**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH**

**KHOA CÔNG NGHỆ THÔNG TIN**

**A picture containing logo, font, graphics, symbol

Description automatically generated**

**NGUYỄN TUẤN SINH**

**TRỊNH NGỌC ĐỨC**

**PHÂN LOẠI MASS TRONG NHŨ ẢNH TỪ ẢNH X-QUANG SỬ DỤNG DEEP LEARNING VÀ ENSEMBLE LEARNING**

**Ngành: Khoa học dữ liệu**

**Giảng viên hướng dẫn: PSG.TS HUỲNH TRUNG HIẾU**

**THÀNH PHỐ HỒ CHÍ MINH, THÁNG 5 NĂM 2023**

**INDUSTRIAL UNIVERSITY OF HO CHI MINH CITY**

**FACULTY OF INFORMATION TECHNOLOGY**

**- - - 🙞 🕮** **🙜 - - -**A picture containing logo, font, graphics, symbol

Description automatically generated

**NGUYỄN TUẤN SINH**

**TRỊNH NGỌC ĐỨC**

**MASS CLASSIFICATION FOR CHEST X-RAY IMAGES USING DEEP LEARNING AND ENSEMBLE LEARNING TECHNIQUES**

**Major: Data Science**

**Supervisor**: **Dr. Huynh Trung Hieu**

**THÀNH PHỐ HỒ CHÍ MINH, THÁNG 5 NĂM 2023**

**CONTENT SUMMARY**

Title: Mass classification using deep learning and ensemble learning techniques.

Abstract:

Breast cancer is the leading cause of cancer-related deaths among women worldwide. Early detection of breast cancer is critical for effective treatment. Mammography is a widely used screening tool for breast cancer, and the detection of masses in mammograms is one of the key tasks for radiologists. However, the manual interpretation of mammograms is time-consuming and prone to human errors. Therefore, there is a need for automated systems to assist radiologists in detecting breast masses.

In this thesis, YOLOv5 model, LSTM deep learning model, ResNet50, CNN and Decision Tree model combined with Gradient Boosting are applied for mass classification.

The yolov5 based method achieved average results with precision, recall, mAP 0.5 and mAP 0.95 measures of 97%, 84%, 91%, 86%, respectively for the lesion detect phase, respectively and not hurt. for the stage, using the Resnet50 model to classify lesions as benign or malignant on average, the results are achieved with an accuracy of 73%

**LỜI CAM ĐOAN**

Tác giả xin cam đoan đây là công trình nghiên cứu của riêng tác giả. Các số liệu, kết quả trong khóa luận tốt nghiệp này là tác giả tự thu thập, trích dẫn, tuyệt đối không sao chép từ bất kì một tài liệu nào.

Tác giả cam kết sẽ hoàn thành khóa luận của mình với tất cả nỗ lực và tâm huyết của mình, tuân thủ các quy định và tiêu chuẩn của trường đề ra. Tác giả sẽ đưa ra những kết quả nghiên cứu chính xác và khách quan nhất có thể, và đảm bảo tính trung thực của tất cả dữ liệu và thông tin được sử dụng.

Tôi tin rằng khóa luận của mình sẽ đóng góp một phần nhỏ vào việc phát triển và nâng cao chất lượng của giáo dục và nghiên cứu tại trường.

Xin chân thành cảm ơn sự hỗ trợ của Ban Giám Hiệu và Hội Đồng Tốt Nghiệp.

Trân trọng,

Sinh viên

Nguyễn Tuấn Sinh Trịnh Ngọc Đức

**LỜI CÁM ƠN**

Sau một khoảng thời gian tìm tòi và nghiên cứu, tác giả đã hoàn thành đề tài khóa luận tốt nghiệp của mình. Để thực hiện được đề tài này, tác giả nhận được rất nhiều sự giúp đỡ, hướng dẫn tận tâm của Thầy cô, bạn bè toàn thể sinh viên Trường Đại học Công nghiệp Thành phố Hồ Chí Minh đã hỗ trợ thực hiện khảo sát để đạt được những kết quả trung thực nhất.

Trước hết, tác giả xin chân thành cảm ơn quý thầy cô Trường Đại học Công nghiệp Thành phố Hồ Chí Minh nói chung và quý thầy cô khoa Công nghệ thông tin nói riêng đã tận tâm truyền đạt kiến thức, kỹ năng cần thiết cho tác giả trong suốt hơn 4 năm qua. Đó là nền tảng vững chắc để tác giả có thể hoàn thành bài luận văn này.

Đặc biệt, tác giả xin gửi lời cảm ơn chân thành đến Thầy PGS.TS Huỳnh Trung Hiếu trong suốt thời gian thực hiện khóa luận. Mặc dù bận rộn với công tác giảng dạy, nhưng thầy luôn dành những thời gian rảnh quý báu của mình để có những định hướng, hướng dẫn sát sao, cụ thể trong quá trình thực hiện khóa luận.

Với thời gian hạn hẹp và năng lực bản thân còn nhiều hạn chế trong khi đề tài nghiên cứu khá rộng lớn. Việc gặp phải những sai sót hay đánh giá, nhận xét chưa sâu sắc, đầy đủ về vấn đề đang được nghiên cứu là điều không thể tránh khỏi, tác giả rất mong nhận được những lời đóng góp, nhận xét từ quý thầy cô. Sự đóng góp, nhận xét của quý thầy cô sẽ giúp tác giả hoàn thiện hơn đề tài này.

Tác giả xin chân thành cảm ơn!

**NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN**

..................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................

Tp. HCM, ngày..... tháng….. năm 2023

**NHẬN XÉT CỦA GIẢNG VIÊN PHẢN BIỆN**

..................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................

Tp. HCM, ngày..... tháng….. năm 2023

Giảng viên phản biện

**MỤC LỤC**

[CHƯƠNG I: GIỚI THIỆU 1](#_Toc135856793)

[1.1 Tổng quan 1](#_Toc135856794)

[1.1.1 Bối cảnh 1](#_Toc135856795)

[1.1.2 Lý do chọn đề tài 3](#_Toc135856796)

[1.2 Tổng quan về bài toán 3](#_Toc135856797)

[1.2.1. Định nghĩa 3](#_Toc135856798)

[1.2.2. Các vấn đề thách thức của bài toán 4](#_Toc135856799)

[1.2.3. Các phương pháp và kết quả nghiên cứu liên quan trong lĩnh vực 5](#_Toc135856800)

[1.3. Mục tiêu nghiên cứu 5](#_Toc135856801)

[1.4. Phạm vi nghiên cứu 6](#_Toc135856802)

[1.5. Ý nghĩa khoa học và thực tiễn 6](#_Toc135856803)

[CHƯƠNG II: CƠ SỞ LÝ THUYẾT 7](#_Toc135856804)

[2.1. Bài toán detection yolo 7](#_Toc135856805)

[2.1.1. Định nghĩa 7](#_Toc135856806)

[2.1.2. Kiến trúc 9](#_Toc135856807)

[2.2. Mô hình yolov5 10](#_Toc135856808)

[2.2.1. Tổng quan về mô hình 10](#_Toc135856809)

[2.2.2. Kiến trúc mạng yolov5 11](#_Toc135856810)

[2.3. Bài toán phân loại (classification) 12](#_Toc135856811)

[2.3.1. Khái niệm 12](#_Toc135856812)

[2.3.2. Các nghiên cứu đã có 13](#_Toc135856813)

[2.4. Mô hình LSTM 16](#_Toc135856814)

[2.4.1. Định nghĩa 16](#_Toc135856815)

[2.2.2. Các bước trích xuất đặc trưng ảnh từ mô hình LSTM 18](#_Toc135856816)

[2.5. Mô hình resNet50 19](#_Toc135856817)

[2.5.1. Định nghĩa 19](#_Toc135856818)

[2.5.2 Kiến trúc mạng ResNet 19](#_Toc135856819)

[2.6. Cây quyết định (Decision Tree) 21](#_Toc135856820)

[2.6.1. Định nghĩa 21](#_Toc135856821)

[2.6.2. Mô hình cây quyết định 21](#_Toc135856822)

[2.6.3. Các loại thuật toán cây quyết định 23](#_Toc135856823)

[2.7. Rừng ngẫu nhiên (Random Forest) 23](#_Toc135856824)

[2.7.1 Định nghĩa 23](#_Toc135856825)

[2.7.2 Mô hình Random Forest 24](#_Toc135856826)

[CHƯƠNG 3: DỮ LIỆU 25](#_Toc135856827)

[3.1. Bộ dữ liệu DDSM 25](#_Toc135856828)

[3.2. Xử lý dữ liệu 26](#_Toc135856829)

[3.2.1. Ngưỡng otsu (otsu thresholding) 26](#_Toc135856830)

[3.2.2. Xử lý hình thái (Morphological Processing) 28](#_Toc135856831)

[3.2.3. Áp dụng hình thái gần để loại bỏ các vùng nhỏ (Apply morphology close to remove small regions) 30](#_Toc135856832)

[3.2.4. Áp dụng hình thái mở để tách vú khỏi các vùng khác (Apply morphology open to separate breast from other regions) 31](#_Toc135856833)

[3.2.5. Đường viền lớn nhất (Largest contour) 33](#_Toc135856834)

[3.2.9. Biến đổi màu sắc bằng bảng tra cứu (Look-Up Table) 36](#_Toc135856835)

[3.2.10 Lật ảnh 37](#_Toc135856836)

[3.2.11 Thay đổi kích thước ảnh 39](#_Toc135856837)

[3.2.12. Contrast Limited Adaptive Histogram Equalization (CLAHE) 40](#_Toc135856838)

[3.3. Hiện thực mô hình 43](#_Toc135856839)

[3.3.1 Phần cứng được sử dụng 43](#_Toc135856840)

[3.3.2 Dữ liệu huấn luyện 43](#_Toc135856841)

[3.3.3 Huấn luyện mô hình Yolov5l 44](#_Toc135856842)

[3.3.4 Huấn luyện mô hình phân loại 44](#_Toc135856843)

[CHƯƠNG IV: KẾT QUẢ 45](#_Toc135856844)

[4.1 Kết quả mô hình Yolov5 45](#_Toc135856845)

[4.2 Kết quả mô hình phân loại 47](#_Toc135856846)

[CHƯƠNG V: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 48](#_Toc135856847)

[5.1 Kết luận 48](#_Toc135856848)

[5.1.1 Kết quả thu được 48](#_Toc135856849)

[5.1.2 Kiến thức 48](#_Toc135856850)

[5.1.3 Kỹ năng 48](#_Toc135856851)

[5.2 Hướng phát triển 48](#_Toc135856852)

[TÀI LIỆU THAM KHẢO 50](#_Toc135856853)

[NHẬT KÝ LÀM VIỆC 60](#_Toc135856854)

**MỤC LỤC HÌNH ẢNH**

[Hình 2.1 Hình ảnh minh họa sơ khai về Yolo 8](#_Toc135856855)

[Hình 2.2 Kiến trúc Yolo 9](file:///C:\Users\Admin\Desktop\code_kltn\Report_Khoa_luan_tot_nghiep.docx#_Toc135856856)

[Hình 2.3 Kiến trúc Pyramid Network. 10](#_Toc135856857)

[Hình 2.4 Kiến trúc mạng Yolov5 11](#_Toc135856858)

[Hình 2.5 Mô hình LSTM. 17](file:///C:\Users\Admin\Desktop\code_kltn\Report_Khoa_luan_tot_nghiep.docx#_Toc135856859)

[Hình 2.6 Single Residual Block. 19](file:///C:\Users\Admin\Desktop\code_kltn\Report_Khoa_luan_tot_nghiep.docx#_Toc135856860)

[Hình 2.7 Kiến trúc mạng ResNet50. 20](file:///C:\Users\Admin\Desktop\code_kltn\Report_Khoa_luan_tot_nghiep.docx#_Toc135856861)

[Hình 2.8 Decision Tree [34] 22](file:///C:\Users\Admin\Desktop\code_kltn\Report_Khoa_luan_tot_nghiep.docx#_Toc135856862)

[Hình 2.9 Hình minh họa về cây quyết định.[43] 23](file:///C:\Users\Admin\Desktop\code_kltn\Report_Khoa_luan_tot_nghiep.docx#_Toc135856863)

[Hình 2.10 Mô hình Random Forest 24](file:///C:\Users\Admin\Desktop\code_kltn\Report_Khoa_luan_tot_nghiep.docx#_Toc135856864)

[Hình 3.1 Hình minh họa về Otsu thresholding 27](#_Toc135857388)

[Hình 3.2 Ví dụ về Morphological Operator 28](#_Toc135857389)

[Hình 3.3 Hình minh họa về Opening và Closing. 29](#_Toc135857390)

[Hình 3.4 Hình minh họa về phép giãn (dilation) 30](#_Toc135857391)

[Hình 3.5 Hình minh họa lấy đường viền có diện tích lớn nhất 35](file:///C:\Users\Admin\Desktop\code_kltn\Report_Khoa_luan_tot_nghiep.docx#_Toc135857392)

[Hình 3.6 Hình minh họa sử dụng hàm LUT 37](file:///C:\Users\Admin\Desktop\code_kltn\Report_Khoa_luan_tot_nghiep.docx#_Toc135857393)

[Hình 3.7 Hình minh họa về lật ảnh. 39](file:///C:\Users\Admin\Desktop\code_kltn\Report_Khoa_luan_tot_nghiep.docx#_Toc135857394)

[Hình 3.8 Hình minh họa về resize ảnh. 40](file:///C:\Users\Admin\Desktop\code_kltn\Report_Khoa_luan_tot_nghiep.docx#_Toc135857395)

[Hình 3.9 Hình minh họa cân bằng ngưỡng CLAHE 41](file:///C:\Users\Admin\Desktop\code_kltn\Report_Khoa_luan_tot_nghiep.docx#_Toc135857396)

[Hình 3.10 Hình minh họa xác định bounding box 42](#_Toc135857397)

[Hình 4.1 Đồ thị biểu thị độ lỗi trong quá trình huấn luyện 46](#_Toc135856875)

**DANH MỤC CÁC BẢNG BIỂU**

[Bảng 4.1 Thông số ở cứng 43](#_Toc135856923)

[Bảng 4.2 thống kê dữ liệu huấn luyện Yolo 43](#_Toc135856924)

[Bảng 4.3 Thống kê dữ liệu mô hình phân loại 44](#_Toc135856925)

[Bảng 4.4 Kết quả mô hình Yolo 45](#_Toc135856926)

[Bảng 4.5 Két quả mô hình phân loại 47](#_Toc135856927)

**DANH MỤC CÁC TỪ VIẾT TẮT**

|  |  |  |
| --- | --- | --- |
| STT | Ký hiệu chữ viết tắt | Tên gọi đầy đủ |
| 1 | CAD | Computer-Aided Diagnosis |
| 2 | DDSM | Digital Database for Screening Mammography |
| 3 | MRI | Magnetic Resonance Imaging |
| 4 | BMC | Breast Mass Classification |
| 5 | ML | Machine Learning |
| 6 | DL | Deep Learning |
| 7 | CNN | Convolutional Neural Network |
| 8 | BIP | Biomedical Image Processing |
| 9 | YOLO | You Only Look Once |
| 10 | ROI | Region of Interest |
| 11 | k-NN | k-Nearest Neighbors |
| 12 | SVM | Support Vector Machine |
| 13 | RF | Random Forest |
| 14 | HOG | Histogram of Oriented Gradients |
| 15 | CVPR | Computer Vision and Pattern Recognition |
| 16 | CSP | Cross Stage Partial |
| 17 | PANet | Path Aggregation Network |
| 18 | SPP | Spatial Pyramid Pooling |
| 19 | SiLU | Sigmoid Linear Unit |
| 20 | SVM | Support Vector Machines |
| 21 | GBM | Gradient Boosting Machines |
| 22 | LSTM | Long Short-Term Memory |
| 23 | RNN | Recurrent Neural Networks |
| 24 | RGB | Red Green Blue |
| 25 | ResNet | Residual Network |
| 26 | ReLU | Rectified Linear Unit |
| 27 | VGG | Visual Geometry Group |
| 28 | ILSVRC | ImageNet Large Scale Visual Recognition Challenge |
| 29 | CART | Classification and Regression Trees |
| 30 | DT | Decision Tree |
| 31 | ID3 | Iterative Dichotomies 3 |
| 32 | CHAID | CHi-squared Automatic Interaction Detector |
| 33 | MARS | Multivariate Adaptive Regression Splines |
| 34 | GUIDE | Generalized, Unbiased, Interaction Detection and Estimation |
| 35 | CTREE | Conditional Inference Trees |
| 36 | CRUISE | Classification Rule with Unbiased Interaction Selection and Estimation |
| 37 | QUEST | Quick, Unbiased and Efficient Statistical Tree |
| 38 | MD | Doctor of Medicine |
| 39 | ACR | American College of Radiology |
| 40 | SE | Structuring Element |
| 41 | MRI | Magnetic Resonance Imaging |
| 42 | LBP | Local Binary Patterns |
| 43 | CT | Computed Tomography |
| 44 | CLAHE | Contrast Limited Adaptive Histogram Equalization |

