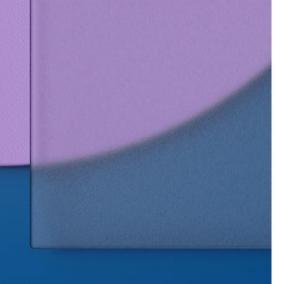
K2 AOIxAI CIMS AI 進度報告

MAMC & ADT





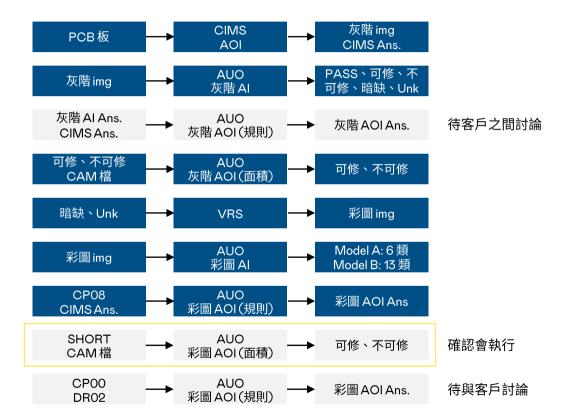
Content

- 流程改善
- 色調指標測試
- 資料清理
- 多模型實驗結果比較
- 壓縮比影響模型準度測試
- 灰階 Model 更新
- Next Step
- 彩圖 Model 更新 (請芊婷說明)

流程改善

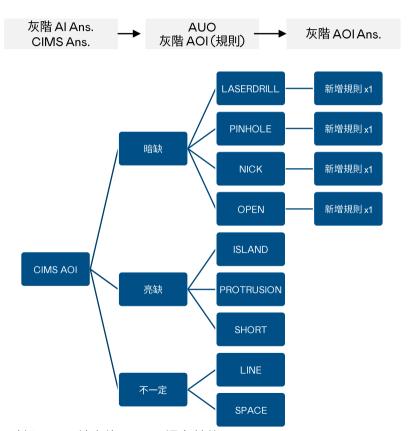


流程改善





流程改善(經雙方討論後)





CIMS Ans.	灰階 Al Ans.	資料更新
LASERDRILL		Unknown
PINHOLE		Unknown
NICK		Unknown
OPEN		Unknown
(其他暗缺)		Unknown

今日已在線上更新



流程改善





VRS 機台參數				
參數	數值			
光源	48			
曝光時間	13 秒			
R	9			
G	6			
В	13			

模型收圖參數					
規則	上限	下限			
亮度	188	130			
B 值差	43	19			
R、G差	15	3			
R、B差	49	21			
G、B差	37	15			

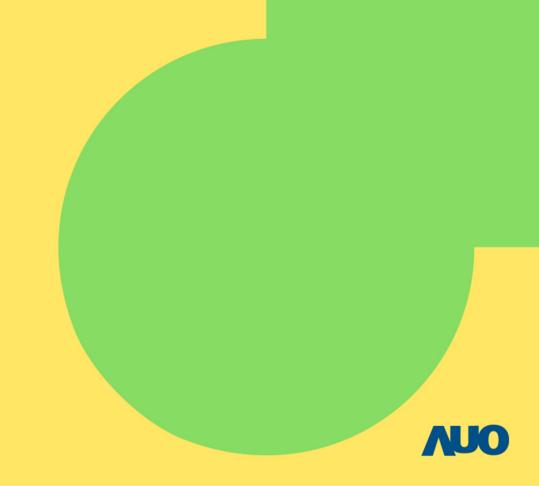
建議根據不同機台進行調整

相同機台、相同參數、不同料號,拍攝影像顏色不同不同機台、相同參數、相同料號,拍攝影像顏色不同Next Step:

- 1. 建立各機台"料號-參數對應表",加入 Auto VRS 流程自動讀取
- 2. 建立樣本至於各機台,加入Auto VRS 自動校正參數至顏色相同



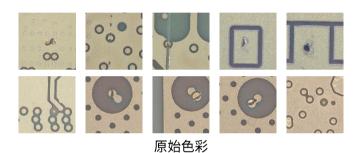
色調指標測試

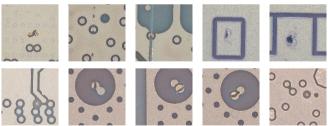


色調指標測試(針對是否可監控單一值)

