

基於多輪影像修復之零樣本工業異常檢測

Zero-Shot Industrial Anomaly Detection via Multi-Round Image Repair

指導教授：許舒涵

組員：黃睿淞

研究動機與目的

工業瑕疵檢測難以蒐集足量的瑕疵樣本，僅憑「正常外觀」就要找出未知瑕疵的 zero-shot 方法因而格外重要，其中 DIVAD[1] 利用 Diffusion model 把測試圖還原成正常的樣子後與原圖相減，得到 anomaly map。與多數需要正常樣本訓練的重建式方法、以及需預先列舉大量「異常」描述（刮痕、破裂、污點…）的 CLIP- / DINO-based 方法不同，DIVAD 是 training-free 且不依賴異常 prompt 的，對未知瑕疵更通用。然而它仍有兩個結構性限制：其一，全圖重建會讓正常區也被重新合成出細微差異、使背景充滿低強度雜訊而大量 false positive；其二，單輪、低解析度（512）的比較使瑕疵即使被偵測到也定位粗糙，且一張圖內散佈的同型瑕疵會被最強的一兩處主導、其餘被壓低而漏檢。

為此，本專題在 DIVAD 之上串接 FLUX.1-Fill-dev[2] inpaint 模型，組成一條「重建 → 比對 → 修復」的 mask-guided 兩輪局部修復 pipeline：只重繪可疑區、其餘保留原像素以壓制背景 false positive，以 1024 高解析 inpaint 讓瑕疵定位更精準，再用 DINO 全圖檢索找回散佈在他處的同型瑕疵。整套方法全程維持 zero-shot，只需一句正常外觀描述即可切換類別。

系統架構

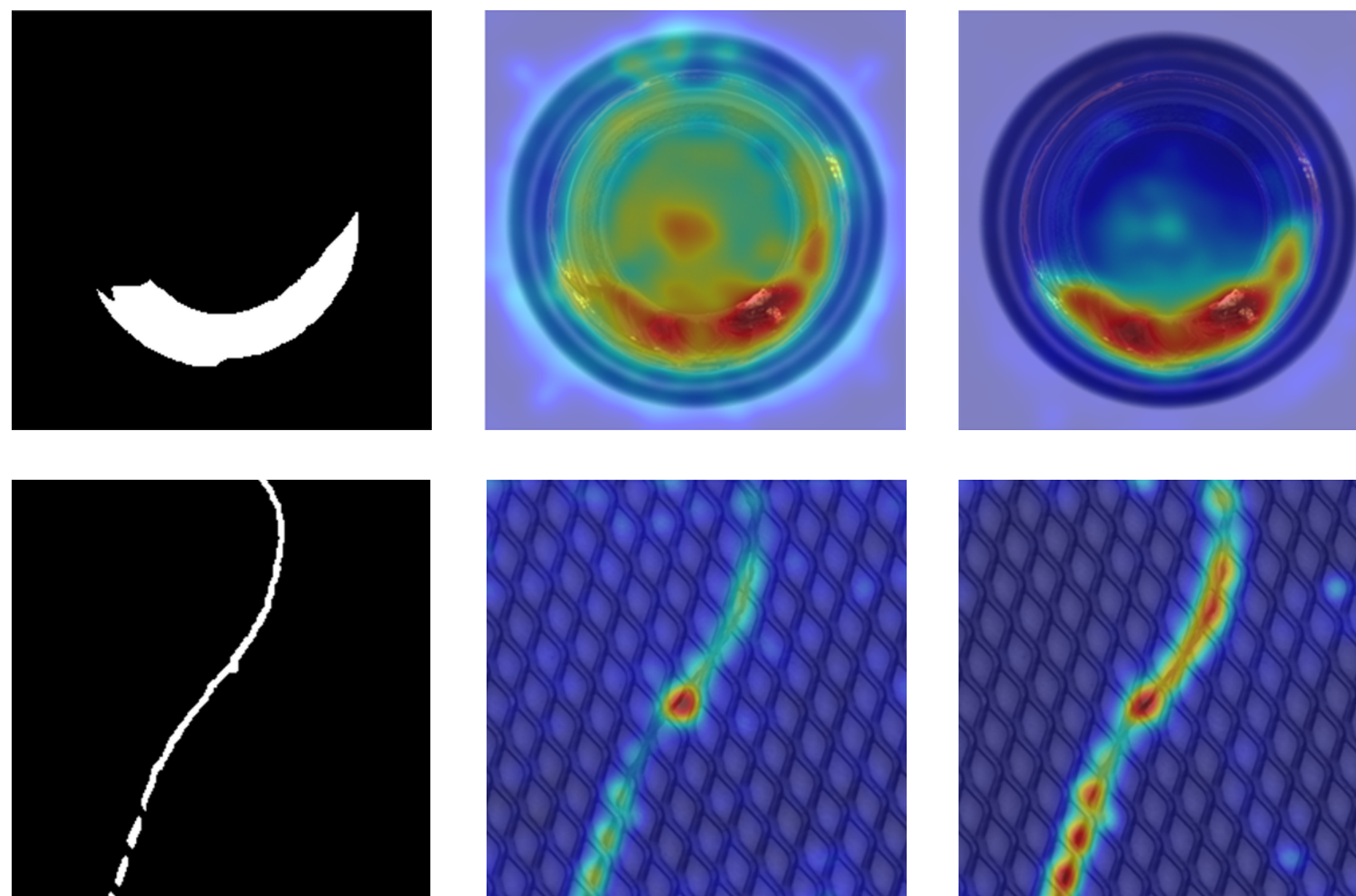
本方法的核心邏輯在於 FLUX inpaint 的「修復一致性」：把可疑區交給 FLUX 重畫成正常外觀後，原本就正常的地方修復前後幾乎不變，真正有瑕疵的地方則會被重畫成正常、與原圖的差距明顯擴大，因此只要量測「原圖與修復圖的差異」就能精準定位異常。整體流程分為兩輪：

