

DOI:10.22144/ctujos.2025.096

PHÂN LOẠI UNG THƯ VÚ SỬ DỤNG KẾT HỢP THUẬT TOÁN PHÂN TÍCH CHỤM MỜ VÀ MẠNG HỌC SÂU INCEPTION RESNET - V2

Võ Văn Tài¹, Lê Thị Xuân Mai^{2,3}, Nguyễn Lê Minh Nhật⁴, Trần Lê Hoàng Bảo⁴ và Phạm Toàn Định^{4*}

¹Khoa Khoa học Tự nhiên, Trường Đại học Cần Thơ, Việt Nam

²Đại học Quốc gia Thành phố Hồ Chí Minh, Việt Nam

³Khoa Toán - Tin, Trường Đại học Khoa học Tự nhiên, Thành phố Hồ Chí Minh, Việt Nam

⁴Khoa Kỹ thuật Cơ – Điện và Máy tính, Trường Công nghệ, Đại học Văn Lang, Thành phố Hồ Chí Minh, Việt Nam

*Tác giả liên hệ (Corresponding author): ding.pt@vlu.edu.vn

Thông tin chung (Article Information)

Nhận bài (Received): 24/09/2024

Sửa bài (Revised): 16/10/2024

Duyệt đăng (Accepted): 13/01/2025

Title: Breast cancer classification using a combination of fuzzy clustering algorithm and the Inception ResNet V2 deep learning network

Author(s): Vo Van Tai¹, Le Thi Xuan Mai^{2,3}, Nguyen Le Minh Nhat⁴, Tran Le Bao Hoang⁴ and Pham Toan Dinh^{4*}

Affiliation(s): ¹College of Natural Science, Can Tho University, Can Tho City, Viet Nam; ²Faculty of Mathematics and Computer Science, University of Science, Ho Chi Minh City, Viet Nam; ³Vietnam National University, Ho Chi Minh City, Viet Nam; ⁴Faculty of Mechanical - Electrical and Computer Engineering, School of Technology, Van Lang University, Ho Chi Minh City, Viet Nam

TÓM TẮT

Phân loại ảnh cộng hưởng từ (MRI) là một trong những ứng dụng thực tiễn đã và đang được quan tâm bởi nhiều nhà khoa học trên thế giới. Trong nghiên cứu, thuật toán phân loại ảnh ung thư vú được đề xuất dựa trên sự kết hợp của thuật toán phân tích cụm mờ và kiến trúc Inception ResNet V2 (FCIR-V2). Đầu tiên, thuật toán phân tích cụm mờ được sử dụng để phân đoạn các ảnh thành các lớp tách biệt. Sau đó, kiến trúc học sâu Inception ResNet V2 được áp dụng để phân loại các ảnh sau phân đoạn. Dựa trên các tham số cải tiến mới được thiết lập từ thuật toán phân tích cụm mờ và kiến trúc Inception ResNet V2, một mô hình phân loại mới hiệu quả cho vấn đề phân loại ảnh được thiết lập. Ngoài ra, mô hình này được ứng dụng cụ thể cho tập dữ liệu ung thư vú (Breast Cancer, 400x). Kết quả cho thấy phương pháp sử dụng thuật toán cho mô hình này hiệu quả hơn các phương pháp khác thông qua các chỉ số ACC và F1-Score.

Từ khóa: Inception ResNet-V2, phân loại, phân tích cụm mờ, ung thư vú

ABSTRACT

Magnetic Resonance Imaging (MRI) classification is one of the practical applications that has garnered significant attention from many scientists around the world. This research proposes a breast cancer image classification algorithm based on the combination of fuzzy clustering analysis and the Inception ResNet V2 architecture (FCIR-V2). First, the fuzzy clustering analysis algorithm is used to segment the images into separate classes. Then, the deep learning Inception ResNet V2 architecture is applied to classify the segmented images. Based on newly established improved parameters from the fuzzy clustering analysis and Inception ResNet V2 architecture, an efficient new classification model for image classification is developed. Additionally, this model is specifically applied to the breast cancer dataset (Breast Cancer, 400x). The results showed that the algorithm used for this model is more effective than other methods, based on the ACC and F1-Score metrics.

Keywords: Breast Cancer, classification, fuzzy clustering, inception ResNet-V2

1. GIỚI THIỆU

Phân đoạn ảnh là quá trình phân chia ảnh thành các khu vực không chồng lấn sao cho các điểm ảnh của chúng trong cùng khu vực có đặc điểm tương tự nhau (Lai et al., 2009; Tran, 2016). Phân đoạn là nghiên cứu cần thiết cho các thay đổi cấu trúc và định lượng của ảnh cộng hưởng từ (MRI). Đây cũng là điều kiện tiên quyết cho mô hình hóa trong chuẩn đoán các tổn thương hoặc khối u, vì tổn thương có thể lan rộng theo các tốc độ khác nhau tùy thuộc vào các mô xung quanh. Sự tồn tại của nhiều, trường độ lệch và các hiệu ứng bên trong ảnh tạo ra nhiều thách thức lớn cho việc phân đoạn ảnh MRI nhằm phát hiện các khối u bên trong. Các kỹ thuật phân đoạn ảnh MRI có thể được chia thành nhiều hướng khác nhau. Một trong những hướng nghiên cứu phổ biến hiện nay là sử dụng thuật toán phân tích chùm mờ (Lai et al., 2009; Tran, 2016). Phương pháp này hoạt động dựa trên đa dạng các tiêu chí để nhóm các đối tượng có điểm tương đồng vào cùng một nhóm và xác định số mờ thuộc về mỗi nhóm. Từ đặc điểm này, một số nghiên cứu đã áp dụng trong việc phân đoạn ảnh MRI, trong đó có khả năng xử lý nhiễu (Saha et al., 2016). Các nghiên cứu về phân tích chùm được áp dụng trong xử lý ảnh MRI sử dụng chủ yếu là Fuzzy c-means (Vasuda et al., 2010; Le et al., 2011; Monalisa et al., 2018; Zotin et al., 2018; Özyurt et al., 2020), mô hình hỗn hợp (Weisenfeld et al., 2009; Despotović et al., 2015; Akkus et al., 2017). Những nghiên cứu nổi bật có thể liệt kê như Tohka et al. (2004) đã giới thiệu một phương pháp phân đoạn dựa trên Markov Random Field (MRF) cục bộ. Các mô hình MRF đồng nhóm được định nghĩa dựa trên các miền ảnh cục bộ. Sau đó, chúng được sử dụng để xây dựng một mô hình dựa trên MRF toàn cục. Ji et al. (2014) đã đề xuất một thuật toán Fuzzy c-means có ràng buộc không gian mạnh mẽ (RSCFCM) cho phân đoạn ảnh MRI của não, thành công vượt qua nhược điểm của các phương pháp phân tích chùm như FCM và mô hình hàm hợp lý cực đại (EM). Tuy nhiên, tiêu chí đánh giá giữa các pixel trong thuật toán này là khoảng cách Euclide không thể xác định rõ các pixel nhiễu, vì vậy kết quả của phân tích chùm thu được là chưa tốt.

