

BỘ GIÁO DỤC VÀ ĐÀO TẠO

TRƯỜNG ĐẠI HỌC SƯ PHẠM TP.HCM

KHOA CÔNG NGHỆ THÔNG TIN



Giảng viên hướng dẫn: **TS. NGUYỄN THỊ HUỲNH TRÂM**

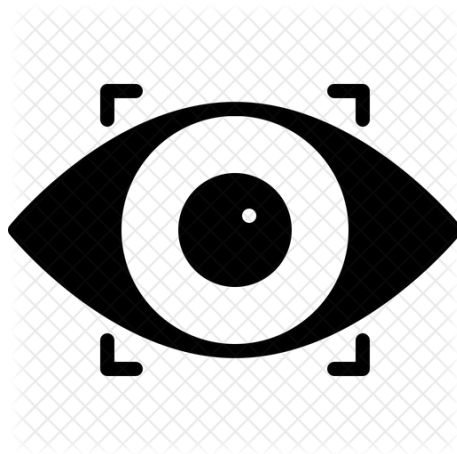
Sinh viên thực hiện: **TRƯƠNG ĐỨC HẢI NGUYỄN**

Lớp: **CNTT.C**

MSSV: **44.01.104.155**

BÁO CÁO ĐỒ ÁN TRÍ TUỆ NHÂN TẠO

TRAFFIC TRACKING



Thành phố Hồ Chí Minh, tháng 8 năm 2020

Lời cảm ơn

Với những lời đầu tiên, tôi xin được gửi đến TS. Nguyễn Thị Huỳnh Trâm – Giảng viên bộ môn Trí Tuệ Nhân Tạo Trường Đại học Sư Phạm Tp.HCM, cảm ơn cô đã truyền đạt những kiến thức đầu tiên để làm tiền đề, cơ sở hoàn thành bài báo cáo này và đã, đang, sẽ trở thành những kinh nghiệm quý báu cho việc học tập và nghiên cứu của bản thân tôi trong tương lai.

Đồng thời, tôi cũng bày tỏ lòng biết ơn với quý Thầy Cô Trường Đại học Sư phạm Tp.HCM đã tận tình truyền dạy kiến thức trong quá trình học tập và thực hiện bài báo cáo này.

Cuối cùng, xin gửi lời cảm ơn tới những người thân trong gia đình và bạn bè luôn bên cạnh động viên, hỗ trợ về mặt tinh thần để tôi vượt qua khó khăn và hoàn thiện bài báo cáo này.

Lời cam đoan

Tôi xin cam đoan kết quả đạt được trong bài báo cáo là sản phẩm của riêng cá nhân, là kết quả của quá trình học tập và nghiên cứu khoa học độc lập. Trong toàn bộ nội dung của bài báo cáo, những điều được trình bày hoặc là của cá nhân hoặc là được tổng hợp từ nhiều nguồn tài liệu. Tất cả các nguồn tài liệu tham khảo đều có xuất xứ rõ ràng và được trích dẫn hợp pháp.

Tôi xin hoàn toàn chịu trách nhiệm và các hình thức kỷ luật theo quy định cho lời cam đoan của mình.

Thành phố Hồ Chí Minh, tháng 08 năm 2020

Tác giả

Trương Đức Hải Nguyên

Mục lục

Lời cảm ơn.....	2
Lời cam đoan	3
I. MỞ ĐẦU	5
1. Lý do chọn đề tài.....	5
2. Mục tiêu đề tài.....	5
3. Nội dung thực hiện	5
4. Phương pháp thực hiện	5
II. Cách hoạt động của thuật toán YOLOv3	6
III. Giới thiệu sơ lược về Keras và Tensorflow	6
IV. Dán nhãn data.....	6
V. Bounding box	8
VI. Huấn luyện	8
VII. Đánh giá	9
VIII. Kết luận	10
IX. Trích dẫn	11

I. MỞ ĐẦU

1. Lý do chọn đề tài

Kỉ nguyên của cuộc cách mạng công nghiệp 4.0 đã và đang mang đến cho nhân loại cơ hội để thay đổi và tiến hoá trong nhiều lĩnh vực thiết yếu trong đời sống từ chiếc điện thoại thông minh cho đến mạng vệ tinh toàn cầu,... Trong đó, trí tuệ nhân tạo được xem như là cốt lõi của cuộc Cách mạng công nghiệp này. Một trong những lĩnh vực đang được quan tâm hiện nay chính là việc nhận diện hình ảnh, cụ thể hơn là nhận diện gương mặt, vóc dáng, đồ vật, con vật,...

