

Mục lục

1	Lý do chọn đề tài	2
2	Tìm hiểu vấn đề	2
2.1	Câu hỏi nghiên cứu	2
2.1.1	Một số phương pháp hiện tại	2
2.1.2	Mục tiêu dự án	3
2.2	Ý nghĩa của dự án	3
2.2.1	Ý nghĩa khoa học	3
2.2.2	Ý nghĩa xã hội	3
3	Hệ thống	4
3.1	Dữ liệu	4
3.2	Model	6
3.2.1	Trên dataset gốc	7
3.2.2	Trên dataset đã xử lý	9
3.2.3	Kết luận	11
3.3	Ứng dụng Di động	11
4	Đánh giá chung	14
4.1	Ưu điểm	14
4.2	Nhược điểm	14
4.3	Hướng phát triển	15
5	Kết luận	15
6	Ghi nhận đóng góp	15
7	Tài liệu tham khảo	16

1 Lý do chọn đề tài

Kon Tum là một tỉnh thuộc khu vực Tây Nguyên, nổi tiếng với những loài cây công nghiệp lâu năm. Trong đó, cà phê có vai trò quan trọng đối với người dân nơi đây, đặc biệt là với đồng bào dân tộc thiểu số, giúp xóa đói, giảm nghèo, mang lại nguồn thu nhập lớn cho nhân dân tỉnh nhà.

“...Cà phê là sản phẩm chiếm tỷ trọng lớn trong xã hội và kim ngạch xuất khẩu hằng năm của tỉnh Kon Tum, tạo ra nguồn thu nhập chính cho người dân các huyện sinh sống trên địa bàn có lợi thế phát triển cà phê, đặc biệt là vùng đồng bào dân tộc thiểu số trồng cà phê chè tại các xã vùng Đông Trường Sơn...”

- *Bộ Nông nghiệp và Phát triển Nông thôn (2017)*

Tuy nhiên, hằng năm, sâu bệnh trên cây cà phê đã gây thiệt hại nặng nề cho bà con và các doanh nghiệp sản xuất, chế biến cà phê bản địa.

Vì vậy, tác giả đã nghiên cứu và phát triển dự án: **Nhận diện một số sâu bệnh thường gặp trên cây cà phê qua hình thái của lá**. Đây là một hệ thống nhận diện thông minh, tương đối chính xác, nhằm giúp nhân dân, nhất là đồng bào dân tộc thiểu số nhận diện một số sâu bệnh thường gặp trên cây cà phê dễ dàng, thuận tiện để có các biện pháp phòng trừ sâu bệnh hiệu quả, góp phần cải thiện ngành trồng trọt cây cà phê tại tỉnh Kon Tum.

2 Tìm hiểu vấn đề

2.1 Câu hỏi nghiên cứu

Trước khi nghiên cứu, tác giả đưa ra một số câu hỏi nghiên cứu:

- Hiện tại, có những phương pháp nào để chẩn đoán sâu bệnh trên cây cà phê?
- Nhận dạng các đặc điểm của sâu bệnh bằng mắt thường như thế nào? Làm thế nào để máy có thể nhận dạng tương tự?

2.1.1 Một số phương pháp hiện tại

Theo tìm hiểu của tác giả, hiện nay có hai phương pháp chính để chẩn đoán sâu bệnh.

Phương pháp đầu tiên: chẩn đoán thông qua đặc điểm hình thái trên lá. Ưu điểm là nhanh chóng, thuận tiện, chi phí thấp... Tuy nhiên, nhược điểm rõ ràng nhất là khó phát hiện được các sâu bệnh phức tạp, ít thể hiện trên hình thái mà phải phân tích chuyên sâu.

Phương pháp thứ hai: phân tích cấu trúc lá trong phòng thí nghiệm. Phương pháp này cho độ chính xác cao hơn, xác định được những sâu bệnh phức tạp hơn mà không thể quan sát bằng mắt thường. Nhưng việc nghiên cứu trong phòng thí nghiệm đòi hỏi nhiều thời gian, công sức để phân tích, chi phí cho cơ sở vật chất...

Với mong muốn xây dựng một giải pháp nhanh chóng và thuận tiện, phương pháp nghiên cứu thông qua đặc điểm hình thái là phương pháp được lựa chọn làm cơ sở cho việc xây dựng giải pháp.

2.1.2 Mục tiêu dự án

Dự án nghiên cứu nhằm:

- Xây dựng một hệ thống giúp nhân dân, đặc biệt là đồng bào dân tộc thiểu số (đa số có trình độ dân trí thấp, chưa biết ứng dụng khoa học kỹ thuật vào canh tác) nhận diện các loại sâu bệnh trên cây cà phê qua hình ảnh với độ chính xác cao;
- Ứng dụng hệ thống vào thực tiễn tại địa phương, giúp giảm thiểu thiệt hại do sâu bệnh gây ra, góp phần phát triển kinh tế - xã hội trên địa bàn, giúp đồng bào dân tộc thiểu số vươn lên thoát nghèo, làm giàu trên chính quê hương của mình.

