

CVPDL HW1

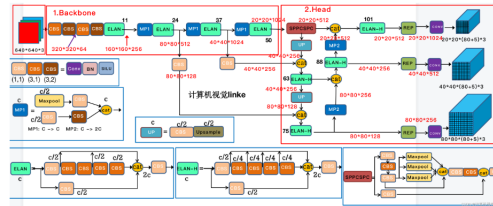
洪郡辰

R11944050

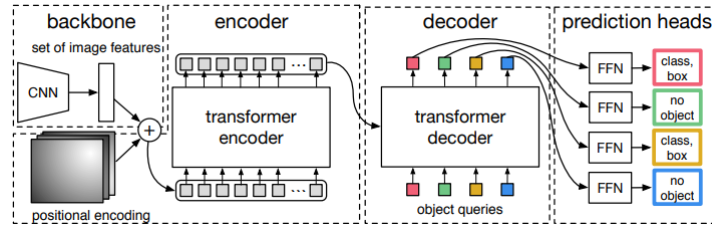
April 18, 2023

Problem

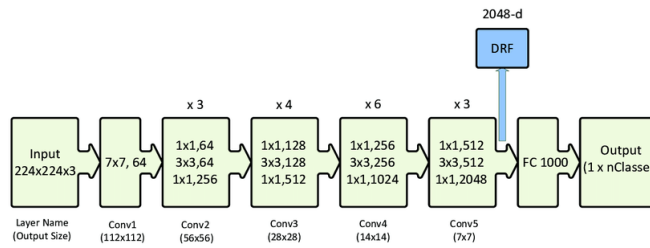
1.



(a) yolov7



(b) DETR



(c) resnet50(cnn for DETR)

Figure 1: model architecture

2.

Table 1: Public score of 4 tasks using multi-label and multi-class classification

Model	mAP@50	mAP@75	mAP@[50:5:95]
YOLOv7	0.7868	0.5157	0.5065
DETR-resnet50	0.5826	0.2889	0.3069

yolov7 的訓練時間約為 40min，DETR-resnet50 的訓練時間約為 12h，兩個 model 都有使用 pretrained-weight。可以發現 yolov7 的表現在各個指標都贏過 detr-resnet50 且訓練速度快很多。

3.

以下是兩個 method 的比較，先討論 yolov7:

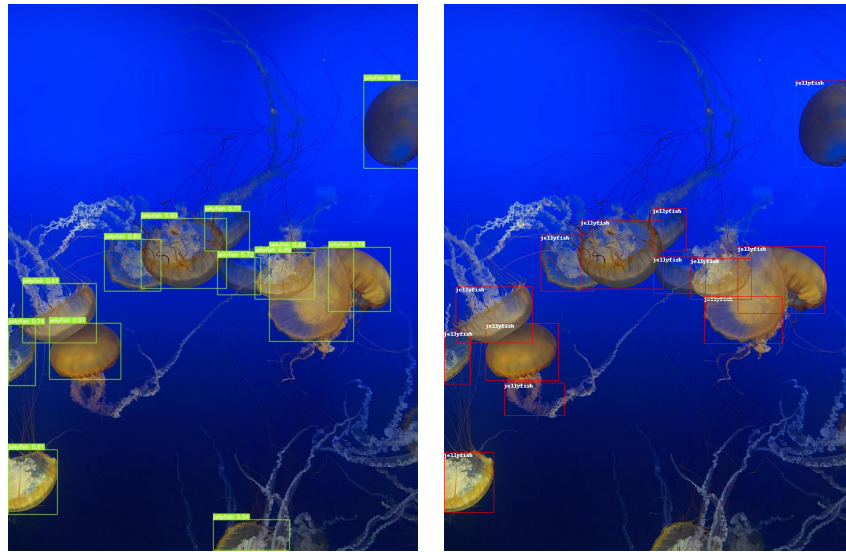
- augmentation
 - image resize: 將 image resize 成 640*640 進行訓練。
 - Mosaic Augmentation: 讀取四張圖片，對四張圖片做 image transform 然後進行拼接，進行 data augmentation。
 - Mixup Augmentation: 讓其他訓練圖片疊加在目前的 image 上，如同殘影一樣，進行 data augmentation。
- loss function
 - objectness loss: 使用 Binary Cross Entropy 計算 predicted objectness probability and CIOU with match target 的誤差。
 - box loss: 定義為 $\text{mean}(1-\text{CIOU})$ 。
 - classification loss: 使用 binary cross entropy 計算 predicted class probability for each anchor box and true class 的誤差。
 - loss weight: 可以對誤差以 predefined-weight 進行調整。
- scheduler
 - cosine learning rate:yolo 使用 cosine scheduler，可以透過 k hyperparameter 調整 lr 的變化方式。
 - linear warmup: 在一開始時 yolo 使用 linear warmup。
- structure
 - FPN:yolo 使用 cosine scheduler，可以透過 k hyperparameter 調整 lr 的變化方式。
 - linear warmup: 在一開始時 yolo 使用 linear warmup。
- optimizer
 - Adam

接下來討論 DETR:

- augmentation
 - image resize: 將 image resize 成 480*480 進行訓練。

- loss function
 - bipartite matching loss: 計算預測結果中的 classification prob. 及 bbox 對 ground truth 的誤差，會有一一對應的情況。
- scheduler
 - linear scheduler learning rate:yolo 使用 linear scheduler 調整 lr 的變化方式。
 - linear warmup: 在一開始時我有設定使用 linear warmup。
- optimizer
 - Adam

4.



(a) yolov7 result

(b) DETR result

Figure 2: img result

reference

- yolov7:<https://github.com/WongKinYiu/yolov7>
- huggingface:<https://huggingface.co/facebook/detr-resnet-50>
- huggingface script:https://huggingface.co/docs/transformers/tasks/object_detection
- huggingface script:https://colab.research.google.com/github/huggingface/notebooks/blob/main/examples/image_classification.ipynb
- yolo structure: https://github.com/akashAD98/yolov7_custom_data_trainig
- detr structure: <https://arxiv.org/pdf/2005.12872.pdf>

- resnet structure: [click here](#)
- yolov7 introduction: [click here](#)