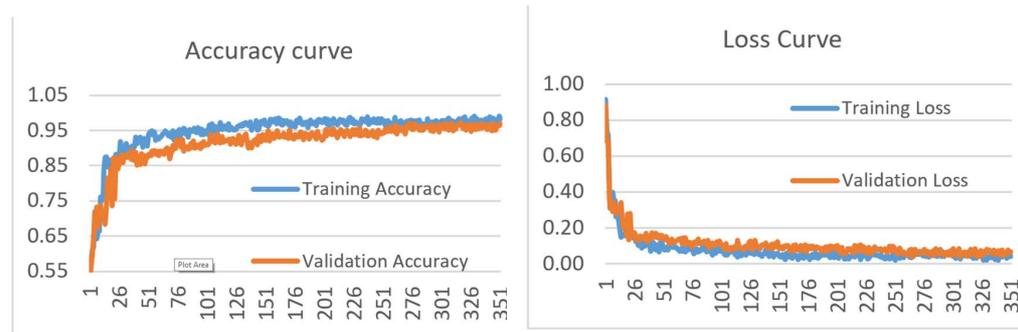
Chương 5 PHÂN LOẠI ẢNH UTV SỬ DỤNG MẠNG HỌC SÂU VÀ HỆ THỐNG HỖ TRỢ CHẨN ĐOÁN UTV TỪ XA SỬ DỤNG MẠNG HỌC SÂU



(a) Đường cong độ chính xác

(b) Đường cong mất mát

Hình 5.22. Các Đường cong độ chính xác và mất mát của mô hình

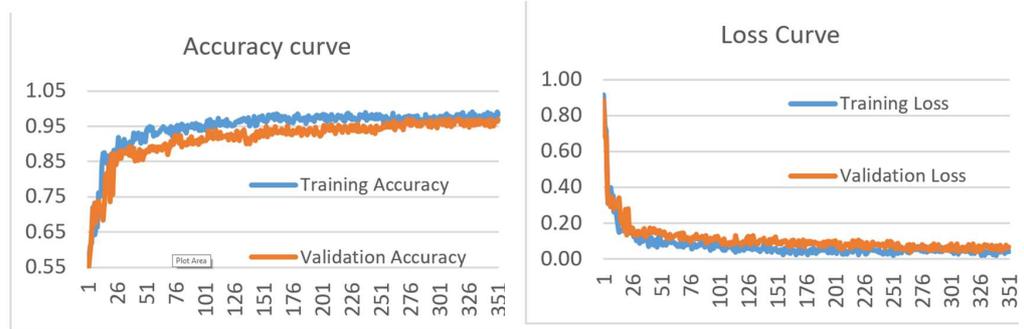


(a) Đường cong độ chính xác

(b) Đường cong mất mát

Hình 5.22 hiển thị các đường cong cho mô hình phân loại, bao gồm Accuracy curve và Loss curve. Đường cong huấn luyện hội tụ suôn sẻ từ giai đoạn đầu tiên đến giai đoạn cuối cùng và phù hợp với mong đợi, với độ chính xác tăng dần. Hơn nữa, khoảng cách giữa đường cong Accuracy validation (màu xanh lá cây) và đường cong training (màu xanh) cho thấy mô hình không xảy ra overfitting trong quá trình huấn luyện. Tương tự như đường cong huấn luyện, đường cong tổn thất thể hiện trong

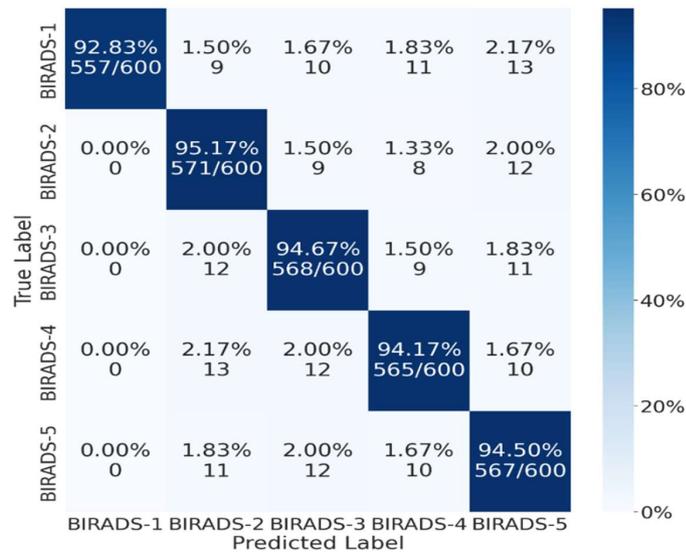
Chương 5 PHÂN LOẠI ẢNH UTV SỬ DỤNG MẠNG HỌC SÂU VÀ HỆ THỐNG HỖ TRỢ CHẨN ĐOÁN UTV TỪ XA SỬ DỤNG MẠNG HỌC SÂU



(a) Đường cong độ chính xác

(b) Đường cong mất mát

Hình 5.22 (b) hội tụ đều đặn. Với những kết quả trên chúng ta thấy mô hình đề xuất có hiệu quả cao trong việc phân loại.



Hình 5.23. Ma trận nhầm lẫn

Hình 5.23 thể hiện kết quả sử dụng ma trận nhầm lẫn cho mô hình phân loại sử dụng mạng EfficientNet với độ chính xác cao nhất. Cụ thể, các giá trị hàng biểu thị nhãn thực tế của các bộ ảnh, từ BIRADS-1 đến BIRADS-5, các giá trị cột mô tả nhãn dự đoán của 5 loại bộ ảnh này và các đường chéo biểu thị các giá trị True Positive (TP). Mô hình đề xuất cho kết quả phân loại tốt nhất đối với mức BIRADS-2.

Bảng 5.14 Kết quả so sánh khi có áp dụng/ không có áp dụng các bước XLA

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

Method	Train Accuracy (%)	Validating Accuracy (%)
Original image without processing	75.6	73.2
Background & Artefacts Removal	87.6	82.3
Adaptive Gamma Correction	87.8	83.7
Histogram Equalization	87.4	84.1
Applying of Colormap	94.2	93.6

Kết quả cho thấy, nhìn chung, tiền XLA giúp gia tăng đáng kể độ chính xác chung của việc phân loại ảnh UTV theo thang đo BIRADS. Cụ thể hơn bước phân đoạn ảnh để loại bỏ background ảnh và các thành phần không mong muốn giúp cải thiện đáng kể độ chính xác với sự gia tăng hơn 8%, việc hiệu chỉnh Gamma có giúp cải thiện độ chính xác nhưng không nhiều (khoảng 1%) trong khi cân bằng Histogram cho kết quả cải thiện không đáng kể. Cuối cùng, việc sử dụng Colormap trên ảnh X-quang để chuyển ảnh xám thành ảnh màu trước khi đưa vào bộ phân loại cũng giúp tăng đáng kể độ chính xác (khoảng 8%). Điều này cho thấy rằng, mô hình EfficientNet rất phù hợp với những dữ liệu ảnh màu.

Trong những năm gần đây, nhiều kết quả nghiên cứu về bộ ảnh chụp nhũ ảnh để phân loại, phát hiện và phân đoạn đã được công bố với độ chính xác ngày càng tăng, trong đó mạng học sâu đã được sử dụng nhiều [172]. CNN có thể phát hiện ra hình ảnh UTV siêu âm với Diện tích dưới đường cong (AUC) là 96% [173]. Tuy nhiên, các nghiên cứu dựa trên siêu âm này có những hạn chế về số lượng các loại hình ảnh siêu âm khác nhau [174].

Trong Bảng 5.15, một nghiên cứu cho thấy độ chính xác trung bình của ba mô hình CNN, bao gồm Inception-V3, ResNet 50 và AlexNet) cho 70% là 97,81% [107]. Các nghiên cứu khác nhau cho thấy việc sử dụng hình ảnh siêu âm mang lại kết quả rất tốt với độ chính xác từ 89,4% đến 99,1% [148], [175], [176]. Mô hình VGG16 đạt độ chính xác 97,9% với các bộ ảnh chụp nhũ ảnh phóng to 100 lần do Khoa Chẩn đoán hình ảnh của BV Nhân dân Ganzhou thu được [177]. Trong khi một mô hình khác đạt được độ chính xác 99,17% và 98,42% khi sử dụng tập dữ liệu CBIS-DDSM và MIAS tương ứng [178]. Đánh giá hiệu suất của giai đoạn CNN được đề xuất là

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

AUC 91,9% (95% CI: 0,852, 0,971) cho nhóm đào tạo và AUC 87,7% (95% CI: 0,805, 0,949) cho nhóm xác thực [36]. Một nghiên cứu cho thấy độ chính xác 89% cũng đạt được khi phân loại theo thang BIRADS với bộ phân loại truyền thống MLF [179].

**Bảng 5.15. Thống kê và đánh giá mức độ tổn thương của hình ảnh với BIRADS**

Papers	Dataset	Method	Output	Results
A. S. Elkorany and Z. F. Elsharkawy – 2023 [107]	MIAS	Inception-V3, ResNet 50 và AlexNet	Normal Benign Malignant	97.8% Acc
S.-M. Hsu và cộng sự – 2019 [175]	Ultrasound	Handcrafted ML	Benign Malignant	89.4% Acc
B. J. E. L. H. S. K. S. H. Choi Ji-Hye Kang Bong Joo – 2018 [176]	Ultrasound	CNN transfer learning	Benign Malignant	99.1% Acc
Z. Liu, J. Peng, X. Guo, S. Chen, and L. Liu – 2024 [177]	Local Mammograms	Deep learning radiomics nomogram	Normal Benign Malignant	97.9% Acc
M. A. K. Raiaan và cộng sự – 2024 [178]	CBIS-DDSM MIAS	Mammo-Light	Normal Benign Malignant	99.2% Acc
E. Tarighati Sereshkeh, H. Keivan, K. Shirbandi, F. Khaleghi, and M. M. Bagheri Asl – 2024 [179]	Local Mammograms	CNN	Normal Benign Malignant	87.7% Acc
E. BOROUMANDZADEH, Mostafa; PARVINNIA – 2021 [180]	Local Mammograms	MLF Classifier	BIRADS 0 to 6	89.0% Acc
H. T. X. Nguyen và cộng sự – 2022 [181]	VinDr-Mammo	LightGBM	BIRADS 1 to 5	67.0% F1 score
Nghiên cứu này	VinDr-Mammo	EfficientNet-B7 With CROI	BIRADS 1 to 5	92.6% Acc 92.33% F1 score

Trong nghiên cứu của chúng tôi, độ chính xác 92,6% đạt được trên trung bình năm bộ ảnh BIRADS cân bằng của tập dữ liệu VinDr-Mammo như thể hiện trong

## 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

---

Bảng 5.15. Ngoài ra, các ảnh có CROI đạt được độ chính xác trung bình của điểm F1 là 92,33%. Trong khi một nghiên cứu trước đây với cùng một bộ ảnh sử dụng bộ phân loại Light Gradient Boosting Machine (LightGBM) đạt được độ chính xác của điểm F1 là 67,0% [181] như mô tả trong Bảng 5.15.

Bài viết này đề xuất các phương pháp xử lý năm bộ ảnh chụp nhũ ảnh không cân bằng của tập dữ liệu VinDr-Mammo. Đặc biệt, phân đoạn Otsu được áp dụng để xác định ngưỡng tối ưu để trích xuất các ảnh có GROI được chuyển đổi thành các ảnh có CROI trước khi phân loại. Ngoài ra, EfficientNet-B7 đã được áp dụng để đánh giá phân loại từng phương pháp đề xuất của hai trường hợp (GROI và CROI), trong đó tập ảnh có CROI đạt độ chính xác cao hơn 92,6% so với GROI. Để đạt được độ chính xác cao này, các tập ảnh cũng được cân bằng bằng các phép biến đổi như xoay và lật. Điều này có nghĩa là các kết quả đã chứng minh để minh họa cho hiệu quả của phương pháp đề xuất. Do đó, phương pháp đề xuất này có thể phát triển để áp dụng cho việc phân loại nhiều tập ảnh UTV khác nhau và cũng cho các ứng dụng thực tế.

### 5.2.5 Hợp nhất bộ dữ liệu để phân loại tổn thương với EfficientNet kết hợp với Multihead Attention

Kiến trúc của phương pháp đề xuất được thiết kế đặc biệt để tận dụng dữ liệu chụp nhũ ảnh đa góc nhìn nhằm phân loại tổn thương vú một cách chính xác. Phương pháp đề xuất sử dụng hai tập dữ liệu riêng biệt: góc nhìn MLO và góc nhìn CC, mỗi góc nhìn bao gồm một số loại tổn thương như Bình thường, Khối u, Vôi hóa và Bất đối xứng cục bộ. Với mỗi loại ảnh MLO-CC, ROI được trích xuất và cải thiện thông qua việc tô màu hình ảnh, tạo ra các đầu vào thông tin cho phân tích tiếp theo.

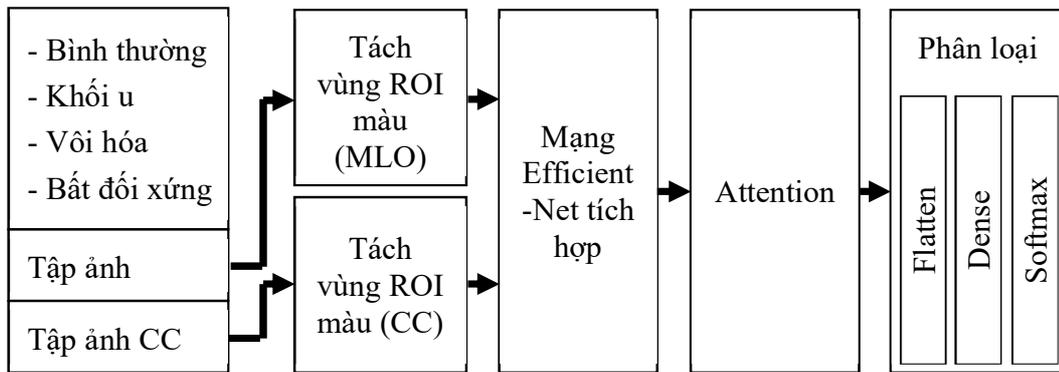
Cả ảnh MLO và CC ROI được tô màu sau đó được đưa đồng thời vào mạng EfficientNet-B7 kết hợp, cho phép trích xuất đặc điểm sâu với khả năng mở rộng phức hợp và hiệu quả mô hình tối ưu. Mạng này có khả năng nắm bắt các mẫu không gian và ngữ cảnh phức tạp từ mỗi góc nhìn chụp nhũ ảnh. Các đặc điểm được trích xuất từ cả hai nhánh được hợp nhất để tạo thành một biểu diễn chung toàn diện.

Để tăng cường hơn nữa khả năng phân biệt và khả năng thích ứng, một cơ chế

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

Attention được áp dụng. Cơ chế này ưu tiên động các vùng liên quan và loại bỏ nhiễu nền, đảm bảo các khía cạnh thông tin nhất của hình ảnh được làm nổi bật cho quyết định cuối cùng.

Các đặc trưng được chú ý sau đó được chuyển vào bộ phân loại, bao gồm một chuỗi các lớp: flatten, dropout, dense và một hàm softmax cuối cùng để tạo ra đầu ra đa lớp. Quy trình thống nhất này cho phép phát hiện và phân loại tổn thương vú với độ chính xác cao bằng cách tích hợp tối ưu thông tin đa nguồn, đa góc nhìn và các kỹ thuật học sâu tiên tiến.



**Hình 5.24 Kiến trúc của một hệ thống phân loại mới**

Trong nghiên cứu, hai loại tập ảnh, MLO và CC, đã được chọn từ cơ sở dữ liệu VinDr-Mammo [41] và một tập khác được thu thập từ Bệnh viện Ung bướu TP.HCM tại Việt Nam. Đặc biệt, đối với tập ảnh ung thư vú từ Bệnh viện Ung bướu TP.HCM, việc hợp tác nghiên cứu với các bác sĩ tại bệnh viện đã được thực hiện trong nhiều tháng và các hình ảnh đã được các chuyên gia xem xét và phân loại trước khi sử dụng cho nghiên cứu này. Để tích hợp hai tập ảnh, ảnh DICOM đã được xử lý sao cho có cùng định dạng với kích thước 224x224.

Hai loại, MLO và CC, được kết hợp từ các bộ ảnh VD và OH để đảm bảo số lượng ảnh đủ cho từng loại tổn thương. Số lượng ảnh sau khi chọn lọc và loại bỏ các ảnh chất lượng thấp được mô tả trong Bảng 2. Cụ thể, bốn loại được sử dụng: Bình thường, Khối u, Vôôi hóa và Bất đối xứng cục bộ.

