



**BỘ LAO ĐỘNG - THƯƠNG BINH VÀ XÃ HỘI**  
**TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT VĨNH LONG**

---

**KHÓA LUẬN TỐT NGHIỆP**  
**ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**PHÁT HIỆN VÀ CHẨN ĐOÁN UNG THƯ VÚ BẰNG**  
**HÌNH ẢNH CT SỬ DỤNG KỸ THUẬT HỌC SÂU**

<b>CBHD:</b>	<b>TS. Phan Anh Cang</b>
<b>Sinh viên:</b>	<b>Lý Trần Hoàng Hiếu</b>
<b>Mã số sinh viên:</b>	<b>19004060</b>

**Vĩnh Long - Năm 2023**



**BỘ LAO ĐỘNG - THƯƠNG BINH VÀ XÃ HỘI**  
**TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT VĨNH LONG**

---

**KHÓA LUẬN TỐT NGHIỆP**  
**ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**PHÁT HIỆN VÀ CHẨN ĐOÁN UNG THƯ VÚ BẰNG**  
**HÌNH ẢNH CT SỬ DỤNG KỸ THUẬT HỌC SÂU**

<b>CBHD:</b>	<b>TS. Phan Anh Cang</b>
<b>Sinh viên:</b>	<b>Lý Trần Hoàng Hiếu</b>
<b>Mã số sinh viên:</b>	<b>19004060</b>

## LỜI CAM ĐOAN

Tôi xin cam đoan kết quả đạt được trong luận văn là sản phẩm của riêng cá nhân, không sao chép lại của người khác. Trong toàn bộ nội dung của luận văn, những điều được trình bày hoặc là của cá nhân hoặc là được tổng hợp từ nhiều nguồn tài liệu. Tất cả các tài liệu tham khảo đều có xuất xứ rõ ràng và được trích dẫn hợp pháp. Đây là công trình nghiên cứu của cá nhân tôi và được sự hướng dẫn thực hiện của Tiến sĩ Phan Anh Cang. Các nội dung nghiên cứu trong đề tài **“Phát hiện và chẩn đoán ung thư vú bằng hình ảnh CT sử dụng kỹ thuật học sâu”** của tôi là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây và dựa trên một số tài liệu, đề tài trước. Tôi xin hoàn toàn chịu trách nhiệm và chịu mọi hình thức kỷ luật theo quy định cho lời cam đoan của mình.

*Vĩnh Long, ngày    tháng    năm 2023*

Người thực hiện đề tài

Lý Trần Hoàng Hiếu

## LỜI CẢM ƠN

Trước hết em xin bày tỏ lòng biết ơn đối với tất cả các thầy cô trong Khoa Công nghệ thông tin - Trường Đại học Sư phạm Kỹ thuật Vĩnh Long. Đặc biệt nhất là thầy Phan Anh Cang, người đã dành cho em rất nhiều thời gian quý báu, trực tiếp hướng dẫn tận tình giúp đỡ, chỉ bảo và góp ý cho em trong suốt quá trình làm đề tài khóa luận tốt nghiệp này.

Em xin chân thành cảm ơn các thầy cô tham gia giảng dạy và truyền đạt những kiến thức quý báu về học tập cũng như về cuộc sống trong suốt thời gian em học tập tại trường, đã đọc và phản biện bài báo cáo của em, giúp em hiểu rõ hơn các vấn đề mà mình nghiên cứu.

Tuy có nhiều cố gắng trong quá trình học tập cũng như trong quá trình làm đề tài khóa luận nhưng không thể tránh khỏi những thiếu sót, em rất mong được sự góp ý quý báu của tất cả các thầy, cô cũng như tất cả các bạn để kết quả của em được hoàn thiện ngày một tốt hơn.

Cuối cùng chúng em xin kính chúc quý thầy, cô dồi dào sức khỏe và thành công trong sự nghiệp cao quý.

Một lần nữa, em xin chân thành cảm ơn!

*Vĩnh Long, ngày    tháng    năm 2023*

Người thực hiện đề tài

Lý Trần Hoàng Hiếu

## MỤC LỤC

LỜI CAM ĐOAN .....	i
LỜI CẢM ƠN .....	ii
DANH MỤC CÁC TỪ VIẾT TẮT .....	v
DANH MỤC HÌNH ẢNH .....	vi
DANH MỤC CÁC BẢNG.....	vii
MỞ ĐẦU.....	1
CHƯƠNG 1: CƠ SỞ KHOA HỌC CỦA ĐỀ TÀI .....	4
1.1 Giới thiệu.....	4
1.2 Lý do chọn đề tài.....	6
1.3 Cơ sở lý luận .....	7
1.4 Cơ sở thực tiễn .....	8
1.5 Những nghiên cứu liên quan .....	8
1.6 Vấn đề đặt ra .....	14
1.7 Hướng giải quyết của đề tài .....	14
CHƯƠNG 2: CƠ SỞ LÝ THUYẾT .....	15
2.1 Bệnh ung thư vú .....	15
2.2 Machine Learning .....	22
2.3 Deep Learning.....	26
2.4 Mạng học sâu và vai trò .....	30
2.5 Vision transformer .....	43
2.6 Phương pháp đánh giá.....	46
CHƯƠNG 3: PHƯƠNG PHÁP ĐỀ XUẤT.....	50
3.1 Tập dữ liệu .....	50
3.2 Môi trường thực nghiệm .....	50
3.3 Phương pháp đề xuất.....	51
3.4 Kịch bản huấn luyện.....	51
CHƯƠNG 4: KẾT QUẢ ĐẠT ĐƯỢC .....	53

4.1 Kết quả huấn luyện.....	53
4.2 Kết quả kiểm thử.....	61
CHƯƠNG 5: KẾT LUẬN .....	63
5.1 Kết quả đạt được .....	63
5.2 Hạn chế.....	63
5.3 Hướng phát triển .....	63
TÀI LIỆU THAM KHẢO.....	64

## DANH MỤC CÁC TỪ VIẾT TẮT

STT	Kí hiệu	Nội dung đầy đủ
1	AI	Artificial Intelligent
2	ViT	Vision Transformer
3	BCH	Hình ảnh mô bệnh học tế bào vú
4	ROI	Vùng quan tâm
5	WDMR	Vùng đa quy mô
6	DSSM	Mô hình phân tách chiến lược kép
7	CLAH	Contrast Limited Adaptive Histogram Equalization
8	CNN	Convolutional Neural Networks
9	CSS	Không gian tỷ lệ cong
10	SVM	Máy vectơ hỗ trợ
11	DL	Deep Learning

## DANH MỤC HÌNH ẢNH

Hình 1.1 Các giai đoạn ung thư vú (Tâm Anh Hospital, 2022) .....	5
Hình 2.1 Hướng dẫn tự khám vú tại nhà (Tâm Anh Hospital, 2022) .....	22
Hình 2.2 Mô tả Machine Learning Workflow (Deepika Koundal, 2022) .....	25
Hình 2.3 Kiến trúc mạng Resnet (Phạm Đình Khánh, 2020) .....	30
Hình 2.4 Cộng trực tiếp đầu vào của khối với nhánh còn lại trong khối Identity block (Phạm Đình Khánh, 2020) .....	31
Hình 2.5 Kiến trúc tóm tắt của mạng ResNet50 (Phạm Đình Khánh, 2020).....	32
Hình 2.6 Kiến trúc DesneNet201 (Phạm Đình Khánh, 2022) .....	33
Hình 2.7 Kiến trúc mạng Unet (Bùi Tiến Tùng, 2020).....	33
Hình 2.8 Residual learning: a building block (Phạm Đình Khánh, 2022).....	34
Hình 2.9 MobileNetV2 building block (Phạm Đình Khánh, 2022).....	35
Hình 2.10 Impact of non-linearity bottleneck layer (Phạm Đình Khánh, 2022) .....	36
Hình 2.11 Squeeze and Excitation networks (Phạm Đình Khánh, 2022).....	37
Hình 2.12 MobileNetV3 block (Phạm Đình Khánh, 2022) .....	37
Hình 2.13 Kiến trúc mạng InceptionV1 (Phạm Đình Khánh, 2022) .....	38
Hình 2.14 Kiến trúc mạng InceptionV3 (Phạm Đình Khánh, 2022) .....	39
Hình 2.15 Mô hình mạng VGG-19 (Phạm Đình Khánh, 2022).....	41
Hình 2.16 Kiến trúc mạng DeeplabV3+ (Phạm Đình Khánh, 2022).....	41
Hình 2.17 Mô tả Atrous Spatial Pyramid Pooling (Phạm Đình Khánh, 2022).....	42
Hình 2.18 Kiến trúc Vision transformer (Lê Minh Chiến, 2021) .....	43
Hình 2.19 Chia batch ảnh gốc (Lê Minh Chiến, 2021).....	44
Hình 2.20 Linear projection (Lê Minh Chiến, 2021).....	44
Hình 2.21 Positional Embeeding (Lê Minh Chiến, 2021) .....	45
Hình 2.22 Kiến trúc của Transformer Encoder (Lê Minh Chiến, 2021).....	45
Hình 2.23 Confusion matrix (Minh Vũ, 2021) .....	48
Hình 3.1 Kiến trúc đề xuất phân loại, phân vùng ung thư vú .....	51
Hình 4.1 Biểu đồ so sánh độ chính xác của các mô hình.....	56
Hình 4.2 Biểu đồ so sánh độ mất mát của các mô hình .....	59
Hình 4.3 Biểu đồ so sánh thời gian huấn luyện của các mô hình .....	59



## DANH MỤC CÁC BẢNG

Bảng 3.1 Bảng thông số môi trường huấn luyện .....	50
Bảng 3.2 Kịch bản huấn luyện .....	51
Bảng 4.1 Kết quả độ chính xác của các mô hình .....	53
Bảng 4.2 Kết quả độ mất mát của các mô hình.....	56
Bảng 4.3 Bảng tổng hợp số liệu Confusion matrix .....	60
Bảng 4.4 Kết quả kiểm thử phân loại ung thư vú .....	61
Bảng 4.5 Kết quả kiểm thử phân vùng ung thư vú .....	61

## MỞ ĐẦU

### 1. TÍNH CẤP THIẾT CỦA ĐỀ TÀI

Ung thư vú là một trong những căn bệnh ung thư thường gặp ở phụ nữ, khi mắc bệnh, mô tuyến vú sẽ xuất hiện tế bào ác tính. Sau một thời gian, tế bào ung thư này phát triển với tốc độ nhanh chóng lan rộng ra toàn bộ vú và di căn sang các cơ quan khác trên cơ thể. Trong những năm trở lại đây, tỷ lệ phụ nữ mắc ung thư vú có xu hướng tăng nhanh.

Trước khi tìm hiểu dấu hiệu ung thư vú, chúng ta cần nắm được một số dạng bệnh thường gặp. Loại ung thư vú thường gặp nhất là ung thư tuyến vú, nguyên nhân là do tế bào gây bệnh ung thư phát triển từ ống tuyến. Ngoài ra, còn có ung thư tiểu thùy hoặc ung thư vú dạng viêm. So với ung thư tuyến vú, hai dạng bệnh kể trên hiếm gặp hơn.

Bệnh ung thư vú phát triển qua 5 giai đoạn. Ở những giai đoạn đầu, tế bào ác tính đã xuất hiện nhưng vẫn chỉ tập trung 1 chỗ chưa lan ra các mô kế cận. Càng ở các giai đoạn sau, khối u càng lớn, và có khả năng lan ra các cơ quan xung quanh. Bước sang giai đoạn cuối, tế bào gây bệnh ung thư đã di căn tới nhiều cơ quan trên cơ thể. Lúc này, việc chữa trị bệnh thường không đem lại hiệu quả cao, bệnh nhân có nguy cơ tử vong.

Các đối tượng có nguy cơ mắc ung thư vú: Người có tiền sử gia đình mắc bệnh ung thư vú; Người gặp vấn đề sinh sản (vô sinh, hiếm muộn hoặc có con đầu lòng khi trên 35 tuổi). Người có tiền sử bản thân mắc bệnh u nang hoặc u xơ tuyến vú, ung thư buồng trứng, ung thư tử cung, ... Người phải tiếp xúc nhiều với hóa chất độc hại, tia bức xạ.

Nguyên nhân gây ung thư vú: Bệnh ung thư vú thường mắc phải ở những người sinh con muộn, không có khả năng sinh sản hoặc không cho con bú. Do gen di truyền: Nếu trong gia đình có mẹ hoặc bà, anh chị em mắc bệnh này thì bạn cũng nên đi bệnh viện kiểm tra, bởi bệnh này có thể di truyền trong các thành viên trong gia đình. Có kinh nguyệt sớm hay mãn kinh muộn cũng có thể là nguyên nhân gây ra bệnh. Những người có tiền sử các bệnh liên quan đến vú như xơ nang tuyến vú,... Sống trong môi trường độc hại, ô nhiễm cũng tạo điều kiện phát sinh ung thư vú. Béo phì, lười vận động, ăn thức ăn nghèo vitamin, hút thuốc lá, uống rượu cũng có nguy cơ bị ung thư.

Từ những tổn hại trên, cần có một giải pháp giúp các bác sĩ chẩn đoán chính xác hơn về Ung thư vú, việc thực hiện đề tài “Phát hiện và chẩn đoán ung thư vú bằng hình ảnh CT sử dụng kỹ thuật học sâu” là cần thiết, nhằm tạo nhiều thuận lợi cho các bác sĩ trong quá trình phân tích và điều trị.

## **2. MỤC TIÊU NGHIÊN CỨU**

- Thu thập dữ liệu hình ảnh y khoa liên quan đến bệnh ung thư vú.
- Tìm hiểu đặc điểm tổn thương của vú do bệnh ung thư hay do một số bệnh khác gây ra.
- Tìm hiểu các mô hình mạng dùng để phân tích và nhận dạng như Densenet201, Unet, Resnet50, InceptionV3, MobileNetV3Small, ViT, DeeplabV3+.
- Phát hiện và khoanh vùng tổn thương ở vú do ung thư theo các mô hình Deep Learning và các tập dữ liệu đã thu thập.

## **3. ĐỐI TƯỢNG NGHIÊN CỨU**

- Các ảnh CT và mô tế bào về bệnh ung thư vú.
- Các tập dữ liệu và các mô hình học sâu.

## **4. PHẠM VI NGHIÊN CỨU**

- Đề tài nghiên cứu về phương pháp phát hiện và khoanh vùng tổn thương vú vào nghiên cứu và các dữ liệu khoa học được công bố.
- Phân loại tổn thương phổi do bệnh ung thư hoặc do các bệnh khác.
- Nghiên cứu một số mạng như: Unet, Resnet50, InceptionV3Small, MobileNetV3, ViT, DeeplabV3+.

## **5. PHƯƠNG PHÁP NGHIÊN CỨU**

- Phương pháp lý thuyết: Tìm hiểu, phân tích các vấn đề liên quan về bệnh ung thư vú và những nghiên cứu khoa học liên quan về đề tài.
- Phương pháp thực nghiệm đề tài: Xây dựng lập trình bằng ngôn ngữ lập trình chuyên về AI là Python kết hợp với các thư viện đã được phát triển và nguồn mở.

## **6. Ý NGHĨA KHOA HỌC VÀ THỰC TIỄN**

- Tập hợp dữ liệu ảnh CT và ảnh mô tế bào ung thư vú.
- Phân loại và khoanh vùng tự động tổn thương phổi do bệnh ung thư gây ra.
- Đề xuất cải tiến, so sánh đánh giá giữa các mô hình.
- So sánh đánh giá một số phương pháp: Unet, Resnet50, InceptionV3Small, MobileNetV3, ViT, Densenet201, DeeplabV3+.

## **7. CẤU TRÚC CỦA KHÓA LUẬN**

Trong quyển báo cáo khóa luận bao gồm những nội dung được trình bày từ cơ sở lý thuyết, nguyên cứu liên quan, phương pháp đề xuất và kết quả thực nhiệm được trình bày vào các chương theo bố cục sau:

**MỞ ĐẦU:** Trình bày nội dung tính cấp thiết đề tài, mục tiêu, phạm vi, phương pháp và ý nghĩa thực hiện của đề tài.

**CHƯƠNG 1 - CƠ SỞ KHOA HỌC CỦA ĐỀ TÀI:** Trình bày nội dung bệnh ung thư vú, cơ sở phát hiện, khoanh vùng và phân loại ung thư vú.

**CHƯƠNG 2 - CƠ SỞ LÝ THUYẾT:** Trình bày nội dung về cơ sở lý thuyết được sử dụng trong các phương pháp được đề xuất.

**CHƯƠNG 3 - PHƯƠNG PHÁP ĐỀ XUẤT:** Trình bày công việc cụ thể từng pha xử lý được thực hiện từ kiến trúc đề xuất. Môi trường thực nghiệm và phương pháp đề xuất cũng được trình bày ở chương này.

**CHƯƠNG 4 - KẾT QUẢ ĐẠT ĐƯỢC:** Trình bày kết quả thực nghiệm đạt được.

**CHƯƠNG 5 - KẾT LUẬN:** Trình bày tóm tắt về kết quả đạt được, những đóng góp cũng như hướng phát triển tiếp theo sau khi thực hiện đề tài.

## CHƯƠNG 1: CƠ SỞ KHOA HỌC CỦA ĐỀ TÀI

### 1.1 Giới thiệu

Ung thư vú là dạng u vú ác tính. Một khối u có thể là lành tính (không ung thư) hoặc ác tính (ung thư). Đa số các trường hợp ung thư vú bắt đầu từ các ống dẫn sữa, một phần nhỏ phát triển ở túi sữa hoặc các tiểu thùy. Ung thư vú nếu phát hiện và điều trị muộn có thể đã di căn vào xương và các bộ phận khác, đau đớn sẽ càng nhân lên.

- Dấu hiệu cảnh báo ung thư vú:

+ Đau vùng ngực: Vùng ngực có cảm giác đau âm ỉ, không có quy luật rõ ràng. Có nhiều khả năng đây là tín hiệu cảnh báo ung thư vú ác tính giai đoạn sớm. Nếu vùng ngực đau, nóng rát liên tục hoặc ngày càng dữ dội thì bạn nên cần đi khám ngay.

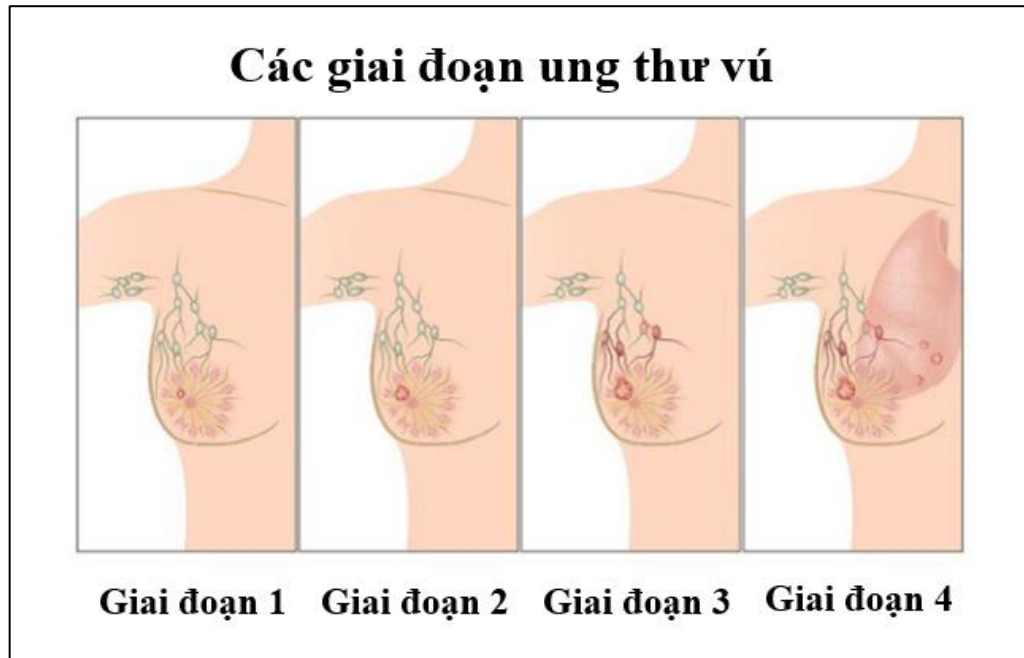
+ Thay đổi vùng da: Hầu hết những người mắc phải căn bệnh này thường thay đổi màu sắc và tính chất da ở vùng ngực. Vùng da thường có thể xuất hiện nhiều nếp nhăn hoặc lõm giống như lúm đồng tiền, vùng da xung quanh thường có mụn nước, ngứa lâu không dứt điểm.

+ Sưng hoặc nổi hạch: Sưng hạch bạch huyết không chỉ là dấu hiệu của các bệnh thông thường như cảm cúm, nhiễm trùng mà còn cả bệnh ung thư vú. Nếu có khối u hoặc vết sưng đau dưới vùng da kéo dài trong vài ngày mà bạn không rõ nguyên nhân thì đây có thể là dấu hiệu cảnh báo bệnh ung thư vú.

+ Đau lưng, vai hoặc gáy: Một số phụ nữ khi mắc phải bệnh này thay vì đau ngực thì thường có cảm giác đau lưng hoặc vai gáy. Những cơn đau thường xảy ra ở phía lưng trên hoặc giữa 2 bả vai, có thể nhầm lẫn với việc giãn dây chằng hoặc các bệnh liên quan trực tiếp đến cột sống.

- Các giai đoạn ung thư vú: Không giống các loại ung thư khác, ung thư vú là một bệnh hoàn toàn có thể chữa được nếu được phát hiện sớm, khoảng 80% bệnh nhân được chữa khỏi hoàn toàn nếu phát hiện bệnh ở những giai đoạn đầu.

+ Ung thư vú giai đoạn 0 (Giai đoạn đầu): Ở giai đoạn đầu này, các tế bào ung thư vú được bác sĩ phát hiện ra trong các ống dẫn sữa. Ta gọi đây là ung thư vú không xâm lấn, hay có cách gọi khác là ung thư biểu mô tuyến vú tại chỗ. Bệnh nhân sẽ được điều trị bệnh ung thư vú để ngăn chặn sự di căn của bệnh. Thường thì bệnh nhân chỉ cần cắt bỏ khối u và sử dụng thêm phương pháp xạ trị.



*Hình 1.1 Các giai đoạn ung thư vú (Tâm Anh Hospital, 2022)*

+ Ung thư vú giai đoạn 1: Ở giai đoạn 1A, khối u vẫn có kích thước nhỏ từ 2cm và các hạch bạch huyết chưa bị ảnh hưởng. Còn khi bệnh đã chuyển sang giai đoạn 1B sẽ không chỉ có khối u ở vú mà còn tìm thấy khối u tại các hạch bạch huyết ở nách. Đây vẫn là 2 giai đoạn phát hiện bệnh sớm. Bác sĩ sẽ sử dụng phương pháp phẫu thuật kết hợp một số liệu pháp để điều trị bệnh.

+ Ung thư vú giai đoạn 2: Ở giai đoạn 2, các khối u có kích thước từ 2 – 5cm và có thể chưa lây lan sang các hạch bạch huyết hoặc sang các hạch nách. Giai đoạn này được chia làm 2 giai đoạn nhỏ: 2A và 2B.

Giai đoạn 2A: Chưa xuất hiện u nguyên phát và chưa đến 4 hạch bạch huyết. Kích thước khối u nhỏ hơn 2cm và chưa đến 4 hạch bạch huyết. Khối u từ 2 – 4cm và chưa lan tới hạch bạch huyết và hạch dưới cánh tay.

Giai đoạn 2B: Ở giai đoạn, khối u có kích thước từ 2 đến 4cm và tìm thấy các cụm tế bào ung thư trong hạch bạch huyết, từ 1-3 hạch bạch huyết ở nách hoặc gần xương ức. Hoặc, kích thước khối u lớn hơn 5cm và chưa xâm lấn đến các hạch bạch huyết.

Bệnh nhân phát hiện bệnh ở giai đoạn 2 nên kết hợp các liệu pháp điều trị phẫu thuật, xạ trị, hóa trị và kích thích tố.

+ Ung thư vú giai đoạn 3: Khi bạn phát hiện bệnh ở giai đoạn 3, thì các khối u trong cơ thể đã lan rộng 4 – 9 hạch bạch huyết ở nách hoặc phủ các hạch bạch huyết

bên trong vú. Ở giai đoạn 3, liệu pháp điều trị cho bệnh nhân cũng giống với giai đoạn 2. Nếu bác sĩ phát hiện ra có khối u nguyên phát lớn, thì bạn sẽ phải dùng biện pháp hóa trị để làm nhỏ khối u trước khi tiến hành phẫu thuật.

+ Ung thư vú giai đoạn 4 (Giai đoạn cuối): Giai đoạn này được xem là giai đoạn cuối của bệnh ung thư. Tế bào ung thư đã lan rộng và di căn đến nhiều cơ quan trong cơ thể người bệnh. Ung thư vú thường di căn đến xương, não, phổi và gan. Ở giai đoạn này, bác sĩ sẽ chỉ định cho bạn điều trị toàn thân tích cực, đây là phương pháp phổ biến đối với bệnh nhân đang điều trị bệnh ung thư vú giai đoạn cuối.

## 1.2 Lý do chọn đề tài

Ung thư vú là một trong những căn bệnh ung thư thường gặp ở phụ nữ, khi mắc bệnh, mô tuyến vú sẽ xuất hiện tế bào ác tính. Sau một thời gian, tế bào ung thư này phát triển với tốc độ nhanh chóng lan rộng ra toàn bộ vú và di căn sang các cơ quan khác trên cơ thể. Trong những năm trở lại đây, tỷ lệ phụ nữ mắc ung thư vú có xu hướng tăng nhanh.

Trước khi tìm hiểu dấu hiệu ung thư vú, chúng ta cần nắm được một số dạng bệnh thường gặp. Loại ung thư vú thường gặp nhất là ung thư tuyến vú, nguyên nhân là do tế bào gây bệnh ung thư phát triển từ ống tuyến. Ngoài ra, còn có ung thư tiểu thùy hoặc ung thư vú dạng viêm. So với ung thư tuyến vú, hai dạng bệnh kể trên hiếm gặp hơn, song chúng ta không thể chủ quan.

