

基於多分支架構結合深度監督的循環腫瘤細胞分類模型

A Circulating Tumor Cell Classification Model with Multi-Branch Architecture and Deep Supervision

研究生：資訊工程系

指導教授：資訊工程系 張傳育 特聘教授

摘要

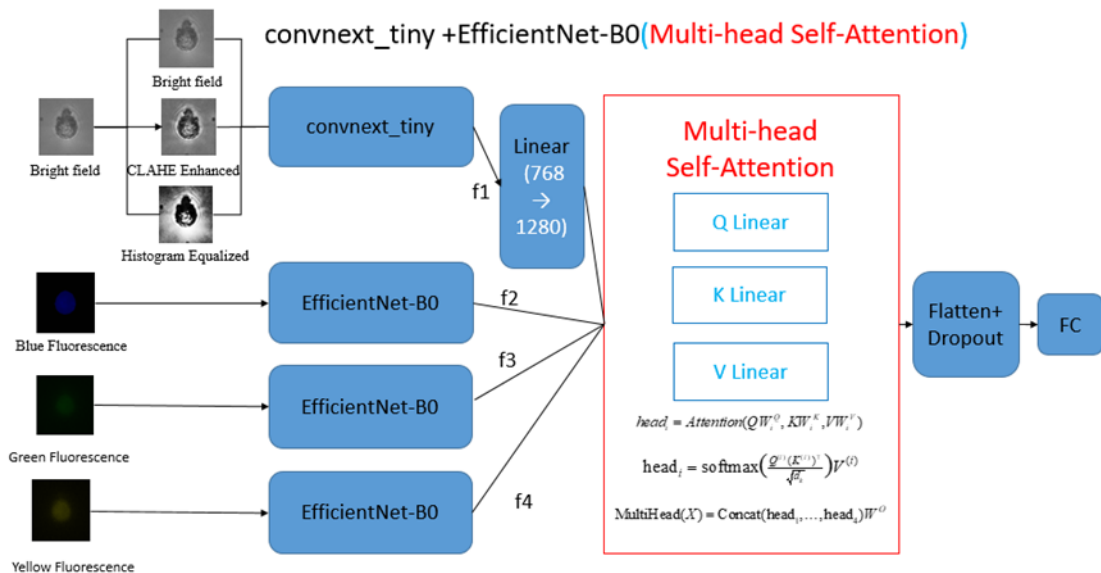
深度學習技術在影像分析領域的醫療應用有了飛快發展,但在循環腫瘤細胞(Circulating Tumor Cell)經常會面臨資料不足與不平衡的挑戰。CTC(Circulating Tumor Cell)在血液中的稀少性(每毫升僅 1 - 10 個)以及形態異質性,導致影像資料集高度不平衡,CTC 單顆和 CTC 團塊的樣本數遠少於組織碎片,影響模型對少數類別的學習效果。

此外,CTC 影像的多模態特性(亮視野影像及藍、綠、黃螢光影像)增加了影像判斷的複雜性,進一步限制了模型的臨床應用潛力。因此,本論文提出一個基於多模態卷積神經網路 Multi-Branch CNN 的分類模型,結合深度監督 Deep Supervision 分類頭與跨模態的多頭自注意力機制 Cross-Modality Multi-Head Attention,用以提升 CTC 分類的準確性與穩定性。

該模型針對四種不同細胞樣態的模態分別設計獨立特徵提取分支,並加上跨模態多頭自注意力機制動態調整模態間的特徵權重,同時利用深度監督增加少數類別的辨識能力。實驗結果顯示,本研究所提出的模型在驗證集和測試集上的表現皆優於 EfficientNet-B0, ConvNeXt, Vision Transformer (ViT), RESNET18 等方法,在準確率(Accuracy)達到 96.4%,提升了對 CTC 單顆和 CTC 團塊的識別能力。

因此,本論文所提出的方法不僅可以有效地解決資料不平衡與多模態融合的挑戰,還為臨床 CTC 分類提供了一個有效且穩定的 AI 輔助工具。

研究方法



在本研究中，我們提出了一個基於多模態卷積神經網路（Multi-Branch CNN）的分類模型，旨在對來自不同影像模態的循環腫瘤細胞進行分類。每個樣本包括四張影像，分別來自亮視野影像與三個螢光影像模態（藍色、綠色與黃色螢光）。其中亮視野影像 convnext_tiny 作為特徵提取模型三張螢光影像則是使用 EfficientNet-B0 作為特徵提取模型。