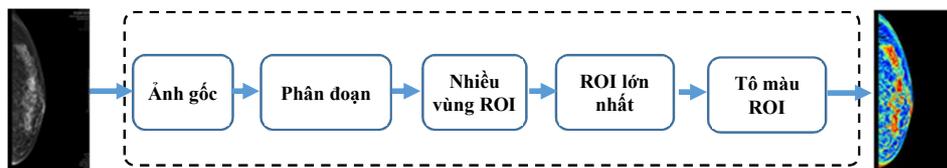
**Bảng 5.16 Mô tả các tập dữ liệu và lớp tổn thương vú**

TT	Bộ dữ liệu	Các loại tổn thương vú
----	------------	------------------------

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

		Bình thường		Khối		Vôi hóa		Bất đối xứng cục bộ	
		MLO	CC	MLO	CC	MLO	CC	MLO	CC
1	Bệnh viện Ung bướu TP.HCM	382	382	198	198	560	560	542	542
2	VinDr-Mammo	393	393	561	561	201	201	116	116
Tổng cộng		775	775	775	759	759	761	761	658
		1550		1518		1522		1316	

Cấu trúc xử lý hình ảnh vú này tuân theo phương pháp từng bước được thiết kế để nâng cao khả năng phân tích và trực quan hóa các vùng quan tâm (ROI). Bắt đầu với hình ảnh vú ban đầu, các thuật toán phân đoạn được áp dụng để phân biệt và cô lập các vùng giải phẫu có ý nghĩa trong quá trình quét. Quá trình này thường tạo ra một số ROI ứng viên. Thông qua phân tích đa ROI, mỗi vùng được phân đoạn được đánh giá và ROI lớn nhất - thường đại diện cho mô vú chính - được chọn để xử lý tiếp theo. Để tinh chỉnh kết quả, ROI đã chọn được sử dụng để che các vùng không liên quan, đảm bảo rằng phân tích tiếp theo chỉ tập trung vào mô quan tâm. Cuối cùng, một thuật toán tô màu được áp dụng cho ROI lớn nhất, chuyển đổi dữ liệu thang độ xám thành hình ảnh tô màu giúp khuếch đại độ tương phản và cấu trúc mô, làm cho ranh giới tổn thương nổi bật hơn về mặt thị giác. Quy trình này không chỉ hợp lý hóa việc phát hiện và đánh giá các tổn thương vú mà còn cung cấp cho các bác sĩ X quang và các mô hình AI thông tin đầu vào nâng cao để chẩn đoán chính xác hơn.



**Hình 5.25 Biểu diễn việc tạo ra một hình ảnh ngực được tô màu.**

Để xác định ROI từ ảnh gốc  $I(x, y)$ , một thuật toán phân đoạn được áp dụng để xác định ngưỡng nhị phân hóa  $T$ , từ đó xác định ngưỡng ảnh  $I_{Th}(x, y)$ . Do đó, việc trích xuất đường viền của các thành phần liên thông  $C = \{C_1, C_2, \dots, C_n\}$  được tính toán, trong đó mỗi thành phần  $C_i$  là một tập hợp tọa độ điểm ảnh để tạo thành một đường viền khép kín trong  $I_{Th}(x, y)$ .

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

---

Để tìm ROI lớn nhất, diện tích đường viền  $C_A$  được xác định dựa trên một  $I_{Th}(x, y) = 1$ , nếu  $I(x, y) \geq T$  và phương trình của nó được mô tả như sau:

$$C_A(C_i) = \sum_{(x,y) \in C_i} 1 \quad (5.11)$$

và ROI lớn nhất được chọn bằng cách sử dụng phương trình sau:

$$C_L = \arg \max_{C_i \in C} C_A(C_i) \quad (5.12)$$

Do đó, hình ảnh có ROI lớn nhất được gọi là  $I_L = (x, y)$ (2)

Giả sử rằng hình ảnh ROI lớn nhất  $I_L = (x, y)$  là hình ảnh có kênh  $R, G, B$ , nó có thể được chuyển đổi thành hình ảnh không gian màu HSV để đào tạo mạng học sâu bằng cách sử dụng phương trình sau:

$$I_{L\_HSV}(x, y) = Con2HSV(I_L(x, y)) \quad (5.13)$$

trong đó,

$$I_{L\_HSV}(x, y) = (I_{L\_H}(x, y), I_{L\_S}(x, y), I_{L\_V}(x, y)) \quad (4)$$

Để phát hiện một trong các vùng  $H, S, V$ , các phương trình sau được sử dụng để trích xuất mặt nạ cho các ngưỡng trên  $\tau_H, \tau_S, \tau_V$  của  $H, S, V$ :

$$I_{L\_H}(x, y) = \begin{cases} 1, & \text{if } H(x, y) \geq \tau_H \\ 0, & \text{otherwise} \end{cases} \quad (5.14)$$

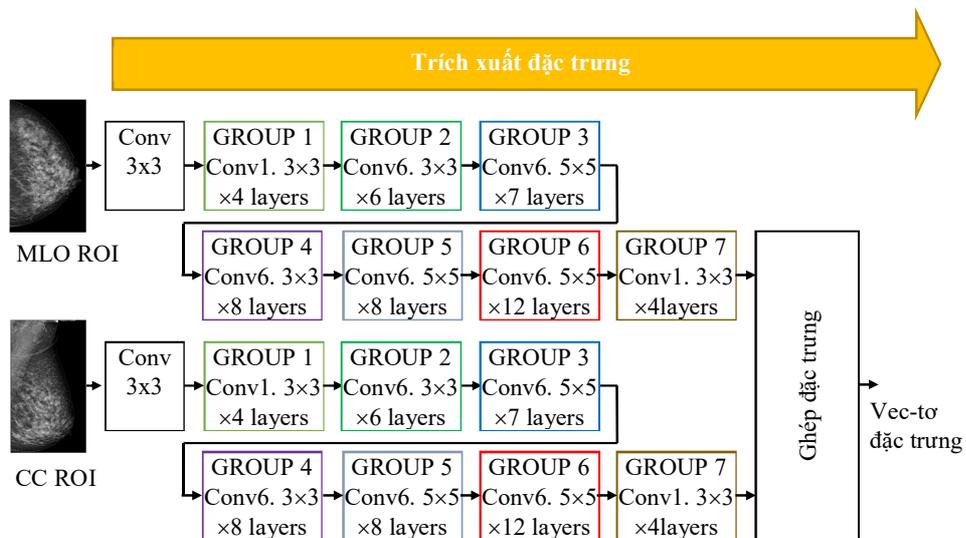
$$I_{L\_S}(x, y) = \begin{cases} 1, & \text{if } S(x, y) \geq \tau_S \\ 0, & \text{otherwise} \end{cases} \quad (5.15)$$

$$I_{L\_V}(x, y) = \begin{cases} 1, & \text{if } V(x, y) \geq \tau_V \\ 0, & \text{otherwise} \end{cases} \quad (5.16)$$

Cấu trúc được mô tả minh họa quy trình trích xuất đặc trưng dựa trên EfficientNet-B7 đường dẫn kép, được thiết kế riêng cho phân tích chụp nhũ ảnh. Mỗi ROI, một từ góc nhìn MLO và một từ góc nhìn CC, trước tiên được truyền qua lớp tích chập 3x3 ban đầu, hoạt động như một bộ tiền xử lý đặc trưng cơ bản. Sau đó, dữ

## 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

liệu từ mỗi đường dẫn đi qua một tập hợp các nhóm tích chập tuần tự được sắp xếp cẩn thận, trong đó mỗi nhóm được xác định bởi các kích thước hạt nhân cụ thể và một số lớp lặp lại được chỉ định. Các nhóm này dần dần trích xuất cả các đặc trưng phân cấp cục bộ và toàn cục, duy nhất cho mỗi góc nhìn. Sau khi trích xuất đặc trưng hoàn tất cho cả hai đầu vào, các bản đồ đặc trưng kết quả sẽ được nối hoặc kết hợp, tạo thành một vectơ đặc trưng toàn diện duy nhất. Sự kết hợp này nắm bắt các chi tiết giải phẫu bổ sung từ cả góc nhìn MLO và CC, điều này rất quan trọng để phát hiện và phân loại tổn thương vú chính xác. Toàn bộ kiến trúc minh họa khả năng mạnh mẽ của EfficientNet-B7 trong việc cân bằng chiều sâu và chiều rộng của mô hình, tạo ra biểu diễn đặc trưng đa góc nhìn mạnh mẽ, phù hợp cho các tác vụ chẩn đoán tiếp theo.



**Hình 5.26 Kiến trúc kết hợp EfficientNet-B7 để tạo ra vectơ đặc trưng của MLO và CC.**

Multihead là một phần mở rộng của cơ chế Self Attention trong mô hình Transformer, cho phép mạng học cách tập trung vào nhiều khía cạnh hoặc khía cạnh của dữ liệu đầu vào cùng một lúc. Cụ thể, Multihead Attention hoạt động như sau: Đặc trưng đầu vào được chia thành các phần nhỏ hơn, mỗi phần được gọi là một head attention. Mỗi head có một tập hợp các ma trận trọng số riêng để biến đổi đầu vào thành các vectơ trong các không gian con riêng biệt. Mỗi head tính toán sự chú ý của

## 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

---

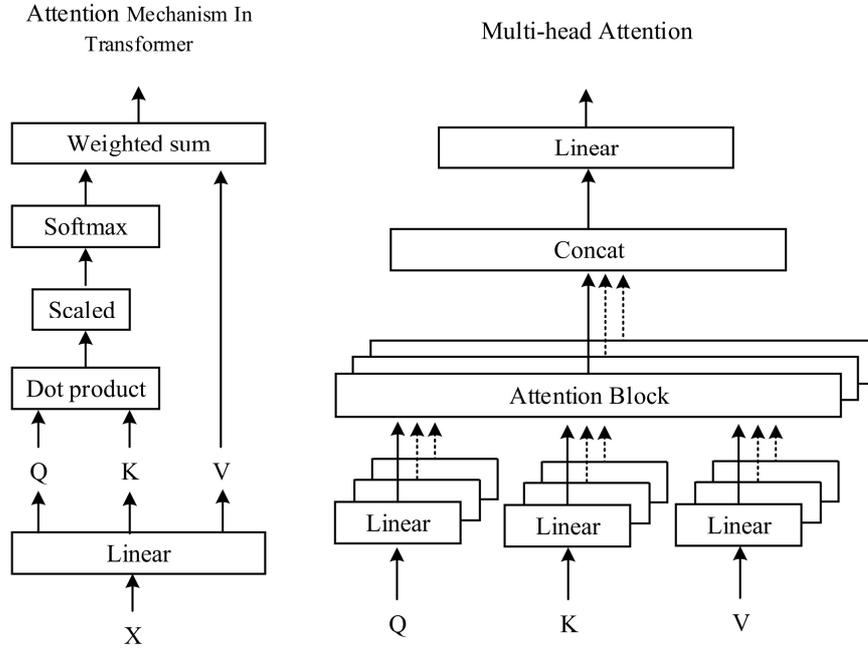
riêng mình bằng phương pháp chú ý tích vô hướng tỷ lệ; nghĩa là, nó đánh giá mức độ liên quan (điểm chú ý) giữa các phần tử đầu vào thông qua Q và K, sau đó sử dụng điểm này để tạo ra một vectơ đặc trưng có trọng số cho Giá trị.

Đầu ra từ các đầu đọc độc lập được nối lại thành một tenxơ lớn hơn. Tenxơ này sau đó được biến đổi bằng một ma trận trọng số cuối cùng để đạt được kích thước cần thiết cho mô hình. Nhờ phương pháp này, mô hình không chỉ tập trung vào một loại mối quan hệ hoặc đặc điểm duy nhất mà còn có thể học song song nhiều "khía cạnh" chú ý. Trong hình ảnh, các đầu đọc có thể học các vùng nổi bật khác nhau hoặc các mức độ chi tiết khác nhau của các đặc điểm.

Các đặc điểm đáng chú ý của mô hình Multihead Attention bao gồm: Nâng cao khả năng biểu diễn và học tập của mô hình khi nó xử lý thông tin trên nhiều không gian con và góc nhìn. Mô hình này cũng tăng hiệu quả học tập bằng cách thực hiện các thao tác chú ý song song. Sử dụng kết nối dư và chuẩn hóa lớp giúp lưu giữ thông tin và ổn định gradient.

Tóm lại, Multihead Attention cho phép mạng kết hợp và lựa chọn thông tin đa chiều, phong phú và phức tạp so với sự chú ý tự thân của một đầu.

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



**Hình 5.27 Cấu trúc của mạng Multihead Attention**

Trong mô hình EfficientNet-B7 kết hợp với hai lớp đầu vào cho hai ROI lớn nhất được tô màu của MLO và CC, hai tính năng EfficientNet-B7 được mô tả như sau:

$$\mathbf{F}_{MLO} = f_{EffB7}(I_{CL\_MLO}(x, y)) \quad (5.17)$$

$$\mathbf{F}_{CC} = f_{EffB7}(I_{CL\_CC}(x, y)) \quad (5.18)$$

trong đó  $I_{CL\_MLO}(x, y)$  hoặc  $I_{CL\_CC}(x, y)$  là một trong các  $I_{L\_V}(x, y)$ ,  $I_{L\_S}(x, y)$ ,  $I_{L\_H}(x, y)$  và  $f_{EffB7}$  biểu thị toàn bộ ngăn xếp các khối EfficientNet-B7 để trích xuất các vectơ đặc trưng.

Do đó, việc kết hợp các tính năng từ cả hai chế độ xem MLP-CC được thể hiện như sau:

$$\mathbf{F}_{Comb} = \mathbf{F}_{MLO} + \mathbf{F}_{CC} \quad (5.19)$$

Để tính toán mạng Attention trên các tính năng kết hợp  $\mathbf{F}_{Comb}$ , một cơ chế chú ý chung được áp dụng để tạo ra nhóm toàn cầu cho kênh Attention như sau:

$$z_{GP} = GAP(\mathbf{F}_{Comb}) \quad (5.20)$$

## 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

---

trong đó  $GAP$  biểu thị tổng hợp trung bình toàn cầu.

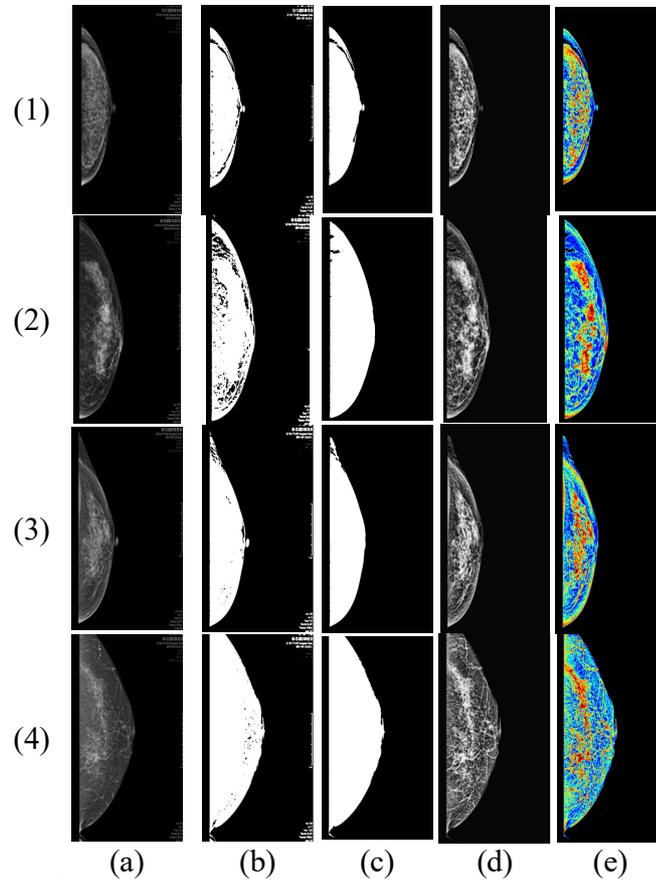
Do đó, trọng số chú ý được tính bằng phương trình sau:

$$\varpi = \sigma(W_2 \bullet \text{ReLU}(W_1 \bullet z_{GAP})) \quad (5.21)$$

trong đó  $W_1$ ,  $W_2$  là các tham số đã học và  $\sigma$  là hàm sigmoid.

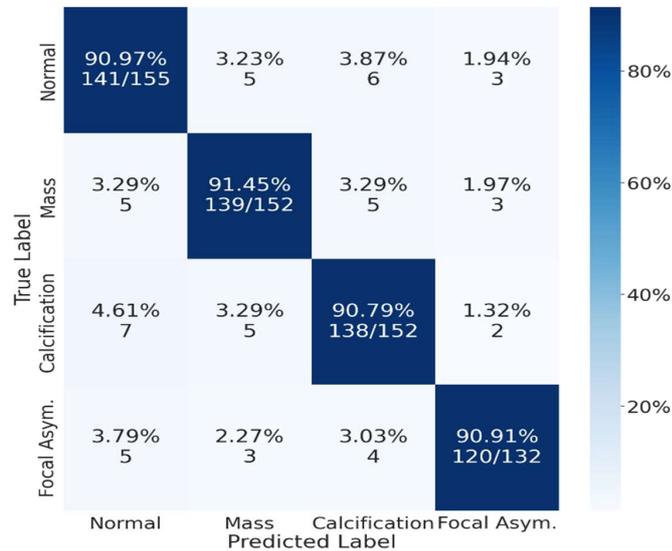
Hình 5.28 cho thấy kết quả của các bước xử lý hình ảnh, minh họa bằng hình ảnh từ bốn loại tổn thương điển hình: (1) bình thường; (2) khối u; (3) vôi hóa; (4) bất đối xứng cục bộ. Cột (a) hiển thị các hình ảnh đầu vào ban đầu; các hình ảnh trong cột (b) hiển thị các vùng được phân đoạn được biểu diễn dưới dạng mặt nạ nhị phân, trong đó nhiều vùng xuất hiện, bao gồm vùng vú và các vùng khác được tô sáng trên nền, chẳng hạn như chú thích trên hình ảnh; (c) hiển thị kết quả sau khi chọn vùng lớn nhất, tức là vùng vú cần quan tâm; (d) là kết quả sau khi sử dụng mặt nạ nhị phân của ROI lớn nhất để loại bỏ các vùng không cần thiết; (e) là kết quả sau khi áp dụng thuật toán tô màu.

