

MÔ HÌNH AI PHÁT HIỆN VÙNG TỤ MÁU DO XUẤT HUYẾT NÃO TRÊN ẢNH CT NỘI SỌ

Nguyễn Thị Bích Hà¹, Phạm Trần Thiện², Lê Thị Kim Nga^{2*}

¹Trường Cao đẳng Kỹ thuật Công nghệ Quy Nhơn

²Khoa Công nghệ thông tin, Trường Đại học Quy Nhơn

nguyenthibichha@cdkctcnqn.edu.vn, phamtranthien@qnu.edu.vn, kimle@qnu.edu.vn

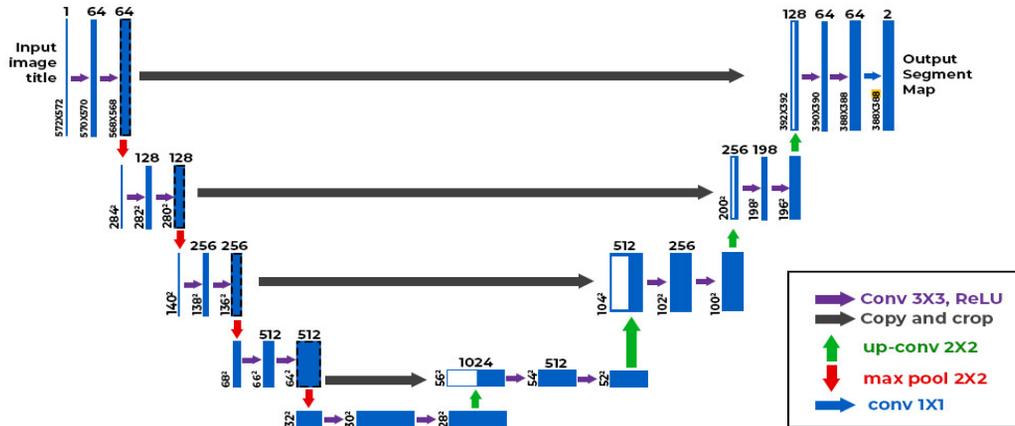
TÓM TẮT— Nghiên cứu này đề xuất một mô hình trí tuệ nhân tạo (AI) dựa trên kiến trúc Viola-UNet nhằm phát hiện vùng tụ máu trên ảnh CT sọ não của bệnh nhân tại Việt Nam. Viola-UNet kết hợp giữa kiến trúc U-Net và các kỹ thuật học sâu hiện đại, bao gồm module chú ý Viola, nhằm nâng cao khả năng trích xuất đặc trưng và cải thiện độ chính xác phân đoạn. Kết quả kiểm định chéo 5 lần cho thấy Viola-UNet cho kết quả cải thiện đáng kể so với nnUNet, đặc biệt ở chỉ số Dice Similarity Coefficient (DSC). Ngoài ra, mô hình cũng cho kết quả tốt hơn ở các chỉ số đánh giá khác như Hausdorff Distance (HD), Normalised Surface Dice (NSD) và Relative Volume Difference (RVD). Những kết quả này cho thấy tiềm năng ứng dụng của Viola-UNet trong thực hành lâm sàng, góp phần hỗ trợ phát hiện và đánh giá tổn thương do xuất huyết não một cách hiệu quả và kịp thời.

Từ khóa— AI y tế, Viola-UNet, CT sọ, xuất huyết não, phân đoạn tụ máu.

I. ĐẶT VẤN ĐỀ

Trí tuệ nhân tạo (AI), đặc biệt là học sâu (deep learning) đã phát triển mạnh mẽ và len lỏi vào nhiều lĩnh vực trong đời sống. Cụ thể trong lĩnh vực y khoa, các mô hình xử lý ảnh đã đạt được những tiến bộ đáng kể trong việc hỗ trợ chẩn đoán. Phương pháp chẩn đoán hình ảnh y khoa truyền thống của bác sĩ là một công việc khó khăn, đòi hỏi sự cẩn thận và chính xác. Trong những trường hợp phức tạp như xuất huyết não hay khối u não, bệnh nhân cần được chẩn đoán nhanh chóng. Ngoài ra, sự quá tải ở những bệnh viện lớn cũng là một yếu tố có thể làm cho tình trạng bệnh trở nặng nếu không kịp thời can thiệp. Sự ra đời của các mô hình học sâu, đặc biệt là mạng nơ-ron tích chập (convolutional neural network – CNN) đã cung cấp những giải pháp nhanh chóng, hiệu quả và chính xác cho bài toán phân đoạn ảnh y khoa [1], [2].

