



DOI:10.22144/ctujos.2024.393

PHÁT HIỆN VÀ HIỂN THỊ 3D VÙNG BẤT THƯỜNG TRÊN ẢNH MRI NÃO VỚI CÔNG DỊCH VỤ BILLOW AISA

Lê Minh Lợi¹, Trần Nguyễn Minh Thu^{2*}, Nguyễn Thiện Hùng², Hồ Quốc An² và Phạm Nguyên Khang²

¹Trường Đại học Y Dược Cần Thơ

²Trường Đại học Cần Thơ

*Tác giả liên hệ (Corresponding author): tnmthu@ctu.edu.vn

Thông tin chung (Article Information)

Nhận bài (Received): 11/05/2024

Sửa bài (Revised): 22/07/2024

Duyệt đăng (Accepted): 24/08/2024

Title: Detecting and displaying 3D abnormal areas on MRI images with service Billow AISA

Author(s): Le-Minh Loi¹, Tran Nguyen Minh Thu^{2*}, Nguyen Thien Hung², Ho Quoc An² and Pham Nguyen Khang²

Affiliation(s): ¹Can Tho University of Medicine and Pharmacy, ²Can Tho University

TÓM TẮT

Việc phát hiện kịp thời khối u hỗ trợ các bác sĩ trong quá trình chẩn đoán và điều trị cho bệnh nhân được thực hiện hiệu quả trong tình trạng các bệnh viện luôn quá tải là rất cần thiết. Ứng dụng Slicer cho phép dựng hình ảnh 2D vùng tổn thương thành dữ liệu khối 3D giúp các bác sĩ có cái nhìn trực quan hơn trong việc chẩn đoán và điều trị. Tuy nhiên, ứng dụng Slicer chưa cho phép phát hiện tự động vùng bất thường và yêu cầu máy tính đủ mạnh để thực thi các mô hình này. Trong nghiên cứu này, tiện ích mở rộng Billow AISA cho Slicer được đề xuất nhằm xây dựng một công dịch vụ phân tích, dự đoán từ dữ liệu ảnh do người dùng cung cấp. Chức năng phân tích, dự đoán được thử nghiệm trong nghiên cứu này là phát hiện vùng bất thường trên ảnh MRI não với mô hình Swin-Unet. Kết quả thực nghiệm trên tập dữ liệu thu thập từ Bệnh viện Trường Đại học Y Dược Cần Thơ cho thấy tính khả thi và hiệu quả của mô hình Billow AISA.

Từ khóa: Phát hiện vùng bất thường, hiển thị hình ảnh ba chiều, Swin-Unet

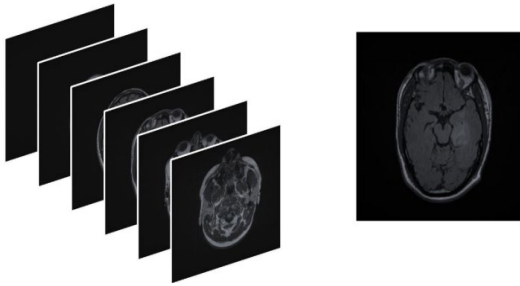
ABSTRACT

Timely detection of tumors to support doctors in effectively diagnosing and treating patients in overloaded hospitals is essential. The Slicer application allows the rendering of 2D images of damaged areas into 3D volume data, helping doctors obtain a more intuitive view of diagnosis and treatment. However, the Slicer application does not yet allow automatic detection of abnormal regions and requires a powerful computer to execute these models. In this research, Slicer's Billow AISA extension is proposed to build an analysis and prediction service portal from image data provided by users. The analysis and prediction function experimented in this research is the detection of abnormal areas on brain MRI images using the Swin-Unet model. Experimental results on the dataset collected from Can Tho University of Medicine and Pharmacy Hospital demonstrate the feasibility and effectiveness of the Billow AISA model.

Keywords: Detecting abnormalities, 3D rendering, Swin-Unet

1. GIỚI THIỆU

Ảnh MRI não của một người là một tập ảnh xám đơn lẻ được tạo thành bằng cách chụp nhiều ảnh cắt ngang sọ não người từ đỉnh đầu xuống vừa tới cổ, mỗi một ảnh đơn lẻ gọi là một lát cắt (slice) như Hình 1. Vùng tổn thương thường có cấu trúc, hình dạng hoặc màu sắc khác với vùng ảnh thông thường. Để có thể chẩn đoán vùng bất thường, các bác sĩ phải khảo sát bằng mắt thường qua từng ảnh một trong các tập ảnh kết quả của mỗi bệnh nhân. Việc khảo sát thủ công như vậy rất phức tạp và tốn thời gian để phát hiện được vùng bất thường.



Hình 1. Một số lát cắt MRI của một người

Dựa vào các giải thuật máy học, các nhà nghiên cứu (Held et al., 1997; Tsai et al., 2003) đã hỗ trợ bác sĩ phát hiện vùng bất thường. Với sự phát triển của các mô hình học sâu, Unet được đề xuất (Ronneberger et al., 2015; Lợi và cv., 2020) để phân đoạn hình ảnh nhằm phát hiện các vùng bất thường. Với sự hiệu quả của cấu trúc hình chữ U, các biến thể Unet được đề xuất như ResNet (Xiao et al., 2018), Unet ++ (Zhou et al., 2018), DenseUNet (Li et al., 2018) và 3dUnet (Çiçek et al., 2016).

