

考卷手寫字英文數字識別



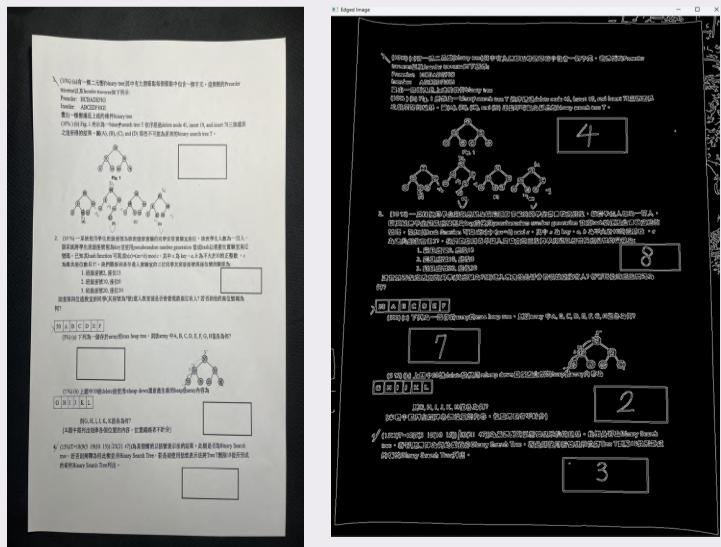
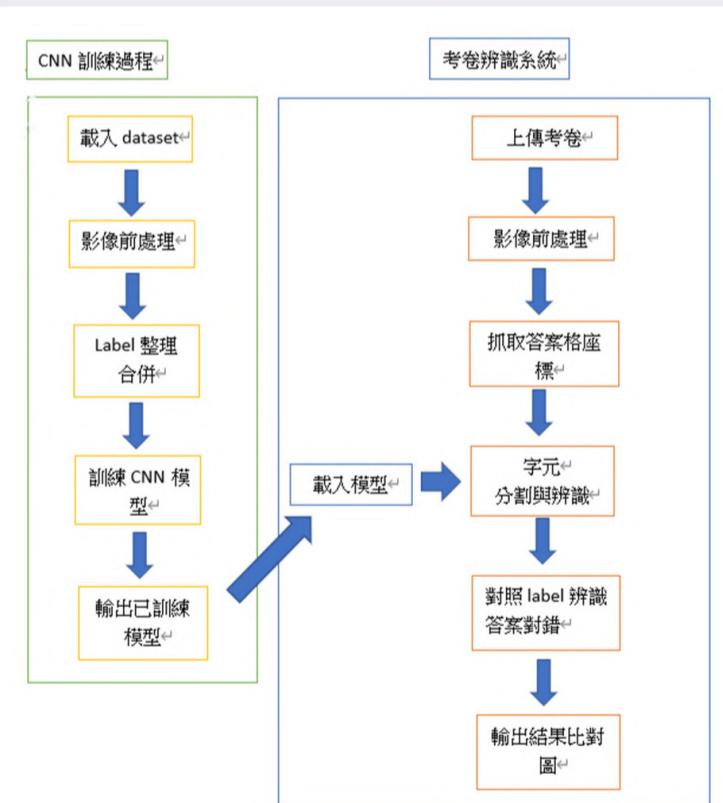
指導教授：余松年 教授

專題生：鄭敬寰 王晟亦

摘要

本專題旨在開發一套整合深度學習與影像處理技術的手寫文字辨識與比對系統，應用於學術考卷的自動閱卷任務。系統以公開 EMNIST 字元資料集為基礎，使用 EfficientNetB0、CNN 及 LeNet 三種卷積神經網路模型進行訓練與融合，完成考卷影像的幾何校正、答案格偵測、文字切割與多模型預測，最後比對學生作答與標準答案，直觀地以紅綠色視覺化標示正誤。整體系統具備高度準確率與自動化辨識能力，適用於教育領域之數位閱卷應用。

