

DOI:10.22144/ctujos.2024.237

NHẬN DẠNG MỘT SỐ BỆNH TRÊN BÔNG XOÀI SỬ DỤNG KỸ THUẬT HỌC SÂU CHUYÊN TIẾP

Phan Tấn Phước¹, Ngô Quang Hiếu² và Trương Quốc Bảo^{2*}

¹Khoa Điện, Trường Cao đẳng nghề An Giang

²Trường Bách Khoa, Trường Đại học Cần Thơ

*Tác giả liên hệ (Corresponding author): tqbao@ctu.edu.vn

Thông tin chung (Article Information)

Nhận bài (Received): 21/08/2023

Sửa bài (Revised): 13/09/2023

Duyệt đăng (Accepted): 08/10/2023

Title: Identification of some diseases on mango flowers using deep transfer learning

Author(s): Phan Tan Phuoc¹, Ngo Quang Hieu² and Truong Quoc Bao^{2*}

Affiliation(s): ¹An Giang Vocational College, ²Can Tho University

TÓM TẮT

Nước ta có điều kiện khí hậu phù hợp với việc trồng cây xoài. Đây là loại cây ăn trái mang lại giá trị kinh tế cho người nông dân. Do đó, nhận dạng được bệnh hại trên bông trong giai đoạn ra hoa, đậu trái là rất quan trọng, ảnh hưởng lớn đến năng suất mùa vụ. Vì vậy, ứng dụng điện thoại thông minh để nhận dạng bệnh trên bông xoài thông qua hình ảnh của bông xoài là rất cần thiết. Nghiên cứu sử dụng tập dữ liệu của 4 loại bông xoài: bông xoài thân thư, bông xoài bị cháy (bệnh đốm đen), bông xoài nghẹn, bông xoài không bệnh, với 733 hình ảnh được thu thập. Nghiên cứu đã sử dụng kỹ thuật học sâu chuyên tiếp để tiến hành huấn luyện trên mô hình EfficientNetB0 và MobileNetV2, kết quả độ chính xác của mô hình EfficientNetB0 đạt 99,84% cao hơn so với mô hình MobileNetV2 chỉ đạt 95,21%. Kết quả nghiên cứu có thể giúp các nhà nông học, nhà nghiên cứu có các biện pháp hỗ trợ người nông dân định hướng xử lý bông xoài mang bệnh hại phù hợp, hiệu quả.

Từ khóa: Bệnh trên bông xoài, điện thoại thông minh, Grad-CAM, học chuyên tiếp, mạng EfficientNet

ABSTRACT

Our country has suitable climatic conditions for growing mango trees. This is a fruit tree that brings economic value to farmers. Therefore, identifying diseases in the flower during the flowering and fruiting stages is very important, greatly affecting to productivity of the season. Therefore, a program on smartphones to identify diseases in mango flowers through images of mango flowers is very necessary. The research used a set of data of 4 types of mango flowers: anthracnose mango flowers, burnt mango flowers (black spot disease), choked mango flowers, and disease-free mango flowers, with 733 images collected. The research already used deep transfer learning techniques to jog the training on model EfficientNetB0 and model MobileNetV2. As a result, the accuracy of the EfficientNetB0 model reached 99.84%, higher than the MobileNetV2 model, which only reached 95.21%. The results of the research can help agronomists and researchers have solutions to support farmers and guide the treatment of disease-carrying mango flowers appropriately and effectively.

Keywords: Android smartphone, EfficientNet network, Grad-CAM, mango flower disease, transfer learning

1. GIỚI THIỆU

Hiện nay, ở Việt Nam, nông nghiệp vẫn là một trong những ngành sản xuất chủ yếu. Quá trình sản xuất thu hoạch, phân loại và đánh giá chất lượng các loại sản phẩm nông nghiệp, đặc biệt là nhận dạng và phân loại bệnh trên hoa, quả chủ yếu còn phải thực hiện bằng các phương pháp thủ công. Đây là công việc không quá khó, nhưng tiêu tốn nhiều thời gian, công sức của con người, chỉ phù hợp với các vườn nhỏ lẻ và là rào cản đối với mở rộng phát triển quy mô sản xuất nông nghiệp nói chung và đặc biệt là sản xuất nông sản nói riêng (Hầu, 2009) như với các cánh đồng mẫu lớn hay vườn cây ăn trái có diện tích tính bằng hecta. Vì vậy, việc nhận dạng được một số loại bệnh hại trên bông kịp thời trong giai đoạn ra hoa, đậu trái nhờ đó có những phương pháp xử lý, phòng ngừa bệnh hại trên bông là hết sức cần thiết nhằm đạt được sản phẩm trái đẹp, đạt yêu cầu đem lại hiệu quả kinh tế cao (Báo điện tử Đảng Cộng Sản Việt Nam, 2011).

