

BỘ GIÁO DỤC
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



HOÀNG HẢI LONG

**NGHIÊN CỨU MÔ HÌNH HỌC SÂU VÀ ỨNG DỤNG HỖ TRỢ
CHUẨN ĐOÁN BỆNH VIÊM DA CƠ ĐỊA**

LUẬN VĂN THẠC SĨ HỆ THỐNG THÔNG TIN

Hà nội - 2026

BỘ GIÁO DỤC
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



HOÀNG HẢI LONG

**NGHIÊN CỨU MÔ HÌNH HỌC SÂU VÀ ỨNG DỤNG HỖ TRỢ
CHUẨN ĐOÁN BỆNH VIÊM DA CƠ ĐỊA**

LUẬN VĂN THẠC SĨ HỆ THỐNG THÔNG TIN

Ngành: Hệ thống thông tin

Mã số: 8 48 01 04

NGƯỜI HƯỚNG DẪN KHOA HỌC :

1. PGS. TS. Nguyễn Long Giang

2. PGS. TS. Vũ Văn Hiệu

Hà Nội - 2026

LỜI CAM ĐOAN

Tôi xin cam đoan rằng đề tài mà tôi nghiên cứu trong luận văn này là nghiên cứu của chính tôi. Đề tài được tôi dựa trên các số liệu, dữ liệu do chính tôi học hỏi, nghiên cứu, tìm hiểu. Do đó, các kết quả có đề cập trong luận văn này được tôi cam đoan chưa từng xuất hiện trong bất cứ một nghiên cứu nào đã có trước đây.

Tất cả các số liệu, khái niệm, kiến thức được đề cập trong luận văn nếu được tham khảo sẽ có nguồn trích dẫn đầy đủ, nếu có bất kỳ sai sót nào tôi xin hoàn toàn chịu trách nhiệm.

Tác giả luận văn ký và ghi rõ họ tên



Hoàng Hải Long

LỜI CẢM ƠN

Trong quá trình tôi học tập, tìm hiểu và nghiên cứu kiến thức chuyên ngành, áp dụng kiến thức đã nghiên cứu vào thực hiện đề tài luận văn, tôi đã nhận được sự quan tâm, hướng dẫn và hỗ trợ quý báu từ các thầy cô Học viện Khoa học và Công nghệ.

Đầu tiên, tôi xin gửi lời cảm ơn sâu sắc tới thầy PGS. TS. Nguyễn Long Giang và thầy TS. Vũ Văn Hiệu. Những người thầy truyền lửa, tận tình chỉ bảo, hướng dẫn, tạo điều kiện rất nhiều cho tôi trong quá trình nghiên cứu và hoàn thành luận văn này. Những kiến thức chuyên môn của các thầy là vô cùng quý báu và tôi rất trân trọng.

Tôi cũng xin gửi lời cảm ơn đến các Phòng ban của Học viện Khoa học và Công nghệ đã tạo điều kiện tốt nhất về học tập và nghiên cứu, cũng như các thủ tục trong suốt thời gian vừa qua.

Trong quá trình hoàn thiện luận văn, không thể tránh khỏi những sai sót không đáng có, rất mong nhận được những sự nhận xét quý báu từ các thầy cô.

Tôi xin trân trọng cảm ơn!

Hà Nội, ngày 19, tháng 05, năm 2026



Hoàng Hải Long

MỤC LỤC

LỜI CAM ĐOAN.....	3
LỜI CẢM ƠN.....	4
MỤC LỤC.....	5
DANH MỤC CÁC BẢNG.....	7
DANH MỤC HÌNH VẼ, ĐỒ THỊ.....	8
MỞ ĐẦU.....	9
Lý do chọn đề tài.....	9
Mục đích nghiên cứu.....	10
Nội dung nghiên cứu.....	11
Cơ sở khoa học.....	12
Tính thực tiễn của đề tài.....	12
Những đóng góp của luận văn.....	13
Chương 1: TỔNG QUAN NGHIÊN CỨU.....	14
1.1. Tổng quan về bệnh Viêm da cơ địa.....	14
1.2. Tổng quan về trí tuệ nhân tạo và học sâu trong lĩnh vực Y tế.....	15
1.2.1. Trí tuệ nhân tạo.....	15
1.2.2. Học máy và học sâu.....	15
1.2.3. Chẩn đoán bệnh da liễu sử dụng học sâu.....	16
1.3. Các mô hình học sâu trong xử lý y khoa.....	16
1.3.1. Tổng quan về xử lý ảnh y khoa bằng học sâu.....	16
1.3.2. Mạng neural tích chập.....	17
1.3.3. Một số mô hình CNN phổ biến.....	21
1.3.4. Thách thức trong ứng dụng học sâu cho ảnh da liễu.....	29
1.3.5. Ý nghĩa của việc nghiên cứu.....	29
1.4. Tổng quan các nghiên cứu liên quan.....	30
1.4.1. Các nghiên cứu trên thế giới.....	30
1.4.2. Các nghiên cứu tại Việt Nam.....	31
1.4.3. Những vấn đề còn tồn tại trong các nghiên cứu hiện nay.....	32
1.4.4. Định hướng nghiên cứu của đề tài.....	32
1.5. Kết luận chương 1.....	33
Chương 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU.....	34
2.1. Đối tượng nghiên cứu.....	34
2.1.1. Đối tượng dữ liệu nghiên cứu.....	34
2.1.2. Đối tượng mô hình nghiên cứu.....	37
2.1.3. Đối tượng đánh giá nghiên cứu.....	38

2.2. Phương pháp nghiên cứu.....	40
2.2.1. Phương pháp nghiên cứu lý thuyết.....	40
2.2.2. Phương pháp thu thập dữ liệu.....	41
2.2.3. Phương pháp tiền xử lý dữ liệu.....	42
2.2.4. Phương pháp xây dựng mô hình học sâu.....	43
2.2.5. Phương pháp huấn luyện mô hình.....	44
2.2.6. Phương pháp đánh giá phân tích và so sánh kết quả.....	44
2.3. Kết luận chương 2.....	45
Chương 3. KẾT LUẬN VÀ THẢO LUẬN.....	46
3.1. Xây dựng và huấn luyện mô hình học sâu.....	46
3.1.1. Chuẩn bị dữ liệu thực nghiệm.....	46
3.1.2. Xây dựng và huấn luyện mô hình.....	47
3.2. Đánh giá và thảo luận kết quả thực nghiệm.....	49
3.3. Triển khai mô hình.....	51
3.4. Kết luận chương 3.....	52
KẾT LUẬN VÀ KIẾN NGHỊ.....	54
DANH MỤC TÀI LIỆU THAM KHẢO.....	56

DANH MỤC CÁC BẢNG

Bảng 1.1: Bảng so sánh các mô hình CNN ứng dụng trong lĩnh vực Y tế.....	27
Bảng 3.1: Bảng kết quả thu được của các mô hình sau khi huấn luyện.....	49

DANH MỤC HÌNH VẼ, ĐỒ THỊ

Hình 1.1: Mạng neural tích chập (CNN).....	17
Hình 1.2: Lớp tích chập (Convolution Layer).....	18
Hình 1.3: Hàm kích hoạt (Activation Function).....	19
Hình 1.4: Lớp gộp (Pooling Layer).....	20
Hình 1.5: Lớp kết nối đầy đủ (Fully Connected Layer).....	21
Hình 1.6: Mô hình AlexNet.....	22
Hình 1.7: Mô hình VGGNet.....	23
Hình 1.8: Mô hình ResNet.....	24
Hình 1.9: Mô hình DenseNet.....	25
Hình 1.10: Mô hình Efficient Net.....	26
Hình 2.1: Dataset “20 Skin Diseases Dataset” trên Kaggle.....	36
Hình 2.2: Dataset ảnh bệnh về Viêm da cơ địa đã được đánh nhãn.....	36
Hình 3.1. Kết quả đánh giá của các mô hình.....	54

MỞ ĐẦU

Lý do chọn đề tài

Trong thời đại công nghệ ngày nay, với sự phát triển rất mạnh mẽ của công nghệ thông tin và trí tuệ nhân tạo, đặc biệt là lĩnh vực học sâu, nhiều ứng dụng hỗ trợ trong y tế đã được nghiên cứu và triển khai nhằm nâng cao hiệu quả chẩn đoán và điều trị bệnh [1]. Đặc biệt là các mô hình học sâu hiện đại có khả năng phân tích dữ liệu hình ảnh y tế với độ chính xác cao rất cao [2].

Đề tài luận văn có nghiên cứu về bệnh Viêm da cơ địa, một loại bệnh da liễu mãn tính cực kỳ phổ biến, có xu hướng gia tăng trên toàn thế giới [3]. Bệnh gây ảnh hưởng lớn đến chất lượng cuộc sống của người bệnh do tình trạng ngứa kéo dài, tổn thương da tái phát nhiều lần và nguy cơ biến chứng nếu không được điều trị kịp thời[4].

Tuy nhiên, việc chẩn đoán bệnh hiện nay chủ yếu dựa trên kinh nghiệm lâm sàng của bác sĩ chuyên khoa. Trong nhiều trường hợp, các biểu hiện bệnh có thể tương đồng với những bệnh da liễu khác, dẫn đến khó khăn trong việc phân biệt và đưa ra chẩn đoán chính xác.

Trên thế giới, đã có nhiều nghiên cứu ứng dụng trí tuệ nhân tạo và học sâu trong phân tích ảnh y khoa, đặc biệt là trong nhận diện các bệnh về da [2]. Các mô hình mạng neural tích chập như CNN, ResNet, Efficient Net hay Vision Transformer đã cho thấy hiệu quả cao trong việc phân loại hình ảnh da liễu [5].

Một số nghiên cứu đạt kết quả tương đương với bác sĩ chuyên khoa trong việc nhận diện các bệnh lý da phổ biến. Tuy nhiên, phần lớn các nghiên cứu tập trung vào các bệnh như ung thư da, vẩy nến hoặc mụn trứng cá, trong khi đó, các nghiên cứu chuyên sâu về hỗ trợ chẩn đoán bệnh viêm da cơ địa vẫn còn hạn chế, đặc biệt đối với dữ liệu bệnh nhân không nhiều ở nước ta [6].

Từ những vấn đề đã được đề cập phía trên, đề tài mà tôi chọn nghiên cứu nhằm tìm hiểu, xây dựng và đánh giá các mô hình học sâu và áp dụng hỗ trợ chẩn đoán bệnh viêm da cơ địa từ bộ dữ liệu hình ảnh da liễu. Kết quả của đề tài có ý nghĩa thực tiễn có thể hỗ trợ bác sĩ trong nâng cao độ chính xác trong công tác chẩn đoán lâm sàng bệnh.

Mục đích nghiên cứu

Đề tài trên được thực hiện nhằm nghiên cứu, tìm hiểu, thử nghiệm các mô hình học sâu sẵn có trong bài toán hỗ trợ chẩn đoán bệnh y khoa, cụ thể là bệnh viêm da cơ địa dựa trên bộ hình ảnh da liễu. Qua đó, đề tài hướng tới việc xây dựng một hệ thống hỗ trợ có khả năng nhận diện và phân loại tổn thương da với độ chính xác cao, góp phần hỗ trợ bác sĩ trong quá trình chẩn đoán và điều trị bệnh.

Đề tài được viết và tập trung vào một số mục đích sau.

Thứ nhất, nghiên cứu tổng quan về bệnh viêm da cơ địa, các phương pháp chẩn đoán bệnh được sử dụng hiện nay và các nghiên cứu về việc áp dụng một số ứng dụng của trí tuệ nhân tạo vào lĩnh vực y khoa, cụ thể là lĩnh vực da liễu.

Hai là, tìm hiểu cơ sở lý thuyết của học sâu và các mô hình mạng neural tích chập và các mô hình CNN hiện đại và ứng dụng của các mô hình đó trong xử lý ảnh bệnh ngoài da.

Ba là, thu thập dữ liệu và tiền xử lý dữ liệu về ảnh bệnh ngoài da trước khi đưa vào huấn luyện và đánh giá mô hình.

Bốn là, xây dựng mô hình, đánh giá mô hình thông qua các hàm mất mát cơ bản, chỉnh sửa các tham số, qua các kết quả đó so sánh độ hiệu quả của các mô hình đã chọn.

Năm là, đưa ra kết luận, đề xuất mô hình tối ưu trong bài toán mà đề tài cần giải quyết.

Nội dung nghiên cứu

Với mục tiêu đã đề ra của đề tài nghiên cứu, luận văn tập trung nghiên cứu các nội dung chính sau.

Một là nghiên cứu tổng quan về bệnh viêm da cơ địa, bao gồm nguyên nhân, triệu chứng, phương pháp chẩn đoán và điều trị hiện nay. Đồng thời khảo sát các nghiên cứu hiện có ở trong và ngoài nước có liên quan đến lĩnh vực da liễu.

Hai là nghiên cứu cơ sở lý thuyết về học máy, học sâu và các mô hình mạng neural tích chập trong xử lý ảnh, đặc biệt là các mô hình CNN hiện đại đã được chứng minh về tính ứng dụng của nó trong y khoa đặc biệt là các bài toán phân loại ảnh.

Ba là thu thập và xây dựng tập dữ liệu hình ảnh bệnh viêm da cơ địa và các bệnh ngoài da khác. Thực hiện các bước tiền xử lý dữ liệu như chuẩn hóa ảnh, tăng cường dữ liệu và phân chia dữ liệu phục vụ quá trình huấn luyện và đánh giá mô hình.

Bốn là xây dựng và triển khai các mô hình học sâu được huấn luyện dựa trên bộ dữ liệu ảnh về da liễu nhằm đóng góp giải quyết bài toán phát hiện bệnh viêm da cơ địa.

Năm là huấn luyện, đánh giá và so sánh các mô hình thông qua các chỉ số như Accuracy, Precision, Recall, F1-score nhằm tìm ra mô hình có hiệu năng tốt nhất.

Sáu là phân tích kết quả thực nghiệm, đánh giá ưu điểm và hạn chế của từng mô hình, từ đó đề xuất mô hình phù hợp cho bài toán hỗ trợ chẩn đoán bệnh viêm da cơ địa.

Cuối cùng là công bố kết quả và đề xuất định hướng phát triển và khả năng ứng dụng hệ thống trong thực tế nhằm hỗ trợ bác sĩ trong quá trình chẩn đoán bệnh về da liễu.

Cơ sở khoa học

Lĩnh vực trí tuệ nhân tạo là cơ sở khoa học để thực hiện đề tài này, ngoài ra học máy và học sâu trong xử lý ảnh số, trong đó mạng neural tích chập là một mô hình vô cùng đặc biệt. Mô hình học sâu như CNN có khả năng trích xuất đặc trưng từ dữ liệu hình ảnh và đã chứng minh hiệu quả cao trong nhiều bài toán nhận dạng và phân loại ảnh y khoa [7].

Trong lĩnh vực da liễu, hình ảnh tổn thương da chứa nhiều đặc trưng quan trọng phục vụ cho việc chẩn đoán bệnh [8]. Việc ứng dụng các mô hình học sâu vào phân tích hình ảnh da liễu cho phép tự động nhận diện các dấu hiệu bất thường, các đặc trưng của bệnh. Nhiều nghiên cứu trên thế giới đã cho thấy các mô hình học sâu như hiện đại như ResNet, Efficient Net có hiệu năng cao trong lĩnh vực xử lý ảnh, đặc biệt là trong y khoa [9].

Ngoài ra, đề tài còn dựa trên các cơ sở khoa học liên quan đến xử lý dữ liệu ảnh, sử dụng các chỉ số như Accuracy, Precision, Recall và F1-score để tối ưu hoá mô hình. Đây là những nền tảng quan trọng để xây dựng hệ thống đánh giá hỗ trợ chuẩn đoán bệnh viêm da cơ địa.