- 第一輪 | 粗定位 → 高解析局部修復**：先以 DIVAD（Stable Diffusion 2.1[3] + DDIM inversion[4]）把原圖反演後重新去噪、得到一張「正常版」重建圖，再用 DINO ViT-S/8[5] 抽取 patch 特徵、以 $1 - \cos$ 算出原圖與重建圖的初步異常熱圖；此熱圖會經過自適應 M0 mask（percentile + 絕對下限 + Otsu[6] 多重門檻，連通分量過濾）最後交給 FLUX.1-Fill-dev 在 1024×1024 高解析度把 M0 區域 inpaint 重畫成正常外觀。
- 第二輪 | DINO bank 全圖檢索 → 再修復**：進一步解決散佈型同型瑕疵的漏檢：先比對原圖與第一輪修復圖，把真正被改變的 patch 視為已驗證缺陷、建成 DINO query bank，再用它對整張原圖做 max-cosine 相似度檢索，找回第一輪沒框到、散佈在他處的同型瑕疵；檢索結果經自適應 sim mask（median + $k \cdot \text{MAD}$ 抗重複紋理、連通分量抗雜訊、覆蓋率上限）與第一輪的差異區聯集並做第二輪 FLUX inpaint。最終異常圖取原圖與第二輪修復圖的 DINO patch $1 - \cos$ 距離，非重繪區因保留原像素、特徵幾乎相同而趨近 0，異常分數自然集中在真正修不回去之處。

實驗結果

Method	iAUC	pAUC	PRO	pF1	pAP
DIVAD (論文)	76	88.6	73.3	31.6	35.7
Ours	80.52	84.11	76.46	35.88	31.51
絕對差距 Δ	4.52	-4.49	3.16	4.28	-4.19
相對提升	5.95%	-5.07%	4.31%	13.54%	-11.74%

