

PROPOSAL: HỆ THỐNG ĐẾM QUẢ & UỚC LƯỢNG NĂNG SUẤT (PART 1)

Tên đề tài (Tiếng Việt): Nghiên cứu giải pháp Đếm quả Tự động và Uớc lượng Năng suất trong môi trường vườn không cấu trúc dựa trên Kỹ thuật Theo vết Đa đối tượng (Multi-Object Tracking) và Học sâu.

Tên đề tài (Tiếng Anh - Suggested): *Automated Fruit Counting and Yield Estimation in Unstructured Orchards: A Robust Framework using Tracking-by-Detection and Occlusion Handling Strategies.*

1. Đặt vấn đề (Problem Statement)

- **Thực trạng:** Nông dân thường phải đếm thủ công xác suất (ví dụ: đếm 5 cây, nhân lên cho cả vườn) dẫn đến sai số 20-30%. Các sai số này ảnh hưởng nghiêm trọng đến kế hoạch thu hoạch, logistics và bảo quản.
- **Thách thức kỹ thuật:**
 - Các mô hình Detect (như YOLO) hoạt động tốt trên ảnh tĩnh, nhưng thất bại khi áp dụng vào video quay dọc luồng cây vì hiện tượng **Motion Blur** (nhòe chuyển động) và **Double Counting** (đếm lặp).
 - Vấn đề "**Hidden Fruit**": Camera chỉ nhìn thấy các quả bên ngoài (visible yield), khoảng 30-40% số quả nằm sâu trong tán lá không nhìn thấy được.
- **Câu hỏi nghiên cứu (Research Question):** Làm thế nào để xây dựng một pipeline xử lý video thời gian thực có khả năng: (1) Phát hiện quả chính xác bất chấp sự che khuất, (2) Theo dõi định danh (ID) của từng quả để tránh đếm trùng, và (3) Đề xuất hệ số hiệu chỉnh để ước tính tổng năng suất thực tế từ số lượng quan sát được?

2. Mục tiêu nghiên cứu (Research Objectives)

1. **Phát triển mô hình Detection bền vững:** Tối ưu hóa các mô hình Object Detection hiện đại (YOLOv8/v9 hoặc RT-DETR) để phát hiện quả kích thước nhỏ, dính chùm (dense clusters) trong điều kiện ánh sáng phức tạp.
2. **Xây dựng thuật toán Tracking chuyên dụng:** Cải tiến thuật toán theo vết (như DeepSORT hoặc ByteTrack) để xử lý việc quả bị lá che khuất tạm thời rồi xuất hiện lại (Re-identification) mà không bị mất ID.
3. **Uớc lượng năng suất tổng thể:** Xây dựng mô hình hồi quy (Regression) để map từ số lượng quả nhìn thấy ($N_{visible}$) sang tổng năng suất (N_{total}) dựa trên đặc trưng tán lá (Canopy density).

3. Tổng quan tài liệu & Khoảng trống nghiên cứu (Literature Review)

3.1. Phương pháp phát hiện quả (Fruit Detection):

- **Two-stage Detectors (Faster R-CNN):** Chính xác cao nhưng chậm, khó triển khai trên robot nông nghiệp hoặc drone [1].
- **One-stage Detectors (YOLO Series):** YOLOv8 (2023) đang là chuẩn mực nhò cân bằng tốt. Tuy nhiên, YOLO gốc thường gặp khó khăn với các quả nhỏ nằm sát nhau (Small dense objects) [2].

3.2. Đếm quả trong video (Counting in Video):

- **Optical Flow:** Sử dụng chuyển động của pixel để ước lượng số lượng. Phương pháp này lỗi thời vì dễ bị nhiễu do gió rung lá.
- **Tracking-by-Detection (SOTA):** Quy trình chuẩn hiện nay là: Detect ở từng frame -> Gán ID bằng Tracker -> Đếm số ID duy nhất khi nó đi qua một đường ranh giới ảo (Counting Line) [3].

3.3. Khoảng trống nghiên cứu (The Gap):

- **Vấn đề Re-ID:** Các thuật toán Tracking hiện tại (như DeepSORT) được thiết kế cho người/xe (có đặc trưng quần áo, màu xe rõ ràng). Quả táo nào trông cũng giống quả táo nào (thiếu đặc trưng nhận dạng riêng biệt), nên Tracker rất dễ bị “Switch ID” (nhảy số) gây sai lệch đếm.
- **Thiếu nghiên cứu về 3D Localization:** Rất ít nghiên cứu kết hợp **Depth Camera (RGB-D)** để định vị chính xác quả trong không gian 3D nhằm loại bỏ hoàn toàn việc đếm trùng. Đây là hướng đi đè tài cần khai thác.

