

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



MÔN HỌC: HỌC MÁY

TÊN ĐỀ TÀI: XÂY DỰNG MÔ HÌNH HỌC SÂU CHO BÀI TOÁN

PHÂN LOẠI BỆNH TRÊN GÀ TỪ ẢNH PHÂN GÀ

Giảng viên:	ThS. Vũ Thị Hạnh
Sinh viên thực hiện:	Nguyễn Văn Tân Phát Nguyễn Hoàng Lộc Trần Thiên Bảo Phạm Thành Doanh
Lớp:	S26-65TTNT

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

NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN

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

LỜI CẢM ƠN.....	2
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
2.3. Tiền xử lý (Preprocessing).....	13
2.3.1. Chuẩn hóa giá trị pixel.....	13
2.3.2. Chia tập dữ liệu.....	13
2.3.3. Xử lý mất cân bằng dữ liệu.....	13
2.3.4. Tăng cường dữ liệu (Data Augmentation).....	13
2.3.5. Data Generator.....	14
CHƯƠNG 3. PHƯƠNG PHÁP/MÔ HÌNH ÁP DỤNG (METHODOLOGY).....	14
3.1. Mạng nơ-ron tích chập (Convolutional Neural Network - CNN).....	14
3.1.1. Lớp tích chập (Convolutional Layer)	15
3.1.2. Lớp gộp (Pooling Layer)	15
3.1.3. Lớp kết nối đầy đủ (Fully Connected Layer).....	15
3.2. Hàm mất mát trong bài toán phân loại.....	15
3.3. Học chuyển giao (Transfer Learning).....	16
3.4. Điều chỉnh tốc độ học (Learning Rate Scheduling).....	16
3.5. Regularization và chống overfitting.....	16
3.6. Độ đo đánh giá mô hình.....	16
3.7. Các mô hình áp dụng.....	17
3.7.1. Mô hình CNN tự xây dựng (CNN from Scratch).....	17
3.7.2. Mô hình MobileNetV3.....	18
3.7.3. Mô hình EfficientNetB3	18
3.7.4 Mô hình Resnet50.....	19
CHƯƠNG 4. KẾT QUẢ BUỚC ĐẦU VÀ NHẬN XÉT (RESULTS & DISCUSSION).....	21
4.1. Bảng so sánh kết quả	21

4.2. Phân tích ma trận nhầm lẫn.....	22
4.2.1. Mô hình CNN from Scratch.....	22
4.2.2. Mô hình EfficientNet.....	23
4.2.3. Mô hình ResNet50.....	24
4.2.4. Mô hình MobileNetV3.....	25
4.3. Đánh giá mô hình tốt nhất.....	26
4.4. So sánh hiệu năng giữa các mô hình.....	26
4.5. Hiện tượng quan sát được.....	27
4.6. Đánh giá khả năng ứng dụng thực tế	27
CHƯƠNG 5. ĐỊNH HƯỚNG PHÁT TRIỂN.....	28
KẾT LUẬN.....	30
Tài Liệu Tham Khảo.....	31
Bảng phân chia công việc	32

DANH MỤC HÌNH ẢNH

<i>Hình 1 Biểu đồ phân bố số lượng ảnh theo từng lớp</i>	11
<i>Hình 2 Phân bố độ sáng (Brightness) và độ tương phản (Contrast) của ảnh</i> ..	11
<i>Hình 3 Phân bố kênh màu RGB</i>	12
<i>Hình 4 Ma trận nhầm lẫn (Confusion Matrix) của mô hình CNN from Scratch trên tập test</i>	22
<i>Hình 5 Ma trận nhầm lẫn (Confusion Matrix) của mô hình EfficientNet trên tập test</i>	23
<i>Hình 6 Ma trận nhầm lẫn (Confusion Matrix) của mô hình ResNet50 trên tập test</i>	24
<i>Hình 7 Ma trận nhầm lẫn (Confusion Matrix) của mô hình MobileNetV3 trên tập test</i>	25

DANH MỤC BẢNG

<i>Bảng 1 So sánh các độ đo Accuracy, Precision, Recall và F1-score của các mô hình trên tập test</i>	21
---	----

CHƯƠNG 1. GIỚI THIỆU BÀI TOÁN

1.1. Đặt vấn đề

Việt Nam là quốc gia có ngành chăn nuôi gia cầm phát triển mạnh, với tổng đàn ước tính khoảng 500–600 triệu con, trong đó gà chiếm tỷ trọng lớn. Ngành này không chỉ đóng góp quan trọng cho kinh tế nông nghiệp mà còn cung cấp nguồn protein thiết yếu và tạo sinh kế cho hàng triệu hộ gia đình. Tuy nhiên, dịch bệnh trên đàn gà vẫn là bài toán nan giải, gây thiệt hại kinh tế nghiêm trọng và tiềm ẩn rủi ro đối với sức khỏe cộng đồng thông qua chuỗi thực phẩm.

Trong thực tế chăn nuôi, một số bệnh phổ biến và nguy hiểm thường gặp ở gà có thể kể đến như: bệnh Cầu trùng (Coccidiosis) do ký sinh trùng Eimeria gây ra, khiến gà tiêu chảy có máu, suy nhược và có thể chết hàng loạt ở gà con; bệnh Salmonella do vi khuẩn Salmonella gây ra, không chỉ ảnh hưởng đến đàn gà mà còn có khả năng lây sang người; bệnh Newcastle do virus Paramyxovirus gây nên, với tỷ lệ chết rất cao và thuộc nhóm bệnh bắt buộc phải công bố dịch. Bên cạnh đó, phân của gà khỏe mạnh thường có màu sắc và hình dạng ổn định, không xuất hiện các dấu hiệu bất thường.

Hiện nay, chẩn đoán bệnh gà chủ yếu dựa trên quan sát triệu chứng lâm sàng, xét nghiệm phòng thí nghiệm hoặc mổ khám bệnh tích. Các phương pháp này đều có những hạn chế nhất định: quan sát lâm sàng phụ thuộc nhiều vào kinh nghiệm và dễ nhầm lẫn; xét nghiệm cho kết quả chính xác nhưng tốn thời gian và chi phí; trong khi mổ khám chỉ thực hiện được khi gà đã chết, không còn giá trị phòng ngừa. Vì vậy, khi bệnh được phát hiện thì dịch thường đã lây lan rộng, gây thiệt hại lớn cho người chăn nuôi.

Sự phát triển mạnh mẽ của trí tuệ nhân tạo, đặc biệt là Deep Learning và Computer Vision, đã mở ra hướng tiếp cận mới cho bài toán chẩn đoán bệnh. Các mô hình mạng nơ-ron tích chập (CNN) cho thấy khả năng nhận dạng hình ảnh với độ chính xác cao, tốc độ xử lý nhanh và chi phí triển khai thấp. Việc ứng dụng CNN trong nông nghiệp cho phép chẩn đoán bệnh tự động chỉ từ hình ảnh, phù hợp với điều kiện thực tế của các hộ chăn nuôi nhỏ lẻ.

Từ những lý do trên, đề tài “Xây Dựng Mô Hình Học Sâu Cho Bài Toán Phân Loại Bệnh Trên Gà Từ Ảnh Phân Gà” được thực hiện với mục tiêu xây dựng mô hình Deep Learning có khả năng phân loại bốn trạng thái sức khỏe của gà, bao gồm bệnh Cầu trùng, bệnh Salmonella, bệnh Newcastle và tình trạng khỏe mạnh, góp phần hỗ trợ phát hiện sớm bệnh và giảm thiểu thiệt hại trong chăn nuôi gia cầm.

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 bốn tình trạng sức khỏe của gà từ hình ảnh phân 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 phân gà, 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 Kaggle Notebook, cho phép viết mã, chạy thử nghiệm và ghi chú kết quả một cách linh hoạt. Kaggle 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.