Tính thực tiễn của đề tài

Viêm da cơ địa là bệnh da liễu phổ biến, có xu hướng gia tăng và ảnh hưởng lớn đến sức khỏe cũng như chất lượng cuộc sống của người bệnh [3]. Tuy nhiên, việc chẩn đoán hiện nay vẫn phụ thuộc nhiều vào kinh nghiệm của bác sĩ chuyên khoa, trong khi số lượng bác sĩ da liễu còn hạn chế, đặc biệt tại các cơ sở y tế tuyến dưới [5].

Việc nghiên cứu và xây dựng mô hình học sâu có ý nghĩa thực tiễn cao, góp phần hỗ trợ bác sĩ trong việc phát hiện bệnh da liễu như Viêm da cơ địa nhanh chóng và chính xác. Hệ thống có thể được sử dụng như một công cụ hỗ trợ sàng lọc ban đầu, giảm áp lực cho đội ngũ y tế và nâng cao hiệu quả khám chữa bệnh.

Bên cạnh đó, đề tài còn góp phần thúc đẩy ứng dụng trí tuệ nhân tạo trong lĩnh vực y tế tại Việt Nam, tạo tiền đề cho việc phát triển các hệ thống hỗ trợ chẩn đoán thông minh đối với các bệnh da liễu khác trong tương lai.

Những đóng góp của luận văn

Luận văn đã tổng hợp và phân tích một cách có hệ thống các cơ sở lý thuyết liên quan đến học sâu, xử lý ảnh y khoa và các phương pháp ứng dụng trí tuệ nhân tạo trong lĩnh vực da liễu. Đặc biệt, nghiên cứu tập trung vào bài toán hỗ trợ chẩn đoán bệnh viêm da cơ địa, qua đó làm rõ vai trò và tiềm năng của các kỹ thuật học sâu trong việc nâng cao hiệu quả chẩn đoán bệnh từ hình ảnh da liễu.

Bên cạnh đó, luận văn đã xây dựng một quy trình xử lý dữ liệu hoàn chỉnh, bao gồm các bước thu thập dữ liệu, tiền xử lý ảnh và tăng cường dữ liệu. Quy trình này góp phần nâng cao chất lượng bộ dữ liệu, giúp cải thiện khả năng học và hiệu quả hoạt động của các mô hình trong quá trình huấn luyện.

Trên cơ sở đó, nghiên cứu đã tiến hành khảo sát, triển khai và đánh giá một số mô hình học sâu tiêu biểu cho bài toán nhận diện và phân loại bệnh viêm da cơ địa từ hình ảnh da liễu. Việc thực nghiệm trên các mô hình khác nhau giúp đánh giá được khả năng ứng dụng của từng phương pháp trong điều kiện dữ liệu thực tế.

Ngoài ra, luận văn thực hiện so sánh hiệu năng giữa các mô hình dựa trên các chỉ số đánh giá phổ biến như Accuracy, Precision, Recall và F1-score. Kết quả so sánh là cơ sở để xác định mô hình phù hợp nhất đối với bài toán nghiên cứu, đồng thời cung cấp những nhận định khách quan về ưu điểm và hạn chế của từng mô hình.

Từ các kết quả đạt được, luận văn đã đề xuất hướng ứng dụng mô hình học sâu trong việc xây dựng hệ thống hỗ trợ chẩn đoán bệnh viêm da cơ địa. Hệ thống này có tiềm năng hỗ trợ các bác sĩ trong quá trình khám và chẩn đoán bệnh, góp phần nâng cao chất lượng dịch vụ y tế cũng như giảm tải áp lực trong công tác chuyên môn.

Chương 1: TỔNG QUAN NGHIÊN CỨU

1.1. Tổng quan về bệnh Viêm da cơ địa

Viêm da cơ địa là một bệnh viêm da mạn tính, tái phát nhiều lần và thường xuất hiện ở mọi lứa tuổi, đặc biệt phổ biến ở trẻ em [3, 4]. Bệnh đặc trưng bởi tình trạng da khô, ngứa kéo dài, nổi mẩn đỏ và có thể xuất hiện các tổn thương dạng chàm trên nhiều vùng khác nhau của cơ thể. Đây là một trong những bệnh da liễu phổ biến nhất hiện nay và có xu hướng gia tăng trong những năm gần đây [3].

Nguyên nhân gây bệnh viêm da cơ địa khá phức tạp, liên quan đến nhiều yếu tố như di truyền, rối loạn miễn dịch, môi trường sống và các tác nhân kích thích từ bên ngoài [4]. Người mắc bệnh thường có hàng rào bảo vệ da suy yếu, khiến da dễ bị kích ứng và nhiễm khuẩn [4]. Ngoài ra, các yếu tố như thời tiết, bụi bẩn, thực phẩm, hóa chất hoặc căng thẳng tâm lý cũng có thể làm bệnh tái phát hoặc trở nên nghiêm trọng hơn [4].

Việc chẩn đoán bệnh viêm da cơ địa hiện nay chủ yếu dựa trên khám lâm sàng và kinh nghiệm của bác sĩ chuyên khoa da liễu. Bác sĩ thường đánh giá dựa trên vị trí tổn thương, đặc điểm tổn thương da, tiền sử dị ứng và các triệu chứng đi kèm [4]. Tuy nhiên, trong nhiều trường hợp, biểu hiện của bệnh có thể tương đồng với các bệnh da liễu khác [8] như vẩy nến, viêm da tiếp xúc hoặc nấm da, gây khó khăn trong việc chẩn đoán chính xác. Đặc biệt là trong bối cảnh các bác sĩ có chuyên môn cao còn hạn chế ở các cơ sở y tế tuyến dưới.

Sự phát triển của trí tuệ nhân tạo, đặc biệt là học sâu, đã mở ra nhiều hướng nghiên cứu mới trong hỗ trợ chẩn đoán bệnh da liễu thông qua hình ảnh [1]. Việc học và trích xuất những đặc trưng từ dữ liệu ảnh được các mô hình học sâu học tự động từ đó đạt hiệu quả cao trong các bài toán phân loại ảnh y khoa [2, 7]. Điều này cho thấy tiềm năng lớn trong việc ứng dụng các mô hình học sâu để hỗ trợ nhận diện và chuẩn đoán bệnh viêm da cơ địa.

Do đó, việc nghiên cứu các mô hình học sâu phục vụ hỗ trợ chuẩn đoán bệnh viêm da cơ địa có ý nghĩa quan trọng cả về mặt khoa học và thực tiễn, góp phần nâng cao hiệu quả khám chữa bệnh và thúc đẩy ứng dụng trí tuệ nhân tạo trong lĩnh vực y tế.

1.2. Tổng quan về trí tuệ nhân tạo và học sâu trong lĩnh vực Y tế

1.2.1. Trí tuệ nhân tạo

Trí tuệ nhân tạo là lĩnh vực khoa học máy tính nghiên cứu việc xây dựng các hệ thống có khả năng mô phỏng trí thông minh của con người. Trong những năm gần đây, AI đã được ứng dụng rộng rãi trong nhiều lĩnh vực như tài chính, giáo dục, giao thông và đặc biệt là y tế [1].

Trong lĩnh vực y tế, trí tuệ nhân tạo được sử dụng nhằm hỗ trợ bác sĩ trong việc phân tích dữ liệu, chẩn đoán bệnh, dự đoán nguy cơ và đưa ra các quyết định điều trị [1]. AI có khả năng xử lý lượng lớn dữ liệu y khoa trong thời gian ngắn, từ đó giúp nâng cao độ chính xác và giảm tải áp lực cho đội ngũ y tế [2].

Một trong những ứng dụng nổi bật của AI trong y tế là phân tích ảnh y khoa. Các hệ thống AI có thể hỗ trợ phát hiện tổn thương bất thường trên ảnh X-quang, CT, MRI hoặc ảnh da liễu với độ chính xác cao [2]. Bệnh đã được phát hiện lâm sàng một cách nhanh chóng hơn trước và hiệu quả hơn trước rất nhiều.

1.2.2. Học máy và học sâu

Học máy là một nhánh của trí tuệ nhân tạo cho phép máy tính học từ dữ liệu và cải thiện hiệu suất mà không cần lập trình tường minh cho từng trường hợp cụ thể [7]. Trong học máy, mô hình được huấn luyện dựa trên dữ liệu để nhận diện quy luật và đưa ra dự đoán hoặc phân loại.

Học sâu là một nhánh của học máy, sử dụng các mạng neural nhân tạo nhiều lớp để học đặc trưng dữ liệu ở mức độ phức tạp cao hơn. So với các phương pháp truyền thống, học sâu có khả năng tự động trích xuất đặc trưng từ dữ liệu mà không cần thiết kế thủ công [7].

Trong xử lý ảnh, các mô hình học sâu, đặc biệt là mạng neural tích chập, đã đạt được nhiều thành tựu vượt trội trong các bài toán nhận dạng và phân loại hình ảnh. CNN có khả năng học các đặc trưng quan trọng của ảnh như cạnh, hình dạng, màu sắc và cấu trúc tổn thương, giúp nâng cao hiệu quả phân loại ảnh y khoa [7].

1.2.3. Chẩn đoán bệnh da liễu sử dụng học sâu

Các bệnh da liễu thường được chẩn đoán thông qua quan sát trực tiếp tổn thương trên da. Vì vậy, hình ảnh da liễu đóng vai trò rất quan trọng trong quá trình chẩn đoán. Với sự phát triển của học sâu, nhiều nghiên cứu đã áp dụng các mô hình CNN [5] để nhận diện và phân loại bệnh da từ hình ảnh .

Một số mô hình học sâu phổ biến như AlexNet [10], VGGNet [11], ResNet [12], DenseNet [13], Efficient Net [9] và Vision Transformer đã được sử dụng trong nhiều nghiên cứu về phân loại bệnh các bệnh ngoài da [5]. Các mô hình này cho thấy khả năng đạt độ chính xác cao, thậm chí trong một số trường hợp có thể đạt hiệu quả tương đương bác sĩ chuyên khoa.

Tuy nhiên, việc ứng dụng học sâu trong hỗ trợ chẩn đoán bệnh viêm da cơ địa vẫn còn gặp nhiều khó khăn do hạn chế về dữ liệu, sự đa dạng của tổn thương da và điều kiện chụp ảnh khác nhau [3,4]. Do đó, việc nghiên cứu và đánh giá các mô hình học sâu phù hợp cho bài toán này là cần thiết nhằm nâng cao hiệu quả ứng dụng trong thực tế.

1.3. Các mô hình học sâu trong xử lý y khoa

1.3.1. Tổng quan về xử lý ảnh y khoa bằng học sâu

Xử lý ảnh y khoa là một trong những lĩnh vực ứng dụng quan trọng của trí tuệ nhân tạo trong y tế [2]. Các ảnh y khoa như ảnh X-quang, CT, MRI, siêu âm hay ảnh da liễu chứa nhiều thông tin quan trọng phục vụ cho việc phát hiện và chẩn đoán bệnh.

Tuy nhiên, việc phân tích ảnh theo phương pháp truyền thống thường phụ thuộc nhiều vào kinh nghiệm của bác sĩ chuyên môn và có thể gặp sai sót trong một số trường hợp phức tạp.

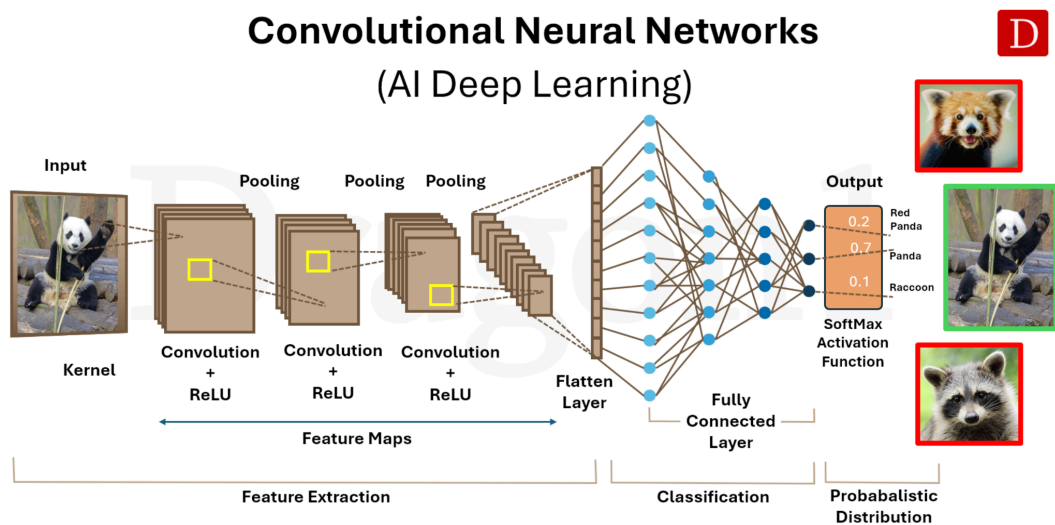
Sự phát triển của học sâu đã tạo ra bước tiến lớn trong lĩnh vực xử lý ảnh y khoa. Khác với các phương pháp xử lý ảnh truyền thống cần thiết kế thủ công đặc trưng ảnh, các mô hình học sâu có khả năng tự động học và trích xuất đặc trưng trực tiếp từ dữ liệu đầu vào [7]. Nhờ đó, mô hình có thể nhận diện các đặc điểm phức tạp của tổn thương mà con người khó quan sát bằng mắt thường.

Trong các bài toán y khoa, học sâu được ứng dụng trong nhiều nhiệm vụ khác nhau bao gồm phân loại ảnh bệnh lý, phát hiện các tổn thương bất thường, phân đoạn vùng tổn thương, hỗ trợ chẩn đoán và dự đoán bệnh, phân tích dữ liệu đa phương thức.

Đối với bài toán bệnh da liễu, hình ảnh tổn thương da đóng vai trò đặc biệt quan trọng vì bác sĩ thường dựa trên hình dạng, kích thước, màu sắc và cấu trúc vùng tổn thương da để đưa ra chẩn đoán [8]. Vì vậy, việc ứng dụng học sâu vào phân tích ảnh da liễu có tiềm năng lớn trong hỗ trợ chẩn đoán tự động.

1.3.2. Mạng neural tích chập

Mạng neural tích chập hay CNN là một trong những mô hình học sâu được sử dụng phổ biến nhất trong xử lý ảnh. CNN được thiết kế đặc biệt để xử lý dữ liệu dạng lưới như hình ảnh và đã đạt hiệu quả cao trong các bài toán nhận dạng đối tượng, phân loại ảnh và xử lý ảnh y khoa [10].



Hình 1.1: Mạng neural tích chập.

Cấu trúc cơ bản của một mạng CNN truyền thống gồm các thành phần chính sau:

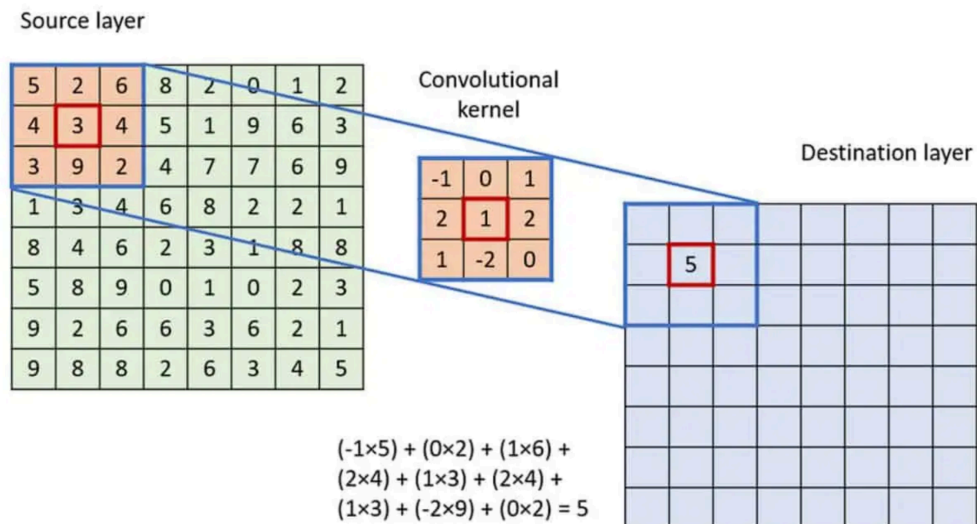
1. Lớp đầu vào (Input Layer)

Nhận dữ liệu ảnh đầu vào dưới dạng ma trận pixel. Hình ảnh đầu vào có thể là nhiều loại ảnh như ảnh xám (ảnh 2 chiều) hoặc ảnh màu (ảnh 3 chiều). Tuy nhiên phải đảm bảo ảnh đầu vào của mô hình phải là ảnh có cùng kích thước (ví dụ ảnh RGB $224 \times 224 \times 3$).

