Deep Learning for Climate Action: Natural Disaster Image Classification

Introduction

使用卷積神經網路（CNN）進行自然災害影像分類，對應聯合國永續發展目標（SDGs）中的「氣候行動」目標。研究選用多種類型的自然災害影像資料進行訓練與測試，並以深度學習模型自動辨識災害類別，提升災情回報效率與災害應變能力，期望於實務應用中發揮預警與決策輔助的價值。

Background

近年極端氣候日益加劇，導致自然災害發生頻率與強度不斷上升，對全球人民生命與財產造成重大威脅。如何有效提升災害應變速度與精準度，是智慧防災領域的核心課題。本研究結合人工智慧技術，透過圖像分類模型實作自然災害自動辨識系統，對應 SDG 目標中的「氣候行動」（Climate Action），以促進災害資訊數位化並提升智慧應變效能。

Methodology

資料集來源：本研究使用自行透過爬蟲蒐集與整理之自然災害分類資料集，資料涵蓋乾旱、下雨、海嘯、沙塵暴、龍捲風災等災害類型，依據 8:1:1 比例切分為訓練集、驗證集與測試集。

模型訓練方式與評估指標：
採用三種深度學習模型進行訓練與比較，分別為ResNet34（表現最穩定）、 VGG16（在 Tsunami 上分類表現落差大Ef)，與ficientNetB0(有些類別 AUC 高但分類準確率偏低）。

