
# 利用不同物件偵測模型進行SAR image之物件辨識

朱信榮 統計所 R26111036
陳祈均 機械所 N16120145
丁柏仁 機械所 N16124482
王崇瀚 資訊所 P76121445
報告日期 : 2024/05/14

# 目錄

- 題目與資料集介紹
- 相關論文
- 使用模型與結果討論
- 遇到困難點與未來展望

# • 資料集介紹

- 資料集選用：SARDet-100K資料集
- 優點：

1. 將多個SAR(合成孔徑雷達) image資料集檔案彙整成一個大檔案，解決資料集大小不足問題
2. 透過彙整多個資料集，解決單一資料集訓練目標過於單一的問題

# • 資料集介紹

- 資料集組成：

| Dataset | Images | Instances | Target |
|---|---|---|---|
| AIR_SARShip1 | 501 | 1058 | S |
| AIR_SARShip1 | 300 | 2040 | S |
| HRSID | 5604 | 16969 | S |
| MSAR | 30158 | 65202 | A,T,B,S |
| SADD | 883 | 7835 | A |
| SAR-AIRcraft | 18888 | 38475 | A |
| ShipDataset | 39729 | 50885 | S |
| SSDD | 1160 | 2587 | S |
| OGSOD | 18331 | 48589 | B,H,T |
| SIVED | 1044 | 12013 | C |
| SARDet-100K | 116598 | 245653 | S,A,T,B,H,C |

- Target (6類):
  S代表ship,
  A代表aircraft,
  T代表tank,
  B代表bridge,
  H代表harbour,
  C代表car

- 因為資料集較為龐大，模型訓練時間較長。為了縮短訓練時間，訓練時只取其中的 **5000 筆資料作為training data**，**600筆資料作為validation data**

# • 相關論文

## • Multi-Scale Ship Detection Algorithm Based on YOLOv7 for Complex Scene SAR Images

**Zhuo Chen , Chang Liu , V. F. Filaretov and D. A. Yukhimets (2023)**

- 當在海面上處理多尺度船舶時，大多數現代演算法都可以成功實現船舶檢測結果。然而，在複雜場景的SAR影像中進行多尺度船舶目標偵測操作時，仍存在漏檢、誤辨識等問題。
- 此論文根據 YOLOv7 提出了一種複雜場景多尺度船舶偵測模型，稱為CSD-YOLO。
- 對 HRSID 和 SSDD 資料集的測試是為了對 CSD-YOLO 和 baseline YOLOv7 做出比較。

**Table 4.** Experimental results of different methods.

| Model | Dataset | Precision | Recall | mAP 0.5 |
|---|---|---|---|---|
| Faster R-CNN | SSDD | 81.63 | 85.31 | 89.63 |
| | HRSID | 88.81 | 72.57 | 77.98 |
| FCOS | SSDD | 84.15 | 92.52 | 90.61 |
| | HRSID | 75.53 | 73.79 | 77.95 |
| YOLOv3 | SSDD | 89.11 | 85.03 | 91.54 |
| | HRSID | 88.73 | 69.19 | 80.59 |
| YOLOv5s | SSDD | 95.14 | 90.01 | 96.28 |
| | HRSID | 84.69 | 75.11 | 83.34 |
| YOLOv7 | SSDD | 91.05 | 84.92 | 93.68 |
| | HRSID | 85.52 | 74.58 | 83.64 |
| CSD-YOLO | SSDD | 95.9 | 95.9 | 98.60 |
| | HRSID | 93.22 | 80.42 | 86.10 |

# • 相關論文

## • A Lightweight SAR Image Ship Detection Method Based on Improved Convolution and YOLOv7

**Hongdou Tang , Song Gao et al. (2024)**

- 利用SAR影像偵測船舶目標具有廣泛的應用前景。由於SAR影像中船舶的密度大、前景和背景雜波的極度不平衡以及目標尺寸的多樣性，實現輕量且高精度的多尺度船舶目標偵測仍然是一個巨大的挑戰。

- 本文提出了一種 multi-scale receptive fields convolution block（AMMRF）的 attention mechanism 。

