

XÂY DỰNG ỨNG DỤNG DI ĐỘNG NHẬN DIỆN BỆNH TRÊN DÂU TÂY Ở VIỆT NAM DÙNG MACHINE LEARNING

BUILDING A MOBILE APP FOR DETECTING STRAWBERRY DISEASES IN VIETNAM USING DEEP LEARNING

Đỗ Bá Quang Huy¹, Trần Thủy Văn^{1,*},
Đinh Văn Lực¹, Trần Quang Huy¹

DOI: <http://doi.org/10.57001/huiv5804.2024.307>

TÓM TẮT

Các tiến bộ trong thị giác máy tính đã mở ra nhiều giải pháp tiềm năng cho ngành nông nghiệp, giúp nâng cao cả chất lượng lẫn sản lượng cây trồng, tăng cường sức cạnh tranh kinh tế và giảm chi phí lao động. Tuy nhiên, việc phát hiện bệnh trên quả trước thu hoạch vẫn phụ thuộc nhiều vào kinh nghiệm của người nông dân, gây khó khăn trong việc kiểm soát và duy trì chất lượng sản phẩm. Để giải quyết vấn đề này, chúng tôi giới thiệu một ứng dụng di động có sẵn trên cả hai hệ điều hành Android và iOS, sử dụng công nghệ xử lý hình ảnh và trí tuệ nhân tạo để nhận diện bệnh trên quả dâu tây. Mô hình được đề xuất, YOLOv8, với số lượng tham số tối thiểu, được áp dụng để đánh giá và phát hiện bệnh với độ chính xác cao. Phương pháp này được xem là một hướng đi hiệu quả để đáp ứng nhu cầu phát hiện bệnh trên quả dâu tây.

Từ khóa: Thị giác máy tính; YOLOv8; các bệnh trên quả dâu tây.

ABSTRACT

The advancements in computer vision have opened up potential solutions for agriculture, enhancing the quality and yield of agricultural products, boosting economic competitiveness, and reducing labor costs. However, the detection of diseases on fruits before harvest still relies heavily on the experience of farmers, posing challenges in controlling and stabilizing product quality. To address this issue, we propose a mobile application for both Android and iOS operating systems, utilizing image processing technology and artificial intelligence to identify diseases on strawberries. The proposed model, YOLOv8, with minimal parameters, is applied to evaluate and detect diseases with high accuracy. This method is considered a feasible solution to meet the demand for disease detection on strawberries.

Keywords: Computer vision; YOLOv8; diseases on strawberries.

¹Khoa Điện, Trường Đại học Công nghiệp Hà Nội

*Email: vantth@hauivn

Ngày nhận bài: 18/4/2024

Ngày nhận bài sửa sau phản biện: 08/6/2024

Ngày chấp nhận đăng: 27/9/2024

1. GIỚI THIỆU

Dâu tây là một loại quả có giá trị kinh tế cao và được tiêu thụ rộng rãi nhờ giá trị dinh dưỡng phong phú. Năm 2020, sản lượng dâu tây toàn cầu đạt giá trị 14 tỷ USD, trong đó Trung Quốc là nước sản xuất hàng đầu với giá trị 5 tỷ USD, vượt xa quốc gia sản xuất lớn thứ hai là Mỹ [1]. Dâu tây chứa nhiều dưỡng chất quan trọng, bao gồm vitamin C, các chất chống oxy hóa như quercetin và anthocyanins, cũng như chất xơ, mangan, vitamin A, vitamin K, axit folic, vitamin E, vitamin B6 và kali [2]. Những thành phần này mang lại nhiều lợi ích cho sức khỏe, như tăng cường hệ miễn dịch, duy trì sự ổn định của đường huyết, bảo vệ tế bào khỏi tổn thương, và hỗ trợ sức khỏe của mắt, da và xương. Do đó, dâu tây là một trong những loại trái cây được tiêu thụ thường xuyên, không chỉ vì giá trị dinh dưỡng mà còn vì hương vị thơm ngon [3].

