Tạp chí Khoa học Đại học Cần Thơ
Số chuyên đề: Khoa học tự nhiên Đồng bằng sông Cửu Long
ISSN 1859-2333 | e-ISSN 2815-5599

DOI:10.22144/ctujos.2026.185

PHÂN ĐOẠN ẢNH CỘNG HƯỞNG TỪ U NÃO SỬ DỤNG THUẬT TOÁN RES-UNET CẢI TIẾN

Lã Hoàng Quý và Phạm Toàn Định*
Trường Công nghệ Văn Lang, Trường Đại học Văn Lang, Việt Nam
**Tác giả liên hệ (Corresponding author): dinh.pt@vlu.edu.vn*

Thông tin chung (Article Information)

Nhận bài (Received): 08/05/2026
Sửa bài (Revised): 02/06/2026
Duyệt đăng (Accepted): 25/08/2026

Title: *Brain tumor mri segmentation using an improved Res-UNet algorithm*

Author(s): *La Hoang Quy and Pham Toan Dinh**

Affiliation(s) *Van Lang School of Technology, Van Lang University, Viet Nam*

TÓM TẮT

Trong nghiên cứu này, một kiến trúc mạng học sâu Res-UNet cải tiến, tích hợp quá trình học phần dư, cơ chế chú ý và tích chập giãn đã được đề xuất nhằm nâng cao khả năng bảo toàn biên và biểu diễn các đặc trưng của ảnh. Mô hình đề xuất được huấn luyện và đánh giá trên bộ dữ liệu MRI gồm 3.929 ảnh. Kết quả thực nghiệm cho thấy phương pháp đề xuất đạt quả phân đoạn tốt hơn so với một số thuật toán khác bao gồm U-Net truyền thống, Fuzzy C-Means (FCM), Attention U-Net, UNet++, SegNet, Swin-UNet, TransUNet và nnUNet, với BF = 98,0%, SSIM = 99,2%, ACC = 99,7%, Dice coefficient = 86,6% và IoU = 76,6%. Ngoài ra, kết quả nghiên cứu cũng đã giúp phân loại sự hiện diện của khối u dựa trên nhãn phân đoạn đạt độ chính xác 99,7% trên tập kiểm tra. Những kết quả này khẳng định hiệu quả của mô hình đề xuất và các ứng dụng trong phân tích ảnh y tế.

Từ khóa: *Ảnh MRI, chẩn đoán, phân đoạn, Res-UNet, u não*

ABSTRACT

This paper presents a Res-UNet-based deep learning architecture that incorporates residual learning, an attention mechanism, and dilated convolutions to improve boundary preservation and feature representation in medical images. The model was trained and evaluated using an MRI dataset containing 3,929 images. Experimental results indicate that the proposed approach provides better segmentation performance compared with several existing methods, including traditional U-Net, Fuzzy C-Means (FCM), Attention U-Net, UNet++, SegNet, Swin-UNet, TransUNet, and nnUNet. The model achieved BF = 98.0%, SSIM = 99.2%, ACC = 99.7%, Dice coefficient = 86.6%, and IoU = 76.6%. Furthermore, tumor presence classification based on the generated segmentation labels reached an accuracy of 99.7% on the test set. These findings suggest that the proposed architecture is a promising approach for medical image analysis.

Keywords: *Brain tumour, diagnosis, MRI images, Res-UNet, segmentation*

1. GIỚI THIỆU

Một trong những thách thức lớn nhất của hệ thống chăm sóc sức khỏe hiện nay là tính phức tạp và quá trình chẩn đoán chính xác các bệnh mới cũng như các tình trạng y khoa khác thông qua hình ảnh chẩn đoán, vốn tốn nhiều thời gian và nguồn lực (Pasche & Kietzmann, 2020). Trí tuệ nhân tạo (AI) đã được đưa vào lĩnh vực y tế thông qua các hệ thống chẩn đoán hỗ trợ bằng máy tính nhằm hỗ trợ bác sĩ phát hiện sớm các dấu hiệu và triệu chứng bệnh, từ đó giảm chi phí điều trị thông qua kế hoạch điều trị hiệu quả hơn (Paschen & Kietzmann, 2020). U não là một trong những tình trạng y khoa nghiêm trọng nhất, đặt ra nhiều thách thức cho việc chẩn đoán kịp thời và điều trị hiệu quả. Ảnh cộng hưởng từ (Magnetic resonance imaging-MRI) đóng vai trò then chốt trong việc phát hiện và theo dõi các bất thường ở não, cung cấp hình ảnh độ phân giải cao và không xâm lấn về cấu trúc cũng như các tổn thương mô não. Tuy nhiên, việc diễn giải ảnh MRI vẫn là một quá trình tốn nhiều công sức và mang tính chủ quan, đòi hỏi các bác sĩ chẩn đoán hình ảnh phải phân tích tỉ mỉ khối lượng dữ liệu lớn, điều này không chỉ tạo ra sự khác biệt trong chẩn đoán giữa các chuyên gia mà còn hạn chế khả năng mở rộng trong các hệ thống y tế thiếu nguồn lực. Trong 20 năm qua, tỷ lệ phát hiện u não đã tăng hơn 40% ở người trưởng thành thuộc mọi độ tuổi (Farmanfarma et al., 2019), với khoảng 11 ca mới được ghi nhận mỗi ngày. Nếu khối u được phát hiện sớm hơn ít nhất một năm, tỷ lệ sống sót của bệnh nhân có thể tăng lên khoảng 40%; ngược lại, nếu được chẩn đoán ở năm thứ 5 trở đi, tỷ lệ sống chỉ còn khoảng 10% (McKinney, 2004). Trước khi AI được ứng dụng trong hệ thống phân tích chuyên sâu, phương pháp phân đoạn thủ công thường được sử dụng để xác định bất thường trên não, do các bác sĩ chuyên khoa hoặc bác sĩ chẩn đoán hình ảnh thực hiện (Starmans et al., 2020), theo từng lát cắt hoặc trong không gian 3D bằng cách chú thích các voxel hoặc khoanh vùng ROI. Mặc dù tận dụng được kiến thức chuyên môn nhưng phương pháp này rất tốn thời gian và dễ xảy ra sai lệch giữa các lần đánh giá của cùng một người hoặc giữa các chuyên gia khác nhau, đồng thời không khả thi khi số lượng ảnh MRI cần xử lý ngày càng lớn. AI phát triển mạnh mẽ cho những công cụ mang tính cách mạng trong lĩnh vực chẩn đoán hình ảnh y khoa, với khả năng tự động hóa và nâng cao quy trình chẩn đoán thông qua các kỹ thuật như học sâu và mạng nơ-ron tích chập, cho phép hệ thống học các mẫu đặc trưng phức tạp trong ảnh MRI để phát hiện, phân loại và phân đoạn u não với độ chính xác và tính nhất quán cao. Kết quả một nghiên cứu trên Nature Medicine cho thấy AI có thể chẩn đoán chính xác u não chỉ trong hai phút trong quá trình phẫu thuật (Saeedi et al., 2023). Ngoài ra, AI còn giúp giảm 25% thời gian nằm viện và giảm 91% số ca chuyển đến cơ sở điều dưỡng (Humphreys et al., 2025), đồng thời hỗ trợ nâng cao độ chính xác phẫu thuật bằng cách phân biệt rõ mô u và mô lành, giúp bác sĩ loại bỏ tối đa khối u và cải thiện tiên lượng sống cho bệnh nhân (Khalighi et al., 2024).