2.2 Ý nghĩa của dự án

2.2.1 Ý nghĩa khoa học

Đề tài nghiên cứu phương pháp nhận diện qua hình ảnh, tuy phổ biến trên những đối tượng khác nhưng lại ít được quan tâm trên cây cà phê. Vì vậy, tác giả mong muốn đề tài này sẽ mở rộng nghiên cứu nhận diện hình ảnh trên cây cà phê trong tương lai.

2.2.2 Ý nghĩa xã hội

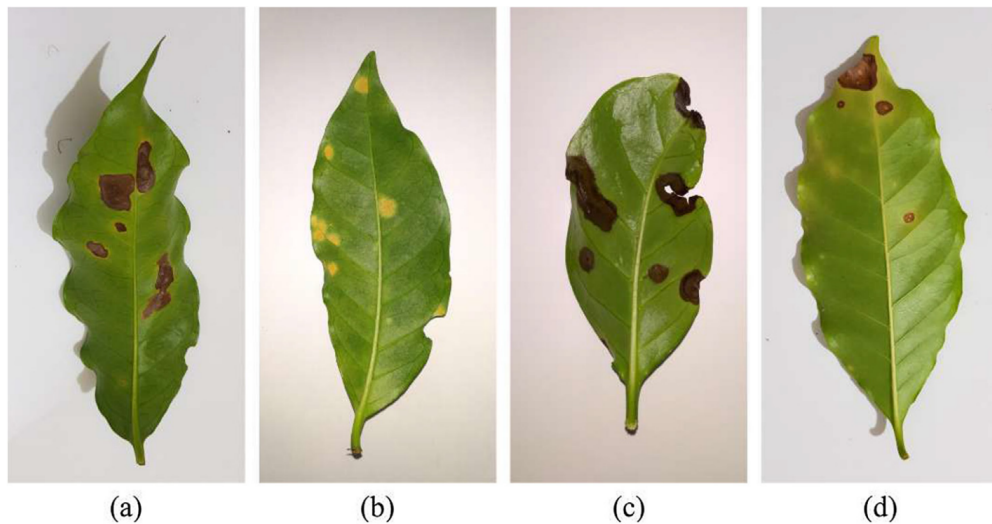
Đối tượng nghiên cứu của dự án là cây cà phê, một trong những cây công nghiệp quan trọng ở tỉnh Kon Tum. Do đó, thông qua việc nghiên cứu dự án, tác giả muốn giúp nhân dân, đặc biệt là đồng bào dân tộc thiểu số dễ dàng nhận diện một số sâu bệnh cơ bản để có cách phòng trừ hiệu quả, góp phần xóa đói, giảm nghèo, phát triển kinh tế - xã hội trên địa bàn.

3 Hệ thống

3.1 Dữ liệu

Dữ liệu được kết hợp thu thập từ thực địa và từ các dataset khác trên mạng. Dữ liệu từ thực địa được thu thập tại các xã, thị trấn thuộc huyện Đăk Hà và xã Ia Chim, thành phố Kon Tum, tỉnh Kon Tum. Ảnh được thu thập từ nhiều thiết bị khác nhau (Samsung Galaxy A02, ASUS Zenfone 2, Xiaomi Redmi 5A, Xiaomi S2, Galaxy S8, và iPhone 6S). Mẫu vật được đặt trên nền trắng, dưới điều kiện kiểm soát một phần.

Một dataset gồm 1685 bức ảnh chụp các lá cà phê (bao gồm lá bình thường và lá sâu bệnh) đã được thu thập và phân loại. Kích thước của mỗi ảnh là 2048 x 1024.



Hình 1: Sâu vẽ bùa (a), Gỉ sắt (b), Đốm nâu (c), Đốm mắt cua (d)

Sâu bệnh	Số lượng ảnh
Bình thường	272
Sâu vẽ bùa	387
Gỉ sắt	531
Đốm nâu	348
Đốm mắt cua	147
Tổng	1685

Bảng 1: Thống kê dữ liệu thu thập được

Tuy nhiên, mỗi bức ảnh có kích thước rất lớn, gây khó khăn trong quá trình xử lý về sau. Do đó, những vùng có sâu bệnh sẽ được cắt ra. Những vùng hoàn toàn bình thường trên lá cũng sẽ được cắt, nhằm mục đích đối

chiều với các vùng sâu bệnh. Sau tiền xử lý, thu được 2722 bức ảnh cắt xung quanh vùng có sâu bệnh và không có sâu bệnh trên lá. Kích thước của mỗi ảnh là 224 x 224.



