

**BỘ NÔNG NGHIỆP VÀ MÔI TRƯỜNG
PHÂN HIỆU TRƯỜNG ĐẠI HỌC THỦY LỢI
BỘ MÔN CÔNG NGHỆ THÔNG TIN**



MÔN HỌC: KHAI PHÁ DỮ LIỆU

**TÊN ĐỀ TÀI: XÂY DỰNG MÔ HÌNH HỌC SÂU CHO BÀI
PHÂN LOẠI BỆNH TRÊN LÁ ĐẬU TỪ HÌNH ẢNH**

Giảng viên:	ThS. Vũ Thị Hạnh
Sinh viên thực hiện:	Nguyễn Văn Tấn Phát Nguyễn Hoàng Lộc Nguyễn Thanh Hoàng
Lớp:	S26-65TTNT

TP. Hồ Chí Minh, ngày ... tháng ... năm 2025

NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN

This image shows a full page of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. There are no margins, text, or other markings on the paper.

TP. Hồ Chí Minh, ngày ... tháng ... năm 2025

Chữ ký của giảng viên

Lời cảm ơn

Trước tiên, với tình cảm sâu sắc và chân thành nhất, nhóm em xin được bày tỏ lòng biết ơn đến tất cả các cá nhân và tổ chức đã tạo điều kiện, hỗ trợ và giúp đỡ nhóm em trong suốt quá trình học tập và thực hiện đề tài này. Trong suốt thời gian học tập tại trường, nhóm em đã nhận được rất nhiều sự quan tâm, chỉ bảo tận tình của quý thầy cô và sự giúp đỡ của bạn bè.

Với lòng biết ơn sâu sắc nhất, nhóm em xin gửi lời cảm ơn chân thành đến quý thầy cô Bộ môn Công nghệ Thông tin – Phân hiệu Trường Đại học Thủy Lợi, những người đã truyền đạt cho nhóm em những kiến thức quý báu trong suốt thời gian học tập. Nhờ có sự giảng dạy, hướng dẫn và động viên của quý thầy cô mà nhóm em có thể hoàn thành tốt đề tài này.

Đặc biệt, nhóm em xin gửi lời cảm ơn sâu sắc đến Cô Vũ Thị Hạnh, người đã trực tiếp hướng dẫn, tận tình giúp đỡ và định hướng cho nhóm trong suốt quá trình thực hiện bài báo cáo.

Do thời gian có hạn và kiến thức còn nhiều hạn chế, bài báo cáo chắc chắn không tránh khỏi những thiếu sót. Nhóm em rất mong nhận được những góp ý quý báu của quý thầy cô để có thể hoàn thiện hơn kiến thức và kinh nghiệm của mình trong lĩnh vực này.

Nhóm em xin chân thành cảm ơn!

Mục Lục

CHƯƠNG 1. GIỚI THIỆU BÀI TOÁN.....	7
1.1. Đặt vấn đề	7
1.2. Mục tiêu	8
1.3. Công cụ sử dụng.....	8
CHƯƠNG 2. MÔ TẢ VÀ TIỀN XỬ LÝ DỮ LIỆU (DATA & PREPROCESSING)	10
2.1. Nguồn dữ liệu.....	10
2.2. Phân tích khám phá dữ liệu (EDA)	10
CHƯƠNG 3. PHƯƠNG PHÁP/MÔ HÌNH ÁP DỤNG (METHODOLOGY)	15
3.1. Mạng nơ-ron tích chập (Convolutional Neural Network - CNN).....	15
3.1.1. Lớp tích chập (Convolutional Layer).....	16
3.1.2. Lớp gộp (Pooling Layer).....	16
3.1.3. Lớp kết nối đầy đủ (Fully Connected Layer).....	16
3.2. Hàm mất mát trong bài toán phân loại	16
3.3. Học chuyển giao (Transfer Learning)	17
3.4. Điều chỉnh tốc độ học (Learning Rate Scheduling)	17
3.5. Regularization và chống overfitting.....	17
3.6. Độ đo đánh giá mô hình	17
3.7. Các mô hình áp dụng.....	18
3.7.1. Mô hình CNN tự xây dựng (CNN from Scratch)	18
3.7.2. Mô hình MobileNetV3	18
3.7.3. Mô hình EfficientNetB3.....	20
3.7.4 Mô hình Resnet50	20
CHƯƠNG 4. KẾT QUẢ VÀ ĐÁNH GIÁ MÔ HÌNH	21
4.1. Tiêu chí và phương pháp đánh giá	21

4.2. Kết quả từ mô hình CNN tự xây của nhóm	21
CHƯƠNG 5. ĐỊNH HƯỚNG PHÁT TRIỂN LÂU DÀI	33
5.1. Mở rộng và đa dạng hóa dữ liệu	33
5.2. Cải tiến mô hình và chiến lược huấn luyện.....	33
5.3. Phát triển giao diện người dùng (UI)	34
5.4. Mở rộng hệ thống trên nền tảng web và di động	34
5.5. Ứng dụng thực tiễn trong nông nghiệp thông minh.....	34

DANH MỤC HÌNH ẢNH

<i>Hình 1 Biểu đồ phân bố số lượng ảnh theo từng lớp.....</i>	<i>11</i>
<i>Hình 2 Phân bố độ sáng (Brightness) và độ tương phản (Contrast) của ảnh ..</i>	<i>12</i>
<i>Hình 3 Phân bố kênh màu RGB</i>	<i>13</i>

DANH MỤC BẢNG

CHƯƠNG 1. GIỚI THIỆU BÀI TOÁN

1.1. Đặt vấn đề

Nông nghiệp đóng vai trò xương sống trong nền kinh tế của nhiều quốc gia đang phát triển, trong đó các loại cây họ Đậu (như đậu xanh, đậu nành, đậu cô ve...) là nguồn cung cấp protein thực vật và dinh dưỡng thiết yếu cho con người. Tuy nhiên, năng suất và chất lượng nông sản đang đối mặt với thách thức lớn từ các loại dịch bệnh gây hại, đặc biệt là các bệnh lý thể hiện trực tiếp qua lá cây.

Trong canh tác cây đậu, các bệnh như Đốm lá do vi khuẩn (Angular Leaf Spot) hay Rỉ sắt (Bean Rust) là những mối đe dọa phổ biến. Nếu không được phát hiện kịp thời, các tổn thương trên lá sẽ làm giảm khả năng quang hợp, khiến cây suy yếu, dẫn đến mất trắng hoặc giảm đáng kể năng suất thu hoạch. Bên cạnh đó, các lá đậu khỏe mạnh cũng cần được nhận diện chính xác để làm cơ sở đối chứng trong quá trình giám sát sự phát triển của cây trồng.

Hiện nay, việc chẩn đoán bệnh trên cây đậu chủ yếu vẫn dựa vào kinh nghiệm quan sát trực quan của người nông dân hoặc các chuyên gia nông nghiệp. Phương pháp này tồn tại nhiều hạn chế:

- Phụ thuộc lớn vào trình độ chuyên môn và dễ gây nhầm lẫn giữa các loại đốm lá có biểu hiện tương đồng.
- Việc mời chuyên gia đến tận vườn để kiểm tra tốn kém chi phí và thời gian, không phù hợp với quy mô sản xuất nhỏ lẻ hoặc ở vùng sâu vùng xa.
- Khi các dấu hiệu bệnh đã biểu hiện rõ rệt trên diện rộng thì dịch thường đã lây lan khó kiểm soát, gây thiệt hại kinh tế nặng nề.

Sự tiến bộ vượt bậc của Trí tuệ nhân tạo, đặc biệt là Học sâu (Deep Learning) và Thị giác máy tính (Computer Vision), đã mở ra giải pháp tự động hóa quá trình này. Các mô hình mạng nơ-ron tích chập (CNN) có khả năng nhận diện các đặc trưng phức tạp về màu sắc, hình dạng và kết cấu của các vết thương trên lá với độ chính xác cao và tốc độ xử lý tức thời. Việc ứng dụng các mô hình này vào nông nghiệp giúp người dân có thể chẩn đoán bệnh cho cây trồng chỉ qua một bức ảnh chụp từ thiết bị di động, từ đó đưa ra biện pháp xử lý kịp thời.

Xuất phát từ thực tiễn đó, đề tài “Xây Dựng Mô Hình Học Sâu Cho Bài Toán Phân Loại Bệnh Trên Lá Đậu Từ Hình Ảnh” được thực hiện nhằm xây dựng một hệ thống có khả năng phân loại chính xác các trạng thái: bệnh Đốm lá, bệnh Rỉ sắt và lá khỏe mạnh. Kết quả nghiên cứu kỳ vọng sẽ trở thành công cụ hỗ trợ đắc lực cho người nông dân trong việc bảo vệ mùa màng và nâng cao hiệu quả canh tác.

1.2. Mục tiêu

Mục tiêu tổng quát của đề tài là xây dựng mô hình học sâu có khả năng phân loại tự động các tình trạng sức khỏe của cây đậu thông qua hình ảnh lá (bao gồm các bệnh lý như đốm lá, rỉ sắt và trạng thái khỏe mạnh) với độ chính xác cao. Về mục tiêu cụ thể, thứ nhất, đề tài thực hiện phân tích và tiền xử lý bộ dữ liệu hình ảnh bệnh lá, bao gồm việc chuẩn hóa kích thước ảnh, tăng cường dữ liệu và xử lý vấn đề mất cân bằng giữa các lớp. Thứ hai, đề tài xây dựng mô hình mạng nơ-ron tích chập từ đầu (CNN from scratch) để đánh giá khả năng học đặc trưng trực tiếp từ dữ liệu. Thứ ba, đề tài áp dụng kỹ thuật Transfer Learning nhằm tận dụng tri thức đã học từ bộ dữ liệu lớn và rút ngắn thời gian huấn luyện. Thứ tư, đề tài tiến hành đánh giá và so sánh hiệu suất hai mô hình thông qua các độ đo Accuracy, Precision, Recall, F1-Score và Confusion Matrix.

1.3. Công cụ sử dụng

Đề tài được thực hiện trên nền tảng ngôn ngữ lập trình Python phiên bản 3.x, một ngôn ngữ phổ biến trong lĩnh vực khoa học dữ liệu và học máy nhờ cú pháp đơn giản và hệ sinh thái thư viện phong phú.

Về framework học sâu, đề tài sử dụng TensorFlow kết hợp với Keras API để xây dựng, huấn luyện và đánh giá các mô hình mạng nơ-ron tích chập. TensorFlow là một thư viện mã nguồn mở được phát triển bởi Google, cung cấp các công cụ mạnh mẽ để xây dựng mô hình học sâu với khả năng tận dụng GPU để tăng tốc quá trình huấn luyện. Keras API cung cấp giao diện lập trình cấp cao, giúp việc thiết kế kiến trúc mô hình trở nên trực quan và dễ dàng hơn.

Về xử lý và phân tích dữ liệu, đề tài sử dụng thư viện NumPy để thực hiện các phép tính ma trận và xử lý mảng đa chiều, Pandas để quản lý và thao tác dữ liệu dạng bảng, OpenCV và Pillow để đọc, xử lý và biến đổi hình ảnh. Thư viện Scikit-learn được sử dụng để chia tập dữ liệu, tính toán các độ đo đánh giá và tạo ma trận nhầm lẫn.

