Designing electrodes and electrolytes for batteries by leveraging deep learning

[1]Chenxi Sui†, [2]Ziyang Jiang†, [1,3]Genesis Higueros†, [2]David Carlson, [1]Po-Chun Hsu*

[1] Pritzker School of Molecular Engineering, University of Chicago, Chicago, IL, USA

[2] Civil and Environmental Engineering, Duke University, Durham, NC, USA

[3] Thomas Lord Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC, USA

†These authors contribute equally

*pochunhsu@uchicago.edu

Abstract

High-performance batteries are poised for electrification of vehicles and therefore mitigate greenhouse gas emissions, which, in turn, promote a sustainable future. However, the design of optimized batteries is challenging due to the nonlinear governing physics and electrochemistry. Recent advancements have demonstrated the potential of deep learning techniques in efficiently designing batteries, particularly in optimizing electrodes and electrolytes. This review provides comprehensive concepts and principles of deep learning and its application in solving battery-related electrochemical problems, which bridges the gap between artificial intelligence and electrochemistry. We also examine the potential challenges and opportunities associated with different deep learning approaches, tailoring them to specific battery requirements. Ultimately, we aim to inspire future advancements in both fundamental scientific understanding and practical engineering in the field of battery technology. Furthermore, we highlight the potential challenges and opportunities for different deep learning methods according to the specific battery demand to inspire future advancement in fundamental science and practical engineering.

1. Introduction

To mitigate greenhouse gas emissions and air pollution, next-generation clean energy and sustainable fuels are urgent to be developed (1, 2). Over the past few decades, renewable technologies such as hydropower, wind power, solar photovoltaics, and bioenergy have gained increasing significance in the quest for green electricity. According to a report, the power generated in the Net Zero Scenario has increased from 19.8% to 28.7% between 2010 and 2021, with an ambitious goal of 60.9% in 2030 (3). Furthermore, the global electric vehicle stock has experienced significant growth, increasing from 11.3 million in 2010 to 2020, indicating a trend towards electrification for Net Zero Emission transportation (4). Additionally, industry-wide cost estimates for battery packs used in electric vehicles have decreased by approximately 14% annually between 2007 and 2014, from above US$1,000 per kWh to around US$410 per kWh, with the cost of battery packs used by market-leading BEV (Battery Electric Vehicle) manufacturers even lower at US$300 per kWh and declining by 8% annually. (5) These trends indicate a major transition towards renewable electricity, and electrochemistry is emerging as a powerful tool for energy and sustainability, with applications ranging from energy storage (6-12), carbon capture(13-16), green energy generation (17-20), and smart buildings (21-26). Among these technologies, battery plays an essential role in storing intermittent solar and wind energy for large-scale electrification (27-29).

The battery is a complex electrochemical system governed by chemical reactions, charge, and ion transport in solid and liquid phases, making it challenging to describe with simple physics formula (30).

Additionally, the electrode microstructures and electrolyte environment, including transfer number, conductivity, and viscosity, significantly impact the battery's performance in terms of charging and discharging rates, cycle life, and energy density. Although researchers have made significant progress in material synthesis through experiments and first-principle modeling, accurate and versatile prediction tools are urgently needed to explore optimal battery performance and reduce EV market prices under various scenarios, crossing large time and length scales and different application conditions.

The rapid development of deep learning (DL) algorithms has stimulated materials discovery by combining accurate first-principle simulations, autonomous synthesis, and testing experiments. (31-44) DL models can predict electronic, thermodynamic, and mechanical properties of the battery and its materials effectively, given a finite number of training datasets generated from experiments and/or simulations. Integrating DL models into traditional experiments and simulations reduces the time and cost required for discovering and characterizing new electrode/electrolyte materials and additives for high-performance batteries, such as large capacity, high cycle life, strong mechanical strength, and improved safety. This contributes to the promising trend of electrification. In this review, we first elucidate basic algorithms in DL and then review recent applications of DL in electrode and electrolyte design. We aim to bridge the gap between the battery design and artificial intelligence (AI) and inspire scientific and technological developments in materials science, computer science, and engineering to harness the battery's potential for sustainability.
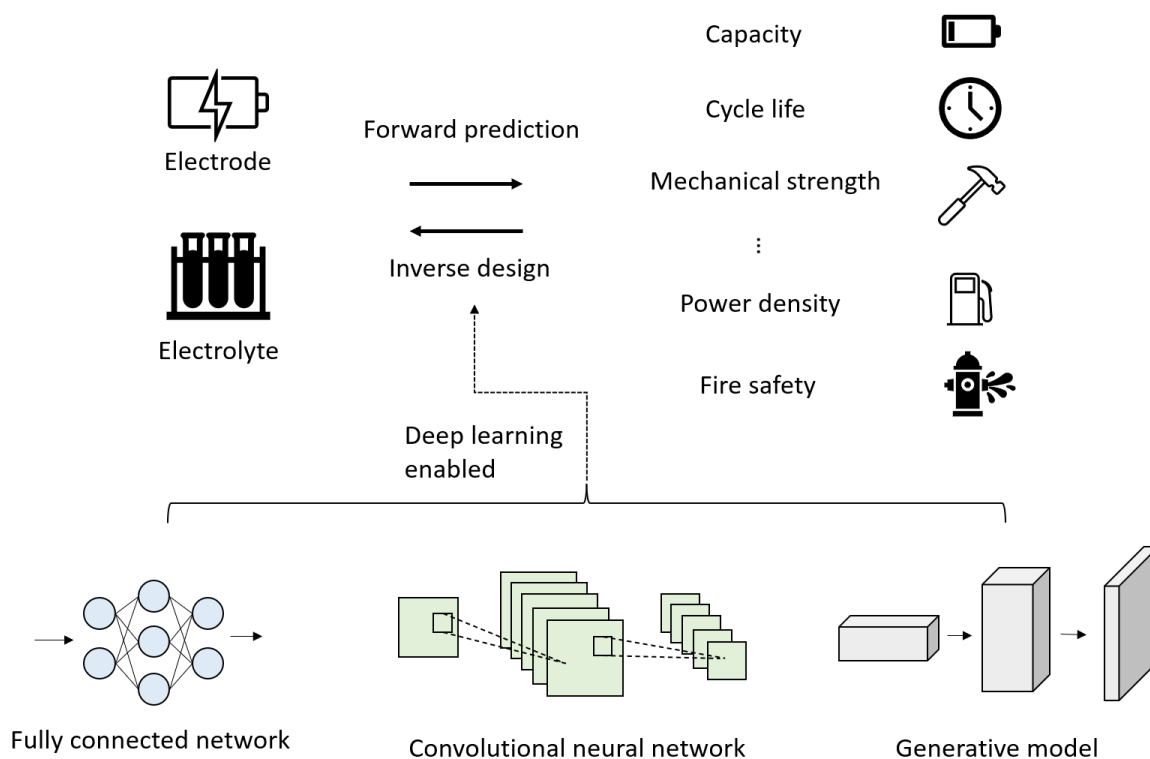

Figure 1. Overview of deep learning enabled battery design.

2. Basic Principles of Machine Learning and Deep Learning

We first present a comprehensive overview of the fundamental principles and key concepts for the commonly used machine learning (ML) or DL algorithms. Then we will elaborate on their applications in electrode and electrolyte design and compare them with traditional design methods in Sections 3 and 4.

2.1 Supervised Learning Algorithms

Supervised learning algorithms aim to fit a function which maps the input features into predictions based on a training dataset consisting of input-target pairs. In most cases, we assume the relationship between the input features and targets is linear. Say we have a training set of independent input features $X = \{x_1, x_2, \ldots, x_N\}$ and target labels $y = \{y_1, y_2, \ldots, y_N\}$. The relationship between X and y is:

$$y_i = \boldsymbol{w}^T \boldsymbol{x}_i + \epsilon \ \text{ for } \ i = 1, 2, \ldots, N,$$

where $\boldsymbol{w}$ is a vector of parameters and ϵ is an unobserved noise term. When $\boldsymbol{x}_i$ is multi-dimensional and y_i is one-dimensional, the method is referred to as Multiple Linear Regression (MLR). In a non-Bayesian paradigm, our objective is to estimate $\boldsymbol{w}$ in such a way that the error between the predicted values $\hat{y}_i$ and the target labels y_i is minimized. To make sure that the trained model generalizes to other data points which do not appear in the training set, we can add penalization terms on $\boldsymbol{w}$, leading to a Penalized Multiple Linear Regression (PMLR) model. It is worth noting that, for MLR, when we set the minimization objective to be the squared loss between the predicted values and the target labels, i.e. $\sum_{i=1}^{N}(y_i - \hat{y}_i)$, this is equivalent to maximizing the likelihood function $p(X, Y|\boldsymbol{w})$ assuming the noise ϵ follows a Gaussian distribution. For PMLR, this is equivalent to maximizing the posterior function $p(\boldsymbol{w}|X, Y) \propto p(X, Y|\boldsymbol{w})p(\boldsymbol{w})$ where L_1 penalization corresponds to assuming a Laplacian distribution on $p(\boldsymbol{w})$ and L_2 penalization corresponds to assuming a Gaussian distribution on $p(\boldsymbol{w})$.

Another supervised learning algorithm we present here is the Support Vector Machine (SVM). Unlike MLR which minimizes the prediction error, SVM aims to find the optimal hyperplane which maximizes the margin between the vectors of two classes in a binary classification problem (*45*). Specifically, fitting a SVM involves solving the following optimization problem for all $i = 1, 2, \ldots, N$:

$$\boldsymbol{w}^*, b^* = \arg\min_{\boldsymbol{w}, b} \|\boldsymbol{w}\|_2^2 \ \text{ subject to } \ y_i(\boldsymbol{w}^T \boldsymbol{x}_i - b) \geq 1,$$

where $\boldsymbol{w}$ is the normal vector to the hyperplane and b controls the offset of the hyperplane to the origin. Besides linear SVM, Guyon et al. (*46*) applied the kernel trick to SVM to construct nonlinear hyperplanes. Furthermore, Drucker et al.(*47*) proposed Support Vector Regression (SVR), extending SVM to regression problems in supervised learning.

2.2 Probabilistic Algorithms

As described in Section 2.1, learning a supervised learning model is sometimes equivalent to maximizing the likelihood function so that the observed data is most probable. However, this approach often leads to excessively complex models and overfitting. To avoid this problem, one common solution is to apply Bayesian treatment, which gives us a probabilistic model and automatically determines its complexity (*48*). Briefly speaking, with Bayesian learning, we assign a prior distribution on the model parameters and then use the observed data to update this prior. For example, to apply Bayesian treatment on the Multiple Linear Regression model in Section 2.1, instead of directly fitting the parameters $\boldsymbol{w}$, we introduce a prior distribution over $\boldsymbol{w}$ and ϵ: $p(\boldsymbol{w}|\alpha) = \mathcal{N}(0, \alpha^{-1}\boldsymbol{I}), p(\epsilon|\beta) = \mathcal{N}(0, \beta^{-1})$ where α and β are precision parameters. Using the observed data $\mathcal{D} = \{X, y\}$, we can calculate the posterior

distribution $p(\boldsymbol{w}|\mathcal{D}, \alpha, \beta) \propto p(\mathcal{D}|\boldsymbol{w}, \alpha, \beta)p(\boldsymbol{w}|\alpha)$ and make inference on a new data point $\boldsymbol{x}^*$ via the predictive posterior below:

$$p(y^*|\boldsymbol{x}^*, \mathcal{D}, \alpha, \beta) = \int p(y^*|\boldsymbol{x}^*, \mathcal{D}, \boldsymbol{w})p(\boldsymbol{w}|\mathcal{D}, \alpha, \beta)d\boldsymbol{w},$$

which not only estimates the value of y^* but also gives its uncertainty. Such Bayesian treatments can also be applied to other types of models such as Logistic Regression and Generalized Linear Models (GLM).