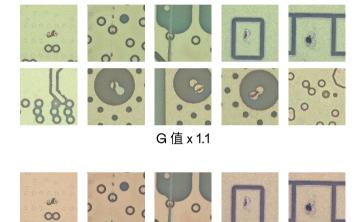
實驗條件:

• 以 CP00、CP08 為主訓練分兩類模型 (準確度 99.4%)





B值 x 1.1



R 值 x 1.1



色調指標測試(針對是否可監控單一值)

不同色調對應之準確率

針對單色 x 倍率	R	G	В
0.5 倍	80%	80%	90%
0.9 倍	100%	90%	100%
1.0 倍	100%	100%	100%
1.1 倍	90%	70%	100%
1.5 倍	30%	40%	80%

單一針對R值進行監控,在微幅降低G值時會檢測不到,顧無法僅針對R值監控。

單一針對G值進行監控,G值和R值沒有可以直接應對的關係,顧無法僅針對G值監控。

單一針對 B 值進行監控,在微幅升高 R 或 G 值時會檢測不到,顧無法僅針對 B 值監控。



資料清理



異常特徵 & 異常顏色移除

		CP00	CP03	CP06	CP08	CP09	DR02	IT03	IT08	IT09
	原始數據	2456	1538	2173	8003	8617	9382	750	4445	1641
	移除異常特徵	1631	1538	1898	3884	8617	7343	749	4445	1631
vN.1.0	移除異常色彩	2004	1069	1930	4334	5341	7623	631	3959	1263
vN.1.1	移除異常特徵+異常色彩	1338	1069	1796	3389	5341	7066	630	3959	1253
vN.1.2	修正異常特徵+移除異常色彩	2497	1069	1796	5930	5341	7066	630	3959	1253

氧化特徵在資料集中較少

PASS:1張暗缺:5張

vN.1.0 色調清理及檔案大小部分過濾 vN.1.1 PASS 中移除暗缺特徵 & 色調清理及檔案大小部分過濾 vN1.2. PASS 中移除暗缺特徵並加入回暗缺 & 色調清理及檔案大小部分過





	線上模型 (V8.5.0)	線上模型 (V2.2.0)	V230713	V230713	V230720 _exp	V9.1.0	V3.1.0	V9.1.1	V3.1.1	V3.1.1 _full_data	V3.1.2	V3.1.2 _full_data
目的	線上	模型		壓縮比影響					混淆特徵			
分類	6	9	6	9	9	6	9	6	9	9	9	9
PASS 混淆特徵	無	無	無	無	移除	無	無	移除	移除	移除	加到各類暗缺	加到各類暗缺
色彩清理	無	無	無	無	無	有	有	有	有	有	有	有
壓縮比	皆有	皆有	未壓縮	未壓縮	壓縮	皆有	皆有	皆有	皆有	皆有	皆有	皆有
資料擴增	各類取(擴增) 1300 張	各類取(擴增) 1000張	各類取(擴增) 100 張	各類取(擴增) 100 張	各類取(擴增) 840張	各類取 (擴增) 1000 張	各類取(擴增) 1000 張	各類取(擴增) 1000 張	各類取(擴增) 1000 張	各類取 (擴增) 1000張 ~ 全	各類取(擴增) 1000 張	各類取(擴增) 1000張 ~ 全
訓練結果	93.6%	90.3%	71.6%	62.2%	74.8%	88.2%	80.7%	92.0%	79.0%	80.6%	79.4%	79.3%
訓練結果 + Unknown	97.4%	94.7%	95.8%	77.5%	95.5%	96.6%	89.5%	96.7%	94.6%	92.4%	91.2%	93.9%
O717-1 盲測	56.6%	8.9%	3.8%	66.3%	62.7%	62%	67.5%	79.4%	67.5%	74.7%	67.5%	63.9%
O717-1 盲測	53.3% (當日估算為 65.0%)	23.5%	0.0%	0.0%	77.3%	50%	94.7%	47.4%	100.0% (但只有1張)	75.0%	77.4%	79.2%
0720 盲測	34.5%	14.5%	11.5%	9.4%	14.5%	30.2%	31.5%	31.6%	29.8%	34.0%	26.8%	23.0%
0720 盲測 + Unknown	38.2%	0.0%	3.8%	3.3%	0.0%	42.4%	22.2%	40.4%	16.7%	40.0%	18.2%	17.2%



