

Xây Dựng Hệ Thống Nhận Dạng Phân Loại Trái Cây Chín Tiếp Cận Mạng Nơron Tích Chập (CNN)

Trịnh Trung Hải¹, Bùi Xuân Thiện², Nguyễn Lê Tùng Khánh³,
Nguyễn Hà Huy Cường⁴

^{1,2,3}Trường Đại học Công nghệ Thông tin và Truyền thông Việt – Hàn
{¹tthai, ²bxthien.18it3, ³nltkhanh}@vku.udn.vn

⁴Trung tâm phát triển phần mềm – Đại học Đà Nẵng

⁴nhhcuong@sdc.udn.vn

Tóm tắt. Xu hướng tự động hóa trong lĩnh vực nông nghiệp hay còn gọi là nông nghiệp thông minh đang được ứng dụng một cách mạnh mẽ trong giai đoạn chuyển đổi số thời kỳ công nghiệp 4.0. Cùng với đó việc ứng dụng khoa học, kỹ thuật để tối ưu hóa các sản phẩm đầu ra của nông nghiệp đang là vấn đề cấp bách được Chính Phủ và Người nông dân quan tâm trong những năm gần đây. Trong đó, việc ứng dụng khoa học, kỹ thuật để phân tích, nhận dạng, bảo quản và chế biến trái cây nhằm mang lại hiệu quả kinh tế cho nông sản Việt là xu hướng nghiên cứu rất mới có tác động tích cực đến hoạt động nông nghiệp công nghệ cao. Trong bài báo này, chúng tôi tập trung nghiên cứu quá trình phân tích, nhận dạng hình ảnh trái cây bằng kỹ thuật thị giác máy tính (Computer Visions), kết hợp mạng nơ-ron tích chập đa lớp (CNN) để xây dựng hệ thống tự động xử lý nhận dạng và phân loại trái cây.

Từ khóa: Hệ thống tự động nhận dạng, nhận dạng trái cây, phân loại trái cây; học máy; học sâu; thị giác máy tính; mạng nơ-ron tích chập.

1 Giới thiệu

Học máy (Machine Learning) là một khái niệm của ứng dụng của trí tuệ nhân tạo (AI) cung cấp cho các hệ thống khả năng tự động học hỏi và tự cải thiện từ kinh nghiệm mà không cần sự chỉ định cụ thể bởi người lập trình. Học sâu (Deep Learning) là một chi của ngành máy học dựa trên một tập hợp các thuật toán để mô hình hóa dữ liệu trừu tượng ở mức cao. [20]

Ngày nay, học máy và học sâu là công nghệ tiên tiến dựa trên ý tưởng cốt lõi của trí tuệ nhân tạo và khả năng tự tư duy được ứng dụng để phát triển các chương trình máy tính nhằm giải quyết các vấn đề của thế giới thực. Công nghệ này tập trung vào khả năng xử lý nguồn dữ liệu lớn, đa lớp với cấu trúc phức tạp để phân tích giải quyết các bài toán phức tạp, mang giá trị công nghệ vượt trội ở nhiều lĩnh vực như: Dịch vụ tài chính, Tiếp thị & Bán hàng, Chăm sóc sức khỏe, Giao thông Vận tải và Nông nghiệp. Một quá trình học máy và học sâu được thực hiện mô hình bắt đầu từ việc quan sát, thu thập dữ liệu đầu vào và huấn luyện tự động để hệ thống đưa ra những dự đoán và quyết định chính xác hơn cho những đối tượng nghiên cứu. [4,5]

Kỹ thuật xử lý hình ảnh và thị giác máy tính ngày nay đang được ứng dụng ở nhiều lĩnh vực và quy mô lớn. Trong đó, nhận dạng, phân loại trái cây từ hình ảnh là một nghiên cứu mới được phát triển ở nhiều Quốc gia. Kỹ thuật nhận dạng trái cây là một nhiệm vụ rất khó khăn vì trái cây có thể chứa nhiều đặc điểm giống nhau giữa chúng. Tuy nhiên với sức mạnh tính toán, phân loại xử dữ liệu lớn trong xử lý ảnh sẽ dễ dàng giải quyết những khó khăn trên. Với mô hình nhận dạng trái cây hiệu quả có thể tạo ra những thiết bị nhận dạng chi phí thấp, mang lại nhiều lợi ích cho người nông dân [2,4,5]

