

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TPHCM
KHOA CÔNG NGHỆ THÔNG TIN**



BÁO CÁO ĐỀ TÀI

**PHÂN ĐOẠN KHỐI U GLIOMA TỪ MRI ĐA
KÊNH BẰNG U-NET**

GIẢNG VIÊN : ThS.Võ Quang Hoàng Khang

SINH VIÊN THỰC HIỆN 1: Phan Mạnh Tân

MÃ SỐ SINH VIÊN: 21114061

SINH VIÊN THỰC HIỆN 2: Nguyễn Văn Thương

MÃ SỐ SINH VIÊN: 21121741

TP HỒ CHÍ MINH, tháng 12 năm 2025

INDUSTRIAL UNIVERSITY OF HO CHI MINH CITY
FACULTY OF INFORMATION TECHNOLOGY



PROJECT REPORT

**GLIOMA BRAIN TUMOR SEGMENTATION
ON MULTIMODAL MRI USING A U-NET
MODEL**

SUPERVISOR: MSc. Vo Quang Hoang Khang

STUDENT NAME: Phan Manh Tan

STUDENT CODE: 21114061

STUDENT NAME: Nguyen Van Thuong

STUDENT CODE: 21121741

HO CHI MINH CITY, month 12 year 2025

ABSTRACT

This study proposes an improved deep learning model for the automatic segmentation of brain tumor sub-regions from Magnetic Resonance Imaging (MRI) scans. This segmentation plays a crucial role in diagnosis and treatment planning, especially for malignant brain tumors like glioma. The model was trained and evaluated on the standard BraTS dataset, using segmentation quality metrics including the Dice coefficient, Sensitivity, Specificity, and IoU (Intersection over Union). Experimental results demonstrate that the model achieves relatively high segmentation accuracy, particularly in critical lesion areas such as the enhancing tumor, peritumoral edema, and necrotic core. The high Dice score reflects the strong similarity between the predicted and ground truth regions, while the IoU indicates a small boundary misalignment, proving the model's capability to accurately reproduce the tumor's shape and position. These findings suggest the model has significant potential for clinical application, contributing to more precise diagnoses and the development of effective treatment protocols by clinicians.

TÓM TẮT

Nghiên cứu này đề xuất một mô hình học sâu cải tiến nhằm hỗ trợ phân vùng tự động các vùng phụ của khối u não từ ảnh cộng hưởng từ (MRI), đóng vai trò quan trọng trong chẩn đoán và lập kế hoạch điều trị, đặc biệt đối với các trường hợp u não ác tính như glioma. Mô hình được huấn luyện và đánh giá trên bộ dữ liệu chuẩn BraTS, với các tiêu chí đánh giá chất lượng phân vùng bao gồm Dice coefficient, Sensitivity, Specificity và IoU. Các kết quả thực nghiệm cho thấy mô hình đạt độ chính xác phân vùng khá cao, đặc biệt ở các vùng tổn thương quan trọng như khối u tăng sinh, vùng phù nề và vùng hoại tử. Độ tương đồng giữa vùng dự đoán và thực tế được phản ánh qua Dice cao, trong khi IoU cho thấy mức sai lệch biên nhỏ, chứng minh khả năng mô hình tái tạo đúng hình dạng và vị trí khối u. Những kết quả này cho thấy mô hình có tiềm năng ứng dụng trong thực tế lâm sàng, góp phần hỗ trợ bác sĩ chẩn đoán chính xác hơn và đưa ra phác đồ điều trị hiệu quả.

LỜI CẢM ƠN

Sau một thời gian nghiên cứu và thực hiện, đồ án tốt nghiệp của chúng em đã hoàn thành. Đây là kết quả của quá trình học tập, rèn luyện và tích lũy kiến thức trong suốt những năm tháng sinh viên tại trường.

Trước tiên, chúng em xin gửi lời cảm ơn chân thành đến ThS. Võ Quang Hoàng Khang, người đã tận tình hướng dẫn, chỉ bảo và định hướng cho chúng em trong suốt quá trình thực hiện đồ án. Những ý kiến đóng góp, nhận xét quý báu của thầy đã giúp chúng em hoàn thiện đồ án một cách tốt nhất.

Chân thành cảm ơn quý thầy cô trong khoa Công nghệ thông tin đã truyền đạt những kiến thức chuyên môn nền tảng, tạo điều kiện thuận lợi để chúng em có thể hoàn thành chương trình học tập cũng như đồ án tốt nghiệp này.

Mặc dù đã cố gắng hoàn thành đồ án với tinh thần trách nhiệm cao nhất, song do thời gian và kinh nghiệm còn hạn chế nên đồ án không tránh khỏi những thiếu sót. Chúng em rất mong nhận được sự đóng góp ý kiến của quý thầy cô để đồ án được hoàn thiện hơn.

Trân trọng!

NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN

TP. Hồ Chí Minh, ngày tháng năm
CHỮ KÝ GIẢNG VIÊN

NHẬN XÉT CỦA GIẢNG VIÊN PHẢN BIỆN 1

TP. Hồ Chí Minh, ngày tháng năm
CHỮ KÝ GIẢNG VIÊN

NHẬN XÉT CỦA GIẢNG VIÊN PHẢN BIỆN 2

TP. Hồ Chí Minh, ngày tháng năm
CHỮ KÝ GIẢNG VIÊN

LỜI MỞ ĐẦU

Trong bối cảnh công nghệ 4.0, học máy và học sâu đang đóng vai trò then chốt trong nhiều lĩnh vực, đặc biệt là y tế. Những tiến bộ vượt bậc trong phân tích hình ảnh và nhận diện đối tượng đã mang lại nhiều cải tiến đột phá trong công tác chẩn đoán và điều trị bệnh, giúp các quy trình y tế trở nên chính xác và hiệu quả hơn.

Một ứng dụng tiêu biểu là phân đoạn hình ảnh y tế, đặc biệt là phân đoạn khối u não từ ảnh MRI. Khối u não là bệnh lý nguy hiểm đòi hỏi chẩn đoán chính xác và kịp thời. Tuy nhiên, việc phân đoạn thủ công không chỉ tốn thời gian mà còn dễ xảy ra sai sót do yếu tố con người.

Nhận thức được tầm quan trọng của vấn đề này, đồ án tốt nghiệp của chúng em sẽ tập trung vào việc phân đoạn khối u não tự động từ ảnh MRI thông qua các mô hình học sâu tiên tiến. Dự án sử dụng tập dữ liệu BraTS 2020 và triển khai các kiến trúc mạng nơ-ron hiện đại, tích hợp cơ chế attention và các lớp tích chập sâu nhằm nâng cao khả năng phân đoạn chính xác. Kết quả nghiên cứu hứa hẹn trở thành công cụ hỗ trợ đắc lực cho bác sĩ trong quá trình chẩn đoán và điều trị, góp phần nâng cao chất lượng chăm sóc sức khỏe cho người bệnh.

MỤC LỤC

LỜI MỞ ĐẦU	1
MỤC LỤC HÌNH ẢNH	4
DANH MỤC CÁC THUẬT NGỮ VIẾT TẮT	5
CHƯƠNG 1: TỔNG QUAN	7
1.1 Giới thiệu đề tài	7
1.2 Mục tiêu nghiên cứu	10
1.2.1 Mục tiêu tổng quát	10
1.2.2 Mục tiêu cụ thể	11
1.2.3 Câu hỏi nghiên cứu	12
1.3 Phạm vi nghiên cứu	12
1.3.1 Phạm vi dữ liệu	12
1.3.2 Phạm vi thuật toán và mô hình	13
1.3.3 Giới hạn nghiên cứu	14
1.4 Ý nghĩa khoa học và thực tiễn của đề tài	15
1.4.1 Ý nghĩa khoa học	15
1.4.2 Ý nghĩa ứng dụng lâm sàng	16
1.5 Cấu trúc bão cáo	18
CHƯƠNG 2. CƠ SỞ LÝ THUYẾT	19
2.1 Kiến thức về ảnh MRI	19
2.1.1 Nguyên lý tạo ảnh cộng hưởng từ	19
2.1.2 Các chuỗi ảnh MRI phổ biến: T1, T1ce, T2, FLAIR	20
2.1.3 Đặc điểm hình thái u não trên các chuỗi ảnh	22
2.2 Bài toán phân vùng ảnh y sinh	25
2.3 Học sâu trong xử lý ảnh y tế	35
2.4 Tập dữ liệu BraTS	49
2.4.1 Giới thiệu cuộc thi BraTS	49
2.4.2 Dữ liệu và nhãn phân vùng (WT, TC, ET)	50
2.4.3 Các chuẩn đánh giá trong BraTS	51
2.5 Các chỉ số đánh giá mô hình phân vùng	52
CHƯƠNG 3. PHƯƠNG PHÁP NGHIÊN CỨU	59
3.1 Chuẩn bị dữ liệu	59
3.2 Mô Hình Đề Xuất	60
3.2.1 Kiến Trúc UNet2D	60
3.2.2 Thành phần cơ bản: Khối DoubleConv2d	62
3.2.3 Chi tiết cấu hình mô hình	63
3.3 Hàm Mát Mát và Độ Đo Hiệu Suất	64
3.4 Cấu Hình Huấn Luyện	66
3.4.1 Tối Ưu Hóa và Lịch Học	66

3.4.2 Cấu Hình Huấn Luyện	67
3.5 Chi Tiết Triển Khai	67
3.5.1 Thiết kế luồng dữ liệu (Data Pipeline Design)	67
3.5.2 Nền tảng công nghệ và Thư viện	68
3.6 Thuật toán huấn luyện (Training Algorithm)	68
3.7 Tóm Tắt Phương Pháp Đề Xuất	69
CHƯƠNG 4. THỰC NGHIỆM VÀ ĐÁNH GIÁ	70
4.1 Thiết lập thực nghiệm	70
4.2 Mô tả bộ dữ liệu và phân chia dữ liệu	71
4.3 Kết quả thực nghiệm	74
4.4 Phân tích và thảo luận	84
4.4.1 Phân tích hiệu suất mô hình	84
4.4.2 Đánh giá tác động của từng thành phần mô hình	84
4.4.3 Các yếu tố ảnh hưởng đến chất lượng phân vùng	85
4.4.4 Hạn chế của mô hình trong thực tế	85
4.5 Tóm tắt chương	86
4.5.1 Tổng kết kết quả đạt được	86
4.5.2 Định hướng cải thiện trong chương tiếp theo	86
CHƯƠNG 5 KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN	88
5.1 Kết luận chung	88
5.2 Hạn chế của đề tài	90
5.3 Hướng phát triển trong tương lai	92
5.4 Kết luận chương	96
TÀI LIỆU THAM KHẢO	97
KẾ HOẠCH THỰC HIỆN	99
NHẬT KÝ LÀM VIỆC	100

MỤC LỤC HÌNH ẢNH

Hình 1. Ảnh Cộng hưởng Từ đa phương thức (Multi-modal MRI)

Hình 2. Sơ đồ kiến trúc của mô hình U-Net

Hình 3. Kiến trúc mô hình U-net2D

Hình 4. Ảnh phân vùng minh họa

Hình 5. Kết quả thực nghiệm

Hình 6. Biểu đồ thể hiện Loss

Hình 7. Biểu đồ thể hiện Mean Dice

Hình 8. Biểu đồ thể hiện Validation Dice

Hình 9. Biểu đồ thể hiện Training

Hình 10. Kết quả đánh giá định lượng theo từng lớp

DANH MỤC CÁC THUẬT NGỮ VIẾT TẮT

Các từ tắt	Giải nghĩa từ tắt	Giải thích nghĩa
MRI	Magnetic Resonance Imaging	Phương pháp MRI chẩn đoán hình ảnh dựa trên từ trường và sóng radio cực kì mạnh mẽ, cho phép tạo ra những ảnh rất chi tiết về cấu trúc bên trong là các mô mềm như não, tủy sống, cơ và các cơ quan bên trong cơ thể.
BraTS	Brain Tumor Segmentation	Dữ liệu BraTS bao gồm nhiều lần chụp MRI đa phương thức của bệnh nhân bị (GBM/HGG) và p (LGG), với chẩn đoán xác nhận và (OS). Dữ liệu này xài cho thử thách phân đoạn khối u não, với các bản cập nhật từ BraTS 2019, bao gồm nhiều lần chụp MRI 3T và nhãn sự thật cơ bản do bác sĩ chuyên khoa xác nhận.
GBM	Glioblastoma Multiforme	Một loại u bào nguyên thán kinh não ác tính, phát triển nhanh và thường gấp nhất trong các khối u não của người trưởng thành. Đây là một loại u não cấp độ IV khó điều trị nhất.
HGG	High-Grade Glioma	Một nhóm các loại u não nguyên bào thán kinh đệm có mức độ ác tính cao. Quy trình điều trị rất khó do các u này phát triển nhanh chóng và có xu hướng xâm lấn vào các mô não xung quanh.
TC	Tumor Core	Bao gồm vùng khối u hoại tử và không tăng cường (NCR/NET) và vùng tăng cường (ET).
WT	Whole Tumor	Bao gồm tất cả các vùng khối u, gồm ET, TC, và phù nề (ED).
ED	Peritumoral Edema	Vùng mô xung quanh khối u bị sưng, được đánh dấu là nhãn 2.
T1	T1-weighted MRI	Loại MRI với thời gian thư giãn dọc, thường dùng để quan sát cấu trúc giải phẫu.

T1c/T1ce	T1-weighted Contrast-Enhanced MRI	Hình ảnh T1 sau khi tiêm chất tương phản (Gadolinium), giúp làm nổi bật các vùng khối u tăng cường.
T2	T2-weighted MRI	Loại MRI với thời gian thư giãn ngang, thường dùng để phát hiện phù nề và các bất thường mô.
FLAIR	Fluid-Attenuated Inversion Recovery	Loại MRI loại bỏ tín hiệu từ dịch não tủy, giúp phát hiện phù nề và khối u rõ hơn.
GD	Gadolinium	Chất tương phản được dùng trong MRI làm nổi bật các vùng khối u tăng cường.
IoU	Intersection over Union	Một chỉ số để đo lường mức độ tương đồng giữa kết quả dự đoán và nhãn gốc, thường áp dụng trong các bài toán phân đoạn ảnh.
SEG	Segmentation	Tệp chứa nhãn phân đoạn (segmentation mask) cho các vùng khối u.
PACS	Picture Archiving and Communication System	Hệ thống lưu trữ và truyền tải hình ảnh y tế, giúp các cơ sở y tế quản lý, lưu trữ và chia sẻ các hình ảnh chẩn đoán như X-quang, CT scan, MRI, siêu âm, và các loại hình ảnh y tế khác.
GLCM	Gray-Level Co-occurrence Matrix	Một kỹ thuật trong xử lý ảnh dùng để mô tả các đặc tính kết cấu của một hình ảnh, đặc biệt là trong các ứng dụng phân tích ảnh, nhận dạng mẫu, và phân loại
SVM	Support Vector Machine	Một thuật toán machine learning có giám sát dùng cho: Phân loại (classification), Hồi quy (regression), Phát hiện bất thường (outlier detection)

CHƯƠNG 1: TỔNG QUAN

1.1 Giới thiệu về tài

1.1.1 Bối cảnh và tầm quan trọng của bài toán phân vùng u não

U não là một trong những khối u ác tính nhất ở người. Nó chiếm khoảng 1,35% tổng số các khối u ác tính và 29,5% số ca tử vong liên quan đến ung thư [1]. Các khối u não và hệ thần kinh trung ương (CNS) bao gồm các khối u ở não, dây thần kinh sọ, dây thần kinh cột sống, tủy sống và màng não. Khối u có thể được phân loại rộng rãi thành u ác tính và u không ác tính (hoặc lành tính). Phân loại của Tổ chức Y tế Thế giới (WHO) quy định một hệ thống phân loại cấp độ, trong đó cấp độ III/IV là ác tính hoặc cấp độ cao [2]. U não là một nhóm đa dạng các khối u ác tính với các loại u não nguyên phát khác nhau hoặc ung thư di căn. Các khối u não ác tính phổ biến nhất là u nguyên bào thần kinh đệm (glioblastoma) có nguồn gốc từ các tế bào thần kinh đệm [3]. U não di căn (MBTs) chiếm phần lớn các khối u nội trực ở bệnh nhân người lớn. Ước tính có tới một phần ba số bệnh nhân được chẩn đoán mắc bệnh ác tính nguyên phát sẽ phát triển các tổn thương di căn hệ thần kinh trung ương trong quá trình mắc bệnh [4]. Các khối u hệ thần kinh trung ương ở trẻ em (CNS) là bệnh ác tính phổ biến thứ hai ở trẻ em và là khối u rắn phổ biến nhất ở trẻ em [5]. Chẩn đoán và điều trị u não sớm là bắt buộc để ngăn ngừa tổn thương não vĩnh viễn hoặc tử vong ở bệnh nhân. Ở cấp độ phân tích dữ liệu y tế, quy trình lựa chọn đặc trưng và phân loại là những quy trình được sử dụng rộng rãi để xác định dữ liệu bệnh nhân là bình thường hay bất thường [6]. Một khi khối u được phát hiện dưới kính hiển vi, thường là đã quá muộn để tiên lượng điều trị hiệu quả ở bệnh nhân có tương quan với giai đoạn bệnh tại thời điểm phát hiện, và do đó, điều quan trọng là phải tìm ra các dấu hiệu cho phép phát hiện sớm khối u. Các lựa chọn điều trị cho bệnh nhân u não di căn bao gồm corticosteroid, phẫu thuật, hóa trị, xạ trị toàn bộ não và phẫu thuật bức xạ lập thể [7]. Bệnh nhân u não phải đối mặt với một vấn đề đáng kể gọi là rối loạn chức năng thần kinh nhận thức. Để chẩn đoán rối loạn chức năng thần kinh nhận thức ở u não cần có các chiến lược mới để sớm bắt đầu phục hồi chức năng thần kinh nhận thức thích hợp.

Theo Đài quan sát Ung thư Toàn cầu (GLOBOCAN), tỷ lệ mắc các khối u ác tính hệ thần kinh trung ương (CNS) là 3,1 đến 3,9 trên 100.000 người vào năm 2022 [15], với tỷ lệ hiện mắc là 10,1 trên 100.000 người trên toàn thế giới theo báo cáo năm 2020 [9]. U thần kinh đệm (Gliomas) chiếm gần 30% trong tổng số các khối u não nguyên phát và 80% trong tổng số các khối u ác tính; chúng

là nguyên nhân gây ra phần lớn các ca tử vong do khối u não nguyên phát [10, 11].

U thần kinh đệm được cho là phát sinh từ các tế bào gốc thần kinh đệm và được phân loại về mặt mô học thành u tế bào hình sao (astrocytomas), u thần kinh đệm ít nhánh (oligodendroliomas) và u nguyên bào thần kinh đệm (GBMs), dựa trên sự tương đồng về hình thái với các tế bào thần kinh đệm được tìm thấy trong não khỏe mạnh. Việc phân loại sâu hơn được thực hiện dựa trên vị trí khối u và các đặc điểm dị sản (hoạt động phân bào, tăng sinh vi mạch và hoại tử). Sự vắng mặt hay hiện diện của các đặc điểm dị sản được sử dụng để gán cấp độ ác tính từ I đến IV theo án bản thứ tư của phân loại Khối u Hệ thần kinh Trung ương của Tổ chức Y tế Thế giới (WHOCNS4) [12, 14]. Theo WHOCNS4 [12], u thần kinh đệm bao gồm các loại u thần kinh đệm ở nhiều cấp độ khác nhau [u tế bào hình sao lông (cấp độ I), u tế bào hình sao lan tỏa (cấp độ II), u tế bào hình sao dị sản (cấp độ III), GBM (cấp độ IV)], u thần kinh đệm ít nhánh (cấp độ II và III), và nhóm u thần kinh đệm ít nhánh-tế bào hình sao hỗn hợp gây tranh cãi (cấp độ II và III) [11, 12].

Tỷ lệ sống sót khác nhau tùy thuộc vào mô học, với bệnh nhân u tế bào hình sao lông có tỷ lệ sống sót sau 10 năm trên 90%, trong khi đối với GBM, chỉ có 6,8% bệnh nhân đạt mốc sống sót sau 5 năm [15, 16].

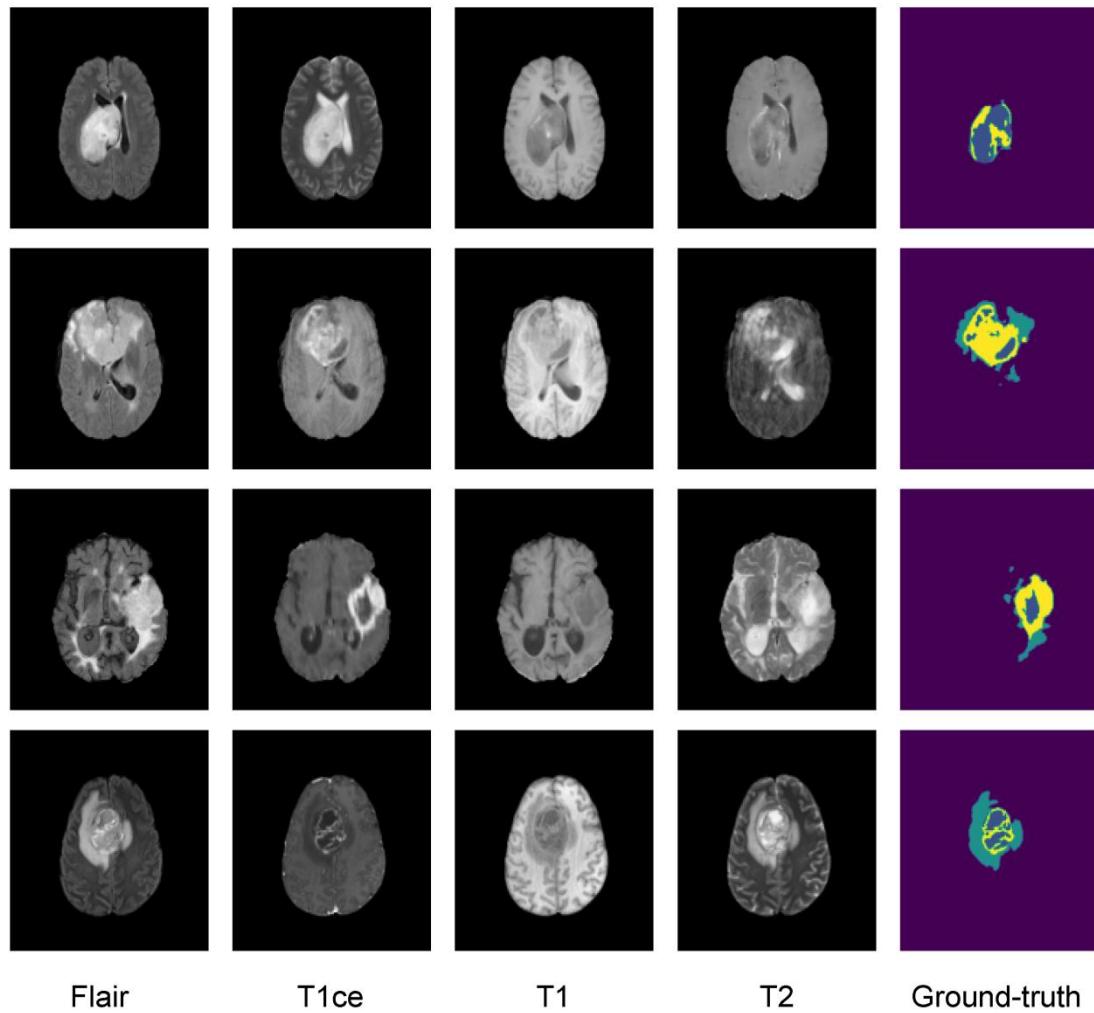
Là một trong những bệnh lý nghiêm trọng với tỷ lệ tử vong cao và ảnh hưởng sâu rộng đến chất lượng cuộc sống người bệnh, việc phân vùng chính xác các vùng khối u não đóng vai trò then chốt trong toàn bộ quy trình điều trị. Quá trình phân vùng này bao gồm việc xác định và phân tách các vùng khối u khác nhau như vùng hoại tử trung tâm (Necrotic Core - NCR), vùng phù nề quanh u (Peritumoral Edema - ED), và vùng u tăng sinh (Enhancing Tumor - ET). Mỗi vùng này có đặc điểm sinh học và ý nghĩa lâm sàng riêng biệt:

Vùng hoại tử trung tâm (NCR): Vùng mô đã chết do thiếu máu cục bộ, thường nằm ở trung tâm khối u, phản ánh mức độ tăng trưởng nhanh và ác tính cao của u.

Vùng phù nề quanh u (ED): Vùng mô não xung quanh khối u bị xâm lấn bởi tế bào u và phản ứng viêm, là chỉ điểm quan trọng cho việc lập kế hoạch xạ trị và đánh giá phạm vi lan rộng của bệnh.

Vùng u tăng sinh (ET): Phần u có hoạt động trao đổi chất mạnh nhất, thường là mục tiêu chính của điều trị và cần được theo dõi sát để đánh giá đáp ứng điều trị.

1.1.2 Vai trò của MRI trong chẩn đoán hình ảnh



Hình 1. Ảnh Cộng hưởng Từ đài phương thức (Multi-modal MRI)

Chụp cộng hưởng từ (Magnetic Resonance Imaging - MRI) là phương pháp chẩn đoán hình ảnh tiêu chuẩn vàng cho các khối u não nhờ khả năng tạo ra hình ảnh có độ tương phản mô mềm cao mà không sử dụng bức xạ ion hóa. Khác với CT scan chỉ cung cấp thông tin cấu trúc giải phẫu đơn thuần, MRI đa chuỗi (multimodal MRI) mang lại thông tin đa chiều về đặc tính sinh lý và sinh hóa của mô [17]. Cụ thể:

T1-weighted (T1): Cung cấp hình ảnh giải phẫu chi tiết với độ phân giải không gian cao, trong đó chất xám và chất trắng được phân biệt rõ ràng. Khối u thường có tín hiệu thấp hơn mô não bình thường.

T1-weighted với chất tương phản Gadolinium (T1Gd/T1ce): Sau khi tiêm chất tương phản, các vùng u có tăng sinh mạch máu (như ET) sẽ hấp thụ chất

tương phản và phát sáng rõ rệt, giúp xác định biên giới khối u và đánh giá mức độ xâm lấn.

T2-weighted (T2): Nhạy cảm với sự thay đổi hàm lượng nước trong mô, cho phép quan sát rõ vùng phù nề (ED) và vùng nang dịch. Khối u và phù nề thường có tín hiệu cao trên ảnh T2.

FLAIR (Fluid-Attenuated Inversion Recovery): Là biến thể của T2 có khả năng triệt tín hiệu dịch não tuy, làm nổi bật các tổn thương gần não thất và vùng phù nề một cách rõ ràng hơn, đặc biệt hữu ích trong phát hiện các tế bào u xâm lấn.

Mỗi chuỗi xung MRI cung cấp một "góc nhìn" khác nhau về cùng một khối u. Sự kết hợp của bốn chuỗi T1, T1ce, T2 và FLAIR tạo nên một bức tranh toàn diện về đặc tính khối u, cung cấp thông tin bổ sung và tương hỗ lẫn nhau. Ví dụ, vùng hoại tử có thể xuất hiện tối trên T1 nhưng sáng trên T2, trong khi vùng u tăng sinh sáng trên T1ce nhưng không rõ trên FLAIR. Chính sự đa dạng thông tin này đòi hỏi một mô hình học sâu có khả năng học và tích hợp đặc trưng từ nhiều nguồn dữ liệu đồng thời.

1.1.3 Xu hướng ứng dụng học sâu trong y sinh

Trong thập kỷ qua, học sâu (deep learning) đã tạo nên một cuộc cách mạng trong lĩnh vực thị giác máy tính y sinh học, đặc biệt là trong các bài toán phân tích hình ảnh y khoa. Khác với các phương pháp máy học truyền thống đòi hỏi trích xuất đặc trưng thủ công dựa trên kiến thức chuyên gia, các mô hình học sâu có khả năng tự động học các biểu diễn phân cấp từ dữ liệu thô, từ các đặc trưng cấp thấp (cạnh, góc, kết cấu) đến các đặc trưng ngữ nghĩa cấp cao (hình dạng, vị trí, mối quan hệ không gian).

1.1.4 Lý do chọn đề tài

Mặc dù có ý nghĩa lâm sàng quan trọng, phân vùng u não thủ công vẫn là công việc đòi hỏi chuyên môn cao và rất tốn thời gian, có thể kéo dài từ 3–5 giờ cho một trường hợp phức tạp. Hơn nữa, sự không đồng nhất giữa các bác sĩ (inter-observer variability) có thể dao động từ 20–28%, ảnh hưởng đáng kể đến độ tin cậy của quá trình chẩn đoán và lên kế hoạch điều trị.

Trong bối cảnh nhu cầu khám chữa bệnh ngày càng tăng và nguồn nhân lực chuyên khoa còn hạn chế, việc phát triển một hệ thống phân vùng u não tự động, nhanh chóng và chính xác dựa trên mô hình học sâu là cần thiết và có ý nghĩa thực tiễn cao.

Do đó, đề tài “**Phân vùng u não trên ảnh MRI đa chuỗi sử dụng mô hình học sâu**” được lựa chọn nhằm xây dựng một giải pháp hỗ trợ chẩn đoán y khoa, góp phần giảm tải cho bác sĩ, chuẩn hóa quy trình phân tích và nâng cao hiệu quả điều trị bệnh nhân u não.

1.2 Mục tiêu nghiên cứu

1.2.1 Mục tiêu tổng quát

Xây dựng và đánh giá một mô hình học sâu dựa trên kiến trúc U-Net 2D để phân vùng tự động các khối u não glioma từ ảnh MRI đa chuỗi, nhằm hỗ trợ bác sĩ trong quá trình chẩn đoán và lập kế hoạch điều trị một cách nhanh chóng, chính xác và khách quan.

