Cross-domain transfer learning for reliable condition monitoring of primary batteries under discharge-only operation

Measurement Science and Technology

OPEN ACCESS

RECEIVED
29 January 2026

REVISED
24 March 2026

ACCEPTED FOR PUBLICATION
9 April 2026

PUBLISHED
17 April 2026

PAPER

Cross-domain transfer learning for reliable condition monitoring of primary batteries under discharge-only operation

Sultan Zeybek[1,*] and **Imen Turki**[2]

[1] Department of Artificial Intelligence and Data Engineering, Istanbul University, 34134 Istanbul, Türkiye
[2] Department of Computer Engineering, Fatih Sultan Mehmet Vakif University, 34015 Istanbul, Türkiye
[*] Author to whom any correspondence should be addressed.

E-mail: sultan.zeybek@istanbul.edu.tr

Keywords: transfer learning, battery condition estimation, domain adaptation, deep learning

Abstract

Primary lithium-based coin-cell batteries are widely used in embedded systems, low-power sensors, and Internet of Things devices due to their long operational life and maintenance-free characteristics. In these applications, accurate estimation of the remaining battery life is essential to ensure system reliability. However, conventional methods for estimating battery health rely on repeated charge and discharge cycles, which are not applicable to primary batteries. This study presents a deep learning-based transfer learning framework to infer the condition of primary coin-cell batteries using discharge-only data. A new dataset is introduced, consisting of voltage versus time profiles obtained under constant current discharge from batteries manufactured by four different brands. The prediction target, referred to as the discharge progression indicator (DPI), is formally defined as the normalised elapsed discharge time and shown to be mathematically equivalent to Depth of Discharge under constant-current operation, making it fully observable from a single discharge event without requiring cycling data or explicit capacity measurements. Neural network models are initially pretrained on large-scale lithium-ion datasets and then adapted to the new dataset through a partial transfer strategy with input layer re-initialisation, enabling cross-chemistry knowledge transfer across fundamentally different battery chemistries. A range of model architectures is evaluated, including fully connected networks, convolutional networks, recurrent memory-based models, and attention mechanisms. The results demonstrate that temporal models, particularly those with memory structures, achieve superior predictive performance and robustness against domain-induced variability. A sensitivity analysis further confirms that standard 8-bit or 10-bit analogue-to-digital converters are sufficient for reliable DPI prediction, supporting deployment in resource-constrained embedded systems. The proposed framework enables early and accurate condition estimation in the absence of charging data or domain-specific calibration.

1. Introduction

Energy storage systems play an increasingly important role in contemporary technological applications, ranging from electric vehicles to distributed energy and microgrid systems. Within this broad landscape, primary lithium coin-cell batteries, such as CR2032 cells, constitute a critical yet relatively underexplored component. These batteries are widely employed in Internet of Things (IoT) applications, including smart metering, medical sensing devices, and safety-critical monitoring systems, where long-term, maintenance-free operation is essential. In contrast to rechargeable lithium-ion batteries (LIBs), which have been extensively investigated in the literature, primary batteries operate under a single, irreversible discharge process without recharge capability. As a result, unexpected battery depletion in such systems can lead to service interruptions, data loss, and increased maintenance costs due to unplanned field replacements [1].

Accurate estimation of battery condition and discharge completion for primary batteries presents several fundamental challenges. First, primary coin cells typically exhibit a long quasi-flat discharge voltage profile followed by a rapid voltage drop near end-of-life (EOL), making early prediction of discharge termination difficult using conventional voltage-threshold-based methods [2]. Second, most data-driven battery health estimation approaches rely on features extracted from repeated charge–discharge cycles, such as capacity fade trends, cycle count, or charging behaviour, which are inherently unavailable in primary battery applications. As a result, widely adopted state-of-health (SOH) estimation frameworks developed for rechargeable batteries [3] cannot be directly applied to primary coin-cell systems operating under real-world conditions [4].

Furthermore, the deployment of prediction models in real-world scenarios is hindered by the phenomenon of domain shift [5]. Batteries from different manufacturers (e.g. Maxell, GP, Duracell) exhibit distinct electrochemical behaviours and voltage-time profiles due to variations in manufacturing processes and material properties [1]. This variability creates a significant distribution discrepancy between the training data (source domain) and the operational data (target domain).

Most existing studies focus on transfer learning between rechargeable batteries of the same chemistry, leaving a gap in methodologies that adapt knowledge from abundant Li-ion datasets to data-scarce primary battery domains. Existing cross-chemistry transfer approaches [6] and small-sample adaptation methods [7] do not address the discharge-only observability constraint, while robustness under large distribution divergence remains a recognised challenge [8]. Conventional battery health assessment methodologies predominantly rely on long-term charge–discharge cycling data, laboratory-controlled testing conditions, and multi-sensor measurements. While effective for rechargeable LIBs, these assumptions are incompatible with primary battery systems, where recharging is not possible and only a single discharge trajectory is available. As a result, widely adopted SOH estimation frameworks, including capacity-based [9] and impedance-based approaches [10, 11], cannot be directly applied to primary coin-cell batteries operating in real-world conditions.

To address these limitations, this study proposes a deep transfer learning framework for discharge-only battery condition inference. Unlike rechargeable LIBs, primary lithium coin-cell batteries operate under a strictly discharge-only regime, which prevents the direct application of conventional SOH estimation methods based on cycling data. In this context, battery condition is inferred solely from voltage-time discharge behaviour under constant current (CC), making the problem fundamentally different from classical SOH prediction.

In this study, a new dataset, termed CR2032-Discharge (CR2032-D), is introduced, comprising constant-current discharge profiles collected from four different CR2032 battery brands. Considering the different type-different charging strategy (DT-DCS) nature of this problem [5], models are first trained on large-scale lithium-ion benchmark datasets (NASA and CALCE) and subsequently adapted to the primary battery domain using transfer learning. This strategy mitigates the effects of limited labelled data and brand-induced domain shift, enabling robust condition inference in realistic IoT deployment scenarios.

The remainder of this paper is organised as follows. Section 2 reviews related work on battery condition monitoring and transfer learning. Section 3 describes the proposed methodology, datasets, and model architectures. Section 4 presents the experimental results, comparative analyses and discusses the limitations of the proposed approach. The final section concludes the study and outlines future research directions.

2. Related work

Battery condition and lifetime estimation have been extensively studied in the context of rechargeable LIBs operating under repeated charge–discharge cycles. In this dominant paradigm, battery SOH is typically defined as the ratio between the current maximum capacity and the nominal capacity, assuming access to long-term cycling data and capacity measurements. Public benchmark datasets such as NASA and CALCE have enabled the development of data-driven SOH estimation models by providing well-instrumented cycling trajectories, capacity labels, and auxiliary sensor signals.

Leveraging these datasets, deep learning architectures, including long short-term memory (LSTM), gated recurrent unit (GRU), convolutional neural networks (CNN), and attention-based models, have demonstrated strong performance in capturing nonlinear degradation dynamics and long-term temporal dependencies. However, the effectiveness of these approaches is fundamentally tied to the availability of rich cycling data and capacity-related health indicators, which limits their applicability to practical deployments involving primary batteries.

To improve generalisation across different batteries, operating conditions, and data distributions, transfer learning has emerged as an effective strategy in cycling-based battery health and lifetime prediction. Recent studies show that models pretrained on large-scale cycling datasets can transfer learned temporal representations to new target domains with limited labelled data, significantly improving prediction accuracy under distribution shift [12, 13]. These methods explicitly address variations in battery chemistry, temperature, and usage profiles, thereby reducing the need for extensive target-domain data.

Several transfer learning frameworks have been proposed for SOH estimation and remaining useful life prediction. Semiparametric adaptive transfer learning approaches enable robust SOH prediction under varying operating conditions [14]. Generalisable transfer learning models have achieved mean absolute error (MAE) below 2% across diverse cell types and temperature regimes [15]. Domain-adversarial learning has been used to align feature distributions between source and target domains, mitigating the effects of domain shift [16]. Other studies combine cycle-based degradation modelling with transfer learning to achieve significant error reductions [17], while LSTM-based architectures [18], individualised prediction frameworks [19], and ensemble transfer learning approaches [20] further demonstrate the effectiveness of these methods.

More recently, multi-source domain adaptation strategies have been introduced to improve robustness under heterogeneous operating conditions [21]. Voltage relaxation-based CNN transfer learning has also shown promising performance in small-sample regimes, achieving accurate capacity estimation using only a small fraction of labelled target-domain data [22]. Collectively, these studies establish transfer learning as a mature and effective solution for cycling-based battery condition estimation.

Despite these advances, existing deep learning and transfer learning approaches remain fundamentally dependent on cycling-based observability and capacity-related health indicators. These assumptions do not hold for primary battery applications, where recharge events are not possible and cycle-based SOH definitions are not observable. In such scenarios, battery condition must be inferred from a single discharge event, often using voltage-time measurements as the only available signal.

Only a limited number of studies have explored discharge-based battery condition and remaining discharge time estimation, particularly for low-power and remote systems. Physics-based and thermodynamic models estimate remaining discharge time under known load conditions, while some data-driven approaches demonstrate that discharge voltage trajectories contain degradation-sensitive information. However, these studies typically rely on handcrafted features, shallow learning models, or application-specific assumptions and rarely address generalisation across different battery brands or chemistries.

In particular, the impact of domain shift in discharge-only settings, arising from manufacturing variability, brand-dependent characteristics, or electrochemical differences, remains largely unexplored. Only a small subset of existing work considers transfer learning under strictly discharge-only observability, despite its practical relevance to primary battery deployments.

Overall, the literature reveals a clear methodological gap. While deep transfer learning has proven highly effective for cycling-based battery health estimation, its applicability to discharge-only primary batteries has not been systematically investigated. Moreover, large distribution discrepancies between source and target domains increase the risk of negative transfer, where knowledge learned from the source domain degrades rather than improves performance in the target domain. Prior studies report that excessive divergence between source and target distributions can lead to transfer failure [8]. This observation motivates the adoption of transfer learning strategies that prioritise the extraction of domain-invariant temporal representations rather than naive fine-tuning.

3. Materials and methods

This section details the proposed framework for discharge-only battery condition inference. The methodology is designed to address the challenge of transferring knowledge from rechargeable Li-ion batteries (source domain) to primary coin-cell batteries (target domain) under significant domain shifts. The framework consists of three main stages: dataset preparation, domain adaptation strategy, and deep learning model architecture.

3.1. Battery datasets and source-target domain overview

This study utilises three battery datasets representing fundamentally different electrochemical behaviours and operational regimes: NASA, CALCE, and the proposed CR2032-Discharge (CR2032-D) dataset. These datasets jointly enable the investigation of transfer learning across battery types, discharge characteristics, and domain shifts.

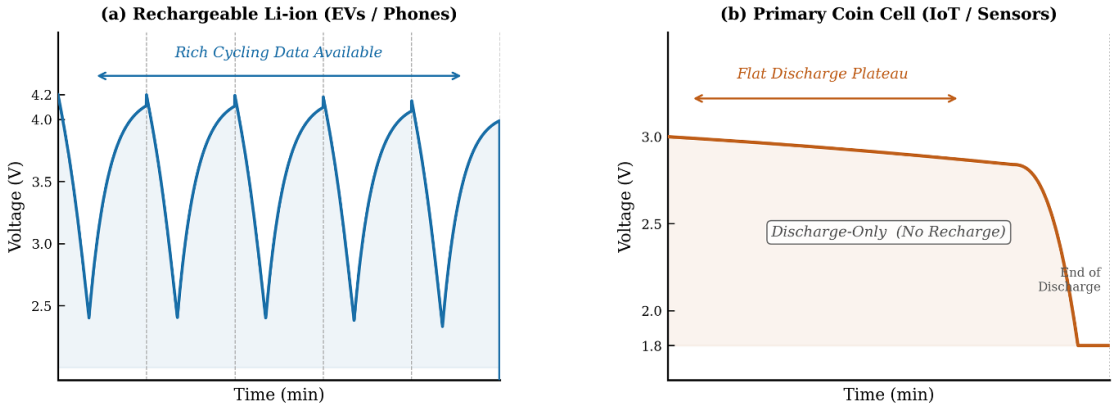


Figure 1. Conceptual comparison between rechargeable lithium-ion batteries and primary lithium coin-cell batteries. Horizontal axis: time (min); vertical axis: voltage (V). The rechargeable cell exhibits repeated charge–discharge cycles with explicit capacity labels, whereas the primary cell operates under a single irreversible discharge trajectory with no recharge events.