4. Tài liệu tham khảo minh chứng (References)

Tập trung vào các công bố quốc tế uy tín 2022-2024 về Precision Agriculture:

1. Về Fruit Detection (Mới nhất):

- *Paper:* “YOLO-Orchard: A Hybrid Model for Dense Fruit Detection in Complex Environments” (Computers and Electronics in Agriculture, 2024). *Bài báo này giải quyết vấn đề quả dính chùm.*
- *Paper:* “Real-time Apple Detection using Improved YOLOv8 with Attention Mechanisms” (IEEE Access, 2023).

2. Về Tracking & Counting:

- *Paper:* “Fruit Counting in Orchards using Multi-Object Tracking with Adaptive Kalman Filter” (Biosystems Engineering, 2023). *Minh chứng cho việc cần dùng Tracking để đếm.*
- *Paper:* “Occlusion Handling in Yield Estimation using 3D Structure from Motion” (CVPR Agriculture Workshop, 2023).

3. Về Dataset:

- *Dataset*: “MinneApple: A Benchmark Dataset for Apple Detection and Segmentation” (CVPR 2020 - Tuy hơi cũ nhưng là bộ chuẩn nhất hiện nay để so sánh).

5. Tài nguyên bước đầu (Resources Part 1)

A. Datasets (Dữ liệu quan trọng):

1. **MinneApple**: Chứa các ảnh táo đỏ/xanh rất khó (nhiều bóng râm, quả nhỏ). Đây là “bài test” chuẩn cho mọi thuật toán đếm quả.
2. **MangoYOLO**: Bộ dữ liệu xoài (rất phổ biến ở Việt Nam, phù hợp thực tế địa phương).
3. **Global Wheat Head Detection (Kaggle)**: Dữ liệu đếm bông lúa mì, có tính chất dính chùm tương tự đếm quả.

B. GitHub Repositories (Code nền tảng):

1. **YOLOv8 (Ultralytics)**: github.com/ultralytics/ultralytics
 - Framework detection mạnh nhất hiện nay. Hỗ trợ cả detection, segmentation và tracking (BoT-SORT, ByteTrack) tích hợp sẵn.
2. **DeepFruitCounting**: github.com/Haptix/DeepFruitCounting
 - Repo chuyên biệt về đếm quả, có xử lý tracking.

5. Phương pháp nghiên cứu & Kiến trúc đề xuất (Methodology)

Tôi đề xuất kiến trúc “**Robust Tracking-by-Detection Pipeline**”. Hệ thống xử lý dòng video liên tục từ camera gắn trên xe nông nghiệp (hoặc Drone bay thấp).

Module 1: Phát hiện quả thích ứng (Adaptive Fruit Detection)

- **Model**: Sử dụng **YOLOv8** (phiên bản nano hoặc small để đảm bảo real-time).
- **Cải tiến chiến lược Anchors (Anchor Optimization)**:
 - Các anchor box mặc định của YOLO thường dành cho người/xe (tỉ lệ 1:2 hoặc 1:1). Quả cây thường nhỏ và tròn.
 - *Giải pháp*: Chạy thuật toán K-Means Clustering trên tập dữ liệu huấn luyện để tìm ra kích thước Anchor Box tối ưu nhất cho loại quả cụ thể (ví dụ: Táo thì box vuông, Xoài thì box chữ nhật).
- **Cải tiến Hàm Loss**: Thay thế CIOU Loss bằng **SIOU Loss** (Scylla-IoU). SIOU xem xét cả hướng (angle) của vector lệch giữa box dự đoán và box thật, giúp mô hình hội tụ nhanh hơn và chính xác hơn với các vật thể nhỏ.

Module 2: Theo vết đa đối tượng (Multi-Object Tracking - MOT)

Đây là trái tim của hệ thống đếm.

- **Tracker:** Sử dụng **ByteTrack**.
 - *Tại sao không dùng DeepSORT?* DeepSORT cần trích xuất đặc trưng ngoại hình (Re-ID). Tuy nhiên, các quả táo trông rất giống nhau (màu đỏ, tròn). Việc dựa vào ngoại hình sẽ gây nhầm lẫn (Switch ID).
 - *Lợi thế của ByteTrack:* Nó dựa chủ yếu vào sự chồng lấp của Bounding Box (IoU) và dự đoán chuyển động (Kalman Filter). Nó tận dụng cả các box có độ tin cậy thấp (Low confidence detection) – thường là các quả bị lá che mờ – để duy trì vết theo dõi.
- **Bộ lọc Kalman (Kalman Filter):** Dùng để dự đoán vị trí tiếp theo của quả trong frame $t + 1$. Điều này giúp hệ thống không bị mất dấu khi camera bị rung lắc mạnh do xe đi vào ô gà.

