ỨNG DỤNG PHƯƠNG PHÁP HỌC ĐA NHÂN TRONG PHÂN LỚP ẢNH MRI BỆNH U NÃO

Từ Ngọc Thảo[1], Nguyễn Đức Thắng[1], Phan Thị Đài Trang[1]

Ngày nhận bài: 18/10/2024; Ngày phản biện thông qua: 27/11/2024; Ngày duyệt đăng: 28/11/2024

TÓM TẮT

Trong vài thập kỷ gần đây, phân tích hình ảnh y tế đã trở thành một lĩnh vực quan trọng nhằm đưa ra các dự đoán định hướng cho nhân viên y tế, đặc biệt có ích cho các công việc chẩn đoán thường xuyên, liên tục với số lượng lớn. Các hình ảnh y tế cùng loại tại cùng một bộ phận có rất nhiều điểm tương đồng, hoàn toàn có thể sử dụng máy vi tính cùng các thuật toán thông minh để phát hiện ra các điểm bất thường trên hình ảnh. Bài báo này trình bày kết quả của một mô hình máy học sử dụng phương pháp học đa nhân (Multiple Kernel Learning) thông qua việc huấn luyện và kiểm thử với bộ ảnh MRI về u não của dự án MICCAI BRATS 2012. Kết quả, sử dụng phương pháp đánh giá theo kiểu thẩm tra chéo (Cross-Validation) cho kết quả dự đoán là F1-score = 82%. Đây là một kết quả khả quan đối với một mô hình phân lớp nhị phân.

***Từ khóa**: Máy học, trí tuệ nhân tạo, phương pháp học đa nhân, phân lớp ảnh, ảnh MRI u não.*

1. MỞ ĐẦU

Các đặc điểm giải phẫu của cơ thể còn được thể hiện dưới dạng hình ảnh hoặc mô hình kết cấu thể hiện tính toàn vẹn của cơ thể và sự hoạt động của từng cơ quan. Những sai lệch, những bất thường của các chức năng sinh lý thường làm thay đổi các đặc điểm giải phẫu tại các cơ quan tương ứng. Và thông qua các phương pháp chụp chiếu, lưu giữ hình ảnh y tế, các bác sĩ có thể thấy được các thay đổi đó để có được những đầu mối quan trọng về tình trạng sức khỏe mà không cần phải mổ hoặc sử dụng bất kỳ phương pháp xâm nhập nào khác. Hình ảnh y tế (X-quang kỹ thuật số, MRI, CT…) đã và đang là những kỹ thuật, công cụ quan trọng để hướng dẫn chẩn đoán các cơ quan bị biến dạng, rối loạn chức năng, chẳng hạn như phát hiện các vết gãy xương, vôi hóa, khối u não, bất thường nhịp tim và tăng trưởng ác tính trong các cơ quan quan trọng khác.

Việc phân tích ảnh y tế bằng các phương pháp máy học có thể đưa ra dự đoán về tình trạng bệnh tật dựa trên những hình ảnh y tế, với hy vọng làm giảm được khối lượng công việc của bác sĩ, giảm giá thành của các dịch vụ y tế cũng như tránh được những sai sót chủ quan không đáng có trong quá trình khám chữa bệnh. Có thể thấy qua các ứng dụng như hỗ trợ chẩn đoán, tầm soát bệnh qua hình ảnh y tế bằng nền tảng Trí tuệ nhân tạo Dr.Aid của công ty VinBrain.

Một số mô hình phổ biến được sử dụng trong phân tích ảnh y tế có thể kể đến như: Mạng nơ-ron tích chập (Convolutional Neural Networks - CNN), là một trong những mô hình phổ biến nhất trong phân tích ảnh y tế, đặc biệt hiệu quả trong việc phát hiện và phân loại các đối tượng trên ảnh. Với các lớp tích chập và pooling, CNN có khả năng nhận diện các đặc điểm khác nhau trong ảnh y tế, từ đó hỗ trợ trong việc phát hiện bệnh hoặc chẩn đoán. Ví dụ: ResNet, VGG, và Inception thường được dùng trong phân loại ảnh hoặc phát hiện tổn thương trên ảnh MRI và X-ray. Nổi bật trong số các mạng CNN là mạng U-Net, được thiết kế đặc biệt cho các nhiệm vụ phân đoạn ảnh y tế. Mô hình này bao gồm các phần encoder và decoder, giúp mô hình có thể trích xuất các đặc điểm và tái tạo lại hình ảnh với các vùng quan tâm được đánh dấu. U-Net đã được ứng dụng rộng rãi trong phân đoạn các mô hoặc khối u trên ảnh MRI và CT. Tiếp theo có thể kể đến phương pháp Học đa nhân (Multiple Kernel Learning - MKL), là phương pháp máy học cho phép sử dụng nhiều hàm nhân (kernel) khác nhau để cải thiện độ chính xác và độ linh hoạt của mô hình. MKL có thể kết hợp thông tin từ nhiều nguồn và phù hợp với những bài toán phân lớp phức tạp trong ảnh y tế. Thường được ứng dụng trong phân loại và phân đoạn khi dữ liệu đầu vào đến từ nhiều nguồn khác nhau (ví dụ: ảnh MRI, ảnh PET).

