JTE HCMUTE
ISSN: 1859-1272

TẠP CHÍ KHOA HỌC GIÁO DỤC KỸ THUẬT
Trường Đại học Sư phạm Kỹ thuật Thành phố Hồ Chí Minh
Website: https://jte.edu.vn
Email: jte@hcmute.edu.vn

Improved Brain Hemorrhage Segmentation Using an Optimized U-Net With Residual Blocks

Hieu Thanh Do[1*], Hung Manh Nguyen[2], Tuan Quoc Hoang[1]
[1]*Hung Yen University of Technology and Education, Vietnam*
[2]*College of Mechanized Construction, Vietnam*

[*]Corresponding author. *Email: dothanhhieukt@gmail.com*

ARTICLE INFO

Received: 24/07/2025
Revised: 08/08/2025
Accepted: 28/08/2025
Published: 28/11/2025

KEYWORDS

Intracranial hemorrhage segmentation;
Computed Tomography;
Convolutional Neural Network;
U-Net;
ResUNet;
Graphics processing unit.

ABSTRACT

Brain hemorrhage is a life-threatening emergency that demands rapid and accurate diagnosis for timely treatment. Currently, computed tomography (CT) is the primary imaging modality, but traditional segmentation techniques still exhibit limited accuracy. In this study, three variants of the U-Net architecture—U-Net integrated with VGG16, U-Net integrated with ResNet-18, and Residual U-Net—were compared to identify the optimal solution. Notably, the Residual U-Net leverages shortcut connections and residual blocks to improve deep learning performance even when training data are scarce. To balance background and hemorrhage segmentation, the training process employs a combined loss function of Binary Cross Entropy and Dice Loss. Evaluation results demonstrate that the Residual U-Net outperforms the other two variants in accuracy and key performance metrics, even on small datasets. Thanks to its effective feature reuse and optimized loss function, the Residual U-Net shows great promise as a powerful clinical support tool for enhancing the effectiveness of cerebral hemorrhage segmentation in CT images.

Nâng cao hiệu quả phân vùng xuất huyết não bằng U-Net được tối ưu hóa với Residual Blocks

Đỗ Thành Hiếu[1*], Nguyễn Mạnh Hùng[2], Hoàng Quốc Tuân[1]
[1]*Trường Đại học Sư phạm Kỹ thuật Hưng Yên, Việt Nam*
[2]*Trường Cao đẳng Cơ giới Xây dựng, Việt Nam*

[*]Tác giả liên hệ. *Email: dothanhhieukt@gmail.com*

THÔNG TIN BÀI BÁO

Ngày nhận bài: 24/07/2025
Ngày hoàn thiện: 08/08/2025
Ngày chấp nhận đăng: 28/08/2025
Ngày đăng: 28/11/2025

TỪ KHÓA

Phân vùng xuất huyết não;
Chụp cắt lớp vi tính (CT);
Mạng nơ-ron tích chập;
U-net;
ResUNet;
Bộ xử lý đồ họa.

TÓM TẮT

Xuất huyết não là một tình trạng cấp cứu nguy hiểm, đòi hỏi chẩn đoán nhanh và chính xác để điều trị kịp thời. Hiện nay, chụp cắt lớp vi tính (CT) là phương pháp hình ảnh chủ yếu, nhưng các kỹ thuật phân vùng truyền thống vẫn còn hạn chế về độ chính xác. Trong nghiên cứu này, ba biến thể của kiến trúc U-Net — gồm U-Net tích hợp VGG16, U-Net tích hợp ResNet-18 và Residual U-Net — được so sánh để tìm ra giải pháp tối ưu. Đặc biệt, Residual U-Net tận dụng các kết nối tắt (shortcut connections) và khối Residual giúp mô hình học sâu hiệu quả ngay cả khi dữ liệu huấn luyện hạn chế. Để cân bằng giữa việc phân vùng vùng nền và vùng xuất huyết, quá trình huấn luyện áp dụng hàm mất mát kết hợp Binary Cross Entropy và Dice Loss. Kết quả đánh giá cho thấy Residual U-Net vượt trội về độ chính xác và các chỉ số quan trọng so với hai biến thể còn lại, ngay cả với tập dữ liệu nhỏ. Nhờ khả năng tái sử dụng đặc trưng và hàm mất mát được tối ưu, Residual U-Net hứa hẹn trở thành một công cụ mạnh mẽ hỗ trợ chẩn đoán lâm sàng, nâng cao hiệu quả phân vùng xuất huyết não trên ảnh CT.

Doi: https://doi.org/10.54644/jte.2025.1968

Copyright © JTE. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License which permits unrestricted use, distribution, and reproduction in any medium for non-commercial purpose, provided the original work is properly cited.

1. Giới thiệu

Xuất huyết não là một dạng đột quỵ nghiêm trọng, xảy ra do sự tràn máu vào nhu mô não, gây tổn thương thần kinh và có thể dẫn đến tử vong nếu không được chẩn đoán và điều trị kịp thời. Phương pháp chẩn đoán ban đầu thường dựa trên hình ảnh chụp cắt lớp vi tính (Computed Tomography - CT) và chụp cộng hưởng từ (Magnetic Resonance Imaging - MRI). Tuy nhiên, các phương pháp này vẫn tồn tại những hạn chế nhất định. Đặc biệt trong các tình huống cấp cứu, việc phát hiện các tổn thương nhỏ hoặc hình ảnh không điển hình có thể gặp khó khăn khi bác sĩ phải đánh giá một lượng lớn hình ảnh liên tục, dẫn đến sự chậm trễ trong quyết định điều trị. Thêm vào đó, độ chính xác của chẩn đoán phụ thuộc nhiều vào kinh nghiệm và trình độ chuyên môn của bác sĩ, có thể gây ra sự khác biệt trong việc đánh giá giữa các chuyên gia.

Sự phát triển vượt bậc của các mô hình học sâu, đặc biệt là mạng nơ-ron tích chập (Convolutional Neural Network - CNN), đã mang lại một cuộc cách mạng trong lĩnh vực trí tuệ nhân tạo. Nhờ những tiến bộ này, nhiều mô hình học sâu đã được nghiên cứu và ứng dụng rộng rãi trong y học để hỗ trợ các tác vụ phát hiện, phân loại và phân vùng tổn thương [1] - [5].

Sự tiến bộ của trí tuệ nhân tạo (AI) và học sâu đã thúc đẩy sự ra đời của các hệ thống hỗ trợ chẩn đoán như CADe (Computer-Aided Detection) và CADx (Computer-Aided Diagnosis), giúp bác sĩ đưa ra quyết định chẩn đoán nhanh chóng và chính xác hơn. Trong lĩnh vực nghiên cứu về xuất huyết não, đã có nhiều công trình tập trung vào việc phát hiện và phân loại. Các nghiên cứu trước đây đã sử dụng cả mạng nơ-ron 2 chiều (2D CNN) [1] và mạng nơ-ron 3 chiều (3D CNN) [2] - [5] cho mục đích này. Ví dụ, một số nhóm nghiên cứu đã áp dụng phương pháp 2D kết hợp học chuyển giao với các kiến trúc như GoogLeNet và Inception-ResNet. Một nghiên cứu khác đã thực hiện phân loại nhị phân xuất huyết bằng mô hình 3D có tên RADnet [5].

