



ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

BÁO CÁO ĐỒ ÁN MÔN HỌC: IE221.L21.MTCL - NHÓM 25.

XÁC ĐỊNH VẬT THỂ VỚI MÔ HÌNH YOLOV3

GIẢNG VIÊN HƯỚNG DẪN: THS. NGUYỄN THANH SƠN

Thành viên nhóm:

18521278 – Trần Phạm Tiến Phương

I. Mô tả đề án:

- Tạo ứng dụng xác định vật thể qua ảnh, sử dụng pre-trained model.
- Các thư viện sử dụng:
 - cv2
 - torch
 - numpy
 - matplotlib
 - utils.py (<https://github.com/Garima13a/YOLO-Object-Detection>)
 - mạng neural darknet (<https://github.com/Garima13a/YOLO-Object-Detection>)
 - tkinter
- Dataset sử dụng : COCO

II. Tìm hiểu mô hình YOLOv3:

You only look once (YOLO) là một mô hình CNN để detect object mà một ưu điểm nổi trội là nhanh hơn nhiều so với những mô hình cũ. Thậm chí có thể chạy tốt trên những IOT device như raspberry pi.

Vì ưu điểm tốc độ cao, độ chính xác tương đối, nên YOLOv3 rất thích hợp cho những ứng dụng xác định vật thể cần real-time.

Trong đề án của chúng em, sử dụng lại CNN đã được train để xác định class của một số object cơ bản. Ta thêm vào 2 anchor box, một chiều dọc và một chiều ngang để có thể xử lý những vật thể bị đè lên nhau.

- ❖ Anchor box: YOLO có thể xử lý tốt với nhiều object, mỗi object được xác định bởi 1 ô vuông (grid), nhưng trong trường hợp một ô vuông chứa đến 2 tâm của vật thể, thì mô hình sẽ phải chọn lựa 1 trong 2 vật thể đó. Với việc sử dụng anchor box, chúng ta có thể cho phép 1 ô vuông xác định được nhiều vật thể.

Mô hình YOLO phân chia bức ảnh thành một ma trận ô vuông $n \times n$ (grid cell) và mạng neural cung cấp các 1 vector đầu ra cho mỗi grid cell. Các vector này chứa các thông tin: có vật thể hay không, vật thể đó thuộc class nào, và bounding box cho vật thể.

Bởi vì chúng ta sử dụng 2 anchor box, chúng ta sẽ có 2 anchor box dự đoán cho mỗi grid cell. mỗi anchor box này sẽ có giá trị PC (xác suất xuất hiện của vật thể)

Một số anchor box sẽ có giá trị PC rất thấp, tức là khả năng xuất hiện vật thể trong grid cell đó thấp, hay nói cách khác là vật thể không xuất hiện trong đó.

Với mỗi class mà các vector output cung cấp:

- Sử dụng thuật toán NMS* để loại bỏ các bounding box mà giá trị PC được xác định bởi anchor box, thấp hơn ngưỡng mà chúng ta đặt ra, ở đây nhóm em sử dụng là 0.6 cho ngưỡng NMS.
- ❖ *Non-Maximal Suppression (NMS): Yolo sử dụng NMS để giữ lại các bounding box tốt nhất, loại bỏ các bounding box thấp hơn ngưỡng NMS.*
- Với bước tiếp theo, IOU* , chọn các bounding box có chỉ số PC cao nhất, và loại bỏ toàn bộ các bounding box khác mà quá giống với bounding box vừa chọn.
- ❖ *Intersection Over Union Threshold (IOU): chọn bounding box với PC cao nhất, và xóa toàn bộ các bounding box có IOU(được tính toán trong mô hình pre-trained) cao hơn ngưỡng mà ta đặt ra ở đây là 0.4.*

Tiếp tục lặp lại 2 bước trên cho đến khi toàn bộ các bounding box không tối ưu bị loại bỏ hết, chỉ còn lại 1 bounding box tối ưu nhất (cho từng class).

Và tiếp tục làm như vậy với các class khác.

III. Quá trình làm đồ án:

[22/4/2021 – 28/4/2021] Lập nhóm, xác định và chọn đề án.

[29/4/2021 – 12/5/2021] Tìm hiểu rõ hơn về YOLOv3, tìm nguồn, test một số model pre-trained có sẵn, sau khi tìm được model pre-trained, thành công demo xác định vật thể.

[15/5/2021] update link github lần 1, chưa có GUI, tiếp tục code GUI.

[16/5/2021] code xong GUI.

[17/5/2021] update link github lần 2, đã có GUI.

[22/5/2021] hoàn thành cải thiện GUI theo hướng dẫn của giảng viên.

[26/5/2021] Hoàn thành viết báo cáo.

IV. Các khó khăn trong quá trình làm đề án:

- Bởi vì em thuộc ngành KTMT, mới tiếp xúc với AI, và vì nhóm chỉ có 1 người, nên vẫn chưa hiểu rõ ràng về các model, cách xây dựng thuật toán AI và train AI, nên phải sử dụng pre-trained để tiết kiệm thời gian.
- Với thuật toán Yolov3, chỉ mới hiểu sơ được cách hoạt động, nhưng chưa thể tự code và xây dựng mô hình được.
- Trong tương lai sẽ tiếp tục tìm hiểu kiến thức về AI, và tiếp tục phát triển đề án lên khóa luận.