Kỹ thuật học sâu và máy học đã đóng một vai trò quan trọng trong lĩnh vực nông nghiệp để dự đoán, phân loại và phát hiện bệnh cây trồng, đồng thời cung cấp các phương tiện phát hiện bệnh cây trồng với chi phí thấp, nhanh chóng và đáng tin cậy. Các nhà nghiên cứu khác nhau đã nghiên cứu chẩn đoán và phát hiện bệnh thực vật, đặc biệt là bệnh xoài. Singh et al. (2019) đề xuất một phương pháp học sâu để xác định lá xoài bị nhiễm bệnh thán thư. Tác giả đã sử dụng mạng nơ-ron tích chập nhiều lớp (Multi Convolutional Neural Network - MCNN) để huấn luyện mô hình phân loại bằng cách sử dụng bộ dữ liệu gồm 1.070 hình ảnh về lá xoài khỏe mạnh và bị nhiễm bệnh thán thư. Trên tập thử nghiệm, độ chính xác 96,89% của MCNN cao hơn đáng kể so với các kỹ thuật trước đó. Ashok and Vinod (2021) cung cấp một cách tiếp cận độc đáo để phát hiện bệnh trên xoài bằng cách sử dụng mô hình học sâu và ứng dụng Android. Một mô hình mạng lưới thần kinh tích chập (Convolutional Neural Network - CNN) được sử dụng trong quy trình này và được huấn luyện bằng cách sử dụng tập dữ liệu hình ảnh những quả xoài có và không có bệnh. Ứng dụng Android sau đó sử dụng mô hình CNN để xác định bệnh tật trong ảnh quả xoài. Một tập dữ liệu gồm 1.000 hình ảnh về quả xoài, trong đó có 500 hình ảnh bị bệnh và 500 hình ảnh không bị bệnh. Độ chính xác của

phương pháp này đạt được là 95% trong việc xác định bệnh trên quả xoài.

Nghiên cứu nhằm đề xuất một giải pháp ứng dụng kỹ thuật học sâu hiện đại để nhận dạng một số bệnh thông thường dựa trên hình ảnh thu thập được trong khi chăm sóc cây xoài đang ra hoa sử dụng mạng nơ-ron tích chập hiện đại EfficientNetB0 đồng thời nghiên cứu cũng tiến hành khảo sát thêm mạng nơ-ron tích chập cải tiến MobileNetV2 để so sánh kết quả đạt được trên một bộ dữ liệu bông xoài một loại cây trồng mang lại nhiều giá trị kinh tế cao ở Việt Nam nói chung và ở tỉnh An Giang nói riêng.

Hình ảnh nhận dạng được thu thập trực tiếp bằng máy ảnh điện thoại hoặc thu thập trước bằng máy ảnh thông dụng. Đối với những sản phẩm cây ăn trái như xoài, chất lượng và độ đẹp hay xấu của sản phẩm đều bị ảnh hưởng trong giai đoạn từ khi ra hoa đến khi đậu trái non. Nghiên cứu này trình bày kỹ thuật học chuyển tiếp (Transfer Learning) sử dụng mô hình mạng EfficientNet-B0 trong bài toán nhận dạng bệnh trên bông xoài, dữ liệu đầu vào là hình ảnh của ba loại bệnh trên bông xoài và một loại bông không bệnh đã được gán nhãn phân loại dữ liệu (data frame). Môi trường huấn luyện mô hình là Google Colab, thử nghiệm mô hình bằng điện thoại thông minh hệ điều hành Android. Kết quả dự đoán đạt được một cách hiệu quả trên tập dữ liệu kiểm tra và nhận dạng với độ chính xác cao bằng điện thoại thông minh hệ điều hành Android.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Mô hình học sâu EfficientNet

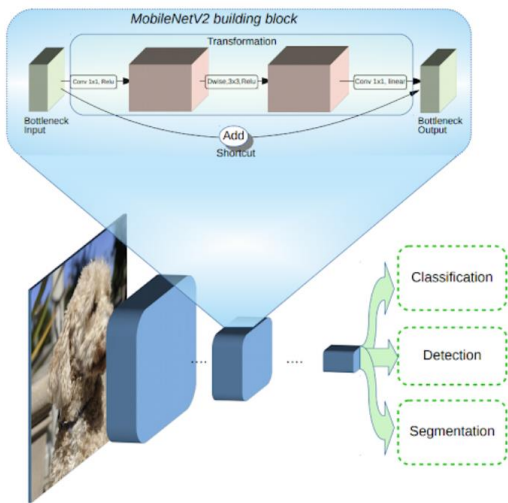
EfficientNet được giới thiệu lần đầu tiên về 2 kỹ sư Google là Mingxing Tan và Quoc V. Le vào tháng 5, năm 2019 (Tan & Le, 2019). Kiến trúc mạng EfficientNet là một mạng CNN đạt state-of-the-art về độ chính xác với số lượng tham số và FLOPS ít hơn nhưng đạt hiệu quả và nhanh hơn các kiến trúc mạng khác. EfficientNets được dựa trên máy học tự động (AutoML) và kết hợp chia tỷ lệ (scaling) để đạt được hiệu suất cao mà không ảnh hưởng đến hiệu quả tài nguyên (Tan & Le, 2019). Các tác giả đã thiết kế một mạng cơ sở có kích thước di động được gọi là EfficientNet-B0, hoạt động bằng cách sử dụng kiến trúc thần kinh đa mục tiêu giúp tối ưu hóa độ chính xác và FLOPS. Mô hình được lấy cảm hứng từ Mnas-Net và có kiến trúc như sau (Hình 1).