Mục tiêu tiếp cận và áp dụng những công nghệ của tương lai cũng như là hoàn thành tốt yêu cầu và mục tiêu của học phần “Trí tuệ nhân tạo” chính yếu tố chính cho đề tài “Traffic Tracking”.

2. Mục tiêu đề tài

Sử dụng thuật toán YOLOv3 cùng ngôn ngữ lập trình Python trên môi trường ảo Google Colab huấn luyện 1 hệ thống AI nhận diện xe ô tô và xe máy.

3. Nội dung thực hiện

Nghiên cứu về các vấn đề trong một mô hình AI bao gồm chiết xuất đặc trưng, data train, data test,...

Nghiên cứu về thuật toán YOLOv3 (You Only Look Once) áp dụng trong việc nhận diện vật thể trong thời gian thực.

Tìm hiểu cách sử dụng các thư viện trong Python như numpy, keras, tensorflow,...
Tìm hiểu cách sử dụng Google Colab.

4. Phương pháp thực hiện

Đầu tiên thử nghiệm với pretrained model của tác giả thuật toán YOLO.

Sau đó huấn luyện bằng tập data đã được dẫn nhãn qua file config_city.json.

Cuối cùng là sử dụng tập dữ liệu nằm ngoài tập dữ liệu đã train để đánh giá model.

II. Cách hoạt động của thuật toán YOLOv3

Sử dụng một mạng neural cho toàn ảnh, mạng neural này sẽ chia ảnh ra thành từng phân vùng và dự đoán các bounding boxes cũng như xác suất cho từng phân vùng đó. Các bounding box này được đánh giá bằng xác suất dự đoán.

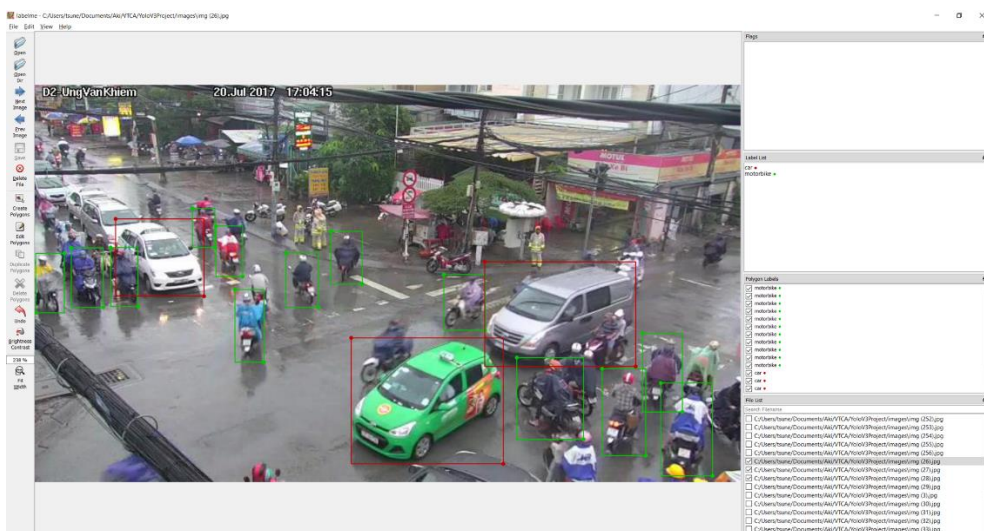
III. Giới thiệu sơ lược về Keras và Tensorflow

Keras là một open source cho Neural Network được viết bằng ngôn ngữ Python. Nhờ vào khả năng xây dựng model deep learning nhanh cùng với việc dễ sử dụng và hỗ trợ xây dựng cả mạng CNN lẫn RNN nên rất được các lập trình viên sử dụng khi xây dựng model.

