

## Nâng cao hiệu quả phát hiện và chuẩn đoán bệnh viêm da cơ địa sử dụng mạng học sâu

**Tóm tắt nội dung**—Bệnh viêm da cơ địa (Atopic Dermatitis-AD) là một triệu chứng nặng của bệnh chàm (Eczema-Ecz)[1]; gây ngứa, đỏ, rạn và ngứa trên da. 70% ca bệnh xảy ra trong độ tuổi trẻ em, trong đó có đến 15-20% trẻ em từ tuổi 13 trở xuống bị ảnh hưởng bởi bệnh[2][3]. Vậy nên cần có một phương pháp ước tính độ phát triển nghiêm trọng của bệnh nhằm đánh giá sự tiến triển của bệnh nhân và hiệu quả của việc điều trị. Cùng lúc đó, công nghệ máy học sâu đang phát triển rất nhanh trong kỷ nguyên số và được sử dụng nhiều trong các lĩnh vực y tế như bệnh học và X quang[4][5]. Nhưng sự áp dụng của công nghệ này trong vùng da liễu còn khá khiêm tốn và nhiều sai sót vì sự phức tạp của các biến chứng bệnh[6]. Bài nghiên cứu này trình bày việc nâng cao hiệu quả chuẩn đoán bệnh viêm da cơ địa dựa trên các công nghệ mạng học sâu tân tiến, bằng cách phân tích dữ liệu ảnh vùng da bị nhiễm bệnh. Kết luận cho thấy rằng phương pháp chứng minh được hiệu quả về cả tốc độ và tính nhất quán thay cho cách chuẩn đoán truyền thống, đạt được kết quả tương đương với chuyên gia chuẩn đoán trong khi hạn chế sự không đồng nhất giữa những người chuẩn đoán.

**Keywords**-Atopic Dermatitis-AD, Deep Learning, Convolutional Neural Networks-CNNs

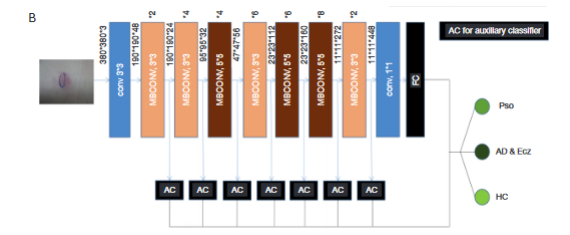
### I. ĐẶT VẤN ĐỀ

AD là một dạng nặng của bệnh chàm gây ngứa mãn tính, ảnh hưởng 15-20% trẻ em nhưng có thể bị nhiễm ở mọi lứa tuổi. Được ước tính rằng có tới 16,5 triệu người lớn tại Mỹ (tương ứng 7,3%) đã từng bị mắc AD khi chỉ mới ngưỡng 2 tuổi trở lên, với 40% trong số đó mắc bệnh vừa hoặc nặng[7]. Ngoài những ảnh hưởng về da, những người mắc phải AD có nguy cơ nhiễm bệnh dị ứng đường hô hấp như hen suyễn hoặc viêm mũi dị ứng, hay những rối loạn về rối loạn thần kinh như trầm cảm hay rối loạn lo âu[8][9]. Thêm vào đó, AD thường dễ bị chuẩn đoán sai do các triệu chứng thường phân bố và biến đổi bất thường và có điểm tương đồng với các loại bệnh viêm da khác. Vì vậy nên cần thiết có phương pháp để hỗ trợ các bác sĩ và chuyên gia trong việc chuẩn đoán, đặc biệt hữu ích cho chuyên gia da liễu trẻ và thiếu kinh nghiệm.

Các mô hình học sâu, cụ thể như CNNs (Convolutional Neural Networks), RNNs (Recurrent Neural Networks) hay GANs (Generative Adversarial Networks) đã cho thấy sự hiệu quả trong việc phân tích các dữ liệu ảnh lâm sàng. Một số ví dụ như việc CNN đã được áp dụng trong hỗ trợ chuẩn đoán và phát hiện sớm dấu hiệu của bệnh Alzheimer's, viêm xương khớp hay vông mạc tiểu đường qua đầu cơ sở dữ liệu các ảnh chụp loại ảnh chụp như MRI, EGG hoặc siêu âm.[10]

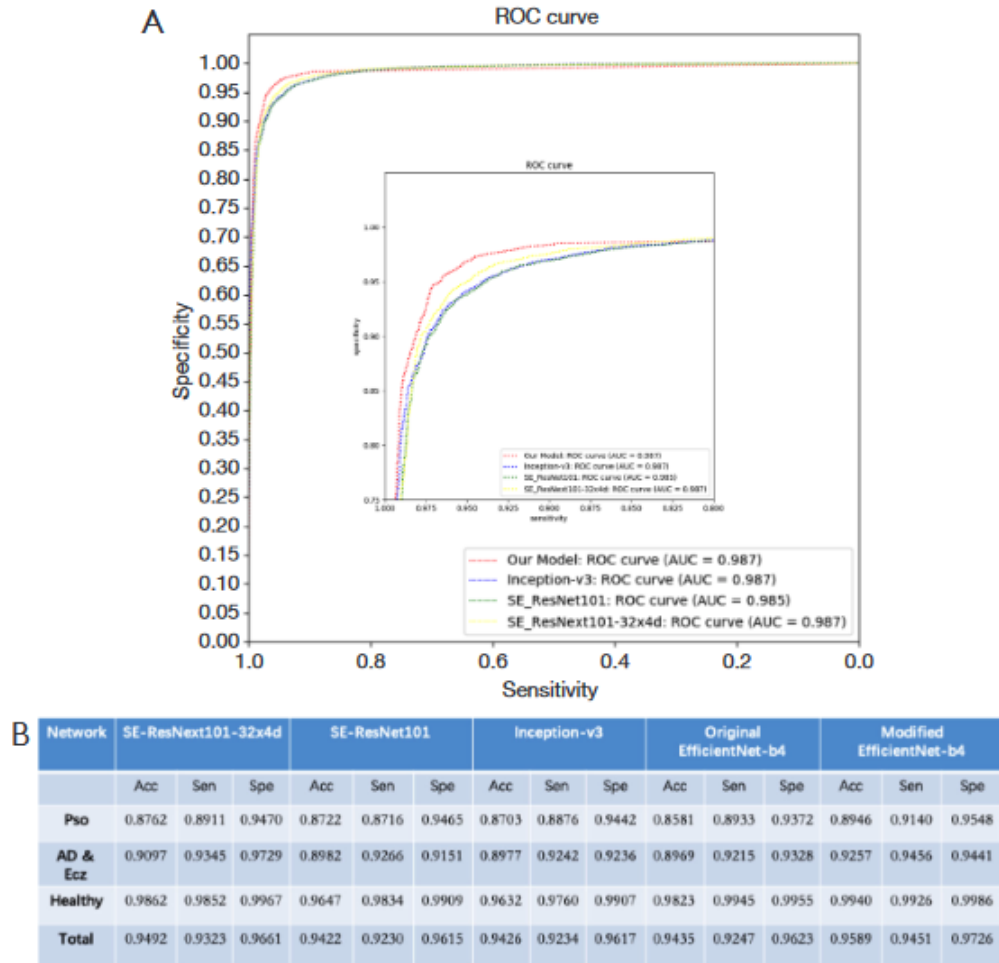
Việc khai triển công nghệ học sâu trong chuẩn đoán da được bắt đầu bởi chuẩn đoán u ác tính. Từ đó, đã có một số nghiên cứu sử dụng trí tuệ nhân tạo (Artificial Intelligence-AI) để xác định bệnh ung thư da dựa trên cơ sở đầu vào là các ảnh chụp mô sinh thiết da và kính soi da. Việc sử dụng công nghệ tương tự trong chuẩn đoán các bệnh da liễu, cụ thể là AD, là hoàn toàn thực tiễn và cần thiết.