Another type of probabilistic model is the Gaussian Process (GP) which is frequently used for regression tasks. By definition, a GP is a collection of random variables, any finite number of which have a joint Gaussian distribution (*49*). The working principle of GP is similar to the Bayesian linear models mentioned above, where we first assume a GP prior by specifying the mean (in most cases, we assume zero mean) and covariance functions and then fit this prior to the observed data by maximizing the marginal likelihood with respect to its parameters. Some commonly used covariance functions include squared exponential (SE), Matérn, rational quadratic (RQ), etc. In practice, GP can be effectively utilized to incorporate prior information into functions or integrated with DL models to improve their predictive performance. (*50*)

Additionally, GP is often used as a surrogate model for Bayesian optimization along with an acquisition function. To give a concrete example, say we want to learn a function $f: \mathcal{X} \to \mathcal{Y}$ to the observed data, i.e. $\hat{y} = f(\boldsymbol{x})$, using a supervised learning model with hyperparameters γ. In most cases, hyperparameters are specified before training and we need to manually tune them if the model performance is unsatisfactory. Here, with Bayesian optimization, we treat γ as the input to a GP and the model performance s (e.g. classification accuracy) as its output, i.e. $s(\gamma) \sim \mathcal{GP}\big(0, k(\gamma, \gamma')\big)$ where k is the covariance function of GP. After fitting the GP, for each γ, we can sample from this GP to obtain the mean performance $\mu_s = \mu(\gamma)$ and its uncertainty $\sigma_s = \sigma(\gamma)$. We then use the acquisition function g to determine the next set of hyperparameters we want to choose. A common choice of acquisition function is the upper confidence bound (UCB) as given below (Some other popular acquisition functions include probability of improvement (PI), expected improvement (EI), etc.):

$$g(\gamma; \lambda) = \mu_s + \lambda\sigma_s,$$

where λ is an acquisition function parameter that controls the tradeoff between exploration and exploitation. Specifically, if λ is small, then $g(\gamma; \lambda)$ will be dominated by μ_s, meaning that we favor the solutions that have higher model performance. If λ is large, then $g(\gamma; \lambda)$ will be dominated by σ_s, meaning that we prefer to explore the domains where we are uncertain about the model performance. We determine the next set of hyperparameters by maximizing this acquisition function:

$$\gamma_{next} = \arg \max_{\gamma} g(\gamma; \lambda).$$

This process is repeated until we find the best set of hyperparameters γ^* which yields the best possible performance. In most cases, Bayesian optimization is much more efficient than standard hyperparameter tuning methods such as grid search.

2.3 Deep Learning Algorithms

An obvious limitation of general machine learning algorithms is that they depend heavily on the representations of the given data and usually fail to give useful predictions when the input x is intricate, such as image, audios, videos, etc.(*51*) . To address this issue, instead of directly learning a mapping from the input space to the output space, it is necessary to learn the representation itself, and this is precisely where DL demonstrates its superiority. Over the past few decades, DL models have been extensively used in the field of computer vision (*52-54*), natural language processing (*55-57*), causal inference (*58-61*), healthcare (*62-64*), and environmental sciences (*65-67*).

2.3.1 Multilayer Perceptrons

Multilayer perceptrons (MLPs), or feedforward neural networks, are one of the most commonly used DL models. It aims to learn a nonlinear functional mapping $y = f_\theta(x)$ which maps an input vector x to an output representation y. Here θ represents the parameters of the neural network. A MLP with L layers are displayed in Figure 2 below, where the network consists of an input layer, an output layer, and $L - 2$ hidden layers.
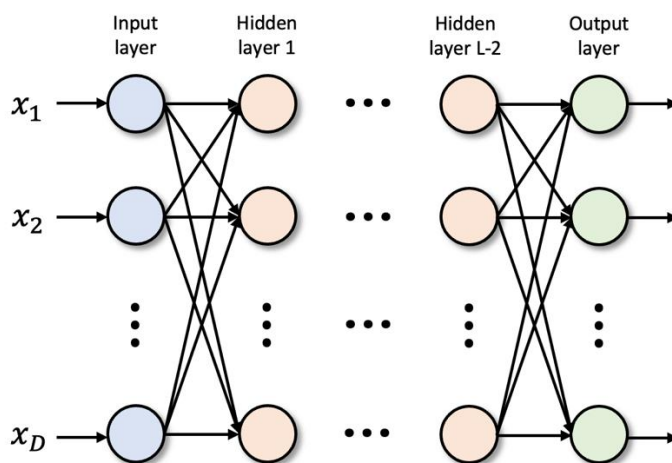

Figure 2. Architecture of a multilayer perceptron

As shown in Figure 2, the input layer has the same number of units as the dimension of the input vector (i.e. D-dimensional in this case). It takes in the input x and passes it to the first hidden layer. Each two consecutive hidden layers are connected by a set of weights and biases along with an activation function. Specifically, let us denote the ith pre-activation and post-activation unit in the lth layer as z_i^l and x_i^l, respectively. Also, we denote the weights and biases in the lth layer as W_{ij}^l and b_i^l, respectively. We can compute z_i^l recursively using the following equations:

$$z_i^l = b_i^l + \sum_{j=1}^{N_l} W_{ij}^l x_j^l, \quad x_j^l = \phi(z_j^{l-1})$$

where N_l is the number of units in the lth layer and ϕ is a nonlinear activation function. As can be seen from the equation above, W_{ij}^l and b_i^l define an affine transformation between the $(l-1)^{th}$ and the lth layers. In the absence of ϕ, the whole MLP is just a sequence of affine transformations which is not capable of capturing nonlinear functional relationships, underscoring the necessity of nonlinearity. Some frequently used nonlinear activations include the Rectified linear units (ReLU), leaky ReLU, hyperbolic

tangent (tanh), etc. The selection of the activation function for the output layer is typically dependent on the specific task at hand. For instance, the identity function may be used for regression, the sigmoid function for binary classification, and the softmax function for multi-class classification.

The training process of neural networks is similar to that of the supervised machine learning models as we have discussed in Section 2.1. For example, we can learn the neural network parameters θ to maximize the likelihood function $p(X, Y|\theta)$, which is equivalent to minimizing a squared error for regression or a cross-entropy error for classification. The key distinction between training neural networks and machine learning models is that the loss surface of neural networks is nonconvex due to the presence of nonlinearity, so convergence to the global minimum is not guaranteed. Gradient-based methods, including mini-batch gradient descent and stochastic gradient descent, are often used to optimize the neural network parameters with respect to the loss function by back-propagating the gradients (*68*).

2.3.2   Convolutional neural networks

Convolutional neural networks (CNNs) (*69, 70*) are a well-established class of DL models, widely recognized for their ability to effectively process complex input data with multiple channels (e.g. audios, images, videos, etc.). Standard CNNs usually include two types of operations: convolution and pooling. An example of 2-D convolution layer is shown in Figure 3 below.
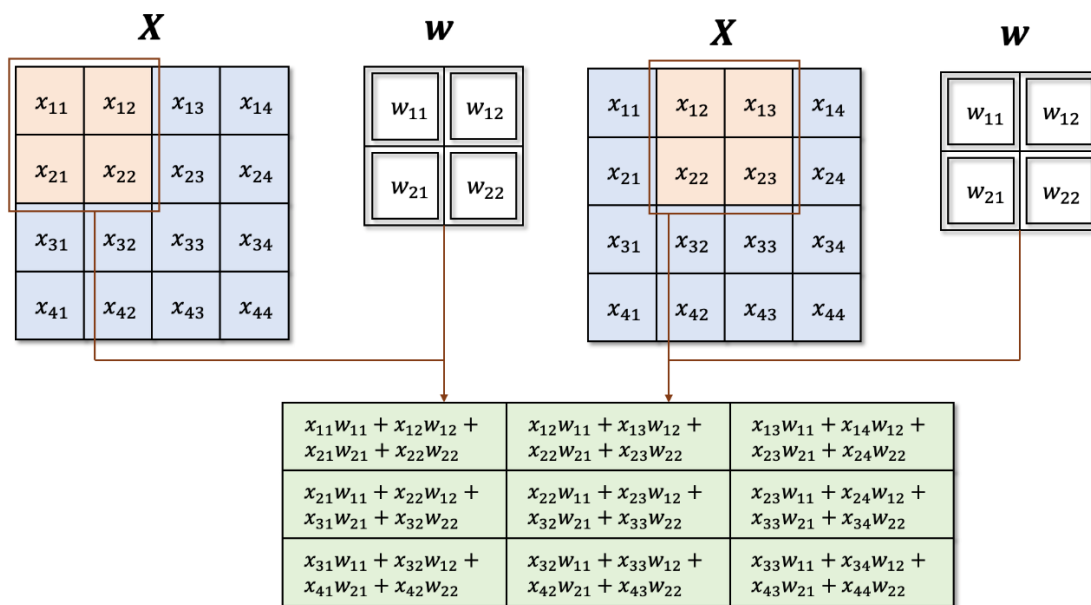

Figure 3. An example of 2-D convolution layer with a kernel size of 2x2

In the figure depicted above, the input and kernel (or filter) of the convolutional layer are represented by X and w, respectively. Notably, the size of the kernel is typically smaller than that of the input, resulting in partial connectivity between each output unit and the input units. This design choice significantly reduces the computational cost and enables the CNN to extract meaningful features from small sub-regions of the input.

The pooling operation, on the other hand, applies a summary statistic (e.g. maximum, average, etc.) to each sub-region of the input by sliding a kernel (or filter) with a certain stride size, resulting in an output

with reduced dimensions compared to the input. Pooling introduces local invariance properties to the output, including translational and rotational invariance. Translational invariance is achieved because the pooling operation is applied independently to each sub-region, allowing it to detect the same features regardless of their position within the sub-region. Rotational invariance is a consequence of the pooling operation being insensitive to the orientation of the features within the sub-region. These local invariance properties make the output more robust to small changes in the input, improving the generalization performance of the CNN.

