

# Ứng dụng học sâu trong phân đoạn khối u ung thư não

Nguyễn Hoàng Thiên Thu<sup>1,2</sup>, Phạm Thị Thu Hiền<sup>1,2</sup>, Lê Thanh Hải<sup>2,3,\*</sup>



Use your smartphone to scan this QR code and download this article

## TÓM TẮT

U thần kinh đệm (glioma) là một trong những bệnh u não phổ biến trên toàn thế giới. Ngày nay, việc xác định và phân đoạn u thần kinh đệm là điều cần thiết khi đưa ra quyết định lâm sàng. Trong khi phân đoạn thủ công tốn nhiều thời gian và mang tính chủ quan, cùng lúc đó nhiệm vụ này cũng rất khó giải quyết đối với các phương pháp phân đoạn tự động. Kiến trúc U-net là một trong những kỹ thuật học sâu phổ biến để phân đoạn trong hình ảnh y sinh. Nghiên cứu này triển khai một mô hình học sâu dạng mạng thần kinh tích chập 2 chiều dựa trên kiến trúc U-net. Mô hình để xuất này được huấn luyện chỉ với các ảnh 2 chiều (slices) từ bộ dữ liệu MRI 3 chiều để phân đoạn các vùng u não khác nhau. Bên cạnh đó, một số phương pháp được áp dụng ở bước tiền xử lý để giúp giảm sự ảnh hưởng của việc mất cân bằng trong dữ liệu giữa các vùng u não khác nhau và áp dụng thêm phần tăng cường dữ liệu mở rộng giúp tránh hiện tượng khớp quá mức (overfitting) khi huấn luyện với bộ dữ liệu nhỏ. Kết quả đánh giá dựa trên bộ dữ liệu Multimodal Brain Tumor Segmentation Challenge (BraTS) 2018 cho thấy độ chính xác cao của mô hình được huấn luyện với điểm Dice cho toàn bộ khối u (whole tumor), khối u lõi (tumor core) và khối u ngấm thuốc (enhancing tumor) lần lượt là 0,88, 0,81 và 0,76. Từ kết quả trên cho thấy khả năng nhận dạng các vùng khối u trên rất tốt khi so sánh với các nghiên cứu khác. Nghiên cứu này đã đưa ra một quy trình phân tích cho u não thần kinh đệm, từ tiền xử lý đến huấn luyện mô hình, từ đó tiến hành đánh giá các yếu tố có thể ảnh hưởng đến kết quả đạt được.

**Từ khóa:** Mạng thần kinh học sâu, Phân đoạn, U não, U thần kinh đệm, U-net

<sup>1</sup>Khoa Kỹ Thuật Y Sinh, Trường Đại học Quốc tế

<sup>2</sup>Đại học Quốc gia TP Hồ Chí Minh, Việt Nam, 700000

<sup>3</sup>Khoa Cơ khí, Trường Đại học Bách Khoa TP.HCM, Việt Nam

## Liên hệ

**Lê Thanh Hải**, Đại học Quốc gia TP Hồ Chí Minh, Việt Nam, 700000

Khoa Cơ khí, Trường Đại học Bách Khoa TP.HCM, Việt Nam

Email: lthai@hcmut.edu.vn

## Lịch sử

- Ngày nhận: 09-12-2021
- Ngày chấp nhận: 07-6-2022
- Ngày đăng: 30-6-2022

DOI : 10.32508/stdjet.v5i2.951



## Bản quyền

© ĐHQG TP.HCM. Đây là bài báo công bố mở được phát hành theo các điều khoản của the Creative Commons Attribution 4.0 International license.



## TỔNG QUAN

U thần kinh đệm (glioma) là một trong những bệnh ung thư não phổ biến nhất, chiếm 15% các trường hợp bệnh u não. Khối u bắt đầu trong tế bào thần kinh đệm nằm ở cột sống và não<sup>1</sup>. Mặc dù u thần kinh đệm không phải là loại ung thư nặng nhất, nhưng nó được coi là khối u ác tính nguyên phát phổ biến của não, chiếm khoảng 80% ung thư não. Có rất nhiều cách để phân loại các loại u thần kinh đệm, chẳng hạn như theo loại tế bào, theo cấp độ và theo vị trí. Việc phân loại theo cấp độ chia u não dựa trên nguyên nhân và ảnh hưởng của khối u thần kinh đệm. Có hai loại u thần kinh đệm là: U thần kinh đệm độ cao (High-grade glioma - HGG), thường là ác tính và tiên lượng xấu hơn, và u thần kinh đệm độ thấp (Low-grade glioma - LGG), biểu hiện xu hướng lành tính và bệnh nhân thường có kết quả tốt hơn<sup>2</sup>. Trong hệ thống của WHO, các khối u được phân loại cấp độ II cho LGG và III-IV cho HGG (phạm vi từ I đến IV tương ứng với tiên lượng từ nhẹ đến nặng nhất)<sup>3</sup>. Về phương pháp điều trị, các bác sĩ thường phải phụ thuộc vào vị trí, loại tế bào và cấp độ u thần kinh đệm để đưa ra phương pháp và liệu trình hiệu quả. Ngày nay, một trong những kỹ thuật phổ biến để theo dõi tiến triển của bệnh là chụp ảnh cộng hưởng từ (Magnetic Resonance Image - MRI)<sup>4,5</sup>. Chụp ảnh cộng hưởng từ

đa phương thức (chuỗi xung) được sử dụng rộng rãi trong lâm sàng để chẩn đoán và theo dõi sự phát triển của khối u. MRI là một trong những kỹ thuật hình ảnh phổ biến vì nó tạo điều kiện thuận lợi cho việc phân tích khối u bằng cách quan sát sự lan rộng của nó; đồng thời MRI cũng cho thấy độ tương phản ở mô mềm so với các kỹ thuật khác như chụp cắt lớp vi tính (Computed Tomography - CT) và ghi hình cắt lớp phát positron (Positron Emission Tomography - PET). Ngoài ra, MRI đa phương thức thường được sử dụng để đánh giá các mô của khối u não vì chúng có khả năng tách các mô khác nhau bằng một trình tự cụ thể dựa trên các đặc tính của mô. Ví dụ, hình ảnh có trong T1 (nguyên bản) tốt cho việc phân tách các mô khỏe mạnh trong não trong khi với T1ce (tăng cường độ tương phản) giúp nhận biết ranh giới khối u do độ sáng lớn hơn bởi tính chất tương phản. Phù nề xung quanh khối u được phát hiện tốt trên ảnh của T2 (tín hiệu ngược với T1), còn hình ảnh Flair (Fluid Attenuated Inversion Recovery) là tốt nhất để phân biệt vùng phù (edema) từ dịch não tủy (Cerebrospinal Fluid - CSF)<sup>6,7</sup>.

U thần kinh đệm có cấu trúc và hình dạng phức tạp. Chúng yêu cầu mô tả chính xác bằng hình ảnh<sup>8</sup>. Các thành phần của khối u thường khuếch tán và có độ tương phản thấp. Đường viền của chúng thường mờ

**Trích dẫn bài báo này:** Thu N H T, Hiền P T T, Hải L T. Ứng dụng học sâu trong phân đoạn khối u ung thư não. *Sci. Tech. Dev. J. - Eng. Tech.*; 5(2):1437-1450.

và khó phân biệt với mô lành (chất trắng, chất xám và dịch não tủy), khiến khối u khó bị phân đoạn<sup>9</sup>. Việc phân đoạn thủ công tốn nhiều thời gian và công sức với phần lớn sai sót của con người. Do đó, phân đoạn tự động trở nên phổ biến và nhận được nhiều sự chú ý gần đây. Tuy nhiên, khối u thần kinh đệm cần phân biệt mức độ ác tính, tiên lượng thay đổi và một số vùng phụ mô học không đồng nhất dựa trên sự khác nhau của cường độ qua các phương thức MRI khác nhau, phản ánh các đặc tính sinh học đa dạng của khối u. Do đó, việc phân đoạn tự động trở nên phức tạp vì các khối u xuất hiện ở các kích thước, vị trí và cường độ khác nhau. May mắn thay, với sự phát triển của kỹ thuật và công nghiệp tiên tiến, học máy (machine learning) - một trong những hướng nghiên cứu hiện đại ngày nay, thu hút sự quan tâm rất lớn của các nhà khoa học - đã phát triển mạnh mẽ và các nhóm nghiên cứu bắt đầu áp dụng các thuật toán đó vào lĩnh vực y tế, bao gồm sự phân đoạn u thần kinh đệm. Có rất nhiều phương pháp cũng như mô hình học máy được áp dụng cho bài toán phân đoạn<sup>10-12</sup>. Trong nghiên cứu này, mô hình U-net - một trong những kiến trúc phổ biến nhất để phân đoạn trong lĩnh vực hình ảnh y tế - được dùng để thực thi nhiệm vụ phân đoạn này<sup>13,14</sup>.