U-Net là một kiến trúc mạng nơ-ron tích chập được thiết kế để phân đoạn hình ảnh y sinh một cách nhanh chóng và chính xác [6]. Được phát triển ban đầu bởi Olaf Ronneberger và cộng sự cho việc phân vùng các cấu trúc nơ-ron trong não người, U-Net với cấu trúc encoder-decoder đặc trưng đã chứng minh được hiệu quả vượt trội trong phân vùng ảnh y tế [6] - [11]. Kiến trúc này cho phép trích xuất hiệu quả các đặc trưng quan trọng từ ảnh y tế đồng thời bảo toàn được các chi tiết nhỏ, từ đó giúp nhận diện chính xác các vùng bệnh lý như xuất huyết não, khối u và tổn thương mô.

Tuy nhiên, dữ liệu hình ảnh y tế vẫn tồn tại nhiều thách thức đối với việc huấn luyện các mô hình học sâu. Trong nghiên cứu này, chúng tôi xác định ba vấn đề chính: số lượng dữ liệu huấn luyện hạn chế, sự mất cân bằng dữ liệu và độ tương phản thấp, ranh giới tổn thương mờ nhạt. Những thách thức này xuất phát từ quy trình thu thập và gán nhãn dữ liệu y tế phức tạp, kích thước nhỏ của tổn thương so với toàn bộ ảnh, và chất lượng hình ảnh bị ảnh hưởng bởi thiết bị và kỹ năng của kỹ thuật viên. Do đó, để phân vùng chính xác vùng xuất huyết trong ảnh CT não, việc cải tiến các kiến trúc mạng truyền thống và lựa chọn các tham số phù hợp là rất cần thiết để mô hình có thể học hiệu quả và đạt được độ chính xác cao.

Trong nghiên cứu này, chúng tôi tiến hành so sánh hiệu suất của ba biến thể U-Net: U-Net cơ bản với backbone VGG16, U-Net với backbone ResNet-18 và Residual U-Net. Quá trình so sánh được thực hiện trên bộ dữ liệu ảnh CT não nhằm xác định mô hình tối ưu cho việc phân vùng xuất huyết não. Dựa trên kết quả phân tích, chúng tôi đề xuất phương pháp cải tiến U-Net bằng cách tích hợp các Residual Blocks, với mục tiêu nâng cao độ chính xác và hiệu quả phân vùng xuất huyết não.

2. Tập dữ liệu sử dụng

Trong nghiên cứu này, chúng tôi sử dụng tập dữ liệu công khai ct-ich [12]. Tập dữ liệu này được xây dựng từ hình ảnh chụp CT của 75 bệnh nhân bị chấn thương sọ não. Đáng chú ý, 36 trong số 75 ca này được chẩn đoán xuất huyết não, được phân loại thành 5 kiểu chính: xuất huyết trong não thất, xuất huyết trong nhu mô, xuất huyết dưới nhện, xuất huyết ngoài màng cứng và xuất huyết dưới màng cứng.

Mỗi ca chụp CT bao gồm trung bình 30 lát cắt, với độ dày mỗi lát là 5 mm. Quá trình chú thích dữ liệu được thực hiện bởi hai bác sĩ chẩn đoán hình ảnh một cách độc lập. Họ đã phân loại từng lát cắt theo kiểu xuất huyết và tình trạng vỡ sọ, với sự đồng thuận cao giữa hai người. Một công cụ chuyên

dụng trong Matlab đã được sử dụng để số hóa các chú thích này, khoanh vùng khu vực xuất huyết và biểu diễn chúng dưới dạng ảnh mask (vùng trắng trên nền đen). Các ảnh mask này sau đó được liên kết tương ứng với từng ảnh CT gốc. Cuối cùng, cả ảnh CT và ảnh mask tương ứng đều được lưu trữ dưới định dạng NIfTI, tổng cộng thu được 75 cặp tệp NIfTI (mỗi cặp gồm ảnh CT và ảnh mask).

Hình 1 minh họa một ví dụ về ảnh CT và ảnh mask (tên tệp 94z27) từ tập dữ liệu [12]. Thông tin chi tiết hơn về các bệnh nhân chấn thương sọ não được trình bày trong Bảng 1 và Bảng 2.

(a) *(b)*

Hình 1. *Hình ảnh lát cắt thứ 27 của ảnh CT số 94 trong tập dữ liệu [15] (a): ảnh CT; (b): ảnh mask*

Bảng 1. *Thông tin về các bệnh nhân bị chấn thương sọ não trong công bố [15]*

Số bệnh nhân (Số ảnh chụp CT)	75
Giới tính	42 Nam; 33 Nữ
Độ tuổi	Từ 1 ngày tuổi đến 72 tuổi
Độ tuổi trung bình	27,74
Số ca bị chẩn đoán xuất huyết não	36
Số ca bị chẩn đoán vỡ sọ	22
Tổng số lát cắt	2814
Số lát cắt trung bình của từng ảnh CT	30

Bảng 2. *Phân loại kiểu xuất huyết não ở tập dữ liệu trong công bố [15]*

Tổng số lát cắt	Xuất huyết ngoài màng cứng	Xuất huyết trong nhu mô	Xuất huyết trong não thất	Xuất huyết dưới nhện	Xuất huyết dưới màng cứng	Tổng số lát cắt bị chẩn đoán xuất huyết	Tổng số lát cắt không có vùng xuất huyết
2814	173	73	24	18	56	318	2496

3. Kiến trúc U-Net được nghiên cứu

3.1. U-Net với backbone VGG16

VGG16, một mạng nơ-ron tích chập sâu, được K. Simonyan và A. Zisserman từ Đại học Oxford đề xuất vào năm 2014 [13]. Kiến trúc này nổi bật với 16 lớp sâu, bao gồm 13 lớp tích chập chuyên biệt để trích xuất đặc trưng hình ảnh và 3 lớp kết nối đầy đủ dùng cho tác vụ phân loại.

Trong nghiên cứu của chúng tôi, chúng tôi sử dụng một biến thể của U-Net được gọi là U-Net với backbone VGG16 (U-Net+VGG16). Trong kiến trúc này, phần mã hóa (Encoder) của U-Net được thay thế bằng các lớp tích chập của VGG16. Việc này giúp tận dụng triệt để khả năng học và trích xuất các đặc trưng mạnh mẽ từ ảnh đầu vào của VGG16. Trong khi đó, phần giải mã (Decoder) của mạng vẫn giữ nguyên cấu trúc và chức năng như trong kiến trúc U-Net cơ bản.

