

Ứng dụng thị giác máy tính cho mobile robot thu hoạch nho

Computer vision of the grape harvesting robot

Lê Đức Hạnh*, Lê Đức Đạo

Trường Đại học Bách khoa - Đại học Quốc gia Thành phố Hồ Chí Minh

*Tác giả liên hệ: ldhanh@hcmut.edu.vn

Ngày nhận bài: 23/2/2024 ; Ngày chấp nhận đăng: 4/4/2024

Tóm tắt:

Nghiên cứu này trình bày phương pháp tự động hóa quá trình thu hoạch và phân loại nho trực tiếp trên cây bằng cách kết hợp mobile robot và thị giác máy tính có ứng dụng trí tuệ thông minh nhân tạo. Đầu tiên, thông tin hình ảnh về nho sống và chín được thu thập, sau đó, đưa vào mạng để huấn luyện nhằm phân biệt nho sống và chín trên cây. Kế tiếp, một mobile robot di chuyển trong khu vực để phát hiện vị trí của quả nho thông qua hệ thống stereo vision, công nghệ xử lý hình ảnh và kết hợp dữ liệu đã huấn luyện để phân biệt nho chín và sống. Vị trí của quả nho được tính toán và tay máy gắn trên robot di động tiến hành thu hoạch. Độ chính xác và tính bền vững của hệ thống và thuật toán được kiểm tra, chứng minh thông qua thực nghiệm thực tế với hệ thống mobile robot được tích hợp tay máy gắn trên nó.

Từ khóa: Thị giác máy tính; Xử lý ảnh; Mobile robot; Stereo vision; Cánh tay robot.

Abstract:

This study presents a method to automate the process of harvesting and sorting grapes directly on a tree by combining mobile robots and computer vision with artificial intelligence applications. First, image information about raw and ripe grapes was collected and then fed into the network for training to distinguish between live and ripe grapes on the tree. Next, a mobile robot moves in the area to detect the location of the grapes through a stereo vision system, image processing technology, and trained data combination to distinguish between ripe and raw grapes. The position of the grapes will be calculated, and the manipulator mounted on the mobile robot will be harvested. The accuracy and sustainability of the system and algorithm were tested and proven through practical experiments with a mobile robot system integrated with a manipulator mounted on it.

Keywords: Computer vision; Image processing; Mobile robot; Stereo vision; Manipulator.

1. Giới thiệu

Nho và các sản phẩm từ nho có nhiều tác dụng rất lớn trong đời sống hàng ngày, ngoài việc tạo nên thu nhập cho nông dân, nho cũng là một nguồn thực phẩm đối với con người. Trong quá trình canh tác và trồng trọt, chất lượng nho bị ảnh hưởng bởi nhiều yếu tố, một trong số đó là các hoạt động trước, sau và trong quá trình thu hoạch được coi là ảnh hưởng đáng kể đến chất lượng sản phẩm. Đặc biệt là rượu vang được làm từ nho. Robot thu hoạch nho được mong

đợi như là một công cụ hỗ trợ nhà nông trong việc thu hoạch – một công việc tốn rất nhiều thời gian và công sức [1]. Gần đây, các công nghệ nông nghiệp đang và sẽ phát triển theo mô hình Nông nghiệp 4.0 [2]. Trong Nông nghiệp 4.0, các công nghệ “mới nổi” như robot, internet vạn vật (IoT), trí tuệ nhân tạo và thị giác máy tính được kết hợp với nhau có mục đích chung là tối ưu, nâng cao nền nông nghiệp [3]. Hơn nữa, với sự ra đời của các hệ thống tự hành thông minh đã dẫn đến sự phát triển của những robot nông nghiệp mạnh mẽ, hay còn được gọi