實驗結果

本研究使用的資料集由公司提供，內容為經過人工標註的循環腫瘤細胞（Circulating Tumor Cells, CTC）影像資料。資料共分為三大類別：CTC 單顆、CTC 團塊與組織碎片，其分類依據為細胞在臨床上的表現與意義，分類目的在於協助判斷癌症的嚴重程度與轉移風險。

CTC 細胞樣式	圖片
亮視野影像	
藍色螢光影像	

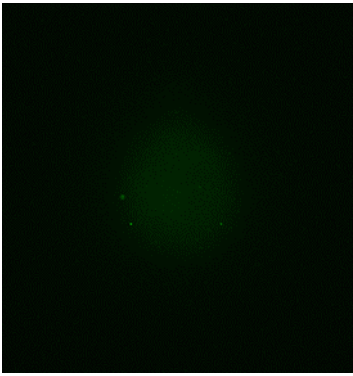
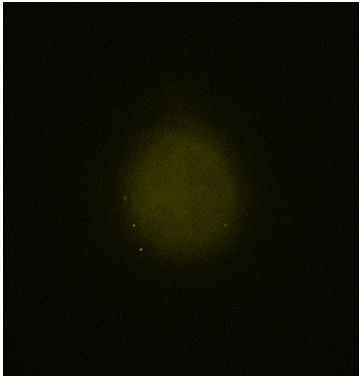
綠色螢光影像	
黃色螢光影像	

表 1 四種不同模態的特徵表示

類別	CTC 各類樣本數
CTC 單顆	298
CTC 團塊	197
組織碎片	493

表 2 資料分布

Models	Model Parameters	Test Acc	AUC	Precision	Recall	F1-score
Multi-resnet18	44,712,195	89.3% ± 0.9%	0.962 ± 0.007	0.872 ± 0.012	0.849 ± 0.015	0.856 ± 0.014
Multi-EfficientNet_B0	16,045,555	93% ± 0.9%	0.982 ± 0.003	0.925 ± 0.013	0.911 ± 0.009	0.916 ± 0.012
Multi-EfficientNet_B2	30,820,875	92.5% ± 1.1%	0.982 ± 0.005	0.927 ± 0.016	0.913 ± 0.017	0.913 ± 0.017

模型	測試準確度	AUC
ResNet18[25]	81.1% ± 2.1%	0.910 ± 0.018
ResNet50[25]	86.4% ± 3.3%	0.929 ± 0.022
ConvNext_tiny[9]	81.3% ± 0.039	0.869 ± 0.057
RegNet_y_400mf[18]	88.9% ± 2.8%	0.914 ± 0.027
vit_b_16[27]	76.8% ±4.3%	0.852 ± 0.104
densenet121[28]	93.3% ± 3.3%	0.953 ± 0.030
Squeezenet[29]	74.9% ± 6.2%	0.821 ± 0.081
mobilenet_v2[26]	75.5% ± 3.8%	0.832 ± 0.024
efficientnet_b0[8]	76.4% ± 1.5%	0.878 ± 0.016
ours	95.6%±0.8%	0.987± 0.002

結論

本研究旨在提出一個有效的深度學習模型，用於循環腫瘤細胞（CTC）的多模態分類，並探索如何結合先進的影像處理技術與深度學習架構以提升分類性能。透過多模態影像資料的整合，我們提出一個多分支卷積神經網路架構，其中針對不同模態分別使用 EfficientNet-B0（處理藍、綠、黃三種螢光影像）與 ConvNeXt-Tiny（處理亮視野影像）作為特徵提取主幹，並結合深度監督（Deep Supervision）與跨模態多頭注意力機制（Cross-Modal Multi-Head Attention），以進行語義融合與強化多層特徵學習。以下為本研究之主要結論：

本研究的核心貢獻 在於提出一個能有效處理 CTC 影像分類任務的多模態深度學習模型，並結合以下技術達成顯著成效：

- (1)多分支特徵提取模型:本論文使用多分支特徵提取模型,對四種不同模態的影像分別使用不同的特徵提取模型提取特徵
- (2)深度監督（Deep Supervision）：本研究設計並比較了多組不同階段的深度監督配置，，能有效提升對少數類別（如單顆 CTC 與 CTC 團塊）的辨識能力，並改善模型的泛化性與穩定性。
- (3)跨模態注意力機制（Cross-Modal Attention）：透過引入跨模態多頭注意力模組，模型能根據各模態之間的語義關聯動態調整權重，成功強化了各分支特徵的互補融合，進一步提升分類準確率與辨識能力。