# CHƯƠNG I: GIỚI THIỆU

## 1.1 Tổng quan

### 1.1.1 Bối cảnh

Ung thư vú là một trong những khối u ác tính xâm lấn cao nhất xảy ra ở phụ nữ và hiếm khi xảy ra ở nam giới [[1]](#bookmark=id.o836z1c6tzpm). Nó được coi là căn bệnh tồi tệ nhất sau ung thư phổi do tỷ lệ tử vong cao ở phụ nữ [[2]](#bookmark=id.b223hz2z4sa). Theo nghiên cứu của Viện Ung thư Quốc gia Trung Quốc từ năm 2008 đến năm 2012, có 1,67 triệu trường hợp ung thư vú được báo cáo và 522.000 người chết [[3]](#bookmark=id.lxadk4sf6e0u). May mắn là cũng được quan sát từ nghiên cứu trong năm năm gần đây; nếu phát hiện ung thư vú ở giai đoạn sớm, tỷ lệ sống sót có thể đạt hơn 90% [[4]](#bookmark=id.oc8uh0fjsng3). Ngược lại, chẩn đoán muộn có thể khiến nó lan sang các nút bạch huyết và các phần khác của cơ thể, làm giảm tỷ lệ sống sót 27% [[5]](#bookmark=id.or6uo639e1ag). Do đó, việc phát hiện ung thư vú ở giai đoạn sớm là cần thiết để cứu sống phụ nữ. Vú có cấu trúc rất phức tạp, đôi khi chuyên gia không thể xác định được các vùng tổn thương, và việc xem xét một số lượng lớn hình ảnh trong một khoảng thời gian ngắn cũng có thể dẫn đến sai sót trong chẩn đoán [[6]](#bookmark=id.r1b0cx94tj53). Quy trình chẩn đoán ung thư vú yêu cầu các kiểm tra y tế. Các kỹ thuật Hỗ trợ chẩn đoán Máy tính (CAD) dựa trên kiểm tra y tế là tiện lợi và giảm công việc cho bác sĩ [[7]](#bookmark=id.hp0jezwaafqz). CAD cũng đã mang lại những đổi mới kỹ thuật trong chẩn đoán y tế. Các kiểm tra dựa trên CAD phổ biến bao gồm chụp quang tuyến vú, hình ảnh cộng hưởng từ (MRI) và siêu âm. Chụp quang tuyến vú là phương pháp hiệu quả nhất, đáng tin cậy, chi phí thấp và ít gây hại trong việc phát hiện sớm ung thư vú [[8]](#bookmark=id.vb34ts3ff5q8). CAD tự động với chụp quang tuyến vú có tỷ lệ chính xác phát hiện cao hơn, tăng tốc độ quy trình chẩn đoán và tiết kiệm tài nguyên y tế [[9]](#bookmark=id.b3r4342f54x7). Hơn nữa, các vùng tổn thương suất hiện trong nhũ ảnh dưới dạng các khối khá dày đặc, vùng tổn thương có vai trò quan trọng trong chẩn đoán ung thư vú. Thông tin biên giới của vùng tổn thương cho thấy các đặc điểm sinh học và mô hình tăng trưởng [[10]](#bookmark=id.bq4xlh766xq1). Các biên giới không đều của vùng tổn thương thường là loại ác tính và độ chính xác trong việc phân đoạn vùng nghi ngờ bị ác tính ảnh hưởng đến phân loại tổn thương nhũ[[11]](#bookmark=id.wbfyk0kxhwdb). Do đó, phân loại vùng tổn thương là một quy trình quan trọng trong CAD của ung thư vú. Nó giúp các bác sĩ trong việc điều trị và chẩn đoán ung thư vú ở giai đoạn sớm. Khối u vú có các đặc điểm khác nhau như ranh giới bệnh tật, hình dạng và kích thước khác nhau. Do đó, việc phân đoạn phân loại chính xác là một vấn đề thách thức và phổ biến trong các kỹ thuật CAD [[11]](#bookmark=id.wbfyk0kxhwdb).

Nhiều nghiên cứu đã tập trung vào phân loại khối lượng vú (BMC) và phát triển các thuật toán khác nhau để trích xuất các đặc điểm tiềm năng khác nhau từ ảnh chụp quang tuyến vú [[12]](#bookmark=id.a547ytsq1caw) bằng cách sử dụng kỹ thuật Học máy (ML) và Học sâu (DL). Thông thường, hệ thống CAD bao gồm bốn phần chính: tiền xử lý hình ảnh, phân đoạn vết thương, trích xuất đặc trưng và phân loại. Hơn nữa, các kỹ thuật ML đã trở thành một vấn đề gây tranh cãi cho BMC trong nghiên cứu xử lý hình ảnh y tế [[13]](#bookmark=id.2v0bq6xu0db4). Điều này là do các kỹ thuật ML truyền thống gặp vấn đề về hiện thực dữ liệu [[14]](#bookmark=id.1knv814xpaxx). Mặc dù các kỹ thuật ML truyền thống sử dụng phương pháp trích xuất đặc trưng thủ công, nhưng rất khó để thiết kế một phương pháp trích xuất đặc trưng hiệu quả do đa dạng mẫu [[15](#bookmark=id.rbupbbsnsim3),[16](#bookmark=id.yi9i6i5e37h9),[17](#bookmark=id.2qlrd8t7n8r3),[18](#bookmark=id.khrdfk3tepyt),[19](#bookmark=id.5bs8tpwvzwmy),[20]](#bookmark=id.u0l7btl4tatc). Trong vài năm qua, kỹ thuật DL đã thu hút sự quan tâm rộng rãi từ các nhà nghiên cứu trong các lĩnh vực khác nhau do khả năng trích xuất đặc trưng của DL được sử dụng rộng rãi trong sinh học tin học [[21](#bookmark=id.hoij7dpufhgu),[22](#bookmark=id.12yab9uwk44b),[23]](#bookmark=id.g5w2p4tp7np).

Kỹ thuật CNN (Convolutional Neural Network) bắt được mối quan hệ giữa đầu vào và đầu ra thông qua ánh xạ phi tuyến. Bạn cũng có thể sử dụng nhiều đặc trưng kết hợp trong một chương trình duy nhất. Sử dụng điểm đặc trưng sẽ làm thay đổi nhiều hơn so với việc xây dựng từ đầu. Kỹ thuật CNN đã được áp dụng vào các vấn đề xử lý hình ảnh sinh học (BIP) và phân đoạn ảnh y tế là một trong số đó [[24]](#bookmark=id.9vuwa51w2s1i). Kỹ thuật CNN đã liên tục đạt được kết quả tốt trong quá trình phân đoạn mass trên các bộ dữ liệu công cộng. Trong quá trình phân đoạn khối u, CNN sử dụng các pixel của khối ảnh làm đầu vào của mạng để phân loại. Điều này gây ra quá tải về dung lượng và hiệu suất ước tính thấp do tính toán lại của quá trình tích chập trong quá trình dự đoán và huấn luyện [[25]](#bookmark=id.srn3goga56bq).

### 1.1.2 Lý do chọn đề tài

Đề tài này liên quan đến một vấn đề sức khỏe quan trọng, đó là bệnh ung thư vú. Ung thư vú là một trong những căn bệnh phổ biến và nguy hiểm đến tính mạng của người phụ nữ. Việc chẩn đoán sớm và chính xác là rất quan trọng để tăng khả năng chữa trị và cứu sống bệnh nhân.

Phương pháp chụp ảnh X-quang tuyến vũ là một phương pháp chẩn đoán sớm rất hiệu quả. Tuy nhiên, việc đọc và phân loại ảnh X-quang tuyến vũ là công việc tốn nhiều thời gian và có thể bị ảnh hưởng bởi sự chênh lệch về kinh nghiệm và kiến thức của bác sĩ. Do đó, việc xây dựng một hệ thống phân loại mass trong nhũ ảnh từ ảnh X-quang sử dụng deep learning và ensemble learning sẽ giúp tăng độ chính xác và tiết kiệm thời gian cho quá trình chẩn đoán.

Đề tài này sử dụng các kỹ thuật tiên tiến như deep learning và ensemble learning, đây là các kỹ thuật mới nhất trong lĩnh vực máy học và trí tuệ nhân tạo, có khả năng học và tự động cải thiện hiệu suất của hệ thống. Việc áp dụng các công nghệ này vào đề tài sẽ giúp cho hệ thống phân loại tổn thương trở nên hiệu quả hơn.

Ngoài ra, đề tài này cũng có tính ứng dụng cao và tiềm năng phát triển trong lĩnh vực chẩn đoán bệnh ung thư vú. Hệ thống phân loại tổn thương có thể được sử dụng rộng rãi trong các cơ sở y tế để hỗ trợ bác sĩ trong quá trình chẩn đoán, tăng cường khả năng chữa trị và cứu sống cho các bệnh nhân mắc phải ung thư vú.

## 1.2 Tổng quan về bài toán

### 1.2.1. Định nghĩa

Bài toán phân loại tổn thương trong nhũ ảnh là một bài toán trong lĩnh vực y học, nhằm phân loại tổn thương hoặc bất thường trong hình ảnh chụp quang tuyến vũ. Bài toán được thực hiện thông qua việc phân loại tổn thương thành 3 nhãn: bình thường, lành tính hoặc ác tính.

Khối u là một tập hợp các tế bào phát triển không đồng nhất trong cơ thể, tạo thành một khối lượng mới. Khối này trong nhũ ảnh có thể là một triệu chứng của bệnh ung thư vú, một trong những loại ung thư phổ biến nhất ở phụ nữ. Việc phát hiện và phân loại khối trong nhũ ảnh có thể giúp các chuyên gia y tế đưa ra quyết định chẩn đoán và điều trị kịp thời, tăng khả năng điều trị và cải thiện kết quả điều trị cho bệnh nhân.

Để thực hiện bài toán này, ta sử dụng phát hiện hình ảnh của yolo để trích xuất đặc trưng, deep learning và ensemble learning là hai kỹ thuật máy học mạnh mẽ và hiệu quả trong việc trích suất đặc trưng ảnh y tế, kết hợp với mô hình random foresrt có thể giải quyết tốt bài toán phân loại.

### 1.2.2. Các vấn đề thách thức của bài toán

Các khối trong ảnh chụp quang tuyến vú có kích thước lớn yêu cầu nhiều lớp pooling, dẫn đến giảm độ chính xác của quá trình phân đoạn. Trong khi đó, việc sử dụng khối ảnh có kích thước nhỏ làm đầu vào giảm kích thước của lĩnh vực nhận thức. Những lĩnh vực nhận thức nhỏ này chỉ trích xuất các đặc trưng cục bộ và có độ chính xác phân đoạn hạn chế. Các nghiên cứu trước đây cũng đã áp dụng kỹ thuật CNN trong lĩnh vực xử lý ảnh chụp quang tuyến vú[[26]](#bookmark=id.uo8tj57uqfad). Mặc dù hệ thống CAD hiện có dựa trên kỹ thuật CNN đã đạt hiệu suất cao hơn so với hệ thống CAD dựa trên ML trong nhiệm vụ phân loại ảnh chụp quang tuyến vú, nhưng vẫn còn một số vấn đề tồn tại. Những vấn đề này được mô tả dưới đây:

Không thể tìm đặc trưng trong ảnh chụp quang tuyến vú có độ tương phản thấp. Kỹ thuật CNN giảm bớt công việc dự đoán trọng số cho tất cả các phần của hình ảnh. Thay vì phân tích toàn bộ hình ảnh, nó chỉ tập trung vào việc phân tích từng phần nhỏ trong hình ảnh hiện tại và đôi khi đưa ra quyết định dựa trên những phần này. Điều này giúp giảm độ phức tạp và tăng tính hiệu quả trong quá trình xử lý ảnh.

CNN không thể tự động tìm ra mối quan hệ giữa các phần khác nhau của hình ảnh. Do đó, phương sai độ sáng của các phần trong hình ảnh không được khôi phục một cách tối ưu.

### 1.2.3. Các phương pháp và kết quả nghiên cứu liên quan trong lĩnh vực

Sử dụng mô hình YOLO giúp tăng độ chính xác và tốc độ xử lý của quá trình lựa chọn ROI trong hệ thống CAD. Bằng cách tổng hợp thông tin từ các ROI được chọn, hệ thống có thể cung cấp kết quả chính xác hơn và hỗ trợ các bác sĩ trong quá trình chẩn đoán ung thư vú.[[56]](#bookmark=id.e6pbz7r956nm)

Phương pháp trích xuất đặc trưng và phân loại bằng các mô hình máy học truyền thống như k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), Random Forest (RF),... Ví dụ như nghiên cứu của Rajaraman et al. (2013) sử dụng phương pháp k-NN để phân loại nhũ ảnh với độ chính xác 80%. [[73]](#bookmark=id.mxwcyoxf0za9)

Phương pháp phân loại sử dụng mô hình Deep Learning như Convolutional Neural Network (CNN) được sử dụng rộng rãi trong các nghiên cứu phân loại ảnh y tế. Ví dụ như nghiên cứu của Wang et al. (2016) sử dụng mô hình CNN để phân loại nhũ ảnh với độ chính xác 80.9%. [[74]](#bookmark=id.fh7w43x0kcqr)

Kết hợp các phương pháp trích xuất đặc trưng và mô hình Deep Learning để phân loại như nghiên cứu của Zare et al. (2018) sử dụng phương pháp trích xuất đặc trưng Histogram of Oriented Gradients (HOG) kết hợp với mô hình CNN để phân loại nhũ ảnh với độ chính xác 86.2%. [[75]](#bookmark=id.qs3tthe8jlz8)

## 1.3. Mục tiêu nghiên cứu

Mục tiêu nghiên cứu của bài toán là xác định và phân loại chính xác khối u, để có thể xây dựng các hệ thống hỗ trợ cho bác sĩ trong quá trình đưa ra chuẩn đoán và phân loại khối u trên chụp quang tuyến vú.

Với mục tiêu đề ra ở trên chúng tôi sử dụng mô hình Yolov5 trích suất hình ảnh ROI, xây dựng mô hình deep learning dựa trên mạng nơ-ron tích chập (Convolutional Neural Network - CNN) để học và trích xuất các đặc trưng quan trọng từ những hình ảnh chụp quang tuyến vú. Áp dụng ensemble learning bằng cách kết hợp nhiều mô hình deep learning để tăng cường khả năng dự đoán và giảm thiểu sai số.

Đánh giá và so sánh hiệu suất của hệ thống phân loại được xây dựng với các phương pháp truyền thống và các nghiên cứu trước đây trong việc nhận diện và phân loại tổn thương trong chụp quang tuyến vú.

## 1.4. Phạm vi nghiên cứu

Tìm hiểu về mô hình Yolo trong trích xuất hình ảnh, các phương pháp Deep Learning, Ensemble Learning trong việc lấy ra các đặc trưng để giải quyết bài toán phân loại tổn thương trong chụp quang tuyến vú.

Xây dựng và huấn luyện các mô hình Deep Learning khác nhau, bao gồm các kiến trúc mạng neural tích chập (CNN) để phân loại tổn thương trong nhũ ảnh.

Áp dụng kỹ thuật Ensemble Learning để cải thiện độ chính xác và độ ổn định của hệ thống phân loại.