là Agrbot [4], [5], [6]. Agrbot linh hoạt, có thể xử lý tốt những nhiệm vụ do sự đa dạng của cây trồng nông nghiệp. Do đó, giảm được những tác động đến môi trường, đồng thời, tăng nguồn cung cấp lương thực, thực phẩm và cải thiện bền vững nền kinh tế. Các nghiên cứu thường tập trung vào những tương tác của robot với cây trồng có những tính năng xử lý linh hoạt thuộc từng nhiệm vụ cụ thể như làm cỏ, thu hoạch, phun thuốc, cắt tỉa, tưới nước,... Rong và cộng sự [7] phát triển robot thu hoạch cà chua bi, hệ thống bao gồm một bộ “stereo-vision” với mục đích thu thập data cho việc xử lý đầu vào, bộ phận công tác cuối “end-effector”, tay máy, bộ phận thu giữ trái cây và bộ điều hướng. Hoặc với robot thu hoạch trái kiwi qua nghiên cứu của Scarfe và cộng sự [8], robot có khả năng truyền dẫn tín hiệu qua liên kết vô tuyến, tự động điều hướng kết hợp với hệ thống định vị toàn cầu (GPS) và thị giác máy tính. Robot thu hoạch dâu tây của Sogaard và cộng sự [9] là một hệ thống bao gồm một camera (RGB-D), gripper và cánh tay robot được gắn trên một mobile robot. Dự án Agrbot của Feng và cộng sự [10] là một hệ thống robot thu hoạch cà chua bi. Robot bao gồm một stereo camera, bộ phận công tác cuối, cánh tay máy, giỏ đựng cà chua và đường. Với cấu hình như trên, stereo camera cho phép robot phát hiện được cà chua trên cây, có thể nhận biết được cà chua sống hay chín để thu hoạch. Bộ phận công tác cuối để giữ và tách cà chua khỏi cuống. Ngoài ra, nghiên cứu robot có thể di chuyển trong vườn nho với quỹ đạo đã được định sẵn cùng các thanh ray. Nhược điểm của robot là hiệu suất thu hoạch còn thấp, chưa phù hợp cho thương mại và tỉ lệ robot thu hoạch thành công chỉ 83%. Robot của Lopez và cộng sự [11] có thể theo dõi được sức khỏe của cây trồng, xác định đường đi trong vườn nho và phân phối các chất dinh dưỡng vi lượng cho nho, tuy nhiên, robot không di chuyển được nhanh và linh hoạt quanh khu vườn. Dos và cộng sự [12] phát triển robot có thể tự hành và điều hướng, thực hiện các nhiệm vụ giám sát và thu hoạch ở những vườn nho có độ dốc lớn. Robot có thể thực hiện công việc ở các điều kiện

ánh sáng khác nhau như ban đêm hoặc ban ngày. Botterill và cộng sự [13] nghiên cứu một robot tự hành được thiết kế để tự động cắt tỉa cành nho, dùng hệ thống “stereo-vision” để trích xuất mô hình không gian ba chiều (3D) của cây nho và một cánh tay robot với nhiệm vụ cắt tỉa. Coad và cộng sự [14] phát triển robot có tên là VineRobot. Đây là một robot tự hành trong vườn nho, được lắp các cảm biến cần thiết để có thể di chuyển theo lối đi, kiểm tra theo dõi sức khỏe, độ ẩm, sự phát triển của nho, đánh giá điều kiện dinh dưỡng và năng suất của cây. Tất cả những thông tin này được gửi đến màn hình điện thoại hoặc máy tính bảng của người nông dân theo thời gian thực. Từ đó, nông dân có thể chủ động trong việc chăm sóc cây trồng. Tuy đã được ứng dụng trong thực tế nhưng nhược điểm của các robot này là giá thành rất cao.

Để làm chủ công nghệ cũng như hạ giá thành sản phẩm phù hợp với điều kiện Việt Nam, nghiên cứu này đề xuất phương pháp sử dụng mobile robot kết hợp với các camera giá rẻ và mạng trí tuệ thông minh nhân tạo để phát hiện và xác định vị trí của nho trên giàn.