2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

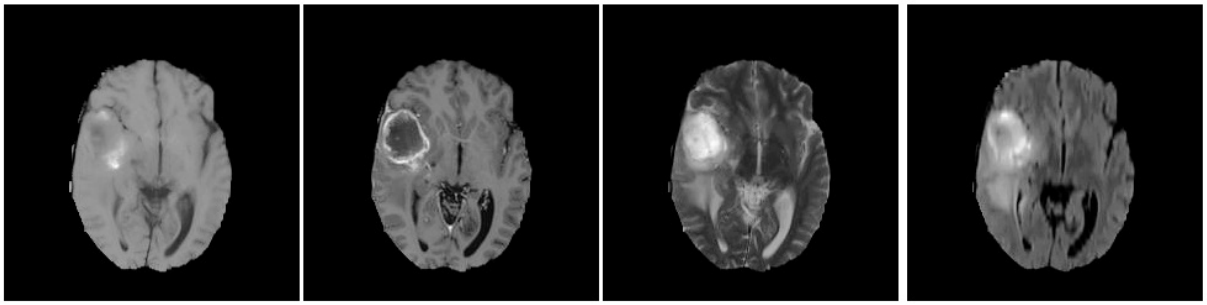
2.1. Vật liệu nghiên cứu

Nghiên cứu dựa trên bộ dữ liệu ảnh MRI về u não của dự án MICCAI BRATS 2012 bao gồm 25 mẫu chụp cộng hưởng từ đã được gán nhãn trước, có 9 mẫu thể hiện u não lành tính và 16 mẫu thể hiện u não ác tính. Trong mỗi mẫu có 4

[1]*Khoa Khoa học Tự nhiên và Công nghệ, Trường Đại học Tây Nguyên;*
Tác giả liên hệ: Từ Ngọc Thảo; ĐT: 0973351560; Email: tungocthao@ttn.edu.vn.

file ảnh MRI có định dạng ".mha" chứa 155 lát cắt chi tiết.

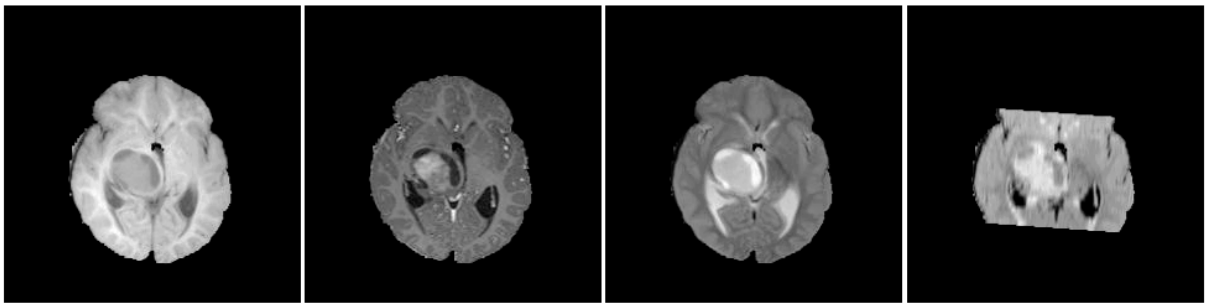
4 file ảnh trong mỗi mẫu gồm ảnh T1, T1ce, T2 và FLAIR. Trong đó, ảnh T1 (T1-weighted Imaging) có màu tối cho dịch não tủy (CSF) và nước, màu sáng đối với mô mỡ và cấu trúc mô mềm với độ tương phản tương đối thấp. Ưu điểm của ảnh T1 là cung cấp chi tiết giải phẫu rõ ràng, tốt cho việc so sánh trước và sau tiêm thuốc cản từ để đánh giá mức độ ngấm thuốc. Ảnh T1 thường được sử dụng để đánh giá cấu trúc giải phẫu của não, xác định biên giới u và tổn thương. Ảnh T1ce (T1-weighted with Contrast Enhancement) là ảnh T1 được chụp sau khi tiêm thuốc cản từ (gadolinium). Thuốc cản từ giúp làm nổi bật các mạch máu và khu vực bất thường trong não, làm tăng độ nhạy trong việc phát hiện các khối u, đặc biệt với u có khả năng ngấm thuốc cản từ. Ảnh T1ce được sử dụng để xác định sự tồn tại của hàng rào máu não bị phá vỡ, biểu hiện qua sự ngấm thuốc của khối u hoặc vùng viêm. Điều này rất quan trọng trong việc đánh giá u não, mức độ lan rộng và tính chất xâm lấn. Ảnh T2 (T2-weighted Imaging) với dịch não tủy và nước có màu sáng, trong khi mô mỡ và các vùng khác thường có màu tối hơn. Cấu trúc u thường xuất hiện sáng hơn mô xung quanh, có thể xác định ranh giới giữa khối u và phù nề giúp dễ dàng phát hiện các vùng mô bị phù nề hoặc viêm nhiễm, rất hữu ích trong việc đánh giá mức độ lan rộng của phù nề quanh khối u. Ảnh FLAIR (Fluid Attenuated Inversion Recovery) là một loại ảnh T2 đã loại bỏ tín hiệu từ dịch não tủy (CSF), giúp tăng cường hiển thị các bất thường ở gần CSF, chẳng hạn như u não hoặc tổn thương viêm. Ảnh FLAIR rất hữu ích để phát hiện các tổn thương nhỏ ở não mà không bị nhiễu bởi dịch não tủy, giúp xác định các bất thường như u hoặc các tổn thương viêm trong chất trắng và xám, làm rõ hơn các tổn thương nhỏ, tổn thương ở rìa của não hoặc bên cạnh não thất mà hình ảnh T2 có thể bỏ qua do bị che mờ bởi dịch não tủy.



