

1. 期末專題題目選擇：情緒辨識

2. 資料集詳細介紹：

● FERPLUS 2013：

包含約 35,887 張人臉圖像，每張圖像皆有情緒標註，情緒分別有 angry、disgust、fear、happy、neutral、sad、surprise，七種。影像尺寸為 48x48 的灰階影像。

資料集來源：<https://www.kaggle.com/datasets/msambare/fer2013/data>

● RAF-DB：

包含約 30,000 張臉部圖片，每張圖有情緒標註，情緒分別有 angry、disgust、fear、happy、neutral、sad、surprise，七種。影像為彩色影像。

資料集來源：<https://www.kaggle.com/datasets/shuvoalok/raf-db-dataset>

● JAFFE：

包含 213 張日本女性臉部影像，展示七種基本情緒 angry、disgust、fear、happy、neutral、sad、surprise，影像為，灰階，標準化構圖。

資料集來源：<https://zenodo.org/records/14974867>

● CK+：

包含 593 個影像序列，有七種情緒，angry、disgust、fear、happy、contempt、sad、surprise，影像為，640x480 的灰階影像

資料集來源：<https://www.kaggle.com/datasets/shuvoalok/ck-dataset>

3. 研究方法：

使用模型：VGG19、ResNet50、MobileNetV2。

訓練流程：先將訓練及切分成訓練集和驗證集再開始訓練，訓練初期凍結大部分預訓練層，隨著訓練進行解凍更多層並降低學習率，避免過擬合現象，讓模型可以泛化，訓練結果會由圖表呈現，並且測試集會進行 Grad-CAM 熱圖分析，再畫出混淆矩陣，視覺化呈現模型訓練結果，每個模型會先從 FERPLUS 2013 開始訓練，再遷移到 RAF-DB，再到 JAFFE，最後遷移學習到 CK+。

4. 預期結果：

對於每個模型的預測都是，再 FERPLUS 2013 訓練時準確率約為 55%~65%，RAF-DB 和 JAFFE 訓練準確應該會提升到 75%~85% 左右，最後 CK+ 訓練準確率應該能達到 95% 以上。並且因為 ResNet50 的特徵學習能力是三個模型中最強，且泛化能力也比其他兩個強，ResNet50 整體的預測表現應該會最好。而 VGG19 因為收斂速度慢，且容易過擬合，表現應該最差，而 MobileNetV2，表現應該介於兩者之間。