

高速公路短期車流量之預測 ——以LSTM-RNN深度學習方法

作者/張軒慈

審者/溫志皓、王中允

提要

- 一、伴隨國人車輛持有數及交通需求的增加，造成塞車問題日益加劇；另用路人行駛時間增長，亦可能產生行車安全問題。交通部於國道設置電子收費系統（Electronic Toll Collection, ETC），除可紓緩並降低人力成本外。並可透過「交通資料庫」提供大量且完整的國道交通流量資訊。
- 二、本研究蒐集平日上班尖峰時段，壅塞機率最高的國道1號「五股至圓山」路段之車流ETC資料庫資料，運用深度學習中適合時間序列預測的長短期記憶演算法（Long Short-Term Memory, LSTM）以建立車流預測模型，透過實驗超參數調校（hyper-parameter turning）使模型優化，最後利用平均絕對百分比誤差（MAPE）以評估模型的準確性。
- 三、透過實驗調整超參數（hyper-parameter）優化模型後，預測5分鐘後之短期車流量，預測結果MAPE值為6.96，可達到相當精確的結果，若將本模型套用至各單位出勤路線之路段，可提供國軍單位派車及路線選擇決策，進而降低行車風險及提高效率。

關鍵詞：交通流量預測、電子收費系統ETC、深度學習、長短期記憶模型LSTM

圖片來源：設計者拍攝



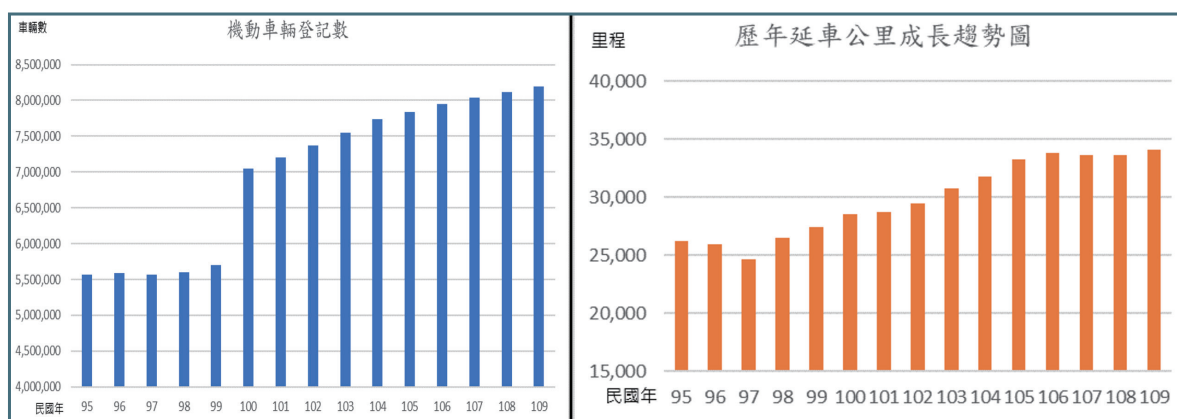
壹、前言

伴隨經濟發展，各城鎮的都市化現象日趨明顯，國人車輛持有數也明顯增加，交通部公路總局統計資料指出，全國車輛總數（含各式客、貨車及特種車輛）自95年5,566,318輛上升至110年8,330,774輛，成長幅度達49.7%。¹此外，高速公路自通車以來交通需求亦逐年增加，根據交通部高速公路局（以下稱高公局）統計資料，國道年度之「百萬車公里」（所有車輛行駛里程加總），亦同樣由95年26,235百萬車公里上升至110年32,400百萬車公里，成長幅度達23.5%。²然高速

公路年報統計顯示，國道總面積自95年7,192.99公頃至110年8,436.83公頃，³僅成長17.3%，遠低於車輛數及行駛里程之增幅，造成塞車問題日益加劇，不僅延長用路人行駛時間，亦容易產生行車安全及能源浪費等問題（如圖一）。⁴

交通需求具有時間集中（如尖峰時段、連續假期）、空間集中（如都會區路段）等特性，在平日有通勤旅次，假日有遊憩旅次及客、貨運輸等情形下，國道部分路段及某些時段容易發生重現性壅塞（recurrent congestion）情形，若遇不可預期事故恐更將加劇回堵狀況。

高公局為紓緩壅塞並降低人力負



圖一 機動車輛及延車公里成長趨勢圖

資料來源：交通部公路總局統計查詢網

- 1 《交通部公路局統計查詢網》，〈<https://stat.thb.gov.tw/hb01/webMain.aspx?sys=100&funid=11300>〉（檢索日期：民國113年8月4日）。
- 2 同註1。
- 3 《交通部高速公路局》，〈網址<http://www.freeway.gov.tw/default.aspx>〉（檢索日期：民國113年8月4日）。
- 4 同註1。

荷，於102年12月30日起導入電子收費系統(Electronic Toll Collection, ETC)，此外，其收費門架的資料累積大量且完整的資料，資料集包含有旅次起訖、時間、各車輛之交通流量與車速，經統計至112年底，ETC供裝客戶數約850.1萬輛，ETC利用率為92.5%。⁵ 為有利促進及活化政府資料應用並提升政府資料品質及價值，高公局自104年6月起，開放「交通資料庫」免費下載應用，協助用路人獲取更多元之服務。

如何獲得有效的交通資訊是智慧運輸的重要課題，精準且可靠的短期交通流量預測是許多智慧運輸系統(Intelligent Transportation Systems, ITS)的基礎，而可靠的長期旅行時間預測將有效支援交通管理，像是交通流量控制或通行費定價。⁶ 預測短期交通流量

之研究已行之有年，許多研究嘗試以不同之方法預測，以達到更精準之預測結果。⁷ 近年來，亦有不少研究使用深度神經網路於城市地區作交通預測，與傳統方法(如統計或機器學習模型)相比，深度學習使模型能使用多層體系結構發現複雜的結構和複雜的模式。⁸

有學者提出若將平均時速低於40公里視為壅塞觀察，國道各路段平日以國道1號「五股至圓山」，其壅塞機率高達94%，且指出七成旅次集中大都會區、旅程20公里以下之都會運輸旅次約占2/3，且逾4成車輛免付通行費，⁹ 顯示高速公路實施每日20公里免費政策，已影響用路人通行習慣，故本研究將針對此路段進行分析並預測。

由交通預測需求以及國內外研究結論可知，交通預測已逐漸轉為基於機器

5 同註3。

6 Che-Ming Chen, Chia-ching Liang, Chih-Peng chu (2020), "Long-term travel time prediction using gradient boosting," Journal of Intelligent Transportation Systems, Vol. 24, No. 2, pp. 109-124.