Hình 1. 4 loại hình ảnh lát cắt 69 của 1 mẫu u não ác tính

Để xác định khối u não là lành tính hay ác tính, bác sĩ thường không chỉ dựa vào một loại ảnh MRI mà kết hợp nhiều chuỗi ảnh để có cái nhìn đầy đủ và chính xác nhất. Cần kết hợp các thông tin như kết cấu khối u: U lành tính thường có cấu trúc đồng nhất và ranh giới rõ ràng hơn, trong khi u ác tính thường có cấu trúc không đồng nhất, ranh giới mờ hoặc xâm lấn. Sự ngấm thuốc: Bác sĩ sẽ đánh giá hình thái và mức độ ngấm thuốc cản từ của khối u để nhận biết dấu hiệu ác tính. Mức độ phù: Mức độ phù nề và ảnh hưởng mô xung quanh cũng là yếu tố gợi ý bản chất khối u. Tuy nhiên, một số loại ảnh MRI đóng vai trò quan trọng trong việc đánh giá bản chất của khối u như: Ảnh T1ce là chuỗi ảnh rất quan trọng để đánh giá mức độ ngấm thuốc cản từ của khối u. Khối u ác tính thường có xu hướng ngấm thuốc mạnh và không đồng nhất, do đó giúp phân biệt với u lành tính vốn thường có mức độ ngấm thuốc ít hơn hoặc không ngấm thuốc vì khối u ác tính có thể làm phá vỡ hàng rào máu não, cho phép chất cản từ thấm vào khối u, tạo nên hình ảnh rõ nét. Trong khi đó, ảnh T2 và FLAIR cung cấp thông tin về sự hiện diện và mức độ phù nề quanh khối u, tính chất và hình thái của mô xung quanh khối u. Khối u ác tính thường gây ra phù nề lớn hơn và ảnh hưởng rộng rãi đến mô lân cận, trong khi u lành tính có thể có ranh giới rõ ràng và ít gây phù nề hơn. FLAIR đặc biệt hữu ích để xác định mức độ ảnh hưởng của khối u và các tổn thương nhỏ liên quan (M. Martucci et al., 2023).

Tóm lại, ảnh T1ce thường là chuỗi ảnh đặc biệt hữu ích trong việc xác định mức độ ác tính của khối u, nhưng bác sĩ cần kết hợp các chuỗi T2 và FLAIR để có cái nhìn toàn diện và đưa ra chẩn đoán chính xác hơn.



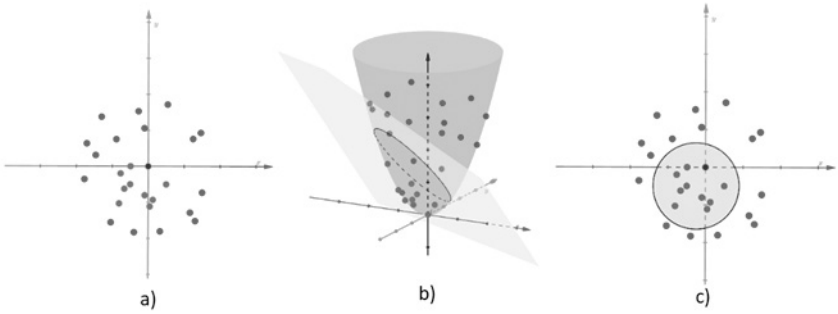
Hình 2. 4 loại hình ảnh lát cắt 65 của 1 mẫu u não lành tính

2.2. Phương pháp nghiên cứu

Bài báo sử dụng hai phương pháp nghiên cứu chính: nghiên cứu lý thuyết và phương pháp thực nghiệm.

2.2.1. Phương pháp nghiên cứu lý thuyết

Phương pháp hàm nhân: Ý tưởng của phương pháp sử dụng hàm nhân là: thay vì thực hiện trên không gian gốc, phương pháp này ánh xạ dữ liệu vào không gian nhiều chiều hơn để dữ liệu bị rời rạc hoàn toàn, sau đó thực hiện các phương pháp phân tích truyền thống trong không gian đích này với hy vọng tìm được các khuôn mẫu thú vị từ việc ánh xạ này. Ánh xạ từ không gian gốc vào không gian đích không cần xác định một cách tường minh mà thông qua một hàm gọi là nhân. Các khuôn mẫu tìm được bằng các phương pháp nhân hàm thường có khuôn dạng phức tạp (phi tuyến, phụ thuộc vào hàm nhân) mặc dù bản chất các khuôn mẫu này là đơn giản (thường là dạng tuyến tính) trong không gian đặc trưng tiềm ẩn.



Hình 3. Ví dụ về hàm nhân

a) Dữ liệu của hai lớp là không phân biệt tuyến tính trong không gian hai chiều.

b) Nếu coi thêm chiều thứ ba là một hàm số của hai chiều còn lại, các điểm dữ liệu sẽ được phân bố trên 1 parabolic và đã trở nên phân biệt tuyến tính. Mặt phẳng màu vàng là mặt phân chia, có thể tìm được bởi Hard/Soft Margin SVM.

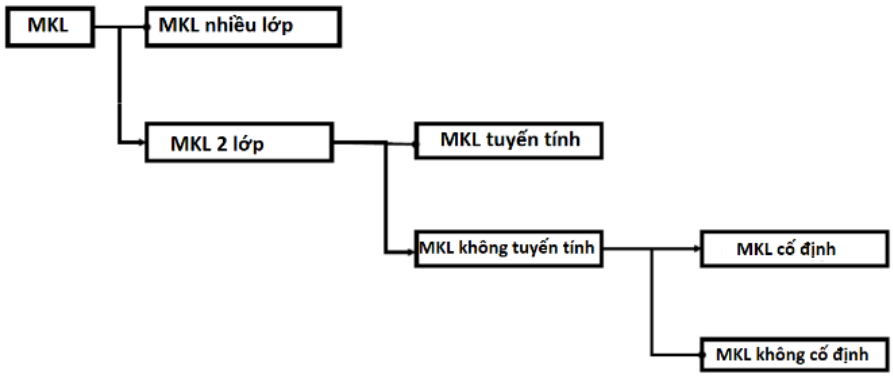
c) Giao điểm của mặt phẳng tìm được và mặt parabolic là một đường ellipse, khi chiếu toàn bộ dữ liệu cũng như đường ellipse này xuống không gian hai chiều ban đầu, ta đã tìm được đường phân chia hai lớp.

Một hàm $k: X \times X \to H$ với phép đặt tương ứng $(x, x') \mapsto k(x, x')$ được gọi là nhân nếu tồn tại một không gian Hilbert H và ánh xạ $\Phi: X \Rightarrow Hsaochok(x, x') = \langle \Phi(x), \Phi(x') \rangle$.

Như vậy nhân là một hàm dùng để biểu diễn mỗi đối tượng $x \in X$ thành một vector $\Phi(x) \in H$ và tính các tích vô hướng. Hay nói cách khác nhân là một hàm dùng để tính nội tích các ảnh $\langle \Phi(x), \Phi(x') \rangle$ trong không gian đặc trưng H mà không cần biết chính xác về hàm ánh xạ Φ.