Sâu bệnh là một thách thức lớn trong nông nghiệp, gây thiệt hại đáng kể cho sản lượng thực phẩm. Gần một nửa sản lượng cây trồng toàn cầu bị ảnh hưởng bởi sâu bệnh [4]. Dâu tây đặc biệt dễ bị nhiễm các loại nấm, vi khuẩn và vi rút gây bệnh [5-7]. Một số mầm bệnh phổ biến ở dâu tây bao gồm *Colletotrichum siamense*, gây bệnh thán thư [8,9]; *Botrytis cinerea*, tác nhân gây mốc xám [10, 11]; *Neopestalotiopsis* spp. [12], gây thối ngọn, thối quả và cháy lá [8]; và các loại nấm khác gây bệnh phấn trắng, ảnh hưởng đến cuống lá, lá và quả dâu tây [13, 14]. Những mầm bệnh này không chỉ làm giảm hiệu suất quang hợp mà còn ảnh hưởng tiêu cực đến chất lượng, sự phát triển và năng suất của quả. Hiện nay, việc xác định bệnh trên dâu tây chủ yếu dựa vào công việc thủ công, tốn nhiều công sức và thời gian. Sự suy giảm lực lượng lao động trong lĩnh vực nông nghiệp làm tăng thêm khó khăn, vì việc dự đoán mức độ nghiêm trọng của

bệnh trên diện rộng trở nên khó khăn. Vì vậy, cần phát triển một phương pháp tự động, nhanh chóng và chính xác để phát hiện sớm các bệnh trên dầu tày.

Vi vậy, nhiều nghiên cứu đã áp dụng phương pháp thị giác máy tính để hỗ trợ phân loại và phát hiện bệnh cây trồng [15, 16]. Trong các công bố, mạng nơ-ron tích chập (CNN) là một trong những kỹ thuật máy học phổ biến nhất để phát hiện bệnh cây trồng. Jeon và Rhee [17] đã sử dụng kỹ thuật CNN trong mô hình GoogLeNet để nhận dạng lá cây. Kỹ thuật này có thể phát hiện những chiếc lá bị hư hỏng với tỷ lệ nhận dạng trên 94%, ngay cả khi chỉ có 30% lá bị tổn thương. Cervantes-Jilaja và cộng sự [18] đã phát triển một phương pháp thị giác máy tính để phát hiện và xác định các khuyết tật trên hạt dẻ dựa trên các đặc điểm bên ngoài như hình dạng, màu sắc, kích thước và kết cấu. Mohanty và cộng sự [19] đã ứng dụng CNN để nhận diện loài cây trồng và phát hiện bệnh từ bộ dữ liệu hình ảnh công khai, sử dụng các mô hình đào tạo như GoogLeNet và AlexNet. Với các đặc trưng về màu sắc, thang độ xám và phân đoạn lá, mô hình của họ đạt độ chính xác lên đến 99,35%.

Dựa trên những hệ thống phân loại và phát hiện tự động sử dụng thị giác máy tính đã được chứng minh hiệu quả đối với các loại nông sản khác. Nghiên cứu này được thực hiện nhằm mục đích ứng dụng thực tiễn trong việc phát hiện bệnh trên quả dầu tày. Chúng tôi đã phát triển một ứng dụng hoạt động trên cả Android và iOS để nhận diện bệnh trên quả dầu tày, tích hợp các tính năng đa chiều và sử dụng mô hình huấn luyện mới là YOLOv8 [20]. Bộ dữ liệu bao gồm số lượng lớn hình ảnh quả dầu tày bị bệnh được thu thập trong phòng thí nghiệm và ngoài vườn, sử dụng các phương pháp như lọc nhiễu và tăng cường độ sắc nét để cải thiện khả năng tổng quát hóa của mô hình. Ngoài ra, chúng tôi cũng đã thử nghiệm mô hình trong điều kiện thiếu sáng để nâng cao khả năng thích ứng của hệ thống trong môi trường tự nhiên và cải thiện hiệu suất phát hiện bệnh trên quả dầu tày.

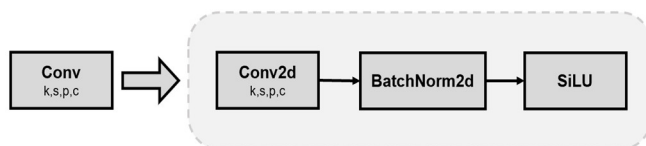
2. CƠ SỞ LÝ THUYẾT

2.1. YOLOv8

Kiến trúc được đề xuất của YOLOv8 để đạt được khả năng trích xuất tính năng tối ưu, xương sống của kiến trúc được đề xuất bao gồm Module Convolutional, mô-đun C2f và mô-đun SPPF, rất quan trọng để trích xuất tính năng.

