The Role of Generative Artificial Intelligence in Internet of Electric Vehicles

Hanwen Zhang, *Student Member, IEEE*, Dusit Niyato, *Fellow, IEEE,* Wei Zhang, *Member, IEEE,* Changyuan Zhao, Hongyang Du, Abbas Jamalipour, *Fellow, IEEE,* Sumei Sun, *Fellow, IEEE,* Yiyang Pei, *Senior Member, IEEE*

***Abstract*—With the advancements of generative artificial intelligence (GenAI) models, their capabilities are expanding significantly beyond content generation and the models are increasingly being used across diverse applications. Particularly, GenAI shows great potential in addressing challenges in the electric vehicle (EV) ecosystem ranging from charging management to cyber-attack prevention. In this paper, we specifically consider Internet of electric vehicles (IoEV) and we categorize GenAI for IoEV into four different layers namely, EV's battery layer, individual EV layer, smart grid layer, and security layer. We introduce various GenAI techniques used in each layer of IoEV applications. Subsequently, public datasets available for training the GenAI models are summarized. Finally, we provide recommendations for future directions. This survey not only categorizes the applications of GenAI in IoEV across different layers but also serves as a valuable resource for researchers and practitioners by highlighting the design and implementation challenges within each layer. Furthermore, it provides a roadmap for future research directions, enabling the development of more robust and efficient IoEV systems through the integration of advanced GenAI techniques.**

***Index Terms*—Generative artificial intelligence, Internet of electric vehicles, scheduling, forecasting, scenarios generation**

Manuscript received January 1, 2000; revised January 1, 2000. This research/project is supported by the National Research Foundation, Singapore and Infocomm Media Development Authority under its Future Communications Research & Development Programme (Grant FCP-SIT-TG-2022-007), A*STAR under its MTC Programmatic (Award M23L9b0052), MTC Individual Research Grants (IRG) (Award M23M6c0113), the Ministry of Education, Singapore, under the Academic Research Tier 1 Grant (Grant ID: GMS 693), and SIT's Ignition Grant (STEM) (Grant ID: IG (S) 2/2023 – 792). (*Corresponding author: Wei Zhang*)

Hanwen Zhang is with both the College of Computing and Data Science, Nanyang Technological University, Singapore 639798, and the Information and Communications Technology Cluster, Singapore Institute of Technology, Singapore 138683 (e-mail: hanwen001@e.ntu.edu.sg and hanwen.zhang@singaporetech.edu.sg)

Wei Zhang, Sumei Sun, and Yiyang Pei are with the Information and Communications Technology Cluster, Singapore Institute of Technology, Singapore 138683 (e-mail: {wei.zhang, sumei.sun, yiyang.pei}@singaporetech.edu.sg).

Dusit Niyato and Changyuan Zhao are with the College of Computing and Data Science, Nanyang Technological University, Singapore 639798 (e-mail: dniyato@ntu.edu.sg and zhao0441@e.ntu.edu.sg).

Hongyang Du is with the Department of Electrical and Electronic Engineering, University of Hong Kong, Pok Fu Lam, Hong Kong (e-mail: duhy@eee.hku.hk).

Abbas Jamalipour is with the School of Electrical and Computer Engineering, the University of Sydney, Australia (e-mail: abbas.jamalipour@sydney.edu.au).

I. Introduction

Electric mobility is the future. This is evidenced by the visions and policies of different nations and companies for achieving sustainable mobility. Electric vehicles (EVs) are gaining popularity rapidly. In the US, the sales of EV reached 1.2 million just in one quarter in 2023, with nearly 8% market share [1]. The numbers are even bigger in China. In 2023, EV's market share was 34% [2]. The trend is not much different in other nations, e.g., the sales of traditional cars with internal combustion engines (ICEs) will be phased out by 2030 and 100% cars will be clean energy based after 2040 in Singapore. Altogether, the global market of EVs reached 392.4 billion USD in 2023 with a predicted compound annual growth rate of 13.9% from 2024 to 2032 [3]. With more and more EVs on the road, they naturally form a IoEV [4]. It shares a similar spectrum of technologies to Internet of things (IoT) and offers new features such as connectivity, grid services, predictive maintenance, and traffic management. IoEV could be regarded as a specialized subset of IoT where EVs, their charging infrastructure, and smart grid interact seamlessly [5]. This creates a network of interconnected devices/systems, leading to various applications, e.g., EV routing problem in IoEV [6], smart EV charging station scheduling [7], EV's battery life prediction [8], and blockchain-based bidirectional energy trading between EVs and charging stations [9].

Same as most new technologies, EV and IoEV have their problems and challenges. Making electricity the main source of power brings constraints related to electricity at the same time. The electricity is stored in batteries, which have charging speed and capacity limits, can degrade over time, and may catch fire occasionally. Those constraints affect EV operations in various aspects, e.g., charging scheduling and battery health monitoring. Furthermore, the impact goes beyond individual EV for IoEV with increased coordination and intelligence demand. One impact is on the power grid, where a significant amount of new electricity load from EVs shall not stress the grid much and hurt the grid's stability. This requires grid-level intelligence such as supply-demand forecasting and matching with the support of IoEV scheduling and vehicle-to-grid (V2G) services.

The problems and challenges have drawn attention from the research communities and industries. Existing efforts can be clustered into application level and technology level. The counterparts of EV and IoEV are traditional cars with ICEs. IoEV is a subset of Internet of vehicle (IoV), with high relevance. Research studies for these applications have the potential of being transferred to IoEV, but this comes with certain constraints and potentially substantial costs for some tasks due to the new challenges of IoEV. For example, charging typically takes hours and cannot be modeled as an instant event, and electricity prices may change hourly or more frequently. As such, the EV charging management system shall schedule and predict the EV loads to minimize cost and avoid overloading the power grid, and charging station installation and operation shall be optimized to promote charging service availability.