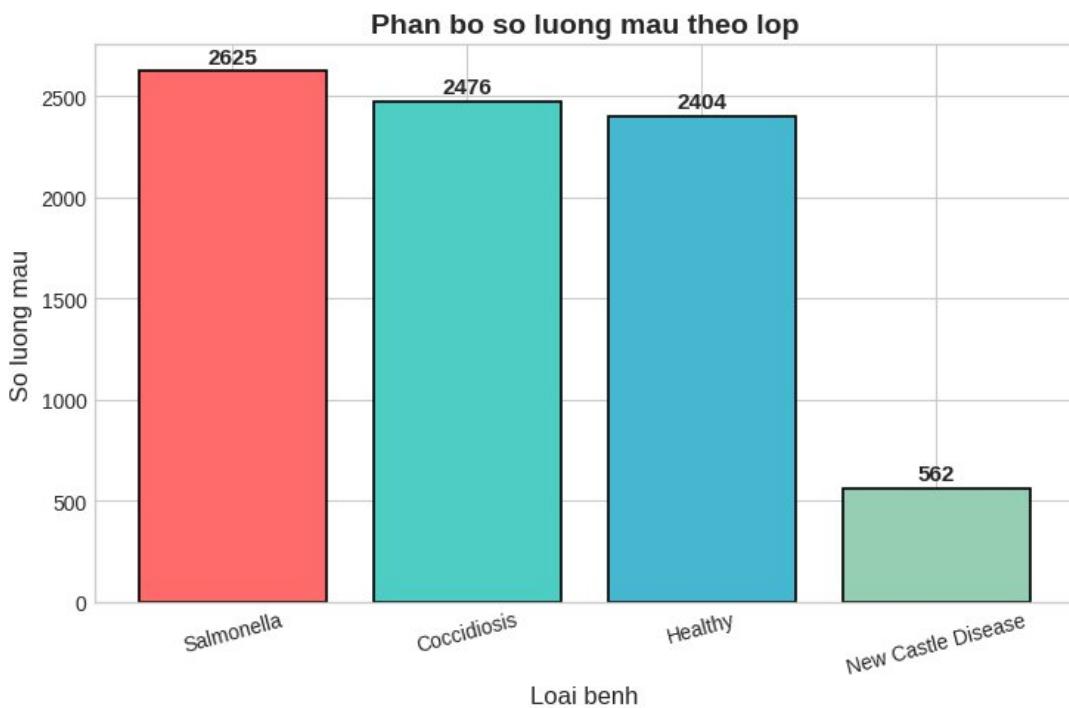
CHƯƠNG 2. MÔ TẢ VÀ TIỀN XỬ LÝ DỮ LIỆU (DATA & PREPROCESSING)

2.1. Nguồn dữ liệu

Bộ dữ liệu được sử dụng trong đề tài là tập hợp các hình ảnh phân gà được thu thập và công bố công khai trên nền tảng Kaggle, một cộng đồng khoa học dữ liệu lớn nhất thế giới. Bộ dữ liệu này được xây dựng với mục đích phục vụ nghiên cứu ứng dụng thị giác máy tính trong lĩnh vực chẩn đoán bệnh gia cầm, bao gồm các hình ảnh chụp phân gà trong điều kiện thực tế tại các trang trại chăn nuôi. Tổng số lượng hình ảnh trong bộ dữ liệu là 8.067 ảnh, được phân chia thành bốn lớp tương ứng với bốn tình trạng sức khỏe khác nhau. Lớp Coccidiosis (bệnh Cầu trùng) chứa khoảng 2.103 hình ảnh với tên file có tiền tố "cocci". Lớp Salmonella chứa khoảng 2.261 hình ảnh với tiền tố "salmo". Lớp Healthy (khỏe mạnh) chứa khoảng 2.057 hình ảnh với tiền tố "healthy". Lớp New Castle Disease (bệnh Newcastle) có số lượng ít nhất với khoảng 394 hình ảnh, bao gồm các file có tiền tố "ncd" và "pcrncd".

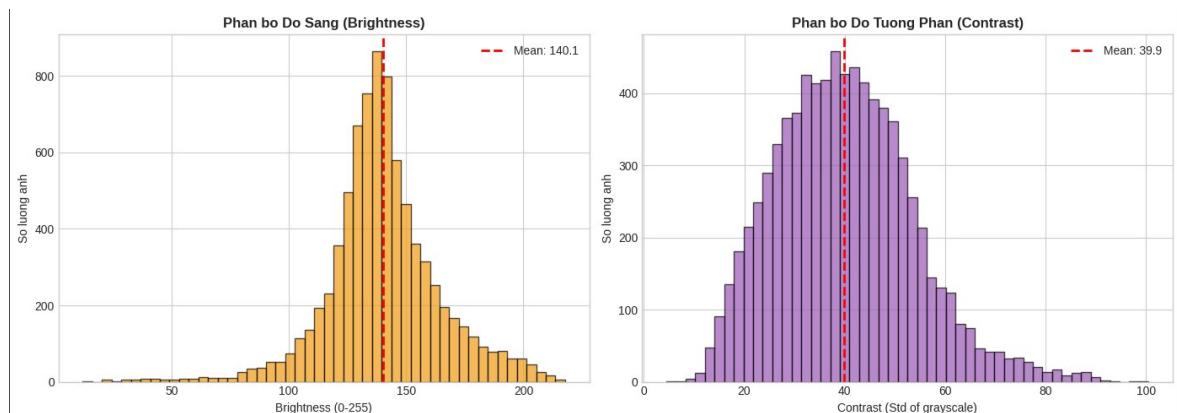
Qua phân tích phân bố dữ liệu, có thể nhận thấy sự mất cân bằng đáng kể giữa các lớp, đặc biệt lớp New Castle Disease chỉ chiếm khoảng 5% tổng số mẫu trong khi các lớp còn lại chiếm từ 25% đến 28%. Đây là một thách thức cần được xử lý trong quá trình huấn luyện mô hình để tránh hiện tượng mô hình thiên vị về các lớp có nhiều mẫu hơn. Các hình ảnh trong bộ dữ liệu có định dạng JPEG với kích thước và độ phân giải đa dạng. Hình ảnh được chụp trong các điều kiện ánh sáng và góc chụp khác nhau, phản ánh tính đa dạng của môi trường chăn nuôi thực tế. Một số hình ảnh có chất lượng cao với độ sắc nét tốt, trong khi một số khác có độ nhiễu hoặc mờ nhất định, đòi hỏi mô hình phải có khả năng tổng quát hóa tốt để xử lý các trường hợp đa dạng này.

2.2. Phân tích khám phá dữ liệu (EDA)



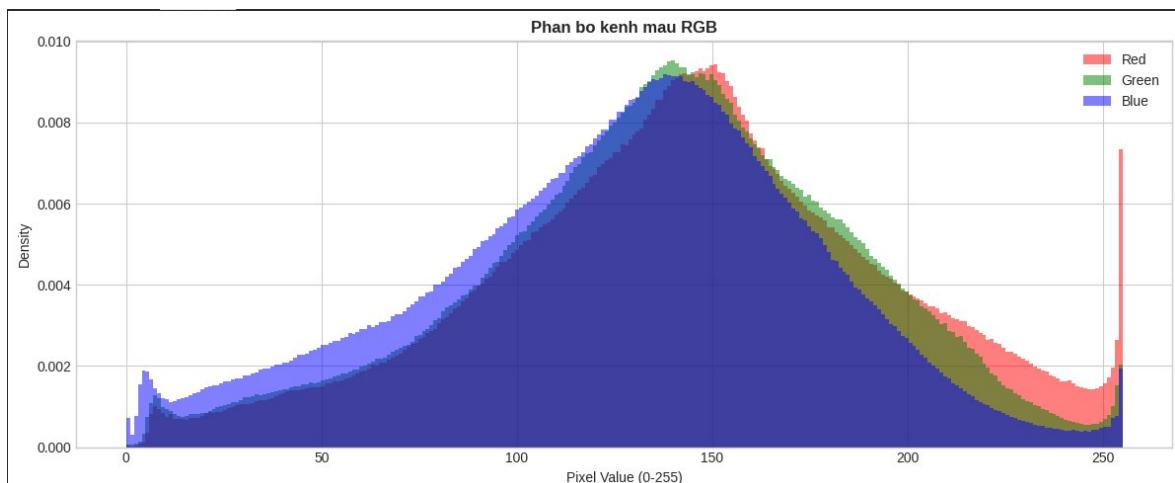
Hình 1 Biểu đồ phân bố số lượng ảnh theo từng lớp

Hình 1 cho thấy sự phân bố số lượng ảnh giữa các lớp bệnh trong bộ dữ liệu. Kết quả thống kê cho thấy bộ dữ liệu tồn tại sự mất cân bằng đáng kể giữa các lớp. Lớp Salmonella chiếm tỷ lệ cao nhất với khoảng 28% tổng số mẫu, tiếp theo là lớp Coccidiosis với khoảng 26%, lớp Healthy với khoảng 25%. Đặc biệt, lớp New Castle Disease chỉ chiếm khoảng 5% tổng số mẫu, tương đương với 394 hình ảnh. Sự chênh lệch này có thể ảnh hưởng đến khả năng học của mô hình đối với lớp thiểu số, do đó cần áp dụng các kỹ thuật xử lý mất cân bằng dữ liệu trong quá trình huấn luyện.



