Ngày nay cùng với sự phát triển của công nghệ hiện đại thì song song hành vi trộm cách danh tính trong kỹ thuật số ngày càng nghiêm trọng. Một trong đó là kẻ xấu có thể sử dụng thông tin từ biển số để thực hiện các hành vi gian lận, truy lùng hoặc xâm phạm riêng tư. Hay có thể là những fan theo dõi hay fan cuồng người nổi tiếng nào đó có thể lấy từ các clip hay ảnh trên mxh mà biết được biển số xe mà có những hành vi quấy rối hay tiêu cực. Chính vì thế cho dù trong ảnh hay video, biển số xe thường bị làm mờ trên TV, trực tuyến và trong phim. Từ những điều đó cho mình động lực làm bài toán về phát hiện và làm mờ biển số xe Việt Nam.

→ input, output

Với bài toán này mình sẽ chọn model theo những ưu tiên sau

Tốc độ xử lý từng video phải nhanh hơn so với phương pháp làm mờ thủ công bằng Photoshop.

Mô hình có khả năng phát hiện biển số tốt tránh bỏ sót hay nhầm lẫn giữa biển số và nền.

Mình tập trung làm mờ biển số trên video ngắn. Nên mình muốn sử dụng 1 mô hình tiết kiệm thời gian và độ chính xác tốt.

Từ những lý do đó mình quyết định chọn yolov5 một trong những objects detector tốt hiện nay.

→ yolov5

Và trong bài toán này mình chọn yolov5l vì mô hình có kết quả, tốc độ tốt nhất chỉ sau YOLOv5x. Hơn nữa YOLOv5l có số lượng tham số chỉ bằng một nửa yolov5x nên tốc độ train cũng khá nhanh.

Tiếp theo đây mình nói về phương pháp mình làm trong bài toán này thì mình sẽ train trên ảnh bằng yolov5 như đã nói trên thì sau khi train mô hình sẽ học được các nhận dạng biển số xe tiếp theo đó khi mình detect 1 vid hay ảnh nào đó thì mình thêm 1 tham số --censor thì ở đây mình sẽ lấy tọa độ bb trả về từ mô hình rồi mình sẽ vẽ 1 hình chữ nhật tương ứng ở đây mình dùng màu xám để che đi biển số xe được detect.

Thì mình sẽ dùng phương pháp này trên bộ dataset được thu thập chủ yếu từ các video trên các kênh YouTube và thư viện mì AI, đồng thời được chụp quay ở các quận của thành phố Hồ Chí Minh, bãi giữ xe trường UIT và Học viện Hành chính Quốc gia TPHCM.

Trong tập dữ liệu này gồm 5 loại biển số là biển số trắng, biển số vàng, biển số xanh, biển số đỏ và được phân bố theo hình sau. Do là trong đời sống hằng ngày chúng ta thường gặp nhất xe biển số trắng và vàng ngược lại với 3 biển số còn lại khá ít gặp nên bộ dữ liệu khá nghiêng về biển số xe trắng và vàng.

Tiếp theo là phần đánh giá kết quả. Việc đánh giá sẽ được thực hiện đối với hai giá trị mà mô hình trả về: - Tọa độ bounding box: một vị trí hình chữ nhật trên ảnh mà mô hình cho rằng có chứa biển số. - Confidence score: phản ánh khả năng bounding box chứa biển số và độ tự tin của mô hình về khả năng đó. Nếu bounding box không chứa biển số, confidence score lý tưởng sẽ bằng 0. Nói chung, confidence score có xu hướng cao hơn đối với các bounding box bao sát vật hơn. Mình sử dụng 4 thang đó IoU, P, R và AP. Đầu tiên là Iou. Chỉ số IoU chỉ ra độ khớp giữa bounding box được mô hình dự đoán và ground truth box do con người gán sẵn. Với mỗi ground truth box ta sẽ tiến hành tính IoU với tất cả các bounding box mà mô hình dự đoán. Khi có được IoU, ta sẽ chọn một threshold để phân loại bounding box thành 3 loại: TP, FP, FN.