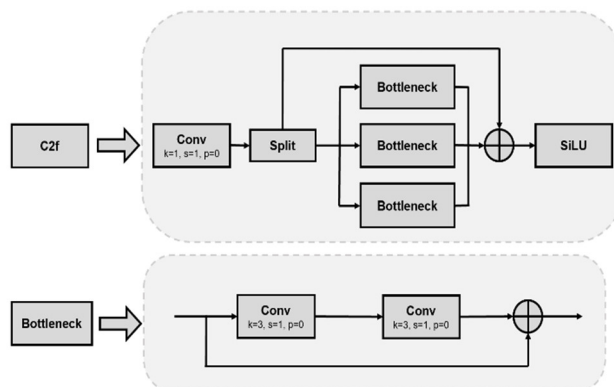
Module Convolutional trong YOLOv8 thường được tích hợp vào mạng cơ sở (backbone network), nơi có nhiệm vụ chuyển đổi ảnh đầu vào thành biểu diễn đặc trưng bằng cách sử dụng các lớp tích chập (convolutional

layers), lớp chuẩn hóa hàng loạt (batch normalization), và hàm kích hoạt SiLU (Sigmoid Linear Unit). Chúng đóng vai trò quan trọng trong việc trích xuất thông tin quan trọng từ ảnh, từ các đặc trưng cấp thấp như cạnh đến các đặc trưng cấp cao như hình dạng và chi tiết của đối tượng. Đồng thời, việc sử dụng hàm kích hoạt SiLU giúp mô hình học được các biểu diễn phi tuyến tính một cách hiệu quả, đóng góp vào khả năng học và nhận diện đối tượng chính xác trong ảnh.



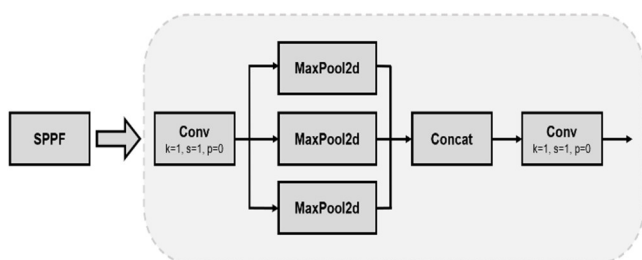
Hình 1. Kiến trúc của Module Convolutional

Khác với YOLOv5, YOLOv8 không đem lại nhiều sự đổi mới đặc biệt, với sự thay đổi đáng chú ý duy nhất là sự giới thiệu của module C2f, thay thế cho module C3 trong YOLOv5. Trong YOLOv5, module C3 đặc trưng ba lớp tích chập chuẩn và nhiều module Bottleneck. Điều đặc biệt là module Bottleneck bao gồm hai nhánh: một nhánh sử dụng nhiều bộ xếp Bottleneck và ba lớp tích chập chuẩn, trong khi nhánh còn lại chỉ đi qua một lớp tích chập cơ bản trước khi ghép hai nhánh lại với nhau. Điều này không chỉ giúp giảm số lượng tham số huấn luyện và tính toán mà còn giải quyết các vấn đề về gradient explosion và gradient disappearance trong mạng sâu, tăng khả năng học của mô hình. Trong khi YOLOv7 cải tiến thông tin gradient thông qua việc thêm nhiều nhánh dòng gradient song song và sử dụng module ELAN để đạt độ chính xác cao hơn và độ trễ hợp lý hơn, YOLOv8 tiếp tục phát triển ý tưởng này bằng cách thiết kế module C2f. Lấy cảm hứng từ module C3 và ELAN, module C2f giúp thu thập đa dạng dòng thông tin gradient trong khi vẫn duy trì cấu trúc nhẹ của mô hình. Chính vì đó đã nâng cao khả năng học và hiệu suất của YOLOv8, đồng thời giảm độ trễ một cách có hiệu quả.



Hình 2. Kiến trúc của Module C2f

Trong YOLOv8, module SPPF đóng vai trò quan trọng trong việc nâng cao độ chính xác của mô hình, đặc biệt là trong nhiệm vụ nhận diện đối tượng. Module này được tích hợp để giải quyết thách thức của việc nhận diện đối tượng ở các vị trí và quy mô khác nhau trong ảnh. SPPF sử dụng phương pháp Spatial Pyramid Pooling để tạo ra biểu diễn đa cấp độ, giúp mô hình chấp nhận và xử lý thông tin từ các vùng có kích thước khác nhau trong ảnh. Do đó, khả năng nhận diện của mô hình được cải thiện, đồng thời giúp giảm thiểu vấn đề về độ chính xác ở các vị trí đa dạng trên ảnh.



