

應用混合深度學習模式於高速公路多時段 交通速率預測¹

APPLICATION OF HYBRID DEEP LEARNING MODELS IN FREEWAY MULTI-PERIOD TRAFFIC SPEEDS PREDICTION

吳姿樺 Zih-Hua Wu²
蘇昭銘 Jau-Ming Su³
林良泰 Liang-Tay Lin⁴
張建彥 Chien-Yen Chang⁵
洪百賢 Pai-Hsien Hung⁶
吳東凌 Tong-Ling Wu⁷
何毓芬 Yu-Fen Ho⁸

(111 年 6 月 17 日收件，111 年 12 月 26 日第一次修正，112 年 2 月 7 日第二次修正，
112 年 2 月 8 日接受)

摘 要

過去傳統依靠經驗法則的判斷方式，常因錯誤的判斷導致不僅無法有效紓解交通問題，嚴重者恐使問題加劇，因此更突顯預測技術與工具掌握

-
1. 本研究係交通部運輸研究所研究計畫 (MOTC-IOT-110-IFB001、MOTC-IOT-111-IFB010) 之部分成果。
 2. 逢甲大學建設規劃與工程博士學位學程博士研究生(聯絡地址：40724 臺中市西屯區文華路 100 號，電話：04-24517250 分機 4535，Email：ronniewu@iciti.com)。
 3. 逢甲大學運輸與物流學系教授。
 4. 逢甲大學運輸與物流學系教授。
 5. 逢甲大學建設學院專任副教授。
 6. 逢甲大學運輸與物流學系助理教授。
 7. 交通部運輸研究所運輸科技及資訊組組長。
 8. 交通部運輸研究所運輸科技及資訊組高級分析師。

的重要性。本研究認為在進行全時段交通預測時，除考量全時段整體預測的準確性外，亦須能夠同時考量尖峰時段預測的準確性。因此提出一新穎之混合模型，結合長短期記憶網路 (LSTM)、注意力機制及雙向長短期記憶網路 (Bi-LSTM)，分別針對 6 種情境，包含全日全時段、全日尖峰時段、工作日全時段、工作日尖峰時段、週末全時段、週末尖峰時段進行交通速率預測。該模式有效地透過 LSTM 記憶及遺忘的特性，將交通流可用的資訊保留下來後，再利用 Attention 機制賦予各個重要資訊權重，最後利用 Bi-LSTM 綜合考量兩個方向（前向、後向）的交通流序列資訊進行預測。結果顯示本研究使用之混合模型，不論於不同日態或不同時段，其預測結果皆比其他模型來得可靠。並且可明顯改善尖峰時段之交通速率預測準確性。

關鍵詞： 交通預測、多時段預測、長短期記憶網路、注意力機制、雙向長短期記憶網路

ABSTRACT

Traffic speed prediction plays a crucial role in traffic congestion mitigation. With the advent of deep learning, transportation researchers are empowered to forecast traffic with an unprecedented accuracy. However, in the context of short-term prediction, such accuracy fails to hold consistent at hourly level; In particular, at peak hour periods. To bridge the gap, this study proposes a novel hybrid model that integrates LSTM, Attention Mechanism, and Bi-LSTM to for traffic prediction in six scenarios: "One week-24 hours", "One week-Peak hours", "Weekday-24 hours", "Weekday - Peak hours", "Weekend-24 hours", "Weekend-Peak hours". By leveraging the strengths of each constituent model, the proposed model demonstrates its capability of capturing peak hour traffic patterns as compared to previous models.

Key Words : *Traffic Prediction ; Multi-period prediction ; LSTM ; Attention Mechanism ; Bi-LSTM*

一、前言

在資訊技術快速發展的時代下，過去交通運輸領域政府投入大量經費佈設先進技術、儀器，目的為蒐集更完善的交通流 (Traffic Stream) 資料進行交通研究，希望能未雨綢繆反應交通狀況。而「預測」為交通管理系統 (Traffic Management System, TMS) 中重要的一部分，亦是路線規劃、動態交通控制與資訊提供之基礎 (Han 等人^[1])。交通預測起源可追溯至 1980 年代，Okutani 與 Stephanede^[2] 使用卡曼濾波器 (Kalman Filtering) 針對日

本名古屋街道進行短期交通流量預測。而交通預測研究以交通流的基本參數，如速率、流量和密度，作為監控交通狀況指標，且根據預測時間長度的不同，一般分為短期（5-30 分鐘內）、中長期（30 分鐘後）兩個尺度。交通管理單位則依據各自不同的策略目標，來選擇要進行短期預測或長期預測。過去相關研究（Cheng 等人^[3]；Dunne 與 Ghosh^[4]；Xia 等人^[5]；Sun 等人^[6]）曾嘗試以各種角度去解析交通預測問題，包含時間序列分析、迴歸與近似分析、聚類分析、圖形識別分析或是利用上述兩種所構成的混合式分析，使用傳統統計方法、機器學習方法對於交通預測皆有不錯成效。

然而，交通流的不確定性及高度複雜性使這些預測方法未必能精確反應真實的交通狀態，例如 Fang 等人^[7]指出當交通數據資料在車流穩定狀態時預測表現良好，而在數據資料波動頻繁的部分預測表現則不佳。另由 Sun 等人^[8]與 Almeida 等人^[9]的研究中亦可看出，圖 1 及圖 2 中大部分的預測值是非常接近真實值的。但在一些局部最大值或最小值處（圖 1 及圖 2 紅框處），則可明顯看到在進行交通預測時所產生的誤差。圖中這些產生局部最大值或最小值之處，則代表於交通尖峰時段時所產生的低車速或高車流量。

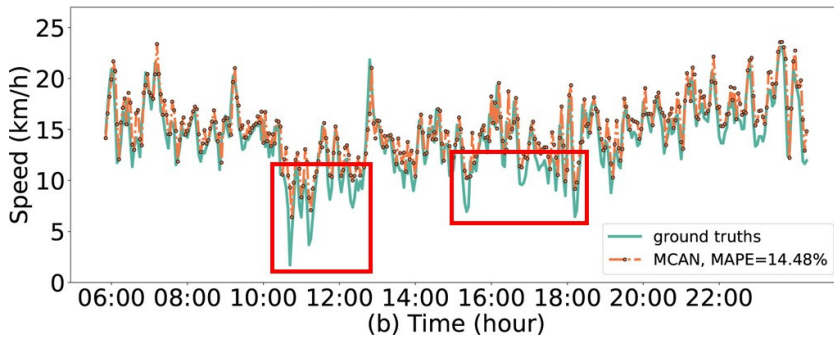


圖 1 交通速率於尖峰時段之預測誤差^[8]

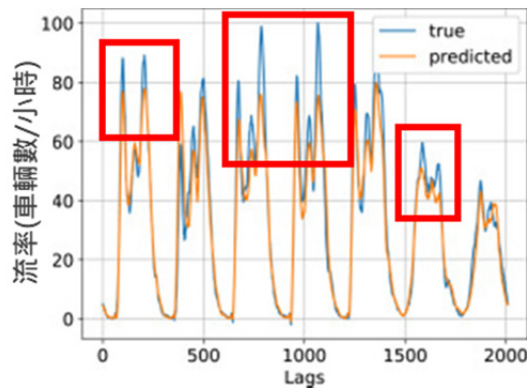


圖 2 交通流量於尖峰時段之預測誤差^[9]

換言之，當需要進行交通流預測時，在尖峰時段的預測效果往往不如預期。且多數的交通流預測相關研究中，評估預測結果良窳的方式，多以全時段的整體預測績效值（如 Cui 等人^[10]；Zheng 等人^[11]）作為判斷之依據。如圖 1 之案例是以每 3 分鐘為一個時間單位進行交通速率預測，預測未來 06:00~22:00 之交通速率（共預測出 320 筆速率資料），而該研究顯示這 320 筆的預測值與真實值的平均絕對值誤差百分比（Mean Absolute Percentage Error, MAPE）為 14.48%。而圖 2 的案例亦是以每 5 分鐘為一個時間單位進行交通流量預測，預測未來一週的交通流量（共預測出 2,016 筆流量資料），而該研究顯示這 2,016 筆的預測值與真實值的整體均方誤差（Mean Square Error, MSE）為 44.857 ± 0.752 。

惟前述兩篇研究中皆無談到，這 320 筆及 2,016 筆資料中，分別有多少筆是屬於尖峰時段之預測值，而這些尖峰時段之預測誤差為多少？亦即若單看整體的誤差值，表面上預測效果看似良好，若是僅針對這些尖峰時期的預測值計算誤差值，由圖 1 及圖 2 中的趨勢看來，其結果可能會不盡理想。

有鑑於此，本研究認為在進行全時段交通預測時，除考量全時段整體預測的準確性外，亦須能夠同時考量尖峰時段預測的準確性。而為了提升交通流預測中的尖峰時段預測效果，本研究提出一種堆疊式深度學習混合模型，結合了長短期記憶模式（Long Short-Term Memory, LSTM）、注意力機制（Attention Mechanism）及雙向長短期記憶模式（Bi-directional LSTM, Bi-LSTM），藉此增強預測模型對於不同時段交通流時空特性的學習力，以提升預測精度。

本研究針對高速公路之交通速率進行預測，而預測之情境分為 3 種日態：全日（不區分工作日及週末）、工作日（星期一至星期五）、週末日（星期六至星期日），且每種日態再分為 2 種時段：全時段（以 24 小時之交通速率資料樣本進行模式構建）、尖峰時段（僅以尖峰時段之交通速率資料樣本進行模式構建），共計 6 種情境，並同時與其他模型進行預測性能比較。期能以此建立較為精確的預測模式，不僅作為交通主管單位提供交通資訊播報之基礎，並能據以擬定更精準之交通管理策略。

本篇研究章節架構如下。第二章首先針對交通流預測之趨勢進行回顧，接著探討過去研究者對於進行交通流預測時所採納之時段情境。第三章簡要說明本研究使用的混合模型架構。第四章則針對本研究選擇的案例地點進行廣泛的實驗並且提出實驗結果。最後綜合前述於第五章提出研究結論與建議。

二、文獻回顧

2.1 交通預測之發展趨勢

近年來，多以數據驅動方法進行交通相關參數預測，亦即藉由感測器蒐集大量交通數據以作為交通參數預測之基礎。惟感測器佈設不足導致無法蒐集到資料，是其面臨的困難

之一。因此 Gentili 與 Mirchandani^[12] 研究如何將感測器的安裝數量與安裝位置進行最佳化藉以在有限資源限制下取得更精準的資料。而研究者能依照交通管理不同的應用需求，對感測器之資料蒐集方式設定不同程度的資料蒐集頻率，例如以時間間隔每 5 分鐘或 10 分鐘為一個單位進行資料蒐集，因此資料蒐集頻率又稱為資料解析度。資料解析度越高，資料雜訊 (Noise) 亦會隨之提高，因此許多針對交通預測前，處理時間序列資料雜訊的方法因應而生。Kashi 與 Akbarzadeh^[13] 提出一種結合小波轉換和 ANN 的混合模型 (WT-ANN)，針對交通流量資料進行降噪 (Denoising) 處理後再進行預測模式建構。研究結果顯示，經過降噪處理後的預測準確性較高。Tang 等人^[14] 則提出 5 種降噪方法，分別與支持向量機 (SVM) 模型結合進行交通流量預測。結果顯示降噪方法中的總體經驗模態分解法 (Ensemble Empirical Mode Decomposition, EEMD) 與 SVM 結合後的預測效果亦優於沒有結合降噪法的模型。