	V9.1.0	V3.1.0				
目的	混淆特徵					
分類	6	9				
PASS 混淆特徵	無	無				
色彩清理	有	有				
壓縮比	皆有	皆有				
資料擴增	各類取(擴增) 1000張	各類取(擴增) 1000 張				
訓練結果	88.2%	80.7%				
訓練結果 + Unknown	96.6%	89.5%				
O717 盲測	62%	67.5%				
0717 盲測 + Unknown	50%	94.7%				
0720 盲測	30.2%	31.5%				
0720 盲測 + Unknown	42.4%	22.2%				

模型架:	I II LL:	ᅓᆓᅉ
作文字等	随下!	野角螈

因 Model B 根據不同特徵細分類, 模型可以更好的學習。

以	Model	В	推詢	侖田	邨	6	類	結果	,
耶	代 Mo	de	A I	可	有軸	交技	仔白	り效り	果

	線上模型 (V2.2.0)	V3.1.0
目的	線上模型	混淆特徵
分類	9	9
PASS 混淆特徵	無	無
色彩清理	無	有
壓縮比	皆有	皆有
資料擴增	各類取(擴增) 1000 張	各類取(擴增) 1000 張
訓練結果	90.3%	80.7%
訓練結果 + Unknown	94.7%	89.5%
0717 盲測	8.9%	67.5%
0717 盲測 +Unknown	23.5%	94.7%
0720 盲測	14.5%	31.5%
0720盲測	0.0%	22.2%

資料清理 - 色調比較實驗

+ Unknown

經過色調清理後的模型可以更好 的學習到缺陷及背景特徵,而非 色調。

色調規範及資料清理是模型訓練 重要的一環

	線上模型 (V2.2.0)	V230713
目的	線上模型	壓縮比影響
分類	9	9
PASS 混淆特徵	無	無
色彩清理	無	無
壓縮比	皆有	未壓縮
資料擴增	各類取(擴增) 1000 張	各類取(擴增) 100張
訓練結果	90.3%	62.2%
訓練結果 + Unknown	94.7%	77.5%
0717 盲測	8.9%	66.3%
0717 盲測 + Unknown	23.5%	0.0%
0720 盲測	14.5%	9.4%
0720 盲測 + Unknown	0.0%	3.3%

壓縮比比較實驗

雖未壓縮影像數量較少,但僅訓 練同盲測資料的影像可在盲測時 保持相同的條件。

故需維持訓練集與上線資料一致

	線上模型 (V2.2.0)	V230713	V230720 _exp
目的	線上模型	壓縮比	上影響
分類	9	9	9
PASS 混淆特徵	無	無	移除
色彩清理	無	無	無
壓縮比	皆有	未壓縮	壓縮
資料擴增	各類取(擴增) 1000張	各類取(擴增) 100張	各類取(擴增) 840張
訓練結果	90.3%	62.2%	74.8%
訓練結果 + Unknown	94.7%	77.5%	95.5%
O717 盲測	8.9%	66.3%	62.7%
0717 盲測 + Unknown	23.5%	0.0%	77.3%
0720 盲測	14.5%	9.4%	14.5%
0720 盲測 + Unknown	0.0%	3.3%	0.0%

壓縮比比較實驗

基於完全未壓縮的影像訓練模型,但因數量較 少,故增加基於完全有壓縮的影像訓練模型。 發現如資料集資料一致,可讓模型專注於學習 缺陷及背景特徵,而非資料差異。 故需維持訓練集資料一致

	V3.1.0	V3.1.1	V3.1.2
目的		混淆特徵	
分類	9	9	9
PASS 混淆特徵	無	移除	加到各類暗缺
色彩清理	有	有	有
壓縮比	皆有	皆有	皆有
資料擴增	各類取(擴增) 1000 張	各類取(擴增) 1000 張	各類取(擴增) 1000 張
訓練結果	80.7%	79.0%	79.4%
訓練結果 + Unknown	89.5%	94.6%	91.2%
O717 盲測	67.5%	67.5%	67.5%
0717 盲測 + Unknown	94.7%	100.0% (但只有 1張)	77.4%
0720 盲測	31.5%	29.8%	26.8%
0720 盲測 + Unknown	22.2%	16.7%	18.2%