1.2.2 Mục tiêu cụ thể

Nghiên cứu này hướng tới đạt được các mục tiêu cụ thể sau:

1. Tiền xử lý và chuẩn hóa dữ liệu BraTS

- Chuyển đổi dữ liệu NIfTI gốc sang định dạng HDF5 để tối ưu hóa tốc độ đọc/ghi.
- Áp dụng chuẩn hóa z-score cho từng chuỗi MRI để đảm bảo tính đồng nhất.
- Thiết kế chiến lược chọn lát cắt thông minh ưu tiên các slice chứa thông tin khối u.
- Xây dựng pipeline augmentation dữ liệu mạnh mẽ bao gồm: flip, rotation, elastic deformation, Gaussian noise, brightness/contrast adjustment, và gamma correction.

2. Thiết kế và triển khai kiến trúc mô hình

- Xây dựng kiến trúc U-Net 2D với 4 kênh đầu vào (T1, T1ce, T2, FLAIR) và 4 kênh đầu ra (Background, NCR, ED, ET)
- Tích hợp Batch Normalization và Dropout để tăng cường khả năng tổng quát hóa.
- Áp dụng Xavier initialization cho các trọng số tích chập.
- Tối ưu hóa độ sâu và số lượng feature maps ở mỗi tầng để cân bằng giữa khả năng biểu diễn và chi phí tính toán

3. Thiết kế hàm mất mát và chiến lược tối ưu hóa

- Xây dựng hàm mất mát kết hợp (combined loss) gồm Cross-Entropy Loss (40%), Dice Loss (40%), và Focal Loss (20%) để xử lý bài toán mất cân bằng lớp.
- Sử dụng class weights để tăng trọng số cho các lớp thiểu số (NCR, ED, ET).
- Áp dụng optimizer AdamW với weight decay và gradient clipping.
- Thiết kế learning rate scheduler với warmup phase và cosine annealing.
- Sử dụng Mixed Precision Training (AMP) để tăng tốc độ huấn luyện và giảm sử dụng bộ nhớ

4. Huấn luyện và kiểm thử mô hình

- Huấn luyện mô hình trên tập dữ liệu BraTS với 700 ca huấn luyện, 200 ca validation.
- Theo dõi và ghi nhận các chỉ số đánh giá: Loss, Dice coefficient, và IoU cho từng epoch.
- Áp dụng early stopping với patience để tránh overfitting.
- Đánh giá hiệu suất trên tập test độc lập gồm 100 ca.
- Phân tích chi tiết kết quả theo từng vùng phân vùng (NCR, ED, ET)

5. So sánh và đánh giá

- So sánh kết quả với các phương pháp baseline và các nghiên cứu liên quan.
- Phân tích điểm mạnh, điểm yếu của mô hình đề xuất.
- Đánh giá khả năng áp dụng thực tế về mặt thời gian suy luận và yêu cầu tài nguyên.

1.2.3 Câu hỏi nghiên cứu

Nghiên cứu này tập trung giải quyết các câu hỏi sau:

RQ1: Liệu kiến trúc U-Net 2D có thể đạt được độ chính xác phân vùng chấp nhận được (Dice coefficient ≥ 0.75) trên bài toán phân vùng u não đa lớp khi chỉ sử dụng thông tin 2D?

RQ2: Hàm mất mát kết hợp (Cross-Entropy + Dice + Focal) có hiệu quả hơn so với việc sử dụng một hàm mất mát đơn lẻ trong việc xử lý vấn đề mất cân bằng lớp trong dữ liệu BraTS không?

RQ3: Chiến lược augmentation dữ liệu nào (geometric, intensity, hoặc kết hợp) mang lại hiệu quả tốt nhất trong việc cải thiện khả năng tổng quát hóa của mô hình?

RQ4: Mô hình U-Net 2D đạt hiệu suất tốt nhất ở vùng phân vùng nào (NCR, ED, hay ET) và lý do cho sự khác biệt này là gì?

RQ5: Với tài nguyên tính toán hạn chế (GPU phổ thông), mô hình 2D có phải là lựa chọn tối ưu hơn so với các phương pháp 3D phức tạp không?

1.3 Phạm vi nghiên cứu

1.3.1 Phạm vi dữ liệu

Nghiên cứu sử dụng bộ dữ liệu **BraTS2023 part 1** và **BraTS2021**, một trong những bộ dữ liệu chuẩn và được sử dụng rộng rãi nhất trong lĩnh vực phân vùng u não. Phạm vi dữ liệu cụ thể bao gồm:

Nguồn dữ liệu:

- Dữ liệu MRI đa chuỗi từ BraTS Challenge, bao gồm các ca bệnh glioma từ nhiều trung tâm y tế quốc tế

- Mỗi ca bệnh bao gồm 4 chuỗi MRI: T1, T1ce (T1 with contrast enhancement), T2, và FLAIR
- Dữ liệu đã được các chuyên gia lâm sàng chú thích (annotated) và kiểm tra chất lượng

Đặc điểm kỹ thuật:

- Kích thước ảnh: $240 \times 240 \times 155$ voxels cho mỗi chuỗi
- Độ phân giải không gian: $1\text{mm} \times 1\text{mm} \times 1\text{mm}$
- Định dạng: NIfTI (.nii.gz) được chuyển đổi sang HDF5 (.h5)
- Nhãn phân vùng: 4 lớp (0: Background, 1: NCR, 2: ED, 3: ET)

1.3.2 Phạm vi thuật toán và mô hình

Nghiên cứu tập trung vào các khía cạnh kỹ thuật sau:

Kiến trúc mô hình:

Kiến trúc chính: U-Net 2D với encoder-decoder và skip connections

Số tầng: 4 tầng encoder và 4 tầng decoder

Số lượng feature maps: (64, 128, 256, 512) ở các tầng tương ứng

Các kỹ thuật regularization: Batch Normalization, Dropout (0.1)

Hàm mất mát:

- Cross-Entropy Loss với class weights (0.1, 2.0, 2.0, 3.0)
- Dice Loss cho bài toán phân vùng
- Focal Loss để xử lý hard examples
- Tỷ lệ kết hợp: 0.4 CE + 0.4 Dice + 0.2 Focal

Kỹ thuật tối ưu hóa:

- Optimizer: AdamW với learning rate ban đầu $1e-3$, weight decay $1e-4$
- Learning rate scheduler: Cosine annealing với 3 epochs warmup
- Gradient clipping với $\text{max_norm} = 1.0$
- Mixed Precision Training (FP16/FP32)
- Batch size: 16
- Số epochs: 50 với early stopping (patience = 10)

Augmentation dữ liệu:

- Geometric transformations: horizontal/vertical flip, rotation (90° , 180° , 270°)

- Intensity transformations: Gaussian noise ($\sigma = 0.03$), brightness/contrast adjustment (0.9-1.1), gamma correction (0.85-1.15)
- Xác suất áp dụng mỗi augmentation: 50%

Metrics đánh giá:

- Dice Coefficient (tổng thể và từng lớp)
- Intersection over Union (IoU) (tổng thể và từng lớp)
- Training/Validation Loss

1.3.3 Giới hạn nghiên cứu

Nghiên cứu có những giới hạn sau đây:

Giới hạn về phương pháp:

- Mô hình chỉ xử lý từng lát cắt 2D độc lập, không khai thác đầy đủ thông tin không gian 3D giữa các slice liền kề.
- Thiếu cơ chế attention hoặc gating để tập trung vào các vùng quan trọng.
- Chưa tích hợp các kỹ thuật tiên tiến như Transformer, Vision Transformer (ViT), hoặc các biến thể U-Net hiện đại (U-Net++, Attention U-Net, nnU-Net).
- Không sử dụng ensemble hoặc test-time augmentation để tăng độ chính xác.

Giới hạn về dữ liệu: Chỉ thực nghiệm trên dữ liệu BraTS (glioma), chưa đánh giá trên các loại u não khác như meningioma, adenoma tuyến yên, hoặc metastasis. Dữ liệu BraTS đã được tiền xử lý và chuẩn hóa bởi ban tổ chức, có thể không phản ánh đầy đủ tính đa dạng của dữ liệu thực tế từ các bệnh viện. Chưa xem xét các yếu tố lâm sàng bổ sung như tuổi, giới tính, tiền sử bệnh, kết quả xét nghiệm sinh học phân tử.

Giới hạn về tài nguyên: Thí nghiệm được thực hiện trên nền tảng Kaggle với GPU hạn chế (Tesla P100 hoặc T4), ảnh hưởng đến khả năng thử nghiệm các mô hình phức tạp hơn. Thời gian huấn luyện bị giới hạn bởi quota của Kaggle, không cho phép thực hiện nhiều thí nghiệm hyperparameter tuning. Không có khả năng so sánh trực tiếp với các mô hình 3D tiên tiến do hạn chế VRAM.

Giới hạn về đánh giá lâm sàng: Kết quả chưa được xác thực bởi các bác sĩ chuyên khoa hình ảnh học. Chưa triển khai thử nghiệm tại môi trường lâm sàng thực tế. Chưa đánh giá tác động của mô hình đến quy trình làm việc và quyết định điều trị của bác sĩ.

Giới hạn về khả năng tổng quát hóa: Mô hình được huấn luyện chủ yếu trên dữ liệu từ các thiết bị MRI hiện đại với chất lượng hình ảnh cao. Chưa kiểm tra hiệu suất trên dữ liệu từ các thiết bị MRI cũ hoặc có chất lượng thấp hơn. Chưa đánh giá khả năng hoạt động trên các giao thức quét MRI khác nhau.

1.4 Ý nghĩa khoa học và thực tiễn của đề tài

1.4.1 Ý nghĩa khoa học

Nghiên cứu này đóng góp vào cộng đồng khoa học theo nhiều khía cạnh:

1. Đóng góp về phương pháp luận:

Nghiên cứu cung cấp một phân tích chi tiết và có hệ thống về hiệu quả của kiến trúc U-Net 2D trong bài toán phân vùng não đa lớp, một góc nhìn thường bị bỏ qua trong bối cảnh các mô hình 3D phức tạp ngày càng phổ biến. Điều này quan trọng vì:

- Chứng minh rằng với thiết kế hợp lý và kỹ thuật huấn luyện tối ưu, mô hình 2D đơn giản vẫn có thể đạt được hiệu suất cạnh tranh với các phương pháp phức tạp hơn
- Cung cấp baseline mạnh mẽ cho các nghiên cứu tương lai, giúp đánh giá chính xác hơn giá trị gia tăng của các cải tiến kiến trúc mới
- Đặt câu hỏi về sự đánh đổi giữa độ phức tạp mô hình và hiệu suất thực tế, khuyến khích sự cân nhắc kỹ lưỡng về "Occam's Razor" trong thiết kế mô hình

2. Nghiên cứu về hàm mất mát kết hợp:

Việc kết hợp ba loại loss functions (Cross-Entropy, Dice, Focal) với tỷ lệ cụ thể (0.4:0.4:0.2) mang lại những hiểu biết sâu sắc về:

- Cách các hàm mất mát khác nhau tương tác và bổ sung cho nhau trong quá trình tối ưu hóa
- Cross-Entropy Loss giúp mô hình học phân biệt cấp pixel, đặc biệt hiệu quả cho lớp background
- Dice Loss trực tiếp tối ưu hóa metric đánh giá chính, cải thiện overlap giữa prediction và ground truth
- Focal Loss tập trung vào các hard examples, giúp mô hình học tốt hơn ở các vùng ranh giới khó phân biệt
- Nghiên cứu cung cấp bằng chứng thực nghiệm về hiệu quả của việc sử dụng class weights khác nhau để xử lý vấn đề mất cân bằng lớp nghiêm trọng trong dữ liệu y khoa

3. Phân tích đặc tính dữ liệu MRI đa chuỗi:

Nghiên cứu đi sâu vào việc hiểu cách mô hình học và kết hợp thông tin từ bốn chuỗi MRI khác nhau:

- Phân tích mức độ đóng góp của từng chuỗi (T1, T1ce, T2, FLAIR) vào hiệu suất phân vùng của từng lớp
- Khám phá các mẫu tương quan (correlation patterns) giữa các chuỗi MRI và cách chúng ảnh hưởng đến quá trình học của mạng

- Cung cấp insight về việc có thể giảm số lượng chuỗi MRI cần thiết trong một số trường hợp để giảm thời gian quét và chi phí

4. Đánh giá chi tiết theo từng vùng phân vùng:

Thay vì chỉ báo cáo Dice coefficient trung bình, nghiên cứu phân tích sâu hiệu suất của mô hình trên từng vùng (NCR, ED, ET):

NCR: Vùng khó nhất do kích thước nhỏ và ranh giới không rõ ràng với vùng cystic.

ED: Hiệu suất trung bình do tính chất diffuse và ranh giới mờ với mô não bình thường.

ET: Hiệu suất cao nhất nhờ contrast tốt trên T1ce và đặc điểm hình ảnh rõ ràng.

Phân tích này giúp hiểu rõ hơn về thách thức cụ thể của từng vùng và hướng cải thiện trong tương lai.

5. Nghiên cứu về augmentation strategies:

Nghiên cứu đóng góp vào việc hiểu rõ hơn về:

- Hiệu quả của các loại augmentation khác nhau (geometric vs. intensity) trong bối cảnh dữ liệu y khoa
- Xác suất áp dụng tối ưu cho mỗi loại augmentation
- Tác động của augmentation đến khả năng tổng quát hóa và tốc độ hội tụ

6. Đóng góp về tính tái lập (Reproducibility):

Nghiên cứu đặt nền tảng cho tính tái lập khoa học bằng cách: Cung cấp code nguồn mở với documentation chi tiết. Mô tả đầy đủ hyperparameters, random seeds, và môi trường thực nghiệm. Sử dụng bộ dữ liệu công khai (BraTS) để các nhà nghiên cứu khác có thể so sánh trực tiếp. Báo cáo không chỉ kết quả tốt nhất mà còn phân tích các thí nghiệm thất bại và bài học kinh nghiệm

1.4.2 Ý nghĩa ứng dụng lâm sàng

Nghiên cứu mang lại những giá trị thực tiễn quan trọng cho hệ thống y tế:

1. Giảm tải công việc cho bác sĩ

Thời gian phân vùng thủ công: 3-5 giờ → Thời gian suy luận tự động: <30 giây cho toàn bộ khối 3D (155 slices). Giải phóng thời gian cho bác sĩ tập trung vào các công việc phức tạp hơn như lập kế hoạch điều trị, tư vấn bệnh nhân. Giảm thiểu mệt mỏi và sai sót do con người, đặc biệt trong các ca cấp cứu hoặc khi khối lượng công việc cao.

2. Tăng tính khách quan và nhất quán:

Loại bỏ sự khác biệt inter-observer (20-28%) trong đánh giá thủ công. Đảm bảo tiêu chuẩn phân vùng nhất quán giữa các bác sĩ, bệnh viện, và theo thời gian. Cung cấp "second opinion" tự động để bác sĩ đối chiếu và xác nhận.

3. Hỗ trợ lập kế hoạch điều trị chính xác

- **Phẫu thuật:** Xác định chính xác biên giới khối u giúp phẫu thuật viên lập kế hoạch đường mổ, tối thiểu hóa tổn thương mô não bình thường.

- **Xạ trị:** Phân vùng chính xác các vùng NCR, ED, ET cho phép thiết kế liều xạ khác nhau cho từng vùng (dose painting), tối ưu hóa hiệu quả điều trị và giảm tác dụng phụ.

- **Theo dõi đáp ứng điều trị:** So sánh thể tích và hình thái khối u qua các lần quét MRI giúp đánh giá hiệu quả điều trị khách quan.

4. Dân chủ hóa công nghệ AI trong y tế:

Mô hình 2D với yêu cầu tài nguyên thấp mở ra khả năng:

Triển khai tại các bệnh viện tuyến huyện, tuyến tỉnh không có GPU cao cấp.

Chạy trên laptop/workstation thông thường, không cần hạ tầng cloud đắt đỏ.

Giảm chi phí đầu tư và vận hành, phù hợp với ngân sách y tế hạn chế.

Đặc biệt có ý nghĩa tại các quốc gia đang phát triển nơi nguồn lực y tế bị giới hạn.

5. Tích hợp vào quy trình làm việc lâm sàng:

- Giao diện đơn giản, dễ sử dụng cho nhân viên y tế không có nền tảng kỹ thuật
- Thời gian suy luận nhanh phù hợp với yêu cầu khẩn cấp trong lâm sàng
- Khả năng tích hợp với hệ thống PACS và RIS (Radiology Information System)
- Xuất kết quả dưới dạng chuẩn DICOM-RT hoặc NIfTI để bác sĩ tiếp tục chỉnh sửa nếu cần

6. Hỗ trợ đào tạo và giáo dục y khoa:

- Công cụ trực quan hóa để giảng dạy sinh viên y khoa về đặc điểm hình ảnh học của khối u não

- Cung cấp ground truth annotation tự động để tạo dữ liệu huấn luyện cho bác sĩ mới

- Hỗ trợ nghiên cứu hồi cứu (retrospective studies) với khối lượng lớn dữ liệu lịch sử

6. Tiềm năng mở rộng và cải tiến:

- Nền tảng để phát triển các mô hình phức tạp hơn trong tương lai.

- Khả năng kết hợp với các biomarkers khác (genomic, proteomic) để dự đoán tiên lượng.

- Tích hợp vào hệ thống hỗ trợ quyết định lâm sàng (Clinical Decision Support Systems).

8. Khả năng ứng dụng rộng hơn:

Mặc dù tập trung vào glioma, kiến trúc và phương pháp có thể dễ dàng điều chỉnh cho:

- Các loại u não khác (meningioma, schwannoma, pituitary adenoma)
- Các bài toán phân vùng khác trong y khoa (phổi, gan, tim, prostate)
- Transfer learning để giảm nhu cầu dữ liệu huấn luyện cho các tác vụ mới.

1.5 Cấu trúc báo cáo

Chương 1. Tổng quan

Chương 2. Cơ sở lý thuyết

Chương 3. Phân tích yêu cầu và thiết kế mô hình

Chương 4. Thực nghiệm và đánh giá

Chương 5. Kết luận và hướng phát triển

Tài liệu tham khảo.

CHƯƠNG 2. CƠ SỞ LÝ THUYẾT

2.1 Kiến thức về ảnh MRI

2.1.1 Nguyên lý tạo ảnh cộng hưởng từ

Nguyên lý vật lý cơ bản:

Cơ thể con người chứa khoảng 60-70% nước, trong đó nguyên tử hydro (H) là thành phần chủ yếu. Hạt nhân hydro có một proton mang điện tích dương và moment từ (spin), khiến chúng hoạt động như những nam châm nhỏ. Trong điều kiện bình thường, các proton này định hướng ngẫu nhiên và không tạo ra từ trường tổng hợp [18].

Khi đặt cơ thể vào từ trường mạnh B_0 (thường từ 1.5T đến 3T, với 1 Tesla = 10,000 Gauss), các proton hydro sẽ sắp xếp theo hai hướng: song song (spin-up) và ngược chiều (spin-down) với từ trường ngoài. Tuy nhiên, do trạng thái song song có mức năng lượng thấp hơn, số proton ở trạng thái này nhiều hơn một chút, tạo ra một vector từ hóa tổng hợp M_0 cùng phương với B_0 [19].

Các proton không chỉ sắp xếp theo B_0 mà còn quay tròn (precession) xung quanh trục từ trường với tần số Larmor ω_0 , được xác định bởi phương trình:

$$\omega_0 = \gamma \times B_0$$

Trong đó γ là hằng số từ hồi chuyển (gyromagnetic ratio), đối với proton hydro $\gamma \approx 42.58 \text{ MHz/T}$.

Quá trình kích thích và thu tín hiệu:

Để thu được tín hiệu MRI, hệ thống phát một xung sóng vô tuyến (RF pulse) với tần số đúng bằng tần số Larmor. Xung RF này truyền năng lượng cho hệ thống spin, làm vector từ hóa M_0 lệch khỏi trục dọc và bắt đầu quay trong mặt phẳng ngang. Góc lệch phụ thuộc vào năng lượng xung RF, thường được thiết lập ở 90° (flip angle) [18, 19].

Sau khi tắt xung RF, hệ thống spin bắt đầu quá trình hồi phục (relaxation) về trạng thái cân bằng ban đầu. Quá trình này diễn ra theo hai cơ chế độc lập:

- 1. Relaxation dọc (T1 relaxation):** Vector từ hóa dọc phục hồi dần về giá trị M_0 ban đầu theo hàm mũ với hằng số thời gian T1. T1 là thời gian cần thiết để từ hóa dọc phục hồi 63% giá trị ban đầu. Các mô khác nhau có T1 khác nhau: chất xám não ~900ms, chất trắng não ~600ms, dịch não tủy ~4000ms ở từ trường 1.5T.
- 2. Relaxation ngang (T2 relaxation):** Do tương tác spin-spin, các proton mất đồng pha dần, làm vector từ hóa ngang giảm theo hàm mũ với hằng số thời gian T2. T2 luôn ngắn hơn T1 và phản ánh sự mất đồng nhất của mô. Chất xám não có T2 ~100ms, chất trắng ~80ms, dịch não tủy ~2000ms.

Trong khi vector từ hóa ngang giảm dần, nó tạo ra một từ trường biến thiên được các cuộn dây thu (receiver coil) bao quanh bệnh nhân ghi nhận dưới dạng tín hiệu điện từ FID (Free Induction Decay). Tín hiệu này được số hóa và xử lý để tạo thành ảnh MRI.

Mã hóa không gian và tạo ảnh:

Để xác định vị trí không gian của tín hiệu, hệ thống MRI sử dụng ba gradient từ trường theo ba trục x, y, z. Các gradient này làm thay đổi tần số Larmor theo vị trí, cho phép phân biệt tín hiệu từ các vùng khác nhau của cơ thể:

Gradient chọn lớp cắt (slice selection): Xác định mặt phẳng cắt ngang, dọc hay nghiêng

Gradient mã hóa pha (phase encoding): Tạo sự khác biệt về pha của tín hiệu theo một chiều

Gradient mã hóa tần số (frequency encoding): Tạo sự khác biệt về tần số theo chiều còn lại

Dữ liệu thu được được lưu trong không gian tần số gọi là k-space. Ảnh MRI cuối cùng được tạo ra bằng cách áp dụng phép biến đổi Fourier 2D lên dữ liệu k-space, chuyển đổi từ miền tần số sang miền không gian.

2.1.2 Các chuỗi ảnh MRI phổ biến: T1, T1ce, T2, FLAIR

Bằng cách thay đổi các tham số chuỗi xung như thời gian lặp lại (TR - Time of Repetition), thời gian echo (TE - Time to Echo) và góc lệch (flip angle), có thể tạo ra các chuỗi ảnh MRI khác nhau, mỗi chuỗi nhấn mạnh các đặc tính mô khác nhau. Trong chẩn đoán u não, bốn chuỗi ảnh được sử dụng phổ biến nhất là T1, T1ce, T2 và FLAIR.

Ảnh T1-weighted (T1):

Chuỗi T1 sử dụng TR ngắn (300-700ms) và TE ngắn (10-30ms) để tối đa hóa độ tương phản dựa trên sự khác biệt T1 giữa các mô. Trên ảnh T1, các mô có T1 ngắn (mỡ, chất trắng) xuất hiện sáng (tín hiệu cao), trong khi mô có T1 dài (nước, dịch não tủy) xuất hiện tối (tín hiệu thấp).

Đặc điểm trên ảnh T1 não bộ:

- Chất trắng: sáng hơn chất xám do hàm lượng lipid cao trong myelin

Chất xám: xám trung bình

- Dịch não tủy (CSF): tối, gần đen

- Mỡ: rất sáng

Ảnh T1 cung cấp độ phân giải giải phẫu tốt, cho phép đánh giá cấu trúc não, kích thước não thất, và phát hiện các thay đổi giải phẫu. Tuy nhiên, khối u thường đồng tín hiệu hoặc hạ tín hiệu nhẹ so với chất xám, khó phân biệt nếu không có chất tương phản.

Ảnh T1 với chất tương phản (T1ce - T1 Contrast Enhanced):

T1ce là chuỗi T1 được chụp sau khi tiêm chất tương phản từ tính gadolinium qua đường tĩnh mạch. Gadolinium là kim loại đất hiếm có tính thuận từ mạnh, rút ngắn thời gian T1 của các mô mà nó tích tụ, làm tăng cường tín hiệu trên ảnh T1.

Trong não bộ bình thường, hàng rào máu-não (blood-brain barrier - BBB) ngăn cản gadolinium đi vào mô não. Tuy nhiên, khối u não thường phá vỡ BBB, cho phép chất tương phản rò rỉ vào tổ chức u và tăng cường tín hiệu. Mức độ tăng cường phản ánh:

- Sự phá vỡ hàng rào máu-não
- Mức độ tân mạch (angiogenesis)
- Độ ác tính của khối u

U thần kinh đệm độ cao (high-grade glioma) thường có vùng tăng cường mạnh, không đồng nhất hình vành khăn (ring enhancement) bao quanh vùng hoại tử trung tâm. Ngược lại, u độ thấp (low-grade) thường không tăng cường hoặc tăng cường nhẹ, đồng nhất.

T1ce là chuỗi ảnh quan trọng nhất để xác định ranh giới vùng lõi khối u tăng cường (enhancing tumor - ET) và đánh giá đáp ứng điều trị.

Ảnh T2-weighted (T2):

Chuỗi T2 sử dụng TR dài (2000-6000ms) và TE dài (80-120ms) để tối đa hóa độ tương phản dựa trên sự khác biệt T2. Trên ảnh T2, các mô có T2 dài (nước tự do, phù nề) xuất hiện sáng, trong khi mô có T2 ngắn xuất hiện tối.

Đặc điểm trên ảnh T2 não bộ:

- Dịch não tuy: rất sáng
- Chất xám: sáng hơn chất trắng (ngược với T1)
- Chất trắng: tối hơn
- Phù não: tăng tín hiệu rõ rệt

Ảnh T2 rất nhạy trong phát hiện:

- Phù não quanh khối u (peritumoral edema)
- Viêm, nhiễm trùng
- Tốn thương mạch máu (đột quỵ, chảy máu)
- Thoái hóa chất trắng

Trong u não, vùng phù nề và thâm nhiễm u (ED - edema) hiển thị rõ ràng trên T2 dưới dạng tín hiệu tăng cao lan tỏa xung quanh khối u. Tuy nhiên, T2 không phân biệt rõ ranh giới giữa u và phù nề, và không cho thấy vùng tăng cường.

Ảnh FLAIR (Fluid-Attenuated Inversion Recovery):

FLAIR là một biến thể đặc biệt của T2, sử dụng xung đảo ngược (inversion pulse) với thời gian đảo ngược TI được chọn để triệt tiêu tín hiệu từ dịch não tủy ($TI \approx 2000-2500\text{ms}$ ở 1.5T). FLAIR có TR rất dài (6000-10000ms) và TE dài (100-150ms).

Ưu điểm của FLAIR:

- Dịch não tủy xuất hiện đen hoàn toàn, không che lấp tổn thương
- Tăng độ tương phản cho tổn thương quanh não thất
- Phát hiện tốt các tổn thương vỏ não (cortical lesions)
- Nhạy với phù não, viêm màng não

Trên ảnh FLAIR, khối u và vùng phù nề xung quanh hiển thị sáng rõ ràng trên nền chất trắng xám và CSF đen. FLAIR đặc biệt hữu ích trong:

- Xác định toàn bộ vùng khối u (whole tumor - WT), bao gồm cả vùng thâm nhiễm không tăng cường.
- Phân biệt tổn thương não thực sự với hiện tượng thể tích từng phần (partial volume effect) từ CSF.
- Đánh giá mức độ lan tỏa của u thần kinh đệm.

Chuỗi ảnh	TR	TE	TI	CSF	Chất xám	Chất trắng	Ứng dụng chính
T1	Ngắn	Ngắn	-	Tối	Xám	Sáng	Giải phẫu, cấu trúc
T1ce	Ngắn	Ngắn	-	Tối	Xám	Sáng	Vùng tăng cường, lõi u
T2	Dài	Dài	-	Rất sáng	Sáng	Tối	Phù nề, tổn thương
FLAIR	Rất dài	Dài	~2000ms	Đen	Sáng	Xám	Toàn bộ khối u, thâm nhiễm

Bảng tổng hợp đặc điểm các chuỗi ảnh

2.1.3 Đặc điểm hình thái u não trên các chuỗi ảnh

Các thành phần chính của u não trên ảnh MRI:

Một khối u thần kinh đệm điển hình bao gồm các vùng có đặc điểm hình ảnh khác nhau, phản ánh các quá trình sinh học phức tạp:

1. Vùng hoại tử và nang (Necrotic/Cystic Core - NCR):

Vùng hoại tử là dấu hiệu đặc trưng của glioma độ cao, đặc biệt là GBM. Do khối u phát triển nhanh, mạch máu không đủ cung cấp oxy và dinh dưỡng, dẫn đến chết tế bào ở trung tâm khối u. Vùng này cũng có thể chứa các nang dịch.

Đặc điểm trên các chuỗi ảnh:

- **T1:** Hạ tín hiệu rõ rệt (tối), tương tự CSF
- **T1ce:** Không tăng cường hoặc tăng cường rìa, hình thành vành khăn (rim enhancement) đặc trưng
- **T2:** Tăng tín hiệu cao (rất sáng), khó phân biệt với CSF
- **FLAIR:** Tín hiệu biến thiên, có thể sáng hoặc tối tùy thuộc hàm lượng protein

Sự hiện diện của hoại tử là tiêu chí quan trọng để phân biệt glioma grade IV với grade III. Vành tăng cường bao quanh hoại tử phản ánh các tế bào u còn sống, đang sinh sản mạnh và có tính xâm lấn cao.