###

- TP (true positive): Bounding box nào có IoU với 1 ground truth box cao nhất trong tất cả các giá trị IoU của ground truth box đó, cao hơn threshold thì sẽ được tính là TP.

- FP (false positive): Bounding box không phải TP sẽ là FP.

- FN (false negative): ground truth box nào không có 1 bounding box được tính là TP giao với nó sẽ được tính là FN.

- TN (true negative): không có ý nghĩa đối với bài toán object detection. TN mang ý nghĩa những vị trí mà mô hình cho là nền và bên ground truth cũng là nền.

###

Số lượng của 3 loại bounding box trên được dùng để tính toán các thang đo còn lại.

###

Các threshold khác nhau thì số lượng 3 loại bounding box vì đó mà thay đổi, dẫn đến thang đo mà nhóm chọn phía dưới cũng thay đổi.

###

Giả sử ta đánh giá trên ảnh sau. Ảnh bên trái là ảnh do mình dán nhãn còn ảnh bên phải là ảnh do model predict với iou = 0.25

LÀM LẠI BẢNG IOU, MATRIX

Thì đây là bảng kết quả mình tình IoU khá cao nên và model đã detect được 4 biển số trong tấm ảnh tương ứng với ảnh grouth truth.

Tiếp theo là precison → precision. Precision có thể không giảm cùng với Recall. Định nghĩa về Precision cho thấy việc giảm threshold có thể làm tăng mẫu số, bằng cách tăng số lượng kết quả trả về. Nếu threshold đang chọn được đặt quá cao, giảm threshold có thể khiến tất cả các kết quả mới là true positive, điều này sẽ làm tăng độ chính xác. Nếu threshold đang chọn phù hợp hoặc quá thấp, việc tiếp tục hạ thấp threshold sẽ khiến false positive tăng, làm giảm độ chính xác.

Về recall → recall. Trong đó, 𝑇𝑃 + 𝐹𝑁 không phụ thuộc vào threshold. Điều này có nghĩa việc giảm threshold có thể tăng Recall bằng cách tăng số lượng kết quả true positive. Ngoài ra, việc hạ thấp threshold có thể không thay đổi Recall, trong khi Precision lại dao động.

Cuối cùng là AP

Mình sẽ minh họa cách vẽ PR curve trên 2 ảnh sau. Hình chữ nhật màu đỏ bao quanh vùng hình ảnh có thể chứa biển số do mô hình dự đoán còn gọi là bounding box - Mô hình dự đoán tổng cộng X bounding box với các mức confidence score khác nhau.

TÍNH LẠI PR

Bước 1: Sắp xếp các bounding box dự đoán theo confidence score giảm dần. Cột detections là thứ tự các bounding box được dự đoán.

Bước 2: Tính TP, FP Cột TP, FP sẽ được tính theo phần định nghĩa IoU ở phía trên nhóm đã trình bày. Nếu là bounding box nào là TP thì sẽ có chỉ số TP = 1 và FP = 0, ngược lại thì TP = 0, FP = 1.

Bước 3: Tính acc TP Acc TP là chỉ số TP tích lũy, tức là sẽ đếm tổng số TP của các bounding box trước đó (các hàng phía trên) và bounding box hiện tại theo chiều giảm confidence score. Nếu TP là 1 sẽ cộng thêm 1 vào chỉ số Acc TP của các hàng trước đó, ngược là giữ nguyên.

Bước 4: Tính Acc FP Ngược lại với cách tính Acc TP.

Bước 5: Tính Precision

Bước 6: Tính Recall

Bước 7: Vẽ biểu đồ với 2 trục tọa độ là Recall và Precision và tính AP.

Cuối cùng đây là conf matrix mình trên tập test với threshold 0.5(trái) và 0.75 (phải)