

THIẾT KẾ VÀ ĐIỀU KHIỂN ROBOT THU HOẠCH KHÓM

DESIGNING AND CONTROLLING A PINEAPPLE - HARVESTING ROBOT

Nguyễn Phạm Thục Anh^{1*}

¹Trường Đại học Bách khoa Hà nội

*Tác giả liên hệ: anh.nguyenphamthuc@hust.edu.vn

(Nhận bài: 18/2/2021; Chấp nhận đăng: 22/4/2021)

Tóm tắt - Bài báo đề cập đến việc phát triển một hệ thống Robot thu hoạch trái khóm tự động. Hệ thống này bao gồm khối thị giác máy, hai tay máy Robot 3DOF dạng Gantry được lắp đặt trên khung công tác có thể di chuyển của máy với cơ cấu tác động cuối được thiết kế chuyên dụng và bộ phận điều khiển thu hoạch dựa trên ảnh phân tích. YOLOv3 (You only look once version 3) - một bộ nhận dạng dựa trên cơ sở mạng Nơ ron tích chập CNN được sử dụng để phát hiện và nhận dạng trái khóm đủ điều kiện thu hoạch. Các kết quả thử nghiệm off -line trên cơ sở dữ liệu 168 ảnh chưa được huấn luyện cho thấy mô hình huấn luyện có độ chính xác trung bình là 90,82%. Hệ thống được tiến hành thử nghiệm thu hoạch tại cánh đồng khóm thuộc tỉnh Tiền giang. Độ chính xác thu hoạch đạt được là 80% với khoảng 25 chu kỳ cắt và 12 giây là thời gian của một chu trình thu hoạch.

Từ khóa - Robot thu hoạch; khóm; học sâu; YOLO; thị giác máy

1. Đặt vấn đề

Khóm là loại quả có vị hương đặc biệt và giàu dinh dưỡng, được ưa chuộng trên thị trường trong nước và ngoài nước. Từ sự thích nghi với đặc điểm sinh thái của vùng đất phèn cùng với hiệu quả kinh tế được khẳng định, cây khóm đã trở thành cây trồng chủ lực tại nhiều vùng đất phèn nhiễm phèn nặng và hoang hóa thuộc đồng bằng sông Cửu long. Một trong những địa phương hàng đầu về canh tác khóm là huyện Tân phước - nơi được coi là trung tâm phèn của tỉnh Tiền giang. Năm 2018 sản lượng khóm của huyện đạt 260.000 tấn/ năm với diện tích canh tác khoảng 16.000 ha, góp phần giảm nghèo và làm giàu cho hàng ngàn hộ nông dân. Tuy nhiên, việc thu hoạch khóm gặp nhiều khó khăn. Phương pháp phổ biến để thu hoạch khóm hiện tại chủ yếu là thủ công. Đây là công việc tiềm ẩn nhiều nguy hiểm cho người công nhân do lá, quả, chồi và tai khóm đều rất sắc và nhiều gai nhọn rất dễ gây thương tích. Mùa khóm chín thường xảy ra lúc nắng nóng và thời gian chín rất nhanh. Điều này dẫn đến phải huy động một lực lượng lao động mùa vụ lớn trong khi lực lượng lao động trẻ có xu hướng chuyển đổi sang làm việc ở các khu công nghiệp lớn. Việc cơ khí hóa và tự động hóa là giải pháp tất yếu cho các khó khăn trên.

Việc cơ khí hóa các hệ thống thu hoạch rau hoặc trái cây đã được nghiên cứu trong nhiều năm trên thế giới. Các hệ thống này thường được phát triển để thu hoạch hàng loạt các loại rau củ được trồng trên diện tích lớn với tốc độ cao. Việc thu hoạch được tiến hành đồng loạt, không lựa chọn chất lượng và độ chín của quả và thường phá hủy cây sau khi thu hoạch. Rõ ràng hệ thống cơ khí hóa nhằm thu hoạch

Abstract - This paper mentions the development of a robotic system to harvest pineapples autonomously. The system contains a machine vision unit, two 3DOF Gantry-type robotic manipulators that mounts on a moveable platform with custom special end-effectors and an image-based harvesting control unit. YOLOv3 (You only look once version 3) detector - that is based on convolutional neural network is used for detecting and recognizing pineapple fruits that meet requirements for harvest. The off-line experiment results on never-trained 168 images demonstrate the success of pineapple recognition with 90.82% mAP. The complete system has been tested in a pineapple field in Tiengiang Province. The success rate of pineapple harvesting is 80% with 25 harvesting cycles and the average time is 12 seconds for a harvesting cycle.

