

國立陽明交通大學

材料科學與工程學系所

碩士論文

基於深度學習進行電池性質預測

Battery property prediction by deep learning networks

研 究 生：許家維 **Chia-Wei Hsu**

指導教授：鄒年棣 **Nien-Ti Tsou**

中 華 民 國 一 百 一 十 年 八 月

基於深度學習進行電池性質預測

Battery property prediction by deep learning networks

研 究 生：許家維 Student：Chia-Wei Hsu

指導教授：鄒年棣 Advisor：Nien-Ti Tsou

國 立 陽 明 交 通 大 學

材 料 科 學 與 工 程 學 系

碩 士 論 文

A Thesis

Submitted to Department of Materials Science and Engineering

College of Engineering

National Yang Ming Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Materials Science and Engineering

August 2021

Hsinchu, Taiwan, Republic of China

中 華 民 國 一 百 一 十 年 八 月

基於深度學習進行電池性質預測

學生：許家維

指導教授：鄒年棣 博士

國立陽明交通大學材料科學與工程學系研究所碩士班

中文摘要

鋰離子電池作為常見的儲能設備，廣泛應用於終端設備上且藉由電池管理系統進行監控確保電池老化程度仍可應付工作所需。然而電池在使用初期並無明顯老化特性的反應，因此對於使用過的電池無法很好評估預期壽命以至於材料的浪費或設備的異常(Early failure)。本研究利用時序資料連續性進行資料擴增更同時對神經網路潛空間進行正則化，並透過包含篩選器與預測器的神經網路架構在僅有少量循環的量測數據下，預測電池產品壽命、剩餘使用壽命、充電所需時間、放電時的電壓電量變化曲線等。其中，僅測量一個充放電完整循環的數據，就能提供僅有57週期方均根誤差的產品壽命預測。本研究亦同時引入注意力機制於此框架中達成僅使用若干個循環的測量資料便可預測整個電池的產品週期放電電量、放電功耗等特性。

關鍵字： 鋰離子電池、老化因子、剩餘壽命、深度學習、特徵篩選、時序資料處理

Battery property prediction by deep learning networks

Student : Chia-Wei Hsu

Advisors : Dr. Nien-Ti Tsou

Department of Materials Science and Engineering

National Yang Ming Chiao Tung University

ABSTRACT

As a general energy storage device, lithium-ion batteries are utilized broadly in edge devices. Battery management systems monitor and ensure the degradation level is still sufficient to meet operational requirements. However, the aging characteristics are obscure at the early stage, which impedes evaluating the remaining useful life accurately of used batteries, causing wasting or abnormalities of equipment. In our research, we propose a method regularized latent space of neural networks with data augmentation. Finally, the model consisting of feature selectors and predictors enables the forecast of battery end of life, remaining useful life, charging time, and voltage v.s. capacity curve during discharge based on only a few cycles of measurement data. Furthermore, the models can estimate the End of life with only 57 cycles of root mean square error based on one cycle data. Moreover, the attention mechanism is introduced in this framework to predict the whole product life of discharge power and discharge power consumption using only a few cycles of measurement data.

Keywords: lithium-ion battery, Aging factor, Remaining useful life, Deep learning, Feature selection, Time-series data processing

目錄

中文摘要	iii
英文摘要	iii
目錄	v
圖目錄	vii
表目錄	x
第一章 緒論	11
1.1 研究動機	11
1.2.1 電池老化現象	11
1.2.2 電池特性的預測	12
第二章 文獻回顧	14
2.1 研究基礎	14
2.1.1 時序訊號處理	14
2.1.2 機器學習應用於時序資料	15
2.2 相關研究	17
2.2.1 基於數據驅動預測電池壽命	17
2.2.2 使用端到端(End to End)深度學習進行電池剩餘壽命預測	19
第 2 章 第三章 理論與研究方法	20
3.1 數據整理和清洗	20

3.1.1 數據介紹	20
3.1.2 特徵工程	21
3.1.3 訓練策略	25
3.2 時序數據上的關聯性	27
3.3 運用語言模型理解電池時序資訊	33
3.4 特徵篩選器和預測器	37
3.5 訓練方法	40
第四章 結果與討論	44
4.1 預測電池產品壽命	44
4.1.1 數據量對神經網路效果的影響	48
4.2 滑動式預測	49
4.2.1 預測餘剩壽命	49
4.2.2 預測電壓衰退狀態	52
4.3 驗證於多段快充策略數據	53
4.4 神經網路即時效能測試	54
第五章 結論與未來展望	58
參考文獻	59
附錄一 神經網路架構	65

圖目錄

圖 1 3 種常見用於時序資料處理的神經網路架構，圖(a)遞迴神經網路、 (b)卷積神經網路、(c) 時間卷積網路，其中藍色代表是運作單位或 是卷積核，而黃色代表是經由單層神經網路處理過後的輸出。.....	17
圖 2 不同電池的 $Q_{100}(V)$ 和 $Q_{10}(V)$ 曲線，其中圖(a) $EoL=335$ cycles (b) $EoL=1836$ cycles	18
圖 3 APR18650M1A 的規格表，充電採用多段式恆定電流充電(CC)並 在建議電壓(3.6V)轉換為恆定電壓充電(CV)，而放電過程則是採用 固定電流 4.4A(4C)的 CC-CV 使用策略。	20
圖 4 電池充電電流與電池電量(state of charge)圖，其中 $C1, C2, C3$ 和 $Q1$ 是可變動的，而圖(b)中的 $C4$ 則是快速充電策略的相依變數 (dependent variable)，會限制充電策略需要 10 分鐘內達到 80%的出 廠設計電荷量。	21
圖 5 溫度記錄異常的電池，線條顏色分別該數據來源於第幾個循環。 ...	22
圖 6 三種測試集篩選策略，其中 s 代表樣本時序起始點， a 代表樣本包 含幾次循環的資訊。	27
圖 7 使用最後填充方式進行資料擴增。	29
圖 8 顯示 1000 條隨機產生的線段藉由梯度下降法在第 j th擬合下的 w, b 誤差，其中黑點為同一條線的收斂路徑。	30
圖 9 不同方法下的訓練過程中的測試集的方均誤差，此誤差僅使用第	

一步的梯度下降法做為預測資訊。	32
圖 10 不同時序長度樣本對於訓練過程中的梯度貢獻.....	33
圖 11 (a)時序資料離散化程度在神經網路表徵距離關係(b)不同循環與 第一百個循環在神經網路表徵上的距離。	36
圖 12 充電時間於每次循環的變化，其中每條線都代表一顆電池。	38
圖 13 特徵篩選器與預測器於本研究框架的示意圖。	39
圖 14 貝葉思最佳化神經網路結構超參數的過程，紅色虛線為當前搜尋 下最佳架構的效能，而黑色點則代表該次搜尋下架構的效能。	41
圖 15 在 5-fold cross validation 下，神經網路對於輸入循環數對於誤差 變化的趨勢圖。	45
圖 16 神經網路預測效果(a)(b)為僅使用第一次循環數據所下的預測與 量測值散點圖，其中誤差使用 MAPE 呈現於圖例中。(c)(d)在測試數 據集中挑選最少和最多老化程度的電池並分別使用第一次、前五次、 前一百次循環數據進行電壓電量預測。	46
圖 17 歷經 20 次不同條件下訓練後的誤差成果圖，其中虛線為平均成 績，而淺藍色區域為誤差分布範圍。	49
圖 18 完全剩餘壽命模型於測試數據集下的誤差著色圖，其中包和大部 分電池所具有的三個範圍，分別為(1) 早期 (1~100循環)、(2) 中期 (4~500循環) 和(3) 晚期(7~800循環)	51

圖 19(a)(b)(c)為測試數據集的誤差著色圖，其中誤差根據右上角的色條著色，若超過 4%則使用黑色標記，(d)為不同階段下進行整個生命周期的最大放電功耗預測曲線，實線為量測值，點則根據不同顏色代表不同階段下的預測值。.....	53
圖 20 完全模型對新式(6 段恆定電流)快充方法的預測效果圖，分別以 (a)整體數據和(b)依據充電策略分別著色呈現。.....	54
圖 21 放電模型中特徵篩選器 1(a)、2(b)架構圖.....	65
圖 22 放電模型中預測器 1 架構圖.....	66
圖 23 放電模型中預測器 2 構圖.....	67
圖 24 完全模型中特徵篩選器 1(a)、2(b)架構圖.....	68
圖 25 完全模型中特徵篩選器 3(a)、4(b)架構圖.....	69
圖 26 完全模型中預測器 3 架構圖.....	70
圖 27 完全模型中預測器 4 架構圖.....	71
圖 28 完全剩餘壽命模型中特徵篩選器 1、2、5 架構圖.....	72
圖 29 完全剩餘壽命模型中特徵篩選器 3、4、6 架構圖.....	73
圖 30 完全剩餘壽命模型中預測器 5 架構圖.....	74
圖 31 完全剩餘壽命模型中預測器 6 架構圖.....	75

表目錄

表格 1 三個神經網路的規格表.....	40
表格 2 本研究框架下，神經網路於不同條件下的誤差整理圖。	47
表格 3 Discharge DNNs 規格表	55
表格 4 Full DNNs 規格表	56
表格 5 Full RUL DNNs 規格表	57

陽明交大
NYCU

第一章 緒論

1.1 研究動機

可充電電池在現代工業設備上皆有廣泛的應用，從移動設備到不斷電系統皆對其有著強烈的需求。這些電池設備上無時無刻都有龐大的能量流通[1]，然而監測設備僅可量測電流、電壓和溫度等外部特徵參數可進行診斷，電池充放電過程中有繁雜的化學反應和物理機制影響其效能，以上主要會受下列幾點影響：(a)電池的化學反應(b)電池的設計與封裝(c)製造參數的調整(d)充電放電的歷程。其中，影響因子(d)會顯著影響到已使用電池的狀態。而電池管理系統(Battery Management System, BMS)能否用少量資源精準判斷出已使用電池的狀態則會帶來巨大的經濟利益和環境效益[2]。

1.2.1 電池老化現象

電池在歷經反覆的循環負載後，最直接的老化現象會體現在電池可放電的電量下降，並通常以此衡量電池是否壽命終結。一般根據經驗法則採用當前放電容量小於設計容量的 80%作為判斷依據，也就是 Stat of health(SOH)=80%作為分界。當電池可放電容量小於該閾值時電池完整放電量會急遽下降。在此基準下有兩個很重要的指標分別是 End of life(EoL)和 Remaining Useful Life(RUL)分別為產品壽命終結的使用週期和剩餘使用週期。

老化過程不僅會使電池可放電電量減少，同時放電電量狀態 Stat of Charge (SOC)所對應的電壓值也會有所下降。這將會造成對工作電壓敏感的設備不穩定外，也會造成 BMS 對於電池當前電量判斷失準，進而產生過度放電、溫度異常

等。此電壓下降程度與老化現象，在 Zhu[3]的研究中提出其對應的動力學性質與對應的化學機制，而在 SEI(Solid-electrolyte interphase)的成長是造成老化現象的關鍵[4]。另一方面，老化帶來的內阻 (Internal resistance)上升也會在充電過程中致使同 SOC 下電壓的提升[5]，帶來提早進入恆壓充電致使完全充電所需時間大幅上升的問題。

1.2.2 電池特性的預測

在電池的性質分析與預測上，有大量研究引入機器學習等方法嘗試歸納或進行預測電池的特性與壽命。這類數據驅動(Data-driven)方法除了上述提及的(a-d)問題外，還存在著數據收集與使用情景不同問題。藉由背景知識合理地選取人為特徵，能有效幫助數學模型更直接與預測目標建立連結，達到在少量數據集下有更強健(robustness)的性能。針對 BMS 系統而言，研究[6]表明有五類特徵與老化指標高度相關，且僅使用電流與電壓的行為狀態即可確認，其分別為 F₁:放電停止後一秒電壓數值的上升量(電池放電內阻的體現)、F₂: 放電停止後一百秒內電壓數值的上升量(電池動態極化的現象)、F₃:進行恆流充電時一秒內的電壓數值的上升量(電池充電內阻的體現)、F₄:恆流(CC)充電電量與恆壓(CV)充電電量之間的比值(電池老化的直接呈現)和 F₅:整段恆流(CC)放電的電量(體現理想工作區間餘剩電量)。其中 F₂ 在其他研究[7]中，就藉由分析此特徵匹配電化學阻抗頻譜法(electrochemical impedance spectroscopy)方法大幅減少估計電阻值所需時間。

然而，上述方法可能面臨到挑選特徵不能完整表達電池狀態的風險，且隨著可訓練樣本的增加，此風險也隨之增加，所以有些研究採用原始數據(Raw data)搭配可學習的特徵擷取器取代人為的特徵挑選。在 Roman 的研究中[8]就利用隨機森林(Random forest)機器學習算法對 Raw data 進行特徵蒐集。另外也有不採用

任何降維(Dimension reduction)投影直接使用放電容量的變化作為特徵，並藉由蒙地卡羅演算法(Monte Carlo method)推測可能的電量下降的可能區間[9]的研究。除上述挑戰外，在利用 LSTM (Long Short-Term Memory)進行老化時放電容量預測的研究[9]中指出，鋰離子電池(Lithium-ion batteries)的剩餘壽命為機率分布函數(Probability distribution function)，所以基於準確數值的預測皆會受到一定程度的不確定性造成學習過程有所偏差，如果當不確定性過大可能導致模型崩潰(Model collapse)。但該研究認為退化特徵可能需要數百個充放電週期體現，所以透過觀察多個循環下老化數據的軌跡並以此降低不確定的程度；除此之外，在遞迴神經網路的層與層之間增加適當的 Dropout[10]可以避免此不確定性或是目前量測技術帶來的誤差。

陽明交大
NYCU

第二章 文獻回顧

2.1 研究基礎

2.1.1 時序訊號處理

- 時序擷取

在常規的預測模型中，輸入的時序資料須具有固定的長度，才能滿足使資料集中全域特徵的學習和對齊。看似簡單的操作對於後續模型的學習十分重要，在深度學習應用於心電圖(Electrocardiography)狀態的分類[11]上都有利用分割的技巧致使全域的特徵對齊達到良好的模型效能。然而除了定義最大時序長度外，多數低於最大時序長度樣本也需要經過處理滿足模型需求。在自然語言處理(Natural Language Processing)中，會藉由填補截止符號區隔有意義的時序區域；在訊號處理領域中，則會考慮填充 0 值或是最後一個時序的值直到樣本滿足最大時序長度。

- 數值校正

在訊號處理上，由於需要長時間的偵測目標特徵並加以保存作為訓練資料，因此可能會出現以下問題需要校正確保模型的合理性。例如，數據可能會受到監測設備的品質影響出現訊號偏移(Signal drift)的狀況，或是數據帶有一定的動態特性且記錄過程具有隨機的噪音。這時常見的做法是利用卡爾曼濾波(Kalman filter)[12]，進行記錄數據的校正。然而該方法是需要參數調整的，換言之，可能在數值校正的過程中引入不必要的人為誤差造成學習偏差。

隨著深度學習領域的發展，深度神經網路可透過多種輔助方法，直接從

原始資料(Raw data)學習特徵並同時獲得很高的穩健性與精確度，這也是本研究嘗試克服的難點。

2.1.2 機器學習應用於時序資料

在長時序任務中，數據特徵可能來自於比較前後時序之間的關係(局部特徵)或是特定時間點上的變化(全域特徵)。若展開時序上所有時間點的資訊並不加修飾地連接所有數據，可能會面臨模型複雜度過高或脫離時序概念等問題。

以下為三種常見的方法，可以建立保有時序關係的預測模型。

- 自迴歸模型(Auto regressive model)

該方法是將歷史軌跡作為預測當前目標變化的特徵。透過建立當前時序和歷史變化量的因果關係，大幅降低了上述提到全連接特徵導致的問題，並在回歸過程中加入高斯噪音(Gaussian noise)增加線性回歸的穩健性。其預測方式為：

$$X_t = c + \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t \quad (1)$$

其中 X 為預測值， c 為常數， ε 為一個平均值為 0 標準差為 1 的隨機誤差， φ 為每個時間步伐的權重， p 為該迴歸模型的階數。該計算模型曾被應用於預測 SoH 的研究[13]。

- 遞迴神經網路(Recurrent neural network)

在類神經網路中，遞迴神經網路是常見用以處理時序任務的模型。因為其具有類似前者自迴歸模型的資料輸入方式，但將只有線性性質的一個權重轉換為具有非線性特性的激勵函數(Activation Function)，整個運作過程可參考示意圖 1 (a)。然而隨著資料的時序長度增長，神經單元需要適時遺忘不相干的特徵

並記憶強相關的特徵，所以 LSTM[14]神經單元藉由遺忘閥(Forget gate)、輸入閥(Input gate)和輸出閥(Output gate)達到上述功能。但因為處理時序過程具有遞迴關係，所以平行化難度相較其他類神經網路高，且訓練時容易造成梯度消散 (Vanishing gradient) 或是梯度爆炸 (Exploding gradient) 等狀況。

- 卷積神經網路(Convolutional Neural Networks)

卷積神經網路是透過多個卷積核(Kernel)對時序資訊進行傅立葉轉換 (Fourier transform)，透過比較卷積核內的資訊得到相對應的頻域特徵。接著依時序方向掃描完整段資料，藉由多個卷積核進行特徵匹配，同時藉由多層卷積層堆將單一卷積核的局部特徵推廣為全域特徵，運作方式可參考示意圖 1 (b)。因為卷積核掃描同時具有權重共享的特性和卷積操作可高度平行化，所以此架構相較於遞迴神經網路更為泛用。然而，此種架構卻丟失時序之間的因果關係，所以後續有改進的 TCN(Temporal Convolution Network)[15]。其結合具時序概念的因果卷積(Causal convolution)和能彈性調整感受野大小(Receptive Field Size)的卷積擴展(Dilated convolution)。其中網路架構單位從神經元變成一個殘差[16]功能塊，如示意圖 1 (c)所展示。Hong et. al[17]的研究結合多個方法優點後，其結果為現今本領域最佳。因此，本研究也採用此高效的神經網路層搭建出神經網路模型。

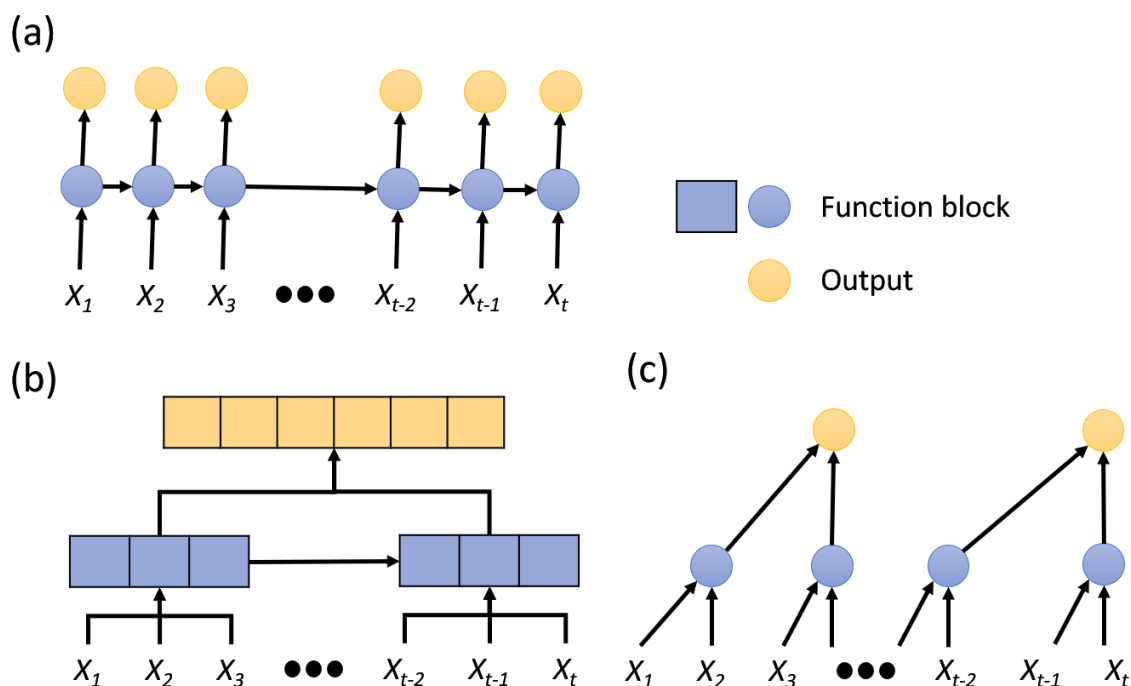


圖 1.3 種常見用於時序資料處理的神經網路架構，圖(a)遞迴神經網路、(b)卷積神經網路、(c) 時間卷積網路，其中藍色代表是運作單位或是卷積核，而黃色代表是經由單層神經網路處理過後的輸出。

2.2 相關研究

2.2.1 基於數據驅動預測電池壽命

Severson 等人的研究[18]基於市面上的常見的電池進行了不同快速充電方法 (電流 $\geq 4.4\text{A}$)的老化測試。研究中總共測試 124 顆磷酸鐵鋰(Lithium iron phosphate/graphite)電池，循環充放電次數高達 10 萬次，產品生命週期從 150 到 2300 個週期不等。除製備數據集外，該研究也建立可以在電池並無明顯老化特徵時(循環週期 ≤ 100)進行預測 EoL 的模型。其採用彈性網路(ElasticNet[19])是一種線性回歸器，但增加 L1、L2 的正則化(Regularization)用以防止過度適應