Các phương pháp kernel được đánh giá là công cụ mạnh mẽ trong phân tích và phân loại hình ảnh y tế. Những phương pháp này có khả năng xử lý các dạng dữ liệu phi tuyến và đa chiều, phù hợp với độ phức tạp của ảnh y tế (Charpiat et al., 2015).

Phương pháp học đa nhân: Trong thực tế, có nhiều hơn một kiểu dữ kiệu được thu thập cho một bài toán. Hơn nữa, dữ liệu được trình bày ở các dạng khác nhau với các độ đo khác nhau. Do đó, giải pháp được đề xuất là sử dụng cùng lúc nhiều nhân (hoặc nhiều tham số của nhân) để xử lý nhiều kiểu dữ liệu khác nhau nhằm tăng tính chính xác của thuật toán.



Hình 4. Minh họa các dạng của MKL

Dù có nhiều loại MKL, có thể nói phương pháp học đa nhân là một dạng của phương pháp hàm nhân.

Một hạt nhân mới có thể được định nghĩa bằng cách sử dụng k1 và k2 với các giá trị tương ứng của chúng Φ_1 (x) và Φ_2 (x). Ta có: $K = k_1 + k_2$ sẽ được biểu diễn bằng $\Phi x = [\Phi_1(x), \Phi_2(x)]$.

Với một tập hợp nhân $\boldsymbol{K} = \{\boldsymbol{k_1}, \boldsymbol{k_2}, \ldots, \boldsymbol{k_m}\}$, có thể được biểu diễn như sau:

$$K = \sum_{i=1}^{m} \mu_i k_i$$

Tùy thuộc vào ràng buộc μ_i (trọng số) là dương, tức là $\mu_i \geq 0$, i = 1 .. .m. Một bộ phân loại thống kê kernel-based (như SVM) tạo ra một đường biên trong không gian đặc trưng, tách hai lớp bằng cách sử dụng thuật toán phân biệt tuyến tính. Để tìm ra được thuật toán phân biệt tuyến tính này, một vấn đề tối ưu hóa cần phải được giải quyết, được biết đến như là một bài toán quy hoạch toàn phương, có thể giải được dễ dàng. Trên cơ sở đường biên này, các thuật toán SVM được phân loại thành hard-margin, 1-norm soft-margin và và 2-norm soft-margin. Nghiên cứu của Sonnenburg và cộng sự (2006, p. 1531-1565) đã mở rộng khả năng áp dụng của MKL, làm cho nó trở thành một phương pháp có giá trị trong phân tích dữ liệu lớn và phức tạp trong nhiều lĩnh vực, bao gồm học máy và sinh học tính toán.

Mô hình phân lớp Support Vector Machines (SVMs): Là một kỹ thuật phân lớp mạnh mẽ và phổ biến trong học máy, mô hình này tối ưu hóa một siêu phẳng để phân chia dữ liệu trong không gian đặc trưng, với mục tiêu tối đa hóa khoảng cách giữa các lớp, được áp dụng rộng rãi trong các ứng dụng như phân loại văn bản, hình ảnh và chẩn đoán y tế. Nghiên cứu đầu tiên giới thiệu về SVMs của Cortes và cộng sự (1995, p.273-297) trình bày nền tảng lý thuyết và ứng dụng của chúng.

Mô hình phân lớp K-Nearest Neighbors (KNNs): Là một mô hình đơn giản nhưng hiệu quả, hoạt động dựa trên việc "bỏ phiếu" từ các điểm dữ liệu gần nhất trong không gian đặc trưng. KNN không yêu cầu quá trình huấn luyện tốn kém nhưng nhược điểm là thời gian dự đoán có thể lớn với dữ liệu lớn. Một trong những tham chiếu quan trọng về KNNs cho đến nay là bài báo giới thiệu thuật toán KNNs của Cover và cộng sự (1967, p.21-27).

Mô hình phân lớp Naïve Bayes: Thuật toán Naïve Bayes dựa trên lý thuyết xác suất Bayes với giả định đơn giản hóa rằng các đặc trưng là độc lập. Mô hình này đặc biệt hiệu quả trong phân loại văn bản như lọc thư rác, và rất nhanh với bộ dữ liệu lớn. McCallum và cộng sự (1998, p.41-48) đưa ra so sánh các mô hình ứng dụng Naïve Bayes trong phân loại văn bản với hai mô hình chính là Multinomial Naïve Bayes và Bernoulli Naïve Bayes, kết quả cho thấy mô hình Multinomial Naïve Bayes thường cho hiệu suất cao hơn đáng kể so với Bernoulli trong các bài toán phân loại văn bản như lọc thư rác hay phân loại chủ đề văn bản. Lý do chính là vì Multinomial xử lý tần suất xuất hiện của các từ (tokens), trong khi Bernoulli chỉ ghi nhận sự xuất hiện hoặc không xuất hiện của một từ trong tài liệu. Kết quả nghiên cứu này đã cung cấp các hướng dẫn rõ ràng cho việc áp dụng Naïve Bayes vào phân loại văn bản hiệu quả hơn, đặc biệt khi xử lý các bộ dữ liệu lớn và phức tạp.

Một số phương pháp trích chọn đặc trưng ảnh: Có rất nhiều phương pháp trích chọn đặc trưng ảnh phù hợp với nhiều loại bài toán khác nhau và hầu hết đều đã được xây dựng thành các bộ thư viện hoặc công cụ chia sẻ miễn phí trên internet. Do đó, trong phạm vi bài báo này, chúng tôi chỉ trình bày khái quát 3 phương pháp sẽ sử dụng để xây dựng mô hình, nhằm nắm được ý tưởng, giá trị đầu vào, đầu ra và tham số cần thiết để sử dụng.