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



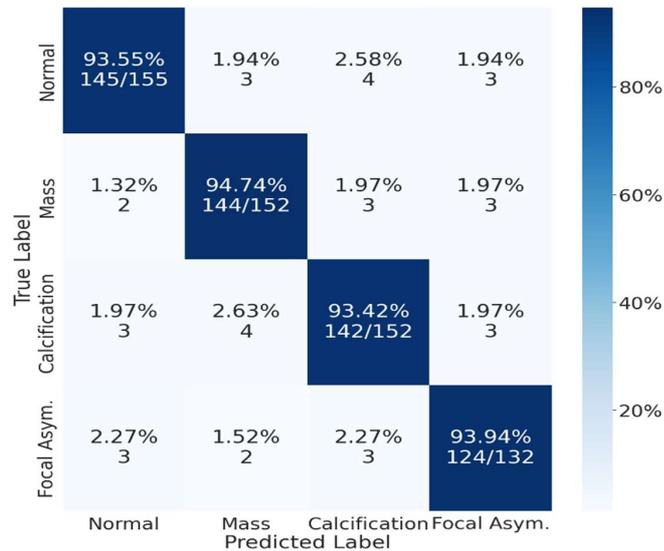
**Hình 5.28** Biểu diễn xử lý hình ảnh: (a) Hình ảnh gốc; (b) Mặt nạ đa ROI; (c) Mặt nạ ROI lớn nhất; (d) ROI lớn nhất RGB (e) ROI được tô màu và (1) Bình thường; (2) Khối lượng; (3) Vôi hóa; (4) Bất đối xứng cục bộ

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



**Hình 5.29 Biểu diễn ma trận nhầm lẫn cho bốn loại tổn thương vú không có Multihead Attention.**

Hình 5.29 là ma trận nhầm lẫn khi không áp dụng phương pháp chú ý đa đầu. Kết quả sử dụng mô hình CNN với hai đầu vào ảnh MLO-CC song song cho thấy độ chính xác trên 90%.



**Hình 5.30 Biểu diễn ma trận nhầm lẫn cho bốn loại tổn thương vú với Multihead Attention**

Hình 5.30 cho thấy độ chính xác tăng đáng kể khi áp dụng phương pháp chú ý đa đầu vào cho hệ thống, với độ chính xác được cải thiện khoảng 2% so với kết quả

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

trước đó.

**Bảng 5.17 Mô tả các tập dữ liệu và lớp tổn thương vú**

Thể loại	Số lượng hình ảnh	Không có Attention				Có Attention			
		Đúng	Đương tính giả			Đúng	Đương tính giả		
Normal	155	141	5 to Mass	6 to Calc.	3 to Focal Asymmetry	145	3 to Mass	4 to Calc.	3 to Focal Asymmetry
Mass	152	139	5 to Normal	5 to Calc.	3 to Focal Asymmetry	144	2 to Normal	3 to Calc.	3 to Focal Asymmetry
Calcification	152	138	7 to Normal	5 to Mass	2 to Focal Asymmetry	142	3 to Normal	4 to Mass	3 to Focal Asymmetry
Focal Asymmetry	132	120	5 to Normal	3 to Mass	4 to Calc.	124	3 to Normal	2 to Mass	3 to Calc.

Bảng 5.17 hiển thị kết quả phân loại chi tiết. Trong cả hai trường hợp, có thể thấy độ chính xác phân loại cho lớp "Khác" thấp hơn đáng kể so với các lớp còn lại. Bảng 5.18 cung cấp tóm tắt chi tiết về độ chính xác, recall và F1 score cho mô hình phân loại được đề xuất trên bốn loại tổn thương vú: Bình thường, Khối u, Vôi hóa và Bất đối xứng cục bộ. Kết quả cho thấy hiệu suất cao liên tục đối với các loại tổn thương điển hình, với giá trị độ chính xác dao động từ khoảng 0,94 đến 0,95 và điểm F1 trên 0,94. Nhìn chung, các số liệu này làm nổi bật hiệu quả của mô hình trong việc phân biệt các lớp quan trọng về mặt lâm sàng đồng thời xác định các cơ hội để cải thiện hơn nữa, đặc biệt là đối với các loại tổn thương khó.

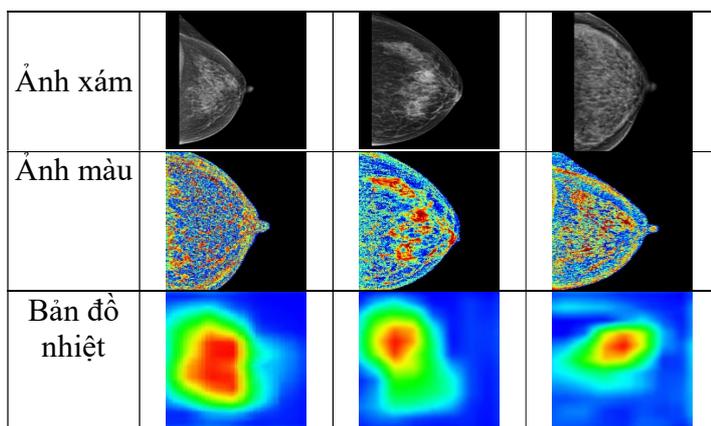
**Bảng 5.18 Biểu diễn độ chính xác và điểm F1**

Loại tổn thương	Độ chính xác	Recall	điểm f1
Bình thường	0.9477	0.9355	0.9416
Khối u	0.9412	0.9474	0.9443
Vôi hóa	0.9342	0.9342	0.9342
Bất đối xứng cục bộ	0.9323	0.9394	0.9358

Kết quả trên minh họa hiệu quả của từng bước trong quy trình xử lý và phân loại hình ảnh để phân tích tổn thương vú. Ban đầu, hình ảnh từ bốn loại đại diện-bình thường, khối u, vôi hóa và các loại khác-được xử lý bằng thuật toán phân đoạn để tạo mặt nạ nhị phân. Chỉ ROI lớn nhất, thường tương ứng với mô vú có liên quan, được chọn để xử lý tiếp theo và sau đó được cải thiện bằng thuật toán tô màu. Chuỗi biến

## 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

đổi này liên tục cải thiện độ nét và khả năng hiển thị của các vùng quan trọng để phát hiện tổn thương.



**Hình 5.31 Biểu diễn bản đồ nhiệt tương ứng với ảnh màu và ảnh xám**

Hình trên trình bày ảnh xám, ảnh màu và bản đồ nhiệt, trong đó bản đồ nhiệt là kết quả phân loại từ mạng EfficientNet-B7. Cụ thể, các bản đồ nhiệt với các màu khác nhau cho thấy mạng học sâu này tập trung vào các đặc điểm tính toán và các vùng tổn thương trên ảnh X-quang vú để tạo ra kết quả chính xác. Hơn nữa, ảnh ROI xám và ảnh màu thể hiện sự tương đồng giữa các vùng tổn thương và các vùng đặc trưng được hiển thị trong bản đồ nhiệt.

Phân tích chi tiết về hiệu suất phân loại được thể hiện thông qua ma trận nhầm lẫn. Không cần sự chú ý của nhiều đầu, kiến trúc CNN đề xuất với đầu vào chế độ xem MLO và CC song song đạt độ chính xác hơn 90% trên hầu hết các lớp. Tuy nhiên, hạng mục tổn thương "Khác", bao gồm các trường hợp bất đối xứng và biến dạng, cho thấy độ chính xác thấp hơn do các đặc điểm đặc trưng đa dạng và kém riêng biệt.

Khi áp dụng phương pháp chú ý đa đầu, độ chính xác phân loại tổng thể của hệ thống đã được cải thiện khoảng 2%, cho thấy cơ chế chú ý động đã nâng cao khả năng học thuộc tính và ra quyết định cho hầu hết các nhóm.

Đường cong độ chính xác huấn luyện và xác thực càng củng cố thêm tính vững

## 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

---

chắc của mô hình đề xuất, vì chúng cho thấy sự hội tụ nhanh và không có dấu hiệu quá khớp. Những kết quả này cùng nhau làm nổi bật điểm mạnh của quy trình phát hiện tổn thương vú và nhấn mạnh các lĩnh vực cần được tinh chỉnh trong tương lai, đặc biệt là trong việc xử lý các trường hợp tổn thương phức tạp hoặc mơ hồ.

Kết quả thu được trong nghiên cứu này cho thấy những cải tiến quan trọng so với các nghiên cứu trước đây về phân loại tổn thương vú đa lớp. Trước đây, nhiều nỗ lực nghiên cứu đã sử dụng chụp nhũ ảnh một góc nhìn hoặc phương pháp CNN tiêu chuẩn, thường cho kết quả tốt đối với các phân loại như "bình thường" và "khối", nhưng độ chính xác lại hạn chế, thường dưới 90%, đối với các loại tổn thương phức tạp hoặc mơ hồ hơn. Các phương pháp trước đây này cũng gặp khó khăn trong việc khái quát hóa và đặc biệt gặp khó khăn với các phân loại có độ biến thiên nội lớp cao.

Bằng cách áp dụng chiến lược đa góc nhìn sử dụng cả tập dữ liệu MLO và CC, đồng thời tích hợp EfficientNet-B7 cùng với các cơ chế chú ý đa đầu, phương pháp hiện tại phù hợp với những tiến bộ gần đây trong học sâu cho hình ảnh y tế. Hợp nhất đa góc nhìn đã được chứng minh là cải thiện tỷ lệ phân loại bằng cách tận dụng các đặc điểm bổ sung từ các góc nhìn giải phẫu khác nhau, với các thí nghiệm của bạn xác nhận độ chính xác tổng thể tăng khoảng 2% khi bổ sung sự chú ý đa đầu-phù hợp với các tiêu chuẩn gần đây trong tài liệu.

Tóm lại, kết quả của bạn có thể so sánh thuận lợi với các phương pháp đa chế độ xem, dựa trên EfficientNet và tăng cường sự chú ý gần đây, chứng minh hiệu suất vượt trội hoặc tương đương đối với hầu hết các danh mục và có lộ trình rõ ràng để cải thiện trong tương lai ở các lớp phức tạp, không đồng nhất.

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

Mục tiêu của nội dung này là xây dựng một hệ thống hỗ trợ chẩn đoán từ xa sử dụng mạng học sâu CNN để giúp phân loại ảnh UTV. Hệ thống cần đảm bảo việc truyền tải và xử lý hình ảnh y tế qua mạng kết hợp với mạng học sâu CNN để phân loại ảnh UTV hỗ trợ cho bác sĩ trong quá trình chẩn đoán.

## 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

---

### 5.3.1 Các yêu cầu thiết kế

Việc xây dựng hệ thống theo đúng tiêu chuẩn và tối ưu hóa là vô cùng cần thiết để đảm bảo hiệu suất tổng thể. Để đáp ứng các mục tiêu đã đặt ra, một số yêu cầu thiết kế cơ bản được xác lập, để từ đó có thể xây dựng bảng các tham số thiết kế chi tiết hơn, định hướng cho việc xây dựng hệ thống như vậy. Các yêu cầu cơ bản này là sự phản ánh của các mục tiêu thiết kế vào trong tính năng kỹ thuật của hệ thống và sẽ là định hướng cho việc triển khai các bước tiếp theo.

Các yêu cầu thiết kế cơ bản bao gồm:

- Hệ thống hoạt động trên máy chủ trên môi trường mạng Internet, cung cấp khả năng truy cập mọi lúc, mọi nơi.
- Hỗ trợ giao thức DICOM, cho phép truyền tải các ảnh X-quang vú, đồng thời cũng hỗ trợ khả năng upload ảnh JPEG.
- Giao tiếp người dùng (BS) dựa trên nền tảng web, cho phép dễ dàng truy cập với nhiều loại thiết bị khác nhau (máy tính, điện thoại thông minh, máy tính bảng, ...) không giới hạn về hệ điều hành, phiên bản phần mềm.
- Các thành phần giao tiếp với nhau thông qua các giao thức và API phù hợp, cho phép dễ dàng phát triển và nâng cấp.
- Cung cấp kết quả ngõ ra bao gồm loại tổn thương, mức độ nguy cơ (lành tính/ ác tính) và khoanh vùng quan tâm trên ảnh.

Nhưng cũng cần lưu ý rằng các kết quả ngõ ra của hệ thống AI có tính chất hỗ trợ và tham khảo cho bác sĩ để ra các kết luận chẩn đoán. Do các vấn đề về luật, các quy định và các khía cạnh y đức, kết quả của hệ thống AI không được xem là, và không thay thế kết quả chẩn đoán của bác sĩ chuyên khoa.

Để đáp ứng các yêu cầu trên, một bảng các tham số chi tiết cho hệ thống được xác định và thiết lập nhằm cụ thể hóa các yêu cầu của hệ thống. Bảng dưới đây cho biết chi tiết các tham số mà hệ thống cần đạt được. Trong đó mỗi thông số được đưa ra với các yêu cầu cụ thể cần đạt được để đảm bảo khả năng vận hành của hệ thống và đáp ứng các mục tiêu ban đầu.

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

**Bảng 5.19 Các thông số yêu cầu hệ thống**

TT	Thông số	Đơn vị	Yêu cầu	Chi tiết
1	Số mô hình AI được áp dụng	Mô hình	$\geq 1$	
2	Số thành phần của hệ thống	thành phần	3	- Server ảnh (DICOM) - Web interface (frontend, backend) - AI processor
3	Định dạng ảnh hỗ trợ	-	2	DICOM, JPEG
4	Giao thức sử dụng	-	-	TCP/IP, DICOM, HTTP,
5	Dung lượng lưu trữ (ảnh)	GB	$\geq 100$ GB	
6	Khả năng truy cập (số kết nối đồng thời)	Kết nối	$\geq 50$	
7	Loại thiết bị đầu cuối hỗ trợ	-	-	- PC - Điện thoại thông minh - Máy tính bảng
8	Kết quả ngõ ra (số loại kết quả)	loại	3	- Loại tổn thương - Nguy cơ ung thư - Vùng quan tâm

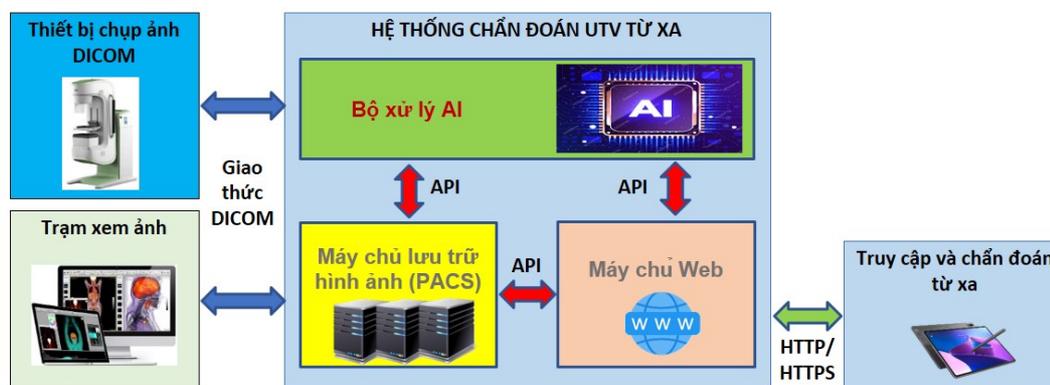
Bảng 5.1 mô tả 8 thông số cơ bản của hệ thống. Các yêu cầu này có thể thay đổi trong quá trình thiết kế cũng như nâng cấp hệ thống trong quá trình hoạt động thực tế sau này.

### 5.3.2 THIẾT KẾ HỆ THỐNG

#### a) Kiến trúc phân cứng của hệ thống y tế từ xa

Hình 5.32. Hình 1 minh họa một hệ thống y tế từ xa được đề xuất để phân loại hình ảnh UTV 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 PACS; 2. Máy chủ web để kết nối dữ liệu với mô hình AI và PACS; 3. Mô hình AI để phân loại hình ảnh.

