

BỘ GIÁO DỤC
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ

.....***.....



Nguyễn Thị Trang

**NGHIÊN CỨU PHƯƠNG PHÁP HỢP NHẤT ĐẶC TRƯNG
ĐA NGUỒN CHO PHÂN LOẠI ẢNH TỔN THƯƠNG DA**

TÓM TẮT LUẬN ÁN TIẾN SĨ MÁY TÍNH

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Hà Nội, 2026

**Công trình được hoàn thành tại Học viện Khoa học và Công nghệ,
Viện Hàn lâm Khoa học và Công nghệ Việt Nam**

Người hướng dẫn khoa học:

1. Người hướng dẫn 1: PGS.TS. Nguyễn Việt Anh, Viện Công nghệ Thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam
2. Người hướng dẫn 2: PGS.TS. Vũ Văn Hiệu, Viện Công nghệ Thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

Phản biện 1:

Phản biện 2:

Luận án được bảo vệ trước Hội đồng đánh giá luận án tiến sĩ cấp Học viện họp tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam vào hồi giờ ngày tháng năm

Có thể tìm hiểu luận án tại:

1. Thư viện Học viện Khoa học và Công nghệ
2. Thư viện Quốc gia Việt Nam

MỞ ĐẦU

1. Tính cấp thiết và khoảng trống nghiên cứu

Phân loại ảnh tổn thương da có ý nghĩa trong hỗ trợ sàng lọc sớm [1, 2, 3], nhưng dữ liệu thực tế biến thiên mạnh về màu sắc, kết cấu, hình dạng, điều kiện thu nhận và phân bố lớp. Các lớp nguy hiểm thường có ít mẫu, nhiều tổn thương khác lớp lại giống nhau về thị giác, trong khi việc gán nhãn cần chuyên gia. Đặc trưng thủ công phản ánh màu sắc, kết cấu và hình dạng nhưng phụ thuộc thiết kế; đặc trưng học sâu có năng lực biểu diễn mạnh nhưng không phải mọi bản đồ đặc trưng đều hữu ích. Nếu giữ toàn bộ phản hồi trung gian, biểu diễn có thể chứa nhiễu và dư thừa.

Hợp nhất ảnh, đặc trưng thủ công và metadata lâm sàng giúp tận dụng thông tin bổ trợ [8, 9, 14], song ghép nối trực tiếp chưa xử lý đầy đủ chênh lệch số chiều, phân bố và mức tin cậy của từng nguồn. Phần lớn mô hình còn xử lý từng ảnh độc lập, chưa khai thác quan hệ tương đồng giữa các mẫu. Trong học bán giám sát, nhãn giả sai có thể được tái sử dụng và khuếch đại sai lệch, đặc biệt đối với lớp hiếm.

Luận án xác định hai khoảng trống liên tiếp. Thứ nhất, còn thiếu một quy trình thống nhất để chọn bản đồ đặc trưng theo khoảng entropy phù hợp, hợp nhất với đặc trưng thủ công và cân bằng biểu diễn. Thứ hai, còn thiếu một khung đa nguồn bán giám sát có khả năng xây dựng quan hệ liên mẫu bằng đồ thị và chỉ sử dụng nhãn giả có độ tin cậy cao.

Về phương diện phương pháp, luận án xem mỗi nguồn dữ liệu là một góc quan sát khác nhau của cùng đối tượng. Ảnh cung cấp dấu hiệu hình thái trực quan; đặc trưng thủ công mã hóa tri thức miền; metadata bổ sung bối cảnh lâm sàng. Vì vậy, bài toán không chỉ là tăng số chiều biểu diễn mà là xác định thông tin nào đáng giữ, cách chuẩn hóa đóng góp giữa các nguồn và cách kiểm soát sai số khi thông tin được truyền giữa các mẫu. Quan điểm này chi phối toàn bộ thiết kế từ BlockNet đến mô hình đồ thị bán giám sát.

2. Mục tiêu, đối tượng và phạm vi nghiên cứu

Mục tiêu tổng quát là phát triển các phương pháp hợp nhất đặc trưng đa nguồn nhằm nâng cao hiệu quả phân loại ảnh tổn thương da trong điều kiện dữ liệu biến thiên lớn, mất cân bằng lớp, hạn chế nhãn và có nhiều nguồn thông tin bổ trợ.

Mục tiêu thứ nhất là đề xuất BlockNet, lựa chọn bản đồ đặc trưng theo khoảng entropy, hỗ trợ cân bằng trong không gian đặc trưng và hợp nhất đặc trưng học sâu với đặc trưng thủ công. Mục tiêu thứ hai là phát triển khung học đa nguồn bán giám sát dựa trên đồ thị, kết hợp đặc trưng học sâu, đặc trưng thủ công và metadata lâm sàng; khai thác quan hệ giữa các mẫu và dữ liệu chưa gán nhãn thông qua gán nhãn giả có kiểm soát độ tin cậy.

Đối tượng nghiên cứu là phương pháp tổ chức, chọn lọc, hợp nhất và khai thác biểu diễn đa nguồn cho phân loại ảnh tổn thương da. Phạm vi thực nghiệm gồm VNCancer, HAM10000 và ISIC 2020. Luận án không xây dựng backbone hoàn toàn mới, không thực hiện thử nghiệm lâm sàng và không xem mô hình là công cụ chẩn đoán độc lập.

Phương pháp nghiên cứu kết hợp khảo sát lý thuyết, mô hình hóa thuật toán, thực nghiệm tính toán và phân tích định lượng. Kết quả được đánh giá bằng Accuracy, Pre-

cision, Recall, F1-score, AUC, Sensitivity, Specificity, ma trận nhầm lẫn và kiểm chứng lặp phù hợp với dữ liệu mất cân bằng.

Các thí nghiệm được tổ chức để trả lời ba câu hỏi: chọn lọc theo entropy có loại bỏ được phản hồi kém hữu ích hay không; hợp nhất đặc trưng đa nguồn có cải thiện độ phân biệt so với từng nguồn riêng lẻ hay không; và dữ liệu chưa gán nhãn có thể được khai thác an toàn đến mức nào. Khi diễn giải kết quả, luận án ưu tiên các chỉ số cân bằng giữa lớp thay vì chỉ dựa vào Accuracy, đồng thời đối chiếu ma trận nhầm lẫn và AUC để tránh kết luận thiên lệch bởi lớp chiếm đa số.

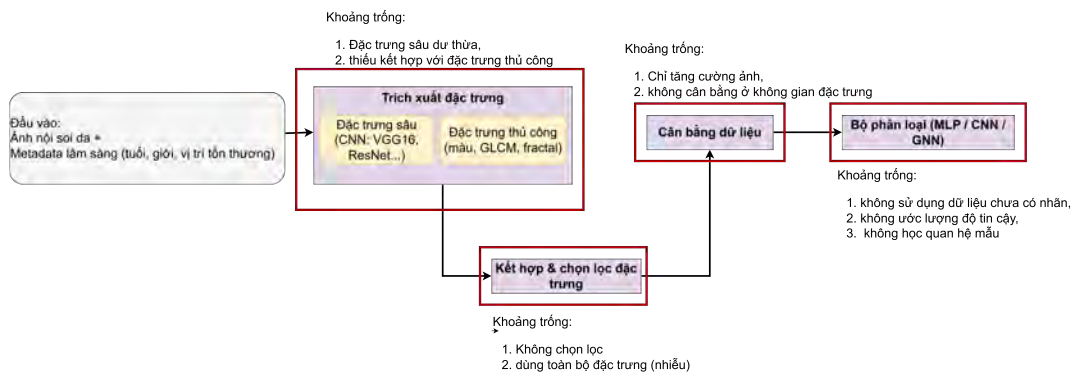
3. Đóng góp và cấu trúc luận án

Đóng góp thứ nhất là BlockNet [CT3]. Mô hình dùng entropy kích hoạt để sàng lọc bản đồ đặc trưng tại nhiều block của VGG16, chuyển các bản đồ được giữ lại thành vector, hợp nhất với đặc trưng thủ công và cân bằng dữ liệu trong không gian đặc trưng. Đóng góp nằm ở quy trình tích hợp chọn lọc, hợp nhất và cân bằng biểu diễn, không phải ở một backbone CNN mới.

Đóng góp thứ hai là khung học bán giám sát đa nguồn trên đồ thị [CT4]. Đặc trưng học sâu, đặc trưng thủ công và metadata được hợp nhất thành biểu diễn nút; đồ thị k láng giềng gần nhất mô hình hóa quan hệ liên mẫu; entropy dự đoán được dùng để phân tầng độ tin cậy và chọn nhãn giả trước khi tái huấn luyện.

Hai đóng góp dùng entropy ở hai cấp khác nhau: Chương 2 tính entropy trên bản đồ kích hoạt để chọn biểu diễn; Chương 3 tính entropy trên phân phối xác suất dự đoán để chọn mẫu chưa gán nhãn. Chương 1 trình bày cơ sở lý thuyết; Chương 2 trình bày BlockNet và thực nghiệm trên VNCancer, HAM10000; Chương 3 trình bày mô hình đa nguồn trên đồ thị và thực nghiệm trên ISIC 2020; phần cuối tổng hợp kết quả, hạn chế và hướng nghiên cứu tương lai.

Hình 1 tóm lược chuỗi vấn đề mà luận án giải quyết.



Hình 1: Khoảng trống nghiên cứu của hệ thống phân loại ảnh tổn thương da.

Mạch nghiên cứu vì thế đi từ biểu diễn nội mẫu đến quan hệ liên mẫu. Chương 2 giải quyết chất lượng biểu diễn của từng ảnh trước khi phân loại; Chương 3 mở rộng sang cấu trúc tập dữ liệu, nơi một mẫu được hiểu trong quan hệ với các mẫu tương đồng. Sự phân chia này giúp hai đóng góp có vai trò rõ ràng và bổ trợ: BlockNet giảm nhiễu ở đầu vào biểu diễn, còn mô hình đồ thị cùng cơ chế độ tin cậy giảm rủi ro khi sử dụng quan hệ và nhãn dự đoán.

4. Câu hỏi nghiên cứu và cách tiếp cận tổng thể

Luận án trả lời bốn câu hỏi liên kết: bản đồ đặc trưng nào chứa lượng thông tin phù hợp; các nguồn đặc trưng cần được hợp nhất thế nào để khai thác tính bổ trợ; quan hệ tương đồng giữa mẫu có thể được mô hình hóa bằng đồ thị hay không; và dữ liệu chưa gán nhãn phải đạt điều kiện tin cậy nào trước khi tái sử dụng. Các câu hỏi này chuyển trọng tâm từ việc chỉ lựa chọn một kiến trúc phân loại sang tổ chức và kiểm soát thông tin trong toàn bộ quy trình học.

Cách tiếp cận gồm hai tầng. Ở tầng biểu diễn, entropy kích hoạt được tính cho từng bản đồ, các phản hồi phù hợp theo block được giữ lại rồi hợp nhất với HSV, GLCM và fractal. Ở tầng dữ liệu, biểu diễn ảnh kết hợp metadata tạo thành nút, đồ thị k -NN mô hình hóa quan hệ liên mẫu và GNN truyền thông tin. Entropy dự đoán được dùng ở pha bán giám sát để phân tầng độ tin cậy trước khi tạo tập huấn luyện mở rộng.

Đánh giá được thực hiện trong phạm vi nghiên cứu phương pháp. So sánh giữa mô hình phải giữ chung cách chia dữ liệu và giao thức huấn luyện. Với dữ liệu mất cân bằng, Macro F1, Macro AUC, Sensitivity và Specificity được xem xét cùng Accuracy. Kết quả định lượng cho biết tính khả thi trong các bộ dữ liệu đã khảo sát, không thay thế đánh giá lâm sàng. Các giới hạn về metadata, sai lệch miền, ngoại kiểm và chất lượng nhãn được dùng để xác định đúng mức độ khái quát của kết luận.

CHƯƠNG 1: TỔNG QUAN VÀ CƠ SỞ LÝ THUYẾT

1.1. Bài toán và biểu diễn đa nguồn

Với tập dữ liệu $D = \{(x_i, y_i)\}_{i=1}^N$, mô hình học ánh xạ từ ảnh hoặc biểu diễn đa nguồn tới một trong C lớp tổn thương. Trong trường hợp có metadata m_i , mẫu được biểu diễn bởi cặp (x_i, m_i) . Khó khăn chính gồm biến thiên nội lớp, tương đồng liên lớp, mất cân bằng dữ liệu, khác biệt thiết bị và thiếu nhãn chuyên gia.