Phương pháp SIFT (Scale-Invariant Feature Transform): Là phương pháp trích chọn đặc trưng từ các điểm hấp dẫn cục bộ (Local Interest Point hay còn gọi là Keypoint), các điểm này có đặc trưng bất biến khi thay đổi tỷ lệ chiều dài và chiều rộng của ảnh, thay đổi góc nhìn hay xoay ảnh và cường độ

chiếu sáng của ảnh. Do đó, độ lớn của đặc trưng cục bộ bất biến phụ thuộc vào số lượng của các keypoint.

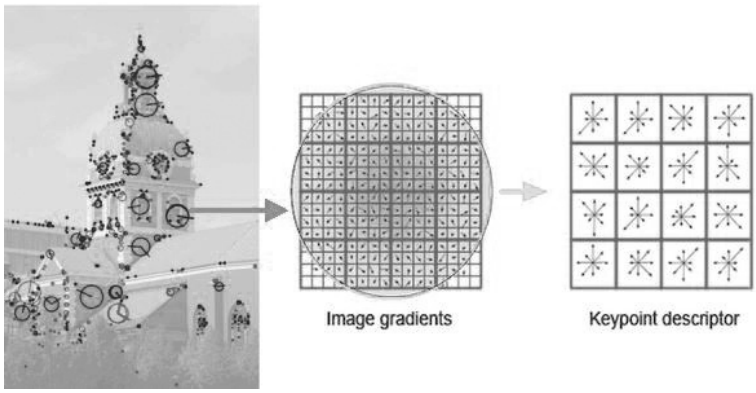
Phương pháp SIFT được thực hiện tuần tự theo 4 bước gồm:

Bước 1: Xác định các điểm bất biến với tỷ lệ, hướng của ảnh. Sử dụng thuật toán different-of-Gaussian, bước đầu tiên này nhằm tìm ra vị trí của các điểm "hấp dẫn" trên các kích thước và góc nhìn khác nhau.

Bước 2: Định vị các keypoint. Bằng cách sử dụng một hàm để kiểm tra, bước này nhằm xác định xem các điểm hấp dẫn tìm ra ở bước 1 có được chọn là keypoint hay không. Các điểm hấp dẫn có độ tương phản thấp hoặc có vị trí dọc theo các cạnh, thay đổi khi ảnh bị nhiễu sẽ bị loại bỏ.

Bước 3: Định hướng cho các keypoint.

Bước 4: Biểu diễn các keypoint dưới dạng vector. Mỗi keypoint sau khi được định hướng sẽ được mô tả dưới dạng một vector có 128 chiều mô tả vị trí, tỷ lệ và hướng của keypoint đó.



Hình 5. Minh họa thuật toán SIFT

Phương pháp HOG (Histogram of Oriented Gradients): Là phương pháp dựa trên biểu đồ (số lần) của các hướng đạo hàm (Gradients) trong các vùng cục bộ của ảnh.

Phương pháp này sẽ chia nhỏ ảnh ra thành các cell, các điểm trong cell sẽ được gán hướng thành các Gradients, từ đó ta có được biểu đồ về các hướng của các Gradients này. Ghép tất cả các biểu đồ của các cell này lại ta sẽ biểu diễn được ảnh ban đầu. Tuy nhiên, để nâng cao khả năng nhận dạng, các biểu đồ của các cell sẽ được chuẩn hóa về độ tương phản bằng cách ghép nhiều cell lại thành block không tách rời nhau (nghĩa là một cell có thể là phần tử của nhiều block), sau đó tính ngưỡng cường độ tương phản của các cell và sử dụng giá trị đó để chuẩn hóa lại các cell trong block. Từ đó, đặc trưng ảnh thu được sẽ ít phụ thuộc vào điều kiện ánh sáng hơn.

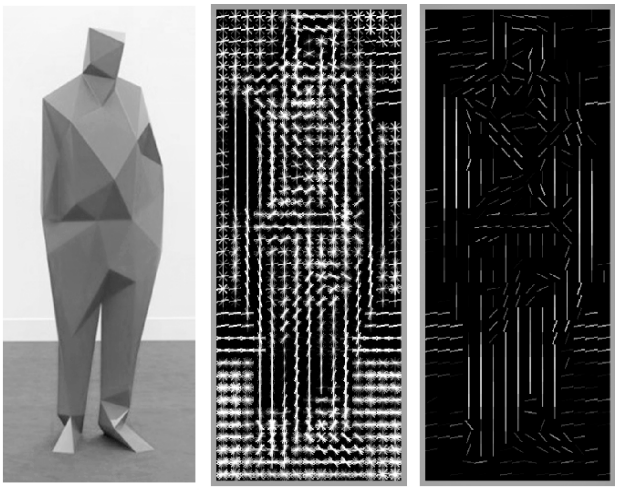
Tóm lại, phương pháp trích chọn đặc trưng HOG sẽ bao gồm các bước:

Bước 1: Tính gradients.

Bước 2: Gán hướng cho các gradients. Theo các nghiên cứu trước đây thì số lượng hướng là 9 sẽ cho là tốt nhất cho công việc nhận dạng các đối tượng trong ảnh.

Bước 3: Tính biểu đồ của các block. Thay vì tính biểu đồ cho các cell, việc tính biểu đồ của các block (có thành phần là các cell đã được chuẩn hóa) sẽ ít bị phụ thuộc vào độ tương phản và ánh sáng hơn.

Bước 4: Chuẩn hóa block.



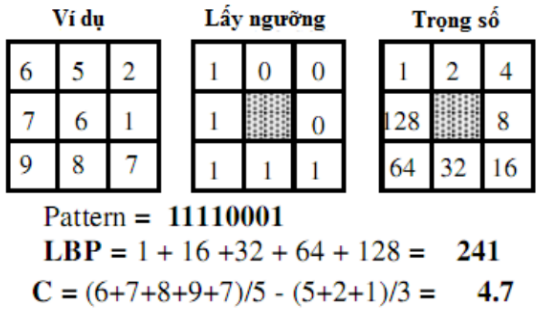
Hình 6. Minh họa thuật toán HOG

a) Hình gốc.

b) Mô tả hình gốc bằng đặc trưng HOG của hình gốc.

c) Hình b sau khi đã gán hướng.

