

Nghiên cứu, phân loại cam tươi, cam hỏng bằng AI & IoT

Đỗ Hải Anh, Bùi Hải Phong, Nguyễn Thu Huyền, Nguyễn Thị Phương Thảo

Nhóm 1, Lớp CNTT16-06, Khoa Công nghệ thông tin

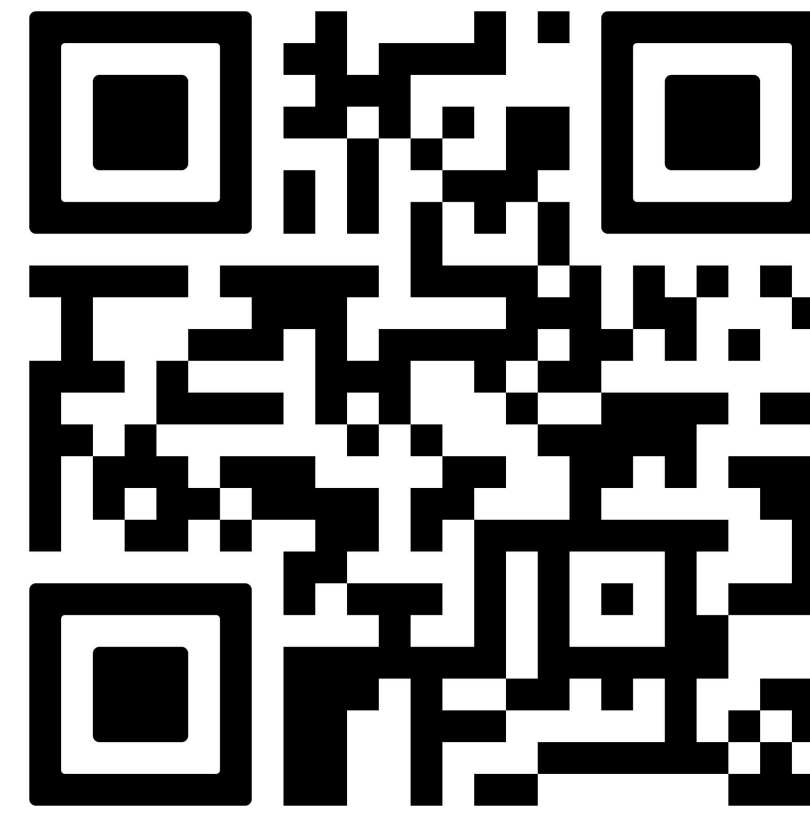
Trường Đại học Đại Nam, Việt Nam

ThS. Lê Trung Hiếu, KS. Nguyễn Thái Khánh

Giảng viên hướng dẫn, Khoa Công nghệ thông tin

Trường Đại học Đại Nam, Việt Nam