Phương pháp sử dụng máy học sâu dựa trên dữ liệu ảnh để tự động chuẩn đoán các bệnh viêm da được giới thiệu trong tài liệu [11], khi các tác giả đã xây dựng một trợ lý trí tuệ nhân tạo chuẩn đoán các bệnh da (AIDDA) dựa trên hình ảnh da lâm sàng để chuẩn đoán bệnh vẩy nến (psoriasis-Pso), Ecz và AD sử dụng mô hình học sâu CNN EfficientNet-b4 đến từ Google (Hình 1).



Hình 1: Khung nền mạng học sâu CNNs.

Nhóm đã sử dụng Pytorch để khai triển khung phần mềm học sâu và vẽ vòng ROC (Receiver Operating Characteristic curve) liên kết chỉ số độ nhạy và độ đặc hiệu thông qua AUC (Area Under



Hình 2: Vòng ROC với AUC và các chỉ số chỉ độ chính xác.

the Curve). Kết quả cho thấy AIDDA đạt độ chính xác là 95.8% với sai số 0.09%; độ nhạy và độ đặc hiệu lần lượt là 97.6% với sai số 0.12% và 95.89% với sai số 0.06%.

Mặc dù AUC của mô hình với những mô hình ConvNets khác được so sánh trong nghiên cứu (gồm SE-ResNeXt101-32x4d, SE-ResNet101 [12] và Inception-v3 [13]) khá tương đồng, mô hình của nhóm nghiên cứu vẫn nhỉnh hơn ở độ nhạy và độ đặc hiệu. (Hình 2).

Mô hình này được áp dụng tạo nên một ứng dụng trên điện thoại (Hình 3), và đã có hơn 7,000 bác sĩ tại các bệnh viện đăng ký sử dụng ứng dụng này. Dữ liệu người dùng cho thấy đã có gần 100,000 ảnh chụp từ các bác sĩ vào ứng dụng để cải thiện quá trình chuẩn đoán, chứng minh sự hiệu quả và cần

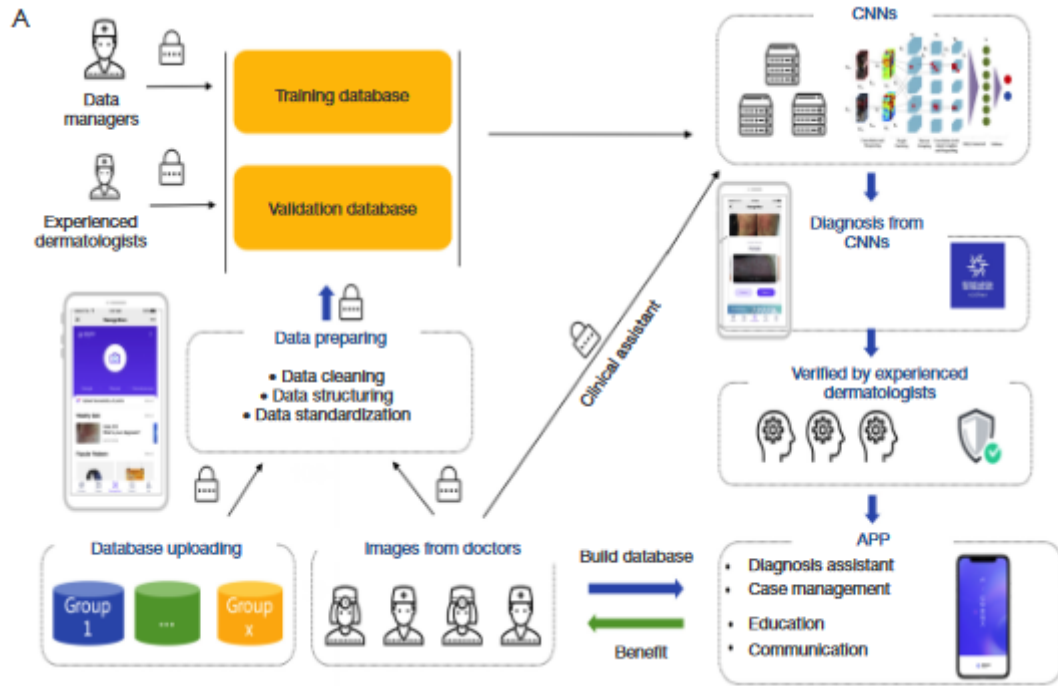
thiết cho một công cụ hỗ trợ sử dụng mạng học sâu như vậy, đặc biệt trong chuẩn đoán các bệnh về da liễu nói chung và viêm da cơ địa nói riêng.

