

基於機器與深度學習之颱風參數預測模型

Machine and Deep Learning-based Prediction Model of Typhoon Parameters

國立中興大學土木工程學系

研究生
楊騏睿
Chi-Jui Yang

教授
陳佳正
Chia-Jeng Chen

摘 要

颱風對東亞和東南亞國家構成重大威脅，因此亟須一個有效的預測模型以利於防災相關作業。目前關於颱風預測模式有所謂動力模式、統計模式及機器學習等方法，而本研究欲建立一組合多種機器學習方法之模型架構，將颱風衛星雲圖、重分析等資料作為機器學習模型主要輸入，對西北太平洋盆地的颱風各項參數與路徑進行預測。為求最佳的預測表現，模型架構組合了卷積神經網路、隨機森林、迴歸模式及長短期記憶網路等方法，進行不同參數的生成和預測；不同機器學習模型組合方式與其內部參數設定也會有系統地進行調校以比較不同組合下之性能。本模型架構進行預測的颱風關鍵氣象參數包括中心氣壓、強度及移動方向等，而使用基準真相資料為日本氣象廳之颱風 Best Track Data。為評估模型架構之預測表現，本研究採用多種評估指標，包括誤差分布圖、F1 score、平均絕對誤差及颱風路徑生成點位等。此研究成果可望進一步應用於颱風防災相關預測評估(如與區域颱風災損評估模型結合)。

關鍵詞：衛星圖像，機器學習，集成學習

Abstract

Typhoons pose a significant threat to East Asian and Southeast Asian countries, necessitating an effective forecasting model to aid in disaster-related operations. Current typhoon forecasting models include dynamic models, statistical models, and machine learning methods. This study aims to establish a model architecture that combines multiple machine learning methods, using typhoon satellite imagery and reanalysis data as the main inputs for machine learning models to predict various parameters and paths of typhoons in the Northwest Pacific Basin.

To achieve optimal predictive performance, the model architecture combines Convolutional Neural Networks (CNNs), Random Forests, regression models, and Long Short-Term Memory (LSTM) networks to generate and predict different parameters. The combinations of different

machine learning models and their internal parameter settings will be systematically tuned to compare the performance of different integrations.

The key meteorological parameters for typhoon prediction in this model architecture include central pressure, intensity, and direction of movement, using benchmark data from the Japan Meteorological Agency's Typhoon Best Track Data. To evaluate the predictive performance of the model architecture, various evaluation indicators will be used, including error distribution maps, F1 score, mean absolute error, and typhoon path generation points.

The results of this study are expected to be further applied in typhoon disaster prevention-related forecasting and assessment, such as combining with regional typhoon damage assessment models.

Keywords: Satellite Image , Machine Learning , Ensemble Learning

一、前言

颱風是全球眾多地區最具破壞力的自然災害之一，東亞與東南亞國家受其影響尤甚，每年對這些地區造成大量的生命財產損失。準確的颱風預測不僅可以幫助政府和相關機構提前做好防災減災工作，還可以為居民提供必要的避險指引，減少災害造成的損害。因此，開發先進且高效的颱風預測模型是氣象水文學研究的重要方向。

傳統的颱風預測方法主要分為動力模式和統計模式。動力模式基於數值天氣預報模型，通過解決大氣運動方程來預測颱風的路徑和強度[8,10,13]。然而，這種方法計算量龐大且對初始條件敏感，預測精度容易受到數據質量和模型分辨率的限制。統計模式則依賴於歷史數據，利用數學統計方法建立預測模型，雖然計算速度快，但較難捕捉颱風的非線性動態特徵，預測準確性相對較低；而隨著機器學習技術的發展，人們開始探索將其應用於颱風預測的方式。機器學習模型具有強大的數據處理和模式識別能力，能夠從大量的歷史數據中學習並提取特徵，從而提高預測的準確性和效率[4]。其中，Convolutional Neural Network(CNN)作為機器學習的一個分支，因其在圖像處理和特徵提取方面的優勢，逐漸成為颱風預測研究中的熱門方向[1,12]。

本研究中應用了 Digital Typhoon Dataset[6]作為資料集應用於 CNN 進行預測研究，為了提升機器學習的精確率和可靠性，本文引入了 Unet3+模型[3]，Unet3+模型是一種先進的深度學習網路架構，最初設計用於醫學圖像分割。相比傳統 CNN，Unet3+引入了更多的跳躍連接和多層特徵融合機制，能夠更好地捕捉圖像中的細節訊息和全局結構。這些特性使 Unet3+也合適於對颱風衛星雲圖進行特徵分析，從中提取關鍵的氣象參數，如中心氣壓、強度和移動方向等。我們希望能夠更準確地識別颱風的結構特徵，並進一步預測其發展趨勢和路徑；具體來說，本研究首先基於以 Unet3+模型作為骨幹架構以及輸出特徵融合[7]的概念提出的新模型架構對資料進行訓練和測試，評估其在颱風預測中的性能，同時利用 EfficientNetV2 模型[9]進行對照比較，接著進一步將模型產出的特徵匯入如 Random Forest 模型進行訓練以求得到最佳的預測表現[11]；最後，本研究結果可進一步應用於颱風防災相關預測評估[2]，如與區域颱風災損評估模型結合，提升防災減災工作的準確性和有效性。