NASA battery dataset: The NASA Ames Prognostics Center of Excellence dataset comprises four lithium-ion battery cells (B0005, B0006, B0007, and B0018) subjected to repeated charge, discharge, and impedance measurement cycles [23]. Charging was performed under a CC mode at 1.5 A until the terminal voltage reached 4.2 V, followed by constant voltage (CV) charging until the current decayed to 20 mA. During discharge, a CC of 2 A was applied until predefined end-of-discharge (EOD) voltages were reached (2.7 V, 2.5 V, 2.2 V, and 2.5 V, respectively). Battery life tests were terminated when the EOL criterion was satisfied, defined as a 30% reduction in rated capacity from 2.0 Ah to 1.4 Ah. This dataset provides rich cycling information and explicit capacity degradation labels.

CALCE battery dataset: The CALCE (Center for Advanced Life Cycle Engineering) dataset includes lithium-ion cells (CS2_35, CS2_36, CS2_37, and CS2_38) tested under controlled laboratory conditions [24]. Cells were charged using a standard CC-CV protocol with a 1 C charging rate up to 4.2 V, followed by a CV phase until the current dropped below 0.05 A. Discharging was conducted at a CC until a cut-off voltage of 2.7 V was reached. Similar to NASA, the CALCE dataset offers long-term cycling trajectories suitable for learning general electrochemical degradation patterns.

CR2032-Discharge (CR2032-D) Dataset: To address the lack of publicly available discharge-only datasets for primary batteries, a new dataset was developed in this study. CR2032-D contains constant-current discharge voltage-time profiles of primary CR2032 lithium coin cells manufactured by four different brands: Maxell, GP, Fujitsu, and Duracell. All cells were discharged at 0.1 A until a cut-off voltage of 1.799 V. Unlike rechargeable lithium-ion datasets, CR2032-D contains no charging cycles and exhibits a characteristically flat discharge profile followed by a sharp terminal voltage drop, posing significant challenges for early-stage condition inference.

As illustrated in figure 1, primary lithium coin-cell batteries fundamentally differ from rechargeable LIBs in terms of observable degradation behaviour. The absence of charge–discharge cycles and the presence of extended quasi-flat discharge plateaus challenge conventional SOH estimation approaches and motivate the need for discharge-only condition inference.

3.2. Target domain characterisation and data preprocessing

The CR2032-Discharge (CR2032-D) dataset constitutes the target domain of this study and represents a fundamentally different operational setting compared to conventional rechargeable battery datasets. CR2032-D is composed of primary (non-rechargeable) lithium coin-cell batteries operating under a strictly discharge-only regime, which introduces a pronounced domain shift relative to cycling-based lithium-ion datasets such as NASA and CALCE.

The dataset includes discharge voltage-time profiles from four different CR2032 battery brands: Maxell, GP, Fujitsu, and Duracell. Although these batteries share the same nominal chemistry and rated capacity, variations in manufacturing processes, electrode materials, and internal resistance lead to distinct electrochemical behaviours across brands [1]. As a result, the target domain exhibits a *brand-induced domain shift*, where batteries from different manufacturers generate systematically different voltage trajectories under identical discharge conditions.

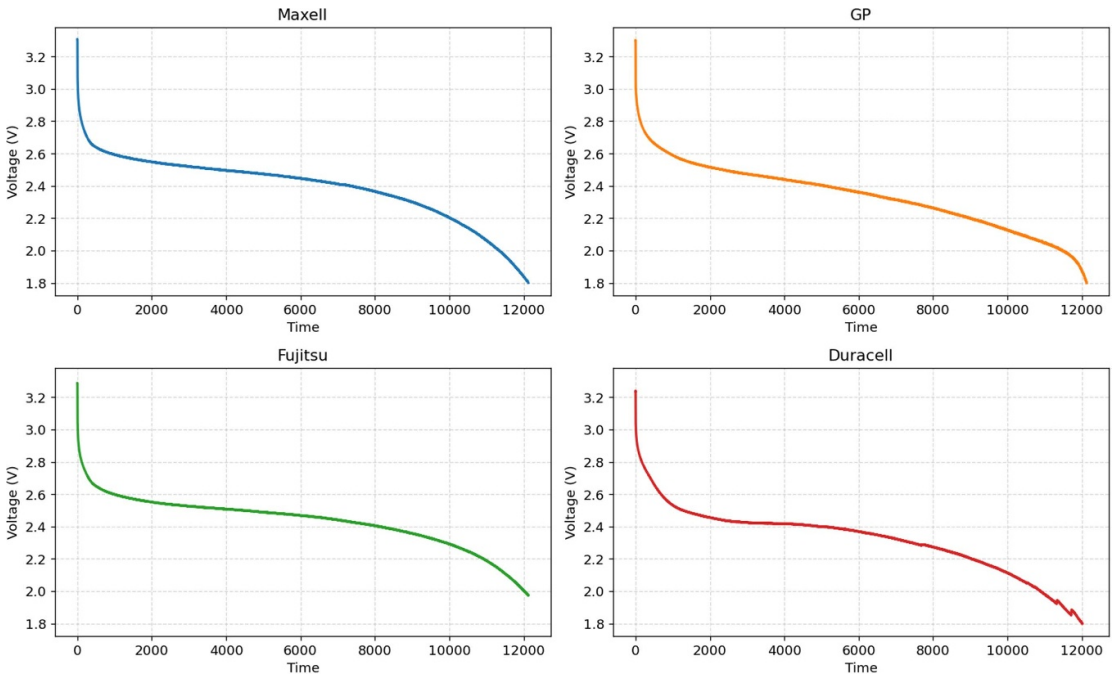


Figure 2. Raw discharge voltage-time profiles of CR2032 coin-cell batteries for different brands, illustrating brand-dependent discharge behaviour under identical load conditions. Horizontal axis: time (min); vertical axis: voltage (V).

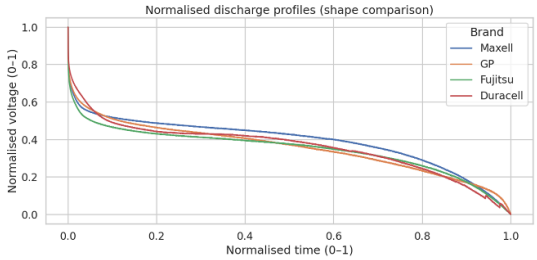


Figure 3. Normalised discharge voltage profiles of CR2032 batteries, highlighting shape-level similarities and deviations across different brands. Horizontal axis: normalised time; vertical axis: normalised voltage.

All CR2032 cells were discharged under a CC protocol at 0.1 A until a cut-off voltage of 1.799 V was reached. Unlike rechargeable LIBs, the CR2032-D domain contains no charging cycles and no explicit capacity measurements. Instead, the discharge voltage remains relatively flat over a large portion of the operating time before undergoing a rapid terminal drop, as illustrated in figure 2. This characteristic flat discharge behaviour significantly complicates early-stage battery condition inference, since degradation-related signatures are weak and difficult to distinguish during the initial discharge phase [2].

To further characterise the structural similarity and divergence across brands, discharge curves were normalised with respect to time and voltage. Figure 3 shows that, while the overall discharge shape remains broadly consistent across brands, noticeable deviations persist in both plateau duration and terminal voltage drop regions. These shape-level discrepancies indicate that brand-induced variability is not limited to absolute voltage levels but also affects the temporal evolution of discharge dynamics.

In addition to temporal differences, statistical distribution mismatch across brands further amplifies the domain shift challenge. Figure 4 presents the probability density functions of voltage values for each brand, revealing shifted peaks and varying spreads. These observations confirm that voltage measurements from different manufacturers are not drawn from the same underlying distribution, violating the common assumption of identical training and deployment data distributions in machine learning models.

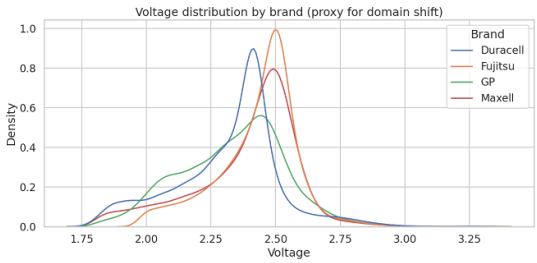


Figure 4. Voltage probability density distributions for different CR2032 brands, illustrating brand-induced distribution shift in the target domain. Horizontal axis: voltage (V); vertical axis: probability density.

Prior to model training, all datasets underwent a unified preprocessing pipeline to ensure compatibility across heterogeneous domains. Voltage and auxiliary measurements were standardised using Z-score normalisation,

$$x' = \frac{x - \mu}{\sigma}, \tag{1}$$

where μ and σ denote the mean and standard deviation of the training data. This normalisation step is critical for mitigating scale discrepancies between high-voltage rechargeable lithium-ion cells (approximately 4.2–2.7 V) and low-voltage primary CR2032 coin cells (approximately 3.0–1.8 V), thereby enabling effective transfer learning across domains [7].

Feature selection was adapted to the characteristics of each dataset. For the CR2032-D target domain, the input feature space is limited exclusively to voltage-time discharge profiles. In contrast, the NASA dataset includes voltage, current, temperature, and capacity measurements, while the CALCE dataset incorporates voltage, current, internal resistance, and capacity. This discrepancy in feature availability further exacerbates the domain shift between source and target domains.

To enable time-series modelling, all signals were reconstructed using a sliding window approach. Each sample was represented as a fixed-length sequence of T timesteps, resulting in three-dimensional input tensors of shape $[N, T, F]$, where N denotes the number of samples and F the number of features. For CR2032-D, $F = 1$, reflecting the voltage-only observability of the target domain. This formulation allows recurrent and attention-based architectures to capture both local voltage decay patterns and longer-term discharge dynamics [3].

Noisy or incomplete measurements were removed during preprocessing to prevent bias in model training. The CR2032-D dataset was pooled across all brands and randomly split into training (70%), validation (15%), and test (15%) subsets. Brand identity was intentionally excluded as an explicit input feature, encouraging the models to learn brand-invariant discharge representations. This setting reflects realistic deployment scenarios in which battery manufacturer information may be unavailable at inference time.

Brand-induced domain shift has a direct and detrimental impact on battery life prediction. Differences in voltage-time trajectories across brands introduce statistical distribution discrepancies between training and deployment data, as evidenced in figure 4. Models trained on one brand therefore tend to exhibit degraded performance when applied to unseen brands, unless appropriate domain adaptation or transfer learning strategies are employed [5].

Given these characteristics, the CR2032-D target domain poses a uniquely challenging learning problem.

In this study, the learning task is formulated as a discharge-only condition inference problem. The prediction target, referred to as the *discharge progression indicator* (DPI), is defined as the normalised elapsed discharge time at the end of each input window:

$$\mathrm{DPI}(t) = \frac{t}{t_{\mathrm{total}}} \in [0, 1], \tag{2}$$

where t denotes the timestamp (in minutes) at the end of the current window and t_{total} is the total discharge duration until the cut-off voltage of 1.799 V is reached. Under the constant-current protocol applied in this study ($I = 0.1$ A), the accumulated charge is directly proportional to elapsed time, i.e. $Q(t) = I \cdot t$. Consequently, the DPI is mathematically equivalent to the normalised consumed capacity,

and thus to the *depth of discharge* (DoD) [25]:

$$\mathrm{DoD}\,(t) = \frac{Q\,(t)}{Q_{\mathrm{total}}} = \frac{I \cdot t}{I \cdot t_{\mathrm{total}}} = \frac{t}{t_{\mathrm{total}}} = \mathrm{DPI}\,(t)\,. \tag{3}$$

A DPI value of 0 corresponds to the very beginning of discharge, while a value of 1 indicates complete discharge at the cut-off voltage. The reported MAE and MSE values are therefore expressed as fractions of total discharge progression, directly proportional to the fraction of charge consumed. This formulation is consistent with recent data-driven approaches that estimate SoC or DoD of primary batteries from voltage-only measurements in IoT applications [1], and with Coulomb-counting-based SoC frameworks that reduce to a time-ratio under constant-current conditions [26].

For each sample, the model receives a fixed-length input sequence $\mathbf{v} = [v_i, v_{i+1}, \ldots, v_{i+T-1}]$ of T consecutive voltage measurements and is trained to predict the DPI at the end of that window, i.e. $\hat{y} = \mathrm{DPI}(t_{i+T-1})$. The sliding-window construction yields input tensors of shape $[N, T, 1]$, where N is the number of samples and the single feature corresponds to the voltage-only observability of the CR2032-D domain.

Unlike conventional SOH definitions that require long-term cycling data, explicit capacity measurements, or charge-event access, the DPI is fully observable from a single discharge event. By learning the mapping from early-stage voltage windows to the DPI target, the proposed models aim to capture subtle degradation-related signatures embedded in the flat discharge plateau without manufacturer-specific calibration [27], enabling reliable condition inference in resource-constrained IoT deployments [1].

