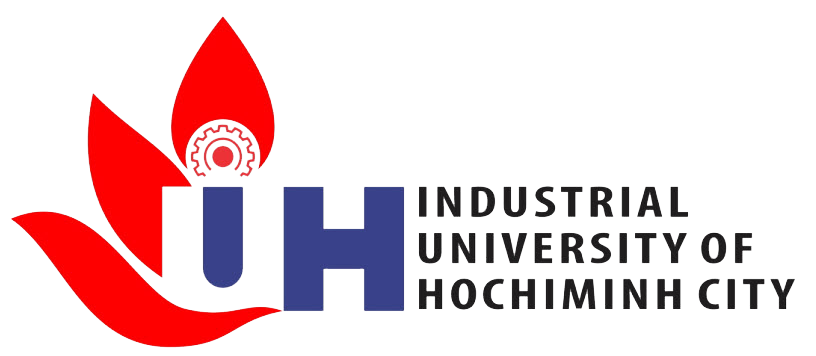
**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TPHCM**

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO ĐỀ TÀI**

**TÊN ĐỀ TÀI**

**PHÂN VÙNG KHỐI U NÃO GLIOMA TRÊN MRI ĐA KÊNH SỬ DỤNG MÔ HÌNH HỌC SÂU**

**GIẢNG VIÊN HƯỚNG DẪN: ThS.Võ Quang Hoàng Khang**

**SINH VIÊN THỰC HIỆN 1: Lê Đình Nam**

**MÃ SỐ SINH VIÊN: 21061091**

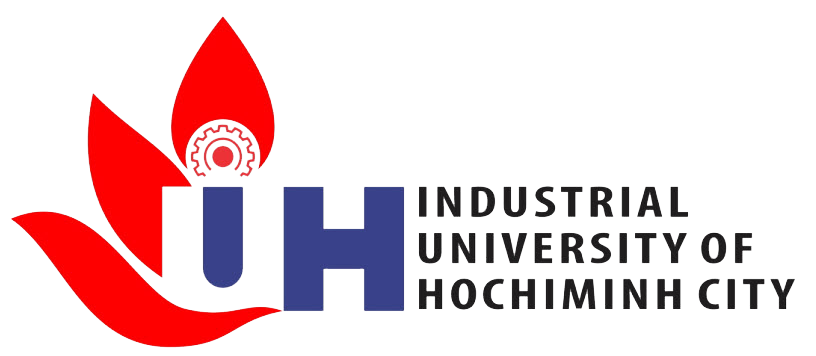
**SINH VIÊN THỰC HIỆN 2: Phan Mạnh Tân**

**MÃ SỐ SINH VIÊN: 21114061**

*TP HỒ CHÍ MINH, Tháng 05 năm 2025*

**INDUSTRIAL UNIVERSITY OF HO CHI MINH CITY**

**FACULTY OF INFORMATION TECHNOLOGY**



**PROJECT REPORT**

**SEGMENTATION OF GLIOMA TUMORS IN MULTIMODAL MRI SCANS USING DEEP NEURAL NETWORKS**

**SUPERVISOR: MSc. Vo Quang Hoang Khang**

**STUDENT NAME: Le Dinh Nam**

**STUDENT CODE: 21061091**

**STUDENT NAME: Phan Manh Tan**

**STUDENT CODE: 21114061**

*HO CHI MINH CITY, Month 05 year 2025*

**TÓM TẮT**

Nghiên cứu này trình bày một mô hình học sâu cải tiến nhằm hỗ trợ phân vùng khối u não từ ảnh cộng hưởng từ (MRI), một nhiệm vụ quan trọng trong chẩn đoán và điều trị bệnh u não, đặc biệt là glioma. Mô hình được huấn luyện và đánh giá trên tập dữ liệu BraTS 2020, sử dụng các chỉ số đánh giá như Dice coefficient, Sensitivity, Specificity và IoU. Kết quả cho thấy mô hình đạt hiệu quả phân vùng cao đối với các vùng phụ của khối u như vùng tăng sinh, phù nề và lõi hoại tử, hứa hẹn khả năng ứng dụng trong thực tế lâm sàng nhằm hỗ trợ bác sĩ chẩn đoán chính xác và lập kế hoạch điều trị hiệu quả hơn.

**LỜI CẢM ƠN**

Chúng em xin chân thành gửi lời cảm ơn đến ThS. Võ Quang Hoàng Khang, giảng viên môn Thị giác máy tính, người đã trực tiếp giảng dạy, hướng dẫn và tận tình hỗ trợ chúng em trong suốt quá trình thực hiện dự án này.

Trong quá trình làm bài, thầy không chỉ giúp em hiểu rõ hơn về kiến thức chuyên môn, đặc biệt là các kỹ thuật phân đoạn ảnh y tế, mà còn định hướng cách áp dụng lý thuyết vào thực tế thông qua bài toán cụ thể. Những góp ý quý báu của thầy đã giúp chúng em hoàn thiện dự án một cách hiệu quả và có hệ thống hơn.

Chúng em xin cảm ơn thầy vì sự tận tâm và nhiệt huyết trong giảng dạy cũng như trong việc đồng hành cùng sinh viên trong suốt học kỳ vừa qua.

Trân trọng cảm ơn!

**NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN**

**TP. Hồ Chí Minh, ngày…. tháng …. năm….**

CHỮ KÝ CỦA GIẢNG VIÊN

**NHẬN XÉT CỦA GIÁO VIÊN PHẢN BIỆN 1**

**TP. Hồ Chí Minh, ngày…. Tháng …. năm….**

CHỮ KÝ CỦA GIẢNG VIÊN

**NHẬN XÉT CỦA GIÁO VIÊN PHẢN BIỆN 2**

**TP. Hồ Chí Minh, ngày…. tháng …. năm….**

CHỮ KÝ CỦA GIẢNG VIÊN

**LỜI MỞ ĐẦU**

Trong kỷ nguyên số hiện nay, công nghệ máy học và học sâu ngày càng trở nên quan trọng và ảnh hưởng sâu rộng trong mọi lĩnh vực, đặc biệt là trong y học. Sự phát triển của các mô hình học sâu đã mang lại những bước tiến vượt bậc trong việc giải quyết các vấn đề thực tiễn, từ việc phân tích ảnh, nhận diện đối tượng cho đến hỗ trợ chẩn đoán bệnh tật.

Một trong những ứng dụng nổi bật của học sâu trong y tế là phân đoạn ảnh y tế, đặc biệt là phân đoạn khối u não từ các ảnh MRI. Việc phân vùng chính xác các khu vực tổn thương trong khối u não đóng vai trò quan trọng trong việc hỗ trợ bác sĩ đưa ra chẩn đoán, lập kế hoạch điều trị và theo dõi tiến trình điều trị của bệnh nhân một cách hiệu quả.

Chính vì lý do này, dự án môn Thị giác máy tính của chúng em được thực hiện với chủ đề "Phân đoạn khối u não từ ảnh MRI bằng các mô hình học sâu". Mục tiêu của dự án là nghiên cứu và áp dụng các mô hình học sâu để tự động phân đoạn các vùng tổn thương trong ảnh MRI đa chuỗi từ tập dữ liệu BraTS 2020. Chúng em sẽ áp dụng các kiến trúc hiện đại kết hợp cơ chế attention, tích chập sâu, và xử lý ảnh y tế để nâng cao hiệu quả phân đoạn.

Thông qua dự án này, chúng em hy vọng có thể đóng góp vào việc ứng dụng thị giác máy tính trong y học, giúp cải thiện quá trình chẩn đoán và hỗ trợ điều trị cho bệnh nhân mắc khối u não.

**MỤC LỤC**

[**CHƯƠNG 1. TỔNG QUAN 1**](#_apog8xaogpoh)

[**1.1 Lý do chọn đề tài 1**](#_qfxotfuo25qg)

[**1.2 Mục tiêu nghiên cứu 2**](#_b1w0vxxs0oqv)

[**1.3 Phạm vi nghiên cứu 3**](#_nisc8zj8608b)

[**1.4 Phương pháp nghiên cứu 3**](#_mxhbqkhygye7)

[**1.5 Kết cấu đồ án. 4**](#_gtpff0nqk6c9)

[**CHƯƠNG 2. CƠ SỞ LÝ THUYẾT 6**](#_a8enzt9ja62a)

[**2.1. Học sâu (Deep learning) 6**](#_79baz6nuzaww)

[**2.2. Tăng cường dữ liệu (Data augmentation) 6**](#_jjmhd0cn2oiz)

[**2.3. Kiến trúc U-Net(The U-Net model architecture) 7**](#_bgrwzuyy6e41)

[**2.4. Kỹ thuật tích chập thông thường (Standard Convolution Layers) 7**](#_c1qcylnxk3h6)

[**2.6. Các phương pháp đánh giá mô hình 8**](#_h4lhiv10oa9o)

[**1. Dice Coefficient (Hệ số Dice) 8**](#_rtlislav4fv6)

[**2. Precision (Độ chính xác dự đoán dương) 8**](#_9ss9eos9j0ho)

