國立雲林科技大學工業工程與管理系


## 摘 要

人流來源探究長期以來都是國家災防議題最重要的一環，再加上近兩年來 COVID-19 爆發，「社交距離」已成為生活中的日常用詞。尤其在本土疫情升溫之時，非必要不外出，成為政府頻頻呼籲的防疫指南。但社交距離政策如何防控疫情？許多新聞裡總是提及部分場所依然人潮洶湧，政府所宣布之三級警戒是否減低人潮移動？有鑑於未來人流趨勢何去何從，假設政府能夠提前得知未來數小時何地人流會出現異常的情形，以及得知影響其的最關鍵因素，則政府即可提前做出相對應措施以避免後續問題出現，甚至是能夠對症下藥來化解人們社交距離過於相近的問題。然而在早期有關這方面的研究，其預測效果非常不佳，因為當時並沒有實際人流數據的支持，所以學者們根本無從了解人在哪裡，更不用說是判斷某地的人流是否處於過於擁擠的情況。

在近年台灣交通大數據逐漸釋出後，越來越多學者嘗試使用這些交通大數據結合機器學習的方法進行各種人流議題分析，並宣稱他們研究的合理性與有效性。但在我們詳細閱讀過這些研究後，我們發現這些研究大多有個共通的盲點-分析中參雜太多專家既有知識，並因此屏蔽了資料本身所帶有的意涵，而這點是現代大數據分析概念中所不贊同的。

一般來說，一個好的大數據分析流程不應該帶有任何既有知識的偏見，應該僅靠分析資料來了解資料內所附帶的知識，如此才能真正從大量數據中了解事情的真相。舉一個學界公認最有貢獻的大數據分析案例來說，過往醫生僅知道九種乳癌影像的特徵，但學者們在透過訓練並拆解 Convolution neural networks (CNN) 後，得知了三種新的乳癌影像特徵，且透過這三種特徵進行的辨識準確率遠較過往九種特徵高。也因此，為了弭平過往人流研究與大數據觀念間的差異，本研究將嘗試新的人流分析概念。

本研究中包含(1)僅使用資料本身，而不參雜任何既有經驗來進行人流來源探討，(2)效仿過往乳癌影像新特徵值找尋的做法，先使用 CNN 模型來對交通資料進行建模，並嘗試拆解這個 CNN 以獲得目標地點的重要人流來源，以及(3)由於拆解 CNN 內部的參數過於困難，我們將採用近期較知名的 Grad-CAM 套件來完成，其中這個套件能夠有效告訴我們每筆輸入的地圖影像中，CNN 究竟判斷哪些網格對目標網格來說是重要的人流來源，並協助我們進行後續分析。藉由上述的多個概念，我們預期這個研究能打破過往人流分析的既定框架，為人流分析開創一條新的路線。


**關鍵字: 人流來源探究、CNN、Grad-CAM**