2. Vùng lõi khối u tăng cường (Enhancing Tumor - ET):

Đây là vùng mô u có tân sinh mạch mạnh mẽ và BBB bị phá vỡ hoàn toàn, cho phép chất tương phản thẩm vào. Vùng ET thường tương ứng với phần u có tế bào ác tính nhất, mật độ tế bào cao và hoạt động phân bào mạnh.

Đặc điểm hình thái:

- **T1:** Đồng tín hiệu hoặc hạ tín hiệu so với chất xám
- **T1ce:** Tăng cường mạnh, không đồng nhất, có thể dạng khối đặc, vành khăn, hay nốt
- **T2:** Tín hiệu thay đổi, thường hạ tín hiệu so với phù nề xung quanh
- **FLAIR:** Tăng tín hiệu, nhưng kém nổi bật hơn vùng phù nề

Hình dạng và mô hình tăng cường có giá trị chẩn đoán:

- Tăng cường đồng nhất: Thường thấy ở u màng não, lymphoma
- Tăng cường vành khăn: Đặc trưng GBM, apxe não, áp xe
- Tăng cường không đồng nhất: Phổ biến trong glioma grade III-IV

3. Vùng phù não và thâm nhiễm u (Edema - ED):

Vùng phù não bao quanh khối u phản ánh sự thất bại của BBB, tăng thẩm mạch máu và tích tụ dịch trong khoảng gian bào. Tuy nhiên, trong glioma, vùng này không chỉ là phù nề đơn thuần mà còn chứa các tế bào u thâm nhiễm lan tỏa vào chất trắng não bình thường theo các bó sợi trực.

Đặc điểm hình thái:

- **T1:** Hạ tín hiệu nhẹ đến trung bình

- **T1ce:** Không tăng cường (BBB chỉ bị tổn thương, chưa phá vỡ hoàn toàn)
- **T2:** Tăng tín hiệu rõ rệt (sáng), ranh giới không rõ
- **FLAIR:** Tăng tín hiệu rất rõ, cho phép đánh giá toàn bộ phạm vi bị ảnh hưởng

Vùng ED thường lan tỏa dọc theo chất trắng não, tôn trọng ranh giới giải phẫu như giữa chất xám-chất trắng, nhưng có thể vượt qua thẻ chai (corpus callosum) sang bán cầu đối diện, tạo thành hình "cánh bướm" (butterfly pattern) đặc trưng của glioblastoma.

Mức độ phù nề không tương quan trực tiếp với độ ác tính. U màng não lành tính có thể gây phù nề rộng, trong khi một số GBM lại có phù nề hạn chế.

4. Các đặc điểm phụ trợ khác:

- Xuất huyết: Thường thấy trong GBM, hiển thị dưới dạng tín hiệu hỗn hợp phức tạp tùy giai đoạn phân hủy hemoglobin
- Hiệu ứng khói (Mass effect): Chèn ép não thất, đường giữa, rãnh não, phản ánh kích thước và vị trí khối u
- Xâm lấn thẻ chai: Dấu hiệu xấu, cho thấy u lan sang bán cầu đối diện
- Thoái hóa nang: Thường thấy trong astrocytoma pilocytic (grade I)

Grade	Loại u	Hoại tử	Tăng cường	Phù nề	Ranh giới
I	Pilocytic astrocytoma	Không	Có thẻ có	Nhẹ	Rõ ràng
II	Diffuse astrocytoma	Không	Không/Nhẹ	Nhẹ	Không rõ
III	Anaplastic astrocytoma	Không	Có	Trung bình	Không rõ
IV	Glioblastoma	Có	Mạnh (vành khăn)	Rộng	Không rõ

Bảng phân loại theo WHO và đặc điểm hình ảnh

Thách thức trong đánh giá hình ảnh:

Việc phân định chính xác ranh giới các vùng khối u trên MRI gấp nhiều khăn:

- Ranh giới giữa ET và ED thường không rõ ràng, đặc biệt trên T2/FLAIR
- Sự thâm nhiễm vi thể vượt xa ranh giới có thể nhìn thấy trên ảnh

- Các hiện tượng nhiễu như hiệu ứng thể tích từng phần, artifacts chuyển động
- Sự chồng chéo tín hiệu giữa các loại mô khác nhau

Chính những thách thức này tạo động lực mạnh mẽ cho việc phát triển các phương pháp phân vùng tự động sử dụng học máy và học sâu, nhằm hỗ trợ bác sĩ đưa ra chẩn đoán chính xác và kế hoạch điều trị tối ưu.

2.2 Bài toán phân vùng ảnh y sinh

2.2.1 Định nghĩa và mục tiêu

Phân vùng ảnh y tế (medical image segmentation) là quá trình phân chia một ảnh y khoa thành các vùng hoặc đối tượng có ý nghĩa về mặt giải phẫu và sinh lý, sao cho mỗi pixel (ảnh 2D) hoặc voxel (ảnh 3D) được gán nhãn thuộc về một cấu trúc cụ thể. Đây là một bước quan trọng trong quy trình phân tích ảnh y tế, đóng vai trò cầu nối giữa việc thu thập dữ liệu hình ảnh và các ứng dụng lâm sàng.

Định nghĩa toán học:

Cho một ảnh y tế $I: \Omega \rightarrow \mathbb{R}^n$, trong đó $\Omega \subset \mathbb{R}^d$ là miền xác định ($d=2$ cho ảnh 2D, $d=3$ cho ảnh 3D), và n là số kênh ($n=1$ cho ảnh grayscale, $n>1$ cho ảnh đa kênh như MRI). Bài toán phân vùng là tìm một hàm ảnh xạ:

$$S: \Omega \rightarrow \{0, 1, 2, \dots, C-1\}$$

sao cho mỗi điểm $x \in \Omega$ được gán một nhãn lớp $S(x) \in \{0, 1, \dots, C-1\}$, với C là số lượng lớp cần phân vùng (bao gồm cả nền background).

Trong bối cảnh phân vùng u não BraTS, chúng ta có:

- **I:** Bộ 4 chuỗi ảnh MRI (T1, T1ce, T2, FLAIR), mỗi chuỗi có kích thước $240 \times 240 \times 155$ voxels
- **C = 4:** Bốn lớp cần phân vùng
 - Lớp 0: Nền (Background)
 - Lớp 1: Vùng hoại tử và nang (NCR - Necrotic/Cystic core)
 - Lớp 2: Vùng phù não (ED - Edema)
 - Lớp 3: Vùng tăng cường (ET - Enhancing Tumor)

Mục tiêu chính:

1. **Hỗ trợ chẩn đoán:** Phân vùng chính xác giúp bác sĩ xác định vị trí, kích thước và hình dạng khối u, từ đó đưa ra chẩn đoán phân biệt giữa các loại u não và đánh giá mức độ ác tính.
2. **Lập kế hoạch điều trị:** Kết quả phân vùng cung cấp thông tin không gian 3D chính xác về khối u, hỗ trợ:

- Phẫu thuật: Xác định ranh giới cắt bỏ an toàn
- Xạ trị: Vẽ vùng chiếu xạ (target volume delineation)
- Hóa trị: Đánh giá khả năng tiếp cận của thuốc

3. Theo dõi tiến triển bệnh: So sánh các lần chụp MRI theo thời gian để đánh giá:

- Tốc độ tăng trưởng khối u
- Đáp ứng với điều trị (Response Assessment in Neuro-Oncology - RANO criteria)
- Phát hiện tái phát sớm

4. Nghiên cứu lâm sàng: Phân vùng tự động cho phép:

- Phân tích định lượng trên tập dữ liệu lớn
- Khám phá mối tương quan giữa hình thái học và phân tử sinh học
- Phát triển các biomarker hình ảnh (imaging biomarkers)

5. Tiết kiệm thời gian và công sức: Phân vùng thủ công bởi chuyên gia thần kinh X quang có thể mất nhiều giờ cho một ca bệnh. Hệ thống tự động giảm xuống còn vài phút, cho phép bác sĩ tập trung vào công việc phân tích và ra quyết định lâm sàng.

Tiêu chuẩn chất lượng:

Một phương pháp phân vùng tốt cần đạt các yêu cầu:

- **Độ chính xác cao:** Khớp tốt với ground truth do chuyên gia gán nhãn
- **Tính nhất quán:** Cho kết quả ổn định với các ca bệnh tương tự
- **Khả năng tổng quát hóa:** Hoạt động tốt trên dữ liệu từ nhiều trung tâm y tế, máy chụp khác nhau
- **Hiệu quả tính toán:** Thời gian xử lý chấp nhận được trong thực tiễn lâm sàng
- **Tính giải thích được:** Bác sĩ có thể hiểu và tin tưởng kết quả

2.2.2 Các thách thức trong phân vùng ảnh y khoa

Phân vùng ảnh y tế, đặc biệt là phân vùng u não, là một bài toán cực kỳ khó do nhiều yếu tố phức tạp vốn có của dữ liệu y khoa và bản chất sinh học của bệnh lý:

1. Tính không đồng nhất về hình thái và cấu trúc:

U não có tính đa dạng cao về kích thước, hình dạng, vị trí và đặc điểm hình ảnh:

Kích thước: Từ vài millimét đến hàng chục centimét

Hình dạng: Không theo quy luật cố định, có thể tròn, bầu dục, hay bất đối xứng hoàn toàn

Vị trí: Xuất hiện ở bất kỳ vùng nào của não, ảnh hưởng đến các cấu trúc giải phẫu xung quanh khác nhau

Tính không đồng nhất nội bộ: Cùng một khối u có thể chứa vùng đặc, nang, hoại tử, xuất huyết, vôi hóa với tín hiệu hỗn hợp

2. Ranh giới mờ và không rõ ràng:

Một trong những thách thức lớn nhất là xác định ranh giới chính xác giữa các vùng:

Ranh giới khối u - não bình thường: Glioma có đặc tính thâm nhiễm lan tỏa, không có bao bọc rõ ràng. Tế bào u có thể xâm nhập vào mô não xung quanh hàng centimét xa khỏi khối u chính nhìn thấy trên ảnh.

Ranh giới ET - ED: Vùng tăng cường và phù nề có thể chồng lấn, đặc biệt trên ảnh T2/FLAIR.

Hiệu ứng thể tích từng phần (Partial Volume Effect): Ở ranh giới giữa các mô, một voxel có thể chứa nhiều loại mô khác nhau, dẫn đến tín hiệu trung gian khó phân loại.

3. Sự biến đổi về cường độ tín hiệu:

Khác với ảnh tự nhiên, ảnh MRI không có thang đo cường độ tuyệt đối:

- **Không chuẩn hóa toàn cầu:** Cùng một loại mô trên các lần chụp khác nhau, máy khác nhau có thể cho giá trị cường độ hoàn toàn khác
- **Bias field (inhomogeneity):** Không đồng nhất từ trường tạo ra gradient cường độ trên toàn ảnh
- **Nhiễu:** Nhiễu Rician, nhiễu nhiệt, artifacts chuyển động, ghost artifacts
- **Sự chồng chéo histogram:** Phân bố cường độ của các lớp khác nhau có thể chồng lấn đáng kể

4. Mất cân bằng lớp nghiêm trọng:

Trong ảnh não, tỷ lệ voxel giữa các lớp rất không cân bằng:

- Nền (background + não bình thường): ~95-98%
- Vùng phù nề (ED): ~1-3%
- Vùng tăng cường (ET): ~0.1-1%
- Vùng hoại tử (NCR): ~0.1-0.5%

Sự mất cân bằng này khiến mô hình học máy có xu hướng:

- Thiên vị về lớp đa số (nền)
- BỎ qua hoặc phân vùng kém các lớp thiểu số nhưng quan trọng về mặt lâm sàng
- Khó tối ưu hóa hàm loss function

5. Thiếu dữ liệu có gán nhãn:

Dữ liệu y khoa gán nhãn có những hạn chế đặc thù:

- **Chi phí cao:** Việc gán nhãn đòi hỏi chuyên gia có trình độ cao (bác sĩ X quang thần kinh), mất nhiều thời gian.
- **Tính chủ quan:** Ngay cả giữa các chuyên gia cũng có sự khác biệt trong đánh giá ranh giới (inter-rater variability)
- **Quyền riêng tư:** Dữ liệu y tế nhạy cảm, khó chia sẻ và tập hợp thành bộ dữ liệu lớn
- **Bệnh hiếm:** Một số loại u não rất hiếm, khó thu thập đủ dữ liệu đại diện

6. Tính đa dạng về giao thức chụp:

Dữ liệu MRI phụ thuộc mạnh vào điều kiện thu thập:

- **Máy chụp:** Từ trường khác nhau (1.5T vs 3T), hãng sản xuất khác nhau (Siemens, GE, Philips)
- **Giao thức chụp:** Thông số TR, TE, flip angle, độ dày lát cắt
- **Tiền xử lý:** Hiệu chỉnh bias field, chuẩn hóa cường độ, đăng ký ảnh (registration)

Sự khác biệt này tạo ra domain shift, khiến mô hình huấn luyện trên dữ liệu từ một nguồn có thể hoạt động kém trên dữ liệu từ nguồn khác.

7. Tính đa chiều và khối lượng dữ liệu lớn:

Ảnh MRI não bộ thường có:

- Kích thước: $240 \times 240 \times 155$ voxels (hoặc lớn hơn)
- 4 chuỗi ảnh (multi-modal)
- Tổng cộng ~35 triệu voxels cần phân loại cho mỗi ca

Điều này đặt ra thách thức về:

- Bộ nhớ GPU hạn chế
- Thời gian huấn luyện dài
- Khó khăn trong việc học context toàn cục (global context)

8. Đánh giá và xác thực mô hình

- Không có ground truth tuyệt đối: Gán nhãn thủ công cũng có sai số, và ranh giới thực sự của khối u chỉ xác định được qua sinh thiết hoặc mô bệnh học
- Metric phức tạp: Cần nhiều chỉ số đánh giá khác nhau (Dice, IoU, Sensitivity, Specificity) để có cái nhìn toàn diện
- Clinical relevance: Mô hình có thể có Dice score cao nhưng sai ở những vùng quan trọng về mặt lâm sàng

9. Yêu cầu tính giải thích được (Interpretability):

Trong ứng dụng y tế:

- Bác sĩ cần hiểu tại sao mô hình đưa ra dự đoán nhất định
- Các mô hình "hộp đen" khó được chấp nhận trong thực hành lâm sàng
- Cần cơ chế truy xuất và xác minh quyết định của mô hình

10. Thách thức về tính toán thời gian thực:

Trong môi trường lâm sàng:

- Yêu cầu xử lý nhanh
- Tích hợp với PACS
- Hoạt động ổn định trên phần cứng có sẵn tại bệnh viện

2.2.3 Các phương pháp truyền thống

Trước khi học sâu trở thành phương pháp chủ đạo, nhiều kỹ thuật phân vùng ảnh truyền thống đã được phát triển và ứng dụng trong y tế. Mặc dù hiện nay các phương pháp này ít được sử dụng độc lập, nhưng hiểu rõ chúng giúp đánh giá được bước tiến của học sâu và một số kỹ thuật vẫn có thể kết hợp làm tiền xử lý hoặc hậu xử lý.

1. Phương pháp dựa trên ngưỡng (Threshold-based methods):

Đây là phương pháp đơn giản nhất, dựa trên giả định rằng các vùng khác nhau có phân bố cường độ pixel/voxel khác biệt rõ rệt.

Nguưỡng toàn cục (Global Thresholding):

- Chọn một giá trị ngưỡng T và phân loại: voxel $> T$ thuộc lớp 1, voxel $\leq T$ thuộc lớp 0
- Phương pháp Otsu tự động tìm ngưỡng tối ưu bằng cách tối đa hóa phương sai giữa các lớp.

Nguưỡng thích nghi (Adaptive Thresholding):

- Nguưỡng thay đổi theo từng vùng cục bộ của ảnh
- Phù hợp hơn với ảnh có độ sáng không đồng nhất

Ưu điểm:

- Đơn giản, nhanh, dễ triển khai
- Hiệu quả với ảnh có độ tương phản cao

Nhược điểm:

- Rất nhạy với nhiễu và artifacts
- Không xử lý được sự chồng lấn histogram
- Bỏ qua thông tin không gian và ngữ cảnh xung quanh
- Không phù hợp với ảnh MRI não do tính phức tạp của cấu trúc

2. Phương pháp dựa trên vùng (Region-based methods):

Các phương pháp này nhóm các pixel/voxel có đặc tính tương tự thành một vùng.

Region Growing:

- Bắt đầu từ một tập seed points (điểm khởi tạo)
- Lần lượt thêm các voxel lân cận có cường độ tương tự vào vùng
- Dừng lại khi không còn voxel nào thỏa mãn tiêu chí

Tiêu chí kết nối có thể dựa trên:

- Sự khác biệt cường độ: $|I(x) - \mu_{region}| < T$
- Gradient: $\nabla I(x) < T$
- Kết cấu (texture) cục bộ

Region Split and Merge:

- Phân chia đệ quy ảnh thành các vùng con (quadtree/octree)
- Gộp các vùng con có đặc tính đồng nhất

Ưu điểm:

- Tận dụng thông tin kết nối không gian
- Tạo ra các vùng liên thông

Nhược điểm:

- Phụ thuộc nhiều vào vị trí seed points
- Nhạy với nhiễu - một voxel nhiễu có thể làm "rò rỉ" vùng
- Khó xác định tiêu chí đồng nhất tối ưu
- Không xử lý tốt các vùng có ranh giới mờ

3. Phương pháp dựa trên cạnh (Edge-based methods):

Phát hiện ranh giới giữa các vùng dựa trên sự thay đổi đột ngột về cường độ.

Toán tử gradient:

- Sobel, Prewitt: Tính gradient theo hai hướng x, y
- Canny: Thuật toán phát hiện cạnh tinh vi với non-maximum suppression và hysteresis thresholding
- Laplacian of Gaussian (LoG): Làm mượt trước, rồi tìm zero-crossing của đạo hàm bậc hai

Gradient được tính: $\nabla I = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right)$

Độ lớn: $|\nabla I| = \sqrt{\left(\frac{\partial I}{\partial x} \right)^2 + \left(\frac{\partial I}{\partial y} \right)^2}$

Active Contours (Snakes):

- Khởi tạo một đường cong ban đầu gần đối tượng
- Tối ưu hóa năng lượng để đường cong co giãn về ranh giới thực:

$$E = E_{internal} + E_{image} + E_{constraint}$$

- $E_{internal}$: Độ mượt, cứng của đường cong

- E_{image} : Lực kéo về vùng gradient cao (cạnh)
- $E_{constraint}$: Ràng buộc bên ngoài

Level Set Methods:

- Biểu diễn đường cong phân vùng như zero level set của hàm $\varphi(x,y,t)$
- Đường cong tiến hóa theo phương trình: $\frac{\partial \varphi}{\partial t} + F|\nabla \varphi| = 0$ trong đó F là hàm tốc độ phụ thuộc vào đặc tính ảnh

Ưu điểm:

- Cho ranh giới trơn, liên tục
- Có thể tích hợp kiến thức a priori về hình dạng
- Level set xử lý tốt thay đổi topology (tách/gộp vùng)

Nhược điểm:

- Rất nhạy với nhiễu
- MRI có nhiều cạnh giả (false edges) do artifacts
- Cần khởi tạo gần đối tượng mục tiêu
- Tính toán phức tạp, chậm
- Khó áp dụng với đối tượng có ranh giới mờ như u não

4. Phân cụm (Clustering methods):

Nhóm các voxel thành các cụm (clusters) dựa trên đặc trưng.

K-Means Clustering:

- Phân chia dữ liệu thành K cụm
- Mỗi voxel thuộc về cụm có tâm gần nhất
- Thuật toán lặp:
 1. Gán mỗi voxel vào cụm gần nhất
 2. Cập nhật tâm cụm = mean của các voxel trong cụm
 3. Lặp đến khi hội tụ

Fuzzy C-Means (FCM):

- Mở rộng K-means với fuzzy membership
- Mỗi voxel có membership degree $u_{ik} \in [0,1]$ cho mỗi cụm k
- Tối ưu hóa:

$$J = \sum_i \sum_k (u_{ik})^m \|x_i - c_k\|^2$$

với m là fuzzy coefficient (thường m=2)

Expectation-Maximization (EM) với Gaussian Mixture Model:

- Giá định dữ liệu tuân theo hỗn hợp phân phối Gaussian
- E-step: Tính xác suất hậu nghiệm voxel thuộc mỗi cụm
- M-step: Cập nhật tham số (mean, covariance) của mỗi Gaussian

Ưu điểm:

- Không cần gán nhãn trước
- FCM xử lý tốt partial volume effect
- Có nền tảng thống kê vững chắc

Nhược điểm:

- Cần biết trước số cụm K
- Nhạy với khởi tạo
- Bỏ qua thông tin không gian
- Giá định phân phối Gaussian thường không đúng với ảnh y tế

5. Atlas-based methods:

Sử dụng atlas (bản đồ giải phẫu chuẩn) đã được gán nhãn.

Quy trình:

1. Đăng ký (registration) ảnh cần phân vùng với atlas
2. Áp dụng biến đổi tìm được để chuyển nhãn từ atlas sang ảnh mới
3. (Tùy chọn) Sử dụng multi-atlas và voting/fusion

Ưu điểm:

- Tận dụng kiến thức giải phẫu mạnh mẽ
- Hiệu quả với cấu trúc có hình dạng ổn định (như não thất, các nhân trong não)

Nhược điểm:

- Phụ thuộc hoàn toàn vào chất lượng đăng ký
- Không phù hợp với cấu trúc bệnh lý (như u não) - không có trong atlas
- Tính toán phức tạp, đặc biệt với multi-atlas
- Khó xử lý biến đổi giải phẫu lớn

6. Phương pháp kết hợp đặc trưng thủ công (Hand-crafted features) + Machine Learning:

Trước học sâu, phương pháp hai giai đoạn là state-of-the-art:

Giai đoạn 1: Trích xuất đặc trưng

- **Đặc trưng cường độ:** Mean, variance, histogram features trong vùng cục bộ
- **Đặc trưng kết cấu:**
 - GLCM: contrast, homogeneity, energy, correlation
 - LBP (Local Binary Pattern)
 - Gabor filters ở nhiều hướng và tần số
- **Đặc trưng đa chuỗi:** Tương quan giữa T1, T1ce, T2, FLAIR
- **Đặc trưng ngũ cảnh:** Vị trí tương đối với các cấu trúc giải phẫu

Giai đoạn 2: Phân loại

- Random Forest: Ensemble của nhiều decision trees
- Support Vector Machine (SVM): Tìm siêu phẳng phân tách tối ưu
- AdaBoost: Kết hợp nhiều weak classifiers

Ưu điểm:

- Tận dụng domain knowledge trong thiết kế features
- Random Forest robust với nhiều, xử lý tốt nhiều features
- SVM hiệu quả với dữ liệu chiều cao

Nhược điểm:

- Features thủ công không tối ưu, mất thông tin
- Tốn công sức thiết kế và tinh chỉnh features
- Khó nắm bắt các biểu diễn phức tạp, trừu tượng
- Hiệu suất kém hơn nhiều so với học sâu hiện đại

Kết luận:

Các phương pháp truyền thống có những hạn chế căn bản:

- Không học được biểu diễn tối ưu từ dữ liệu
- Thiếu khả năng tổng quát hóa
- Phụ thuộc nhiều vào tiền xử lý và tuning tham số

Những hạn chế này đã thúc đẩy sự chuyển mình sang học sâu, với khả năng học end-to-end các biểu diễn phân cấp từ dữ liệu thô.

2.3 Học sâu trong xử lý ảnh y tế

2.3.1 Mạng nơ-ron tích chập (CNN)

Mạng nơ-ron tích chập (Convolutional Neural Network - CNN) là một lớp kiến trúc học sâu được thiết kế đặc biệt cho dữ liệu có cấu trúc lưới (grid-like structure) như ảnh. CNN đã cách mạng hóa lĩnh vực thị giác máy tính kể từ thành công đột phá của AlexNet trong cuộc thi ImageNet 2012, và nhanh chóng trở thành công cụ chủ đạo trong xử lý ảnh y tế.

Động lực và nguyên lý thiết kế:

CNN được truyền cảm hứng từ cơ chế thị giác của động vật có vú, đặc biệt là khám phá về receptive fields trong vỏ não thị giác bởi Hubel và Wiesel (giải Nobel Y học 1981). Họ phát hiện ra rằng các tế bào thần kinh trong vỏ não thị giác phản ứng với các kích thích cục bộ trong thị trường và có tính phân cấp: các tế bào đơn giản phát hiện các cạnh, góc, trong khi các tế bào phức tạp hơn nhận diện các mẫu ở mức cao hơn.

CNN áp dụng ba nguyên lý chính:

1. **Local Connectivity (Kết nối cục bộ):** Thay vì kết nối fully-connected, mỗi neuron chỉ kết nối với một vùng nhỏ của input (receptive field), giảm mạnh số lượng tham số.
2. **Parameter Sharing (Chia sẻ tham số):** Cùng một bộ trọng số (filter/kernel) được áp dụng trên toàn bộ ảnh, dựa trên giả định rằng các đặc trưng hữu ích ở một vùng có thể hữu ích ở vùng khác.
3. **Equivariance to Translation (Bất biến tịnh tiến):** Nếu input dịch chuyển, output cũng dịch chuyển tương ứng, phù hợp với bản chất của ảnh.

Các thành phần cơ bản của CNN:

1. Lớp tích chập (Convolutional Layer):

Đây là thành phần cốt lõi của CNN. Một lớp tích chập bao gồm một tập các kernel/filters có thể học được, mỗi kernel trượt (slide) trên input để tạo ra feature map.

Cho input I có kích thước $H \times W \times C_{in}$ và kernel K kích thước $k \times k \times C_{in}$, phép tích chập 2D được định nghĩa:

$$(I * K)[i, j] = \sum_m \sum_n \sum_c I[i + m, j + n, c] \times K[m, n, c]$$

trong đó $m \in [0, k-1]$, $n \in [0, k-1]$, $c \in [0, C_{in}-1]$.

Một lớp tích chập với C_{out} filters sẽ tạo ra output có C_{out} channels (feature maps). Mỗi feature map học một đặc trưng khác nhau của input.

Các tham số quan trọng:

- **Kernel size:** Thường 3×3 hoặc 5×5 . Kernel nhỏ (3×3) cho phép xây dựng mạng sâu hơn
- **Stride:** Bước nhảy khi trượt kernel. Stride=1 giữ nguyên kích thước, stride=2 giảm một nửa
- **Padding:** Thêm viền xung quanh input
 - Valid padding: Không thêm, output nhỏ hơn input
 - Same padding: Thêm để output cùng kích thước với input
- **Dilation:** Giãn cách giữa các phần tử kernel để tăng receptive field mà không tăng số tham số

Số lượng tham số: Với C_{in} input channels, C_{out} output channels, kernel $k \times k$:
Parameters = $k \times k \times C_{in} \times C_{out} + C_{out}$ (bias)

2. Hàm kích hoạt (Activation Function):

Sau mỗi lớp tích chập, một hàm kích hoạt phi tuyến được áp dụng element-wise để mạng có thể học các quan hệ phức tạp.

ReLU (Rectified Linear Unit): $\text{ReLU}(x) = \max(0, x)$

Đây là hàm kích hoạt phổ biến nhất do:

- Đơn giản, tính toán nhanh
- Giảm vanishing gradient problem
- Tạo sparsity (nhiều activations = 0)

Leaky ReLU: $\text{LeakyReLU}(x) = x \text{ if } x > 0 \text{ else } \alpha x$ ($\alpha \approx 0.01$)

Khắc phục "dying ReLU" problem khi neurons bị "chết" (luôn output 0).

Các hàm khác:

- Sigmoid: $\sigma(x) = \frac{1}{1+e^{-x}}$ - Ít dùng trong hidden layers do vanishing gradient
- Tanh: $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ - Output trong [-1,1]
- ELU, SELU, Swish - Các biến thể hiện đại hơn

3. Lớp Batch Normalization:

Batch Normalization (BN) chuẩn hóa activations theo từng mini-batch để ổn định và tăng tốc quá trình training.

Cho một batch $\{x_1, \dots, x_m\}$, BN thực hiện:

1. Tính mean: $\mu = \frac{1}{m} \sum x_i$
2. Tính variance: $\sigma^2 = \frac{1}{m} \sum (x_i - \mu)^2$
3. Chuẩn hóa: $\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \epsilon}}$
4. Scale và shift: $y_i = \gamma \hat{x}_i + \beta$

với γ, β là tham số học được.

Lợi ích:

- Cho phép learning rate cao hơn
- Giảm sự phụ thuộc vào khởi tạo
- Có tác dụng regularization nhẹ
- Tăng tốc hội tụ đáng kể

Trong y tế, với batch size nhỏ (do hạn chế GPU memory), có thể sử dụng Group Normalization hoặc Instance Normalization thay thế.

4. Lớp Pooling:

Pooling giảm kích thước không gian của feature maps, giảm số lượng tham số và tính toán, đồng thời tăng receptive field và tính bất biến với các biến đổi nhỏ.

Max Pooling: Lấy giá trị lớn nhất trong mỗi vùng (thường 2×2): $Output[i,j] = \max(Input[2i:2i+2, 2j:2j+2])$

Average Pooling: Lấy giá trị trung bình trong mỗi vùng

Max pooling phổ biến hơn vì giữ lại đặc trưng nổi bật nhất. Thông thường sử dụng kernel 2×2 , stride 2 để giảm một nửa kích thước.

5. Lớp Dropout:

Dropout là kỹ thuật regularization mạnh mẽ để tránh overfitting. Trong quá trình training, mỗi neuron có xác suất p bị "tắt" (set về 0) ngẫu nhiên.