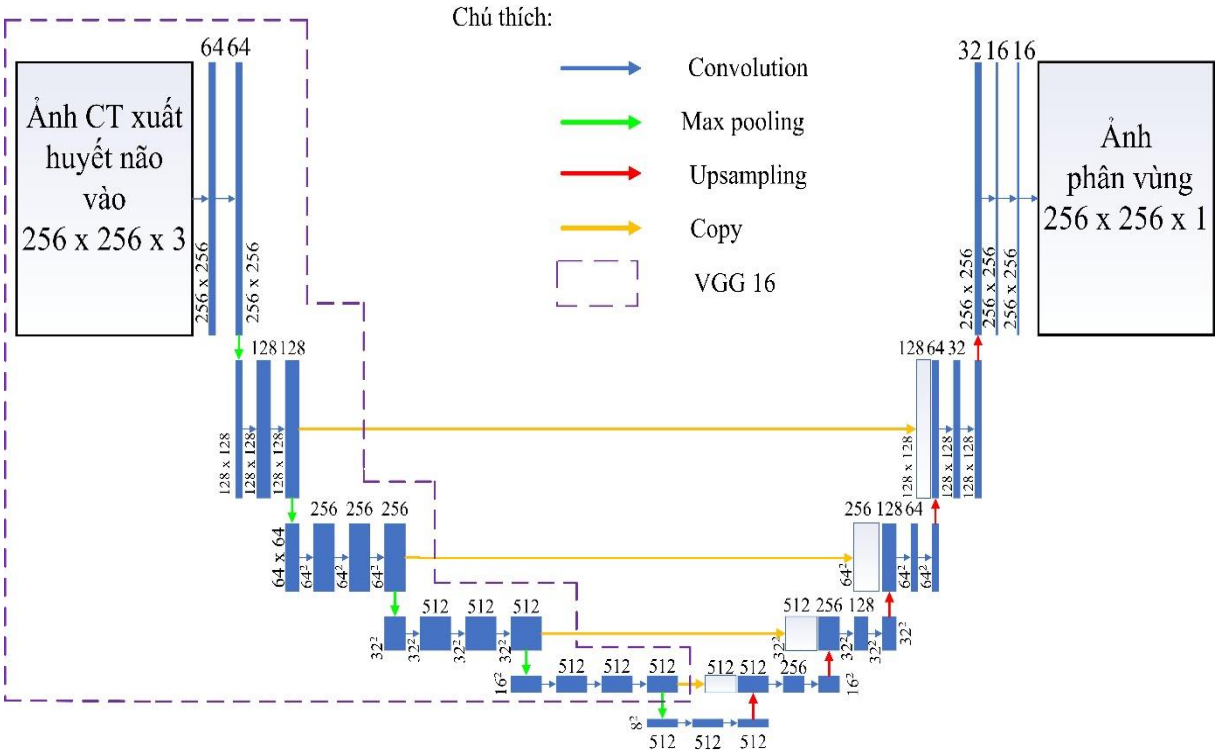
Kiến trúc U-Net+VGG16 được sử dụng trong nghiên cứu này được minh họa chi tiết trong Hình 2.

3.2. U-Net với backbone ResNet-18

Mạng ResNet (Residual Network) được giới thiệu vào năm 2015 bởi nhóm nghiên cứu của He Kaiming, Sun Jian và các đồng nghiệp tại Microsoft Research Asia. ResNet đã cho thấy hiệu quả vượt trội trong các tác vụ thị giác máy tính và có nhiều biến thể như ResNet-18, ResNet-34 và ResNet-50 [14].

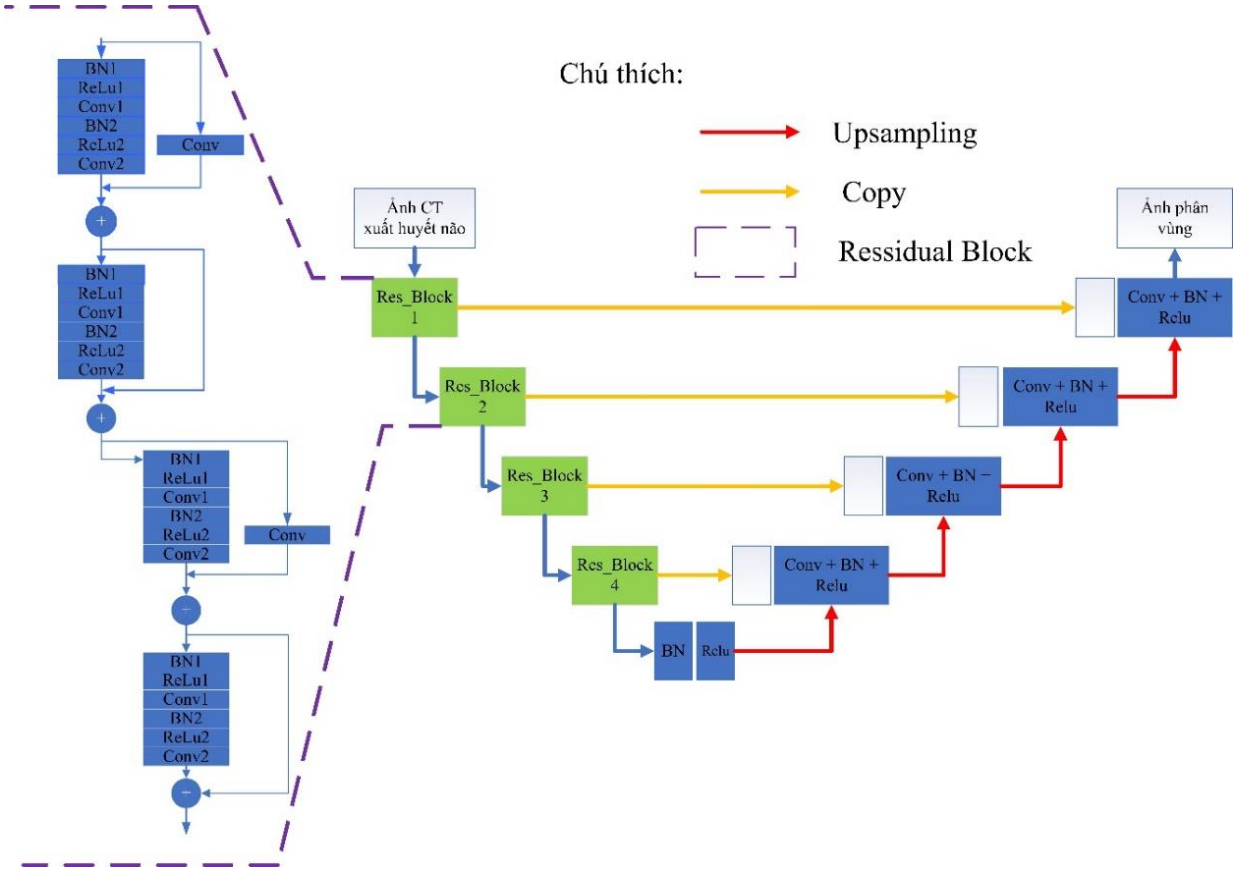
Kỹ thuật cốt lõi làm nên sức mạnh của ResNet là Residual Block. Residual Block giải quyết hiệu quả vấn đề vanishing gradient (khó khăn trong việc cập nhật trọng số của các lớp ban đầu khi mạng quá sâu do gradient giảm dần về 0). Thay vì chỉ cho phép dữ liệu đi qua toàn bộ các lớp và bị biến đổi hoàn toàn, Residual Block duy trì thông tin gốc bằng cách thêm một shortcut connection. Shortcut connection này cộng trực tiếp tín hiệu đầu vào của khối với đầu ra, cho phép thông tin được truyền thẳng, giúp mô hình học nhanh và hiệu quả hơn mà vẫn giữ được thông tin.

ResNet-18, một biến thể phổ biến, có tổng cộng 18 lớp: một lớp Convolution đầu vào, 16 lớp tích chập được tổ chức thành 4 Residual Block, và một lớp Fully Connected ở đầu ra để phân loại.



