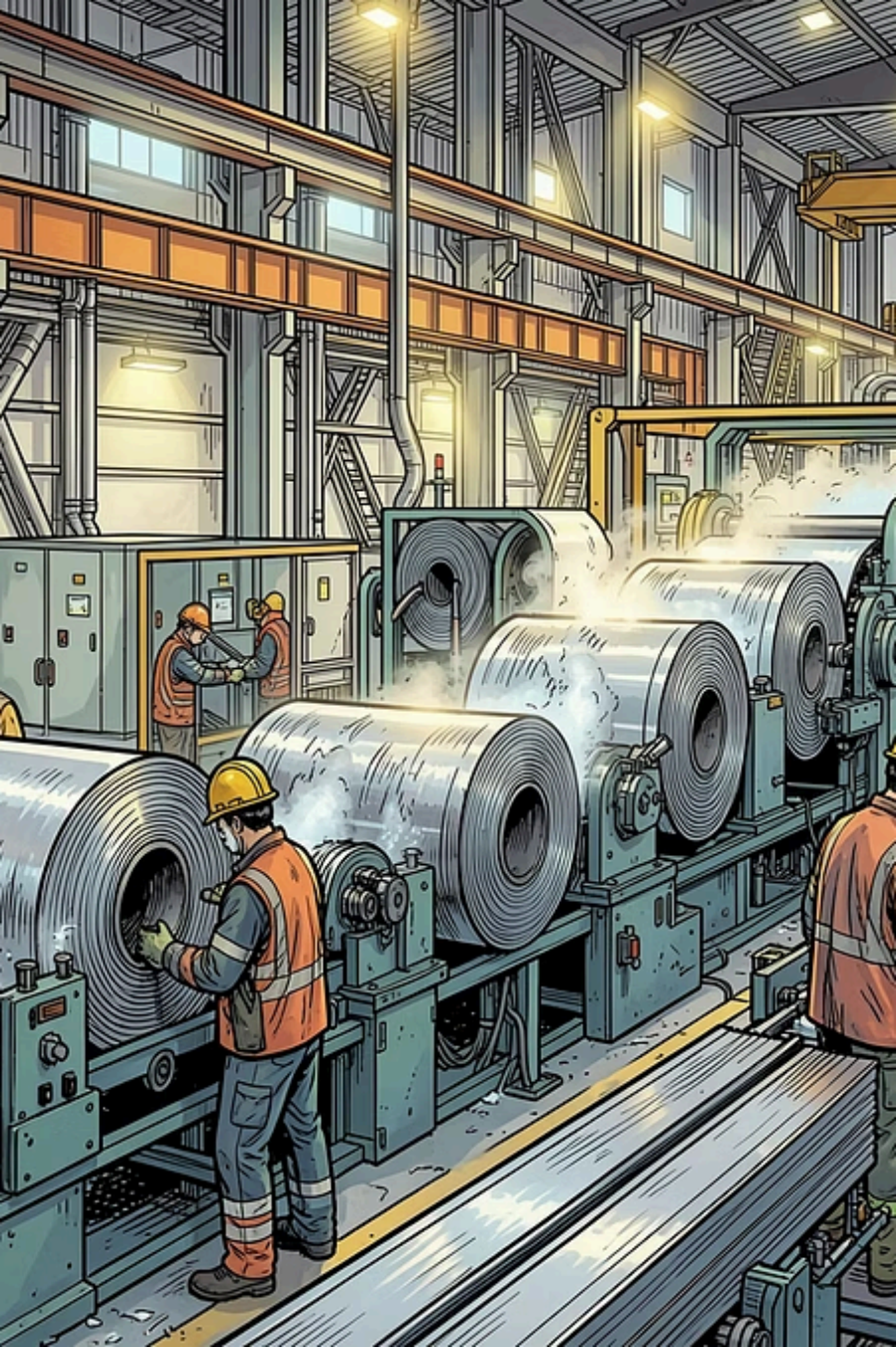




鍍鋅鋼板表面瑕疵影像辨識 深度學習架構之最佳化與實作

Optimization and Implementation of Deep Learning Architectures for
Surface Defect Recognition

侯昱廷、劉奕辰、吳岱昀 | 指導教授：洪昌鈺 教授 | 國立成功大學 資訊工程學系
(CSIE, NCKU)