Về trực quan hóa, thư viện Matplotlib và Seaborn được sử dụng để vẽ các biểu đồ phân tích dữ liệu, đường cong huấn luyện và ma trận nhầm lẫn, giúp quan sát trực quan quá trình huấn luyện và kết quả đánh giá mô hình.

Môi trường phát triển được sử dụng là Jupyter Notebook hoặc Kanggle Notebook, cho phép viết mã, chạy thử nghiệm và ghi chú kết quả một cách linh hoạt. Kanggle Notebook cung cấp GPU miễn phí, hỗ trợ đáng kể trong việc rút ngắn thời gian huấn luyện các mô hình học sâu có kiến trúc phức tạp.

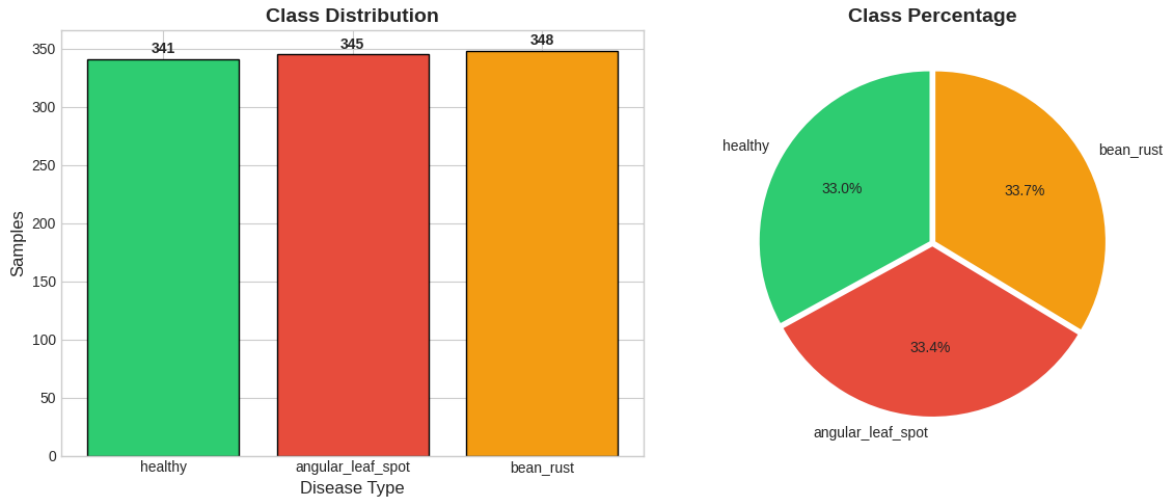
CHƯƠNG 2. MÔ TẢ VÀ TIỀN XỬ LÝ DỮ LIỆU (DATA & PREPROCESSING)

2.1. Nguồn dữ liệu

Bộ dữ liệu được sử dụng trong nghiên cứu này là tập hợp các hình ảnh về tổn thương trên lá đậu, được thu thập và công bố công khai trên nền tảng khoa học dữ liệu Kaggle. Mục tiêu chính của bộ dữ liệu là cung cấp nguồn tài nguyên chuẩn cho các nghiên cứu ứng dụng thị giác máy tính trong lĩnh vực nông nghiệp thông minh, cụ thể là nhận diện bệnh lý cây trồng qua hình ảnh lá. Tổng số lượng hình ảnh trong bộ dữ liệu bao gồm 1034 ảnh trong tập huấn luyện và 133 ảnh trong tập hiệu chỉnh, được phân loại thành ba nhóm chính tương ứng với các tình trạng sức khỏe khác nhau: bệnh Đốm lá (Angular Leaf Spot), bệnh Rỉ sắt (Bean Rust) và lá khỏe mạnh (Healthy). Mỗi lớp dữ liệu đại diện cho một đặc điểm hình thái riêng biệt của vết bệnh, giúp mô hình học cách phân biệt các dấu hiệu bất thường trên bề mặt lá.

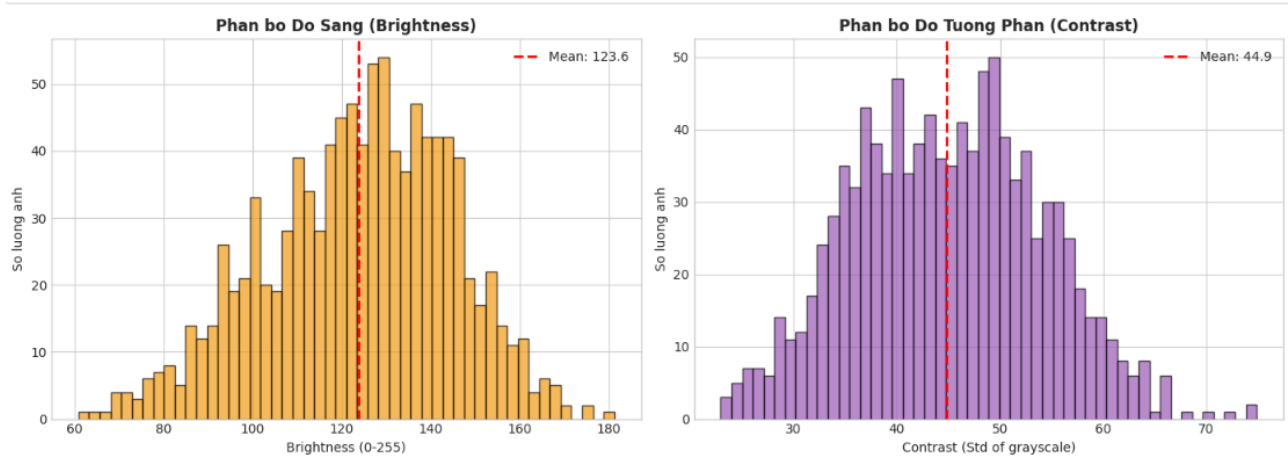
Dữ liệu được cung cấp dưới định dạng JPEG với kích thước chuẩn 500x500 pixel, được chụp trong điều kiện môi trường thực tế tại các cánh đồng đậu, do đó hình ảnh phản ánh sự đa dạng về cường độ ánh sáng, góc chụp và độ phức tạp của hậu cảnh. Qua phân tích sơ bộ, bộ dữ liệu cho thấy sự phân bố tương đối đồng đều giữa các lớp, tuy nhiên vẫn cần được kiểm soát kỹ lưỡng để tránh hiện tượng quá khớp (overfitting) do quy mô tập mẫu không quá lớn. Chất lượng hình ảnh có sự biến thiên từ các ảnh sắc nét đến các ảnh có độ nhiễu nhất định do yếu tố môi trường ngoài trời, đặt ra yêu cầu mô hình học sâu phải có khả năng tổng quát hóa cao để xử lý hiệu quả các trường hợp dữ liệu trong thực tế canh tác.

2.2. Phân tích khám phá dữ liệu (EDA)



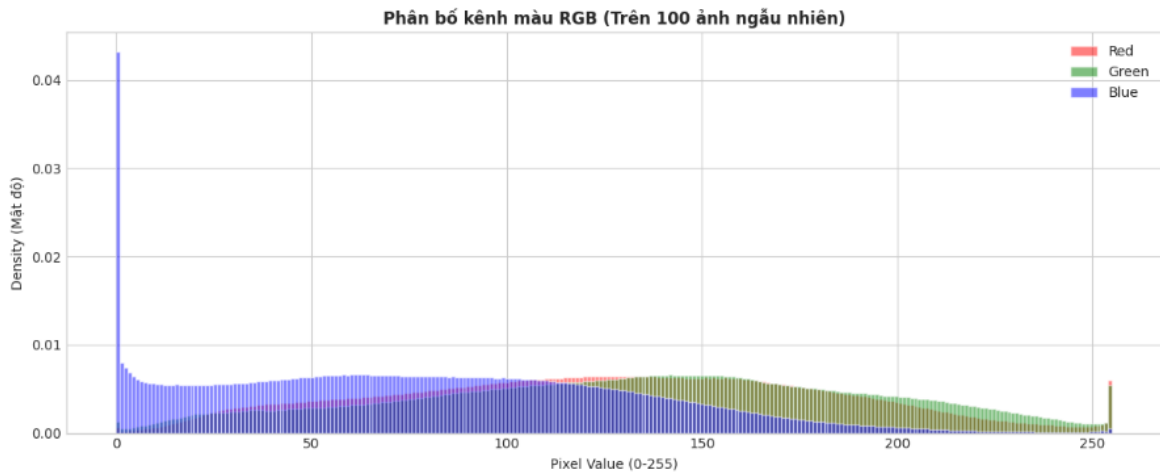
Hình 1 Biểu đồ phân bố số lượng ảnh theo từng lớp

Hình 1 cho thấy sự phân bố số lượng ảnh giữa các lớp bệnh trong bộ dữ liệu. Cụ thể, trong tập train, mỗi lớp có khoảng 345–350 ảnh, sự chênh lệch là không đáng kể, giúp mô hình tránh được hiện tượng thiên lệch về một lớp cụ thể trong quá trình học. Tương tự, tập validation cũng duy trì sự cân bằng với khoảng 44–45 ảnh cho mỗi lớp, đảm bảo quá trình đánh giá hiệu năng mô hình được thực hiện một cách khách quan và đáng tin cậy. Nhìn chung, việc phân chia dữ liệu hợp lý này tạo điều kiện thuận lợi cho mô hình học được đặc trưng đại diện của từng lớp và phản ánh chính xác hơn khả năng tổng quát hóa. Tuy nhiên, quy mô dữ liệu ở mức trung bình, do đó cần kết hợp các kỹ thuật như tăng cường dữ liệu và transfer learning nhằm nâng cao hiệu quả huấn luyện và hạn chế hiện tượng overfitting.



Hình 2 Phân bố độ sáng (Brightness) và độ tương phản (Contrast) của ảnh

Biểu đồ phân bố độ sáng và độ tương phản của ảnh sau xử lý cho thấy chất lượng dữ liệu nhìn chung ở mức ổn định và phù hợp cho bài toán huấn luyện mô hình học sâu. Phân bố độ sáng trải dài từ 60.8 đến 181.2 với giá trị trung bình 123.6, tập trung chủ yếu quanh vùng trung bình, phản ánh điều kiện chiếu sáng tương đối đồng đều trong tập dữ liệu; số lượng ảnh quá tối (brightness < 80) chỉ chiếm 2.6% và ảnh quá sáng (brightness > 180) gần như không đáng kể với 0.1%, cho thấy hiện tượng lệch sáng không nghiêm trọng. Đối với độ tương phản, giá trị dao động từ 22.9 đến 74.9 với trung bình 44.9, trong đó đa số ảnh có mức tương phản trung bình đến khá, đủ để bảo toàn các đặc trưng thị giác quan trọng; tỷ lệ ảnh có tương phản thấp (contrast < 30) chiếm 4.7%, ở mức chấp nhận được và chưa gây ảnh hưởng lớn đến khả năng học của mô hình. Tổng thể, các biểu đồ cho thấy tập dữ liệu có phân bố ánh sáng và tương phản hợp lý, ít nhiễu cực đoan, tạo nền tảng thuận lợi cho quá trình trích xuất đặc trưng và huấn luyện mô hình, đồng thời có thể được cải thiện thêm bằng các kỹ thuật tăng cường dữ liệu liên quan đến điều chỉnh ánh sáng và tương phản nhằm nâng cao khả năng tổng quát hóa.



