

# Phân loại độ tươi của trái cây bằng cách sử dụng Thuật toán mạng nơron hợp pháp

1 Nguyễn Dương Thanh Dự

Cao Nguyễn Gia Hưng thứ 2

Ngày 3 Trương Thị Nhỏ Lê

Thứ 4 Lý Đặng Thái Hưng

5th Lê Đặng Thảo Mai

Tóm tắt – Phân loại độ tươi của trái cây bằng cách sử dụng Mạng nơ-ron Convolutions (CNN) là một trong những ứng dụng nhiều nhất trong thị giác máy tính. Trong bài báo này, một hệ thống nhận biết độ tươi của trái cây sử dụng CNN được đề xuất. Phương pháp được đề xuất sử dụng các kỹ thuật học sâu để phân loại. Chúng tôi đã sử dụng tập dữ liệu của mình cho mục đích đánh giá. Từ tập dữ liệu, chúng tôi đã thiết lập một tập dữ liệu bao gồm 845 hình ảnh từ 20 danh mục khác nhau.

Các hình ảnh được chia thành tập dữ liệu đào tạo và kiểm tra. Cuối cùng, chúng tôi đã đạt được độ chính xác kiểm tra tốt nhất là 96,6% trên dữ liệu thử nghiệm và độ chính xác đào tạo là 97,78%.

Từ khóa– Nhận biết độ tươi của trái cây, phù hợp Mạng thần kinh (CNN), Thị giác máy tính, Phân loại trái cây.

## I. GIỚI THIỆU

Trái cây là một phần thiết yếu của chế độ ăn uống cân bằng và mang lại nhiều lợi ích cho sức khỏe. Nhiều loại trái cây có quanh năm, trong khi những loại khác chỉ có vào một số mùa nhất định. Nông nghiệp tiếp tục là ngành đóng góp lớn cho nền kinh tế Việt Nam. Đất nông nghiệp chiếm 70% diện tích đất ở Việt Nam. Do đó, việc sử dụng các hệ thống học sâu để phân loại trái cây có lợi cho cả nhà tiếp thị và người tiêu dùng. Trong vài năm qua, tầm nhìn máy tính đã được sử dụng rộng rãi trong các phương pháp nhận dạng trái cây. Trong lĩnh vực nhận dạng và phân loại hình ảnh, phương pháp tiếp cận để phát hiện quả bằng cách sử dụng mạng nơron xoắn sâu được trình bày trong bài báo. Trong bài báo này, chúng tôi điều chỉnh một CNN đơn giản thông qua Tensorflow và phát hiện độ tươi của trái cây từ quan sát dọc, từ trên xuống.

Trong quá trình đào tạo, chúng tôi sẽ tách tập dữ liệu huấn luyện và tập dữ liệu thử nghiệm nhỏ hơn. Sau khi lựa chọn thuật toán và huấn luyện mô hình, chúng ta sẽ vẽ biểu đồ để đánh giá mô hình có hiệu quả hay không dựa trên giá trị của độ chính xác. Cuối cùng, chúng tôi sẽ dự đoán trên hình ảnh thực tế không có trong dữ liệu để đánh giá mô hình của chúng tôi.

## II. CÔNG VIỆC LIÊN QUAN

Nhiều hệ thống nhận dạng và phân loại có thể được tìm thấy trên internet để kiểm tra giai đoạn chín của trái cây và nhận biết chúng loại, độ tươi, v.v. [1] Phương pháp được sử dụng để phân loại chuối theo kênh màu và CIElab. Thuật toán tối ưu hóa bầy hạt (PSO) liên quan đến việc điều chỉnh các tham số mờ. [2] Đã sử dụng máy vectơ hỗ trợ nhân nhiều lớp (k-SVM) để phân loại trái cây.