Github: https://github.com/kanchou2420/Nghien_cuu_phan_loai_HQ



Giới thiệu

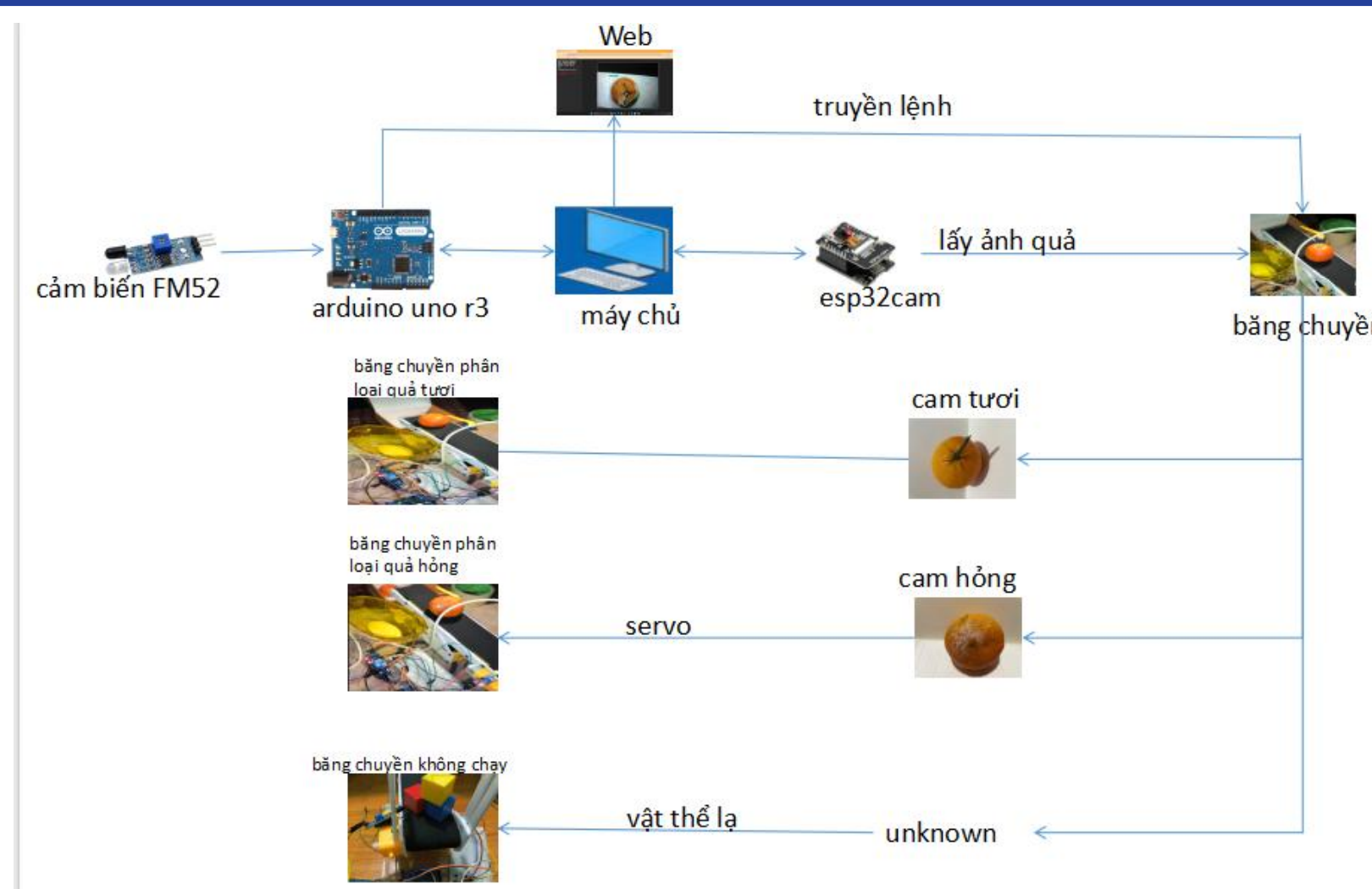
Mô tả vấn đề:

- Sau thu hoạch, việc phân loại cam tươi và cam hỏng chủ yếu dựa vào quan sát thủ công, mất thời gian và dễ sai sót.

Giải pháp:

- Sử dụng YOLOv8 kết hợp AI & IoT để tự động phát hiện, phân loại cam tươi và cam hỏng trên băng chuyền.