Hình 2: Các vùng được cắt sau tiền xử lý

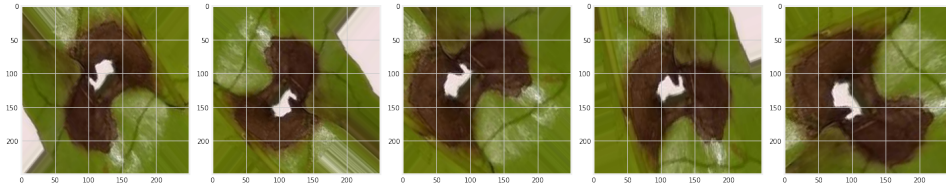
Sâu bệnh	Số lượng ảnh
Bình thường	256
Sâu vẽ bùa	593
Gỉ sắt	991
Đốm nâu	504
Đốm mắt cua	378
Tổng	2722

Bảng 2: Thống kê dữ liệu sau tiền xử lý

Như vậy, có hai dataset: dataset gốc (gồm các ảnh chưa xử lý) và dataset đã qua xử lý. Từng dataset được chia ra thành training set, validation set và test set, theo tỉ lệ 70-15-15. Data Augmentation được áp dụng trên các tập để tăng cường dữ liệu cho các set.

Cấu hình	Giá trị
Rescaling factor	$\frac{1}{255}$
Rotation range	180°
Sheer range	20%
Zoom range	20%
Random flip	true

Bảng 3: Cấu hình Data Augmentation



Hình 3: Ảnh được chỉnh sửa qua Data Augmentation

3.2 Model

Machine Learning đang được ứng dụng rộng rãi, hỗ trợ rất nhiều lĩnh vực trong đời sống. Machine Learning có khả năng học từ data, và ứng dụng những gì học được để giải quyết các bài toán cụ thể.

Deep Learning, một nhánh của Machine Learning, mô phỏng cách não bộ con người tiếp nhận và xử lý thông tin bằng một neural network gồm nhiều tầng xếp lên nhau. Deep Learning có khả năng tự chiết xuất đặc tính từ data, và càng nhiều data thì model Deep Learning càng hiệu quả. Vì vậy, Deep Learning là giải pháp phù hợp để phân loại hình ảnh.

Các model được thử nghiệm trong dự án bao gồm: InceptionV3, MobileNetV2, ResNet50V2, VGG16 và Xception. Những model này đều đã được học qua ImageNet và được tiếp tục luyện trên toàn bộ mạng bằng data mới mà không đóng băng thông số nào.

Kiến trúc	Số lượng tham số (triệu)	Số tầng
InceptionV3	24	48
MobileNetV2	2,2	54
ResNet50V2	25	50
VGG16	138	16
Xception	23	71

Bảng 4: Thông số của các kiến trúc được thử nghiệm

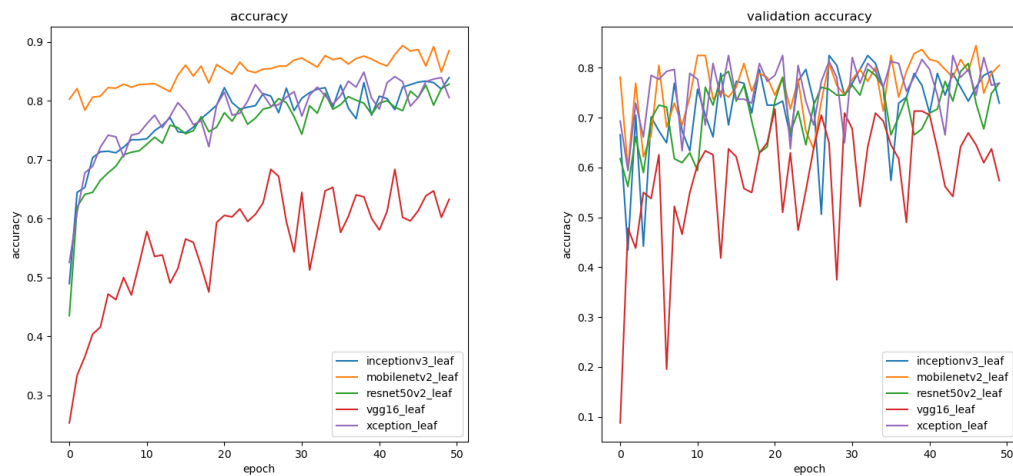
Hyperparameter	Giá trị
Batch size	24
Epochs	50
Loss function	Categorical Cross-Entropy
Optimization function	Stochastic Gradient Descent ($\alpha = 0.001$, $\eta = 0.9$)
Regularization function	L2 Regularization ($\lambda = 5 \times 10^{-4}$)

