



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

Ứng dụng mô hình mạng U-NET
trong phân đoạn khối u não trên ảnh CT/MRI

CBHD: TS. Phan Anh Cang
Sinh viên: Lê Quang Hiếu
Mã số sinh viên: 19004059

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

Ứng dụng mô hình mạng U-NET
trong phân đoạn khối u não trên ảnh CT/MRI

CBHD:	TS. Phan Anh Cang
Sinh viên:	Lê Quang Hiếu
Mã số sinh viên:	19004059

Vĩnh Long - Năm 2023

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 từ nguồn khác. Trong toàn bộ nội dung của luận văn, những thông tin được trình bày đều là ý kiến cá nhân hoặc được tổng hợp từ nhiều nguồn tài liệu khác nhau. Tất cả các tài liệu tham khảo đều có nguồn gốc rõ ràng và được trích dẫn đúng quy định. Đây là công trình nghiên cứu của cá nhân tôi và đã được hướng dẫn thực hiện bởi Tiến sĩ Phan Anh Cang. Các nội dung nghiên cứu trong đề tài "**Ứng dụng mô hình mạng U-NET trong phân đoạn khối u não trên ảnh CT/MRI**" của tôi là trung thực và chưa được công bố trước đây dưới bất kỳ hình thức nào, dựa trên các tài liệu và đề tài trước đây. Tôi hoàn toàn chịu trách nhiệm và sẵn sàng chấp nhận mọi hình thức kỷ luật theo quy định nếu lời cam đoan của tôi có bất kỳ sai sót nào.

Vĩnh Long, ngày tháng năm 2023

Người thực hiện đề tài

Lê Quang 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, những 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

Lê Quang Hiếu

MỤC LỤC

LỜI CAM ĐOAN	i
LỜI CẢM ƠN	ii
DANH SÁCH CÁC TỪ VIẾT TẮT	v
DANH MỤC HÌNH ẢNH	vi
DANH SÁCH BẢNG	vii
CHƯƠNG 1: CƠ SỞ KHOA HỌC CỦA ĐỀ TÀI	1
1.1. Giới Thiệu	1
1.2. Lý do chọn đề tài	2
1.3. Cơ sở lý luận	3
1.4. Cơ sở thực tiễn	4
1.5 Những nghiên cứu liên quan	4
1.6 Vấn đề đặt ra	10
1.7 Hướng giải quyết của đề tài	11
CHƯƠNG 2: CƠ SỞ LÝ THUYẾT	12
2.1 Khối u não	12
2.1.1. <i>Khái quát</i>	12
2.1.2. <i>Dấu hiệu và triệu chứng</i>	14
2.2 Đặc điểm nhận dạng khối u não dựa trên ảnh MRI/CT	15
2.3 Tổng quan về trí tuệ nhân tạo	16
2.3.1. <i>Khái niệm</i>	16
2.3.2. <i>Một số công nghệ Trí tuệ nhân tạo phổ biến</i>	17
2.4. Kỹ thuật máy học (machine learning)	21
2.4.1. <i>Khái niệm Machine Learning</i>	21
2.4.2. <i>Machine learning Workflow</i>	21
2.4.3. <i>Phân loại Machine learning</i>	22
2.4.4. <i>Một số khái niệm cơ bản</i>	23
2.4.5. <i>Ứng dụng của Machine learning</i>	23
2.5. Deep learning	24
2.5.1. <i>Khái niệm Deep Learning</i>	24
2.5.2. <i>Cách thức hoạt động của Deep Learning</i>	25
2.5.3. <i>Ứng dụng của Deep Learning</i>	26
2.5.4. <i>Vai trò của Deep Learning</i>	27
2.6. Một số kiến mạng học sâu phổ biến	29
2.6.1. <i>Kiến trúc mạng ResNet50</i>	29
2.6.2. <i>Kiến trúc mạng VGG-19</i>	31

2.6.3. Kiến trúc mạng DenseNet121	33
2.6.4. Kiến trúc mạng Unet	37
2.7. Phương pháp đánh giá	38
2.7.1. Đánh giá độ chính xác (Accuracy)	38
2.7.2. Chỉ số IoU	39
2.7.3. Chỉ số Dice coefficient	39
CHƯƠNG 3: PHƯƠNG PHÁP NGHIÊN CỨU	41
3.1. Tập dữ liệu	41
3.2 Môi trường	42
3.3. Phương pháp đề xuất	43
3.4. Kịch bản huấn luyện	45
CHƯƠNG 4: KẾT QUẢ	46
4.1. Kết quả tham số huấn luyện	46
4.2. Kết quả huấn luyện	48
4.2.1. Độ chính xác	48
4.2.2. Độ mất mát	51
4.2.3. Thời gian huấn luyện	53
4.3. KẾT QUẢ KIỂM THỬ	55
KẾT LUẬN	60
Kết luận	60
Hướng phát triển	60
TÀI LIỆU KHAM KHẢO	62
PHỤ LỤC	65

DANH SÁCH CÁC TỪ VIẾT TẮT

STT	Kí hiệu	Nội dung đầy đủ
1	AI	Artificial Intelligent
2	CPU	Central Processing Unit
3	GPU	Graphics Processing Unit
4	IOT	Internet Of Things
5	mAP	Mean Average Precision
6	DL	Deep Learning

DANH MỤC HÌNH ẢNH

Hình 2. 1 Đặc điểm nhận dạng khối u não dựa trên ảnh MRI/CT[31]	16
Hình 2. 2. Cộng trực tiếp đầu vào của khối với nhánh còn lại trong khối Identity Block[7]	30
Hình 2. 3. Kiến trúc tổng quan của Resnet[26]	30
Hình 2. 4. Kiến trúc mạng VGG-19[12]	31
Hình 2. 5. Kiến trúc mạng Densenet [9]	34
<i>Hình 2. 6 Kiến trúc DenseNet-121[24]</i>	35
Hình 2. 7. Kiến trúc mạng Unet[30].....	37
Hình 3. 1. Phân phối dữ liệu được nhóm theo chẩn đoán	41
Hình 3. 2. Chia dữ liệu train, Validation, test	42
Hình 3. 3. Mô hình đề xuất.....	44
Hình 4. 1. Biểu đồ so sánh kết quả các mô hình	46
Hình 4. 2. Biểu đồ thời gian huấn luyện	54

DANH SÁCH BẢNG

Bảng 3. 1 Thông số môi trường huấn luyện	43
Bảng 3. 2 bảng kịch bản huấn luyện	45
Bảng 4. 1 Kịch bản mô hình phân loại	46
Bảng 4. 2 bảng chứa các biểu đồ Độ chính xác	48
Bảng 4. 3 Bảng chứa các biểu đồ mất mát	51
Bảng 4. 4 Thời gian huấn luyện	53
Bảng 4. 5 Bảng chứa kết quả kiểm thử phân loại	57
Bảng 4. 6 Bảng chứa kết quả kiểm thử phân đoạn	59

CHƯƠNG 1: CƠ SỞ KHOA HỌC CỦA ĐỀ TÀI

1.1. Giới Thiệu

Theo ThS.BS.CKII Chu Tấn Sĩ - Trưởng khoa Ngoại thần kinh, Trung tâm Thần kinh, Bệnh viện Đa khoa Tâm Anh TP HCM, u não chiếm 2% trong tổng số các ca ung thư từ mọi nhóm tuổi. Trong số các trường hợp tử vong do ung thư ở nhóm trẻ em dưới 15 tuổi và nhóm từ 20-39 tuổi, bệnh u não là nguyên nhân gây tử vong cao thứ hai. Người ngoài 85 tuổi có tỷ lệ bị u não cao nhất. Tại Việt Nam, chỉ tính riêng bệnh viện Việt Đức, mỗi năm có khoảng 2.500 người bệnh được phát hiện mắc u não. Bệnh u não ác tính rất phổ biến ở Đức với hơn 8.000 người mắc các bệnh liên quan. Hơn 500 trường hợp mới được chẩn đoán mắc u não mỗi ngày trên toàn thế giới. Số bệnh nhân có khối u di căn não còn nhiều hơn thế. Đây là loại ung thư phổ biến thứ 2 ở trẻ em (sau ung thư máu) và chiếm 30% tổng số ca tử vong vì ung thư ở trẻ em. U não là căn bệnh không hiếm gặp trên thế giới. Ước tính mỗi năm có khoảng 19.000 bệnh nhân u não mới được phát hiện ở Mỹ (dân số 300 triệu dân). Hiện tại, có tới gần 700.000 người Mỹ đang sống chung với căn bệnh này, trong đó 130.000 người mắc u ác tính.

Một khối u não là sự phát triển của các tế bào trong não hoặc gần nó. Các khối u não có thể xảy ra trong mô não và gần mô não. Các vị trí gần đó bao gồm dây thần kinh, tuyến yên, tuyến tùng và màng bao phủ bề mặt não. Uống thuốc chống ung thư, phẫu thuật và điều trị bằng tia X quang hoặc tia gamma là những phương pháp thông thường để chữa trị khối u não. Tuy nhiên, trước khi thực hiện những phương pháp này, việc phân đoạn chính xác khối u trên ảnh CT/MRI là cần thiết. Điều này giúp xác định vị trí và kích thước của khối u, từ đó hỗ trợ việc lên kế hoạch điều trị và đánh giá hiệu quả của nó. Ứng dụng mô hình mạng U-Net trong phân đoạn khối u não trên ảnh CT/MRI có thể giúp tự động hóa quá trình phân đoạn khối u, giảm thiểu sai số và tăng độ chính xác của quá trình này.

Mạng U-Net là một trong những mô hình mạng nổi tiếng và được sử dụng phổ biến trong lĩnh vực xử lý ảnh y tế. Mô hình này được thiết kế đặc biệt để xử lý các nhiệm vụ phân đoạn hình ảnh, đặc biệt là trong trường hợp có ít dữ liệu huấn luyện. Nghiên cứu và ứng dụng mô hình mạng U-Net trong phân đoạn khối u não trên ảnh CT/MRI có thể mang lại kết quả tốt và ứng dụng thực tiễn trong lĩnh vực y tế.

Nghiên cứu về ứng dụng mô hình mạng U-Net trong phân đoạn khối u não trên ảnh CT/MRI có tính ứng dụng cao trong việc hỗ trợ chẩn đoán và điều trị bệnh nhân ung thư não. Việc tự động hóa phân đoạn khối u trên ảnh CT/MRI giúp tăng tốc độ và độ chính xác của quá trình phân tích hình ảnh, từ đó cải thiện chất lượng chẩn đoán và giúp các bác sĩ và chuyên gia y tế đưa ra quyết định điều trị chính xác hơn.

Đề tài này đưa ra một số thách thức trong việc áp dụng mô hình mạng U-Net vào phân đoạn khối u não trên ảnh CT/MRI, bao gồm xử lý dữ liệu y tế phức tạp, đánh giá hiệu suất của mô hình và tối ưu hóa quá trình huấn luyện. Nghiên cứu về đề tài này có thể đóng góp vào việc giải quyết những thách thức này và mở ra cơ hội phát triển các phương pháp phân đoạn khối u không chỉ trong lĩnh vực y tế mà còn ứng dụng rộng rãi trong lĩnh vực xử lý ảnh và trí tuệ nhân tạo.

1.2. Lý do chọn đề tài

Lý do chúng tôi lựa chọn đề tài "Ứng dụng mô hình mạng U-Net trong phân đoạn khối u não trên ảnh CT/MRI" là bởi những lợi ích và tiềm năng mà mô hình U-Net mang lại trong lĩnh vực phân đoạn hình ảnh y tế. Trước tiên, trí tuệ nhân tạo và học máy đã có những tiến bộ đáng kể trong thời gian gần đây, đặc biệt là trong việc xử lý ảnh và nhận dạng đối tượng. Công nghệ mạng neural convolutional (CNN) đã được chứng minh là một công nghệ quan trọng và hiệu quả trong xử lý ảnh y tế.

Phân đoạn khối u não trên ảnh CT/MRI đóng vai trò quan trọng trong quá trình chẩn đoán và điều trị bệnh nhân ung thư não. Việc xác định chính xác vị trí và kích thước của khối u giúp các chuyên gia y tế đưa ra quyết định điều trị chính xác, từ đó giảm thiểu tác động đến các vùng não bình thường và tăng hiệu quả của quá trình điều trị.

Mô hình U-Net là một mô hình mạng neural convolutional đặc biệt được thiết kế để giải quyết các nhiệm vụ phân đoạn hình ảnh. Với cấu trúc encoder-decoder, mô hình U-Net có khả năng trích xuất thông tin đặc trưng và tái tạo hình ảnh với độ phân giải cao. Đã có nhiều nghiên cứu chứng minh khả năng chính xác của mô hình U-Net trong việc phân đoạn khối u não trên ảnh CT/MRI và các tập dữ liệu y tế khác.

Áp dụng mô hình U-Net vào phân đoạn khối u não trên ảnh CT/MRI mang lại nhiều lợi ích thực tế. Việc tự động hóa quá trình phân đoạn giúp giảm thời gian và công sức của các chuyên gia y tế, đồng thời cải thiện độ chính xác và nhất quán của kết quả phân đoạn.

Điều này góp phần quan trọng vào việc nâng cao chất lượng chẩn đoán và điều trị ung thư não, cung cấp thông tin quan trọng cho quyết định điều trị và tối ưu hóa kết quả cho bệnh nhân.

1.3. Cơ sở lý luận

Cơ sở lý luận để ứng dụng mô hình mạng U-Net trong phân đoạn khối u não trên ảnh CT/MRI là sự phát triển đáng kể của Deep Learning và các thuật toán Neural Networks. Deep Learning là một thuật toán dựa trên các ý tưởng lấy cảm hứng từ cấu trúc và hoạt động của não bộ, cho phép xử lý thông tin ở nhiều tầng biểu đạt khác nhau và hiểu rõ hơn về nghĩa của dữ liệu. Hiện nay, Deep Learning đã được áp dụng thành công trong nhiều lĩnh vực khác nhau.

Thuật toán Deep Learning được xây dựng dựa trên Artificial Neural Networks với nhiều lớp ẩn, tương tự như cấu trúc mạng thần kinh nơ-ron trong hệ thống sinh học. Deep Neural Networks có khả năng mô phỏng các mối quan hệ phức tạp phi tuyến tính trong cả bài toán giám sát và không giám sát. Điều này cho phép nó xử lý các tập dữ liệu để dự đoán kết quả và tìm ra các mẫu mới trong dữ liệu một cách chính xác. Đặc điểm của Deep Learning là khả năng học sâu và rộng, đạt được độ chính xác cao và có thể tự động mô hình hóa mối quan hệ phức tạp trong dữ liệu mà không cần các bước trích xuất đặc trưng thủ công. Điều này rất phù hợp với bài toán phân đoạn khối u não trên ảnh CT/MRI, vì nó có thể xử lý trực tiếp dữ liệu ảnh và tìm hiểu các đặc trưng quan trọng mà không cần các bước trung gian phức tạp.

Trong lĩnh vực Deep Learning, có nhiều kiến trúc mạng được sử dụng, bao gồm mạng neural tích chập (Convolutional Neural Network - CNN), Deep Belief Network và U-Net. Mô hình U-Net là một kiến trúc mạng CNN đặc biệt được thiết kế cho nhiệm vụ phân đoạn hình ảnh. Nó kết hợp cấu trúc encoder-decoder để trích xuất thông tin đặc trưng và tái tạo lại hình ảnh với độ phân giải cao. Đã có nhiều nghiên cứu chứng minh khả năng chính xác và hiệu quả của mô hình U-Net trong việc phân đoạn khối u não trên ảnh CT/MRI.

Việc áp dụng mô hình U-Net trong phân đoạn khối u não trên ảnh CT/MRI mang lại nhiều lợi ích trong thực tế. Nó giúp tự động hóa quá trình phân đoạn, tiết kiệm thời gian và công sức của các chuyên gia y tế. Đồng thời, nó cải thiện độ chính xác và nhất quán của kết quả phân đoạn, giúp xác định chính xác vị trí và kích thước của khối u. Điều này giúp đưa

ra quyết định điều trị chính xác, giảm tác động đến các vùng não bình thường và tăng hiệu quả của quá trình điều trị.

1.4. Cơ sở thực tiễn

Chụp cộng hưởng từ (Magnetic Resonance Imaging - MRI) là một phương pháp thu hình ảnh của các cơ quan trong cơ thể sống bằng cách quan sát lượng nước bên trong của từng cấu trúc của các cơ quan. Nguyên lý của cộng hưởng từ dựa trên một hiện tượng vật lý là hiện tượng cộng hưởng từ hạt nhân.

Bên cạnh đó, có sự khác biệt so với chụp X quang là dùng năng lượng tia bức xạ từ tia X, chụp cộng hưởng từ dùng năng lượng vô tuyến điện. Chính vì thế, nguy cơ nhiễm xạ khi chụp cộng hưởng từ là hoàn toàn không có, trong khi cấu trúc hình ảnh được phản ánh rõ ràng hơn. Do đó, phương tiện chẩn đoán này được ưu tiên áp dụng trong khảo sát các bệnh lý của mô mềm với cấu trúc tinh vi hay đan xen phức tạp.

Nhìn chung, độ nhạy và độ đặc hiệu của MRI rất cao trong chẩn đoán u màng não. MRI đã được chứng minh là vượt trội trong việc phân định khối u và mối quan hệ của nó với các cấu trúc xung quanh. Chính vì thế, nếu có các triệu chứng bất thường gì trên hệ Thần kinh trung ương như đau đầu, nhìn mờ, yếu liệt..., cần đi khám đúng chuyên khoa và thực hiện các kỹ thuật hình ảnh thích hợp để được điều trị sớm, đề phòng các biến chứng có thể xảy ra.

1.5 Những nghiên cứu liên quan

Trong lĩnh vực ứng dụng mô hình mạng U-Net trong phân đoạn khối u não trên ảnh CT/MRI, một tác phẩm khác là bài báo "Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks" [6] của các tác giả Xiao Li, Hao Chen, Xiaomeng Li, Qi Dou và Pheng-Ann Heng. Trong nghiên cứu này, các tác giả đã đề xuất một phương pháp phân đoạn tự động khối u não bằng cách sử dụng mạng U-Net kết hợp với mạng nơ-ron tích chập không đồng nhất (anisotropic convolutional neural networks - AnisoCNN). Các tác giả đã xây dựng một kiến trúc mạng U-Net có cấu trúc nhiều lớp và mở rộng bằng cách kết hợp với mạng AnisoCNN. Mạng AnisoCNN được thiết kế để xử lý các ảnh không đồng nhất về độ phân giải theo các hướng khác nhau, nhằm giải quyết vấn đề về tỷ lệ phân giải không đồng đều trong dữ liệu ảnh CT/MRI của não. Sự kết hợp của mạng U-Net và mạng AnisoCNN cho phép mô hình tự động phân đoạn khối u não một cách

chính xác và hiệu quả. Các tác giả đã tiến hành thử nghiệm và đánh giá phương pháp đề xuất trên tập dữ liệu lớn gồm các ảnh CT/MRI của các bệnh nhân có khối u não. Kết quả thử nghiệm cho thấy mô hình đề xuất đạt được độ chính xác cao trong việc phân đoạn khối u, vượt qua các phương pháp tiền đề hiện có. Điều này cho thấy tiềm năng và hiệu quả của mô hình U-Net kết hợp với mạng AnisoCNN trong phân đoạn khối u não trên ảnh CT/MRI.

Bài báo "The DCT-CNN-ResNet50 architecture to classify brain tumors with super-resolution, convolutional neural network, and the ResNet50"[7]. Tác giả của bài báo là Anand Deshpande, Vania V. Estrela và Prashant Patavardhan. Bài báo tập trung vào việc phân loại khối u não bằng cách sử dụng các phương pháp siêu phân giải, mạng nơ-ron tích chập (CNN) và kiến trúc ResNet50. Việc chẩn đoán khối u não thường được thực hiện dựa trên hình ảnh từ máy chụp cộng hưởng từ (MRI). Các phương pháp phân tích mô được sử dụng để xác định các khối u này. Tuy nhiên, một số yếu tố như chất lượng của thiết bị MRI và độ phân giải hình ảnh thấp có thể làm giảm chất lượng của hình ảnh MRI. Ngoài ra, việc phát hiện khối u trong hình ảnh độ phân giải thấp là một thách thức. Một phương pháp siêu phân giải có thể giúp khắc phục vấn đề này. Bài báo đề xuất việc sử dụng trí tuệ nhân tạo (AI) dựa trên mạng nơ-ron tích chập (CNN) để phân loại khối u não bằng cách sử dụng các tập dữ liệu công khai. Đề xuất trong bài báo này là một kiến trúc kết hợp giữa phương pháp hợp nhất hình ảnh dựa trên Biến đổi Cosine Rời rạc (DCT), mạng nơ-ron tích chập và kiến trúc ResNet50, được gọi là DCT-CNN-ResNet50. Kiến trúc này có khả năng phân biệt các mô u và không có u bằng cách sử dụng các tập dữ liệu công khai. Kết quả thực nghiệm trên hình ảnh MRI cho thấy rằng kiến trúc siêu phân giải được đề xuất dựa trên DCT, CNN và ResNet50 có khả năng cải thiện độ chính xác phân loại, và đã đạt được tỷ lệ độ chính xác 98,14% khi sử dụng siêu phân giải và kiến trúc ResNet50.