Hình 3 Phân bố kênh màu RGB

Biểu đồ phân bố giá trị kênh màu RGB trên 100 ảnh ngẫu nhiên cho thấy sự khác biệt rõ ràng giữa ba kênh màu, phản ánh đặc trưng màu sắc của tập dữ liệu ảnh lá cây. Kênh xanh lá (Green) có phân bố tập trung mạnh ở vùng giá trị trung bình đến cao, phù hợp với đặc điểm sinh học của lá cây khi màu xanh chiếm ưu thế trong ảnh, đồng thời cho thấy thông tin màu sắc quan trọng phục vụ việc nhận diện tình trạng bệnh. Kênh đỏ (Red) có phân bố trải rộng hơn nhưng mật độ thấp hơn kênh Green, cho thấy sự hiện diện của các vùng đổi màu như vàng hoặc nâu, vốn thường liên quan đến các triệu chứng bệnh trên lá. Trong khi đó, kênh xanh dương (Blue) có mật độ cao ở vùng giá trị thấp và giảm dần về phía giá trị lớn, phản ánh việc thành phần màu xanh dương ít chiếm ưu thế trong bối cảnh ảnh tự nhiên ngoài trời. Tổng thể, sự phân bố RGB không đồng đều nhưng hợp lý về mặt ngữ cảnh, cho thấy tập dữ liệu có đặc trưng màu sắc rõ ràng và đủ thông tin để mô hình học sâu khai thác các đặc trưng màu trong quá trình phân loại, đồng thời cũng gợi ý rằng việc chuẩn hóa màu sắc và áp dụng các kỹ thuật tăng cường dữ liệu liên quan đến màu sắc là cần thiết để nâng cao khả năng tổng quát hóa của mô hình.

2.3. Tiền xử lý (Preprocessing)

Để tối ưu hóa quá trình huấn luyện và nâng cao khả năng tổng quát hóa của mô hình đối với dữ liệu lá đậu, nhóm thực hiện các bước tiền xử lý và tăng cường dữ liệu cụ thể như sau:

2.3.1. Chuẩn hóa giá trị pixel

Toàn bộ hình ảnh đầu vào được thực hiện chuẩn hóa lại giá trị pixel (Rescaling). Các giá trị pixel gốc nằm trong khoảng từ 0 đến 255 được chuyển đổi về đoạn $[0, 1]$ thông qua phép chia cho 255. Bước này giúp ổn định quá trình tính toán đạo hàm, giảm thiểu hiện tượng bùng nổ gradient và giúp thuật toán tối ưu hóa hội tụ nhanh hơn.

2.3.2. Tăng cường dữ liệu (Data Augmentation)

Do đặc thù bộ dữ liệu lá đậu có số lượng mẫu không quá lớn, kỹ thuật tăng cường dữ liệu được áp dụng nhằm tạo ra các biến thể mới từ dữ liệu gốc, giúp mô hình học được các đặc trưng bất biến và chống quá khớp (Overfitting). Các phép biến đổi bao gồm:

- Biến đổi hình học: Xoay ảnh ngẫu nhiên trong khoảng 15 độ, dịch chuyển ảnh theo chiều rộng và chiều cao tối đa 10%. Ngoài ra, ảnh được áp dụng phép lật ngẫu nhiên theo cả chiều ngang và chiều dọc để mô phỏng các góc chụp khác nhau trong thực tế.
- Biến đổi tỷ lệ: Áp dụng phép phóng to hoặc thu nhỏ (Zoom) trong khoảng 10% để mô phỏng sự khác biệt về khoảng cách từ thiết bị chụp đến bề mặt lá.
- Điều chỉnh độ sáng: Nhóm thiết lập dải độ sáng từ 0.85 đến 1.15. Mức điều chỉnh này được tính toán nhẹ nhàng nhằm giữ trọn vẹn các đặc trưng màu sắc quan trọng để phân biệt bệnh lý như: sắc xanh tươi của lá khỏe mạnh, màu nâu có viền vàng của bệnh Đốm lá (Angular Leaf Spot), và sắc nâu đỏ gỉ đặc trưng của bệnh Rỉ sắt (Bean Rust).

2.3.3. Cơ chế xử lý vùng trống và Data Generator

Trong quá trình biến đổi hình học (xoay hoặc dịch chuyển), các vùng trống phát sinh tại biên ảnh được xử lý bằng phương pháp reflect (phản chiếu). Kỹ thuật này giúp duy trì cấu trúc vân lá và texture tự nhiên, tránh tạo ra các đường kẻ đen nhân tạo gây nhiễu cho mô hình.

Cuối cùng, dữ liệu được đưa vào mô hình thông qua Data Generator để xử lý theo từng đợt (Batch processing). Trong khi tập huấn luyện (Training set) được áp dụng đầy đủ các kỹ thuật tăng cường, tập kiểm định (Validation set) chỉ thực hiện duy nhất bước chuẩn hóa pixel (rescale) để đảm bảo tính khách quan và chính xác khi đánh giá hiệu năng mô hình trên dữ liệu thực.

CHƯƠNG 3. PHƯƠNG PHÁP/MÔ HÌNH ÁP DỤNG (METHODOLOGY)

Phân loại ảnh (*Image Classification*) là một trong những bài toán cốt lõi của lĩnh vực **thị giác máy tính (Computer Vision)**, với mục tiêu gán một nhãn duy nhất cho mỗi ảnh đầu vào dựa trên nội dung hình ảnh. Trong bài toán phân loại đa lớp, mỗi ảnh chỉ thuộc về **một lớp duy nhất** trong tập các lớp đã được xác định trước.

Trong đề tài này, bài toán được mô hình hóa như một **bài toán phân loại ảnh đa lớp**, trong đó **đầu vào (Input)** là hình ảnh lá cây và **đầu ra (Output)** là một trong ba nhãn tương ứng với tình trạng của lá, bao gồm: **healthy (lá khỏe)**, **angular_leaf_spot (bệnh đốm góc)** và **bean_rust (bệnh gỉ sắt)**.

Đây là một bài toán có độ khó tương đối cao do **sự tương đồng về hình thái, màu sắc và kết cấu tổn thương giữa các loại bệnh**, đồng thời chịu ảnh hưởng bởi **điều kiện chụp ảnh khác nhau trong môi trường thực tế** như ánh sáng, góc chụp và nhiễu nền. Điều này đòi hỏi mô hình học sâu phải trích xuất được các đặc trưng mang tính phân biệt cao để đạt được kết quả phân loại chính xác.

3.1. Mạng nơ-ron tích chập (Convolutional Neural Network - CNN)

Mạng nơ-ron tích chập là một kiến trúc học sâu được thiết kế đặc biệt để xử lý dữ liệu có cấu trúc dạng lưới như hình ảnh. Khác với mạng nơ-ron truyền thống kết nối

đầy đủ giữa các lớp, CNN sử dụng các bộ lọc tích chập để trích xuất đặc trưng cục bộ từ hình ảnh, giúp giảm đáng kể số lượng tham số cần học và tăng khả năng nhận dạng các mẫu hình không phụ thuộc vào vị trí trong ảnh.

Kiến trúc CNN cơ bản bao gồm ba loại lớp chính. Lớp tích chập (Convolutional Layer) thực hiện phép tích chập giữa ảnh đầu vào với các bộ lọc để tạo ra các bản đồ đặc trưng. Lớp gộp (Pooling Layer) thực hiện việc giảm kích thước không gian của bản đồ đặc trưng. Lớp kết nối đầy đủ (Fully Connected Layer) nằm ở cuối mạng để thực hiện phân loại cuối cùng.

3.1.1. Lớp tích chập (Convolutional Layer)

Lớp tích chập sử dụng các bộ lọc (kernel) trượt trên ảnh đầu vào để trích xuất các đặc trưng cục bộ như biên, góc, kết cấu và các mẫu hình phức tạp hơn ở các tầng sâu. Nhờ cơ chế chia sẻ trọng số (weight sharing), CNN có khả năng học các đặc trưng không phụ thuộc vào vị trí, rất phù hợp với dữ liệu hình ảnh.

3.1.2. Lớp gộp (Pooling Layer)

Lớp gộp, phổ biến nhất là MaxPooling, được sử dụng để giảm kích thước không gian của feature map, giúp:

- Giảm số lượng tham số
- Giảm chi phí tính toán
- Tăng khả năng khái quát hóa của mô hình

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

Sau khi trích xuất đặc trưng, các feature map được chuyển thành vector và đưa vào các lớp kết nối đầy đủ để thực hiện phân loại. Lớp cuối cùng thường sử dụng hàm kích hoạt Softmax để tạo ra phân phối xác suất trên các lớp.

3.2. Hàm mất mát trong bài toán phân loại

Trong bài toán phân loại đa lớp, hàm mất mát **Categorical Cross-Entropy** được sử dụng phổ biến để đo lường mức độ sai lệch giữa phân phối xác suất dự đoán của mô hình và nhãn thực tế.

Hàm mất mát này khuyến khích mô hình:

- Tăng xác suất cho lớp đúng
- Giảm xác suất cho các lớp sai

Đối với các mô hình trong đề tài, Categorical Cross-Entropy được kết hợp với **Label Smoothing**, một kỹ thuật giúp giảm sự tự tin thái quá của mô hình, đặc biệt hiệu quả trong trường hợp dữ liệu có nhiễu hoặc nhãn không hoàn toàn chính xác.

3.3. Học chuyển giao (Transfer Learning)

Học chuyển giao là kỹ thuật sử dụng kiến thức đã học được từ một bài toán để áp dụng vào bài toán khác có liên quan. Các mô hình đã được huấn luyện trước trên bộ dữ liệu ImageNet với hàng triệu hình ảnh có thể được tái sử dụng cho các bài toán phân loại hình ảnh cụ thể. Ưu điểm của Transfer Learning bao gồm giảm thời gian huấn luyện, cải thiện hiệu suất trên bộ dữ liệu nhỏ và giảm nguy cơ overfitting.

3.4. Điều chỉnh tốc độ học (Learning Rate Scheduling)

Để tránh việc mô hình hội tụ kém hoặc bị kẹt tại cực trị cục bộ, các kỹ thuật điều chỉnh learning rate như **Warmup Scheduler** và **ReduceLROnPlateau** được áp dụng. Các kỹ thuật này giúp mô hình học ổn định ở giai đoạn đầu và tinh chỉnh tốt hơn ở giai đoạn sau.