Hình 1. Kiến trúc EfficientNet-B0

Hình 1 cho thấy mô hình EfficientNet-B0 cơ sở nhận hình ảnh màu có kích thước [224x224]. Mạng cơ sở EfficientNet-B0, các tác giả đã áp dụng phương pháp chia tỷ lệ hỗn hợp. Đặt $\Phi=1$ và thực hiện tìm kiếm dạng lưới để tìm các tham số α và γ dựa trên các phương trình được đề xuất và theo ràng buộc $\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$. Kết quả như sau $\alpha=1,2$, $\beta=1,1$ và $\gamma=1,15$ [10].

Kể từ khi ra đời, MobileNetV2 là một trong những kiến trúc được ưa chuộng nhất khi phát triển các ứng dụng trí tuệ nhân tạo, MobileNetV2 có một số điểm cải tiến so với MobileNetV1 giúp cho nó có độ chính xác cao hơn, số lượng tham số và số lượng các phép tính ít hơn. Tuy nhiên, kết nối tắt ở MobileNetV2 được điều chỉnh sao cho số kênh (hoặc chiều sâu) ở input và output của mỗi khối dư (block residual) được thắt hẹp lại. Chính vì thế, nó được gọi là các bottleneck layers (bottleneck là một thuật ngữ thường được sử dụng trong deep learning để ám chỉ các kiến trúc thu hẹp kích thước theo một chiều nào đó). Hình 2 cho thấy kiến trúc cải tiến của mạng MobileNetV2 (Mark et al., 2019).



Hình 2. Kiến trúc MobileNetV2

2.2. Thu thập và tiền xử lý tập dữ liệu cho huấn luyện mô hình

Tập dữ liệu bông xoài được chụp thủ công bằng máy ảnh của nhiều hãng điện thoại khác nhau (Hình 3) như: Samsung J7 Prim, Samsung A22 5G, Samsung Note 5, Asus, Iphone 13.



Hình 3. Hình ảnh quá trình thu thập dữ liệu

Bông xoài được chụp trực tiếp trên cây xoài tại các vườn xoài đang ra hoa, bông xoài cần chụp nằm ở trung tâm ảnh chụp, khoảng cách từ máy ảnh đến bông xoài khoảng 30-50 cm đảm bảo ảnh không bị mờ, chụp theo các giai đoạn sinh trưởng của bông (từ nhú cựa gà đến đậu trái non). Kích thước hình ảnh bao gồm [3264x2248] và [3264x1836] pixels. Dữ liệu sau khi thu thập sẽ tiếp tục xử lý gán nhãn dữ liệu bằng cách đưa hình mỗi loại bông xoài vào một thư mục đại diện cho một loại bông xoài cần nhận dạng (Hình 4). Dữ liệu hình ảnh thu thập được thường bị tác động bởi nhiễu. Vì vậy, khâu tiền xử lý dữ liệu sẽ được áp dụng trước khi đưa vào huấn luyện mô hình (được thực hiện ở mục 2.3)

Chi tiết số lượng ảnh của từng loại sau khi tiền xử lý được thể hiện trong Bảng 1. Số lượng hình ảnh trong tập dữ liệu tương đối cân bằng giữa các loại nhãn dữ liệu (Tên bông xoài cần nhận dạng), tạo điều kiện thuận lợi cho việc huấn luyện mô hình.



Bông không bệnh Bông cháy Bông nghẹn Bông thán thư

Hình 4. Các loại bệnh trên bông xoài cần nhận dạng

Bảng 1. Chi tiết phân chia dữ liệu hình ảnh của các loại bông xoài

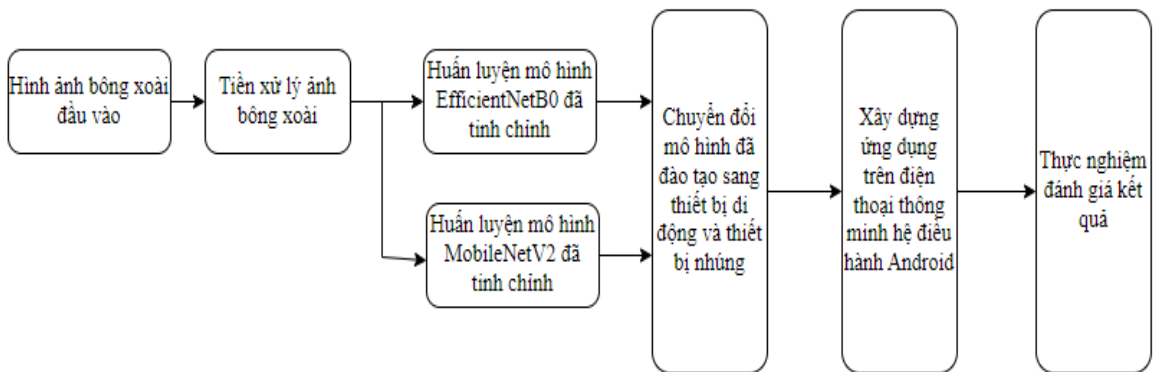
Tên bông xoài cần nhận dạng	Tập dữ liệu huấn luyện	Tập dữ liệu kiểm định	Tập dữ liệu kiểm tra
Bông không bệnh (bông bình thường)	1134	140	126
Bông cháy	1134	140	126
Bông nghẹn	1134	140	126
Bông thán thư	1134	140	126

2.3. Huấn luyện mô hình

Như đã đề cập, mạng nơ-ron học sâu được sử dụng trong thực nghiệm này là mạng EfficientNet-B0. Ưu điểm của mạng này là đã được huấn luyện trước trên tập dữ liệu ImageNet (10.000 ảnh thông

dụng); do đó, khi tiếp tục huấn luyện với tập dữ liệu mới, mạng dễ dàng hội tụ (Chollet et al., 2017).

