



# AI趨勢論壇- Panel Discussion 案例分享

嘉衡科技  
梁百鋒

2021/09/30

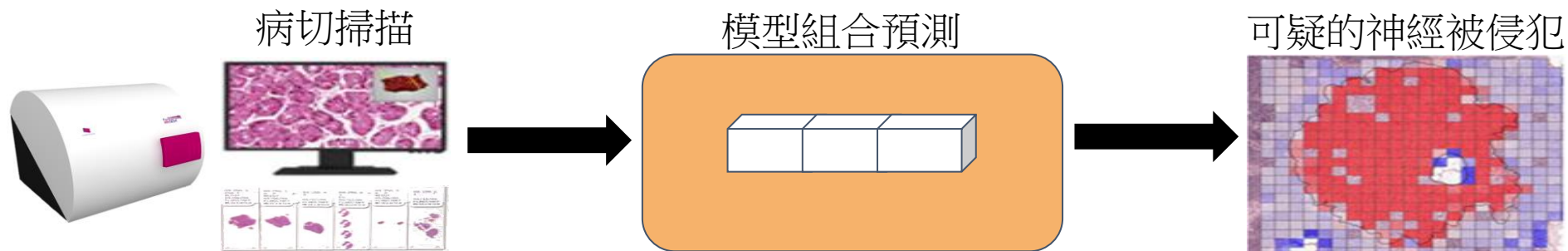
# 企業導入AI技術應用的挑戰與障礙

- 數據基礎建設不足、數據標註人力時間花費高
- AI人才供給不足、人才成本過高
- 數據不斷導入、反覆運算與細作，才能生成可用的模型
- AI模型構建和進化速度達不到生產要求
- 生成的模型與現有系統整合困難，無法直接與現有設備或產線整合

# 案例分享

乳癌不良預後因數之AI模組建置  
病理影像類之分析

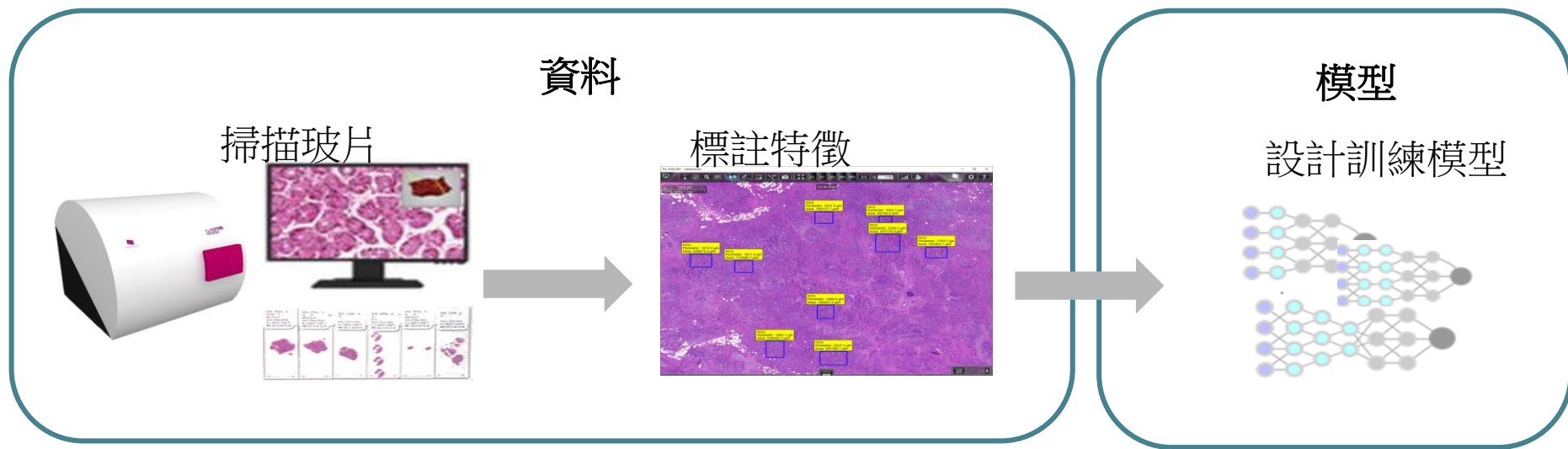
# 專案目標：由模型預測出可疑的神經被侵犯區域，加快醫師診斷速度



## 專案困難點

1. 不同掃片機的解析度不同，正確的對焦也需要人工調整。
2. 大型圖片標註困難，適當的標註方法需要協調。
3. 模型多元，他人發表的方法，不一定適用這個專案。
4. 初期資料量少，如何快速建立模型來評估。

