

深度學習模型於工業表面缺陷檢測之比較研究

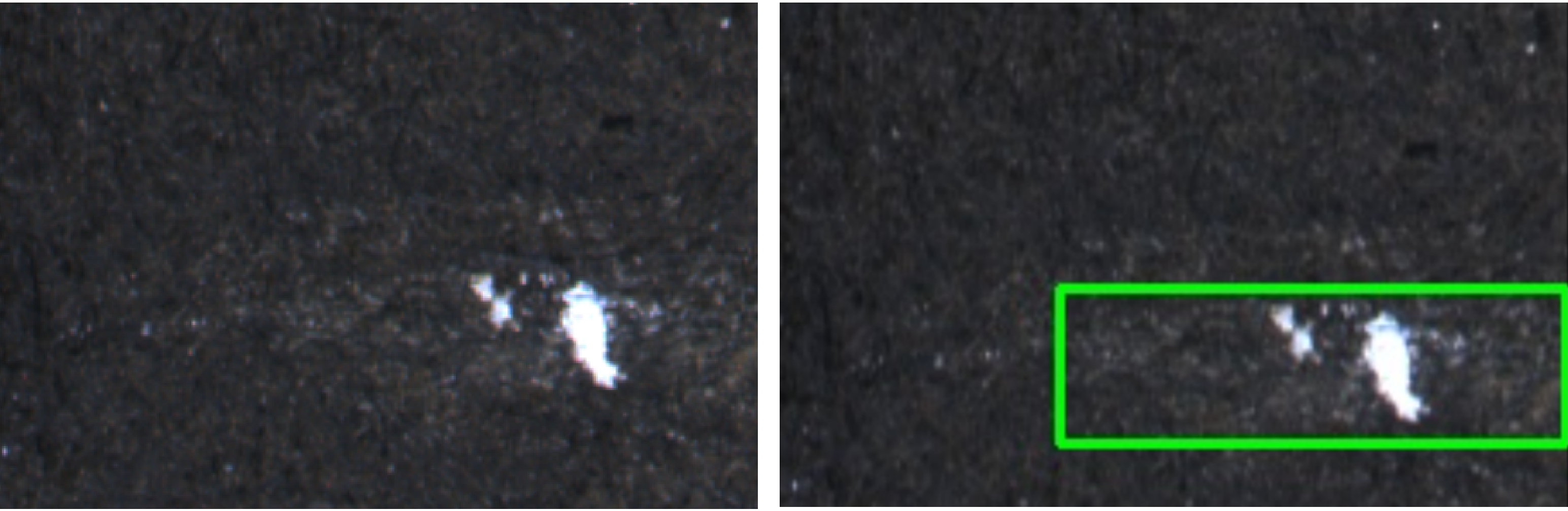
A Comparative Study of Deep Learning Models for Industrial Surface Defect Detection