Phương pháp LBP (Local Binary Pattern): Là phương pháp gắn nhãn các điểm ảnh bằng cách so sánh với các điểm ảnh lân cận. Cụ thể ta lấy ví dụ chọn 1 điểm ảnh và 8 điểm ảnh xung quanh, sử dụng giá trị của điểm ảnh trung tâm làm ngưỡng, giá trị LBP được xác định bằng cách nhân các giá trị ngưỡng với trọng số ứng với mỗi điểm ảnh sau đó cộng tổng lại. Hình dưới minh họa cách tính độ tương phản trực giao (C) là hiệu cấp độ xám trung bình của các điểm ảnh lớn hơn hoặc bằng ngưỡng với các điểm ảnh thấp hơn ngưỡng.



Hình 7. Minh họa thuật toán LBP

Phương pháp LBP có thể tóm tắt gồm các bước:

Bước 1: Xác định số lượng các điểm ảnh lân cận cần xét.

Bước 2: Tìm giá trị LBP điểm ảnh trung tâm dựa trên các điểm ảnh lân cận.

LBP được chứng minh là một công cụ hiệu quả để nhận dạng và phân loại kết cấu nhờ khả năng kháng nhiễu tốt, bất biến với thay đổi độ xám, và khả năng bất biến với xoay vòng, giúp phân tích kết cấu chính xác trong nhiều điều kiện và góc độ khác nhau (Ojala et al., 2002, p.971-987).

2.2.2. Phương pháp thực nghiệm

Sử dụng bộ dữ liệu ảnh MRI về u não của dự án MICCAI BRATS 2012 nêu trên để huấn luyện và kiểm thử mô hình máy học sử dụng phương pháp Học đa nhân. Đánh giá kết quả của các mô hình, đồng thời tinh chỉnh các thông số để đạt được mô hình tối ưu với bài toán phân lớp ảnh MRI bệnh u não.

3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Huấn luyện và kiểm thử mô hình Học đa nhân trong phân lớp ảnh MRI bệnh u não.

Để minh họa cho phương pháp học đa nhân, chúng tôi sử dụng bộ ảnh MRI về u não của dự án MICCAI BRATS 2012 bao gồm 25 mẫu chụp cộng hưởng từ đã được gán nhãn trước, có 9 mẫu thể hiện u não lành tính và 16 mẫu thể hiện u não ác tính. Trong mỗi mẫu có 4 file ảnh MRI lần lượt là T1, T1ce, T2 và FLAIR có định dạng ".mha" chứa 155 lát cắt chi tiết.

Các bước tiến hành thực nghiệm gồm:

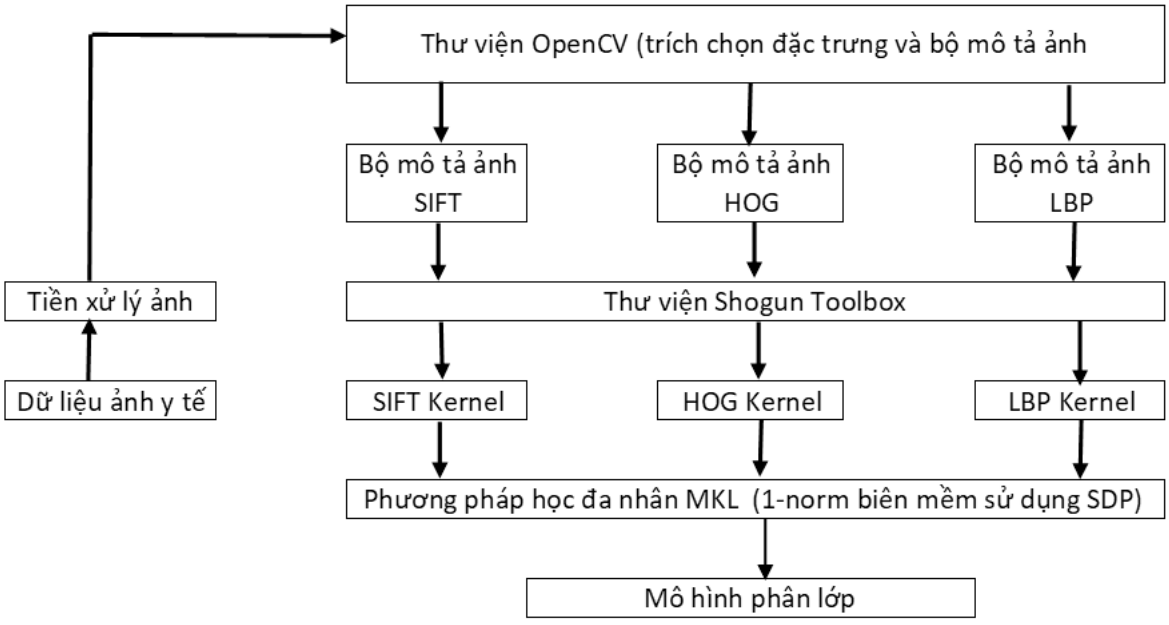
Bước 1: Tiền xử lý ảnh và trích chọn đặc trưng ảnh (sử dụng công cụ OpenCV).

Bước 2: Xây dựng các bộ mô tả ảnh.

Bước 3: Huấn luyện mô hình sử dụng công cụ (Shogun Toolbox).

Bước 4: Kiểm thử mô hình.

Bước 5: Đánh giá mô hình.



Hình 8. Sơ đồ quá trình thực nghiệm

Sau khi tiền xử lý ảnh và sử dụng 3 thuật toán trích chọn đặc trưng ảnh là SIFT, HOG và LBP (sử dụng công cụ OpenCV, thư viện nguồn mở về thị giác máy tính), chúng tôi thu được 3 bộ mô tả ảnh tương đương, mỗi bộ mô tả ảnh là một bộ các vector có cùng kích thước được xếp chồng lên nhau theo chiều dọc để phù hợp với cấu trúc dữ liệu khi sử dụng công cụ Shogun Toolbox, một thư viện nguồn mở về máy học để huấn luyện mô hình.