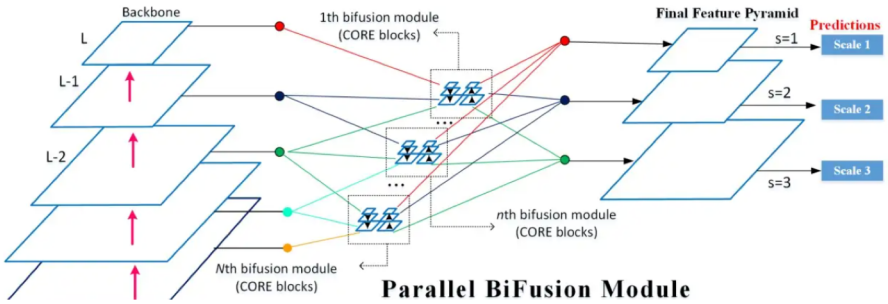
- 提出了一種基於YOLOv7的新型船舶目標檢測方法——You Only Look Once SAR Ship Identification（YOLO-SARSI），該方法獲取從高層卷積中提取的抽象語義信息，同時保留從低層卷積中提取的詳細語意資訊。

**Table 1.** Experimental results for the datasets. Bolded numbers indicate the best indicators. The bolded method is the one proposed in this paper.

| Dataset | Model | $AP_{50}$ (%) | $AP_{50:95}$ (%) | Params (M) |
|---|---|---|---|---|
| HRSID | Cascade R-CNN | 65.1 | 41.2 | 68.93 |
| | Faster R-CNN | 69.8 | 43.6 | 41.12 |
| | Mask R-CNN | 69.9 | 43.9 | 41.12 |
| | SDD 300 | 56.5 | 36.8 | 23.75 |
| | Swin Transformer | 57.1 | 32.6 | 36.82 |
| | RetinaNet | 70.9 | 45.5 | 36.10 |
| | YOLOv7 | 86.7 | 61.8 | 34.79 |
| | **YOLO-SARSI** | **89.3** | **64.0** | **18.43** |
| LS-SSDD-V1.0 | Cascade R-CNN | 55.4 | 20.1 | 68.93 |
| | Faster R-CNN | 63.4 | 23.9 | 41.12 |
| | Mask R-CNN | 63.3 | 24.1 | 41.12 |
| | SDD 300 | 32.5 | 10.1 | 23.75 |
| | Swin Transformer | 37.0 | 10.2 | 36.82 |
| | RetinaNet | 64.9 | 24.8 | 36.10 |
| | YOLOv7 | 69.8 | 27.7 | 34.79 |
| | **YOLO-SARSI** | **73.7** | **28.5** | **18.43** |

# • 使用模型

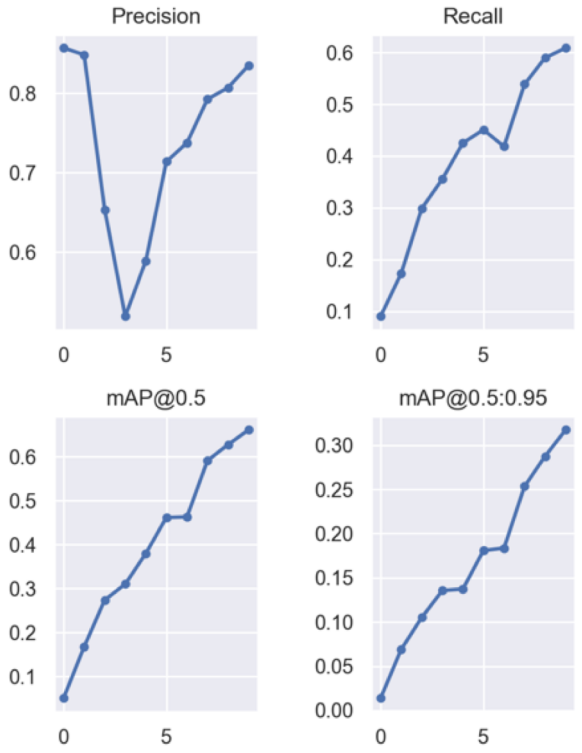
- YOLOv7
- YOLOv7x
- PRB-FPN (取代YOLO中backbone，特別適用於小物件偵測)