Bài báo "A Customized VGG19 Network with Concatenation of Deep and Handcrafted Features for Brain Tumor Detection"[8]. Tác giả của bài báo là Venkatesan Rajinikanth, Alex Noel Joseph Raj, Krishnan Palani Thanaraj và Ganesh R. Naik. Bài báo tập trung vào việc phát hiện khối u não bằng cách sử dụng mạng VGG19 được tùy chỉnh và kết hợp các đặc trưng sâu và được tạo bằng tay. Khối u não là một trong những bất thường về não phát sinh do nhiều nguyên nhân khác nhau. Việc không nhận ra và không điều trị khối u não sẽ tăng tỷ lệ tử vong và tỷ lệ mắc bệnh. Đánh giá cấp độ lâm sàng của

khối u não thường được thực hiện bằng kỹ thuật sinh hình và sàng lọc não hỗ trợ bằng MRI là một trong những kỹ thuật phổ biến. Công việc đề xuất nhằm phát triển một kiến trúc học sâu (DLA) để hỗ trợ phát hiện tự động khối u não bằng cách sử dụng lát cắt MRI hai chiều. Công việc đề xuất các DLA sau đây để phát hiện khối u não: (i) triển khai các DLA được huấn luyện trước, chẳng hạn như AlexNet, VGG16, VGG19, ResNet50 và ResNet101 với bộ phân loại SoftMax dựa trên đặc trưng sâu; (ii) DLA được huấn luyện trước với phân loại dựa trên đặc trưng sâu bằng cách sử dụng cây quyết định (DT), k-nearest neighbor (KNN), SVM-linear và SVM-RBF; và (iii) một mạng VGG19 tùy chỉnh với kết hợp tuần tự các đặc trưng sâu và đặc trưng được tạo bằng tay để cải thiện độ chính xác phát hiện khối u não. Nghiên cứu thực nghiệm đã được thực hiện riêng biệt trên các lát cắt MRI của các modality Flair, T2 và T1C, và đã sử dụng phương pháp kiểm tra chéo mười lần để kiểm chứng hiệu suất của DLA được đề xuất. Kết quả của công việc này xác nhận rằng VGG19 với SVM-RBF đã giúp đạt được độ chính xác phân loại tốt hơn với các lát cắt Flair (>99%), T2 (>98%), T1C (>97%) và hình ảnh lâm sàng (>98%).

Bài báo "DenseNet for Anatomical Brain Segmentation"[9]. Tác giả của bài báo là Ram Deepak Gottapu và Cihan H Dagli. Bài báo tập trung vào việc phân segmen bộ não từ hình ảnh cộng hưởng từ (MRI) bằng một kiến trúc gọi là DenseNet. Phân segmen tự động trong hình ảnh MRI não đóng vai trò quan trọng trong phân tích nhiều bệnh và tình trạng. Công trình đã chỉ ra rằng mạng nơ-ron tích chập (DenseNet) có thể đạt độ chính xác cao hơn và sử dụng ít tham số hơn nếu mỗi lớp trong mạng được kết nối với mọi lớp khác theo kiểu feed forward. Các tác giả áp dụng ý tưởng này và tạo ra một kiến trúc mới có thể gán mỗi điểm ảnh/voxel trong hình ảnh MRI của não với vùng giải phẫu tương ứng. Để đánh giá mô hình, các tác giả sử dụng bộ dữ liệu được cung cấp bởi IBSR 2 (Internet Brain Segmentation Repository), bao gồm 18 hình ảnh MRI của não được phân đoạn thủ công. Theo kiến thức của chúng tôi, phương pháp của chúng tôi là phương pháp đầu tiên sử dụng DenseNet để thực hiện phân đoạn giải phẫu của toàn bộ não.

Bài báo "AResU-Net: Attention Residual U-Net for Brain Tumor Segmentation"[10]. Các tác giả của bài báo là Jianxin Zhang, Xiaogang Lv, Hengbo Zhang và Bin Liu. Bài báo tập trung vào việc phân đoạn tự động khối u não từ hình ảnh cộng hưởng từ (MRI). Việc phân đoạn khối u não trong hình ảnh MRI là một nhiệm vụ khó khăn

do kích thước và hình dạng không đồng đều, không đều và không có cấu trúc của các khối u. Gần đây, các phương pháp phân đoạn khối u não dựa trên kiến trúc U-Net đối xứng đã đạt được hiệu suất khả quan. Đồng thời, hiệu quả của việc cải thiện các phản ứng cục bộ để trích xuất và khôi phục đặc trưng cũng đã được thể hiện trong các nghiên cứu gần đây, điều này có thể khuyến khích hiệu suất tốt hơn của bài toán phân đoạn khối u não. Được truy cảm hứng bởi điều này, chúng tôi đã thử nghiệm việc giới thiệu cơ chế chú ý vào kiến trúc U-Net hiện có để khám phá tác động của các phản ứng quan trọng cục bộ đối với bài toán này. Cụ thể hơn, chúng tôi đề xuất mạng phân đoạn khối u não 2D từ đầu đến cuối, tức là AResU-Net, kết hợp cơ chế chú ý và đơn vị còn lại vào U-Net để cải thiện hiệu suất phân đoạn khối u não. AResU-Net thêm một loạt các đơn vị chú ý vào quá trình down-sampling và up-sampling tương ứng và điều chỉnh tỷ lệ các đặc trưng một cách linh hoạt để cải thiện hiệu suất các phản ứng cục bộ của các đặc trưng phục hồi trong quá trình up-sampling. Chúng tôi đánh giá một cách chi tiết hiệu suất của AResU-Net trên hai bộ dữ liệu thử nghiệm về phân đoạn khối u não từ hình ảnh MRI là BraTS 2017 và BraTS 2018. Kết quả thực nghiệm cho thấy AResU-Net đề xuất vượt trội so với các phương pháp cơ bản và đạt được hiệu suất tương đương với các phương pháp phân đoạn khối u não tiêu biểu.

Bài báo tựa đề "A hybrid DenseNet121-UNet model for brain tumor segmentation from MR Images"[11]. Các tác giả của bài báo là Necip Cinar, Alper Ozcan và Mehmet Kaya. Bài báo tập trung vào việc phân đoạn khối u não từ hình ảnh cộng hưởng từ (MRI) bằng các phương pháp học sâu. Mạng neural UNet, một trong các mạng học sâu, được sử dụng như một mô hình kết hợp với mô hình DenseNet121 được huấn luyện trước để thực hiện quá trình phân đoạn. Trong quá trình huấn luyện và kiểm tra mô hình, tác giả tập trung vào các vùng con nhỏ hơn của khối u mà bao gồm cấu trúc phức tạp. Mô hình đề xuất được xác thực trên tập dữ liệu về khối u não BRATS 2019, chứa các khối u glioma cấp độ cao và cấp độ thấp. Kết quả thực nghiệm cho thấy mô hình đề xuất hoạt động tốt hơn so với các phương pháp tiên tiến khác được trình bày trong lĩnh vực này. Cụ thể, chỉ số Dice Similarity Coefficient (DSC) tốt nhất được đạt bằng cách sử dụng phương pháp đề xuất để phân đoạn toàn bộ khối u (WT), nhân trung tâm của khối u (CT) và khối u tăng cường (ET).

Bài báo có tựa đề "Performance Comparison between VGG16 & VGG19 Deep Learning Method with CNN for Brain Tumor Detection"[12]. Tác giả của bài báo là Sami

Ul Hoque. Bài báo này tập trung vào việc so sánh hiệu suất giữa hai phương pháp học sâu VGG16 và VGG19 với mạng neural tích chập (CNN) để phát hiện khối u não. Nghiên cứu này giải quyết vấn đề phân đoạn các mô não bất thường và mô não bình thường như chất xám (GM), chất trắng (WM) và chất lỏng não tủy (CSF) từ hình ảnh từ cộng hưởng từ (MR) bằng kỹ thuật trích xuất đặc trưng và bộ phân loại máy vector hỗ trợ (SVM). Khối u là sự phát triển không kiểm soát của các tế bào ung thư ở bất kỳ phần nào của cơ thể, trong khi khối u não là sự phát triển không kiểm soát của các tế bào ung thư trong não. Một khối u não có thể lành tính hoặc ác tính. Khối u não lành tính có tính đồng nhất về cấu trúc và không chứa các tế bào ung thư (hoạt động), trong khi khối u não ác tính không đồng nhất về cấu trúc (khác nhau) và chứa các tế bào hoạt động. Gliomas và meningiomas là ví dụ về khối u cấp độ thấp, được phân loại là khối u lành tính, và glioblastoma và astrocytomas là lớp khối u cấp độ cao, được phân loại là khối u ác tính. Theo Tổ chức Y tế Thế giới (WHO) và Hiệp hội Khối u Não Mỹ (ABTA), hệ thống phân loại phổ biến nhất sử dụng một thang điểm từ cấp độ I đến cấp độ IV để phân loại các loại khối u lành tính và ác tính. Có nhiều phương pháp được sử dụng để phát hiện khối u não và phân loại chúng, trong đó VGG16 và VGG19 được sử dụng để so sánh. Trong bài báo này, tác giả mô tả sự khác biệt so sánh giữa hai phương pháp và cho thấy kết quả tốt hơn so với các phương pháp khác. Điều này hữu ích cho nghiên cứu tương lai trong lĩnh vực y học. So sánh cho thấy cho cả VGG16 và VGG19, độ chính xác thấp. Tuy nhiên, từ phân tích của chúng tôi, chúng tôi nhận thấy rằng VGG16 hoạt động tốt hơn so với VGG19. Độ chính xác tối đa cho VGG16 là 56,5% và cho VGG19 là 53%.

Bài báo mang tựa đề "Optimized U-Net for Brain Tumor Segmentation"[13] và có các tác giả là Michał Futrega, Alexandre Milesi, Michał Marcinkiewicz và Pablo Ribalta. Nghiên cứu này đề xuất một kiến trúc U-Net được tối ưu hóa cho nhiệm vụ phân segmet khối u não trong thử thách BraTS21. Để tìm kiến trúc mô hình tối ưu và lịch học, các tác giả đã tiến hành một nghiên cứu hồi tả rộng để kiểm tra: hàm mất mát giám sát sâu, Focal loss, sự chú ý của bộ giải mã, drop block và kết nối dư thừa. Ngoài ra, họ đã tìm kiếm độ sâu tối ưu của bộ mã hóa U-Net, số kênh tích chập và chiến lược xử lý sau khi phân segmet. Phương pháp của họ đã giành chiến thắng trong giai đoạn xác thực và đạt vị trí thứ ba trong

giai đoạn kiểm tra. Họ đã công khai mã nguồn để tái tạo kết quả của mình trong cuộc thi BraTS21 trên NVIDIA Deep Learning Examples GitHub Repository.

Tiêu đề bài báo: "A comparison between the VGG16, VGG19, and ResNet50 architecture frameworks for classification of normal and CLAHE processed medical images"[14] của tác giả: Kamal Kamal, Hamid EZ-ZAHRAOUI. Bài báo này tập trung vào việc tìm kiếm mô hình mạng neural nổi bật nhất cho nhiệm vụ phân loại nhị phân trong hình ảnh y tế với kiểu ảnh nền thông thường và được cải thiện bằng phương pháp CLAHE (Contrast Limited Adaptive Histogram Equalization). Bằng cách so sánh kiến trúc của VGG16, VGG19 và ResNet50, bài báo đánh giá độ chính xác, F1 score và Recall cho một tập hình ảnh được chọn về não bộ với các trường hợp ảnh nền thông thường và ảnh nền được cải thiện bằng CLAHE. Sau khi thực hiện 10 thí nghiệm với hai chế độ: có và không có gia tăng dữ liệu, và giữ nguyên các tham số mô phỏng, bài báo kết luận rằng VGG16 là kiến trúc tốt nhất cho việc phân loại hình ảnh y tế. Kết quả cho trường hợp sử dụng dữ liệu gia tăng và ảnh nền thông thường cho ba mô hình là 0.78, 0.66 và 0.69 tương ứng cho VGG16, VGG19 và ResNet50. Trong trường hợp sử dụng ảnh nền được cải thiện bằng CLAHE, độ chính xác của ba mô hình lần lượt là 0.78, 0.76 và 0.72. Đối với chế độ không có gia tăng dữ liệu và ảnh nền thông thường, độ chính xác của ba mô hình lần lượt là 0.88, 0.79 và 0.75. Trong trường hợp sử dụng ảnh nền được cải thiện bằng CLAHE, độ chính xác của ba mô hình lần lượt là 0.86, 0.87 và 0.65.

Tiêu đề bài báo: "Comparison of Backbones for Semantic Segmentation Network"[15] của tác giả: Rongyu Zhang, Lixuan Du, Qi Xiao, Jiaming Liu. Bài báo này tập trung vào việc so sánh hiệu suất của các mạng lõi (backbones) khác nhau trong mạng phân đoạn ngữ nghĩa (semantic segmentation). Bài báo lựa chọn tập dữ liệu trích xuất đường phố từ CVPR DeepGlobe và so sánh sự khác biệt về hiệu suất giữa VGG-16 làm mạng lõi của mạng Unet, ResNet34, ResNet101 và Xception làm mạng lõi của AD-LinkNet. Khi sử dụng VGG-16 làm mạng lõi trong mạng phân đoạn ngữ nghĩa, nó thể hiện hiệu suất tốt hơn trong việc trích xuất đường dài và rộng. Với vai trò là mạng lõi trong mạng phân đoạn ngữ nghĩa, ResNet có khả năng cao hơn trong việc trích xuất đường nhỏ. Khi sử dụng Xception làm mạng lõi trong mạng phân đoạn ngữ nghĩa, nó không chỉ giữ lại các đặc

điểm của ResNet34 mà còn có thể xử lý hiệu quả các tình huống phức tạp trong việc trích xuất mục tiêu bị che khuất.

1.6 Vấn đề đặt ra

Vấn đề đặt ra trong Cơ sở lý luận Ứng dụng mô hình mạng U-Net trong phân đoạn khối u não trên ảnh CT/MRI là áp dụng một phương pháp hiệu quả và chính xác để phân đoạn khối u não trên ảnh CT/MRI. Vấn đề này là một thách thức quan trọng trong lĩnh vực y tế, đặc biệt là trong chẩn đoán và điều trị bệnh nhân ung thư não.

Trước đây, quá trình phân đoạn khối u trên ảnh CT/MRI thường được thực hiện thủ công bởi các chuyên gia y tế. Tuy nhiên, phương pháp này tốn nhiều thời gian và công sức, đồng thời còn tồn tại nguy cơ sai sót và không đồng nhất giữa các chuyên gia. Do đó, ứng dụng mô hình mạng U-Net để tự động hóa quá trình phân đoạn khối u trên ảnh CT/MRI trở thành một giải pháp hứa hẹn.

Mô hình mạng U-Net được thiết kế đặc biệt để xử lý nhiệm vụ phân đoạn hình ảnh, với cấu trúc mạng tổng quát và khả năng học từ ít dữ liệu huấn luyện. Qua việc sử dụng kiến trúc encoder-decoder và các kỹ thuật như skip connections, mô hình U-Net có khả năng học được các đặc trưng cục bộ và toàn cục của khối u, từ đó tạo ra các dự đoán phân đoạn chính xác.

Ứng dụng mô hình mạng U-Net trong phân đoạn khối u não trên ảnh CT/MRI có thể đem lại nhiều lợi ích. Đầu tiên, nó giúp tăng tốc độ và độ chính xác của quá trình phân tích hình ảnh, giảm công sức và thời gian cho các chuyên gia y tế. Thứ hai, nó cung cấp thông tin chính xác về vị trí và kích thước của khối u, từ đó hỗ trợ việc lên kế hoạch điều trị và đánh giá hiệu quả của nó. Cuối cùng, ứng dụng mô hình mạng U-Net trong phân đoạn khối u não có tiềm năng để cải thiện chất lượng chẩn đoán và giúp các bác sĩ và chuyên gia y tế đưa ra quyết định điều trị chính xác hơn.

Tuy nhiên, vấn đề đặt ra là cần giải quyết những thách thức liên quan đến ứng dụng mô hình mạng U-Net trong phân đoạn khối u não trên ảnh CT/MRI. Đầu tiên, xử lý dữ liệu y tế phức tạp và đa dạng là một thách thức đáng kể, đòi hỏi sự xử lý và tiền xử lý dữ liệu kỹ lưỡng để đảm bảo chất lượng đầu vào cho mô hình. Thứ hai, cần đánh giá hiệu suất của mô hình mạng U-Net trong việc phân đoạn khối u, đo lường độ chính xác, độ nhạy và đặcificity của mô hình. Cuối cùng, tối ưu hóa quá trình huấn luyện mô hình để đạt được kết

quả tốt nhất và đảm bảo tính ổn định của mô hình trong việc áp dụng thực tế cũng là một thách thức quan trọng.

Vì vậy, nghiên cứu về ứng dụng mô hình mạng U-Net trong phân đoạn khối u não trên ảnh CT/MRI là cần thiết để giải quyết những vấn đề trên và mang lại những cải tiến quan trọng trong lĩnh vực chẩn đoán và điều trị bệnh nhân ung thư não.

1.7 Hướng giải quyết của đề tài

Thu thập và tiền xử lý dữ liệu: Để đảm bảo chất lượng đầu vào cho mô hình mạng U-Net, cần thu thập và tiền xử lý dữ liệu ảnh CT/MRI chính xác và đa dạng. Việc này bao gồm loại bỏ nhiễu, chuẩn hóa độ sáng, cắt ảnh thành các mẫu nhỏ hơn và áp dụng các kỹ thuật tiền xử lý khác để cải thiện chất lượng và đồng nhất dữ liệu đầu vào.

Xây dựng mô hình mạng U-Net: Thiết kế một mô hình mạng U-Net phù hợp với bài toán phân đoạn khối u não trên ảnh CT/MRI. Mô hình U-Net có cấu trúc encoder-decoder với các skip connections để học được cả thông tin cục bộ và toàn cục của khối u. Đồng thời, cần tinh chỉnh các tham số và kiến trúc mạng để đạt được hiệu suất tốt và khả năng phân đoạn chính xác.

Huấn luyện mô hình: Sử dụng dữ liệu huấn luyện đã tiền xử lý, tiến hành huấn luyện mô hình mạng U-Net. Quá trình huấn luyện bao gồm việc chia dữ liệu thành tập huấn luyện và tập kiểm tra, áp dụng các kỹ thuật như data augmentation để tăng cường dữ liệu huấn luyện và tránh overfitting. Đồng thời, tối ưu hóa hàm mất mát và cập nhật trọng số mô hình để đạt được kết quả tốt nhất.

Đánh giá hiệu suất mô hình: Để đánh giá hiệu suất của mô hình mạng U-Net trong phân đoạn khối u não trên ảnh CT/MRI, cần sử dụng các phép đo như độ chính xác, độ nhạy và độ đo F1-score để đo lường khả năng phân đoạn chính xác của mô hình. So sánh kết quả của mô hình với các phương pháp khác và các chuyên gia y tế để đánh giá hiệu suất và độ tin cậy của mô hình.

Tối ưu hóa và ứng dụng thực tế: Tối ưu hóa quá trình huấn luyện và áp dụng mô hình mạng U-Net vào thực tế trong lĩnh vực chẩn đoán và điều trị bệnh nhân ung thư não. Đảm bảo tính ổn định và độ tin cậy của mô hình, kiểm tra và đánh giá kết quả trên các tập dữ liệu mới và đối tượng khác nhau. Đồng thời, phát triển các công cụ và giao diện hỗ trợ để tích hợp mô hình mạng U-Net vào quy trình làm việc của bác sĩ và chuyên gia y tế.

CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

2.1 Khối u não

2.1.1. *Khái quát*

U não [21] là một tập hợp số lượng lớn các tế bào não phát triển bất thường vượt ngoài tầm kiểm soát của cơ thể. Các u não có thể bắt đầu trực tiếp từ tế bào não, tế bào đệm của hệ thần kinh trung ương, hoặc cũng có thể bắt đầu từ các bộ phận khác (ví dụ như phổi, thận...) rồi theo máu đến não, được gọi là u di căn não.

Theo ThS.BS.CKII Chu Tấn Sĩ - Trưởng khoa Ngoại thần kinh, Trung tâm Thần kinh, Bệnh viện Đa khoa Tâm Anh TP HCM U não chiếm 2% trong tổng số các ca ung thư từ mọi nhóm tuổi. Bác sĩ Sĩ dẫn các nghiên cứu cho thấy, u não chiếm 2% trong tổng số các ca ung thư từ mọi nhóm tuổi. Trong số các trường hợp tử vong do ung thư ở nhóm trẻ em dưới 15 tuổi và nhóm từ 20-39 tuổi, bệnh u não là nguyên nhân gây tử vong cao thứ hai. Người ngoài 85 tuổi có tỷ lệ bị u não cao nhất.

Tại Việt Nam, chỉ tính riêng bệnh viện Việt Đức, mỗi năm có khoảng 2.5000 người bệnh được phát hiện mắc u não.

Bệnh u não ác tính rất phổ biến ở Đức với hơn 8.000 người mắc các bệnh liên quan. Hơn 500 trường hợp mới được chẩn đoán mắc u não mỗi ngày trên toàn thế giới. Số bệnh nhân có khối u di căn não còn nhiều hơn thế. Đây là loại ung thư phổ biến thứ 2 ở trẻ em (sau ung thư máu) và chiếm 30% tổng số ca tử vong vì ung thư ở trẻ em.

U não là căn bệnh không hiếm gặp trên thế giới. Ước tính mỗi năm có khoảng 19.000 bệnh nhân u não mới được phát hiện ở Mỹ (dân số 300 triệu dân). Hiện tại, có tới gần 700.000 người Mỹ đang sống chung với căn bệnh này, trong đó 130.000 người mắc u ác tính.

2.1.2. **Khái quát Nguyên nhân tạo ra khối u và các yếu tố liên quan**

Nguyên nhân gây u não chính xác trong hầu hết mọi trường hợp là không thể xác định. Tuy nhiên, nguyên nhân tiềm ẩn thì có rất nhiều. Bất cứ điều gì làm tăng khả năng mắc bệnh u não đều là có thể được xem là một yếu tố nguy cơ của bệnh (tức nguyên nhân tiềm ẩn).

Việc bạn có một trong số các nguy cơ gây u não dưới đây không chắc chắn là bạn sẽ bị u não trong tương lai. Các yếu tố nguy cơ làm tăng rủi ro u não có thể là:

Tuổi tác

Người càng lớn tuổi càng có nguy cơ bị u não. Hầu hết các khối u não xảy ra ở người lớn tuổi từ 85 đến 89, mặc dù vẫn có một số loại u não phổ biến hơn ở trẻ em dưới 15 tuổi.