2. Lớp tích chập (Convolution Layer)

Lớp tích chập có nhiệm vụ trích xuất đặc trưng từ ảnh đầu vào bằng cách sử dụng các bộ lọc (Kernel/Filter). Mỗi bộ lọc sẽ quét qua ảnh để phát hiện các đặc trưng như cạnh, góc, màu sắc hoặc kết cấu [10].

Ở các lớp đầu, CNN thường học các đặc trưng đơn giản như đường biên hoặc màu sắc. Ở các lớp sâu hơn, mô hình có thể học được các đặc trưng phức tạp hơn như hình dạng tổng thể hoặc cấu trúc bất thường trên da [10].



Hình 1.2: Lớp tích chập (Convolution Layer)

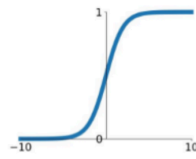
3. Hàm kích hoạt (Activation Function)

Activation Function là hàm giúp mô hình học được các đặc trưng phi tuyến của hình ảnh da liễu. Nếu không có activation function thì dù mạng có nhiều layer, nó vẫn chỉ giống một phép biến đổi tuyến tính lớn [10].

Nói đơn giản, để tăng khả năng biểu diễn phi tuyến của mô hình thì dữ liệu thường được đưa qua hàm này. Hàm ReLU là hàm kích hoạt phổ biến nhất trong mạng neural tích chập do có khả năng giảm hiện tượng mất gradient và giúp mô hình học nhanh hơn.

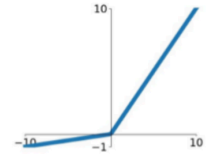
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



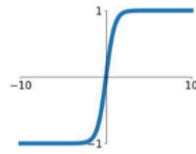
Leaky ReLU

$$\max(0.1x, x)$$



tanh

$$\tanh(x)$$

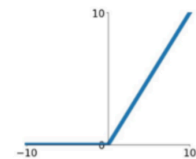


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

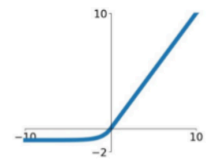
ReLU

$$\max(0, x)$$



ELU

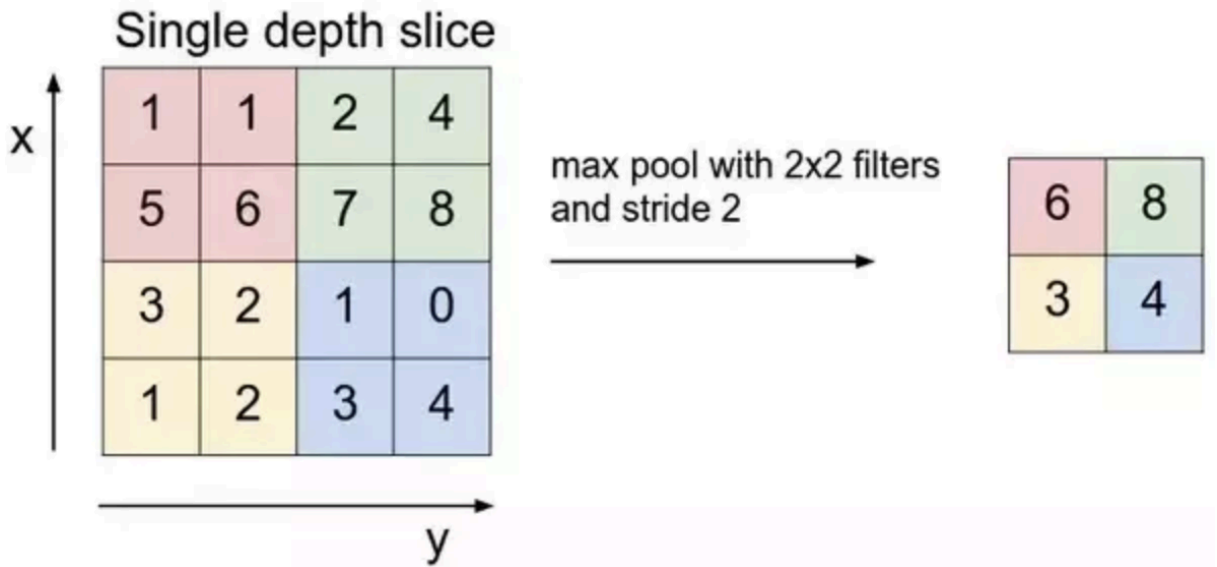
$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



Hình 1.3: Các hàm kích hoạt phổ biến (Activation Function)

4. Lớp gộp (Pooling Layer)

Lớp này giúp giảm kích thước dữ liệu đầu ra, từ đó giảm số lượng tham số và chi phí tính toán. Đồng thời, chúng ta thường muốn giữ lại các đặc trưng quan trọng của ảnh và hạn chế hiện tượng quá khớp. Hai phương pháp Pooling phổ biến gồm Max Pooling và Average Pooling [10].



Hình 1.4: Lớp gộp (Pooling Layer)

5. Lớp chuẩn hóa (Batch Normalization Layer)

Chuẩn hóa dữ liệu giữa các lớp giúp mô hình học ổn định và nhanh hơn. Lớp này còn giúp giảm hiện tượng dịch chuyển phân phối dữ liệu nội bộ, từ đó cải thiện khả năng hội tụ của mạng [14].

6. Lớp loại bỏ ngẫu nhiên (Dropout Layer)

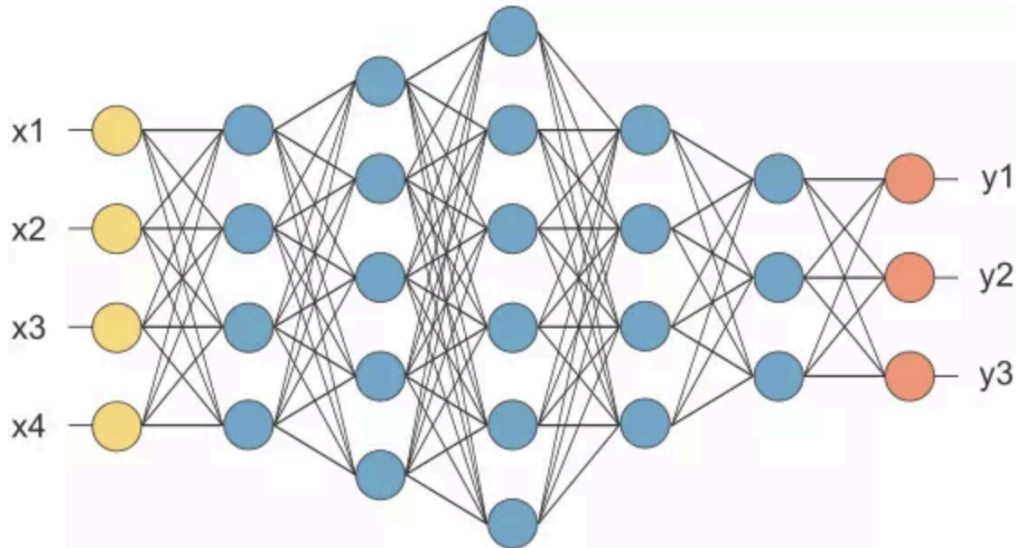
Một số neuron sẽ được lớp này loại bỏ ngẫu nhiên trong huấn luyện để giảm hiện tượng quá khớp. Mục đích khiến cho mô hình tránh bị phụ thuộc quá vào một số neuron đa số, làm tăng tính tổng quát hoá của mô hình [15].

7. Lớp làm phẳng (Flatten Layer)

Chuyển ma trận đặc trưng 2D/3D thành vector 1D trước khi đưa vào lớp kết nối đầy đủ. Quá trình này giúp kết nối các đặc trưng đã được trích xuất với các lớp phân loại ở cuối mạng để đưa ra dự đoán.

8. Lớp kết nối đầy đủ (Fully Connected Layer)

Sau quá trình trích xuất đặc trưng, dữ liệu được đưa vào các lớp này để thực hiện phân loại. Hàm Softmax thường được dùng ở đây để đưa ra xác suất dự đoán cho từng lớp bệnh [10].



Hình 1.5: Lớp kết nối đầy đủ (Fully Connected Layer)

9. Lớp đầu ra (Output Layer)

Đưa ra kết quả cuối cùng, thường sử dụng các hàm Softmax cho phân loại nhiều lớp hoặc Sigmoid cho phân loại nhị phân.

1.3.3. Một số mô hình CNN phổ biến

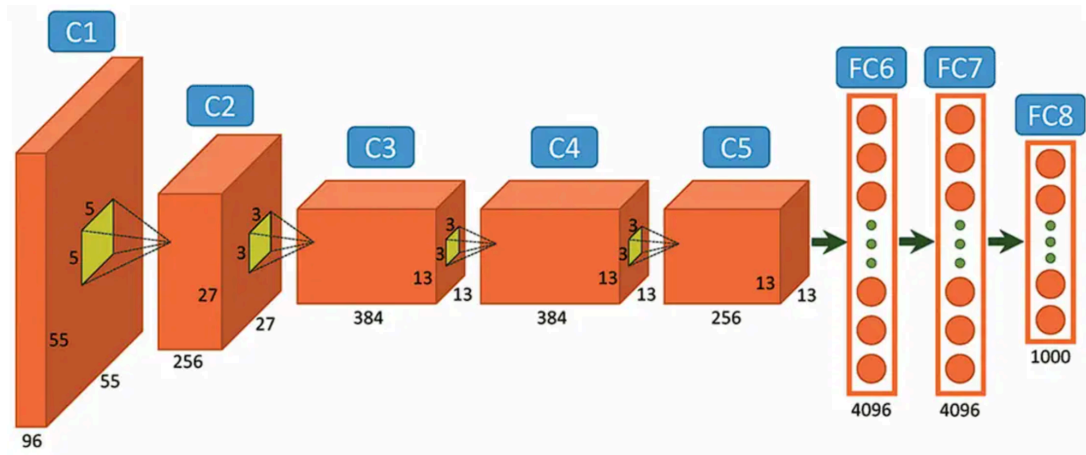
Hạn chế trong quá trình huấn luyện khi số lớp ảnh quá sâu vẫn là yếu điểm của mô hình mạng neural tích chập truyền thống, hiện tượng mất gradient, số lượng tham số lớn và yêu cầu tài nguyên tính toán cao [7]. Ngoài ra, khả năng tái sử dụng đặc trưng giữa các lớp còn hạn chế, trên các hình ảnh y khoa phức tạp sẽ bị giảm tính hiệu quả.

Để khắc phục những nhược điểm này, nhiều kiến trúc CNN cải tiến đã được đề xuất và đạt hiệu quả cao hơn trong bài toán phân loại và nhận dạng ảnh. Một số mô hình CNN phổ biến hiện đại hiện nay bao gồm AlexNet, VGGNet, ResNet, DenseNet và Efficient Net.

1. AlexNet

AlexNet là một trong những mô hình CNN nổi bật đầu tiên đạt thành công lớn trong cuộc thi ImageNet năm 2012. Mô hình gồm nhiều lớp tích chập và lớp kết nối đầy đủ, sử dụng hàm ReLu và kỹ thuật Dropout nhằm giảm hiện tượng quá khớp [10].

AlexNet đánh dấu bước phát triển quan trọng của học sâu trong xử lý ảnh và mở đường cho nhiều kiến trúc CNN hiện đại sau này.



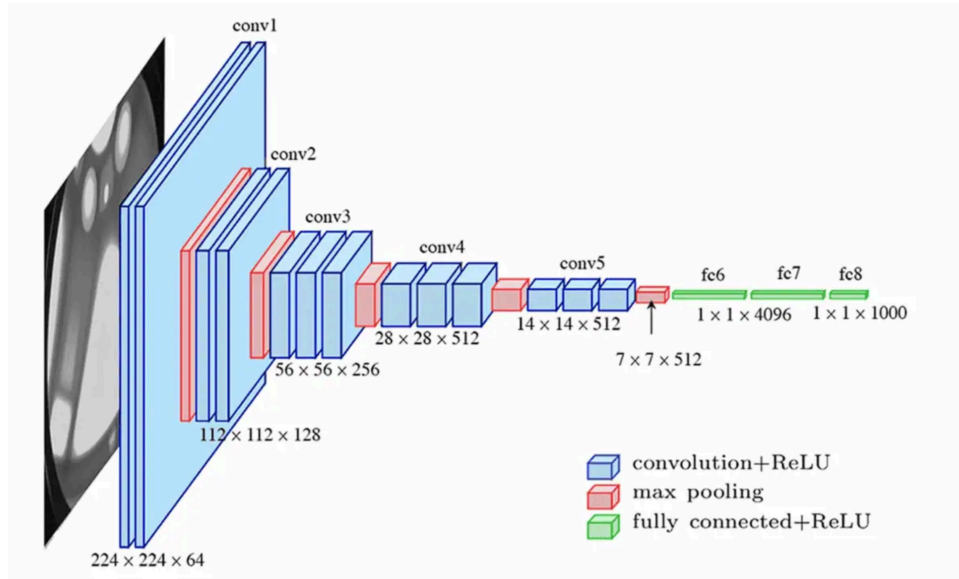
Hình 1.6: Mô hình AlexNet

Hình 6 minh họa kiến trúc của mô hình AlexNet gồm 5 lớp tích chập từ C1 đến C5 và 3 lớp kết nối đầy đủ từ FC6 đến FC8. So với mô hình CNN truyền thống, AlexNet được cải tiến theo hướng xây dựng mạng sâu hơn với số lượng lớp lớn hơn, sử dụng hàm kích hoạt ReLu để tăng tốc độ huấn luyện, áp dụng kỹ thuật Dropout nhằm giảm hiện tượng quá khớp và tận dụng GPU để nâng cao hiệu năng tính toán.

2. VGGNet

VGGNet được phát triển với kiến trúc sâu hơn AlexNet, sử dụng nhiều lớp tích chập kích thước nhỏ 3x3. Mô hình có khả năng học đặc trưng tốt hơn nhờ số lượng lớp lớn [11].

Ưu điểm của VGGNet là kiến trúc đơn giản và hiệu quả cao trong phân loại ảnh. Tuy nhiên, mô hình có số lượng tham số lớn nên yêu cầu tài nguyên tính toán cao.