	V3.1.1	V3.1.1 _full_data
目的	混淆	特徵
分類	9	9
PASS 混淆特徵	移除	移除
色彩清理	有	有
壓縮比	皆有	皆有
資料擴增	各類取(擴增) 1000 張	各類取(擴增) 1000張~全
訓練結果	79.0%	80.6%
訓練結果 + Unknown	94.6%	92.4%
O717 盲測	67.5%	74.7%
0717 盲測 + Unknown	100.0% (但只有 1張)	75.0%
0720 盲測	29.8%	34.0%
0720 盲測 + Unknown	16.7%	40.0%

混淆特徵比較實驗

將 PASS 類別中類似暗缺的特徵(僅因背景不同)與暗缺類別中類似PASS的特徵挑選出,並進行移除或是根據特徵分類,結果並沒有太大差異,針對混淆特徵處理沒有明確的方向。 不過直接移除可讓模型更專注學習

超大量資料實驗

基於壓縮比仍存在不同的情況, 使用超大量資料訓練可幫助模型 學習。

在壓縮比未完全清理的情況下, 先使用所有資料訓練模型

實驗總結

- 1. 以 Model B 推論時取 6 類結果取代 Model A
- 2. 收集資料時,需遵循色調規範進行資料清理
- 3. 將逐步排除壓縮資料,使訓練集與上線資料格式一致
- 4. 針對混淆特徵 (PASS 與暗缺),以目前實驗沒有太大 的直接關係,但為使模型更專注學習特徵以達到更高 的準確度,需根據各類特徵去分,次要才是實際成因
- 5. 在壓縮比未完全清理的情況下,先使用所有資料訓練 模型



壓縮比影響模型準度測試



壓縮比影響模型準度測試

實驗條件:

• 使用有壓縮影像進行模型訓練,測試集分為 A. 有壓縮、B. 無壓縮

資料集

資料集	是否壓縮	CP00	CP03	CP06	CP08	CP09	DR02	IT03	IT08	IT09	Total
原資料數	是	7,005	627	1,377	4,025	3,179	1,409	1,411	8,323	1,585	28,941
原資料數	否	254	127	313	705	294	338	122	420	222	2,795
訓練集	是	600	387*	600	600	600	600	600	600	600	5,187
測試集A	有	120	120	120	120	120	120	120	120	120	1,080
測試集 B	否	120	120	120	120	120	120	120	120	120	1,080
*擴增至6	00										

模型基本準確度(未加 Unknown Filter) 差距約 9.1% 影響程度以色調為主因 但壓縮比仍存在準確度降低的問題

測試結果 A 有壓縮

						AI模型預測					Accuracy
		CP00	CP03	CP06	CP08	CP09	DR02	IT03	IT08	ITO9	79.4%
	CP00	99	0	1	10	2	3	0	3	2	82.5% ¹²⁰
	CP03	4	104	2	1	3	0	2	0	4	86.7% ¹⁰⁰
	CP06	1	0	99	0	3	0	8	1	8	82.5%
E-4/	CP08	7	0	2	105	3	3	0	0	0	87.5%
原始 標記	CP09	4	2	2	2	95	2	2	8	3	79.2%
1木6し	DR02	2	1	0	1	3	112	1	0	0	93.3% 20
	IT03	1	3	3	0	1	4	66	0	42	55.0%
	ITO8	5	1	1	5	9	1	1	95	2	79.2%
	IT09	1	6	1	0	1	0	29	0	82	68.3%



測試結果 B 無壓縮

						AI模型預測					Accuracy
		CP00	CP03	CP06	CP08	CP09	DR02	IT03	IT08	IT09	70.3%
	CP00	92	0	4	16	2	4	0	2	0	76.7% 120.0
	CP03	4	95	4	1	4	3	5	1	3	79.2% 100.0
	CP06	4	1	101	4	0	2	3	2	3	84.2% 80.0
	CP08	6	1	1	102	1	3	0	6	0	85.0% 60.0
原始 標記	CP09	4	3	15	9	43	23	8	10	5	35.8% 40.0
作品に	DR02	7	2	0	1	7	96	5	1	1	80.0% 20.0
	IT03	0	1	2	0	2	0	74	2	39	61.7% o.o
	IT08	1	0	7	7	10	3	1	91	0	75.8%
	IT09	0	2	3	1	0	2	46	1	65	54.2%