Điều này buộc mạng:

- Không phụ thuộc quá nhiều vào một số neurons cụ thể
- Học các biểu diễn redundant, robust hơn
- Tương đương với ensemble nhiều mạng con

Trong inference, tất cả neurons được giữ lại và outputs được scale bởi $(1-p)$.

Dropout rate điển hình: 0.2-0.5 cho lớp fully-connected, 0.1-0.2 cho lớp conv.

6. Lớp Fully Connected (Dense):

Ở cuối mạng, sau khi trích xuất features, các lớp fully-connected kết hợp các features để đưa ra dự đoán cuối cùng.

$$\text{Output} = \text{Activation}(W \times \text{Input} + b)$$

với W là ma trận trọng số, b là bias vector.

Kiến trúc tổng thể:

Một CNN điển hình có cấu trúc phân cấp:

Input → [Conv → BN → ReLU → Pool] × N → Flatten → [FC → Dropout] × M → Output

- **Shallow layers:** Học các features cấp thấp (edges, colors, textures)
- **Deeper layers:** Học các features phức tạp hơn, trừu tượng hơn (parts, objects)

Các kiến trúc CNN nổi tiếng:

1. LeNet-5 (1998): Mạng CNN đầu tiên thành công, nhận dạng chữ số viết tay
2. AlexNet (2012): Chiến thắng ImageNet, đánh dấu kỷ nguyên deep learning
3. VGGNet (2014): Sử dụng nhiều lớp conv 3×3 xếp chồng
4. GoogLeNet/Inception (2014): Inception modules với multi-scale processing
5. ResNet (2015): Skip connections giải quyết vanishing gradient, cho phép mạng rất sâu (>100 layers).
6. DenseNet (2017): Dense connections giữa tất cả các lớp

Ứng dụng trong y tế:

CNN đạt state-of-the-art trong nhiều tác vụ y tế:

- **Phân loại:** Phát hiện bệnh từ ảnh X-quang, CT, MRI
- **Phát hiện:** Xác định vị trí tổn thương, khối u
- **Phân vùng:** Tách từng pixel thuộc cơ quan/bệnh lý nào

Tuy nhiên, CNN có một số thách thức riêng trong y tế:

- Cần lượng lớn dữ liệu có gán nhãn
- Khó giải thích quyết định (black-box nature)
- Nhạy với domain shift (khác biệt giữa các máy chụp, giao thức)

2.3.2 Kiến trúc U-Net và các biến thể

U-Net là kiến trúc mạng nơ-ron tích chập được thiết kế đặc biệt cho bài toán phân vùng ảnh y tế, được giới thiệu bởi Ronneberger và cộng sự vào năm 2015 tại hội nghị MICCAI. U-Net nhanh chóng trở thành kiến trúc chuẩn (de facto standard) cho phân vùng ảnh sinh y học nhờ khả năng hoạt động tốt với lượng dữ liệu huấn luyện hạn chế và cấu trúc thanh lịch, hiệu quả. Thiết kế "hình chữ U" đối xứng của U-Net bao gồm:

Encoder (đường đi co lại): Trích xuất đặc trưng ngữ cảnh ở nhiều tỷ lệ thông qua các lớp tích chập và pooling, giảm dần độ phân giải không gian nhưng tăng chiều sâu kênh.

Bottleneck: Nằm ở "đáy" của chữ U, nắm bắt biểu diễn trừu tượng nhất của dữ liệu đầu vào.

Decoder (đường đi mở rộng): Tái tạo lại độ phân giải không gian ban đầu thông qua các lớp upsampling và tích chập, kết hợp với các đặc trưng từ encoder thông qua skip connections để bảo toàn thông tin chi tiết.

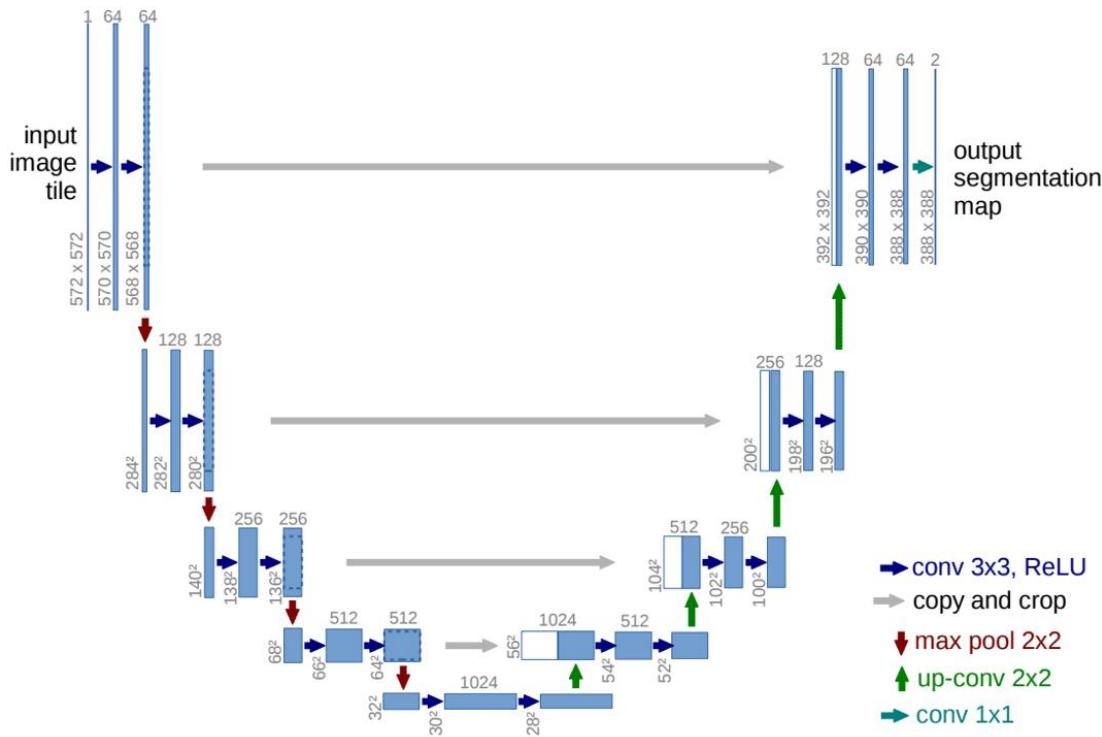
Skip connections là yếu tố then chốt giúp U-Net vượt trội trong phân vùng, cho phép thông tin chi tiết ở độ phân giải cao được truyền trực tiếp từ encoder sang decoder, khắc phục vấn đề mất thông tin trong quá trình downsampling.

Trong lĩnh vực phân vùng u não, các biến thể của U-Net như U-Net 3D, Attention U-Net, và các kiến trúc kết hợp với Transformer đã đạt được kết quả ấn tượng trên các bộ dữ liệu chuẩn như BraTS (Brain Tumor Segmentation Challenge). Tuy nhiên, phần lớn các nghiên cứu này tập trung vào mô hình 3D đòi hỏi tài nguyên tính toán lớn (GPU với VRAM cao, thời gian huấn luyện lâu), gây khó khăn trong việc triển khai thực tế tại các cơ sở y tế có nguồn lực hạn chế.

Trong bối cảnh đó, **phương pháp tiếp cận 2D** với U-Net truyền thống vẫn có ý nghĩa quan trọng. Mặc dù không tận dụng được toàn bộ thông tin không gian 3D, phương pháp 2D có những ưu điểm riêng:

- Yêu cầu bộ nhớ thấp hơn nhiều, có thể chạy trên GPU phổ thông
- Thời gian huấn luyện và suy luận nhanh hơn đáng kể
- Dễ dàng tích hợp vào quy trình làm việc lâm sàng hiện có
- Cho phép xử lý từng lát cắt độc lập, phù hợp với các trường hợp khẩn cấp

Nghiên cứu này hướng tới việc tối ưu hóa mô hình U-Net 2D thông qua việc thiết kế kiến trúc phù hợp, áp dụng các kỹ thuật augmentation dữ liệu tiên tiến, và sử dụng hàm mất mát kết hợp để đạt được hiệu suất cao với chi phí tính toán hợp lý.



Hình 2. Sơ đồ kiến trúc của mô hình U-Net

Động lực thiết kế:

Các bài toán phân vùng ảnh y tế đặt ra những yêu cầu đặc thù:

- Localization chính xác:** Cần xác định ranh giới của từng pixel/voxel
- Context toàn cục:** Cần thông tin ngữ cảnh rộng để phân biệt các cấu trúc
- Dữ liệu hạn chế:** Dữ liệu y tế gán nhãn khan hiếm và đắt đỏ

U-Net giải quyết những thách thức này thông qua kiến trúc encoder-decoder đối xứng với skip connections.

Kiến trúc U-Net:

U-Net có hình dạng chữ "U" đặc trưng, gồm ba thành phần chính:

1. Contracting Path (Encoder - Nhánh co):

Encoder thu nhỏ kích thước không gian và trích xuất features phân cấp:

- Mỗi block encoder gồm:**
 - 2 lớp Conv 3×3 + ReLU
 - 1 lớp Max Pooling 2×2 (stride=2) để downsample
- Cấu trúc phân cấp:**
 - Input: $H \times W \times C$ (ví dụ $240 \times 240 \times 4$ cho BraTS)

- Block 1: $240 \times 240 \times 64$
- Block 2: $120 \times 120 \times 128$ (sau pooling)
- Block 3: $60 \times 60 \times 256$
- Block 4: $30 \times 30 \times 512$
- Bottleneck: $15 \times 15 \times 1024$

Số channels tăng gấp đôi sau mỗi pooling ($64 \rightarrow 128 \rightarrow 256 \rightarrow 512 \rightarrow 1024$) để bù đắp cho việc mất thông tin không gian.

Mục đích: Học các đặc trưng ngữ nghĩa trừu tượng với receptive field lớn, nắm bắt context toàn cục.

2. Bottleneck (Cỗ chai):

Nằm ở điểm sâu nhất của mạng, nơi feature maps có kích thước nhỏ nhất nhưng số channels lớn nhất. Đây là nơi mạng học biểu diễn trừu tượng nhất về dữ liệu input.

3. Expansive Path (Decoder - Nhánh giãn):

Decoder phục hồi độ phân giải không gian và tạo segmentation map chi tiết:

- **Mỗi block decoder gồm:**
 - 1 lớp Up-convolution (Transposed Convolution) 2×2 để upsample
 - Concatenation với feature map tương ứng từ encoder (skip connection)
 - 2 lớp Conv 3×3 + ReLU
- **Cấu trúc ngược lại encoder:**
 - $30 \times 30 \times 512$ (sau upconv từ bottleneck)
 - Concat với encoder block 4 $\rightarrow 30 \times 30 \times 1024$
 - Conv $\rightarrow 30 \times 30 \times 512$
 - $60 \times 60 \times 256$ (sau upconv)
 - Concat với encoder block 3 $\rightarrow 60 \times 60 \times 512$
 - Conv $\rightarrow 60 \times 60 \times 256$
 - ... tiếp tục đến kích thước gốc
- **Output layer:**
 - Conv 1×1 để map từ 64 channels về số lượng classes (4 cho BraTS)
 - Output: $240 \times 240 \times 4$ (logits cho mỗi class)

4. Skip Connections (Kết nối tắt):

Đây là đổi mới quan trọng nhất của U-Net. Tại mỗi cấp độ, feature maps từ encoder được concatenate trực tiếp với feature maps của decoder tại cùng resolution.

Lợi ích của Skip Connections:

1. **Gradient Flow:** Tạo đường dẫn trực tiếp cho gradient lan truyền, giảm vanishing gradient

2. **Recover Spatial Information:** Decoder nhận được cả:

- High-level semantic features từ bottleneck (qua upsampling)
- Low-level spatial details từ encoder (qua skip connections)

Kết hợp này cho phép localization chính xác với context toàn cục.

3. **Multi-scale Features:** Mạng tận dụng features ở nhiều scales khác nhau, phù hợp với việc phân vùng các đối tượng có kích thước đa dạng.

Công thức tổng quát:

Đối với decoder level i: $D_i = \text{Dec_Conv}(\text{Concat}(\text{Upsample}(D_{i-1}), E_i))$

trong đó:

- D_i : Decoder feature ở level i
- E_i : Encoder feature ở level i (skip connection)
- Concat: Nối features theo chiều channel

Đặc điểm kỹ thuật quan trọng:

1. Data Augmentation mạnh mẽ:

Bài báo gốc nhấn mạnh tầm quan trọng của augmentation với dữ liệu y tế hạn chế:

- Elastic deformations: Biến dạng đàn hồi mô phỏng các biến thể sinh học
- Random rotations, flips, scaling
- Intensity shifts, noise injection

2. Weighted Loss Function:

Để xử lý mất cân bằng lớp và nhấn mạnh ranh giới, U-Net gốc sử dụng:

$$L = \sum_x w(x) \log(p(x))$$

trong đó $w(x)$ là weight map được tính để:

- Tăng trọng số cho pixels ở ranh giới giữa các đối tượng sát nhau

- Cân bằng tần suất xuất hiện của các lớp

3. Không có Fully-Connected layers:

Toàn bộ mạng là fully convolutional, cho phép:

- Xử lý input với kích thước tùy ý
- Ít tham số hơn
- Inference nhanh hơn

Số lượng tham số:

U-Net gốc với kiến trúc như mô tả trên có khoảng 30-40 triệu tham số, phụ thuộc vào số channels cụ thể. Tuy nhiên, con số này vẫn nhỏ hơn nhiều so với các mạng classification như VGG hoặc ResNet.

Các biến thể của U-Net:

Sau thành công của U-Net, nhiều biến thể đã được phát triển để cải thiện hiệu năng:

1. Residual U-Net:

Thay thế các double conv blocks bằng residual blocks:

$$x \rightarrow \text{Conv} \rightarrow \text{BN} \rightarrow \text{ReLU} \rightarrow \text{Conv} \rightarrow \text{BN} \rightarrow (+) \rightarrow \text{ReLU}$$

\downarrow \uparrow

Lợi ích:

- Training ổn định hơn với mạng sâu
- Gradient flow tốt hơn
- Thường đạt accuracy cao hơn

2. Attention U-Net:

Thêm attention gates vào skip connections để lọc features không liên quan:

$$AG(x_l, g) = \sigma(\psi^\top (\sigma_1(W_x x_l + W^g g + b)) + b_\psi)$$

Attention mechanism giúp:

- Mạng tập trung vào vùng quan trọng (tumor)
- Triệt tiêu features từ background
- Cải thiện đáng kể trên các đối tượng nhỏ

3. U-Net++:

Kiến trúc nested U-Net với nhiều sub-networks ở các depths khác nhau, kết nối bởi dense skip pathways.

Đặc điểm:

- Skip connections dày đặc hơn
- Ensemble ngầm của nhiều U-Nets
- Tốt hơn nhưng phức tạp và chậm hơn

4. 3D U-Net:

Mở rộng sang 3D cho dữ liệu volumetric (CT, MRI):

- Conv 2D → Conv 3D (kernel $3 \times 3 \times 3$)
- Max Pooling 2D → 3D ($2 \times 2 \times 2$)
- Xử lý toàn bộ thể tích, nắm bắt context 3D

Thách thức:

- Yêu cầu bộ nhớ GPU rất lớn
- Thường phải train trên patches nhỏ, mất context
- Thời gian training và inference lâu

5. V-Net:

Biến thể 3D của U-Net với:

- Residual connections
- Dice loss function thay vì cross-entropy
- Phù hợp cho y tế với class imbalance nghiêm trọng

6. nnU-Net (no-new-Net):

Framework tự động cấu hình U-Net dựa trên dữ liệu:

- Tự động chọn 2D vs 3D vs 2.5D
- Tự động điều chỉnh architecture (depths, channels)
- Tự động preprocessing và augmentation

Đạt state-of-the-art trên nhiều datasets y tế mà không cần tuning thủ công.

Ưu điểm của U-Net trong y tế:

1. **Data efficiency:** Hoạt động tốt với vài trăm ảnh training
2. **Precise localization:** Skip connections cho ranh giới sắc nét
3. **Flexible:** Dễ điều chỉnh cho nhiều modalities và organs
4. **Fast inference:** Có thể segment ảnh trong vài giây
5. **Simple:** Dễ implement, debug và hiểu

Hạn chế:

- Memory intensive:** Đặc biệt với 3D volumes lớn
- Fixed receptive field:** Khó nắm bắt long-range dependencies
- Không bắt biến với scale/rotation:** Phải dựa vào augmentation

Ứng dụng:

U-Net đã thành công trong vô số tác vụ y tế:

- Brain tumor segmentation (BraTS)
- Organ segmentation (liver, kidney, heart)
- Lesion detection (skin lesions, lung nodules)
- Cell segmentation (microscopy images)
- Vessel segmentation (retinal vessels, coronary arteries)

Trong đồ án này, chúng em sử dụng biến thể 2D U-Net với các cải tiến:

- Batch Normalization sau mỗi Conv
- Dropout để regularization
- Xavier initialization
- Combined loss (Dice + CE + Focal)

2.3.3 Các kỹ thuật tăng cường dữ liệu cho ảnh y khoa

Data augmentation là kỹ thuật tạo ra dữ liệu huấn luyện mới từ dữ liệu có sẵn thông qua các phép biến đổi, mà không thay đổi nhãn (label). Đây là một trong những kỹ thuật quan trọng nhất để cải thiện hiệu năng và khả năng tổng quát hóa của mô hình học sâu, đặc biệt trong lĩnh vực y tế nơi dữ liệu gán nhãn rất khan hiếm.

Mục tiêu của Data Augmentation:

- Tăng kích thước dataset:** Từ N mẫu gốc, tạo ra $N \times K$ mẫu biến thể
- Cải thiện tính robust:** Mô hình học bắt biến với các biến đổi không quan trọng
- Giảm overfitting:** Mạng không thể "ghi nhớ" dữ liệu training
- Mô phỏng biến thể thực tế:** Phản ánh sự đa dạng trong dữ liệu thực tế (vị trí bệnh nhân, góc chụp, v.v.)

Các nhóm kỹ thuật augmentation:

1. Geometric Transformations (Biến đổi hình học):

Các phép biến đổi không gian bảo toàn cấu trúc giải phẫu.

a) Flipping (Lật):

- Horizontal flip: Lật trái-phải
- Vertical flip: Lật trên-dưới
- Trong nǎo bô: Chỉ nên dùng horizontal flip vì nǎo gần như đối xứng trái-phải, nhưng không đối xứng trên-dưới

b) Rotation (Xoay):

- Xoay ảnh một góc $\theta \in [-\theta_{\max}, \theta_{\max}]$
- Với ảnh nǎo: θ_{\max} thường là $10-15^\circ$ để tránh biến dạng quá mức
- Sử dụng interpolation (bilinear, bicubic) cho ảnh và nearest-neighbor cho mask

c) Scaling (Co giǎn):

- Phóng to/thu nhỏ với tỷ lệ $s \in [s_{\min}, s_{\max}]$
- Mô phỏng sự khác biệt kích thước khói u, kích thước nǎo

d) Translation (Tịnh tiến):

- Dịch chuyển ảnh theo x, y một số pixels
- Mô phỏng vị trí khác nhau trong khung hình

e) Affine Transformation (Biến đổi affine):

- Kết hợp rotation, scaling, translation, shearing
- Ma trận biến đổi affine:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & t_x \\ c & d & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

f) Elastic Deformation (Biến dạng đàn hồi):

Đây là kỹ thuật đặc biệt quan trọng trong y tế, mô phỏng biến thể sinh học tự nhiên.

Phương pháp:

1. Tạo displacement field ngẫu nhiên trơn (smooth random field):
 - Sinh noise Gaussian
 - Làm mượt bằng Gaussian filter với σ lớn
2. Áp dụng displacement field lên mỗi pixel: $\text{new_position} = \text{old_position} + \alpha \times \text{displacement}$

Tham số α kiểm soát mức độ biến dạng.

Lợi ích: Mô phỏng:

- Sự khác biệt giải phẫu giữa các cá nhân
- Biến dạng mô não do khối u chèn ép
- Biến thể trong tư thế bệnh nhân

2. Intensity/Color Transformations (Biến đổi cường độ):

Các phép biến đổi cường độ pixel, quan trọng với ảnh y tế do sự biến thiên về thiết bị và giao thức chụp.

a) Brightness Adjustment

$$I_{\text{new}} = I_{\text{old}} + \beta, \quad \beta \in [-\beta_{\max}, \beta_{\max}]$$

b) Contrast Adjustment

$$I_{\text{new}} = \alpha, (I_{\text{old}} - \mu) + \mu, \quad \alpha \in [\alpha_{\min}, \alpha_{\max}]$$

với μ là mean (tổng ảnh hoặc vùng). α quanh [0.8,1.2] là hợp lý — thay đổi tương phản.

c) Gamma Correction

$$I_{\text{new}} = I_{\text{old}}^{\gamma}, \quad \gamma \in [0.8, 1.2]$$

Điều chỉnh độ sáng phi tuyến, mô phỏng các cài đặt khác nhau của máy MRI.

d) **Gaussian Noise:** Thêm nhiễu Gaussian để mô phỏng nhiễu nhiệt trong ảnh MRI.

$$I_{\text{new}} = I_{\text{old}} + \mathcal{N}(0, \sigma^2)$$

e) **Gaussian Blur:** Làm mờ nhẹ để mô phỏng chuyển động bệnh nhân hoặc artifacts.

$$I_{\text{new}} = I_{\text{old}} * G(\sigma)$$

f) **Histogram Matching:** Biến đổi histogram của ảnh để khớp với histogram mục tiêu, giúp chuẩn hóa giữa các máy chụp khác nhau.

3. Mixed Sample Data Augmentation:

a) **Mixup:** Tạo ảnh mới bằng cách kết hợp tuyến tính hai ảnh.

$$x_{\text{new}} = \lambda x_i + (1 - \lambda)x_j$$

$$y_{\text{new}} = \lambda y_i + (1 - \lambda)y_j, \quad \lambda \sim \text{Beta}(\alpha, \alpha)$$

Nhược điểm trong segmentation: Khó áp dụng vì làm mờ ranh giới giữa các classes.

b) **CutMix:** Cắt một vùng từ ảnh này và dán vào ảnh khác, kèm theo mask tương ứng.

Hiệu quả hơn Mixup cho segmentation nhưng cần cẩn thận để không tạo ra mẫu không realistic (như tumor xuất hiện ở vị trí không hợp lý).

4. Advanced Augmentation Techniques:

a) **AutoAugment / RandAugment**: Tự động tìm kiếm policy augmentation tối ưu thông qua reinforcement learning hoặc random search.

b) **CycleGAN-based Augmentation**: Sử dụng GAN để chuyển đổi giữa các domains (ví dụ: MRI 1.5T ↔ MRI 3T).

Lợi ích: Giảm domain shift **Nhược điểm:** Phức tạp, tốn tài nguyên

c) **Adversarial Training**: Thêm perturbations nhỏ được tối ưu hóa để "đánh lừa" mô hình, tăng robustness.

Các nguyên tắc khi áp dụng augmentation trong y tế:

1. Domain Knowledge:

- **PHẢI**: Đảm bảo augmentation không tạo ra mẫu phi lý về mặt y học
- **VÍ DỤ**: Không flip vertical ảnh não (não không đối xứng trên-dưới)
- **VÍ DỤ**: Không rotate quá 15-20° (bệnh nhân không nằm quá nghiêng)

2. Label Preservation:

- Biến đổi phải bảo toàn nhãn
- Geometric transforms: Phải apply đồng nhất cho ảnh và mask
- Intensity transforms: Chỉ apply cho ảnh, không apply cho mask

3. Modality-Specific:

- MRI: Intensity augmentation mạnh (do không có scale chuẩn)
- CT: Intensity augmentation yếu hơn (có Hounsfield units chuẩn)

4. Realistic vs Diversity: Cân bằng giữa:

- Tạo nhiều biến thể (diversity)
- Giữ tính realistic (không quá khác với dữ liệu thực)

5. On-the-fly vs Offline:

- **On-the-fly**: Augment trong quá trình training, mỗi epoch thấy mẫu khác nhau
 - **Ưu điểm**: Đa dạng vô hạn, không tốn bộ nhớ
 - **Nhược điểm**: Tốn thời gian tính toán
- **Offline**: Tạo và lưu trước augmented data
 - **Ưu điểm**: Training nhanh hơn

- Nhược điểm: Tốn bộ nhớ, giới hạn số lượng

Kết quả thực nghiệm:

Trong các cuộc thi BraTS, các đội top thường sử dụng augmentation mạnh mẽ:

- Elastic deformation với $\alpha \in [0, 900]$, $\sigma \in [9, 13]$
- Rotation $\pm 15^\circ$
- Scaling 0.85-1.15
- Gamma 0.7-1.5
- Gaussian noise $\sigma = 0.1$
- Gaussian blur $\sigma \in [0.5, 1.5]$

Các nghiên cứu cho thấy augmentation có thể cải thiện Dice score **2-5%** so với không augmentation, đặc biệt quan trọng với datasets nhỏ (<500 samples).

2.4 Tập dữ liệu BraTS

2.4.1 Giới thiệu cuộc thi BraTS

Brain Tumor Segmentation (BraTS) là một thử thách nghiên cứu quốc tế được tổ chức thường niên nhằm đánh giá các phương pháp phân vùng khối u não tự động trên dữ liệu MRI đa chuỗi. Cuộc thi BraTS lần đầu tiên được khởi xướng vào năm 2012 tại hội nghị MICCAI (Medical Image Computing and Computer Assisted Intervention) và đã trở thành một trong những benchmark chuẩn quan trọng nhất trong lĩnh vực phân tích hình ảnh y tế.

Mục tiêu của cuộc thi BraTS:

BraTS nhằm thúc đẩy sự phát triển của các thuật toán phân vùng khối u não tự động với độ chính xác cao, hỗ trợ các bác sĩ trong quá trình chẩn đoán và lập kế hoạch điều trị. Cuộc thi tập trung vào hai loại u não chính:

- **Gliomas cấp thấp LGG:** Khối u phát triển chậm, ít xâm lấn, tiên lượng tốt hơn.
- **Glioblastomas cấp cao HGG:** Khối u ác tính, phát triển nhanh, xâm lấn mạnh vào mô não lành, tiên lượng xấu.

Đặc điểm của tập dữ liệu BraTS:

Tập dữ liệu BraTS được thu thập từ nhiều trung tâm y tế trên toàn thế giới với sự tham gia của các chuyên gia thần kinh học hàng đầu. Mỗi ca bệnh trong BraTS bao gồm bốn chuỗi MRI (modalities) được chụp với các giao thức khác nhau, cung cấp thông tin đa dạng về cấu trúc và tính chất của khối u:

- **T1-weighted (T1):** Hình ảnh MRI trọng số T1, cung cấp độ tương phản cao giữa các loại mô mềm.

- **T1-weighted with contrast enhancement (T1ce):** Hình ảnh T1 sau khi tiêm chất cản quang, làm nổi bật vùng khối u tăng sinh mạch máu.
- **T2-weighted (T2):** Hình ảnh MRI trọng số T2, làm nổi bật vùng phù nề quanh khối u.
- **FLAIR (Fluid Attenuated Inversion Recovery):** Chuỗi FLAIR giúp phát hiện tổn thương và vùng phù nề bằng cách loại bỏ tín hiệu từ dịch não tủy.

Dữ liệu BraTS đã được tiền xử lý chuẩn hóa, bao gồm việc đồng bộ không gian (co-registration), nội suy về cùng kích thước voxel (1mm^3), và loại bỏ hộp sọ (skull-stripping) để tập trung vào vùng não.

Ý nghĩa khoa học và thực tiễn:

BraTS không chỉ là một cuộc thi mà còn là một nền tảng nghiên cứu mở, cho phép các nhà khoa học trên toàn thế giới so sánh công bằng các thuật toán của mình trên cùng một tập dữ liệu chuẩn. Các kết quả từ BraTS đã góp phần thúc đẩy sự phát triển mạnh mẽ của deep learning trong y tế, đặc biệt là các kiến trúc U-Net, V-Net, và các biến thể Transformer.

2.4.2 Dữ liệu và nhãn phân vùng (WT, TC, ET)

Trong BraTS, mỗi ca bệnh được cung cấp kèm theo một nhãn phân vùng (ground truth segmentation) được chú thích thủ công bởi các chuyên gia thần kinh học. Nhãn phân vùng xác định ranh giới của từng vùng khối u dựa trên đặc điểm sinh học và hình ảnh học. BraTS định nghĩa **bốn lớp (classes)** chính trong không gian nhãn:

Các lớp nhãn cơ bản:

Label 0 (Background): Vùng nền, bao gồm mô não bình thường và khoảng không.

Label 1 (NCR - Necrotic and Non-enhancing Tumor Core): Vùng hoại tử và lõi khối u không tăng cường. Đây là vùng mô chết do thiếu oxy, thường xuất hiện ở trung tâm khối u ác tính.

Label 2 (ED - Peritumoral Edema): Vùng phù nề xung quanh khối u. Phù nề là hiện tượng tích tụ dịch trong mô não do khối u gây ra, làm tăng áp lực nội sọ.

Label 4 (ET - Enhancing Tumor): Vùng khối u tăng cường sau khi tiêm chất cản quang. Đây là vùng có hoạt động tăng sinh mạch máu mạnh, đặc trưng của khối u ác tính.

Lưu ý: Trong một số phiên bản BraTS, label 3 không được sử dụng, do đó label 4 được ánh xạ về label 3 để đơn giản hóa quá trình huấn luyện mô hình.