Tiền sử gia đình (di truyền)

- Theo báo cáo, chỉ có từ 5-10% tổng số ca ung thư là do di truyền. U não chỉ chiếm 2% tổng số ca ung thư trên toàn thế giới, do đó tỉ lệ khối u não được di truyền là rất thấp.
- Một số tình trạng di truyền được biết là làm tăng nguy cơ mắc khối u não, bao gồm: bệnh xơ cứng củ, bệnh u sợi thần kinh loại 1, loại 2, hội chứng Turner, hội chứng Li-Fraumeni, hội chứng Turcot, hội chứng Gorlin,...

Chế độ ăn thiếu khoa học

- Một số nghiên cứu đã chỉ ra rằng các hợp chất N-nitroso trong chế độ ăn uống có thể ảnh hưởng đến nguy cơ mắc các khối u não ở trẻ em và người lớn.
- Gần đây, Tiến sĩ. Lee Wrensch phát hiện ra rằng người mắc bệnh u thần kinh đệm có tỉ lệ lớn tiêu thụ chế độ ăn ít trái cây, ít rau quả, ít vitamin C mà chứa nhiều nitrit như phô mai, cá, thịt xông khói, thức ăn đã qua chế biến, lên men, ủ muối qua đêm (cá khô), đồ đóng hộp.

Thừa cân và béo phì

- Thừa cân hoặc béo phì làm tăng nguy cơ mắc bệnh u màng não. Khoảng 2% tổng số ca được chẩn đoán u não ở Anh mỗi năm là do thừa cân hoặc béo phì.
- Cố gắng giữ cân nặng lành mạnh bằng cách tham gia các hoạt động thể chất và ăn uống điều độ là một việc làm cần thiết.

Không có tiền sử bệnh thủy đậu

Dựa theo một báo cáo năm 2016 được xuất bản trên tạp chí Cancer Medicine, những người chưa có tiền sử mắc bệnh thủy đậu ở thời thơ ấu nguy cơ phát triển u thần kinh đệm cao hơn 21% so với người đã nhiễm bệnh thủy đậu.

Phơi nhiễm hóa chất

Một số ngành nghề do môi trường làm việc đặc thù cần tiếp xúc với nhiều hóa chất có thể làm tăng nguy cơ ung thư não, chẳng hạn như:

- Người làm nông nghiệp phải tiếp xúc nhiều với thuốc trừ sâu.

- Công nhân làm việc trong môi trường nhiều kim loại nặng (niken, thủy ngân).
- Người làm ngành vật liệu xây dựng, tấm lợp, gạch lát, đóng tàu thuyền,..do tiếp xúc nhiều với chất amiăng có thể gây u não.
- Người sống gần nguồn nước thải công nghiệp, bãi rác tập thể chứa Vinyl Clorua.
- Nhân viên xăng dầu không bảo hộ kỹ hít phải hợp chất vòng thơm như benzene, ethylbenzene trong thời gian dài.

Tiếp xúc với bức xạ

- Bức xạ ion hóa là một loại bức xạ được sử dụng bởi một số quy trình quét y tế, chẳng hạn như chụp X-quang và chụp CT. Những người đã tiếp xúc với bức xạ ion hóa có nguy cơ mắc các khối u não cao hơn người bình thường.
- Do đó, nếu bạn đã có tiền sử xạ trị trước đây với các bệnh ung thư khác thì cũng có thể làm tăng nguy cơ bị u não của bạn lên một chút. Tuy nhiên, u não do tiếp xúc với bức xạ xảy ra với tỉ lệ rất hiếm (dưới 1%).

2.1.2. Dấu hiệu và triệu chứng

Tùy theo vị trí khối u mà triệu chứng của bệnh u não cũng rất đa dạng và có thể ảnh hưởng đến các bộ phận cơ thể khác nhau. Cụ thể:

- Các dấu hiệu u não thùy trán: Thùy trán kiểm soát chuyển động như đi bộ và là một phần tính cách của bạn. Một khối u ở thùy trán có thể gây ra: Tê tay, tê chân, run tay, run chân, chuyển động tay không thể kiểm soát, đi lại khó khăn; Suy giảm khả năng nghe, vị giác hoặc khứu giác. Dễ có cảm giác buồn nôn hoặc nôn mửa liên tục. Thay đổi giọng nói, ăn không ngon, mất mùi, sụt cân; Tê hoặc ngứa ran ở một bên của cơ thể, yếu một tay, một chân hoặc co giật (đặc biệt ở người lớn); Thay đổi tính cách, cư xử theo cách mà bạn không thường làm.
- Các triệu chứng u não thùy thái dương: Thùy thái dương là nơi bạn xử lý âm thanh và là nơi bạn lưu giữ những ký ức. Một khối u ở khu vực này có thể gây ra: Mất trí nhớ ngắn hạn; Khó nghe và nói.
- Các dấu hiệu u não thùy đỉnh: Thùy đỉnh cho phép bạn nhận ra các đối tượng và lưu trữ kiến thức đó. Một khối u ở khu vực này có thể gây ra: Khó viết, khó đọc, khó nuốt, khó hiểu những gì người khác đang nói; Mất cảm giác ở một phần cơ thể.

- Các triệu chứng u não thùy chẩm: Thùy chẩm xử lý những gì bạn có thể nhìn thấy. Một khối u nằm trong khu vực này có thể gây ra các vấn đề về thị lực như: Tầm nhìn bị tối, hoa mắt, ảnh bị nhòe; Khó xác định màu sắc và kích thước của vật thể.
- Các dấu hiệu u não vùng tiểu não: Tiểu não kiểm soát sự cân bằng và tư thế của chúng ta. Vì vậy, một khối u ở khu vực này có thể gây ra: Vấn đề với sự phối hợp đi đứng và cân bằng; Chuyển động không kiểm soát của mắt như giật mí mắt.
- Các triệu chứng u não vùng thân não: Thân não kiểm soát các chức năng quan trọng của cơ thể như hô hấp. Một khối u ở khu vực này có thể gây ra: Khó nuốt và khó nói; Loạng choạng và đi lại khó khăn.
- Các dấu hiệu u não tuyến yên: Tuyến yên tạo ra các hormone quan trọng cho cơ thể bạn hoạt động. Một khối u trong khu vực này của não có thể gây ra: Tăng cân, lượng đường trong máu cao (tiểu đường); Vô sinh, rò rỉ sữa từ vú dù bạn đang không cho con bú; Thay đổi tâm trạng, tăng huyết áp.
- Các triệu chứng u não tuyến tùng: Tuyến tùng tạo ra một loại hormone gọi là melatonin phụ trách việc kiểm soát giấc ngủ. Các khối u trong khu vực này có thể gây ra: Mất ngủ, mệt mỏi, đau đầu; Mất thăng bằng khi đi bộ.
- Các dấu hiệu u não tủy sống: Tủy sống là một bó dây thần kinh dài trải dài từ não đến phần dưới của lưng. Một khối u trong tủy sống có thể gây đau và tê hoặc yếu ở các bộ phận khác nhau của cơ thể. Bạn cũng có thể mất kiểm soát bàng quang hoặc ruột, ảnh hưởng đến bài tiết và tiêu hóa.

2.2 Đặc điểm nhận dạng khối u não dựa trên ảnh MRI/CT

Đặc điểm nhận dạng khối u não dựa trên ảnh MRI/CT có thể bao gồm những yếu tố sau:

1. Kích thước và hình dạng: Khối u não có thể có kích thước và hình dạng đa dạng. Các đặc điểm này được phân tích để xác định mức độ phát triển và tương tác với các cấu trúc xung quanh.
2. Mật độ: Mật độ của khối u trong ảnh MRI/CT có thể đa dạng, từ mờ nhạt đến sáng tương phản cao. Đặc điểm này có thể liên quan đến loại khối u và cung cấp thông tin về tính chất mô và tương tác với mô xung quanh.

3. Biên: Đường biên của khối u trong ảnh MRI/CT có thể cung cấp thông tin về việc lan rộng của khối u và mức độ tách biệt với các cấu trúc xung quanh.

4. Mô cấu trúc nội bộ: Một số khối u não có cấu trúc nội bộ rõ ràng, ví dụ như các mạch máu, búi chất, hoặc các thành phần tương tự. Các đặc điểm này có thể giúp xác định loại khối u và đánh giá tính chất mô bên trong.

5. Vị trí: Vị trí của khối u trong não cung cấp thông tin về vị trí của nó đối với các cấu trúc não khác và có thể gợi ý về khả năng gây ảnh hưởng lên chức năng não.

6. Tái phát: Sự xuất hiện của khối u trong các hình ảnh theo thời gian (nếu có) có thể cho thấy tính chất tái phát và sự phát triển của khối u.

Những đặc điểm trên có thể được sử dụng để phân loại, đánh giá mức độ nghiêm trọng và theo dõi sự phát triển của khối u não trong quá trình chẩn đoán và điều trị. Tuy nhiên, việc nhận dạng khối u và đưa ra chẩn đoán cuối cùng thường yêu cầu sự kết hợp của các phương pháp hình ảnh, kiến thức lâm sàng và hỗ trợ từ các mô hình máy học và trí tuệ nhân tạo.



Hình 2. 1 Đặc điểm nhận dạng khối u não dựa trên ảnh MRI/CT[31]

2.3 Tổng quan về trí tuệ nhân tạo

2.3.1. Khái niệm

Trong khoa học máy tính, trí tuệ nhân tạo hay AI [2] (tiếng Anh: artificial intelligence), đôi khi được gọi là trí thông minh nhân tạo, là trí thông minh được thể hiện

bằng máy móc, trái ngược với trí thông minh tự nhiên của con người. Thông thường, thuật ngữ "trí tuệ nhân tạo" thường được sử dụng để mô tả các máy móc (hoặc máy tính) có khả năng bắt chước các chức năng "nhận thức" mà con người thường phải liên kết với tâm trí, như "học tập" và "giải quyết vấn đề".

2.3.2. Một số công nghệ Trí tuệ nhân tạo phổ biến

2.3.2.1. Công nghệ AI phản ứng

Đây là loại có khả năng phân tích những động thái khả thi nhất của chính mình và của đối thủ, từ đó đưa ra được những giải pháp tối ưu nhất.

Ví dụ của công nghệ AI phản ứng là Deep Blue. Đây là một chương trình chơi cờ vua tự động, được tạo ra bởi IBM, với khả năng xác định các nước cờ cũng như dự đoán những bước đi tiếp theo của đối thủ.

2.3.2.2. Công nghệ AI với bộ nhớ hạn chế

Đây là loại có khả năng sử dụng những kinh nghiệm trong quá khứ để đưa ra những quyết định trong tương lai. Công nghệ AI này thường kết hợp với cảm biến môi trường xung quanh nhằm mục đích dự đoán những trường hợp có thể xảy ra và đưa ra quyết định tốt nhất cho thiết bị.

2.3.2.3. Lý thuyết trí tuệ nhân tạo

Đây là loại có thể học hỏi cũng như tự suy nghĩ, sau đó áp dụng những gì học được để thực hiện một việc cụ thể. Hiện nay, công nghệ AI này vẫn chưa trở thành một phương án khả thi.

2.3.2.4. Tự nhận thức

Chúng có khả năng tự nhận thức về bản thân, có ý thức và hành xử như con người vậy. Đây được xem là bước phát triển cao nhất của công nghệ AI cho đến thời điểm hiện tại, thế nhưng công nghệ này vẫn chưa khả thi.

2.3.3. Ứng dụng của trí tuệ nhân tạo trong cuộc sống hiện nay

Ứng dụng trí tuệ nhân tạo ở Việt Nam cũng như trên toàn cầu là vô cùng quan trọng. Lợi ích cũng như vai trò của trí tuệ nhân tạo vô cùng rộng khắp ở nhiều lĩnh vực, ngành nghề như:

2.3.3.1. *Vai trò của trí tuệ nhân tạo trong y tế*

AI trong robots y học: Các ứng dụng của công nghệ AI y tế cũng bao gồm các robot và thiết bị y tế hỗ trợ. Ví dụ, robot điều phối có thể tạo điều kiện giao tiếp giữa bệnh nhân với các chuyên gia y tế; thiết bị hỗ trợ di chuyển có thể giúp điều phối việc đi, đứng hoặc ngồi; và robot giống như động vật có thể giao tiếp và xác định bệnh nhân. Robot cũng có thể được sử dụng trong phẫu thuật với tư cách là trợ lý bác sĩ phẫu thuật, hệ thống phẫu thuật da Vinci là một trong những hệ thống phẫu thuật robot được sử dụng phổ biến nhất.

Hệ thống khai thác dữ liệu bệnh nhân và các nguồn dữ liệu khác để tạo thành một giả thuyết, sau đó nó đưa ra một lược đồ chấm điểm tin cậy.

AI trong ra quyết định lâm sàng: Với sự phát triển nhanh chóng của công nghệ, lượng thông tin mới từ các công trình nghiên cứu ngày càng nhiều và nhanh hơn. Hiện nay, lượng thông tin y khoa đã tăng gấp đôi cứ sau mỗi 3 năm. Ước tính rằng nếu một bác sĩ muốn cập nhật toàn bộ thông tin y khoa thì phải đọc 29 giờ mỗi ngày. Ngoài ra, nguồn dữ liệu lớn (big data), bao gồm các dữ liệu từ hồ sơ sức khỏe điện tử (EHR), các dữ liệu "omic" - dữ liệu về di truyền học (genomics), dữ liệu về chuyển hóa (metabolomics) và dữ liệu về protein (proteomics), và dữ liệu về xã hội học và lối sống là những nguồn dữ liệu sẽ không có ích nếu không được phân tích toàn diện. Giải pháp duy nhất để có thể tiếp cận và sử dụng khối lượng thông tin khổng lồ trong lĩnh vực y tế đó là sử dụng công nghệ trí tuệ nhân tạo (Artificial Intelligence - AI). IBM Watson là một trong những hệ thống hàng đầu hỗ trợ chăm sóc sức khỏe bằng AI, giúp các bác sĩ đưa ra quyết định hiệu quả. Với khả năng máy học (machine learning) và khả năng xử lý ngôn ngữ tự nhiên, hệ thống này giúp các bác sĩ xem xét hồ sơ sức khỏe điện tử của bệnh nhân và cập nhật thông tin hướng dẫn điều trị từ các công trình nghiên cứu y khoa. Một nghiên cứu mù đôi đã so sánh các quyết định của hội đồng các chuyên gia ung thư với quyết định của trí tuệ nhân tạo từ hệ thống ung thư Watson. Kết quả cho thấy 90% các khuyến nghị mà hệ thống đưa ra là phù hợp với các khuyến nghị của hội đồng các chuyên gia, nhưng hệ thống chỉ mất 40 giây để hoàn tất quy trình.

AI trong chẩn đoán: Lỗi chẩn đoán là một mối đe dọa nghiêm trọng đối với chất lượng và an toàn trong chăm sóc sức khỏe. Tại Mỹ, ước tính tỷ lệ lỗi chẩn đoán ngoại trú là 5,08%, tương đương 12 triệu người mỗi năm. Khoảng một nửa trong số các lỗi này có

khả năng gây hại. Công nghệ AI đã được sử dụng để cải thiện chất lượng chẩn đoán, đặc biệt là trong X-quang. AI dựa trên nguồn dữ liệu 129.450 hình ảnh lâm sàng để chẩn đoán bệnh ngoại da, kết quả đã chứng minh rằng hệ thống này có thể phân loại ung thư da ở mức tương đương với các bác sĩ da liễu. Một thuật toán dựa trên hình ảnh cộng hưởng từ của chuyển động tim cho phép dự đoán chính xác kết quả bệnh nhân bị tăng áp phổi; một phương pháp phân loại nhịp tự động trong phân tích điện tâm đồ liên tục (ECGs) ở những bệnh nhân mắc bệnh nghiêm trọng. Một nghiên cứu khác đã xem xét các kết quả đầy hứa hẹn sử dụng AI trong hình ảnh đột quy và cho rằng công nghệ AI có thể đóng một vai trò quan trọng trong việc quản lý bệnh nhân đột quy.

AI trong y học cá thể: Y học cá thể (Personalized medicine) là một mô hình chăm sóc sức khỏe mới, trong đó việc điều trị và phòng ngừa bệnh dựa trên các điều kiện cá nhân, bao gồm thông tin di truyền, đặc điểm tâm lý xã hội, môi trường và lối sống. Tất cả thông tin này tạo ra một lượng dữ liệu khổng lồ, chỉ có thể được phân tích và tích hợp bởi công nghệ AI.

2.3.3.2. Vai trò quan trọng của trí tuệ nhân tạo trong kinh doanh

Riêng trong lĩnh vực kinh doanh, trí tuệ nhân tạo (AI) góp phần mang đến những trải nghiệm tốt hơn cho hệ thống khách hàng. Thông qua việc thu thập cũng như phân tích dữ liệu, trí tuệ AI có thể nắm bắt thông tin cũng như hành vi sử dụng dịch vụ của khách hàng, từ đó mang lại những giải pháp hữu ích trong kinh doanh.

Còn với lĩnh vực truyền thông, trí tuệ nhân tạo trong marketing đóng vai trò quan trọng, góp phần làm thay đổi những cách thức tiếp cận với khách hàng mục tiêu. Với những ưu điểm của công nghệ AI, các công ty, tổ chức có thể triển khai những chiến dịch quảng cáo đúng thời điểm, đúng khách hàng mục tiêu dựa trên những phân tích chính xác về nhân khẩu học, thói quen, hành vi và nhu cầu.

Tự động hóa quá trình robot đang được áp dụng cho các nhiệm vụ lặp đi lặp lại thường được thực hiện bởi con người.

Các thuật toán Machine Learning đang được tích hợp vào các nền tảng phân tích và CRM để khám phá thông tin về cách phục vụ khách hàng tốt hơn. Chatbots đã được kết hợp vào các trang web để cung cấp dịch vụ ngay lập tức cho khách hàng. Tự động hóa các

vị trí công việc cũng đã trở thành một điểm nói chuyện giữa các học giả và các nhà phân tích.

2.3.3.3. Ứng dụng của trí tuệ nhân tạo trong giáo dục

Trí tuệ nhân tạo đã tạo nên những thay đổi lớn trong ngành giáo dục hiện nay. Cụ thể hơn, phải kể đến:

- Các hoạt động giáo dục như chấm điểm, dạy kèm có thể được tự động hóa nhờ công nghệ AI.
- Rất nhiều trò chơi, phần mềm giáo dục ra đời phổ biến để đáp ứng những nhu cầu của học sinh, sinh viên để cải thiện và nâng cao trình độ học tập của mỗi người.
- Chúng còn có thể chỉ ra những vấn đề mà các khóa học nên cải thiện: Nếu nhiều học sinh gửi đáp án sai cho bài tập, hệ thống sẽ thông báo cho các giáo viên đồng thời gửi thông điệp đến với các em học sinh để chỉnh sửa đáp án cho phù hợp nhất.
- Trí tuệ nhân tạo thay thế con người bởi còn có khả năng theo dõi sự tiến bộ của học sinh để từ đó đưa ra những thông báo đến với các giáo viên khi phát hiện ra nhiều vấn đề đối với kết quả học tập của học sinh.
- Trí tuệ nhân tạo cung cấp nhiều dữ liệu bổ ích để giúp các sinh viên lựa chọn được những khóa học tốt nhất cho mình.

2.3.3.4. Ứng dụng của trí tuệ nhân tạo trong tài chính

AI trong các ứng dụng tài chính cá nhân, như Mint hoặc Turbo Tax, đang phá vỡ các tổ chức tài chính. Các ứng dụng như thu thập dữ liệu cá nhân và cung cấp tư vấn tài chính. Các chương trình khác, như IBM Watson, đã được áp dụng cho quá trình mua nhà.

2.3.3.5. Ứng dụng của trí tuệ nhân tạo trong pháp luật

Tự động hóa quá trình này là sử dụng thời gian hiệu quả hơn. Các công ty khởi nghiệp cũng đang xây dựng các trợ lý máy tính hỏi và trả lời có thể sàng lọc các câu hỏi được lập trình để trả lời bằng cách kiểm tra phân loại và bản thể học liên quan đến cơ sở dữ liệu.

2.3.3.6. Ứng dụng của trí tuệ nhân tạo trong sản xuất

Đây là một lĩnh vực đã đi đầu trong việc kết hợp robot vào quy trình làm việc. Robot công nghiệp được sử dụng để thực hiện các nhiệm vụ đơn lẻ và được tách ra khỏi công nhân của con người, nhưng như công nghệ nâng cao mà thay đổi.

2.3.3.7. Ứng dụng của trí tuệ nhân tạo trong dịch vụ khách hàng

Cuối cùng nhưng không kém phần quan trọng, Google đang làm việc với một trợ lý AI có thể thực hiện các cuộc gọi giống như con người để thực hiện các cuộc hẹn tại, giả sử, tiệm tóc của khu phố. Ngoài các từ, hệ thống hiểu ngữ cảnh và sắc thái.

2.4. Kỹ thuật máy học (machine learning)

2.4.1. Khái niệm Machine Learning

Machine Learning (ML)[22] hay máy học là một nhánh của trí tuệ nhân tạo (AI), nó là một lĩnh vực nghiên cứu cho phép máy tính có khả năng cải thiện chính bản thân chúng dựa trên dữ liệu mẫu (training data) hoặc dựa vào kinh nghiệm (những gì đã được học). Machine learning có thể tự dự đoán hoặc đưa ra quyết định mà không cần được lập trình cụ thể. Bài toán machine learning thường được chia làm hai loại là dự đoán (prediction) và phân loại (classification). Các bài toán dự đoán như dự đoán giá nhà, giá xe... Các bài toán phân loại như nhận diện chữ viết tay, nhận diện đồ vật... Machine Learning (ML) hay máy học là một nhánh của trí tuệ nhân tạo (AI), nó là một lĩnh vực nghiên cứu cho phép máy tính có khả năng cải thiện chính bản thân chúng dựa trên dữ liệu mẫu (training data) hoặc dựa vào kinh nghiệm (những gì đã được học). Machine learning có thể tự dự đoán hoặc đưa ra quyết định mà không cần được lập trình cụ thể. Bài toán machine learning thường được chia làm hai loại là dự đoán (prediction) và phân loại (classification). Các bài toán dự đoán như dự đoán giá nhà, giá xe... Các bài toán phân loại như nhận diện chữ viết tay, nhận diện đồ vật...