Đặc trưng thủ công mô tả những thuộc tính có ý nghĩa trực quan [36, 37, 39, 41]: histogram RGB/HSV cho màu sắc; GLCM, LBP, Gabor và wavelet cho kết cấu; HOG, biên và fractal cho hình dạng hoặc độ phức tạp. Đặc trưng học sâu được trích từ các tầng tích chập [50, 51, 52], nơi mỗi kênh tạo một bản đồ đặc trưng. Biểu diễn sâu tổng quát có thể viết:

$$F_i^{(l)} = \{f_{i,1}^{(l)}, f_{i,2}^{(l)}, \dots, f_{i,K_l}^{(l)}\}, \quad (1.26)$$

Không phải mọi $F_k^{(l)}$ đều hữu ích. Một số phản hồi gần như hằng, một số quá phân tán hoặc nhạy với nhiễu. Đây là cơ sở để luận án đánh giá bản đồ đặc trưng trước bước hợp nhất thay vì sử dụng toàn bộ đầu ra trung gian.

Trong biểu diễn đa nguồn, yêu cầu quan trọng là bảo toàn thông tin bổ trợ nhưng hạn chế sự lẫn át của nguồn có số chiều lớn. Chuẩn hóa và phép chiếu về không gian chung được xem là bước tiền xử lý có ý nghĩa mô hình, bởi chúng quyết định khoảng cách giữa các mẫu và ảnh hưởng trực tiếp đến bộ phân loại hoặc cấu trúc đồ thị ở các chương sau. Do đó, chất lượng hợp nhất phải được đánh giá qua cả hiệu năng cuối và khả năng giải thích vai trò của từng nguồn.

1.2. Hợp nhất đặc trưng và entropy

Hợp nhất đa nguồn nhằm tạo biểu diễn chung từ các nguồn bổ trợ. Với đặc trưng học sâu $z_i^{(d)}$ và đặc trưng thủ công $z_i^{(h)}$, phép ghép nối có dạng:

$$x_i^{(dh)} = \mathcal{F}(x_i^{(d)}, x_i^{(h)}), \quad (1.28)$$

Với phân phối xác suất dự đoán $\mathbf{p} = (p_1, \dots, p_C)$, entropy dự đoán được xác định bởi:

$$H_{\text{pred}}(p_{\theta}(x)) = - \sum_{c=1}^C p_{\theta}(c|x) \ln p_{\theta}(c|x). \quad (1.33)$$

Entropy thấp biểu thị phân phối tập trung hơn, còn entropy cao biểu thị dự đoán phân tán hơn. Trên bản đồ đặc trưng, entropy được tính từ histogram chuẩn hóa của các mức kích hoạt:

$$H_{\text{act}}(f) = - \sum_{r=1}^{K_f} p_r \log_2 p_r. \quad (1.34)$$

Hai dạng entropy khác nhau về đối tượng và mục đích: entropy bản đồ đặc trưng dùng để sàng lọc biểu diễn ở Chương 2; entropy dự đoán dùng để kiểm soát nhãn giả ở Chương 3. Entropy không trực tiếp chứng minh khả năng phân biệt lớp, vì vậy mọi lựa

chọn đều phải được kiểm chứng bằng hiệu năng phân loại.

Entropy trong luận án không được diễn giải như một thước đo chất lượng tuyệt đối. Giá trị quá thấp có thể tương ứng với phản hồi gần như đồng nhất, trong khi giá trị quá cao có thể phản ánh hoạt hóa phân tán hoặc nhiễu. Vì vậy, tiêu chuẩn phù hợp là một khoảng entropy phụ thuộc block thay vì ngưỡng toàn cục. Cách tiếp cận này giữ lại các bản đồ có mức biến thiên vừa đủ để mô tả cấu trúc tổn thương, đồng thời giảm số phản hồi dư thừa trước khi hợp nhất.

1.3. Học trên đồ thị và học bán giám sát

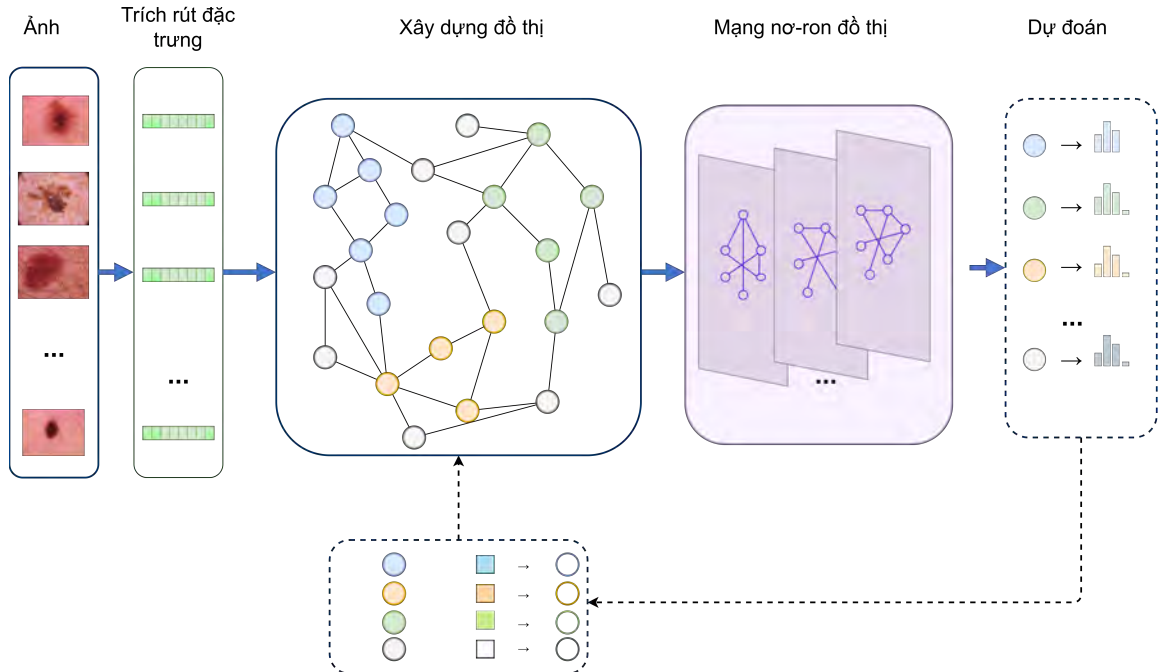
Đồ thị $G = (V, E)$ biểu diễn mỗi mẫu bằng một nút [20, 21, 22], còn cạnh mô tả quan hệ tương đồng. Ma trận kề được xác định từ độ tương đồng và ngưỡng hoặc quan hệ k láng giềng gần nhất:

$$A_{ij} = \text{sim}(x_i, x_j), \quad (1.36)$$

Tại mỗi tầng GNN, biểu diễn nút được cập nhật từ chính nút và các nút lân cận:

$$h_i^{(l+1)} = \sigma \left(\sum_{j \in \mathcal{N}(i)} W^{(l)} h_j^{(l)} \right), \quad (1.37)$$

Học bán giám sát sử dụng đồng thời tập có nhãn và tập chưa gán nhãn. Mô hình ban đầu được huấn luyện trên dữ liệu có nhãn, dự đoán dữ liệu chưa gán nhãn, chọn các dự đoán đáng tin cậy rồi tái huấn luyện. Nguy cơ chính là confirmation bias: nhãn giả sai nhưng tự tin có thể làm lệch biên quyết định. Vì vậy, Chương 3 chỉ chọn nhóm entropy thấp và không xem nhãn giả là nhãn tham chiếu.



Hình 1.7: Minh họa pipeline học trên đồ thị cho bài toán phân loại ảnh tổn thương da. Các đặc trưng ảnh và metadata được sử dụng để xây dựng đồ thị, trong đó mỗi nút biểu diễn một mẫu và các cạnh phản ánh mức độ tương đồng giữa các mẫu. Mô hình GNN lan truyền thông tin giữa các nút lân cận nhằm học biểu diễn có tính ngữ cảnh, tạo cơ sở cho việc khai thác quan hệ giữa các mẫu trong các thiết lập học bán giám sát.

Đối với đồ thị, lựa chọn biểu diễn nút và quy tắc tạo cạnh có quan hệ chặt chẽ. Nếu biểu diễn nút chưa phản ánh đúng tương đồng lâm sàng, đồ thị k -NN có thể nổi các mẫu khác lớp và làm lan truyền thông tin sai. Vì vậy, luận án xây dựng nút sau bước hợp nhất đa nguồn, dùng độ tương đồng cosine và kiểm soát số láng giềng. Học bán giám sát chỉ được thực hiện sau khi mô hình có giám sát đã tạo phân phối dự đoán đủ ổn định để đánh giá độ tin cậy.

1.4. Cơ sở hình thành phương pháp của luận án

Tổng quan cho thấy bốn vấn đề liên kết: chất lượng biểu diễn đầu vào, hợp nhất nguồn không đồng nhất, quan hệ giữa các mẫu và khai thác dữ liệu chưa gán nhãn. Các phép hợp nhất trực tiếp không bảo đảm loại bỏ nhiễu hoặc cân bằng đóng góp của từng nguồn. GNN có thể khai thác cấu trúc liên mẫu nhưng phụ thuộc chất lượng biểu diễn nút và cạnh. Gán nhãn giả mở rộng dữ liệu huấn luyện nhưng phải đi cùng cơ chế kiểm soát độ tin cậy.

Từ đó, luận án tổ chức hai hướng. BlockNet xử lý biểu diễn ở cấp bản đồ đặc trưng: trích xuất nhiều block, tính entropy, giữ các phản hồi trong khoảng phù hợp, hợp nhất với đặc trưng thủ công và cân bằng biểu diễn. Hướng thứ hai xử lý dữ liệu ở cấp mẫu: hợp nhất ảnh với metadata, xây dựng đồ thị k -NN, học GNN và chọn nhãn giả theo entropy dự đoán.

Các chỉ số macro được ưu tiên vì Accuracy có thể bị chi phối bởi lớp đa số. Macro F1 và Macro AUC phản ánh cân bằng giữa các lớp; Sensitivity và Specificity làm rõ khả năng nhận diện mẫu ác tính và lành tính. Kết quả được diễn giải trong phạm vi giao thức thực nghiệm, không suy rộng thành hiệu quả lâm sàng.

Kết luận Chương 1

Chương 1 xác lập nền tảng cho hai đóng góp và làm rõ ranh giới giữa entropy kích hoạt với entropy dự đoán.

Từ các cơ sở trên, hai vai trò của entropy được phân biệt rõ. Entropy kích hoạt là thuộc tính của bản đồ trung gian và phục vụ chọn đặc trưng; entropy dự đoán là thuộc tính của phân phối xác suất đầu ra và phục vụ chọn mẫu. Không đồng nhất hai đại lượng này giúp tránh suy diễn rằng bản đồ có entropy cao tất yếu tạo dự đoán không chắc chắn. Chúng tác động ở hai tầng khác nhau nhưng cùng thực hiện nguyên tắc kiểm soát thông tin trước khi đưa vào bước học tiếp theo.

Phân tích mối liên hệ giữa các thành phần lý thuyết

Đặc trưng học sâu và đặc trưng thủ công khác nhau cả về nguồn gốc, kích thước và ý nghĩa. Đặc trưng học sâu được tối ưu từ dữ liệu và có khả năng mô tả cấu trúc phức tạp, trong khi HSV, GLCM và fractal cung cấp các đại lượng gắn với màu sắc, kết cấu và độ phức tạp hình học. Nếu chỉ ghép nối trực tiếp, nhánh có số chiều lớn có thể chi phối hàm khoảng cách và gradient. Vì vậy, chuẩn hóa, chọn lọc và phép chiếu không chỉ là thao tác kỹ thuật mà là điều kiện để các nguồn đóng góp trên thang đo tương thích.

Entropy được lựa chọn vì có thể mô tả mức phân tán thông tin mà không cần nhãn ở cấp bản đồ. Tuy nhiên, entropy không cho biết trực tiếp phản hồi thuộc cấu trúc tổn thương hay nền ảnh. Luận án khắc phục một phần bằng cách xác định khoảng phù hợp theo từng block và kiểm chứng bằng hiệu năng phân loại. Cách tiếp cận này tạo cầu nối giữa tiêu chuẩn không giám sát ở tầng biểu diễn và mục tiêu có giám sát ở đầu ra, đồng

thời giữ được khả năng truy vết số bản đồ được chọn tại mỗi tầng.