Bảng 5: Bảng giá trị của các Hyperparameter

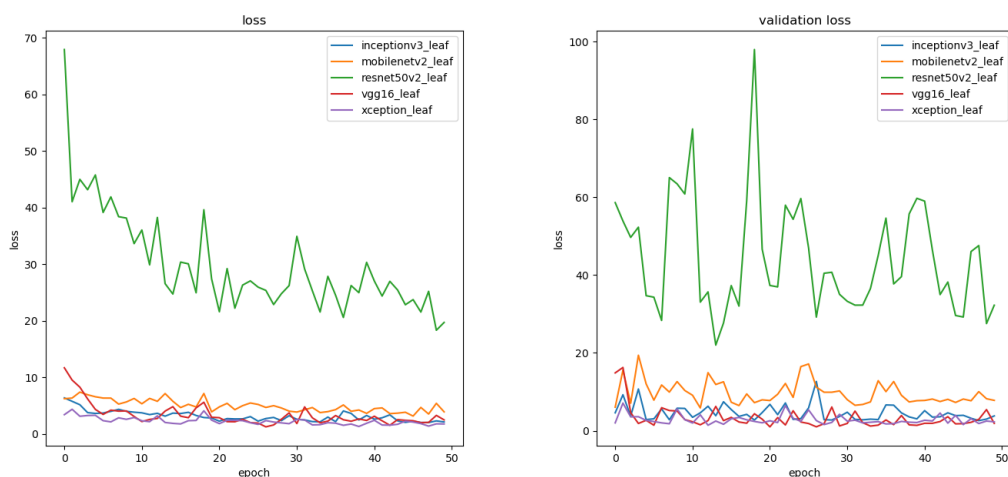
Các model được huấn luyện trên môi trường Google Colab (NVIDIA Tesla K80). Thư viện Keras được sử dụng để xây dựng và huấn luyện cho model.

3.2.1 Trên dataset gốc

Các model ban đầu được train trên dataset gốc. Kích thước gốc của các ảnh (2048 x 1024) quá lớn, gây tràn stack trong quá trình train. Vì vậy, ảnh được downsample (kích thước mỗi chiều chỉ còn $\frac{1}{4}$ kích thước ban đầu).



Hình 4: Biểu đồ accuracy trên training set và validation set (gốc)



Hình 5: Biểu đồ loss trên training set và validation set (gốc)

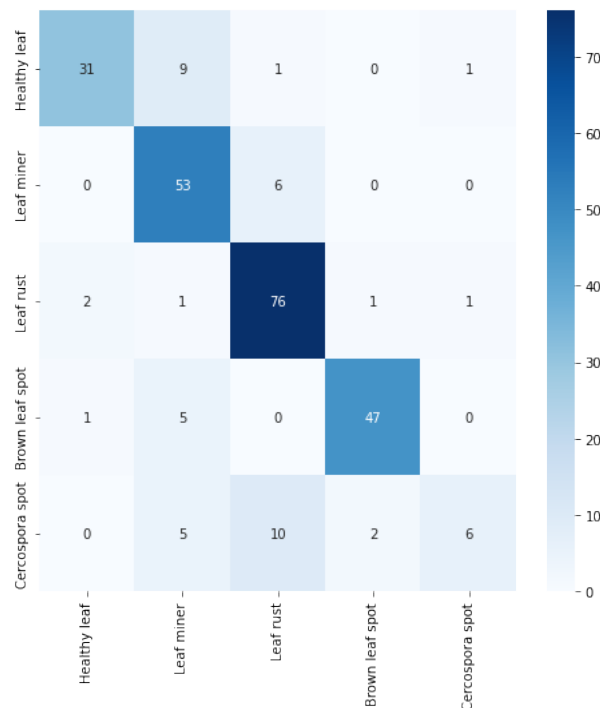
Kiến trúc	Thời gian (s)
InceptionV3	76.54
MobileNetV2	76.28
ResNet50V2	77.94
VGG16	78.82
Xception	88.88

Bảng 6: Thời gian trung bình cho mỗi epoch trên dataset (gốc)

Kiến trúc	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
InceptionV3	72.09	72.09	72.09	72.09
MobileNetV2	82.55	82.55	82.55	82.55
ResNet50V2	75.58	75.58	75.58	75.58
VGG16	56.20	56.20	56.20	56.20
Xception	75.96	75.96	75.96	75.96

Bảng 7: Kết quả đo các model trên test set (gốc)