Những tiến bộ gần đây trong AI và học sâu đã cải thiện đáng kể độ chính xác và hiệu quả của phân tích ảnh y khoa, đặc biệt trong việc phát hiện và phân đoạn khối u não bằng MRI. Kết quả nhiều nghiên cứu đã khám phá các mô hình dựa trên AI như mạng nơ-ron tích chập (CNN), Máy vector hỗ trợ (SVM) và các mạng nơ-ron truyền thống. Chẳng hạn, các mô hình do Amin et al. (2018) và Abdalla and Esmail (2018) đề xuất đạt độ chính xác cao (lên đến 99%) nhưng được huấn luyện và kiểm thử trên các tập dữ liệu tương đối nhỏ (thường dưới 100 ảnh), làm hạn chế khả năng tổng quát hóa và độ tin cậy trong các bối cảnh lâm sàng thực tế. Chẳng hạn, Vo and Pham (2021) đã đề xuất một phương pháp phân đoạn ảnh MRI của bệnh nhân Covid-19 sử dụng thuật toán phân cụm mờ dựa trên hạt nhân tự động thích nghi. Phương pháp này không chỉ tự động xác định số lượng cụm phù hợp mà còn cải thiện quá trình hình thành cụm mờ bằng cách nâng cấp kỹ thuật phân cụm mờ dựa trên kernel ban đầu. Mặc dù kết quả nghiên cứu này đã đóng góp đáng kể cho học máy không giám sát trong phân đoạn ảnh nhưng nó vẫn tồn tại một số hạn chế, bao gồm độ phức tạp tính toán cao và thời gian phân đoạn dài. Các kiến trúc hiện đại hơn như U-Net và các biến thể của nó hiện đã được sử dụng rộng rãi trong phân đoạn ảnh y khoa nhờ khả năng mạnh mẽ trong việc trích xuất đặc trưng không gian. Tuy nhiên, các mô hình U-Net truyền thống vẫn gặp khó khăn khi xử lý các đặc trưng sâu và các mẫu phức tạp, đặc biệt trong các tập dữ liệu lớn. Để khắc phục những hạn chế này, các nhà nghiên cứu đã phát triển nhiều mở rộng vượt ra ngoài U-Net cổ điển. Chẳng hạn, Wu et al. (2024) đã giới thiệu MedSegDiff-V2, một khung mô hình khuếch tán dựa trên bộ biến đổi mới, tích hợp các mô hình xác suất khuếch tán với bộ biến đổi thị giác. Mặc dù các mô hình khuếch tán dựa trên U-Net đã cho thấy hiệu năng mạnh mẽ nhưng việc kết hợp trực tiếp với transformer ban đầu lại cho kết quả chưa tối ưu. MedSegDiff-V2 đã giải quyết thành công hạn chế này, đạt được kết quả vượt trội trên 20 bài toán phân đoạn ảnh y khoa với nhiều loại dữ liệu hình ảnh khác nhau. Ngoài ra, Velastegui et al.

(2024) đã thực hiện một khảo sát toàn diện bao quát các kiến trúc phân đoạn, từ các mô hình cổ điển như FCN và PSPNet đến các phương pháp hiện đại dựa trên bộ biến đổi như SegFormer và FAN. Trong cùng hướng nghiên cứu, Ruan and Xiang (2025) đã đề xuất VM-UNet, một mô hình phân đoạn ảnh y khoa đầu tiên được xây dựng hoàn toàn dựa trên các mô hình không gian trạng thái (SSMs). Bằng cách tích hợp khối Visual State Space (VSS), VM-UNet có thể nắm bắt hiệu quả các phụ thuộc dài hạn với độ phức tạp tính toán tuyến tính, khắc phục hạn chế về khả năng mô hình hóa toàn cục của CNN và nhu cầu tài nguyên cao của transformer. Kết quả thực nghiệm trên các tập dữ liệu như ISIC17, ISIC18 và Synapse cho thấy hiệu năng cạnh tranh, nhấn mạnh SSMs là một hướng nghiên cứu đầy tiềm năng cho phân đoạn ảnh y khoa trong tương lai.

Dựa vào các hạn chế của các hướng nghiên cứu trên, chúng tôi đề nghị mô hình kết hợp mới, tập trung vào việc tích hợp các kết nối dư vào kiến trúc U-Net nhằm cải thiện khả năng lan truyền đặc trưng, tăng cường Gradient và nâng cao độ chính xác phân đoạn.