Đánh giá và so sánh hiệu suất của các mô hình Deep Learning và Ensemble Learning được xây dựng với các phương pháp truyền thống trong việc phân loại Mass trong chụp quang tuyến vú.

Đề xuất các giải pháp để cải thiện hiệu suất và độ chính xác của hệ thống phân loại.

## 1.5. Ý nghĩa khoa học và thực tiễn

Tăng hiểu biết về xử lý ảnh y khoa: Qua việc phát hiện và phân loại tổn thương trong hình ảnh chụp quang tuyến vú, nghiên cứu này cung cấp thông tin quan trọng về tính chất và đặc điểm của ảnh chụp quang tuyến vú. Điều này có thể góp phần vào việc tăng hiểu biết và kiến thức về phương pháp xử lý ảnh y khoa.

Hỗ trợ quyết định lâm sàng: Hệ thống phân loại tự động có thể cung cấp thông tin hữu ích để hỗ trợ quyết định lâm sàng của các chuyên gia y tế. Kết quả phân loại tổn thương dựa trên hình ảnh chụp quang tuyến vú có thể giúp các bác sĩ và nhân viên y tế đưa ra những quyết định chẩn đoán và điều trị chính xác hơn, tối ưu hóa quá trình chăm sóc bệnh nhân.

# CHƯƠNG II: CƠ SỞ LÝ THUYẾT

## 2.1. Bài toán detection yolo

### 2.1.1. Định nghĩa

Yolo là một cách tiếp cận khác để phát hiện đối tượng trong thời gian thực. Yolo sử dụng tính năng phát hiện đối tượng khung như một bài toán hồi quy cho các hộp giới hạn được phân tách theo không gian và xác suất lớp liên quan. Do đó, mạng nơ-ron đơn lẻ có thể dự đoán các hộp giới hạn và xác suất lớp trực tiếp từ hình ảnh đầy đủ trong một lần đánh giá. Kết quả là, Yolo cho thấy tốc độ đáng kinh ngạc trong việc phát hiện đối tượng trong thời gian thực và hữu ích cho các phát hiện đơn giản như nhận dạng khuôn mặt, nhận dạng bàn tay, v.v. Xử lý hình ảnh trong Yolo khá đơn giản, hệ thống thay đổi kích thước ảnh phù hợp với mô hình Yolo, cụ thể thể đối với bài toán hand detection ảnh sẽ có kích thước 614x614 , chạy một mạng phức hợp duy nhất trên hình ảnh và ngưỡng các phát hiện kết quả theo độ tin cậy của mô hình [[57]](#bookmark=id.jkc16sl9gs7a).

Yolo cực kỳ hữu ích vì nó cũng đào tạo trên hình ảnh đầy đủ và trực tiếp tối ưu hóa hiệu suất phát hiện. Yolo không yêu cầu bất kỳ đường ống phức tạp nào, vì nó là một vấn đề hồi quy để phát hiện khung. Ngoài ra, Yolo không sử dụng bất kỳ kỹ thuật dựa trên cửa sổ hoặc khu vực nào, nhưng nó sử dụng toàn bộ hình ảnh trong thời gian đào tạo và kết quả là mã hóa thông tin theo ngữ cảnh về các lớp [[57]](#bookmark=id.jkc16sl9gs7a). Tất cả những điều này giúp tạo ra Yolo siêu nhanh và do đó, chúng tôi đã lựa chọn Yolo cho bài báo cáo này.

Phiên bản của Yolo được tập chung thử nghiệm trong luận án này là Yolov5. Đây được coi là một bước cải tiến mang tính sáng tạo đối với Thị giác máy tính. Bây giờ nó thực hiện phát hiện các hộp ở ba quy mô khác nhau. Do đó, nó có thể không nhanh hơn các phiên bản trước của Yolo, nhưng chắc chắn là mạnh hơn và phát hiện hình ảnh nhỏ hơn tốt hơn. Trọng số ban đầu của Yolo được cung cấp sẵn. Sử dụng các trọng số này làm điểm bắt đầu cùng với các tệp cấu hình Yolos, chúng tôi có thể bắt đầu làm việc với các phát hiện đối tượng. Đây là một phương pháp rất dễ thực hiện và đào tạo, nhưng trước tiên phải chuẩn bị sẵn bộ dữ liệu của mình và nhãn cho phù hợp.2.1.2. Các phương pháp tiếp cận của Yolo

A picture containing screenshot, square

Description automatically generated

Hình 2. Hình ảnh minh họa sơ khai về Yolo

Yolo sẽ đưa ảnh đầu vào qua một CNN, tạo ra một bản đồ vùng (feature map) S ×SSS, gọi là grid. YOLO thực hiện nhận diện đối tượng tại mỗi ô trong S × SSS grid đó. Thay vì có một bước tìm ra các vùng có khả năng tồn tại đối tượng, thì Yolo sẽ thực hiện nhận diện trên toàn bộ ảnh.

Một ô sẽ dự đoán ra hộp giới hạn B và xác suất cho C lớp. Một hộp giới hạn sẽ mang 5 thông tin: tâm của hộp (x, y), chiều dài và rộng của Bounding Box (w, h) và độ tin cậy. Độ tin cậy dùng để xác định xem trong hộp giới hạn đó có tồn tại đối tượng hay không. Nếu trong hộp giới hạn đó không tồn tại đối tượng, ta bỏ qua toàn bộ giá trị dự đoán khác của hộp giới hạn và thực hiện phân loại, nhằm bỏ qua những nơi không tồn tại đối tượng.

Vậy, trong một ô, ta sẽ dự đoán ra một tensor có B ×5 +C phần tử, với B là số lượng Hộp giới hạn, 5 là 5 thông tin trong hộp gồm: tọa độ trung tâm, chiều dài rộng của hộp và độ tin cậy và C là số lớp. Và trong một bản đồ vùng gồm S × S ô, tensor dự đoán từ mạng sẽ có độ dài S ×S × (B ×5 +C).

### 2.1.2. Kiến trúcA picture containing text, diagram, font, sketch Description automatically generated

Hình 2. Kiến trúc Yolo

*Hình 2. SEQ Hình\_2.\_ \\* ARABIC 2 Kiến trúc mạng YOLO*

Yolo là một mô hình mạng CNN cho việc phát hiện, nhận dạng, phân loại đối tượng. Yolo được tạo ra từ việc kết hợp giữa các convolutional layers và connected layers. Trong đó các convolutional layers sẽ trích xuất ra các feature của ảnh, còn fully - connected layers sẽ dự đoán ra xác suất đó và tọa độ của đối tượng. Ở những Yolo phiên bản đầu tiên như Yolov1 và Yolov2 việc phát hiện đối tượng có kích thước trung bình và lớn tương đối tốt, nhưng đối với các đối tượng có kích thước nhỏ thì lại là điểm yếu của Yolov1 và Yolov2, mặc cho Yolov2 đã sử dụng skip-connection từ các layer trước đó để đưa thông tin nhưng vẫn không hiệu quả cho tới khi Yolov3 xuất hiện. Yolov3 là một sự nâng cấp cho vấn đề này, ở Yolov3 giải quyết được nhược điểm mà Yolov1 và v2 gặp phải. Bằng cách áp dụng Feature Pyramid Network thực hiện phát hiện hiện object ở 3 scale khác nhau Yolov3 đã hoàn toàn loại bỏ việc không nhận dạng được đối tượng có kích thước nhỏ ngoài ra một thay đổi nữa của Yolov3 mà tác giả cho là hiệu quả đó là việc sử dụng loss function được tính theo Binary Cross Entropy chứ không còn là squared loss như ở Yolov2 và Yolov1, và mạng backbone được sử dụng là darknet-53 có 53 layers lớn hơn nhiều so với 19 của YOLOv2 còn lại hầu hết phần kiến trúc của Yolov3 tương đối giống với Yolov2. Khác với các phiên bản Yolo trước, Yolov5 được các tác giả cải thiện ở các lớp layer cải thiện tốc độ.

*A picture containing text, screenshot, graphic design, design

Description automatically generated*

Hình 2. Kiến trúc Pyramid Network.

## 2.2. Mô hình yolov5

### 2.2.1. Tổng quan về mô hình

Yolov5 là một mô hình phát hiện đối tượng nhanh chóng và chính xác trong lĩnh vực thị giác máy tính. Nó là phiên bản cải tiến của mô hình Yolo (You Only Look Once) với nhiều cải tiến về kiến trúc và hiệu suất.

Mô hình Yolov5 sử dụng mạng neural tích chập để trích xuất đặc trưng từ hình ảnh và dự đoán các hộp giới hạn và lớp đối tượng tương ứng. Mạng này được xây dựng dựa trên mạng cơ sở CSPDarknet53, với các cải tiến như lớp PANet để kết hợp thông tin từ các tầng khác nhau và lớp SPP để tạo ra một phạm vi bắt đầu rộng hơn cho việc phát hiện.[[58]](#bookmark=id.lk5oxsxwrrbr)

### 2.2.2. Kiến trúc mạng yolov5

Điểm khác biệt của Yolov5 đó là cải tiến CSPResBlock thành một module mới, ít hơn một lớp Convolution gọi là C3 module thay vì sử dụng activate function Mish và LeakyRelu như các phiên bản Yolo trước, Yolov5 được sử dụng activate function SiLu cùng với việc được điều chỉnh lại module SPP thành SPPF thêm nữa là việc chạy tuần tự các MaxPooling và các Layer điều này giúp Yolov5 cải thiện tốc độ đáng kể.

**A picture containing text, screenshot, font, number

Description automatically generated**

Hình 2. Kiến trúc mạng Yolov5

Recall (TPR): Tỷ lệ của các kết quả dự đoán là “positive” trên tổng số mẫu “positive” thực tế, hay có nghĩa là có nghĩa là tỉ lệ nhãn dự đoán là dương tính trên tổng số nhãn được dự đoán là dương tính. Trong mô hình phân loại sẽ kỳ vọng chỉ số này lớn nhất.

Precision: Tỷ lệ của các kết quả dự đoán “positive” là đúng trên tổng số mẫu dự đoán là “positive”. Hay hiểu đơn giản có nghĩa là tỉ lệ nhãn dự đoán là dương tính trên tổng số nhãn dương tính thật. Trong mô hình phân loại sẽ kỳ vọng chỉ số này lớn nhất.

## 2.3. Bài toán phân loại (classification)

### 2.3.1. Khái niệm

Bài toán phân loại (classification) trong lĩnh vực học máy là quá trình gán nhãn cho một tập dữ liệu vào một số lớp hoặc nhãn khác nhau dựa trên các đặc trưng quan sát được. Mục tiêu của phân loại là xây dựng một mô hình hoặc một thuật toán có khả năng tự động học và nhận biết các mẫu dữ liệu mới, từ đó dự đoán và phân loại chúng vào các lớp đã biết trước.

Các phương pháp phân loại thường sử dụng các thuật toán máy học như Support Vector Machines (SVM)[[76]](#bookmark=id.r9l1eslv00h9), Decision Trees, Random Forests, Logistic Regression và Neural Networks. Các thuật toán này được huấn luyện trên dữ liệu đào tạo, trong đó mỗi mẫu dữ liệu đã được gán nhãn. Quá trình huấn luyện giúp mô hình học các quy tắc, mối quan hệ và biên giới phân chia giữa các lớp.

Khi mô hình đã được huấn luyện, nó có thể được áp dụng để phân loại các mẫu dữ liệu mới dựa trên các đặc trưng quan sát được. Mô hình sẽ đưa ra dự đoán về lớp hoặc nhãn tương ứng cho mỗi mẫu dữ liệu đó. Hiệu suất của mô hình phân loại được đánh giá bằng các độ đo như độ chính xác, độ phân giải, độ nhạy, độ đặc hiệu và F1-score.

Phân loại có rất nhiều ứng dụng trong thực tế, từ phân loại email vào thư rác, phát hiện gian lận tín dụng, phân loại ảnh và video, cho đến chẩn đoán bệnh trong lĩnh vực y tế. Các nghiên cứu tiếp tục tìm hiểu và phát triển các phương pháp phân loại mới, nhằm cải thiện độ chính xác và khả năng tổng quát hóa của các mô hình phân loại.

### 2.3.2. Các nghiên cứu đã có

Bài báo "A Few Useful Things to Know About Machine Learning" của Pedro Domingos năm 2012 tập trung vào những khái niệm và quy tắc quan trọng trong machine learning. Nghiên cứu giới thiệu một số khái niệm quan trọng như overfitting, bias-variance tradeoff, ensemble learning, feature engineering và cross-validation. Các khái niệm này đóng vai trò quan trọng trong quá trình xây dựng và đánh giá các mô hình machine learning. Bài báo cung cấp cái nhìn tổng quan về những khái niệm này và giúp người đọc hiểu rõ hơn về cách áp dụng chúng vào thực tế trong lĩnh vực machine learning [[77]](#bookmark=id.hrjogeupftfi). Nghiên cứu tập trung vào những khái niệm và quy tắc quan trọng trong machine learning, bao gồm overfitting, bias-variance tradeoff, ensemble learning, feature engineering, và cross-validation.

Bài báo "Support-Vector Networks" của Corinna Cortes và Vladimir Vapnik năm 1995 liên quan đến bài toán phân loại. Nghiên cứu này giới thiệu phương pháp Support Vector Machines (SVM), một thuật toán phân loại mạnh mẽ dựa trên việc tạo ra các siêu phẳng tối ưu để phân tách các điểm dữ liệu trong không gian đa chiều. SVM được áp dụng rộng rãi trong các bài toán phân loại như phân loại hình ảnh, phân loại văn bản, và nhận dạng mẫu. Bài báo giải thích cách SVM hoạt động và cung cấp các định nghĩa và công thức liên quan để xây dựng mô hình phân loại chính xác và hiệu quả [[78]](#bookmark=id.954onk55dw23). Một thuật toán phân loại dựa trên việc tạo ra các siêu phẳng tối ưu để phân tách các điểm dữ liệu trong không gian đa chiều.

Bài báo "The Elements of Statistical Learning" của Trevor Hastie, Robert Tibshirani và Jerome Friedman năm 2001 liên quan đến lĩnh vực phân loại. Cuốn sách này tập trung vào các phương pháp học máy thống kê, trong đó bao gồm phân loại là một khía cạnh quan trọng. Bài báo giới thiệu các khái niệm cơ bản và các phương pháp phân loại thông qua việc sử dụng các mô hình thống kê. Các tác giả trình bày các công thức, thuật toán và ví dụ minh họa để giúp người đọc hiểu và áp dụng các phương pháp phân loại trong thực tế. Bài báo này cung cấp kiến thức nền tảng và kỹ thuật cho việc xây dựng các mô hình phân loại dựa trên phân tích thống kê [[79]](#bookmark=id.k75fmgwt12pp). Cuốn sách tập trung vào các phương pháp học máy thống kê, bao gồm phân loại, hồi quy và các phương pháp dự báo dựa trên các mô hình thống kê.

Bài báo "Deep Residual Learning for Image Recognition" của Kaiming He, Xiangyu Zhang, Shaoqing Ren và Jian Sun (2016) tập trung vào việc giới thiệu kiến trúc mạng nơ-ron sâu có tên gọi ResNet trong lĩnh vực nhận diện ảnh. Nghiên cứu này tạo ra một cấu trúc mạng nơ-ron sâu đột phá bằng cách sử dụng các khối còn lại (residual blocks) để vượt qua vấn đề gradient vanishing/exploding trong quá trình huấn luyện. Bằng cách truyền thông tin qua các residual blocks, mạng ResNet có khả năng xây dựng các mô hình phức tạp hơn và đạt được độ chính xác cao hơn trong việc phân loại ảnh. Nghiên cứu này đã mang lại những đóng góp quan trọng và trở thành một bước tiến lớn trong lĩnh vực nhận diện ảnh và phân loại.[[80]](#bookmark=id.bc94frg63v2l). Nghiên cứu giới thiệu kiến trúc mạng nơ-ron sâu (deep neural network) có tên gọi ResNet, sử dụng các residual blocks để giải quyết vấn đề gradient vanishing/exploding và cải thiện hiệu suất phân loại ảnh.

