Self-Driving Cars Radar Detection Report

Graduate Degree Program in Robotics
Student ID : 312605001
Name : Ou, Ting-Wei

1. Results

Model	GPU	Parameter	Training Time	Score
YOLO v8n	RTX3060	Epoch = 2000	11hr	0.7930
YOLO v8s	RTX3060	Epoch = 500	18hr	0.9189
YOLO v8m	RTX4060*2	Epoch = 2000	54hr	0.9073
YOLO v8m	RTX4060*2	Epoch = 3000	68hr	0.7140
YOLO v8l	RTX3060	Epoch = 1000	73hr	0.8680
YOLO v8l	RTX3060	Epoch = 2000	92hr	0.8045

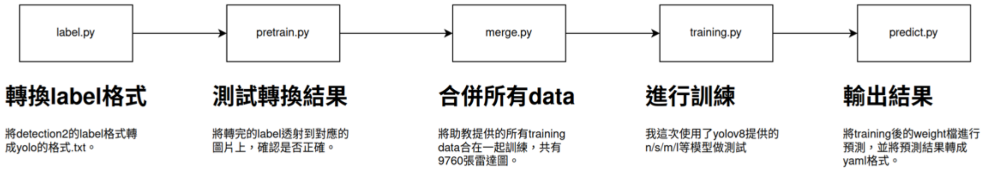
2. Contribution
   a. Using YOLO V8

      i. Why I choose YOLO V8

      YOLO is a widely recognized image recognition model, so in this project, I aim to assess its performance in radar detection tasks. Additionally, YOLO v8 introduced in 2023, may benefit from updated techniques, potentially improving task performance. Another advantage of the YOLO model is its ability to resume training if unexpectedly interrupted, ultimately saving time.

      Following my experiments, I observed that the accuracy of the task detection consistently maintained a level of around 0.8 or above, which is quite impressive.

      ii. Flow Chart



轉換label格式

將detection2的label格式轉成yolo的格式.txt。

測試轉換結果

將轉完的label透射到對應的圖片上，確認是否正確。

合併所有data

將助教提供的所有training data合在一起訓練，共有9760張雷達圖。

進行訓練

我這次使用了yolov8提供的n/s/m/l等模型做測試

輸出結果

將training後的weight檔進行預測，並將預測結果轉成yaml格式。

b. I provide the code to the teaching assistant for YOLO JSON conversion.

電控碩 / ICN 王可萱 2023年12月29日,17:23
收件人 機器人碩士 / ROB 歐庭維

感謝您！！

參考

機器人碩士 / ROB 歐庭維
2023年12月29日,14:50

助教好

昨天有提到關於yolo的部份

可以提供一下基於yolo訓練格式的detection程式碼，如附件所示。

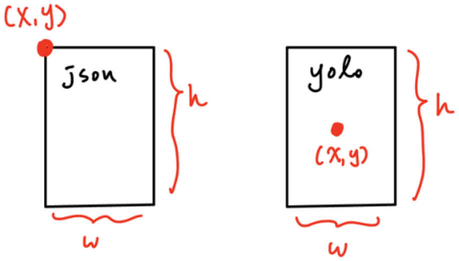
我確認過是可以正常執行的，謝謝。

predict.py 2.5KB

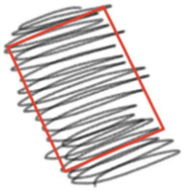
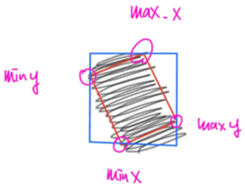
3. Problem & Solution

a. Label

There are some differences between the labels in training data and those used in YOLO. In annotations, the (x, y) represents the left-upper corner of the bounding box, but in yolo, whereas in YOLO, it donates the center of the bounding box.

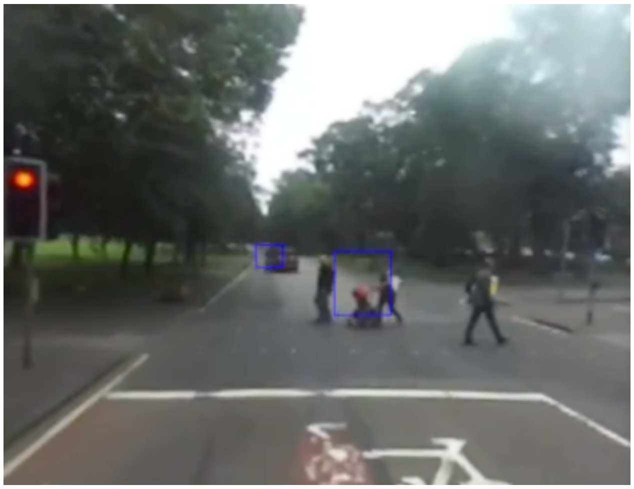


There is an issue when considering the rotation of the label, the object must be perfectly enclosed by the BBox to minimize the noise. However, according to YOLO's definition, the label must be a rectangle aligned with the image orientation. This implies that it is not possible to accurately outline objects with rotation using YOLO. To address this I enlarge the BBox to ensure all four sides of the object fit within it. While this approach may introduce some noise, the training process is designed to handle such uncertainties.

