|  |
| --- |
| 第15組  **元利盛醫療零件瑕疵檢測** |

|  |  |  |  |
| --- | --- | --- | --- |
| 組長 | 機械所 | 113323116 | 王康鋕 |
| 組員 | 機械所 | 113323098 | 葉律旻 |
| 組員 | 機械所 | 113323096 | 張富勝 |

指導教師： 李朱育 教授

陳怡呈 教授

林智揚 教授

廖展誼 教授

－第15組工作分配表－

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 姓名 | 負責工作 | 權重\* | 簽名 |
| 組長 | 王康鋕 | 資料整合、標記 | 0.9 |  |
| 組員 | 葉律旻 | SAM訓練 | 1.2 |  |
| 組員 | 張富勝 | Yolo訓練 | 0.9 |  |

註:

\*權重由1.4~0.6，但總和必須為人數。例: 5人之組權重總和須為5.0。

各組員之書面報告分數 = 各組書面報告分數乘以權重。

－目錄－

1. 摘要 3
2. 前言 5
3. 專題研究方法 9
4. 結果與討論 11
5. 結論 13
6. 未來研究方向
7. 心得 15
8. 參考文獻 17
9. 附錄 19

(所有內文字體12點：中文標楷體、英文Times New Roman)

1. 摘要

(300字內，列出方法、目的和主要結果)

本次研究以醫療零件瑕疵檢測為主題，針對表面亮度不均與瑕疵顏色太淡的挑戰，提出以AI進行分割的解決方案，目標是比較 YOLOv11 與 SAM 在瑕疵分割能力上的表現。

使用方法: 資料標記、數據擴增、模型訓練與評估。其中 YOLO 使用 labelme 標記瑕疵位置與類型，SAM 使用多邊形註解和 Bounding Box 作為提示。兩種方法分別進行模型調優及結果驗證；在 YOLO 模型中，透過數據分配至訓練與驗證集，採用 yolo11n-seg.pt 作為基礎，設定200個 epoch 進行訓練；SAM 則針對圖像建立 Ground Truth Mask 以優化模型。

主要結果: YOLO的mAP50達到0.97、mAP50-95達到0.52。SAM 的mAP50達到0.9225、mAP50-95達到0.73，為醫療零件檢測提供更高效的解決方案。YOLO 與 SAM 在瑕疵分割各有優勢。

1. 前言

(包含研究背景、動機與目的等)

(1) **研究背景**：在醫療零件的生產過程中，產品表面的瑕疵可能直接影響其功能和使用安全。然而，瑕疵檢測面臨著由於零件表面曲面造成亮度不均，以及部分瑕疵顏色太淡難以區分的挑戰。傳統的 AOI (自動光學檢測) 技術在過濾背景與瑕疵時，效果有限，因此亟需新穎的技術來解決此問題。

(2) **研究動機**：隨著人工智慧技術的快速發展，深度學習模型在影像分割和檢測領域展現出卓越的能力。將這些技術應用於醫療零件瑕疵檢測，預期能大幅提高檢測準確率，降低人工檢測成本，並提升產品品質和市場競爭力。因此，本研究旨在探討深度學習技術在醫療零件檢測的應用潛力。

(3) **研究目的**：本研究以 YOLOv11 與 Segment-Anything Model (SAM) 為核心技術，透過影像數據標記和模型訓練，進行瑕疵分割能力的比較。目標是針對表面瑕疵難以檢測的問題，建立一套可靠且高效的解決方案，進一步提升檢測準確率，預期 YOLO mAP50達 80% 以上，SAM 則達 85% 以上。

1. 專題研究方法

(例如：依序列出影像處理、Yolov8、機器人操作、手眼校正、路徑規劃、流程圖、分類任務等，可自行加小節，如3.1、3.2等)

**(3.1) 資料標註與增強**: (I) 使用 Labelme 工具對瑕疵進行標記，標記包含瑕疵位置與類型，為深度學習模型提供準確的標注數據。(II) 進行資料增強，包括水平翻轉、垂直翻轉、隨機旋轉、馬賽克增強、混和增強、色調、飽和度、亮度增強、隨機平移、隨機縮放、隨機剪切、隨機透視變換，增強模型的泛化能力並平衡資料分布。(III) 將標記生成的 JSON 文件轉換為模型所需的 TXT 格式，並按 8:2 比例分配至訓練集和驗證集。

**(3.2) YOLOv11 模型訓練:** (I) 採用 YOLOv11 的 Segmentation 模型進行訓練，使用 Yolo11n-seg.pt 作為基礎模型。(II) 使用學習率調度器達成快速收斂，並添加早停機制及正則化防止過擬合。(III)進行模型調參，以提升瑕疵檢測的準確率。