Phân loại là việc gán một phần tử vào các nhóm đã biết một cách thích hợp nhất. Phân loại có thể thực hiện cho dữ liệu ảnh và dữ liệu số. Phân loại cho dữ liệu ảnh phức tạp hơn rất nhiều so với dữ liệu số, vì nó đòi hỏi quá trình trích xuất thông tin đặc trưng rõ ràng. Khi các đặc trưng được trích xuất, chúng cần được biểu diễn dưới dạng các phần tử đại diện để nhận dạng. Trong hầu hết các nghiên cứu

hiện có, phần tử được biểu diễn trong ảnh thông thường là một ma trận số (Pham et al., 2024). Để phân loại hình ảnh theo cách biểu diễn này, chúng ta cần đề xuất một số phép đo khoảng cách phù hợp. Mặc dù có nhiều loại phép đo khoảng cách khác nhau, nhưng cho đến nay vẫn chưa có phương pháp nào được coi là tối ưu cho tất cả các loại dữ liệu (Monalisa et al., 2018). Vấn đề này là một thách thức đáng kể đối với người dùng khi tìm kiếm một độ đo phù hợp khi áp dụng phân loại ảnh trong các vấn đề thực tế. Hơn nữa, đối với dữ liệu ma trận hoặc vector đa chiều được trích xuất từ ảnh, việc phân loại trở nên công kềnh, đặc biệt là đối với các tập dữ liệu lớn (Vasuda et al., 2010). Gần đây, các nhà nghiên cứu đã thực hiện trích xuất đặc trưng của các ảnh trong không gian hai chiều dựa vào bài toán phân tích chùm (Singh et al., 2012; Zotin et al., 2018). Một số hướng nghiên cứu cũng đã chuyển đổi các ảnh trong không gian hai chiều cho mục đích phân loại. Tuy nhiên, các thuật toán này chỉ phù hợp với các trường hợp riêng của tập dữ liệu ảnh nhất định (Singh et al., 2012). Nhận thấy ưu điểm từ thuật toán phân tích chùm mờ trong phân đoạn ảnh, nên thuật toán được sử dụng trong nghiên cứu để tiền xử lý cho bài toán phân loại. Đây được xem là một phương pháp xử lý hiệu quả các nhiễu của ảnh đầu vào nhằm làm giảm lượng thông tin không có đóng góp cho vấn đề trích xuất đặc trưng của ảnh.

Về các phương pháp phân loại, chúng có thể được chia thành hai nhóm chính là dựa vào thống kê và học máy. Theo hướng thống kê, một số kỹ thuật thông dụng bao gồm hồi quy Logistic (Tohka et al., 2004), phương pháp Fisher (Ji et al., 2014), Naive Bayes (áp dụng cho các biến rời rạc) (Özyurt et al., 2020) và phương pháp Bayes cải tiến (áp dụng cho các biến liên tục) (Despotović et al., 2015). Phương pháp Fisher xây dựng một hàm tuyến tính để đưa các biến quan sát đa chiều thành một chiều nhằm phân chia các nhóm một cách hiệu quả, từ đó thiết lập các nguyên tắc phân loại. Tuy nhiên, trong phương pháp này, ma trận hiệp phương sai của các nhóm được yêu cầu phải bằng nhau, do đó, tính hiệu quả của phương pháp này không cao (Akkus et al., 2017; Weisenfeld et al., 2009). Naive Bayes là phương pháp với nguyên tắc phân loại lý thuyết đơn giản, tuy nhiên nó chỉ áp dụng cho các dữ liệu rời rạc và không phân loại được cho nhiều đối tượng khác nhau, đặc biệt là ảnh.

Cùng với sự phát triển của học máy, kỹ thuật học sâu ngày càng trở nên mạnh mẽ. Nhiều hướng nghiên cứu đã được cải tiến trong những năm gần đây. Chẳng hạn, Nawaz et al. (2018) đã đề xuất một phiên bản cải tiến của kiến trúc AlexNet để phân loại

ảnh ung thư vú. Kết quả thực nghiệm cho thấy phương pháp của họ đạt độ chính xác tổng thể là 57% trên tập kiểm tra. Trong cùng hướng phát triển, Hossain et al. (2022) đã đề xuất mô hình ResNet50 để phân loại bệnh nhân COVID-19 từ cơ sở dữ liệu chụp X-quang COVID-19. Họ thu được độ chính xác là 99,17%, trong đó, độ chính xác của quá trình huấn luyện là 99,95%, độ nhạy là 99,03% và chỉ số F1-Score 99,17% trên tập dữ liệu ImageNet_ChestX-ray14 (ChexPert, ChestX-ray14) Kong et al. (2022) thu được 97,3% độ chính xác khi phân loại ảnh X-quang ngực dựa trên sự kết hợp giữa các đặc trưng của kiến trúc mạng DenseNet và kiến trúc VGG16. Những mô hình được liệt kê ở trên đều có những ưu điểm và phù hợp cho từng loại dữ liệu riêng, nhưng chúng sẽ khó đạt được hiệu quả cho tất cả các dữ liệu được xem xét. Một trong những yếu tố chính của AlexNet là chúng có một số lượng lớn tham số cần tính toán. Điều này có thể dẫn đến chi phí tính toán và sử dụng bộ nhớ tăng cao (Lu et al., 2019). Đối với kiến trúc ResNet50, việc huấn luyện các mạng sâu vẫn có thể gặp thách thức và đòi hỏi phải tinh chỉnh cẩn thận các siêu tham số. DenseNet có phần vượt trội hơn so với các kiến trúc AlexNet và ResNet50. Do việc tái sử dụng đặc trưng hiệu quả và dòng Gradient hiệu quả, DenseNet có xu hướng tổng quát hóa tốt trên dữ liệu ẩn, giảm thiểu hiện tượng quá khớp. Tuy nhiên, việc triển khai DenseNet có thể phức tạp hơn, dễ bị lỗi do các kết nối dày đặc và yêu cầu quản lý cẩn thận các mạng đặc trưng.