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

Phân tích đặc trưng hình ảnh cho thấy độ sáng (brightness) của bộ dữ liệu có phân bố gần dạng chuẩn, với giá trị tập trung chủ yếu trong khoảng 120–150 và giá trị trung bình đạt 140,1 trên thang đo 0–255, cho thấy chất lượng ánh sáng nhìn chung là tốt và ổn định, rất ít ảnh quá tối hoặc quá sáng. Do đó, việc điều chỉnh độ sáng mạnh là không cần thiết, chỉ cần áp dụng các biến đổi nhẹ trong giai đoạn tăng cường dữ liệu để tạo sự đa dạng. Ngược lại, phân bố độ tương phản (contrast) cho thấy xu hướng lệch phải, với khoảng 25% số ảnh nằm trong vùng contrast thấp (10–30) và giá trị trung bình chỉ khoảng 39,9, phản ánh sự tồn tại của nhiều ảnh mờ, ít chi tiết, có thể do điều kiện ánh sáng và thiết bị chụp trong môi trường chăn nuôi thực tế. Đặc điểm này vừa là thách thức đối với quá trình học đặc trưng của mô hình, vừa phản ánh tính thực tiễn của bộ dữ liệu.



Hình 3 Phân bố kênh màu RGB

Phân tích phân bố kênh màu RGB cho thấy ba kênh có mức chồng lấp mạnh trong khoảng giá trị 100–170, phản ánh việc hình ảnh chủ yếu mang các tông màu nâu và be, đặc trưng cho phân gà và nền trấu trong môi trường chăn nuôi. Ở vùng giá trị cao (150–250), cường độ kênh đỏ (Red) lớn hơn kênh xanh lá (Green) và xanh dương (Blue), cho thấy bộ dữ liệu có xu hướng nghiêng về tông màu ám, phù hợp với màu sắc tự nhiên của phân gà như nâu, cam hoặc đỏ nâu. Ngoài ra, sự xuất hiện của các đỉnh giá trị cao ở kênh đỏ tại mức 255 cho thấy một số hình ảnh có vùng bão hòa màu

đỏ, có thể liên quan đến sự hiện diện của máu trong phân bệnh hoặc ảnh hưởng từ điều kiện chiếu sáng. Trong khi đó, kênh xanh dương xuất hiện một đỉnh nhỏ tại giá trị gần 0, phản ánh sự tồn tại của các vùng tối thiếu ánh sáng xanh. Các đặc điểm này cho thấy màu sắc của bộ dữ liệu phản ánh đúng điều kiện thực tế và đóng vai trò quan trọng trong việc phân biệt các trạng thái bệnh, do đó đề tài không áp dụng các kỹ thuật tăng cường màu sắc mạnh nhằm tránh làm sai lệch thông tin màu quan trọng phục vụ cho quá trình phân loại.

2.3. Tiền xử lý (Preprocessing)

2.3.1. Chuẩn hóa giá trị pixel

Bộ dữ liệu hình ảnh phân gà đã được chuẩn hóa sẵn với kích thước đồng nhất 224x224 pixel cho tất cả các hình ảnh, phù hợp với đầu vào tiêu chuẩn của kiến trúc MobileNetV3. Giá trị pixel của hình ảnh ban đầu nằm trong khoảng từ 0 đến 255 được chuẩn hóa về khoảng từ 0 đến 1 thông qua phép chia cho 255, giúp quá trình huấn luyện ổn định hơn và tăng tốc độ hội tụ của thuật toán tối ưu hóa.

2.3.2. Chia tập dữ liệu

Bộ dữ liệu được chia thành ba tập con theo tỷ lệ 70% cho tập huấn luyện, 15% cho tập kiểm định và 15% cho tập kiểm tra. Việc chia dữ liệu được thực hiện theo phương pháp phân tầng (Stratified Split) để đảm bảo tỷ lệ các lớp trong mỗi tập con tương đương với tỷ lệ trong bộ dữ liệu gốc, đặc biệt quan trọng đối với lớp New Castle Disease có số lượng mẫu ít.

2.3.3. Xử lý mất cân bằng dữ liệu

Để giải quyết vấn đề mất cân bằng dữ liệu với lớp New Castle Disease chỉ chiếm khoảng 5% tổng số mẫu, kỹ thuật oversampling được áp dụng trên tập huấn luyện bằng cách sao chép ngẫu nhiên các hình ảnh thuộc lớp thiểu số

2.3.4. Tăng cường dữ liệu (Data Augmentation)

Kỹ thuật tăng cường dữ liệu được áp dụng trên tập huấn luyện nhằm tăng tính đa dạng của dữ liệu và giảm hiện tượng overfitting. Các phép biến đổi bao gồm xoay ảnh ngẫu nhiên trong khoảng 0 đến 40 độ, lật ảnh theo chiều ngang và chiều dọc, dịch chuyển

ảnh theo chiều rộng và chiều cao với tỷ lệ tối đa 20%, phóng to hoặc thu nhỏ ảnh với tỷ lệ từ 80% đến 120%, và điều chỉnh độ sáng ngẫu nhiên.

2.3.5. Data Generator

Data Generator được sử dụng để tải và xử lý dữ liệu theo từng batch với kích thước 32 trong quá trình huấn luyện, giúp tiết kiệm bộ nhớ và cho phép huấn luyện trên các máy tính có cấu hình hạn chế.

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 đúng một lớp trong tập các lớp đã 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 phân gà và **đầu ra (Output)** là một trong bốn nhãn tương ứng với tình trạng sức khỏe của gà: *Coccidiosis*, *Salmonella*, *Newcastle Disease* hoặc *Healthy*. Đây là bài toán có độ khó cao do sự tương đồng về hình thái giữa các lớp và sự nhiễu lớn từ môi trường chăn nuôi thực tế.

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 nhiều 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 CNN do nhóm xây dựng từ đầu là một kiến trúc mạng nơ-ron tích chập tùy chỉnh, được thiết kế dựa trên nguyên lý của Residual Network (ResNet) nhằm nâng cao khả năng học sâu trên bộ dữ liệu hình ảnh phân gà. Thay vì chỉ xếp chồng các lớp tích chập theo kiểu tuần tự, vốn dễ gặp vấn đề vanishing gradient khi mạng trở nên sâu hơn, mô hình tích hợp các kết nối tắt (skip connections) thông qua các phép cộng (Add). Cơ chế này cho phép thông tin và gradient được truyền trực tiếp qua các khối mạng, giúp quá trình huấn luyện ổn định và hiệu quả hơn khi tăng độ sâu của kiến trúc.

Mô hình có tổng số tham số khoảng 11,2 triệu, nhận đầu vào là ảnh màu kích thước chuẩn $224 \times 224 \times 3$ và thực hiện trích xuất đặc trưng theo cách phân cấp. Kiến trúc bắt đầu với một Stem block gồm các lớp Conv2D và MaxPooling nhằm nhanh chóng giảm kích thước không gian của ảnh xuống 56×56 . Sau đó, dữ liệu lần lượt đi qua bốn giai đoạn xử lý chính (stages) với số lượng bộ lọc tăng dần là 64, 128, 256 và 512, cho phép mô hình học các đặc trưng từ mức thấp đến mức cao.

Một điểm nổi bật của kiến trúc là việc áp dụng Batch Normalization ngay sau các lớp tích chập, giúp chuẩn hóa phân phối dữ liệu trung gian, ổn định quá trình huấn luyện và đẩy nhanh tốc độ hội tụ. Ở phần cuối của mạng trích xuất đặc trưng, thay vì sử dụng lớp Flatten truyền thống dễ làm tăng số lượng tham số và gây overfitting, mô hình sử dụng Global Average Pooling (GAP) để nén feature map cuối cùng từ kích thước $7 \times 7 \times 512$ thành một vector đặc trưng duy nhất. Cách tiếp cận này giúp giảm độ phức tạp của mô hình và cải thiện khả năng tổng quát hóa.

Khối phân loại (Classification Head) bao gồm một lớp Dense với 128 đơn vị, kết hợp Dropout nhằm giảm hiện tượng overfitting, trước khi đưa ra dự đoán thông qua lớp Softmax tương ứng với bốn lớp bệnh gà. Quá trình huấn luyện mô hình sử dụng bộ tối ưu Adam với tốc độ học $1e^{-4}$, kết hợp Gradient Clipping để ngăn chặn hiện tượng bùng nổ gradient, đảm bảo tính ổn định và độ tin cậy cho mô hình CNN tự xây dựng.