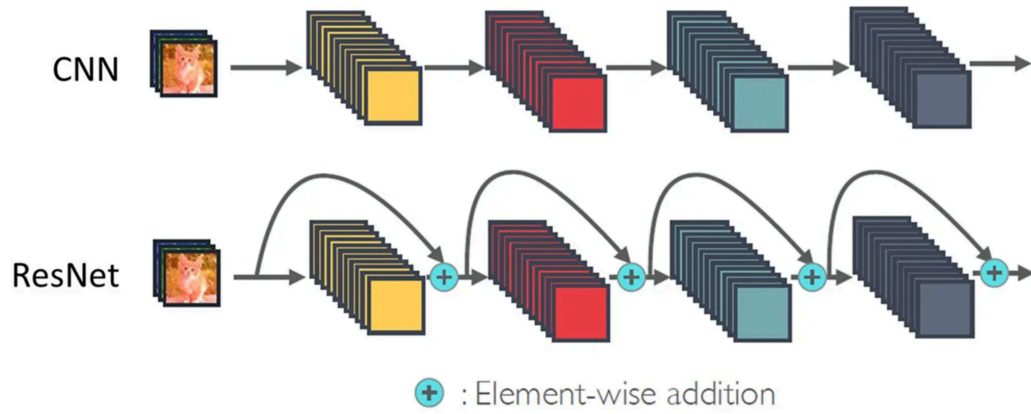
Hình 1.7: Mô hình VGGNet

Hình 7 minh họa kiến trúc của mô hình VGGNet, gồm nhiều lớp tích chập nhỏ kích thước 3×3 được xếp chồng liên tiếp nhằm tăng khả năng học đặc trưng. So với AlexNet, VGGNet được cải tiến theo hướng tăng độ sâu của mạng nhưng sử dụng kernel nhỏ hơn. Sau mỗi nhóm lớp tích chập là lớp Pooling để giảm kích thước dữ liệu, cuối cùng là các lớp Fully Connected dùng cho quá trình phân loại ảnh.

3. ResNet

ResNet được phát triển nhằm giải quyết vấn đề mất gradient khi huấn luyện mạng sâu. Điểm nổi bật của ResNet là sử dụng kết nối tắt, cho phép thông tin truyền trực tiếp giữa các lớp [12].

Nhờ cơ chế này, ResNet có thể xây dựng mạng với hàng chục hoặc hàng trăm lớp mà vẫn đảm bảo hiệu quả huấn luyện. Đây là một trong những mô hình được ứng dụng rộng rãi trong xử lý ảnh y khoa hiện nay.



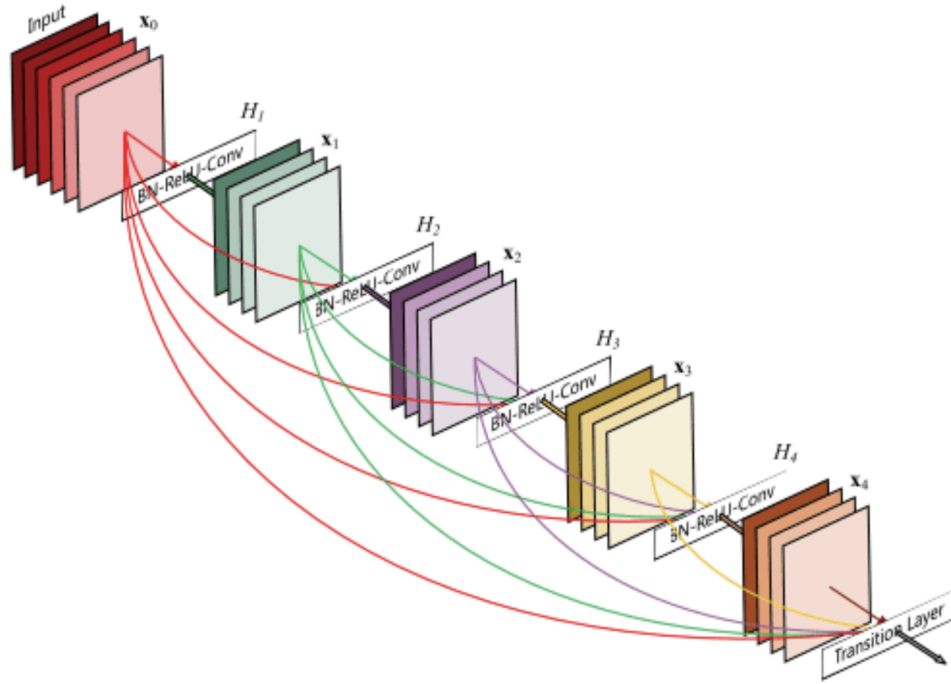
Hình 1.8: Mô hình ResNet

Hình 8 minh họa cơ chế hoạt động của ResNet so với CNN truyền thống. Trong CNN, dữ liệu được truyền tuần tự qua các lớp nên khi mạng quá sâu dễ xảy ra hiện tượng mất gradient, làm giảm hiệu quả huấn luyện.

ResNet khắc phục vấn đề này bằng cách sử dụng kết nối tắt, cho phép thông tin từ lớp trước được truyền trực tiếp đến lớp sau thông qua phép cộng phần tử. Nhờ đó, mô hình duy trì khả năng lan truyền gradient tốt hơn, giúp huấn luyện các mạng rất sâu với hiệu quả cao và cải thiện độ chính xác trong bài toán phân loại ảnh.

4. DenseNet

DenseNet kết nối tất cả các lớp với nhau nhằm tăng khả năng truyền thông tin và tái sử dụng đặc trưng. Điều này giúp mô hình học hiệu quả hơn và giảm số lượng tham số cần huấn luyện [13].



Hình 1.9: Mô hình DenseNet

Hình 9 minh họa kiến trúc của mô hình DenseNet, trong đó mỗi lớp đều được kết nối trực tiếp với tất cả các lớp phía trước thông qua cơ chế dense connection. Khác với CNN chỉ truyền dữ liệu tuần tự qua từng lớp, DenseNet cho phép tái sử dụng đặc trưng đã học bằng cách ghép các feature map của những lớp trước vào đầu vào của lớp hiện tại [13].

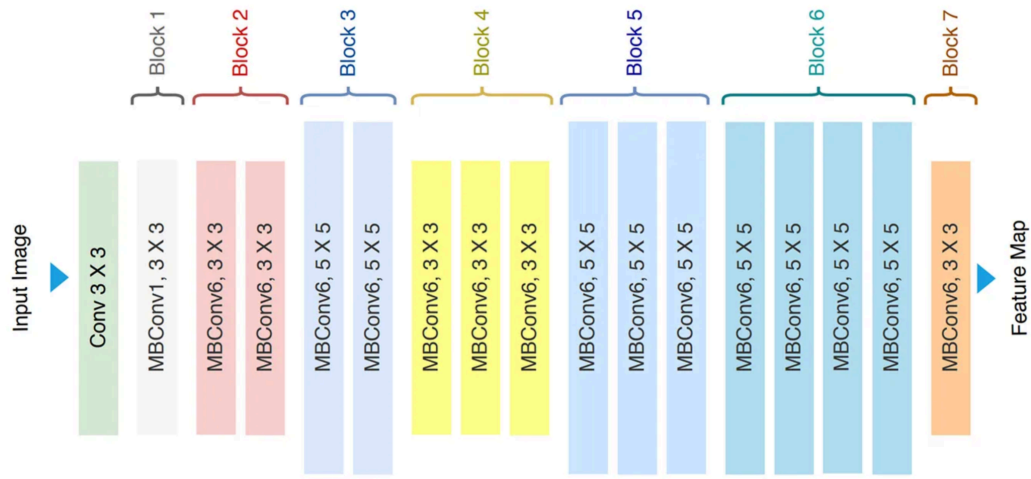
Mỗi khối trong mạng thường bao gồm Batch Normalization, hàm kích hoạt ReLU và lớp tích chập nhằm trích xuất đặc trưng hiệu quả hơn. So với CNN truyền thống, DenseNet được cải tiến theo hướng tăng khả năng lan truyền thông tin và gradient, giúp giảm hiện tượng mất gradient, đồng thời giảm số lượng tham số cần huấn luyện nhờ tái sử dụng đặc trưng.

Ngoài ra, mô hình còn sử dụng Transition Layer để giảm kích thước dữ liệu và số lượng feature map giữa các dense block. Nhờ khả năng học đặc trưng hiệu quả và tối ưu tham số, DenseNet đạt hiệu suất cao trong các bài toán nhận dạng và phân loại ảnh

5. EfficientNet

EfficientNet là mô hình tối ưu hóa đồng thời chiều sâu, chiều rộng và độ phân giải ảnh đầu vào nhằm đạt hiệu suất cao với số lượng tham số thấp hơn.

Mô hình này phù hợp với các bài toán yêu cầu độ chính xác cao nhưng hạn chế tài nguyên tính toán, đặc biệt trong các ứng dụng y tế thực tế [9].



Hình 1.10: Mô hình EfficientNet

Hình 10 minh họa kiến trúc của mô hình EfficientNet, bao gồm nhiều khối MBConv được sắp xếp theo từng block với kích thước kernel 3×3 và 5×5 [9]. Mỗi MBConv block sử dụng cơ chế mở rộng kênh, tích chập depthwise và kết nối tắt nhằm giảm số lượng tham số nhưng vẫn đảm bảo khả năng học đặc trưng hiệu quả.

So với CNN, EfficientNet được cải tiến theo hướng tối ưu đồng thời độ sâu của mạng, chiều rộng của các lớp và độ phân giải đầu vào thông qua phương pháp compound scaling, giúp mô hình đạt độ chính xác cao hơn với chi phí tính toán thấp hơn.

Ngoài ra, EfficientNet sử dụng các MBConv block kế thừa từ MobileNet nhằm tăng hiệu quả xử lý và giảm tài nguyên phần cứng cần thiết. Nhờ khả năng tối ưu hiệu năng và tốc độ, EfficientNet được sử dụng nhiều trong các bài toán phân loại ảnh.

Bảng 1.1: Bảng so sánh các mô hình CNN ứng dụng trong lĩnh vực Y tế

Tiêu chí	AlexNet	VGGNet	ResNet	DenseNet	EfficientNet
Năm ra mắt	2012	2014	2015	2017	2019
Ý tưởng chính	CNN sâu đầu tiên đạt hiệu quả cao trên ImageNet	Tăng độ sâu bằng nhiều kernel nhỏ 3×3	Thêm skip connection để giải quyết mất gradient	Kết nối mọi layer với tất cả layer trước đó	Tối ưu đồng thời depth, width và resolution
Kiến trúc nổi bật	Conv lớn + Max Pooling + FC	Nhiều Conv 3×3 xếp chồng	Residual Block	Dense Block	MBCConv Block + Compound Scaling
Cải tiến so với CNN truyền thống	Chứng minh mạng sâu học đặc trưng tốt hơn	Kernel nhỏ giúp học đặc trưng chi tiết hơn	Gradient truyền tốt hơn khi mạng rất sâu	Tái sử dụng feature giữa các layer	Hiệu năng cao nhưng ít tham số
Cơ chế kết nối đặc biệt	Không có	Không có	Skip Connection (cộng)	Dense Connection (concatenate)	Skip Connection + MBCConv
Độ sâu mạng	Khoảng 8 layer	16–19 layer	Có thể >100 layer	Rất sâu nhưng ít tham	Thay đổi tùy phiên bản

				số	
Số lượng tham số	Nhiều	Rất nhiều	Trung bình	Ít hơn ResNet/VGG	Tối ưu, khá ít
Ưu điểm	Đặt nền móng cho Deep CNN	Kiến trúc đơn giản, hiệu quả	Huấn luyện mạng sâu rất tốt	Tái sử dụng đặc trưng hiệu quả	Chính xác cao, tiết kiệm tài nguyên
Nhược điểm	Tham số lớn, dễ overfit	Rất nặng, tốn bộ nhớ	Kiến trúc phức tạp hơn VGG	Tốn RAM do nhiều kết nối	Thiết kế phức tạp
Ứng dụng phổ biến	Phân loại ảnh cơ bản	Nhận dạng ảnh, transfer learning	Computer Vision hiện đại	Medical Imaging, nhận dạng ảnh	Thiết bị di động, AI hiệu năng cao
Đặc điểm nhận diện	Kernel lớn 11×11	Toàn Conv 3×3	Có shortcut	Nhiều kết nối chéo giữa layer	MBCConv và scaling
Mức độ tối ưu	Thấp theo chuẩn hiện nay	Trung bình	Cao	Cao	Rất cao

1.3.4. Thách thức trong ứng dụng học sâu cho ảnh da liễu

Mặc dù các mô hình học sâu đã đạt được nhiều kết quả tích cực trong lĩnh vực xử lý ảnh y khoa [2], việc ứng dụng chúng vào bài toán nhận diện và phân loại bệnh da liễu vẫn phải đối mặt với nhiều thách thức.

Một trong những khó khăn lớn nhất là việc thu thập dữ liệu ảnh y khoa do liên quan trực tiếp đến quyền riêng tư và bảo mật thông tin của bệnh nhân [2]. Bên cạnh đó, số lượng ảnh về bệnh viêm da cơ địa còn tương đối hạn chế [3, 4].

Ngoài vấn đề về số lượng dữ liệu, chất lượng hình ảnh cũng là một yếu tố ảnh hưởng đáng kể đến hiệu quả của mô hình. Hình ảnh da liễu thường bị tác động bởi nhiều yếu tố như điều kiện ánh sáng, góc chụp, khoảng cách chụp và chất lượng của thiết bị ghi hình [8]. Những khác biệt này có thể làm thay đổi đặc điểm trực quan của tổn thương da, gây khó khăn cho quá trình nhận diện và phân loại tự động.

Một thách thức khác là nhiều bệnh da liễu có biểu hiện lâm sàng tương đối giống nhau, dẫn đến sự chồng lấn về đặc trưng hình ảnh giữa các nhóm bệnh [8]. Điều này làm tăng độ phức tạp của bài toán phân loại và đòi hỏi mô hình phải có khả năng học được các đặc trưng tinh vi để phân biệt chính xác giữa các bệnh lý khác nhau.

Ngoài ra, các bộ dữ liệu ảnh trong lĩnh vực da liễu thường tồn tại tình trạng mất cân bằng dữ liệu nghiêm trọng, khi số lượng mẫu giữa các nhóm bệnh có sự chênh lệch lớn [2]. Đồng thời, quá trình thu thập, xác thực và gán nhãn dữ liệu cũng đòi hỏi nhiều thời gian, chi phí và sự tham gia của các chuyên gia y tế, khiến việc xây dựng các bộ dữ liệu chất lượng cao trở nên khó khăn hơn [2].

1.3.5. Ý nghĩa của việc nghiên cứu

Việc nghiên cứu và ứng dụng học sâu trong lĩnh vực y tế nói chung, cũng như trong lĩnh vực da liễu nói riêng, mang lại nhiều ý nghĩa quan trọng cả về mặt khoa học và thực tiễn. Đối với bài toán chẩn đoán bệnh viêm da cơ địa, các mô hình học sâu có khả năng hỗ trợ bác sĩ trong việc phát hiện và nhận diện bệnh một cách nhanh chóng và chính xác hơn thông qua việc phân tích hình ảnh da liễu [1,2]. Điều này góp phần nâng cao hiệu quả chẩn đoán và giảm thiểu nguy cơ sai sót trong quá trình đánh giá tình trạng bệnh.

Bên cạnh đó, việc ứng dụng các hệ thống hỗ trợ chẩn đoán dựa trên học sâu giúp rút ngắn thời gian phân tích [1] và xử lý dữ liệu hình ảnh y khoa. Thay vì phải thực hiện hoàn toàn bằng phương pháp thủ công, bác sĩ có thể nhận được các kết quả đánh giá sơ bộ [1] từ hệ thống, từ đó tập trung nhiều hơn vào việc đưa ra quyết định điều trị phù hợp cho bệnh nhân.

1.4. Tổng quan các nghiên cứu liên quan

1.4.1. Các nghiên cứu trên thế giới

Trong những năm gần đây, cùng với sự phát triển mạnh mẽ của trí tuệ nhân tạo và học sâu, nhiều nghiên cứu trên thế giới đã tập trung ứng dụng các mô hình học sâu trong lĩnh vực y khoa [2]. Các phương pháp này chủ yếu khai thác dữ liệu hình ảnh tổn thương da nhằm xây dựng các hệ thống hỗ trợ tự động cho bác sĩ trong việc phát hiện và phân loại bệnh da liễu.