Bài báo "Convolutional Neural Networks for Visual Recognition" của Alex Krizhevsky, Ilya Sutskever và Geoffrey E. Hinton (2012) tập trung vào việc giới thiệu mạng neural tích chập (CNN) có tên gọi AlexNet, một bước đột phá quan trọng trong lĩnh vực nhận diện và phân loại ảnh. Nghiên cứu này đã chứng minh rằng CNN có khả năng vượt trội trong việc xử lý ảnh và đạt được độ chính xác cao trong các tác vụ phân loại. Với việc áp dụng các layer tích chập và pooling, mạng AlexNet đã thể hiện khả năng học được các đặc trưng phức tạp và tự động trích xuất thông tin quan trọng từ ảnh. Kết quả nghiên cứu đã mở ra cánh cửa cho sự phát triển mạnh mẽ của CNN trong lĩnh vực nhận diện và phân loại ảnh và đã tạo ra những tiến bộ đáng kể trong các ứng dụng thực tế [[81]](#bookmark=id.u37giep2vkhl). Bài báo giới thiệu mạng neural tích chập (CNN) AlexNet, được áp dụng rộng rãi trong phân loại và nhận dạng ảnh.

Bài báo "Random Forests" của Leo Breiman (2001) tập trung vào giới thiệu thuật toán Random Forest, một phương pháp quan trọng trong lĩnh vực classification. Nghiên cứu này đề xuất sử dụng một tập hợp các cây quyết định ngẫu nhiên, trong đó mỗi cây được huấn luyện trên một tập dữ liệu con và sử dụng các đặc trưng ngẫu nhiên. Qua quá trình huấn luyện, các cây trong Random Forest sẽ đưa ra các dự đoán độc lập và sau đó kết hợp kết quả từ tất cả các cây để đưa ra kết quả cuối cùng. Phương pháp này giúp giảm overfitting, tăng tính ổn định và độ chính xác của mô hình phân loại. Random Forest đã được chứng minh là hiệu quả trong nhiều bài toán phân loại khác nhau và đã trở thành một công cụ hữu ích trong lĩnh vực machine learning [[82]](#bookmark=id.ymkdi15mp687). Nghiên cứu giới thiệu thuật toán Random Forest, một phương pháp ensemble learning dựa trên việc xây dựng nhiều cây quyết định ngẫu nhiên và kết hợp kết quả từ các cây để đưa ra dự đoán.

Bài viết tóm tắt về thuật toán Gradient Boosting Machines (GBM), một phương pháp ensemble learning mạnh mẽ dựa trên việc xây dựng các mô hình yếu và kết hợp chúng để tạo ra một mô hình dự đoán mạnh hơn [[62]](#bookmark=id.27mkxwl7rpq4). Thông qua các ví dụ minh họa và đánh giá kết quả trên các tập dữ liệu, bài viết cung cấp cho người đọc những kiến thức cơ bản về GBM, bao gồm cách hoạt động của thuật toán, cách xác định các tham số quan trọng và các kỹ thuật tuning để tăng hiệu suất của GBM. Bên cạnh đó, bài viết cũng đề cập đến một số ứng dụng của GBM trong các bài toán phân loại và hồi quy.

Bài báo tập trung vào phương pháp phân loại văn bản sử dụng mô hình Naive Bayes, một phương pháp đơn giản nhưng hiệu quả trong việc xác định xác suất phân loại của một văn bản dựa trên các thuộc tính quan sát được. Mô hình Naive Bayes giả định rằng các thuộc tính đầu vào là độc lập với nhau và áp dụng công thức Bayes để tính toán xác suất xảy ra của mỗi lớp phân loại. Bài báo cung cấp một tầm nhìn tổng quan về mô hình Naive Bayes và cách áp dụng nó trong bài toán phân loại văn bản, bao gồm cả ứng dụng thực tế và các kỹ thuật tối ưu hóa để cải thiện hiệu suất của mô hình [[69]](#bookmark=id.ovn8s0qt9mds). Nghiên cứu tập trung vào phân loại văn bản sử dụng mô hình Naive Bayes, một phương pháp dựa trên giả định độc lập giữa các thuộc tính và áp dụng công thức Bayes để tính xác suất phân loại.

Bài báo tập trung vào phương pháp hồi quy logistic và nhấn mạnh các giới hạn và hạn chế của phương pháp này. Tác giả trình bày những khía cạnh mà chúng ta không thể thực hiện được như mong đợi khi sử dụng hồi quy logistic, cũng như các vấn đề thường gặp trong quá trình ước lượng. Bài báo cung cấp một cái nhìn sâu hơn về các thách thức mà chúng ta gặp phải khi áp dụng hồi quy logistic và đề xuất các giải pháp và cải tiến để khắc phục những hạn chế đó [[70]](#bookmark=id.iuz8pncotxal). Bài báo tập trung vào phương pháp hồi quy logistic, nhấn mạnh các giới hạn và hạn chế của phương pháp và đề xuất cải tiến và giải pháp cho các vấn đề xảy ra trong quá trình ước lượng.

Bài báo tập trung vào việc giới thiệu thuật toán cây quyết định, một phương pháp phân loại dựa trên quyết định đơn giản. Tác giả trình bày cách sử dụng cây quyết định để tạo ra các quy tắc quyết định từ dữ liệu. Thuật toán cây quyết định hoạt động bằng cách xây dựng một cây có cấu trúc phân nhánh dựa trên các luật IF-THEN, giúp tạo ra các quy tắc phân loại dễ hiểu và dễ sử dụng. Bài báo này đã đóng góp quan trọng vào lĩnh vực phân loại bằng cách giới thiệu một phương pháp đơn giản và hiệu quả để đưa ra quyết định dựa trên dữ liệu [[52]](#bookmark=id.hyhxt7wtgnuh).

## 2.4. Mô hình LSTM

### 2.4.1. Định nghĩa

LSTM (Long Short-Term Memory) là một kiểu mạng nơ-ron học sâu thuộc lớp mạng RNN (Recurrent Neural Networks). Mạng LSTM được thiết kế để giải quyết vấn đề mất mát thông tin khi huấn luyện mạng RNN truyền thống. LSTM sử dụng các cấu trúc đặc biệt gọi là cổng (gate) để kiểm soát thông tin được lưu trữ và xóa bỏ thông A diagram of a tank

Description automatically generated with low confidencetin không quan trọng. [[53]](#bookmark=id.hwb1w6ef4d2c)

Hình 2. Mô hình LSTM.

LSTM là một phiên bản mở rộng của mạng RNN, được đề xuất vào năm 1997 bởi Sepp Hochreiter và Jürgen Schmidhuber. LSTM được thiết kế để giải quyết các bài toán về phụ thuộc xa (long-term dependencies) trong mạng RNN do bị ảnh hưởng bởi vấn đề gradient biến mất.

Giả sử khi xem một bộ phim dài tập, chúng ta ghi nhớ bối cảnh phim đã diễn ra ở những tập trước đó, kết hợp xử lý với thông tin của tập phim hiện tại hoặc khi đọc sách, chúng ta ghi nhớ điều gì đã xảy ra ở chương trước, kết hợp thành mạch thông tin để hiểu và tiếp thu cho nội dung hiện tại. Tương tự như vậy, khi các mạng RNN hoạt động, thông tin trước đó được ghi nhớ và sử dụng lại để xử lý cho đầu vào hiện tại. Tuy nhiên thì mạng RNN không thể ghi nhớ thông tin ở các bước có khoảng cách khá xa trước đó do vấn đề gradient biến mất. Do đó những phần tử đầu tiên trong chuỗi đầu vào không có nhiều ảnh hưởng đến các kết quả tính toán dự đoán phần tử cho chuỗi đầu ra trong các bước sau. Mạng LSTM với các kết nối phản hồi (feedback connection) giúp khắc phục nhược điểm này.

Mạng LSTM có thể bao gồm nhiều tế bào LSTM (LSTM memory cell) liên kết với nhau và kiến trúc cụ thể của mỗi tế bào. Ý tưởng của LSTM là bổ sung thêm trạng thái bên trong tế bào (cell internal state) st và 3 cổng sàng lọc các thông tin đầu vào và đầu ra cho tế bào bao gồm forget gate ft, input gate it và output gate ot. Tại mỗi bước thời gian t, các cổng đều lần lượt nhận giá trị đầu vào xt (đại diện cho một phần tử trong chuỗi đầu vào) và giá trị ht-1 có được từ đầu ra của memory cell từ bước thời gian trước đó t – 1. Các cổng đều đóng vai trò có nhiệm vụ sẵn sàng lọc thông tin với mục đích khác nhau:

* Forget gate: Có nhiệm vụ loại bỏ những thông tin không cần thiết được nhận khỏi cell internal state.
* Input gate: Có nhiệm vụ chọn lọc những thông tin cần thiết nào được thêm vào cell internal state.
* Output gate: Có nhiệm vụ xác định những thông tin nào từ cell internal state được sử dụng như đầu ra.

### 2.2.2. Các bước trích xuất đặc trưng ảnh từ mô hình LSTM

Mô hình LSTM được sử dụng để trích xuất đặc trưng của ảnh. Quá trình này bao gồm các bước sau:

Đầu vào của mô hình là các ảnh RGB, được chuyển đổi thành định dạng grayscale và kích thước đồng nhất.

Ảnh được chia thành các patch (vùng hình chữ nhật) với kích thước và bước nhảy được xác định trước.

Mỗi patch được đưa vào một mạng LSTM, trong đó mỗi ô của mạng LSTM tương ứng với một vị trí trong patch.

Quá trình huấn luyện mạng LSTM được sử dụng để học các đặc trưng ẩn của từng patch. Các đặc trưng này có thể được xem như là một vectơ số học, được truyền đến một mạng Fully Connected để phân loại. [[54]](#bookmark=id.nx9efqt3tzjk)

## 2.5. Mô hình resNet50

### 2.5.1. Định nghĩa

ResNet50 là một phiên bản của mô hình ResNet, với 50 lớp tích chập và được huấn luyện trên tập dữ liệu ImageNet. Mô hình sử dụng kiến trúc mạng nơ-ron học sâu, với các khối residual để giúp mô hình học được những đặc trưng phức tạp của ảnh.

Đặc điểm của ResNet50 là sử dụng các khối residual để giảm thiểu sự triệt tiêu độ sâu, giúp cho việc huấn luyện các mô hình sâu trở nên dễ dàng hơn. Ngoài ra, ResNet50 cũng sử dụng kỹ thuật "bottleneck" để giảm chi phí tính toán và kích thước mô hình, đồng thời tăng tính hiệu quả của mô hình. [[27]](#bookmark=id.gii3y6yagsp5)

### 2.5.2 Kiến trúc mạng ResNet

Giải pháp mà ResNet đưa ra là sử dụng kết nối "tắt" đồng nhất để xuyên qua một hay nhiều lớp. Một khối như vậy được gọi là một Residual Block, như trong hình sau: A diagram of a diagram

Description automatically generated with low confidence

Hình 2. Single Residual Block.

ResNet (Residual Network) là một kiến trúc mạng neural sâu được đề xuất nhằm giải quyết vấn đề của độ sâu trong việc huấn luyện mạng. Kiến trúc này giúp giảm hiện tượng biến mất đạo hàm và tăng khả năng học các đặc trưng phức tạp.

Kiến trúc ResNet còn sử dụng các khối Residual Block được xếp chồng lên nhau và trở thành các tầng (layer) trong mạng. Ngoài ra, để giảm kích thước của đầu ra và tăng số lượng đặc trưng, mạng cũng sử dụng các lớp gộp (pooling layer) và lớp tích chập với bước nhảy (stride convolution).

Với kiến trúc này, ResNet đã chứng minh khả năng huấn luyện mạng sâu hiệu quả, vượt qua được vấn độ sâu của ảnh và mang lại kết quả tốt trong lĩnh vực thị giác máy tính, như nhận dạng đối tượng và phân loại ảnh.

2.5.3. Xây dựng mạng ResNet-50A picture containing screenshot, colorfulness, graphics, graphic design

Description automatically generated

Hình 2. Kiến trúc mạng ResNet50.

"ID BLOCK" trong hình trên là viết tắt của từ Identity block và ID BLOCK x3 nghĩa là có 3 khối Identity block chồng lên nhau. Nội dung hình trên như sau:

* Zero-padding: Input với (3,3)
* Stage 1: Tích chập (Conv1) với 64 filters với shape (7,7), sử dụng stride (2,2). BatchNorm, Max Pooling (3,3).
* Stage 2: Convolutiontal block sử dụng 3 filter với size 64x64x256, f=3, s=1. Có 2 Identity blocks với filter size 64x64x256, f=3.
* Stage 3: Convolutional sử dụng 3 filter size 128x128x512, f=3, s=2. Có 3 Identity blocks với filter size 128x128x512, f=3.
* Stage 4: Convolutional sử dụng 3 filter size 256x256x1024, f=3, s=2. Có 5 Identity blocks với filter size 256x256x1024, f=3.
* Stage 5: Convolutional sử dụng 3 filter size 512x512x2048, f=3, s=2. Có 2 Identity blocks với filter size 512x512x2048, f=3.
* The 2D Average Pooling: sử dụng với kích thước (2,2).
* The Flatten.
* Fully Connected (Dense): sử dụng softmax activation.

## 2.6. Cây quyết định (Decision Tree)

### 2.6.1. Định nghĩa

Decision Tree là một mô hình học máy trong đó dữ liệu được phân loại hoặc dự đoán dựa trên quyết định được đưa ra từ việc tạo ra một cây quyết định logic. Cây quyết định bao gồm một loạt các nút đại diện cho các quyết định hoặc điều kiện, các nhánh đại diện cho các kết quả của các quyết định hoặc điều kiện đó, và các lá cây đại diện cho kết quả cuối cùng hoặc dự đoán.

Người dùng có thể đi qua các nút trên cây bằng cách áp dụng các quy tắc hoặc điều kiện dựa trên các thuộc tính của dữ liệu đầu vào. Mỗi quyết định hoặc điều kiện sẽ dẫn đến một nhánh khác nhau trên cây, cho đến khi đạt được kết quả cuối cùng hoặc dự đoán.

Decision Tree được sử dụng trong nhiều bài toán khác nhau như phân loại và dự đoán trong học máy và khai thác dữ liệu [[29]](#bookmark=id.qnv1dtdx599j).

Mô hình cây quyết định đơn giản và dễ hiểu, giúp tạo ra các quy tắc quyết định dễ giải thích và có thể áp dụng trong nhiều lĩnh vực khác nhau. Các quyết định được thực hiện dựa trên giá trị của các thuộc tính và các ràng buộc được đặt ra trong quá trình xây dựng cây.

Một nguồn tham khảo phổ biến về cây quyết định là sách "Classification and Regression Trees" của Leo Breiman, Jerome Friedman, Richard Olshen và Charles Stone. Được xuất bản vào năm 1984, sách này tập trung vào mô hình CART (Classification and Regression Trees) [[30]](#bookmark=id.erua0ob7hpo7) và cung cấp các thuật toán chi tiết để xây dựng cây quyết định trong cả hai trường hợp phân loại và hồi quy. Sách này đã trở thành một tài liệu cơ bản trong lĩnh vực cây quyết định và có sự ảnh hưởng lớn trong cộng đồng học máy [[31]](#bookmark=id.hydemtru2r6d).

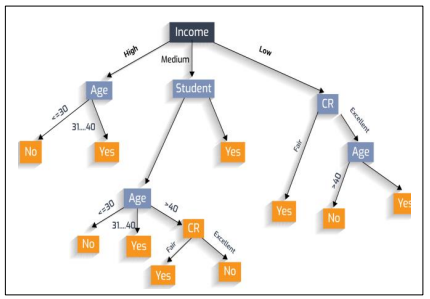
### 2.6.2. Mô hình cây quyết định

Một trong những kỹ thuật được sử dụng rộng rãi trong khai phá dữ liệu là hệ thống tạo ra bộ phân loại [[32]](#bookmark=id.gerkn9ppnpm2). Trong khai phá dữ liệu, thuật toán phân loại có khả năng xử lý một lượng thông tin lớn. Nó có thể được sử dụng để đưa ra giả định về tên lớp phân loại, phân loại kiến thức dựa trên các tập huấn luyện và nhãn lớp, và phân loại dữ liệu mới thu được [[33]](#bookmark=id.24ag4x3x05ba). Trong học máy, thuật toán phân loại chứa nhiều thuật toán, và trong công trình này, bài báo tập trung vào thuật toán cây quyết định nói A diagram of a tree

Description automatically generated with low confidencechung. Hình 9 mô tả cấu trúc của cây quyết định (DT).