工業檢測

深度學習

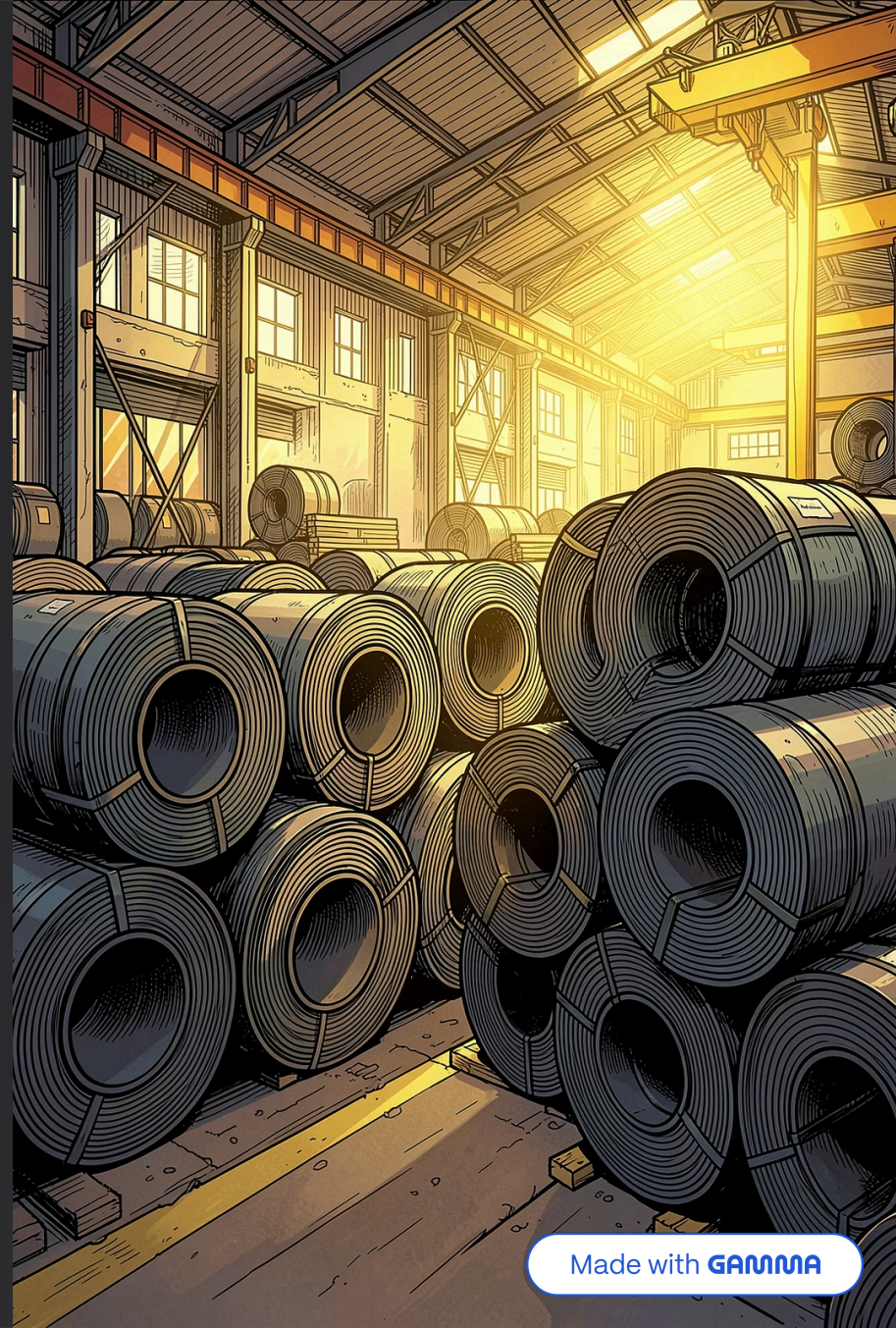
研究動機與目的

核心痛點

重大瑕疵（Sliver 與 A26）視覺特徵極度相似，辨識瓶頸嚴重

研究目標

提出改良式分類架構，克服視覺混淆，性能超越實務產線需求



工業表面檢測系統



(a) Upper surface system



(b) lower surface system

8

線掃描相機

4K SIPPIS40160

280

最高速度

mpm

0.25

辨識精度

mm (X & Y)