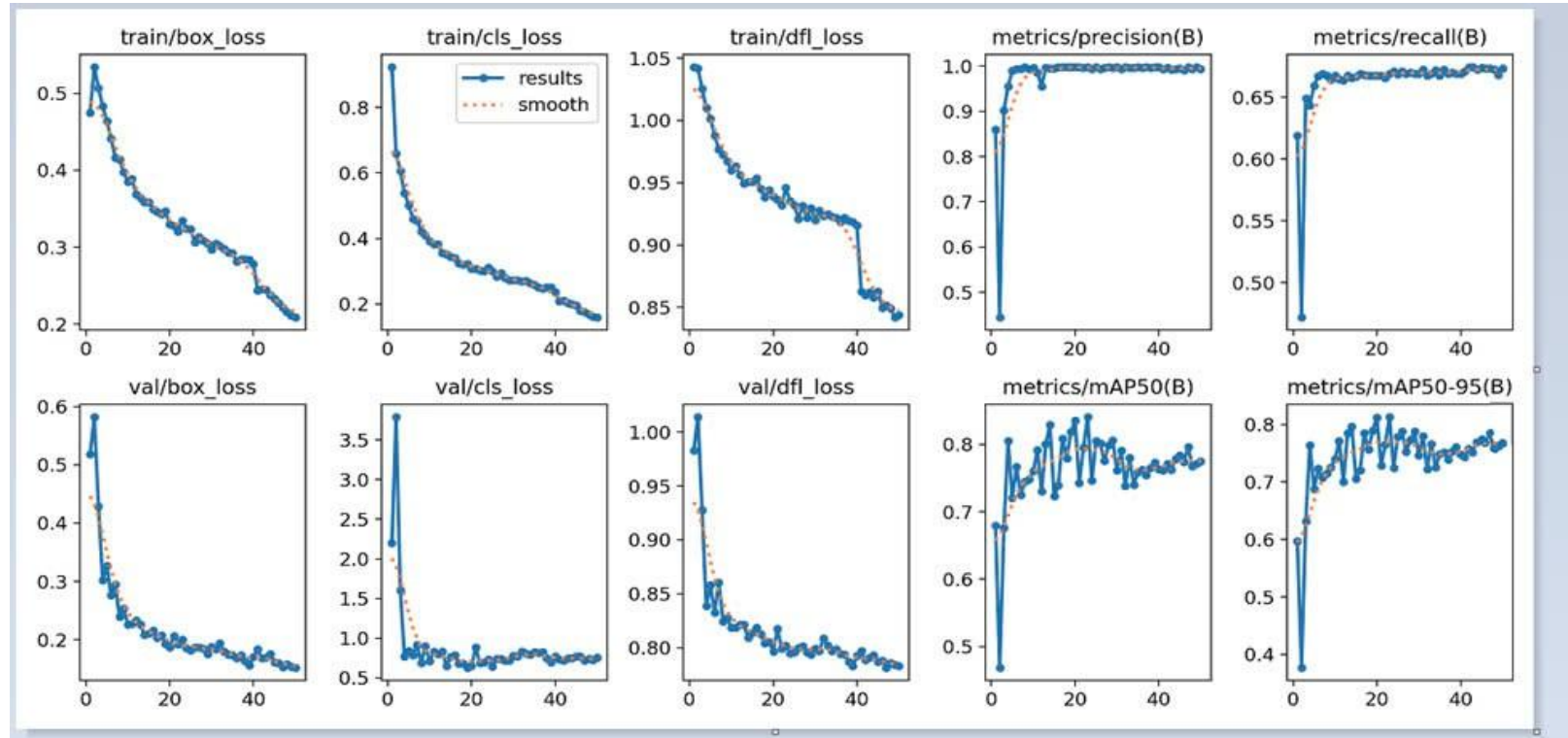
Kiến trúc hệ thống



Phương pháp đề xuất

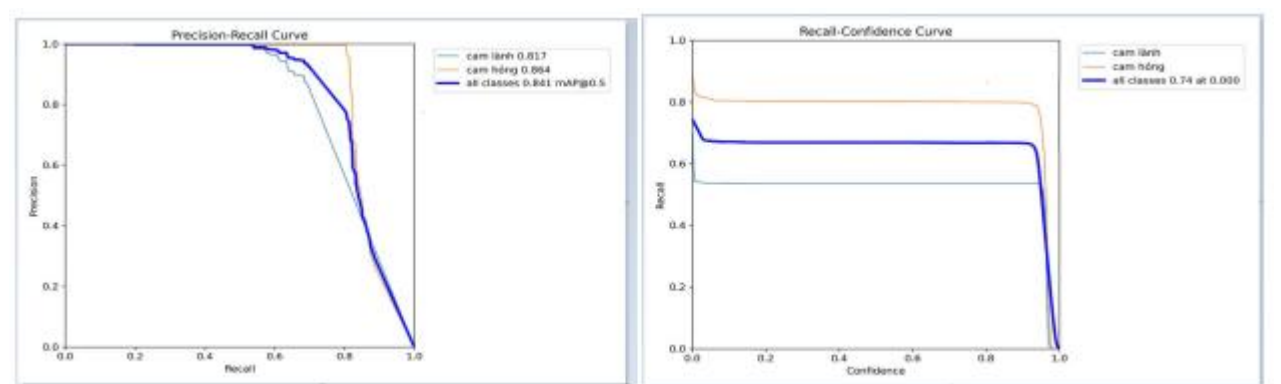
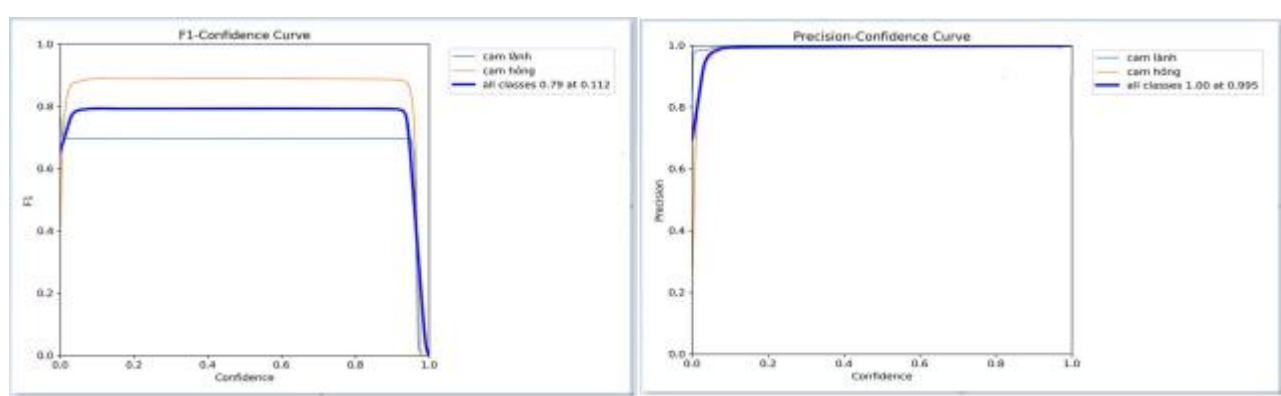
Khung bao gồm 4 thành phần quan trọng:

- Dữ liệu đầu vào: Hình ảnh cam tươi & cam hỏng thu thập từ camera băng chuyền, đa dạng ánh sáng & góc độ.
- Huấn luyện mô hình YOLOv8: Dùng YOLOv8, tối ưu Adam, huấn luyện trên GPU, đánh giá bằng precision, recall, F1-score, mAP.
- Kết nối với IoT: Gửi kết quả đến IoT để loại cam hỏng, hỗ trợ giám sát từ xa qua màn hình.



Quá trình huấn luyện mô hình YOLOv8 và đánh giá hiệu suất:

- Hàng trên (Train Metrics - Chỉ số huấn luyện): train/box_loss, train/cls_loss, train/df_l_loss, metrics/precision(B), metrics/recall(B)
- Hàng dưới (Validation Metrics - Chỉ số đánh giá): val/box_loss, val/cls_loss, val/df_l_loss, metrics/mAP50(B), metrics/mAP50-95(B)



Hiệu suất của mô hình YOLOv8 trong việc phân loại cam tươi (cam lành) và cam hỏng:

a)F1-Confidence Curve b)Precision-Confidence Curve
c)Precision-Recall Curve d)Recall-Confidence Curve