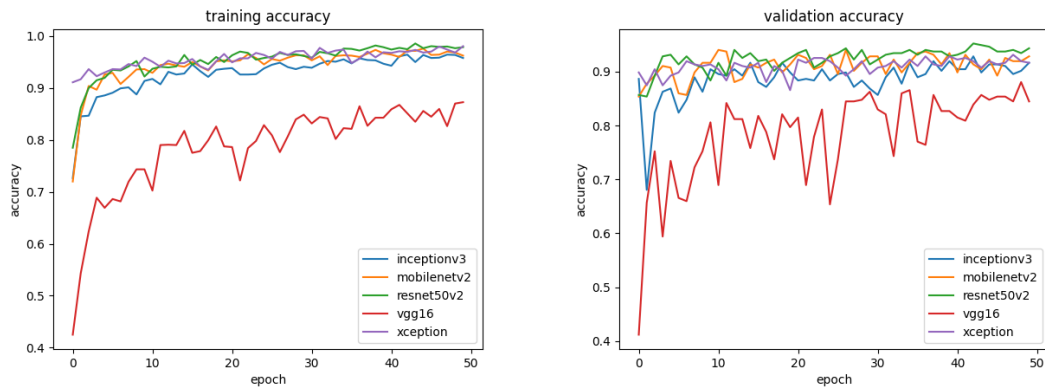
MobileNetV2 đạt được F1-score 82.55%, vượt trội so với các model khác. Tốc độ train của model này cũng nhanh nhất, đạt 76.28s/epoch. Với lợi thế yêu cầu tính toán và yêu cầu bộ nhớ thấp nhờ vào lượng trainable parameter ít hơn đáng kể, MobileNetV2 xử lý khá ổn với input lớn và chưa qua xử lý. Cũng vì không cần nhiều sức mạnh tính toán, nên MobileNetV2 vốn là model thích hợp để tích hợp trực tiếp trên các thiết bị di động.



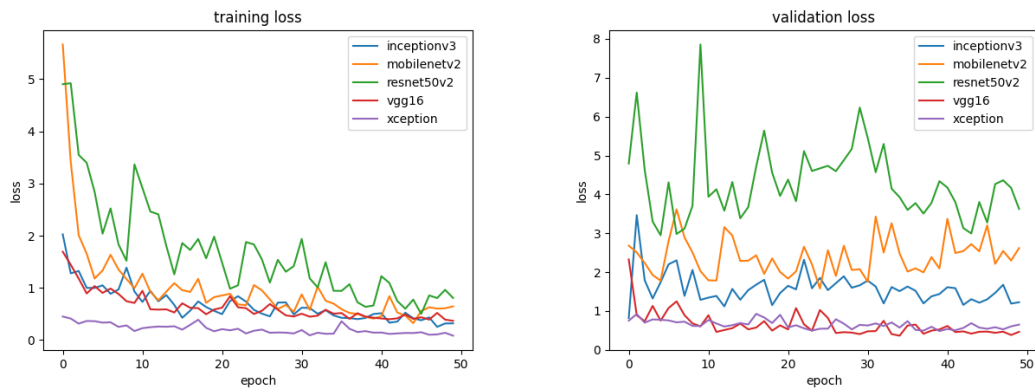
Hình 6: Confusion Matrix của MobileNetV2 trên test set (gốc)

Kết quả đo trên dataset gốc không thực sự tốt. Việc giữ nguyên ảnh gốc khiến cho khối lượng input rất lớn, trong khi downsample sẽ giảm độ chi tiết, gây khó khăn trong nhận diện. Tốc độ train vẫn còn rất chậm, dù kích thước của input đã giảm.

3.2.2 Trên dataset đã xử lý



Hình 7: Biểu đồ accuracy trên training set và validation set (đã xử lý)



Hình 8: Biểu đồ loss trên training set và validation set (đã xử lý)

Kiến trúc	Thời gian (s)
InceptionV3	24.20
MobileNetV2	21.28
ResNet50V2	24.58
VGG16	27.20
Xception	27.42

Bảng 8: Thời gian trung bình cho mỗi epoch trên dataset (đã xử lý)

Kiến trúc	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
InceptionV3	93	93	93	93
MobileNetV2	90	90	90	90
ResNet50V2	93	93	93	93
VGG16	84	84	84	84
Xception	93	93	93	93