Tương tự [11], trong các năm gần đây nhiều nghiên cứu, thử nghiệm đã tiến hành sử dụng và kết hợp CNN trong chuẩn đoán bệnh viêm da cơ địa, kết quả cho thấy độ chính xác tăng cao, tốc độ tính toán cũng được cải thiện.

## II. CÁC NGHIÊN CỨU LIÊN QUAN

### A. Ứng dụng tự động hóa đánh giá, phân loại và chuẩn đoán các bệnh da liễu

Trong [7], các tác giả đã thực nghiệm với công nghệ học sâu CNN kết hợp thước đo đánh giá mức độ nghiêm trọng của viêm da cơ địa (SCORAD) để tạo thành thuật toán tính SCORAD tự động



Hình 3: Mô hình luồng làm việc trên nền tảng di động của AIDDA.

(ASCORAD). Họ cũng dùng Pytorch để xây dựng mô hình và 3 bộ dữ liệu (dataset) từ 3 nguồn chuyên gia da liễu khác nhau, dựa trên hai chỉ tiêu: phân đoạn bề mặt bị thương và đánh giá mức độ nghiêm trọng của dấu hiệu trực quan.

Kết quả đạt được cho thấy RMAE tổng thể đạt được một ngưỡng khá là 13,0% và xuất sắc đạt được AUC là 0,93 và IoU là 0,75 khi ước tính bề mặt tổn thương trên da sáng. Điều này chứng minh rằng mạng học sâu có thể là một phương pháp thay thế thuận lợi và khách quan cho tự động đánh giá về AD với nhiều tiềm năng, có thể so sánh với kết quả đánh giá của chuyên gia người, đồng thời giảm sự không tương đồng của việc nhiều chuyên gia đánh giá và giảm sự hao tổn thời gian.

Trong [14], các nhà nghiên cứu đã sử dụng một phương pháp tương tự thông qua hai mô hình học sâu CNN và LSTM (long short-term memory) để chuẩn đoán bệnh vẩy nến (Psoriasis-Pso), một loại bệnh có nhiều đặc điểm tương đồng với AD-cùng thuộc nhóm bệnh viêm da, có quan hệ tương tự với Ecz. Kết quả cho thấy độ chính xác thu được ở CNN và LSTM lần lượt là 84.2% and 72.3%, cho thấy tiềm năng của việc áp dụng mô hình học sâu trong chuẩn đoán các bệnh về da liễu và là bằng

chứng cho sự cải thiện đáng kể về độ chính xác của mô hình trong [11].

Việc sử dụng các công nghệ học sâu liên quan như R-CNN dưới khung phần mềm Tensorflow để chuẩn đoán các bệnh như AD, mụn trứng cá và ghẻ lở được nhóm nghiên cứu trong [15] khai thác. Các tác giả đã xây dựng một hệ thống dựa trên nền tảng web để chuẩn đoán nhanh và hiệu quả hơn những bệnh về da kể trên (medilab-plus). Kết quả cho thấy mô hình dự đoán chính xác 88% cho AD, 85% cho mụn trứng cá và 84.7% cho ghẻ lở. Thông qua so sánh có thể thấy kết quả sự tương đồng kết quả với nhiều nghiên cứu khác như trong [11] và [14].

Nghiên cứu [16] giới thiệu một cách đánh giá khác sử dụng CNN và ảnh RSOM 3D thô để phân loại AD dựa trên các cách tiếp cận học máy truyền thống như SVM (support vector machine) và RF (Random Forest). Các tác giả sử dụng package Scikit-learn làm thực nghiệm và đánh giá trên chỉ số tự tạo EVSI (Eczema Vascular Severity Index) bao gồm (i) da khỏe vs. AD và (ii) Triệu chứng nhẹ vs. Vừa-Nặng của AD. Kết quả cho thấy mẫu hình CNN đạt được độ chính xác là 97% và RF là 65%. Điều này lần nữa chứng minh hiệu suất của công cụ máy học sâu, đặc biệt là CNN, trong đó cho thấy khả năng phát triển đáng mong chờ trong việc áp

dụng RF vào phân loại cùng với đó là chuẩn đoán các bệnh da liễu, cụ thể là AD.

### B. Ứng dụng sử dụng dữ liệu hệ phiên mã và quần thể vi khuẩn

Các tác giả trong [10] và [17] xây dựng 2 thuật toán phân loại học máy (machine learning classifier) sử dụng hệ phiên mã của tế bào biểu mô ruột (i) và dữ liệu quần thể vi khuẩn đường ruột (ii) để chuẩn đoán AD với độ chính xác cao một cách tự động. Khác với các nghiên cứu trước, tác giả sử dụng Scikit-learn thay vì Pytorch, do package này phù hợp hơn để học máy.

Tác giả sử dụng SVM làm mô hình lựa chọn đặc trưng chính (feature selection), và cho ra chỉ số: ở (i) với F1 là 0.76 với lựa chọn đặc trưng, hơn đó xấp xỉ 0.5 với tất cả đặc trưng và AUC tốt nhất là 0.75, có lựa chọn đặc trưng là 0.72, thấp hơn SVM có tất cả đặc trưng là 0.3; ở (ii) với F1 cải thiện so với (i) là 0.02, cho kết quả 0.78 và AUC tương đồng với (i) là 0.75. Kết quả cho thấy máy phân loại đạt được độ Precision và Recall cao (0.70 và 0.88). Nghiên cứu đã tạo bước đầu trong quá trình chuẩn đoán và đưa ra những thấu hiểu mặt sinh học về sự phát triển dấu ấn sinh học của loại bệnh này.