# Yolo v7

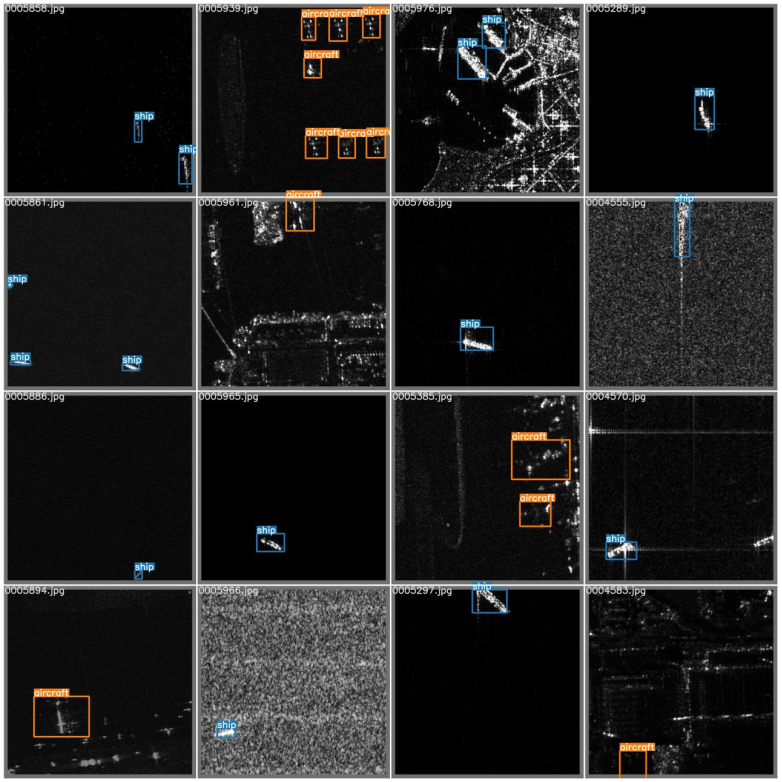
Train : 5000 筆; Val : 600 筆; Epoch : 10; Batch size : 32
Model summary : 415 layers; 37,223,526 parameters; 105.2 GFLOPS

```
     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       6/9     23.6G   0.04416  0.006385  0.003363   0.05391        26       640: 100%|| 157/157 [30:35<00:00, 11.69s/
                 Class      Images      Labels           P           R      mAP@.5  mAP@.5:.95: 100%|| 10/10 [00:52<00:0
                   all         600        1217       0.737       0.418       0.463       0.184

     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       7/9     23.6G   0.04337  0.006352  0.003193   0.05292        27       640: 100%|| 157/157 [30:19<00:00, 11.59s/
                 Class      Images      Labels           P           R      mAP@.5  mAP@.5:.95: 100%|| 10/10 [00:52<00:0
                   all         600        1217       0.792       0.539       0.591       0.253

     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       8/9     23.6G   0.04107  0.006157  0.002308   0.04954        20       640: 100%|| 157/157 [30:19<00:00, 11.59s/
                 Class      Images      Labels           P           R      mAP@.5  mAP@.5:.95: 100%|| 10/10 [00:52<00:0
                   all         600        1217       0.807        0.59       0.627       0.287

     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       9/9     23.6G   0.03958  0.005864  0.001797   0.04724        19       640: 100%|| 157/157 [30:10<00:00, 11.53s/
                 Class      Images      Labels           P           R      mAP@.5  mAP@.5:.95: 100%|| 10/10 [00:55<00:0
                   all         600        1217       0.835       0.609       0.661       0.317
                  ship         600         719       0.805       0.782       0.838       0.422
              aircraft         600         353       0.743       0.581       0.679       0.299
                   car         600          35       0.804       0.743        0.83       0.484
                  tank         600         104       0.823       0.938        0.93       0.373
                bridge         600           6           1           0       0.027     0.00873
10 epochs completed in 5.265 hours.
```