Những nghiên cứu tiêu biểu phải kể đến là công trình của Esteva và cộng sự tại Stanford University. Trong nghiên cứu này, mô hình CNN được huấn luyện trên tập dữ liệu gồm hơn 129.000 hình ảnh da liễu thuộc nhiều loại bệnh khác nhau. Kết quả thực nghiệm hiệu suất của mô hình tương đương với các bác sĩ chuyên khoa trong việc phát hiện ung thư da. Nghiên cứu khẳng định tiềm năng ứng dụng của học sâu trong lĩnh vực y tế, mở ra nhiều hướng nghiên cứu mới trong chẩn đoán tự động các bệnh lý về da [5].

Codella và cộng sự vào năm 2018 đã nghiên cứu các phương pháp học sâu trên bộ dữ liệu ISIC (International Skin Imaging Collaboration) nhằm giải quyết bài toán phân loại tổn thương da. Nó tập trung vào việc so sánh nhiều kiến trúc CNN khác nhau như ResNet và Inception, đồng thời cho thấy các mô hình học sâu vượt trội hơn so với các phương pháp học máy truyền thống trong việc nhận diện các bệnh da liễu từ ảnh [16].

Nhiều kiến trúc mạng hiện đại như ResNet [12], DenseNet [13], EfficientNet [9] và Vision Transformer đã được áp dụng. Các đặc trưng phức tạp của ảnh là một trong những điều khó nhưng nó được các mô hình này học tự động với độ chính xác cao hơn nhiều so với các mô hình truyền thống.

Trong khi đó, một số nghiên cứu tập trung vào việc cải thiện chất lượng dữ liệu đầu vào thông qua các kỹ thuật tiền xử lý dữ liệu gồm xoay ảnh, lật ảnh, điều

chỉnh độ sáng, cắt vùng tổn thương, chuẩn hoá màu sắc ảnh [2].

Các kỹ thuật này giúp tăng khả năng tổng quát hóa của mô hình và giảm hiện tượng quá khớp khi dữ liệu huấn luyện hạn chế. Ngoài ra, nhiều nghiên cứu còn kết hợp học sâu với Transfer Learning [10] nhằm tận dụng các mô hình đã được huấn luyện trước trên tập dữ liệu lớn như ImageNet [17].

Phần lớn các nghiên cứu quốc tế tập trung nhiều vào các bệnh như ung thư da, u hắc tố, vẩy nến, mụn trứng cá, nấm da [3,4]. Trong khi đó, các nghiên cứu chuyên sâu về hỗ trợ chuẩn đoán bệnh viêm da cơ địa vẫn còn ít.

1.4.2. Các nghiên cứu tại Việt Nam

Tại Việt Nam, nghiên cứu ứng dụng trí tuệ nhân tạo trong y tế đang ngày càng được quan tâm, đặc biệt trong các lĩnh vực xử lý ảnh y khoa và hỗ trợ chẩn đoán bệnh. Một số trường đại học, viện nghiên cứu và cơ sở y tế đã triển khai các nghiên cứu liên quan đến ứng dụng học máy và học sâu trong phân tích ảnh y khoa.

Trong lĩnh vực da liễu, các nghiên cứu bước đầu đã áp dụng CNN để nhận diện một số bệnh ngoài da thông qua hình ảnh. Tuy nhiên, do hạn chế về tập dữ liệu y khoa và điều kiện nghiên cứu, số lượng công trình chuyên sâu vẫn chưa nhiều.

Các khó khăn phổ biến trong nghiên cứu tại Việt Nam bao gồm thiếu tập dữ liệu ảnh da liễu có chất lượng cao và được gán nhãn đầy đủ, dữ liệu phân tán tại nhiều cơ sở y tế khác nhau, điều kiện chụp ảnh chưa đồng nhất và hạn chế về tài nguyên tính toán, hạn chế về hệ thống lưu trữ dữ liệu.

Ngoài ra, nhiều nghiên cứu mới chỉ dừng lại ở mức thử nghiệm mô hình cơ bản mà chưa đánh giá toàn diện khả năng ứng dụng thực tế trong môi trường y tế.

Đối với bệnh viêm da cơ địa, hiện nay số lượng nghiên cứu ứng dụng học sâu tại Việt Nam còn ít [3,4]. Điều này cho thấy đây là một hướng nghiên cứu còn nhiều tiềm năng phát triển và có ý nghĩa thực tiễn cao.

1.4.3. Những vấn đề còn tồn tại trong các nghiên cứu hiện nay

Mặc dù đã đạt được nhiều kết quả tích cực, các nghiên cứu ứng dụng học sâu trong những bài toán phân loại ảnh vẫn còn tồn tại một số hạn chế:

1. Hạn chế về dữ liệu

Các mô hình học sâu thường yêu cầu lượng dữ liệu lớn [7] để đạt hiệu quả cao. Tuy nhiên, dữ liệu ảnh y khoa thường khó thu thập [2] do liên quan đến quyền riêng tư của bệnh nhân và yêu cầu chuyên môn trong quá trình gán nhãn dữ liệu.

Đối với bệnh viêm da cơ địa, dữ liệu hiện có thường chưa đủ đa dạng về độ tuổi bệnh nhân, giai đoạn bệnh, loại da, điều kiện ánh sáng, góc chụp ảnh.

2. Hiện tượng mất cân bằng dữ liệu

Trong các tập dữ liệu y khoa đã thu thập được, số lượng ảnh giữa các nhóm bệnh thường không đồng đều [2]. Một số lớp bệnh có rất ít dữ liệu, khiến mô hình dễ thiên lệch khi huấn luyện và làm giảm hiệu quả phân loại.

3. Độ phức tạp của tổn thương da

Các tổn thương viêm da cơ địa có thể thay đổi theo thời gian và biểu hiện khác nhau giữa các bệnh nhân [4]. Ngoài ra, nhiều bệnh da liễu có đặc điểm tương đồng [8] khiến mô hình khó phân biệt chính xác.

4. Khả năng ứng dụng thực tế

Nhiều nghiên cứu đạt kết quả tốt trong môi trường thử nghiệm nhưng chưa được triển khai thực tế do các vấn đề liên quan đến tốc độ xử lý, khả năng hoạt động trên thiết bị cấu hình thấp [9], độ ổn định của mô hình, tính minh bạch trong quá trình dự đoán.

1.4.4. Định hướng nghiên cứu của đề tài

Xuất phát từ những hạn chế trong các nghiên cứu trước đây, đề tài tập trung nghiên cứu và đánh giá các mô hình học sâu phù hợp cho bài toán hỗ trợ chuẩn đoán bệnh viêm da cơ địa.

Đề tài hướng tới xây dựng quy trình tiền xử lý các dữ liệu về ảnh sao cho đạt hiệu quả cao nhất trước khi đưa vào mô hình. Sử dụng một số kỹ thuật tăng cường

dữ liệu hình ảnh giúp cải thiện chất lượng dữ liệu trước khi huấn luyện. Thử nghiệm và đánh giá và so sánh tính hiệu quả của nhiều mô hình học sâu khác nhau. So sánh hiệu năng giữa các mô hình. Cuối cùng là đề xuất mô hình phù hợp có khả năng ứng dụng thực tế.

Kết quả nghiên cứu kỳ vọng sẽ góp phần nâng cao hiệu quả hỗ trợ chẩn đoán các bệnh y khoa đồng thời thúc đẩy trí tuệ nhân tạo trong lĩnh vực da liễu.

1.5. Kết luận chương 1

Chương 1 trình bày tổng quan về bệnh viêm da cơ địa, bao gồm đặc điểm, nguyên nhân, triệu chứng và những khó khăn trong quá trình chẩn đoán bệnh. Bên cạnh đó, luận văn cũng giới thiệu tổng quan về trí tuệ nhân tạo, học máy và học sâu trong lĩnh vực y tế, đặc biệt là mô hình mạng neural tích chập được ứng dụng rộng rãi trong xử lý ảnh y khoa.

Chương này đã phân tích một số mô hình học sâu phổ biến như AlexNet, VGGNet, ResNet, DenseNet và EfficientNet, đồng thời trình bày những ưu điểm và hạn chế của các mô hình trong bài toán phân loại ảnh da liễu. Các nghiên cứu trong nước và ngoài nước liên quan đến lĩnh vực da liễu có sử dụng học sâu được đưa ra khảo sát và đánh giá.

Qua quá trình tổng quan, có thể nhận thấy rằng việc ứng dụng học sâu trong hỗ trợ chẩn đoán bệnh là hướng nghiên cứu có tiềm năng lớn, tuy nhiên vẫn còn tồn tại nhiều khó khăn như hạn chế dữ liệu, sự đa dạng của tổn thương da và yêu cầu cao về độ chính xác mô hình. Đây cũng chính là cơ sở để luận văn tiếp tục nghiên cứu, xây dựng và đánh giá các mô hình học sâu phù hợp trong các chương tiếp theo.

Chương 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng nghiên cứu

Đối tượng chính được tập trung khảo sát là các mô hình học sâu được ứng dụng trong các bài toán y tế qua dữ liệu hình ảnh da liễu. Các mô hình học sâu có khả năng học đặc trưng tự động sau đó đưa ra kết quả phân loại bệnh với độ chính xác và hiệu quả cao, từ đó hỗ trợ bác sĩ trong quá trình chẩn đoán bệnh.

Đối tượng nghiên cứu của luận văn bao gồm hai thành phần chính là dữ liệu ảnh da liễu và các mô hình học sâu xử lý ảnh.

2.1.1. Đối tượng dữ liệu nghiên cứu

Dữ liệu nghiên cứu là các hình ảnh tổn thương da liên quan đến bệnh viêm da cơ địa. Đây là loại dữ liệu đóng vai trò quan trọng trong quá trình huấn luyện và đánh giá mô hình học sâu.

Các hình ảnh da liễu thường chứa nhiều thông tin quan trọng bao gồm màu sắc vùng tổn thương, hình dạng tổn thương, mức độ viêm đỏ, tình trạng bong tróc da, mật độ tổn thương trên bề mặt da [4].

Thông qua việc phân tích các đặc điểm này, các đặc trưng của bệnh viêm da cơ địa sẽ được mô hình học và phân biệt với các trường hợp da bình thường hoặc các bệnh da khác [2, 5].

Dữ liệu ảnh sử dụng trong nghiên cứu này được lấy từ bộ dữ liệu “20 Skin Diseases Dataset” trên Kaggle [18].

Bộ dữ liệu bao gồm các hình ảnh đã được gán nhãn thuộc nhiều nhóm bệnh da liễu khác nhau như:

1. Acne and Rosacea
2. Atopic Dermatitis
3. Eczema
4. Psoriasis
5. Melanoma Skin Cancer
6. Lupus
7. Urticaria
8. Vasculitis

9. Warts và các bệnh nhiễm virus khác

10. Các bệnh nhiễm nấm, vi khuẩn, rối loạn sắc tố da và nhiều bệnh da liễu khác.

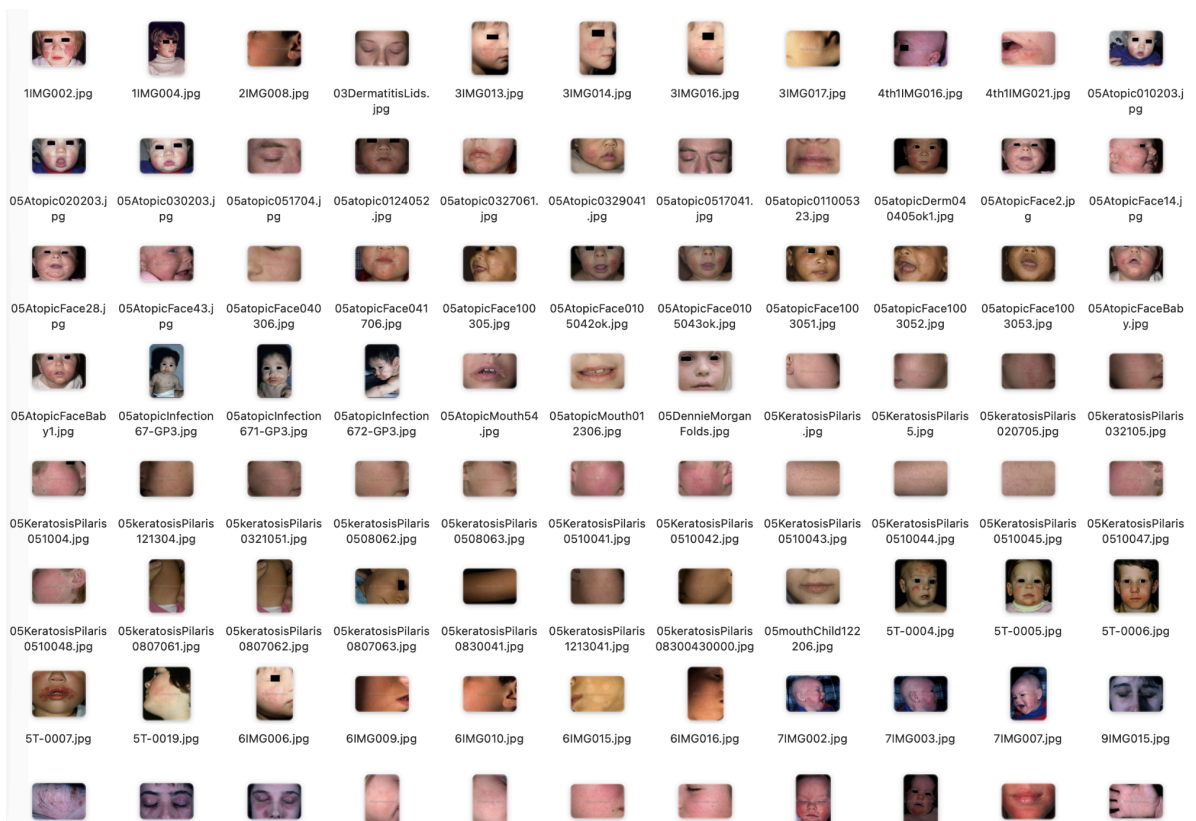
Tổng cộng bộ dữ liệu gồm 19 nhóm bệnh da liễu khác nhau với các hình ảnh được phân loại sẵn theo từng nhãn bệnh. Trong bộ dữ liệu, số lượng ảnh giữa các lớp có sự chênh lệch đáng kể. Một số nhóm bệnh có số lượng ảnh lớn như “Acne and Rosacea Photos” với 840 ảnh, “Actinic Keratosis Basal Cell Carcinoma and other Malignant Lesions” với 322 ảnh và “Light Diseases and Disorders of Pigmentation” với 299 ảnh. Trong khi đó, một số lớp có số lượng ảnh khá ít như “Urticaria Hives” chỉ có 3 ảnh, “Vasculitis Photos” có 15 ảnh và “Exanthema and Drug Eruptions” có 20 ảnh.

Đối với bệnh viêm da cơ địa, nhóm dữ liệu “Atopic Dermatitis Photos” gồm 100 ảnh da liễu đã được gán nhãn. Các hình ảnh này chứa nhiều đặc điểm đặc trưng của bệnh như vùng da viêm đỏ, bong tróc và tổn thương bề mặt da, giúp mô hình học sâu có thể học và nhận diện đặc trưng bệnh hiệu quả hơn.

Sự chênh lệch số lượng ảnh giữa các lớp cho thấy bộ dữ liệu có hiện tượng mất cân bằng dữ liệu. Hướng đến hiệu quả huấn luyện mô hình và độ chính xác phân loại giữa các nhóm bệnh. Vì vậy, trong quá trình nghiên cứu cần xem xét áp dụng các phương pháp tiền xử lý hoặc tăng cường dữ liệu nhằm cải thiện khả năng học của mô hình.