成員: 魏靖豪、黃苡愷、呂紹恩

指導老師：洪昌鈺

研究動機

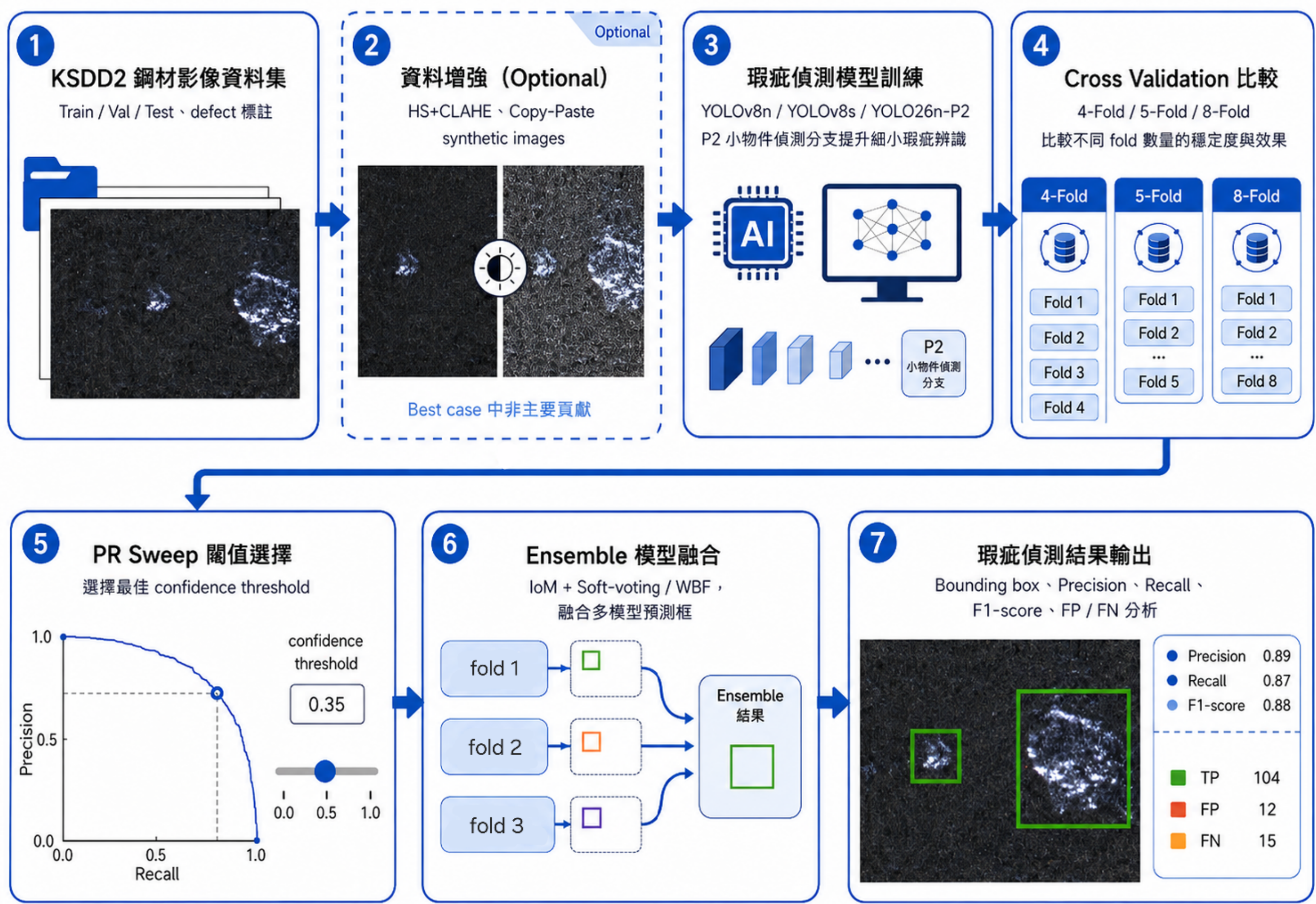
工業產品的表面瑕疵檢測通常依賴人工目視檢查，但人工檢查容易受到疲勞、主觀判斷與檢測速度限制影響，導致漏檢或誤檢問題。因此，本專題以 KSDD2 工業表面瑕疵資料集為基礎，設計一套基於深度學習的自動化瑕疵偵測模型，目標是從影像中判斷是否存在瑕疵，並以邊界框標示瑕疵位置。



資料集瑕疵範例原始圖片

資料集瑕疵範例標記

研究方法



- 1.KSDD2 官方訓練集共有 2,331 張影像，其中 246 張含有可見瑕疵、2,085 張為無瑕疵影像；測試集共有 1,004 張，其中 110 張含瑕疵。
- 2.資料增強方面，我們使用了Copy-Paste方法，將訓練集中的缺陷框，隨機進行縮放後貼在任一訓練資料中，並將此建立為一筆新的資料圖。Histogram Stretching(HS)+CLAHE用於強化暗部細節與對比度。
- 3.訓練yolo系列模型比較不同版本的差異，並且嘗試微調模型架構來強化訓練成果。
- 4.將訓練資料切割多份，做交叉驗證優化模型表現。

- 6.透過交叉驗證方法會產生多份的模型預測結果，透過比較不同融合方法(WBF/NMS/IoM/Soft-voting)，得到的結果為IoM+Soft-voting的結果最好，IoM是Intersection over Minimum:

$$IoM(A, B) = \frac{|A \cap B|}{\min(|A|, |B|)}$$
if $IoM \geq \text{threshold}$, then the two boxes are considered the same defect.

$$S_{\text{fused}} = \frac{1}{K} \sum_{i \in G} s_i$$

K = total number of fold models, G = matched box group

IoM 以較小框面積作為基準，判斷不同模型的偵測框是否代表同一個缺陷。Soft-Voting 將匹配框的分數除以總模型數，降低單一模型誤檢的影響，並保留多模型共識較高的缺陷。最後透過 confidence-weighted average 融合框座標，得到更穩定的偵測結果。

成果

表一展示了各個yolo系列模型在預設狀態下的表現

Method	F1-score	Precision	Recall
YOLO26n baseline	0.8340	0.8223	0.8461
YOLO11n baseline	0.8207	0.7803	0.8655
YOLO26s baseline	0.8072	0.7723	0.8455
YOLOv8s baseline	0.7606	0.8182	0.7105
YOLOv8n baseline	0.7383	0.8316	0.6639