Mặc dù các thuật toán phân loại dựa trên học sâu đã cho thấy nhiều ưu điểm so với các phương pháp học máy và thống kê, chúng cũng có một số nhược điểm, chẳng hạn như: phụ thuộc vào nguồn dữ liệu huấn luyện lớn, vấn đề nội suy, vấn đề quá khớp, yêu cầu tài nguyên tính toán, thiếu tính giải thích và thách thức về khả năng mở rộng (Ramezani et al., 2023). Tuy nhiên, trong những năm gần đây, nhiều nhà khoa học đặc biệt quan tâm tới kiến trúc mạng học sâu Inception ResNet-V2. Nó là một biến thể của mạng nơ-ron tích chập (CNN) kết hợp hai kiến trúc mạnh mẽ là Inception và ResNet. Được giới thiệu bởi Google, Inception ResNet-V2 là một trong những mô hình tiên tiến nhất trong lĩnh vực học sâu, đặc biệt là trong nhiệm vụ phân loại ảnh. Các khối Inception được xen kẽ bởi các kết nối tắt của ResNet giúp mô hình vừa khai thác được khả năng phân tích đa cấp độ của Inception vừa duy trì độ sâu hiệu quả của ResNet. Ngoài ra, kiến trúc Inception ResNet-V2 đã được huấn luyện cho 1000 tập dữ liệu khác nhau với sự đa dạng các đối tượng. Do đó, việc sử

dụng kiến trúc này trong phân loại ảnh tạo ra nhiều thuận lợi hơn.

Một số nghiên cứu tiêu biểu như Neshat et al. (2024) đã so sánh hiệu suất của mô hình học sâu Inception-ResNet trong việc chẩn đoán viêm phổi từ ảnh chụp X-quang ngực. Kết quả cho thấy Inception ResNet-V2 có độ chính xác phân loại vượt trội so với các mô hình khác, bao gồm: ResNet152V2, MobileNet-V3 (Large và Small), EfficientNetV2 (Large và Small), InceptionV3 và NASNet-Mobile, với mức chênh lệch lần lượt là 2,6%, 6,5%, 7,1%, 13%, 16,1%, 3,9%, và 1,6% cho mỗi mô hình. Manivannan et al. (2023) đề xuất giải pháp phân loại 120 giống chó sử dụng kiến trúc Inception ResNet-V2 đã được huấn luyện trước. Hiệu suất và độ bền của mô hình đã được đánh giá qua việc thử nghiệm với các ảnh từ internet. Mạng dự đoán đúng giống chó với độ chính xác là 88,92%. Các nghiên cứu này có độ chính xác thấp và chưa thật sự tối ưu. Nguyên nhân chủ yếu là dữ liệu ảnh đầu vào chưa được xử lý đúng cách và số lượng đặc trưng không cần thiết có thể được sử dụng trong quá trình huấn luyện. Do đó, các kết quả thu được là không tối ưu.

Dựa trên các phân tích, một thuật toán cải tiến được đề xuất cho bài toán phân loại dựa trên sự khắc phục một số nhược điểm của các thuật toán tồn tại với những đóng góp như sau:

- (i) Đề nghị thuật toán phân đoạn ảnh dựa trên kỹ thuật phân tích chùm mờ trước khi trích xuất các đặc trưng của ảnh, thay vì sử dụng toàn bộ dữ liệu ảnh để trích xuất như nhiều nghiên cứu khác.
- (ii) Tăng cường các kết nối ảnh thông qua phương pháp phân tích chùm mờ với hàm mục tiêu tối ưu.
- (iii) Xây dựng kiến trúc Inception ResNet-V2 cho dữ liệu thu được từ (ii) để phân loại ảnh.

Với những đóng góp mới này, một thuật toán phân loại hiệu quả cho ảnh được thiết lập. Thuật toán đã được giới thiệu thông qua ứng dụng thực tế và thực thi bằng chương trình Matlab. Hiệu quả của thuật toán đề xuất đã được kiểm tra trên các tập dữ liệu ảnh ung thư vú (400x). Ngoài ra, các chỉ số đánh giá đã được sử dụng để so sánh hiệu suất của mô hình này so với nhiều mô hình khác nhau, bao gồm các mô hình từ thống kê, học máy và học sâu.

2. TIÊU CHUẨN ĐÁNH GIÁ MÔ HÌNH

Định nghĩa 1: Độ chính xác (ACC) là một chỉ số phổ biến được sử dụng để đánh giá các thuật toán phân loại. Nó được định nghĩa như sau:

$$ACC = \frac{\text{Số phân loại đúng}}{\text{Tổng số phần tử phân loại}} .$$

Đối với bài toán phân loại cho hai lớp, chỉ số ACC được tính như sau:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} ,$$

trong đó $TP =$ Dương đúng, $TN =$ Âm đúng, $FP =$ Dương sai, và $FN =$ Âm sai.

Chỉ số ACC càng lớn gần tới 1 thì độ chính xác và độ nhạy của thuật toán sẽ càng tốt, ngược lại nếu chỉ số ACC càng gần 0 thì hiệu suất của thuật toán sẽ càng kém.