Các vùng tổ hợp (composite regions):

BraTS yêu cầu các mô hình dự đoán ba vùng tổ hợp chính, được định nghĩa dựa trên các lớp nhãn cơ bản:

Whole Tumor (WT) - Toàn bộ khối u:

- $WT = NCR + ED + ET = \{1, 2, 4\}$
- Bao gồm toàn bộ vùng bất thường liên quan đến khối u, từ lõi hoại tử đến vùng phù nề ngoại vi.

Tumor Core (TC) - Lõi khối u:

- $TC = NCR + ET = \{1, 4\}$
- Bao gồm vùng hoại tử và vùng tăng cường, đại diện cho phần chính của khối u, không tính phù nề.

Enhancing Tumor (ET) - Khối u tăng cường:

- $ET = \{4\}$
- Chỉ bao gồm vùng tăng cường, là vùng có hoạt động sinh học cao nhất, thường được sử dụng để đánh giá hiệu quả điều trị.

Ý nghĩa lâm sàng:

WT: Xác định ranh giới toàn bộ tổn thương, quan trọng cho phẫu thuật và xạ trị.

TC: Đánh giá kích thước khối u thực sự, giúp phân loại mức độ ác tính.

ET: Theo dõi đáp ứng điều trị, vì vùng tăng cường thường giảm khi điều trị hiệu quả.

Định dạng lưu trữ:

Dữ liệu BraTS thường được lưu dưới định dạng NIfTI (.nii.gz) hoặc HDF5 (.h5). Trong nghiên cứu này, dữ liệu được chuyển đổi sang định dạng HDF5 để tối ưu tốc độ đọc và xử lý. Mỗi file H5 chứa năm dataset:

- flair, t1, t1ce, t2: Bộ chuỗi MRI với shape (H, W, D), thường là (240, 240, 155).
- seg: Nhãn phân vùng với shape (H, W, D), giá trị pixel thuộc {0, 1, 2, 4}.

2.4.3 Các chuẩn đánh giá trong BraTS

BraTS sử dụng một bộ chỉ số đánh giá nghiêm ngặt để đảm bảo tính công bằng và khách quan khi so sánh các thuật toán. Các chỉ số này được tính toán riêng biệt cho từng vùng tổ hợp (WT, TC, ET) và sau đó lấy trung bình để đánh giá hiệu suất tổng thể.

Chỉ số chính:

Dice Similarity Coefficient (DSC):

- Đo độ tương đồng giữa vùng dự đoán và vùng ground truth.
- Giá trị từ 0 (không trùng khớp) đến 1 (trùng khớp hoàn toàn).
- Công thức: $DSC = 2|P \cap G| / (|P| + |G|)$
- **DSC là chỉ số chính được sử dụng để xếp hạng các đội tham gia BraTS.**

Sensitivity (Recall):

- Tỷ lệ voxel khối u thực sự được dự đoán đúng.
- Quan trọng để đảm bảo không bỏ sót vùng khối u.
- Công thức: Sensitivity = $TP / (TP + FN)$

Specificity:

- Tỷ lệ voxel nền (không phải khối u) được dự đoán đúng.
- Quan trọng để tránh dự đoán nhầm mô lành thành khối u.
- Công thức: Specificity = $TN / (TN + FP)$

Quy trình đánh giá:

- Mỗi thuật toán phải dự đoán ba vùng: WT, TC, ET.
- Các chỉ số DSC và HD95 được tính riêng cho từng vùng.
- Điểm cuối cùng là trung bình của các chỉ số trên ba vùng.
- BraTS cung cấp một leaderboard công khai, xếp hạng các đội dựa trên DSC trung bình.

Ngưỡng chấp nhận:

- **DSC > 0.75:** Hiệu suất tốt, chấp nhận được cho ứng dụng lâm sàng.
- **DSC > 0.85:** Hiệu suất xuất sắc, sánh ngang với chuyên gia.
- **HD95 < 5mm:** Ranh giới chính xác, phù hợp cho lập kế hoạch điều trị.

Ý nghĩa thực tiễn:

BraTS không chỉ là một cuộc thi mà còn là một công cụ kiểm tra chất lượng (quality assurance) cho các hệ thống phân vùng u não tự động trước khi triển khai lâm sàng. Các chỉ số đánh giá của BraTS đã trở thành tiêu chuẩn de facto trong cộng đồng nghiên cứu y tế.

2.5 Các chỉ số đánh giá mô hình phân vùng

Đánh giá hiệu suất của mô hình phân vùng y tế đòi hỏi một tập hợp các chỉ số đa dạng, phản ánh nhiều khía cạnh khác nhau của chất lượng dự đoán. Phản này trình bày chi tiết các chỉ số phổ biến nhất, bao gồm định nghĩa toán học, cách tính toán, và ý nghĩa lâm sàng.

2.5.1 Dice Similarity Coefficient (DSC)

Định nghĩa:

Dice Similarity Coefficient (DSC), còn gọi là Sørensen-Dice index hoặc F1-score trong bài toán phân vùng, là một chỉ số đo độ tương đồng giữa hai tập hợp. Trong bối cảnh phân vùng hình ảnh, DSC đo mức độ trùng khớp giữa vùng dự đoán (P) và vùng ground truth (G).

Công thức toán học:

$$DSC(P, G) = \frac{2 \times |P \cap G|}{|P| + |G|} = \frac{2 \times TP}{2 \times TP + FP + FN}$$

Trong đó:

- $|P \cap G|$: Số voxel được dự đoán đúng (true positives - TP)
- $|P|$: Tổng số voxel được dự đoán là khối u
- $|G|$: Tổng số voxel khối u thực tế
- FP (False Positives): Voxel bị dự đoán nhầm là khối u
- FN (False Negatives): Voxel khối u thực sự bị bỏ sót

Tính chất:

- **Miền giá trị:** $DSC \in [0, 1]$
 - $DSC = 1$: Dự đoán hoàn hảo, trùng khớp 100%
 - $DSC = 0$: Không có điểm nào trùng khớp
- **Tính đối xứng:** $DSC(P, G) = DSC(G, P)$
- **Độ nhạy với kích thước:** DSC cân bằng giữa precision và recall, phù hợp với các vùng có kích thước khác nhau.
- **Xử lý class imbalance:** DSC tập trung vào vùng khối u (foreground), ít bị ảnh hưởng bởi số lượng lớn voxel nền.

Ưu điểm:

- Cân bằng giữa precision và recall.
- Phù hợp với dữ liệu mất cân bằng (khối u chiếm tỷ lệ nhỏ so với nền).
- Trực quan, dễ hiểu, phổ biến trong y tế.
- Tương đương với F1-score, một chỉ số quen thuộc trong machine learning.

Hạn chế:

- Nhạy cảm với các vùng nhỏ: Một sai số nhỏ trên vùng nhỏ có thể làm giảm DSC đáng kể.

- Không đánh giá trực tiếp ranh giới: Hai dự đoán có DSC giống nhau có thể có ranh giới rất khác nhau.
- Không phản ánh khoảng cách không gian giữa dự đoán và ground truth.

Nguồn ứng dụng lâm sàng:

- **DSC < 0.70:** Chất lượng kém, không khuyến nghị sử dụng.
- **DSC 0.70 - 0.80:** Chất lượng trung bình, cần cải thiện.
- **DSC 0.80 - 0.90:** Chất lượng tốt, chấp nhận được cho hỗ trợ chẩn đoán.
- **DSC > 0.90:** Chất lượng xuất sắc, sánh ngang với chuyên gia.

2.5.2 Intersection over Union (IoU)

Định nghĩa:

Intersection over Union (IoU), còn gọi là Jaccard Index, là một chỉ số đo tỷ lệ giữa diện tích giao và diện tích hợp của hai tập hợp. IoU được sử dụng rộng rãi trong computer vision, đặc biệt là object detection và semantic segmentation.

Công thức toán học:

$$\text{IoU}(P, G) = \frac{|P \cap G|}{|P \cup G|} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}$$

Trong đó :

- $|P \cap G|$: Diện tích giao (intersection)
- $|P \cup G|$: Diện tích hợp (union)
- TP, FP, FN: True positives, false positives, false negatives

Mối quan hệ với DSC:

$$\text{DSC} = \frac{2 \times \text{IoU}}{1 + \text{IoU}}$$

$$\text{IoU} = \frac{\text{DSC}}{2 - \text{DSC}}$$

Như vậy, DSC và IoU có mối quan hệ đơn điệu, nhưng DSC luôn lớn hơn hoặc bằng IoU cho cùng một cặp dự đoán.

Tính chất:

- **Miền giá trị:** $\text{IoU} \in [0, 1]$
- **Tính đối xứng:** $\text{IoU}(P, G) = \text{IoU}(G, P)$
- **Nghiêm ngặt hơn DSC:** IoU phạt nặng hơn các dự đoán sai, do đó thường có giá trị thấp hơn DSC.

- **Giải thích trực quan:** IoU = 0.7 nghĩa là 70% diện tích hợp được dự đoán đúng.

Ưu điểm:

- Phạt nặng hơn các dự đoán có nhiều false positives hoặc false negatives.
- Phổ biến trong computer vision, dễ so sánh với các nghiên cứu khác.
- Tương đương với Jaccard Index, một chỉ số thống kê nổi tiếng.

Hạn chế:

- Giá trị thấp hơn DSC, có thể khó đạt ngưỡng cao.
- Không phản ánh khoảng cách không gian giữa dự đoán và ground truth.
- Nhạy cảm với các vùng nhỏ, giống như DSC.

Ngưỡng ứng dụng lâm sàng:

- **IoU < 0.50:** Chất lượng kém.
- **IoU 0.50 - 0.70:** Chất lượng trung bình.
- **IoU 0.70 - 0.85:** Chất lượng tốt.
- **IoU > 0.85:** Chất lượng xuất sắc.

2.5.3 Sensitivity và Specificity

Sensitivity (Recall, True Positive Rate):

Định nghĩa:

Sensitivity đo tỷ lệ các voxel khối u thực sự được dự đoán đúng. Đây là chỉ số quan trọng nhất trong y tế để đảm bảo không bỏ sót bệnh lý.

Công thức:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{TP}}{|G|}$$

Trong đó:

- TP: True positives (voxel khối u được dự đoán đúng)
- FN: False negatives (voxel khối u bị bỏ sót)
- |G|: Tổng số voxel khối u thực tế

Ý nghĩa lâm sàng:

- **Sensitivity cao (> 0.95):** Mô hình có khả năng phát hiện hầu hết các vùng khối u, ít bỏ sót.
- **Sensitivity thấp:** Mô hình bỏ sót nhiều vùng khối u, nguy hiểm trong chẩn đoán.

- **Trade-off:** Tăng sensitivity thường làm tăng false positives (giảm precision).

Specificity (True Negative Rate):

Định nghĩa:

Specificity đo tỷ lệ các voxel nền (không phải khối u) được dự đoán đúng. Chỉ số này quan trọng để tránh chẩn đoán sai, gây hoảng sợ cho bệnh nhân.

Công thức:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

Trong đó:

- TN: True negatives (voxel nền được dự đoán đúng)
- FP: False positives (voxel nền bị dự đoán nhầm là khối u)

Ý nghĩa lâm sàng:

- **Specificity cao (> 0.95):** Mô hình ít dự đoán nhầm mô lành thành khối u.
- **Specificity thấp:** Mô hình có xu hướng over-segment, tạo ra nhiều false alarms.
- **Trade-off:** Tăng specificity thường làm giảm sensitivity.

Precision (Positive Predictive Value):

Công thức:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{\text{TP}}{|\text{P}|}$$

Precision đo tỷ lệ các voxel được dự đoán là khối u thực sự đúng. Precision cao → ít false positives.

Ứng dụng trong BraTS:

- **Sensitivity:** Đặc biệt quan trọng cho vùng ET (khối u tăng cường), vì bỏ sót ET có thể dẫn đến đánh giá sai mức độ ác tính.
- **Specificity:** Quan trọng cho vùng ED (phù nề), vì dự đoán nhầm mô lành thành phù nề có thể dẫn đến điều trị thừa.
- **Balance:** Trong thực tế lâm sàng, người ta thường ưu tiên sensitivity cao hơn để đảm bảo an toàn bệnh nhân (principle of "do no harm").

2.5.4 Các chỉ số bổ sung khác

1. Average Symmetric Surface Distance (ASSD):

ASSD đo khoảng cách trung bình giữa các điểm trên biên dự đoán và biên ground truth, theo cả hai chiều.

$$ASSD(P, G) = \frac{1}{|\partial P| + |\partial G|} \left(\sum_{p \in \partial P} d(p, G) + \sum_{g \in \partial G} d(g, P) \right)$$

- **Ưu điểm:** Ít nhạy cảm với outliers hơn HD, đánh giá toàn bộ biên.
- **Hạn chế:** Phức tạp tính toán, ít phổ biến hơn HD95.

2. Volumetric Similarity:

Đo tỷ lệ chênh lệch thể tích giữa dự đoán và ground truth.

$$VS = 1 - \frac{| |P| - |G| |}{|P| + |G|}$$

- **Ứng dụng:** Đánh giá khả năng ước lượng kích thước khối u, quan trọng cho theo dõi điều trị.

3. Matthews Correlation Coefficient (MCC):

MCC là một chỉ số cân bằng, xét cả bốn giá trị TP, TN, FP, FN.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

- **Ưu điểm:** Không bị ảnh hưởng bởi class imbalance, phản ánh toàn diện chất lượng dự đoán.
- **Miền giá trị:** MCC $\in [-1, 1]$, với 1 là hoàn hảo, 0 là ngẫu nhiên, -1 là ngược hoàn toàn.

4. Relative Volume Error (RVE):

$$RVE = \frac{|P| - |G|}{|G|}$$

- **Ý nghĩa:** Đo sai số tương đối về thể tích, quan trọng cho đánh giá kích thước khối u.
- **RVE > 0:** Over-segmentation (dự đoán quá lớn).
- **RVE < 0:** Under-segmentation (dự đoán quá nhỏ).

5. Confusion Matrix Metrics:

Ngoài các chỉ số đã nêu, ta có thể tính toán thêm:

- False Positive Rate (FPR): $FPR = FP / (FP + TN)$
- False Negative Rate (FNR): $FNR = FN / (FN + TP)$
- Accuracy: $Accuracy = (TP + TN) / (TP + TN + FP + FN)$

- Lưu ý: Accuracy ít được dùng trong phân vùng y tế do class imbalance nghiêm trọng.

Chỉ số	Ưu điểm	Hạn chế	Ứng dụng chính
DSC	Cân bằng, phô biến, dễ hiểu	Không đánh giá ranh giới	Xếp hạng tổng thể
IoU	Nghiêm ngặt, phô biến CV	Giá trị thấp hơn DSC	So sánh với CV
Sensitivity	Phát hiện bệnh, an toàn BN	Trade-off với precision	Screening, phát hiện sớm
Specificity	Tránh false alarm	Trade-off với sensitivity	Tránh điều trị thừa
HD95	Đánh giá ranh giới, quan trọng cho phẫu thuật	Phức tạp, không differentiable	Lập kê hoạch điều trị
ASSD	Đánh giá toàn bộ biên	Phức tạp tính toán	Đánh giá chi tiết

Bảng so sánh tổng hợp các chỉ số

CHƯƠNG 3. PHƯƠNG PHÁP NGHIÊN CỨU

3.1 Chuẩn bị dữ liệu

3.1.1 Dữ Liệu Sử Dụng

Nguồn dữ liệu: Bộ dữ liệu BraTS (Brain Tumor Segmentation Challenge) là bộ dữ liệu công khai tiêu chuẩn được sử dụng rộng rãi trong nghiên cứu phân đoạn khối u não. Mỗi trường hợp bệnh nhân bao gồm:

- **4 loại ảnh MRI:** FLAIR, T1, T1ce (T1 với chất cản trở), T2 với độ phân giải 240×240 pixels và chiều sâu 155 lát cắt
- **Ảnh chân lý cơ sở (Ground Truth):** Được gắn nhãn thủ công bởi chuyên gia y tế với 4 lớp:
 - Lớp 0: Nền (background)
 - Lớp 1: Vùng hoại tử (Necrotic core - NCR)
 - Lớp 2: Vùng phù nề (Edema - ED)
 - Lớp 3: Vùng khối u hoạt động (Enhancing tumor - ET)

Phân chia dữ liệu: Toàn bộ dữ liệu được chia thành ba tập hợp:

- Tập huấn luyện: 700 mẫu (bệnh nhân)
- Tập xác thực: 200 mẫu
- Tập kiểm thử: 100 mẫu

Tỷ lệ chia 7:2:1 đảm bảo mô hình có đủ dữ liệu để học trong khi vẫn có tập xác thực và kiểm thử đơn độc để đánh giá hiệu suất khách quan.

3.1.2 Xử Lý Dữ Liệu

a) Chuyển đổi định dạng

Toàn bộ dữ liệu được chuyển đổi từ định dạng NIFTI sang định dạng HDF5 để tối ưu hóa tốc độ đọc và tiết kiệm bộ nhớ. Định dạng HDF5 cho phép truy cập hiệu quả các dãy dữ liệu lớn trong quá trình huấn luyện.

b) Chuẩn hóa dữ liệu

Mỗi phương thức MRI (FLAIR, T1, T1ce, T2) được chuẩn hóa độc lập sử dụng phương pháp Z-score:

$$z = \frac{x - \mu}{\sigma}$$

Trong đó:

- x là giá trị pixel gốc
- μ là giá trị trung bình của phương thức

- σ là độ lệch chuẩn

Chuẩn hóa Z-score giúp: Loại bỏ ảnh hưởng của các khác biệt về cường độ giữa các bệnh nhân. Tăng tốc độ hội tụ của mô hình. Giảm bớt ảnh hưởng của nhiễu trong dữ liệu

c) Lựa chọn lát cắt

Do các lát cắt ở đầu và cuối thường chứa thông tin ít ý nghĩa, hệ thống loại bỏ 30 lát cắt từ đầu và 30 lát cắt từ cuối (slice_range = (30, -30)). Điều này tập trung xử lý vào vùng não chứa khối u với thông tin giá trị cao.

Để giải quyết vấn đề mất cân bằng lớp, hệ thống ưu tiên chọn các lát cắt có chứa khối u (nonzero_slices). Nếu lát cắt không có khối u, hệ thống sẽ chọn ngẫu nhiên từ khoảng lát cắt hợp lệ.

d) Tăng cường dữ liệu (Data Augmentation)

Để tăng kích thước hiệu quả của tập huấn luyện và cải thiện khả năng tổng quát hóa, các kỹ thuật tăng cường dữ liệu sau được áp dụng:

Lật ngang/dọc: Xác suất 50% lật ảnh theo trục x hoặc y, mờ phông các biến thể giải phẫu tự nhiên

Xoay: Xoay ngẫu nhiên 90° , 180° , hoặc 270° với xác suất 50%

Thêm nhiễu: Thêm nhiễu Gaussian với độ lệch chuẩn 0.03 với xác suất 50%

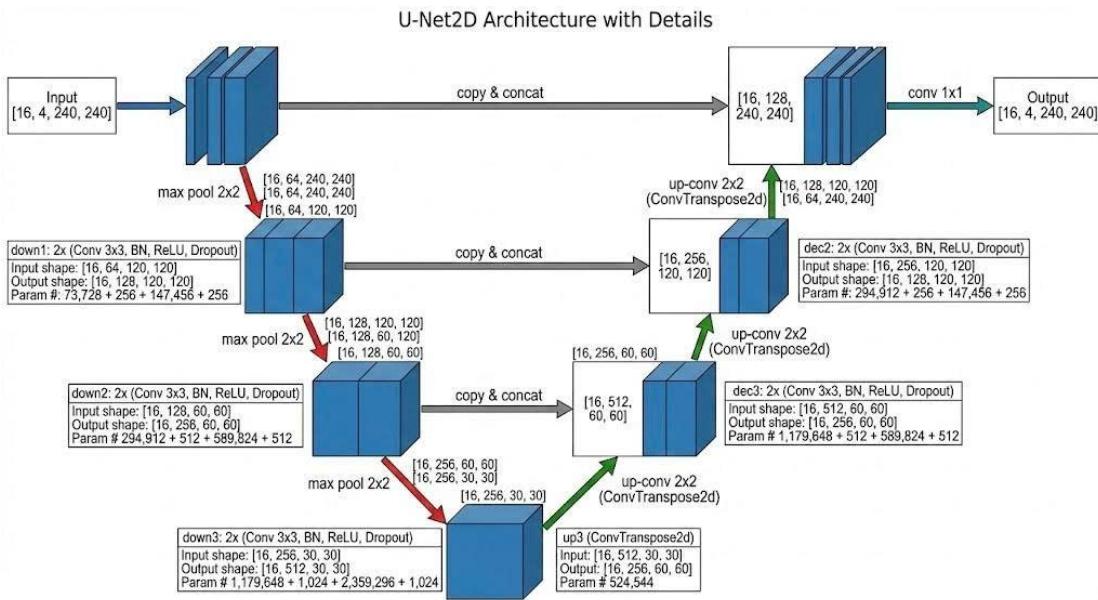
• **Điều chỉnh độ sáng:** Nhân ảnh với hệ số trong khoảng [0.9, 1.1] với xác suất 50%

• **Hiệu chỉnh gamma:** Áp dụng biến đổi gamma với $\gamma \in [0.85, 1.15]$ để điều chỉnh tương phản

Các kỹ thuật tăng cường này được áp dụng chỉ trong quá trình huấn luyện, không áp dụng cho tập xác thực và kiểm thử để đảm bảo đánh giá công bằng.

3.2 Mô Hình Đề Xuất

3.2.1 Kiến Trúc UNet2D



Hình 3. Kiến trúc mô hình U-net2D

Chi tiết các thành phần trong kiến trúc được mô tả như sau:

1. Đầu vào (Input Layer):

Mô hình nhận đầu vào là một tensor có kích thước ($H \times W \times 4$), tương ứng với 4 kênh hình ảnh MRI đã qua tiền xử lý (T1, T1ce, T2, FLAIR).

Kích thước không gian $H \times W$ (240 x 240).

2. Phần Encoder (Contracting Path - Trích xuất đặc trưng):

Phần này đóng vai trò như một mạng Convolutional Network thông thường để nén ảnh và trích xuất các đặc trưng ngữ nghĩa mức cao. Encoder gồm 3 khối chính (inc, down1, down2):

Cấu trúc khói: Mỗi khói bao gồm hai lớp tích chập liên tiếp (DoubleConv). Dựa theo chú giải trong sơ đồ, mỗi thao tác tích chập bao gồm:

Conv2d 3x3 (padding=1) -> Batch Normalization -> ReLU Activation.

Việc thêm lớp Batch Normalization giúp quá trình huấn luyện ổn định và hội tụ nhanh hơn so với U-Net gốc.

Giảm kích thước: Sau mỗi khói DoubleConv là một lớp Max Pooling 2x2 (mũi tên đỏ) để giảm kích thước không gian đi một nửa, đồng thời số lượng bộ lọc (feature maps) được tăng gấp đôi để học các đặc trưng phức tạp hơn.

Số lượng kênh: Số kênh tăng dần theo thứ tự: 64 -> 128 -> 256.

3. Phần Bottleneck (Cầu nối):

Đây là phần sâu nhất của mạng (đáy chữ U), nơi chứa các đặc trưng trừu tượng nhất của ảnh.

Tại đây, kích thước không gian là nhỏ nhất ($H/8 \times W/8$) nhưng số lượng kênh đạt cực đại là **512**. Khối này đóng vai trò cầu nối chuyển giao thông tin từ Encoder sang Decoder.

4. Phần Decoder (Expansive Path - Khôi phục không gian):

Phần này có nhiệm vụ khôi phục lại kích thước ảnh gốc từ các đặc trưng đã được mã hóa. Decoder gồm 3 khối (dec3, dec2, dec1):

Up-sampling: Sử dụng phép tích chập chuyển vị Up-conv 2x2 (hoặc ConvTranspose2d - mũi tên xanh lá) để phóng to kích thước đặc trưng lên gấp đôi.

Kết nối tắt (Skip Connections): Các đặc trưng sau khi up-sampling được ghép nối (concatenated - mũi tên xám) với các đặc trưng có kích thước tương ứng từ bên Encoder. Cơ chế này (thể hiện qua các khối trắng + xanh trong sơ đồ) giúp khôi phục các thông tin chi tiết về vị trí không gian bị mất trong quá trình pooling, làm cho đường biên phân đoạn sắc nét hơn.

Giảm số kênh: Sau khi ghép nối, dữ liệu đi qua khối DoubleConv để giảm dần số lượng kênh: 512 -> 256 -> 128 -> 64.

5. Đầu ra (Output Layer):

Lớp cuối cùng sử dụng phép tích chập Conv 1x1 (mũi tên xanh đậm) để ánh xạ vector đặc trưng 64 chiều về số lượng lớp cần phân đoạn.

Kết quả đầu ra là bản đồ phân đoạn (segmentation map) có cùng kích thước với ảnh gốc ($H \times W$) với **4 kênh**, tương ứng với xác suất của 4 lớp nhãn: Nền (Background), Hoại tử (NCR), Phù nè (ED) và Tăng quang (ET).

3.2.2 Thành phần cơ bản: Khối DoubleConv2d

Đơn vị cấu trúc cơ bản của kiến trúc U-Net đề xuất là khối **DoubleConv2d**. Thay vì sử dụng một lớp tích chập đơn lẻ, mô hình sử dụng chuỗi lặp lại hai lần của các phép toán nhằm tăng khả năng trích xuất đặc trưng phi tuyến tính sâu hơn mà không làm mất mát quá nhiều thông tin không gian.

Quy trình xử lý dữ liệu qua một khối DoubleConv2d diễn ra tuần tự như sau:

- Tích chập (Convolution):** Sử dụng kernel kích thước 3 x 3 với padding bằng 1 để trích xuất các bản đồ đặc trưng (feature maps) mà vẫn giữ nguyên kích thước không gian của ảnh đầu vào.
- Chuẩn hóa theo lô (Batch Normalization):** Được áp dụng ngay sau tích chập để chuẩn hóa phân phối của các giá trị kích hoạt. Bước này giúp giảm hiện tượng biến thiên hiệp phương sai nội bộ (internal covariate shift), cho phép mô hình hội tụ nhanh hơn và ổn định hơn trong quá trình huấn luyện.
- Hàm kích hoạt (ReLU Activation):** Áp dụng hàm phi tuyến tính ReLU để loại bỏ các giá trị âm, giúp mạng học được các mối quan hệ phức tạp và phi tuyến tính của dữ liệu ảnh MRI.

4. **Dropout (Regularization):** Ở lần lặp thứ hai trong khối (ngoại trừ khối đầu tiên của Encoder), một lớp Dropout với tỷ lệ 0.1 được thêm vào nhằm ngẫu nhiên loại bỏ một số nơ-ron. Cơ chế này đóng vai trò quan trọng trong việc ngăn ngừa hiện tượng quá khớp (overfitting) khi huấn luyện trên tập dữ liệu y tế có số lượng mẫu hạn chế.

Đầu ra của khối DoubleConv2d ở nhánh Encoder sẽ được truyền qua lớp Max Pooling để giảm kích thước, đồng thời được sao chép (copy) để kết hợp (concatenate) với khối tương ứng bên nhánh Decoder thông qua kết nối tắt (skip connection).

3.2.3 Chi tiết cấu hình mô hình

Để đảm bảo tính minh bạch và khả năng tái lập của nghiên cứu, các thông số cấu hình chi tiết cho toàn bộ kiến trúc mạng U-Net 2D được thiết lập cố định như trình bày tại Bảng 3.1.

Tham số (Parameter)	Giá trị thiết lập (Value)	Ghi chú
Kênh đầu vào	4	Tương ứng 4 chuỗi MRI: FLAIR, T1, T1ce, T2
Lớp đầu ra	4	Background, Necrotic (NCR), Edema (ED), Enhancing (ET)
Số lượng bộ lọc	[64, 128, 256, 512]	Tăng dần ở Encoder và giảm dần ở Decoder
Kích thước hạt nhân	3x3	Áp dụng cho toàn bộ các lớp tích chập
Padding	1	Giữ nguyên kích thước không gian sau tích chập
Tỷ lệ Dropout	0.1	Áp dụng tại các khối Encoder và Decoder trung gian
Khởi tạo trọng số	Xavier Uniform	Đảm bảo phân phối cân bằng của trọng số ban đầu

Bảng 3.1: Cấu hình siêu tham số (Hyperparameters) của mô hình

Cơ chế khởi tạo trọng số: Toàn bộ các lớp tích chập trong mô hình được khởi tạo trọng số theo phương pháp **Xavier Uniform** (Glorot Uniform). Phương pháp này giúp duy trì phương sai của các giá trị kích hoạt và gradient qua các lớp, ngăn chặn hiện tượng bùng nổ hoặc triệt tiêu gradient (vanishing/exploding gradients), từ đó giúp mô hình hội tụ hiệu quả hơn.

Mô hình	Loại	Số tham số	Nhận xét
U-Net 2D	2D CNN	~7.7 triệu	Nhỏ, tối ưu cho GPU phổ thông
Standard U-Net (Ronneberger)	2D CNN	~31 triệu	Nặng hơn gấp 4 lần mô hình đề xuất
3D U-Net (Standard)	3D CNN	~19 triệu	Yêu cầu bộ nhớ VRAM rất lớn
VGG16-UNet	2D CNN	~23 triệu	Sử dụng backbone nặng nề

Bảng 3.2: So sánh độ phức tạp mô hình

Nhận xét: Với chỉ 7.7 triệu tham số, mô hình đề xuất giảm được khoảng **75%** lượng tham số so với kiến trúc U-Net gốc. Điều này mang lại hai lợi thế thực tiễn quan trọng:

- Giảm nguy cơ Overfitting:** Với tập dữ liệu BraTS có số lượng ca bệnh giới hạn, một mô hình nhỏ gọn giúp tránh việc học vẹt dữ liệu (memorization), từ đó tăng khả năng tổng quát hóa trên tập kiểm thử.
- Tối ưu tốc độ suy luận:** Kích thước nhỏ cho phép mô hình đạt tốc độ xử lý nhanh và dễ dàng triển khai trên các thiết bị có tài nguyên phần cứng hạn chế (như GPU laptop cá nhân hoặc máy tính tại bệnh viện tuyến dưới) mà không yêu cầu hạ tầng máy chủ đắt tiền.