Những đóng góp chính của nghiên cứu này như sau:

(1) Một kiến trúc Res-UNet cải tiến, tích hợp các khối dư, cơ chế chú ý và tích chập giãn đã được đề xuất nhằm nâng cao khả năng xác định ranh giới khối u trong ảnh MRI não. Mô hình đề xuất sử dụng các khối residual kết hợp dilated convolution để mở rộng receptive field và khai thác thông tin đa tỷ lệ mà vẫn duy trì độ phân giải không gian của đặc trưng. Đồng thời, cơ chế attention được tích hợp trên các kết nối skip connection nhằm tăng cường khả năng tập trung vào vùng phổi và giảm ảnh hưởng của các cấu trúc nền không liên quan. Sự kết hợp đồng thời giữa residual learning, attention mechanism và dilated convolution tạo nên một kiến trúc có khả năng khai thác hiệu quả cả thông tin cục bộ lẫn ngữ cảnh toàn cục.

(2) Việc đánh giá toàn diện đã được thực hiện trên một tập dữ liệu MRI đã được công bố, có số lượng lớn các ảnh.

(3) Việc phân tích và so sánh với các phương pháp phân đoạn truyền thống và các phương pháp học sâu khác thông qua các chỉ số đánh giá như chỉ số BF , SSIM và độ chính xác (ACC) đã được tiến hành.

2. TIÊU CHUẨN ĐÁNH GIÁ KẾT QUẢ PHÂN ĐOẠN

Trong nghiên cứu này, suất phân đoạn của các thuật toán được đánh giá định lượng bằng các chỉ số như chỉ số F1 theo biên (BF) và chỉ số tương đồng cấu trúc (SSIM). Các chỉ số này được sử dụng để đánh giá mức độ chồng lấp vùng, độ chính xác biên và tính tương đồng cấu trúc giữa các mặt nạ dự đoán và nhãn chuẩn thật trong tập dữ liệu.

2.1. Định nghĩa 1: chỉ số BF

BF được định nghĩa tương tự như chỉ số F1 truyền thống, nhưng thay vì so sánh toàn bộ mặt nạ của điểm ảnh, nó đánh giá các đường biên được trích xuất từ nhãn chuẩn thật và kết quả dự đoán. Công thức tính cho chỉ số này được cho trong (1).

$$BF = \frac{2 \cdot P \cdot R}{P + R} \qquad (1)$$

Trong đó, $P = \frac{|B_{pred} \cap B_{gt}|}{|B_{pred}|}$ là tỷ lệ các điểm biên dự đoán đúng trên tổng số điểm biên dự đoán,

$R = \frac{|B_{pred} \cap B_{gt}|}{|B_{gt}|}$ là tỷ lệ các điểm biên dự đoán đúng trên tổng số điểm biên thực, B_{gt} là tập các điểm biên của nhãn chuẩn, B_{pred} là tập các điểm biên của kết quả dự đoán.

2.2. Định nghĩa 2: chỉ số tương đồng cấu trúc

Chỉ số chỉ số tương đồng cấu trúc (SSIM) là một thước đo được dùng để đánh giá mức độ tương đồng giữa hai hình ảnh. Khác với các chỉ số truyền thống như sai số bình phương trung bình (MSE) hay tỷ số tín hiệu trên nhiễu đỉnh, chúng chỉ so sánh sự khác biệt theo từng điểm ảnh, SSIM đánh giá thông tin cấu trúc, phù hợp hơn với cách con người cảm nhận chất lượng hình ảnh.

SSIM được sử dụng rộng rãi để đánh giá chất lượng tái tạo ảnh hoặc kết quả phân đoạn ảnh, đặc biệt trong các bài toán như khử nhiễu, nén ảnh và phân tích ảnh y khoa. Công thức tính toán cho chỉ số này được cho trong (2)

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \qquad (2)$$

Trong đó, μ_x and μ_y là giá trị trung bình của hai ảnh x và y, σ_x^2 và σ_y^2 là phương sai, σ_{xy} là hiệp phương sai, C_1 , C_2 là các hằng số được sử dụng để ổn định phép chia.

3. MÔ HÌNH ĐỀ NGHỊ

3.1. Mô hình đề nghị

Khung phương pháp đề xuất gồm hai giai đoạn:

• Giai đoạn 1: Phân đoạn khối u bằng mô hình Res-UNet cải tiến.

•Giải đoạn 2: Phân loại sự hiện diện của khối u dựa trên các nhãn mặt nạ phân đoạn đã dự đoán.

Mô hình phân đoạn hoạt động ở mức pixel để tạo ra các mặt nạ nhị phân của khối u. Mỗi ảnh MRI trong bộ dữ liệu đều đi kèm với một mặt nạ phân đoạn được chuyên gia gán nhãn thủ công, trong đó các điểm ảnh thuộc vùng khối u được gán giá trị khác 0 và các điểcm ảnh nền được gán giá trị 0. Từ mặt nạ này, nhãn phân loại nhị phân được xác định theo nguyên tắc như sau:

• Nếu mặt nạ phân đoạn chứa ít nhất một điểm ảnh thuộc vùng khối u (tức tổng số điểm ảnh của vùng khối u lớn hơn 0) thì ảnh được gán nhãn Tumor (1).

• Nếu toàn bộ mặt nạ chỉ bao gồm các điểm ảnh nền (không tồn tại vùng khối u) thì ảnh được gán nhãn Normal (0).

Nhãn phân loại được xây dựng trực tiếp từ mặt nạ phân đoạn tương ứng. Nhãn phân loại nhị phân y_i được xác định như sau:

$$y_i = \begin{cases} 1, & \text{nếu } \sum_{p=1}^{H}\sum_{q=1}^{W} M_i(p,q) > 0, \\ 0, & \text{Ngược lại,} \end{cases}$$

