**Bài thực hành về nhà**

*Nguyễn Hồng Đăng*

*Video kết quả đặt tại link google drive:*

<https://drive.google.com/file/d/17mIW-CQf1rwqIrRTqdAiN0M2Vcm1U9b5/view?usp=sharing>

**Nội dung:**

Sử dụng mô hình coconut detection đã thực hành ở Bài 14, kết hợp với thuật toán Kalman filter, hãy xây dựng thuật toán đếm số dừa trên băng chuyền trên video tại link sau:

<https://husteduvn-my.sharepoint.com/:v:/g/personal/cuong_ltb140542_sis_hust_edu_vn/EZxB_icbI-VJkci_Yt80yugBrjDKswPx5KcJ3AluLJ3Yyw?e=o4DeBB>

**1. Mô hình coconut detection**

Ở bài 14, ta đã thử nghiệm thuật toán YOLO và đạt được kết quả tương đối tốt:

- Input: ảnh RGB cần detect, kích thước 720\*1280

- Output: Bao gồm các bounding boxes tương ứng với các quả dừa, dạng (x1,y1,x2,y2) trong đó (x1,y1) là tọa độ góc trái trên (upper-left) và (x2,y2) là tọa độ góc phải dưới (bottom- right).

**2. Mô hình đếm dừa sử dụng thuật toán Kalmal filter**

Sử dụng thuật toán Kalmal filter để tracking sự di chuyển của các quả dừa. Mỗi quả dừa sẽ có một Kalmal filter tương ứng với nó.

**2.1. Kalmal filter**

Sử dụng Kalmal filter với các tham số như sau:

- Hidden state độ dài 7 bao gồm:

+ x: tọa độ tâm box theo trục ngang

+ y: tọa độ tâm box theo trục thẳng đứng

+ s: diện tích bouding box

+ r: tỉ lệ chiều dài/ chiều rộng của box

+ vx: vận tốc của box theo trục ngang trên một đơn vị thời gian (dt)

+ vy: vận tốc của box theo trục dọc trên một đơn vị thời gian (dt)

+ vs: tốc độ thay đổi diện tích box trên một đơn vị thời gian (dt) (vì hướng di chuyển của dừa từ xa tới gần đối với camera)

- Observation (measurement): 4 trường đầu tiên của hidden state: x,y,s,r

- Các giá trị error được khởi tạo phù hợp

Kalmal filter trong trường hợp này sử dụng các trạng thái x,y,s,r; do đó cần chú ý các thao tác cần chuyển đổi qua lại giữa bouding box ở dạng (upper-left, bottom-right) với (x,y,s,r) nói trên.

**2.2. Tracker**

Ý tưởng thuật toán trong source code cung cấp (Tracker.py) được mô tả như sau:

- Khởi tạo danh sách objects khi danh sách theo dõi là trống (ví dụ, ở frame đầu tiên).

- Tại một thời điểm, mỗi qủa dừa được theo dõi bởi một đối tượng Track, sử dụng một Kalmal filter

- Tại mỗi thời điểm t+1, so sánh tương quan giữa các detections của Kalmal sau thời điểm t, với các boudding box detect được ở thời điểm hiện tại, sử dụng thuật toán Hungarian để ghép cặp các cặp (prediction, boudding box) với IoU cao nhất.

- Kiểm tra nếu có detections chưa được ghép cặp, đó là object mới cần tracking, +1 vào tổng số dừa cần đếm.

- Nếu object biến mất sau một số frame, xóa bỏ object đó.

- Update trạng thái của toàn bộ các Kalmal filter ứng với mỗi object

**3. Cải thiện thuật toán phù hợp hơn**

Chọn threshold cho score phù hợp, cụ thể chọn threshold= 0.5

Đối với thuật toán nêu trên, quá trình đếm bị dư rất nhiều (nhiều hơn số dừa thực tế). Để đạt được độ chính xác cao, thuật toán cần được cải thiện.