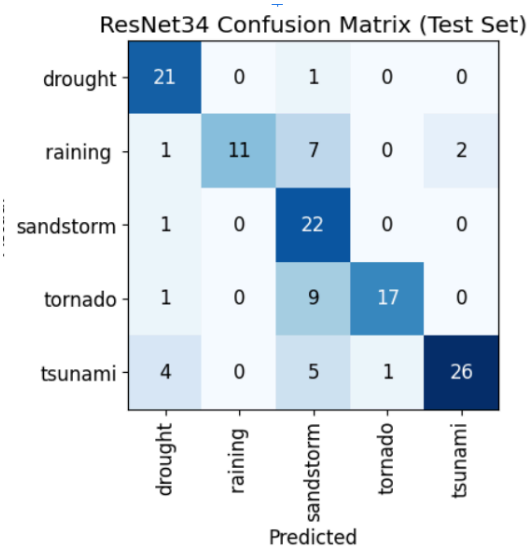
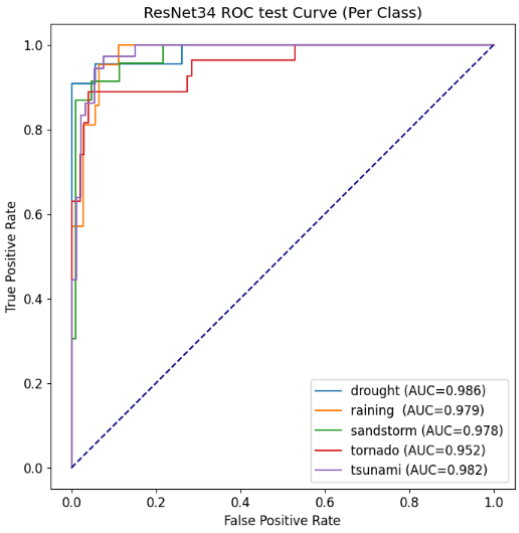
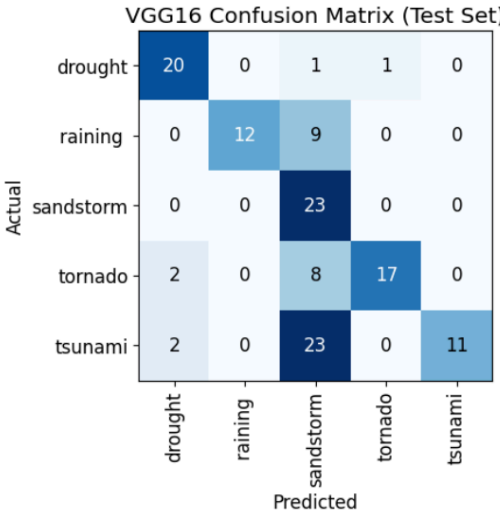
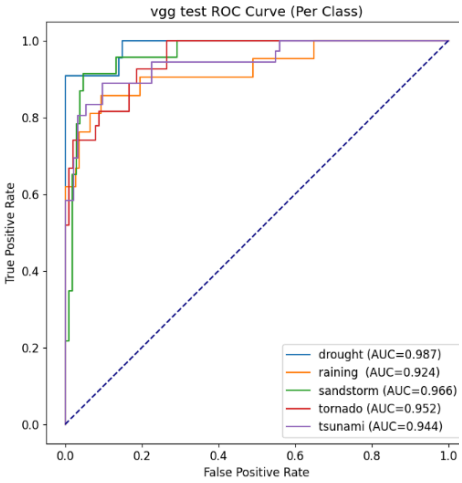
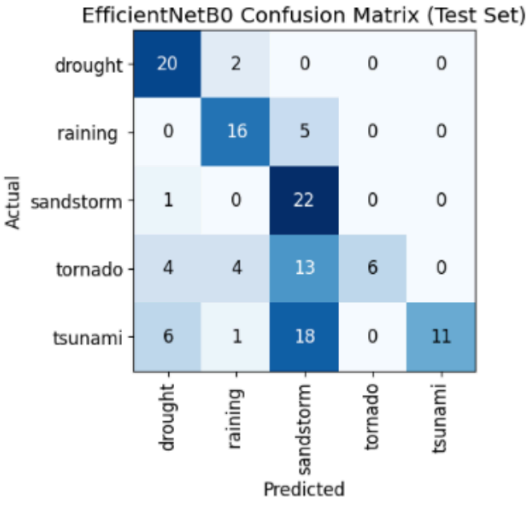
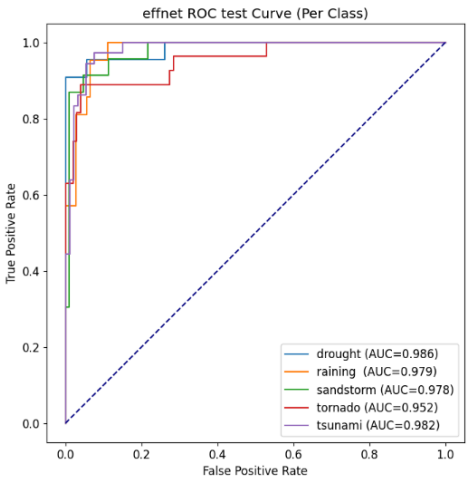
使用 FastAI 架構與 transfer learning 技術，從 .pkl 模型進行 fine-tune。並透過Confusion Matrix（混淆矩陣），Accuracy、Precision、Recall、F1-score（Macro / Micro）ROC Curve（每類別 AUC）還有Gradcam等指標，量化各模型在災害分類任務上的表現，找出最適合應用於實際情境的分類架構。

Result

ResNet34 (final choosen)

在準確率（Accuracy）、召回率（Recall）、F1-score 與 AUC 等指標中皆表現優異，整體分類穩定性最佳，適合用於廣泛的災害影像辨識任務。VGG16 雖然整體準確率略低，但在 Precision（精確率）上維持良好，顯示其對於部分災害類型仍具辨識優勢。 EfficientNetB0 在本次資料上表現相對較弱，雖然 AUC 值尚可，惟整體分類能力不足，F1-score 僅達 0.50。

Analysis

model	Confusion Matrix	ROC
ResNet		
VGG		
EFFNet		

Resnet	Output
Resnet	

epoch	train_loss	valid_loss	accuracy	error_rate	time
0	0.576374	0.209653	0.932773	0.067227	00:45
1	0.631889	0.725434	0.840336	0.159664	00:23
2	0.633412	0.403531	0.890756	0.109244	00:21
3	0.560248	0.233083	0.957983	0.042017	00:24
4	0.463282	0.150771	0.957983	0.042017	00:23
5	0.380964	0.129039	0.957983	0.042017	00:21

