

NGHIÊN CỨU MÔ HÌNH NÔNG NGHIỆP CÔNG NGHỆ CAO HỖ TRỢ NÔNG TRẠI TRỒNG MĂNG CỤT

A RESEARCH ON HIGH-TECH AGRICULTURAL MODEL FOR MANGOSTEEN FARMS

SV: *Bùi Xuân Thiện**, *Nguyễn Đặng Khuê Văn*, *Trần Thanh Hoàng*
Lớp 18IT3, Khoa Khoa học máy tính; Email: *bxthien.18it3@vku.udn.vn

GVHD: TS. *Nguyễn Hà Huy Cường*
Khoa Khoa học máy tính; Email: nhhcuong@vku.udn.vn

Tóm tắt: Trước xu hướng công nghiệp hóa và tự động hóa trong các lĩnh vực công nghiệp và dịch vụ, hiện nay các nhà khoa học, các nhà nghiên cứu đang đưa nền nông nghiệp bước vào kỉ nguyên 4.0, hay còn gọi là nền nông nghiệp thông minh, mà ở đó công nghệ cao đang được ứng dụng một cách mạnh mẽ. Trong đó, việc ứng dụng khoa học, kỹ thuật để phân tích, nhận dạng, bảo quản và chế biến trái cây nhằm đem lại hiệu quả kinh tế cho nông sản Việt là xu hướng nghiên cứu rất mới, có tác động trực tiếp đến hoạt động nông nghiệp công nghệ cao. Nhận thấy được những lợi thế trên, nhóm tác giả quyết định nghiên cứu và phát triển dự án “Ứng dụng mô hình nông nghiệp công nghệ cao hỗ trợ nông trại trồng măng cụt”, với mục đích phục vụ cho người dân sản xuất nông nghiệp nói chung và ngành nông nghiệp trồng măng cụt nói riêng. Đề tài của chúng tôi phát triển nhằm kết nối người dân với máy móc hiện đại, giúp người dân có thể theo dõi và quan sát quá trình phát triển của trái măng cụt.

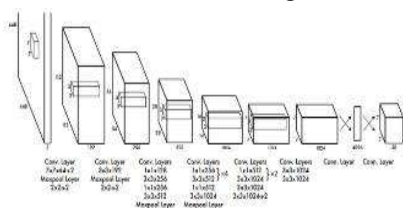
Từ khóa: Trí tuệ nhân tạo, phát hiện đối tượng, Thời kỳ chín, Học máy, Học sâu

1. Tổng quan về đề tài nghiên cứu

1.1. Cơ sở lý luận

1.1.1. Mô hình phát hiện đối tượng YOLO

You Only Look Once, hay còn gọi tắt là YOLO, là một thuật toán nhận diện đối tượng (Object Detection) trong lĩnh vực thị giác máy tính (Computer Vision). YOLO là một mô hình mạng neural tích chập (Convolutional Neural Network – CNN) được phát triển cho việc phát hiện, nhận dạng, và phân loại đối tượng. YOLO được tạo ra bằng sự kết hợp giữa các lớp tích chập (convolutional layers) và các lớp kết nối (connected layers). Trong đó các convolutional layers sẽ trích xuất ra các đặc điểm của ảnh, còn full-connected layers sẽ dự đoán xác suất và tọa độ của đối tượng.



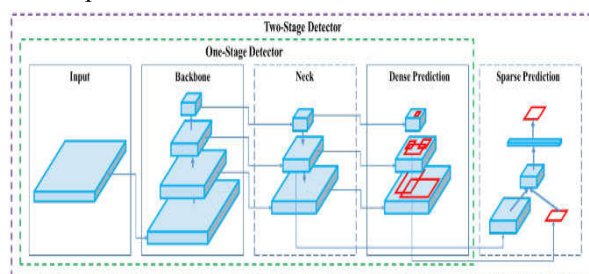
Hình 1: Cấu trúc CNN

Abstract: Facing the trend of industrialization and automation in the fields of industry and services, scientists and researchers are now bringing agriculture into Industry 4.0, also known as smart agriculture in which high technology is being highly utilized. In particular, the application of science and technology to analyze, identify, preserve and process fruits in order to bring economic efficiency to Vietnamese agricultural products is a very new research trend that has a direct impact on hi-tech agricultural activities. Foreseeing those advantages, the authors decided to research and develop the project "A research on hi-tech agricultural model for mangosteen farms" for farmers and and particularly the mangosteen farming. Our project is developed to connect people with modern machines, so that people can monitor and observe the development of mangosteen

Keywords: Artificial Intelligence, Object Detection, Growth stage, Machine Learning, Deep Learning

Kiến trúc của YOLO gồm 4 phần:

- Backbone (Dense Block & DenseNet, CSP, CSPDarknet53);
- Neck (FPN, SPP);
- Head (Dense Prediction);
- Sparse Prediction;



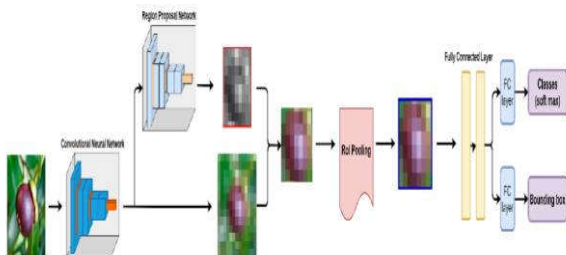
Hình 2: Mô hình kiến trúc You Only Look Once

1.1.2. Nhận dạng đối tượng Faster R-CNN

Hiệu suất cao do học sâu mang lại, một tập hợp con của học máy, nó đang trở nên rất phổ biến trên nhiều loại dữ liệu. CNN được xây dựng để phân loại hình ảnh sử