處理完高解析度資料後，如何操作資料聚合 (Data Aggregation) 是資料分析的關鍵步驟。Suvin 等人^[15] 研究顯示，與交通參數相關的特徵，其時間聚合 (Temporal Aggregation) 程度，會對之後使用的數學模型所產出之績效有顯著影響。因此在交通預測的領域中，預測方法的有效性取決於研究者對於資料聚合程度的選擇，且合適的資料聚合程度能夠盡可能的消除不必要的資料雜訊。在空間聚合 (Spatial Aggregation) 的部分，Cassdei 等人^[16] 研究亦表明，使用聚合技術能夠降低大規模交通路網的複雜性，該研究將交通速度相似的相鄰路段進行聚合，能夠在保留重要資訊的同時，大幅減小路網的複雜性，且對於後續進行分析時產生之誤差，仍可維持在 25% 以內。另一方面，資料品質亦是進行交通預測的挑戰之一，由於存在於資料集中的缺失值會影響預測的準確度，研究者必須選用適當的資料填補策略來克服。過去研究顯示，將資料的缺失值進行填補後再用於交通預測，能夠有效地提升預測準確度。而資料遺失值填補的方法包含圖形馬可夫網路法 (Graph Markov network, GMN) (Cui 等人^[10])、模糊神經網路 (Fuzzy Neuro Network, FNN) (Zhang 等人^[17])、灰時間序列模式 (Grey Time Series) 及灰關聯最近鄰法 (Grey-Relational-Based Nearest Neighbor Method) (Huang 與 Lee^[18])、貝葉斯張量分解法 (Bayesian tensor decomposition) (Chen 等人^[19]) 等。

隨著資料蒐集的新興技術發展，交通資料量與日俱增，科技運算能力亦逐步提升。許多研究者提出了各式各樣交通流預測的方法 (Lippi 等人^[20])，而交通流預測方法已從傳統的統計模式逐漸轉變為以資料驅動的機器學習方法 (Vlahogianni 等人^[21])。傳統統計方法主要多為參數方法，包含線性迴歸 (高雄市政府交通局^[22])、ARIMA (Williams 與 Hoel^[23])、卡曼濾波器 (Guo 等人^[24]) 以及其他時間序列 (Time series) 模式等 (Ghosh 等人^[25])。而由於人工智慧的興起，機器學習方法也由淺層學習 (Shallow Learning) 模式，如 KNN (Yu 等人^[26]; Cai 等人^[27])、SVM (Wu 等人^[28]; Liu 與 Yao^[29])、隨機森林 (Leshem 與 Ritov^[30]) 等等，逐步發展為深度學習 (Deep Learning) 模式，也因此再次刺激了交通預測研究領域的發展。

許多深度神經網路模型，如循環神經網路 (Recurrent Neural Network, RNN) (Lai 等人

^[31])、卷積神經網路 (Convolutional Neural Networks, CNN) (Almeida 等人 ^[9]; 黃冠維 ^[32]; Zhao 等人 ^[33]; Benkraouda 等人 ^[34]) 或堆疊自動編碼器 (Stacked Auto Encoders, SAE) (Jin 等人 ^[35])、長短期記憶 (Zhao 等人 ^[36]; Ma 等人 ^[37]) 等方法, 已廣泛應用於交通預測研究中。惟現今受到各種外部因素影響, 使得道路交通環境更趨複雜, 更難以掌握交通流之趨勢, 研究者開始使用混合式的神經網路模型, 如 Wu 等人 ^[38] 提出的結合 GRU、Attention Mechanism 及 CNN 模式的 DNN-BTF 混合模型; Yang 等人 ^[39] 提出的 ARIMA-SA-BPNN 模型; Bogaerts 等人 ^[40] 提出的結合 Graph CNN、LSTM 及 FFNN 模式的 Graph CNN-LSTM 混合模型, 皆是透過綜合各模型之優點以提升交通流預測之準確性。

2.2 交通預測之時段特性

交通流雖具有一定之週期性, 但隨不同地區社經發展之差異, 使得交通流之週期性變化並不是絕對的規律, 因此難以完全掌握其趨勢。Zhang 與 Wang ^[41] 交通狀態一般可分為尖峰期與非尖峰期, 而在工作日期間, 一天通常會有兩個尖峰。且 Peng 等人 ^[42] 表示, 即便是工作日, 每一天的交通流高峰也會有明顯的差異, 週末或連續假期期間更是難以掌握其尖峰交通量變化的差異。

本研究以預測績效誤差為基礎, 分為 6 種時段情境, 包含全日 (不區分工作日及週末) 全時段 (以 24 小時之交通速率資料樣本進行模式構建)、全日尖峰時段 (僅以尖峰時段之交通速率資料樣本進行模式構建)、工作日 (交通速率資料樣本僅採用星期一至星期五) 全時段、工作日尖峰時段、週末 (交通速率資料樣本僅採用星期六至星期日) 全時段、週末尖峰時段, 進行交通預測相關文獻蒐集, 並彙整如表 1 所示。由表中發現多數研究者, 主要僅針對單一情境進行交通預測研究。且單一情境中又以「全日全時段」情境為最多 (Lai 等人 ^[31]; Ma 等人 ^[37]; Wu 等人 ^[38]; Bogaerts 等人 ^[40]; 吳俊緯 ^[43]; 張軒慈 ^[44]; Xing 與 Liu ^[45]; Guo 等人 ^[46]; Ma 等人 ^[47]), 其次為「工作日尖峰時段」(Do 等人 ^[48]; Qi 及 Ishak ^[49]; Qi 及 Ishak ^[50]; Chan 等人 ^[51]; Hong ^[52])。而少數研究者 (Almeida 等人 ^[9]; Lippi 等人 ^[20]; Ma 等人 ^[53]; Shen 等人 ^[54]; Tang 等人 ^[55]; Du 等人 ^[56]; 周諺鴻等人 ^[57]) 有同時考量兩種時段情境進行預測。此外, 幾乎沒有研究者有針對「週末尖峰時段」進行交通預測。

迄今為止, 無論於工作日、週末, 或是不區分日態的全日交通預測中, 能夠同時考量全時段及尖峰時段預測性能之研究鮮少受到關注。為了提升全時段交通流預測中的尖峰時段預測效果, 本研究參考許多過去的研究, 發現許多研究者會於預測模型中加入 attention 機制以增強模型對於時間或空間相關特徵之學習力, 能夠有效提升預測準確性。如 Guo 等人 ^[46]、Do 等人 ^[48]、Du 等人 ^[56]、徐志榮 ^[59]、Yang 等人 ^[60] 分別提出的 Graph Attention Temporal Convolutional Networks (GATCN)、Spatial-Temporal Attention based Neural Network (STANN)、Sequence-to-Sequence Spatial-Temporal Attention (STSSTA)、Attention-Residual-LSTM、LSTM-Attention 等模型。

表 1 六種交通預測時段情境相關文獻彙整

作者 (按年份排序)	全日		工作日		週末		預測目標	預測模型	文獻編號
	24H	PH	24H	PH	24H	PH			
吳俊緯 (2022.08)	V						捷運運量	H-LSTM	[43]
張軒慈 (2022.06)	V						車流量	BiLSTM	[44]
Almeida et al. (2022.03)			V		V		車流量	CNN	[9]
Xing 及 Liu (2021.07)	V						車流量	DFBD-BiLSTM	[45]
Ma et al. (2021.05)			V		V		車流量	CNN LSTM	[53]
Ma et al. (2021.02)	V						車流量	LSTM_BiLSTM	[47]
Shen et al. (2020.12)			V		V		車流量	EMD-ARIMA	[54]
Guo et al. (2020.10)	V						車速率	GATCN	[46]
Du et al. (2020.06)	V		V				車流量	STATF	[56]
Bogaerts et al. (2020.03)	V						車速率	Graph CNN-LSTM	[40]
Do et al. (2019.11)				V			車流量	STANN	[48]
Huang et al. (2019.07)		V					公車旅行時間	LSTM	[58]
徐志榮 (2019.07)		V					計程車搭乘需求	AR-LSTM	[59]
Yang et al. (2019.03)			V				車流量	LSTM-Attention	[60]
周諺鴻等人 (2018.10)			V		V		車流量、速率	Gradient Descent	[57]
Wu et al. (2018.05)	V						車流量	DNN-BTF	[38]
Tang et al. (2017.09)			V		V		車速率	EFNN	[55]
Lai et al. (2016.02)	V						車速率	BPNN	[31]

作者 (按年份排序)	全日		工作日		週末		預測目標	預測模型	文 獻 編 號
	24H	PH	24H	PH	24H	PH			
Tian 及 Pan (2015.12)			V				車流量	LSTM RNN	[61]
Ma et al. (2015.05)	V						車速率	LSTM NN	[37]
Qi 及 Ishak (2014.06)				V			車流量	HMM	[49]
Qi 及 Ishak (2013.06)				V			車速率	Stochastic Approach	[50]
Lippi et al. (2013.06)	V	V					車流量	SARIMA _{Kal}	[20]
Chan et al. (2012.06)				V			車流量	NN-EXP-LM	[51]
Hong (2011.06)				V			車流量	SSVRCSA	[52]
本研究	V	V	V	V	V	V	車速率	LSTM-Attention-BiLSTM	-

註：24H 表示 24 小時；PH 表示尖峰小時 (Peak Hour)；H-LSTM: Hybrid Long Short-Term Memory；CNN: Convolutional Neural Network；DFBD-BiLSTM: Data Fusion powered Bi-Directional Long Short Term Memory；EMD-ARIMA: Empirical Mode Decomposition Auto Regressive Integrated Moving Average；GATCN: Graph Attention Temporal Convolutional Networks；STATF: Spatial-Temporal Attention Traffic Forecasting；STANN: Spatial and Temporal Attentions Neural Network；AR-LSTM: Attention、Recurrent Neural Network、LSTM；DNN-BTF: Deep neural network Based Traffic Flow；EFNN: Evolving Fuzzy Neural Network；BPNN: Back Propagation Neural Network；HMM: Hidden Markov Models；SARIMAKal: Seasonal Auto Regressive Integrated Moving Average with Kalman filter；NN-EXP-LM: Neural Network Exponential Smoothing and Levenberg - Marquardt；SSVRCSA: Seasonal SVR(Support Vector Regression) with Chaotic Simulated Annealing。

2.3 小結

由回顧過去文獻可知交通流參數資料具有高度非線性、動態和隨機之特性，過去傳統統計方法較無法有效捕捉到上述特性，亦即時間和空間的傳遞特徵，因此其預測性能往往不甚理想。時至今日，許多研究將目光集中在如何利用機器學習方法解決複雜的交通問題。在交通預測的背景下，會根據不同的交通管理策略，來決定要進行短期或是長期交通預測。若是要掌握交通流之變化趨勢，可以選擇長期預測，若是要掌握交通流異常狀態藉以擬定疏解策略，則是選擇短期預測。

綜合前述發現：(1) 多數研究使用混合模型之預測結果較單一模型來得可靠；(2) LSTM、Bi-LSTM 與注意力機制之混合模式鮮少應用於交通流預測領域中；(3) 鮮少研究