Quá trình huấn luyện sử dụng hàm mất mát Categorical Cross-Entropy. Chiến lược tối ưu hóa được thiết kế theo hướng ổn định và hiệu quả, bao gồm Warmup Scheduler để tăng dần tốc độ học trong các epoch đầu nhằm tránh hiện tượng sốt trọng số, kết hợp với ReduceLROnPlateau để tự động giảm learning rate khi mô hình đạt trạng thái bão hòa. Cách tiếp cận này giúp mô hình hội tụ ổn định và đạt hiệu suất phân loại cao trên tập dữ liệu bệnh gà.

3.7.2. Mô hình MobileNetV3

Để giải quyết bài toán phân loại bệnh gà với yêu cầu cân bằng giữa độ chính xác và hiệu năng tính toán, nhóm nghiên cứu lựa chọn mô hình MobileNetV3Large làm kiến trúc chính. Đây là mạng nơ-ron tích chập hiện đại được tối ưu cho các hệ thống có tài nguyên hạn chế, thông qua việc sử dụng các khối Inverted Residual kết hợp với cơ chế Squeeze-and-Excitation, giúp giảm chi phí tính toán trong khi vẫn duy trì khả năng trích xuất đặc trưng hiệu quả. Thay vì huấn luyện mô hình từ đầu, chiến lược Transfer Learning được áp dụng bằng cách khởi tạo MobileNetV3Large với trọng số huấn luyện sẵn trên bộ dữ liệu ImageNet, cho phép tận dụng tri thức đã học từ hàng triệu hình ảnh đa dạng. Trong quá trình huấn luyện, toàn bộ các lớp của mô hình nền đều được mở đóng băng (unfreeze) để thực hiện fine-tuning toàn phần, giúp mô hình thích nghi tốt hơn với đặc điểm riêng của dữ liệu hình ảnh phân gà.

Về mặt kiến trúc, các đặc trưng trích xuất từ MobileNetV3Large được đưa qua lớp Global Average Pooling (GAP) nhằm nén thông tin không gian và giảm số lượng tham số so với việc sử dụng lớp Flatten truyền thống. Phần phân loại được thiết kế lại với một lớp Dense 128 đơn vị, sử dụng hàm kích hoạt Swish nhằm cải thiện khả năng lan truyền đạo hàm trong mạng sâu. Để kiểm soát hiện tượng quá khớp, mô hình kết hợp L2 Regularization và Dropout với tỷ lệ 0.5 trước khi đi qua lớp Softmax cho bài toán phân loại bốn lớp.

Quá trình huấn luyện sử dụng hàm mất mát Categorical Cross-Entropy có tích hợp Label Smoothing với hệ số 0.05, giúp giảm sự tự tin thái quá của mô hình đối với các mẫu nhiễu. Chiến lược tối ưu hóa được thiết kế theo hướng ổn định và hiệu quả, bao gồm Warmup Scheduler để tăng dần tốc độ học trong các epoch đầu nhằm tránh hiện tượng sốc trọng số, kết hợp với ReduceLROnPlateau để tự động giảm learning rate khi mô hình đạt trạng thái bão hòa. Cách tiếp cận này giúp mô hình hội tụ ổn định và đạt hiệu suất phân loại cao trên tập dữ liệu bệnh gà.

3.7.3. Mô hình EfficientNetB3

Tiếp nối quá trình thử nghiệm các kiến trúc học sâu tiên tiến, nhóm nghiên cứu triển khai mô hình EfficientNetB3 trên nền tảng PyTorch nhằm khai thác các kỹ thuật tối ưu hóa hiện đại. Kiến trúc EfficientNetB3 làm backbone chính nhờ khả năng cân bằng tối ưu giữa hiệu suất và chi phí tính toán thông qua nguyên lý Compound Scaling, với kích thước đầu vào chuẩn được thiết lập là 300x300 pixel giúp nắm bắt

tốt các chi tiết nhỏ trên ảnh phân gà . Chiến lược Transfer Learning được áp dụng bằng cách khởi tạo trọng số từ ImageNet nhưng cho phép tinh chỉnh toàn bộ các lớp (unfreeze all layers) ngay từ đầu, giúp mô hình nhanh chóng thích nghi với các đặc trưng chuyên biệt của bộ dữ liệu bệnh gà thay vì chỉ huấn luyện lớp phân loại cuối cùng.

Để xử lý triệt để tình trạng mất cân bằng dữ liệu giữa các lớp bệnh, nghiên cứu áp dụng kỹ thuật Random Oversampling ngay trên tập huấn luyện nhằm nhân bản ngẫu nhiên các mẫu thiểu số cho đến khi cân bằng số lượng với lớp đa số.. Bên cạnh đó, tính tổng quát hóa của mô hình được củng cố bằng chiến lược tăng cường dữ liệu mạnh mẽ (Heavy Augmentation) bao gồm xoay ảnh 90 độ, lật, biến đổi hình học Affine và điều chỉnh màu sắc (ColorJitter), giúp mô hình hoạt động bền vững trước các thay đổi về góc chụp và điều kiện ánh sáng trong thực tế.

Cấu hình huấn luyện sử dụng hàm mất mát CrossEntropyLoss có tích hợp Label Smoothing để hạn chế hiện tượng overfitting và giảm sự tự tin thái quá của mô hình đối với các mẫu nhiễu. Bộ tối ưu AdamW được sử dụng weight decay kết hợp cùng cơ chế điều chỉnh tốc độ học ReduceLROnPlateau để tự động giảm learning rate khi loss không cải thiện sau 3 epochs, đồng thời áp dụng Early Stopping với độ kiên nhẫn 7 epochs để dừng huấn luyện sớm nhằm đảm bảo mô hình hội tụ tại điểm tối ưu nhất.

3.7.4 Mô hình Resnet50

Mô hình được xây dựng dựa trên kiến trúc mạng nơ-ron tích chập sâu ResNet50 theo phương pháp Transfer Learning, trong đó bộ trọng số huấn luyện trước trên tập dữ liệu ImageNet được sử dụng làm nền tảng trích xuất đặc trưng (backbone). Để phù hợp với bài toán phân loại bệnh gà, các lớp phân loại gốc của ResNet50 được loại bỏ và thay thế bằng một khối phân loại tùy chỉnh (Classification Head). Khối này bao gồm lớp GlobalAveragePooling2D nhằm giảm chiều không gian của đặc trưng, tiếp theo là lớp Batch Normalization để ổn định quá trình huấn luyện. Đặc trưng sau đó được đưa qua một lớp Dense gồm 128 nơ-ron với hàm kích hoạt ReLU, kết hợp L2 Regularization ($1e^{-4}$) và Dropout với tỷ lệ 0.3 nhằm hạn chế hiện tượng quá khớp, trước khi đi vào lớp đầu ra sử dụng hàm kích hoạt Softmax cho bài toán phân loại bốn lớp.

Chiến lược huấn luyện mô hình được chia thành hai giai đoạn nhằm đảm bảo quá trình hội tụ ổn định và hiệu quả. Ở giai đoạn thứ nhất, chỉ các lớp phân loại mới được huấn luyện trong khi toàn bộ backbone ResNet50 được đóng băng, nhằm bảo toàn các đặc trưng tổng quát đã học từ ImageNet. Giai đoạn này sử dụng bộ tối ưu Adam với tốc độ học $3e^{-4}$, kết hợp gradient clipping để tránh hiện tượng bùng nổ gradient. Sang giai đoạn thứ hai, kỹ thuật fine-tuning được thực hiện bằng cách mở đóng băng 30 lớp cuối của backbone, cho phép mô hình điều chỉnh các đặc trưng cấp cao phù hợp hơn với dữ liệu bệnh gà, với tốc độ học được giảm xuống $3e^{-5}$.

Trong cả hai giai đoạn, hàm mất mát Categorical Cross-Entropy được sử dụng cùng với Label Smoothing (0.1) nhằm cải thiện khả năng tổng quát hóa và giảm sự tự tin thái quá của mô hình. Quá trình huấn luyện được kiểm soát tự động thông qua các cơ chế EarlyStopping và ReduceLROnPlateau, giúp ngăn chặn overfitting và tối ưu hiệu suất mô hình.

CHƯƠNG 4. KẾT QUẢ BUỚC ĐẦU VÀ NHẬN XÉT (RESULTS & DISCUSSION)

4.1. Bảng so sánh kết quả

Sau quá trình huấn luyện và đánh giá các mô hình học sâu trên tập dữ liệu kiểm tra, nhóm tiến hành so sánh hiệu năng của từng mô hình thông qua các độ đo phổ biến trong bài toán phân loại, bao gồm Accuracy, Precision, Recall và F1-score. Đây là các chỉ số phản ánh toàn diện khả năng dự đoán của mô hình, đặc biệt trong bối cảnh bộ dữ liệu tồn tại sự mất cân bằng giữa các lớp bệnh.