7 吳佩倫，〈深度學習應用於短期國道交通流量預測〉，中華民國運輸學會107年學術論文國際研討會(臺中市：中華民國運輸學會，民國107年12月6日)，頁2466-2483。

8 Hongsuk Yi, Khac-Hoai Nam Bui, Heejin Jung, "Implementing a deep learning framework for short term traffic flow prediction." the 9th international conference on web intelligence, mining, and semantics(2019), pp. 1-8.

9 郭昌儒，〈探勘交通統計大數據(Big Data)—高速公路易壅塞路段概況分析〉，《主計月刊》，第710期，西元2015年2月，頁78-85。

學習甚至進階至架構較為複雜之深度學習方法，不僅在運算速度及考慮資料特徵屬性較為廣泛，同時也能獲得較佳的預測結果。儘管國際已有許多交通預測相關研究，惟國內研究仍為少數，據此，本研究規劃運用深度學習方法於時間序列資料分析，進而發展短期交通流量預測模型，並針對國道車流量及模型架構進行探討。

國軍車輛出勤任務執行期間，若遇交通壅塞，除造成無法在時限內完成任務，甚或肇生行車安全問題，致使無法如期如質完成任務，故本研究期能找出適合預測交通流量之模型，並透過超參數之調整，藉由先期預測車流量的方法，可準確預估運勤任務執行所需之時間，並適當避開易壅塞之路段，有利於出勤前完成良好的出發時間及路線之規劃，並將本模型套用至各單位出勤路線之路段，可提供國軍單位派車及路線選擇決策，進而降低行車風險及提高效率。

貳、文獻探討

準確預測未來的交通狀況（如交通流量，旅行速度和旅行時間）是智慧運輸系統的關鍵要求，有助於管理者採取適當的預防措施以應對壅塞，並讓用路人做出更明智的決定。自1980年代以來，短期交通預測已成為大多數智慧運輸系統研究和應用不可或缺的部分，大部分研究致力於對交通特徵進行模式建立並產生預期交通狀況的方法，而短期交通預測的兩個挑戰是模型選擇和模型性能的判斷，¹⁰以下對短期交通流量預測方法及深度學習方法進行延伸探討。

一、交通流量管理

旅行時間預測分為短期和長期兩類，短期旅行時間即幾分鐘（或幾個小時）前收集的數據，可用於判斷即時交通狀況，並允許用路人重新選擇路線。幾天或幾周後的長期旅行時間預測，則依賴幾天前或幾周前獲得的關鍵變數之數據，亦是政府交通規劃的重要參考。¹¹準確即時的預測交通流量訊息可協助用路人判斷出發時間及路線，進而節省時間、減少能源損耗並降低污染。

二、短期交通流量預測方法

10 Eleni I. Vlahogianni, Matthew G. Karlaftis, John C. Golias (2014). "Short-term traffic forecasting: Where we are and where we're going," Transportation Research Part C: Emerging Technologies, Vol. 43, pp. 3-19.

11 同註6。

交通流量預測於1970年代起即受到重視，亦是智慧運輸系統最關鍵的環節之一，¹²而短期交通流量預測為時間序列的問題，是一種量化的預測方法，其根據某一變項過去的變動以預測未來可能產生的變化。

在過去已有許多學者提出多種預測交通流量之方法，像是基於自回歸移動平均模型（Autoregressive Integrated Moving Average Model，ARIMA）的交通流量預測。然許多研究發現流量數據是隨機且非線性的，而因ARIMA是基於線性關係，故難以分析非線性流量數據。此外，還提出如時間序列、支持向量機（Support Vector Machine, SVM）、K-近鄰演算法（K-nearest Neighbors, KNN）等機器學習方法。¹³

現有的交通流量預測方法主要使用

淺層交通預測模型，惟無法滿足於許多實際應用上，這種情況激發研究需重新思考基於具有大量交通數據的深度架構模型的交通流量預測問題。¹⁴相較傳統統計方法短期交通流量預測建構模型，深度學習之技術能達到更精準之預測。¹⁵近來研究和開發人員則聚焦於深度神經網路在城市地區的交通預測，相較傳統方法（如統計或機器學習模型），深度學習使模型能使用多層體系結構發現複雜的結構與模式。¹⁶由於交通流量的隨機性和非線性性質，參數方法無法有效擷取交通特徵，故研究人員在交通流量預測領域更加關注非參數方法。¹⁷

綜上可知，交通問題通常有較為複雜、隨機和非線性的性質，大部分研究認為參數的傳統統計方法較難擷取交通特徵，而機器學習使用非參數方法能使模

12 Bailin Yang, Shulin Sun, Jianyuan Li, Xianxuan Lin, Yan Tian (2019). "Traffic flow prediction using LSTM with feature enhancement," *Neurocomputing*, Vol. 332, pp. 320-327.

13 Yipeng Liu, Haifeng Zheng, Xinxin Feng, Zhonghui Chen, "Short-term traffic flow prediction with Conv-LSTM," the 9th International Conference on Wireless Communications and Signal Processing (China: WCSP, Oct 11-13, 2017).

14 Zhongsheng Hou, Xingyi Li (2016). "Repeatability and similarity of freeway traffic flow and long-term prediction under big data," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17, No. 6, pp. 1786-1796.

15 同註7。

16 同註8。

17 同註10。

型藉由多個隱藏層的深度結構及算法對複雜且隨機的模式進行預測，從而達到較佳的預測結果。

短期交通流量預測已發展數十年，在現今資料量大增的狀況下，以深度學習模型架構預測交通流量的效果普遍較佳，且近期的深度學習方法相較以往的方法已取得卓越的成果，故後續擬針對臺灣國道以深度學習之架構模型作流量預測，接續對深度學習方法進行探討。

三、深度學習方法

深度學習是一種新興的機器學習方法，近期引起學術界和工業界的廣泛關注，相較簡單的神經網路結構，其具有更高的層次深度，可學習數據的更多特徵；¹⁸深度學習具有較少的計算複雜性和更好的特徵提取等優勢，並在自然語言處理、語音識別和圖像識別等各個領域取得成功。¹⁹

深度學習方法使用多層結構提取數據中的固有特徵，可在數據中發現大量的結構。由於交通流量過程性質複雜，深

度學習方法無需事先瞭解即可表示流量特徵，對流量預測具有良好的性能。²⁰在智慧運輸系統中，深度學習方法應用於準確且及時的交通流量訊息，係改善交通系統的重點，而近期研究人員和開發人員專注於都市地區深度神經網路進行交通預測，相較傳統方法（如統計或機器學習模型），深度學習使模型能使用多層架構發現複雜的結構和模式。²¹

深度學習有足夠的模型複雜度和表達能力可解決較困難的工作，現階段已被大量應用於機器視覺、語音辨識、圖形辨識和語言翻譯等方面。而依不同對象又變化出更多不同種類的類神經網路，如適合圖型和聲音辨識的卷積神經網路（Convolutional Neural Networks, CNN）、以時間序列為重要學習特徵的遞歸神經網路（Recurrent Neural Networks, RNN）等，這些有特定解決目標的類神經網路又被改良為更多不同形式，使效能和辨識率大幅提升，而這些不同類型的深度神經網路亦可互相搭配使

18 同註10。

19 Yoshua Bengio, Learning deep architectures for AI (Now Foundations and Trends, 2009).

20 Yisheng Lv, Yanjie Duan, Wenwen Kang, Zhengxi Li, Fei-Yue Wang (2015), "Traffic flow prediction with big data: a deep learning approach," IEEE Transactions on Intelligent Transportation Systems, Vol. 16, No. 2, pp. 865-873.