Trong mô hình đồ thị, giả định đồng dạng cho rằng các mẫu gần nhau trong không gian biểu diễn có xu hướng chia sẻ nhãn. Giả định chỉ đáng tin khi biểu diễn hợp nhất đã loại bớt nhiễu và mã hóa được thông tin bổ trợ. Truyền thông điệp có thể tăng tính nhất quán cục bộ nhưng cũng có nguy cơ làm trơn quá mức hoặc lan truyền sai số qua cạnh khác lớp. Vì vậy, số láng giềng, độ sâu GNN và cơ chế chọn nhãn giả phải được đánh giá cùng nhau.

Từ đó, cơ sở lý thuyết dẫn đến ba nguyên tắc thiết kế. Thứ nhất, chọn thông tin trước khi hợp nhất. Thứ hai, xây dựng quan hệ liên mẫu sau khi đã có biểu diễn đa nguồn. Thứ ba, chỉ tái sử dụng dự đoán khi phân phối đầu ra đạt mức tin cậy xác định. Ba nguyên tắc này tạo thành mạch logic xuyên suốt cho hai đóng góp thực nghiệm của luận án.

Các điều kiện đánh giá và nguy cơ sai lệch

Sai lệch có thể xuất hiện từ thiết bị chụp, điều kiện chiếu sáng, tiền xử lý, tỷ lệ lớp và cách chia bệnh nhân. Nếu ảnh của cùng bệnh nhân có mặt ở cả tập huấn luyện và tập kiểm tra, mô hình có thể ghi nhớ đặc điểm cá nhân và tạo kết quả lạc quan. Vì vậy, mọi bước chọn ngưỡng entropy, siêu tham số đồ thị và điều kiện gán nhãn giả chỉ được dựa trên tập huấn luyện hoặc xác thực; tập kiểm tra phải được giữ độc lập đến bước đánh giá cuối.

Mất cân bằng lớp đòi hỏi phân biệt hiệu năng tổng thể với hiệu năng theo lớp. Accuracy có thể tăng khi mô hình ưu tiên lớp phổ biến trong khi Recall của lớp nguy hiểm vẫn thấp. Macro F1 gán trọng số ngang nhau cho các lớp; AUC đánh giá khả năng xếp hạng ở nhiều ngưỡng; Sensitivity và Specificity phản ánh hai phía của quyết định. Ma trận nhầm lẫn giúp nhận diện cặp lớp thường bị nhầm và liên hệ sai số với hình thái hoặc phân bố dữ liệu.

Khả năng tái lập phụ thuộc việc công bố cấu hình trích xuất, chuẩn hóa, số kênh được chọn, hạt giống ngẫu nhiên, số folds và tiêu chuẩn dừng. Với GNN, cần ghi cách dựng cạnh, giá trị k , quy tắc đối xứng hóa và số tầng truyền thông điệp. Với nhãn giả, cần báo cáo ngưỡng entropy, số mẫu từng mức tin cậy và số mẫu thực sự được sử dụng. Khung báo cáo thống nhất giúp tránh quy kết cải thiện cho một thành phần khi nhiều yếu tố thay đổi đồng thời.

Một yêu cầu khác là phân biệt ý nghĩa thống kê với ý nghĩa thực tiễn. Chênh lệch trung bình nhỏ có thể ổn định qua nhiều folds nhưng chưa chắc tạo cải thiện đáng kể cho lớp bệnh nguy hiểm; ngược lại, cải thiện Recall của lớp hiếm có thể quan trọng dù Accuracy tổng thể thay đổi ít. Vì vậy, kết quả cần được trình bày cùng độ lệch chuẩn, khoảng tin cậy và phân tích theo lớp. Khi so sánh nhiều cấu hình, cần hạn chế lựa chọn hậu nghiệm chỉ dựa vào kết quả cao nhất.

Các khái niệm trong chương cũng xác định ranh giới diễn giải. Trọng số chú ý thể hiện mức đóng góp trong mô hình nhưng không chứng minh quan hệ nhân quả. Entropy thấp cho biết phân phối tập trung nhưng không bảo đảm dự đoán đúng. Cạnh gần trong đồ thị biểu thị tương đồng theo biểu diễn đã chọn, không đồng nghĩa hai mẫu giống nhau về mọi đặc điểm lâm sàng. Việc nêu rõ các ranh giới này giúp kết quả của Chương 2 và Chương 3 được đánh giá đúng bản chất phương pháp.

CHƯƠNG 2: BLOCKNET: CHỌN LỌC ENTROPY VÀ HỢP NHẤT ĐẶC TRƯNG

2.1. Phát biểu bài toán và ý tưởng

Cho tập ảnh có nhãn, mục tiêu là học bộ phân loại từ biểu diễn hợp nhất nhưng không đưa toàn bộ bản đồ đặc trưng vào mô hình. Với tập bản đồ từ các block CNN và vector thủ công, BlockNet tìm tập bản đồ hợp lệ theo entropy, chuyển chúng thành vector, hợp nhất và cân bằng dữ liệu trước khi huấn luyện.

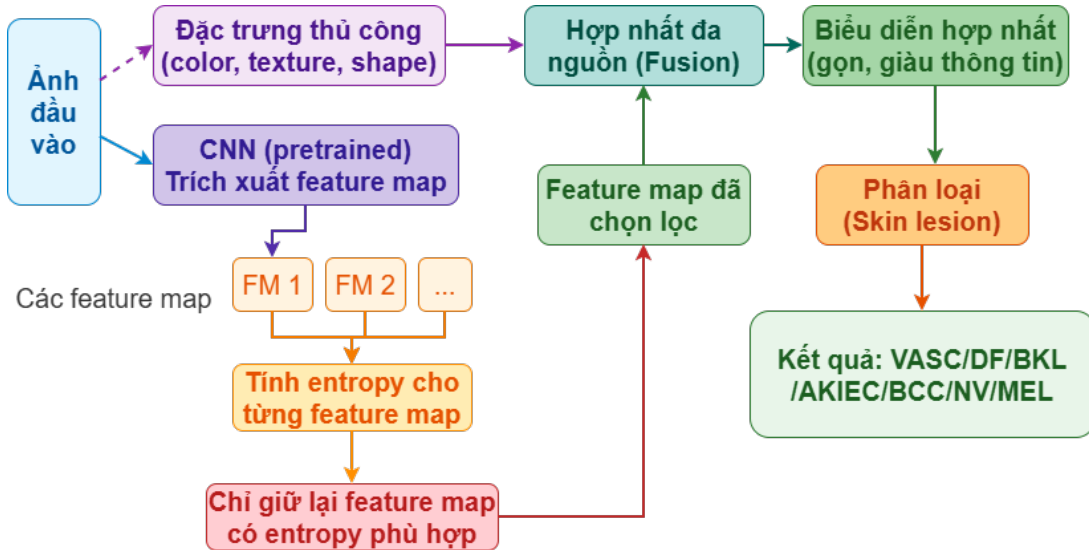
Entropy của một bản đồ F được tính từ phân phối histogram chuẩn hóa:

$$H_{\text{act}}(F_m^{(b)}(x)) = - \sum_{r=1}^{K_m} p_r \log_2 p_r, \quad (2.16)$$

Giá trị này thỏa mãn:

$$0 \leq H_{\text{act}}(F_m^{(b)}(x)) \leq \log_2 K_m \leq \log_2 N_m. \quad (2.17)$$

Khoảng entropy được xác định riêng cho từng block. Bản đồ quá ít biến thiên thường chứa phản hồi nghèo; bản đồ quá phân tán có thể chứa nhiễu. Phương pháp không mặc định entropy càng cao càng tốt mà lựa chọn một khoảng được kiểm chứng qua phân loại.



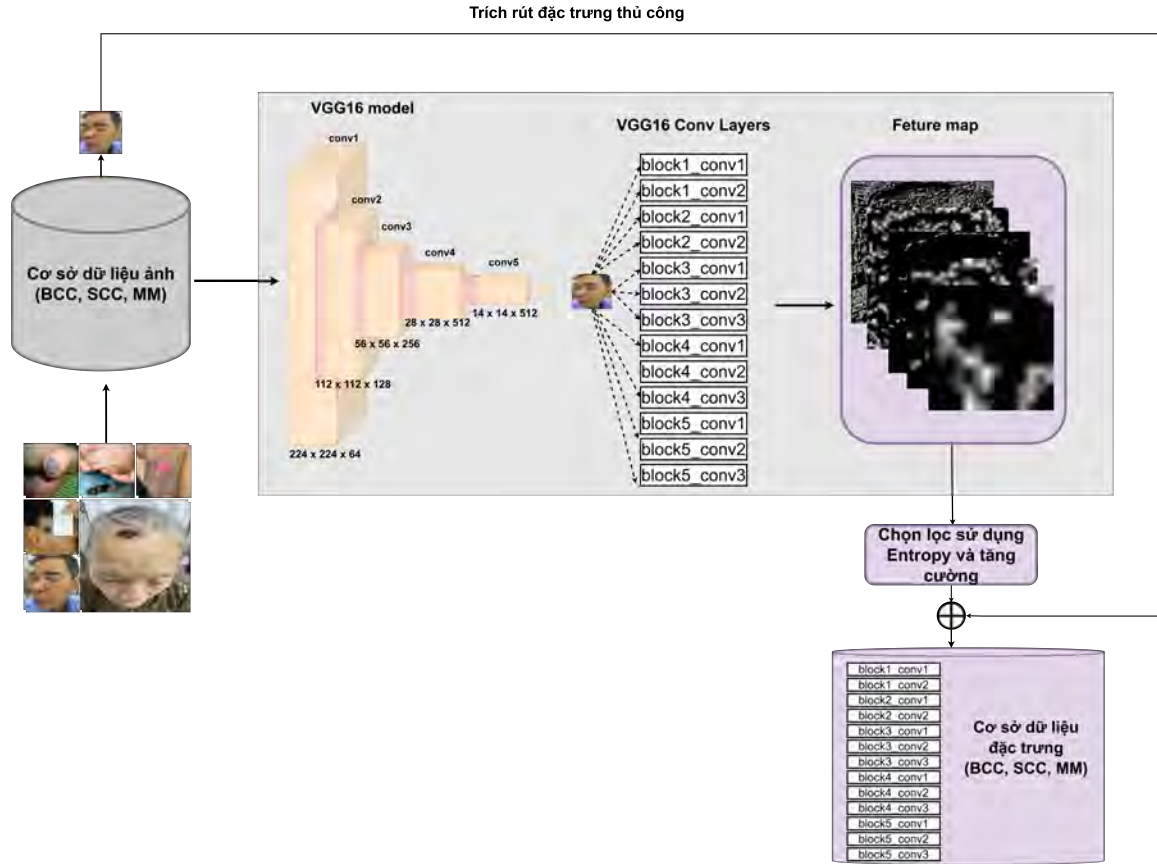
Hình 2.1: Ý tưởng phương pháp đề xuất: chọn lọc bản đồ đặc trưng dựa trên entropy và hợp nhất đặc trưng đa nguồn.

Đầu vào BlockNet kết hợp ảnh đã tiền xử lý với đặc trưng thủ công; chọn lọc được thực hiện riêng tại từng block để phù hợp với mức trừu tượng của phản hồi.

2.2. Kiến trúc tổng quan của BlockNet

BlockNet gồm năm khối: trích xuất bản đồ đặc trưng từ VGG16; tính entropy và chọn bản đồ; tạo vector sâu; trích xuất đặc trưng thủ công; hợp nhất và cân bằng trong không gian đặc trưng trước bộ phân loại MLP. Backbone được giữ cố định về nguyên lý, do đó đóng góp tập trung vào tổ chức biểu diễn.

Các block nông bảo toàn màu sắc, biên và kết cấu cục bộ; block sâu mô tả cấu trúc ngữ nghĩa hơn nhưng có thể mất chi tiết. Khai thác nhiều block cho phép đánh giá riêng từng mức biểu diễn. Các bản đồ được chọn được vector hóa và chuẩn hóa trước khi hợp nhất với histogram HSV, GLCM và fractal.



Hình 2.3: Kiến trúc tổng quan của mô hình BlockNet.

Kiến trúc được triển khai theo chuỗi: trích xuất bản đồ, tính entropy, xác định tập hợp hợp lệ, vector hóa, ghép với đặc trưng thủ công, cân bằng biểu diễn và phân loại. Việc đặt cân bằng sau hợp nhất bảo đảm mẫu tổng hợp được tạo trong cùng không gian mà bộ phân loại sử dụng. Thiết kế này cũng tránh tạo ảnh giả trực tiếp, vốn có thể sinh cấu trúc tổn thương không hợp lý và làm khó việc kiểm soát ý nghĩa lâm sàng.