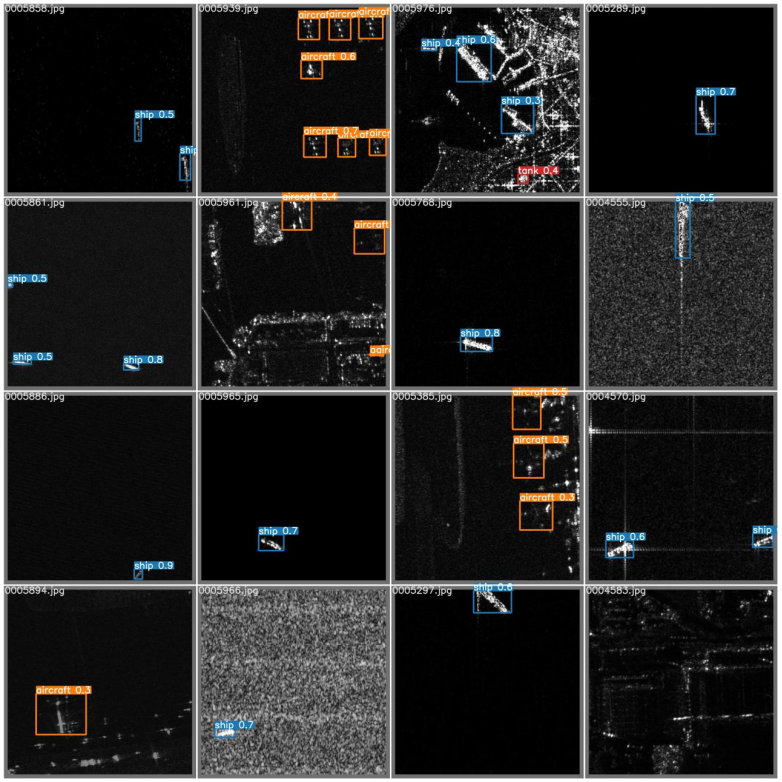


# Yolo v7

Ground truth



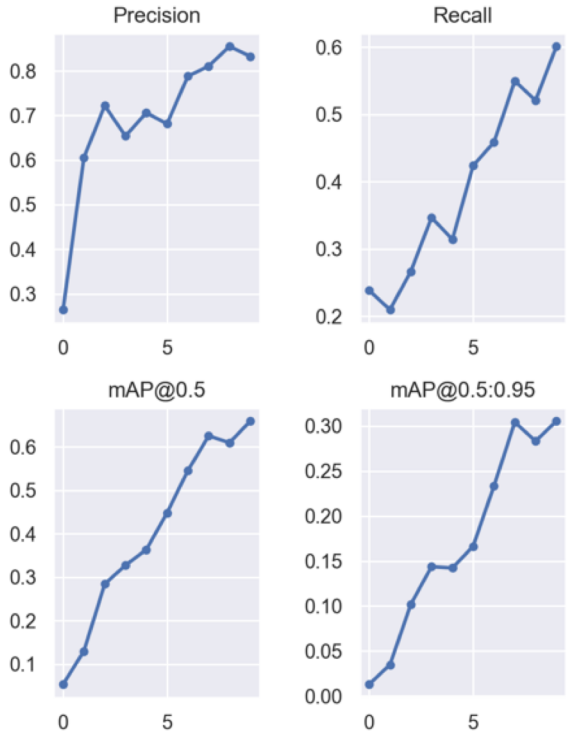
pred



# Yolo v7x

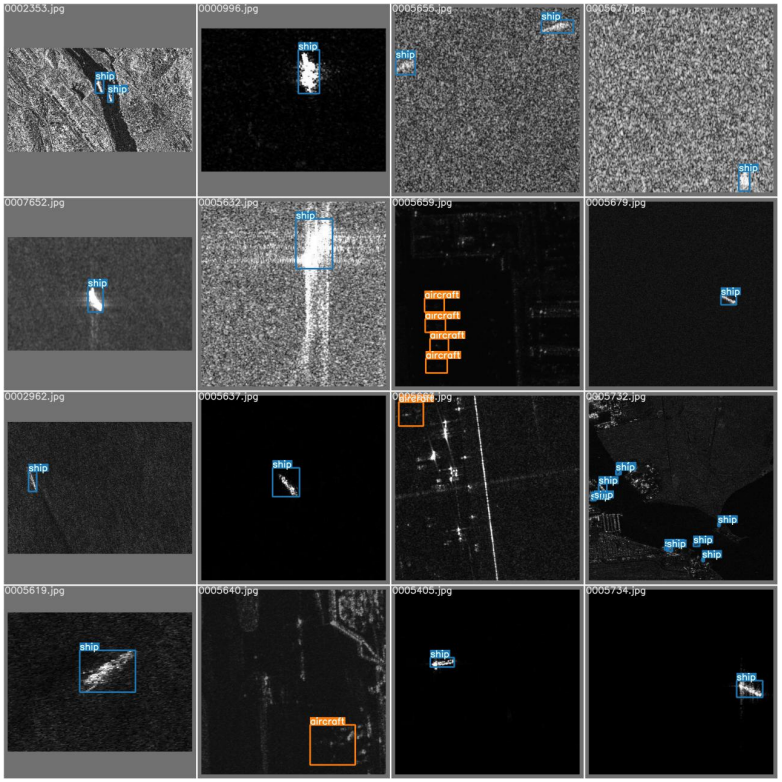
Train : 5000 筆; Val : 600 筆; Epoch : 10; Batch size : 24
Model summary : 467 layers; 70,848,782 parameters; 189.0 GFLOPS

```
     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       6/9     21.8G   0.04412  0.006602  0.003416   0.05414        38       640: 100%|| 209/209 [39:49<00:00, 11.43s/
                 Class     Images     Labels          P          R     mAP@.5  mAP@.5:.95: 100%|| 13/13 [01:33<00:0
                   all        600       1217      0.788      0.458      0.545      0.233

     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       7/9     21.8G   0.04176  0.006288  0.002428   0.05047        25       640: 100%|| 209/209 [38:56<00:00, 11.18s/
                 Class     Images     Labels          P          R     mAP@.5  mAP@.5:.95: 100%|| 13/13 [01:33<00:0
                   all        600       1217      0.811       0.55      0.625      0.304

     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       8/9     21.8G   0.04018  0.006237  0.002315   0.04873        25       640: 100%|| 209/209 [36:32<00:00, 10.49s/
                 Class     Images     Labels          P          R     mAP@.5  mAP@.5:.95: 100%|| 13/13 [01:32<00:0
                   all        600       1217      0.855       0.52      0.609      0.283

     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       9/9     21.8G   0.03899  0.006068  0.001791   0.04685        23       640: 100%|| 209/209 [36:25<00:00, 10.45s/
                 Class     Images     Labels          P          R     mAP@.5  mAP@.5:.95: 100%|| 13/13 [01:37<00:0
                   all        600       1217      0.833      0.601      0.659      0.305
                  ship        600        719      0.776       0.79      0.829      0.418
              aircraft        600        353      0.815      0.587      0.753      0.335
                   car        600         35      0.706      0.771       0.79      0.444
                  tank        600        104      0.866      0.856      0.911      0.328
                bridge        600          6          1          0      0.012    0.00216
10 epochs completed in 6.525 hours.
```