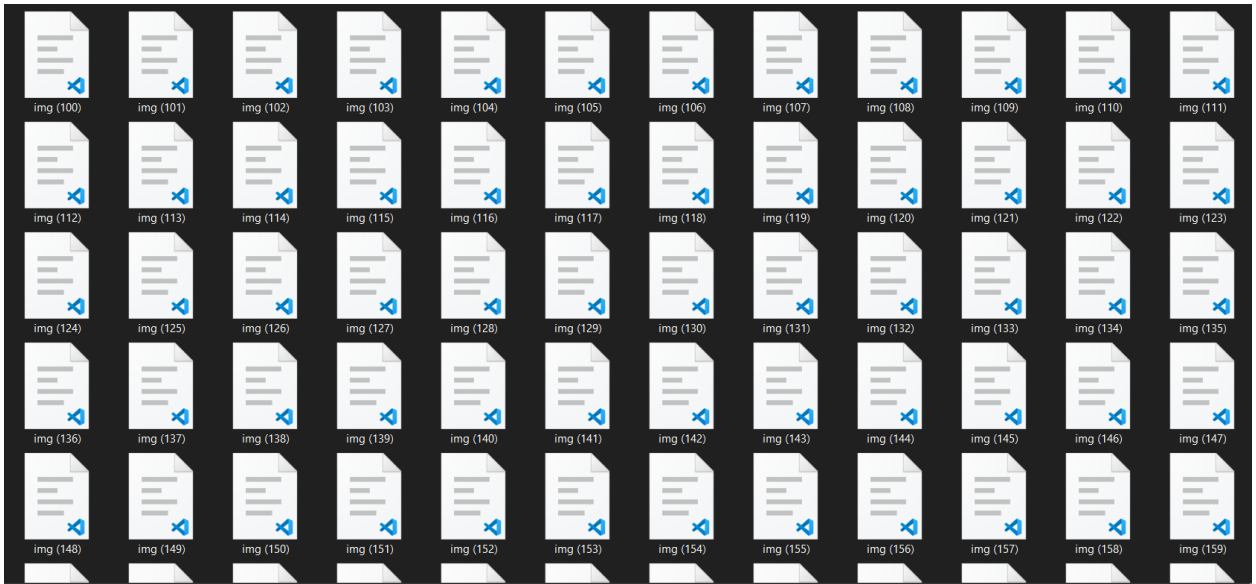
Tensorflow là một open source thường được dùng chung với các thư viện khác như Keras, Theano, PyTorch,... Được dùng để xử lý các tính toán số học.

IV. Dán nhãn data

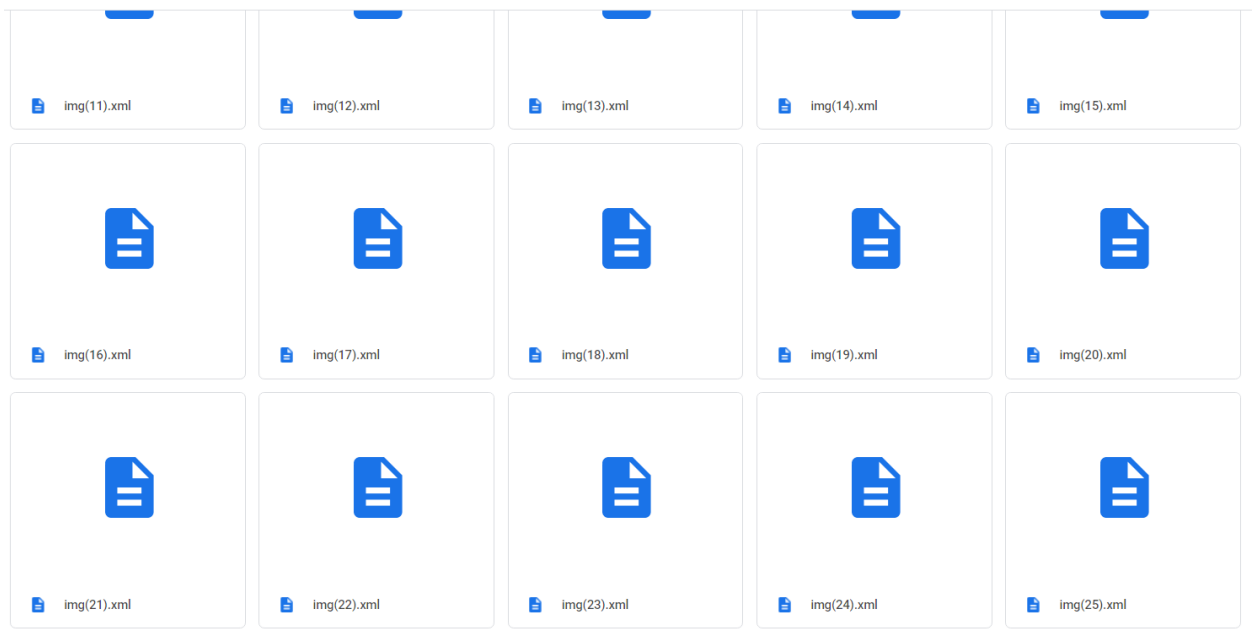
Sử dụng labelme để dán nhãn data, trong trường hợp này ta sẽ có 2 nhãn là [car] [bike]



Sau khi dán nhãn xong, ta sẽ có được các file json.