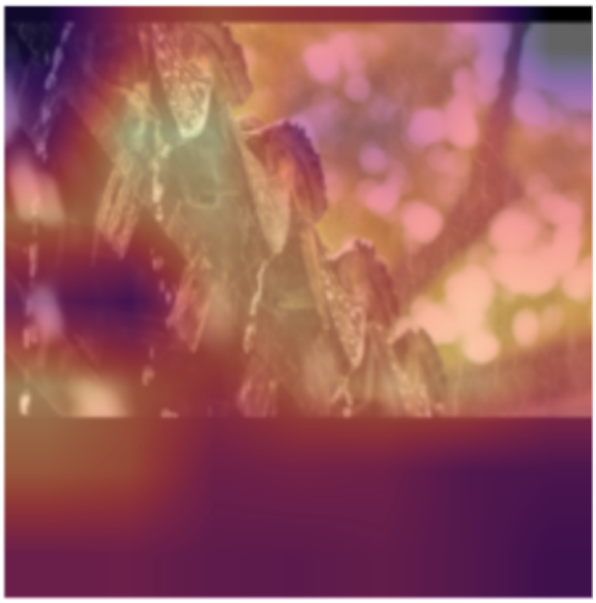
VGG

epoch	train_loss	valid_loss	accuracy	error_rate	time
0	1.952788	0.594819	0.815126	0.184874	00:27
1	1.252023	0.467647	0.857143	0.142857	00:25
2	0.844081	0.269365	0.932773	0.067227	00:25
3	0.670646	0.208916	0.957983	0.042017	00:23
4	0.491985	0.149094	0.957983	0.042017	00:25
5	0.431932	0.135193	0.966387	0.033613	00:23

EFFnet

epoch	train_loss	valid_loss	accuracy	error_rate	time
0	2.289543	0.863984	0.647059	0.352941	00:22
1	1.445138	0.366771	0.899160	0.100840	00:23
2	0.942167	0.219021	0.941176	0.058824	00:23
3	0.713237	0.180706	0.949580	0.050420	00:21
4	0.538148	0.125111	0.974790	0.025210	00:24
5	0.398370	0.118293	0.983193	0.016807	00:22

Hugging face ORcode



ResNet34 在 Tsunami 分類上的整體穩定性與泛化能力最優，而 Tornado 為所有模型中最難預測的類別

Model	Accuracy	Precision Micro	Precision Macro	Recall Micro	Recall Macro	F1 Micro	F1 Macro	AUC Weighted	AUC Macro
ResNet34	0.75	0.75	0.82	0.75	0.76	0.75	0.75	0.98	0.98
VGG16	0.64	0.64	0.83	0.64	0.68	0.64	0.67	0.95	0.95
EfficientNetB0	0.58	0.58	0.74	0.58	0.63	0.58	0.57	0.89	0.89

Conclusion

本研究透過 ResNet34、VGG16 與 EfficientNetB0 三種 CNN 架構，針對五類自然災害影像進行分類與比較，最終在測試集上進行準確率、F1-score、AUC 以及混淆矩陣與 ROC 曲線的分析，並得到以下觀察與結論：