Ở mỗi block, đầu ra có kích thước và số kênh khác nhau nên bước vector hóa phải bảo toàn thứ tự xử lý và tạo biểu diễn có kích thước xác định. Các vector sau chọn lọc được chuẩn hóa trước khi ghép, nhờ đó đặc trưng có miền giá trị lớn không tự động chi phối khoảng cách. Nhánh thủ công gồm các đại lượng màu HSV, kết cấu GLCM và fractal, tương ứng với ba nhóm dấu hiệu thường được quan sát khi mô tả tổn thương. Biểu diễn hợp nhất sau cùng được đưa vào cùng một bộ phân loại để việc so sánh giữa các cấu hình phản ánh thay đổi ở bước chọn lọc, thay vì thay đổi đồng thời kiến trúc dự đoán.

2.3. Xác định khoảng entropy và chọn bản đồ

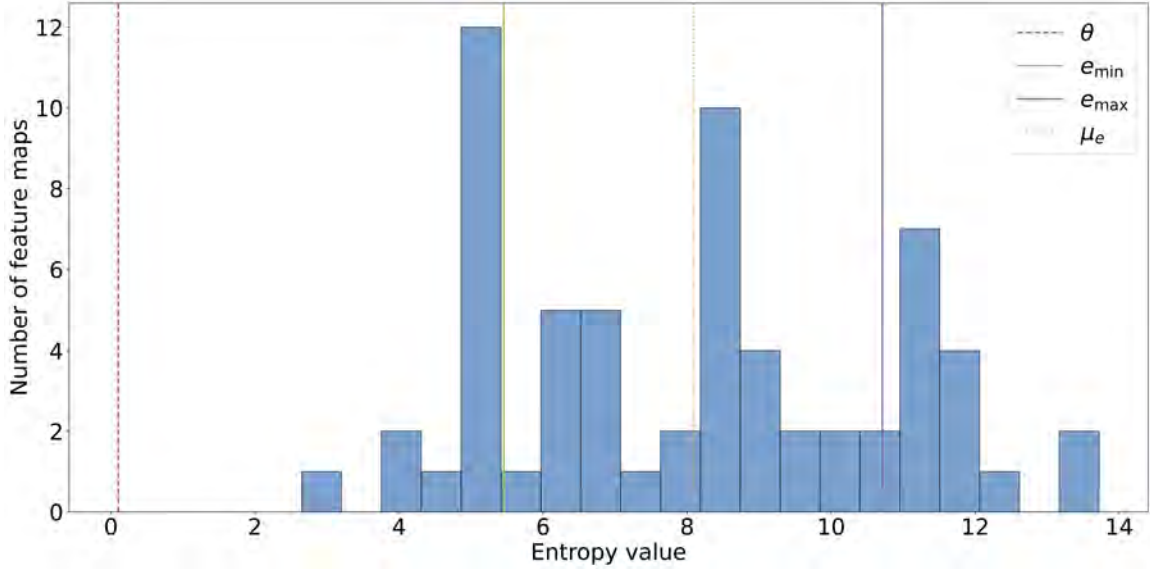
Với block b , tập bản đồ hợp lệ được xác định bởi hai ngưỡng $E_{\min}^{(b)}$ và $E_{\max}^{(b)}$:

$$E^{(b)} = \left\{ H(F_m^{(b)}(x_i)) \mid H(F_m^{(b)}(x_i)) > \theta^{(b)}, i = 1, 2, \dots, N, m = 1, 2, \dots, M_b \right\}, \quad (2.23)$$

Điều kiện chọn một bản đồ là:

$$e_{\min}^{(b)} \leq H(F_m^{(b)}(x)) \leq e_{\max}^{(b)}. \quad (2.26)$$

Các ngưỡng không dùng chung cho mọi block vì kích thước không gian, số kênh và phân bố kích hoạt thay đổi theo độ sâu. Histogram entropy cung cấp quan sát định lượng để loại phản hồi gần hằng và phản hồi quá phân tán.



Hình 2.8: Phân bố entropy của các bản đồ đặc trưng tại lớp `block1_conv1` và các ngưỡng dùng để xác định khoảng entropy phù hợp.

Khoảng entropy được xác định từ phân bố thực nghiệm của từng block. Điều này phù hợp với thực tế số kênh, kích thước không gian và độ thưa hoạt hóa thay đổi theo độ sâu mạng. Một bản đồ chỉ được giữ khi entropy nằm trong miền đã chọn; tập hợp thu được có thể khác nhau giữa block và giữa bộ dữ liệu. Cơ chế này làm giảm giả định rằng cùng một ngưỡng có thể đại diện cho mọi tầng và mọi miền ảnh.

Histogram được chuẩn hóa thành phân phối xác suất trước khi tính entropy, vì vậy phép đo không phụ thuộc trực tiếp vào số điểm ảnh của bản đồ. Các ngưỡng dưới và trên loại đồng thời phản hồi gần hằng và phản hồi quá phân tán. Việc khảo sát theo block còn cho phép nhận biết tầng nào đóng góp nhiều bản đồ hợp lệ nhất. Trong thực nghiệm, cấu hình ngưỡng được lựa chọn trên dữ liệu huấn luyện và giữ cố định khi đánh giá, bảo đảm kết quả kiểm tra không tham gia quyết định chọn đặc trưng.

2.4. Hợp nhất và cân bằng trong không gian đặc trưng

Sau chọn lọc, tập bản đồ hợp lệ được gom thành biểu diễn sâu. Vector cuối cùng của mẫu x_i là phép ghép nối biểu diễn sâu và đặc trưng thủ công:

$$z_i^{(b)} = \psi\left(z_i^{(d,b)}, z_i^{(h)}\right), \quad (2.28)$$

Để giảm ảnh hưởng mất cân bằng lớp, số lượng biểu diễn cần bổ sung cho lớp k được xác định theo lớp lớn nhất. Hệ số nhân và số mẫu tăng cường lần lượt là:

$$m_k^{(b)} = \left\lceil \frac{n_{\max}^{(b)}}{n_k^{(b)}} \right\rceil. \quad (2.29)$$

$$n_k^{(b)} = n_k^{(b)} \cdot m_k^{(b)}. \quad (2.30)$$

Việc cân bằng được thực hiện sau hợp nhất, trong không gian đặc trưng, nhằm tạo thêm biểu diễn cho lớp thiểu số mà không thay đổi ảnh gốc. Bộ phân loại tối ưu cross-entropy trên tập mở rộng. Luận án lưu ý rằng tác động riêng của cân bằng cần được đối chứng thêm với class-weighted loss, focal loss [115] và SMOTE [116].

Sau phép ghép, mỗi mẫu vẫn giữ một nhãn duy nhất và toàn bộ thao tác cân bằng được thực hiện trong tập huấn luyện. Tập kiểm tra không được dùng để xác định số mẫu bổ sung hoặc điều chỉnh tham số. Đây là điều kiện cần để tránh rò rỉ dữ liệu. Bộ phân loại tối ưu hàm mất mát cross-entropy; các lựa chọn như trọng số lớp, focal loss hay SMOTE được xem là đối chứng cần thiết khi phân tích riêng đóng góp của bước cân bằng.

Hệ số $m_k^{(b)}$ xác định mức thiếu hụt tương đối của lớp k so với lớp lớn nhất, còn $n_k^{(b)}$ biểu diễn số phần tử cần tạo thêm. Cách viết tách hai đại lượng giúp phân biệt tỷ lệ nhân với quy mô bổ sung thực tế. Các mẫu tổng hợp chỉ có vai trò điều chỉnh phân bố huấn luyện; số lượng mẫu thật và cơ cấu tập đánh giá được giữ nguyên để chỉ số cuối phản ánh dữ liệu quan sát ban đầu.

2.5. Dữ liệu và thiết lập thực nghiệm

VNCancer [119] gồm ba lớp BCC, MM và SCC, phản ánh dữ liệu thu nhận trong nước với khác biệt thiết bị và điều kiện chụp. HAM10000 [3] gồm bảy lớp, mất cân bằng mạnh, trong đó lớp NV chiếm tỷ lệ lớn. Hai bộ dữ liệu cho phép đánh giá phương pháp trong bối cảnh ba lớp và đa lớp.

Thực nghiệm dùng stratified five-fold cross-validation. Các cấu hình so sánh sử dụng cùng bộ phân loại và giao thức huấn luyện để chèn lệch chủ yếu phản ánh chất lượng biểu diễn. Chỉ số báo cáo gồm Accuracy, Macro Precision, Macro Recall, Macro F1 và Macro AUC; phân tích theo lớp sử dụng ROC và ma trận nhầm lẫn.

Bảng 2.2: Các lớp tổn thương da và phân bố số lượng trong bộ dữ liệu VNCancer

Ký hiệu	Tên tiếng Anh	Tên tiếng Việt	Số lượng
BCC	Basal Cell Carcinoma	Ung thư biểu mô tế bào đáy	2719
MM	Malignant Melanoma	U hắc tố ác tính	314
SCC	Squamous Cell Carcinoma	Ung thư biểu mô tế bào vảy	833
Tổng số ảnh			3866

VNCancer phản ánh điều kiện thu nhận gần với ảnh thông thường, còn HAM10000 chủ yếu gồm ảnh soi da và có bảy lớp với mất cân bằng rõ. Hai bộ dữ liệu tạo ra phép kiểm tra bổ sung về miền ảnh và số lớp. Tiền xử lý, chia tập và lựa chọn siêu tham số phải được thực hiện nhất quán trong từng giao thức; các chỉ số giữa hai bộ chỉ dùng để nhận diện xu hướng chứ không so sánh trực tiếp độ khó tuyệt đối.

2.6. Ảnh hưởng của chọn lọc entropy

Trên VNCancer, cấu hình tốt nhất không dùng entropy đạt Accuracy 80.97% và F1-score 82.11%. Khi áp dụng chọn lọc entropy, cấu hình block2_conv2 đạt Accuracy 88.75%, F1-score 88.16% và Specificity 93.42%. Mức tăng tương ứng là 7.78 và 6.05

Bảng 2.3: Các lớp tổn thương da và phân bố số lượng trong bộ dữ liệu HAM10000

Ký hiệu	Tên tiếng Anh	Tên tiếng Việt	Số lượng
MEL	Melanoma	U hắc tố	1113
NV	Melanocytic nevus	Nốt ruồi hắc tố	6705
BCC	Basal cell carcinoma	Ung thư biểu mô tế bào đáy	514
AKIEC	Actinic keratoses and intraepithelial carcinoma	Dày sừng quang hóa và ung thư biểu mô nội biểu bì	327
BKL	Benign keratosis-like lesions	Tổn thương dày sừng lành tính	1099
DF	Dermatofibroma	U xơ da	115
VASC	Vascular lesions	Tổn thương mạch máu	142
Tổng số ảnh			10015

Ghi chú: Dữ liệu dermoscopy, các lớp được sắp xếp theo số lượng giảm dần.

điểm phần trăm.

Ở các block sâu, cải thiện rõ hơn: block4_conv3 tăng Accuracy từ 49.73% lên 86.85%. Kết quả cho thấy tập phản hồi sau sàng lọc phù hợp hơn với bộ phân loại trong cấu hình đang xét, nhưng không chứng minh entropy riêng lẻ là độ đo trực tiếp của tính hữu ích.

Bảng 2.7: Kết quả trên bộ dữ liệu VNCancer khi sử dụng chọn lọc entropy

Block	Accuracy	Precision	F1-score	Specificity
block1_conv2	88.41	88.02	87.58	93.12
block2_conv2	88.75	88.40	88.16	93.42
block3_conv1	88.24	87.95	87.57	93.05
block4_conv3	86.85	86.44	86.04	92.63

Kết quả VNCancer cho thấy lựa chọn bản đồ ảnh hưởng đồng thời đến Accuracy, Precision và F1-score. Mức cải thiện lớn nhất xuất hiện ở cấu hình phù hợp, trong khi không phải mọi block đều tăng như nhau. Quan sát này ủng hộ việc chọn theo phân bố thay vì giữ toàn bộ phản hồi. Tuy nhiên, kết quả phải được hiểu là tác động của cả quy trình, vì đặc trưng thủ công, cân bằng và bộ phân loại cùng tham gia tạo ra dự đoán cuối.

Kiểm định chéo năm folds được dùng để giảm phụ thuộc vào một lần chia dữ liệu. Giá trị trung bình mô tả xu hướng chung, còn độ lệch chuẩn cho biết mức ổn định giữa các folds. Khi hai cấu hình có trung bình gần nhau, độ phân tán và kết quả trên lớp thiếu số cần được xem xét trước khi lựa chọn. Cách báo cáo này phù hợp hơn việc chỉ nêu cấu hình đạt điểm cực đại trong một lần chạy.