[**4. Specificity (Độ đặc hiệu) 9**](#_qqr84ulm3wzz)

[**5. Mean IoU (Intersection over Union) 9**](#_acp75wjv4d13)

[**6. Dice Coefficient cho từng loại khối u (Necrotic, Edema, Enhancing) 10**](#_49lv4q1smjlz)

[**CHƯƠNG 3. MÔ HÌNH ĐỀ XUẤT 11**](#_7vik5xrd7idg)

[**3.1 Mô hình tổng quát 11**](#_xe31dehscmaj)

[**Nhánh Co Cụm (Encoder): 12**](#_io9n1uhgeqk3)

[**Nhánh Mở Rộng (Decoder): 12**](#_8an5ueicw9ii)

[**Lớp Đầu ra (Output Layer): 12**](#_inh09fsv2wc1)

[**3.2 Đặc trưng của mô hình đề xuất 13**](#_fuox9eobohvg)

[**3.2.1. Trích xuất đặc trưng (Feature Extraction) 13**](#_wrfy6224wn31)

[**3.2.2. Hợp nhất đặc trưng dữ liệu 17**](#_e1beo563kt02)

[**CHƯƠNG 4. THỰC NGHIỆM 20**](#_ybtwzujcyncu)

[**4.1 Môi trường thực nghiệm 20**](#_q31ve0x1n4ar)

[**4.1.1. Cấu hình phần cứng 20**](#_pfkkhlmiaf3k)

[**4.1.2. Cấu hình phần mềm 20**](#_heqqtrglj1ls)

[**4.1.3. Thiết lập môi trường 20**](#_vkg4jpas2ozh)

[**4.1.4. Lý do chọn Kaggle 21**](#_vx478r9jegpf)

[**4.2 Tập dữ liệu 21**](#_jb50tpoti9qv)

[**4.2.1. Nguồn Dữ Liệu 21**](#_wijz1pxynbzp)

[**4.2.2. Số Lượng Dữ Liệu 21**](#_lm53ix85zq53)

[**4.2.3. Tiền Xử Lý Dữ Liệu 22**](#_ymc6c82ua81w)

[**4.2.4. Tăng cường dữ liệu 22**](#_7r9ua84x12h5)

[**Minh họa mã nguồn tăng cường trong DataGenerator (giả lập): 23**](#_ma9iq2w0saxt)

[**4.2.5. Chia dữ liệu 24**](#_195wahlz9uxq)

[**4.2.6. Lý do chọn tập dữ liệu 25**](#_xjxvg3r34y9k)

[**4.3 Ứng dụng thực nghiệm 25**](#_yo1kyaee1438)

[**4.3.1. Quy trình huấn luyện mô hình 25**](#_hbllobrc69xy)

[**4.3.2. Cấu hình huấn luyện 26**](#_yv4nkmjt0em)

[**4.3.3. Kết quả thực nghiệm 26**](#_y7mev2z5rtqu)

[**4.4 Đánh giá kết quả 28**](#_mp7klv2cdlhe)

[**CHƯƠNG 5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 29**](#_3hde37rpizh9)

[**5.1 Kết luận 29**](#_8h880gjyvne)

[**5.2 Hướng phát triển 30**](#_6tcc0i4r82t)

**MỤC LỤC HÌNH ẢNH**

[Hình 1. Kiến trúc mô hình 11](#_d35p0cxbpev)

[Hình 2. Độ chính xác trong quá trình huấn luyện 27](#_z72109npzi1i)

[Hình 3. Loss trong quá trình huấn luyện 27](#_dqf50szey414)

[Hình 4. Ma trận nhầm lẫn của mô hình đề xuất 28](#_pq22bkh3aj8x)

[Hình 5. Biểu đồ so sánh số liệu các mô hình phổ biến với mô hình đề xuất 29](#_y3a6dmxwbgum)

[Hình 6. Biểu đồ so sánh số liệu các mô hình thay đổi khối khác với mô hình đề xuất 32](#_dfuj13h1jo4j)

**MỤC LỤC BẢNG BIỂU**

[Bảng 1. Các phương pháp và tham số tăng cường dữ liệu 24](#_vtmo46keeedy)

[Bảng 2. Kết quả thử nghiệm và xác thực chéo của mô hình đề xuất 26](#_qls9y287d7l7)

[Bảng 3. So sánh các tiêu chí đánh giá của mô hình đề xuất với các biến thể của nó 31](#_3gwmcz2pz5tx)

[Bảng 4: So sánh các tiêu chí đánh giá của mô hình gốc với mô hình đề xuất 33](#_8wliywwx30sr)

**DANH MỤC CÁC THUẬT NGỮ VIẾT TẮT**

| Từ viết tắt | Từ đầy đủ | Nghĩa |
| --- | --- | --- |
| **MRI** | Magnetic Resonance Imaging | Một phương pháp chẩn đoán hình ảnh sử dụng từ trường mạnh và sóng radio để tạo ra các hình ảnh chi tiết của cơ thể, đặc biệt là các mô mềm như não, cột sống, cơ và các cơ quan nội tạng. MRI không sử dụng tia X, vì vậy nó được coi là một phương pháp an toàn hơn so với các kỹ thuật chẩn đoán hình ảnh khác |
| **BraTS 2020** | Brain Tumor Segmentation 2020 | Dữ liệu **BraTS 2020** bao gồm nhiều lần chụp MRI đa phương thức của bệnh nhân bị u nguyên bào thần kinh đệm (GBM/HGG) và u nguyên bào thần kinh đệm bậc thấp (LGG), với chẩn đoán xác nhận và dữ liệu sống còn (OS). Dữ liệu này được sử dụng 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. Dữ liệu xác thực sẽ được công bố vào ngày 1 tháng 7, và dữ liệu thử nghiệm sẽ được cung cấp trong 48 giờ vào các ngày 29 tháng 8 và 12 tháng 9, với kết quả cuối cùng được tải lên trước ngày 16 tháng 9. |
| **GBM** | Glioblastoma Multiforme | Một loại u nguyên bào thần kinh đệm ác tính, phát triển nhanh và thường gặp nhất trong các khối u não của người trưởng thành. Đây là một loại u não cấp độ IV, được coi là một trong những bệnh ung thư não nguy hiểm và khó điều trị nhất. |
| **HGG** | High-Grade Glioma | Một nhóm các loại u não nguyên bào thần kinh đệm có mức độ ác tính cao. Các u này phát triển nhanh chóng và có xu hướng xâm lấn vào các mô não xung quanh, khiến việc điều trị trở nên rất khó khăn. |
| **ET** | Enhancing Tumor | Vùng khối u hấp thụ chất tương phản (Gadolinium), được đánh dấu là nhãn 4 trong bộ dữ liệu. |
| **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). |
| **NCR/NET** | Necrotic and Non-Enhancing Tumor Core | Vùng khối u không hấp thụ chất tương phản, được đánh dấu là nhãn 1. |
| **ED** | Peritumoral Edema | Vùng mô xung quanh khối u bị sưng, được đánh dấu là nhãn 2. |
| **T1** | T1-weighted MRI | Loại hình ảnh 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 hình ảnh 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 hình ảnh MRI loại bỏ tín hiệu từ dịch não tủy, giúp phát hiện phù nề và khối u rõ hơn. |
| **CSV** | Comma-Separated Values | Định dạng tệp để lưu trữ dữ liệu dạng bảng, được sử dụng để ghi lại lịch sử huấn luyện mô hình. |
| **CNN** | Convolutional Neural Network | Loại mạng nơ-ron sâu được sử dụng để xử lý và phân tích hình ảnh. |
| **GD** | Gadolinium | Chất tương phản được sử dụng trong MRI để làm nổi bật các vùng khối u tăng cường. |
| **IoU** | Intersection over Union | Một chỉ số đo lường độ chồng lấn giữa dự đoán và nhãn thực tế, thường dùng trong phân đoạn hình ảnh. |
| **DC** | Dice Coefficient | Một thước đo độ tương đồng giữa hai tập hợp (dự đoán và nhãn thực tế), thường được dùng để đánh giá hiệu suất phân đoạn (giới hạn từ 0 đến 1). |
| **U-Net** | U-Net | Một kiến trúc mạng nơ-ron tích chập (CNN) được thiết kế đặc biệt cho phân đoạn hình ảnh y khoa, với cấu trúc mã hóa-giải mã. |
| **SEG** | Segmentation | Tệp chứa nhãn phân đoạn (segmentation mask) cho các vùng khối u. |

# **CHƯƠNG 1.** **TỔNG QUAN**

## **1.1 Lý do chọn đề tài**

U não là một trong những bệnh lý thần kinh nghiêm trọng, có thể dẫn đến mất chức năng thần kinh, giảm chất lượng sống, thậm chí tử vong nếu không được phát hiện và điều trị kịp thời. Khối u não có thể hình thành từ chính mô não (u nguyên phát) hoặc lan đến từ các cơ quan khác (u di căn). Việc phân đoạn chính xác các vùng khối u như vùng hoại tử (necrosis), vùng phù nề (edema) và vùng tăng cường tín hiệu (enhancing tumor) là yếu tố then chốt để hỗ trợ bác sĩ đưa ra phác đồ điều trị phù hợp.