Mô hình EfficientNet-B0 cần được huấn luyện lại và huấn luyện thêm với tập dữ liệu về bông xoài đã được thu thập. Hình ảnh trước khi được vào mô hình cần phải qua bước tiền xử lý. Bước tiền xử lý đầu tiên là chuyển đổi kích thước hình ảnh (Resize) thuộc cả 3 tập dữ liệu đã phân chia về kích thước chuẩn của mô hình là [224x224]. Đối với tập dữ liệu huấn luyện, ngoài áp dụng thay đổi kích thước hình ảnh, tập dữ liệu này còn được áp dụng thêm các phương pháp khác để tăng độ đa dạng cho hình ảnh huấn luyện, giúp mô hình được học tốt hơn, giúp gia tăng độ chính xác bao gồm: horizontal_flip, rotation_range, width_shift_range, height_shift_range, zoom_range. Hình ảnh sau khi tiền xử lý được đưa vào mô hình để huấn luyện, chuyển đổi mô hình và xây dựng ứng dụng theo các bước như Hình 5.



Hình 5. Các bước chính để nhận dạng bệnh trên bông xoài

Nghiên cứu sử dụng Google Colab để làm môi trường huấn luyện và kiểm tra đánh giá mô hình với ngôn ngữ lập trình Python và hỗ trợ sử dụng các nền tảng học sâu phổ biến như Tensorflow, Keras. Ngoài ra, việc huấn luyện với sự hỗ trợ của GPU (Graphics

Processing Unit) giúp quá trình này được thực hiện nhanh hơn nhờ khả năng tính toán song song.

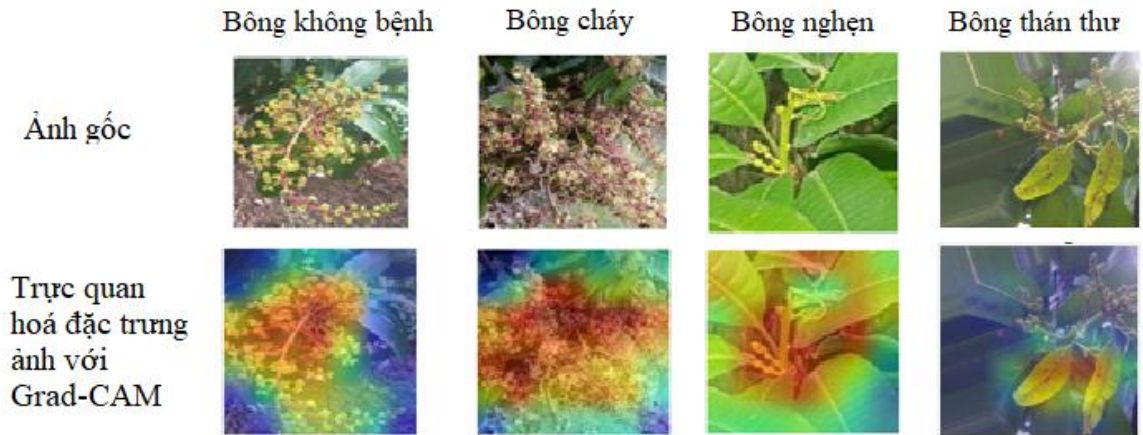
Một điểm quan trọng là nghiên cứu sử dụng mô hình EfficientNet-B0 đã được huấn luyện sẵn với

tập dữ liệu ImageNet có khả năng phân loại cho 1000 đối tượng khác nhau. Nghiên cứu này tiến hành nhận dạng cho đối tượng bông xoài với số lớp cần nhận dạng là 4 lớp. Chi tiết mô hình EfficientNet-B0 sau khi đã thay đổi, tinh chỉnh lớp phân loại cuối Dense_output (num_class=4) được xây dựng trên nền tảng Tensorflow. Thông số cơ bản được cài đặt để huấn luyện mạng gồm: batch_size=32, epochs=100, patience=1, stop_patience=3, threshold=.9, factor=.5, dwell=True, freeze=False, ask_epoch=10.

2.4. Trục quan hóa đặc trưng ảnh với Grad-CAM

Mặc dù phương pháp học sâu đã chứng tỏ độ chính xác cao trong phân loại ảnh, nhưng một trong những vấn đề của phương pháp này là khả năng diễn giải mô hình, thành phần cốt lõi để thông hiểu và gỡ lỗi. Mô hình học sâu thường được xem là “hộp đen” và vì vậy chúng ta không có thông tin về: 1) Mạng đang “nhìn” ở đâu trong ảnh đầu vào; 2) Các nơ-ron nào được kích hoạt để chuyển tiếp trong quá trình suy luận hoặc dự đoán; 3) Làm thế nào mạng đưa ra quyết định cuối cùng. Do đó, để đảm bảo mô hình mạng hoạt động đúng đắn, chính xác cần có công cụ