21 同註8。

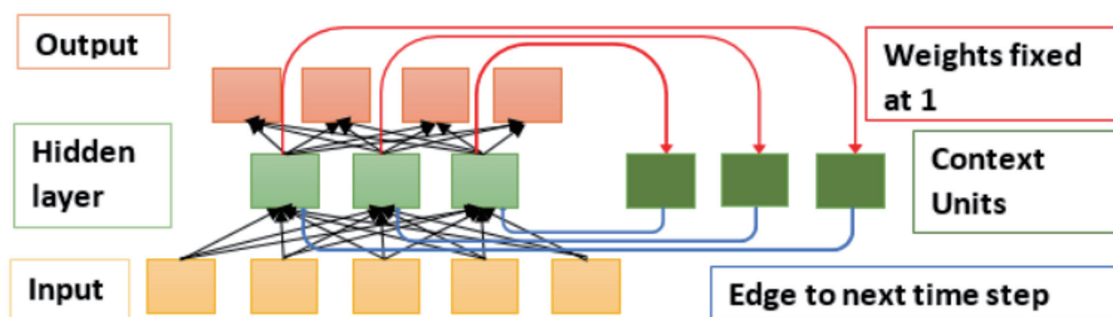
用，使模型更為複雜且穩健，同時也能解決更困難的問題，²²深度學習應用於時間序列分析，最常見的便是遞歸神經網路，而本研究主要為時間序列的交通流量分析，故後續將深入探討遞歸神經網路的相關文獻。

（一）遞歸神經網路

遞歸神經網路是一種監督式的深度學習模型，通常用於擷取動態數據序列，透過引入反饋連接解釋數據，其可存儲並捕捉最近輸入的數據序列。此功能可使存儲器有選擇地跨序列步驟傳遞訊息，以在特定時間處理數據。因此，每

個狀態均取決於當前輸入和先前時間的網路狀態。在此體系結構中，與每個隱藏單元相關聯的是一個上下文單元，該上下文單元將上一時間對應單元的確切狀態輸入，並在下一步將其與學習到的權重一起重新饋送到同一單元（如圖二）。²³

一般的遞歸神經網路無法長期記憶，若模型的隱藏層增加時，則可能造成梯度消失 (Vanishing Gradient) 或梯度爆炸 (Gradient Exploding) 的問題。²⁴梯度信號沿著隱藏層前進，並乘以神經元的權重矩陣，若權重矩陣太小或太大，則



圖二 遞歸神經網路結構圖

資料來源：Deep learning, nature

- 22 Yann LeCun, Yoshua Bengio, Geoffrey Hinton (2015), "Deep learning," nature, Vol. 521, pp. 436-444.
- 23 Arya Ketabchi Haghighat, Varsha Ravichandra-Mouli, Pranamesh Chakraborty, Yasaman Esfandiari, Saeed Arabi, Anuj Sharma (2020), "Applications of Deep Learning in Intelligent Transportation Systems," Journal of Big Data Analytics in Transportation, Vol. 2, No. 2, pp. 115-145.
- 24 同註22。

梯度信號最終將變得太小而使學習停止工作，或因太大而使學習分散。²⁵

(二) 長短期記憶神經網路

為解決前述問題，Hochreiter and Schmidhuber於1997年提出長短期記憶網路(LSTM)。²⁶相比基本RNN，其具有輸入閘(Input Gate)、輸出閘(Output Gate)，而後Gers, Schmidhuber, and Cummins於1999年發現當輸入時間序列延長時，兩個閘的LSTM模型將受到限制。²⁷此種LSTM傾向對線性過程進行建模，原始數據集中的非線性方面將被遺忘，為處理此問題，而將遺忘閘(Forget Gate)加入LSTM的架構中。²⁸

LSTM的關鍵思想是一種稱為存儲單元的新結構。其包含一個具有自循環連接的神經元及與之關聯的三個閘(輸入、輸出和遺忘閘)，自循環連接使存儲

單元的狀態從一個時間步至另一個時間步保持恆定，而不受外界干擾。輸入和輸出閘可使信號進入或離開神經元或阻止它。遺忘閘可調節存儲單元的自循環連接，從而使該單元能記住或忘記其先前狀態。²⁹Wang et al.(2017)以長短期記憶(LSTM)遞歸神經網路為基礎，透過殘差連接、深度層次網路和雙向流量提出深度雙向長短期記憶(DBL)，使用迴歸層進行監督預測，並採用丟棄(dropout)的訓練法避免過度擬合問題，該模型能挖掘流量的深層特徵，並充分利用時間感知流量數據，且證明模型有良好的普適能力。³⁰

近年來深度學習的演算法在交通預測上已廣泛應用，以長短期記憶演算法為基礎不在少數，也證實預測準確率較傳統的參數方法更佳，考量國內利用高

25 Hongxin Shao; Boon-Hee Soong(2016).“Traffic flow prediction with long short-term memory networks (LSTMs),”the 2016 IEEE Region 10 Conference(Singapore, Nov 22-25, 2016).

26 Sepp Hochreiter, Jürgen Schmidhuber(1997), “Long short-term memory,” Neural computation, Vol. 9, No. 8, pp. 1735-1780.

27 F. A. Gers, J. Schmidhuber, F. Cummins, Learning to forget: continual prediction with LSTM, the 1999 Ninth International Conference on Artificial Neural Networks ICANN 99 (Sept 7-10,1999).

28 Azzedine Boukerche, Jiahao Wang (2020). “Machine Learning-based traffic prediction models for Intelligent Transportation Systems,” Computer Networks, Vol 181.

29 同註25。

30 Jingyuan Wang, Fei Hu, Li Li, Deep bi-directional long short-term memory model for short-term traffic flow prediction, Neural information processing(2017), pp. 306-316.