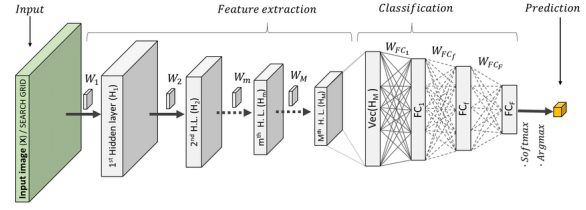
Từ kết quả đạt được tại [10] và [17], các tác giả trong [18] sử dụng dữ liệu hệ phiên mã ở mô hữu cơ trong da dựa trên Random Forest và mạng nơ-ron nhân tạo (neural networks) để phát hiện AD một cách chính xác hơn là AUC bằng 0.9802 ở bộ dữ liệu huấn luyện (training set) và bằng 1.0000, 0.8500 ở lần lượt hai tập kiểm duyệt GSE5667 và GSE102628. Kết quả cho thấy sự phát triển đáng kể từ những kết quả của các báo cáo trước, và sự khôn khéo trong việc lần đầu áp dụng Random Forest để chuẩn đoán AD cũng như làm tiền nhiệm cho những nghiên cứu sau ([16]). Mặc dù vậy, còn một số hạn chế như việc mẫu thử còn khá ít để thực sự chứng minh tiềm năng của công này và còn thiếu các tập dữ liệu để dùng trong khuôn kiểm duyệt.

## III. MỘT SỐ KHÁI NIỆM CƠ BẢN

### A. Mạng nơ-ron tích chập (Convolutional Neural Networks -CNN)

CNN là một lớp của mạng nơ-ron nhân tạo thường được dùng trong các bài toán phân lớp, được áp dụng phổ biến nhất để phân tích hình ảnh trực quan. CNN Còn nhiều ứng dụng phổ biến khác trong thị giác máy tính [19], hệ thống gợi ý, xử lý ngôn ngữ tự nhiên hay phân loại hình ảnh, tính toán dữ liệu ảnh y tế. [20]. Hệ thống CNN thường gồm

3 phần: đầu vào, gồm nguồn dữ liệu được chuẩn bị sẵn từ một mô hình huấn luyện bên thứ 3; phần ẩn, gồm phần chuẩn lọc đặc trưng và các lớp phân loại; và đầu ra, gồm dữ liệu chuẩn đoán và các chỉ số liên quan tới dữ liệu (Hình 4).



Hình 4: Mô hình mẫu đơn giản cấu tạo CNN.

Trong đó, các lớp tách và phân loại đặc trưng thường có vai trò quyết định trong việc đưa ra kết quả cuối cùng; vì vậy, chúng được chia thành nhiều nhánh nhỏ, gồm nhiều lớp tách và phân loại tương ứng. Trong các thử nghiệm ở phần trên, mỗi CNN đều bao gồm một số lượng đáng kể của tầng lớp phân loại, chúng giúp các tác giả nhìn nhận được sự thay đổi chỉ số ở mỗi tầng, từ đó đưa ra được kết luận và mô hình áp dụng chính xác nhất.

Hiện nay, công nghệ này đã có nhiều ứng dụng thực tiễn trong ngành y, sử dụng để đánh giá, phân loại và chuẩn đoán các loại bệnh thông qua dữ liệu phim và ảnh. Đặc biệt trong các khoa ngành như khoa bệnh lý, khoa X quang hay khoa da liễu. Vì vậy, CNN có nhiều tiềm năng để cải thiện độ chính xác chuẩn đoán, qua cải ảnh giải phẫu và lâm sàng.

### B. Đường cong ROC (Receiver Operating Characteristic-ROC Curve)

ROC Curve là một dạng đồ thị được dùng khá phổ biến trong khoa kiểm chứng (validation), dựa trên chỉ số độ nhạy & độ chuyên (Sensitivity-TPR & Specificity-FPR) của một mô hình học sâu. Nghiên cứu chính và các nghiên cứu liên quan đều sử dụng ROC để đánh giá tính khách quan mô hình học sâu của họ đến một ngưỡng nào đó. TPR và FPR cũng là hai tham số ở trên hai trục của đồ thị, được tính như dưới đây:

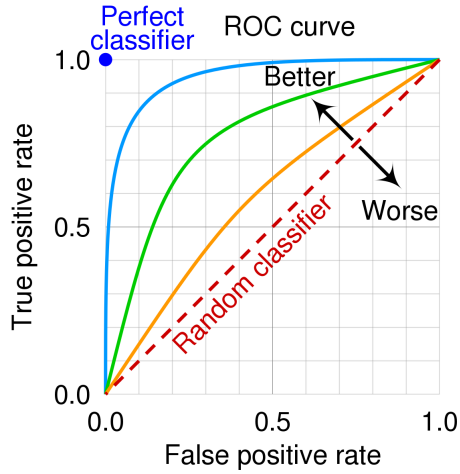
- TPR (True Positive Rate), đồng nghĩa với chỉ số Recall được thấy trong một số nghiên cứu, được định nghĩa:

$$TPR = \frac{TP}{TP + FN}$$

- FPR (False Positive Rate), được định nghĩa:

$$FPR = \frac{FP}{FP + TN}$$

Hình 5 là ví dụ cho một biểu đồ đường cong ROC.

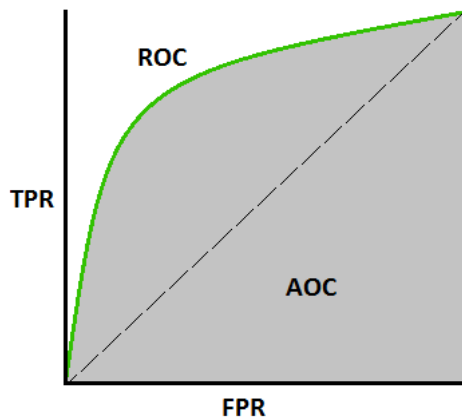


Hình 5: Biểu đồ đặc trưng một đường cong ROC

#### C. Diện tích Dưới Đường Cong (Area Under the Curve-AUC)

AUC là một chỉ số chỉ diện tích dưới một đường cong ROC. Chỉ số đo toàn bộ diện tích hai chiều ở dưới đường cong, và cho ta một cái nhìn toàn diện về hiệu suất xuyên suốt tất cả các ngưỡng phân loại. AUC có miền từ 0.0 đến 1.0, với 1.0 tương ứng 100% độ chính xác và ngược lại.

