

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG
KHOA CÔNG NGHỆ THÔNG TIN 1



BÀI CÁO BÀI TẬP LỚN
MÔN XỬ LÝ ẢNH

ĐỀ TÀI

**Phân tích thuật toán xử lý ảnh trong bài báo
“Automated Segmentation of Brain Tumor MRI
Images using Deep Learning”**

Giáo viên hướng dẫn: TS. Phạm Hoàng Việt

Nhóm học phần: 02

Nhóm bài tập lớn: 21

Danh sách sinh viên:

STT	Họ và tên	Mã SV
1	Nguyễn Quang Dũng	B22DCCN133
2	Cao Thị Thu Hương	B22DCCN422

Hà Nội, 2025

Mục lục

Danh mục hình ảnh	1
1 Giới thiệu	2
1.1 Thông tin bài báo	2
1.2 Giới thiệu bài toán và nghiệp vụ liên quan	2
1.2.1 Ảnh MRI Laf gì?	2
1.2.2 Các loại ảnh MRI (Modalities)	2
1.2.3 Giới thiệu về bài toán Segmentation u não	2
1.3 Tổng quan	3
2 Cơ sở lý thuyết	4
2.1 Ma trận đồng xuất hiện mức xám (GLCM)	4
2.2 Cây điểm thuận lợi (VPT)	5
2.3 Mạng Nơ-ron Tích chập 3D (3D-CNN)	6
2.4 Mạng U-Net 3D	6
3 Phương pháp	8
4 Demo và Kết quả	10
4.1 Demo	10
4.2 Kết quả	10
5 Kết luận	10

Danh mục hình ảnh

1	Ảnh 4 kênh màu của MRI	2
2	GLCM's Direction Distance	4
3	Vantage Point Tree Construction	5
4	Architecture 3 D-Convolution Neural Network	6
5	Architecture of 3D U-Net	7
6	Full pipeline	8
7	Cơ chế Ensemble	9
8	Kết quả 1	10
9	Kết quả 2	10

1 Giới thiệu

1.1 Thông tin bài báo

- Tiêu đề: **Automated Segmentation of Brain Tumor MRI Images using Deep Learning**
- Tác giả: *Surendran Rajendran, Suresh Kumar Rajagopal, Tamilvizhi Thanarajan, K. Shankar, Sachin Kumar, Najah Alsubaie, Mohamad Khairi Ishak, Samih M. Mostafa*
- Tạp chí: IEEE Access
- Thời gian công bố: 20/07/2023
- Điểm Scopus: 0.849 - Q1, nguồn: SJR
- Link bài báo: Automated Segmentation of Brain Tumor MRI Images using Deep Learning

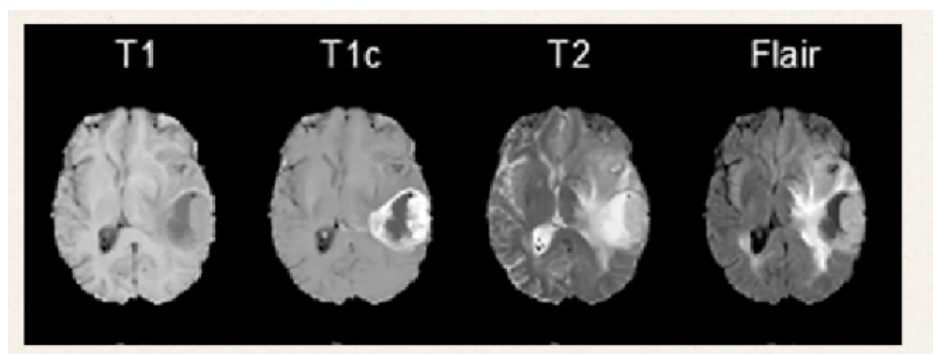
1.2 Giới thiệu bài toán và nghiệp vụ liên quan

1.2.1 Ảnh MRI Là gì?

- Kỹ thuật chẩn đoán hình ảnh không dùng tia X, cho hình ảnh não 3D độ phân giải cao.
- Rất hiệu quả trong phát hiện và phân tích u não.