để trực quan hóa và kiểm tra xem mô hình mạng đang “nhìn” và “xác định” các vị trí nào trong ảnh đầu vào. Vì vậy, Zhou et al. (2016) đã đề xuất một kỹ thuật được gọi là Grad-CAM (Gradient-Class Activation Map), cho phép các phương pháp rút trích đặc trưng ảnh tự động bằng cách sử dụng mạng nơ-ron tích chập học sâu (Depth Convolutional Neural Network - DCNN) thực hiện định vị đối tượng mà không cần sử dụng bất kỳ một khung giới hạn bao quanh đối tượng (bounding box). Grad-CAM (LeCun et al., 1998) sử dụng tín hiệu của gradient để kết hợp các bản đồ đặc trưng (feature maps) mà không cần phải thay đổi kiến trúc của mạng, chúng ta có thể kiểm tra một cách trực quan vị trí mạng đang “nhìn”, xác định chính xác đối tượng trong hình ảnh và các kích hoạt xung quanh nó (Bazzani et al., 2016). Nghiên cứu này sẽ trình bày các hình ảnh trực quan Grad-CAM của EfficientNet-B0 với mô hình tinh chỉnh của nó trên các ảnh bông xoài. Hình 6 cho thấy EfficientNet-B0 có thể “nhìn” đúng vào khu vực của đối tượng cần nhận dạng thông qua trích xuất đặc trưng mà mô hình đã học. Các hình ảnh mẫu này được chọn ngẫu nhiên từ tập dữ liệu kiểm thử.



Hình 6. Kết quả của Grad-CAM trực quan vùng cần trích rút đặc trưng trên bông xoài

2.5. Chuyển đổi mô hình sang TFLite và thiết kế ứng dụng Android

2.5.1. Chuyển đổi mô hình sang TFLite

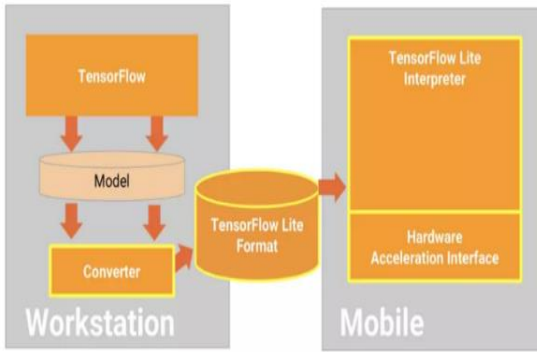
Vì đặc thù của các thiết bị như IoT (Internet of Things), mobile là cấu hình phần cứng hạn chế nên không thể đưa hẳn một mô hình trí tuệ nhân tạo (Artificial Intelligence model) lên thiết bị di động được mà cần phải xử lý để giảm kích thước của mô hình xuống, giảm được độ trễ xử lý, tiết kiệm không gian bộ nhớ. TensorFlow Lite được thiết kế để chạy các mô hình trên thiết bị di động và thiết bị nhúng (Hình 7), TensorFlow Lite là phiên bản đơn giản, ít

tốn tài nguyên bộ nhớ hơn của TensorFlow cũng do Google phát triển.

Mô hình sau khi chuyển đổi để triển khai trên thiết bị di động được trình bày ở Bảng 2.

Bảng 2. Kích thước mô hình trước và sau khi chuyển đổi

Model Tensorflow	Kích thước (Mb)
flowers_model.h5	51MB
flowers_model.tflite	3.2MB
Flowers_labels.txt	76KB



Hình 7. Mô hình triển khai AI model trên thiết bị di động và thiết bị nhúng

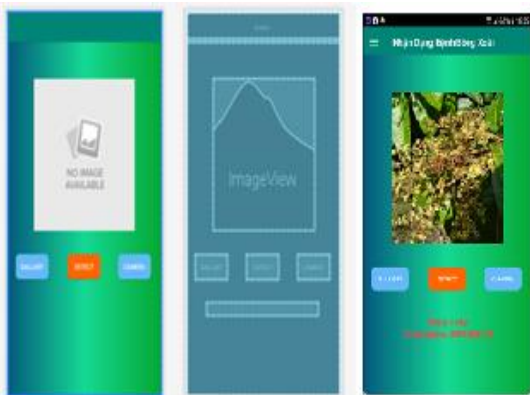
2.5.2. Thiết kế xây dựng ứng dụng Android

a. Các phương pháp triển khai mô hình trí tuệ nhân tạo (AI model)

Việc đưa ứng dụng đến với người sử dụng thực tế là một thành công lớn trong lĩnh vực máy học (Machine Learning). Để có thể sử dụng mô hình trí tuệ nhân tạo (AI model) trên điện thoại thông minh thì có hai cách có thể sử dụng, một là có thể dựng Server, viết một phương thức trung gian kết nối các ứng dụng và thư viện khác nhau (Application Programming Interface - API) cho ứng dụng gọi đến server, truyền hình ảnh vào input. Server xử lý xong trả kết quả. Cách hai là chuyển đổi mô hình sang dạng TFLite để đưa trực tiếp lên trên thiết bị xử lý, cũng tùy thuộc vào bài toán mà áp dụng làm theo cách thích hợp. Nghiên cứu này sẽ xây dựng ứng dụng với cách thứ hai, là đưa trực tiếp mô hình đã chuyển đổi lên xử lý trực tiếp trên điện thoại thông minh Android.

b. Cài đặt và thiết kế giao diện phần mềm

Về ý tưởng của ứng dụng thì nghiên cứu sẽ thiết kế giao diện người dùng được trình bày như Bảng 3 và Hình 8.