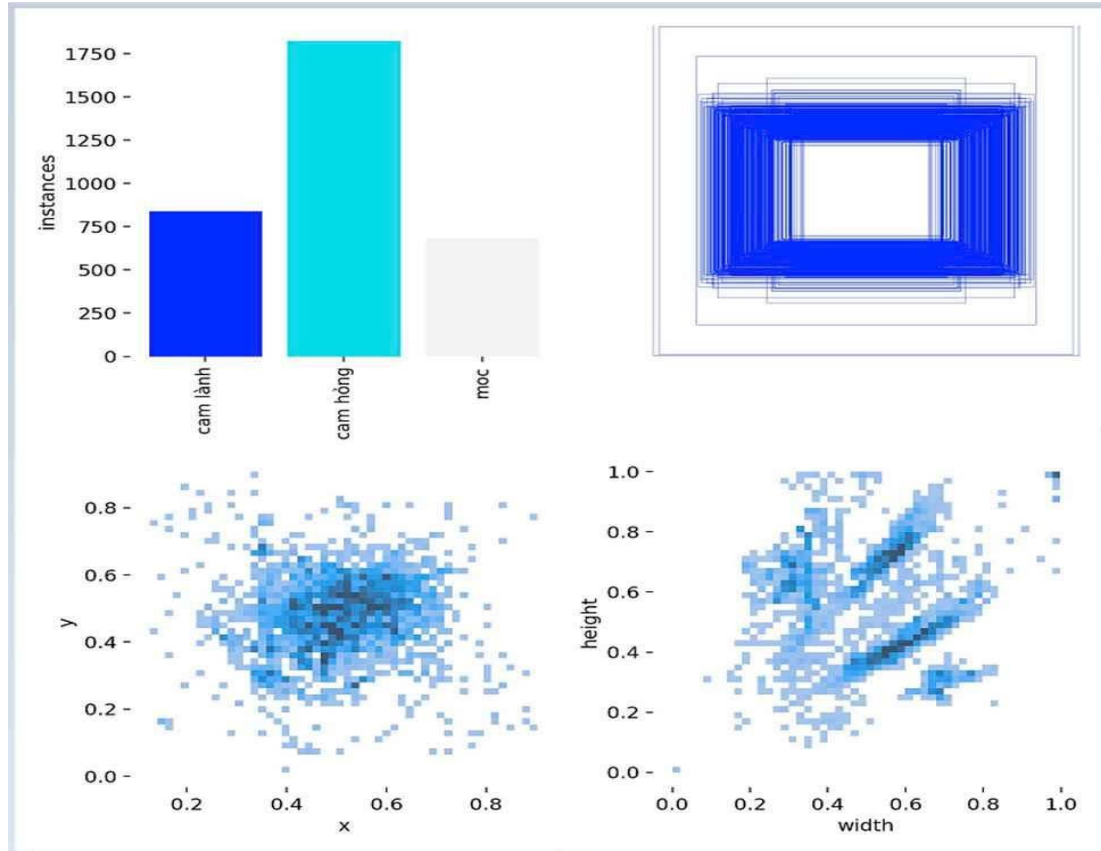
| Mô hình | Số lượng tham số(triệu) | Thời gian suy luận(ms) | Độ chính xác(%) |
|---------------------------------|-------------------------|------------------------|-----------------|
| YOLOv8 | 30 - 40 | 10 - 20 | 85 - 90 |
| YOLOv8 + GADF | 35 - 45 | 15 - 25 | 90 - 93 |
| YOLOv8 + GADF + CoAtNet - light | 25 - 35 | 8 - 15 | 93 - 96 |

Dữ liệu huấn luyện

Bộ dữ liệu các loại cam:

- Dữ liệu được thu thập bao gồm ảnh của cam tươi và cam hỏng.
- Dữ liệu được sử dụng để huấn luyện và đánh giá mô hình YOLOv8 trong việc phân loại cam.