Bệnh ung thư vú phát triển qua 5 giai đoạn. Ở những giai đoạn đầu, tế bào ác tính đã xuất hiện nhưng vẫn chỉ tập trung 1 chỗ chưa lan ra các mô kế cận. Càng ở các giai đoạn sau, khối u càng lớn, và có khả năng lan ra các cơ quan xung quanh. Bước sang giai đoạn cuối, tế bào gây bệnh ung thư đã di căn tới nhiều cơ quan trên cơ thể. Lúc này, việc chữa trị bệnh thường không đem lại hiệu quả cao, bệnh nhân có nguy cơ tử vong.

Các đối tượng có nguy cơ mắc ung thư vú: Người có tiền sử gia đình mắc bệnh ung thư vú; Người gặp vấn đề sinh sản (vô sinh, hiếm muộn hoặc có con đầu lòng khi trên 35 tuổi). Người có tiền sử bản thân mắc bệnh u nang hoặc u xơ tuyến vú, ung thư buồng trứng, ung thư tử cung, ... Người phải tiếp xúc nhiều với hóa chất độc hại, tia bức xạ.

Nguyên nhân gây ung thư vú: Bệnh ung thư vú thường mắc phải ở những người sinh con muộn, không có khả năng sinh sản hoặc không cho con bú. Do gen di truyền:

Nếu trong gia đình có mẹ hoặc bà, anh chị em mắc bệnh này thì bạn cũng nên đi bệnh viện kiểm tra, bởi bệnh này có thể di truyền trong các thành viên trong gia đình. Có kinh nguyệt sớm hay mãn kinh muộn cũng có thể là nguyên nhân gây ra bệnh. Những người có tiền sử các bệnh liên quan đến vú như xơ nang tuyến vú,... Sống trong môi trường độc hại, ô nhiễm cũng tạo điều kiện phát sinh ung thư vú. Béo phì, lười vận động, ăn thức ăn nghèo vitamin, hút thuốc lá, uống rượu cũng có nguy cơ bị ung thư.

Từ những tổn hại trên, cần có một giải pháp giúp các bác sĩ chẩn đoán chính xác hơn về Ung thư vú, việc thực hiện đề tài “Phát hiện và chẩn đoán ung thư vú bằng hình ảnh CT sử dụng kỹ thuật học sâu” là cần thiết, nhằm tạo nhiều thuận lợi cho các bác sĩ trong quá trình phân tích và điều trị.

### **1.3 Cơ sở lý luận**

Deep Learning là một thuật toán dựa trên một số ý tưởng từ não bộ tới việc tiếp thu nhiều tầng biểu đạt, cả cụ thể lẫn trừu tượng, qua đó làm rõ nghĩa của các loại dữ liệu. Hiện nay Deep Learning được ứng dụng trong rất nhiều lĩnh vực.

Thuật toán Deep Learning được tạo thành từ Artificial Neural Networks với các layers ẩn, dựa trên mạng thần kinh nơ-ron sinh học. Deep Natural Networks có thể được sử dụng để mô phỏng các mối quan hệ phức tạp phi tuyến tính trong cả giám sát (tập dữ liệu được sử dụng để dự đoán kết quả sau này) và không giám sát (gom nhóm và tìm ra các mẫu mới và khác thường trong dữ liệu). Deep Learning có độ chính xác cao, có thể học sâu rộng và đạt được độ chính xác ở mức công nhận vô cùng cao. Dù mới phát triển nhưng các kiến trúc Deep Learning có thể mô hình hoá một cách tự động các mối quan hệ phi tuyến phức tạp của dữ liệu. Do đó các kỹ thuật Deep Learning có thể hoạt động với dữ liệu thô mà không cần thông qua các bước trích xuất đặc trưng thủ công, phức tạp, chỉ phù hợp với từng vấn đề nhỏ. Có nhiều kiến trúc Deep Learning như mạng nơ-ron tích chập (Convolutional Neural Network), Deep Belief Network, Deep Learning đã đem lại nhiều thành công mang tính đột phá trong nhiều lĩnh vực như thị giác máy tính, xử lý ngôn ngữ tự nhiên.

Mô hình Vision Transformer (ViT) đã được giới thiệu trong một bài báo nghiên cứu được xuất bản dưới dạng báo cáo hội nghị tại ICLR 2021, có tiêu đề “An Image is Worth 16\*16 Words: Transformers for Image Recognition in Scale”. Nó được phát triển và xuất bản bởi Neil Houlsby, Alexey Dosovitskiy, và 10 tác giả khác của Google Research Brain Team.



Mã tinh chỉnh và các mô hình ViT đã pre-trained (tạm dịch: tiền huấn luyện) hiện có sẵn trên GitHub của Google Research. Quá trình pre-train này được tiến hành qua bộ dữ liệu ImageNet và ImageNet-21k.

Transformer trong học máy là một mô hình học sâu sử dụng các cơ chế của Cơ chế Attention, cân nhắc kỹ lưỡng tầm quan trọng của từng phần dữ liệu đầu vào. Transformer trong học máy bao gồm nhiều lớp Cơ chế Self-attention, chủ yếu được sử dụng trong các lĩnh vực AI của xử lý ngôn ngữ tự nhiên (NLP) và thị giác máy tính (CV).

Transformer trong học máy sở hữu những tiềm năng đầy hứa hẹn về một phương pháp học tập tổng quát có thể được áp dụng cho các phương thức dữ liệu khác nhau trong thị giác máy tính, nhằm đạt được độ chính xác tiêu chuẩn hiện đại với hiệu quả tham số tốt hơn.

Trong phạm vi đề tài này sẽ tìm hiểu kiến trúc và thực nghiệm huấn luyện mô hình mạng với tập dữ liệu ảnh y tế được tham khảo từ Kaggle và một số liên kết liên quan để phân loại và phân vùng tế bào ung thư vú trên môi trường Kaggle với mô hình mạng Densenet, Unet, Mobilenet, Inception, Resnet, ViT, DeeplabV3+.

#### **1.4 Cơ sở thực tiễn**

Hình ảnh CT - ảnh mô đã và đang được áp dụng rộng rãi trong việc phát hiện một số bệnh lý. Việc lý giải các hình ảnh thường gặp nhiều sai lầm do kinh nghiệm, cảm giác hoặc do nhận thức cá nhân. Vì thế khi phân tích, các bác sĩ phải luôn kết hợp với đặc điểm dịch tễ và biểu hiện lâm sàng để đưa ra chẩn đoán chính xác. Ngoài ra, Deep Learning ứng dụng trong lĩnh vực y tế đã ghi nhận được nhiều nghiên cứu mang đến những kết quả khả quan. Đó là tiền đề thực hiện nghiên cứu phát hiện bệnh ung thư vú dựa trên dấu hiệu tổn thương của vú.

#### **1.5 Những nghiên cứu liên quan**

Trong bài báo “Effective Image Processing and Segmentation-Based Machine Learning Techniques for Diagnosis of Breast Cancer” [18] năm 2022 của tác giả Deepika Koundal. Những người bị ung thư vú có nhiều khả năng là phụ nữ, nhưng nó cũng có thể xảy ra với nam giới. Chỉ có một số khác biệt nhỏ tách biệt ngực nam và nữ về cấu trúc. Bất cứ ai muốn tránh ung thư vú đều không thể thực hiện với bất kỳ phương pháp nào hiện có. Bệnh nhân được chẩn đoán và điều trị ung thư vú ở giai đoạn sớm sẽ có cơ hội điều trị thành công và phục hồi tốt hơn nếu họ được chăm sóc

nhANH chóng. Mọi người đều đồng ý rằng chụp nhũ ảnh kỹ thuật số là một cách rất tốt để phát hiện ung thư vú ở giai đoạn đầu. Chúng ta có thể sử dụng các kỹ thuật xử lý hình ảnh để phát hiện ung thư vú nhanh hơn, điều này có thể giúp chúng ta sống và được điều trị tốt hơn. Bài báo này nói về một hệ thống máy học và xử lý hình ảnh ung thư vú đã được tạo ra. Một ví dụ về dữ liệu đầu vào cho khuôn khổ này là một tập hợp các hình ảnh chụp nhũ ảnh. Những hình ảnh này được sử dụng làm dữ liệu đầu vào. Hình ảnh sau đó được xử lý để trông đẹp hơn bằng cách sử dụng phương pháp CLAHE. Điều này giúp loại bỏ nhiễu khỏi ảnh đồng thời làm cho ảnh đẹp hơn. Vẫn còn rất nhiều việc phải làm đối với khung trước khi nó có thể bắt đầu chia tách hình ảnh. Bởi vì các điểm ảnh được gắn nhãn vào thời điểm này, một hình ảnh được chia thành các phần riêng biệt. Điều này giúp xác định các đối tượng và vẽ ranh giới. Các kỹ thuật như SVM mờ, bộ phân loại Bayesian và rừng ngẫu nhiên được sử dụng để nhóm các hình ảnh được xử lý trước này thành các nhóm. Năm tham số, độ chính xác, độ nhạy, độ đặc hiệu, độ chụm và thu hồi, được sử dụng trong phân tích thực nghiệm. Fuzzy SVM đang hoạt động tốt hơn thuật toán phân loại Bayesian và rừng ngẫu nhiên.

Nghiên cứu tiếp theo được đăng trong bài báo “Automatic cell nuclei segmentation and classification of breast cancer histopathology images” [28] của các tác giả Pin Wang, Xianling Hu, Yongming Li, Qianqian Liu, Xinjian Zhu. Ung thư vú là loại khối u ác tính hàng đầu được quan sát thấy ở phụ nữ và việc điều trị hiệu quả phụ thuộc vào chẩn đoán sớm. Chẩn đoán từ hình ảnh mô bệnh học vẫn là "tiêu chuẩn vàng" cho ung thư vú. Sự phức tạp của hình ảnh mô bệnh học tế bào vú (BCH) làm cho việc phân đoạn và phân loại đáng tin cậy trở nên khó khăn. Trong bài báo này, một kỹ thuật phân tích hình ảnh định lượng tự động của hình ảnh BCH được đề xuất. Đối với phân đoạn hạt nhân, phép biến đổi mũ trên-dưới được áp dụng để nâng cao chất lượng hình ảnh. Phân tách sóng con và tăng trưởng vùng đa quy mô (WDMR) được kết hợp để thu được các vùng quan tâm (ROI) nhờ đó nhận ra vị trí chính xác. Mô hình phân tách chiến lược kép (DSSM) chứa hình thái toán học thích ứng và phương pháp phát hiện góc Không gian tỷ lệ cong (CSS) được áp dụng để phân tách các ô chồng chéo để có độ chính xác và mạnh mẽ hơn. Để phân loại nhân tế bào, 4 đặc điểm dựa trên hình dạng và 138 đặc điểm kết cấu dựa trên không gian màu được trích xuất. Bộ tính năng tối ưu thu được bằng máy vectơ hỗ trợ (SVM) với

thuật toán di truyền tác nhân giống như chuỗi (CAGA). Phương pháp đề xuất đã được thử nghiệm trên 68 ảnh BCH chứa hơn 3600 ô. Kết quả thực nghiệm cho thấy độ nhạy phân đoạn trung bình là 91,53% ( $\pm 4,05\%$ ) và độ đặc hiệu là 91,64% ( $\pm 4,07\%$ ). Hiệu suất phân loại của hình ảnh tế bào bình thường và tế bào ác tính có thể đạt được độ chính xác 96,19% ( $\pm 0,31\%$ ), độ nhạy 99,05% ( $\pm 0,27\%$ ) và độ đặc hiệu 93,33% ( $\pm 0,81\%$ ).

Trong bài báo “Breast Cancer Segmentation Methods: Current Status and Future Potentials” [26] của nhóm các tác giả Paul Harrison, Epimack Michael, He Ma, Hong Li, Frank Kulwa, and Jing Li đã đề xuất một số phương pháp phân đoạn dựa trên vùng ung thư vú. Cụ thể như sau: Dehghani và Dezfooli đã trình bày một phương pháp để cải thiện quá trình tiền xử lý hình ảnh chụp quang tuyến vú. Phương pháp này có hai giai đoạn: (a) các phần hình ảnh dư thừa được loại bỏ bằng cách sử dụng độ sáng pixel và (b) hình ảnh chụp quang tuyến vú được đặt theo một hướng. Tổng cộng có 60 hình ảnh thu được từ cơ sở dữ liệu MIAS đã được sử dụng để kiểm tra thuật toán. Ngoài ra, nhiều từ hình ảnh chụp quang tuyến vú đã được loại bỏ bằng cách sử dụng giới hạn ngưỡng. Phương pháp đề xuất tạo ra độ chính xác phân đoạn 99,0%. Senthilkumar và cộng sự xây dựng một phương pháp cho thuật toán phân vùng phát triển để phát hiện ung thư vú. Tổng cộng có 40 hình ảnh chụp quang tuyến vú được lấy từ cơ sở dữ liệu MIAS và bộ lọc trung bình được sử dụng để loại bỏ nhiễu khỏi hình ảnh chụp quang tuyến vú. Để tăng độ chính xác của phân đoạn, hình ảnh chụp quang tuyến vú được tăng cường bằng cách sử dụng cân bằng biểu đồ thích ứng hạn chế tương phản (CLAHE) và góc Harris. Phương pháp được đề xuất tạo ra độ chính xác phân đoạn là 93,0%. Berber đã đề xuất một phương pháp phân đoạn đường viền khối lượng vú cho chụp quang tuyến vú kỹ thuật số. Phương pháp đề xuất dựa trên sự tăng trưởng vùng hạt giống cổ điển. Hơn nữa, phương pháp này được đánh giá bằng cách sử dụng 260 khối chụp quang tuyến vú thu được từ bộ chụp nhũ ảnh Dokuz Eylul (DEMS). Phương pháp đề xuất đạt độ chính xác 95,06%. Petrick và cộng sự đã đề xuất một phương pháp kết hợp tăng cường thích ứng và phân vùng phát triển theo khu vực của khối vú trên phim chụp quang tuyến vú. Các hình ảnh được cải thiện bằng cách sử dụng tăng cường độ tương phản theo trọng số mật độ và loại bỏ nhiễu bằng bộ lọc Gaussian. Phương pháp này đã được thử nghiệm bằng cách sử

dụng 253 bức ảnh chụp quang tuyến vú thu được từ Bệnh viện Đại học Michigan (UMH). Phương pháp đề xuất tạo ra độ chính xác phân đoạn 98,0%.

Trong bài báo “Medical breast ultrasound image segmentation by machine learning” [40] của tác giả Yuan Xu. Họ đã đề xuất sử dụng mạng thần kinh tích chập (CNN) để phân đoạn hình ảnh siêu âm vú thành bốn mô chính: da, mô sợi, khối và mô mỡ, trên hình ảnh siêu âm vú ba chiều (3D). Các chỉ số định lượng để đánh giá kết quả phân đoạn bao gồm Độ chính xác, Độ chính xác, Thu hồi đều đạt trên 80%, điều này cho thấy phương pháp được đề xuất có khả năng phân biệt các mô chức năng trong hình ảnh siêu âm vú. Một số liệu khác được gọi là chỉ số tương tự Jaccard (JSI) mang lại giá trị 85,1%, vượt trội so với nghiên cứu trước đây của họ bằng cách sử dụng thuật toán đầu nguồn với giá trị JSI 74,54%. Do đó, phương pháp đề xuất của họ có thể có khả năng cung cấp các phân đoạn cần thiết để hỗ trợ chẩn đoán lâm sàng ung thư vú và cải thiện hình ảnh ở các chế độ khác trong siêu âm y tế.

Trong bài báo “Automated 3D ultrasound image segmentation to aid breast cancer image interpretation” [27] của tác giả Peng Gu. Họ đã đề xuất một thuật toán tự động để phân chia khối lượng siêu âm 3D thành ba loại mô chính: u nang/khối u, mô mỡ và mô tuyến xơ. Để kiểm tra hiệu quả và tính nhất quán của nó, phương pháp tự động được đề xuất đã được sử dụng trên cơ sở dữ liệu gồm 21 trường hợp siêu âm toàn bộ vú. Kết quả thử nghiệm cho thấy rằng phương pháp đề xuất của họ không chỉ phân biệt chính xác các mô mỡ và không béo mà còn thực hiện tốt việc phân loại nang/khối. So sánh đánh giá tỷ trọng giữa phương pháp tự động và phân đoạn thủ công cho thấy tính nhất quán tốt với độ chính xác 85,7%. So sánh định lượng các thể tích mô tương ứng, sử dụng tỷ lệ chồng chéo, cho độ tương đồng trung bình là 74,54%. Do đó, phương pháp đề xuất của họ thể hiện tiềm năng lớn như một cách tiếp cận tự động để phân chia khối lượng siêu âm toàn bộ vú 3D thành các mô riêng biệt về chức năng có thể giúp điều chỉnh tốc độ siêu âm của quang sai và hỗ trợ tiên lượng ung thư vú dựa trên mật độ.

Tiếp theo là bài báo “Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors” [30] của Ruey-Feng Chang. Trong nghiên cứu này, ban đầu các khối u được phân đoạn bằng cách sử dụng phương pháp thiết lập cấp độ mới được phát triển và sau đó sáu đặc điểm hình thái được sử dụng để phân biệt các trường hợp lành tính và ác tính. Máy vectơ hỗ trợ (SVM) được sử dụng

để phân loại các khối u. Có 210 hình ảnh siêu âm của khối u vú lành tính đã được chứng minh bệnh lý từ 120 bệnh nhân và ung thư biểu mô từ 90 bệnh nhân trong cơ sở dữ liệu hình ảnh siêu âm. Cơ sở dữ liệu chỉ chứa một hình ảnh từ mỗi bệnh nhân. Hình ảnh siêu âm được chụp ở đường kính lớn nhất của khối u. Các hình ảnh được sưu tầm liên tục từ ngày 1 tháng 8 năm 1999 đến ngày 31 tháng 5 năm 2000; độ tuổi của bệnh nhân dao động từ 18 đến 64 tuổi. Siêu âm được thực hiện bằng hệ thống ATL HDI 3000 với đầu dò bộ phận nhỏ L10-5. Trong thực nghiệm, độ chính xác của SVM với thông tin hình dạng để phân loại khối u ác tính là 90,95% (191/210), độ nhạy 88,89% (80/90), độ đặc hiệu 92,5% (111/120), giá trị tiên đoán dương tính là 89,89% (80/89) và giá trị tiên đoán âm tính là 91,74% (111/121).

Trong bài báo “Segmentation of Breast Using Ultrasound Image for Detection Breast Cancer” [36] của Uswatun Khasana. Thuật toán biến đổi đầu nguồn được sử dụng trong quá trình phân đoạn để tạo ra vị trí của ung thư và có thể phân biệt các đối tượng dựa trên nền. Kết quả của việc phân đoạn sử dụng lưu vực là quá trình phân đoạn thứ hai sử dụng các mã nhị phân ngưỡng để tách hình ảnh của bệnh ung thư thành đối tượng được quan sát. Bước cuối cùng là quá trình tính diện tích ung thư. Kết quả của bài báo này là so sánh cách tính diện tích ung thư giữa dữ liệu của Bệnh viện và kết quả thử nghiệm, từ kết quả của các thử nghiệm này, độ chính xác của hệ thống đạt 88,65% trên tất cả các dữ liệu được thử nghiệm và sai số là 11,35%. Từ những thử nghiệm này, có thể kết luận rằng phương pháp được sử dụng là thuật toán biến đổi đầu nguồn có khả năng phân đoạn hình ảnh của hình ảnh siêu âm vú.

Một nghiên cứu khác từ “Development of a Deep-Learning-Based Method for Breast Ultrasound Image Segmentation” [29] của tác giả Rania Almajalid. Bài báo này đã phát triển một khung phân đoạn mới dựa trên kiến trúc học sâu u-net, để chụp ảnh siêu âm vú. U-net là một kiến trúc mạng thần kinh tích chập được thiết kế để phân đoạn hình ảnh sinh học với dữ liệu huấn luyện hạn chế. Ban đầu nó được đề xuất để phân đoạn cấu trúc nơ-ron trong hình ảnh kính hiển vi. Trong công việc của mình, họ đã sửa đổi và cải tiến phương pháp phân đoạn ảnh BUS. Trên cơ sở dữ liệu gồm 221 hình ảnh BUS, trước tiên họ áp dụng các kỹ thuật tiền xử lý bao gồm tăng cường độ tương phản và giảm nhiễu để cải thiện chất lượng hình ảnh. Sau đó, mô hình u-net đã được đào tạo và thử nghiệm thông qua xác thực chéo hai lần. Để tăng kích thước của tập huấn luyện, các chiến lược tăng dữ liệu bao gồm xoay và biến dạng đàn

hồi đã được áp dụng. Cuối cùng, một bước hậu xử lý đã loại bỏ (các) vùng nhiễu khối kết quả phân đoạn đã hoàn thiện toàn bộ phương pháp. Các số liệu lỗi khu vực, hệ số xúc xắc và tỷ lệ tương tự đã được tính toán để đánh giá hiệu suất trên các bộ thử nghiệm. Họ đã so sánh phương pháp của mình với hai phương pháp phân đoạn hoàn toàn tự động khác trên cùng một tập dữ liệu. Phương pháp của họ vượt trội so với hai phương pháp kia một cách đáng kể với hệ số tương đồng = 0,825 và tỷ lệ tương tự = 0,698. Kết quả thực nghiệm cho thấy phương pháp u-net cải tiến mạnh mẽ và chính xác hơn trong việc phân vùng khối u vú cho hình ảnh siêu âm.

Một nghiên cứu khác từ “Deep integrated pipeline of segmentation guided classification of breast cancer from ultrasound images” [23] của tác giả Muhammad Sakib Khan Inan. Họ đã phát triển một hệ thống tích hợp từ đầu đến cuối để phân loại hình ảnh siêu âm vú, họ đã tiến hành phân tích toàn diện các phương pháp tiền xử lý hình ảnh như K Means++ và SLIC, cũng như bốn mô hình học chuyển đổi như VGG16, VGG19, DenseNet121 và ResNet50. Với điểm hệ số tương đồng là 63,4 trong giai đoạn phân đoạn và độ chính xác cũng như Điểm F1 (Lợi tính) là 73,72% và 78,92% trong giai đoạn phân loại, sự kết hợp của SLIC, UNET và VGG16 vượt trội hơn tất cả các kết hợp tích hợp khác. Cuối cùng, họ đã đề xuất một khung đường ống tự động tích hợp từ đầu đến cuối, bao gồm tiền xử lý với SLIC để nắm bắt các tính năng siêu pixel từ tạo tác phức tạp của hình ảnh siêu âm, bổ sung phân đoạn ngưỡng nghĩa với U-Net đã sửa đổi, dẫn đến phân loại khối u vú bằng cách sử dụng phương pháp học chuyển giao. cách tiếp cận với VGG16 được đào tạo trước và mạng lưới thần kinh được kết nối dày đặc. Đường ống tự động được đề xuất có thể được triển khai một cách hiệu quả để hỗ trợ các bác sĩ trong việc chẩn đoán ung thư vú chính xác và kịp thời hơn.

Một nghiên cứu khác từ “SHA-MTL: soft and hard attention multi-task learning for automated breast cancer ultrasound image segmentation and classification” [21] của tác giả Guisheng Zhang. Họ đã đề xuất mô hình SHA-MTL bao gồm một bộ mã hóa CNN dày đặc và một bộ giải mã upsampling, được kết nối bởi các đơn vị chú ý (AG) với cơ chế chú ý mềm. Các thử nghiệm xác thực chéo được thực hiện trên bộ dữ liệu BUS có nhãn danh mục và mặt nạ, đồng thời thực hiện nhiều phân tích toàn diện trên hai tác vụ. Kết quả họ đánh giá mô hình SHA-MTL trên bộ dữ liệu hình ảnh BUS công khai. Đối với tác vụ phân đoạn, độ nhảy và DICE của mô hình SHA-MTL

đối với các vùng tổn thương lần lượt tăng 2,27% và 1,19% so với mô hình tác vụ đơn lẻ. Độ chính xác phân loại và điểm F1 tăng lần lượt là 2,45% và 3,82%. Kết luận Kết quả xác nhận tính hiệu quả của mô hình của họ và chỉ ra rằng mô hình SHA-MTL yêu cầu ít kiến thức tiên nghiệm hơn để đạt được kết quả tốt hơn bằng cách so sánh với các mô hình gần đây khác. Do đó, họ đã rút ra kết luận rằng việc chú ý nhiều hơn đến vùng tổn thương của BUS sẽ có lợi cho việc phân biệt các loại tổn thương.

Trong bài báo “Dual-Intended Deep Learning Model for Breast Cancer Diagnosis in Ultrasound Imaging” [24] của tác giả Nicolle Vigil. Hệ thống do ông đề xuất đã được đào tạo và xác nhận bằng cách sử dụng 780 hình ảnh siêu âm để phân đoạn và chẩn đoán các tổn thương vú được phân đoạn. Độ chính xác của mô hình rừng ngẫu nhiên thu được thông qua xác thực chéo loại bỏ một lần mang lại 78,5% (65,1–84,1%) cho mô hình xác thực chéo tối đa (đa biến), trong khi mô hình cho truyền thống và phóng xạ sâu cho kết quả độ chính xác lần lượt là 78,8% (64,7–85,5%) và 78,5% (64,5–83,9%). Trong công việc trong tương lai, chúng tôi sẽ mở rộng phân tích của mình cho một tập dữ liệu độc lập bằng cách kết hợp nó với các loại dữ liệu khác để giải quyết khả năng khái quát hóa và độ tin cậy của hệ thống.

## **1.6 Vấn đề đặt ra**

Nghiên cứu và phát triển đề tài “Phát hiện và chẩn đoán ung thư vú bằng hình ảnh CT sử dụng kỹ thuật học sâu” nhằm mục đích:

- Phát hiện tổn thương và khoanh vùng tổn thương do ung thư vú.
- Tinh chỉnh mô hình đạt độ chính xác cao.