(overfitting)的發生[20]。

然而受限於該模型僅能處理線性關係，並無法直接將目標與複雜的原始數據做特徵連結。因此作者整理出 20 個候選的電池老化的特徵，主要包括 $\Delta Q_{100-10}(V)$ 相關(也就是放電過程中電壓對電池電量曲線在第 10 和第 100 次循環的差距，如圖 2 所示)、放電容量衰退曲線相關和其他相關特徵。其中研究發現 $\Delta Q_{100-10}(V)$ 的變異數(Variance)與 EoL 相關係數(Correlation coefficient)高達 -0.92。隨後作者根據目的與資料收集的難度提出 3 種模型，分別是從 20 個候選特徵中挑選若干個特徵搭建模型。其中這三種模型皆採用 $\Delta Q_{100-10}(V)$ 的特徵，然而其效能在測試(Testing)和驗證(validating)數據集上都無明顯效能差異，足以說明該特徵的重要性。此特徵在其他研究[21]中也被報導可能與(1)集電體的腐蝕(Corrosion of current collector)(2)過渡金屬的溶解(Transition metal dissolution)(3)電極顆粒的開裂(Electrode particle cracking)有關。

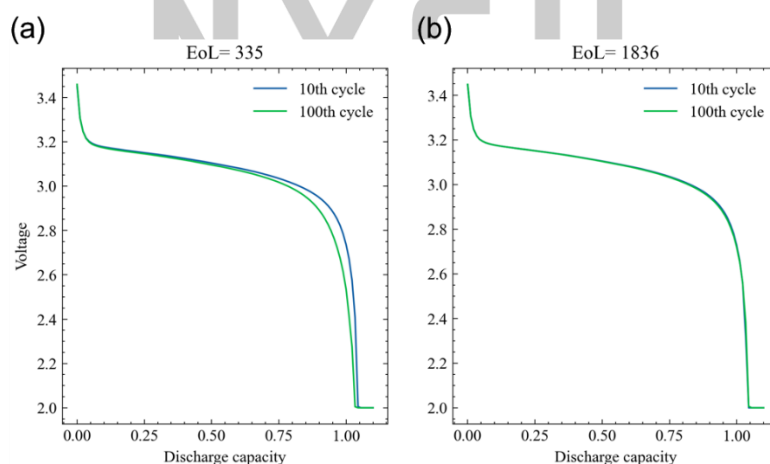


圖 2 不同電池的 $Q_{100}(V)$ 和 $Q_{10}(V)$ 曲線，其中圖(a)EoL=335 cycles (b)EoL=1836 cycles

2.2.2 使用端到端(End to End)深度學習進行電池剩餘壽命預測

Hong et. al[17]使用時間卷積網絡架構處理電池時序資料，其中包含完整循環過程的電壓、電流和溫度，即可在如此長時序的資訊中找出相關因子並進行有效推論。研究成果指出模型僅需 4 個完整循環週期(本文將使用 $a = 4$ 表示，其中代表意義演示於圖 6)的資料作為推論依據，即可預測 RUL 同時誤差僅有 76 cycles。其中此架構可以擴展至對於使用過電池的預測，跟過往採用相同數據集的模型相比加快 25 倍。除此之外，藉由將此框架的網路改為貝葉思神經網路(Bayesian neural networks)[22]，搭配損失目標函數 NLL(negative log likelihood):

$$\text{Loss} = -\frac{\log V_{\theta}(x)}{2} + \frac{(y - M_{\theta}(x))^2}{2V_{\theta}(x)} \quad (2)$$

其中 V_{θ} 、 M_{θ} 代表多個神經網路方程式所得的變異數與平均值、 x 為輸入的數據、 y 為相對應的輸出值。藉由此函數可得到各樣本統計的不確定性，若發生模型推論的不確定性過高即可切換為傳統的特徵篩選方法。這增強神經網路運用於時間電池管理系統上的穩健性，因為神經網路模型視為一個黑盒模型[23]，其可能在被攻擊[24]或是遇到離群值時，預測值將會變得不可靠。面對神經網路黑盒問題，作者也利用 t-SNE (t-distributed stochastic neighbor embedding)作為解釋神經網路的工具。將數據經過模型時，紀錄每層神經網的表徵(Representation)並加以收集，嘗試利用歸類表徵整理樣本與樣本間的關係，以表現何種特徵對於 EoL 有高度相關性。同時也可利用此歸納神經網路判斷 RUL 的準則與檢查其規則的合理性。

第三章 理論與研究方法

3.1 數據整理和清洗

3.1.1 數據介紹

在近期，電池應用與儲能領域都在積極累積數據作為機器學習的學習樣本[25]。然而，除了電池用途廣泛與使用環境的多樣性外，在累積數據過程中亦需滿足能長時間進行監測同時保有數據監測的精度，這使得樣本數增長緩慢，也使機器學習等數據驅動方法在應用上舉步維艱。2019 年，Severson[18]釋出當前最大的磷酸鋰鐵/石墨電池數據集，此數據是基於 124 顆 A123 公司製造的 APR18650M1A(參數如圖 3 所展示)，進行完全的循環充放電測試。過程中根據電容量的變化量，固定記錄時間、電壓、電流、電量和溫度，並於完全循環充放電後記錄內阻、總充電量、總放電量、平均溫度、最低溫度、最高溫度和總花費充電時間，而充放電過程都是透過恆流轉恆充電(CC-CV)的方式確保電池完全充/放電。

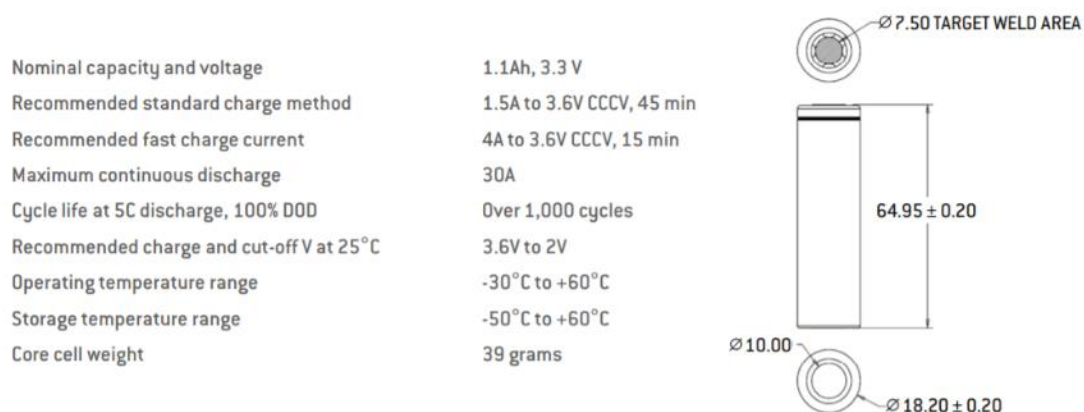


圖 3 APR18650M1A 的規格表，充電採用多段式恆定電流充電(CC)並在建議電壓(3.6V)轉換為恆定電壓充電(CV)，而放電過程則是採用固定電流 4.4A(4C)的

CC-CV 使用策略。

Severson[18]也考量到電池實際使用狀況，設計多種快速充電方式於恆流充電階段。充電方式是基於 3 個決定參數： C_1, C_2 為第一、二階段快充電流數值， Q_1 為何時切換一二階充電策略的臨界電量狀態。圖 4(a) 是電池充電時的電量與電流變化圖。而此種多步階快充方法在後續研究中[26]進一步將步階增加到五種電流狀態，但改為固定每當有 20% 電量增長就切換下一步階的狀態，如圖 4(b)。

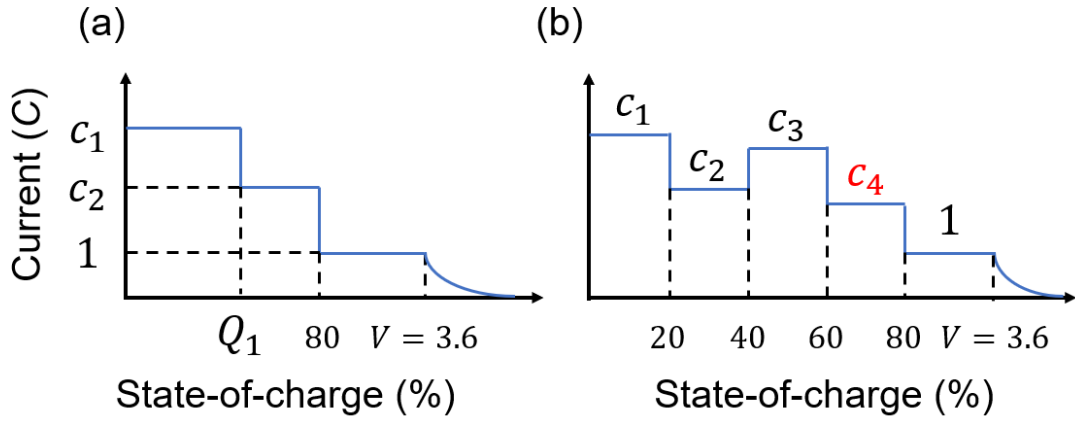


圖 4 電池充電電流與電池電量(state of charge)圖，其中 C_1, C_2, C_3 和 Q_1 是可變動的，而圖(b)中的 C_4 則是快速充電策略的相依變數 (dependent variable)，會限制充電策略需要 10 分鐘內達到 80% 的出廠設計電荷量。

3.1.2 特徵工程

為除去異常的數據集進行合理的模型訓練，我們參考了數據來源[18]的實驗記錄，刪除以下清條件的數據：(1) 未到截止電量卻停止循環測試的樣本 (2) 溫度控制失效或是錯誤紀錄的樣本 (3) 受額外電磁干擾的樣本。此外在進行資料可視化疊圖分析過程中，發現仍有部分電池，溫度紀錄仍有違常理，如圖 5 所示。因此我們將此類樣本從數據集中刪除。從一開始的 144 顆電池數據，最後有 115 個樣本可以在前一百圈放電測試中保留，95 個樣本可以在前一百圈充電與放

電測試作為數據集，而僅有 81 顆電池能在到達 EoL 前充放電皆是無異常的數據可作為訓練用。

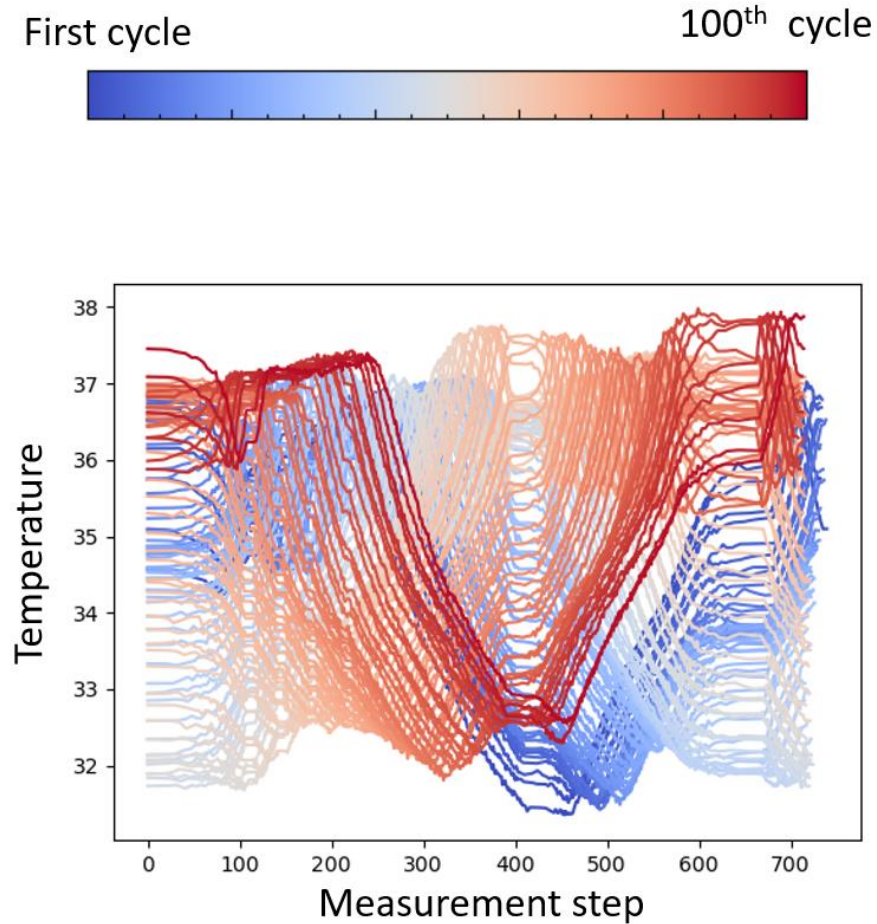


圖 5 溫度記錄異常的電池，線條顏色分別該數據來源於第幾個循環。

另一方面，原始數據集是根據偵測 SoC 的變動量作為紀錄依據，數據與數據點之間的時間間格是不一致的，且每次循環充放電的數據點數目也並非恆定。因此，我們分別將充電開始、結束、放電開始和結束的時間點做紀錄，並在開始與結束之間根據上述時間點進行線性插值，將電壓、電流、溫度和帶電量等特徵經轉換後作為循環的特徵。此外，我們也一併紀錄一次循環內總充電量、總放電量、平均溫度、最低溫度、最高溫度總花費充電時間。

接著，我們發現到各特徵之間的標準差與平均值有較大的差異，此種數據分

布狀態具有風險造成神經網路在進行梯度下降法最佳化權重時受限於有較大標準差的特徵影響。在此我們引入 Standard scaler 標準化方法，可表示為：

$$Z \equiv \frac{x - x_{\mu}}{x_{\sigma}} \quad (3)$$

其中 Z 為標準化後的數據， x 原始數據， x_{μ} 和 x_{σ} 為原始數據的平均和變異數。經由式(3)的轉換後，每 1 單位即為該特徵的標準差，除了可避免上述問題外，同時會使神經網路在激勵函數中有良好學習區間。也就是激勵值不會過大或過小，造成反向傳播算法過程中梯度消散或爆炸。

本研究有兩種型態的輸入資料分別為：單次循環下所記錄的數值，例如充放電過程的溫度變化情形；對於單次循環進行的概括描述，例如此次循環內溫度的平均值。另外有兩種預測目標的形式分別為：單一數值表明該電池的狀態，例如電池剩餘壽命；線段描述電池在充放電負載下的行為，例如電壓電量曲線。本研究建立明確的符號定義表，進一步闡述所使用的特徵與待預測項目的確切類型，如表格 1 所整理。該表整理上述提到的輸入與輸出資料類別，並在表格 1 和表格 2 列出。該表整理上述提到的類別，並列出所有相關框架會使用的資料類型與相對應的代表意義。

表格 1 本研究所使用到的符號定義表之一

Symbol	Component	Definition	Symbol	Component	Definition
n		Cycle number in the dataset, $n = 1 \dots \text{EoL}$	F_{im}^{human}		Human-picked features in cycle m
s		Number of shifted cycles from $n = 1$ used as the first cycle of input to the DNNs.	F_{1m}^{human}	P_m	Charge capacity in cycle m
m		Cycle number in the input dataset to the DNNs, where $m = 1 \dots a$.	F_{2m}^{human}	Q_m	Discharge capacity in cycle m
a		Total number of cycles used as input to the DNNs, i.e. $a = \max(m)$.	F_{3m}^{human}	T_m^{avg}	Average temperature in cycle m
t		Time step	F_{4m}^{human}	T_m^{min}	Temperature minimum in cycle m
EoL		End-of-life	F_{5m}^{human}	T_m^{max}	Temperature maximum in cycle m
RUL		Residual useful life	F_{6m}^{human}	t_m^{charge}	Total charge time in cycle m
D_n		Discharge power of cycle number n			

表格 2 本研究所使用到的符號定義表之二

x_{itm}^{dis}		Continuous- t data in discharge half-cycle m	$F_{im}^{\text{data, dis}}$		Data-driven features in discharge half-cycle m
x_{1tm}^{dis}	$Q(t)_m^{\text{dis}}$	Capacity curve in discharge half-cycle m	$F_{1m}^{\text{data, dis}}$	$\widetilde{\text{EoL}}_m^{\text{dis}} / \widetilde{\text{RUL}}_m^{\text{dis}}$	Data-driven feature for end-of-life/ residual useful life in discharge half-cycle m
x_{2tm}^{dis}	$V(t)_m^{\text{dis}}$	Voltage curve in discharge half-cycle m	$F_{2m}^{\text{data, dis}}$	$\left(\tilde{t}_{n=\text{EoL}}^{\text{charge}}\right)_m^{\text{dis}}$	Data-driven feature for total charge time in discharge half-cycle m
x_{3tm}^{dis}	$I(t)_m^{\text{dis}}$	Current curve in discharge half-cycle m	$F_{3m}^{\text{data, dis}}$	\tilde{s}_m^{dis}	Data-driven feature for age in discharge half-cycle m
x_{4tm}^{dis}	$T(t)_m^{\text{dis}}$	Temperature curve in discharge half-cycle m			
x_{itm}^{char}		Continuous- t data in charge half-cycle m	$F_{im}^{\text{data, char}}$		Data-driven features in charge half-cycle m
x_{1tm}^{char}	$Q(t)_m^{\text{char}}$	Capacity curve in charge half-cycle m	$F_{1m}^{\text{data, char}}$	$\widetilde{\text{EoL}}_m^{\text{char}} / \widetilde{\text{RUL}}_m^{\text{char}}$	Data-driven feature for end-of-life/ residual useful life in charge half-cycle m
x_{2tm}^{char}	$V(t)_m^{\text{char}}$	Voltage curve in charge half-cycle m	$F_{2m}^{\text{data, char}}$	$\left(\tilde{t}_{n=\text{EoL}}^{\text{charge}}\right)_m^{\text{char}}$	Data-driven feature for total charge time in charge half-cycle m
x_{3tm}^{char}	$I(t)_m^{\text{char}}$	Current curve in charge half-cycle m	$F_{3m}^{\text{data, char}}$	$\tilde{s}_m^{\text{char}}$	Data-driven feature for age in charge half-cycle m
x_{4tm}^{char}	$T(t)_m^{\text{char}}$	Temperature curve in charge half-cycle m			

3.1.3 訓練策略

電池作為一個廣泛的能源供應設備，需滿足各種使用情景和狀態。然而，電池內的反應機制不僅複雜，還夾雜著大量的動力學機制，致使在電池壽命問題上同時具有龐雜的影響因子且可能為相關。例如，NASA(National Aeronautics and Space Administration)[27]的數據集中，即是使用脈衝式放電的方式進行老化試驗，然而利用循環負載工具進行該測試時並無法在過程中達到完全的充電與放電，測試設備先因為電池無法穩定供電而失效。

在 Severson[18]的數據集中，電池皆使用精密觀測設備進行全充全放和老化測試且應用於電池上的使用方式皆為固定行為直到壽命終止。本研究將選用三種資料篩選方式來驗證模型的一般化(Generalization)能力。其中，樣本的屬性可歸納成 3 個自由度:不同的電池(每顆電池只有一組對應的充電策略)、樣本當前進行第幾次循環測試和測試時的量測特徵，如圖 6 所展示。第一種訓練策略是為了驗證模型在未經過訓練的電池上，是否能藉由已知的數據集推論出未經學習過的樣本。訓練過程中，我們會隨機抽取某幾顆電池的整段循環充電序列作為測試數據集，也就是圖 6(a)。第二種訓練策略，則是考量到電池充電策略對於電池壽命有著顯著的影響，因此將每個獨特的充電策略作為樣本，如圖 6 (b)所展示。前兩種策略的一般化能力測試皆是以早期前 100 筆循環數據 ($a \leq 100$)，並且皆假設電池是未使用過的 (本文將使用 $s = 0$ 表示，其中代表意義演示於圖 6)。第三種訓練策略則更貼近電池實際使用狀況，需要神經網路有能力對於已使用過 ($EoL > s > 0$) 的電池具有良好的預測能力，此策略會先隨機抽取若干個電池作為測試數據，同時對電池從初始直到壽命終結前 100 循環的數據採用滾動式訓練(Convolutional training)，如圖 6 (c)演示。此處會如此做是考量到壽命終結前 100 個循環直到壽命終結之間，電池會因為急速老化而產生大量不穩定的量測數值。這些離群值(Outlier)會造成特徵工程的失效[28, 29]，因此排除於數據集之外。此策略可幫助我們推廣模型適用範圍，將原本僅能作為出場電池的良率品質檢測，進一步提升到可對於電池使用狀況做實時監測，抑或是電池剩餘價值評估等。