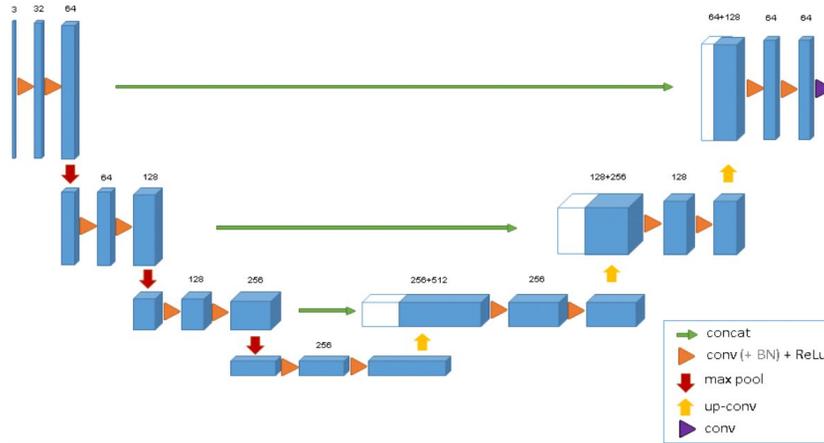
Một trong những mô hình tiêu biểu và được ứng dụng rộng rãi nhất trong phân đoạn hình ảnh y khoa là U-Net được phát triển bởi Ronneberger và cộng sự [3]. U-Net là một mô hình nơ-ron sâu có cấu trúc đối xứng đặc biệt, với khả năng kết hợp thông tin từ cả lớp mã hóa (encoder) và lớp giải mã (decoder) (Hình 1). Điều này giúp U-Net đặc biệt hiệu quả trong việc xử lý các bài toán phân đoạn ảnh với lượng dữ liệu giới hạn.



Hình 1. Kiến trúc U-Net

Một cải tiến đáng chú ý là sự ra đời của 3D U-Net (Hình 2), được phát triển bởi Çiçek và cộng sự [4], mở rộng U-Net lên không gian ba chiều. Nhờ vào cấu trúc đặc biệt này, 3D U-Net đã trở thành mô hình chủ đạo trong các bài toán phân đoạn y khoa 3D, giúp cải thiện đáng kể hiệu suất trong các bài toán phức tạp như phân đoạn khối u não. Ngoài ra, nnU-Net, một phiên bản tự động của U-Net, đã được Isensee và cộng sự phát triển [5] nhằm tối ưu hóa hiệu suất của mô hình cho từng loại dữ liệu cụ thể mà không cần can thiệp thủ công.

* Corresponding Author



Hình 2. Kiến trúc 3D-UNet

Bên cạnh nnU-Net, V-Net, một mô hình khác dựa trên CNN, cũng đã được Milletari và cộng sự [6] phát triển với khả năng phân đoạn các dữ liệu hình ảnh y khoa thể tích (volumetric medical imaging), đem lại độ chính xác cao trong việc phát hiện tổn thương. Trong các bài toán phân đoạn hình ảnh y khoa thì bài toán phát hiện vùng tụ máu do xuất huyết não là một thử thách nhận được nhiều sự quan tâm. Xuất huyết não là một trong những bệnh lý có nguy cơ gây tử vong hàng đầu. Việc phát hiện sớm giúp bác sĩ đánh giá và can thiệp kịp thời, cải thiện được tình trạng bệnh nhân. Phát hiện khối máu tụ còn là thách thức đối với các mô hình phân đoạn, do hình dạng phức tạp và không đồng nhất của khối máu giữa các ca lâm sàng. Đã có nhiều công trình liên quan đề xuất giải pháp cho bài toán này.

Trong nghiên cứu của Lee công bố năm 2020 [7], CNN được sử dụng để phân loại và phát hiện xuất huyết não trên hình ảnh CT đạt độ chính xác hơn 91%. Nghiên cứu của Hoang và cộng sự (2024) [8] cũng chứng minh việc sử dụng CNN có thể cải thiện độ chính xác của việc phát hiện các vùng tổn thương xuất huyết trên ảnh CT. Tuy nhiên, hạn chế của CNN truyền thống là khó phân đoạn các vùng tổn thương nhỏ và dễ bị nhiễu do các yếu tố không mong muốn trong dữ liệu y khoa.

Để vượt qua những hạn chế của các phương pháp truyền thống, các mô hình học sâu hiện đại như mạng nơ-ron hồi quy (Recurrent Neural Networks – RNN) đã được đưa vào ứng dụng trong phân tích ảnh y khoa. Kadam và cộng sự (2021) [9] đã đề xuất một mô hình kết hợp giữa mạng tích chập (CNN) và RNN, không chỉ có khả năng phát hiện mà còn dự đoán sự tiến triển của các vùng xuất huyết. Tương tự, nghiên cứu của Rajpurkar và đồng nghiệp [10] đã tích hợp CNN với mạng bộ nhớ dài ngắn hạn (Long Short-Term Memory – LSTM) nhằm phân tích hình ảnh y tế và dự báo quá trình phát triển của xuất huyết não, cho thấy tiềm năng của các mô hình lai trong việc nâng cao độ chính xác và khả năng dự đoán trong các bài toán y sinh phức tạp. Một trong những bước tiến đáng kể để giải quyết bài toán này là mô hình U-Net có khả năng phân đoạn chính xác các vùng xuất huyết nhỏ trên hình ảnh CT. Nghiên cứu của Hssayeni và cộng sự [11] đã sử dụng U-Net để phân đoạn các vùng xuất huyết trong não và đạt được kết quả ấn tượng về độ chính xác và độ nhạy. Ngoài ra, nghiên cứu của Ker và cộng sự [12] đã cải tiến U-Net với lớp chú ý (attention layer) để cải thiện hiệu suất phân đoạn.