於全時段交通預測過程中能夠同時考量尖峰時段預測之性能。基於此，本研究改良傳統 LSTM 模式，提出考量時空特性之 LSTM-Attention-BiLSTM 模型，並針對上述 6 種時段情境（全日全時段、全日尖峰時段、工作日全時段、工作日尖峰時段、週末全時段、週末尖峰時段）進行高速公路交通速率預測，並同時與其他基準模型進行預測性能比較。

三、交通預測模式架構

本研究提出的模式架構主要分為 5 個部分，分別為輸入層 (Input Layer)、LSTM 層 (LSTM Layer)、注意力層 (Attention Layer)、Bi-LSTM 層 (Bi-LSTM Layer) 及全連接與輸出層 (Full Connected & Output Layer)，前述簡易架構及各部分功能如圖 3 所示。交通資料輸入後先透過 LSTM 將可用的資訊保留下來後，接著利用注意力機制將前述保留下來的資訊賦予權重，最後利用兩層雙向 LSTM 綜合考量兩個方向(前向：交通序列資訊由開始到結束、反向：交通序列資訊由結束到開始)的交通流量序列資訊進行預測，以全連接層提取前述兩次預測間之關聯後輸出最終預測結果。模式詳細架構如圖 4 所示，本節將依序詳細說明各部分的組成。

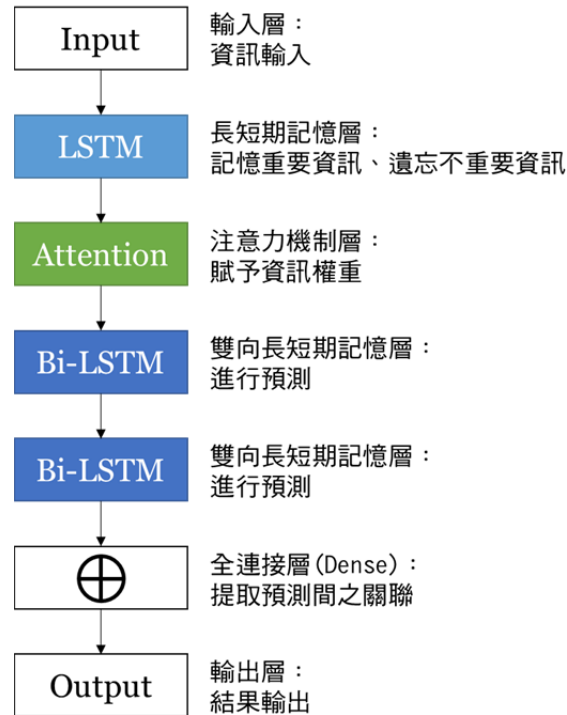


圖 3 本研究交通速率預測模式簡易架構與各層功能

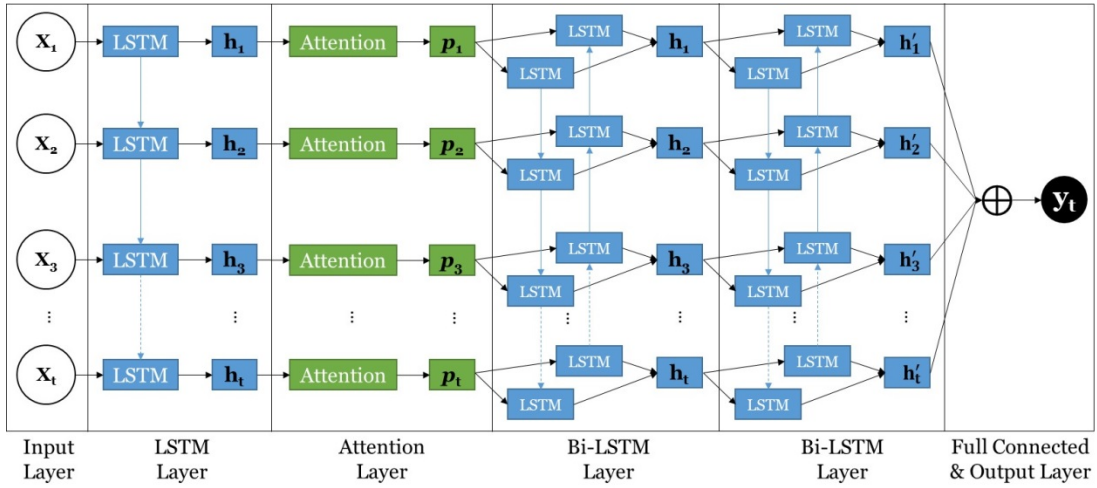


圖 4 本研究交通速率預測模式詳細架構

1. 輸入層(Input Layer)

考量交通流具強烈之週期性，且會受時空特性之影響。週期性是由於工作日和非工作日、尖峰和離峰而導致交通流所產生之規律性，而時空特性中之時間特性係指前一時刻的交通流會對下一時刻的交通流產生影響，空間特性係指相鄰路段的交通流間亦會相互影響，許多研究著力於整合上下游車流資訊以提升預測準確度 (Chandra 與 Al-Deek^[62]；Kamarianakis 等人^[63])。基於此，本研究將週期性特徵 (星期別、小時、分鐘) 及時間特性特徵 (目標路段 t 時刻平均速率、前一刻 ($t-1$) 平均速率) 以及空間特性特徵 (上游路段前一刻 ($t-1$) 速率) 作為時間序列輸入。輸入集合表示如下：

$$\begin{aligned}
 X = & \{(S_1(s, t), S_2(s, t), S_3(s, t), \dots, S_n(s, t)), (B_1(s, t-1), B_2(s, t-1), B_3(s, t-1), \dots, B_n(s, t-1)), \\
 & (U_1(s-1, t-1), U_2(s-1, t-1), U_3(s-1, t-1), \dots, U_n(s-1, t-1)), (W_1, W_2, W_3, \dots, W_n), \\
 & (H_1, H_2, H_3, \dots, H_n) | S, B, U \in \mathbb{Q}\}
 \end{aligned} \quad (1)$$

式中， S 代表 s 路段 t 時刻之平均速率； B 為 s 路段前一刻 ($t-1$) 平均速率； U 為上游路段 ($s-1$) 前一刻 ($t-1$) 平均速率； W 為星期別； H 為小時； n 為速率樣本編號。

2. LSTM 層(Long Short-Term Memory Layer)

LSTM 首次發表於 1997 年，適合於處理和預測時間序列中間隔和延遲較長之重要資訊 (Hochreiter 與 Schmidhuber^[64])。而交通流序列即是一種時間序列資料，前後輸入的交通速度資料是互相關聯的，在 LSTM 神經元的輸出會在下一個時間傳遞給自己，在前後相互關聯的時間序列預測中具有優勢，且相較傳統的 RNN 神經網路，可以有效解決梯度爆

炸和梯度消失問題。典型 LSTM 單元主要由輸入閥、遺忘閥、記憶細胞、輸出閥構成，其網路結構如圖 5 所示。在本研究模型架構中，LSTM 計算過程表示如下：

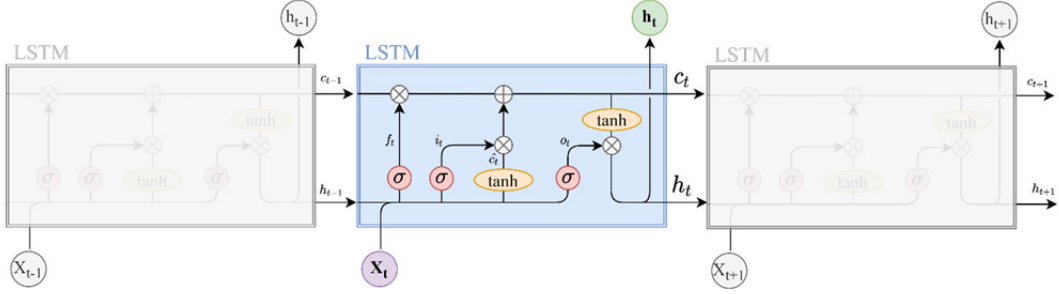


圖 5 LSTM 網路結構

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t, c_{t-1}] + b_i) \quad (2)$$

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t, c_{t-1}] + b_f) \quad (3)$$

$$\hat{c}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \hat{c}_t \quad (5)$$

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t, c_{t-1}] + b_o) \quad (6)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (7)$$

式中， i_t 為輸入閥； f_t 為遺忘閥； o_t 為輸出閥； h_t 為隱藏層在 t 時刻的輸出向量； $\hat{c}_t$ 為記憶單元； c_t 為記憶細胞在 t 時刻的輸出向量； w_i 、 w_f 、 w_c 、 w_o 、 b_i 、 b_f 、 b_c 、 b_o 各為輸入閥、遺忘閥、記憶細胞、輸出與隱藏層、 x_t 、細胞單元的權重及偏移量； σ (Sigmoid function) 和 $\tanh$ (hyperbolic tangent function) 皆為激活函數 (Activation function)； $\cdot$ 為向量內積。

3. 注意力機制層(Attention Mechanism Layer)

注意力機制最早是應用於圖像辨識領域，於 2014 年由 Google DeepMind 團隊所提出 (Mnih 等人^[65])。同年，Bahdanau 等人^[66] 便將注意力機制應用於時間序列中之自然語言 (NLP) 領域，並取得不錯之成效，自此開啟了注意力機制於時間序列預測之發展。

將上一階段 LSTM 之輸出向量輸入注意力層，藉由注意力機制針對重要資訊則賦予較

高權重，無關資訊則賦予較低權重。在本研究模型架構中，注意力機制層之計算過程表示如下：

$$e_t = \tanh(w_e v_t + b_e) \quad (8)$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{t=1}^n \exp(e_t)} \quad (9)$$

$$p_t = \sum_{i=1}^n \alpha_i v_i \quad (10)$$

式中： e_t 為 t 時刻輸入交通速率序列 (v_t) 的注意力概率分佈值 (Attention Distribution)； w_e 和 b_e 為權重值和偏移量，接著透過 softmax 函數將 e_t 值進行歸一化 (Normalization) 至 $[0, 1]$ 區間後得到 α_t ，再計算其權值得到注意力機制之輸出向量 (p_t)。

4. 雙向長短期記憶層(Bidirectional LSTM Layer)

將 Attention 層之輸出結果，輸入到兩個 Bi-LSTM 網路中，Bi-LSTM 是由前向(Forward)和後向 (Backward) 兩個單向 LSTM 組合而成，其中前向 LSTM 表示輸入從開始到結束的前向交通速率序列來進行預測，後向 LSTM 表示輸入從結束到開始的後向交通速率序列來進行預測。綜合考量兩個方向的交通速率序列資訊，能提升模型預測的精準度。

在本研究模型架構中，Bi-LSTM 的輸出計算過程表示如下：

$$h'_t = w_h [h_t^f, h_t^b] + b_h \quad (11)$$

式中： h_t^f 為前向 LSTM 網路的輸出向量； h_t^b 為後向 LSTM 網路的輸出向量； w_h 和 b_h 為權重值和偏移量。

5. 全連接與輸出層(Fully Connected & Output Layer)