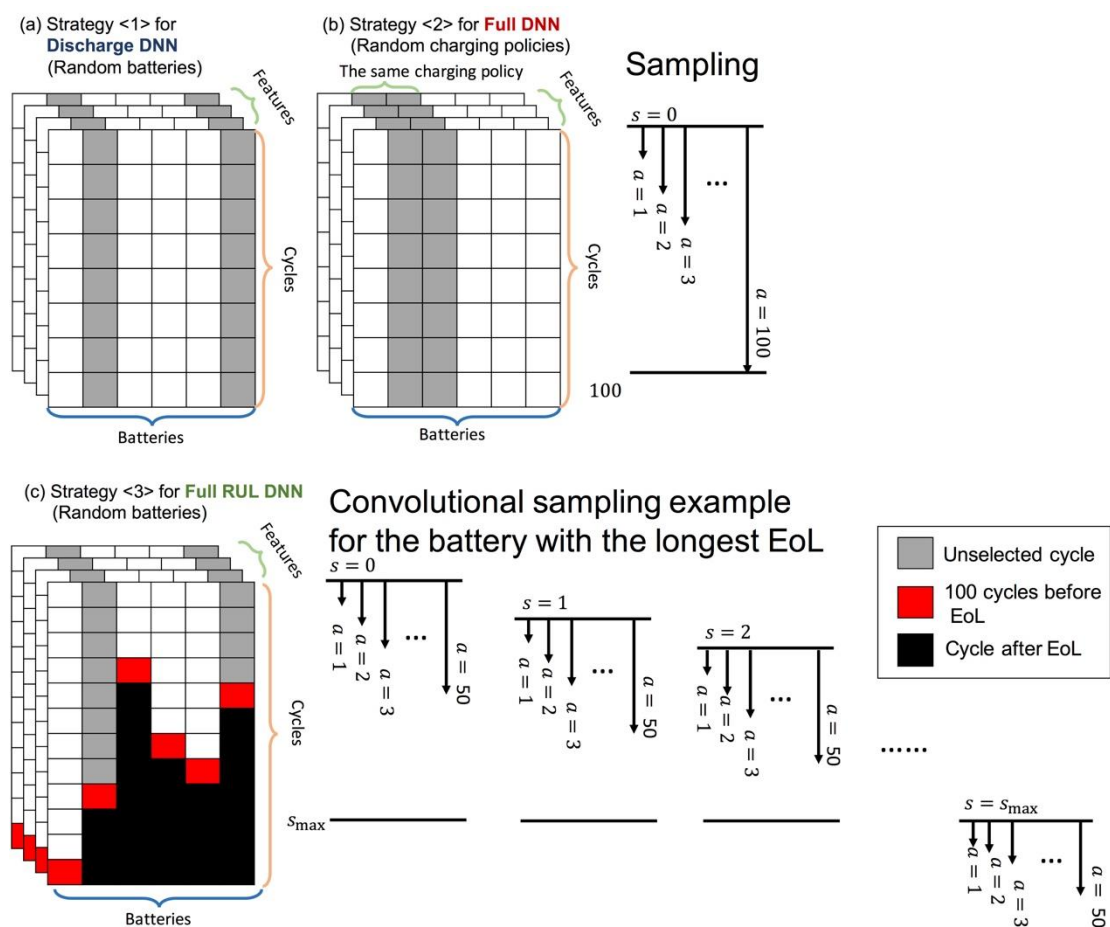


圖 6 三種測試集篩選策略，其中 s 代表樣本時序起始點， a 代表樣本包含幾次循環的資訊。

3.2 時序數據上的關聯性

考量到判斷電池的老化狀態與電化學機制不僅需要單次循環充放下的數據，也需要跨循環的多次比較每次循環中老化的特徵，所以本研究提出能整合這兩種時序資訊的神經網路架構。先利用特徵篩選器擷取單一次充放電與老化相關參數之間的相關性，也就是利用數據驅動方式藉由一次循環內的數據粗估目標數值。此方法能將一段長時序的資訊投射在目標參數的主成分(principal components /axes)上，隨後藉由多個循環的粗估結果堆疊作為預測器的輸入資訊並令模型精確的預測目標。為了驗證本研究提出框架的有效性和其對於數據學習機制的改動，我們設計相關的測

試實驗分別提出時序正則概念、檢視特徵篩選器的有效性和測量精度對於數據解釋力的影響。

如上所述，在充放電過程所造成的老化現象可顯現於單次循環週期內(循環內差異)，例如在放電過程中比較恆定電流與恆定電壓所釋放的電量可簡單得出電池的老化狀態。老化現象也可以顯現於不同次循環過程中的差異(循環間差異)，例如透過比較 $\Delta Q_{100-10}(V)$ 的變異數即可得到與 EoL 高達 0.92 的相關係數。這在多個研究中證明其良好的穩健性與預測精準度，但因為需要比較多次循環之間的特徵，所以在資料蒐集上較前者費時。

電池老化亦是動態行為，所以循環間的差異需要累積一定循環後才能藉由觀察差異軌跡判斷電池特性。電池再生現象(Regeneration Phenomena)的研究[30]中就有展示出，若有規律地打亂這種動態過程，會使原本因老化損失的最大放電容量提升。但仍會在多次循環下快速衰減至未經過電池再生現象的水準。考量到循環間差異需要透過較多週期測量數據用以降低此不確定性或隨機性造成的偏差，本研究提出可藉由將一段完整的時序拆分成前一個時間步驟、前兩個時間步驟等，隨後使用最後填充(Last padding)的方式，將每個小段時序最後有效紀錄的數據重複直到每個時序樣本的長度一致。圖 7 所展示上述的過程，其中 F 代表數據特徵。圖 7(a)表示樣本狀況為($s = 0, a = 5$)時，如何處理樣本資料並應用於放電神經網路(Discharge DNN)中；圖 7(b)則展示樣本狀況為($s = 5, a = 48$)時，如何應用最後填充方法處理資料並輸入於完整剩餘壽命神經網路(Full RUL DNN)中。藉由上述方法，我們可以將原始僅有 N 筆的資料進行資料擴增(Data augmentation)後產生 $N \times t$ 筆(t 代表每個時序中有多少步階)，達到資料筆數顯著增長的效果。且因為在訓練模型的過程中指定輸入輸出分別為每個資料擴增後的樣本與該樣本來自電池的 EoL，以至於在神經網路的潛在空間(Latent space)會連結相同電池但不同資料擴增方式的樣本，等價於該網路被施加於一正則，並在隨機批次訓練下建立更為合理的時序關係。

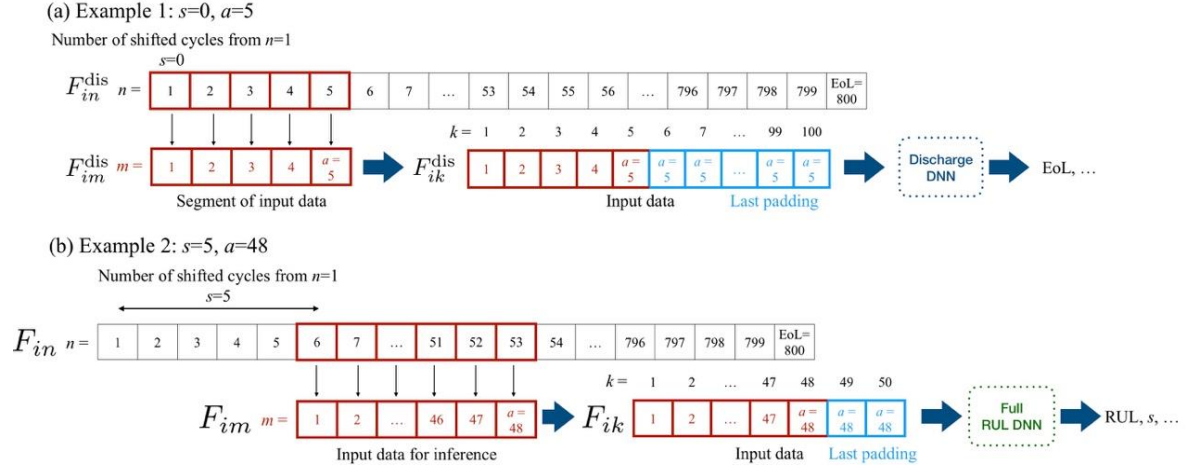


圖 7 使用最後填充方式進行資料擴增。

根據上述的假設，本研究設計一個小實驗作為參照判斷時序正則的有效性，並同時以傳統的端到端訓練方式作為對照組。我們分別讓實驗和對照組學習取代牛頓梯度下降法(Gradient descent)多次迭代，並根據前幾次梯度下降法回饋的結果快速預測出最佳化值。演算法如 Algorithm 1 所示。首先我們在 x, y 空間中隨機產生 1000 條可以被 $y = wx + b$ 方程式描述的線，其中 $0 \leq w, b \leq 10$ 且均勻分布在此範圍。同時每個方程式皆會以 $x = [1, 2, 3, 4, 5]$ 和其相對應的 y 描述並回饋給神經網路，然後每個方程式在進行梯度下降法時都有初始猜測 $w^0 = 0, b^0 = 0$ ，並以此起點開始進行線性擬合。在過程中我們會紀錄 (w^j, b^j) ，其中 $j = 1 \dots 10$ ，而 (w^{10}, b^{10}) 則是經過 10 次迭代最靠近最佳解 (w^*, b^*) 的預測數值。為了明確此項任務並非簡單的線性組合，我們將 $(w^j - w^*, b^j - b^*)$ 的結果呈現並依據 j 分別著色，圖 8 所示。其中，同一樣本(以黑點做事例)在每次進行算法迭代時，權重會於在於權重分布範圍反覆震盪直到逼近最佳解。

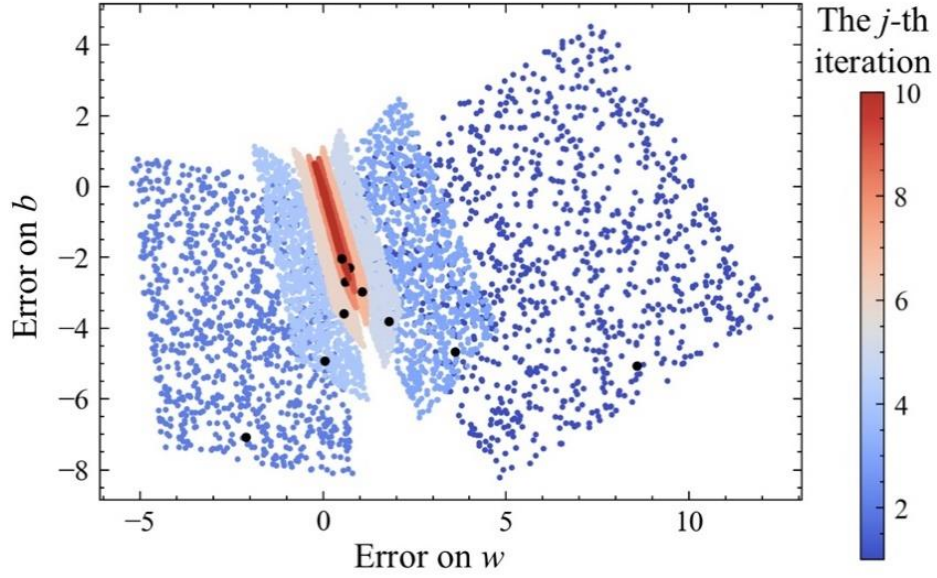


圖 8 顯示 1000 條隨機產生的線段藉由梯度下降法在第 j^{th} 擬合下的 w, b 誤差，其中黑點為同一條線的收斂路徑。

在實驗數據集產生完善後，我們將用其訓練三種候選的神經網路架構：(1) 端到端訓練方法的人工神經網路 (2) 具有時序正則的人工神經網路 (3) 具有時序正則的卷積神經網路。在此我們將這三種神經網路的參數量控制在 10^5 以確保候選模型有相似的表達能力。首先第一個模型直接將輸入 (w^1, b^1) 而 (w^*, b^*) 視為理想輸出，並在 $[2, 128, 128, 2]$ 的全連接層的架構下，且層與層之間都有執行批量歸一化 (Batch Normalization) 以減少特徵在層與層傳遞過程產生的 Internal Covariate Shift [31] 現象。第二個模型則是遵循 $[20, 128, 128, 2]$ 的全連接層排列，並且特徵皆使用 last padding 擴展到 10 個步階，其中每個步階都包含著對應的 (w, b) ，換句話說，如果僅取得該線段進行擬合的第一步數據 (w^1, b^1) ，則我們會將此結果重複 9 次作為特徵進行神經網路訓練。而第三個模型則是跟第二個模型有著相同的輸入特徵與輸出預測項，但是使用卷積神經層組合而成，其目的在於卷積具有共享權重並強化特徵與特徵之間的特性，從而加強時序正則的效果。最後我們將這三種模型的歷史訓練效能對於訓練資料歷遍 (Epoch) 次數繪圖，如圖 9 所示。其中訓練效能是透過使用測試集上的僅有的第一步數據。預測與實際值之間的均方誤差來衡量。我們可以看到具有時序正則的神經網路相較於對照組，不僅在訓練期間的收斂速度上有顯著優勢，連最終預測精度都優於對照組，且此效果隨著使用時序敏感的神經網路而提升。

Algorithm 1

Require: n : amount of sample

Require: t : time steps

For $i = 1$ to n

$w_i^* \leftarrow \text{random}(\min = 0, \max = 10)$ random generate target w uniformly

$b_i^* \leftarrow \text{random}(\min = 0, \max = 10)$ random generate target b uniformly

$f(x) = w_i^* \times x + b_i^*$ setup the objective function

$y_1 \dots y_5 \leftarrow f(1) \dots f(5)$ generate data to describe function

$w_i^0 \leftarrow 0$ set initial guess of w

$b_i^0 \leftarrow 0$ set initial guess of b

For $j = 1$ to t

$w_i^j, b_i^j \leftarrow \text{Newton}(1 \dots 5, y_1 \dots y_5)$ use Newton method update $w \& b$

End to End method

Minimize $|(f(w_i^1, b_i^1) - [w_i^* \ b_i^*])|$

Our method

Minimize $|(f(w_i^{1 \dots R}, b_i^{1 \dots R}) - [w_i^* \ b_i^*])|, \ R \in [0, t]$

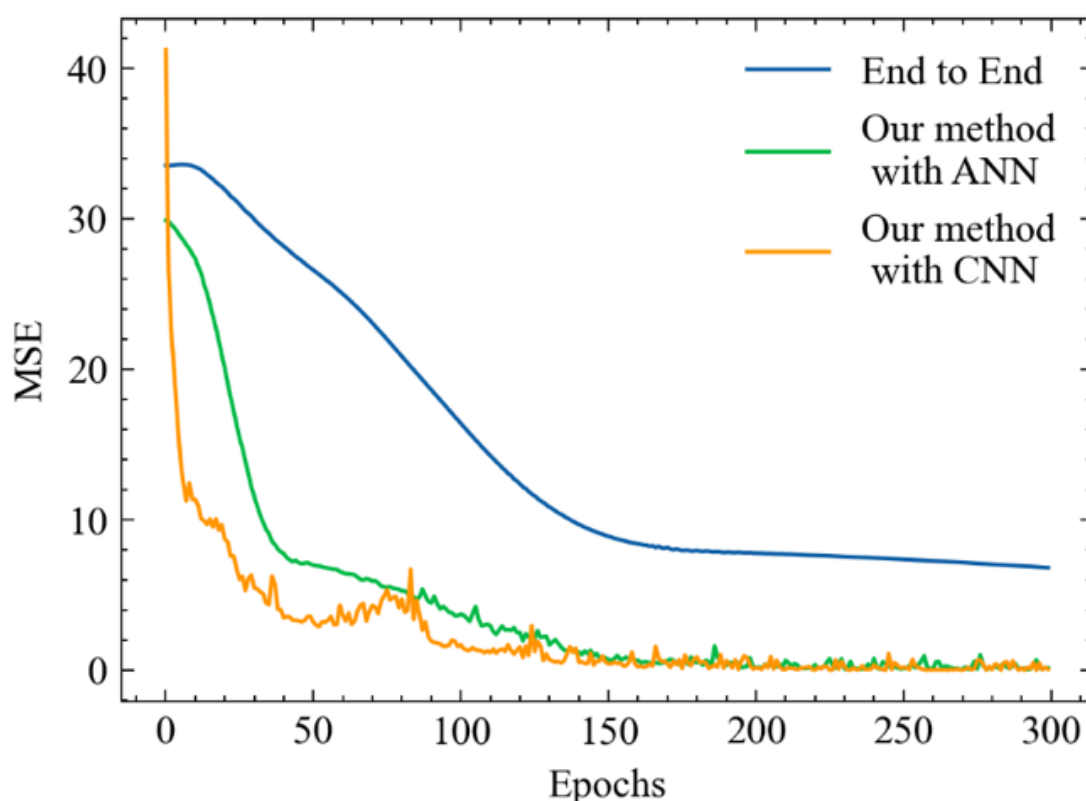


圖 9 不同方法下的訓練過程中的測試集的方均誤差，此誤差僅使用第一步的梯度下降法做為預測資訊。

除了透過最終性能的比較外，我們也觀察每批次資料進行反向傳播算法(Back propagation)時，依照不同的時序長度(w^1, b^1) ... (w^j, b^j)進行分類並記錄該長度類別下所產生的梯度總和，如圖 10 所示。可以觀察到在前期(也就是大約 100 個批次)，對於長時序(紅色線段)造成的梯度實際改變量較短時序的多，而短時序的資料(藍色線段)所造成權重更新的變化量較小；但在訓練後期，模型已充分學習長時序特徵之間的關係，轉而積極的對於短時序特徵進行參數調整，並透過短時序不全的資料進行合理的連結至長時序的特徵，且該連結在測試數據集上並未出現明顯的過擬合現象。所以此類方法相較於傳統端到端的訓練方法可以在僅利用很少的資訊下得到更準確的預測結果。這也是本研究為何能將過去需要 100 個循環測試才能較精準測定的電池產品壽命，縮短為僅使用 1 個循環測試就得到精確的預測值的原因。

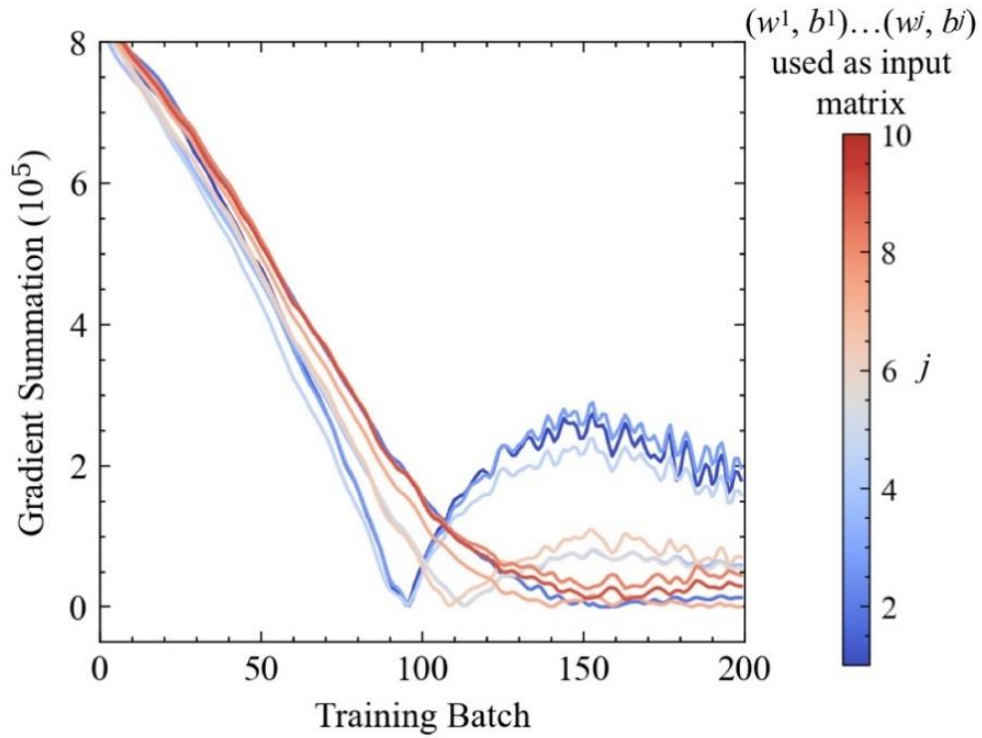


圖 10 不同時序長度樣本對於訓練過程中的梯度貢獻

3.3 運用語言模型理解電池時序資訊

上述小實驗已證明本研究可根據循環間差異有效擴增數據，然而循環內的差異，也就是單純藉由單次循環就得出能判斷電池的特性是很艱鉅的。特別當具有低壽命的樣本與長壽命樣本在測量前期幾乎不具差異，電池老化狀態在可量測循環還不明顯時。針對於我們所使用的數據集，先前分別有團隊嘗試過僅使用前 4[17]或 100[18]次循環資料進行產品壽命的預測。若要繼續突破該限制就須先探討前期甚至第一次循環過程中對於呈現樣本或老化差異的可行性。所以本節我們引入自然語言處理的方式，理解未老化數據中微小變化(Pattern)是否會在其他處理時序的模型中反映出樣本之間的獨特性。而本節的小實驗將會展示如何使用已經預訓練好(Pre-trained)的多種語言模型，在未經電池數據集重新訓練的狀態下，即可分辨電池循環間的差異，來證明確實可行。

在語言模型中有兩大關鍵部分，一個是語言訊息的嵌入(Tokenizer)，另一個則是處理詞嵌入的類神經網路。前者先簡單地將單詞轉換為獨一無二的 Word vectors[32]，接

著再考量詞句出現順序與頻率關係的 Position Encoding[33]，其編碼方式可以表示為：

$$\begin{cases} PE(pos, 2i) = \sin \frac{pos}{10000^{2i/d_{model}}} \\ PE(pos, 2i + 1) = \cos \frac{pos}{10000^{2i/d_{model}}} \end{cases} \quad (4)$$

其中 pos 表示該詞句在句子中的位置， d_{model} 則為模型的維度， i 表示詞向量的位置。詞向量的奇偶採用不同的週期函數，以此填滿整個 Position Encoding 矩陣作為詞相關訊息的嵌入(Embedding)並做為輸入資料送入變換器 (Transformer[33])。此變換器是時序處理任務中另一關鍵部分，係透過多頭注意力機制(Multi-head attention)擷取若干個可能出現於系統中的全域重要資訊並線性轉換輸出為分類或迴歸問題。這兩者搭配表徵的處理即為基於變換器的雙向編碼器表示技術(Bidirectional Encoder Representations from Transformers, BERT)[34]。此種方法在時序資料領域取得多項最先進水平的認可，同時這項能力可幫助我們探索循環內特徵的描述是否足以提供資訊差異性用以判斷電池特性。

如上所述，本研究設計一個小實驗除了觀察各個循環內的數據與其他循環之間差異性是否存在可以被描述成方程式的關係也同時探討測量精度對於時序變化數據所能提供資訊的影響。演算法如 Algorithm 2 所示。首先，我們將量測到的數據點(電量、電壓、電流、溫度)進行線性插值確保每個電池每個循環中放電部分皆有 500 個數據點描述整個充放電過程。然後將 115 個電池前 100 個循環的樣本蒐集產生 5750000 個數據點，並根據指定描述的詞彙量(Vocabulary size)使用 K-平均演算法(K-means++ clustering)[35]將整個域劃分出相應詞彙量的沃羅諾伊圖(Voronoi Diagram)。其中不同的詞彙量代表時序資料用不同解析度進行離散化處理(可對應成電池老化實驗的量測精度)，致使後續預測模型僅根據離散狀態於時間之間的排列組合分辨樣本與樣本之間的差異。然後我們會將這段離散化數據做為輸入語句餵入 ALBERT(A Lite BERT for Self-supervised Learning of Language Representations)的 Tokenizer 與變換器中。ALBERT 是利用多國語言數據集以自監督算法進行表徵訓練的知名語言處理框架，同時該模型藉由跨層參數的共享(Cross-layer parameter sharing)可以維持每一層的輸出的向量距離接近，致使我們得以在穩定狀態下分析每個樣本的表徵分布差異。

Algorithm 2

Require: v : Vocabulary size

Require: C, V, I, T : Standardized Capacity, Voltage, Current, Temperature

Require: n : number of cells

Require: a : number of cells

$ID_t \leftarrow$ K-mean clustering ($[C_t^{i=1\dots n, j=1\dots a} \ C_t^{i=1\dots n, j=1\dots a} \ C_t^{i=1\dots n, j=1\dots a} \ C_t^{i=1\dots n, j=1\dots a}]$) with maximum vocabulary size v , where i indicate the cell index and j represent the j^{th} full charge and discharge cycles.