Vì Keras YOLOv3 sử dụng dạng Pascal VOC nên ta phải convert từ JSON sang XML.



V. Bounding box

Trong YOLOv3, các mô hình được xây dựng cũng sử dụng anchor boxes, là những binding boxes được xác định trước với hình dạng và kích thước hợp lý được tùy chỉnh trong quá trình huấn luyện. Sự lựa chọn các bounding boxes cho hình ảnh được xử lý trước bằng cách sử dụng thuật toán phân cụm K-means trên tập dữ liệu huấn luyện.

VI. Huấn luyện

Trước khi huấn luyện, ta cấu hình lại file config_city.json theo ý ta sau đó huấn luyện qua file train.py

```
/usr/local/lib/python3.6/dist-packages/keras/callbacks/callbacks.py:998: UserWarning: `epsilon` argument is deprecated and will be removed, u
warnings.warn("`epsilon` argument is deprecated and ")
WARNING:tensorflow:From /usr/local/lib/python3.6/dist-packages/keras/backend/tensorflow_backend.py:422: The name tf.global_variables is depre
WARNING:tensorflow:From /usr/local/lib/python3.6/dist-packages/keras/backend/tensorflow_backend.py:431: The name tf.is_variable_initialized is
WARNING:tensorflow:From /usr/local/lib/python3.6/dist-packages/keras/backend/tensorflow_backend.py:438: The name tf.variables_initializer is
WARNING:tensorflow:From /usr/local/lib/python3.6/dist-packages/keras/callbacks/tensorboard_v1.py:200: The name tf.summary.merge_all is deprec
WARNING:tensorflow:From /usr/local/lib/python3.6/dist-packages/keras/callbacks/tensorboard_v1.py:203: The name tf.summary.FileWriter is depre

Epoch 1/100
2020-08-06 02:43:17.184690: W tensorflow/core/framework/cpu_allocator_impl.cc:81] Allocation of 154140672 exceeds 10% of system memory.
2020-08-06 02:43:17.551925: W tensorflow/core/framework/cpu_allocator_impl.cc:81] Allocation of 154140672 exceeds 10% of system memory.
2020-08-06 02:43:18.093992: I tensorflow/stream_executor/platform/default/dso_loader.cc:44] Successfully opened dynamic library libcudnn.so.7
2020-08-06 02:43:22.273872: I tensorflow/stream_executor/platform/default/dso_loader.cc:44] Successfully opened dynamic library libcublas.so.
WARNING:tensorflow:From /content/gdrive/My Drive/AI/Learning/Collab/YoloV3/callbacks.py:19: The name tf.Summary is deprecated. Please use tf.

resizing: 384 384
resizing: 320 320
resizing: 320 320
resizing: 384 384
resizing: 448 448
resizing: 320 320
2020-08-06 02:45:27.995048: W tensorflow/core/common_runtime/bfc_allocator.cc:239] Allocator (GPU_0_bfc) ran out of memory trying to allocate