## **1.7 Hướng giải quyết của đề tài**

Để thực hiện các mục đích được đề ra ở phần vấn đề đặt ra cần tiến hành:

- Tìm hiểu về bệnh ung thư vú.
- Tìm hiểu về những nghiên cứu liên quan về tổn thương vú do ung thư gây ra.
- Thu thập tập dữ liệu về ảnh CT và ảnh mô tế bào vú.
- Xây dựng mô hình Deep Learning để phát hiện tổn thương vú do ung thư từ ảnh CT và ảnh mô.

## CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

### 2.1 Bệnh ung thư vú

Ung thư vú (breast cancer) là bệnh lý u vú ác tính khi các tế bào ác tính hình thành từ trong mô tuyến vú. Các tế bào ung thư sau đó có thể phát triển lan rộng ra toàn bộ vú và di căn sang các bộ phận khác trên cơ thể. Đây là loại ung thư thường gặp nhất ở nữ giới. [5]

Theo ghi nhận tỷ lệ mắc mới trên toàn cầu của ung thư vú ở nữ giới là 24,5% cao nhất trong số các loại ung thư ở phụ nữ (theo Trung tâm ghi nhận ung thư toàn cầu Globocan, số liệu năm 2020). [10]

Loại ung thư vú thường gặp nhất là ung thư ống tuyến vú (ductal carcinoma), xuất phát từ tế bào của ống tuyến. Ung thư xuất phát từ tiểu thùy và thùy tuyến vú được gọi là ung thư tiểu thùy (lobular carcinoma). Ung thư vú dạng viêm thường có biểu hiện sưng, nóng và đỏ, đây là dạng ung thư vú ít gặp.

#### 2.1.1 Dấu hiệu ung thư vú

Nếu có những triệu chứng này, bạn nên đến gặp bác sĩ ngay để được thăm khám và chẩn đoán kịp thời: [9]

- Xuất hiện khối u ở vú, gần xung quanh vú hoặc dưới nách;
- Dịch từ núm vú đặc biệt dịch có máu;
- Vết lõm da vú hoặc dày da vú;
- Đau nhức vùng vú hoặc núm vú;
- Biểu hiện tụt núm vú;
- Vú có sự thay đổi về kích thước và hình dáng;
- Da vùng vú, quầng vú hoặc núm vú có vảy, đỏ hoặc sưng;
- Vết lõm da vú giống như da quả cam gọi là sần da cam.

#### 2.1.2 Nguyên nhân gây ung thư vú

Theo quy luật tất yếu của tự nhiên, trong cơ thể con người ai cũng mang một nguy cơ ung thư vú tiềm tàng trong suốt cuộc đời và nguy cơ này tăng lên theo độ tuổi và được tính bằng cách cộng dồn nguy cơ trong từng thập niên. Tuy nhiên, có nhiều yếu tố khác làm tăng hoặc thúc đẩy sự phát triển của ung thư vú. Những yếu tố nguy cơ ung thư vú quan trọng có thể kể đến là:

- Lớn tuổi.
- Bản thân có các bệnh ác tính hoặc lành tính tại vú.



- Ít tham gia các hoạt động thể chất, chế độ ăn uống không lành mạnh, sử dụng nhiều rượu bia, thuốc lá.

- Thừa cân, béo phì.

- Phụ nữ có quá trình dậy thì sớm hoặc bước vào mãn kinh muộn.

- Phụ nữ lớn tuổi mới có con đầu lòng hoặc không sinh con hoặc không nuôi con bằng sữa mẹ.

- Phụ nữ một số đột biến về gen như BRCA 1 hoặc BRCA 2. Theo thống kê khoảng 10% số ca ung thư vú có yếu tố gen đột biến.

- Phụ nữ có sử dụng các liệu pháp nội tiết tố thay thế.

- Tiền sử gia đình có người mắc ung thư vú, đặc biệt là quan hệ thế hệ 1 (mẹ, chị em gái, con gái).

- Từng xạ trị vào vùng vú hoặc vùng ngực.

### **2.1.3 Mức độ nguy hiểm của ung thư vú**

Tỷ lệ tử vong do bệnh ung thư vú đứng hàng đầu trong nhóm nguyên nhân tử vong do ung thư ở nữ giới trên toàn thế giới. Mỗi năm nước ta có khoảng hơn 21.555 ca mắc mới và hơn 9.315 ca tử vong. Ung thư vú ở nam giới chiếm khoảng 1% các trường hợp. [8]

Một số dấu hiệu gợi ý ung thư vú bao gồm:

- Vú có sự thay đổi về hình dạng và kích thước;

- Vùng da ở ngực, núm vú hoặc quầng vú xuất hiện vảy, đỏ hoặc sưng;

- Khi sờ phát hiện có khối u ở vú, hoặc vùng xung quanh vú như dưới nách;

- Núm vú bị tụt;

- Vùng vú hoặc núm vú xuất hiện đau nhức;

- Có vết lõm ở da hoặc phần da xung quanh vú dày lên;

- Tiết dịch hoặc dịch có lẫn máu ở đầu vú.

Với những vết lõm ở da như những vết lõm da vú giống như da quả cam được gọi là sần da cam. Ung thư vú ở giai đoạn đầu, phần tổn thương chỉ ảnh hưởng đến ở xung quanh vú, nếu không được nhận biết và thăm khám sớm kết hợp điều trị kịp thời, bệnh sẽ chuyển biến sang các giai đoạn nặng hơn và ảnh hưởng đến nhiều cơ quan trong cơ thể.

Tế bào ung thư khi lây lan vào trong cơ thể sẽ thông qua 3 con đường chính là: mô tế bào lân cận, hệ thống bạch huyết và qua đường máu. Việc các tế bào ung thư xâm lấn và lây lan trong cơ thể được gọi là di căn:

- Mô tế bào lân cận: Lúc này tế bào ung thư đã lan rộng từ cơ quan khởi bệnh và phát triển xâm lấn sang các mô tế bào lân cận.

- Hệ thống bạch huyết: Lúc này tế bào ung thư đã lan rộng từ cơ quan khởi phát bệnh vào hệ thống bạch huyết sau đó di chuyển trong mạch bạch huyết để đến với các cơ quan khác của cơ thể.

- Máu: Sau khi tế bào ung thư xâm lấn từ cơ quan khởi phát bệnh vào máu và di chuyển ngay trong mạch máu để lan đến các cơ quan khác của cơ thể.

Đặc điểm của khối u di căn là có đặc tính giống với khối u nguyên phát (khối u ban đầu). Ví dụ nếu ung thư vú di căn xâm lấn đến xương, tế bào ung thư ở xương sẽ hoạt động tương tự như tế bào ung thư ở vú. Trường hợp này được gọi là ung thư vú di căn xương, không phải là ung thư xương nguyên phát.

Ở giai đoạn khối u nguyên phát di căn đến các mô và cơ quan trong cơ thể ở trường hợp nặng có thể dẫn đến tử vong. Vì vậy việc khám tầm soát rất quan trọng trong việc tìm ra các dấu hiệu bệnh lý trước khi người bệnh xuất hiện các triệu chứng. Mục tiêu của việc tầm soát là phát hiện ra ung thư ở giai đoạn sớm để giúp cho điều trị khỏi hẳn bệnh.

#### **2.1.4 Phân loại ung thư vú**

Ung thư vú không xâm lấn: Với ung thư vú không xâm lấn còn được gọi là ung thư biểu mô tại chỗ hoặc đôi khi được xem là tổn thương tiền ung thư. Với ung thư vú không xâm lấn các tế bào ung thư chỉ giới hạn trong các tiểu thùy vú hoặc trong các ống dẫn sữa. Với ung thư này, các tế bào ung thư chưa phát triển hoặc xâm lấn và mô vú bình thường. Có 2 dạng ung thư vú không xâm lấn là ung thư ống tuyến vú tại chỗ và Ung thư ô tiểu thùy tại chỗ.

Ung thư vú xâm lấn: Khi tế bào ung thư đã lan ra ngoài các ống tuyến hoặc các tiểu thùy của vú đến mô vú xung quanh chúng được gọi là ung thư vú xâm lấn. Các dạng ung thư vú xâm lấn bao gồm:

- Bệnh Paget của núm vú;
- Ung thư biểu mô ống tuyến xâm lấn;
- Ung thư vú di căn;

- Ung thư vú dạng viêm;
- Ung thư biểu mô tiểu thùy xâm lấn;
- Ung thư vú tiến triển tại chỗ;
- U Phyllodes vú.

Ung thư vú khác: Các dạng khác của ung thư vú dựa trên các gen mà bệnh ung thư biểu hiện, như:

- Ung thư vú thụ thể hormone dương tính;
- Ung thư vú HER2 dương tính;
- Ung thư vú bộ ba âm tính.

### ***2.1.5 Các giai đoạn ung thư vú***

+ Ung thư vú giai đoạn 0 (Giai đoạn đầu): Ở giai đoạn đầu này, các tế bào ung thư vú được bác sĩ phát hiện ra trong các ống dẫn sữa. Ta gọi đây là ung thư vú không xâm lấn, hay có cách gọi khác là ung thư biểu mô tuyến vú tại chỗ. Bệnh nhân sẽ được điều trị bệnh ung thư vú để ngăn chặn sự di căn của bệnh. Thường thì bệnh nhân chỉ cần cắt bỏ khối u và sử dụng thêm phương pháp xạ trị. [8]

+ Ung thư vú giai đoạn 1: Ở giai đoạn 1A, khối u vẫn có kích thước nhỏ từ 2cm và các hạch bạch huyết chưa bị ảnh hưởng. Còn khi bệnh đã chuyển sang giai đoạn 1B sẽ không chỉ có khối u ở vú mà còn tìm thấy khối u tại các hạch bạch huyết ở nách. Đây vẫn là 2 giai đoạn phát hiện bệnh sớm. Bác sĩ sẽ sử dụng phương pháp phẫu thuật kết hợp một số liệu pháp để điều trị bệnh.

+ Ung thư vú giai đoạn 2: Ở giai đoạn 2, các khối u có kích thước từ 2 – 5cm và có thể chưa lây lan sang các hạch bạch huyết hoặc sang các hạch nách. Giai đoạn này được chia làm 2 giai đoạn nhỏ: 2A và 2B.

Giai đoạn 2A: Chưa xuất hiện u nguyên phát và chưa đến 4 hạch bạch huyết. Kích thước khối u nhỏ hơn 2cm và chưa đến 4 hạch bạch huyết. Khối u từ 2 – 4cm và chưa lan tới hạch bạch huyết và hạch dưới cánh tay.

Giai đoạn 2B: Ở giai đoạn, khối u có kích thước từ 2 đến 4cm và tìm thấy các cụm tế bào ung thư trong hạch bạch huyết, từ 1-3 hạch bạch huyết ở nách hoặc gần xương ức. Hoặc, kích thước khối u lớn hơn 5cm và chưa xâm lấn đến các hạch bạch huyết.

Bệnh nhân phát hiện bệnh ở giai đoạn 2 nên kết hợp các liệu pháp điều trị phẫu thuật, xạ trị, hóa trị và kích thích tố.

+ Ung thư vú giai đoạn 3: Khi bạn phát hiện bệnh ở giai đoạn 3, thì các khối u trong cơ thể đã lan rộng 4 – 9 hạch bạch huyết ở nách hoặc phủ các hạch bạch huyết bên trong vú. Ở giai đoạn 3, liệu pháp điều trị cho bệnh nhân cũng giống với giai đoạn 2. Nếu bác sĩ phát hiện ra có khối u nguyên phát lớn, thì bạn sẽ phải dùng biện pháp hóa trị để làm nhỏ khối u trước khi tiến hành phẫu thuật.

+ Ung thư vú giai đoạn 4 (Giai đoạn cuối): Giai đoạn này được xem là giai đoạn cuối của bệnh ung thư. Tế bào ung thư đã lan rộng và di căn đến nhiều cơ quan trong cơ thể người bệnh. Ung thư vú thường di căn đến xương, não, phổi và gan. Ở giai đoạn này, bác sĩ sẽ chỉ định cho bạn điều trị toàn thân tích cực, đây là phương pháp phổ biến đối với bệnh nhân đang điều trị bệnh ung thư vú giai đoạn cuối.

### ***2.1.6 Các phương pháp chẩn đoán ung thư vú***

- Khám vú: Việc khám vú như khám ở 2 bên vú cùng phần nách, thượng đòn để kiểm tra hạch các vùng lân cận, kết hợp thăm khám các cơ quan bộ phận khác. Một số triệu chứng ung thư vú cần lưu ý như đau đầu, đau xương, khó thở vì đó là dấu hiệu cho thấy có di căn xa.

- Siêu âm vú: Siêu âm vú thông thường hoặc siêu âm 3D, siêu âm đàn hồi, siêu âm quét thể tích vú tự động sẽ được bác sĩ chỉ định tùy trường hợp.

- Chụp nhũ ảnh: Chụp nhũ ảnh (còn gọi là chụp X-quang tuyến vú) là phương pháp dùng tia X để ghi hình ảnh tuyến vú, với mục đích phát hiện sớm các bất thường nghi ngờ và bệnh ung thư vú. Bác sĩ Albert Salomon (1883 – 1976) là một phẫu thuật viên người Đức. Ông đã chụp X-quang 3000 mẫu bệnh phẩm ung thư vú và nhận thấy rằng đa số các trường hợp ung thư vú sẽ được biểu hiện bằng các nốt vôi hóa nhỏ và hạch nách. Đây là một thành quả hết sức quan trọng, tạo nên nền móng cho việc sử dụng tia X để phát hiện sớm bệnh ung thư vú trên những người chưa có triệu chứng.

Chụp nhũ ảnh thường được chỉ định với phụ nữ trên 40 tuổi với mục đích chẩn đoán khi họ có các triệu chứng như sờ thấy khối u, da vú dày lên hoặc lõm vào, xuất hiện dịch tiết hoặc núm vú bị thụt vào bất thường, các vết loét xuất hiện làm ăn mòn ở núm vú, gây đau đớn.

Với độ nhạy cao đến 70% (tức là nếu có 100 người mắc bệnh ung thư vú, thì nhũ ảnh có thể phát hiện được 70 người) và giá thành hợp lý, cho đến hiện tại, nhũ ảnh được xem là phương pháp tầm soát ung thư vú hiệu quả nhất. Các phương tiện

khác như Siêu âm tuyến vú (Ultrasound) và Cộng hưởng từ tuyến vú (MRI) sẽ đóng vai trò hỗ trợ trong một số tình huống cụ thể và sẽ do bác sĩ chỉ định.

Có 3 loại nhũ ảnh:

- Nhũ ảnh tầm soát: Đối tượng là người chưa có triệu chứng, nhằm phát hiện sớm ung thư vú (giai đoạn tiền lâm sàng).

- Nhũ ảnh chẩn đoán: Đối tượng là người đến khám vì có dấu hiệu bất thường ở vú, mà thường gặp nhất là sờ thấy khối ở vú hoặc tiết dịch ở núm vú.

- Nhũ ảnh theo dõi: Là sau khi chụp nhũ ảnh (tầm soát, chẩn đoán), bác sĩ vẫn chưa kết luận được một cách rõ ràng tình trạng có hay không có ung thư, đồng thời các phương pháp hỗ trợ làm ngay sau đó (ví dụ siêu âm, nhũ ảnh đặc biệt, cộng hưởng từ,...) vẫn không cho kết luận rõ ràng, thì bạn có thể được cho về và hẹn chụp lại nhũ ảnh 6 tháng sau để đối chiếu với lần đầu.

- Chụp cộng hưởng từ: Ở một số trường hợp bác sĩ có thể chỉ định làm cộng hưởng từ khi chưa xác định được bản chất của tổn thương sau khi chụp nhũ ảnh, phần vú đã trải qua phẫu thuật tạo hình trước đó....

- Xét nghiệm dịch từ núm vú: Việc dịch tiết bất thường ở núm vú cũng là 1 trong những dấu hiệu của ung thư vú. Việc xét nghiệm tế bào học dịch tiết ở núm vú có mục đích truy tìm các tế bào ung thư bong ra giúp ích cho việc chẩn đoán.

- Sinh thiết vú: Sinh thiết vú là phương pháp lấy một phần khối u hoặc lấy hoàn toàn khối u vú dưới sự phân tích của các chuyên gia giải phẫu bệnh cùng kính hiển vi về các chỉ số hình thái tế bào và cấu trúc mô vú.

### ***2.1.7 Các phương pháp điều trị ung thư vú***

Phẫu thuật: Phần lớn bệnh nhân ung thư vú đều sẽ được phẫu thuật để cắt bỏ khối u tuyến vú. Một số phương pháp phẫu thuật bao gồm:

- Phẫu thuật cắt vú bảo tồn (breast-conserving surgery): Phẫu thuật cắt bỏ khối u ác tính và một ít tế bào lành tính xung quanh khối u, và không cắt bỏ toàn bộ mô vú còn lại.

- Phẫu thuật cắt toàn bộ vú (total mastectomy): Phẫu thuật cắt toàn bộ vú. Nó còn được gọi là phẫu thuật cắt vú đơn giản (simple mastectomy). Có thể kết hợp lấy một vài hạch nách.

- Phẫu thuật cắt vú triệt để cải biên (modified radical mastectomy): Phẫu thuật cắt bỏ toàn bộ vú, hạch nách và một phần cơ ngực.

**Hóa trị:** Sử dụng thuốc hóa chất để tiêu diệt các tế bào ung thư hoặc làm chậm sự phát triển của tế bào ung thư. Khi thực hiện hóa trị (bằng đường uống hoặc đường tiêm truyền), thuốc sẽ theo đường máu để tìm diệt tế bào ung thư trong cơ thể. Hóa trị có thể thực hiện trước mổ, sau mổ hoặc lúc tế bào ung thư đã cho di căn xa.

**Xạ trị:** Điều trị ung thư vú bằng xạ trị là phương pháp sử dụng các chùm tia năng lượng cao để tiêu diệt tế bào ung thư hoặc ngăn chặn sự phát triển của chúng. Cách thức xạ trị được thực hiện tùy thuộc loại và giai đoạn của ung thư.

**Liệu pháp nhắm trúng đích:** Là sử dụng các loại thuốc hoặc chế phẩm có tác động “chọn lọc” lên một hay nhiều đặc tính riêng biệt của các loại tế bào ung thư. So với hóa trị hoặc xạ trị, liệu pháp trúng đích thường ít gây tác dụng phụ lên tế bào lành.

**Liệu pháp nội tiết tố:** Là phương pháp điều trị để loại bỏ nội tiết tố hoặc ức chế sự hoạt động của nội tiết tố từ đó làm ngưng sự phát triển của tế bào ung thư.

**Liệu pháp miễn dịch:** Là biện pháp điều trị sử dụng hệ thống miễn dịch của bệnh nhân để chống lại bệnh ung thư.

### ***2.1.8 Các cách phòng ngừa ung thư vú***

**Dinh dưỡng:** Hiện nay, không có một loại thực phẩm hay chế độ ăn có thể ngăn chặn ung thư vú. Tuy nhiên việc lựa chọn đúng thực phẩm và chế độ ăn có thể giúp cơ thể bạn khỏe mạnh nhất có thể, tăng cường hệ thống miễn dịch của bạn và giúp giảm nguy cơ mắc bệnh ung thư vú ở mức thấp nhất có thể. Và không có thực phẩm hoặc chế độ ăn uống nào có thể chữa khỏi ung thư, mặc dù một số chúng có thể giúp kiểm soát tác dụng phụ của việc điều trị hoặc giúp cơ thể bạn khỏe lại sau khi điều trị. Một số lựa chọn thực phẩm có thể giúp điều trị ung thư hiệu quả hơn hoặc có thể giúp bạn khỏe mạnh. [9]

- Giữ trọng lượng cơ thể của bạn trong mức phù hợp.
- Ăn nhiều rau và trái cây.
- Hạn chế lượng chất béo bão hòa tiêu thụ dưới 10% tổng lượng calo mỗi ngày và lượng chất béo tiêu thụ được khuyến cáo nên ở khoảng 30gr mỗi ngày.
- Sử dụng thêm các thực phẩm giàu omega-3 và axit béo, hạn chế sử dụng các chất béo chuyển hóa, thực phẩm chế biến sẵn chứa nhiều phụ gia.

**Tập thể dục:** Phụ nữ tập thể dục hơn 4 tiếng mỗi tuần được ghi nhận có nguy cơ mắc ung thư vú thấp hơn phụ nữ không tập thể dục. Hiệu quả của việc tập thể dục

trong việc làm giảm nguy cơ mắc ung thư vú được ghi nhận rõ rệt ở phụ nữ tiền mãn kinh có cân nặng bình thường hoặc thấp.

Tự khám vú tại nhà: Theo bác sĩ Trần Vương Thảo Nghi, tự khám vú là việc làm quan trọng giúp phát hiện sớm ung thư vú nhờ sớm nhận ra những dấu hiệu bất thường xuất hiện ở vú.

Việc kiểm tra vú hoàn toàn có thể tự thực hiện tại nhà và định kỳ mỗi tháng bạn nên kiểm tra 1 lần, thường là vào ngày thứ 7 – 10 của chu kỳ kinh (Ngày 1 của chu kỳ kinh = ngày bắt đầu thấy kinh) vì ở thời điểm này phần vú mềm nhất, bạn sẽ dễ dàng tự khám để phát hiện bất thường. Việc tự khám có thể dễ dàng thực hiện theo hướng dẫn ở hình dưới.



Hình 2.1 Hướng dẫn tự khám vú tại nhà (Tâm Anh Hospital, 2022)

Tầm soát: Việc chẩn đoán ung thư vú thường dựa trên nền tảng là tầm soát bệnh, bắt đầu với việc tự khám vú đều đặn hàng tháng để phát hiện sớm các bất thường. Phần lớn các khối u này là những thay đổi lành tính của tuyến vú, chỉ có khoảng 10 – 20% khối u vú không may là ác tính.

Với những phụ nữ từ 40-49 tuổi nếu không có triệu chứng, việc chụp nhũ ảnh có thể được cân nhắc thực hiện hàng năm và mỗi 1-2 năm/lần với người trên 50 tuổi.

## 2.2 Machine Learning

### 2.2.1 Khái niệm Machine Learning

Machine Learning là một nhánh của trí tuệ nhân tạo (AI) và khoa học máy tính, tập trung vào việc sử dụng dữ liệu và thuật toán để bắt chước hành động của con người, dần dần cải thiện độ chính xác của nó.

Machine Learning còn là một thành phần quan trọng của lĩnh vực khoa học dữ liệu đang phát triển. Thông qua việc sử dụng các phương pháp thống kê, các thuật toán được đào tạo để đưa ra các phân loại hoặc dự đoán và khám phá những thông tin chi tiết từ chính các dự án khai thác dữ liệu.

Thông qua các thông tin chi tiết có được để thúc đẩy việc đưa ra quyết định đối với các ứng dụng và doanh nghiệp, tác động mạnh đến các chỉ số tăng trưởng. Khi dữ liệu lớn tiếp tục nhu cầu mở rộng và phát triển đòi hỏi nhu cầu tuyển dụng các nhà khoa học dữ liệu sẽ tăng lên. Họ sẽ được yêu cầu giúp xác định các câu hỏi kinh doanh có liên quan nhất và dữ liệu để trả lời chúng.

Bài toán của Machine Learning thường được chia làm hai loại là dự đoán (prediction) và phân loại (classification). Các bài toán dự đoán thường là giá nhà, giá xe, v.v, còn các bài toán phân loại thường là nhận diện chữ viết tay, đồ vật, v.v.

## **2.2.2 Các phương pháp Machine Learning**

### *2.2.2.1 Supervised learning*

Supervised learning còn được gọi là máy học có giám sát, được định nghĩa bằng cách sử dụng các tập dữ liệu được gắn nhãn để huấn luyện các thuật toán phân loại dữ liệu hoặc dự đoán kết quả một cách chính xác.

Khi dữ liệu đầu vào được đưa vào mô hình, mô hình sẽ điều chỉnh trọng lượng của nó cho đến khi nó được lắp một cách thích hợp. Điều này xảy ra như một phần của quá trình xác nhận chéo để đảm bảo rằng mô hình tránh trang bị quá nhiều hoặc trang bị thiếu thông tin.

Supervised Machine Learning giúp các tổ chức giải quyết nhiều vấn đề trong thế giới thực trên quy mô lớn, chẳng hạn như phân loại thư rác trong một thư mục riêng biệt từ hộp thư đến của bạn.

Một số phương pháp được sử dụng trong Supervised Machine Learning bao gồm: logistic regression, neural networks, linear regression, naive bayes, random forest, và support vector machine (SVM).

### *2.2.2.2 Unsupervised learning*

Unsupervised Machine Learning là phương pháp sử dụng các thuật toán máy học để phân tích và phân cụm các tập dữ liệu không được gắn nhãn.

Không cần sự can thiệp của con người, các thuật toán này có thể phát hiện ra các mẫu hoặc nhóm dữ liệu ẩn. Khả năng phát hiện ra những điểm tương đồng và



khác biệt trong thông tin của phương pháp này khiến nó trở nên lý tưởng cho việc phân tích dữ liệu khám phá, chiến lược bán chéo (cross-sell), phân khúc khách hàng cũng như nhận dạng hình ảnh và mẫu.

Unsupervised Machine Learning cũng được sử dụng để giảm số lượng các tính năng trong một mô hình thông qua quá trình giảm kích thước. Phân tích thành phần chính (PCA) và phân tích giá trị đơn lẻ (SVD) là hai cách tiếp cận phổ biến cho việc này.

Các thuật toán khác được sử dụng trong học tập không giám sát bao gồm: k-means clustering, neural networks, và probabilistic clustering methods.

### *2.2.2.3 Semi-supervised learning*

Semi-supervised learning cung cấp một phương pháp hiệu quả giữa học tập có giám sát và không giám sát. Trong quá trình đào tạo, nó sử dụng một tập dữ liệu có nhãn nhỏ hơn để hướng dẫn phân loại và trích xuất tính năng từ một tập dữ liệu lớn hơn, không được gắn nhãn.