SVM đã được đào tạo bằng cách sử dụng xác thực chéo phân tầng 5 lần với vector đặc trưng giảm. Sự kết hợp các đặc điểm màu sắc, kết cấu và hình dạng đã được sử dụng trong quá trình phân loại. [3] Học sâu nhận dạng hình ảnh dạng trái cây để xác định tươi và thối của sơ đồ phân loại cho bộ dữ liệu tiêu chuẩn. Trong công trình này, 60 giống trái cây khác nhau được

được nhận dạng từ 38409 hình ảnh. Trong các nghiên cứu trước đây, màu sắc, kết cấu và các đặc tính cạnh được coi là để phân loại trái cây. [4] Garcia và cộng sự. đã sử dụng thị giác nhân tạo để nhận dạng trái cây và chiết xuất các đặc điểm về hình dạng, màu sắc và kết cấu.

Mô hình phân loại trái cây đa mô hình được phát triển bằng cách sử dụng các ứng dụng học sâu. Nhận dạng trái cây có thể được coi là một bài toán phân đoạn hình ảnh. Một số công trình có sẵn trong tài liệu đề cập đến vấn đề nhận dạng trái cây như một vấn đề phân đoạn hình ảnh. [5]

Wang và cộng sự. thành lập một hệ thống phát hiện táo dựa trên màu sắc của chúng. [6] Byoungjun Kim và cộng sự, đề xuất sử dụng học sâu để cải thiện khả năng phát hiện chính xác quả Dâu trong thử nghiệm môi trường huyền bí, xen lẫn trong lá. [7] Jose Luis Rojas-Aranda và cộng sự. đã sử dụng phân loại trái cây học sâu cho cửa hàng bán lẻ, với màu đơn sắc RGB và biểu đồ RGB từ phân nhóm K-mean được sử dụng sớm trong mô hình.

## III. KIẾN TRÚC TỔNG THỂ CỦA MÔ HÌNH CNN

Học sâu đã được chứng minh gần đây là một kỹ thuật phát hiện hình ảnh rất mạnh mẽ và CNN là một phương pháp tiếp cận hiện đại để học sâu. CNN cho thấy độ chính xác cao hơn so với các phương pháp dựa trên máy-vector-hỗ trợ truyền thống với các tính năng thủ công. Đối với phát hiện hình ảnh độ tươi của trái cây, CNN cũng cho thấy độ chính xác cao hơn đáng kể so với một phương pháp thông thường. Bên cạnh đó, cách tiếp cận này cũng nhanh chóng hơn để triển khai.

Các bước thực hiện: Thu nhận hình ảnh, xử lý trước và trích xuất trong tương lai, đào tạo mô hình dựa trên CNN, dự đoán, đánh giá mô hình.

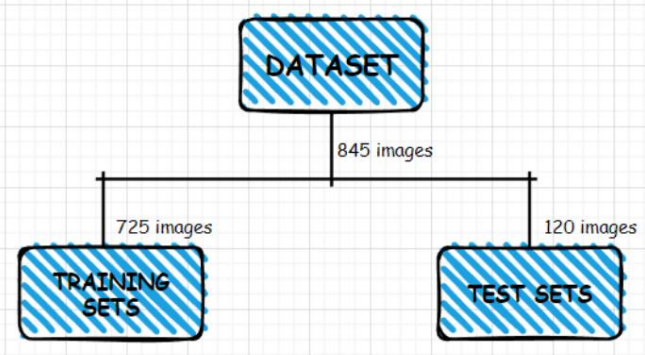
### A. Tập dữ liệu

Tất cả dữ liệu trong tập dữ liệu được thu thập bằng camera của điện thoại thông minh với mọi góc độ quan sát từ trên xuống dưới. Bộ dữ liệu có 5 loại trái cây (táo, chuối, xoài, ổi, cam) với 4 mức độ tươi (100%, 75%, 50%, 25%) cho mỗi loại trái cây.