Các mô hình Transformer trong lĩnh vực xử lý ngôn ngữ tự nhiên (Natural Language Processing - NLP) với những thành công đáng kể cũng được ứng dụng trong lĩnh vực thị giác máy tính (Carion et al., 2020). Trong nghiên cứu của Dosovitskiy et al. (2021), tác giả đã đề xuất một Transformer tiên phong (Vision Transformer - ViT) để phân đoạn hình ảnh. ViT đạt kết quả khá tốt so với các phương pháp dựa trên mạng tích chập (Convolutional Neuron Network - CNN). Tuy nhiên, nhược điểm của ViT là mô hình phải được huấn luyện trước với lượng dữ liệu lớn mới đạt được kết quả tốt. Việc áp dụng ViT vào trong lĩnh vực y tế sẽ gặp nhiều khó khăn vì dữ liệu ảnh y tế được gán nhãn sẵn có số lượng hạn chế. Để khắc phục nhược điểm này, mô hình Swin Unet Transformer (Hatamizadeh et al., 2022) được đề xuất để phân đoạn ảnh sử dụng kiến trúc chữ U với Swin Transformer làm bộ mã hóa kết nối với bộ giải mã CNN tại các giai đoạn khác nhau

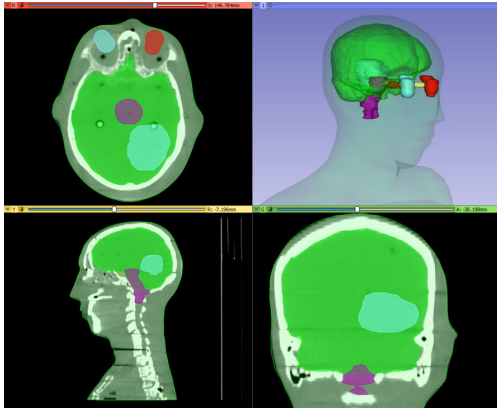
thông qua các kết nối. Mô hình được dùng để phân đoạn khối u não từ tập ảnh 3D BraTS 2021. Trong nghiên cứu của Lợi và ctv. (2024), mô hình Swin-Unet được thực nghiệm và so sánh, đánh giá với mô hình Unet và ResNet nhằm phát hiện các vùng bất thường trên ảnh MRI não. Mô hình đề xuất được huấn luyện trên tập dữ liệu LGG của 110 bệnh nhân với 3929 ảnh MRI (2437 ảnh không chứa vùng bất thường, 1373 ảnh có chứa vùng bất thường) kết hợp với tập dữ liệu đã được thu thập và gán nhãn của 106 bệnh nhân (1551 ảnh bình thường và 604 ảnh bất thường) tại Bệnh viện Trường Đại học Y Dược Cần Thơ (Thư và ctv., 2022). Kết quả cho thấy mô hình Swin-Unet với một số tùy chỉnh khi áp dụng vào ảnh MRI não cũng đạt hiệu quả cao hơn nghiên cứu trước đó (Cao et al., 2023) cho ảnh CT bụng và MRI tim.

Việc phát hiện kịp thời khối u hỗ trợ các bác sĩ trong quá trình chẩn đoán và điều trị cho bệnh nhân được thực hiện hiệu quả trong tình trạng các bệnh viện luôn quá tải là rất cần thiết. Vùng tổn thương sau khi được khoan vùng nếu được thể hiện dưới góc nhìn ba chiều (3D) giúp các bác sĩ có cái nhìn trực quan hơn trong việc chẩn đoán và điều trị. Các công cụ dựng ảnh 3D từ nhiều ảnh 3D là những nghiên cứu thiết thực hỗ trợ quá trình chẩn đoán được hiệu quả hơn. Một số ứng dụng, thư viện sử dụng các giải thuật hình học cơ bản như Marching-Cube, Ray-Casting, Texture rendering để xây dựng, hiển thị hình ảnh với định dạng ba chiều trực quan hoá thông tin như VIT, Santosoft, 3D-DOCTOR,... Mô hình đối tượng của VTK gồm 2 thành phần: mô hình đồ họa (graphics model) và mô hình trực quan hóa (visualization model).

Các mô hình học sâu cũng được ứng dụng để tái tạo các mô hình 3D. Phương pháp sử dụng mạng nơron tái tạo hồi quy và mạng nơron tích chập nhằm tạo một kiến trúc mới là mạng nơron tái tạo hồi quy 3D (3D Recurrent Reconstruction Neural Network - 3D-R2N2) (Chris et al., 2016). Mạng 3D-R2N2 dùng một hoặc nhiều hình ảnh của một đối tượng từ các góc nhìn khác nhau và đưa ra một bản dựng lại của đối tượng dưới dạng lưới. Ưu điểm lớn nhất của mạng R2N2 trong cả huấn luyện và thử nghiệm là không yêu cầu có bất kỳ nhãn của lớp và cho phép tái tạo 3D các đối tượng trong các tình huống thiếu thông tin kết cấu. Trong một nghiên cứu khác (Nhường và ctv., 2021), các tác giả sử dụng mạng nơron tái tạo hồi quy để tái tạo mô hình 3D của đối tượng thông qua bản phác thảo 2,5D. Mô hình gồm ba bước: ước tính các bản phác thảo 2,5D, ước tính hình dạng 3D và tinh chỉnh hình dạng 3D. Kết quả thực nghiệm được thực hiện trên bộ dữ liệu ShapNet

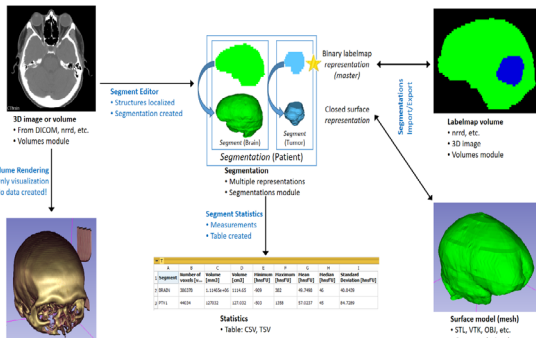
Core55 được cung cấp bởi học viện MIT, tổng số đối tượng trong bộ dữ liệu cả huấn luyện và thử nghiệm là 5.652 đối tượng, trong đó ghê là 1.816, xe hơi là 1.906, máy bay là 1.930.