灰階 Model 更新



AOI Gray Model 資料集 0724

No.	Date	PASS	REPAIRABLE	NONREPAIR	NG	備註
1	7/21	27	295	125	0	奕陞提供
2	7/21	12	77	19	92	奕陞提供
3	7/21	3	134	51	84	奕陞提供
4	7/21	32	71	77	487	奕陞提供
5	7/24	168	445	405	211	奕陞提供
6	7/24	137	0	0	0	奕陞提供
7	7/24	123	0	0	0	奕陞提供
8	7/24	0	2,033	112	0	騰賢提供
9	7/25	0	-1	-1	0	有孔洞的圖移除
	7/25	0	+3	-2	-1	類別修正
	Total	502	3057	783	873	



AOI Gray Model 資料集 0724

從 7/21~ 7/24 開始蒐集								
PASS REPAIRABLE NONREPAIR NG								
502	3055	786	874					

訓練標籤	PASS	REPAIRABLE	NONREPAIR	NG
Train	402	402	402	402
Valid	50	50	50	50
Test	50	50	50	50
Total	502	502	502	502

• 各類皆取 PASS 最多 502 張

• Train : Valid : Test = 8 : 1 : 1



Gray Model 0724 結果

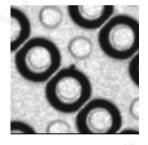
	Unknown Filter: 0.0%		Accuracy			
	Unknown Rate: 0.0%				NG	94.0%
		45	3	0	2	90.0%
原始		3	46	2	0	90.2%
		0	1	49	0	98.0%
	NG	1	2	0	46	93.9%

	Unknown Filter: 99.45%		Accuracy			
	Unknown Rate: 21.50%	PASS	REPAIRABLE	NONREPAIR	NG	98.7%
	PASS	36	0	0	1	97.3%
原始	REPAIRABLE	0	39	0	0	100.0%
原始標記	NONREPAIRABLE	0	1	41	0	97.6%
	NG	0	0	0	39	100.0%%



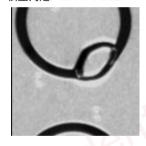
Gray Model 0724 結果

■ 原始標記: PASS ■ 模型判定: NG



NG 中找到相似特徵 疑似標記錯誤

■ 原始標記: NONREPAIR ■ 模型判定: REPAIRABLE



模型誤判



彩圖 Model 更新



ModelB v3.1.1_full BlindTest_0720

	Unknown Filter: 0%	AI模型預測							
	Unknown Rate: 0%	PASS	SHORT	CP08	CP00	CP03	DR02	Unknown	35.59%
	PASS	3	106	14	0	0	2	0	2.40%
原始	SHORT	1	73	4	0	2	6	0	84.88%
標記	CP08	0	0	4	0	0	0	0	100.00%
	CP00	0	2	7	0	0	0	0	0.00%
	CP03	0	0	0	0	0	0	0	-
	DR02	0	4	0	0	0	0	0	0.00%

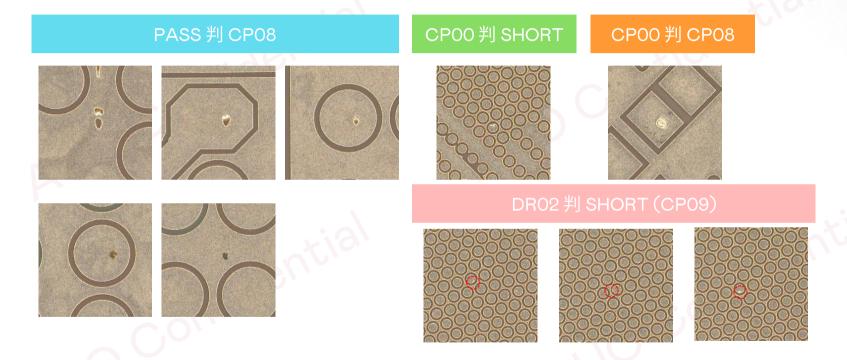
Unkno	own Filter: 99.45%	AI模型預測								
Unknown Rate: 38.16%		PASS	SHORT	CP08	CP00	CP03	DR02	Unknown	41.13%	
	PASS	0	73	5	0	0	0	47	0.00%	
原始	SHORT	0	56	0	0	0	0	30	100.00%	
標記	CP08	0	0	2	0	0	0	2	100.00%	
	CP00	0	1	1	0	0	0	7	0.00%	
	CP03	0	0	0	0	0	0	0	-	
	DR02	0	3	0	0	0	0	1	0.00%	