將雙向長短期記憶網路層之輸出 (h'_t) 透過扁平層 (Flatten Layer) 全數降為一維後，再經由全連接層 (Dense Layer) 進行疊加計算出最終交通速率 (y_t) 之預測結果。計算過程表示如下：

$$y_t = \sigma(w_g h'_t + b_g) \quad (12)$$

四、模式構建與實驗結果

交通預測需要根據不同的交通管理策略目的來決定預測的時間跨度，若要執行瞬間或

短時間內的交通控制與管理策略，此類交通預測所選擇的時間跨度會偏好 1 分鐘、3 分鐘、5 分鐘、10 分鐘等短期時間單位；若要掌握交通流發展的趨勢，藉以擬定長期的交通管理策略，此類交通預測所選擇的時間跨度會偏好 30 分鐘、60 分鐘甚至更長分鐘數之長期時間單位。而本計畫高速公路交通參數預測之目的較偏向後者，屬於長期預測之範疇，預測結果主要用以輔助交通管理單位掌握交通流趨勢並做相關視覺化資訊揭露呈現。

4.1 資料蒐集與預處理

1. 研究範圍說明

本計畫研究範圍選定之目標路段為國道 1 號南下 166.4~172.5K，路段涵蓋臺中系統-豐原-大雅，如前所述，為考量時空特性之影響，遂將目標路段之上游路段範圍一併納入作為預測模式構建使用，而上游路段範圍為國道 1 號南下 162.1~166.4K，路段涵蓋后里-臺中系統-豐原，研究範圍彙整如圖 6 所示。以下茲就高速公路交通參數預測精進模式構建內容與步驟進行說明，包含資料蒐集、資料預處理、預測模式構建與實驗結果。



圖 6 交通預測研究涵蓋範圍

2. 資料蒐集

本研究以高速公路 ETC 系統所蒐集之交通資料進行交通速率預測模型構建，資料來源為高速公路局歷史資料庫 M05A，資料共 6 個欄位，包含 TimeInterval、GantryFrom、GantryTo、VehicleType、SpaceMeanSpeed 及 Volume，原始資料架構與內容如表 2 及圖 7 所示。資料蒐集量為 365 天，其中包含了 261 個工作日及 104 個週末日。此外，本研究另蒐集歷史降水資料，藉以探討是否降雨對於交通預測之影響，資料來源為 CWB Observation Data Inquire System (CODiS) 臺中市觀測站。

表 2 M05A 資料架構與內容

欄 位	欄位名稱	資料名稱	型態及計算方式
A	TimeInterval	報表產製時間	● 顯示方式：2021-04-12 22:55:00 ● 以 5 分鐘為一統計時階
B	GantryFrom	上游偵測站名稱	—
C	GantryTo	下游偵測站名稱	—
D	VehicleType	車種	31：小客車 32：小貨車 41：大客車 42：大貨車 5：聯結車
E	SpaceMeanSpeed	空間平均速率	● 顯示方式：平均車速：該時階之各車輛車速加總/交通量。 ● 單一車輛計算車速：相鄰上下游偵測站之門架距離/旅行時間(單位 km/ hr)。
F	Volume	交通量	計算上下游偵測站在此張表之時階範圍內所經過之車流量總量。

資料來源：高速公路局 M05A 開放資料、本計畫彙整

3. 資料預處理

巨量資料下載完畢後即進行資料預處理，共分為八個步驟，分別為資料讀取、資料篩選、資料重新命名與組合、資料聚合、遺漏值填補、資料融合（雨量資料、時空相關特徵融合）、週期性特徵編碼轉換、特徵分析，如圖 8 所示。茲就各步驟執行內容進行說明。

(1) 資料讀取：首先利用程式語言 Python 3.8 讀入 105,120 個檔案約 20 多億筆之巨量資料。

而本計畫所有的實驗均是在以下環境中進行：Intel(R) Core(TM) i7-9700 CPU @ 3.00 GHz、NVIDIA GeForce GT 1030、64.0GB RAM。

- (2) 資料篩選：接著進行資料篩選，將每個檔案針對目標路段 01F1664S~01F1725S 及其上游路段 01F1621S~01F1664S 之範圍資料進行採樣，而每個檔案中針對各路段範圍分別取出 5 筆資料，如圖 9 所示，目標路段之資料筆數共計 525,600 筆(5 筆資料*105,120 個檔案)，上游路段資料筆數亦同。
- (3) 資料重新命名與組合：將各路段的 105,120 個檔案進行資料組合，留下 TimeInterval、VehicleType、SpaceMeanSpeed、Volume 共 4 個欄位，並分別重新命名為時間、車種、速率、流量，此時資料型態如圖 10 所示，而資料總筆數為 525,600 筆。
- (4) 資料聚合：亦即時間解析度轉換，將每 5 分鐘為一筆之資料型態聚合成每 1 小時為一筆，亦即將各時段 12 筆資料 (如 0:00、0:05、0:10、0:15、...、0:55) 加權平均為 1 小時資料(00:00、01:00、02:00、03:00、...、23:00)。以圖 10 之資料型態為例，每 5 分鐘有 5 筆 (5 個車種) 速率及流量資料，因此每 1 小時共有 60 筆資料，本研究以式 (13) 進行加權平均後將資料聚合，此時每個路段之資料總筆數各為 8,760 筆。

$$\frac{\sum_{i=1}^{12} \sum_{j=1}^5 (\bar{v}_{ij} \times f_{ij})}{\sum_{i=1}^{12} \sum_{j=1}^5 f_{ij}} \quad (13)$$

	A	B	C	D	E	F	G	H	I	J	K	L
1	2019/4/5 18:10 01F0017N	01F0005N	31	93	69							
2	2019/4/5 18:10 01F0017N	01F0005N	32	91	15							
3	2019/4/5 18:10 01F0017N	01F0005N	41	86	1							
4	2019/4/5 18:10 01F0017N	01F0005N	42	86	1							
5	2019/4/5 18:10 01F0017N	01F0005N	5	0	0							
6	2019/4/5 18:10 01F0029N	01F0017N	31	90	82							
7	2019/4/5 18:10 01F0029N	01F0017N	32	90	19							
8	2019/4/5 18:10 01F0029N	01F0017N	41	83	4							
9	2019/4/5 18:10 01F0029N	01F0017N	42	84	1							
10	2019/4/5 18:10 01F0029N	01F0017N	5	0	0							
11	2019/4/5 18:10 01F0005S	01F0017S	31	87	77							
12	2019/4/5 18:10 01F0005S	01F0017S	32	89	19							
13	2019/4/5 18:10 01F0005S	01F0017S	41	84	2							
14	2019/4/5 18:10 01F0005S	01F0017S	42	0	0							
15	2019/4/5 18:10 01F0005S	01F0017S	5	0	0							
16	2019/4/5 18:10 01F0061N	01F0029N	31	96	109							
17	2019/4/5 18:10 01F0061N	01F0029N	32	96	31							
18	2019/4/5 18:10 01F0061N	01F0029N	41	92	4							
19	2019/4/5 18:10 01F0061N	01F0029N	42	0	0							
20	2019/4/5 18:10 01F0061N	01F0029N	5	0	0							
21	2019/4/5 18:10 01F0017S	01F0029S	31	90	83							
22	2019/4/5 18:10 01F0017S	01F0029S	32	90	21							
23	2019/4/5 18:10 01F0017S	01F0029S	41	96	1							
24	2019/4/5 18:10 01F0017S	01F0029S	42	0	0							
25	2019/4/5 18:10 01F0017S	01F0029S	5	0	0							
26	2019/4/5 18:10 01F0099N	01F0061N	31	92	125							

資料來源：高速公路局 M05A 開放資料

圖 7 M05A 原始資料型態

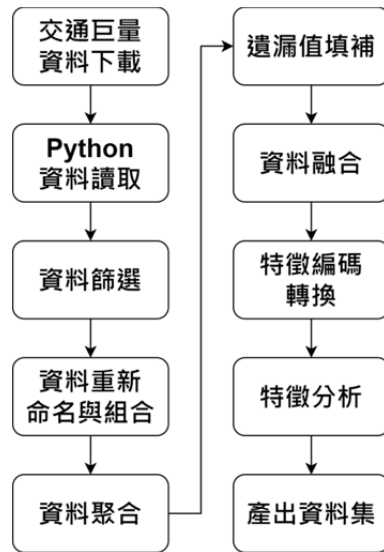


圖 8 資料預處理流程

	A	B	C	D	E	F	G
1	2019/4/5 19:00 01F1621S	01F1664S		31	99	138	
2	2019/4/5 19:00 01F1621S	01F1664S		32	97	28	
3	2019/4/5 19:00 01F1621S	01F1664S		41	93	4	
4	2019/4/5 19:00 01F1621S	01F1664S		42	61	2	
5	2019/4/5 19:00 01F1621S	01F1664S		5	64	2	
6	2019/4/5 19:00 01F1664S	01F1725S		31	97	194	
7	2019/4/5 19:00 01F1664S	01F1725S		32	97	41	
8	2019/4/5 19:00 01F1664S	01F1725S		41	90	7	
9	2019/4/5 19:00 01F1664S	01F1725S		42	90	1	
10	2019/4/5 19:00 01F1664S	01F1725S		5	75	3	
11							

圖 9 篩選後之資料型態示意

式中， i 為單位時段數； j 為車種； $\bar{v}_{ij}$ 為 i 時段 j 車種之平均速率（公里/小時）； f_{ij} 為 i 時段 j 車種之流率（車輛數/5 分鐘）。

- (5) 遺漏值填補：將資料進行組合後，發現目標路段 01F1664S~01F1725S 資料共有 9 筆遺漏值，本研究使用歷史平均法 (Historical Average) 以相同星期別及相關時段之平均數進行遺漏值填補，而上游路段 01F1621S~01F1664S 資料則無遺漏值。
- (6) 資料融合：將歷史是否降水資料、上游路段交通資料及目標路段交通資料進行融合。
- (7) 特徵編碼：

原始的日期欄位資料型態（如 2019/1/1 20:00）轉換成星期別（1~7）及小時（0~23），分別作為日間（Interday）週期性特徵及日內（Intraday）週期性特徵。另因在構建機器學

	A	B	C	D
1	時間	車種	速率	流量
2	2019/1/1 00:00	31	107	63
3	2019/1/1 00:00	32	107	10
4	2019/1/1 00:00	41	100	5
5	2019/1/1 00:00	42	105	4
6	2019/1/1 00:00	5	107	1
7	2019/1/1 00:05	31	106	50
8	2019/1/1 00:05	32	111	8
9	2019/1/1 00:05	41	95	2
10	2019/1/1 00:05	42	104	6
11	2019/1/1 00:05	5	0	0
12	2019/1/1 00:10	31	112	57
13	2019/1/1 00:10	32	110	16
14	2019/1/1 00:10	41	107	2
15	2019/1/1 00:10	42	113	5
16	2019/1/1 00:10	5	106	1
17	2019/1/1 00:15	31	107	119
18	2019/1/1 00:15	32	108	28
19	2019/1/1 00:15	41	98	3
20	2019/1/1 00:15	42	104	3
21	2019/1/1 00:15	5	0	0
22	2019/1/1 00:20	31	104	168
23	2019/1/1 00:20	32	104	40