2.4.2. Machine learning Workflow

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 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.

2.4.3. Phân loại Machine learning

Có rất nhiều cách phân loại machine learning, thông thường thì Machine Learning sẽ được phân làm hai loại chính sau:

- Supervised learning: học có giám sát.
- Unsupervised learning: học không giám sát.

Ngoài ra, machine learning còn có thể phân làm các loại sau:

- Semi-supervised learning: học bán giám sát.
- Deep learning: học sâu (về một vấn đề nào đó).
- Reinforce learning: học củng cố/tăng cường.

Supervised learning

Supervised learning là việc cho máy tính học trên dữ liệu đã được gán nhãn (label), hay nói cách khác, với mỗi đầu vào X_i , chúng ta sẽ có nhãn Y_i tương ứng.

Unsupervised learning

Unsupervised learning là cho máy tính học trên dữ liệu mà không được gán nhãn, các thuật toán machine learning sẽ tìm ra sự tương quan dữ liệu, mô hình hóa dữ liệu hay chính là làm cho máy tính có kiến thức, hiểu về dữ liệu, từ đó chúng có thể phân loại các

dữ liệu về sau thành các nhóm, lớp (clustering) giống nhau mà chúng đã được học hoặc giảm số chiều dữ liệu (dimension reduction).

2.4.4. Một số khái niệm cơ bản

Dataset (còn gọi là data corpus hay data stock): là tập dữ liệu ở dạng nguyên thủy chưa qua xử lý mà bạn đã thu thập được ở bước data collection. Một dataset sẽ bao gồm nhiều data point.

Data point: là điểm dữ liệu, mỗi điểm dữ liệu biểu diễn cho một quan sát. Mỗi data point sẽ có nhiều đặc trưng hay thuộc tính khác nhau, chúng được chia làm hai loại: loại dữ liệu số (numerical) và loại dữ liệu không phải số (ví dụ như chuỗi) (nonnumerical/categorical). Data point được biểu diễn thành dòng tương ứng, mỗi dòng có thể có 1 hoặc nhiều dữ liệu (chính là các đặc trưng).

Training data và test data: dataset thường sẽ được chia làm 2 tập này, training data dùng để huấn luyện cho mô hình, test data dùng để dự đoán kết quả và đánh giá mô hình. Có bài toán người ta sẽ cho sẵn hai tập này thì bạn không cần phải chia nữa, đối với bài toán chỉ cho mỗi dataset thôi thì phải chia ra. Thường tỷ lệ giữa tập train và test sẽ là 8/2.

Features vector: là vector đặc trưng, mỗi vector này sẽ biểu diễn cho một điểm dữ liệu trong dataset. Mỗi vector có n chiều biểu diễn các đặc trưng của điểm dữ liệu, mỗi đặc trưng là một chiều và phải là dữ liệu số. Các mô hình chỉ có thể huấn luyện được từ các vector đặc trưng này, do đó dataset cần phải chuyển về dạng một tập các vector đặc trưng (features vectors).

Model: là các mô hình được dùng để training trên một training data theo thuật toán của mô hình đó. Sau đó mô hình có thể dự đoán hoặc đưa ra các quyết định dựa trên những gì chúng đã được học.

2.4.5. Ứng dụng của Machine learning

Machine learning được ứng dụng cực kỳ nhiều trong đời sống hiện nay trong mọi lĩnh vực:

- Tài chính - ngân hàng.
- Sinh học, hóa học.
- Nông nghiệp.
- Tìm kiếm, trích xuất thông tin.

- Tự động hóa.
- Robotics.
- Mạng máy tính.
- Khoa học vũ trụ.
- Quảng cáo.
- Xử lý ngôn ngữ tự nhiên.
- Thị giác máy tính.

Và còn rất rất nhiều lĩnh vực mà machine learning có thể được áp dụng, machine learning tỏ ra cực kỳ hiệu quả, hơn hẳn con người trong cụ thể các lĩnh vực mà chúng được áp dụng.

Ví dụ đơn giản như dự báo thời tiết, người ta sẽ dùng các phép tính và những quan sát, ghi nhận về thời tiết trong quá khứ để dự báo về thời tiết của những ngày kế tiếp. Tuy nhiên sẽ thế nào nếu như có cực kỳ nhiều quan sát được thực hiện, có thể lên đến hàng triệu, hàng tỉ quan sát, lúc đó con người không thể nào thực hiện được việc tính toán trên dữ liệu lớn như vậy. Hơn nữa, việc tính toán với dữ liệu lớn như vậy có thể gặp sai sót và dẫn đến kết quả dự đoán bị sai.

Khi này, việc áp dụng machine learning vào để cho máy tính học các quan sát được ghi nhận trong quá khứ, chúng có thể dự đoán được thời tiết trong tương lai với độ chính xác cao hơn rất nhiều so với con người dự đoán.

2.5. Deep learning

2.5.1. Khái niệm Deep Learning

Deep Learning[1] là tập hợp con của Machine Learning và nó có tác dụng hỗ trợ cho máy tính tự huấn luyện chính nó để có thể thực hiện tác vụ tương tự như con người. Điều này chính là giúp máy tính bắt chước con người cách học hỏi và suy nghĩ. Deep Learning hoạt động dựa vào mạng lưới thần kinh nhân tạo và nó bao gồm nhiều lớp chứa các dữ liệu mô phỏng cách thức vận hành của não người. Các hệ thống của Deep Learning có khả năng cải thiện được những hiệu suất của chúng với quyền truy cập vào dữ liệu sẽ được nhiều hơn.

Thông thường, phiên bản máy sẽ có nhiều kinh nghiệm hơn; những máy móc đã có đủ kinh nghiệm thì sẽ được mang đi phục vụ cho những công việc như: phát hiện chẩn đoán trong y khoa, lái xe, các hệ thống trợ lý ảo,...

Deep Learning có hỗ trợ cho việc dịch ngôn ngữ, phân loại các hình ảnh, nhận dạng giọng nói. Chính vì thế, nó có thể được ứng dụng để giải quyết mọi nhu cầu cần nhận dạng mẫu mà không cần đến sự can thiệp của con người.

Bộ phận mạnh thần kinh nhân tạo này giống với não người bao gồm: Các node (đơn vị thần kinh trong mạng thần kinh nhân tạo) là những neuron thần kinh. Hầu hết các node bản thân thường chỉ có khả năng trả lời cho những câu hỏi đơn giản và cơ bản nhất; với những tác vụ khó thì chúng sẽ tiến hành liên kết với nhau để trả lời.

Có thể dạy hoặc có thể huấn luyện chúng bằng những thuật toán cụ thể. Những node giải đáp các câu hỏi phức tạp thì sẽ được gọi là mạng lưới thần kinh sâu (DNN), nó được định nghĩa như sau: Mạng lưới thần kinh sâu có khả năng thực hiện các hoạt động phức tạp như biểu diễn, trừu tượng có mang các ý nghĩa về âm thanh, văn bản và hình ảnh. Chúng được đánh giá là lĩnh vực phát triển nhất trong Machine Learning.

2.5.2. Cách thức hoạt động của Deep Learning

Deep Learning được xem là một phương pháp của học máy theo phương pháp: lập trình AI sẽ được huấn luyện để dự đoán đầu ra dựa vào một tập đầu vào. Ví dụ cụ thể: Dự đoán hành động của con mèo khi gặp chuột rồi tiến hành huấn luyện nó bằng phương pháp học có giám sát.

Khi ta muốn dự đoán được hành động của nó vào các đầu vào sẽ được thực hiện như sau:

- Lựa chọn con mèo phù hợp.
- Lúc này: các bộ phận trên cơ thể của con mèo như mắt, móng vuốt, tai,.. của con mèo sẽ trở nên rất nhanh nhạy.
- Con chuột sẽ xuất hiện ở đâu.

Về cơ bản thì Deep Learning sẽ không khác với học máy thông thường. Tuy nhiên, với ví dụ trên thì thời gian cần dành khá nhiều mới có thể thiết kế được các tính năng đại diện cho con mèo. Những gì cần làm là cung cấp cho hệ thống một số lượng hình ảnh con mèo, video con mèo bắt chuột thì hệ thống có thể tự học được những tính năng đại diện dành cho 1 con

mèo. Với những tác vụ như thị giác máy tính, nhận dạng giọng nói, robot hoặc dịch máy thì hiệu suất của Deep Learning có thể vượt xa những hệ thống máy khác. Tuy nhiên, việc xây dựng được hệ thống Deep Learning lại không mấy dễ dàng so với hệ thống máy móc thông thường.

Cách thức hoạt động của thuật toán Deep Learning diễn ra như sau: Các dòng thông tin sẽ được trải qua nhiều lớp cho đến lớp sau cùng. Lấy quy trình học của con người làm ví dụ cụ thể. Qua các lớp đầu tiên sẽ tập trung vào việc học các khái niệm cụ thể hơn trong khi các lớp sâu hơn sẽ sử dụng thông tin đã học để nghiên cứu và phân tích sâu hơn trong các khái niệm trừu tượng. Quy trình xây dựng biểu diễn dữ liệu này được gọi là trích xuất tính năng. Kiến trúc phức tạp của việc học sâu được cung cấp từ mạng lưới thần kinh sâu với khả năng thực hiện trích xuất tính năng tự động. Ngược lại, trong học máy thông thường còn gọi là học nông, nhiệm vụ này được thực hiện khi truy xuất các thuật toán cụ thể.

2.5.3. Ứng dụng của Deep Learning

Ứng dụng trong ngành công nghệ cao: Ứng dụng vượt bậc của Deep Learning không thể không kể đến đó chính là chế tạo Robot. Hiện nay, các phiên bản Robot giống con người với khả năng cảm nhận, phản ứng với môi trường đang dần được ra đời. Hiện nay, Robot còn có thể hợp tác với các hoạt động của con người và chúng có thể thực hiện được các nhiệm vụ riêng biệt phù hợp với thể mạnh của chúng. Robot đang góp phần thay thế con người trong việc thực hiện các công việc khó khăn hơn. Đây chính là phát minh tuyệt vời nhờ ứng dụng Deep Learning.

Ứng dụng trong nông nghiệp: Hiện nay, nhờ có Deep Learning mà người nông dân có thể triển khai những thiết bị có khả năng phân biệt được cỏ dại với cây trồng. Từ đó, các loại máy móc phun thuốc diệt cỏ có thể chọn lọc phun lên cỏ dại để đảm bảo cho cây trồng không bị ảnh hưởng. Ngoài vai trò loại bỏ cỏ dại bằng thuốc diệt cỏ thì nhờ Deep Learning mà sản lượng nông nghiệp ngày càng được cải thiện. Bên cạnh đó, Deep Learning đang được mở rộng hơn nữa sang các hoạt động như: thu hoạch, tưới tiêu, bón phân, gieo trồng.

Ứng dụng trong điện tử: Deep Learning được ứng dụng cho công việc nhận diện không mất trong các thiết bị điện tử như điện thoại, các phần mềm như facebook, messenger,... Thêm vào đó, Deep Learning còn được dùng trong nghe dịch và nói tự động mà trang web google dịch đang sử dụng. Ngoài ra, một số thiết bị trợ giúp tại nhà cũng

được ứng dụng Deep Learning để đáp ứng giọng nói và hiểu được sở thích của bản (đây chính là ứng dụng học sâu).

Ứng dụng cho lĩnh vực không gian vũ trụ và quốc phòng: Deep Learning được dùng với các mục đích như: xác định các vật thể như vệ tinh, xác định khu vực trung tâm, xác định khu vực an toàn và không an toàn trong quân đội. AI kết hợp với phân tích không gian địa lý có thể hỗ trợ việc trích xuất thông tin hữu ích từ các thiết bị được liên kết như radar và hệ thống nhận dạng tự trị. Dữ liệu này có thể hỗ trợ việc phát hiện bất kỳ hoạt động bất hợp pháp hoặc đáng ngờ nào, cũng như cảnh báo cho các cơ quan chức năng thích hợp. Việc nhận dạng và phân loại mục tiêu cũng có thể được hỗ trợ bởi rô bốt hỗ trợ AI và thị giác máy tính hỗ trợ IoT.

Ứng dụng trong y tế và chăm sóc sức khỏe: Deep Learning phát huy được hiệu quả siêu nổi bật của mình bằng cách hiển thị dữ liệu chất lượng cao và đo lường khả năng của các mạng thần kinh rồi thực hiện phân loại hình ảnh. Đồng thời, các nhà nghiên cứu về căn bệnh ung thư cũng đang ứng dụng DL để có thể tự động phát hiện ra các tế bào phát bệnh, nó được thực hiện như sau: chế tạo kính hiển vi tiên tiến để có thể tạo ra bộ dữ liệu có chiều cao phù hợp, được dùng để đào tạo một ứng dụng học sâu để có thể xác định tế bào ung thư. Ngoài ra, Deep Learning còn có bước tiến quan trọng trong việc cải thiện chất lượng sức khỏe của bệnh nhân bằng sự dự đoán với các sự kiện y tế từ dữ liệu hồ sơ sức khỏe điện tử.

Ứng dụng trong công nghiệp: Deep Learning có khả năng cải thiện sự an toàn cho các công nhân xung máy móc hạng nặng bằng cách: tự động phát hiện nếu như người hoặc vật nằm trong khoảng cách không an toàn đối với máy. Sản xuất thông minh hỗ trợ hệ thống thông tin đã tăng năng suất và chất lượng của các tổ chức công nghiệp. Sử dụng phân tích dữ liệu, mô hình thống kê và thuật toán dự đoán đã tăng lên theo các bước nhảy vọt, vì chất lượng và xu hướng của dữ liệu do máy tạo và con người tạo ra được cải thiện theo thời gian.

2.5.4. Vai trò của Deep Learning

Học sâu (deep learning), một nhánh của trí tuệ nhân tạo (AI), là một phương pháp tái tạo khả năng học hỏi và phát triển của con người với máy móc. Học sâu cho phép máy tính tiếp nhận thông tin mới, giải mã và tạo thông tin đầu ra mà không cần con người tham

gia vào quá trình. Lĩnh vực này có ý nghĩa rất lớn đối với các công nghệ của tương lai, bao gồm: xe tự lái, phần mềm nhận dạng khuôn mặt, y học cá nhân hóa...

Dạy máy tính tìm mẫu: Mục tiêu cuối cùng của việc học sâu là dạy cho máy tính cách đưa ra một tập hợp dữ liệu phi cấu trúc, để nhận ra các mẫu. Một ví dụ đơn giản về dữ liệu phi cấu trúc là hình ảnh của một cảnh trong thế giới thực, trong đó những thứ như bầu trời, cây cối và con người không được đánh dấu sẵn cho máy tính. Một thuật toán được đào tạo bởi học sâu sẽ xác định các thành phần riêng lẻ đó. Nó sẽ cho bạn biết những điểm ảnh (pixel) nào trong hình ảnh tạo nên một người, một cái cây và là một phần của bầu trời.

Ở quy mô rộng hơn, khả năng nhận dạng mẫu này có thể được áp dụng cho nhiều thứ khác. Ví dụ, trong một chiếc xe tự lái, máy tính sẽ có thể nhận ra một biển hiệu dừng lại và sau đó kích hoạt chiếc xe dừng lại một cách thích hợp. Trong y học, một thuật toán học sâu sẽ có thể nhìn vào hình ảnh của các tế bào qua kính hiển vi và xác định xem đó có phải là tế bào ung thư không.

Học sâu đóng vai trò ngày càng lớn trong công nghệ hiện đại. Về mặt lý thuyết, nó có thể vượt quá khả năng của bộ não con người, vì chúng có các thuật toán nền tảng rộng lớn, được gọi là mạng lưới thần kinh.

Kết nối của mạng lưới thần kinh: Mạng lưới thần kinh được đặt tên như vậy vì về cơ bản chúng bắt chước hoạt động của các tế bào thần kinh trong não người. Các mạng này được tạo thành từ 3 lớp tế bào thần kinh kỹ thuật số: lớp đầu vào, lớp ẩn và lớp đầu ra.

Lớp đầu vào là một chuỗi các nơ-ron kỹ thuật số có thể nhìn thấy thông tin mà máy tính được cung cấp. Ví dụ, một tế bào thần kinh có thể phát sáng khi màu xanh lục hiện diện trong một hình ảnh, trong khi một tế bào khác có thể phát sáng khi có hình dạng cụ thể, một đặc tính cụ thể trong dữ liệu.

Lớp đầu ra cho máy tính biết phải làm gì để đáp ứng với dữ liệu đầu vào. Trong một chiếc xe tự lái, đây sẽ là những tế bào thần kinh kỹ thuật số cuối cùng bảo máy tính quyết định tăng tốc, phanh hoặc quay đầu.

“Phép thuật” của mạng lưới thần kinh nằm ở lớp ẩn. Lớp này lấy các nơ-ron thần kinh từ lớp đầu vào và chuyển hướng để các nơ-ron lớp đầu ra lấy được. Lớp ẩn bao gồm hàng ngàn hoặc hàng triệu nơ-ron riêng lẻ, mỗi hàng được kết nối với nhau trong mạng lưới.

Công nghệ của tương lai: Học sâu hiện đang nắm giữ tiềm năng rất lớn cho các công nghệ tự động mới. Xe tự lái có lẽ là ứng dụng nổi bật nhất của thuật toán học sâu, nhưng có nhiều ứng dụng hơn trong thế giới kinh doanh và hơn thế nữa.

Ví dụ, học sâu có thể có ý nghĩa lớn đối với ngành tài chính. Các ngân hàng có thể sử dụng học sâu để giúp bảo vệ tài khoản trực tuyến bằng cách dạy một mô hình để xác định xem nỗ lực đăng nhập mới nhất của bạn có giống với đăng nhập thông thường hay không. Hoặc, các ngân hàng có thể áp dụng các thuật toán học sâu để đối phó tốt hơn các hoạt động lừa đảo và rửa tiền. Một tiềm năng đáng giá là ngân hàng và các nhà đầu tư có thể sử dụng học sâu để dự đoán khi nào giá cổ phiếu sắp tăng hay giảm.

Ứng dụng khác của công nghệ học sâu là nhận dạng khuôn mặt. Các thuật toán học sâu được đào tạo về hình ảnh khuôn mặt của bạn, sẽ cho phép phần mềm nhận dạng khuôn mặt bất kể bạn trông như thế nào vào một ngày nhất định. Điều đó sẽ khiến người khác không thể cố gắng xâm nhập tài khoản của bạn.

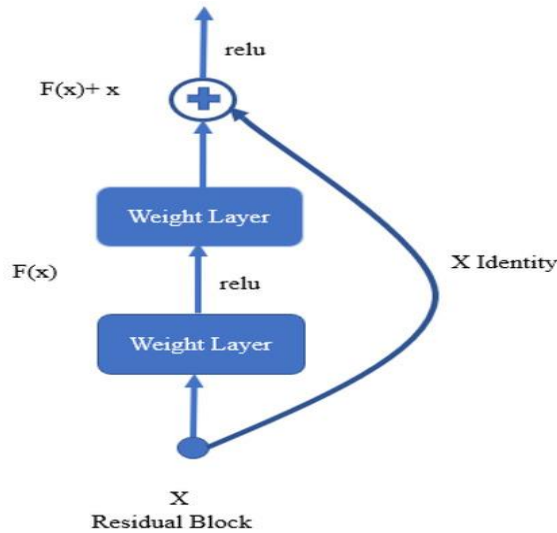
Học sâu cũng có thể giúp các nhà khoa học dự đoán động đất và các thảm họa thiên nhiên khác. Các mô hình học sâu có thể được huấn luyện về loại rung lắc nào xảy ra trước trận động đất trong quá khứ và sau đó phát báo động khi những mẫu tương tự được phát hiện trong tương lai.

Trong khi công nghệ học tập sâu tiếp tục được cải thiện, danh sách các ứng dụng tiềm năng ngày càng dài và ấn tượng hơn. Chúng ta có thể dạy cho máy tính nhận ra các mẫu, nhưng sự sáng tạo của con người sẽ rất cần thiết trong việc tìm ra cách tốt nhất để học sâu có thể làm việc cho xã hội.

2.6. Một số kiến mạng học sâu phổ biến

2.6.1. Kiến trúc mạng ResNet50

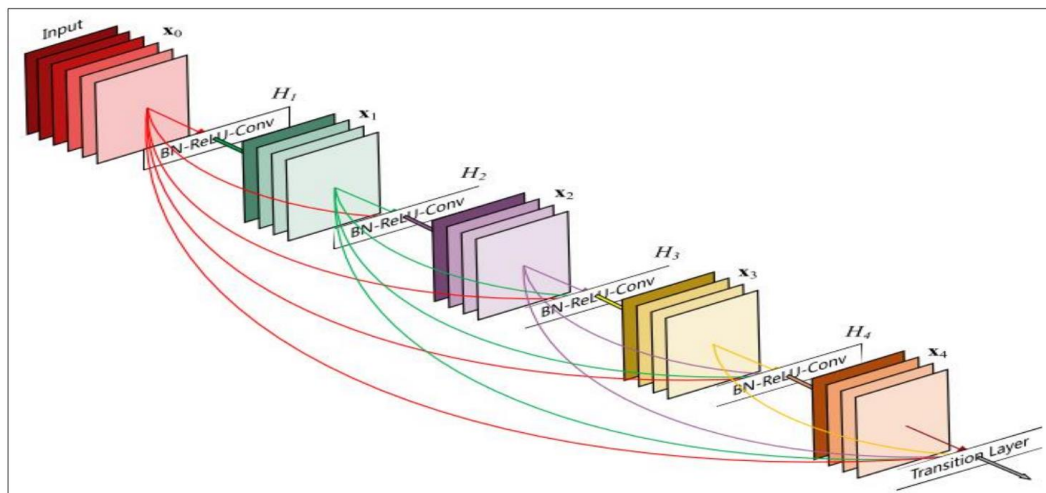
ResNet (viết tắt của Residual Network) [7], là mạng học sâu nhận được quan tâm từ những năm 2012 và trở nên phổ biến trong lĩnh vực thị giác máy tính. ResNet khiến cho việc huấn luyện hàng trăm thậm chí hàng nghìn lớp của mạng nơ-ron trở nên khả thi và hiệu quả. Ý tưởng chính của ResNet là sử dụng kết nối tắt đồng nhất để xuyên qua một hay nhiều lớp. Một khối như vậy được gọi là một residual block. ResNet-50 [14] là một mạng neural tích chập gồm 50 lớp (bao gồm 48 lớp tích chập, một lớp MaxPool và một lớp average pool).