# 進流程



資料的蒐集與標註曠日費時

# 進行中的挑戰

掃描玻片：高倍影像處理不易

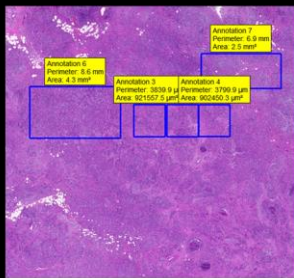


厚薄高低不同，需要對焦清晰  
不同掃描器出來的解析度必須調整到相同  
一整片影像約為：10萬 x 20萬像素 (2.5 GB)，需要大型的儲存空間

標註方法："腫瘤"邊界難以界定、"神經"稀少難以尋找

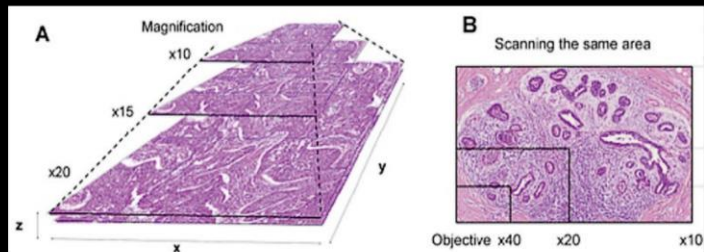


若以區域完整標註一整片  
將會花費大量時間



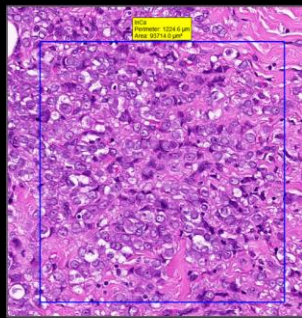
若以方框標註部分區域  
方框大小將難以抉擇  
也難標註不規則形狀的神經

標註特徵：範圍廣大、不同組織在不同倍率判斷、標註非常耗時

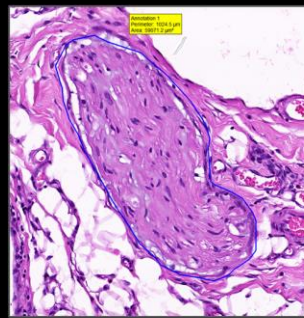


神經需在高倍下搜尋；腫瘤需在中倍下確認

標註特徵：找到折衷的標註方式，加快專案的進行



"腫瘤"於10x視野標註方框  
邊型

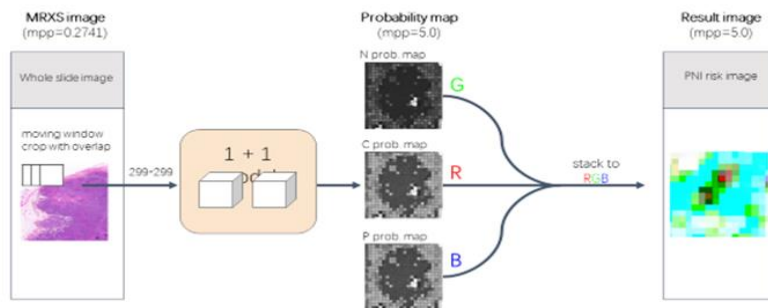


"神經"以S100輔助尋找標註多

# 執行方法

1. 模型種類確認:以圖像目標檢測(image object detection)的方式做訓練
2. 影像標註與訓練資料:針對不同目標做不同形式的標註
3. 圖像切割大小:玻片影像需要切割為小區域作為分類模型的輸入，不同的切割大小將直接影響對物件的分類正確率
4. 模型建置、訓練與調整:將自動從標註區域取樣出 299\*299 pixel 的圖像進入AutoML平台自動生成模型

Whole Slide Image (WSI) inference





# 案例分享

票據內容辨識



## 專案困難點

1. 票據品質不一，易有折痕、破損、污漬……。
2. 票據種類繁多，但特定種類數據量不足。
3. 拍攝票據影像時，不同人使用不同工具會產生不同效果，且易有歪斜、陰影、模糊……。
4. 手寫資料難以辨識。

# 進行中的挑戰

有些手寫票據是連人都難以辨識

地 址：		縣市	鄉鎮市區	路街	段	巷	弄
品名	數量	單價	金額	備			
大同m	1	18000	18000	PU=C			
				營業人			
良新 50%			90500				
銷 售 額 合 計				90500			
營 業 稅	應 稅	零稅率	免 稅				
	✓			4950			
總 計				99500			
總計新臺幣 (中文大寫) 貳 仟 佰 拾 玖 萬 玖 仟 伍 佰 拾 元							