Hình 1: Ví dụ về tập dữ liệu

Những hình ảnh được thu thập từ chợ Gò Vấp. Từ tập dữ liệu, chúng tôi đã thiết lập một tập dữ liệu chứa 845 hình ảnh với kênh RGB, chia thành 725 hình ảnh cho bộ đào tạo và 120 hình ảnh cho bộ thử nghiệm.



Hình 2: Cấu trúc tập dữ liệu

B. Tiền xử lý

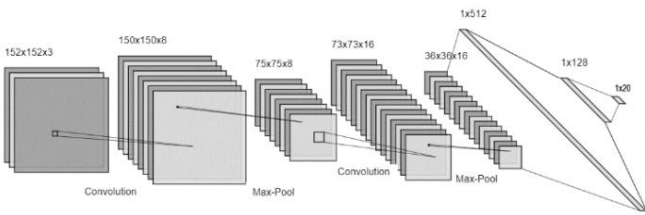
Khi quá trình thu thập dữ liệu bằng điện thoại thông minh, các hình ảnh có hình dạng và kích thước khác nhau và việc đào tạo một mạng nơ-ron phức tạp trên các bức ảnh thô sơ rất có thể sẽ dẫn đến các cuộc triển lãm phân loại khủng khiếp. Vì vậy, hình ảnh được thay đổi kích thước thành hình vuông (152 x 152 pixel) và loại bỏ các đối tượng không cần thiết khỏi hình ảnh.

Đối với mô hình nhận dạng độ tuổi trái cây này, mỗi hình ảnh được xoay 20 độ, dịch chuyển chiều rộng và chiều cao 20% ngẫu nhiên, thay đổi tỷ lệ và thu phóng 20%, lật ngang và cắt với phạm vi 20%.

Chúng tôi chuẩn hóa hình ảnh để mỗi kênh trở về phân phối chuẩn chuẩn N (0, 1). Cuối cùng, tập dữ liệu sẽ được gắn nhãn và lưu dưới dạng định dạng tập dữ liệu Tensorflow.

C. Mô hình phân loại độ tuổi của trái cây

Chúng tôi sử dụng các thuật toán học sâu của CNN để nhận dạng độ tuổi của trái cây. Mạng bao gồm hai lớp tích tụ, mỗi lớp được theo sau bởi các lớp tổng hợp và hai lớp được kết nối đầy đủ được thể hiện trong hình 3.



Hình 3: Kiến trúc mạng trung tính hợp lệ

Lớp đầu vào của mạng chứa 69.312 nơ-ron làm dữ liệu đầu vào, đại diện cho hình ảnh RGB tiêu chuẩn có kích thước 152 x 152 pixel. Lớp ẩn đầu tiên là lớp chập 1 có 8 bộ lọc với nhân có kích thước 3 x 3 pixel và các đơn vị tuyến tính đã chỉnh lưu (ReLU) (1) như một chức năng kích hoạt.

$$() = (0, ) \tag{1}$$

Lớp chập thứ hai là lớp chập 2 trong đó 16 bộ lọc với kích thước hạt nhân là 3 x 3 pixel và ReLU đã được sử dụng như trên lớp chập đầu tiên. Lớp phức hợp được sử dụng để trích xuất tính năng từ dữ liệu đầu vào. Cũng thực hiện thao tác tích chập đối với các khu vực nhỏ được bản địa hóa bằng cách kết hợp một bộ lọc với lớp trước đó.