Định nghĩa 2: F1-Score là một thước đo độ chính xác của mô hình trên một tập dữ liệu. Nó được sử dụng để đánh giá hiệu suất phân loại nhị phân, trong đó các trường hợp được phân loại là “Dương” hoặc “Âm”. F1-Score kết hợp độ chính xác (*Precision*) và độ nhạy (*Recall*) bằng cách sử dụng trung bình điều hòa của chúng. Tối ưu hóa F1-Score đồng nghĩa với việc tối ưu hóa đồng thời cả độ chính xác và độ nhạy. F1-Score được tính như sau:

$$F1 - \text{Score} = \frac{2(\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} ,$$

trong đó

$$\text{Precision} = \frac{TP}{TP + FP} , \text{Recall} = \frac{TP}{TP + FN} .$$

Trong mô hình phân loại hai lớp, chỉ số F1-Score càng gần tới 1 thì độ chính xác và độ nhạy của mô hình càng cao, ngược lại, chỉ số F1-Score càng gần tới 0 thì hiệu suất của mô hình sẽ càng kém.

3. MÔ HÌNH PHÂN LOẠI ĐỀ NGHỊ

3.1. Thuật toán phân đoạn ảnh MRI

Cho $\{y_1, y_2, \dots, y_N\}$ là tập hợp N ảnh. Mô hình đề nghị được trình bày theo các bước sau:

Bước 1. Chuẩn hóa hình ảnh thuộc khoảng $[0, 1]$ cho bởi.

$$P_{N_x, N_y}(\hat{y}_j) = \frac{p(y_j) - \min(p(y_j))}{\max(p(y_j)) - \min(p(y_j))} ,$$

trong đó:

• $P_{N_x, N_y}(\hat{y}_j)$ là giá trị pixel của ảnh đã được chuẩn hoá. Đây là giá trị của pixel tại tọa độ (N_x, N_y) trên một ảnh. Quá trình chuẩn hóa này nhằm mục đích đưa các giá trị pixel (từ 0 đến 255) về một phạm vi cụ thể nhỏ hơn (từ 0 đến 1) giúp thuận lợi

cho các tính toán, đặc biệt là quá trình phân đoạn các ảnh,

- $p(y_j)$ là giá trị pixel của ảnh gốc,
- $\min(p(y_j))$ là giá trị pixel tối thiểu trong ảnh,
- $\max(p(y_j))$ là giá trị pixel tối đa trong ảnh.

Bước 2. Chuyển đổi giá trị pixel của ảnh đã chuẩn hóa thành các vector cột $\hat{I} = [\hat{I}_1, \hat{I}_2, \dots, \hat{I}_N]$ với

$$\hat{I}_h = [P_{1,i}(\hat{y}_j), P_{2,i}(\hat{y}_j), \dots, P_{N_x,i}(\hat{y}_j), \dots, P_{N_x, N_y}(\hat{y}_j)] ,$$

trong đó N_x và N_y là kích thước $N_x \times N_y$ của một ảnh ($1 \leq h \leq N$).

Bước 3. Khởi tạo ma trận phân vùng ngẫu nhiên $U^{(0)}$ với k hàng và N cột như sau:

$$U^{(0)} = [\mu_{ij}^{(0)}]_{k \times N} ,$$

trong đó $\mu_{ij}^{(0)} \in [0, 1]$, với $1 \leq i \leq k, 1 \leq j \leq N$.

Bước 4. Tìm phần tử trọng tâm của cụm C_i theo phương trình sau:

$$C_i = \frac{\sum_{j=1}^N [\mu_{ij}^{(0)}]^2 \cdot \hat{I}_j}{\sum_{j=1}^N [\mu_{ij}^{(0)}]^2} , \quad (5)$$

Bước 5. Cập nhật ma trận mờ mới $U^{(t=1)}$, mỗi ma trận được tính theo nguyên tắc sau:

$$\text{Nếu } D_E(\hat{I}_j, C_m) > 0 ,$$

$$\mu_{ij}^{(t=1)} = \frac{1}{\sum_{m=1}^k \left[\frac{D_E(\hat{I}_j, C_i)}{D_E(\hat{I}_j, C_m)} \right]^2} , \quad (6)$$

và ngược lại $\mu_{ij}^{(1)} = 0$, trong đó

$$1 \leq i, m \leq k, 1 \leq j \leq N .$$

Bước 6. Lặp lại Bước 4 và Bước 5 t lần, cho đến khi thỏa điều kiện sau:

$$|J^{(t+1)} - J^{(t)}| < \varepsilon , \quad (7)$$

trong đó:

$$J^{(t)} = \sum_{i=1}^N \sum_{j=1}^k [\mu_{ij}^{(t)}]^2 \cdot D_E^2(\hat{I}_j, C_i) ,$$

$D_E(\hat{I}_j, C_i)$ là khoảng cách Euclide của \hat{I}_j và C_i .

Bước 7. Lấy chỉ số cụm cho từng pixel

$$H_j = \arg \max_{1 \leq j \leq k} U^{(t+1)} = \arg \max_{1 \leq j \leq k} [\mu_{ij}^{(t+1)}],$$

Ở bước này, thuật toán lấy phân vùng bằng cách phân tích các cụm và ảnh gốc để tạo ra ảnh được phân đoạn.

3.2. Kiến trúc mạng Inception ResNet-V2

- **Cấu trúc Stem:** Đây là bước quan trọng cho mạng Inception ResNet V2. Nó giúp giảm kích thước dữ liệu đầu vào thông qua các lớp tích chập (Conv3x3). Cấu trúc này gồm các lớp tích chập liên tiếp giúp phát hiện các đặc trưng cơ bản như cạnh, góc và các hình dạng đơn giản trong hình ảnh. Sau khi đi qua cấu trúc Stem, dữ liệu đã được trích xuất các đặc trưng cơ bản và giảm sâu kích thước đầu ra. Kết quả đầu ra của khối này là dữ liệu ảnh có kích thước 35x35x256 (Đầu vào là ảnh có kích thước 299x299x3).