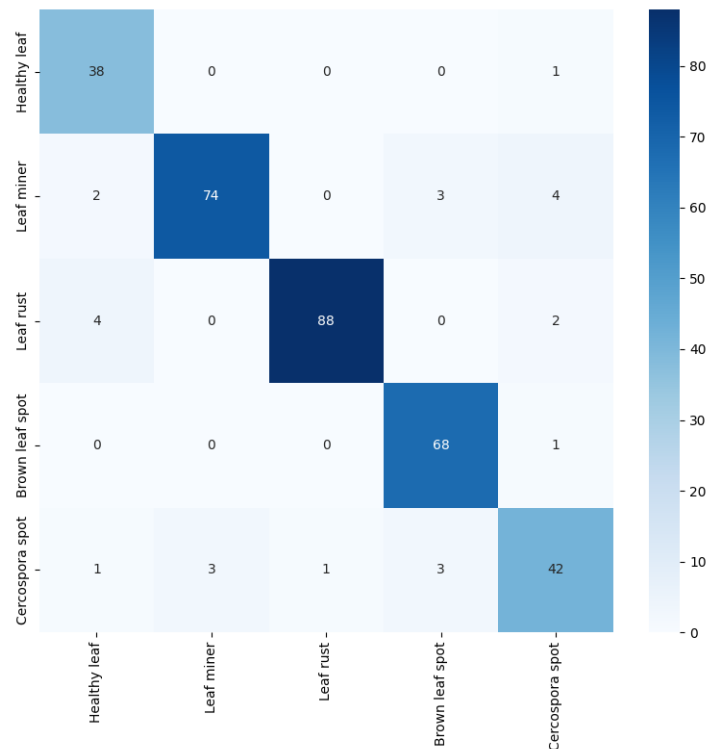
Bảng 9: Kết quả đo các model trên test set (đã xử lý)

Trên dataset đã xử lý, khả năng tính toán của các model có phần khác biệt so với trên dataset gốc. Các kết quả đo đều có sự cải thiện rõ rệt một cách đáng kể.

Ba model InceptionV3, ResNet50V2 và Xception đều có F1-score đạt 93%, tốt hơn MobileNetV2 (90%) và VGG16 (84%). Cả ba model đều có số lượng parameter và số layer tương đối tương đồng.

MobileNetV2 vẫn là model có tốc độ train nhanh nhất, dù hiệu năng có phần thua thiệt hơn. Sức mạnh của VGG16 kém xa các model còn lại, trong khi thời gian train chỉ nhanh hơn Xception.

Mặc dù thời gian train chậm nhất, nhưng với lợi thế là model có nhiều tầng nhất, Xception có phần nhỉnh hơn khi chỉ mất vài epoch để các thông số đạt trạng thái gần tối ưu. Xuyên suốt quá trình train, biểu đồ của Xception cũng ổn định hơn nhiều.



Hình 9: Confusion Matrix của Xception trên test set (đã xử lý)

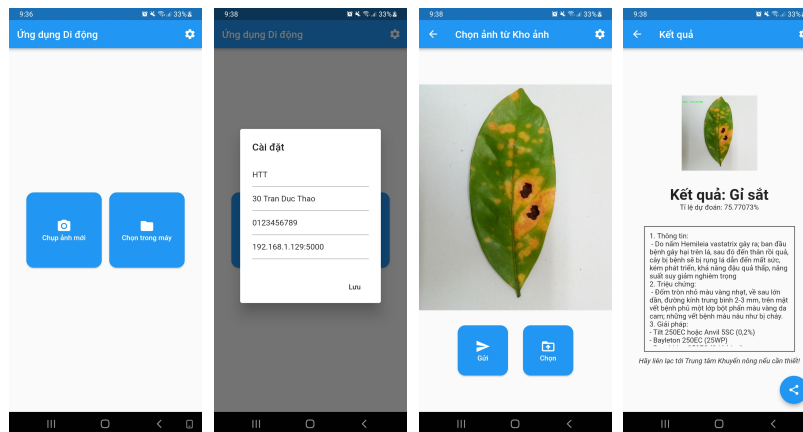
3.2.3 Kết luận

Trên dataset gốc, kết quả thu thập được không quá tốt khi chỉ MobileNetV2 đạt F1-score trên 80%. Vì vậy, việc xử lý data trước khi train là cần thiết. Kết quả train trên dataset đã xử lý khả quan hơn hoàn toàn. Ba model InceptionV3, ResNet50V2 và Xception thể hiện tương đồng trên dataset này. InceptionV3 có tốc độ train nhanh nhất, còn Xception đạt trạng thái tối ưu sớm nhất.

Sau quá trình nghiên cứu và thử nghiệm, **Xception** là model phù hợp nhất với bài toán và phương pháp xử lý dữ liệu.