Con người có thể vận dụng trí tuệ, kinh nghiệm của mình để nhận biết bất kỳ đối tượng hay bất kỳ vật phẩm nào trong thực tế nhờ việc xác định tính chất, đặc điểm của từng đối tượng. Để nhận biết, phân loại trái cây người ta có thể đo lường các thông số đặc điểm, tính chất hóa, lý cụ thể của từng loại trái cây, nhưng phương pháp này chỉ là cảm tính, không nhất quán, kết quả đưa ra chỉ mang tính chất tương đối, không chính xác đối với việc xử lý số lượng lớn. Hơn nữa với phương pháp thủ công này sẽ tốn kém thời gian, công sức và không có hiệu quả kinh tế cao. Công đoạn thủ công kiểm tra phân loại trái cây có thể thay thế bằng hệ thống nhận dạng tự động dựa vào việc phân tích hình ảnh (thị giác máy tính) sẽ cho ra kết quả chính xác hơn. Trong nghiên cứu này, chúng tôi ứng dụng kỹ thuật thị giác máy tính để phân tích, xử lý nguồn dữ liệu lớn hình ảnh của nhiều loại trái cây được thu thập từ thực tế. công đoạn kiểm tra chất lượng hình ảnh bằng kỹ thuật thị giác máy tính bao gồm 4 bước chính: đó là thu thập, phân đoạn, trích xuất và phân loại tính năng. Kỹ thuật nhận dạng trái cây bằng công nghệ thị giác máy tính được coi là một nghiên cứu mới mẻ đầy tiềm năng. Tuy nhiên để mang lại kết quả chính xác cao, kỹ thuật này còn nhiều thách thức do hình ảnh nhận diện của các đối tượng (trái cây) tạo ra nhiều điểm tương đồng về nhận diện giữa hình ảnh các loại trái cây và các giai đoạn phát triển của một loại trái cây do tác động bởi nhiều yếu tố môi trường khác nhau như: ánh sáng, dịch bệnh... [3, 5]. Vì vậy, để khắc phục các vấn đề nêu trên, với mục đích là phát triển một hệ thống tự động nhận dạng, phân loại loại trái cây một cách tối ưu và cho kết quả chính xác cần có sự kết hợp kỹ thuật thị giác máy tính và mô hình học máy với thuật toán học sâu mạng nơ-ron tích chập đa lớp (CNN).

Trong bài báo này chúng tôi trình bày các nội dung như sau: phần một giới thiệu tổng quan, phần hai trình bày các hướng nghiên cứu liên quan, phần ba thiết kế hệ thống cài đặt thuật toán, phần thứ tư đánh giá các kết quả thử nghiệm, phần thứ năm 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.... Trong thực tế có rất nhiều tình huống hệ thống phân loại hình ảnh các đối tượng như một ảnh kỹ thuật số, các

phép đo mô tả các kết quả đầu ra của mỗi pixel trong hình ảnh được cung cấp bởi 3 thành phần màu chính là RGB (Red, Green, Blue) [1, 2, 17].