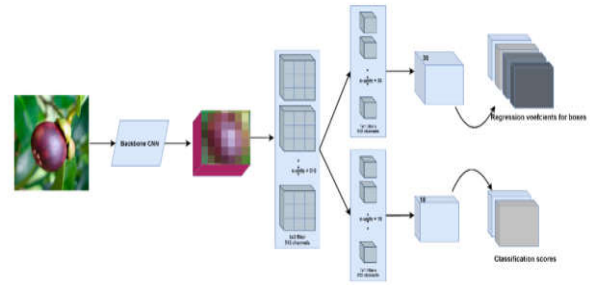
dùng học sâu. Xây dựng mô hình CNN dễ dàng hơn với việc sử dụng thư viện Keras trong Python. Mô hình sử dụng các lớp CNN như Conv2D & MaxPooling2D. Conv2D còn được gọi là 2D Convolution Layer. Lớp này tạo ra hàng chục đầu ra bằng cách tạo ra nhân chập trùng với lớp đầu vào. MaxPooling 2D là dùng cho hoạt động tổng hợp tối đa cho dữ liệu không gian. Dữ liệu không gian có thể được định nghĩa là biểu diễn thông tin về một đối tượng vật lý bằng các giá trị số. Việc chọn phần tử tối đa từ vùng của bản đồ điện tử được bao phủ bởi bộ lọc là toán hạng được thực hiện bởi lớp tổng hợp tối đa. Để giảm kích thước của bản đồ đối tượng, các lớp gộp được sử dụng. Do đó, có thể kết luận rằng việc sử dụng các lớp tổng hợp làm giảm số lượng các tham số cần tìm hiểu và khả năng tính toán được thực hiện trong mạng nhân tạo.

Faster R-CNN tiếp tục sử dụng tìm kiếm có chọn lọc. Faster R-CNN, không giống như R-CNN, không tách đề xuất vùng 2000 khỏi hình ảnh và thực hiện phân loại hình ảnh cho từng hình ảnh riêng biệt. Nó sẽ tạo ra một bản đồ đặc trưng phức tạp, tải toàn bộ hình ảnh vào ConvNet (lớp chập + lớp tổng hợp tối đa). Sau khi kết hợp CNN, chúng ta sẽ có một bản đồ đặc trưng nhỏ hơn đáng kể so với hình ảnh gốc, nhờ lớp ROI Pooling, lớp này sẽ đưa các vùng hình ảnh về kích thước hình vuông và định hình lại với cùng kích thước đầu ra. Tiếp tục qua các lớp Được kết nối đầy đủ cho đến khi đạt đến đầu ra vector tính năng ROI. Cuối cùng, để có được kết quả cuối cùng, mô hình sẽ dự đoán lớp và độ lệch của hộp giới hạn cho vùng phân mảnh đó.



Hình 3: Cấu trúc của Faster R-CNN

Faster R-CNN là một cải tiến so với hai người họ hàng gần là R-CNN và Fast R-CNN. Nó tránh tìm kiếm Chọn lọc (thuật toán sẽ tạo ra khoảng 2000 vùng cho mỗi hình ảnh), là những vùng có thể chứa nhiều thứ và kích thước. Cách tiếp cận này phân tích khoảng 2000 vùng gợi ý cho mỗi hình ảnh, việc này tốn nhiều thời gian và nó được tạo ra từ một mô hình khác, tốn kém chi phí đào tạo. Sau khi hoàn thành quá trình đào tạo trước CNN để trích xuất tính năng hình ảnh, Faster R-CNN sẽ sử dụng mô hình RPN (mạng đề xuất vùng) tinh chỉnh cho nhiệm vụ vùng đề xuất, được khởi tạo bởi bộ phân loại hình ảnh đào tạo trước. Nó phụ trách phân loại các mặt tích cực (khu vực chứa đối tượng) và tiêu cực (nền). Công việc tiếp theo trong mô hình đào tạo được thực hiện giống như Fast R-CNN, sử dụng các lĩnh vực đề xuất của RPN. RPN và mạng phát hiện hiện đang trao đổi các lớp tích chập. Faster R-CNN kết hợp các thuật toán khuyến nghị RPN và Fast R-CNN để tăng tốc độ xử lý tính năng.



Hình 4: Region Proposal của Faster R-CNN

1.1.3. Hệ thống thông tin địa lý WebGIS

Cùng với sự bùng nổ về công nghệ Internet, GIS cũng đã phát triển công nghệ cho phép chia sẻ thông tin thông qua Internet bằng cách kết hợp GIS và Web để tạo thành WebGIS. WebGIS là một giải pháp client – server cho phép quản lý, phân tích, cập nhật, phân phối thông tin bản đồ và GIS trên mạng Internet, giảm thiểu chi phí đầu tư phần mềm, phần cứng cho người dùng cuối; giao diện thân thiện, đơn giản phù hợp với nhiều người dùng.

WebGIS thích hợp với các cơ sở dữ liệu bản đồ, GIS từ rất bé cho đến rất lớn, có khả năng tùy biến cao, phù hợp với nhiều loại hình tổ chức. WebGIS là xu hướng phổ biến thông tin mạnh mẽ trên Internet không chỉ dưới góc độ thông tin thuộc tính thuần túy mà nó kết hợp được với thông tin không gian hữu ích cho người sử dụng. WebGIS được xem là một hệ thống thông tin địa lý được phân bố qua các môi trường mạng máy tính để tích hợp, phân phối và truyền tải thông tin địa lý trực tuyến thông qua Internet. Người dùng Internet có thể truy cập đến các ứng dụng của GIS mà không phải mua phần mềm. Ngoài ra, WebGIS còn cho phép thêm các chức năng của GIS trong các ứng dụng về quản lý, tra cứu thông tin... của giao thông, du lịch, hành chính, nông nghiệp, môi trường, v.v.

- Cho phép quản lý nhiều bản đồ
- Cho phép chọn lọc, tìm kiếm thông tin
- Cho phép cập nhật thông tin
- Quản trị hệ thống



Hình 5: Hệ thống thông tin địa lý WebGIS

Xác định vị trí của phòng khám, phòng thí nghiệm, dịch vụ chăm sóc sức khỏe gần nhất. Xác định loại trường (đại học, cao đẳng, trung cấp ...) và địa chỉ xung quanh đó để tìm kiếm, tham khảo dữ liệu nền đường phố.