2. Huấn luyện và nhận dạng

2.1. Chuẩn bị dữ liệu

YOLOv5 là một mô hình thuộc mạng CNN cho việc phát hiện, nhận dạng, phân loại đối tượng. YOLO được tạo ra từ việc kết hợp giữa các convolutional layers và connected layers. Trong đó, các convolutional layers trích xuất các feature của ảnh, full-connected layers dự đoán xác suất và tọa độ của đối tượng. Yolo chia một bức ảnh thành $S \times S$ ô lưới (grid cell), thường là 7×7 . Trọng tâm của vật thể được tìm trong các ô lưới đó (được học thông qua việc gán nhãn và đào tạo). Nếu trọng tâm của đối tượng nằm trong ô lưới bất kỳ, thì nó (ô lưới chứa trọng tâm của đối tượng) sẽ chịu trách nhiệm tìm vật thể đó. YOLO dự đoán trong mỗi ô, liệu tồn tại object với điểm trung tâm rơi vào ô. Và dự đoán điểm trung tâm, kích thước và xác suất object đó là object cần xác định. Trong nghiên cứu này, dựa trên khuyến nghị của nông

dân: 1500 hình ảnh nho sống và chín được thu thập, tiến hành gán nhãn thông qua phần mềm RoboFlow như Hình 1.



Hình 1. Quá trình gán nhãn chùm nho chín và sống trên phần mềm RoboFlow.

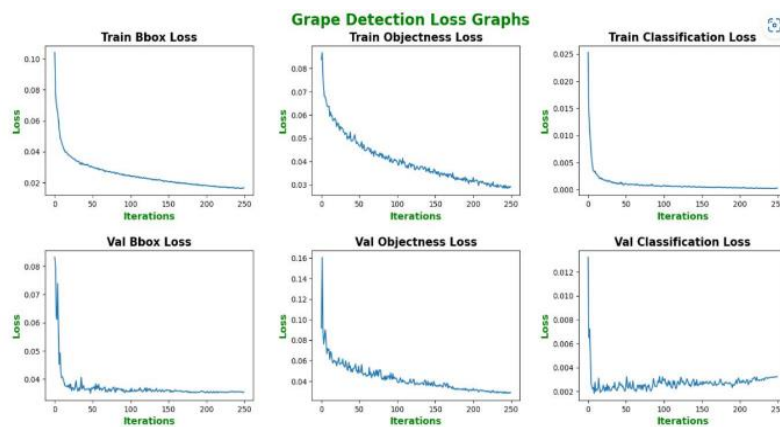
2.2. Đánh giá mô hình huấn luyện

Các dữ liệu sau khi dán nhãn được đưa vào mô hình mạng để huấn luyện, các đường cong mô tả đường đi của hàm loss trong quá trình huấn luyện (train) và xác thực (validation) được thể hiện qua ba hàm mất mát. Giá trị hàm mất mát càng nhỏ, lỗi càng nhỏ và càng nhiều dự đoán

chính xác. Đồ thị “Train Bbox_Loss” đo sai số giữa khung box dự đoán và nhãn dán. Đồ thị “Train Objectness_Loss” đo xác suất khi mục tiêu nho tồn tại trong vùng quan tâm. Đồ thị “Train Classification_Loss” thể hiện khả năng của thuật toán dự đoán chính xác loại nho. Các thông số của mô hình huấn luyện và kết quả của quá trình huấn luyện được trình bày tại Bảng 1 và Hình 2.

Bảng 1. Thông số của mạng huấn luyện.

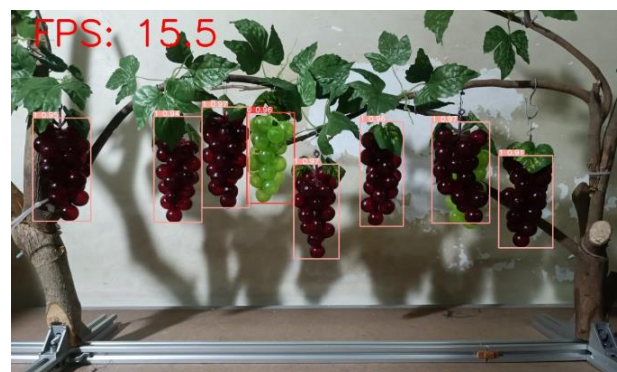
Tham số	Giá trị
Input size	640x640
Batch size	64
Class	2
Epoch	250
Learning rate	0.001
Conf-thres	0.8
IoU- thres	0.5



Hình 2. Các kết quả của quá trình huấn luyện.