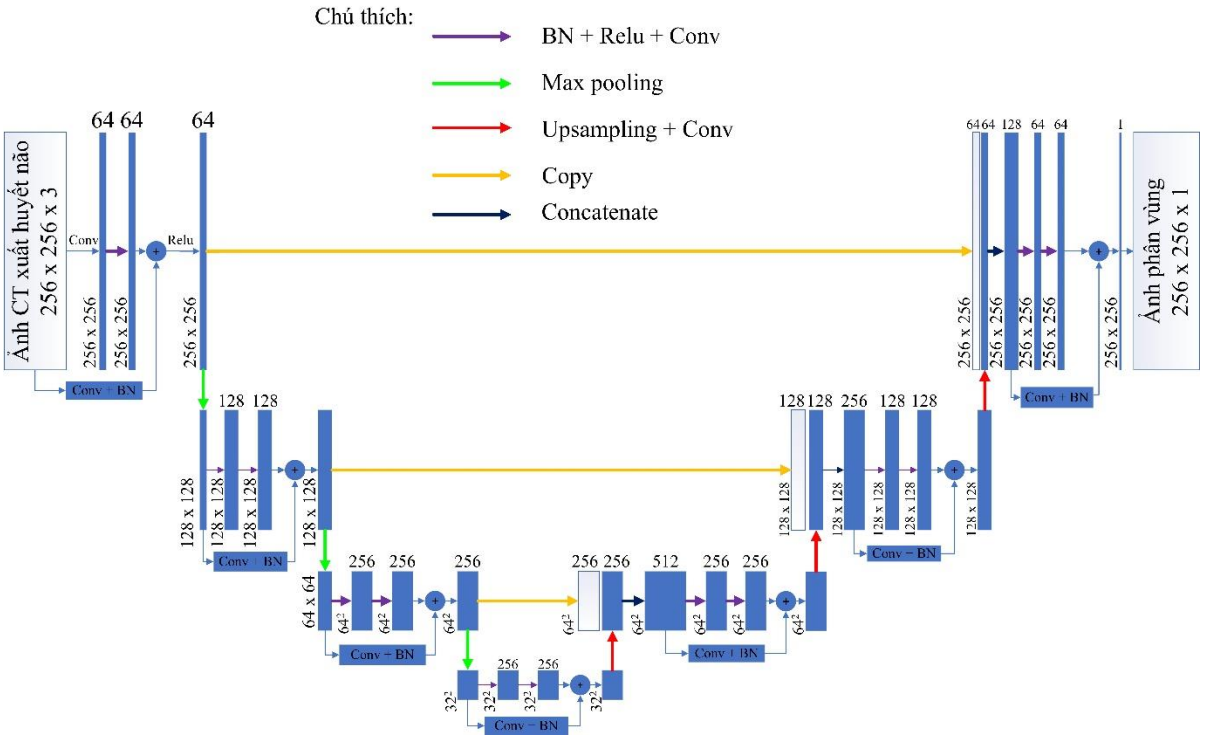
Hình 2. *Kiến trúc U-Net +VGG16 được sử dụng trong nghiên cứu*

Hình 3 trình bày sơ đồ kiến trúc U-Net + ResNet-18, được áp dụng trong nghiên cứu của chúng tôi. Trong thực nghiệm, chúng tôi sử dụng một biến thể của U-Net gọi là U-Net với backbone ResNet-18 (U-Net + ResNet-18). Kiến trúc này tận dụng ResNet-18 để cải thiện hiệu suất của U-Net trong việc trích xuất đặc trưng và giảm thiểu mất thông tin. Cụ thể, phần encoder của U-Net truyền thống được thay thế bằng ResNet-18, giúp tăng cường khả năng trích xuất đặc trưng mạnh mẽ hơn. Phần bottleneck của mạng chính là tầng cuối cùng của ResNet-18, đóng vai trò kết nối với phần đầu của decoder U-Net. Phần decoder của kiến trúc này giữ nguyên kiến trúc của mạng U-Net cơ bản. Sự kết hợp này giúp tối ưu hóa quá trình huấn luyện và nâng cao độ chính xác của mô hình.



Hình 3. *Kiến trúc U-Net + ResNet-18 được sử dụng trong nghiên cứu*

3.3. Residual U-Net



Hình 4. *Kiến trúc ResUNet được sử dụng trong nghiên cứu*

Residual U-Net (ResUNet) là một biến thể của U-Net, trong đó cả hai phần encoder và decoder đều được cải tiến bằng cách sử dụng các Residual Blocks thay vì các lớp convolution truyền thống. Việc tích hợp Residual Blocks vào cả encoder và decoder giúp nâng cao khả năng học đặc trưng, đồng thời bảo tồn thông tin xuyên suốt quá trình trích xuất và tái tạo, từ đó nâng cao độ chính xác trong phân vùng.

Hình 4 trình bày sơ đồ kiến trúc của ResUNet được áp dụng trong nghiên cứu thực nghiệm của chúng tôi. Mạng bao gồm 7 khối Residual, trong đó 6 khối được phân bổ đều cho hai phần Encoder và Decoder. Khối Residual thứ 4 đóng vai trò là Bottleneck trong kiến trúc của mạng U-Net. Các khối Residual trong ResNet có cấu trúc tương tự như trong mạng U-Net+ResNet18, bao gồm 2 lớp Batch Normalization để ổn định giá trị đầu vào, 2 lớp ReLU Activation giúp cải thiện khả năng học phi tuyến tính và 2 lớp Conv dùng để trích xuất đặc trưng. Shortcut Connection bao gồm 1 lớp Conv và 1 lớp Batch Normalization ở tất cả các khối, giúp cộng trực tiếp tín hiệu đầu vào với đầu ra. Ở phần Decoder, ngay sau khi thực hiện Upsampling, một lớp tích chập được áp dụng trước khi tiến hành phép Concat.

4. Thực nghiệm

4.1. Tiền xử lý dữ liệu

Trong nghiên cứu của chúng tôi, chúng tôi sử dụng công cụ Python có sẵn cùng với tập dữ liệu "ct-ich" để giải nén các tệp NIfTI thành ảnh có định dạng "*.png" với kích thước 512 x 512 pixel. Kết quả thu được 2814 ảnh (tương ứng với 2814 lát cắt), trong đó 2496 ảnh không chứa vùng xuất huyết, 318 ảnh có xuất huyết, được phân thành 5 loại khác nhau. Phần lớn ảnh chỉ có một loại xuất huyết, nhưng cũng có một số ảnh có nhiều hơn một loại xuất huyết. Ảnh mask thu được có nền đen (pixel 0) và vùng xuất huyết màu trắng (pixel 255) [như thể hiện ở Hình 1].

Mục tiêu của cả hai thực nghiệm là mô hình U-Net chỉ cần thực hiện phân vùng xuất huyết não (có hoặc không có xuất huyết), mà không cần phân biệt loại xuất huyết. Vì vậy, chúng tôi đã chuyển đổi các pixel trong ảnh mask, với pixel 0 tương ứng với vùng bình thường và pixel 1 đại diện cho vùng xuất huyết. Trong cả hai thực nghiệm, chúng tôi sử dụng một công cụ có sẵn trong thư viện Sklearn để chia ngẫu nhiên tập dữ liệu thành hai phần: tập huấn luyện và tập xác thực, với tỷ lệ 80% và 20% tương ứng.

Ở thực nghiệm thứ nhất, 318 ảnh CT và ảnh mask tương ứng được resize về kích thước 256 x 256 pixel mà không thực hiện thêm bất kỳ biện pháp xử lý nào khác. Ở thực nghiệm thứ hai, chúng tôi áp dụng một số biện pháp gia tăng dữ liệu đối với 318 ảnh có kích thước 256 x 256 từ thực nghiệm đầu tiên. Giả thiết rằng khi bệnh nhân chụp CT, tư thế đầu của họ có thể thay đổi (ngửa lên, cúi xuống, hoặc quay trái, quay phải), và kích thước hộp sọ của bệnh nhân cũng có sự khác biệt. Do đó, chúng tôi thực hiện các biện pháp sau: xoay ảnh 10 độ theo chiều ngược kim đồng hồ, dịch ảnh xuống 50 pixel và thu nhỏ ảnh 5%. Những biện pháp xử lý này được áp dụng đồng nhất cho cả ảnh CT và ảnh mask để đảm bảo sự nhất quán. Kết quả thu được tổng cộng 1272 ảnh CT và 1272 ảnh mask, tất cả đều có kích thước 256x256 và được lưu dưới định dạng *.png."