- **Cấu trúc Inception ResNet A:** Khối này kết hợp giữa ý tưởng của mô-đun Inception truyền thống với các kết nối Residual từ kiến trúc ResNet để tăng cường hiệu suất và khả năng hội tụ của mạng. Inception-ResNet A tận dụng khả năng của mô-đun Inception trong việc học đa dạng các đặc trưng thông qua các nhánh khác nhau. Nhờ đó mạng tránh được vấn đề mất Gradient khi đi qua nhiều lớp, giúp mạng có thể học được sâu hơn. Kết quả đầu ra của khối này là dữ liệu ảnh có kích thước 5x35x256.

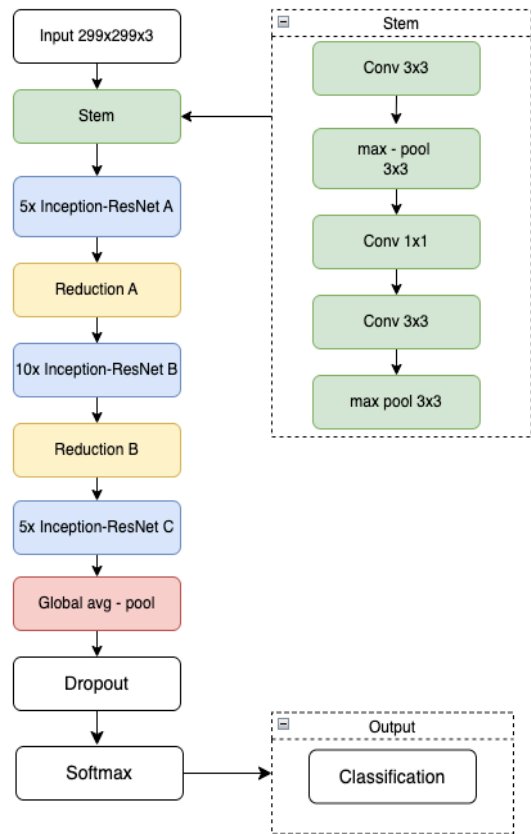
- **Cấu trúc Reduction A:** Khối Reduction A giảm kích thước không gian của Feature map một cách hiệu quả, thường là giảm một nửa (về chiều cao và chiều rộng), từ đó giúp giảm thiểu khối lượng tính toán trong các lớp sau. Nhờ sự kết hợp của nhiều nhánh với các phép biến đổi khác nhau, khối này không chỉ giảm kích thước không gian mà còn bảo toàn được nhiều thông tin quan trọng của đặc trưng. Kết quả đầu ra của khối này là dữ liệu ảnh có kích thước 17x17x1024

- **Cấu trúc Inception ResNet B:** Khối này cho phép mô hình học các đặc trưng từ các thang đo và góc nhìn khác nhau nhờ sử dụng các phép tích chập không đối xứng (asymmetric convolutions) và các nhánh khác nhau. Kết quả đầu ra của khối này là dữ liệu ảnh có kích thước 17x17x1024.

- **Cấu trúc Reduction B:** Khối này được thiết kế để giảm mạnh kích thước không gian của các đặc trưng (feature maps) trong khi tiếp tục trích xuất các đặc trưng phức tạp từ dữ liệu đầu vào. Kết quả đầu

ra của khối này là dữ liệu ảnh có kích thước 28x8x1536.

- **Cấu trúc Inception ResNet C:** Khối này tiếp tục phát triển ý tưởng kết hợp giữa mô-đun Inception với các kết nối Residual từ ResNet, giúp mô hình học các đặc trưng phức tạp một cách hiệu quả và ổn định. Inception-ResNet C được thiết kế để trích xuất các đặc trưng không gian ở các giai đoạn cuối của mạng, nơi kích thước không gian của feature map đã được giảm đáng kể. Kết quả đầu ra của khối này là dữ liệu ảnh có kích thước 8x8x1536.



Hình 1. Sơ đồ minh họa cho kiến trúc Inception ResNet-V2

- **Cấu trúc khối Global average Pooling:** Khối này được sử dụng để giảm kích thước không gian của các đặc trưng (Feature maps) xuống thành một giá trị duy nhất cho mỗi Channel, giúp chuẩn bị dữ liệu cho lớp đầu ra của mạng, thường là lớp fully connected hoặc softmax trong các tác vụ phân loại. Kết quả đầu ra của khối này là dữ liệu ảnh có kích thước 1x1536.

- **Cấu trúc Khối Softmax:** Khối này là một hàm kích hoạt được sử dụng phổ biến trong các mô

hình học máy, đặc biệt trong các bài toán phân loại đa lớp. Hàm Softmax chuyển đổi đầu ra của một mô hình thành xác suất, sao cho tổng xác suất của tất cả các lớp bằng 1. Thông thường hàm softmax được sử dụng là:

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_{i=1}^n e^{x_i}}, \quad (8)$$

trong đó x_i là giá trị đầu vào của vector (x_1, x_2, \dots, x_n) , n là tổng số lớp.



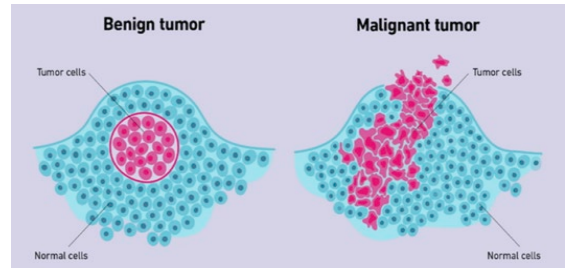
Hình 2. Sơ đồ minh họa các giai đoạn của mô hình đề nghị

Nhận xét: Trong Bước 6, giá trị của ϵ là một số rất nhỏ được chọn một cách tùy ý. ϵ càng nhỏ, số lần lặp và thời gian tính toán càng nhiều. Trong các ví dụ số, chúng tôi chọn $\epsilon = 10^{-4}$. Tất cả chương trình được thực thi bằng phần mềm Matlab 2024 (Online).