Hình 8. Màn hình giao diện người dùng

Bảng 3. Danh mục các thiết kế có trong giao diện

Tên	Số lượng	Mô tả chức năng
Vùng hiển thị ảnh	1	Vùng hiển thị hình ảnh do người dùng chọn.
Vùng hiển thị chữ	1	Vùng hiển thị kết quả nhận dạng trên màn hình ứng dụng.
Nút nhấn (GALLERY)	1	Dùng để chọn ảnh đã lưu trong thư viện trên điện thoại.
Nút nhấn (CAMERA)	1	Dùng để chụp ảnh sử dụng camera điện thoại chụp ảnh.
Nút nhấn (DETECT)	1	Dùng để nhận dạng ảnh đã chọn.

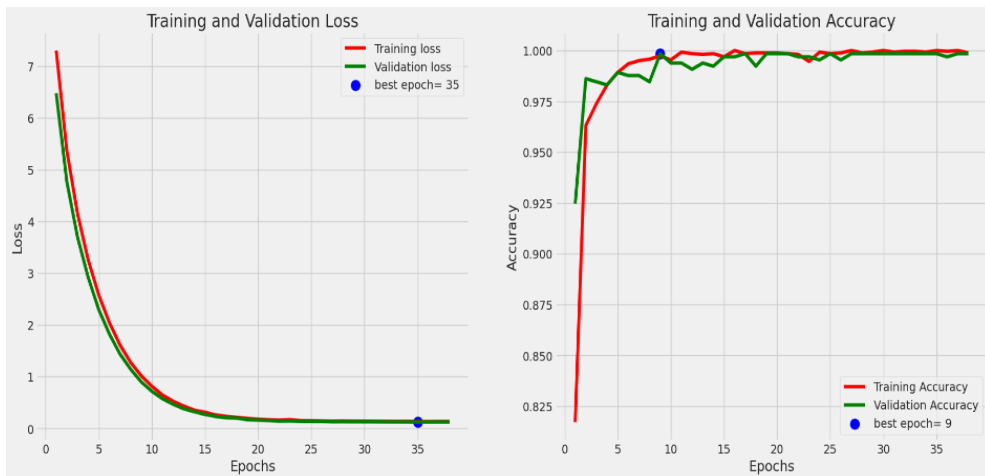
3. KẾT QUẢ VÀ THẢO LUẬN

Sau quá trình huấn luyện trên Google Colab với tập dữ liệu huấn luyện là 4.536 hình ảnh, tập kiểm thử là 560 hình ảnh, tập kiểm tra là 504 hình ảnh của bốn loại bông xoài trong Bảng 1 được học chuyên tiếp với mô hình EfficientNet-B0 và MobileNetV2 mà nghiên cứu đã thực hiện. Kết quả từ đường cong học tập cho thấy mô hình EfficientNet-B0 (Hình 9) có độ hội tụ tốt hơn mô hình MobileNetV2 (Hình 10) và độ chính xác trên tập kiểm tra (Test Accuracy) đạt 99,84% cũng cao hơn so với mô hình MobileNetV2 chỉ 95,21%.

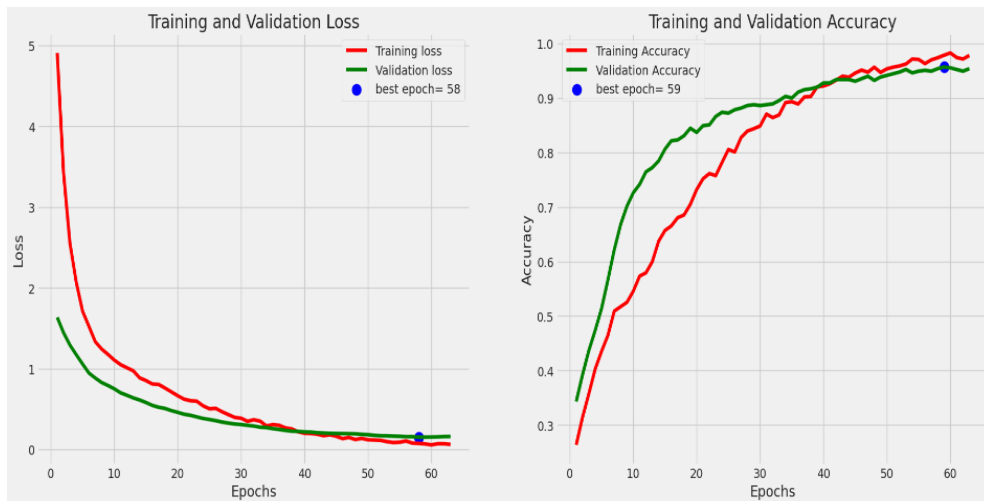
Từ kết quả kiểm tra độ chính xác (Test Accuracy) chỉ cho biết được bao nhiêu phần trăm lượng dữ liệu được phân loại đúng mà không chỉ ra được cụ thể mỗi loại được phân loại như thế nào, lớp nào được phân loại đúng nhiều nhất và dữ liệu thuộc lớp nào thường bị phân loại nhầm vào lớp khác. Để có thể đánh giá được các giá trị này, chúng ta sử dụng một ma trận được gọi là ma trận hỗn loạn (confusion matrix).

