

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN



BÁO CÁO MÔN HỌC
PHÁT HIỆN VÀ LÀM MỜ BIỂN SỐ XE

Giảng viên hướng dẫn: **Thầy Phạm Nguyễn Trường An**

Thầy Lê Đình Duy

Lớp: **CS114.M11**

Sinh viên thực hiện: **MSSV:**

Phạm Thành Luân **19520154**

Lê Quang Nha **19520195**

Đỗ Minh Trí **19520307**

Tp Hồ Chí Minh, tháng 1 năm 2021

Mục lục

I)	Giải trình chỉnh sửa vấn đáp:.....	1
II)	Giới thiệu bài toán	1
1.	Hoàn cảnh ứng dụng.....	1
2.	Mô tả bài toán.....	2
3.	Mô tả bộ dữ liệu	2
III)	Training và đánh giá mô hình:.....	3
1)	Training:	3
2)	Đánh giá mô hình	4
3)	Kết quả thực nghiệm:	8
IV)	Kết luận:	9
1)	Ưu điểm:.....	9
2)	Khuyết điểm:	9
3)	Hướng phát triển.....	10
V)	Tài liệu tham khảo	11

I) Giải trình chỉnh sửa vấn đề:

Sau khi vấn đề thì nhóm chúng em đã bổ sung về ý nghĩa của các số bounding box trong yolov4 được bổ sung ở phần Giới thiệu bài toán ở mục mô tả bộ dữ liệu, và lý do sử dụng yolov4 của nhóm ở phần đầu của mục training và đánh giá mô hình.

[1] Link train: https://drive.google.com/drive/u/1/folders/1g6_sMzaRsyTOu-aJAhpOAN0vzawcYp-Q (bao gồm dataset, lastweight, tập train, tập validation)

[2] Link colab: [Train with yolov4.ipynb - Colaboratory \(google.com\)](https://colab.research.google.com/github/hoanghuy1999/yolov4/blob/master/train.ipynb)

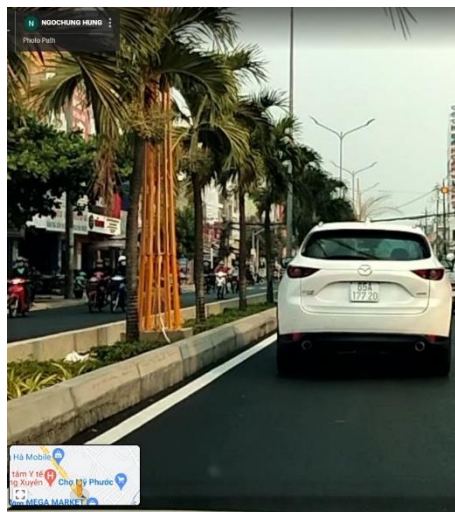
[3] Link github của nhóm: [nhalq/CS114.M11: CS114.M11 - Machine Learning \(github.com\)](https://github.com/nhalq/CS114.M11)

II) Giới thiệu bài toán

1. Hoàn cảnh ứng dụng

Cách mạng 4.0 phát triển dẫn đến việc phát triển của nhiều ngành nghề 1 trong số đó là các nghề như youtuber, vlogger,... Các nghề youtuber, vlogger là những nghề hay quay những video, clip sau đó đăng lên mạng xã hội. Tuy nhiên khi quay cảnh đường phố thì các biển số xe thường không được làm mờ. Điều này có thể để lộ ra thông tin người trong ảnh đến nơi đó vào thời điểm nào, có 1 số người sẽ khó chịu về điều này.

Ngoài ra ứng dụng Google map cũng che biển số của người dùng. Ứng dụng này được sử dụng rất tốt với các khu vực của nước ngoài, nhưng ở Việt Nam thì nó chưa hoạt động được tốt lắm.



Chính vì vậy nhóm tụi em đã vận dụng kiến thức của máy học, để xây dựng 1 ứng dụng giúp tự động phát hiện biển số xe trong bức ảnh và tự động làm mờ chúng

2. Mô tả bài toán

Input: 1 bức ảnh có phương tiện giao thông có độ phân giải tối thiểu 480p.