4. ỨNG DỤNG CHO DỮ LIỆU UNG THƯ

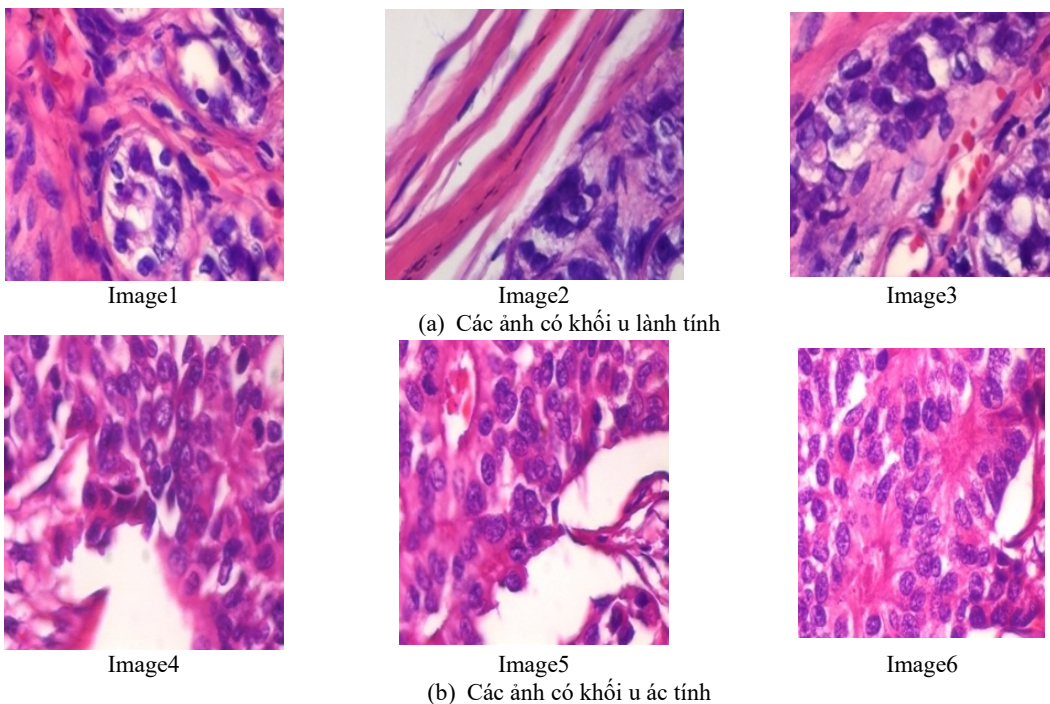
Tập dữ liệu ảnh MRI của hai nhóm (ảnh có khối u lành và ác tính) được sử dụng để minh họa các giai

đoạn của thuật toán đề nghị. Giới thiệu về sự khác biệt giữa khối u lành tính và ác tính ở cấp độ tế bào được cho trong Hình 3.



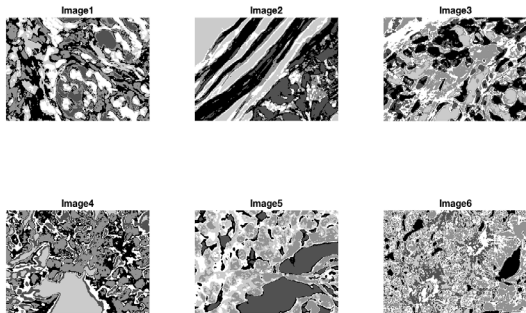
Hình 3. Sự khác biệt giữa khối u lành tính và khối u ác tính. Khối u lành tính sẽ giới hạn bởi vòng tròn màu hồng, trong khi ảnh có khối u ác tính sẽ có những phần xâm lấn không theo một cấu trúc nào

Cơ sở dữ liệu BreakHis chứa các ảnh sinh thiết vi mô của các khối u lành tính và ác tính của bệnh nhân ung thư vú. Bộ dữ liệu xem xét là bộ dữ liệu có độ phóng đại quang học 400x. Tập dữ liệu được chia thành hai phần, trong đó 80% dữ liệu được sử dụng để huấn luyện với 1148 ảnh (371 ảnh có khối u lành tính, 777 ảnh có khối u ác tính) và 20% dữ liệu ảnh làm tập kiểm tra với 545 ảnh (176 ảnh có khối u lành tính, 369 ảnh có khối u ác tính). Một số mẫu ảnh đại diện cho tập kiểm tra được cho trong Hình 4.



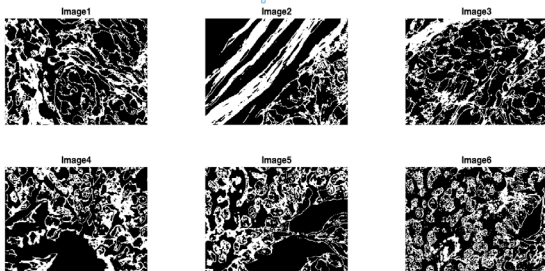
Hình 4. Một số mẫu ảnh đầu vào của dữ liệu BreakHis

Trước tiên, giai đoạn 1 của mô hình đề nghị được thực hiện là phân đoạn các ảnh của tập dữ liệu. Kết quả của giai đoạn này được giới thiệu trong Hình 4 với một số mẫu ảnh trong tập.



Hình 5. Một số mẫu ảnh được phân đoạn trong tập dữ liệu kiểm tra

Kết quả được thể hiện ở Hình 5 là quá trình phân đoạn các ảnh trong vòng lặp đầu. Chúng ta có thể thấy rằng các cấu trúc bên trong ảnh có sự phân chia rõ ràng hơn.



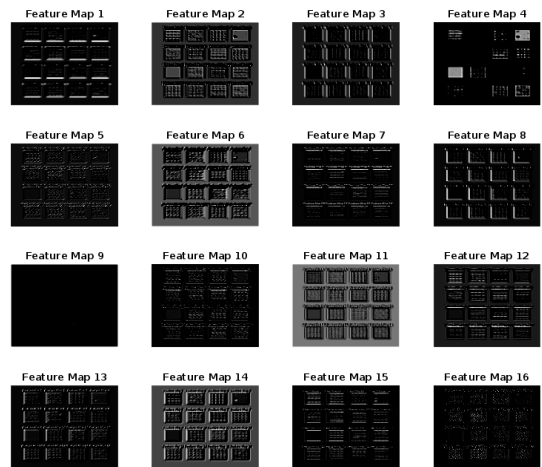
Hình 6. Kết quả phân đoạn các ảnh ở lần lặp cuối

Hình 6 thể hiện rằng các ảnh giữa nhóm có khối u lành tính và ác tính được phân vùng rõ ràng hơn. Các tế bào trong Image1, Image2 và Image3 có sự tập trung theo dạng tròn và rất dễ nhìn thấy bằng mắt thường, trong khi tế bào của các ảnh còn lại có sự phân chia không thống nhất về màu sắc. Điều này cho thấy chúng không phân bố theo một trật tự nào, có nghĩa rằng chúng là các tế bào có khả năng xâm lấn, mở rộng.