## 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



**Hình 5.32 Minh họa hệ thống y tế từ xa để phân loại hình ảnh UTV sử dụng mô hình AI.**

Do đó, hệ thống y tế từ xa này kết nối với các thiết bị ngoại vi như thiết bị chụp X-quang, thiết bị xem hình ảnh và thiết bị đầu cuối để truy cập từ xa. Chi tiết từng thành phần trong hệ thống được mô tả như sau:

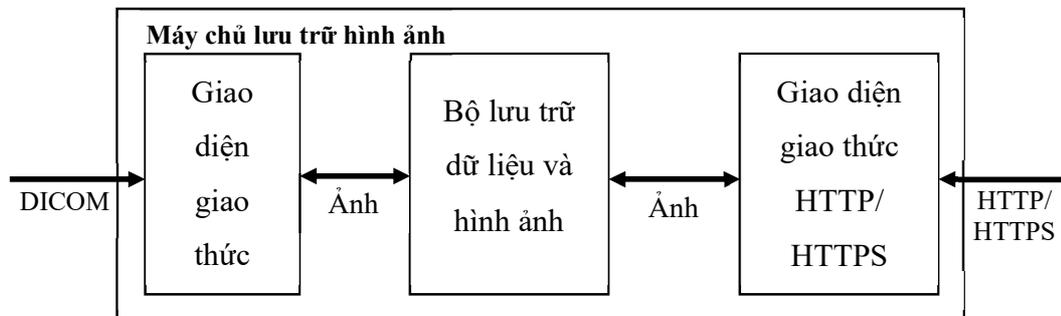
- Máy chủ PACS có thể tiếp nhận và lưu trữ hình ảnh, đặc biệt là hình ảnh được tải lên từ các thiết bị hình ảnh y tế theo chuẩn DICOM. Hơn nữa, hệ thống cho phép tải lên hình ảnh ở định dạng JPEG để lưu trữ và xử lý. Máy chủ này cũng giao tiếp với các thiết bị đầu cuối, các thành phần khác trong hệ thống thông qua các giao thức và API.
- Máy chủ web: cung cấp dịch vụ web để tạo ra giao diện truy cập cho người dùng. Giao diện web được xây dựng bằng công nghệ web hiện đại cung cấp khả năng thích ứng cao đối với nhiều loại thiết bị đầu cuối khác nhau.
- Mô hình AI được áp dụng mô hình EfficientNet-B7 đã được đào tạo để phân loại ảnh chụp nhũ ảnh. Cụ thể, khi nhận được yêu cầu từ máy chủ web, mô hình AI sẽ thực hiện quá trình phân loại và sau đó trả về kết quả đã phân loại thông qua các giao thức nội bộ của hệ thống.
- Các thiết bị ngoại vi được kết nối với hệ thống trên bao gồm: DICOM Modalities là các thiết bị chụp X-quang vù tương thích với chuẩn DICOM; Máy trạm hình ảnh để BS xem hình ảnh; người dùng từ xa như BS hoặc bệnh nhân có thể truy cập từ xa thông qua website.

## 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

Các thành phần của hệ thống trên kết nối với nhau thông qua các giao thức tiêu chuẩn như TCP/IP, DICOM, HTTP/HTTPS, và các API nội bộ được xây dựng riêng để cung cấp các dịch vụ được truy vấn giữa các thành phần chính trong hệ thống.

### b) Máy chủ lưu trữ hình ảnh

Để lưu trữ hình ảnh và tiếp nhận hình ảnh được tải lên từ các cơ sở y tế, một máy chủ lưu trữ hình ảnh được xây dựng bên cạnh hệ thống xử lý và chẩn đoán dựa trên AI như được mô tả trong Hình 5.33.

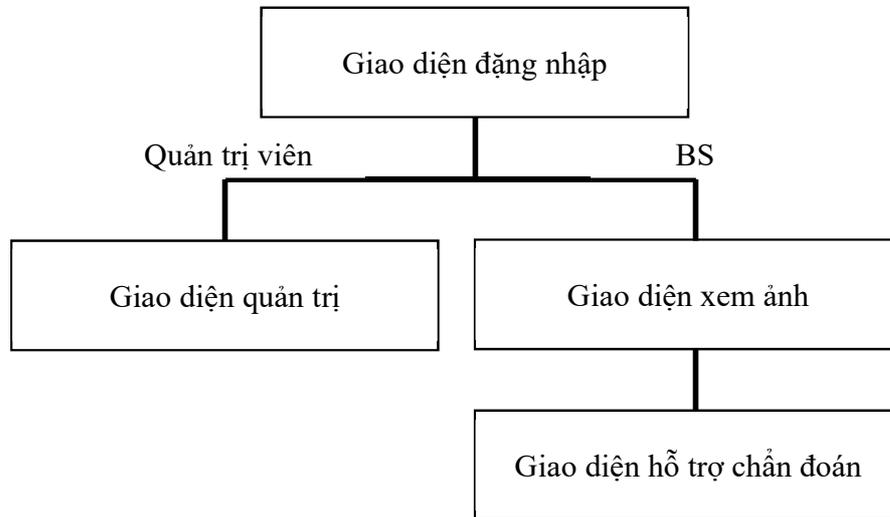


**Hình 5.33. Mô hình hệ thống máy chủ lưu trữ hình ảnh (PACS)**

Trung tâm của máy chủ này là bộ nhớ lưu trữ dữ liệu và hình ảnh. Các ảnh cần xử lý có thể được chuyển đến bộ lưu trữ thông qua giao thức DICOM (đối với ảnh DICOM) hoặc giao thức HTTP/HTTPS (đối với ảnh JPEG). Nhiệm vụ chính của server này là tiếp nhận và lưu trữ hình ảnh để xử lý ở các bước tiếp theo cũng như cho phép xem lại các hình ảnh đã lưu trữ.

### c) Cổng truy cập dựa trên nền tảng web

Thành phần tiếp theo của hệ thống là giao diện truy cập dựa trên nền tảng web. Hệ thống hỗ trợ chẩn đoán UTV thông qua ảnh X-quang ví dụ có giao diện người dùng được thiết kế dựa trên nền tảng web để phù hợp với việc triển khai trực tuyến và cho phép truy cập từ xa. Các trang giao diện dự kiến được thể hiện ở Hình 5.34 dưới đây.



**Hình 5.34. Sơ đồ các giao diện phần mềm chính của hệ thống**

Các giao diện hệ thống như trong Hình 5.34 sẽ được thiết kế trên nền tảng web html, sử dụng ngôn ngữ PHP hoặc python (cho phép tương thích tốt với hệ thống AI). Các kỹ thuật mới nhất như kỹ thuật thiết kế web thích ứng (tương thích với nhiều kích thước hiển thị) sẽ được ứng dụng.

Như đã trình bày ở trên, phần giao diện web sẽ bao gồm 2 thành phần chính là Frontend và Backend. Frontend trong website là một phần giao diện giúp người dùng có thể trực tiếp tương tác với hệ thống. Nó giúp cho người dùng có một trải nghiệm thú vị khi truy cập vào Website bao gồm giao diện, hình ảnh, nút bấm, ... Fontend được phát triển bằng ngôn ngữ lập trình HTML, CSS và Javascript, ngoài ra Fontend còn có chức năng lấy dữ liệu bằng cách thu thập tác động tương tác của người dùng và đưa ra tín hiệu để Backend xử lý theo bước lập trình định sẵn.

Phần Frontend của hệ thống được thiết kế trực quan và tương thích trên nhiều nền tảng. Với yêu cầu của hệ thống, triển khai một giao diện người dùng cần có những thuộc tính sau:

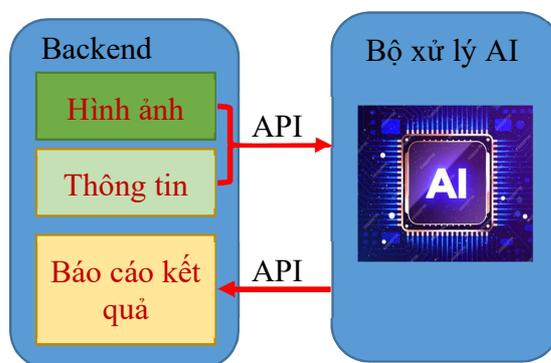
- Tính thẩm mỹ và dễ sử dụng: Giao diện ứng dụng dành cho người dùng cần phải trực quan, bố cục hợp lý đồng thời đảm bảo việc đơn giản khi thực hiện các thao tác trong quá trình sử dụng. Các thành phần của giao diện, nút nhấn,

## 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

label, v.v. nên được sắp xếp một cách logic, rõ ràng. Bên cạnh đó phải đảm bảo màu sắc, kích thước của các đơn vị được hiển thị một cách hài hoà và dễ nhìn. Giao diện cũng cần có khả năng đáp ứng để hiển thị lên các thiết bị, màn hình khác nhau.

- Thông qua API để tương tác với Backend: Giao diện cần thông qua các API để tương tác với Backend trong việc gửi và nhận yêu cầu từ người dùng. Các yêu cầu được sử dụng bao gồm việc gửi hình ảnh X quang và nhận kết quả trả về. Giao diện cần triển khai chức năng gọi API tương ứng và xử lý phản hồi mà Backend trả về.
- Sử dụng Firebase – Storage trong truyền dữ liệu: Giao diện ứng dụng cần cung cấp dữ liệu dữ liệu để hệ thống có thể xử lý để từ đó trả về kết quả. Trong ứng dụng này, dữ liệu cần cung cấp là các file ảnh mà người dùng cần chẩn đoán, chúng được lưu trữ ở Storage trong Firebase, để từ đó thông qua API Backend có thể lấy về xử lý.

Backend là phần ứng dụng của Website, đảm bảo hoạt động trơn tru và xử lý dữ liệu được truyền từ FrontEnd thông qua tương tác của người dùng. Backend được xây dựng bằng các ngôn ngữ lập trình (như Node.js, Python, Java, PHP) là phần mà người dùng không nhìn thấy và chịu trách nhiệm xử lý dữ liệu, CSDL, logic xử lý của Website.



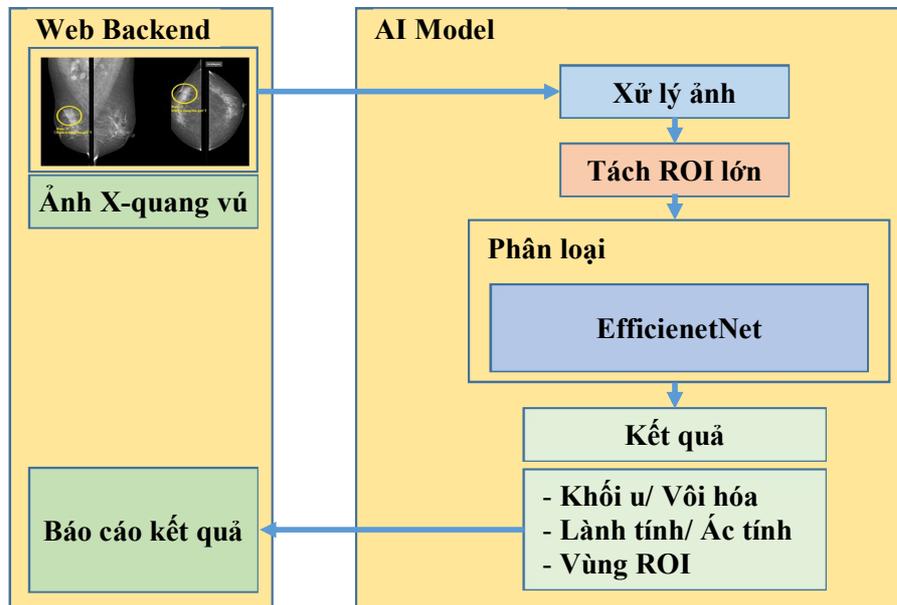
Hình 5.35. Sơ đồ kết nối Backend và Bộ xử lý AI

## 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

Sơ đồ ở Hình 5.35 thể hiện nguyên tắc kết nối giữa Backend và Module AI để thực hiện các nhiệm vụ phân loại ảnh. Trong hệ thống này, Backend đóng vai trò quan trọng trong việc thực hiện các chức năng chính của hệ thống. Backend sẽ được xây dựng bằng việc sử dụng Pycharm và các API, kết nối với Firebase, sử dụng module để tạo nên chương trình thực hiện các chức năng theo yêu cầu của đề tài. Ngoài ra, với việc backend là nơi thực hiện các yêu cầu từ người dùng. Yêu cầu backend cũng được chạy đồng bộ với frontend.

### d) Bộ xử lý thuật toán AI

Bộ xử lý AI chính là thành phần thực thi các thuật toán AI trên các ảnh nhận được từ Backend của web chuyển sang để thực hiện việc phân loại và trả kết quả lại cho Backend để từ đó tạo ra các report cuối cùng cho người dùng. Mô hình của thành phần này được mô tả ở Hình 5.36.



**Hình 5.36. Mô hình AI trong hệ thống**

Hình 5.36 cho thấy cấu trúc cụ thể của mô hình AI và mối liên hệ của nó với các thành phần khác trong hệ thống. Nhiệm vụ chính của nó là thực thi các thuật toán AI trên các ảnh được phân giao diện web gửi sang và trả kết quả lại sau khi chạy các

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

---

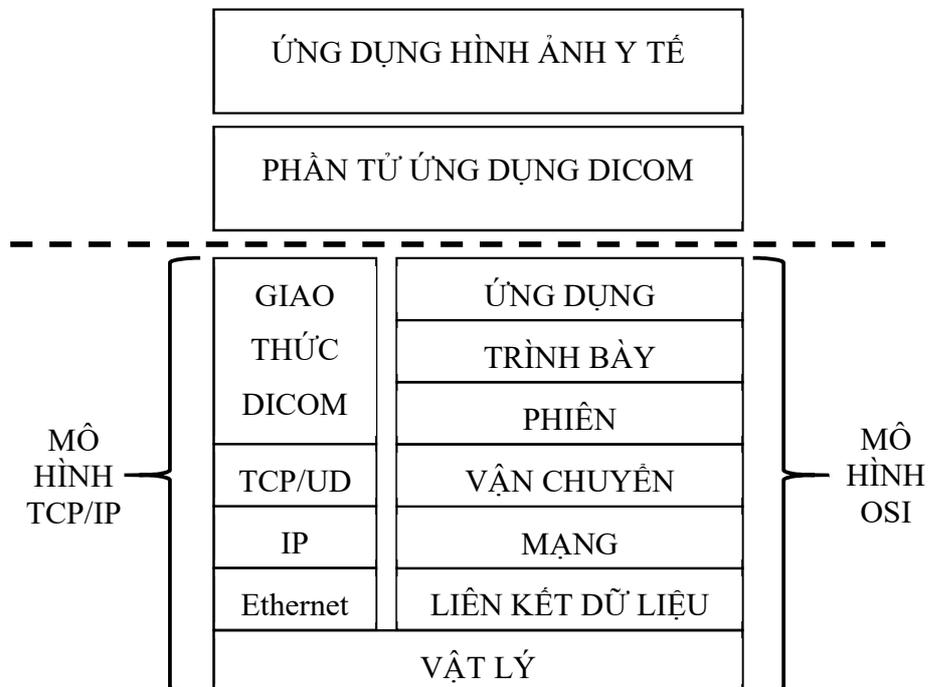
thuật toán tương ứng.

**e) Các giao thức và API trao đổi dữ liệu**

Để các thành phần trong hệ thống trên có thể kết nối, trao đổi thông tin và làm việc được với nhau, cũng như các ngoại vi khác và thiết bị đầu cuối của người dùng, cần có các giao thức và các API được thiết lập cho các mục đích này. Trong hệ thống sử dụng cả các giao thức tiêu chuẩn để tăng khả năng tương thích của hệ thống, cũng như các API được thiết kế riêng cho các chức năng cụ thể.

**f) Giao thức truyền tải hình ảnh y tế DICOM**

Giao thức truyền dữ liệu là hình ảnh y tế được sử dụng trong các hệ thống thông tin y tế là DICOM. Đây cũng là giao thức để lưu trữ, in ấn, xử lý và truyền tải hình ảnh y tế. Giao thức này bao gồm cấu trúc tập tin và phương pháp truyền các bản tin dựa trên nền tảng TCP/IP.



**Hình 5.37. Mô hình giao thức DICOM**

Hình 5.36 mô tả các lớp của giao thức DICOM cũng như mối liên hệ giữa chúng với các giao thức khác trong hệ thống.

## 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

---

Giao thức DICOM cho phép dễ dàng truyền tải hình ảnh từ các thiết bị thu nhận hình ảnh y tế lên các server lưu trữ và xử lý hình ảnh, và cho phép tích hợp các trạm làm việc, máy in ảnh và các thiết bị tương thích từ các nhà sản xuất khác nhau vào trong hệ thống hình ảnh y tế. Hình dưới đây mô tả giao thức DICOM.

Trong nội dung của đề tài, do hình ảnh cần truyền tải và lưu trữ là ảnh X-quang vú, thuộc lớp hình ảnh DICOM, nên giao thức DICOM là phù hợp với hệ thống này và cũng là tiêu chuẩn được yêu cầu để đáp ứng các qui định về hình ảnh y tế như mô tả trong Hình 5.36.

Chương 5 đã đi sâu vào thực hiện bài toán phân loại với các mô hình mạng học sâu tiên tiến. Từ các kết quả thực nghiệm trong Chương 5 cho thấy EfficientNet-B7 đạt độ chính xác cao hơn VGG16 một cách ổn định trên nhiều tập dữ liệu (MIAS, CBIS-DDSM, VinDr-Mammo), đặc biệt trong các bài toán phân loại tổn thương phức tạp và đánh giá mức độ BIRADS. Điều này có thể được giải thích từ cả khía cạnh kiến trúc mạng lẫn đặc thù của ảnh X-quang vú.

Thứ nhất, EfficientNet-B7 sử dụng chiến lược compound scaling, đồng thời mở rộng độ sâu (depth), độ rộng (width) và độ phân giải (resolution) của mạng theo một tỉ lệ tối ưu. Nhờ đó, mô hình có khả năng học các đặc trưng đa cấp (multi-scale features) hiệu quả hơn so với VGG16 – vốn chỉ mở rộng mạng theo chiều sâu với các khối tích chập  $3 \times 3$  lặp lại. Trong ảnh X-quang vú, tổn thương có thể xuất hiện ở nhiều kích thước khác nhau, từ vi vôi hóa rất nhỏ đến khối u có biên mờ, nên khả năng học đặc trưng đa tỉ lệ là yếu tố then chốt giúp EfficientNet vượt trội.