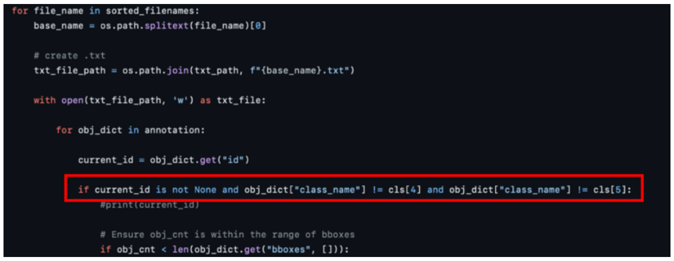
Without rotation	With rotation	Enlarge the BBox
		

b. Object Name

There are various types of objects in the training data. However, the objective is to detect the vehicles in radar image. Therefore, it is necessary to exclude pedestrians and groups of pedestrians from the training data.



This can be easily achieved by employing an “if” loop to filter out these specific labels.

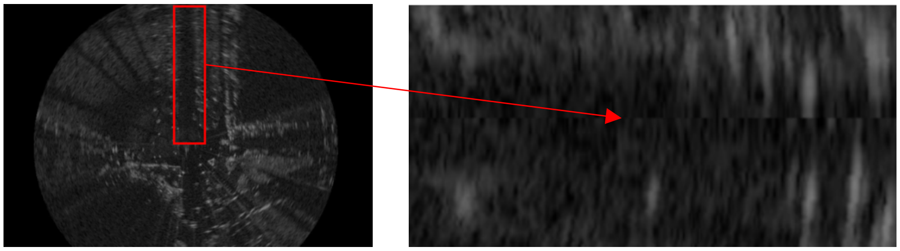


```
for file_name in sorted_filenames:
    base_name = os.path.splitext(file_name)[0]

    # create .txt
    txt_file_path = os.path.join(txt_path, f"{base_name}.txt")

    with open(txt_file_path, 'w') as txt_file:

        for obj_dict in annotation:

            current_id = obj_dict.get("id")

            if current_id is not None and obj_dict["class_name"] != cls[4] and obj_dict["class_name"] != cls[5]:
                #print(current_id)

                # Ensure obj_cnt is within the range of bboxes
                if obj_cnt < len(obj_dict.get("bboxes", [])):
```

4. Other Discussion or Findings

a. Radar Cartesian Image

When reviewing the training data, I observed that the radar cartesian coordinates do not appear to be compensated for the velocity of the cars, This issue can lead to a situation where the detector may not identify objects approaching from the heading direction of the cars.

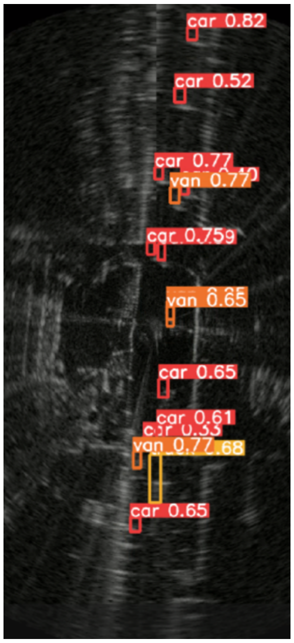


I came up with a potential solution to address this issue. When velocity compensation is not possible, the angle range of converting polar coordinates to cartesian coordinates can be adjusted from 0 to 360 to -180 to 180. This is because, for the detection of moving objects in autonomous vehicles, objects behind are generally less critical.

b. If I use all the object ID for training

I also wandering if I use all the labels(cars, vans , etc.) from the frame IDs, can the model's detection performance on radar images be comparable to that on normal RGB images?

However, the results show that there are some problems. Firstly, the object id may fluctuate due to the uncertainty in the radar shape, such as switching from a bus to a car and then back to a bus. Additionally, there are instances of misidentification, where many non-vehicle objects are falsely detected.



c. How to improve model Performance

I also discovered that, in terms of adjusting the model or the number of iterations, the highest accuracy I achieved was only 91%. This was because I initially set the 'imgsize' too small during training. Later on, when testing with a bonus dataset, I attempted training with the original size for 'imgsize,' resulting in slightly better performance than the default setting. To further enhance the model's performance, one can consider adjusting the learning rate and initializing parameters multiple times to prevent the model from getting stuck in local minima during convergence.