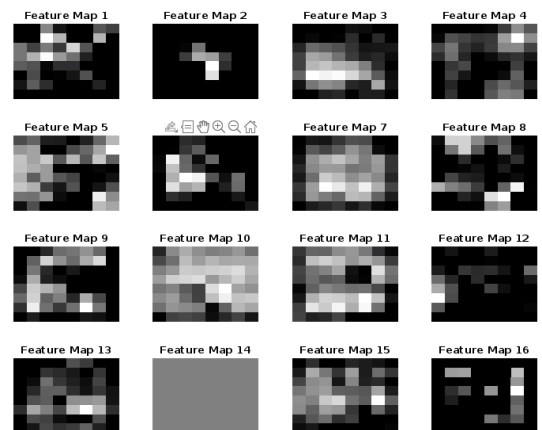
Sau quá trình phân đoạn ảnh bằng thuật toán phân tích chùm mờ, các ảnh này được phân loại bằng mạng Inception ResNet-V2. Kết quả thu được từ tập dữ liệu ảnh sau Block Stem được cho trong Hình 7.

Hình 7 giới thiệu 16 thông tin đặc trưng của 16 ma trận trong 256 ma trận cấp 35×35 . Như vậy, mỗi Features Map sẽ chứa ma trận 35×35 . Sau khối Stem, thuật toán thu được 256 ma trận đặc trưng cấp 35×35 cho mỗi ảnh. Tương tự, thuật toán thực hiện

cho đến Khối Inception ResNet C, thuật toán sẽ thu được 1536 ma trận cấp 8×8 .



Hình 7. Kết quả trích xuất các đặc trưng của ảnh Image1 sau Khối Stem



Hình 8. Kết quả của 16 ma trận đặc trưng cấp 8×8 của ảnh Image1 sau khối Inception ResNet C

Sau Khối Inception ResNet C, thì khối average Pooling và Fully Connected được áp dụng. Kết quả thu được sau khối Fully Connected là vector đặc trưng gồm 1000 giá trị. Đây là vector đặc trưng sau cùng của mạng Inception ResNet-V2 trước khi đưa vào lớp Softmax để tính toán giá trị xác suất theo công thức (8). Sau đó, khối Classification Layer được áp dụng để thực hiện quá trình phân loại các đặc trưng thu được từ tập kiểm tra. Trong nghiên cứu này, hai chỉ số được sử dụng là ACC và F1-Score để đánh giá kết quả phân loại đầu ra của thuật toán đề nghị. Kết quả của nó được cho trong Bảng 2.

Bảng 2. Kết quả tính toán các chỉ số đánh giá của mô hình phân loại

| Mô hình | ACC (%) | F1-Score | Hạng |
|---------------------------------|-------------|-------------|----------|
| FCIR-V2 | 99,6 | 99,3 | 1 |
| Support vector machine (SVM) | 97,7 | 99,2 | 2 |
| Quadratic Linear Analysis (QLA) | 97,7 | 99,2 | 2 |
| DenseNet | 95,4 | 98,4 | 3 |
| VGG16 | 93,0 | 93,5 | 4 |
| CNN | 93,0 | 93,5 | 5 |
| Resnet50 | 81,8 | 83,9 | 6 |

Kết quả tính toán các chỉ số cho thấy mô hình FCIR-V2 thu được kết quả tốt hơn các mô hình khác. Cụ thể, **FCIR-V2** đạt hiệu suất cao nhất với **ACC = 99,6%** và **F1-Score = 99,3%**, đứng ở vị trí đầu tiên trong Bảng 2, trong khi ba mô hình tốt đạt giá trị ACC lần lượt là 97,7% (SVM và QLA) và 95,4% (DenseNet). **ResNet50** có hiệu suất kém nhất với **ACC là 81,8%** và **F1-Score là 83,9%**, đứng cuối bảng. Điều này cho thấy mô hình này không hoạt động hiệu quả trong tập dữ liệu này.

5. KẾT LUẬN VÀ MỘT SỐ THẢO LUẬN

Nghiên cứu này có những đóng góp đáng kể cho mô hình phân loại ảnh MRI có giám sát thông qua sự kết hợp của thuật toán phân tích chùm mờ và kiến trúc mạng học sâu Inception ResNet-V2.

Đầu tiên, thuật toán phân tích chùm mờ cải tiến được giới thiệu như một ý tưởng mới cho bài toán phân đoạn các ảnh đầu vào của mạng học sâu. Thuật

toán này tập trung vào việc làm nổi bật sự khác biệt giữa các nhóm trong mỗi ảnh. Sau đó, sử dụng kiến trúc học sâu Inception ResNet-V2 để huấn luyện cho các ảnh đã được phân đoạn và phân loại các ảnh của tập dữ liệu kiểm tra. Hiệu suất của FCIR-V2 đã được chứng minh thông qua thí nghiệm kiểm chứng với 99,6% độ chính xác (ACC) và 99,3% chỉ số F1-Score. Kết quả thí nghiệm cho thấy được ưu điểm vượt trội của FCIR-V2 so với các thuật toán khác như CNN, DenseNet, SVM, QLA, ResNet50 and VGG16.

Thí nghiệm của FCIR-V2 được áp dụng trong phần mềm Matlab 2024 (Online). Chương trình thực hiện cho các thí nghiệm do chúng tôi xây dựng, riêng kiến trúc mạng học sâu Inception ResNet-v2 là thư viện có sẵn của Matlab. Mặc dù có sự cải tiến đáng kể về kết quả phân loại nhưng mô hình này cũng có một số hạn chế nhất định bởi quá trình phân đoạn các ảnh mất khá nhiều thời gian tính toán và độ phức tạp của nó cũng ở mức cao.