錯圖分析(未含Unknown)





錯圖分析(未含Unknown)





錯圖分析(含Unknown)





ModelB v3.1.1_full

BlindTest_0717

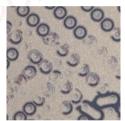
Unknown Filter: 0%		AI模型預測						Accuracy	
	Unknown Rate: 0%	PASS	SHORT	CP08	CP00	CP03	DR02	Unknown	78.48%
	PASS	0	3	0	0	1	O	0	0.00%
原始	SHORT	0	61	1	0	0	1	0	96.83%
標記	CP08	0	2	1	0	0	0	0	33.33%
	CP00	0	9	0	0	0	0	0	0.00%
	CP03	0	0	0	0	0	0	0	-
	DR02	0	0	0	0	0	0	0	-

Unknown Filter: 99.45% Unknown Rate: 84.81%		AI模型預測							Accuracy
		PASS	SHORT	CP08	CP00	CP03	DR02	Unknown	75.00%
	PASS	0	0	0	0	0	0	4	-
原始	SHORT	0	9	0	0	0	0	54	100.00%
標記	CP08	0	0	0	0	0	0	3	-
	CP00	0	3	0	0	0	0	6	0.00%
	CP03	0	0	0	0	0	0	0	-
	DR02	0	0	0	0	0	0	0	· ALIC

錯圖分析(未含Unknown)

CP00判SHORT (IT08)





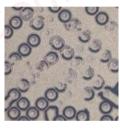


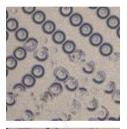


錯圖分析(含Unknown)

CP00 判 SHORT (IT08)

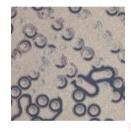


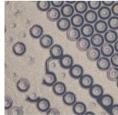


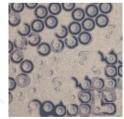










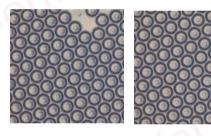




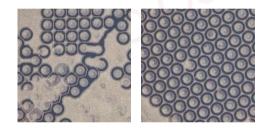


錯圖分析(含Unknown)

PASS 判 SHORT (CP09)



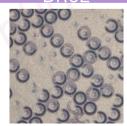
CP08 判 SHORT (IT08)



PASS 判 CP03



SHORT 判 DR02





結論

PASS/CP00/CP03/DR02 誤判 SHORT (CP09/IT08)

- 當影像出現大量小圓圖樣,模型容易判為CPO9/ITO8
- 由熱圖可見,模型因背景小圓圖樣做判斷,而非正確瑕疵 位置

PASS誤判其他類別

加強資料集清潔髒污影像,讓髒污影像降低信心分數,進 而判為Unknown











Next Step



Next Step - Al Model

Issue	項目	目標	日期與進度			
流程更新	AOI 灰階	加入灰階 AOI 規則,輔助灰階 Model 判斷結果		與客戶討論 新增規則		
	收圖	提供色調規範		提供 Sample 資料 基於 Sample 計算色調規範,取 1.5 標準差為上下限 基於訓練集計算色調規範,取 1 標準差為上下限 提供色調規範,目前 User 已較好收圖,但仍有準確 度低落問題		
2		提供 "料號 - 機台參數對照表"	07/19	機台參數差異大,只能每台機台都有一個對照表		
色調		Auto VRS 加入參數讀取功能(需 AOI 改機)	07/19	待評估		
		拍攝色調與影像色調之間如何校正(需 AOI 改機)	07/20	新問題		
	資料清理	使用色調規範整理訓練集	07/18 07/20	基於色調規範逐步限縮訓練集範圍 與線上收圖同步色調規範並進行資料清理		
	實驗	單色監控評估	07/20	無法透過單色數值監控實驗		
料號	盲測	確認料號影響比例		0717-2 固定料號實驗,但因瑕疵多為 DR02 尚不能確定是模型既有問題或料號因素提供新盲測資料,但因色調問題,尚無法定論		
	收圖	固定料號	07/18 07/19	待實驗結束後評估 7/E上線先以大料號(AO6)為主		