2.3.3   Recurrent neural networks

Recurrent neural networks (RNNs) (*68, 71*) are a family of DL models for processing sequential data such as time series and text streams. The hidden units in RNNs are usually referred to as states (denoted as variable h). The computation graph, both before and after unfolding, of a fully recurrent neural network (FRNN) is illustrated in Figure 4, where each output unit $o^{(t)}$ at time step t is connected to each input unit $x^{(t)}$. This connectivity enables the network to retain and utilize information from all previous time steps, allowing it to capture temporal dependencies and patterns in the input data. The training procedure for RNNs is relatively straightforward, as the gradients are propagated backward along the unfolded computation graph. This algorithm is commonly known as backpropagation through time (BPTT).
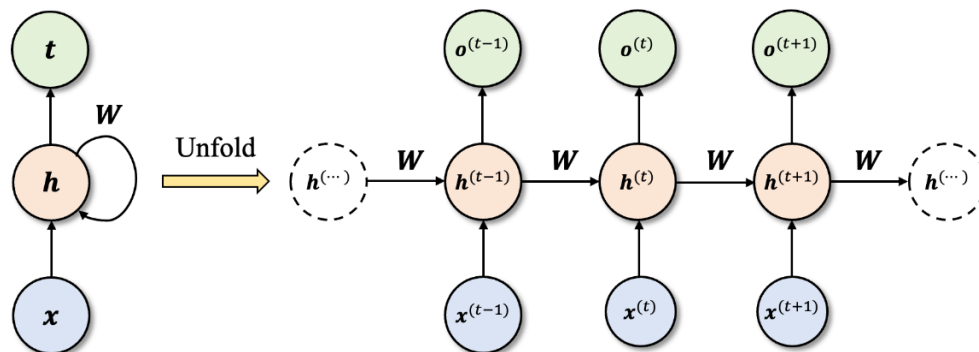

Figure 4. Computation graph of a fully recurrent neural network

One of the significant challenges associated with RNNs is their inability to handle long-term dependencies when the unfolded computational graph becomes extremely deep. Specifically, when the forward pass of an RNN involves repeatedly multiplying a weight matrix W to the state h at each time step, as depicted in Figure 4, the gradient may either vanish or explode, rendering the RNN difficult to train. The vanishing gradient problem occurs when the gradient becomes extremely small, making it difficult for the network to learn from past information. On the other hand, the exploding gradient problem occurs when the gradient becomes exceedingly large, making the network learning unstable. These issues are particularly problematic in deep architectures and can severely limit the performance of RNNs.

Several techniques have been proposed to address the vanishing and exploding gradient problem in RNNs. For instance, gradient clipping can prevent the gradient from exploding, while parameter regularization can help with the vanishing gradient issue. One of the most notable solutions is the introduction of the long short-term memory (LSTM) model by Hochreiter and Schmidhuber (*72*). The

LSTM model allows the RNN to learn whether to remember or forget relevant information using gated memory cells, making it particularly well-suited for processing long sequences of data. Another variant of the LSTM model is the gated recurrent unit (GRU), which has a simpler structure but performs comparably well to LSTM (*73, 74*).

Despite these advances, recent developments in machine learning have shown that attention mechanisms (*75*) and transformer architectures (*57*) outperform RNNs in many cases. Attention mechanisms allow for modeling dependencies between different parts of a sequence without regard to their distance in the input or output sequences. The transformer architecture, which is based solely on attention mechanisms and does not use recurrent connections, has achieved state-of-the-art results in many natural language processing and computer vision tasks. Therefore, while RNNs continue to be a valuable tool for machine learning applications, attention-based models have emerged as a promising alternative for modeling sequential data.

2.3.4   Deep generative models

Unlike regression-based models or conventional neural networks that directly estimate the conditional probability of the target variable given the input, i.e., $p(y|\boldsymbol{x})$, generative models estimate the joint distribution $p(\boldsymbol{x}, y)$ on both the input and the target variable. One of the most widely used deep generative models is the autoencoder (AE) which is a neural network comprising an encoder function $\boldsymbol{z} = f(\boldsymbol{x})$ that maps the input $\boldsymbol{x}$ into a latent feature $\boldsymbol{z}$ and a decoder function $\hat{\boldsymbol{x}} = g(\boldsymbol{z}) = g(f(\boldsymbol{x}))$ that reconstructs the input $\boldsymbol{x}$. In practice, AEs have been successfully applied to many tasks including dimensionality reduction, representation learning, and information retrieval.

An alternative perspective is to regard the latent feature $\boldsymbol{z}$ as an unobserved continuous random variable with a prior distribution $p(\boldsymbol{z})$. The observed variable $\boldsymbol{x}$ is then generated from a conditional distribution $p(\boldsymbol{x}|\boldsymbol{z})$. However, computing the marginal distribution $p(\boldsymbol{x}) = \int p(\boldsymbol{x}|\boldsymbol{z})p(\boldsymbol{z})d\boldsymbol{z}$ can be infeasible, which also makes the posterior of the latent variable $p(\boldsymbol{z}|\boldsymbol{x}) = p(\boldsymbol{x}|\boldsymbol{z})p(\boldsymbol{z})/p(\boldsymbol{x})$ computationally intractable. Kingma and Welling (*76*) introduced a variational Bayes approach to optimize an approximation of the posterior, $q(\boldsymbol{z}|\boldsymbol{x}) \approx p(\boldsymbol{z}|\boldsymbol{x})$, resulting in the variational autoencoder (VAE) framework.

Another category of deep generative model is the Generative Adversarial Network (GAN) (*77*) which consists of two models: a generator $G(\boldsymbol{z}; \boldsymbol{\theta}_g)$, which maps an input noise variable $\boldsymbol{z}$ into the data space, and a discriminator $D(\boldsymbol{x}; \boldsymbol{\theta}_d)$, which evaluates whether an input $\boldsymbol{x}$ comes from real data or from the generator by outputting a probability $0 \leq D(\boldsymbol{x}) \leq 1$. Here $\boldsymbol{\theta}_g$ and $\boldsymbol{\theta}_d$ represents the parameters of the generator and the discriminator, respectively. The generator and discriminator models are trained concurrently in an adversarial game until the generator produces samples that cannot be distinguished from real data by the discriminator. Specifically, D and G are optimized by playing the following minimax game with value function $V(G, D)$:

$$\min_G \max_D V(D, G) = \mathbb{E}_{\boldsymbol{x} \sim p_{data}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z} \sim p(\boldsymbol{z})}\left[\log\left(1 - D(G(\boldsymbol{z}))\right)\right]$$

Given arbitrary functions G and D, it is theoretically established that a unique solution exists, where G completely recovers the original data distribution, and $D(\boldsymbol{x})$ is universally equal to 1/2. The applications of GAN are diverse, including image synthesis, style transfer, image-to-image translation, and text-to-image generation, etc.

3. Deep learning-assisted electrode design

Developing fast-charging batteries is critical for the electrification of transportation to mitigate carbon emissions and climate change. The U.S. Department of Energy has set an ambitious goal of enabling electric vehicles to recharge under 10 minutes, covering a distance of 200 miles.(*28, 78*) This objective presents significant challenges and opportunities for scientists and engineers to overcome conventional obstacles in different aspects and scales. The main hurdle is balancing the battery's capacity and charging rate. Most batteries struggle to maintain high capacity when charged quickly due to the mass transport limitation and the ineffective utilization of deep electrode materials. Addressing this issue requires rational designs of the electrode microstructures and materials. However, with increasing resolutions of electrode design and characterization, more complex datasets are generated, leading to a need for advanced analysis techniques to extract detailed insights about samples. As a result, DL models become viable tools to assist in the analysis and deduction of complex datasets and parameters space and bridge the gap between experimental data and multiphysics modeling.

3.1 Governing physics and traditional methods

Within battery systems, the solid-state matrices (or active materials) play a crucial role in facilitating the transport of electrons, and the electrolyte is responsible for enabling the flow of ions between the positive and negative electrodes. However, the tortuous path of the porous electrode, due to the randomly packed particles, can hinder the ionic transport, resulting in elevated resistance and a reduction in battery performance. (Figure 5a) Thus, the degree of turning of the ionic path in a porous electrode is defined as its tortuosity (τ). This parameter is crucial as it directly impacts capacity under high current density. The tortuosity can be mathematically represented as:

$$\tau = \varepsilon \frac{D_0}{D_{eff}}$$

where the ε is the overall porosity of the porous electrode, D_0 is the diffusivity of electrolytes and D_{eff} is the effective diffusivity of the ions through the whole porous electrode. In general, an elevated level of tortuosity results in increased ion transport resistance, thereby negatively affecting the performance of the battery. Conversely, a lower degree of tortuosity promotes faster ion transport, which enhances the battery's overall performance and enhances its efficiency.

In addition to tortuosity, the reaction rate distribution is another essential factor governing electrochemical kinetics. The nonlinearity of the Nernst-Planck equation causes a significant gradient in ionic concentration and reaction rate along the depth of the electrode. These gradients lead to poor utilization of deep (far from the separator) electrode materials when attempting to improve energy and power density via thick electrodes. (Figure 5b) This non-uniform distribution is highly dependent on the electrode's porous structures. To elucidate this phenomenon, we introduce a physical term called "reaction penetration depth" based on John Newman's theory for a one-dimensional (1D) porous electrode. (*79*) This theory simplifies the three-dimensional (3D) partial differential equation groups to a 1D form. If we solve equations by inserting boundary conditions and physical assumptions, we can get a relationship between the reaction penetration depth (PD) and the porosity of the electrode:

$$PD = a_1 \sqrt{\frac{1}{1 - \varepsilon}}$$

where a_1 is the constant related to materials properties (i.e., electronic, and ionic conductivity and size of active materials particles) and ε is the porosity of the electrode. The equation indicates that increasing the electrode porosity can facilitate deeper reaction penetration, resulting in improved specific capacity during high C-rate charging. Figure 5c demonstrates the improvement of penetration depth through an increase in average porosity. Varying a_1, or the material properties, implies that this penetration depth relationship applies to different electrode materials and electrolytes. Hence, it is crucial to develop an optimal electrode structure and select suitable materials to balance out the gradient distribution of the reaction rate and alleviate the effects of tortuosity. This is necessary to enhance the performance of the battery during fast-charging.
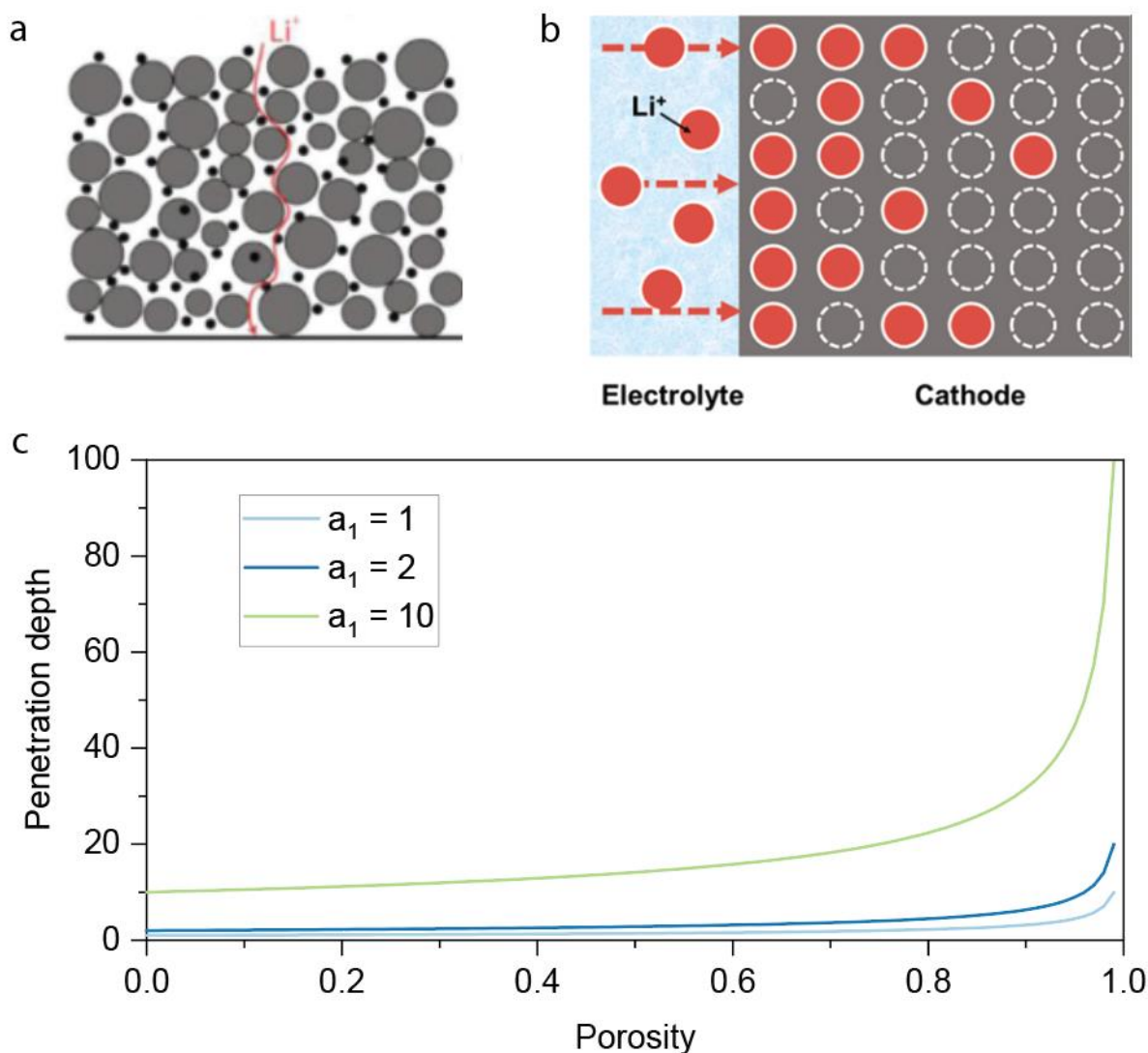