2.7. Kết quả hợp nhất đa nguồn

Sau khi chọn bản đồ, BlockNet hợp nhất biểu diễn sâu với đặc trưng thủ công. Trên VNCancer, cấu hình đầy đủ đạt kết quả cao hơn các biểu diễn đơn nguồn trong cùng giao thức. Kết quả xác nhận tính bổ trợ giữa phản hồi CNN và mô tả miền, đồng thời cho thấy việc tổ chức biểu diễn quan trọng hơn phép ghép nối đơn thuần.

Trên HAM10000, chọn lọc entropy cải thiện F1-score tốt nhất từ 63.78% lên 65.00%; mức tăng nhỏ hơn VNCancer do số lớp lớn hơn, phân bố mất cân bằng mạnh và tương đồng hình thái cao. Phân tích ROC ghi nhận AUC của VASC tăng từ 0.8681 lên 0.9179, DF từ 0.8294 lên 0.8440, AKIEC từ 0.7587 lên 0.7727 và NV từ 0.8409 lên 0.8465.

Bảng 2.10: So sánh hiệu năng trên bộ dữ liệu VNCancer sử dụng kiểm định chéo 5 folds. Kết quả được trình bày dưới dạng trung bình $\pm$ độ lệch chuẩn (%).

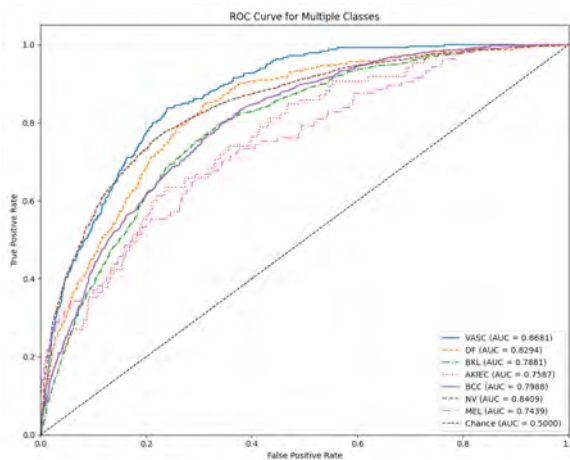
Mô hình	Accuracy	Precision	F1-score	Specificity
VGG16	92.19 $\pm$ 0.48	92.16 $\pm$ 0.49	92.15 $\pm$ 0.47	89.78 $\pm$ 1.54
ResNet101	93.59 $\pm$ 0.78	93.56 $\pm$ 0.71	93.53 $\pm$ 0.76	91.13 $\pm$ 0.44
EfficientNetB7	92.55 $\pm$ 0.54	92.65 $\pm$ 0.54	92.57 $\pm$ 0.54	91.15 $\pm$ 1.64
BlockNet-block1_conv1	95.58$\pm$0.69	95.59$\pm$0.69	95.59$\pm$0.69	97.53$\pm$0.37

Ở HAM10000, chênh lệch nhỏ hơn phản ánh bài toán nhiều lớp và mức mất cân bằng nghiêm trọng hơn. AUC theo lớp và ma trận nhầm lẫn cho thấy các lớp hiếm vẫn khó nhận diện dù Accuracy tổng thể cao. Do đó, hiệu quả của BlockNet không nên được mô tả bằng một chỉ số duy nhất. Sự cải thiện ổn định trên nhiều phép đo có giá trị hơn một mức tăng lớn nhưng chỉ xuất hiện ở một cấu hình hoặc một lớp.

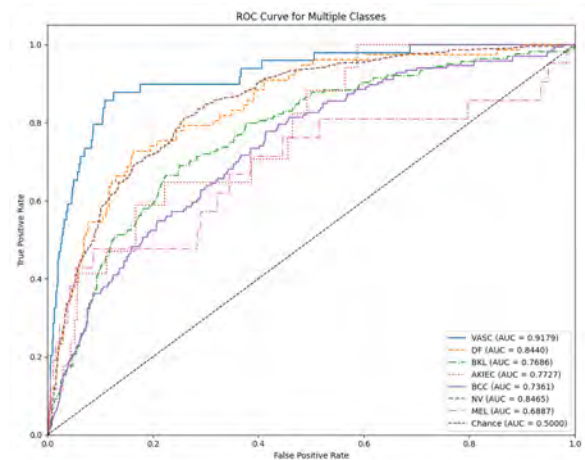
Đường ROC trước và sau chọn lọc cho thấy thay đổi khả năng xếp hạng mẫu ở nhiều ngưỡng quyết định. Với dữ liệu nhiều lớp, AUC được xem xét theo từng lớp và ở mức tổng hợp; điều này giúp phát hiện trường hợp chỉ số chung che khuất lớp hiếm. Các kết quả không thuận lợi vẫn được giữ trong phân tích để xác định giới hạn của phương pháp và tránh lựa chọn báo cáo theo kết quả tốt nhất.

2.8. Thảo luận kết quả BlockNet

Thứ nhất, chọn lọc entropy làm thay đổi rõ chất lượng biểu diễn, đặc biệt tại các block sâu. Thứ hai, hiệu quả không đồng đều giữa hai bộ dữ liệu, cho thấy khoảng entropy phải gắn với block và phân bố dữ liệu. Thứ ba, đặc trưng thủ công vẫn hỗ trợ cho đặc trưng CNN trong các trường hợp màu sắc, kết cấu và hình dạng có ý nghĩa miền. Thứ tư, cân bằng trong không gian đặc trưng giúp giảm mức chi phối của lớp lớn nhưng chưa được tách hoàn toàn khỏi các thành phần khác bằng thí nghiệm loại trừ.



(a) Không sử dụng entropy, block2_conv2



(b) Sử dụng entropy, block2_conv1

Hình 2.15: So sánh ROC trên bộ dữ liệu HAM10000 trước và sau khi sử dụng chọn lọc entropy.

Kết luận Chương 2

BlockNet hình thành quy trình chọn lọc, hợp nhất và cân bằng biểu diễn. Kết quả trên VNCancer và HAM10000 cung cấp bằng chứng cho giả thuyết rằng không cần sử dụng toàn bộ bản đồ đặc trưng để đạt biểu diễn tốt hơn.

Phân tích tổng hợp cho thấy chọn lọc entropy đóng vai trò như một cơ chế điều chuẩn biểu diễn có thể diễn giải được. Số bản đồ giảm không đồng nghĩa mất thông tin nếu các phản hồi bị loại chủ yếu là đồng nhất hoặc phân tán kém hữu ích. Hạn chế còn lại là khoảng entropy mang tính thực nghiệm và chưa được tối ưu chung với hàm mất mát. Đây là cơ sở để nghiên cứu cơ chế chọn mềm hoặc học ngưỡng trong các mô hình tiếp theo.

Về khả năng tái lập, mỗi cấu hình cần lưu backbone, block, khoảng entropy, số bản đồ được giữ, phương pháp chuẩn hóa, bộ phân loại và hạt giống ngẫu nhiên. Việc báo cáo đầy đủ các thành phần này cho phép phân biệt cải thiện do chọn lọc với cải thiện do thay đổi thiết lập huấn luyện. Đây cũng là điều kiện để mở rộng BlockNet sang backbone khác hoặc miền ảnh khác mà không mặc định giữ nguyên ngưỡng đã tìm trên dữ liệu hiện tại.

Phân tích bổ sung về thiết kế và kết quả BlockNet

BlockNet được thiết kế theo hướng mô-đun để có thể quan sát tác động của từng block. Với mỗi ảnh, VGG16 tạo nhiều bản đồ ở các độ sâu khác nhau. Histogram của giá trị kích hoạt được chuẩn hóa, sau đó entropy được tính cho từng kênh. Tập bản đồ hợp lệ không cố định về số lượng mà phụ thuộc phân bố entropy của block. Cơ chế này tránh giả định rằng toàn bộ kênh của backbone đều đóng góp ngang nhau cho bài toán tổn thương da.

Sau chọn lọc, phép vector hóa chuyển phản hồi không gian thành biểu diễn có thể hợp nhất. Đặc trưng HSV hỗ trợ mô tả khác biệt màu; GLCM phản ánh quan hệ mức xám và kết cấu; fractal bổ sung thông tin về độ phức tạp. Các nhóm được chuẩn hóa trước khi ghép để hạn chế sai lệch thang đo. Cân bằng lớp được thực hiện trên biểu diễn hợp nhất của tập huấn luyện, không tác động vào tập kiểm tra và không tạo ảnh tổn thương nhân tạo.

Kết quả trên VNCancer cho thấy cấu hình chọn lọc phù hợp có thể tăng đồng thời Accuracy và F1-score. Điều này đáng chú ý vì F1-score phản ánh sự cân bằng giữa Precision và Recall, phù hợp hơn Accuracy khi số mẫu giữa các lớp khác nhau. Tuy nhiên, độ cải thiện không đồng đều ở mọi block. Các tầng nông dễ giữ phản hồi màu và biên, còn tầng sâu mang thông tin ngữ nghĩa nhưng có thể chứa kênh chuyên biệt không phù hợp với dữ liệu đích.

Trên HAM10000, bài toán bảy lớp làm rõ giới hạn của chỉ số tổng hợp. Lớp NV chiếm ưu thế có thể làm Accuracy cao trong khi các lớp hiếm vẫn có Recall thấp. Vì vậy, luận án đối chiếu thêm AUC theo lớp, F1-score và ma trận nhầm lẫn. Đường ROC trước và sau chọn lọc cho phép quan sát thay đổi khả năng xếp hạng ở nhiều ngưỡng, thay vì phụ thuộc một ngưỡng quyết định duy nhất.

Việc so sánh cần giữ chung cách chia dữ liệu, tiền xử lý, bộ phân loại và lịch huấn luyện. Nếu đồng thời thay backbone hoặc siêu tham số, không thể quy chênh lệch cho entropy. Do đó, kết luận phù hợp là toàn bộ quy trình BlockNet tạo biểu diễn hiệu quả hơn trong các cấu hình đã kiểm chứng; đóng góp riêng của từng thành phần cần tiếp tục

được lượng hóa bằng ablation có kiểm soát.

Khả năng mở rộng của phương pháp nằm ở việc tiêu chuẩn entropy không phụ thuộc một lớp bệnh cụ thể. Dẫu vậy, khoảng entropy không nên chuyển nguyên trạng sang backbone hoặc miền ảnh mới. Mỗi thay đổi về kích thước bản đồ, hàm kích hoạt và điều kiện thu nhận đều có thể làm dịch chuyển phân bố. Quy trình tái hiệu chỉnh trên tập huấn luyện và đánh giá ngoại kiểm là yêu cầu bắt buộc trước khi áp dụng rộng hơn.

Một hướng hoàn thiện trực tiếp là xây dựng thí nghiệm loại trừ theo ma trận thành phần: backbone cơ sở, chọn entropy, đặc trưng thủ công và cân bằng biểu diễn. Mỗi cấu hình cần được lập với cùng hạt giống và folds, sau đó báo cáo chênh lệch cùng khoảng tin cậy. Thiết kế này sẽ xác định phần cải thiện do từng thành phần, kiểm tra tương tác giữa chọn lọc với cân bằng và làm rõ trường hợp entropy có lợi nhất cho lớp hiếm.

Ý nghĩa của kết quả đối với lựa chọn biểu diễn

Kết quả thực nghiệm cho thấy số lượng bản đồ được giữ lại không tỷ lệ tuyến tính với hiệu năng. Một tập biểu diễn nhỏ hơn vẫn có thể đạt kết quả tốt hơn nếu giảm được phản hồi đồng nhất, nhiễu hoặc dư thừa. Vì vậy, mục tiêu của BlockNet không phải nén đặc trưng bằng mọi giá mà là duy trì cân bằng giữa lượng thông tin và khả năng phân biệt. Sự khác nhau giữa các block cũng chỉ ra rằng tầng nông và tầng sâu cần tiêu chuẩn lựa chọn riêng, do chúng mã hóa các mức đặc trưng khác nhau.

Khi chuyển sang dữ liệu mới, quy trình nên bắt đầu bằng khảo sát histogram entropy trên tập huấn luyện, xác định khoảng ứng viên rồi đánh giá bằng kiểm chứng chéo. Không được chọn khoảng dựa trên tập kiểm tra. Cùng với hiệu năng, cần báo cáo số bản đồ được giữ, mức giảm chiều biểu diễn và thời gian tính toán. Ba nhóm thông tin này giúp đánh giá liệu cải thiện có đi kèm giảm chi phí hay chỉ chuyển độ phức tạp sang bước chọn lọc.