Mặc dù có nhiều công trình nghiên cứu với những kết quả đáng khích lệ trong lĩnh vực hình ảnh y tế, việc phân đoạn nhanh và chính xác khối u não 3 chiều (3D - 3 dimensional space) bằng máy tính cho đến nay vẫn là một quá trình đầy thử thách và là một nhiệm vụ khó khăn vì khối u não có thể xuất hiện ở kích thước, hình dạng và vị trí bất kỳ với cường độ sáng khác nhau. Nhiều nghiên cứu gần đây đã sử dụng các phương pháp học sâu (Deep learning - DL), đặc biệt là mạng thần kinh tích chập (Convolutional Neural Network - CNN) đã cho thấy hiệu quả của chúng và chứng minh thành công trong việc tự động quét ảnh MRI não bình thường và bệnh lý. Nhiều phương pháp phân đoạn khối u não dựa trên mô hình học sâu đã được đề xuất và đạt được thành công lớn. Havaei và cộng sự<sup>9</sup> đã đề xuất một kiến trúc hai đường với một đường cục bộ và một đường toàn phần, có thể khai thác đồng thời cả các đặc trưng cục bộ và các đặc trưng ngữ cảnh toàn phần. Kamnitsas và cộng sự<sup>13</sup> đã đề xuất một kiến trúc CNN đa cấp độ hiệu quả có tên là Deepmedic sử dụng các phép tính tích chập 3D và tích hợp đường dẫn có độ phân giải cao và độ phân giải thấp để có được kết quả phân đoạn. Ngoài ra, họ còn sử dụng trường ngẫu nhiên 3D có điều kiện để loại bỏ các kết quả dương tính giả một cách hiệu quả. Chang và cộng sự<sup>15</sup> đề xuất một mạng thần kinh học sâu xếp tầng bao gồm hai mạng con không chỉ xác định được toàn bộ vùng khối u mà còn có thể phân đoạn chính xác

cấu trúc bên trong khối u. Tương tự, nhóm nghiên cứu của Casamitjana<sup>16</sup> cũng đề xuất dùng hai mạng V-Net với các phép tích chập 3D để giúp phân đoạn với các vùng khối u não. Đặc biệt, có một mạng V-Net sẽ được dùng để phân đoạn vùng khối u lớn nhất trong ảnh MRI từ đó sẽ chuyển tất cả vùng bên ngoài thành vùng mô bình thường. Hai mô hình V-Net này có thể được huấn luyện riêng lẻ sau đó kết hợp lại để cho kết quả tốt hơn. Trong khi nhóm của Alqazzaz<sup>17</sup> dùng bốn mạng SegNet để huấn luyện với bốn phương thức của MRI (Flair, T1, T1ce và T2) rồi sau đó tổng hợp kết quả lại bằng thuật toán cây quyết định (Decision tree). Việc này đã giúp cho mô hình của nhóm có thể nhận dạng chính xác các vùng khối u có kích thước nhỏ bên trong vùng khối u lớn. Trong những nghiên cứu trên, độ chính xác của toàn bộ khối u đạt được kết quả tốt là trên 0,85, trong khi khối u ngấm thuốc luôn cho thấy kết quả thấp hơn trong khoảng từ 0,67 đến 0,79. Trong nghiên cứu này, chúng tôi đề xuất sử dụng một mạng thần kinh tích chập 2D duy nhất với kiến trúc U-net cùng với các bước xử lý dữ liệu ảnh MRI cho bốn phương thức (Flair, T1, T1ce và T2) để phân đoạn chính xác ba vùng khối u gồm toàn bộ khối u (whole tumor - WT), khối u lõi (tumor core - TC) và khối u ngấm thuốc (enhancing tumor - ET).

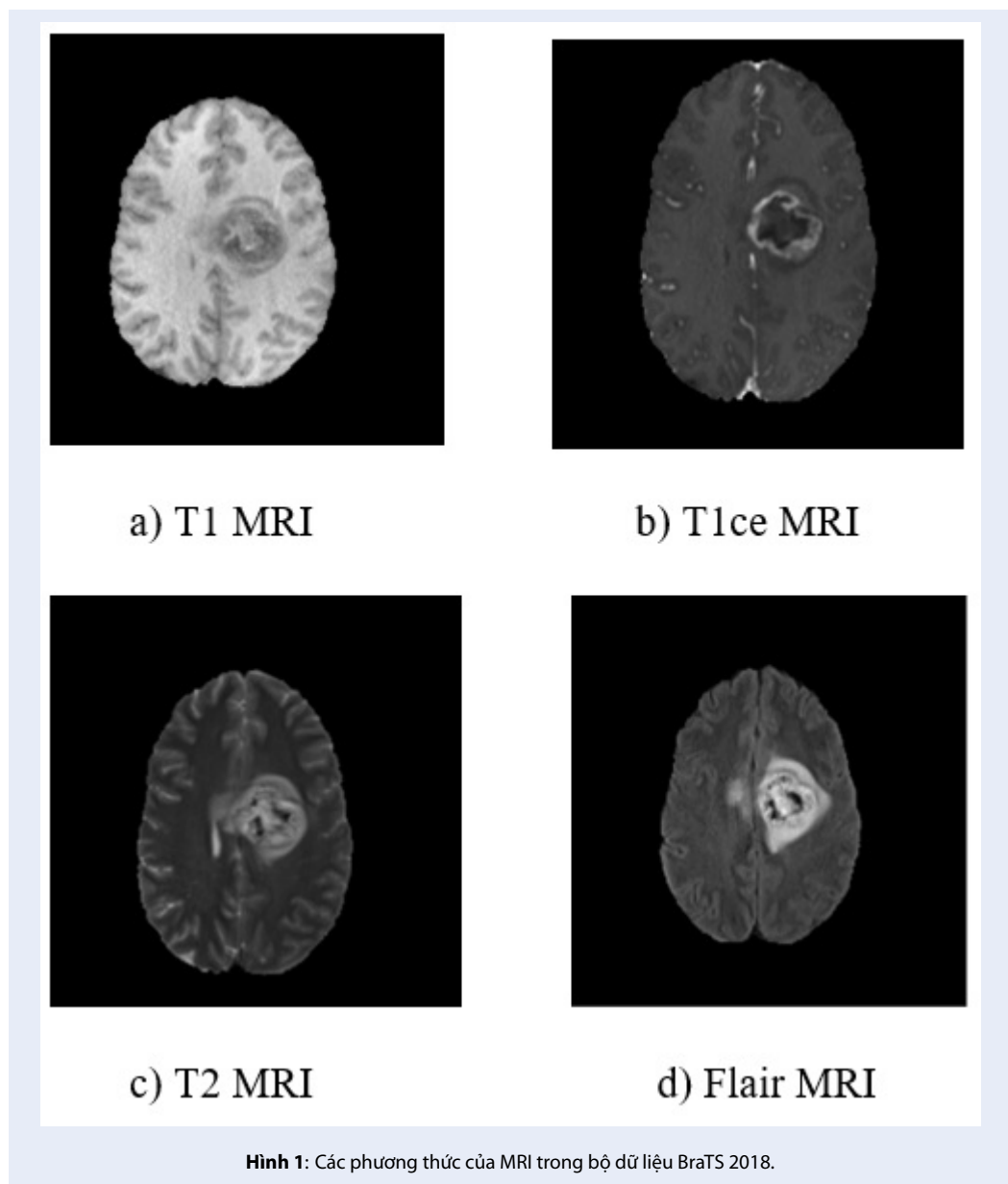
## **PHƯƠNG PHÁP NGHIÊN CỨU**

### **Bộ dữ liệu**

Nghiên cứu này sử dụng bộ dữ liệu lấy từ cuộc thi Multimodal Brain Tumor Segmentation Challenge (BraTS) 2018<sup>18,19</sup>. Đây là bộ dữ liệu được công khai, gồm 2 nhóm u thần kinh đệm là HGG và LGG. HGG có 210 mẫu MRI u não, trong khi LGG gồm 75 mẫu. Dữ liệu này được thu thập với các quy trình lâm sàng và các máy quét khác nhau từ nhiều bệnh viện trên thế giới. Trong mỗi mẫu MRI, có bốn loại phương thức: T1 nguyên bản (T1), T1 tăng độ tương phản (T1ce), T2 có tín hiệu ngược với T1 (T2), T2 xung xóa dịch (Fluid Attenuated Inversion Recovery - Flair) như được thể hiện trong Hình 1 và ảnh sự thật (Ground truth - GT) đánh dấu khối u não do các chuyên gia thực hiện.

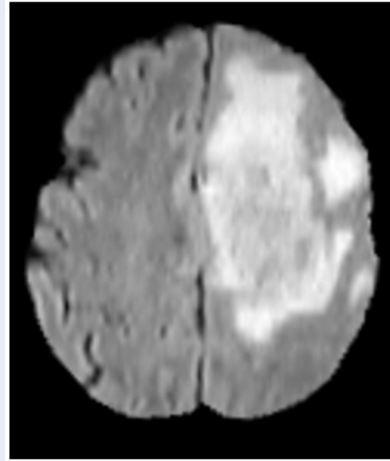
Ảnh sự thật (GT) được gán 4 nhãn đại diện cho các vùng u não khác nhau: 1 cho khối u hoại tử và không ngấm thuốc (necrotic and non-enhancing tumor), 2 cho khối u phù phúc mạc (peritumoral edema), 4 cho khối u ngấm thuốc (enhancing tumor) và 0 cho không phải khối u, được dán nhãn thủ công bởi các bác sĩ X quang chuyên về thần kinh đã được hội đồng chuyên gia chứng nhận như minh họa trong Hình 2.