Key words - Harvesting robot; pineapple; deep learning; You Only Look Once (YOLO); computer vision

hàng loạt trái khóm không phải là giải pháp phù hợp, do khóm thường được trồng xen kẽ, các quả khóm chín không đều. Hơn nữa, một cây khóm sẽ ra trái khoảng 3 lần, vì vậy cần có lựa chọn trái chín để thu hoạch nhưng không làm tổn hại đến cây là một trong các yêu cầu cơ bản. Tự động hóa và nông nghiệp chính xác là xu hướng tất yếu của nông nghiệp. Thu hoạch có chọn lọc yêu cầu sự tích hợp giữa công nghệ Robot và thị giác máy trong hệ thống thu hoạch tự động. Để tiếp cận chính xác đối tượng thu hoạch, hệ thống Robot cần được trang bị để có khả năng phát hiện, nhận biết và ước lượng khoảng cách tới đối tượng đó. Sự phát triển của công nghệ AI đã dẫn đến sự phát triển chưa từng có của các hệ thống Robot và máy thu hoạch chọn lọc tự động các loại hoa, quả, rau khác nhau. Máy thu hoạch tự động đầu tiên được phát triển bởi Schretz và Brown [1]. Trong những năm gần đây, các nghiên cứu về các Robot thu hoạch tự động được công bố, bao gồm Robot thu hoạch trái chanh [2], Robot thu hoạch táo [3], Robot thu hoạch anh đào [4], Robot thu hoạch dâu [5]. Mỗi loại quả có những đặc điểm riêng nên các Robot thu hoạch cần đáp ứng các yêu cầu cụ thể để có thể thực hiện hiệu quả công việc của mình. Mặc dù, có rất nhiều các công bố trong hàng chục năm qua nhưng cho đến nay trên thế giới chưa có Robot thu hoạch trái cây nào trở thành sản phẩm thương mại trên thị trường. Môi trường làm việc không xác định trước trên thực địa là bài toán khó khăn đối với các Robot thu hoạch, trong đó Robot thu hoạch khóm dường như là một trong những hệ thống khó khăn nhất và chưa nhận được nhiều sự quan tâm nghiên cứu, một phần đây là loại trái cây nhiệt đới, chỉ hợp với thổ nhưỡng ở một số lượng

¹ Hanoi University of Science and Technology (Anh P.T. Nguyen)

hạn chế các quốc gia trên thế giới. Rất ít nghiên cứu liên quan về Robot này trên thế giới và hầu hết chỉ quan tâm đến bài toán phát hiện và nhận dạng trái khóm trong phạm vi phòng thí nghiệm [6, 7]. Một hệ thống Robot thu hoạch khóm hoàn chỉnh đã được thiết kế và chạy thử nghiệm trong phòng thí nghiệm có kết quả tốt [8], tuy nhiên chưa đủ điều kiện để đưa ra hoạt động trên cánh đồng khóm.

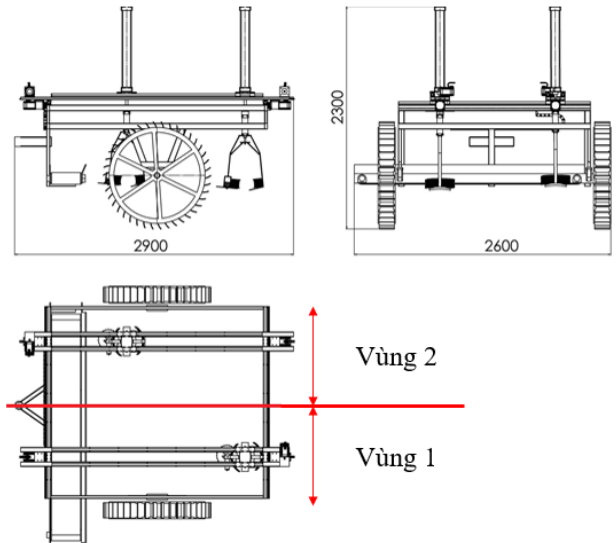
Bài báo đề cập đến việc phát triển một Robot thu hoạch khóm hoàn chỉnh trong môi trường thực tế là những cánh đồng khóm. Bộ phận thị giác máy có nhiệm vụ phát hiện các quả khóm đồng thời xác định vị trí của chúng và gửi về bộ phận điều khiển để thực hiện việc di chuyển các cơ cấu thu hoạch tiếp cận và cắt trái khóm khỏi thân cây. Cơ cấu thao tác là hai Robot dạng Gantry gồm ba bậc tự do được truyền động bởi ba động cơ bước. Để nhận dạng trái khóm, YOLO (You only look once) - một phần mềm nhận dạng dựa trên cơ sở mạng Nơ ron tích chập CNN được tích hợp với hệ thống cơ cấu tay cắt. Các kết quả thử nghiệm cho thấy sự hiệu quả của mô hình huấn luyện nhận dạng thu hoạch cũng như độ chính xác của hệ thống điều khiển trong thu hoạch.