Technology-wise, there are solutions specialized to IoEV, with or without using machine learning (ML). Non-ML solutions often use model-based statistical methods for various IoEV related tasks, e.g., user behavior understanding [10] and charging scheduling [11]. A common problem of these methods is that the real system dynamics are modeled in a highly summarized way, e.g., mean and variance, and formulated with simplification to make the optimization tractable. As a result, they are often impractical and fail to offer sufficient accuracy and effectiveness in modeling and charging scheduling. ML-based solutions are becoming the trend. ML models learn from data and continuously improve with more data available. For example, the long short-term memory (LSTM) is often used in the prediction of EV load [12] and voltage changes of battery [13] because of its superior capability of handling long-term dependencies in sequential data, yet the prediction accuracy decreases when uncertainty in the real systems increases.

In this paper, we propose to use GenAI to advance IoEV technologies. GenAI techniques applied in IoEV mirror the same advancements in IoT where a large amount of data from various sources are analyzed and utilized to enhance efficiency, safety, and user experience. The integration of GenAI within IoEV not only enhances its specific applications but also contributes to the overarching goals of IoT by enabling smarter, more autonomous, and connected systems. Moreover, GenAI has the potential to address the challenges mentioned above and case studies for certain IoEV tasks are available, e.g., charging demand forecasting [14], [15] and data augmentation [16], [17]. We aim to go beyond case studies and provide a comprehensive discussion of GenAI's roles in the IoEV ecosystem. Specifically, we structure the system into four layers as shown in Fig. 1. The bottom layer is the battery, which is a critical component of an EV and brings many new constraints such as long charging time and battery degradation to IoEV compared to traditional IoV. Next to the battery layer is the EV layer, where EVs are considered individually for various aspects such as EV charging behaviors and load, as well as an EV routing problem. Then, we consider the existence of many EVs that collectively share a few charging stations connected to the smart grid, which becomes the third layer. This layer features the aggregated demand and aims to achieve the optimal charging scheduling of many EVs. Finally, we provide an investigation of the security layer that is located vertically across the above three layers. For reader's convenience, we also present a list of common abbreviations for reference in Table I. In summary, we make the following main contributions in this paper.

TABLE I
LIST OF ABBREVIATIONS.

Abbreviation	Description
AE	Autoencoder
AM	DRL with general attention model
ANN	Artificial neural network
AutoML	Automated machine learning
BMS	Battery management system
CAN	Controller area network
CNN	Convolutional neural network
DBN	Deep belief network
DDIM	Denoising diffusion implicit model
DDPM	Denoising diffusion probabilistic model
DNN	Deep neural network
DoS	Denial of service
DRL	Deep reinforcement learning
ESS	Energy storage system
EV	Electric vehicle
EVRP	Electric vehicle routing problem
FDIAs	False data injection attacks
FGSM	Fast gradient sign method
FNN	Feed-forward neural network
GAE	Graph autoencoder
GenAI	Generative artificial intelligence
GAN	Generative adversarial network
GDM	Generative diffusion model
GMM	Gaussian mixture model
GPR	Gaussian process regression
EMS	Energy management system
IoEV	Internet of electric vehicles
IoT	Internet of things
IoV	Internet of vehicle
KNN	k-nearest neighbor
LLMs	Large language models
LSTM	Long short-term memory
MADRL	Multi-agent deep reinforcement learning
MAE	Mean absolute error
MARL	Multi-agent reinforcement learning
ML	Machine learning
NLP	Natural language processing
PPO	Proximal policy optimization
PV	Photovoltaic
RL	Reinforcement learning
RNN	Recurrent neural network
SAC	Soft actor-critic
SoC	State of charge
SoH	State of health
SVM	Support vector machine
VAE	Variational autoencoder
VRP	Vehicle routing problem

- We present a detailed survey of the latest GenAI techniques across all layers of IoEV, including an in-depth exploration of various GenAI models. We also highlight their roles in solving various IoEV problems such as data scarcity charging load prediction.
- We systematically categorize GenAI-enabled IoEV applications into four distinct layers for battery, individual EV, the grid, and security. We describe each layer with the specific GenAI techniques for their respective IoEV applications, and provide a holistic view of how GenAI can be integrated within the IoEV ecosystem.
- We provide a summary of publicly available datasets for training GenAI models within the IoEV context. It