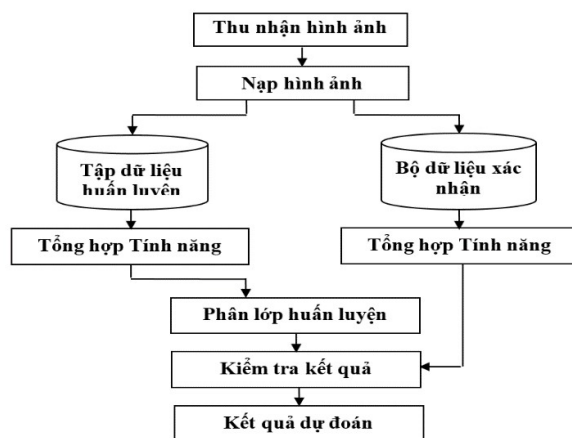
Trong lĩnh vực nông nghiệp công nghệ cao các nhà nghiên cứu Việt Nam, Ấn Độ, Nhật Bản, Trung Quốc quan tâm đến vấn đề nông nghiệp công nghệ cao rất nhiều, có thể kể ra như: Nhóm nghiên cứu Đặng Thị Phương Chung và Đinh Văn Tài [3] đã đề xuất hệ thống nhận dạng trái cây bằng cách sử dụng mạng nơ-ron tích 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 như: lớp đầu vào, lớp ẩn và lớp đầu ra bao gồm trong kiến trúc của ANN. Mỗi lớp được cấu tạo bởi nhiều tế bào thần kinh và dựa trên hàm chi phí, sản lượng được so sánh với các giá trị mục tiêu. Nhóm nghiên cứu Horea và cộng sự 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. Nhóm Susovan Jana và cộng sự. [6] đề xuất hệ thống tự động nhận dạng trái cây và rau quả sử dụng công nghệ thị giác máy tính và máy học. Hệ thống này được sử dụng 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 một cách chính xác và thực tế hơn. Anuja Bhargava và Atul Bansal [7] đã đề xuất một hệ thống quy định việc sử dụng công nghệ xử lý hình ảnh và thị giác máy tính trong lĩnh vực công nghiệp thực phẩm và nông nghiệp. Lei Hou và cộng sự [8] đề xuất một hệ thống thuật toán nhận dạng trái cây dựa trên việc sử dụng Mạng nơ ron Convolution (CNN). Với kỹ thuật CNN này hệ thống được huấn luyện để tìm ra loại trái cây gốc từ nhiều hình ảnh được đưa trực tiếp vào mạng, kết hợp đặc tính khác để tính toán mà không cần trích xuất tính năng nào trên đó. Chang Liu và cộng sự [9] đề xuất hệ thống học sâu để đánh giá, cách tiếp cận để nhận biết thực phẩm trong cơ sở hạ tầng dịch vụ điện toán biên. Hệ thống nhận dạng thực phẩm theo thời gian thực được thực hiện với việc sử dụng dịch vụ điện toán biên. Tian, Y., Yang, G., Wang, Z [10] và nhóm của Redmon, J., Farhadi [11] đã nghiên cứu và đề xuất mô hình YOLO v3 cải tiến phát hiện các giai đoạn tăng trưởng khác nhau của táo trong vườn cây ăn quả. Stajanko, Denis and Lakota và các cộng sự [12] đã nghiên cứu, đề xuất hệ thống ước tính số lượng và đường kính của quả táo trong vườn trong mùa sinh trưởng bằng hình ảnh nhiệt. Phương pháp này dựa trên những đặc tính về hình thể, kích thước và số lượng, mật độ quả đối với từng loại cây ăn trái để so sánh với hình ảnh thực tế ghi nhận được tại vườn để xác định loại quả và thời điểm thu hoạch. Zhong-Qiu Zhao và cộng sự [13] đã nghiên cứu tổng quan về phát hiện đối tượng bằng mô hình học sâu, qua đó chứng minh sự phát triển nhanh chóng của học sâu, các công cụ tìm hiểu ngữ nghĩa, các tính năng cấp cao, sâu hơn, được giới thiệu để giải quyết các vấn đề tồn tại trong các kiến trúc truyền thống, cải tiến hiệu suất trong phát hiện và nhận dạng đối tượng. Shadman Sakib và cộng sự [14] đã đề xuất hệ thống nhận dạng trái cây bằng cách sử dụng Mạng Nơ-ron tích chập (CNN) và thị giác máy tính. Mô hình toán học được xây dựng và thực thi trong python bao gồm việc sử dụng TensorFlow. Ross Girshick, Jeff Donahue [15] và nhóm cộng sự của Shaoqing Ren [16] đã đề xuất hệ thống phân cấp phát hiện đối tượng chính xác và phân đoạn ngữ nghĩa một cách đa dạng. Phương pháp này đề xuất thuật toán mạng nơ-ron tích chập cải tiến (R-CNN) và mạng đề xuất khu vực (RPN) để chia sẻ các tính năng tổng hợp hình ảnh đầy đủ với mạng phát hiện đối tượng trên bộ dữ liệu PASCAL VOC để cho ra hiệu suất đo đạt chính xác hơn với thời gian thực chính xác hơn. Nhóm của Kaiming He đã nghiên

cứ kỹ thuật Mark R-CNN, với phương pháp này chỉ thêm một chi phí nhỏ vào Faster R-CNN, chạy ở tốc độ 5 khung hình / giây và mô phỏng phương pháp tiếp cận phân đoạn cá thể đối tượng trong một hình ảnh để tạo mặt nạ phân đoạn chất lượng cao.