隨著颱風預測技術的不斷進步，利用深度學習網路對颱風衛星雲圖進行特徵分析與參數預測將成為未來研究的主要方向之一。我們期待本研究能夠為相關領域的研究提供有價值的參考，並為減少颱風災害的影響做出貢獻。

二、研究區域與資料

2.1 研究區域

本研究以西北太平洋為整體區域，採用 2002 至 2022 年之紅外線衛星雲圖產品與其 Best track data(詳如 2.2 至 2.3 節)。

2.2 Digital Typhoon 衛星影像產品

本研究採用 Digital Typhoon 在其 Digital Typhoon: Long-term Satellite Image Dataset for the Spatio-Temporal Modeling of Tropical Cyclones 研究下之西北太平洋衛星觀測颱風雲圖資料集。該資料集是基於日本向日葵衛星紅外線雲圖經由剪裁後之圖片集(以颱風中心擴展 1250km, 512x512 像素)與 Japan Meteorological Agency (JMA)之 Best Track Data 整合而成(詳如圖 1.)。

2.3 各颱風環境因子預測方式與資料前處理

表 1. 展示了於此研究中使用到的各項颱風參數及其單位；圖片集首先經由 Resize 轉換為 256x256 的圖像，再以圖片集之颱風中心於圖像上點位之平均作為中心點進行裁剪至 224x224 之圖像，最後進行正規化。

表 1. 颱風參數來源及各參數單位

Sources	Year	Best Track Input	Unit/Range
JMA Best Track Satellite Image Dataset	2002 - 2022	Lat and Lon	Range: North West Pacific
		TC Class	Range: Class2 To Class8 Using Class 2 To 5
		TC Pressure	Unit: Hpa
		TC Wind	Unit: Knot (1.852km/h)
		TC Direction	Range: From 0 to 9
		TC Radius	Unit: Nautical Mile

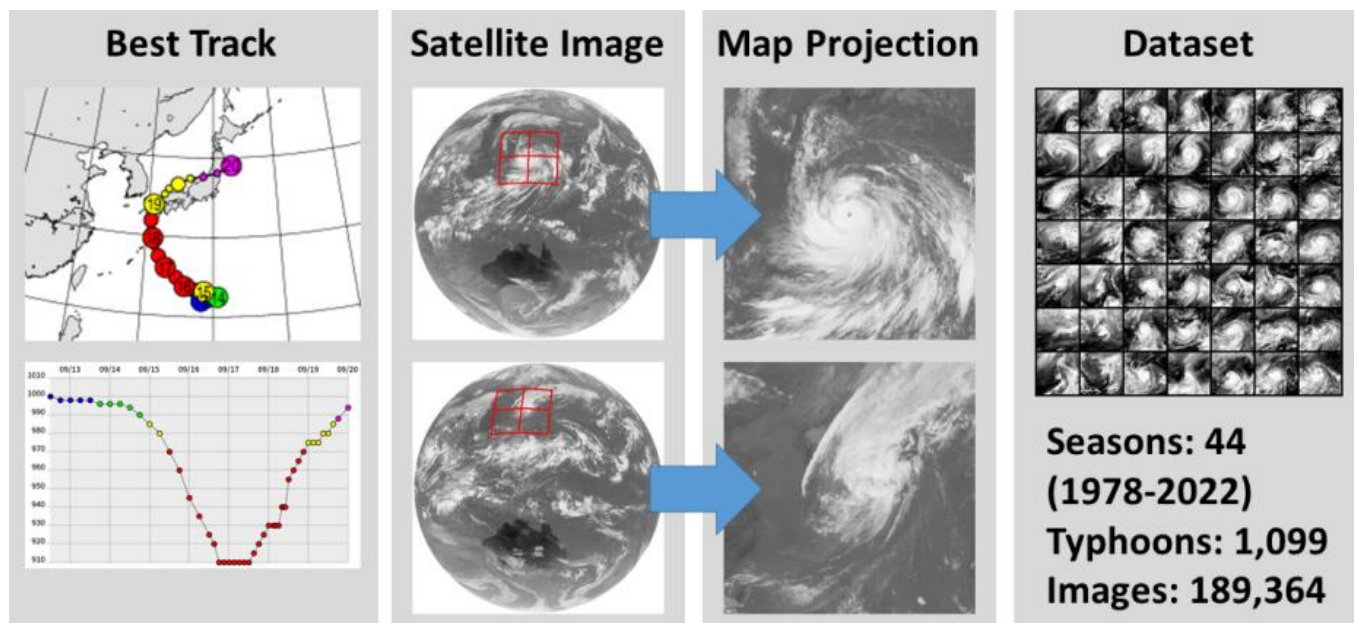


圖 1. Digital Typhoon 圖像資料介紹[6]