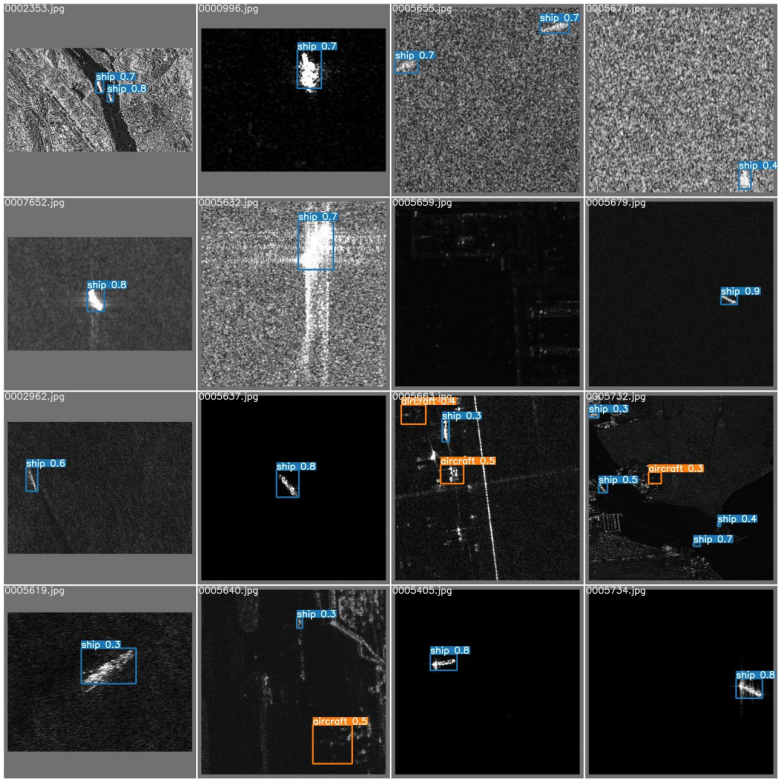


# Yolo v7x

Ground truth



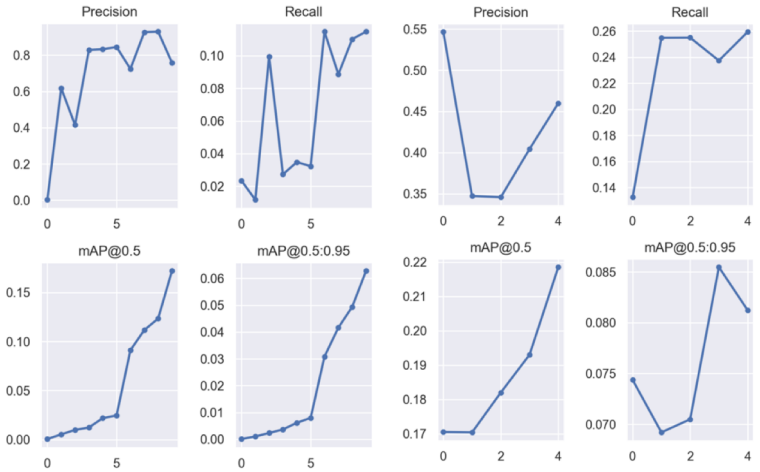
pred



# PRB-FPB-CSP

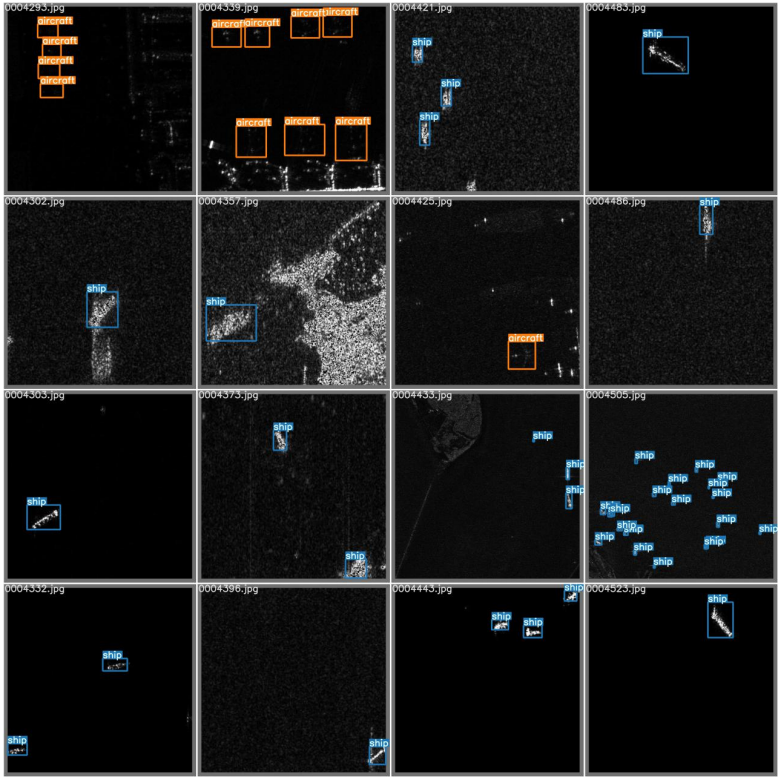
Train : 5000 筆; Val : 600 筆; Epoch : 15; Batch size : 24
Model summary : 698 layers; 61,303,171 parameters; 153.3 GFLOPS

```
     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       0/4     8.99G   0.05304  0.008204   0.01113   0.07238        15       640: 100%|| 209/209 [1:45:57<00:00, 30.42
                Class      Images      Labels           P           R      mAP@.5  mAP@.5:.95: 100%|| 13/13 [00:16<00:0
                  all         600        1217       0.547       0.133       0.171      0.0744

     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       1/4       21G   0.05262   0.00825    0.0107   0.07157        39       640: 100%|| 209/209 [1:23:43<00:00, 24.03
                Class      Images      Labels           P           R      mAP@.5  mAP@.5:.95: 100%|| 13/13 [01:32<00:0
                  all         600        1217       0.347       0.255        0.17      0.0692

     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       2/4       21G   0.05313  0.008046   0.01006   0.07124        24       640: 100%|| 209/209 [1:10:10<00:00, 20.14
                Class      Images      Labels           P           R      mAP@.5  mAP@.5:.95: 100%|| 13/13 [01:33<00:0
                  all         600        1217       0.346       0.255       0.182      0.0705

     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       3/4       21G   0.05265  0.007977   0.01038     0.071        25       640: 100%|| 209/209 [1:13:16<00:00, 21.04
                Class      Images      Labels           P           R      mAP@.5  mAP@.5:.95: 100%|| 13/13 [01:34<00:0
                  all         600        1217       0.404       0.237       0.193      0.0855

     Epoch   gpu_mem       box       obj       cls     total    labels  img_size
       4/4       21G   0.05187   0.00784   0.01008   0.06979        17       640: 100%|| 209/209 [1:06:08<00:00, 18.99
                Class      Images      Labels           P           R      mAP@.5  mAP@.5:.95: 100%|| 13/13 [01:35<00:0
                  all         600        1217       0.46       0.259       0.218      0.0812
                 ship         600         719       0.314       0.807       0.701       0.265
             aircraft         600         353       0.224       0.266       0.157      0.0407
                  car         600          35       0.704       0.137       0.193      0.0896
                 tank         600         104      0.0578      0.0865      0.0191     0.00666
               bridge         600           6           1           0      0.0219     0.00438
5 epochs completed in 6.768 hours.
```