Hình 3. Kiến trúc của Module SPPF

Tóm lại, mô hình YOLOv8 là tích hợp của các thành phần quan trọng như Module Convolutional, mô-đun C2f và mô-đun SPPF để hiệu quả trích xuất đặc trưng và đối mặt với thách thức của kích thước đầu vào không đồng đều trong ảnh. Đặc biệt, mô-đun C2f đóng vai trò quan trọng trong việc giảm số lượng tham số và tính toán, giải quyết vấn đề về gradient explosion và gradient disappearance trong mạng học sâu, đồng thời tăng khả năng học của mô hình. Kết quả thử nghiệm đã khẳng định hiệu suất xuất sắc của YOLOv8, đạt độ chính xác cao hơn đáng kể so với các phiên bản YOLO trước đó. Đây là lý do YOLOv8 được chọn cho nghiên cứu phát hiện bệnh của dâu tây.

2.2. Bộ dữ liệu về bệnh trên quả dâu tây

Trước những thách thức của môi trường thực tế, chúng tôi đã phát triển các mô hình học máy được đào tạo trên hình ảnh thực tế, sử dụng nguồn dữ liệu đa dạng từ Google Hình ảnh, cũng như các bức ảnh chụp tại trang trại và từ Bộ Nông nghiệp. Quá trình này bao gồm việc tải xuống hình ảnh từ internet, đặc biệt trong khuôn khổ một nghiên cứu lớn về bệnh dâu tây được thực hiện trên thực địa. Chúng tôi không chỉ dựa vào nguồn dữ liệu trực tuyến mà còn tự thu thập 551 hình ảnh bằng cách sử dụng cả tên khoa học và tên thông thường của 5 loại dâu tây có trong tập dữ liệu.

Để tạo ra một tập dữ liệu chất lượng, một quy trình lọc kỹ lưỡng được áp dụng. Tiêu chí chọn lọc bao gồm thông tin siêu dữ liệu trên trang web và các nguyên tắc được đề

ra bởi Bộ Nông nghiệp. Màu sắc, diện tích, mật độ của vùng bị nhiễm bệnh và hình dạng của từng loại được xác định là các yếu tố quan trọng nhất để phân loại hình ảnh thành từng nhóm. Ngoài ra, việc loại bỏ những hình ảnh không chính xác cũng được thực hiện, ví dụ như những hình ảnh không phải là quả dâu tây do phòng thí nghiệm kiểm soát và nằm ngoài phạm vi. Đồng thời, để đảm bảo tính chính xác của từng loại, hình ảnh trùng lặp trong các lớp thông qua quá trình tìm kiếm cũng được loại bỏ.

Các hình ảnh của tập dữ liệu đều được kiểm tra bởi hai cá nhân theo hướng dẫn cụ thể, nhằm giảm thiểu khả năng xuất hiện lỗi ghi nhãn, và đảm bảo độ chính xác cũng như chất lượng của dữ liệu. Kết quả, đã thu thập được một tập dữ liệu gồm 3 class, trong đó 2 class về bệnh của quả dâu tây và 1 class về quả dâu tây không bị bệnh được mô tả trong bảng 1.

Bảng 1. Mô tả bệnh cùng với số lượng hình ảnh của mỗi loại quả dâu tây

Trình trạng	Mô tả	Số lượng mẫu	Hình ảnh
Normal	Dâu tây có màu đỏ sáng, đồng đều trên toàn bộ. Nó không có vết thương lớn hoặc tổn thương, có bề mặt mịn và không có dấu hiệu của các đốm đen hoặc sự thối rữa.	280	
Gray mold disease	Dâu tây có thể xuất hiện các vùng màu xám hoặc trắng trên bề mặt. Những vùng này có xu hướng lan rộng dần và theo thời gian, chúng tạo thành một lớp nấm mốc màu xám.	90	
Powdery mildew disease	Dâu tây có những mảng phấn màu trắng nhạt trên bề mặt. Những lớp màu trắng này giống như bột mịn, phủ lên quả, làm giảm vẻ bóng bẩy tự nhiên của quả.	90	