速公路局ETC開放資料，並應用深度學習方法建立交通流量預測模型之研究較少，故本研究後續將利用長短期記憶演算法設計模型架構，以ETC資料集訓練模型，並透過調整超參數與測試模型架構優化模型，期能得到精確的預測結果。

參、研究方法

本研究使用長短期記憶演算法，針對國道壅塞路段建立模型，第一節為研究架構。第二節為研究資料及對象，以高公局之數據資料針對所選擇之壅塞路段，因2020年Covid-19疫情於臺灣逐漸爆發，車流量較以往大幅下降，為避免研究偏差，且機器學習演算法需達一定數量之樣本數，故本研究選擇疫情趨緩後2021年7至9月共三個月之車流資料，並考量本研究路設「五股至圓山」平日壅塞機率高達94%，故僅以平日做為研究範圍。

第三節為模型建立及優化，透過長短期記憶演算法建立模型並針對各項超參數進行測試，並比較各種優化器之效能，將資料分為前兩個月訓練集及後一個月測試集以驗證模型。第四節為模

型評估，以平均絕對百分比誤差（Mean Absolute Percentage Error，MAPE）衡量指標做為評估各模型之預測效能。

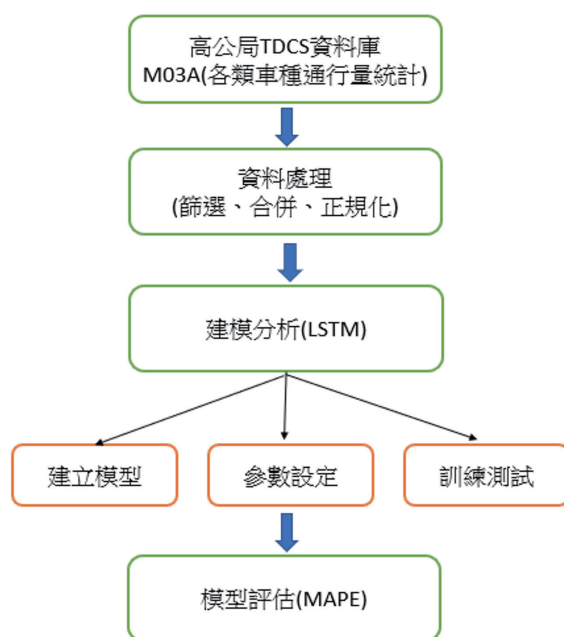
一、研究架構

本研究透過文獻探討，建構一個交通流量預測模型，以高公局「高速公路電子收費交通資料蒐集支援系統」（Traffic Data Collection System, TDCS），所提供之高速公路各個E-Tag感應門架所蒐集到的流量、速度、車行路徑等原始資料集，針對國道易壅塞路段進行資料篩選後，以遞歸神經網路中長短期記憶演算法，實驗並找出最適之超參數建構以及優化模型，透過MAPE值做為模型評估指標，以確認模型優劣，期望模型的預測效果可有效且即時提供用路人快速做出行駛決策，並達到分散車流、壅塞紓緩之成效（如圖三）。

二、研究對象

本研究參考郭昌儒於2015年研究提出平日上班尖峰時段，其壅塞機率最高的國道1號「五股至圓山」路段，³¹為確認該路段現況，利用2021年4月19日Etag_M03A各類車種通行量統計資料內之01F0233N（臺北至圓山）門架，檢視車速變化情形，發現「五股至圓山」路段在平日尖峰時段車速及流量仍維持高峰狀

31 同註9。



圖三 研究架構圖

資料來源:本研究整理

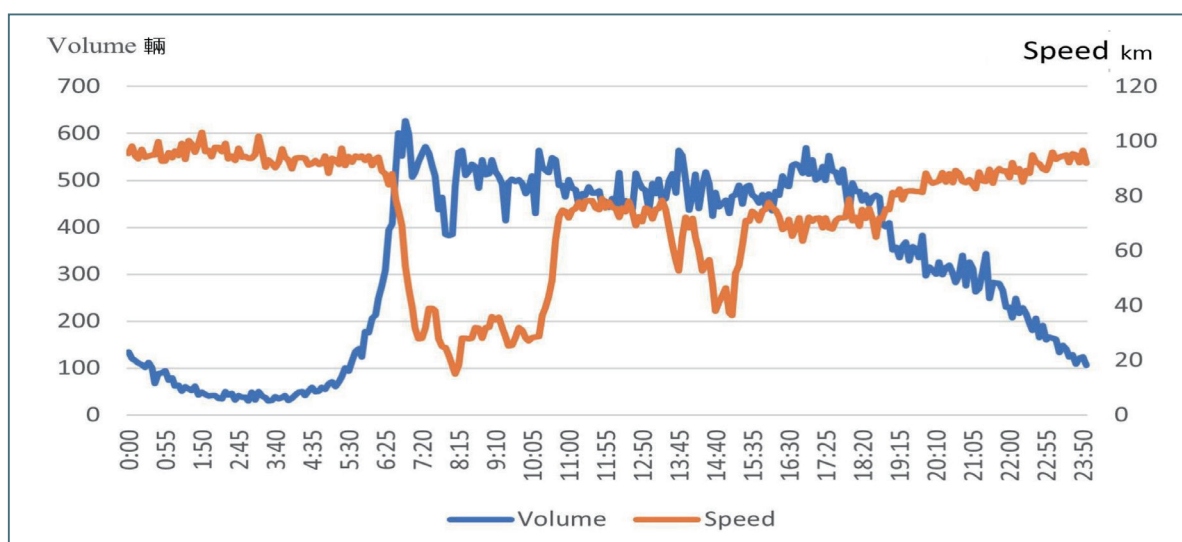
態，故設定以此為研究對象（如圖四）。

本研究使用高公局「交通資料庫（TDCS）」，之eTag_M03A（各類車種通

行量統計）內之01F0233N門架資料，蒐集2021年7至9月上班日將各類車種之交通量（Volume）加總後做為研究範圍，而TDCS資料間隔為每5分鐘乙次，每日288筆，研究期間共64日，總計18,432筆資料，分為前兩個月為訓練集（training dataset）及後一個月為測試集（test dataset）做為資料之劃分。

三、模型建立與優化

本研究欲建構交通流量預測模型，而交通流量具有穩定的週期變化趨勢，且能以過去時間的資料分析後，據以預測未來之變化，符合時間序列之特性，由於遞歸神經網路中的循環允許將訊息從網路的一個步驟傳遞到下一個步驟，但傳統的遞歸神經網路無法記憶較長遠的資料，而衍生長短期記憶模型改善此問