Figure 5. Key challenges of ionic transport in porous electrodes under fast-charging. (a) High tortuosity in the traditional porous electrode.(80) Copyright 2019, Wiley (b) Non-uniform ions distribution in the porous electrode. (81) Copyright 2022, ACS (c) Penetration depth as the function of porosity.

There are many previous progresses to improve the battery fast-charging performance by designing the microstructures and materials. Ramadesigan et al. first tried to solve the gradient issue theoretically by

optimizing the local porosity of the electrode along the depth.(*82, 83*) It has been observed that gradual porosity across the electrode in an optimal manner for a specific amount of active material can result in a 15%-33% reduction in the ohmic resistance. In lithium-ion battery design, single objective optimization such as reducing overall electrode resistance through a graded design has a modest effect of 4-6%. Multiple objective optimizations considering resistance and overpotential variance and average, allows for a more varied design space to achieve multiple goals. Then, inspired by theoretical progress, Liu et al. developed a bilayer electrode with gradual porosity. (*84*) The research showed that gradual porosity can decrease capacity fade by about 8.285% in full cell and 5.29% in half-cell. The increase in porosity enhances the conductivity and diffusivity of lithium-ions in the electrode and allows for control of solid electrolyte interphase (SEI) formation. Additionally, Zhao et al. introduced a gradient electrode design with vertically aligned porous channels that have smaller openings on one end and larger openings on the other. (*85*) It is found that faster kinetics occur in larger openings with more concentrated active material near the separator. Similarly, Huang et al. employed the ice templating technique to produce thick cathodes (with a thickness of 900 µm) based on LiFePO$_4$. The cathodes were designed to have a pore structure gradient and fast ion transport pathways, which enable high energy densities at fast rates.(*86*) Furthermore, Kim et al. also created stratified electrodes with Li[Ni$_{0.6}$Co$_{0.2}$Mn$_{0.2}$]O$_2$, which improved the cycle life.(*87*)

Low-tortuosity or graded porosity electrodes can also be fabricated by sacrificial template method. (Figure 6) The designed pore structures form after the templates in the slurry are removed. Bae et al. used co-extrusion to produce templates and thus macro-pore channels, with controlled channel spacings. The resulting low-tortuosity electrodes have tunable channel spacing down to ~15 µm and showed 3 times areal capacity under 2C charging rate. (*88*) Sander et al. utilized magnetic field to align the sacrificial templates to create low-tortuous electrodes. The magnetic field can not only align the magnetized nylon rod, but also the magnetic emulsion droplets.(*89*) Billaud applied a similar method and showed the enhanced ionic transport in the battery.(*90*) Zhang et al. also combined the magnetic alignment and ice-template method for low-tortuous electrode fabrication. Under 10 mA/cm^2, the electrodes exhibit an areal capacity of approximately 3.6 mAh/cm^2. (*91*) Although physics-guided research has made significant progress, there is still a gap between the optimal structure, particularly in varying application scenarios.
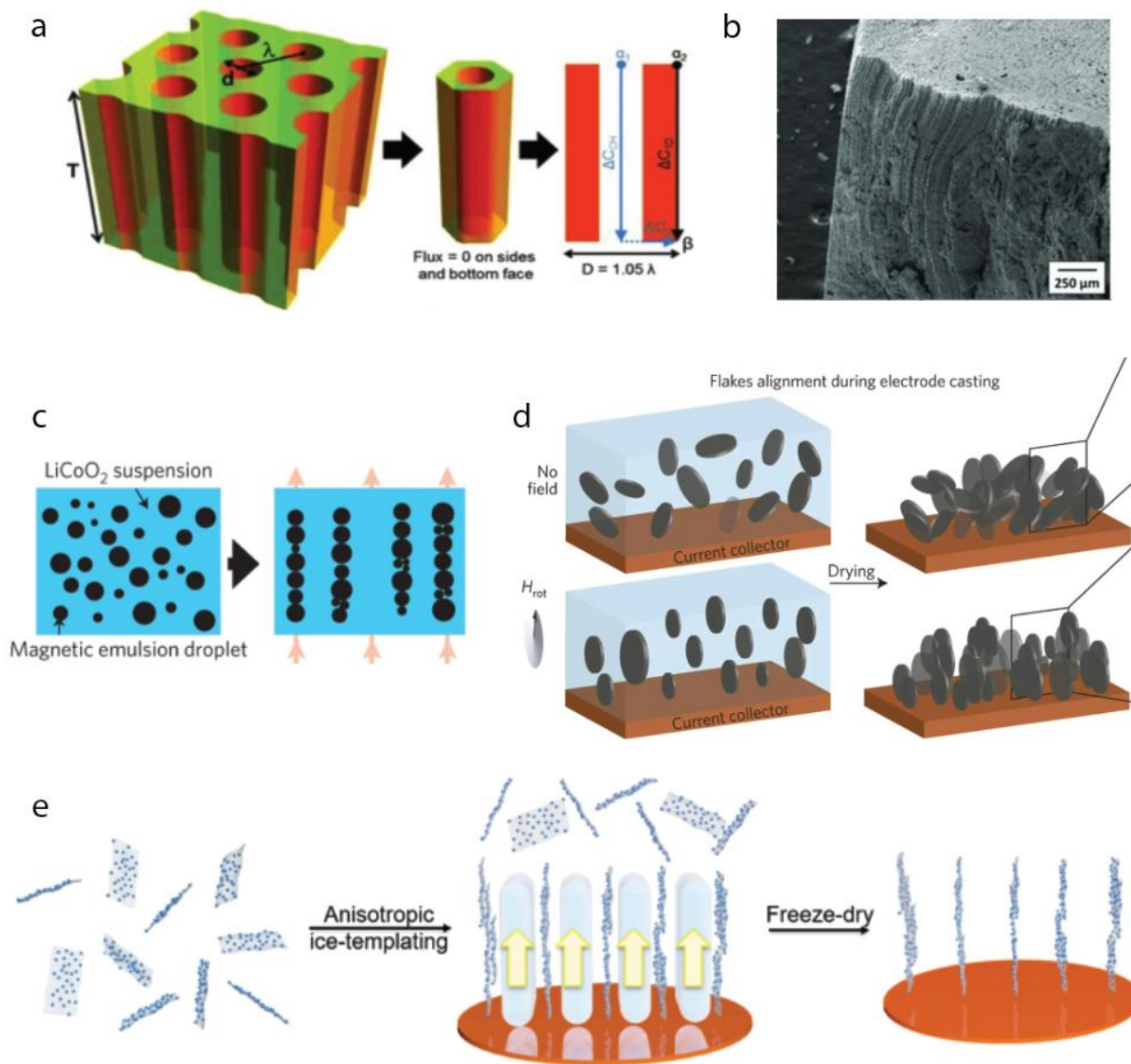
Figure 6. Low-tortuous electrodes fabrication. (a) Low-tortuous channels in the electrodes and geometric control factors.(88) Copyright 2013, Wiley (b) SEM image of the fabricated electrode. (88) Copyright 2013, Wiley (c, d, e) creating the low-tortuous channels by magnetic field alignment method and ice-template method. (89-91) Copyright 2016, Springer Nature and 2019, ACS

3.2 Deep learning method and results

It is apparent that the future of electrode design is heavily linked to the development of DL, given the challenges of identifying optimal structures for diverse application scenarios in complex parameter spaces. Employing DL across the materials synthesis, structure design, and characterization can significantly enhance design efficiency and accuracy. In this context, we will discuss recent progress in DL-assisted electrode optimization and analysis.

As mentioned earlier, one of the significant challenges in battery design is the slow ionic transport in porous electrodes. While progress has been made in designing low-tortuosity batteries through vertical