Ứng dụng Slicer cho phép dựng hình ảnh 2D thành dữ liệu khối 3D và kết hợp các tính năng dựng cấu trúc 3D một cách trực quan với nhiều góc nhìn. Thư viện VTK (The Visualization ToolKit) (Ken et al., 2015) được sử dụng trong Slicer để tạo hình ảnh ba chiều. Thư viện VTK là một bộ công cụ mã nguồn mở cung cấp nhiều hàm, thuật toán xử lý ảnh và trực quan hóa. Slicer hỗ trợ rất nhiều phương pháp chuyển đổi và hiển thị ảnh như Marching cube, Ray casting; Texture-base rendering,... Hình 2 là một ví dụ minh họa hiển thị ba chiều của ảnh MRI. Ảnh MRI hai chiều được trích ra từ tập các lát cắt MRI thông qua Slicer tạo thành dữ liệu ba chiều. Slicer có tính năng để hiển thị cấu trúc 3D một cách trực quan với nhiều góc nhìn 2D, 3D khác nhau.



Hình 2. Hình ảnh hiển thị bởi ứng dụng Slicer

Ứng dụng Slicer cho phép chỉnh sửa các phân vùng ảnh thủ công khá hiệu quả và trực quan. Trong Hình 3, Slicer cho phép người dùng đánh dấu nhãn, phân vùng cho từng ảnh và hiển thị ba chiều theo nhãn đã đánh dấu cùng với một số bộ lọc.



Hình 3. Các bước chuyển đổi từ Ảnh MRI sang mô hình 3D bằng Slicer

Ứng dụng Slicer cũng cho phép mở rộng, tích hợp các tính năng bên ngoài theo dạng module mở rộng tạo điều kiện cho việc nhúng và phát triển các chương trình phát hiện, đánh dấu các phân vùng trên ảnh. Hiện tại, ứng dụng Slicer đã có nhiều tiện ích và thư viện hỗ trợ máy học, ... để phân tích các vùng bộ phận trên cơ thể người như extension TotalSegmentator, SurfaceWrapSolidify (Weidert et al., 2020), NVIDIA-AIAX, ...

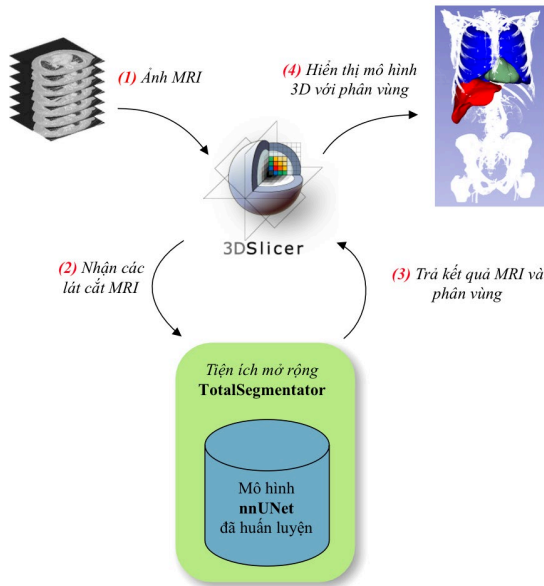
Tuy nhiên, khi sử dụng các thư viện tích hợp này thì việc phân tích dữ liệu hầu hết được xử lý tại máy tính cá nhân của người dùng. Quá trình thực thi này yêu cầu tài nguyên máy tính cá nhân phải có cấu hình mạnh. Ví dụ, nếu phân tích một tập dữ liệu ở vùng bụng với module TotalSegmentator (chế độ phân tích nhanh) thì cần khoảng hơn 45 phút và chế độ đầy đủ thì hơn 1 ngày khi chạy trên CPU (mất vài phút khi chạy trên GPU). Thông thường, người dùng không trang bị máy tính có cấu hình đáp ứng yêu cầu về tài nguyên. Và nếu trang bị để đáp ứng yêu cầu nhưng không sử dụng thường xuyên thì không tận dụng hiệu quả nguồn tài nguyên.

Ngoài việc dựng ảnh 3D, các công cụ trên cũng cung cấp chức năng xuất ra ảnh 3D để có thể khảo sát ảnh y khoa với nhiều góc nhìn khác nhau. Tuy nhiên, tích hợp quá trình phát hiện vùng bất thường và hiển thị 3D vùng bất thường thì chưa có nhiều nghiên cứu. Với những đặc điểm đó của Slicer và nhu cầu thực tế của người dùng, trong nghiên cứu này, tiện ích mở rộng Billow AISA của Slicer được thực hiện nhằm xây dựng một công cụ phân tích, dự đoán từ dữ liệu ảnh do người dùng cung cấp. Tiện ích mở rộng giúp người dùng thực thi những phân tích, mô hình phức tạp trên các máy chủ mạnh của các dịch vụ và chỉ hiển thị hình ảnh 3D trên máy cá nhân. Chức năng phân tích, dự đoán được thử nghiệm trong nghiên cứu này là phát hiện vùng bất thường trên ảnh MRI não với mô hình Swin-Unet.