Hình 2. 2. Cộng trực tiếp đầu vào của khối với nhánh còn lại trong khối Identity Block[7]

Việc xếp chồng các lớp sẽ không làm giảm hiệu suất mạng. Chúng ta có thể đơn giản xếp chồng các ánh xạ đồng nhất lên mạng hiện tại và hiệu quả của kiến trúc không thay đổi. Điều này giúp cho kiến trúc sâu ít nhất là không kém hơn các kiến trúc nông. Hơn nữa, với kiến trúc này, các lớp ở phía trên có được thông tin trực tiếp hơn từ các lớp dưới nên sẽ điều chỉnh trọng số hiệu quả hơn.

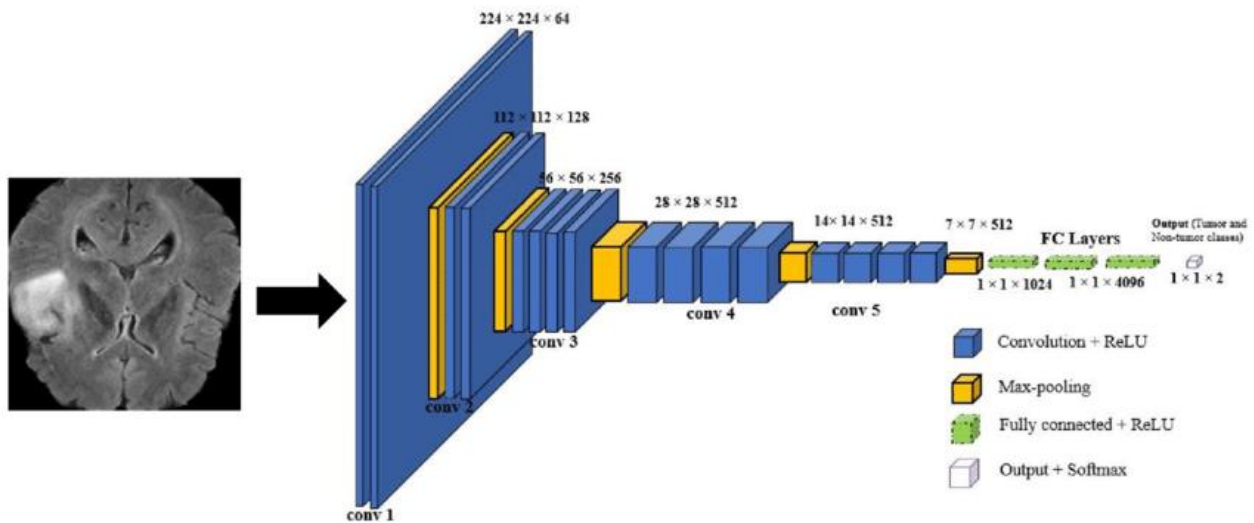
ResNet có khối tích chập (Convolutional Block, chính là Conv block trong hình) sử dụng bộ lọc kích thước 3x3 giống với của InceptionNet. Khối tích chập bao gồm 2 nhánh tích chập trong đó một nhánh áp dụng tích chập 1x1 trước khi cộng trực tiếp vào nhánh còn lại. Khối xác định (Identity block) thì không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



Hình 2. 3. Kiến trúc tổng quan của Resnet[26]

Phương pháp phân trực giác là thả ngẫu nhiên các lớp trong quá trình đào tạo và sử dụng mạng đầy đủ trong thử nghiệm. Sử dụng khối còn lại làm khối xây dựng mạng của họ, do đó, trong quá trình đào tạo, khi một khối còn lại cụ thể được kích hoạt, đầu vào của nó chảy qua cả phím tắt nhận dạng và các lớp trọng lượng, nếu không đầu vào chỉ chảy qua phím tắt nhận dạng. Trong thời gian đào tạo, mỗi lớp có một "xác suất sống sót" và được thả ngẫu nhiên. Trong thời gian thử nghiệm, tất cả các khối được duy trì hoạt động và hiệu chỉnh lại theo xác suất sống sót của nó trong quá trình đào tạo.

2.6.2. Kiến trúc mạng VGG-19



Hình 2. 4. Kiến trúc mạng VGG-19[12]

Giống như AlexNet và LeNet, mạng VGG được phân chia thành hai phần: phần đầu tiên gồm các tầng tích chập và tầng gộp, còn phần thứ hai bao gồm các tầng kết nối đầy đủ. Phần tích chập của mạng gồm các mô-đun vgg-block kết nối liên tiếp với nhau. Mạng VGG gốc có 5 khối tích chập, trong đó hai khối đầu tiên bao gồm một tầng tích chập ở mỗi khối, ba khối còn lại chứa hai tầng tích chập ở mỗi khối. Khối đầu tiên có 64 kênh đầu ra, mỗi khối tiếp theo nhân đôi số kênh đầu ra cho tới khi đạt giá trị 512. Vì mạng này sử dụng 8 tầng tích chập và 3 tầng kết nối đầy đủ.

Kiến trúc của VGG16 [27] bao gồm 16 layer :13 layer Conv (2 layer conv-conv,3 layer conv-conv-conv) đều có kernel 3×3 , sau mỗi layer conv là maxpooling downsize xuống 0.5, và 3 layer fully connection. VGG19 tương tự như VGG16 nhưng có thêm 3 layer convolution ở 3 layer conv cuối (thành 4 conv stack với nhau). VGG19 có 16 lớp tích chập, 3 lớp fully connected, 5 lớp MaxPool và 1 lớp SoftMax.

Mô hình VGG19 [8] là một trong những mô hình mạng nơ-ron tích chập (CNN) nổi tiếng và được sử dụng rộng rãi trong lĩnh vực xử lý ảnh. Nó được đặt tên theo Viện Nghiên cứu Visual Geometry (VGG) tại Đại học Oxford, người đã phát triển mô hình này.

Kiến trúc của VGG19 gồm tổng cộng 19 layer, trong đó bao gồm 16 layer tích chập (convolutional) và 3 layer fully connected. Mô hình VGG19 được chia thành hai phần chính: phần tích chập và phần kết nối đầy đủ.

Phần tích chập của VGG19 bao gồm sự sắp xếp liên tiếp các tầng tích chập và tầng gộp (pooling). Mỗi khối tích chập trong VGG19 gồm nhiều tầng tích chập liên tiếp. Đầu tiên, có hai khối với một tầng tích chập trong mỗi khối, sau đó là ba khối với hai tầng tích chập trong mỗi khối. Số lượng kênh đầu ra tăng gấp đôi sau mỗi khối tích chập cho đến khi đạt giá trị 512. Các tầng gộp (pooling) được sử dụng để giảm kích thước của đầu vào.

Cụ thể, kiến trúc của VGG19 là như sau:

- a. 13 tầng tích chập với kernel size là 3×3 và padding là 1, có stride mặc định là 1.
- b. Sau mỗi tầng tích chập, có một tầng gộp (max pooling) với kernel size là 2×2 và stride là 2. Điều này giúp giảm kích thước của đầu vào xuống một nửa.
- c. Cuối cùng, sau 13 tầng tích chập và 5 tầng gộp, dữ liệu được làm phẳng và đưa vào 3 tầng kết nối đầy đủ (fully connected).

Mô hình VGG19 có tổng cộng 143,667,240 tham số huấn luyện, là một mô hình rất lớn và mạnh mẽ. Nó thường được sử dụng cho các nhiệm vụ nhận dạng và phân loại ảnh trên các tập dữ liệu lớn như ImageNet.

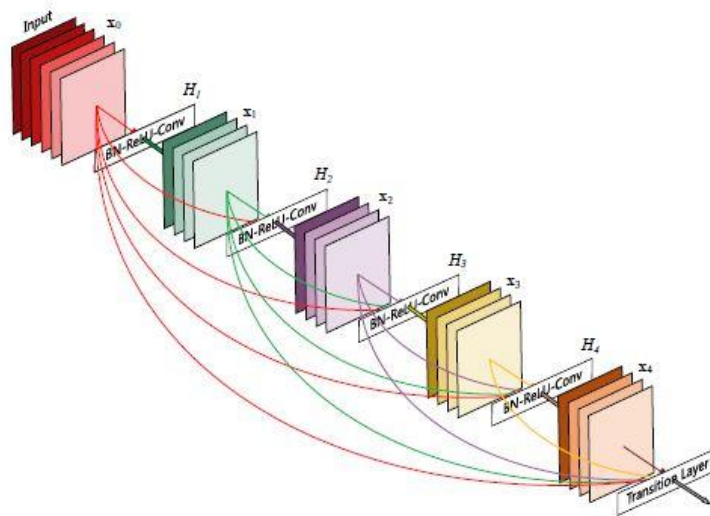
Để đánh giá hiệu suất của mô hình VGG19, các chỉ số thông thường như Accuracy (độ chính xác) và Loss (mất mát) được sử dụng. Accuracy đo lường tỷ lệ phân loại đúng của mô hình trên tập dữ liệu kiểm tra, trong khi Loss đo lường mức độ mất mát của mô hình trong quá trình huấn luyện.

2.6.3. Kiến trúc mạng DenseNet121

Densenet(Dense connected convolutional network) [9] là một trong những network mới nhất cho visual object recognition. Nó cũng gần giống Resnet nhưng có một vài điểm khác biệt. Densenet có cấu trúc gồm các dense block và các transition layers.

Với dense block là một khối xây dựng cấu trúc chính của mạng. Nó được sử dụng để kết hợp các lớp convolutional và tạo ra các kết nối mật thiết giữa các lớp trong khối đó. Mỗi lớp trong dense block nhận tất cả các đầu vào từ các lớp trước đó trong khối và sau đó chuyển đầu ra của nó cho tất cả các lớp sau đó trong khối. Điều này tạo ra một mạng nối tiếp mật thiết giữa các lớp và cho phép truyền thông tin đa chiều qua mạng. Bằng cách thực hiện việc này, dense block giúp tăng cường khả năng học và khả năng tái sử dụng đặc trưng, đồng thời giảm số lượng tham số cần học trong mạng; transition layers (còn được gọi là bottleneck layers) được sử dụng để thực hiện việc giảm kích thước của feature maps giữa các dense blocks. Chức năng chính của transition layers là thay đổi số kênh (number of channels) của feature maps và thực hiện thao tác down-sampling để giảm kích thước của chúng. Cụ thể, transition layers thường bao gồm các phép toán convolution và pooling để giảm kích thước không gian của feature maps. Một phương pháp thông thường là sử dụng một lớp convolution có kích thước kernel nhỏ (ví dụ: kernel 1x1) để giảm số kênh của feature maps xuống còn một nửa. Sau đó, một phép toán pooling (thường là max pooling) được áp dụng để giảm kích thước không gian của feature maps. Transition layers cho phép điều chỉnh số lượng kênh và kích thước của feature maps giữa các dense blocks. Điều này giúp kiểm soát độ sâu của mạng và cải thiện khả năng học và tính toán của mạng DenseNet.

DenseNets giải quyết vấn đề này bằng cách sửa đổi kiến trúc CNN tiêu chuẩn và đơn giản hóa mẫu kết nối giữa các lớp. Trong kiến trúc DenseNet, mỗi lớp được kết nối trực tiếp với mọi lớp khác, do đó có tên là Mạng tích chập được kết nối dày đặc. Đối với các lớp 'L', có các kết nối trực tiếp $L(L+1)/2$.



Hình 2. 5. Kiến trúc mạng Densenet [9]

DenseNet sử dụng kiến trúc “chuẩn hóa theo batch, hàm kích hoạt và phép tích chập” đã qua sửa đổi của ResNet. Về mặt kết nối xuyên tầng, không giống như trong ResNet khi đầu vào và đầu ra được cộng lại với nhau, DenseNet nối các đầu vào và đầu ra theo chiều kênh. Các thành phần chính tạo nên DenseNet là các khối dày đặc và các tầng chuyển tiếp.

Kích thước map đối tượng người dùng giống nhau trong khối rậm rạp để chúng hoàn toàn có thể được nối với nhau một cách thuận tiện. Ở cuối khối rậm rạp sau cuối, một tổng hợp trung bình toàn thể giới được thực thi và sau đó một bộ phân loại softmax được đính kèm.

Nhiều khối dày đặc: Chuyển đổi 1×1 theo sau là gộp trung bình 2×2 được sử dụng làm lớp chuyển tiếp giữa hai khối rậm rạp liền kề. Một khối Dense-block là một nhóm các layers kết nối tới tất cả các layer trước nó. Một layer bao gồm: Batch normalization, ReLU activation, 3×3 Conv.

Một khối dày đặc, nơi tất cả các đơn vị được kết nối hoàn toàn và mỗi tính năng được xử lý riêng biệt, phát triển nhanh chóng trong việc sử dụng bộ nhớ với kích thước đầu vào. Điều này có nghĩa là các khối dày đặc không phù hợp với dữ liệu như hình ảnh trong đó đầu vào là $height \times width \times number\ of\ channels$, đối với ảnh độ nét tiêu chuẩn thì ít nhất là 1 triệu. Các khối chuyển đổi, ví dụ như khối chuyển đổi 2D, tận dụng lợi thế tốt hơn của cấu trúc bên trong dữ liệu hình ảnh và cần ít bộ nhớ hơn để hoạt động.

Việc sử dụng pre-activation (BN và ReLU trước Conv) [20] hiệu quả hơn so với post-activation. Không giống như ResNet, DenseNet ghép tất cả các feature map lại với nhau. Mặt khác không thể ghép khi feature map khác kích thước. Do đó trong các Dense block, feature map ở mỗi layer có cùng kích thước. Tuy nhiên việc downsampling vốn do CNN thực hiện. Trong DenseNet sẽ do các transition layer đảm nhiệm. Kiến trúc gồm: Batch normalization, 1×1 Conv, 2×2 Average pooling. DenseNet có một số dạng như sau: DenseNet standard, DenseNets-B: là dạng chuẩn, sử dụng 1×1 Conv để giảm kích thước feature map trước khi đưa vào 3×3 Conv. Chữ B chính là Bottleneck trong mạng ResNet, DenseNets-BC: là một sự cải tiến của DenseNets-B trong đó giảm bớt số lượng feature maps. Chức năng phân loại, kích hoạt này phụ thuộc vào loại nhiệm vụ phân loại: Nếu dự đoán một trong nhiều nhãn là đúng, hãy sử dụng Softmax, nếu dự đoán nhiều nhãn độc lập, hãy sử dụng Sigmoid.

Kiến trúc DenseNet-121

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112×112	7×7 conv, stride 2			
Pooling	56×56	3×3 max pool, stride 2			
Dense Block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv			
	28×28	2×2 average pool, stride 2			
Dense Block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv			
	14×14	2×2 average pool, stride 2			
Dense Block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14×14	1×1 conv			
	7×7	2×2 average pool, stride 2			
Dense Block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1×1	7×7 global average pool			
		1000D fully-connected, softmax			

Hình 2. 6 Kiến trúc DenseNet-121[24]

Một tóm tắt về các kiến trúc khác nhau được triển khai cho cơ sở dữ liệu ImageNet đã được cung cấp trong bảng trên. Stride là số pixel dịch chuyển trên ma trận đầu vào. Stride của 'n' (giá trị mặc định là 1) chỉ ra rằng bộ lọc được di chuyển 'n' pixel mỗi lần.

Sử dụng kiến trúc DenseNet-121 để hiểu bảng trên, chúng ta có thể thấy rằng mỗi dense block có số lượng lớp (lặp lại) khác nhau, mỗi lớp gồm hai convolution; một lớp bottleneck với kernel kích thước 1×1 và lớp convolution với kernel kích thước 3×3 để thực hiện phép convolution.

Ngoài ra, mỗi transition layer trong DenseNet bao gồm một lớp convolution kích thước 1×1 và một lớp average pooling kích thước 2×2 với bước đi (stride) là 2. Do đó, các lớp có trong mạng DenseNet được sắp xếp như sau:

- Một lớp convolution cơ bản với 64 bộ lọc kích thước 7×7 và bước đi là 2.
- Một lớp pooling cơ bản với max pooling kích thước 3×3 và bước đi là 2.
- Dense Block 1 với 2 lớp convolution lặp lại 6 lần.
- Transition layer 1 (1 Conv + 1 AvgPool).
- Dense Block 2 với 2 lớp convolution lặp lại 12 lần.
- Transition layer 2 (1 Conv + 1 AvgPool).
- Dense Block 3 với 2 lớp convolution lặp lại 24 lần.
- Transition layer 3 (1 Conv + 1 AvgPool).
- Dense Block 4 với 2 lớp convolution lặp lại 16 lần.
- Lớp Global Average Pooling - nhận tất cả feature maps của mạng để thực hiện phân loại.
- Lớp đầu ra (Output layer).

Vì vậy, DenseNet-121 bao gồm các lớp sau:

- 1 lớp Convolution kích thước 7×7 .
- 58 lớp Convolution kích thước 3×3 .
- 61 lớp Convolution kích thước 1×1 .
- 4 lớp AvgPool.
- 1 lớp Fully Connected Layer.

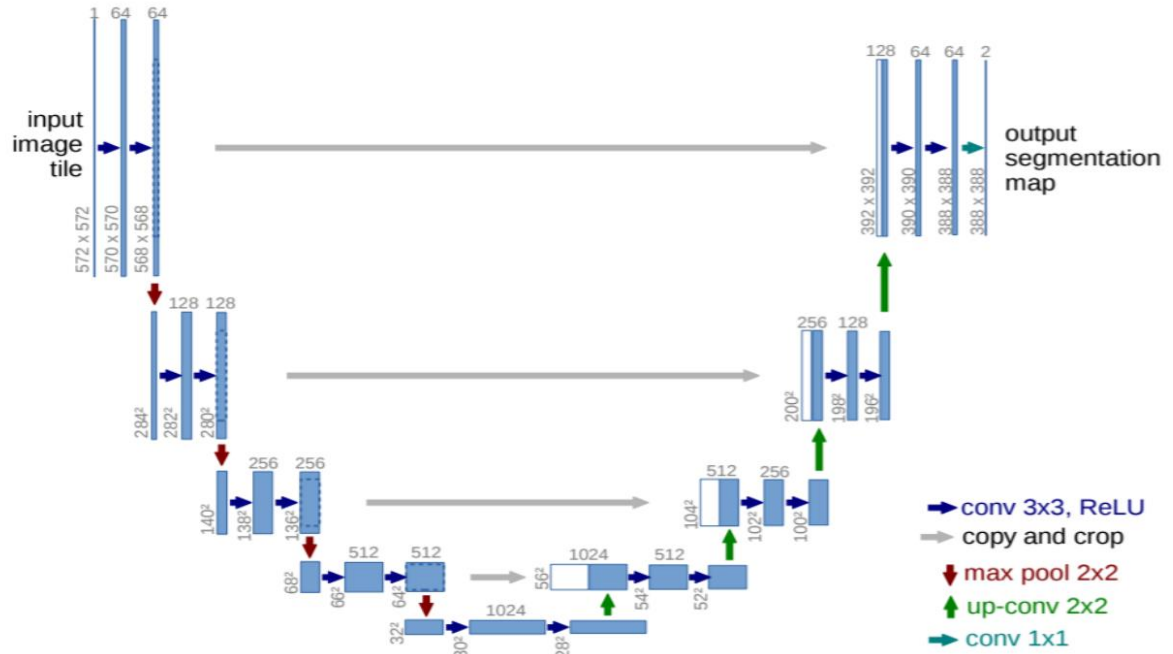
Tóm lại, DenseNet-121 có tổng cộng 120 lớp Convolution và 4 lớp AvgPool.

Tất cả các lớp, bao gồm cả trong cùng một dense block và transition layers, chia sẻ trọng số của mình qua nhiều đầu vào, cho phép các lớp sâu hơn sử dụng các đặc trưng được trích xuất từ sớm hơn trong quá trình huấn luyện.

Do các transition layer đầu ra nhiều đặc trưng trùng lặp, các lớp trong dense block thứ hai và thứ ba gán trọng số ít nhất cho đầu ra của các transition layer.

Ngoài ra, mặc dù trọng số của toàn bộ dense block được sử dụng bởi các lớp cuối cùng, vẫn có thể có nhiều đặc trưng cấp cao hơn được tạo ra sâu hơn vào mô hình vì có xu hướng tập trung cao hơn vào các bản đồ đặc trưng cuối cùng trong các thí nghiệm.

2.6.4. Kiến trúc mạng Unet



Hình 2. 7. Kiến trúc mạng Unet[30]

Mạng Unet [13] là một kiến trúc mạng nơ-ron tích chập (CNN) được sử dụng chủ yếu trong các nhiệm vụ phân đoạn hình ảnh (image segmentation). Kiến trúc của Unet được thiết kế để giải quyết vấn đề phân đoạn hình ảnh bằng cách kết hợp các phần thu hẹp (contraction) và mở rộng (expansion).