1.2.2 Các loại ảnh MRI (Modalities)

- T1: Thông tin giải phẫu, mô mỡ sáng, dịch tối.
- T1ce: T1 có tiêm tương phản → vùng u tăng sinh sáng rõ.
- T2: Nhạy với nước → phù (edema) sáng.
- FLAIR: T2 loại bỏ dịch CSF → rõ tổn thương cạnh não thất.



Hình 1: Ảnh 4 kênh màu của MRI

1.2.3 Giới thiệu về bài toán Segmentation u não

- Vai trò của Segmentation:

- Tự động xác định ranh giới và vùng bất thường trong ảnh MRI.
- Hỗ trợ chẩn đoán, đánh giá mức độ ác tính, theo dõi tiến triển điều trị.
- Quan trọng trong lập kế hoạch phẫu thuật, xạ trị và đo đặc kích thước khối u.

- Thách thức trong phân đoạn u não:

- Hình dạng – kích thước – vị trí thay đổi lớn giữa các bệnh nhân.
- Cường độ tín hiệu không đồng nhất (inhomogeneity) do MRI từ nhiều máy, nhiều điều kiện khác nhau.
- Nhiều, artifact và cấu trúc não phức tạp gây khó khăn cho mô hình.
- Dữ liệu hạn chế, đặc biệt trong y tế \rightarrow khó huấn luyện mô hình sâu.

1.3 Tổng quan

Bài báo trình bày một phương pháp mới để tự động phân đoạn khối u não từ dữ liệu MRI, điều này rất quan trọng cho việc điều tra và theo dõi bệnh.

- **Vấn đề:** Các khối u thần kinh đệm (gliomas) có tính chất xâm lấn và đa dạng, khiến việc phân đoạn thủ công trở nên tốn kém, tốn thời gian và đòi hỏi bác sĩ có chuyên môn cao.
- **Giải pháp:** Các tác giả đề xuất một kỹ thuật gọi là ASBTCNN, sử dụng chiến lược kết hợp giữa hai mạng nơ-ron riêng biệt để mang lại kết quả chính xác hơn. Cụ thể ở đây là một mạng U-Net và một mạng 3-D CNN.

2 Cơ sở lý thuyết

Phần này trình bày chi tiết các thành phần cốt lõi của khung làm việc được đề xuất: trích xuất đặc trưng kết cấu sử dụng GLCM, lựa chọn đặc trưng thông qua VPT, và kiến trúc học sâu kép (3D-CNN và U-Net) được sử dụng cho phân đoạn thể tích.

2.1 Ma trận đồng xuất hiện mức xám (GLCM)

Để nắm bắt sự không đồng nhất về kết cấu của các mô khối u, chúng tôi sử dụng Ma trận đồng xuất hiện mức xám (GLCM). GLCM là một phương pháp thống kê bậc hai mô tả mối quan hệ không gian giữa các pixel bằng cách tính toán tần suất xuất hiện của một cặp pixel với các giá trị cụ thể và hướng không gian xác định trong ảnh.

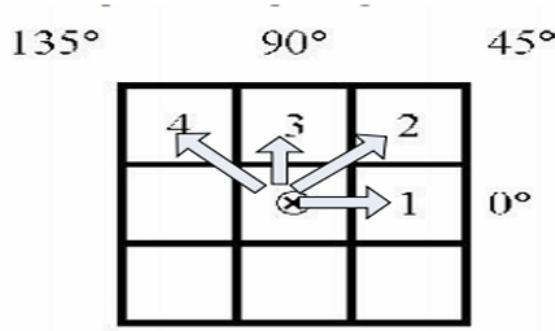


FIGURE 1. GLCM's Direction Distance

Hình 2: GLCM's Direction Distance

Cho một ảnh I , GLCM $P(i, j|d, \theta)$ biểu diễn xác suất chuyển từ mức xám i sang mức xám j tại khoảng cách d theo hướng θ . Các hướng thường được xem xét là $0^\circ, 45^\circ, 90^\circ$, và 135° . Từ ma trận GLCM đã chuẩn hóa, các đặc trưng thống kê được trích xuất để đại diện cho kết cấu của các lát cắt MRI:

- **Năng lượng (Energy):** Đo lường tính đồng nhất của kết cấu.

$$Energy = \sum_{i,j} P(i, j)^2 \quad (1)$$

- **Độ đồng nhất (Homogeneity):** Đo lường sự gần gũi của phân bố các phần tử trong GLCM so với đường chéo chính.

$$Homogeneity = \sum_{i,j} \frac{P(i, j)}{1 + |i - j|} \quad (2)$$

- **Độ tương phản (Contrast):** Đo lường các biến thiên cục bộ hiện diện trong ảnh.

$$Contrast = \sum_{i,j} (i - j)^2 P(i, j) \quad (3)$$

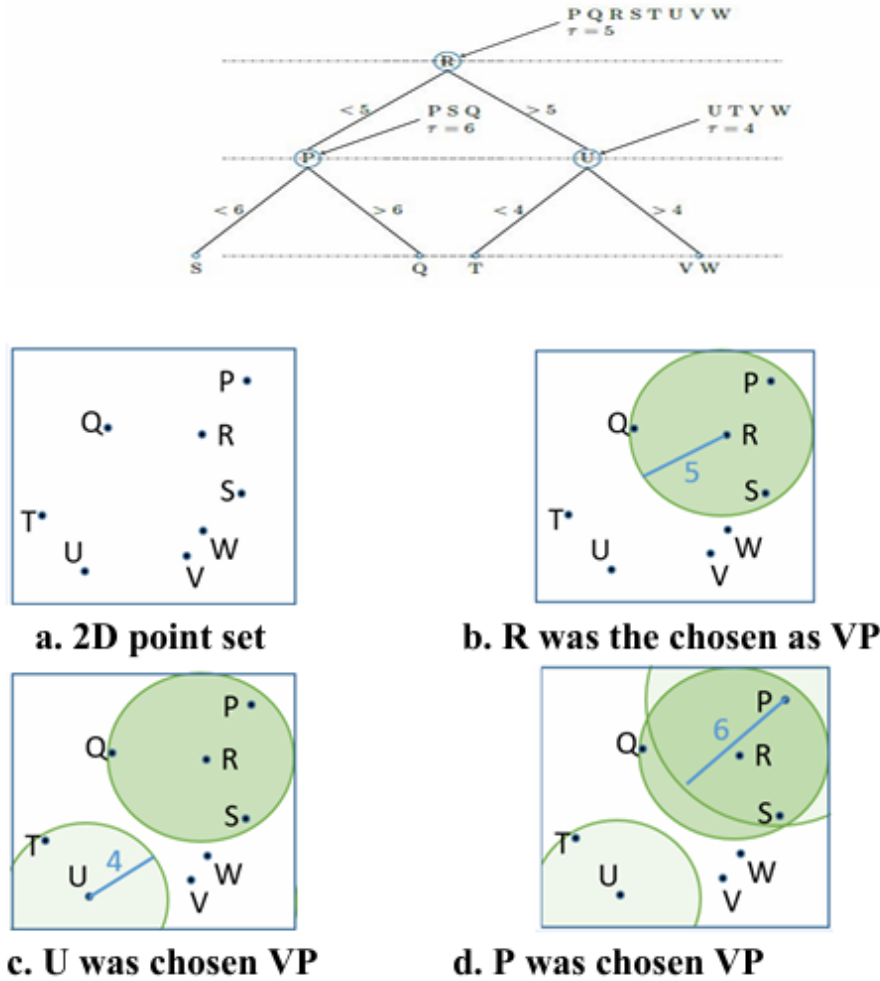
- **Độ tương quan (Correlation):** Đo lường xác suất xuất hiện đồng thời của các cặp pixel được chỉ định.

$$Correlation = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)P(i, j)}{\sigma_i \sigma_j} \quad (4)$$

Các đặc trưng này được nối (concatenate) với các kênh MRI gốc để tạo thành một tensor đặc trưng nhiều chiều, cung cấp cho mạng thông tin kết cấu tường minh nhằm cải thiện độ chính xác phân loại.