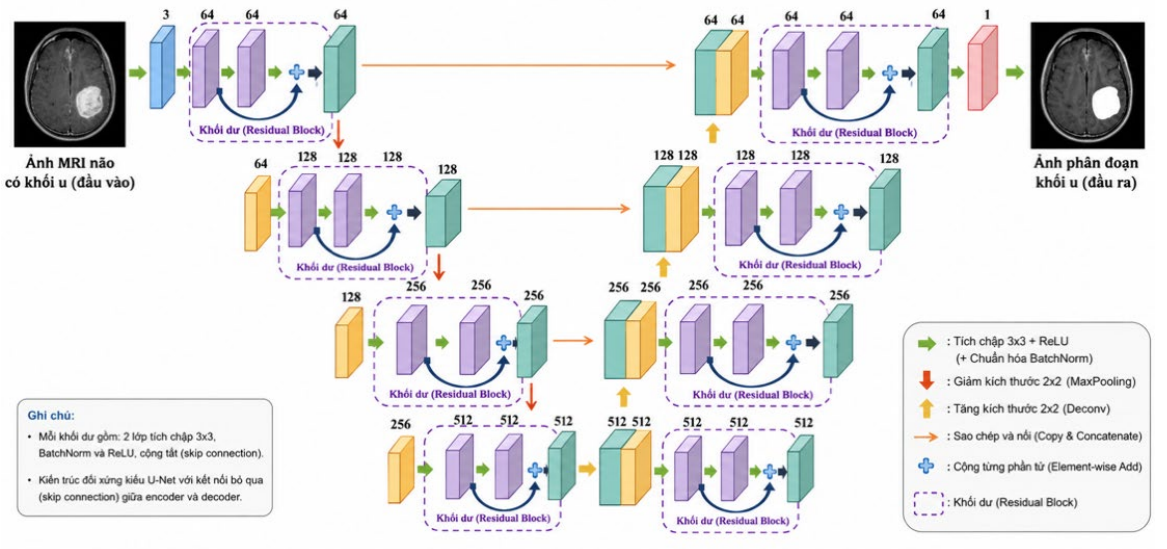
Trong đó $y_i \in \{0,1\}$ là nhãn phân loại của ảnh thứ i, $y_i = 1$ biểu thị ảnh có khối u $y_i = 0$ biểu thị ảnh không có khối u, $M_i(p,q)$ là giá trị của mặt nạ phân đoạn tại điểm ảnh (p,q), H và W lần lượt là chiều cao và chiều rộng của mặt nạ phân đoạn. Như vậy, nếu mặt nạ phân đoạn chứa ít nhất một điểm ảnh thuộc vùng khối u thì ảnh MRI tương ứng được gán nhãn $y_i = 1$. Ngược lại, nếu toàn bộ mặt nạ chỉ bao gồm các điểm ảnh nền thì ảnh được gán nhãn $y_i = 0$. Kiến trúc mô hình đề nghị được thể hiện như sau:

• **Encoder**: Ảnh đầu vào được xử lý qua các lớp tích chập và giảm kích thước để trích xuất đặc trưng đa mức.

• **Decoder**: Các đặc trưng được tăng kích thước và tái cấu trúc nhằm khôi phục độ phân giải ban đầu, tạo ra mặt nạ phân đoạn.

• **Kết nối dư** (ResNet): Việc bổ sung các kết nối dư giúp cải thiện lan truyền Gradient, tăng khả năng học và hiệu năng của mô hình.

Kiến trúc của mô hình đề xuất được minh họa trong Hình 1.



Hình 1. Sơ đồ minh họa của mô hình Res-UNet cải tiến

3.2. Kết nối dư

Để khắc phục hiện tượng mất Gradient và cải thiện khả năng học đặc trưng sâu, mô hình đề xuất sử dụng các khối Residual trong cả bộ mã hóa và bộ giải mã. Đầu ra của một khối Residual được xác định bởi (3).

$$y = F(x, W) + x, \qquad (3)$$

Trong đó, x là đầu vào của khối residual, $F(x, W)$ là hàm ánh xạ được học thông qua các lớp tích chập với tập trọng số W và y là đầu ra của khối.

3.2.1. Cơ chế chú ý (Attention)

Nhằm tăng cường khả năng tập trung vào vùng khối u và giảm ảnh hưởng của các cấu trúc nền, một Attention Gate được tích hợp trên mỗi kết nối của Encoder và Decoder. Hệ số Attention được tính bởi công thức (4).

$$\alpha = \sigma\left(\psi^T\left(\mathrm{ReLU}\left(W_x x + W_g g + b\right)\right)\right), \quad (4)$$

Trong đó, x là đặc trưng từ encoder, g là tín hiệu điều khiển từ decoder, W_x và W_g là các ma trận trọng số học được, b là hệ số dịch chuyển, $\sigma(.)$là hàm sigmoid và α là bản đồ attention. Đặc trưng đầu ra sau Attention Gate được xác định bởi:

$$x' = \alpha \odot x, \qquad (5)$$

*T*rong đó, $\odot$ biểu diễn phép nhân từng phần tử. Cơ chế này giúp mô hình tập trung vào các vùng chứa khối u và loại bỏ các đặc trưng không cần thiết.

3.2.2. Tích chập giãn (Dilated Convolution)

Để mở rộng vùng cảm thụ (receptive field) mà không làm giảm độ phân giải không gian của đặc trưng, tích chập giãn (Dilated Convolution) được sử dụng tại các tầng sâu của encoder. Công thức tổng quát của phép tích chập giãn được biểu diễn như sau:

$$y(i) = \sum_{k=1}^{K} x(i + r \cdot k) w(k)$$

Trong đó, *r* là hệ số giãn (dilation rate), $w(k)$ là trọng số của bộ lọc và *K* là kích thước kernel. Trong nghiên cứu này, các hệ số giãn $r = 2$ và $r = 4$ được sử dụng nhằm khai thác hiệu quả thông tin ngữ cảnh ở nhiều thang đo khác nhau.

3.2.3. Hàm mất mát (Loss Function)

Để tối ưu hóa chất lượng phân đoạn, hàm mất mát kết hợp giữa Dice Loss và Binary Cross-Entropy (BCE) được sử dụng, như trong (6).

$$L = \lambda L_{Dice} + (1 - \lambda) L_{BCE}, \qquad (6)$$

Trong đó, λ là hệ số cân bằng giữa hai thành phần mất mát,

$$L_{Dice} = 1 - \frac{2\sum_{i=1}^{N} p_i g_i + \varepsilon}{\sum_{i=1}^{N} p_i + \sum_{i=1}^{N} g_i + \varepsilon},$$

với p_i là xác suất dự đoán của pixel thứ i, g_i là nhãn thực tế tương ứng, *N* là tổng số pixel trong ảnh và ε là một hằng số nhỏ nhằm tránh hiện tượng chia cho 0 và

$$L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}[g_i \log(p_i) + (1 - g_i)\log(1 - p_i)].$$

Việc kết hợp $L_{Diçe}$ và L_{BCE} giúp mô hình vừa tối ưu mức độ chồng lấp giữa vùng dự đoán và vùng thực tế, vừa cải thiện tính ổn định trong quá trình huấn luyện đối với các tập dữ liệu có sự mất cân bằng giữa vùng khối u và nền ảnh.