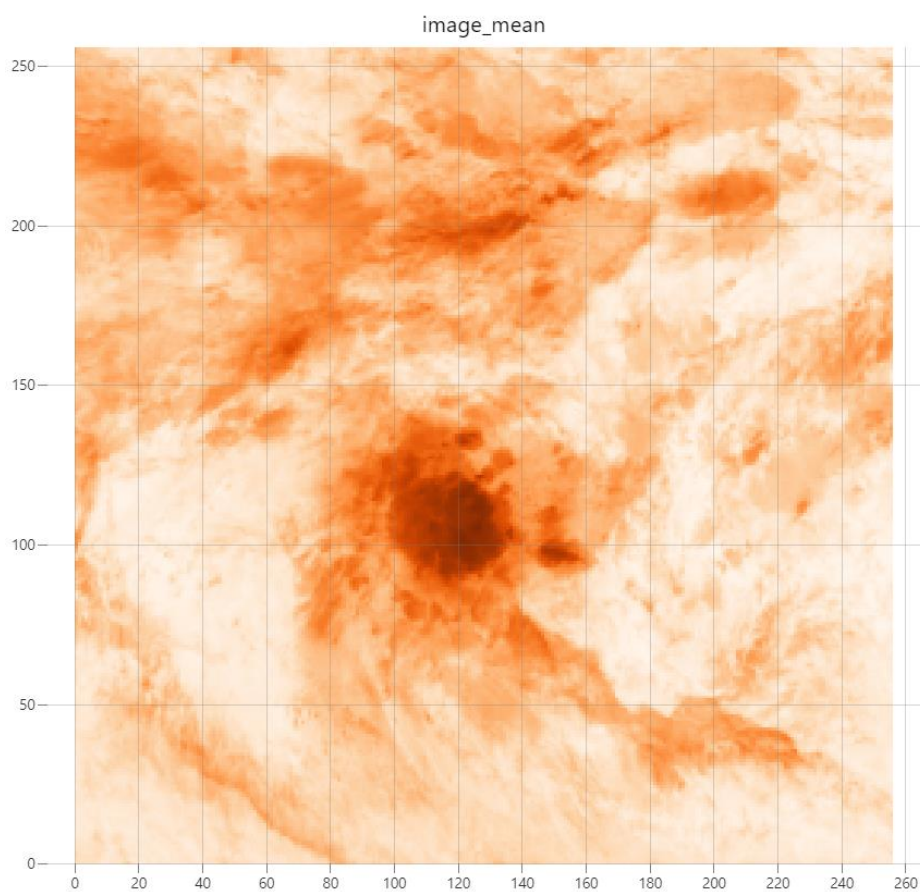


圖 2. 圖片集之平均圖像

三、研究方法

本研究首先將資料進行整合(圖片集與參數資訊結合在將其以 8.4:1.6 之比例分為訓練集與測試集)、正規化及分割後，為進行不同颱風參數中圖片特徵的獲取訓練以及結果比較，圖片訓練集與測試集將進行以下步驟：

- (1) 進行 CNN 模型的特徵提取訓練，本文採用了三種模型(Unet3+轉換模型兩種、EfficientNetV2B0 一種)作為相互對照組。
- (2) 提取 CNN 訓練後產生之特徵，其中兩種 Unet3+模型為 2048 個特徵/每張圖，EfficientNetV2 為 1280 個特徵/每張圖。
- (3) 依照不同的颱風參數進行相應的處理，例如颱風強度是將特徵資料匯入 Random Forest 進行二次訓練後再與 CNN 模型本身架構輸出之預測結果進行諸如準確率與 F1 Score 等指標進行比較。