Hệ thống được đưa ra nhằm quản lý và giám sát các thông tin chi nhánh, đại lý, cửa hàng của các đơn vị quản lý trên một nền thông tin địa lý phong phú và đa dạng.

Quản lý hệ thống phân phối hàng hóa, các mặt hàng, khả năng tiêu thụ, doanh thu lợi nhuận thu được.

1.2. Tổng quan các nghiên cứu liên quan trong và ngoài nước

Sự phát triển nhanh chóng của khoa học và công nghệ (KH&CN) làm cho nền nông nghiệp dần thay đổi, mang đến xu hướng áp dụng các thành tựu công nghệ vào sản xuất nhằm nâng cao năng suất và sản lượng sản phẩm cũng như thúc đẩy phát triển nông nghiệp bền vững. Việt Nam là một nước có thế mạnh về nông nghiệp, do đó cần phải có sự thay đổi trong quá trình sản xuất, tiêu thụ sản phẩm. Hiện nay, vấn đề an toàn thực phẩm đang đặt ra rất nhiều thách thức mới, đòi hỏi chúng ta phải nâng cao năng suất cũng như chất lượng cây trồng, đồng thời phải thích nghi với biến đổi khí hậu ngày càng khắc nghiệt. Chính vì vậy ngành nông nghiệp luôn luôn vận động, thay đổi sáng tạo để đáp ứng các nhu cầu đó. Một trong những hướng phát triển hiện nay là dựa trên nền tảng công nghệ truyền thông, trí tuệ nhân tạo để sản xuất nhiều nông sản hơn, ít tốn công sức lao động hơn, giảm thiểu tác động đến môi trường và ngược lại.

Ngày nay, trí tuệ nhân tạo phát triển rất nhanh, nó được hiểu là trí tuệ do con người tạo ra và dùng máy móc để vận hành. Trí tuệ nhân tạo có cách cư xử, sự học hỏi và khả năng thích ứng thông minh. Đây là một thành tựu của loài người, ứng dụng trong mọi lĩnh vực đời sống và ngành nông nghiệp cũng vậy. Mặc dù nhiều thách thức nhưng những kết quả đạt được cho thấy tiềm năng to lớn khi áp dụng vào nông nghiệp.

- Ứng dụng AI trong máy móc, robot nông nghiệp, thao tác bằng máy móc từ gieo trồng, chăm sóc, tưới tiêu, theo dõi đến thu hoạch và phân loại sản phẩm.

- Ứng dụng AI trong việc giám sát chất lượng đất, năng suất mùa vụ thông qua camera, vệ tinh, máy bay không người lái.

- Ứng dụng AI trong phân tích dự báo thời tiết, dự báo dịch bệnh và hệ thống phòng trừ để nâng cao độ chính xác và nhanh chóng.

1.2.1. Ngoài nước

Neural Networks được lấy cảm hứng từ sự hiểu biết về sinh học của bộ não loài người – sự liên kết giữa các nơ-ron. Tuy nhiên, không giống như một bộ não sinh học nơi mà bất kỳ nơ-ron nào cũng có thể liên kết với các nơ-ron khác trong một khoảng cách vật lý nhất định, các mạng thần kinh nhân tạo này có các lớp rời rạc, các kết nối, và các hướng truyền dữ liệu.

Tại Mỹ, được biết tới là nơi xuất khẩu nông nghiệp top đầu thế giới. Với sản lượng đạt tới hàng trăm, hàng nghìn tấn mỗi năm. Đây là một trong những nơi có lượng đất đai trồng cây nông nghiệp rộng lớn nhất lên tới hàng trăm héc ta. Việc trồng cây và thu hoạch đòi hỏi một lượng nhân công cực kỳ lớn và trải qua hàng tháng trời. Đòi hỏi các nhân công lành nghề lâu năm. Nhưng việc này lại kéo theo giá thành của nông phẩm cao. Khiến người tiêu dùng phải băn khoăn.

Phát hiện trái cây trong vườn cây đã được giải quyết bởi suchet Bargoti et al. [3], thực hiện một RCNN nhanh hơn để bản địa hóa các loại trái cây (xoài, hạnh nhân và táo), nhưng không phát hiện các giai đoạn chín.

Một cách tiếp cận khác để phát hiện quả trong vườn được Joseph Redmon et al. [17,18,19,20,21,26] đưa ra, phát triển kiến trúc MangoYolo (dựa trên YOLO-v3, YOLO-v2 [7,8]) để phát hiện xoài trong vườn với độ chính xác trung bình là 0,983%.

D.Stajanko, M.Lakota, M.Hočevar [28] đã sử dụng máy ảnh nhiệt để chụp ảnh táo và áp dụng các thuật toán xử lý hình ảnh khác nhau để ước tính năng suất và kích thước của táo trong vườn cây.

Phát hiện giai đoạn chín của trái cây trong vườn cây đã được thảo luận bởi Y. Tian et al. [7], sử dụng YOLO-v3 và phát hiện các giai đoạn phát triển khác nhau của táo trong vườn cây. Vấn đề phát hiện chín trong [11] yêu cầu đào tạo các giai đoạn phát triển táo khác nhau thành các đối tượng riêng biệt và do đó, nó sẽ gây ra chi phí đào tạo lớn nếu một giai đoạn phát hiện bổ sung được thêm vào mô hình do S. Kim et al. [20]. Một số nhà nghiên cứu đã nghiên cứu những thay đổi về tính chất vật lý và hóa học của các loại trái cây khác nhau trong quá trình chín [21, 22].

Nhiệm vụ của những chiếc máy bay này là chụp những khu vực trồng trọt của nông dân, với mục đích tối ưu hóa năng suất cây trồng và giảm chi phí. Để đạt được điều này, họ sử dụng máy tính để phân tích những hình ảnh được chụp từ vệ tinh, máy bay không người lái và điện thoại di động. Các kết quả được đưa vào ứng dụng Aeroview- 1 app của công ty.