3.5. Regularization và chống overfitting

Để hạn chế hiện tượng quá khớp, đề tài sử dụng nhiều kỹ thuật regularization:

- **Dropout**: ngẫu nhiên loại bỏ một phần nơ-ron trong quá trình huấn luyện
- **L2 Regularization**: phạt các trọng số lớn
- **Data Augmentation**: tăng tính đa dạng của dữ liệu huấn luyện
- **Early Stopping**: dừng huấn luyện khi mô hình không còn cải thiện

3.6. Độ đo đánh giá mô hình

Để đánh giá hiệu quả của mô hình phân loại, đề tài sử dụng các độ đo phổ biến:

- **Accuracy**: tỷ lệ dự đoán đúng trên tổng số mẫu
- **Precision**: mức độ chính xác của các dự đoán dương

- **Recall:** khả năng phát hiện đúng các mẫu dương
- **F1-score:** trung bình điều hòa giữa Precision và Recall

Trong bối cảnh dữ liệu mất cân bằng, **F1-score** được xem là độ đo quan trọng hơn Accuracy vì phản ánh tốt hơn hiệu năng của mô hình trên các lớp thiểu số.

3.7. Các mô hình áp dụng

3.7.1. Mô hình CNN tự xây dựng (CNN from Scratch)

Mô hình được xây dựng từ đầu với tên gọi LiteEfficientNet, lấy cảm hứng từ kiến trúc EfficientNet nhưng được tinh gọn để phù hợp với việc huấn luyện trực tiếp trên tập dữ liệu lá đậu mà không cần pre-trained weights. Kiến trúc mạng bao gồm một khối khởi tạo và 4 khối trích xuất đặc trưng chính, sử dụng chuỗi các lớp tích chập (Conv2d) kết hợp với Normalization (BatchNorm) và hàm kích hoạt hiện đại SiLU (Swish). Ảnh đầu vào kích thước 350x350 được xử lý qua các tầng để giảm chiều không gian và tăng chiều sâu đặc trưng lên 512 kênh tại lớp cuối cùng. Để ngăn chặn overfitting – một vấn đề phổ biến khi huấn luyện từ đầu – mô hình tích hợp lớp Dropout (tỷ lệ 0.3) và kỹ thuật Label Smoothing (0.1) vào hàm mất mát CrossEntropyLoss. Quá trình huấn luyện kéo dài 80 epochs sử dụng optimizer AdamW kết hợp với CosineAnnealingLR scheduler để tối ưu hóa learning rate. Kết quả thực nghiệm cho thấy mô hình hội tụ rất tốt, đạt độ chính xác trên tập validation lên tới ~97% (tại epoch 69), chứng minh tính hiệu quả của kiến trúc tự thiết kế này.

3.7.2. Mô hình MobileNetV3

Bài toán phân loại bệnh lá đậu sử dụng kiến trúc MobileNetV3Large - một mạng neural network nhẹ nhưng hiệu quả, được pretrain trên tập ImageNet. Mô hình được thiết kế với base model MobileNetV3Large kết hợp cùng một Custom Classification Head bao gồm: GlobalAveragePooling2D để giảm chiều dữ liệu và tránh overfitting, tiếp theo là một lớp Dense với 256 units áp dụng L2 Regularization ($1e-4$), lớp BatchNormalization để chuẩn hóa phân phối features, hàm kích hoạt Swish cho hiệu quả cao hơn ReLU truyền thống, Dropout với tỷ lệ 0.3 để ngăn ngừa overfitting, và cuối cùng là lớp Softmax output cho 3 classes (healthy, angular_leaf_spot,

bean_rust). Tổng cộng mô hình có khoảng 3.2 triệu tham số, trong đó chỉ có 247K tham số trainable ở phần head và 3M tham số frozen ở base model.

Về kỹ thuật tăng cường dữ liệu, bài toán áp dụng chiến lược Disease-Aware Augmentation - tức là augmentation nhẹ để bảo toàn các đặc trưng quan trọng của bệnh lá như đốm, màu sắc và texture. Cụ thể, các tham số augmentation được thiết lập ở mức thấp với rotation tối đa 10 độ, dịch chuyển theo chiều rộng và chiều cao chỉ 5%, zoom range 5%, horizontal và vertical flip, và brightness thay đổi trong khoảng $\pm 10\%$. Với tập validation, chỉ thực hiện rescale mà không áp dụng augmentation để đảm bảo đánh giá chính xác.

Quá trình huấn luyện được chia thành hai giai đoạn (Two-Phase Training). Trong Phase 1, mục tiêu là train phần Classification Head trong khi giữ base model frozen hoàn toàn. Giai đoạn này sử dụng optimizer AdamW với weight decay $1e-2$, learning rate khởi đầu $5e-4$ kết hợp Warmup Schedule tăng dần từ 0.0001 đến 0.0005 trong 5 epochs đầu để tránh gradient explosion. Hàm loss sử dụng CategoricalCrossentropy với Label Smoothing 0.1 để tạo soft labels giúp model không overconfident. Các callbacks bao gồm EarlyStopping với patience 8 epochs theo dõi val_loss, ModelCheckpoint lưu model tốt nhất, và ReduceLROnPlateau giảm learning rate 50% khi val_loss không cải thiện sau 3 epochs.

Trong Phase 2, mô hình thực hiện Fine-Tuning bằng cách unfreeze các layer từ layer 50 trở đi của base model, giữ nguyên các low-level features đã học và chỉ train lại high-level features. Điểm quan trọng là tất cả các lớp BatchNormalization trong base model được giữ frozen để tránh unstable statistics khi làm việc với tập dữ liệu nhỏ. Learning rate được giảm xuống còn $1e-5$ (thấp hơn 50 lần so với Phase 1) để bảo vệ các pretrained weights. EarlyStopping được tăng patience lên 15 epochs và theo dõi val_accuracy thay vì val_loss.

Tổng hợp các kỹ thuật chống overfitting được sử dụng bao gồm: L2 Regularization áp dụng trên Dense layer với hệ số $1e-4$, Dropout với tỷ lệ 30%, BatchNormalization chuẩn hóa phân phối, GlobalAveragePooling giảm số tham số, Label Smoothing 0.1 tạo soft labels, Weight Decay $1e-2$ tích hợp trong AdamW optimizer, EarlyStopping dừng training khi không cải thiện, Learning Rate Warmup tránh gradient explosion ở

giai đoạn đầu, freeze BatchNorm layers khi fine-tune để giữ ổn định statistics, và Conservative Augmentation bảo toàn đặc trưng bệnh. Kết quả cuối cùng đạt Best Validation Accuracy 98.50% tương đương 131/133 samples được phân loại đúng.

3.7.3. Mô hình EfficientNetB3

sử dụng mô hình EfficientNet-B3 áp dụng kỹ thuật Transfer Learning từ bộ trọng số pre-trained trên ImageNet. Kiến trúc mạng được tùy chỉnh bằng cách thay thế lớp classifier cuối cùng bằng một lớp Linear mới có số đầu ra tương ứng với 3 lớp bệnh (healthy, angular_leaf_spot, bean_rust). Quá trình huấn luyện cho phép fine-tuning toàn bộ các lớp đặc trưng (features) ngay từ đầu để mô hình thích nghi tốt nhất với tập dữ liệu lá đậu.

Chiến lược huấn luyện được thiết lập trong 30 epochs với Batch Size là 32. Các kỹ thuật tối ưu hóa bao gồm:

Optimizer: Sử dụng AdamW với learning rate $3e-4$ và weight decay $1e-2$ để giúp hội tụ nhanh và tránh overfitting.

Loss Function: Sử dụng CrossEntropyLoss tiêu chuẩn cho bài toán phân loại đa lớp.

Learning Rate Scheduler: Áp dụng CosineAnnealingLR ($T_{max}=30$) để điều chỉnh learning rate theo hình sin giảm dần, giúp mô hình thoát khỏi các điểm cực tiểu cục bộ và hội tụ tốt hơn ở giai đoạn cuối.

Early Stopping: Cơ chế dừng sớm với patience=5 được sử dụng để ngăn chặn overfitting khi validation loss không còn cải thiện.

3.7.4 Mô hình Resnet50

Mô hình ResNet50 được triển khai sử dụng thư viện TensorFlow/Keras cho bài toán phân loại bệnh lá đậu với 3 lớp (angular_leaf_spot, bean_rust, healthy). Đây là giải pháp Transfer Learning tận dụng kiến trúc ResNet50 đã được huấn luyện trước (thường là trên ImageNet). Mô hình bao gồm phần base là ResNet50 để trích xuất đặc trưng, kết hợp với một lớp phân loại tùy chỉnh gồm GlobalAveragePooling2D và các lớp Dense để đưa ra dự đoán cuối cùng. Dữ liệu đầu vào được resize về kích thước 224×224 , xử lý qua hàm preprocess_input chuẩn của ResNet và huấn luyện với

batch size 32 trong tối đa 150 epochs. Chiến lược huấn luyện sử dụng hàm mất mát CategoricalCrossentropy và optimizer Adam, đồng thời áp dụng các kỹ thuật chống overfitting và tối ưu hóa như EarlyStopping (dừng sớm khi không cải thiện) và ReduceLROnPlateau (giảm learning rate khi loss đi ngang). Tập dữ liệu được phân bố cân bằng với khoảng 345 ảnh mỗi lớp cho tập train, đảm bảo mô hình học đều các đặc trưng của từng loại bệnh.

CHƯƠNG 4. KẾT QUẢ VÀ ĐÁNH GIÁ MÔ HÌNH

4.1. Tiêu chí và phương pháp đánh giá

Các mô hình CNN cơ bản, ResNet và EfficientNet được đánh giá trên tập test gồm 132 ảnh với 3 lớp, bao gồm lá khỏe (healthy), bệnh đốm góc (angular leaf spot) và bệnh gỉ sắt (bean rust).

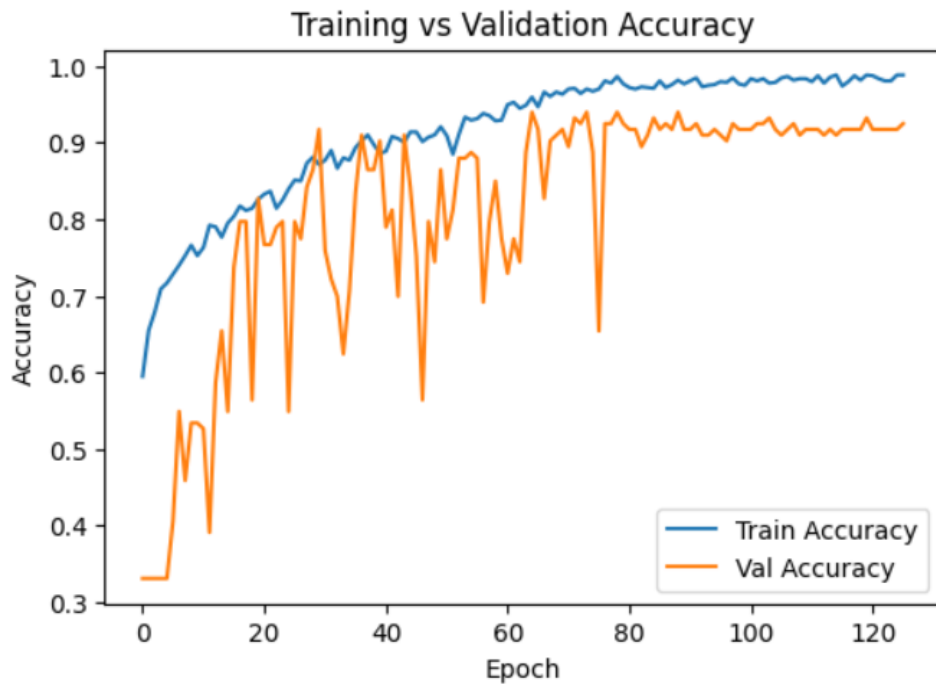
Hiệu năng của các mô hình được đánh giá thông qua các chỉ số Accuracy, Precision, Recall, F1-score và ma trận nhầm lẫn (Confusion Matrix).