Next Step – Al Model

Issue	項目	目標	日期與進度			
壓縮比	資料清理	逐步增加 Auto VRS 固定壓縮比資料量	07/18	逐步限縮訓練集範圍		
	驗證	壓縮比模型測試(PASS、亮缺、暗缺): 壓縮模型驗證未壓縮 資料	07/20 07/21	資料清理、模型訓練 以有壓縮資料進行訓練,測試分為有壓縮及無壓縮 資料,前者準確度高後者 9.1%		
	驗證	基於 0717-1 盲測資料確認 v9.1模型改善效度	07/21 07/24	實際推論:準確率為59.0% (原65.0%) 實際推論:準確率為75.0%		
		基於 0720-1 盲測資料確認 v9.1模型改善效度	07/24	實際推論:準確率為 40.0% (原 11.8%)		
	特徵定義	將易誤判的 PASS 挑出 (氧化),模型訓練實驗 A. 移出 PASS – 專一 PASS 學習;B. 新增氧化類別 – 專一氧化學習		基於特徵相似度,進行特徵挑選 特徵分類結束,待資料統計(氧化在資料集中稀缺) 氧化特徵不足,無法進行		
彩色模型 	問題分析	SHORT 易判 DR02 問題分析	07/19 07/25	待盲測實驗變因確認後進行分析 基於 0720 盲測,與結果易判 SHORT 一同分析		
		結果易判 SHORT 問題分析	07/25	基於 0720 盲測,多種特徵易判為 SHORT,進行實驗找出問題		
	模型評估	彩圖模型 6 或 7 類訓練評估 (彩圖 AOI 會分可修、不可修, 是否 AI 模型也要分 – 可判度與準確度的 trade-off)		待誤判 SHORT 實驗結束進行評估,但目前應 A、B Model 合併,如需細分類則 Model B 將提高至 16類且要再請客戶協助分圖		
灰階模型	模型更新	模型資料過時,待加入資料更新(已可修不可修為主)		圖片收集 (過去訓練集為 bmp,資料集改為 jpg) 小量資料訓練 V2.1基於 502 張 (PASS 最大數量) 進行訓練		

Next Step - Al Model

是否上線?

	優點	缺黑占
上線	1. 提供基礎數據:可以提供初步的預測結果或分析數據,為後續的優化改進提供參考。2. 發現新問題:不穩定的模型反映出的問題,可引領我們發現一些原本未被察覺的新問題,對改進模型或理解問題有所幫助。	 預測不準確:不穩定的模型預測結果可能會有很大的誤差,且範圍難以掌握,導致決策失準或導向錯誤。 影響使用者體驗:不穩定的模型可能會導致產品效能不穩、產生錯誤的輸出結果,嚴重影響使用者體驗。
不上線	 避免正確性問題:當模型預測不準確,可能會產生不正確的結論或決策,結果可能導致實際執行後的結果與預期不符。 保護企業形象:如果是客戶面向的服務或產品,模型不穩定可能會影響到企業的聲譽和客戶滿意度。 提供修正與優化的時間:可以讓開發者有更多的時間來修正並優化該模型,提升其穩定性和預測精準性。 	 延遲反饋:上線可以讓模型在真實環境中運行,獲得更加真實的反饋數據,有助於模型的調整和優化。 錯失機會:就算模型不穩定,但對部分任務還是有一定的預測能力,則不上線可能會錯過一些機會,影響業務的進行。 資源浪費:對於已經開發完畢的模型,如果不上線,則相當於浪費了之前開發模型所投入的資源。

結語

像是知識語義搜尋等對即時性要求較高的情境,即使模型不穩定,上線運行也可能比完全沒有服務要好。但是在醫療、金融等領域,模型的準確性和穩定性則是至關重要的。

因此,關於模型是否上線,需要綜合考慮模型的準確性、所處領域、使用情景等諸多因素做出決定。