Output: Hình ảnh đầu vào với biển số xe được nhận diện và làm mờ (nếu có biển số xe xuất hiện trong ảnh)



3. Mô tả bộ dữ liệu

Bộ dữ liệu cho bài toán tập trung hoàn toàn vào các biển số xe Việt Nam được nhóm tự thu thập, cắt khung hình từ các videos nhóm quay được trong quá trình tham gia giao thông và một số video về chủ đề giao thông trên youtube.

Tổng số lượng ảnh trong dữ liệu gồm 2470 ảnh. Sau đó nhóm phân chia thành 2 tập dữ liệu train và validation theo tỷ lệ là 8:2 tức có tổng cộng 2014 ảnh trong tập train dùng để huấn luyện mô hình và 456 ảnh trong tập validation dùng để đánh giá mô hình trong quá trình huấn luyện.

Trong đó có 2 class là biển số xe trắng và biển số xe vàng với số lượng mỗi class là:

Trong tập validation có 1075 biển số trắng và 44 biển số vàng

Trong tập train có 5828 biển số trắng và 224 biển số vàng

Do luật về xe mang biển số vàng mới có vài năm gần đây do đó việc thu thập dữ liệu mới có sự chênh lệch lớn như vậy.

- Cách gán nhãn là cố gắng gán nhãn đủ tất cả những biển số xe thấy rõ và có khả năng nhìn thấy xuất hiện trong ảnh. Các bounding box được bao quanh các biển số xe tức chứa các biển số xe cần nhận diện
- Mỗi bounding box sẽ có 5 giá trị
Số đầu tiên chỉ nhãn của đối bounding box có thể là biển số vàng 1, biển số trắng 0
2 số tiếp theo là tâm của bounding box
2 số cuối cùng là chiều dài và chiều rộng của ảnh đã chuẩn hoá.

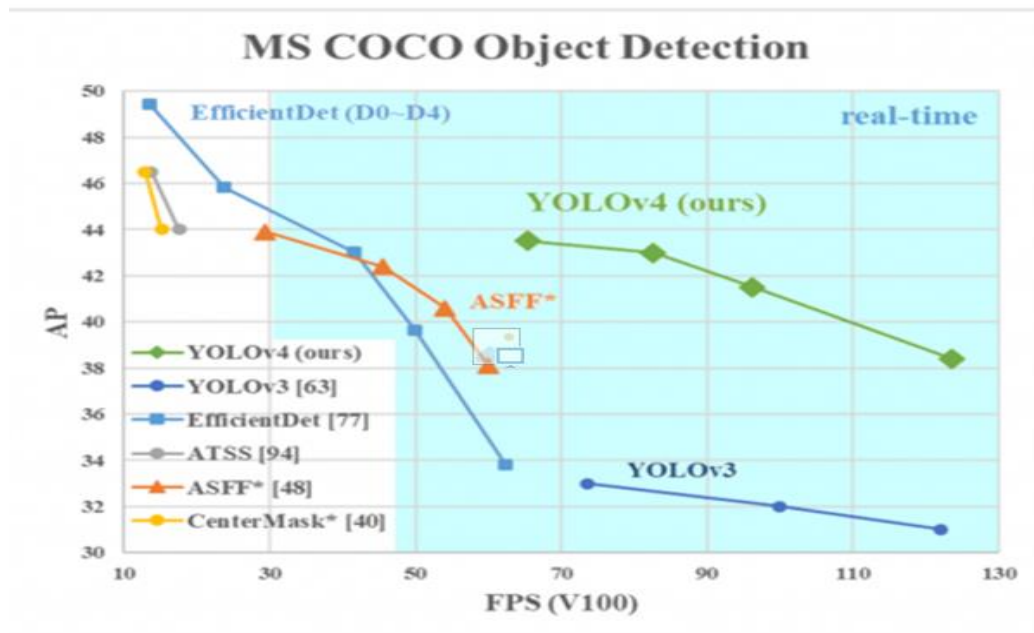
```
class_id center_x center_y width height
```

```
1 0.661000 0.601124 0.106000 0.117978
0 0.135000 0.530899 0.046000 0.044944
```

Lưu ý các số được chuẩn hoá thành các tọa độ thuộc miền giá trị $[0;1]$ ta chia các giá trị thuộc chiều dài ($center_x$, $width$) cho chiều dài thực của bức ảnh, các giá trị thuộc chiều cao ($center_y$, $height$) cho chiều cao thực của bức ảnh