3.3. Deep learning models and transfer learning strategy

In this study, deep learning models, including multilayer perceptron (MLP), CNN, LSTM, GRU, and attention-based architectures were trained for battery condition inference and their performance was compared. The performance comparisons were based on the study conducted by Wang *et al* [28]. During the experiments, pre-trained models were utilised to apply transfer learning on the NASA, CALCE, and CR2032-D datasets.

3.3.1. MLP

A MLP is a type of feedforward artificial neural network consisting of an input layer, multiple hidden layers, and an output layer. Each neuron in a layer is fully connected to neurons in the adjacent layers. The output of the ith neuron in layer l is computed as:

$$z_i^{[l]} = \sum_{j=1}^{n^{[l-1]}} W_{i,j}^{[l]} a_j^{[l-1]} + b_i^{[l]} \tag{4}$$

$$a_i^{[l]} = f\left(z_i^{[l]}\right). \tag{5}$$

Here, $W_{i,j}^{[l]}$ denotes the weight between neuron j in layer $(l-1)$ and neuron i in layer l, $b_i^{[l]}$ is the bias term, $a_j^{[l-1]}$ is the activated output from the previous layer, and $f(\cdot)$ represents the activation function. ReLU, sigmoid, or tanh functions are commonly used as activation functions.

Due to its lack of explicit temporal modelling, the MLP serves as a baseline to evaluate the importance of sequential information in discharge-only battery condition inference.

3.3.2. CNN

A 1D CNN is used for sequential or time-series data to capture local temporal patterns. For a single filter k, the convolution operation at position t is defined as:

$$z_t^{(k)} = \sum_{i=0}^{F-1} W_i^{(k)} x_{t+i} + b^{(k)} \tag{6}$$

$$a_t^{(k)} = f\left(z_t^{(k)}\right). \tag{7}$$

Here, F is the filter size, x_{t+i} is the input at time step $t+i$, $W_i^{(k)}$ is the filter weight, $b^{(k)}$ is the bias term, and $f(\cdot)$ is the activation function. The outputs are downsampled using MaxPooling1D layers, followed by fully connected layers for the final prediction.

While effective at capturing local temporal patterns, CNNs lack explicit long-term memory, making them sensitive to flat discharge plateaus commonly observed in primary batteries. The CNN architecture employed in this study uses MaxPooling1D following Conv1D layers, consistent with standard 1D CNN designs for time-series regression. While global average pooling (GAP) and large convolutional kernels are valid architectural alternatives, they address different inductive biases. GAP collapses the temporal dimension entirely, discarding the sequential ordering that is essential for tracking the monotonic voltage decay embedded in CR2032-D's flat discharge plateau. Large kernels widen the local receptive field but still extract fixed-width features without accumulating state across the full sequence. Neither modification addresses the fundamental architectural limitation: the absence of recurrent state propagation across the entire discharge trajectory. The experimental superiority of LSTM over all CNN variants is therefore a consequence of this architectural constraint rather than a hyperparameter choice.

3.3.3. LSTM

LSTM networks are designed to learn long-term dependencies in sequential data. Each LSTM cell maintains a cell state C_t and hidden state h_t, updated using input, forget, and output gates:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{8}$$

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{9}$$

$$\tilde{C}_t = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{10}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{11}$$

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \tag{12}$$

$$h_t = o_t \odot \tanh\left(C_t\right). \tag{13}$$

Here, σ is the sigmoid activation, $\odot$ denotes element-wise multiplication, and x_t is the input at time step t.

The explicit memory mechanism of LSTM networks makes them particularly suitable for capturing subtle early-stage degradation signatures embedded within long discharge plateaus.

3.3.4. GRU

GRU is a simplified variant of LSTM with only reset and update gates:

$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t] + b_z\right) \tag{14}$$

$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t] + b_r\right) \tag{15}$$

$$\tilde{h}_t = \tanh\left(W_h \cdot [r_t \odot h_{t-1}, x_t] + b_h\right) \tag{16}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t. \tag{17}$$

Here, z_t is the update gate controlling information retention, and r_t is the reset gate controlling how much past information to forget.

GRU provides a computationally efficient alternative to LSTM while retaining the ability to model long-term temporal dependencies in voltage-time sequences.

3.3.5. Attention mechanism

Attention allows the model to focus on the most relevant parts of the input sequence. For input matrices Query Q, Key K, and Value V, the attention output is computed as:

$$\text{Attention}\left(Q, K, V\right) = \text{softmax}\left(\frac{QK^{\top}}{\sqrt{d_k}}\right) V. \tag{18}$$

Here, d_k is the dimensionality of the key vectors, and the softmax function converts similarity scores into attention weights, allowing the model to emphasise important time steps in sequential data.

Model performance was evaluated using the following regression metrics. MAE represents the average of absolute differences between predicted and actual values, treating all errors equally:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{19}$$

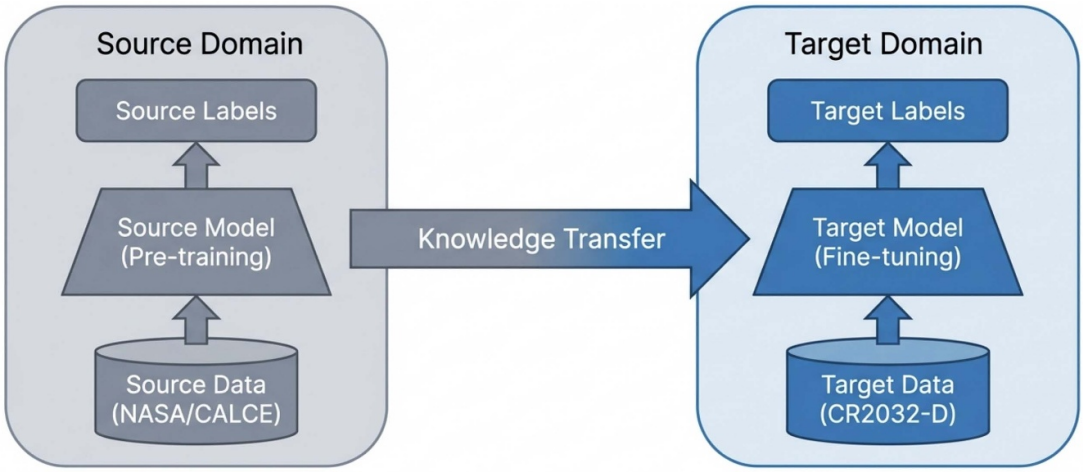


Figure 5. Illustration of the transfer learning paradigm adopted in this study.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{20}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{21}$$

where y_i denotes the true value of the ith sample, $\hat{y}_i$ represents the predicted value, $\bar{y}$ is the mean of the true values, and n corresponds to the total number of samples. Attention mechanisms enable adaptive weighting of informative time steps, complementing recurrent architectures under heterogeneous discharge dynamics.

3.3.6. Implementation of transfer learning

Transfer learning strategies can generally be categorised based on how knowledge from the source domain is adapted to the target domain. direct transfer learning (DTL), which applies a pre-trained source model to the target domain without any modification, often suffers from negative transfer when substantial discrepancies exist between source and target distributions [5, 8]. In the present case, the fundamental differences between rechargeable LIBs and primary coin-cell batteries make DTL unsuitable. To address this, a *partial transfer with input layer re-initialisation* strategy is adopted, as illustrated in figures 5 and 6.

3.3.6.1 Theoretical justification for cross-chemistry transfer

Despite their operational differences, rechargeable Li-ion and primary lithium coin-cell batteries share a common electrochemical foundation that supports cross-domain knowledge transfer. Both chemistries are governed by the same fundamental processes: lithium-ion intercalation and de-intercalation at electrode interfaces, solid electrolyte interphase dynamics, and diffusion-limited charge transport through the electrolyte [3]. These processes produce voltage-time discharge profiles that, while differing in absolute magnitude and shape, follow structurally analogous patterns: a characteristic open-circuit voltage determined by electrode potential, a quasi-flat discharge plateau governed by diffusion kinetics, and a sharp terminal voltage drop near EOD arising from reactant depletion. From a representation learning perspective, models pretrained on Li-ion cycling data learn to detect these electrochemical signatures in the temporal structure of voltage sequences. Z-score normalisation prior to fine-tuning aligns the statistical distributions of source and target voltage signals, enabling effective reuse of learned temporal representations. This is consistent with the DT-DCS transfer paradigm [5], which establishes that cross-chemistry transfer is viable when the shared observable signal (voltage) encodes analogous degradation dynamics.

In the first stage, all model architectures (MLP, CNN, LSTM, GRU, and Attention) are pretrained on the NASA and CALCE cycling datasets. For NASA, four input features are used per timestep: measured

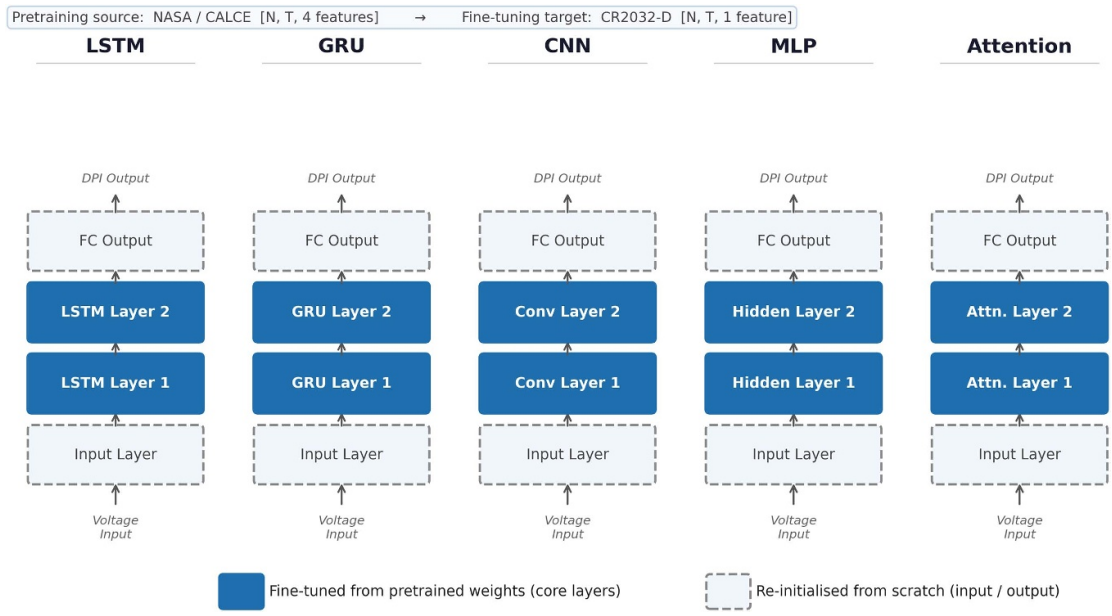


Figure 6. Transfer learning adaptation strategy adopted in this study. The input layer is re-initialised from scratch to accommodate the single-feature (voltage-only, $F = 1$) CR2032-D domain, while the pretrained core temporal layers, trained on four-feature NASA/CALCE data ($F = 4$), are transferred and fully fine-tuned on the target domain. The fully connected output layer is also re-initialised for DPI regression. Solid blue blocks indicate layers fine-tuned from pretrained weights; dashed blocks indicate layers re-initialised from scratch.

voltage, current, temperature, and capacity. For CALCE, the four features are voltage, current, internal resistance, and capacity. All signals are normalised using Z-score standardisation and reconstructed into fixed-length sequences using a sliding window of size $T = 4$, yielding input tensors of shape $[N, T, 4]$. This pretraining stage enables the models to learn general electrochemical degradation representations, including nonlinear voltage decay trends and long-term temporal dependencies common across lithium-based chemistries [3].

The CR2032-D target domain provides only voltage measurements, resulting in input tensors of shape $[N, T, 1]$. Since the input dimensionality differs from the source domain ($F = 4$ versus $F = 1$), the input projection layer of each pretrained model is discarded and re-initialised from scratch for the target domain. Crucially, voltage is the one signal present across all three datasets, and the internal temporal layers of the pretrained models were primarily trained to track voltage decay dynamics, the dominant signal driving capacity degradation. Consequently, re-initialising only the input layer does not discard the core transferred knowledge; rather, it removes the dependency on signals unavailable in the target domain (current, temperature, resistance) while preserving the voltage-based temporal representations that transfer directly to primary battery discharge profiles [7].

In the second stage, the pretrained temporal weights serve as a structured initialisation point for adaptation to CR2032-D. All temporal layers are subsequently fine-tuned on the target domain, allowing the model to adjust its internal representations to the unique characteristics of primary batteries: the extended quasi-flat voltage plateau and the abrupt terminal voltage drop near end-of-discharge. The fully connected output regression layer is also re-initialised and trained from scratch to align with the DPI prediction target defined in equation (2).