圖 10 資料重新命名與組合後之資料型態示意

習模式過程中，需要將非連續型的特徵（亦即分類型變數）進行轉換後再做為模型之輸入值，因此需將上述特徵以獨熱編碼(One-Hot Encoding)形式引入模型中使用，例如以 7 碼的型式區分不同星期別，工作日中的星期二為[0,1,0,0,0,0,0]、週末中的星期六為[0,0,0,0,0,1,0]；以 24 碼型式區分不同小時別，藉此考量尖峰時段與離峰時段之差異，例如 17:00 中的 17 表示為[0,0]。

(8) 特徵分析：

該步驟亦即特徵提取(Feature Extraction)，當模型輸入的資料太過龐大冗餘，會大幅降低模型訓練之效率，亦即分析過多特徵變數會耗費大量的記憶體以及計算能力，且可能也會產生過度擬合之問題，需要透過特徵提取刪除大量冗餘或無關之特徵變量來降低維度，達到降低模型計算量從而減少模型訓練時間，亦可減少資料雜訊對於模型分析的影響。

本計畫之目標特徵為速率 (speed)，其他採納之特徵共計 7 項（詳表 3），其中屬於連續型特徵包含目標路段前一時階流量 (bf60_volume)、目標路段前一時階速率 (bf60_speed)、上游路段前一時階流量 (us_bf60_volume)、上游路段前一時階速率 (us_bf60_speed) 四

項；屬於分類型特徵包含星期別 (day_of_week)、小時別 (hourly)、是否降雨 (precipitation)。

針對連續型特徵，利用皮爾森相關係數 (Pearson Correlation Coefficient) 分析方法，分析四項連續型特徵變數與預測目標(速率)間之相關係數，係數絕對值越大代表該特徵越重要，基本上絕對值介於 0.0~0.2 間就屬於極弱相關或無相關；針對分類型特徵，利用變異數 (ANOVA) 分析方法，分析三項分類型特徵變數與預測目標 (速率) 間之顯著性 (有無影響) 及效果性 (影響程度)，p 值小於 0.05 則代表該特徵對於預測目標有顯著影響，反之則無顯著影響； η_p^2 (Partial Eta Squared, 淨 eta 平方) 越大表示自變數 (特徵變數) 對依變數(預測目標)就越重要，影響效果越大， $0.01 \leq \eta_p^2 < 0.058$ 為小效果， $0.058 \leq \eta_p^2 < 0.138$ 為中效果， $0.138 \leq \eta_p^2$ 為大效果。

表 3 本研究模式考量之時空特性特徵彙整

變數名稱	變數定義	屬性
speed	速率	Target Label
bf60_volume	目標路段前一時階流量	Spatio-Temporal Feature
us_bf60_volume	上游路段前一時階流量	Spatio-Temporal Feature
bf60_speed	目標路段前一時階速率	Spatio-Temporal Feature
us_bf60_speed	上游路段前一時階速率	Spatio-Temporal Feature
day_of_week	星期別	Periodic Feature of Interday
hourly	小時別	Periodic Feature of Intraday
precipitation	是否降雨	Other Feature

依據表 4 之四項特徵與兩種分析情境 (工作日速率、週末速率) 之相關性分析彙整可知，相關性皆大於 0.2，本計畫將上述相關性分析結果繪製成熱力圖以視覺化呈現如圖 11 所示，而前述相關係數絕對值>0.2 之特徵將予以保留。

依據表 5 之三項特徵與兩種分析情境之顯著性分析彙整可知，三項分類型特徵 p 值皆小於 0.05，表示對於工作日及週末之速率存在顯著影響關係，亦即三項分類特徵對於預測目標是有影響。而特徵中的小時別 (hourly) 之 η_p^2 值皆超過 0.138，表示 hourly 此項特徵對於速率影響效果高，因此保留作為後續預測模式構建使用。星期別 (day_of_week) 特徵變數之 η_p^2 值介於 0.058 ~ 0.138 間，為中效果之影響程度，亦可保留作為後續預測模式構建使用。而是否降雨 (precipitation) 特徵變數之 η_p^2 值介於 0.008 ~ 0.016 間，為小效果甚至為低於 0.01 之無效果影響程度，因此該項特徵將予以剔除。

表 4 不同日態之速率與連續類型特徵相關性分析

特徵 相關係數		目標標籤 (Target Label)-速率	
		工作日	週末
特徵 (Feature)	bf60_volume	-0.53	-0.67
	us_bf60_volume	-0.61	-0.71
	bf60_speed	0.73	0.80
	us_bf60_speed	0.33	0.70

註：

bf60_volume、bf60_speed：目標路段前一時階流量、速率

us_bf60_volume、us_bf60_speed：上游路段前一時階流量、速率

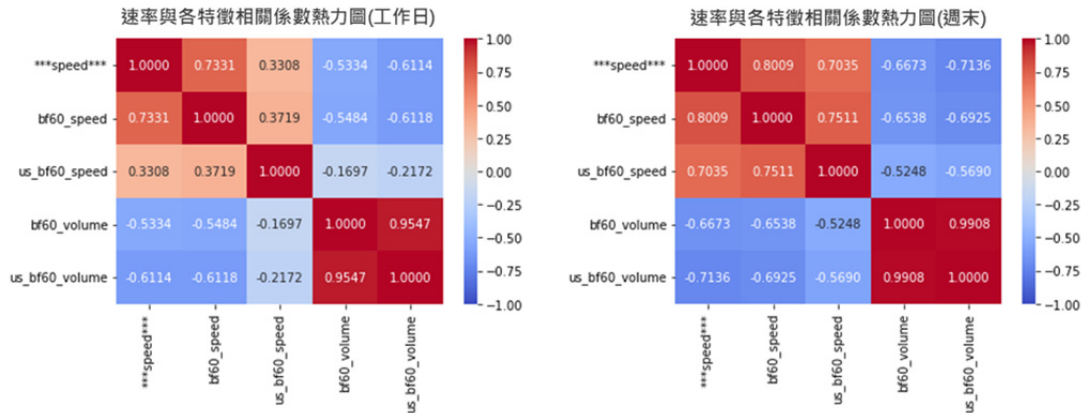


圖 11 速率與各特徵之相關係數熱力圖-工作日與週末

表 5 不同日態之速率與分類型特徵變異數分析

預測目標	日態	指標	星期別 (day_of_week)	小時別 (hourly)	是否降雨 (precipitation)
速率	工作日	F	100.325	338.718	98.237
		p	0.000	0.000	0.000
		η_p^2	0.062	0.564	0.016*
	週末	F	79.349	183.687	19.619
		p	0.000	0.000	0.000
		η_p^2	0.032	0.638	0.008*

註：0.01 $\leq \eta_p^2 < 0.058$ 為小效果；0.058 $\leq \eta_p^2 < 0.138$ 為中效果；0.138 $\leq \eta_p^2$ 為大效果。

綜合前述，本研究之預測目標 (Target Label) 為速率，用於後續模式構建之特徵包含前一時階流量 (bf60_volume)、目標路段前一時階速率 (bf60_speed)、上游路段前一時階流量(us_bf60_volume)、上游路段前一時階速率 (us_bf60_speed)、星期別 (day_of_week)、小時別 (hourly)，最後匯出上述內容之資料集。

4.2 模式構建與預測結果

1. 模式構建

本研究於速率預測模式建構流程如圖 12，經由上述資料預處理過程產出模式建構資料集後，首先進行資料集劃分，而機器學習主要係由資料中進行學習並建構模型，該過程稱為訓練，建構模型的過程中，須將資料劃分為訓練集 (Training Dataset)、驗證集 (Validation Dataset)、測試集 (Test Dataset)，訓練集主要為訓練階段時使用、驗證集主要為調整模型超參數時使用、測試集主要為模型最終評估使用。而本研究將 1 年共 52 週之資料集，劃分為訓練集和測試集，訓練集為前 48 週資料，測試集資料為後 4 週資料，又將訓練集資料中的前 10%作為驗證資料、後 90%則為訓練資料。

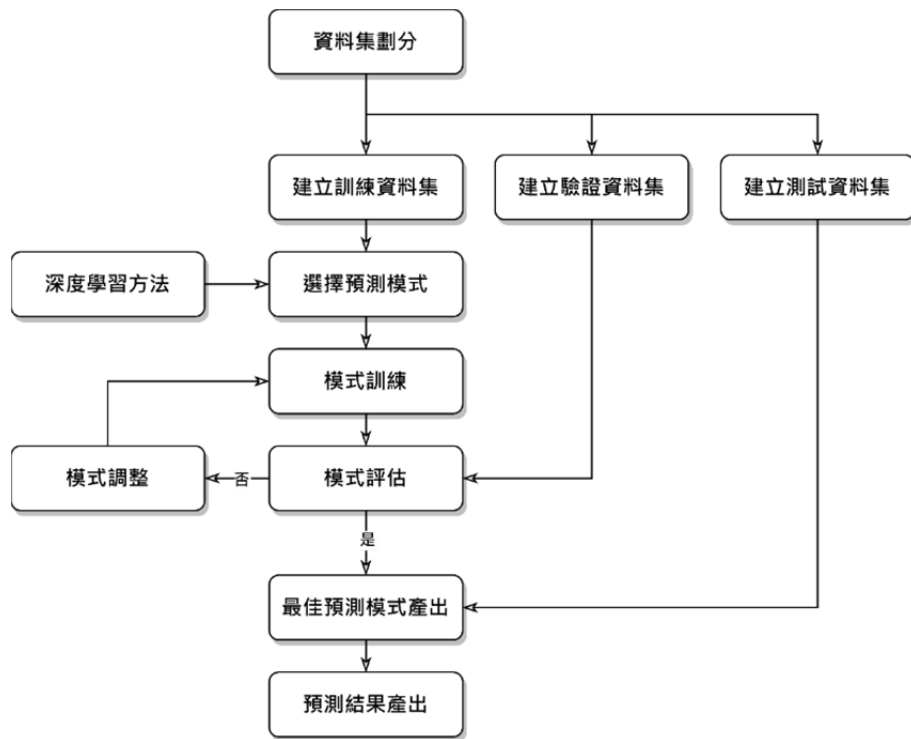


圖 12 預測模式構建流程

資料集劃分完畢後，選擇建構預測模型使用之模式與方法，本研究提出一種深度學習

混合模型，結合了 LSTM、注意力機制、Bi-LSTM，藉此增強預測模型對於不同時段交通流時空特性的學習力，以提升預測精度。使用訓練集用以訓練模式，而驗證集用以評估模式之表現，若表現不好則返回訓練模式進行相關參數調整後，再次將驗證集資料引入訓練模式進行評估，表現良好則代表該訓練模式為最佳預測模式，最後將實際待預測之測試資料集輸入最佳預測模式中進行實際預測，並產出實際預測結果。

本研究所有的實驗均是在以下環境中進行：Intel(R) Core(TM) i7-9700 CPU @ 3.00 GHz、NVIDIA GeForce GT 1030、64.0GB RAM，並且以兩種常用的指標評估預測之準確性，包含平均絕對值誤差百分比 (Mean Absolute Percentage Error, MAPE)、平均絕對誤差 (Mean Absolute Error, MAE)。MAPE 係由 Lewis^[67] 於 1982 年提出，為評估預測精準度最常用的指標，該值為相對數值，較不受到預測值與真實值的單位和大小影響，因此得到預測值與真實值之差異程度較客觀，其公式如下所示，MAPE 值的範圍為 $(0, +\infty)$ ，MAPE 為 0%，表示為完美模型。上述指標定義如下，速率真實值以 y 表示，速率預測值以 $\hat{y}$ 表示。

$$MAPE(y, \hat{y}) = \frac{100}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (14)$$