Tập dữ liệu huấn luyện chiếm 70% của bộ dữ liệu BraTS 2018 (bao gồm 285 mẫu bao gồm HGG và

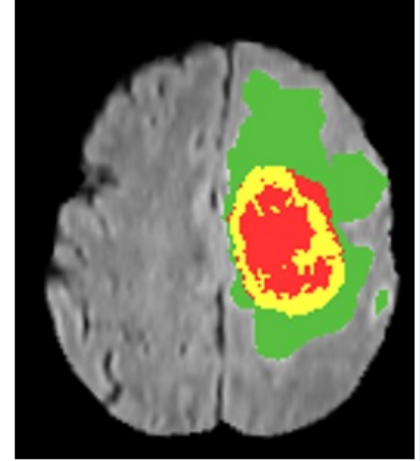


**Bảng 1:** Danh sách các mẫu và cách phân chia dữ liệu từ bộ dữ liệu BraTS 2018.

Loại u não thần kinh đệm	Số lượng mẫu (Dataset)	Tập dữ liệu huấn luyện (Training set)	Tập dữ liệu kiểm tra (Testing set)
HGG	210	147	63
LGG	75	53	22
Tổng cộng	285	200	85

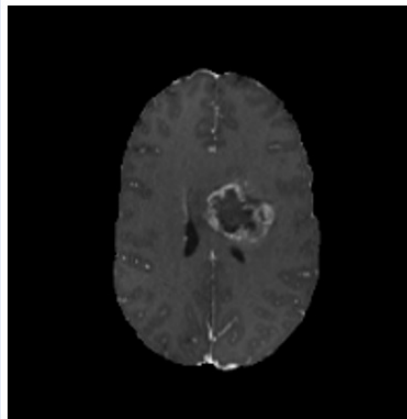


a) T1 MRI

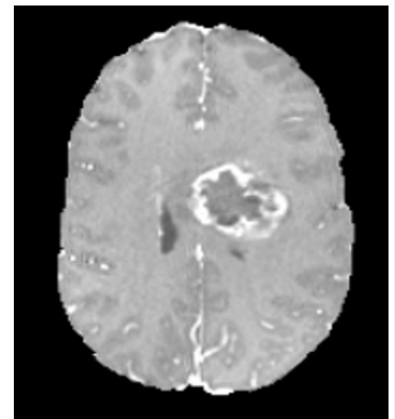


b) T1 kết hợp với GT

**Hình 2:** Các vùng u thần kinh đệm: hình ảnh T1 (a) và hình ảnh T1 kết hợp với các nhãn khối u được áp dụng cho toàn bộ dữ liệu BraTS 2018 (b): toàn bộ khối u (màu xanh lá cây), khối u ngấm thuốc (màu vàng) và khối u lõi - khối u không ngấm thuốc (màu đỏ).



a) Trước



b) Sau

**Hình 3:** Ảnh MRI trước (a) và sau (b) khi chuẩn hóa và giảm bớt kích thước ảnh

LGG) và phần còn lại là tập dữ liệu kiểm tra dùng để đánh giá mô hình như mô tả trong Bảng 1. Trong giai đoạn huấn luyện, tập dữ liệu sẽ được chia thành 5 tập dữ liệu nhỏ để áp dụng phương pháp 5-fold cross validation giúp huấn luyện và đánh giá mô hình đã thiết lập tốt hơn cho nhiệm vụ phân đoạn. Kết quả đầu ra sẽ là ảnh được gán nhãn theo dự đoán của mô hình U-net dựa trên các MRI đa phương thức. Mặt khác, trong giai đoạn kiểm tra, ảnh đầu vào của mô hình được đào tạo trước đó sẽ là ảnh đầy đủ (full slice) của

bộ dữ liệu MRI.

### Quá trình tiền xử lý

Mỗi phương thức MRI là một tập hợp 155 hình ảnh (hay còn gọi là slice) với mỗi ảnh có kích thước  $240 \times 240$  điểm ảnh. Do đó, số lượng điểm ảnh của một tập hình ảnh có thể lên đến hàng triệu trong không gian 3 chiều ( $240 \times 240 \times 155$ ) đòi hỏi công suất tính toán rất lớn. Một trong những giải pháp cho vấn đề này là trích xuất cửa sổ nhỏ trong ảnh (patch extrac-

tion). Phương pháp này giúp giảm kích thước của dữ liệu huấn luyện mà không làm mất thông tin. Trong phương pháp cửa sổ, các hình ảnh đầu vào chia thành nhiều cửa sổ dọc theo các ảnh (slice) được trích xuất với một kích thước nhất định. Các ảnh cửa sổ nhỏ được đưa vào mô hình riêng lẻ và việc xác định lớp (loại u não) phụ thuộc vào nhân của điểm ảnh 3 chiều (voxel) trung tâm của cửa sổ này. Ngoài ra, bộ dữ liệu BraTS 2018 cho thấy sự mất cân bằng nghiêm trọng giữa các lớp với số lượng điểm ảnh không phải khối u (nhân 0) lớn hơn nhiều lần so với các điểm ảnh khối u (nhân 1, 2 và 4). Điều này gây ra khó khăn khi huấn luyện mô hình sử dụng những dữ liệu này, vì các điểm ảnh không phải khối u ảnh hưởng đến hàm mất mát (loss function) mạnh hơn nhiều so với điểm ảnh của những khối u. Do đó, cửa sổ được trích xuất bằng cách sử dụng phân phối đồng đều (uniform distribution) đảm bảo rằng hầu hết số lượng lớp của mỗi cửa sổ là như nhau. Điều này giúp loại bỏ ảnh hưởng gây ra bởi sự mất cân bằng lớp vì chỉ cần lấy mẫu ngẫu nhiên các cửa sổ. Khi đó số lượng cửa sổ chứa phần lớn điểm ảnh có cường độ bằng 0 và không chứa thông tin cần thiết của khối u sẽ bị loại bỏ khỏi tập hợp các cửa sổ mẫu.

Bên cạnh đó, kích thước cửa sổ có ảnh hưởng lớn đến khả năng huấn luyện của mô hình. Nếu kích thước quá nhỏ, thì đặc trưng cục bộ trong vùng xung quanh điểm ảnh trung tâm có thể không đủ thông tin để mô hình hoạt động tốt, trong khi kích thước lớn hơn thì ảnh hưởng tới tốc độ tính toán. Trong nghiên cứu này, cửa sổ có kích thước  $128 \times 128$  điểm ảnh được chọn vì kích thước này có thể chứa đủ thông tin để huấn luyện mô hình U-net<sup>20-22</sup>. Thêm nữa, cần tiến hành chuẩn hóa các kích thước dữ liệu cũng như khoảng cường độ sáng trong các ảnh để chúng có cùng tỷ lệ được minh họa như trong Hình 3. Tại mỗi ảnh, dữ liệu được chia theo độ lệch chuẩn của nó, để chuyển ảnh thành dữ liệu có tâm là 0 (zero center).

Khi sử dụng hình ảnh y tế cho mô hình phân đoạn, đặc biệt là MRI, rất khó khăn trong việc đánh dấu nhân, nên nguồn dữ liệu thường khá hạn chế. Tuy nhiên, có một nhận thức chung là mô hình học sâu càng truy cập nhiều dữ liệu thì nó càng có thể học hiệu quả hơn. Vì vậy, để có thêm dữ liệu cho việc huấn luyện và cũng để giải quyết vấn đề mất cân bằng lớp trong ảnh sự thật, chúng tôi đã áp dụng các phép biến đổi để từ một hình ảnh thành tạo ra nhiều hình ảnh có nhãn, nhưng khác vị trí<sup>23</sup>. Dữ liệu hình ảnh có thể được tăng giả tạo bằng cách sử dụng lật ngang / dọc, xoay, thay đổi độ sáng của hình ảnh, dịch chuyển ngang / dọc, v.v. Kỹ thuật này được gọi là tăng cường dữ liệu (data augmentation). Tăng cường dữ liệu là một cách để tăng hiệu suất của các mô hình học sâu,

đặc biệt là mạng thần kinh tích chập (CNN), và có thể giảm xác suất mô hình khớp quá mức (overfitting), bằng cách thêm lớp bỏ học (dropout layer), đối với dữ liệu nhỏ<sup>24</sup>.