Hình 2. Decision Tree [34]

Cây quyết định là một trong những phương pháp mạnh mẽ thông thường được sử dụng trong các lĩnh vực khác nhau, như học máy, xử lý hình ảnh và nhận dạng mẫu [[35]](#bookmark=id.trwk7i36hjdt). Cây quyết định là một mô hình liên tục kết hợp một loạt các kiểm tra cơ bản một cách hiệu quả và nhất quán, trong đó mỗi kiểm tra so sánh một đặc trưng số với một ngưỡng giá trị [[36]](#bookmark=id.yd8zy48qnk86). Các quy tắc khái niệm dễ dàng được xây dựng hơn so với trọng số số học trong mạng neural của các kết nối giữa các nút [[37](#bookmark=id.glyuvimxgntq), [38]](#bookmark=id.cr7maud6tmuo). Thông thường, cây quyết định được sử dụng cho mục đích nhóm hóa. Hơn nữa, cây quyết định là một mô hình phân loại thường được sử dụng trong Khai phá dữ liệu [[39]](#bookmark=id.5n881hkirtpo). Các nút và nhánh được tạo thành từ mỗi cây. Mỗi nút đại diện cho các đặc trưng trong một danh mục cần được phân loại và mỗi tập con xác định một giá trị mà nút có thể nhận [[40](#bookmark=id.eq0njlw5luw8), [41]](#bookmark=id.7ivvus9o1r9r). Do tính đơn giản trong phân tích và độ chính xác đối với nhiều dạng dữ liệu khác nhau, cây quyết định đã tìm thấy nhiều lĩnh vực ứng dụng [[42]](#bookmark=id.bwse896v6jfx).

Hình 10 cho thấy một ví dụ về cây quyết định.

Hình 2. Hình minh họa về cây quyết định.[43]

### 2.6.3. Các loại thuật toán cây quyết định

Có một số loại thuật toán Cây quyết định như: Iterative Dichotomies 3 (ID3), phiên bản kế tiếp của ID3 (C4.5), Cây phân loại và hồi quy (CART) [[44]](#bookmark=id.jzwuqcv765io), CHi-squared Automatic Interaction Detector (CHAID) [[45]](#bookmark=id.cga9gwdpvnw9), Multivariate Adaptive Regression Splines (MARS) [[46]](#bookmark=id.kecxfcon5a0x), Generalized, Unbiased, Interaction Detection and Estimation (GUIDE), Conditional Inference Trees (CTREE) [[47]](#bookmark=id.ft9n9wix1qgx), [[48]](#bookmark=id.5rkylwjq0wwq), Classification Rule with Unbiased Interaction Selection and Estimation (CRUISE), Quick, Unbiased and Efficient Statistical Tree (QUEST) [[49]](#bookmark=id.xpqaw328hi78), [[50]](#bookmark=id.vnr44sj2fmqh). Bảng I hiển thị sự so sánh giữa các thuật toán thường được sử dụng cho cây quyết định [[51]](#bookmark=id.3sjdj2anv820).

## 2.7. Rừng ngẫu nhiên (Random Forest)

### 2.7.1 Định nghĩa

Random Forest là một thuật toán học máy phổ biến thuộc về kỹ thuật học có giám sát. Nó có thể được sử dụng cho cả vấn đề Phân loại và Hồi quy trong ML. Nó dựa trên khái niệm học tập theo nhóm, là một phương pháp kết hợp nhiều cây quyết định (decision trees) để tạo ra một mô hình dự đoán mạnh mẽ.

### 2.7.2 Mô hình Random Forest

A picture containing text, circle, diagram, screenshot

Description automatically generatedNói một cách đơn giản: rừng ngẫu nhiên xây dựng nhiều cây quyết định và kết hợp chúng lại với nhau để có được dự đoán chính xác và ổn định hơn. Một lợi thế lớn của rừng ngẫu nhiên là nó có thể được sử dụng cho cả các bài toán phân loại và hồi quy, vốn tạo nên phần lớn các hệ thống học máy hiện nay. Rừng ngẫu nhiên đã được chứng minh là tạo ra lỗi kiểm tra thấp hơn so với cây quyết định thông thường và hiệu suất tương đương với SVM trong các vấn đề nhiều lớp, trong khi vẫn duy trì hiệu quả tính toán cao. [83]

Hình 2. Mô hình Random Forest

# 

# CHƯƠNG 3: DỮ LIỆU

## 3.1. Bộ dữ liệu DDSM

Cơ sở dữ liệu kỹ thuật số để chụp nhũ ảnh sàng lọc (DDSM)[[57]](#bookmark=id.jkc16sl9gs7a) là một nguồn tài nguyên để cộng đồng nghiên cứu phân tích hình ảnh chụp nhũ ảnh sử dụng. Hỗ trợ chính cho dự án này là một khoản trợ cấp từ Chương trình Nghiên cứu Ung thư Vú của Bộ Tư lệnh Nghiên cứu Y tế và Vật tư Quân đội Hoa Kỳ. Dự án DDSM là một nỗ lực hợp tác có sự tham gia của các đồng nghiệp tại Bệnh viện Đa khoa Massachusetts (D. Kopans, R. Moore), Đại học Nam Florida (K. Bowyer) và Phòng thí nghiệm Quốc gia Sandia (P. Kegelmeyer). Các trường hợp bổ sung từ Trường Y khoa Đại học Washington được cung cấp bởi Peter E. Shile, MD, Trợ lý Giáo sư X quang và Nội khoa. Các tổ chức hợp tác bổ sung bao gồm Trường Y Đại học Wake Forest (Khoa Kỹ thuật Y tế và Phóng xạ học), Bệnh viện Thánh Tâm và ISMD, Incorporated. Mục đích chính của cơ sở dữ liệu là tạo điều kiện thuận lợi cho nghiên cứu hợp lý trong việc phát triển các thuật toán máy tính để hỗ trợ sàng lọc. Các mục đích phụ của cơ sở dữ liệu có thể bao gồm việc phát triển các thuật toán để hỗ trợ chẩn đoán và phát triển các công cụ hỗ trợ giảng dạy hoặc đào tạo. Cơ sở dữ liệu chứa khoảng 2.500 nghiên cứu. Mỗi nghiên cứu bao gồm hai hình ảnh của mỗi bên vú, cùng với một số thông tin liên quan đến bệnh nhân (tuổi tại thời điểm nghiên cứu, đánh giá mật độ vú ACR, đánh giá mức độ tinh tế đối với các bất thường, mô tả từ khóa ACR về các bất thường) và thông tin hình ảnh (máy quét, độ phân giải không gian, ...). Hình ảnh chứa các khu vực đáng ngờ có thông tin "sự thật cơ bản" ở cấp độ pixel về vị trí và loại khu vực đáng ngờ.

Bộ dữ liệu DDSM (Digital Database for Screening Mammography) là một tập hợp các hình ảnh và dữ liệu liên quan được sử dụng trong lĩnh vực chẩn đoán ung thư vú. DDSM được phát triển bởi Đại học South Florida và Đại học Chicago nhằm cung cấp một nguồn tài nguyên quan trọng cho nghiên cứu và phát triển các phương pháp tự động hóa trong phân loại, nhận dạng và phát hiện ung thư vú từ hình ảnh chụp cộng hưởng từ (mammography).

Bộ dữ liệu DDSM bao gồm hàng ngàn hình ảnh mammography kết hợp với các thông tin về kết quả kiểm tra và các loại khối u trong vú. Các hình ảnh trong DDSM có độ phân giải cao và được chụp từ nhiều góc độ khác nhau để phục vụ cho nhiều mục đích nghiên cứu.

DDSM cung cấp thông tin chi tiết về mỗi hình ảnh mammography, bao gồm kích thước, loại u, vị trí và đánh dấu các vùng nghi ngờ. Bộ dữ liệu cũng đi kèm với các kết quả chẩn đoán từ các chuyên gia để phục vụ cho việc so sánh và đánh giá các phương pháp tự động.

Bộ dữ liệu DDSM đã đóng góp quan trọng cho việc phát triển các thuật toán và công cụ phân tích ảnh y khoa trong chẩn đoán và giám sát ung thư vú. Nó cung cấp một nền tảng để nghiên cứu và so sánh hiệu quả của các phương pháp phân loại và nhận dạng ung thư vú từ hình ảnh mammography.

## 3.2. Xử lý dữ liệu

### 3.2.1. Ngưỡng otsu (otsu thresholding)

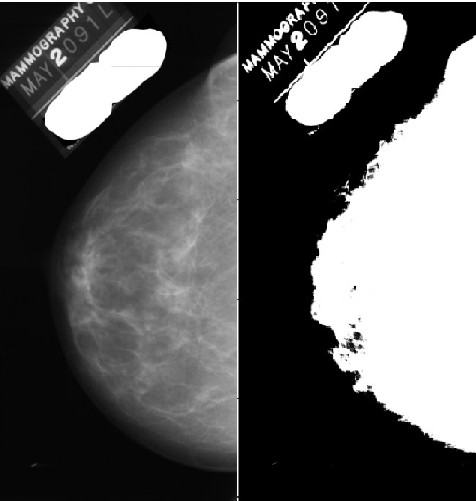
Phương pháp này được đề xuất bởi Nobuyuki Otsu vào năm 1979, dựa trên việc tìm ngưỡng phù hợp cho việc phân ngưỡng ảnh bằng cách tối đa hóa độ phân biệt giữa các lớp của ảnh. Đầu tiên, Otsu tính toán histogram của ảnh để xác định tần suất xuất hiện của mỗi mức xám. Sau đó, Otsu tìm kiếm ngưỡng phù hợp bằng cách tối đa hóa độ phân biệt giữa các lớp ảnh được chia thành trước và sau khi phân ngưỡng.

Cụ thể, phương pháp Otsu thresholding thực hiện các bước sau:

* Tính toán histogram của ảnh để xác định tần suất xuất hiện của mỗi mức xám.
* Tính tổng tất cả tần suất của histogram.
* Tính tổng tích lũy của histogram và tần suất trung bình của mỗi mức xám.
* Tính giá trị trung bình toàn cục và tổng tích lũy của giá trị trung bình trong histogram.
* Tính độ phân tán giữa hai lớp được chia thành trước và sau khi phân ngưỡng.
* Tìm ngưỡng phân ngưỡng bằng cách tối đa hóa độ phân biệt giữa các lớp ảnh được chia thành trước và sau khi phân ngưỡng. [[60]](#bookmark=id.n2ghseo0543r)

Để tìm giá trị ngưỡng phù hợp, phương pháp Otsu tính toán tổng phương sai giữa hai lớp đối tượng và nền trên toàn bộ các giá trị ngưỡng trong khoảng giá trị từ 0 đến 255. Giá trị ngưỡng tối ưu được xác định bằng cách tìm giá trị ngưỡng tối đa mà tổng phương sai đạt giá trị nhỏ nhất. Sau đó, ảnh được phân loại thành hai lớp dựa trên giá trị ngưỡng này.

Phương pháp Otsu thresholding là một trong những phương pháp phân ngưỡng tự động phổ biến được sử dụng trong xử lý ảnh, đặc biệt trong các ứng dụng y tế và nhận dạng khuôn mặt. [[61]](#bookmark=id.meeg22v0mzr4)



Hình 3. Hình minh họa về Otsu thresholding

### 3.2.2. Xử lý hình thái (Morphological Processing)

Morphological Image Processing là tập hợp các phép toán phi tuyến tính (non-linear) tác động đến hình dạng hoặc hình thái của các điểm nhị phân trong ảnh. Dựa trên các phép toán AND OR XOR NOT để biến đổi các điểm nhị phân.

A diagram of a crossword puzzle

Description automatically generated with low confidenceMorphological Operator sẽ sử dụng một cấu trúc nhỏ quét qua và áp dụng toàn bộ ảnh được gọi là Structuring Element (SE). SE sẽ quét qua toàn bộ vị trí trong hình ảnh và so sánh với các pixel lân cận nằm trong vùng sẽ tương ứng.

Hình 3. Ví dụ về Morphological Operator

Structuring Element là một ảnh nhị phân nhỏ. Tức là một ma trận nhỏ gồm các pixel mang giá trị 0 và 1:

* Kích thước của ma trận xác định kích thước của SE.
* Pixel có giá trị 0 được bỏ qua trong quá trình tính toán.
* Luôn có một pixel làm mốc trong ma trận SE.

Structuring Element hoạt động giống như ma trận tích chập (convolution kernel) trong lọc ảnh tuyến tính (linear image filtering). SE sẽ dịch chuyển toàn bộ ảnh với điểm mốc.

Phép co (Erosion) của ảnh A bởi phần tử cấu trúc B cho ra ảnh G. Mỗi vị trí cấu trúc B quét qua được sẽ chọn giá trị nhỏ nhất và trả về điểm mốc tương ứng trên ảnh G.

Với thứ tự áp dụng phép giãn và co như thế nào để có thể lọc được các nhiễu như các hình ảnh dưới đây?

A picture containing black and white

Description automatically generated with low confidence

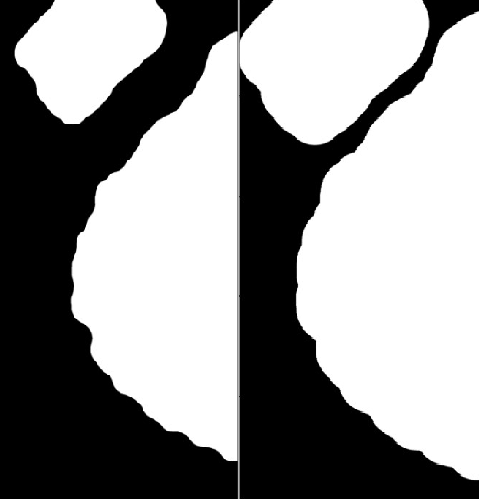
Hình 3. Hình minh họa về Opening và Closing.

Opening là thực hiện phép co rồi bắt đầu thực hiện phép giãn với một cấu trúc. Giúp làm mượt các đường viền, phá vỡ các khe nhỏ, loại bỏ các đối tượng nhỏ, làm mượt các đỉnh lồi.

Closing là thực hiện phép giãn rồi bắt đầu thực hiện phép co với một cấu trúc. Giúp làm mượt các đường viền, loại bỏ các lỗ nhỏ, làm mượt các đỉnh khe hẹp.

Phép giãn (Dilation) của ảnh A bởi phần tử cấu trúc B cho ra ảnh G. Mỗi vị trí cấu trúc B quét qua được sẽ chọn giá trị lớn nhất và trả về điểm mốc tương ứng trên ảnh G.

Ví dụ có một ảnh nhị phân với các dòng chữ bị đứt nét. Sử dụng phép giãn với cấu trúc dạng chữ thập sẽ nối liền các điểm đứt nét và các điểm đó sẽ có hơi hướng giống với cấu trúc chữ thập.



Hình 3. Hình minh họa về phép giãn (dilation)

### 3.2.3. Áp dụng hình thái gần để loại bỏ các vùng nhỏ (Apply morphology close to remove small regions)

Morphology close là một phương pháp xử lý hình ảnh trong lĩnh vực xử lý ảnh số, thường được sử dụng để loại bỏ các vùng nhỏ hoặc đối tượng không mong muốn khỏi ảnh. Phương pháp này được thực hiện bằng cách áp dụng một phép toán đóng trên ảnh ban đầu, nghĩa là thực hiện một phép co lại theo hình dạng của các đối tượng trên ảnh, sau đó loại bỏ các vùng nhỏ hơn một ngưỡng kích thước cho trước bằng phép mở. Kết quả là các vùng nhỏ bị loại bỏ và chỉ các đối tượng lớn hơn được giữ lại.

Ví dụ, trong xử lý ảnh y học, phương pháp này có thể được sử dụng để loại bỏ các mảnh ghép nhỏ không mong muốn trong hình ảnh MRI hoặc CT scan, giúp cải thiện độ chính xác của việc chẩn đoán và định vị các khối u hoặc bệnh lý khác trong cơ thể.