CHƯƠNG 3: HỌC ĐA NGUỒN BÁN GIÁM SÁT DỰA TRÊN ĐỒ THỊ

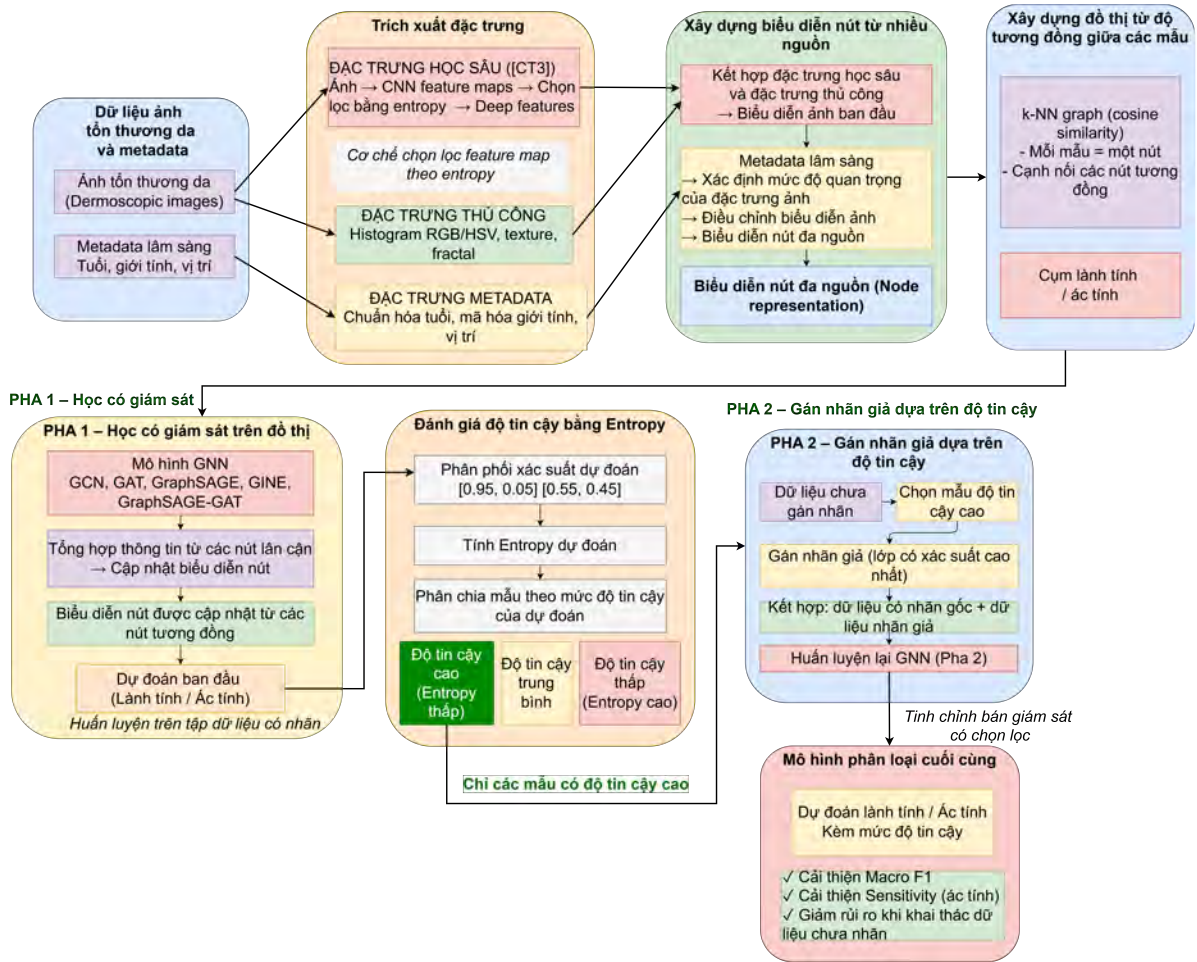
3.1. Bài toán và mô hình tổng quan

Tập có nhãn D_L và tập chưa gán nhãn D_U gồm ảnh cùng metadata lâm sàng. Mục tiêu là học biểu diễn hợp nhất, xây dựng đồ thị liên mẫu và khai thác có kiểm soát D_U . Biểu diễn tổng quát của mẫu là:

$$\mathbf{x}_i^{img} = [\mathbf{x}_i^{deep}; \mathbf{x}_i^{hand}], \quad (3.4)$$

$$\mathbf{h}_i = \Phi(\mathbf{x}_i^{img}, \mathbf{x}_i^{meta}).$$

Đồ thị $G = (V, E)$ được xây dựng từ biểu diễn hợp nhất; mỗi nút là một mẫu, cạnh nối các mẫu gần nhau theo cosine. Quy trình gồm pha 1 huấn luyện GNN có giám sát và pha 2 tính entropy, chọn nhãn giả tin cậy cao rồi tái huấn luyện.



Hình 3.4: Tổng quan mô hình đề xuất.

Khác với Chương 2, mỗi mẫu ở Chương 3 không được phân loại độc lập mà trở thành một nút trong đồ thị. Biểu diễn nút phải kết hợp dấu hiệu ảnh, đặc trưng thủ công và metadata trước khi xác lập láng giềng. Thiết kế này cho phép thông tin từ các mẫu tương đồng hỗ trợ nhau, nhưng đồng thời đặt ra yêu cầu kiểm soát cạnh sai và sự lan

truyền của dự đoán thiếu tin cậy.

Quy trình gồm hai pha có ranh giới rõ. Pha thứ nhất chỉ dùng tập có nhãn để học bộ mã hóa, phép hợp nhất và bộ phân loại đồ thị. Mô hình tốt nhất của pha này được dùng suy luận trên tập chưa gán nhãn. Pha thứ hai tính entropy từ phân phối dự đoán, phân tầng mẫu theo độ tin cậy và chỉ bổ sung nhóm High Trust vào tập huấn luyện. Tập kiểm tra được giữ độc lập trong cả hai pha. Cách tổ chức này giúp đánh giá mức cải thiện do dữ liệu bổ sung mà không đưa thông tin kiểm tra vào quá trình lựa chọn nhãn giả.

3.2. Hợp nhất ảnh và metadata có chú ý

Đặc trưng học sâu được lấy từ vector sau pooling của VGG19, DenseNet121 và MobileNet [50, 52, 121]. Đặc trưng thủ công gồm histogram RGB, histogram HSV và fractal. Hai nhóm thị giác được ghép thành z_i^{img} ; metadata gồm tuổi, giới tính và vị trí tổn thương được chuẩn hóa và mã hóa.

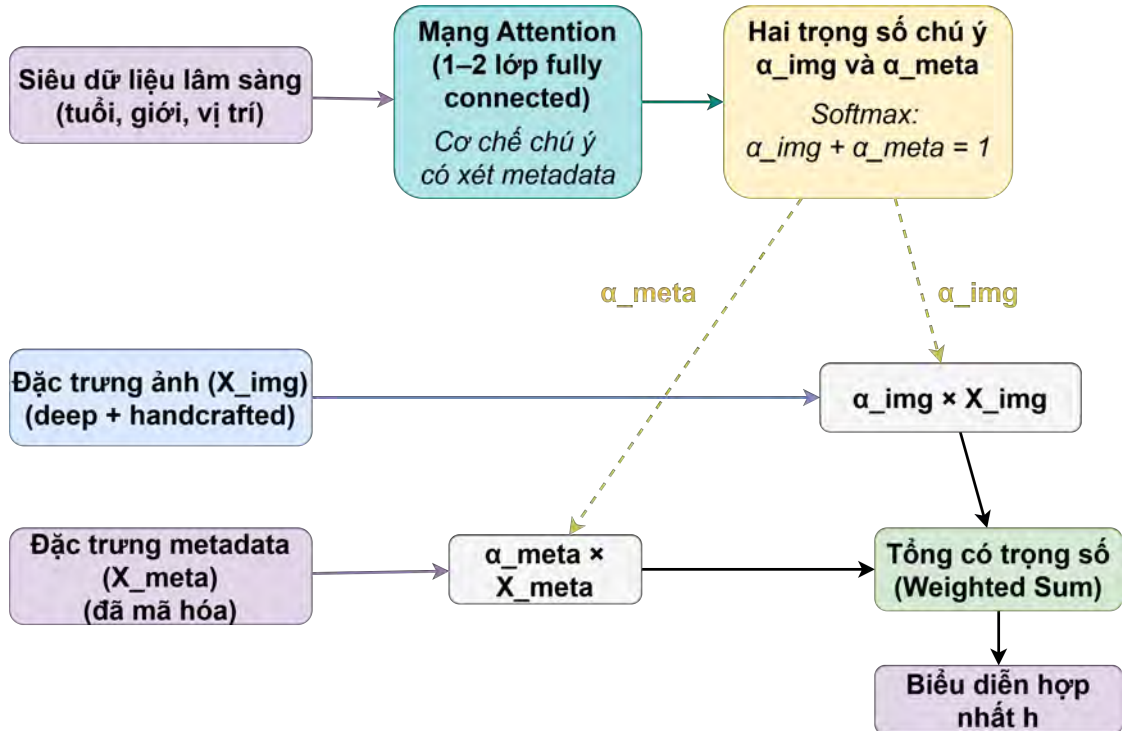
Hai trọng số chú ý cho đường ảnh và metadata được tính bởi:

$$\left[\alpha_i^{img}, \alpha_i^{meta} \right] = \text{softmax} \left(W_a x_i^{meta} + b_a \right), \quad (3.14)$$

Biểu diễn có xét metadata là:

$$h_i = \alpha_i^{img} x_i^{img} + \alpha_i^{meta} x_i^{meta}, \quad (3.16)$$

Cơ chế này không sinh ba trọng số độc lập cho đặc trưng sâu, thủ công và metadata; đặc trưng sâu và thủ công đã được ghép thành một đường ảnh trước khi kết hợp metadata.



Hình 3.6: Hợp nhất đặc trưng đa nguồn theo cơ chế chú ý có xét metadata.

Cơ chế chú ý hai đường học trọng số tương đối cho nhánh ảnh và nhánh metadata. Metadata không chỉ được ghép ở cuối mà tham gia điều chỉnh biểu diễn hợp nhất. Khi trường dữ liệu bị thiếu hoặc có độ tin cậy thấp, trọng số học được có thể giảm đóng góp của nguồn tương ứng. Tuy nhiên, chú ý không tự động đồng nghĩa với giải thích nhân

quả; nó chỉ phản ánh mức đóng góp trong phép hợp nhất của mô hình.

Các vector từ nhiều backbone được đưa về miền kích thước tương thích trước khi hợp nhất. Việc dùng nhiều kiến trúc nhằm thu nhận các thiên hướng biểu diễn khác nhau, không phải cộng trực tiếp mọi đầu ra chưa chuẩn hóa. Metadata được mã hóa theo loại biến và ghép sau bước xử lý giá trị thiếu. Nhờ đó, trọng số chú ý được học trên biểu diễn đã chuẩn hóa và có thể thay đổi theo từng mẫu thay vì cố định cho toàn bộ tập dữ liệu.

3.3. Xây dựng đồ thị và huấn luyện GNN

Tập láng giềng k -NN của nút i được xác định từ độ tương đồng cosine:

$$\mathcal{N}_5(v_i) = \text{TopK}_{j \neq i}(s_{ij}, 5), \quad (3.18)$$

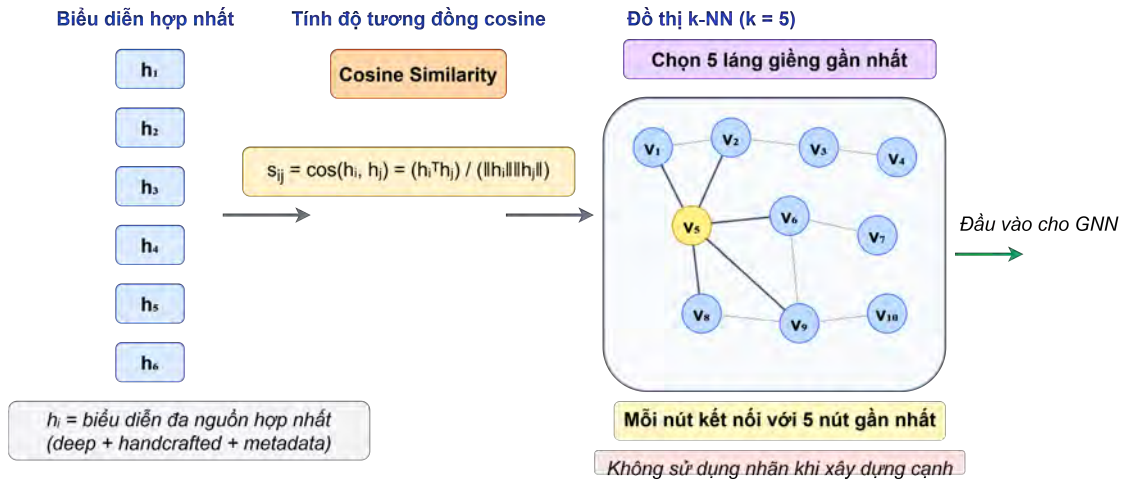
Tập cạnh được tạo bởi:

$$E = \{(v_i, v_j) \mid v_j \in \mathcal{N}_5(v_i)\}, \quad (3.19)$$

Ở mỗi tầng, nút tổng hợp thông điệp từ lân cận và cập nhật biểu diễn:

$$\mathbf{h}_i^{(k+1)} = \text{UPDATE}^{(k)}\left(\mathbf{h}_i^{(k)}, \text{AGGREGATE}^{(k)}\left(\left\{\mathbf{h}_j^{(k)} : v_j \in \mathcal{N}_5(v_i)\right\}\right)\right), \quad (3.20)$$

Các kiến trúc GCN, GAT, GraphSAGE, GINE và GraphSAGE–GAT được khảo sát. Pha 1 tối ưu cross-entropy trên các nút có nhãn; chất lượng phụ thuộc đồng thời vào biểu diễn nút, độ đo tương đồng và k .