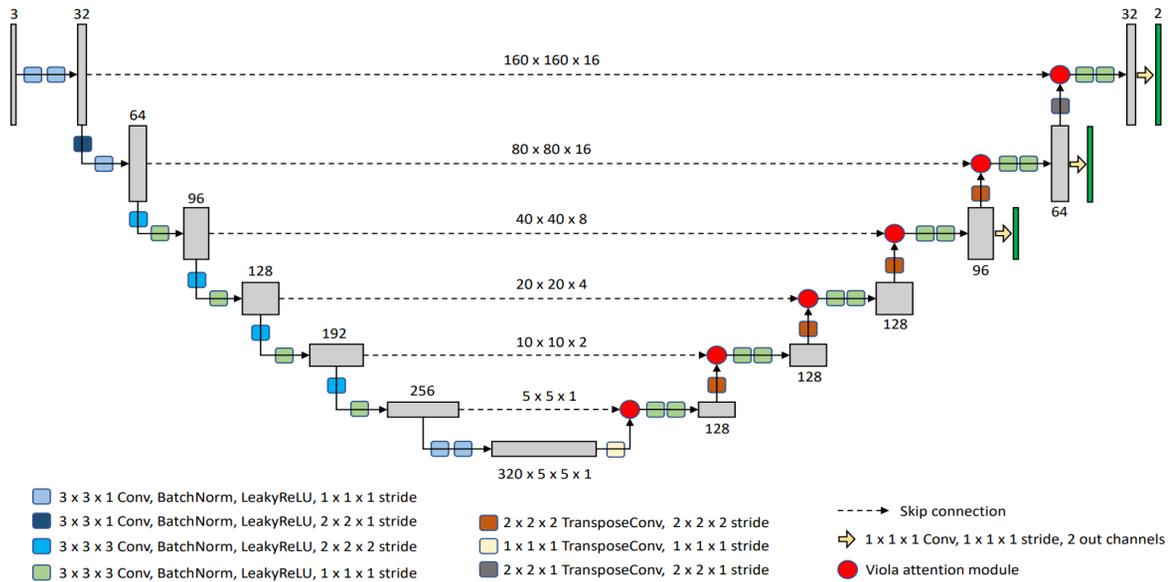
Gần đây, nghiên cứu của Walluscheck và cộng sự [13] sử dụng mô hình lai giữa CNN và GAN (Generative Adversarial Networks) để phân đoạn xuất huyết não đã tạo ra những cải tiến lớn trong việc phát hiện các tổn thương nhỏ và tinh vi. Mô hình được ghi nhận có cải tiến đáng kể nhất trong bài toán phân đoạn vùng xuất huyết não là mô hình Viola-UNet. Mô hình Viola-UNet, được giới thiệu bởi Liu và cộng sự [14], là một biến thể cải tiến của kiến trúc U-Net truyền thống, được thiết kế đặc biệt cho nhiệm vụ phân đoạn xuất huyết nội sọ từ ảnh CT không tiêm thuốc cản quang. Viola-UNet sử dụng cơ chế chú ý (attention mechanism) để tối ưu hóa quá trình phân đoạn, giúp hệ thống đạt được độ chính xác cao trong việc xác định các vùng tổn thương do xuất huyết não. Mô hình này đã vượt qua các mô hình chuẩn khác như nnU-Net và chiến thắng trong thử thách INSTANCE 2022 trong nhiệm vụ phân đoạn xuất huyết nội sọ.