Phần thu hẹp của Unet, tương tự như một bộ mã hóa (encoder), được sử dụng để trích xuất đặc trưng và hiểu bối cảnh của hình ảnh. Nó bao gồm một mạng CNN sâu để trích xuất thông tin từ hình ảnh. Kích thước của các lớp (layers) trong phần này được giảm dần từ 572x572 xuống 32x32 và độ sâu tăng từ 3 lên 512. Quá trình này giúp mô hình nắm bắt các đặc trưng cấp cao và toàn cục của hình ảnh.

Phần mở rộng của Unet, tương tự như một bộ giải mã (decoder), giúp khôi phục lại kích thước ban đầu của ảnh và tạo ra một mặt nạ (mask) đánh dấu nhãn dự báo cho từng pixel. Nó bao gồm các lớp đối xứng với các lớp trong phần thu hẹp và sử dụng kỹ thuật

tăng kích thước (upsampling) để phóng to ảnh. Một đặc trưng đáng chú ý của Unet là sự áp dụng kết nối tắt (skip connections) đối xứng giữa các lớp trong phần thu hẹp và mở rộng. Điều này giúp truyền thông tin từ các lớp trước đến các lớp sau, cho phép mô hình tiếp cận cùng một lúc các đặc trưng cấp thấp và cấp cao.

Mặc dù mô hình Unet có độ chính xác cao trong việc phân đoạn hình ảnh, tuy nhiên, nó có tốc độ thấp và không phù hợp cho các ứng dụng yêu cầu xử lý thời gian thực như xe tự hành. Tuy nhiên, Unet vẫn được sử dụng rộng rãi trong các tác vụ không yêu cầu xử lý thời gian thực do độ chính xác cao và dễ dàng triển khai.

Trong phần mở rộng của Unet, có sử dụng up-conv (upsampling convolution) để tăng kích thước của ảnh. Điều này tương tự như max pooling layer trong quá trình thu hẹp, nhưng thay vì giảm kích thước, up-conv được sử dụng để tăng kích thước của ảnh. Các đường màu xám trong kiến trúc Unet thể hiện việc kết nối tắt giữa các lớp trước và sau. Điều này được sử dụng phổ biến trong các mạng CNN hiện đại như DenseNet để tránh hiện tượng gradient biến mất và truyền thông tin hiệu quả từ lớp trước đến lớp sau. Sử dụng các feature map tương tự như trong quá trình tạo vector giúp mở rộng vector thành hình ảnh được phân đoạn. Điều này giúp bảo vệ tính toàn vẹn cấu trúc của hình ảnh và giảm thiểu hiện tượng méo biến (distortion).

2.7. Phương pháp đánh giá

2.7.1. Đánh giá độ chính xác (Accuracy)

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.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FB + FN} = \frac{TP + TN}{\text{Số lượng mẫu}} \quad (1)[7]$$

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 “không” nhưng thực tế “có”.
- TN (True negative): là âm tính thật, mô hình dự đoán là “không” và thực tế là “không”.

Giả sử độ accuracy = 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 hiểm nhầm.

2.7.2. Chỉ số IoU

Intersection over Union (IoU) là tỷ lệ giữa phần giao và phần hợp của vùng dự đoán và vùng đối tượng thật.

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \quad (2)[13]$$

Trong đó:

- Area of Overlap là diện tích vùng giao nhau giữa predicted bounding box và ground truth.
- Area of Union bao gồm diện tích phần hợp của predicted bounding box và ground truth.

Kết quả của IoU là những giá trị trong khoảng (0,1) mỗi detection sẽ có một giá trị IoU riêng. Để xác định đó là wrong detection hay correct detection, chúng ta dựa vào một ngưỡng (threshold) cho trước (có thể là 0.5, 0.75, 0.95 tùy vào bài toán), nếu IoU lớn hơn hoặc bằng ngưỡng thì đó là correct detection, còn lại là wrong detection.

Dựa vào những khái niệm trên chúng ta định nghĩa True/false positive/negative:

- True Positive (TP): IoU lớn hơn hoặc bằng ngưỡng, là một correct detection.
- False Positive (FP): IoU bé hơn ngưỡng, là một wrong detection.
- False Negative (FN): ground truth không có predicted bounding box.

2.7.3. Chỉ số Dice coefficient

Chỉ số Dice coefficient (còn được gọi là Sorensen-Dice coefficient) là một phép đo sự tương đồng giữa hai tập hợp hoặc đối tượng dựa trên sự trùng lặp giữa chúng. Chỉ số này được sử dụng rộng rãi trong các bài toán liên quan đến xử lý ảnh, nhận dạng đối tượng, và phân đoạn hình ảnh.

Dice coefficient được tính bằng cách so sánh tập hợp các điểm hay pixel được phân loại là "positive" trong hai tập kết quả (predicted) và tập thực tế (ground truth). Công thức tính Dice coefficient như sau:

$$\text{Dice coefficient} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} \quad (3)[13]$$

Trong đó:

- TP (True Positives) là số lượng điểm được phân loại chính xác là "positive" trong cả tập predicted và ground truth.
- FP (False Positives) là số lượng điểm bị phân loại sai thành "positive" trong tập predicted.
- FN (False Negatives) là số lượng điểm bị phân loại sai thành "negative" trong tập predicted.

Giá trị của Dice coefficient nằm trong khoảng từ 0 đến 1, với 1 biểu thị sự trùng khớp hoàn toàn giữa hai tập, còn 0 biểu thị không có sự trùng lặp nào.

Dice coefficient được sử dụng phổ biến trong các bài toán phân đoạn hình ảnh, trong đó chúng ta muốn so sánh độ tương đồng giữa kết quả phân đoạn (predicted segmentation) và mẫu chuẩn (ground truth) để đánh giá chất lượng của mô hình.

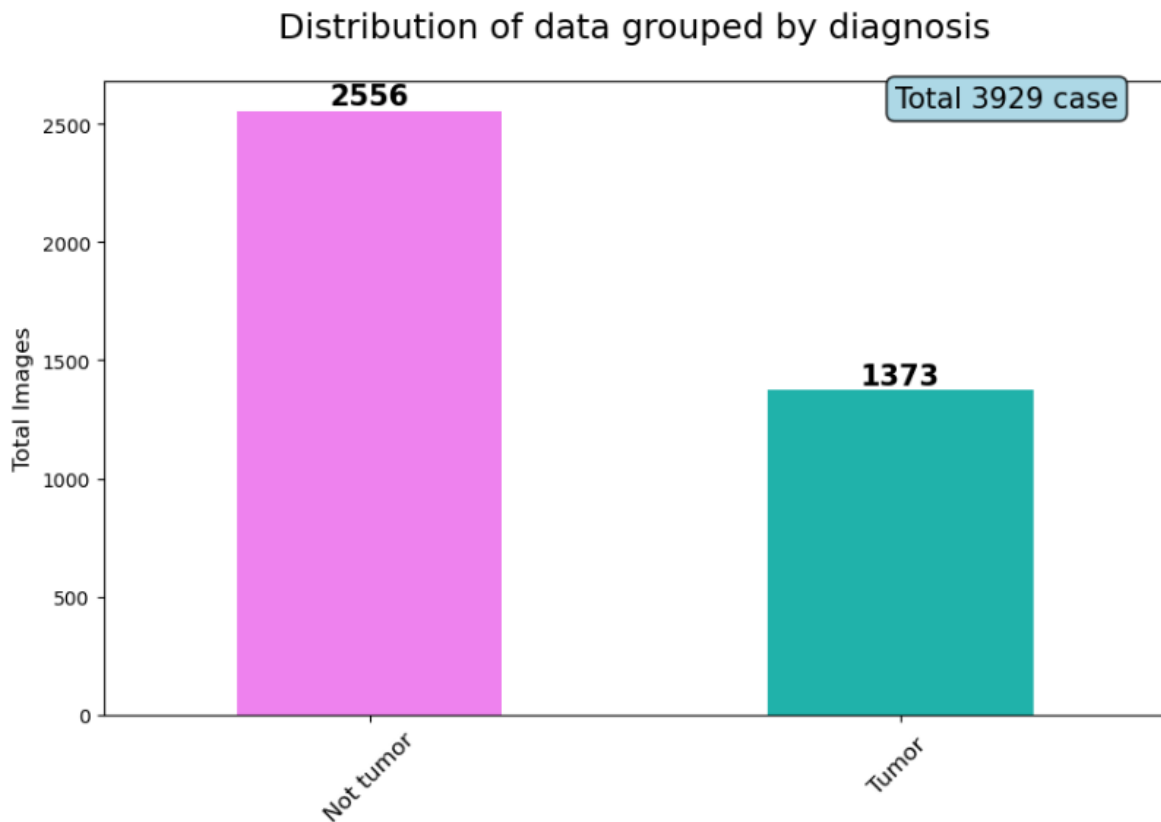
CHƯƠNG 3: PHƯƠNG PHÁP NGHIÊN CỨU

3.1. Tập dữ liệu

Bộ dữ liệu này được thu thập từ The Cancer Imaging Archive (TCIA) và bao gồm các hình ảnh MRI não từ 110 bệnh nhân được bao gồm trong bộ sưu tập glioma cấp độ thấp của The Cancer Genome Atlas (TCGA). Những bệnh nhân này có ít nhất một chuỗi FLAIR (fluid-attenuated inversion recovery) và dữ liệu về phân loại di truyền khối u.

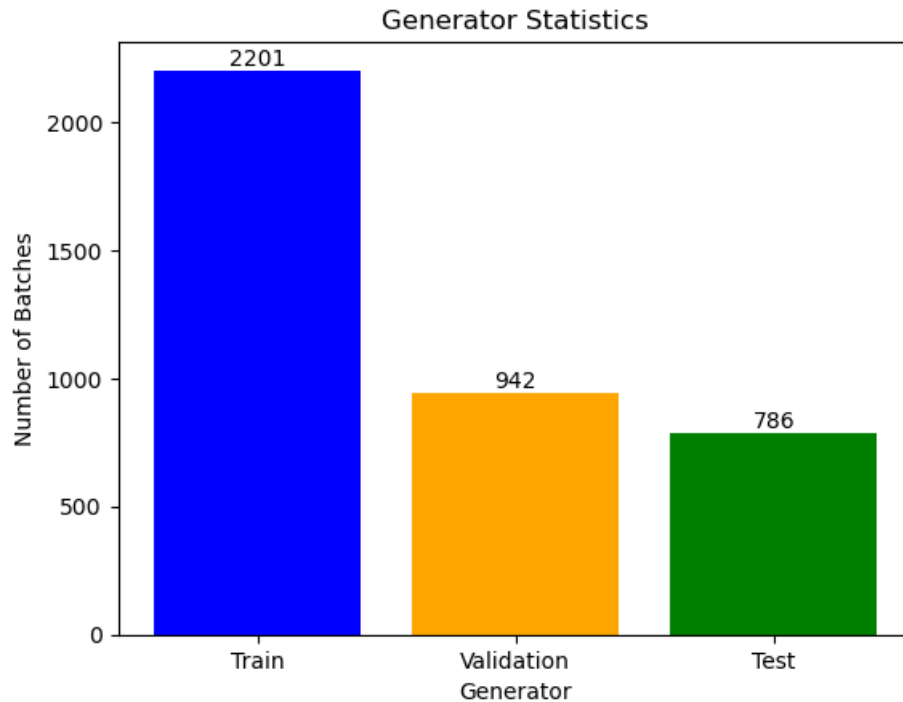
Bộ dữ liệu tên: Brain MRI segmentation[30], được chia sẻ công khai trên kaggle.

Có 7,858 ảnh, gồm 3929 trường hợp, 2556 trường hợp không bệnh và 1373 trường hợp bệnh.



Hình 3. 1. Phân phối dữ liệu được nhóm theo chẩn đoán

Chia ra 2201 trường hợp train, 942 trường hợp Validation, 786 trường hợp test.



Hình 3. 2. Chia dữ liệu train, Validation, test

Bộ dữ liệu được sử dụng trong: Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm [16] của Mateusz Buda, Ashirbani Saha, Maciej A. Mazurowski và Radiogenomics of lower-grade glioma: algorithmically-assessed tumor shape is associated with tumor genomic subtypes and patient outcomes in a multi-institutional study with The Cancer Genome Atlas data [17] của Maciej A. Mazurowski, Kal Clark, Nicholas M. Czarnek, Parisa Shamsesfandabadi, Katherine B. Peters, Ashirbani Saha.

3.2 Môi trường

Trong quá trình thực hiện đề tài, em thực hiện thực nghiệm trên 6 kịch bản cho 6 mạng như sau:

- Kịch bản 1: Xử lý mạng ResNet50 để phân loại trong môi trường máy tính đơn trên nền tảng cung cấp sẵn có của Kaggle với CPU Intel Xeon 2.20 GHz đi kèm với 13GB ram cho phép huấn luyện với batch size lớn hơn, ngoài ra việc cung cấp ngẫu nhiên các GPU có hiệu suất cao như GPU NVIDIA Tesla P100, NVIDIA T4 x2 giúp rút ngắn đáng kể thời gian tính toán.

- Kịch bản 2: Xử lý các mạng VGG19 để phân loại trong môi trường trên.

- Kịch bản 3: Xử lý các mạng Densenet121 để phân loại trong môi trường trên.
- Kịch bản 4: Xử lý các mạng ResNet50-Unet để phân đoạn trong môi trường trên.
- Kịch bản 5: Xử lý các mạng VGG19 -Unet để phân đoạn trong môi trường trên.
- Kịch bản 6: Xử lý các mạng Densenet121-Unet để phân đoạn trong môi trường trên.

Bảng 3. 1 Thông số môi trường huấn luyện

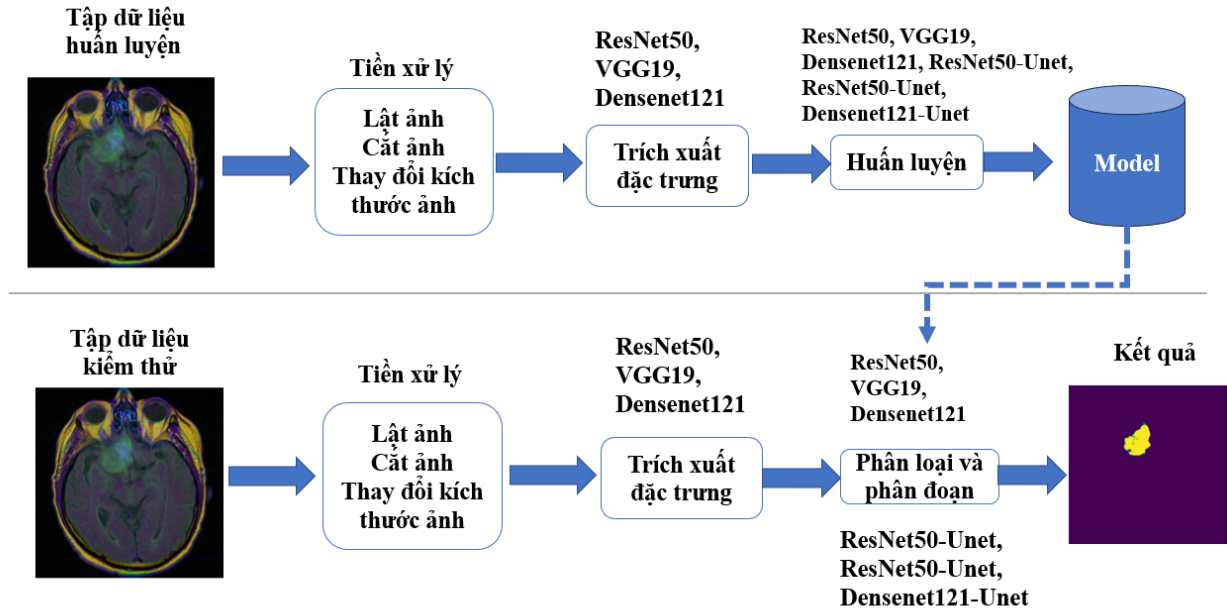
Tên môi trường	Cấu hình	Thư viện
Kaggle	CPU: Intel Xeon 2.20 GHz GPU: NVIDIA Tesla P100 RAM: 13GB	Tensorflow Kares

3.3. Phương pháp đề xuất

Để nghiên cứu và xây dựng một hệ thống phát hiện khối u não, chúng tôi lựa chọn và kết hợp các kiến trúc mạng nổi tiếng như Unet, VGG19, ResNet50, Densenet121. Mục tiêu của hệ thống này là cung cấp một phương pháp chẩn đoán nhanh chóng và chính xác cho việc phát hiện khối u não, từ đó giúp giảm áp lực đối với các bác sĩ, phát hiện sớm căn bệnh và cung cấp điều trị kịp thời để tránh những tác động nguy hiểm tính mạng cho bệnh nhân.

Mô hình hệ thống đề xuất của chúng tôi được xây dựng dựa trên các kiến trúc mạng Unet, VGG19, ResNet50, Densenet, với mục đích tối ưu hóa khả năng phát hiện và phân đoạn khối u trong hình ảnh MRI của não. Các kiến trúc mạng này đã được chứng minh là hiệu quả trong việc xử lý hình ảnh y khoa và có khả năng học và tổng hợp các đặc trưng quan trọng.

Bằng cách kết hợp những kiến trúc mạng này, hệ thống đề xuất có khả năng xử lý hình ảnh MRI của não và phát hiện khối u với độ chính xác cao. Hệ thống này có thể giúp các bác sĩ và chuyên gia y tế trong việc chẩn đoán và điều trị khối u não một cách hiệu quả, đồng thời đảm bảo sự an toàn và tối ưu hóa quy trình chăm sóc sức khỏe cho bệnh nhân.



Hình 3. 3. Mô hình đề xuất

Quá trình huấn luyện:

Sau khi tiến hành tiền xử lý dữ liệu, tập dữ liệu huấn luyện sẽ được nhập vào mô hình mạng để thực hiện quá trình huấn luyện trên nền tảng Kaggle, sử dụng một gói thư viện hỗ trợ huấn luyện.

Các bước trong quá trình huấn luyện như sau:

1. Tiền xử lý dữ liệu huấn luyện: Đây là giai đoạn thực hiện các thao tác như dịch ảnh, zoom ảnh, và lật ảnh để tạo ra sự đa dạng cho các mẫu dữ liệu ảnh.

2. Trích xuất đặc trưng bằng cách sử dụng các kiến trúc tích chập và pooling: Ở bước này, mô hình sẽ sử dụng các lớp tích chập và lớp pooling để trích xuất những đặc trưng quan trọng từ dữ liệu huấn luyện.

3. Huấn luyện các tham số của mô hình: Sau khi đã trích xuất được các đặc trưng từ dữ liệu huấn luyện, mô hình sẽ được huấn luyện bằng cách điều chỉnh các tham số để tối ưu hóa việc phân lớp.

Kết quả của quá trình huấn luyện là một mô hình phân lớp đã được huấn luyện.

Quá trình nhận dạng/chẩn đoán:

Quá trình nhận dạng/chẩn đoán được thực hiện như sau:

1. Tiền xử lý dữ liệu ảnh cần nhận dạng: Ảnh cần nhận dạng sẽ được tiền xử lý bằng cách thay đổi kích thước của nó để có cùng kích thước với các ảnh trong tập dữ liệu huấn luyện. Điều này giúp đồng nhất kích thước và đảm bảo mô hình có thể xử lý ảnh đầu vào một cách chính xác.

2. Phân lớp/chẩn đoán dựa vào mô hình đã được huấn luyện: Sau khi đã tiền xử lý ảnh, mô hình phân lớp/chẩn đoán được áp dụng lên ảnh để đưa ra kết quả phân lớp/chẩn đoán dựa trên những gì đã học từ quá trình huấn luyện.

Kết quả cuối cùng của quá trình nhận dạng/chẩn đoán là kết quả phân lớp/chẩn đoán cho ảnh đầu vào.