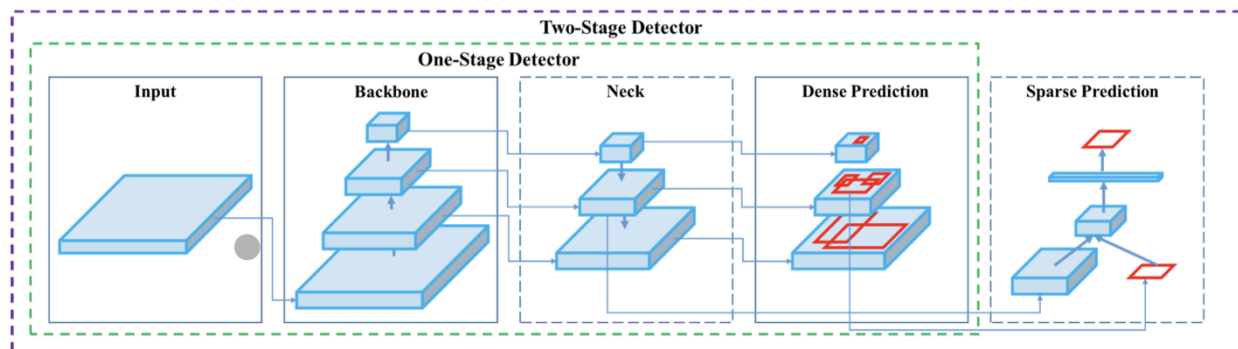
III) Training và đánh giá mô hình:

1) Training:



Dựa vào biểu đồ MS CoCo, nhóm em sử dụng model yolov4, bởi vì yolov4 là sự cải tiến của yolov3 nhờ vào những cải tiến các layer giúp cho độ chính xác mAP và số khung hình trên giây được cải thiện đáng kể (mAP tăng 10%, FPS nhanh hơn 12%)

Các Layer của mô hình yolov4:



2) Đánh giá mô hình

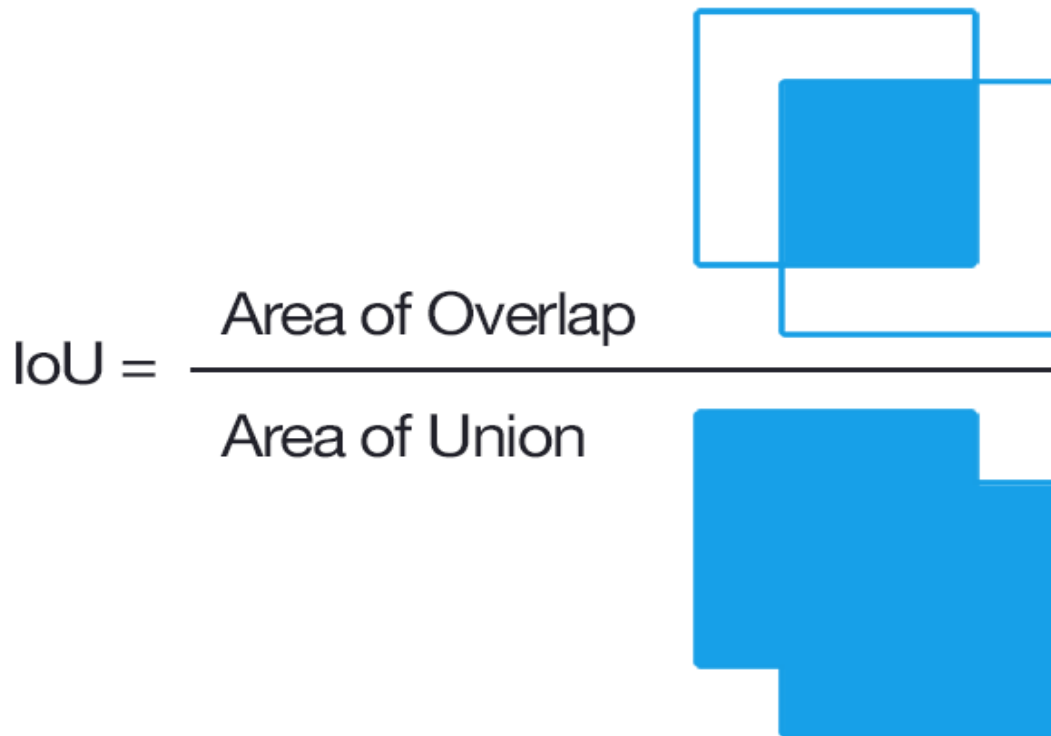
2.1) mAP:

Để đánh giá chất lượng của hai mô hình huấn luyện trên, nhóm sử dụng thông số mean Average Precision (mAP) để so sánh kết quả

Trước tiên, chúng ta hãy cùng có cái nhìn tổng quan về ý nghĩa của chỉ số mAP và mAP50 đối với bài toán nhận diện biển số xe qua các định nghĩa có liên quan sau:

IoU (Intersection over Union) là giá trị thể hiện sự giao nhau giữa vùng xuất hiện thực tế của vật thể cần nhận diện (growth truth box) và vùng mô hình dự đoán là có sự xuất hiện của vật thể (predicted box). Khi tính toán, mô hình sẽ được gán một ngưỡng IoU (IoU threshold) từ trước – trong trường hợp của chúng em, IoU threshold được gán là 0.5. Ngưỡng này có tác dụng xác định các trường hợp nhận diện vật thể của mô hình là đúng hay sai. Nếu IoU của một trường hợp lớn hơn bằng IoU threshold cho trước thì đó được xem là một trường hợp mô hình dự đoán nhận diện vật thể đúng, ngược

lại là sai.



Minh họa cách tính toán chỉ số IoU, ô vuông xanh thể hiện growth truth box và ô vuông đỏ thể hiện predicted box. IoU là tỉ lệ giữa diện tích giữa phần giao nhau của growth truth box và predicted box và diện tích phần hợp của growth truth box và predicted box. IoU càng tiến về 1 chứng tỏ mô hình dự đoán càng tốt.

Các chỉ số tiếp theo ta cần quan tâm chính là **True Positive, False Positive, True Negative, False Negative**:

True Positive (TP): Một trường hợp mà mô hình dự đoán được xem là True Positive khi mô hình nhận diện vật thể đó đúng. Tức IoU của trường hợp đó lớn hơn hoặc bằng IoU Threshold.

True Negative (TN): Một trường hợp mà mô hình dự đoán được xem là True Negative khi không có sự xuất hiện của vật thể trong hình ảnh đầu vào và mô hình đã hoàn thành tốt việc dự đoán đó khi không phát hiện được bất cứ vật thể nào cần nhận diện trong hình ảnh.

False Postive (FP): False Positive được tính khi mô hình nhận diện sai vật thể xuất hiện trong hình ảnh. Tức IoU của trường hợp đó bé hơn IoU Threshold.

False Negative (FN): Một trường hợp được tính là False Negative khi có sự xuất hiện của vật thể cần nhận diện trong hình ảnh nhưng mô hình không nhận diện được vật thể đó.

Từ các giá trị trên, mô hình sẽ có cơ sở để tính toán hai giá trị có liên quan và ý nghĩa trực tiếp đến mAP đó chính là **Precision** và **Recall**. Về cơ bản, trong các bài toán Object Detection, ta có thể hiểu Precision là tỉ lệ phần trăm mức độ chính xác các dự đoán của mô hình, trong khi đó, Recall thể hiện khả năng nhận diện chính xác các vật thể của mô hình. Mối liên hệ giữa Precision và Recall đối với TP, FN, FP được thể hiện qua hai công thức sau:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

Sau khi tính toán được Precision và Recall ta sẽ tính toán được **Average Precision (AP)** là tổng các tích giữa Precision và Recall tại các ngưỡng theo công thức sau đây.

$$AP = \sum_n (R_n + R_{n-1}) P_n \quad (3)$$

Trong đó:

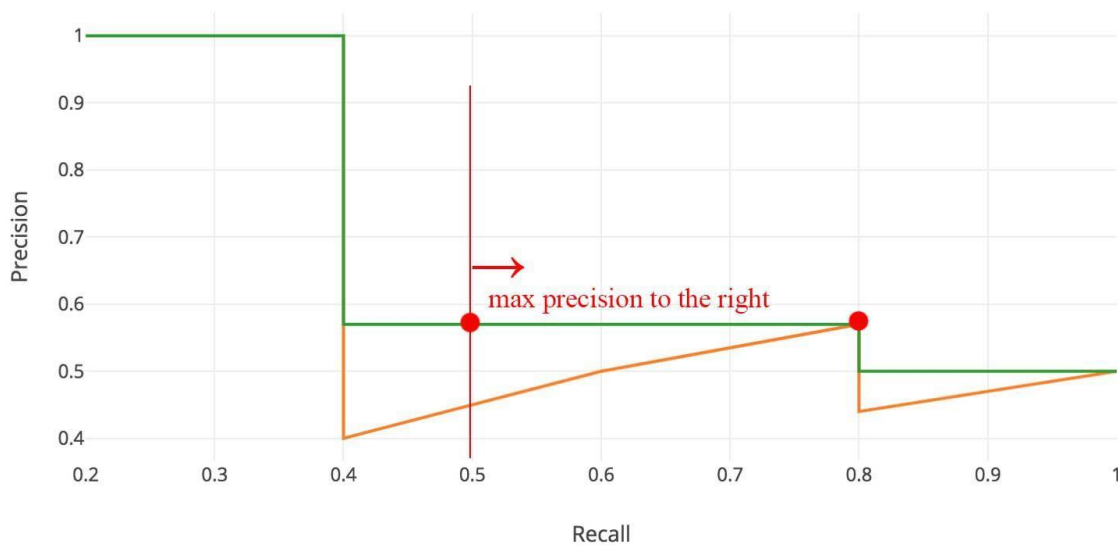
N là ngưỡng đã được đặt từ trước.

$n = 1, 2, 3, \dots, N$ là các ngưỡng, mà tại các ngưỡng này ta sẽ tính giá trị Precision và Recall.

R_n là giá trị Recall tại ngưỡng đang xét n

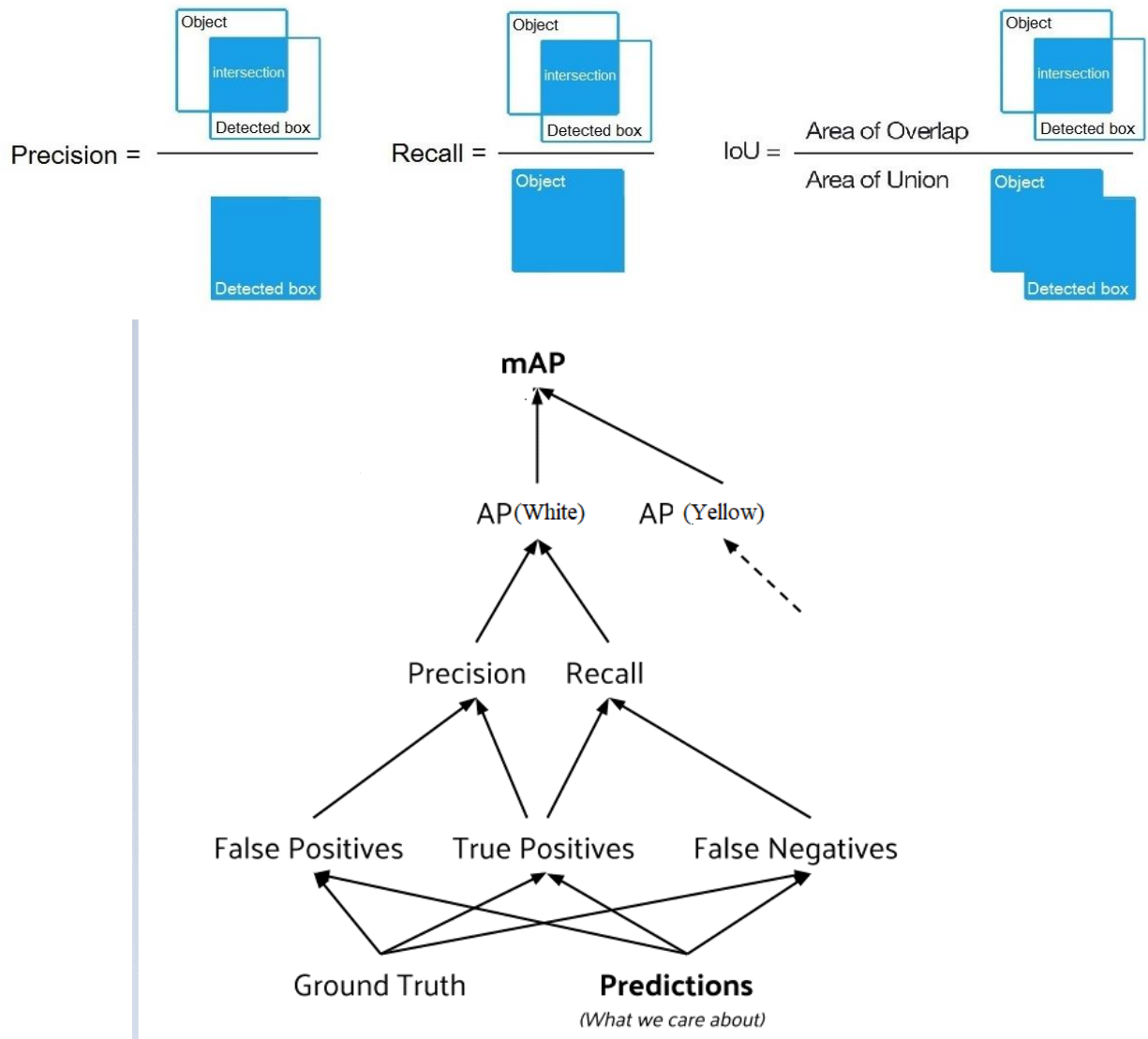
R_{n-1} là giá trị Recall tại ngưỡng ngay trước ngưỡng đang xét.