Ngoài ra, mô hình còn được trực quan hóa bằng Grad-CAM nhằm phân tích và đánh giá các vùng ảnh mà mạng nơ-ron tập trung khi đưa ra quyết định phân loại bệnh lá cây.

4.2. Kết quả từ mô hình CNN tự xây của nhóm

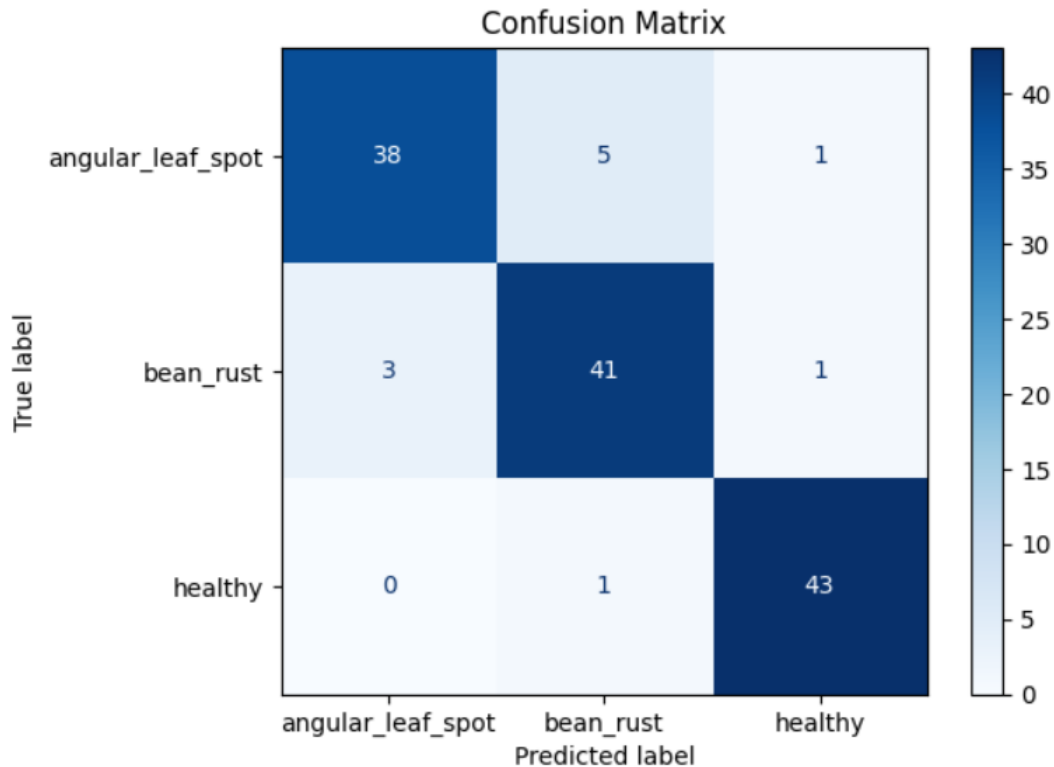
4.2.1. Kết quả huấn luyện mô hình CNN tự xây

Mô hình CNN cơ bản được huấn luyện trong 10 epoch với dữ liệu đầu vào gồm ba lớp: healthy, angular leaf spot và bean rust. Các chỉ số đánh giá bao gồm Accuracy, Precision, Recall, F1-score và Confusion Matrix.



- Độ chính xác trên tập huấn luyện (training accuracy) tăng nhanh trong các epoch đầu và tiếp tục cải thiện ổn định theo thời gian, đạt mức xấp xỉ 98–99% ở các epoch cuối.
- Trong khi đó, độ chính xác trên tập kiểm định (validation accuracy) tăng dần nhưng dao động mạnh trong giai đoạn đầu và giữa quá trình huấn luyện, sau đó ổn định quanh mức 90–93% ở các epoch cuối.
- Sự dao động của validation accuracy cho thấy mô hình chịu ảnh hưởng của kích thước tập dữ liệu kiểm định còn nhỏ và sự khác biệt giữa các mẫu, tuy nhiên xu hướng chung vẫn tăng và hội tụ, cho thấy mô hình có khả năng tổng quát hóa tương đối tốt.

4.2.2. Ma trận nhầm lẫn của CNN tự xây



- Mô hình ResNet phân loại đúng 122/132 mẫu, đạt độ chính xác cao trên tập kiểm tra.
- Đối với lớp angular leaf spot, mô hình nhận diện đúng 38/44 mẫu, trong đó có 5 mẫu bị nhầm sang bean rust và 1 mẫu bị nhầm sang healthy.
- Với lớp bean rust, mô hình dự đoán đúng 41/45 mẫu, chỉ có 3 mẫu bị nhầm sang angular leaf spot và 1 mẫu bị nhầm sang healthy.
- Lớp healthy được nhận diện chính xác nhất với 43/44 mẫu đúng, chỉ có 1 mẫu bị nhầm sang bean rust.
- Các lỗi phân loại chủ yếu xảy ra giữa hai lớp bệnh có đặc trưng hình thái tương đồng, trong khi tỷ lệ nhầm lẫn với lá khỏe là rất thấp, cho thấy mô hình ResNet có khả năng tổng quát hóa tốt trên tập dữ liệu kiểm tra.

4.3 Kết quả từ mô hình ResNet của nhóm

4.3.1 Cấu trúc mô hình

Mô hình được xây dựng dựa trên kiến trúc ResNet50 (Residual Network với 50 tầng), sử dụng trọng số ImageNet để tận dụng đặc trưng học sẵn,

4.3.2 Kết quả huấn luyện của ResNet50

Mô hình ResNet50 được huấn luyện trong 50 epoch trên bộ dữ liệu gồm 3 lớp: *angular_leaf_spot*, *bean_rust* và *healthy*.

```

... Epoch 1/50
33/33 ————— 74s 2s/step - accuracy: 0.5348 - loss: 1.0526 - val_accuracy: 0.8045 - val_loss: 0.5967
Epoch 2/50
33/33 ————— 65s 2s/step - accuracy: 0.7727 - loss: 0.6290 - val_accuracy: 0.8797 - val_loss: 0.4803
Epoch 3/50
33/33 ————— 62s 2s/step - accuracy: 0.8095 - loss: 0.5633 - val_accuracy: 0.8797 - val_loss: 0.4119
Epoch 4/50
33/33 ————— 61s 2s/step - accuracy: 0.8511 - loss: 0.5011 - val_accuracy: 0.8797 - val_loss: 0.3854
Epoch 5/50
33/33 ————— 57s 2s/step - accuracy: 0.8656 - loss: 0.4668 - val_accuracy: 0.9173 - val_loss: 0.3559
Epoch 6/50
33/33 ————— 57s 2s/step - accuracy: 0.8888 - loss: 0.4323 - val_accuracy: 0.9248 - val_loss: 0.3401
Epoch 7/50
33/33 ————— 57s 2s/step - accuracy: 0.8714 - loss: 0.4322 - val_accuracy: 0.9098 - val_loss: 0.3335
Epoch 8/50
33/33 ————— 59s 2s/step - accuracy: 0.8723 - loss: 0.4414 - val_accuracy: 0.9173 - val_loss: 0.3208
Epoch 9/50
33/33 ————— 58s 2s/step - accuracy: 0.9130 - loss: 0.3843 - val_accuracy: 0.9323 - val_loss: 0.3176
Epoch 10/50
33/33 ————— 57s 2s/step - accuracy: 0.9207 - loss: 0.3764 - val_accuracy: 0.9474 - val_loss: 0.3093
Epoch 11/50
33/33 ————— 65s 2s/step - accuracy: 0.9110 - loss: 0.3673 - val_accuracy: 0.9398 - val_loss: 0.2992
Epoch 12/50
33/33 ————— 58s 2s/step - accuracy: 0.9284 - loss: 0.3504 - val_accuracy: 0.9248 - val_loss: 0.2959
Epoch 13/50
...
Epoch 49/50
33/33 ————— 52s 2s/step - accuracy: 0.9826 - loss: 0.2481 - val_accuracy: 0.9624 - val_loss: 0.2692
Epoch 50/50
33/33 ————— 51s 2s/step - accuracy: 0.9787 - loss: 0.2491 - val_accuracy: 0.9624 - val_loss: 0.2639
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...

```

Quá trình huấn luyện cho thấy mô hình hội tụ ổn định, không xuất hiện dấu hiệu quá khớp rõ rệt.

- Độ chính xác trên tập huấn luyện (training accuracy) tăng đều theo các epoch và đạt mức trên 97% ở giai đoạn cuối.
- Độ chính xác trên tập kiểm định (validation accuracy) duy trì ổn định quanh mức 96%, với khoảng chênh lệch nhỏ so với tập huấn luyện.
- Giá trị loss trên cả hai tập đều giảm dần và hội tụ, cho thấy mô hình học hiệu quả và có khả năng tổng quát hóa tốt.

4.3.3 Kết quả huấn luyện của mô hình ResNet50

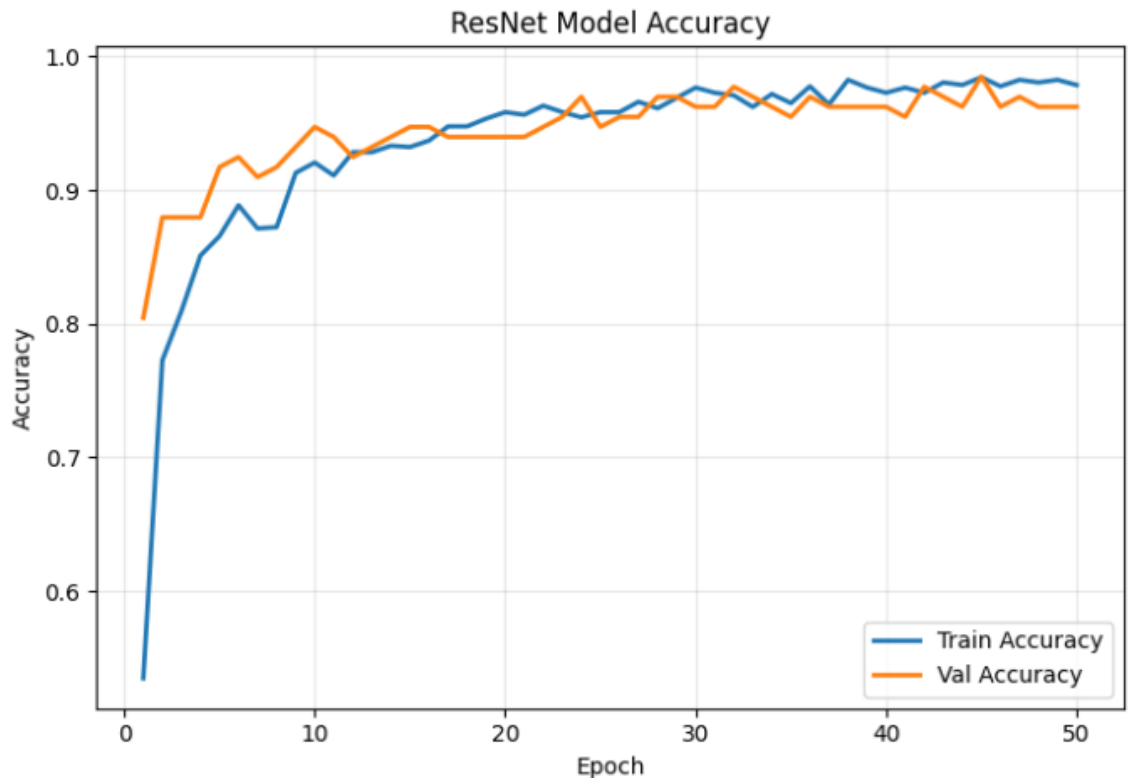
Mô hình ResNet được huấn luyện trong 50 epoch với kích thước ảnh đầu vào $224 \times 224 \times 3$.