2. Các thiết bị chính của Robot thu hoạch khóm

2.1. Lựa chọn cấu hình Robot

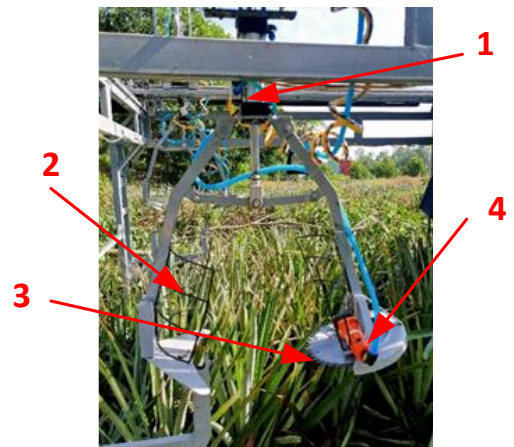
Cấu hình Robot được đề xuất thông qua việc tìm hiểu về canh tác và khảo sát tại các cánh đồng khóm tại xã Mỹ Phước - tỉnh Tiền Giang. Khóm được trồng trên liếp có chiều rộng khoảng 2m. Dựa trên một số tài liệu về canh tác cây khóm, để đảm bảo năng suất cao nhất, cây khóm (loại Queen) được canh tác theo hàng rộng, mỗi hàng cách nhau khoảng 40 - 50 cm, khoảng cách giữa hai quả khóm khoảng 40 - 50 cm. Cây khóm trưởng thành có thân cao khoảng từ 100 - 120 cm, đường kính tán rộng từ 30 - 50 cm. Không giống như phần lớn các loại cây ăn quả mọc trong các chùm lá (như cam, táo, dâu...), khóm được mọc lên từ gốc như các loại rau củ trên cánh đồng. Vì vậy, cấu hình Robot 3 bậc tự do dạng Đề các (3DOF PPP Gantry Robot) là phù hợp nhất để thực hiện nhiệm vụ thu hoạch khóm. Robot gồm ba khớp trượt, có thể chuyển động tịnh tiến nhờ ba khớp trượt theo các trục x -, y -, z của hệ trục tọa độ Đề các. Để nâng cao năng suất cắt, hai Robot được lắp đặt trên khung công tác, mỗi Robot sẽ đảm nhiệm cắt $\frac{1}{2}$ vùng thu hoạch và không xung đột vùng hoạt động nhờ các công tắc hành trình lắp đặt trên khung công tác. Hệ thống đầu tiên được thiết kế và thử nghiệm trong phòng thí nghiệm chưa tính chọn đầy đủ đến các yêu cầu thực nghiệm nên chưa đạt được thông số động học phù hợp, độ cứng vững của hệ thống chưa cao dẫn đến hạn chế về độ chính xác thu hoạch tại thực nghiệm [8]. Việc tính toán thiết kế khung công tác cần phải dựa trên kích thước của thửa khóm trung bình và độ cao lớn nhất của cây khóm. Cụ thể, hệ thống mới được xây dựng với khung công tác cao 1,5m, chiều rộng 2,6m và chiều dài 2,9 m. Khung công tác được lắp bánh xe để di chuyển trên thực địa dễ dàng. Để phù hợp với điều kiện làm việc trên cánh đồng, các động cơ cũng như hệ dẫn động cần được lựa chọn phù hợp. Hai động cơ bước với hệ thống dẫn động dạng con lăn được sử dụng để di chuyển cơ cấu thu hoạch theo phương x . Một động cơ bước với dẫn động bánh răng - thanh răng truyền động cơ cấu chuyển động

theo phương y . Chuyển động đi xuống theo trục z của tay cắt được thực hiện nhờ động cơ khí nén.



Hình 1. Cấu hình Robot thu hoạch khóm

2.2. Cơ cấu tác động cuối



Hình 2. Cơ cấu tác động cuối

1. Xylanh khí nén; 2. Lòng đỡ; 3. Dao cắt; 4. Động cơ quay dao

Cơ cấu tác động cuối hay còn gọi là tay cắt là thiết bị cơ khí quan trọng, là bộ phận trực tiếp làm nhiệm vụ thu hoạch trái khóm. Các quả khóm được mọc lên từ gốc, cuống quả khóm thông thường dài 30 - 50 cm. Khối lượng của quả khóm (dạng Queen) khi trưởng thành khoảng 0,8 - 1,5 kg. Các quả khóm đạt tiêu chuẩn về kích cỡ và màu sắc sẽ được thu hoạch. Để cắt khóm, người nông dân sẽ dùng dao cắt cuống, khoảng từ 5 - 10 cm từ đáy quả khóm. Vết cắt phải phẳng và đặc biệt tránh dập và gãy chồi và cuống. Mỗi cây khóm sẽ cho ra trái ba lần, các quả ra sau sẽ nghiêng với các góc lớn hơn theo phương thẳng đứng. Khóm sau khi thu hoạch cần được vận chuyển về kho sớm do đây là loại trái nhạy cảm với nhiệt độ. Các đặc tính này cần được quan tâm trong việc thiết kế cơ cấu cắt khóm của Robot cắt khóm tự động.

Tay cắt được gắn với thanh nối cuối của Robot, thực hiện chuyển động đi xuống để tiếp cận quả khóm được lựa chọn và cắt cuống chính xác mà không làm hư hỏng quả

khóm và cây khóm. Đây là một công việc khó khăn do một số không nhỏ quả khóm nằm thấp hơn hẳn các tán lá và lá khóm rất cứng cản trở chuyển động của tay cắt. Do vậy, tay cắt thu hoạch được thiết kế chuyên dụng, bao gồm hai phần chính là một lồng đỡ và dao cắt. Lồng cắt được thiết kế theo kích thước quả khóm lớn nhất, được truyền động đóng mở nhờ cơ cấu khí nén. Khi thanh nối cuối của Robot đi xuống, lồng cắt cũng mở ra và ép những chiếc lá bao quanh quả khóm xuống, để lộ cuống khóm cho dao cắt phía dưới lồng dễ dàng thực hiện việc cắt cuống, khoảng 7 -10cm từ đáy quả khóm lên. Sau khi cắt xong, lồng đỡ vẫn duy trì đóng để giữ quả khóm bên trong, tay cắt được điều khiển để quay về vị trí ban đầu, lồng kẹp mở ra và thả quả khóm về hệ thống băng tải.