2. PHƯƠNG PHÁP NGHIÊN CỨU

Ứng dụng Slicer cung cấp nhiều tiện ích mở rộng hỗ trợ phân vùng ảnh y tế tự động. TotalSegmentator là module dùng để phân vùng một số bộ phận ở vùng bụng bằng mô hình U-net. Hình 4 mô tả các bước xử lý dữ liệu từ một bộ ảnh MRI chuyển đổi thành mô hình 3D thông qua tiện ích mở rộng TotalSegmentator trên Slicer. Đầu tiên, người dùng cung cấp ảnh MRI vào ứng dụng Slicer. Tiếp theo, Slicer sẽ gửi tập ảnh MRI tới tiện ích mở rộng TotalSegmentator. Mô hình nnU-Net đã được huấn luyện trước tiến hành khoanh vùng các bộ phận trong ảnh MRI và trả kết quả về cho ứng dụng Slicer.

Slicer sử dụng kết quả này để hiển thị mô hình 3D với các phân vùng đã được phát hiện bởi nnU-net.



Hình 4. Quá trình xử lý của module TotalSegmentator trong Slicer

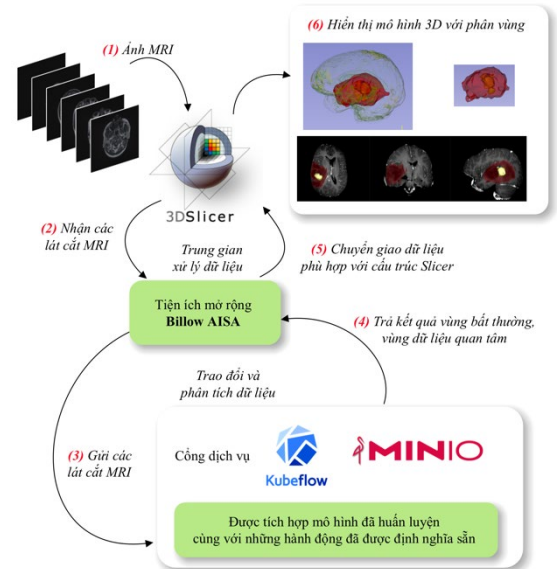
Thư viện TotalSegmentator cho phép thực hiện nhiều chế độ phân tích khác nhau tùy theo hệ thống và cài đặt thích hợp. Tuy nhiên, việc cài đặt tiện ích mở rộng TotalSegmentator đòi hỏi phải có kiến thức về hệ thống nên một số người dùng chưa thể tiếp cận được dễ dàng và thư viện này hiện chưa có chức năng phát hiện vùng bất thường.

Để giải quyết vấn đề tài nguyên khi phân tích xử lý dữ liệu ở máy tính cá nhân người dùng, hệ thống Billow AISA cho phép nhận và xử lý dữ liệu từ người dùng được đề xuất. Hệ thống sẽ xử lý những công việc cần nhiều tài nguyên mà không cần phải thực thi ở máy tính cá nhân như phân tích ảnh, dự đoán vùng bất thường, ... Hệ thống hỗ trợ chia sẻ tài nguyên cho nhiều người dùng khác cùng sử dụng cũng như làm giảm sự phức tạp của tiện ích mở rộng TotalSegmentator.

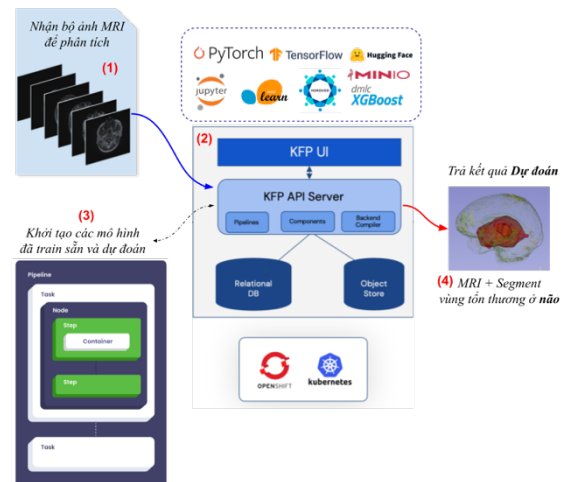
2.1. Billow AI-Segment Annotation (Billow AISA)

Billow AISA gồm có hai phần: (1) Module tích hợp trong ứng dụng Slicer; (2) Công dịch vụ phân tích, dự đoán bộ ảnh do người dùng gửi tới. Trong Hình 5, module Billow AISA được tích hợp trong ứng dụng Slicer để làm trung gian trao đổi dữ liệu giữa Slicer và công dịch vụ đã được tích hợp mô hình huấn luyện có sẵn. Đầu tiên, người dùng đưa bộ ảnh MRI vào Slicer; tiếp theo, từ ứng dụng 3D Slicer người dùng mở bảng điều khiển của module

Billow AISA để tiến hành chọn những lát cắt đã có trong Slicer để gửi đến công dịch vụ; Công dịch vụ phân tích (xử lý) các ảnh được gửi từ Slicer và gửi kết quả trở về cho module Billow AISA. Khi module Billow AISA nhận được kết quả từ công dịch vụ (5) thì module sẽ tiến hành nhận và hiển thị kết quả dưới định dạng ba chiều theo định dạng yêu cầu của ứng dụng 3D Slicer.



Hình 5. Quy trình tích hợp module Billow AISA trong Slicer



Hình 6. Sơ đồ công dịch vụ trên mô hình KubeFlow & MinIO để dự đoán kết quả

Mô hình công dịch vụ của Billow AISA được xây dựng dựa trên cấu trúc KubeFlow kết hợp với MinIO như mô tả trong Hình 6. Từ dữ liệu đầu vào do người dùng cung cấp thông qua module Billow AISA (1), công dịch vụ (2) sẽ tiến hành khởi tạo

hoặc tái sử dụng lại những “pipeline” có những “công việc” đã được cài đặt sẵn (3). Sau khi dữ liệu đã được phân tích và xử lý ở các pipeline (3), kết quả sẽ được công dịch vụ trả về cho người dùng (4) thông qua module Billow AISAs. Các “công việc” của các “pipeline” trong nghiên cứu này là mô hình Swin-Unet dùng để phát hiện các vùng bất thường trong ảnh MRI não. Nếu quá trình thực thi cần nhiều “pipeline” hay nhiều công việc cần xen kẽ thì công dịch vụ (2) kết nối và bổ sung thêm tài nguyên để đáp ứng nhiều công việc của người dùng.