2.2 Cây điểm thuận lợi (VPT)

Cây điểm thuận lợi (VPT) được sử dụng để tìm kiếm lân cận gần nhất hiệu quả và lựa chọn đặc trưng trong không gian nhiều chiều được tạo ra bởi GLCM. Nó tổ chức các điểm dữ liệu vào một cấu trúc cây metric dựa trên khoảng cách của chúng đến một "điểm thuận lợi"(vantage point) được chọn.



Hình 3: Vantage Point Tree Construction

Quá trình xây dựng VPT bao gồm các bước đệ quy:

1. Một điểm cụ thể, gọi là Điểm thuận lợi (VP), được chọn từ tập dữ liệu.
2. Khoảng cách từ VP đến tất cả các điểm khác được tính toán bằng thước đo Euclidean.
3. Không gian dữ liệu được phân chia thành các không gian con dựa trên bán kính ngưỡng trung vị (median threshold radius).

Về mặt toán học, độ tương đồng giữa các điểm dữ liệu trong không gian đặc trưng được xác định bằng hàm kernel Gaussian:

$$D(k, l) = \exp\left(-\frac{\|U_k - U_l\|^2}{2\sigma^2}\right) \quad (5)$$

Trong đó U_k và U_l là các vector đặc trưng, và σ kiểm soát độ rộng của vùng lân cận. Bậc (degree) của một điểm dữ liệu được định nghĩa là:

$$M_k = \sum_{l=1}^n D_{k,l} \quad (6)$$

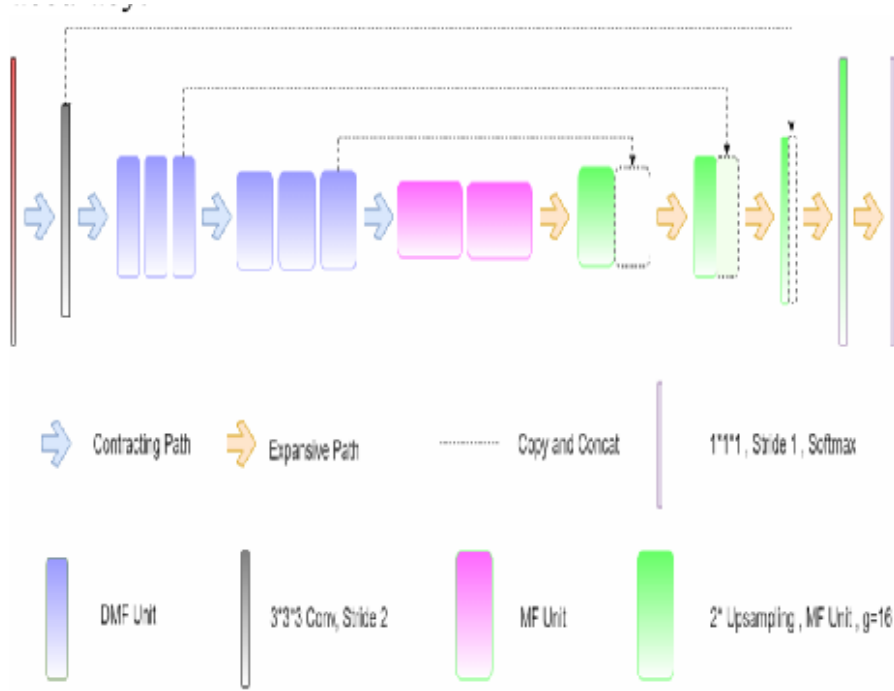
Dựa trên các giá trị này, ma trận Laplacian chưa chuẩn hóa L được tạo ra:

$$L = D_I - M \quad (7)$$

Trong đó D_I là ma trận đường chéo. Cấu trúc này cho phép hệ thống lọc bỏ các đặc trưng không liên quan và chọn ra các vector đại diện nhất cho các giai đoạn phân loại tiếp theo.

2.3 Mạng Nơ-ron Tích chập 3D (3D-CNN)

Thành phần học sâu đầu tiên trong mô hình tập hợp là một mạng 3D-CNN được thiết kế để xử lý các vùng (patches) thể tích. Kiến trúc bao gồm các lớp tích chập theo sau là các lớp gộp cực đại (max-pooling) và mạng nơ-ron kết nối đầy đủ.