Hình 6 là ví dụ cho một biểu đồ ROC curve chứa AUC.



Hình 6: Biểu đồ đường cong ROC chứa AUC

#### D. Precision, Recall & F1-score

Precision, Recall và F1-score đều là những chỉ số dựa trên TPR và FPR đã được định nghĩa ở phần trên. Theo đó,

- Precision được định nghĩa:

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

- Recall, đồng nghĩa với TPR, được định nghĩa:

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

- Trong mô hình, nhằm phục vụ độ chính xác chúng ta mong muốn hai chỉ số này càng cao càng tốt. Tuy nhiên trong thực tế, nếu ta điều chỉnh mô hình tăng Recall quá mức có thể làm giảm Precision và ngược lại. Nhiệm vụ của chúng ta là phải cân bằng 2 đại lượng này. Và để giải quyết bài toán này khi chọn cách điều chỉnh mô hình, ta có chỉ số F1-score được định nghĩa:

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

- Trong điều kiện lý tưởng, F1-score = 1, từ đó ta có thể căn cứ chọn mô hình thông qua F1-score.

#### IV. THỰC NGHIỆM

##### A. Miêu tả dữ liệu

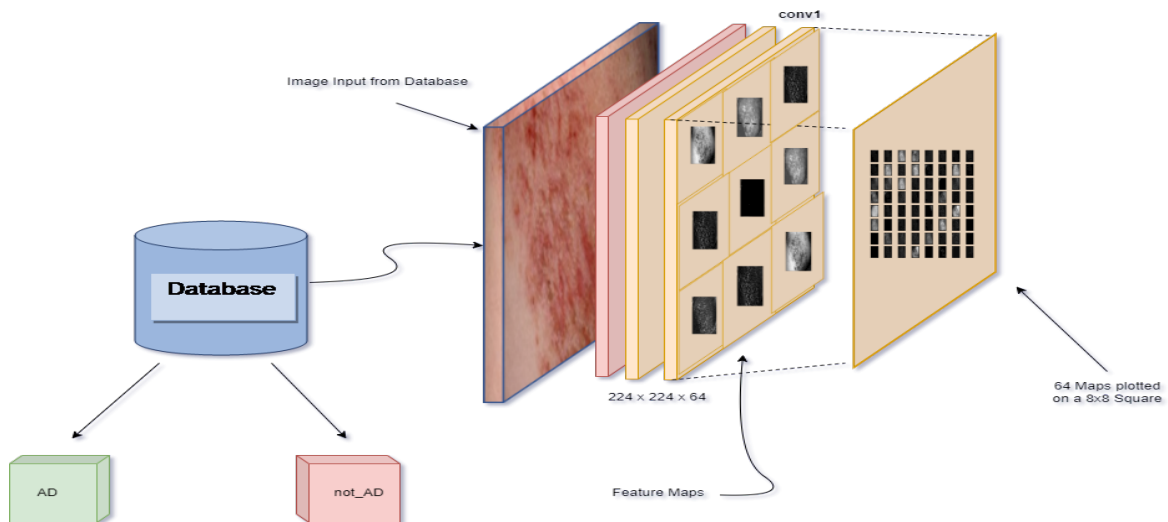
Bộ dữ liệu chúng tôi dùng để thực nghiệm là gồm hai thư mục chính: AD (ảnh chụp bị mắc viêm da cơ địa) và not\_AD (ảnh chụp không phải viêm da cơ địa).

Thông tin bộ dữ liệu:

- Bộ dữ liệu gốc có tổng cộng 207 mẫu ảnh được chia vào hai thư mục AD và not\_AD. Trong đó, có 90 ảnh nằm trong AD và 117 ảnh nằm trong not\_AD. Tập tin trong AD gồm ảnh chụp những vùng bị nhiễm viêm da cơ địa ở nhiều phần bộ phận cơ thể khác nhau. Tập tin trong not\_AD gồm ảnh chụp những vùng bị nhiễm những bệnh viêm da hay da liễu khác như vẩy nến, bệnh chàm và viêm da dầu. Nhằm đạt độ chính xác, các ảnh đã được chuẩn hóa dưới dạng 294x222, dưới dạng ảnh PNG và được chuyển đen trắng khi cho vào thực nghiệm.

- Bộ dữ liệu làm sạch được sử dụng gồm hai ảnh đại diện cho hai thư mục chính ở bộ dữ liệu gốc. Thông qua mô hình tách lọc đặc trưng trong [21], thực nghiệm tách 2 ảnh thành 64 lớp ảnh đặc trưng cho mỗi ảnh ở từng tầng cuối của từng lớp tích chập trong mô hình VGG16. Kết quả cho bộ dữ





Hình 7: Cách thức phân khúc & triết xuất ảnh của các tầng tích chập

liệu cuối cùng đạt chuẩn 320 ảnh đặc trưng chia ra tương ứng với 5 thư mục được đặt dưới tên của lớp tích chập đó. Ngoài ra còn có một thư mục chứa 5 ảnh 8x8 gồm tất cả các ảnh đặc trưng của từng lớp nhằm đánh giá toàn diện chất lượng từng tầng. (Hình 7)

### B. Mô hình thực nghiệm

Về tổng thể, mô hình gồm đầu vào từ bộ dữ liệu đã được chuẩn lọc từ tầng CNN là VGG16 (Hình 8), một tầng gồm phân lớp ảnh HOG (Histogram of Oriented Gradients) và qua hồi quy LR (Logistic Regression) để cho đầu ra chỉ số độ chuẩn xác (Accuracy), độ chính xác phép đo (Precision) và điểm F1.