Sau khi xây dựng được tập dữ liệu về bệnh của quả dâu tây, chúng ta cần các vùng giới hạn chính xác chứa quả dâu tây trong ảnh đầy đủ. Do đó, đã sử dụng Roboflow để tạo các hộp giới hạn xung quanh những chiếc lá trong tất cả các hình ảnh. Trong tình huống thực tế, hình ảnh có thể có nhiều quả dâu tây hoặc sự kết hợp giữa những quả dâu tây bị bệnh và khỏe mạnh. Sau đó,

thực hiện gắn nhãn rõ ràng cho tất cả các lá trong ảnh bằng các lớp cụ thể của chúng. Trong khi dán nhãn cho các hộp, đảm bảo rằng toàn bộ đầu tây phải ở bên trong hộp và diện tích của hộp bao quanh không được nhỏ hơn 1/8 (xấp xỉ) kích thước hình ảnh.

Kết quả là một tập dữ liệu mới có tên STRAWBERRY dataset đã được xây dựng. Nó chứa 460 hình ảnh và 2260 trường hợp, trong đó 80% (368 hình ảnh) đã được chọn ngẫu nhiên làm tập dữ liệu huấn luyện, 10% (46 hình ảnh) làm tập dữ liệu huấn luyện tập dữ liệu xác thực và 10% còn lại (46 hình ảnh) làm tập dữ liệu thử nghiệm. Tập dữ liệu thử nghiệm chỉ được dùng cho đánh giá hiệu suất của mô hình sau khi huấn luyện, như thể hiện trong bảng 2.

Bảng 2. Số lượng hình ảnh đã ghi chú cho mỗi loại quả dâu tây

STRAWBERRY dataset	Normal	Gray mold disease	Black spot disease	Số lượng mẫu
Train	914	220	164	368
Test	204	80	78	46
Valid	123	60	52	46

3. KẾT QUẢ NGHIÊN CỨU

3.1. Môi trường thí nghiệm

Mô hình đề xuất trong nghiên cứu này được huấn luyện bằng tập dữ liệu tự xây dựng, được mô tả ở trên bằng cách sử dụng Google Colab với cấu hình High RAM Colab Runtime và GPU Tesla V100. Sau khi quá trình huấn luyện hoàn thành, thu được các bộ trọng số tương ứng cho mỗi mô hình. Tiếp theo, đánh giá hiệu quả của mỗi mô hình dựa trên tập dữ liệu kiểm thử. Cuối cùng, so sánh các kết quả thu được giữa các phiên bản YOLOv8l, YOLOv8m, YOLOv8, YOLOv8x.

3.2. Các chỉ số đánh giá hiệu suất

Để đánh giá hiệu quả các phiên bản của mô hình YOLOv8 đối với phát hiện bệnh của quả dâu tây, các chỉ số đánh giá được sử dụng bao gồm GFLOPS (Giga Floating-point Operations Per Second), Precision, Recall và MAP (Mean Average Precision).

GFLOPS là số lượng tỷ lệ các phép tính điểm dấu chấm động trên giây, thường được dùng như một tham số hiệu năng GPU và có thể được quan sát thông qua GFLOPs. Kích thước tham số của mô hình có thể được sử dụng để xác định độ phức tạp của mô hình bằng cách kiểm tra các tham số. Trong việc cải tiến mô hình, đôi khi GFLOPs và các tham số tăng lên không tránh khỏi. Nhìn chung, chúng ta mong muốn GFLOPs và các tham số được càng nhỏ càng tốt.

Tỷ lệ chính xác được xác định bằng phương trình bên dưới. Nó được xác định là tỷ lệ giữa số lượng mẫu tích cực

được dự đoán chính xác bởi mô hình đến tổng số mẫu tích cực được dự đoán:

$$\text{precision} = \frac{TP}{TP+FP} \tag{1}$$

Recall đại diện cho số lượng mẫu tích cực được dự đoán chính xác bởi mô hình là một phần trăm của tất cả các mục tiêu. Công thức tính tỷ lệ recall được hiển thị trong phương trình:

$$\text{recall} = \frac{TP}{TP + FN} \tag{2}$$

Đường cong precision-recall đánh giá độ chính xác và khả năng nhận diện của mô hình, trong đó khu vực dưới đường này (AP) thường được sử dụng để đánh giá hiệu suất, với giá trị càng cao cho thấy mô hình càng chính xác và có khả năng phát hiện tốt hơn.

$$AP = \int_0^1 P(R)dR \tag{3}$$

mAP là độ chính xác trung bình trên tất cả các lớp. Đối với toàn bộ mô hình, giá trị mAP càng cao thì hiệu quả phát hiện tổng thể của mô hình càng tốt và độ chính xác phát hiện càng cao.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \tag{4}$$

3.3. Kết quả thử nghiệm và thảo luận

Kết quả thử nghiệm dựa trên dữ liệu bảng 3 nhằm mục đích so sánh kết quả và đánh giá hiệu suất của 4 mô hình là YOLOv5, YOLOv6, YOLOv7 và YOLOv8. Bảng bao gồm dữ liệu về độ chính xác cho 3 lớp: Normal, Gray mold disease, Black spot disease.