Hình 4: Architecture 3 D-Convolution Neural Network

Mạng sử dụng các kernel 3D để trích xuất đặc trưng không gian-thời gian. Giá trị đầu ra y tại vị trí (x, y, z) trên bản đồ đặc trưng thứ j của lớp thứ i được tính toán như sau:

$$\gamma_{j,xyz}^i = \text{ReLU} \left(b_j^{(i)} + \sum_m \sum_{u,v,w} w_{jm,uvw}^{(i)} \gamma_{m,(x+u)(y+v)(z+w)}^{(i-1)} \right) \quad (8)$$

Trong đó w đại diện cho trọng số kernel và b là bias. Mỗi quan hệ giữa hai lớp lân cận được mô tả bởi:

$$\gamma^{(i)} = \sigma(W^{(i)}\gamma^{(i-1)} + b^{(i)}) \quad (9)$$

Mạng được tối ưu hóa bằng cách giảm thiểu hàm mất mát Sai số Bình phương Trung bình (MSE):

$$L(W, b|D) = \frac{1}{n} \sum_{k=1}^n ||y_{truth}^{(k)} - y_{pred}^{(k)}||^2 \quad (10)$$

2.4 Mạng U-Net 3D

Mô hình thứ hai là 3D U-Net, có cấu trúc bộ mã hóa-giải mã (encoder-decoder) với các kết nối nhảy (skip connections), được tối ưu hóa cho phân đoạn hình ảnh y sinh. Mạng bao gồm:

- **Đường dẫn thu hẹp (Contracting Path):** Các khối tích chập lặp lại theo sau là max-pooling để nắm bắt thông tin ngữ cảnh.

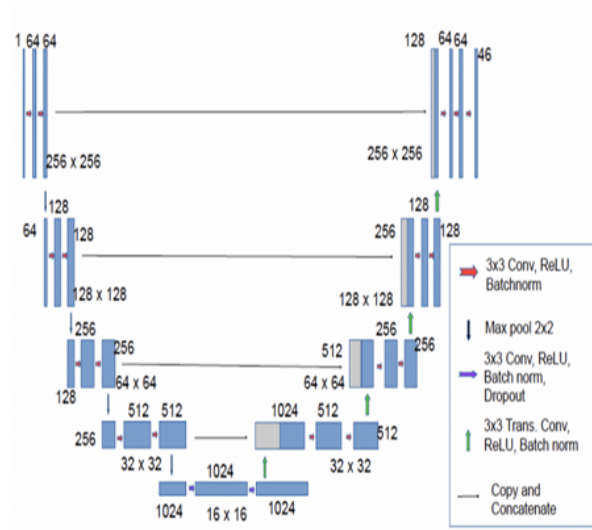
- **Đường dẫn mở rộng (Expansive Path):** Các hoạt động lấy mẫu lên (up-sampling) theo sau bởi tích chập để khôi phục độ phân giải không gian.
- **Kết nối nhảy (Skip Connections):** Nối các bản đồ đặc trưng từ đường dẫn thu hẹp với đường dẫn mở rộng để bảo tồn các chi tiết cục bộ.
- **Cải tiến:** Thay thế ReLU bằng Leaky ReLU và Batch Normalization bằng Instance Normalization.

Xác suất hậu nghiệm của một voxel i được gán nhãn với lớp l cụ thể được tính toán bằng hàm Softmax:

$$p(y_i = l | M_i) = \frac{\exp(f_l(M_i))}{\sum_{k'} \exp(f_{k'}(M_i))} \quad (11)$$