Và một trong những sáng kiến nổi trội nữa đó là những nhà sáng chế đã tạo ra những cỗ máy khổng lồ có khả năng gieo hạt, bón phân, tưới nước thậm chí thu hoạch cây trồng trong một diện tích lớn. Những cỗ máy này được điều khiển chỉ bằng một người nông dân. Giờ đây, người nông dân có thể tiết kiệm được số tiền thuê người lao động và bán ra với giá thành ổn định. Những cỗ máy có khả năng thực hiện nhiệm vụ chỉ trong một ngày so với đội ngũ nhân công làm việc trong 1 tuần.

1.2.2. Trong nước

Trong những năm gần đây, chúng ta chứng kiến được nhiều thành tựu vượt bậc trong ngành Thị giác máy tính (Computer Vision). 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.

Ở Việt Nam thời gian gần đây đang chú trọng đến lĩnh vực hẹp của công nghệ Trí tuệ nhân tạo (AI), Machine Learning là phương pháp phân tích dữ liệu từ đó tự động hóa việc xây dựng mô hình phân tích. Có thể nói đây là công nghệ rất hứa hẹn mang lại những hỗ trợ tối ưu cho các doanh nghiệp với nhiều ứng dụng trong thế giới thực, ví dụ như: nhận dạng giọng nói và nhận diện hình ảnh.

Với nhận diện hình ảnh, có rất nhiều tình huống hệ thống phân loại hình ảnh các đối tượng như một ảnh kỹ thuật số. Đối với hình ảnh số, các phép đo mô tả các kết quả đầu ra của mỗi pixel trong hình ảnh. Với ảnh đen trắng, mật độ của mỗi pixel được tính như 1 đơn vị đo. Vì thế, nếu một hình ảnh đen trắng có $N * N$ pixels, tổng số

pixel và các phép đo là N2. Với ảnh màu, mỗi pixel được tính như cung cấp 3 phép đo với mật độ của 3 thành phần màu chính là RGB. Vì vậy ảnh màu $N \times N$ có 3 phép đo N2.

Về phần nhận diện khuôn mặt – là các hạng mục có thể yêu cầu trình diện hoặc không trình diện khuôn mặt. Có thể có hạng mục riêng cho từng người trong cơ sở dữ liệu của một số cá nhân.

Về phần nhận dạng ký tự – Các đoạn văn bản có thể được phân tách thành các hình ảnh nhỏ hơn, mỗi hình có chứa một ký tự đơn. Các mục có thể bao gồm 26 chữ cái trong bảng chữ cái tiếng Anh, 10 chữ số, và một số ký tự đặc biệt.

Là một trong những nước xuất khẩu hàng đầu của thế giới, đứng thứ 2 về xuất khẩu gạo thế giới. Tuy nhiên về chất lượng gạo thì thua Thái Lan và các nước khác và chưa tìm được nguồn cầu ổn định. Ngành nông nghiệp chủ yếu là sản xuất truyền thống, chưa tiếp thu được các máy móc công nghệ cao. Sử dụng nguồn nhân lực là chính. Người dân chưa áp dụng một số kỹ thuật cho việc chăm sóc và theo dõi nông nghiệp. Làm cho chất lượng đầu ra thấp hơn các nước khác.

Tại Việt Nam, Nhóm nghiên cứu của TS. Lê Đình Duy và đồng sự tại PTN Truyền thông Đa Phương tiện, ĐH Công nghệ Thông tin có những nghiên cứu về tái nhận dạng nhân vật trong khuôn khổ của đề tài về phát hiện trộm dùng mạng cảm biến camera. Phát hiện sự kiện trong kho dữ liệu video; Phát hiện thông tin bạo lực trong video; Tái xếp hạng trong tìm kiếm ảnh trong kho dữ liệu lớn; Nghiên cứu kỹ thuật hỗ trợ phân loại ảnh nhân khoa trong chẩn đoán hình ảnh;

Nhóm nghiên cứu của TS. Huỳnh Hữu Hưng và cộng sự nghiên cứu về “Xử lý tín hiệu video và nhận dạng cử chỉ hỗ trợ người khuyết tật hòa nhập cộng đồng”;

Nhóm nghiên cứu của PGS.TS Đỗ Năng Toàn và cộng sự nghiên cứu về kỹ thuật vùng quan sát và phát hiện bất thường của các đối tượng trong hệ thống camera giám sát”; Hướng nghiên cứu giải quyết việc chuyển tiếp các camera trong hệ thống giám sát tự động.

Nổi bật nhất là tỉnh Lâm Đồng với hơn 17 trang trại nông nghiệp thông minh, sử dụng công nghệ kết nối vạn vật (IoT) và trí tuệ nhân tạo để kết nối với các thông số và chăm sóc cà chua, dâu tây, rau sạch. Ứng dụng AI trong nuôi tôm ở Cà Mau, Bạc Liêu giúp tính toán, phân tích các yếu tố ngoại lai tác động vào và kích bản ứng phó và hỗ trợ máy móc. Từ đó phát triển và gia tăng năng suất lao động và nguồn kinh tế ổn định.

2. Nội dung nghiên cứu

2.1. Đối tượng nghiên cứu

Bảng 1: Đối tượng nghiên cứu

Đối tượng nghiên cứu	Khách thể nghiên cứu
Khó khăn của nông dân trồng măng cụt trong quá trình canh tác	Các nông hộ trồng măng cụt tại địa phương
Nhu cầu cải thiện canh tác	Các nông hộ trồng măng cụt

của nông dân trồng măng cụt tại địa phương	
Cách ứng dụng mô hình học sâu vào phát hiện thời kỳ chín của măng cụt	Tài liệu về các công nghệ, thuật toán liên quan
Áp dụng các công nghệ mới trong quá trình canh tác	Nông trại trồng măng cụt tại địa phương

2.1.1. Phạm vi nghiên cứu

Nghiên cứu của nhóm tác giả thực hiện triển khai tại trang trại trồng măng cụt tại Tỉnh Bình Dương. Trong nghiên cứu, nhóm quan tâm đến việc thu thập dữ liệu phục vụ cho nghiên cứu. Xây dựng hệ thống có khả năng nhận dạng, theo dõi, tính tổng sản lượng và đánh giá được thời kỳ chín của trái măng cụt có tại vườn trồng. Kết quả được cập nhật và đánh dấu các giai đoạn trên bản đồ web. Giúp nông dân có thể quản lý dễ dàng hơn.