Kết quả so sánh hiệu năng của các mô hình được trình bày trong Bảng 1:

Bảng 1 So sánh các độ đo Accuracy, Precision, Recall và F1-score của các mô hình trên tập test

Mô hình	Accuracy	Precision	Recall	F1-score
CNN from Scratch	96.78%	96%	97%	97%
MobileNetV3	97.36%	98%	98%	98%
ResNet50	97.44%	97%	97%	97%
EfficientNetB3	98%	98%	98%	98%

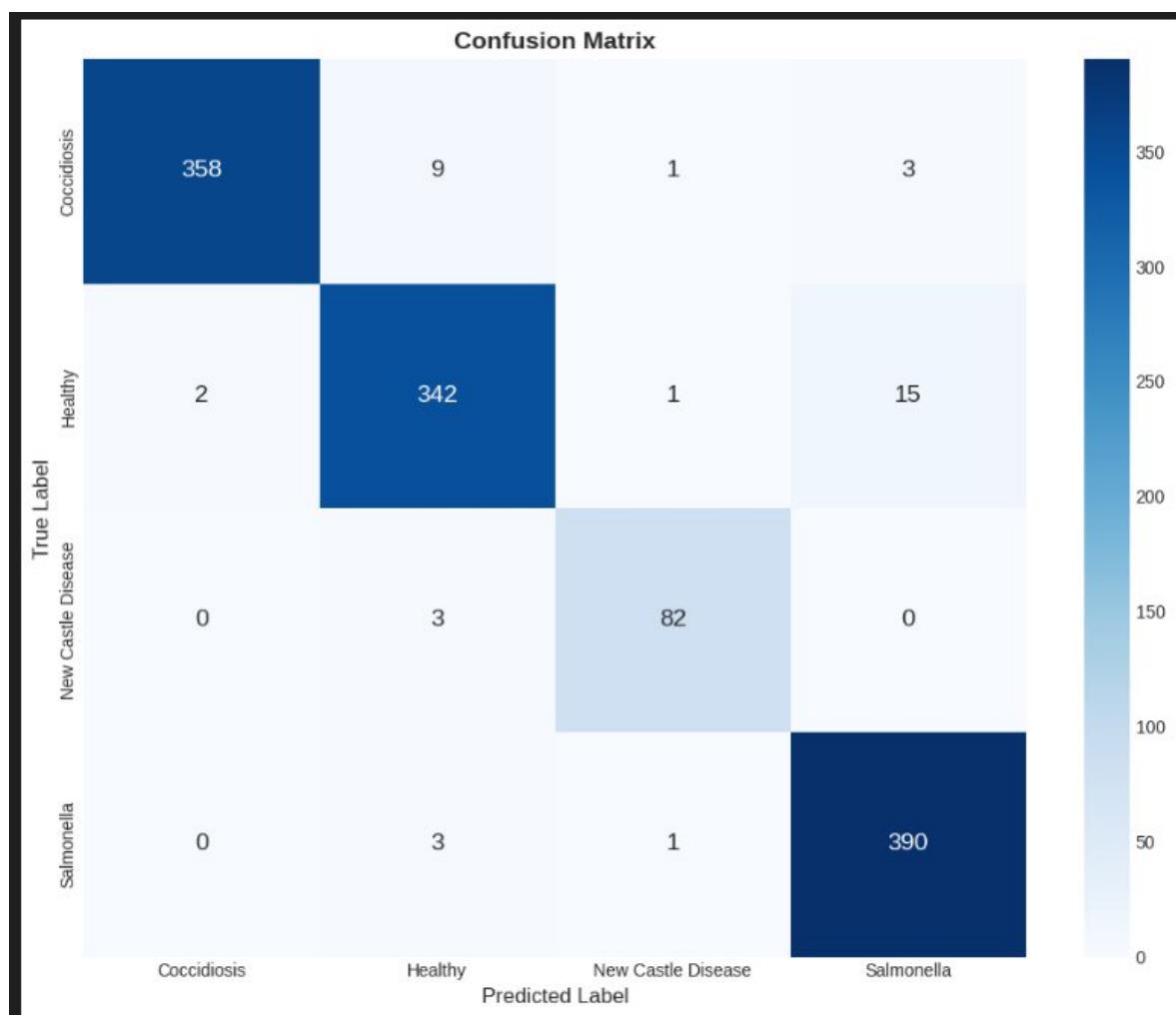
Kết quả thực nghiệm được trình bày trong Bảng 1, cho thấy sự khác biệt rõ rệt về hiệu năng giữa mô hình CNN xây dựng từ đầu và các mô hình sử dụng chiến lược học chuyển giao (Transfer Learning).

Nhìn chung, tất cả các mô hình đều đạt độ chính xác cao (trên 96%), chứng tỏ bài toán phân loại bệnh gà từ ảnh phân gà là khả thi và các đặc trưng hình ảnh chứa thông tin đủ mạnh để mô hình học sâu khai thác. Tuy nhiên, mức độ ổn định và khả năng tổng quát hóa của từng mô hình có sự khác nhau đáng kể, đặc biệt khi xét đến các lớp thiểu số như Newcastle Disease.

Trong số các mô hình được thử nghiệm, EfficientNetB3 đạt kết quả tốt nhất với Accuracy và F1-score xấp xỉ 98%, cho thấy sự cân bằng tốt giữa khả năng nhận diện đúng các lớp phổ biến và các lớp hiếm. MobileNetV3 và ResNet50 cũng cho kết quả rất khả quan, trong khi mô hình CNN from Scratch có hiệu năng thấp hơn, dù vẫn đạt mức chấp nhận được.

4.2. Phân tích ma trận nhầm lẫn

4.2.1. Mô hình CNN from Scratch

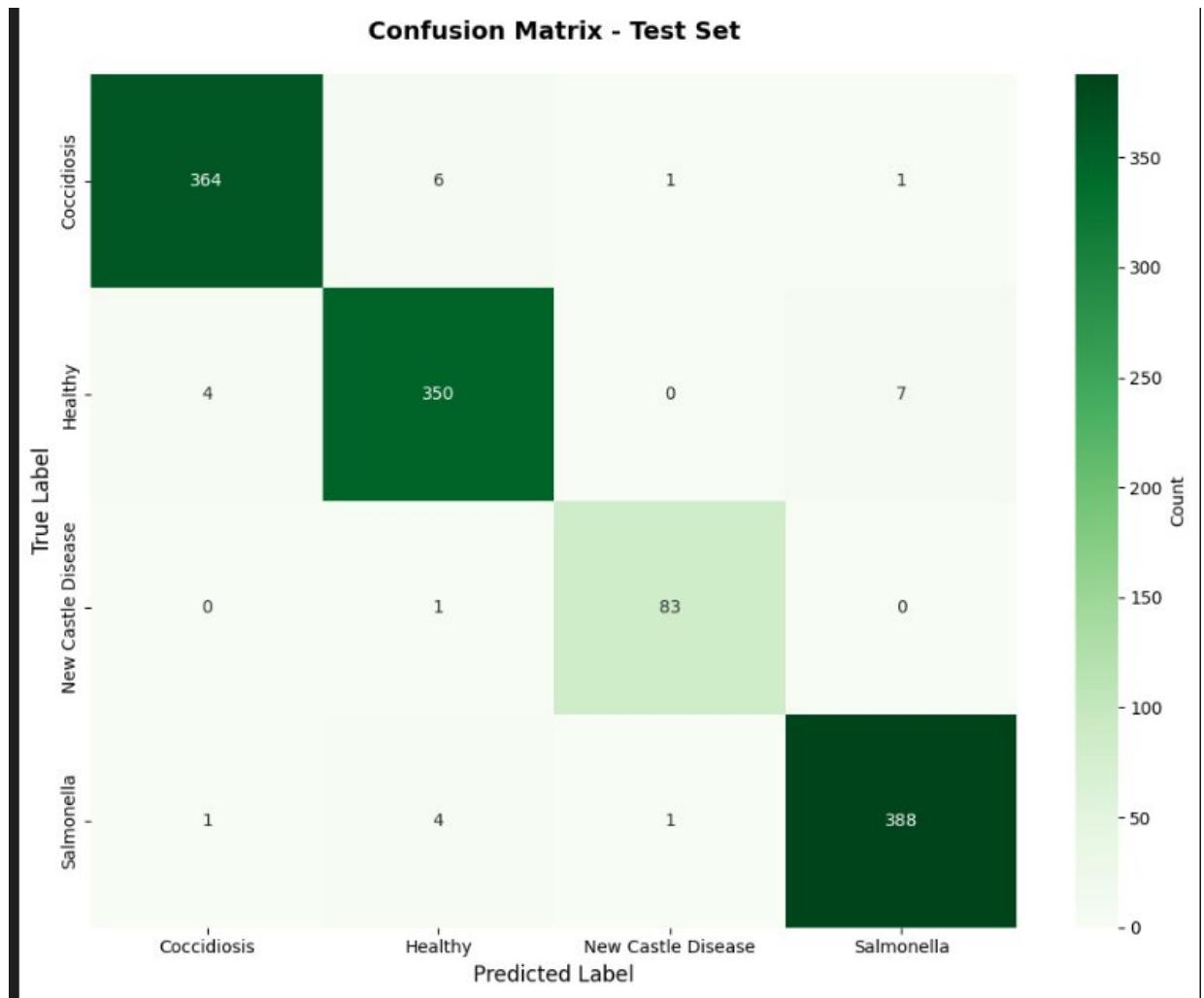


Hình 4 Ma trận nhầm lẫn (Confusion Matrix) của mô hình CNN from Scratch trên tập test