Đây là một trong những thước đo đánh giá hiệu suất của mô hình học máy dựa trên phân loại. Tập dữ liệu thử nghiệm sẽ được sử dụng để đánh giá hiệu suất của mô hình. Một trong những chỉ số được kiểm tra là độ chính xác đo lường phần dự đoán mà mô hình đã đúng. Các chỉ số khác như sau (Bảng 4):

Độ chính xác P (Precision): Tỷ lệ kết quả dương tính thực sự TP (True Positives- dự đoán chính xác) từ tổng số kết quả có liên quan, tức là tổng TP và kết quả dương tính giả FP (False Positives). Đối với các vấn đề phân loại nhiều lớp, P được tính trung bình giữa các lớp. Sau đây là công thức cho độ chính xác (1).



Hình 9. Đường cong học tập và dự đoán trên tập test_images mô hình EfficientNet-B0



Hình 10. Đường cong học tập và dự đoán trên tập test_images mô hình MobileNetV2

Ma trận hỗn loạn

Nhãn thật	Bông không bệnh	126	0	0	0
	Bông cháy	0	126	0	0
	Bông nghẹn	0	0	126	0
	Bông thán thư	0	0	3	123
		Bông không bệnh	Bông cháy	Bông nghẹn	Bông thán thư

Nhãn dự đoán

Hình 11. Ma trận hỗn loạn của mô hình EfficientNet-B0

		Ma trận hỗn loạn			
		Bông không bệnh	Bông cháy	Bông nghẹn	Bông thán thư
Niên thật	Bông không bệnh	123	3	0	0
	Bông cháy	16	110	0	0
	Bông nghẹn	0	0	123	3
	Bông thán thư	5	2	10	109
		Bông không bệnh	Bông cháy	Bông nghẹn	Bông thán thư
		Nhân dự đoán			

Hình 12. Ma trận hỗn loạn của mô hình MobileNetV2

$$P = \frac{TP}{TP + FP} \tag{1}$$

Thu hồi R (Recall): Tỷ lệ TP từ tổng số TP và âm tính giả FN (False Negatives). Đối với các vấn đề phân loại nhiều lớp, R được tính trung bình trong số tất cả các lớp. Sau đây là công thức để thu hồi (2).

$$R = \frac{TP}{TP + FN} \tag{2}$$

Điểm F1 (F1): Ý nghĩa hài hòa của độ chính xác và thu hồi. Đối với các vấn đề phân loại nhiều lớp, F1 được tính trung bình trong số tất cả các lớp. Sau đây là công thức tính điểm F1 (3).

$$F1 = \frac{2 * TP * FP}{TP + FP} \tag{3}$$

Hỗ trợ (Support): Số lần xuất hiện thực tế của lớp trong tập dữ liệu. Nó không khác nhau giữa các mô hình, nó chỉ chẩn đoán quá trình đánh giá hiệu suất.

Macro Avg (Macro averaged Precision): Độ chính xác trung bình vĩ mô được tính bằng mức trung bình của độ chính xác của tất cả các lớp.

Weighted avg (Weighted averaged Precision): Độ chính xác trung bình có trọng số cũng được tính

toán dựa trên độ chính xác trên mỗi lớp nhưng có tính đến số lượng mẫu của từng lớp trong dữ liệu.

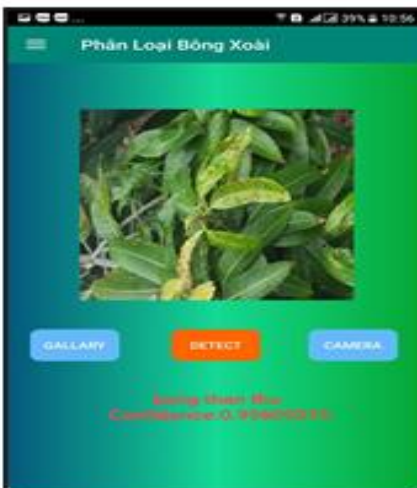
Kết quả từ ma trận hỗn loạn cho thấy, mô hình EfficientNet-B0 (Hình 11) dự đoán nhầm lẫn giữa các lớp thấp hơn so với mô hình MobileNetV2 (Hình 12) và báo cáo dự đoán của ma trận hỗn loạn trên tập ảnh test được trình bày như Bảng 4 trên hai mô hình MobileNetV2 và EfficientNet-B0.

Kết quả nhận dạng thực tế trên điện thoại hệ điều hành Android Samsung Note 5 có độ chính xác đạt 99,6% với mô hình EfficientNet-B0 như Hình 12 (link video: <https://youtu.be/8gwjYfWB-Ho?si=0zhHtVWBOESBQIC>)

Kết quả ở Bảng 4 và Hình 9, Hình 10, Hình 11, Hình 12 cho thấy tỷ lệ mô hình nhận diện đúng bệnh bông xoài trên tập dữ liệu kiểm tra độc lập (test_images) dao động từ 98% đến 100% và kiểm nghiệm thực tế chạy ứng dụng trên điện thoại thông minh Android Samsung Note 5 nhận dạng chính xác từ 96,48% đến 99,60% như ở Hình 12 với các bệnh cần nhận dạng. Bước đầu cho thấy khả năng ứng dụng công nghệ trí tuệ nhân tạo để nhận dạng bệnh bông xoài trong giai đoạn bắt đầu ra hoa đến khi đầu trái non trên điện thoại thông minh thông qua việc phân loại ảnh màu là hoàn toàn khả thi.