- > Acne and Rosacea Photos
- > Actinic Keratosis Basal Cell Carcinoma and other Malignant Lesions
- > Atopic Dermatitis Photos
- > Cellulitis Impetigo and other Bacterial Infections
- > Eczema Photos
- > Exanthems and Drug Eruptions
- > Herpes HPV and other STDs Photos
- > Light Diseases and Disorders of Pigmentation
- > Lupus and other Connective Tissue diseases
- > Melanoma Skin Cancer Nevi and Moles
- > Poison Ivy Photos and other Contact Dermatitis
- > Psoriasis pictures Lichen Planus and related diseases
- > Seborrheic Keratoses and other Benign Tumors
- > Systemic Disease
- > Tinea Ringworm Candidiasis and other Fungal Infections
- > Urticaria Hives
- > Vascular Tumors
- > Vasculitis Photos
- > Warts Molluscum and other Viral Infections

Hình 2.1: Dataset “20 Skin Diseases Dataset” trên Kaggle



Hình 2.2: Dataset ảnh bệnh về Viêm da cơ địa đã được đánh nhãn

Các yếu tố để một bộ dữ liệu đạt chuẩn bao gồm hình ảnh rõ nét, có đầy đủ vùng tổn thương, hạn chế ảnh bị nhiễu hoặc mờ, có độ phân giải phù hợp.

2.1.2. Đối tượng mô hình nghiên cứu

Đối tượng tiếp theo của đề tài là các mô hình học sâu trong lĩnh vực xử lý ảnh, đặc biệt là các mô hình mạng neural tích chập. CNN là nhóm mô hình được sử dụng phổ biến trong các bài toán nhận dạng và phân loại ảnh nhờ khả năng tự động học đặc trưng từ dữ liệu đầu vào [7].

Khác với các phương pháp xử lý ảnh truyền thống cần thiết kế thủ công đặc trưng ảnh, CNN có thể tự động trích xuất các đặc điểm quan trọng như màu sắc, kết cấu, hình dạng và vùng tổn thương da. Chúng giúp mô hình đạt hiệu quả cao trong các bài toán phân loại ảnh.

Tuy nhiên, CNN cơ bản vẫn tồn tại một số hạn chế bao gồm hiệu quả giảm khi mạng quá sâu do hiện tượng mất gradient, số lượng tham số lớn dẫn đến chi phí tính toán cao, dễ xảy ra hiện tượng quá khớp khi dữ liệu huấn luyện không đủ lớn, khả năng tối ưu tài nguyên và tốc độ xử lý còn hạn chế đối với các thiết bị có cấu hình thấp [12].

Để khắc phục các hạn chế trên, đề tài tiến hành nghiên cứu và sử dụng các mô hình CNN hiện đại bao gồm ResNet, DenseNet, MobileNet và EfficientNet.

ResNet sử dụng cơ chế skip connection giúp giảm hiện tượng mất gradient và hỗ trợ huấn luyện mạng rất sâu hiệu quả hơn [12].

DenseNet áp dụng cơ chế kết nối dày đặc giữa các lớp nhằm tái sử dụng đặc trưng, tăng khả năng lan truyền thông tin và giảm số lượng tham số cần huấn luyện [13].

MobileNet được thiết kế tối ưu cho thiết bị di động và hệ thống có tài nguyên hạn chế bằng cách sử dụng depthwise separable convolution để giảm chi phí tính toán [19].

EfficientNet sử dụng phương pháp compound scaling nhằm tối ưu đồng thời độ sâu, chiều rộng và độ phân giải ảnh đầu vào, giúp đạt hiệu suất cao với số lượng tham số thấp hơn nhiều mô hình CNN truyền thống [9].

Việc nghiên cứu và so sánh nhiều mô hình khác nhau giúp đánh giá khả năng

học đặc trưng, hiệu quả phân loại và chi phí tính toán của từng kiến trúc mạng. Lựa chọn được mô hình phù hợp cho đề tài hỗ trợ chẩn đoán bệnh từ ảnh da liễu.

2.1.3. Đối tượng đánh giá nghiên cứu

Ngoài mô hình và dữ liệu, đề tài còn nghiên cứu các phương pháp đánh giá hiệu quả mô hình học sâu. Các chỉ số đánh giá được sử dụng nhằm xác định mức độ chính xác và khả năng tổng quát hóa của mô hình trên dữ liệu mới.

Trước khi trình bày công thức, cần làm rõ các khái niệm trong ma trận nhầm lẫn gồm TP, TN, FP và FN như sau [20]:

1. TP (True Positive): Số mẫu thuộc lớp dương tính và được mô hình dự đoán đúng là dương tính. Ví dụ: ảnh bị viêm da cơ địa và mô hình cũng dự đoán là viêm da cơ địa [20].
2. TN (True Negative): Số mẫu thuộc lớp âm tính và được mô hình dự đoán đúng là âm tính. Ví dụ: ảnh không bị viêm da cơ địa và mô hình dự đoán không bị bệnh [20].
3. FP (False Positive): Số mẫu thuộc lớp âm tính nhưng bị mô hình dự đoán nhầm là dương tính. Ví dụ: ảnh da bình thường nhưng mô hình dự đoán bị viêm da cơ địa [20].
4. FN (False Negative): Số mẫu thuộc lớp dương tính nhưng bị mô hình dự đoán nhầm là âm tính. Ví dụ: ảnh bị viêm da cơ địa nhưng mô hình dự đoán là không bị bệnh [20].

Các chỉ số đánh giá bao gồm:

1. Accuracy

Đây là chỉ số cơ bản nhất dùng để đánh giá thể hiện tỷ lệ dự đoán đúng trên tổng số bộ dữ liệu mẫu [21].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

2. Precision

Precision thể hiện mức độ chính xác của các mẫu được dự đoán là dương tính [21].

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

3. Recall

Recall thể hiện khả năng phát hiện đúng các mẫu dương tính thực tế [21].

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

4. F1-score

F1-score là trung bình điều hòa giữa Precision và Recall [21].

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

Thông qua các chỉ số này, ta có thể đánh giá: Khả năng nhận diện đúng bệnh, khả năng hạn chế dự đoán sai, hiệu quả tổng thể của mô hình.

Accuracy thể hiện tỷ lệ dự đoán đúng được thể hiện trên tổng số bộ dữ liệu đưa vào, là một trong những hàm mất mát cực kỳ đơn giản. Tuy nhiên, chỉ số này có thể không phản ánh chính xác hiệu quả mô hình khi dữ liệu bị mất cân bằng giữa các lớp bệnh [21].

Precision thể hiện tỷ lệ dự đoán đúng trên các mẫu được mô hình dự đoán là dương tính. Chỉ số này quan trọng trong trường hợp cần giảm các dự đoán nhầm bệnh [21].

Recall thể hiện khả năng phát hiện đúng các mẫu thực sự thuộc lớp bệnh. Đây là chỉ số quan trọng trong bài toán y tế vì giúp hạn chế bỏ sót bệnh nhân mắc bệnh [21].

F1-score cân bằng giữa cả hai chỉ số Precision và Recall, giúp đánh giá tổng quát hiệu quả mô hình, cũng là chỉ số được ưa chuộng khi nó giải quyết được những bài toán khi dữ liệu bị mất cân bằng [21].

Trong bộ dữ liệu đang có về nhận diện viêm da cơ địa từ ảnh da, có sự mất cân bằng dữ liệu rất lớn giữa các lớp bệnh. Như lớp bệnh “Acne and Rosacea Photos” có 840 ảnh, trong khi lớp bệnh “Urticaria Hives” chỉ có 3 ảnh và “Atopic Dermatitis

Photos” có 100 ảnh.

Trong đề tài sử dụng chỉ số F1-score làm tiêu chí đánh giá chính cho mô hình học sâu. Nguyên nhân là do bộ dữ liệu có sự mất cân bằng giữa các lớp bệnh, khi một số nhóm có số lượng ảnh rất lớn trong khi nhiều nhóm khác chỉ có số lượng ảnh nhỏ. Trong trường hợp này, Accuracy có thể đạt giá trị cao nhưng không phản ánh chính xác hiệu quả thực tế của mô hình.

F1-score cân bằng Precision và Recall, giúp đánh giá đồng thời khả năng dự đoán đúng và khả năng phát hiện đầy đủ các mẫu bệnh. Chỉ số này đặc biệt phù hợp trong các bài toán y tế vì giúp hạn chế cả dự đoán sai và bỏ sót bệnh.

2.2. Phương pháp nghiên cứu

Để thực hiện đề tài, luận văn sử dụng kết hợp phương pháp nghiên cứu lý thuyết và phương pháp thực nghiệm. Quá trình nghiên cứu được triển khai theo từng bước giữ tính khoa học, tính khách quan và khả năng đánh giá chính xác hiệu quả của các mô hình học sâu.

Phương pháp nghiên cứu của luận văn tập trung vào việc thu thập dữ liệu, xử lý dữ liệu ảnh, xây dựng mô hình học sâu, huấn luyện mô hình và đánh giá kết quả thực nghiệm.

2.2.1. Phương pháp nghiên cứu lý thuyết

Phương pháp nghiên cứu lý thuyết được sử dụng nhằm xây dựng cơ sở khoa học cho đề tài. Trong quá trình này, luận văn tiến hành thu thập, tổng hợp và phân tích các tài liệu liên quan đến bệnh viêm da cơ địa, trí tuệ nhân tạo, học sâu và xử lý ảnh y khoa.

1. Nghiên cứu tổng quan về bệnh viêm da cơ địa

Luận văn tiến hành tìm hiểu các kiến thức liên quan đến bệnh viêm da cơ địa như khái niệm bệnh, nguyên nhân gây bệnh, triệu chứng lâm sàng, đặc điểm tổn thương da [3].

Việc nghiên cứu bệnh lý giúp xác định các đặc điểm quan trọng trên hình ảnh da liễu phục vụ cho bài toán phân loại ảnh [4].

2. Nghiên cứu về trí tuệ nhân tạo, học máy và học sâu

Luận văn nghiên cứu cơ sở lý thuyết của trí tuệ nhân tạo, học máy và học sâu nhằm hiểu rõ nguyên lý hoạt động của các mô hình xử lý ảnh [1].

Các nội dung được nghiên cứu bao gồm: Cấu trúc mạng neural nhân tạo, nguyên lý huấn luyện mô hình, thuật toán các mô hình CNN hiện đại, hàm mất mát, các phương pháp tối ưu mô hình [7].

Ngoài ra, luận văn cũng nghiên cứu vai trò của học sâu trong xử lý ảnh y khoa và khả năng ứng dụng trong lĩnh vực da liễu.

3. Nghiên cứu các mô hình học sâu trong xử lý ảnh

Luận văn tiến hành khảo sát các kiến trúc CNN hiện đại và phổ biến được sử dụng trong phân loại ảnh bao gồm: ResNet [12], DenseNet [13], MobileNet [18], EfficientNet [9].

Mỗi mô hình được nghiên cứu về: Kiến trúc mạng, Số lượng lớp, Cơ chế hoạt động, Ưu điểm và hạn chế, Khả năng áp dụng cho dữ liệu ảnh y khoa.

Thông qua quá trình nghiên cứu lý thuyết, luận văn lựa chọn các mô hình phù hợp để triển khai thực nghiệm.

2.2.2. Phương pháp thu thập dữ liệu

Dữ liệu là thành phần quan trọng quyết định hiệu quả của mô hình học sâu. Vì vậy, luận văn thực hiện quá trình thu thập và lựa chọn dữ liệu ảnh da liễu phục vụ nghiên cứu.

Các dữ liệu được sử dụng bao gồm hình ảnh liên quan đến bệnh viêm da cơ địa và các hình ảnh phục vụ đối chiếu phân loại.

Quá trình thu thập dữ liệu được thực hiện thông qua: Các bộ dữ liệu công khai, nguồn dữ liệu nghiên cứu học thuật và dữ liệu hình ảnh da liễu trực tuyến phù hợp với nghiên cứu.

Sau khi thu thập, dữ liệu được kiểm tra nhằm loại bỏ: Ảnh mờ, ảnh có chất lượng thấp, ảnh không rõ vùng tổn thương, ảnh trùng lặp. Việc làm sạch dữ liệu giúp nâng cao chất lượng huấn luyện mô hình.

Trong nghiên cứu này, bộ dữ liệu “20 Skin Diseases Dataset” trên Kaggle được lựa chọn do chứa nhiều hình ảnh da liễu đã được gán nhãn sẵn theo từng nhóm bệnh, thuận lợi cho quá trình huấn luyện và đánh giá mô hình học sâu.

Bộ dữ liệu bao gồm nhiều loại bệnh da khác nhau, trong đó có nhóm “Atopic Dermatitis Photos” phù hợp với mục tiêu nghiên cứu bệnh viêm da cơ địa, dễ dàng phân biệt với các dữ liệu ảnh về các bệnh ngoài da khác. Ngoài ra, dữ liệu ở dạng ảnh thực tế giúp các mô hình CNN học được các đặc trưng quan trọng như màu sắc, hình dạng và mức độ tổn thương da.

Bên cạnh đó, đây là bộ dữ liệu công khai, dễ truy cập và phù hợp cho việc nghiên cứu, thử nghiệm và so sánh hiệu quả giữa các mô hình học sâu trong bài toán của đề tài.

2.2.3. Phương pháp tiền xử lý dữ liệu

Dữ liệu ảnh ban đầu thường có sự khác biệt về kích thước, độ sáng và chất lượng ảnh. Vì vậy, trước khi đưa vào mô hình học sâu, dữ liệu cần được tiền xử lý nhằm đảm bảo tính đồng nhất.

Các bước tiền xử lý được thực hiện bao gồm:

1. Chuẩn hóa kích thước ảnh

Nhằm phù hợp với đầu vào của mô hình các hình ảnh sẽ được tiền xử lý về cùng kích thước. Việc chuẩn hóa kích thước giúp mô hình xử lý dữ liệu hiệu quả hơn. Các hình ảnh trong đề tài này được đưa về dạng: 224×224 [12].

Kích thước 224×224 được lựa chọn vì đây là kích thước đầu vào tiêu chuẩn của nhiều mô hình CNN phổ biến như ResNet [12], DenseNet [13], MobileNet [9] và EfficientNet [8]. Việc sử dụng kích thước này giúp tối ưu các mô hình đang nghiên cứu, đồng thời đảm bảo giữ được các đặc trưng quan trọng của ảnh da liễu mà vẫn tối ưu thời gian huấn luyện và tài nguyên tính toán.

2. Chuẩn hóa giá trị các điểm ảnh

Được đưa về khoảng phù hợp nhằm giúp mô hình học ổn định hơn. Thông thường, giá trị ảnh RGB được chuẩn hóa trong khoảng: $0 \leq x \leq 1$ hoặc $-1 \leq x \leq 1$.

3. Loại bỏ nhiễu dữ liệu

Một số ảnh có thể chứa nhiễu, ánh sáng không đồng đều hoặc chất lượng thấp. Quá trình tiền xử lý giúp giảm ảnh hưởng của các yếu tố này đến kết quả huấn luyện.

4. Phương pháp tăng cường dữ liệu

Trong bài toán ảnh y khoa, số lượng dữ liệu thường hạn chế. Để tăng khả năng tổng quát hóa của mô hình, luận văn sử dụng các kỹ thuật tăng cường dữ liệu (Data Augmentation).

Các kỹ thuật được áp dụng bao gồm: Xoay ảnh, lật ngang ảnh, lật dọc ảnh, thay đổi độ sáng, dịch chuyển ảnh, thu phóng ảnh [22].

Tăng cường dữ liệu giúp tăng số lượng mẫu huấn luyện, giảm hiện tượng quá khớp và giúp mô hình học được nhiều biến thể dữ liệu khác nhau.

2.2.4. Phương pháp xây dựng mô hình học sâu

Sau quá trình tiền xử lý dữ liệu, ta nghiên cứu và xây dựng một số mô hình học sâu phù hợp với bài toán hiện tại. Quá trình xây dựng mô hình bao gồm lựa chọn kiến trúc mạng phù hợp, thiết lập tham số huấn luyện, xây dựng lớp đầu vào và đầu ra, thiết lập hàm mất mát, lựa chọn thuật toán tối ưu.

