

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC CẦN THƠ
KHOA CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG**



**LUẬN VĂN TỐT NGHIỆP ĐẠI HỌC
NGÀNH CÔNG NGHỆ THÔNG TIN**

Đề tài

**TÌM HIỂU ỨNG DỤNG CỦA HỌC SÂU TRONG XỬ LÝ ẢNH Y KHOA
VÀ ỨNG DỤNG KỸ THUẬT SEGMENTATION ĐỂ PHÁT HIỆN
KHỐI U NÃO TRONG ẢNH MRI**

**Sinh viên: Lê Quang Sang
Mã số: B1606927
Khóa: K42**

Cần Thơ, 01/2021

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC CẦN THƠ
KHOA CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG
BỘ MÔN CÔNG NGHỆ THÔNG TIN**



**LUẬN VĂN TỐT NGHIỆP ĐẠI HỌC
NGÀNH CÔNG NGHỆ THÔNG TIN**

Đề tài

**TÌM HIỂU ỨNG DỤNG CỦA HỌC SÂU TRONG XỬ LÝ ẢNH Y KHOA
VÀ ỨNG DỤNG KỸ THUẬT SEGMENTATION ĐỂ PHÁT HIỆN
KHỐI U NÃO TRONG ẢNH MRI**

**Người hướng dẫn
TS Trần Công Ân**

**Sinh viên thực hiện
Lê Quang Sang
Mã số: B1606927
Khóa: K42**

Cần Thơ, 01/2021

[illegible]

Cần Thơ, ngày ... tháng ... năm 2021
(Chữ ký giảng viên)

[illegible]

Cần Thơ, ngày ... tháng ... năm 2021
(Chữ ký giảng viên)

Lời cảm ơn

Đầu tiên em gửi lời cảm ơn đến thầy TS.Trần Công Ân khoa Công nghệ Thông tin và Truyền thông, trường Đại học Cần Thơ. Em xin chân thành cảm ơn thầy đã tận tình hướng dẫn góp ý, truyền đạt cảm hứng, kinh nghiệm cho em trong suốt quá trình nghiên cứu.

Tiếp theo em xin gửi lời cảm ơn chân thành đến các thầy cô trường Đại học Cần Thơ đặc biệt thầy cô Khoa Công Nghệ Thông Tin và Truyền Thông Đại học Cần Thơ đã nhiệt tình giảng dạy giúp đỡ em từ những học phần đại cương cho đến những học phần chuyên ngành cho em có đủ kiến thức để làm nền tảng thực hiện và hoàn thành luận văn của mình.

Cuối cùng em xin gửi lời cảm ơn đến gia đình, chỗ dựa tinh thần vững chắc và là động lực để em cố gắng học tập trong suốt quá trình học tại trường. Cảm ơn các bạn, anh chị đã chia sẻ những kinh nghiệm, kiến thức giúp đỡ em trong suốt quá trình vừa qua.

Mặc dù đã cố gắng hoàn thành luận văn của mình một cách chính chu nhất nhưng không thể tránh khỏi những thiếu sót, rất mong được sự góp ý của thầy cô bạn bè để em có thêm kinh nghiệm để thực hiện tốt hơn các nghiên cứu sau này.

Xin chân thành cảm ơn.

Cần Thơ, Ngày 26 tháng 12 năm 2020

Người viết

Lê Quang Sang

MỤC LỤC

PHẦN GIỚI THIỆU.....	1
1. LÝ DO CHỌN ĐỀ TÀI.....	2
2. MỤC TIÊU ĐỀ TÀI.....	2
2.1. Mục tiêu chung	2
2.2. Mục tiêu riêng	2
3. NỘI DUNG NGHIÊN CỨU	3
4. BỐ CỤC LUẬN VĂN	3
4.1. Chương 1: Cơ sở lý thuyết.....	3
4.2. Chương 2 : ứng dụng kỹ thuật Segmentation để phát hiện khối u trong ảnh MRI não.....	3
PHẦN NỘI DUNG	5
CHƯƠNG 1 – CƠ SỞ LÝ THUYẾT	6
1. GIỚI THIỆU ẢNH Y KHOA	6
2. CÁC LOẠI ẢNH Y KHOA	6
2.1. Ảnh X-Quang	6
2.2. Ảnh CT (Chụp cắt lớp vi tính).....	7
2.3. MRI (chụp cộng hưởng từ MRI)	7
3. CÁC NGHIÊN CỨU LIÊN QUAN	9
3.1. Tổng quan các nghiên cứu về ảnh y khoa.....	9
3.2. Sử dụng mạng học sâu dựa trên CNN (U-Net).....	11
3.3. Sử dụng mạng GAN	14
4. GIỚI THIỆU HỌC SÂU (DEEP LEARNING)	17
4.1. Tổng quan các phương pháp học sâu được ứng dụng trong xử lý ảnh y khoa	17
4.2. Mạng nơ-ron và các mô hình mạng nơ-ron.....	17
4.2.1. Mạng nơ-ron nhân tạo.....	17
4.2.2. Mạng nơ-ron tích chập	18
4.2.3. Các thành phần của mạng tích chập	19
a. Lớp tích chập	19
b. Lớp Pooling	19
c. Lớp kết nối đầy đủ.....	20
CHƯƠNG 2 – ỨNG DỤNG KỸ THUẬT SEGMENTATION ĐỂ PHÁT HIỆN KHỐI U TRONG ẢNH MRI NÃO	21
1. CÁC GIAO THỨC MRI ĐƯỢC SỬ DỤNG	21
2. GIỚI THIỆU TẬP DỮ LIỆU BRATS-2018.....	21
3. PHƯƠNG PHÁP THỰC HIỆN	23
3.1. Tăng cường dữ liệu	23
3.2. Xây dựng mô hình	24
3.3. Huấn luyện và tối ưu mô hình	25
3.4. Chỉ số đánh giá DSC.....	26
3.5. Kết quả đạt được và thảo luận	27
PHẦN KẾT LUẬN	31
Tài liệu tham khảo.....	33

DANH MỤC HÌNH

Hình 1: Ảnh X-quang ngực [4]	6
Hình 2: Chụp cắt lớp vi tính trên lồng ngực [5].....	7
Hình 3: Các chuỗi ảnh não MRI [7].....	8
Hình 4: Khảo sát số lượng nghiên cứu về ảnh y khoa, nhiệm vụ giải quyết, phương thức hình ảnh, các khu vực được ứng dụng [3]	9
Hình 5: Kết quả dự đoán của mô hình U-Net của Hao Dong và cộng sự [7]	13
Hình 6: Mạng SegGAN của Yuan Xue [13].....	15
Hình 7: Kết quả dự đoán của mô hình SegGAN của Yuan Xue và cộng sự [8]	16
Hình 8: Mạng CNN với nhiều lớp tích chập [8]	19
Hình 9: Tính toán tích chập trong CNN [8].....	19
Hình 10: Lớp pooling trong mạng CNN [8]	20
Hình 11: Lớp kết nối đầy đủ trong CNN [8]	20
Hình 12: Tổng quan toàn bộ mạng CNN [8]	20
Hình 13: Các chuỗi xung T1, T1ce, T2, FLAIR của một lát cắt của bệnh nhân HGG	22
Hình 14: Các chuỗi xung T1, T1ce, T2, FLAIR của một lát cắt của bệnh nhân LGG	22
Hình 15: Kiến trúc mạng U-Net [13]	25
Hình 16: Ảnh minh họa DSC [17]	26
Hình 17: Một số kết quả mô hình sự đoán trên HGG	Error! Bookmark not defined.
Hình 18: Một số kết quả mô hình sự đoán trên LGG.....	29

DANH MỤC BẢNG

Bảng 1: Bảng chỉ số đánh giá mô hình U-Net của Hao Dong và cộng sự.....	13
Bảng 2: Kết quả mô hình SegGAN của Yuan Xue.....	16
Bảng 3: Tổng kết các biến đổi trong tăng cường dữ liệu mô hình [13].....	24
Bảng 4: Các thông số cài đặt cho mô hình U-Net	25
Bảng 5: Kết quả ứng dụng mô hình U-Net.....	27

TÓM TẮT

Trong những năm gần đây máy học đặt biệt được chú ý đến, sự bùng nổ bắt đầu vào khoảng năm 2009 trở lại đây với sự xuất hiện của các mô hình học sâu vượt trội về hiệu suất và dần thay thế cho các mô hình máy học truyền thống. Các mô hình học sâu hiện tại là những mô hình nổi bật và hiện đại nhất trong nhiều lĩnh vực, từ xử lý ảnh đến xử lý ngôn ngữ tự nhiên và được áp dụng rộng rãi trong nghiên cứu và thực tế. Những sự phát triển mạnh mẽ trên đã thúc đẩy các nghiên cứu ứng dụng các mô hình học sâu vào lĩnh vực y khoa, phân tích và chuẩn đoán ảnh y khoa đang dần được hiện thực hóa bằng các mô hình học sâu. Trong nghiên cứu này tôi sẽ cung cấp một cách tổng quan ngắn gọn về sự phát triển gần đây của việc ứng dụng các mô hình máy học vào việc xử lý dữ liệu ảnh y khoa và phân tích hình ảnh. Vì điều này đã trở nên quá rộng và phát triển một cách nhanh chóng cho nên trong luận văn này chúng tôi không bao quát hết các vấn đề mà ứng học sâu vào hình ảnh y khoa đã và đang được áp dụng, do đó trong nghiên cứu này chúng tôi chỉ tập trung vào ứng dụng học sâu vào ảnh MRI cụ thể hơn là bài toán phân đoạn hình ảnh. Sau đó là trình bày việc ứng dụng mạng học sâu U-Net để phát hiện và phân đoạn khối u não trong ảnh MRI. Mô hình cài đặt của chúng tôi được đánh giá trên tập dữ liệu BRATS 2018, bao gồm 220 trường hợp HGG (High-grade brain tumor) và 75 trường hợp LGG (Low-grade brain tumor). Mô hình có thể mang lại phân đoạn một cách tích cực khả quan.

ABSTRACT

In recent years machine learning has received special attention, The current boom started around 2009 when so-called deep artificial neural networks began outperforming other established models on a number of important benchmarks. Deep neural networks are now the state-of-the-art machine learning models across a variety of areas, from image analysis to natural language processing, and widely deployed in academia and industry. These developments have a huge potential for medical imaging technology, medical data analysis, medical diagnostics and healthcare in general, slowly being realized. In this study we provide a short overview of recent advances and some associated challenges in machine learning applied to medical image processing and image analysis. As this has become a very broad and fast expanding field we will not survey the entire landscape of applications, but particular focus on deep learning in MRI. More specifically than the image segmentation problem. And then presented the application of deep learning network U-Net to detect and segment brain tumors in MRI images. Our method was evaluated on Multimodal Brain Tumor Image Segmentation (BRATS 2018) datasets, which contain 220 high-grade brain tumor and 75 low-grade tumor cases. Method can obtain promising segmentation efficiently.

PHẦN GIỚI THIỆU

1. LÝ DO CHỌN ĐỀ TÀI

Hình ảnh y khoa nhằm tìm kiếm các cấu trúc bên trong được che giấu bởi da và xương cũng như chẩn đoán và điều trị. Các bác sĩ lâm sàng có kinh nghiệm sẽ chẩn đoán bệnh thông qua các hình ảnh y khoa này. Với nhiều cách diễn giải ảnh y khoa khác nhau dẫn đến chẩn đoán bệnh sẽ khác nhau. Bên cạnh đó, diễn giải ảnh y khoa làm mất nhiều thời gian, chi phí và dễ gây mệt mỏi cho các bác sĩ lâm sàng dẫn đến việc chẩn đoán sai lầm và thiếu sót.