### Mô hình phân đoạn khối u não

Trong nghiên cứu này, kiến trúc U-net nâng cấp được xây dựng để xử lý đầu vào là tập hợp các ảnh cửa sổ hai chiều được trích xuất dọc theo trục z từ bốn phương thức MRI. Đầu ra sẽ là ảnh sự thật có cùng kích thước với đầu vào và mỗi điểm ảnh được gán nhãn 1, 2, 4 tương ứng với các loại u não và 0 là không phải u não. Mô hình này có ba loại khối tương ứng với hai phần của mô hình: khối rút gọn (contraction block) và khối mở rộng (expansion block). Phần rút gọn sử dụng ba khối rút gọn. Mỗi khối có đầu vào là hai lớp tích chập (convolutional layer)  $3 \times 3$  theo sau là hàm kích hoạt tuyến tính được dùng là PReLU (Parametric Rectified Linear Unit)<sup>25</sup>, thay vì hàm ReLU (Rectified Linear Unit) được sử dụng trong kiến trúc ban đầu. Giữa hai khối kế tiếp, một lớp tích chập khác có bước là 2 được sử dụng để giảm số lượng mẫu đầu vào được thể hiện ở Hình 4(a). Sau mỗi khối, số lượng bộ lọc (kernel) tăng gấp đôi để mô hình có thể học các cấu trúc phức tạp một cách hiệu quả. Tuy nhiên, trung tâm của mô hình này nằm ở phần mở rộng. Tương tự như phần rút gọn, phần này cũng bao gồm một số khối mở rộng. Mỗi khối chuyển đầu vào đến hai lớp tích chập  $3 \times 3$ , tiếp theo là lớp tăng số lượng mẫu lên 2 như minh họa trong Hình 4(b). Ngoài ra, sau mỗi khối, số lượng đặc trưng sẽ bị giảm đi một nửa trước khi đưa vào lớp tích chập tiếp theo. Sau đó, ở đầu vào của lớp tích chập kế tiếp sẽ được thêm vào các đặc trưng của lớp rút gọn tương ứng. Hành động này sẽ đảm bảo rằng các đặc trưng được học trong quá trình rút gọn ảnh sẽ được sử dụng để tái tạo lại hình ảnh đó. Số lượng khối mở rộng cũng bằng với số lượng khối rút gọn. Sau cùng, các kết quả đưa qua một lớp tích chập  $3 \times 3$  để xuất ra số lượng ảnh đặc trưng bằng số lượng nhãn mong muốn. Cả khối rút gọn và khối mở rộng đều có khả năng thực hiện chuẩn hóa hàng loạt (batch normalization) trên các đầu vào cũng như bao gồm việc cộng dồn các đặc trưng được đưa vào các khối tích chập thông qua kết nối tắt (short skip connection) trong kiến trúc mạng dư (Residual network - ResNet).

Tóm lại, mô hình đề xuất gồm 7 khối dư, 25 lớp tích chập và 15 lớp chuẩn hóa hàng loạt như trình bày ở Hình 5. Mô hình này được huấn luyện với các hình ảnh được trích xuất từ bộ dữ liệu MRI dành cho phần huấn luyện trong BraTS 2018, bao gồm các trường hợp HGG và LGG, và các ảnh sự thật tương ứng. Mục tiêu huấn luyện mô hình là xác định các tham

số cho mạng U-net (trọng số - weight và độ lệch - bias) nhằm giảm thiểu hàm mất mát (loss function). Để làm được điều này, thuật toán Stochastic Gradient Descent (SGD) được sử dụng để cập nhật các tham số theo hướng ngược lại của gradient ở mỗi lần chạy. Đầu vào của mô hình là tập hợp bốn ảnh ở cùng vị trí từ bốn phương thức MRI (Flair, T1, T1ce và T2) và mỗi lần đưa dữ liệu vào mô hình để huấn luyện sẽ là 16 (batch size) với 100 lần chạy hết toàn bộ dữ liệu huấn luyện (epoch) theo tỉ lệ học (learning rate) 0,001. Ảnh output là tổng hợp giá trị xác suất cho từng nhãn gồm 0, 1, 2 và 4 (đại diện bởi thứ tự ma trận) và sau đó hàm argmax sẽ được dùng để tìm ra nhãn có giá trị lớn nhất tại cùng tọa độ pixel để xuất ra kết quả cuối cùng là ảnh 2D với kích thước  $128 \times 128 \times 1$ . Toàn bộ mô hình và thuật toán xử lý được viết với ngôn ngữ Python và thư viện Pytorch. Còn bộ xử lý đồ họa (Graphics Processing Unit - GPU) được sử dụng để huấn luyện và đánh giá mô hình là GTX 1060 8GB.

### Bộ thước đo đánh giá mô hình

Trong nghiên cứu này, bốn thước đo tiêu chuẩn được sử dụng để đánh giá độ chính xác các kết quả phân đoạn và đo lường mức độ giống nhau giữa phân đoạn A và B. Các thước đo này có thể được xác định bằng các biểu thức (1) như trình bày sau đây:

$$Dice(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (1)$$

Thước đo *Dice* được xem là chỉ số đo độ chính xác. Nó đo lường sự giống nhau giữa hai tập A và B, tương ứng với phân đoạn đầu ra của mô hình và ảnh sự thật được đánh dấu bởi các chuyên gia.

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

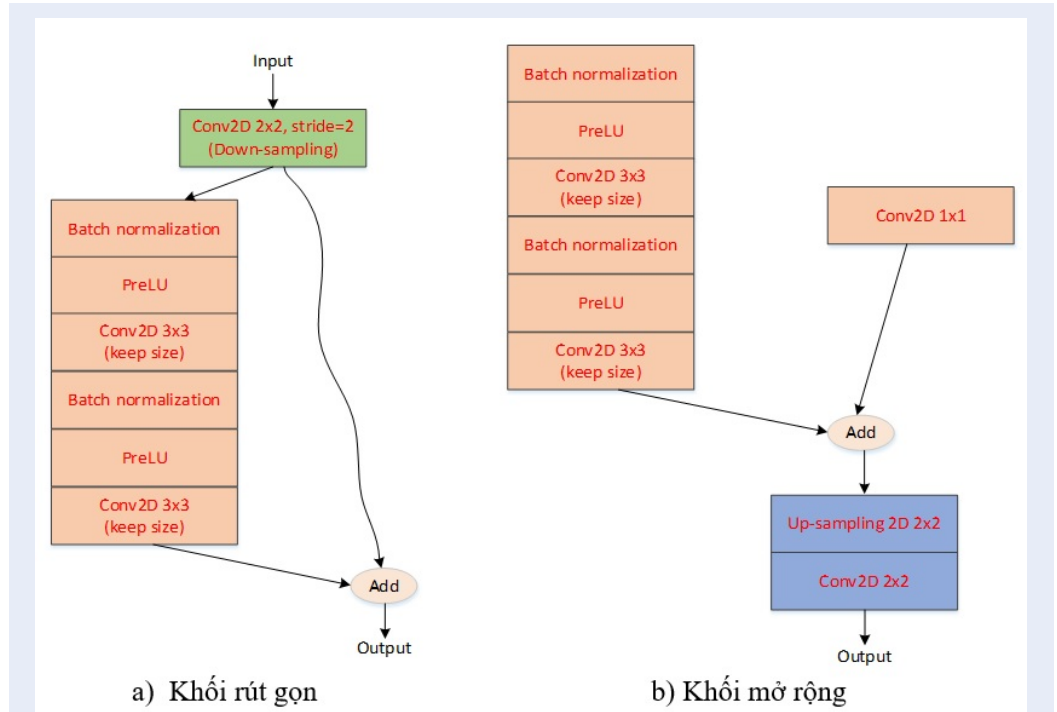
Thước đo *Sensitivity* và *Specificity* là các chỉ số thống kê được sử dụng để đánh giá các dự đoán và tỷ lệ True Positive (TP), False Negative (FN), False Positive (FP) và True Negative (TN) của các điểm ảnh không gian. Thước đo *Sensitivity*, còn được gọi là True Positive Rate (TPR), đo tỷ lệ TP của điểm ảnh không gian trong khối u não được phân đoạn chính xác, trong khi thước đo *Specificity*, còn được gọi là True Negative Rate (TNR), cho biết mức độ TN được dự đoán.