4.2. Môi trường tiến hành thực nghiệm

Chúng tôi sử dụng Google Colaboratory với GPU để huấn luyện ba mô hình. Cấu hình phần cứng của Google Colab được trình bày trong Bảng 3.

Bảng 3. *Cấu hình phần cứng của Google Colab*

Thông số	Google Colaboratory	Các module, thư viện được sử dụng
No CPU cores	2	Numpy OpenCV Matplotlib Tensorflow Segmentation_models [15] Keras Sklearn, ski-image
RAM	12 GB	
Disk space	358 GB	
GPU	Tesla T4	
GPU Memory clock	0.82/1.59 GHz	
GPU Memory	16 GB	
No GPU cores	40	

4.3. Thực nghiệm

Đầu tiên, chúng tôi xây dựng ba mô hình U-Net+VGG16, U-Net+ResNet18 và ResUNet theo kiến trúc đã trình bày trong phần 3. Để xác định mô hình phân vùng tối ưu nhất cho tập dữ liệu, chúng tôi tiến hành huấn luyện cả ba mô hình trong các điều kiện giống nhau. Dựa trên kết quả thu được, chúng tôi tiến hành đánh giá, phân tích và so sánh hiệu quả của ba mô hình này.

Bảng 4. *Bảng thông số cấu hình của ba kiến trúc Unet được sử dụng trong thực nghiệm*

	U-Net +VGG16	**U-Net + ResNet-18**	**ResUNet**
Số lớp tích chập	26	32	25
Shortcut Connection	Không	Chỉ có ở Encoder	Có cả ở Encoder và Decoder
Tổng số tham số	23 752 273 (90,61 MB)	14 340 570 (54,70 MB)	7 598 081 (28,98 MB)
Tham số có thể huấn luyện	23 748 241 (90,59 MB)	14 330 644 (54,67 MB)	7 590 785 (28,96 MB)

Trong quá trình huấn luyện mô hình học sâu, hàm mất mát (Loss Function) đóng vai trò then chốt. Hàm này định lượng sự khác biệt giữa kết quả dự đoán của mô hình và nhãn thực tế, từ đó cung cấp tín hiệu để mô hình điều chỉnh các tham số nhằm tối ưu hóa hiệu suất. Trong nghiên cứu này, chúng tôi đề xuất một hàm mất mát hỗn hợp, kết hợp Binary Cross Entropy (BCE) và Dice Loss.

$$Loss = L_{BCE} + Diceloss \tag{1}$$

BCE là một trong những hàm mất mát cơ bản và phổ biến nhất trong học sâu, đặc biệt hiệu quả với các mô hình CNN trong các bài toán phân loại nhị phân, ví dụ như phân biệt giữa trạng thái "có xuất huyết" và "không xuất huyết". BCE đo lường mức độ khác biệt giữa phân phối xác suất dự đoán của mô hình và phân phối xác suất thực tế của các lớp.

$$L_{BCE} = \frac{-1}{n}\sum_{i=1}^{n} Y_i.\log Y_i + (1-Y_i).\log(1-Y_i) \tag{2}$$

Dice Loss là một hàm mất mát được sử dụng rộng rãi trong các bài toán phân vùng ảnh, đặc biệt là trong lĩnh vực y tế. Ưu điểm của Dice Loss là khả năng giúp mô hình tối ưu hóa việc xác định các vùng tổn thương có kích thước nhỏ so với tổng thể ảnh. Giá trị của Dice Loss dao động từ 0 đến 1, trong đó giá trị gần 0 cho thấy sự chồng lấp hoàn hảo giữa vùng dự đoán và vùng thực tế, đồng nghĩa với việc mô hình hoạt động hiệu quả.

$$Diceloss\left(y, \overline{p}\right) = 1 - \frac{2y\overline{p}+1}{y+\overline{p}+1} \tag{3}$$

Trong bài toán phân vùng ảnh, mục tiêu không chỉ là gán nhãn đúng cho từng pixel mà còn phải đảm bảo mối liên hệ giữa các pixel lân cận để tạo thành một hoặc nhiều vùng có biên dạng rõ ràng. Do đó, các chỉ số đánh giá dựa trên diện tích thường được ưu tiên sử dụng để đo lường hiệu quả của mô hình.

Dice Coefficient là một chỉ số rất phổ biến trong phân vùng ảnh y tế, giúp đánh giá mức độ chồng lấp giữa vùng dự đoán và vùng thực tế. Giá trị Dice Score càng gần 1, hiệu suất của mô hình càng tốt. IoU cũng được sử dụng rộng rãi trong cả phân vùng ảnh và nhận diện đối tượng. Giá trị IoU càng cao thể hiện khả năng phân vùng càng chính xác. Ngưỡng IoU thường được coi là tốt khi lớn hơn 0,5 và các mô hình mạnh có thể đạt giá trị trên 0,75.

Trong cả hai thực nghiệm, chúng tôi đã sử dụng hai chỉ số chính là Dice Coefficient và IoU để đánh giá hiệu suất của các mô hình.

$$Dice\ Coefficient = \frac{2|A \cap B|}{|A| + |B|} \quad (4)$$

$$IoU = \frac{A \cap B}{A \cup B} \quad (5)$$

Trong đó: A: Tập hợp các điểm ảnh nằm trong vùng xuất huyết do mô hình dự đoán.

B: Tập hợp các điểm ảnh trong vùng xuất huyết thực tế.

Thông số khởi tạo mô hình và dữ liệu ảnh đầu vào trong hai thực nghiệm được trình bày ở bảng 5:

Bảng 5. *Thông số khởi tạo mô hình và dữ liệu ảnh đầu vào trong hai thực nghiệm*

Thông số	Giá trị	
	Thực nghiệm thứ nhất	***Thực nghiệm thứ hai***
Kích thước ảnh	256 x 256 pixel	256 x 256 pixel
Số lượng ảnh	318	1272
Batch size	16	16
Optimizer	Adam	Adam
Learning Rate	0,0001	0,0001
Kernel initializer	He_normal	He_normal
Loss Function	BCE + Dice	BCE + Dice
Metric	Dice Coefficient, IoU Score	Dice Coefficient, IoU Score
Epoch	50	50

5. Kết quả và thảo luận

Chúng tôi đã tiến hành thu thập các thông số đánh giá như Dice coefficient, IoU Scores và Loss từ ba mô hình. Đồng thời hiển thị kết quả phân vùng xuất huyết não đạt được trên các ảnh CT não.