Phương pháp Semi-supervised learning có thể giải quyết vấn đề không có đủ dữ liệu được gắn nhãn cho thuật toán học có giám sát. Nó cũng hữu ích nếu quá tốn kém để gắn nhãn đủ dữ liệu.

### **2.2.3 Các thuật toán phổ biến của Machine Learning**

Neural networks: Mô phỏng cách thức hoạt động của bộ não con người, với một số lượng khổng lồ các nút xử lý được liên kết. Neural networks là thuật toán được dùng trong việc nhận dạng các mẫu và đóng một vai trò quan trọng trong các ứng dụng bao gồm dịch ngôn ngữ tự nhiên, nhận dạng hình ảnh, nhận dạng giọng nói và tạo hình ảnh.

Linear regression: Thuật toán này được sử dụng để dự đoán các giá trị số, dựa trên mối quan hệ tuyến tính giữa các giá trị khác nhau.

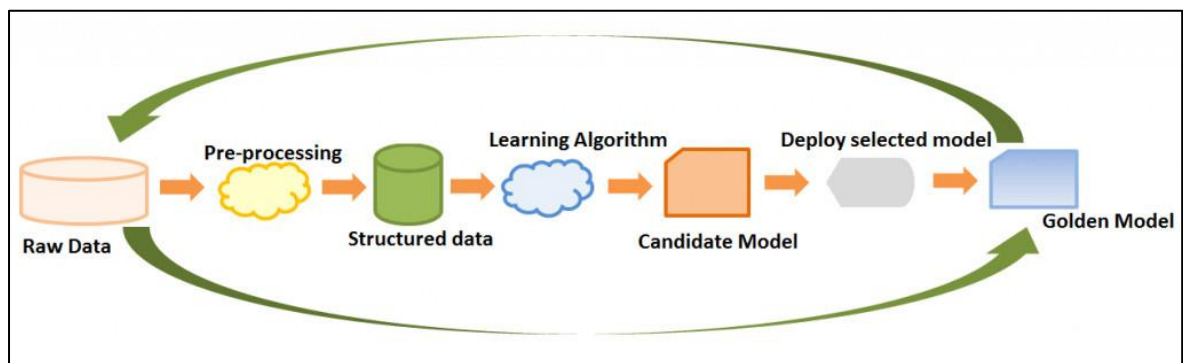
Logistic regression: Thuật toán giúp đưa ra dự đoán cho các biến phản hồi phân loại, chẳng hạn như câu trả lời “có/không” cho các câu hỏi. Nó có thể được sử dụng cho các ứng dụng như phân loại thư rác và kiểm soát chất lượng trên dây chuyền sản xuất.

Clustering: Các thuật toán phân cụm có thể xác định các mẫu trong dữ liệu để nó có thể được nhóm lại. Máy tính có thể giúp các nhà khoa học dữ liệu bằng cách xác định sự khác biệt giữa các mục dữ liệu mà con người đã bỏ qua.

Decision trees: Là thuật toán được sử dụng để dự đoán giá trị số (hồi quy) và phân loại dữ liệu. Decision trees sử dụng một chuỗi phân nhánh của các quyết định được liên kết có thể được biểu diễn bằng sơ đồ cây. Một trong những ưu điểm của decision trees là chúng dễ xác thực và kiểm tra, không giống thuật toán Neural networks.

Random forests: Trong một khu rừng ngẫu nhiên, thuật toán máy học dự đoán một giá trị hoặc danh mục bằng cách kết hợp các kết quả từ một số cây quyết định.

#### 2.2.4 Machine Learning Workflow



Hình 2.2 Mô tả Machine Learning Workflow (Deepika Koundal, 2022)

MLW sẽ cho thấy quy trình để làm việc với Machine Learning như thế nào. Cụ thể từng bước trong Machine Learning Workflow như sau:

- Data collection - thu thập dữ liệu: để máy tính có thể học được bạn cần có một bộ dữ liệu (dataset), bạn có thể tự thu thập chúng hoặc lấy các bộ dữ liệu đã được công bố trước đó. Lưu ý là bạn phải thu thập từ nguồn chính thống, có như vậy dữ liệu mới chính xác và máy có thể học một cách đúng đắn và đạt hiệu quả cao hơn.

- Preprocessing - tiền xử lý: bước này dùng để chuẩn hóa dữ liệu, loại bỏ các thuộc tính không cần thiết, gán nhãn dữ liệu, mã hóa một số đặc trưng, trích xuất đặc trưng, rút gọn dữ liệu nhưng vẫn đảm bảo kết quả các kỹ thuật tiền xử lý dữ liệu phổ biến hiện nay bao gồm: xử lý dữ liệu bị khuyết (missing data), mã hóa các biến nhóm (encoding categorical variables), chuẩn hóa dữ liệu (standardizing data), co giãn dữ liệu (scaling data),... Bước này tốn thời gian nhất tỉ lệ thuận với số lượng dữ liệu bạn có. Bước 1 và 2 thường chiếm hơn 70% tổng thời gian thực hiện.

- Training model - huấn luyện mô hình: bước này là bước bạn huấn luyện cho mô hình hay chính là cho nó học trên dữ liệu bạn đã thu thập và xử lý ở hai bước đầu.

- Evaluating model - đánh giá mô hình: sau khi đã huấn luyện mô hình xong, chúng ta cần dùng các độ đo để đánh giá mô hình, tùy vào từng độ đo khác nhau mà mô hình cũng được đánh giá tốt hay không khác nhau. Độ chính xác của mô hình đạt trên 80% được cho là tốt.

- Improve - cải thiện: sau khi đã đánh giá mô hình, các mô hình đạt độ chính xác không tốt thì cần được train lại, chúng ta sẽ lặp lại từ bước 3, cho đến khi đạt độ chính xác như kỳ vọng. Tổng thời gian của 3 bước cuối rơi vào khoảng 30% tổng thời gian thực hiện

### **2.2.5 Ứng dụng thực tế của Machine Learning**

Speech recognition: Dùng để nhận dạng giọng nói tự động (ASR), nhận dạng giọng nói máy tính hoặc chuyển giọng nói thành văn bản. Đây là một khả năng sử dụng xử lý ngôn ngữ tự nhiên (NLP) để dịch giọng nói của con người sang định dạng viết.

Customer service: Chatbots trực tuyến đang thay thế các tác nhân con người trong hành trình của khách hàng, thay đổi cách chúng ta nghĩ về sự tương tác của khách hàng trên website và nền tảng xã hội.

Computer vision: Công nghệ AI này cho phép máy tính lấy thông tin có ý nghĩa từ video, hình ảnh kỹ thuật số và các đầu vào trực quan khác, sau đó thực thi hành động thích hợp.

Recommendation engines: Sử dụng dữ liệu hành vi tiêu dùng trong quá khứ, các thuật toán AI learning có thể giúp khám phá các xu hướng dữ liệu có thể được sử dụng để phát triển các chiến lược cross-sell hiệu quả hơn.

Automated stock trading: Được thiết kế để tối ưu hóa danh mục đầu tư chứng khoán, các nền tảng giao dịch tần suất cao do AI điều khiển để hàng triệu giao dịch mỗi ngày mà không cần đến sự can thiệp của con người.

Fraud detection: Các ngân hàng và các tổ chức tài chính có thể sử dụng máy học để phát hiện các giao dịch đáng ngờ.

## **2.3 Deep Learning**

### **2.3.1 Khái niệm Deep Learning**

Deep Learning là một tập hợp con của học máy, về cơ bản là một mạng nơ-ron có ba lớp trở lên. Những mạng lưới thần kinh này cố gắng mô phỏng hành vi của não người cho phép Deep Learning “học” từ một lượng lớn dữ liệu. Trong khi mạng nơ-

ron có một lớp vẫn có thể đưa ra các dự đoán gần đúng, các lớp ẩn bổ sung có thể giúp tối ưu hóa và tinh chỉnh để có độ chính xác.

Deep Learning thúc đẩy nhiều ứng dụng và dịch vụ trí tuệ nhân tạo (AI) nhằm cải thiện tự động hóa, thực hiện các tác vụ phân tích và vật lý mà không cần sự can thiệp của con người.

Công nghệ Deep Learning được ứng dụng rộng rãi trong các sản phẩm và dịch vụ hàng ngày chẳng hạn như trợ lý kỹ thuật số, điều khiển từ xa hỗ trợ giọng nói và phát hiện gian lận thẻ tín dụng cũng như các công nghệ mới nổi chẳng hạn như ô tô tự lái.

### ***2.3.2 Cách thức hoạt động của Deep Learning***

Học sâu Deep Learning hoạt động bằng cách khám phá các cấu trúc phức tạp trong dữ liệu mà chúng trải nghiệm. Cụ thể là xây dựng các mô hình tính toán bao gồm nhiều lớp xử lý, mạng có thể tạo ra nhiều mức trừu tượng để biểu diễn dữ liệu.

Ví dụ, một mô hình học sâu được gọi là mạng nơ-ron phức hợp có thể được đào tạo bằng cách sử dụng số lượng lớn (hàng triệu) hình ảnh, chẳng hạn như những hình ảnh có chứa mèo.

Loại mạng nơ-ron này thường học hỏi từ các pixel có trong hình ảnh mà nó thu được. Nó có thể phân loại các nhóm pixel đại diện cho các đặc điểm của mèo, với các nhóm đặc điểm như móng vuốt, tai và mắt cho biết sự hiện diện của mèo trong hình ảnh.

Đối với nhiều tác vụ, chẳng hạn như thị giác máy tính, nhận dạng giọng nói còn được gọi là xử lý ngôn ngữ tự nhiên, dịch máy và robot, hiệu suất của các hệ thống Deep Learning vượt xa so với các hệ thống Machine Learning.

Điều này không có nghĩa là việc xây dựng các hệ thống học sâu là tương đối dễ dàng so với các hệ thống học máy thông thường. Mặc dù nhận dạng tính năng tự trị trong học sâu, hàng nghìn siêu tham số (nút bấm) cần được điều chỉnh để mô hình học sâu Deep Learning trở nên hiệu quả.

### ***2.3.3 Lợi thế khi sử dụng Deep Learning so với phương pháp Machine Learning cổ điển***

#### ***2.3.3.1 Tự động hóa các tính năng***

Một trong những lợi thế khi sử dụng Deep Learning là khả năng tự động hóa các tính năng, cụ thể các thuật toán học sâu có thể tự tạo ra các tính năng mới từ một số lượng hạn chế ngay trong dữ liệu đào tạo mà không cần con người can thiệp.

Có nghĩa là học sâu có thể thực hiện những nhiệm vụ phức tạp đòi hỏi yêu cầu kỹ thuật cao với các tính năng mở rộng. Đối với các doanh nghiệp, lợi thế tự động hóa các tính năng của học sâu cho phép triển khai ứng dụng hoặc công nghệ nhanh, ổn định hơn, đem đến độ chính xác cao.

#### *2.3.3.2 Tương thích tốt với kiểu dữ liệu phi cấu trúc*

Ưu điểm lớn nhất của Deep Learning là khả năng hoạt động tốt với dữ liệu phi cấu trúc. Điều này đặc biệt phù hợp khi phần lớn dữ liệu kinh doanh hiện nay đều không có cấu trúc. Hình ảnh, văn bản và giọng nói được cho là những định dạng dữ liệu phổ biến nhất được các doanh nghiệp sử dụng.

Khác với Deep Learning, các thuật toán Machine Learning cổ điển bị hạn chế về khả năng phân tích dữ liệu phi cấu trúc, có nghĩa là lượng thông tin này không được khai thác thường xuyên. Tuy nhiên, Deep Learning sẽ đem đến những tác động lớn nhất đối với dữ liệu phi cấu trúc.

Đào tạo mạng lưới Deep Learning với dữ liệu phi cấu trúc và ghi nhãn phù hợp sẽ giúp các doanh nghiệp có thể tối ưu hóa mọi chức năng từ bán hàng, tiếp thị cho đến tài chính.

#### *2.3.3.3 Khả năng tự học tốt hơn*

Các lớp nơ-ron trong Deep Learning cho phép các mô hình trở nên hiệu quả hơn, cụ thể là thực hiện được các tính năng phức tạp, tác vụ tính toán chuyên sâu. Cụ thể là có thể thực hiện đồng thời nhiều hoạt động phức tạp.

Deep Learning thể hiện sự vượt trội hơn, điều này thể hiện rõ trong các nhiệm vụ nhận thức máy, còn được gọi là khả năng hiểu được các đầu vào như: hình ảnh, âm thanh và video giống như con người, liên quan đến các bộ dữ liệu phi cấu trúc.

Điều này là do khả năng của các thuật toán học sâu để cuối cùng, giúp học hỏi từ các lỗi của chính nó. Deep Learning còn hỗ trợ việc xác minh tính chính xác, cụ thể là các dự đoán/kết quả đầu ra và thực hiện điều chỉnh trong trường hợp cần thiết. Trong một số trường hợp Deep Learning yêu cầu sự hỗ trợ của con người nhằm xác định dữ liệu đầu ra được chính xác nhất.

#### 2.3.3.4 Lợi thế về thuật toán phân tán và song song

Một mạng nơ-ron điển hình hoặc mô hình học sâu cần nhiều ngày để tìm hiểu các tham số xác định mô hình. Các thuật toán song song và phân tán giải quyết điểm khó khăn này bằng cách cho phép các mô hình học sâu được đào tạo nhanh hơn nhiều. Deep Learning sẽ được đào tạo thông qua phương pháp đào tạo cục bộ, GPU hoặc kết hợp cả hai phương thức này với nhau.

Tuy nhiên, một khối lượng lớn các bộ dữ liệu đào tạo có liên quan đến nhau không thể lưu trữ cùng một máy, điều này đã tạo nên dữ liệu song song. Việc dữ liệu hoặc Deep Learning được phân bổ trên nhiều máy giúp quá trình đào tạo được hiệu quả hơn.

#### 2.3.3.5 Tối ưu chi phí

Mặc dù đào tạo mô hình học sâu có thể tốn kém nhiều chi phí, tuy nhiên nếu được đào tạo bài bản, Deep Learning sẽ giúp doanh nghiệp tối ưu những chi tiêu không cần thiết. Trong các ngành công nghiệp như sản xuất, tư vấn hoặc thậm chí bán lẻ, chi phí cho một dự đoán không chính xác hoặc lỗi sản phẩm là rất lớn. Nó thường lớn hơn chi phí đào tạo mô hình học sâu.

Các thuật toán học sâu có thể tạo ra sự thay đổi giữa các tính năng học tập để giảm tỷ lệ sai sót đáng kể giữa các ngành. Điều này đặc biệt đúng khi bạn so sánh các hạn chế của mô hình học máy cổ điển với các thuật toán học sâu.

#### 2.3.3.6 Phân tích nâng cao

Học sâu, khi được áp dụng vào khoa học dữ liệu, có thể đưa ra các mô hình xử lý tốt hơn và hiệu quả hơn. Việc Deep Learning có thể giám sát giúp thúc đẩy sự cải thiện liên tục, nhờ đó đem đến độ chính xác và đạt được kết quả cao. Nó cũng cung cấp cho các nhà khoa học dữ liệu bằng các kết quả phân tích ngắn gọn và đáng tin cậy hơn.

Công nghệ này hỗ trợ hầu hết các phần mềm dự đoán ngày nay với các ứng dụng từ tiếp thị đến bán hàng, nhân sự, tài chính, v.v. Giống với bộ tự động hóa bán hàng và tiếp thị thông minh cũng sử dụng các thuật toán Deep Learning để đưa ra những dự đoán dựa trên dữ liệu lịch sử.

### 2.3.3.7 Khả năng mở rộng

Học sâu có khả năng mở rộng cao do khả năng xử lý một lượng lớn dữ liệu và thực hiện nhiều phép tính đem lại kết quả tốt nhất, giúp tối ưu về chi phí và thời gian. Điều này ảnh hưởng trực tiếp đến năng suất, tính mô-đun và tính di động.

### 2.3.4 Ứng dụng của Deep Learning

Một số ứng dụng của Deep Learning hiện nay, cụ thể:

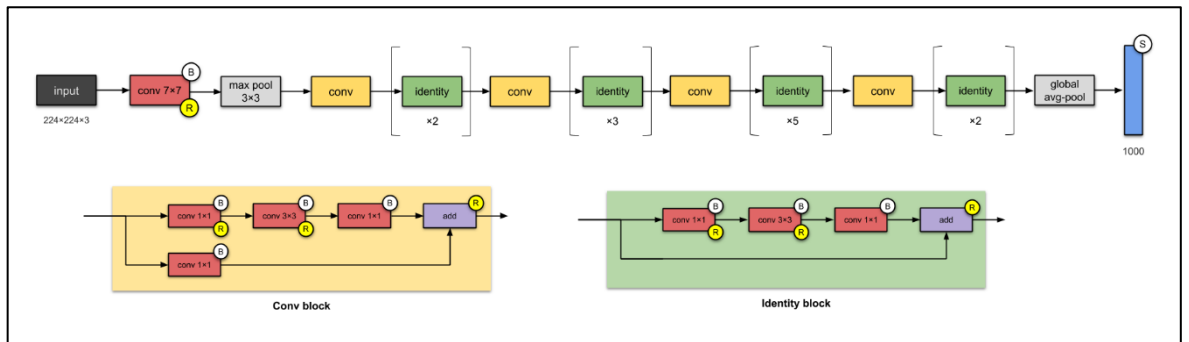
Người máy: Sự phát triển của khoa học công nghệ trong thời gian gần đây, đặc biệt trong lĩnh vực người máy được thúc đẩy bởi những tiến bộ trong AI và Deep Learning.

Nông nghiệp: Ngày nay, Deep Learning cho phép nông dân triển khai thiết bị có thể nhìn thấy và phân biệt giữa cây trồng và cỏ dại. Khả năng này cho phép máy làm cỏ phun thuốc diệt cỏ một cách có chọn lọc lên cỏ dại và để các cây khác không bị ảnh hưởng.

Hình ảnh y tế và chăm sóc sức khỏe: Học sâu đang có những bước tiến đáng kể vào việc cải thiện chất lượng, chăm sóc sức khỏe bằng cách dự đoán các sự kiện y tế từ dữ liệu hồ sơ sức khỏe điện tử.

## 2.4 Mạng học sâu và vai trò

### 2.4.1 Kiến trúc mạng ResNet50



Hình 2.3 Kiến trúc mạng Resnet (Phạm Đình Khánh, 2020)

ResNet là kiến trúc được sử dụng phổ biến nhất ở thời điểm hiện tại. ResNet cũng là kiến trúc sớm nhất áp dụng batch normalization. Kiến trúc với ít tham số nhưng hiệu quả của ResNet đã mang lại chiến thắng trong cuộc thi ImageNet năm 2015.

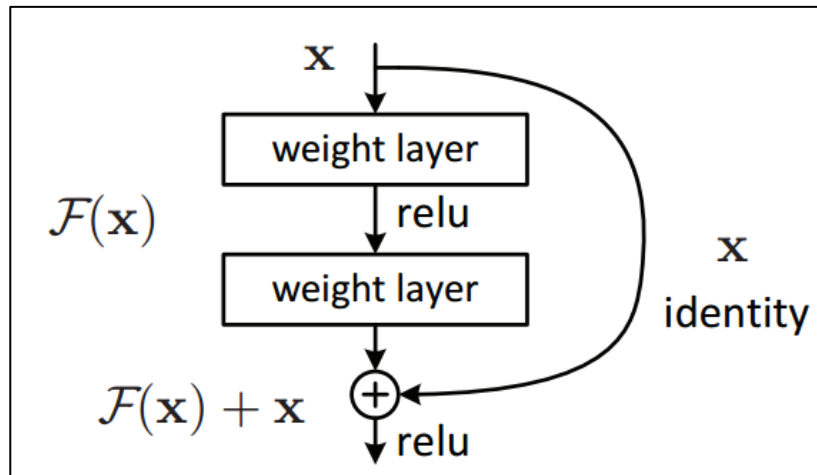
Những kiến trúc trước đây thường cải tiến độ chính xác nhờ gia tăng chiều sâu của mạng CNN. Nhưng thực nghiệm cho thấy đến một ngưỡng độ sâu nào đó thì độ chính xác của mô hình sẽ bão hòa và thậm chí phản tác dụng và làm cho mô hình

kém chính xác hơn. Khi đi qua quá nhiều tầng độ sâu có thể làm thông tin gốc bị mất đi thì các nhà nghiên cứu của Microsoft đã giải quyết vấn đề này trên ResNet bằng cách sử dụng kết nối tắt.

Các kết nối tắt (skip connection) giúp giữ thông tin không bị mất bằng cách kết nối từ layer sớm trước đó tới layer phía sau và bỏ qua một vài layers trung gian. Trong các kiến trúc base network CNN của các mạng YOLOv2, YOLOv3 và gần đây là YOLOv4 bạn sẽ thường xuyên thấy các kết nối tắt được áp dụng.

ResNet có khối tích chập (Convolutional Block, chính là Conv block trong hình) sử dụng bộ lọc kích thước 3x3 giống với của InceptionNet. Khối tích chập bao gồm 2 nhánh tích chập trong đó một nhánh áp dụng tích chập 1x1 trước khi cộng trực tiếp vào nhánh còn lại.

Khối xác định (Identity block) thì không áp dụng tích chập 1 x 1 mà cộng trực tiếp giá trị của nhánh đó vào nhánh còn lại.

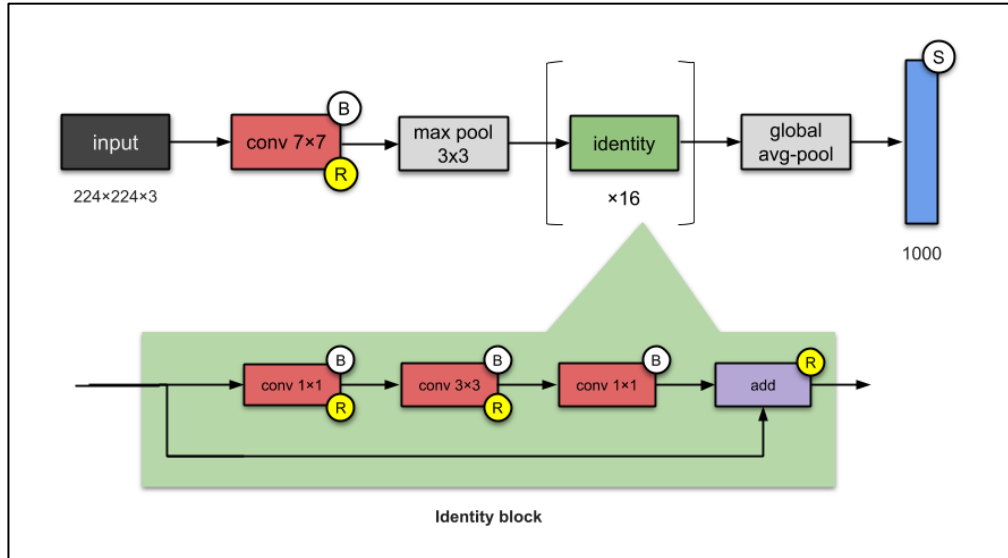


Hình 2.4 Cộng trực tiếp đầu vào của khối với nhánh còn lại trong khối Identity block (Phạm Đình Khánh, 2020)

Giả sử chúng ta có  $x$  là đầu vào của khối xác định. Chúng ta cần ánh xạ đầu vào  $x$  thành hàm  $f(x)$ . Để tìm ra ánh xạ chuẩn xác tương đương với hàm  $f(x)$  là một việc khá khó. Nhưng nếu cộng thêm ở đầu ra thành  $x + f(x)$  thì chúng ta sẽ qui về tham số hóa độ lệch, tức cần tham số hóa phần dư  $f(x)$ . Tìm ánh xạ theo phần dư sẽ dễ hơn nhiều vì chỉ cần tìm giá trị  $f(x)$  sao cho nó gần bằng 0 là có thể thu được một ánh xạ chuẩn xác. Tại một khối xác định, chúng ta sẽ áp dụng một layer activation ReLU sau mỗi xen kẽ giữa những tầng trọng số.



Mặc dù có kiến trúc khối kế thừa lại từ GoogleNet nhưng ResNet lại dễ tóm tắt và triển khai hơn rất nhiều vì kiến trúc cơ sở của nó chỉ gồm các khối tích chập và khối xác định. Ta có thể đơn giản hóa kiến trúc của ResNet-50 như hình bên dưới:



Hình 2.5 Kiến trúc tóm tắt của mạng ResNet50 (Phạm Đình Khánh, 2020)

#### 2.4.2 Kiến trúc mạng DenseNet201

Ở ResNet chúng ta phân tách hàm số thành một hàm xác định và một hàm phi tuyến:  $f(x) = x + g(x)$ . Cùng nhắc lại công thức khai triển Taylor tại  $x = 0$ :

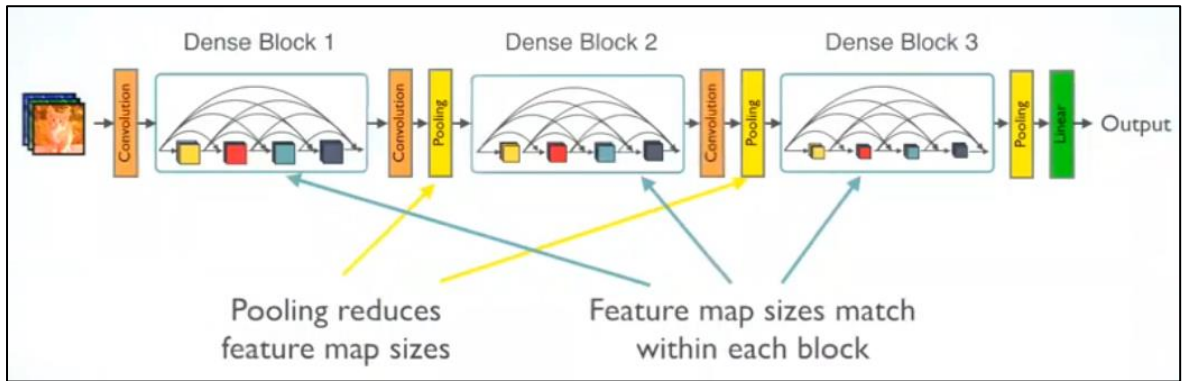
$$f(x) = f(0) + f'(x)x^1 + \frac{f''(x)}{2!}x^2 + \dots + \frac{f^{(n)}(x)}{n!}x^n + o(x^n)$$

Ta có thể thấy công thức của ResNet cũng gần tương tự như khai triển Taylor tại đạo hàm bậc nhất, tương ứng với thành phần số dư. Khai triển Taylor sẽ càng chuẩn xác nếu chúng ta phân rã được số dư thành nhiều đạo hàm bậc cao hơn.