**(3.3) Segment-Anything Model (SAM)訓練:** (I) 使用多邊形註解生成 Bounding Box 作為模型提示，提供分割瑕疵的基礎數據。(II) 使用多邊形註解構建Ground Truth Mask，用以評估模型輸出的分割結果與實際標注相符。(III) 結合細化調整策略，進一步優化模型表現，提升分割精度，流程圖如圖1所示，其中多邊形註解、Bounding Box與Ground Truth Mask如圖2所示。

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

圖1、Segment-Anything Model (SAM)訓練流程圖

一張含有 圓形, 螢幕擷取畫面, 設計 的圖片

自動產生的描述

圖2、Segment-Anything Model (SAM)訓練流程圖

**(3.3) Segment-Anything Model (SAM)應用:** (I) 訓練YOLO Bounding box model (II) 使用測試集圖像輸入至以訓練好的YOLO Bounding box model得到Bounding box作為SAM模型的prompt。(III) 使用Bounding box作為prompt與測試圖像一同輸入至微調完成的SAM，得到最終的分割遮罩，流程圖如圖3所示。

一張含有 文字, 螢幕擷取畫面, 字型, Rectangle 的圖片

自動產生的描述

圖3、Segment-Anything Model (SAM)應用流程圖

**(3.4) 評估與比較**: (I) 利用評估指標(如精確率、召回率、mAP50與mAP50-95)及其他圖表（如圖2、3(a)、3(b)、4），對 YOLOv11、SAM Bounding box (如圖6、7、8、9、10、11、12) 與SAM的分割結果進行量化評估。(II) 比較兩種模型在不同瑕疵條件下的表現，與模型預測所耗費的時間。(如圖5、13)

一張含有 文字, 行, 圖表, 繪圖 的圖片

自動產生的描述

圖2、Yolo segmentation黃原始數據分析

一張含有 螢幕擷取畫面, 文字, Rectangle 的圖片

自動產生的描述一張含有 文字, 螢幕擷取畫面, Rectangle, 圖表 的圖片

自動產生的描述

圖3(a)、Yolo segmentation之混淆矩陣 圖3(b)、Yolo segmentation之正規化混淆矩陣

一張含有 螢幕擷取畫面 的圖片

自動產生的描述

圖4、Yolo segmentation預測與實際之對比圖

一張含有 文字, 字型, 圖形, 螢幕擷取畫面 的圖片

自動產生的描述

圖5、Yolo segmentation之預測時間

一張含有 螢幕擷取畫面, 文字, Rectangle, 正方形 的圖片

自動產生的描述

圖6、SAM Bounding box之混淆矩陣

一張含有 文字, 行, 圖表, 字型 的圖片

自動產生的描述

圖7、SAM Bounding box數據分析

一張含有 文字, 行, 繪圖, 圖表 的圖片

自動產生的描述

圖8、SAM loss curve

一張含有 文字, 螢幕擷取畫面, 圖表, 設計 的圖片

自動產生的描述

圖9、SAM 微調後基於像素的混淆矩陣

一張含有 文字, 字型, 螢幕擷取畫面, 印刷術 的圖片

自動產生的描述

圖10、SAM 微調後的評估指標

一張含有 文字, 螢幕擷取畫面, 圖表, 行 的圖片

自動產生的描述

圖11、SAM 微調後的評估指標的圖形化顯示

一張含有 螢幕擷取畫面, 圓形 的圖片

自動產生的描述

一張含有 螢幕擷取畫面, 圓形, 設計 的圖片

自動產生的描述

一張含有 螢幕擷取畫面, 圓形 的圖片

自動產生的描述

一張含有 螢幕擷取畫面, 圓形, 設計 的圖片

自動產生的描述

圖12、SAM 微調後預測與實際之對比圖



圖13、SAM預測總圖片數及總時間

1. 結果與討論

(根據所學的知識，針對期末專題，你們覺得特別困難或是有挑戰性的地方提出討論，可讓你們的報告加分)