2. MÔ HÌNH VIOLA-UNET

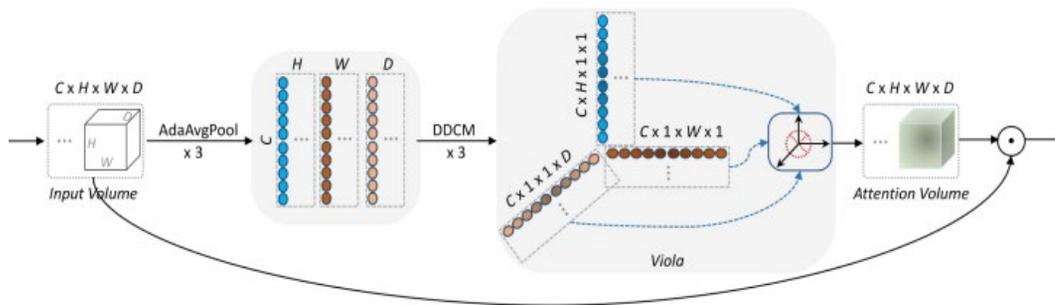
Viola-UNet dựa trên cấu trúc encoder-decoder của U-Net, kết hợp với module chú ý Viola (Viola attention module). Module này áp dụng các phép chiếu trục giao để khai thác các đặc trưng trong không gian 3D (Hình 3). Viola tận dụng các voxel trong không gian đặc trưng giao nhau dọc theo các mức trục giao (Voxels Intersecting Along Orthogonal Levels – Viola) để tạo ra bản đồ chú ý (attention map) (Hình 4).

Module chú ý Viola được xây dựng trên cơ sở mô hình hóa mối quan hệ giữa các chiều khác nhau của không gian ảnh. Cụ thể, module Viola sử dụng cơ chế nén trung bình thích ứng (adaptive average pooling) để nén khối lượng đặc trưng đầu vào $X \in R^{C \times H \times W \times D}$, (trong đó C , H , W , và D lần lượt là số lượng kênh, chiều cao, chiều rộng và độ

sâu của không gian đặc trưng) thành ba không gian biểu diễn dọc theo các trục: $X_h \in R^{C \times H}$, $X_w \in R^{C \times W}$, và $X_d \in R^{C \times D}$.



Hình 3. Kiến trúc Viola U-Net



Hình 4. Module Voxels Intersecting along Orthogonal Levels Attention (Viola)

Tiếp theo, module sử dụng mạng gộp tích chập giãn dày đặc (dense dilated convolutions merging networks) để kết hợp thông tin giữa các kênh và thông tin ngữ cảnh phi cục bộ trên mỗi phương trục giao với các hạt nhân (kernel) thích ứng.

Các không gian biểu diễn ẩn được chuyển kích thước (reshape) như $X_h \in R^{C \times H \times 1 \times 1}$, $X_w \in R^{C \times 1 \times W \times 1}$, $X_d \in R^{C \times 1 \times 1 \times D}$ được dùng để cấu tạo bản đồ chú ý $A_{Viola} \in R^{C \times H \times W \times D}$ bằng các phương trình toán học sau.

$$\begin{aligned}
 \tilde{X}_h, \tilde{X}_w, \tilde{X}_d &= \sigma_s(X_h, X_w, X_d) \\
 \hat{X}_h, \hat{X}_w, \hat{X}_d &= \sigma_{gt}(X_h, X_w, X_d) \\
 A &= \sigma_{gt}(\tilde{X}_h + \hat{X}_h + \tilde{X}_w + \hat{X}_w + \tilde{X}_d + \hat{X}_d) + \tilde{X}_h \otimes \tilde{X}_h + \tilde{X}_w \otimes \tilde{X}_w + \tilde{X}_d \otimes \tilde{X}_d + \tilde{X}_h \otimes \tilde{X}_h \otimes \tilde{X}_h \\
 A_{Viola} &= (\alpha + \|flatten(A)\|^2 - 1)A + \beta \\
 X &= X \odot A_{Viola}
 \end{aligned}$$

Trong đó, σ_s là hàm kích hoạt Sigmoid, σ_{gt} là sự kết hợp giữa chuẩn hóa nhóm (group normalization) và hàm Tanh, σ_r là hàm kích hoạt ReLU, và \otimes là phép nhân tensor. Hệ số chú ý $\alpha=0,1$ và $\beta=0,3$ được sử dụng để cân bằng trọng số của các bản đồ chú ý và các bản đồ đặc trưng dư (residual feature maps).

Viola-UNet có những ưu điểm vượt trội so với các mô hình phân đoạn khác như nnU-Net. Việc sử dụng module chú ý không chỉ giúp tăng cường độ chính xác trong phân đoạn mà còn tối ưu hóa việc xử lý thông tin không gian 3D, đặc biệt là trong các vùng xuất huyết nội sọ phức tạp. Kết quả của cuộc thi INSTANCE 2022 cho thấy mô hình vượt trội ở các chỉ số NSD và RVD so với những mô hình cùng tham gia. Ngoài ra, kết hợp giữa mô hình nnUNet và mô hình Viola-UNet cũng vượt trội hơn ở chỉ số DSC và HD trong cuộc thi.