This full fine-tuning strategy from pretrained weights consistently outperforms both random initialisation (no transfer) and strict layer-freezing approaches on the CR2032-D domain. The pretrained temporal weights provide a significantly better optimisation starting point, accelerating convergence and improving generalisation under the limited-data conditions of the target domain, in line with established transfer learning principles [5, 7].

Overall, the adopted strategy effectively addresses the DT-DCS domain shift identified by Chen *et al* [5]. By preserving transferable voltage-based temporal knowledge while re-initialising domain-incompatible input and output layers, the framework ensures reliable and reproducible generalisation to primary battery applications despite the absence of cycling data and extensive target-domain labels.

Table 1. Training hyperparameters used in all experiments.

Parameter	Value
Maximum epochs	100
Learning rate	0.0001
Learning rate scheduler	StepLR
Scheduler step size	30 epochs
Gamma	0.5
Early stopping patience	10 epochs

Table 2. Batch sizes used for different models and datasets.

Model	CR2032-D	NASA	CALCE
GRU	128	128	128
Attention	128	128	128
MLP	512	128	128
CNN	64	64	256
LSTM	128	128	128

4. Experiments

This section reports on the experimental setting and comparable performance results of the proposed algorithm.

4.1. Experimental setup and parameter tuning

The training of the deep learning-based models was carried out with specific hyperparameter settings. For all models, the training process was limited to a maximum of 100 epochs, with an initial learning rate set to 0.0001. To gradually decrease the learning rate, it was updated by a factor of 0.5 every 30 epochs (StepLR: step size = 30, gamma = 0.5), ensuring a more stable learning process. To prevent overfitting, an early stopping mechanism was employed. The model training continued uninterrupted for the first 10 epochs; thereafter, if no improvement in validation loss was observed, the training was terminated early. The details of these settings are presented in table 1.

In addition, the batch sizes were optimised for each model depending on the structural differences of the models and datasets. The detailed configurations are presented in table 2.

4.2. Results and discussion

This section presents a comprehensive evaluation of the proposed discharge-only battery condition inference framework across three datasets: the newly introduced CR2032-D target domain and the NASA and CALCE lithium-ion source domains. The results are analysed with particular emphasis on model generalisation behaviour under domain shift, architectural robustness, and suitability for discharge-only primary battery scenarios based on performance metrics (MAE, MSE, and R^2).

Table 3 reports the training, validation, and test performance of all models on the CR2032-D dataset, which represents the most challenging experimental setting due to its discharge-only nature, flat voltage characteristics, limited observability, and brand-induced domain shift. Among all evaluated architectures, the LSTM model clearly outperforms the alternatives, achieving the lowest test MAE (0.05 957), lowest MSE (0.00 638), and the highest coefficient of determination ($R^2 = 0.99164$).

The experimental results provide clear and consistent evidence on the role of temporal modelling in discharge-only battery condition inference under domain shift. Across all evaluated datasets, LSTM, GRU, and attention-based models are capable of learning long-term temporal dependencies. They systematically outperform non-sequential approaches, highlighting the importance of exploiting the temporal structure embedded in voltage–time profiles.

This observation is particularly pronounced in the CR2032-D dataset. Primary lithium coin-cell batteries exhibit long voltage plateau regions followed by abrupt terminal drops, where degradation-related information is subtle and sparsely distributed over time. Capturing these early-stage variations requires memory-aware architectures that can preserve long-range temporal information. LSTM networks, by design, are well-suited to this task and consistently achieve the highest accuracy, confirming their ability

Table 3. Train, validation, and test performance of deep learning models on the CR2032-D dataset. Bold values indicate the best performance (lowest MAE/MSE and highest R^2) among the compared models.

Model	MAE	MSE	R^2
Train			
GRU	0.09 375	0.03 377	0.96 540
MLP	0.12 026	0.12 996	0.86 681
LSTM	**0.02 290**	**0.00 962**	**0.99 014**
Attention	0.05 624	0.02 400	0.97 541
CNN	0.11 442	0.05 644	0.94 216
Validation			
GRU	0.31 483	0.23 020	0.84 765
MLP	0.13 069	0.14 818	0.90 193
LSTM	**0.05 578**	**0.02 274**	**0.98 495**
Attention	0.15 548	0.06 179	0.95 911
CNN	0.69 710	0.84 746	0.43 915
Test			
GRU	0.39 534	0.26 361	0.65 486
MLP	0.14 060	0.03 729	0.95 118
LSTM	**0.05 957**	**0.00 638**	**0.99 164**
Attention	0.11 422	0.02 173	0.97 155
CNN	0.21 648	0.08 044	0.89 468

to extract degradation-sensitive patterns that remain indistinguishable to models relying on local or static representations. Attention-based models also demonstrate strong performance on CR2032-D, achieving a test R^2 of 0.97 155, indicating that adaptive temporal weighting is beneficial. However, their slightly inferior performance relative to LSTM suggests that explicit recurrent memory remains critical under flat discharge dynamics.

In contrast, CNN models exhibit poor generalisation on CR2032-D, with large validation and test errors. Although convolutional filters can effectively capture local temporal patterns, they fail to model the global discharge evolution of primary batteries. The absence of recurrent state accumulation means that even with architectural modifications such as GAP or larger kernels, CNN-based models cannot track the subtle monotonic voltage decay that characterises the flat discharge plateau of CR2032-D cells. Similarly, MLP architectures, which lack any temporal modelling mechanism, struggle to generalise across operating regimes, reinforcing the necessity of sequential representations for reliable condition inference.

The results obtained on the NASA lithium-ion dataset (table 4) further support these conclusions. Across batteries B0005, B0006, B0007, and B0018, recurrent architectures consistently achieve high test accuracy, with R^2 values typically exceeding 0.95. While CNN models attain near-perfect training performance, their validation and test accuracy deteriorate substantially, indicating severe overfitting and limited robustness to distribution shift. Such behaviour is particularly undesirable in transfer learning scenarios, where adaptability to unseen domains is essential. Attention-based models show competitive but less stable performance, suggesting sensitivity to dataset-specific characteristics.

A similar performance hierarchy is observed on the CALCE dataset (table 5). GRU, LSTM, and attention-based models maintain strong generalisation across different cells, whereas CNN and MLP architectures again exhibit pronounced degradation during testing. These consistent trends across three fundamentally different datasets confirm that local pattern extraction or static mappings alone are insufficient for robust battery condition inference.

Crucially, these findings directly validate the proposed discharge-only transfer learning framework. Although lithium-ion source datasets provide rich cycling-based information, successful adaptation to primary battery domains requires models that can abstract domain-invariant temporal representations rather than memorising dataset-specific voltage patterns. As illustrated by the convergence behaviour in figure 7, recurrent models not only achieve lower prediction errors but also demonstrate more stable optimisation under pronounced domain shift. Additional experimental results on multiple battery datasets, including NASA B0005, B0006, B0007, B0018, and CALCE, are presented in the appendix (see appendix) to demonstrate the robustness and generalisation performance of the proposed models.

Table 4. Performance comparison on the NASA dataset (B0005–B0018). Bold values indicate the best performance (lowest MAE/MSE and highest R^2) among the compared models.

	Train			Validation			Test		
Model	MAE	MSE	R^2	MAE	MSE	R^2	MAE	MSE	R^2
B0005									
GRU	0.06 039	0.04 197	0.95 807	**0.05 950**	0.02 788	0.97 304	**0.07 683**	**0.03 099**	**0.96 963**
MLP	0.11 455	0.05 971	0.94 036	0.08 994	0.03 923	0.96 206	0.15 836	0.07 506	0.92 642
LSTM	0.06 633	0.04 235	0.95 770	0.06 989	**0.02 715**	**0.97 374**	0.10 969	0.03 859	0.96 217
Attention	0.07 472	0.03 888	0.96 116	0.09 952	0.03 852	0.96 275	0.19 816	0.08 114	0.92 047
CNN	**0.04 134**	**0.00 548**	**0.99 452**	0.39 073	0.31 567	0.69 469	0.37 612	0.38 722	0.62 043
B0006									
GRU	0.05 671	0.06 353	0.93 649	**0.04 590**	0.02 724	0.97 209	**0.05 186**	**0.02 758**	**0.97 371**
MLP	0.11 403	0.07 683	0.92 318	0.08 429	0.03 444	0.96 470	0.15 173	0.05 743	0.94 370
LSTM	0.06 583	0.06 475	0.93 526	0.06 571	0.02 905	0.97 023	0.11 747	0.04 486	0.95 723
Attention	0.07 607	0.04 993	0.95 008	0.06 865	**0.02 537**	**0.97 400**	0.14 406	0.04 611	0.95 605
CNN	**0.04 761**	**0.00 816**	**0.99 184**	0.41 309	0.27 803	0.71 507	0.54 887	0.45 764	0.56 374
B0007									
GRU	0.07 176	0.08 222	0.91 790	**0.06 232**	0.04 137	0.96 207	**0.08 953**	**0.06 832**	**0.93 775**
MLP	0.11 481	0.10 226	0.89 789	0.08 145	0.05 449	0.95 004	0.15 054	0.10 741	0.90 213
LSTM	0.07 341	0.08 248	0.91 764	0.08 784	0.04 735	0.95 658	0.13 850	0.07 441	0.93 221
Attention	0.07 578	0.05 318	0.94 690	0.06 409	**0.03 783**	**0.96 532**	0.12 437	0.08 920	0.91 872
CNN	**0.04 937**	**0.01 292**	**0.98 710**	0.35 335	0.36 166	0.66 840	0.43 879	0.52 108	0.52 521
B0018									
GRU	**0.06 335**	0.04 434	0.95 563	0.07 287	0.04 373	0.96 049	**0.07 190**	**0.04 846**	**0.95 748**
MLP	0.08 928	0.04 920	0.95 077	**0.07 259**	**0.03 762**	**0.96 601**	0.10 787	0.05 973	0.94 758
LSTM	0.06 669	0.04 324	0.95 673	0.07 737	0.04 461	0.95 969	0.08 053	0.05 130	0.95 498
Attention	0.08 754	0.04 492	0.95 505	0.10 145	0.04 165	0.96 236	0.14 574	0.07 084	0.93 783
CNN	0.06 690	**0.02 285**	**0.97 714**	0.37 959	0.34 248	0.69 053	0.49 460	0.52 205	0.54 188

Table 5. Performance comparison on the CALCE dataset (CS2_35–CS2_38). Bold values indicate the best performance (lowest MAE/MSE and highest R^2) among the compared models.

	Train			Validation			Test		
Model	MAE	MSE	R^2	MAE	MSE	R^2	MAE	MSE	R^2
CS2_35									
GRU	**0.03 683**	0.02 220	0.97 780	0.06 919	0.02 863	0.97 165	0.10 382	0.05 063	0.95 200
MLP	0.09 171	0.06 328	0.93 671	0.17 118	0.28 882	0.71 397	0.37 406	0.58 215	0.44 808
LSTM	0.03 900	0.03 414	0.96 585	**0.04 588**	0.03 703	0.96 333	0.08 078	0.05 814	0.94 488
Attention	0.04 453	0.00 956	0.99 044	0.05 013	**0.01 378**	**0.98 636**	**0.05 495**	**0.02 204**	**0.97 911**
CNN	0.03 864	**0.00 461**	**0.99 539**	0.12 307	0.09 599	0.90 493	0.20 753	0.23 985	0.77 261
CS2_36									
GRU	**0.02 344**	0.01 653	0.98 347	**0.04 664**	**0.02 183**	**0.97 995**	**0.09 101**	**0.02 576**	**0.97 693**
MLP	0.13 450	0.20 496	0.79 497	0.26 302	0.31 899	0.70 695	0.47 818	0.52 863	0.52 649
LSTM	0.03 894	0.03 283	0.96 716	0.05 813	0.04 266	0.96 081	0.10 914	0.03 943	0.96 468
Attention	0.03 360	**0.00 429**	**0.99 571**	0.07 467	0.04 880	0.95 516	0.21 841	0.23 910	0.78 583
CNN	0.06 019	0.01 119	0.98 881	0.15 884	0.25 578	0.76 502	0.45 652	0.76 929	0.31 093
CS2_37									
GRU	**0.02 277**	0.01 316	0.98 683	0.02 823	**0.02 453**	0.97 696	0.05 662	**0.03 097**	**0.97 278**
MLP	0.10 589	0.12 831	0.87 164	0.16 241	0.24 707	0.76 794	0.30 326	0.39 234	0.65 514
LSTM	0.02 298	0.01 932	0.98 068	**0.02 673**	0.02 940	0.97 239	**0.04 146**	0.03 352	0.97 054
Attention	0.03 932	0.01 132	0.98 867	0.06 276	0.02 259	**0.97 878**	0.10 934	0.06 172	0.94 574
CNN	0.04 813	**0.00 742**	**0.99 258**	0.10 799	0.16 207	0.84 777	0.29 875	0.53 486	0.52 986
CS2_38									
GRU	0.04 085	0.02 942	0.97 057	0.05 094	0.02 772	0.97 385	0.09 483	0.04 460	0.95 908
MLP	0.10 106	0.06 320	0.93 679	0.14 475	0.24 096	0.77 276	0.31 876	0.54 491	0.50 006
LSTM	**0.02 607**	0.02 341	0.97 658	**0.02 988**	0.02 579	0.97 568	**0.05 009**	0.03 897	0.96 424
Attention	0.03 784	**0.01 319**	**0.98 681**	0.04 119	**0.01 691**	**0.98 405**	0.06 057	**0.02 457**	**0.97 746**
CNN	0.07 226	0.01 337	0.98 662	0.09 462	0.16 909	0.84 054	0.16 181	0.36 282	0.66 712