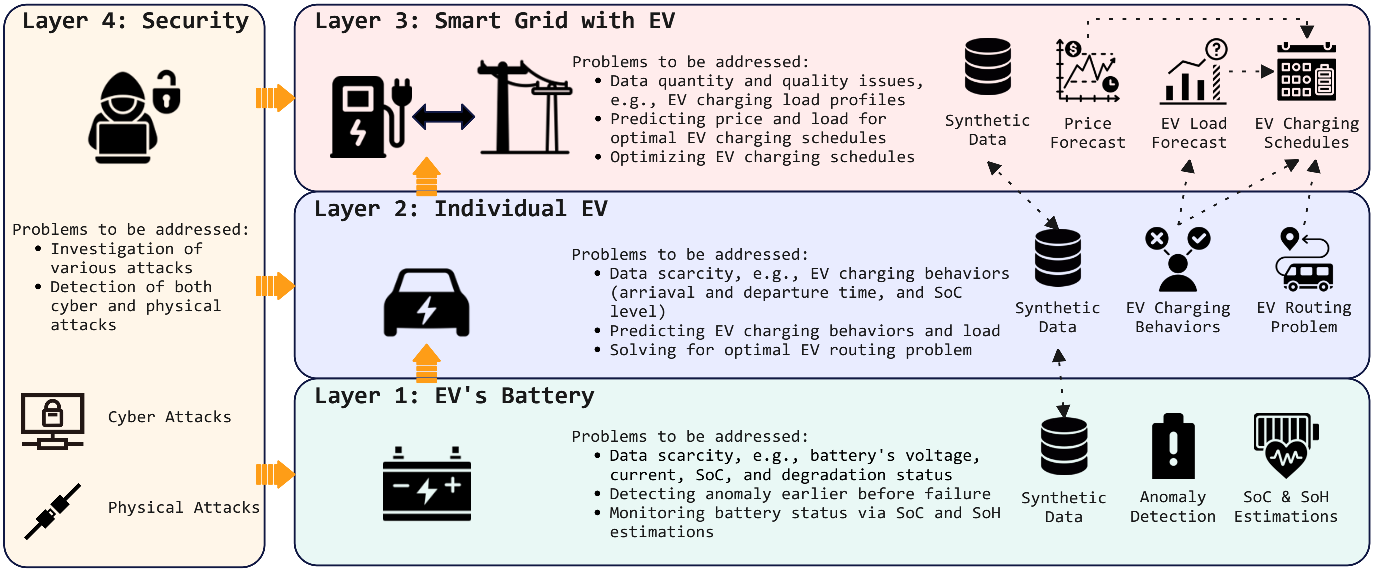

Fig. 1. GenAI for IoEV applications can be categorized into four layers: Layer 1's problems are anomaly detection, SoC estimation, and SoH evaluation. Layer 2 primarily focuses on data augmentation in EV charging behaviors, prediction of an EV load at home, and solving the optimal EV routing problem. Layer 3 concentrates on: Forecasting and augmenting the EV charging load profiles; Optimizing EV charging schedules based on the constraints from either EV charging stations or the smart grid; Predicting the electricity price is necessary for EV charging station operators. Layer 4 studies various attacks which may be harmful to the EV and the charging system. Both cyber and physical attacks need to be studied and detected such as adversarial attacks, false data injection attacks, denial of service attacks, fuzzy attacks, and impersonation attacks.

facilitates future research on GenAI for IoEV applications.

- We identify and discuss the critical challenges and gaps in the existing GenAI applications in IoEV. After that, we suggestion future research directions aiming at solving existing challenges and providing new opportunities.
- We bridge the gap between multiple disciplines including: computer science, electrical engineering, and transportation. This comprehensive survey could serve as a valuable resource for a wide audience.

The subsequent sections of this paper are organized as follows. In Section II, the basic concepts of GenAI and EV charging system are initially conducted to give a brief background. Following this, a comprehensive review of GenAI techniques as they pertain to IoEV is undertaken in Section III. Subsequently, the available public datasets are summarised in Section IV. Next, the future directions are recommended in Section V. Lastly, a conclusion is presented in Section VI.

II. Background

In this section, we introduce the concept of EV and IoEV, the basics of GenAI models, as well as a brief introduction of GenAI's industry adoption.

A. Electric Vehicle (EV) and Internet of EV (IoEV)

Fig. 2 shows the concept of an EV charging system in an electrical distribution network. The network consists of charging stations, residential buildings (e.g., smart homes), load, and distributed resources (e.g., renewable energies, and grid-scale energy storage systems). A distribution system operator (DSO) being one of the grid operators, manages the electrical distribution network. The DSO can coordinate with charging station operators (CSOs) and smart home owners to optimally schedule the charging/discharging of connected EVs. In this case, the DSO sends requests to the contracted CSOs and smart home owners to increase/decrease the energy consumption in a certain period of time. Then, the CSOs and smart home owners respond to the DSO's request and optimally schedule the EV charging/discharging, considering photovoltaic (PV) generation and an energy storage system (ESS) for achieving different objectives such as charging cost minimization. To participate in the day-ahead energy market, CSOs often need to forecast the electricity price and EV load for optimal EV scheduling in advance. Moreover, EV charging behaviors such as arrival time, departure time, charging duration, and unplugging time can influence EV load forecasting. From the EV users' perspective, they may not only be interested in saving charging costs through smart home systems but also need to understand EV's battery status. Two important aspects are state of charge (SoC) and state of health (SoH). The former refers to the remaining quantity of electricity available in the EV battery, which implies the remaining range. The latter indicates the aging status of the battery for making decisions about battery maintenance and retirement [18]. Furthermore, an IoEV ecosystem, is formed by connecting DSO, CSOs, and EV users, and the optimal EV scheduling can be achieved through coordinated efforts among the IoEV components. The approach ensures that the concerns of a DSO, CSOs, and EV users are well considered and addressed based on their respective interests and constraints.

B. Basics of Generative Artificial Intelligence (GenAI)

GenAI is proposed based on traditional ML, which are discriminative models that learn the probability distribution $p(y|x)$ in Bayes' theorem for input x and output y. The

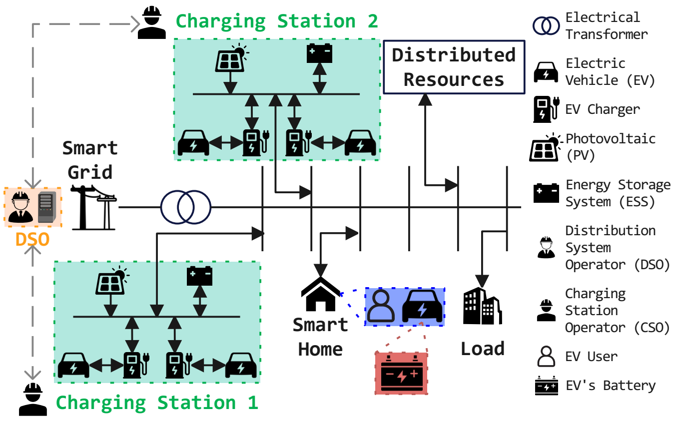

Fig. 2. Concept of EV charging system in an electrical distribution network: the network consists of charging stations, smart home, distributed resources, and load. Distribution system operator (DSO) and charging station operators (CSOs) manage the distribution network and the charging stations respectively. Smart homes and charging stations with PV and ESS can coordinate with DSO to ensure a stable and robust smart grid operation.

discriminative models categorize the data space into different classes by learning the decision boundaries. They often focus on distinguishing between different classes or outcomes. In the context of computer vision applications, the discriminative models are incapable of processing unknown inputs and it is required to provide label distributions for every image. As such, traditional ML models are mainly used for classification, regression, clustering, etc.