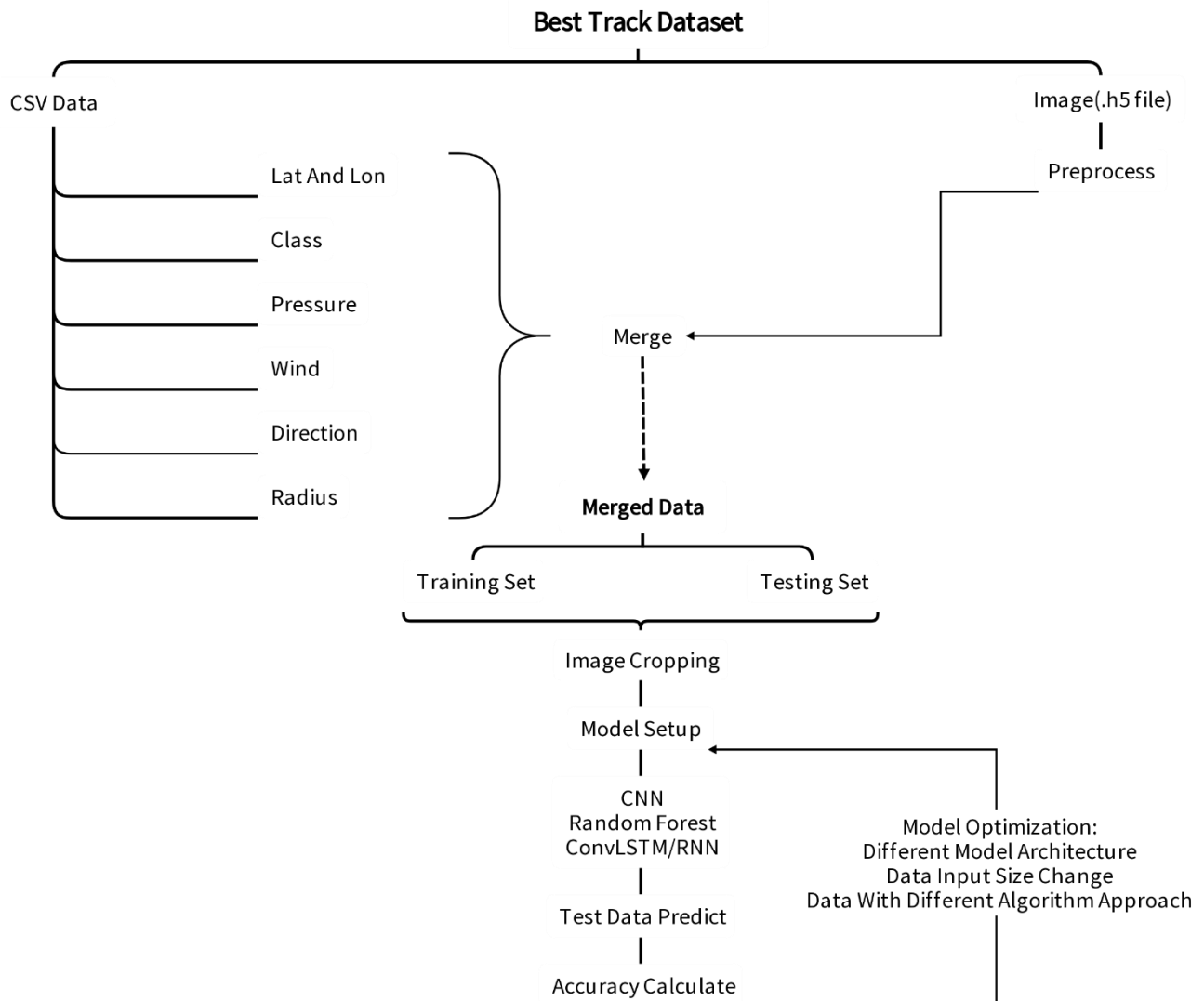


圖 3. 研究流程架構圖

3.1 Unet3+轉換模型

本研究基於 Unet3+模型架構(圖 4.)於特徵獲取的極佳能力，將 Unet3+作為骨幹(backbone)提出了一種新的模型架構，將經由 Unet3+處理輸出後的圖像與原圖像進行資訊的合成以得到更加清晰的訓練特徵，其架構如圖 5.所示。

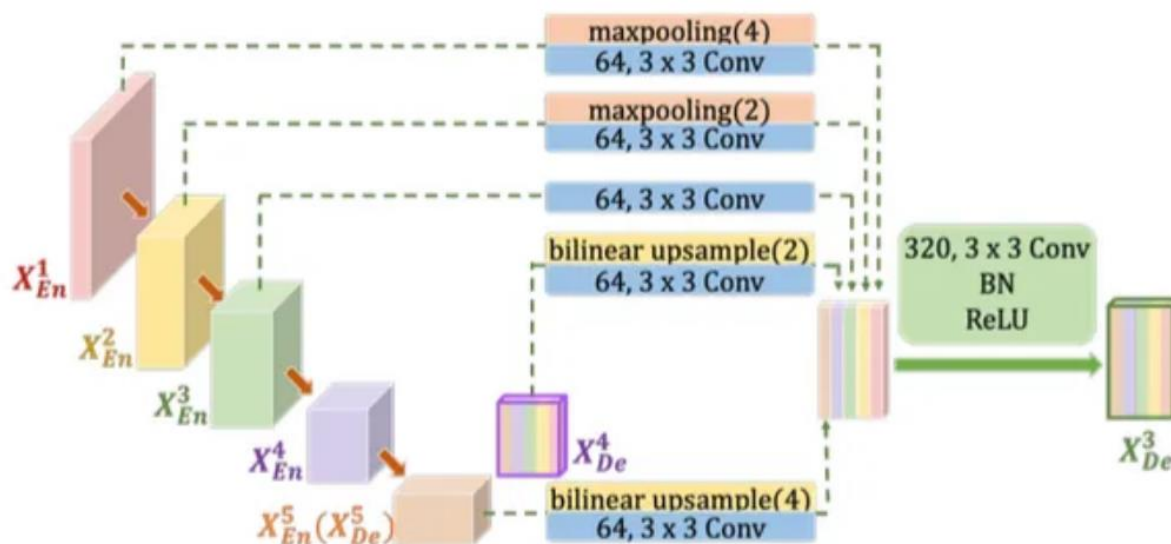


圖 4. Unet3+架構圖[3]