Tại vị trí (2) trong sơ đồ Hình 6, dữ liệu đầu vào/đầu ra được trao đổi tuân theo định dạng NifTI. Các ứng dụng Minio, DirectPV và JuiceFS được sử dụng kết hợp để có thể tùy chọn các dịch vụ lưu trữ bằng CSI (Container Storage Interface). Quá trình này hoạt động tương tự như các server AWS, Google được tích hợp giao thức S3 (Simple Storage Service) và CSI tương ứng nhằm cung cấp khả năng trao đổi và giao tiếp dữ liệu qua S3. Giao thức S3 thông qua ứng dụng Minio được sử dụng để trao đổi dữ liệu giữa các luồng trên máy chủ phân tích và xử lý yêu cầu được gửi lên. Kubernetes định nghĩa POD như là một đại diện cho một nhóm gồm một hoặc nhiều ứng dụng cũng như một số tài nguyên được chia sẻ trong quá trình thực thi, POD trong nghiên cứu này đảm nhiệm một quá trình xử lý trong một “Pipeline”. Các POD giúp xác định các tập ảnh CT/MRI sử dụng trong quá trình trao đổi dữ liệu này.

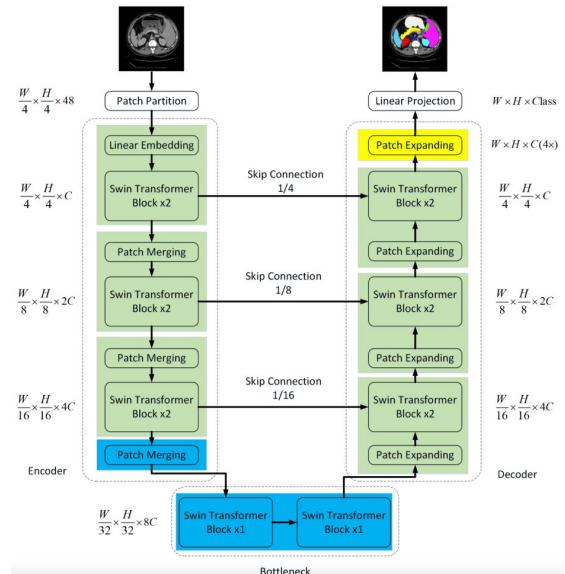
Tại vị trí xử lý (3) của sơ đồ Hình 6, tập ảnh được chỉ định được đưa vào một Pipeline để kiểm tra và khoanh vùng bất thường. Pipeline nhận xử lý và truyền tham số đầu vào cho POD xử lý dữ liệu đầu vào trước khi thực hiện mô hình Swin-Unet được xây dựng như đặt lại kích thước ảnh cho phù hợp với kích thước đầu vào 256x256 của mô hình; chuyển đổi các vùng bất thường từ ảnh có kích thước 256x256 thành 512x512 phù hợp với kích thước ảnh đầu vào; chuẩn hoá ảnh,... Sau khi dữ liệu thực thi xong các yêu cầu như chuẩn hoá và phát hiện vùng bất thường được trả về cho máy tính cá nhân của người dùng để hiển thị 3D các vùng tổn thương.

Nhằm tăng tốc độ thực thi và dễ bảo trì, sao lưu dữ liệu với ứng dụng Minio, ứng dụng DirectPV cùng với ổ đĩa giả lập bằng phương thức iSCSI trên hệ điều hành Windows được cài đặt. Bên cạnh đó, ứng dụng JuiceFS được sử dụng để kết nối đường dẫn S3 vào trong các POD để lưu trữ kết quả sau khi xử lý hoàn tất nhằm tăng hiệu suất trao đổi dữ liệu. Ứng dụng JuiceFS sẽ đảm nhiệm việc kết nối thư mục (PVC - Persistent Volume Claim) chứa các kết

quả đầu ra để việc trao đổi dữ liệu (ghi, xuất dữ liệu) được thực hiện nhanh chóng. JuiceFS cũng có tác dụng kết nối thư mục cho nhiều mô hình huấn luyện khác mà không cần phải xây dựng lại mô hình đã huấn luyện cho phù hợp với mô hình trao đổi dữ liệu theo phương thức S3.

2.2. Mô hình Swin-Unet

Mô hình Swin-Unet được cài đặt trong nghiên cứu này là sự kết hợp giữa cấu trúc của Swin Transformer và kiến trúc mạng Unet như Hình 7. Mạng nhận đầu vào là ảnh có kích thước 256x256x3 và sử dụng kiến trúc Swin-T cho bộ mã hóa để trích xuất thông tin đặc trưng từ ảnh đầu vào.