Thông qua Hình 2 có thể thấy, đồ thị hàm loss giảm rất nhanh từ 10 đến 50 epoch đầu tiên và cũng hội tụ khá nhanh. Trong giai đoạn này, độ chính xác, thu hồi và mAP tăng nhanh, cho thấy hiệu quả học mô hình cao. Với việc đào tạo lâu hơn (số epoch > 50), độ dốc của đường cong đào tạo giảm dần và về cơ bản ổn định sau 200 chu kỳ đào tạo. Giá trị mất mát dao động quanh mức 0,011. Do đó, đầu ra của mô hình sau 250 chu kỳ đào tạo được xác định đạt được mục tiêu đào tạo.

Kết quả nhận dạng ảnh thật sau khi huấn luyện mô hình được minh họa như Hình 3.



Hình 3. Kết quả của quá trình huấn luyện.

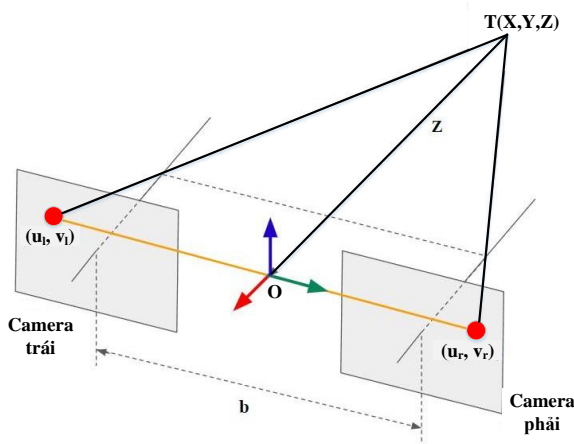
Từ Hình 3 nhận thấy, mô hình có thể nhận dạng được chum nho chín và sổng trên cây có độ chính xác cao, đồng thời, tốc độ nhận dạng một lúc cho nhiều chum nho một cách nhanh chóng với 15.5 FPS, tốc độ này phù hợp với việc thu hoạch nho thật trên giàn.

3. Chuyển đổi tọa độ ảnh sang tọa độ robot

Sau khi nhận dạng, sử dụng kỹ thuật AI, kết quả có được là tọa độ của vật trong hệ tọa độ ảnh. Để tay máy trên robot di động có thể sử dụng, các tọa độ ảnh này cần được chuyển sang hệ tọa độ thật trong hệ tọa độ cánh tay máy của robot. Quá trình này gồm hai bước chuyển đổi hệ trục tọa độ là từ ảnh sang camera gắn trên robot và từ camera sang tay máy.

3.1. Tính toán tọa độ ảnh sang tọa độ camera

Stereo vision là kỹ thuật sử dụng hai camera để đo khoảng cách giữa các đối tượng. Cấu hình đơn giản nhất (cấu hình stereo chuẩn) sử dụng hai camera thẳng hàng và cách nhau theo phương ngang. Sử dụng camera stereo này, có thể thu được hình ảnh của đối tượng tại hai vị trí khác nhau: Ảnh bên trái và ảnh bên phải của đối tượng (sự chênh lệch). Các ảnh của các camera được phân tích để tìm những điểm chung. Sử dụng quy tắc tam giác đồng dạng và độ lệch của các điểm chung để xác định khoảng cách (độ sâu) so với camera như Hình 4.



Hình 4. Cấu hình stereo camera.