GenAIs are different from traditional ML and they learn the probability distribution of data $p(x)$ for unconditional generative models or $p(x|y)$ for conditional generative models. This allows GenAIs to understand the underlying data distribution and generate new samples from the distribution. The generated data/contents could be statistically similar to the input data and the similarity is useful for data augmentation, simulation, and creative tasks. Overall, the generative nature of GenAI is useful for developing more dynamic and innovative solutions across various domains.

1) Transformer: The Transformer architecture was first introduced by the influential article "Attention is all you need" [19] which utilizes a self-attention mechanism to capture long-range dependencies without relying on sequential processing. Fig. 3 (a) shows a single-layer Transformer that includes a typical self-attention module and feedforward layers with residual connections where each layer begins with the application of self-attention. The resulting output from the attention mechanism is then processed by feedforward layers, where the same feedforward weight matrices are used independently for each position. After processing through the first feedforward layer, a nonlinear activation function, e.g., `ReLU`, is applied.

The Transformer structure is widely used in large language models (LLMs) [20], image processing [21], automatic speech recognition [22], visual question answering [23], sentiment analysis [24], etc. Besides the traditional ML applications such as audio, computer vision, and natural language processing (NLP) applications, Transformer is also extended to other domains such as the anomaly detection of the EV battery [25], EV routing [26], and EV charging load forecasting [27]. The unique strength of capturing long-range dependencies makes Transformer suitable for parallel computing, scalable according to the different tasks, and transferable with a pretrained model. However, the Transformer is not perfect. For example, the Transformer used in LLMs often requires significant computational and memory resources during the training. The performance of the Transformer may be compromised when there is only a small amount of dataset available for training.

2) Generative Adversarial Network (GAN): The GAN model [28] consists of two deep neural network (DNN): a generator and a discriminator. Both networks engage in an adversarial training process; one network generates new data while the other assesses whether the data is real or fake. Fig. 3 (b) shows the principle of GAN. Assume that the data x_i is extracted from the data distribution $p_{data}(x)$, with the goal of sampling it according to p_{data}. The sample z, the latent variable, drawn from the simple prior $p(z)$ is fed into the generator network. Then, the generated sample is drawn from the data distribution of the generator p_g. Subsequently, the joint training of the generator and discriminator is carried out until p_g converges to p_{data}, i.e., $p_g \approx p_{data}$. In the training process, the generator network is trained to "deceive" the discriminator that concurrently learns to classify whether the generated data is real or fake with generator and discriminator loss functions. From the mathematical point of view, it is similar to a two-player minimax game with an objective function.

GAN was first proposed in 2014 [28]. Some evolved versions are deep convolutional GAN (DCGAN) [29] in 2016, Progressive GAN [30] and Wasserstein GAN [31] in 2017, StyleGAN [32] in 2019 as well as its adoption in semantics communication [33], [34], and MaskGAN [35] in 2020. Besides its applications in computer vision, GAN is also extended to other domains, e.g., data augmentation for battery's SoC estimation [36], data augmentation for EV charging behaviors [17], data augmentation for EV charging load profile [37], EV load forecasting or changing scenarios generation [38], as well as generation and detection of adversarial attacks [39]. Moreover, the idea of GAN combined with imitation learning formed a new term called generative adversarial imitation learning (GAIL) [40] which is particularly useful in reinforcement learning (RL) when defining a reward function is explicitly difficult. Besides, a variant of GAN is generative adversarial imputation network (GAIN) [41], which is particularly useful for missing data imputation and accordingly improves performance. This is important for IoEV applications that involve incomplete datasets, e.g., with missing entries in charging behavior logs or irregularities in electricity consumption records. While GAN is capable of learning complex and high-dimensional data distribution as well as generating high-quality data, there are challenges, such as balancing generator and discriminator, avoiding mode collapse, and accelerating convergence.

3) Autoencoder (AE) and Variational Autoencoder (VAE): An AE is an unsupervised approach that extracts feature vectors from raw data x without labeled examples. It consists of the encoder and decoder during the training as shown in the left side of Fig. 3 (c). The encoder learns the useful information,