Một số kỹ thuật giúp cải thiện khả năng học và giảm hiện tượng quá khớp được sử dụng:

Dropout là một kỹ thuật giảm hiện tượng quá khớp bằng cách vô hiệu hóa ngẫu nhiên một vài neuron trong quá trình huấn luyện [15]. Tăng khả năng tổng quát hóa của mô hình, không bị phụ thuộc quá mức vào một số đặc trưng nhất định.

Batch Normalization là kỹ thuật giúp chuẩn hóa dữ liệu đầu ra giữa các lớp trong mạng neural, từ đó tăng tốc độ hội tụ và giúp quá trình huấn luyện ổn định hơn [14]. Ngoài ra, kỹ thuật này còn hỗ trợ giảm hiện tượng mất gradient khi mạng có nhiều lớp sâu.

Early Stopping là kỹ thuật được sử dụng để ngăn mô hình huấn luyện quá mức trên dữ liệu huấn luyện [23]. Quá trình huấn luyện sẽ dừng lại khi độ chính xác trên tập validation không còn được cải thiện sau một số epoch nhất định, giúp hạn chế overfitting và tiết kiệm thời gian huấn luyện.

2.2.5. Phương pháp huấn luyện mô hình

Trước khi huấn luyện mô hình ta cần chia dữ liệu thành 3 tập:

Tập huấn luyện (Training Set) là tập dữ liệu được sử dụng để huấn luyện mô hình học máy hoặc học sâu. Trong quá trình này, mô hình sẽ học các đặc trưng và mối quan hệ từ dữ liệu đầu vào để tối ưu các tham số bên trong. [24]

Tập kiểm tra (Validation Set) là tập dữ liệu được tách ra từ tập huấn luyện nhằm đánh giá mô hình trong quá trình huấn luyện. Tập này giúp điều chỉnh siêu tham số, lựa chọn mô hình tốt nhất và hạn chế hiện tượng quá khớp. [24]

Tập đánh giá (Test Set) là tập dữ liệu hoàn toàn độc lập, chỉ được sử dụng sau khi mô hình đã huấn luyện xong. Từ những dữ liệu mà mô hình chưa từng nhìn thấy đánh giá chính xác hiệu năng tổng quát của mô hình. [24]

Trong bộ dataset “20 Skin Diseases Dataset” của kaggle đã được chia sẵn tập dữ liệu có đánh nhãn gồm train set và test set. Cụ thể, tập train bao gồm 2609 ảnh và tập test bao gồm 693 ảnh. Trong đó, lớp viêm da cơ địa (Atopic Dermatitis) có 100 ảnh trong tập train và 24 ảnh trong tập test, tổng cộng 124 ảnh. Các lớp bệnh da liễu khác chiếm phần lớn dữ liệu với 2509 ảnh trong tập train và 669 ảnh trong tập test.

Tập validation không được cung cấp sẵn, do đó trong nghiên cứu này, tập train được tiếp tục chia tách thành tập huấn luyện và tập validation theo tỷ lệ 80% cho tập huấn luyện và 20% cho tập validation, nhằm phục vụ quá trình tinh chỉnh siêu tham số và đánh giá mô hình trong quá trình huấn luyện.

Trong quá trình huấn luyện có nhiều epoch nhằm tối ưu trọng số mạng nơ-ron. Trong quá trình huấn luyện, luận văn theo dõi chỉ số F1-score. Quá trình này giúp lựa chọn mô hình có hiệu năng tốt nhất.

2.2.6. Phương pháp đánh giá phân tích và so sánh kết quả

Sau quá trình huấn luyện, các mô hình học sâu được đánh giá và so sánh nhằm xác định mô hình phù hợp nhất cho bài toán chẩn đoán bệnh từ ảnh da liễu.

Trong nghiên cứu này, chỉ số F1-score được sử dụng làm để đánh giá chính về độ hiệu quả của mô hình [20]. Nguyên nhân là do bộ dữ liệu có sự mất cân bằng

giữa các lớp bệnh, khi một số nhóm có số lượng ảnh rất lớn trong khi nhiều nhóm khác có số lượng ảnh thấp. Trong trường hợp này, Accuracy không phản ánh đầy đủ hiệu quả thực tế của mô hình.

F1-score là chỉ số kết hợp giữa Precision và Recall, giúp đánh giá đồng thời khả năng dự đoán đúng và khả năng phát hiện đầy đủ các mẫu bệnh.

Thông qua chỉ số F1-score, đề tài tiến hành so sánh hiệu năng giữa các mô hình như ResNet, DenseNet, MobileNet và EfficientNet. Đánh giá khả năng nhận diện chính xác bệnh viêm da cơ địa và phân biệt với các bệnh da liễu khác. Phân tích khả năng tổng quát hóa của mô hình trên tập dữ liệu kiểm tra. Đánh giá từng kiến trúc mạng về độ chính xác, khả năng học đặc trưng và chi phí tính toán.

Kết quả phân tích giúp xác định mô hình có hiệu quả tốt nhất đối với bài toán phân loại bệnh da liễu. Đồng thời, đây cũng là cơ sở để đề xuất các hướng cải tiến và khả năng ứng dụng mô hình học sâu trong hỗ trợ chẩn đoán bệnh da trong thực tế.

2.3. Kết luận chương 2

Chương 2 đã trình bày đối tượng và phương pháp nghiên cứu được sử dụng trong đề tài. Nội dung chương đã xác định rõ đối tượng nghiên cứu là các mô hình học sâu ứng dụng trong phân tích ảnh da liễu và dữ liệu hình ảnh liên quan đến bệnh viêm da cơ địa.

Bên cạnh đó, chương cũng đã trình bày chi tiết các phương pháp nghiên cứu được áp dụng trong quá trình thực hiện đề tài, bao gồm nghiên cứu lý thuyết, thu thập dữ liệu, tiền xử lý dữ liệu, tăng cường dữ liệu, xây dựng mô hình học sâu, huấn luyện và đánh giá mô hình. F1-score được sử dụng nhằm đánh giá khách quan hiệu quả của mô hình trong bài toán phân loại ảnh da liễu.

Ngoài ra, chương này cũng đã trình bày quy trình thực nghiệm và các kỹ thuật hỗ trợ nhằm nâng cao hiệu quả của mô hình học sâu. Đây là cơ sở quan trọng để triển khai thực nghiệm, phân tích kết quả và đánh giá hiệu quả mô hình trong các chương tiếp theo.

Chương 3. KẾT LUẬN VÀ THẢO LUẬN

3.1. Xây dựng và huấn luyện mô hình học sâu

Trong chương này, luận văn tiến hành xây dựng, huấn luyện và đánh giá các mô hình học sâu phục vụ bài toán hỗ trợ chuẩn đoán bệnh viêm da cơ địa thông qua dữ liệu ảnh da liễu. Quá trình thực nghiệm đánh giá khả năng nhận diện và phân loại bệnh của các mô hình học sâu khác nhau, từ đó lựa chọn mô hình phù hợp nhất cho bài toán nghiên cứu.

3.1.1. Chuẩn bị dữ liệu thực nghiệm

Dữ liệu thực nghiệm đóng vai trò quan trọng trong quá trình xây dựng mô hình học sâu. Chất lượng dữ liệu ảnh ảnh hưởng trực tiếp đến khả năng học và hiệu quả phân loại của mô hình.

1. Thu thập dữ liệu và phân chia dữ liệu

Như đã nói ở Chương 2, bộ dataset “20 Skin Diseases Dataset” của kaggle đã được chia sẵn tập dữ liệu có đánh nhãn gồm train set và test set. Cụ thể, tập train bao gồm 2609 ảnh và tập test bao gồm 693 ảnh. Trong đó, lớp viêm da cơ địa (Atopic Dermatitis) có 100 ảnh trong tập train và 24 ảnh trong tập test, tổng cộng 124 ảnh. Các lớp bệnh da liễu khác chiếm phần lớn dữ liệu với 2509 ảnh trong tập train và 669 ảnh trong tập test.

Tập validation không được cung cấp sẵn, do đó trong nghiên cứu này, tập train được tiếp tục chia tách thành tập huấn luyện và tập validation theo tỷ lệ 80% cho tập huấn luyện và 20% cho tập validation, nhằm phục vụ quá trình tinh chỉnh siêu tham số và đánh giá mô hình trong quá trình huấn luyện.

2. Tiền xử lý dữ liệu ảnh

Do dữ liệu ảnh da liễu được thu thập từ nhiều nguồn khác nhau nên tồn tại sự khác biệt về kích thước, ánh sáng và chất lượng ảnh. Vì vậy, trước khi đưa vào huấn luyện, dữ liệu cần được tiền xử lý nhằm đảm bảo tính đồng nhất và tăng hiệu quả học của mô hình.

Tất cả ảnh được chuyển về cùng kích thước nhằm phù hợp với đầu vào của CNN là (224×224) [12, 13]. Việc chuẩn hóa giúp giảm chi phí tính toán và tăng tính ổn định trong huấn luyện.

Giá trị điểm ảnh RGB được chuẩn hóa về khoảng [0,1] bằng cách chia cho 255 thông qua lớp Rescaling(1./255). Việc chuẩn hóa giúp mô hình hội tụ nhanh hơn và cải thiện độ ổn định trong quá trình huấn luyện [7].

$$x' = \frac{x}{255}$$

Trong đó: x là giá trị điểm ảnh ban đầu, x' là giá trị điểm ảnh sau chuẩn hóa.

Các ảnh không đúng định dạng hoặc không thể đọc được sẽ không được sử dụng nhằm đảm bảo chất lượng dữ liệu đầu vào cho mô hình.

3. Tăng cường dữ liệu ảnh

Do dữ liệu y khoa thường có số lượng hạn chế và mất cân bằng giữa các lớp bệnh, nghiên cứu sử dụng kỹ thuật tăng cường dữ liệu nhằm tăng khả năng tổng quát hóa của mô hình và giảm hiện tượng quá khớp [22].

Các kỹ thuật tăng cường dữ liệu được áp dụng trực tiếp trong quá trình huấn luyện bao gồm: lật ngang ảnh, xoay ảnh ngẫu nhiên, thu phóng ảnh [22].

Các kỹ thuật này giúp tạo ra nhiều biến thể khác nhau của ảnh da liễu nhưng vẫn giữ được đặc trưng của vùng tổn thương. Nhờ đó, mô hình có thể học được nhiều đặc điểm đa dạng hơn và cải thiện hiệu quả phân loại trên dữ liệu thực tế.

3.1.2. Xây dựng và huấn luyện mô hình

Sau khi hoàn thành quá trình chuẩn bị và tiền xử lý dữ liệu, nghiên cứu tiến hành xây dựng và huấn luyện các mô hình học sâu phục vụ bài toán.

1. Lựa chọn mô hình

Trong nghiên cứu này, bốn mô hình học sâu phổ biến trong lĩnh vực xử lý ảnh được lựa chọn bao gồm ResNet [12], DenseNet [13], MobileNet [18], EfficientNet [9].

Các mô hình này đều thuộc nhóm mạng neural tích chập và đã được chứng minh đạt hiệu quả cao trong các bài toán phân loại ảnh y khoa.

Các mô hình được sử dụng dưới dạng pretrained trên tập dữ liệu ImageNet nhằm tận dụng các đặc trưng đã được học trước thông qua phương pháp Transfer Learning [25]. Phần fully connected cuối cùng của các mô hình được thay đổi để phù hợp với bài toán phân loại 19 nhóm bệnh da liễu trong nghiên cứu.

2. Thiết lập tham số

Trong quá trình huấn luyện, các tham số được thiết lập như sau:

- Kích thước ảnh đầu vào: 224×224 pixel
- Batch size: 32
- Số epoch: 10
- Optimizer: Adam [26]
- Loss Function: Sparse Categorical Crossentropy [24]

Hàm mất mát được sử dụng cho bài toán phân loại nhiều lớp. Optimizer Adam được lựa chọn do có khả năng tối ưu nhanh ổn định trong quá trình huấn luyện. Ngoài ra, tập test độc lập được sử dụng để đánh giá mô hình sau khi hoàn thành quá trình huấn luyện.

3. Quá trình huấn luyện

Quá trình huấn luyện được thực hiện bằng cách đưa dữ liệu ảnh qua nhiều epoch để mô hình học các đặc trưng.

Trong mỗi epoch mô hình thực hiện dự đoán nhãn ảnh. Tính toán sai số thông qua hàm mất mát. Cập nhật trọng số bằng thuật toán lan truyền ngược.

Mục tiêu của quá trình huấn luyện là tối thiểu hóa hàm mất mát và nâng cao khả năng phân loại ảnh bệnh da. Sau khi hoàn thành huấn luyện, từng mô hình được đánh giá trên tập test nhằm kiểm tra khả năng tổng quát hóa đối với dữ liệu mới chưa xuất hiện trong quá trình học.

4. Kỹ thuật chống quá khớp

Để giảm hiện tượng quá khớp và tăng khả năng tổng quát hóa của mô hình, nghiên cứu áp dụng một số kỹ thuật bao gồm Dropout [15], Data Augmentation [22], Early Stopping [23].

Các kỹ thuật này giúp các mô hình hoạt động ổn định hơn và nâng cao hiệu quả phân loại ảnh da liễu trên dữ liệu thực tế.

3.2. Đánh giá và thảo luận kết quả thực nghiệm

Sau quá trình huấn luyện, các mô hình được đánh giá nhằm xác định hiệu quả trong bài toán hỗ trợ chẩn đoán bệnh viêm da cơ địa. Hiệu năng mô hình được đánh giá thông qua các chỉ số thống kê phổ biến trong bài toán phân loại ảnh.

Các mô hình ResNet50, DenseNet121, EfficientNetB0, MobileNetV2 được huấn luyện với tập dữ liệu bao gồm 2609 ảnh, trong đó có 100 ảnh viêm da cơ địa, các ảnh có kích thước được chuẩn hoá 240x240px. Việc huấn luyện mô hình được thực hiện với 50 epochs, sử dụng batch size 32. Sau đó mô hình được test trên tập dữ liệu 693 ảnh, trong đó gồm 30 ảnh viêm da cơ địa. Optimizer Adam và Loss Function, Sparse Categorical Crossentropy.

Bảng 3.1: Bảng kết quả thu được của các mô hình sau khi huấn luyện.

	ResNet50	DenseNet121	EfficientNetB0	MobileNetV2
Loss trên tập kiểm thử	0.58	0.83	0.71	1.92
Độ chính xác (accuracy) (%)	68.75	50.00	59.38	12.50
Precision (%)	71.43	61.54	63.16	5.94
Recall (%)	62.50	33.33	50.00	25.00
F1-score (%)	66.67	43.24	55.81	9.60

Kết quả thực nghiệm cho thấy ResNet50 đạt hiệu năng tốt nhất trong số các mô hình được khảo sát với Precision đạt 71.43%, Recall đạt 62.50% và F1-score đạt 66.67%. Các chỉ số này cho thấy mô hình có khả năng nhận diện bệnh viêm da cơ địa tương đối chính xác và cân bằng giữa khả năng phát hiện đúng bệnh và hạn chế các dự đoán sai. Với kiến trúc sử dụng các kết nối tắt, ResNet50 có khả năng trích

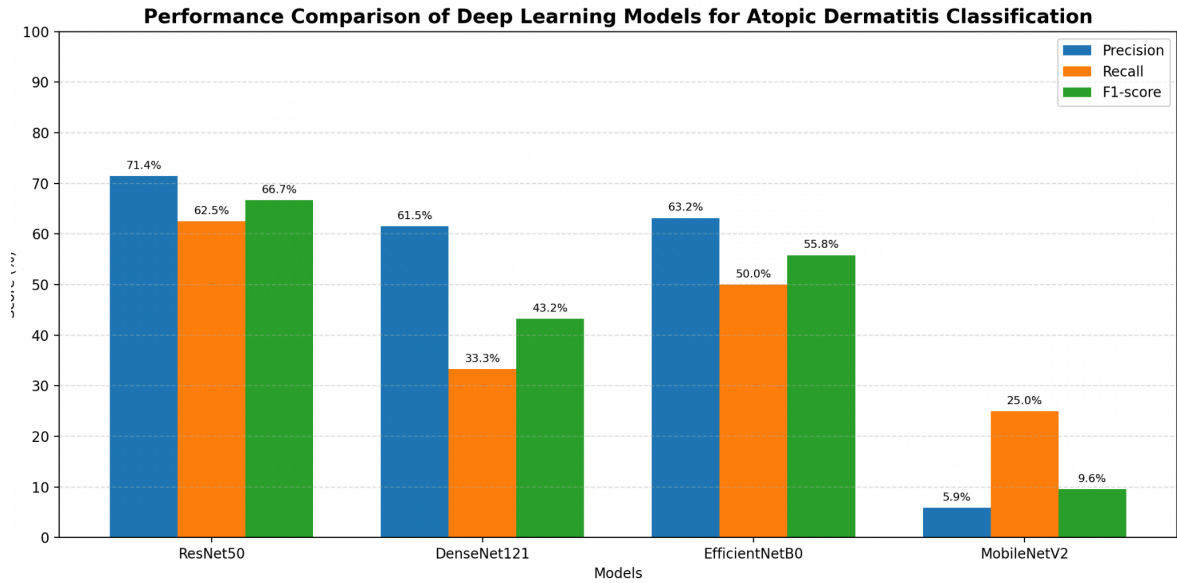
xuất đặc trưng hiệu quả và phù hợp với bài toán phân loại ảnh da liễu trong nghiên cứu này [12].