2.3. Bộ điều khiển thu hoạch

2.3.1. Phân cứng của bộ điều khiển thu hoạch



Hình 3. Khối điều khiển thu hoạch sử dụng các PLC

Cấu trúc phân cứng của khối điều khiển được minh họa trên Hình 3. Phần tử cốt lõi của khối điều khiển là một máy tính công nghiệp có nhiệm vụ tích hợp các ngoại vi điều khiển và các phần mềm ứng dụng để điều khiển và giám sát toàn bộ hoạt động thu hoạch khóm. Nhiệm vụ chính của khối là điều khiển hai tay cắt chuyên động để tiếp cận quả khóm được nhận dạng và đến lượt theo thứ tự cắt. Vị trí của quả khóm trong mặt phẳng ảnh và các độ sâu ảnh để tính toán được biến đổi và tính toán thành số xung của các động cơ bước để di chuyển tay cắt trong mặt phẳng OXY song song với bề mặt ruộng và độ tịnh tiến của tay cắt đi xuống tiếp cận quả khóm. Các thông tin về chuyển động này sẽ được máy tính chuyển xuống hai PLC qua giao thức truyền thông RS-485, hoạt động của mỗi tay cắt sẽ do một PLC đảm nhiệm. Đầu ra của PLC nối với các drivers truyền động cho các động cơ bước di chuyển tay cắt trong mặt phẳng ngang, các van khí nén truyền động cơ cấu lồng kẹp và động cơ cắt cuống khóm.

2.3.2. Thuật toán điều khiển của quá trình cắt

Hoạt động của máy thu hoạch được thực hiện qua các bước như sau: Bắt đầu chu trình, các modul của máy sẽ được kích hoạt. Máy tính khởi tạo truyền thông với các PLC và kích hoạt mô hình nhận dạng trái khóm được huấn luyện từ trước. Khi các tay cắt được kiểm tra để đảm bảo đã về vị trí gốc và sẵn sàng cho việc thu hoạch, các PLC sẽ gửi tín hiệu yêu cầu chụp ảnh vùng khóm tới máy tính để phát lệnh cho camera. Ảnh chụp sẽ được gửi về máy tính và hiển thị trên $\frac{1}{2}$ cửa sổ màn hình. Khối xử lý ảnh sẽ tự động phát hiện và nhận dạng các quả khóm xuất hiện trong ảnh. Các quả khóm được nhận dạng ra sẽ được bao quanh

bởi hình chữ nhật và được hiển trên $\frac{1}{2}$ cửa sổ màn hình còn lại. Tọa độ của các quả khóm trong ảnh sẽ được tính toán và quy đổi thành tọa độ trong hệ trục tọa độ gốc gắn trên khung công tác và được lưu lại trong bộ nhớ máy tính. Thuật toán xác định thứ tự thu hoạch các quả khóm được nhận dạng sẽ được thực hiện và tại một thời điểm sẽ chỉ có tọa độ của một quả khóm được gửi tới PLC khi có tín hiệu tay cắt ở vị trí gốc sẵn sàng cho thu hoạch. Việc gửi tọa độ này sẽ tuần tự và diễn ra cho đến khi toàn bộ các tọa độ được hết xuống các PLC. Máy tính sẽ tự động phát lệnh chụp ảnh lần hai để đảm bảo có thể thu hoạch được những quả bị che khuất bởi các quả được thu hoạch trước. Sau khi thực hiện xong hai lần, xe đầu kéo sẽ được di chuyển tới vùng khóm tiếp theo. Máy tính sẽ lại tiếp tục chờ tín hiệu sẵn sàng từ PLC gửi đến để tiếp tục chu trình mới.