▲表一、整體五項指標對比（Ours vs DIVAD，單位 %）



▲圖二、Ground truth / DIVAD / Ours

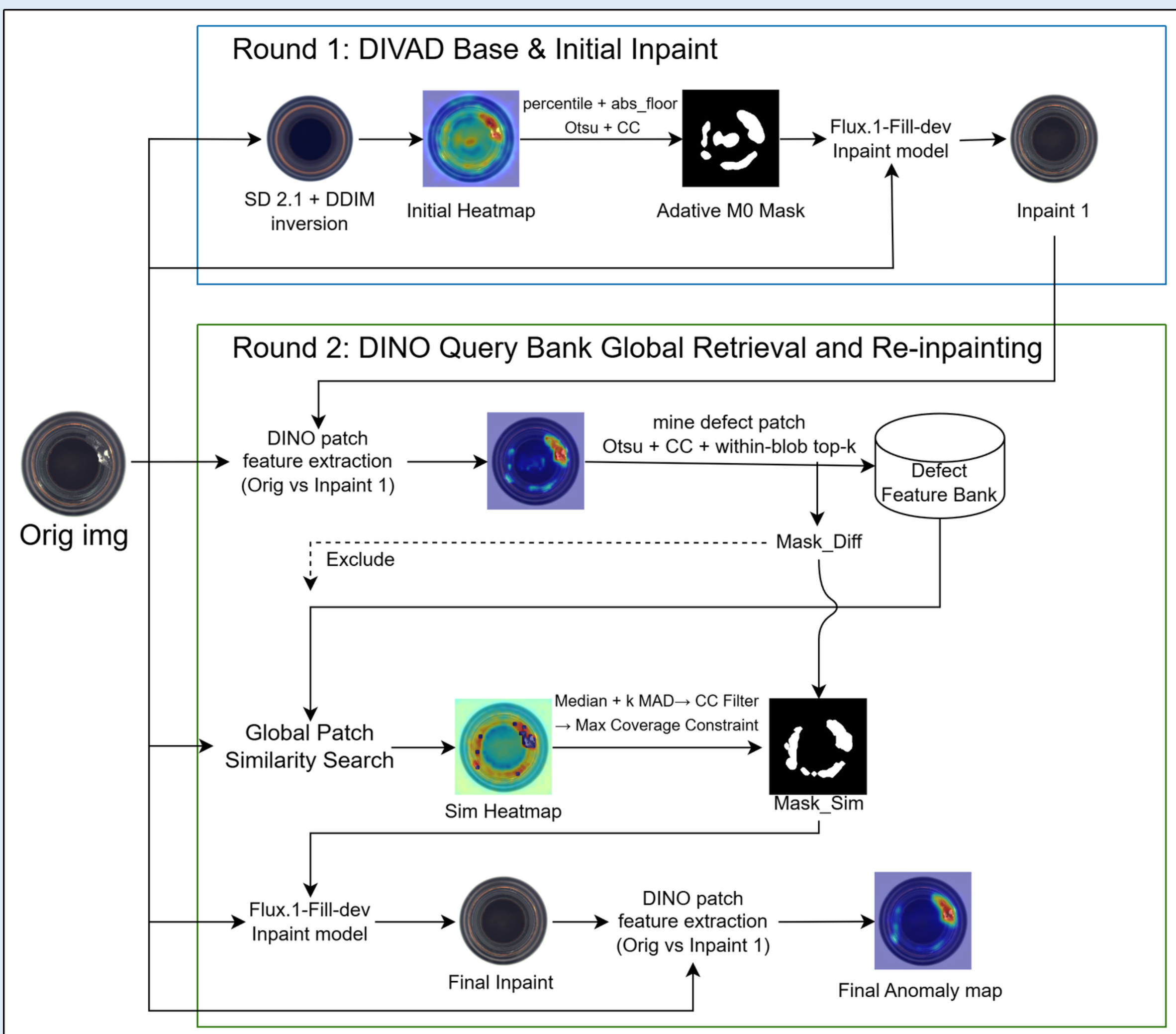
本研究於 MVTec AD[7] 完整 15 類驗證，在 5 項核心指標中取得 3 勝 2 負 的佳績，核心指標 pF1 大幅提升 +13.54%。得益於創新的 Mask-Guided 重繪機制，系統能精準聚焦可疑區域並完整保留背景，大幅壓低產線最棘手的背景誤報率（False Positive）。

全閾值指標（pAUC、pAP）的輕微降幅，屬於預期內的技術權衡。在像素極度不平衡下，pAUC 本就對 DIVAD 殘留之背景雜訊具備高寬容度，致使基準分數偏高；而 pAP 的微降，則是因 Mask 外部未被重繪的真實瑕疵分數趨近於 0，導致在極端閾值下評分受限。這兩項全閾值指標的修正，正是模型為了在固定閾值下達到最佳表現所做出的合理收斂。

相較於理論上的全閾值排序，本方法成功在產線實務最關鍵的「單一固定閾值」下，換取了定位純度。此項突破有效解決了背景雜訊干擾，展現出極佳的穩定度與實用性，符合工業量產落地部署之需求。

參考文獻

- [1] S. Hicsonmez et al., ICPR 2026. (DIVAD)
- [2] Black Forest Labs, FLUX.1 Fill 2024.
- [3] R. Rombach et al., CVPR 2022. (LDM/SD)
- [4] J. Song et al., ICLR 2021. (DDIM)
- [5] M. Caron et al., ICCV 2021. (DINO)
- [6] N. Otsu, IEEE Trans. SMC, 1979. (Otsu)
- [7] P. Bergmann et al., CVPR 2019. (MVTec AD)



▲圖一、系統架構圖