DenseNet121 đạt Precision 61.54%, Recall 33.33% và F1-score 43.24%. Mặc dù mô hình có khả năng đưa ra các dự đoán tương đối chính xác khi nhận diện được bệnh, nhưng Recall thấp cho thấy mô hình bỏ sót khá nhiều trường hợp mắc bệnh viêm da cơ địa. Kết quả này cho thấy DenseNet121 chưa khai thác hiệu quả các đặc trưng của bộ dữ liệu hiện tại và cần được tối ưu thêm về dữ liệu hoặc siêu tham số để nâng cao hiệu năng.

EfficientNetB0 đạt Precision 63.16, Recall 50.00% và F1-score 55.81%, cho kết quả đứng thứ hai sau ResNet50. Mô hình thể hiện khả năng cân bằng tương đối tốt giữa độ chính xác và khả năng phát hiện bệnh. Với ưu điểm về số lượng tham số ít và hiệu quả tính toán cao, EfficientNetB0 là một lựa chọn phù hợp cho các hệ thống hỗ trợ chẩn đoán cần tối ưu tài nguyên nhưng vẫn đảm bảo chất lượng dự đoán [9].

MobileNetV2 cho kết quả thấp nhất với Precision 5.94%, Recall 2.5% và F1-score 9.60%. Các chỉ số này cho thấy mô hình chưa học được hiệu quả các đặc trưng cần thiết để phân biệt bệnh viêm da cơ địa trên bộ dữ liệu nghiên cứu. Mặc dù MobileNetV2 được thiết kế tối ưu cho các thiết bị có tài nguyên hạn chế, nhưng trong trường hợp này hiệu năng của mô hình chưa đáp ứng được yêu cầu của bài toán chẩn đoán bệnh da liễu [18].

Dựa trên các kết quả thực nghiệm, ResNet50 là mô hình đạt hiệu năng cao nhất và phù hợp nhất cho bài toán nhận diện bệnh viêm da cơ địa trong nghiên cứu này. EfficientNetB0 cũng cho kết quả khả quan và có thể được xem là một lựa chọn thay thế nhờ ưu điểm về hiệu quả tính toán. Trong khi đó, DenseNet121 và đặc biệt là MobileNetV2 cho kết quả thấp hơn đáng kể, cho thấy cần có các phương pháp tối ưu bổ sung hoặc bộ dữ liệu lớn hơn để cải thiện hiệu năng của các mô hình này.



Hình 3.1: Kết quả đánh giá của các mô hình

Kết quả thực nghiệm cho thấy các mô hình học sâu có khả năng học và nhận diện đặc trưng của bệnh viêm da cơ địa từ hình ảnh da liễu với hiệu quả cao.

Các mô hình sâu như ResNet và EfficientNet đạt hiệu năng tốt hơn so với CNN cơ bản nhờ khả năng học đặc trưng sâu hơn, giảm hiện tượng mất gradient, tối ưu số lượng tham số.

Tuy nhiên, quá trình thực nghiệm cũng cho thấy một số khó khăn gồm dữ liệu chưa đủ lớn. Một số ảnh có chất lượng không đồng đều. Biểu hiện tổn thương da đa dạng. Có sự tương đồng giữa các bệnh da liễu khác nhau.

Ngoài ra, mô hình vẫn có thể xuất hiện hiện tượng dự đoán sai trong các trường hợp tổn thương nhẹ, ảnh bị nhiễu, điều kiện ánh sáng không phù hợp.

Mặc dù vậy, kết quả nghiên cứu cho thấy tiềm năng lớn của học sâu trong hỗ trợ chuẩn đoán bệnh viêm da cơ địa và có thể phát triển thành hệ thống hỗ trợ bác sĩ trong tương lai.

3.3. Triển khai mô hình

Sau khi hoàn thành quá trình huấn luyện và lựa chọn được mô hình có hiệu năng tốt nhất, mô hình được triển khai thành một hệ thống hỗ trợ chẩn đoán bệnh viêm da cơ địa nhằm phục vụ người dùng cuối. Mục tiêu của hệ thống là cho phép người dùng tải lên hình ảnh vùng da cần kiểm tra và nhận kết quả dự đoán từ mô hình học sâu một cách nhanh chóng.

Quá trình triển khai bao gồm việc lưu mô hình đã huấn luyện dưới dạng tệp mô hình *.keras*. Tệp mô hình này được nạp vào một máy chủ ứng dụng (Application Server) để thực hiện suy luận (Inference) trên các hình ảnh được người dùng gửi lên.

Kiến trúc tổng thể của hệ thống gồm ba thành phần chính: giao diện người dùng (Frontend), máy chủ ứng dụng (Backend) và mô hình học sâu. Giao diện người dùng được xây dựng dưới dạng website, cho phép người dùng tải ảnh da liễu lên hệ thống. Máy chủ ứng dụng tiếp nhận yêu cầu, thực hiện các bước tiền xử lý ảnh tương tự như trong quá trình huấn luyện, sau đó gửi dữ liệu đến mô hình học sâu để thực hiện dự đoán. Kết quả dự đoán sẽ được trả về giao diện người dùng dưới dạng xác suất hoặc nhãn phân loại.

Các công nghệ như HTML, CSS, JavaScript kết hợp với các framework web như Flask, FastAPI được sử dụng để xây dựng hệ thống. Mô hình học sâu được tích hợp trực tiếp vào Backend thông qua TensorFlow hoặc PyTorch. Đối với các hệ thống có số lượng người dùng lớn, mô hình có thể được đóng gói bằng Docker và triển khai trên các nền tảng điện toán đám mây như Google Cloud, AWS hoặc Microsoft Azure để đảm bảo khả năng mở rộng và tính ổn định.

Về yêu cầu phần cứng, quá trình suy luận của mô hình không đòi hỏi tài nguyên lớn như quá trình huấn luyện. Một máy chủ có cấu hình tối thiểu gồm CPU 4 nhân, RAM 8 GB và dung lượng lưu trữ khoảng 50 GB là có thể đáp ứng nhu cầu sử dụng ở quy mô nhỏ và trung bình. Trong trường hợp cần xử lý đồng thời nhiều yêu cầu hoặc sử dụng các mô hình phức tạp, hệ thống có thể được trang bị thêm GPU để tăng tốc độ dự đoán.

3.4. Kết luận chương 3

Chương 3 đã trình bày quá trình xây dựng, huấn luyện và đánh giá các mô hình học sâu phục vụ bài toán. Các bước thực nghiệm được thực hiện thông qua các bước chuẩn bị dữ liệu, tiền xử lý ảnh, tăng cường dữ liệu và huấn luyện mô hình trên môi trường hỗ trợ GPU.

Luận văn đã tiến hành thử nghiệm và đánh giá một số mô hình học sâu phổ biến như ResNet, DenseNet và EfficientNet. Kết quả thực nghiệm cho thấy các mô hình học sâu có khả năng nhận diện và phân loại bệnh viêm da cơ địa với độ chính xác tương ổn, đặc biệt là các mô hình có kiến trúc sâu và khả năng tối ưu đặc trưng tốt.

Các chỉ số đánh giá như F1-score cho thấy mô hình đạt hiệu quả tích cực trong bài toán phân loại ảnh da liễu. Đồng thời, quá trình thực nghiệm cũng chỉ ra một số khó khăn như hạn chế về dữ liệu, chất lượng ảnh chưa đồng đều và sự tương đồng giữa các bệnh khác nhau.

Mặc dù vẫn còn tồn tại một số hạn chế, kết quả nghiên cứu đã cho thấy tiềm năng ứng dụng của học sâu trong hỗ trợ chuẩn đoán bệnh viêm da cơ địa. Đây là cơ sở quan trọng để tiếp tục nghiên cứu, cải tiến mô hình và phát triển các hệ thống trong tương lai.

KẾT LUẬN VÀ KIẾN NGHỊ

Kết luận

Luận văn đã nghiên cứu và ứng dụng các mô hình học sâu trong bài toán phân loại ảnh da liễu nhằm hỗ trợ chuẩn đoán bệnh viêm da cơ địa. Trong quá trình thực hiện, luận văn đã tiến hành khảo sát cơ sở lý thuyết liên quan đến trí tuệ nhân tạo, học sâu, xử lý ảnh y khoa và các mô hình CNN phổ biến trong phân tích ảnh.

Luận văn đã trình bày tổng quan về bệnh viêm da cơ địa, các đặc điểm tổn thương da và những khó khăn trong quá trình chẩn đoán bệnh. Đồng thời, các nghiên cứu trong và ngoài nước liên quan đến ứng dụng học sâu trong lĩnh vực da liễu cũng đã được phân tích và đánh giá nhằm xác định hướng nghiên cứu phù hợp cho đề tài.

Trong phần thực nghiệm, luận văn đã tiến hành thu thập, tiền xử lý và tăng cường dữ liệu ảnh da liễu phục vụ quá trình huấn luyện mô hình. Một số mô hình học sâu như CNN, ResNet, DenseNet và EfficientNet đã được triển khai và đánh giá thông qua các chỉ số Accuracy, Precision, Recall và F1-score.

Kết quả thực nghiệm cho thấy các mô hình học sâu có khả năng nhận diện và phân loại bệnh viêm da cơ địa với hiệu quả tương đối cao. Đặc biệt, các mô hình như ResNet và EfficientNet cho kết quả tốt hơn nhờ khả năng học đặc trưng hiệu quả và tối ưu hóa quá trình huấn luyện.

Bên cạnh những kết quả đạt được, luận văn vẫn còn một số hạn chế như số lượng dữ liệu chưa lớn, chất lượng dữ liệu chưa hoàn toàn đồng nhất và phạm vi nghiên cứu mới chỉ tập trung vào bài toán phân loại ảnh. Tuy nhiên, kết quả nghiên cứu đã cho thấy tiềm năng lớn của việc ứng dụng học sâu trong hỗ trợ chuẩn đoán bệnh, góp phần nâng cao hiệu quả khám chữa bệnh và thúc đẩy ứng dụng trí tuệ nhân tạo trong lĩnh vực y tế.

Kiến nghị

Để nâng cao hiệu quả và khả năng ứng dụng thực tế của hệ thống hỗ trợ chuẩn đoán bệnh viêm da cơ địa bằng học sâu, luận văn đề xuất một số kiến nghị và hướng phát triển.

Tiếp tục mở rộng và xây dựng tập dữ liệu ảnh da liễu với số lượng lớn hơn, đa dạng hơn về độ tuổi, loại da và mức độ tổn thương nhằm nâng cao khả năng tổng quát hóa của mô hình.

Kết hợp dữ liệu hình ảnh với các dữ liệu lâm sàng khác như triệu chứng, tiền sử bệnh và kết quả xét nghiệm để nâng cao độ chính xác trong hỗ trợ chẩn đoán. Nghiên cứu và áp dụng các mô hình học sâu hiện đại hơn như Vision Transformer hoặc các mô hình kết hợp đa kiến trúc nhằm cải thiện hiệu năng phân loại.

Nghiên cứu các phương pháp giải thích mô hình AI nhằm tăng tính minh bạch và hỗ trợ bác sĩ hiểu rõ cơ sở đưa ra dự đoán của hệ thống. Triển khai thử nghiệm hệ thống trong môi trường thực tế tại các cơ sở y tế nhằm đánh giá khả năng ứng dụng và hiệu quả hỗ trợ chẩn đoán.

Phát triển hệ thống hỗ trợ chuẩn đoán trên nền tảng web hoặc thiết bị di động nhằm tăng khả năng tiếp cận và hỗ trợ người dùng trong thực tế. Tiếp tục nghiên cứu mở rộng ứng dụng học sâu cho các bệnh da liễu khác nhằm xây dựng các hệ thống hỗ trợ chẩn đoán thông minh phục vụ lĩnh vực y tế.

DANH MỤC TÀI LIỆU THAM KHẢO

- [1] E. J. Topol, *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. New York: Basic Books, 2019, pp. 1–25.
- [2] G. Litjens et al., “A Survey on Deep Learning in Medical Image Analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [3] J. I. Silverberg, “Public Health Burden and Epidemiology of Atopic Dermatitis,” *Dermatologic Clinics*, vol. 35, no. 3, pp. 283–289, 2017.
- [4] T. Bieber, “Atopic Dermatitis,” *New England Journal of Medicine*, vol. 358, no. 14, pp. 1483–1494, 2008.
- [5] A. Esteva et al., “Dermatologist-level Classification of Skin Cancer with Deep Neural Networks,” *Nature*, vol. 542, pp. 115–118, 2017.
- [6] N. Brinker et al., “Deep Learning Outperformed 136 Dermatologists in a Head-to-Head Dermoscopic Melanoma Image Classification Task,” *European Journal of Cancer*, vol. 113, pp. 47–54, 2019.
- [7] Y. LeCun, Y. Bengio, and G. Hinton, “Deep Learning,” *Nature*, vol. 521, pp. 436–444, 2015.
- [8] H. Kittler et al., *Diagnostic Accuracy of Dermoscopy*, Springer, 2011.
- [9] M. Tan and Q. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” ICML, 2019.
- [10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 25, pp. 1097–1105, 2012.
- [11] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *International Conference on Learning Representations (ICLR)*, 2015.
- [12] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *CVPR*, pp. 770–778, 2016.
- [13] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely

- Connected Convolutional Networks,” *CVPR*, pp. 4700–4708, 2017.
- [14] S. Ioffe and C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” *ICML*, 2015.
- [15] N. Srivastava et al., “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014.
- [16] N. C. F. Codella et al., “Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging,” *ISBI*, pp. 168–172, 2018.
- [17] J. Deng et al., “ImageNet: A Large-Scale Hierarchical Image Database,” *CVPR*, pp. 248–255, 2009.
- [18] “Ahmad, “20 Skin Diseases Dataset,” Kaggle Dataset, 2023. Available: <https://www.kaggle.com/datasets/haroonalam16/20-skin-diseases-dataset?resource=download>
- [19] A. G. Howard et al., “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” arXiv:1704.04861, 2017.
- [20] T. Fawcett, “An Introduction to ROC Analysis,” *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [21] D. M. Powers, “Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation,” *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.
- [22] L. Perez and J. Wang, “The Effectiveness of Data Augmentation in Image Classification using Deep Learning,” arXiv:1712.04621, 2017.
- [23] L. Prechelt, “Early Stopping — But When?,” *Neural Networks: Tricks of the Trade*, Springer, pp. 55–69, 1998.
- [24] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [25] S. J. Pan and Q. Yang, “A Survey on Transfer Learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

- [26] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” ICLR, 2015.