**(4.1) 結果分析:** (I) 在 YOLOv11 的模型訓練中，模型對部分瑕疵的檢測precision達到約 90%，mAP50 為 0.97、mAP50-95 為 0.52。並且資料集預測平均耗時7.59/351 = 0.02 sec/image。然而，由於瑕疵顏色淡化和曲面反光的影響，模型在細微瑕疵的分割效果上仍有提升空間。(II) SAM 模型使用 Bounding Box 作為提示時，對瑕疵分割的pixel precision達到了約82.63%(如圖8)，mAP50為0.9225、mAP50-95為0.73。YOLO具備更快的檢測速度且在寬鬆IOU閾值下的準確率略高於SAM。SAM則是在高IOU閾值下有優勢，適用於對切割細節要求嚴格的任務。資料集預測平均耗時240/351 = 0.68 sec/image (如圖6)，耗時相當久。不過SAM的混淆矩陣是基於每個像素來進行分類(如圖9、10)，因此無法直接與YOLO之混淆矩陣比較(如圖11)。不過我們也發現到: SAM的最終結果受到 YOLO bounding box預測效能的限制，因為SAM的分割過程依賴YOLO所預測的bounding box作為其提示。

**(4.2) 主要挑戰**: (I) 標註困難：由於瑕疵顏色太淡且範圍不明確，標記的準確性受到限制，導致模型訓練數據的可靠性降低。我們嘗試了不同的標注方式，如顏色深淺、範圍大小的標準化，但仍難以完全解決問題。(II) 模型複雜性：在SAM 的訓練中，一個epoch訓練耗時約一個小時，若要調整到最佳超參數，需耗費許多時間。

**(4.3) 討論與建議**: (I) 我們建議未來可邀請貴公司的專業人士參與標註工作或使用更精細的標注工具，如 GAN (生成對抗網路)，生成更多樣化的瑕疵樣本，提升模型的學習效果。(II) 數據增強:本專題中並未在微調SAM之前做資料增強，未來可以添加有提升模型的穩健性並防止過擬合。(III) 在模型選擇上，結合 YOLO 的速度優勢與 SAM 的分割能力，開發混合模型可能是一個可行方向。(IV) 針對SAM預測時間過長的問題，未來可著重於模型輕量化與高效推理技術的研究，實現在邊緣設備（如嵌入式系統或工廠設備）上的實時部署，降低硬體成本，提升產線效率。

**(4.4) 未來展望**: 本研究為醫療零件瑕疵檢測提供了實驗基礎，但在模型性能與數據處理上仍有許多改進空間。未來若能克服標註與訓練數據的挑戰，AI 技術在實務中的應用前景將更加廣闊。

1. 結論

(主要成果及討論，兼顧定性和定量，定量就是使用數字表示，例如成功率80%)

**(5.1) 主要成果:** (I) 本研究針對醫療零件表面瑕疵的檢測，使用 YOLOv11 與 Segment-Anything Model (SAM) 進行分割能力的比較，成功開發了一套以深度學習為核心的瑕疵檢測方法。(II) YOLOv11在訓練後，達成precision約 90%；SAM 模型則透過多邊形註解與 Bounding Box 提示，pixel precision達到約 84%。YOLO的mAP50達到0.97、mAP50-95達到0.52。SAM 的mAP50達到0.9225、mAP50-95達到0.73，為醫療零件檢測提供更高效的解決方案。YOLO 與 SAM 在瑕疵分割各有優勢。

**(5.2) 成果分析:** (I) 定性分析表明，深度學習模型能有效提升醫療零件瑕疵檢測的效率，尤其是對傳統 AOI 技術難以檢測的淡色或邊界模糊瑕疵。(II) 定量分析顯示，YOLO具備更快的檢測速度且在寬鬆IOU閾值下的準確率略高於SAM。SAM則是在高IOU閾值下有優勢，適用於對切割細節要求嚴格的任務。

**(5.3) 研究貢獻:** (I) 本研究驗證了深度學習在醫療零件瑕疵檢測中的應用潛力，為產業界解決實際檢測問題提供了新思路。(II) 通過對模型性能的量化比較，提供了 YOLO 與 SAM 模型的性能參考，為未來模型選擇與優化提供依據。

**(5.4) 研究局限:** (I) 瑕疵標記的精準度與數據不均衡問題對模型的表現仍有影響，未來需進一步優化數據標註與增強技術。(II) 模型在部分特殊光線或反光條件下的表現仍需加強，以提高檢測穩定性。

**(5.5) 展望:** 未來研究可結合 YOLO 與 SAM 的特長，開發混合式模型以進一步提升檢測準確率與速度。此技術不僅在醫療零件檢測中有廣泛應用潛力，亦可拓展至其他高精度製造領域。

1. 未來研究方向

(本次專題未完成但留有遺憾的部分，未來若有時間想繼續做的內容和改善方向，以及在產業界、生活、實務上的應用)