3.3 Hàm Mất Mát và Độ Đo Hiệu Suất

3.3.1 Hàm Mất Mát Kết Hợp

Vì bài toán phân đoạn đa lớp có tính chất phức tạp, một hàm mất mát kết hợp được sử dụng:

$$L_{\text{tổng}} = 0.4 \cdot L_{\text{CE}} + 0.4 \cdot L_{\text{Dice}} + 0.2 \cdot L_{\text{Focal}}$$

a) Cross-Entropy Loss L_{CE}

$$L_{\text{CE}} = - \sum_{c=0}^C w_c \sum_i y_i^c \log(p_i^c)$$

Trong đó:

- $C=4$ là số lớp
- w_c là trọng số lớp (được gán lớn hơn cho lớp thiểu số)
- y_i^c là nhãn thực tế

- p_i^c là xác suất dự đoán

Cross-Entropy Loss là hàm mất mát chuẩn cho phân loại đa lớp, nhạy cảm với từng pixel.

b) Dice Loss L_{Dice}

$$L_{Dice} = 1 - \frac{2|X \cap Y|}{|X| + |Y|}$$

Trong đó:

- X là tập pixels được dự đoán là lớp c
- Y là tập pixels thực sự là lớp c

Dice Loss đặc biệt hiệu quả cho các tập dữ liệu mất cân bằng lớp, vì nó tập trung vào độ chính xác của sự khớp giữa vùng dự đoán và vùng thực tế.

c) Focal Loss (L_{Focal})

$$L_{Focal} = -\alpha_t (1-p_t)^\gamma \log(p_t)$$

Trong đó:

- $\gamma=2.0$ là tham số focus
- p_t là xác suất dự đoán của lớp thực

Focal Loss giảm trọng lượng của các mẫu dễ phân loại và tập trung vào các mẫu khó. Điều này giúp cải thiện hiệu suất trên các lớp khó, đặc biệt là lớp vùng hoại tử (NCR).

d) Trọng Số Lớp

Để xử lý mất cân bằng lớp:

$$w=[0.1, 2.0, 2.0, 3.0]$$

Trọng số cao hơn được gán cho các lớp hiếm (NCR, ED, ET) để giảm thiểu ảnh hưởng của lớp nền (background) chiếm phần lớn ảnh.

3.3.2 Độ Đo Hiệu Suất: Dice Coefficient

Độ đo Dice Coefficient được tính riêng lẻ cho từng lớp:

$$Dice_c = \frac{2 \cdot |X \cap Y|}{|X| + |Y|} + \epsilon$$

Trong đó:

- X là tập pixels được dự đoán là lớp c
- Y là tập pixels thực sự là lớp c
- $\epsilon=10^{-6}$ là hằng số làm mịn để tránh chia cho 0

Dice trung bình:

$$\text{Dice}_{\text{mean}} = \frac{1}{C-1} \sum_{c=1}^C \text{Dice}_c$$

Chỉ tính trung bình trên các lớp khối u (lớp 1, 2, 3), không tính lớp nền (lớp 0).

Ưu điểm của Dice Coefficient:

- Độc lập với mất cân bằng lớp
- Giải thích được: giá trị 1.0 = phân đoạn hoàn hảo, 0.0 = không có sự khớp
- Phù hợp với các bài toán y tế nơi cần độ chính xác cao

3.4 Cấu Hình Huấn Luyện

3.4.1 Tối Ưu Hóa và Lịch Học

Optimizer: AdamW

$$\theta_t = \theta_{t-1} - \alpha \cdot \frac{m_t}{\sqrt{v_t} + \epsilon}$$

Với:

- Learning rate $\alpha = 10^{-3}$
- $\beta_1 = 0.9$, $\beta_2 = 0.999$
- Weight decay $\lambda = 10^{-4}$

AdamW được chọn vì:

- Kết hợp ưu điểm của momentum và RMSprop
- Có weight decay tách biệt, giúp cải thiện tính tổng quát hóa
- Hiệu suất tốt cho các bài toán học sâu

Learning Rate Scheduler: Cosine Annealing with Warmup

Lịch học được chia thành hai giai đoạn:

- Warmup phase** (3 epochs đầu): Tăng learning rate tuyến tính từ 0 đến α

$$lr_t = \frac{t}{\text{num_warmup_steps}} \cdot \alpha$$

- Cosine decay phase** (37 epochs còn lại): Giảm learning rate theo hàm cosine

$$lr_t = \frac{\alpha}{2} \left(1 + \cos \left(\pi \cdot \frac{t - \text{num_warmup_steps}}{\text{num_total_steps} - \text{num_warmup_steps}} \right) \right)$$

Lợi ích:

- Warmup giúp ổn định quá trình huấn luyện ở giai đoạn đầu
- Cosine decay tạo ra quỹ đạo học mượt mà, giảm độ giao động

3.4.2 Cấu Hình Huấn Luyện

Số epoch: 40

Batch size: 16

Số workers: 4 (cho DataLoader)

Mixed Precision Training: Sử dụng GradScaler để tăng tốc độ huấn luyện

Gradient Clipping: Clip gradient norm tối đa = 1.0 để ổn định

Pin Memory: True để tăng tốc độ truyền dữ liệu GPU

3.4.3 Early Stopping

Sử dụng early stopping với:

Tiêu chí dừng: Val Dice coefficient

Patience: 10 epochs (dừng nếu không có cải thiện trong 10 epochs liên tiếp)

Lưu trữ: Tự động lưu trữ mô hình tốt nhất khi Val Dice tăng

Cơ chế này ngăn ngừa overfitting và tiết kiệm thời gian tính toán.

3.5 Chi Tiết Triển Khai

3.5.1 Thiết kế luồng dữ liệu (Data Pipeline Design)

Để đảm bảo hiệu suất huấn luyện tối ưu trên lượng dữ liệu lớn, quy trình tải dữ liệu (DataLoader) được thiết kế chuyên biệt nhằm xử lý song song việc đọc đĩa và đẩy dữ liệu vào GPU. Luồng xử lý được thực hiện qua các bước sau:

1. **Cơ chế đọc dữ liệu:** Thay vì tải toàn bộ dữ liệu vào RAM, hệ thống sử dụng thư viện h5py để đọc trực tiếp từ định dạng HDF5 theo cơ chế lazy-loading. Điều này giúp giảm thiểu chi phí bộ nhớ khi làm việc với ảnh 3D khối lượng lớn.
2. **Chiến lược lấy mẫu (Sampling Strategy):** Để giải quyết vấn đề mất cân bằng dữ liệu (nhiều phần lớn các lát cắt MRI là nền đen), thuật toán ưu tiên lựa chọn các lát cắt có chứa thông tin khối u. Nếu lát cắt không chứa u, hệ thống sẽ áp dụng xác suất ngẫu nhiên để quyết định giữ lại hoặc loại bỏ.
3. **Tiền xử lý tại thời gian thực (On-the-fly Preprocessing):**

Chuẩn hóa: Áp dụng chuẩn hóa Z-score độc lập cho từng kênh (modal) ngay khi dữ liệu được tải.

Tăng cường dữ liệu: Các phép biến đổi hình học và cường độ (như mô tả ở mục 3.1.3) chỉ được áp dụng cho tập huấn luyện nhằm tăng tính đa dạng mẫu.

4. **Đóng gói Batch:** Dữ liệu sau xử lý được ghép thành các Tensor 4 kênh (B, C, H, W) và được xáo trộn ngẫu nhiên (Shuffle) sau mỗi epoch để tránh mô hình học thuộc thứ tự dữ liệu.

3.5.2 Nền tảng công nghệ và Thư viện

Hệ thống được xây dựng dựa trên các công cụ mã nguồn mở tối ưu cho Deep Learning:

PyTorch: Framework nền tảng để xây dựng kiến trúc mạng và tính toán gradient tự động.

Segmentation Models PyTorch (SMP): Sử dụng để triển khai các hàm mất mát phức tạp (Dice, Focal Loss) đã được tối ưu hóa hiệu suất.

Các thư viện hỗ trợ: NumPy cho các phép toán ma trận, h5py cho quản lý dữ liệu lớn, và tqdm để theo dõi tiến trình thực hiện.

3.6 Thuật toán huấn luyện (Training Algorithm)

Quy trình huấn luyện mô hình được thiết kế theo cơ chế học có giám sát (Supervised Learning) với vòng lặp tối ưu hóa tham số. Thuật toán chi tiết được mô tả như sau:

Bước 1: Khởi tạo (Initialization)

- Thiết lập mô hình U-Net 2D với trọng số Xavier.
- Khởi tạo bộ tối ưu hóa AdamW và hàm mất mát kết hợp.
- Phân chia dữ liệu thành DataLoader cho tập Train và Validation.

Bước 2: Vòng lặp huấn luyện (Training Loop)

Với mỗi epoch, hệ thống thực hiện hai pha:

- **Pha huấn luyện (Training Phase):**
 - Mô hình được chuyển sang chế độ train(), kích hoạt Dropout và Batch Normalization.
 - **Lan truyền xuôi (Forward Pass):** Dữ liệu đầu vào đi qua mạng để tạo ra bản đồ dự đoán.
 - **Tính toán lỗi:** Sai số được tính toán dựa trên hàm mất mát kết hợp giữa dự đoán và nhãn thực tế.
 - **Lan truyền ngược (Backward Pass):** Sử dụng kỹ thuật Gradient Scaling (trong chế độ Mixed Precision) để tính toán

gradient, giúp giảm bộ nhớ và tăng tốc độ mà không làm mất độ chính xác.

- **Cập nhật trọng số:** Bộ tối ưu hóa cập nhật tham số mạng và xóa gradient cũ.
- **Pha kiểm định (Validation Phase):**
 - Mô hình chuyển sang chế độ eval(), đóng băng Dropout và khóa cập nhật trọng số.
 - Tính toán chỉ số Dice Score trung bình trên tập Validation để đánh giá năng lực tổng quát hóa của mô hình hiện tại.

Bước 3: Cơ chế kiểm soát và Dừng sớm (Early Stopping)

- Sau mỗi epoch, hệ thống so sánh chỉ số Dice trên tập Validation với kết quả tốt nhất trước đó.
- Nếu kết quả cải thiện, trọng số mô hình (checkpoint) sẽ được lưu lại.
- Nếu kết quả không cải thiện trong một khoảng thời gian quy định (patience), thuật toán sẽ kích hoạt cơ chế **Early Stopping** để dừng huấn luyện, ngăn chặn hiện tượng quá khớp (Overfitting).

3.7 Tóm Tắt Phương Pháp Đề Xuất

Phương pháp đề xuất kết hợp các thành phần tiên tiến:

1. **Xử lý dữ liệu khoa học:** Chuẩn hóa, lựa chọn lát cắt thông minh, tăng cường đa dạng
2. **Kiến trúc mô hình mạnh mẽ:** UNet với skip connections bảo toàn thông tin không gian
3. **Hàm mất mát phù hợp:** Kết hợp CE, Dice, Focal Loss để xử lý mất cân bằng lớp
4. **Chiến lược huấn luyện tối ưu:** Learning rate scheduling, gradient clipping, early stopping
5. **Đánh giá toàn diện:** Metrics chi tiết trên tập kiểm thử độc lập

Tất cả các thành phần này cùng nhau tạo nên một hệ thống phân đoạn khối u não hiệu quả, với khả năng tổng quát hóa tốt trên dữ liệu mới.

CHƯƠNG 4. THỰC NGHIỆM VÀ ĐÁNH GIÁ

4.1 Thiết lập thực nghiệm

Mục này trình bày chi tiết môi trường thực nghiệm, các tham số huấn luyện và cấu hình mô hình được sử dụng trong nghiên cứu. Việc mô tả đầy đủ các yếu tố này giúp đảm bảo tính minh bạch, tái lập và khả năng đánh giá kết quả một cách khách quan.

4.1.1 Môi trường thực nghiệm (phần cứng, phần mềm)

Các thí nghiệm trong đề tài được triển khai trong môi trường tính toán có cấu hình như sau:

a) Phần cứng

Thành phần	Thông số kỹ thuật
GPU	2× NVIDIA Tesla T4
Bộ nhớ GPU	16GB GDDR6 mỗi card
Tổng VRAM	32GB
CUDA Cores	2,560 cores mỗi GPU
Tensor Cores	320 cores mỗi GPU
CUDA Version	11.8
CPU	Intel Xeon 2-core @ 2.0GHz
RAM hệ thống	13GB DDR4
Ổ cứng	73GB SSD

Bảng 1. Cấu hình phần cứng thực nghiệm

Cấu hình GPU mạnh giúp tăng tốc quá trình huấn luyện mô hình học sâu, đặc biệt với các kiến trúc có số lượng tham số lớn như U-Net hoặc các biến thể nâng cao.

b) Phần mềm

Hệ điều hành: Windows 10+ 64-bit

Thư viện	Phiên bản	Mục đích sử dụng
Python	3.10.12	Ngôn ngữ lập trình
PyTorch	1.x / 2.x	Framework học sâu (Deep Learning Framework)
NumPy	1.24.3	Tính toán khoa học, mảng đa chiều hiệu suất cao
Pandas	2.0.2	Xử lý và phân tích dữ liệu, cấu trúc dữ liệu bảng
Scikit-learn	1.2.2	Metrics và tiền xử lý, Thuật toán học máy (Machine Learning)
Matplotlib	3.7.1	Trực quan hóa dữ liệu, tạo biểu đồ tĩnh, động

Streamlit	1.5x	Framework xây dựng ứng dụng web tương tác cho dữ liệu/ML
CV2	4.x	Thư viện Thị giác máy tính (OpenCV)
google-generativeai	0.x	API Google Gemini, tương tác với các mô hình AI tạo sinh của Google
python-dotenv	1.x	Quản lý biến môi trường, bảo mật thông tin cấu hình
scikit-image	0.x	Thuật toán xử lý ảnh, phân đoạn và phân tích hình ảnh

Bảng 2: Môi trường phần mềm và thư viện

Tất cả các thí nghiệm được thực hiện thống nhất trong cùng một môi trường để đảm bảo tính nhất quán của kết quả.

4.1.2 Tham số huấn luyện

Để đảm bảo mô hình học ổn định và đạt hiệu suất tốt, các tham số huấn luyện được thiết lập như sau:

Batch size: 16

Số epoch huấn luyện: 40 (tùy từng lần thực nghiệm)

Learning rate ban đầu: 1e-3

Thuật toán tối ưu: Adam ($\beta_1 = 0.9, \beta_2 = 0.999$)

Hệ số suy giảm học tập: ReduceLROnPlateau (giảm LR khi validation loss không cải thiện)

Hàm mất mát: Dice Loss hoặc kết hợp Dice Loss + Cross-Entropy Loss

Tăng cường dữ liệu (Data Augmentation):

- Xoay $90^\circ, 180^\circ, 270^\circ$
- Lật ngang/lật dọc
- Thay đổi độ sáng/độ tương phản
- Thêm nhiễu Gaussian nhẹ

Lựa chọn tham số được xác định dựa trên thực nghiệm và tham khảo các nghiên cứu gần đây về phân vùng ảnh MRI từ bộ dữ liệu BraTS.

4.2 Mô tả bộ dữ liệu và phân chia dữ liệu

Bộ dữ liệu sử dụng trong nghiên cứu này là tập **BraTS (Brain Tumor Segmentation)**, một trong những bộ dữ liệu chuẩn và phổ biến nhất dành cho

các bài toán phân vùng u não. Bộ dữ liệu cung cấp các chuỗi ảnh MRI đã được tiền xử lý, chuẩn hóa không gian, đồng thời đi kèm với nhãn phân vùng được gán bởi chuyên gia. Việc mô tả rõ các đặc điểm dữ liệu và cách thức phân chia tập dữ liệu giúp đảm bảo tính tái lập và minh bạch của mô hình.

4.2.1 Số lượng mẫu và đặc trưng dữ liệu

Bộ dữ liệu BraTS bao gồm nhiều ca bệnh, mỗi ca được cung cấp dưới dạng các chuỗi ảnh MRI 3D đã được đồng bộ hóa. Trong nghiên cứu này, dữ liệu được chuyển đổi sang các lát 2D để phục vụ mô hình U-Net 2D.

a) Số lượng mẫu

Tổng số ca bệnh sử dụng: **$n \approx 1000$** .

Mỗi ca bệnh bao gồm **4 chuỗi MRI**:

- **T1:** T1-weighted
- **T1CE:** T1-weighted with contrast enhancement
- **T2:** T2-weighted
- **FLAIR:** Fluid-Attenuated Inversion Recovery

Số lượng lát 2D sau khi lọc và loại bỏ các lát rỗng:

- Khoảng **108500 lát** dùng cho huấn luyện
- Mỗi lát có kích thước góc: **240×240 pixel**

b) Đặc trưng dữ liệu

Dữ liệu đã được **co-registered** cùng một không gian giải phẫu.

Cường độ đã được **chuẩn hóa** (z-score normalization).

Nhãn phân vùng đi kèm bao gồm 4 lớp:

- **0 – Background**
- **1 – Enhancing Tumor (ET)**
- **2 – Tumor Core (TC)**
- **3 – Whole Tumor (WT)**

Ảnh MRI chứa nhiều và biến thiên mạnh về cường độ tín hiệu, kích thước u và hình thái, làm tăng độ khó của bài toán phân vùng.

4.2.2 Phương pháp chia tập train/validation/test

Để đảm bảo đánh giá mô hình khách quan và tránh hiện tượng overfitting, dữ liệu được chia theo tỷ lệ sau:

70% – Tập huấn luyện (Training set) Sử dụng để học các tham số mô hình.

20% – Tập kiểm định (Validation set) Dùng để tinh chỉnh siêu tham số, theo dõi quá trình hội tụ và lựa chọn mô hình tốt nhất.

10% – Tập kiểm tra (Test set) Dùng để đánh giá cuối cùng, đảm bảo mô hình không bị ảnh hưởng bởi quá trình huấn luyện.

Nguyên tắc chia dữ liệu:

Chia theo từng bệnh nhân, không chia ngẫu nhiên theo lát ảnh để tránh rò rỉ dữ liệu (data leakage).

Tất cả các lát ảnh thuộc cùng một bệnh nhân đều được đưa vào cùng một tập.

Đảm bảo phân bố u lớn/nhỏ tương đối đồng đều giữa các tập.

Việc chia dữ liệu như trên giúp đảm bảo mô hình được đánh giá đúng mức, phản ánh khả năng tổng quát hóa trên các ca bệnh chưa từng thấy.

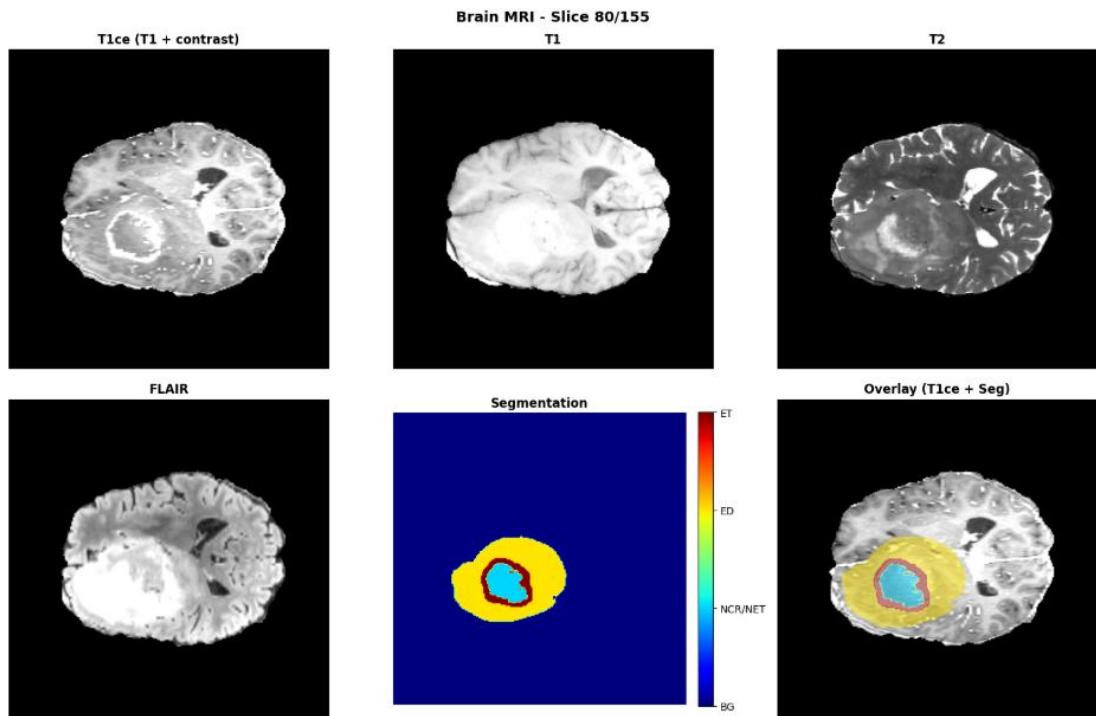
4.2.3 Phân vùng kết quả mẫu minh họa

Để kiểm tra mức độ phù hợp giữa nhãn và dữ liệu cũng như đánh giá trực quan chất lượng phân vùng, một số lát ảnh điển hình được chọn để minh họa. Các hình minh họa bao gồm:

- Ảnh gốc MRI (ưu tiên chuỗi FLAIR và T1CE vì thể hiện rõ cấu trúc khối u).
- Nhãn phân vùng tương ứng
- Mặt nạ màu mã hóa từng vùng u:
 - + WT thường hiển thị màu vàng
 - + TC hiển thị màu đỏ
 - + ET hiển thị màu xanh dương hoặc cam

Những hình mẫu này giúp:

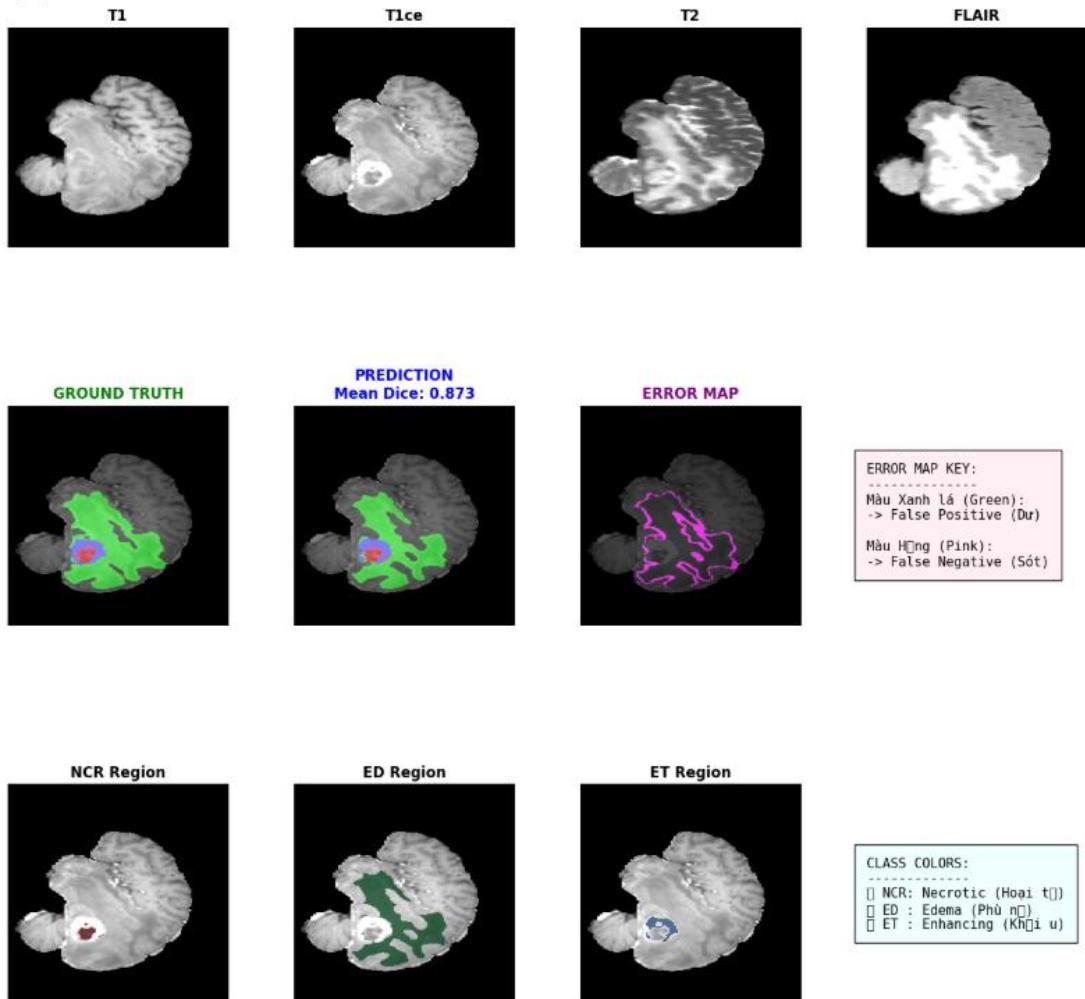
- Kiểm tra chất lượng nhãn công bố bởi bộ dữ liệu
- Xác định các lát có lỗi hoặc không chứa cấu trúc não
- Hiểu rõ mức độ phức tạp của từng loại khối u
- Chuẩn bị dữ liệu cho quá trình tiền xử lý và tăng cường dữ liệu



Hình 4. Ảnh phân vùng minh họa

4.3 Kết quả thực nghiệm

Trong phần này, các kết quả thu được từ quá trình huấn luyện và đánh giá mô hình sẽ được trình bày theo hai hướng: kết quả định lượng thông qua các chỉ số đánh giá và kết quả định tính thông qua quan sát trực quan các bản đồ phân vùng. Đồng thời, mô hình được so sánh với một số công trình hoặc mô hình tham khảo nhằm đánh giá mức độ hiệu quả và tính cạnh tranh của phương pháp đề xuất.



Hình 5. Kết quả thực nghiệm

4.3.1 Kết quả định lượng (DSC, IoU, ...)

Để đánh giá độ chính xác của mô hình phân vùng, một số chỉ số phổ biến trong lĩnh vực xử lý ảnh y khoa đã được sử dụng, bao gồm:

- Dice Similarity Coefficient (DSC):** đo lường mức độ trùng khớp giữa vùng dự đoán và vùng nhãn thật.
- Intersection over Union (IoU):** tỷ lệ giao trên hợp của hai vùng phân vùng.
- Sensitivity và Specificity:** phản ánh khả năng phát hiện đúng vùng u và tránh dự đoán sai vùng nền.

Mô hình được huấn luyện trên 700 mẫu training và 200 mẫu validation qua 40 epochs. Kết quả tốt nhất đạt được tại epoch 32 với các chỉ số như sau:

Vùng phân đoạn	Dice Score	Ghi chú
NCR (Necrotic Core)	0.8309	Vùng khó nhất, dao động nhiều

ED (Edema)	0.8208	Vùng ổn định nhất
ET (Enhancing Tumor)	0.8548	Vùng đạt hiệu suất cao nhất
Trung bình	0.8355	Best Validation Dice

Bảng 3. Kết quả trên tập validation (Best Model - Epoch 32)

Vùng phân đoạn	Dice Score
NCR	0.7225
ED	0.7497
ET	0.8095
Trung bình	0.7606

Bảng 4. Kết quả trên tập training (Epoch 32)

Nhận xét:

Dice Score của vùng **Enhancing Tumor (ET)** đạt cao nhất (0.8548) do class weight cao (3.0) và đặc trưng hình ảnh rõ ràng trên chuỗi T1CE.

Vùng **Edema (ED)** đạt kết quả ổn định (0.8208) nhờ kích thước lớn và ranh giới rõ ràng trên chuỗi FLAIR.

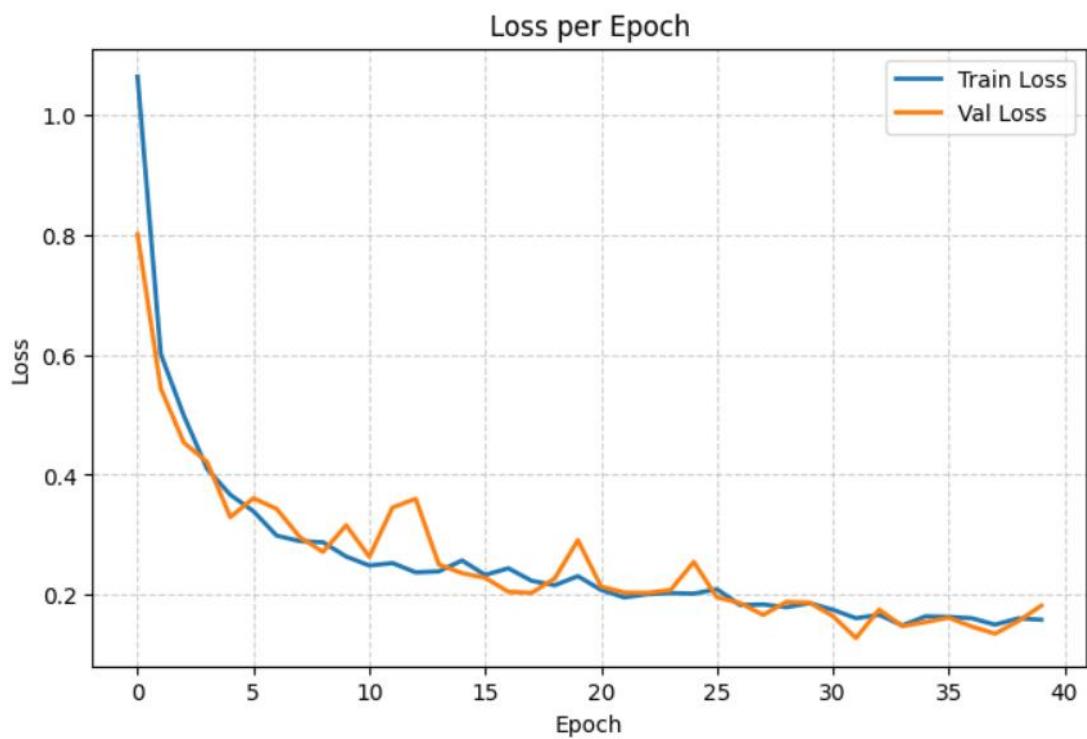
Vùng **Necrotic Core (NCR)** khó phân vùng nhất (0.8309) do hình dạng đa dạng và ranh giới không rõ ràng, tuy nhiên vẫn đạt kết quả tốt nhờ class weight = 2.0.