模型特色&結果關聯		
ResNet34： 特徵學習穩定、分類表現最均衡 模型特性： 使用殘差結構（Residual Connections），能避免深層網路的梯度消失問題。 深層穩定、能學習結構性特徵，容易擷取影像中具區辨性的深層特徵。 訓練穩定，即便類別樣本略不均，也能維持不錯表現。 對照結果：Raining / Tsunami 較準，整體平衡 測試集準確率最高（0.75），Macro F1 與 AUC 都達 0.75 / 0.98。特別對「Raining」與「Tsunami」分類最準，表示模型能正確抓住這兩類的關鍵特徵（例如雲層、水流、地形）。 誤分類多集中於 Tornado → Sandstorm 推測原因: 這兩類影像在輪廓與色調上過於接近（如灰黃塵霧、模糊遠景）。	VGG16： 基礎卷積網路，結構簡單但容易過度擬合 模型特性： 結構規則（重複 3x3 卷積 + max pooling），但缺乏 shortcut 結構輔助，結構簡單、對模糊圖敏感。 容易學習到影像整體形狀，但在複雜特徵區分力不足。 對於某些視覺風格明顯的類別有優勢（如 Drought 或 Sandstorm） 對照結果：Tsunami 誤判嚴重，整體落差大 Tsunami 分類效果最差，出現 23 張誤分類為 Sandstorm，是所有模型中最大誤分類案例。雖 AUC 達 0.95，但 F1 Macro 僅 0.67，代表預測信心高，但真正命中率偏低。 推測原因： VGG16 無法分辨 Tsunami 特徵與背景海霧、沙塵的差異，模型對細節敏感度不夠。	EfficientNetB0： 參數最少、效能精簡，對資料敏感 模型特性： 強調網路縮放（depth/width/resolution scaling）以減少參數。 訓練快速，但精簡高效、易受背景干擾影響，所以在特徵不明顯或數據量較小時，容易表現不穩。 適合資料結構明確、類別邊界清楚的任務。 對照結果：Tornado 與 Sandstorm 混淆明顯 準確率最低（0.53），F1 Macro 僅 0.57，但 AUC 保持 0.89。Tsunami 與 Tornado 極易被誤分為 Sandstorm（共31張），顯示模型雖能對高信號做反應，但無法細緻區分背景。 推測原因： 特徵分布不夠密集的類別（如 Tornado 含雲層、沙霧、建築混合），導致模型縮放策略失效。

哪些類別最容易被錯誤分類？—沙塵暴 Sandstorm

分析三組混淆矩陣結果可發現，「沙塵暴 Sandstorm」是最容易被錯誤分類、或將其他類別誤分類為 Sandstorm 的類別。具體情形包括：

1.在 VGG16 模型中有多達 23 張 Tsunami 圖像被錯判為 Sandstorm，為整體誤判最多的單一案例。

2.在 EfficientNetB0 中，也有 18 張 Tsunami 與 13 張 Tornado 被誤判為 Sandstorm。

3,就連 ResNet34，也出現 Tsunami、Tornado、Raining 被誤判為 Sandstorm 的狀況（合計 21 張）。

顯示 Sandstorm 本身在影像上的辨識難度較高，同時也容易與其他類別混淆。

是哪些特徵導致模型難以分辨？

1.視覺特徵重疊性高：
Sandstorm 的典型視覺特徵為 灰黃、模糊、不清楚背景的構圖，這些特徵與海嘯（Tsunami）的遠景霧氣、龍捲風（Tornado）的塵煙旋流以及降雨（Raining）的模糊灰階區域具有高重疊性。

2.邊界與背景不明確：
Sandstorm 圖像常缺乏明確輪廓（例如建物、地標），使得模型難以抓取區分邊界，導致特徵向量之間高度接近。

3.資料集中樣本分布不平衡或類別內變異過高：

Sandstorm 類別內部可能包含來自不同地區的場景（城市沙塵、沙漠、道路起塵等），但模型學習時仍視為單一類別，造成學習模糊化。

如何改進數據或模型來提升效能？

1.Data Augmentation：

使用如亮度調整、對比度增強、視角旋轉、模糊模擬 等方法，來模擬不同光照、視野與拍攝條件，幫助模型學習更多元的沙塵暴表徵。特別針對 Tsunami、Tornado 與 Sandstorm 這幾組容易混淆的類別，進行有針對性的合成增強。

2.導入Attention Mechanism：在分類模型中加入 Spatial 或 Channel Attention，可以讓模型更專注於局部關鍵區域，例如「水線」、「沙流紋理」、「建築輪廓」等特徵，降低類別間誤判。

3.Ensemble Learning：

結合 ResNet34 的穩定性、VGG16 對簡單圖形的敏感度與 EfficientNet 的高效特性，建構加權平均的模型集成架構，彌補單一模型對特定類別的盲點。

Reference

https://ithelp.ithome.com.tw/m/articles/10356469

摘要內容使用ChatGPT協助統整

討論內容使用NotebookLM協助統整