For $m = 1$ to n

$PE_t^{i=m, j=1} \leftarrow$ Positional Encoding($ID_t^{i=m, j=1}$)

(id sequence of the 1th cycle form n cell)

$Code_t^{i=m, j=1} \leftarrow ALBERT(PE_t^{i=m, j=1})$

(Representations of encoding feature by ALBERT)

For $l = 2$ to a

$PE_t^{i=m, j=l} \leftarrow$ Positional Encoding($ID_t^{i=m, j=l}$)

(id sequence of the 100th cycle form n cell)

$Code_t^{i=m, j=l} \leftarrow ALBERT(PE_t^{i=m, j=l})$

(Representations of encoding feature by ALBERT)

$$D^{i=m, j=l} = \sum_t |Code_t^{i=m, j=1} - Code_t^{i=m, j=l}|$$

Return mean (D)

試驗中依序使用從 100 到 10000 單詞量離散化原始測量數據(也就代表電池實驗量測精度由粗至細)，再將其與該單詞量第一圈完整循環下與第一百圈循環下的所產生的表徵差異繪製成圖 11(a)。其中表徵的距離考量到離群值和穩健性等因素，我們採用曼哈頓距離來量測。由圖可以發現就算未進行電池數據集的重新訓練且電容量並無明顯的老化現象發生，仍可藉由增加描述的字彙量提升電池老化特性的辨別性，我們進一步分析在固定詞彙量下(Vocabulary size=1000)每個循環的數據在表徵上與第一百個循環的數據差異，如圖 11(b)所展示。可以觀察到任意循環與第一百圈數據的循環數差異與神經網路表徵上的差異略呈現線性關係。也就是就算在未明顯老化的循環，神經網路也能透過其中時序的排列得知樣本間的差異，值得一提的是，在圖 11(b)中 50-60th 循環的與 100th 循環距離陡增是因為溫度擾動異常造成數據錯誤所造成。本小實驗的結果證明單依循環資料包含充足的資訊可供神經網路分辨不同電池並預測效能，且可透過量測實驗數據的精度，提升預測準度。不僅如此，若將 1 至 100th 的資訊混合訓練，則在僅輸入第一個循環時，該模型可藉由轉換出可能的第二個循環...至第一百個的預測結果，並集成(Ensemble)取得最後輸出值，這做法可以有效避免在僅使用少量循環數據集進行預測所產生的不確定性(Epistemic uncertainty)[36]。

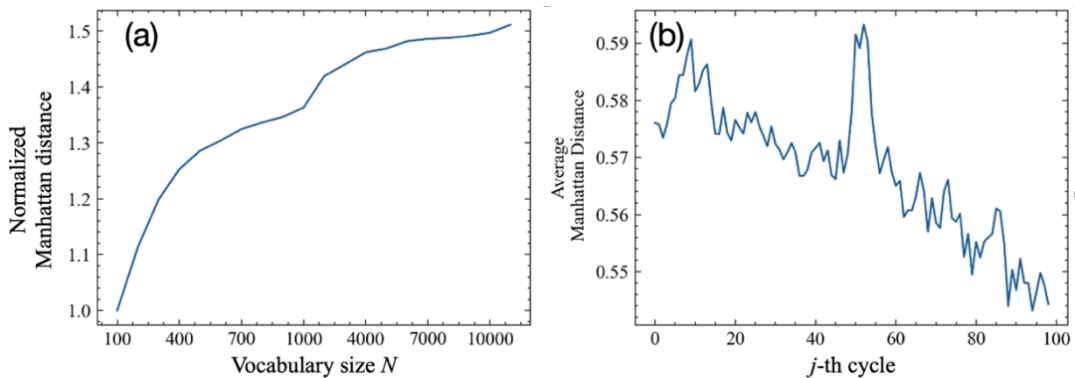


圖 11 (a)時序資料離散化程度在神經網路表徵距離關係(b)不同循環與第一百個循環在神經網路表徵上的距離。

3.4 特徵篩選器和預測器

在深度學習中，隨著數據集的完善與神經網路計算效能的上升，端到端(End-to-end)的神經網路設計架構成為主流。這是因為該架構能直接藉由一個模型即可得到目標輸出，具有可大幅簡化預測框架複雜度等的優點。但是，在分析端到端與非端到端架構的深度學習研究[37]指出，端到端存在找不到最優解的風險，並且該方法相對上要求更多數量的數據。然而，受限於鋰電池老化的數據筆數不足等原因，端到端學習在電池特性領域有所受限。因此，本研究將採用分層框架的方式，分別處理『循環內』的數據與『循環間』的數據，藉此解決兩者資訊具有不同時間單位的問題。另一方面，若數據集允許，進行多任務學習(Multi-task learning)[38]同時預測多個相關特性，不僅可以擴展模型適用範圍，且可以因為於訓練過程中引入更多資訊使得單一預測任務的精準度上也可以提升。這種透過多輸出潛在學習一些特徵表達可以視為輔助學習(Auxiliary task)，也會被本研究框架所採用。

本研究框架將由『特徵篩選器』與『預測器』構成，分別對應『循環內』與『循環間』的特徵。特徵篩選器將把每個樣本中的每個循環作為獨立樣本進行訓練，輸出則為產品壽命與壽命完結時的充電時間。其中，後者不僅可以使模型適用於限制電池充電時間的用途，例如應用於閉環式最佳化電池壽命的充電策略[26]上，且在資料分析過程中，也能顯現老化特徵於恆流充電與恆壓充電之間充電的電量比例。接著再綜合恆壓充電速度較慢等原因，可以發現電池壽命與充電時間是息息相關的，如圖 12 所呈現。所以本研究根據這兩個不同的輸出建立兩個獨立的神經網路，同時也會針對資料來源區分採用充電段或是放電段的資訊獨立訓練網路。在放電模型中會有 2 個特徵篩選器；完整模型中會有 4 個特徵篩選器；剩餘壽命模型會有 6 個特徵篩選器(多兩個分別使用充放電段資訊進行已經使用多少循環的預測)。整個框架的示意可參考圖 13 和表格 3 的總整理表，分別表述電池特性預測和與之對應的模型。其中包含產品壽命(EoL)、在壽命完結時的放電時電壓與電量曲線($V(Q)_{n=EoL}$)、在壽命完結時的充電所需時間($t_{n=EoL}^{charge}$)、任意循環 n 放電時電壓與電量曲線($V(Q)_n$)、剩餘壽命(RUL)、已被使用

的循環數(s)、任意循環 n 放電時總放電量(Q_n)和任意循環 n 放電時總放電功耗(D_n)。總而言之，本研究會有三種主要的模型架構，分別是針對僅有放電數據進行預測的放電模型(Discharge model)和能同時獲得充電與放電過程資訊的完全模型(Full model)，和可變起始點的完全剩餘壽命模型(Full RUL model)。此規劃使本研究的適用範圍擴大到已使用的電池。另一方面，預測器則需要處理萃取出來的表徵，按照最後填充的方式理解時序相關性。達到從時間概念的角度說明特徵篩選器將每個循環視為獨立樣本，讓樣本之間時間關係互相獨立，強迫網路只能透過資料的聚類在非監督狀況下重建時序關係；預測器則藉由後續對時序的處理來建立正確時間關係。

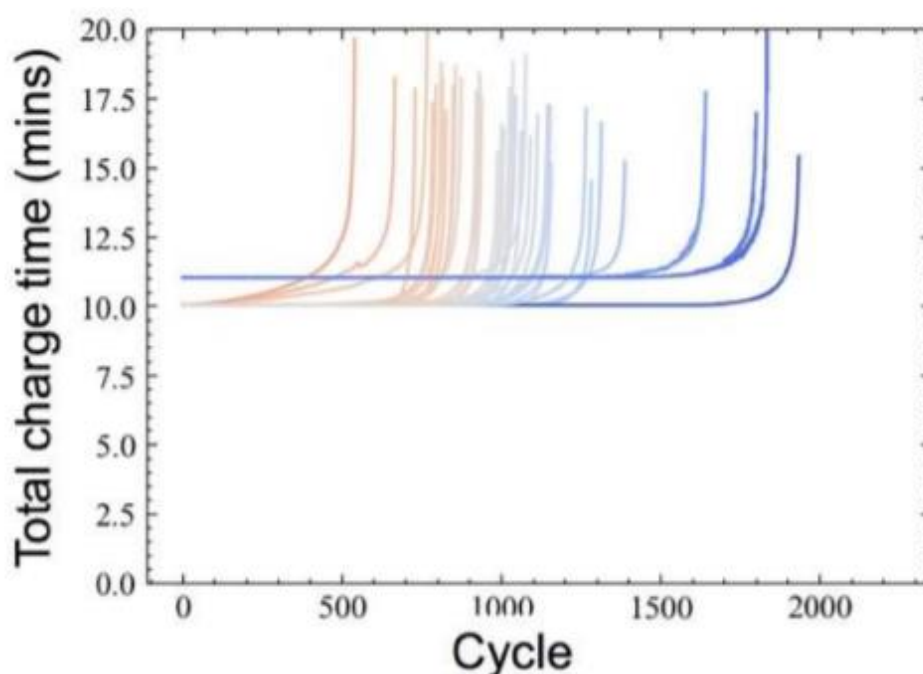


圖 12 充電時間於每次循環的變化，其中每條線都代表一顆電池。

近年來自監督學習(Self-supervised learning)領域蓬勃發展。該方法可利用資料擴增方式在未知標籤(Label)的狀態下，進行拉近『正樣本』(該樣本與該樣本進行資料擴增的副本)和遠離『負樣本』(該樣本與其他樣本)的學習達到分群。隨後僅需藉由 K-近鄰演算法(k-nearest neighbors algorithm)即可快速學習且具有與監督式學習相近的準確度。例如近期發表僅透過資料擴充即可運行的 SimCLR[39]；與不需負樣本也可學習的 BYOL[40]；抑或是藉由將表徵投影至超球面(Hypersphere)上的對比式學習[41]。其

中，在分析對比式學習的研究中[42]，提出正樣本是對於同一類別的對齊(Alignment)或親近(Closeness)，而負樣本則需拉開目標類別與資料集中其餘類別的距離。因此為了獲得最優解，需控制每個類別間的距離，致使在超球面上每個類別分佈均勻(Uniformity)。此概念也可套用在特徵篩選器中，以產品壽命的模型為例子，透過將同一電池的所有循環數據規為同一類(因為具有共通的产品壽命)，再同時疏遠具有不同產品壽命的循環數據。這樣產品壽命的分布能幫助我們控管數據於潛空間的均勻度，使神經網路可以較一般監督式學習有更好的表徵能力。

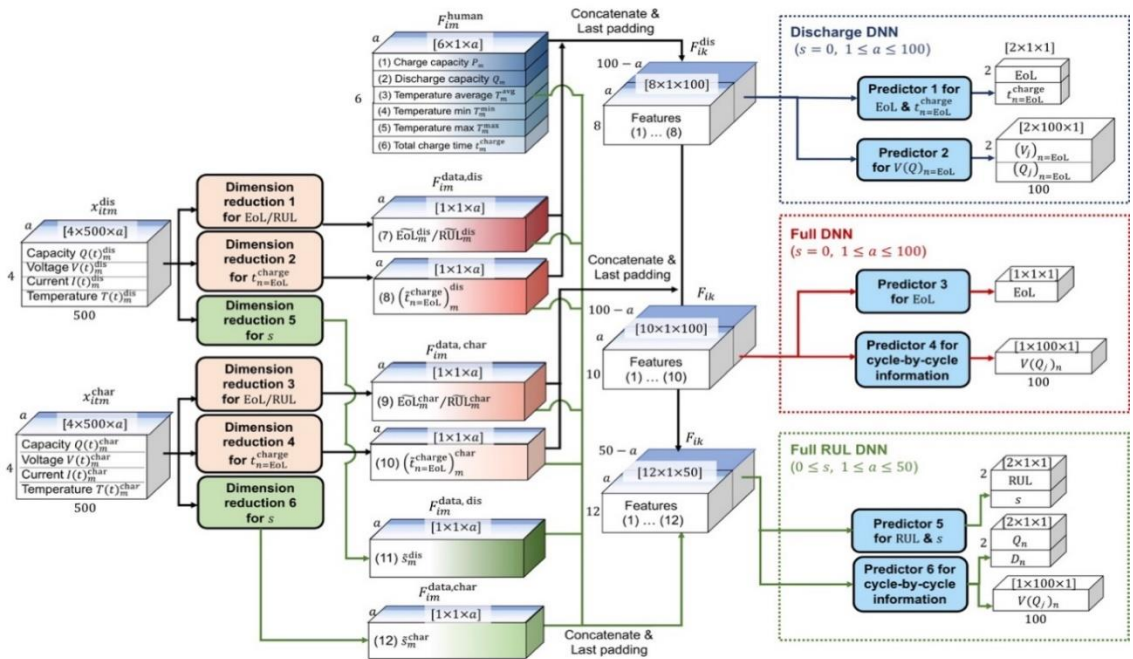


圖 13 特徵篩選器與預測器於本研究框架的示意圖。

表格 3 三個神經網路的規格表

Name of DNN	Main objectives	Auxiliary learning tasks	Testing strategy	Purposes of DNN
Discharge DNN	EoL, $V(Q)_{n=EoL}$	$t_{n=EoL}^{charge}$	<1> Unselected random batteries, only trained with 100 cycles.	Predicting for unknown batteries.
Full DNN	EoL, $V(Q)_n$		<2> Unselected random charging policies, only trained with 100 cycles.	Predicting for unknown charging policies.
Full RUL DNN	RUL, Q_n , D_n , $V(Q)_n$	s	<3> Unselected random batteries, trained with 1...EoL-100 cycles.	Predicting for unknown age of unknown used batteries.

3.5 訓練方法

我們在框架固定後，依序進行框架下的神經網路訓練。然而，神經網路的架構設計仍需要大量的試驗 (Trial and error) 才能取得良好的表現。本研究透過先建立候選的功能方塊，搭配貝葉斯最佳化[43](Bayesian Optimization)方法進行方塊的選取與詳細數據的調整。其中候選功能方塊包括(1)Inception blocks[44]:利用不同大小的卷積核整合不同視野下的資訊。(2)Residual blocks[16]:透過 skip connection 連結輸入與輸出端的資訊，在深度神經網路中仍可有效進行學習。(3)Luong-style attention layers[45]、(4)Network in network (NiN)[46]。在最佳化的過程中總共會有 80 次的試驗，每次試驗中會以貝葉斯演算法建議的超參數(hyper-parameter)進行 80 次歷遍訓練數據集的權重更新。在訓練完成後會以神經網路在訓練數據集的適應性(Fitness)作為回饋(Feedback)給貝葉斯演算法進行下一次超參數的建議。

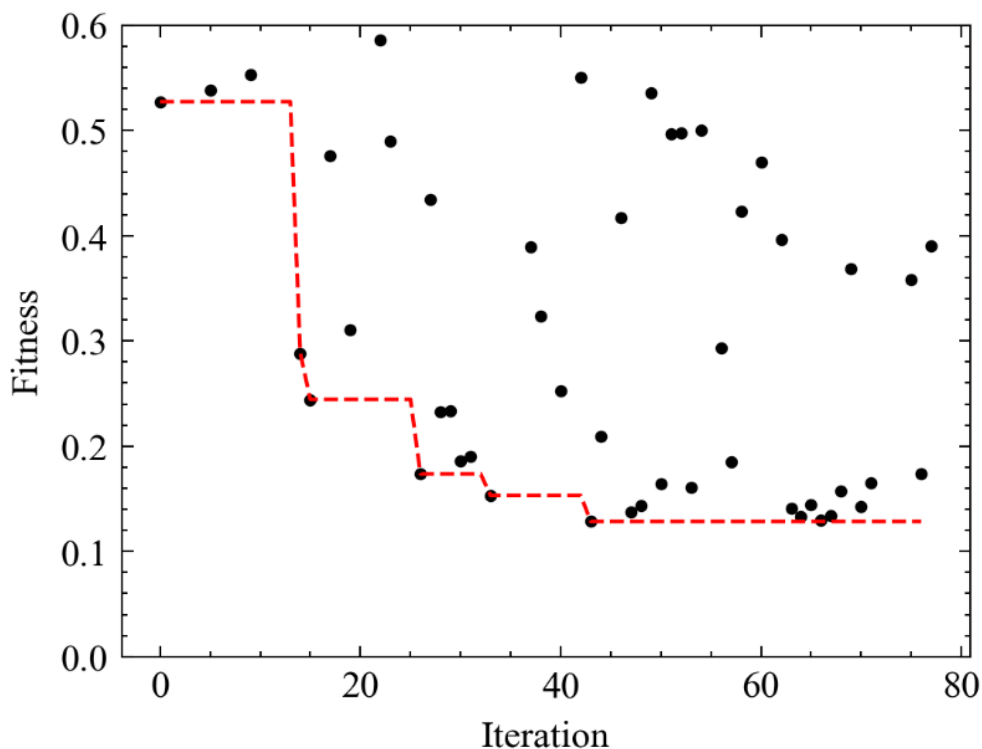


圖 14 貝葉思最佳化神經網路結構超參數的過程，紅色虛線為當前搜尋下最佳架構的效能，而黑色點則代表該次搜尋下架構的效能。

在每次訓練過程中，皆使用 AMSGrad[47]演算法作為神經網路權重的最佳化器，該方法不僅能夠自適應學習率(Adaptive learning rate)[48]，且能獲得相較於常見的自適應最佳化器更嚴謹的收斂保證。損失函數(Loss function)會因為不同預測目標特性或分布而不同。在一般的迴歸問題中，通常會採用均方誤差作為損失函數，可表示為：

$$\text{MSE}(\hat{y}_i, y_i) = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (5)$$

其中， \hat{y}_i 代表神經網路的預測值， y_i 代表數據標記數值， n 代表該次預測多少樣本。該方法會考慮誤差 >1 的樣本所造成的影響，同時誤差越大此放大的效應也越大。因此該函數具有風險會造成神經網路受離群值的影響反而偏離數據分布該有的行為。另一方面，該函數也會縮小誤差 <1 的樣本所造成的影響，這可能造成模型仍會留有部分殘差。即使該函數有眾多缺點，但仍被採納為多數神經網路的損失函數，這是因為該方法具有前期收斂快速的效果。但若預測目標的數據峰度(Kurtosis)過高，可能會多數樣本誤差皆落於小於1的區間，造成誤差過小而梯度消失，所以在預測器2和4中本研究將訓練過程的誤

差函數換成平均絕對誤差，如下所示：

$$\text{MAE}(\hat{y}_i, y_i) = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (6)$$

該函數對於所有誤差大小皆無任何縮放效果，可應用於過於集中的目標分布。而鋰電池不管在放電或是充電過程中都會有很大佔比的區間電壓幾乎不隨電量而改變(可參考圖 2)，因此在預測電壓時，本研究改用式(6)作為模型損失函數。除此之外，因為需要和相似研究[17, 18]進行準確度的比較，我們沿用其他研究所量測模型性能的指標，其中包含均方根誤差，如下方程式：

$$\text{RMSE}(\hat{y}_i, y_i) = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (7)$$

與平均絕對百分比誤差：

$$\text{MAPE}(\hat{y}_i, y_i) = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (8)$$