Bảng 3. Kết quả trên tập dữ liệu kiểm tra

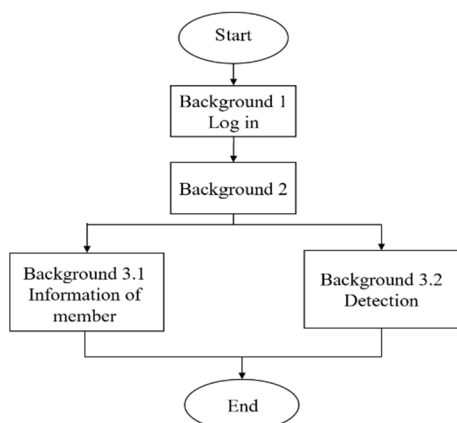
Model	Precision	Recall	AP			mAP
			Normal	Gray mold disease	Powdery mildew disease	
YOLOv5	0,709	0,699	0,769	0,763	0,757	0,763
YOLOv6	0,668	0,728	0,779	0,766	0,759	0,768
YOLOv7	0,719	0,725	0,772	0,757	0,745	0,758
YOLOv8	0,750	0,745	0,790	0,772	0,774	0,779

Dựa trên kết quả bảng 3, mô hình YOLOv8 đã đạt độ chính xác cao nhất là 77,9%, vượt trội so với các mô hình khác như YOLOv5 (76,3%), YOLOv6 (76,8%) và YOLOv7 (75,8%). Mặc dù các phiên bản của YOLO khi kết hợp các mô hình để cải thiện độ chính xác, chúng thường phức tạp hơn, xử lý chậm hơn và đòi hỏi nhiều tính toán hơn.

Tuy nhiên, YOLOv8 có lợi thế với số lượng tham số nhỏ hơn so với các mô hình khác, nhưng vẫn duy trì hiệu suất cao. Tham số (Parameters) bao gồm trọng số (weights) và độ lệch (biases) cần được học trong quá trình huấn luyện. Mặc dù số lượng tham số lớn có thể giúp mô hình học được các đặc trưng phức tạp hơn, nó cũng đòi hỏi nhiều tài nguyên và thời gian huấn luyện hơn. Dù YOLOv7 có số lượng tham số cao nhất nhưng lại có độ chính xác thấp nhất, điều này có thể do overfitting, underfitting, vấn đề tối ưu hóa, thiếu dữ liệu, hoặc thiếu các kỹ thuật regularization. Hơn nữa, cấu trúc mô hình cũng rất quan trọng, một mô hình thiết kế tốt với ít tham số có thể hoạt động hiệu quả hơn một mô hình phức tạp nhưng không tối ưu. Điều này cho thấy YOLOv8 đã đạt được sự cân bằng tốt giữa số lượng tham số và hiệu suất. Đặc biệt, YOLOv8 duy trì độ ổn định và đạt độ chính xác cao nhất ở mỗi lớp trong 4 mô hình, với các điểm số tương ứng là: 79,0% cho "Normal", 77,2% cho "Gray mold disease", 77,4% cho "Powdery mildew disease". Ngoài ra ở các tham số như Precision và Recall, YOLOv8 cũng vượt trội hơn các model phiên bản trước đó lần lượt là 75,0% và 74,5%. Những điều này chứng tỏ YOLOv8 có hiệu suất chính xác cao nhất trong số các phiên bản YOLO còn lại.