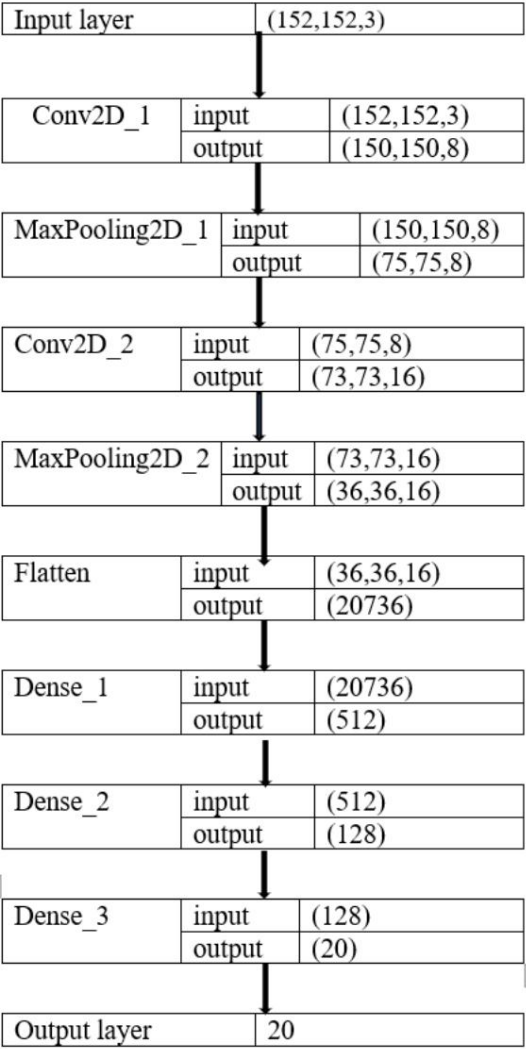
Kích thước hạt nhân xác định vị trí của các bộ lọc. ReLU được sử dụng như một chức năng kích hoạt để nâng cao hiệu suất ở cuối tất cả các lớp chập và các lớp được kết nối đầy đủ.

Trong gộp các lớp 1 và 2, chúng tôi sử dụng gộp tối đa với kích thước gộp là 2 x 2 với độ dài sải chân là 1. Các lớp gộp sẽ giảm thiểu kích thước không gian của đầu ra. Bước tiến xác định cách hoạt động của phép tích chập với một hạt nhân khi kích thước lớn hơn của một hình ảnh và các hạt nhân phức tạp được sử dụng để cải thiện hiệu suất của mạng bằng cách làm cho nó mạnh mẽ hơn. Điều này làm cho mạng trở nên có khả năng tổng quát hóa tốt hơn.

Cuối cùng, sự xuất hiện của phương trình tích chập sau:

$$(\text{---} + 2 + 1) \quad (\text{---} + 2 + 1) \quad ( \quad )$$

Đầu vào nào có kích thước (H, W, D). Kernel có kích thước (F, F, D), sải chân là S và phần đệm là P và K là bao nhiêu hạt nhân được sử dụng cho lớp tích chập này.



Hình 4: Kiến trúc nhận dạng hình ảnh trái cây tuổi

Sau khi trích xuất đặc trưng từ lớp chập, một lớp phẳng được sử dụng để chuyển ma trận bộ lọc 2D thành vectơ đặc trưng 1D với 20,736 tế bào thần kinh sẽ đi vào các lớp được kết nối đầy đủ. Lớp ẩn tiếp theo là lớp 1 được kết nối đầy đủ đủ bao gồm 20,736 tế bào thần kinh với và ReLU

(1). Cuối cùng, lớp 2 được kết nối đầy đủ hoặc lớp đầu ra chứa 20 tế bào thần kinh nơi kích hoạt bộ phân loại SoftMax (3)

được sử dụng để báo trước đầu ra của mô hình và thể hiện 4 mức độ tuổi khác nhau của 5 loại trái cây.

$$= \frac{\dots}{\dots}, \quad = 1, 2, \dots, \quad (3)$$

Trong phương trình này, ai đại diện cho xác suất là loại i. C là số đầu ra của lớp, trong mạng này C bằng 20.

Để đào tạo mạng của chúng tôi, chúng tôi đã sử dụng trình tối ưu hóa Adam. Adam là viết tắt của cụm từ Ước tính thời điểm thích ứng, tính toán tỷ lệ học tập thích ứng cho siêu tham số. Thuật toán tối ưu hóa Adam rất dễ triển khai, thích hợp cho các vấn đề có độ dốc thưa thớt và hiệu quả về mặt tính toán. Đối với mô hình CNN của chúng tôi, chúng tôi đã sử dụng trình tối ưu hóa Adam với tỷ lệ học tập là 0,001, kích thước lô là 16 và 25 kỷ nguyên. Tỷ lệ học tập được mặc định là 0,001 đối với bắt đầu tự nhiên tăng hoặc giảm bằng cách kiểm tra độ chính xác của xác nhận.