Mô hình Viola-UNet cho thấy khả năng vượt trội khi xử lý bài toán phân đoạn vùng xuất huyết não trên tập dữ liệu

của cuộc thi. Mô hình này có tiềm năng để có thể ứng dụng trong lâm sàng và cần được huấn luyện và đánh giá trên bộ dữ liệu thực tế. Nghiên cứu này huấn luyện và đánh giá mô hình Viola – UNet trên tập dữ liệu ảnh CT sọ não tại bệnh viện Việt Nam.

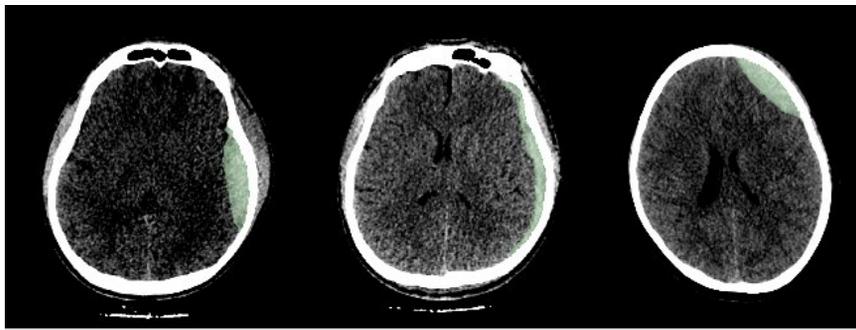
3. THỰC NGHIỆM

A. DỮ LIỆU

Quy trình thu thập và xử lý dữ liệu CT Scan cho nghiên cứu này được thực hiện với sự hợp tác chặt chẽ giữa nhóm nghiên cứu và các bác sĩ tại Bệnh viện Đa khoa tỉnh Bình Định. Dữ liệu CT Scan được thu thập từ các bệnh nhân được chẩn đoán có nguy cơ hoặc đã bị xuất huyết não. Việc thu thập dữ liệu tuân thủ các quy định về bảo mật thông tin và đạo đức y khoa, đảm bảo rằng tất cả bệnh nhân tham gia đều đã được thông báo và đồng ý sử dụng dữ liệu y tế cho mục đích nghiên cứu.

1. THU THẬP DỮ LIỆU CT SCAN

Dữ liệu CT được thu thập bằng cách sử dụng các máy chụp CT tiên tiến tại Bệnh viện Đa khoa tỉnh Bình Định, với các thông số kỹ thuật chuẩn trong chẩn đoán hình ảnh não bộ. Dữ liệu được thu thập dựa trên kỹ thuật chụp quét xoắn ốc, lát cắt ngang dày 1,25 mm từ bờ dưới xương chẩm đến bờ trên vòm sọ bằng máy CT Scan Optima 660, 128 lát cắt của GE. Các bệnh nhân có các loại xuất huyết nội sọ khác nhau (chẳng hạn như tụ máu ngoài màng cứng, dưới màng cứng, xuất huyết trong mô não) được đưa vào bộ dữ liệu để đảm bảo tính đa dạng (Hình 5).



Hình 5. Dữ liệu ảnh CT sọ não thu thập tại Bệnh viện Bình Định (khối máu được đánh dấu bằng vùng màu xanh)

2. GÁN NHÃN DỮ LIỆU (LABELING)

Sau khi thu thập dữ liệu, bước tiếp theo là quá trình gán nhãn, trong đó các bác sĩ chuyên khoa tại Bệnh viện Đa khoa tỉnh Bình Định tiến hành phân tích và xác định vị trí, kích thước của vùng xuất huyết trên từng ảnh CT. Việc gán nhãn này được thực hiện thủ công, dựa trên kinh nghiệm lâm sàng của các bác sĩ, đảm bảo rằng dữ liệu được gán nhãn với độ chính xác cao nhất có thể. Quá trình gán nhãn sẽ bao gồm việc xác định biên giới của vùng xuất huyết và các thông tin liên quan như thể tích của khối máu tụ.

3. TIỀN XỬ LÝ DỮ LIỆU (PREPROCESSING)

Trước khi đưa vào mô hình Viola-UNet, dữ liệu CT sẽ trải qua các bước tiền xử lý để đảm bảo tính nhất quán và chất lượng của dữ liệu. Quá trình này bao gồm:

a) Chuẩn hóa ảnh (normalization)

Dữ liệu CT Scan từ các bệnh nhân khác nhau có thể có độ phân giải và cường độ tín hiệu khác nhau. Do đó, các ảnh sẽ được chuẩn hóa để đảm bảo rằng tất cả dữ liệu đều có cùng mức độ xám và kích thước, giúp mô hình học tập hiệu quả hơn.

b) Loại bỏ nhiễu (denoising)

Quá trình thu nhận ảnh CT có thể bị ảnh hưởng bởi nhiễu phát sinh từ môi trường hoặc thiết bị chụp, làm giảm chất lượng hình ảnh đầu vào. Việc áp dụng các kỹ thuật loại bỏ nhiễu trong giai đoạn tiền xử lý giúp cải thiện độ rõ nét của ảnh, đồng thời giảm thiểu các yếu tố gây nhiễu không mong muốn. Nhờ đó, mô hình học sâu có thể phân tích dữ liệu hình ảnh một cách chính xác và ổn định hơn.