Thứ hai, EfficientNet-B7 tích hợp các khối MBConv (Mobile Inverted Bottleneck Convolution) cùng cơ chế Squeeze-and-Excitation (SE), cho phép mô hình tự động tái cân bằng tầm quan trọng giữa các kênh đặc trưng. Điều này đặc biệt phù hợp với ảnh X-quang vú, nơi các dấu hiệu bệnh lý thường chỉ chiếm một vùng nhỏ và có độ tương phản thấp so với mô nền. Ngược lại, VGG16 xử lý các kênh đặc trưng một cách đồng đều, dễ dẫn đến việc mô hình bị chi phối bởi các vùng nền lớn không mang nhiều thông tin chẩn đoán.

## 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

---

Thứ ba, EfficientNet-B7 cho thấy khả năng tổng quát hóa tốt hơn khi huấn luyện trên các tập dữ liệu dị nguồn (đa bệnh viện, đa thiết bị chụp). Điều này xuất phát từ việc kiến trúc EfficientNet được thiết kế tối ưu hóa giữa hiệu năng và số lượng tham số, giúp mô hình giảm nguy cơ overfitting so với VGG16 – một kiến trúc cũ, có số lượng tham số lớn nhưng khả năng biểu diễn đặc trưng không còn phù hợp với các bài toán y sinh phức tạp hiện nay.

Trong các thí nghiệm phân loại đa ảnh (multi-view) sử dụng hai tư thế chụp CC và MLO, kết quả cho thấy mô hình kết hợp EfficientNet-B7 + Multi-head Attention cho độ chính xác cao hơn rõ rệt so với phương pháp cộng hoặc trung bình đặc trưng từ hai ảnh. Hiệu quả này có thể được giải thích dựa trên bản chất của dữ liệu đa góc nhìn và cơ chế học của Attention.

Trước hết, trong thực tế lâm sàng, mỗi tư thế chụp (CC, MLO) cung cấp thông tin bổ sung chứ không hoàn toàn tương đương. Có những tổn thương chỉ biểu hiện rõ ở một tư thế nhất định do chồng lấp mô tuyến hoặc hướng chiếu tia X. Phương pháp cộng/trung bình đặc trưng giả định rằng mọi góc nhìn đều có mức độ đóng góp như nhau, từ đó có thể làm suy giảm thông tin quan trọng nếu một tư thế chứa dấu hiệu bệnh lý rõ ràng hơn tư thế còn lại.

Ngược lại, Multi-head Attention cho phép mô hình học trọng số động cho từng đặc trưng và từng góc nhìn. Cơ chế này giúp mạng tập trung mạnh hơn vào những vùng và những tư thế chụp mang nhiều thông tin chẩn đoán, đồng thời giảm ảnh hưởng của các đặc trưng nhiễu hoặc ít giá trị. Về mặt nhận thức, cách tiếp cận này mô phỏng quá trình đọc ảnh của bác sĩ X-quang: bác sĩ không đánh giá các ảnh một cách máy móc mà luôn ưu tiên những góc nhìn thể hiện tổn thương rõ nhất.

Ngoài ra, Multi-head Attention còn giúp mô hình khai thác mối quan hệ tương hỗ giữa các đặc trưng không gian và ngữ cảnh từ nhiều ảnh đầu vào. Điều này đặc biệt quan trọng đối với các dạng tổn thương khó phân biệt như bất đối xứng cục bộ hoặc biến dạng cấu trúc, vốn đòi hỏi phải so sánh thông tin giữa nhiều góc chụp để đưa ra kết luận chính xác.

## 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

---

Do đó, kết quả thực nghiệm cho thấy việc sử dụng Multi-view Attention kết hợp với EfficientNet không chỉ cải thiện độ chính xác tổng thể mà còn giúp mô hình ổn định hơn trước các trường hợp khó, chứng minh tính hợp lý của kiến trúc đề xuất trong luận án.

Đối với hệ thống chẩn đoán ung thư vú từ xa dựa trên AI, trong thực tế tại Việt Nam, hệ thống y tế đang đối mặt với nhiều thách thức trong công tác sàng lọc và chẩn đoán ung thư vú, bao gồm: sự quá tải tại các bệnh viện tuyến trung ương, sự thiếu hụt bác sĩ chẩn đoán hình ảnh tại tuyến dưới, và sự chênh lệch về khả năng tiếp cận dịch vụ y tế giữa khu vực thành thị và vùng sâu, vùng xa. Trong bối cảnh đó, một hệ thống chẩn đoán từ xa tích hợp trí tuệ nhân tạo có khả năng hỗ trợ bác sĩ là một giải pháp mang tính thực tiễn cao và phù hợp với định hướng chuyển đổi số ngành y tế.

Hệ thống được đề xuất trong luận án không nhằm thay thế bác sĩ, mà đóng vai trò như một công cụ hỗ trợ quyết định lâm sàng (Clinical Decision Support System – CDSS), giúp nâng cao độ chính xác, rút ngắn thời gian đọc ảnh và giảm tải công việc cho đội ngũ y tế. Một ưu điểm quan trọng của hệ thống là khả năng tích hợp linh hoạt với hạ tầng y tế hiện hành, đặc biệt là:

Hỗ trợ chuẩn DICOM: hệ thống cho phép tiếp nhận trực tiếp ảnh X-quang vú từ các thiết bị chụp hiện có tại bệnh viện, đảm bảo không làm thay đổi quy trình kỹ thuật chẩn đoán đang được sử dụng.

Tương thích với hệ thống PACS: dữ liệu ảnh được lưu trữ và quản lý tập trung, giúp bác sĩ dễ dàng truy xuất, so sánh và theo dõi lịch sử bệnh nhân.

Giao diện web-based: cho phép truy cập từ nhiều thiết bị (máy tính, máy tính bảng) mà không yêu cầu cài đặt phần mềm phức tạp, phù hợp với điều kiện kỹ thuật tại các bệnh viện tuyến dưới.

Nhờ đó, hệ thống có thể được triển khai theo mô hình từng bước, không đòi hỏi đầu tư lớn ban đầu và không gây gián đoạn hoạt động chuyên môn của bệnh viện.

Một trong những giá trị thực tiễn nổi bật của hệ thống là khả năng hỗ trợ y tế từ xa cho các cơ sở thiếu bác sĩ chuyên khoa: Các bệnh viện tuyến huyện hoặc trung tâm

## 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

---

y tế khu vực có thể tải ảnh X-quang vù lên hệ thống; Hệ thống AI phân tích và trả kết quả sơ bộ trong thời gian ngắn; Bác sĩ tuyến trên có thể truy cập dữ liệu từ xa để hội chẩn khi cần thiết. Cách tiếp cận này giúp thu hẹp khoảng cách chất lượng chẩn đoán giữa các tuyến y tế, đồng thời giảm số ca phải chuyển tuyến không cần thiết, tiết kiệm chi phí và thời gian cho bệnh nhân.

Về mặt ứng dụng, hệ thống được thiết kế với các nguyên tắc đảm bảo an toàn và độ tin cậy vì kết quả AI mang tính hỗ trợ, không tự động đưa ra chẩn đoán cuối cùng, giúp tránh rủi ro pháp lý và y khoa. Những yếu tố này giúp hệ thống có thể được sử dụng trong môi trường lâm sàng một cách thận trọng và có kiểm soát.

Hệ thống chẩn đoán ung thư vú từ xa dựa trên AI được đề xuất trong luận án thể hiện khả năng áp dụng thực tế cao nhờ tính linh hoạt, khả năng tích hợp với hạ tầng hiện có và vai trò hỗ trợ hiệu quả cho bác sĩ trong quy trình chẩn đoán. Việc triển khai hệ thống này có tiềm năng góp phần nâng cao chất lượng sàng lọc ung thư vú, giảm tải cho tuyến trên và cải thiện khả năng tiếp cận y tế cho người dân, đặc biệt tại các khu vực còn hạn chế về nguồn lực chuyên môn.

## Chương 6

# KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

### 6.1. KẾT LUẬN

Luận án đã hoàn thành mục tiêu xây dựng một hệ thống chẩn đoán bệnh ung thư vú dựa trên trí tuệ nhân tạo, kết hợp giữa tiền xử lý ảnh X-quang, trích xuất vùng ROI và ứng dụng các mô hình học sâu hiện đại. Các đóng góp của luận án được triển khai xuyên suốt từ việc chuẩn hóa dữ liệu, tối ưu hóa khả năng nhận dạng tổn thương đến việc xây dựng một hệ thống chẩn đoán từ xa dựa trên AI có khả năng ứng dụng thực tiễn tại các bệnh viện.

Trước hết, luận án đã xây dựng được quy trình tiền xử lý ảnh hoàn chỉnh, bao gồm tăng cường ảnh, lọc nhiễu, hiệu chỉnh mức xám, phân đoạn bằng ngưỡng Otsu và chuẩn hóa kích thước vùng ROI. Đặc biệt, kết quả cho thấy việc tô màu ROI (CROI) đã giúp cải thiện hiệu suất mô hình từ 3–5% so với ROI xám – đây là phát hiện có giá trị thực tiễn cao trong bối cảnh ảnh X-quang vú thường có độ tương phản thấp.

Tiếp theo, luận án đã phân tích sâu các đặc trưng (GLCM), histogram mức xám và đặc trưng hình dạng, từ đó minh chứng vai trò quan trọng của các đặc trưng này trong phân loại ung thư vú. Cụ thể, thuật toán ma trận đồng mức xám (GLCM) cho lựa chọn các đặc trưng tối ưu để thực hiện phân loại ảnh UTV sử dụng bộ phân loại SVM cũng đã được thực hiện [141]. Luận án còn kiến nghị sử dụng thuật toán Otsu cho tìm ngưỡng tối ưu để trích vùng ROI trên ảnh X-quang UTV. Các kết quả ban đầu cho thấy hiệu quả của thuật toán trong việc phân biệt giữa vùng có tổn thương và không có tổn thương [77]. Các kết quả này đã được công bố trong các bài báo khoa học liên quan đến luận án [77], [141]. Các bộ phân loại truyền thống như SVM và kNN đã được thử nghiệm để kiểm chứng giả thuyết, cung cấp nhận định rõ ràng về mức độ ảnh hưởng của từng nhóm đặc trưng.

## Chương 6 KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

---

Đóng góp quan trọng nhất của luận án nằm ở việc tinh chỉnh và huấn luyện các mô hình học sâu như VGG16, EfficientNet-B0/B7 trên các tập dữ liệu chuẩn hóa. Kết quả nghiên cứu cho thấy các bước tiền XLA đóng vai trò quan trọng trong việc cải thiện hiệu suất và độ chính xác của các thuật toán học sâu. Đặc biệt, với phương pháp tách vùng GROI và tô màu cho ảnh xám X-quang vú (CROI) hiệu suất của việc phân loại đã tăng đáng kể với độ chính xác đạt gần 95% đối với phân loại tổn thương và đạt khoảng 92.6% đối với phân loại đánh giá nguy cơ ung thư theo thang đo BIRADS (từ BIRADS 1 đến BIRADS 5). Việc áp dụng các thuật toán trích các ROI và loại bỏ các thành phần không mong muốn, cũng như tăng cường tập ảnh giúp cải thiện độ chính xác lên tới 6-8%. Việc tô màu ảnh (CROI) giúp cải thiện thêm 5-8% độ chính xác, tùy thuộc từng tập ảnh. Hơn nữa, việc cải thiện các thông số trong các mạng học sâu như VGG16, EfficientNet và một số mạng khác đã cải thiện đáng kể hiệu quả của việc phân loại. Các kết quả trên cũng chỉ ra rằng các bước tiền XLA X-quang vú trước khi đưa vào huấn luyện và phân loại đóng vai trò hết sức quan trọng trong việc cải thiện độ chính xác, độ tin cậy và hiệu quả của thuật toán mạng học sâu được triển khai sau đó. Các kết quả này cũng đã được trình bày trong 2 bài báo khoa học [142] [171].

Ngoài việc xây dựng mô hình, luận án còn phát triển hệ thống chẩn đoán từ xa gồm PACS, mô-đun AI và giao diện web, cho phép bác sĩ tải ảnh DICOM/JPEG từ xa và nhận kết quả phân loại tự động. Đây là một đóng góp có giá trị ứng dụng cao, đặc biệt đối với các cơ sở y tế thiếu bác sĩ X-quang hoặc ở vùng nông thôn.

Kết quả nghiên cứu cũng cho thấy các yếu tố ảnh hưởng đến hiệu suất phân loại bao gồm chất lượng hình ảnh, số lượng ảnh sử dụng để huấn luyện, các loại tổn thương khác nhau cũng có độ chính xác phân loại khác nhau tùy thuộc vào tính chất và mức độ thể hiện của tổn thương trên hình ảnh. Bên cạnh đó, các phương pháp tiền xử lý ảnh đóng vai trò đáng kể trong việc nâng cao hiệu suất phân loại của hệ thống.

So với các kết quả nghiên cứu gần đây, luận án đã đạt được các kết quả đáng kể trên tập dữ liệu nghiên cứu. Do sự khác biệt về tập dữ liệu và mục tiêu phân loại ngớ ra, việc so sánh trực tiếp độ chính xác với các kết quả nghiên cứu khác là không hợp

## Chương 6 KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

---

lý, tuy nhiên kết quả nhìn chung cho thấy sự đóng góp của luận án vào nghiên cứu phân loại ảnh UTV, đúng với mục tiêu đề ra của luận án.

Cuối cùng, việc sử dụng dữ liệu thực từ Bệnh viện Ung Bướu TP.HCM để huấn luyện và đánh giá mô hình đã nâng cao tính ứng dụng thực tiễn của hệ thống, đảm bảo mô hình phù hợp với dữ liệu và điều kiện khám chữa bệnh tại Việt Nam. Các quy trình chặt chẽ cho việc thu thập, chọn lọc, xử lý và gán nhãn ảnh cũng đã được xây dựng cho mục đích này. Tập ảnh này dự kiến sẽ được sử dụng và công bố trong một bài báo khoa học tiếp theo.

Luận án đã chứng minh tính khả thi của việc triển khai AI vào chẩn đoán hình ảnh vú, góp phần quan trọng vào việc hỗ trợ phát hiện sớm ung thư vú và giảm tải cho bác sĩ X-quang. Các kết quả này góp phần vào các kết quả nghiên cứu trong chẩn đoán và phát hiện sớm bệnh UTV, cải thiện hiệu suất và độ chính xác chẩn đoán của BS, cũng như có khả năng áp dụng cho các hệ thống y tế từ xa.

### 6.2. HƯỚNG PHÁT TRIỂN

Trong phạm vi mục tiêu ban đầu của luận án, một số khía cạnh vẫn còn chưa được xem xét hết như việc ứng dụng các mạng học sâu trong phân đoạn ảnh UTV, kết hợp các kết quả chẩn đoán lâm sàng khác và thông tin bệnh nhân trong chẩn đoán. Bên cạnh đó một số loại tổn thương hiếm gặp chưa được phân loại do không có đủ số lượng ảnh cần thiết, cũng như việc giải quyết bài toán mất cân bằng dữ liệu. Các nghiên cứu sau này của NCS sau khi hoàn thành LATS có thể tiếp tục khám phá thêm các hướng này.

Hướng phát triển của luận án có thể bao gồm việc tiếp tục đào sâu nghiên cứu cải tiến các thuật toán và mô hình mạng học sâu đã có, cũng như xem xét việc ứng dụng các mạng học sâu khác vào bài toán phân loại ảnh UTV. Tập dữ liệu từ BV Ung Bướu (và có thể các BV khác) cũng sẽ được thu thập hoàn thiện để công bố và sử dụng trong các nghiên cứu tiếp theo.