## KẾT QUẢ VÀ THẢO LUẬN

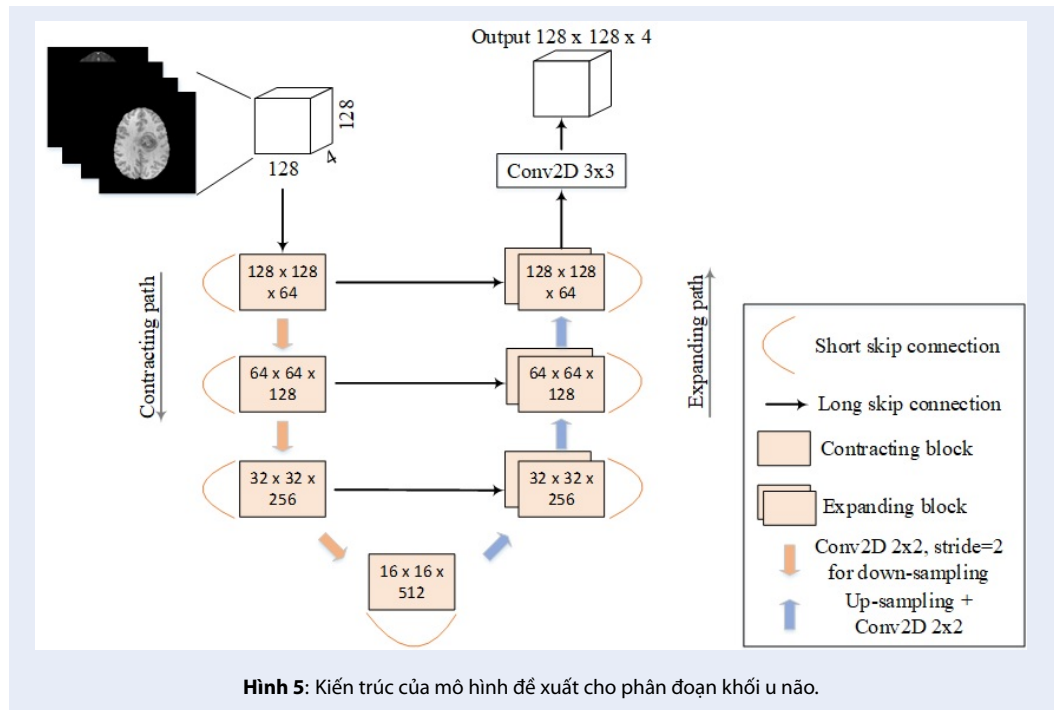
Phần đánh giá kết quả sẽ dựa trên việc phân đoạn chính xác ba vùng khối u gồm: toàn bộ khối u (whole tumor - WT) là bao gồm tất cả khối u với nhãn 1, 2 và 4, khối u lõi (tumor core -TC) là gồm khối u phủ phúc mạc và khối u ngấm thuốc với nhãn 1 và 4 và cuối cùng là khối u ngấm thuốc (enhancing tumor - ET) với nhãn 4 trong ảnh sự thật của bộ dữ liệu BraTS 2018.

Trong Bảng 2, có thể thấy toàn bộ khối u (WT) đạt được điểm Dice cao nhất là 0.88, thể hiện hiệu suất tốt nhất trong ba loại phân đoạn, trong khi khối u ngấm thuốc (ET) là vùng khó phân đoạn, được chứng minh bằng điểm Dice thấp nhất với 0.76. Lý do là do sự mất cân bằng rất lớn trong sự phân bố của khối u ngấm thuốc ở cả hai loại u thần kinh đệm HGG và LGG. Khối u ngấm thuốc có tồn tại chủ yếu trong các loại u thần kinh đệm HGG, trong khi u thần kinh đệm LGG hầu như không chứa vùng khối u này nên dẫn tới thông tin của vùng khối u này ít hơn nhiều so với toàn bộ khối u và khối u lõi (TC) khi đưa vào huấn luyện mô hình. Do đó, độ chính xác của việc phân đoạn khối u ngấm thuốc thường không cao bằng hai loại khối u khác. Độ lệch chuẩn của khối u lõi và khối u ngấm thuốc là khá lớn, gần 0.3, thể hiện sự biến đổi lớn của dữ liệu khối u. Trong khi đó ở Bảng 3 và Bảng 4, khi so sánh kết quả của thước đo Sensitivity và Specificity, rõ ràng là Sensitivity có giá trị trung bình gần với giá trị trung bình của thước đo Dice, và độ lệch chuẩn của Sensitivity cũng tương tự với Dice, trong khi Specificity đạt được hiệu suất tốt, trong khoảng từ 0,98 đến 1, với phân bố tập trung do trong hình ảnh phân đoạn, số lượng điểm ảnh có giá trị là 0 (TN) tương ứng với lớp 0 (không phải khối u) chiếm số lượng lớn.

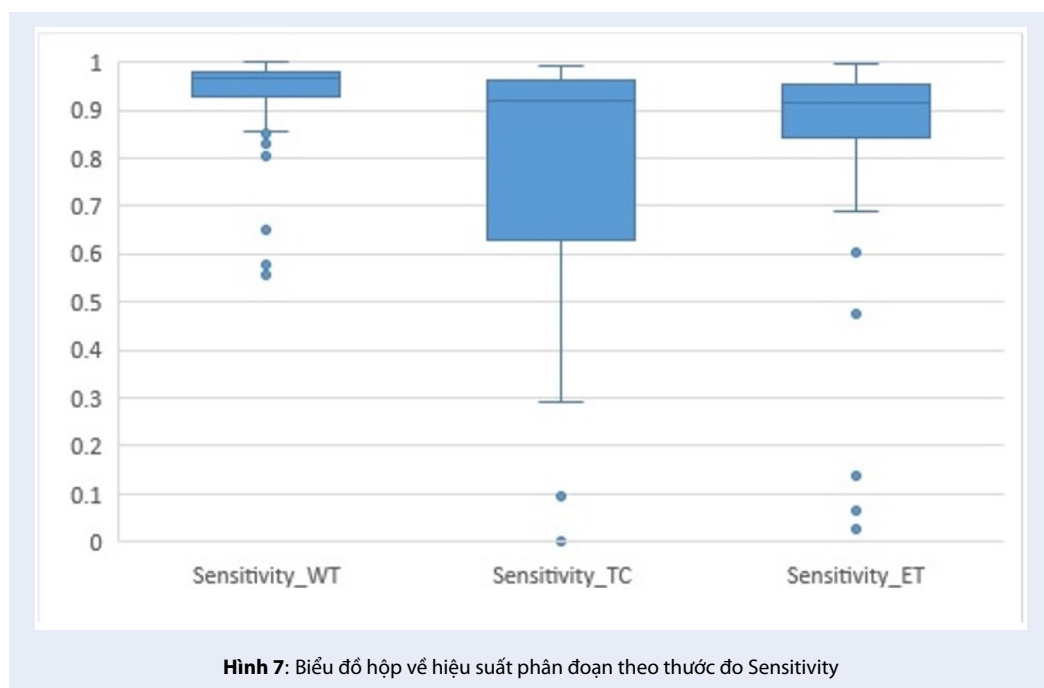
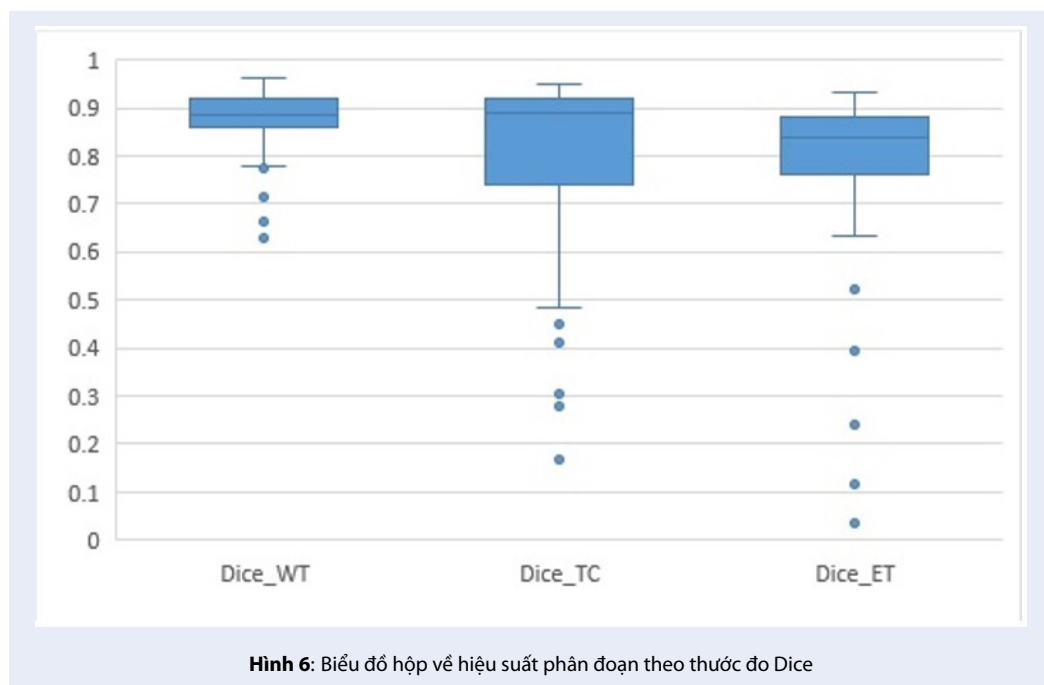
Hình 6 thể hiện sự phân bố kết quả điểm Dice của ba vùng khối u gồm toàn bộ khối u, khối u lõi và khối u ngấm thuốc. Từ hình vẽ, chúng ta có thể thấy rằng mặc dù phần lớn kết quả phân phối trên 0.75, nhưng có một số điểm ngoại lai (outliers) trong khối u lõi và khối u ngấm thuốc nằm dưới 0.5, điều này dẫn đến sự khác biệt lớn trong kết quả và cho thấy có một số trường hợp mô hình không thể phân đoạn chính xác. Từ các Hình 7 và Hình 8, có thể nhận ra rằng mô hình dễ dàng phát hiện vùng không phải khối u (giá trị Specificity cao có nghĩa số lượng TN thực sự nhiều), nhưng rất khó phân đoạn được khối u não thực sự, và thường nhầm lẫn giữa các điểm ảnh thuộc khối u và không phải khối u trong vùng khối u như một số hình ảnh minh họa trong Hình 9. Do đó, nếu muốn cải thiện hiệu suất của điểm Dice (độ chính xác) trong phân đoạn, thì việc cải thiện chỉ số Sensitivity là cần