圖四 2021年4月19日五股至圓山（北上）車速及流量變化圖

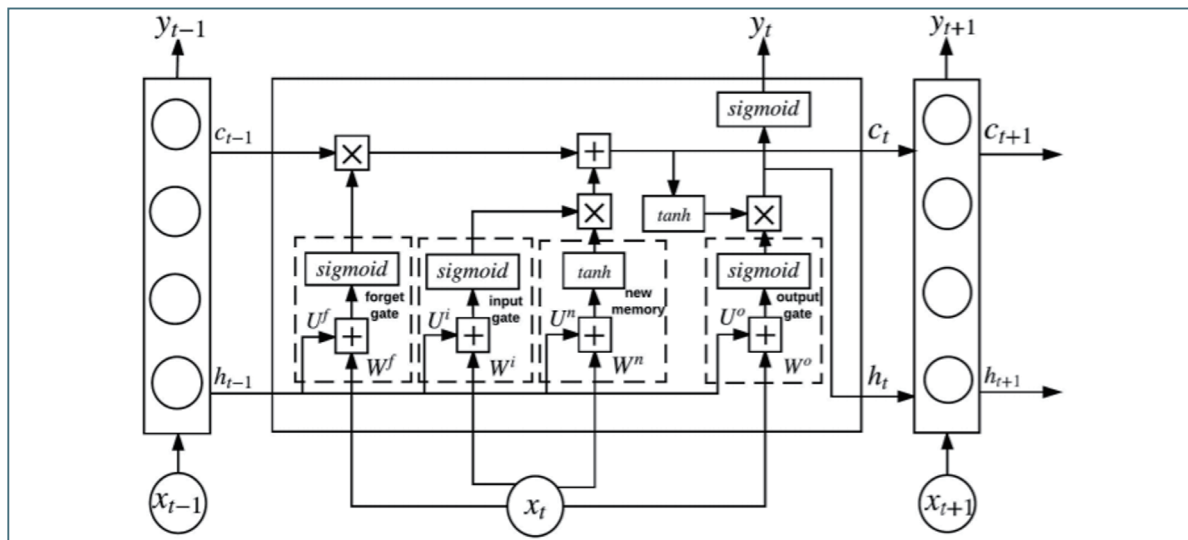
資料來源：本研究繪製

題，後續將以LSTM-RNN架構建立交通流量預測模型。

LSTM為一具有特殊結構的時間遞迴神經網路，其在處理資料期數較長之長期時間序列數據上呈現優異表現，在各領域中應用相當成功。在LSTM結構中設有各種閘道（gate），緩解梯度消失及梯度爆炸等問題，同時能決定輸入及輸出端的資訊是否能被儲存或輸出而LSTM中閘道依序排列為遺忘閘（forget gate）、輸入閘（input gate）、和輸出閘（output gate）其分別具有不同功能（如圖五）。³²

LSTM包含一個存儲單元狀態的單

元，以及三個閘，分別控制該單元狀態輸入和輸出的更新。首先，LSTM單元執行「遺忘閘」以確定從該單元的先前狀態中丟失哪部分訊息；其次，該單元需確定哪些訊息應保留在輸入數據中。輸入閘確定LSTM模塊中的哪些值被更新；第三，tanh激活函數用於創建中間狀態矢量 S_t ，此時具有加速度數據；最後具有LSTM模塊的輸出，然後單元格值將更新。將「遺忘閘」的值與 $t-1$ 時刻的像元值相乘，以獲得最後時刻保留的值。將輸入門的值和 $\tanh(\cdot)$ 函數生成的狀態向量相乘以獲得新值，該新值將添加到該單元中。這兩個值被加到單元格的更新值



圖五 LSTM神經網路架構圖

資料來源: Zhang, Wang, and Liu

32 Lei Zhang, Shuai Wang, Bing Liu(2018), “Deep learning for sentiment analysis: A survey,” Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, Vol. 8, No. 4.

上。最後，LSTM的輸出是通過輸出Gate獲得。輸出僅與當前單元格值相關，且當前單元格值取決於先前的單元格值。³³

數學式運算如公式(1)至(6)，模型以 x_t 表示在時間 t 對儲存單元的輸入，以 c_t 代表單元儲存狀態，以 h_t 代表隱藏層狀態， i_t 、 f_t 、 o_t 代號分別為輸入閘、遺忘閘以及輸出閘的啟動向量。其執行步驟順序為遺忘階段、選擇記憶階段以及輸出階段，遺忘階段主要針對上一個節點的結果進行選擇性遺忘，如公式(1)，以前一次輸出的結果 h_{t-1} 與當前輸入的資料 x_t 相結合，乘上權重值 U^f 及 W^f ，再加上偏差值 b_f ，並使用Sigmoid激活函數，公式中以 $\sigma(\cdot)$ 符號表示，其輸出判於0至1之間，以決定通過的訊息量；選擇記憶階段又分為兩步驟，第一為對當前的輸入做選擇性的記憶，決定將使用哪個值更新，如公式(2)，同樣以前一次輸出的結果 h_{t-1} 與當前輸入的資料 x_t 相結合，並乘上權重值 U^i 及 W^i ，再加上偏差值 b_i ，使用Sigmoid激活函數，以決定要留住訊息的程度，第二為決定是否添加到儲存單元，如公式(3)，使用 $\tanh(\cdot)$ 激活函數，為雙曲線的切線非線性函數，其輸出介於1~-1之間，為一相關係數，說明需加強或減弱的狀態。通過上述兩階段，利用 f_t 與

前一次的狀態 c_{t-1} 相乘，決定前一次保留的訊息，並丟棄不要的資訊，加上 i_t 與 $\tilde{c}_t$ 相乘，需更新的部分，以產出新的儲存單元值 c_t ，如公式(4)；輸出階段則是選擇哪些結果要被輸出，如公式(5)，同樣以前一次輸出的結果 h_{t-1} 與當前輸入的資料 x_t 相結合，並乘上權重值 U^o 及 W^o ，再加上偏差值 b_o ，使用Sigmoid激活函數，以控制訊息的輸出，而最後輸出的結果，如公式(6)，儲存單元的資訊 c_t 會先通過激活函數 $\tanh$ 後，在與 o_t 相乘，以得到最終的結果 h_t ，以做為下一個隱藏層的輸入，而 y_t 即輸出結果的目標函數。

$$f_t = \sigma(W^f x_t + U^f h_{t-1} + b_f) \quad (1)$$

$$i_t = \sigma(W^i x_t + U^i h_{t-1} + b_i) \quad (2)$$

$$\tilde{c}_t = \tanh(W^n x_t + U^n h_{t-1} + b_c) \quad (3)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{c}_t \quad (4)$$

$$o_t = \sigma(W^o x_t + U^o h_{t-1} + b_o) \quad (5)$$

$$h_t = \tanh(C_t) * o_t \quad (6)$$

所謂權重值以及偏差值，是透過訓練過程學習到的參數，而要改進模型的學習方式，則須藉由修改超參數優化模型，係訓練前手動設定的，將依訓練結果的好壞進行調整，而這些超參數在學習過程中有很多方法，主要分為激活函數(Activation Function)、損失函數(Loss Function)、優化器(Optimizers)…等。