5.1. Kết quả thực nghiệm thứ nhất

Bảng 6. *Bảng kết quả thực nghiệm thứ nhất*

Thông số	U-Net +VGG16	U-Net + ResNet-18	ResUNet
Loss Funnction	0,4914	0,9836	0,2432
Dice Coefficient	0,5254	0,0998	0,7821
IoU Score	0,2802	0,0442	0,5055
Thời gian huấn luyện	12m49s	9m14s	19m6s

5.2. Kết quả thực nghiệm thứ hai

Bảng 7. *Bảng kết quả thực nghiệm thứ hai*

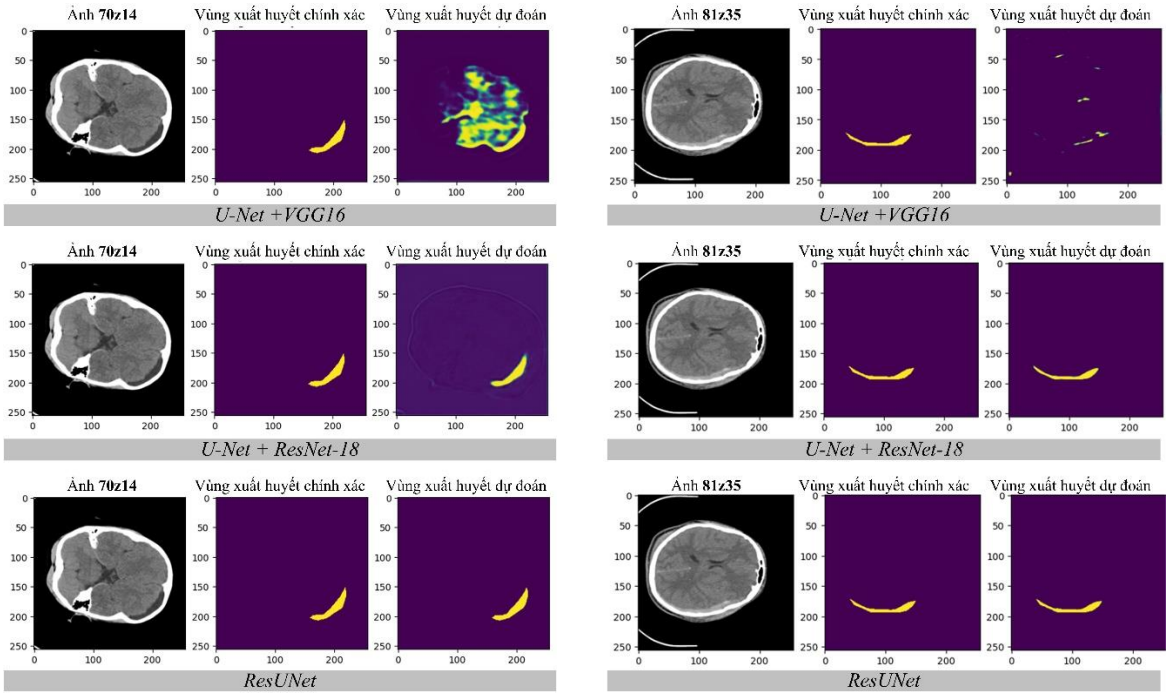
Thông số	U-Net +VGG16	U-Net + ResNet-18	ResUNet
Loss Funnction	0,0905	0,1514	0,1290
Dice Coefficient	0,9152	0,8634	0,8842
IoU Score	0,7594	0,6605	0,7246
Thời gian huấn luyện	38m4s	31m34s	1h8m48s

Với U-Net+VGG16, khả năng phân vùng vùng xuất huyết kém rõ rệt: trong thực nghiệm thứ nhất, hàm mất mát còn cao (0,4914) và hệ số Dice cùng IoU thấp (0,5254 và 0,2802). Dù ở thực nghiệm thứ hai, sau khi tăng cường dữ liệu, hàm mất mát giảm xuống 0,0905, Dice cải thiện lên 0,9152 và IoU lên 0,7594, kết quả thực tế cho thấy mô hình không học được đặc trưng, nhất là với bộ dữ liệu nhỏ, mất cân đối cấu trúc và đối tượng mờ nhạt.

Kiến trúc UNet+ResNet18 cũng cho thấy những hạn chế đáng kể khi đối mặt với dữ liệu ít. Trong thực nghiệm thứ nhất, mô hình gần như không học được đặc trưng của dữ liệu, thể hiện qua giá trị loss rất cao (0,9836) và các chỉ số hiệu suất phân vùng như Dice và IoU cực thấp (lần lượt là 0,0998 và 0,0442). Tuy nhiên, kết quả từ thực nghiệm thứ hai đã chứng minh rằng mô hình này có khả năng cải thiện mạnh mẽ khi được huấn luyện trên một tập dữ liệu lớn hơn. Cụ thể, giá trị loss giảm đáng kể xuống 0,1514, trong khi hệ số Dice và IoU tăng vượt trội, đạt 0,8634 và 0,6605. Điều này cho thấy UNet+ResNet18 là một kiến trúc tiềm năng, nhưng đòi hỏi một bộ dữ liệu đủ lớn để phát huy tối đa hiệu quả.

Ngược lại, Residual UNet nổi bật như là mô hình hiệu quả nhất. Ngay trong thực nghiệm đầu tiên với dữ liệu hạn chế, mô hình đã hội tụ nhanh chóng chỉ sau 5 epoch, cho thấy khả năng học mạnh mẽ. Với cấu trúc residual block được tích hợp trong cả encoder và decoder, mô hình đã đạt hiệu quả phân vùng ấn tượng với giá trị loss là 0,2432, cùng các hệ số Dice và IoU lần lượt là 0,7821 và 0,5055. Khi áp dụng phương pháp tăng cường dữ liệu trong thực nghiệm thứ hai, hiệu suất của Residual UNet tiếp tục được cải thiện rõ rệt. Giá trị loss giảm xuống chỉ còn 0,1290, trong khi các chỉ số Dice và IoU tăng lên mức 0,8842 và 0,7246. Điều này minh chứng cho tính hiệu quả của mô hình trong việc khai thác tối đa lợi thế từ dữ liệu được mở rộng.

Về thời gian huấn luyện, ở thực nghiệm thứ nhất Residual UNet tiêu tốn nhiều nhất (19 phút 6 giây) trong khi U-Net+VGG16 và U-Net+ResNet18 lần lượt là 12 phút 49 giây và 9 phút 14 giây. Khi khối lượng dữ liệu tăng gấp bốn vào thực nghiệm thứ hai, thời gian của Residual UNet kéo dài lên 1 giờ 8 phút 48 giây, so với 38 phút 4 giây của U-Net+VGG16 và 31 phút 34 giây của U-Net+ResNet18.