3.3. Các tham số đã thiết lập

Mô hình Res-UNet sử dụng ảnh đầu vào dạng vuông có kích thước 256 × 256 tương thích với kiến trúc Encoder và Decoder. Lớp đầu ra được cấu hình theo số lượng lớp phân đoạn, trong đó nghiên cứu này sử dụng phân đoạn nhị phân. Encoder được cấu thành từ các khối dư kết hợp các lớp tích chập, chuẩn hóa và hàm kích hoạt ReLU, giúp trích xuất đặc trưng hiệu quả và duy trì sự lan truyền Gradient ổn định. Quá trình tăng kích thước được thực hiện bằng tích chập chuyển vị, đồng thời có thể sử dụng các trọng số của mạng ResNet được huấn luyện trước cho mục đích học chuyển giao. Các tham số huấn luyện cho mô hình đề nghị được tóm tắt trong Bảng 1.

Bảng 1. Một số tham số của mạng Res U-Net

Tham số	Giá trị
Tỷ lệ học	0,0001
Kích thước Batch	16
Số epochs	50
Hàm mất mát	Binary Cross-Entropy + Dice Loss
Tỷ lệ xác thực	0,2
Dừng sớm	Patience = 5
Hàm kích hoạt	ReLU / Sigmoid
Tỷ lệ Dropout	0,5
Trọng số khởi tạo	He normal
Thuật toán tối ưu	Adam

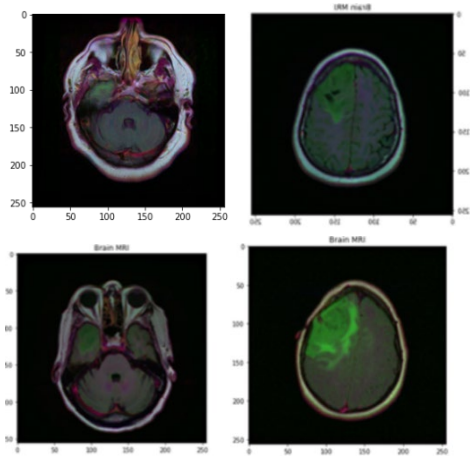
Bảng 2. Mô tả kiến trúc mô hình Res-UNet đề nghị

Tầng	Kích thước đầu ra	Số filters	Thành phần
Input	256×256×1	-	Ảnh MRI não
Encoder-1	256×256×64	64	Residual Block
MaxPool-1	128×128×64	-	Max Pooling (2×2)
Encoder-2	128×128×128	128	Residual Block
MaxPool-2	64×64×128	-	Max Pooling (2×2)
Encoder-3	64×64×256	256	Residual Block + Dilated Conv (r=2)
MaxPool-3	32×32×256	-	Max Pooling (2×2)
Encoder-4	32×32×512	512	Residual Block + Dilated Conv (r=4)
MaxPool-4	16×16×512	-	Max Pooling (2×2)
Bottleneck	16×16×1024	1024	Residual Block
Decoder-4	32×32×512	512	UpSampling + Attention Gate + Skip Connection
Decoder-3	64×64×256	256	UpSampling + Attention Gate + Skip Connection
Decoder-2	128×128×128	128	UpSampling + Attention Gate + Skip Connection
Decoder-1	256×256×64	64	UpSampling + Attention Gate + Skip Connection
Output	256×256×1	1	Conv (1×1) + Sigmoid

4. KẾT QUẢ ĐÁNH GIÁ VÀ THẢO LUẬN

4.1. Tập dữ liệu

Dữ liệu được sử dụng gồm có 3.929 ảnh MRI não người kèm theo thông tin vị trí khối u của 110 bệnh nhân (Buda et al., 2019; Akkus et al., 2017). Trong nghiên cứu này, 80% ảnh MRI não được sử dụng làm tập huấn luyện và 20% còn lại được sử dụng làm tập kiểm tra. Bộ dữ liệu này được công bố tại National Cancer Institute (n.d.). Một số ảnh tiêu biểu của bộ dữ liệu được minh họa trong Hình 2.



Hình 2. Ảnh MRI cắt ngang của não người

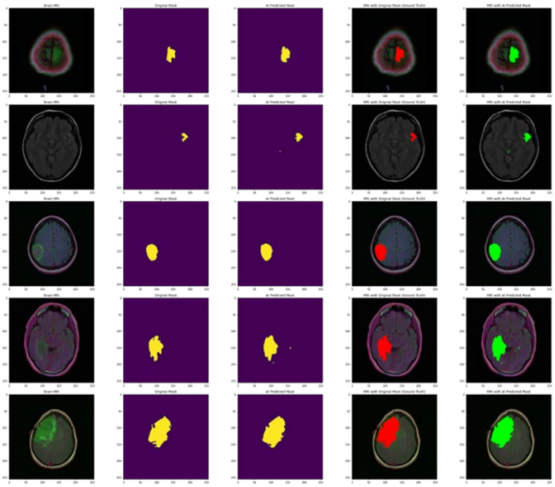
Ghi chú: Sự khác biệt về cường độ màu sắc cho thấy sự khác nhau về mật độ và thành phần mô, thường được sử dụng cho mục đích phân tích và chẩn đoán y khoa.