## TÀI LIỆU THAM KHẢO

- [1] World Health Organization, “Vietnam Source: Globocan 2022,” 2022. [Online]. Available: <https://gco.iarc.who.int/media/globocan/factsheets/populations/704-vietnam-fact-sheet.pdf>.
- [2] Vietnamnews, “Bach Mai Hospital leads digital transformation in healthcare,” *Vietnamnews*, 2025. <https://vietnamnews.vn/opinion/1693875/bach-mai-hospital-leads-digital-transformation-in-healthcare.html> (accessed Jun. 18, 2025).
- [3] R. L. Siegel, A. N. Giaquinto, and A. Jemal, “Cancer statistics, 2024,” *CA. Cancer J. Clin.*, vol. 74, no. 1, pp. 12–49, 2024, doi: <https://doi.org/10.3322/caac.21820>.
- [4] G. A. Colditz, K. A. Kaphingst, S. E. Hankinson, and B. Rosner, “Family history and risk of breast cancer: nurses’ health study,” *Breast Cancer Res. Treat.*, vol. 133, no. 3, pp. 1097–1104, 2012, doi: [10.1007/s10549-012-1985-9](https://doi.org/10.1007/s10549-012-1985-9).
- [5] S. Sieri *et al.*, “Sex Hormone Levels, Breast Cancer Risk, and Cancer Receptor Status in Postmenopausal Women: the ORDET Cohort,” *Cancer Epidemiol. Biomarkers Prev.*, vol. 18, no. 1, pp. 169–176, Jan. 2009, doi: [10.1158/1055-9965.EPI-08-0808](https://doi.org/10.1158/1055-9965.EPI-08-0808).
- [6] B. Smolarz, A. Z. Nowak, and H. Romanowicz, “Breast Cancer—Epidemiology, Classification, Pathogenesis and Treatment (Review of Literature),” *Cancers (Basel)*, vol. 14, no. 10, 2022, doi: [10.3390/cancers14102569](https://doi.org/10.3390/cancers14102569).
- [7] N. Harbeck *et al.*, “Breast cancer,” *Nat. Rev. Dis. Prim.*, vol. 5, no. 1, p. 66, 2019, doi: [10.1038/s41572-019-0111-2](https://doi.org/10.1038/s41572-019-0111-2).
- [8] G. Beral, V.; Bull, D.; Doll, R.; Peto, R.; Reeves, “Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease.,” *Lancet (London, England)*, vol. 360, no. 9328, pp. 187–195, Jul. 2002, doi: [10.1016/S0140-6736\(02\)09454-0](https://doi.org/10.1016/S0140-6736(02)09454-0).
- [9] T. Mahmood, J. Li, Y. Pei, F. Akhtar, A. Imran, and K. U. Rehman, “A Brief Survey on Breast Cancer Diagnostic With Deep Learning Schemes Using Multi-Image Modalities,” *IEEE Access*, vol. 8, pp. 165779–165809, 2020, doi: [10.1109/ACCESS.2020.3021343](https://doi.org/10.1109/ACCESS.2020.3021343).
- [10] J.-Y. Chiao, K.-Y. Chen, K. Y.-K. Liao, P.-H. Hsieh, G. Zhang, and T.-C. Huang, “Detection and classification the breast tumors using mask R-CNN on sonograms,” *Medicine (Baltimore)*, vol. 98, no. 19, 2019, doi: [10.1097/MD.00000000000015200](https://doi.org/10.1097/MD.00000000000015200).
- [11] J. Yan, Z. Liu, S. Du, J. Li, L. Ma, and L. Li, “Diagnosis and Treatment of Breast Cancer in the Precision Medicine Era BT - Precision Medicine,” T. Huang, Ed. New York, NY: Springer US, 2020, pp. 53–61.
- [12] “Breast Cancer,” 2022. <https://www.who.int/news-room/fact-sheets/detail/breast-cancer> (accessed Aug. 08, 2024).
- [13] M. Kashif, K. R. Malik, S. Jabbar, and J. Chaudhry, “Application of machine learning and image processing for detection of breast cancer,” in *Next Gen Tech Driven Personalized Med&Smart Healthcare*, M. D. Lytras and A. B. T.-I. in H. I. Sarirete, Eds. Academic Press, 2020, pp. 145–162.

## TÀI LIỆU THAM KHẢO

---

- [14] K. N. Ramanto and A. A. Parikesit, “The usage of deep learning algorithm in medical diagnostic of breast cancer,” *Malaysian J. Fundam. Appl. Sci.*, vol. 15, no. 2, 2019, doi: <https://doi.org/10.11113/mjfas.v15n2.1231>.
- [15] Y. Deng *et al.*, “A New Framework to Reduce Doctor’s Workload for Medical Image Annotation,” *IEEE Access*, vol. 7, pp. 107097–107104, 2019, doi: [10.1109/ACCESS.2019.2917932](https://doi.org/10.1109/ACCESS.2019.2917932).
- [16] N. Ngọc Trung, “Updating and upgrading the features of artificial intelligence (AI) applications in early diagnosis of breast cancer with Korean experts at Thai Binh University of Medicine and Pharmacy Hospital,” *Thai Binh University of Medicine and Pharmacy web Portal*, 2025. <https://tbump.edu.vn/tin-hoat-dong/cap-nhat-nang-cap-tinh-nang-cua-ung-dung-tri-tue-nhan-tao-ai-trong-chan-doan-som-ung-thu-vu-cung-cac-chuyen-gia-han-quoc-tai-benh-vien-dai-hoc-y-thai-binh-298.html> (accessed Jun. 16, 2025).
- [17] S. A. A. Karim, U. H. Mohamad, and N. E. N. Puteri, “Discovery of Interpretable Patterns of Breast Cancer Diagnosis via Class Association Rule Mining (CARM) With SHAP-Based Explainable AI (XAI),” *Malaysian J. Fundam. Appl. Sci.*, vol. 21, no. 3, 2025, doi: <https://doi.org/10.11113/mjfas.v21n3.3792>.
- [18] K. K. Evans, R. L. Birdwell, and J. M. Wolfe, “If you don’t find it often, you often don’t find it: why some cancers are missed in breast cancer screening,” *PLoS One*, vol. 8, no. 5, p. e64366, 2013, doi: [10.1371/journal.pone.0064366](https://doi.org/10.1371/journal.pone.0064366).
- [19] M. Issaiy, D. Zarei, and A. Saghazadeh, “Artificial Intelligence and Acute Appendicitis: A Systematic Review of Diagnostic and Prognostic Models,” *World J. Emerg. Surg.*, vol. 18, no. 1, p. 59, Dec. 2023, doi: [10.1186/s13017-023-00527-2](https://doi.org/10.1186/s13017-023-00527-2).
- [20] A. Rodriguez-Ruiz *et al.*, “Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists,” *JNCI J. Natl. Cancer Inst.*, vol. 111, no. 9, pp. 916–922, 2019, doi: [10.1093/jnci/djy222](https://doi.org/10.1093/jnci/djy222).
- [21] M. A. Aswathy and M. Jagannath, “Detection of breast cancer on digital histopathology images: Present status and future possibilities,” *Informatics Med. Unlocked*, vol. 8, no. November 2016, pp. 74–79, 2017, doi: [10.1016/j.imu.2016.11.001](https://doi.org/10.1016/j.imu.2016.11.001).
- [22] V. Nemade, S. Pathak, and A. K. Dubey, “A Systematic Literature Review of Breast Cancer Diagnosis Using Machine Intelligence Techniques,” *Arch. Comput. Methods Eng.*, vol. 29, no. 6, pp. 4401–4430, 2022, doi: [10.1007/s11831-022-09738-3](https://doi.org/10.1007/s11831-022-09738-3).
- [23] Z. Jafari and E. Karami, “Breast Cancer Detection in Mammography Images: A CNN-Based Approach with Feature Selection,” *Information*, vol. 14, no. 7, 2023, doi: [10.3390/info14070410](https://doi.org/10.3390/info14070410).
- [24] D. Ribli, A. Horváth, Z. Unger, P. Pollner, and I. Csabai, “Detecting and classifying lesions in mammograms with Deep Learning,” *Sci. Rep.*, vol. 8, no. 1, p. 4165, 2018, doi: [10.1038/s41598-018-22437-z](https://doi.org/10.1038/s41598-018-22437-z).
- [25] F. F. Ting, Y. J. Tan, and K. S. Sim, “Convolutional neural network improvement for breast cancer classification,” *Expert Syst. Appl.*, vol. 120, pp. 103–115, 2019, doi: [10.1016/j.eswa.2018.11.008](https://doi.org/10.1016/j.eswa.2018.11.008).
- [26] C. Gupta, N. S. Gill, P. Gulia, N. Alduaiji, J. Shreyas, and P. K. Shukla, “Applying YOLOv6 as an ensemble federated learning framework to classify breast cancer

- pathology images.,” *Sci. Rep.*, vol. 15, no. 1, p. 3769, Jan. 2025, doi: 10.1038/s41598-024-80187-7.
- [27] A. Kapoor, P. Nambisan, and E. Baker, “Mobile applications for breast cancer survivorship and self-management: A systematic review,” *Health Informatics J.*, vol. 26, no. 4, pp. 2892–2905, 2020, doi: 10.1177/1460458220950853.
- [28] J. V. Johansson and E. Engström, “‘Humans think outside the pixels’ – Radiologists’ perceptions of using artificial intelligence for breast cancer detection in mammography screening in a clinical setting,” *Health Informatics J.*, vol. 30, no. 3, p. 14604582241275020, 2024, doi: 10.1177/14604582241275020.
- [29] N. Zhou *et al.*, “Concordance Study Between IBM Watson for Oncology and Clinical Practice for Patients with Cancer in China.,” *Oncologist*, vol. 24, no. 6, pp. 812–819, Jun. 2019, doi: 10.1634/theoncologist.2018-0255.
- [30] Y. Shen *et al.*, “Artificial Intelligence System Reduces False-Positive Findings in the Interpretation of Breast Ultrasound Exams,” *medRxiv*, 2021, doi: 10.1101/2021.04.28.21256203.
- [31] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *arXiv e-prints*, p. arXiv:1409.1556, Sep. 2014, doi: 10.48550/arXiv.1409.1556.
- [32] M. Tan and Q. V Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.” 2020.
- [33] M. Arnold *et al.*, “Current and future burden of breast cancer: Global statistics for 2020 and 2040.,” *Breast*, vol. 66, pp. 15–23, Dec. 2022, doi: 10.1016/j.breast.2022.08.010.
- [34] J. D. Fackenthal and O. I. Olopade, “Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations,” *Nature Reviews Cancer*, vol. 7, no. 12, pp. 937–948, 2007, doi: 10.1038/nrc2054.
- [35] A. El-Nabawy, N. El-Bendary, and N. A. Belal, “A feature-fusion framework of clinical, genomics, and histopathological data for METABRIC breast cancer subtype classification,” *Appl. Soft Comput.*, vol. 91, p. 106238, 2020, doi: <https://doi.org/10.1016/j.asoc.2020.106238>.
- [36] R. Aggarwal *et al.*, “Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis,” *npj Digit. Med.*, vol. 4, no. 1, p. 65, 2021, doi: 10.1038/s41746-021-00438-z.
- [37] A. B. Nassif, M. A. Talib, Q. Nasir, Y. Afadar, and O. Elgendy, “Breast cancer detection using artificial intelligence techniques: A systematic literature review,” *Artif. Intell. Med.*, vol. 127, p. 102276, May 2022, doi: 10.1016/j.artmed.2022.102276.
- [38] H. Yao, X. Zhang, X. Zhou, and S. Liu, “Parallel Structure Deep Neural Network Using CNN and RNN with an Attention Mechanism for Breast Cancer Histology Image Classification,” *Cancers (Basel)*, vol. 11, no. 12, 2019, doi: 10.3390/cancers11121901.
- [39] A. Bhushan, A. Gonsalves, and J. U. Menon, “Current State of Breast Cancer Diagnosis, Treatment, and Theranostics,” *Pharmaceutics*, vol. 13, no. 5, 2021, doi:

- 10.3390/pharmaceutics13050723.
- [40] S. J. Mambou, P. Maresova, O. Krejcar, A. Selamat, and K. Kuca, “Breast Cancer Detection Using Infrared Thermal Imaging and a Deep Learning Model,” *Sensors*, vol. 18, no. 9, 2018, doi: 10.3390/s18092799.
- [41] S. J. S. Gardezi, A. Elazab, B. Lei, and T. Wang, “Breast Cancer Detection and Diagnosis Using Mammographic Data: Systematic Review,” *J. Med. Internet Res.*, vol. 21, no. 7, pp. e14464–e14464, Jul. 2019, doi: 10.2196/14464.
- [42] K. Mohaideen, A. Negi, D. K. Verma, N. Kumar, K. Sennimalai, and A. Negi, “Applications of artificial intelligence and machine learning in orthognathic surgery: A scoping review,” *J. Stomatol. Oral Maxillofac. Surg.*, vol. 123, no. 6, pp. e962–e972, 2022, doi: <https://doi.org/10.1016/j.jormas.2022.06.027>.
- [43] A. Derevianko *et al.*, “The Use of Artificial Intelligence (AI) in the Radiology Field: What Is the State of Doctor–Patient Communication in Cancer Diagnosis?,” *Cancers (Basel)*, vol. 15, no. 2, 2023, doi: 10.3390/cancers15020470.
- [44] D. Al-karawi *et al.*, “An Evaluation of the Effectiveness of Image-based Texture Features Extracted from Static B-mode Ultrasound Images in Distinguishing between Benign and Malignant Ovarian Masses,” *Ultrason. Imaging*, vol. 43, no. 3, pp. 124–138, Feb. 2021, doi: 10.1177/0161734621998091.
- [45] D. Al-karawi *et al.*, “Prospective clinical evaluation of texture-based features analysis of ultrasound ovarian scans for distinguishing benign and malignant adnexal tumors,” *Australas. J. Ultrasound Med.*, vol. 22, no. 2, p. 144, 2019, doi: <https://doi.org/10.1002/ajum.12143>.
- [46] M. Wu, C. Yan, H. Liu, and Q. Liu, “Automatic classification of ovarian cancer types from cytological images using deep convolutional neural networks,” *Biosci. Rep.*, vol. 38, no. 3, p. BSR20180289, 2018, doi: 10.1042/BSR20180289.
- [47] S. P and T. R, “Aiding the Digital Mammogram for Detecting the Breast Cancer Using Shearlet Transform and Neural Network.,” *Asian Pac. J. Cancer Prev.*, vol. 19, no. 9, pp. 2665–2671, Sep. 2018, doi: 10.22034/APJCP.2018.19.9.2665.
- [48] K. Karacan, T. Uyar, B. Tunga, and M. A. Tunga, “A novel multistage CAD system for breast cancer diagnosis,” *Signal, Image Video Process.*, vol. 17, no. 5, pp. 2359–2368, 2023, doi: 10.1007/s11760-022-02453-3.
- [49] H. Rahman, T. F. Naik Bukht, R. Ahmad, A. Almadhor, and A. R. Javed, “Efficient Breast Cancer Diagnosis from Complex Mammographic Images Using Deep Convolutional Neural Network,” *Comput. Intell. Neurosci.*, vol. 2023, no. 1, p. 7717712, 2023, doi: <https://doi.org/10.1155/2023/7717712>.
- [50] D. Abdelhafiz, C. Yang, R. Ammar, and S. Nabavi, “Deep convolutional neural networks for mammography: advances, challenges and applications,” *BMC Bioinformatics*, vol. 20, no. 11, p. 281, 2019, doi: 10.1186/s12859-019-2823-4.
- [51] P. Hamet and J. Tremblay, “Artificial intelligence in medicine,” *Metab. - Clin. Exp.*, vol. 69, pp. S36–S40, Apr. 2017, doi: 10.1016/j.metabol.2017.01.011.
- [52] M. A. Al-Antari, M. A. Al-Masni, M.-T. Choi, S.-M. Han, and T.-S. Kim, “A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification.,” *Int. J. Med. Inform.*, vol. 117,

- pp. 44–54, Sep. 2018, doi: 10.1016/j.ijmedinf.2018.06.003.
- [53] A. Rodriguez-Ruiz *et al.*, “Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study,” *Eur. Radiol.*, vol. 29, no. 9, pp. 4825–4832, 2019, doi: 10.1007/s00330-019-06186-9.
- [54] S. L. Thanh-Nghia Nguyen, Thanh-Hai Nguyen, Manh-Hung Nguyen, “Wavelet-Based Kernel Construction for Heart Disease Classification,” *Adv. Electr. Eng.*, vol. 17, no. 3, pp. 306–319, 2019.
- [55] T.-N. N. Manh-Hung Nguyen, Vu-Hoang-Tran, Thanh-Hai Nguyen, “A Deep Learning Framework for Inter-Patient ECG Classification,” *Int. J. Comput. Sci. Netw. Secur.*, vol. 19, no. 1, pp. 74–84, 2019.
- [56] L. Wei, Q. Gan, and T. Ji, “Skin Disease Recognition Method Based on Image Color and Texture Features,” *Comput. Math. Methods Med.*, vol. 2018, p. 8145713, 2018, doi: 10.1155/2018/8145713.
- [57] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, “Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1240–1251, 2016, doi: 10.1109/TMI.2016.2538465.
- [58] M. A. S. Al Husaini, M. H. Habaebi, S. A. Hameed, M. R. Islam, and T. S. Gunawan, “A Systematic Review of Breast Cancer Detection Using Thermography and Neural Networks,” *IEEE Access*, vol. 8, pp. 208922–208937, 2020, doi: 10.1109/ACCESS.2020.3038817.
- [59] Z. Zhu, Y. Sun, and B. Honarvar Shakibaei Asli, “Early Breast Cancer Detection Using Artificial Intelligence Techniques Based on Advanced Image Processing Tools,” *Electronics*, vol. 13, no. 17, 2024, doi: 10.3390/electronics13173575.
- [60] A. Yücel, B. Değirmenci, M. Acar, H. Ellidokuz, R. Albayrak, and A. Haktanır, “Knowledge about breast cancer and mammography in breast cancer screening among women awaiting mammography,” *Turkish J. Med. Sci.*, vol. 35, pp. 35–42, 2005.
- [61] I. Domingues, P. H. Abreu, and J. Santos, “BIRADS Classification of Breast Cancer: A New Pre-Processing Pipeline for Deep Models Training,” in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 1378–1382, doi: 10.1109/ICIP.2018.8451510.
- [62] R. K. Samala, H.-P. Chan, L. Hadjiiski, M. A. Helvie, C. D. Richter, and K. H. Cha, “Breast Cancer Diagnosis in Digital Breast Tomosynthesis: Effects of Training Sample Size on Multi-Stage Transfer Learning Using Deep Neural Nets,” *IEEE Trans. Med. Imaging*, vol. 38, no. 3, pp. 686–696, Mar. 2019, doi: 10.1109/TMI.2018.2870343.
- [63] D. et al. Suckling, J., Parker, J., Dance, *Mammographic Image Analysis Society (MIAS) database v1.21*. 2015.
- [64] M. Heath *et al.*, “Current Status of the Digital Database for Screening Mammography BT - Digital Mammography: Nijmegen, 1998,” N. Karssemeijer, M. Thijssen, J. Hendriks, and L. van Erning, Eds. Dordrecht: Springer Netherlands, 1998, pp. 457–460.