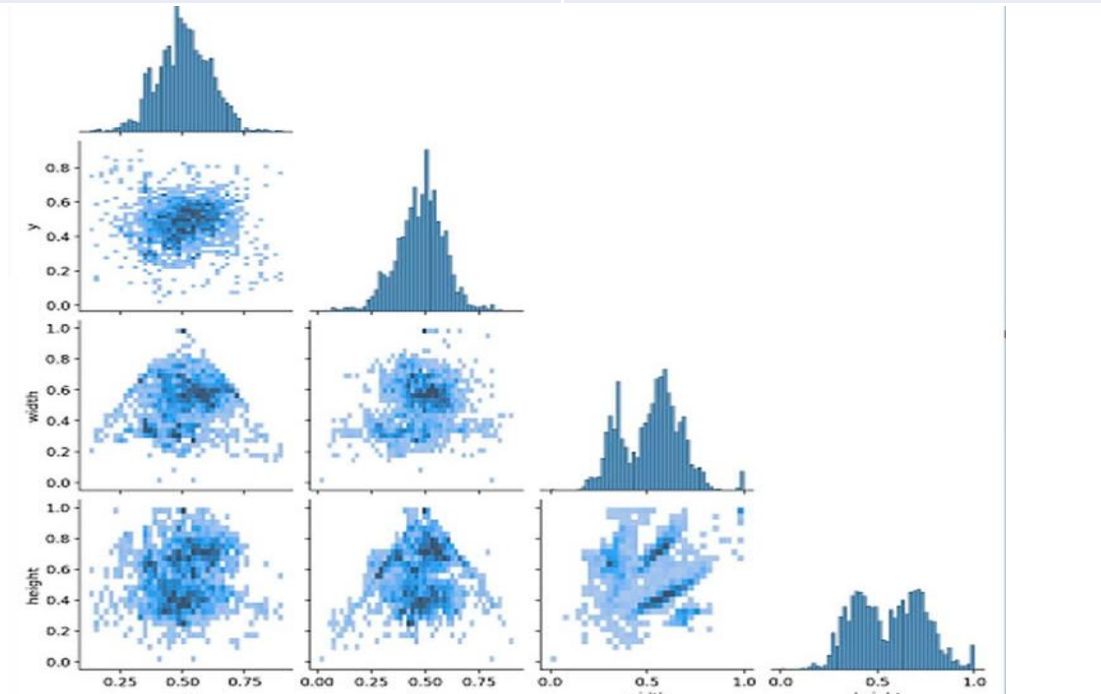
| Thuộc tính | Số lượng |
|----------------------------|----------------------------|
| Loại cam | cam tươi, cam hỏng |
| Số lượng ảnh cam tươi | 750 |
| Số lượng ảnh cam hỏng | 2550 |
| Tổng số ảnh | 3300 |
| Tiêu chí đánh giá cam hỏng | vết thâm, nấm mốc, dập nát |
| Nguồn dữ liệu | tự thu thập |
| Chia tập dữ liệu | 80% train, 20% test |



Bộ dữ liệu cam hỏng:

- Chứa hình ảnh cam hỏng thu thập thực tế, dùng để huấn luyện YOLOv8 phân loại theo mức độ hư hỏng.dựa trên mức độ hư hỏng.
- Bộ dữ liệu này được sử dụng để phát hiện cam hỏng, giúp hệ thống tự động loại bỏ cam không đạt tiêu chuẩn.

| Loại cam hỏng | Số lượng ảnh |
|---------------|--------------|
| Vết thâm nhẹ | 25 |
| Vết thâm nặng | 20 |
| Mốc xanh | 1250 |
| Mốc đen | 1255 |