3.4. Kịch bản huấn luyện

Bảng 3. 2 bảng kịch bản huấn luyện

Stt	Mạng sử dụng	Mạng rút trích đặc trưng	Số epoch	Batch size
1	ResNet50	-	100	16
2	VGG19	-	100	16
3	Densenet121	-	100	16
4	UNet	ResNet50	100	16
5	UNet	VGG19	100	16
6	UNet	Densenet121	100	16

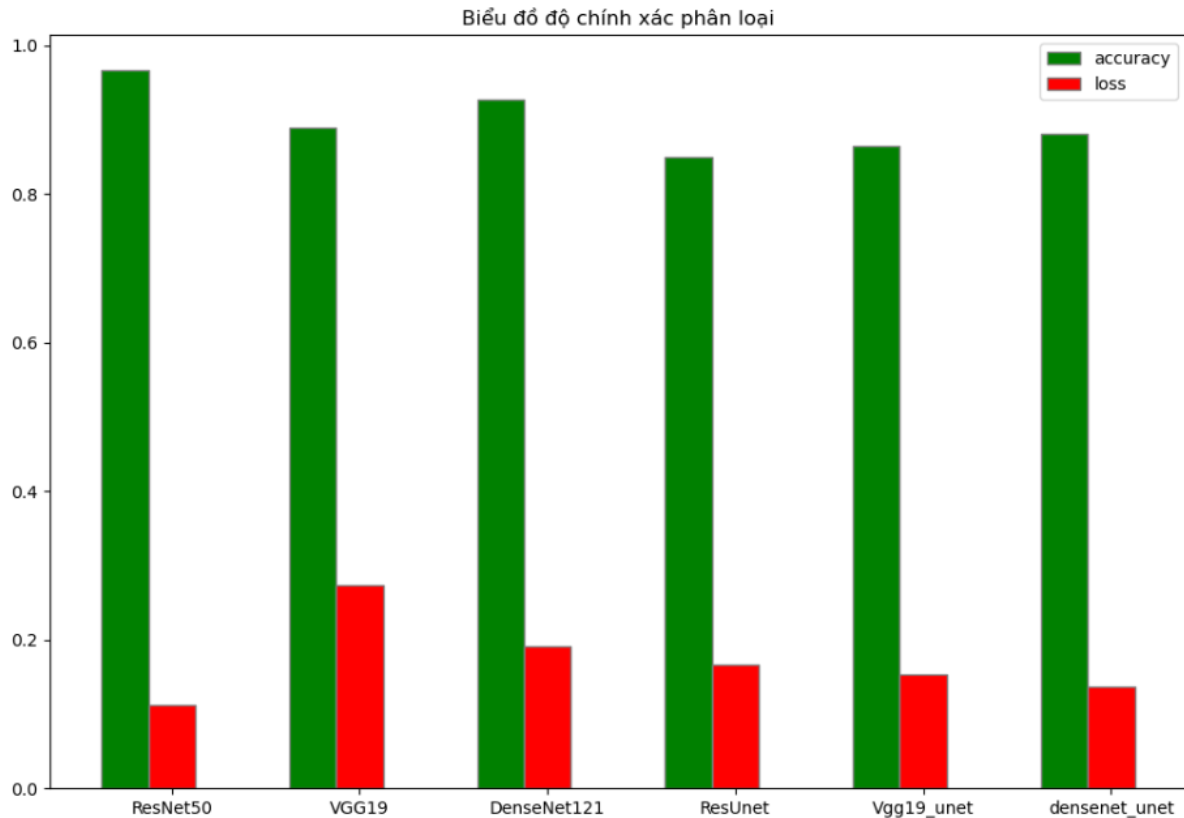
CHƯƠNG 4: KẾT QUẢ

4.1. Kết quả tham số huấn luyện

Bảng 4. 1 Kịch bản mô hình phân loại

Kịch bản	Mạng huấn luyện	Num step (epoch)		Thời gian mỗi epoch	Accuracy		Loss
1	ResNet50	100		69 giây	0.97		0.11
2	VGG19	100		67 giây	0.89		0.27
3	Densenet121	100		66 giây	0.92		0.19
Kịch bản	Mạng phân vùng	Mạng rút trích đặc trưng	Num step (epoch)	Thời gian mỗi epoch	Dice Coefficient	IoU	Loss
1	UNet	ResNet50	100	59 giây	0.85	0.75	0.17
2	UNet	VGG19	100	82 giây	0.86	0.77	0.15
3	UNet	Densenet121	100	76 giây	0.88	0.79	0.14

So sánh kết quả chính xác và mất mát các mô hình ta được biểu đồ:



Hình 4. 1. Biểu đồ so sánh kết quả các mô hình

Các mô hình đều đã được huấn luyện qua 100 bước (epoch). Thời gian huấn luyện cho mỗi epoch của từng mô hình cũng đã được ghi lại. Độ chính xác (Accuracy) của mô hình là một chỉ số đánh giá hiệu suất, trong khi hàm mất mát (Loss) đo lường sự sai khác giữa đầu ra dự đoán và giá trị thực tế.

ResNet50 có độ chính xác cao nhất với 0.97 và hàm mất mát thấp nhất với 0.11, cho thấy hiệu suất tốt của mô hình trong việc phân loại. VGG19 đạt được độ chính xác 0.89 và hàm mất mát 0.27, trong khi Densenet121 đạt được độ chính xác 0.92 và hàm mất mát 0.19.

Các mạng phân vùng sử dụng kiến trúc UNet cùng với các mạng rút trích đặc trưng khác nhau để thực hiện phân vùng. Thời gian huấn luyện cho mỗi epoch của từng mạng phân vùng cũng đã được ghi lại. Dice Coefficient và IoU đo lường mức độ chính xác của phân vùng, trong khi hàm mất mát đo lường sự sai khác giữa phân vùng dự đoán và phân vùng thực tế.

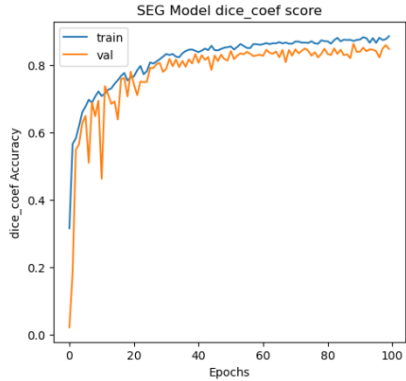
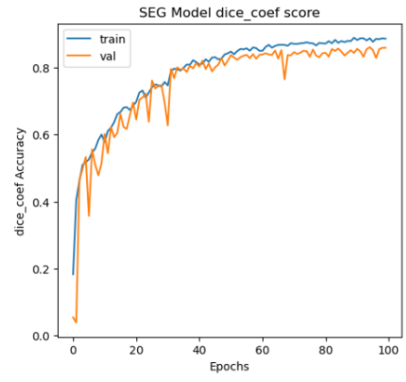
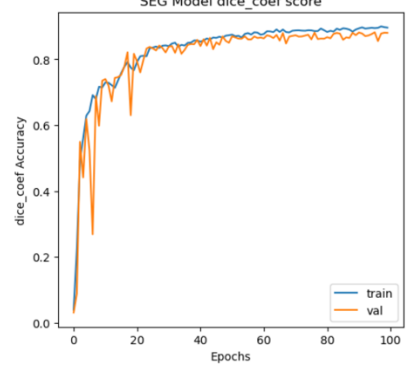
Các kết quả cho thấy UNet với mạng rút trích đặc trưng Densenet121 đạt được kết quả tốt nhất với Dice Coefficient là 0.88, IoU là 0.79 và hàm mất mát là 0.14. UNet với VGG19 và ResNet50 cũng đạt được kết quả tương đối tốt với Dice Coefficient trên 0.85 và IoU trên 0.75.

4.2. Kết quả huấn luyện

4.2.1. Độ chính xác

Bảng 4. 2 bảng chứa các biểu đồ Độ chính xác

Stt	Mô hình mạng	Biểu đồ độ chính xác
1	ResNet50	
2	VGG19	
3	DenseNet121	

4	ResNet50-Unet	
5	VGG19-Unet	
6	DenseNet121-Unet	

Kịch bản 1:

Sau khi huấn luyện với tập dữ liệu qua 100 epoch thì ta có kết quả như sau: Accuracy của mô hình mạng ResNet50 cho ta thấy được độ chính xác của mô hình đạt đến 96,56%.

Kịch bản 2:

Sau khi huấn luyện với tập dữ liệu qua 100 epoch thì ta có kết quả như sau: Accuracy của mô hình mạng VGG19 cho ta thấy được độ chính xác của mô hình đạt đến 88,93%.

Kịch bản 3:

Sau khi huấn luyện với tập dữ liệu qua 100 epoch thì ta có kết quả như sau: Accuracy của mô hình mạng Densenet121 cho ta thấy được độ chính xác của mô hình đạt đến 92,62%.

Kịch bản 4:

Sau khi huấn luyện với tập dữ liệu qua 100 epoch thì ta có kết quả như sau: Dice coef của mô hình mạng ResNet50-Unet cho ta thấy được độ chính xác của mô hình đạt đến 85.00%.

Kịch bản 5:

Sau khi huấn luyện với tập dữ liệu qua 100 epoch thì ta có kết quả như sau: Dice coef của mô hình mạng VGG19-Unet cho ta thấy được độ chính xác của mô hình đạt đến 86.4%.

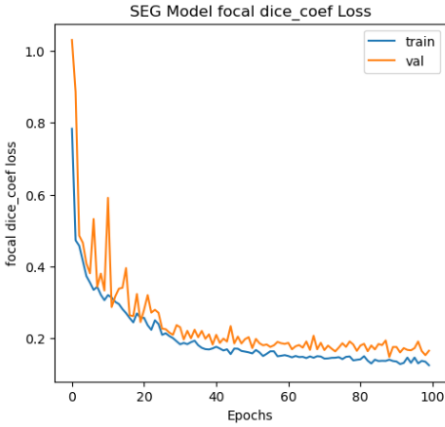
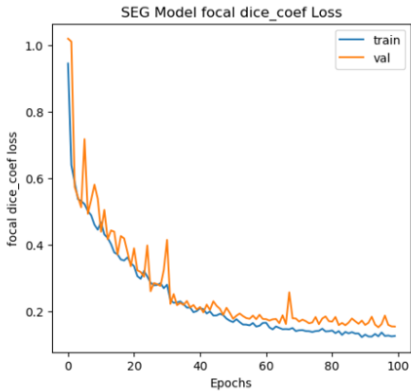
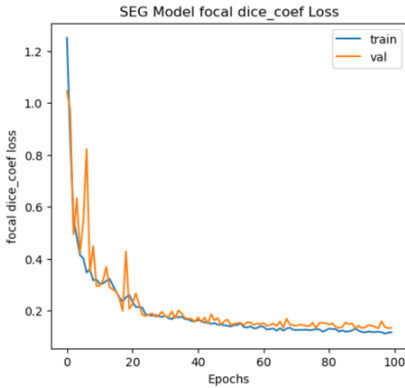
Kịch bản 6:

Sau khi huấn luyện với tập dữ liệu qua 100 epoch thì ta có kết quả như sau: Dice coef của mô hình mạng DenseNet121-Unet cho ta thấy được độ chính xác của mô hình đạt đến 88.06%..

4.2.2. Độ mất mát

Bảng 4. 3 Bảng chứa các biểu đồ độ mất mát

Stt	Mô hình mạng	Biểu đồ độ mất mát
1	ResNet50	<p>Classification Model LOSS</p> <p>loss</p> <p>train</p> <p>val</p>
2	VGG19	<p>Classification Model LOSS</p> <p>loss</p> <p>train</p> <p>val</p> <p>Epochs</p>
3	DenseNet121	<p>Classification Model LOSS</p> <p>loss</p> <p>train</p> <p>val</p> <p>Epochs</p>

4	ResNet50-Unet	 <p>SEG Model focal dice_coef Loss</p> <p>Y-axis: focal dice_coef loss (0.2 to 1.0)</p> <p>X-axis: Epochs (0 to 100)</p> <p>Legend: train (blue line), val (orange line)</p>
5	VGG19-Unet	 <p>SEG Model focal dice_coef Loss</p> <p>Y-axis: focal dice_coef loss (0.2 to 1.0)</p> <p>X-axis: Epochs (0 to 100)</p> <p>Legend: train (blue line), val (orange line)</p>
6	DenseNet121-Unet	 <p>SEG Model focal dice_coef Loss</p> <p>Y-axis: focal dice_coef loss (0.2 to 1.2)</p> <p>X-axis: Epochs (0 to 100)</p> <p>Legend: train (blue line), val (orange line)</p>

Kịch bản 1:

Sau khi huấn luyện với tập dữ liệu qua 100 epoch thì ta có kết quả như sau: Loss của mô hình mạng ResNet50 cho ta thấy được sự mất mát của mô hình chỉ khoảng 11,15%.

Kịch bản 2:

Sau khi huấn luyện với tập dữ liệu qua 100 epoch thì ta có kết quả như sau: Loss của mô hình mạng VGG19 cho ta thấy được sự mất mát của mô hình chỉ khoảng 27,29%.

Kịch bản 3:

Sau khi huấn luyện với tập dữ liệu qua 100 epoch thì ta có kết quả như sau: Loss của mô hình mạng Densenet cho ta thấy được sự mất mát của mô hình chỉ khoảng 19,17%.

Kịch bản 4:

Sau khi huấn luyện với tập dữ liệu qua 100 epoch thì ta có kết quả như sau: Loss của mô hình mạng ResNet-Unet cho ta thấy được sự mất mát của mô hình chỉ khoảng 16.61%.

Kịch bản 5:

Sau khi huấn luyện với tập dữ liệu qua 100 epoch thì ta có kết quả như sau: Loss của mô hình mạng VGG19-Unet cho ta thấy được sự mất mát của mô hình chỉ khoảng 15.27%.

Kịch bản 6:

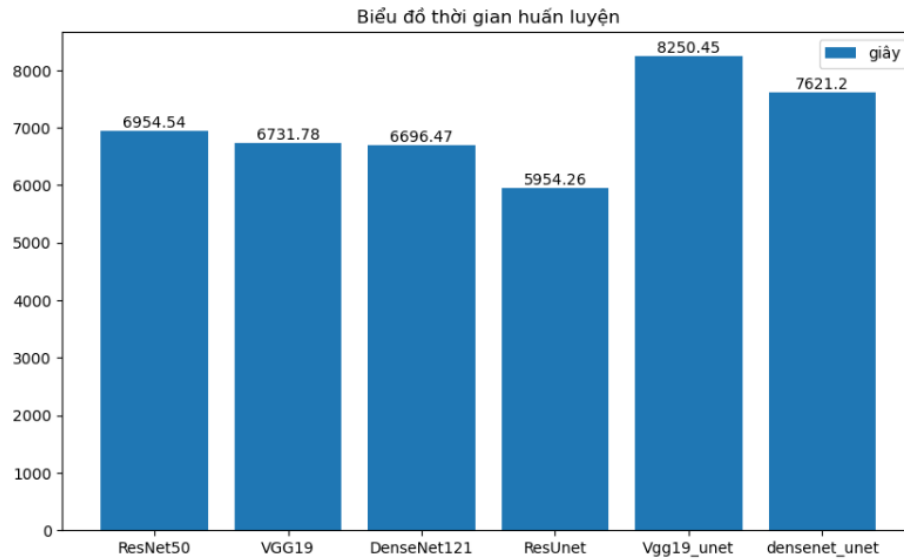
Sau khi huấn luyện với tập dữ liệu qua 100 epoch thì ta có kết quả như sau: Dice coef Loss của mô hình mạng DenseNet121-Unet cho ta thấy được sự mất mát của mô hình chỉ khoảng 13.63%.

4.2.3. Thời gian huấn luyện

Bảng 4. 4 Thời gian huấn luyện

Kịch bản	Mô hình	Thời gian huấn luyện
1	ResNet50	6954.54 giây (1.93 giờ)
2	VGG19	6731.78 giây (1.87 giờ)
3	Densenet121	6696.47 giây (1.86 giờ)
4	ResNet UNet	5954.26 giây (1.65 giờ)
5	VGG19 UNet	8250.45 giây (2.29 giờ)
6	Densenet121 UNet	7621.20 giây (2.11 giờ)

Khi thực nghiệm so sánh về thời gian huấn luyện mô hình mạng có sự chênh lệch về thời gian huấn luyện do mỗi mô hình có mỗi thông số và kiến trúc khác nhau.



Hình 4. 2. Biểu đồ thời gian huấn luyện

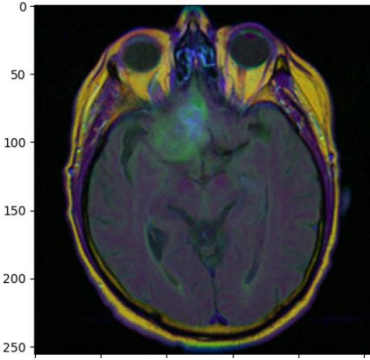
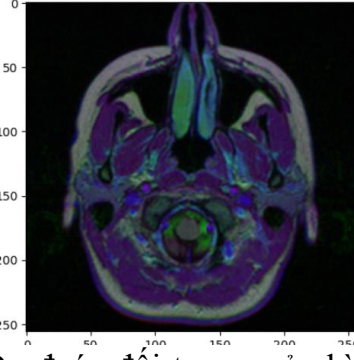
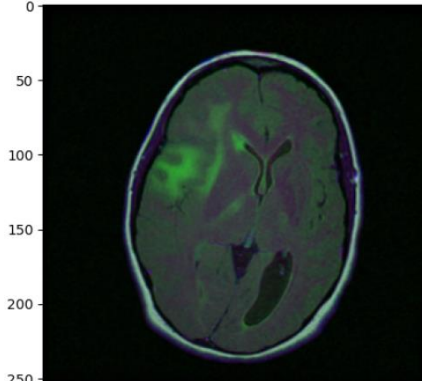
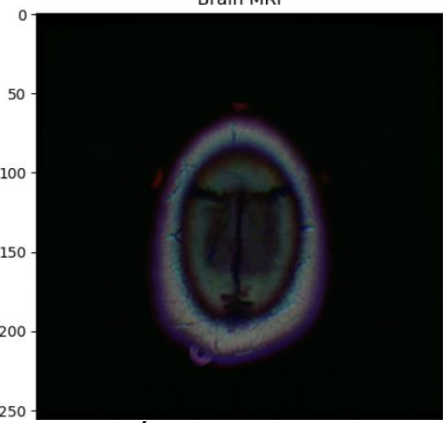
Thời gian huấn luyện là một yếu tố quan trọng trong việc xây dựng mô hình học máy. Trong danh sách các mô hình được liệt kê, có thể thấy rằng các mô hình phức tạp và sâu hơn như ResNet50, VGG19 và Densenet121 yêu cầu nhiều thời gian huấn luyện hơn so với các mô hình kết hợp UNet.

Với ResNet50, thời gian huấn luyện là 6954.54 giây (1.93 giờ), trong khi VGG19 và Densenet121 lần lượt mất 6731.78 giây (1.87 giờ) và 6696.47 giây (1.86 giờ). Đây là những mô hình có hiệu suất tốt và đạt được kết quả ấn tượng trong nhiều tác vụ như phân loại hình ảnh. Tuy nhiên, việc huấn luyện chúng mất nhiều thời gian do độ phức tạp và số lượng lớp.

Trong khi đó, các mô hình kết hợp UNet như ResNet UNet, VGG19 UNet và Densenet121 UNet có thời gian huấn luyện thấp hơn. ResNet UNet mất 5954.26 giây (1.65 giờ), VGG19 UNet mất 8250.45 giây (2.29 giờ) và Densenet121 UNet mất 7621.20 giây (2.11 giờ). Sự kết hợp giữa các mô hình UNet và mô hình gốc giúp giảm thiểu thời gian huấn luyện mà vẫn đảm bảo hiệu suất và độ chính xác của mô hình.

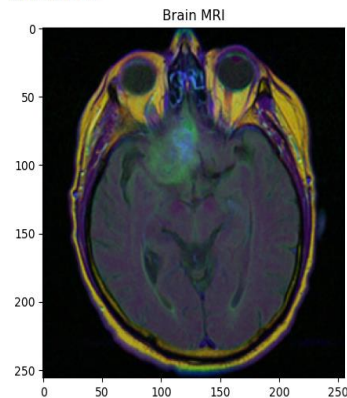
4.3. KẾT QUẢ KIỂM THỬ

Đánh giá này chỉ áp dụng cho hình ảnh đã được kiểm thử và không đại diện cho tất cả các trường hợp.

Phân loại		
STT	Tumor	Not tumor
Kịch bản 1	<p>Predictions: [[0.01654099 0.983459]] Độ chính xác: (98.346 %) Dự đoán: Tumor ResNet50 Prediction</p>  <p>Dự đoán đối tượng của hình ảnh là "Tumor" với độ chính xác 98.346%. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân loại chính xác hình ảnh. Tuy nhiên, cần chú ý rằng có 1.654% khả năng đó là một lớp khác.</p>	<p>Predictions: [[9.9942744e-01 5.7253568e-04]] Độ chính xác: (99.943 %) Dự đoán: Not Tumor ResNet50 Prediction</p>  <p>Dự đoán đối tượng của hình ảnh là "Not Tumor" với độ chính xác 99.943%. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân loại chính xác hình ảnh. Tuy nhiên, cần chú ý rằng có 0.057% khả năng đó là một lớp khác.</p>
	<p>Predictions: [[3.6585846e-04 9.9963415e-01]] Độ chính xác: (99.963 %) Dự đoán: Tumor VGG19 Prediction</p>  <p>Dự đoán đối tượng của hình ảnh là "Tumor" với độ chính xác 99.963%. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân loại chính xác hình ảnh. Tuy nhiên, cần chú ý rằng có 0.037% khả năng đó là một lớp khác.</p>	<p>Predictions: [[0.50534517 0.4946549]] Độ chính xác: (50.535 %) Dự đoán: Not Tumor VGG19 Prediction</p>  <p>Dự đoán đối tượng của hình ảnh là "Not Tumor" với độ chính xác 50.535%. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân loại chính xác hình ảnh. Tuy nhiên, cần chú ý rằng có 49.465% khả năng đó là một lớp khác.</p>

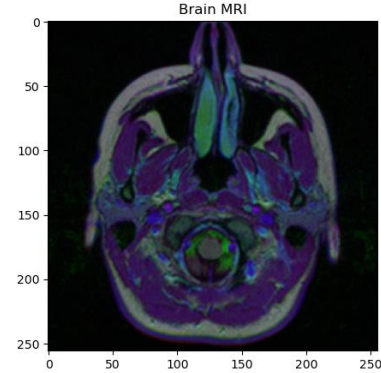
Kịch
bản 2

Predictions: [[0.4989203 0.5010797]]
Độ chính xác: (50.108 %)
Dự đoán: Tumor
VGG19 Prediction



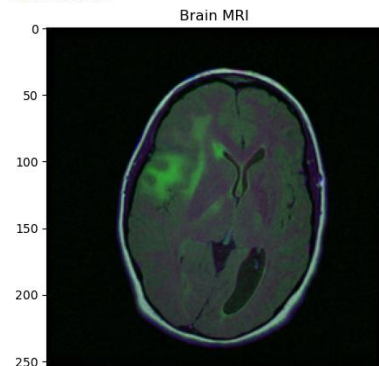
Dự đoán đối tượng của hình ảnh là "Tumor" với độ chính xác 50.108%. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân loại chính xác hình ảnh. Tuy nhiên, cần chú ý rằng có 49.892% khả năng đó là một lớp khác.

Predictions: [[0.9876665 0.0123333]]
Độ chính xác: (98.767 %)
Dự đoán: Not Tumor
VGG19 Prediction



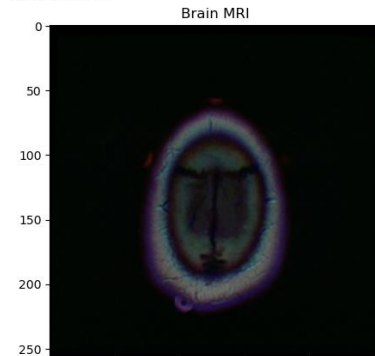
Dự đoán đối tượng của hình ảnh là "Not Tumor" với độ chính xác 98.767%. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân loại chính xác hình ảnh. Tuy nhiên, cần chú ý rằng có 1.233% khả năng đó là một lớp khác.