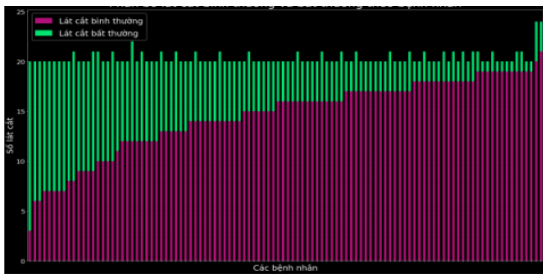
Hình 7. Kiến trúc mô hình Swin-Unet

Trong quá trình thiết kế mạng Swin-Unet, nghiên cứu đã áp dụng trọng số từ mô hình Swin Transformer V2 phiên bản 'tiny' (swin2_tiny_256) làm điểm khởi đầu. Mô hình này đã được huấn luyện trước trên tập dữ liệu ImageNet, chứa đựng kiến thức về một loạt đặc trưng và mối quan hệ giữa các đối tượng trong lĩnh vực thị giác máy tính. Sự lựa chọn này nhằm mục đích tận dụng kiến thức đã học được từ một trong những tập dữ liệu lớn và đa dạng hiện nay, nhằm cải thiện khả năng của Swin-Unet trong việc trích xuất thông tin từ hình ảnh, đòi hỏi sự hiểu biết sâu sắc về cấu trúc và hình thái. Trọng số của 'swin2_tiny_256' được áp dụng là điểm khởi đầu cho bộ mã hóa của SwinUnet. Điều này bao gồm việc khởi tạo các khối Swin Transformer với các trọng số đã học được từ ImageNet, giúp mô hình nhanh chóng thích ứng với dữ liệu mới và tiết kiệm đáng kể thời gian cần thiết cho việc huấn luyện mạng từ đầu.

Hàm tối ưu hóa Adam với tỉ lệ học là $5e-5$ và epsilon là $1e-5$ được áp dụng trong nghiên cứu này. Kích thước lô (batch_size) đặt là 8. Ở đây mô hình tự dừng lại sau 150 epoch không có sự giảm về val_loss. Hàm mất mát sử dụng để huấn luyện ở đây là hàm tversky_loss

3. KẾT QUẢ THỰC NGHIỆM

Trong bài báo này, máy khách được sử dụng có cấu hình: MacBook Pro M1 2020, RAM 16G. Máy chủ để thực thi các mô hình được xử lý trên một máy tính đơn: CPU Intel(R) Xeon(R) CPU E5-2670 v3 @2.30 GHz đi kèm với 64GB ram cho phép huấn luyện với batch có kích thước lớn, GPU có hiệu suất cao như GPU Nvidia Tesla M40 giúp rút ngắn đáng kể thời gian tính toán.



Hình 8. Phân bố lát cắt bình thường và bất thường theo bệnh nhân

Tập dữ liệu thu thập tại Bệnh viện Trường Đại học Y Dược Cần Thơ (BV TĐHYDCT) gồm 139 ảnh MRI não sử dụng chuỗi xung FLAIR cắt theo phương ngang trung bình mỗi ảnh có 20 lát cắt. Do bộ dữ liệu thu từ BV TĐHYDCT cũng như đa phần các bệnh viện ở các huyện xã tại Việt Nam đều sử

dụng là ảnh MRI xám trong khi LGG là ảnh màu, vì vậy bước đầu tiên là chuyển ảnh LGG thành ảnh xám. Tỷ lệ ảnh bình thường trên tổng số ảnh của BV TĐHYDCT là khá lớn 3,7 trong khi LGG là 1,8. Vì vậy trong nghiên cứu này, một số ảnh bình thường trong tập dữ liệu BV TĐHYDCT đã được loại bỏ, kết quả phân bố ảnh bình thường và bất thường như Hình 8. Sau khi xử lý dữ liệu BV TĐHYDCT có 593 lát cắt bất thường và 1065 lát cắt bình thường.

3.1. Kết quả hiển thị vùng bất thường của bệnh nhân với tiện ích mở rộng đã đề xuất: Billow AISA

Hình 9. là kết quả hiển thị ba chiều của ảnh MRI não chứa một vùng bất thường lớn và một vùng bất thường lớn cùng hai vùng bất thường nhỏ (màu cam đỏ) khi sử dụng tiện ích mở rộng Billow AISA để chuyển và nhận dữ liệu sau khi phát hiện vùng bất thường bằng mô hình Swin-Unet đã được huấn luyện với tập dữ liệu ảnh MRI não. Tập các ảnh MRI não của một bệnh nhân được chuyển đến máy tính đủ mạnh đã tích hợp Kubeflow và các mô hình để thực hiện việc phát hiện vùng bất thường thông qua tiện ích Billow AISA. Sau khi phát hiện được vùng bất thường, các hình ảnh được chuyển về lại Slicer trên máy tính cá nhân để thể hiện dưới dạng 3D. Toạ độ không gian trong một bộ ảnh DCOM hay MRI sẽ giúp tính toán kích thước để xuất ảnh kết quả về định dạng hiển thị ba chiều của Slicer. Trong nghiên cứu này, ảnh đầu vào được thử nghiệm với kích thước 256×256 và 512×512 . Với mô hình đề xuất, các vùng bất thường thể hiện trực quan trên ảnh MRI não của bệnh nhân, giúp dễ dàng định vị được vùng bất thường của bệnh nhân.



Hình 9. Kết quả hiển thị vùng bất thường sau khi sử dụng Swin_Unet để phát hiện vùng bất thường và Billow AISA

3.2. Đánh giá thời gian xác định tạo ra mô hình 3D với Billow AISA

Tiện ích mở rộng Billow AISA có khả năng hỗ trợ cho các thiết bị không đủ mạnh để thực thi các mô hình phát hiện bất thường. Để đánh giá tính khả thi của tiện ích đề xuất, thời gian thực thi được thực nghiệm và đánh giá qua tất các bước kết nối, thực thi việc phân vùng, chuyển ảnh về định dạng hiển thị của Slicer.