Tọa độ chiều (mm) của điểm T trên camera trái:

$$(X_L, Y_L, Z_L) = (X - b/2, Y, Z) \quad (1)$$

Tọa độ chiều (mm) của điểm T trên camera phải:

$$(X_R, Y_R, Z_R) = (X + b/2, Y, Z) \quad (2)$$

Tọa độ điểm ảnh của điểm T trên camera trái và phải:

$$\Rightarrow u_l = \left(X + \frac{b}{2}\right) \frac{f}{Z} \quad u_r = \left(X - \frac{b}{2}\right) \frac{f}{Z} \quad (3)$$

Tọa độ thực điểm T (X, Y, Z) được tính từ tọa độ ảnh pixel:

$$Z = \frac{bf}{u_l - u_r} = \frac{bf}{d} \Rightarrow \begin{cases} X = \frac{b(u_L + u_R)}{2(u_L - u_R)} = \frac{b(u_L + u_R)}{2d} \\ Y = \frac{by}{u_L - u_R} = \frac{by}{d} \end{cases} \quad (4)$$

Với d là sự khác nhau giữa hai điểm ảnh (disparity).

3.2. Tính toán tọa độ camera sang tọa độ tay máy

Do camera được đặt cố định trên mobile robot có gắn tay máy nên khi điều khiển tay máy có thể gặp được vật, cần chuyển đổi hệ tọa độ camera về hệ tọa độ tay máy. Một cách tổng quát, để chuyển một điểm từ hệ tọa độ này về một hệ tọa độ khác, cần tìm ma trận chuyển M . Lần lượt thu thập ít nhất bốn điểm trong tọa độ tay máy tương ứng với tọa độ của camera, từ đó suy ra công thức sau:

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix}_R = \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_X \\ R_{21} & R_{22} & R_{23} & T_Y \\ R_{31} & R_{32} & R_{33} & T_Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix}_C \quad (5)$$

$$\Leftrightarrow P_R = MP_C$$

Trong đó:

i : Có giá trị từ 1 đến 4;

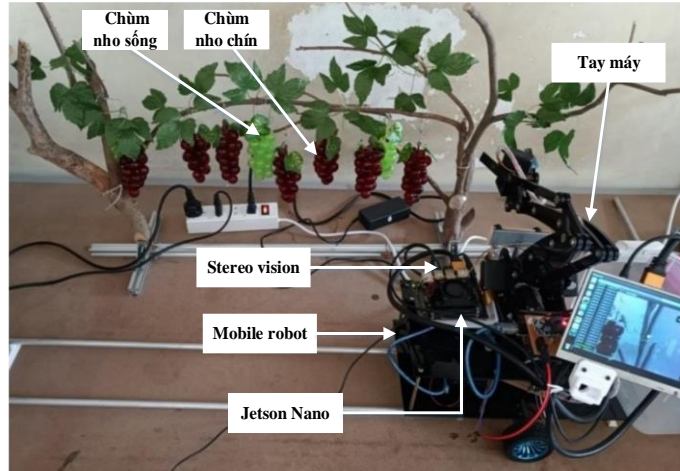
P_R : Tọa độ điểm trung gian đã biết trước trong hệ tọa độ robot;

P_C : Tọa độ điểm trung gian đã biết trước trong hệ tọa độ camera.

4. Thực nghiệm

Tổng quan không gian thực nghiệm: Thực nghiệm trong môi trường ánh sáng bình thường ban ngày. Khoảng cách giữa mobile và giàn nho là 300 mm. Chiều cao giàn nho là 340 mm. Chiều dài giàn nho là 1000 mm. Số lượng nho chín là 07, nho sống là