4.2. Tiền xử lý ảnh MRI

Trước khi huấn luyện mô hình, tất cả các ảnh MRI đều được đưa qua một quy trình tiền xử lý thống nhất nhằm đảm bảo tính nhất quán của dữ liệu và nâng cao độ ổn định của mô hình đề xuất. Trước tiên, tất cả các ảnh được thay đổi kích thước về 224 × 224 pixel, tương ứng với kích thước đầu vào của kiến trúc đề xuất. Để giảm hiện tượng quá khớp và cải thiện khả năng tổng quát hóa của mô hình, các kỹ thuật tăng cường dữ liệu được áp dụng. Các kỹ thuật này bao gồm lật ảnh ngang ngẫu nhiên (random horizontal flipping), xoay ngẫu nhiên trong phạm vi ±15°, phóng to hoặc thu nhỏ ngẫu nhiên với tỷ lệ tối đa 20%, dịch chuyển ảnh theo chiều ngang và chiều dọc với tỷ lệ tối đa 10%, cùng với điều chỉnh nhẹ độ sáng của ảnh. Những phép biến đổi này giúp tạo ra nhiều mẫu huấn luyện đa dạng hơn trong khi vẫn bảo toàn các đặc điểm giải phẫu quan trọng của vùng phổi.

4.3. Kết quả và thảo luận

Trong phần này, phương pháp đề xuất được áp dụng để phân đoạn vùng có khối u não trong ảnh MRI. Sau đó, các nhãn mặt nạ phân đoạn dự đoán được sử dụng để xác định kết quả phân loại sự hiện diện của khối u trên tập kiểm tra. Ngoài ra, để đánh giá hiệu quả của phương pháp đề xuất, kết quả của mô hình được so sánh với một số phương pháp học sâu như VGG16, ResNet50 và Inception V3. Hiệu quả phân loại được đánh giá định lượng bằng các chỉ số ACC và F1-Score. Các ảnh minh họa về kết quả phân đoạn được cho trong Hình 3.



Hình 3. Minh họa 10 ảnh MRI não cùng với các mặt nạ đã được phân đoạn

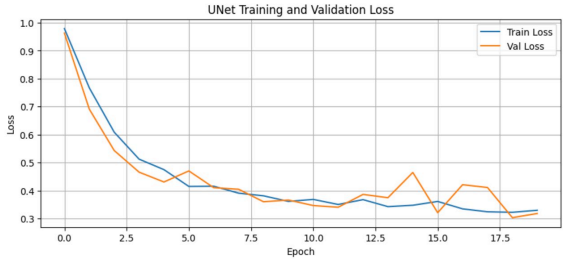
Kết quả thực thi các phương pháp phân đoạn ảnh được thể hiện trong Hình 4.

Kết quả đánh giá định lượng đối với đầu ra của các thuật toán, dựa trên các chỉ số chất lượng đã được thiết lập, được tóm tắt trong Bảng 3.

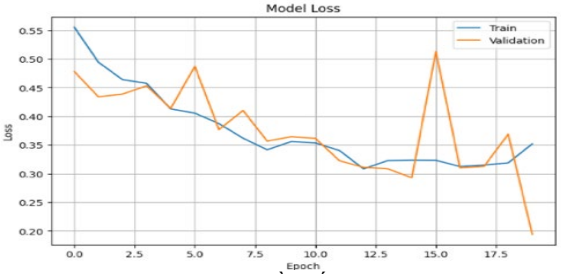
Bảng 3 thể hiện thuật toán U-Net thu được kết quả tốt hơn so với thuật toán FCM trên các chỉ số đánh giá như BF, SSIM và ACC . Tuy nhiên, phương pháp đề nghị thu được kết quả cao nhất ở tất cả các chỉ số so sánh, với BF = 98,0%, SSIM = 97,0% , ACC = 99,7%, Dice coefficient = 86,6% và IoU = 76,6%.

Bảng 3. Kết quả chỉ số BF Score, SSIM và ACC cho các thuật toán phân đoạn cùng loại

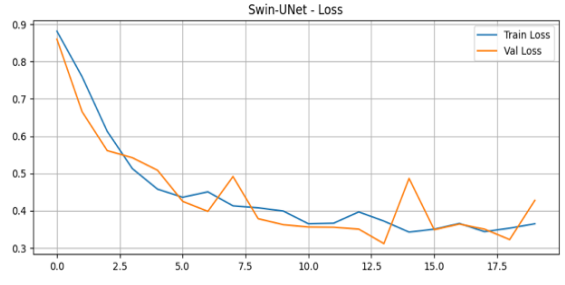
Thuật toán	BF	SSIM	ACC	Dice coefficient	IoU	Hausdorff Distance
Đề xuất	98,0	99,2	99,7	86,6	76,6	4,3
FCM	85,0	86,0	88,0	76,3	67,2	11,8
U-Net	88,0	89,0	90,0	81,0	73,2	7,9
Attention U-Net	86,5	99,2	99,7	86,5	76,2	4,6
UNet++	86,4	99,2	99,7	86,4	76,0	4,8
SegNet	85,4	99,1	99,7	85,4	74,5	5,5
Swin-UNet	77,2	99,0	99,6	77,1	62,7	10,9
Trans-UNet	82,4	99,0	99,6	82,4	70,1	7,1
NN-UNet	86,3	99,2	99,7	86,3	75,8	4,9