c) Tái cấu trúc và chuẩn hóa không gian (spatial normalization)

Ảnh CT sọ não của bệnh nhân có thể được thu nhận với các hướng khác nhau tùy thuộc vào quy trình chụp và tư thế bệnh nhân. Để đảm bảo tính nhất quán trong toàn bộ tập dữ liệu, quá trình tái cấu trúc và chuẩn hóa không gian được thực hiện nhằm đưa tất cả ảnh về cùng một hệ tọa độ chuẩn. Bước tiền xử lý này đóng vai trò quan trọng trong việc chuẩn hóa đầu vào, giúp mô hình xử lý dữ liệu một cách ổn định và nâng cao độ chính xác trong các tác vụ phân tích hình ảnh.

4. PHÂN CHIA DỮ LIỆU

Dữ liệu CT sau khi được gán nhãn và tiền xử lý tiếp tục được phân chia thành các tập dữ liệu huấn luyện (training), kiểm thử (validation), và kiểm tra (testing). Tập huấn luyện sẽ chiếm phần lớn dữ liệu, được sử dụng để huấn

luyện mô hình Viola-UNet, trong khi tập kiểm thử và kiểm tra sẽ được sử dụng để đánh giá hiệu suất của mô hình sau khi huấn luyện. Quy trình thu thập và xử lý dữ liệu một cách kỹ lưỡng giúp mô hình Viola-UNet học được các đặc trưng chính xác của xuất huyết não từ hình ảnh CT thực tế, qua đó hỗ trợ hiệu quả cho bác sĩ trong công tác chẩn đoán và điều trị.

B. HUẤN LUYỆN

1. THIẾT LẬP HUẤN LUYỆN

Mô hình Viola-UNet được huấn luyện trên tập dữ liệu ảnh CT sọ não với kích thước đầu vào là $3 \times 256 \times 256 \times 128$. Quá trình huấn luyện sử dụng thuật toán tối ưu hóa Stochastic Gradient Descent (SGD) kết hợp với động lượng Nesterov, được thiết lập ở mức 0,98 nhằm tăng tốc độ hội tụ. Tốc độ học ban đầu được đặt là 5×10^{-3} và được điều chỉnh giảm dần theo lịch trình cosine trong suốt quá trình huấn luyện. Để đảm bảo sự ổn định trong giai đoạn đầu, tốc độ học được tăng dần trong 500 bước đầu tiên, giúp mô hình tránh hiện tượng dao động lớn khi bắt đầu học.

Để đảm bảo tính ổn định và đánh giá khách quan hiệu suất của mô hình, quá trình huấn luyện được triển khai theo phương pháp kiểm định chéo 5 lần (5-fold cross-validation). Mỗi vòng huấn luyện bao gồm 60.000 bước, trong đó các mô hình được lưu lại tại các thời điểm đạt giá trị trung bình Dice Similarity Coefficient (DSC) cao nhất trên tập kiểm thử.

2. KẾT QUẢ HUẤN LUYỆN

Trong quá trình huấn luyện, Viola-UNet đã đạt hiệu suất cao, vượt trội hơn so với các mô hình chuẩn khác như nnU-Net, khi được đánh giá trên tập dữ liệu CT từ Bệnh viện Đa khoa tỉnh Bình Định. Kết quả 5 lần kiểm tra chéo được trình bày trong Bảng 1.

Bảng 1. Bảng đánh giá so sánh các mô hình

Mô hình	Lần 1	Lần 2	Lần 3	Lần 4	Lần 5	DSC trung bình
nnU-Net	0,7482	0,7254	0,7689	0,7421	0,7510	0,7471
Viola-UNet (s)	0,7725	0,7460	0,7912	0,7836	0,7618	0,7710

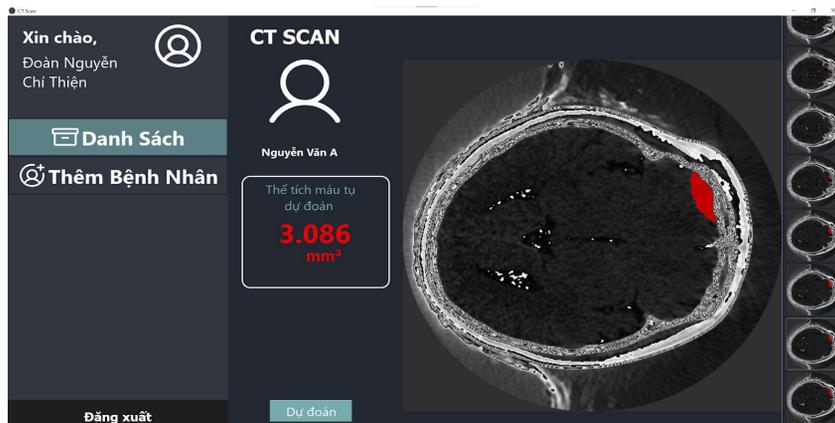
Kết quả cho thấy Viola-UNet đạt được điểm DSC trung bình là 0,7710, cao hơn so với nnU-Net với 0,7471, tương đương với mức cải thiện khoảng 3,2%. Điều này cho thấy Viola-UNet có khả năng phân đoạn chính xác các vùng xuất huyết não hơn so với nnU-Net trên cùng một bộ dữ liệu.