**3.1. Theo dõi các objects trong một vùng chính giữa góc quay**

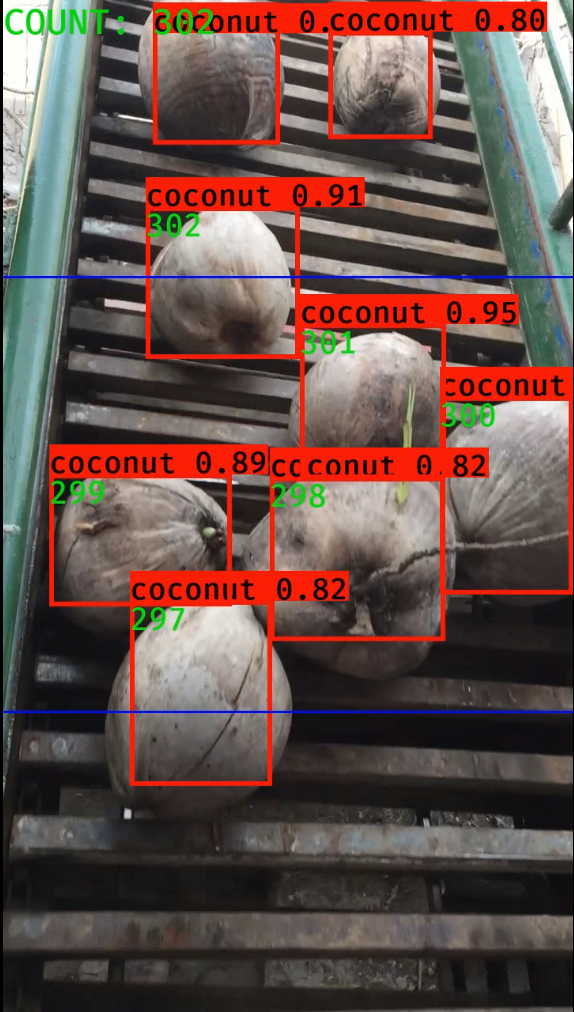
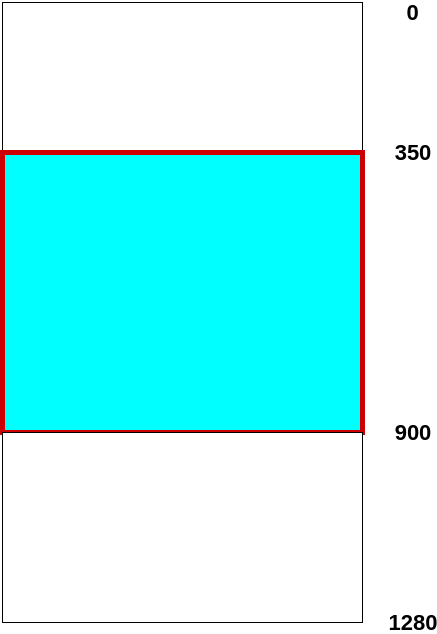
Nhận xét rằng nếu theo dõi tất cả các objects trong toàn bộ khung hình, dễ rất dễ gây nhầm lẫn trong các trường hợp:

- Object bắt đầu xuất hiện chưa rõ ràng, kích thước object quá bé so với kích thước thực.

- Một số object lăn qua khung hình nhưng lại lăn lại, gây đếm sai.

Giới hạn lại vùng detection, còn giúp giảm thiểu sai lệch do bước detection sử dụng YOLO chưa chính xác.

Giới hạn các detections xem xét, có tọa độ dọc trong khoảng (350,900), như sau:



**3.2. Loại bỏ bớt các sai lệch do bước detections**

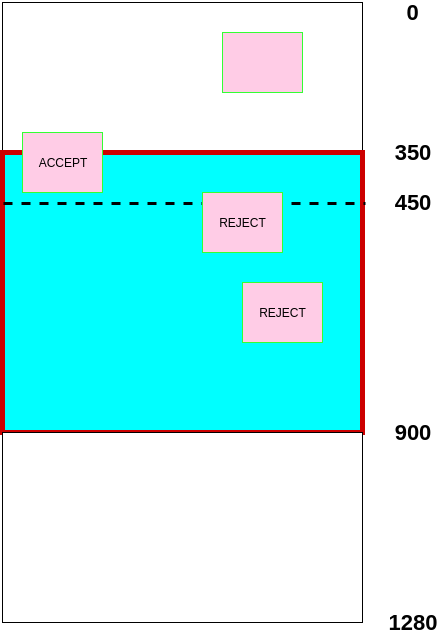
**Nguyên nhân:**

- Một số object đã được detect, nhưng sau đó bị detect thiếu, sau một khoảng thời gian lại xuất hiện lại

- Một số object bị detect sai, ví dụ: chỉ có 1 quả dừa nhưng lại có tới 2 boudding box. Điều này có thể giảm thiểu nhờ điều chỉnh threshold của score tăng lên, nhưng vô tình làm ảnh hưởng tới các kết quả detect khác.