接著將針對能進行剩餘壽命預測模型(Full RUL DNN)的開發。此模型雖然有 81 顆電池通過資料篩選，但在這些電池的循環過程中仍存在部分不合理的量測數值，考量到這些紀錄異常或是隨機性造成的噪音，本研究引入 Huber loss[49]作為神經網路的損失函數，該函數如下：

$$\text{Huber loss}(\hat{y}_i, y_i) = \begin{cases} \frac{1}{2} \text{MSE}(\hat{y}_i, y_i) & \text{if } \text{MSE}(\hat{y}_i, y_i) \leq \delta \\ \delta \left(\text{MAE}(\hat{y}_i, y_i) - \frac{\delta}{2} \right), O.W. & \text{otherwise} \end{cases} \quad (9)$$

其中 δ 作為調和平均平方誤差與平均絕對誤差的適用範圍。該方法不僅可以藉由轉換成平均絕對誤差處理具有大誤差的樣本，使神經網路在訓練過程中不容易受離群值影響，同時該精心設計過的函數能夠使 δ 附近的值仍然可微分，也增強該損失函數的圓滑性。最後，基於電壓電量曲線在達到產品壽命時，電壓變化十分劇烈也存在大量不確定性在曲線中。這使可能在僅有電容量微小差距的狀況下，即有龐大的電壓變化量，該現象會

造成神經網路曲線擬合上的困難，因此在針對預測器需要預測電壓電量曲線的模型，本研究採用另一個指標用以衡量曲線貼合度。該指標透過將整個線段的誤差投影至線段的垂直方向，如下所展示：

$$\mathcal{M}_n = \frac{\sqrt{\frac{1}{100} \sum_{j=1}^{100} ((\Delta V_{jn}/(V^{\text{upper}} - V^{\text{lower}}))^2 + (\Delta Q_{jn}/Q_{\text{nameplate}})^2)}{\int |V'(q)_n| dq} \quad (10)$$

其中 V^{upper} 和 V^{lower} 分別為高低截止電壓(3.6V/2V)，而 ΔV 和 ΔQ 分別代表電壓電量的誤差值。

陽明交大
NYCU

第四章 結果與討論

4.1 預測電池產品壽命

在利用上述貝葉斯方法最佳化神經網路架構後，將進行模型中神經網路的權重訓練。其中訓練數據集(Training set)為對應完整數據集的隨機選取 80%的樣本，而其餘樣本被收集為測試數據集(Testing set)。也就是放電模型(Discharge DNN)有 92 顆電池可以進行訓練，同時也會有與前者互斥的 23 顆電池進行過度擬合的測試；另一方面在完全模型(Full DNN)中，因為採樣方式改為針對任意充電策略，如圖 6(b)所示，所以會在 55 種不同的充電策略中選取 11 種策略作為測試數據集。

為了呈現神經網路的預測效能，本研究繪製預測-測量值散點圖，觀察整個預測分佈範圍的狀況，同時根據樣本來源而分別標記藍色與橘色代表訓練與測試數據集，如圖 16(a)(b)所呈現。此外在圖標位置有標記該類別的平均絕對百分比誤差供參考，值得注意的是，本預測結果僅使用了第一次循環所記錄的數據進行預測，而測試數據集上的表現(6.73%與 7.78%)皆優於數據提供者所擬合的預測模型[18]，(且藉由前一百筆的循環資料，得到MAPE = 9.1%的誤差)。且在整個數據可能涵蓋區間沒有出現明顯的低估或高估狀況，說明了若納入更多可參考的資料可有效降低數據不確定性或隨機不確定性(Aleatory uncertainty)。更進一步，雖然圖中沒成現，但我們也若採用前一百次循環的資料進行預測($a = 100$)神經網路可以取得僅有 3.97%的 MAPE，其餘條件下的效能測試可參考表格 4。但相較於預測性能的提升幅度，連續進行高品質無噪音的數據紀錄更難以達成顯得不符合效益。我們根據該現象進行邊際效應的探討，透過 5 叢交叉驗證(5-fold cross validation)訓練五個神經模型後，利用與該訓練數據互斥的樣本進行預測性能的評估，成果呈現於圖 15 中。本研究針對放電模型(Discharge DNN)與完全模型(Full DNN)所交叉驗證的效能，依序採用前一、前二至前一百次循環的測試數據進行誤差量測。我們使用淺藍色區域表示預測誤差的上下界限並同時繪製黑色虛線代表這五個模型的平均效能，可以觀察到大約在前 20 個循環後模型的誤差並未有顯著下

降。

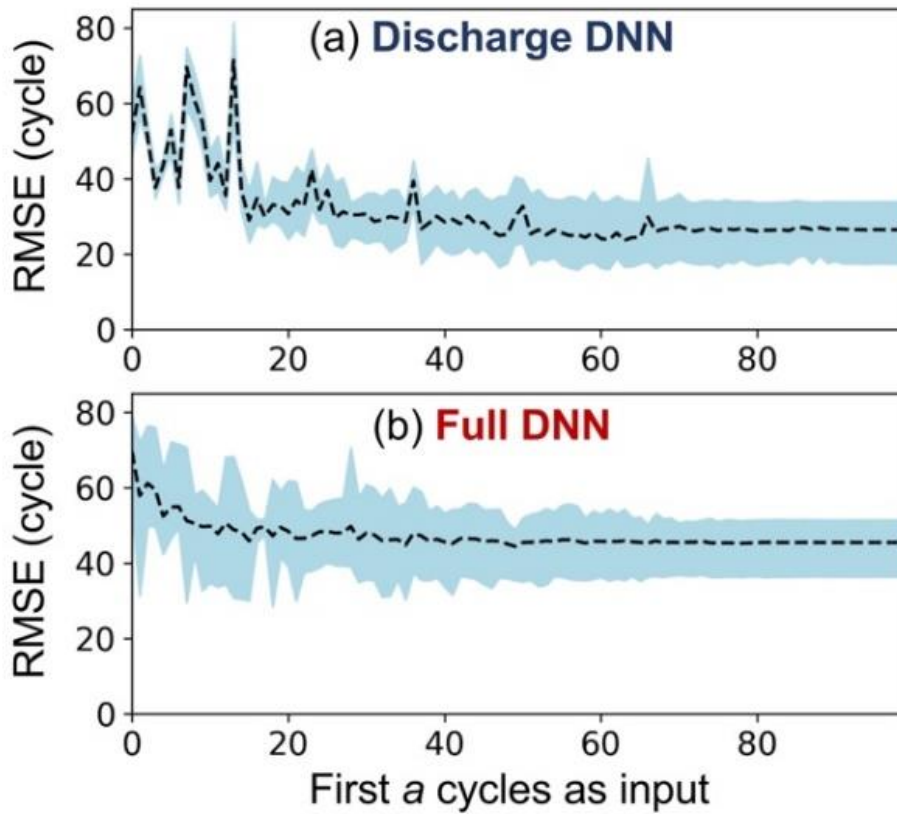


圖 15 在 5-fold cross validation 下，神經網路對於輸入循環數對於誤差變化的趨勢圖。

除了預測電池壽命外，本研究也進行了電壓精準預測，如圖 16(c)(d)所示。為了呈現預測與量測值的擬合情形，我們選用在測試數據集中老化程度最大與最小的樣本作為參考範例。其中老化程度是根據初始狀態與產品壽命時的電壓電量曲線之間面積而定，當此面積越大說明該電池電壓降低程度大和老化嚴重。電池被選定後，框架中也針對放電模型(Discharge DNN)與完全模型(Full DNN)的資訊量進行預測可行性的評估。最後預測電壓的部分在放電模型中僅顯示 $V(Q)_n$ ，也就是壽命完結時的電壓電量曲線；完全模型則因為多提供充電部分的資訊，使 $V(Q_j)_n$ 預測成為可能，也就是可以針對 n 循環(以任意百分位置表示)預測電壓電量曲線。在圖 16(d)呈現分別選取生命週期於 1、33、66、100%時的結果呈現。

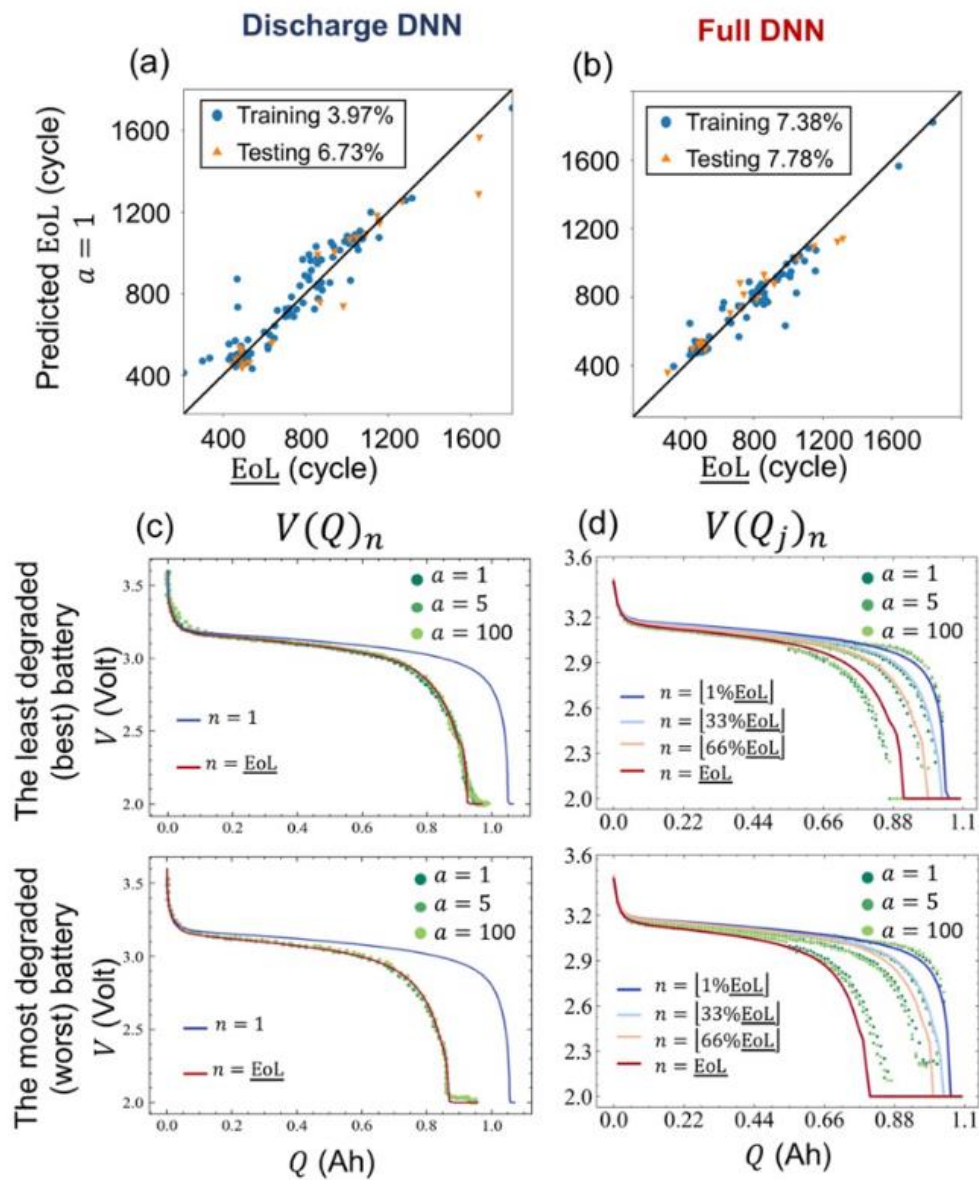


圖 16 神經網路預測效果(a)(b)為僅使用第一次循環數據所下的預測與量測值散點圖，其中誤差使用 MAPE 呈現於圖例中。(c)(d)在測試數據集中挑選最少和最多老化程度的電池並分別使用第一次、前五次、前一百次循環數據進行電壓電量預測。

表格 4 本研究框架下，神經網路於不同條件下的誤差整理圖。

	First cycle $s = 0$ $a = 1$	First 5 cycles $s = 0$ $a = 5$	First 100 cycles $s = 0$ $a = 100$	First 100 cycles $s = 0$ $a = 100$
	Train/Test	Train/Test	Train/Test	Train/Test
Discharge DNN				Discharge Model in [18]
RMSE(EoL, <u>EoL</u>) (cycles)	44/57	38/49	15/33	76/122
RMSE($t_{n=\text{EoL}}^{\text{charge}}, t_{\underline{n=\text{EoL}}}^{\text{charge}}$) (mins)	0.37/0.6	0.38/0.53	0.17/0.45	N/A
$\mathcal{M}_{n=\underline{\text{EoL}}}$	32/43	23/36	17/16	N/A
Full DNN				Full Model in [18]
RMSE(EoL, <u>EoL</u>) (cycles)	50/58	42/48	35/39	51/167
$\langle \mathcal{M}_n \rangle$	15/26	12/20	10/22	N/A

4.1.1 數據量對神經網路效果的影響

本研究是基於數據驅動(data-driven)方法進行電池使用壽命預測，並同時預測多項相關參數。此時是深度學習算法將十分依賴龐大的數據取得更精確的類神經網路[50]。在過往進行老化循環測試電池的壽命是十分耗時的，例如在本研究採用的數據上，電池平均的產品壽命為八百次循環測試，此過程需耗費至少一個月才能得出結果。另一方面，Severson 等人[18]透過後驗(Posteriori)知識先行篩選出電池老化會有的重要特徵，並使用參數較少的線性回歸模型 ElasticNet，才可達成在僅提供 41 個樣本作為訓練的條件下，藉由前一百個循環充放電的特徵判斷出產品壽命 (EoL)，得到在測試數據上有 167 個循環的均方根誤差 (RMSE)。

在上述的背景下，若能使用低數據樣本、低需求特徵，但又能高精度的預測電池特徵，對於整個電池領域會有重大的幫助，尤其是在其他種電化學反應的電池上，抑或是不同封裝或製成。所以本研究藉由控制訓練的樣本數，進一步探討樣本數對於神經網路效能的影響，我們選用放電模型(Discharge DNN)作為測試模型，並分別隨機抽取[10, 20, 30, 40, 50, 60, 70, 80, 92]的樣本個數作為訓練樣本，同時測試數據集會有[105, 95, 85, 75, 65, 55, 45, 35, 23] 樣本並互斥於上述訓練樣本。除此之外，透過考量到訓練因為生成網路參數造成的隨機性，本研究在相同訓練樣本下重複 20 次試驗並將結果取平均且同時紀錄，結果如圖 17 所示。而此試驗的誤差值是基於前一、前二到前一百次循環的數據分別得出的結果進行 RMSE 的計算。根據此研究成果，可以看到隨著訓練樣本的增大，類神經網路的誤差逐漸減小，儘管訓練樣本只有 10 個但在預測電池壽命上仍保有 160 循環次數的誤差，也展現我們提出模型具有高的強健性 (Robustness)。雖然 Severson 等人[18]可以在僅有 40 個訓練電池下藉由前 100 次循環特徵也取得 160 個循環誤差，但本研究提出的框架可分別基於 1~100 個樣本分別進行預測，顯示採用深度學習擴展的方式在小樣本下，仍可接近甚至超越線性回歸模型的效果。

除此之外，我們的研究成果能成為一個參考指標，說明在電池壽命預測上，需要多少個訓練樣本才能達到相對應的精準度。這對於要將此深度學習方法推廣到其他種類的

電池是一個很重要的基準點，例如在訓練樣本數大於 80 個後，模型的誤差大幅下降。

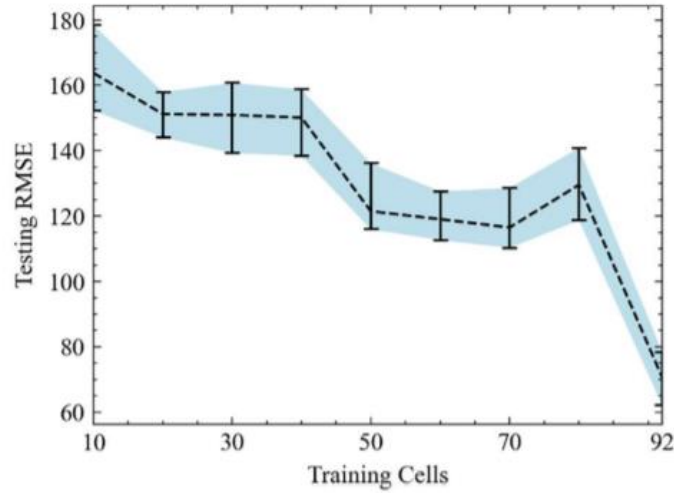


圖 17 歷經 20 次不同條件下訓練後的誤差成果圖，其中虛線為平均成績，而淺藍色區域為誤差分布範圍。

4.2 滑動式預測

本研究雖達成對於新電池的特性預測，然而若能進一步將適用範圍擴張應用於使用中的儲能設備並達到潛在風險的預測，則會顯著地提升能源利用率[2]。在上述願景下，本研究將預測的範圍拓寬至已使用過的樣本($s \neq 0$)，以下將詳細介紹本框架下的預測結果。

4.2.1 預測餘剩壽命

在完全剩餘壽命模型(Full RUL DNN)中的預測器 5，能透過結合資料驅動與人為挑選特徵並預測剩餘壽命與該狀況下電池已被使用的循環數。這是因為訓練資料集的筆數由之前的 100 個循環擴增至 EoL-100，在數據量倍增同時預測的目標域也可隨之擴增(任意狀態下的電池皆可考量)。但令人意外的是，本研究提出的神經網路的適應能力極強，以至於在如此複雜的問題上仍有相當傑出的效果。在呈現神經網路的效果上，因為該任務具有 2 個描述狀態(s, a)，所以我們根據這兩個變數建立座標軸，並將符合該狀況的測試樣本計算出神經網路預測(\hat{x})與量測值(x)的均方根誤差進行相對應的著

色，如圖 19 所描述。其中圖的 x 軸為已使用過的循環數 s (模型並不知道此值)。在圖中分別以(a)(b)表示預測 \widehat{RUL} 和已使用過的循環數 \hat{s} 之 RMSE，然而由於電池壽命(具有 1 至 1684 個週期範圍)分佈十分廣泛使 s ，因此本研究著重呈現 3 個相對重要的時區。也就是(1)早期:(所有電池均無明顯老化現象發生)(2)中期:(部分電池達到產品壽命)(3)晚期:(約過半電池達到產品壽命)。根據上述呈現方法，可以歸納出隨著老化特徵的出現對於剩餘壽命的預測誤差會隨之下降。此結論與利用加速充電速率進行快速老化試驗(有助於在少量循環下得知電池特性)的研究[51]相符。另一方面，對於已循環週期的預測結果中，可以看出隨著老化作用加劇，預測的誤差隨之加大，甚至出現大於 90 個循環週期的誤差。該現象可能來自於對於降解變化並無函數可以形容，也就是函數會發生一對多的情形。電池品質較差的電池可能僅在少量循環下會與高品質電池但歷經多次循環負載下具有相似特徵。然而隨著神經網路可觀察資料的增加(a 上升)，可藉由分析觀測輸入資料時序中前後老化程度變化分離上述提到的降解特徵重疊問題。此現象可以在圖 19(b)Old 區域中發現。除此之外，基於滑動式預測方法的資料處理手段會造成不同已使用循環狀態下樣本不平衡的問題，如圖 18 所展示。圖中分別以平均已循環次數作為分界標註高於和低於平均樣本數於圖標中，高於平均樣本數佔整體樣本的 15%。在樣本數目不平衡時，神經網路會因為訓練隨機性而忽略少樣本數的特徵。

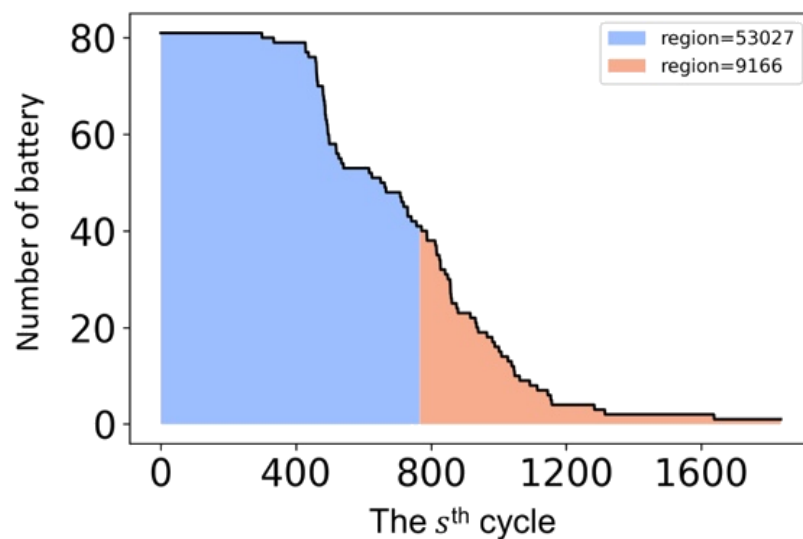


圖 18 黑色線段標示符合已使用循環次數條件下的電池數量，而線段下面積則代表訓練樣本個數。

值得注意的是，圖 19(a)Baby 區域中具有 A、B 兩個區域誤差相對較高。這是因為 A 為製備電池數據進行負載測試時恆溫控制器失效所導致的；B 則是因為初期並無明顯老化特徵，造成模型判斷困難所導致。若我們可以在待預測數據中，分離出子時序，而這些子時序可以被以不同狀態(s, a)所描述，該現象可被下列函式所演示：

$$s + a = c \quad (11)$$

其中 c 為一定值。該函式關係多次出現於神經網路預測數據集效能的結果圖中，也就是圖 19 三個區域中都存在斜率為 1 的高誤差尖銳線段。這是因為本研究框架所採用最後填充資料擴增方式(重複填充最後一個測量數據)，造成數據過分依賴測量時序的最後一個元素。若剛好被重複的循環資料數據測量發生問題，就有可能造成神經網路預測失真的情形；但若下一個循環資料便恢復正常數據紀錄時，數據擴增會對正常測量的資料進行複製，並輸入於神經網路中得到較精確的結果。