架構介紹

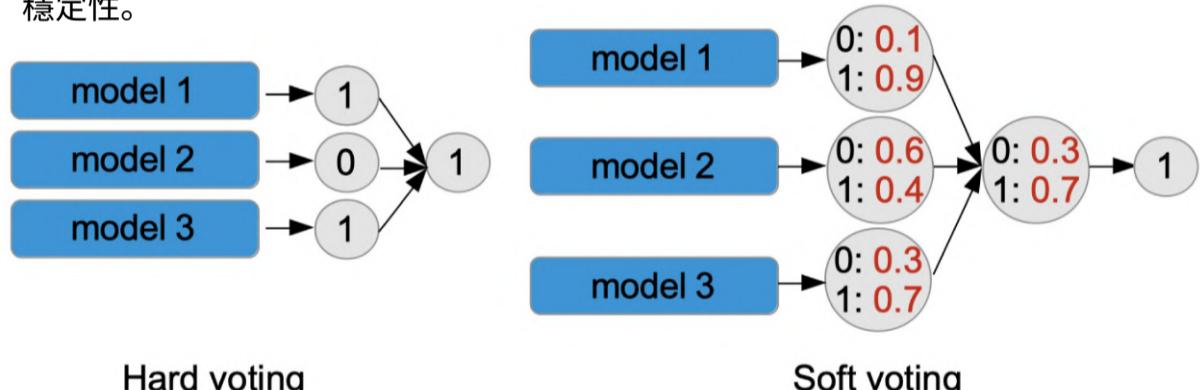


實驗方法

本研究以公開資料集 EMNIST (Extended MNIST) 為基礎，針對 36 類手寫英文字母與數字進行訓練與測試。首先，從資料集中擷取影像樣本，進行前處理 (Preprocessing)：將原始 28×28 灰階影像進行旋轉修正、尺寸調整 (至 112×112 或 28×28)、灰階轉 RGB、正規化至 0-1，以符合各模型輸入需求。

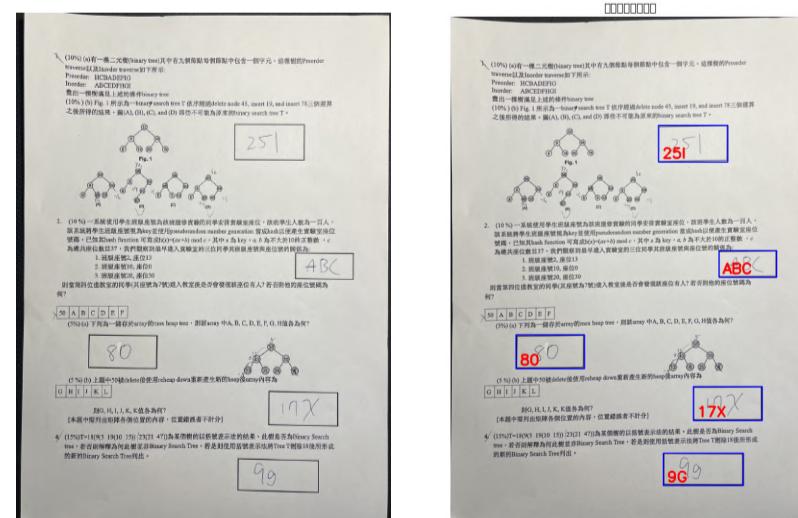
接著，建立四種卷積神經網路模型：AlexNet、EfficientNetB0、LeNet 以及專用 CNN。每個模型皆使用相同的訓練設定：損失函數採用 Categorical Crossentropy，batch 大小 16，訓練 10 個 Epoch。各模型皆以 EMNIST 影像為輸入，學習筆劃、邊緣、形狀等特徵，輸出 36 類 Softmax 機率分布。

為進一步提升辨識率與穩定性，本系統導入多專家模型融合 (Ensemble Learning) 機制。每張影像會同時輸入至多個模型，各自產生預測機率，之後以 Soft Voting 平均法 對機率分布進行加權整合，取最高信心度作為最終分類結果。此方法可有效結合不同模型的判斷特性，減少單一模型誤判對整體準確率的影響，並提升在筆跡不規則或模糊情況下的辨識穩定性。



在應用階段，系統針對拍攝或掃描之考卷影像進行多階段處理：先以 高斯模糊 (Gaussian Blur) 去除雜訊，再利用 Canny 邊緣偵測 找出考卷外框，經透視校正 (Perspective Transform) 將影像拉正。校正後偵測並擷取所有答案格，若單格中包含多字元，則進行輪廓偵測與切割，將每個字元獨立置中並統一尺寸。最後，透過多模型融合進行預測，並與標準答案比對，自動標示正確（綠色）與錯誤（紅色）結果，實現高準確率之自動化閱卷流程。

結果圖



結論

本研究以 EMNIST 手寫字元資料集 為基礎，針對傳統紙本考卷批改作業中耗時費力與容易誤判的問題，提出一套結合深度學習模型訓練與電腦視覺影像處理技術的自動化辨識系統。整體流程涵蓋：影像預處理與透視校正、答案區塊定位、多字元分割、多模型融合辨識，以及與標準答案比對的視覺化呈現。整體還算滿意但仍具有能加強的部分。