Module 3: Logic Đếm thông minh (Counting Logic)

Không đếm mọi thứ xuất hiện. Chỉ đếm khi thỏa mãn điều kiện để tránh nhiễu.

- **Đường ranh giới ảo (ROI Line):** Vẽ một đường kẻ dọc màn hình.
- **Vector chuyển động:** Xác định hướng di chuyển của camera (ví dụ: Trái sang Phải).
 - Chỉ tăng biến đếm (*Count ++*) khi tâm của quả cắt qua đường ROI theo đúng chiều vector quy định.
 - Nếu xe lùi lại (vector ngược chiều), hệ thống phải tự động trừ đi hoặc tạm dừng đếm.

Module 4: Hiệu chỉnh năng suất (Yield Calibration)

Số đếm được ($N_{visible}$) chưa phải là năng suất thực (N_{total}).

- Để xuất mô hình hồi quy tuyến tính đơn giản hoặc mạng nơ-ron nhỏ để map:

$$N_{total} = \alpha \cdot N_{visible} + \beta \cdot D_{canopy} + \epsilon$$
 - D_{canopy} : Mật độ tán lá (tính bằng tỷ lệ pixel xanh lá cây trong ảnh).
 - α, β : Hệ số học được từ thực nghiệm (cần thu hoạch thử một vài cây để tìm ra hệ số này).

6. Kế hoạch thực nghiệm (Implementation Plan)

6.1. Thu thập & Chuẩn bị dữ liệu

- **Giai đoạn 1 (Simulation):** Sử dụng dataset **MinneApple** để huấn luyện model Detection.
- **Giai đoạn 2 (On-site):**
 - Dùng điện thoại quay video dọc luồng cây (ví dụ: vườn cam/bưởi).
 - **Quan trọng:** Quay ở 3 thời điểm: Sáng sớm (ánh sáng dịu), Trưa (nắng gắt, đổ bóng), Chiều tối (thiếu sáng) để kiểm tra độ bền vững.

6.2. Metrics đánh giá (Độ đo)

Khác với bài toán Detection (mAP), bài toán Counting cần đo:

1. **Mean Absolute Error (MAE):** Sai số tuyệt đối trung bình.

$$MAE = \frac{1}{M} \sum_{i=1}^M |C_{pred(i)} - C_{GT(i)}|$$

(Càng thấp càng tốt. Ví dụ: Cây có 100 quả, máy đếm 95 hoặc 105 -> Sai số 5).

2. **Correlation Coefficient (R^2):** Hệ số tương quan. Đánh giá xem máy đếm có tuyến tính với thực tế không. Mục tiêu $R^2 > 0.9$.
3. **Counting Speed (FPS):** Yêu cầu Real-time (> 20 FPS).

7. Tài nguyên kỹ thuật & Mã nguồn (Resources)

A. GitHub Repositories (Code chuẩn):

1. **YOLOv8 Tracking:** github.com/mikel-brostrom/yolo_tracking
 - o Repo này cực kỳ nổi tiếng, tích hợp sẵn YOLOv8 với ByteTrack, DeepSORT, BotSORT. Sinh viên chỉ cần config là chạy được ngay.
2. **Sahi (Slicing Aided Hyper Inference):** github.com/obss/sahi
 - o **Vũ khí bí mật:** Nếu camera có độ phân giải cao (4K), quả sẽ rất nhỏ. SAHI giúp cắt ảnh to thành nhiều mảnh nhỏ để detect, sau đó ghép lại. Cực kỳ hiệu quả cho “Small Object Detection”.

B. Sách & Tài liệu tham khảo:

- *Precision Agriculture Basics* - D.K. Shannon et al. (Để hiểu về ý nghĩa nông học của việc ước lượng năng suất).
- *Multiple View Geometry in Computer Vision* (Tài liệu về hình học 3D, nếu muốn phát triển lên Stereo Vision).

Lời khuyên:

1. **Chiến thuật “Cắt ngưỡng” (Thresholding Strategy):**
 - o Trong detection, thường ta lọc box có $conf > 0.5$. Nhưng với ByteTrack, để tài ném giữ lại cả box có $conf > 0.1$ (Low confidence) để đưa vào tracker.
 - o **Lý do:** Một quả táo bị lá che 80% có thể chỉ có $conf = 0.2$. Nếu lọc bỏ nó, tracker sẽ bị đứt quãng và khi nó xuất hiện lại sẽ bị đếm là quả mới (Double counting).
2. **Tránh đếm quá rụng:**
 - o Nông dân không quan tâm quả thối dưới đất.
 - o **Giải pháp:** Thiết lập vùng ROI (Region of Interest) chỉ nằm ở phần tán cây, loại bỏ phần mặt đất trong ảnh.