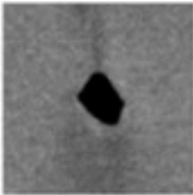
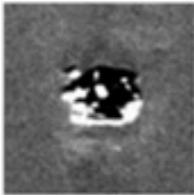
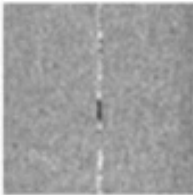
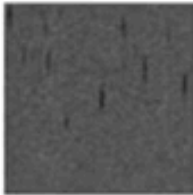
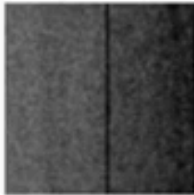
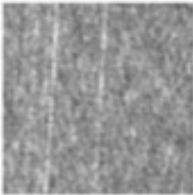
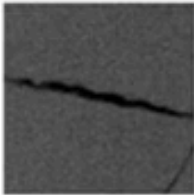
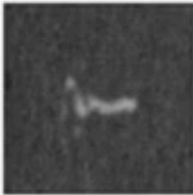
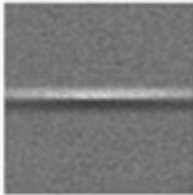
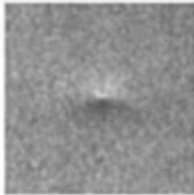
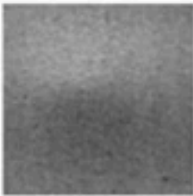
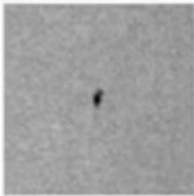
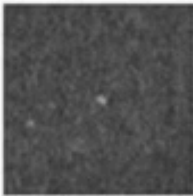
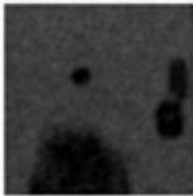
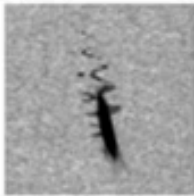
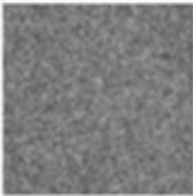
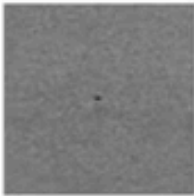
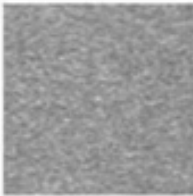
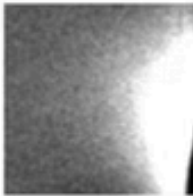
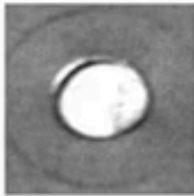
1800

檢測範圍

mm

20 類瑕疵介紹

鍍鋅鋼板瑕疵分為 3 類重大瑕疵 和 17 類輕微瑕疵。

Hole	Sliver	A26	Black_scratch _short	Black_scratch _long
				
White_Scratch	Black_linewide	White_linewide	Fold	Dent
				
Bulge	Black_point	White_point	Dirt	Insect
				
Back_ground	Noise_black	Noise_white	Edge	Weldpoint
				

重大瑕疵

Hole、Sliver、A26 —— 需要整卷重建的深度凹陷和刮痕。

輕微瑕疵

黑/白刮痕（短、長、寬）、褶皺、凹陷、隆起、黑點、白點、污垢、昆蟲、背景 —— 可修復或無需處理。

實驗方法與策略

資料預處理

29,329 張影像進行尺寸統一，先調整為 256×256，再隨機裁剪為 224×224，並完成資料清理與格式標準化。

參數、優化器與資料增量

採用 **AdamW**、**dropout 0.1** 與低學習率，降低過度擬合風險。

模型架構

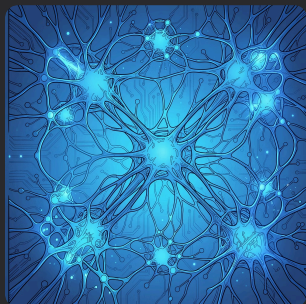
策略一：分層分割

先將視覺相似的 Sliver 與 A26 合併為一類進行初始訓練，然後再訓練專用于模型將其分類。結果：子模型表現差 —— 某些樣本即使對人眼也無法區分，較難加強。