Ma trận nhầm lẫn của mô hình CNN xây dựng từ đầu cho thấy mô hình có khả năng phân loại tương đối tốt đối với các lớp có số lượng mẫu lớn như Coccidiosis, Salmonella và Healthy. Phần lớn các mẫu thuộc ba lớp này được dự đoán đúng, phản

ánh việc mô hình đã học được những đặc trưng cơ bản về màu sắc và hình dạng của phân gà. Tuy nhiên, khi quan sát kỹ các ô ngoài đường chéo chính, có thể nhận thấy tỷ lệ nhầm lẫn vẫn còn tương đối cao so với các mô hình sử dụng học chuyển giao.

4.2.2. Mô hình EfficientNet

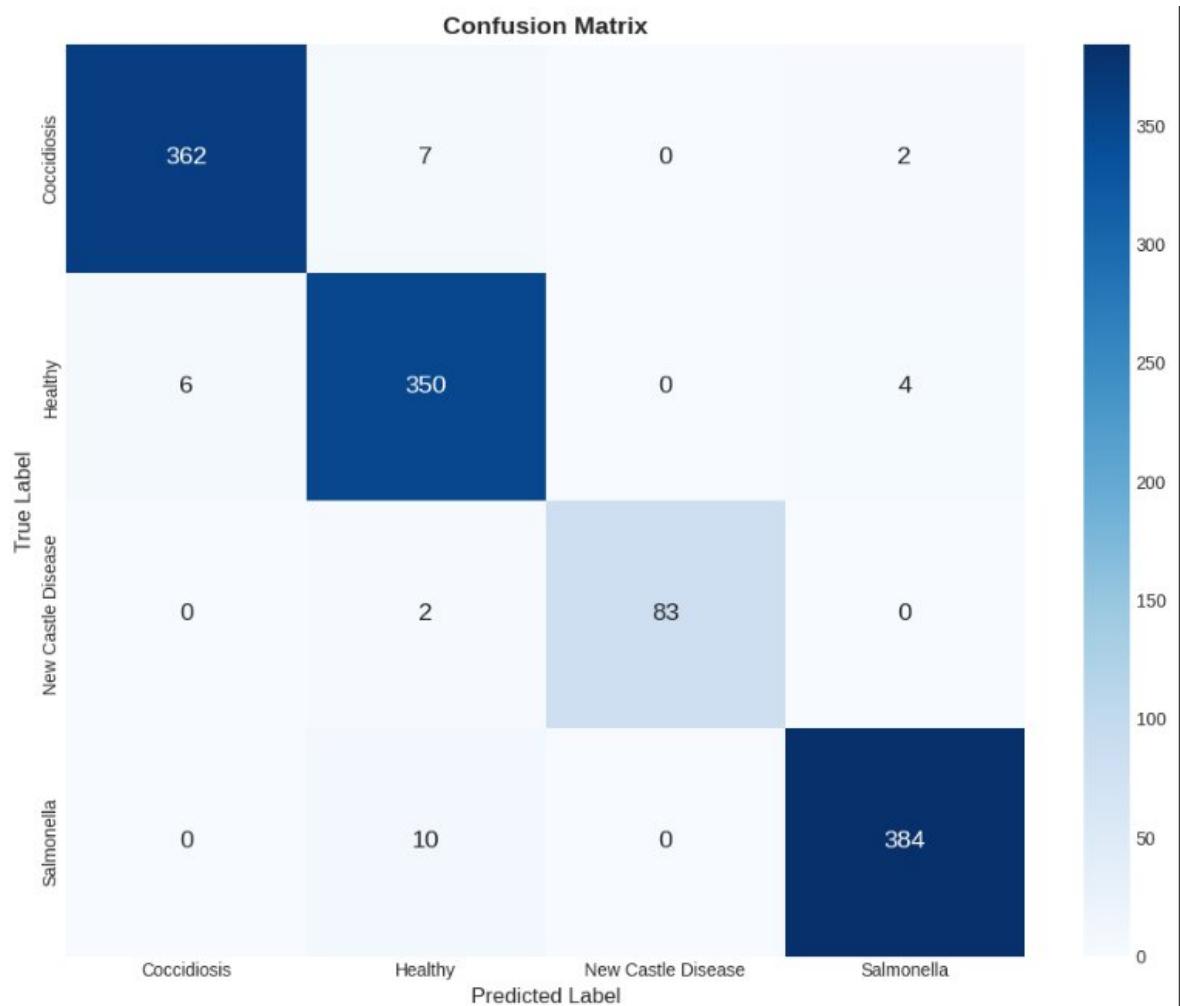


Hình 5 Ma trận nhầm lẫn (Confusion Matrix) của mô hình EfficientNet trên tập test

Ma trận nhầm lẫn (Confusion Matrix) của mô hình EfficientNetB3 trên tập test, được minh họa trong Hình 7, cho thấy phần lớn các mẫu thuộc bốn lớp đều được phân loại chính xác. Đặc biệt, mô hình thể hiện khả năng nhận diện tốt lớp Newcastle Disease, dù đây là lớp có số lượng mẫu ít nhất trong bộ dữ liệu.

Một số ít trường hợp nhầm lẫn chủ yếu xảy ra giữa lớp Healthy và các lớp bệnh ở giai đoạn nhẹ. Điều này là hợp lý về mặt sinh học, bởi trong thực tế, phân gà ở giai đoạn đầu của bệnh có thể chưa biểu hiện rõ ràng các dấu hiệu bất thường về màu sắc hoặc hình dạng. Hiện tượng này cho thấy giới hạn tự nhiên của dữ liệu hình ảnh, chứ không hoàn toàn xuất phát từ sai sót của mô hình.