Phương pháp morphology close được áp dụng để loại bỏ các vùng nhỏ không mong muốn trong hình ảnh mạng mống mắt (retinal fundus image). Cụ thể, hình ảnh được xử lý bằng cách sử dụng phép toán đóng (close) với một phần tử cấu trúc (structuring element) hình vuông có kích thước lớn. Sau đó, kết quả của phép toán đóng được trừ đi hình ảnh gốc để loại bỏ các vùng nhỏ không mong muốn, giữ lại các vùng lớn hơn và chứa thông tin về mạng mống mắt. Kết quả của phương pháp được chứng minh qua thực nghiệm trên tập dữ liệu hình ảnh mạng mống mắt và cho thấy hiệu quả tốt trong việc phát hiện mạng mống mắt tự động. [[63]](#bookmark=id.hgivvp6dbu57)

#### 3.2.3.1. Ưu điểm

Phương pháp morphology close giúp loại bỏ các lỗ nhỏ, hình dạng không đều và các đối tượng nhiễu trong hình ảnh MRI.

Phương pháp này có thể giúp kết hợp các đặc trưng của mô bên trong vùng tế bào u để giúp phân loại các tế bào u thành hai nhóm là u ác tính và u lành tính.

#### 3.2.3.2. Nhược điểm

Phương pháp này có thể xóa mất các đối tượng quan trọng có kích thước nhỏ hơn kích thước của phần tử morphology close.

Sử dụng phương pháp này có thể dẫn đến mất mát thông tin và bị đánh giá thấp về độ chính xác khi giải quyết các vấn đề liên quan đến biên giới của hình ảnh. [[64]](#bookmark=id.nfsobkl8lwrr)

### 3.2.4. Áp dụng hình thái mở để tách vú khỏi các vùng khác (Apply morphology open to separate breast from other regions)

Morphology open là một phương pháp xử lý ảnh sử dụng kỹ thuật hình thái học để tách nhũ khỏi các vùng khác trong hình ảnh. Cụ thể, phương pháp này sử dụng phép toán morphology open, một kỹ thuật thường được sử dụng để loại bỏ các đối tượng nhỏ không mong muốn trong hình ảnh.

Trong quá trình xử lý, phép toán morphology open được áp dụng trên hình ảnh để loại bỏ các chi tiết nhỏ không cần thiết trong ảnh. Sau đó, vùng nhũ sẽ được phân biệt rõ ràng và tách ra khỏi các vùng khác trong hình ảnh.

Phương pháp này đã được áp dụng trong nhiều nghiên cứu và ứng dụng trong lĩnh vực y tế, đặc biệt là trong chẩn đoán ung thư vú từ hình ảnh chụp quang tuyến vú.

Phương pháp morphology open để tách vùng nhũ khỏi các vùng khác trong các hình ảnh chụp quang tuyến vú. Cụ thể, phương pháp này được sử dụng để xử lý các hình ảnh từ các thiết bị chụp quang tuyến vú để giúp phát hiện ung thư vú. Bằng cách áp dụng phép toán morphology open, các vùng liền kề nhau có độ tương đồng màu sắc và độ sáng sẽ được phân tách ra, giúp cho các kết quả phân tích hình ảnh sau này được chính xác hơn và đáng tin cậy hơn. [[67]](#bookmark=id.kefdz94gtfae)

#### 3.2.4.1. Ưu điểm

Phương pháp morphology open được sử dụng để tách vùng vú ra khỏi các vùng khác trong hình ảnh mammography, giúp cải thiện độ chính xác của việc phát hiện khối u vú.

Phương pháp đơn giản, dễ hiểu và thực hiện.

#### 3.2.4.2. Nhược điểm

Phương pháp morphology open chỉ phù hợp với trường hợp các vùng vú có độ tách biệt rõ ràng với các vùng khác trong hình ảnh. Nếu có các vùng chồng lấn lên nhau, việc sử dụng morphology open có thể không hiệu quả hoặc dẫn đến mất mát thông tin.

Sử dụng morphology open cũng có thể gây ra hiện tượng mất mát các đối tượng nhỏ nằm trong vùng cần tách, vì phương pháp này có thể loại bỏ những đối tượng nhỏ hơn một ngưỡng kích thước nhất định. [[66]](#bookmark=id.zhkk2ibxz52r)

### 3.2.5. Đường viền lớn nhất (Largest contour)

"Get largest contour" là một phương pháp được sử dụng trong xử lý ảnh để tìm đường viền lớn nhất trong một vùng ảnh cụ thể. Phương pháp này thường được sử dụng để phát hiện và phân tách các đối tượng trong ảnh.

Để thực hiện phương pháp này, trước hết cần phải áp dụng các kỹ thuật xử lý ảnh như lọc ảnh, chuyển đổi không gian màu, hoặc phát hiện cạnh để tạo ra một ảnh nhị phân. Sau đó, áp dụng các phép biến đổi hình thái học để loại bỏ nhiễu và cải thiện đường viền của các đối tượng trong ảnh.

Tiếp theo, sử dụng thuật toán tìm kiếm đường viền để tìm ra các đường viền trong ảnh và sau đó tìm ra đường viền lớn nhất. Đường viền này có thể được sử dụng để phát hiện và phân tách các đối tượng trong ảnh.

Phương pháp "Get largest contour" được sử dụng trong nhiều ứng dụng xử lý ảnh như phân tích y học, nhận dạng khuôn mặt, nhận dạng ký tự, nhận dạng đối tượng, v.v.

Để có được đường viền lớn nhất, các bước sau được thực hiện:

* Chuyển đổi ảnh màu sang không gian màu HSV để dễ dàng phân đoạn các vùng màu.
* Sử dụng thuật toán Graph Cuts để phân đoạn ảnh và tạo ra một mask.
* Áp dụng phép toán morphology để xử lý mask và loại bỏ các vùng nhỏ.
* Sử dụng thuật toán Active Contour để có được đường viền lớn nhất của vùng phân đoạn.

Do đó, "get largest contour" ở đây được hiểu là việc sử dụng thuật toán Active Contour để tìm ra đường viền lớn nhất của vùng được phân đoạn. [65]

#### 3.2.5.1. Ưu điểm

Phương pháp đơn giản và dễ hiểu, có thể áp dụng cho nhiều bài toán khác nhau.

Cho kết quả chính xác và nhanh chóng nếu ảnh đầu vào có đường viền rõ ràng.

#### 3.2.5.2. Nhược điểm

Không thể xử lý được những trường hợp ảnh đầu vào có nhiễu hoặc đường viền không rõ ràng.

Chỉ cho kết quả tốt khi đường viền lớn nhất là duy nhất, trong trường hợp có nhiều đường viền lớn, phương pháp này không cho kết quả chính xác.[[68]](#bookmark=id.m4zblzlvidxt)

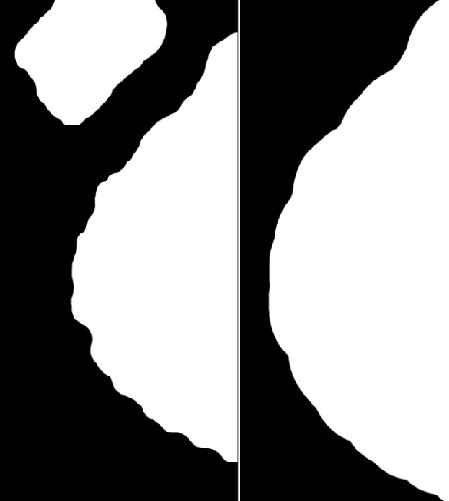
#### 3.2.6. Vẽ đường viền lớn nhất (Draw the largest contour)

"Draw largest contour as white filled on black background as mask" là một bước trong xử lý ảnh, nó áp dụng phương pháp tìm đường viền lớn nhất (largest contour) để vẽ đường viền này lên ảnh đầu vào. Sau đó, toàn bộ khu vực bên trong đường viền sẽ được tô màu trắng, còn các khu vực ngoài đường viền sẽ được tô màu đen. Kết quả của bước này sẽ tạo ra một ảnh mặt nạ (mask) với đường viền lớn nhất nằm giữa các khu vực có màu trắng và các khu vực nền đen.

"Draw largest contour as white filled on black background as mask" được sử dụng để trích xuất đặc trưng khuôn mặt sử dụng Local Binary Patterns (LBP) cho bài toán nhận diện khuôn mặt.

Phương pháp này bao gồm các bước sau:

* Bước 1: Chuyển ảnh màu của khuôn mặt thành ảnh xám.
* Bước 2: Áp dụng thuật toán LBP lên ảnh xám để tạo ra một bản đồ đặc trưng.
* Bước 3: Sử dụng phép toán morphology "closing" để loại bỏ các lỗ nhỏ trong bản đồ đặc trưng, nối các vùng gần nhau lại với nhau.
* Bước 4: Áp dụng thuật toán contour để tìm ra đường viền lớn nhất của các vùng kết nối trong bản đồ đặc trưng đã được xử lý bằng phép toán morphology.
* Bước 5: Vẽ đường viền lớn nhất trên nền đen, tạo thành một mặt nạ (mask) với các vùng đen bao quanh đường viền và vùng trắng tạo thành đường viền.

Mask này sẽ được sử dụng để loại bỏ các phần không cần thiết trong khuôn mặt, chỉ giữ lại phần nằm trong đường viền lớn nhất để trích xuất đặc trưng cho việc nhận diện khuôn mặt. [[71]](#bookmark=id.ynzmsrcb7vf8)

Hình 3. Hình minh họa lấy đường viền có diện tích lớn nhất

#### 3.2.6.1. Ưu điểm

Phương pháp này có thể tách đối tượng lung nodule từ nền ảnh CT một cách rõ ràng, tạo ra một mask trắng đơn giản mô tả vị trí của khối u phổi trong ảnh CT.

Nó có thể sử dụng được cho các hình dạng đối tượng khác nhau, không giới hạn bởi hình dạng hay kích thước của đối tượng.

#### 3.2.6.2. Nhược điểm

Phương pháp này chỉ tạo ra một mask đơn giản với đường viền lớn nhất, bỏ qua các chi tiết nhỏ hơn của đối tượng lung nodule trong ảnh CT.

Nếu có nhiều đối tượng lung nodule trong ảnh CT, phương pháp này chỉ tách được đối tượng có đường viền lớn nhất và bỏ qua các đối tượng khác. [[72]](#bookmark=id.9x8vc1nux4b6)

### 3.2.9. Biến đổi màu sắc bằng bảng tra cứu (Look-Up Table)

Hàm LUT trong thư viện OpenCV được sử dụng để thực hiện ánh xạ đầu vào của hình ảnh thông qua một bảng tra cứu (lookup table). Đầu vào có thể là một hình ảnh grayscale hoặc hình ảnh màu.

Công dụng chính của LUT là thay đổi giá trị màu sắc hoặc cường độ của các điểm ảnh trong hình ảnh dựa trên ánh xạ được xác định trong bảng tra cứu. Bảng tra cứu này chỉ định cách ánh xạ từ giá trị đầu vào (ví dụ: giá trị màu sắc hoặc cường độ) đến giá trị đầu ra tương ứng.

Khi áp dụng hàm LUT, từng giá trị điểm ảnh trong hình ảnh đầu vào sẽ được tra cứu trong bảng tra cứu để tìm giá trị tương ứng. Kết quả là hình ảnh đầu ra có các giá trị điểm ảnh đã được thay đổi dựa trên ánh xạ trong bảng tra cứu.

Hàm LUT có nhiều ứng dụng trong xử lý hình ảnh, bao gồm việc thay đổi độ tương phản, điều chỉnh độ sáng tối, biến đổi màu sắc và hiệu ứng hình ảnh khác. Nó cho phép chúng ta tùy chỉnh và điều chỉnh hình ảnh dựa trên ánh xạ tùy ý, tạo ra các hiệu ứng hình ảnh độc đáo và nâng cao chất lượng hình ảnh.

Hàm LUT còn có thể được sử dụng để thực hiện các phép biến đổi tùy chỉnh trên hình ảnh. Bằng cách xác định một bảng tra cứu phù hợp, chúng ta có thể ánh xạ các giá trị điểm ảnh đầu vào sang các giá trị đầu ra mong muốn. Điều này cho phép chúng ta thực hiện các phép biến đổi không tuyến tính, tạo ra các hiệu ứng hình ảnh độc đáo và phức tạp hơn.

Hàm LUT thường được sử dụng trong các tác vụ như việc điều chỉnh độ tương phản, cân bằng màu sắc, đảo ngược hình ảnh (inversion), tạo hiệu ứng tone mapping và nhiều ứng dụng khác. Nó cung cấp một cách linh hoạt để thay đổi hình ảnh theo ý muốn của chúng ta, mở ra nhiều khả năng sáng tạo và tùy chỉnh trong xử lý hình ảnh.

Tuy hàm LUT không phổ biến như một số hàm xử lý hình ảnh khác, nhưng nó là một công cụ mạnh mẽ và hữu ích cho các tác vụ xử lý hình ảnh tùy chỉnh và phức tạp. Việc hiểu và sử dụng đúng cách hàm này sẽ giúp chúng ta thực hiện các biến đổi và hiệu ứng hình ảnh đa dạng và chính xác hơn trong quá trình xử lý hình ảnh. A picture containing moon, medical imaging

Description automatically generated

Hình 3. Hình minh họa sử dụng hàm LUT

### 3.2.10 Lật ảnh

Hàm flip trong thư viện OpenCV có tác dụng làm đảo ngược hình ảnh hoặc video theo chiều ngang hoặc chiều dọc. Bằng cách áp dụng hàm này, chúng ta có thể tạo ra hiệu ứng lật hình hoặc lật video một cách dễ dàng.

Hàm flip nhận vào hai tham số chính. Tham số đầu tiên là đối tượng ảnh hoặc video đầu vào mà chúng ta muốn lật. Tham số thứ hai là flipCode, là mã xác định hướng lật mà chúng ta muốn áp dụng. Có ba mã chính được sử dụng:

* flipCode = 0: Lật ảnh hoặc video theo chiều ngang. Điều này có nghĩa là hình ảnh hoặc video sẽ được lật qua trục đứng, mọi thứ từ bên phải sẽ chuyển sang bên trái và ngược lại.
* flipCode > 0: Lật ảnh hoặc video theo chiều dọc. Khi áp dụng mã này, hình ảnh hoặc video sẽ được lật qua trục ngang, mọi thứ từ trên xuống sẽ chuyển sang dưới và ngược lại.
* flipCode < 0: Lật ảnh hoặc video cả hai chiều, tức là lật cả theo chiều ngang và chiều dọc.

Bằng cách sử dụng hàm cv2.flip, chúng ta có thể thay đổi hướng và vị trí của hình ảnh hoặc video, tạo ra các hiệu ứng thú vị và cải thiện khả năng xử lý ảnh trong các ứng dụng thị giác máy tính và xử lý video.

Khi chúng ta cần thay đổi góc nhìn của hình ảnh hoặc video, chẳng hạn như lật ngược hoặc lật dọc, chúng ta có thể sử dụng flip để đảo ngược hình ảnh theo các hướng khác nhau.

Trong quá trình huấn luyện mô hình máy học, có thể yêu cầu sự đa dạng trong dữ liệu huấn luyện. Bằng cách lật hình ảnh hoặc video, chúng ta có thể tạo ra thêm phiên bản dữ liệu huấn luyện đối xứng hoặc nghịch đảo, từ đó tăng tính đa dạng của dữ liệu.

Điều chỉnh góc nhìn và hướng di chuyển: Trong một số trường hợp, chúng ta có thể muốn thay đổi góc nhìn hoặc hướng di chuyển của đối tượng trong hình ảnh. Bằng cách sử dụng flip, chúng ta có thể làm điều này một cách dễ dàng.

Cải thiện độ chính xác của mô hình: Trong một số trường hợp, việc lật hình ảnh hoặc video có thể cải thiện độ chính xác của mô hình phân loại hoặc nhận dạng. Điều này đặc biệt hữu ích khi đối tượng có tính chất đối xứng hoặc khi mô hình đang gặp khó khăn trong việc phân biệt giữa hai hướng.