**Hình 4:** Các lớp trong một khối rút gọn và mở rộng: (a) khối màu cam thể hiện khối rút gọn, màu xanh lá thể hiện khối giảm mẫu và (b) khối màu cam đại diện cho khối mở rộng, màu xanh đại diện cho khối tăng số lượng mẫu



**Hình 5:** Kiến trúc của mô hình đề xuất cho phân đoạn khối u não.





**Bảng 2: Đánh giá kết quả phân đoạn theo thước đo Dice**

	Dice		
	WT	TC	ET
Mean	0,880	0,807	0,760
Std. Dev	0,069	0,261	0,283
Median	0,886	0,890	0,835
25 quantile	0,861	0,749	0,774
75 quantile	0,917	0,919	0,877

**Bảng 3: Đánh giá kết quả phân đoạn theo thước đo Sensitivity**

	Sensitivity		
	WT	TC	ET
Mean	0,930	0,772	0,825
Std. Dev	0,094	0,281	0,254
Median	0,967	0,919	0,915
25 quantile	0,931	0,639	0,835
75 quantile	0,978	0,962	0,953

**Bảng 4: Đánh giá kết quả phân đoạn theo thước đo Specificity**

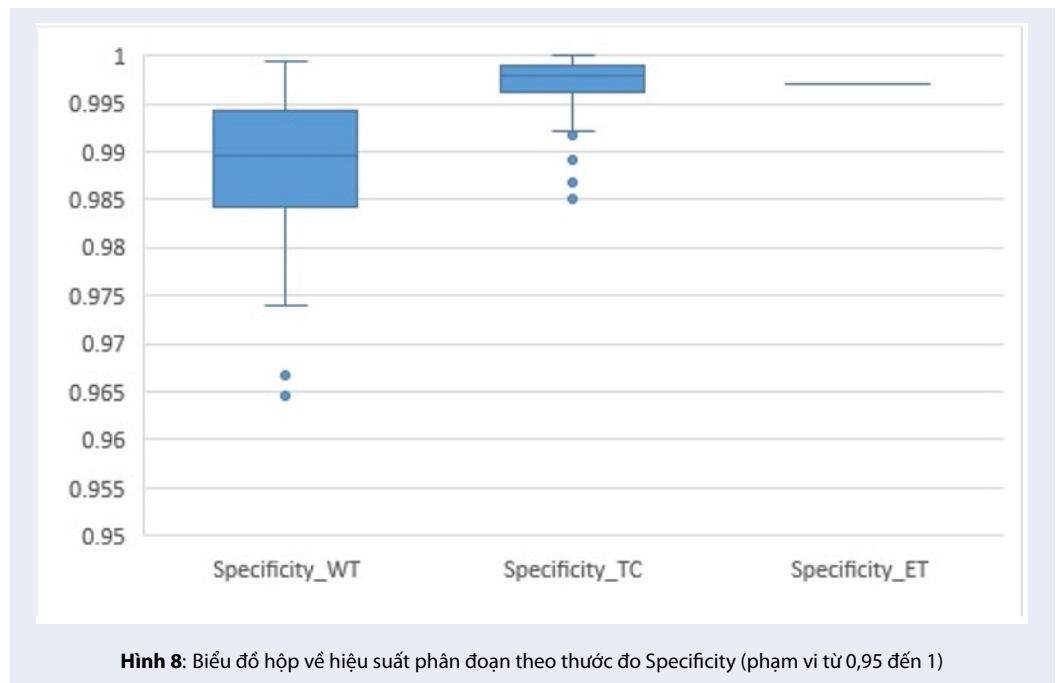
	Specificity		
	WT	TC	ET
Mean	0,988	0,997	0,996
Std. Dev	0,00746	0,003	0,005
Median	0,990	0,998	0,998
25 quantile	0,994	0,996	0,996
75 quantile	0,994	0,999	0,999

thiết trong trường hợp này, bởi vì thước đo Sensitivity được tính dựa trên số lượng TP và FN nên ảnh hưởng lớn đến độ chính xác của mô hình.

Hình 9 và Hình 10 cho thấy một số kết quả phân đoạn của mô hình đã đề xuất cho các trường hợp HGG và LGG trong tập dữ liệu kiểm tra đã đề cập ở Bảng 1. Các kết quả này cho thấy mô hình đề xuất của nhóm có thể nhận diện tốt vùng khối u và không phải khối u. Tuy nhiên, phân đoạn cho các vùng bên trong thì còn sai sót do các vùng này rất nhỏ, đặc biệt là vùng biên giữa khối u lõi và khối u ngấm thuốc đối với loại u thần kinh đệm HGG và vùng biên giữa toàn bộ khối u và khối u lõi đối với loại u thần kinh đệm LGG. Đặc biệt, nếu vùng khối u ngấm thuốc rất nhỏ thì mô hình đề xuất có thể phân đoạn sai thành toàn bộ khối u như ví dụ cuối cùng trong Hình 9.

Phương pháp của chúng tôi đạt được điểm Dice đối với WT, TC và ET trung bình lần lượt là 0,88, 0,81 và

0,76. So với các công trình khác về phân đoạn u thần kinh đệm, những điểm số này tương đương với các kết quả thu được của các phương pháp hàng đầu hiện nay. Bảng 5 thể hiện kết quả của mô hình U-net có cải tiến của chúng tôi so với kết quả của bốn mô hình gần đây dựa trên thước đo Dice cho mỗi vùng phân đoạn. Đối với khối u lõi, mô hình của chúng tôi đạt được điểm cao thứ hai với 0,81. Mặt khác, điểm Dice của toàn bộ khối u và khối u ngấm thuốc cùng xếp thứ ba, tương ứng là 0,88 và 0,76. Kết quả này thấp hơn kết quả của Chang và cộng sự<sup>15</sup> với mô hình dựa trên hai mạng 2D-CNN còn với các nhóm nghiên cứu khác, như kết quả của Kamnitsas và cộng sự<sup>13</sup> tốt hơn ở phần nhận dạng toàn bộ khối u nhưng thấp hơn kết quả của chúng tôi ở khối u lõi và khối u ngấm thuốc; Alqazzaz và cộng sự<sup>17</sup>, sử dụng mô hình SegNet cao hơn với điểm Dice 0,79 cho khối u ngấm thuốc nhưng



Hình 8: Biểu đồ hộp về hiệu suất phân đoạn theo thước đo Specificity (phạm vi từ 0,95 đến 1)

có kết quả thấp hơn hoặc bằng cho hai loại khối u còn lại. So với Casamitjana và cộng sự<sup>16</sup>, kết quả của chúng tôi vượt trội hoàn toàn, đặc biệt ở phân đoạn hai vùng khối u lõi và khối u ngấm thuốc.

Theo kết quả so sánh ở Bảng 5 có thể nhận thấy mô hình đề xuất của chúng tôi tuy không vượt trội hoàn toàn kết quả của các nhóm nghiên cứu khác. Tuy nhiên, như đã giới thiệu ở phần Tổng quan, các mô hình của các nhóm nghiên cứu trên hầu hết đều sử dụng mạng thần kinh tích chập 3D<sup>13,16,17</sup> và áp dụng huấn luyện các mạng riêng lẻ cho từng phương thức của bộ dữ liệu MRI (Flair, T1, T1ce và T2)<sup>16,17</sup> hoặc phải qua nhiều bước (với mạng thần kinh khác nhau) xử lý<sup>13,15</sup>. Sau đó, kết quả phân đoạn với từng phương thức này sẽ được tổng hợp thành kết quả phân đoạn cuối cùng cho các vùng khối u với các thuật toán học máy khác<sup>13,15,17</sup>. Vì vậy, trong nghiên cứu này chúng tôi đề xuất mô hình U-net phân đoạn khối u dựa trên phép toán tích chập cho ảnh 2D với việc tổng hợp các ảnh (slices) từ bốn phương thức MRI và kết quả phân đoạn cuối cùng được tính thông qua hàm argmax. Cho nên, mô hình đề xuất của chúng tôi được huấn luyện và đưa ra kết quả nhanh và đơn giản không phải qua nhiều bước (multi-stage framework) như các nghiên cứu trên. Đặc biệt, mô hình chúng tôi vẫn đạt được kết quả tốt khi phân đoạn vùng khối u lõi và khối u ngấm thuốc là hai vùng có kích thước nhỏ và nằm lẫn với vùng toàn bộ khối u.