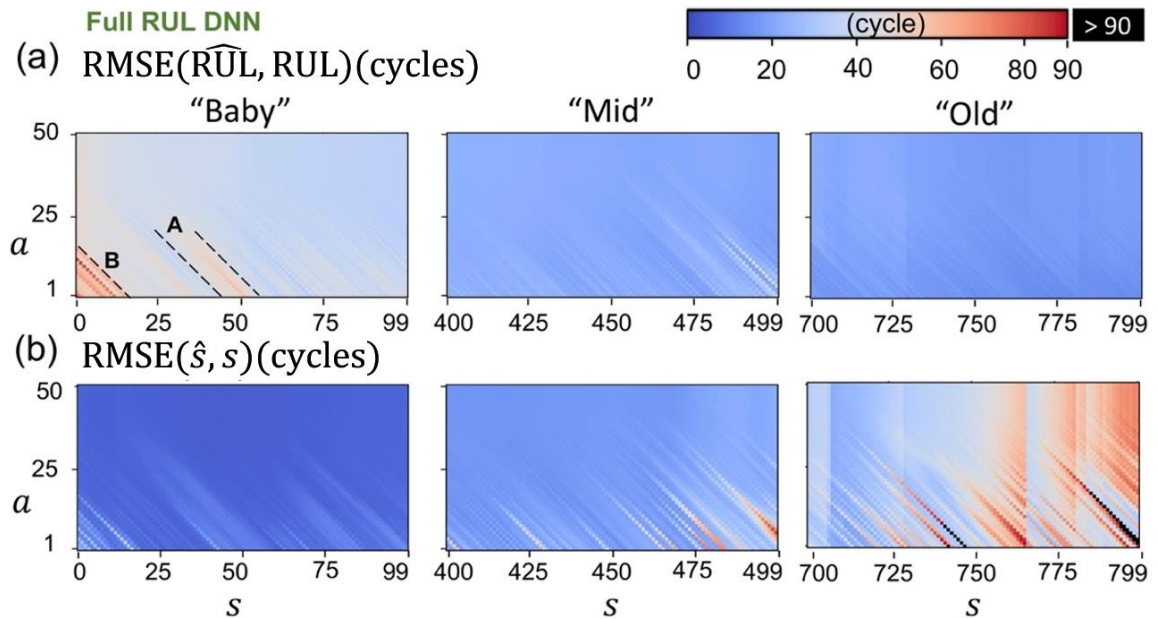


圖 19 完全剩餘壽命模型於測試數據集下的誤差著色圖，其中包和大部分電池所具有的三個範圍，分別為(1) 早期 (1~100循環)、(2) 中期(4~500循環) 和(3) 晚期(7~800循環)

4.2.2 預測電壓衰退狀態

完全剩餘壽命模型中的預測器 6 中是根據 Encoder-Decoder 架構進行設計，該神經網路在藉由一小片段的時序資訊，預測出整段電池生命週期中任意時刻的特性包含該次放電的功耗、電量和電壓電量曲線。描述任意時刻的方式是採用選取電池生命週期中的 100 個等距時間點，在圖表中會使用 $\left\lfloor \frac{n}{\text{EoL}} \% \right\rfloor$ 表示。根據上述預測剩餘壽命的方法，分別針對 3 個輸出值進行全盤的測試誤差結果呈現如圖 20(a)(b)(c) 所展示。且誤差計算方式分別註記在子圖上方。除了接近電池壽命終止時具有高度不確定性所造成的誤差過大外，其餘區間的預測精度在僅使用一個循環數據時誤差皆小於 4%。值得注意的是，若輸入的時序來源與需要預測的目標時刻相近，會有顯著的誤差下降情形。這來自於模型僅需在原始輸入上進行小幅度加工便可推論出精確答案。例如我在一個電池週期有 800 次循環的電池上，給予 400 次循環時量測到的充放電資訊，期望模型進行 $\left\lfloor \frac{n}{\text{EoL}} \% \right\rfloor = 50$ 時的特性預測，在此條件下因為輸入與輸出結果十分相近或是直接相關，導致出現明顯誤差偏低的現象，也就是圖 20(a)(b)(c) 中從左下到右上角的一條深色帶狀區域。另一方面，針對放電最大功耗隨著循環次數增加的變化情形也呈現於圖 20(d)。其中呈現的樣本均為測試數據集，隨後藉由輸入 4 個階段的單次循環資料進行預測，這些階段分別來自於 $s = 0, \lfloor (\text{EoL} - 100) \times 33\% \rfloor, \lfloor (\text{EoL} - 100) \times 66\% \rfloor, \text{EoL} - 100$ ，並在圖中分別對應“Baby,” “Teen,” “Mid,” and “Old” ages。由圖可看出本研究的預測效能相當優越，且利用不同時期的循環資料輸入，預測能力相近。

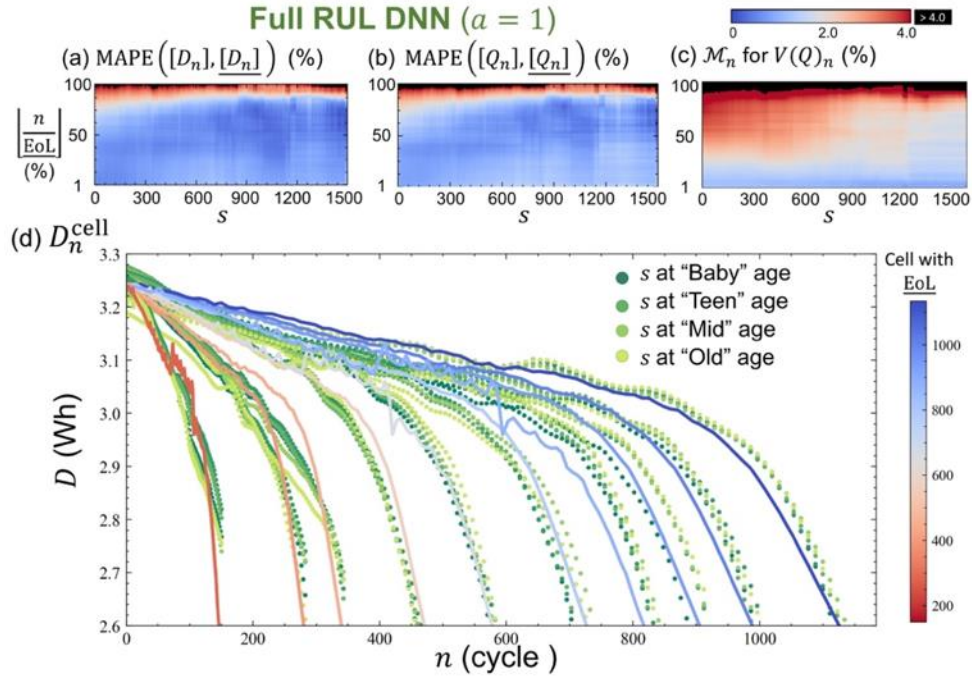


圖 20(a)(b)(c)為測試數據集的誤差著色圖，其中誤差根據右上角的色條著色，若超過4%則使用黑色標記，(d)為不同階段下進行整個生命周期的最大放電功耗預測曲線，實線為量測值，點則根據不同顏色代表不同階段下的預測值。

4.3 驗證於多段快充策略數據

本實驗使用的數據集視電池在當前2段可選恆定電流的方法快速充電(圖4(a))。對此數據，本研究的神經網路模型能顯現出極高的靈活性(對資料循環次數要求少)與精確度。然而，此快速充電策略已經有使用者習慣假設，因此仍無法貼近實際使用狀況，造成歸納偏置(Inductive bias)，實際使用場景可能接近維納過程(Wiener process)等類似的連續時間隨機過程。為了驗證本框架所提出的神經網路是否針對該快速充電設計方針產生過適應，我們採用另一基於同型號電池所進行的老化試驗數據集[26]作為新的測試數據集。在該研究中，總共包含三類不同的充電策略選取方式:(1)進行閉環式訓練所得(CLO top3)、(2)根據其他文獻所啟發設計(Literature-inspired)和(3)其他。透過此多樣的額外充電策略對本研究的模型進行泛化能力檢測，在僅使用第一次循環資料整體的平均絕對百分比誤差僅9.6%，預測結果如圖21所呈現。值得注意得是，因為CLO top3的充電策略與其餘兩類相比，控制電流切換的時間相對不一致，造成該數據

與訓練時的數據有微幅出入，導致預測結果較其餘兩類差。

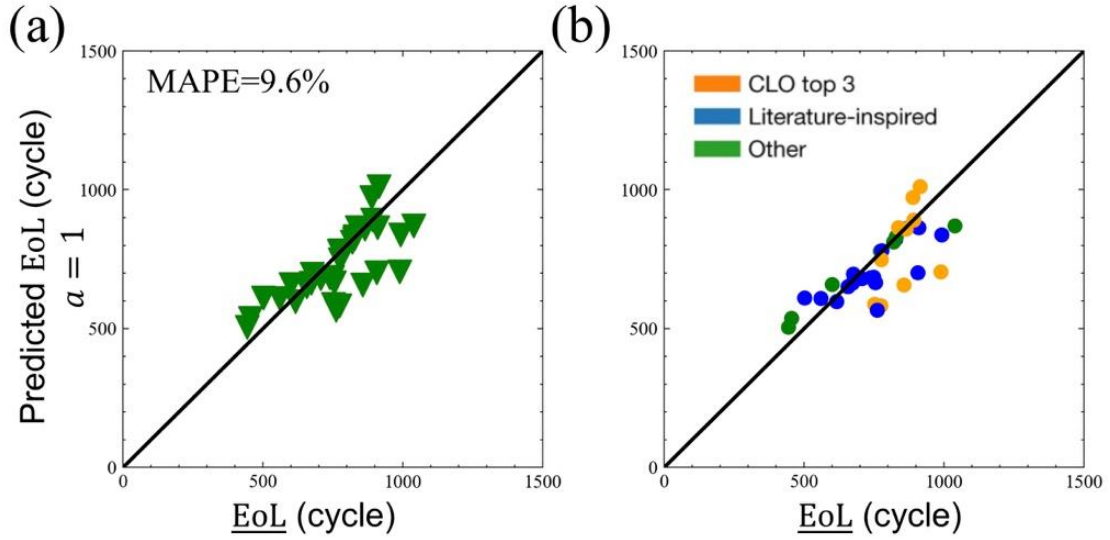


圖 21 完全模型對新式(6 段恆定電流)快充方法的預測效果圖，分別以(a)整體數據和(b)依據充電策略分別著色呈現。

4.4 神經網路即時效能測試

在實時(real-time)預測任務中，因計算效能受摩爾定律的限制需在進行預測精度與計算速度之間做出抉擇平衡兩者。其中在終端設備即時推論的研究[52]指出，經過結構最佳化的模型通常可訓練參數會介於 $10^6 \sim 10^8$ 之間。另一方面，在設備計算能力調查的研究[53]中表示對於常見 IOS & Android 終端設備而言，若將模型參數控制在 10^7 左右可以讓設備具有每秒預測 25~40 次樣本的能力。此外，參考於最先進物件檢測模型 EfficinetDet[54]的設計方針，建議將模型參數設定為 10^7 的數量級且具有小於 0.2 秒的 GPU 計算時間，即適合用以推廣到各領域所使用。

根據上述原因，本研究在設計模型架構時，會透過貝葉斯最佳化的搜尋方法找尋效能突出的若干個神經網路設計與超參數 (hyperparameter)候選組合，同時在保證可以實時推論的前題下，我們會篩選掉大幅超出計算範圍的模型使模型總參數量控制在 10^7 ，如表格 5、表格 6 和表格 7 所示。表中分別呈現我們提出類神經網路具有的參數量與每秒浮點運算次數 (Floating-point operations per second)。其中可以看到若任務目標越

複雜困難，例如需預測完整放電過程中電壓變化的 $V(Q)_{n=EoL}$ ，或是在已經過任意循環次數下預測該電池 RUL 的 Full RUL DNN，都需要較多的參數量才能達到參數與精度的甜蜜點，也就是參數量的增加並不能顯著增加模型預測的效能。

表格 5 Discharge DNNs 規格表

Component	Objective	Number of Parameters(M)	Number of Floating-point operation (M)
Feature selector	EoL/RUL (Discharge)	1.3	4.1
	Charge time at EoL (Discharge)	0.7	6.3
Predictor	EoL & Charge time at EoL (60 cycle)	1.3	12.6
	EoL & Charge time at EoL (80 cycle)	1.3	25.2
	EoL & Charge time at EoL (100 cycle)	1.3	37.8
	$V(Q)_{n=EoL}$	4.2	32.3

表格 6 Full DNNs 规格表

Component	Objective		Parameters(M)	FLOPs(M)
Feature selector	EoL (Discharge)		3.1	11.7
	Charge time at EoL (Discharge)		3.4	20.6
	EoL (Charge)		2.1	29.5
	Charge time at EoL (Charge)		0.8	32.9
Predictor	Integrated DML Predictor	Student 1	0.1	
		Student 2	0.2	
		Student 3	0.4	
		Total	0.8	6
	$V(Q)_{n=EoL}$		1.3	21.2
Test run to predict RUL	RUL ($0 \leq s \leq 100, 1 \leq a \leq 100$)		2.2	28.3

表格 7 Full RUL DNNs 規格表

Component	Objective	Number of Parameters(M)	Number of Floating-point operation (M)
Feature selector	Discharging part (1&2&5)	5.9	33.5
	Charging part (3&4&6)	5.9	23.8
Predictor	RUL & s ($0 \leq s, a \leq 50$)	2.7	37.9

陽明交大
NYCU

第五章 結論與未來展望

本研究提出兩個小實驗分別佐證對於時序的假設。假設一是透過時序正則的手段改善神經網路遇到具有不確定性樣本時的行為；假設二則是透過語言模型的方式處理電池上的時序資料，並輸入於已訓練過的模型得知，就算電池在未明顯降解的狀態下，仍會在表徵上顯現出老化的差異並同時發現測量數值會對神經網路的辨別該差異能力造成影響。隨後，本研究就利用該假設建立特徵篩選器與預測器模型，並預測多個電池控制系統高度關注的重要特性。其中在 EoL 的預測上超越現今最先進水平，並配合注意力機制使電池的片段足以推論電池整個產品週期的電池特性。

在本研究的框架下，仍有兩個分面可以改進與增強，分別是藉由此模型快速進行電池特性的最佳化和增強模型與電池管理系統的整合。第一個方面，後續研究可利用此模型的預測省去測試電池老化所需要時間，藉此快速調整製程參數以最大化電池電量；同時最小化電池充電所需時間等。抑或是將原本框架下的模型改為逆問題(Inverse Problems)，固定住神經網路的權重，透過更新輸入的參數(電流)求得最佳的充電策略同時在固定充電時間下保有高壽命。第二方面，則可以透過自我一致性(Self-consistent)的檢查，對模型的輸出進行管控進而增強系統穩健性。在本框架下，目前仍具有輸出參數之間非獨立的狀況，例如預測的 RUL 加 S1 並不等於 EoL($EoL \neq RUL + S$)，或是電壓-電量曲線進行積分並不等於最大放電功耗等。因此可透過比較多個網路之間的輸出用判斷輸出值的可信程度，同時可以進行環境條件的變化與是否存在惡意攻擊的情形。

參考文獻

1. Z.Zhu, D.Yu, Z.Shi, R.Gao, X.Xiao, I.Waluyo, M.Ge, Y.Dong, W.Xue, G.Xu, W. K.Lee, A.Hunt, J.Li, “Gradient-morph LiCoO₂ single crystals with stabilized energy density above 3400 W h L⁻¹,” *Energy Environ. Sci.*, **13**, pp. 1865–1878, 2020.
2. “Remaining the battery experience - Qnovo. (2020). Available at <https://qnovo.com/>.”
3. J.Zhu, M. S.Dewi Darma, M.Knapp, D. R.Sørensen, M.Heere, Q.Fang, X.Wang, H.Dai, L.Mereacre, A.Senyshyn, X.Wei, H.Ehrenberg, “Investigation of lithium-ion battery degradation mechanisms by combining differential voltage analysis and alternating current impedance,” *J. Power Sources*, **448**, 2020.
4. J. S.Edge, S.O’Kane, R.Prosser, N. D.Kirkaldy, A. N.Patel, A.Hales, A.Ghosh, W.Ai, J.Chen, J.Yang, S.Li, M. C.Pang, L.Bravo Diaz, A.Tomaszewska, M. W.Marzook, K. N.Radhakrishnan, H.Wang, Y.Patel, B.Wu, G.J.Offer, “Lithium ion battery degradation: what you need to know,” vol. , 23, 2021.
5. H. G.Schweiger, O.Obeidi, O.Komesker, A.Raschke, M.Schiemann, C.Zehner, M.Gehnen, M.Keller, P.Birke, “Comparison of several methods for determining the internal resistance of lithium ion cells,” *Sensors*, **10**, 2010.
6. R.Xiong, *Battery management algorithm for electric vehicles*, 2019.
7. A.Barai, K.Uddin, W. D.Widanage, A.McGordon, P.Jennings, “A study of the influence of measurement timescale on internal resistance characterisation methodologies for lithium-ion cells,” *Sci. Rep.*, **8**, 2018.
8. D.Roman, S.Saxena, V.Robu, M.Pecht, D.Flynn, “Machine learning pipeline for battery state-of-health estimation,” *Nat. Mach. Intell.*, **3**, 2021.
9. Y.Zhang, R.Xiong, H.He, M. G.Pecht, “Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries,” *IEEE Trans. Veh. Technol.*, **67**, 2018.

10. N.Srivastava, G.Hinton, A.Krizhevsky, I.Sutskever, R.Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, **15**, 2014.
11. G.Sannino, G.DePietro, "A deep learning approach for ECG-based heartbeat classification for arrhythmia detection," *Futur. Gener. Comput. Syst.*, **86**, 2018.
12. Z.Xue, Y.Zhang, C.Cheng, G.Ma, "Remaining useful life prediction of lithium-ion batteries with adaptive unscented kalman filter and optimized support vector regression," *Neurocomputing*, **376**, 2020.
13. Z.Chen, Q.Xue, R.Xiao, Y.Liu, J.Shen, "State of Health Estimation for Lithium-ion Batteries Based on Fusion of Autoregressive Moving Average Model and Elman Neural Network," *IEEE Access*, **7**, 2019.
14. S.Hochreiter, J.Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, **9**, pp. 1735–1780, 1997.
15. C.Lea, M. D.Flynn, R.Vidal, A.Reiter, G. D.Hager, "Temporal convolutional networks for action segmentation and detection," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vols. , 2017-January, 2017.
16. K.He, X.Zhang, S.Ren, J.Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vols. , 2016-December, 2016.
17. J.Hong, D.Lee, E. R.Jeong, Y.Yi, "Towards the swift prediction of the remaining useful life of lithium-ion batteries with end-to-end deep learning," *Appl. Energy*, **278**, pp. 115646, 2020.
18. K. A.Severson, P. M.Attia, N.Jin, N.Perkins, B.Jiang, Z.Yang, M. H.Chen, M.Aykol, P. K.Herring, D.Fraggedakis, M. Z.Bazant, S. J.Harris, W. C.Chueh, R. D.Braatz, "Data-driven prediction of battery cycle life before capacity degradation," *Nat. Energy*, **4**, pp. 383–391, 2019.