channels and gradient active materials,(*88-91*) optimal solutions have yet to be found. Although topological optimization has been applied to search for suitable electrode structures, time-dependent charging problems and the time-consuming nature of solving 3D finite element models remain obstacles. (*92*) To address these issues, Sui et al. drew inspiration from nature and introduced efficient vascular channels into the porous electrode to enhance ionic transport and mitigate non-uniform reaction rate distribution. (Figure 7) (*78*) The vertical central channels and gradient branch channels reduced tortuosity and increased reaction penetration depth. Given the large number of possible vascular channel designs, the researchers developed a DL pipeline to accelerate computation speed. The neural network, informed by the geometric factors of the electrode, speeded up computation for all possible electrode designs and corresponding charging curves by 84 times, compared to the conventional finite element method. Furthermore, the researchers developed an inverse design workflow to find the optimal electrode structure from the total library of computed designs under different application scenarios, such as varying charging rates. The workflow delivered a customized package containing electrode geometric factors, charging curves, charging capacity, energy density, etc. With the aid of bio-inspired vascularized design and DL-based prediction, the optimized porous electrode demonstrated a 66% improvement in capacity under 3.2C constant current charging. This work could potentially inspire advancements in the experimental and theoretical aspects of fast-charging batteries in the future.
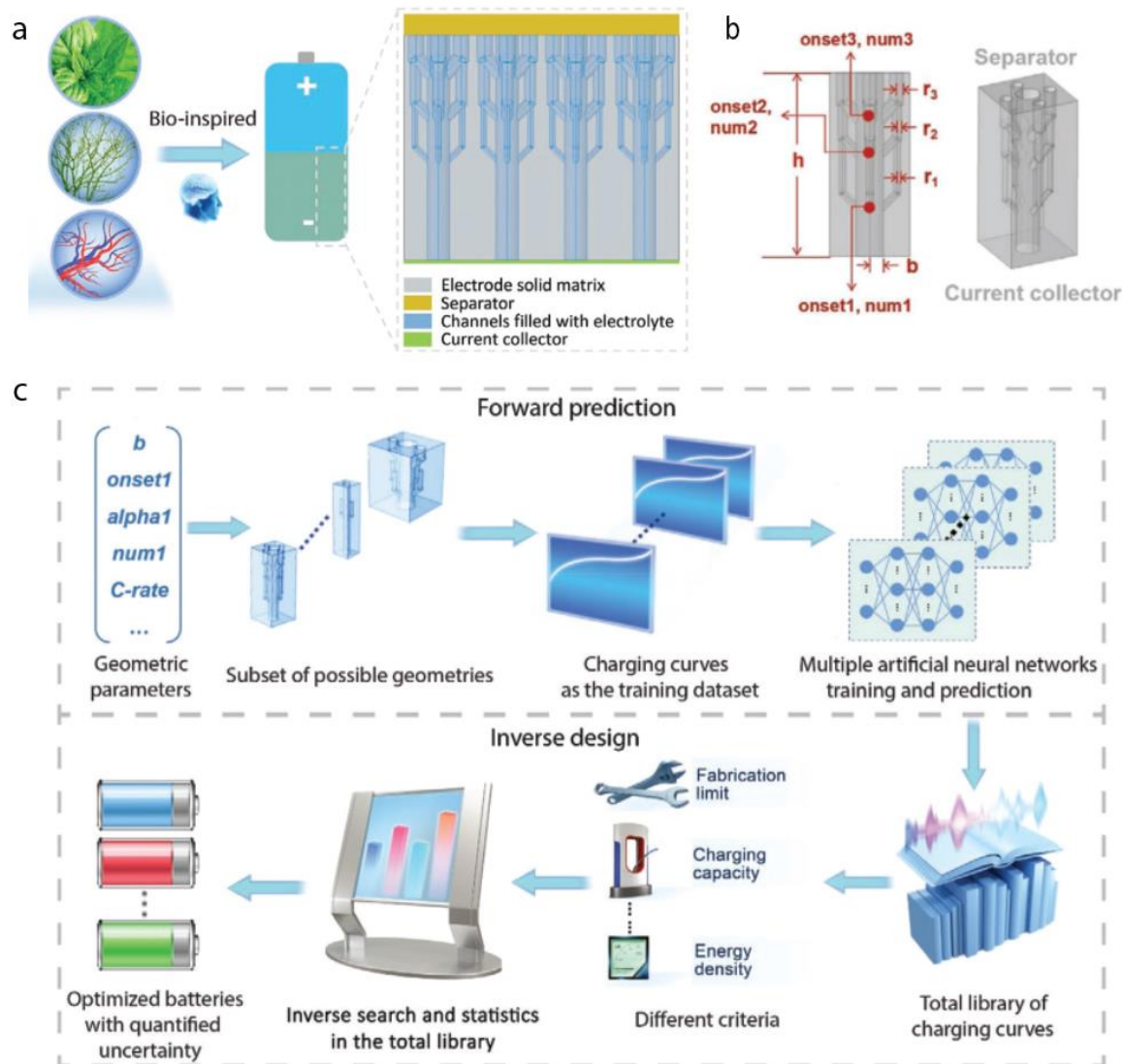
Figure 7. Bio-inspired vascularized electrode designed by deep learning. (a) Bio-inspired concept. (b) Geometric factors of vascular electrodes. (c) Deep learning optimization loop. (78) Copyright 2022, Wiley

Besides, GANs and CNNs can also be used to design the microstructure in the electrodes. Niu et al. proposed a performance-informed learning framework, called π learning, to generate electrode microstructures informed by electrochemical performance. (93) This is achieved by integrating generative adversarial neural networks and deep neural networks with physical knowledge to accurately predict current density. The framework was demonstrated in two design philosophies, inverse and forward design, for solid oxide fuel cell (SOFC) anodes. The results showed that π learning can generate electrode microstructures with the globally optimal electrode microstructure. The physical and electrochemical insights obtained by prediction can guide the rational design of SOFC electrodes. The framework can be easily transferred to the design of other porous electrodes in various electrochemical devices, such as fuel cells and batteries. The proposed π learning can be further enhanced by involving high-dimensional multi-physics computation to inform the generative model.

In addition to electrode design, characterization is a crucial tool for understanding physical mechanisms and evolution to advance the design of batteries. The performance of batteries depends on the transport of lithium ions and the kinetic process of electrodes, which are influenced by microstructure parameters such as porosity and tortuosity. These parameters play a vital role in predicting the fast-charging performance. Visualizing electrode microstructure during the service process is imperative for optimizing electrode structure and diagnosing potential safety issues. Fortunately, significant progress has been made in the field of picture processing and complex structure analysis, providing a robust framework for electrode microstructure visualization and analysis. (*94-97*) Yang et al. studied the evolution of electrode microstructure in a battery by using a modified U-Net convolutional neural network for high-precision segmentation. (*98*) DL is used to obtain the porosity and thickness of the negative and positive electrodes at different states of charge, and the relationship between the evolution of porosity and thickness during charging. The method could be extended to measure additional microstructural parameters using broad ion beam-SEM/FIB-SEM for 3D models in the future, providing an approach to exploring microstructure evolution and aiding in electrode structure optimization. Additionally, Gayon-Lombardo presented a method for creating synthetic three-dimensional microstructures consisting of several material phases using DC-GANs. This approach enables the model to represent the statistical and morphological characteristics of actual microstructures. The study used two open-source microstructural datasets, and various microstructural properties were calculated for the real data and compared to the synthetic structures created by the trained generator. The results showed excellent agreement, although the synthetic structures had a smaller variance than the training data. (*99*) Moreover, Petrich et al. developed a classifier that distinguishes between three causes for particle separation using the shape of the gap between particles. (*100*) Simulated anode material was used to generate correctly labeled sample data for developing the classifier. The classifier was then tested using hand-labeled data from a real electrode, achieving an overall accuracy of 73%.

a

Inverse Design (Specify Target Performance)

Frequency

Desired Value

Real *J* distribution

Generated distribution

Generated distribution

Specific *J*

Conditional generation

Measure distance

Label training samples

True data

Physical labels

$L_G = L_{gen,org} + L_{performance}$ **Physical loss term**

Predicted *J-V* curves

GAN Critic *C*

Phys-DNN or Phys-CNN

Microstructures

Label *J* only for training

π learning

$G(z)$

$C(G(z))$

Latent space (Gaussian noise)

Physical labels Porosity, active TPB length and current density

GAN Generator *G*

Synthetic microstructure

Active TPB
Inactive TPB
Pore Ni YSZ
Connectivity
+
Active TPB extraction

Feature engineering

Operating *J*

η_{total}

$\psi_{ele} = \eta_{total}$ $H_2 + H_2O$

$\psi_{ion} = 0$ Electrolyte (O^{2-})

Pore-scale model (OpenFOAM)

Physics-informed surrogate models

Design space z (4×16²)

Microstructure generation
Pretrained Generator *G(z)*

Physical prediction
Phys-DNN or phys-CNN
S(G(z))

If NO

If converge YES

Global optimisation algorithm
Max(*S(G(z))*)

Optimal structure with the highest performance

Forward Design (Global Optimisation)

b

Experiment

Deep learning

Tube

e- beam

Ar-ion beams

Collector

Thickness (µm)
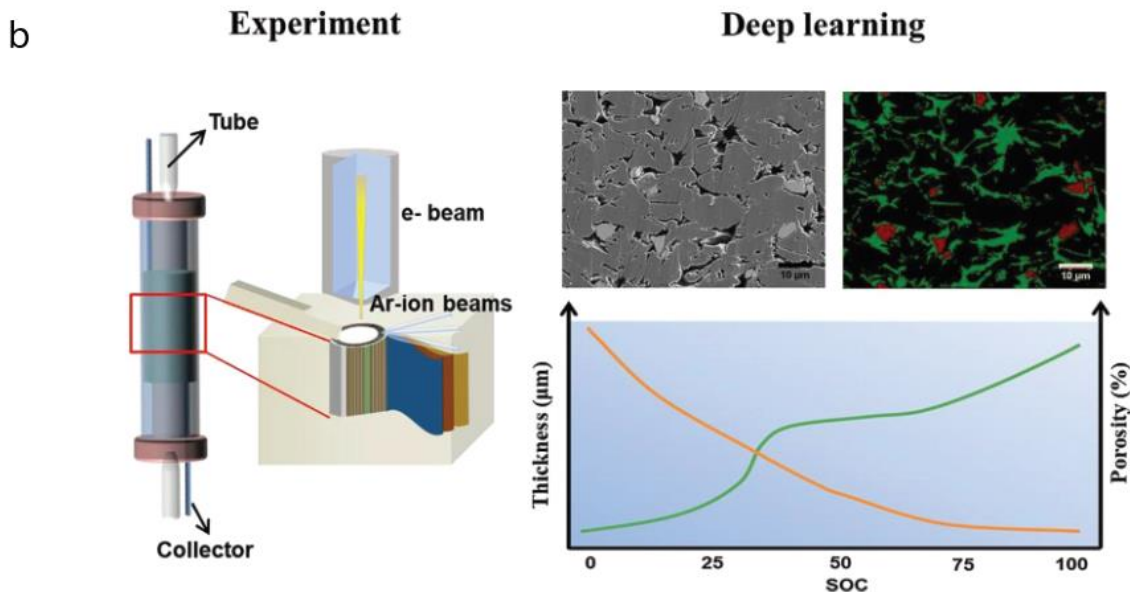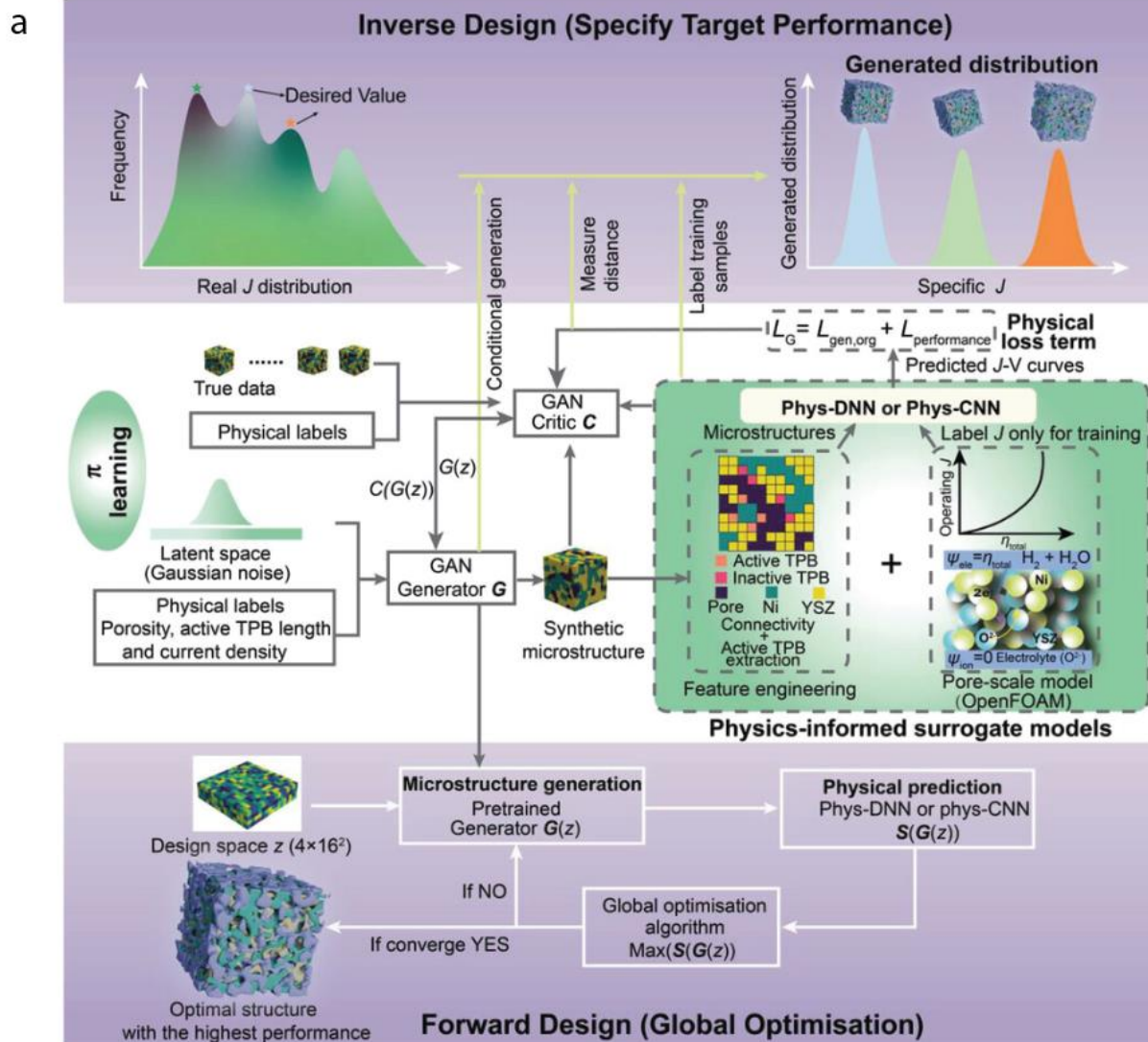
Porosity (%)

