

SO SÁNH HIỆU SUẤT CỦA MẠNG NƠ-RON PHÂN ĐOẠN NGŨ NGHĨA DEEPLAB V3+ KHI SỬ DỤNG RESNET-50 BACKBONE VÀ MOBILENET V2 BACKBONE TRONG VIỆC XÁC ĐỊNH CÁC BỘ PHẬN THÂN CÀNH LÁ CỦA CÂY CÀ CHUA

Phan Thị Đài Trang¹, Trương Thị Hương Giang¹

Ngày nhận bài: 07/12/2022; Ngày phản biện thông qua: 17/3/2023; Ngày duyệt đăng: 31/5/2023

TÓM TẮT

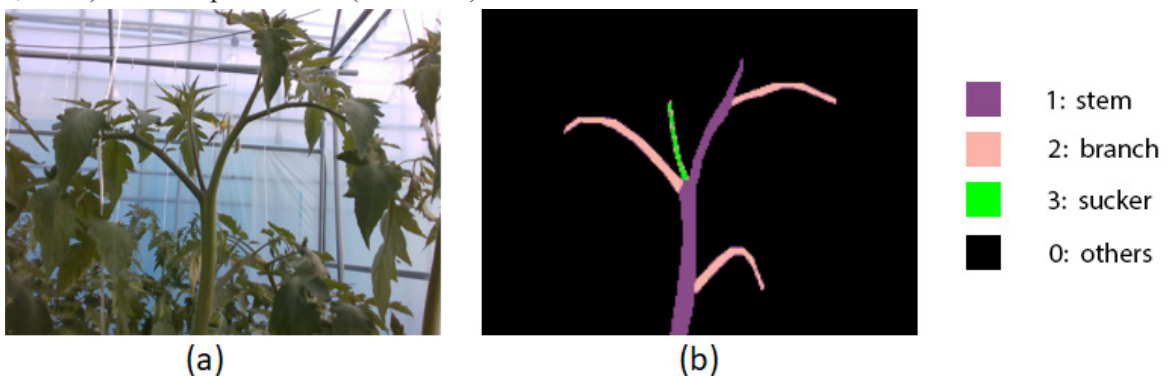
Trong dự án nghiên cứu tự động hóa tại các trang trại thông minh, cụ thể là trồng cây cà chua trong nhà lưới, chúng tôi hướng tới việc tự động cắt tỉa cây cà chua bằng robot. Để đạt được điều này, chúng tôi cần một hệ thống phân tích xác định được vị trí cành, chồi của cây sau đó tiến hành chọn lựa và cắt tỉa. Sau quá trình tìm hiểu, chúng tôi nhận thấy rằng Semantic segmentation neural network (mạng nơ-ron phân đoạn ngữ nghĩa) có thể giải quyết được vấn đề trên. Deeplab V3+ là một trong các mạng nơ-ron có kết quả tốt mà chúng tôi hướng đến sử dụng. Tuy nhiên đây là hệ thống sẽ được thực thi trên máy tính nhúng hoặc máy tính có tài nguyên thấp, chúng tôi cần đánh giá lại hiệu suất cũng như tốc độ thực thi của Deeplab V3+ với backbone là Resnet và MobileNet để có sự lựa chọn phù hợp. Trong nghiên cứu này, chúng tôi sẽ trình bày về đánh giá này trên cùng một dataset về các chỉ số thời gian cũng như độ chính xác trong dự đoán.

Keywords: *Semantic segmentation neural network, Deeplab V3+, cắt tỉa cà chua.*

1. MỞ ĐẦU

Hiện nay việc tự động hóa trong nông nghiệp đã trở thành một xu hướng nổi trội. Với sự phát triển mạnh mẽ của Deep learning (Học sâu) nói chung và Deep Convolution Neural Network (mạng nơ-ron tích chập) nói riêng đã thúc đẩy việc tự động hóa này ngày càng mạnh mẽ. Trong nghiên cứu này chúng tôi chú trọng đến việc chăm sóc cây cà chua trong nhà lưới, mà cụ thể là việc cắt tỉa cây cà chua một cách tự động. Trong hệ thống này chúng tôi sử dụng camera RGB-D để lấy Hình cây cà chua, sử dụng mạng nơ-ron phân đoạn ngữ nghĩa để dự đoán hình Hình xác định các thành phần cây cà chua như trong Hình 1 dưới đây. Trong quá trình nghiên cứu chúng tôi nhận thấy Deeplab V3+ (Chen et al., 2018b) là một trong những mạng nơ-ron phân đoạn ngữ nghĩa có hiệu suất cao. Trong cấu trúc được đề xuất, Deeplab V3+ sử dụng Resnet 50 (He et al., 2016) là thành phần chính (backbone) để trích

xuất đặc trưng hình Hình. Tuy nhiên Resnet-50 làm mạng nơ-ron Deeplab V3+ trở nên nặng nề với số lượng parameter lớn. Từ sự ra đời của MobileNet V2 (Zhou et al., 2020), các mạng phân đoạn ngữ nghĩa nơ-ron có thể thực thi trong thời gian thực trên máy tính có tài nguyên thấp ra đời. Cũng từ nghiên cứu này, Deeplab V3+ cũng có phiên bản mới đó là áp dụng mobileNet V2 làm backbone thay vì Resnet-50. Với phiên bản mới này, số lượng parameter (tham số) của mạng nơ-ron được giảm xuống, tốc độ cải thiện, tuy nhiên hiệu quả trên việc dự đoán sẽ có thể giảm xuống. Thêm vào đó, chưa có báo cáo đánh giá nào về kết quả của áp dụng mạng nơ-ron phân đoạn ngữ nghĩa nhận biết các bộ phận cây cà chua, cũng như so sánh đánh giá giữa hai mạng nơ-ron đã đề cập ở trên. Chúng tôi cần một đánh giá chi tiết trên hai phiên bản này về hiệu năng cũng như tốc độ để có sự lựa chọn phù hợp trong hệ thống cắt tỉa cây cà chua trong tương lai.



Hình 1. Ví dụ về việc tạo hình Hình phân đoạn ngữ nghĩa.

Ghi chú: (a) là hình Hình gốc, (b) là hình Hình được vẽ ra theo ngữ nghĩa, thường được dùng trong huấn luyện mạng nơ-ron.

¹Khoa Khoa học Tự nhiên và Công nghệ, Trường Đại học Tây Nguyên;

Tác giả liên hệ: Phan Thị Đài Trang; ĐT: 0943087474; Email: ptdtrang@ttn.edu.vn.