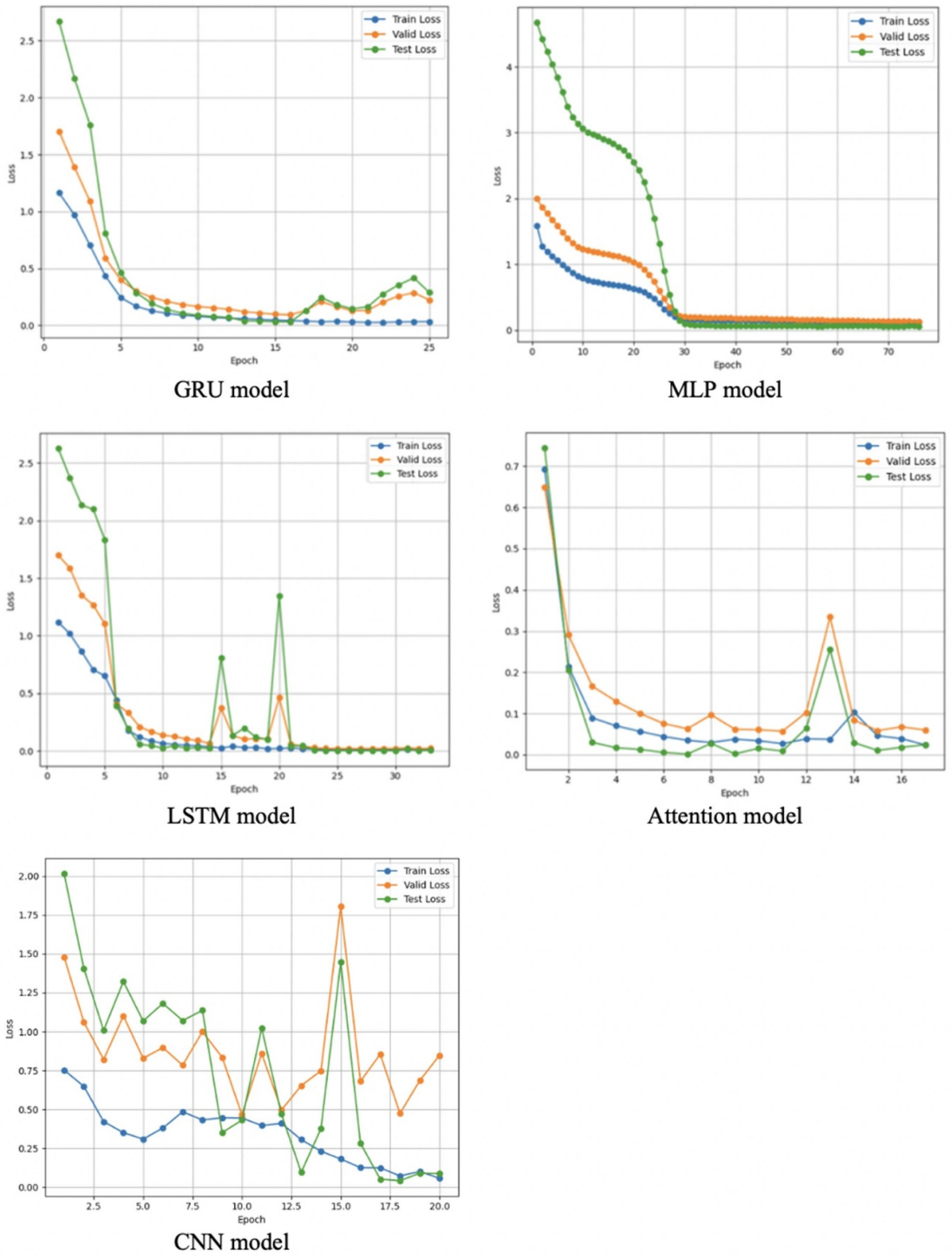


Figure 7. Training, validation, and test loss curves of different deep learning models for CR2032-D primary batteries.

Overall, the results confirm the feasibility of early-stage discharge-based condition inference for primary batteries, even in the absence of recharge cycles, capacity measurements, or manufacturer-specific labels. Transfer learning from cycling-based lithium-ion datasets is shown to be not only viable but essential for enabling data-efficient and robust deployment in realistic primary battery monitoring scenarios.

4.3. Sensitivity to voltage measurement resolution

In resource-constrained IoT deployments, voltage measurements are typically acquired using low-cost analogue-to-digital converters (ADCs) with limited bit resolution. To assess the robustness of the proposed framework under such conditions, a controlled sensitivity analysis was conducted by uniformly quantising CR2032-D voltage measurements to 12-bit, 10-bit, 8-bit, and 6-bit resolution over the

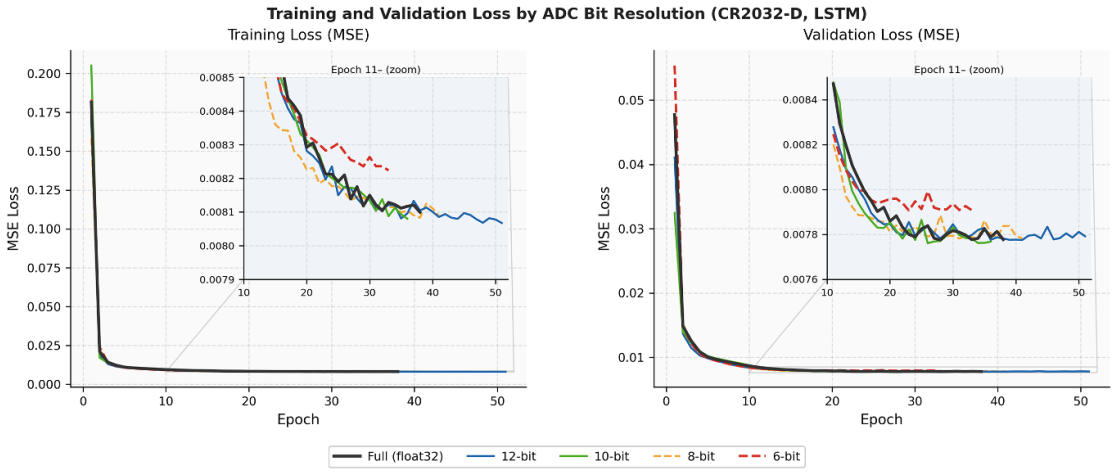


Figure 8. Training and validation loss curves by ADC bit resolution (CR2032-D, LSTM). The zoomed inset (epoch 11 onwards) highlights the marginal separation between full-precision and quantised configurations; 6-bit resolution exhibits slightly higher and noisier validation loss relative to higher-resolution counterparts.

Table 6. Effect of ADC bit resolution on LSTM DPI prediction performance (CR2032-D test set).

Resolution	MAE	MSE	R^2
Full (float32)	0.07 008	0.00 781	0.90 826
12-bit	0.06 977	0.00 780	0.90 841
10-bit	0.06 946	0.00 779	0.90 848
8-bit	0.06 973	0.00 779	0.90 852
6-bit	0.07 021	0.00 802	0.90 585

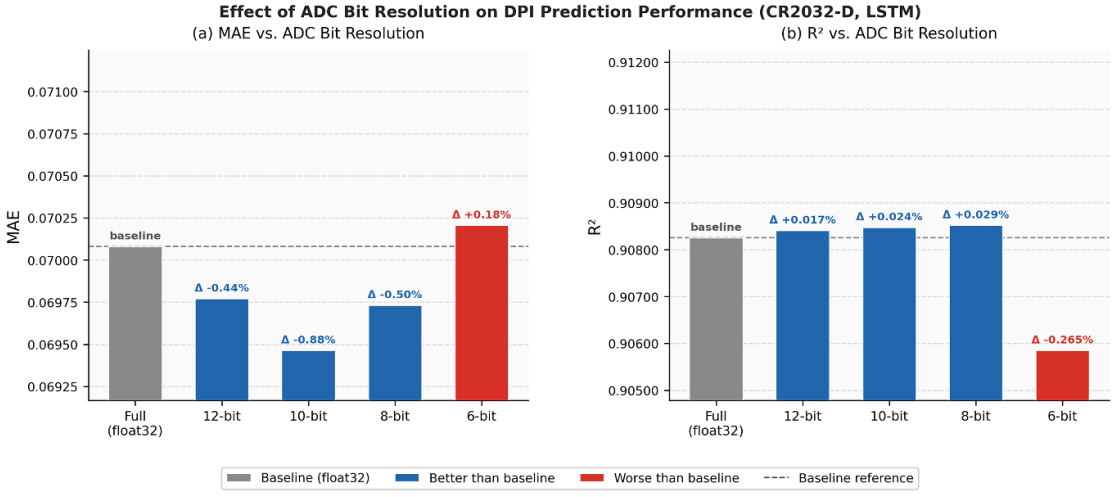


Figure 9. Effect of ADC bit resolution on DPI prediction performance (CR2032-D, LSTM). Panel (a): MAE; Panel (b): R^2. Blue bars indicate configurations that match or outperform the full-precision baseline; the red bar (6-bit) indicates marginal degradation. Δ values denote relative change from the float32 baseline. The dashed line marks the baseline reference.

observed operating range [1.799 V, 3.307 V], and evaluating LSTM prediction performance under each configuration.

The results, summarised in table 6 and Figure 9, demonstrate strong robustness across all tested resolutions. Notably, 12-bit, 10-bit, and 8-bit configurations achieve MAE values marginally below the full-precision baseline, suggesting that moderate quantisation introduces a mild regularisation effect without degrading predictive performance. Even at 6-bit resolution, the performance degradation remains marginal ($R^2 = 0.906$ versus 0.908 at baseline). These findings confirm that standard 8-bit or 10-bit ADCs, widely used in embedded IoT microcontrollers, are fully sufficient for reliable DPI prediction without meaningful accuracy loss.

The convergence behaviour under all quantisation levels is illustrated in figure 8. Training and validation loss curves are nearly identical across resolutions during the first ten epochs. The zoomed inset reveals that 6-bit quantisation produces marginally higher and more oscillatory validation loss in later epochs, consistent with the slight performance decline observed in table 6. This confirms that reduced voltage resolution does not adversely affect model training dynamics within practically relevant bit depths.

4.4. Limitations

The generalisability of the proposed framework is subject to certain boundary conditions that inform directions for future work. First, all experiments are conducted under a single fixed discharge current of 0.1 A, representative of the typical low-power operating regime of CR2032 cells in IoT applications. The DPI formulation is inherently normalised with respect to total discharge duration, which partially mitigates C-rate dependency for trajectories with analogous voltage shapes; however, systematic validation under higher discharge rates or variable-current scenarios remains a priority direction for future work. Second, CR2032-D encompasses four battery brands and does not exhaustively represent all manufacturers or cell generations; performance on unseen brands may vary and requires further validation. Finally, the current models produce point estimates of the DPI without associated confidence intervals; uncertainty-aware prediction would be valuable for safety-critical monitoring applications.

5. Conclusion

This study presented a cross-domain transfer learning framework for condition inference of primary lithium coin-cell batteries operating under discharge-only conditions. To reflect realistic deployment scenarios in long-life IoT and sensing applications, a new dataset, CR2032-Discharge (CR2032-D), was introduced, capturing brand-dependent variability in constant-current discharge behaviour. The prediction target, the DPI, was formally defined as the normalised elapsed discharge time and shown to be mathematically equivalent to the DoD under constant-current operation, making it fully observable from a single discharge event without requiring cycling data or explicit capacity measurements.

By leveraging transfer learning from established lithium-ion benchmark datasets, the proposed framework enables effective knowledge transfer across battery chemistries and operating regimes. Extensive experiments demonstrate that sequential deep learning models, particularly LSTM and GRU architectures, achieve the most reliable and robust performance under pronounced domain shift. These models consistently outperform convolutional and fully connected networks, which exhibit limited generalisation capability due to their inability to accumulate state across the full discharge trajectory.