4.2.3. Mô hình ResNet50

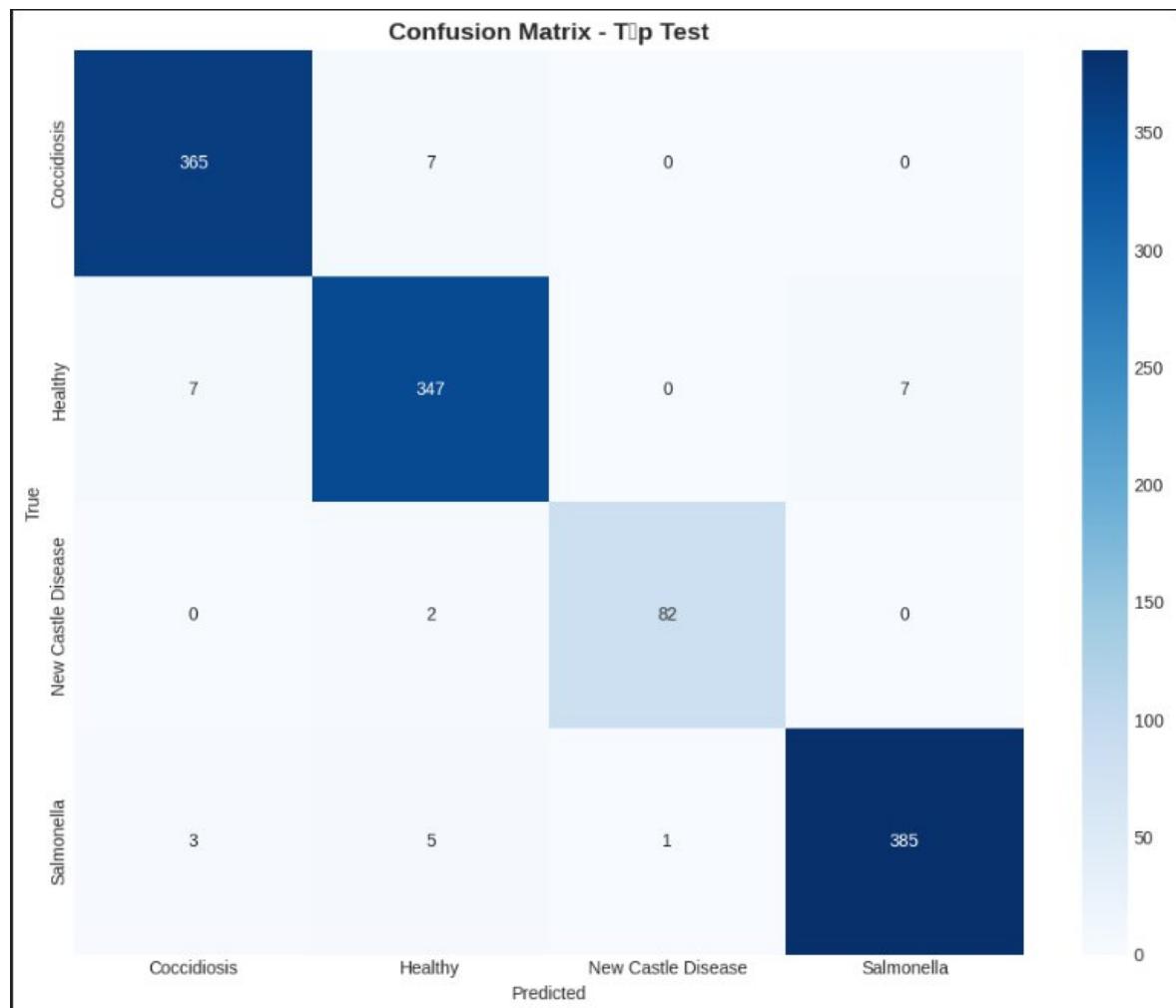


Hình 6 Ma trận nhầm lẫn (Confusion Matrix) của mô hình ResNet50 trên tập test

Ma trận nhầm lẫn của mô hình ResNet50 cho thấy sự cải thiện rõ rệt so với mô hình CNN from Scratch. Hầu hết các mẫu thuộc bốn lớp đều được phân loại chính xác, với số lượng dự đoán đúng trên đường chéo chính chiếm tỷ lệ lớn. Điều này chứng tỏ ResNet50 đã khai thác hiệu quả các đặc trưng thị giác tổng quát được học từ ImageNet và thích nghi tốt với dữ liệu phân gà thông qua quá trình fine-tuning.

Đối với các lớp Coccidiosis và Salmonella, ResNet50 đạt tỷ lệ phân loại đúng rất cao, với số trường hợp nhầm lẫn là không đáng kể. Lớp Healthy cũng được nhận diện ổn định, tuy vẫn tồn tại một số ít mẫu bị nhầm sang các lớp bệnh, chủ yếu trong các trường hợp ảnh có biểu hiện không rõ ràng hoặc điều kiện ánh sáng kém. Đây là hiện tượng thường gặp và mang tính thực tế, không phải là lỗi nghiêm trọng của mô hình.

4.2.4. Mô hình MobileNetV3



Hình 7 Ma trận nhầm lẫn (Confusion Matrix) của mô hình MobileNetV3 trên tập test

Ma trận nhầm lẫn của mô hình MobileNetV3 cho thấy một hiệu năng rất cân bằng giữa độ chính xác và tính ổn định, mặc dù đây là mô hình có kiến trúc nhẹ và được tối ưu cho các thiết bị hạn chế tài nguyên. Phần lớn các mẫu thuộc các lớp Coccidiosis,

Salmonella và Healthy đều được phân loại chính xác với tỷ lệ cao, thể hiện khả năng trích xuất đặc trưng hiệu quả của MobileNetV3 mặc dù số lượng tham số thấp hơn đáng kể so với ResNet50.

4.3. Đánh giá mô hình tốt nhất

Dựa trên các độ đo định lượng và phân tích định tính, EfficientNetB3 được xác định là mô hình cho hiệu năng tốt nhất trong nghiên cứu này. Ưu thế của EfficientNet đến từ chiến lược Compound Scaling, cho phép mở rộng đồng thời chiều sâu, chiều rộng và độ phân giải đầu vào một cách tối ưu. Nhờ đó, mô hình có khả năng trích xuất các đặc trưng tinh vi hơn từ hình ảnh phân gà, đặc biệt là các chi tiết nhỏ liên quan đến màu sắc và kết cấu.

Ngoài ra, việc kết hợp Label Smoothing, AdamW, Early Stopping và Heavy Data Augmentation đã giúp mô hình giảm hiện tượng overfitting, đồng thời cải thiện khả năng tổng quát hóa trên tập dữ liệu kiểm tra.

4.4. So sánh hiệu năng giữa các mô hình

Trong quá trình thực nghiệm, mô hình CNN tự xây dựng (CNN from Scratch) cho thấy hiệu năng thấp nhất trong số các kiến trúc được thử nghiệm. Do không được thừa hưởng các đặc trưng thị giác từ dữ liệu lớn (ImageNet) thông qua cơ chế Transfer Learning, mô hình này gặp khó khăn trong việc trích xuất các đặc trưng phức tạp của ảnh phân gà và dễ bị quá khớp (overfitting) khi đối mặt với nhiều hoặc sự đa dạng của dữ liệu. Mặc dù vậy, mô hình này đóng vai trò quan trọng làm baseline (mức cơ sở), giúp khẳng định tính hiệu quả và sự cần thiết của việc áp dụng các kỹ thuật học chuyển tiếp cho bài toán này.

Chuyển sang các mô hình tiền huấn luyện, MobileNetV3 nổi lên là ứng cử viên sáng giá cho các bài toán yêu cầu tốc độ phản hồi nhanh. Nhờ kiến trúc sử dụng các lớp tích chập tách biệt theo chiều sâu (depthwise separable convolutions), MobileNetV3 đạt được sự cân bằng ánh tượng giữa độ chính xác. Mặc dù độ chính xác có thể thấp hơn một chút so với các mô hình lớn, nhưng với thời gian suy luận cực nhanh và kích thước mô hình nhẹ, đây là lựa chọn tối ưu nhất để triển khai lên ứng dụng web demo hoặc các thiết bị di động có tài nguyên phần cứng hạn chế.

Trong khi đó, ResNet50 thể hiện sự vượt trội về độ ổn định và khả năng học sâu nhờ vào kiến trúc Residual Block (khối thặng dư). Cơ chế skip connection giúp mô hình giải quyết triệt để vấn đề biến mất đạo hàm (vanishing gradient), cho phép huấn luyện mạng nơ-ron rất sâu mà không bị suy giảm hiệu năng. Kết quả là ResNet50 đạt độ chính xác cao hơn MobileNetV3, tuy nhiên, sự đánh đổi nằm ở thời gian huấn luyện

và suy luận của ResNet50 lâu hơn đáng kể, đòi hỏi hạ tầng phần cứng mạnh mẽ hơn để vận hành trơn tru.

Cuối cùng, EfficientNet (cụ thể là phiên bản B3 được sử dụng) đã chứng minh vị thế vượt trội nhất trong thực nghiệm này. Nhờ phương pháp Compound Scaling giúp tối ưu hóa đồng thời cả chiều sâu, chiều rộng và độ phân giải của mạng, EfficientNet không chỉ đạt độ chính xác cao nhất mà còn thể hiện khả năng "đọc" dữ liệu xuất sắc. Mô hình này xử lý rất tốt các trường hợp ảnh khó như điều kiện ánh sáng yếu, độ tương phản không đồng đều hoặc khi các đặc trưng bệnh lý bị mờ nhạt - những tình huống mà các mô hình khác thường dự đoán sai. Do đó, EfficientNet được lựa chọn là mô hình tốt nhất để làm nòng cốt cho hệ thống chẩn đoán bệnh.