2.1.2. Phương pháp nghiên cứu

Trong nghiên cứu này, nhóm thực hiện qua 2 bước chính: Huấn luyện mô hình và xây dựng hệ thống.

Huấn luyện mô hình: Sử dụng tập dữ liệu mà nhóm đã thu thập được tại nông trại trồng măng cụt trên địa bàn tỉnh Bình dương. Tiến hành đánh nhãn trên 10000 hình ảnh ban đầu. Sử dụng mô hình huấn luyện YOLOv4 để đào tạo ra mô hình.

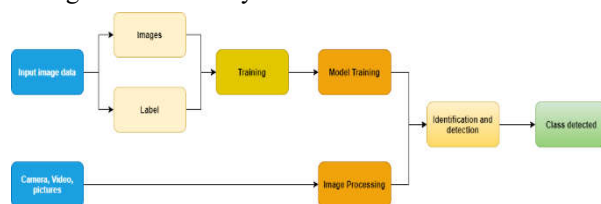
-Xây dựng hệ thống: Xây dựng hệ thống nhận dạng, phát hiện, đánh giá thời kỳ chín và tính tổng sản lượng trái măng cụt có trong vườn trồng.

•Hệ thống sử dụng 2 phương pháp nhận dạng là: nhận dạng qua hình ảnh và nhận dạng qua video.

•Hệ thống tích hợp giám sát địa lý bằng WebGIS: cập nhật thông tin của chủ sở hữu, diện tích, địa điểm, khoanh vùng đất trồng, ...

2.2. Thiết kế hệ thống

Hệ thống nhận dạng và phát hiện thời kỳ chín của trái măng cụt dựa trên cơ sở ứng dụng công nghệ thị giác máy tính và kỹ thuật xử lý hình ảnh, sử dụng mô hình học sâu cụ thể là mạng nơ-ron tích chập. Hệ thống ban đầu gán nhãn cho những hình ảnh đầu vào, sau đó sử dụng mô hình học sâu đào tạo trích xuất những đặc trưng của hình ảnh. Kết quả của quá trình huấn luyện, mô hình đã được học sẽ được đưa vào để dự đoán hình ảnh. Trên cơ sở dựa vào màu sắc, hình dạng, kích thước và kết cấu, việc phân tích và đánh giá trái măng cụt được thực hiện để có thể nhận dạng và phát hiện được thời kỳ chín của trái măng cụt. Kiến trúc hệ thống nhận dạng trái măng cụt được mô tả trong hình 7 dưới đây.

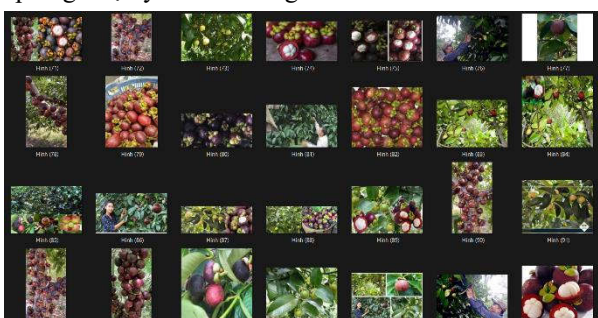


Hình 6: Sơ đồ hoạt động của hệ thống

3. Triển khai xây dựng hệ thống

3.1. Bộ dữ liệu đầu vào

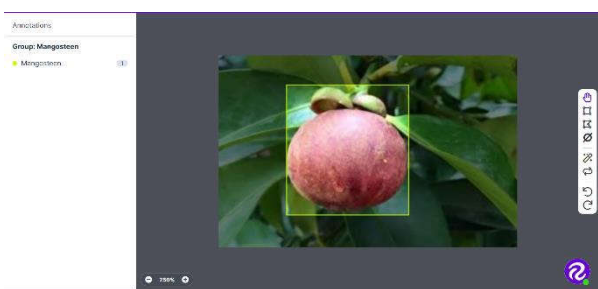
Bộ dữ liệu được sử dụng trong đề tài nghiên cứu được nhóm tác giả tự thu thập trong quá trình học tập tại trường. Hình ảnh được lấy từ các vườn trồng măng cụt trên địa bàn tỉnh Bình Dương. Bộ dữ liệu bao gồm 10000 ảnh trái măng cụt trong các giai đoạn. Trong đó, bộ dữ liệu được chia thành 3 tập dữ liệu nhỏ lần lượt là tập huấn luyện – chịu trách nhiệm cho quá trình huấn luyện, sử dụng cho máy học trích xuất những đặc điểm thông tin quan trọng có trong hình ảnh. Tập xác thực – chịu trách nhiệm đối sánh với mô hình huấn luyện trong quá trình đào tạo để đảm bảo mô hình chạy đúng ngưỡng cho phép. Tập kiểm thử - chịu trách nhiệm kiểm chứng độ chính xác sau khi mô hình đào tạo hoàn tất quá trình huấn luyện có đáp ứng được yêu cầu không.



Hình 7: Hình ảnh của bộ dữ liệu

3.2. Gán nhãn dữ liệu

Gán nhãn dữ liệu là giai đoạn quan trọng nhất trong quá trình huấn luyện có đạt được độ chính xác hay không. Sau khi thu thập tập dữ liệu đầu vào, nhóm nghiên cứu tiến hành gán nhãn cho từng ảnh để trích xuất được những vùng đặc trưng của trái măng cụt trong hình ảnh. Gán nhãn được nhóm sử dụng phần mềm website có sẵn Roboflow. Roboflow được sử dụng để gán nhãn dữ liệu hình ảnh với giao diện trực quan tiếp cận người dùng. Hệ thống tự động lấy khung hình đã được chọn cho việc trích xuất thông tin trên hình ảnh và lưu giữ bằng file txt cùng tên với hình ảnh.