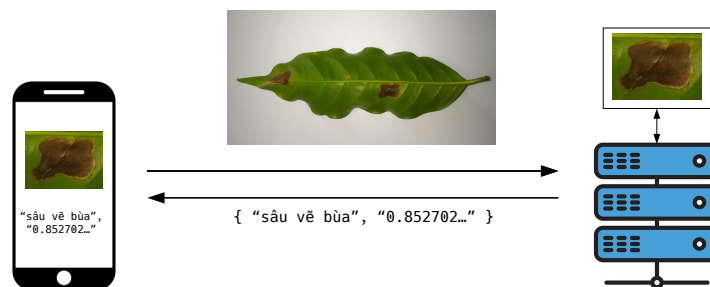
3.3 Ứng dụng Di động

Để ứng dụng kết quả từ model trên vào thực tế, tác giả đã phát triển một ứng dụng di động, liên lạc với máy chủ để xử lý.



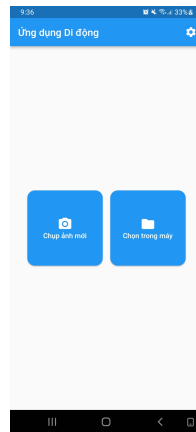
Hình 10: Giao diện của ứng dụng di động

Ứng dụng được thiết kế đơn giản, gọn nhẹ, dễ sử dụng với bà con, đặc biệt là đồng bào dân tộc thiểu số.



Hình 11: Sơ đồ hoạt động

Ứng dụng cho phép người dùng gửi hình ảnh (ảnh có sẵn trong máy hoặc ảnh chụp mới) lên máy chủ. Máy chủ tiếp nhận ảnh, xử lý ảnh bằng model được xây dựng, rồi trả kết quả chẩn đoán về lại ứng dụng. Các kết quả trả về bao gồm tên bệnh, một số triệu chứng, đặc điểm đặc trưng và tên một vài thuốc đặc trị cho sâu bệnh.



Hình 12: Màn hình chính

Ở màn hình chính, có 2 lựa chọn: **Chụp ảnh mới** (chụp ảnh từ camera để lấy ảnh) và **Chọn trong máy** (chọn ảnh có sẵn trong bộ nhớ của máy).



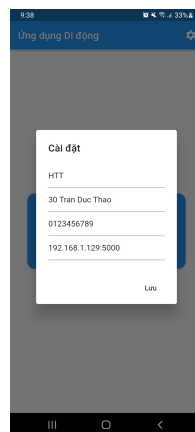
Hình 13: Màn hình chọn ảnh

Ban đầu, sẽ không có ảnh nào được chọn. Nhấn **Chọn** để chọn ảnh, sẽ dẫn tới trình chụp ảnh hoặc trình chọn ảnh trong máy (tùy thuộc vào lựa chọn ở màn hình trước đó). Sau khi chọn ảnh, ảnh sẽ xuất hiện trên màn hình. Nhấn **Gửi** để gửi ảnh sang hệ thống xử lý.



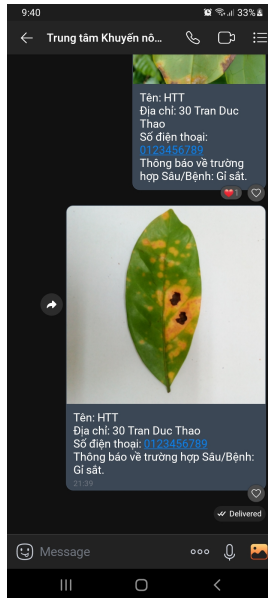
Hình 14: Màn hình kết quả

Sau khi chờ máy chủ xử lý xong, kết quả sẽ được trả về. Thông tin trả về bao gồm tên sâu bệnh, triệu chứng, một vài thông tin và một vài thuốc chữa. Ở góc phải dưới của màn hình, có nút **Chia sẻ** qua các ứng dụng, mạng xã hội hay e-mail, để bà con có thể liên lạc tới các Trung tâm về nông nghiệp để được tư vấn thêm.



Hình 15: Cài đặt

Nút **Cài đặt** có hình bánh răng cưa, ở góc phải trên ở mọi màn hình. Người dùng có thể nhập Tên, Địa chỉ, Số điện thoại để được đính kèm khi **Chia sẻ** ở Màn hình Kết quả.



Hình 16: Tính năng chia sẻ (trong ảnh: Zalo)

4 Đánh giá chung

Đề tài này nghiên cứu việc ứng dụng Trí tuệ Nhân tạo vào việc nhận diện một số sâu bệnh thường gặp, thông qua biểu hiện của lá. Dữ liệu cho việc rèn luyện được thu thập từ khu vực bản địa của tác giả, song song với tìm kiếm trên mạng. Một số Mạng Tích chập đã được thử nghiệm, trong đó Xception mang lại kết quả tốt nhất. Để ứng dụng model vào thực tế, một ứng dụng di động đơn giản đã được xây dựng, có chức năng gửi và tiếp nhận thông tin từ model.