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (15)$$

2. 預測結果說明

本研究使用 LSTM-Attention-BiLSTM 模型進行 6 種情境 (全日全時段、全日尖峰時段、工作日全時段、工作日尖峰時段、週末全時段、週末尖峰時段) 之交通速率預測，並與 LSTM、BiLSTM、LSTM-Attention、BiLSTM-Attention 模型以及 Transformer 模型進行比較。Transformer 為 Google 團隊 Vaswani 等人^[68] 於 2017 年在 Attention Is All You Need 一文中提出，該文使用 Attention 替換了原先 Seq2Seq 模型中的迴圈結構，Transformer 的模型結構是由 Self-attention 和前饋神經網路 (Feed Forward Neural Network) 組成，在時間序列之自然語言領域取得良好之預測成效。而本研究模型亦同樣以 Attention 作為模型主體，再分別連結神經網路 LSTM (前饋) 及 Bi-LSTM (前饋+反饋) 所組成，與 Transformer 類似。而本研究為確認所發展之模式是否亦能於時間序列相關之預測方面優於其他預測模式，故選取 Transformer 進行比較。

為確保實驗之公平性，本研究的 LSTM-Attention-BiLSTM 模型以及其他用於比較的基準模型 (LSTM、BiLSTM、LSTM-Attention、BiLSTM-Attention、Transformer) 在訓練過程中，皆是選擇均方誤差 (MSE) 作為模型訓練的損失函數 (Loss Function)，選擇 Adam 函數作為模型訓練的優化器 (Optimizer)，迭代次數 (epoch) 為 40，批次處理大小 (batch size) 為 64，而模型輸出層的激活函數 (Activation Function) 則是採用 softmax，時間窗大小 (time

window size) 設置為 24。

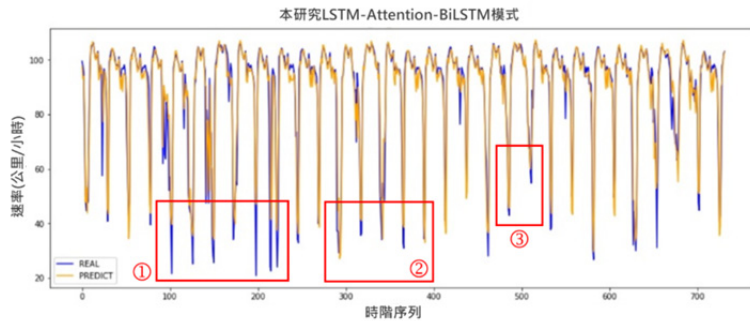
實驗結果彙整如表 6，可看出本研究所提出之 LSTM-Attention-BiLSTM 模型於 6 種情境之預測準確性皆優於其他 5 種模型，在全時段預測中之 MAPE 介於 6.852%~6.946%、MAE 介於 3.953~4.384；在尖峰時段預測中 MAPE 介於 15.762%~16.354%、MAE 介於 7.798~9.869。另外，本研究將 LSTM-Attention-BiLSTM 與 Transformer 之預測結果與真實值繪製如圖 13 所示，可發現本研究模式明顯更符合現實趨勢。

圖 14 為本研究彙整混合模型 LSTM-Attention-BiLSTM 優於 Google 之 Transformer 模式之幅度。於各日態之全時段預測（藍色長條）優於 Transformer 模式 43.447%~99.385%；於各日態之尖峰時段預測（紅色長條）優於 Transformer 模式 60.497%~68.196%。且本研究模式於全日及工作日之尖峰時段預測優良幅度明顯高於全日時段，顯示 LSTM-Attention-BiLSTM 模型較能夠有效捕捉到尖峰時段時交通流劇烈變化之特性。

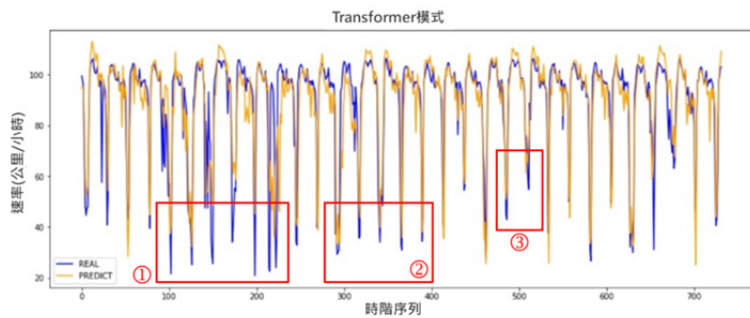
表 6 不同模型之預測性能比較彙整

日態	模式	全時段		尖峰時段	
		MAPE(%)	MAE	MAPE(%)	MAE
全日	LSTM	8.026	4.671	18.032	9.404
	BiLSTM	7.759	4.655	17.273	8.857
	LSTM-Attention	8.004	4.722	17.998	9.650
	BiLSTM-Attention	10.890	6.152	22.924	13.413
	Transformer	10.166	6.447	26.248	12.839
	LSTM-Attention-BiLSTM*	6.880	4.163	16.354	7.798
工作日	LSTM	9.192	5.491	19.170	10.784
	BiLSTM	8.103	4.888	19.010	9.277
	LSTM-Attention	8.038	4.655	18.503	9.844
	BiLSTM-Attention	8.959	5.361	19.124	11.012
	Transformer	9.829	6.440	26.511	12.666
	LSTM-Attention-BiLSTM*	6.852	3.953	15.762	7.916
週末	LSTM	10.663	6.144	22.668	14.613
	BiLSTM	8.950	5.482	18.718	11.215
	LSTM-Attention	8.530	5.234	19.283	11.153
	BiLSTM-Attention	11.230	6.573	23.092	14.522
	Transformer	13.849	7.570	26.765	19.782
	LSTM-Attention-BiLSTM*	6.946	4.384	15.988	9.869

註：*本研究使用模型



(a)本研究預測模式



(b)Transformer 預測模式

說明：

- 本研究使用同一批交通速率資料進行各預測模式構建，上圖為使用本研究模式產出之預測值(橘色線)與真實值(藍色線)之疊圖；下圖為使用Google Transformer模式產出之預測值(橘色線)與真實值(藍色線)之疊圖。
- 將上圖①號框處及下圖①號框處進行比較，發現上圖之橘色線明顯較越貼近藍色線，表示上圖之預測值越接近真實值，代表上圖模式之預測效果優於下圖之預測模式。②號與③號框處亦同。

圖 13 本研究模式及 Transformer 模式之預測值及真實值比較趨勢圖

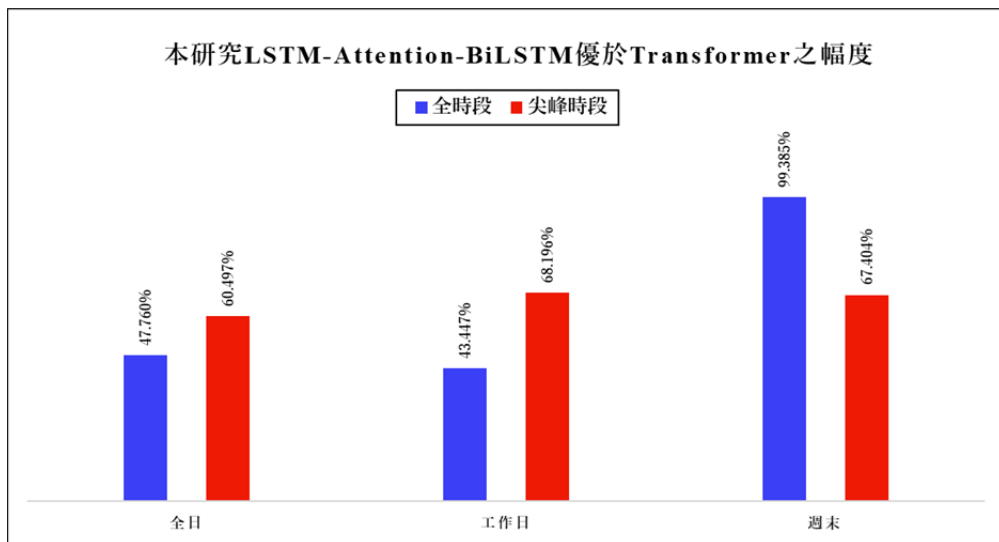


圖 14 本研究模式於各日態預測結果優於 Transformer 模式之幅度

此外，在判斷預測模型之良窳時，除考量預測準確性外，模型於訓練過程中所耗費之時間亦會是使用者考量是否採納該模型進行應用之關鍵，資料量多寡、資料維度（亦即特徵數量）、電腦效能、模型複雜度皆會影響訓練時間，本研究於資料量、資料維度及每次訓練使用之電腦皆為固定之情形下，比較採用不同模型之訓練時間差異並彙整如圖 15 所示。由圖中可知，本研究模式所需之訓練時間，無論於何種日態下皆優於 Google 團隊所提出之 Transformer 模式。與其他同系列單一模型（LSTM、BiLSTM）及兩種模型之混合模型（LSTM-Attention、BiLSTM-Attention）之訓練時間進行比較，於全日日態下多了 37~156 秒（0.6~2.6 分鐘）訓練時間；工作日日態下多了 44~85 秒（0.7~1.4 分鐘）；週末日態下多了 158~312 秒（2.6~5.2 分鐘），惟本研究模式所增加之模型訓練時間尚在可接受範圍，未來亦可透過硬體設備之升級降低訓練所需之時間。

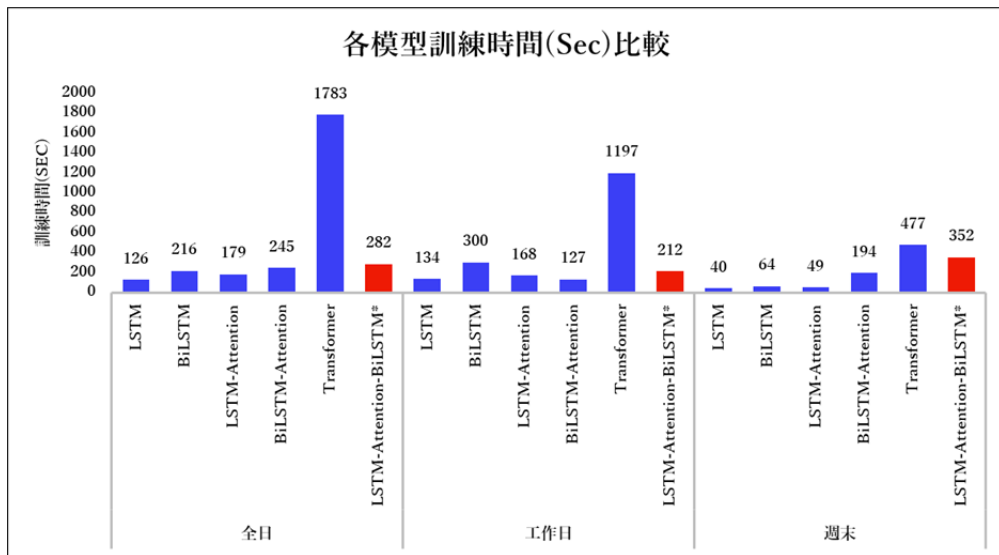


圖 15 不同模型之訓練時間比較彙整

前述無論是本研究的 LSTM-Attention-BiLSTM 模型，或是其他用於比較的基準模型，包含 LSTM、BiLSTM、LSTM-Attention、BiLSTM-Attention 及 Transformer，皆屬於機器學習中之深度學習模式，而過去文獻亦曾使用過傳統進階統計模式以及機器學習中之淺層學習模式進行交通預測，本研究為比較上述三類型模式之差異，遂以傳統進階統計模式中之時間序列 ARIMA 模式及淺層學習模式中之支持向量迴歸 (Support Vector Regression, SVR) 模式進行預測，並與本研究使用之 LSTM-Attention-BiLSTM 深度學習模式進行資料需求、計算耗費時間及預測誤差之差異比較，彙整如表 7 所示。本研究採用之深度學習模式與傳統進階統計模式相比，模式計算時間增加約 269 秒（約 4.5 分鐘），預測誤差改善 60.5%；與淺層學習模式相比，模式計算時間增加約 267 秒（約 4.5 分鐘），預測誤差改善 29.8%。