(優化)策略二：重大 vs. 輕微

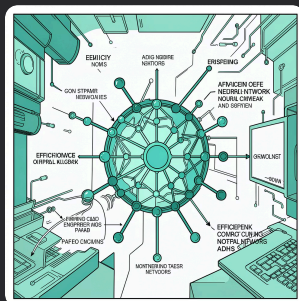
將所有瑕疵分為重大（3 類合併）和輕微（17 類合併）兩類，也測試了將 3 類重大合併 + 17 個輕微分開，共18類來訓練，更符合生產部署的實際需求。

模型選擇



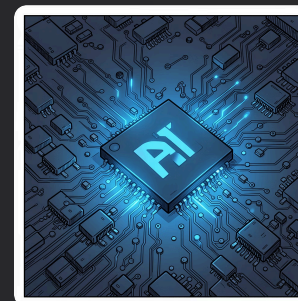
ResNet50

深度殘差學習骨幹



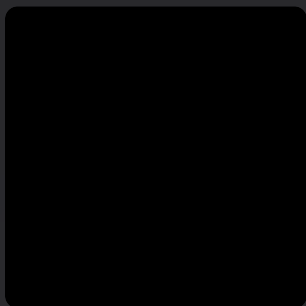
EfficientNet

最優效率的縮放 CNN



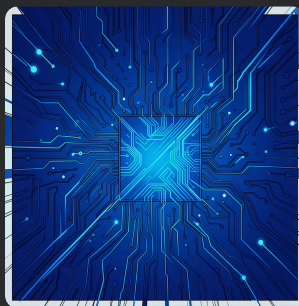
MobileNet

輕量級行動視覺架構



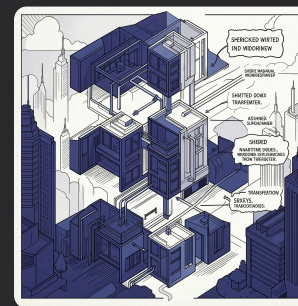
ViT

大規模視覺 Transformer



ConvNeXt

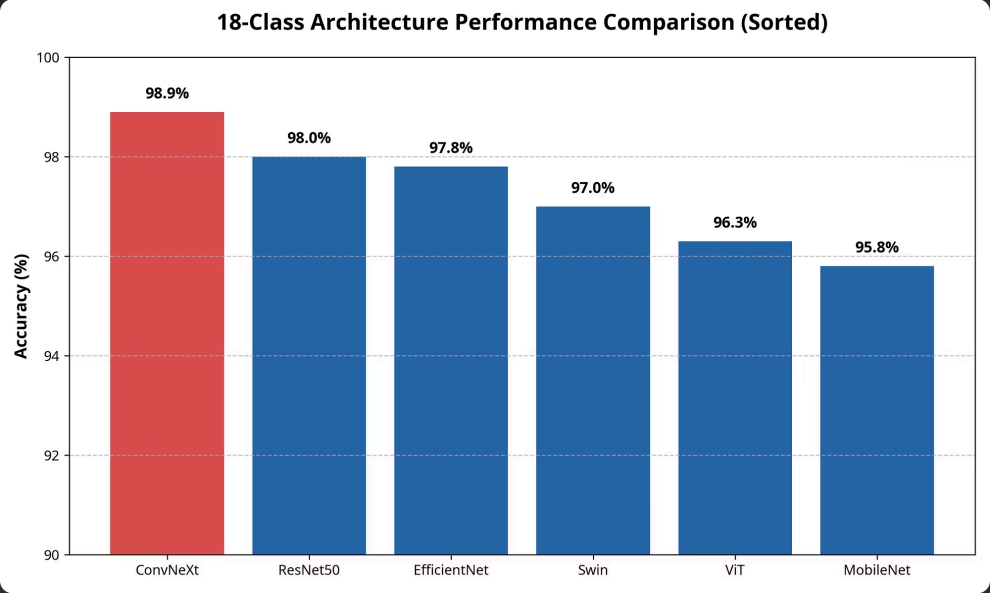
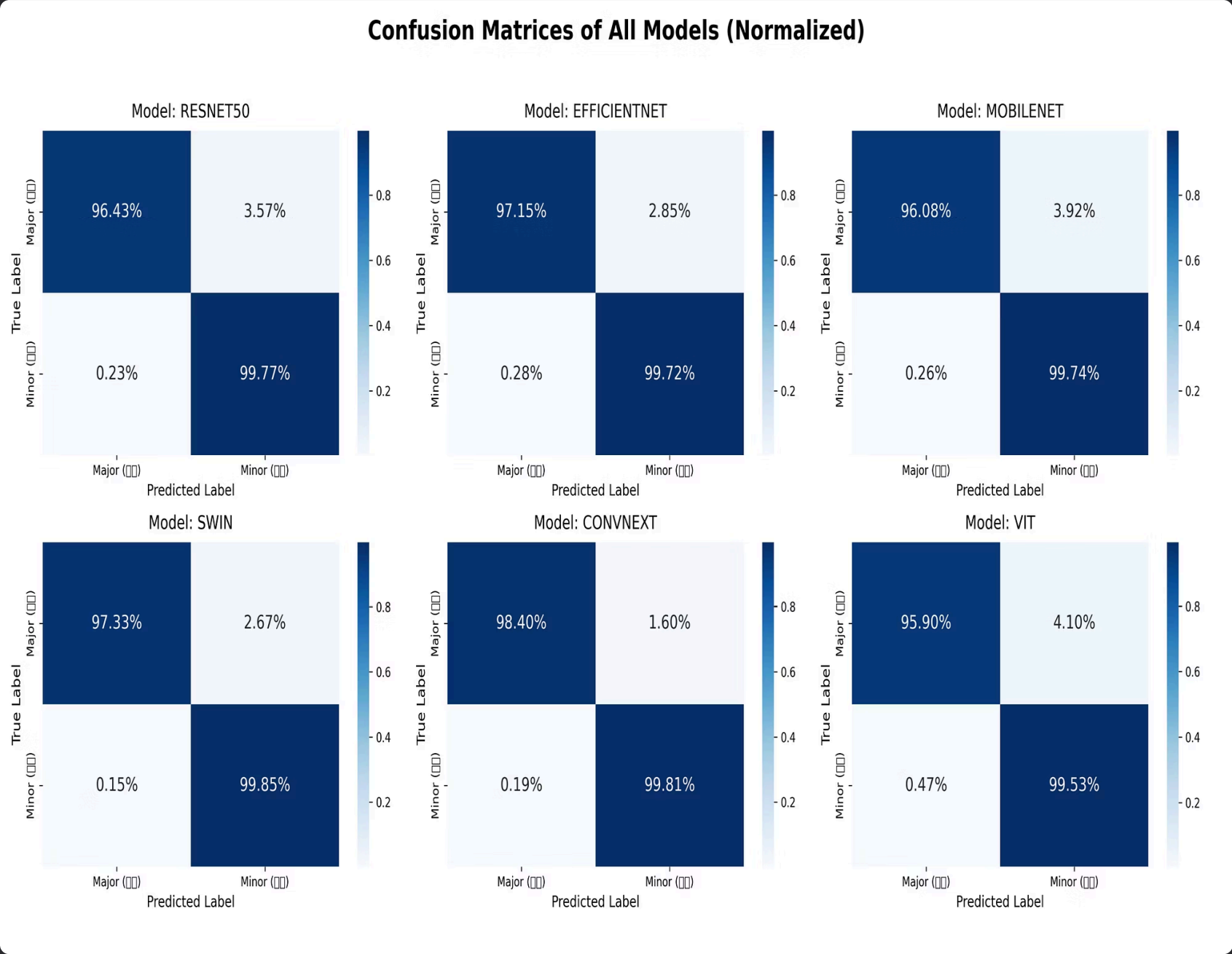
2020 年代的現代化 ConvNet