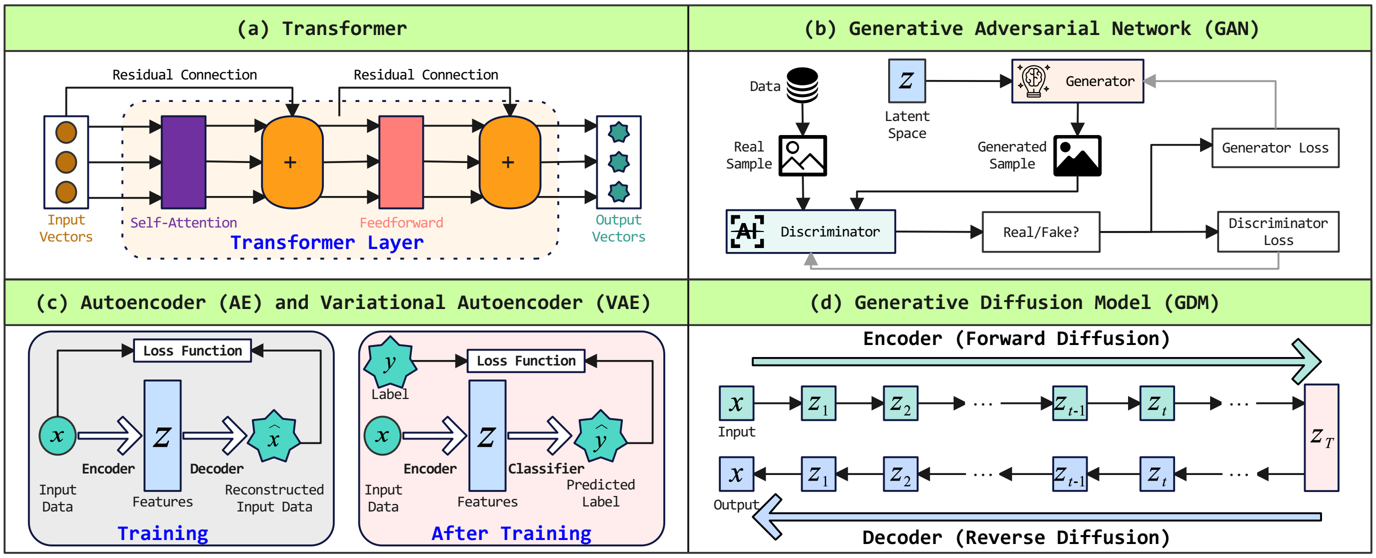

Fig. 3. Concept of basic GenAI models — Transformer, GAN, AE, VAE, and GDM: (a) shows a single-layer Transformer process where the output vectors are achieved by passing the input vectors through self-attention and feedforward layers with residual connections; (b) illustrates principles of GAN where the generator competes with the discriminator by producing increasingly realistic samples to "fool" the discriminator, while the discriminator attempts to differentiate between real and fake data; (c) depicts principles of AE: the left side shows the training process of AE, while the right side depicts its usage once the model is completely trained; VAE is a subcategory of AE, but it differs slightly in that VAE uses the mean and diagonal covariance to generate samples in both the encoder and decoder; (d) displays the processes of GDM consisting of forward diffusion and reverse diffusion.

i.e., features Z, from the input raw data x. The encoder could be `Sigmoid`, fully connected, or `ReLU` convolutional neural network (CNN). After that, the decoder utilizes the learned features to reconstruct the input data $\hat{x}$. The decoder could be `Sigmoid`, fully connected, or `ReLU` CNN (up-convolution or transposed convolution). The loss function of the training could be the L2 distance between input and reconstructed data. After the training, the decoder is removed and the trained encoder is useful for the downstream tasks as shown on the right side of Fig. 3 (c). For example, a supervised classification model can be initialized using the encoder which is often fine-tuned jointly with the classifier and a task-specific loss function.

AE can be used as the context encoder in semantic inpainting tasks [42], the temporal context encoder for video applications [43], and representation learning [44]. Besides, AE is useful in EV related applications, e.g., detection of false data injection attacks (FDIAs) that may pose a threat to the EV charging process [45]; cyber and physical anomaly detection for abnormal behaviors within EV charging stations [46]; detection of denial of service (DoS), fuzzy, and impersonation attacks to the controller area network (CAN) protocol communications of EVs connected to EV charging system [47]. AE offers several benefits. It can effectively reduce the dimensionality of the data and learn the compact representations, which enables it to be useful for data preprocessing and feature extraction. It allows unsupervised learning since no labeled data is required for training, and is suitable for identifying anomalies by learning to reconstruct normal data well. However, vanilla AE lacks generative capabilities compared to VAE and GAN, and cannot capture complex data distributions effectively.

The VAE is a type of directed model that relies on approximate inference learned during training and can be optimized solely through gradient-based methods. The concept of VAE is similar to AE, as it represents a specific subset of AE. The encoder and decoder of the original AE are modified in the VAE where the sampling processes are achieved using means and diagonal covariances. The VAE has been used for common ML applications such as facial expression editing [48], future forecasting from static images [49], and point cloud completion [50]. Its usage has also been extended to IoEV applications, e.g., anomaly detection for EV's battery [51] and data augmentation for EV load profiles [14]. Generally, VAE has better generative capabilities compared to vanilla AE, and can generate new samples similar to the training dataset. The capabilities make VAE suitable for data augmentation and image synthesis, and useful for the clustering and interpolation tasks. However, using a Gaussian prior and a reconstruction loss function, VAE may not be able to capture fine details well, e.g., in images with blurry outputs. Moreover, VAE may suffer from mode collapse where the model generates a few types of outputs despite having diverse data. Its generative capabilities can be also limited by the latent space Gaussian assumption.

4) Generative Diffusion Model (GDM): Fig. 3 (d) shows the principle of the GDM. In its forward diffusion process, i.e., encoder, the model transforms x through a sequence of latent variables $z_1, \ldots, z_T$. The procedure is predefined and progressively blends the data with noise until only noise persists at z_T. Given a sufficient number of steps, both the conditional distribution $q(z_T|x)$ and the marginal distribution $q(z_T)$ of the final latent variable converge to the standard normal distribution. All the learned parameters are included in the decoder as predefined. In the reverse diffusion process, the data is processed through the latent variables by the decoder which is trained to eliminate noise progressively at each stage.

The backward mapping between each pair of adjacent latent variables z_t and z_{t-1} is achieved through the training of a sequence of networks. Each network is guided by the loss function to perform an inversion of its associated encoder step. Following the training process, the new examples are created by sampling vectors of noise z_T and then processing these through the decoder.