## KẾT LUẬN

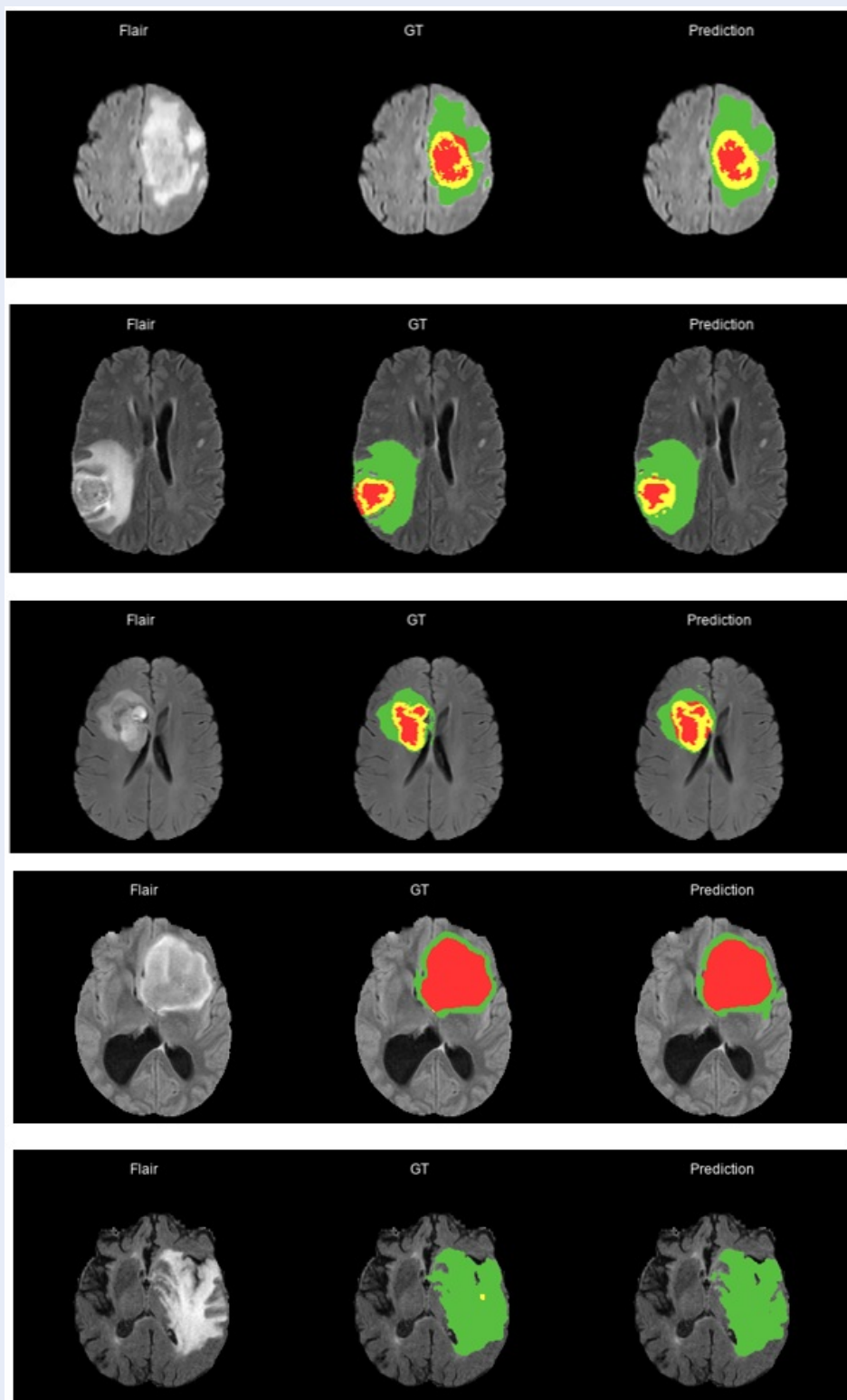
Bài báo này đề xuất mô hình U-net có cải tiến cho phân đoạn u thần kinh đệm với việc thêm các kết nối tắt, lớp chuẩn hóa hàng loạt và thay đổi hàm tuyến tính. Đồng thời, mô hình này được huấn luyện với kỹ thuật k-fold cross validation giúp tìm ra bộ tham số tốt nhất cho các bộ dữ liệu nhỏ. Cùng với đó, các dữ liệu đầu vào được áp dụng các kỹ thuật trích xuất, chuẩn hóa và tăng cường số lượng nhằm tránh cho mô hình bị khớp quá mức trong quá trình huấn luyện với bộ dữ liệu nhỏ. Bộ dữ liệu về u thần kinh đệm BraTS 2018 được sử dụng nhằm đánh giá hiệu suất của mô hình cho bước huấn luyện và kiểm tra. Kết quả cho thấy, độ chính xác với điểm Dice của toàn bộ khối u đạt giá trị cao nhất là 0,88 và của khối u lõi và khối u ngấm thuốc lần lượt là 0,81 và 0,76, điều này cho thấy hiệu suất của mô hình đề xuất tốt khi so với kết quả của các nhóm nghiên cứu khác. Trong tương lai, chúng tôi sẽ thêm một số giải pháp để cải thiện kết quả phân đoạn và đặc biệt là cho khối u ngấm thuốc.

## LỜI CẢM ƠN

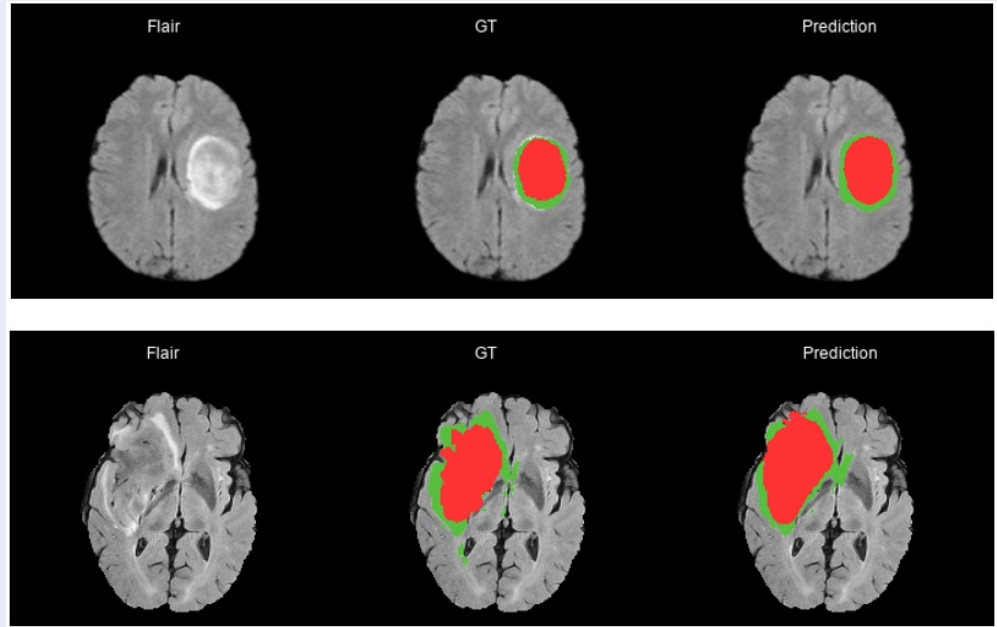
Nghiên cứu được tài trợ bởi Đại học Quốc gia Thành phố Hồ Chí Minh (ĐHQG-HCM) trong khuôn khổ Đề tài mã số DS2020-28-02.

## DANH MỤC TỪ VIẾT TẮT

BraTS: Multimodal Brain Tumor Segmentation Challenge



Hình 9: Một số ví dụ về kết quả phân đoạn dẫn của các mẫu HGG



Hình 10: Một số ví dụ về kết quả phân đoạn của các mẫu LGG

Bảng 5: So sánh kết quả phân đoạn với các nghiên cứu khác gần đây

Các nhóm nghiên cứu	Dice		
	WT	TC	ET
Kamnitsas và cộng sự <sup>13</sup>	0,90	0,75	0,73
Chang và cộng sự <sup>15</sup>	0,89	0,82	0,77
Casamitjana và cộng sự <sup>16</sup>	0,86	0,68	0,67
Alqazzaz và cộng sự <sup>17</sup>	0,85	0,81	0,79
Mô hình đề xuất	0,88	0,81	0,76

HGG: High-grade glioma  
 LGG: Low-grade glioma  
 MRI: Magnetic Resonance Image  
 CT: Computed Tomography  
 PET: Positron Emission Tomography  
 FLAIR: Fluid Attenuated Inversion Recovery  
 CSF: Cerebrospinal Fluid  
 DL: Deep Learning  
 CNN: Convolutional Neural Network  
 WT: Whole Tumor  
 ET: Enhancing Tumor  
 TC: Tumor Core  
 GT: Ground Truth

### XUNG ĐỘT LỢI ÍCH

Nhóm tác giả xin cam đoan rằng không có bất kỳ xung đột lợi ích nào trong công bố bài báo này.