Chúng tôi sử dụng hàm CombinedKernel() trong thư viện của Shogun toolbox để nối các bộ mô tả SIFT, HOG, LBP lại thành một Kernel hợp nhất sau đó đưa vào cùng với nhãn để huấn luyện mô hình.

Sau đó, để đánh giá hiệu quả của mô hình, chúng tôi sử dụng phương pháp đánh giá theo kiểu thẩm tra chéo (cross-validation), với tập dữ liệu kiểm tra được chia ngẫu nhiên thành k tập con không giao nhau (gọi là "fold") có kích thước xấp xỉ nhau. Mỗi lần (trong số k lần) lặp, một tập con được sử dụng kiểm thử, và (k-1) tập con còn lại được dùng để huấn luyện.

Với mỗi lần huấn luyện và kiểm thử, một ma trận nhầm lẫn dùng để đánh giá tỉ lệ lỗi của mô hình được xây dựng như sau:

Bảng 1. Ma trận nhầm lẫn

		Nhãn lớp thực tế	
		1	0
Nhãn lớp dự đoán	1	TP	FP
	0	FN	TN

Trong đó:

TP: Những nhãn "1" hệ thống dự đoán đúng

FP: Những nhãn "1" hệ thống dự đoán sai

FN: Những nhãn "0" hệ thống dự đoán sai

TN: Những nhãn "0" hệ thống dự đoán đúng

Các chỉ số để đánh giá độ tốt của mô hình phân lớp được tính toán từ Ma trận nhầm lẫn như sau:

$$\text{Precision} = \frac{TP}{TP+FP}$$

Precision thể hiện mô hình hay dự đoán sai nhãn "1" như thế nào. Khi Precision lớn tức là False Positive nhỏ, mô hình dự đoán sai ít nhãn "1" còn khi Precision nhỏ tức là False Positive lớn. Precision càng lớn khả năng dự đoán sai nhãn "1" của hệ thống càng thấp nên Precision càng cao càng tốt.

$$\text{Recall} = \frac{TP}{TP+FN}$$

Recall thể hiện mô hình hay dự đoán sai nhãn "0" như thế nào. Khi Recall lớn tức là False Negative nhỏ, mô hình dự đoán sai ít nhãn "0" còn khi Recall nhỏ tức là False Negative lớn. Recall càng lớn khả năng dự đoán sai nhãn "0" của mô hình càng thấp nên Recall càng cao càng tốt.

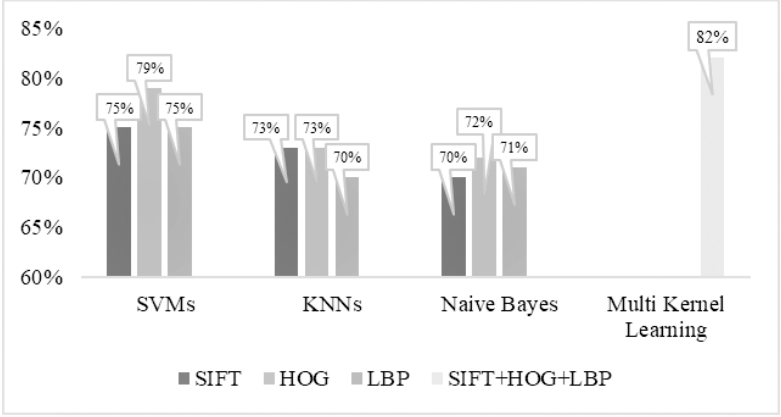
$$\text{F1-score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

F1-score chính là trung bình điều hòa (Harmonic Mean) của Precision và Recall. Nó phản ánh cho ta thấy tổng lượng FP và FN là bao nhiêu. Khi F1 score lớn thì số lượng nhãn dự đoán sai là nhỏ, mô hình hiệu quả và ngược lại khi F1 score nhỏ thì số lượng dự đoán sai là lớn, mô hình không tốt.

3.2. Kết quả

Sử dụng bộ dữ liệu nêu trên, chúng tôi tiến hành huấn luyện và kiểm thử mô hình với phương pháp Học đa nhân (MKL), sử dụng các hàm có sẵn trong bộ công cụ Shogun. Chúng tôi đã sử dụng các hàm nhân khác nhau và nhận thấy hàm nhân Gaussian cho kết quả tốt nhất (cả ba tập dữ liệu đều sử dụng hàm nhân Gausian), ngoài ra các tham số khác được sử dụng các giá trị mặc định. Kết quả cho thấy, trên tập dữ liệu huấn luyện, phương pháp MKL cho kết quả dự đoán là F1-score = 82%. Đây là một kết quả khả quan đối với một mô hình phân lớp nhị phân.

Ngoài ra, để đánh giá được mô hình, chúng tôi cũng sử dụng bộ dữ liệu trên cho các mô hình phân lớp khác như SVMs, KNNs và Naïve Bayes với 3 tập đặc trưng riêng biệt (SIFT, HOG, LBP). Chúng tôi sử dụng phần mềm Weka cho các thực nghiệm này. Kết quả dự đoán của các phương pháp được thể hiện trong Hình 9, trong đó chỉ số được dùng để so sánh là F1-score.



Hình 9. So sánh kết quả dự đoán của các mô hình phân lớp bằng chỉ số F1-score

3.3. Thảo luận

Đối với kết quả thực nghiệm của bài báo này, có thể thấy thuật toán trích chọn đặc trưng ảnh HOG cho độ chính xác cao hơn SIFT và LBP, khi kết hợp với mô hình phân lớp SVM cho ra kết quả dự đoán lên đến 79%. Kết quả này cho thấy HOG là thuật toán trích chọn đặc trưng mạnh mẽ và phổ biến, đặc biệt phù hợp sử dụng chung với mô hình phân lớp SVM (G. Tsai, 2010).

