

Tóm tắt báo cáo

I/ Bài toán

Phân loại thanh long với input gồm 3 ảnh chụp 3 mặt của cùng một quả thanh long có độ phân giải 720p và output là một trong ba loại được miêu tả chi tiết trong báo cáo. Bước tiền xử lý dữ liệu xóa ngoại cảnh (background) sử dụng semantic segmentation model (thử nghiệm mạng UNet và ENet và nhóm thống nhất sử dụng ENet). Quá trình training cho model classify được chia làm 2 giai đoạn trước và sau khi tăng cường dữ liệu, trong đó:

- Giai đoạn 1: Thử nghiệm 3 mạng là InceptionResnetv2, Resnet50 và MobileNetv2. Sau khi đánh giá kết quả, nhóm giữ 2 mạng InceptionResnetv2 và MobileNetv2.
- Giai đoạn 2: Thử nghiệm tiếp InceptionResnetv2 và MobileNetv2 đã được tăng cường dữ liệu. MobileNetv2 độ chính xác kém hơn InceptionResnetv2 2-5% nhưng tốc độ xử lý nhanh hơn khoảng 2.5 lần.

II/ Thống kê dữ liệu

Dữ liệu khoảng 200 quả thanh long chia cho 3 class với mỗi class khoảng 60-70 quả.

- Bộ dữ liệu ban đầu thu thập được khoảng 676 tập dữ liệu trái thanh long (1 tập dữ liệu có 3 ảnh) tương ứng với 2028 ảnh 1280x720x3.
 - o 246 tập thuộc label 1
 - o 210 tập thuộc label 2
 - o 220 tập thuộc label 3
- Dữ liệu sau khi tăng cường bằng cách xoay thủ công trái thanh long trong quá trình quay video thu được khoảng 1120 tập dữ liệu trái thanh long tương ứng với khoảng 3360 ảnh 1280x720x3.
 - o 406 tập thuộc label 1
 - o 412 tập thuộc label 2
 - o 301 tập thuộc label 3

III/ Cập nhật thay đổi so với báo cáo vấn đáp:

- Bổ sung kết quả training với dữ liệu thô (chưa xóa background).

- Bổ sung số lượng trái thanh long đã thu thập dùng làm bộ dữ liệu.
- Bổ sung link demo và cấu hình máy tính sử dụng chạy demo ứng dụng.