Do đó, nếu công việc diễn giải ảnh y khoa có sự giúp đỡ của máy học và kỹ thuật xử lý ảnh sẽ giúp các bác sĩ lâm sàng tiết kiệm được thời gian, chi phí y tế cũng như công sức giúp cho nhiều người được điều trị bệnh hơn mà không cần tốn thời gian vào việc diễn giải hình ảnh y khoa.

Hiện nay lĩnh vực phân tích ảnh y khoa đang được nhiều nhà nghiên cứu quan tâm và thực hiện các nghiên cứu về nó. Với sự trợ giúp của máy học đặc biệt là học sâu giúp cho việc chẩn đoán bệnh qua hình ảnh y khoa trở nên tiện lợi hơn, nâng cao độ chính xác. Giúp tiết kiệm thời gian chi phí của bác sĩ trong quá trình chuẩn đoán hình ảnh để tập trung vào các thao tác lâm sàng khác.

2. MỤC TIÊU ĐỀ TÀI

2.1. Mục tiêu chung

Hiện nay với sự phát triển mạnh mẽ của Machine Learning cũng như các mô hình Deep Learning, đã đến lúc các nghiên cứu được ứng dụng thử nghiệm trên thực tế để đánh giá độ chính xác cũng như khả năng của các mô hình trên dữ liệu thực tế.

Nhằm khảo sát một cách ngắn gọn về các nghiên cứu hiện tại về ảnh y khoa cũng như một số phương pháp kỹ thuật có liên quan đến bài toán phát hiện và phân đoạn khối u trên ảnh MRI não.

2.2. Mục tiêu riêng

Nhằm tìm ra các phương pháp phát hiện và phân đoạn hình ảnh một cách tối ưu, cụ thể là phát hiện và phân đoạn khối u trên ảnh MRI não, tiến hành cài đặt và thử nghiệm kỹ thuật Segmentation để phát hiện và phân đoạn khối u não trên ảnh MRI.

Với một số chiến thuật huấn luyện mô hình khác nhau nhằm đánh giá tính ổn định của bộ dữ liệu và độ chính xác của mô hình. Mô hình được cài đặt và thử nghiệm trên tập dữ liệu BRATS 2018.

3. NỘI DUNG NGHIÊN CỨU

Đầu tiên chúng tôi tìm hiểu về các phương pháp học sâu, lý thuyết về học sâu tiếp đến là việc ứng dụng học sâu trong lĩnh vực y khoa. Các vấn đề mà học sâu đang được nghiên cứu ứng dụng để giải quyết các bài toán phân tích xử lý ảnh y khoa. Một số vấn đề tiêu biểu như bài toán: phân loại hình ảnh, phát hiện vùng bất thường trên hình ảnh y khoa, bài toán phân đoạn vùng, bộ phận, cơ quan, cấu trúc trên ảnh y khoa.

Phần tiếp theo là ứng dụng kỹ thuật Segmentation để phát hiện và phân đoạn khối u trên ảnh MRI, trong phần này chúng tôi tìm hiểu về mô hình về nghiên cứu phân đoạn hình ảnh, tìm hiểu các chuẩn ảnh thường được sử dụng và mang lại kết quả cao trong quá trình huấn luyện và dự đoán. Tiến hành cài đặt mô hình và tinh chỉnh một số thao tác huấn luyện để tăng độ chính xác dự đoán của mô hình.

4. BỐ CỤC LUẬN VĂN

Luận văn bao gồm 2 chương chính: Cơ sở lý thuyết và Ứng dụng kỹ thuật Segmentation để phát hiện khối u trong ảnh MRI não

4.1. Chương 1: Cơ sở lý thuyết

Đầu tiên giới thiệu về các khái niệm ảnh y khoa, các loại ảnh y khoa cũng như là các chuẩn lưu trữ của các chuẩn ảnh trong ngành y khoa. Các loại ảnh y khoa hiện tại đã và đang được sử dụng rộng rãi trong chuẩn đoán y tế, ưu nhược điểm của từng loại ảnh.

Tiếp theo đó là giới thiệu về các nghiên cứu xử lý ảnh y khoa hiện tại đang được áp dụng, các thống kê liên quan số lượng nghiên cứu về ảnh y khoa trên các phương diện như: tác vụ cần giải quyết, phương thức ảnh sử dụng, bộ phận trên cơ thể được nghiên cứu,...

Sau đó là trình bày các nghiên cứu liên quan bài toán phân đoạn khối u não trên ảnh MRI đang được thực hiện và đạt kết quả khả quan trong đó có 2 nghiên cứu nổi tiếng nhất tại thời điểm hiện tại là: sử dụng mạng U-Net của Hao Dong và cộng sự năm 2017, sử dụng mạng SegAN của Yuan Yue và cộng sự năm 2018. Trong phần này chúng tôi sẽ trình bày cụ thể các kỹ thuật mà 2 phương pháp trên áp dụng và các tập dữ liệu họ đã tiến hành thử nghiệm.

Cuối cùng giới thiệu về các cơ sở lý thuyết liên quan đến các phương pháp học sâu thường được áp dụng trong huấn luyện các bài toán phân tích ảnh y khoa. Sau đó là giới thiệu về mạng nơ-ron nhân tạo cũng như là mạng nơ-ron tích chập, các thành phần của mạng tích chập được sử dụng trong kiến trúc mạng U-Net.

4.2. Chương 2 : ứng dụng kỹ thuật Segmentation để phát hiện khối u trong ảnh MRI não

Đầu tiên giới thiệu về giao thức MRI được sử dụng trong mô hình thực nghiệm, đặc điểm của chuỗi xung FLAIR sau đó giới thiệu về tập dữ liệu tiêu chuẩn BRATS-2018 cũng như là cấu trúc tập tin bên trong tập dữ liệu được cung cấp.

Sau đó chúng tôi giới thiệu và trình bày các phương pháp thực nghiệm cài đặt và huấn luyện mô hình cũng như các thao tác tiền xử lý dữ liệu, tăng cường dữ liệu.

Tiếp đến là các chỉ số đánh giá mô hình đặt biệt cho bài toán Segmentation, các thông số cài đặt mô hình, số lượng dữ liệu huấn luyện và kiểm tra mô hình và một số phương pháp huấn luyện mô hình, đánh giá mô hình.

Cuối cùng là phần kết quả đạt được của mô hình và phân thảo luận những phần mà mô hình chưa thực hiện được, cũng như đánh giá tổng quan về mô hình và hướng phát triển tiếp theo của mô hình hiện tại.

PHẦN NỘI DUNG

CHƯƠNG 1 – CƠ SỞ LÝ THUYẾT

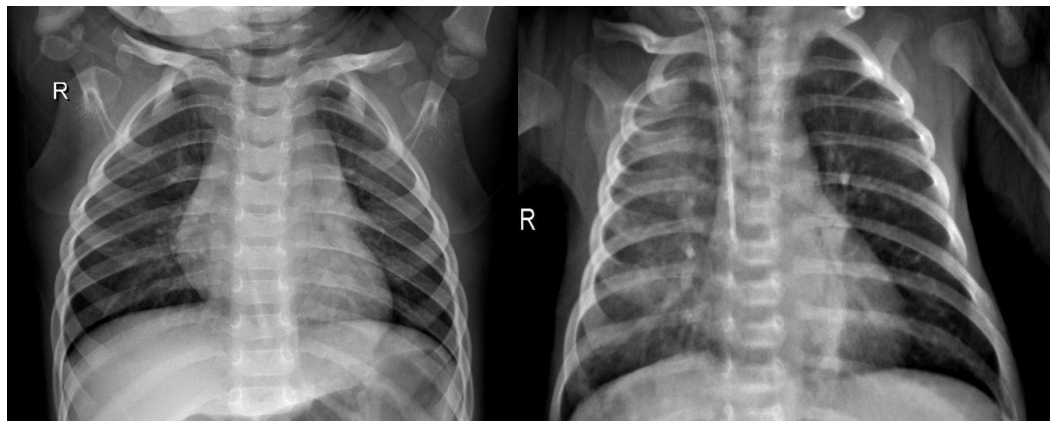
1. GIỚI THIỆU ẢNH Y KHOA

Ảnh y khoa hay hình ảnh y khoa là kỹ thuật và quy trình tạo hình ảnh trực quan về bên trong của cơ thể để phân tích lâm sàng và can thiệp y tế, cũng như biểu thị trực quan chức năng của một số cơ quan hoặc mô sinh lý học. Hình ảnh y khoa cũng thiết lập một cơ sở dữ liệu giải phẫu học và sinh lý học bình thường để phục vụ việc xác định các bất thường trong mô sinh học. Hình ảnh y khoa theo nghĩa rộng nhất, nó là một phần của hình ảnh sinh học và kết hợp với hình ảnh y khoa sử dụng các công nghệ hình ảnh như: chụp X-quang, chụp ảnh cộng hưởng từ (MRI), chụp cắt lớp vi tính (CT),...

2. CÁC LOẠI ẢNH Y KHOA

2.1. Ảnh X-Quang

Sử dụng bức xạ ion để tạo nên hình ảnh của cơ thể, một chùm tia X xuyên qua cơ thể một số bị hấp thụ một số bị phân tán, phần còn lại sẽ được truyền đến máy để ghi nhận hình ảnh, những phần hấp thụ lại trên cơ thể tùy vào mức độ và lượng bức xạ mà một tế bào trên cơ thể hấp thụ sẽ tạo ra những điểm ảnh đậm hay nhạt khác nhau trên ảnh. Ví dụ như xương sẽ hấp thụ nhiều các tia bức xạ hơn thì hình ảnh thu được từ chụp x-quang chúng ta sẽ dễ dàng thấy phần xương có màu trắng hơn các mô khác, cũng như dựa vào lượng bức xạ mà một tế bào hấp thụ thể hiện trên hình ảnh mà bác sĩ dễ dàng nhận dạng được sự bất thường trong cơ thể, cũng như các mô hình máy học có thể trích xuất được các đặc trưng trên ảnh.

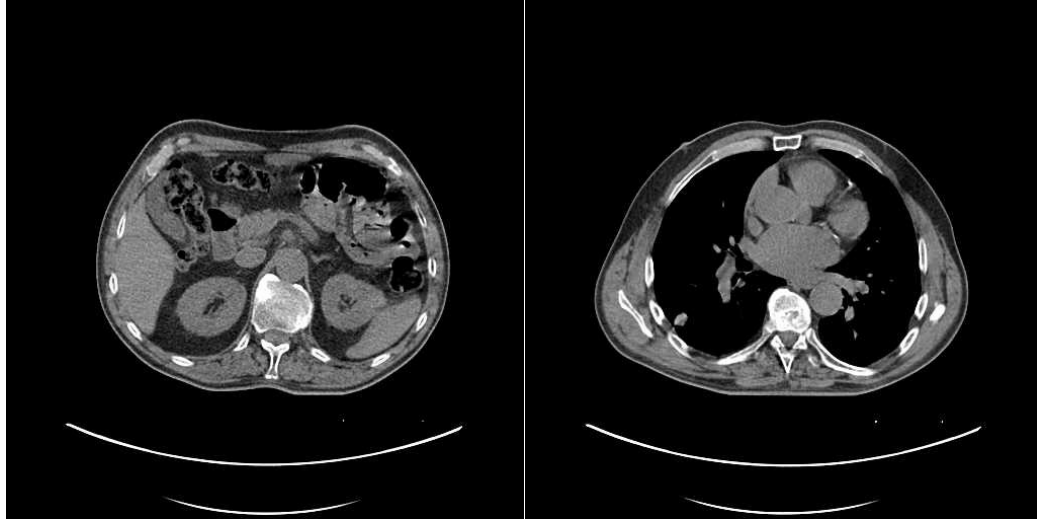


Hình 1: Ảnh X-quang ngực [4]

Trên đây ảnh bên trái là ảnh của một bệnh nhân không bị viêm phổi, ảnh bên phải là của một bệnh nhân viêm phổi, do các tế bào xấu hấp thụ nhiều lượng bức xạ hơn ta dễ dàng nhận thấy hình ảnh phần phổi bên phải có phần trắng hơn.