TÀI LIỆU THAM KHẢO (REFERENCES)

Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D. L., & Erickson, B. J. (2017). Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. *Journal of Digital Imaging*, 30, 449-459. <https://doi.org/10.1007/s10278-017-9983-4>.

Despotović, I., Goossens, B., & Philips, W. (2015). MRI segmentation of the human brain: Challenges, methods, and applications. *Computational and Mathematical Methods in Medicine*, (1), 450341. <https://doi.org/10.1155/2015/450341>

Hossain, M. B., Iqbal, S. H. S., Islam, M. M., Akhtar, M. N., & Sarker, I. H. (2022). Transfer learning with fine-tuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images. *Informatics in Medicine Unlocked*, 3(0), 100916. <https://doi.org/10.1016/j.imu.2022.100916>

Ji, Z., Liu, J., Cao, G., Sun, Q., & Chen, Q. (2014). Robust spatially constrained fuzzy c-means algorithm for brain MR image segmentation. *Pattern recognition*, 47(7), 2454-2466. <https://doi.org/10.1016/j.patcog.2014.01.017>

Kong, L., & Cheng, J. (2022). Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion. *Biomedical Signal Processing and Control*, 77, 103772. <https://doi.org/10.1016/j.bspc.2022.103772>

Lai, C. C., & Chang, C. Y. (2009). A hierarchical evolutionary algorithm for automatic medical image segmentation. *Expert Systems with Applications*, 36(1), 248-259. <https://doi.org/10.1016/j.eswa.2007.09.003>

Le, C. H., and Frélicot, C. (2011). A cluster-validity index combining an overlap measure and a separation measure based on fuzzy-aggregation operators. *IEEE Trans. Fuzzy Syst*, 19(3), 580-588. <https://doi.org/10.1109/TFUZZ.2011.2106216>

Lu, S., Lu, Z., & Zhang, Y. D. (2019). Pathological brain detection based on AlexNet and transfer learning. *Journal of Computational Science*, 30, 41-47. <https://doi.org/10.1016/j.jocs.2018.11.008>

- Manivannan, S., & Venkateswaran, N. (2023). Dog Breed Classification using Inception ResNet-V2. In *2023 International Conference for Advancement in Technology*, 1-5. <https://doi.org/10.1109/ICONAT57137.2023.10080065>
- Monalisa, A., Swathi, D., Karuna, Y., & Saladi, S. (2018). Robust Intuitionistic Fuzzy c-Means Clustering Algorithm for Brain Image Segmentation. In *Proceedings of the 2018 IEEE International Conference on Communication and Signal Processing*, 0781-0785. <https://doi.org/10.1109/ICCSP.2018.8524360>
- Nawaz, W., Ahmed, S., Tahir, A., & Khan, H. A. (2018). Classification of breast cancer histology images using alexnet. In *Image Analysis and Recognition: 15th International Conference, ICLAR 2018, Póvoa de Varzim, Portugal, June 27–29, 2018, Proceedings 15* (pp. 869-876). Springer International Publishing. https://doi.org/10.1007/978-3-319-93000-8_99
- Neshat, M., Ahmed, M., Askari, H., Thilakarathne, M., & Mirjalili, S. (2024). Hybrid Inception Architecture with Residual Connection: Fine-tuned Inception-ResNet Deep Learning Model for Lung Inflammation Diagnosis from Chest Radiographs. *Procedia Computer Science*, 235, 1841-1850. <https://doi.org/10.1016/j.procs.2024.04.175>
- Özyurt, F., Sert, E., & Avcı, D. (2020). An expert system for brain tumor detection: Fuzzy C-means with super resolution and convolutional neural network with extreme learning machine. *Med. Hypotheses*, 134, 109433. <https://doi.org/10.1016/j.mehy.2019.109433>
- Pham, T. B., Vo, T. H. N., Vu, Q. T., & Pham, T. D. (2024). Improving fuzzy clustering algorithm for overlapping elements and its application. *International Journal of Information Technology*, 16(4), 2595-2602. <https://doi.org/10.1007/s41870-024-01745-w>
- Ramezani, S. B., Cummins, L., Killen, B., Carley, R., Amirlatifi, A., Rahimi, S., Seale, M., & Bian, L. (2023). Scalability, Explainability and Performance of Data-Driven Algorithms in Predicting the Remaining Useful Life: A Comprehensive Review. *IEEE Access*, 11, 41741-41769. <https://doi.org/10.1109/ACCESS.2023.3267960>
- Saha, S., Alok, A. K., & Ekbal, A. (2016). Brain image segmentation using semi-supervised clustering. *Expert Systems with Applications*, 52, 50-63. <https://doi.org/10.1016/j.eswa.2016.01.005>
- Singh, D., & Kaur, K. (2012). Classification of Abnormalities in Brain MRI Images Using GLCM, PCA and SVM. *Int. J. Eng. Adv. Technol*, 1(6), 243-248.
- Tohka, J., Zijdenbos, A., & Evans, A., (2004). Fast and robust parameter estimation for statistical partial volume models in brain MRI. *Neuroimage*, 1(6), 243-248. <https://doi.org/10.1016/j.neuroimage.2004.05.007>
- Tran, M. T. (2016). A cooperative semi-supervised fuzzy clustering framework for dental X-ray image segmentation. *Expert Systems with Applications*, 46, 380-393. <https://doi.org/10.1016/j.eswa.2015.11.001>
- Vasuda, P., & Satheesh, S. (2010). Improved Fuzzy C-Means Algorithm for MR Brain Image Segmentation. *Int. J. Comput. Sci. Eng*, 2(5), 1713.
- Weisenfeld, N. I., & Warfield, S. K. (2009). Automatic segmentation of newborn brain MRI. *Neuroimage*, 47(2), 564-572. <https://doi.org/10.1016/j.neuroimage.2009.04.068>
- Zotin, A., Simonov, K., Kurako, M., Hamad, Y., & Kirillova, S. (2018). Edge detection in MRI brain tumor images based on fuzzy C-means clustering. In *Procedia Computer Science*, 126, 1261-1270. <https://doi.org/10.1016/j.procs.2018.08.069>