Có thể khẳng định rằng những nhà nghiên cứu trong nước về trí tuệ nhân tạo, học máy, học sâu và thị giác máy tính đang chú trọng những hướng nghiên cứu xử lý ảnh liên quan đến con người như nhận dạng hình ảnh; nghiên cứu xử lý ảnh liên quan đến lĩnh vực Y khoa như chuẩn đoán trên ảnh X-quang, trong khi đó vấn đề nghiên cứu xử lý hình ảnh từ các lĩnh vực khác như hình ảnh nhận dạng trái, quả, vùng canh tác nông nghiệp còn rất hạn chế và rời rạc.

3 Thiết kế hệ thống

Hệ thống nhận dạng trái cây tự động dựa trên cơ sở ứng dụng công nghệ thị giác máy tính và kỹ thuật xử lý hình ảnh trong nông nghiệp. Đồng thời tiếp cận mô hình học sâu với mạng nơ-ron tích chập, xử lý phân lớp dữ liệu nhằm đơn giản hóa việc phân tích và trích xuất hình ảnh và bộ mô tả các thuộc tính của các loại trái cây. Trên cơ sở màu sắc, hình dạng, kích thước và kết cấu, việc phân tích chất lượng của quả được thực hiện [22]. Đề xuất sơ đồ kiến trúc của hệ thống nhận dạng trái cây được mô tả trong hình 1.



Hình 1. Kiến trúc hệ thống phân loại trái cây

Đầu vào của hệ thống là tập dữ liệu hình ảnh được thu thập thực tế theo các loại quả nhất định, sau khi đưa vào hệ thống theo đường dẫn xác định. Kết quả đầu ra, người dùng sẽ nhận được tên của trái cây cùng với hình ảnh đầu vào. Hệ thống nhận dạng tự động trái cây được thực hiện dựa trên mô hình học sâu, sử dụng thuật toán Mạng thần kinh phân lớp kế thừa thư viện Keras. Có hai mô hình là một để nhận dạng trái cây tươi và mô hình thứ hai để nhận dạng trái cây hỏng.

3.1 Mô hình nhận dạng trái cây tươi

Có tổng cộng 8 loại trái cây được sử dụng cho mục đích đào tạo và xác nhận. 8 loại trái cây này bao gồm Táo, Dưa leo, Cam, Lựu, Xoài, Dứa, Lê và Dâu. Tập dữ liệu bắt buộc của những trái này được lấy từ tập dữ liệu Fruit-360 của Kaggle. Số lượng hình ảnh hoa quả được thu thập cho mục đích đào tạo và xác nhận được thể hiện trong bảng 1.

Bảng 1. Số lượng hình ảnh trái cây được dùng để huấn luyện và xác nhận mô hình trái cây tươi

<i>STT</i>	<i>Loại trái cây</i>	<i>Thông số huấn luyện:</i> <i>7427</i>	<i>Thông số xác nhận:</i> <i>1387</i>
1	Táo	3295	1101
2	Dưa leo	745	257
3	Cam	479	160
4	Lựu	492	164
5	Xoài	490	166
6	Dứa	490	166
7	Lê	492	164
8	Dâu	492	164

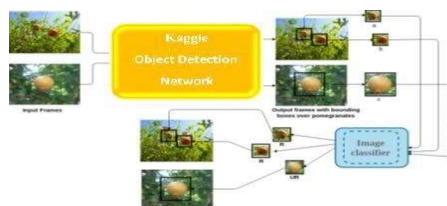
Bước đầu tiên: Cung cấp tập dữ liệu hình ảnh thu thập cho hệ thống. Tập dữ liệu được đưa vào xử lý gồm 8 đối tượng cần nhận dạng.

Bước thứ hai: Bao gồm việc lưu trữ tập dữ liệu vào bộ nhớ của hệ thống. Tập dữ liệu hình ảnh này sẽ được xử lý thô như: loại bỏ những tạp ảnh để có được hình ảnh rõ ràng trước khi xử lý.

Bước thứ ba: Bao gồm trích xuất đặc tính của hình ảnh. Các đặc tính như phát hiện và loại bỏ viền ảnh, kiểu mẫu và màu sắc. Kiểu mẫu của ảnh được sử dụng nhờ nhận biết qua bề mặt da và các khuyết tật (lỗ lấm) có trên quả. Các đặc điểm màu sắc được sử dụng để so sánh, xác định màu sắc của trái cây bao gồm giá trị trung bình và phương sai của mô hình màu RGB. Dựa vào các tính năng trên để trích xuất và huấn luyện bằng thuật toán Mạng nơ-ron tích chập (CNN) trong thư viện Keras.