Ý tưởng của DenseNet cũng như vậy, chúng ta sẽ sử dụng một mạng lưới các kết nối tắt dày đặc để liên kết các khối với nhau.

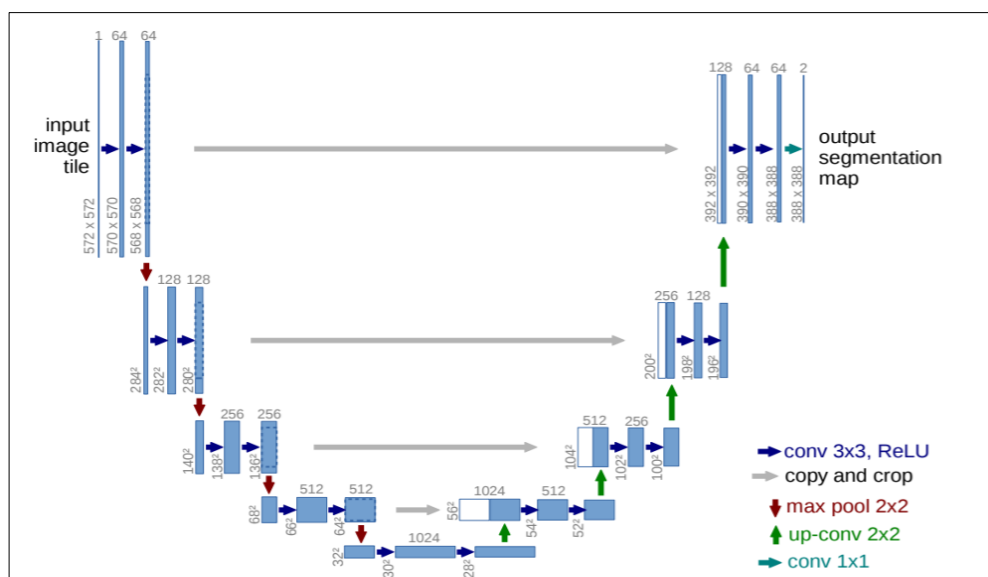
Từ đầu vào  $x$  ta sẽ áp dụng liên tiếp một chuỗi các ánh xạ liên tiếp với cấp độ phức tạp tăng dần:  $x \rightarrow f_1(x) \rightarrow f_2(x, f_1(x)) \dots x \rightarrow f_4(x, f_3(x, f_2(x, f_1(x))))$

DenseNet sẽ khác so với ResNet đó là chúng ta không cộng trực tiếp  $x$  vào  $f(x)$  mà thay vào đó, các đầu ra của từng phép ánh xạ có cùng kích thước dài và rộng sẽ được concatenate với nhau thành một khối theo chiều sâu. Sau đó để giảm chiều dữ liệu chúng ta áp dụng tầng chuyển tiếp (transition layer). Tầng này là kết hợp của một layer tích chập giúp giảm độ sâu và một max pooling giúp giảm kích thước dài và rộng. Các bạn sẽ dễ dàng hình dung hơn qua hình vẽ bên dưới:



Hình 2.6 Kiến trúc DenseNet201 (Phạm Đình Khánh, 2022)

### 2.4.3 Kiến trúc mạng Unet



Hình 2.7 Kiến trúc mạng Unet (Bùi Tiến Tùng, 2020)

Unet là một kiến trúc được phát triển bởi Olaf Ronneberger và các cộng sự nhằm phân vùng các cấu trúc nơ ron thần kinh trong não người. Kiến trúc này lần đầu áp dụng đã dành được chiến thắng trong cuộc thi EM segmentation challenge at ISBI 2012.

Mạng Unet bao gồm 2 nhánh đối xứng nhau hình chữ U nên được gọi là Unet.

Kiến trúc mạng Unet bao gồm 2 phần là phần thu hẹp (contraction) ở bên trái và phần mở rộng (expansion) ở bên phải. Mỗi phần sẽ thực hiện một nhiệm vụ riêng như sau:

+ Phần thu hẹp: Làm nhiệm vụ trích lọc đặc trưng để tìm ra bối cảnh của hình ảnh. Vai trò của phần thu hẹp tương tự như một Encoder. Một mạng Deep CNN sẽ đóng vai trò trích lọc đặc trưng. Lý do nhánh được gọi là thu hẹp vì kích thước dài và

rộng của các layers giảm dần. Từ input kích thước 572x572 chỉ còn 32x32. Đồng thời độ sâu cũng tăng dần từ 3 lên 512.

+ Phần mở rộng: Gồm các layer đối xứng tương ứng với các layer của nhánh thu hẹp. Quá trình Upsampling được áp dụng giúp cho kích thước layer tăng dần lên. Sau cùng ta thu được một ảnh mask đánh dấu nhãn dự báo của từng pixel.

Đặc trưng riêng trong cấu trúc của Unet đó là áp dụng kết nối tắt đối xứng giữa layer bên trái với layer bên phải.

Mặc dù có độ chính xác khá cao nhưng Unet có tốc độ thấp. Do đó nó không phù hợp để áp dụng vào các tác vụ yêu cầu realtime như xe tự hành. Tuy nhiên, Unet lại thường được sử dụng khá phổ biến trong các tác vụ không đòi hỏi realtime vì accuracy của nó cũng không tồi và kiến trúc dễ implement.

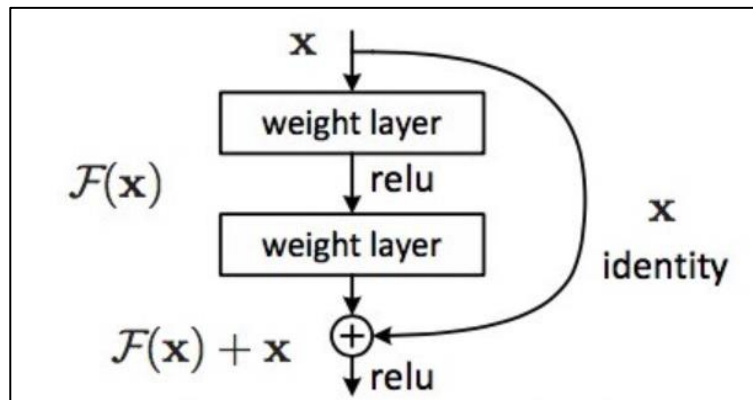
#### 2.4.4 Kiến trúc mạng MobileNetV3Small

Kể từ khi ra đời, MobileNetV2 là một trong những kiến trúc được ưa chuộng nhất khi phát triển các ứng dụng AI trong computer vision. Rất nhiều các kiến trúc sử dụng backbone là MobileNetV2 như SSDLite trong object detection và DeepLabV3+ trong image segmentation.

MobileNetV2 có một số điểm cải tiến so với MobileNetV1 giúp cho nó có độ chính xác cao hơn, số lượng tham số và số lượng các phép tính ít hơn mà chúng ta sẽ tìm hiểu ngay sau đây.

##### 2.4.4.1 Inverted Residual Block

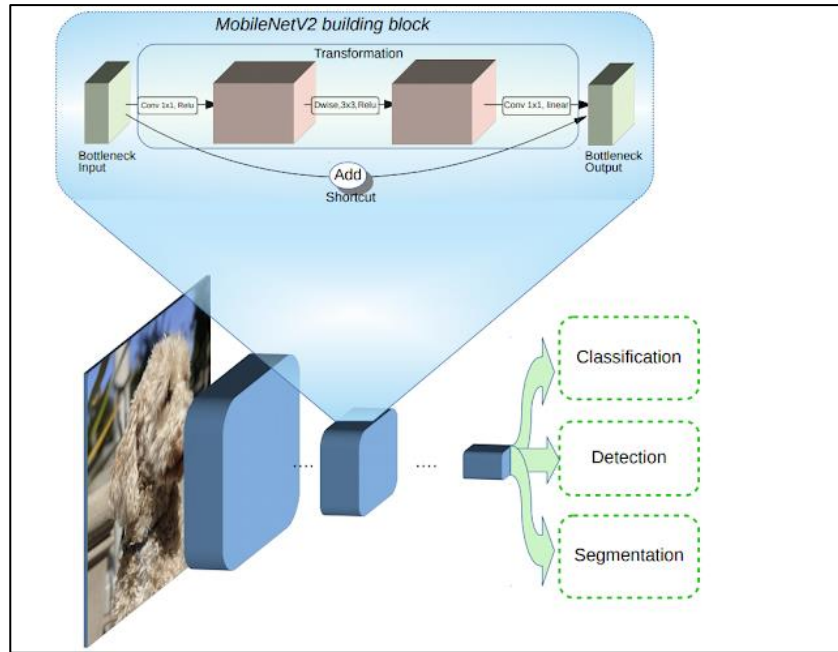
MobileNetV2 cũng sử dụng những kết nối tắt như ở mạng ResNet. Các khối ở layer trước được cộng trực tiếp vào layer liền sau. Nếu coi layer liền trước là  $x$ , sau khi đi qua các xử lý tích chập hai chiều ta thu được kết quả  $F(x)$  thì output cuối cùng là một residual block có giá trị  $x + F(x)$ .



Hình 2.8 Residual learning: a building block (Phạm Đình Khánh, 2022)

Tuy nhiên kết nối tắt ở MobileNetV2 được điều chỉnh sao cho số kênh (hoặc chiều sâu) ở input và output của mỗi block residual được thắt hẹp lại. Chính vì thế nó được gọi là các bottleneck layers (bottleneck là một thuật ngữ thường được sử dụng trong Deep Learning để ám chỉ các kiến trúc thu hẹp kích thước theo một chiều nào đó).

Cụ thể hơn chúng ta theo dõi hình minh họa bên dưới:



Hình 2.9 MobileNetV2 building block (Phạm Đình Khánh, 2022)

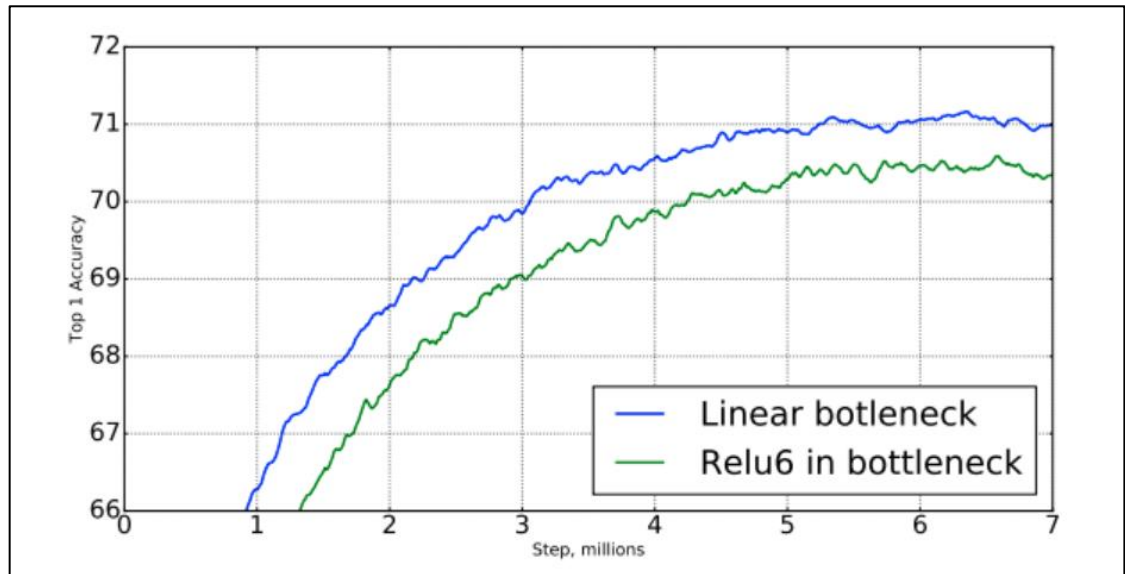
Kiến trúc residual này ngược lại so với các kiến trúc residual truyền thống vì kiến trúc residual truyền thống có số lượng kênh ở input và output của một block lớn hơn so với các layer trung gian. Chính vì vậy nó còn được gọi là kiến trúc inverted residual block.

Tác giả cho rằng các layer trung gian trong một block sẽ làm nhiệm vụ biến đổi phi tuyến nên cần dày hơn để tạo ra nhiều phép biến đổi hơn. Kết nối tắt giữa các block được thực hiện trên những bottleneck input và output chứ không thực hiện trên các layer trung gian. Do đó các layer bottleneck input và output chỉ cần ghi nhận kết quả và không cần thực hiện biến đổi phi tuyến.

Ở giữa các layer trong một block inverted residual block chúng ta cũng sử dụng những biến đổi tích chập tách biệt chiều sâu để giảm thiểu số lượng tham số của mô hình. Đây cũng chính là bí quyết giúp họ các model MobileNet có kích thước giảm nhẹ.

#### 2.4.4.2 Loại bỏ non-linear

Một trong những thực nghiệm được tác giả ghi nhận đó là việc sử dụng các biến đổi phi tuyến (như biến đổi qua ReLu hoặc sigmoid) tại input và output của các residual block sẽ làm cho thông tin bị mất mát. Cụ thể cùng xem kết quả thực nghiệm bên dưới:

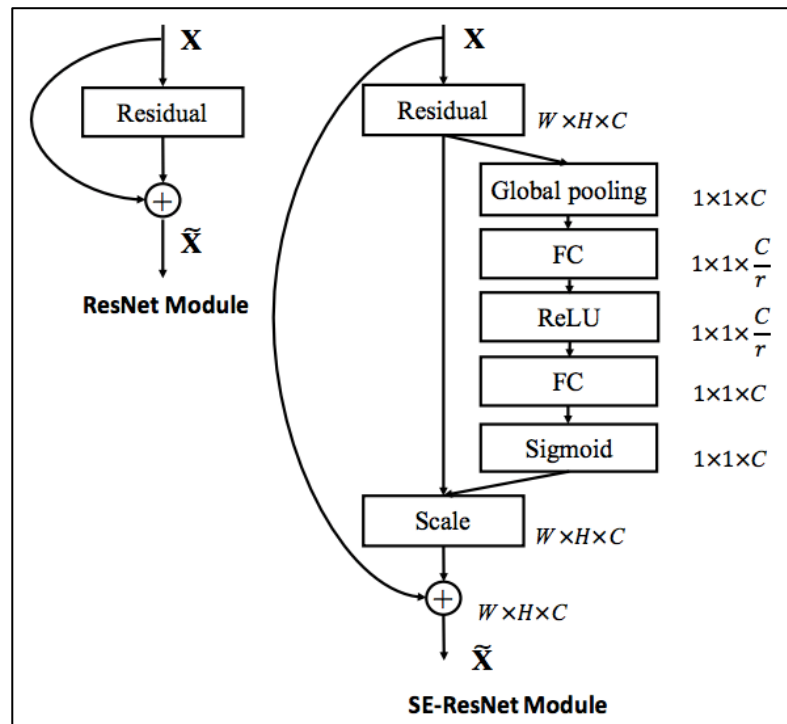


Hình 2.10 Impact of non-linearity bottleneck layer (Phạm Đình Khánh, 2022)

Chính vì thế trong kiến trúc của residual block tác giả đã loại bỏ hàm phi tuyến tại layer input và output và thay bằng các phép chiếu tuyến tính.

#### 2.4.4.3 MobileNetV3

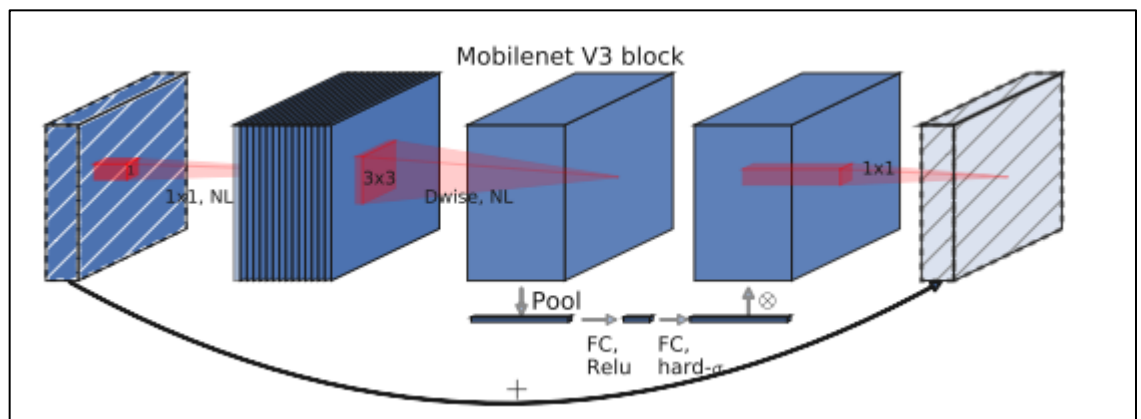
Thêm Squeeze and Excitation (SE) vào block Residual để tạo thành một kiến trúc có độ chính xác cao hơn. Chúng ta cùng phân tích hình minh họa bên dưới để thấy sự khác biệt giữa kiến trúc Residual của mạng ResNet thông thường và SE-ResNet (áp dụng thêm Squeeze and Excitation).



Hình 2.11 Squeeze and Excitation networks (Phạm Đình Khánh, 2022)

SE-ResNet áp dụng thêm một nhánh Global pooling có tác dụng ghi nhận bối cảnh của toàn bộ layer trước đó. Kết quả sau cùng ở nhánh này ta thu được một véc tơ global context được dùng để scale đầu vào  $X$ .

Tương tự như vậy SE được tích hợp vào kiến trúc của một residual block trong mobilenetV3 như sau:



Hình 2.12 MobileNetV3 block (Phạm Đình Khánh, 2022)

Tại layer thứ 3 có một nhánh Squeeze and Excitation có kích thước (width x height) bằng  $1 \times 1$  có tác dụng tổng hợp global context. Nhánh này lần lượt đi qua các biến đổi FC  $\rightarrow$  Relu  $\rightarrow$  FC  $\rightarrow$  hard sigmoid (FC là fully connected layer). Cuối cùng được nhân trực tiếp vào nhánh input để scale input theo global context. Các kiến trúc còn lại hoàn toàn giữ nguyên như MobileNetV2. Để khởi tạo một inverted

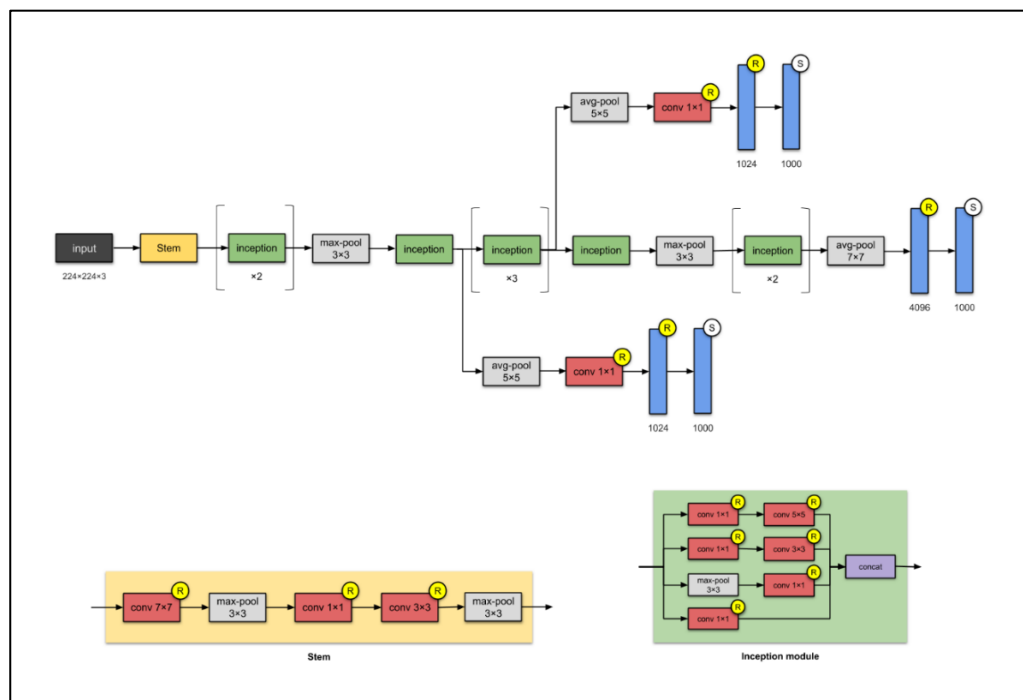
residual block trong mobilenetv3 chúng ta sẽ sử dụng lại code của MobileNetV2 và thêm vào một nhánh SE.

### 2.4.5 Kiến trúc mạng InceptionV3

#### 2.4.5.1 InceptionV1

Mạng Inception-V1 đã giành chiến thắng ở cuộc thi ImageNet vào năm 2015. Kiến trúc này đã giải quyết một câu hỏi lớn trong mạng CNN đó là sử dụng kernel\_size với kích thước bao nhiêu thì hợp lý. Các kiến trúc mạng nơ ron trước đó đều sử dụng các bộ lọc với đa dạng các kích thước 11x11, 5x5, 3x3 cho tới nhỏ nhất là 1x1. Một khám phá được đưa ra bởi bài báo đó là việc cùng kết hợp đồng thời các bộ lọc này vào cùng một block có thể mang lại hiệu quả đó chính là kiến trúc khối Inception.

Khối Inception sẽ bao gồm 4 nhánh song song. Các bộ lọc kích thước lần lượt là 1x1, 3x3, 5x5 được áp dụng trong Inception Module giúp trích lọc được đa dạng đặc trưng trên những vùng nhận thức có kích thước khác nhau.



Hình 2.13 Kiến trúc mạng InceptionV1 (Phạm Đình Khánh, 2022)

Ở đầu các nhánh 1, 2, 4 từ trên xuống, phép tích chập 1x1 được sử dụng trên từng điểm ảnh như một kết nối fully connected nhằm mục đích giảm độ sâu kênh và số lượng tham số của mô hình. Ví dụ: Ở block trước chúng ta có kích thước width x height x channels = 12 x 12 x 256. Sau khi áp dụng 32 bộ lọc kích thước 1x1 sẽ không làm thay đổi width, height và độ sâu giảm xuống 32, output shape lúc này có kích



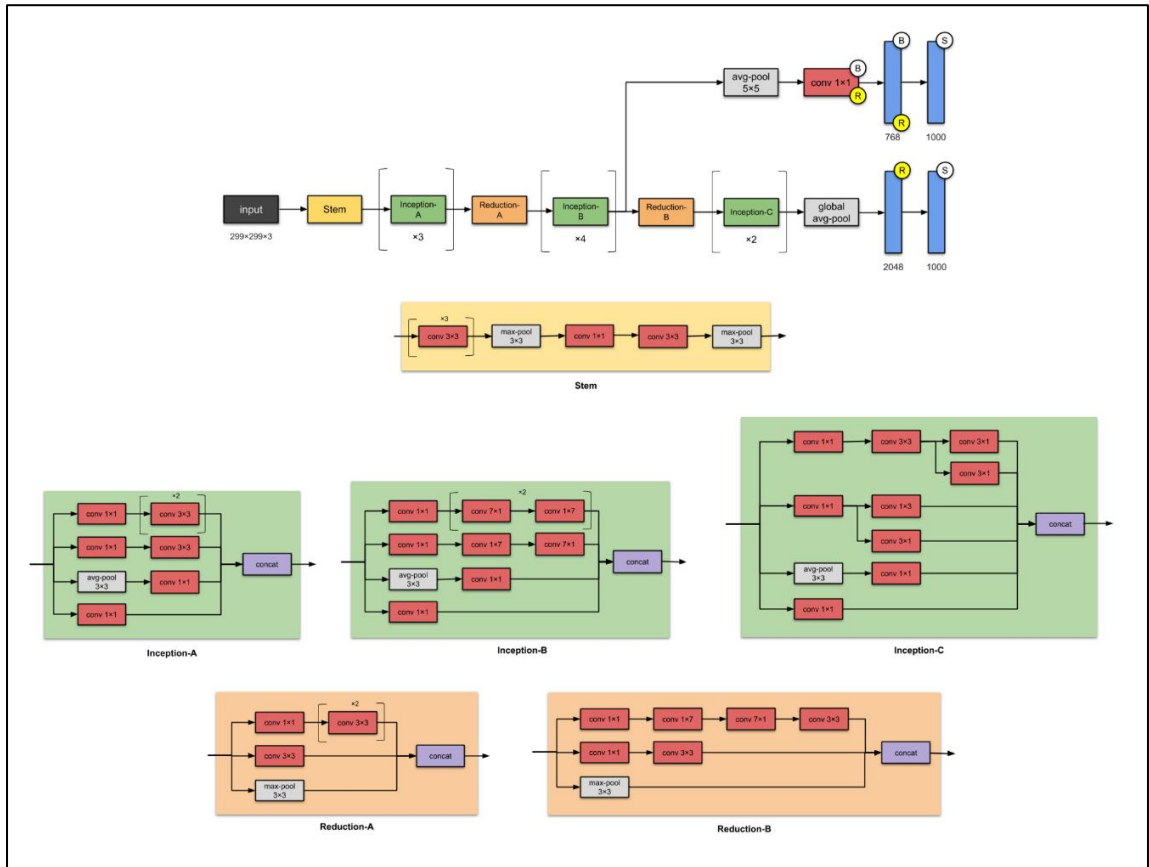
thước là  $12 \times 12 \times 32$ . Ở layer liền sau, khi thực hiện tích chập trên toàn bộ độ sâu, chúng ta chỉ khởi tạo các bộ lọc có độ sâu 32 thay vì 256. Do đó số lượng tham số giảm đi một cách đáng kể.