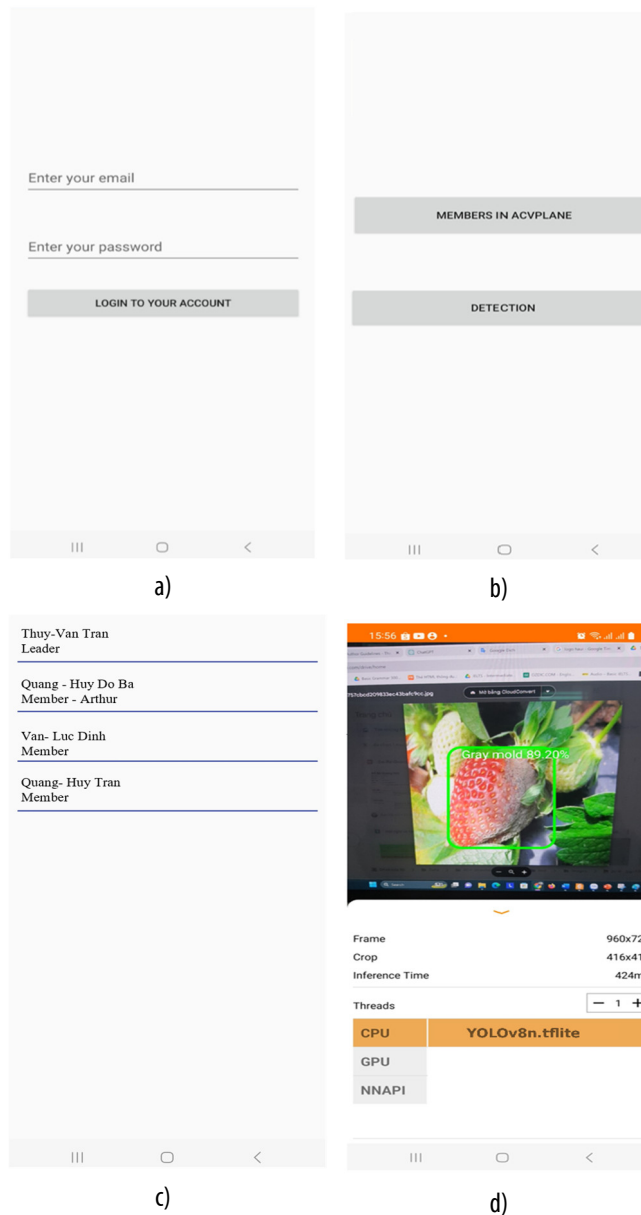
3.4. Kết quả của ứng dụng

Sau khi huấn luyện xong mô hình, thấy rằng mô hình YOLOv8 có hiệu suất cao nhất so với các phiên bản còn lại. Do đó, file.yaml của YOLOv8 được sử dụng để tiếp tục phát triển ứng dụng nhận diện bệnh cho quả dâu tây. Điều này đặt ra một cơ hội mới và quan trọng trong nghiên cứu và ứng dụng công nghệ để hỗ trợ ngành nông nghiệp và giám sát sức khỏe của cây trồng. Sự kết hợp giữa độ chính xác của mô hình YOLOv8 và khả năng ứng dụng thực tế sẽ mang lại những đóng góp tích cực cho cộng đồng nông dân và người làm nghiên cứu trong lĩnh vực này.



Hình 4. Sơ đồ khối

Cấu trúc của ứng dụng theo sơ đồ khối đã được xây dựng như hình 5.



Hình 5. Hình ảnh kết quả của ứng dụng

a) Background 1; b) Background 2; c) Background 3.1; d) Background 3.2

4. KẾT LUẬN

Trong bài báo này, nhóm tác giả đã xây dựng và thiết kế một ứng dụng tích hợp được YOLOv8 để phát hiện bệnh của quả dâu tây. Sử dụng bộ dữ liệu tự tạo gồm 551 hình ảnh RGB của quả dâu tây (bao gồm cả quả bị bệnh và không bị bệnh), kết hợp với thuật toán YOLOv8 với tỷ lệ chính xác lên tới 77,9% với 4 loại bệnh phổ biến. Thông qua các quan sát được đề cập trước đó, có thể chứng minh được độ chính xác vượt trội của mô hình YOLOv8 trong việc phát hiện lỗi của quả dâu tây. Tuy nhiên kết quả vẫn chưa đạt được độ chính xác trên 80%, do trong

khi tạo tập dữ liệu, các biểu hiện trên hai loại bệnh “Gray mold disease” và “Powdery mildew disease” là rất giống nhau gây ra khó khăn trong việc nhận dạng. Do đó, cần tiếp tục nghiên cứu và xây dựng tập dữ liệu mở rộng hơn trong tương lai. Ngoài ra, kết quả nghiên cứu này có thể được phát triển và ứng dụng thực tế với các loại hạt và trái cây khác trong nông nghiệp.