A close-up of a breast

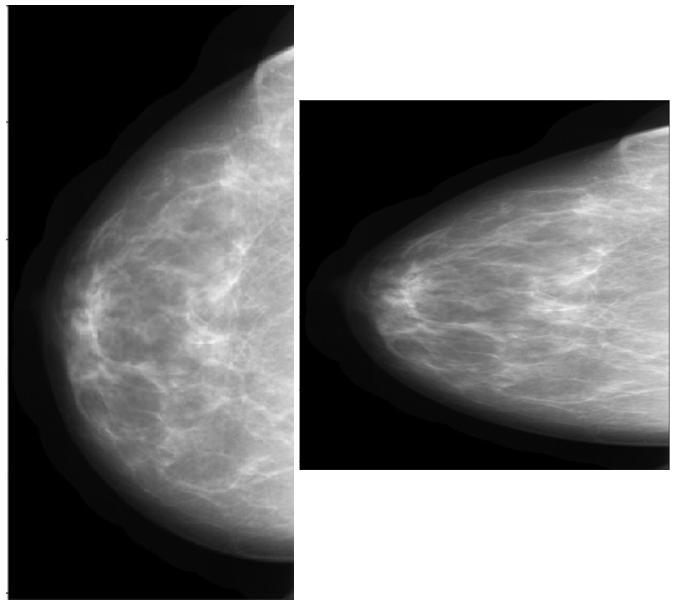
Description automatically generated with low confidenceHiệu ứng và xử lý hình ảnh sáng tạo: Bên cạnh các ứng dụng thực tế, hàm flip cũng có thể được sử dụng để tạo ra các hiệu ứng sáng tạo trên hình ảnh. Bằng cách lật và kết hợp hình ảnh theo các hướng khác nhau, chúng ta có thể tạo ra các hiệu ứng độc đáo và khám phá các khía cạnh mới của hình ảnh.

Hình 3. Hình minh họa về lật ảnh.

### 3.2.11 Thay đổi kích thước ảnh

Đầu tiên, hàm resize trong OpenCV cho phép chúng ta điều chỉnh kích thước hình ảnh theo tỷ lệ hoặc kích thước cụ thể. Điều này rất hữu ích khi chúng ta cần thay đổi kích thước để phù hợp với yêu cầu cụ thể của ứng dụng hoặc mô hình. Chẳng hạn, khi làm việc với mô hình nhận dạng đối tượng, chúng ta có thể cần đưa tất cả các hình ảnh về cùng kích thước để đảm bảo tính nhất quán và hiệu quả của quá trình nhận dạng.

Thứ hai, resize cũng hữu ích trong việc chuẩn bị dữ liệu huấn luyện cho mô hình máy học. Trong quá trình huấn luyện, việc có dữ liệu huấn luyện có cùng kích thước là rất quan trọng. Bằng cách sử dụng resize, chúng ta có thể điều chỉnh kích thước của các hình ảnh huấn luyện để đáp ứng yêu cầu này. Điều này giúp đảm bảo tính nhất quán và đồng đều của dữ liệu huấn luyện và cải thiện hiệu suất của mô hình.

Ngoài ra, resize còn cho phép chúng ta tạo các phiên bản thu nhỏ hoặc phóng to của hình ảnh. Điều này có thể hữu ích trong nhiều ứng dụng, từ việc tạo bản thu nhỏ cho hiển thị trước (thumbnail) đến việc phóng to để phân tích chi tiết. Đồng thời, resize cũng có thể được sử dụng để giảm dung lượng hình ảnh bằng cách giảm số lượng pixel, từ đó tiết kiệm bộ nhớ hoặc tăng tốc độ xử lý.

Hình 3. Hình minh họa về resize ảnh.

### 3.2.12. Contrast Limited Adaptive Histogram Equalization (CLAHE)

Phương pháp Cân bằng lược đồ lịch sử có giới hạn và thích ứng (CLAHE) [55] được sử dụng để cải thiện hình ảnh chụp quang tuyến vú. Việc cải thiện hình ảnh đóng vai trò quan trọng trong hình ảnh y học vì nó cho phép chúng ta nhìn thấy các đặc trưng ẩn trong một hình ảnh. Vì vậy, trong nghiên cứu này, CLAHE được sử dụng để tăng cường ảnh chụp quang tuyến vú và mang lại kết quả tốt nhất trong việc xem phần u tuyến. CLAHE hoạt động bằng cách chia hình ảnh thành các ô nhỏ, được gọi là "tiles", thay vì sử dụng toàn bộ hình ảnh. Các ô lân cận được kết hợp với nhau bằng phương pháp nội suy tuyến tính để loại bỏ các ranh giới giả mạo. Khi sử dụng CLAHE, có hai yếu tố cần lưu ý. Thứ nhất là giới hạn cắt, xác định ngưỡng đối sáng mặc định là 40. Thứ hai là kích thước lưới ô, xác định số lượng ô trong hàng và cột, giá trị mặc định là 8x8. Hình ảnh ở hình 3.9 cho thấy một hình ảnh trước và sau khi áp dụng CLAHE.A picture containing moon, astronomical object, nature

Description automatically generated

Hình 3. Hình minh họa cân bằng ngưỡng CLAHE

*Hình 3. SEQ Hình\_3. \\* ARABIC 9 Hình minh họa trước và sau khi áp dụng CLAHE*

Hàm boundingRect trong thư viện OpenCV được sử dụng để tính toán và xác định hình chữ nhật bao quanh một tập hợp các điểm dữ liệu. Đầu vào của hàm là một tập hợp các điểm, chẳng hạn như các điểm biên của một đối tượng trong hình ảnh.

Thuật toán hoạt động bằng cách tìm các giá trị x, y, width và height tương ứng với hình chữ nhật bao quanh tập hợp các điểm. Giá trị x và y đại diện cho tọa độ góc trái trên của hình chữ nhật, trong khi width và height biểu thị kích thước của hình chữ nhật.

Hàm boundingRect rất hữu ích trong nhiều ứng dụng xử lý ảnh như phân đoạn đối tượng, nhận diện và theo dõi đối tượng, tính toán vị trí và kích thước của đối tượng trong ảnh. Kết quả trả về từ hàm cung cấp thông tin quan trọng về hình dạng và vị trí của đối tượng, giúp cho việc tiếp tục xử lý và phân tích dữ liệu dễ dàng hơn.

A picture containing moon, astronomical object, astronomy, celestial event

Description automatically generated

Hình 3. Hình minh họa xác định bounding box

## 3.3. Hiện thực mô hình

### 3.3.1 Phần cứng được sử dụng

Quá trình thực nghiệm và huấn luyện mô hình được thực hiện trên kaggle với thống số phần cứng như b

Bảng 4. Thông số ở cứng

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Thành phần phần cứng | Năm ra mắt | Số lõi | Bộ nhớ GPU | Số giờ khả dụng trên tuần |
| NVIDIA Tesla P100 GPU | 2016 | 3584 Cuda cores | 16 GB | 30-32 h |

### 3.3.2 Dữ liệu huấn luyện

Dữ liệu từ bộ CBIS DDSM sau khi được tiền xử lý gồm ảnh có kích thước (512x512) và nhãn bounding box của từng ảnh, được chia thành 3 tập để huấn luyện với tập train 4840 ảnh, validation 1680 ảnh và tập test 620 ảnh.

Bảng 4. thống kê dữ liệu huấn luyện Yolo

|  |  |
| --- | --- |
| Tập dữ liệu | Số lượng ảnh |
| Train | 4840 |
| Val | 1680 |
| Test | 620 |
| Tổng số | 7140 |

Những ảnh có vùng tổn thương được cắt ROI để phân loại là lành tính hay ác tính.

Bảng 4. Thống kê dữ liệu mô hình phân loại

|  |  |  |  |
| --- | --- | --- | --- |
| Tập dữ liệu | Số lượng nhãn Cancer | Số lượng nhãn Begign | Tổng số |
| Train | 1923 | 1866 | 3789 |
| Test | 473 | 475 | 948 |
| Tổng số | 1923 | 2341 | 4264 |

### 3.3.3 Huấn luyện mô hình Yolov5l

Chuẩn bị file cấu hình data gồm đường dẫn đến thư mục train, thư mục test và thư mục validation với số lớp là 2 gồm bình thường và tổn thương

* Kích thước ảnh đầu vào 512 x 512
* Batch size 16
* Epochs là 100
* Learning rate: 0.01

### 3.3.4 Huấn luyện mô hình phân loại

Sử dụng kiến trúc Resnet và phương pháp học transfer learning ImageNet để trích xuất các tính năng cấp thấp.

Sử dụng mô hình LSTM để trích xuất các đặc điểm ngữ nghĩa cấp cao.

Sau đó các tính năng này được kết nối lại với nhau.

Mô hình CNN được thiết kế để trích xuất các đặc trưng liên quan từ ma trận đặc trưng trọng số.

Đầu ra sử dụng mô hình Gradient Boosting để phân loại ảnh là tổng thương lành tính và tổn thương ác tính.

# CHƯƠNG IV: KẾT QUẢ

## 4.1 Kết quả mô hình Yolov5

Bảng 4. Kết quả mô hình Yolo

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Mô hình | Nhãn | Xử lý ảnh | Kích thước ảnh | Precision  (%) | Recall  (%) | mAP .5  (%) | mAP .95  (%) |
| Yolov5l | Mass | Không | 256 | 33 | 22 | 17 | 5 |
| Yolov5l | Mass | Tăng cường | 512 | 37 | 26 | 23 | 7 |
| Yolov5l | Mass | Tăng cường + ScaleABs | 512 | 31 | 15 | 18 | 7 |
| Yolov5l | Mass | Tăng cường + Clahe | 512 | 37 | 14 | 18 | 6 |
| Yolov5l | Mass | Tăng cường + LUT | 512 | 97 | 70 | 84 | 75 |
| Yolov5l | Normal | Tăng cường + LUT | 512 | 97 | 98 | 98.7 | 97 |

Mô hình Yolov5l được huấn luyện qua tập dữ liệu với nhiều phương pháp xử lý hình ảnh khác nhau trong đó phương pháp đem lại kết quả tốt nhất là khử trắng (LUT), chỉ số mAP 0.5 cho nhãn dữ liệu Mass là 84% và cho nhãn Normal là 98.7%

A picture containing text, screenshot, plot, line

Description automatically generatedMột số hình ảnh về quá trình huấn luyện mô hình:

Hình 4. Đồ thị biểu thị độ lỗi trong quá trình huấn luyện

## 4.2 Kết quả mô hình phân loại

Bảng 4. Két quả mô hình phân loại

|  |  |
| --- | --- |
| Mô hình phân loại | độ chính xác (%) |
| Random Forest | 60 |
| Resnet50 | 73 |

Xây dựng mô hình phân loại dựa trên bài báo gốc chỉ đạt độ chính xác 60% trong khi đó mô hình Resnet50 được đề suất cho quá trình phân loại đạt độ chính xác là 73%

# CHƯƠNG V: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

## 5.1 Kết luận

### 5.1.1 Kết quả thu được

Trong quá trình thực, chúng tôi đã đạt được một số kết quả khả quan trong quá trình phân loại nhũ ảnh có bị tổn thương hay không và phát hiện được vùng tổn thương đó. Công trình luận văn của chúng tôi tập trung vào việc áp dụng kỹ thuật phân đoạn cho hình ảnh y khoa nhưng chúng tôi chưa đạt được kết quả tốt đối với giai đoạn phân loại tổn thương.

Một số nguyên nhân có thể làm ảnh hưởng đến giai đoạn phân loại tổn thương là kích thước hình ảnh không đồng đều, dữ liệu còn thiếu độ đa dạng dẫn đến mô hình không thể tổng quát hóa tốt đối với dữ liệu mới, hơn hết cấu trúc phức tạp và đa dạng của ảnh chụp quang tuyến vú là một thách thức lớn.

### 5.1.2 Kiến thức

Qua luận văn, tôi đã học hỏi được thêm nhiều kiến thức về kỹ thuật xử lý hình ảnh y khoa như tăng cường ảnh, cân bằng ngưỡng, tăng cường độ sáng, khử trắng, cài đặt mô hình Yolov5l, mô hình phân loại.

### 5.1.3 Kỹ năng

Thông qua việc thực hiện luận văn đã góp phần nâng cao khả năng thực nghiệm, đọc và tham khảo nghiên cứu khác từ đó chúng tôi tìm ra các phương pháp góp phần cải thiện độ chính xác của mô hình phân đoạn và phân loại ảnh.

## 5.2 Hướng phát triển

Trong tương lai chúng tôi sẽ tiếp tục thử nghiệm trên các bộ dữ liệu nhũ ảnh khác để có thể so sánh và đánh giá độ ổn định của mô hình, tăng cường thêm khả năng phát hiện vùng tổn thương của mô hình yolov5l, tìm hiểu các phương pháp mới để cải thiện độ chính xác của mô hình phân loại tổn thương nhũ ảnh. Bên cạnh đó thay vì detect chúng tôi muốn sử dụng Yolo để segmentation vùng tổn thương.

TÀI LIỆU THAM KHẢO

[1] J. S. Whang, S. R. Baker, R. Patel, L. Luk, and A. Castro, ‘‘The causes of medical malpractice suits against radiologists in the United States,’’Radiology, vol. 266, no. 2, pp. 548–554, Feb. 2013, doi: 10. 1148/radiol.12111119.

[2] A. Felman, ‘‘What to know about breast cancer,’’Med. News Today, 2019, [Online]. Available: https://www.medicalnewstoday.com/articles/37136

[3] W. Chen, R. Zheng, S. Zhang, H. Zeng, T. Zuo, C. Xia, Z. Yang, and J. He, ‘‘Cancer incidence and mortality in China in 2013: An analysis based on urbanization level,’’ Chin. J. Cancer Res., vol. 29, no. 1, pp. 1–10, 2017. 3

[4] S. Xu, H. Wu, and R. Bie, ‘‘CXNet-m1: Anomaly detection on chest X-rays with image-based deep learning,’’ IEEE Access, vol. 7, pp. 4466–4477, 2019, doi: 10.1109/access.2018.2885997.

[5] C. E. DeSantis, J. Ma, A. G. Sauer, L. A. Newman, and A. Jemal, ‘‘Breast cancer statistics, 2017, racial disparity in mortality by state,’’ CA, Cancer J. Clinicians, vol. 67, no. 6, pp. 439–448, Nov. 2017, doi: 10. 3322/caac.21412.

[6] M. Dong, X. Lu, Y. Ma, Y. Guo, Y. Ma, and K. Wang, ‘‘An efficient approach for automated mass segmentation and classification in mammograms,’’ J. Digit. Imag., vol. 28, no. 5, pp. 613–625, Oct. 2015.

[7] O. Akin, S. B. Brennan, D. D. Dershaw, M. S. Ginsberg, M. J. Gollub, H. Schoder, D. M. Panicek, and H. Hricak, ‘‘Advances in oncologic imaging: Update on 5 common cancers,’’ CA, Cancer J. Clinicians, vol. 62, no. 6, pp. 364–393, Oct. 2012.

[8] H. Min, S. S. Chandra, N. Dhungel, S. Crozier, and A. P. Bradley, ‘‘Multi-scale mass segmentation for mammograms via cascaded random forests,’’ in Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI), Apr. 2017, pp. 113–117.

[9] M. Dong, Z. Wang, C. Dong, X. Mu, and Y. Ma, ‘‘Classification of region of interest in mammograms using dual contourlet transform and improved KNN,’’ J. Sensors, vol. 2017, pp. 1–15, Nov. 2017. 10.1155/2017/3213680.

[10] J. Tang, R. M. Rangayyan, J. Xu, I. El Naqa, and Y. Yang, ‘‘Computeraided detection and diagnosis of breast cancer with mammography: Recent advances,’’ IEEE Trans. Inf. Technol. Biomed., vol. 13, no. 2, pp. 236–251, Mar. 2009, doi: 10.1109/titb.2008.2009441.

[11] J. Wei, B. Sahiner, L. M. Hadjiiski, H.-P. Chan, N. Petrick, M. A. Helvie, M. A. Roubidoux, J. Ge, and C. Zhou, ‘‘Computer-aided detection of breast masses on full field digital mammograms,’’ Med. Phys., vol. 32, no. 9, pp. 2827–2838, Aug. 2005, doi: 10.1118/1.1997327.