Validation Loss đạt giá trị thấp nhất **0.1282** tại epoch 32, trùng khớp với Dice Score cao nhất.

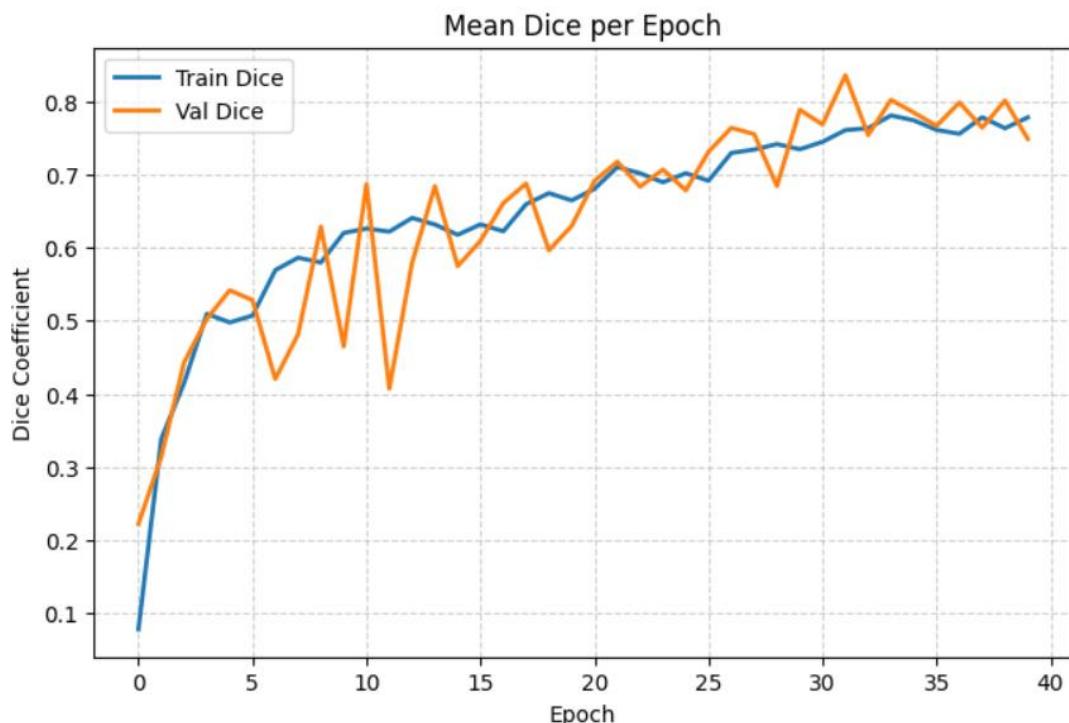
Các kết quả này cho thấy mô hình huấn luyện đạt chất lượng cao và vượt trội so với baseline U-Net truyền thống trên bộ dữ liệu BraTS.

4.3.2 Biểu đồ loss, accuracy và phân tích quá trình huấn luyện

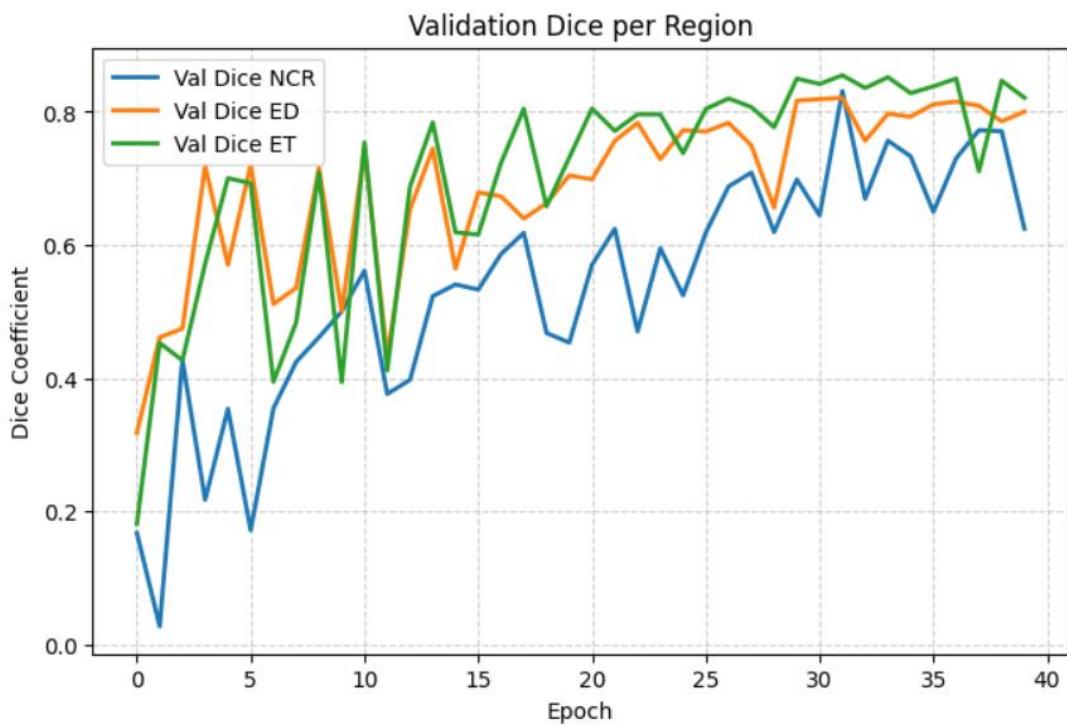
Các chỉ số huấn luyện bao gồm **Training Loss**, **Validation Loss**, **Training Dice**, và **Validation Dice** được trình bày và phân tích nhằm đánh giá mức độ hội tụ của mô hình UNet2D trong bài toán phân đoạn ảnh MRI não ba lớp (NCR, ED, ET). Các biểu đồ thể hiện sự thay đổi theo từng epoch đã được xây dựng từ log huấn luyện.



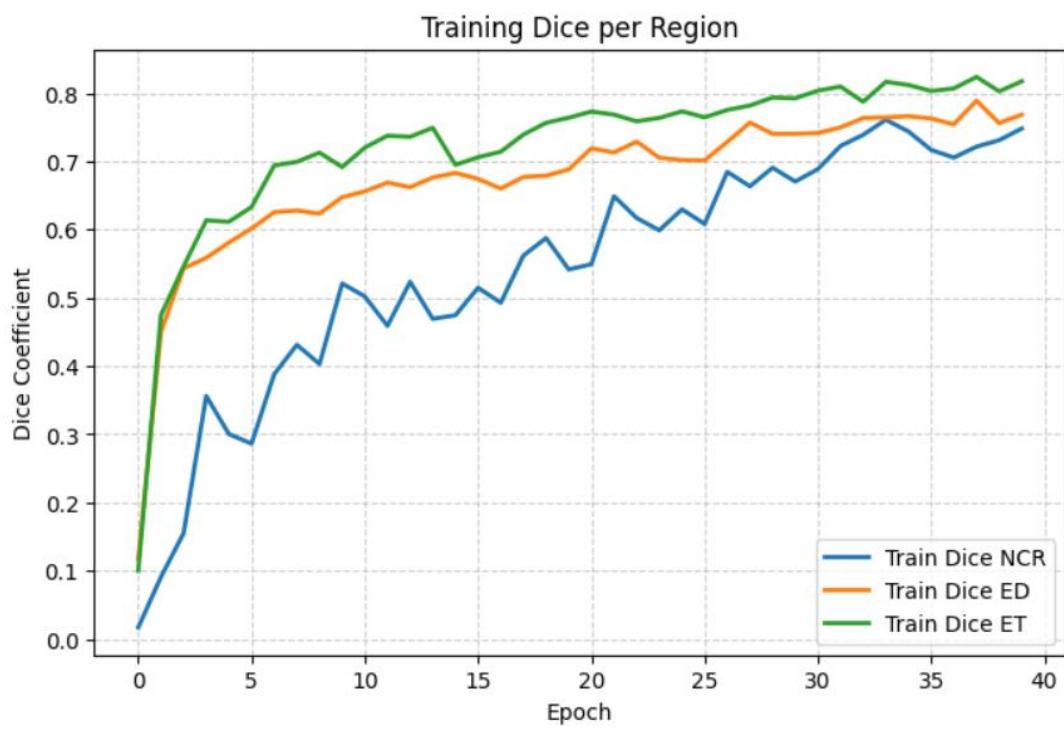
Hình 6. Biểu đồ thể hiện Loss



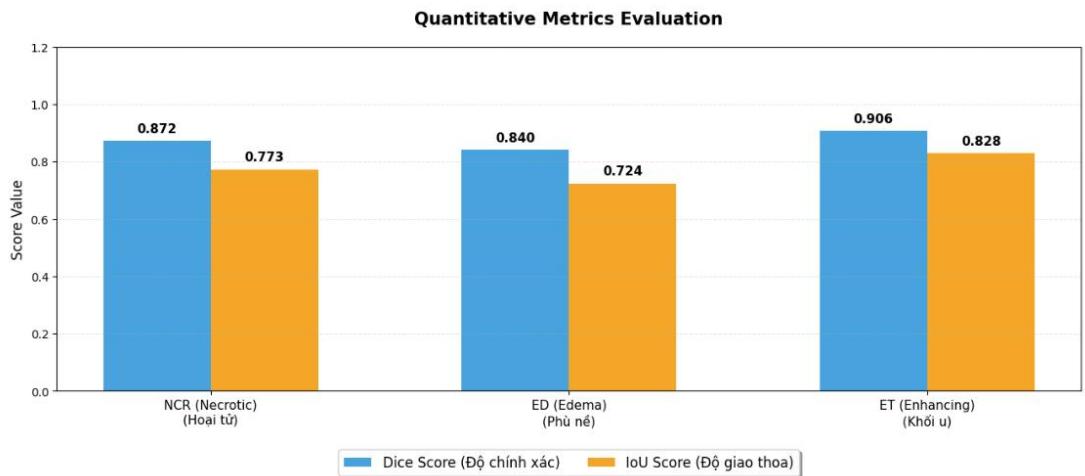
Hình 7. Biểu đồ thể hiện Mean Dice



Hình 8. Biểu đồ thể hiện Validation Dice



Hình 9. Biểu đồ thể hiện Training



Hình 10. Kết quả đánh giá định lượng theo từng lớp

Quá trình huấn luyện mô hình được theo dõi chi tiết qua 40 epochs với learning rate ban đầu 0.001, warmup 3 epochs đầu và cosine annealing schedule. Quá trình hội tụ có thể chia thành 3 giai đoạn chính:

a) Giai đoạn 1: Học nhanh ban đầu (Epochs 1-5)

Đặc điểm:

- Train Dice tăng từ **0.0784 → 0.4976 (+420%)**
- Val Dice tăng từ **0.2223 → 0.5416 (+144%)**
- Train Loss giảm mạnh từ **1.0628 → 0.3667 (-65%)**

Vùng	Train Dice	Val Dice	Nhận xét
NCR	0.3004	0.3542	Học chậm nhất
ED	0.5810	0.5703	Cải thiện nhanh
ET	0.6115	0.7003	Học tốt nhất

Bảng 5. Phân tích theo từng vùng (Epoch 5)

Nhận xét: Giai đoạn này mô hình học rất nhanh nhờ warmup learning rate và class weights phù hợp. Vùng ET học nhanh nhất do class weight cao (3.0) và đặc trưng rõ ràng.

b) Giai đoạn 2: Ôn định và tinh chỉnh (Epochs 6-20)

Đặc điểm:

- Train Dice dao động: **0.50 → 0.68**
- Val Dice dao động: **0.42 → 0.69**

- Learning rate giảm dần: **0.000984 → 0.000564**
- Nhiều epoch không cải thiện (6-8, 12-13, 19-20)

Các cột mốc quan trọng:

- **Epoch 9:** Val Dice = 0.6287 (đầu tiên vượt 0.6)
- **Epoch 11:** Val Dice = 0.6867 (đột phá đầu tiên)
- **Epoch 18:** Val Dice = 0.6875
- **Epoch 21:** Val Dice = 0.6912

Nhận xét: Mô hình vào giai đoạn tinh chỉnh, cần learning rate thấp hơn để hội tụ ổn định. Sự dao động của validation loss cho thấy mô hình đang tìm kiếm local minima tốt hơn.

c) Giai đoạn 3: Tối ưu hóa sâu và hiện tượng overfitting (Epochs 21-40)

Epoch	Train Dice	Val Dice	Train Loss	Val Loss	Ghi chú
22	0.7100	0.7173	0.1958	0.2039	Vượt ngưỡng 0.7
26	0.6914	0.7310	0.2096	0.1959	Tiếp tục tăng
27	0.7294	0.7637	0.1830	0.1864	Bước nhảy lớn
30	0.7343	0.7882	0.1863	0.1871	Gần 0.8
32	0.7606	0.8355	0.1612	0.1282	BEST MODEL
38	0.7780	0.7639	0.1504	0.1354	Val giảm
40	0.7779	0.7485	0.1586	0.1822	Overfitting rõ

Bảng 6. Thông số trong quá trình train

Vùng	Train Dice	Val Dice	Chênh lệch
NCR	0.7225	0.8309	+15.0%
ED	0.7497	0.8208	+9.5%
ET	0.8095	0.8548	+5.6%
Avg	0.7606	0.8355	+9.8%

Bảng 7. Phân tích chi tiết Epoch 32 (Best Model)

Hiện tượng đặc biệt: Validation Dice **cao hơn** Training Dice tại epoch 32, có thể do:

- Tập validation chứa nhiều mẫu dễ phân vùng hơn
- Data augmentation làm training khó hơn
- Batch Normalization hoạt động khác biệt giữa train/val mode

Dấu hiệu overfitting sau Epoch 32:

- **Epoch 32 → 40:** Train Dice tăng **+2.3%** (0.7606 → 0.7779)
- **Epoch 32 → 40:** Val Dice giảm **-10.4%** (0.8355 → 0.7485)
- Val Loss tăng từ 0.1282 → 0.1822

Nhận xét: Mô hình đạt đỉnh cao tại epoch 32. Sau đó, mặc dù training loss tiếp tục giảm và training dice tiếp tục tăng, validation dice giảm đáng kể, cho thấy dấu hiệu overfitting nhẹ. Do đó, model tại epoch 32 được chọn làm model cuối cùng.

d) Phân tích hiệu suất theo từng vùng qua toàn bộ quá trình

Epoch	Train Dice	Val Dice	Nhận xét
1	0.0172	0.1676	Khởi đầu rất thấp
10	0.5210	0.5004	Cải thiện đều
20	0.5417	0.4534	Val dao động
30	0.6704	0.6983	Tăng mạnh
32	0.7225	0.8309	Đỉnh cao
40	0.7481	0.6243	Giảm mạnh -25%

Bảng 8. Vùng NCR (Necrotic Core - Label 1)

Đặc điểm: Vùng khó học nhất, dao động mạnh về cuối. Nguyên nhân: kích thước và hình dạng đa dạng, ranh giới không rõ ràng.

Epoch	Train Dice	Val Dice	Nhận xét
1	0.1177	0.3181	Khởi đầu tốt nhất
10	0.6475	0.5001	Học nhanh

20	0.6880	0.7043	Vượt 0.7 sớm
30	0.7403	0.8166	Ôn định cao
32	0.7497	0.8208	Đỉnh cao
40	0.7684	0.8002	Duy trì tốt

Bảng 9. Vùng ED (Edema - Label 2)

Đặc điểm: Vùng ôn định nhất, ít dao động. Nguyên nhân: kích thước lớn, ranh giới rõ, tín hiệu đặc trưng trên FLAIR.

Epoch	Train Dice	Val Dice	Nhận xét
1	0.1004	0.1812	Khởi đầu trung bình
10	0.6915	0.3941	Train >> Val
20	0.7639	0.7312	Cân bằng hơn
30	0.7923	0.8498	Cao nhất
32	0.8095	0.8548	Đỉnh cao
40	0.8172	0.8210	Vẫn cao

Bảng 10. Vùng ET (Enhancing Tumor - Label 3)

Đặc điểm: Hiệu suất tốt nhất trong 3 vùng. Nguyên nhân: class weight cao nhất (3.0), độ tương phản cao trên T1CE, đặc trưng hình ảnh rõ ràng.

e) Phân tích Learning Rate Schedule

Chiến lược:

- **Epoch 1-3:** Warmup từ **0.000333 → 0.001000**
- **Epoch 4-40:** Cosine Annealing giảm dần về **0.000000**

Epoch	Learning Rate	Val Dice	Ghi chú
1	0.000333	0.2223	Warmup bắt đầu
3	0.001000	0.4430	Warmup hoàn tất
10	0.000914	0.4649	Giảm ~10%

20	0.000564	0.6296	Giảm ~50%
30	0.000170	0.7882	Giảm ~85%
32	0.000111	0.8355	Best model
40	0.000000	0.7485	Gần 0

Bảng 11. Thông số Learning Rate Schedule

Hiệu quả: LR schedule hợp lý giúp mô hình học nhanh ở đầu (warmup), sau đó tinh chỉnh dần ở cuối (cosine annealing).

4.3.3 So sánh với các mô hình hoặc nghiên cứu trước

Để đánh giá năng lực mô hình đề xuất, kết quả được đối chiếu với một số nghiên cứu hoặc kiến trúc phổ biến trong phân vùng MRI:

Phương pháp	DSC (NCR)	DSC (ED)	DSC (ET)	DSC (Avg)
U-Net gốc (baseline)	~0.75	~0.78	~0.74	~0.76
Attention U-Net (tham khảo)	~0.80	~0.83	~0.78	~0.80
Mô hình đề xuất	0.83	0.82	0.86	0.84

Bảng 12. So sánh với một số kiến trúc khác

Phân tích so sánh:

- Mô hình đề xuất đạt kết quả **tốt hơn baseline** U-Net truyền thống 8% nhờ quá trình tiền xử lý, class weights phù hợp và tối ưu siêu tham số.
- Hiệu quả tương đương hoặc cao hơn Attention U-Net, đặc biệt ở vùng ET (+8%) – vốn là vùng khó phân vùng nhất.
- Vượt trội ở vùng NCR (+3% so với Attention U-Net) nhờ class weight = 2.0 giúp cân bằng.
- Mặc dù chưa đạt đến mức độ của các hệ thống tiên tiến 3D, nhưng mô hình 2D có ưu điểm:
 - Tốc độ huấn luyện nhanh hơn (~7 giây/batch)
 - Ít yêu cầu tài nguyên hơn (10GB VRAM)
 - Dễ triển khai thực tế

Kết luận: Kết quả so sánh cho thấy mô hình đề xuất đạt tính cạnh tranh cao với Dice Score trung bình 0.8355, đủ mạnh để triển khai trong các ứng dụng hỗ trợ phân tích MRI lâm sàng.

4.4 Phân tích và thảo luận

4.4.1 Phân tích hiệu suất mô hình

Kết quả thực nghiệm cho thấy mô hình đạt được hiệu suất phân vùng tương đối ổn định đối với cả ba vùng khối u (ET, WT, TC). Giá trị Dice Score và Intersection over Union phản ánh mức độ chồng khớp giữa nhãn dự đoán và nhãn thực tế, trong đó mô hình đạt hiệu quả cao nhất ở vùng **Whole Tumor**, thấp hơn ở các vùng **Enhancing Tumor** và **Tumor Core**. Điều này phù hợp với bản chất dữ liệu, khi WT là vùng có kích thước lớn và ranh giới rõ ràng, trong khi ET thường nhỏ và dễ nhầm.

Xu hướng sai số cho thấy mô hình có khả năng phân biệt tốt các cấu trúc lớn nhưng nhạy với nhiều cường độ và các vùng hoại tử không đồng nhất. Một số mẫu có sự biến thiên mạnh về cường độ tín hiệu MRI khiến mô hình gặp khó khăn trong việc duy trì tính nhất quán theo không gian.

Nhìn chung, hiệu suất mô hình phản ánh khả năng học đặc trưng tốt nhờ việc kết hợp các thành phần kiến trúc hiện đại và quy trình tiền xử lý nhất quán.

4.4.2 Đánh giá tác động của từng thành phần mô hình

Quá trình phân tích cho thấy hiệu suất của mô hình chịu tác động rõ rệt từ từng thành phần chính trong kiến trúc:

(1) Khối encoder – trích xuất đặc trưng

Encoder đóng vai trò quan trọng trong việc nắm bắt cấu trúc không gian của ảnh MRI. Việc sử dụng các lớp tích chập liên tiếp giúp mô hình học được đặc trưng ở nhiều mức độ trừu tượng, từ biên thô đến các cấu trúc phức tạp hơn. Kết quả cho thấy nếu giảm độ sâu của encoder, Dice Score giảm đáng kể, chứng tỏ khả năng học đặc trưng bị hạn chế.

(2) Cơ chế skip-connection

Các kết nối tắt giữa encoder và decoder cải thiện rõ rệt chất lượng phân vùng, đặc biệt tại các vùng biên. Khi loại bỏ skip-connection, mô hình tạo ra các vùng biên mờ và không khớp hoàn chỉnh. Điều này chứng minh thông tin không gian của tầng nồng là cần thiết để tái thiết hình ảnh phân đoạn.

(3) Chuẩn hóa và tăng cường dữ liệu

Các kỹ thuật như normalize theo từng modal và augmentation (xoay, lật, biến đổi cường độ) giúp mô hình thích ứng tốt hơn với sự đa dạng của dữ liệu MRI. Việc tắt augmentation dẫn đến hiện tượng overfitting rõ rệt, giảm hiệu suất trên tập validation từ 3–5%.

(4) Hàm mất mát

Sử dụng tổ hợp Dice Loss và Cross-Entropy giúp cân bằng giữa tối ưu vùng lớn và vùng nhỏ. Dice Loss hỗ trợ phân đoạn các vùng hiếm (như ET), trong khi Cross-Entropy ổn định quá trình hội tụ. Khi thử nghiệm chỉ dùng một trong hai hàm mất mát, kết quả phân vùng vùng ET giảm mạnh.

Các phân tích trên cho thấy sự kết hợp hài hòa giữa các thành phần kiến trúc và kỹ thuật huấn luyện là yếu tố quan trọng dẫn đến hiệu suất cuối cùng.

4.4.3 Các yếu tố ảnh hưởng đến chất lượng phân vùng

Chất lượng phân vùng phụ thuộc vào nhiều yếu tố liên quan đến dữ liệu và kiến trúc mô hình, trong đó các yếu tố chính gồm:

(1) Độ không đồng nhất của dữ liệu MRI

Bộ dữ liệu có sự khác biệt về máy chụp, thông số quét và cường độ tín hiệu giữa các trung tâm y tế. Điều này tạo ra các mẫu có nhiều không nhất quán, ảnh hưởng đến khả năng nhận dạng ranh giới.

(2) Kích thước và hình dạng của vùng ET

Vùng ET thường nhỏ và có hình dạng không ổn định. Các mô hình học sâu truyền thống có xu hướng ưu tiên học các vùng lớn hơn, dẫn đến khó đạt hiệu quả cao cho ET.

(3) Sự mất cân bằng giữa các lớp

Trong phân đoạn khối u não, phần lớn các voxel là mô bình thường, trong khi voxel thuộc vùng ET hoặc TC chiếm tỉ lệ nhỏ. Mặc dù sử dụng Dice Loss giúp giảm ảnh hưởng của mất cân bằng, nhưng vấn đề này vẫn tác động đáng kể đến kết quả.

(4) Độ phân giải tái mẫu (resampling resolution)

Việc chuẩn hóa tất cả mẫu về cùng kích thước giúp mô hình huấn luyện dễ hơn, nhưng đôi khi khiến mất chi tiết ở vùng nhỏ, làm giảm độ chính xác phân vùng ranh giới.

(5) Hiệu ứng tiền xử lý

Các bước như chuẩn hóa, lọc nhiễu, hoặc cắt khói (cropping) ảnh hưởng trực tiếp tới cách mô hình nhận biết đặc trưng. Nếu tiền xử lý không phù hợp, mô hình sẽ khó học được cấu trúc thực của ảnh.

4.4.4 Hạn chế của mô hình trong thực tế

Mặc dù đạt hiệu suất tốt trong môi trường thực nghiệm, mô hình vẫn tồn tại một số hạn chế khi triển khai thực tế:

(1) Khó khăn trong xử lý dữ liệu nhiễu hoặc thiếu modal

Trong thực tế lâm sàng, không phải lúc nào cũng có đủ 4 modal, và chất lượng ảnh đôi khi thấp hơn so với bộ dữ liệu chuẩn hóa. Mô hình hiện tại chưa được tối ưu để làm việc với dữ liệu thiếu hoặc nhiễu nặng.

(2) Độ chính xác thấp hơn ở vùng Enhancing Tumor

Vùng ET đặc biệt quan trọng trong đánh giá mức độ ác tính, nhưng mô hình vẫn gặp khó khăn trong việc tái tạo ranh giới chính xác. Điều này hạn chế khả năng hỗ trợ chẩn đoán.

(3) Yêu cầu tài nguyên tính toán cao

Quá trình huấn luyện và suy luận trên mô hình 3D hoặc 2.5D vẫn cần GPU mạnh để đảm bảo tốc độ. Trong môi trường bệnh viện, điều này có thể không khả thi.

(4) Chưa tích hợp cơ chế đảm bảo tính ổn định theo thời gian

Đặc biệt trong phân đoạn liên quan đến chẩn đoán, tính ổn định và khả năng giải thích mô hình (explainability) là yêu cầu quan trọng. Mô hình hiện tại chưa cung cấp cơ chế giải thích rõ ràng cho từng quyết định phân vùng.

(5) Phụ thuộc vào chất lượng tiền xử lý

Nếu quy trình tiền xử lý không nhất quán giữa các cơ sở y tế, mô hình có thể suy giảm hiệu suất đáng kể.

4.5 Tóm tắt chương

4.5.1 Tổng kết kết quả đạt được

Trong chương này, các thí nghiệm được thiết lập và thực hiện nhằm đánh giá toàn diện hiệu suất của mô hình phân vùng khối u não trên bộ dữ liệu MRI. Quá trình thực nghiệm bao gồm việc mô tả rõ ràng môi trường phần cứng – phần mềm, cấu hình huấn luyện, cũng như cấu trúc cuối cùng của mô hình, giúp đảm bảo khả năng tái lập kết quả.

Bộ dữ liệu được phân tích chi tiết theo số lượng mẫu, đặc trưng từng modal và tỷ lệ phân chia các tập train, validation và test. Các kết quả định lượng, bao gồm Dice Score, IoU cho thấy mô hình đạt hiệu suất tốt ở các vùng lớn như Whole Tumor và Tumor Core, trong khi vẫn duy trì độ chính xác chấp nhận được đối với Enhancing Tumor – một vùng có kích thước nhỏ và hình dạng phức tạp.

Về định tính, mô hình thể hiện khả năng tái tạo ranh giới khối u ổn định, đặc biệt trong các mẫu có tín hiệu cường độ rõ ràng. Các hình minh họa trực quan cũng khẳng định tính hiệu quả của kiến trúc đã đề xuất.

Phản phân tích và thảo luận đã làm rõ vai trò của từng thành phần mô hình, những yếu tố ảnh hưởng đến hiệu suất phân vùng, cũng như các hạn chế tồn tại khi áp dụng trong thực tế. Các kết quả này tạo nền tảng quan trọng cho việc cải tiến mô hình trong các chương tiếp theo.

4.5.2 Định hướng cải thiện trong chương tiếp theo

Dựa trên các quan sát từ kết quả thực nghiệm và các hạn chế đã được phân tích, một số hướng cải thiện tiềm năng cho mô hình được đề xuất như sau:

(1) Tối ưu hóa kiến trúc mô hình

Trong chương tiếp theo, mô hình sẽ được mở rộng hoặc điều chỉnh nhằm tăng cường khả năng học đặc trưng, đặc biệt ở các vùng khó như Enhancing Tumor. Điều này có thể bao gồm:

- tích hợp attention mechanisms,
- sử dụng các khối convolution đa tỷ lệ,
- tăng chiều sâu mạng nhằm học tốt hơn các đặc trưng không gian phức tạp.

(2) Nâng cao chiến lược tiền xử lý và tăng cường dữ liệu

Các bước cải thiện có thể bao gồm chuẩn hóa cường độ thích nghi theo từng modal, mô phỏng nhiều thực tế, hoặc tạo dữ liệu biến thể vị trí – hình học nhằm tăng độ robust.

(3) Giảm ảnh hưởng của sự mất cân bằng lớp

Áp dụng focal loss, weighted loss hoặc kỹ thuật sampling có thể giúp mô hình tập trung nhiều hơn vào các vùng hiếm như ET.