Nhánh thứ 3 từ trên xuống chúng ta giảm chiều dữ liệu bằng một layer max-pooling kích thước  $3 \times 3$  và sau đó áp dụng bộ lọc kích thước  $1 \times 1$  để thay đổi số kênh.

Các nhánh áp dụng padding và stride sao cho đầu ra có cùng kích cỡ chiều dài và chiều rộng. Cuối cùng ta concatenate toàn bộ kết quả đầu ra của các khối theo kênh để thu được output có kích thước bằng với input.

Khối Inception được lặp lại 7 lần trong kiến trúc Inception-V1. Toàn bộ mạng bao gồm 22 Layers, lớn hơn gần gấp đôi so với VGG-16.

#### 2.4.5.2 InceptionV3



Hình 2.14 Kiến trúc mạng InceptionV3 (Phạm Đình Khánh, 2022)

Inception-V3 là kế thừa của Inception-V1 bao gồm 24 triệu tham số. Toàn bộ các layer tích chập của Inception-V3 được theo sau bởi một layer batch normalization và một ReLU activation. Batch normalization là kỹ thuật chuẩn hóa đầu vào theo từng minibatch tại mỗi layer theo phân phối chuẩn hóa, giúp cho quá trình huấn luyện thuật toán nhanh hơn.



Inception-V3 giải quyết được vấn đề thắt cổ chai (representational bottlenecks). Tức là kích thước của các layers không bị giảm một cách đột ngột. Đồng thời Inception-V3 có một cách tính toán hiệu quả hơn nhờ sử dụng phương pháp nhân tố (factorisation methods).

Hiện tại Inception module bao gồm 3 version. Chúng ta hãy cùng xem qua các điểm đặc biệt ở từng version.

- Inception-A: Cải tiến so với Inception module V1. Tại nhánh thứ nhất thay 1 layer tích chập  $5 \times 5$  bằng 2 layer tích chập  $3 \times 3$  liên tiếp giúp giảm số lượng tham số từ 25 về 18 và tăng độ sâu cho mô hình.

- Inception-B: Cải tiến so với Inception-A. Thay tích chập  $3 \times 3$  bằng tích chập  $7 \times 7$  ở nhánh thứ nhất và nhánh thứ 2. Đồng thời chúng ta phân tích nhân tố tích chập  $7 \times 7$  thành 2 tích chập liên tiếp  $7 \times 1$  và  $1 \times 7$  số lượng tham số sẽ ít hơn so với tích chập  $7 \times 7$  thành 2 tích chập  $3 \times 3$  liên tiếp. Nhờ đó số lượng tham số giảm từ 18 xuống còn 14.

- Inception-C: Cải tiến so với Inception-B. Thay tích chập  $7 \times 1$  bằng tích chập  $3 \times 1$  và  $1 \times 7$  bằng  $1 \times 3$  và đồng thời thay vì đặt layer  $3 \times 1$  và  $1 \times 3$  liên tiếp thì đặt chúng song song. Kiến trúc này giúp giảm số lượng tham số từ 14 về còn 6.

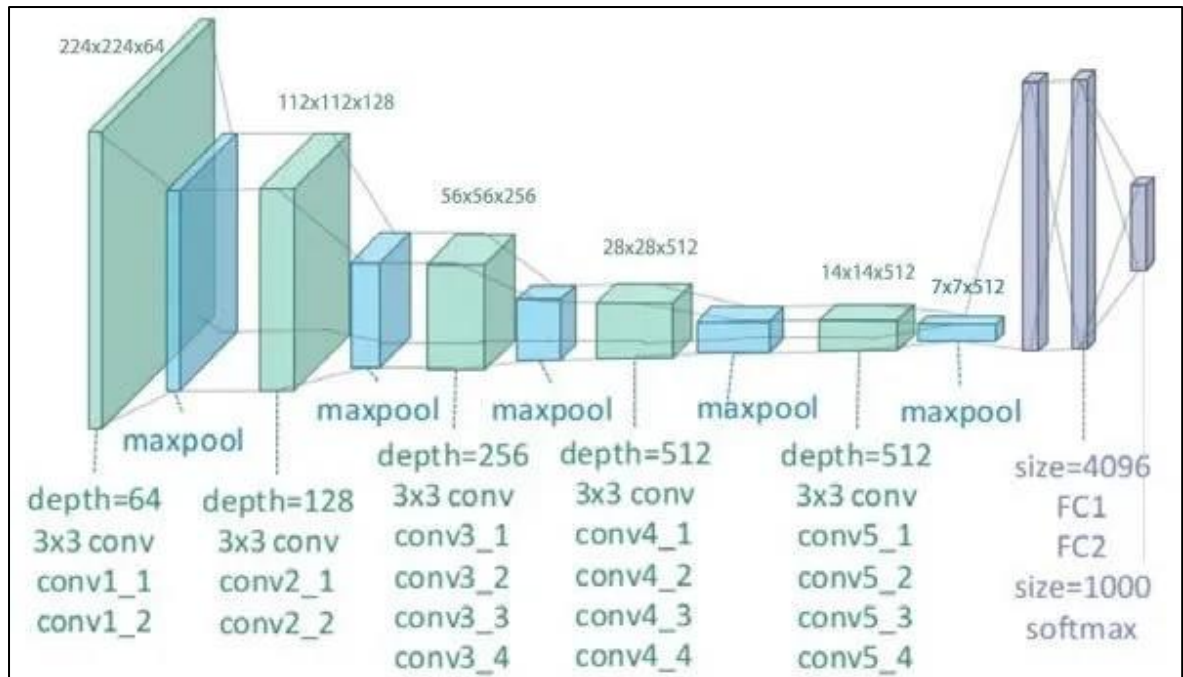
Ngoài ra ở Inception-V3 chúng ta còn sử dụng 2 kiến trúc giảm chiều dữ liệu là Reduction-A và Reduction-B.

#### **2.4.6 Kiến trúc mạng VGG-19**

VGG là một mạng thần kinh tích chập (CNN) có độ sâu 19 lớp. Nó được xây dựng và đào tạo bởi Karen Simonyan và Andrew Zisserman tại Đại học Oxford vào năm 2014. Bạn có thể truy cập tất cả thông tin từ bài báo của họ, Very Deep Convolutional Networks for Large-Scale Image Recognition, được xuất bản vào năm 2015. VGG-19 được đào tạo bằng cách sử dụng hơn 1 triệu hình ảnh từ cơ sở dữ liệu ImageNet, trong đó có các ảnh màu  $224 \times 224$  pixel. Đương nhiên, bạn có thể nhập vào mô hình các trọng số được huấn luyện bởi ImageNet. Điểm đặc biệt, VGG-19 có thể phân loại tới 1000 đối tượng. Dưới đây là thông tin ngắn gọn về kích thước và hiệu suất của nó:

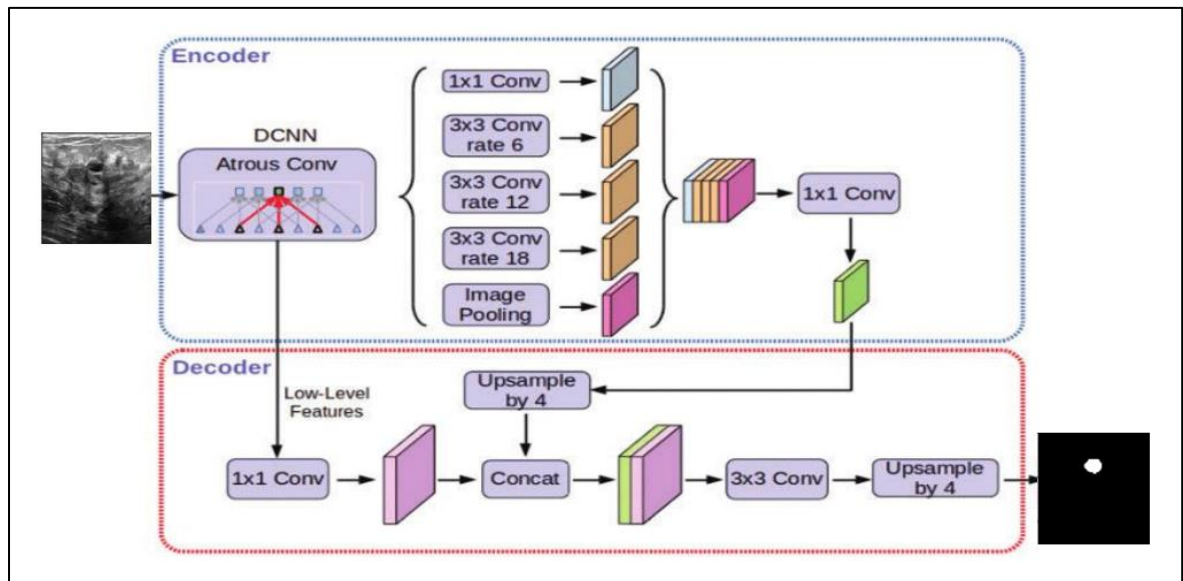
- Kích thước: 549 MB
- Top-1: Độ chính xác: 71,3%
- Top-5: Độ chính xác: 90,0%
- Số tham số: 143.667.240

- Độ sâu: 26



Hình 2.15 Mô hình mạng VGG-19 (Phạm Đình Khánh, 2022)

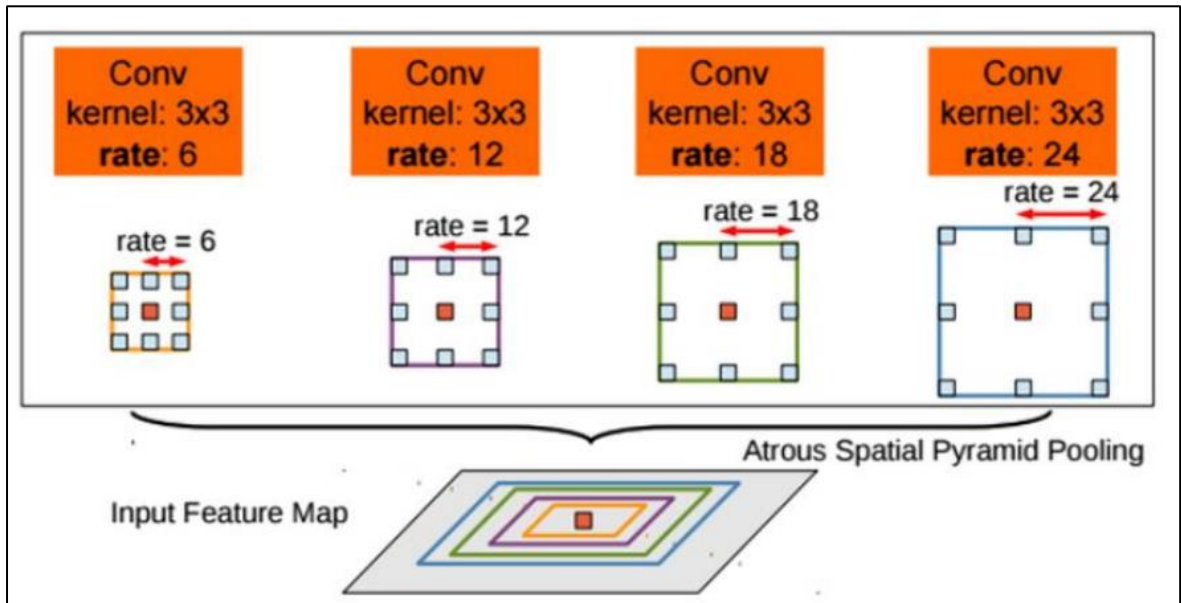
#### 2.4.7. Kiến trúc mạng DeeplabV3+



Hình 2.16 Kiến trúc mạng DeeplabV3+ (Phạm Đình Khánh, 2022)

DeepLabV3+ là một phiên bản nâng cấp của kiến trúc DeepLab được phát triển để giải quyết bài toán phân đoạn hình ảnh (image segmentation) với độ chính xác cao. Nó kết hợp các thành phần quan trọng từ các phiên bản trước đó của DeepLab và áp dụng một số cải tiến để nâng cao hiệu suất và chất lượng phân đoạn. Dưới đây là mô tả tổng quan về kiến trúc DeepLabV3+:

Backbone Network: DeepLabv3+ sử dụng một mạng nơ-ron tích chập (convolutional neural network - CNN) làm phần cốt lõi (backbone) của mô hình. Thông thường, ResNet hoặc Xception được sử dụng làm backbone network. Mạng này được sử dụng để trích xuất các đặc trưng từ hình ảnh đầu vào và tạo ra các biểu diễn đặc trưng có độ phân giải thấp.



Hình 2.17 Mô tả Atrous Spatial Pyramid Pooling (Phạm Đình Khánh, 2022)

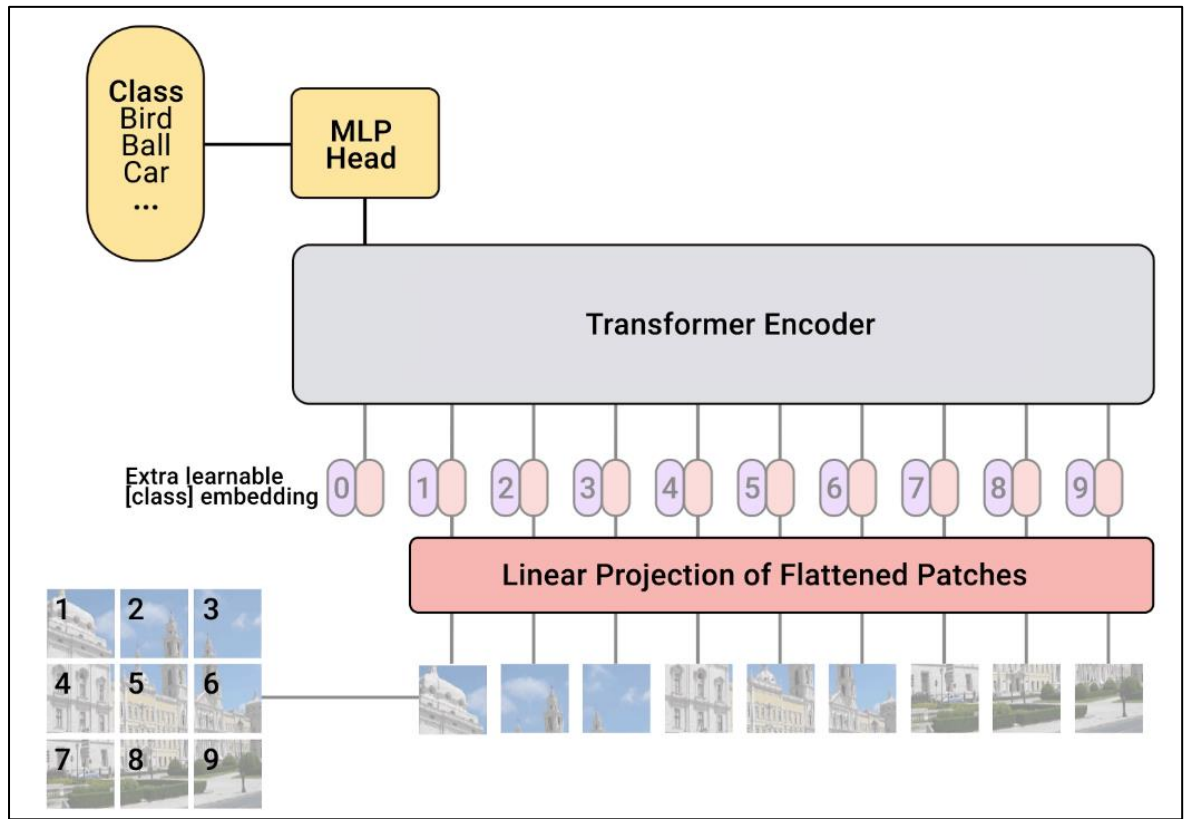
Atrous Spatial Pyramid Pooling (ASPP): DeepLabv3+ sử dụng kiến trúc Atrous Spatial Pyramid Pooling (ASPP) để thu thập thông tin từ các vùng bối cảnh ở nhiều tỷ lệ khác nhau. ASPP sử dụng các lớp atrous convolution ở các tỷ lệ atrous rate khác nhau để tạo ra các biểu diễn đặc trưng với các lĩnh vực nhìn rộng và chi tiết. Điều này giúp mô hình có khả năng phân đoạn các đối tượng có kích thước và hình dạng khác nhau.

Encoder-Decoder Architecture: DeepLabv3+ sử dụng kiến trúc encoder-decoder để kết hợp thông tin chi tiết từ các biểu diễn đặc trưng có độ phân giải thấp và thông tin bối cảnh rộng từ các biểu diễn đặc trưng có độ phân giải cao. Mô hình sử dụng lớp upsampling và skip connections để kết hợp các đặc trưng từ các lớp phân giải thấp và lớp phân giải cao, từ đó tạo ra kết quả phân đoạn chi tiết và chính xác.

DeepLabv3+ là một kiến trúc mạng mạnh mẽ cho bài toán phân đoạn hình ảnh. Nó kết hợp các thành phần quan trọng như ASPP và kiến trúc encoder-decoder để đạt được kết quả phân đoạn chính xác và chi tiết. Cải tiến refinement cũng cải thiện chất lượng kết quả. DeepLabv3+ đã đạt được kết quả ấn tượng trên nhiều bộ dữ liệu phân

đoạn và trở thành một trong những kiến trúc phổ biến trong lĩnh vực này.

## 2.5 Vision transformer



Hình 2.18 Kiến trúc Vision transformer (Lê Minh Chiến, 2021)

Kiến trúc của mô hình gồm 3 thành phần chính:

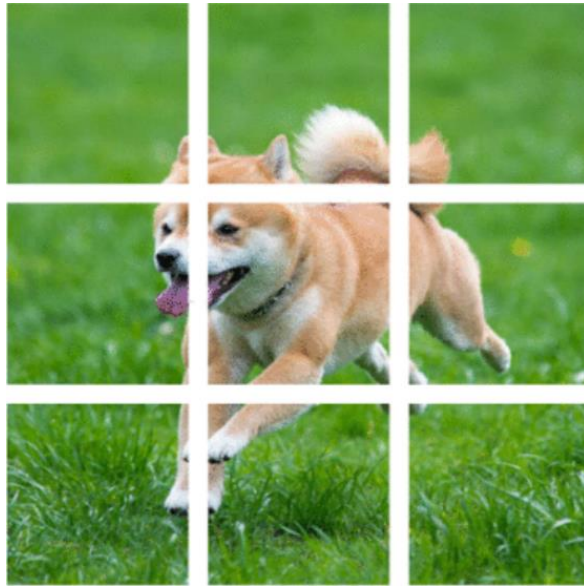
- Linear Projection of Flattened Patches
- Transformer encoder.
- Classification head.

### 2.5.1 Linear Projection and Flattened Patches

#### 2.5.1.1 Patch Embedding

Với các mô hình CNN cho bài toán image classification, ảnh input đầu vào cho mô hình CNN đó là toàn bộ ảnh với kích thước cố định. Tuy nhiên ViT có một cách xử lý khác.

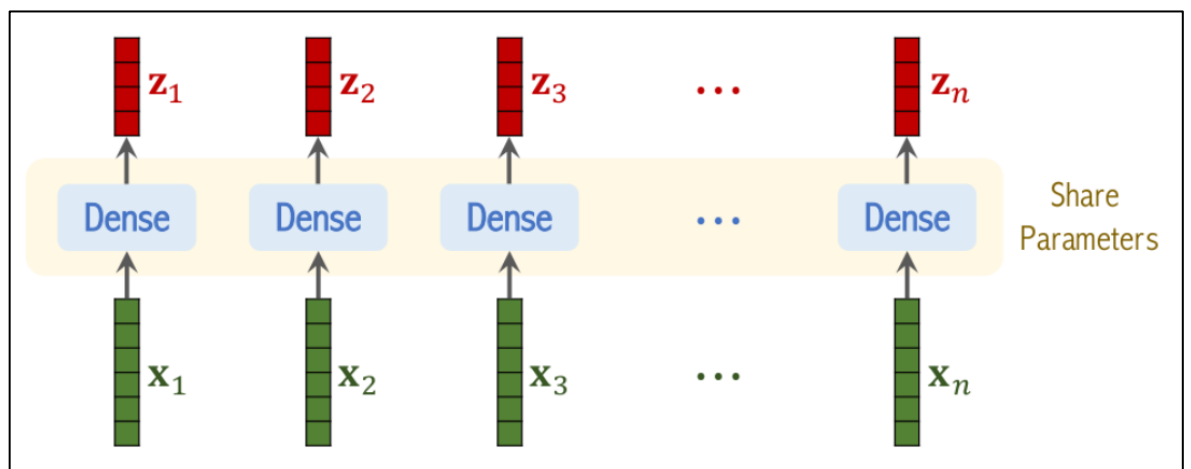
Với mỗi ảnh đầu vào, ViT xử lý bằng cách chia ảnh ra thành các phần có kích thước bằng nhau (patch).



Hình 2.19 Chia batch ảnh gốc (Lê Minh Chiến, 2021)

Ảnh gốc có kích thước là 48x48, ViT sẽ chia ảnh gốc này ra thành các patch có kích thước 16x16. Sau khi chia nhỏ ảnh đầu vào ra ta sẽ có 9 patches tất cả.

Bước tiếp theo, đưa các patches này về dạng vector bằng cách flattend các patches này ra.



Hình 2.20 Linear projection (Lê Minh Chiến, 2021)

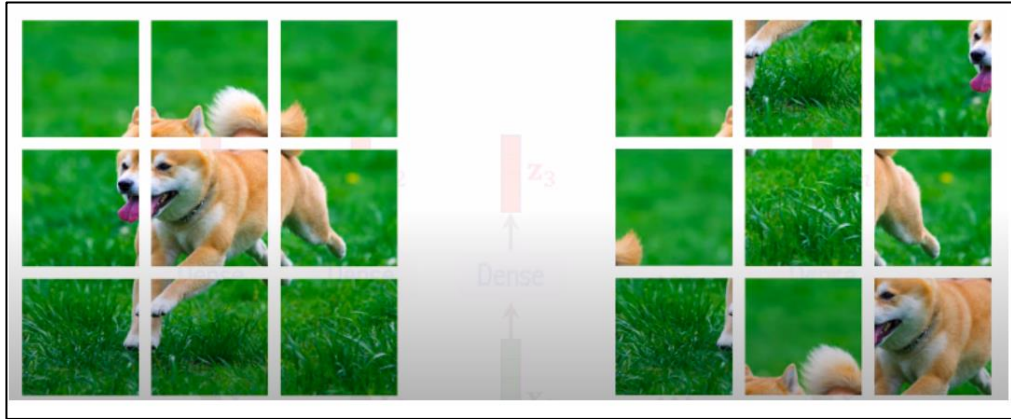
Hình trên mô tả phần Linear Projection. Thực chất Linear Projection là một lớp Dense với đầu vào là flattend vector của các patches, đầu ra sẽ là embedding vector tương ứng với từng patch:  $z_i = W \cdot x_i + b$

Trong đó:

- $x_i$  là flattend vector của patch thứ  $i$
- $z_i$  là output tương ứng của  $x_i$  khi qua Linear Projection
- $W$  được gọi là ma trận embedding.

### 2.5.1.2 Positional Embedding

Ý tưởng tương tự với mô hình Transformer gốc. Positional embedding trong mô hình ViT sẽ chứa thông tin về vị trí của patch trong ảnh (spatial information). Vậy tại sao với dữ liệu là ảnh mà ta vẫn cần spatial information. Ví dụ như hình dưới đây:

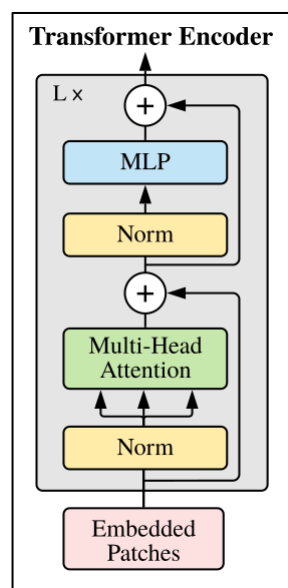


Hình 2.21 Positional Embedding (Lê Minh Chiến, 2021)

Nếu như ta chỉ Embedding các patch và đưa vào mô hình Transformer thì với 2 ảnh ở bên trên sẽ hoàn toàn không có sự khác biệt. Do đó ta cần thêm thông tin về vị trí cho mỗi patch.

Sau khi có vector positional embedding cho mỗi patch ta sẽ cộng các vector này tương ứng với embedding vector của từng patch đã tính ở trên và thu được các vector embedding vừa chứa thông tin của vùng ảnh vừa chứa thông tin về vị trí của nó trong ảnh.

### 2.5.2 Transformer Encoder



Hình 2.22 Kiến trúc của Transformer Encoder (Lê Minh Chiến, 2021)



### 2.5.2.1 Self Attention layer

Self attention layer là thành phần chính để tạo nên một block trong Transformer Encoder.

Đầu vào của Self attention layer là một chuỗi  $X = [x_1, x_2, x_3, \dots, x_m]$

Đầu ra của Self Attention layer là một context vector  $C$  chứa những thông tin quan trọng nhất của chuỗi đầu vào  $C = [c_1, c_2, c_3, \dots, c_m]$ .

Các parameters của layer này bao gồm  $W_Q, W_K, W_V$ .

Cách thức hoạt động của Self attention layer:

- Bước 1: Ứng với mỗi  $x_i$  của chuỗi đầu vào  $X$  tính toán các giá trị  $q_i, k_i, v_i$  tương ứng theo công thức  $q_i = W_Q x_i, k_i = W_K x_i, v_i = W_V x_i$

- Bước 2: Tính alignment score tương ứng với  $x_i$ , theo công thức  $\alpha_i = \text{Softmax}(K^T q_i)$

- Bước 3: Tính context vector  $C$  tương ứng với  $x_i$  theo công thức  $c_i = \alpha_{i1} v_1 + \alpha_{i2} v_2 + \dots + \alpha_{im} v_m = V \alpha_i$

### 2.5.2.2 Multi-head Attention

Multi-head Attention đơn giản là sự xếp chồng các lớp self attention. Ví dụ 1 lớp Multi-head Attention có  $L$  lớp self attention. Đầu ra của mỗi lớp self attention có kích thước  $d \times m$  thì đầu ra của multi-head attention sẽ là  $(Ld) \times m$ .