33 同註13。

激活函數又稱為啟動函數或轉換函數，主要為一個可微分的非線性函數，神經網路結構後最重要的是透過激活函數，以控制神經元的輸出幅度 (Karlik & Olgac, 2011)。³⁴神經網路透過調整神經元和資料流之間的權重學習。若無轉換，深度學習的效果將和一般神經網路的效果沒有區別。常用的非線性激活函數有Sigmoid、Tanh、ReLU等三種 (Chen, Lee, & Chen, 2019)。³⁵優化器又稱為優化演算法，每個模型最終的目標即透過不斷的訓練建立最佳化模型，類神經網路是由許多的神經元所組成，每個神經元都有自己的權重，神經網路模型即在不斷的批次訓練中，不斷的更新權重與誤差 (bias)，使損失函數的誤差值最小化，並在最終找出誤差值最小的組合，³⁶常見的優化器有包含隨機梯度下降法 (Stochastic Gradient Descent, SGD)、Adagrad (Adaptive gradient descent)、Adadelat、RMSprop、Adam等。由於調整超參數對模型的影響顯著，因此本研究將針對各參數優化進

行探討。

四、模型評估

為量化模型性能，計算出實際值和預測值之間的誤差，許多研究已提出許多計算誤差之方法。本研究選擇以平均絕對百分比誤差 (MAPE) 做為性能評估指標。MAPE是一組預測中誤差的平均大小的絕對值，因其採用百分比形式故易於理解，當MAPE值愈接近0時，表示預測的效果越佳，如公式 (7) 所示。

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - y'_i}{y_i} \right| \quad (7)$$

肆、實驗結果

本研究實驗係利用深度學習建立一個預測模型，分為以下兩小節，第一節針對實驗設計方式及資料蒐集和處理過程做說明；第二節針對模型分析並評估，探討各項超參數對模型優化之影響，找出最適合的模型架構，針對各項超參數實施測試及調整，最後以MAPE為模型評估

34 Bekir Karlik, A. Vehbi Olgac (2011), "Performance analysis of various activation functions in generalized MLP architectures of neural networks," International Journal of Artificial Intelligence and Expert Systems, Vol. 1, No. 4, pp. 111-122.

35 Liang-chu Chen, Chia-Meng Lee, Mu-Yen Chen(2019), "Exploration of social media for sentiment analysis using deep learning," Soft Computing, Vol. 24, No.11, pp. 8187-8197.

36 林大貴，〈TensorFlow+Keras深度學習人工智慧實務應用〉（新北市：博碩文化股份有限公司，西元2017年6月）。

指標，以確認模型準確程度。

一、實驗設計

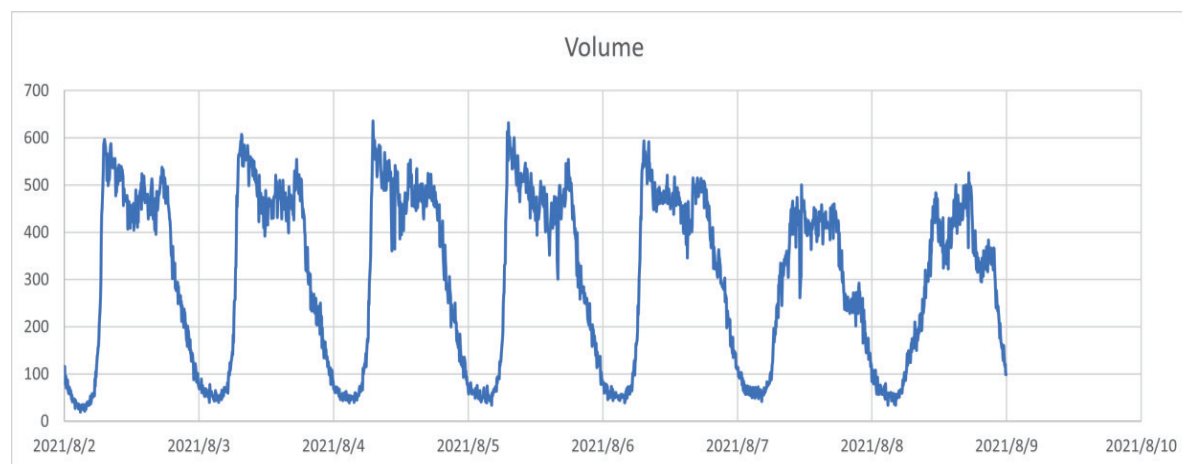
本研究利用高公局ETC資料庫，「交通資料庫(TDCS)」，之eTag_M03A(各類車種通行量統計)內原始資料，內容包含「資料時間」、「偵測站編號」(目前共計三百三十七個ETC門架)、「車行方向」、「車種」(31-小客車、32-小貨車、41-大客車、42-大貨車、5-聯結車)及「交通量」等五類資訊，對象為國道壅塞機率最高的國道1號臺北至圓山(門架編號：01F0233N)偵測器所蒐集之資料做為模型之訓練資料，蒐集2021年7至9月各類車種之交通量(Volume)加總後做為研究範圍，而檢視8月2日至8月9日一周的各車種加總後之車流量統計，可發現前5天為上班日，車流量相似，而後2日為周末假期，車流較平日少，且高峰期出現

之時間也與平日有差異，故確立本研究僅用平日為分析對象(如圖六)。

本研究將上班日資料間隔為每5分鐘乙次，每日288筆，研究期間共64日，總計18,432筆資料，分為前兩個月為訓練集(trained dataset)及後一個月為測試集(test dataset)，據以建立模型，並將模型進行綜合評估，後續調整激活函數及優化器以驗證其影響模型的程度。透過平均絕對百分比誤差(MAPE)評估指標實施評估，將LSTM透過模型調校後之最佳結果，依此評估指標檢視其模型效能。

二、模型分析與評估

本研究將探討激活函數以及優化器的變化是否會導致模型效能改變，使用python 3程式的tensorflow套件撰寫LSTM模型，模型的結構複雜程度及訓



圖六 2021年8月2日至9日車流量變化圖

資料來源：本研究整理

練次數將影響模型的效率及表現，儘管訓練次數愈多可減少誤失，惟同時運算時間也將增加；此外，各種超參數亦影響模型預測，進而產生不同效果，故找到最佳的模型架構是神經網路研究的重要課題。

本研究將模型之隱藏層神經元設定為33，而訓練輪數 (Epochs) 設定為70，預設優化器 (Optimizers) 為Adam、學習率 (Learning Rate) 0.001、損失函數 (Loss function) 為均方誤差 (MSE) 及批次大小 (Batch size) 為64，以找出最適之模型架構。