Ngoài ra, mô hình Viola-UNet cũng được đánh giá dựa trên các chỉ số khác như HD, NSD, và RVD. Kết quả đạt được thể hiện ở Bảng 2.

Bảng 2. Bảng kết quả đánh giá mô hình đề xuất

Mô hình	DSC	HD	NSD	RVD
nnU-Net	0,7471	25,034	0,5541	0,2067
Viola-UNet (s)	0,7710	22,540	0,5688	0,1920

Viola-UNet không chỉ đạt điểm DSC cao hơn, mà còn có HD và RVD thấp hơn so với nnU-Net, cho thấy mô hình này có khả năng phân đoạn các vùng xuất huyết chính xác hơn và sai số thấp hơn về mặt thể tích. Điều này đặc biệt quan trọng trong các ứng dụng lâm sàng vốn yêu cầu độ chính xác cao. Kết quả phân đoạn và tính thể tích máu tụ được thể hiện trên Hình 6.



Hình 6. Kết quả phân đoạn và tính thể tích máu tụ (khối máu được đánh dấu bằng vùng màu đỏ)

4. BÀN LUẬN

Viola-UNet đã được kiểm chứng không chỉ trên các chỉ số về độ chính xác mà còn về tính ổn định và khả năng tổng quát hóa trên nhiều tập dữ liệu khác nhau. Mô hình này có khả năng phân đoạn chính xác các vùng xuất huyết với sai số thể tích nhỏ, giúp bác sĩ đưa ra các quyết định lâm sàng nhanh chóng và chính xác. Quá trình đánh giá cho thấy Viola-UNet có thể áp dụng một cách hiệu quả trên các tập dữ liệu y khoa phức tạp, cải thiện quá trình chẩn đoán và điều trị cho bệnh nhân. Mô hình có thể trở thành tiền đề để xây dựng hệ thống phát hiện tự động vùng tụ máu do xuất huyết não.

5. KẾT LUẬN

Trong nghiên cứu này, chúng tôi đã phát triển và đánh giá thành công mô hình Viola U-Net nhằm phát hiện và phân đoạn chính xác các vùng tụ máu do xuất huyết não trên ảnh CT sọ não, với dữ liệu thu thập tại Việt Nam. Kết quả cho thấy Viola U-Net đạt độ chính xác cao hơn so với các mô hình phổ biến như nnU-Net, đặc biệt ở các chỉ số DSC, HD, NSD và RVD, chứng minh hiệu quả trong việc xử lý các trường hợp xuất huyết phức tạp. Mô hình không chỉ nâng cao độ tin cậy trong hỗ trợ chẩn đoán hình ảnh mà còn mở ra tiềm năng ứng dụng trong lâm sàng, góp phần giảm tải công việc cho bác sĩ và hỗ trợ phát hiện sớm xuất huyết não. Trong tương lai, việc mở rộng bộ dữ liệu, tối ưu thuật toán và triển khai thử nghiệm trên diện rộng sẽ là những hướng nghiên cứu cần thiết để hoàn thiện và ứng dụng mô hình trong thực tiễn.