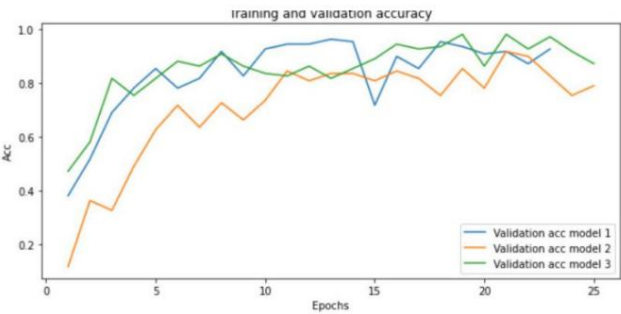
Để tính toán hiệu suất của mô hình, cần phải tính toán tỷ lệ lỗi. Chúng tôi đã sử dụng hàm chỉ phí phân loại-xuyên tâm làm hàm tổn thất.

IV. KẾT QUẢ THỰC NGHIỆM

Trong bài báo này, chúng tôi đã áp dụng một mạng nơ ron phức hợp để tìm hiệu suất phân loại độ tuổi của trái cây của mạng. Để xác định độ chính xác của phân loại tổng thể, chúng tôi đã thực hiện bốn trường hợp trong đó chúng tôi áp dụng các kết hợp khác nhau của các lớp ẩn (tích chập và nhóm) cho 25 kỷ nguyên với kích thước lô 16 và tính toán độ chính xác trên bộ thử nghiệm và đào tạo.

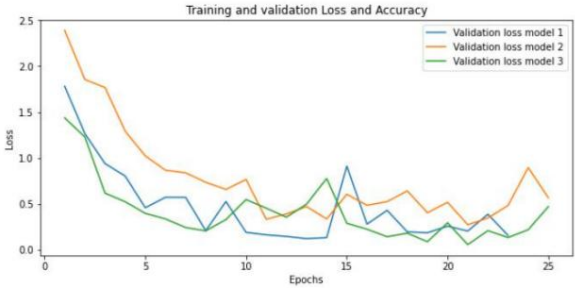
A. Hiệu suất mô hình

Hình 5 so sánh độ chính xác của quá trình đào tạo và kiểm tra tổng thể đối với các kết hợp khác nhau của các lớp ẩn. Độ chính xác phân loại cao nhất trên các hình ảnh thử nghiệm được tìm thấy là 96% đạt được bởi mô hình 3. Điều này cho thấy hiệu suất lớn nhất của mạng trên bộ thử nghiệm xảy ra liên tiếp ở 18 đến 23 kỷ nguyên. Mặt khác, mô hình 2 cũng thu được độ chính xác nhận dạng cao nhất trên các hình ảnh đào tạo, là 94% ở kỷ nguyên 18 đến 23 kỷ nguyên. Đào tạo CNN bằng cách kết hợp các lớp ẩn và tăng số lượng kỷ nguyên dẫn đến độ chính xác của bài kiểm tra cao nhất so với độ chính xác đào tạo.



Hình 5: Đào tạo đường cong độ chính xác cho các kỷ nguyên khác nhau kèm theo các kết hợp

Hình 6 cho thấy tổn thất cho các kỷ nguyên khác nhau đối với các kết hợp khác nhau của tích chập và gộp. Các phản hồi khác nhau được tìm thấy trong sự kết hợp khác nhau, mô hình 3 cho thấy sự mất mát hiệu suất ít hơn so với mô hình khác.