本研究探討各項超參數對模型優化之影響，找出最適之模型架構，分別針對批次大小 (Batch size)、激活函數 (Activation function)、損失函數 (Loss function)、及優化器 (Optimizers) 進行測試，期能得知最佳組合。

(一) 批次大小

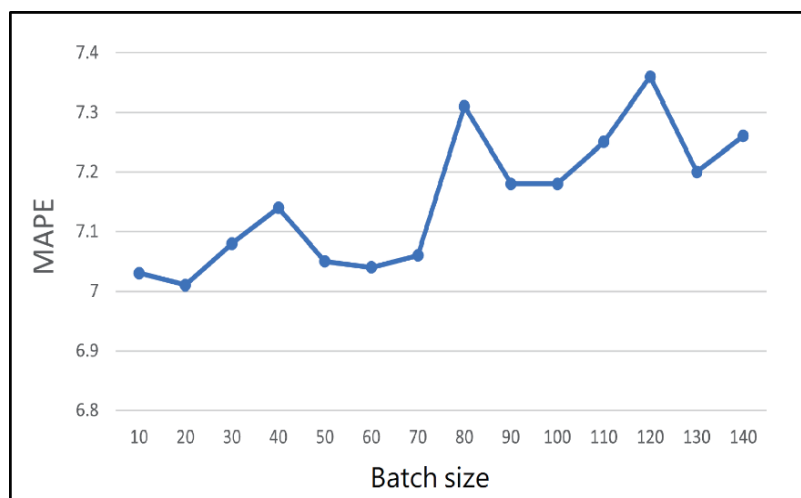
批次大小 (Batch size) 為一次訓練的資料筆數，若Batch size過小，訓練資料將極難收斂，造成預測效果不彰，而增大Batch size相

對處理速度較快，惟所需記憶體容量也將增加，而epoch的次數亦需增加，以達到最好的結果，故正確選擇Batch size是為在內存效率和內存容量之間尋找最佳平衡。

本研究將LSTM模型的Batch size設為10、20、30、……、140之大小，發現Batch size為20時，MAPE值較原Batch size預設值64時7.04進一步降低至7.01；另為確認批量大小 (Batch size) 10-30間是否有更小之MAPE值，發現在Batch size值為16時，MAPE值達到最低之6.96 (如圖七)。

(二) 激活函數

激活函數 (Activation function) 對模型學習、理解非常複雜和非線性的函數而言，具有十分重要的作用。其將非線



圖七 LSTM各批次大小預測結果

資料來源：本研究繪製

性特性引入模型中。若不運用激活函數，則輸出信號將僅是一個簡單的線性函數。

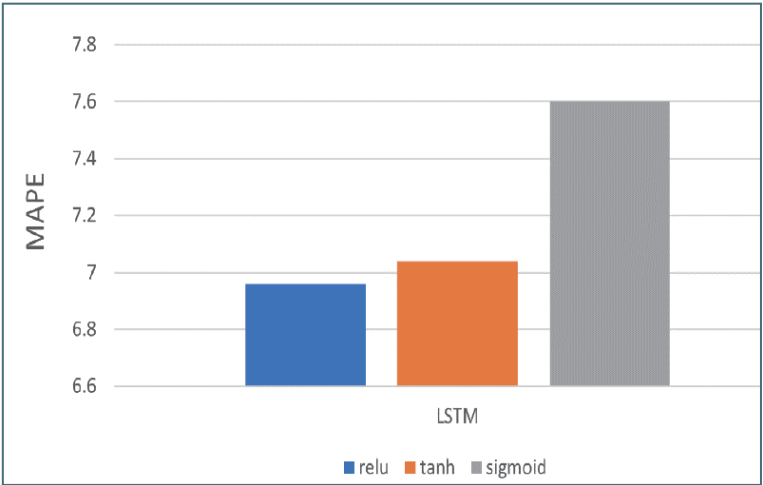
本研究使用Sigmoid、Tanh和Relu等三種神經網路較常使用的激活函數，使用於已最優化Batch size的LSTM模型中，發現選擇Relu為激活函數時MAPE值為6.96，效果較其它兩者更好（如圖八）。

（三）損失函數選擇

損失函數 (loss function) 是用以評估模型的預測值與真實值不一致的程度，也是神經網路中優化的目標函數，模型訓練和優化的過程即最小化損失函數的過程，損失函數越小，則說明模型之預測值越接近真實值。本研究將均方誤差 (MSE)、平均絕對值誤差 (MAE) 等兩種損失函數，使用於優化Batch size及激活函數後的LSTM模型中，發現選擇均方誤差 (MSE) 為損失函數時MAPE值最低，為6.96（如圖九）。

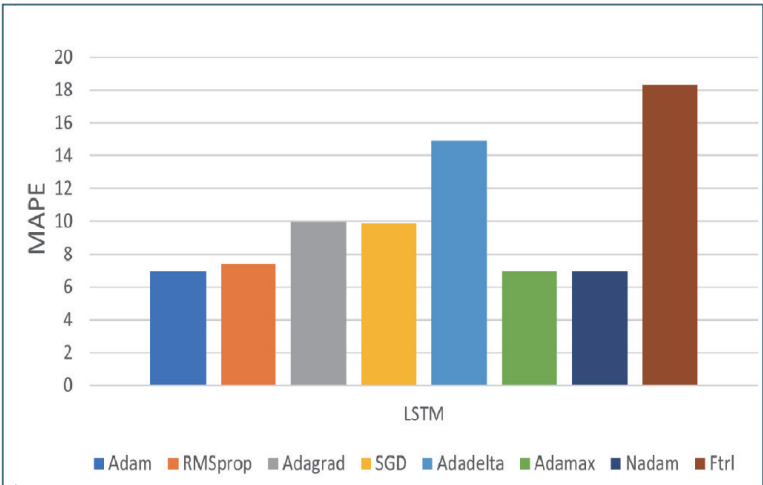
（四）優化器選擇

優化器 (Optimizer) 是幫助神經網路調整參數，目的在使損失函數最小化，期望預測數據的函數能更符合要求；本研究選擇Adam、RMSprop、Adagrad、SGD、Adadelata、Adamax、Nadam、Ftrl等八種優化器（如表一），當優化器設置為



圖八 LSTM使用各激活函數差異比較

資料來源：本研究繪製



圖九 LSTM使用各激活函數差異比較

資料來源：本研究繪製

表一 LSTM使用各優化器參數設定及預測結果

Optimizer	Adam	RMSprop	Adagrad	SGD	Adadelata	Adamax	Nadam	Ftrl
MAPE	6.96	7.41	9.94	9.85	14.9	6.98	6.98	18.32