0 25 50 75 100
SOC

Figure 8. Deep learning assist microstructure of electrode design and characterization. (a) π learning for electrode microstructure design. (93) Copyright 2023, Wiley (a) Electrode microstructure characterization with the help of deep learning. (98) Copyright 2022, Elsevier

4. Deep learning-assisted electrolyte design

The electrolyte is another crucial component that significantly affects the performance of batteries, in addition to the electrode. In particular, it plays a vital role in facilitating ion transport in rechargeable batteries. However, the design of electrolytes for fast-charging batteries poses a significant challenge. Unlike solid electrodes, electrolytes are highly disordered, making it difficult to identify the governing physical factors and optimize their properties. Additionally, the vast number of possible combinations of salts and solvents makes it challenging to achieve desirable performance, because even slight variations in composition can lead to noticeable performance degradation. For fast-charging batteries, the need for rapid ionic transport throughout electrolyte to compensate for high current density exacerbates this challenge. This complex transportation problem is strongly correlated with various physical properties of electrolytes, including ionic conductivity, viscosity, diffusivity, and transfer number. Investigating such a high-dimensional parameter space is impractical using traditional trial-and-error experimentation. Furthermore, since optimal battery design is heavily dependent on application scenarios, identifying optimal electrolytes faces more obstacles.

To tackle this challenge, DL can be useful for accelerating the electrolyte design, both for simulation and experiment. Here we first show some cases for DL-accelerated numerical simulations. Among all simulations methods, molecular dynamic (MD) is the most popular tool to simulate liquid electrolytes. In the electrolyte system, ions lead to a strong local electric field and cause the polarization of the solvent molecules, especially for highly concentrated electrolytes. These polarization terms influence the electrolyte transport properties a lot and therefore need to be predicted by MD. (101) This prediction can be effectively accelerated by implementing NN to learn the atomic polarizabilities and charges (102). Such an algorithm is only dependent on connectivity of the atoms within a molecule, so the dependencies on the 3D conformation can be avoided. Apart from this, researchers also informed the neural networks with the surrounding environment of the atoms or molecules to improve the accuracy and speed of the prediction. (Figure 9) (103, 104) These methods offer a widely applicable and automated tool to comprehend atomic-level dynamics in material systems, considering the vast quantities of molecular dynamics data produced daily in almost all areas of materials design. This provides more insights of the electrolyte properties without completely screening the whole parameter space or synthesizing all possible electrolyte candidates, even in shortage of some physical interactions. To get closer to the cutting-edge battery technologies, highly concentrated electrolyte is more crucial for Zinc ion batteries because it can mitigate the hydrogen evolution reaction. Hence, equivalent methods have been utilized to examine aqueous Zinc electrolyte, which demonstrates that utilizing neural network to learn a functional physical potential is feasible even for extremely disordered systems. (105) They confirmed that the computational results accurately reproduce the observed radial distribution function and X-ray absorption near edge structure spectrum of zinc-water obtained experimentally. There were also other works that demonstrated the implementation of ML in the computational study of electrolytes. Nakayama et al. utilized an exhaustive search with a Gaussian process (ES-GP) to gauge the coordination energy. (106) This interaction energy will provide physical insights into the ionic transport of the electrolyte, which is directly related to the performance of fast-charging Li-ion batteries. Moreover, the exhaustive search can also be coupled with linear regression to become an ES-LiR model.

(*107*) By using the melting point as a target property for battery operational temperature windows, ES-LiR showed the most accurate estimation compared to multiple linear regression (MLR) and the least absolute shrinkage and selection operator (LASSO) approaches.
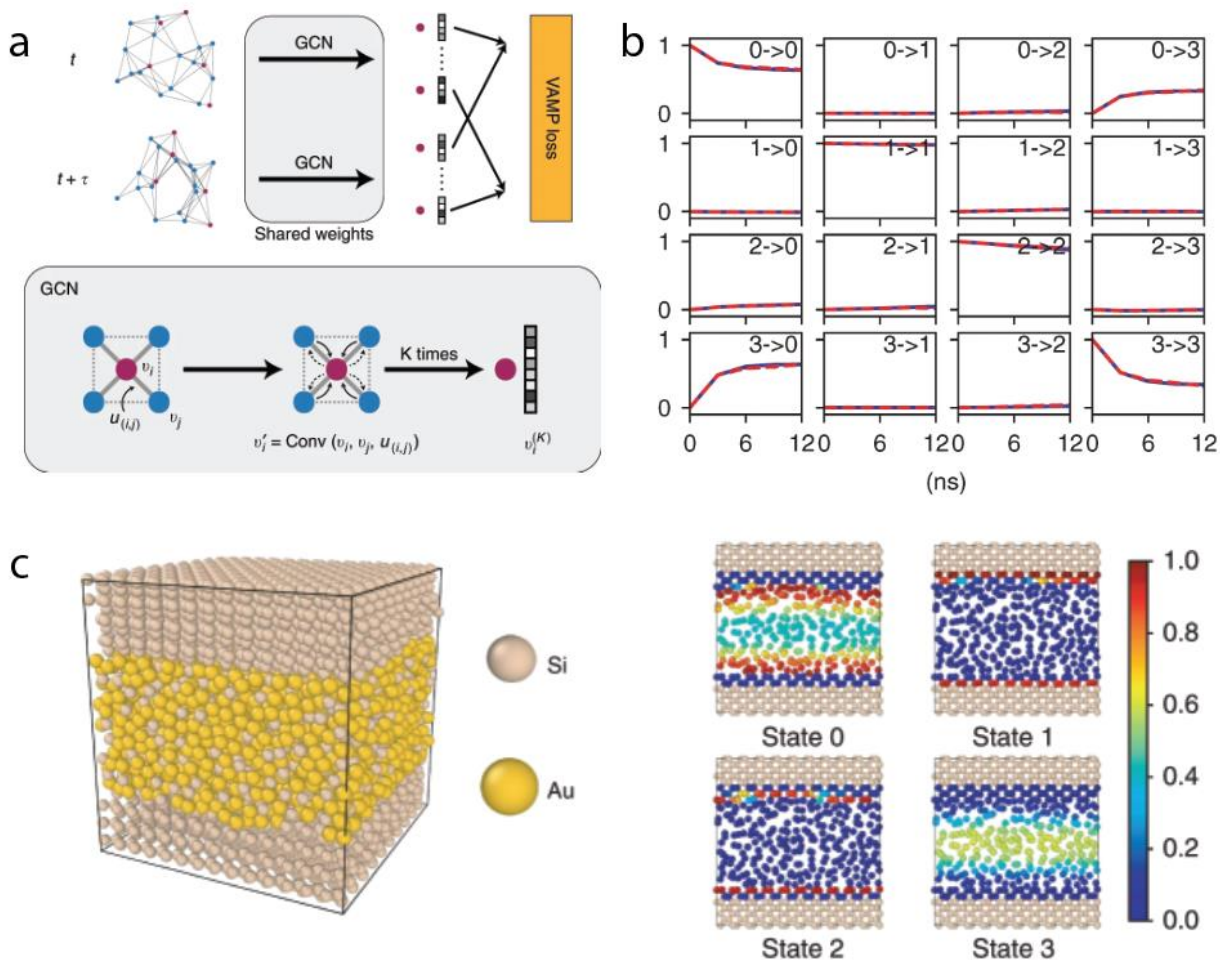

Figure 9. Graph dynamical networks for unsupervised learning of atomic scale dynamics in materials. (a) Schematics of the graph dynamical networks architecture. (b) CK test comparing the long-time dynamics predicted by Koopman models. (104) Copyright 2019, Springer

Despite the improved accuracy and efficiency of numerical simulation, empirical data plays a crucial role in the battery research domain due to the presence of experimental variations that cannot be adequately described by physical models. Considering this, researchers have begun utilizing ML techniques to facilitate the guidance of electrolyte synthesis experiments. The Cui group at Stanford University employs linear regression, random forest, and bagging models to identify key features for predicting Coulombic efficiency (CE), by utilizing the elemental composition of electrolytes as model features. They finally create fluorine-free solvent-based electrolyte formulations that achieve a remarkable CE of 99.70% according to the as-trained ML model. (*108*) The Viswanathan group presented an autonomous method for optimizing battery electrolytes using machine learning and a robot, where hundreds of sequential experiments are carried out. (Figure 10) A Bayesian optimization technique is employed to explore aqueous electrolyte salt mixtures with excellent electrochemical stability. After