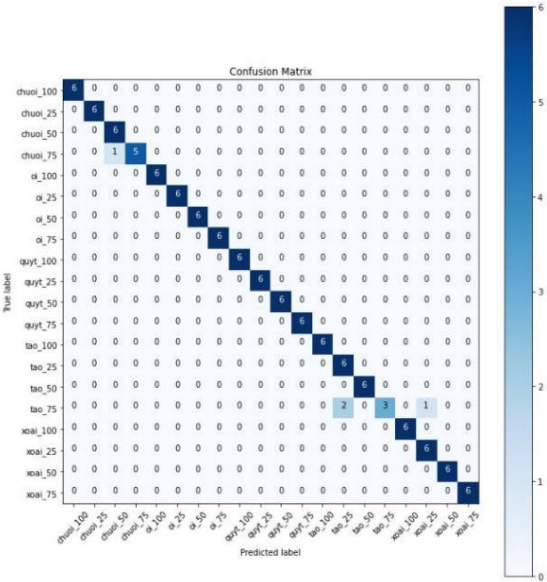


Hình 6: Đường cong suy hao cho CNN trong các kỷ nguyên khác nhau đi kèm với sự kết hợp

B. Ma trận nhầm lẫn

Về mặt hiệu suất của mô hình phân loại, ma trận nhầm lẫn được sử dụng rộng rãi. Nó hoạt động trên một tập hợp các tập dữ liệu thử nghiệm trong đó các giá trị thực là quen thuộc. Các mục nhập trong ma trận là tỷ lệ Đúng Dương tính (TP), Tỷ lệ Dương tính Sai (FP), Tỷ lệ Âm tính Đúng (TN), Tỷ lệ Âm tính Sai (FN) cho từng loại tập dữ liệu. Độ chính xác của ma trận nhầm lẫn được tính theo phương trình:

$$= \frac{(+)}{(+ + +)} \quad (4)$$



Hình 7: Ma trận nhầm lẫn từ mô hình của chúng tôi

Trong hình 7, chúng ta có thể thấy rằng nhãn thực và nhãn dự đoán là giống nhau đối với nhiều lớp. Nó có nghĩa là mô hình của chúng tôi là giá trị rất tốt của ma trận nhầm lẫn của nhiều lớp chỉ ra số lượng lớn thì nó có vẻ là kết quả tốt cho dữ liệu thử nghiệm.

C. Hiệu suất của từng lớp

Classification Report				
	precision	recall	f1-score	support
chuoai_100	1.00	1.00	1.00	6
chuoai_25	1.00	1.00	1.00	6
chuoai_50	0.86	1.00	0.92	6
chuoai_75	1.00	0.83	0.91	6
oi_100	1.00	1.00	1.00	6
oi_25	1.00	1.00	1.00	6
oi_50	1.00	1.00	1.00	6
oi_75	1.00	1.00	1.00	6
quyt_100	1.00	1.00	1.00	6
quyt_25	1.00	1.00	1.00	6
quyt_50	1.00	1.00	1.00	6
quyt_75	1.00	1.00	1.00	6
tao_100	1.00	1.00	1.00	6
tao_25	0.75	1.00	0.86	6
tao_50	1.00	1.00	1.00	6
tao_75	1.00	0.50	0.67	6
xoai_100	1.00	1.00	1.00	6
xoai_25	0.86	1.00	0.92	6
xoai_50	1.00	1.00	1.00	6
xoai_75	1.00	1.00	1.00	6
accuracy			0.97	120
macro avg	0.97	0.97	0.96	120
weighted avg	0.97	0.97	0.96	120

Hình 8: Biểu diễn của mỗi lớp

Độ chính xác và thu hồi của bộ dữ liệu thử nghiệm được trình bày trong hình 8. Kết quả cho thấy mô hình có hiệu suất tốt ở tất cả các loại trái cây. Và mô hình có khả năng nhận biết màu sắc quả tốt, có thể xác định được các giống khác nhau của cùng một loại quả. Ví dụ: tỷ lệ thu hồi của Banana\_50% và Apple\_25% lần lượt là 86% và 75%, và tỷ lệ thu hồi của Mango\_75%, Banana\_75% lần lượt là 100% và 83%.