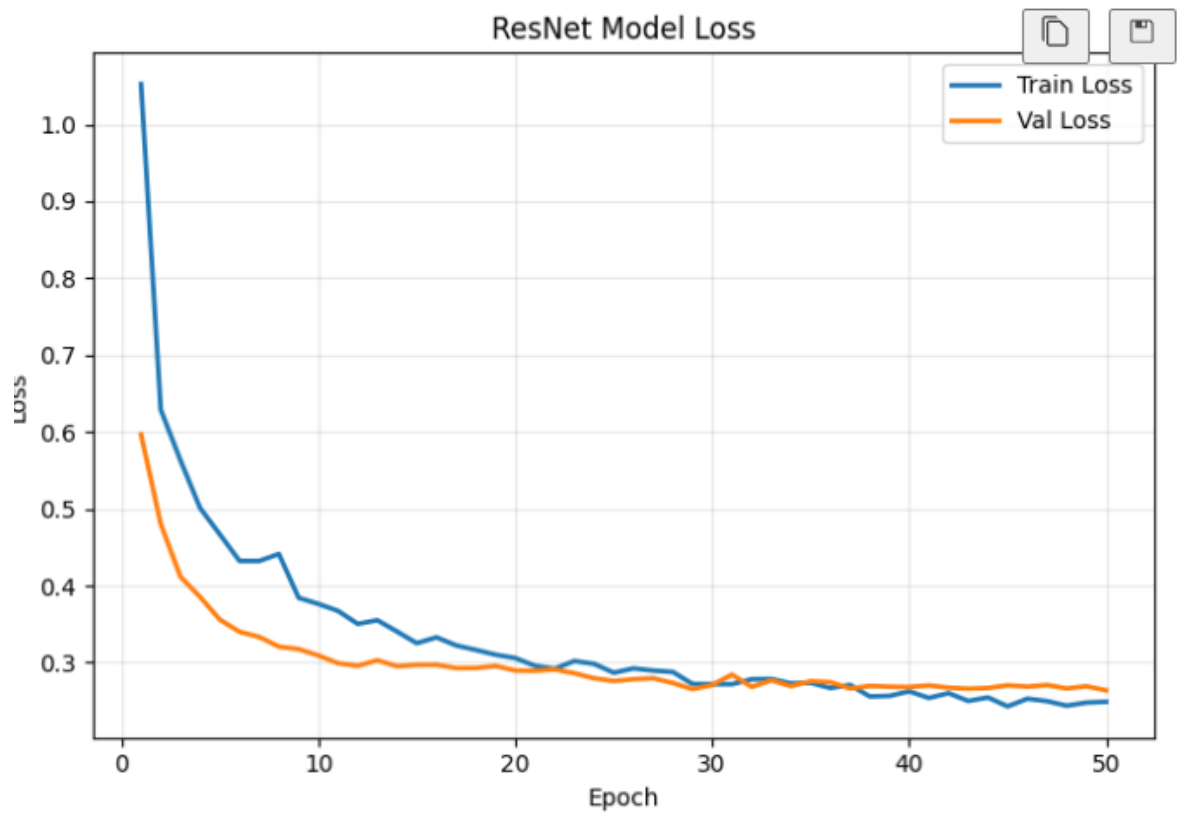
Biểu đồ Accuracy và Loss dưới đây thể hiện rõ quá trình học và khả năng hội tụ của mô hình.

Sau giai đoạn huấn luyện, mô hình hội tụ ổn định với các đặc điểm sau:

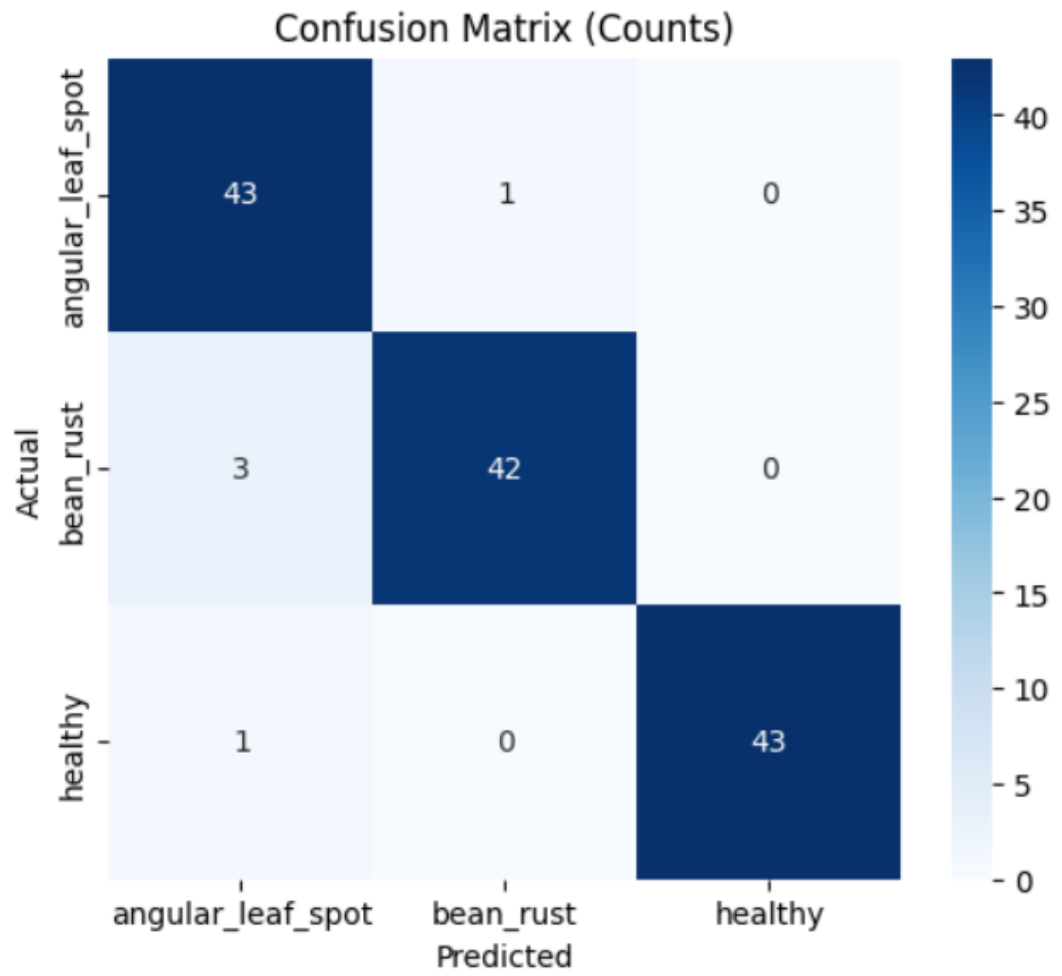
- Training Accuracy tăng nhanh trong các epoch đầu và tiếp tục cải thiện đều, đạt mức xấp xỉ 98–99% ở các epoch cuối, cho thấy mô hình học tốt các đặc trưng của dữ liệu huấn luyện.

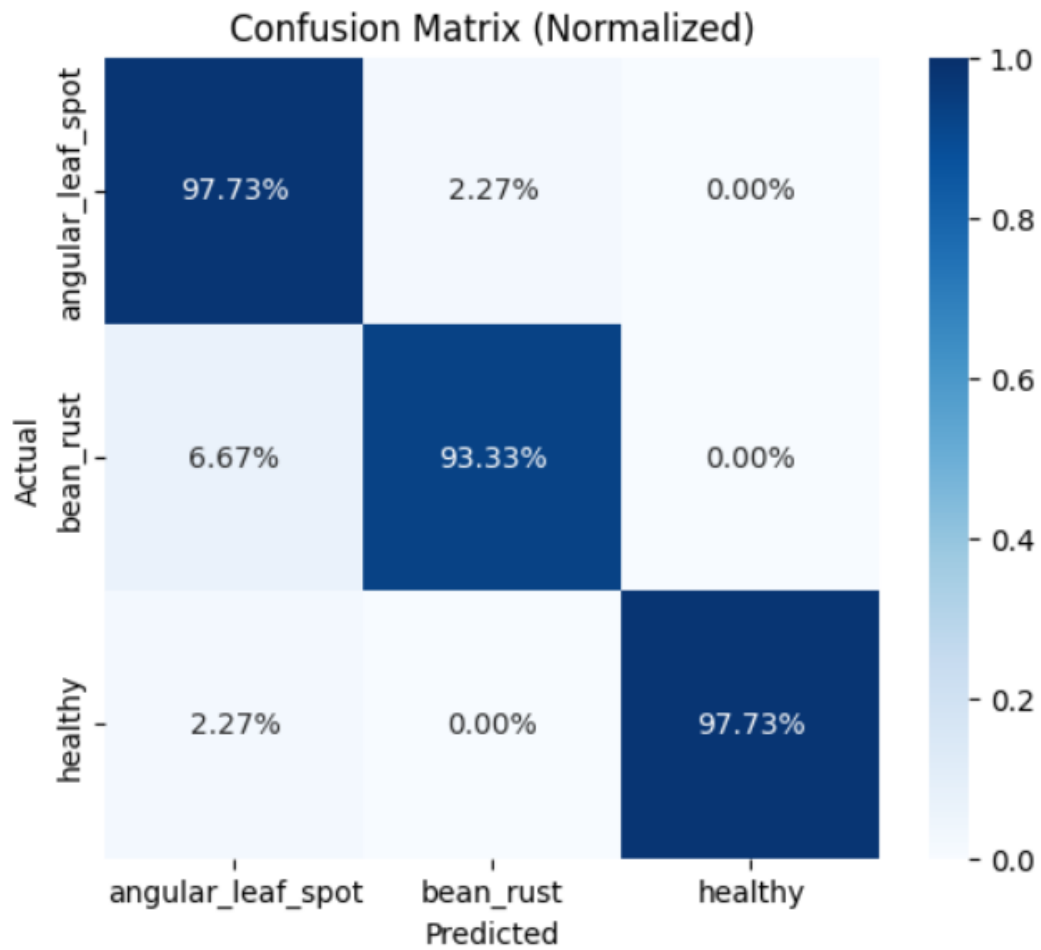
- Validation Accuracy tăng ổn định và duy trì quanh mức 95–97%, với khoảng chênh lệch nhỏ so với tập huấn luyện, phản ánh khả năng tổng quát hóa tốt của mô hình.
- Training Loss giảm mạnh từ giá trị ban đầu trên 1.0 xuống khoảng 0.25, trong khi Validation Loss giảm và ổn định quanh mức 0.26–0.28, cho thấy mô hình không xuất hiện dấu hiệu quá khớp rõ rệt.