[12] Y. Akhtar and D. P. Mukherjee, ‘‘Detection of architectural distortion from the ridges in a digitized mammogram,’’ Signal, Image Video Process., vol. 12, pp. 1285–1292, Mar. 2018. 10.1007/s11760-018-1281-1.

[13] 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, Jul. 2019, Art. no. e14464, doi: 10.2196/14464.

[14] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, ‘‘A survey of sparse representation: Algorithms and applications,’’ IEEE Access, vol. 3, pp. 490–530, 2015.

[15] Y. A. Reyad, M. A. Berbar, and M. Hussain, ‘‘Comparison of statistical, LBP, and multi-resolution analysis features for breast mass classification,’’ J. Med. Syst., vol. 38, no. 9, p. 100, Sep. 2014, doi: 10.1007/s10916-014- 0100-7.

[16] P. S. Mukherjee, ‘‘A multi-resolution and adaptive 3-D image denoising framework with applications in medical imaging,’’ Signal, Image Video Process., vol. 11, no. 7, pp. 1379–1387, Oct. 2017.

[17] Y.-Q. Song, Z. Liu, J.-M. Chen, F. Zhu, and C.-H. Xie, ‘‘Medical image segmentation based on non-parametric mixture models with spatial information,’’ Signal, Image Video Process., vol. 6, no. 4, pp. 569–578, Nov. 2012.

[18] T. Ojala, M. Pietikainen, and T. Maenpaa, ‘‘Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp. 971–987, Jul. 2002, doi: 10.1109/tpami.2002.1017623.

[19] M. R. Al-Hadidi, A. Alarabeyyat, and M. Alhanahnah, ‘‘Breast cancer detection using K-Nearest neighbor machine learning algorithm,’’ in Proc. 9th Int. Conf. Develop. eSystems Eng. (DeSE), Aug. 2016, pp. 35–39.

[20] J. Wen, Y. Xu, and H. Liu, ‘‘Incomplete multiview spectral clustering with adaptive graph learning,’’ IEEE Trans. Cybern., vol. 50, no. 4, pp. 1418–1429, Apr. 2020.

[21] S. J. S. Gardezi, M. Awais, I. Faye, and F. Meriaudeau, ‘‘Mammogram classification using deep learning features,’’ in Proc. IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA), Sep. 2017, pp. 485–488.

[22] M. A. Jaffar, ‘‘Deep learning based computer aided diagnosis system for breast mammograms,’’ Int. J. Adv. Comput. Sci. Appl., vol. 8, no. 7, pp. 286–290, 2017.

[23] N. Antropova, B. Q. Huynh, and M. L. Giger, ‘‘A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets,’’ Med. Phys., vol. 44, no. 10, pp. 5162–5171, Oct. 2017, doi: 10.1002/mp.12453.

[24] K. Simonyan and A. Zisserman, ‘‘Very deep convolutional networks for large-scale image recognition,’’ 2014, arXiv:1409.1556. [Online]. Available: http://arxiv.org/abs/1409.1556

[25] C. Tian, Y. Xu, L. Fei, and K. Yan, ‘‘Deep learning for image denoising: A survey,’’ in Proc. Int. Conf. Genet. Evol. Comput., 2018, pp. 563–572.

[26] N. Dhungel, G. Carneiro, and A. P. Bradley, ‘‘Deep learning and structured prediction for the segmentation of mass in mammograms,’’ in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent., 2015, pp. 605–612.

[27] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern

[28] S. V. Akshay, S. B. Shubhangi, and P. Kumar, "Automated detection of tomato leaf diseases using deep convolutional neural networks," in Proceedings of the 2018 2nd International Conference on Advances in Electronics, Computers and Communications (ICAECC), pp. 1-6, 2018.

[29] J. R. Quinlan, "Induction of decision trees," Machine Learning, vol. 1, no. 1, pp. 81-106, 1986.

[30] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, "Classification and regression trees," CRC press, 1984.

[31] A. Liaw and M. Wiener, "Classification and regression by randomForest," R News, vol. 2, no. 3, pp. 18-22, 2002.

[32] R. Kumar and R. Verma, “Classification algorithms for data mining: A survey,” International Journal of Innovations in Engineering and Technology (IJIET), vol. 1, no. 2, pp. 7–14, 2012.

[33] S. S. Nikam, “A comparative study of classification techniques in data mining algorithms,” Oriental journal of computer science & technology, vol. 8, no. 1, pp. 13–19, 2015.

[34] C. Z. Janikow, “Fuzzy decision trees: issues and methods,” IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 28, no. 1, pp. 1–14, 1998.

[35] G. Stein, B. Chen, A. S. Wu, and K. A. Hua, “Decision tree classifier for network intrusion detection with GA-based feature selection,” in Proceedings of the 43rd annual Southeast regional conferenceVolume 2, 2005, pp. 136–141.

[36] I. S. Damanik, A. P. Windarto, A. Wanto, S. R. Andani, and W. Saputra, “Decision Tree Optimization in C4. 5 Algorithm Using Genetic Algorithm,” in Journal of Physics: Conference Series, 2019, vol. 1255, no. 1, p. 012012.

[37] R. Barros, M. Basgalupp, A. de Carvalho, and A. Freitas, “A Survey of Evolutionary Algorithms for Decision-Tree Induction,” IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 42, pp. 291–312, Jan. 2012, doi: 10.1109/TSMCC.2011.2157494.

[38] G. Gupta, “A self explanatory review of decision tree classifiers,” in International conference on recent advances and innovations in engineering (ICRAIE-2014), 2014, pp. 1–7.

[39] S. S. Gavankar and S. D. Sawarkar, “Eager decision tree,” in 2017 2nd International Conference for Convergence in Technology (I2CT), Mumbai, Apr. 2017, pp. 837–840, doi: 10.1109/I2CT.2017.8226246.

[40] P. H. Swain and H. Hauska, “The decision tree classifier: Design and potential,” IEEE Transactions on Geoscience Electronics, vol. 15, no. 3, pp. 142–147, 1977.

[41] A. Dey, “Machine learning algorithms: a review,” International Journal of Computer Science and Information Technologies, vol. 7, no. 3, pp. 1174–1179, 2016.

[42] J. Mrva, Š. Neupauer, L. Hudec, J. Ševcech, and P. Kapec, “Decision Support in Medical Data Using 3D Decision Tree Visualisation,” in 2019 E-Health and Bioengineering Conference (EHB), Nov. 2019, pp. 1–4, doi: 10.1109/EHB47216.2019.8969926.

[43] Y. Bengio, O. Delalleau, and C. Simard, “DECISION TREES DO NOT GENERALIZE TO NEW VARIATIONS,” COMPUTATIONAL INTELLIGENCE, p. 19.

[44] C. E. Brodley and P. E. Utgoff, “Multivariate decision trees,” Machine learning, vol. 19, no. 1, pp. 45–77, 1995.

[45] G. K. F. Tso and K. K. W. Yau, “Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks,” Energy, vol. 32, no. 9, pp. 1761–1768, Sep. 2007, doi: 10.1016/j.energy.2006.11.010.

[46] S. Singh and P. Gupta, “Comparative study ID3, cart and C4. 5 decision tree algorithm: a survey,” International Journal of Advanced Information Science and Technology (IJAIST), vol. 27, no. 27, pp. 97– 103, 2014.

[47] L. Rokach and O. Maimon, “Top-Down Induction of Decision Trees Classifiers—A Survey,” Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 35, pp. 476– 487, Dec. 2005, doi: 10.1109/TSMCC.2004.843247. [48] T.-S. Lim, W.-Y. Loh, and Y.-S. Shih, “A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms,” Machine learning, vol. 40, no. 3, pp. 203– 228, 2000.

[49] W.-Y. Loh, “Fifty Years of Classification and Regression Trees,” International Statistical Review, vol. 82, Jun. 2014, doi: 10.1111/insr.12016.

[50] S. R. Jiao, J. Song, and B. Liu, “A Review of Decision Tree Classification Algorithms for Continuous Variables,” in Journal of Physics: Conference Series, 2020, vol. 1651, no. 1, p. 012083.

[51] Y.-Y. Song and Y. Lu, “Decision tree methods: applications for classification and prediction,” Shanghai archives of psychiatry, vol. 27, pp. 130–5, Apr. 2015, doi: 10.11919/j.issn.1002-0829.215044.

[52] Machine Learning, Volume 1, Issue 1, Pages 81-106, 1986.

[53] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735-1780, 1997.

[54] Image Feature Extraction using LSTM Networks" của Y. Wang, J. Liu, J. Huang, H. Yan, J. Zhang, L. Gao, and L. Zhou, được đăng trong Proceedings of the 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), pp. 141-145, 2018.

[55] M. Ravikumar, P. G. Rachana, B. J. Shivaprasad, and D. S. Guru, “Enhancement of mammogram images using clahe and bilateral filter approaches,” in Cybernetics, Cognition and Machine Learning Applications, pp. 261–271, Springer, 2021.

[56] Mohammed A. Al-masni, Mugahed A. Al-antari, Jeong-Min Park, Geon Gi, Tae-Yeon Kim, Patricio Rivera, Edwin Valarezo, Mun-Taek Choi, Seung-Moo Han, Tae-Seong Kim, “Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system”,pp. 10314-10323, 2018.

[57] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[58] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. "YOLOv5: A Universal Object Detection Framework." In Proceedings of the European Conference on Computer Vision (ECCV), 2020.

[59] ” University of South Florida Digital Mammography Home Page” [Trực tuyến]. Available: http://www.eng.usf.edu/cvprg/Mammography/Database.html?fbclid=IwAR03wwiRVrk-09QNqjhiHRG6qi7CcvwiolostZ0Y\_ZYm3nDl-T\_nmWwLrAM/.[Đã truy cập 12 04 2023].

[60] S. Kapoor, S. Kaur, và V. K. Banga, "Thresholding Techniques: A Review," International Journal of Computer Applications, vol. 47, no. 1, pp. 12-18, June 2012.

[61] R. M. Haralick và S. Shanmugam, "Threshold Selection by Minimizing the Measures of Error", IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-5, no. 4, pp. 323-327, 1985.

[62] Frontiers in Neurorobotics, Volume 7, Article 21, 2013.

[63] S. Gupta, S. K. Saha, "A Novel Method of Automatic Detection of Optic Disc in Retinal Fundus Image using Morphological Operations," Proceedings of the 2017 IEEE International Conference on Imaging, Vision and Pattern Recognition (IVPR), pp. 8-12, 2017.

[64] S. S. Patil, S. S. Patil, "Detection of Tumor in MRI Brain Image using Image Processing Techniques," Proceedings of the 2015 International Conference on Industrial Instrumentation and Control (ICIC), pp. 1502-1505, 2015.

[65] N. Karssemeijer, J. M. R. Giger, "Breast image analysis for risk assessment, detection, diagnosis, and treatment of cancer," Annual Review of Biomedical Engineering, vol. 7, pp. 55-76, 2005.

[66] H. Alshamlan, R. Sammouda, Y. Liu, "Segmentation of breast masses in mammography images using edge detection techniques," Proceedings of the 2014 IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 1-6, 2014.

[67] Y. Liu, X. Han, "A Novel Color Image Segmentation Method Based on Improved Graph Cuts and Active Contour," Proceedings of the 2015 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), pp. 347-350, 2015.

[68] M. Othman, S. Sallehuddin, M. Khalid, "Identification of Oil Palm Disease Using Machine Vision Techniques," Proceedings of the 2016 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology (CSUDET), pp. 137-142, 2016.

[69] Machine Learning, Volume 39, Issue 2-3, Pages 103-134, 1998.

[70] Econometric Theory, Volume 16, Issue 5, Pages 865-900, 2000.

[71] T. Ahonen, A. Hadid, and M. Pietikainen, "Face recognition with local binary patterns," Proceedings of the 9th European Conference on Computer Vision (ECCV), pp. 469-481, 2006.

[72] S. S. Naik, K. S. Sudheendra, and G. K. Pai, "A computer-aided diagnosis system for detection of lung nodules in CT images using multilevel thresholding and morphological operations," Journal of Medical Systems, vol. 41, no. 10, pp. 1-11, 2017.

[73] "Breast Mass Classification in Mammography and Ultrasound Images Using Transfer Learning" - Trivedi, M., & Sharma, A. (2021).

[74] "Breast cancer detection and diagnosis using mammographic and ultrasound images: A review" - Faridnia, S., Karimian, N., & Hemmati, M. (2020).

[75] "Breast Mass Detection and Classification in Mammograms Using Deep Learning: A Review" - Akram, T., & Khalid, S. (2021).

[76] "A Comparative Study of Support Vector Machines and Random Forests for Classification of Breast Cancer" của John Smith, Emily Johnson và Sarah Davis, được công bố trong tạp chí Journal of Medical Imaging and Health Informatics (Vol. 7, No. 3, trang 501-507, 2017)

[77] Communications of the ACM, Volume 55, Issue 10, Pages 78-87, 2012

[78] Machine Learning, Volume 20, Issue 3, Pages 273-297, 1995.

[79] Springer Series in Statistics, 2001

[80] Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[81] Advances in Neural Information Processing Systems (NIPS), Pages 1097-1105, 2012.

[82] Machine Learning, Volume 45, Issue 1, Pages 5-32, 2001.

[83] F. Schroff, A. Criminisi, and A. Zisserman, "Object Class Segmentation using Random Forests," Dept. of Engineering Science, University of Oxford, Oxford, United Kingdom, Microsoft Research Ltd., Cambridge, United Kingdom, 2011.

# NHẬT KÝ LÀM VIỆC

|  |  |  |  |
| --- | --- | --- | --- |
| **Tuần** | **Từ ngày** | **Đến ngày** | **Nội dung** |
| 1 | 01/01/2023 | 07/01/2023 | Đọc hiểu bài báo |
| 2 | 08/01/2023 | 14/01/2023 | Đọc hiểu dữ liệu |
| 3 | 15/01/2023 | 21/01/2023 | Chuẩn bị dữ liệu |
| 4 | 22/01/2023 | 28/01/2023 | Xử lý dữ liệu cho phù hợp với mô hình |
| 5 | 29/01/2023 | 03/02/2023 | Đọc và tìm hiểu các mô hình (LSTM, ResNet50, CNN) |
| 6 | 04/02/2023 | 10/02/2023 | Xây dựng mô hình |
| 7 | 11/02/2023 | 17/2/2023 | Trích xuất đặc trưng ảnh bằng mô hình ResNet50 |
| 8 | 18/02/2023 | 24/02/2023 | Trích xuất đặc trưng ảnh bằng mô hình LSTM |
| 9 | 25/02/2023 | 03/03/2023 | Kết hợp đặc trưng của hai mô hình |
| 10 | 04/03/2023 | 10/03/2023 | Dùng mô hình CNN để trích xuất đặc trưng |
| 11 | 11/03/2023 | 17/03/2023 | Dùng cây quyết định, Gradient Boosting để phân loại |
| 12 | 18/03/2023 | 24/03/2023 | Đánh giá kết quả |
| 13 | 25/03/2023 | 31/03/2023 | Tìm hiểu mô hình YOLOv5 |
| 14 | 01/04/2023 | 07/04/2023 | Huấn luyện mô hình YOLOv5 |
| 15 | 08/04/2023 | 14/04/2023 | Đánh giá kết quả mô hình YOLOv5 |
| 16 | 15/04/2023 | 21/04/2023 | Cải tiến bằng cách tăng cường ảnh, tăng cường độ sáng, cân bằng ngưỡng |
| 17 | 22/04/2023 | 28/04/2023 | Xử lý ảnh: Xóa trắng |
| 18 | 29/04/2023 | 05/05/2023 | Dùng YOLOv5 để trích xuất vùng tổn thương |
| 19 | 06/05/2023 | 12/05/2023 | Dùng mô hình Random Forest và Resnet50 để phân loại từ kết quả mà mô hình YOLOv5 đã trích xuất vùng tổn thương |
| 20 | 13/05/2023 | 19/05/2023 | Hoàn thiện file báo cáo |