Ngoài thành phần chính là Multi-head Attention thì Transformer Encoder còn được tạo bởi các lớp khác như Add & Norm, Feed Forward, Add & Norm.

### 2.5.3 Classification Head

Phần này đơn giản là một khối MLP (Multilayer perceptron) nhận đầu vào là context vector  $c$  trả về từ Transformer Encoder và đưa ra kết quả cuối cùng là xác suất tương ứng với các class.

## 2.6 Phương pháp đánh giá

### 2.6.1 Đánh giá độ chính xác

Trong bài toán phát hiện và phân lớp đối tượng đặc biệt là bài toán phân loại đa lớp áp dụng nhiều mô hình mạng nơ-ron khác nhau, việc lựa chọn một phương pháp phù hợp để đánh giá và so sánh là rất cần thiết. Một trong các phương pháp được sử dụng phổ biến là xác định độ đo Loss và Accuracy.

Accuracy (độ chính xác) chỉ đơn giản đánh giá mô hình thường xuyên dự đoán đúng đến mức nào. Độ chính xác là tỉ lệ giữa số điểm dữ liệu được dự đoán đúng và tổng số điểm dữ liệu.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FB+FN} = \frac{TP+TN}{\text{Số lượng mẫu}} \quad (1)$$

Trong đó:

- TP (True Positive): dương tính thật, mô hình dự đoán là “có” và thực tế là “có”. FP (False Positive): dương tính giả, mô hình dự đoán là “có” nhưng thực tế là “không”. FN (False Negative): âm tính giả, mô hình dự đoán là “không” nhưng thực tế là “có”. TN (True Negative): âm tính thật, mô hình dự đoán là “không” và thực tế là “không”.

Giả sử độ độ chính xác = 90% có nghĩa là trong số 100 mẫu thì có 90 mẫu được phân loại chính xác. Tuy nhiên đối với tập dữ liệu kiểm thử không cân bằng (nghĩa là số positive lớn hơn rất nhiều so với negative) thì đánh giá có thể gây hiểu nhầm.

### 2.6.2 Confusion matrix

Trong những bài toán phân loại lớp, sau khi đã xử lý dữ liệu và đưa vào mô hình học máy, đầu ra của mô hình sẽ là một vector xác suất tương ứng của từng lớp. Ta có thể đánh giá mức độ chính xác của mô hình thông qua chỉ số về accuracy (độ chính xác): là phần trăm các lớp đã phân loại đúng / tổng số dự đoán, thế nhưng với chỉ số các lớp đã phân loại sai thì sẽ ảnh hưởng rất lớn đến việc chẩn đoán. Giả sử với bài toán chẩn đoán ung thư với việc phân loại một người thành 2 lớp: bị ung thư và không bị ung thư. Nếu trong 100 người cần phân loại có 10 người bị ung thư mà mô hình dự đoán cả 100 người đều không bị ung thư thì độ chính xác của mô hình đã là 0.9 khá cao. Nhưng với 10 người bị ung thư đã bị bỏ qua thì rõ ràng mô hình này không đáng tin cậy. Cần phải có một phương pháp đánh giá tốt hơn là độ chính xác trong bài toán phân loại và một trong đó là sử dụng Confusion matrix.



		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	<b>TP</b>	<b>FP</b>
	Negative (0)	<b>FN</b>	<b>TN</b>

Hình 2.23 Confusion matrix (Minh Vũ, 2021)

Confusion matrix là một phương pháp đánh giá kết quả của những bài toán phân loại với việc xem xét cả những chỉ số về độ chính xác và độ bao quát của các dự đoán cho từng lớp. Một confusion matrix gồm 4 chỉ số sau đối với mỗi lớp phân loại.

Để đơn giản hóa, ta sẽ sử dụng lại bài toán về chẩn đoán tế bào ung thư để giải thích 4 chỉ số này. Trong bài toán chuẩn đoán tế bào ung thư ta có 2 lớp: lớp tế bào ung thư ác tính được chuẩn đoán Negative và lớp tế bào ung thư lành tính được chuẩn đoán là Positive:

- TP (True Positive): Số lượng dự đoán chính xác. Là khi mô hình dự đoán đúng một người bị ung thư lành tính.

- TN (True Negative): Số lượng dự đoán chính xác một cách gián tiếp. Là khi mô hình dự đoán đúng một người bị ung thư ác tính, tức là việc không chọn trường hợp bị ung thư lành tính là chính xác.

- FP (False Positive): Số lượng các dự đoán sai lệch. Là khi mô hình dự đoán một người bị ung thư lành tính nhưng người đó bị ung thư ác tính.

- FN (False Negative): Số lượng các dự đoán sai lệch một cách gián tiếp. Là khi mô hình dự đoán một người không bị ung thư lành tính nhưng người đó bị ung thư ác tính, tức là việc không chọn trường hợp bị ung thư lành tính là sai. Từ 4 chỉ số này, ta có 2 con số để đánh giá mức độ tin cậy của một mô hình:

- Precision: Trong tất cả các dự đoán Positive được đưa ra, bao nhiêu dự đoán là chính xác. Chỉ số này được tính theo công thức:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

- Recall: Trong tất cả các trường hợp Positive, bao nhiêu trường hợp đã được dự đoán chính xác. Chỉ số này được tính theo công thức:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

- F-score: Kết hợp Precision và Recall thành một chỉ số duy nhất, được tính theo công thức:

$$\text{F-score} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad (4)$$

- Một mô hình có chỉ số F-score cao chỉ khi cả 2 chỉ số Precision và Recall đều cao. Một trong 2 chỉ số này thấp đều sẽ kéo điểm F-score xuống. Trường hợp xấu nhất khi 1 trong hai chỉ số Precision và Recall bằng 0 sẽ kéo điểm F-score về 0. Trường hợp tốt nhất khi cả điểm chỉ số đều đạt giá trị bằng 1, khi đó điểm F-score sẽ là 1.

### 2.6.3 Dice coefficient

Dice coefficient (hệ số Dice) là một phép đo sự tương đồng (similarity measure) thường được sử dụng trong lĩnh vực xử lý hình ảnh và phân đoạn hình ảnh (image segmentation). Nó được sử dụng để đánh giá mức độ khớp (overlap) giữa hai khu vực phân đoạn hoặc đối tượng trên hình ảnh.

Hệ số Dice được tính bằng cách so sánh kích thước khu vực chung giữa hai đối tượng (hoặc vùng phân đoạn) với tổng kích thước của hai đối tượng. Kết quả của Dice coefficient dao động từ 0 đến 1, với 0 cho thấy không có sự khớp giữa hai đối tượng và 1 cho thấy sự khớp hoàn toàn.

Công thức tính Dice coefficient cho hai vùng (A và B) được biểu diễn như sau:

$$\text{Dice} = \frac{2 | A \cap B |}{| A | + | B |} \quad (5)$$

Trong đó:

- $| A \cap B |$  đại diện cho kích thước (diện tích) khu vực chung giữa A và B.
- $| A |$  đại diện cho kích thước (diện tích) của vùng A.
- $| B |$  đại diện cho kích thước (diện tích) của vùng B.

Dice coefficient là một phép đo thường được sử dụng để đánh giá chất lượng phân đoạn hình ảnh trong các bài toán như phân đoạn tế bào, phân đoạn đối tượng, hay phân đoạn cấu trúc ảnh. Càng gần 1 hệ số Dice, thì sự tương đồng giữa hai vùng phân đoạn càng cao.

### CHƯƠNG 3: PHƯƠNG PHÁP ĐỀ XUẤT

#### 3.1 Tập dữ liệu

##### 3.1.1 Tập dữ liệu 1

Dữ liệu hình ảnh được cung cấp bởi “Laboratório Visão Robótica e Imagem” từ Spanhol, F., Oliveira, L. S., Petitjean, C. and Heutte, L [35]. Tập dữ liệu bao gồm 10.000 hình ảnh hiển vi của mô khối u vú. Trong đó, có 5.000 hình ảnh là khối u lành tính, 5.000 hình ảnh là khối u ác tính. Trong bài nghiên cứu này, em sử dụng 8.000 hình ảnh để huấn luyện và 2.000 hình ảnh để thẩm định.

##### 3.1.2 Tập dữ liệu 2

Dữ liệu hình ảnh được cung cấp bởi Walid Al-Dhabyani a, Mohammed Gomaa b, Hussien Khaled b, Aly Fahmy [37]. Tập dữ liệu bao gồm 780 hình ảnh CT của mô khối u vú. Trong đó, có 487 hình ảnh ung thư lành tính, 210 hình ảnh ung thư ác tính và 133 hình ảnh bình thường. Trong bài nghiên cứu này, em sử dụng 624 hình ảnh để huấn luyện và 156 hình ảnh để thẩm định.

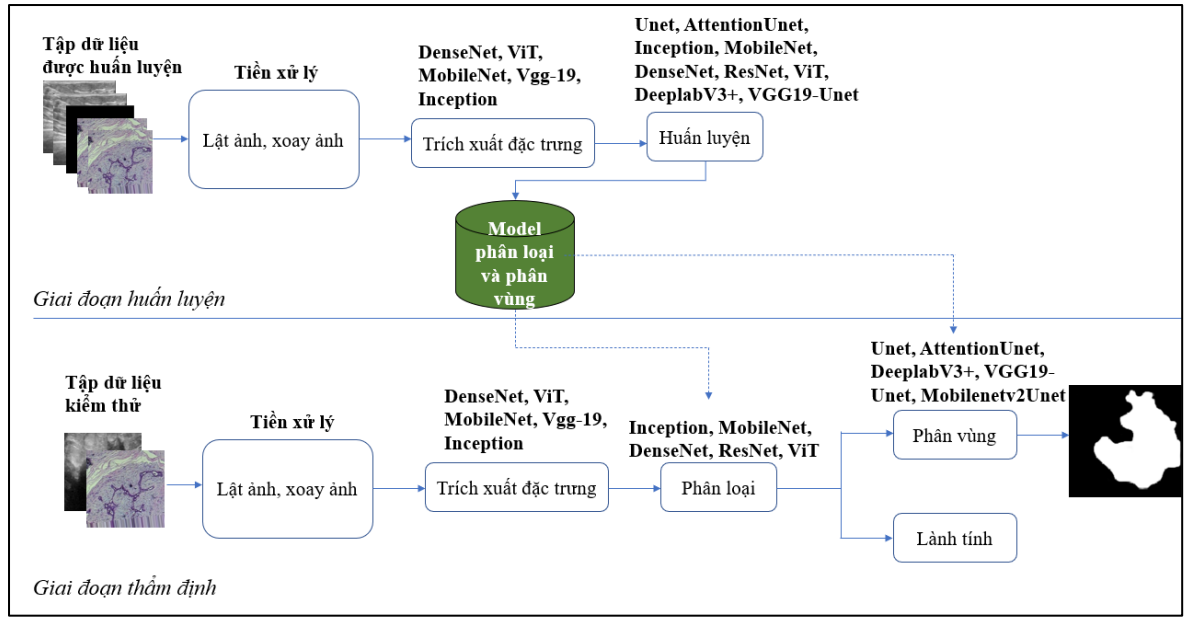
#### 3.2 Môi trường thực nghiệm

Môi trường máy tính đơn trên nền tảng cung cấp sẵn có của Kaggle với GPU NVIDIA-SMI 470.161.03 đi kèm với 13GB RAM và GPU Memory 16GB cho phép huấn luyện với batch size lớn hơn, ngoài ra dung lượng ổ cứng lớn hơn 73 GB cho phép khả năng lưu trữ tốt.

*Bảng 3.1 Bảng thông số môi trường huấn luyện*

Tên môi trường	Cấu hình	Thư viện
Kaggle GPU P100	VRAM: 15.9 GB GPU: NVIDIA-SMI 470.161 RAM: 13GB Disk: 73.1 GB	Tensorflow 2.4.0 Keras

### 3.3 Phương pháp đề xuất



Hình 3.1 Kiến trúc đề xuất phân loại, phân vùng ung thư vú

- Giai đoạn huấn luyện: Đầu tiên, cơ sở dữ liệu sẽ trải qua bước tiền xử lý để tăng cường dữ liệu. Sau đó, tiến hành huấn luyện với các mô hình mạng đề xuất.

- Giai đoạn thẩm định: Tiến hành tiền xử lý ảnh để tăng chất lượng ảnh. Sau đó, đưa vào mô hình phân loại, nếu kết quả là ác tính thì tiếp tục tiến hành phân vùng để xác định khối u, ngược lại thì kết thúc.

### 3.4 Kịch bản huấn luyện

Bảng 3.2 Kịch bản huấn luyện

Kịch bản	Mạng phân vùng	Mạng rút trích đặc trưng và huấn luyện	Số epochs	Learning rate	Loss function	Batch size
1	—	DenseNet201	200	0.001	Binary Crossentropy	64
2	—	ResNet50	200	0.001	Binary Crossentropy	64
3	—	MobileNetV3Small	200	0.001	Binary Crossentropy	64
4	—	InceptionV3	200	0.001	Binary Crossentropy	64

5	—	ViT	200	0.001	Binary Crossentropy	64
6	—	ViT-InceptionV3	200	0.001	Binary Crossentropy	64
7	Unet	—	200	0.001	Mean Square Error	64
8	AttentionUnet	—	200	0.001	Mean Square Error	64
9	DeeplabV3+	ResNet50	200	0.001	Mean Square Error	64
10	Unet	InceptionResNetV2	200	0.001	Mean Square Error	64
11	Unet	MobileNetV2	200	0.001	Mean Square Error	64
12	Unet	VGG19	200	0.001	Mean Square Error	64

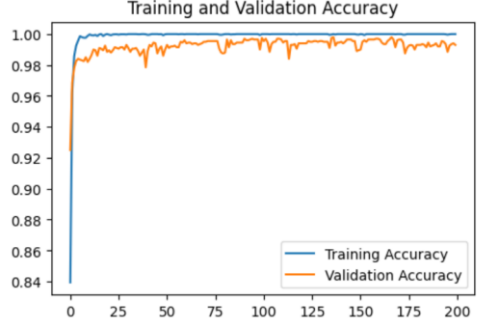
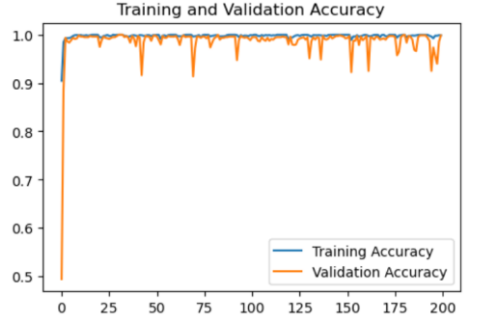
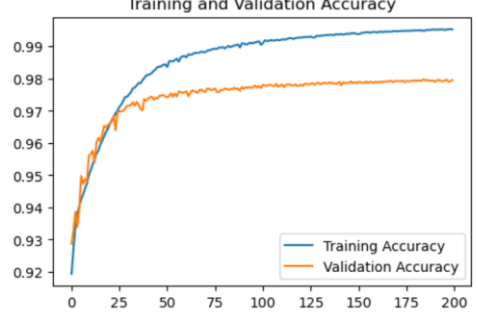
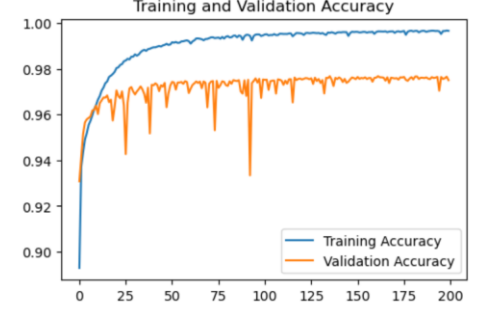
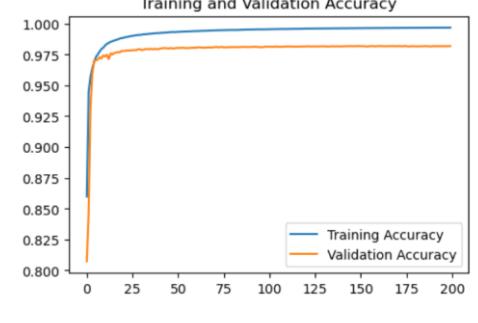
## CHƯƠNG 4: KẾT QUẢ ĐẠT ĐƯỢC

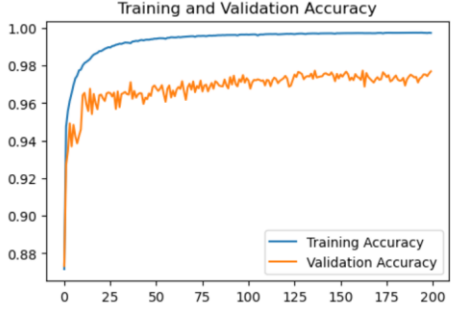

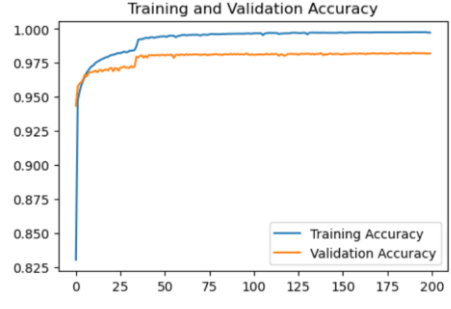
### 4.1 Kết quả huấn luyện

#### 4.1.1 Độ chính xác

Bảng 4.1 Kết quả độ chính xác của các mô hình

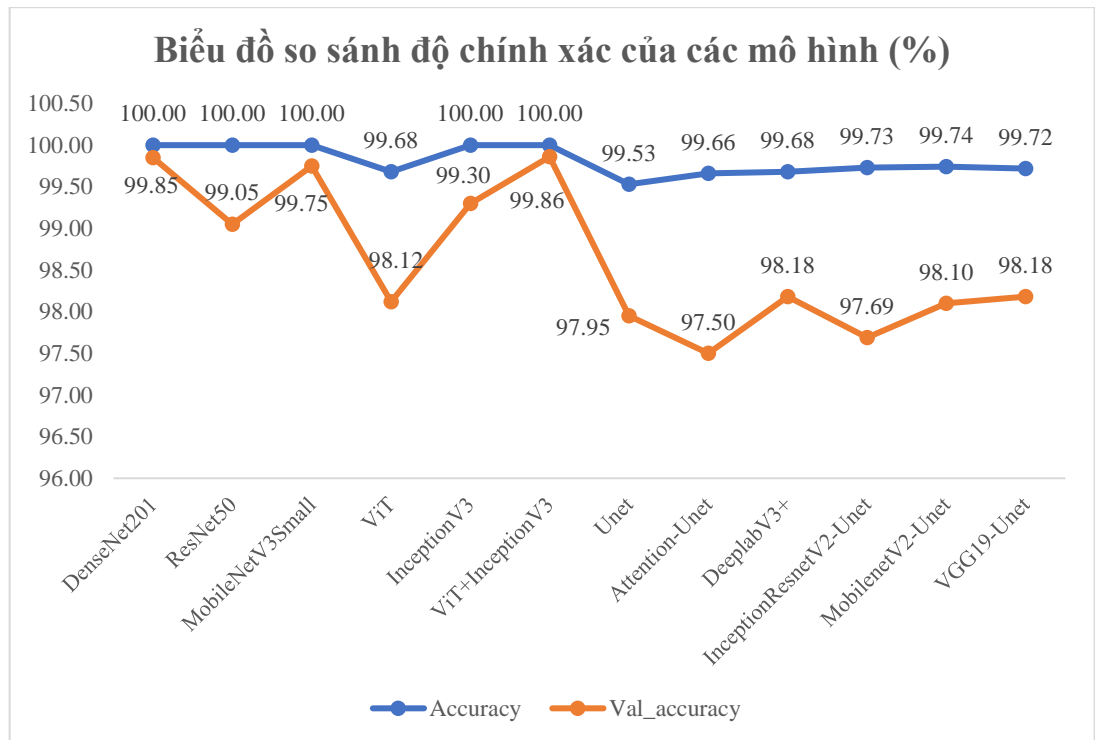
Kịch bản	Mô hình mạng	Biểu đồ độ chính xác
1	DenseNet201	<p>Training and Validation Accuracy for DenseNet201. The x-axis represents epochs from 0 to 200, and the y-axis represents accuracy from 0.93 to 1.00. The training accuracy (blue line) rises sharply from approximately 0.93 to 1.00 within the first 10 epochs and remains stable. The validation accuracy (orange line) also rises sharply, reaching approximately 0.99 by epoch 10, and then fluctuates slightly around 0.995 for the remainder of the training process.</p>
2	ResNet50	<p>Training and Validation Accuracy for ResNet50. The x-axis represents epochs from 0 to 200, and the y-axis represents accuracy from 0.84 to 1.00. The training accuracy (blue line) rises sharply from approximately 0.85 to 1.00 within the first 10 epochs and remains stable. The validation accuracy (orange line) rises sharply, reaching approximately 0.98 by epoch 10, and then fluctuates slightly around 0.99 for the remainder of the training process.</p>
3	MobileNetV3Small	<p>Training and Validation Accuracy for MobileNetV3Small. The x-axis represents epochs from 0 to 200, and the y-axis represents accuracy from 0.75 to 1.00. The training accuracy (blue line) rises sharply from approximately 0.75 to 1.00 within the first 10 epochs and remains stable. The validation accuracy (orange line) rises sharply, reaching approximately 0.99 by epoch 10, and then fluctuates slightly around 0.995 for the remainder of the training process.</p>
4	ViT	<p>Training and Validation Accuracy for ViT. The x-axis represents epochs from 0 to 200, and the y-axis represents accuracy from 0.75 to 1.00. The training accuracy (blue line) rises sharply from approximately 0.75 to 1.00 within the first 10 epochs and remains stable. The validation accuracy (orange line) rises sharply, reaching approximately 0.98 by epoch 10, and then fluctuates slightly around 0.99 for the remainder of the training process.</p>

5	InceptionV3	 <p>Training and Validation Accuracy for InceptionV3. The x-axis represents epochs from 0 to 200, and the y-axis represents accuracy from 0.84 to 1.00. The training accuracy (blue line) rises sharply from approximately 0.84 to 1.00 within the first 10 epochs and remains stable. The validation accuracy (orange line) also rises sharply, reaching approximately 0.98 by epoch 10, and then fluctuates slightly between 0.97 and 0.99 for the remainder of the training process.</p>
6	ViT-InceptionV3	 <p>Training and Validation Accuracy for ViT-InceptionV3. The x-axis represents epochs from 0 to 200, and the y-axis represents accuracy from 0.5 to 1.0. The training accuracy (blue line) rises sharply from approximately 0.5 to 1.0 within the first 10 epochs and remains stable. The validation accuracy (orange line) also rises sharply, reaching approximately 0.98 by epoch 10, and then fluctuates significantly between 0.92 and 1.00 for the remainder of the training process.</p>
7	Unet	 <p>Training and Validation Accuracy for Unet. The x-axis represents epochs from 0 to 200, and the y-axis represents accuracy from 0.92 to 0.99. The training accuracy (blue line) rises steadily from approximately 0.92 to 0.995 over the 200 epochs. The validation accuracy (orange line) also rises steadily, reaching approximately 0.98 by epoch 100 and remaining stable thereafter.</p>
8	AttentionUnet	 <p>Training and Validation Accuracy for AttentionUnet. The x-axis represents epochs from 0 to 200, and the y-axis represents accuracy from 0.90 to 1.00. The training accuracy (blue line) rises sharply from approximately 0.90 to 0.98 within the first 10 epochs and then continues to rise slowly towards 1.00. The validation accuracy (orange line) also rises sharply, reaching approximately 0.96 by epoch 10, and then fluctuates between 0.94 and 0.98 for the remainder of the training process.</p>
9	DeeplabV3+	 <p>Training and Validation Accuracy for DeeplabV3+. The x-axis represents epochs from 0 to 200, and the y-axis represents accuracy from 0.800 to 1.000. The training accuracy (blue line) rises sharply from approximately 0.800 to 0.98 within the first 10 epochs and then continues to rise slowly towards 1.00. The validation accuracy (orange line) also rises sharply, reaching approximately 0.97 by epoch 10, and then remains stable around 0.975 for the remainder of the training process.</p>

10	InceptionResNetV2Unet	
11	MobileNetV2Unet	
12	VGG19-Unet	

Sau khi huấn luyện qua 200 epochs thì ta thu được kết quả sau: Độ chính xác của các mô hình DenseNet201, ResNet50, MobileNetV3Small, InceptionV3, ViT, ViT-InceptionV3, Unet, AttentionUnet, DeeplabV3+, InceptionResNetV2Unet, MobileNetV2Unet, VGG19-Unet ở tập thử lần lượt đạt 99.85%, 99.05%, 99.75%, 98.12%, 99.30%, 99.86%, , 97.95%, 97.50%, 98.18%, 97.69%, 98.10%, 98.18%. Đối với quá trình phân loại, mô hình ViT-Inception cho kết quả huấn luyện cao nhất với 98.86%, mô hình ViT cho kết quả huấn luyện thấp nhất với 98.12%. Đối với quá trình phân vùng, mô hình DeeplabV3+ và VGG19-Unet cùng cho kết quả huấn luyện cao nhất với 98.18%, mô hình AttentionUnet cho kết quả huấn luyện thấp nhất với 97.50%.