2.2. Ảnh CT (Chụp cắt lớp vi tính)

Chụp cắt lớp vi tính sử dụng một chùm tia X xoay xung quanh cơ thể bệnh nhân để ghi nhận hình ảnh cắt lớp 2D, kết hợp nhiều lát cắt xếp chồng lên nhau hình thành ảnh 3D. Ảnh CT là một hình ảnh 3D bao gồm nhiều lát cắt vi tính trên cơ thể các hình ảnh cắt lớp này sử dụng cho việc chẩn đoán và trị liệu, cũng như khi chụp ảnh X-quang thì độ đậm nhạt, trắng đen trên ảnh phụ thuộc vào sự hấp thụ các tia bức xạ của các bộ phận trên cơ thể.



Hình 2: Chụp cắt lớp vi tính trên lồng ngực [5]

Trên đây là hai lát cắt từ một ảnh CT ngực 3D, tùy vào mục đích nghiên cứu mà chúng ta có thể chọn các lát cắt phù hợp với nghiên cứu và có thể chọn tất cả các lát cắt.

2.3. MRI (chụp cộng hưởng từ MRI)

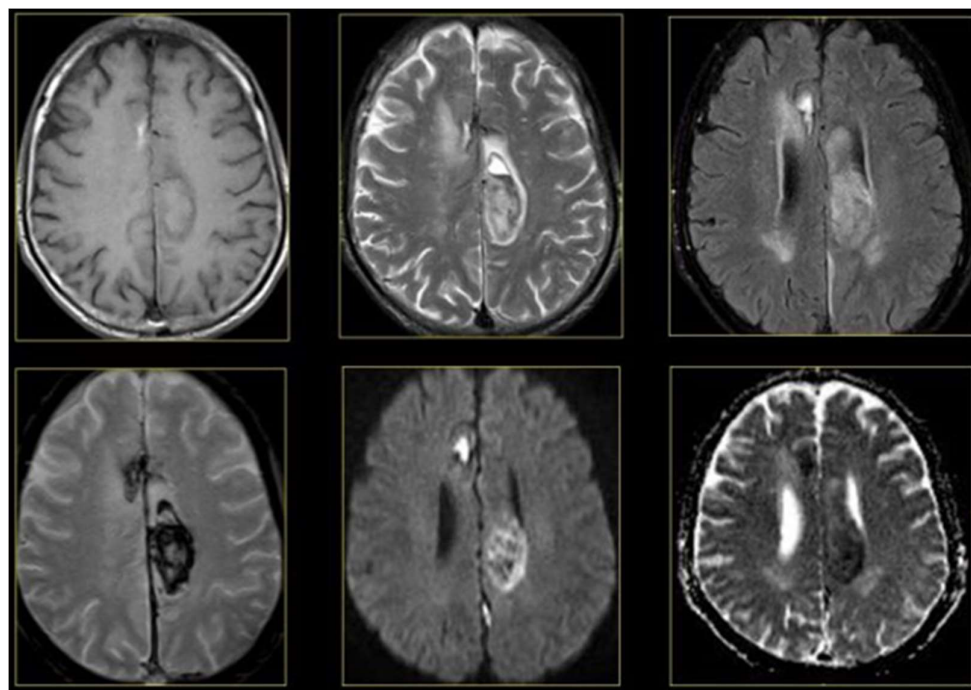
Chụp cộng hưởng từ MRI là một trong những phương pháp chẩn đoán hình ảnh cận lâm sàng hiện đại và hiệu quả với hình ảnh rõ nét và giúp cho việc chẩn đoán chính xác tình hình bệnh. Chụp cộng hưởng từ MRI giúp chẩn đoán nhiều bệnh lý hiệu quả hơn các phương pháp cận lâm sàng khác. Cộng hưởng từ MRI là một kỹ thuật tạo hình cắt lớp sử dụng sóng từ trường và sóng Radio. Khi các nguyên tử Hydrogen trong cơ thể người dưới tác động của từ trường và sóng Radio, hấp thụ và phóng thích năng lượng RF (Radiologist Frequency) [6]. Quá trình phóng thích này được máy thu nhận xử lý và chuyển đổi các tín hiệu thành hình ảnh. Hình ảnh cộng hưởng từ MRI có độ tương phản cao, sắc nét và rõ ràng, chi tiết, giải phẫu tốt và có khả năng tái tạo 3D mang lại hiệu quả chẩn đoán cho bác sĩ đối với bệnh lý của bệnh nhân. Trong rất nhiều trường hợp, hiệu quả chẩn đoán của MRI tốt hơn rất nhiều so với chụp X-

quang, chụp siêu âm hay chụp cắt lớp CT,... Do độ phân giải của hình ảnh MRI cao hơn nhiều so với chụp CT hoặc là X-Quang do đó ảnh MRI có thể dễ dàng chuẩn đoán các vùng tổn thương có kích thước nhỏ, ví dụ tổn thương xơ hóa rải rác trong não, nhồi máu thân não, các khối u nhỏ trong não,...

Chuẩn ảnh MRI không phải là ảnh JPG hoặc PNG thông thường các giá trị trong từng điểm ảnh MRI không phải là giá trị Grayscale từ 0 – 255 như ảnh trắng đen thông thường. Giá trị của mỗi điểm ảnh trên ảnh MRI rơi vào khoảng từ 7.6×10^{-6} cho đến 720. Ảnh MRI là kết quả của quá trình thu thập sóng Radio kết hợp với sóng điện từ sau đó mới cho ra được hình ảnh MRI chứ không phải là kết quả thu thập từ sóng ánh sáng thông thường.

Do đó với một số tác vụ đặt biệt cụ thể thì chúng ta không thể nào chuyển đổi từ chuẩn ảnh MRI sang chuẩn ảnh JPG hoặc PNG thông thường để phục vụ việc chuẩn đoán cho các tác vụ đặt biệt. Vì làm như vậy thì chúng ta sẽ bị mất mát thông tin rất nhiều trên từng điểm ảnh.

Bên cạnh đó chụp cộng hưởng từ MRI rất an toàn, không sử dụng tia bức xạ nên được các bác sĩ chuyên môn đánh giá cao trong chỉ định chụp và chẩn đoán bệnh.



Hình 3: Các chuỗi ảnh não MRI [7]

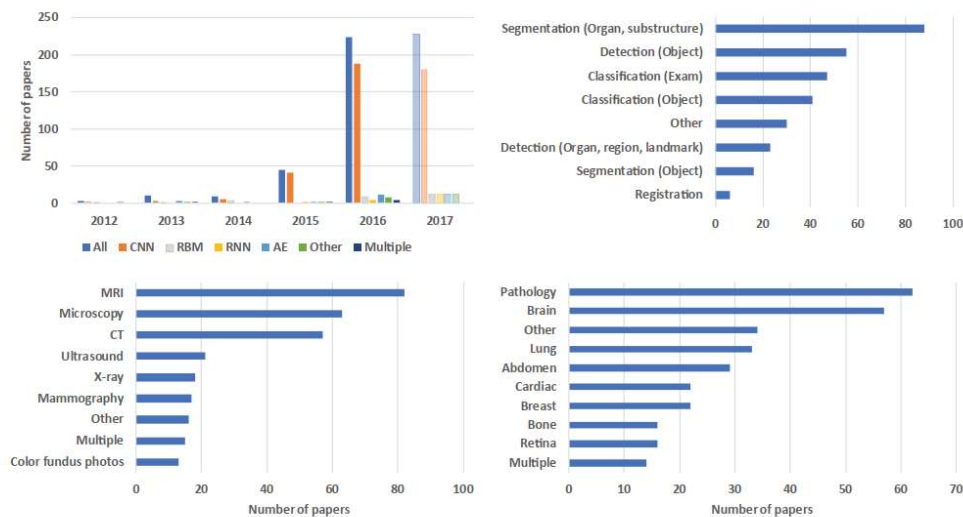
3. CÁC NGHIÊN CỨU LIÊN QUAN

3.1. Tổng quan các nghiên cứu về ảnh y khoa

Các mô hình máy học ngày nay đã tạo ra các phương pháp tiếp cận hiện đại đối với nhiều vấn đề trong thị giác máy tính, mô hình ngôn ngữ, tương tác người máy. Học sâu đã vươn lên vị trí nổi bật khi mạng nơ-ron nhân tạo vượt trội hơn các mạng truyền thống trên nhiều tiêu chuẩn đánh giá hình ảnh. Nổi bật nhất là trong thách thức nhận dạng hình ảnh ImageNet trong năm 2012 [1] khi một mô hình học sâu (một mạng nơ-ron tích chập) đã giảm một nửa tỷ lệ lỗi tốt thứ hai so với các mô hình trong thử thách đó.

Máy tính có thể nhận dạng đối tượng trong hình ảnh tự nhiên là một nhiệm vụ rất khó khăn cho đến thời điểm gần đây, nhưng hiện tại thì các mạng nơ-ron tích chập đã vượt qua hiệu suất của cả con người trên ImageNet. Học sâu đã trở thành một kỹ thuật tiêu chuẩn cho các vấn đề của thị giác máy tính.

Ứng dụng của học sâu trong phân tích ảnh y khoa dần được xuất hiện trong các hội thảo cho đến các nghiên cứu bắt đầu rộng rãi trong những năm 2015 và 2016. Hình 1 bên dưới thể hiện số lượng bài báo, nhiệm vụ giải quyết, phương thức hình ảnh sử dụng, khu vực được áp dụng đặt biệt trên IEEE. Một khảo sát đánh giá việc ứng dụng các mô hình học trong lĩnh vực y tế đầu tiên đã được Shen và các cộng sự xuất bản năm 2017 [2].



Hình 4: Khảo sát số lượng nghiên cứu về ảnh y khoa, nhiệm vụ giải quyết, phương thức hình ảnh, các khu vực được ứng dụng [3]

Mặc dù học sâu được ứng dụng trong nhiều bài toán xử lý ảnh y khoa như phân loại hình ảnh, nhận dạng đối tượng, phân đoạn hình ảnh, nhưng trong phạm vi nghiên cứu của luận văn này chỉ tập trung vào việc khảo sát các kỹ thuật phân đoạn hình ảnh phát hiện đối tượng trong hình ảnh y khoa cụ thể là MRI não.

Trong đề tài này chúng tôi sẽ trình bày một số nghiên cứu gần đây liên quan đến việc phân đoạn hình ảnh cũng như phát hiện khối u, vùng bất thường trên ảnh y tế MRI và giới thiệu các mô hình và kỹ thuật đang được áp dụng trong xử lý ảnh y khoa, tiếp theo là tiến hành phân tích và cài đặt mô hình U-Net để phát hiện phân đoạn khối u não trên ảnh MRI.

Trong phần tiếp theo chúng tôi sẽ trình bày hai hướng nghiên cứu khác nhau trong bài toán phân đoạn ảnh y khoa. Hướng nghiên cứu thứ nhất sử dụng mô hình học sâu dựa trên CNNs truyền thống và hướng nghiên cứu thứ hai là sử dụng mạng đối lập GAN (Generative adversarial network) [9]

3.2. Sử dụng mạng học sâu dựa trên CNN (U-Net)

Một thách thức lớn trong việc thiết kế phát đồ điều trị khối u não và đánh giá phân lượng của khối u là việc xác định mức độ xâm lấn của khối u. Kỹ thuật chụp không xâm lấn cộng hưởng từ (MRI) phát triển như một công cụ chuẩn đoán hàng đầu, nhưng đối với các khối u não không có bức xạ ion hóa, việc phân đoạn thủ công khối u não là một việc cực kì khó khăn, phụ thuộc và kinh nghiệm của người chuẩn đoán trong bối cảnh này thì một phương pháp phân đoạn khối u não một cách tự động là rất cần thiết để đo lường, đánh giá mức độ xâm lấn của khối u. Trong nghiên cứu của Hao Dong và cộng sự trong năm 2017 [10] đã phát triển một mạng học sâu dựa trên mạng U-Net [11] để phân đoạn tự động khối u não.