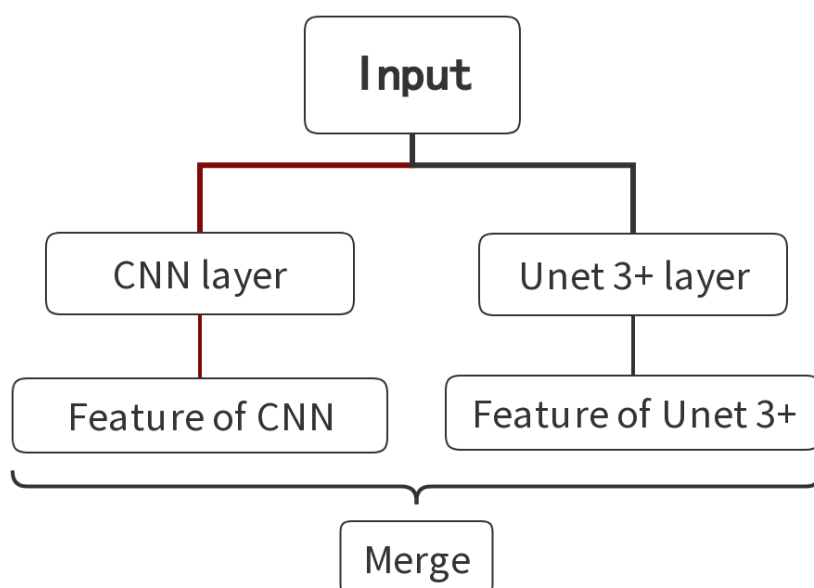


圖 5. Unet3+轉換模型示意圖

3.2 卷積神經網絡 Convolutional Neural Network(CNN)

卷積神經網路是一種機器學習模型類別，擅長於處理具有連續性網格結構的數據，如圖像(CNN2D)或影片片段(CNN3D)。CNN 通過模仿人類視覺系統的工作原理，從輸入數據中自動提取特徵，這使其在圖像分類、物體檢測和圖像分割等任務中表現出色；3.1 的 Unet3+即是應用 CNN 的架構進行延伸的機器學習模型。CNN 的核心構件是卷積層、池化層和連接層。卷積層通過應用多個濾波器(也稱為

卷積核)掃描輸入數據，提取出不同層次的特徵，如邊緣、鋒點和更高級的結構。池化層則負責縮小特徵圖的尺寸，減少計算量，並提高模型對於位置變化的穩定性。最後，全連接層將提取出的特徵組合並輸出最終的分類結果或其他預測結果。

與傳統的手工特徵提取方法相比，CNN 能夠自動學習數據中的有用特徵，且能夠很好地處理高維度和大規模的數據，這使它在計算機視覺等多種領域得到了廣泛應用。

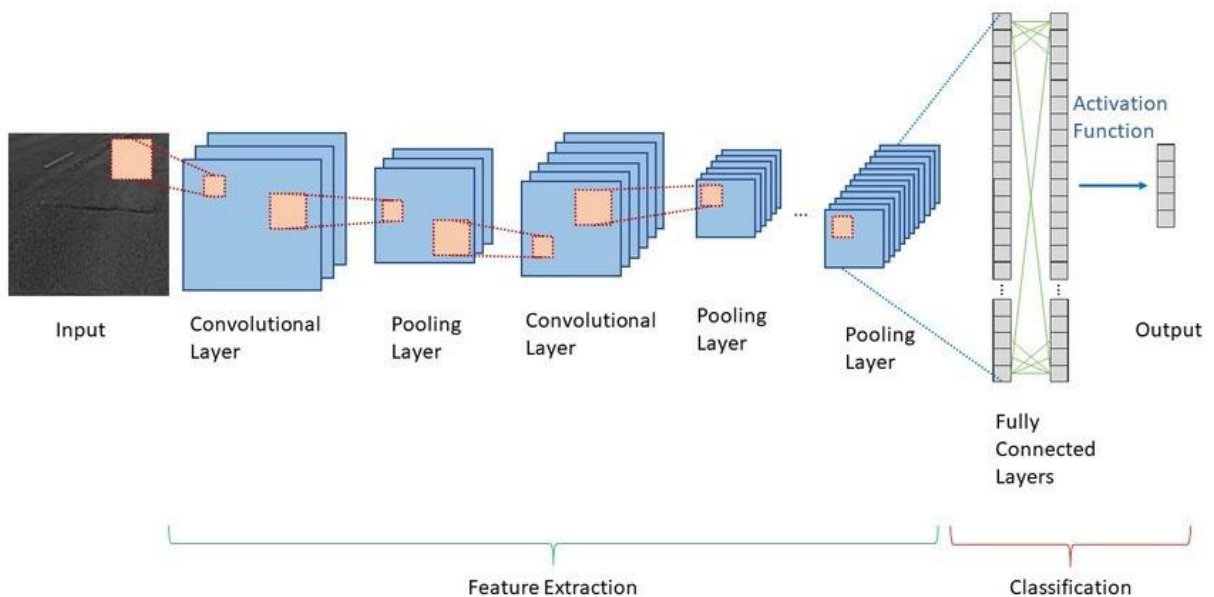
圖 6. CNN 訓練步驟視覺化[5]