表 7 不同類型模式之預測比較彙整

模式類型	模式名稱	輸入資料需求	模式計算時間	預測誤差 (MAPE)
進階統計	ARIMA	● 僅能輸入單一特徵。 ● 以本研究為例，資料的輸入為速率。	13 秒 (0.2 分鐘)	16.972%
機器學習-淺層學習	SVR	● 可輸入多特徵，亦可只輸入單一特徵。	15 秒 (0.3 分鐘)	9.799%
機器學習-深度學習	LSTM-Attention-BiLSTM	● 以本研究為例，資料的輸入包含速率、前一時階流量 (bf60_volume)、目標路段前一時階速率 (bf60_speed)、上游路段前一時階流量 (us_bf60_volume)、上游路段前一時階速率 (us_bf60_speed)、星期別 (day_of_week)、小時別 (hourly)。	282 秒 (4.7 分鐘)	6.880%

五、結論與建議

本研究提出一新穎之混合模式，名為 LSTM-Attention-BiLSTM，進行交通速率預測，該方法透過 LSTM 提取交通時空特性中之重要資訊，再透過 Attention 機制針對各重要資訊賦予其權重，最後以 BiLSTM 綜合考量兩個方向的交通流量序列資訊進行預測。綜合前述，臚列相關研究結論如下：

1. 回顧過去相關研究，鮮少研究能夠同時考量不同日態之全時段及尖峰時段之預測準確性，基於此，本研究之交通速率預測係針對 6 種時段情境，包含全日全時段、全日尖峰時段、工作日全時段、工作日尖峰時段、週末全時段、週末尖峰時段進行。
2. 時空特性係指交通流中，前一刻的交通流會對下一刻的交通流產生影響，意即現在或未來的交通狀態會受歷史交通狀態影響。另相鄰路段的交通流之間也會相互影響，下游車流之變化趨勢與上游車流變化趨勢有關。因此本研究除採用常見的週期性特徵(星期別、小時別)外，亦加入研究範圍上游路段前一時階之速率以及目標路段前一時階速率作為模式構建考量之時空特徵。並且蒐集歷史降水資料，藉以探討降雨對於交通預測之影響，然經相關性分析後，「是否降雨」對於本研究預測無顯著影響，因此予以剔除。
3. 本研究預測模式之預測結果，以 MAPE 作為評估指標，於全日全時段為 6.880%、全日尖峰時段為 16.354%；於工作日全時段為 6.852%、尖峰時段為 15.762%；週末全時段為 6.946%、尖峰時段為 15.988%。以 Lewis 於 1982 年提出之劃分標準來看，MAPE<10%

屬高準確預測、10~20%屬於優良預測。

4. 該模式與 Transformer 模式相比，於各日態全時段預測之 MAPE 改善幅度約為 39.609~59.804%，於尖峰時段之改善幅度約為 31.784~39.123%。尤其於週末日之預測改善幅度更為明顯，改善幅度為 59.804%，顯示本研究提出之模式能有效捕捉到週末日時段交通流劇烈變化之特性。
5. 本研究採用之深度學習模式與傳統進階統計模式相比，模式計算時間增加約 269 秒（約 4.5 分鐘），預測誤差改善 60.5%；與淺層學習模式相比，模式計算時間增加約 267 秒（約 4.5 分鐘），預測誤差改善 29.8%，惟本研究模式所增加之模型訓練時間尚在可接受範圍，未來亦可透過硬體設備之升級降低訓練所需之時間。

此外，在後續研究上，建議可朝三個方向進行：

1. 無論哪個領域，預測的重要意義在於透過大量的歷史數據，從中找尋客觀之規律後進行預測未來的事件，藉以事先掌握未來趨勢達到預警目的，並研擬相關策略。因此本研究發展之預測模式，並不侷限於交通參數之預測，未來研究亦可蒐集其他交通領域之資料，如公共運輸、停車等進行客流、停車需求相關預測。
2. 由於本研究預測模式並未考量特殊事件對交通造成之影響，例如事故、施工、大型活動等。鑒於交通事件影響預測為一項重要且困難之研究課題，尤其隨時空範圍的擴大，預測困難度亦會隨之提升，主因為交通事件在時空範圍中相對稀疏，且在不同時空環境背景下所發生之事故同質性低，難以掌握其趨勢。故後續隨 5G、車聯網等資通訊技術之發展，相關交通資料之蒐集能較全面且多元，可嘗試透過本研究模式進行交通事件影響之預測。
3. 本研究在研究範圍部分係以國道 1 號南下 166.4~172.5K 為範圍，後續可納入車行速率之空間特性選取不同路段進行分析探討。

致 謝

本研究係於交通部運輸研究所 MOTC-IOT-110-IFB001 及 MOTC-IOT-111-IFB010 計畫支持下之部分研究成果，特此感謝。

參考文獻

1. Han, T. Y., Tang, K. H., Oguchi, T., "Short-Term Travel Speed Prediction for Urban Expressways Using Convolutional Neural Network and Tensor Decomposition", *Transportation Research Procedia*, Vol.48, 2020, pp.962-974.
2. Okutani, I., Stephanedes, Y. J., "Dynamic Prediction of Traffic Volume Through Kalman Filtering Theory", *Transportation Research Part B: Methodological*, Vol.18, Iss.1, 1984,

pp.1-11.

3. Cheng, T., Haworth, J., Wang, J., “Spatio-Temporal Autocorrelation of Road Network Data”, *Journal of Geographical System*, Vol.14, 2012, pp.389-413.
4. Dunne, S., Ghosh, B., “Regime-Based Short-Term Multivariate Traffic Condition Forecasting Algorithm”, *Journal of Transportation Engineering*, Vol.138, Iss.4, 2012, pp.455-466.
5. Xia, J., Huang, W., Guo, J., “A Clustering Approach to Online Freeway Traffic State Identification Using ITS Data”, *KSCE Journal of Civil Engineering*, Vol.16, Iss.3, 2012, pp. 426-432.
6. Sun, S., Huang, R., Gao, Y., “Network-Scale Traffic Modeling and Forecasting with Graphical Lasso and Neural Networks”, *Journal of Transportation Engineering*, Vol.138, Iss.11, 2012, pp. 1358-1367.
7. Fang, W., Zhuo, J., Song, Y., Jiang, D., Zhou, T., “Attention Meets Long Short-Term Memory: A Deep Learning Network for Traffic Flow Forecasting”, *Physica A: Statistical Mechanics and its Applications*, Vol.587, 2022, Article 126485.
8. Sun, Y., Jiang, G., Lam, S. K., He, P., Ning F., “Multi-fold Correlation Attention Network for Predicting Traffic Speeds with Heterogeneous Frequency”, *Applied Soft Computing*, Vol.124, 2022, 108977.
9. Almeida, A., Brás, S., Oliveira, I., Sargento, S., “Vehicular Traffic Flow Prediction Using Deployed Traffic Counters In A City”, *Future Generation Computer Systems*, Vol.128, 2022, pp. 429-442.
10. Cui, Z. Y., Lin, L. F., Pu, Z., Wang, Y. H., “Graph Markov Network for Traffic Forecasting with Missing Data”, *Transportation Research Part C: Emerging Technologies*, Vol.117, 2020, 102671.
11. Zheng, Y., Dong, C., Dong, D., Wang, S., “Traffic Volume Prediction: A Fusion Deep Learning Model Considering Spatial-Temporal Correlation”, *Sustainability*, Vol.13, Iss.19, 2021, 10595.
12. Gentili, M., Mirchandani, P. B., “Locating Sensors on Traffic Networks: Models, Challenges And Research Opportunities”, *Transportation Research Part C: Emerging Technologies*, Vol.24, 2012, pp.227-255.
13. Kashi, S. O., Akbarzadeh, M., “A Framework for Short-Term Traffic Flow Forecasting Using The Combination of Wavelet Transformation and Artificial Neural Networks”, *Journal of Intelligent Transportation Systems*, Vol.23, Iss.1, 2019, pp.60-71.
14. Tang, J., Chen, X., Hu, Z., Zong, F., Han, C., Li, L., “Traffic Flow Prediction Based on Combination of Support Vector Machine and Data Denoising Schemes”, *Physica A: Statistical Mechanics and its Applications*, Vol.534, No.15, 2019, 1210642.
15. Suvin, P. V., Ankita, R. K., Sabina, Y. I., Mallikarjuna, C., “Criteria for Temporal Aggregation of the Traffic Data from a Heterogeneous Traffic Stream”, *Transportation Research Procedia*, Vol.48, 2020, pp.3401-3412.