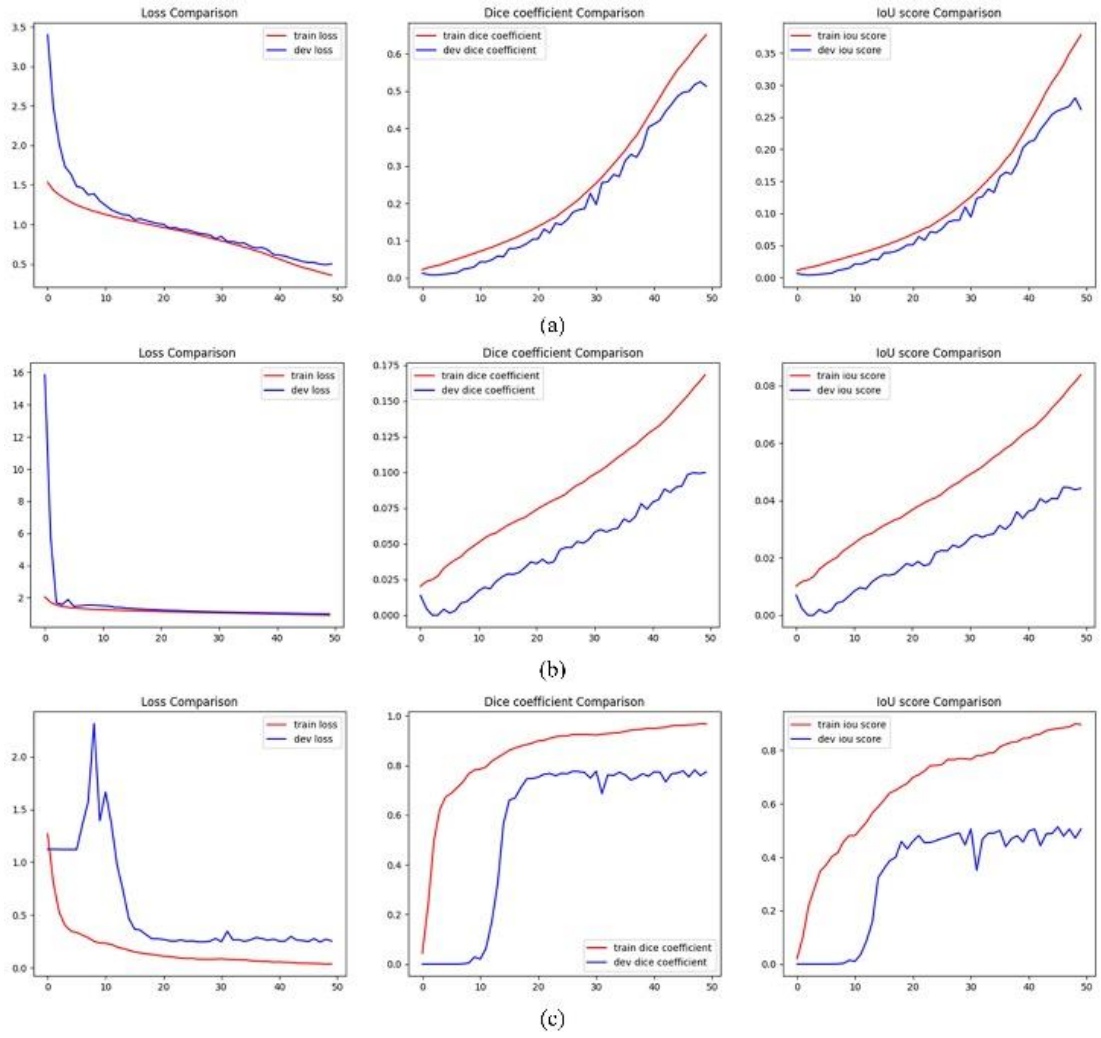
(a) *(b)*

Hình 5. *Kết quả phân vùng xuất huyết não của 3 mô hình*
(a) Thực nghiệm thứ nhất; (b) Thực nghiệm thứ hai

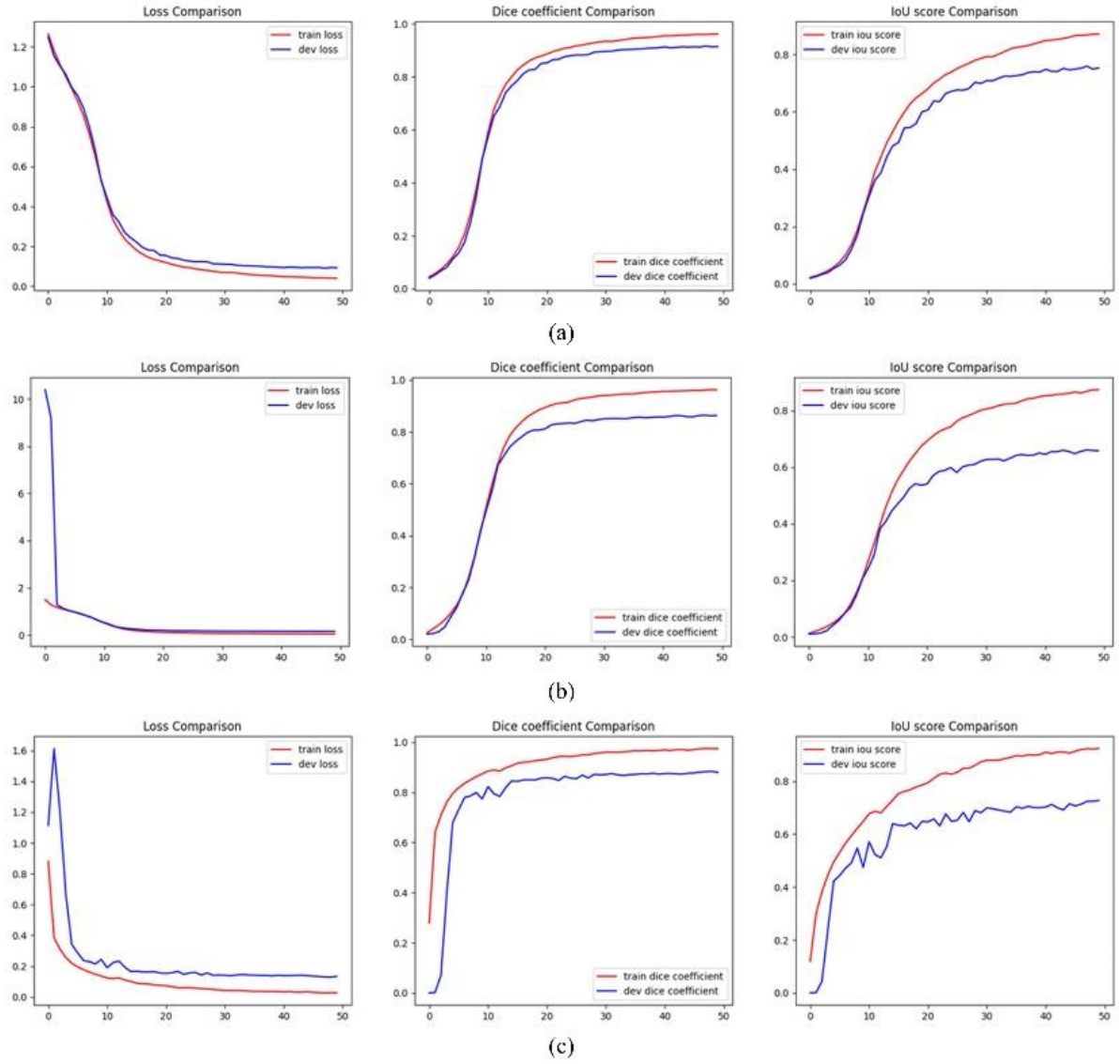
Ngoài ra, để đánh giá toàn diện hiệu quả của mô hình, chúng tôi đã tiến hành so sánh kết quả của Residual UNet với các nghiên cứu điển hình về phân vùng xuất huyết não. Kết quả, được tổng hợp trong Bảng 8, cho thấy các chỉ số hiệu suất của Residual UNet đạt mức tương đương, thậm chí vượt trội so với các phương pháp hiện có. Điều này không chỉ khẳng định tiềm năng lớn của mô hình trong việc nâng cao hiệu quả phân vùng mà còn chỉ ra rằng Residual UNet là một kiến trúc phù hợp và dễ dàng triển khai cho các bài toán phân tích hình ảnh y tế khác.