Trong quá trình thực nghiệm, mạng cục bộ được sử dụng để kiểm tra thời gian trao đổi dữ liệu giữa máy chủ và máy khách cho thấy dữ liệu chuyển đến và đi giữa máy chủ và máy khách gần như không đáng kể, quá trình trao đổi tập tin giữa máy khách

và máy chủ chỉ mất khoảng 0,2 ~ 0,5 giây cho một tập ảnh CT/MRI. Kết quả đánh giá thời gian thực thi được thể hiện trong Bảng 1. Để xác định vùng bất thường cho 1 bệnh nhân với cấu hình của máy chủ như đã trình bày ở đầu mục 3, thời gian thực thi là từ 1 đến chưa tới 2 phút tùy tập ảnh của bệnh nhân. Quá trình chuyển ảnh về định dạng hiển thị 3 chiều theo định dạng Slicer là chưa tới 30 giây cho 1 POD. Tổng thời gian xử lý cho toàn bộ quá trình nhận, trả kết quả của 1 POD mất khoảng gần 3 phút. Khi thực hiện với 10 POD thì thời gian thực thi cho xác định vùng bất thường cũng như chuyển ảnh về định dạng hiển thị của Slicer tăng lên khoảng 7-8 lần.

Bảng 1. Đánh giá thời gian thực thi

Mẫu	1 POD	10 POD
Kết nối	30 giây	3-6 phút
Xác định vùng bất thường tập ảnh MRI (256x256x155) 2,7Mb	1phút – 1phút 56 giây	14phút – 16phút 23 giây
Chuyển ảnh 3D về định dạng của cấu trúc Slicer	10~30 giây	2 phút 10 giây – 5 phút 12 giây

3.3. Đánh giá khả năng phát hiện và hiển thị 3D với Billow AISA

Bên cạnh những kết quả đánh giá về thời gian thực thi, các bác sĩ tại BV TĐHYDCT cũng đã thực hiện việc đánh giá khả năng hiển thị của Billow AISA. Kết quả khảo sát được đánh giá bởi 5 bác sĩ khác nhau (thực hiện độc lập) trên 10 tập ảnh MRI

của 10 bệnh nhân khác nhau. Thông tin tập ảnh của 10 bệnh nhân sử dụng để đánh giá kết quả của giải pháp đề xuất được thể hiện trong Bảng 2. Trung bình tập ảnh của mỗi bệnh nhân gồm có 20-21 lát cắt, trong đó có từ 8 đến 13 lát cắt chứa vùng bất thường. Tùy bệnh nhân, thể tích vùng bất thường có thể lớn nhỏ khác nhau như mô tả trong Bảng 2.

Bảng 2. Thông tin tập ảnh của bệnh nhân được sử dụng để đánh giá

Bệnh nhân	Thể Tích Vùng Bất Thường (cm3)	Số lát cắt của một bệnh nhân	Số lát cắt chứa vùng bất thường
1	158,843	21	12
2	89,002	21	13
3	25,078	20	10
4	12,503	21	11
5	115,175	20	11
6	18,321	21	8
7	111,300	20	12
8	52,080	20	13
9	178,622	20	13
10	107,486	20	10

Các bác sĩ phản hồi độ chính xác của các ảnh hiển thị vùng bất thường dưới định dạng ba chiều sau khi sử dụng Billow AISA với các tiêu chí: vị trí, hình dáng, màu sắc và độ rõ nét của vùng bất thường được hiển thị. Mỗi tiêu chí sẽ có 5 mức độ đánh giá từ xấu nhất (mức 1) đến tốt nhất (mức 5). Bên cạnh đó, thể tích của vùng bất thường cũng được ước

lượng và đánh giá bởi các bác sĩ. Kết quả khảo sát trung bình trên 10 bệnh nhân của mỗi bác sĩ được hiển thị như Bảng 3. Các vùng bất thường có thể tích lớn đa phần được đánh giá ở mức 4, một số vùng bất thường có thể tích nhỏ được đánh giá ở mức 3. Trung bình đánh giá của các bác sĩ ở các tiêu chí đều trên 3,5. Về phần ước lượng thể tích, độ chính xác được đánh giá đạt gần 75%

Bảng 3. Đánh giá độ chính xác hiển thị của Billow AISA

Tiêu chí	BS 1	BS 2	BS 3	BS 4	BS 5	TB
Vị trí	3,6	3,9	3,9	3,6	3,4	3,68
Hình dáng	3,4	3,6	3,8	3,5	3,3	3,52
Màu sắc	3,8	3,7	3,6	3,6	3,3	3,6
Độ rõ nét	3,5	3,3	3,9	3,6	3,5	3,56

4. KẾT LUẬN

Hiện thị hình ảnh dưới dạng ba chiều hỗ trợ rất nhiều cho các bác sĩ trong quá trình chẩn đoán và điều trị cho bệnh nhân trong tình trạng các bệnh viện luôn quá tải. Trong nghiên cứu này, tiện ích mở rộng Billow AISA được xây dựng và thực nghiệm trong việc phát hiện và hiển thị 3D vùng bất thường trên ảnh MRI não. Mô hình Swin-Unet được huấn luyện trên tập dữ liệu LGG kết hợp với tập dữ liệu đã được thu thập và gán nhãn của 106 bệnh nhân tại BV TĐHYDCT để phát hiện vùng bất thường. Tiện ích Billow AISA cho phép trao đổi dữ liệu giữa máy chủ và máy khách để thực thi việc phát hiện vùng bất

thường và kết xuất dữ liệu đúng định dạng hiển thị 3D của ứng dụng Slicer. Thời gian thực thi trên 1POD và 10POD cũng được thực thi và đánh giá cho thấy tính khả thi và hiệu quả của Billow AISA.

Trong tương lai, các mô hình phát hiện vùng bất thường khác cũng có thể đưa vào một cách dễ dàng, việc xác định tọa độ tương ứng với các kích thước đầu vào cũng sẽ được bổ sung để quá trình hiển thị 3D của ảnh trên Slicer được đa dạng thay vì chỉ có 2 kích thước 256×256 và 512×512 như hiện nay. Thêm vào đó, độ chính xác trong việc hiển thị ba chiều các vùng bất thường cũng cần cải thiện thêm để tăng độ tin cậy trong triển khai thực tế.