The results confirm that early-stage discharge data contain sufficient information for meaningful battery condition inference, even in the absence of recharge cycles or extensive labelled datasets. A sensitivity analysis further demonstrates that the framework remains robust under low-cost ADC conditions, with negligible performance degradation down to 8-bit voltage resolution, supporting its applicability to resource-constrained embedded systems. This finding significantly broadens the applicability of data-driven battery monitoring techniques to primary batteries, where conventional health estimation approaches are not feasible. Future work will focus on incorporating uncertainty-aware prediction, extending the framework to additional primary battery chemistries and multi-rate discharge protocols, and exploring online adaptation strategies for real-time deployment.

Data availability statement

The datasets used in this study are publicly available. The NASA Li-ion Battery Aging dataset can be accessed via the NASA Open Data Portal [23] at https://data.nasa.gov/dataset/li-ion-battery-aging-datasets, and the CALCE battery dataset is available from the Center for Advanced Life Cycle Engineering (CALCE) [24] at https://calce.umd.edu/battery-data. The CR2032-D dataset generated and analysed during the current study is available from the corresponding author upon reasonable request, since additional analyses using this dataset are currently in progress.

Conflict of interest

The authors declare that they have no conflict of interest.

Consent for publication

The authors hereby give their consent for the publication of this research article.

Author contributions

Sultan Zeybek 0000-0002-1298-9499
Conceptualization (lead), Data curation (lead), Formal analysis (lead), Investigation (lead), Methodology (equal), Project administration (lead), Resources (lead), Software (equal), Supervision (lead), Validation (lead), Visualization (lead), Writing – original draft (lead), Writing – review & editing (lead)

Imen Turki
Data curation (supporting), Formal analysis (supporting), Methodology (supporting), Software (supporting), Validation (supporting), Writing – original draft (supporting)

Appendix. Additional experiments

This appendix presents additional experimental results obtained from multiple battery datasets to further evaluate the generalisation performance of the proposed deep learning models. The results include training, validation, and test loss convergence behaviours for the NASA B0005, B0006, B0007, B0018, and CALCE datasets.

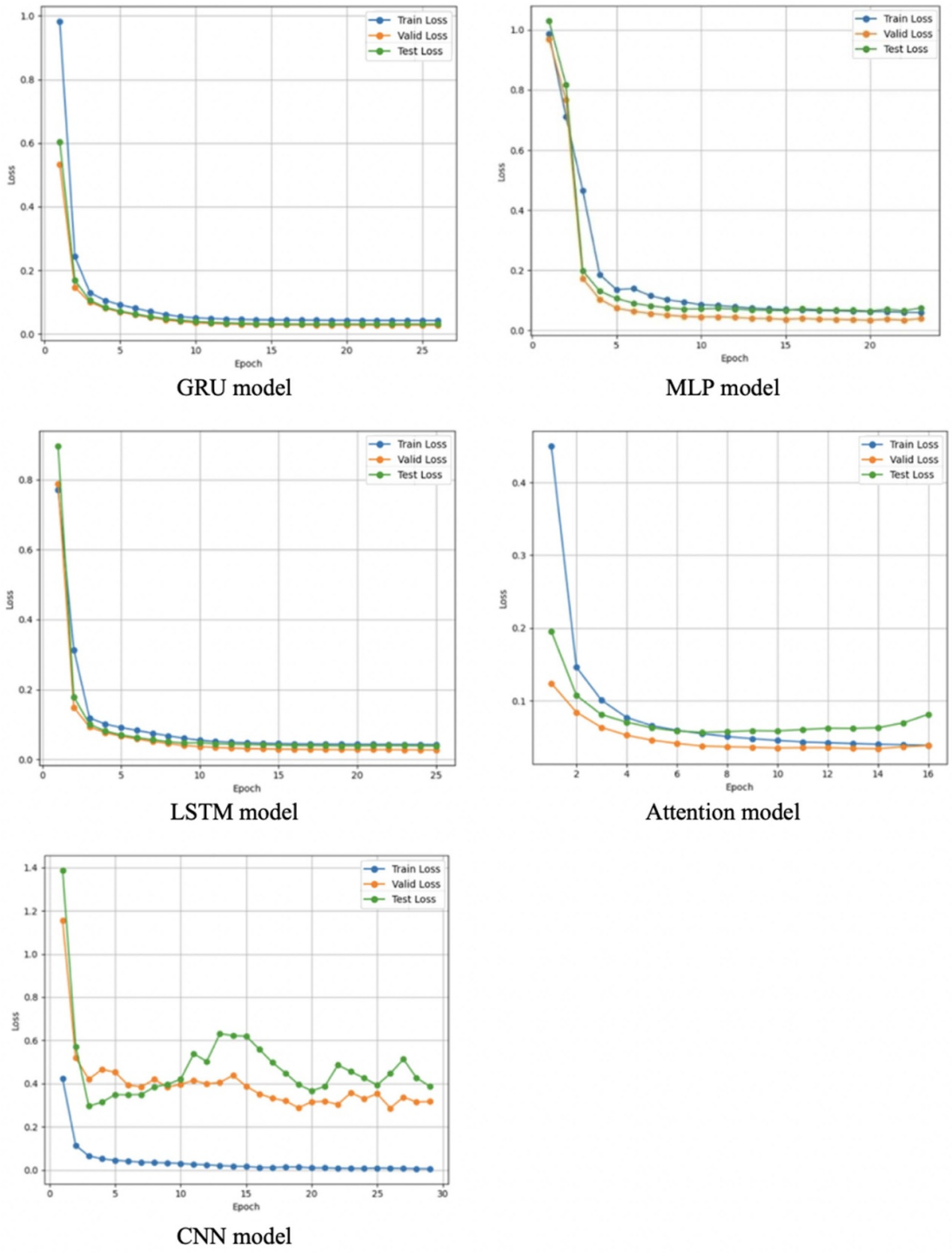


Figure 10. Training, validation, and test loss curves of different deep learning models for the NASA B0005 battery.

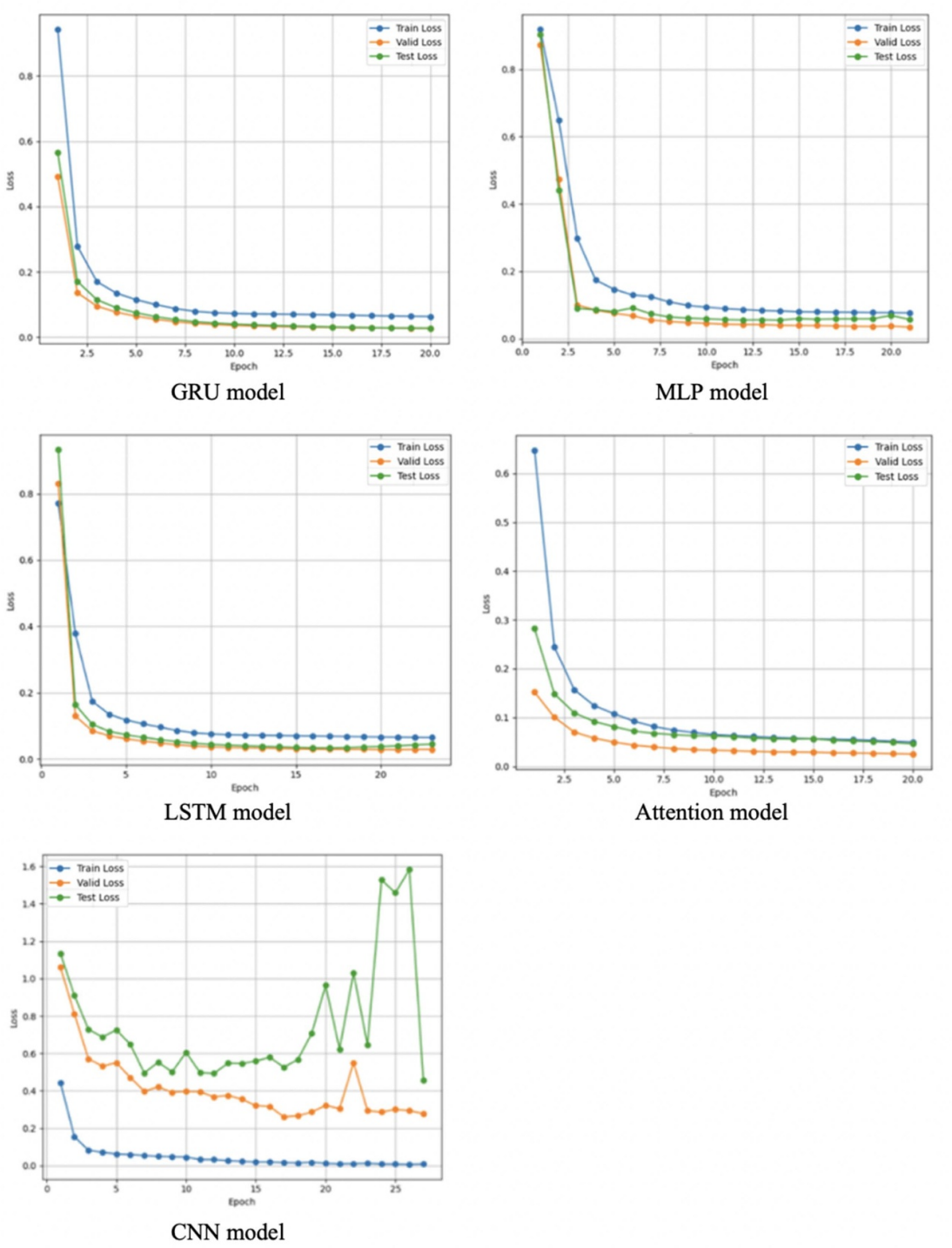


Figure 11. Training, validation, and test loss curves of different deep learning models for the NASA B0006 battery.

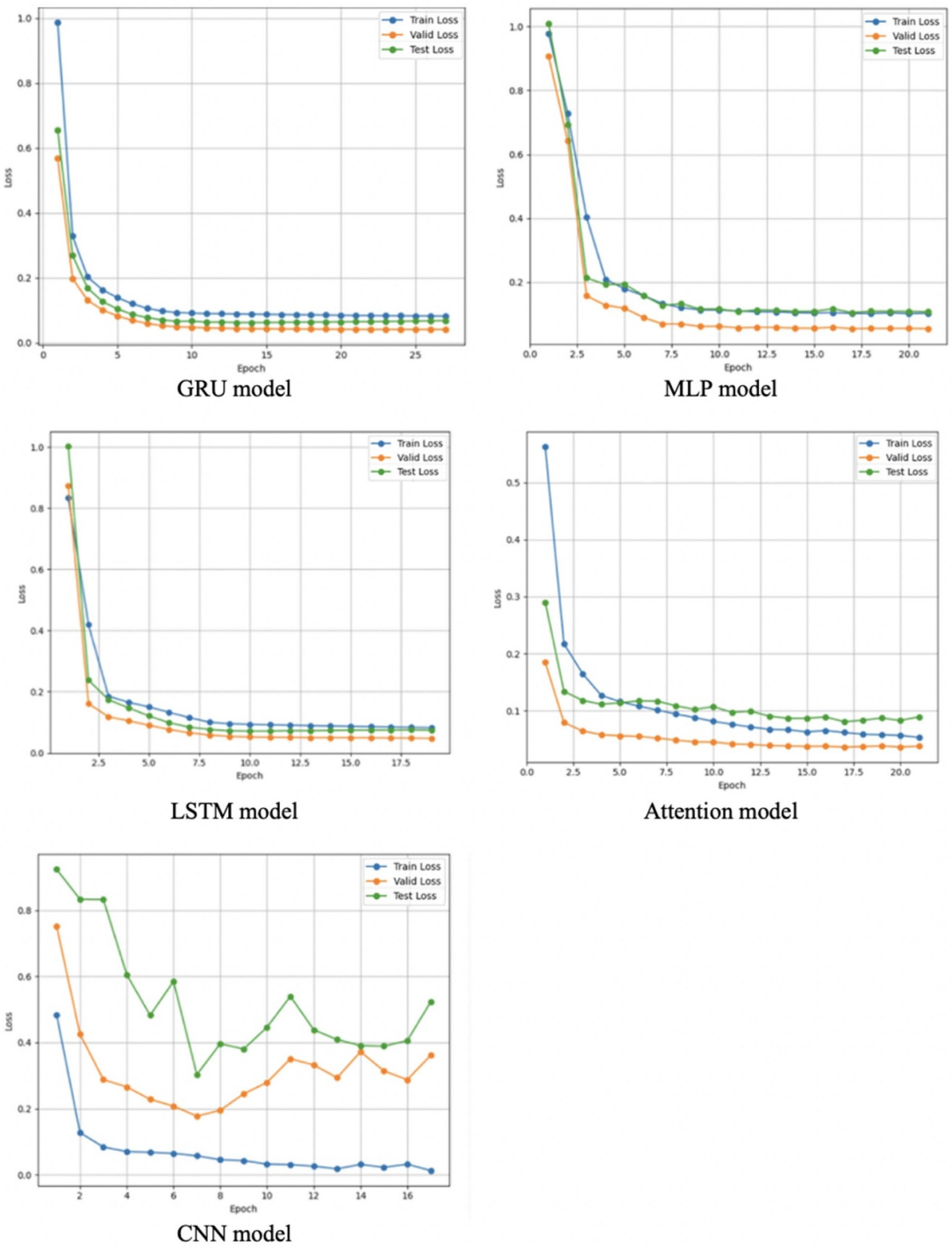


Figure 12. Training, validation, and test loss curves of different deep learning models for the NASA B0007 battery.