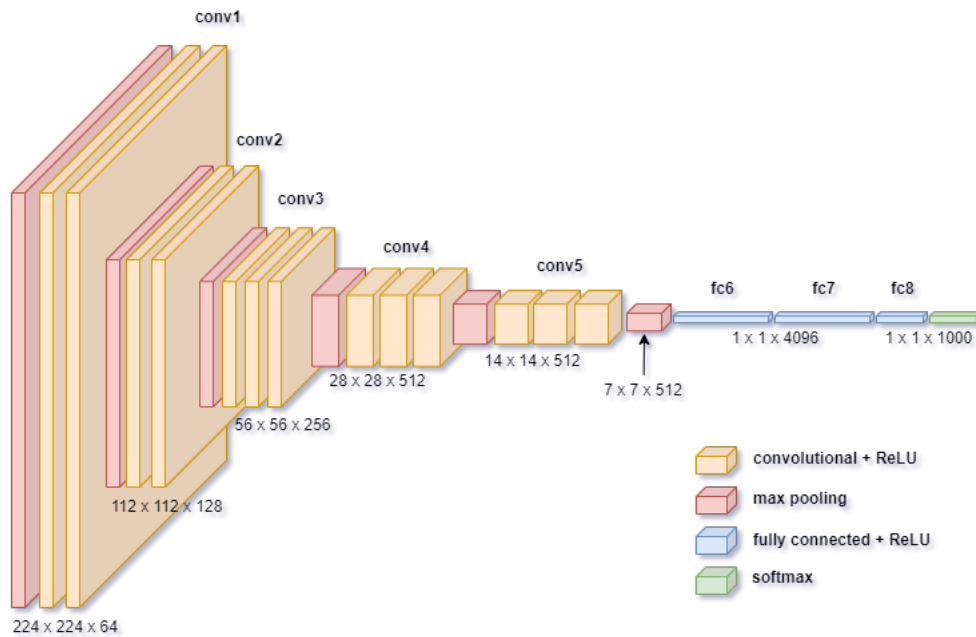
Đầu tiên, mô hình nhận nguồn dữ liệu bằng cách thực hiện phép lặp tìm tệp tin có dạng PNG ở từng thư mục. Sau đó, mô hình sẽ chiết xuất ảnh HOG của một tệp, gán mã tên và in số lượng ảnh tương ứng của từng thư mục. Sử dụng matplotlib, có thể thu được mẫu ảnh ban đầu so sánh với ảnh sau khi được tách đặc trưng HOG. Khi đã có ảnh đặc trưng HOG và gán mã cho ảnh, chúng tôi tạo dataset được chia ra thành 90% tập huấn luyện (training set) và 10% tập kiểm thử (testing set). Cuối cùng là thiết lập một hồi quy LR để phân loại đa tầng và huấn luyện tầng phân loại vừa được tạo. Khi đó, ta có thể dự đoán được các gán mã trên bộ dữ liệu vừa được tạo cũng như đưa ra những chỉ số về độ chính xác của các dự đoán đó. Hình 9 cho thấy cách thức hoạt động của phân loại HOG.

Để đạt được độ chính xác cao nhất, chúng tôi đã chiết xuất ảnh HOG của từng khối lớp tích chập ở mô hình trong phần Miêu tả bộ dữ liệu và tạo thành thư mục riêng lần lượt là Conv1, Conv2, Conv3, Conv4 và Conv5 tương ứng với các lớp tích chập của VGG16. Ở trong mỗi thư mục là hai thư mục con gán mã AD và not\_AD đại diện cho đặc trưng bệnh AD và các bệnh da liễu khác. Điều này nhằm mục đích chọn ra khối tầng hay thư mục có đặc trưng có thể được chuẩn đoán một cách chính xác và hiệu quả nhất để làm dữ liệu đầu vào cho các nghiên cứu trong tương lai.

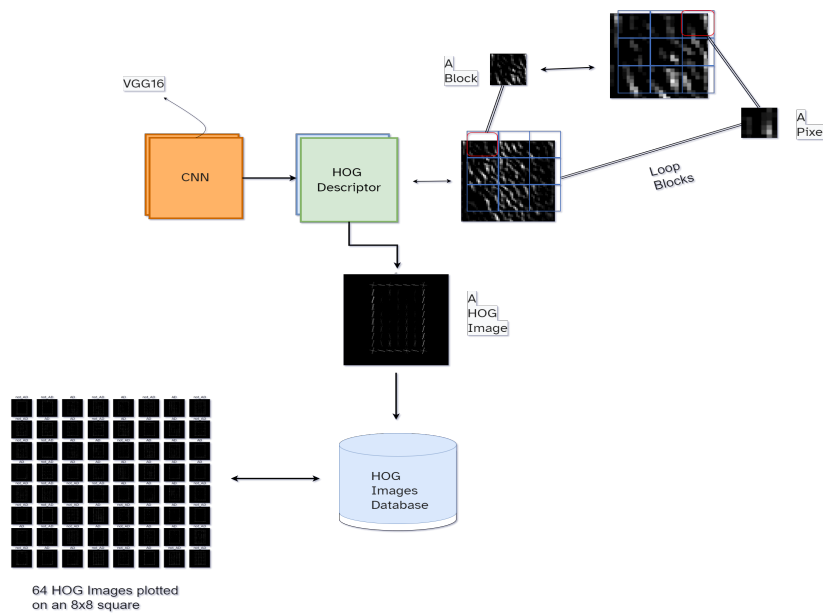
### C. Kết quả thực nghiệm

Khi lấy những thư mục kể trên làm bộ dữ liệu đầu vào cho tầng giữa HOG, sau hồi quy LR kết quả đạt được là một độ chính xác tuyệt đối với chỉ số 1.0 cho Conv1 và Conv2, sau đó giảm nhẹ còn 0.923 ở Conv3 và tụt dốc tại Conv4 với 0.69, khi ở tầng cuối cùng của khối Conv5 chỉ đạt 0.53. Điều này chứng minh rằng sẽ là phù hợp nhất nếu lấy các đặc trưng, cụ thể là ảnh, ở Conv3 trở lên để giữ được chất lượng và độ tính khách quan cho thí nghiệm.