資料來源：本研究整理

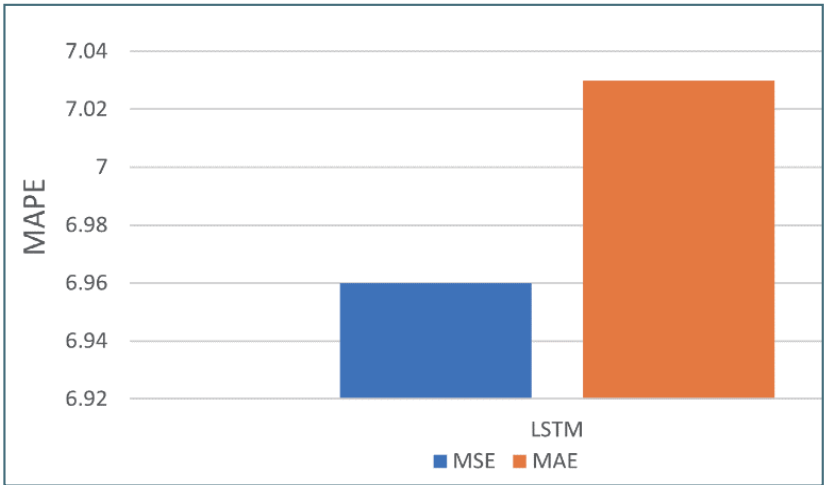
Adam時，發現MAPE值為最低，為6.96，另除Adam外，Adamax及Nadam優化器預測效果亦相當接近，達到6.98之水準，而Adadelata及Ftrl優化器之效果較不理想，故優化器以Adam為最佳（如圖十）。

（五）丟棄率比較

模型容量過多或是訓練資料過少極易造成過擬合，丟棄率（Dropout）是一種對抗過擬合的正則化方法，在訓練時

每一次的Epoch皆以一定的機率丟棄隱藏層神經元，而被丟棄的神經元不會傳遞訊息，故在訓練時不會過度依賴某些神經元，藉此達到對抗過擬合的效果，本研究將丟棄率設定為0.0001至0.01之間，觀察並使用於模型中，當丟棄率設置為0.0001時，發現無論使用何種模型，MAPE值均為最低，分別為7及6.96，研究發現若丟棄率值越小時甚至未使用

丟棄率之模型，預測效果更佳，推測原因係丟棄率為避免模型因變數較多而過度配適，導致模型效果不佳，而本研究資料集變數僅車流量一項，若Dropout又將模型內神經元傳遞之特徵丟棄，反造成學習效果下降（如表二）。



圖十 LSTM使用優化器差異比較圖

資料來源：本研究繪製

表二 LSTM及Bi-LSTM使用各丟棄率參數設定及預測結果（MAPE）

Dropout	0.0001	0.001	0.01	0.1
LSTM	7	7.01	7.04	7.17

資料來源：本研究整理

(六) 模型評估

本研究將數據集運用LSTM演算法，將各種超參數優化後所找出最適之模型架構，以預測5分鐘後之車流量，預測結果MAPE值為6.96，依Lewis於1982年所提出之MAPE評估指標區間標準，MAPE小於10之預測為「高準確的預測」，³⁷顯示此演算法所建立之模型適用於交通流量預測等時間序列問題方面之研究，可達到即時且精準的交通預測（如表三）。

伍、結論與建議

一、結論

本研究為有效且迅速地預測國道經常壅塞路段之車流量，以深度學習方法建立模型，並調整各項超參數使其優化，提升預測效能，並在各項超參數調整及優化過程發現，優化器之選擇對模型準確度有較高之影響，透過遞歸神經網路的LSTM演算法為基礎之深度學習模型，能提供國道壅塞路段短期交通流量準確且即時之預測。若將本模型套用至各單位出勤路線之路段，可提供國軍單位派車及路線選擇決策，進而降低行車風險及提高效率；另本研究以高公局交通資料庫針對國道平日壅塞情形較為嚴重之

表三 模型架構及超參數組合

輸入大小(訓練集)	[12672,40,1]
隱藏層輸出維度(units)	33
全連接層輸出維度	1
訓練輪數(epochs)	70
輸出大小(測試集)	[5760,1]
激活函數	Relu
優化器	Adam
損失函數	MSE
批次大小(batch size)	16
丟棄率(dropout)	0.0001

資料來源：本研究整理

37 Lewis Colin David, Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting(London Boston: Butterworth Scientific, 1982).

臺北至圓山路段，做為本研究預測車流量之對象，實驗結果利用LSTM演算法建構車流量預測模型，並利用模型優化提升其效能，所建構之模型能準確預測後續車流量。此外，研究中亦曾嘗試以其它路段或時間，均可達成模型適配，故可推斷本研究建立之模型及最佳化之架構可適用於高速公路各路段及時間之車流量預測。

二、建議

國軍運勤任務往往受限於任務期限及時間，較難避開壅塞時段，若遇交通壅塞，除造成無法在時限內完成任務，甚或肇生行車安全問題，致使無法如期如質完成任務。交通流量資料除高速公路ETC資料庫外，臺灣主要路段亦設有車流偵測裝置，若將派遣系統連線至相關的交通資料庫，如交通部公路局之公路局省道交通控制系統及各縣市政府，並於資料平臺公布各重要路段偵測之車輛偵測器資料，藉由先期預測車流量的方法，準確預估運勤任務執行所需之時間，並適當避開易壅塞之路段，有利於出勤前完成良好的出發時間及路線之規劃，故若能透過快速且準確的模型，選擇該時段最佳之路線，並將本模型套用至各單位出勤路線之路段，可提供國軍單位派車及路線選擇決策，進而降低行車風

險及提高效率。

另針對模型架構及超參數各種組合實施探討，可供爾後時間序列預測方面相關研究模型設計架構之參考，甚至除短期車流量，可進一步拉長預測時間至中期或長期。在實務方面短期車流量預測及透過分析數據及車流規則性後，因公路為軍民共用，於平時可透過交管行控之手段，適當調節車流量，減少壅塞情形、提升行車品質並增進交通順暢程度；於戰時根據教則，由聯合運輸管制組實施公路調節，針對易壅塞路段，協調交通部公路總局、高公局及縣市政府交通局，結合交控中心可變標誌和匝道儀控等手段，更精準控制車流，適當調節車流量，在時限內達成運輸（補）任務。鑑此，訂定有效的運輸管理策略，將是後續國軍派車管理之目標，故精準且可靠的交通流量預測結果係至關重要之課題。

作者簡介

張軒慈中校，國防大學管理學院運籌管理學系碩士班畢業，國防大學管理學院採購正規班102年班。曾任排長、隊長、教官，現為陸軍後勤訓練中心學員生二大隊大隊長。