conducting 140 electrolyte formula tests over a period of 40 hours, an optimal electrolyte, which was not intuitive, is finally obtained. (*109*) An impressive result is that they presented a dataset of 251 aqueous electrolytes and their conductivities, pH values, and electrochemical reactions on platinum. (*109*) In the year 2022, a comparable methodology was employed to analyze non-aqueous Li-ion batteries, resulting in a six-fold increase in time efficiency compared to an arbitrary search conducted by the identical automated experiment. (*110*) In the experimental validation test, it was observed that all the pouch cells that were filled with electrolytes developed by the robot showed improved capability of fast-charging. (*110*) Their robotic platform, real-time machine learning optimization, and integration with device testing, tailored to the specific requirement, have the potential to optimize other self-sufficient discovery platforms for energy and sustainability application.
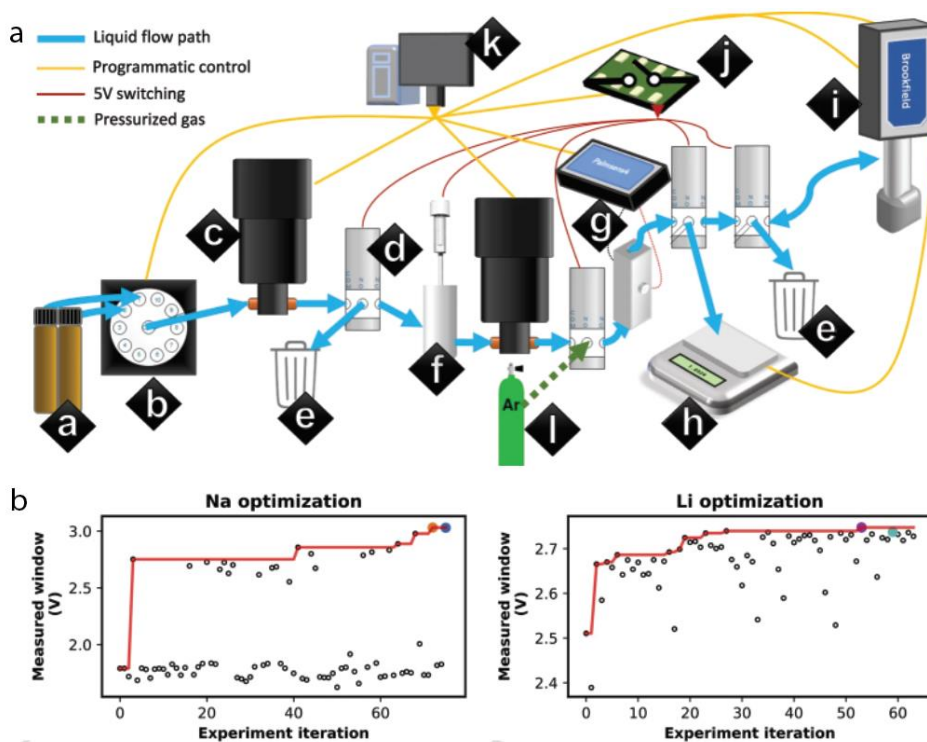

Figure 10. Autonomous optimization loop for electrolytes. (a) Schematics of automated electrolyte experiment set-up. (b) optimization routine for sodium and lithium electrolytes. (109, 110) Copyright 2020, Elsevier and 2022, Springer Nature

5.   Conclusion and outlook

In conclusion, the application of DL techniques has demonstrated considerable potential to revolutionize the field of battery design. By integrating DL algorithms with traditional experiments and simulations, researchers can accelerate the discovery and characterization of new electrodes and electrolytes, leading to the development of high-performance batteries with improved capacity and cycle life. The combination of accurate first-principle simulation, autonomous synthesis, characterizations, and DL optimization loop has emerged as a promising approach for enhancing the efficiency of battery design and reducing the cost of electric vehicles. Continued exploration of DL algorithms in battery design holds significant promise for advancing the field of materials science, computer science, and engineering to mitigate the impact of climate change.

Nevertheless, there are challenges in the application of DL to battery design. The efficacy and precision of DL models hinge heavily on the availability of high-quality and reliable data, which is often proprietary and not easily obtainable. Furthermore, these models can occasionally encounter the issue of overfitting, where they become excessively tailored to the training data, resulting in poor generalization and inaccurate predictions. This challenge is particularly pronounced in battery design, given the intricate interplay of various components, making it arduous to encompass all the pertinent variables within a single model. It is imperative to go beyond training-based machine learning and extrapolate insights from existing designs to new ones with limited data. Last but not least, there is a pressing need to expedite not only the research and development phase but also the scaling-up and manufacturing processes.

Acknowledgments

Competing interests:

The authors declare no competing financial interests.

Data and materials availability:

All data needed to support the conclusions in the paper are present in the manuscript and/or the Supplementary Materials. Additional data related to this paper may be requested from the corresponding author upon request.

References

1. H. L. van Soest, M. G. J. den Elzen, D. P. van Vuuren, Net-zero emission targets for major emitting countries consistent with the Paris Agreement. *Nature Communications* **12**, 2140 (2021).
2. United Nations Sustainable Development Goals. *https://sdgs.un.org/goals*.
3. Renewable Electricity – Analysis. *International Energy Agency*.
4. Global EV Outlook 2021 [PDF]. *International Energy Agency*.
5. B. Nykvist, M. Nilsson, Rapidly falling costs of battery packs for electric vehicles. *Nature Climate Change* **5**, 329-332 (2015).
6. W. Deng *et al.*, Quantification of reversible and irreversible lithium in practical lithium-metal batteries. *Nature Energy* **7**, 1031-1041 (2022).
7. M. S. Kim *et al.*, Revealing the Multifunctions of Li3N in the Suspension Electrolyte for Lithium Metal Batteries. *ACS Nano* **17**, 3168-3180 (2023).
8. R. Shimizu *et al.*, Unraveling the Stable Cathode Electrolyte Interface in all Solid-State Thin-Film Battery Operating at 5 V. *Advanced Energy Materials* **12**, 2201119 (2022).
9. C.-Y. Wang *et al.*, Fast charging of energy-dense lithium-ion batteries. *Nature* **611**, 485-490 (2022).
10. H. Wang *et al.*, Liquid electrolyte: The nexus of practical lithium metal batteries. *Joule* **6**, 588-616 (2022).
11. J. Xiao, F. Shi, T. Glossmann, C. Burnett, Z. Liu, From laboratory innovations to materials manufacturing for lithium-based batteries. *Nature Energy* **8**, 329-339 (2023).

12. G. Zhou, H. Chen, Y. Cui, Formulating energy density for designing practical lithium–sulfur batteries. *Nature Energy* **7**, 312-319 (2022).

13. C. Chen *et al.*, Exploration of the bio-analogous asymmetric C–C coupling mechanism in tandem CO2 electroreduction. *Nature Catalysis* **5**, 878-887 (2022).

14. C. Choi *et al.*, Highly active and stable stepped Cu surface for enhanced electrochemical CO2 reduction to C2H4. *Nature Catalysis* **3**, 804-812 (2020).

15. D. Le, T. S. Rahman, On the role of metal cations in CO2 electrocatalytic reduction. *Nature Catalysis* **5**, 977-978 (2022).

16. Y. Xie *et al.*, High carbon utilization in CO2 reduction to multi-carbon products in acidic media. *Nature Catalysis* **5**, 564-570 (2022).

17. G. Han, G. Li, Y. Sun, Electrocatalytic dual hydrogenation of organic substrates with a Faradaic efficiency approaching 200%. *Nature Catalysis* **6**, 224-233 (2023).

18. T. Lazaridis, B. M. Stühmeier, H. A. Gasteiger, H. A. El-Sayed, Capabilities and limitations of rotating disk electrodes versus membrane electrode assemblies in the investigation of electrocatalysts. *Nature Catalysis* **5**, 363-373 (2022).

19. I. T. McCrum, M. T. M. Koper, The role of adsorbed hydroxide in hydrogen evolution reaction kinetics on modified platinum. *Nature Energy* **5**, 891-899 (2020).

20. B. Y. Tang, R. P. Bisbey, K. M. Lodaya, W. L. Toh, Y. Surendranath, Reaction environment impacts charge transfer but not chemical reaction steps in hydrogen evolution catalysis. *Nature Catalysis* **6**, 339-350 (2023).

21. Radiative electrochromism for energy-efficient buildings. *Nature Sustainability* **6**, 358-359 (2023).

22. C. J. Barile *et al.*, Dynamic Windows with Neutral Color, High Contrast, and Excellent Durability Using Reversible Metal Electrodeposition. *Joule* **1**, 133-145 (2017).

23. C. J. Barile, D. J. Slotcavage, M. D. McGehee, Polymer–Nanoparticle Electrochromic Materials that Selectively Modulate Visible and Near-Infrared Light. *Chemistry of Materials* **28**, 1439-1445 (2016).

24. T. S. Hernandez *et al.*, Bistable Black Electrochromic Windows Based on the Reversible Metal Electrodeposition of Bi and Cu. *ACS Energy Letters* **3**, 104-111 (2018).

25. D. C. Madu *et al.*, Investigating Formate, Sulfate, and Halide Anions in Reversible Zinc Electrodeposition Dynamic Windows. *ACS Applied Materials & Interfaces* **14**, 47810-47821 (2022).

26. C. Sui *et al.*, Dynamic electrochromism for all-season radiative thermoregulation. *Nature Sustainability* **6**, 428-437 (2023).

27. Research for greener batteries. *Nature Sustainability* **4**, 373-373 (2021).

28. Y. Liu, Y. Zhu, Y. Cui, Challenges and opportunities towards fast-charging battery materials. *Nature Energy* **4**, 540-550 (2019).

29. L. Trahey *et al.*, Energy storage emerging: A perspective from the Joint Center for Energy Storage Research. *Proceedings of the National Academy of Sciences* **117**, 12550-12557 (2020).

30. V. Ramadesigan *et al.*, Modeling and simulation of lithium-ion batteries from a systems engineering perspective. *Journal of the electrochemical society* **159**, R31 (2012).

31. A. Agrawal, A. Choudhary, Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science. *Apl Materials* **4**, 053208 (2016).

32. M. Aykol *et al.*, Network analysis of synthesizable materials discovery. *Nature Communications* **10**, 2018 (2019).

33. P. V. Balachandran, B. Kowalski, A. Sehirlioglu, T. Lookman, Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning. *Nature Communications* **9**, 1668 (2018).

34. K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, A. Walsh, Machine learning for molecular and materials science. *Nature* **559**, 547-555 (2018).

35. J. Carrete, W. Li, N. Mingo, S. Wang, S. Curtarolo, Finding Unprecedentedly Low-Thermal-Conductivity Half-Heusler Semiconductors via High-Throughput Materials Modeling. *Physical Review X* **4**, 011019 (2014).

36. C. W. Coley *et al.*, A graph-convolutional neural network model for the prediction of chemical reactivity. *Chemical science* **10**, 370-377 (2019).

37. J. M. Granda, L. Donina, V. Dragone, D.-L. Long, L. Cronin, Controlling an organic synthesis robot with machine learning to search for new reactivity. *Nature* **559**, 377-381 (2018).

38. O. Isayev *et al.*, Universal fragment descriptors for predicting properties of inorganic crystals. *Nature Communications* **8**, 15679 (2017).

39. C. Kim, G. Pilania, R. Ramprasad, Machine Learning Assisted Predictions of Intrinsic Dielectric Breakdown Strength of ABX3 Perovskites. *The Journal of Physical Chemistry C* **120**, 14575-14580 (2016).

40. P. Raccuglia *et al.*, Machine-learning-assisted materials discovery using failed experiments. *Nature* **533**, 73-76 (2016).

41. J. Schmidt, M. R. G. Marques, S. Botti, M. A. L. Marques, Recent advances and applications of machine learning in solid-state materials science. *npj Computational Materials* **5**, 83 (2019).

42. A. Y.-T. Wang *et al.*, Machine Learning for Materials Scientists: An Introductory Guide toward Best Practices. *Chemistry of Materials* **32**, 4954-4965 (2020).

43. L. Ward *et al.*, Strategies for accelerating the adoption of materials informatics. *MRS Bulletin* **43**, 683-689 (2018).

44. D. Xue *et al.*, Accelerated search for materials with targeted properties by adaptive design. *Nature Communications* **7**, 11241 (2016).