Trong nghiên cứu của Hao Dong và cộng sự đã đánh giá mô hình trên tập dữ liệu phân đoạn hình ảnh não đa phương thức BRATS 2015 [12] bao gồm 220 bệnh nhân HGG và 54 bệnh nhân LGG.

Phân đoạn hình ảnh là một tác vụ phổ biến trong cả phân tích hình ảnh tự nhiên và hình ảnh y tế, mục tiêu của phân đoạn ảnh là phân loại từng điểm ảnh (pixel) trong ảnh với một lớp tương ứng với những gì đang được biểu diễn trên ảnh. Bởi vì chúng ta dự đoán cho mỗi điểm ảnh trong ảnh nên tác vụ này thường được gọi là dự đoán dày đặc (dense prediction). Điều đáng chú ý là việc phân đoạn hình ảnh không giống như việc phân loại hình ảnh hoặc phát hiện đối tượng trong ảnh, đơn giản chỉ là gán nhãn hoặc bounding box (trong bài toán phát hiện đối tượng trong ảnh). Bản thân đầu ra của bài toán phân đoạn là một hình ảnh có độ phân giải cao thường là giống độ phân giải ảnh đầu vào, trên mỗi điểm ảnh được phân loại vào một lớp cụ thể nào đó. vì vậy bài toán phân đoạn nó là một loại phân loại hình ảnh cấp độ điểm ảnh.

Bất kỳ một kiến trúc phân đoạn hình ảnh nào cũng có thể nói là có một phần mã hóa (encoder) và theo sau là phần giải mã (decoder). Phần mã hóa thường là một mạng phân loại được đào tạo trước đó (pre-trained) như là ResNet/VGG có nhiệm vụ phân loại từng điểm ảnh vào một danh mục nhất định. Mặt khác bộ giải mã lấy các điểm ảnh có độ phân giải thấp đó đã được phân loại chiếu các thuộc tính đó lên mặt phẳng để có được hình ảnh phân đoạn chất lượng cao.

Một trong những mạng phân đoạn ảnh tốt nhất đặt biệt trong ảnh y khoa là U-Net được Ronneberger và cộng sự giới thiệu năm 2015, U-Net được thiết kế để xử lý hình ảnh y khoa. Các mạng thông thường yêu cầu số lượng lớn dữ liệu và chất lượng cao để huấn luyện mô hình là điều không khả dụng trong lĩnh vực này. U-Net ra đời để giải quyết tất cả các khó khăn của mạng CNNs trước đây cũng như đầu ra mong muốn các điểm ảnh sẽ được phân loại theo các nhãn tương ứng.

Trong nghiên cứu của Hao Dong và cộng sự họ đã sử dụng mô hình dựa trên mô hình truyền thống ban đầu của U-Net và không sử dụng các mô hình đã được huấn luyện trước đó để làm bộ trích xuất đặt trung. Mỗi phần lấy mẫu xuống và lấy mẫu lên đều bao gồm 5 khối tích chập để giảm kích thước từ 240 x 240 xuống 15 x 15 và tăng kích thước trở lại 15 x 15 lên 240 x 240.

Sự khác biệt ở mô hình của Hao Dong và cộng sự so với U-Net của Ronneberger là phần đệm bằng không cho cả hai chiều giảm mẫu và tăng mẫu của 2 phần trong mô hình U-Net, sau đó sử dụng một lớp tích chập với bộ lọc 1x1 để giảm giảm số lượng bản đồ đặt trung xuống bằng 2 cho phân biệt tiền cảnh và hậu cảnh tương ứng.

Trong quá trình đào tạo mô hình Hao Dong và cộng sự đã sử dụng đánh giá Soft Dice [13] như là hàm chi phí của mô hình, hàm này tốt hơn hàm mất mát thông tin hoặc sai số bình phương và đặt biệt phù hợp cho các bài toán phân loại trên cấp độ điểm ảnh. Bên cạnh đó họ cũng sử dụng Adam như một hàm tối ưu các tham số.

$$DSC = \frac{2TP}{FP + 2TP + FN}$$

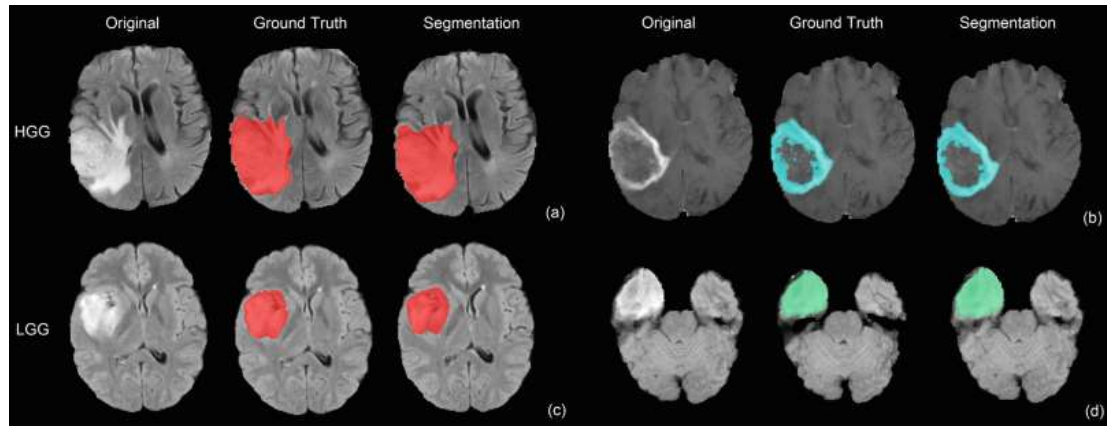
Trong đó:

- TP: phần đúng tích cực
- FP: phần sai tích cực
- FN: Phần sai tiêu cực

Hao Dong và cộng sự đã dùng ảnh Flair để huấn luyện cho mạng phân đoạn toàn bộ khối u, và ảnh T1ce để huấn luyện cho mạng phân đoạn phần lõi (phần hoại tử) của khối u. Hao Dong và cộng sự đã sử dụng đánh giá chéo để đánh giá mô hình trên cả bệnh nhân HGG và LGG sau đây là kết quả nghiên cứu của Hao Dong và các mô hình cùng thời điểm.

DSC					
Method	Data	Grade	Complete	Core	Enhancing
Proposed	Cross-validation on BRATS 2015 training datasets	HGG	0.88	0.87	0.81
		LGG	0.84	0.85	0.00
		Combined	0.86	0.86	0.65
Pereira16	BRATS 2013 leaderboard	HGG	0.88	0.76	0.73
		LGG	0.65	0.53	0.00
		Combined	0.84	0.72	0.62
	BRATS 2013 challenge	HGG	0.88	0.83	0.77
	BRATS 2015 challenge	Combined	0.79	0.65	0.75
Havaei16	BRATS 2013 training	Combined	0.88	0.79	0.73
	BRATS 2013 challenge	Combined	0.88	0.79	0.73
	BRATS 2013 leaderboard	Combined	0.84	0.71	0.57
	BRATS 2015 challenge	Combined	0.79	0.58	0.69
Kamnitsas17	BRATS 2015 training	Combined	0.90	0.76	0.73
	BRATS 2015 training	Combined	0.90	0.76	0.73
	BRATS 2015 challenge	Combined	0.85	0.67	0.63

Bảng 1: Bảng chỉ số đánh giá mô hình U-Net của Hao Dong và cộng sự



Hình 5: Kết quả dự đoán của mô hình U-Net của Hao Dong và cộng sự [7]

3.3. Sử dụng mạng GAN

Hiện tại những kiến mạng học sâu nổi bật nhất dựa trên CNNs cho bài toán phân đoạn hình ảnh là U-Net như đã trình bày ở trên, nhưng phần lớn các mạng U-Net sử dụng cross-entropy như là hàm lỗi để giảm thiểu vấn đề mất cân bằng của dữ liệu, tuy nhiên việc lựa chọn trọng số trong hàm lỗi của U-Net là một nhiệm vụ khó khăn và khó tối ưu.

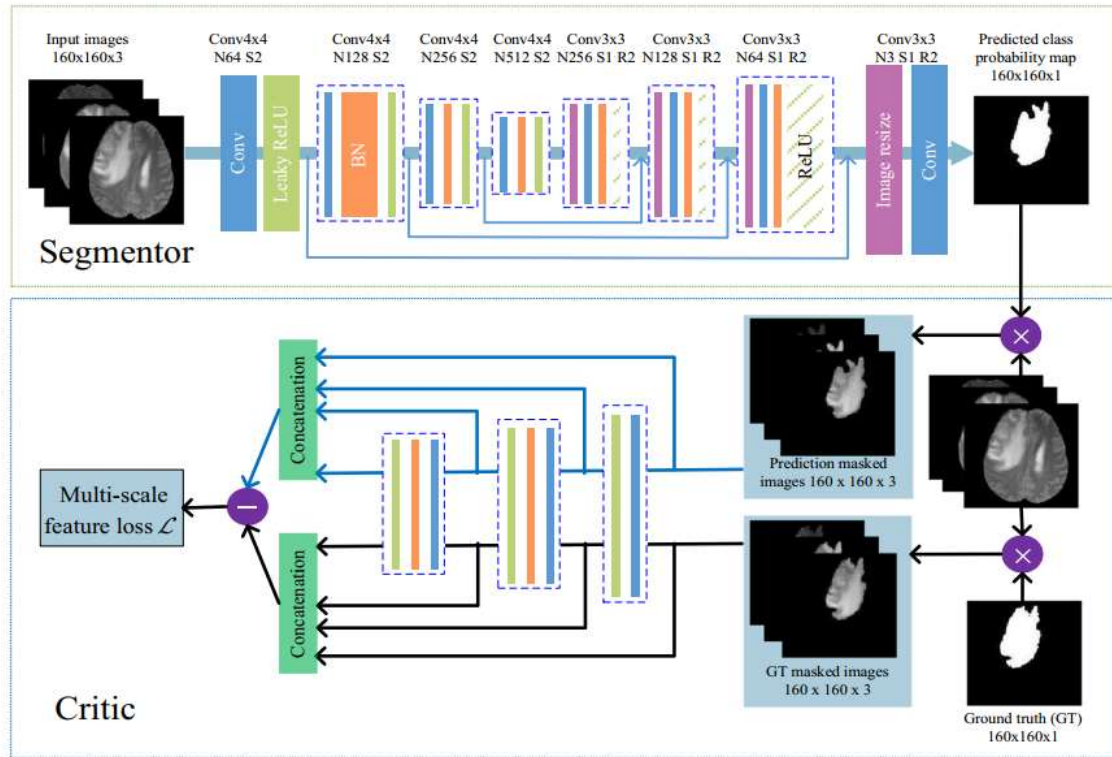
Trong nghiên cứu của Yuan Xue và cộng sự trong năm 2018 [14], họ đã sử dụng một mạng tích chập kết nối đầy đủ như là một mạng để tạo ra các phân đoạn của khối u não, và đề xuất một mạng mới dựa trên mạng GAN với hàm lỗi là Multi-scale L1 [15] để tập trung vào việc học các đặc trưng cục bộ và toàn cục của quan hệ các điểm ảnh trong không gian.

Trong mô hình trên thì mạng phân đoạn và mạng chỉ trích thì được đào tạo theo kiểu trò chơi min-max. Mạng chỉ trích được đào tạo tối đa hóa hàm lỗi multi-scale, mạng tạo phân đoạn khối u thì được đào tạo tối thiểu hóa hàm lỗi multi-scale. Như hình minh họa bên dưới mạng SegAN được đề xuất gồm 2 phần: mạng phân đoạn (Segmentor) và mạng chỉ trích (Critic).