TÀI LIỆU THAM KHẢO

Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-End Object Detection with Transformers. *In: Vedaldi, A., Bischof, H., Brox, T., Frahm, JM. (eds) Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science (12346)*. Springer, Cham. https://doi.org/10.1007/978-3-030-58452-8_13

Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Wang, M. (2023). Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation. *In: Karlinsky, L., Michaeli, T., Nishino, K. (eds) Computer Vision – ECCV 2022 Workshops. ECCV 2022. Lecture Notes in Computer Science (13803)*. Springer, Cham. https://doi.org/10.1007/978-3-031-25066-8_9

Chris, C., Danfei, X., Young, Y. G., Kevin, C., & Silvio, S. (2016). 3D-R2N2: A Unified Approach for Single and Multi-view 3D Object Reconstruction. *European Conference on Computer Vision* (pp. 628-644). 10.1007/978-3-319-46484-8_38.

Çiçek, Ö., Abdulkadir, A., Lienkamp, S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. *In: Ourselin, S., Joskowicz, L., Sabuncu, M., Unal, G., Wells, W. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. MICCAI 2016. Lecture Notes in Computer Science (pp. 424–432)*. Springer, Cham. https://doi.org/10.1007/978-3-319-46723-8_49

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. *in International Conference on Learning Representations*.

Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H. R., & Xu, D. (2022). Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images. *In: Crimi, A., Bakas, S. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2021. Lecture Notes in Computer Science (12962)*. Springer, Cham. https://doi.org/10.1007/978-3-031-08999-2_22

Held, K., Kops, E., Krause, B., Wells, W., Kikinis, R., & Muller-Gartner, H. W. (1997). Markov random field segmentation of brain MRI images. *IEEE Transactions on Medical Imaging* (pp. 878–886). <https://doi.org/10.1109/42.650883>

Ken, M., Will, S., Bill, L. (2015). *The Visualization Toolkit*. <https://github.com/Kitware/VTK?tab=readme-ov-file>

Li, X., Chen, H., Qi, X., Dou, Q., Fu, C. W., & Heng, P. A. (2018). H-denseunet: Hybriddensely connected unet for liver and tumor segmentation from ct volumes. *IEEE Transactions on Medical Imaging* (pp. 2663–2674). <https://doi.org/10.1109/TMI.2018.2845918>

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swintransformer:

- Hierarchical vision transformer using shifted windows. *IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada* (pp. 9992-10002). doi: 10.1109/ICCV48922.2021.00986
- Lợi, L. M., Thur, T. N. M., Nguyễn, H. T., & Khang, P. N. (2020). Ứng dụng mô hình U-net phát hiện vùng bất thường trên ảnh MRI Não”. *Kỷ yếu Hội nghị KHCN Quốc gia lần thứ XIII về Nghiên cứu cơ bản và ứng dụng Công nghệ thông tin (FAIR)*, Nha Trang (trang 656-662). <http://dx.doi.org/10.15625/vap.2020.00225>
- Lợi, L. M., Thur, T. N. M., An, H. Q., & Khang, P. N. (2024). Phát hiện vùng bất thường trên ảnh MRI não với mô hình Swin-Unet. *Tạp chí Khoa học Đại học Thái Nguyên*, 229(7). (trang 111-120). <https://doi.org/10.34238/tnu-jst.10053>
- Nhường, Q. T. B., Ninh, T. V., Thịnh, Đ. P., & Thường, P. M. (2021). Tái tạo mô hình 3D của đối tượng từ ảnh phác thảo 2.5D. *Hội nghị KHCN Quốc gia lần thứ XV về Nghiên cứu cơ bản và ứng dụng Công nghệ thông tin (FAIR)*. (trang 69-74).
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *In Medical Image Computing and Computer Assisted Intervention (MICCAI), ser. LNCS* (pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28
- Thur, T. N. M., Liêm, B. T., Lợi, L. M., Nghĩa, P. T., & Khang, P. N. (2022). Ứng dụng mạng Gan có điều kiện phát hiện vùng bất thường trên ảnh MRI não. *Hội nghị KHCN Quốc gia lần thứ XV về Nghiên cứu cơ bản và ứng dụng Công nghệ thông tin (FAIR)*, Học viện Kỹ thuật Mật mã, Hà Nội, 2022. (trang 462-469). <https://doi.org/10.15625/vap.2022.0256>
- Tsai, A. Y., Wells, W., Tempany, C., Tucker, A. F., Grimson, & Willsky, W. A. (2003). A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 22(2). (pp. 137–154). <https://doi.org/10.1109/TMI.2002.808355>
- Weidert, S., Address, S., Linhart, C., Suero, E. M., Greiner, A., Böcker, W., Kammerlander, C., & Becker, C.A. (2020). 3D printing method for next-day acetabular fracture surgery using a surface filtering pipeline: feasibility and 1-year clinical results. *International Journal of Computer Assisted Radiology and Surgery*, 15, 565–575. <https://doi.org/10.1007/s11548-019-02110-0>
- Xiao, X., Lian, S., Luo, Z., & Li, S. (2018). Weighted res-unet for high-quality retina vesselsegmentation. *9th International Conference on Information Technology in Medicine and Education (ITME)* (pp. 327–331). <https://doi.org/10.1109/ITME.2018.00080>
- Zhou Z., Siddiquee M. R., Tajbakhsh N., & Liang, J. (2018). *Unet++: A nested u-net architecture for medical image segmentation* (pp. 3–11). Springer Verlag. https://doi.org/10.1007/978-3-030-00889-5_1