Hình 3.8: Đồ thị k -NN với $k=5$ được xây dựng từ biểu diễn hợp nhất dựa trên độ tương đồng cosine.

Đồ thị k -NN được tạo từ biểu diễn hợp nhất và độ tương đồng cosine. Giá trị k quá nhỏ có thể làm đồ thị phân mảnh; giá trị quá lớn có thể nối các vùng khác lớp và gây làm trơn quá mức. Quá trình truyền thông điệp tổng hợp thông tin láng giềng qua nhiều tầng, sau đó bộ phân loại nút tạo phân phối xác suất. Vì vậy, lựa chọn k , chiều biểu diễn và độ sâu GNN cần được xem xét đồng thời.

Tập cạnh được xây dựng mà không sử dụng nhãn của nút đích, qua đó hạn chế rò rỉ thông tin. Khi quan hệ láng giềng không đối xứng, có thể đối xứng hóa cạnh để bảo đảm truyền thông điệp theo hai chiều. Các biến thể GCN, GAT, GraphSAGE và GINE [20, 21, 22, 91] được đối chiếu nhằm đánh giá cơ chế tổng hợp; sự khác biệt giữa chúng phản ánh cách chuẩn hóa, lấy mẫu hoặc gán trọng số cho thông tin lân cận.

3.4. Gán nhãn giả dựa trên độ tin cậy

Với xác suất dự đoán p_{jc} của mẫu chưa gán nhãn x_j , entropy dự đoán là:

$$H(\hat{\mathbf{y}}_i) = - \sum_{c=1}^C \hat{y}_{i,c} \log(\hat{y}_{i,c}), \quad (3.24)$$

Hai ngưỡng α_1, α_2 chia mẫu thành High, Medium và Low Trust:

$$T(\hat{\mathbf{y}}_i) = \begin{cases} \text{High}, & H(\hat{\mathbf{y}}_i) \leq \alpha_1, \\ \text{Medium}, & \alpha_1 < H(\hat{\mathbf{y}}_i) \leq \alpha_2, \\ \text{Low}, & H(\hat{\mathbf{y}}_i) > \alpha_2. \end{cases} \quad (3.26)$$

Chỉ mẫu High Trust thỏa điều kiện entropy được gán nhãn giả:

$$\tilde{y}_j = \arg \max_c \hat{y}_{j,c}, \quad \text{if } H(\hat{\mathbf{y}}_j) \leq \alpha_1. \quad (3.28)$$

Tập huấn luyện mở rộng là:

$$D_{\text{new}} = D_L \cup D_{\text{pseudo}}^{\text{High}}, \quad (3.29)$$

Cách chọn lọc này giảm nguy cơ đưa dự đoán kém chắc chắn vào tái huấn luyện nhưng không bảo đảm mọi nhãn High Trust đều đúng.

Entropy dự đoán đo mức phân tán của xác suất giữa các lớp. Mẫu có entropy thấp được xếp High Trust và mới đủ điều kiện nhận nhãn giả; nhóm Medium và Low Trust được giữ ngoài tập tái huấn luyện. Ngưỡng lựa chọn giúp đánh đổi giữa số lượng và chất lượng nhãn giả. Nếu ngưỡng quá lỏng, sai số được khuếch đại; nếu quá chặt, lượng dữ liệu bổ sung không đủ tạo thay đổi đáng kể.

Nhãn giả được lấy từ lớp có xác suất cực đại nhưng chỉ được chấp nhận sau điều kiện entropy. Hai tiêu chuẩn này bổ sung nhau: argmax xác định lớp, còn entropy đánh giá mức tập trung của toàn bộ phân phối. Một xác suất cực đại tương đối cao vẫn có thể đi cùng phân phối cạnh tranh giữa nhiều lớp; entropy giúp nhận diện tình huống đó tốt hơn việc chỉ dùng một ngưỡng xác suất duy nhất.

3.5. Thiết lập và kết quả pha có giám sát

ISIC 2020 được tổ chức thành tập có nhãn và tập chưa gán nhãn; bài toán chính là phân loại lành tính và ác tính trong điều kiện mất cân bằng. Metadata thiếu được xử lý, biến liên tục chuẩn hóa z-score, biến phân loại one-hot. Các nhóm đặc trưng được lựa chọn bằng Chi-square trong chính thiết lập Chương 3.

Kết quả pha 1 cho thấy các mô hình học máy truyền thống có Specificity cao nhưng Sensitivity thấp ở một số cấu hình. Các GNN tích hợp metadata thay đổi cân bằng giữa hai chỉ số; hiệu quả phụ thuộc cấu trúc đồ thị. Vì vậy, đánh giá sử dụng Macro F1, Macro Recall, Macro AUC, Sensitivity và Specificity thay vì chỉ Accuracy.

Pha có giám sát tạo đường cơ sở và cung cấp mô hình sinh nhãn giả. So sánh giữa mô hình học máy truyền thống và các kiến trúc GNN cho thấy giá trị của quan hệ liên mẫu khi biểu diễn nút đủ tốt. Các chỉ số Sensitivity, Specificity, Macro F1 và Macro AUC được báo cáo cùng nhau nhằm phản ánh cả khả năng phát hiện lớp dương và kiểm soát dự đoán sai trong dữ liệu mất cân bằng.

Bảng 3.10: So sánh kết quả pha huấn luyện có giám sát giữa các mô hình học máy truyền thống và các kiến trúc nền GNN trên bộ dữ liệu ISIC 2020.

Mô hình	Sensitivity	Specificity	Macro F1	Macro AUC
Logistic Regression	53.8±2.9	99.1±0.1	75.6±1.2	92.6±1.1
XGBoost	51.3±3.7	99.7±0.1	80.6±1.6	96.0±0.8
GCN	56.7±20.4	70.8±8.4	44.2±2.5	71.2±3.6
GINE	76.0±22.7	61.2±13.7	40.8±5.0	75.9±5.2
GraphSAGE-GAT	70.2±17.7	65.8±6.9	42.8±2.1	74.8±5.7

Các mô hình truyền thống sử dụng cùng biểu diễn đầu vào để tạo đối chứng công bằng ở mức bộ phân loại. Với GNN, cấu trúc đồ thị bổ sung thông tin lân cận nhưng cũng làm tăng phụ thuộc giữa mẫu; vì vậy, phép chia dữ liệu và quá trình dựng đồ thị phải được kiểm soát. Kết quả pha 1 là mốc tham chiếu bắt buộc để xác định liệu pha nhãn giả có tạo cải thiện nhất quán hay chỉ thay đổi do dao động huấn luyện.

3.6. Kết quả pha gán nhãn giả

Sau pha 2, GINE cải thiện một số chỉ số cân bằng lớp, nổi bật ở Macro F1, Macro Recall và Macro AUC. Kiểm định cặp trên 15 lần kiểm chứng được dùng để đánh giá mức thay đổi trước và sau tái huấn luyện. Kết quả phải được hiểu trên phần dữ liệu có nhãn; ISIC 2020 Test không có nhãn phân loại công khai nên nhãn giả không phải nhãn thật.

Bảng 3.11: So sánh kết quả của mô hình GINE trước và sau pha gán nhãn giả dựa trên độ tin cậy.

Chỉ số	Pha 1 (%)	Pha 2 (%)
Accuracy	78.5±1.8	89.2±1.1
Macro AUC	72.6±2.1	88.5±1.4
Macro F1	43.0±1.9	60.9±1.7
Macro Recall	65.0±2.3	82.1±1.6
Precision	97.4±0.6	95.7±0.7
Recall	67.2±2.5	85.1±1.8

Sau khi bổ sung nhóm High Trust, mô hình được tái huấn luyện trên tập mở rộng nhưng vẫn đánh giá trên tập kiểm tra độc lập. Mức thay đổi giữa hai pha cho biết lợi ích thực tế của dữ liệu chưa gán nhãn. Việc chỉ chọn khoảng một phần ba số mẫu cho thấy phương pháp ưu tiên chất lượng hơn số lượng. Các nhãn giả không được dùng để báo cáo như nhãn thật và không thay thế quy trình gán nhãn chuyên gia.

Sự gia tăng của Macro F1 và Macro Recall quan trọng hơn riêng mức tăng Accuracy vì hai chỉ số này cho trọng số cân bằng hơn giữa các lớp. Precision giảm nhẹ trong một số cấu hình cho thấy tái huấn luyện tạo ra sự đánh đổi: mô hình phát hiện thêm mẫu dương nhưng cũng có thể phát sinh dự đoán dương sai. Do đó, ngưỡng entropy cần được lựa chọn trên tập xác thực và báo cáo cùng khoảng biến thiên qua các lần kiểm chứng. Kết quả pha 2 chỉ được xem là bằng chứng hỗ trợ cho cơ chế chọn mẫu, không phải chứng cứ rằng toàn bộ dữ liệu chưa gán nhãn đều có thể được tự động chuyển thành dữ liệu huấn luyện.

3.7. Phân tích độ tin cậy và thảo luận

Trong 10,982 mẫu chưa gán nhãn, High Trust có 3,623 mẫu, Medium Trust 3,621 và Low Trust 3,738. Chỉ khoảng 33% dữ liệu chưa gán nhãn được đưa vào tái huấn luyện; khoảng hai phần ba bị loại để đổi quy mô lấy độ tin cậy.

Bảng 3.13: Phân bố số lượng mẫu theo mức độ tin cậy trên tập chưa gán nhãn D_U

Mức tin cậy	Số mẫu	Tỷ lệ	Sử dụng
High Trust	3,623	33.0%	Có
Medium Trust	3,621	33.0%	Không
Low Trust	3,738	34.0%	Không

Entropy chỉ phản ánh mức chắc chắn tương đối của mô hình, chưa thay thế ECE, Brier score hoặc calibration curve. Hạn chế còn lại gồm metadata thiếu, cạnh đồ thị có thể nổi mẫu khác ngữ nghĩa, chưa khảo sát đầy đủ backbone hiện đại và chưa ngoại kiểm đa trung tâm.

Kết luận Chương 3

Khung đề xuất kết hợp biểu diễn đa nguồn, quan hệ đồ thị và nhãn giả được lọc bằng entropy. Kết quả khẳng định tiềm năng của tổ chức dữ liệu ở cấp mẫu nhưng không phải bằng chứng về hiệu quả chẩn đoán lâm sàng.

Phân bố độ tin cậy cung cấp thêm thông tin ngoài Accuracy. Tỷ lệ lớn mẫu Medium hoặc Low Trust có thể chỉ ra vùng giao nhau giữa lớp, thiếu metadata, sai lệch miền hoặc biểu diễn chưa đủ phân biệt. Vì vậy, entropy vừa là bộ lọc vừa là công cụ phân tích lỗi. Trong triển khai tương lai, cần kết hợp với hiệu chỉnh xác suất để xác nhận rằng mức tin cậy dự đoán tương ứng với xác suất đúng quan sát được.