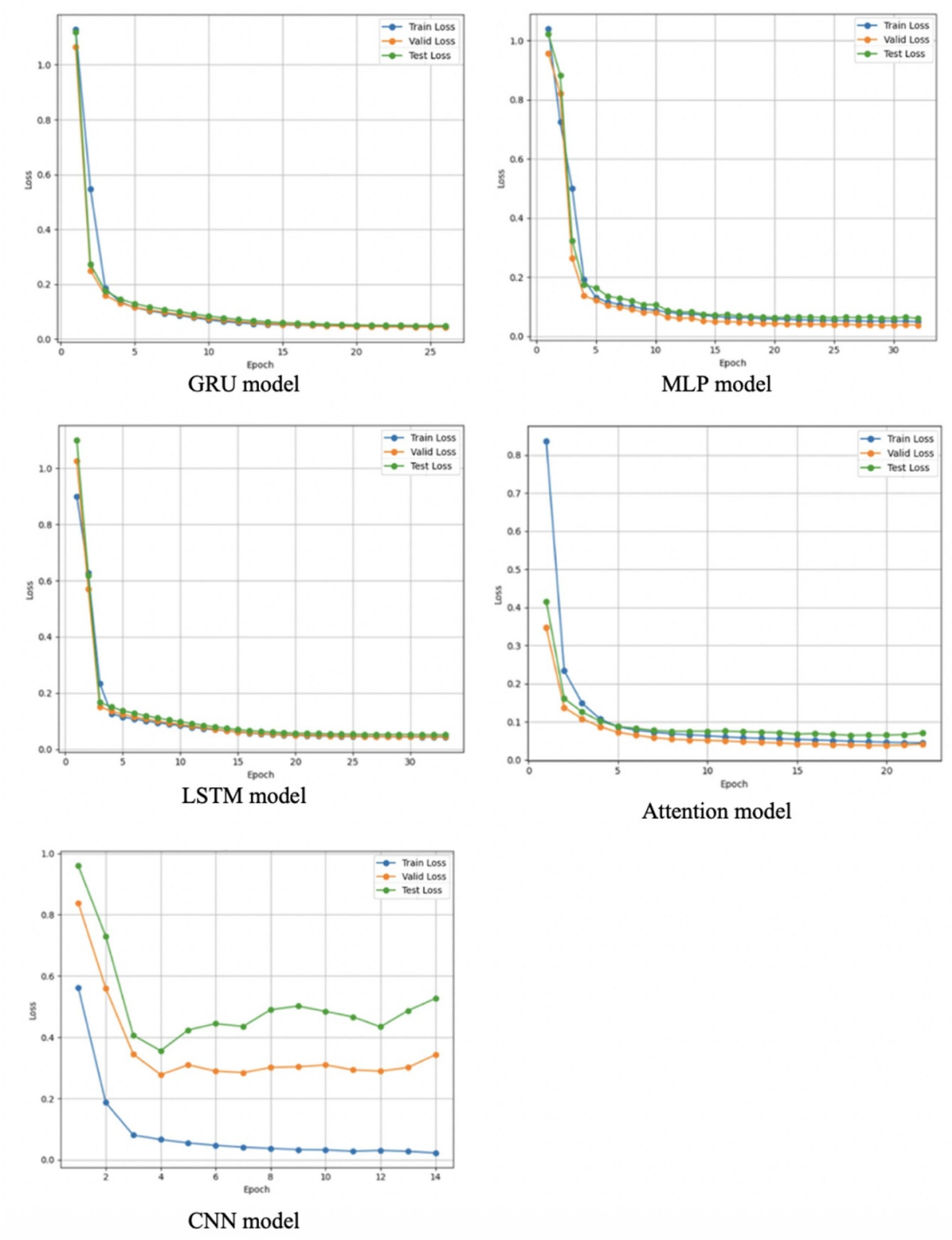


Figure 13. Training, validation, and test loss curves of different deep learning models for the NASA B0018 battery.

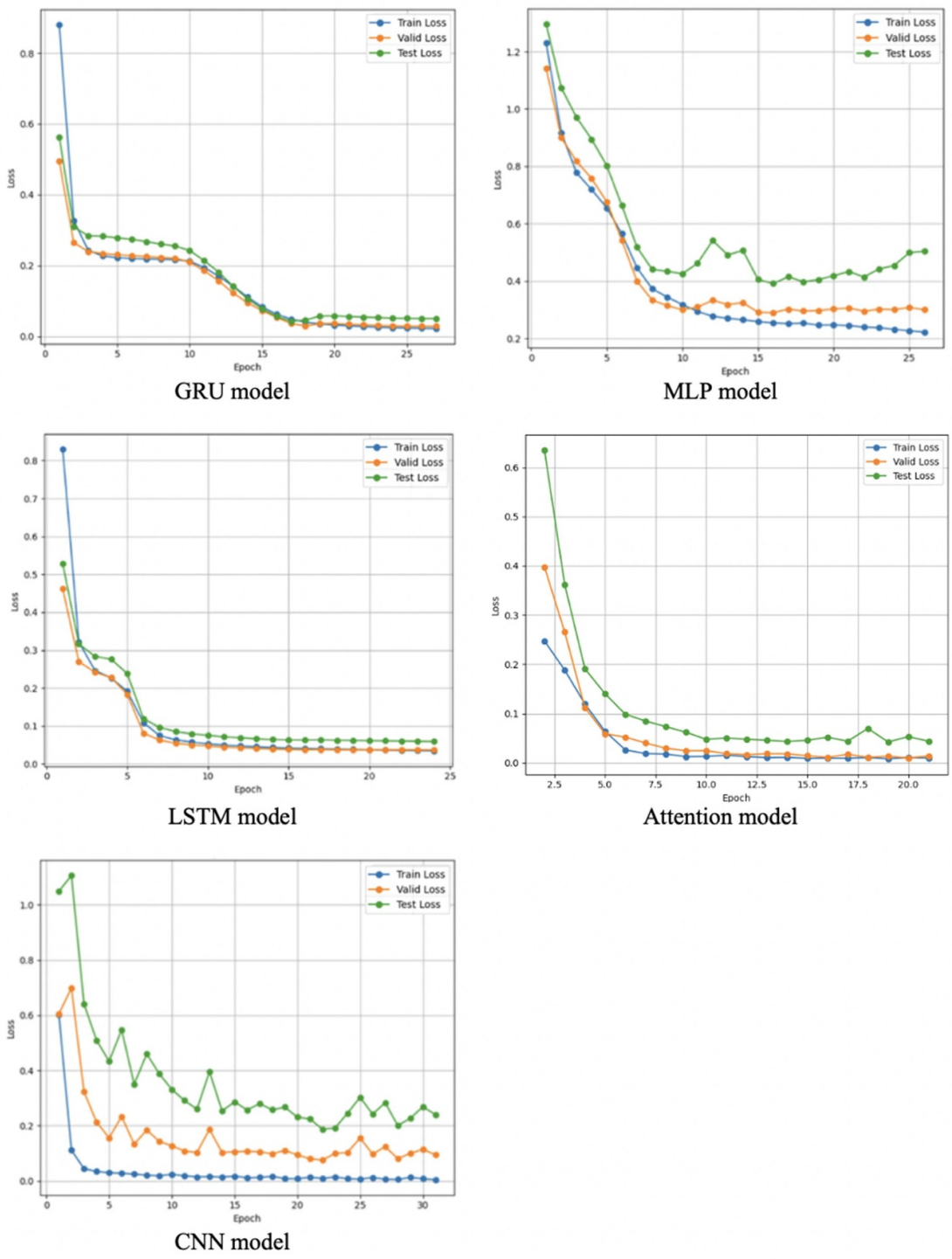


Figure 14. Training, validation, and test loss curves of different deep learning models for the CALCE CS2_35 dataset.

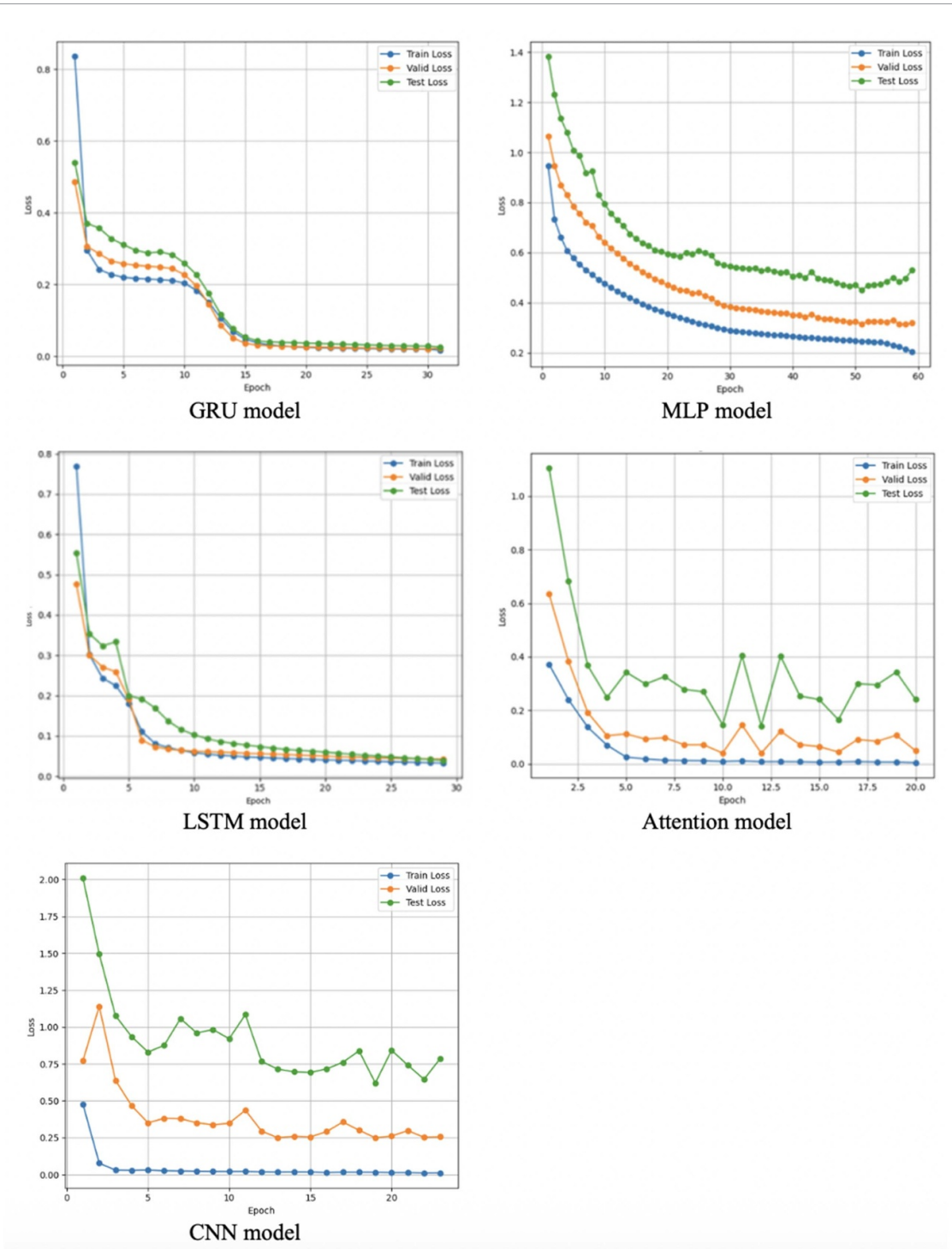
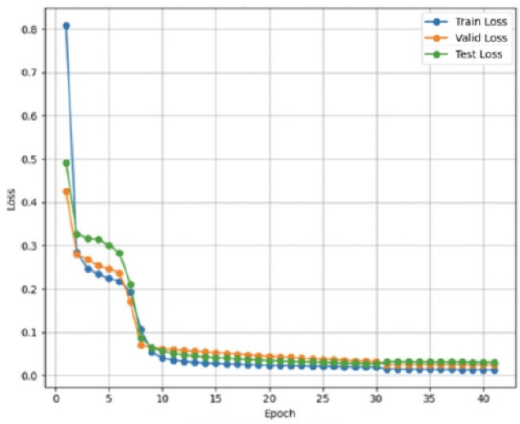
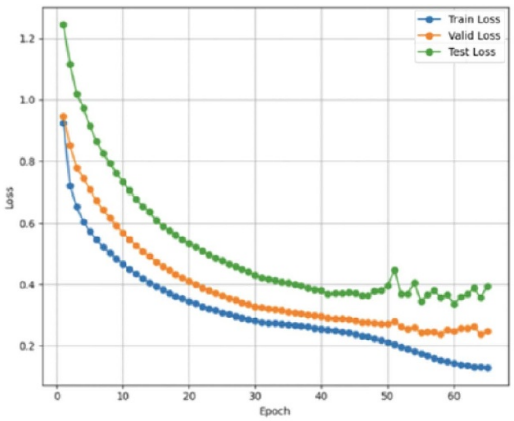


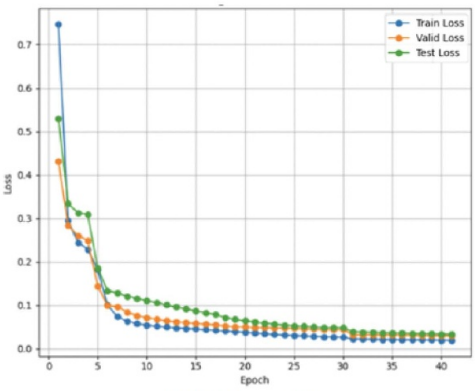
Figure 15. Training, validation, and test loss curves of different deep learning models for the CALCE CS2_36 dataset.