Trí tuệ Nhân tạo là một phương pháp khả thi để nhận diện sâu bệnh qua hình thái của lá cà phê, và hoàn toàn có thể tiếp tục phát triển trong tương lai, mở rộng sang nhiều giống cây trồng khác, phục vụ người nông dân ở khắp mọi miền đất nước.

4.1 Ưu điểm

- Tỷ lệ nhận diện mang lại hiệu quả tương đối chính xác;
- Tiết kiệm thời gian, công sức, chi phí cho việc chẩn đoán so với các phương pháp hiện có;
- Ứng dụng di động gọn nhẹ, đơn giản, dễ sử dụng.

4.2 Nhược điểm

- Model vẫn chưa thực sự tối ưu, quá trình rèn luyện chưa đạt kết quả tốt nhất, tốc độ nhận diện còn chậm;

- Kích thước dataset để rèn luyện còn hạn chế, cần được mở rộng và làm phong phú thêm;
- Cần bổ sung thêm cơ sở dữ liệu về sâu bệnh và các thông tin liên quan.

4.3 Hướng phát triển

- Cải tiến model, cải tiến phương pháp luyện để khai thác hết tiềm năng của model, thử nghiệm với những model mới;
- Thu thập thêm dữ liệu, xây dựng một bộ cơ sở dữ liệu trung tâm;
- Hoàn thiện, phát triển thêm tính năng cho ứng dụng di động;
- Hợp tác với các nhà nghiên cứu về cây cà phê để tiếp tục hoàn thiện dự án.

5 Kết luận

Mục đích nghiên cứu của dự án xuất phát từ tình hình thực tiễn tại chính địa phương nơi tác giả đang học tập và sinh sống; với hy vọng đề tài sẽ góp phần cải thiện ngành trồng trọt cà phê ở Kon Tum, giúp người dân, đặc biệt là đồng bào dân tộc thiểu số vươn lên thoát nghèo, làm giàu chính đáng trên quê hương của mình; góp phần phát triển kinh tế của địa phương.

6 Ghi nhận đóng góp

Xin gửi lời cảm ơn đặc biệt đến mẹ và em gái của tác giả vì đã luôn ủng hộ, dõi theo và tin tưởng tác giả trên con đường nghiên cứu khoa học.

Xin chân thành cảm ơn:

- Thầy Hồ Hữu Sơn, trường THPT Chuyên Nguyễn Tất Thành, tỉnh Kon Tum
- Thạc sĩ Trần Văn Cao Sơn, Sở Nông nghiệp & Phát triển nông Thôn tỉnh Kon Tum
- Anh Bùi Đình Nguyên Khoa, trường Đại học Khoa học Tự nhiên - ĐHQG TP.HCM
- Anh Nguyễn Hoàng Khang, trường Đại học Khoa học Tự nhiên - ĐHQG TP.HCM
- Anh Hồ Khánh Duy, trường Đại học Khoa học Tự nhiên - ĐHQG TP.HCM
- Gia đình ông Mai Văn Khá, Thị trấn Đăk Hà, huyện Đăk Hà, tỉnh Kon Tum

- Gia đình ông Nguyễn Đức Vũ, Thị trấn Đắk Hà, huyện Đắk Hà, tỉnh Kon Tum
 - Gia đình ông Hoàng Nguyên Chiến, xã Ia Chim, thành phố Kon Tum, tỉnh Kon Tum
 - Chị Phạm Xuân My, trường Đại học Luật TP.HCM
- đã hỗ trợ tác giả hoàn thành đề tài nghiên cứu này.

7 Tài liệu tham khảo

- [1] Suhartono, Derwin & Aditya, Wahyu & Lestari, Miranty & Yasin, Muhammad. (2013). Expert System in Detecting Coffee Plant Diseases. International Journal of Electrical Energy. 156-162. 10.12720/ijoe.1.3.156-162.
- [2] M. Kumar, P. Gupta, P. Madhav and Sachin, "Disease Detection in Coffee Plants Using Convolutional Neural Network," 2020 5th International Conference on Communication and Electronics Systems (ICCES), 2020, pp. 755-760, doi: 10.1109/ICCES48766.2020.9138000.
- [3] Esgario, J. G., Krohling, R. A., & Ventura, J. A. (2020) "Deep learning for classification and severity estimation of coffee leaf biotic stress" Computers and Electronics in Agriculture 169, 105162, doi:10.1016/j.compag.2019.105162
- [4] Krohling, Renato; esgario, José; Ventura, Jose A. (2019), "BRACOL - A Brazilian Arabica Coffee Leaf images dataset to identification and quantification of coffee diseases and pests", Mendeley Data, V1, doi: 10.17632/yy2k5y8mxg.1
- [5] François Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions", arXiv:1610.02357v3 [cs.CV], Apr 2017.