Swin Transformer

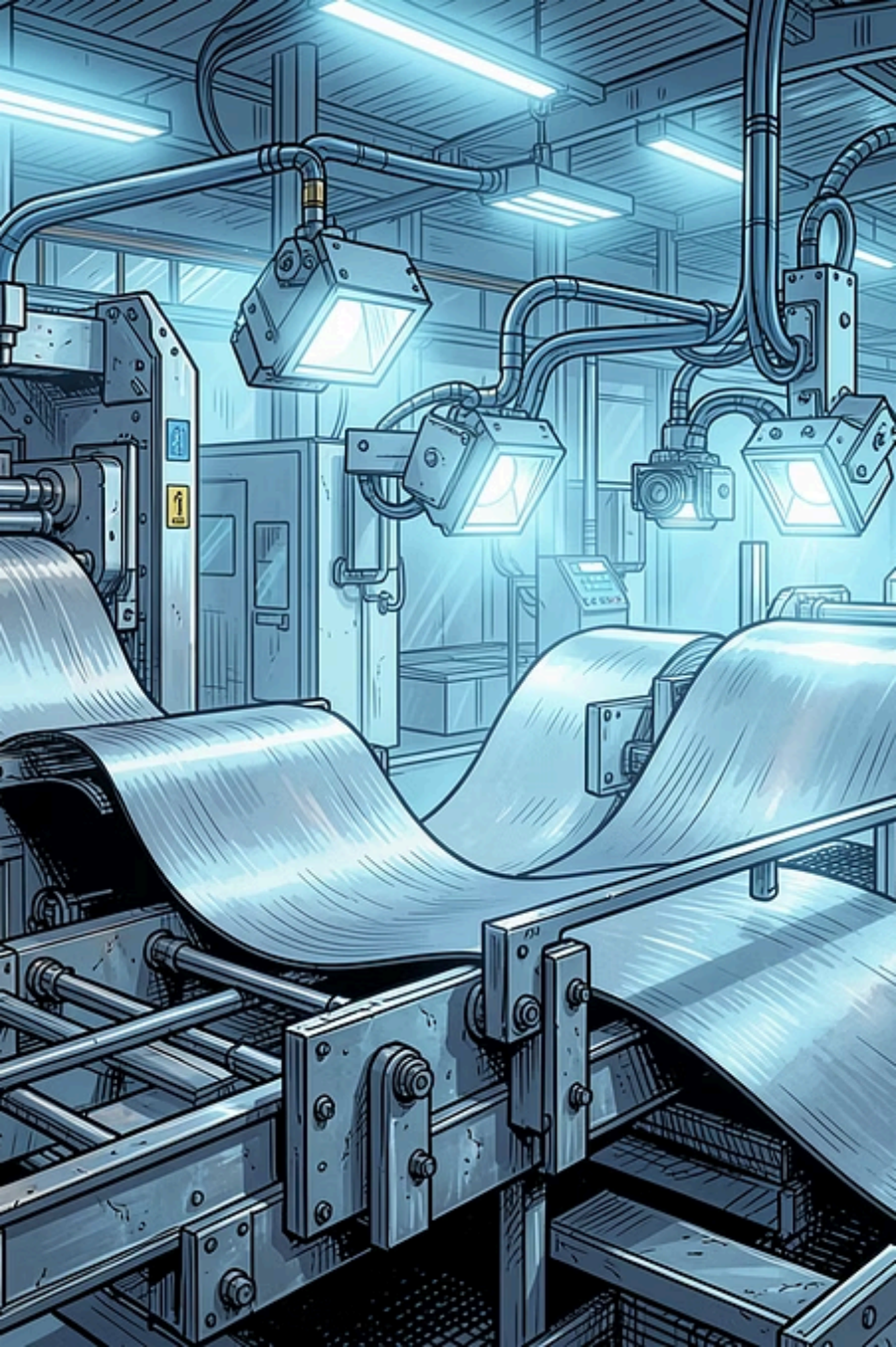
分層視覺 Transformer

實驗結果



主要發現

- **ConvNeXt** 領先 18 類分類，準確率達 **98.9%**
- **Swin Transformer** 在二元設置中最佳，達 **97.33%**
- **EfficientNet** 在二元設置中次佳，達 **97.15%**
- 所有結果遠超工業接受門檻 **85%**



結論



最佳模型：ConvNeXt
— 98.4% & 98.9%

不管是二元分類還是18類分類，表現皆最佳，遠超 85% 的工業基準。



Sliver vs. A26：視覺瓶頸

極度視覺相似性使得此區分無法通過模型調整解決 — 且不必要，因為兩者都是重大瑕疵，需要相同的處理。