Mạng phân đoạn là một mạng tích chập encoder-decoder kết nối đầy đủ như là mạng để sinh ra bản đồ nhãn xác suất từ ảnh đầu vào. Mạng chỉ trích nhận vào 2 đầu vào: ảnh gốc được đánh dấu bởi bản đồ nhãn chính xác và ảnh gốc được đánh dấu bởi bản đồ nhãn sinh ra từ mạng phân đoạn.

Mạng phân đoạn và mạng chỉ trích được đào tạo theo kiểu đối đầu. Segmentor sử dụng một mạng tích chập encoder-decoder kết nối đầy đủ, Yuan Xue sử dụng bộ lọc 4×4 với bước nhảy 2 cho phần encoder và với phần decoder sử dụng kernel 3×3 với bước nhảy là 1. SegAN cũng theo kiến trúc của U-Net là bỏ qua một số kết nối tương ứng giữa các lớp của phần mã hóa và giải mã.

Critic kiến trúc phần critic tương tự phần encoder của mạng phân đoạn ở trên. Các đặt trưng được trích xuất theo từng cấp qua nhiều lớp để tính chi phí multi-scale L1. Hàm mục tiêu này có thể nắm bắt được các mối quan hệ không gian trong phạm vi dài ngắn của các đặt trưng theo từng thứ bậc: cấp độ điểm ảnh đặt trưng (pixel-level features), đặt trưng cấp thấp (low-level features), đặt trưng mức trung gian (middle-level features).



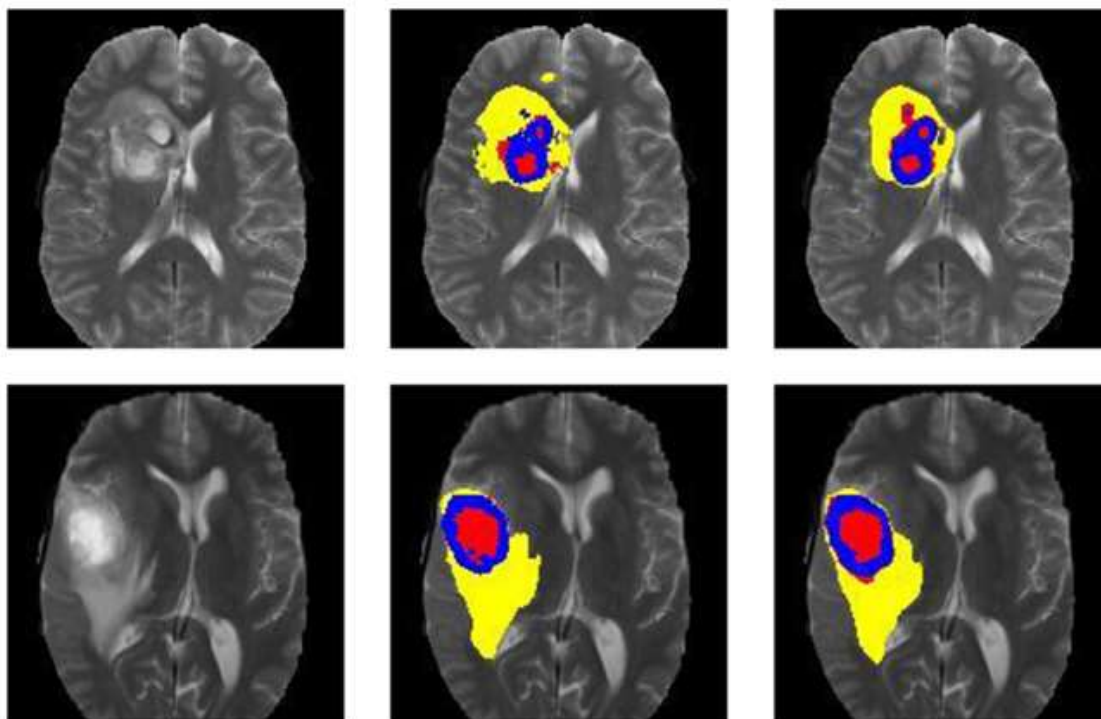
Hình 6: Mạng SegGAN của Yuan Xue [13]

Yuan Xue sử dụng 4 khối convolutional cho mỗi phần giảm mẫu và tăng cường mẫu trong mạng sinh (phân đoạn của SegAN). Bên cạnh 3 khối tích chập bên trong mạng chỉ trích (Critic). Kích thước bỏ đưa vào huấn luyện là 64, tốc độ học 0.00002. Yuan Xue thực hiện việc đánh giá mô hình đề xuất trên tập dữ liệu BRATS 2015, họ đã thay đổi kích thước ảnh từ 240 x 240 xuống còn 180 x 180 để lấy phần trung tâm của ảnh.

Bảng bên dưới so sánh kết quả của mô hình với một số mô hình cùng thời điểm, trong đó có kết quả của mô hình U-Net trong nghiên cứu của Hao Dong như đã trình bày ở trên.

Data	Methods	Dice			Precision			Sensitivity		
		Whole	Core	Enhancing	Whole	Core	Enhancing	Whole	Core	Enhancing
BRATS 2013 Leaderboard	Havaei et al. (2017)	0.84	0.71	0.57	0.88	0.79	0.54	0.84	0.72	0.68
	Pereira et al. (2016)	0.84	0.72	0.62	0.85	0.82	0.60	0.86	0.76	0.68
	SegAN	0.84	0.70	0.65	0.87	0.80	0.68	0.83	0.74	0.72
BRATS 2015 Test	Kamnitsas et al. (2017)	0.85	0.67	0.63	0.85	0.86	0.63	0.88	0.60	0.67
	U-Net	0.80	0.63	0.64	0.83	0.81	0.78	0.80	0.58	0.60
	SegAN	0.85	0.70	0.66	0.92	0.80	0.69	0.80	0.65	0.62

Bảng 2: Kết quả mô hình SegGAN của Yuan Xue



Hình 7: Kết quả dự đoán của mô hình SegGAN của Yuan Xue và cộng sự [8]

4. GIỚI THIỆU HỌC SÂU (DEEP LEARNING)

Hiện nay các mô hình máy học được sử dụng rộng rãi trong việc phân tích hình ảnh y khoa, các nhà nghiên cứu có thể sử dụng ảnh y khoa để huấn luyện cho các mô hình máy học giải quyết các bài toán như phân loại bệnh, phát hiện bệnh hay thậm chí có thể khoanh vùng chỉ biết các cơ quan, vùng tổn thương. Một số nghiên cứu về ảnh y khoa hiện tại, cũng như các bài toán mà ứng dụng máy học vào lĩnh vực y khoa có thể được giải quyết.

Chuẩn đoán chính xác bệnh phụ thuộc vào việc thu nhận hình ảnh và phân tích hình ảnh, các thiết bị thu nhận hình ảnh được cải thiện trong những năm gần đây các công nghệ X-ray hay CT, MRI có khả năng cung cấp ảnh Radio-Logical với độ phân giải cao. Một trong những ứng dụng học máy tốt nhất là thị giác máy tính. Máy học đã phát triển một cách vượt trội trong vài năm qua, đạt được khả năng sàng lọc, trích xuất kĩ càng thông qua dữ liệu lớn và phức tạp. Học sâu rất được quan tâm trong mọi lĩnh vực khoa học đặc biệt là trong phân tích hình ảnh y khoa.

4.1. Tổng quan các phương pháp học sâu được ứng dụng trong xử lý ảnh y khoa

Theo phương pháp học các thuật toán học sâu được ứng dụng trong y khoa thường được chia làm hai nhóm chính: học có giám sát và học không giám sát. Học có giám sát (Supervised Learning) là thuật toán dự đoán đầu ra của một dữ liệu mới dựa trên các cặp (đầu vào, đầu ra) đã biết từ trước. Supervised learning là nhóm phổ biến nhất trong các thuật toán Machine Learning.

Một cách toán học, học giám sát là khi chúng ta có một tập hợp biến đầu vào $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ và một tập hợp đầu ra tương ứng $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\}$, trong đó $\mathbf{x}_i, \mathbf{y}_i$ là các vector. Các cặp dữ liệu biết trước $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbf{X} \times \mathbf{Y}$ được gọi là tập dữ liệu huấn luyện. Từ dữ liệu huấn luyện này chúng ta cần tạo ra một hàm số ánh xạ mỗi phần tử từ tập \mathbf{X} sang một phần tử (xấp xỉ) tương ứng của tập \mathbf{Y} :

$$\mathbf{y}_i \approx \mathbf{f}(\mathbf{x}_i), \forall i = 1, 2, \dots, N$$

Mục đích là xấp xỉ hàm số \mathbf{f} thật tốt để khi có một dữ liệu \mathbf{x} mới, chúng ta có thể tính được nhãn tương ứng của nó $\mathbf{y} = \mathbf{f}(\mathbf{x})$.

Học không giám sát (Unsupervised Learning) là chúng ta không biết được đầu ra mà chỉ có dữ liệu đầu vào thôi. Thuật toán học không giám sát sẽ dựa vào cấu trúc của dữ liệu để thực hiện một công việc nào đó để mô hình hóa cấu trúc nền tảng hoặc sự phân bố trong dữ liệu từ đó hiểu rõ hơn về dữ liệu

4.2. Mạng nơ-ron và các mô hình mạng nơ-ron

4.2.1. Mạng nơ-ron nhân tạo

Mạng nơ-ron nhân tạo (ANN) là một mô hình lập trình được lấy cảm hứng từ mạng nơ-ron thần kinh bao gồm số lượng lớn các nơ-ron thần kinh được gắn kết để xử lý thông tin. ANN đơn giản nhất là Perceptron, chỉ bao gồm một lớp nơ-ron là bộ phân loại tuyến tính, tức là một thuật toán phân loại đầu vào bằng cách tách hai loại bằng một đường thẳng. Tuy nhiên, không phải tất cả các nhiệm vụ phân loại đều có thể

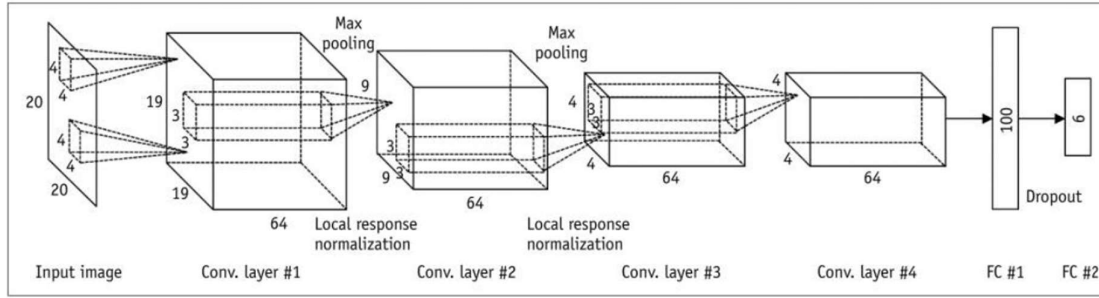
phân tách tuyến tính. Bằng cách thêm các lớp nơ ron, mô hình được gọi là Perceptron nhiều lớp (Multi-layered Perceptron - MLP) thường được áp dụng trong các tình huống mà dữ liệu không thể phân tách tuyến tính. Các lớp trung gian được gọi là các lớp ẩn và có thể có số lượng các nơ-ron tùy ý. MLP có thể huấn luyện một cách hiệu quả dựa trên một quy trình gọi là lan truyền ngược (backpropagation) hay còn gọi là quy tắc chuỗi trong tính toán đạo hàm. Việc tính toán đạo hàm của một hàm số phức tạp mô tả mối quan hệ của đầu vào và đầu ra là rất quan trọng vì hầu hết các thuật toán tối ưu đều thực hiện thông qua việc tính đạo hàm, Gradient descent là một ví dụ. Việc này giúp cho mạng nơ ron MLP thoát ra sự hạn chế của Perceptron về việc chỉ biểu diễn được các quan hệ tuyến tính. Để biểu diễn các quan hệ phi tuyến tính, phía sau mỗi lớp ẩn là một hàm kích hoạt phi tuyến ví dụ: hàm sigmoid hoặc tanh, ReLU,... Khi mạng có nhiều lớp ẩn nó thường được gọi là mạng nơ-ron sâu (DNN). Hiện nay các mô hình phổ biến được đào tạo từ đầu đến cuối theo cách học có giám sát để đơn giản hóa quá trình đào tạo. Các kiến trúc phổ biến nhất là mạng nơ ron tích chập (CNNs) và mạng nơ ron hồi tiếp (RNNs) nhưng mạng CNNs đang được sử dụng rộng rãi trong phân tích hình ảnh y tế, mặc dù RNNs cũng dần đang được phổ biến.