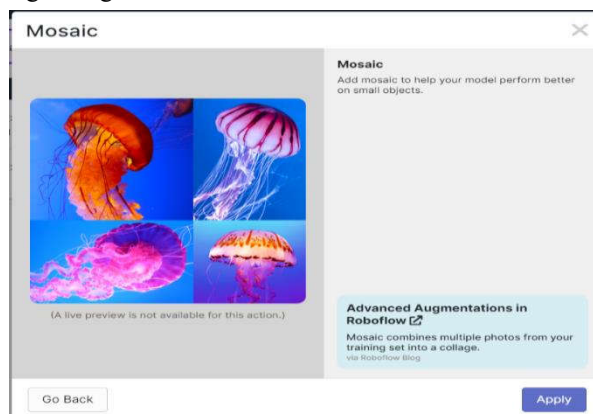
Hình 8: Gán nhãn dữ liệu

3.3. Tăng cường dữ liệu

Bộ dữ liệu ban đầu còn khá ít, điều này ảnh hưởng tới việc huấn luyện mô hình để tạo ra được một kết quả tốt trong việc dự đoán. Vì vậy nhóm nghiên cứu đã đưa ra đề xuất là tăng cường dữ liệu để phục vụ cho việc nghiên cứu ít dữ liệu mà vẫn có thể tạo ra được nhiều dữ liệu hơn dựa trên những dữ liệu đã có. Hiện nay có nhiều cách để tăng cường dữ liệu, mỗi chức năng tăng cường dữ liệu lại

có các yếu tố điều khiển riêng như:

- Phương thức tăng cường dữ liệu: lật ảnh, xoay, cắt ảnh, đổi màu, tăng độ sáng, tăng nhiễu, v.v.
- Yếu tố điều khiển tăng cường: mỗi tăng cường sẽ có những yếu tố điều khiển riêng. Như xoay bao nhiêu độ, lật ảnh bao nhiêu lần hay cắt ảnh ngẫu nhiên khoảng bao nhiêu lần, ...
- Tăng cường cho từng giai đoạn huấn luyện hay epoch. Epoch đầu ở learning rate lớn có thể khác với các epoch cuối ở learning rate nhỏ.
- Tăng cường cho từng lớp khác nhau, tần suất dùng tăng cường có thể khác nhau.



Hình 9: Tăng cường dữ liệu

3.4. Huấn luyện mô hình

Trước khi bắt đầu huấn luyện, mô hình yêu cầu một số chức năng cần phải thực hiện. Trong đề tài nghiên cứu lần này, nhóm sử dụng mô hình huấn luyện YOLOv4. Đây là một mô hình đã được cải thiện đáng kể về độ chính xác và tốc độ huấn luyện. Sau đây là các tệp cần thiết để phục vụ quá trình huấn luyện dữ liệu.

-Tập các lớp nhận dạng: tệp có chức năng lưu trữ những thông tin các lớp dùng cho việc nhận dạng. Trong nghiên cứu này, nhóm thực hiện duy nhất trái măng cụt.

-Tập đào tạo: Lưu trữ các thông tin bao gồm tên, định dạng và đường dẫn lưu nơi chứa hình ảnh phục vụ cho việc nhận dạng

-Tập kiểm thử: Lưu trữ các thông tin bao gồm tên, định dạng và đường dẫn lưu nơi chứa hình ảnh phục vụ cho việc kiểm thử.

-Tập thông tin gán nhãn: Lưu trữ các thông số liên quan đến quá trình gán nhãn ở mục c bao gồm định danh class, thông số định dạng tọa độ vị trí gán nhãn trên hình ảnh

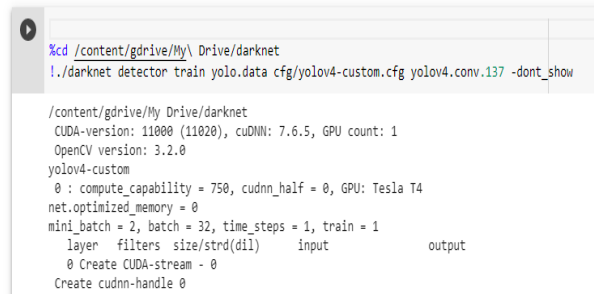
-Tập cấu hình: tệp chứa thông tin, cấu trúc, cấu hình của mô hình huấn luyện YOLOv4.

-Thư mục data: Lưu trữ hình ảnh phục vụ cho việc đào tạo và kiểm thử cho mô hình huấn luyện YOLOv4.

Để thực hiện quá trình huấn luyện một cách đầy đủ và hoàn chỉnh. Nhóm sử dụng Google Colab trên Drive làm môi trường phục vụ đào tạo vì tính hữu dụng và các thiết lập sẵn của nó. Hơn hết, Google Colab cho phép sử dụng GPU để phục vụ cho mục đích huấn luyện.



Hình 10: Kết nối drive với Google Colaboratory



Hình 11: Sử dụng mô hình huấn luyện yolo từ darknet

3.5. Xây dựng hệ thống nhận dạng

Sau quá trình hoàn thiện và đánh giá mô hình, nhóm đã xây dựng được một hệ thống có chức năng phân tích hình ảnh và video. Hệ thống nhận dữ liệu đầu vào xử lý và phân tích, sau đó so sánh với mô hình huấn luyện đã được đào tạo.

Hệ thống được tích hợp các chức năng để nhận dạng và phát hiện được trái mãng cụt dựa trên hình ảnh và video đầu vào. Ngoài ra hệ thống có thể đếm số lượng giúp cho việc tính tổng sản lượng có thể dễ dàng hơn.

Nhóm nghiên cứu phát triển thêm hệ thống thông tin địa lý bằng website giúp quản lý được thông tin vùng đất của các chủ hộ và các thông số sản lượng và diện tích của vùng đất của chủ hộ.