2.4. Khối thị giác máy

Khối thị giác máy là phần tử cốt yếu của các robot thu hoạch. Khối này bao gồm Camera và phần mềm xử lý ảnh. Chức năng chính của khối này là lấy ảnh vùng khóm và xử lý ảnh, cụ thể là phát hiện và nhận dạng các quả khóm trong ảnh nhận được từ camera, sau đó tính khoảng cách từ camera tới quả khóm và thông qua các biến đổi tính chính xác tọa độ của các quả khóm nhận dạng trong hệ trục tọa độ gốc để chỉ dẫn cho tay cắt của Robot có thể tiếp cận và thu hoạch chúng. Có thể nói bài toán phát hiện và nhận dạng quả khóm rất quan trọng, đòi hỏi sự hỗ trợ của các công nghệ xử lý ảnh. Trước khi công nghệ học sâu (deep-learning) ra đời, các phương pháp xử lý ảnh truyền thống sử dụng các đặc trưng như màu sắc, hình dạng, kích thước, góc cạnh, kết cấu và các công nghệ phân loại được sử dụng cho nhận dạng một số loại quả [9 -11]. Tuy nhiên, các công nghệ dựa trên các đặc trưng này không thực sự hiệu quả với môi trường nông nghiệp nhiều biến đổi, điều kiện ánh sáng không ổn định, các quả bị che lấp bởi lá, các quả mọc theo chùm... Để vượt qua các hạn chế này, đòi hỏi những công nghệ tổng quát hiệu quả có khả năng bền vững và linh hoạt với các điều kiện biến đổi của môi trường và biểu diễn các đặc trưng với độ chuẩn xác cao. Trong những năm gần đây, công nghệ deep-learning là một giải pháp tiên tiến cho việc học các đặc trưng ảnh. Các mạng Nơ-ron tích chập (CNN) được phát triển và ứng dụng cho nhận dạng các đối tượng. Các mạng điển hình có thể kể đến bao gồm YOLO (You Only Look Once) [12], SSD (Single Shot Detector) [13], và Faster R-CNN [14]. Các mạng Nơ-ron này đều cho các kết quả tốt khi được ứng dụng nhận dạng các loại rau và quả [15, 16]. Trong nghiên cứu này, nhóm tác giả sử dụng YOLO phiên bản 3 cho việc phát hiện, nhận dạng và phân loại trái khóm. YOLO là mạng nơ-ron tích chập thông minh nhận dạng vật thể với tốc độ cao và có khả năng nhận dạng trong thời gian thực với độ chính xác cao. YOLO được huấn luyện trên cơ sở hàm tổn thất liên quan trực tiếp đến đặc tính huấn luyện. Thuật toán “You only look once” trên ảnh cần nhận dạng đòi hỏi chỉ một lần lan truyền thuận trên toàn bộ mạng để dự đoán. Kết quả đưa ra sẽ là vật thể được nhận dạng với bounding boxes. Thuật toán áp dụng một mạng nơ-ron cho ảnh toàn bộ, và sau đó chia ảnh thành từng vùng và dự đoán các bounding boxes và xác suất cho mỗi vùng. Các bounding boxes này được đánh giá trọng số bởi xác suất dự đoán. Nhóm tác giả

huấn luyện dữ liệu với YOLO phiên bản số 3 (YOLOv3) sử dụng kỹ thuật chuyển giao tri thức (transfer learning).

2.4.1. Phần cứng

Trong môi trường thị giác máy, các camera là thiết bị cảm biến để cung cấp thông tin về đối tượng và môi trường làm việc. Trong hệ thống này, camera công nghiệp D415i được gắn cố định trên khung công tác của máy thu hoạch và có nhiệm vụ chụp ảnh vùng làm việc. Các ảnh chụp này sẽ được chuyển tới máy tính chủ thông qua công USB type C để thực hiện các quá trình xử lý tiếp theo.

2.4.2. Phân tích ảnh và nhận dạng trái khóm

Khi ảnh chụp vùng khóm cần thu hoạch từ camera gửi về máy tính chủ, khối thị giác máy sẽ tiến hành phát hiện các quả dưa xuất hiện trong khung ảnh, tạo các đường bao hình chữ nhật xung quanh các quả khóm được phát hiện và thực hiện việc tính tọa độ của chúng trong hệ trục tọa độ Robot. Như đã đề cập ở trên, YOLOv3 được sử dụng để phát hiện trái khóm. Cấu hình của YOLOv3 được trình bày chi tiết trong [12]. Quá trình huấn luyện được thực hiện qua các bước như sau:

Bước 1. Thu thập dữ liệu cho huấn luyện

Một trong những công đoạn quan trọng của deep - learning đó là thu thập cơ sở dữ liệu cho huấn luyện. Điều kiện cốt yếu để nhận được mô hình huấn luyện tốt đó là cơ sở dữ liệu đủ lớn và chất lượng cao. Trong nghiên cứu này, nhóm tác giả đã thu thập rất nhiều mẫu khóm trên cánh đồng trong nhiều khoảng thời gian khác nhau từ 9 giờ sáng đến 5 giờ chiều. Tổng cộng số lượng ảnh thu được khoảng 840 ảnh chia thành 12 tập con theo điều kiện sáng khác nhau, mỗi ảnh trung bình có 16 trái khóm với các loại khác nhau: quả xanh, quả chín và quả non. Để đảm bảo tính chính xác cao cho thử nghiệm trên hệ thống thực, camera trong quá trình thu thập ảnh cũng chính là camera lắp trên mạch thực nghiệm với vị trí và độ nghiêng đồng nhất với mạch thực nghiệm. Các ảnh chụp được chuyển về lưu tại máy tính gồm một ảnh JPEG với độ phân giải 1280x720 và một ảnh độ sâu.

Bước 2. Gán nhãn.

Khi cơ sở dữ liệu đã sẵn sàng, bước tiếp theo đó là gán nhãn cho các trái khóm có mặt trong ảnh. Các ô chữ nhật sẽ được vẽ bao quanh những trái khóm có mặt trong ảnh cùng với phân loại của nó. Trên cơ sở quan sát bằng mắt thường và dựa trên kinh nghiệm của nhóm tác giả về màu sắc và kích thước của quả khóm. Về cơ bản, quả khóm chín là khi vỏ khóm chuyển từ màu xanh sang màu vàng, khoảng 1/3 độ dài quả kể từ dưới lên. Cụ thể, ở đây trái khóm được phân loại thành ba loại: Trái chín, trái xanh và trái non được gán theo các nhãn ripe pineapple, green pineapple và baby pineapple tương ứng. Công cụ sử dụng để gán nhãn là labelImg (<http://github.com/tzutalin/labelImg>). Sau khi gán nhãn xong, chúng ta sẽ thu được bộ cơ sở dữ liệu gồm ảnh chứa các trái khóm đã được gán nhãn, đi kèm với mỗi ảnh là một file.xml chứa thông tin của những trái khóm trong ảnh bao gồm tên lớp gán cho trái khóm, tọa độ theo trục x - và y - (theo pixel) tương ứng của tâm hình bao chữ nhật, chiều rộng và chiều cao (theo pixel) tương ứng của hình bao chữ nhật, chiều cao và chiều rộng (theo pixel) tương ứng của toàn bộ ảnh. Các file.xml sau đó cần được chuyển về

dạng file.txt để phù hợp với việc huấn luyện mạng.