4.5. Hiện tượng quan sát được

Trong quá trình đánh giá mô hình trên tập kiểm thử (Test Set), nhóm nghiên cứu đã ghi nhận những hiện tượng đặc thù liên quan mật thiết đến sự phân bố dữ liệu đầu vào và kiến trúc mạng. Đầu tiên, hiện tượng mất cân bằng dữ liệu ở lớp "Newcastle Disease" đã ảnh hưởng rõ rệt đến Output của mô hình; cụ thể, Ma trận nhầm lẫn (Confusion Matrix) cho thấy tỷ lệ dự đoán sai (False Negative) của lớp này cao hơn các lớp khác, thường bị nhầm lẫn sang các bệnh có biểu hiện phân lồng tương tự như Salmonella. Điều này phản ánh rằng khi Input thiếu đa dạng mẫu, mô hình chưa học đủ các đặc trưng tinh tế để phân tách ranh giới quyết định (decision boundary) rõ ràng. Thứ hai, sự chênh lệch hiệu năng giữa CNN from Scratch và các mô hình Pretrained là minh chứng cho sức mạnh của việc khởi tạo trọng số. Với Input là các ảnh phân gà có độ nhiễu nền cao, mô hình tự xây dựng (Random weight initialization) gặp khó khăn trong việc hội tụ (Loss giảm chậm và dao động), trong khi các mô hình như ResNet50 hay EfficientNet tận dụng được các bộ lọc đặc trưng đã học từ ImageNet để trích xuất thông tin hữu ích ngay từ những epoch đầu tiên. Cuối cùng, một quan sát quan trọng về kỹ thuật là sự đánh đổi giữa kích thước Input và hiệu năng: EfficientNet-B3 yêu cầu Input đầu vào có độ phân giải cao hơn (thường là 300x300 so với 224x224 tiêu chuẩn), giúp mô hình "nhìn" rõ hơn các chi tiết nhỏ trong kết cấu phân, từ đó Output đạt độ chính xác cao hơn nhưng đồng thời cũng làm tăng độ tốn tính toán so với MobileNetV3.

4.6. Đánh giá khả năng ứng dụng thực tế

Để kiểm chứng khả năng hoạt động của mô hình trong môi trường thực tế, nhóm sẽ xây dựng một ứng dụng demo trên nền tảng Streamlit, mô phỏng quy trình chẩn đoán

đầu cuối (End-to-End). Quy trình xử lý Input của ứng dụng được thiết kế đồng bộ hoàn toàn với quy trình huấn luyện: ảnh người dùng tải lên sẽ được tự động Resize về kích thước tiêu chuẩn và chuẩn hóa (Normalize) theo các chỉ số trung bình và độ lệch chuẩn của tập ImageNet trước khi đưa vào mô hình. Tại đầu ra (Output), thay vì chỉ đưa ra nhãn dự đoán cứng (Hard label), hệ thống sử dụng hàm Softmax để trích xuất vector xác suất, từ đó hiển thị "Độ tin cậy" (Confidence Score) cho người dùng. Kết quả thử nghiệm thực tế cho thấy MobileNetV3 phản hồi gần như tức thời (Low Latency) nhờ kiến trúc nhẹ, rất phù hợp cho các thiết bị di động của người chăn nuôi tại chuồng trại nơi kết nối internet có thể không ổn định. Ngược lại, EfficientNet dù có thời gian suy luận lâu hơn đôi chút nhưng lại đóng vai trò quan trọng trong việc tham chiếu (Second opinion) cho các ca bệnh khó nhò độ tin cậy vượt trội. Sự kết hợp này chứng minh hệ thống hoàn toàn khả thi để phát triển thành một trợ lý ảo hỗ trợ bác sĩ thú y và người chăn nuôi trong việc sàng lọc bệnh sớm.

CHƯƠNG 5. ĐỊNH HƯỚNG PHÁT TRIỂN

Trong các nghiên cứu và triển khai tiếp theo, nhóm dự định mở rộng đề tài theo các hướng sau:

- **Triển khai hệ thống dưới dạng ứng dụng web hoàn chỉnh**, cho phép người dùng tải ảnh phân gà trực tiếp thông qua trình duyệt và nhận kết quả dự đoán theo thời gian thực. Hệ thống có thể được mở rộng theo mô hình client–server, trong đó mô hình học sâu được triển khai phía server để phục vụ nhiều người dùng đồng thời.
- **Thử nghiệm và so sánh thêm các kiến trúc học sâu khác**, đặc biệt là các mô hình hiện đại trong lĩnh vực thị giác máy tính như Vision Transformer (ViT), ConvNeXt hoặc các biến thể EfficientNet mới, nhằm đánh giá khả năng cải thiện hiệu năng và độ ổn định của mô hình.
- **Nâng cao khả năng giải thích mô hình**, thông qua việc áp dụng các kỹ thuật trực quan hóa như Grad-CAM để phân tích các vùng ảnh mà mô hình tập trung khi đưa ra dự đoán, từ đó tăng mức độ tin cậy và khả năng ứng dụng trong thực tế.

- **Tối ưu hiệu năng triển khai**, hướng tới việc rút gọn mô hình (model compression) hoặc sử dụng các mô hình nhẹ để có thể triển khai trên thiết bị có tài nguyên hạn chế, phù hợp với điều kiện ứng dụng trong nông nghiệp thông minh.

KẾT LUẬN

Đề tài "Xây dựng mô hình học sâu cho bài toán phân loại bệnh trên gà từ ảnh phân gà" đã hoàn thành các mục tiêu nghiên cứu và triển khai đề ra ban đầu, bao gồm việc xây dựng bộ dữ liệu, huấn luyện các mô hình Deep Learning và phát triển ứng dụng demo minh họa. Qua quá trình thực nghiệm trên bốn kiến trúc mạng nơ-ron tích chập (CNN from Scratch, MobileNetV3, ResNet50, và EfficientNet-B3), nhóm nghiên cứu đã rút ra được những kết luận quan trọng về tính hiệu quả của kỹ thuật học chuyển tiếp (Transfer Learning) trong bài toán chẩn đoán hình ảnh y sinh. Kết quả cho thấy, việc tận dụng các trọng số đã được huấn luyện trên tập ImageNet giúp mô hình hội tụ nhanh hơn và đạt độ chính xác vượt trội so với mô hình tự xây dựng, đặc biệt trong bối cảnh dữ liệu thu thập còn hạn chế và nhiễu nền lớn.

Về mặt hiệu năng, EfficientNet-B3 đã khẳng định vị thế là mô hình tối ưu nhất về độ chính xác, thể hiện khả năng trích xuất đặc trưng mạnh mẽ đối với các lớp bệnh có biểu hiện phức tạp như Coccidiosis (Càu trùng) hay Salmonella. Ngược lại, MobileNetV3 tuy có độ chính xác thấp hơn đôi chút nhưng lại chứng minh được ưu thế vượt trội về tốc độ suy luận (Inference time) và kích thước mô hình, mở ra tiềm năng to lớn cho việc triển khai trên các thiết bị di động hoặc hệ thống nhúng tại các trang trại. Bên cạnh đó, nhóm sẽ xây công ứng web demo trên nền tảng Streamlit, cho phép thực hiện quy trình chẩn đoán đầu cuối (End-to-End) từ khâu tải ảnh, xử lý tiền kỳ đến hiển thị kết quả và độ tin cậy, chứng minh tính khả thi của việc đưa mô hình vào ứng dụng thực tế.

Tuy nhiên, nghiên cứu vẫn còn tồn tại một số hạn chế nhất định, điển hình là vấn đề mất cân bằng dữ liệu dẫn đến khả năng dự đoán chưa đồng đều ở lớp bệnh Newcastle, cũng như sự phụ thuộc vào chất lượng ảnh đầu vào. Những hạn chế này đã được nhóm nhận diện và đề xuất các hướng giải quyết cụ thể trong phần định hướng phát triển, bao gồm việc mở rộng tập dữ liệu, áp dụng kỹ thuật học tập kết hợp (Ensemble Learning) và triển khai mô hình lên các nền tảng. Tổng kết lại, kết quả của đề tài không chỉ đóng góp một giải pháp công nghệ cụ thể cho ngành chăn nuôi gia cầm mà

còn tạo tiền đề vững chắc cho việc phát triển các hệ thống nông nghiệp thông minh, hỗ trợ người nông dân giám sát và quản lý dịch bệnh hiệu quả hơn trong tương lai.

Tài Liệu Tham Khảo

TensorFlow, "TensorFlow Core Documentation," [Online]. Available: https://www.tensorflow.org/api_docs.

PyTorch, "PyTorch Documentation," [Online]. Available: <https://pytorch.org/docs/stable/index.html>.

Kaggle, "Chicken Disease Dataset," [Online]. Available: <https://www.kaggle.com/datasets/allandclive/chicken-disease-1>.

Bảng phân chia công việc

Tên	Công việc	Tỉ lệ hoàn thành
Nguyễn Văn Tân Phát	Eda, xây dựng mô hình cnn from scratch, mobilenetv3	100%
Phạm Thành Doanh	Xây dựng mô hình ResNet50	100%
Trần Thiên Bảo	Xây dựng mô hình mobilenetv3, tìm hiểu công nghệ trong bài	Chưa hoàn thành
Nguyễn Hoàng Lộc	Xây dựng mô hình EfficientNetB3	100%