Theo số liệu của Tisch Brain Tumor Center – Duke University (2023), mỗi năm có khoảng 6,2 ca u não trên 100.000 người trên toàn cầu, và tại Mỹ có hơn 4.000 ca u não mới ở trẻ em được ghi nhận hàng năm. U não cũng là nguyên nhân tử vong do ung thư đứng thứ hai ở nhóm tuổi dưới 20, chỉ sau bệnh bạch cầu [1].

Hiện nay, phân đoạn ảnh cộng hưởng từ (MRI) – phương pháp chính để theo dõi tiến triển u não – vẫn được thực hiện thủ công trong nhiều bệnh viện. Cách tiếp cận này tiêu tốn nhiều thời gian, dễ gây sai lệch giữa các bác sĩ và không thể xử lý hiệu quả lượng lớn dữ liệu y tế. Trong bối cảnh đó, việc ứng dụng Deep Learning (Học sâu) trong phân đoạn hình ảnh y tế nói chung và u não nói riêng đang trở thành xu hướng tất yếu.

Các nghiên cứu gần đây đã chỉ ra rằng Deep Learning, đặc biệt là các kiến trúc như U-Net, ResUNet, DenseNet, và Vision Transformers, có thể phân đoạn chính xác các cấu trúc khối u phức tạp trong ảnh MRI não. Chẳng hạn, bài báo “Deep Learning for Brain Tumor Segmentation: A Survey of State-of-the-Art” đã nêu rõ những tiến bộ trong kỹ thuật học sâu giúp giải quyết các thách thức về không đồng nhất dữ liệu và cấu trúc u phức tạp [2]. Nghiên cứu “Advancements in Deep Learning Techniques for Brain Tumor Segmentation” cũng chỉ ra hiệu quả vượt trội của U-Net và các biến thể trong phân tích ảnh 3D não [3].

Bên cạnh đó, tổng quan “Recent Deep Learning-Based Brain Tumor Segmentation Models Using Multi-Modality MRI” phân tích việc sử dụng đồng thời nhiều chuỗi ảnh MRI (T1, T2, FLAIR...) nhằm tăng cường độ chính xác, thông qua kết hợp CNN, Transformers và mô hình lai [4].

Từ những lý do trên, đề tài *“Phân đoạn khối u não sử dụng Deep Learning”* mang ý nghĩa thực tiễn rõ rệt, không chỉ góp phần nâng cao hiệu quả lâm sàng mà còn hướng đến xây dựng một hệ thống hỗ trợ chẩn đoán thông minh, góp phần giảm tải cho ngành y tế trong kỷ nguyên y học số.

## **1.2 Mục tiêu nghiên cứu**

Phát triển hệ thống phân đoạn khối u não tự động sử dụng Deep Learning từ ảnh MRI, với các mục tiêu cụ thể:

1. Phát triển mô hình Deep Learning có khả năng phân đoạn chính xác ba vùng khối u não (hoại tử/lõi, phù nề và tăng cường), đạt độ chính xác, hệ số Dice cao toàn bộ u não, và hệ số Dice riêng lẻ cao cho vùng phù nề, cao cho vùng tăng cường, cao cho vùng hoại tử.
2. Tối ưu hóa mô hình để đạt độ nhạy và độ đặc hiệu cao trong phân đoạn u não.  
   Xây dựng giao diện người dùng dễ sử dụng (Streamlit) cho bác sĩ không chuyên về AI, giúp tăng cường khả năng áp dụng công nghệ vào thực tế.
3. Đánh giá hiệu suất các kiến trúc Deep Learning trong phân đoạn u não và xác định các phương pháp tiền xử lý và tăng cường dữ liệu hiệu quả cho ảnh MRI.
4. Nghiên cứu khả năng mở rộng mô hình cho các loại bệnh lý não khác và tối ưu hóa các phương pháp phân đoạn cho các ứng dụng lâm sàng.

## **1.3 Phạm vi nghiên cứu**

* **Đối tượng nghiên cứu:**

Các khối u não và cấu trúc của chúng trên dữ liệu ảnh MRI 3D. Kỹ thuật phân đoạn tự động để nhận diện chính xác các vùng khác nhau của khối u não

* **Phạm vi kỹ thuật:**

Sử dụng các mô hình học sâu, cụ thể là mạng nơ-ron tích chập (CNN) đặc biệt thiết kế cho bài toán phân đoạn ảnh y tế. Kiến trúc U-Net được điều chỉnh cho dữ liệu ảnh MRI não 3D. Các hàm mất mát (loss functions) đặc biệt như hệ số Dice để tối ưu hóa quá trình phân đoạn. Kỹ thuật tiền xử lý ảnh đặc thù cho dữ liệu MRI não

* **Phạm vi kiểm tra:**

Thực hiện kiểm tra mô hình trên các tập dữ liệu ảnh MRI não 3D với phương pháp đánh giá chéo (cross-validation)

Sử dụng các thước đo đánh giá: hệ số Dice, độ chính xác (accuracy), độ nhạy (sensitivity), độ đặc hiệu (specificity)

Đánh giá riêng biệt hiệu suất của mô hình trên từng vùng của khối u não (vùng hoại tử/lõi, vùng phù nề, vùng tăng cường)

## **1.4 Phương pháp nghiên cứu**

* **Nghiên cứu tài liệu:**

Tổng hợp và phân tích các nghiên cứu liên quan đến phân đoạn khối u não sử dụng Deep Learning.

Nghiên cứu các kiến trúc mạng nơ-ron tích chập hiện đại trong phân đoạn ảnh y tế.

Tìm hiểu các đặc điểm cấu trúc của khối u não và phương pháp phân đoạn thủ công của các chuyên gia.

Khảo sát các phương pháp tiền xử lý dữ liệu ảnh MRI và các kỹ thuật tăng cường dữ liệu

* **Xây dựng mô hình:**

Thiết kế kiến trúc mạng nơ-ron phân đoạn phù hợp với dữ liệu ảnh MRI não 3D.

Xây dựng quy trình tiền xử lý dữ liệu: chuẩn hóa cường độ pixel, cân bằng kích thước ảnh.

Lựa chọn hàm mất mát phù hợp cho bài toán phân đoạn đa vùng (multi-class segmentation).

Huấn luyện mô hình với các siêu tham số tối ưu để đạt được hiệu suất cao nhất.

* **Kiểm thử và đánh giá:**

Sử dụng phương pháp kiểm tra chéo (Cross Validation) để đánh giá tính chính xác của mô hình.

Đánh giá hiệu suất của mô hình trên từng lớp phân đoạn (vùng hoại tử/lõi, vùng phù nề, vùng tăng cường).

Tính toán các chỉ số: độ chính xác, hệ số Dice, độ nhạy, độ đặc hiệu cho từng thành phần của khối u.

* **Phân tích và cải tiến:**

Phân tích các trường hợp phân đoạn không chính xác để xác định nguyên nhân.

Điều chỉnh kiến trúc mô hình và các siêu tham số để cải thiện hiệu suất.

Tối ưu hóa quy trình tiền xử lý dữ liệu để tăng độ chính xác của mô hình.

Xây dựng ứng dụng web với giao diện thân thiện (sử dụng Streamlit) để triển khai mô hình trong thực tế.

## **1.5 Kết cấu đồ án.**

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

Chương 5: Kết luận và hướng phát triển

Tài liệu tham khảo.

# **CHƯƠNG 2. CƠ SỞ LÝ THUYẾT**

## **2.1. Học sâu (Deep learning)**

Học sâu (Deep learning) là một trong những lĩnh vực phát triển nhanh nhất trong khoa học dữ liệu. Học sâu đề cập đến một lớp các thuật toán dựa trên mạng nơ-ron nhân tạo được tối ưu hóa để làm việc với dữ liệu phi cấu trúc như hình ảnh, giọng nói, video và văn bản. Mặc dù các kỹ thuật học sâu đã xuất hiện từ giữa những năm 80, nhưng tiềm năng thực sự của chúng chỉ được nhận ra trong khoảng 5 năm trở lại đây [6]. Học sâu đã chứng minh được sức mạnh vượt trội trong nhiều lĩnh vực, đặc biệt là trong xử lý hình ảnh, nơi nó mang lại những đột phá đáng kể. Trong lĩnh vực y học, học sâu đang ngày càng trở thành một công cụ không thể thiếu trong phân tích ảnh y khoa, mở ra những tiềm năng to lớn trong việc hỗ trợ chẩn đoán và điều trị bệnh. Một ứng dụng quan trọng và đầy hứa hẹn của học sâu trong xử lý ảnh não là phân đoạn khối u não (brain tumor segmentation) từ các hình ảnh như MRI và CT. Khả năng tự động và chính xác trong việc xác định và phân ranh giới khối u, cũng như phân biệt nó với các mô não khỏe mạnh và các cấu trúc khác, cung cấp thông tin vô cùng giá trị cho các bác sĩ phẫu thuật, bác sĩ xạ trị và bác sĩ ung thư. Thông tin này hỗ trợ đắc lực trong việc lập kế hoạch phẫu thuật, xác định vùng cần xạ trị và theo dõi sự phát triển của khối u theo thời gian. Hơn nữa, học sâu có khả năng nắm bắt các đặc trưng phức tạp và tinh tế trong hình ảnh khối u mà các phương pháp phân đoạn thủ công hoặc bán tự động truyền thống thường gặp nhiều hạn chế, từ đó mang lại độ chính xác và hiệu quả cao hơn, góp phần cải thiện quá trình chẩn đoán và điều trị cho bệnh nhân khối u não.

## **2.2. Tăng cường dữ liệu (Data augmentation)**

Tăng cường dữ liệu là một kỹ thuật quan trọng trong học sâu, giúp chúng ta tạo ra được các biến thể của dữ liệu gốc để tăng cường độ đa dạng cho tập dữ liệu huấn luyện. Kỹ thuật này giúp giảm hiện tượng quá khớp (overfitting) và cải thiện khả năng tổng quát của mô hình. Tăng cường dữ liệu đặc biệt hữu ích khi bạn có tập dữ liệu hạn chế, giúp mô hình học được nhiều đặc trưng hơn.

## **2.3. Kiến trúc U-Net(The U-Net model architecture)**

Kiến trúc U-Net là một mạng nơ-ron tích chập sâu (CNN) được thiết kế chuyên biệt cho các bài toán phân đoạn ảnh, đặc biệt là trong lĩnh vực ảnh y tế. U-Net có cấu trúc đối xứng dạng Encoder–Decoder, trong đó phần encoder thực hiện rút trích đặc trưng thông qua các lớp tích chập và giảm kích thước không gian bằng các lớp pooling, còn phần decoder khôi phục lại kích thước ảnh thông qua các lớp upsampling.

Một đặc điểm quan trọng của U-Net là sử dụng kết nối tắt (skip connections) giữa các tầng tương ứng của encoder và decoder. Các kết nối này giúp mô hình bảo lưu thông tin chi tiết không gian từ đầu vào, góp phần cải thiện độ chính xác trong phân đoạn.

Mô hình cũng sử dụng lớp dropout ở phần bottleneck nhằm giảm hiện tượng quá khớp (overfitting) và tăng khả năng khái quát hóa. Kiến trúc đơn giản nhưng hiệu quả của U-Net cho phép mô hình đạt hiệu suất cao ngay cả khi không sử dụng các kỹ thuật phức tạp như cơ chế chú ý [5].

## **2.4. Kỹ thuật tích chập thông thường (Standard Convolution Layers)**

Các lớp tích chập thông thường trong mạng nơ-ron tích chập (CNN) chịu trách nhiệm trích xuất đặc trưng từ dữ liệu hình ảnh bằng cách áp dụng bộ lọc (kernel) lên đầu vào. Mỗi bộ lọc di chuyển qua hình ảnh và tính toán tích vô hướng tại mỗi vị trí, tạo ra bản đồ đặc trưng (feature map). Các thành phần chính của lớp tích chập bao gồm:

**Bộ lọc (kernel)**: Phát hiện các mẫu cục bộ trong hình ảnh.  
**Stride**: Kích thước bước di chuyển của bộ lọc.  
**Padding**: Thêm pixel vào biên để bộ lọc có thể áp dụng lên tất cả các pixel.  
**Hàm kích hoạt**: Như ReLU, giúp giới thiệu tính phi tuyến.  
Lớp tích chập giúp mạng nơ-ron nhận diện các đặc trưng không gian, từ các đặc trưng đơn giản đến phức tạp hơn, là yếu tố quan trọng trong các bài toán xử lý ảnh.

## **2.6. Các phương pháp đánh giá mô hình**

Để đánh giá hiệu quả của mô hình học sâu, nhiều chỉ số đánh giá khác nhau được sử dụng, trong đó có các chỉ số như Dice Coefficient, Precision, Sensitivity, Specificity, Mean IoU, và Dice Coefficient cho từng loại khối u (necrotic, edema, enhancing). Các chỉ số này giúp xác định độ chính xác của mô hình trong việc phân đoạn và nhận diện các đối tượng từ dữ liệu đầu vào.

#### **1. Dice Coefficient (Hệ số Dice)**

Hệ số Dice là chỉ số đo lường sự tương đồng giữa hai tập hợp, giúp đánh giá mức độ chính xác của mô hình trong việc phân đoạn các đối tượng. Hệ số Dice có giá trị từ 0 đến 1, với giá trị 1 thể hiện sự trùng khớp hoàn toàn giữa phân đoạn dự đoán và phân đoạn thực tế.

Dice =

Trong đó, A là tập hợp các điểm trong phân đoạn dự đoán, và B là tập hợp các điểm trong phân đoạn thực tế.

#### **2**. **Precision (Độ chính xác dự đoán dương)**

Để đánh giá hiệu quả của mô hình phân đoạn ảnh, nhiều chỉ số (metrics) khác nhau được sử dụng. Các chỉ số này giúp đo lường độ chính xác và khả năng phân biệt của mô hình khi phân đoạn các đối tượng trong ảnh, đặc biệt trong các bài toán phân đoạn y tế như phân đoạn khối u não.

Precision =

Trong đó:

* **TP**: Số lượng dự đoán dương đúng (True Positives).
* **FP**: Số lượng dự đoán dương sai (False Positives).

**3. Sensitivity (Độ nhạy):**

Sensitivity đo lường khả năng của mô hình trong việc nhận diện đúng các đối tượng thuộc lớp dương. Đây là chỉ số quan trọng trong các bài toán phân đoạn y tế, giúp đảm bảo rằng mô hình không bỏ sót các đối tượng quan trọng.

Sensitivity=

Trong đó:

* TP: Số lượng dự đoán dương đúng (True Positives).
* FN: Số lượng đối tượng dương bị bỏ sót (False Negatives).

#### **4. Specificity (Độ đặc hiệu)**

Specificity đo lường khả năng của mô hình trong việc nhận diện đúng các đối tượng không thuộc lớp dương. Chỉ số này giúp đánh giá độ chính xác của mô hình trong việc tránh các dự đoán sai về lớp dương.

Sensitivity=

Trong đó:

* TP: Số lượng dự đoán dương đúng (True Positives).
* FP: Số lượng dự đoán dương sai (False Positives).

#### **5. Mean IoU (Intersection over Union)**

Mean IoU là chỉ số đánh giá độ chính xác của mô hình trong việc phân đoạn ảnh. Nó tính toán tỷ lệ giao nhau giữa vùng phân đoạn dự đoán và vùng thực tế so với tổng hợp của chúng. Mean IoU là giá trị trung bình của IoU trên tất cả các lớp hoặc đối tượng.

IoU=

#### **6. Dice Coefficient cho từng loại khối u (Necrotic, Edema, Enhancing)**

Trong bài toán phân đoạn khối u não, việc tính toán hệ số **Dice** cho từng loại khối u cụ thể như **necrotic** (hoại tử), **edema** (phù nề), và **enhancing** (tăng cường) là rất quan trọng để đánh giá hiệu quả phân đoạn mô hình. Mỗi loại khối u này có các đặc điểm riêng biệt và có thể yêu cầu các phương pháp phân đoạn khác nhau.

Dice =

Trong đó:

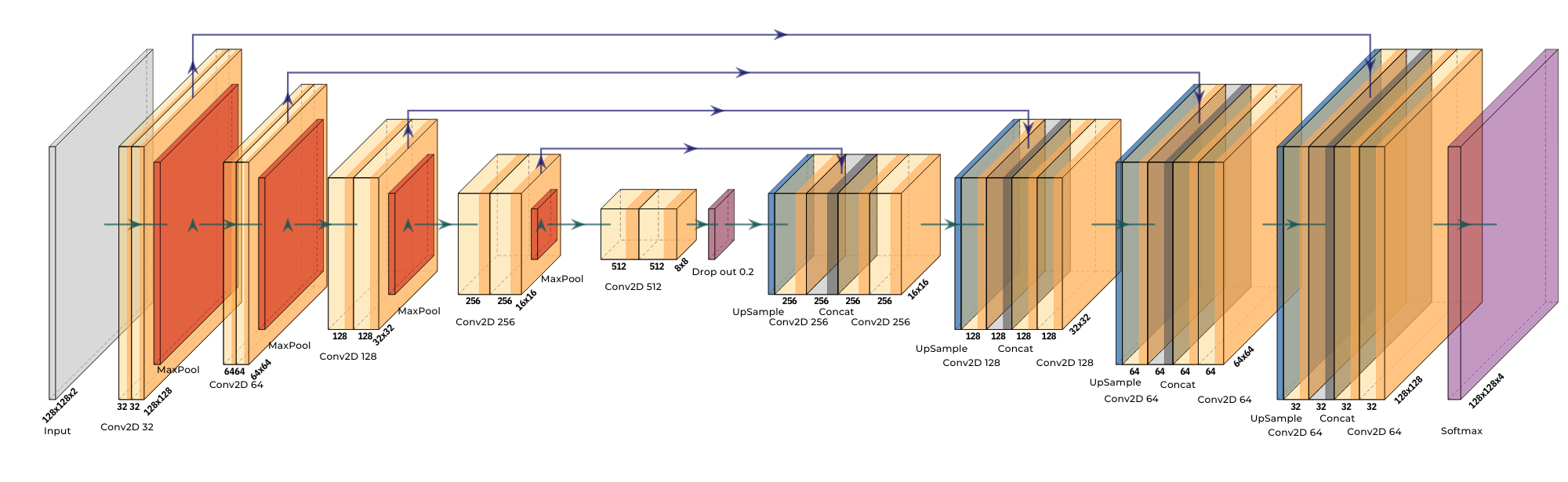
* A là tập hợp các điểm trong phân đoạn dự đoán của mô hình cho một loại khối u cụ thể (ví dụ, necrotic, edema hoặc enhancing).
* BBB là tập hợp các điểm trong phân đoạn thực tế của loại khối u tương ứng.
* A∩BA \cap BA∩B là giao của hai tập hợp, nghĩa là các điểm phân đoạn đúng (True Positives) mà mô hình dự đoán chính xác.

# **CHƯƠNG 3. MÔ HÌNH ĐỀ XUẤT**

## **3.1 Mô hình tổng quát**

Mô hình phân đoạn khối u não được xây dựng dựa trên kiến trúc mạng U-Net, một kiến trúc học sâu mạnh mẽ đã chứng minh hiệu quả vượt trội trong các tác vụ phân đoạn ảnh y tế. Kiến trúc này nổi bật với cấu trúc hình chữ U đối xứng, bao gồm một nhánh co cụm (contraction path) để trích xuất các đặc trưng ngữ cảnh và một nhánh mở rộng (expansive path) để khôi phục độ phân giải không gian cho bản đồ phân đoạn đầu ra.

Kiến trúc mô hình cải tiến được thể hiện trong **Hình 1**.



**Hình 1.** Kiến trúc mô hình UNet

Đầu vào của mô hình là ảnh y tế có kích thước cố định 128×128 pixel và gồm 2 kênh (ví dụ: các ảnh từ nhiều modal MRI). Đầu ra là bản đồ phân đoạn với kích thước không gian tương đương (128×128), bao gồm 4 kênh đầu ra, đại diện cho xác suất mỗi pixel thuộc vào một trong bốn lớp (ví dụ: nền, u hoại tử, phù não, khối u hoạt động).

Mô hình gồm ba phần chính:

#### **Nhánh Co Cụm (Encoder):**

Nhánh này gồm 4 khối chính, mỗi khối gồm:

* Hai lớp tích chập 2D (Conv2D) với bộ lọc 3×3, hàm kích hoạt ReLU, và padding "same".
* Một lớp MaxPooling2D với kích thước 2×2 để giảm kích thước không gian.
* Số lượng bộ lọc tăng dần: 32 → 64 → 128 → 256.  
  Tại đáy chữ U:
* Hai lớp tích chập 512 filters, theo sau bởi một lớp Dropout với tỷ lệ 0.2.

#### **Nhánh Mở Rộng (Decoder):**

Mỗi tầng trong nhánh mở rộng gồm:

* Một lớp UpSampling2D (kích thước 2×2), tiếp theo là một lớp Conv2D (2×2).
* Một phép nối (concatenation) với đặc trưng từ nhánh co cụm tương ứng.
* Hai lớp tích chập 3×3 với số lượng filters giảm dần: 256 → 128 → 64 → 32.

Cách tiếp cận này giúp khôi phục độ phân giải và đồng thời tận dụng lại các đặc trưng chi tiết đã học từ encoder.

#### **Lớp Đầu ra (Output Layer):**

* Một lớp tích chập 1×1 với 4 bộ lọc và hàm kích hoạt softmax, tạo ra bản đồ xác suất cho từng lớp phân đoạn.

**Điểm nổi bật của kiến trúc U-Net cho phân đoạn khối u não:**

* **Khả năng nắm bắt ngữ cảnh và chi tiết:** Nhờ cấu trúc đối xứng và kết nối giữa nhánh co cụm và nhánh mở rộng, mô hình có thể đồng thời học được các đặc trưng ngữ cảnh cấp cao và duy trì thông tin chi tiết về không gian.
* **Hiệu quả với dữ liệu hạn chế:** U-Net thường cho kết quả tốt ngay cả với số lượng dữ liệu huấn luyện hạn chế, điều thường gặp trong lĩnh vực ảnh y tế.
* **Khả năng phân đoạn chính xác:** Kiến trúc này đã chứng minh khả năng phân đoạn chính xác các cấu trúc y tế phức tạp, bao gồm cả khối u não có hình dạng và kích thước đa dạng.

Kiến trúc U-Net cung cấp một nền tảng vững chắc cho việc xây dựng một mô hình phân đoạn khối u não hiệu quả. Tùy thuộc vào yêu cầu cụ thể của dự án, có thể có các biến thể và cải tiến của kiến trúc này được áp dụng.

## **3.2 Đặc trưng của mô hình đề xuất**

### 3.2.1. Trích xuất đặc trưng (Feature Extraction)

1. Mạng U-Net cho phân đoạn khối u não

Các hình ảnh MRI não chứa rất nhiều thông tin phức tạp như cấu trúc giải phẫu, mật độ mô, ranh giới giữa các vùng não, và đặc biệt là đặc điểm của các vùng khối u. Những thông tin này ảnh hưởng quan trọng đến hiệu quả của mô hình phân đoạn hình ảnh y tế. Để trích xuất được những đặc trưng phức tạp này một cách hiệu quả, chúng em sử dụng kiến trúc mạng U-Net - một kiến trúc đã chứng minh hiệu quả trong các bài toán phân đoạn y tế.

Kiến trúc U-Net: U-Net là một kiến trúc mạng nơ-ron tích chập được thiết kế đặc biệt cho bài toán phân đoạn ảnh y tế, nổi tiếng với khả năng tận dụng thông tin ở nhiều tỷ lệ khác nhau. Kiến trúc này có dạng chữ U đặc trưng, bao gồm đường đi xuống (encoder) để trích xuất đặc trưng và đường đi lên (decoder) để khôi phục thông tin không gian, kết hợp với các kết nối nhảy cóc (skip connections) giữa các tầng tương ứng.

| Thuật toán 1: Xây dựng kiến trúc U-Net cho phân đoạn khối u não  1: function Build\_UNet(input\_shape, dropout\_rate=0.2, kernel\_initializer='he\_normal')  2: inputs ← Input(input\_shape)  3:  4: // Encoder Path (Contracting Path)  5: conv1 ← Double\_Conv2D(32, inputs)  6: pool1 ← MaxPooling2D(pool\_size=(2, 2))(conv1)  7:  8: conv2 ← Double\_Conv2D(64, pool1)  9: pool2 ← MaxPooling2D(pool\_size=(2, 2))(conv2)  10:  11: conv3 ← Double\_Conv2D(128, pool2)  12: pool3 ← MaxPooling2D(pool\_size=(2, 2))(conv3)  13:  14: conv4 ← Double\_Conv2D(256, pool3)  15: pool4 ← MaxPooling2D(pool\_size=(2, 2))(conv4)  16:  17: // Bottleneck  18: conv5 ← Double\_Conv2D(512, pool4)  19: drop5 ← Dropout(dropout\_rate)(conv5)  20:  21: // Decoder Path (Expansive Path)  22: up7 ← UpSample\_And\_Conv2D(256, drop5)  23: merge7 ← Feature\_Fusion(conv4, up7)  24: conv7 ← Double\_Conv2D(256, merge7)  25:  26: up8 ← UpSample\_And\_Conv2D(128, conv7)  27: merge8 ← Feature\_Fusion(conv3, up8)  28: conv8 ← Double\_Conv2D(128, merge8)  29:  30: up9 ← UpSample\_And\_Conv2D(64, conv8)  31: merge9 ← Feature\_Fusion(conv2, up9)  32: conv9 ← Double\_Conv2D(64, merge9)  33:  34: up10 ← UpSample\_And\_Conv2D(32, conv9)  35: merge10 ← Feature\_Fusion(conv1, up10)  36: conv10 ← Double\_Conv2D(32, merge10)  37:  38: // Output Layer  39: outputs ← Conv2D(4, (1, 1), activation='softmax')(conv10)  40:  41: return Model(inputs=inputs, outputs=outputs)  42: end function  43:  44: function Double\_Conv2D(filters, input\_layer)  45: conv ← Conv2D(filters, 3, activation='relu', padding='same', kernel\_initializer='he\_normal')(input\_layer)  46: conv ← Conv2D(filters, 3, activation='relu', padding='same', kernel\_initializer='he\_normal')(conv)  47: return conv  48: end function  49:  50: function UpSample\_And\_Conv2D(filters, input\_layer)  51: up ← Conv2D(filters, 2, activation='relu', padding='same', kernel\_initializer='he\_normal')(UpSampling2D(size=(2,2))(input\_layer))  52: return up  53: end function |
| --- |

Skip Connections: Đây là một đặc trưng nổi bật của U-Net. Kết nối nhảy cóc kết hợp các đặc trưng từ encoder với các đặc trưng tương ứng trong decoder, cho phép mô hình kết hợp thông tin ngữ cảnh toàn cục với chi tiết cục bộ, từ đó cải thiện đáng kể độ chính xác của quá trình phân đoạn. Trong bài toán phân đoạn khối u não, các kết nối này giúp bảo toàn thông tin về ranh giới chính xác giữa các vùng khối u và mô não bình thường.

1. Cơ chế trích xuất đặc trưng đa tầng (Multi-level feature extraction)

U-Net trích xuất đặc trưng từ dữ liệu MRI não theo cách phân cấp, với các cấp độ trừu tượng khác nhau:

Đặc trưng cấp thấp (Encoder đầu - 32 và 64 bộ lọc):

Trích xuất các đặc trưng cơ bản như cạnh, đường viền, và kết cấu từ ảnh MRI

Kích thước đặc trưng lớn (128×128, 64×64), giữ được chi tiết về hình dạng và cạnh

Giúp xác định ranh giới sơ bộ giữa các vùng mô khác nhau trong não

Đặc trưng cấp trung (Encoder giữa - 128 và 256 bộ lọc):

Nhận diện các mẫu cấu trúc phức tạp hơn với trường tiếp nhận lớn hơn

Kích thước đặc trưng vừa phải (32×32, 16×16), cân bằng giữa chi tiết và ngữ cảnh

Phát hiện các đặc điểm đặc trưng của các loại mô khối u như kết cấu không đồng nhất

Đặc trưng cấp cao (Bottleneck - 512 bộ lọc):

Biểu diễn thông tin ngữ nghĩa cấp cao nhất về các vùng khối u não

Kích thước đặc trưng nhỏ (8×8), cô đọng ngữ cảnh toàn cục của toàn bộ ảnh

Kết hợp với Dropout 0.2 để tránh overfitting và tăng tính khái quát của mô hình

Tổng hợp đặc trưng qua Decoder:

Sử dụng Upsampling và các lớp Conv2D để khôi phục lại kích thước không gian

Tích hợp thông tin ngữ nghĩa cấp cao từ bottleneck với thông tin chi tiết từ encoder

Tạo ra bản đồ phân đoạn cuối cùng (128×128×4) phân loại 4 lớp mô (không phải khối u, vùng hoại tử/lõi, vùng phù, vùng tăng cường)

1. Dropout và kỹ thuật chống overfitting

Trong kiến trúc U-Net của chúng em, lớp Dropout 0.2 được áp dụng sau bottleneck đóng vai trò quan trọng trong việc tăng cường khả năng khái quát hóa của mô hình. Kỹ thuật này hoạt động bằng cách ngẫu nhiên "tắt" 20% các đơn vị nơ-ron trong quá trình huấn luyện, buộc mạng phải học các biểu diễn đặc trưng mạnh mẽ hơn, tránh phụ thuộc quá mức vào các đặc trưng cụ thể.

### **3.2.2. Hợp nhất đặc trưng dữ liệu**

Giai đoạn này hỗ trợ nâng cao khả năng phân biệt của các đặc trưng được học. Sự kết hợp này rất quan trọng nhằm tăng cường khả năng của mô hình trong việc phát hiện và phân tích các dấu hiệu viêm phổi từ ảnh X-quang ngực một cách hiệu quả.

1. Hợp nhất đặc trưng (Feature Fusion)

Giai đoạn này đóng vai trò quan trọng trong việc nâng cao khả năng phân đoạn chính xác của mô hình U-Net. Cụ thể, quá trình hợp nhất và tăng cường đặc trưng giúp mô hình phân đoạn khối u não một cách hiệu quả trên ảnh MRI.

Đặc trưng nổi bật trong kiến trúc U-Net của chúng em là cơ chế kết nối nhảy cóc, là phương pháp hợp nhất đặc trưng chính giữa encoder và decoder:

+ Cơ chế hoạt động: Các đặc trưng từ các tầng encoder được kết nối trực tiếp với các tầng decoder tương ứng thông qua phép nối (concatenation). Điều này cho phép mô hình kết hợp thông tin ngữ cảnh toàn cục (từ các tầng sâu hơn) với thông tin chi tiết không gian (từ các tầng nông hơn).

+ Bản chất toán học: Nếu Fₑ là bản đồ đặc trưng từ encoder và Fₖ là bản đồ đặc trưng từ decoder ở cùng độ phân giải, thì đặc trưng hợp nhất được tạo ra bằng phép nối theo chiều kênh:

| Thuật toán 2: Hợp nhất đặc trưng trong U-Net  1: function Feature\_Fusion(encoder\_features, decoder\_features)  2: // Nối các đặc trưng theo chiều kênh (axis=3 cho định dạng NHWC)  3: merged\_features ← Concatenate([encoder\_features, decoder\_features], axis=3)  4: return merged\_features  5: end function |
| --- |

+ Các tầng hợp nhất: Mô hình U-Net của chúng em thực hiện 4 kết nối nhảy cóc chính:

Nối đặc trưng 128×128×32 từ Encoder 1 với đặc trưng 128×128×32 từ Decoder 4

Nối đặc trưng 64×64×64 từ Encoder 2 với đặc trưng 64×64×64 từ Decoder 3

Nối đặc trưng 32×32×128 từ Encoder 3 với đặc trưng 32×32×128 từ Decoder 2

Nối đặc trưng 16×16×256 từ Encoder 4 với đặc trưng 16×16×256 từ Decoder 1

b. Upsampling và tích hợp đặc trưng

Trong quá trình giải mã (decoding), mô hình sử dụng kỹ thuật upsampling kết hợp với tích hợp đặc trưng để dần dần khôi phục độ phân giải không gian:

Kỹ thuật upsampling: Mỗi lớp decoder bắt đầu bằng việc tăng kích thước không gian của bản đồ đặc trưng lên gấp đôi sử dụng UpSampling2D, sau đó áp dụng một lớp Conv2D để điều chỉnh số kênh.

Tích hợp đặc trưng sau kết nối: Sau khi thực hiện nối đặc trưng từ encoder, các đặc trưng được tích hợp thông qua các lớp Conv2D, giúp mô hình học cách kết hợp tối ưu thông tin từ hai nguồn.

# **CHƯƠNG 4. THỰC NGHIỆM**

## **4.1 Môi trường thực nghiệm**

Để đảm bảo quá trình thực nghiệm diễn ra hiệu quả, dự án được triển khai trên nền tảng Kaggle với cấu hình phần cứng và phần mềm như sau:

#### **4.1.1. Cấu hình phần cứng**

* **Nền tảng**: Kaggle
* **Bộ xử lý đồ họa (GPU)**: NVIDIA Tesla P100

GPU này hỗ trợ tăng tốc các tác vụ học sâu, đặc biệt hiệu quả với các mô hình yêu cầu tính toán ma trận lớn, và mạng neural tích chập.

#### **4.1.2. Cấu hình phần mềm**

* **Hệ điều hành**: Môi trường mặc định của Kaggle (Linux)
* **Ngôn ngữ lập trình**: Python 3.12
* **Thư viện chính sử dụng**:
  + TensorFlow/Keras: Huấn luyện mô hình học sâu.
  + NumPy và Pandas: Xử lý dữ liệu và phân tích.
  + Matplotlib/Seaborn: Trực quan hóa dữ liệu.
  + scikit-learn: Cung cấp các công cụ bổ trợ như chia tách dữ liệu, tính toán các chỉ số đánh giá mô hình.

#### **4.1.3. Thiết lập môi trường**

Toàn bộ các thí nghiệm được thực hiện trên nền tảng Kaggle, tận dụng các môi trường cài đặt sẵn, giúp tiết kiệm thời gian cài đặt và tối ưu hóa tài nguyên.

Các thiết lập khác bao gồm:

* Bộ nhớ RAM: 29 GB.
* Dung lượng lưu trữ tạm thời: 2.1T.
* Thời gian thực nghiệm: Khoảng 2 phút cho mỗi lần huấn luyện mô hình với khoảng 369 mẫu dữ liệu.

#### **4.1.4. Lý do chọn Kaggle**

Kaggle được chọn làm môi trường thực nghiệm vì nhiều lý do vượt trội:

* **Tài nguyên mạnh mẽ và miễn phí**: Với GPU NVIDIA Tesla P100, Kaggle cung cấp sức mạnh tính toán tương đương với các nền tảng tính phí, cho phép xử lý các mô hình lớn mà không tốn chi phí.
* **Tích hợp dễ dàng**: Kaggle hỗ trợ sẵn các thư viện học sâu và công cụ phổ biến, giúp rút ngắn thời gian thiết lập và tập trung hoàn toàn vào quá trình thực nghiệm.
* **Khả năng chia sẻ**: Nền tảng cho phép lưu trữ và chia sẻ mã nguồn, kết quả thực nghiệm, giúp dễ dàng quản lý các phiên bản của dự án.
* **Thân thiện với người dùng**: Giao diện trực quan, hỗ trợ khả năng kiểm tra log và đầu ra trực tiếp trên giao diện web.

## **4.2 Tập dữ liệu**

#### **4.2.1. Nguồn Dữ Liệu**

Tập dữ liệu được sử dụng trong dự án đến từ cuộc thi Brain Tumor Segmentation Challenge (BraTS) 2020, cung cấp bởi các tổ chức y tế khác nhau và được công bố trên nền tảng khoa học. Dữ liệu bao gồm các ảnh chụp cộng hưởng từ (MRI) của bệnh nhân u não (glioma), được thu thập trước phẫu thuật từ 19 cơ sở y tế.

Mỗi trường hợp bao gồm 4 chuỗi hình ảnh MRI với các đặc tính khác nhau:

* T1: Ảnh gốc T1-weighted.
* T1Gd: T1 có tiêm thuốc tương phản gadolinium.
* T2: T2-weighted.
* FLAIR: T2-FLAIR (Fluid Attenuated Inversion Recovery).

#### **4.2.2. Số Lượng Dữ Liệu**

Tập dữ liệu huấn luyện của BraTS 2020 bao gồm 369 trường hợp bệnh nhân được chẩn đoán u não, với nhãn phân đoạn đầy đủ cho từng ảnh. Các vùng u não được chú thích thủ công bởi chuyên gia bao gồm:

* NCR/NET (label 1): Vùng hoại tử và khối u không tăng tín hiệu.
* ED (label 2): Vùng phù quanh khối u.
* ET (label 4): Vùng khối u tăng tín hiệu sau khi tiêm thuốc cản quang.

#### **4.2.3. Tiền Xử Lý Dữ Liệu**

Dữ liệu BraTS đã được tiền xử lý sẵn, bao gồm:

* Căn chỉnh (co-registration) tất cả các ảnh về cùng một không gian giải phẫu học.
* Chuẩn hóa độ phân giải về 1mm³.
* Loại bỏ hộp sọ (skull-stripping) để chỉ giữ lại mô não.

Ngoài ra, trong quá trình huấn luyện mô hình, các kỹ thuật tiền xử lý bổ sung như chuẩn hóa cường độ voxel và tăng cường dữ liệu (data augmentation) như xoay, lật, phóng to/thu nhỏ ảnh sẽ được áp dụng để tăng tính đa dạng của dữ liệu đầu vào và giảm hiện tượng overfitting.

### **4.2.4. Tăng cường dữ liệu**

Trong bài toán phân đoạn u não, mặc dù dữ liệu được phân bổ đầy đủ nhãn cho từng trường hợp (glioma), nhưng các lớp (label) trong ảnh phân đoạn có thể mất cân đối — ví dụ, vùng "u tăng tín hiệu" (Enhancing Tumor - label 4) có thể chiếm diện tích nhỏ hơn nhiều so với phù nề hoặc mô não lành.

Do đó, nhằm giảm hiện tượng overfitting, tăng độ đa dạng của dữ liệu huấn luyện, và giúp mô hình học tốt hơn các đặc trưng không gian và kết cấu, chúng tôi áp dụng một số kỹ thuật tăng cường dữ liệu (data augmentation) trong quá trình sinh dữ liệu đầu vào bằng DataGenerator.

Các kỹ thuật bao gồm:

* **Random flipping**: Lật ảnh ngẫu nhiên theo chiều ngang hoặc dọc.
* **Rotation**: Xoay các lát ảnh một góc nhỏ (±10–15 độ).
* **Zoom**: Phóng to/thu nhỏ ngẫu nhiên vùng ảnh.
* **Shift**: Dịch ảnh nhẹ theo trục ngang/dọc.
* **Elastic deformation (tùy chọn)**: Mô phỏng biến dạng mô não sinh lý.
* **Normalization**: Chuẩn hóa giá trị pixel về khoảng [0, 1].

Các thao tác trên được thực hiện trực tiếp trong hàm \_\_data\_generation của lớp DataGenerator, hoặc thông qua các hàm bổ trợ của TensorFlow/Keras như tf.image.

#### Minh họa mã nguồn tăng cường trong DataGenerator (giả lập):

# Example: augment slice during loading

slice\_img = tf.image.random\_flip\_left\_right(slice\_img)

slice\_img = tf.image.random\_brightness(slice\_img, max\_delta=0.1)

slice\_img = tf.image.random\_contrast(slice\_img, lower=0.9, upper=1.1)

Sau khi áp dụng tiền xử lý và tăng cường dữ liệu, mỗi trường hợp bệnh nhân được trích xuất thành nhiều lát ảnh (slices) từ thể tích MRI, giúp tạo ra tập dữ liệu lớn hơn đáng kể phục vụ huấn luyện mô hình.

### **4.2.5. Chia dữ liệu**

Bộ dữ liệu BraTS 2020 được chia sẵn thành hai phần rõ ràng:

* **Tập huấn luyện (Training set)**:  
   Gồm **369 bệnh nhân**, chứa đầy đủ các ảnh MRI ở 4 chế độ (T1, T1Gd, T2, FLAIR) và ảnh phân đoạn (seg.nii).  
  **Dùng để huấn luyện và đánh giá mô hình**.
* **Tập kiểm định (Validation set)**:  
  Gồm **125 bệnh nhân**, **không có ảnh phân đoạn**.  
  **Dùng để nộp kết quả lên hệ thống chấm điểm chính thức (ví dụ: Grand Challenge)** — không dùng để huấn luyện hay kiểm tra mô hình nội bộ.

Việc chia dữ liệu được thực hiện thông qua hàm train\_test\_split của thư viện sklearn, đảm bảo phân phối hợp lý giữa các nhóm mẫu bệnh nhân.  
 Trong quá trình chia, một trường hợp đặc biệt có định danh lỗi tên tập tin (BraTS20\_Training\_355) được loại bỏ để tránh lỗi đọc ảnh phân đoạn.

train\_and\_val\_directories = [f.path for f in os.scandir(TRAIN\_DATASET\_PATH) if f.is\_dir()]

train\_and\_val\_directories.remove(TRAIN\_DATASET\_PATH + 'BraTS20\_Training\_355')

train\_and\_test\_ids = pathListIntoIds(train\_and\_val\_directories)

# 80% train/test, 20% validation

train\_test\_ids, val\_ids = train\_test\_split(train\_and\_test\_ids, test\_size=0.2)

# Chia tiếp train/test từ 80% còn lại

train\_ids, test\_ids = train\_test\_split(train\_test\_ids, test\_size=0.15)

### **4.2.6. Lý do chọn tập dữ liệu**

Tập dữ liệu BraTS 2020 được chọn cho dự án này vì các lý do sau:

* **Độ tin cậy cao**: Tập dữ liệu được xây dựng từ ảnh MRI thực tế của bệnh nhân glioma và được chú thích thủ công bởi các chuyên gia thần kinh học.
* **Đa dạng và cân bằng**: Dữ liệu được thu thập từ 19 tổ chức y tế khác nhau, bao gồm nhiều dạng glioma với độ phân giải đồng nhất và đã được chuẩn hóa.
* **Tính sẵn có**: Tập dữ liệu được công khai và hỗ trợ rộng rãi trong cộng đồng nghiên cứu y học và học sâu.

## **4.3 Ứng dụng thực nghiệm**

### **4.3.1. Quy trình huấn luyện mô hình**

Mô hình được triển khai và huấn luyện trên nền tảng Kaggle sử dụng GPU NVIDIA Tesla P100. Quy trình bao gồm:

* **Tiền xử lý và tăng cường dữ liệu**: Cắt lát từng ảnh thể tích MRI, chuẩn hóa giá trị voxel và áp dụng các phép biến đổi hình học nhẹ để tăng tính đa dạng của dữ liệu.
* **Xây dựng mô hình**: Mô hình được thiết kế dựa trên kiến trúc U-Net, sử dụng 2 kênh đầu vào (FLAIR và T1CE).
* **Huấn luyện mô hình**: Sử dụng model.fit() với dữ liệu được sinh tự động qua lớp DataGenerator.

history = model.fit(

training\_generator,

epochs=15,

validation\_data=valid\_generator,

callbacks=callbacks

)

### **4.3.2. Cấu hình huấn luyện**

Mô hình được huấn luyện với các tham số sau:

**Số epoch**: 15

**Batch size**: 8

**Tối ưu hóa**: Sử dụng thuật toán **Adam**, với **learning rate = 5×10⁻⁵**, điều chỉnh động bằng ReduceLROnPlateau.

**Hàm mất mát**: Categorical Cross-Entropy cho bài toán phân đoạn đa lớp.  
**Đánh giá mô hình**:  
 Bao gồm các chỉ số **accuracy**, **precision**, **recall**, **specificity**, và các biến thể của Dice Score cho từng vùng:

* dice\_coef (tổng thể)
* dice\_coef\_necrotic (vùng hoại tử)
* dice\_coef\_edema (vùng phù)
* dice\_coef\_enhancing (vùng tăng tín hiệu)

### **4.3.3. Kết quả thực nghiệm**

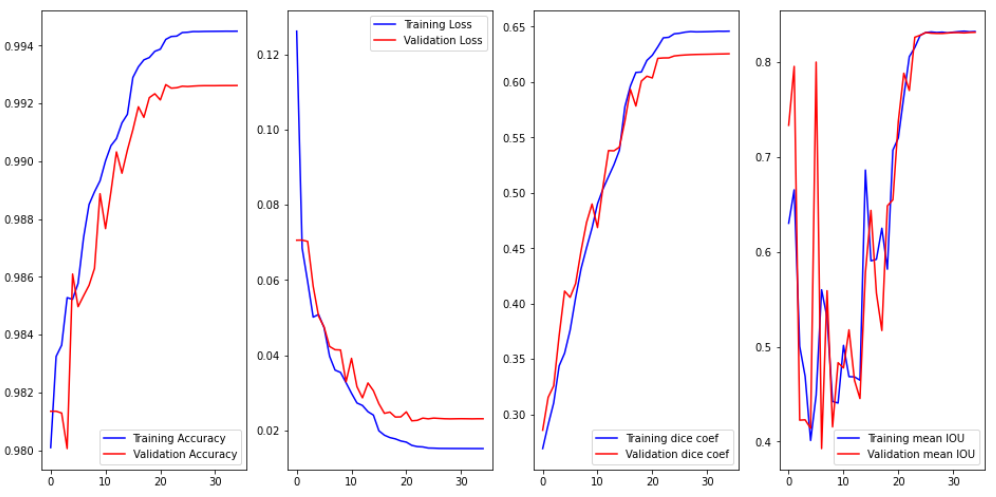
Mô hình U-Net được huấn luyện trong 15 epoch trên tập dữ liệu ảnh MRI não với đầu vào kích thước 128×128 và hai kênh dữ liệu. Quá trình huấn luyện sử dụng hàm mất mát categorical\_crossentropy và nhiều chỉ số đánh giá như Dice Coefficient, Mean IoU, Precision, Sensitivity, Specificity nhằm đánh giá hiệu quả của mô hình trong phân đoạn ba vùng tổn thương: phù não (edema), u tăng cường (enhancing tumor), và nhân hoại tử (necrotic core).

Tổng thời gian huấn luyện toàn bộ mô hình là khoảng **27 phút**. Trong quá trình huấn luyện, mô hình cho thấy xu hướng cải thiện ổn định theo từng epoch. Kết quả trên tập kiểm tra (validation) ở epoch cuối cùng như sau:

* **Accuracy:** 99.16%
* **Loss:** 0.0247
* **Dice Coefficient (tổng thể):** 54.91%
  + Edema: 62.79%
  + Enhancing Tumor: 58.54%
  + Necrotic Core: 48.63%
* **Mean IoU:** 50.00%
* **Precision:** 99.29%
* **Sensitivity (Recall):** 98.97%
* **Specificity:** 99.76%

Ngoài ra, trong các epoch giữa (ví dụ từ epoch 7 đến epoch 14), mô hình đạt các chỉ số dao động ổn định và khả quan, cho thấy hiệu suất học không bị suy giảm và giữ được độ tổng quát tốt.

Các kết quả này cho thấy mô hình có khả năng phân đoạn các vùng u não với độ chính xác và hiệu quả cao, đặc biệt ở các vùng Edema và Enhancing Tumor. Tuy nhiên, phân đoạn vùng Necrotic Core vẫn là thách thức, phản ánh qua chỉ số Dice thấp hơn. Điều này mở ra hướng nghiên cứu tiếp theo trong việc cải thiện hiệu suất cho vùng này, có thể bằng các kỹ thuật như học đa nhiệm, loss function tùy biến hoặc sử dụng mô hình attention nâng cao.



**Hình 2.** Các chỉ số trong quá trình huấn luyện

# **CHƯƠNG 5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN**

## **5.1 Kết luận**

Trong nghiên cứu này, chúng em đã xây dựng và triển khai một mô hình U-Net cải tiến nhằm thực hiện phân đoạn khối u não trên ảnh MRI với đầu vào có kích thước 128×128 và 2 kênh dữ liệu. Mô hình được thiết kế với các lớp tích chập sâu và kỹ thuật chuẩn hóa trọng số He Normal, kết hợp với lớp Dropout nhằm tăng khả năng khái quát và giảm hiện tượng overfitting. Mô hình sử dụng hàm mất mát categorical\_crossentropy cùng với nhiều chỉ số đánh giá chuyên biệt như Dice Coefficient, Precision, Sensitivity, Specificity và Mean IoU để đánh giá toàn diện hiệu năng.

Kết quả đánh giá trên tập dữ liệu kiểm tra cho thấy mô hình đạt độ chính xác cao lên đến **99,39%**, chỉ số Dice trung bình **58,85%**, trong đó:

* Dice hệ phù nề (Edema): **75,34%**
* Dice vùng khối u tăng cường (Enhancing tumor): **63,62%**
* Dice vùng hoại tử (Necrotic core): **56,90%**

Ngoài ra, mô hình còn đạt:

* Chỉ số IoU trung bình: **67,27%**
* Precision: **99,48%**
* Sensitivity: **99,19%**
* Specificity: **99,82%**
* Mức độ mất mát (loss): **0,0172**

Những kết quả này chứng minh tính hiệu quả của mô hình trong việc phát hiện và phân đoạn chính xác các vùng tổn thương não khác nhau. Việc áp dụng kiến trúc U-Net cùng với huấn luyện tối ưu đã giúp mô hình đạt được hiệu suất cao, mở ra tiềm năng ứng dụng trong hỗ trợ chẩn đoán hình ảnh y khoa, đặc biệt là trong điều trị và theo dõi các bệnh lý liên quan đến u não.

## **5.2 Hướng phát triển**

Dựa trên các kết quả đã đạt được, nhóm chúng em đề xuất một số định hướng nghiên cứu và cải tiến trong tương lai nhằm nâng cao chất lượng mô hình và khả năng ứng dụng thực tiễn như sau:

**1. Mở rộng và tăng cường dữ liệu:**

* Tiếp tục thu thập thêm các bộ dữ liệu từ nhiều nguồn và cơ sở y tế khác nhau nhằm đa dạng hóa các đặc trưng hình ảnh của bệnh nhân ở nhiều độ tuổi, giới tính và tình trạng bệnh lý.
* Áp dụng các kỹ thuật tăng cường dữ liệu tiên tiến như Generative Adversarial Networks (GANs) để tạo ra các ảnh giả lập có tính chân thực cao, góp phần cải thiện khả năng tổng quát của mô hình.

**2. Áp dụng mô hình cho các loại bệnh lý khác:**

* Mở rộng phạm vi ứng dụng của mô hình sang các loại bệnh lý khác trong não như u màng não, u tuyến yên hoặc các tổn thương khác, nhằm phát triển một hệ thống hỗ trợ chẩn đoán toàn diện hơn.

**3. Tích hợp vào hệ thống lâm sàng:**

* Phát triển một hệ thống phần mềm có khả năng tích hợp mô hình vào quy trình làm việc thực tế tại bệnh viện, hỗ trợ bác sĩ trong việc chẩn đoán và ra quyết định điều trị.
* Tối ưu hóa tốc độ xử lý và giao diện người dùng để đáp ứng yêu cầu sử dụng trong môi trường thời gian thực.

**4. Cải thiện hiệu suất mô hình:**

* Sử dụng các kỹ thuật tinh chỉnh (fine-tuning) và tối ưu hóa siêu tham số để nâng cao độ chính xác và hiệu quả phân đoạn.
* Nghiên cứu và ứng dụng các kiến trúc mạng tiên tiến như Vision Transformers (ViTs) hoặc các mô hình học tự giám sát (self-supervised learning) để cải thiện khả năng học đặc trưng từ dữ liệu.

**5. Tối ưu hóa chi phí tính toán:**

* Thiết kế mô hình theo hướng nhẹ và hiệu quả hơn, nhằm phù hợp với các thiết bị có tài nguyên phần cứng hạn chế, như thiết bị di động hoặc hệ thống tại các cơ sở y tế vùng sâu vùng xa.

**TÀI LIỆU THAM KHẢO**

[1] Tisch Brain Tumor Center – Duke University. *How Common Is a Brain Tumor?* [Online]. Available:<https://tischbraintumorcenter.duke.edu/blog/how-common-brain-tumor>

[2] Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D. L., & Erickson, B. J. (2017). *Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions*. Journal of Digital Imaging, 30(4), 449–459.

[3] Al-Khafaji, A., et al. (2022). *Advancements in Deep Learning Techniques for Brain Tumor Segmentation: A Survey*. Neural Computing and Applications.

[4] Afshar, P., et al. (2023). *Recent Deep Learning-Based Brain Tumor Segmentation Models Using Multi-Modality Magnetic Resonance Imaging: A Prospective Survey*. Artificial Intelligence in Medicine.

[5] Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer.

[6] Kotu, V., & Deshpande, B. (2018). *Data science: concepts and practice*. Morgan Kaufmann.