Ba mức tin cậy cũng tạo một cơ chế ưu tiên cho gán nhãn chuyên gia. Nhóm Low Trust có thể được chuyển vào hàng đợi xem xét vì chứa nhiều trường hợp khó hoặc ngoài phân bố; nhóm Medium Trust phù hợp cho học chủ động; nhóm High Trust phục vụ mở rộng huấn luyện có kiểm soát. Cách sử dụng này khai thác toàn bộ tập chưa gán nhãn mà không buộc mọi mẫu phải nhận nhãn giả.

3.8. So sánh với phương pháp liên quan và phạm vi áp dụng

Mô hình đề xuất đạt Accuracy 89.2% trong giao thức thực nghiệm của luận án, cạnh tranh với các phương pháp được đối chiếu. Tuy nhiên, khác biệt về cách chia dữ liệu, tiền xử lý, backbone và chỉ số báo cáo không cho phép kết luận ưu thế tuyệt đối. Giá trị chính của phương pháp là kết hợp ba thành phần trong một quy trình: biểu diễn đa nguồn, đồ thị liên mẫu và nhãn giả có kiểm soát entropy.

Phương pháp mới được kiểm chứng trên dữ liệu công khai, chưa ngoại kiểm đa trung tâm và chưa đánh giá lâm sàng. Vì vậy, mô hình chỉ có ý nghĩa nghiên cứu và hỗ trợ sàng lọc tiềm năng, chưa phải công cụ chẩn đoán độc lập.

So sánh với nghiên cứu liên quan được thực hiện thận trọng vì mỗi công trình có thể dùng cách chia dữ liệu và tiền xử lý khác nhau. Kết quả của luận án chứng minh tính khả thi của chuỗi hợp nhất đa nguồn, học đồ thị và chọn nhãn giả, nhưng chưa chứng minh ưu thế tuyệt đối ngoài giao thức đã dùng. Ngoại kiểm đa trung tâm và đánh giá dưới dịch chuyển phân bố là bước cần thiết trước khi xem xét ứng dụng hỗ trợ lâm sàng.

Bảng 3.14: So sánh mô hình đề xuất với một số phương pháp liên quan trên ISIC 2020 theo chỉ số Accuracy

Phương pháp	Hướng tiếp cận	Accuracy (%)
Gupta và cộng sự	Knowledge distillation	87.3
Alhudhaif và cộng sự	Mạng CNN sâu	85.5
Choudhary và cộng sự	Mask-RCNN kết hợp GRU	88.0
Rana và cộng sự	Mô hình CNN xếp chồng	86.7
Mô hình đề xuất	Đa nguồn + đồ thị + nhãn giả tin cậy	89.2

Phân tích bổ sung về học đồ thị và nhãn giả

Biểu diễn nút quyết định trực tiếp chất lượng đồ thị. Các backbone cung cấp những mô tả thị giác bổ trợ, trong khi đặc trưng thủ công và metadata giúp phân biệt các mẫu có hình thái gần nhau. Sau chuẩn hóa và hợp nhất có chú ý, độ tương đồng cosine được dùng để xác định láng giềng. Nếu dựng đồ thị trước hợp nhất, quan hệ có thể chỉ phản ánh ảnh mà bỏ qua tuổi, giới tính hoặc vị trí tổn thương.

Giá trị k điều khiển mức kết nối. Đồ thị quá thưa làm hạn chế truyền thông tin; đồ thị quá dày tăng cạnh khác lớp và nguy cơ làm trơn biểu diễn. Các kiến trúc GCN, GAT, GraphSAGE và GINE thực hiện tổng hợp lân cận theo cơ chế khác nhau, nên được so sánh trong cùng giao thức. Kết quả pha có giám sát xác lập mô hình nền trước khi bất kỳ mẫu chưa gán nhãn nào được đưa vào huấn luyện.

Ở pha thứ hai, entropy được tính trên toàn bộ phân phối xác suất. So với chỉ dùng xác suất cực đại, entropy còn phản ánh mức cạnh tranh giữa các lớp còn lại. Nhóm High Trust có phân phối tập trung và được gán nhãn theo argmax; hai nhóm còn lại không tham gia tái huấn luyện. Cơ chế này giảm lượng dữ liệu sử dụng nhưng hạn chế rủi ro confirmation bias, trong đó dự đoán sai ban đầu được học lại như sự thật.

Mức tăng Macro F1, Macro Recall và Macro AUC cho thấy lợi ích không chỉ đến từ lớp chiếm đa số. Dù vậy, Precision có thể giảm nhẹ khi mô hình phát hiện thêm mẫu dương. Sự đánh đổi này cần được diễn giải cùng độ lệch chuẩn qua các lần kiểm chứng. Nhãn giả chỉ là công cụ huấn luyện và không được dùng như nhãn tham chiếu khi tính chỉ số trên tập kiểm tra.

Phân tầng độ tin cậy còn hỗ trợ phân tích dữ liệu. Mẫu Low Trust có thể là trường hợp khó, thiếu metadata hoặc nằm ngoài phân bố; mẫu Medium Trust phù hợp cho quy trình học chủ động có chuyên gia. Trong nghiên cứu tiếp theo, entropy cần kết hợp với ECE, Brier score, temperature scaling hoặc ensemble để đánh giá cả độ sắc nét lẫn độ hiệu chỉnh xác suất.

Giới hạn quan trọng là đồ thị được xây dựng từ biểu diễn học trên một miền dữ liệu cụ thể. Khi chuyển cơ sở y tế hoặc thiết bị thu nhận, cấu trúc láng giềng có thể thay đổi dù bộ phân loại vẫn cho xác suất cao. Ngoại kiểm đa trung tâm, kiểm tra dịch chuyển phân bố và phân tích sai số theo nhóm bệnh nhân là các bước cần thiết trước khi xem xét tích hợp vào quy trình hỗ trợ sàng lọc.

Ngoài hiệu năng dự đoán, nghiên cứu tương lai cần đo độ ổn định của cạnh khi backbone, phép tăng cường hoặc metadata thay đổi. Tỷ lệ láng giềng được bảo toàn và

mức đồng nhất nhân trong lân cận có thể dùng để chẩn đoán chất lượng đồ thị. Kết hợp các phép đo này với độ tin cậy dự đoán sẽ giúp phân biệt mẫu khó do bản thân ảnh với mẫu khó do quan hệ đồ thị không phù hợp.

Ý nghĩa của kết quả đối với khai thác dữ liệu chưa gán nhãn

Phân tầng entropy cho thấy dữ liệu chưa gán nhãn không phải một khối đồng nhất. Nhóm High Trust cung cấp tín hiệu đủ tập trung để mở rộng huấn luyện; nhóm Medium Trust chứa các trường hợp cần thêm bằng chứng; nhóm Low Trust có thể phản ánh mẫu khó hoặc dịch chuyển phân bố. Việc chỉ sử dụng nhóm High Trust làm giảm quy mô dữ liệu bổ sung nhưng tạo cơ chế kiểm soát rõ ràng hơn so với gán nhãn giả cho toàn bộ tập.

Để đánh giá đầy đủ, cần theo dõi tỷ lệ nhãn giả theo lớp, mức thay đổi ma trận nhầm lẫn và độ ổn định qua nhiều lần chạy. Nếu nhãn giả tập trung chủ yếu ở lớp đa số, hiệu năng tổng thể có thể tăng trong khi mất cân bằng trở nên nghiêm trọng hơn. Vì vậy, ngưỡng entropy có thể được hiệu chỉnh theo lớp hoặc kết hợp hạn ngạch nhằm ngăn một lớp chi phối tập mở rộng. Các biện pháp này cần được kiểm chứng độc lập trước khi đưa vào phiên bản phát triển tiếp theo.

Trong bối cảnh hỗ trợ lâm sàng, các mẫu không đạt ngưỡng không nên bị loại bỏ khỏi quy trình mà có thể được ưu tiên cho chuyên gia xem xét. Cách tổ chức này biến độ tin cậy thành cơ chế phân luồng: mẫu chắc chắn hỗ trợ tái huấn luyện, mẫu trung gian phục vụ học chủ động và mẫu khó được chuyển sang đánh giá chuyên sâu. Đây là hướng khai thác an toàn hơn so với coi mọi dự đoán tự động là nhãn thật.

KẾT LUẬN VÀ HƯỚNG NGHIÊN CỨU TƯƠNG LAI

Kết luận

Luận án đã giải quyết hai vấn đề liên kết trong phân loại ảnh tổn thương da. Block-Net tổ chức lại biểu diễn ở cấp bản đồ đặc trưng bằng chọn lọc entropy, hợp nhất với đặc trưng thủ công và cân bằng trong không gian đặc trưng. Khung thứ hai tổ chức dữ liệu ở cấp mẫu bằng hợp nhất ảnh với metadata, xây dựng đồ thị k -NN và khai thác dữ liệu chưa gán nhãn có kiểm soát.

Ở đóng góp thứ nhất, entropy kích hoạt được tính từ histogram của từng bản đồ và dùng để sàng lọc theo khoảng riêng của block. Các bản đồ được giữ lại được vector hóa, hợp nhất với HSV, GLCM và fractal. Trên VNCancer, Accuracy tốt nhất của thí nghiệm chọn lọc tăng từ 80.97% lên 88.75%; F1-score tăng từ 82.11% lên 88.16%. Trên HAM10000, mức tăng nhỏ hơn nhưng xuất hiện ở F1-score và AUC của một số lớp khó. Kết quả xác nhận giá trị của quy trình đầy đủ trong phạm vi giao thức thực nghiệm, không quy toàn bộ cải thiện cho riêng entropy.

Ở đóng góp thứ hai, biểu diễn ảnh từ VGG19, DenseNet121, MobileNet và đặc trưng thủ công được hợp nhất với metadata bằng chú ý hai đường. Đồ thị mô hình hóa quan hệ liên mẫu; entropy dự đoán lọc nhãn giả. Chỉ 3,623/10,982 mẫu High Trust được dùng để tái huấn luyện. Một số chỉ số Macro được cải thiện sau pha 2, nhưng nhãn giả không được xem là nhãn tham chiếu.

Hạn chế và hướng nghiên cứu tương lai

Các giới hạn chính gồm chưa tách hoàn toàn ảnh hưởng của từng thành phần Block-Net; chưa đối chứng đầy đủ các cơ chế mất cân bằng; metadata còn thiếu; chất lượng đồ thị phụ thuộc biểu diễn nút và k ; entropy chưa thay thế đánh giá hiệu chỉnh; chưa ngoại kiểm đa trung tâm. Hướng tiếp theo là thực hiện ablation đầy đủ, phân tích lớp thiểu số, phát triển cross-attention hoặc Transformer đa nguồn, khảo sát graph transformer và đồ thị không đồng nhất, kết hợp entropy với ECE, Brier score và deep ensembles, tích hợp XAI, ngoại kiểm độc lập và thẩm định lâm sàng.

Nhìn chung, hai đóng góp thống nhất ở mục tiêu tổ chức và khai thác biểu diễn đa nguồn đáng tin cậy hơn. Mô hình là kết quả nghiên cứu phương pháp, chưa phải công cụ chẩn đoán độc lập.

DANH MỤC CÁC CÔNG TRÌNH CÔNG BỐ LIÊN QUAN ĐẾN LUẬN ÁN

- [CT1]. Nguyen, Thi Trang; Nguyen, Ha Phuong; Vu, Van Hieu. "Hybrid Feature Fusion: Enhancing Scene Image Classification with Handcrafted and Deep Learning Approaches". In: Kỷ yếu Hội nghị Khoa học Công nghệ Quốc gia lần thứ XVII về Nghiên cứu cơ bản và ứng dụng Công nghệ thông tin (FAIR), Hà Nội, Việt Nam, 08--09/08/2024. 2024. DOI: 10.15625/vap.2024.0226.
- [CT2]. Vu, Van Hieu; Nguyen, Thi Trang; Nguyen-Thi, Thao. "Ensemble Deep Learning Models for Enhanced Skin Lesion Diagnosis Using the HAM10000 Dataset". In: Hội thảo quốc gia lần thứ XXVII: Một số vấn đề chọn lọc của Công nghệ thông tin và truyền thông, Nha Trang, Việt Nam, 11--12/10/2024. 2024.
- [CT3] Nguyen, T.T., Vu, V.H. "Entropy-based feature fusion via BlockNet for robust skin cancer classification". Iran Journal of Computer Science. Volume 8, pages 2871–2891 (2025). DOI:10.1007/s42044-025-00351-8 (Scopus, Q2).
- [CT4] Nguyen, T.T., Vu, V.H. \& Nguyen, V.A. Semi-supervised GNN for Multi-modal Skin Lesion Classification. International Journal of Computational Intelligence Systems. Volume 19, article number 208 (2026). <https://doi.org/10.1007/s44196-026-01296-x> (SCI-E/Q1).