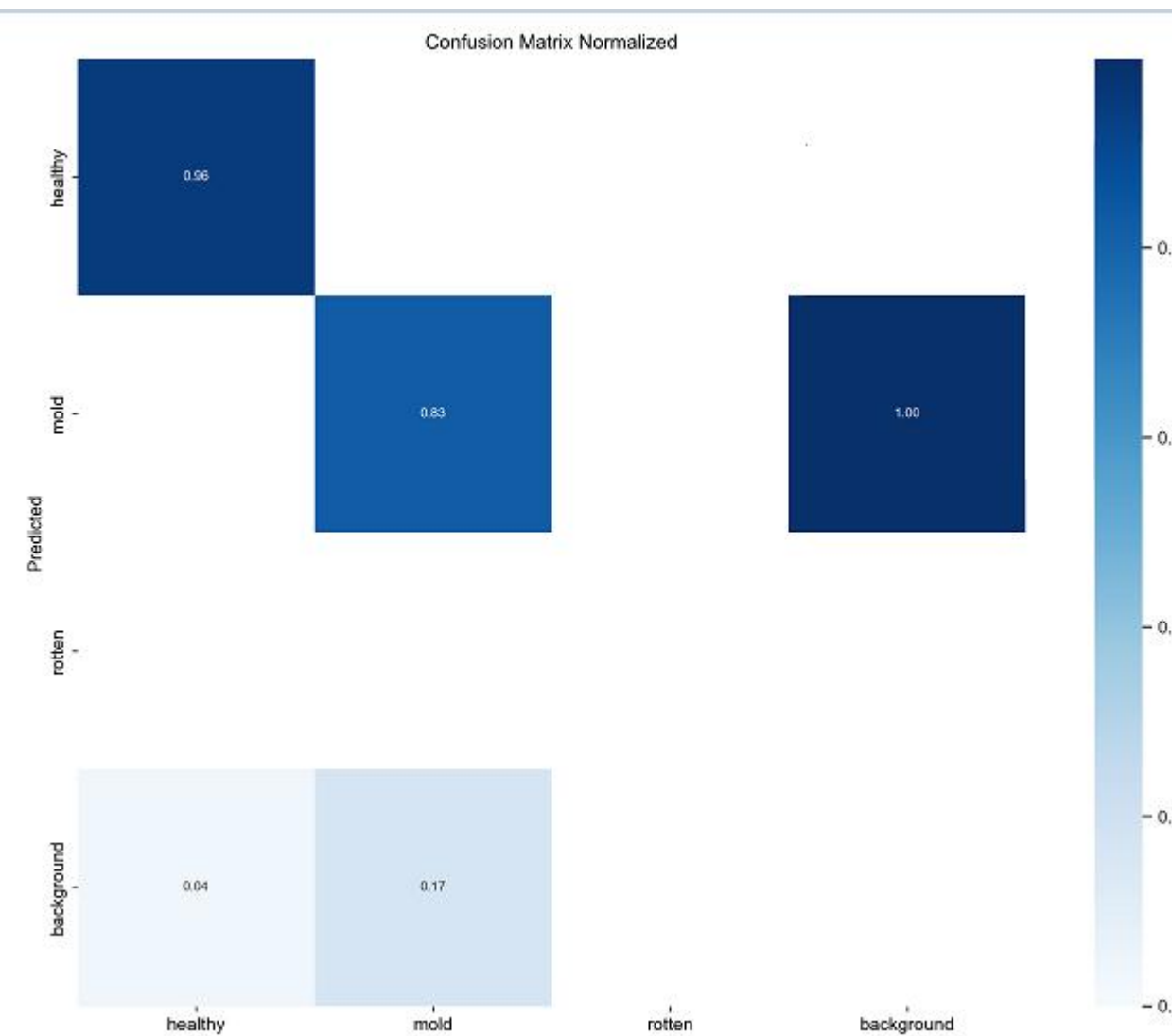


Kết quả

So sánh hiệu suất trên của ma trận nhầm lẫn chuẩn hóa

| Phương pháp | Độ chính xác(%) | Độ chính xác(%) | Độ thu hồi(%) | F1-score(%) |
|---------------------------|-----------------|-----------------|---------------|-------------|
| YOLOv8+ GADF+ Transformer | 98 | 96 | 92 | 97 |