Bước 3. Tiền xử lý

Dữ liệu sau đó được đưa vào Google Drive. Các ảnh được chuyển về kích thước 416x416 pixels nhưng vẫn giữ nguyên tỷ lệ bằng cách đưa thêm các viền đen. Thuật toán K -means được sử dụng để lựa chọn số anchor -boxes. Thuật toán K -means hoạt động theo vòng lặp:

Bước 1: Chọn ngẫu nhiên k điểm làm tâm của k cụm.

Bước 2: Cập nhật cụm cho từng bounding box: Mỗi bounding box thuộc về cụm có tâm gần nó nhất.

Bước 3: Kiểm tra khoảng cách trung bình từ bounding box đến tâm cụm của nó. Nếu giá trị < ngưỡng, chuyển đến bước 5.

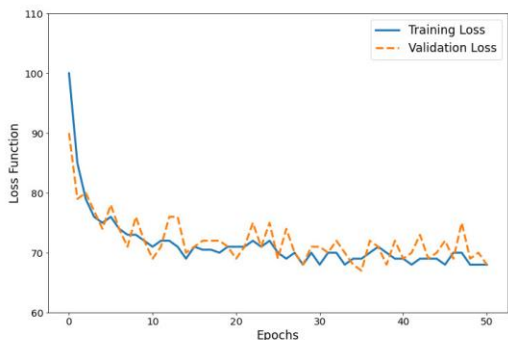
Bước 4: Cập nhật tâm cụm: Tâm mỗi cụm = trung bình tọa độ (w, h) của các điểm trong cụm. Quay lại 2.

Bước 5: Anchor boxes là tâm các cụm.

Nhóm tác giả sử dụng Google Colaboratory để huấn luyện cơ sở dữ liệu trên cấu hình GPU Tesla K80. Cơ sở dữ liệu được tải lên Google Drive và đường dẫn chỉ thư mục cơ sở dữ liệu với Google Colaboratory được thiết lập. Các ảnh trong cơ sở dữ liệu được phân chia thành hai tập với tỷ lệ 75:25: (1) training set (train.txt) dùng cho huấn luyện mô hình và (2) validating set (valid.txt) để kiểm tra mô hình.

Bước 4. Xây dựng mô hình nhận dạng trái khóm sử dụng công nghệ học chuyển giao qua tinh chỉnh (transfer learning via fine-tuning)

Với tất cả các ảnh đã gán nhãn và các file yêu cầu, chúng ta có thể tiến hành xây dựng mô hình nhận dạng bằng YOLOv3 sử dụng công nghệ học chuyển giao thông qua tinh chỉnh. Để khởi tạo base network cho mô hình, Darknet53.conv.74 được sử dụng làm pretrained -model. Trước tiên, các lớp tích chập CNN đầu được đóng băng để các trọng số của nó không đổi và chỉ huấn luyện trên hai lớp kết nối hoàn toàn (fully-connected layers). Mục đích của việc đóng băng các lớp CNN là đảm bảo không phá hủy các đặc trưng bậc cao đã được học từ pretrained -model trên cơ sở dữ liệu lớn và có độ chính xác cao hơn so với khởi tạo hệ số ngẫu nhiên. Pha huấn luyện đầu tiên được thực hiện qua 30 epochs với tốc độ học (learning rate) là $1e-3$ với kích thước mẻ (batch size) là 32. Mục đích của pha huấn luyện đầu là đạt được ổn định của hàm tổn thất, giúp mô hình hội tụ về tối ưu toàn cục. Sau khi mô hình đạt ngưỡng tối ưu trên các fully-connected layers, rất khó để chúng ta tăng độ chính xác cao hơn nữa. Với pha huấn luyện đầu tiên, chúng ta có thể nhận dạng ra các trái khóm. Tuy nhiên, để phân loại các trái khóm, tiếp đó chúng ta cần thực hiện quá trình huấn luyện thứ 2 - finetuning. Trong quá trình này, các lớp tích chập CNN của base network được phá băng và quá trình huấn luyện được thực hiện trên toàn bộ các layers của mạng qua 60 epochs với tốc độ học là $1e-4$ với kích thước mẻ là 8. Quá trình huấn luyện kết thúc khi hàm tổn thất hội tụ về ngưỡng sau khoảng 20 epochs. Kết thúc quá trình huấn luyện, trọng số của mô hình sẽ được ghi vào file.h5 trên máy tính chủ và chúng ta có thể sử dụng mô hình huấn luyện trên để nhận dạng và phân loại trái khóm.