3.6. Đơn vị đánh giá mô hình

Nhóm sử dụng 4 đại lượng đánh giá mô hình bao gồm:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}$$

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

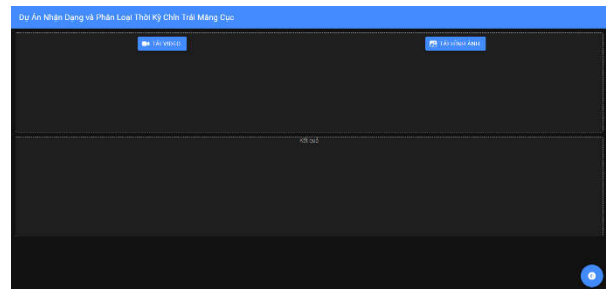
$$IoU = \frac{TP}{(TP + FP + FN)}$$

Hình 12: Đại lượng đánh giá mô hình

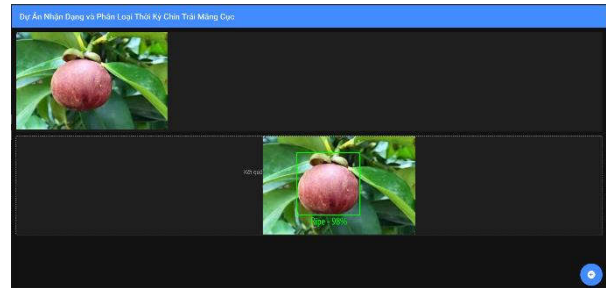
4. Kết quả nghiên cứu

Sau quá trình hoàn thiện và đánh giá mô hình, nhóm đã xây dựng được một hệ thống có chức năng phân tích hình ảnh và video. Hệ thống nhận dữ liệu đầu vào xử lý và phân tích, sau đó so sánh với mô hình huấn luyện đã được đào tạo. Kết quả đánh giá được hiển thị sau khi phân tích. Hệ thống dễ dàng sử dụng cho những người dân

canh tác tại các khu vườn. Sau đây là một số hình ảnh của hệ thống.



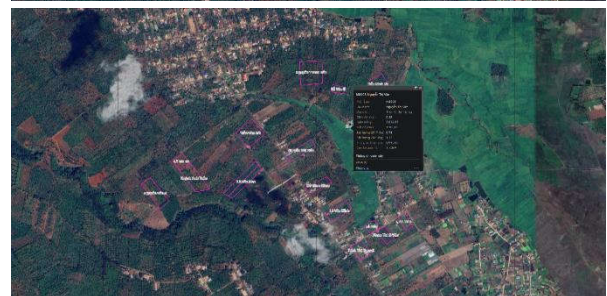
Hình 13: Giao diện của hệ thống



Hình 14: Giao diện hệ thống khi nhận dạng trái

Ngoài ra, nhóm nghiên cứu đã xây dựng một hệ thống bản đồ địa lý có khả năng lưu trữ những thông tin về nông trại. Lưu những thông tin như, ranh giới đất đai nông trại, chủ nông trại, thông tin về vùng đất, tọa độ đất, sản lượng của các năm, ...

Các chỉ dẫn địa lý bao gồm các thông tin về nguồn gốc của vị trí địa lý và ngành nông sản: Từ ngữ, dấu hiệu, biểu tượng, hình ảnh để chỉ, vùng lãnh thổ, hàng hóa, sản lượng, tháng năm, quy mô, diện tích, chủ trang trại, màu sắc lãnh thổ, v.v. Nhóm đã sử dụng một vị trí ngẫu nhiên trên bản đồ Việt Nam để giới thiệu về bản đồ địa lý WebGIS. Từ đó triển khai website bằng cách đăng kí tên miền và hosting. Người dùng truy cập internet để truy cập địa chỉ WebGIS và thực hiện các thao tác trên bản đồ.



Hình 15: Giao diện hệ thống thông tin địa lý

4.1. Đánh giá

Nhóm đã xây dựng được một hệ thống khá hoàn chỉnh trong quá trình nghiên cứu. Hệ thống đã cho thấy được một kết quả khả quan. Tuy nhiên, trong tình hình Covid-19 kéo dài nên việc thu thập dữ liệu gây khó khăn trong việc thực tế tại các vườn trồng măng cụt tại Tỉnh Bình Dương. Nhóm chỉ nghiên cứu và huấn luyện với một bộ dữ liệu nhỏ. Đây là tiền đề để nhóm tiếp tục phát triển dự sau này.

Nhóm đã phát triển các chức năng mới và phát triển thêm hệ thống cải tiến để có thể đánh giá chi tiết hơn về quá trình nhận dạng và phân loại. Cập nhập các thông tin này lên website. Với hệ thống có giao diện mới và dễ sử dụng đối với người dùng. Bản đồ địa lý có thể chuyển đổi các bản đồ vệ tinh khác nhằm cho người dùng có cái nhìn khách quan nhất về vùng đất của mình.

4.2. Hướng phát triển

Trong thời gian sắp tới, nhóm nghiên cứu sẽ nâng cấp và cải tiến hệ thống của mình và đưa ra những bản thử nghiệm để đem tới cho những chủ trang trại để tiến hành thực nghiệm và kiểm tra hệ thống với các chức năng vận hành. Nhằm thử nghiệm và kiểm tra để đem sản phẩm tới người nông dân sớm nhất.

5. Kết luận

Đề tài đã nghiên cứu được về quá trình phát triển của trái măng cụt. Từ đó nhóm có nhiều kiến thức hơn để thực hiện quá trình hoàn thiện sản phẩm. Hệ thống đã cho thấy được quá trình giám sát và theo dõi trái măng cụt trong thời kỳ chín. Giúp cho người dân có thể tiết kiệm được thời gian và nhân công.

Trong quá trình nghiên cứu hệ thống, nhóm đã phát hiện 1 số khó khăn nhất định. Hình ảnh chưa được đa dạng cho các hình dạng của trái măng cụt. Tùy vào thời tiết mà hệ thống có thể nhận dạng chính xác đạt mức cho phép. Nghiên cứu đang ở mức nhận dạng trái măng cụt tại 1 vùng địa lý nhất định.

Hiện nhóm chưa nghiên cứu được hết tất cả chức năng trong đề tài nghiên cứu của nhóm, những chức năng chưa hoàn thiện đầy đủ và chưa tích hợp với các máy móc kỹ thuật cao. Hệ thống đã có thể thực hiện một số chức năng nhất định. Đóng góp trong quá trình hoàn thiện của sản phẩm. Tăng năng suất và hiệu quả trong những mô hình trồng măng cụt diện tích lớn. Cải thiện được chất lượng và theo dõi được tình hình của trái măng cụt.