The diffusion model has been widely utilized in image applications, e.g., the denoising diffusion probabilistic model (DDPM) [52] for generating high-quality image samples without adversarial training, denoising diffusion implicit model (DDIM) [53] for improving the sampling speed of DDPM, stable diffusion [54] for generating images from text, and ControlNet [55] allowing model being trained with a small dataset of image pairs. Furthermore, the diffusion model has been extended for other applications, e.g., network optimization [56], generating optimal pricing strategies [57], repairing and enhancing extracted signal features in wireless sensing [58], estimating the signal direction of arrival in near-field scenarios [59], estimating battery's SoH [60], and generating EV charging scenarios [61]. GDM can produce high-resolution and realistic samples, achieving similar performance of GAN generated data or even better, e.g., outperforming the traditional Gaussian mixture model (GMM) model by 91% for charging scenarios generation. Compared to GAN, GDM shows better training stability in versatile applications, e.g., optimization [56], [57]. Moreover, the clear and iterative process of refining the generated data from noise to coherent output makes the generation process of GDM more interpretable. Nevertheless, the iterative nature of the process can be the drawback of GDM. It requires a significant amount of computational resources for the training and inference of GDM. This nature also slows down the sampling process of GDM compared to the VAE and GAN models that can generate samples in a single pass. When designing and tuning the GDM, the noise schedules and model architecture have to be carefully considered. This increases the complexity of training well-performed models. Moreover, GDM lacks controllability in specific attributes or features.

5) GenAI Development and Deployment: Three important stages are training, fine-tuning, and deploying the above-mentioned GenAI models in practice. First, GenAI models are trained on large and diverse datasets to learn broad patterns and general features. Foundation GenAI models are produced in this stage and serve as the backbone of various applications. However, such models are general-purposed, so a fine-tuning stage is needed to customize the models for specific applications, such as EV charging, load forecasting, and route optimization. With the support of the backbone and given the realistic constraints, the stage involves small-scale datasets only which, however, shall be domain-specific. Finally, the fine-tuned models are deployed in practice for real-time inference which is significantly less computing intensive compared to training and fine-tuning models. Overall, the three stages, with different resource demands and objectives, orchestrate GenAI development and deployment.

6) Summary: Transformer, GAN, AE, VAE, and GDM are foundational models of GenAI which have demonstrated their versatility across a wide range of applications. They were first invented for traditional ML tasks such as computer vision and NLP and gradually adapted for electric mobility applications. For EV batteries, GenAI helps estimate battery capacity and health and detect potential anomalies. For EVs, GenAI can be used to understand charging behavior, energy management, and routing. With many EVs forming an IoEV, data augmentation and large-scale charging scheduling can be supported by GenAI. Furthermore, GenAI's role in detecting cyber and physical attacks on IoEV components can be explored. Overall, adapting foundational GenAI models for various IoEV aspects from batteries to security requires careful consideration of GenAI's respective strengths and limitations.

C. GenAI for Industry

With the basics of GenAI, we present adoption of AI and GenAI in EV industry as well as other domains.

1) EV Industry: GenAI has not yet been specifically presented in the EV industry but AI has been among the strategic focuses of the big EV players and we introduce some latest progress from two major players. AI has been adopted in various aspects of Tesla's business, e.g., EV manufacturing and autonomous driving. The large-scale AI adoption is enabled by Tesla's Dojo supercomputer which offers abundant computing resources and realizes computational-intensive tasks such as high-throughput EV video processing. The company also owns an overarching aspiration to develop artificial general intelligence (AGI). BYD has introduced its XUANJI Architecture, an intelligent vehicular framework, integrating electrification with intelligent functionalities and functioning as the EV's cognitive core. EV's internal and external environments are monitored in real-time and the collected information is used to make decisions about the EV's operation to improve safety and comfort. The industry favors the integration of AI and physical systems and we foresee an increasing adoption of GenAI in the EV industry.

2) Other Industry Sectors: Besides EV and the transportation sector, GenAI has found widespread adoption in other industry sectors and we introduce a few sectors as follows. For business and finance, a survey [62] is available with a description of GenAI's practical applications and cutting-edge tools in the sector. The survey [63] is about GenAI's applications, advantages, and obstacles for the healthcare sector. The education sector is witnessing GenAI's significant impact and the authors in [64] discuss GenAI's ability to boost learner engagement and motivation, emphasizing the need for ethical guidelines and human oversight as well as GenAI's impact on critical thinking. Compared to different industry sectors, the IoEV industry has various unique features, e.g., battery and grid integration. The features require GenAI algorithms to be customized and specialized for performance maximization.

III. Technical Reviews: Generative Artificial Intelligence (GenAI) for Internet of Electric Vehicles (IoEV)

In this section, we provide an overview and discussions of GenAI's application in different layers of IoEV.

A. Layer 1: Electric Vehicle (EV)'s Battery

The battery is the core component of an EV. It powers an electric motor and has a direct impact on EV's range, performance, and efficiency. It is also the most expensive part of most EVs, due to the fact of which, battery's operating condition and longevity play a crucial role in the overall user experience and sustainability of the EVs. We specifically would like to survey three important aspects of batteries, including anomaly detection [25], [51], SoC [36], [65], [66], and SoH [60]. For anomaly detection, various detection algorithms can be integrated with the battery management system (BMS) for proactive battery maintenance before any potential failures cause hazardous battery damage. The SoC and SoH indicate the battery's short-term energy capacity and long-term health condition, respectively. Specifically, SoC measures the stored energy relative to the maximum capacity and SoH reflects the battery's maximum capacity which degrades over time. Both aspects are influenced by different factors such as temperature, charging voltage, and cycling history [18].