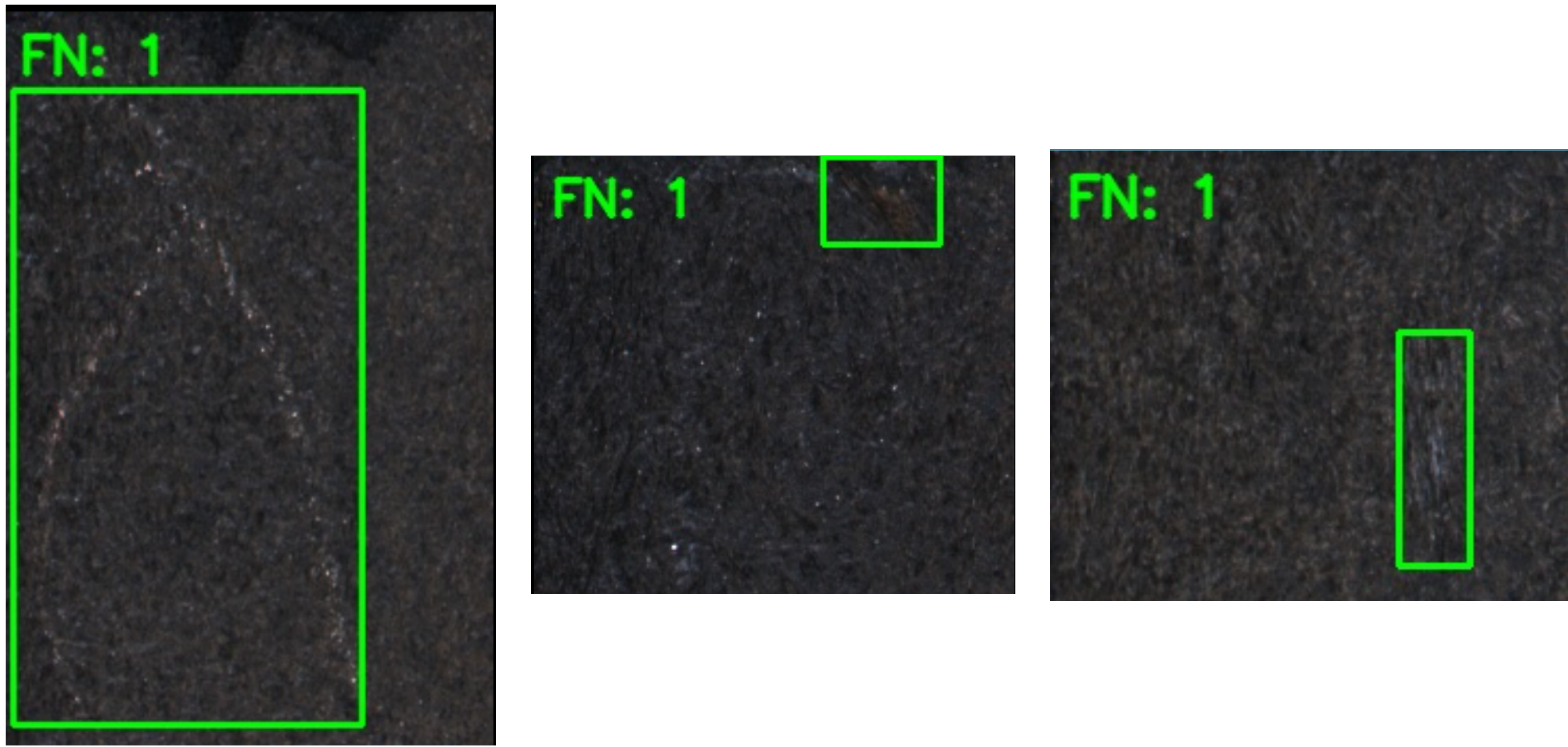
表二展示在yolov8n模型在不同的架構與實驗設定中的結果比較

Method	F1-score	Precision	Recall
YOLOv8n + 5-fold IoM + Soft-voting	0.8636	0.9406	0.7983
YOLOv8n + 8-fold IoM + Soft-voting	0.8609	0.8919	0.8319
YOLOv8n + 5-fold IoM + Soft-voting + Copy-Paste	0.7556	0.8019	0.7143
YOLOv8n baseline	0.7383	0.8316	0.6639
YOLOv8n + Copy-Paste	0.7342	0.7073	0.7632
YOLOv8n + 4-fold WBF	0.7124	0.7281	0.6975