TÀI LIỆU THAM KHẢO

- [1]. Hernández-Martínez, Nelda R., et al., “Current state and future perspectives of commercial strawberry production: A review,” *Scientia Horticulturae*, 312: 111893, 2023.
- [2]. Skrovankova Sona, et al., “Bioactive compounds and antioxidant activity in different types of berries,” *International Journal of Molecular Sciences*, 16.10: 24673-24706, 2015.
- [3]. Tylewicz Urszula, et al., “Chemical and physicochemical properties of semi-dried organic strawberries enriched with bilberry juice-based solution,” *Lwt*, 114: 108377, 2019.
- [4]. Li Yanfen, et al., “Crop pest recognition in natural scenes using convolutional neural networks,” *Computers and Electronics in Agriculture*, 169: 105174, 2020.
- [5]. Pan Leiqing, et al., “Early detection and classification of pathogenic fungal disease in post-harvest strawberry fruit by electronic nose and gas chromatography-mass spectrometry,” *Food Research International*, 62: 162-168, 2014.
- [6]. Maas J. L., “Strawberry diseases and pests-progress and problems,” In *VII International Strawberry Symposium 1049*, 133-142, 2012.
- [7]. Paulus Albert O., “Fungal diseases of strawberry,” *HortScience*, 25.8: 885-889, 1990.
- [8]. Chung P.C., et al., “First report of anthracnose crown rot of strawberry caused by *Colletotrichum siamense* in Taiwan,” *Plant disease*, 103.7: 1775, 2019.
- [9]. Chen X. Y., et al., “Genetic diversity of *Colletotrichum* spp. causing strawberry anthracnose in Zhejiang, China,” *Plant Disease*, 104.5: 1351-1357, 2020.
- [10]. Feliziani Erica, Romanazzi Gianfranco, “Postharvest decay of strawberry fruit: Etiology, epidemiology, and disease management,” *Journal of Berry Research*, 6.1: 47-63, 2016.
- [11]. Petrasch Stefan, et al., “Grey mould of strawberry, a devastating disease caused by the ubiquitous necrotrophic fungal pathogen *Botrytis cinerea*,” *Molecular plant pathology*, 20.6: 877-892, 2019.
- [12]. Chamorro M., Aguado A., De Los Santos B., “First report of root and crown rot caused by *Pestalotiopsis clavispora* (*Neopestalotiopsis clavispora*) on strawberry in Spain,” *Plant Dis*, 100.7: 1495, 2016.
- [13]. Amsalem Liat, et al., “Effect of climatic factors on powdery mildew caused by *Sphaerotheca macularis* f. sp. *fragariae* on strawberry,” *European journal of plant pathology*, 114: 283-292, 2006.
- [14]. Rebollar-Alviter Angel, et al., “An emerging strawberry fungal disease associated with root rot, crown rot and leaf spot caused by *Neopestalotiopsis rosae* in Mexico,” *Plant Disease*, 104.8: 2054-2059, 2020.
- [15]. Mahmud Md Sultan, et al., “Development of an artificial cloud lighting condition system using machine vision for strawberry powdery mildew disease detection,” *Computers and electronics in agriculture*, 158: 219-225, 2019.
- [16]. Jayawardena R. S., et al., “An account of *Colletotrichum* species associated with strawberry anthracnose in China based on morphology and molecular data,” *Mycosphere*, 7(8) 1147-1163, 2016.
- [17]. Ferentinos Konstantinos P., “Deep learning models for plant disease detection and diagnosis,” *Computers and electronics in agriculture*, 145: 311-318, 2018.
- [18]. Cervantes-Jilaja Claudia, et al., “Optimal Selection and Identification of Defects in Chestnuts Processing, through Computer Vision, Taking Advantage of its Inherent Characteristics,” In *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, IEEE, 513-520, 2019.
- [19]. Cheng Xi, et al., “Pest identification via deep residual learning in complex background,” *Computers and Electronics in Agriculture*, 141: 351-356, 2017.
- [20]. Diao Zhihua, et al., “Navigation line extraction algorithm for corn spraying robot based on improved YOLOv8s network,” *Computers and Electronics in Agriculture*, 212: 108049, 2023.

AUTHORS INFORMATION

Do Ba Quang Huy, Tran Thuy Van, Dinh Van Luc, Tran Quang Huy

Faculty of Electrical Engineering, Hanoi University of Industry, Vietnam