Bước thứ tư: Dự đoán hình ảnh dựa trên mô hình được đào tạo. Đường dẫn tập dữ liệu hình ảnh đưa vào chính xác sẽ được phân tích, so khớp với dữ liệu hình ảnh và thuộc tính của từng loại trái cây đã được lưu trữ trong hệ thống để dự đoán kết quả theo tên nhãn và hình ảnh trái cây. Kích thước, màu sắc, hình dạng, kết cấu và khuyết tật là những đặc điểm chất lượng quan trọng nhất của trái cây.

Hàm mất mát categorical_crossentropy và trình tối ưu adamax được sử dụng để làm phép đo khi biên dịch mô hình nhận dạng tối ưu hóa. [18] Sự kết hợp này đã được sử dụng để kiểm thử trong mô hình biên dịch đã cho ra kết quả tốt nhất so với một số hàm func và trình tối ưu hóa mất mát khác. Hình 2 là mô hình tổng quát trong quá trình nhận dạng trái chín.



Hình. 2. Tóm tắt mô hình huấn luyện nhận dạng trái cây tươi



Hình. 3. Kết quả hình mẫu đầu vào và đầu ra để nhận dạng trái cây tươi

3.2 Nhận dạng trái cây hồng

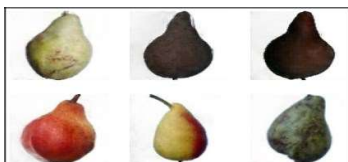
Trong mô hình nhận dạng trái cây hồng này chúng tôi sử dụng dữ liệu đầu vào gồm 3 loại trái cây: Táo, Cam và Chuối cho mục đích đào tạo và xác nhận. Bộ dữ liệu yêu cầu của những loại trái cây này được lấy từ bộ dữ liệu Fruit-360 của Kaggle. Số lượng hình ảnh trái cây được chụp cho mục đích đào tạo và xác nhận được trình bày trong bảng 2.

Bảng 2. Số lượng hình ảnh trái cây được sử dụng để đào tạo và xác nhận mô hình trái cây hồng

Loại trái cây	Thông số huấn luyện: 7427	Thông số xác nhận: 1387
Táo	2432	700
Cam	2500	648
Chuối	1609	532

Quy trình được sử dụng cho mô hình nhận biết trái cây hồng giống như quy trình nhận biết trái cây tươi. Hàm mất mát categorical_crossentropy và trình tối ưu adamax được

sử dụng để làm phép đo khi biên dịch mô hình nhận dạng tối ưu hóa [8]. Hai hàm này được sử dụng vì sau khi kiểm thử với một số hàm mất mát và trình tối ưu hóa khác, trong số đó, chúng cho kết quả tốt nhất. Kết quả đầu vào và đầu ra của mô hình nhận dạng trái cây hồng được thể hiện trong hình 4.



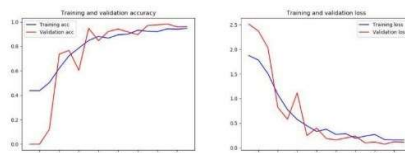
Hình 4. Đầu vào và đầu ra mẫu để nhận dạng trái cây hồng

4 Đánh giá hiệu suất

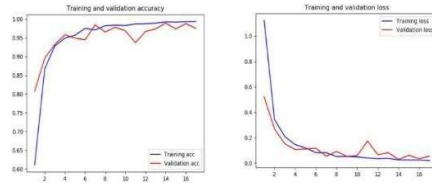
Thuật toán đã được đề xuất đã được kiểm tra hiệu suất và độ chính xác với sự hỗ trợ của chương trình tính toán tối ưu. Điều này được thực hiện bằng cách đánh giá các kết quả đầu ra của quá trình huấn luyện, hàm chi phí đào tạo, độ chính xác nhận dạng, mất mát nhận dạng tương ứng.

Hiệu suất cao do học sâu mang lại là một tập hợp con của học máy, nó đã được thực hiện trên các nguồn dữ liệu lớn và áp dụng rộng rãi trong nhiều lĩnh vực. Thuật toán CNN được xây dựng để phân loại hình ảnh sử dụng 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 như vậy. 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.