## TÀI LIỆU THAM KHẢO

---

- [65] D. R. R. S. Lee, F. Gimenez, A. Hoogi, “Curated Breast Imaging Subset of DDSM,” *Cancer Imaging Arch.*, 2016.
- [66] R. S. Lee, F. Gimenez, A. Hoogi, K. K. Miyake, M. Gorovoy, and D. L. Rubin, “A curated mammography data set for use in computer-aided detection and diagnosis research,” *Sci. data*, vol. 4, p. 170177, Dec. 2017, doi: 10.1038/sdata.2017.177.
- [67] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, “INbreast: Toward a Full-field Digital Mammographic Database,” *Acad. Radiol.*, vol. 19, no. 2, pp. 236–248, 2012, doi: <https://doi.org/10.1016/j.acra.2011.09.014>.
- [68] M. A. G. Lopez *et al.*, “BCDR : A BREAST CANCER DIGITAL REPOSITORY,” 2012, [Online]. Available: <https://api.semanticscholar.org/CorpusID:16423079>.
- [69] J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, and M. A. Guevara Lopez, “Representation learning for mammography mass lesion classification with convolutional neural networks,” *Comput. Methods Programs Biomed.*, vol. 127, pp. 248–257, 2016, doi: <https://doi.org/10.1016/j.cmpb.2015.12.014>.
- [70] B. R. N. Matheus and H. Schiabel, “Online mammographic images database for development and comparison of CAD schemes,” *J. Digit. Imaging*, vol. 24, no. 3, pp. 500–506, Jun. 2011, doi: 10.1007/s10278-010-9297-2.
- [71] A. Badano *et al.*, “Evaluation of Digital Breast Tomosynthesis as Replacement of Full-Field Digital Mammography Using an In Silico Imaging Trial,” *JAMA Netw. open*, vol. 1, no. 7, p. e185474, Nov. 2018, doi: 10.1001/jamanetworkopen.2018.5474.
- [72] M. D. Halling-Brown *et al.*, “OPTIMAM Mammography Image Database: A Large-Scale Resource of Mammography Images and Clinical Data,” *Radiol. Artif. Intell.*, vol. 3, no. 1, p. e200103, Jan. 2021, doi: 10.1148/ryai.2020200103.
- [73] M. H. Yap *et al.*, “Automated Breast Ultrasound Lesions Detection Using Convolutional Neural Networks,” *IEEE J. Biomed. Heal. Informatics*, vol. 22, no. 4, pp. 1218–1226, 2018, doi: 10.1109/JBHI.2017.2731873.
- [74] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy, “Dataset of breast ultrasound images,” *Data Br.*, vol. 28, p. 104863, 2020, doi: <https://doi.org/10.1016/j.dib.2019.104863>.
- [75] Y. Zhang *et al.*, “BUSIS: A Benchmark for Breast Ultrasound Image Segmentation,” *Healthcare*, vol. 10, no. 4, 2022, doi: 10.3390/healthcare10040729.
- [76] M. A. Zanona, “Image Processing & Neural Network Based Breast Cancer Detection,” *Comput. Inf. Sci.*, vol. 12, no. 2, pp. 146–, 2019, doi: 10.5539/cis.v12n2p146.
- [77] 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.
- [78] S. M. Shaaban, M. Nawaz, Y. Said, and M. Barr, “An Efficient Breast Cancer Segmentation System based on Deep Learning Techniques,” *Eng. Technol. Appl. Sci. Res.*, vol. 13, no. 6 SE-, pp. 12415–12422, Dec. 2023, doi: 10.48084/etasr.6518.
- [79] A. I. Dumachi and C. Buiu, “Applications of Machine Learning in Cancer Imaging:

## TÀI LIỆU THAM KHẢO

---

- A Review of Diagnostic Methods for Six Major Cancer Types,” *Electronics*, vol. 13, no. 23, 2024, doi: 10.3390/electronics13234697.
- [80] E. Justaniah, A. Alhothali, and G. Aldabbagh, “Mammogram segmentation techniques: A review,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 5, 2021.
- [81] T.-H. Nguyen, T.-N. Nguyen, and B.-V. Ngo, “A VGG-19 Model with Transfer Learning and Image Segmentation for Classification of Tomato Leaf Disease,” *AgriEngineering*, vol. 4, pp. 871–887, 2022, doi: 10.3390/agriengineering4040056.
- [82] Katherine, R. Rulaningtyas, and K. Ain, “CT scan image segmentation based on hounsfield unit values using Otsu thresholding method,” *J. Phys. Conf. Ser.*, vol. 1816, no. 1, p. 12080, Feb. 2021, doi: 10.1088/1742-6596/1816/1/012080.
- [83] K. Loizidou, G. Skouroumouni, C. Nikolaou, and C. Pitris, “Automatic Breast Mass Segmentation and Classification Using Subtraction of Temporally Sequential Digital Mammograms,” *IEEE J. Transl. Eng. Heal. Med.*, vol. 10, p. 1801111, 2022, doi: 10.1109/JTEHM.2022.3219891.
- [84] H. D. Cheng, X. J. Shi, R. Min, L. M. Hu, X. P. Cai, and H. N. Du, “Approaches for automated detection and classification of masses in mammograms,” *Pattern Recognit.*, vol. 39, no. 4, pp. 646–668, 2006, doi: <https://doi.org/10.1016/j.patcog.2005.07.006>.
- [85] M. Hmida, K. Hamrouni, B. Solaiman, and S. Boussetta, “Mammographic mass segmentation using fuzzy contours,” *Comput. Methods Programs Biomed.*, vol. 164, pp. 131–142, 2018, doi: <https://doi.org/10.1016/j.cmpb.2018.07.005>.
- [86] K. L. Kashyap, M. K. Bajpai, and P. Khanna, “Globally supported radial basis function based collocation method for evolution of level set in mass segmentation using mammograms,” *Comput. Biol. Med.*, vol. 87, pp. 22–37, 2017, doi: <https://doi.org/10.1016/j.compbiomed.2017.05.015>.
- [87] S. Hussain *et al.*, “A Discriminative Level Set Method with Deep Supervision for Breast Tumor Segmentation,” *Comput. Biol. Med.*, vol. 149, p. 105995, 2022, doi: <https://doi.org/10.1016/j.compbiomed.2022.105995>.
- [88] H. Pezeshki, “Breast tumor segmentation in digital mammograms using spiculated regions,” *Biomed. Signal Process. Control*, vol. 76, p. 103652, 2022, doi: <https://doi.org/10.1016/j.bspc.2022.103652>.
- [89] W. M. Salama and M. H. Aly, “Deep learning in mammography images segmentation and classification: Automated CNN approach,” *Alexandria Eng. J.*, vol. 60, no. 5, pp. 4701–4709, 2021, doi: <https://doi.org/10.1016/j.aej.2021.03.048>.
- [90] S. Li, M. Dong, G. Du, and X. Mu, “Attention Dense-U-Net for Automatic Breast Mass Segmentation in Digital Mammogram,” *IEEE Access*, vol. 7, pp. 59037–59047, 2019, doi: 10.1109/ACCESS.2019.2914873.
- [91] Z. Sani, R. Prasad, and E. K. M. Hashim, “Grouped mask region convolution neural networks for improved breast cancer segmentation in mammography images,” *Evol. Syst.*, vol. 15, no. 1, pp. 25–40, 2024, doi: 10.1007/s12530-023-09527-8.
- [92] M. Biltawi, N. Al-Najdawi, and S. Tedmori, “Mammogram Enhancement and Segmentation Methods: Classification, Analysis, and Evaluation,” 2012.
- [93] P. Teare, M. Fishman, O. Benzaquen, E. Toledano, and E. Elnkave, “Malignancy

- Detection on Mammography Using Dual Deep Convolutional Neural Networks and Genetically Discovered False Color Input Enhancement,” *J. Digit. Imaging*, vol. 30, no. 4, pp. 499–505, 2017, doi: 10.1007/s10278-017-9993-2.
- [94] K. A. A. Saifullah Harith Suradi and N. A. M. Isa, “Improvement of image enhancement for mammogram images using Fuzzy Anisotropic Diffusion Histogram Equalisation Contrast Adaptive Limited (FADHECAL),” *Comput. Methods Biomech. Biomed. Eng. Imaging & Vis.*, vol. 10, no. 1, pp. 67–75, 2022, doi: 10.1080/21681163.2021.1972344.
- [95] A. Dounis, A.-N. Avramopoulos, and M. Kallergi, “Advanced Fuzzy Sets and Genetic Algorithm Optimizer for Mammographic Image Enhancement,” *Electronics*, vol. 12, no. 15, 2023, doi: 10.3390/electronics12153269.
- [96] H. G. Welch, P. C. Prorok, A. J. O’Malley, and B. S. Kramer, “Breast-Cancer Tumor Size, Overdiagnosis, and Mammography Screening Effectiveness,” *N. Engl. J. Med.*, vol. 375, no. 15, pp. 1438–1447, 2016, doi: 10.1056/NEJMoa1600249.
- [97] M. Alsheh Ali, M. Eriksson, K. Czene, P. Hall, and K. Humphreys, “Detection of potential microcalcification clusters using multivendor for-presentation digital mammograms for short-term breast cancer risk estimation,” *Med. Phys.*, vol. 46, no. 4, pp. 1938–1946, Apr. 2019, doi: 10.1002/mp.13450.
- [98] P. S. K. V. N, and S. S, “Breast Cancer Detection using Crow Search Optimization based Intuitionistic Fuzzy Clustering with Neighborhood Attraction,” *Asian Pac. J. Cancer Prev.*, vol. 20, no. 1, pp. 157–165, Jan. 2019, doi: 10.31557/APJCP.2019.20.1.157.
- [99] Y. Jiang, M. F. Inciardi, A. V Edwards, and J. Papaioannou, “Interpretation Time Using a Concurrent-Read Computer-Aided Detection System for Automated Breast Ultrasound in Breast Cancer Screening of Women With Dense Breast Tissue,” *Am. J. Roentgenol.*, vol. 211, no. 2, pp. 452–461, May 2018, doi: 10.2214/AJR.18.19516.
- [100] M. Fan, Y. Li, S. Zheng, W. Peng, W. Tang, and L. Li, “Computer-aided detection of mass in digital breast tomosynthesis using a faster region-based convolutional neural network,” *Methods*, vol. 166, pp. 103–111, 2019, doi: <https://doi.org/10.1016/j.ymeth.2019.02.010>.
- [101] M. G. Ertosun and D. L. Rubin, “Probabilistic visual search for masses within mammography images using deep learning,” in *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2015, pp. 1310–1315, doi: 10.1109/BIBM.2015.7359868.
- [102] M. A. Al-masni *et al.*, “Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system,” *Comput. Methods Programs Biomed.*, vol. 157, pp. 85–94, 2018, doi: <https://doi.org/10.1016/j.cmpb.2018.01.017>.
- [103] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.
- [104] W. Zhu, Q. Lou, Y. S. Vang, and X. Xie, “Deep Multi-instance Networks with Sparse Label Assignment for Whole Mammogram Classification BT - Medical Image Computing and Computer Assisted Intervention,” in *MICCAI 2017*, 2017, pp. 603–

611.

- [105] L. Bouzar-Benlabiod, K. Harrar, L. Yamoun, M. Y. Khodja, and M. A. Akhloufi, “A novel breast cancer detection architecture based on a CNN-CBR system for mammogram classification,” *Comput. Biol. Med.*, vol. 163, p. 107133, 2023, doi: <https://doi.org/10.1016/j.combiomed.2023.107133>.
- [106] F. Gao *et al.*, “SD-CNN: A shallow-deep CNN for improved breast cancer diagnosis,” *Comput. Med. Imaging Graph.*, vol. 70, pp. 53–62, 2018, doi: <https://doi.org/10.1016/j.compmedimag.2018.09.004>.
- [107] A. S. Elkorany and Z. F. Elsharkawy, “Efficient breast cancer mammograms diagnosis using three deep neural networks and term variance,” *Sci. Rep.*, vol. 13, no. 1, p. 2663, 2023, doi: [10.1038/s41598-023-29875-4](https://doi.org/10.1038/s41598-023-29875-4).
- [108] X. Yu, W. Pang, Q. Xu, and M. Liang, “Mammographic image classification with deep fusion learning,” *Sci. Rep.*, vol. 10, no. 1, p. 14361, 2020, doi: [10.1038/s41598-020-71431-x](https://doi.org/10.1038/s41598-020-71431-x).
- [109] A. A. Mohamed, W. A. Berg, H. Peng, Y. Luo, R. C. Jankowitz, and S. Wu, “A deep learning method for classifying mammographic breast density categories,” *Med. Phys.*, vol. 45, no. 1, pp. 314–321, Jan. 2018, doi: [10.1002/mp.12683](https://doi.org/10.1002/mp.12683).
- [110] C. K. Ahn, C. Heo, H. Jin, and J. H. Kim, “A novel deep learning-based approach to high accuracy breast density estimation in digital mammography,” in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Mar. 2017, vol. 10134, p. 101342O, doi: [10.1117/12.2254264](https://doi.org/10.1117/12.2254264).
- [111] N. Wu *et al.*, *Breast Density Classification with Deep Convolutional Neural Networks*. 2018.
- [112] M. Kallenberg *et al.*, “Unsupervised Deep Learning Applied to Breast Density Segmentation and Mammographic Risk Scoring,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1322–1331, 2016, doi: [10.1109/TMI.2016.2532122](https://doi.org/10.1109/TMI.2016.2532122).
- [113] A. Rampun, B. W. Scotney, P. J. Morrow, and H. Wang, “Breast Mass Classification in Mammograms using Ensemble Convolutional Neural Networks,” in *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)*, 2018, pp. 1–6, doi: [10.1109/HealthCom.2018.8531154](https://doi.org/10.1109/HealthCom.2018.8531154).
- [114] M. M. Jadoon, Q. Zhang, I. U. Haq, S. Butt, and A. Jadoon, “Three-Class Mammogram Classification Based on Descriptive CNN Features,” *Biomed Res. Int.*, vol. 2017, p. 3640901, 2017, doi: [10.1155/2017/3640901](https://doi.org/10.1155/2017/3640901).
- [115] B. Q. Huynh, H. Li, and M. L. Giger, “Digital mammographic tumor classification using transfer learning from deep convolutional neural networks,” *J. Med. Imaging (Bellingham, Wash.)*, vol. 3, no. 3, p. 34501, Jul. 2016, doi: [10.1117/1.JMI.3.3.034501](https://doi.org/10.1117/1.JMI.3.3.034501).
- [116] E. Wu, K. Wu, D. Cox, and W. Lotter, “Conditional Infilling GANs for Data Augmentation in Mammogram Classification,” Jul. 2018, doi: [10.48550/arXiv.1807.08093](https://doi.org/10.48550/arXiv.1807.08093).
- [117] S. Shams, R. Platania, J. Zhang, J. Kim, K. Lee, and S.-J. Park, “Deep Generative Breast Cancer Screening and Diagnosis BT - Medical Image Computing and Computer Assisted Intervention – MICCAI 2018,” 2018, pp. 859–867.

## TÀI LIỆU THAM KHẢO

---

- [118] S. M. McKinney *et al.*, “International evaluation of an AI system for breast cancer screening,” *Nature*, vol. 577, no. 7788, pp. 89–94, 2020, doi: 10.1038/s41586-019-1799-6.
- [119] Y.-S. Sun *et al.*, “Risk Factors and Preventions of Breast Cancer,” *Int. J. Biol. Sci.*, vol. 13, no. 11, pp. 1387–1397, 2017, doi: 10.7150/ijbs.21635.
- [120] R. D. Nindrea, T. Aryandono, L. Lazuardi, and I. Dwiprahasto, “Diagnostic Accuracy of Different Machine Learning Algorithms for Breast Cancer Risk Calculation: a Meta-Analysis,” *Asian Pac. J. Cancer Prev.*, vol. 19, no. 7, pp. 1747–1752, Jul. 2018, doi: 10.22034/APJCP.2018.19.7.1747.
- [121] M. Sepandi, M. Taghdir, A. Rezaianzadeh, and S. Rahimikazerooni, “Assessing Breast Cancer Risk with an Artificial Neural Network,” *Asian Pac. J. Cancer Prev.*, vol. 19, no. 4, pp. 1017–1019, Apr. 2018, doi: 10.22034/APJCP.2018.19.4.1017.
- [122] K. Dembrower *et al.*, “Comparison of a Deep Learning Risk Score and Standard Mammographic Density Score for Breast Cancer Risk Prediction,” *Radiology*, vol. 294, no. 2, pp. 265–272, Feb. 2020, doi: 10.1148/radiol.2019190872.
- [123] D. Arefan, A. A. Mohamed, W. A. Berg, M. L. Zuley, J. H. Sumkin, and S. Wu, “Deep learning modeling using normal mammograms for predicting breast cancer risk,” *Med. Phys.*, vol. 47, no. 1, pp. 110–118, Jan. 2020, doi: 10.1002/mp.13886.
- [124] A. Yala, C. Lehman, T. Schuster, T. Portnoi, and R. Barzilay, “A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction,” *Radiology*, vol. 292, no. 1, pp. 60–66, 2019, doi: 10.1148/radiol.2019182716.
- [125] A. Yala *et al.*, “Toward robust mammography-based models for breast cancer risk,” *Sci. Transl. Med.*, vol. 13, no. 578, p. eaba4373, 2021, doi: 10.1126/scitranslmed.aba4373.
- [126] R. Ha *et al.*, “Convolutional Neural Network Based Breast Cancer Risk Stratification Using a Mammographic Dataset,” *Acad. Radiol.*, vol. 26, no. 4, pp. 544–549, Apr. 2019, doi: 10.1016/j.acra.2018.06.020.
- [127] Y. M. George, H. H. Zayed, M. I. Roushdy, and B. M. Elbagoury, “Remote Computer-Aided Breast Cancer Detection and Diagnosis System Based on Cytological Images,” *IEEE Syst. J.*, vol. 8, no. 3, pp. 949–964, 2014, doi: 10.1109/JSYST.2013.2279415.
- [128] L. K. Singh, M. Khanna, and R. Singh, “Artificial intelligence based medical decision support system for early and accurate breast cancer prediction,” *Adv. Eng. Softw.*, vol. 175, p. 103338, 2023, doi: <https://doi.org/10.1016/j.advengsoft.2022.103338>.
- [129] M. B. Morgan and J. L. Mates, “Applications of Artificial Intelligence in Breast Imaging,” *Radiol. Clin. North Am.*, vol. 59, no. 1, pp. 139–148, Jan. 2021, doi: 10.1016/j.rcl.2020.08.007.
- [130] G. Quellec, M. Lamard, M. Cozic, G. Coatrieux, and G. Cazuguel, “Multiple-Instance Learning for Anomaly Detection in Digital Mammography,” *IEEE Trans. Med. Imaging*, vol. 35, no. 7, pp. 1604–1614, 2016, doi: 10.1109/TMI.2016.2521442.
- [131] E. B. Mendelson, “Artificial Intelligence in Breast Imaging: Potentials and Limitations,” *Am. J. Roentgenol.*, vol. 212, no. 2, pp. 293–299, 2019, doi: 10.2214/AJR.18.20532.