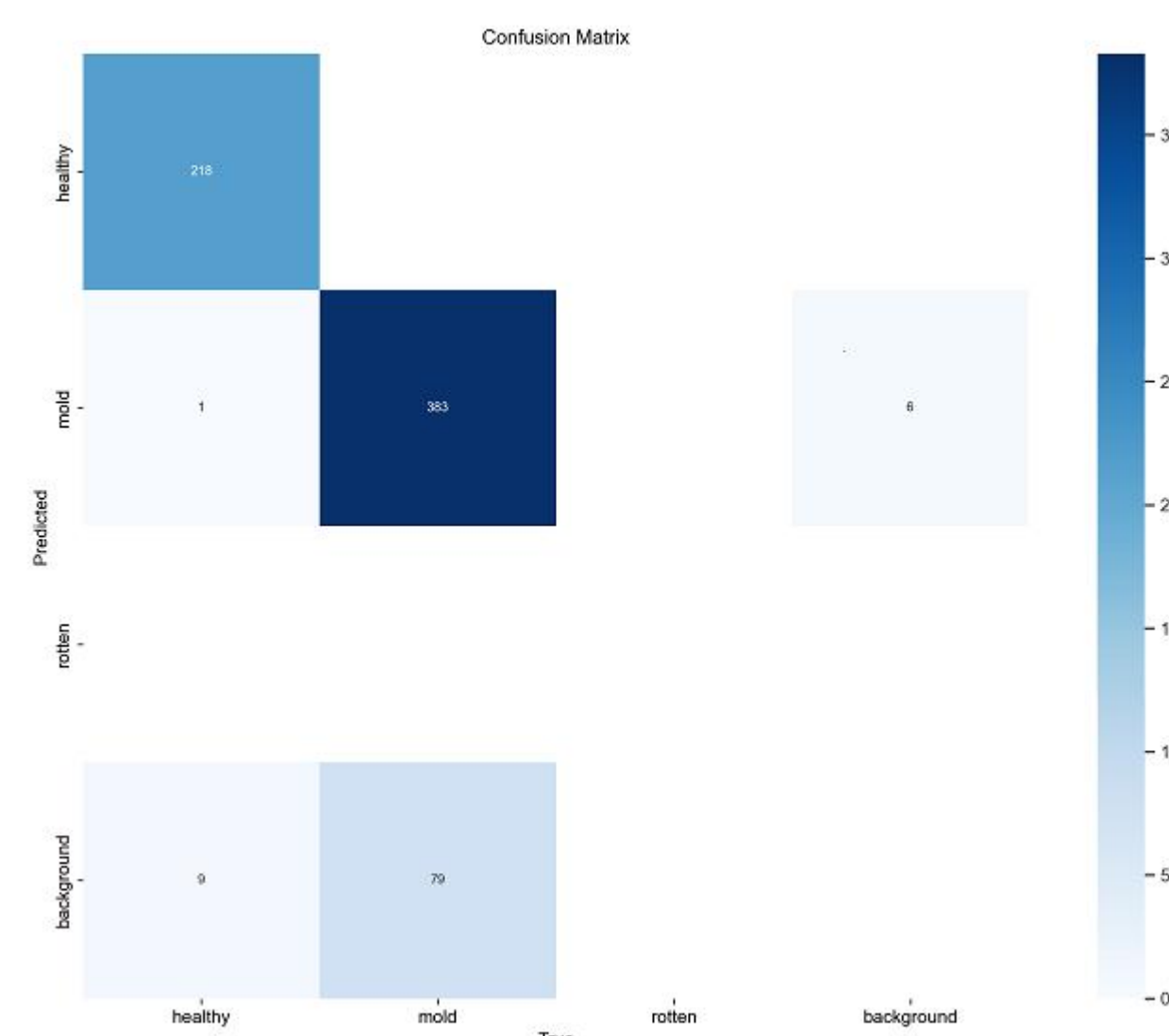
Ma trận nhầm lẫn chuẩn hóa (Normalized Confusion Matrix)



So sánh hiệu suất trên của ma trận nhầm lẫn chưa chuẩn hóa

| Phương pháp | Độ chính xác(%) | Độ chính xác(%) | Độ thu hồi(%) | F1-score(%) |
|---------------------------|-----------------|-----------------|---------------|-------------|
| Baseline | 80 | 82 | 78 | 80 |
| YOLOv8 | 90 | 92 | 88 | 90 |
| YOLOv8+ GADF+ Transformer | 95 | 96 | 94 | 95 |

Ma trận nhầm lẫn chưa chuẩn hóa (Confusion Matrix)



Kết luận & Hướng phát triển

Kết luận:

- Hệ thống AI & IoT ứng dụng YOLOv8 giúp nhận diện, phân loại cam hỏng – cam tươi, tối ưu kiểm soát chất lượng, giảm hao hụt, nâng cao giá trị nông sản và tăng thu nhập cho nông dân.

Hướng phát triển:

- Phát triển AI tiên tiến hơn, tích hợp cảm biến đo chỉ số cam, xây dựng hệ thống quản lý dữ liệu thông minh, và mở rộng phân loại sang các loại trái cây khác.