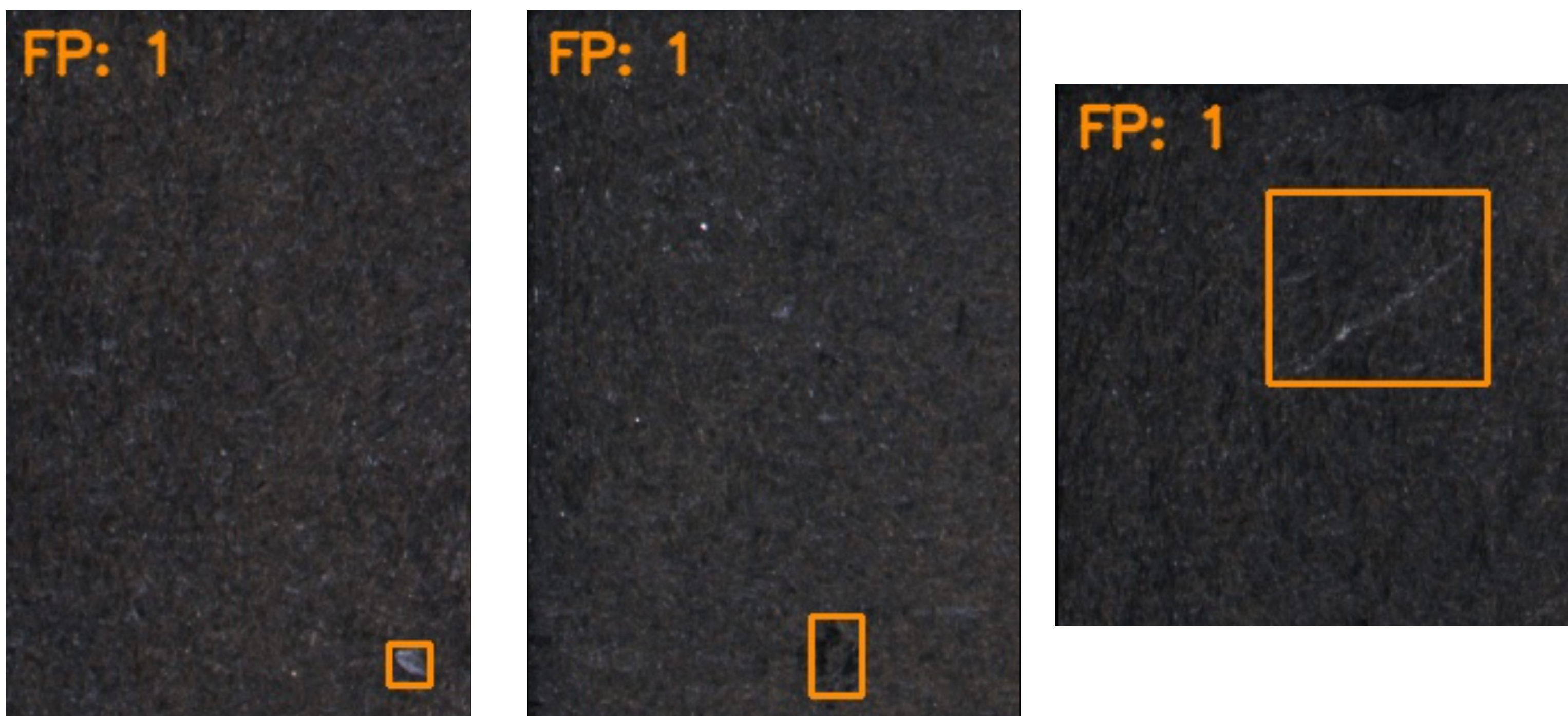
表三展示yolo26n在架構比較與元件消融，包括 P2 head、輸入解析度及不同偵測器，與分割任務下的表現。

Method	F1-score	Precision	Recall
YOLO26n + 5-fold IoM + Soft-voting	0.8851	0.8900	0.8700
YOLO26n + 5-fold IoM + Soft-voting + P2 head	0.8800	0.9300	0.8319
Faster R-CNN ResNet50-FPN	0.8571	0.9490	0.7815
YOLO26n + 8-fold IoM + Soft-voting	0.8430	0.8293	0.8571
YOLO26n + 5-fold WBF + Copy-Paste	0.8370	0.8796	0.7983
YOLO26n baseline	0.8340	0.8223	0.8461
YOLO26n + 5-fold IoM + Soft-voting + 1280 image size	0.7004	0.7034	0.6975
Mask R-CNN + ResNet-50-FPN*	0.8374	0.7900	0.8909

*This result is reported under the segmentation task, while the other methods are evaluated under the detection task.



模型漏檢代表案例



模型誤檢代表案例