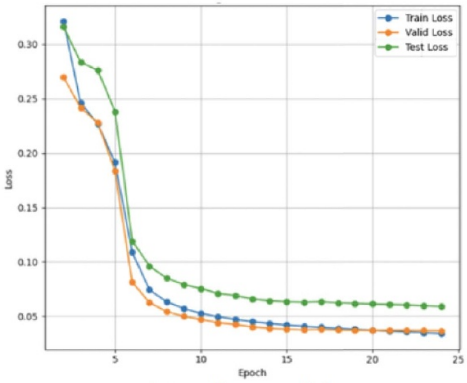
GRU model



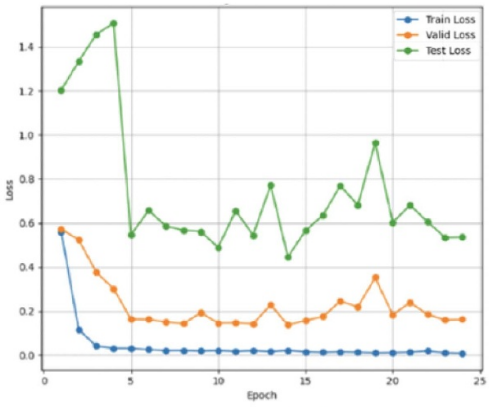
MLP model



LSTM model



Attention model



CNN model

Figure 16. Training, validation, and test loss curves of different deep learning models for the CALCE CS2_37 dataset.

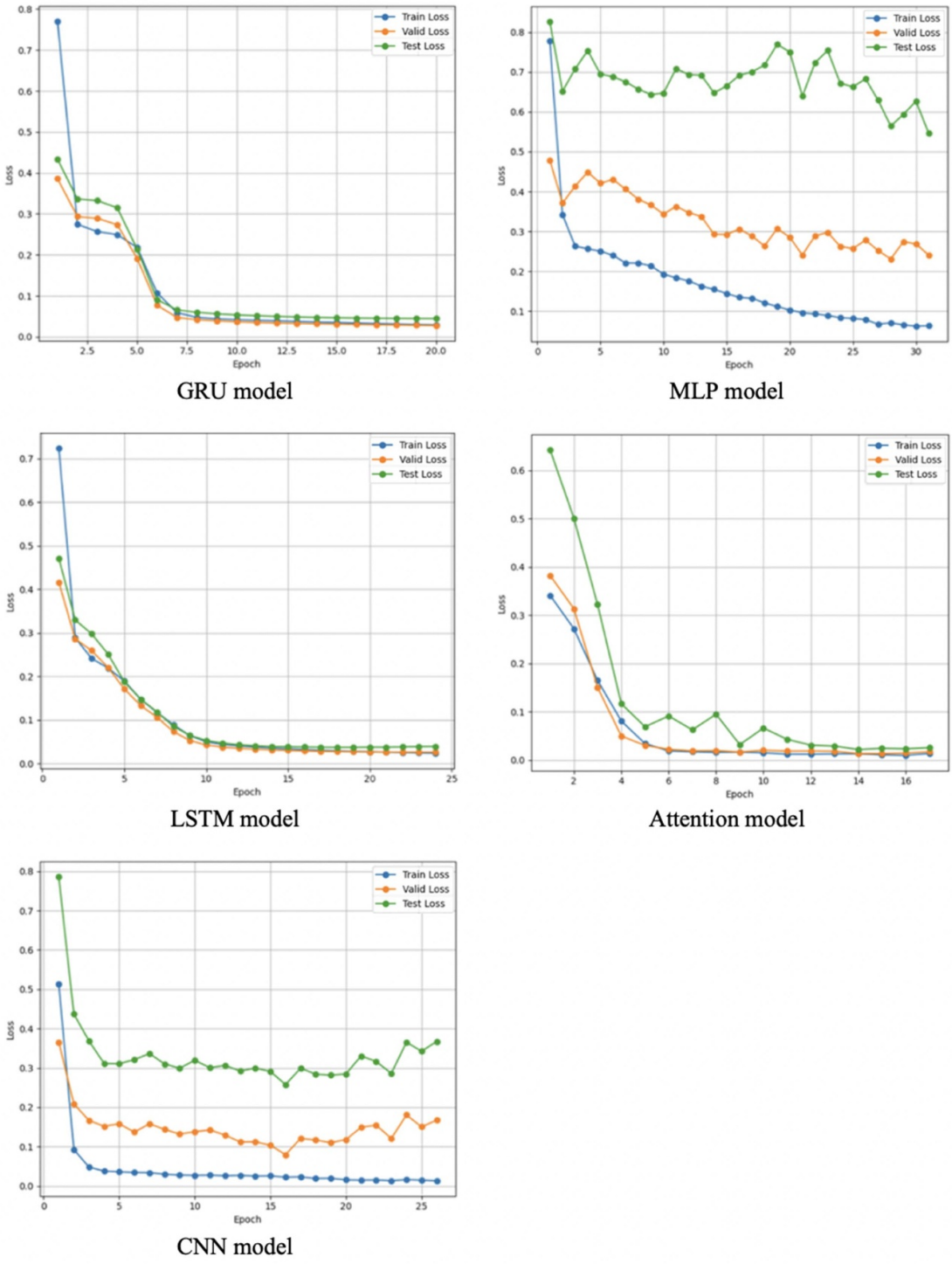


Figure 17. Training, validation, and test loss curves of different deep learning models for the CALCE CS2_38 dataset.

References

[1] Yong T, Lee C, Kim S and Kim J 2025 Battery life prediction for ensuring robust operation of IoT devices in remote metering *Appl. Sci.* **15** 2968
[2] Duan Z, Yuan Z, Jiang Y, Zhao Q, Huang Q, Zhang Y, Liu B and Tai H 2022 Power generation humidity sensor based on primary battery structure *Chem. Eng. J.* **446** 136910
[3] Yang Y, Yuchen X, Nie Y, Jianming Li, Liu S, Zhao L, Quanqing Y and Zhang C 2024 Deep transfer learning enables battery state of charge and state of health estimation *Energy* **294** 130779
[4] Chen K, Luo Y, Long Z, Wang H, Yang Li, Gao G and Guangning W 2025 Battery state of health estimation using deep transfer learning on short-term charging data *Measurement* **256** 118233
[5] Chen X, Sun T, Lai X, Zheng Y and Han X 2024 Transfer learning strategies for lithium-ion battery capacity estimation under domain shift differences *J. Energy Storage* **90** 111860
[6] Xiong Y, Tiangang Lv, Gao L, Jingtian H, Zhang Z and Liu H 2025 Transfer learning-based LRCNN for lithium battery state of health estimation with small samples *Processes* **13** 2223
[7] Xiaoming L, Yang X, Wang X, Shi Y, Wang J, Yao Y, Gao X, Xie H and Chen S 2025 Small-sample battery capacity prediction using a multi-feature transfer learning framework *Batteries* **11** 62
[8] Kang W, Wang D, Jongbloed G, Jiawen H and Chen P 2025 Robust transfer learning for battery lifetime prediction using early cycle data *IEEE Trans. Ind. Inform.* **21** 4639–48
[9] Sun Z, Weilin H, Wang J and Xin H 2024 State of health estimation for lithium-ion batteries with deep learning approach and direct current internal resistance *Energies* **17** 6
[10] Fan Y, Xiao F, Chaoran Li, Yang G and Tang X 2020 A novel deep learning framework for state of health estimation of lithium-ion battery *J. Energy Storage* **32** 12
[11] Schmitz M and Kowal J 2024 A deep learning approach for online state of health estimation of lithium-ion batteries using partial constant current charging curves *Batteries* **10** 206
[12] Zhang W, Pranav R S B, Wang R, Lee C, Zeng J, Cho M and Shim J 2024 Lithium-ion battery life prediction using deep transfer learning *Batteries* **10** 434
[13] Lin T, Chen S, Harris S J, Zhao T, Liu Y and Wan J 2024 Investigating explainable transfer learning for battery lifetime prediction under state transitions *Escience* **4** 100280
[14] Mondal A, Routray A and Puravankara S 2024 State-of-health estimation of li-ion batteries using semiparametric adaptive transfer learning *IEEE Trans. Transp. Electrificat.* **10** 1080–8
[15] Sahoo S, Hariharan K S, Agarwal S, Swernath S B, Bharti R, Han S and Lee S 2022 Transfer learning based generalized framework for state of health estimation of li-ion cells *Sci. Rep.* **12** 13173
[16] Zhuang Y and Jianbo Y 2022 State-of-health estimation for lithium-ion batteries using domain adversarial transfer learning *IEEE Trans. Power Electron.* **37** 3528–43
[17] Zhou K Q, Qin Y and Yuen C 2023 Transfer-learning-based state-of-health estimation for lithium-ion battery with cycle synchronization *IEEE/ASME Trans. Mechatronics* 28 692–702
[18] Tan Y and Zhao G 2020 Transfer learning with long short-term memory network for state-of-health prediction of lithium-ion batteries *IEEE Trans. Ind. Electron.* **67** 8723–31
[19] Guijun M, Songpei X, Yang T, Zhenbang D, Zhu L, Ding H and Yuan Y 2024 A transfer learning-based method for personalized state of health estimation of lithium-ion batteries *IEEE Trans. Neural Netw. Learn. Syst.* **35** 759–69
[20] Shen S, Sadoughi M, Meng Li, Wang Z and Chao H 2020 Deep convolutional neural networks with ensemble learning and transfer learning for capacity estimation of lithium-ion batteries *Appl. Energy* **260** 114296
[21] Zhuang Y, Jianbo Y and Mao L 2021 Multisource domain adaption for health degradation monitoring of lithium-ion batteries *IEEE Trans. Transp. Electr.* **7** 2279–92
[22] Zhang S, Zhu H, Jun W and Chen Z 2023 Voltage relaxation-based state-of-health estimation of lithium-ion batteries using convolutional neural networks and transfer learning *J. Energy Storage* **73** 12
[23] NASA Ames Prognostics Center of Excellence Li-ion battery aging datasets (available at: https://data.nasa.gov/dataset/li-ion-battery-aging-datasets) (Accessed 29 January 2026)
[24] Center for Advanced Life Cycle Engineering (CALCE) 2026 Battery data (available at: https://calce.umd.edu/battery-data)
[25] Hannan M A, Lipu M S H, Hussain A and Mohamed A 2017 A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: challenges and recommendations *Renew. Sustain. Energy Rev.* **78** 834–54
[26] Zine B, Bia H, Benmouna A, Becherif M and Iqbal M 2022 Experimentally validated coulomb counting method for battery State-of-Charge estimation under variable current profiles *Energies* **15** 8172
[27] Hannan M A, How D N T, Mansor M B, Lipu M S H, Ker P J and Muttaqi K M 2021 State-of-charge estimation of Li-ion battery using gated recurrent unit with one-cycle learning rate policy *IEEE Trans. Ind. Appl.* **57** 2964–71
[28] Wang F, Zhai Z, Liu B, Zheng S, Zhao Z and Chen X 2024 Open access dataset, code library and benchmarking deep learning approaches for state-of-health estimation of lithium-ion batteries *J. Energy Storage* **77** 109884