Bảng 8. *So sánh hiệu quả phân vùng xuất huyết não của Residual UNet với các nghiên cứu điển hình khác*

Nghiên cứu	IoU score	Dice Coefficient
H.S. Bhadauria [7]	0,78	0,87
Manas Kumar Nag [8]	0,55	0,7
Hulin Kuang [9]	x	0,65
Murtadha Hssayeni [10]	x	0,31
Quoc Tuan Hoang [11]	0,8075	x
Giải pháp đề xuất	0,72	0,88



Hình 6. *Biểu đồ Loss, Dice và IoU của mô hình trong quá trình huấn luyện và xác thực – Thực nghiệm thứ nhất (a): U-Net +VGG16; (b): U-Net + ResNet-18; (c) ResUNet*



Hình 7. *Biểu đồ Loss, Dice và IoU của mô hình trong quá trình huấn luyện và xác thực – Thực nghiệm thứ hai (a): U-Net +VGG16; (b): U-Net + ResNet-18; (c) ResUNet*

6. Kết luận

Phân vùng chính xác xuất huyết não trên ảnh CT đóng vai trò then chốt trong quá trình chẩn đoán và lập kế hoạch điều trị, giúp bác sĩ xác định rõ vùng tổn thương để đưa ra quyết định can thiệp tối ưu. Tuy nhiên, đây là một thách thức đáng kể, đặc biệt khi vùng xuất huyết có kích thước nhỏ hoặc dữ liệu ảnh y tế bị hạn chế. Trong nghiên cứu này, chúng tôi đã đề xuất và so sánh hiệu suất của ba kiến trúc học sâu: U-Net với backbone VGG16, U-Net với backbone ResNet-18 và Residual U-Net. Kết quả thực nghiệm cho thấy Residual U-Net đạt được độ chính xác vượt trội, ngay cả trên dữ liệu có số lượng hạn chế và bị mất cân đối, nhờ vào việc tích hợp các khối residual trong cả phần encoder và decoder. Đồng thời, việc kết hợp hàm mất mát BCE và Dice Loss đã cải thiện khả năng học cho các vùng tổn thương nhỏ và trong trường hợp dữ liệu không cân bằng.

Trong tương lai, chúng tôi dự định mở rộng thử nghiệm trên các bộ dữ liệu lớn hơn và đa dạng hơn. Bên cạnh đó, việc áp dụng học chuyển giao hoặc tích hợp các cơ chế attention như Squeeze-and-Excitation (SE), Self-Attention hoặc Transformer Attention hứa hẹn sẽ nâng cao khả năng nhận diện các vùng xuất huyết nhỏ. Những cải tiến này được kỳ vọng sẽ tiếp tục thúc đẩy hiệu quả của các phương pháp phân vùng ảnh y tế, góp phần quan trọng vào việc nâng cao chất lượng chăm sóc sức khỏe.

Xung đột lợi ích

Các tác giả tuyên bố không có xung đột lợi ích trong bài báo này.

TÀI LIỆU THAM KHẢO

[1] P. D. Chang *et al.*, "Hybrid 3D/2D convolutional neural network for hemorrhage evaluation on head CT," *Am. J. Neuroradiol.*, vol. 39, no. 9, pp. 1609–1616, 2018.
[2] P. Inkeaw *et al.*, "Automatic hemorrhage segmentation on head CT scan for traumatic brain injury using 3D deep learning model," *Comput. Biol. Med.*, vol. 146, 2022.
[3] K. Jnawali, M. R. Arbabshirani, N. Rao, and A. A. Patel, "Deep 3D convolution neural network for CT brain hemorrhage classification," in *Proc. SPIE Med. Imaging: Comput.-Aided Diagnosis*, 2018, vol. 10575, Art. no. 105751C.
[4] H. Ye *et al.*, "Precise diagnosis of intracranial hemorrhage and subtypes using a three-dimensional joint convolutional and recurrent neural network," *Eur. Radiol.*, vol. 29, pp. 6191–6201, 2022.
[5] M. Grewal, M. M. Srivastava, P. Kumar, and S. Varadarajan, "RADnet: Radiologist level accuracy using deep learning for hemorrhage detection in CT scans," in *Proc. IEEE Int. Symp. Biomed. Imaging (ISBI)*, Washington, DC, USA, 2018, pp. 281–284, doi: 10.1109/ISBI.2018.8363574.
[6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2015.
[7] H. S. Bhadauria, A. Singh, and M. L. Dewal, "An integrated method for hemorrhage segmentation from brain CT imaging," *Comput. Electr. Eng.*, vol. 39, no. 5, pp. 1527–1536, 2013.
[8] M. K. Nag, S. Chatterjee, A. K. Sadhu, J. Chatterjee, and N. Ghosh, "Computer-assisted delineation of hematoma from CT volume using autoencoder and Chan Vese model," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, pp. 259–269, 2019.
[9] H. Kuang, B. K. Menon, and W. Qiu, "Segmenting hemorrhagic and ischemic infarct simultaneously from follow-up non-contrast CT images in patients with acute ischemic stroke," in *Lecture Notes in Computer Science*, vol. 11766, pp. 856–863, 2019.
[10] M. D. Hssayeni, M. S. Croock, A. D. Salman, H. F. Al-khafaji, Z. A. Yahya, and B. Ghoraani, "Intracranial hemorrhage segmentation using a deep convolutional model," *Data*, vol. 5, no. 1, p. 14, 2020.
[11] Q. T. Hoang, X. H. Pham, X. T. Trinh, A. V. Le, M. V. Bui, and T. T. Bui, "An efficient CNN-based method for intracranial hemorrhage segmentation from computerized tomography imaging," *J. Imaging*, vol. 10, no. 4, p. 77, Mar. 25, 2024.
[12] PhysioNet, "CT Intracranial Hemorrhage dataset (version 1.3.1)," [Online]. Available: https://physionet.org/content/ct-ich/1.3.1
[13] S. Liu and W. Deng, "Very deep convolutional neural network based image classification using small training sample size," in *Proc. IAPR Asian Conf. Pattern Recognit. (ACPR)*, Kuala Lumpur, Malaysia, 2015, pp. 730–734.
[14] M. Shafiq and Z. Gu, "Deep residual learning for image recognition: A survey," *Appl. Sci.*, vol. 12, p. 8972, 2022.
[15] P. Yakubovskiy, *segmentation_models* (Version 1.0.1) [Software]. GitHub, 2022. [Online]. Available: https://github.com/qubvel/segmentation_models

Hieu Thanh Do was born in Hai Duong, Vietnam, in 1985. He received the B.S. degree in electrical engineering from the Hung Yen University of Technology and Education, Hung Yen, in 2008, the M.S. degree in automation engineering from Le Quy Don University, Ha Noi, in 2010, and the Ph.D. degree in electrical engineering and automation from the Hefei University of Technology, Hefei, China, in 2016. Since 2017, he has been a Faculty Member with the Hung Yen University of Technology and Education. His current research interest includes power electronic and its application in renewable energy systems, artificial intelligence.
Email: dothanhhieukt@gmail.com. ORCID: https://orcid.org/0009-0002-2124-9326

Hung Manh Nguyen was born in 1999 in Hai Duong, Vietnam. He graduated from Mechatronics engineering technology at Hung Yen University of Technical Education in 2021. He is currently studying for a Master's degree at Hung Yen University of Technical Education, class code H60232 (2023-2025). He is interested in PLC programing, computer vision and image segmentation.
Email: nguyenmanhhung.aos@gmail.com. ORCID: https://orcid.org/0009-0005-8748-7679

Tuan Quoc Hoang received a engineering's degree in automation engineering from Hung Yen University of Technology and Education, Vietnam, in 2008. Master's degree in automation engineering at Le Quy Don Technical University, Vietnam, in 2012. Ph.D. degree in Electronic engineering at Hung Yen University of Technology and Education, Vietnam, in 2024. His research interests include automation, embedded systems, image processing and digital signal processing
Email: hqtcdt@gmail.com. ORCID: https://orcid.org/0000-0002-7821-248X