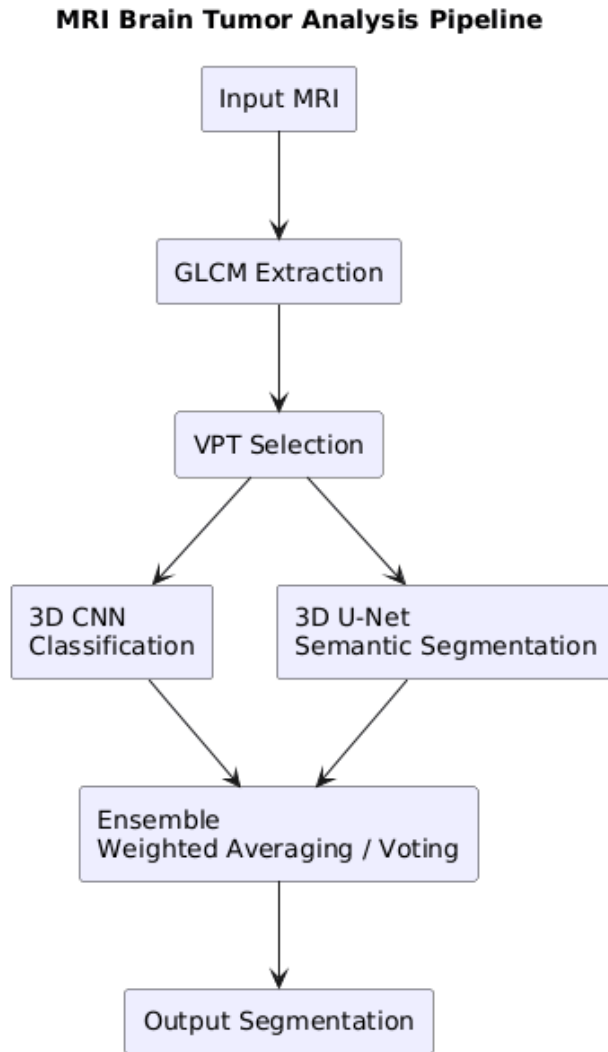
Trong đó $f(\cdot)$ là hàm tính toán của CNN. Mô hình được huấn luyện sử dụng hàm mất mát cross-entropy có trọng số:

$$Loss = - \sum_i \log(p(y_i = \text{groundtruth} | X(M_i))) \quad (12)$$



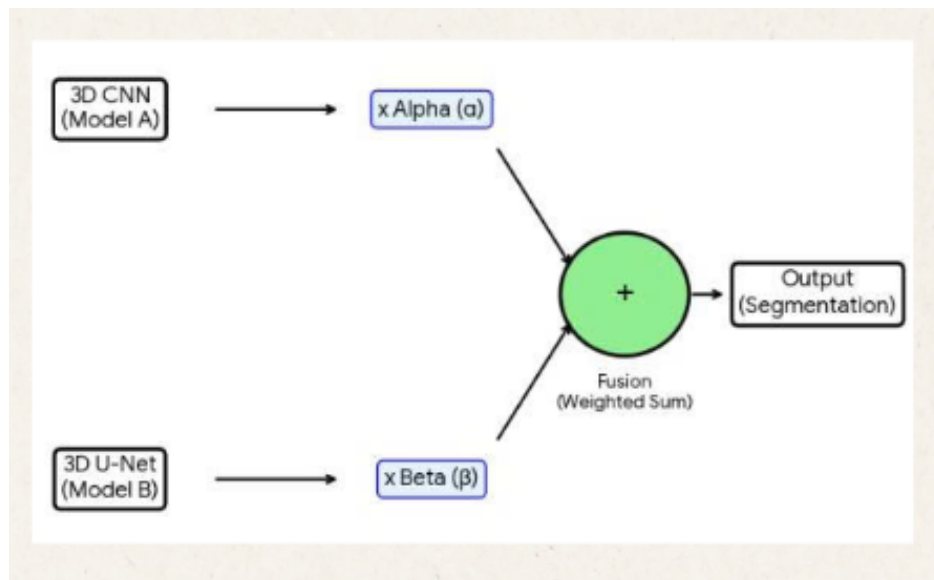
Hình 5: Architecture of 3D U-Net

3 Phương pháp



Hình 6: Full pipeline

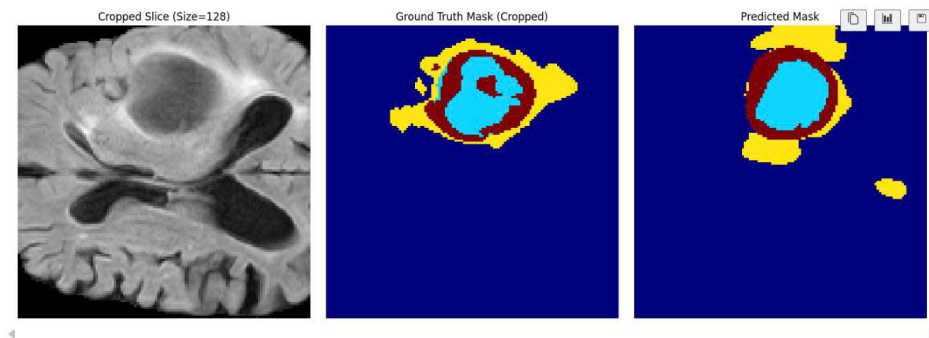
1. **Tiền xử lý ảnh MRI:** Chuẩn hóa kích thước, tăng cường dữ liệu (crop, rotate, mirror).
2. **Trích xuất đặc trưng kết cấu bằng GLCM:** Tính Energy, Homogeneity, Contrast,... giúp làm nổi bật mô bất thường
3. **VPT – Chọn vùng quan trọng:** Tìm lân cận gần và lọc đặc trưng nhanh, giúp giảm nhiễu.
4. **3D-CNN:** Học đặc trưng thể tích (3 chiều) của ảnh MRI.
5. **3D U-Net:** Mạng segmentation chính với skip-connections, tái tạo chi tiết tốt.
6. **Ensemble:** Kết hợp 2 mô hình -> tăng Dice và Sensitivity.



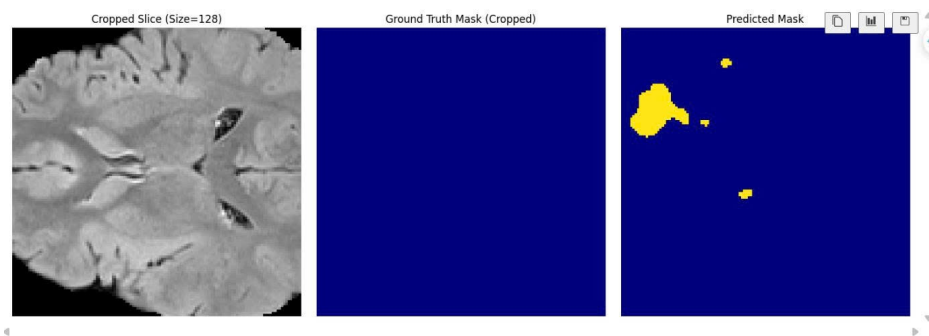
Hình 7: Cơ chế Ensemble

4 Demo và Kết quả

4.1 Demo



Hình 8: Kết quả 1



Hình 9: Kết quả 2

4.2 Kết quả

- F-SCORE:

- Whole Tumor: 68.07%
- Tumor Core: 73.34%
- Enhancing Tumor: 70.23%

- DICE COEFFICIENT: 68.12% Nhận xét kết quả:

- 3D U-Net phân đoạn tốt vùng Tumor Core nhờ skip-connections.
- 3D-CNN mạnh ở nhận diện cấu trúc khối 3D → tốt cho Enhancing Tumor.
- Ensemble cho kết quả cao nhất, ổn định hơn từng mô hình đơn lẻ.

5 Kết luận

Tóm tắt

- Bài báo đề xuất phương pháp ASBTCNN, kết hợp GLCM → VPT → 3D-CNN → 3D U-Net → Ensemble.
- Mục tiêu: phân đoạn tự động 3 vùng u não trên MRI với độ chính xác cao.

Hiệu quả đạt được

- F-Score & Sensitivity $> 65\%$.
- Ensemble giữa 3D-CNN và U-Net tạo kết quả ổn định và chính xác hơn.

ý nghĩa

- Khả năng áp dụng cho hỗ trợ chẩn đoán y khoa.
- Là hướng tiếp cận mạnh trong segmentation ảnh y sinh.