Mô hình phân lớp SVM cũng cho kết quả tốt hơn so với KNNs và Naïve Bayes trên cả 3 bộ mô tả ảnh (tương ứng với 3 tập đặc trưng), vì mô hình phân lớp SVM đặc biệt hiệu quả với các bài toán phân loại nhị phân trên các tập dữ liệu nhỏ và có độ phức tạp cao. Kết quả này cũng tương đồng với nghiên cứu của Wei và cộng sự (2005, p. 371-380) khi cũng cho thấy SVM có hiệu quả tốt với độ chính xác cao trong phân loại vi vôi hóa dạng cụm trên hình ảnh nhũ ảnh (một yếu tố quan trọng trong phát hiện ung thư vú lành tính và ác tính) nhờ khả năng tìm ra siêu phẳng tối ưu phân chia dữ liệu.

Với bộ dữ liệu không đồng nhất như trên, mô hình Học đa nhân cho thấy khả năng kết hợp tốt thông tin từ nhiều nguồn hoặc nhiều đặc trưng khác nhau khi cho ra chỉ số F1-score lên đến 82%. Tuy nhiên, với một số nhược điểm như độ phức tạp tính toán cao dẫn tới chi phí tính toán cao, quá trình lựa chọn và kết hợp các nhân (kernel) phức tạp, nguy cơ quá khớp (overfitting) là những thách thức cần được cân nhắc. Để sử dụng MKL hiệu quả, cần có sự đầu tư về tài nguyên, chuyên môn, và các biện pháp tối ưu hóa phù hợp. Trong nhiều trường hợp, các phương pháp đơn giản hơn có thể là lựa chọn thay thế khả thi hơn.

4. KẾT LUẬN

Bài báo mô tả chi tiết bộ dữ liệu ảnh MRI về u não của dự án MICCAI BRATS 2012, cơ sở chuẩn đoán bệnh u não là lành tính hay ác tính dựa trên hình ảnh MRI, từ đó tiến hành thực nghiệm xây dựng mô hình học máy bằng phương pháp Học đa nhân để phân lớp các hình ảnh MRI bệnh u não.

Kết quả dự đoán khả quan từ mô hình Học đa nhân (có độ chính xác 82%) cho thấy rõ ràng việc kết hợp nhiều cách mô tả ảnh thành một sẽ giúp cho mô hình phân lớp hoạt động hiệu quả hơn. Đồng thời cũng có thể thấy rằng, mô hình học đa nhân cho phép tích hợp được nhiều nguồn thông tin về các đối tượng cần phân lớp hơn, làm cho mô hình dự đoán giàu thông tin hơn, nên kết quả dự đoán cao hơn các mô hình khác. Vì vậy, có thể sử dụng mô hình học đa nhân cho bài toán mà các đối tượng dữ liệu được mô tả bởi các tập thuộc tính khác nhau. Kết quả của thực nghiệm này thể hiện rõ sức mạnh của phương pháp học đa nhân khi đã nâng cao độ chính xác trong việc phân loại bằng cách sử dụng cùng lúc nhiều bộ mô tả ảnh, từ đó hy vọng sẽ có nhiều nghiên cứu về việc ứng dụng phương pháp này trong môi trường y tế hơn nữa, vì ở đó chắc chắn sẽ có tập dữ liệu hình ảnh lớn hơn, nhiều loại hình ảnh hơn, nhiều thông tin thu thập được từ người bệnh hơn, nên độ chính xác sẽ cao hơn.

APPLYING MACHINE LEARNING ON CLASSIFICATION OF BRAIN TUMORS FROM MRI IMAGES

Tu Ngoc Thao[1], Nguyen Duc Thang[1], Phan Thi Dai Trang[1]

Received Date: 18/10/2024; Revised Date: 27/11/2024; Accepted for Publication: 28/11/2024

ABSTRACT

In recent decades, medical image analysis has become an essential field, providing directional predictions for healthcare professionals, especially valuable for regular, continuous diagnostic work with large volumes of data. Medical images of the same type and from the same department share many similarities, making it feasible to use computers and intelligent algorithms to detect abnormalities in these images. This paper presents the results of a machine learning model using Multiple Kernel Learning through training and testing with the MRI brain tumor image set from the MICCAI 2012 project: Multimodal Brain Tumor Segmentation. The result, evaluated through Cross-Validation, yielded an F1-score prediction of 82%, which is a fairly notable outcome for a binary classification model.

Keywords: *Machine learning, Artificial Intelligence, Multiple Kernel Learning, Image classification, Brain tumor MRI images.*

TÀI LIỆU THAM KHẢO

L. Wei, Y. Yang, R.M. Nishikawa, Y. Jiang (2005). *A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications*, IEEE Trans. Med. Imaging 24, pp. 371-380.

T. Ojala, M. Pietikainen, T. Maenpaa (2002). *Multiresolution gray-scale and rotation invariant texture classification with local binary patterns*, IEEE Trans. Pattern Anal. Mach. Intell. 24, pp. 971-987.

Cortes, C., & Vapnik, V. (1995). *Support-vector networks,* Machine Learning, 20(3), pp. 273-297.

Cover, T. M., & Hart, P. E. (1967). *Nearest neighbor pattern classification*, IEEE Transactions on Information Theory, 13(1), pp. 21-27.

McCallum, A., & Nigam, K. (1998). *A comparison of event models for Naive Bayes text classification*, AAAI-98 workshop on learning for text categorization, 752(1), pp. 41-48.

S. Sonnenburg, G. Raatsch, C. Schaafer, B. Schoalkopf (2006). *Large scale multiple kernel learning*, J. Mach. Learn. Res. 7, pp. 1531-1565.

G. Tsai (2010). *Histogram of Oriented Gradients*, University of Michigan.

G. Charpiat, M. Hofmann, B. Scholkopf (2015). *Kernel Methods in Medical Imaging*, Springer, US.

[1]Faculty of Natural Science and Technology, Tay Nguyen University;
Corresponding author: Tu Ngoc Thao; Tel: 0973351560; Email: tungocthao@ttn.edu.vn.