4.2.2. Mạng nơ-ron tích chập

Mạng tích chập là một trong những mô hình học sâu phổ biến nhất có ảnh hưởng nhiều nhất trong lĩnh vực thị giác máy tính. CNN được ứng dụng nhiều trong các bài toán xử lý ảnh. CNNs tốt hơn MLP thông thường vì nó khả năng trích xuất được đặt trung trong không gian ảnh thông qua các bộ lọc (filter) hai chiều. Hơn nữa, các bộ lọc này nhỏ nên việc lưu trữ và tính toán cũng tốt hơn so với MLP thông thường. Về cơ bản CNN bao gồm một vài lớp tích chập với các hàm kích hoạt phi tuyến được áp dụng ở đầu ra của lớp tích chập. Trong mạng MLP mỗi nơ ron đầu vào được kết nối đến tất cả nơ ron của lớp kế tiếp. Ngược lại, ở mạng CNN mỗi nơ ron trong một lớp chỉ liên kết với một số nơ ron của lớp trước đó chúng liên kết theo cơ chế convolution, layer tiếp theo là kết quả tích chập của layer trước đó nhờ vậy mà ta có được các kết nối cục bộ. Nghĩa là mỗi nơ ron ở lớp tiếp theo sinh ra từ bộ lọc áp đặt lên một vùng ảnh cục bộ của nơ ron lớp trước nó.

Lớp tích chập được thực hiện thông qua các bộ lọc: mỗi bộ lọc cho phép trích xuất một thuộc tính, như tần số, hướng,... từ đó tạo nên các bản đồ thuộc tính. Thông tin được lan truyền theo các lớp từ trước ra sau. Lớp cuối cùng thực hiện đánh giá để đưa ra quyết định ở đầu ra. Mạng CNN thường nhận đầu vào là một ảnh, máy tính nhìn thấy như là một mảng của các pixel số lượng pixel phụ thuộc vào độ phân giải của ảnh thông thường với ảnh màu thì nó sẽ có dạng $H \times W \times D$ (H = chiều dài, W = chiều rộng, D = số kênh màu). Ví dụ một ma trận $6 \times 6 \times 3$ (3 ở đây thể hiện cho 3 giá trị RGB của ảnh), $6 \times 6 \times 1$ là ma trận thể hiện cho ảnh xám với giá trị trong khoảng 0 – 1.

Trong mạng CNN mỗi đầu vào sẽ đi qua một chuỗi các lớp tích chập với các bộ lọc, các kết nối đầy và áp dụng một hàm kích hoạt thích hợp ở đầu ra cho bài toán cụ thể. Hình ảnh bên dưới thể hiện đầy đủ một dòng chảy của một mạng CNN trong việc xử lý dữ liệu đầu vào và phân loại đối tượng dựa trên các giá trị của đầu vào.

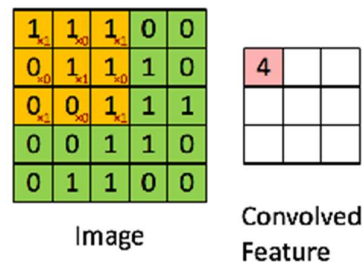


Hình 8: Mạng CNN với nhiều lớp tích chập [8]

4.2.3. Các thành phần của mạng tích chập

a. Lớp tích chập

Lớp tích chập là lớp đầu tiên để trích xuất đặc trưng từ ảnh đầu vào, lớp tích chập thể hiện mối quan hệ giữa các điểm ảnh bằng cách học các thuộc tính của ảnh sử dụng một số bộ lọc nhỏ để tích chập lên ảnh.



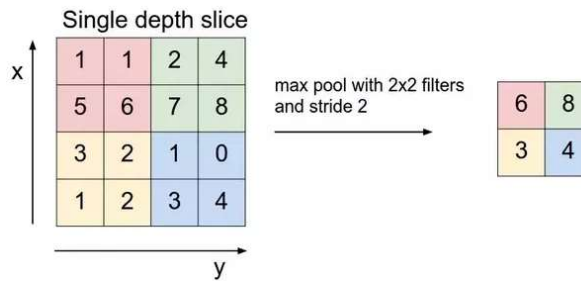
Hình 9: Tính toán tích chập trong CNN [8]

Chúng ta di chuyển mỗi bộ lọc đi qua chiều rộng và chiều dài của ảnh đồng thời tính toán tích chập của bộ lọc và phần ảnh xạ của bộ lọc trên ảnh tại mỗi vị trí mà bộ lọc đi qua, sau khi bộ lọc đi qua toàn bộ ảnh thì chúng ta thu được một bản đồ đặc trưng (feature map). Khi áp dụng các bộ lọc khác nhau sẽ cho kết quả khác nhau tùy thuộc vào mục đích, như bộ lọc phát hiện cạnh (edge detection), làm mờ (blur), danh tính (identity),...

b. Lớp Pooling

Pooling layer dùng để giảm số lượng tham số mô hình khi ảnh đầu vào quá lớn, tổng hợp không gian hay còn được gọi là lấy mẫu con hoặc giảm mẫu là giảm chiều của mẫu mà vẫn giữ lại được những thông tin quan trọng của ảnh đầu vào nhờ vào các loại giảm chiều sau: Max pooling, Average pooling, Sum pooling,...

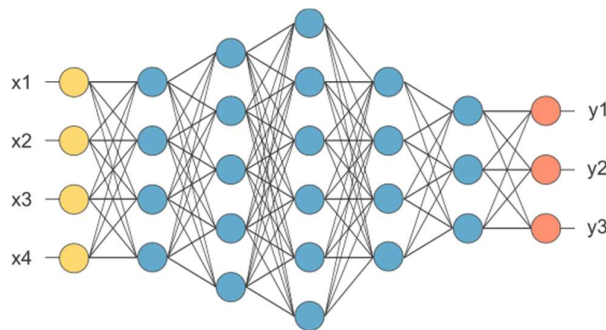
Max pooling lấy giá trị lớn nhất trong khung pooling hay là filter được áp dụng lên bản đồ thuộc tính thu được từ lớp tích chập trước đó, tương tự như là lấy giá trị trung bình, hoặc là lấy tổng.



Hình 10: Lớp pooling trong mạng CNN [8]

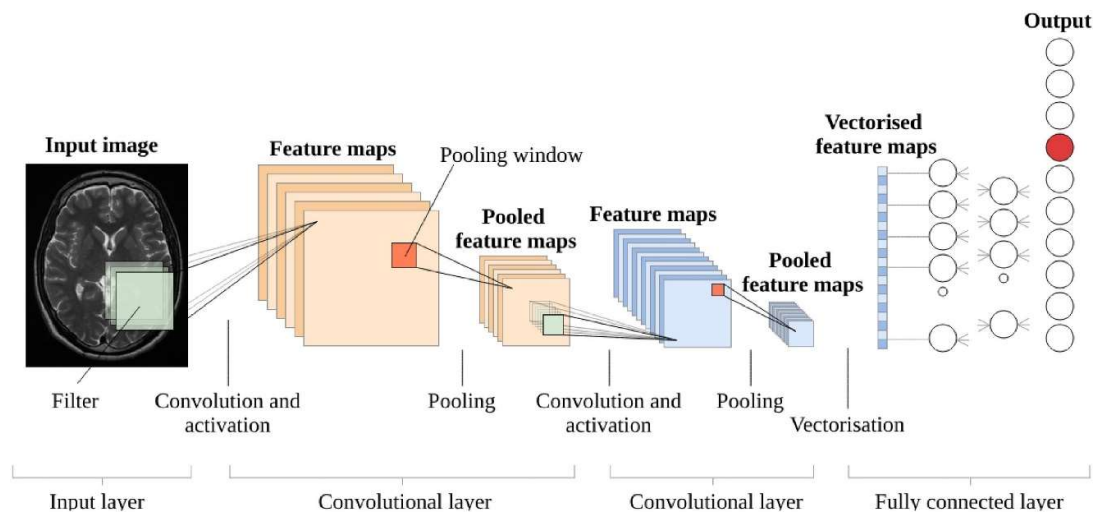
c. Lớp kết nối đầy đủ

Chúng ta làm phẳng các feature map của chúng ta thành một vector và đưa nó vào một lớp kết nối đầy đủ như là một mạng nơ-ron truyền thống.



Hình 11: Lớp kết nối đầy đủ trong CNN [8]

Trong ảnh trên ma trận đặt trưng được làm phẳng như là một vector (x_1, x_2, x_3, \dots) với một lớp kết nối đầy đủ, chúng ta kết hợp chúng lại với nhau để tạo thành một mô hình. Cuối cùng chúng ta có một hàm kích hoạt như là softmax hoặc sigmoid để phân loại đầu ra.



Hình 12: Tổng quan toàn bộ mạng CNN [8]

CHƯƠNG 2 – ỨNG DỤNG KỸ THUẬT SEGMENTATION ĐỂ PHÁT HIỆN KHỐI U TRONG ẢNH MRI NÃO

Trong thực tế hiện tại phân đoạn chuẩn đoán khối u não vẫn được thực hiện thủ công bởi con người. Việc phân đoạn là một công việc mất rất nhiều công sức thông thường phải chuẩn đoán trên từng lát cắt MRI, mỗi một bệnh nhân thì thường sẽ có khoảng 155 lát cắt cho nên bác sĩ phải chuẩn đoán phần lớn tất cả các lát cắt đó và việc chuẩn đoán phụ thuộc vào kinh nghiệm của bác sĩ. Hơn thế nữa kết quả chuẩn đoán khó có thể lặp lại giống nhau thậm chí trên cùng một bác sĩ chuẩn đoán. Do đó rất cần một phương pháp có thể thực hiện việc chuẩn đoán này một cách tự động và chính xác nhất. Sau đây chúng tôi sẽ tiến hành xây dựng một mô hình học sâu có thể phát hiện và chuẩn đoán khối u não trên ảnh MRI.

1. CÁC GIAO THỨC MRI ĐƯỢC SỬ DỤNG

Các giao thức MRI não điển hình và được sử dụng thường xuyên bao gồm T1-weighted, T2-weighted, FLAIR, và T1ce [15] đây là các chuỗi xung MRI thường được sử dụng trong chuẩn đoán hình ảnh y khoa. Với mỗi chuỗi xung có thể phục vụ cho các mục đích chuẩn đoán khác nhau. Ví dụ tín hiệu trên chuỗi FLAIR thì dễ dàng nhất cho việc nhận dạng toàn bộ khối u, phần hoại tử của khối u thì xuất hiện rõ nhất trên chuỗi T2.