resizing: 288 288
resizing: 320 320
resizing: 448 448
resizing: 448 448
- 200s - loss: 16.3167 - yolo_layer_1_loss: 0.6037 - yolo_layer_2_loss: 3.5597 - yolo_layer_3_loss: 12.1532
```



```

# compare the seen labels with the given labels in config.json
if len(labels) > 0:
    overlap_labels = set(labels).intersection(set(train_labels.keys()))

    print('Seen labels: \t' + str(train_labels) + '\n')
    print('Given labels: \t' + str(labels))

    # return None, None, None if some given label is not in the dataset
    if len(overlap_labels) < len(labels):
        print('Some labels have no annotations! Please revise the list of labels in the config.json.')
        return None, None, None
    else:
        print('No labels are provided. Train on all seen labels.')
        print(train_labels)
        labels = train_labels.keys()

max_box_per_image = max([len(inst['object']) for inst in (train_ints + valid_ints)])

return train_ints, valid_ints, sorted(labels), max_box_per_image

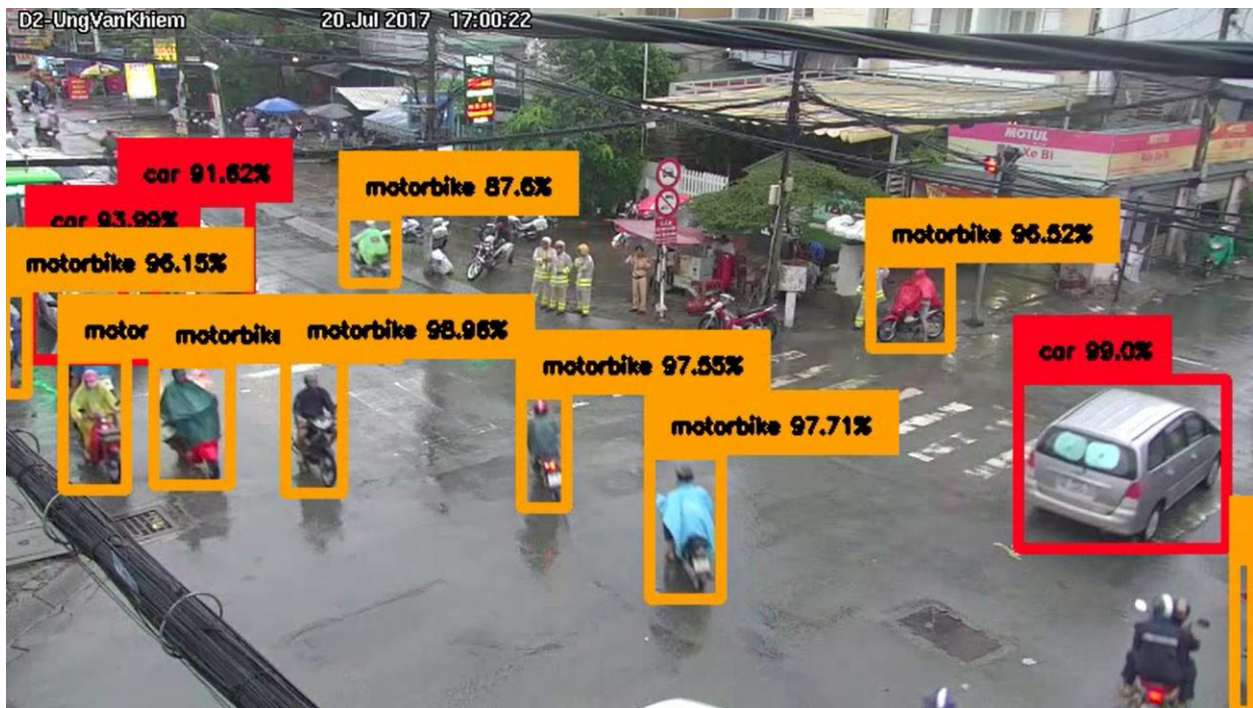
def create_callbacks(saved_weights_name, tensorboard_logs, model_to_save):
    makedirs(tensorboard_logs)

    early_stop = EarlyStopping(
        monitor = 'loss',
        min_delta = 0.01,
        patience = 7,
        mode = 'min',
        verbose = 1
    )
    checkpoint = CustomModelCheckpoint(
        model_to_save = model_to_save,
        filepath = saved_weights_name, # + '{epoch:02d}.h5',
        monitor = 'loss',
        verbose = 1,
        save_best_only = True,
        mode = 'min',
        period = 1
    )

```

VII. Đánh giá

Sau khi huấn luyện xong, ta đánh giá model bằng cách sử dụng dữ liệu bên ngoài, ở trường hợp này là video định dạng mp4.



VIII. Kết luận

Về cơ bản mô hình đã có thể nhận diện được xe ô tô và xe máy với độ chính xác có thể chấp nhận được. Tuy vậy vẫn còn nhiều hạn chế và sai sót, cần nâng cấp và phát triển trong tương lai để cải thiện độ chính xác nhằm áp dụng được vào điều kiện thực tế.

Về hướng phát triển, có thể thêm nhiều nhãn dữ liệu, kết hợp với việc đếm phương tiện giao thông, xa hơn là kết hợp NLP nhằm nhận diện biển số xe.

IX. Trích dẫn

- [1] YOLOv3 <https://pjreddie.com/darknet/yolo/>
- [2] Keras <https://en.wikipedia.org/wiki/Keras>
- [3] TensorFlow <https://en.wikipedia.org/wiki/TensorFlow>