45. C. Cortes, V. Vapnik, Support-vector networks. *Machine Learning* **20**, 273-297 (1995).

46. I. M. Guyon, B. E. Boser, V. N. Vapnik, in *NIPS*. (1992).

47. H. Drucker, C. J. Burges, L. Kaufman, A. Smola, V. Vapnik, Support vector regression machines. *Advances in neural information processing systems* **9**, (1996).

48. C. M. Bishop, N. M. Nasrabadi, *Pattern recognition and machine learning*. (Springer, 2006), vol. 4.

49. C. Rasmussen, C. Williams. (Cambridge, MA: MIT press, 2006).

50. Z. Jiang, T. Zheng, D. Carlson, Incorporating Prior Knowledge into Neural Networks through an Implicit Composite Kernel. *arXiv preprint arXiv:2205.07384*, (2022).

51. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436-444 (2015).

52. K. He *et al.*, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2022), pp. 16000-16009.

53. A. Ramesh *et al.*, in *International Conference on Machine Learning*. (PMLR, 2021), pp. 8821-8831.

54. J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, S. Ganguli, in *International Conference on Machine Learning*. (PMLR, 2015), pp. 2256-2265.

55. T. Brown *et al.*, Language models are few-shot learners. *Advances in neural information processing systems* **33**, 1877-1901 (2020).

56. J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, (2018).

57. A. Vaswani *et al.*, Attention is all you need. *Advances in neural information processing systems* **30**, (2017).

58. T. Hatt, S. Feuerriegel, in *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. (2021), pp. 680-689.

59.   Z. Jiang *et al.*, Estimating Causal Effects using a Multi-task Deep Ensemble. *arXiv preprint arXiv:2301.11351*, (2023).

60.   C. Louizos *et al.*, Causal effect inference with deep latent-variable models. *Advances in neural information processing systems* **30**, (2017).

61.   U. Shalit, F. D. Johansson, D. Sontag, in *International Conference on Machine Learning*. (PMLR, 2017), pp. 3076-3085.

62.   M. A. Ahmad, C. Eckert, A. Teredesai, in *Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics*. (2018), pp. 559-560.

63.   B. J. Erickson, P. Korfiatis, Z. Akkus, T. L. Kline, Machine learning for medical imaging. *Radiographics* **37**, 505-515 (2017).

64.   R. Miotto, F. Wang, S. Wang, X. Jiang, J. T. Dudley, Deep learning for healthcare: review, opportunities and challenges. *Briefings in bioinformatics* **19**, 1236-1246 (2018).

65.   Z. Jiang, T. Zheng, M. Bergin, D. Carlson, Improving spatial variation of ground-level PM2. 5 prediction with contrastive learning from satellite imagery. *Science of Remote Sensing* **5**, 100052 (2022).

66.   T. Zheng, M. Bergin, G. Wang, D. Carlson, Local PM2. 5 Hotspot Detector at 300 m Resolution: A Random Forest–Convolutional Neural Network Joint Model Jointly Trained on Satellite Images and Meteorology. *Remote Sensing* **13**, 1356 (2021).

67.   T. Zheng, M. H. Bergin, S. Hu, J. Miller, D. E. Carlson, Estimating ground-level PM2. 5 using micro-satellite images by a convolutional neural network and random forest approach. *Atmospheric Environment* **230**, 117451 (2020).

68.   D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning representations by back-propagating errors. *Nature* **323**, 533-536 (1986).

69.   Y. L. Cun *et al.*, Handwritten digit recognition: applications of neural network chips and automatic learning. *IEEE Communications Magazine* **27**, 41-46 (1989).

70.   Y. LeCun, Y. Bengio, Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks* **3361**, 1995 (1995).

71.   M. I. Jordan, in *Advances in psychology*. (Elsevier, 1997), vol. 121, pp. 471-495.

72.   S. Hochreiter, J. Schmidhuber, Long short-term memory. *Neural computation* **9**, 1735-1780 (1997).

73.   J. Chung, C. Gulcehre, K. Cho, Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, (2014).

74.   R. Jozefowicz, W. Zaremba, I. Sutskever, in *International conference on machine learning*. (PMLR, 2015), pp. 2342-2350.

75.   D. Bahdanau, K. Cho, Y. Bengio, Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, (2014).

76.   D. P. Kingma, M. Welling, Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, (2013).

77.   I. Goodfellow *et al.*, Generative adversarial networks. *Communications of the ACM* **63**, 139-144 (2020).

78.   C. Sui *et al.*, Bio-Inspired Computational Design of Vascularized Electrodes for High-Performance Fast-Charging Batteries Optimized by Deep Learning. *Advanced Energy Materials* **12**, 2103044 (2022).

79.   J. Newman, N. P. Balsara, *Electrochemical systems*. (John Wiley & Sons, 2021).

80.   L. Li, R. M. Erb, J. Wang, J. Wang, Y.-M. Chiang, Fabrication of Low-Tortuosity Ultrahigh-Area-Capacity Battery Electrodes through Magnetic Alignment of Emulsion-Based Slurries. *Advanced Energy Materials* **9**, 1802472 (2019).

81.   J. Wu *et al.*, Low-Tortuosity Thick Electrodes with Active Materials Gradient Design for Enhanced Energy Storage. *ACS Nano* **16**, 4805-4812 (2022).

82.   Y. Qi, T. Jang, V. Ramadesigan, D. T. Schwartz, V. R. Subramanian, Is there a benefit in employing graded electrodes for lithium-ion batteries? *Journal of The Electrochemical Society* **164**, A3196 (2017).

83.   V. Ramadesigan, R. N. Methekar, F. Latinwo, R. D. Braatz, V. R. Subramanian, Optimal porosity distribution for minimized ohmic drop across a porous electrode. *Journal of The Electrochemical Society* **157**, A1328 (2010).

84.   L. Liu, P. Guan, C. Liu, Experimental and simulation investigations of porosity graded cathodes in mitigating battery degradation of high voltage lithium-ion batteries. *Journal of The Electrochemical Society* **164**, A3163 (2017).

85.   X. Zhang *et al.*, Gradient Architecture Design in Scalable Porous Battery Electrodes. *Nano Letters* **22**, 2521-2528 (2022).

86.   C. Huang, M. Dontigny, K. Zaghib, P. S. Grant, Low-tortuosity and graded lithium ion battery cathodes by ice templating. *Journal of Materials Chemistry A* **7**, 21421-21431 (2019).

87.   H. Kim *et al.*, Failure mode of thick cathodes for Li-ion batteries: Variation of state-of-charge along the electrode thickness direction. *Electrochimica Acta* **370**, 137743 (2021).

88.   C.-J. Bae, C. K. Erdonmez, J. W. Halloran, Y.-M. Chiang, Design of Battery Electrodes with Dual-Scale Porosity to Minimize Tortuosity and Maximize Performance. *Advanced Materials* **25**, 1254-1258 (2013).

89.   J. S. Sander, R. M. Erb, L. Li, A. Gurijala, Y. M. Chiang, High-performance battery electrodes via magnetic templating. *Nature Energy* **1**, 16099 (2016).

90.   J. Billaud, F. Bouville, T. Magrini, C. Villevieille, A. R. Studart, Magnetically aligned graphite electrodes for high-rate performance Li-ion batteries. *Nature Energy* **1**, 16097 (2016).

91.   X. Zhang *et al.*, Promoting Transport Kinetics in Li-Ion Battery with Aligned Porous Electrode Architectures. *Nano Letters* **19**, 8255-8261 (2019).

92.   T. Xu, Topology optimization of lithium ion batteries: How to maximize the discharge capacity by changing the geometry. (2015).

93.   Z. Niu *et al.*, π Learning: A Performance-Informed Framework for Microstructural Electrode Design. *Advanced Energy Materials* **13**, 2300244 (2023).

94.   I. Arganda-Carreras *et al.*, Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics* **33**, 2424-2426 (2017).

95.   H.-P. Chan, R. K. Samala, L. M. Hadjiiski, C. Zhou, Deep learning in medical image analysis. *Deep Learning in Medical Image Analysis: Challenges and Applications*, 3-21 (2020).

96.   S. Kench, S. J. Cooper, Generating three-dimensional structures from a two-dimensional slice with generative adversarial network-based dimensionality expansion. *Nature Machine Intelligence* **3**, 299-305 (2021).

97.   H. Xu *et al.*, Guiding the Design of Heterogeneous Electrode Microstructures for Li-Ion Batteries: Microscopic Imaging, Predictive Modeling, and Machine Learning. *Advanced Energy Materials* **11**, 2003908 (2021).

98.   Y. Yang *et al.*, Microstructure evolution of lithium-ion battery electrodes at different states of charge: Deep learning-based segmentation. *Electrochemistry Communications* **136**, 107224 (2022).

99.   A. Gayon-Lombardo, L. Mosser, N. P. Brandon, S. J. Cooper, Pores for thought: generative adversarial networks for stochastic reconstruction of 3D multi-phase electrode microstructures with periodic boundaries. *npj Computational Materials* **6**, 82 (2020).

100.   L. Petrich *et al.*, Crack detection in lithium-ion cells using machine learning. *Computational Materials Science* **136**, 297-305 (2017).

101. D. Bedrov *et al.*, Molecular Dynamics Simulations of Ionic Liquids and Electrolytes Using Polarizable Force Fields. *Chemical Reviews* **119**, 7940-7995 (2019).

102. E. Heid, M. Fleck, P. Chatterjee, C. Schröder, A. D. MacKerell, Jr., Toward Prediction of Electrostatic Parameters for Force Fields That Explicitly Treat Electronic Polarization. *Journal of Chemical Theory and Computation* **15**, 2460-2469 (2019).

103. K. T. Schütt, H. E. Sauceda, P. J. Kindermans, A. Tkatchenko, K. R. Müller, SchNet – A deep learning architecture for molecules and materials. *The Journal of Chemical Physics* **148**, 241722 (2018).

104. T. Xie, A. France-Lanord, Y. Wang, Y. Shao-Horn, J. C. Grossman, Graph dynamical networks for unsupervised learning of atomic scale dynamics in materials. *Nature Communications* **10**, 2667 (2019).

105. M. Xu, T. Zhu, J. Z. H. Zhang, Molecular Dynamics Simulation of Zinc Ion in Water with an ab Initio Based Neural Network Potential. *The Journal of Physical Chemistry A* **123**, 6587-6595 (2019).

106. T. Nakayama, Y. Igarashi, K. Sodeyama, M. Okada, Material search for Li-ion battery electrolytes through an exhaustive search with a Gaussian process. *Chemical Physics Letters* **731**, 136622 (2019).

107. K. Sodeyama, Y. Igarashi, T. Nakayama, Y. Tateyama, M. Okada, Liquid electrolyte informatics using an exhaustive search with linear regression. *Physical Chemistry Chemical Physics* **20**, 22585-22591 (2018).

108. S. C. Kim *et al.*, Data-driven electrolyte design for lithium metal anodes. *Proceedings of the National Academy of Sciences* **120**, e2214357120 (2023).

109. A. Dave *et al.*, Autonomous Discovery of Battery Electrolytes with Robotic Experimentation and Machine Learning. *Cell Reports Physical Science* **1**, 100264 (2020).

110. A. Dave *et al.*, Autonomous optimization of non-aqueous Li-ion battery electrolytes via robotic experimentation and machine learning coupling. *Nature Communications* **13**, 5454 (2022).