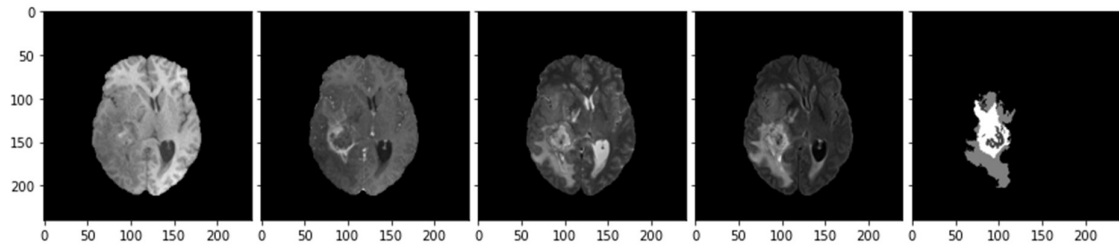
Trong nghiên cứu thực nghiệm này chúng tôi thực nghiệm trên tập dữ liệu BRATS-2018 bao gồm 210 bệnh nhân HGG, 75 bệnh nhân LGG mỗi bệnh nhân đều bao gồm 4 chuỗi xung như trên.

Trong phần tiến hành cài đặt và thử nghiệm chúng tôi sử dụng mô hình U-Net của Hao Dong và cộng sự để tiến hành huấn luyện và kiểm tra trên tập BRATS-2018 bao gồm 210 HGG và 75 LGG.

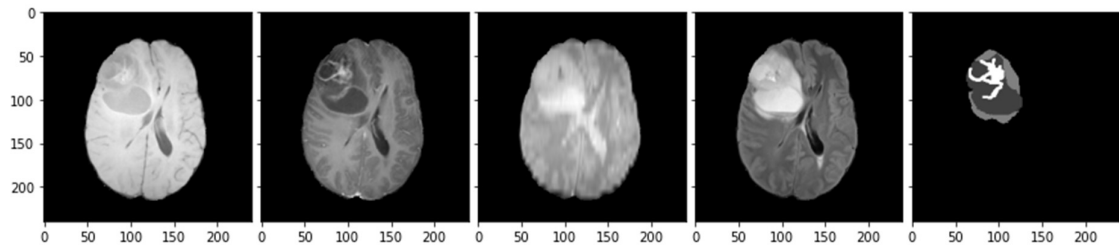
2. GIỚI THIỆU TẬP DỮ LIỆU BRATS-2018

Tập dữ liệu BRATS-2018 được công bố bởi ngành y dược của trường đại học Pennsylvania, trong thách thức phân đoạn hình ảnh đa phương thức 2018. Tất cả các bản sao trong BRATS đều đầy đủ 4 chuỗi xung: T1, T2, T1ce, FLAIR đạt được từ các phát đồ lâm sàng khác nhau và nhiều người quét ($n=19$). Tất cả các dữ liệu đều được gắn nhãn bằng tay bởi 1 đến 4 người xếp hạng và tuân theo cùng một giao thức và các chú thích đã được các bác sĩ lâm sàng có kinh nghiệm xác nhận. Các chú thích bao gồm các khối u tăng cường (GD-enhancing nhãn 4), phù khúc mạc (ED nhãn 2), phần hoại tử và không tăng cường hay là phần lõi (NCR/NET nhãn 1). Trong mô hình huấn luyện này chúng tôi huấn luyện để phát hiện phần toàn bộ khối u có nghĩa là bao gồm cả 3 nhãn. Dữ liệu được cung cấp đã được tiền xử lý trước, tức là được đồng

đăng ký với cùng một mẫu giải phẫu, được nội suy đến cùng một độ phân giải (1mm^3) và đã được tách hộp sọ. Tất cả bệnh nhân HGG và LGG trên mỗi dữ liệu bệnh nhân dùng huấn luyện đều có 155 lát cắt tương tự phân nhân cũng có 155 lát cắt tương ứng.



Hình 13: Các chuỗi xung T1, T1ce, T2, FLAIR của một lát cắt của bệnh nhân HGG



Hình 14: Các chuỗi xung T1, T1ce, T2, FLAIR của một lát cắt của bệnh nhân LGG

3. PHƯƠNG PHÁP THỰC HIỆN

Chúng tôi tiến hành chuẩn hóa dữ liệu cho mỗi chuỗi xung trên mỗi lát cắt bằng cách trừ đi trung bình của chuỗi và chia lại cho độ lệch chuẩn của chuỗi xung đó. Mục đích để chuẩn hóa lại dữ liệu, giảm độ phức tạp trong quá trình tính toán cũng như không gian lưu trữ của bộ nhớ.

Ngoài ra phân đoạn thủ công với 3 lớp: hoại tử (1), phù nề (2), tăng cường (4). Các phân đoạn thủ công được thực hiện trên sử dụng như là nhãn của dữ liệu trên cả tập huấn luyện và tập đánh giá mô hình. Trong cài đặt trước đó các chuỗi xung đã được xếp chồng lên nhau như là đầu vào đa kênh RGB.

Trong cài đặt này chúng tôi sử dụng chuỗi xung FLAIR để nhận dạng khu vực của toàn bộ khối u đã được chứng minh một cách hiệu quả [16]. Làm như vậy bởi vì mô hình yêu cầu ít đầu vào hơn và trong thực tế số lượng chuỗi xung đầy đủ của một bệnh nhân có thể không có sẵn do các triệu chứng và thời gian thu thập dữ liệu.

3.1. Tăng cường dữ liệu

Mục đích của tăng cường dữ liệu là để tăng hiệu suất của mô hình để cho mô hình có thể học được các mẫu khó hơn và nhiều hơn so với các mẫu ban đầu. Mô hình áp dụng một bộ các thao tác tăng cường dữ liệu như bảng 2 bên dưới, một vài sự biến đổi dữ liệu như: lật, xoay, phóng to thu nhỏ, nhưng sẽ không làm thay đổi rất khác so với mẫu ban đầu. Tất cả các phạm vi áp dụng từng thao tác tăng cường dữ liệu bên dưới được áp dụng ngẫu nhiên và không cố định, Nghĩa là ảnh thứ nhất có thể lật dọc và tinh chỉnh độ sáng là 0.8, nhưng ảnh thứ hai có thể là lật ngang và tinh chỉnh độ sáng có thể là 0.9, các thông số còn lại cũng thực hiện tương tự như miêu tả trong ảnh bên dưới:

Phương pháp	Phạm vi
Flip horizontally	50% probability
Flip vertically	50% probability
Rotation	$\pm 20^\circ$
Shift	10% on both horizontal and vertical direction
Shear	20% on horizontal direction
Zoom	$\pm 10\%$
Brightness	$\gamma = 0.8-1.2$
Elastic distortion	$\alpha = 720, \sigma = 24$

Bảng 3: Tổng kết các biến đổi trong tăng cường dữ liệu mô hình [13]

Trong đó γ điều khiển độ sáng, α điều khiển góc độ, σ điều khiển độ nghiêng.

3.2. Xây dựng mô hình

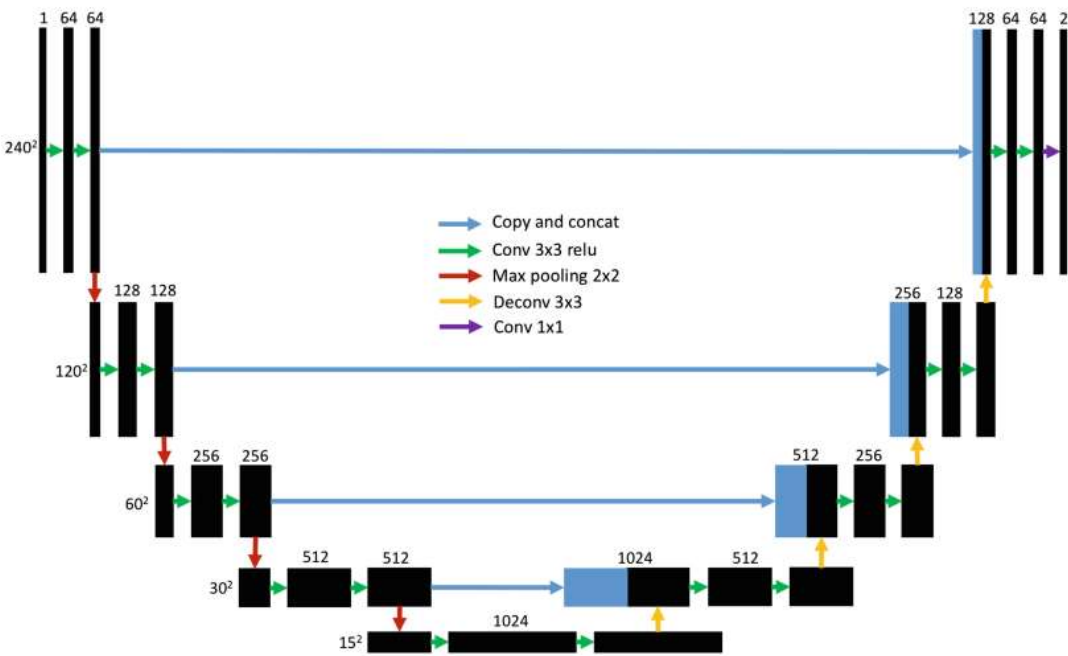
Trong mô hình U-Net chúng tôi sử dụng 5 khối tích chập cho mỗi phần giảm mẫu và tăng mẫu của mạng. Mỗi khối tích chập có 2 lớp tích chập với bộ lọc 3×3 , bước trượt là 1 cho cả 2 chiều và hàm kích hoạt là ReLu. Để tăng số lượng bản đồ đặt trung (feature maps) từ 1 lên 1024. Trong phần giảm mẫu, Max Pooling với bộ lọc 2×2 được áp dụng cho mỗi khối trừ khối cuối cùng, để giảm kích thước của bản đồ đặt trung (feature maps) từ 240×240 xuống 15×15 .

Trong phần tăng mẫu mỗi khối bắt đầu với một lớp có thể gọi là tích chập ngược (deconvolutional) với bộ lọc 3×3 và bước trượt 2×2 cái mà có thể làm gấp đôi kích thước của bản đồ thuộc tính. Vì vậy kích thước bản đồ thuộc tính tăng từ 15×15 ở lớp cuối cùng của phần giảm mẫu lên 240×240 như kích thước đầu vào ban đầu. Trong mỗi khối tích chập của phần tăng mẫu, gồm 2 lớp tích chập được áp dụng để giảm số lượng bản đồ thuộc tính của phần tổng cộng tích chập ngược (deconvolutional) và tích chập xuôi (convolutional) tương ứng của phần giảm mẫu.

Sự khác biệt so với mô hình U-Net ban đầu của Olaf Ronnerbeger, trong quá trình cài đặt chúng tôi sử dụng phần đệm bằng 0, để giữ số chiều đầu ra cho tất cả các lớp tích chập bao gồm cả phần down-sampling và up-sampling.

Cuối cùng sử dụng một lớp tích chập 1×1 giảm số lượng feature map xuống còn 2 để phân biệt phần khối u và phần không phải khối u một cách tương ứng. Không sử

dùng một lớp kết nối đầy đủ nào trong cái đặt kiến trúc mạng này, cũng như là không sử dụng thêm các lớp chuẩn hóa như: dropout hay batch normalization.



Hình 15: Kiến trúc mạng U-Net [13]

Parameters	Value
Number of convolutional blocks	[4-6]
Number of deconvolutional blocks	[4-6]

Bảng 4: Các thông số cài đặt cho mô hình U-Net

3.3. Huấn luyện và tối ưu mô hình

Trong quá trình huấn luyện chúng tôi sử dụng đo lường Soft Dice [18] như là hàm mất mát của mô hình tốt hơn cross-entropy hoặc là sai số bình phương, Soft Dice có thể được coi như là một biến thể của Dice Similarity Coefficient truyền thống [18].

Việc huấn luyện mô hình học sâu yêu cầu cần tối ưu hóa dựa trên gradient ngẫu nhiên để giảm thiểu hàm chi phí liên quan đến các tham số mô hình. Chúng tôi sử dụng Adam như là một hàm tối ưu, Adam có thể coi là tận dụng được các ưu thế của RMSprop và Stochastic Gradient Decent với giá tốc, Adam được chứng minh là hàm tối ưu có ưu điểm vượt trội so với các hàm tối ưu khác [17].