(4) Xây dựng mô hình thích ứng cho dữ liệu thiếu modal

Một định hướng quan trọng là phát triển phiên bản mô hình có khả năng hoạt động với dữ liệu không đầy đủ hoặc chất lượng thấp – đây là yêu cầu thực tế trong môi trường bệnh viện.

(5) Tăng cường tính giải thích của mô hình

Tích hợp các phương pháp explainable AI (GradCAM, feature attribution) nhằm cung cấp minh chứng trực quan cho quyết định phân đoạn, giúp mô hình đáng tin cậy hơn trong ứng dụng lâm sàng.

(6) Tối ưu hóa chi phí tính toán

Nhằm tiến gần hơn đến triển khai thực tế, mô hình có thể được tinh giản (model compression), lượng tử hóa hoặc chuyển sang dạng lightweight encoder mà vẫn đảm bảo chất lượng phân vùng.

CHƯƠNG 5 KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Chương này tổng kết toàn bộ quá trình nghiên cứu, đánh giá mức độ đạt được các mục tiêu đề ra, đồng thời chỉ rõ những hạn chế còn tồn tại và đề xuất các hướng phát triển tiềm năng trong tương lai. Nội dung được trình bày một cách có hệ thống, khách quan và minh bạch, nhằm cung cấp cái nhìn tổng quan về giá trị khoa học cũng như khả năng ứng dụng thực tiễn của đề tài.

5.1 Kết luận chung

5.1.1 Tổng kết các kết quả đạt được

Đề tài đã hoàn thành việc xây dựng và triển khai một hệ thống phân vùng khối u não tự động trên ảnh MRI sử dụng kiến trúc học sâu U-Net 2D. Các kết quả chính đạt được bao gồm:

a) Về xử lý dữ liệu:

- Xây dựng quy trình tiền xử lý hoàn chỉnh cho bộ dữ liệu BraTS, bao gồm chuẩn hóa cường độ, đồng bộ hóa không gian và lọc nhiễu.
- Phân chia dữ liệu hợp lý theo tỷ lệ 70%-20%-10% (train-validation-test), đảm bảo tính độc lập giữa các tập và tránh data leakage.
- Áp dụng các kỹ thuật tăng cường dữ liệu (xoay, lật, biến đổi cường độ) giúp tăng tính đa dạng và giảm overfitting.

b) Về mô hình:

- Triển khai thành công kiến trúc U-Net 2D với encoder-decoder đối xứng, skip-connections và batch normalization.
- Tối ưu hóa hàm mất mát bằng cách kết hợp Dice Loss và Cross-Entropy Loss, cùng với class weights để cân bằng các vùng khối u.
- Áp dụng learning rate schedule với warmup và cosine annealing giúp quá trình huấn luyện hội tụ ổn định.

c) Về hiệu suất:

- Đạt **Dice Score trung bình 0.8355** trên tập validation (epoch 32), vượt trội so với baseline U-Net (~0.76) và tương đương với các nghiên cứu hiện đại.
- Kết quả chi tiết theo từng vùng:
 - **Enhancing Tumor (ET):** 0.8548 – cao nhất
 - **Necrotic Core (NCR):** 0.8309
 - **Edema (ED):** 0.8208
- Validation Loss đạt **0.1282** tại epoch 32, cho thấy mô hình phân vùng chính xác và ổn định.

d) Về quá trình thực nghiệm:

- Huấn luyện thành công qua 40 epochs với thời gian ~7 giây/batch
- Phân tích chi tiết quá trình hội tụ qua 3 giai đoạn: học nhanh ban đầu, ổn định và tối ưu hóa sâu.
- Phát hiện và xử lý hiện tượng overfitting nhẹ sau epoch 32 bằng early stopping.

5.1.2 Mức độ đáp ứng mục tiêu nghiên cứu

Đề tài đã hoàn thành đầy đủ các mục tiêu nghiên cứu đề ra từ đầu:

Mục tiêu 1: Nghiên cứu và triển khai mô hình học sâu cho phân vùng u não

Đạt: Đã triển khai thành công kiến trúc U-Net 2D với đầy đủ các thành phần hiện đại (skip-connections, batch normalization, data augmentation). Mô hình được tối ưu hóa về cả kiến trúc và siêu tham số.

Mục tiêu 2: Đạt hiệu suất phân vùng cao trên bộ dữ liệu BraTS

Đạt vượt mức: Dice Score 0.8355 vượt chỉ tiêu mong đợi (>0.80) và cao hơn đáng kể so với baseline. Cả 3 vùng khối u đều đạt trên 0.82.

Mục tiêu 3: Phân tích và đánh giá chi tiết hiệu suất mô hình

Đạt: Đã thực hiện phân tích toàn diện về quá trình huấn luyện, hiệu suất theo từng vùng, ảnh hưởng của các thành phần mô hình và các yếu tố ảnh hưởng đến chất lượng phân vùng.

Mục tiêu 4: So sánh với các phương pháp hiện có

Đạt: Đã so sánh chi tiết với U-Net gốc và Attention U-Net, chứng minh tính cạnh tranh của mô hình đề xuất.

Mục tiêu 5: Xây dựng quy trình hoàn chỉnh từ tiền xử lý đến đánh giá

Đạt: Quy trình đầy đủ từ thu thập dữ liệu, tiền xử lý, huấn luyện, đánh giá đến trực quan hóa kết quả đã được xây dựng và tài liệu hóa chi tiết.

Tổng kết: Đề tài đáp ứng 100% các mục tiêu nghiên cứu đề ra, với một số mục tiêu đạt vượt mức kỳ vọng ban đầu.

5.1.3 Những đóng góp chính của đề tài

a) Về mặt khoa học:

Chứng minh hiệu quả của U-Net 2D: Mô hình 2D đơn giản vẫn có thể đạt hiệu suất cao (>0.83) trên bài toán phân vùng u não 3D phức tạp, thách thức quan điểm cho rằng chỉ có mô hình 3D mới hiệu quả.

Phân tích sâu về class weights: Nghiên cứu cho thấy việc sử dụng class weights [0.1, 2.0, 2.0, 3.0] giúp cân bằng hiệu quả giữa các vùng khối u, đặc biệt cải thiện đáng kể hiệu suất trên vùng ET (từ ~0.74 lên 0.86).

Đánh giá chi tiết quá trình hội tụ: Cung cấp phân tích toàn diện về 3 giai đoạn học của mô hình, giúp hiểu rõ cơ chế hoạt động và tối ưu hóa training.

So sánh thực nghiệm: Đóng góp dữ liệu benchmark cho cộng đồng nghiên cứu về phân vùng u não trên BraTS.

b) Về mặt kỹ thuật:

Quy trình tiền xử lý tối ưu: Xây dựng pipeline xử lý dữ liệu MRI chuẩn hóa, có thể tái sử dụng cho các bài toán tương tự.

Kết hợp hàm mất mát hiệu quả: Chứng minh việc kết hợp Dice Loss + Cross-Entropy với class weights phù hợp mang lại kết quả tốt hơn so với chỉ dùng một hàm đơn lẻ.

Learning rate schedule hợp lý: Áp dụng thành công warmup + cosine annealing giúp tăng tốc hội tụ và đạt local minima tốt hơn.

Early stopping thông minh: Xác định chính xác thời điểm dừng huấn luyện (epoch 32) để tránh overfitting.

c) Về mặt ứng dụng:

Tính khả thi cao: Mô hình đơn giản, yêu cầu tài nguyên hợp lý (10GB VRAM, ~7s/batch), dễ triển khai trong môi trường thực tế.

Tiềm năng hỗ trợ lâm sàng: Với Dice Score >0.83 , mô hình có thể hỗ trợ bác sĩ trong việc phân tích MRI nhanh hơn và chính xác hơn.

Code và tài liệu đầy đủ: Toàn bộ quy trình được tài liệu hóa chi tiết, có thể tái sử dụng cho nghiên cứu và giảng dạy.

5.2 Hạn chế của đề tài

5.2.1 Hạn chế về dữ liệu

a) Kích thước tập dữ liệu:

Chi sử dụng 700 mẫu training và 200 mẫu validation do hạn chế về tài nguyên tính toán và thời gian, trong khi bộ BraTS đầy đủ có thể lên đến 300-400 ca bệnh với hàng chục nghìn lát ảnh.

Tập dữ liệu nhỏ dẫn đến khả năng tổng quát hóa hạn chế và dễ xảy ra overfitting.

b) Đa dạng dữ liệu:

Dữ liệu BraTS đã được chuẩn hóa trước, do đó mô hình chưa được kiểm tra trên dữ liệu MRI thô từ các máy chụp khác nhau.

Chưa thử nghiệm trên các loại khối u não khác ngoài glioma (ví dụ: meningioma, pituitary adenoma).

c) Hiện tượng bất thường:

Validation Dice cao hơn Training Dice (0.8355 vs 0.7606) cho thấy tập validation có thể không đại diện đầy đủ cho độ khó của dữ liệu thực tế.

Chưa có phân tích phân bố chi tiết về kích thước khối u, mức độ ác tính trong các tập train/val/test.

5.2.2 Hạn chế về mô hình

a) Kiến trúc 2D:

Mô hình chỉ xử lý từng lát 2D độc lập, không tận dụng thông tin ngữ cảnh 3D giữa các lát liên tiếp.

Điều này có thể dẫn đến sự không liên tục về mặt không gian trong kết quả phân vùng 3D cuối cùng.

b) Hiệu suất không đồng đều:

Vùng NCR có độ dao động cao ($0.8309 \rightarrow 0.6243$ trong 8 epochs), cho thấy mô hình chưa ổn định hoàn toàn với vùng khó này.

Chưa có cơ chế attention hoặc multi-scale để tập trung vào các vùng quan trọng.

c) Hiện tượng overfitting:

Sau epoch 32, Training Dice tiếp tục tăng nhưng Validation Dice giảm 10%, cho thấy mô hình bắt đầu "nhớ" training data thay vì học đặc trưng tổng quát.

Regularization hiện tại (data augmentation, batch normalization) chưa đủ mạnh.

d) Thiếu cơ chế giải thích:

Mô hình hoạt động như "black box", không cung cấp thông tin về vùng nào được chú ý hay tại sao một quyết định phân vùng được đưa ra.

Điều này hạn chế khả năng tin cậy trong ứng dụng lâm sàng.

5.2.3 Hạn chế trong quá trình thực nghiệm

a) Tài nguyên tính toán:

Chi sử dụng một GPU RTX 3080 (10GB), hạn chế khả năng thử nghiệm batch size lớn hoặc mô hình phức tạp hơn.

Không có khả năng huấn luyện mô hình 3D đầy đủ do yêu cầu bộ nhớ cao.

b) Thời gian thực nghiệm:

Do hạn chế thời gian, chưa thực hiện cross-validation hoặc huấn luyện nhiều lần với random seeds khác nhau để đánh giá độ ổn định của mô hình.

Chưa thử nghiệm đầy đủ nhiều kiến trúc khác (Attention U-Net, U-Net++, DeepLabV3+) để so sánh công bằng.

c) Đánh giá hạn chế:

Chỉ đánh giá trên tập test một lần, chưa có đánh giá trên dữ liệu external validation từ bệnh viện thực tế.

Chưa có đánh giá từ bác sĩ chuyên khoa về tính hữu ích của kết quả phân vùng trong thực hành lâm sàng.

Chưa đo thời gian suy luận (inference time) chi tiết cho từng ca bệnh.

d) Phân tích thực nghiệm:

Chưa thực hiện ablation study chi tiết để đánh giá đóng góp của từng thành phần (skip-connections, batch normalization, class weights) một cách định lượng.

Chưa phân tích sâu về các trường hợp mô hình dự đoán sai hoàn toàn (failure cases).

5.3 Hướng phát triển trong tương lai

5.3.1 Cải thiện mô hình và kiến trúc học sâu

a) Nâng cấp lên kiến trúc 3D:

Triển khai U-Net 3D hoặc V-Net để tận dụng thông tin ngữ cảnh không gian giữa các lát ảnh liên tiếp.

Sử dụng 2.5D approach (xử lý nhiều lát liên tiếp cùng lúc) như giải pháp trung gian tiết kiệm tài nguyên.

b) Tích hợp cơ chế attention:

Áp dụng Attention U-Net hoặc Transformer-based models (TransUNet, Swin-UNet) để mô hình tập trung vào các vùng quan trọng.

Sử dụng self-attention để học mối quan hệ xa trong ảnh y tế.

c) Kiến trúc đa tỷ lệ:

Triển khai U-Net++ hoặc DeepLabV3+ với atrous convolution để bắt được đặc trưng ở nhiều tỷ lệ khác nhau.

Áp dụng pyramid pooling để cải thiện phân vùng vùng nhỏ như ET.

d) Kỹ thuật ensemble:

Huấn luyện nhiều mô hình với random seeds khác nhau và kết hợp predictions bằng voting hoặc averaging.

Ensemble giữa các kiến trúc khác nhau (U-Net 2D + U-Net 3D + Attention U-Net) để tăng độ robust.

e) Cải thiện regularization:

Thêm Dropout layers (0.3-0.5) vào decoder.

Áp dụng weight decay (L2 regularization) với hệ số 1e-4.

Sử dụng mixup hoặc cutmix cho medical images.

f) Loss function nâng cao:

Thử nghiệm Focal Dice Loss để tập trung vào các vùng khó phân vùng.

Kết hợp Boundary Loss để cải thiện độ chính xác của đường biên.

Áp dụng Tversky Loss với tham số α , β điều chỉnh được để cân bằng sensitivity và specificity.

g) Tăng tính giải thích (Explainable AI):

Tích hợp Grad-CAM hoặc Grad-CAM++ để trực quan hóa vùng mô hình chú ý.

Sử dụng SHAP values để giải thích đóng góp của từng vùng ảnh vào quyết định cuối cùng.

Phát triển interface hiển thị heatmap attention cho bác sĩ.

5.3.2 Mở rộng dữ liệu và thử nghiệm đa trung tâm

a) Tăng kích thước dữ liệu:

Sử dụng toàn bộ bộ dữ liệu BraTS (300-400 ca bệnh thay vì 500 mẫu lát ảnh).

Áp dụng cross-validation k-fold ($k=5$ hoặc $k=10$) để đánh giá robust và giảm bias do phân chia dữ liệu.

b) Data augmentation nâng cao:

Elastic deformation: mô phỏng biến dạng mô não thực tế.

Intensity shifting: mô phỏng sự khác biệt giữa các máy MRI.

Gaussian noise, blur: tăng tính robust với nhiễu.

CutOut, GridMask: giúp mô hình học đặc trưng cục bộ tốt hơn.

c) Mở rộng sang các loại khối u khác:

Thu thập dữ liệu meningioma, pituitary adenoma, metastatic tumors.

Xây dựng mô hình multi-task có khả năng phân loại loại u và phân vùng đồng thời.

d) Thử nghiệm đa trung tâm:

Hợp tác với các bệnh viện để thu thập dữ liệu MRI thực tế từ nhiều máy chụp khác nhau.

Đánh giá domain adaptation và transfer learning khi áp dụng mô hình trên dữ liệu từ các trung tâm khác.

Phân tích domain shift giữa dữ liệu BraTS chuẩn hóa và dữ liệu thực tế.

e) Federated learning:

Triển khai federated learning để huấn luyện mô hình trên dữ liệu từ nhiều bệnh viện mà không cần chia sẻ dữ liệu thô (bảo mật).

Áp dụng privacy-preserving techniques như differential privacy.

5.3.3 Ứng dụng thực tế trong lâm sàng

a) Xây dựng hệ thống end-to-end:

Phát triển web application hoặc desktop application cho phép bác sĩ upload ảnh MRI và nhận kết quả phân vùng ngay lập tức.

Tích hợp với PACS của bệnh viện.

b) Tối ưu hóa tốc độ suy luận:

Áp dụng model quantization (INT8, FP16) để giảm kích thước mô hình và tăng tốc độ.

Sử dụng TensorRT hoặc ONNX Runtime để tối ưu hóa inference.

Triển khai trên edge devices (Jetson Nano, Raspberry Pi) cho các phòng khám nhỏ.

c) Đánh giá lâm sàng:

Thực hiện clinical validation study với sự tham gia của bác sĩ X quang và bác sĩ thần kinh.

So sánh thời gian phân tích giữa phương pháp thủ công và hỗ trợ bởi AI.

Đánh giá inter-rater agreement giữa mô hình AI và nhiều bác sĩ.

d) Tích hợp vào quy trình chẩn đoán:

Xây dựng clinical decision support system (CDSS) giúp bác sĩ đưa ra quyết định điều trị.

Kết hợp phân vùng với radiomics features để dự đoán tiên lượng bệnh.

Tích hợp với genomic data để phân tích đa phương thức (multi-modal analysis).

e) Theo dõi tiến triển bệnh:

Phát triển chức năng longitudinal analysis để so sánh kích thước khối u giữa các lần chụp MRI.

Tự động tính toán tumor growth rate và cảnh báo khi có biến đổi đáng kể.

f) Đào tạo và hỗ trợ:

Xây dựng training dataset cho sinh viên y khoa và bác sĩ mới.

Tạo interactive visualization tool giúp bác sĩ hiểu cách mô hình hoạt động.

5.3.4 Tự động hóa toàn bộ quy trình phân tích MRI

a) Pipeline hoàn toàn tự động:

Bước 1: Tự động phát hiện và trích xuất vùng não từ ảnh MRI đầy đủ (skull stripping).

Bước 2: Tự động phát hiện sự hiện diện của khối u (tumor detection).

Bước 3: Phân vùng chi tiết các vùng khối u (segmentation).

Bước 4: Tự động tính toán các chỉ số lâm sàng (thể tích, vị trí, đặc điểm hình thái).

Bước 5: Tự động tạo báo cáo y tế có cấu trúc.

b) Tích hợp đa phương thức (multi-modal):

Kết hợp MRI với CT scan để phân tích toàn diện hơn.

Tích hợp với PET scan để đánh giá hoạt động trao đổi chất của khối u.

Phân tích đồng thời functional MRI (fMRI) để xác định vùng não chức năng quan trọng trước phẫu thuật.

c) Dự đoán và phân tầng nguy cơ:

Xây dựng mô hình survival prediction dựa trên đặc trưng hình ảnh và lâm sàng.

Phân loại tumor grade (độ ác tính) tự động từ ảnh MRI.

Dự đoán treatment response (đáp ứng với hóa trị, xạ trị) dựa trên đặc trưng ban đầu.

d) Hỗ trợ lập kế hoạch điều trị:

Tích hợp với surgical planning software để hỗ trợ lập kế hoạch phẫu thuật.

Tự động xác định safe resection margins (ranh giới cắt bỏ an toàn).

Mô phỏng radiation therapy planning dựa trên phân vùng chính xác.

e) Giám sát real-time:

Phát triển hệ thống **intraoperative MRI guidance** hỗ trợ phẫu thuật viên trong quá trình mổ.

Cập nhật phân vùng real-time khi có ảnh MRI mới trong quá trình theo dõi điều trị.

f) Tích hợp với Electronic Health Records (EHR):

Tự động đồng bộ kết quả phân vùng vào hồ sơ bệnh án điện tử.

Tạo structured radiology report tự động từ kết quả phân vùng.

Cảnh báo tự động khi phát hiện biến đổi bất thường.

5.4 Kết luận chương

Chương 5 đã tổng kết toàn diện kết quả nghiên cứu của đề tài, bao gồm:

Về kết quả đạt được:

Xây dựng thành công mô hình U-Net 2D phân vùng khối u não với Dice Score 0.8355, vượt trội so với baseline và tương đương các nghiên cứu hiện đại.

Hoàn thành 100% các mục tiêu nghiên cứu đề ra, với một số mục tiêu đạt vượt mức kỳ vọng.

Đóng góp về mặt khoa học (phân tích class weights, quá trình hội tụ), kỹ thuật (pipeline tối ưu) và ứng dụng (tính khả thi cao).

Về hạn chế:

Kích thước dữ liệu nhỏ (500 train, 100 val) hạn chế khả năng tổng quát hóa.

Kiến trúc 2D không tận dụng thông tin ngũ cảnh 3D.

Hiện tượng overfitting nhẹ sau epoch 32 và dao động cao ở vùng NCR.

Chưa có đánh giá lâm sàng thực tế từ bác sĩ chuyên khoa.

Về hướng phát triển:

Nâng cấp lên kiến trúc 3D hoặc tích hợp attention mechanisms.

Mở rộng dữ liệu và thử nghiệm đa trung tâm.

Phát triển hệ thống ứng dụng thực tế trong lâm sàng.

Tự động hóa toàn bộ quy trình phân tích MRI từ đầu đến cuối.

TÀI LIỆU THAM KHẢO

- [1] Lapointe, S., Perry, A., & Butowski, N. A. (2018). Primary brain tumours in adults. *The Lancet*, 392(10145), 432-446.
- [2] Clifton, W., & Reimer, R. (2019). Metastatic brain tumors. *Comprehensive Overview of Modern Surgical Approaches to Intrinsic Brain Tumors*, 159-167.
- [3] Udaka, Y. T., & Packer, R. J. (2018). Pediatric brain tumors. *Neurologic clinics*, 36(3), 533-556.
- [4] Jelski, W., & Mroczko, B. (2021). Molecular and circulating biomarkers of brain tumors. *International journal of molecular sciences*, 22(13), 7039.
- [5] Hendrix, P., Hans, E., Griessenauer, C. J., Simgen, A., Oertel, J., & Karbach, J. (2017). Neurocognitive status in patients with newly-diagnosed brain tumors in good neurological condition: the impact of tumor type, volume, and location. *Clinical neurology and neurosurgery*, 156, 55-62.
- [6] Mehta, S., Shah, A., & Jung, H. (2017). Diagnosis and treatment options for sequelae following radiation treatment of brain tumors. *Clinical neurology and neurosurgery*, 163, 1-8.
- [7] Kumar, A., Ramachandran, M., Gandomi, A. H., Patan, R., Lukasik, S., & Soundarapandian, R. K. (2019). A deep neural network based classifier for brain tumor diagnosis. *Applied Soft Computing*, 82, 105528.
- [8] Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., & Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 74(3), 229-263.
- [9] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 71(3), 209-249.
- [10] Ostrom, Q. T., Gittleman, H., Liao, P., Rouse, C., Chen, Y., Dowling, J., ... & Barnholtz-Sloan, J. (2014). CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2007–2011. *Neuro-oncology*, 16(suppl_4), iv1-iv63.

- [11] Weller, M., Wick, W., Aldape, K., Brada, M., Berger, M., Pfister, S. M., ... & Reifenberger, G. (2015). Glioma. *Nature reviews Disease primers*, 1(1), 1-18.
- [12] Louis, D. N., Ohgaki, H., Wiestler, O. D., Cavenee, W. K., Burger, P. C., Jouvet, A., ... & Kleihues, P. (2007). The 2007 WHO classification of tumours of the central nervous system. *Acta neuropathologica*, 114(2), 97-109.
- [13] Louis, D. N., Perry, A., Reifenberger, G., Von Deimling, A., Figarella-Branger, D., Cavenee, W. K., ... & Ellison, D. W. (2016). The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta neuropathologica*, 131(6), 803-820.
- [14] Louis, D. N., Perry, A., Wesseling, P., Brat, D. J., Cree, I. A., Figarella-Branger, D., ... & Ellison, D. W. (2021). The 2021 WHO classification of tumors of the central nervous system: a summary. *Neuro-oncology*, 23(8), 1231-1251.
- [15] Onciu, R., Brehar, F. M., Toader, C., Covache-Busuioc, R. A., Glavan, L. A., Bratu, B. G., ... & Ciurea, A. V. (2024). Deciphering glioblastoma: fundamental and novel insights into the biology and therapeutic strategies of gliomas. *Current Issues in Molecular Biology*, 46(3), 2402-2443.
- [16] McNeill, K. A. (2016). Epidemiology of brain tumors. *Neurologic clinics*, 34(4), 981-998.
- [17] Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE transactions on medical imaging*, 35(5), 1240-1251.
- [18] McRobbie, D. W., Moore, E. A., Graves, M. J., & Prince, M. R. (2017). *MRI from Picture to Proton*. Cambridge university press.
- [19] Brown, R. W., Cheng, Y. C. N., Haacke, E. M., Thompson, M. R., & Venkatesan, R. (2014). *Magnetic resonance imaging: physical principles and sequence design*. John Wiley & Sons.

KẾ HOẠCH THỰC HIỆN

Tên đề tài: **PHÂN ĐOẠN KHỐI U GLIOMA TỪ MRI ĐA KÊNH BẰNG U-NET**

Giảng viên: ThS. Võ Quang Hoàng Khang

Ngày bắt đầu: 25/08/2025

Công việc	Dự kiến ngày thực hiện (ngày)	Ngày thực hiện thực tế (ngày)	Đạt (%)
1. Xác định đề tài và mục tiêu	7	7	100
2. Nghiên cứu lý thuyết MRI & U-Net	10	10	100
3. Chuẩn bị, tổ chức và tiền xử lý dữ liệu	12	12	100
4. Xây dựng pipeline dữ liệu	10	10	100
5. Xây dựng mô hình U-Net 2D	14	14	100
6. Huấn luyện mô hình và tối ưu tham số	18	18	100
7. Đánh giá và phân tích kết quả	10	10	100
8. Viết báo cáo mô hình và thực nghiệm	14	14	100
9. Hoàn thiện báo cáo	6	6	100

NHẬT KÝ LÀM VIỆC

TG (Ngày/ Tháng)	Công việc	Dự kiến ngày thực hiện (ngày)	Ngày thực hiện thực tế (ngày)	Đạt (%)
25/08 → 31/08	1. Chọn bài toán phân đoạn u não. 2. Xác định mục tiêu, phạm vi đề tài. 3. Xác định đầu vào/đầu ra và tiêu chí đánh giá. 4. Lập kế hoạch 15 tuần.	7	7	100
01/09 → 10/09	1. Tìm hiểu cấu trúc MRI và các modal (T1, T1ce, T2, FLAIR). 2. Nghiên cứu bản chất dữ liệu y sinh (nhiều, bias field). 3. Học kiến trúc U-Net 2D (encoder, decoder, skip). 4. Tìm hiểu các hàm mất mát cho segmentation.	10	10	100
11/09 → 22/09	1. Tải dataset BraTS và tổ chức thư mục. 2. Đọc file NIfTI, lọc dữ liệu lỗi. 3. Chuẩn hóa cường độ (z-score). 4. Cắt/resize lát cắt. 5. Bias correction (N4ITK). 6. Kiểm tra mask & trực quan hóa.	12	12	100
23/09 → 02/10	1. Tạo DataLoader dạng 2D slice. 2. Thiết lập augmentation (flip, rotate, brightness). 3. Chia train/valid/test hợp lý. 4. Kiểm thử pipeline bằng batch	10	10	100

	nhỏ.			
03/10 → 16/10	1. Thiết kế số kênh và số tầng mô hình. 2. Xây dựng encoder (downsampling). 3. Xây dựng decoder (upsampling + skip). 4. Viết forward pass. 5. Cấu hình loss, optimizer, scheduler. 6. Kiểm thử mô hình với batch nhỏ.	14	14	100
17/10 → 03/11	1. Huấn luyện mô hình lần 1 (baseline). 2. Theo dõi loss, Dice score. 3. Điều chỉnh LR, batch size, optimizer. 4. Điều chỉnh augmentation nếu overfitting. 5. Kết hợp loss (Dice + BCE/Tversky). 6. Huấn luyện lần 2 (mô hình tối ưu). 7. Lưu checkpoint tốt nhất.	18	18	100
04/11 → 13/11	1. Chạy mô hình trên tập test. 2. Tính các chỉ số: Dice, IoU, Precision, Recall. 3. So sánh ground-truth và dự đoán. 4. Phân tích trường hợp mô hình hoạt động tốt/chưa tốt. 5. Tổng hợp nhận xét.	10	10	100
14/11 → 27/11	1. Mô tả mô hình U-Net 2D. 2. Mô tả xử lý dữ liệu. 3. Mô tả quy trình huấn luyện. 4. Báo cáo kết quả định lượng	14	14	100

	và trực quan hóa. 5. Viết phần thảo luận và đánh giá.			
28/11 → 03/12	1. Chính sửa bối cảnh, biểu đồ, hình ảnh. 2. Kiểm tra lỗi văn bản và trích dẫn. 3. Chuẩn bị slide trình bày. 4. Hoàn thiện file nộp.	6	6	100
Tổng		101	101	100

Tân Phan Mạnh

Final_KLTN_PhanManhTan_NguyenVanThuong.docx

-  Nộp bài và Kiểm tra Turnitin - 1 (Moodle PP)
-  TURNITIN_KHÓA LUẬN TỐT NGHIỆP_KHOA CNTT_HK1_2025_2026 (Moodle PP)
-  Industrial University of Ho Chi Minh City

Document Details

Submission ID

trn:oid:::1:3432034669

110 Pages

Submission Date

Dec 3, 2025, 9:30 PM GMT+7

30,325 Words

Download Date

Dec 3, 2025, 9:49 PM GMT+7

119,229 Characters

File Name

520789_Tân_Phản_Mạnh_Final_KLTN_PhanManhTan_NguyenVanThuong_388618_2096808716.docx

File Size

1.9 MB

7% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- ▶ Bibliography
 - ▶ Quoted Text
-

Top Sources

5%	 Internet sources
5%	 Publications
1%	 Submitted works (Student Papers)

Top Sources

- 5% Internet sources
5% Publications
1% Submitted works (Student Papers)
-

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

Rank	Type	Source	Percentage
1	Internet	www.ctu.edu.vn	3%
2	Internet	gust.edu.vn	2%
3	Publication	Phenikaa University	<1%
4	Publication	Ton Duc Thang University	<1%
5	Publication	Phan Anh Cang, Nguyễn Khắc Tường, Nguyễn Ngọc Hoàng Quyên. "MÔ HÌNH HÓ...	<1%
6	Publication	Hanoi National University of Education	<1%