凱威利真  
統  
台北縣  
25  
負  
TE

# 執行方法

1. 訓練資料收集、增生與標註
2. 影像預處理: 將票據影像透過預處理排除干擾的因素
3. 模型建置、訓練與調整
4. 與現有系統介接

AN 42777003 統一發票 (三聯式)

買受人: 崑崙崙管飲有限公司 中華民國 107 年 3 月 25 日

統一編號: 288704779

品名	數量	單位	單價	小計
阿瑪蘇	400	17		32200
此司	3	37		111
厚片	2	37		74
厚片	30	55		165

買受人宜用統一發票專用章

銷售額合計: 32550

營業稅: 1628

總計: 34178

崑崙崙管飲有限公司 統一發票專用章 統一編號: 08005372 屏東縣

## PHASE I

1. 憑證號碼: (前2碼英文字母 + 8碼數字) AN42777003
2. 賣方統編: (8碼數字) 28890222
3. 憑證日期: (西元格式) 2018-03-25 + 發票上未註明日期者, 應為印刷之日起算之日期。如例為「一〇七〇三二五」, 即為 2018 03 30。
4. 單位價 (NonTaxRate): 32550
5. 稅額: 營業、附加、贈與、贈與稅。
6. 稅點 (TaxRate): 1628
7. 總金額 (TotalMoney): 34178 (也可從中文大寫方式取得)。  
\* TotalMoney = TaxRate + NonTaxRate
8. 買方統編: (8碼數字) 08005372

## PHASE II

1. 買方公司名稱: 崑崙崙管飲有限公司
2. 賣方公司名稱: 崑崙崙管飲有限公司
3. 交易細項: (北碼碼, 4600, 7, NT32200); (世界, 3, 37, NT111); ...

收銀機統一發票 (三聯式 扣抵聯)

中華民國 107 年 3 月 25 日 AP 01056094

買受人為非營業人者本聯作廢

明宏藥品有限公司  
高雄市路竹區延平路520號  
NO 88704779 2

日期: 107 / 04 / 10 3  
買受人: 統一編號: 09047168 4  
賣方統編: 德昌藥師藥局

品名	數量	單位	單價	小計
大保力達B (箱x12罐)	60	箱x	761.91	45,714.6
小保力達B (箱x24罐)	20	箱x	809.52	16,190.4

受託代銷

銷售額: 61,905.5  
營業稅: 3,095.6  
總計: 65,000.7

買受人區分所得及費用固定資產  
註記欄得扣抵不得扣抵

## PHASE I

1. 憑證號碼: (前2碼英文字母 + 8碼數字) AP01056094
2. 賣方統編: 88704779
3. 憑證日期: 2018-04-10
4. 買方統編: 09047168
5. 未稅額 (銷售額, NonTaxRate): 61905
6. 稅額 (營業稅, TaxRate): 3095
7. 總金額 (Total Money): 65000  
\* Total Money = NonTaxRate + TaxRate 或  
Total Money = NonTaxRate \* 1.05 未必成立!  
\* 一張憑證裡可能有部份項目為零稅或免稅商品

## PHASE II

1. 買方公司名稱: 德昌藥師藥局
2. 賣方公司名稱: 明宏藥品有限公司
3. 交易細項: { 大保力達 B (箱 x12罐), 60, 761.91, 45714.6 }, ...

# 導入AI的挑戰

# Lessons learned

- 數據基礎建設不足、數據標註人力時間花費高
  - 有些數據必須由專業人士標註，時間與人力成本高，特別是醫療數據
  - 數據預處理關係到模型準確率，是專案成敗重要的因素
- AI人才供給不足、人才成本過高
  - 導入AI是否符合經濟效益來提升一定的競爭力
- 數據不斷導入、反覆運算與細作，才能生成可用的模型
  - 並非單一模型能解決所有問題，需考慮不同模型的組合
  - 嘗試模型的選擇與數據導入的方式以提升準確率
- AI模型構建和進化速度達不到生產要求
  - 模型需透過再訓練以提升或維持模型準確率
- 生成的模型與現有系統整合困難，無法直接與現有設備或產線整合
  - 除了現有系統及流程的調整，仍需考慮基礎建設設備與模型再訓練的環境

A wide-angle, nighttime photograph of a city, likely Taipei, viewed from an elevated position. The city is densely packed with lights from buildings and streets, creating a vibrant, glowing effect against the dark night sky. The Taipei 101 tower is a prominent feature on the left side of the image. The overall atmosphere is serene yet dynamic, capturing the energy of the city at night.

# Thank you

嘉衡科技

JUSTIN@COIGN.COM.TW