3.3 研究成果評估

為了將不同颱風參數預測結果進行準確的比較，於本研究中主要採用了下列的評估方式：

(1)F1 Score

F1 Score 是一種用於衡量模型分類效果的評估指標，特別適合在類別分佈不均衡的



情況下使用。它是由精確率(Precision)和召回率(Recall)的調和平均值；精確率表示在模型預測為正例的樣本中，實際為正例的比例；召回率則表示在所有實際為正例的樣本中，模型正確預測的比例。F1 Score 介於 0 到 1 之間，值越高表示模型的分類效果越好，達到 1 則表示完美分類。F1 Score 能夠平衡精確率和召回率，避免僅關注單一指標所可能導致的偏頗，公式如式 1 所示。

(2) Mean Squared Error (MSE)

MSE(均方誤差)是一種用於衡量回歸模型預測效果的評估指標。它通過計算預測值與實際值之間的差異平方的平均值來衡量模型的預測誤差。具體而言，MSE 將所有預測誤差平方後取平均，結果越小，表示模型的預測越準確。MSE 的單位與目標變數的單位相同，且對於大誤差有較高的懲罰作用，因此在某些情況下會放大模型的錯誤影響。由於其易於理解和計算，MSE 是回歸分析中最常用的評估指標之一，公式如式 2 所示。

(3) 誤差常態分佈

誤差常態分佈是指在統計和數據分析中，模型預測誤差(即預測值與實際值的差異)呈現出常態分佈的情況。這種分佈的特點是誤差大多集中在平均值附近，隨著誤差絕對值增大，其出現的頻率逐漸減少，形成一個對稱的鐘形曲線。常態分佈假設在回歸分析中尤為重要，因為它允許使用許多標準統計方法來進行推論；當誤差符合常態分佈時，模型的預測可靠性通常更高。

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (1)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

四、初步研究結果

研究目前進度為颱風中心氣壓以及颱風強度的預測，颱風中心氣壓預測離散圖(x軸為氣壓之真值，y軸為氣壓之預測值，虛線為回歸基準值)、分布柱狀圖(分為上下兩部分，上半x軸為預測氣壓值，y軸為預測數量，下半則是x軸真值及y軸真值數量)以及預測離散分布曲線圖(評估回歸模型是否符合常態分布，x軸為離散百分比，y軸為不同離散率之數量)分別為圖 7.至圖 9.；颱風強度之示意圖像(藉由將資料集不同強度之颱風圖像進行平均後得出)以及 Unet3+模型產出之特徵圖像(heatmap)範例則為圖 10.、11.，強度訓練未採用 Unet3+模型而是直接進行 Random forest 之預測結果為表 2. 及表 3.。