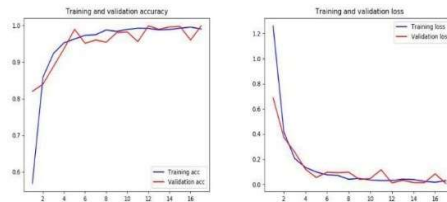
Các biến thể được tính toán để nhận dạng trái cây tươi với việc sử dụng bộ tối ưu như sau.



Hình 5. Các biến thể đối với trình tối ưu hóa sử dụng hàm SGD



Hình 6. Các biến thể đối với trình tối ưu hóa sử dụng hàm Adamax

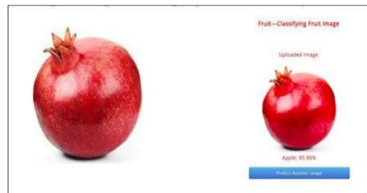


Hình 7. Các biến thể đối với trình tối ưu hóa sử dụng hàm RMSprop

Sau khi so sánh với các hàm mục tiêu, chúng ta thấy hàm mục tiêu Adamax cho ra kết quả tối ưu hơn những hàm mục tiêu khác. Vì thế việc sử dụng chương trình tính toán tối ưu với hàm mục tiêu adamax để nhận dạng trái cây tươi là rất hiệu quả. Kết quả nghiên cứu vẫn còn hạn chế với những dự đoán nhận dạng đối với một số loại trái cây có kích thước nhỏ.

Hình ảnh của lựu đã được nhập vào như là hình ảnh thử nghiệm như trong hình 8 (a), và đầu ra được hiển thị như quả Táo với 83,88% độ chính xác như trong hình 9 (b). Do đó dự đoán là sai vì mô hình được coi là tính năng màu của hình ảnh đầu vào và dự đoán cho màu của quả Táo đã được nhiều hơn theo tỷ lệ phần trăm như so với lựu.

Sử dụng hệ thống nhận dạng để nhận dạng và phát hiện được trái cây.



Hình 8 (a). Hình ảnh đầu vào

Hình 8 (b). Hình ảnh đầu ra

Hình ảnh quả xoài xanh đã được nhập vào như là hình ảnh thử nghiệm như trong hình 9 (a), và đầu ra được hiển thị như quả Lê với 60,76% độ chính xác như trong hình 9 (b). Do đó dự báo là sai vì mô hình được coi là tính năng màu của hình ảnh đầu vào và dự đoán cho màu quả Lê được nhiều hơn trong tỷ lệ phần trăm như so với quả xoài xanh.

**Hình 9 (a).** Hình ảnh đầu vào**Hình 9 (b).** Hình ảnh đầu ra

Hai trường hợp trên là những dự đoán sai lệch khi thực hiện các mô hình với những hạn chế nhất định. Lý do cho điều này có thể là do chất lượng hình ảnh thấp hoặc do tỷ lệ quyết định của yếu tố màu sắc quá cao trong huấn luyện. Do đó cần phải kết hợp thông số huấn luyện nhận dạng về hình thể, kích thước của từng đối tượng (loại quả) trong mô hình sử dụng để huấn luyện. Sau khi so sánh với trình tối ưu hóa, kết quả tốt nhất thu được từ trình tối ưu hóa Adamax hơn những phép thử khác. Do đó, việc sử dụng trình tối ưu hóa Adamax để nhận dạng trái cây hồng đã được thực hiện phù hợp nhất trong trường hợp này.

Với những kết quả thử nghiệm trên cùng một bộ CSDL ảnh chất lượng tốt, đã được tiền xử lý và gán nhãn như trên thì mô hình huấn luyện nhận dạng và phân loại đối tượng được ứng dụng phương pháp học sâu (mạng nơron tích chập - CNN) sẽ cho kết quả vượt trội hơn phương pháp học máy truyền thống. Đặc biệt đối với các loại trái cây có nhiều trích chọn đặc trưng phức tạp, sẽ giúp đơn giản hóa quá trình huấn luyện mô hình nhận dạng tự động, đồng thời cho phép tái sử dụng các mô hình đã huấn luyện trước để giảm thời gian cài đặt giải pháp cho các bài toán nhận dạng mới.