V. KẾT LUẬN

Bài báo này đã trình bày một hệ thống nhận dạng độ tươi của trái cây cho dữ liệu hình ảnh được chụp bởi điện thoại thông minh bằng CNN. Hệ thống nhận dạng bốn loại trái cây: táo, xoài, ổi và chuối, có bốn mức độ tươi. Tổng thể hiệu suất của mô hình phát hiện hình ảnh trái cây thực sự tốt, đạt kết quả mong muốn với độ chính xác khoảng 96%. Phương pháp CNN được áp dụng là một phương pháp rất mạnh mẽ cho các phương pháp tiếp cận máy học nhận dạng thành công các kết quả của hình ảnh cho mô hình được đề xuất và thuật toán áp dụng của CNN được thực hiện thành công để phân loại hình ảnh và

phát hiện đối tượng. Trong tương lai, cơ sở dữ liệu hiện có sẽ được tăng cường với nhiều loại hình ảnh trái cây khác nhau.

VI. NGƯỜI GIỚI THIỆU

[1] S. Marimuthu và S. Mohamed Mansoor Roomi, “Mô hình mờ được tối ưu hóa bằng hạt để phân loại độ chín của chuối,” IEEE Sens. J., vol. 17, không. 15, trang 4903–4915, tháng 8 năm 2017, doi: 10.1109 / JSEN.2017.2715222.

[2] J. Suriya Prakash, K. Annamalai Vignesh, C. Ashok và R. Adithyan, “Bộ phân loại máy vectơ hỗ trợ đa lớp cho ứng dụng thị giác máy” 2012 Int. Lời thú nhận. Mach. Số lượt truy cập Quy trình Hình ảnh. doi: 2012, 197-199, 2012, pp. MVIP 10.1109 / MVIP.2012.6428794.

[3] H. Mureşan và M. Oltean, “Nhận dạng hiệu quả từ hình ảnh bằng cách sử dụng học sâu”, Acta Univ. Sapientiae, Inform., Vol. 10, không. 1, trang 26–42, 2018, doi: 10.2478 / ausi-2018-0002.

[4] F. Garcia, J. Cervantes, A. Lopez, và M. Alvarado, “Phân loại trái cây bằng cách chiết xuất các đặc điểm về màu sắc, hình dạng và kết cấu: Hướng tới ứng dụng cho siêu thị,” IEEE Lat. Là. Trans., Vol. 14, không. 7, pp. 3434–3443, Jul. 2016, doi: 10.1109 / TLA.2016.7587652.

[5] Q. Wang, S. Nuske, M. Bergerman và S. Singh, “Ước tính năng suất cây trồng tự động cho vườn táo”, không. Tháng 6, trang 745–758, 2013, doi: 10.1007 / 978-3-319-00065-7\_50.

[6] B. Kim, YK Han, JH Park và J. Lee, “Cải thiện khả năng phát hiện bệnh dâu tây dựa trên thị lực bằng cách sử dụng mạng lưới thần kinh sâu”, Front. Khoa học thực vật, tập. 11, không. Tháng 1, trang 1-11, năm 2021, doi: 10.3389 / fpls.2020.559172.

[7] JL Rojas-Aranda, JI Nunez-Varela, JC Cuevas Tello và G. Rangel-Ramirez, “Phân loại trái cây cho các cửa hàng bán lẻ bằng cách sử dụng học sâu”, Bài giảng. Tính toán Ghi chú. Khoa học. (bao gồm Công cụ phụ. Bài giảng. Ghi chú Artif. Giới thiệu. Bài giảng. Ghi chú Tin sinh học), vol. 12088 LNCS, trang 3–13, 2020, doi: 10.1007 / 978-3-030-49076-8\_1.