4.3.4 Ma trận nhầm lẫn





- Đối với lớp *angular_leaf_spot*, mô hình dự đoán đúng 43/44 ảnh (97.73%), chỉ có 1 ảnh bị nhầm sang lớp *bean_rust*, cho thấy khả năng nhận diện bệnh đốm góc rất tốt.
- Lớp *bean_rust* được nhận diện đúng 42/45 ảnh (93.33%), trong đó 3 ảnh bị nhầm sang *angular_leaf_spot*. Điều này cho thấy sự nhầm lẫn chủ yếu xảy ra giữa hai loại bệnh có đặc trưng hình thái và màu sắc tương đồng.
- Với lớp *healthy*, mô hình dự đoán chính xác 43/44 ảnh (97.73%), chỉ có 1 ảnh bị nhầm sang *angular_leaf_spot*, cho thấy khả năng phân biệt lá khỏe với lá bệnh là rất cao.

4.3.6 Nhận xét tổng quan

- Tổng thể, mô hình ResNet cho thấy hiệu năng phân loại cao và ổn định trên bài toán nhận diện bệnh lá cây với ba lớp *angular_leaf_spot*, *bean_rust* và *healthy*

- Kết quả huấn luyện cho thấy mô hình hội tụ nhanh và ổn định, với training accuracy đạt gần 99% và validation accuracy duy trì ở mức 95–97%, đồng thời giá trị loss trên cả hai tập đều giảm và hội tụ, không xuất hiện dấu hiệu quá khớp rõ rệt.
- Phân tích ma trận nhầm lẫn cho thấy mô hình nhận diện rất tốt lớp healthy và *angular_leaf_spot*, với tỷ lệ dự đoán đúng trên 97%. Các lỗi phân loại chủ yếu xảy ra giữa hai lớp bệnh *angular_leaf_spot* và *bean_rust*, nguyên nhân có thể đến từ sự tương đồng về màu sắc và hình thái tổn thương trên lá.
- Nhìn chung, mô hình ResNet thể hiện khả năng tổng quát hóa tốt hơn so với CNN cơ bản, đặc biệt trong việc giảm nhầm lẫn với lớp lá khỏe, cho thấy kiến trúc sâu với các kết nối dư (residual connections) giúp trích xuất đặc trưng hiệu quả hơn cho bài toán phân loại bệnh lá cây.

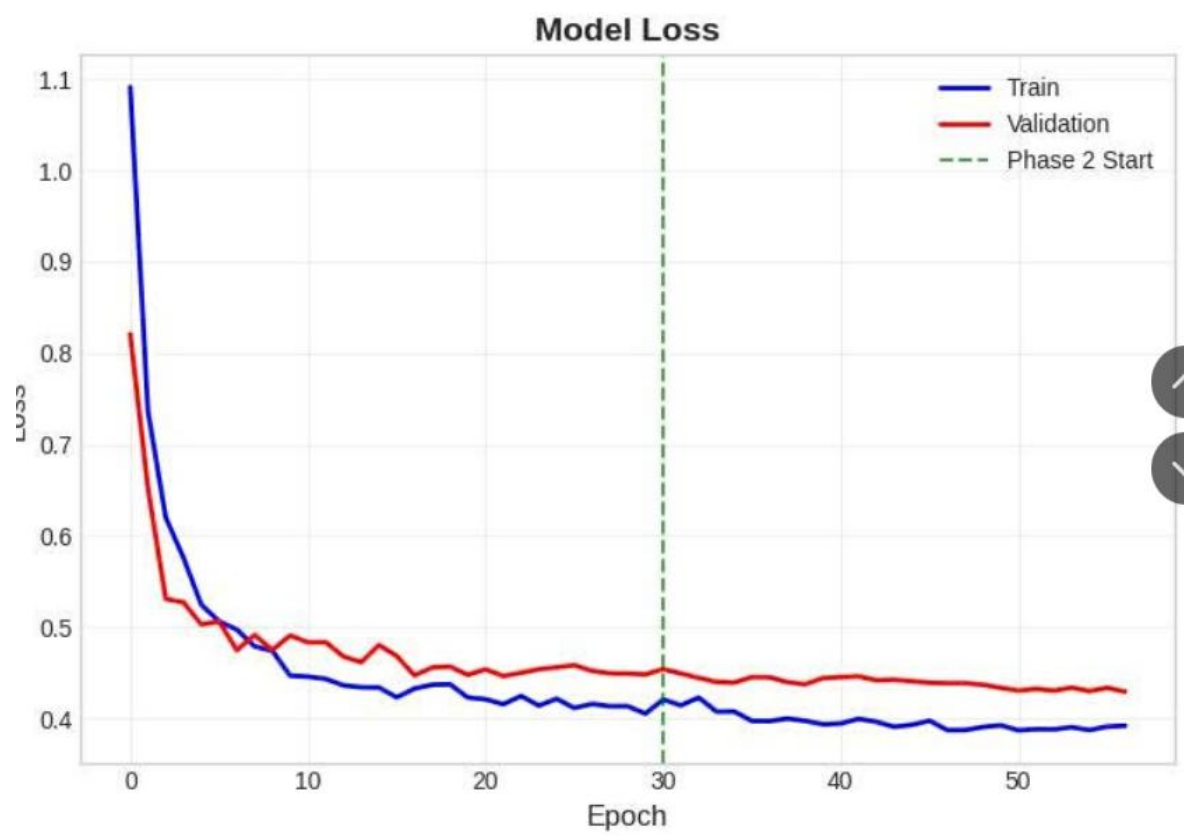
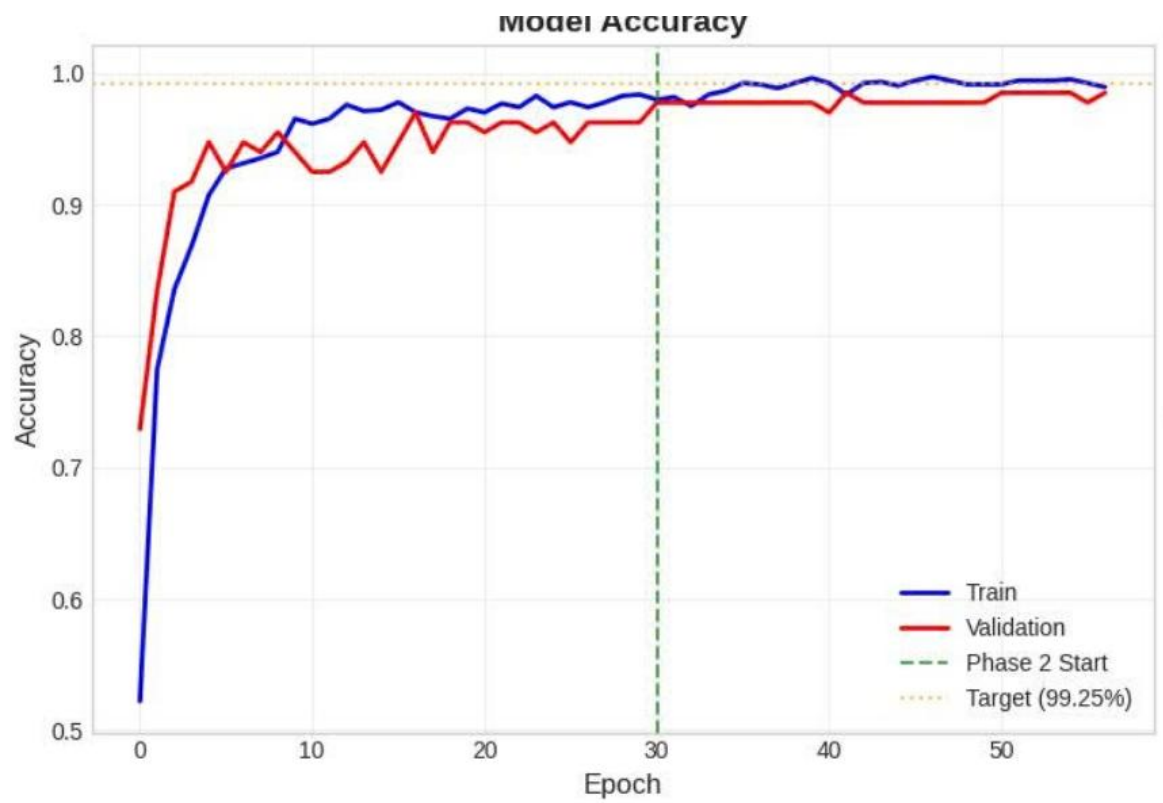
4.4 Kết quả từ mô hình MobileNet của nhóm

4.4.1 Kết quả huấn luyện của mô hình

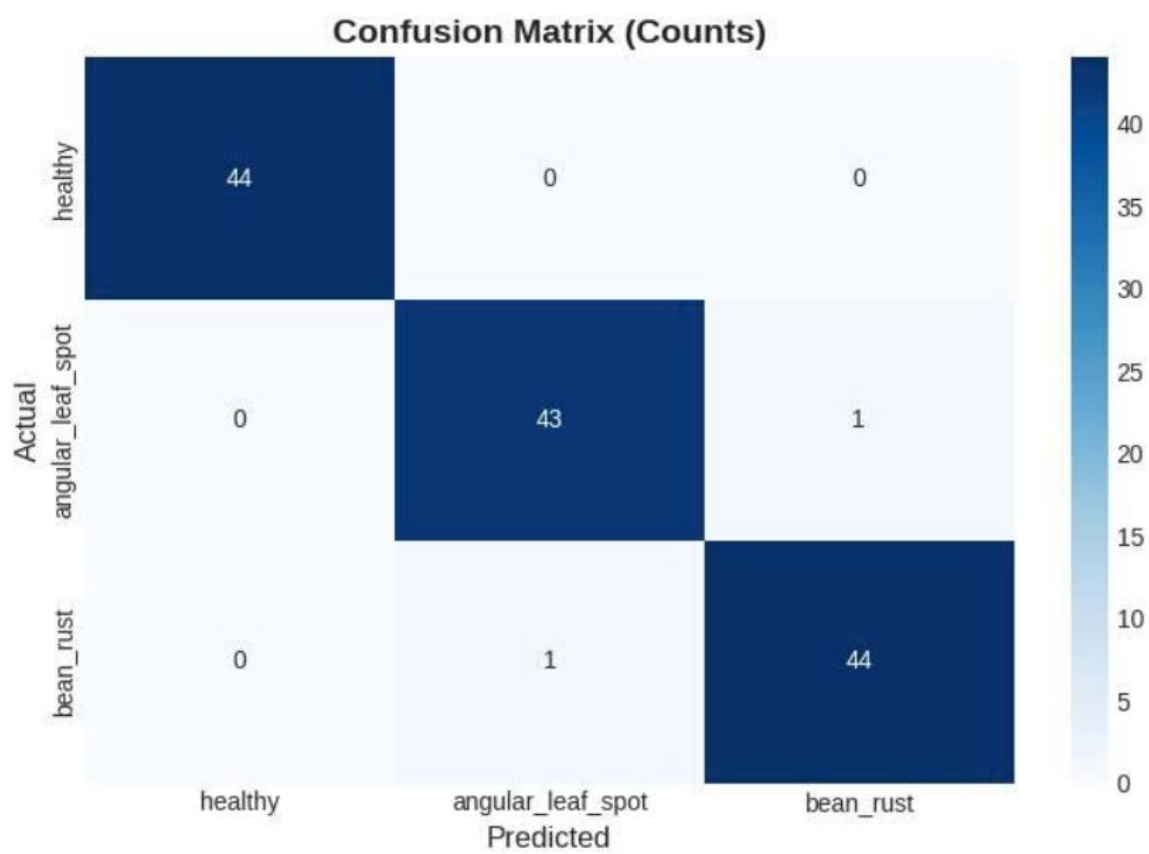
Mô hình MobileNet được huấn luyện trong 50 epoch với kích thước ảnh đầu vào $224 \times 224 \times 3$.