1) Anomaly Detection: Despite continuous technological progress in the past years, battery safety remains a big concern for EV owners and customers. One of the most hazardous issues is battery fire, the occurrence of which has raised debates and doubts about battery safety and highlights the necessity of early anomaly detection to prevent potential safety breaches and irreversible damage [67].

a) Traditional Machine Learning (ML) Approaches: The early effort of battery anomaly detection and diagnosis is based on traditional ML algorithms, e.g., the random forest [68], multiclass relevance vector machine [69], and finite-element-based models [70]. These ML algorithms are generally simple to implement but the performance suffers when the raw data is used directly. Domain knowledge complements the capability of the algorithms by guiding the extraction of domain-specific and useful features as the ML input, with improved correlation with battery anomalies. The advancements of deep learning offer new methods for battery anomaly detection. Among the methods, LSTM should be the most widely used architecture [71], [72], which is capable of predicting battery voltage with multiple inputs [71] and forecasting parameters such as voltage, temperature, and SoC simultaneously [72]. However, LSTM being a type of recurrent neural networks (RNNs) presented challenges in practical training scenarios, with poor training stability and issues such as vanishing or exploding gradients.

b) AE and VAE-based Approaches: GenAI can potentially address the above-mentioned challenges and the reconstruction-based models have been studied. AE as a basic reconstruction model is shown to be ineffective for generating diverse and high-quality data. This is largely due to the deterministic nature of the latent codes produced by the AE encoder. A more suitable model is VAE, which adeptly learns the probability distribution of multivariate time series (MVTS) to be robust against perturbations and noise [73]. In [51], a semi-supervised VAE-based anomaly detection model was proposed for early anomaly detection in battery packs. The model detects different anomalies such as irregular terminal voltage, differences between all bricks, and abnormal temperature fluctuations. The model, named GRU-VAE, consists of a gated recurrent unit (GRU) and a VAE, where the former captures MVTS and the latter reconstructs the input samples. Worth mentioning that the paper is based on an EV operation dataset from the National Service and Management Center for Electric Vehicles (NSMC-EV) in Beijing [74]. The dataset includes 13-dimensional time series signals such as data acquisition time, vehicle speed and state, charging state, voltage, current, mileage accumulated, SoC, temperature, insulation resistance, and DC–DC state. GRU-VAE shows an improvement in anomaly detection compared to the AE-based models, achieving a 25% increase in F1-score [51]. Nevertheless, the model may need to be updated, given the changed data patterns over time.

c) Transformer-based Approach: Anomaly detection can also be transformer-based [25]. The authors in [25] developed BERTtery, a transformer-based model for battery fault diagnosis and failure prognosis. BERTtery is able to capture early-warning signals across multiple spatial–temporal scales in various operational conditions, predict battery system fault evolution using onboard sensor data, and avoid faults leading to thermal runaways [75]. The model is based on a dataset with various battery faults and failures, e.g., internal short circuits, lithium plating faults, overcharging/overdischarging, abnormal self-discharge, abnormal capacity degradation, abnormal voltage fluctuations, abnormal temperature behaviors, electrolyte leakages, cell balancing issues, and thermal runaways [76]. Specifically, the model's input includes the time series of voltage, current, and temperature, sampled every 10 seconds from real-world EV operations. The output of the model is the predicted safety labels. For future work, the generalization ability of the proposed model could be improved by considering different battery types and operational conditions, to enhance the early warning capabilities with reduced false alarms.

2) State of Charge (SoC) Estimation: SoC indicates how much battery energy remains and directly affects the EV's range [77]. Thus EV owners monitor SoC for trip planning, charging scheduling, and battery usage optimization.

a) Traditional ML-based Approaches: Many ML algorithms have been adopted for SoC modeling, e.g., the random forest [78], Gaussian process regression (GPR) [79], and support vector machine (SVM) [80]. These algorithms are commonly adopted partially because of their simplicity, which however is one of the reasons that they cannot handle the complex battery operating conditions. For relatively more complex algorithms, the CNNs [81] capture local feature representation for time series forecasting and may overlook distant variable correlations, limiting the ability to capture remote topological structures. The long dependencies can be learned by LSTM [82], which however is limited with extended sequences due to the inherent constraint of recurrent models, where signals must traverse both forward and backward. Moreover, traditional ML models typically require extensive data for model training; otherwise, the model could be overfitted with reduced accuracy. Unfortunately, data is often scarce. One idea is to use GenAI for data augmentation, and this idea has been studied for SoC estimation, as detailed below.

b) GAN-based Approaches: Efforts to utilize GenAI for general-purpose data augmentation have yielded successes, e.g., convolutional recursive GAN (CR-GAN) [83], recurrent conditional GAN (RC-GAN) [84], time-series GAN (TimeGAN) [85], and global trend diffusion algorithms [86]. Among the GAN variants, CR-GAN and RC-GAN struggle to capture the temporal dynamics encompassed within the entirety of a time series sometimes, owing to gradient vanishing or explosion that arose when processing long sequences with RNNs. TimeGAN divides long time series into smaller segments, potentially leading to the loss of crucial cross-segment information. The global trend diffusion algorithm lacks sufficient adaptability to different scenarios because of its triangle distribution and fails to adequately represent the complexity of the original data.

Several GAN-based models are specially designed for SoC estimation. The Wasserstein GAN (WGAN) is a promising base model, which is shown to be robust by generating the underlying real data distribution, enhancing the generation quality of vanilla GAN, and accelerating convergence [87]. Hence, the authors in [65] developed a time-series Wasserstein GAN (TS-WGAN) based on WGAN for SoC estimation. The new model consists of data pre-processing and a Wasserstein GAN with gradient penalty (WGAN-GP) architecture [87]. Besides, the model is trained by the EV dataset [88] and LG 18650HG2 Li-ion Battery dataset [89]. The former includes EV information during charging/discharging such as timestamp, vehicle speed, voltage, current, cell temperature, motor controller input voltage and current, mileage, and SoC. The latter consists of the battery's performance data during charging, discharge cycle measurements, and drive cycles. Note that TS-WGAN may suffer from convergence issues as it requires complex index modifications due to the nature of GANs.