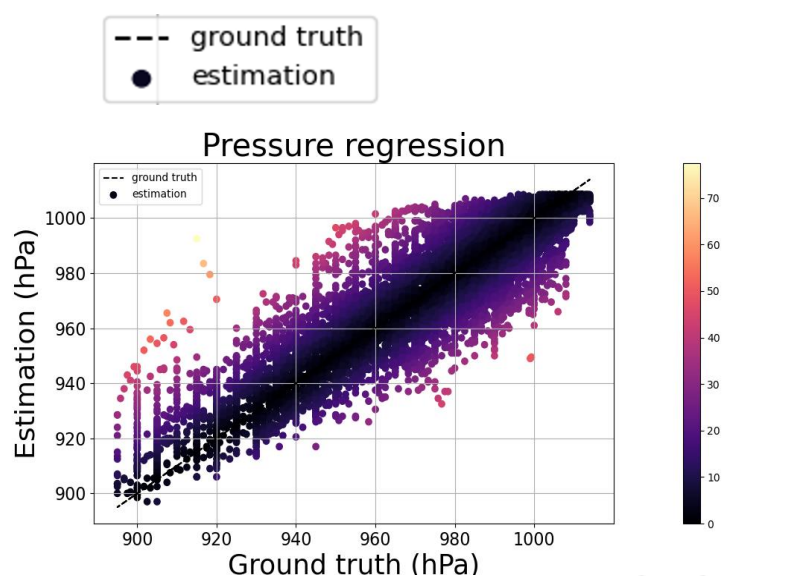


圖 7. 颱風中心氣壓預測離散圖

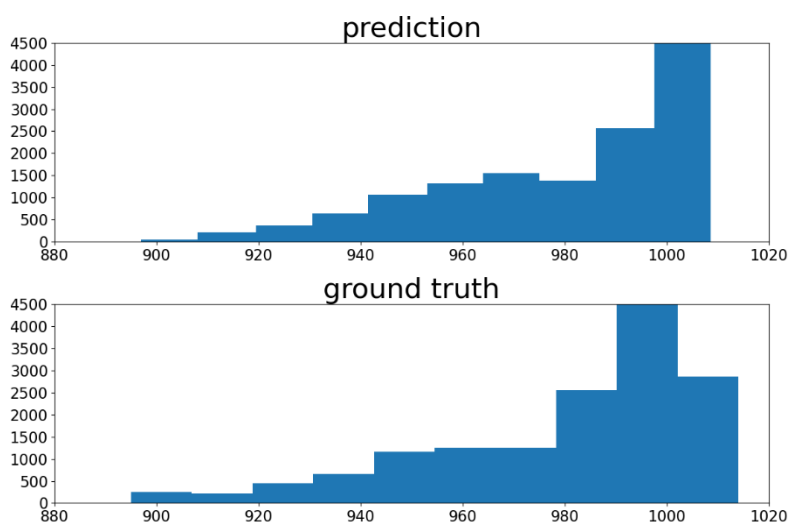


圖 8. 颱風中心氣壓預測分布柱狀圖

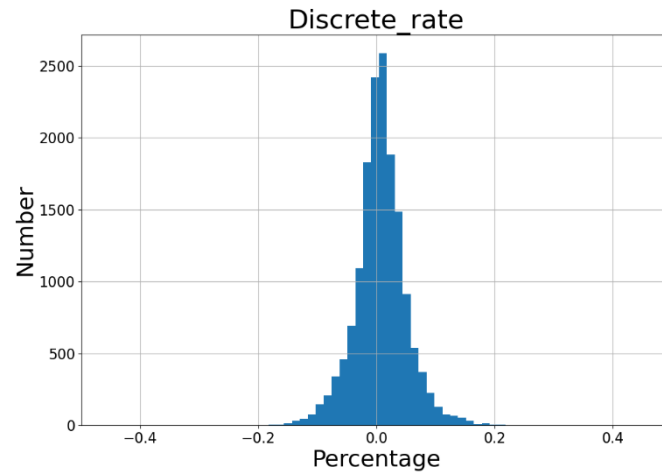


圖 9. 颱風中心氣壓預測離散分布曲線圖

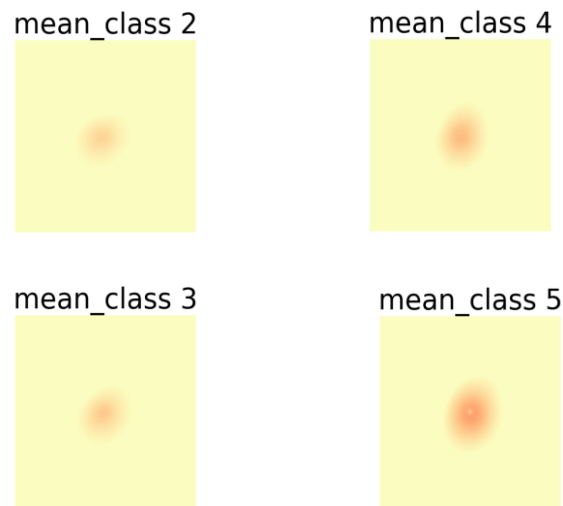


圖 10. 颱風強度之平均圖像

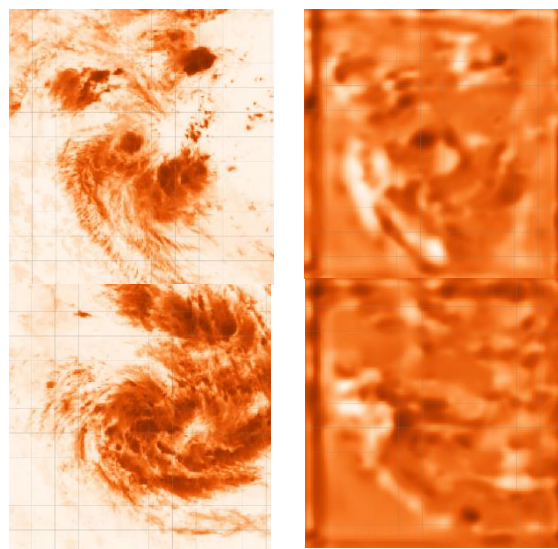


圖 11. Unet3+模型產出之特徵圖範例
，原圖像為左，輸出圖像為右

表 2. Random Forest 經由資料平衡處理後之強度預測混淆矩陣

Balanced Data	Pred. Class 2	Pred. Class 3	Pred. Class 4	Pred. Class 5
Obs. Class 2 (3826)	2441	808	308	269
Obs. Class 3 (3826)	1139	1293	854	540
Obs. Class 4 (3826)	383	846	1383	1214
Obs. Class 5 (3826)	108	182	662	2874