So sánh với kết quả ở các nghiên cứu liên quan như [17] và [10], chỉ số F1-score cũng như hai chỉ số liên hệ là precision và recall của chúng tôi đặc biệt ở 3 tầng đầu-đều là cải thiện một cách đáng kể. Riêng ở hai tầng Conv1 và Conv2, tương ứng với độ chính xác, chúng đều cho kết quả F1, Precision và Recall là 1.00. Tuy vậy, chúng tôi khuyến khích



Hình 8: Tầng CNN đầu vào-VGG16, một công nghệ phát triển từ CNN dùng trong phân loại và tách đặc trưng



Hình 9: Tầng phân loại đặc trưng HOG

sử dụng Conv3 để tiếp tục những nghiên cứu sau này, vì tầng mang những đặc trưng sâu hơn của ảnh để đem so sánh mà vẫn giữ được độ chính xác cao (1.00 Precision, 0.83 Recall) (Hình 10), thay vì chỉ

giới hạn đánh giá bề mặt như 2 tầng đầu Conv1 & Conv2 cũng như tránh việc bị giảm độ chính xác đáng kể ở 2 tầng cuối Conv4 & Conv5.

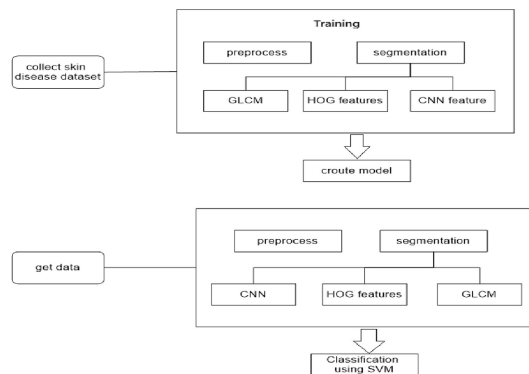
Với kết quả đạt được, nhiều ứng dụng thực tiễn

Accuracy: 0.9230769230769231

	precision	recall	f1-score	support
AD	0.88	1.00	0.93	7
not_AD	1.00	0.83	0.91	6
accuracy			0.92	13
macro avg	0.94	0.92	0.92	13
weighted avg	0.93	0.92	0.92	13

Hình 10: Kết quả chỉ số thu được khối tầng Conv3.

có thể được phát triển thông qua thực nghiệm. Ví dụ cụ thể nhất là kết hợp mô hình này làm lớp giữa cho một hệ thống hay DNN/CNN khác, đóng vai trò vừa là một lớp phân loại cho bộ dữ liệu ở tầng DNN trước vừa là một trụ sở dữ liệu cho một tầng DNN sau. Ngoài ra còn có thể kết hợp với một số công nghệ phân loại khác, như SVM hay KNN (K-nearest neighbors algorithm), để cường độ hóa năng suất của cả CNN. Điều này được thể hiện sự tiềm năng cũng như linh hoạt của CNN, đặc biệt trong việc ứng dụng khoa da liễu, khi dữ liệu mà thực nghiệm sử dụng dựa trên cơ sở bệnh viêm da cơ địa.



Hình 11: Mô hình gồm (A), tầng huấn luyện dữ liệu đầu vào; (B), CNN chỉ rõ lớp chuẩn lọc và phân loại (gồm HOG và SVM).

Thực nghiệm cho độ chính xác tương đối so với các nghiên cứu khác, như trong [11] khi mô hình cũng đạt kết quả về độ chính xác ngưỡng 0.86. Một số thực nghiệm khác cũng áp dụng VGG16 làm lớp tích hợp phân loại trước khi cho ảnh vào phân loại đặc trưng HOG. Một trong số đó có thể kể đến là [20], khi thí nghiệm thay vì phân loại và đánh giá ảnh y, các tác giả phân tích ảnh logo thương hiệu các tập đoàn đa quốc gia và đạt được những kết quả

rất khả quan. Thông thường, những nghiên cứu về tách biểu trưng hay nhân của hình ảnh như logo sẽ kết hợp sử dụng SVM sau khi đã thu được kết quả từ tính toán đặc trưng ảnh HOG, một mô hình đặc trưng như vậy có thể được thấy trong Hình 11 [21]. Điều này cũng chứng minh sự khả thi của việc áp dụng SVM với lớp DNN như CNN và phân loại HOG một cách tương tự trong đánh giá và chuẩn đoán AD và viêm da liễu nói riêng, thậm trí kể đến ung thư da và các bệnh da khác nói chung.

## V. KẾT LUẬN

Bài nghiên cứu cùng với thực nghiệm của chúng tôi đã cho thấy một số kết luận nhất định trong nâng cao hiệu quả phát hiện và chuẩn đoán bệnh viêm da cơ địa sử dụng mạng học sâu. Mạng học sâu được sử dụng chính trong các bài nghiên cứu và thực nghiệm của chúng tôi là CNN, cho ra kết quả chuẩn đoán và đánh giá cao các đầu ảnh dữ liệu đạt độ chính xác cao. Điều này cho thấy tiềm năng của công nghệ này không chỉ trong phát hiện viêm da cơ địa nói riêng mà cả các bệnh da liễu nói chung, thậm chí có thể áp dụng trong chuẩn đoán ung thư da với bộ dữ liệu đủ lớn để đưa ra những chỉ số có thể so sánh được với chuẩn đoán của chuyên gia. Một số nhóm nghiên cứu đã bắt đầu áp dụng những công cụ phân loại như RF, hồi quy LR hay SVM để tăng độ chính xác cho mô hình của mình. Thực nghiệm của chúng tôi cũng cho thấy việc sử dụng phân loại đặc trưng HOG đạt được kết quả tốt khi kết hợp với một tầng CNN làm đầu vào dữ liệu nhằm phục vụ chuẩn đoán AD. Tương tự như vậy, việc kết hợp CNN với nhiều tầng lớp phân loại và DNN khác là hoàn toàn khả thi, không chỉ dừng lại trong chuẩn đoán ảnh y tế mà công việc này đã bắt đầu được thực thi trong nhận diện khuôn mặt hay logo đại diện thương hiệu. Mặc dù những điểm tích cực kể trên, còn một số hạn chế nhất định khi sử dụng mạng học sâu để chuẩn đoán AD. Một trong số đó là hạn chế về số lượng và chiều sâu của các bộ dữ liệu, và sự thiếu nhất quán trong các bộ dữ liệu ảnh đó. Tuy nhiên, công nghệ đã đạt được mục đích nâng cao hiệu quả toàn diện và đã có những ứng dụng thực tiễn tại các bệnh viện trung ương cũng như tạo nền móng cho các nghiên cứu tương tự.