Hình 4. Đồ thị hàm loss sau khi huấn luyện trên toàn mạng

Bước 5. Kiểm tra

Với mô hình huấn luyện vừa nhận được, chúng ta có thể thực hiện công việc nhận dạng trái khóm trên các tấm ảnh hoặc các video. Ở đây nhóm tác giả sử dụng 168 ảnh chụp trái khóm, các ảnh này đều không dùng trong quá trình huấn luyện. Mô hình nhận dạng sẽ dự đoán các trái khóm, vẽ vùng bao hình chữ nhật (bounding -box) quanh các trái khóm, phân loại các trái khóm và xác định độ tự tin về dự đoán (predicted confidence score).



Hình 5. Ảnh nhận dạng khóm trên ảnh với các quả khóm được khoanh bằng các hình chữ nhật với phân loại và độ tự tin

Hình 4 chỉ ra rằng, các quả khóm trong ảnh đã được nhận dạng và bao bởi hình chữ nhật, gán nhãn theo phân loại xanh hoặc chín với độ tự tin tương ứng. Rõ ràng mô hình huấn luyện có khả năng nhận dạng tốt ngay cả khi các quả dứa bị che khuất bởi lá hoặc trái khóm khác. Kết quả nhận dạng off -line trên 168 ảnh được thống kê trên Bảng 1 cho thấy độ chính xác nhận dạng trung bình mAP lên tới 90.82%. Các kết quả này đã tạo tiền đề cho việc thực nghiệm nhận dạng và cắt trái khóm tự động trong phòng thí nghiệm và trên thực địa được trình bày trong Mục 3.

Bảng 1. Kết quả nhận dạng

Lớp	Số trái trên ảnh kiểm tra	Số trái được dự đoán	Số trái dự đoán đúng	Số trái dự đoán sai	Độ chính xác
Trái xanh	554	569	501	68	88,39%
Trái chín	880	951	853	98	93,26%
Độ chính xác trung bình (mAP)					90,82%

3. Thực nghiệm

Khởi thị giác máy được nghiên cứu trong mục trên được tích hợp với hệ thống Robot cắt khóm để thực hiện các thử

thực nghiệm cắt khóm tự động trong phòng thí nghiệm và thực nghiệm trên cánh đồng khóm. Hình 7 mô tả toàn bộ hệ thống tích hợp và được tiến hành thử nghiệm trên cánh đồng khóm thuộc xã Mỹ Phước - huyện Tân Phước - tỉnh Tiền Giang. Camera được gắn trên khung công tác của hệ thống, cách mặt đất khoảng 2,5m, thích hợp để chụp được các ảnh từ trên xuống trong khoảng diện tích 1,8 m x 1,8m của vườn khóm. Quá trình thực nghiệm được tiến hành trong nhiều điều kiện sáng khác nhau. Để đảm bảo quá trình thu hoạch có thể được thực hiện ngay cả vào buổi đêm, hệ thống đèn LED chiếu sáng được bố trí trên khung công tác.



Hình 6. Robot thu hoạch khóm tại thực địa



Hình 7. Kết quả nhận dạng và thu hoạch từ thực địa được đưa từ camera về máy tính chủ (Ảnh phải: Các trái khóm được nhận dạng; Ảnh trái: Hai tay cắt tiếp cận hai trái khóm)

Toàn bộ hệ thống sẽ được di chuyển trên cánh đồng nhờ xe đầu kéo do người vận hành lái. Khi hệ thống khởi động, các tay cắt sẽ được kiểm tra để đảm bảo đã về vị trí xuất phát. Camera chụp và gửi ảnh về máy tính chủ. Mô hình giám sát được tạo lập và màn hình máy tính chủ sẽ gồm hai cửa sổ, một cửa sổ hiển thị ảnh chụp từ camera, một cửa sổ hiển thị kết quả nhận dạng là ảnh tương ứng với các quả khóm được khoanh hình bao hình chữ nhật với phân loại chín, non, hoặc xanh cùng với độ tự tin dự đoán (Hình 7 - ảnh phải). Để đảm bảo việc nhận dạng online, các cửa sổ được cập nhật sau mỗi 0.005 giây. Hai chế độ cắt được lựa chọn tùy theo yêu cầu: chỉ cắt quả khóm chín hoặc cắt tất cả khóm xanh và chín. Trong lần thử nghiệm đầu tiên vào tháng 8/2020, hệ thống đã thực hiện khoảng 50 chu kỳ cắt khóm tự động. Với 385 quả khóm được nhận dạng thành công bởi mô hình huấn luyện, có 308 quả khóm được thu hoạch thành công. Thống kê trung bình hiệu suất hệ thống đạt mức 80%. Trung bình

một chu kỳ cắt một trái khóm khoảng 12 giây kể từ khi tay cắt từ vị trí ban đầu đi đến thu hoạch trái khóm và kết thúc bằng việc thả trái khóm tới băng tải.

Hình 7 (ảnh trái) minh họa hình ảnh hai tay cắt được điều khiển di chuyển tiếp cận hai quả khóm. Có thể thấy, hai quả khóm được chọn nằm gọn trong lòng đỡ để chuẩn bị cho hoạt động cắt cuống. Với kết quả thực nghiệm như vậy, có thể đánh giá hệ thống nhận dạng và điều khiển đã đạt được độ tin cậy ở mức tương đối cao.