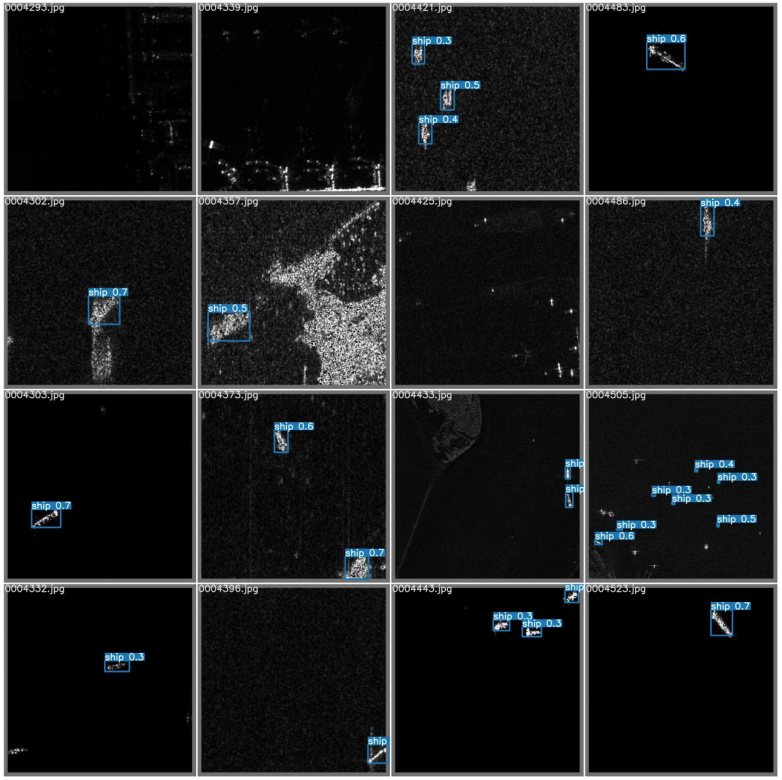


# PRB-FPB-CSP

Ground truth



pred



# • 結果討論

- YOLOv7 與 YOLOv7x 表現差不多，且與原始論文相比在 mAP 上表現較佳

- PRB-FPN 表現較差，可能原因:
  1. YOLOv7 有 pretrain weight，但 PRB-FPN 直接以初始權重作訓練
  2. PRB-FPN主要用於自駕車魚眼，可能不適合SAR Image

- PRB-FPN在船隻判別上較YOLO 為佳

**Table 4:** Comparison of different pretrain strategies using Faster-RCNN and Res50 as the detection model.

| ID | Model Input | Pretrain | | | mAP ↑ |
|---|---|---|---|---|---|
| | | Multi-stage | Dataset | Component | |
| 1 | SAR (Raw pixels) | ✗ | ImageNet | Backbone | 49.0 |
| 2 | | ✓ | ImageNet + DIOR | Framework | 49.5 |
| 3 | | ✓ | ImageNet + DOTA | Backbone | 49.3 |
| 4 | | | | Framework | **50.2** |

[Yuxuan Li et al., 2024]

# • 遇到困難點

- 背景知識不足
- 資料量過大難以處理
- 理解目標花費時間較多
- 實作方面

# • 未來展望

- 將所有資料執行訓練
- 嘗試其他特徵處理
- PRB-FPN先進行預訓練