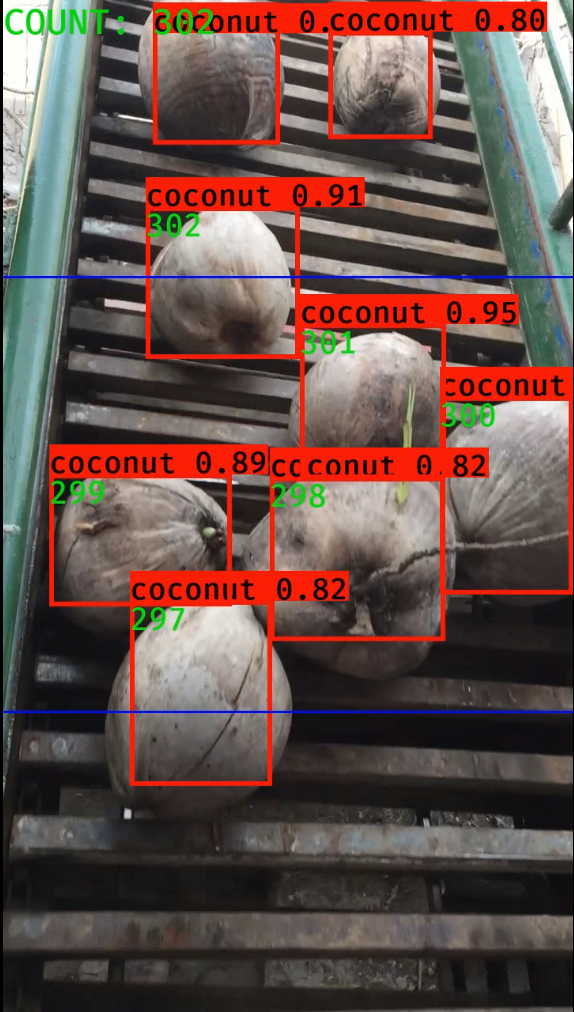
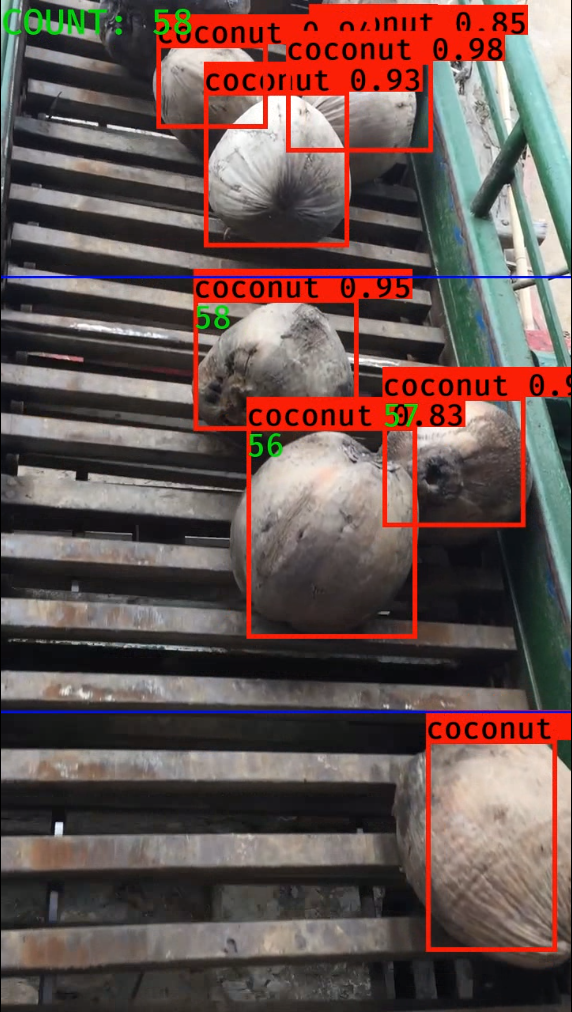
**Giải pháp:**

Khi một detection chưa được ghép cặp với một track, thay vì ngay lập tức coi đó là object mới (+1 vào số dừa), ta thực hiện kiểm tra ràng buộc: tâm box của detection không được vượt quá 350+100=450. Ngắn gọn hơn, lần đầu tiên xuất hiện của một object, thì tọa độ của nó không được quá cao( bất thường).



**4. Kết quả:**

Thử nghiệm với video cắt từ video ban đầu (do dung lượng lớn), video dùng để đánh giá kết quả là từ 1:30 tới 4:00 của video ban đầu.



**Số dừa thực tế: 400**

**Số dừa đếm được: 406**

Chi tiết kết quả mỗi lượt đếm 100 quả dừa ( đếm được/ thực tế):

- Lượt 1: 100/100

- Lượt 2: 100/100

- Lượt 3: 102/100

- Lượt 4: 104/100

**Video kết quả đặt tại link google drive:**

<https://drive.google.com/file/d/17mIW-CQf1rwqIrRTqdAiN0M2Vcm1U9b5/view?usp=sharing>