Việc kết hợp nhận dạng và phân loại thời kỳ chín của trái măng cụt và hệ thống thông tin địa lý WebGIS là một bước tiến mới trong quá trình phát triển một hệ thống nông nghiệp công nghệ cao. Sản phẩm có thể áp dụng tại các nông trại quy mô lớn lên tới trăm hecta. Thay thế việc quản lý và chăm sóc truyền thống. WebGIS có thể đưa tất cả các thông tin về chủ hộ và vùng đất của chủ hộ. Từ đó có thể so sánh sản lượng hàng năm và điều chỉnh hợp lý.

Đề tài này là tiền đề để nhóm nghiên cứu trong thời gian tới. Hệ thống nhận dạng và phát hiện thời kỳ chín của trái măng cụt không dừng ở việc giám sát. Hệ thống có thể nâng cấp thêm, tích hợp với camera tại vườn hay sử dụng máy bay không người lái gửi dữ liệu về hệ thống.

Khi đã có dữ liệu đầy đủ, việc phát triển thêm chức năng theo dõi tình trạng bệnh có trên trái măng cụt, theo dõi độ ẩm của đất, độ pH có trong đất. Phân tích và đánh giá để có thể đem lại chất lượng cho người dân. Đem lại năng suất cao và giá thành ổn cho thị trường. Đề tài nghiên cứu cũng có thể mở rộng trong các mô hình trồng bưởi, chuối, đu đủ, ...

Tuy nhiên, điều kiện hiệu quả của bản thân tôi là sinh viên nên còn nhiều hạn chế, việc nghiên cứu khoa học, song hành với việc học chính khóa nên chắc chắn không tránh khỏi những thiếu sót. Vì thế, tôi rất mong nhận được sự đóng góp ý kiến của quý thầy cô và các bạn sinh viên để bài nghiên cứu của tôi được hoàn thiện hơn.

Tài liệu tham khảo

- [1] Inkyu Sa, ZongYuan Ge, Feras Dayoub, B. Upcroft, Tristan Perez, C. McCool. DeepFruits: A Fruit Detection System Using Deep Neural Networks (2016).
- [2] YunongTian, GuodongYang, ZheWang, HaoWang, EnLi, ZizeLiang. Apple detection during different growth stages in orchards using the improved YOLO-V3 model (2012).
- [3] Byoungjun Kim, You-Kyoung Han, Jong-Han Park and Joonwhoan Leel. Improved Vision-Based Detection of Strawberry Diseases Using a Deep Neural Network (2021).
- [4] Jose Luis Rojas-Aranda, Jose Ignacio Nunez-VarelaJ. C. Cuevas-TelloGabriela Rangel-Ramirez. Fruit Classification for Retail Stores Using Deep Learning (2020).
- [5] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection
- [6] Guoxu Liu, Joseph Christian Nouaze, Philippe Lyonel Touko Mbouembe and Jae Ho Kim. YOLO-Tomato: A Robust Algorithm for Tomato Detection Based on YOLOv3 (2020).
- [7] Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27 June 2016; pp. 779–788.
- [8] Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 26 July 2017; pp. 7263–7271.
- [9] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster R-CNN: Towards RealTime Object Detection with Region Proposal Networks," arXiv:1506.01497v3 [cs.CV], Jan. 2016.
- [10] Redmon, J.; Farhadi, A. YOLOv3: An incremental improvement. arXiv 2018, arXiv:1804.02767.
- [11] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao. YOLOv4: Optimal Speed and Accuracy of Object Detection.
- [12] Xiongwei Wu, Doyen Sahoo, Steven C.H. Hoi, "Recent Advances in Deep Learning for Object

- Detection,” arXiv:1908.03673v1 [cs.CV], Aug. 2019
- [13] Yuting Zhang, Kihyuk Sohn, Ruben Villegas, Gang Pan, Honglak Lee, “Improving Object Detection with Deep Convolutional Networks via Bayesian Optimization and Structured Prediction,” arXiv:1504.03293 [cs.CV], Jan. 2016
- [14] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In ICLR, 2015.
- [15] Dumitru Erhan, Christian Szegedy, Alexander Toshev, Dragomir Anguelov (2014) Scalable Object Detection using Deep Neural Networks. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2147-2154
- [16] Yongxi Lu, Tara Javidi, Svetlana Lazebnik, “Adaptive Object Detection Using Adjacency and Zoom Prediction,” arXiv:1512.07711 [cs.CV], Apr. 2016
- [17] Sun Q., Pfahringer B. (2011) Bagging Ensemble Selection. In: Wang D., Reynolds M. (eds) AI 2011: Advances in Artificial Intelligence. AI 2011. Lecture Notes in Computer Science, vol 7106. Springer, Berlin, Heidelberg
- [18] Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, Xindong Wu, “Object Detection with Deep Learning: A Review,” arXiv:1807.05511 [cs.CV], Apr. 2019
- [19] Keiron O’Shea, Ryan Nash, “An Introduction to Convolutional Neural Networks,” arXiv:1511.08458 [cs.NE], Dec. 2015
- [20] Munera, S., Amigo, J. M., Blasco, J.; Cubero, S.; Talens, P., Alexios, N.(2017). Ripeness monitoring of two cultivars of nectarine using VIS-NIR hyperspectral reflectance imaging. *Journal of Food Engineering*, 214(3), 29-39.
- [21] Santagapita, P.R., Tylewicz, U. Panarese, V., Rocculi, P., Dalla, Rosa, M. (2016). Nondestructive assessment of kiwifruit physic-chemical parameters to optimize the osmotic dehydration process, A study on FT-NIR spectroscopy. *Journal of Biosyst. Eng.* 142,
- [22] S. Kim, Y. Ji and K. Lee, ” An Effective Sign Language Learning with Object Detection Based ROI Segmentation,” 2018 Second IEEE International Conference on Robotic Computing (IRC), Laguna Hills, CA, 2018, pp. 330-333.