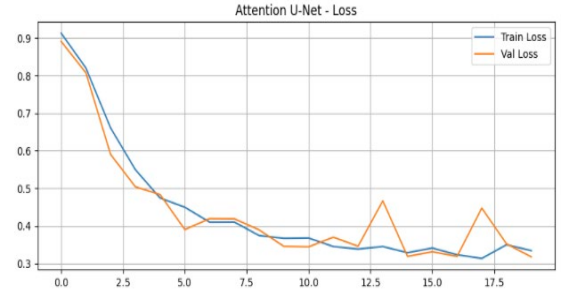
(a) Unet



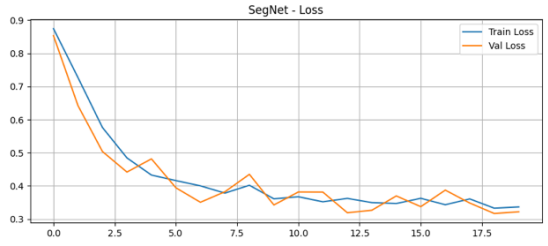
(b) Đề xuất



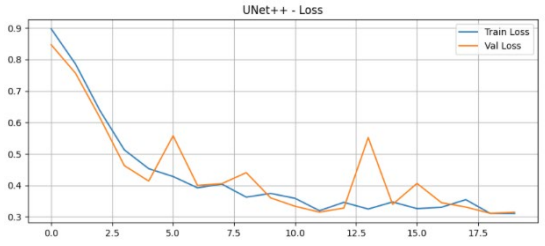
(c) Swin-UNet



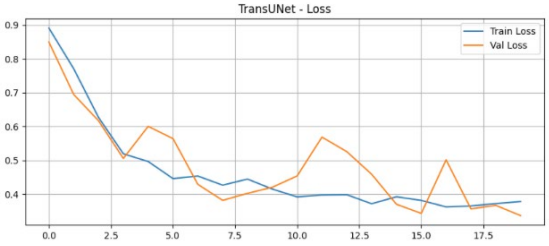
(d) Attention U-Net



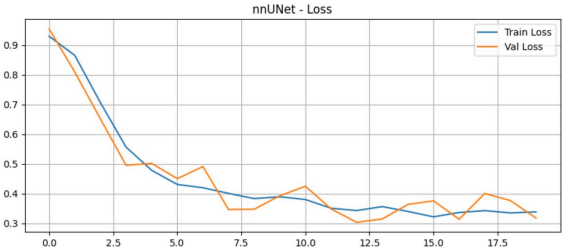
(e) SegNet



(f) UNet++

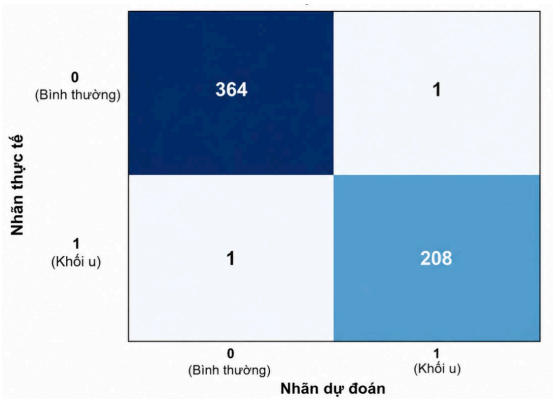


(g) TransUNet



(h) nnUNet

Hình 4. Đồ thị của hàm mất mát của các thuật toán



Hình 5. Kết quả phân loại được biểu diễn thông qua ma trận nhầm lẫn, trong đó 0 thể nhóm bình thường và 1 thể hiện nhóm ảnh có khối u

Ngoài ra, để kiểm chứng thêm về kết quả phân đoạn của thuật toán đề nghị, việc phân tích bài toán phân loại có sự hiện diện của khối u trong ảnh MRI đã được thực hiện bằng cách so sánh kết quả bệnh của dữ liệu có các nhãn mặt nạ trong tập kiểm tra và kết quả dự đoán của thuật toán đề nghị. Kết quả ma trận nhầm lẫn và các tính toán chỉ số đánh giá được tóm tắt trong Hình 5 và Bảng 4.

Bảng 4 cho thấy các phương pháp do Amin et al. (2018) và Abdalla & Esmail (2018) đề xuất đạt hiệu năng tốt, với giá trị ACC và F1-được báo cáo là 99%. Tuy nhiên, phương pháp đề xuất thu được kết quả tốt hơn với chỉ số ACC = 99,7% và F1-score = 99,5%

Bảng 4. Kết quả phân loại nhãn mặt na vùng khối u của các thuật toán

Thuật toán	ACC (%)	Precision (%)	Recall (%)	F1-Score (%)
Datta et al. (2020)	70,00	70,30	69,70	70,00
Abdalla and Esmail (2018)	99,00	99,20	98,80	99,00
Tamilselvan et al. (2013)	98,66	98,80	98,52	98,66
ResNet50	98,50	98,70	98,30	98,50
VGG16	95,28	95,50	94,90	95,20
Inception V3	92,40	92,70	92,10	92,40
Amin et al. (2018)	99,00	99,10	98,90	99,00
Đề xuất	99,70	99,80	99,20	99,50

5. KẾT LUẬN

Trong nghiên cứu này, một mô hình phân đoạn ảnh MRI não cải tiến đã được đề xuất bằng cách tích hợp các khối dư của ResNet vào kiến trúc U-Net kết hợp với cơ chế chú ý. Mô hình tận dụng khả năng học đặc trưng sâu của ResNet, đồng thời khai thác hiệu quả cấu trúc Encoder và Decoder của U-Net nhằm tăng cường khả năng trích xuất đặc trưng, cải thiện quá trình hội tụ và nâng cao chất lượng phân đoạn. Cơ chế chú ý được sử dụng để tăng cường khả

năng tập trung vào vùng khối u và giảm ảnh hưởng của các vùng nền không liên quan. Kết quả thực nghiệm trên bộ dữ liệu MRI não cho thấy mô hình đạt độ chính xác **99,7%.** Khi so sánh với các phương pháp như FCM, U-Net, Attention U-Net, UNet++, SegNet, Swin-UNet, TransUNet và nnUNet, mô hình đề xuất đạt các chỉ số **BF = 98,0%, SSIM = 99,2%, ACC = 99,7%, Dice coefficient = 86,6%** và **IoU = 76,6%,** cho thấy hiệu năng phân đoạn cạnh tranh và đạt kết quả tốt trên bộ dữ liệu được sử dụng. Bên cạnh đó, mô hình cũng đạt kết quả khả quan trong bài toán phân loại ảnh chứa khối u với **ACC = 99,7%** và **F1-score = 99,5%**, cao hơn các phương pháp được lựa chọn để so sánh như Datta et al. (2020), ResNet50, VGG16, Inception V3, Amin et al. (2018), Abdalla and Esmail (2018) và Tamilselvan et al. (2013). Mặc dù các kết quả thực nghiệm cho thấy hiệu quả của mô hình trên bộ dữ liệu nghiên cứu nhưng khả năng tổng quát hóa của mô hình vẫn cần được đánh giá thêm thông qua các thí nghiệm trên các bộ dữ liệu độc lập và các chiến lược xác thực mở rộng như **k-fold cross-validation**. Đây cũng là hướng nghiên cứu sẽ được triển khai trong các nghiên cứu tiếp theo nhằm đánh giá toàn diện hơn tính ổn định và khả năng ứng dụng của mô hình trong các điều kiện thực tế.