Predictions: [[0.12284943 0.87715054]]
Độ chính xác: (87.715 %)
Dự đoán: Tumor
VGG19 Prediction



Dự đoán đối tượng của hình ảnh là "Tumor" với độ chính xác 99.963%. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân loại chính xác hình ảnh. Tuy nhiên, cần chú ý rằng có 0.037% khả năng đó là một lớp khác.

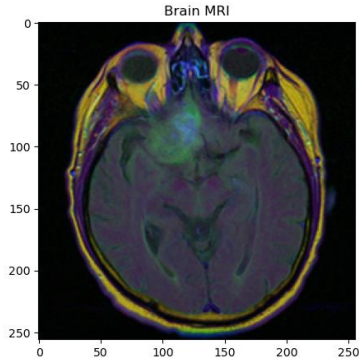
Predictions: [[0.766242 0.2337579]]
Độ chính xác: (76.624 %)
Dự đoán: Not Tumor
VGG19 Prediction



Dự đoán đối tượng của hình ảnh là "Not Tumor" với độ chính xác 50.535%. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân loại chính xác hình ảnh. Tuy nhiên, cần chú ý rằng có 49.465% khả năng đó là một lớp khác.

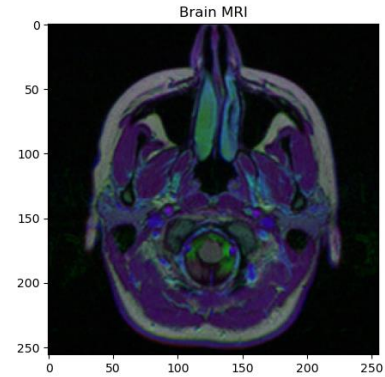
Kịch
bản 3

Predictions: [[0.00148213 0.99851793]]
Độ chính xác: (99.852 %)
Dự đoán: Tumor
Densenet Prediction



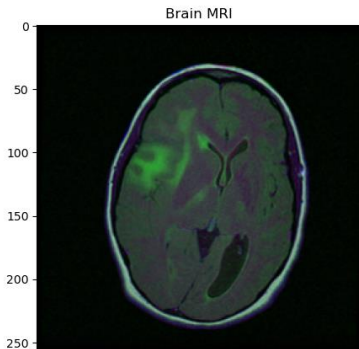
Dự đoán đối tượng của hình ảnh là "Tumor" với độ chính xác 99.852%. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân loại chính xác hình ảnh. Tuy nhiên, cần chú ý rằng có 0.148% khả năng đó là một lớp khác.

Predictions: [[9.999186e-01 8.142594e-05]]
Độ chính xác: (99.992 %)
Dự đoán: Not Tumor
Densenet Prediction



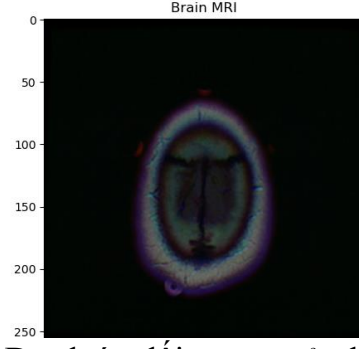
Dự đoán đối tượng của hình ảnh là "Not Tumor" với độ chính xác 99.992%. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân loại chính xác hình ảnh. Tuy nhiên, cần chú ý rằng có 0.008% khả năng đó là một lớp khác.

Predictions: [[0.01719379 0.9828062]]
Độ chính xác: (98.281 %)
Dự đoán: Tumor
VGG19 Prediction



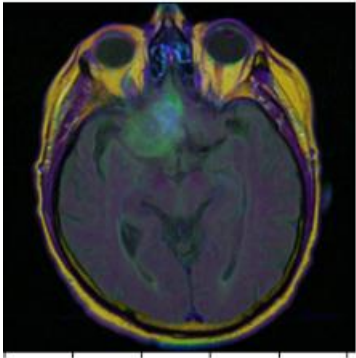


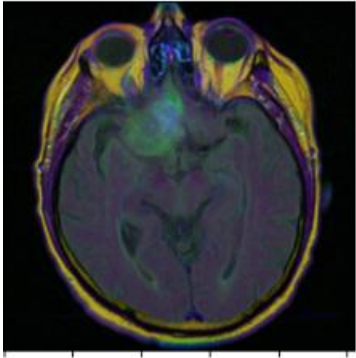


Dự đoán đối tượng của hình ảnh là "Tumor" với độ chính xác 98.281%. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân loại chính xác hình ảnh. Tuy nhiên, cần chú ý rằng có 1.719% khả năng đó là một lớp khác.

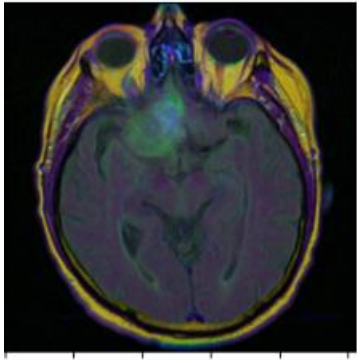

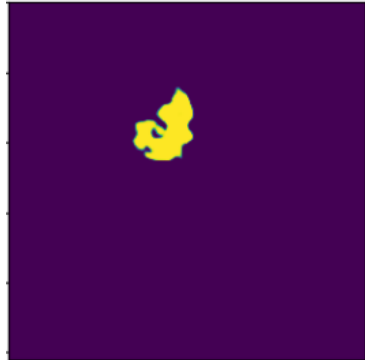
Predictions: [[0.99091107 0.009089]]
Độ chính xác: (99.091 %)
Dự đoán: Not Tumor
VGG19 Prediction



Dự đoán đối tượng của hình ảnh là "Not Tumor" với độ chính xác 99.091%. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân loại chính xác hình ảnh. Tuy nhiên, cần chú ý rằng có 0.909% khả năng đó là một lớp khác.

Bảng 4. 5 Bảng chứa kết quả kiểm thử phân loại

Phân vùng			
STT	Ảnh gốc	Mặt nạ	Mặt nạ dự đoán
Kịch bản 4			 <p>Dice Coefficient đạt giá trị 0.92. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân đoạn một cách chính xác và tương đồng với mặt nạ phân đoạn thực tế.</p>
Kịch bản 5			 <p>Dice Coefficient đạt giá trị 0.93. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân đoạn một cách chính xác và tương đồng với mặt nạ phân đoạn thực tế.</p>

Kịch bản 6	<p>Brain MRI</p> 	<p>Mask</p> 	<p>Pred-Mask</p>  <p>Dice Coefficient đạt giá trị 0.90. Đây là một kết quả rất cao, cho thấy mô hình có khả năng phân đoạn một cách chính xác và tương đồng với mặt nạ phân đoạn thực tế.</p>
------------	--	--	--

Bảng 4. 6 Bảng chứa kết quả kiểm thử phân đoạn

Nhận xét

Trong việc các model phân loại: Sau khi huấn luyện ta rút ra được mô hình mạng ResNet50 là mô hình mạng đạt được tham số Acc cao nhất (0.97), Loss thấp nhất (0.11). Bất quá xét về thời gian thì mô hình ResNet50 dùng nhiều thời gian nhất (6954.54 giây) để huấn luyện. Nhưng ResNet50 có tham số Acc cao nhất, Loss thấp nhất nên việc sử dụng mạng ResNet50 để tiếp tục nghiên cứu và phát triển hệ thống.

Trong việc các model phân vùng: Sau khi huấn luyện ta rút ra được mô hình mạng Densenet121-Unet là mô hình mạng đạt được tham số Dice coef cao nhất, Loss thấp nhất và thời gian huấn luyện thứ hai. Nên việc sử dụng mạng Densenet121-Unet để tiếp tục nghiên cứu và phát triển hệ thống.

KẾT LUẬN

Kết luận

Sau một khoảng thời gian nghiên cứu và thực hiện, đề tài quản lý "Ứng dụng mô hình mạng U-Net trong phân đoạn khối u não trên ảnh CT/MRI" đã được hoàn thành theo yêu cầu và nội dung đã đề ra. Trong quá trình thực hiện đề tài, dù có quyết tâm cao nhưng do hạn chế về kinh nghiệm và kiến thức, đề tài này chắc chắn không tránh khỏi những thiếu sót. Mong muốn được nhận sự đóng góp từ các thầy cô và các bạn để nâng cao chất lượng và hoàn thiện đề tài.

Trong quá trình nghiên cứu và thực hiện đề tài, tôi đã thực hiện một số đóng góp sau:

- Thu thập dữ liệu về các khối u não trên ảnh CT/MRI.
- Tiền xử lý ảnh bằng cách lật ảnh dọc, lật ảnh ngang, xoay ảnh, cắt ảnh để tăng tính chính xác cho mô hình.
- Cài đặt và sử dụng môi trường và các thư viện hỗ trợ để huấn luyện và kiểm thử.
- Thực hiện các mạng rút trích đặc trưng, phân đoạn và phân lớp sử dụng mô hình Desenet121, VGG-19, ResNet50, ResNet50-Unet, VGG19-Unet, Desenet121-Unet.
- Cung cấp kết quả thực hiện, đánh giá độ chính xác, độ mất mát và thời gian huấn luyện của các mô hình.
- So sánh và đánh giá các mô hình trong quá trình thực hiện.

Tuy nhiên, đề tài cũng gặp một số hạn chế. Do kiến thức còn hạn hẹp, việc thực hiện đề tài vẫn còn thiếu một số ý kiến và không thể đạt được tối đa các mục tiêu đã đề ra.

Hướng phát triển

Sau quá trình nghiên cứu và thực hiện đề tài về Ứng dụng mô hình mạng U-NET trong phân đoạn khối u não trên ảnh CT/MRI, tôi nhận thấy còn nhiều khía cạnh cần được phát triển và hoàn thiện. Để đạt được một kết quả tốt hơn, tôi đề ra một số hướng phát triển trong tương lai.

Đầu tiên, tôi sẽ tiếp tục thu thập và xây dựng các tập dữ liệu mới hơn để nâng cao độ đa dạng và độ phong phú của dữ liệu. Việc này sẽ giúp mô hình mạng U-NET học được nhiều đặc trưng và đánh giá chính xác hơn về phân đoạn khối u não trên ảnh CT/MRI.

Tiếp theo, tôi sẽ thử nghiệm và sử dụng các mô hình mạng khác nhau để so sánh kết quả và hiệu suất của mô hình U-NET. Bằng cách so sánh và đánh giá các mô hình khác nhau, tôi sẽ có cái nhìn rõ hơn về sự hiệu quả của mô hình U-NET trong ứng dụng phân đoạn khối u não trên ảnh CT/MRI.

Một yếu tố quan trọng khác cần được cải thiện là thời gian huấn luyện mô hình. Tôi dự định sử dụng và áp dụng kỹ thuật huấn luyện và tính toán mô hình song song trong môi trường Apache Spark. Điều này sẽ giúp tăng tốc quá trình huấn luyện mô hình và nâng cao hiệu suất của ứng dụng.

Ngoài ra, tôi cũng sẽ tìm hiểu và áp dụng các kỹ thuật và công nghệ mới nhất trong lĩnh vực phân đoạn ảnh như mô hình U-NET có sử dụng các kiến trúc mạng nâng cao như U-NET++, U-Net with Attention Mechanism, hoặc các biến thể khác để tăng cường khả năng phân đoạn và chính xác của mô hình.

Tóm lại, trong tương lai, tôi sẽ tiếp tục nghiên cứu và phát triển đề tài Ứng dụng mô hình mạng U-NET trong phân đoạn khối u não trên ảnh CT/MRI bằng cách thu thập dữ liệu mới, thử nghiệm các mô hình khác nhau, cải thiện thời gian huấn luyện và áp dụng các công nghệ mới nhất. Hy vọng rằng những công việc này sẽ mang lại kết quả tốt hơn và đóng góp vào việc nâng cao khả năng phân đoạn và chẩn đoán khối u não trong lĩnh vực y học.

TÀI LIỆU KHAM KHẢO

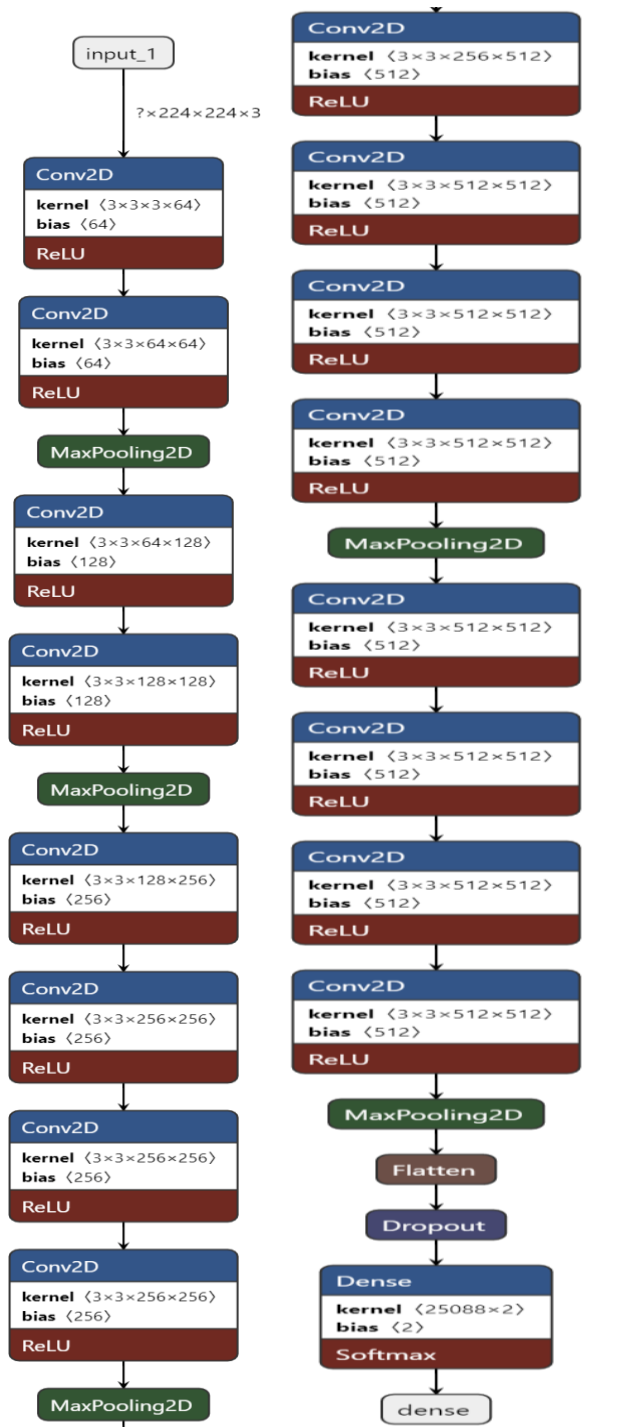
- [1] Nguyễn Thanh Tuấn (2019). *Sách Deep Learning Cơ Bản*. NXB Giáo Dục, Hà Nội.
- [2] Đỗ Trung Tuấn (1997). *Trí Tuệ Nhân Tạo*. NXB Giáo Dục, Hà Nội.
- [3] Nguyễn Thanh Thủy (1999). *Trí tuệ nhân tạo : Các phương pháp giải quyết vấn đề và kỹ thuật xử lý tri thức*. NXB Giáo Dục, Hà Nội.
- [4] Đinh Mạnh Tường (2005). *Trí tuệ nhân tạo*. NXB khoa học và kỹ thuật, Hà Nội.
- [5] Phan Huy Khánh (2004). *Lập trình logic trong Prolog*. NXB Đại học quốc gia Hà Nội, Hà Nội.
- [6] Xiao Li (2017). *Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks*. International MICCAI Brainlesion Workshop.
- [7] Anand Deshpande, Vania V. Estrela và Prashant Patavardhan (2021). *The DCT-CNN-ResNet50 architecture to classify brain tumors with super-resolution, convolutional neural network, and the ResNet50*. Department of Electronics & Communication Engineering, Angadi Institute of Technology and Management, Belagavi, Karnataka, India.
- [8] Venkatesan Rajinikanth, Alex Noel Joseph Raj, Krishnan Palani Thanaraj và Ganesh R. Naik (2020). *A Customized VGG19 Network with Concatenation of Deep and Handcrafted Features for Brain Tumor Detection*. Department of Electronics and Instrumentation Engineering, St. Joseph's College of Engineering, Chennai 600119, India; MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Milperra, NSW 2560, Australia.
- [9] Ram Deepak Gottapu and Cihan H Dagli (2018). *DenseNet for Anatomical Brain Segmentation*. Missouri University of Science and Technology, Rolla, MO, USA.
- [10] Jianxin Zhang, Xiaogang Lv, Hengbo Zhang and Bin Liu (2020). *AResU-Net: Attention Residual U-Net for Brain Tumor Segmentation*. Key Lab of Advanced Design and Intelligent Computing (Ministry of Education), Dalian University, Dalian 116622, China; School of Computer Science and Engineering, Dalian Minzu University, Dalian 116600, China.
- [11] Necip Cinar, Alper Ozcan và Mehmet Kaya (2022). *A hybrid DenseNet121-UNet model for brain tumor segmentation from MR Images*. Biomedical Signal Processing and Control.

- [12] Sami Ul Hoque (2022). *PERFORMANCE COMPARISON BETWEEN VGG16 & VGG19 DEEP LEARNING METHOD WITH CNN FOR BRAIN TUMOR DETECTION*. Lecturer, Electrical and Electronic Engineering, Feni University, Feni, Bangladesh.
- [13] Michał Futrega, Alexandre Milesi, Michał Marcinkiewicz and Pablo Ribalta . Optimized U-Net for Brain Tumor Segmentation. NVIDIA Deep Learning Examples GitHub Repository.
- [14] Kamal Kamal, Hamid EZ-ZAHRAOUI (2023). A comparison between the VGG16, VGG19, and ResNet50 architecture frameworks for classification of normal and CLAHE processed medical images. Condensed matter and interdisciplinary sciences laboratory, faculty of sciences rabat, morocco.
- [15] Rongyu Zhang, Lixuan Du, Qi Xiao, Jiaming Liu (2020). Comparison of Backbones for Semantic Segmentation Network. International School, Beijing University of Post and Telecommunication, BEIJING, 100876 CHINA.
- [16] Anand Deshpande, Vania V. Estrela and Prashant Patavardhan (2019). Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm. Computers in Biology and Medicine.
- [17] Maciej A. Mazurowski, Kal Clark, Nicholas M. Czarnek, Parisa Shamsesfandabadi, Katherine B. Peters, Ashirbani Saha (2017). *Radiogenomics of lower-grade glioma: algorithmically-assessed tumor shape is associated with tumor genomic subtypes and patient outcomes in a multi-institutional study with The Cancer Genome Atlas data*. Journal of Neuro-Oncology.
- [18] Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-net: Convolutional networks for biomedical image segmentation*. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 234-241). Springer.
- [19] Milletari, F., Navab, N., & Ahmadi, S. A. (2016). *V-net: Fully convolutional neural networks for volumetric medical image segmentation*. In 2016 Fourth International Conference on 3D Vision (3DV) (pp. 565-571). IEEE.
- [20] Khened, M., Kavurucu, Y., & Ciompi, F. (2020). *Multi-scale and multi-task U-Net architectures for biomedical image segmentation*. Neural Networks, 121, 74-87.

- [21] Isensee, F., Petersen, J., Kohl, S. A., Jäger, P. F., Maier-Hein, K. H., & Neher, P. F. (2018). *nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation*. *Nature methods*, 18(2), 203-211.
- [22] Müller, A.C., & Guido, S. (2016). *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media.
- [23] Hai loại u não thường gặp, <<https://vnexpress.net/hai-loai-u-nao-thuong-gap-4548055.html#:~:text=U%20não%20là%20một%20tập,phổi%2C%20thận...>>, xem 10/02/2023.
- [24] <<https://iq.opengenus.org/architecture-of-densenet121/>>, xem 17/02/2023.
- [25] Sơ lược về Transfer Learning, <<https://dlapplications.github.io/2018-07-15-Transfer-Learning-Basic/>>, xem 17/02/2023.
- [26] ResNet – Mạng học sâu đúng nghĩa, <<https://trituenhantao.io/kien-thuc/resnet-mang-hoc-sau-dung-nghia/>>, xem 10/02/2023.
- [27] <<https://obatambeienwasirherbal.com/vgg16-la-gi/>>, xem 18/02/2023.
- [28] <<https://dlapplications.github.io/2018-07-06-CNN/>>, xem 20/02/2023.
- [29] <<https://myknowlog.wordpress.com/2019/09/17/bai-1-image-classification/>>, xem 19/02/2023.
- [30] Image segmentation với U-Net, <<https://nttuan8.com/bai-12-image-segmentation-voi-u-net/>>, xem 15/02/2023.
- [31] <<https://viblo.asia/p/tensorflow-cho-nguoi-moi-bat-dau>>, xem 22/02/2023.
- [32] Brain MRI segmentation, <<https://www.kaggle.com/datasets/mateuszbuda/lgg-mri-segmentation>>, xem 15/01/2023.
- [33] U não: dấu hiệu, nguyên nhân và cách chẩn đoán bệnh <<https://tamanhhospital.vn/u-nao/>>, xem 23/02/2023.
- [34] Xây dựng mạng Unet cho bài toán Segmentation, <<https://viblo.asia/p/xay-dung-mang-unet-cho-bai-toan-segmentation-gDVK2QOX5Lj>>, xem 24/02/2023.

PHỤ LỤC

1. Tổng kết kiến trúc mô hình VGG-16:



2. Tổng kết kiến trúc mô hình Unet:

