

# Xác định độ chín của trái Dứa hỗ trợ quá trình phân loại, bảo quản trái chín sau thu hoạch

Trịnh Trung Hải<sup>1</sup>, Bùi Xuân Thiện<sup>2</sup>, Nguyễn Hà Huy Cường<sup>3</sup>, Ninh Khánh Duy<sup>4</sup>

<sup>1,2,3</sup>Trường Đại học Công nghệ Thông tin và Truyền thông Việt - Hàn

<sup>1</sup>Email: [tthai@vku.udn.vn](mailto:tthai@vku.udn.vn), <sup>2</sup>Email: [bxthien.18it3@vku.udn.vn](mailto:bxthien.18it3@vku.udn.vn),

<sup>3</sup>Email: [nhhcuong@vku.udn.vn](mailto:nhhcuong@vku.udn.vn)

<sup>4</sup>Trường Đại học Bách khoa – Đại học Đà Nẵng

<sup>4</sup>Email: [nkduy@dut.udn.vn](mailto:nkduy@dut.udn.vn)

## Tóm tắt.

**Tóm tắt:** Trái cây sau khi thu hoạch cần được phân loại, bảo quản đúng cách là khâu quyết định chất lượng và giá trị sản phẩm trước khi phân phối ra thị trường. Do đó cần tăng cường áp dụng giải pháp khoa học kỹ thuật vào các khâu này sẽ góp phần hỗ trợ người nông dân đạt hiệu quả kinh tế cao đối với đầu ra của trái cây Việt nói chung và trái Dứa trồng tại Việt Nam nói riêng. Trong nghiên cứu này, chúng tôi xin trình bày giải pháp xác định độ chín của trái Dứa tại các vựa Dứa sau khi thu hoạch bằng kỹ thuật nhận dạng theo mô hình YOLO-v5, đồng thời đề xuất giải pháp tăng cường dữ liệu (data augmentation). Tăng cường số lượng dữ liệu huấn luyện và tăng số lớp trên mô hình YOLO-v5; với mỗi lớp chứa bộ dữ liệu gồm 10.000 hình ảnh sẽ cải thiện độ chính xác trong nhận dạng và tăng tốc độ huấn luyện khi sử dụng mô hình YOLO-v5. Chúng tôi đã thực hiện thử nghiệm trên bộ dữ liệu gồm 50.000 hình ảnh trái Dứa chín cho kết quả phân loại có độ nhạy chính xác là 94,5%. Từ đó phát triển xây dựng hệ thống tự động xử lý phân loại trái Dứa chín sau thu hoạch.

**Từ khóa:** Xác định độ chín trái cây, nhận dạng trái Dứa chín, phân loại trái Dứa chín; tăng cường dữ liệu nhận dạng, mô hình YOLO-v5 cải tiến.

The fruits after harvesting need to be classified, stored properly which is the determination in quality and fresh before distributing to the market. Therefore, it is necessary to increase the application of scientific and technical solutions, which will contribute to supporting farmers to achieve high economic efficiency for the first part of Vietnamese fruits in general and pineapple growing in Vietnam in particular. In this study, we would like to present a solution to determine the ripeness of pineapple fruits after harvesting by identification technique according to YOLO-v5 model, and the data augmentation solutions as well. Enhance training data and increase the number of classes on the YOLO-v5 model; with each layer contains data of 10,000 images will improve the recognition accuracy and enhance training speed when using the YOLO-v5 model. We have performed a test on a dataset of 50,000 images of ripe Pineapple, which is given the accuration of 94.5%. Then, developing and building an automatic system for handling and classifying post-harvest pineapples.

Keywords: Deep Convolution Networks, YOLO-v5 improve, Identification of ripe Pineapples, classification of pineapple.

## 1. Đặt vấn đề

Ứng dụng của thành tựu khoa học kỹ thuật, trí tuệ nhân tạo cho ngành nông nghiệp chính xác, nông nghiệp thông minh đã và đang được triển khai mạnh mẽ, góp phần vào phát triển kỹ nguyên số của nhân loại và công cuộc chuyển đổi số Quốc gia. Dịch bệnh Covid 19 đã kiểm chứng thành tựu công nghệ trong thực tế, ngày càng có nhiều ứng dụng tiện ích trong lĩnh vực nông nghiệp để thay thế nông nghiệp truyền thống như: Máy kéo không người lái, máy thu hoạch tự động, hệ thống đo độ ẩm của đất phục vụ cho tưới tiêu nông sản... việc đánh giá chất lượng, tối ưu quy trình

xử lý đầu ra của nông sản Việt có vai trò quan trọng trong việc gia tăng giá trị và phát triển bền vững nền kinh tế nông nghiệp.

Trái dứa là một loại cây ăn quả nhiệt đới hiện được trồng nhiều ở các nước Nam Mỹ và một số nước nhiệt đới, cận nhiệt đới có sông ngòi nhiều như đảo Hawaii, Đài Loan, Indonesia... Việt Nam đứng thứ 10 Thế giới về sản lượng Dứa, được trồng nhiều ở các tỉnh, Tiền Giang, Kiên Giang, Thanh hóa, Nam Định và Quảng Nam; trái Dứa dễ thâm canh, cho năng suất cao có nhiều tiềm năng phát triển. Tuy nhiên hiện nay có tỷ trọng xuất khẩu chưa cao, giá bán thị trường thấp, do phần lớn các hoạt động chăm sóc, thu hoạch, phân loại, bảo quản trái Dứa đang được thực hiện thủ công, mất nhiều thời gian, chi phí, ảnh hưởng đến sản lượng và chất lượng của trái cây, làm giảm tỷ trọng nông nghiệp của nước ta.

Để nhận biết, phân loại trái Dứa chín người ta có thể đo lường các thông số đặc điểm, tính chất hóa, lý cụ thể của nó, nhưng phương pháp này đòi hỏi phải sử dụng các tính chất và phương pháp kiểm tra hóa, lý bên trong sản phẩm rất phức tạp và khó triển khai thực tế. Theo kinh nghiệm của người nông dân để nhận biết độ chín của trái Dứa và sản phẩm trái cây chín thường là lựa chọn thủ công bằng mắt thường, theo cảm tính, đặc biệt đối với số lượng lớn sẽ cho kết quả không thực sự chính xác và không nhất quán và tốn rất nhiều công sức, thời gian thực hiện. Để mang lại hiệu quả cao trong phân loại sản phẩm trái cây chín phục vụ bảo quản, phân phối và chế biến đối với Trái Dứa cần có hệ thống tự động nhận dạng chính xác độ chín của trái Dứa dựa trên các tính chất đặc trưng bề ngoài của hình ảnh trái Dứa chín như: màu sắc, kích thước, đặc điểm bề mặt trái Dứa, cuống Dứa và các đặc điểm tương đồng về đặc tính bề ngoài do tác động bởi nhiều yếu tố thời tiết, môi trường như: ánh sáng, dịch bệnh... [3, 5]. Trong nghiên cứu này, chúng tôi ứng dụng mô hình học sâu hiện đại bằng kỹ thuật YOLO-v5 để nhận dạng độ chín của trái Dứa, đồng thời đề xuất giải pháp gia tăng nguồn dữ liệu lớn hình ảnh trái Dứa chín để cải thiện độ chính xác và tăng tốc độ nhận dạng của mô hình YOLO-v5 để xác định độ chín của trái Dứa trồng tại các khu vực Đại Lộc Quảng Nam. Hình ảnh trái Dứa chín thu thập từ thực tế và kế thừa nguồn dữ liệu hình ảnh sẵn có từ các cơ sở dữ liệu mạng. Qua đó hỗ trợ phân loại, bảo quản, phân phối, xuất khẩu và chế biến sản phẩm từ trái Dứa.

Trong bài báo này chúng tôi trình bày các nội dung như sau: phần thứ nhất: giới thiệu tổng quan; phần 2: trình bày các hướng nghiên cứu liên quan; phần 3: phương pháp thu thập và tổ chức dữ liệu; phần 4: giải pháp kỹ thuật; phần 5: đánh giá các kết quả thử nghiệm, phần 6: Kết luận và cuối cùng là tài liệu tham khảo.

## 2. Các nghiên cứu liên quan

Trong những năm gần đây, chúng ta đã chứng kiến được nhiều thành tựu vượt bậc trong ngành Thị giác máy tính (Computer Visions) và những tiến bộ trong học máy (Machine Learning), học sâu (Deep Learning) để phát hiện và nhận dạng đối tượng, qua đó được nhiều nhà nghiên cứu trong nước và quốc tế quan tâm nghiên cứu và có những công bố liên quan như: Các hệ thống xử lý ảnh lớn như Facebook, Google hay Amazon đã đưa vào sản phẩm của mình có những chức năng thông minh như nhận diện khuôn mặt người dùng, phát triển xe hơi tự lái hay giao hàng tự động [1].

Trong lĩnh vực nông nghiệp công nghệ cao các nhà nghiên cứu Việt Nam, Ấn Độ, Nhật Bản, Trung Quốc.. đã nghiên cứu, ứng dụng công nghệ AI vào nông nghiệp, có thể kể ra như: Nhóm nghiên cứu Alex Krizhevsky và cộng sự tại đại học Toronto [2] đã nghiên cứu sử dụng mạng nơ ron phức hợp, tích chập sâu để phân loại hình ảnh có độ phân giải cao với kết quả thử nghiệm có độ chính xác và đào tạo nhanh hơn khi sử dụng các mạng nơ ron không bão hòa và kết hợp phương pháp “dropout”; bỏ qua dữ liệu. Các nghiên cứu của nhóm Widyawati và Kim Jun-Hwa [3], [4], [5] đã áp dụng kỹ thuật YOLO để phát hiện và nhận dạng độ chín của trái cây và cải tiến YOLO để phát hiện các đối tượng khác đạt hiệu quả rất cao sau khi thuật toán YOLO v4 và YOLO v5 ra đời. Nhóm của Đặng Thị Phương Chung và Đinh Văn Tài [6] đã đề xuất hệ thống nhận dạng trái cây bằng cách sử dụng mạng nơ-ron tích chập (CNN) và kiến trúc EfficientNet. Dựa trên các lớp tổng hợp và tích chập CNN được sử dụng để phân loại hình ảnh đầu vào 2-D và nhận dạng các đối tượng

theo mô hình ba lớp. Một số nghiên cứu của nhóm Zaw Min Khaing, nhóm Nguyễn Hà Huy Cường và các cộng sự [7], [8], [9] đã nghiên cứu tối ưu ứng dụng mạng nơ ron tích chập (CNN), đề xuất cải tiến, điều chỉnh các tham số lớp tích chập nhằm tối ưu hóa kỹ thuật phát hiện và nhận dạng trái cây hỗ trợ ra quyết định, khuyến nghị trong nông nghiệp. Nhóm nghiên cứu Horea và cộng sự [10] đã đề xuất hệ thống huấn luyện theo kỹ thuật mạng thần kinh sâu để tạo ra bộ phân loại trái cây được xác định từ hình ảnh của nhiều loại trái cây khác nhau. Nghiên cứu của Abien Fred Agarap [11] điều chỉnh hàm ReLU tuyến tính trong kỹ thuật học sâu để tăng khả năng phân loại đối tượng ảnh trong mạng Nơ ron sâu (DNN). Nhóm Susovan Jana và cộng sự. [12] đề xuất hệ thống tự động nhận dạng trái cây và rau quả sử dụng công nghệ thị giác máy tính và học máy để phân loại trái cây từ hình ảnh dựa trên nội dung của chúng và phát hiện các đối tượng từ hình ảnh tự nhiên. Ngoài ra các nghiên cứu [13] – [20] đã đề xuất các giải pháp nghiên cứu nhận dạng nông sản, thực phẩm và các hệ thống tự động nhận dạng hỗ trợ chăm sóc, thu hoạch trái cây rất đa dạng trong việc ứng dụng kỹ thuật học sâu hiện đại trong xử lý hình ảnh hỗ trợ phát triển lĩnh vực nông nghiệp thông minh, nông nghiệp chính xác.

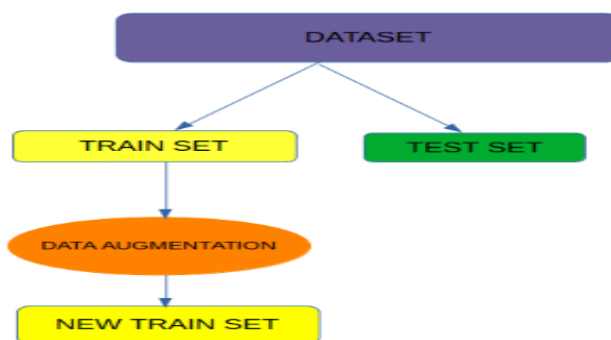
### 3. Giải pháp tổ chức dữ liệu tăng cường

Để tổ chức tập dữ liệu hình ảnh trái Dứa chín về các loại giống Dứa khác nhau, chúng tôi đã tiếp cận, thu thập hình ảnh trái Dứa từ nhiều nguồn khác nhau, sử dụng thiết bị điện tử, giải pháp IoT... quay chụp, lấy hình ảnh từ các vựa Dứa ở khu vực Đại Lộc, Quảng Nam để xử lý và gán nhãn cho hơn 5000 hình ảnh trái Dứa chín và kết hợp kế thừa nguồn dữ liệu đến 50000 hình ảnh trái Dứa từ các cơ sở dữ liệu đã có trên COCO, Kaggle [21] để xây dựng tập huấn luyện hỗ trợ.

Sử dụng công cụ Roboflow Annotate để tự động gán nhãn dữ liệu giúp đơn giản hóa quá trình xử lý ảnh thô thành mô hình thị giác máy tính để triển khai huấn luyện, sẽ được sử dụng để đặt tên cho các ảnh đã tập hợp. Với công cụ này, có thể tạo một khung hình chữ nhật xung quanh quả Dứa để nhìn thấy được, đồng thời xác định được các pixel tương ứng với mỗi trái cây. Công việc của chú thích là trích xuất các vùng đặc điểm của quả dứa dựa trên các đặc tính hình thái đặc trưng bề ngoài của quả Dứa như mắt, màu sắc, hình dạng, v.v. Đầu ra cuối cùng sẽ là một tập dữ liệu có chú thích được lưu ở định dạng Microsoft COCO từ Roboflow Annotate.



**Hình 1.** Một số hình ảnh mẫu từ tập dữ liệu

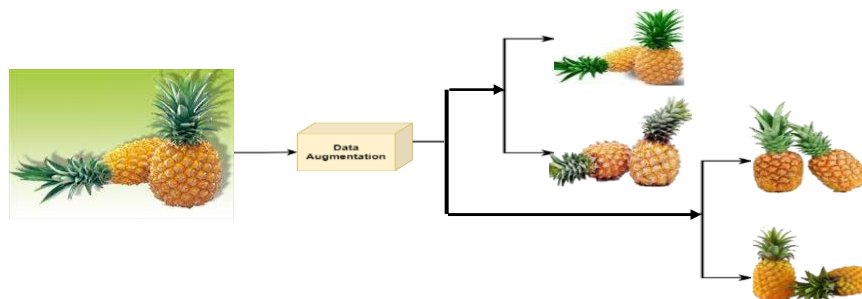


**Hình 2.** Mô hình cơ bản chung áp dụng gia tăng dữ liệu huấn luyện

Sử dụng kỹ thuật tăng cường dữ liệu là một trong những phương pháp hiệu quả nhất để cải thiện tăng cường độ chính xác cho quá trình huấn luyện một mô hình nhận dạng đối tượng ảnh số từ (Data\_Train\_Set → Data Augmentation → New\_Data\_Train\_Set) gồm các phép biến đổi:

- (1) Các phép biến đổi hình học ( Geometry based ) – lật, cắt, xoay hoặc dịch hình ảnh một cách ngẫu nhiên

- (2) Chuyển đổi không gian màu ( Color based) – thay đổi các kênh màu RGB, tăng hoặc giảm bất kỳ màu nào, tăng độ sắc nét, tăng tính tương phản, .v.v
- (3) Thêm nhiễu ( Noise/occlusion ) – thêm nhiễu cho ảnh như nhiễu ngẫu nhiên, nhiễu có mẫu, nhiễu do nén ảnh, .v.v
- (4) Xóa ngẫu nhiên (Random crop) – xóa ngẫu nhiên một phần của hình ảnh ban đầu; Biến đổi thời tiết (Weather) thêm tác dụng của thời tiết như mưa, tuyết, sương mờ, ...

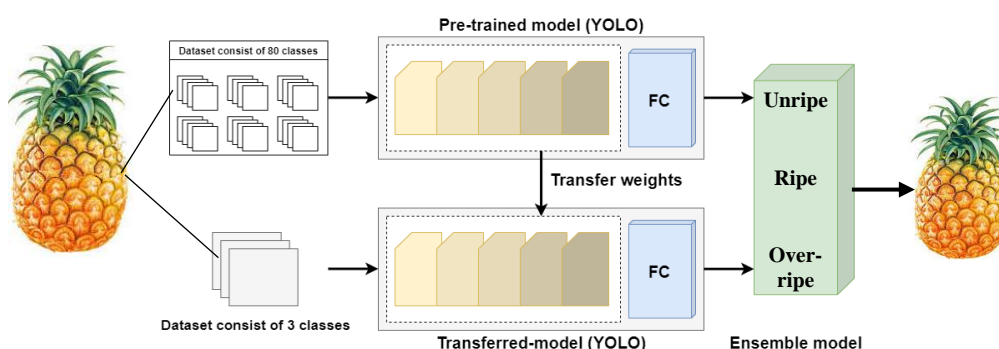


**Hình 3.** Tăng cường và biến đổi hình ảnh quả Dứa

#### 4. Đề xuất cải tiến mô hình

Các nghiên cứu trước đây đã chỉ ra rằng mô hình được huấn luyện trước của YOLO đã mang lại kết quả khả quan trong các kịch bản thử nghiệm như VGG, Snapchat v.v. Trong bài báo này, kiến trúc Resnet được sử dụng làm nền tảng cơ bản để phân tích, phân dạng hình ảnh trái Dứa từ tập dữ liệu hình ảnh đã thu thập được. Với kỹ thuật YOLO-v5 đã nghiên cứu trong thời gian gần đây đã góp phần tăng tốc độ huấn luyện ảnh mẫu cho trước, giảm độ mất mát trong huấn luyện và có độ chính xác khá cao. Đã tiến hành thử nghiệm với kiến trúc Resnet100 cho nhiệm vụ phân loại các hình ảnh huấn luyện thu được.

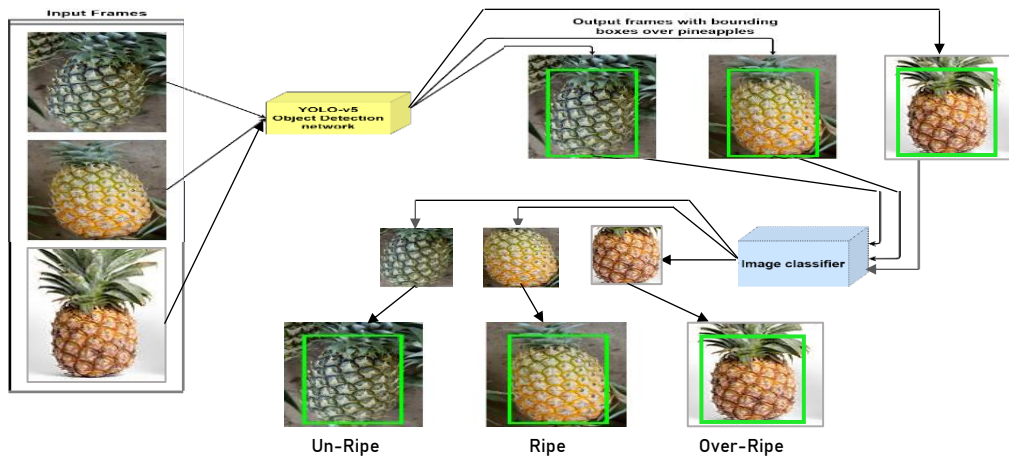
Việc phát hiện Dứa chín trong thời gian thực được sử dụng kỹ thuật nhận dạng theo mô hình YOLO-v5, có thể xử lý dữ liệu đầu vào, sắp xếp và huấn luyện các tập dữ liệu; xây dựng quy trình phân loại mức độ dựa trên kiến trúc CNN resnet100 và phân loại quả Dứa chín hay chưa chín. Trong một kịch bản thực tế, khi thực hiện thí nghiệm tại các vườn trồng Dứa ở Việt Nam, tập dữ liệu đầu vào là hình ảnh Dứa đã được thu thập, xử lý tăng cường và gửi đến mạng YOLO-v5 cải tiến để huấn luyện xác định độ chín của trái Dứa.



**Hình 5.** Mô hình gia tăng dữ liệu để phát hiện trái Dứa chín với YOLO-v5

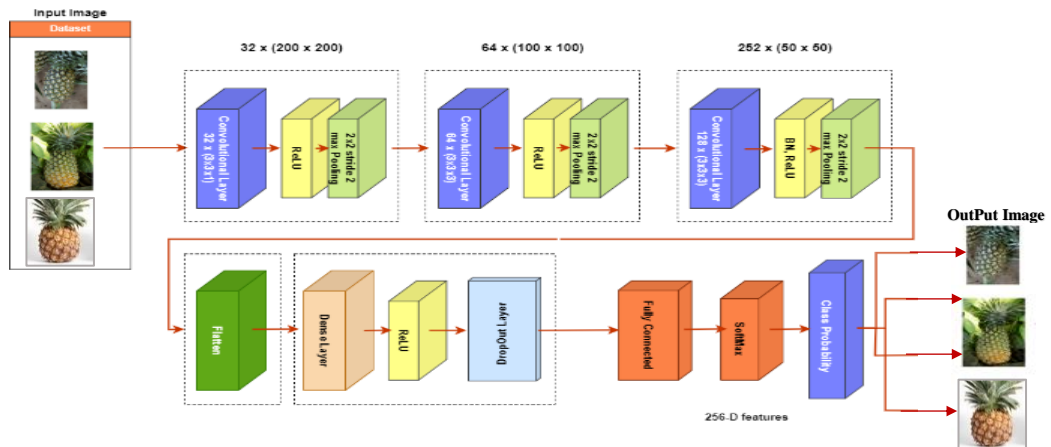
YOLO-v5 sử dụng lưới ma trận ô vuông (SxS) để thu thập thông tin dữ liệu và tự động xác định tọa độ tâm của đối tượng trong vùng thể hiện ma trận lưới. Tâm của đối tượng cần xác định nằm trong ô vuông nào thì ô vuông đó chứa đối tượng đó. Bằng cách tính toán các giá trị hàng và cột, các đối tượng đã kiểm tra trong YOLO-v5, chọn từng đối tượng được đánh dấu riêng biệt để loại bỏ nó. Nếu nhiều hộp giới hạn phát hiện cùng một đối tượng, giải pháp chọn hộp giới hạn thích hợp nhất. Cũng như các hộp giới hạn cho mỗi lớp, chúng tôi tính xác suất tìm thấy các loại đối tượng khác nhau. Đặc điểm của YOLO-v5 là sử dụng kỹ thuật hồi quy xử lý hộp giới hạn để xác định xác suất lớp và tăng tốc độ tính toán để đạt được kết quả kiểm tra cao. Các phiên bản mới của YOLO-v5 đã được công bố [6], Mô hình phát hiện một đối tượng là quả Dứa chín khi màu da chuyển sang

màu vàng bóng với xác suất là 0,9868. Các kết quả sau đó được thử nghiệm và tiếp tục sau đó cho kết quả khả quan.



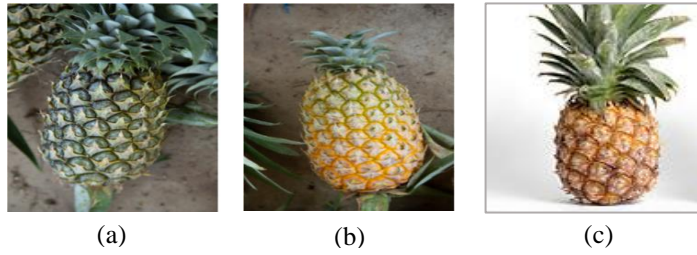
**Hình 5.** Mô hình YOLO-v5 phát hiện độ chín của trái Dứa

Bài báo là kế thừa kỹ thuật ưu việt của mô hình YOLO-v5 được huấn luyện trước để đánh dấu, phát hiện đối tượng trái Dứa, đồng thời sử dụng phương pháp gia tăng dữ liệu đưa vào huấn luyện, thử nghiệm bằng cách nhân bản thêm nhiều hình ảnh từ nguồn dữ liệu hình ảnh đã thu thập, tăng thêm nhiều phép biến đổi dữ liệu theo góc nhìn, độ sáng, kích thước, trạng thái ... của dữ liệu ảnh số, bao gồm phóng to, chiếu sáng, lật, xoay và làm cong.... Mô hình đề xuất phát triển sử dụng ba lớp phức hợp và hai lớp kết nối để phát hiện và quyết định, gồm mô hình được huấn luyện trước, mô hình chuyển giao và mô hình tổng hợp. Đối với một hình ảnh thực, mô hình tạo ra bốn hình ảnh ngẫu nhiên được tự động nhân bản. Sử dụng các tính năng được trích xuất từ hình ảnh đầu vào, ConvNet tạo ra các bản đồ tính năng ở đầu ra. Một bộ lọc  $16 \times 16$  và bốn lớp phức hợp được đề xuất. Các khung hình hợp lệ có kích thước  $9 \times 9$  và  $25 \times 25$  vượt trội hơn để tăng cường phân lớp nhận dạng chi tiết độ chín của trái Dứa như: chưa chín, chín và quá chín.



**Hình 6.** Kiến trúc khung huấn luyện đề xuất để phân loại trái Dứa

Một phần trong đề xuất của chúng tôi về học sâu là cài đặt hàm ReLU và hàm Softmax để kích hoạt tăng cường các lớp ẩn khi gia tăng dữ liệu huấn luyện trong mạng CNN với mô hình YOLO-v5, sử dụng kích hoạt phi tuyến tính và hỗ trợ tính toán hồi quy cho phép cải tiến mô hình huấn luyện được nhanh hơn. Trong mạng đề xuất đã cài đặt một mô đun tích chập tạo ra các vector đặc trưng với 128 chiều. Các lớp hợp thành kết nối, phân loại tuyến tính được xây dựng lồng ghép trong các lớp mạng.



Hình 7. (a) Dứa chưa chín, (b) Dứa chín vừa, (c) Dứa chín quá

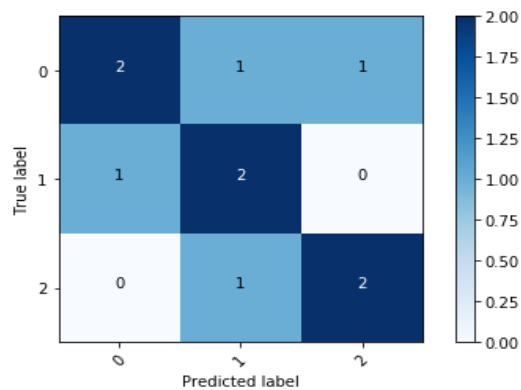
## 5. Đánh giá kết quả

Bộ sưu tập cơ sở dữ liệu này sẽ gặp phải nhiều khó khăn khi sử dụng nguồn dữ liệu huấn luyện để so sánh nhận dạng với các hình ảnh thực tế có nhiều sự thay đổi do chịu tác động của môi trường như: ánh sáng, sâu bệnh, sự thay đổi sắc tố vỏ trái cây sau thời gian thu hoạch, môi trường lưu trữ, bảo quản sản phẩm .... Do đó yêu cầu phải sử dụng kỹ thuật tiền xử lý ảnh để giảm bớt các biến thể màu, ánh sáng, tách hình ảnh khỏi nền, khử độ sáng, độ nhiễu màu và các tính chất khác của ảnh, xử lý các kịch bản soi hình ảnh trái cây trong không gian thực tế, trong siêu thị, cửa hàng trái cây... Tất cả dữ liệu được lưu trữ trong không gian màu RGB, 8 bit trên mỗi kênh. Các hình ảnh được chụp vào các thời điểm khác nhau trong ngày cho cùng một danh mục. Do đó, tập dữ liệu cũng phải gồm nhiều kịch bản là thực tế hơn, các hình ảnh rất đa dạng về chất lượng và ánh sáng. Chúng tôi đã thu thập dữ liệu trái cây trong một môi trường tương đối không hạn chế.. Bảng 1 cho thấy số lượng mẫu đã sử dụng để huấn luyện trên mô hình đề xuất của mình.

Bảng 1. Danh sách các hình ảnh cho các loại trái cây khác nhau

| STT   | Mức độ chín của trái Dứa | Số lượng hình ảnh |
|---|--------------------------|-------------------|
| 1   | Dứa chưa chín            | 10.000            |
| 2   | Dứa chín                 | 20.000            |
| 3   | Dứa chín quá             | 20.000            |
| Tổng số hình ảnh cho 3 nhóm dữ liệu là 50.000 ảnh |                          |                   |

Chúng ta đều biết rằng ma trận nhầm lẫn là một ma trận  $M \times N$  được sử dụng để phân tích hiệu suất của một mô hình nhận dạng, trong đó  $N$  đại diện cho số lượng các lớp đầu ra. Sử dụng phương pháp này, chúng ta có thể có được một bức tranh tổng thể về khung phân loại này đang hoạt động tốt như thế nào và những loại lỗi mà nó đang mắc phải. Ma trận so sánh các giá trị mục tiêu thực tế với các giá trị dự đoán được xác định bởi khung đào tạo sâu. Ma trận có trục  $X$  đại diện cho nhãn trái cây đề xuất, trong khi trục  $Y$  đại diện cho nhãn trái cây thực tế, trong khi các giá trị đường chéo của ma trận độ không tăng, điều đó thể hiện tỷ lệ dự đoán chính xác từ mô hình đề xuất. Việc thu thập thông tin trong Bảng 3 làm rõ lý do của những mất mát, tỷ lệ chu kỳ chuẩn bị với tốc độ học tập là 0,0005. Độ chính xác mà thí nghiệm đạt được là 97%.



Hình 8. Mô hình cơ sở sử dụng ma trận nhầm lẫn trong phân lớp đánh giá 3 nhóm dữ liệu



Trong những năm trở lại đây đã có nhiều nghiên cứu về việc sử dụng mạng CNN hiện đại để triển khai nhận dạng linh hoạt các kiến trúc: R-CNN, R-FCN, Fast R-CNN, Faster R-CNN, SSD trong Tensorflow đã sử dụng để thử nghiệm theo dõi accuracy/speed qua feature extractor, image resolution, etc. Với mô hình Faster RCNN khi dùng ít proposal hơn có thể tăng đáng kể tốc độ mà không làm thay đổi quá nhiều tới độ chính xác giúp nó có thể cạnh tranh với R-FCN và SSD. Tuy nhiên đối với nhiều ứng dụng đòi hỏi thực hiện theo thời gian thực cho từng đối tượng (trái Dứa chín) riêng lẻ đối với hình ảnh cận cảnh để xác định phân loại đối tượng tức thời, do đó nhóm đề xuất sử dụng mô hình YOLO-v5 với những ưu điểm về tốc độ nhận dạng phù hợp [20].

Để đánh giá mô hình trên, chúng tôi so sánh kết quả từ mô hình YOLO-v5 trên tập xác nhận đã thu thập từ trước với chỉ số IoU-Intersection Over Union đo độ chính xác của nhận dạng đối tượng được xác định bằng cách chia số pixel trong giao điểm cho toàn bộ pixel.

$$IoU = \frac{\text{Area of Overlab}}{\text{Area of Union}} ; IoU = \frac{TP}{(TP + FP + FN)}$$

Độ chính xác trung bình (AP) là diện tích của biểu đồ bao gồm Precisions, Recall và mAP là diện tích trung bình của toàn bộ biểu đồ; trong đó đại lượng Precisions, Recall và F1 được sử dụng để đánh giá hiệu suất mô hình và được tính theo các công thức bên dưới:

$$Precision = \frac{TP}{TP + FP} ; Recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Trong đó: F1 là đại lượng đánh giá độ chính xác trung bình điều hòa giữa precision và recall; Precision phản ánh mức độ chuẩn xác của mô hình dự đoán đúng (nhóm positive); tỷ lệ precision càng cao thì kết quả dự đoán của mô hình càng tốt trong việc phát hiện trái Dứa chín. Recall: đo lường tỷ lệ dự báo chính xác các trường hợp positive trên toàn bộ các dữ liệu mẫu Dứa chín (positive). Để tính được recall thì chúng ta phải biết trước nhãn của dữ liệu. Do đó recall có thể được dùng để đánh giá trên tập train và validation vì chúng ta đã biết trước nhãn. Trên tập test khi dữ liệu được coi như mới hoàn toàn và chưa biết nhãn thì chúng ta sẽ sử dụng precision.

Các chỉ số TP, TN đại diện cho giá trị dự đoán chính xác, FP, FN đại diện cho giá trị dự đoán sai lệch. Cụ thể:

TP: Giá trị dự đoán chính xác; khi mô hình dự đoán đúng (lớp Positive) cho trái Dứa chín

FP: Giá trị dự đoán sai lệch; khi mô hình dự đoán đúng (lớp Positive) cho trái Dứa chín.

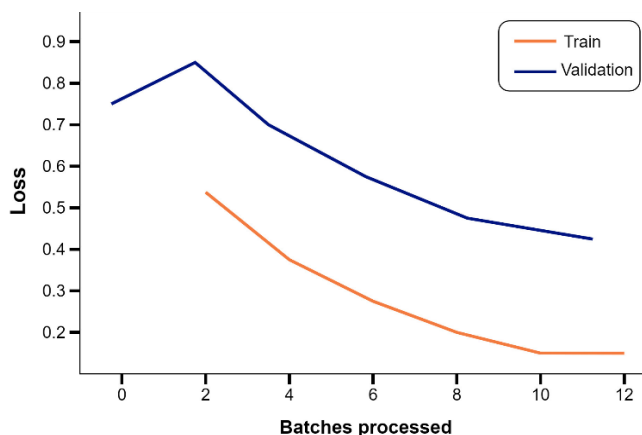
TN: Giá trị dự đoán đúng; khi mô hình dự đoán sai (lớp Negative) cho trái Dứa chín.

FN: Giá trị dự đoán sai lệch ; khi mô hình dự đoán sai (lớp Negative) cho trái Dứa chín.

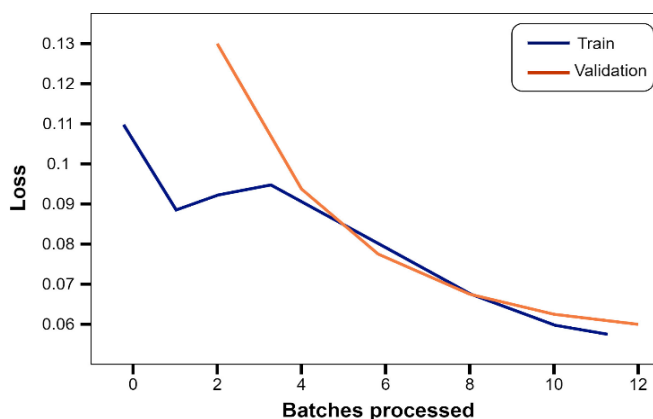
Bảng 2. Tổng hợp tỷ lệ mất mát và độ chính xác trong phân loại

| Kích thước ảnh | Lô dữ liệu | Tỷ lệ huấn luyện | Tỷ lệ xác nhận | Độ chính xác |
|----------------|------------|------------------|----------------|--------------|
| 1366 x 768     | 1          | 0.87             | 0.16           | 0.9          |
| 1360 x 768     | 2          | 0.92             | 0.13           | 0.94         |
| 1280 x 768     | 3          | 0.90             | 0.12           | 0.93         |
| 1280 x 720     | 4          | 0.88             | 0.16           | 0.95         |
| 1280 x 600     | 5          | 0.87             | 0.13           | 0.93         |
| 1024 x 768     | 6          | 0.8              | 0.16           | 0.93         |
| 800 x 600      | 7          | 0.85             | 0.11           | 0.96         |
| 1280 x 600     | 8          | 0.87             | 0.13           | 0.95         |
| 1024 x 768     | 9          | 0.92             | 0.18           | 0.93         |
| 800 x 600      | 10         | 0.93             | 0.15           | 0.94         |
| 1280 x 768     | 11         | 0.90             | 0.14           | 0.93         |
| 1280 x 720     | 12         | 0.88             | 0.16           | 0.95         |

Khi đánh giá kết quả thử nghiệm với kịch bản kết quả là giá trị tiên đoán và độ nhạy kiểm tra lại bằng 0,967. Hình 9 kết quả đánh giá hiệu suất các giai đoạn huấn luyện và xác nhận các lớp được xử lý trong mô hình cơ sở.



**Hình 9.** Kết quả đánh giá hiệu suất cho các giai đoạn huấn luyện và xác thực trên mô hình cơ sở



**Hình 10.** Kết quả đánh giá hiệu suất cho các giai đoạn huấn luyện và xác thực trên mô hình tinh chỉnh

Cuối cùng là tiến hành đánh giá kết quả nhận dạng để đưa ra quyết định: trái Dứa chín, chưa chín hay quá chín. Trong bài báo, các kỹ thuật tinh chỉnh các lớp ẩn và giải phóng các nút trong lớp mạng được áp dụng cho tất cả các lớp của bộ dữ liệu dữ liệu hình ảnh đầu vào được đào tạo từ mô hình YOLO-v5 cải tiến. Để duy trì kết quả tính toán qua các bước được đào tạo, tập hợp các lớp điển hình của mô hình cơ sở được đào tạo. Các thí nghiệm được tiến hành để trích xuất một số dữ liệu được chọn làm tập hợp con so với các lớp được chia ban đầu. Thử nghiệm vẫn giữ tốc độ học là 0,00005 cho các lớp đầu tiên và 0,00046 cho các lớp cuối cùng. Sau 20 phút huấn luyện, độ chính xác đạt được là 94,2%.

## 6. Kết luận

Trong bài báo này, chúng tôi đã có được kết quả thử nghiệm cải thiện độ chính xác và tăng tốc độ trong nhận dạng độ chín của trái Dứa thông qua giải pháp tăng cường dữ liệu hình ảnh nhóm Dứa để đưa vào mô hình huấn luyện YOLO-v5 nhằm xác định độ chín của trái Dứa tại các vựa Dứa sau khi thu hoạch với kết quả thử nghiệm chính xác khoảng 94,2%. Trên cơ sở xác định độ chín chính xác của từng trái Dứa theo mô hình huấn luyện đã hoàn thiện, chúng tôi đã phát triển một hệ thống có khả năng phân tích và đánh giá hình ảnh đầu vào để đưa ra các chỉ số về độ chín của trái Dứa qua ảnh để hỗ trợ người dân trong khâu phân loại, bảo quản và phân phối trái Dứa ra thị trường một cách phù hợp, góp phần nâng cao năng suất và giá trị kinh tế cho trái Dứa ở vùng trồng Đại Lộc Quảng Nam. Giải pháp nghiên cứu này có thể phát triển các tập dữ liệu khác nhau đối với các loại trái cây khác và tiếp tục nghiên cứu kết hợp các mô hình học sâu khác như R-CNN, FAST R-CNN. để phân loại, đánh giá chất lượng cho nhiều loại trái cây khác nhau góp phần giải quyết hiệu quả đầu ra của trái cây chín và tăng giá trị kinh tế cho nông sản Việt.



## Tài liệu tham khảo

- [1] S. Afaq and S. Rao, "Significance of epochs on training a neural network," *International Journal of Scientific and Technology Research*, vol. 19, no. 6, pp. 485–488, 2020.
- [2] A. Krizhevsky, I. Sutskever, and G. E. Hinton: "Imagenet classification with deep convolutional neural networks". In *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1097–1105, Tahoe City, CA, USA, 3–8 December 2012.
- [3] J. Redmon, S. Divvala, R. Girshick, Ali Farhadi: "You Only Look Once: Unified, Real-Time Object Detection". arXiv:1506.02640 [cs.CV], May (2016).
- [4] W. Widyawati, R. Febriani: "Real-time detection of fruit ripeness using the YOLOv4 algorithm". In *Teknika: Jurnal Sains Dan Teknologi*, Vol 17 No 02. 205–210, (2021).
- [5] K. Jun-Hwa, K. Namho, P. Y. Woon Park, W. C. Sun: "Object Detection and Classification Based on YOLO-V5 with Improved Maritime Dataset". In *Journal of Marine Science and Engineering*, *J. Mar. Sci. Eng.* 2022, 10, 377. <https://doi.org/10.3390/jmse10030377>, (2022).
- [6] D. T. P. Chung and D. Van Tai, "A fruits recognition system based on a modern deep learning technique," in *Journal of physics: conference series*, vol. 1327, p. 012050, IOP Publishing, 2019.
- [7] Z. M. Khaing, Y. Naung, and P. H. Htut: "Development of control system for fruit classification based on convolutional neural network,". In *Proceedings of the 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, Moscow and St. Petersburg, Russia, 29 Jan.-1 Feb. 2018.
- [8] Nguyen H.H.C., Luong A.T., Trinh T.H., Ho P.H., Meesad P., Nguyen T.T: "Intelligent Fruit Recognition System Using Deep Learning". In: Meesad P., Sodsee D.S., Jitsakul W., Tangwannawit S. (eds) *Recent Advances in Information and Communication Technology. IC2IT 2021. Lecture Notes in Networks and Systems*, vol 251. Springer, Cham (2021).
- [9] Nguyen H.H.C., Nguyen, D.H., Nguyen, V.L., Nguyen, T.T.: "Smart solution to detect images in limited visibility conditions based convolutional neural networks". In *Advances in Computational Collective Intelligence. ICCCI 2020. Communications in Computer and Information, Science*, vol. 1287, pp. 641{650. Springer, Cham (2020).
- [10] Horea Muresan an M. Oltean, "Fruit recognition from images using deep learning", *Acta Univ. Sapientiae, Informatica* 10, pp. 26-42, June 2018.
- [11] A. F. Agarap, "Deep learning using rectified linear units (RELU)," arXiv preprint arXiv:1803.08375, 2018.
- [12] S. Jana, Saikat B. and Ranjan P., "Automatic Fruit Recognition from Natural Images using Color and Texture Features", *IEEE, Devices for Integrated Circuit (DevIC)*, pp. 620-624, March 2017
- [13] K. Yamamoto, W. Guo, Y. Yoshioka, and S. Ninomiya: "On plant detection of intact tomato fruits using image analysis and machine learning methods". *Sensors*, vol. 14, no. 7, pp. 12191–12206, 2014.
- [14] S. Bargoti and J. Underwood: "Deep fruit detection in orchards". *IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, 3626-3633. doi: 10.1109/ICRA.2017.7989417 (2017).
- [15] Y. Zhang, K. Sohn, R. Villegas, Gang P., Honglak Lee: "Improving Object Detection with Deep Convolutional Networks via Bayesian Optimization and Structured Prediction". arXiv:1504.03293 [cs.CV], Jan. (2016).
- [16] K. Simonyan and A. Zisserman: "Very deep convolutional networks for large-scale image recognition". In *Computer Vision and Pattern Recognition (cs.CV)*, ICLR <https://doi.org/10.48550/arXiv.1409.1556>, (2015).
- [17] F. Valentino, T. W. Cenggoro, and B. Pardamean, "A design of deep learning experimentation for fruit freshness detection," in *IOP Conference Series: Earth and Environmental Science*, vol. 794, p. 012110, IOP Publishing, 2021.
- [18] D. Karakaya, O. Ulucan, and M. Turkan, "A comparative analysis on fruit freshness classification," in *2019 Innovations in Intelligent Systems and Applications Conference (ASYU)*, pp. 1–4, IEEE, 2019.
- [19] J. Feng, L. Zeng, and L. He, "Apple fruit recognition algorithm based on multi-spectral dynamic image analysis," *Sensors*, vol. 19, no. 4, p. 949, 2019.
- [20] J. Huang, V. Rathod, C. Sun, Menglong Z. A. Korattikara, S, Sergio, "Speed/accuracy trade-offs for modern convolutional object detectors", *Computer Vision and Pattern Recognition (cs.CV)*, <https://doi.org/10.48550/arXiv.1611.10012>. 2017.
- [21] <https://images.cv/dataset/pineapple-image-classification-dataset>.