Tất cả các thí nghiệm được thực hiện trên nền tảng Google Colaboratory (Google Colab). Mô hình đề xuất được triển khai bằng Python 3.10 sử dụng khung học sâu PyTorch. Trong quá trình thực nghiệm, môi trường Google Colab cung cấp GPU NVIDIA Tesla T4 (16 GB VRAM), kết hợp với bộ xử lý Intel Xeon và khoảng 12 GB bộ nhớ RAM.

TÀI LIỆU THAM KHẢO

Abdalla, H. E. M., & Esmail, M. (2018). Brain tumor detection by using artificial neural network. In Yousif, A. F., Ahmed, S. Y. M., & Elhag, N. A. A., *2018 International Conference on Computer, Control, Electrical, and Electronics Engineering* (pp. 1-6). IEEE. https://doi.org/10.1109/ICCCEEE.2018.8515763

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. https://doi.org/10.1007/s10278-017-9983-4

Amin, J., Sharif, M., Yasmin, M., & Fernandes, S. L. (2018). Big data analysis for brain tumor detection: Deep convolutional neural networks. *Future Generation Computer Systems*, *87*, 290-297. https://doi.org/10.1016/j.future.2018.04.065

Buda, M., Saha, A., & Mazurowski, M. A. (2019). Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm, *Computers in Biology and Medicine*, *109*, 218-225. https://doi.org/10.1016/j.compbiomed.2019.05.002

Datta, S., Barua, R., & Das, J. (2020). Application of artificial intelligence in modern healthcare. In Leonel Pereira*, Alginates: Recent uses of this natural polymer* (pp. 121-134). IntechOpen. https://doi.org/10.5772/intechopen.90454

Farmanfarma, K. K., Mohammadian, M., Shahabinia, Z., Hassanipour, S., & Salehiniya, H. (2019). Brain cancer in the world: an epidemiological review. *World Cancer Research Journal*, *6*(5), 1-5.

Khalighi, S., Reddy, K., Midya, A., Pandav, K. B., Madabhushi, A., & Abedalthagafi, M. (2024). Artificial intelligence in neuro-oncology: advances and challenges in brain tumor diagnosis, prognosis, and precision treatment. *NPJ Precision Oncology*, *8*(1), 80. https://doi.org/10.1038/s41698-024-00575-0

Humphreys, S. C., Block, J. E., Sivaganesan, A., Nel, L. J., Peterman, M., & Hodges, S. D. (2025). Optimizing the clinical adoption of total joint replacement of the lumbar spine through imaging, robotics and artificial intelligence*. Expert Review of Medical Devices*, *22*(5), 405-413. https://doi.org/10.1080/17434440.2025.2484252

McKinney, P. A. (2004). Brain tumours: incidence, survival, and aetiology. *Journal of Neurology, Neurosurgery & Psychiatry*, *75*(2), 12-17. https://doi.org/10.1136/jnnp.2004.040741

National Cancer Institute. (n.d.). *The Cancer Genome Atlas Program (TCGA).* https://www.cancer.gov/ccg/research/genome-sequencing/tcga

Tamilselvan, K. S., Murugesan, G., & Gnanasekaran, B. (2013). Brain tumor detection from clinical CT and MRI images using WT-FCM algorithm. In P. Rangarajan, *2013 International Conference on Green Computing, Communication and Conservation of Energy* (pp. 260-263). IEEE. https://doi.org/10.1109/ICGCE.2013.6823441

Paschen, U. P. C., & Kietzmann, J. (2020). Artificial intelligence: Building blocks and an innovation typology. *Business Horizons*, *63*(2), 147-155. https://doi.org/10.1016/j.bushor.2019.10.004

Ruan, J., Li, J., & Xiang, S. (2025). VM-UNet: Vision Mamba UNet for medical image segmentation. *ACM Transactions on Multimedia*

Computing, Communications, and Applications, 1-17. https://doi.org/10.1145/3767748

Saeedi, S., Rezayi, S., Keshavarz, H., & R. Niakan Kalhori, S. (2023). MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. *BMC Medical Informatics and Decision Making*, *23*(1), 1-17. https://doi.org/10.1186/s12911-023-02114-6

Starmans, M. P. A., van der Voort, S.R., Castillo Tovar, J.M., Veenland, J.F., Klein, S., & Niessen, W.J. (2020). Radiomics: Data mining using quantitative medical image features. In Zhou, K. S., Rueckert, D., & Fichtinger, G., *Handbook of Medical Image Computing and Computer Assisted Intervention* (pp. 429-456). Academic Press. https://doi.org/10.1016/B978-0-12-816176-0.00023-5

Velastegui, R., Tatarchenko, M., Karaoglu, S., & Gevers, T. (2024). Image semantic segmentation of indoor scenes: A survey. *Computer Vision and Image Understanding*, *248*, 104102. https://doi.org/10.1016/j.cviu.2024.104102

Vo, V. T., & Pham, T. D. (2021). Segmentation for the magnetic resonance images of Covid-19 patients using adaptive automatic kernel-based fuzzy clustering algorithm. In Dinh, S. K., Nguyen, X. H., Nursanty, E., Van, T. C., Truong, V. M., Ton, T. L., Rahaman, H., Narvios, W. M. O., Galindo, R., Pepito, J., Pham, T. K., Ngo, T.T.T., Hai, D., Ngo, T. H., Nguyen, H.. T., Phuoc, H., Do, T. H. D., Dang, Q. N., & Lam, T. M. D., *1st van lang international conference on heritage and technology conference proceeding (p.* 040005*). AIP Conference Proceedings*. https://doi.org/10.1063/5.0066463

Wu, J., Ji, W., Fu, H., Xu, M., Jin, Y., & Xu, Y. (2024). Medsegdiff-v2: Diffusion-based medical image segmentation with transformer. In Wooldridge, M., Dy, J., & Natarajan, S., *The Thirty-Eighth AAAI Conference on Artificial Intelligence* (pp. 6030–6038). AAAI Press. https://doi.org/10.1609/aaai.v38i6.28418