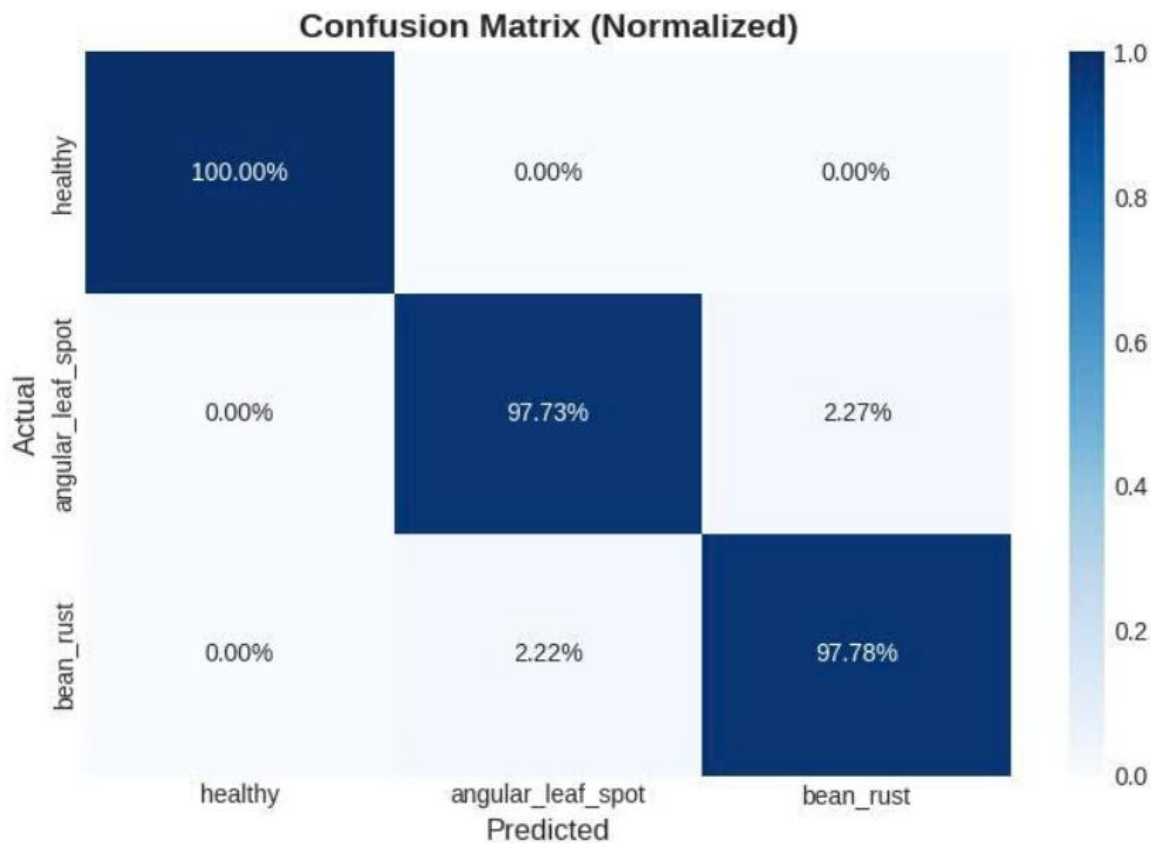
Biểu đồ Accuracy và Loss thể hiện rõ quá trình học và khả năng hội tụ của mô hình.

- Training Accuracy tăng mạnh trong các epoch đầu và tiếp tục cải thiện đều, đạt mức xấp xỉ 98–99% ở các epoch cuối, cho thấy MobileNet học hiệu quả các đặc trưng chính của dữ liệu.
- Validation Accuracy tăng nhanh và duy trì ổn định quanh mức 95–97%, với độ chênh lệch nhỏ so với tập huấn luyện, phản ánh khả năng tổng quát hóa tốt dù số lượng tham số của mô hình tương đối thấp.
- Training Loss giảm mạnh từ giá trị ban đầu trên 1.0 xuống khoảng 0.25, trong khi Validation Loss giảm và ổn định quanh mức 0.26–0.28, cho thấy mô hình không xuất hiện dấu hiệu quá khớp rõ rệt.



4.4.2 Ma trận nhầm lẫn





- Đối với lớp `angular_leaf_spot`, mô hình MobileNet dự đoán đúng 43/44 ảnh (97.73%), chỉ có 1 ảnh bị nhầm sang lớp `bean_rust`. Kết quả này cho thấy MobileNet có khả năng nhận diện bệnh đốm góc rất tốt.
- Lớp `bean_rust` được mô hình phân loại đúng 44/45 ảnh (97.78%), chỉ có 1 ảnh bị nhầm sang lớp `angular_leaf_spot`. Điều này cho thấy mức độ nhầm lẫn giữa hai lớp bệnh là rất thấp, dù chúng có một số đặc trưng hình thái tương đồng.
- Với lớp `healthy`, mô hình dự đoán chính xác 44/44 ảnh (100%), không có trường hợp nhầm lẫn sang các lớp bệnh, chứng tỏ MobileNet phân biệt rất hiệu quả giữa lá khỏe và lá bị bệnh.

4.4.3 Nhận xét tổng quan

- Tổng thể, mô hình MobileNet cho thấy hiệu năng phân loại rất cao và ổn định trên bài toán nhận diện bệnh lá cây với ba lớp `healthy`, `angular_leaf_spot` và `bean_rust`.
- Kết quả từ ma trận nhầm lẫn cho thấy mô hình dự đoán chính xác hầu hết các mẫu trong tập kiểm tra, với tỷ lệ đúng trên 97% cho cả ba lớp và đặc biệt đạt

100% độ chính xác đối với lớp healthy. Điều này chứng tỏ MobileNet phân biệt rất tốt giữa lá khỏe và lá bị bệnh.

- Các lỗi phân loại còn lại là rất ít và chủ yếu xảy ra giữa hai lớp bệnh `angular_leaf_spot` và `bean_rust`, nguyên nhân có thể do sự tương đồng về màu sắc và hình thái tổn thương trên lá. Tuy nhiên, mức độ nhầm lẫn này là không đáng kể và không ảnh hưởng lớn đến hiệu năng tổng thể của mô hình.
- Nhìn chung, MobileNet không chỉ đạt độ chính xác cao mà còn duy trì được khả năng tổng quát hóa tốt với số lượng tham số nhỏ, cho thấy mô hình rất phù hợp cho các ứng dụng thực tế yêu cầu tốc độ xử lý nhanh và tài nguyên tính toán hạn chế.

CHƯƠNG 5. ĐỊNH HƯỚNG PHÁT TRIỂN LÂU DÀI

Sau khi mô hình nhóm em đã đạt được kết quả khả quan trong việc phân loại bệnh lá cây với độ chính xác cao, hệ thống vẫn còn nhiều tiềm năng thì nhóm bọn em sẽ tiếp tục bàn, hoàn thiện và mở rộng trong tương lai nhằm nâng cao tính ứng dụng thực tiễn.

5.1. Mở rộng và đa dạng hóa dữ liệu

Trong giai đoạn tiếp theo, hệ thống có thể được cải thiện bằng cách:

- Thu thập thêm dữ liệu ảnh lá cây từ nhiều điều kiện môi trường khác nhau (ánh sáng, góc chụp, nền ảnh, thời tiết).
- Mở rộng tập dữ liệu với nhiều giống cây trồng và nhiều loại bệnh hơn, giúp mô hình học được các đặc trưng tổng quát hơn.
- Kết hợp dữ liệu từ thực tế sản xuất nông nghiệp để tăng khả năng ứng dụng trong môi trường thực.

5.2. Cải tiến mô hình và chiến lược huấn luyện

Bên cạnh các kiến trúc đã sử dụng, trong tương lai có thể:

- Thử nghiệm các mô hình tiên tiến hơn như EfficientNet, Vision Transformer (ViT) hoặc các kiến trúc lai (CNN + Attention).
- Áp dụng fine-tuning sâu hơn cho các mô hình pretrained nhằm khai thác tốt hơn đặc trưng của dữ liệu.
- Nghiên cứu các kỹ thuật ensemble learning để kết hợp nhiều mô hình, từ đó cải thiện độ chính xác và độ ổn định của hệ thống.

5.3. Phát triển giao diện người dùng (UI)

Để tăng tính thân thiện và khả năng sử dụng:

- Xây dựng giao diện người dùng trực quan cho phép người dùng tải ảnh lá cây và nhận kết quả phân loại một cách nhanh chóng.
- Hiển thị kết quả kèm theo xác suất dự đoán của từng lớp và gợi ý tình trạng bệnh.
- Tích hợp trực quan hóa Grad-CAM giúp người dùng hiểu được vùng ảnh mà mô hình tập trung khi đưa ra dự đoán.

5.4. Mở rộng hệ thống trên nền tảng web và di động

Hệ thống có thể được triển khai rộng rãi thông qua:

- Phát triển ứng dụng web sử dụng các framework như Flask, FastAPI hoặc Django để phục vụ nhiều người dùng cùng lúc.
- Tích hợp mô hình vào ứng dụng di động nhằm hỗ trợ nông dân chẩn đoán bệnh lá cây trực tiếp tại ruộng.
- Triển khai mô hình trên các thiết bị biên (edge devices) nhờ sử dụng các mô hình nhẹ như MobileNet, giúp giảm phụ thuộc vào máy chủ trung tâm.

5.5. Ứng dụng thực tiễn trong nông nghiệp thông minh

Trong tương lai, hệ thống có thể trở thành một phần của:

- Hệ thống giám sát cây trồng thông minh, kết hợp với camera hoặc drone để phát hiện bệnh sớm trên diện rộng.
- Hệ thống hỗ trợ ra quyết định trong nông nghiệp, giúp người trồng lựa chọn phương án xử lý bệnh phù hợp, kịp thời.
- Nền tảng quản lý dữ liệu nông nghiệp, lưu trữ lịch sử bệnh và theo dõi sự phát triển của cây trồng theo thời gian.