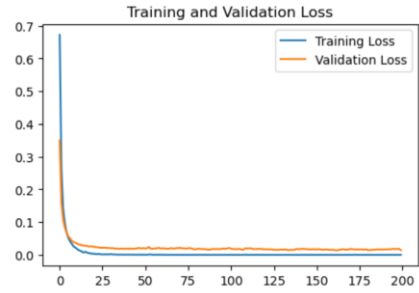
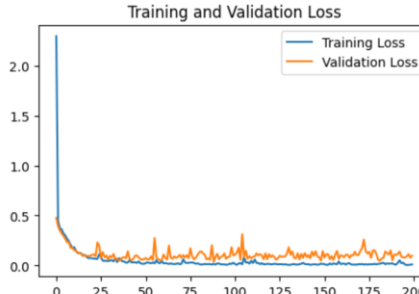
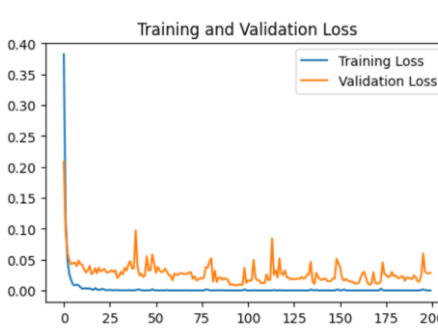
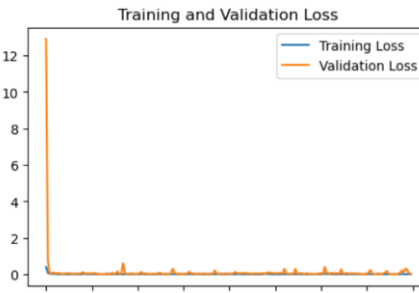



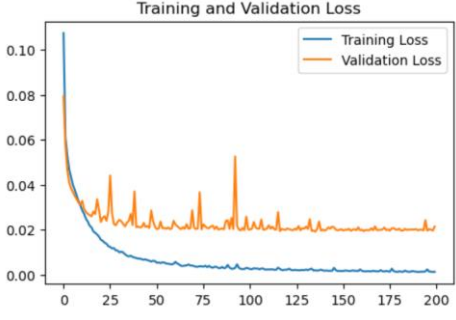
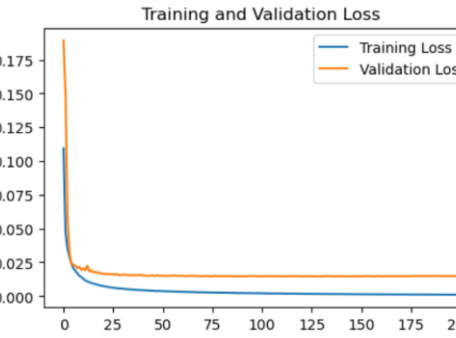
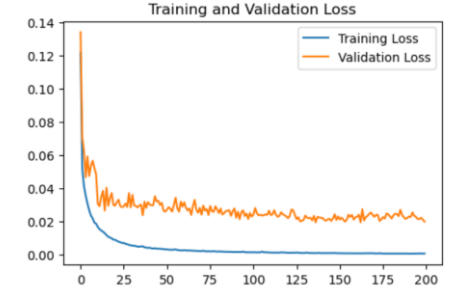
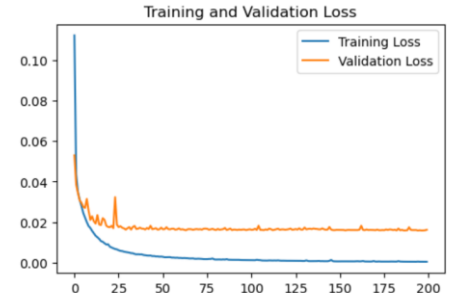
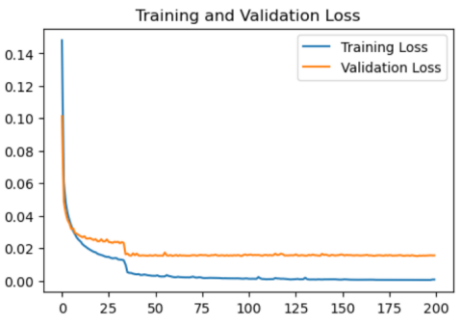
Hình 4.1 Biểu đồ so sánh độ chính xác của các mô hình

#### 4.1.2 Độ mất mát

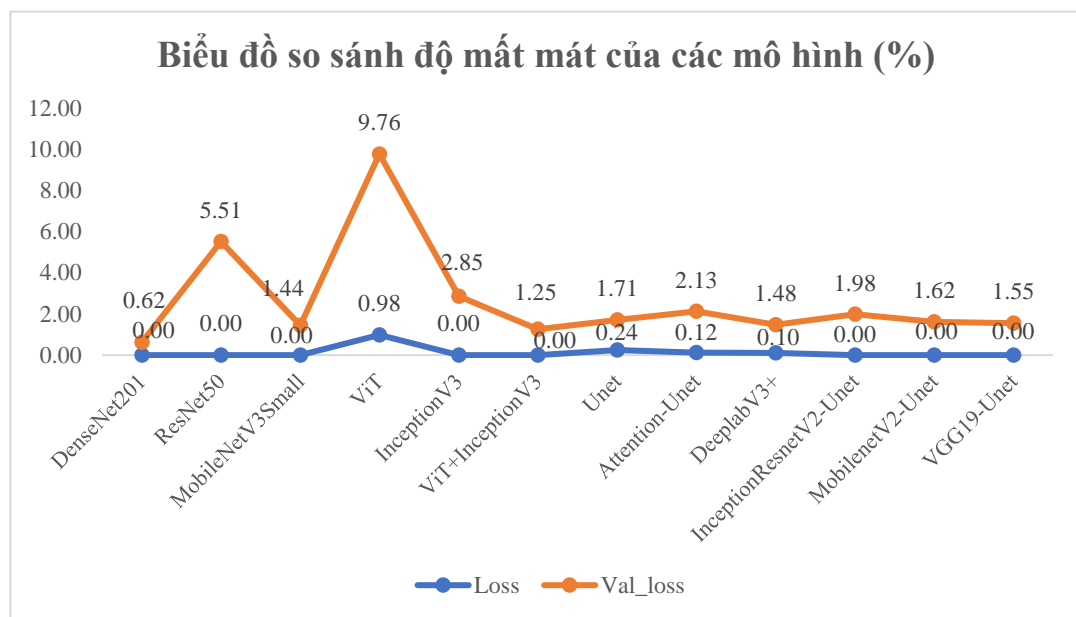
Bảng 4.2 Kết quả độ mất mát của các mô hình

Kịch bản	Mô hình mạng	Biểu đồ độ mất mát
1	DenseNet201	
2	ResNet50	

3	MobileNetV3Small	 <p>Training and Validation Loss for MobileNetV3Small. The Training Loss (blue line) starts at approximately 0.65 and drops sharply to near 0.0 by epoch 25. The Validation Loss (orange line) starts at approximately 0.35 and drops to near 0.0 by epoch 25, remaining stable thereafter.</p>
4	ViT	 <p>Training and Validation Loss for ViT. The Training Loss (blue line) starts at approximately 2.2 and drops sharply to near 0.0 by epoch 25. The Validation Loss (orange line) starts at approximately 0.5 and drops to near 0.0 by epoch 25, remaining stable thereafter.</p>
5	InceptionV3	 <p>Training and Validation Loss for InceptionV3. The Training Loss (blue line) starts at approximately 0.35 and drops sharply to near 0.0 by epoch 25. The Validation Loss (orange line) starts at approximately 0.20 and drops to near 0.0 by epoch 25, remaining stable thereafter.</p>
6	ViT-InceptionV3	 <p>Training and Validation Loss for ViT-InceptionV3. The Training Loss (blue line) starts at approximately 12.0 and drops sharply to near 0.0 by epoch 25. The Validation Loss (orange line) starts at approximately 12.0 and drops sharply to near 0.0 by epoch 25, remaining stable thereafter.</p>
7	Unet	 <p>Training and Validation Loss for Unet. The Training Loss (blue line) starts at approximately 0.07 and drops sharply to near 0.0 by epoch 25. The Validation Loss (orange line) starts at approximately 0.06 and drops to near 0.0 by epoch 25, remaining stable thereafter.</p>

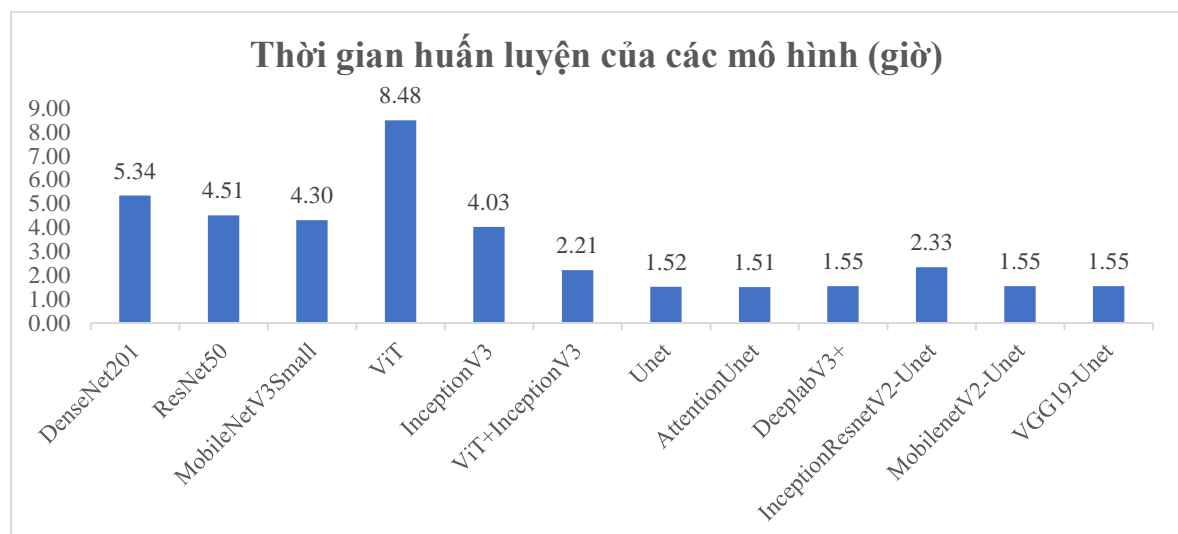
8	AttentionUnet	 <p>Training and Validation Loss for AttentionUnet. The x-axis represents epochs from 0 to 200, and the y-axis represents loss from 0.00 to 0.10. The Training Loss (blue line) decreases rapidly from approximately 0.10 to near 0.00 by epoch 50. The Validation Loss (orange line) starts at approximately 0.08, drops to about 0.02 by epoch 25, and then fluctuates between 0.01 and 0.03 for the remainder of the training process.</p>
9	DeeplabV3+	 <p>Training and Validation Loss for DeeplabV3+. The x-axis represents epochs from 0 to 200, and the y-axis represents loss from 0.000 to 0.175. The Training Loss (blue line) decreases rapidly from approximately 0.11 to near 0.00 by epoch 50. The Validation Loss (orange line) starts at approximately 0.18, drops to about 0.02 by epoch 25, and then remains relatively stable around 0.02 for the remainder of the training process.</p>
10	InceptionResNetV2Unet	 <p>Training and Validation Loss for InceptionResNetV2Unet. The x-axis represents epochs from 0 to 200, and the y-axis represents loss from 0.00 to 0.14. The Training Loss (blue line) decreases rapidly from approximately 0.13 to near 0.00 by epoch 50. The Validation Loss (orange line) starts at approximately 0.13, drops to about 0.03 by epoch 25, and then fluctuates between 0.02 and 0.03 for the remainder of the training process.</p>
11	MobileNetV2Unet	 <p>Training and Validation Loss for MobileNetV2Unet. The x-axis represents epochs from 0 to 200, and the y-axis represents loss from 0.00 to 0.10. The Training Loss (blue line) decreases rapidly from approximately 0.10 to near 0.00 by epoch 50. The Validation Loss (orange line) starts at approximately 0.05, drops to about 0.02 by epoch 25, and then remains relatively stable around 0.02 for the remainder of the training process.</p>
12	VGG19-Unet	 <p>Training and Validation Loss for VGG19-Unet. The x-axis represents epochs from 0 to 200, and the y-axis represents loss from 0.00 to 0.14. The Training Loss (blue line) decreases rapidly from approximately 0.14 to near 0.00 by epoch 50. The Validation Loss (orange line) starts at approximately 0.10, drops to about 0.02 by epoch 25, and then remains relatively stable around 0.02 for the remainder of the training process.</p>

Sau khi huấn luyện qua 200 epochs thì ta thu được kết quả sau: Độ mất mát của các mô hình DenseNet201, ResNet50, MobileNetV3Small, InceptionV3, ViT, ViT-InceptionV3, Unet, AttentionUnet, DeeplabV3+, InceptionResNetV2Unet, MobileNetV2Unet, VGG19-Unet ở tập thử định lần lượt đạt 0.62%, 5.51%, 1.44%, 9.76%, 2.85%, 1.25%, 1.71%, 2.13%, 1.48%, 1.62%, 1.55%. Đối với quá trình phân loại, mô hình DenseNet201 cho kết quả huấn luyện thấp nhất với 0.62%, mô hình ViT cho kết quả huấn luyện cao nhất với 9.76%. Đối với quá trình phân vùng, mô hình DeeplabV3+ cho kết quả huấn luyện thấp nhất với 1.48%, mô hình AttentionUnet cho kết quả huấn luyện cao nhất với 2.13%.



Hình 4.2 Biểu đồ so sánh độ mất mát của các mô hình

#### 4.1.3 Thời gian huấn luyện



Hình 4.3 Biểu đồ so sánh thời gian huấn luyện của các mô hình

Sau khi huấn luyện qua 200 epochs thì ta thu được kết quả sau: Thời gian huấn luyện của các mô hình DenseNet201, ResNet50, MobileNetV3Small, InceptionV3, ViT, ViT-InceptionV3, Unet, AttentionUnet, DeeplabV3+, InceptionResNetV2Unet, MobileNetV2Unet, VGG19-Unet lần lượt là 5.34 giờ, 4.51 giờ, 4.30 giờ, 8.48 giờ, 4.03 giờ, 2.21 giờ, 1.52 giờ, 1.51 giờ, 1.55 giờ, 2.33 giờ, 1.15 giờ, 1.15 giờ. Đối với quá trình phân loại, mô hình ViT-InceptionV3 có thời gian huấn luyện thấp nhất với 2.21 giờ, mô hình ViT có thời gian huấn luyện cao nhất với 8.48 giờ. Đối với quá trình phân vùng, mô hình AttentionUnet có thời gian huấn luyện thấp nhất với 1.51 giờ, mô hình InceptionResNetV2Unet có thời gian huấn luyện cao nhất với 2.33 giờ.

#### 4.1.4 Confusion matrix

Số lượng ảnh trong tập thử định là 2000 ảnh, trong đó có 1000 ảnh là lành tính và 1000 ảnh là ác tính. Số lượng dự đoán chính xác ảnh lành tính (TP) trung bình là 914 ảnh. Số lượng dự đoán sai lệch ảnh lành tính (FP) trung bình là 90 ảnh. Số lượng dự đoán chính xác ảnh ác tính (TN) trung bình là 912 ảnh. Số lượng dự đoán sai lệch ảnh ác tính (FN) trung bình là 85 ảnh. Độ chính xác (Accuracy) trung bình của quá trình thử định đạt 0.9128. Chỉ số dự đoán sai lệch ảnh lành tính (FP) là chính xác (Precision) trung bình đạt 0.91. Chỉ số các trường hợp được dự đoán chính xác (Recall) trung bình đạt 0.915. Chỉ số đáng tin cậy hiệu năng (F-score) trung bình đạt 0.910. Trong đó, mô hình MobileNetV3Small có lần lượt các chỉ số cao hơn so với các mô hình còn lại. Mô hình DenseNet201 có lần lượt các chỉ số thấp nhất so với các mô hình còn lại.

Bảng 4.3 Bảng tổng hợp số liệu Confusion matrix

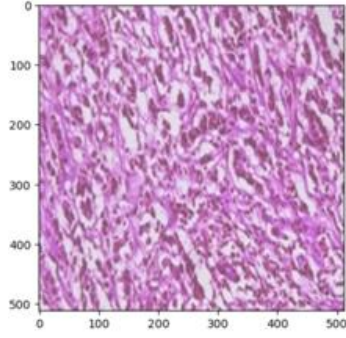
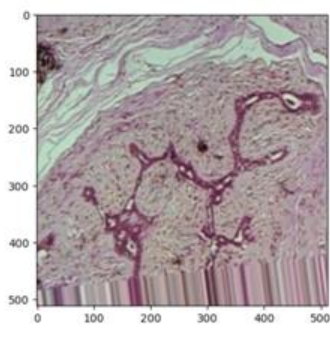
Tên mô hình	Confusion matrix							
	TP	FP	FN	TN	Accuracy	Precision	Recall	F-score
DenseNet201	512	488	485	515	0.5135	0.512	0.513	0.512
ResNet50	989	11	8	992	0.9905	0.989	0.9919	0.9814
MobileNetV3Small	<b>996</b>	<b>4</b>	<b>1</b>	<b>999</b>	<b>0.9975</b>	<b>0.996</b>	<b>0.9989</b>	<b>0.9974</b>
ViT	987	25	2	986	0.9865	0.975	0.9979	0.9863
Inception	994	6	8	992	0.993	0.994	0.992	0.9929
ViT-Inception	1006	6	2	986	0.996	0.994	0.998	0.9959

## 4.2 Kết quả kiểm thử

### 4.2.1 Phân loại khối u trên ảnh mô sinh ung thư vú

- Đối với ảnh mô tế bào ung thư lành tính kết quả sẽ hiển thị là: benign.
- Đối với ảnh mô tế bào ung thư ác tính kết quả sẽ hiển thị là: malignant.




*Bảng 4.4 Kết quả kiểm thử phân loại ung thư vú*

Ảnh kiểm thử Mô hình		
DenseNet201	malignant	benign
ResNet50	malignant	benign
MobileNetV3Small	malignant	benign
ViT	malignant	benign
InceptionV3	malignant	benign
ViT-InceptionV3	malignant	benign

Sau quá trình kiểm thử, ta thu được kết quả sau: Tất cả 6 mô hình đều cho kết quả kiểm thử đúng.

### 4.2.2 Phân vùng vùng bị tổn thương trên ảnh chụp CT

*Bảng 4.5 Kết quả kiểm thử phân vùng ung thư vú*

Mô hình	Mặt nạ dự đoán	Độ tương đồng so với mặt nạ thật
Unet		97.39%
Attention Unet		97.38%
DeeplabV3+		98.01%

InceptionResnetV2-Unet		99.16%
MobilenetV2-Unet		99.12%
VGG19-Unet		99.12%

Sau quá trình kiểm thử, ta thu được kết quả sau: Tất cả 6 mô hình đều cho kết quả kiểm thử khá tốt. Trong đó, mô hình InceptionResnetV2-Unet đưa ra mặt nạ dự đoán có độ tương đồng so với mặt nạ thật cao nhất với 99.16%. Mô hình AttentionUnet đưa ra mặt nạ dự đoán có độ tương đồng so với mặt nạ thật thấp nhất với 97.38%.

## CHƯƠNG 5: KẾT LUẬN

### 5.1 Kết quả đạt được

- Tiến hành tiền xử lý ảnh qua việc lật ảnh và xoay ảnh để có thể tạo ra bộ dữ liệu mới nhằm tăng tính chính xác cho mô hình hơn.
- Tiến hành cài đặt, sử dụng môi trường và các thư viện để tiến hành huấn luyện và kiểm thử các mô hình.
- Thực hiện các mạng rút trích đặc trưng cũng như phân vùng và phân lớp.
- Cung cấp các kết quả thực nghiệm, đánh giá các mô hình DenseNet201, Unet, ResNet50, InceptionV3, MobileNetV3Small, ViT, DeeplabV3+, độ chính xác, độ mất mát và thời gian huấn luyện mô hình.
- So sánh đánh giá các mô hình mạng trong quá trình thực hiện.

### 5.2 Hạn chế

- Chưa giải quyết được bài toán phân lớp và phân vùng trong một model duy nhất và một tập dữ liệu duy nhất.
- Chưa sử dụng nhiều mạng khác.

### 5.3 Hướng phát triển

- Dẫn hoàn thiện về mặt thu thập các tập dữ liệu mới hơn.
- Sử dụng thay thế nhiều mạng để xem kết quả, cải thiện thêm mô hình.
- Sử dụng huấn luyện và tính toán mô hình song song trong môi trường Apache Spark để cải thiện thời gian huấn luyện mô hình.



## TÀI LIỆU THAM KHẢO

- [1] Phan Anh Cang (2021). *Giáo trình Trí tuệ nhân tạo*. Trường Đại học Sư phạm Kỹ thuật Vĩnh Long.
- [2] Phan Anh Cang (2021). *Giáo trình Xử lý ảnh*. Trường Đại học Sư phạm Kỹ thuật Vĩnh Long.
- [3] Văn Bình Doãn (2022). *Tổng quan về Vision Transformer (ViT)*. VinBigData.
- [4] Phạm Đình Khánh (2020). *Làm chủ kỹ thuật học sâu*. NXB Thế Giới.
- [5] Nguyễn Thị Mai Lan (2020). *Nghiên cứu tỉ lệ mắc mới ung thư vú ở phụ nữ Hà nội giai đoạn 2014 - 2016*. Trường Đại học Y Hà Nội.
- [6] Thúy Hà (2018). *Ung thư vú có thể chữa khỏi nếu phát hiện sớm*. NXB Pháp luật Việt Nam.
- [7] Phạm Thọ Hoàn (2019). *Giáo trình Trí tuệ nhân tạo*. Trường Đại học Sư phạm Hà Nội.
- [8] Nguyễn Thị Hoa Mai (2022). *Sàng lọc, chẩn đoán sớm ung thư vú*. Bộ y tế.
- [9] Hà Minh Mẫn (2022). *Ung thư vú: dấu hiệu, nguyên nhân và cách phòng ngừa*. NXB Bệnh viện Tâm Anh.
- [10] Nguyễn Thanh (2022). *Hơn 21.000 ca mới ung thư vú tại Việt Nam*. NXB Tuổi trẻ. Số 121.
- [11] Lê Đình Roanh (2019). *Bệnh học ung thư vú*. Đại học Y Hà Nội, NXB Y học.
- [12] Aakash Kaushik (2020). *Understanding ResNet50 architecture*. OpenGenus.
- [13] Ademola Enitan Ilesanmi (2020). *A method for segmentation of tumors in breast ultrasound images using the variant enhanced deep learning*. Science Direct.
- [14] Alan Fuad Jahwar (2022). *Segmentation and Classification for Breast Cancer Ultrasound Images Using Deep Learning Techniques: A Review*. IEEE Explore.
- [15] Chris Solomon (2017). *Fundamentals of Digital Image Processing*. John Wiley & Sons, Ltd.
- [16] Daniel J Bell (2021). *Dice similarity coefficient*. Radiopaedia.

- [17] Deepika Koundal (2021). *Breast Cancer Calcifications: Identification Using a Novel Segmentation Approach*. Medical Data Analysis for Neurodegenerative Disorders Diagnosis using Computational Techniques.
- [18] Deepika Koundal (2022). *Effective Image Processing and Segmentation-Based Machine Learning Techniques for Diagnosis of Breast Cancer*. Medical Data Analysis for Neurodegenerative Disorders Diagnosis using Computational Techniques.
- [19] Detechtron2 (2021). Architecture of *DeeplabV3+*. Paper with code.
- [20] Khalid Salama (2018). *Image classification with Vision Transformer*. Science.
- [21] Guisheng Zhang (2022). *SHA-MTL: soft and hard attention multi-task learning for automated breast cancer ultrasound image segmentation and classification*. International Journal of Computer Assisted Radiology and Surgery
- [22] Haoran Lu (2022). *Half-UNet: A Simplified U-Net Architecture for Medical Image Segmentation*. Frontiersin.
- [23] Muhammad Sakib Khan Inan (2022). *Deep integrated pipeline of segmentation guided classification of breast cancer from ultrasound images*. Science Direct.
- [24] Nicolle Vigil (2022). *Dual-Intended Deep Learning Model for Breast Cancer Diagnosis in Ultrasound Imaging*. Cancers MPDI.
- [25] T. Nagalakshmi (2022). *Breast Cancer Semantic Segmentation for Accurate Breast Cancer Detection with an Ensemble Deep Neural Network*. Springer Link.
- [26] Paul Harrison (2021). *Breast Cancer Segmentation Methods: Current Status and Future Potentials*. BioMed Research International.
- [27] Peng Gu (2022). *Automated 3D ultrasound image segmentation to aid breast cancer image interpretation*. Scientist Direct.
- [28] Pin Wang (2020). *Automatic cell nuclei segmentation and classification of breast cancer histopathology images*. Signal Processing.
- [29] Rania Almajalid (2022). *Development of a Deep-Learning-Based Method for Breast Ultrasound Image Segmentation*. IEEE Explore

- [30] Ruey-Feng Chang (2022). *Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors*. Breast Cancer Research and Treatment.
- [31] Ruey-Feng Chang (2022). *Is the Dice coefficient the same as accuracy*. Stackexchange.
- [32] Ruhuls (2022). *Understanding Jaccard's Index and Dice Coefficient in Object Detection and Image Segmentation*. Science Direct.
- [33] Shashwati Mishra (2018). *Medical image retrieval using self-organising map on texture features*. Future Computing and Informatics Journal.
- [34] Sik-Ho Tsang (2018). *Review: DenseNet — Dense Convolutional Network (Image Classification)*. Toward Data Science.
- [35] Spanhol, F., Oliveira, L. S., Petitjean, C. and Heutte, L (2022). *Breast Cancer Histopathological Database (BreakHis)*. Laboratório Visão Robótica e Imagem.
- [36] Uswatun Khasana (2022). *Segmentation of Breast Using Ultrasound Image for Detection Breast Cancer*. IEEE Explore.
- [37] Walid Al-Dhabyani a, Mohammed Gomaa b, Hussien Khaled b, Aly Fahmy (2022). *Dataset of breast ultrasound images*. Science Direct.
- [38] Xiaozheng (2022). *Breast Ultrasound Image Classification and Segmentation Using Convolutional Neural Networks*. Advances in Multimedia Information Processing
- [39] Yang Lie (2022). *Breast tumor segmentation in 3D automatic breast ultrasound using Mask scoring R-CNN*. American Association of Phycists in medicine.
- [40] Yuan Xu (2020). *Medical breast ultrasound image segmentation by machine learning*. Science Direct.