P_n là giá trị Precision cao nhất được xét tại ngưỡng n



Minh hoạ thuật toán tính toán Average Precision qua biểu đồ. Các giá trị Recall và Precision được đánh dấu bằng toạ độ của các giá trị tại các ngưỡng, tại mỗi ngưỡng ta sẽ tính diện tích hình chữ nhật có chiều dài và chiều cao lần lượt là $(R_n - R_{n-1})$ và P_n . AP sẽ là tổng diện tích tất cả các hình chữ nhật “dưới đường kẻ màu xanh”.

mAP là một thước đo cơ bản và có ý nghĩa quan trọng đối với các bài toán nhận diện vật thể. Nó chính là giá trị trung bình của AP tại tất cả các vòng lặp trong quá trình huấn luyện mô hình



Hình mô tả cách tính mAP

3) Kết quả thực nghiệm:

Sau khi huấn luyện tập dữ liệu ta có kết quả trên tập test với [mAP@0.50](#) là 0.920128 với ngưỡng IoU là 0.5



IV) Kết luận:

1) Ưu điểm:

- Kết quả huấn luyện cho độ chính xác khá cao

2) Khuyết điểm:

- Dữ liệu bị mất cân bằng (số lượng class biển số trắng nhiều hơn số lượng biển số vàng rất nhiều)

- Bộ dữ liệu chưa đa dạng các biển báo bị ố màu hay bị che mất 1 phần chưa có nhiều trong tập data
- Khả năng nhận diện các biển số xe nhỏ vẫn còn hạn chế và thiếu sót, các biển xe của cơ quan, tổ chức, cá nhân trong nước (biển màu xanh),... mô hình huấn luyện vẫn chưa nhận diện được
- Các chỉ số mAP và mAP50 thể hiện độ chính xác vẫn chưa phản ánh đúng được hiệu năng của mô hình huấn luyện

3) Hướng phát triển

Để khắc phục các hạn chế trên thì chúng em hướng tới việc bổ sung dữ liệu và bổ sung các class như biển số xanh, biển số đỏ, và các biển số trong điều kiện khó khăn như biển số bị ố màu, bị che 1 phần,...

Ngoài ra nhóm chúng em còn muốn tạo ra chương trình xử lý che mờ biển số trong điều kiện thực tế để có thể hỗ trợ cho các nhu cầu cần thiết của cuộc sống.

V) Tài liệu tham khảo

[1] yolov4: <https://paperswithcode.com/method/yolov4>

[2]darknet: [AlexeyAB/darknet: YOLOv4 / Scaled-YOLOv4 / YOLO - Neural Networks for Object Detection \(Windows and Linux version of Darknet \) \(github.com\)](https://github.com/AlexeyAB/darknet)

[3] Đánh giá mAP: <https://devai.info/2020/12/17/tim-hieu-mapmean-average-precision-danh-gia-mo-hinh-object-detection-su-dung-yolov4/?fbclid=IwAR0c-KdH9w3YJykmp6EFNSQDySY32sbHzubEn17O2wi4f-W4vYMA8fioghM>

[4] Tính Ap: <https://blog.paperspace.com/mean-average-precision/>