Bảng 4. Bảng so sánh báo cáo kết quả mô hình MobileNetV2 và mô hình EfficientNetB0

		Precision	Recall	F1-Score	Support
MobileNetV2	Bông không bệnh	0,91	0,99	0,95	126
	Bông cháy	0,98	0,93	0,95	126
	Bông nghẹn	0,95	0,98	0,97	126
	Bông thán thư	0,98	0,91	0,95	126
	Accuracy			0,95	504
	Macro avg	0,95	0,95	0,95	504
	Weighted avg	0,95	0,95	0,95	504
EfficientNet-B0	Bông không bệnh	1,00	1,00	1,00	126
	Bông cháy	1,00	1,00	1,00	126
	Bông nghẹn	0,98	1,00	0,99	126
	Bông thán thư	1,00	0,99	0,99	126
	Accuracy			1,00	504
	Macro avg	1,00	1,00	1,00	504
	Weighted avg	1,00	1,00	1,00	504



Hình 13. Kết quả nhận dạng bông xoài trên điện thoại Android

4. KẾT LUẬN

Nghiên cứu đã xây dựng được mô hình ứng dụng kỹ thuật học sâu trong bài toán phân loại (nhận dạng) ảnh màu, thông qua phương pháp học chuyển tiếp kết hợp với tinh chỉnh mô hình EfficientNet-B0 đã cho kết quả nhận dạng bông xoài với độ chính xác cao 99,60%. Vì vậy, thông qua hình ảnh bông xoài sử dụng kỹ thuật xử lý ảnh, kỹ thuật học sâu chuyển tiếp và tinh chỉnh, xây dựng ứng dụng (.apk). Nghiên cứu đã đáp ứng được yêu cầu về nhận dạng bệnh trên bông xoài trong giai đoạn bắt đầu ra hoa đến lúc đậu trái non.

Đề xuất nghiên cứu phát triển thêm ứng dụng trên thiết bị di động hệ điều hành IOS, Raspbian, Linux giúp hỗ trợ tốt hơn cho việc nhận dạng bệnh trên bông xoài trong thực tế. Thu thập thêm bộ dữ liệu huấn luyện lớn hơn, không chỉ dữ liệu của ba loại bệnh đã thực hiện trong nghiên cứu mà các loại bệnh phổ biến khác trên bông xoài.

LỜI CẢM ƠN

Nhóm nghiên cứu xin chân thành cảm ơn gia đình và cảm ơn đến các bác, các cô có vườn xoài tại Cù Lao Giêng, xã Mỹ Hiệp, xã Tấn Mỹ, xã Bình Phước Xuân, huyện Chợ Mới, tỉnh An Giang đã hỗ trợ nghiên cứu trong quá trình thu thập dữ liệu.

TÀI LIỆU THAM KHẢO

- Ashok, V., & Vinod, D. S. (2021). A Novel Fusion of Deep Learning and Android Application for Real-Time Mango Fruits Disease Detection. In S. C. Satapathy, V. Bhateja, B. Janakiramaiah, & Y. W. Chen (Eds.), *Intelligent System Design* (pp. 781–791). Springer Singapore. https://doi.org/10.1007/978-981-15-5400-1_74
- Báo điện tử Đảng Cộng Sản Việt Nam. (2011). *Đông Tháp: Cần phát triển cây xoài Nam bộ theo hướng bền vững*. <https://dangcongsan.vn/kinh-te/dong-thap-can-phat-trien-cay-xoai-nam-bo-theo-huong-ben-vung-89060.html>.
- Bazzani, L., Bergamo, A., Anguelov, D., & Torresani, L. (2016). Self-taught object localization with deep networks. *Proceedings of the IEEE 2016 Winter Conference on Applications of Computer Vision (WACV 2016)* (pp.1-13). <https://doi.org/10.48550/arXiv.1409.3964>
- Chollet, F. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. *Computer Vision and Pattern Recognition* (pp. 1251-1258). <https://doi.org/10.48550/arXiv.1610.02357>
- Hâu, T. V. (2009). *Giáo trình xử lý ra hoa cây ăn trái*. Nhà xuất bản Đại học Quốc gia Thành phố Hồ Chí Minh. <https://doi.org/10.48550/arXiv.1905.11946>
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, 86(11), 2278-2323.
- Mark, S., Andrew, H., Menglong Z., Andrey, Z., & Liang-Chieh, C. MobileNetV2 (2019): Dự lượng đảo ngược và tắc nghẽn tuyến tính. *Computer Vision and Pattern Recognition (cs.CV)* arXiv:1801.04381 [cs.CV]
- Singh, U. P., Chouhan, S. S., Jain, S., & Jain, S. (2019). Multilayer Convolution Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease. *IEEE Access*, 7, 43721-43729. <https://doi.org/10.1109/ACCESS.2019.2907383>
- Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks, *36th International Conference on Machine Learning, ICML 2019* (pp. 10691-10700). <https://doi.org/10.48550/arXiv.1905.11946>
- Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning Deep Features for Discriminative Localization, *Proceedings of the IEEE Computer Society Conference on CVPR* (pp. 2921-2929). <https://doi.org/10.1109/CVPR.2016.319>