4. Kết luận và phương hướng phát triển

Trong nghiên cứu này, nhóm tác giả đã thiết kế và điều khiển hệ thống Robot cắt khóm tự động. Hệ thống bao gồm hai tay cắt đặt trên khung công tác được thiết kế chuyên dụng để phù hợp với việc thu hoạch khóm. Công nghệ deep learning được ứng dụng huấn luyện mạng Nơ-ron nhận dạng trái khóm nhờ công cụ YOLOv3 với độ chính xác trung bình lên đến 90,82%, đảm bảo việc nhận dạng được thực hiện online trong các điều kiện môi trường thay đổi. Tính hiệu quả và độ chính xác của hệ thống nói chung đã được kiểm nghiệm nhờ kết quả khi hệ thống tác nghiệp trên thực địa tại cánh đồng khóm thuộc Tiền giang. Trong các nghiên cứu sắp tới, nhóm tác giả sẽ tập trung nâng cấp phần cơ khí và điều khiển của hệ thống để đạt được độ chính xác và năng suất cao hơn.

Lời cảm ơn: Nghiên cứu này được tài trợ bởi Chương trình Khoa học và Công nghệ Phục vụ phát triển bền vững Tây Nam bộ trong đề tài “Nghiên cứu thiết kế chế tạo một số thiết bị cơ giới hóa, tự động hóa một số khâu trong thu hoạch một số loại cây ăn quả tại vùng Tây Nam bộ” mã số KHCHN -TNB/14 -19/C30.

Tác giả trân trọng cảm ơn các kỹ sư Nguyễn Minh Huy, Đặng Xuân Vương và nhóm Xử lý ảnh đã trợ giúp kỹ thuật cho nghiên cứu này.

TÀI LIỆU THAM KHẢO

- [1] Schertz C E, Brown GK, “Basic considerations in mechanizing citrus harvest”, *Trans. of the ASAE*, pp. 343 -346, 1968.
- [2] Mehta, S.S. T.F. Burksauthor, “Vision -based control of robotic manipulator for citrus harvesting”, *Computers and electronics in agriculture*, v.102, pp. 146 -158, 2014.
- [3] Silwal, A., Davidson, J.R., Karkee, M., Mo, C., Zhang, Q., Lewis, K., “Design, integration, and field evaluation of a robotic apple harvester”, *Journal of Field Robotics*, pp. 1140 -1159, April 2017.
- [4] Tanigaki, K., Fujura, T., Akase, A., Imagawa, J., “Cherry -harvesting robot”, *Computers and Electronics in Agriculture* 63(1), pp. 65 -72, 2008.
- [5] Kondo N., “Strawberry harvesting robot”, *Journal of the Japan Society of Mechanical Engineers* 103(976): 148 -149, 2000.
- [6] Bin Li, Maohua Wang, “In -field recognition and navigation path extraction for pineapple harvesting robots”, *Intelligent Automation and Soft Computing*, Vol. 19, No.1, pp. 99 -107, 2013.
- [7] Bin Li, Maohua Wang, Ning Wang, “Development of a real -time Fruit recognition system for pineapple harvesting robots”, An ASABE Meeting Presentation, 2010.
- [8] Nguyen Pham Thục Anh, Hoàng Sơn, Duong Van Tài, Bui Le Cuong Quoc, “Developing robotic system for harvesting pineapples”, *Proc. of International Conference of Advanced Mechanic Systems*, Hanoi 2020.
- [9] Tabb, A., Peterson, D., Park, J., “Segmentation of Apple fruit from video via background modeling”, *ASABE paper* No. 063060. St. Joseph, Mich.: ASABE, 2006.
- [10] A. J. Scarfe, R. C. Flemmer, H. H. Bakker and C. L. Flemmer, Development of an autonomous kiwifruit picking robot, *Proceedings of the 4th International Conference on Autonomous Robots and Agents*, pp. 380 -384, Feb 10 -12, 2009, Wellington, New Zealand.
- [11] P. Rajendra, N. Kondo, K. Ninomiya, J. Kamata, M. Kurita, T. Shiig, S. Hayashi, H. Yoshida, Y. Kohno. “Machine Vision Algorithm for Robots to Harvest Strawberries in Tabletop Culture Greenhouse”, *Engineering in Agriculture, Environment and Food*, Volume 2, Issue 1, 2009, pp. 24 -30.
- [12] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real -time object detection”, in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 779-788.
- [13] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. -Y. Fu, A. C. Berg, “SSD: Single shot multibox detector”, in *Proc. Eur. Conf. Comp. Vis.* New York, NY, USA: Springer, 2016, pp. 21-37.
- [14] R. Girshick, “Fast R -CNN”, in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 1440-1448.
- [15] Halstead, M., McCool, C., Denman, S., Perez, T., Fookes, C., “Fruit Quantity and Quality Estimation using a Robotic Vision System”, arXiv:1801.05560v1 [cs.CV] 17 Jan 2018.
- [16] Zhang, J., Gao, G., Karkee, M., Zhang, X., He, L. & Zhang, Q. “Branch Detection for Apple Trees Trained in Fruiting Wall Architecture using Depth Features and Regions -Convolutional Neural Network (R -CNN)”. *Computers and Electronics in Agriculture*. Volume 155, 2018, pp. 386 -393.