19. H.Zou, T.Hastie, “Regularization and variable selection via the elastic net,” *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **67**, 2005.
20. R.Caruana, S.Lawrence, L.Giles, “Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping,” *Advances in Neural Information Processing Systems*, 2001.
21. C. R.Birkl, M. R.Roberts, E.McTurk, P. G.Bruce, D. A.Howey, “Degradation diagnostics for lithium ion cells,” *J. Power Sources*, **341**, 2017.
22. A.Kendall, Y.Gal, “What uncertainties do we need in Bayesian deep learning for computer vision?,” *Advances in Neural Information Processing Systems*, vols. , 2017-December, 2017.
23. D.Castelvecchi, “Can we open the black box of AI?,” *Nature*, **538**, 2016.
24. I. J.Goodfellow, J.Shlens, C.Szegedy, “Explaining and harnessing adversarial examples,” 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, 2015.
25. O.Russakovsky, J.Deng, H.Su, J.Krause, S.Satheesh, S.Ma, Z.Huang, A.Karpathy, A.Khosla, M.Bernstein, A. C.Berg, L.Fei-Fei, “Large Scale Visual Recognition Challenge ({ILSVRC}),” *Int. J. Comput. Vis.*, **115**, 2015.
26. P. M.Attia, A.Grover, N.Jin, K. A.Severson, T. M.Markov, Y. H.Liao, M. H.Chen, B.Cheong, N.Perkins, Z.Yang, P. K.Herring, M.Aykol, S. J.Harris, R. D.Braatz, S.Ermon, W. C.Chueh, “Closed-loop optimization of fast-charging protocols for batteries with machine learning,” *Nature*, **578**, pp. 397–402, 2020.
27. B.Saha, K.Goebel, “Battery Data Set,” 2007.
28. T.Hastie, R.Tibshirani, M.Wainwright, *Statistical learning with sparsity: The lasso and generalizations*, 2015.
29. P. J.Huber, “Robust Statistics,” *International Encyclopedia of Statistical Science*, 2011.
30. T.Qin, S.Zeng, J.Guo, Z.Skaf, “A rest time-based prognostic framework for state of

- health estimation of lithium-ion batteries with regeneration phenomena,” *Energies*, **9**, 2016.
31. S.Ioffe, C.Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” 32nd International Conference on Machine Learning, ICML 2015, vol. , 1, 2015.
 32. T.Mikolov, K.Chen, G.Corrado, J.Dean, “Efficient estimation of word representations in vector space,” 1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings, 2013.
 33. A.Vaswani, N.Shazeer, N.Parmar, J.Uzkoreit, L.Jones, A. N.Gomez, Ł.Kaiser, I.Polosukhin, “Attention is all you need,” Advances in Neural Information Processing Systems, vols. , 2017-December, 2017.
 34. J.Devlin, M. W.Chang, K.Lee, K.Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference, vol. , 1, 2019.
 35. D.Arthur, S.Vassilvitskii, “K-means++: The advantages of careful seeding,” Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms, vols. , 07-09-January-2007, 2007.
 36. E.Hüllermeier, W.Waegeman, “Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods,” *Mach. Learn.*, **110**, 2021.
 37. T.Glasmachers, “Limits of end-to-end learning,” Journal of Machine Learning Research, vol. , 77, 2017.
 38. S.Ruder, “An Overview of Multi-Task Learning for Deep Learning,” 2017.
 39. T.Chen, S.Kornblith, M.Norouzi, G.Hinton, “A simple framework for contrastive learning of visual representations,” 37th International Conference on Machine Learning, ICML 2020, vols. , PartF168147-3, 2020.

40. J. B.Grill, F.Strub, F.Altché, C.Talleg, P. H.Richemond, E.Buchatskaya, C.Doersch, B. A.Pires, Z. D.Guo, M. G.Azar, B.Piot, K.Kavukcuoglu, R.Munos, M.Valko, “Byol,” arXiv, **200**, 2020.
41. P.Khosla, P.Teterwak, C.Wang, A.Sarna, Y.Tian, P.Isola, A.Maschinot, C.Liu, D.Krishnan, “Supervised Contrastive Learning,” 2020.
42. T.Wang, P.Isola, “Understanding contrastive representation learning through alignment and uniformity on the hypersphere,” 37th International Conference on Machine Learning, ICML 2020, vols. , PartF168147-13, 2020.
43. J.Snoek, H.Larochelle, R. P.Adams, “Practical Bayesian optimization of machine learning algorithms,” Advances in Neural Information Processing Systems, vol. , 4, 2012.
44. C.Szegedy, V.Vanhoucke, S.Ioffe, J.Shlens, Z.Wojna, “Rethinking the Inception Architecture for Computer Vision,” Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vols. , 2016-December, 2016.
45. M. T.Luong, H.Pharm, C. D.Manning, “Effective approaches to attention-based neural machine translation,” Conference Proceedings - EMNLP 2015: Conference on Empirical Methods in Natural Language Processing, 2015.
46. M.Lin, Q.Chen, S.Yan, “Network in network,” 2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings, 2014.
47. S. J.Reddi, S.Kale, S.Kumar, “On the convergence of Adam and beyond,” 6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings, 2018.
48. M.Zaheer, S. J.Reddi, D.Sachan, S.Kale, S.Kumar, “Adaptive methods for nonconvex optimization,” Advances in Neural Information Processing Systems, vols. , 2018-December, 2018.

49. P. J. Huber, "Robust Estimation of a Location Parameter," *Ann. Math. Stat.*, **35**, 1964.
50. A. Tang, R. Tam, A. Cadrin-Chênevert, W. Guest, J. Chong, J. Barfett, L. Chepelev, R. Cairns, J. R. Mitchell, M. D. Cicero, M. G. Poudrette, J. L. Jaremko, C. Reinhold, B. Gallix, B. Gray, R. Geis, T. O'Connell, P. Babyn, D. Koff, D. Ferguson, S. Derkatch, A. Bilbily, W. Shabana, "Canadian Association of Radiologists White Paper on Artificial Intelligence in Radiology," vol. , 69, 2018.
51. X. Tang, C. Zou, K. Yao, J. Lu, Y. Xia, F. Gao, "Aging trajectory prediction for lithium-ion batteries via model migration and Bayesian Monte Carlo method," *Appl. Energy*, **254**, pp. 113591, 2019.
52. X. Xu, Y. Ding, S. X. Hu, M. Niemier, J. Cong, Y. Hu, Y. Shi, "Scaling for edge inference of deep neural networks," *Nat. Electron.*, **1**, pp. 216–222, 2018.
53. A. Sehgal, N. Kehtarnavaz, "Guidelines and benchmarks for deployment of deep learning models on smartphones as real-time apps," 2019.
54. M. Tan, R. Pang, Q. V. Le, "EfficientDet: Scalable and efficient object detection," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2020.