5 Kết luận

Việc ứng dụng công nghệ xử lý hình ảnh và thị giác máy tính kết hợp mô hình học máy tự động phân tích, nhận dạng trái cây để thay thế quá trình xử lý thủ công nhằm tăng hiệu suất phân loại trái cây tươi và trái cây hồng hỗ trợ bảo quản, chế biến các sản phẩm nông nghiệp là hết sức cấp thiết trong thời điểm hiện nay khi mà quá trình Covid 19 vẫn còn diễn ra trong thời gian dài. Khâu thiết yếu trong quá trình nhận và bảo quản trái chín đưa sau đó đưa vào kho lưu trữ giữ gìn trong thời gian kéo dài vài tháng có khi cả năm vừa tiết kiệm được nguồn lực nhân công vừa mang lại hiệu quả kinh tế cao. Giải pháp nghiên cứu kết hợp giữa thuật toán học sâu qua phương pháp CNN để tối ưu hóa quá trình xử lý hình ảnh và tập dữ liệu thuộc tính đối tượng như: kích thước, màu sắc, hình dạng, kết cấu và khuyết tật là những đặc điểm chất lượng quan trọng nhất của trái cây. Kết quả đạt được sau khi triển khai thử nghiệm cho phép đánh giá được điểm mạnh, điểm yếu và tính tối ưu của các thuật toán, đồng thời có thể ứng dụng xây dựng hệ thống nhận dạng, phân loại trái tự động từ hình ảnh thực tế theo phương pháp học sâu.

6 Acknowledgment

The authors wish to express their appreciation to the Ministry of Education and Training for supporting this research project as part of the Ministerial Program of Science

and Technology CTB.2021.DNA. "Research on applying deep learning model to recognize ripe pineapple period in Quang Nam - Da Nang"

References

1. Xiongwei Wu, Doyen Sahoo, Steven C.H. Hoi, "Recent Advances in Deep Learning for Object Detection", arXiv:1908.03673v1 [cs.CV], Aug. 2019.
2. Taesung Park, Ming-Yu Liu, Ting-Chun Wang, Jun-Yan Zhu, "Semantic Image Synthesis with Spatially-Adaptive Normalization", *Journal of Computer Vision Pattern Recognition (JCVPR)* - ISSN: 2432-5457, p-ISSN: 2196-9170), 18 Mar 2019.
3. Dang Thi Phuong Chung and Dinh Van Tai, "A fruits recognition system based on a modern deep learning technique", *IOP Conf. Series: Journal of Physics: Conf. Series* 1327, pp. 1-5, 2019.
4. Horea MURESAN an Mihai OLTEAN, "Fruit recognition from images using deep learning", *Acta Univ. Sapientiae, Informatica* 10, pp. 26-42, June 2018.
5. Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, Xindong Wu, "Object Detection with Deep Learning: A Review," arXiv:1807.05511 [cs.CV], Apr. 2019
6. Susovan Jana, Saikat Basak and Ranjan Parekh, "Automatic Fruit Recognition from Natural Images using Color and Texture Features", *IEEE, Devices for Integrated Circuit (DevIC)*, pp. 620-624, March 2017.
7. Anuja Bhargava and Atul Bansal, "Fruits and vegetables quality evaluation using computer vision: A review", *Journal of King Saud University – Computer and Information Sciences*, pp. 1-15, 8, May 2018.
8. Tian, Y., Yang, G., Wang, Z., et al., 2019. "Apple detection during different growth stages in orchards using the improved YOLO-V3 model[J]". *Computers and Electronics in Agriculture* 157, 417–426.
9. Redmon, J., Farhadi, A., 2018. "YOLOv3: An incremental improvement. In: *IEEE conference on Computer Vision and Pattern Recognition*", arXiv:1804.0276.
10. Stajanko, Denis and Lakota, Miran and Hočevcar, Marko. (2004). Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging. *Computers and Electronics in Agriculture*. 42. 31-42.
11. Shadman Sakib, Zahidun Ashrafi and Md. Abu Bakr Siddique, "Implementation of Fruits Recognition Classifier using Convolutional Neural Network Algorithm for Observation of Accuracies for Various Hidden Layers", *ArXiv e-Journal*, pp. 1-4, 1 April, 2019
12. Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," arXiv:1311.2524 [cs.CV], Oct. 2014
13. Ross Girshick, "Fast R-CNN," arXiv:1504.08083 [cs.CV], Sep. 2015
13. 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
14. Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick, "Mask R-CNN," arXiv:1703.06870 [cs.CV], Jan. 2018
15. <https://www.expertsystem.com/machine-learning-definition/>