表 3. Random forest 未經由資料平衡處理之強度預測混淆矩陣

Unbalanced Data Class/Predicted	Predicted 2	Predicted 3	Predicted 4	Predicted 5
Class 2 (7121)	4664	1733	50	674
Class 3 (6142)	2055	2652	322	1113
Class 4 (4366)	605	1428	431	1902
Class 5 (7521)	380	574	148	6419

五、小結與後續工作

5.1 小結

本研究分析發現 Unet3+網路在颱風衛星雲圖特徵分析和參數預測方面的確具有優勢，與傳統方法相比，Unet3+能夠更加準確地捕捉雲圖中的細微變化，並將這些變化轉化為有價值的預測特徵。具體而言，利用 Unet3+預測的颱風路徑和強度變化與實際情況的吻合度提高，同時預測誤差也減小。此外，Unet3+模型還具有良好的泛化能力，能夠適應不同時期的颱風數據，這使得該模型在實際應用中具有很高的可行性和可靠性。

5.2 後續工作

本研究後續將針對以下問題進行深究探討：首先，如何進一步提高模型的訓練效率和預測精度仍然具有挑戰性，接續會引入更多對於颱風參數預測有關的環境因子(例如海平面溫度)等資料作為輸入變量，以求達到更佳的預測精度。再者，如何將 Unet3+與除了 Random forest 以外的其餘機器學習模型(如 LSTM 等)進行整合，以形成更為強大及泛化的預測系統，也是未來研究的方向之一。最後，颱風預測不僅需要高精度的模型，還需要考慮實際應用中的實時性和穩定性。因此，如何將 Unet3+模型部署到實時預測系統中，並確保其在不同運行環境下的穩定性和可靠性，也是需要重點關注的問題。

六、參考文獻

1. Chen, Rui, Xiang Wang, Weimin Zhang, Xiaoyu Zhu, Aiping Li, and Chao Yang. "A hybrid CNN-LSTM model for typhoon formation forecasting." *GeoInformatica* 23, 375-396, 2019.
2. Haghroosta, Tahereh, and Wan Ruslan Ismail. "Typhoon activity and some important parameters in the South China Sea." *Weather and Climate Extremes* 17, 29-35, 2017.
3. Huang, Huimin, Lanfen Lin, Ruofeng Tong, Hongjie Hu, Qiaowei Zhang, Yutaro Iwamoto, Xianhua Han, Yen-Wei Chen, and Jian Wu. "Unet 3+: A full-scale connected unet for medical image segmentation." In *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 1055-1059. IEEE, 2020.
4. Jiang, Guo-Qing, Jing Xu, and Jun Wei. "A deep learning algorithm of neural network for the parameterization of typhoon-ocean feedback in typhoon forecast models." *Geophysical Research Letters* 45, no. 8, 3706-3716, 2018.
5. Khan, Salman, Hossein Rahmani, Syed Afaq Ali Shah, Mohammed Bennamoun, Gerard Medioni, and Sven Dickinson. "A guide to convolutional neural networks for computer vision.", 1-207, 2018.
6. Kitamoto, Asanobu, Jared Hwang, Bastien Vuillod, Lucas Gautier, Yingtao Tian, and Tarin Clanuwat. "Digital typhoon: Long-term satellite image dataset for the spatio-temporal modeling of tropical cyclones." *Advances in Neural Information Processing Systems* 36, 2024.
7. Ma, Wenping, Na Li, Hao Zhu, Licheng Jiao, Xu Tang, Yuwei Guo, and Biao Hou. "Feature split-merge-enhancement network for remote sensing object detection." *IEEE Transactions on Geoscience and Remote Sensing* 60, 1-17, 2022.
8. Meng, Yan, Masahiro Matsui, and Kazuki Hibi. "An analytical model for simulation of the wind field in a typhoon boundary layer." *Journal of wind engineering and industrial aerodynamics* 56, no. 2-3, 291-310, 1995.
9. Tan, Mingxing, and Quoc Le. "Efficientnetv2: Smaller models and faster training." In *International conference on machine learning*, pp. 10096-10106. PMLR, 2021.
10. Wada, Akiyoshi, and Norihisa Usui. "Importance of tropical cyclone heat potential for tropical cyclone intensity and intensification in the western North Pacific." *Journal of Oceanography* 63, 427-447, 2007.
11. Wang, Aili, Ying Wang, and Yushi Chen. "Hyperspectral image classification based on convolutional neural network and random forest." *Remote sensing letters* 10, no. 11, 1086-1094, 2019.
12. Wang, Chong, Qing Xu, Xiaofeng Li, and Yongcun Cheng. "CNN-based tropical cyclone track forecasting from satellite infrared images." In *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*, pp. 5811-5814. IEEE, 2020.
13. Yanai, Michio. "A detailed analysis of typhoon formation." *Journal of the Meteorological*

Society of Japan. Ser. II 39, no. 4, 187-214, 1961.