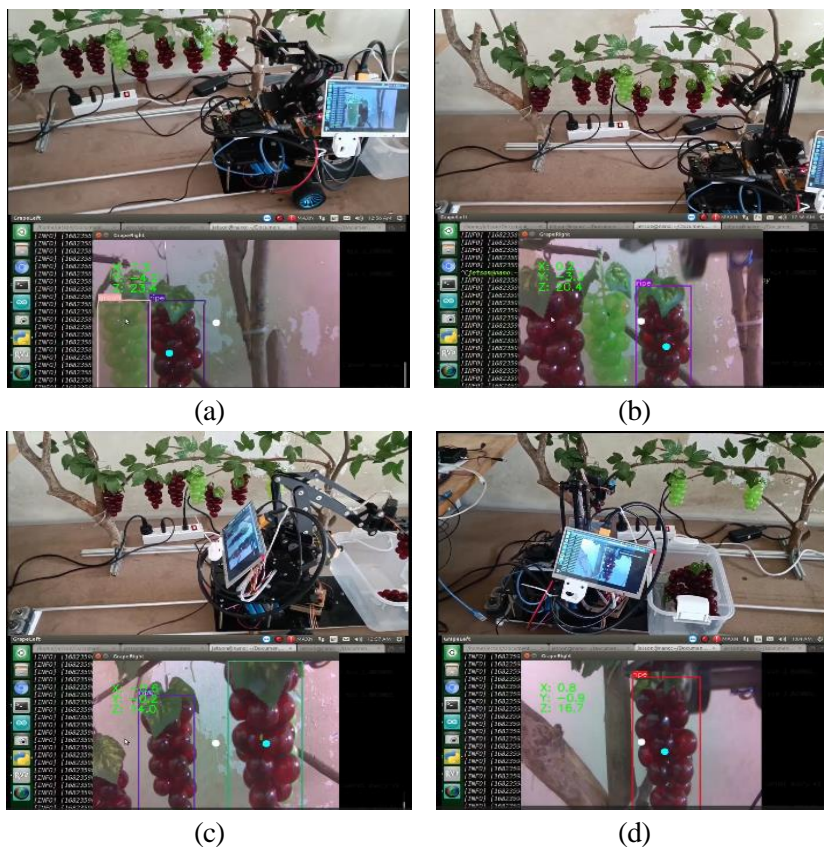
02. Các chùm nho được bố trí ở các vị trí cao, thấp, xa, gần để kiểm tra khả năng thu hoạch nho của robot. Phần cứng của robot bao gồm phần điều khiển sử dụng Jetson Nano B01, camera stereo IMX219-83, tay máy ba bậc tự do để thu hoạch chùm nho như Hình 5.



Hình 5. Môi trường thực nghiệm.

Quá trình nhận dạng và thu hoạch nho trên giàn như sau: Bằng cách sử dụng mạng neural đã được huấn luyện robot đi vào khu vực thu hoạch và nhận dạng chùm nho như Hình 6a, sau đó, hệ thống tính toán vị trí của chùm nho chín và

truyền về cho tay máy gắn trên robot tiến hành cắt cuống như Hình 6b. Robot tiến hành bỏ chùm nho vào khay đặt trên robot như Hình 6c. Quá trình thu hoạch cứ thế tiếp diễn cho đến khi hết chùm nho chín trên giàn.



Hình 6. Quá trình robot nhận dạng và thu hoạch nho trên giàn.

Nhận xét:

- Robot có khả năng thu hoạch nhỏ ở các vị trí khác nhau trên giàn nho;

- Qua quá trình thực nghiệm, thấy được robot thu hoạch được 07 chùm nho trong vòng 4,5 phút. Như vậy, trong vòng một giờ, robot thu hoạch được 105 chùm nho với khối lượng chùm nho trung bình là 50 g, tương đương với 5,25 kg.

Điều này, có thể thấy rằng, robot đáp ứng năng suất đề ra ở đầu bài toán với 5kg/h. Robot có khả năng phân biệt nho chín và nho sống để thu hoạch nhỏ.

5. Kết luận

Nghiên cứu đã trình bày một phương pháp hiệu quả để thay thế người thu hoạch nhỏ, giúp cho quá trình thu hoạch nhanh hơn, hiệu quả hơn. Bằng cách ứng dụng kỹ thuật mới nhất hiện nay là trí tuệ thông minh nhân tạo, chùm nho chín và sống được nhận diện chính xác, dễ dàng. Hệ thống có chi phí thấp, không quá phức tạp, có thể sản xuất ứng dụng hàng loạt. Thời gian cho quá trình nhận dạng và truyền tín hiệu cho robot đáp ứng được yêu cầu tốc độ thu hoạch. Ngoài ra, mô hình vẫn còn nhiều tiềm năng để phát triển như: Thêm các mô hình xác định độ chín, phát hiện bệnh cho cây nho hay kết nối robot với người dùng thông qua các ứng dụng trên điện thoại hoặc máy tính bảng. Từ đó, người dùng có thể theo dõi, giám sát các quá trình hoạt động của robot.