### ĐÓNG GÓP CỦA CÁC TÁC GIẢ

Nguyễn Hoàng Thiên Thư: thu thập dữ liệu, lập trình xử lý và viết bản thảo đầu tiên.

Phạm Thị Thu Hiền: đưa ra ý tưởng, phân tích kết quả và hoàn thiện bản thảo.

Lê Thanh Hải: lập trình xử lý, phân tích kết quả và chỉnh sửa bản thảo.

### TÀI LIỆU THAM KHẢO

- Holland EC. Progenitor cells and glioma formation. *Curr Opin Neurol.* 2001;14(6):683-8; PMID: 11723374. Available from: <https://doi.org/10.1097/00019052-200112000-00002>.
- Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 world health organization classification of tumors of the central nervous system: A summary. *Acta Neuropathol.* 2016;131(6):803-20; PMID: 27157931. Available from: <https://doi.org/10.1007/s00401-016-1545-1>.

3. Ohgaki H, Kleihues P. Population-based studies on incidence, survival rates, and genetic alterations in astrocytic and oligodendroglial gliomas. *J Neuropathol Exp Neurol*. 2005;64(6):479-89;PMID: 15977639. Available from: <https://doi.org/10.1093/jnen/64.6.479>.
4. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. version 1.1. 2009;45(2):228-47;PMID: 19097774. Available from: <https://doi.org/10.1016/j.ejca.2008.10.026>.
5. Wen PY, Macdonald DR, Reardon DA, Cloughesy TF, Sorensen AG, Galanis E, et al. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *J Clin Oncol*. 2010;28(11):1963-72;PMID: 20231676. Available from: <https://doi.org/10.1200/JCO.2009.26.3541>.
6. Juan-Albarracín J, Fuster-García E, Manjón JV, Robles M, Aparici F, Martí-Bonmati L, et al. Automated glioblastoma segmentation based on a multiparametric structured unsupervised classification. *PLOS ONE*. 2015;10(5):e0125143;PMID: 25978453. Available from: <https://doi.org/10.1371/journal.pone.0125143>.
7. Bauer S, Wiest R, Nolte LP, Reyes M. A survey of MRI-based medical image analysis for brain tumor studies. *Phys Med Biol*. 2013;58(13):R97-R129;PMID: 23743802. Available from: <https://doi.org/10.1088/0031-9155/58/13/R97>.
8. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans Med Imaging*. 2015;34(10):1993-2024;PMID: 25494501. Available from: <https://doi.org/10.1109/TMI.2014.2377694>.
9. Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, et al. Brain tumor segmentation with deep neural networks. *Med Image Anal*. 2017;35:18-31;PMID: 27310171. Available from: <https://doi.org/10.1016/j.media.2016.05.004>.
10. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *The 25th International Conference on Advances in neural information processing systems*. 2012;1:1097-105;.
11. Angelini ED, Clatz O, Mandonnet E, Konukoglu E, Capelle L, Duffau H. Glioma dynamics and computational models: a review of segmentation, registration, and in silico growth algorithms and their clinical applications. *Curr Med Imaging*. 2017;3(4):262-76; Available from: <https://doi.org/10.2174/157340507782446241>.
12. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Vol. 2016; 2016. p. 770-8;PMID: 26180094. Available from: <https://doi.org/10.1109/CVPR.2016.90>.
13. Kamnitsas K, Ledig C, Newcombe VFJ, Simpson JP, Kane AD, Menon DK, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal*. 2017;36:61-78;PMID: 27865153. Available from: <https://doi.org/10.1016/j.media.2016.10.004>.
14. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science. International Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI*. 2015:234-41; Available from: [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
15. Chang Y-J, Lin Z-S, Yang T-L, Huang T-Y. Automatic segmentation of brain tumor from 3D MR images using a 2D convolutional neural network. *2018 International MICCAI BraTS Challenge*. 2018:83-90; Available from: [https://doi.org/10.1007/978-3-030-46643-5\\_22](https://doi.org/10.1007/978-3-030-46643-5_22).
16. Casamitjana A, Catà M, Sánchez I, Combalia M, Vilaplana V. Cascaded V-Net using ROI masks for brain tumor segmentation. *Lecture Notes in Computer Science*. 2018;10670:381-91; Available from: [https://doi.org/10.1007/978-3-319-75238-9\\_33](https://doi.org/10.1007/978-3-319-75238-9_33).
17. Alqazzaz S, Sun X, Yang X, Nokes L. Automated brain tumor segmentation on multi-modal MR image using SegNet. *Comp Visual Media*. 2019;5(2):209-19; Available from: <https://doi.org/10.1007/s41095-019-0139-y>.
18. Bakas S, Akbari H, Sotiras A, et al. Segmentation labels and radiomic features for the preoperative scans of the TCGA-GBM collection. *The Cancer Imaging Archive*. 2017;.
19. Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby JS, et al. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci Data*. 2017;4:170117;PMID: 28872634. Available from: <https://doi.org/10.1038/sdata.2017.117>.
20. Imai H, Matzek S, Le TD, et al. Fast and accurate 3D medical image segmentation with data-swapping method. *arXiv preprint: 1812.07816*; 2018; Available from: <https://arxiv.org/abs/1812.07816>.
21. Isensee F, Kickingereder P, Wick W, Bendszus M, Maier-Hein KH. Brain tumor segmentation and radiomics survival prediction: contribution to the BRATS 2017 challenge. *Lecture Notes in Computer Science*. 2018;10670:287-97; Available from: 10.1007/978-3-319-75238-9\_25.
22. Kerimi A, Mahmoudi I, Khadir MT. Deep convolutional neural networks using U-Net for automatic brain tumor segmentation in multimodal MRI volumes. *Lecture Notes in Computer Science*. 2019;11384:37-48; Available from: [https://doi.org/10.1007/978-3-030-11726-9\\_4](https://doi.org/10.1007/978-3-030-11726-9_4).
23. Perez L, Wang J. The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint: 1712.04621*; 2017; Available from: <https://arxiv.org/abs/1712.04621>.
24. Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: A simple way to prevent neural networks from overfitting. *J Mach Learn Res*. 2014;15(56):1929-58;.
25. He K, Zhang X, Ren S, et al. Delving deep into rectifiers: surpassing human-level performance on imagenet classification *IEEE International Conference on Computer Vision (ICCV)*. Vol. 2015; 2015. p. 1026-34; Available from: <https://doi.org/10.1109/ICCV.2015.123>.

# Application of deep learning in brain tumor segmentation

Hoang Thien Thu Nguyen<sup>1,2</sup>, Thi Thu Hien Pham<sup>1,2,\*</sup>, Thanh Hai Le<sup>2,3,\*</sup>



Use your smartphone to scan this QR code and download this article

## ABSTRACT

Glioma is one of the most common brain tumor worldwide. Nowadays, glioma identification and segmentation is essential when making clinical decisions. Obviously, manual segmentation is not only time consuming but also subjective, in addition to this task, it is rather difficult to solve the solution of the automated segmentation methods. U-net architecture is one of the popular deep learning models for segmentation of biomedical images. This study implements a 2 dimensional convolutional neural network based on a U-net architecture. The proposed model is trained only with 2D images (slices) from 3D MRI dataset to segment the different tumor regions. Furthermore, in the preprocessing step some methods are applied to deal with imbalanced data for three brain tumor regions, and then data augmentation is utilized to prevent overfitting when training with a small dataset. The evaluated results based on the Multimodal Brain Tumor Segmentation Challenge (BraTS) 2018 dataset get the high accuracy for Dice scores of whole tumor, tumor core and enhancing tumor being 0.88, 0.81 and 0.76, respectively. The performance of the proposed model show that the brain tumors can be recognized with high accuracy compared to the similar studies. The proposed study has provided an analytical procedure for glioma brain tumors, from preprocessing to the training model, and then assessing factors that may affect the performance of the achieved results.

**Key words:** Brain tumor, Deep neural network, Glioma, Segmentation, U-net

<sup>1</sup>School of Biomedical Engineering, International University, Vietnam

<sup>2</sup>Vietnam National University HCMC, Vietnam, 700000

<sup>3</sup>Faculty of Mechanical Engineering, Ho Chi Minh City University of Technology, Vietnam

## Correspondence

**Thi Thu Hien Pham**, School of Biomedical Engineering, International University, Vietnam

Vietnam National University HCMC, Vietnam, 700000

Email: ptthien@hcmiu.edu.vn

## Correspondence

**Thanh Hai Le**, Vietnam National University HCMC, Vietnam, 700000

Faculty of Mechanical Engineering, Ho Chi Minh City University of Technology, Vietnam

Email: lthai@hcmut.edu.vn

## History

- Received: 09-12-2021
- Accepted: 07-6-2022 •
- Published: 30-6-2022

DOI : 10.32508/stdjet.v5i2.951



**Cite this article :** Nguyen H T T, Pham T T H, Le T H. **Application of deep learning in brain tumor segmentation.** *Sci. Tech. Dev. J. – Engineering and Technology*; 5(2):1437-1450.