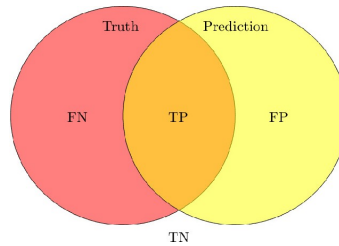
Các tham số của hàm Adam được cài đặt: Tốc độ học = 0.0001, số lượng epoch tối đa = 100, tất cả các tham số được khởi tạo bằng phân phối chuẩn với giá trị trung bình là 0 và độ lệch chuẩn là 0.01, và tất cả bias được khởi tạo là 0.

3.4. Chỉ số đánh giá DSC

DSC (Dice similarity coefficient) là một thống kê được sử dụng để đánh giá mức độ giống nhau của 2 mẫu. Khi ứng dụng cho dữ liệu dạng boolean đã sử dụng các định nghĩa: True Positive (TP), False Positive (FP), và False Negative (FN) nó có thể được viết dưới dạng:

$$DSC = \frac{2TP}{FP + 2TP + FN}$$

Trong đó số lượng TP là số lượng mà mô hình tìm thấy là cụ thể hơn là số điểm ảnh mà mô hình nhận dạng đúng, số lượng FP là số lượng điểm ảnh không phải là khối u nhưng mà mô hình nhận dạng là khối u hay điểm ảnh tích cực. FN là số lượng điểm ảnh mô hình nhận dạng sai không phải là khối u.



Hình 16: Ảnh minh họa DSC [17]

3.5. Kết quả đạt được và thảo luận

Sau khi huấn luyện và kiểm tra trên tập dữ liệu BRATS-2018, sử dụng 79 bệnh nhân HGG và 45 bệnh nhân LGG để huấn luyện mô hình, sau đó dùng 42 bệnh nhân HGG, 15 bệnh nhân LGG để đánh giá mô hình. Tổng cộng huấn luyện trên 19,220 ảnh bao gồm HGG và LGG trên chuỗi xung FLAIR. Mô hình đạt được 91,62 % trên chỉ số đánh giá DSC trên HGG và 92.14 % trên LGG. Bảng 5 bên dưới thống kê các kết quả của mô hình hiện tại so với các mô hình của các đội thi trên bảng xếp hạng BRATS 2018.

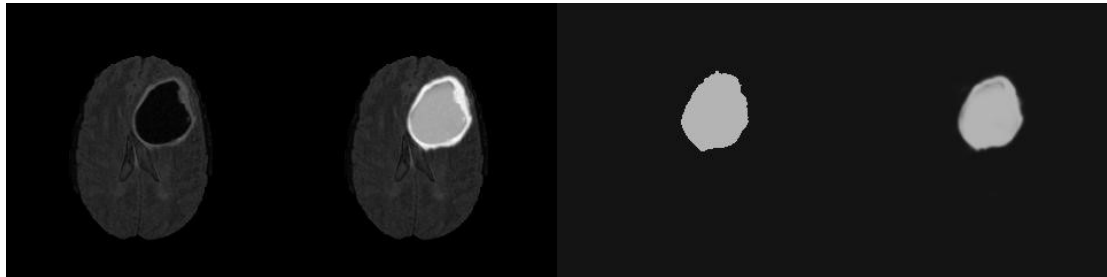
Data	Cases Evaluated	Methods & Team	Whole
BRATST 2018 Training Leaderboard	285	MIC-DKFZ (German Cancer Research Center)	0.90303
	285	DL-86-61 (University of Sydney, Australia)	0.90949
	285	ECONIB	0.92682
	285	Stroer	0.91241
	285	LSA	0.91911
	285	Our Exp	0.92368
BRATST 2018 Testing	57	Our Exp	0.91099

Bảng 5: Kết quả ứng dụng mô hình U-Net

Sau đây là một số kết quả hình ảnh thực nghiệm trong bộ dữ liệu kiểm tra của mô hình. Hình ảnh bao gồm từ trái sang phải là: chuỗi xung Flair là ảnh đầu vào, tiếp theo là phần phản chiếu của khối u mà mô hình dự đoán lên ảnh đầu vào để cho chúng ta có thể dễ dàng hình dung vị trí của khối u trên ảnh đầu vào, ground truth là nhãn đúng của hình ảnh đầu vào, cuối cùng prediction là ảnh mà mô hình dự đoán vị trí và kích cỡ của khối u.

Mỗi hình ảnh bên dưới được trích từ một lát cắt bất kỳ của các bệnh nhân khác nhau trong bộ dữ liệu kiểm tra gồm 45 trường hợp HGG và 15 trường hợp LGG.

Hình 17: Một số kết quả mô hình sự đoán trên HGG



Hình 18: Một số kết quả mô hình sự đoán trên LGG

Mô hình được xây dựng bằng thư viện Tensorflow và Tensorlayer, phần đọc dữ liệu từ tập dữ liệu BRATS 2018 thì được thực hiện bởi thư viện Niblabel. mô hình được huấn luyện trên 35 epochs, và thời gian huấn luyện sắp xỉ 35 giờ trên nền tảng Kaggle sử dụng NVIDIA Titan P100 16GB. Do giới hạn của GPU cho nên số lượng mẫu trong batch_size nhỏ dẫn đến quá trình huấn luyện mất nhiều thời gian. Mô hình mất gần 2 giây để dự đoán trên toàn bộ các lát cắt MRI của một bệnh nhân bất kể là HGG hay là LGG.

Do mô hình được cài đặt dựa trên phiên bản Tensorflow 1.15 và huấn luyện bằng GPU để tăng tốc quá trình huấn luyện chúng ta có thể chuyển đổi chuẩn lưu trữ các ảnh đầu vào từ ma trận của thư viện numpy sang chuẩn định dạng tfrecord của thư viện Tensorflow để có thể sử dụng TPU trên các nền tảng Kaggle hoặc Colab huấn luyện mô hình một cách nhanh chóng.

PHẦN KẾT LUẬN

Trong nghiên cứu thực nghiệm này chúng tôi đã trình bày một các tổng quan ngắn gọn về 2 nghiên cứu ứng dụng học sâu vào xử lý ảnh y khoa cụ thể là bài toán phát hiện và phân đoạn khối u não trên ảnh MRI. Sau đó là trình bày các bước triển khai ứng dụng mô hình U-Net để phát hiện khối u não trên ảnh MRI.

Hai hướng tiếp cận trên có thể được gọi là hiện đại nhất (state-of-the-art) của bài toán phân đoạn hình ảnh y khoa cho đến hiện tại. Cả hai phương pháp tiếp cận đều đạt kết quả đầy hứa hẹn trong chuẩn đoán khối u não.

Trong phần ứng dụng mạng U-Net vào chuẩn đoán phát hiện khối u cũng mang lại kết quả khá tốt tương đương với kết quả mà hai hướng tiếp cận trên mang lại. Mô hình U-Net được huấn luyện và đánh giá trên bộ dữ liệu BRATS-2018 bao gồm cả bệnh nhân HGG và LGG, chúng tôi đã trình bày việc ứng dụng mô hình U-Net vào chuẩn đoán khối u có thể so sánh với sự chuẩn đoán thủ công của các bác sĩ.

Mô hình U-Net được huấn luyện có thể phân đoạn khối u não cho từng bệnh nhân mà không cần sự can thiệp thủ công từ đó cho đánh giá một cách khách quan các nhiệm vụ lâm sàng như chuẩn đoán để lập phát đồ điều trị cho các bệnh nhân.

Hiện tại việc ứng dụng mô hình U-Net để phát hiện khối u não trên ảnh MRI chỉ huấn luyện để phát hiện phần toàn bộ khối u, tiếp theo có thể phát triển và huấn luyện mô hình nhận dạng phần lõi hay là phần hoại tử là phần không lan rộng của khối u (Necrotic hay Core), và phần phù nề của phần lõi khối u (Edema), cuối cùng là phần lan rộng hay phần tăng cường của khối u (Enhancing tumor) dựa trên các chuỗi xung còn lại của tập dữ liệu như là T2 và T1ce, T1.

Do giới hạn phần cứng cũng như thời gian huấn luyện trong tương lai chúng tôi có thể mở rộng phần cứng cũng như thời gian huấn luyện giúp cho mô hình có thể huấn luyện trên nhiều Epoch hơn so với hiện tại để mô hình có thể học tốt hơn các mẫu của dữ liệu.

Bên cạnh đó chúng tôi có thể tinh chỉnh lại các thông số mô hình, ví dụ như: giảm kích thước ảnh đầu vào lại thành 128x128 hoặc là 180x180 để lấy phần trung tâm của ảnh cũng như giảm tham số huấn luyện mô hình và tăng tốc quá trình huấn luyện cũng như đánh giá mô hình.

Tài liệu tham khảo

1. Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2017. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), pp.84-90.
2. Shen, D., Wu, G., Suk, H.-I., Mar. 2017. Deep learning in medical image analysis. *Annu Rev Biomed Eng*
3. Litjens G, Kooi T, Bejnordi BE, Setio AA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. *Medical image analysis*. 2017 Dec 1;42:60-88.
4. Pneumonia. [(accessed on 31 December 2019)]; Available online: <https://www.radiologyinfo.org/en/info.cfm?pg=pneumonia>
5. Hoult, D.I.; Bahkar, B. (1998). "NMR Signal Reception: Virtual Photons and Coherent Spontaneous Emission". *Concepts in Magnetic Resonance*. 9 (5): 277–297. doi:10.1002/(SICI)1099-0534(1997)9:5<277::AID-CMR1>3.0.CO;2-W
6. B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", *IEEE Transactions on Medical Imaging* 34(10), 1993-2024 (2015) DOI: 10.1109/TMI.2014.2377694
7. J. Deng, W. Dong, R. Socher, L. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.
8. Xue Y, Xu T, Zhang H, Long LR, Huang X. Segan: Adversarial network with multi-scale l1 loss for medical image segmentation. *Neuroinformatics*. 2018 Oct 1;16(3-4):383-92.
9. Dong H, Yang G, Liu F, Mo Y, Guo Y. Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. In *annual conference on medical image understanding and analysis* 2017 Jul 11 (pp. 506-517). Springer, Cham.
10. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* 2015 Oct 5 (pp. 234-241). Springer, Cham.

11. Menze et al., The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS), IEEE Trans. Med. Imaging, 2015
12. Sørensen, T. (1948). "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons". Kongelige Danske Videnskabernes Selskab. 5 (4): 1–34
13. Xue Y, Xu T, Zhang H, Long LR, Huang X. SegAN: Adversarial Network with Multi-scale L1 Loss for Medical Image Segmentation. Neuroinformatics. 2018 Oct;16(3-4):383-392. doi: 10.1007/s12021-018-9377-x. PMID: 29725916.
14. Liu, Jin & Li, Min & Wang, Jianxin & Wu, Fangxiang & Liu, Tianming & Pan, Yi. (2014). A Survey of MRI-Based Brain Tumor Segmentation Methods. Tsinghua Science and Technology. 19. 578-595. 10.1109/TST.2014.6961028.
15. Soltaninejad, M., Yang, G., Lambrou, T., Allinson, N., Jones, T.L., Barrick, T.R., Howe, F.A., Ye, X.: Automated brain tumour detection and segmentation using superpixel-based extremely randomized trees in FLAIR MRI. Int. J. Comput. Assist. Radiol. Surg. 12(2), 183–203 (2016)
16. Kingma, D., Ba, J.: Adam: A Method for Stochastic Optimization (2014)
17. Ek, Erika. "Tissue segmentation of head and neck for hyperthermia treatment planning." (2018).
18. Milletari F, Navab N, Ahmadi SA. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In 2016 fourth international conference on 3D vision (3DV) 2016 Oct 25 (pp. 565-571). IEEE.