Lời cảm ơn

Nhóm tác giả xin cảm ơn Trường Đại học Bách khoa - Đại học Quốc gia Thành phố Hồ Chí Minh đã hỗ trợ thời gian, phương tiện và cơ sở vật chất cho nghiên cứu này.

Tài liệu tham khảo

- [1] Y. Jiang et al., “Development of a dual-arm rapid grape-harvesting robot for horizontal trellis cultivation,” *Front. Plant Sci.*, vol. 13, 2022, Art. no. 881904, doi: 10.3389/fpls.2022.881904.
- [2] D. C. Rose, R. Wheeler, M. Winter, M. Loble, and C. -A. Chivers, “Agriculture 4.0: Making it work for people, production, and the planet,” *Land Use Policy*, vol. 100, Jan. 2021, Art. no. 104933, doi: 10.1016/j.landusepol.2020.104933.
- [3] T. Duckett et al., “Agricultural robotics: The future of robotic agriculture,” 2018, *arXiv:1806.06762*.
- [4] R. Sparrow and M. Howard, “Robots in agriculture: Prospects, impacts, ethics, and policy,” *Precis. Agric.*, vol. 22, pp. 818-833, 2021, doi: 10.1007/s11119-020-09757-9.
- [5] A. Bechar and C. Vigneault, “Agricultural robots for field operations. Part 2: Operations and systems,” *Biosyst. Eng.*, vol. 153, pp. 110-128, Jan. 2017, doi: 10.1016/j.biosystemseng.2016.11.004.
- [6] K. G. Fue, W. M. Porter, E. M. Barnes, and G. C. Rains, “An extensive review of mobile agricultural robotics for field operations: Focus on cotton harvesting,” *AgriEng.*, vol. 2, no. 1, pp. 150-174, 2020, doi: 10.3390/agriengineering2010010.
- [7] X. Rong, J. Huanyu, and Y. Yibin “Recognition of clustered tomatoes based on binocular stereo vision,” *Comput. Electron. Agric.*, vol. 106, pp. 75-90, Aug. 2014, doi: 10.1016/j.compag.2014.05.006.
- [8] A. J. Scarfe, R. C. Flemmer, H. H. Bakker, and C. L. Flemmer, “Development of an autonomous kiwifruit picking robot,” in *Proc. 2009 4th Int. Conf. Auton. Robots Agents*, Wellington, New Zealand, 2009, pp. 380–384, doi: 10.1109/ICARA.2000.4804023.
- [9] H. T. Sogaard and I. Lund, “Application accuracy of a machine vision-controlled robotic micro-dosing system,” *Biosyst. Eng.*, vol. 96, no. 3, pp. 315-322, Mar. 2007, doi: 10.1016/j.biosystemseng.2006.11.009.
- [10] Q. Feng, W. Zou, P. Fan, C. Zhang, and X. Wang “Design and test of robotic harvesting system for cherry tomato,” *Int. J. Agric. Biol. Eng.*, vol. 11, no. 1, 2018. [Online]. Available: <https://ijabe.org/index.php/ijabe/article/view/2853>.
- [11] A. Lopez-Castro et al., “Design of a vineyard terrestrial robot for multiple applications as part of the innovation of process and product: Preliminary results,” in *Proc. 2020 IEEE Int. Conf. Eng.*, Boca del Rio, Mexico, 2020, doi: 10.1109/ICEV50249.2020.9289671.

- [12] F. N. dos Santos et al., "Towards a reliable robot for steep slope vineyards monitoring," *J. Intell. Robot. Syst.*, vol. 83, pp. 429-444, 2016, doi: 10.1007/s10846-016-0340-5.
- [13] T. Botterill, "A robot system for pruning grape vines," *J. Field Robot.*, vol. 34, no. 6, pp. 1100-1122, Sep. 2017, doi: 10.1002/rob.21680.
- [14] M. M. Coad et al., "Vine robots," *IEEE Robot. Autom. Mag.*, vol. 27, no. 3, pp. 120-132, Sep. 2020, doi: 10.1109/MRA.2019.2947538.