附錄一 神經網路架構

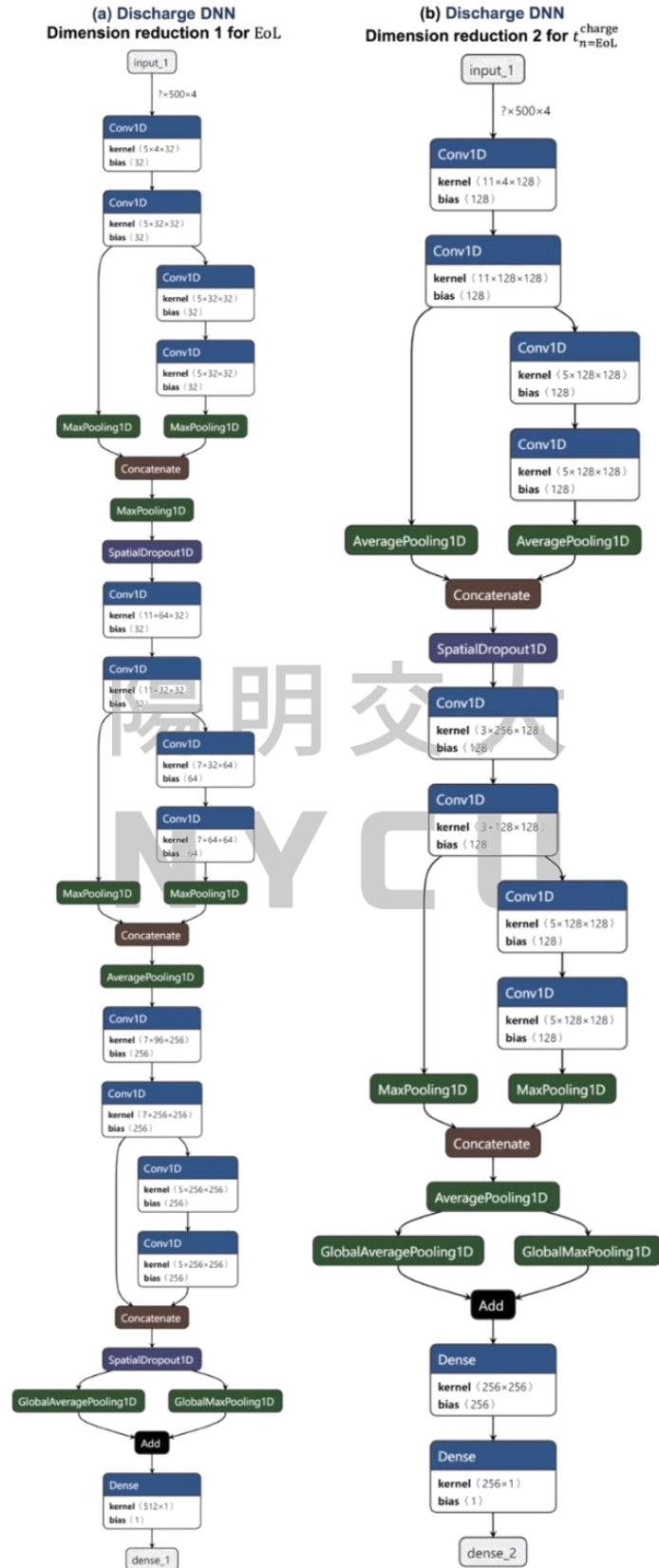


圖 22 放電模型中特徵篩選器 1(a)、2(b)架構圖

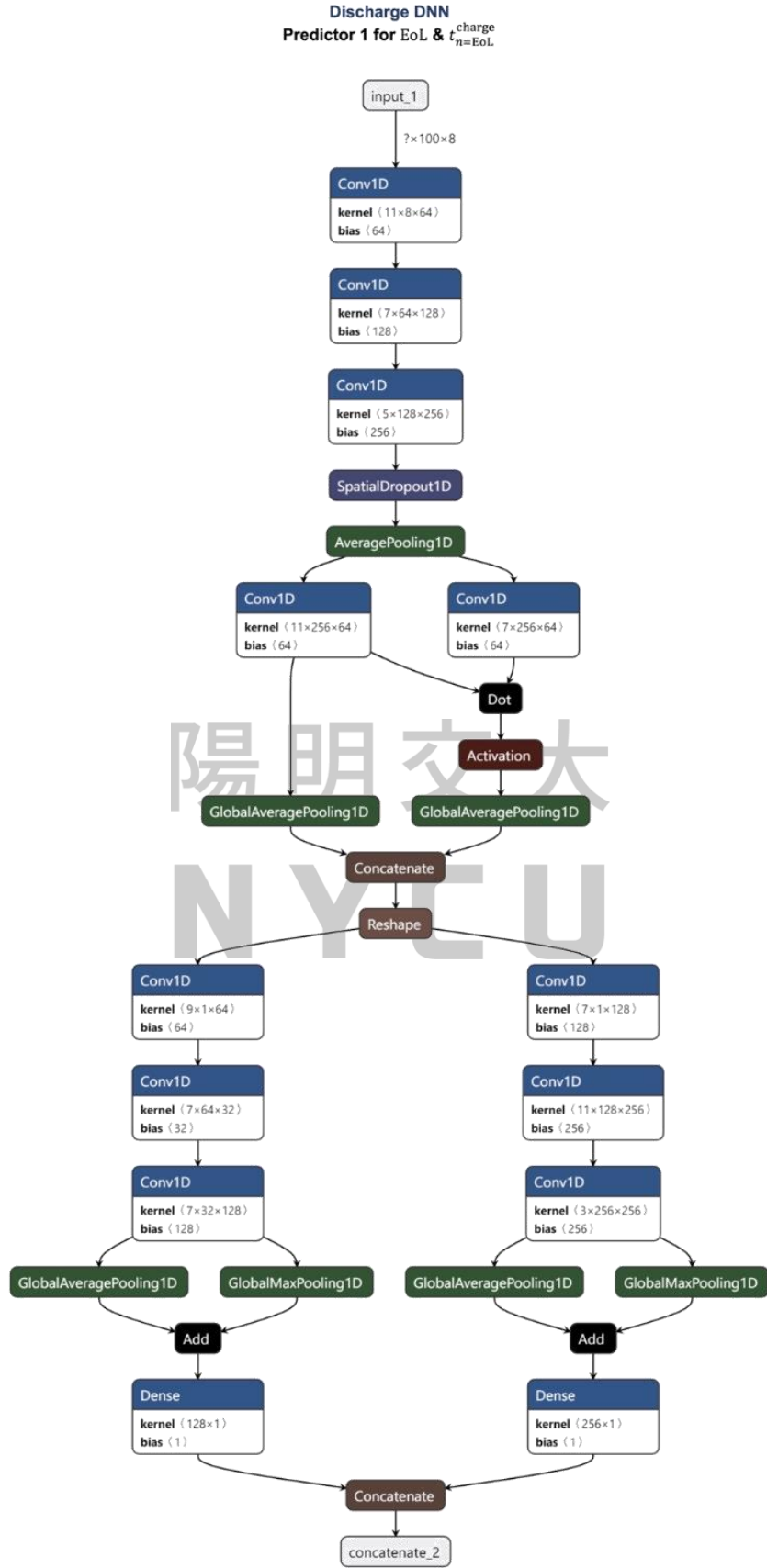


圖 23 放電模型中預測器 1 架構圖

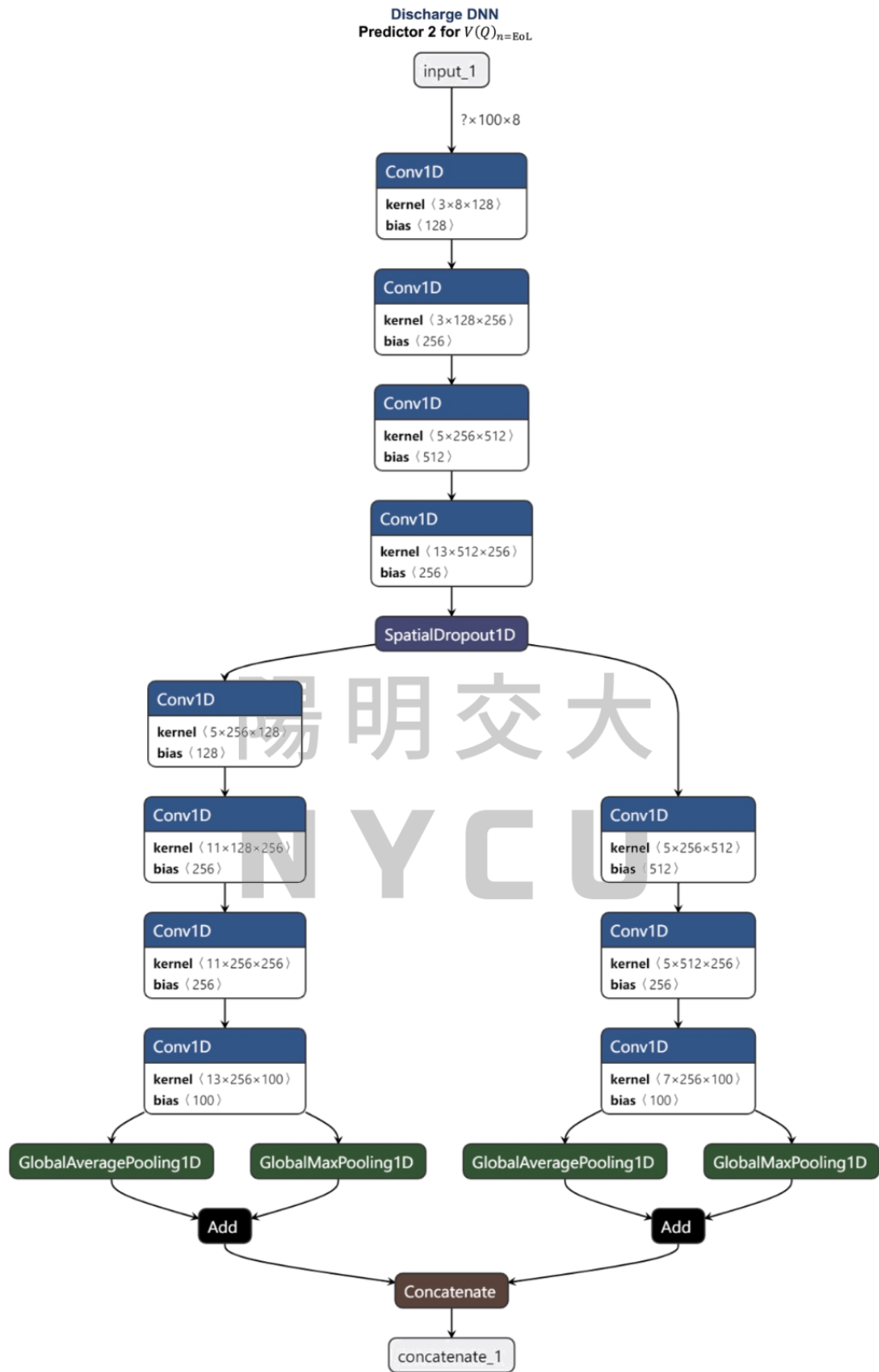


圖 24 放電模型中預測器 2 構圖

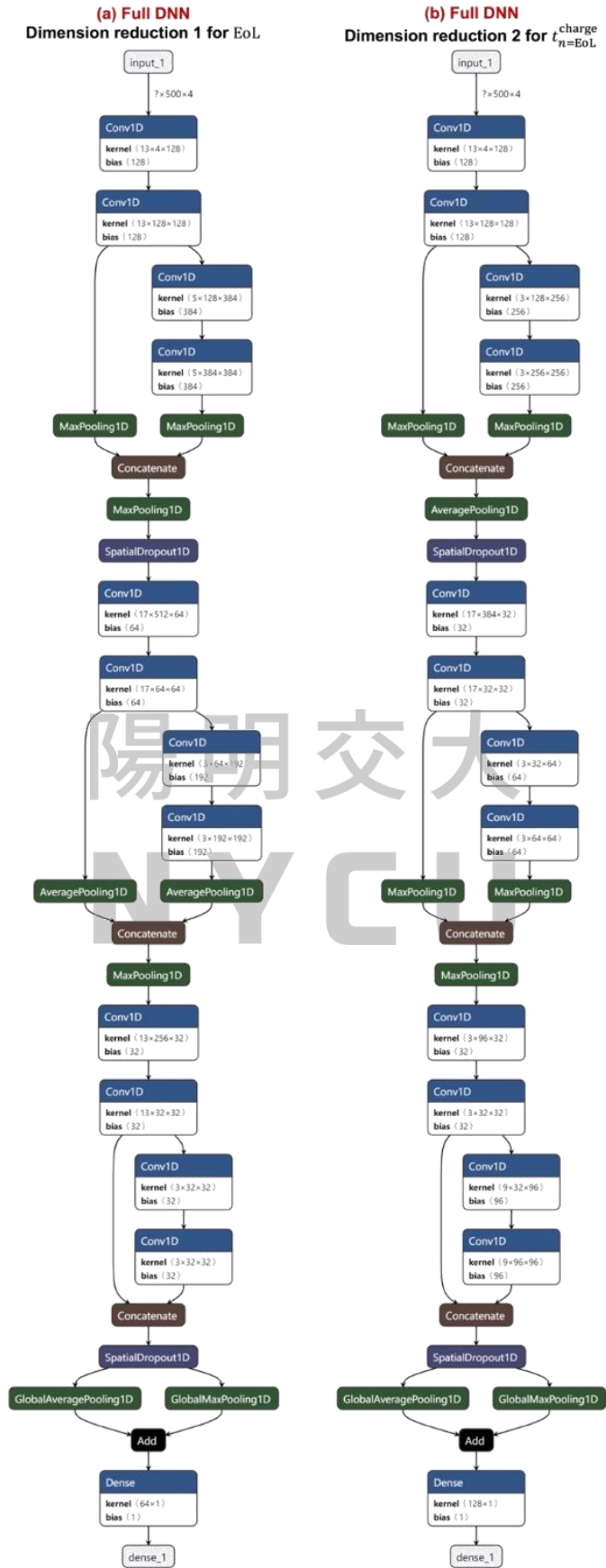


圖 25 完全模型中特徵篩選器 1(a)、2(b)架構圖

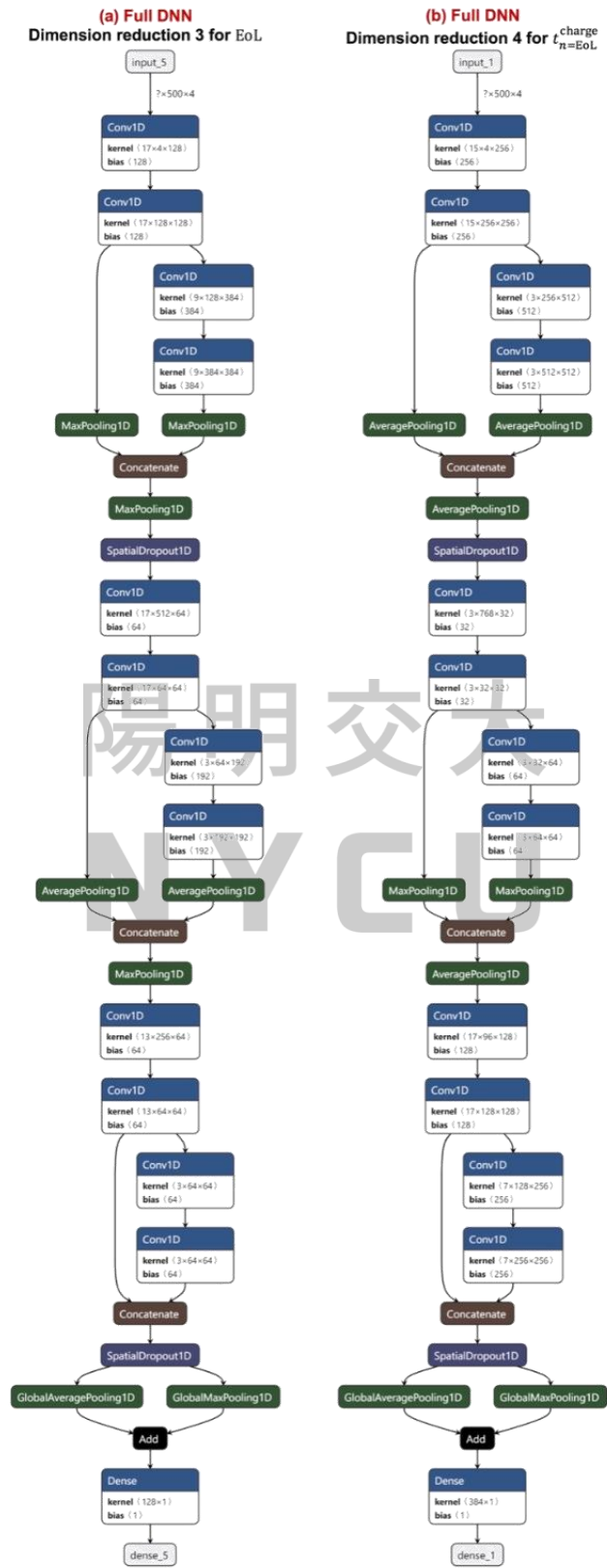


圖 26 完全模型中特徵篩選器 3(a)、4(b)架構圖

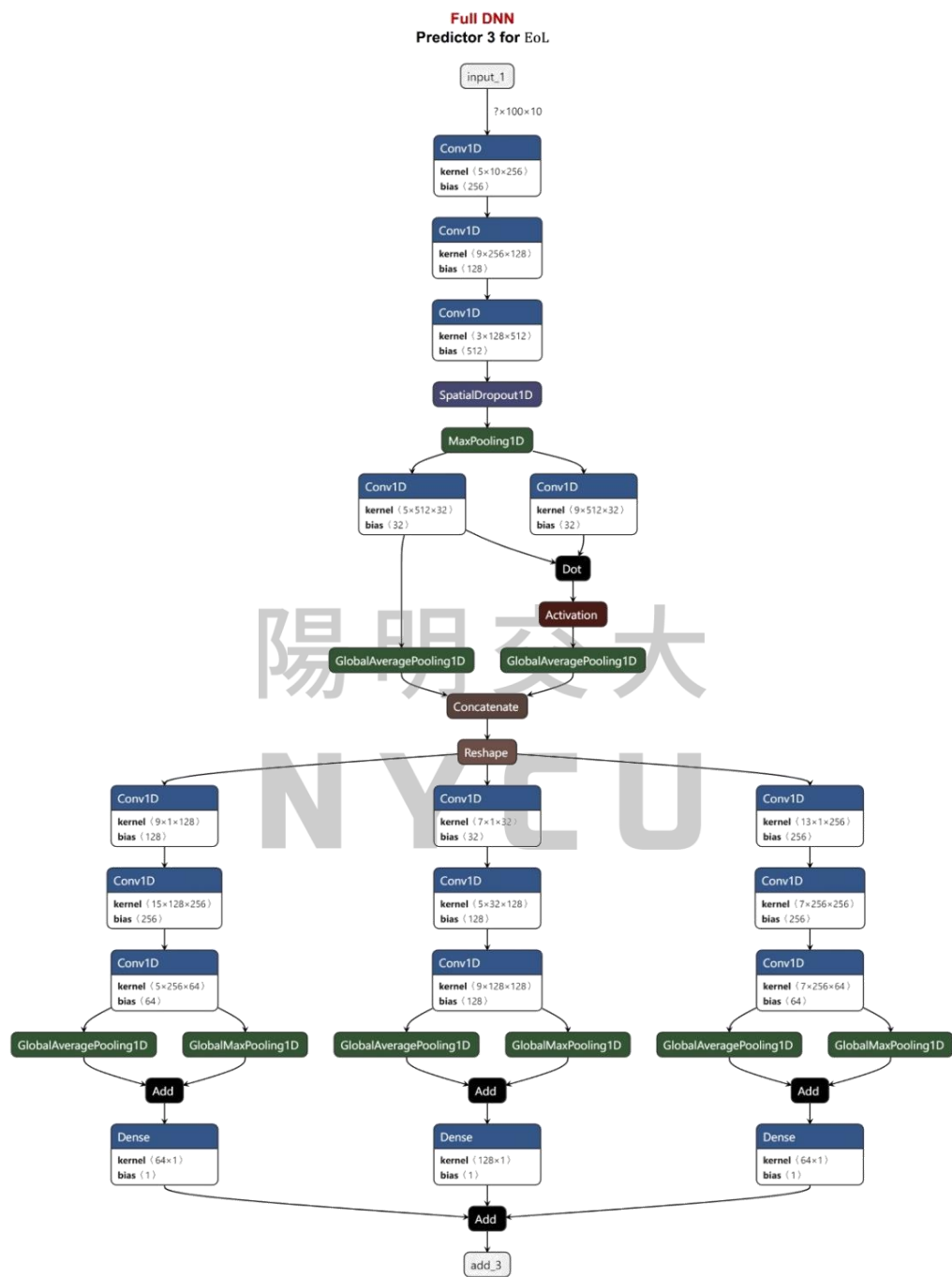


圖 27 完全模型中預測器 3 架構圖

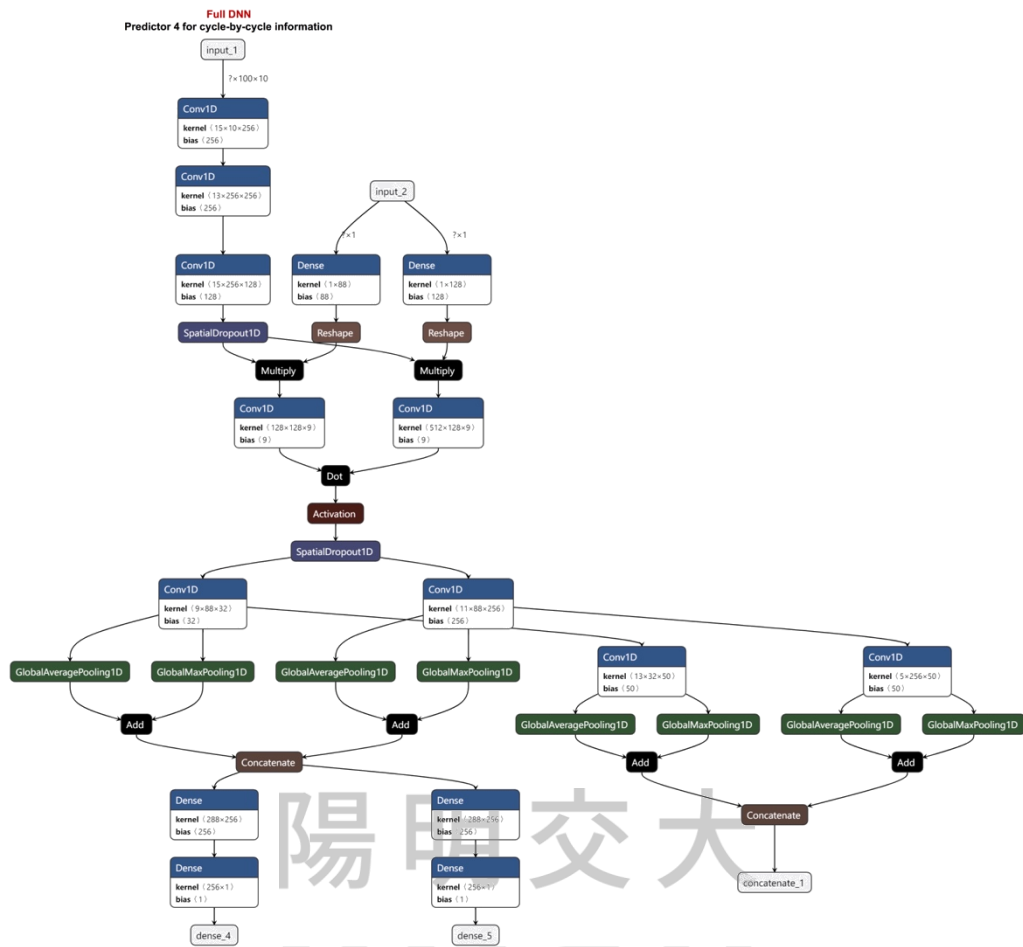


圖 28 完全模型中預測器 4 架構圖

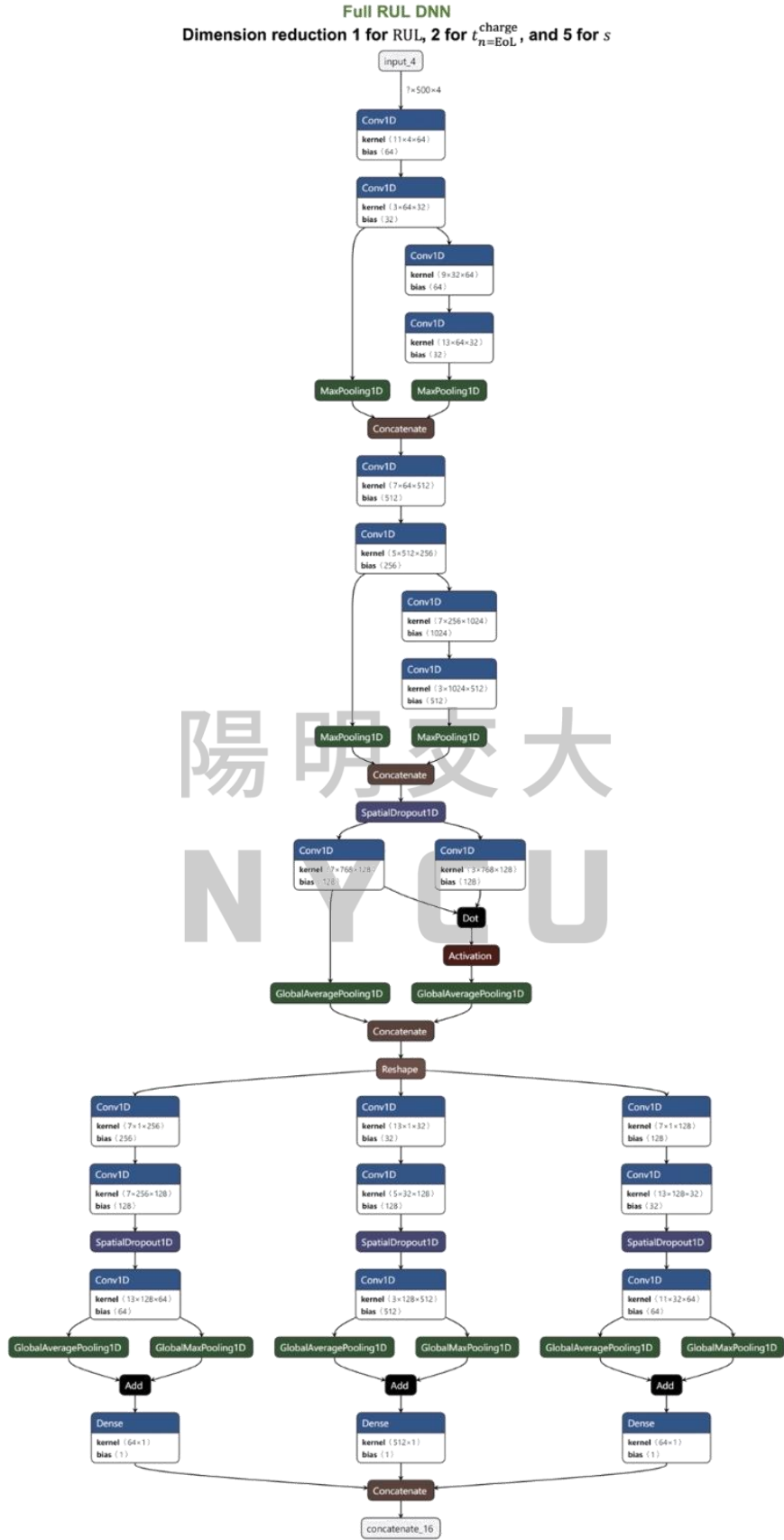


圖 29 完全剩餘壽命模型中特徵篩選器 1、2、5 架構圖

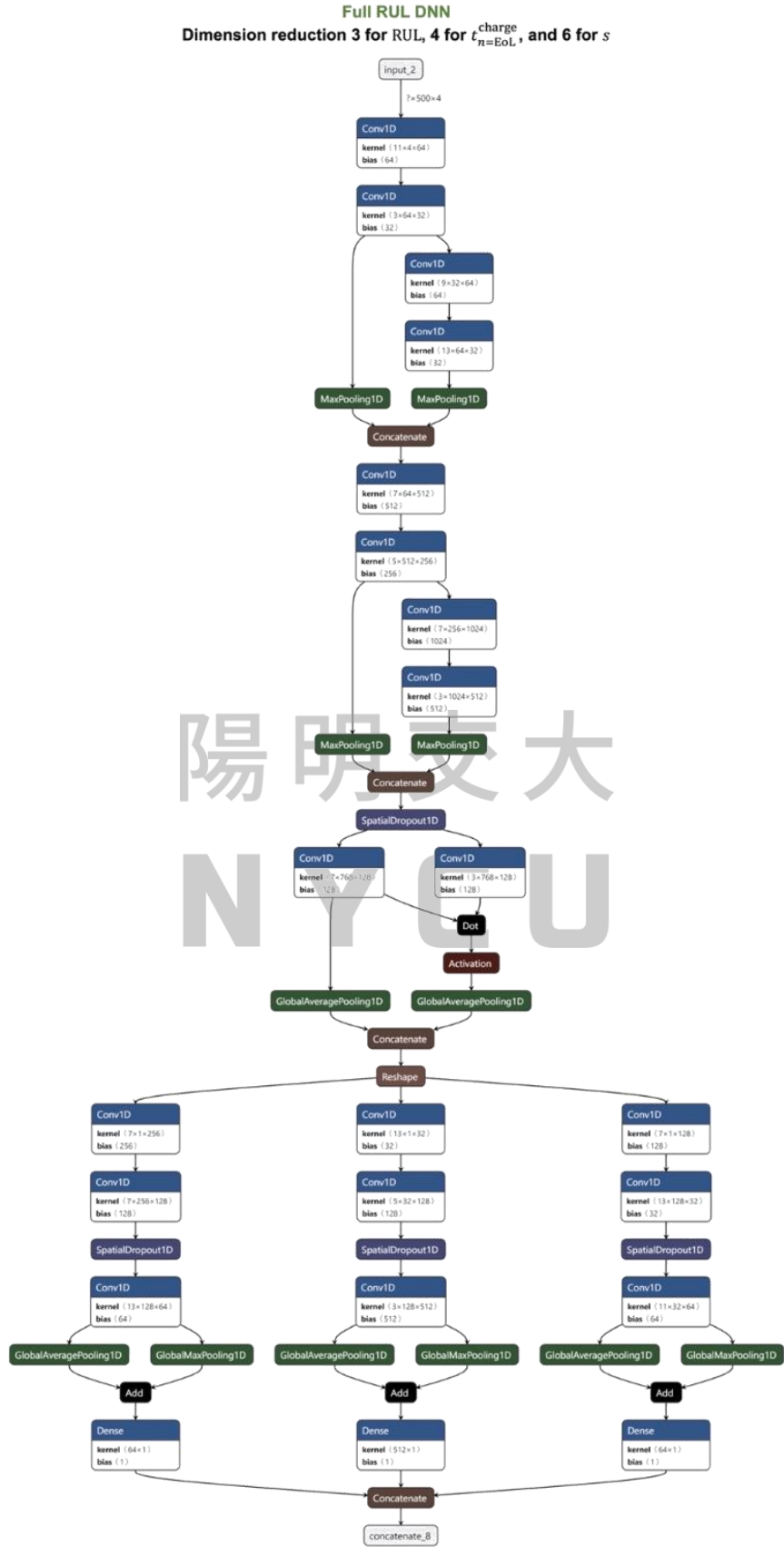


圖 30 完全剩餘壽命模型中特徵篩選器 3、4、6 架構圖

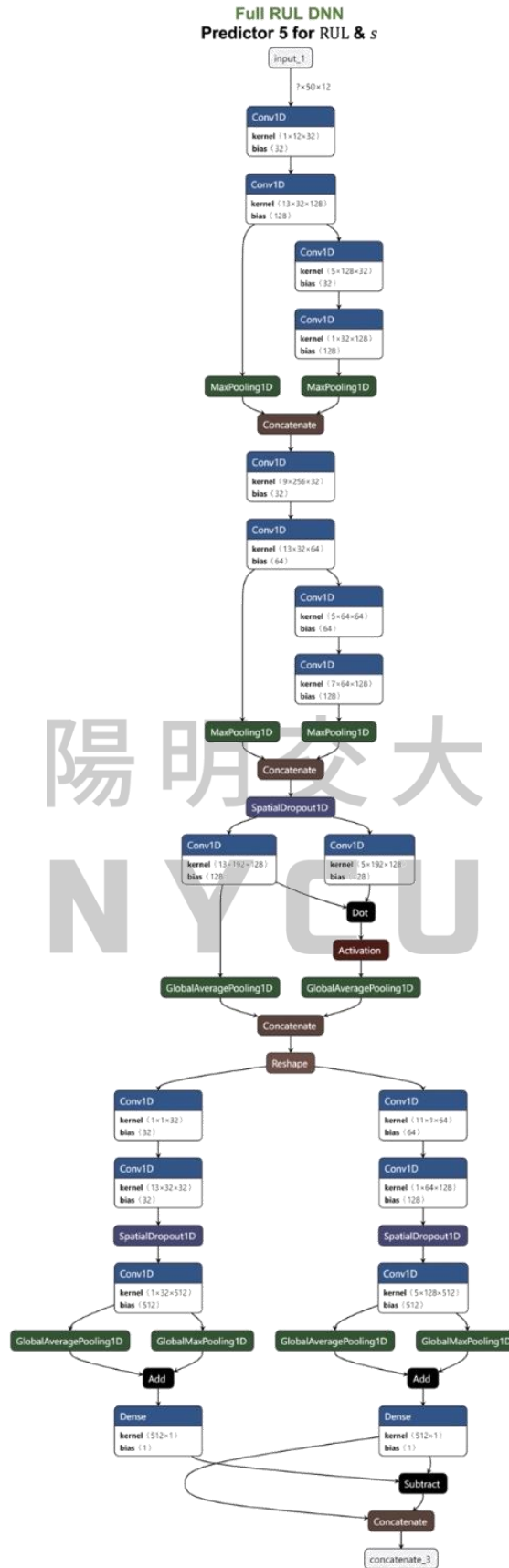


圖 31 完全剩餘壽命模型中預測器 5 架構圖

Full RUL DNN
Predictor 6 for cycle-by-cycle information

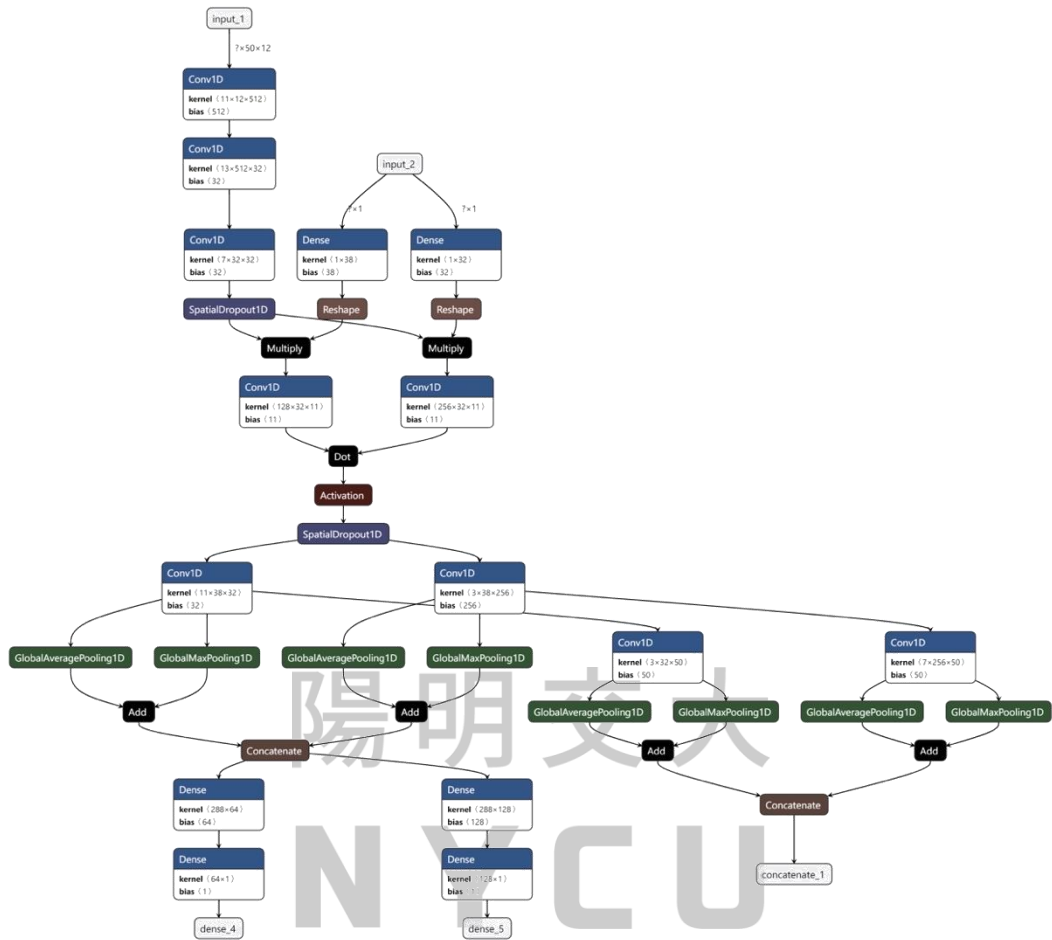


圖 32 完全剩餘壽命模型中預測器 6 架構圖