## CẢM ƠN

## TÀI LIỆU

- [1] T. Aoki, T. Fukuzumi, J. Adachi, K. Endo, and M. Kojima, "Re-evaluation of skin lesion distribution in atopic dermatitis. analysis of cases 0



- to 9 years of age.” *Acta dermato-venereologica. Supplementum*, vol. 176, pp. 19–23, 1992.
- [2] F. S. Larsen and J. M. Hanifin, “Epidemiology of atopic dermatitis,” *Immunology and Allergy Clinics*, vol. 22, no. 1, pp. 1–24, 2002.
  - [3] S. Hong, D. K. Son, W. R. Lim, S. H. Kim, H. Kim, H. Y. Yum, and H. Kwon, “The prevalence of atopic dermatitis, asthma, and allergic rhinitis and the comorbidity of allergic diseases in children,” *Environmental health and toxicology*, vol. 27, 2012.
  - [4] Z. C. C. Fuxench, J. K. Block, M. Boguniewicz, J. Boyle, L. Fonacier, J. M. Gelfand, M. H. Grayson, D. J. Margolis, L. Mitchell, J. I. Silverberg *et al.*, “Atopic dermatitis in america study: a cross-sectional study examining the prevalence and disease burden of atopic dermatitis in the us adult population,” *Journal of Investigative Dermatology*, vol. 139, no. 3, pp. 583–590, 2019.
  - [5] R. Luoma, A. Koivikko, and M. Viander, “Development of asthma, allergic rhinitis and atopic dermatitis by the age of five years: a prospective study of 543 newborns,” *Allergy*, vol. 38, no. 5, pp. 339–346, 1983.
  - [6] J. Silverberg, J. Gelfand, D. Margolis, M. Boguniewicz, L. Fonacier, M. Grayson, P. Ong, Z. Chiesa Fuxench, and E. Simpson, “Symptoms and diagnosis of anxiety and depression in atopic dermatitis in us adults,” *British Journal of Dermatology*, vol. 181, no. 3, pp. 554–565, 2019.
  - [7] A. Medela, T. Mac Carthy, S. A. A. Robles, C. M. Chiesa-Estomba, and R. Grimalt, “Automatic scoring of atopic dermatitis using deep learning: A pilot study,” *JID Innovations*, vol. 2, no. 3, p. 100107, 2022.
  - [8] S. Wang, D. M. Yang, R. Rong, X. Zhan, and G. Xiao, “Pathology image analysis using segmentation deep learning algorithms,” *The American journal of pathology*, vol. 189, no. 9, pp. 1686–1698, 2019.
  - [9] M. A. Mazurowski, M. Buda, A. Saha, and M. R. Bashir, “Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on mri,” *Journal of magnetic resonance imaging*, vol. 49, no. 4, pp. 939–954, 2019.
  - [10] J. Li, “Developing machine learning and statistical methods for the analysis of genetics and genomics,” *UCLA Electronic Theses and Dissertations*, pp. 119–146, 2021.
  - [11] H. Wu, H. Yin, H. Chen, M. Sun, X. Liu, Y. Yu, Y. Tang, H. Long, B. Zhang, J. Zhang *et al.*, “A deep learning, image based approach for automated diagnosis for inflammatory skin diseases,” *Ann Transl Med* 2020, vol. 8(9), p. 581, 2020.
  - [12] J. Hu, L. Shen, S. Albanie *et al.*, “Squeeze-and-excitation networks,” *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7132–7141, 2018.
  - [13] E. Laitinen, E. S. Lohan, J. Talvitie, and S. Shrestha, “Access point significance measures in wlan-based location,” pp. 24–29, 2012.
  - [14] S. F. Aijaz, S. J. Khan, F. Azim, C. S. Shakeel, and U. Hassan, “Deep learning application for effective classification of different types of psoriasis,” *Journal of Healthcare Engineering*, vol. 2022, 2022.
  - [15] P. Dwivedi, A. A. Khan, A. Gawade, and S. Deolekar, “A deep learning based approach for automated skin disease detection using fast r-cnn,” in *2021 Sixth International Conference on Image Information Processing (ICIIP)*, vol. 6. IEEE, 2021, pp. 116–120.
  - [16] S. Park, S. N. Saw, X. Li, M. Paknezhad, D. Coppola, U. Dinish, A. B. E. Attia, Y. W. Yew, S. T. G. Thng, H. K. Lee *et al.*, “Model learning analysis of 3d optoacoustic mesoscopy images for the classification of atopic dermatitis,” *Biomedical optics express*, vol. 12, no. 6, pp. 3671–3683, 2021.
  - [17] Z. Jiang, J. Li, N. Kong, J.-H. Kim, B.-S. Kim, M.-J. Lee, Y. M. Park, S.-Y. Lee, S.-J. Hong, and J. H. Sul, “Accurate diagnosis of atopic dermatitis by combining transcriptome and microbiota data with supervised machine learning,” *Scientific reports*, vol. 12, no. 1, pp. 1–13, 2022.
  - [18] W. Zhou, A. Li, C. Zhang, Y. Chen, Z. Li, and Y. Lin, “Accurate diagnosis of atopic dermatitis by applying random forest and neural networks with transcriptomic data,” *medRxiv*, 2022.
  - [19] S. Khan, H. Rahmani, S. A. A. Shah, and M. Benamoun, “A guide to convolutional neural networks for computer vision,” *Synthesis lectures on computer vision*, vol. 8, no. 1, pp. 1–207, 2018.
  - [20] B. Kayalibay, G. Jensen, and P. van der Smagt, “Cnn-based segmentation of medical imaging data,” *arXiv preprint arXiv:1701.03056*, 2017.
  - [21] J. Brownlee, “How to visualize filters and feature maps in convolutional neural networks,” *Machine Learning Mastery*, 2019.