16. Cassdei, G., Bertand, V., Gouin, B., Canudas-De-Wit, C., "Aggregation and Travel Time Calculation Over Large Scale Traffic Networks: An Empiric Study on The Grenoble City", *Transportation Research Part C: Emerging Technologies*, Vol.95, 2018, pp.713-730.
17. Zhang, T., Zhang, D. G., Yan, H. R., Qiu, J. N., Gao J. X., "A New Method of Data Missing Estimation with FNN-Based Tensor Heterogeneous Ensemble Learning for Internet of Vehicle", *Neurocomputing*, Vol.420, 2021, pp.98-110.
18. Huang, C. and Lee, H., "A Grey-based Nearest Neighbor Approach for Missing Attribute Value Prediction", *Applied Intelligent*, Vol.20, No.3, 2004, pp.239-252.
19. Chen, X., He, Z., Sun, L., "A Bayesian Tensor Decomposition Approach for Spatiotemporal Traffic Data Imputation", *Transportation Research Part C: Emerging Technologies*, Vol.98, 2019, pp.73-84.
20. Lippi, M., Bertini, M., Frasconi, P., "Short-Term Traffic Flow Forecasting: An Experimental Comparison of Time-Series Analysis and Supervised Learning", *IEEE Transactions on Intelligent Transportation System*, Vol.14, No.2, 2013, pp.871-882.
21. Vlahogianni, E. I., Karlaftis, M. G., Golias, J. C., "Short-Term Traffic Forecasting: Where We Are and We'Re Going", *Transportation Research Part C: Emerging Technologies*, Vol.43, 2014, pp.3-19.
22. 高雄市政府交通局，*高雄市區建國路段旅行時間差異性檢定與推估*，民國 104 年。
23. Williams, B. M., Hoel, L. A., "Modeling and Forecast-ing Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results", *Journal of Transportation Engineering*, Vol.129, No.6, 2003, pp.664-672.
24. Guo, J., Huang, W., Williams, B. M., "Adaptive Kalman Filter Approach for Stochastic Short-Term Traffic Flow Rate Prediction and Uncertainty Quantification", *Transportation Research Part C: Emerging Technologies*, Vol.43, 2014, pp.50-64.
25. Ghosh, B., Basu, B., O'Mahony, M., "Multivarite Short-Term Traffic Flow Forecasting Using Time-Series Analysis", *IEEE Transactions on Intelligent Transportation Systems*, Vol.10, Iss.2, 2009, pp.246-254.
26. Yu, B., Song, X., Guan, F., Yang, Z., Yao, B., "k-Nearest Neighbor Model for Multiple-Time-Step Prediction of Short-Term Traffic Condition", *Journal of Transportation Engineering*, Vol. 142, Iss.6, 2016, 04016018.
27. Cai, Y., Wang, Y., Lu, G., Chen, P., Ding, C., Sun, J., "A Spatiotemporal Correlative K-Nearest Neighbor Model for Short-Term Traffic Multistep Forecasting", *Transportation Research Part C: Emerging Technologies*, Vol.62, 2016, pp.21-34.
28. Wu, C. H., Ho, J. M., Lee, D. T., "Travel-Time Prediction with Support Vector Regression", *IEEE Transactions on Intelligent Transportation Systems*, Vol.5, Iss.4, 2004, pp.276-281.
29. Liu, S., Yao, E., "Holiday Passenger Flow Forecasting Based on the Modified Least-Square Support Vector Machine for the Metro System", *Journal of Transportation Engineering Part A Systems*, Vol.143, Iss.2, 2017, 04016005.
30. Leshem, G., Ritov, Y., "Traffic Flow Prediction using Adaboost Algorithm with Random

- Forests as a Weak Learner”, *International Journal of Intelligent Technology*, Vol.2, No.2, 2007, pp.404-409.
31. Lai, W. K., Kuo, T. H., Chen, C. H., “Vehicle Speed Estimation and Forecasting Methods Based on Cellular Floating Vehicle Data”, *Applied Science*, Vol.6, No.2, 2016, pp.47.
 32. 黃冠維, 「利用卷積神經網路預測市區道路路況之研究」, 國立陽明交通大學運輸與物流管理學系碩士論文, 民國 107 年。
 33. Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., Li, H., “T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction”, *IEEE Transactions on Intelligent Transportation Systems*, Vol.21, No.9, 2020, pp.3848-3858.
 34. Benkraouda, O., Thodi, B.T., Yeo, H., Menéndez, M., Jabari S.E., “Traffic Data Imputation Using Deep Convolutional Neural Networks”, *IEEE Access*, Vol.8, 2020, pp.104740-104752.
 35. Jin, Y., Xu, W., Wang, P., Yan, J., “SAE Network: A Deep Learning Method for Traffic Flow Prediction”, In *2018 5th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS)*, 2018, pp.241-246.
 36. Zhao, Z., Chen, W., Wu, X., Chen, P., Liu, J., “LSTM Network: A Deep Learning Approach for Short-Term Traffic Forecast”, *IET Intelligent Transport Systems*, Vol.11, Iss.2, 2017, pp.68-75.
 37. Ma X., Tao Z., Wang Y., Yu H., Wang Y., “Long Short-Term Memory Neural Network for Traffic Speed Prediction Using Remote Microwave Sensor Data”, *Transportation Research Part C: Emerging Technologies*, Vol.54, 2015, pp.187-197.
 38. Wu, Y., Tan H., Qin, L., Ran B., Jiang Z., “A Hybrid Deep Learning Based Traffic Flow Prediction Method and its Understanding”, *Transportation Research Part C: Emerging Technologies*, Vol.90, 2018, pp.166-180.
 39. Yang, H., Li, X., Qiang, W., Zhao, Y., Tang, C., “A Network Traffic Forecasting Method Based on SA Optimized ARIMA-BP Neural Network”, *Computer Networks*, Vol.193, 2021, 108102.
 40. Bogaerts, T., Masegosa, A. D., Angarita-Zapata, J. S., Onieva, E., Hellinckx, P., “A Graph CNN-LSTM Neural Network for Short and Long-Term Traffic Forecasting Based on Trajectory Data”, *Transportation Research Part C: Emerging Technologies*, Vol.112, 2020, pp.62-77.
 41. Zhang, Y., Wang, M., “Peak Traffic Forecasting Using Nonparametric Approaches”, *Journal of Shanghai Jiaotong University (Science)*, Vol.17, 2012, pp.76-81.
 42. Peng, H., Du, B., Liu, M., Ji, S., Wang, S., Zhang, X., He, L., “Dynamic Graph Convolutional Network for Long-Term Traffic Flow Prediction with Reinforcement Learning”, *Information Sciences*, Vol.578, 2021, pp.401-416.
 43. 吳俊緯, 「捷運起訖流量預測之研究：基於混合深度學習方法之應用」, 東吳大學巨量資料管理學院碩士論文, 民國 111 年。
 44. 張軒慈, 「以雙向長短期記憶遞歸神經網路建構國道壅塞路段短期車流量預測模型」,

國防大學運籌管理學系碩士論文，民國 111 年。

45. Xing, L., Liu, W., “A Data Fusion Powered Bi-Directional Long Short Term Memory Model for Predicting Multi-Lane Short Term Traffic Flow”, *IEEE Transactions on Intelligent Transportation Systems*, Vol.23, Iss.9, 2021, pp.1-10.
46. Guo, G., Yuan, W., “Short-Term Traffic Speed Forecasting Based on Graph Attention Temporal Convolutional Networks”, *Neurocomputing*, Vol.410, 2020, pp.387-393.
47. Ma, C., Dai, G., Zhou, J., “Short-Term Traffic Flow Prediction for Urban Road Sections Based on Time Series Analysis and LSTM_BILSTM Method”, *IEEE Transactions on Intelligent Transportation Systems*, Vol.23, Iss.6, 2021, pp.1-10.
48. Do, L.N., Vu, H.L., Vo, B.Q., Liu, Z., Phung, D., “An Effective Spatial-Temporal Attention Based Neural Network for Traffic Flow Prediction”, *Transportation Research Part C: Emerging Technologies*, Vol.108, 2019, pp.12-28.
49. Qi, Y., Ishak, S., “A Hidden Markov Model for Short Term Prediction of Traffic Conditions on Freeways”, *Transportation Research Part C: Emerging Technologies*, Vol.43, Part 1, 2014, pp. 95-111.
50. Qi, Y., Ishak, S., “Stochastic Approach for Short-Term Freeway Traffic Prediction During Peak Periods”, *IEEE Transactions on Intelligent Transportation Systems*, Vol.14, Iss.2, 2013, pp. 660-672.
51. Chan, K. Y., Dillon, T. S., Singh, J., Chang, E., “Neural-Network-Based Models for Short-Term Traffic Flow Forecasting Using A Hybrid Exponential Smoothing and Levenberg-Marquardt Algorithm”, *IEEE Transactions on Intelligent Transportation Systems*, Vol.13, Iss.2, 2012, pp. 644-654.
52. Hong, W., “Traffic Flow Forecasting By Seasonal SVR with Chaotic Simulated Annealing Algorithm”, *Neurocomputing*, Vol.74, Iss.12-13, 2011, pp.2096-2107.
53. Ma, D., Song, X., Li, P., “Daily Traffic Flow Forecasting Through a Contextual Convolutional Recurrent Neural Network Modeling Inter- and Intra-Day Traffic Patterns”, *IEEE Transactions on Intelligent Transportation Systems*, Vol.22, Iss.5, 2021, pp.2627-2636.
54. Shen, L., Lu, J., Geng, D., Deng, L., “Peak Traffic Flow Predictions: Exploiting Toll Data from Large Expressway Networks”, *Sustainability*, Vol.13, No.1, 2020, pp.260.
55. Tang, J., Liu, F., Xou, Y., Zhang, W., Wang, Y., “An Improved Fuzzy Neural Network for Traffic Speed Prediction Considering Periodic Characteristic”, *IEEE Transactions on Intelligent Transportation Systems*, Vol.18, Iss.9, 2017, pp.2340-2350.
56. Du, S., Li, T., Tang, Y., Wang, H., Xie, P., Hong, S., “A Sequence-to-Sequence Spatial-Temporal Attention Learning Model for Urban Traffic Flow Prediction”, *Journal of Computer Research and Development*, Vol.57, No.8, 2020, pp.1715-1728.
57. 周諺鴻、顏郁航、王志綱，「電信與車輛偵測大數據應用-高雄市旅運特性分析與易壅塞道路預報」，*土木水利*，第 45 卷，第 5 期，頁 75-81，民國 107 年。
58. Huang, Z., Xia, J., Li, F., Li, Z., Li, Q., “A Peak Traffic Congestion Prediction Method

- Based on Bus Driving Time”, *Entropy*, Vol.21, No.7, 2019, pp.709.
59. 徐志榮, 「預測交通需求之分佈與數量-基於多重式注意力機制之 AR-LSTMs 模式」, 國立中央大學資訊工程學系軟體工程碩士班碩士論文, 民國 108 年。
60. Yang, B., Sun, S., Li, J., Lin, X., Tian, Y., “Traffic Flow Prediction Using LSTM with Feature Enhancement”, *Neurocomputing*, Vol.332, 2019, pp.320-327.
61. Tian, Y., Pan, L., “Predicting Short-Term Traffic Flow by Long Short-Term Memory Recurrent Neural Network”, 2015 IEEE International Conference on Smart City/Socialcom/Sustaincom (Smartcity), 2015, pp.153-158.
62. Chandra, S.R., Al-Deek, H., “Predictions of Freeway Traffic Speeds and Volumes Using Vector Autoregressive Models”, *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, Vol.13, Iss.2, 2009, pp.53-72.
63. Kamarianakis, Y., Gao, H., Prastacos, P., “Characterizing Regimes in Daily Cycles of Urban Traffic Using Smooth-Transition Regressions”, *Transportation Research Part C: Emerging Technologies*, Vol.18, Iss.5, 2010, pp.821-840.
64. Hochreiter, S., and Schmidhuber, J., “Long Short-Term Memory”, *Neural computation*, Vol.9, Iss.8, 1997, pp.1735-1780.
65. Mnih, V., Heess, N., Graves, A., Kavukcuoglu, K., “Recurrent Models of Visual Attention”, *Proceedings of The 27th International Conference on Neural Information Processing Systems*, Vol. 2, 2014, pp.2204-2212.
66. Bahdanau, D., Cho, K., Bengio, Y., “Neural Machine Translation by Jointly Learning to Align and Translate”, 2014, ArXiv: <https://arxiv.org/abs/1409.0473>
67. Lewis, C. D., *Industrial and Business Forecasting Methods London*, Butterworth Scientific Publishers, 1982.
68. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., “Attention is All You Need”, 31st Conference on Neural Information Processing Systems (NIPS 2017), 2017.