**(6.1) 未完成部分與改進方向:** (I) 標註品質提升：由於瑕疵顏色太淡導致標註準確性不足，未來可嘗試更精細的標記工具或引入自動化標註技術，提升數據準確性。(II) 數據集多樣性：目前的數據集仍不足以覆蓋所有可能的瑕疵情況，未來可透過生成對抗網路 (GAN) 或擴增實際樣本數據，創建更豐富且平衡的數據集。(III) 實時性與邊緣設備應用:因應目前SAM在預測時間過長的問題，未來可著重於模型輕量化與高效推理技術的研究，實現在邊緣設備（如嵌入式系統或工廠設備）上的實時部署，降低硬體成本，提升產線效率。

**(6.2) 應用方向:** (I) 產業應用：深度學習技術可廣泛應用於其他高精度製造領域，如汽車零件、航空材料和電子元件的表面檢測，提升自動化檢測的效率與準確性。(II) 醫療器材發展：進一步應用於更多類型的醫療器材檢測，降低人工檢測的成本，確保產品的安全性與品質。(III) 智慧工廠：將 AI 瑕疵檢測技術整合至智慧製造系統中，形成自動化檢測流程，優化生產效率並降低次品率。

**(6.3) 生活與實務應用:** (I) 日常檢測：可將技術推廣至食品包裝瑕疵、布料瑕疵等日常產品的檢測，改善消費品的品質保障。(II) 教育與科研：提供學術界研究參考，推動更多高效瑕疵檢測技術的發展，並為相關領域的學生提供實際應用案例。

**(6.4) 未來願景:** 本研究希望在技術改進與實務應用中實現多方面突破，未來不僅著眼於解決技術挑戰，更期望促進人工智慧在製造業與生活中的廣泛應用，助力於提升產業競爭力與人類生活品質。

1. 心得

(每個人都要寫一段，包含由本課程的個人收穫、對本課程的建議、對個人未來進修和工作的興趣啟發、對助教/老師的想說的話、對系上的建議等)

(7.1) **王康鋕**  
在這次專題中，我負責資料整合與標記的工作，透過實踐我更加理解了深度學習模型對於資料標記準確性的依賴性，這是我以往未曾深入了解的環節。感謝助教在過程中的指導，讓我們在挑戰中學到了如何解決實際問題。同時，我也建議課程可以多安排實務案例的分享，幫助學生了解理論與應用的連結。未來，我希望能進一步深入研究 AI 與製造業的結合，為自己的職涯探索更多可能。

(7.2) **葉律旻**  
本專題的 SAM 模型訓練工作讓我初步掌握了深度學習模型的調參技巧與實際應用場景，這對我而言是一個全新的挑戰與收穫。我認為課程如果能夠增加更多有關模型選擇與調參的實例分析，會更有助於學生的學習。感謝教授和助教的細心指導，讓我們有機會完成這項專題。未來，我希望在研究生涯中繼續深入學習 AI 技術，並探索其在智慧製造中的潛力。

(7.3) **張富勝**  
在負責 YOLO 訓練的過程中，我深刻體會到深度學習模型需要長時間調試與測試的過程，但也因此感受到成功調教模型後的成就感。我建議課程可進一步強調數據處理的實務操作，這對學生的實際應用能力提升會有很大的幫助。特別感謝助教在我們遇到技術問題時的即時回應與支持，這對我們的研究進展非常重要。未來，我希望能將課堂所學運用在自動化檢測技術的研究上，為產業需求提供更多創新解決方案。

(7.4) **對助教、老師、系上的建議**  
本次課程的專題設計讓我們有機會從理論到實踐進行全方位的學習，這是一個非常有價值的經驗。未來，希望系上能引入更多產業合作專案，並增加與業界專家的互動機會，讓學生能更早接觸到實務問題，進而提升解決問題的能力與視野。

1. 參考文獻

(書籍、講義、paper、網頁連結含存取日期等)

[1] Redmon, J., et al. You Only Look Once: Unified, Real-Time Object Detection. CVPR, 2016.

[2] Zhang, X., et al. Application of YOLO for Real-Time Surface Defect Detection. IEEE Transactions, 2020.

[3] Kim, D., et al. Using YOLO for Automated Wafer Defect Detection. Journal of Manufacturing Science, 2022.

[4] Kirillov, A., et al. Segment Anything. arXiv, 2023.

[5] Lin, T., et al. Hybrid Object Detection and Segmentation with YOLO and SAM. IEEE Transactions, 2023.

1. 附錄

(如影片連結、程式碼、CAD圖等)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Precision | IOU | F1-score | Recall | mAP50 | mAP50-95 |
| Before Finetune | 68.36% | 0.57 | 0.61 | 0.65 | 0.51 | 0.50 |
| After Finetune | 82.63% | 0.83 | 0.89 | 0.91 | 0.9225 | 0.73 |

圖13、SAM 微調前後評估指標比較表