6. TÀI LIỆU THAM KHẢO

- [1] K. He, X. Zhang, S. Ren, and J. Sun (2016). Deep Residual Learning for Image Recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778. Doi: 10.1109/CVPR.2016.90.
- [2] Y. LeCun, Y. Bengio, and G. Hinton (2015). Deep learning, *Nature*, vol. 521, no. 7553, pp. 436–444. Doi: 10.1038/nature14539.
- [3] O. Ronneberger, P. Fischer, and T. Brox (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation, In *International Conference on Medical image computing and computer-assisted intervention. Cham: Springer international publishing*, pp. 234–241. Doi: 10.1007/978-3-319-24574-4_28.
- [4] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI, S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, Eds., Cham: Springer International Publishing*, pp. 424–432. Doi: 10.1007/978-3-319-46723-8_49.
- [5] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation, *Nat. Methods*, vol. 18, no. 2, pp. 203–211. Doi: 10.1038/s41592-020-01008-z.
- [6] F. Milletari, N. Navab, and S.-A. Ahmadi (2016). V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation, in *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571. Doi: 10.1109/3DV.2016.79.
- [7] J. Y. Lee, J. S. Kim, T. Y. Kim, and Y. S. Kim (2020). Detection and classification of intracranial haemorrhage on CT images using a novel deep-learning algorithm, *Sci. Rep.*, vol. 10, no. 1, pp. 20546. Doi: 10.1038/s41598-020-77441-z.
- [8] Q. T. Hoang, X. H. Pham, X. T. Trinh, A. V. Le, M. V. Bui, and T. T. Bui (2024). An Efficient CNN-Based Method for Intracranial Hemorrhage Segmentation from Computerized Tomography Imaging, *J. Imaging*, vol. 10, no. 4, Art. no. 4, pp. 77. Doi: 10.3390/jimaging10040077.
- [9] P. Kadam, J. Raphael, P. Karale, I. D’silva, and K. Sonawane (2021). A CNN-RNN based approach for Simultaneous Detection, Identification and Classification of Intracranial Hemorrhage, in *2021 International Conference on Communication information and Computing Technology (ICCICT)*, pp. 1–6. Doi: 10.1109/ICCICT50803.2021.9510112.
- [10] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, ... & Ng, A. Y. (2017). Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv. arXiv preprint arXiv:1711.05225*, 10. Doi: 10.48550/arXiv.1711.05225.
- [11] M. D. Hssayeni, M. S. Croock, A. D. Salman, H. F. Al-Khafaji, Z. A. Yahya, & B. Ghoraani (2020). Intracranial hemorrhage segmentation using a deep convolutional model. *Data*, 5(1), 14.
- [12] J. Ker, L. Wang, J. Rao, and T. Lim (2018). Deep Learning Applications in Medical Image Analysis, *IEEE Access*, vol. 6, pp. 9375–9389. Doi: 10.1109/ACCESS.2017.2788044.
- [13] S. Walluscheck, A. Gerken, I. Galinovic, K. Villringer, J. B. Fiebach, J. Klein & S. Heldmann (2024). Generative adversarial network-based reconstruction of healthy anatomy for anomaly detection in brain CT scans, *J. Med. Imaging Bellingham Wash*, vol. 11, no. 4, p. 044508. Doi: 10.1117/1.JMI.11.4.044508.
- [14] Q. Liu, B. J. MacIntosh, T. Schellhorn, K. Skogen, K. Emblem, and A. Bjørnerud (2024). Voxels Intersecting along Orthogonal Levels Attention U-Net for Intracerebral Haemorrhage Segmentation in Head CT, *arXiv.org*. [Online]. Available: <https://arxiv.org/abs/2208.06313v2>.

AN AI MODEL FOR DETECTING HEMATOMA REGIONS CAUSED BY INTRACEREBRAL HEMORRHAGE IN BRAIN CT SCANS

Nguyen Thi Bich Ha, Pham Tran Thien, Le Thi Kim Nga

¹Quy Nhon College of Engineering and Technology

²Department of Information Technology, Quy Nhon University.

ABSTRACT— This study proposes an artificial intelligence model based on the Viola-UNet architecture for detecting hematoma regions in brain CT scans of patients in Vietnam. Viola-UNet integrates the U-Net framework with advanced deep learning techniques, including the Viola attention module, to enhance feature extraction and improve segmentation accuracy. Results from five-fold cross-validation show that Viola-UNet significantly outperforms nnUNet, particularly in terms of the Dice Similarity Coefficient (DSC). Additionally, Viola-UNet achieves better results across other evaluation metrics such as Hausdorff Distance (HD), Normalised Surface Dice (NSD), and Relative Volume Difference (RVD). These findings suggest that Viola-UNet holds strong potential for clinical application, supporting the timely and accurate detection of hemorrhagic brain lesions.

Keywords— Medical AI, Viola-UNet, Brain CT, Intracerebral Hemorrhage, Hematoma Segmentation.



Nguyễn Thị Bích Hà là thạc sĩ khoa học máy tính. Hiện đang là giảng viên khoa Điện tử - Tin học, Trường Cao đẳng kỹ thuật công nghệ Quy Nhơn. Lĩnh vực nghiên cứu liên quan đến xử lý ảnh và video



Phạm Trần Thiện là thạc sĩ công nghệ phần mềm. Hiện đang là giảng viên của Khoa Công nghệ thông tin, Trường Đại học Quy Nhơn. Lĩnh vực nghiên cứu liên quan đến xử lý ảnh và video; kỹ thuật phần mềm.



Lê Thị Kim Nga là tiến sĩ chuyên ngành khoa học máy tính. Hiện đang là giảng viên Khoa Công nghệ thông tin của Trường Đại học Quy Nhơn. Lĩnh vực nghiên cứu chính là Thị giác máy tính, Trí tuệ nhân tạo, 3D/VR/AR.