## TÀI LIỆU THAM KHẢO

---

- [132] A. A. Mohamed, Y. Luo, H. Peng, R. C. Jankowitz, and S. Wu, “Understanding Clinical Mammographic Breast Density Assessment: a Deep Learning Perspective,” *J. Digit. Imaging*, vol. 31, no. 4, pp. 387–392, 2018, doi: 10.1007/s10278-017-0022-2.
- [133] T. Kooi *et al.*, “Large scale deep learning for computer aided detection of mammographic lesions,” *Med. Image Anal.*, vol. 35, pp. 303–312, 2017, doi: <https://doi.org/10.1016/j.media.2016.07.007>.
- [134] K. C. K. W. H. Kim Jaeil Kim Hye Jung, “Artificial intelligence in breast ultrasonography,” *Ultrasonography*, vol. 40, no. 2, pp. 183–190, 2021, doi: 10.14366/usg.20117.
- [135] H. T. Nguyen *et al.*, “VinDr-Mammo: A large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography,” *medRxiv*, 2022, doi: 10.1101/2022.03.07.22272009.
- [136] A. J and D. Peter, “Mass segmentation in mammograms using a kernel-based fuzzy level set method,” *Int. J. Biomed. Eng. Technol.*, vol. 9, p. pp.133-153, Jan. 2015, doi: 10.1504/IJBET.2015.072933.
- [137] S. Rahman, M. M. Rahman, M. Abdullah-Al-Wadud, G. D. Al-Quaderi, and M. Shoyaib, “An adaptive gamma correction for image enhancement,” *EURASIP J. Image Video Process.*, vol. 2016, no. 1, p. 35, 2016, doi: 10.1186/s13640-016-0138-1.
- [138] N. S. P. Kong, H. Ibrahim, and S. C. Hoo, “A literature review on histogram equalization and its variations for digital image enhancement,” *Int. J. Innov. Manag. Technol.*, vol. 4, no. 4, p. 386, 2013.
- [139] R.-C. Chen, C. Dewi, Y.-C. Zhuang, and J.-K. Chen, “Contrast Limited Adaptive Histogram Equalization for Recognizing Road Marking at Night Based on Yolo Models,” *IEEE Access*, vol. 11, pp. 92926–92942, 2023, doi: 10.1109/ACCESS.2023.3309410.
- [140] A. Mohd Khuzi, R. Besar, W. Wan Zaki, and N. Ahmad, “Identification of masses in digital mammogram using gray level co-occurrence matrices,” *Biomed. Imaging Interv. J.*, vol. 5, no. 3, p. e17, Jul. 2009, doi: 10.2349/bij.5.3.e17.
- [141] 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.
- [142] T. T. Nguyen, T. H. Nguyen, T. N. Nguyen, and B. V. Ngo, “Largest roi segmentation for breast cancer classification using a VGG16 deep learning network,” *Adv. Electr. Electron. Eng.*, 2024.
- [143] N. Thanh-Tam and N. Thanh-Hai, “Segmentation and Color ROI Extraction from Breast Imaging Datasets for Cancer Classification,” *Eng. Technol. Appl. Sci. Res.*, 2025.
- [144] C. D. Lekamlage, F. Afzal, E. Westerberg, and A. Cheddad, “Mini-DDSM: Mammography-based Automatic Age Estimation,” in *2020 3rd International Conference on Digital Medicine and Image Processing*, Nov. 2020, p. 1◆6, doi: 10.1145/3441369.3441370.

- [145] 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.
- [146] P. Oza, P. Sharma, S. Patel, F. Adedoyin, and A. Bruno, “Image Augmentation Techniques for Mammogram Analysis,” *J. Imaging*, vol. 8, no. 5, 2022, doi: 10.3390/jimaging8050141.
- [147] H. Ulutaş and V. Aslantas, “A Fast and Accurate Method for Classifying Tomato Plant Health Status Using Machine Learning and Image Processing,” vol. 23, pp. 54–68, 2023, doi: 10.5755/j02.eie.33866.
- [148] D. Al-Karawi *et al.*, “A Review of Artificial Intelligence in Breast Imaging,” *Tomography*, vol. 10, no. 5, pp. 705–726, 2024, doi: 10.3390/tomography10050055.
- [149] B. Alane, Y. Terchi, and S. Bouguezel, “New Face Recognition System Based on DCT Pyramid and Backpropagation Neural Network,” *Elektron. ir Elektrotehnika*, vol. 30, no. 1, pp. 68–76, 2024, doi: 10.5755/j02.eie.35897.
- [150] M. Cantone, C. Marrocco, F. Tortorella, and A. Bria, “Convolutional Networks and Transformers for Mammography Classification: An Experimental Study,” *Sensors*, vol. 23, no. 3, 2023, doi: 10.3390/s23031229.
- [151] A. Bekkouche, M. Merzoug, M. Hadjila, and W. Ferhi, “Towards Early Breast Cancer Detection: A Deep Learning Approach,” *Eng. Technol. Appl. Sci. Res.*, vol. 14, no. 5 SE-, pp. 17517–17523, Oct. 2024, doi: 10.48084/etasr.8634.
- [152] A. Krizhevsky, I. Sutskever, and G. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Neural Inf. Process. Syst.*, vol. 25, Jan. 2012, doi: 10.1145/3065386.
- [153] S.-H. Chen, Y.-L. Wu, C.-Y. Pan, L.-Y. Lian, and Q.-C. Su, “Breast ultrasound image classification and physiological assessment based on GoogLeNet,” *J. Radiat. Res. Appl. Sci.*, vol. 16, no. 3, p. 100628, 2023, doi: <https://doi.org/10.1016/j.jrras.2023.100628>.
- [154] S. Montaha *et al.*, “BreastNet18: A High Accuracy Fine-Tuned VGG16 Model Evaluated Using Ablation Study for Diagnosing Breast Cancer from Enhanced Mammography Images,” *Biology (Basel)*, vol. 10, no. 12, 2021, doi: 10.3390/biology10121347.
- [155] M. M. Li Li, Changjie Pan, Ming Zhang, Dong Shen, Guangyuan He, “Predicting Malignancy in Breast Lesions: Enhancing Accuracy with Fine-Tuned Convolutional Neural Network Models,” *Res. Sq.*, 2024, doi: <https://doi.org/10.21203/rs.3.rs-3937557/v1>.
- [156] H. Li, S. Zhuang, D. Li, J. Zhao, and Y. Ma, “Benign and malignant classification of mammogram images based on deep learning,” *Biomed. Signal Process. Control*, vol. 51, pp. 347–354, 2019, doi: 10.1016/j.bspc.2019.02.017.
- [157] R. Raza *et al.*, “Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images,” *Eng. Appl. Artif. Intell.*, vol. 126, p. 106902, 2023, doi: 10.1016/j.engappai.2023.106902.
- [158] A. Abdelrahman and S. Viriri, “EfficientNet family U-Net models for deep learning

- semantic segmentation of kidney tumors on CT images,” *Front. Comput. Sci.*, vol. 5, 2023, doi: 10.3389/fcomp.2023.1235622.
- [159] Y. Sun, Z. Zhu, and B. Honarvar Shakibaei Asli, “Automated Classification and Segmentation and Feature Extraction from Breast Imaging Data,” *Electronics*, vol. 13, no. 19, 2024, doi: 10.3390/electronics13193814.
- [160] D. Albashish, R. Al-Sayyed, A. Abdullah, M. H. Ryalat, and N. Ahmad Almansour, “Deep CNN Model based on VGG16 for Breast Cancer Classification,” in *2021 International Conference on Information Technology (ICIT)*, 2021, pp. 805–810, doi: 10.1109/ICIT52682.2021.9491631.
- [161] S. Shamy and S. Dheeba, “A Research on Detection and Classification of Breast Cancer using k- means GMM & CNN Algorithms,” *Int. J. Eng. Adv. Technol.*, vol. 8, no. 6S, pp. 501–505, 2019, doi: <http://doi.org/10.35940/ijeat.F1102.0886S19>.
- [162] V. S. Vijayan and L. P. L., “Deep Learning based Prediction of Breast Cancer in Histopathological images,” *Int. J. Eng. Res. Technol.*, vol. 8, no. 7, pp. 2278–0181, 2019.
- [163] N. Wu *et al.*, “Deep Neural Networks Improve Radiologists’ Performance in Breast Cancer Screening,” *IEEE Trans. Med. Imaging*, vol. 39, no. 4, pp. 1184–1194, Apr. 2020, doi: 10.1109/TMI.2019.2945514.
- [164] A. Khalid *et al.*, “Breast Cancer Detection and Prevention Using Machine Learning,” *Diagnostics*, vol. 13, no. 19, 2023, doi: 10.3390/diagnostics13193113.
- [165] Z. Wang *et al.*, “Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion With CNN Deep Features,” *IEEE Access*, vol. 7, pp. 105146–105158, 2019, doi: 10.1109/ACCESS.2019.2892795.
- [166] S. U. Khan, N. Islam, Z. Jan, I. Ud Din, and J. J. P. C. Rodrigues, “A novel deep learning based framework for the detection and classification of breast cancer using transfer learning,” *Pattern Recognit. Lett.*, vol. 125, pp. 1–6, 2019, doi: 10.1016/j.patrec.2019.03.022.
- [167] A. Khamparia *et al.*, “Diagnosis of breast cancer based on modern mammography using hybrid transfer learning,” *Multidimens. Syst. Signal Process.*, vol. 32, no. 2, pp. 747–765, 2021, doi: 10.1007/s11045-020-00756-7.
- [168] Shallu and R. Mehra, “Breast cancer histology images classification: Training from scratch or transfer learning?,” *ICT Express*, vol. 4, no. 4, pp. 247–254, 2018, doi: 10.1016/j.icte.2018.10.007.
- [169] Z. Hameed, S. Zahia, B. Garcia-Zapirain, J. Javier Aguirre, and A. María Vanegas, “Breast Cancer Histopathology Image Classification Using an Ensemble of Deep Learning Models,” *Sensors (Basel)*, vol. 20, no. 16, Aug. 2020, doi: 10.3390/s20164373.
- [170] H. Nasir Khan, A. Shahid, B. Raza, A. Dar, and H. Alquhayz, “Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network,” *IEEE Access*, vol. PP, p. 1, Nov. 2019, doi: 10.1109/ACCESS.2019.2953318.
- [171] T. T. Nguyen and T. H. Nguyen, “Classification of Breast Cancer Based on Region of Interest and EfficientNet-B7 Model.”

- [172] D. Uzun Ozsahin, D. Ikechukwu Emegano, B. Uzun, and I. Ozsahin, “The Systematic Review of Artificial Intelligence Applications in Breast Cancer Diagnosis,” *Diagnostics*, vol. 13, no. 1, 2023, doi: 10.3390/diagnostics13010045.
- [173] H. Liu *et al.*, “Deep learning radiomics based prediction of axillary lymph node metastasis in breast cancer,” *npj Breast Cancer*, vol. 10, no. 1, p. 22, 2024, doi: 10.1038/s41523-024-00628-4.
- [174] Q. Sun *et al.*, “Deep Learning vs. Radiomics for Predicting Axillary Lymph Node Metastasis of Breast Cancer Using Ultrasound Images: Don’t Forget the Peritumoral Region,” *Front. Oncol.*, vol. 10, p. 53, 2020, doi: 10.3389/fonc.2020.00053.
- [175] S.-M. Hsu, W.-H. Kuo, F.-C. Kuo, and Y.-Y. Liao, “Breast tumor classification using different features of quantitative ultrasound parametric images,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 4, pp. 623–633, 2019, doi: 10.1007/s11548-018-01908-8.
- [176] B. J. E. L. H. S. K. S. H. Choi Ji-Hye Kang Bong Joo, “Application of computer-aided diagnosis in breast ultrasound interpretation: improvements in diagnostic performance according to reader experience,” *Ultrasonography*, vol. 37, no. 3, pp. 217–225, 2018, doi: 10.14366/usg.17046.
- [177] Z. Liu, J. Peng, X. Guo, S. Chen, and L. Liu, “Breast cancer classification method based on improved VGG16 using mammography images,” *J. Radiat. Res. Appl. Sci.*, vol. 17, no. 2, p. 100885, 2024, doi: <https://doi.org/10.1016/j.jrras.2024.100885>.
- [178] M. A. K. Raiaan, N. M. Fahad, M. S. H. Mukta, and S. Shatabda, “Mammo-Light: A lightweight convolutional neural network for diagnosing breast cancer from mammography images,” *Biomed. Signal Process. Control*, vol. 94, p. 106279, 2024, doi: <https://doi.org/10.1016/j.bspc.2024.106279>.
- [179] E. Tarighati Sereshkeh, H. Keivan, K. Shirbandi, F. Khaleghi, and M. M. Bagheri Asl, “A convolution neural network for rapid and accurate staging of breast cancer based on mammography,” *Informatics Med. Unlocked*, vol. 47, p. 101497, 2024, doi: <https://doi.org/10.1016/j.imu.2024.101497>.
- [180] E. BOROUMANDZADEH, Mostafa; PARVINNIA, “Automated classification of BIRADS in textual mammography reports,” *Turkish J. Electr. Eng. Comput. Sci.*, vol. 29, pp. 632–647, 2021, doi: 10.1016/j.jbi.2017.04.011.
- [181] H. T. X. Nguyen, S. B. Tran, D. B. Nguyen, H. H. Pham, and H. Q. Nguyen, “A novel multi-view deep learning approach for BIRADS and density assessment of mammograms,” in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2022, pp. 2144–2148, doi: 10.1109/EMBC48229.2022.9871564.

## CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ

### Công trình trước khi học NCS:

1. N T Hai, P T Phu, N Q Cuong, N T. Tam, “A SVM Algorithm for Investigation of TriAccelerometer-Based Falling Data,” *American Journal of Signal Processing*, Vol. 6, No. 2, pp. 56-65, 2016.
2. Pham Ty Phu, Nguyen Thanh Hai, Nguyen Thanh Tam, “A Threshold Algorithm in a Fall Alert System for Elderly People,” *The 5th International Conference on the Development of Biomedical Engineering in Vietnam*, 2014.
3. Trung Nguyen, Thanh Hai Nguyen, Tam T. Nguyen, Toi V. Vo, “A Mean Threshold Algorithm for Detection of Human Eye Activities Using EEG Technique,” *the 2012 International Conference on Green Technology and Sustainable Development (GTSD2012)*, UTE, HCMC, Vietnam, pp. 555-560, 2012. (ISBN: 978-604-918-021-7).
4. Thanh N. Nguyen, Thanh H. Nguyen, Tam T. Nguyen, “Wavelet Coefficient Average Value for Prediction of Motor Control Area of Human Brain using fNIRS,” *the BME2012 Conference, IFMBE Proceedings*, Vietnam, Vol. 27, pp. 411-415, 2012. (ISSN:1680-0737)

### **Book Chapter:**

1. Thanh-Hai Nguyen, **Thanh-Tam Nguyen**, Ba-Viet Ngo, “A SVM Algorithm for Falling Detection in an IoTs-based System,” in *Handbook: A Handbook of Internet of Things in Biomedical and Cyber Physical System*, Springer Nature Switzerland AG 2020, Vol. 165, pp. 139-170, July 2019. ISBN 978-3-030-23983-1
2. Thanh-Hai Nguyen, Thanh-Nghia Nguyen, **Thanh-Tam Nguyen**, “A Deep Learning Framework for Heart Disease Classification in an IoTs-based

System,” in *Handbook: A Handbook of Internet of Things in Biomedical and Cyber Physical System*, Springer Nature Switzerland AG 2020, Vol. 165, pp. 217-244, July 2019. ISBN 978-3-030-23983-1.

**Kết quả trong quá trình học NCS:**

1. 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.
2. 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.
3. 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/aeee.v22i4.240303.
4. 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, DOI: 10.48084/etasr.12067.
5. 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.

## CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ

---

6. 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, DOI: 10.54644/jte.2025.1969.