Another WGAN-based model is proposed in [66], named conditional LSTM Wasserstein GAN with gradient penalty (C-LSTM-WGAN-GP). It is an LSTM-based conditional GAN model and owns the capability to generate data that closely resembles actual battery data across various profiles. The training data is obtained from the experiments. Two Li-ion rechargeable cells were set up for the experiments where one is Li-ion with the material of Ni/Co/Mn ternary composites and another one is Li-ion with the material of lithium iron phosphate. The usable battery information, such as terminal voltage, current, and temperature, was monitored and recorded while SoC was estimated after the experiment was done. Future improvements can be architecture and training convergence optimization, as well as system integration with existing BMS.

Extending from the above WGAN-based models, the authors in [36] introduced a time-series deep convolutional GAN (TS-DCGAN) framework. The framework combines the time-frequency domain techniques and DCGAN [29] to train the models for SoC estimation. The models generate synthetic datasets with high fidelity and diversity, effectively capturing the dependencies between multidimensional time series. The training data is obtained from LG INR18650HG2 batteries [90], which are common EV batteries [36]. The dataset consists of battery voltage, current, cell temperature, SoC at different temperatures, and driving cycles. Real data is complemented with the generated synthetic data. The experimental results show that TS-DCGAN successfully reduces the discriminative score of CR-GAN by 53% and TimeGAN by 46% in producing reliable synthetic datasets for the subsequent SoC estimation tasks. Generally, the common issues of the GAN-based models, such as training stability, also apply to TS-DCGAN, which can be further enhanced.

3) State of Health (SoH) Estimation: EV owners are not only interested in knowing the battery's SoC but also SoH [91]. This allows them to plan effectively and schedule maintenance or replacement as needed. It also offers vital insights into EV battery control strategies, protection mechanisms, and sustainable development [92]. The SoH estimation methods can be categorized into model-driven and data-driven approaches. The former includes the electrochemical model [93] and the equivalent circuit model [94]. Relatively, the latter has advantages such as independence from prior knowledge of battery mechanisms and avoidance of subjective intervention, with a focus on latent input-output relationships.

a) Traditional ML Approaches: SoH prediction has been addressed by different traditional ML algorithms, e.g., artificial neural network (ANN) [95], SVM [96], automated machine learning (AutoML) [97], GPR [98], CNN [99], RNN [100], LSTM [101], GRU [102], and Bayesian neural network (BNN) [103]. The above models are regarded as discriminative models, which focus on battery health indices including cycles and capacity. They map input parameters to output variables without prior sample knowledge, and adjust network weights through training with a loss function. However, it is challenging for them to well capture intrinsic characteristics to represent the battery's operational dynamics accurately, and the challenge is often addressed by increasing the quality and quantity of the training dataset.

b) Diffusion-based Approach: Compared to conventional discriminative ML models, GDMs, can capture the distribution characteristics inherent in training data more accurately, thereby offering a more comprehensive understanding of the underlying problem. By leveraging generative diffusion techniques, one could mitigate the risk of introducing significant deviations to the overall distribution of training data, thereby facilitating robust modeling that transcends merely isolated feature representations [104].

The authors in [60] introduced a diffusion-based model namely, DDPM, to predict the SoH of lithium-ion batteries with both offline and online modeling. The dataset used in their experiment is a lithium iron phosphate battery dataset from TOYOTA Research Institute [105] and a nickel cobalt manganese battery dataset from their laboratory [60]. The datasets consist of specifications such as rated capacity, number of cells, charging/discharging current, maximum/minimum cut-off voltage, and number of cycles. The prediction variables are the battery's capacity in the unit of Ah. The proposed DDPM outperforms other ML techniques in terms of several SoH prediction error metrics, e.g., root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). Taking RMSE as an example, DDPM can reduce prediction errors of RNN by 57%, LSTM by 35%, GRU by 10%, transformer [106] by 70%, and CNN-Transformer [107] by 52% [60]. In the future, DDPM's effectiveness for SoH may further be explored by comparing it with other GenAI methods such as VAE and GAN.

Overall for layer 1, we provide a summary of the GenAI

TABLE II
SUMMARY OF GENAI MODELS FOR IOEV IN LAYER 1 FOR BATTERIES.
●: GENAI METHODS; ✓: PROS OF THE METHODS; ✗: CONS OF THE METHODS.

Application	Reference	GenAI Model	Pros & Cons
Anomaly Detection	[51]	GRU-VAE	● A GRU-VAE framework for battery anomaly detection. ✓ Learn probability distribution of multivariate time series data adeptly. ✓ Robust against perturbations and noise. ✗ Adaptability issues and frequent model updates for new data.
	[25]	BERTtery	● A transformer-based method for battery fault diagnosis and failure prognosis. ✓ Learn battery's nonlinear cell behaviors in a self-supervised data-driven manner. ✓ Competitive performance, e.g., above 95% in accuracy, precision, recall and F1 score. ✗ Lack generalization ability to various battery types and operational conditions. ✗ Need for an improved early warning predictions with reduced false positives and negatives.
SoC Estimation	[65]	TS-WGAN	● A GAN-based approach for SoC estimation of lithium-ion batteries. ✓ Robust and able to generate underlying data distribution. ✓ Enhanced generation quality of vanilla GAN and accelerated convergence. ✗ Convergence issue during training and requiring complex index modifications. ✗ Limitations for real-time SoC estimation due to its computational intensity.
	[36]	TS-DCGAN	● A GAN-based approach to generate synthetic data for SoC estimation. ✓ Produce synthetic data of high fidelity and diversity. ✗ Common issues of GAN-based structure, e.g., stability of model training.
	[66]	C-LSTM-WGAN-GP	● A LSTM-based conditional GAN model for data generation and SoC estimation. ✓ Stable model training performance. ✓ Generate realistic battery data cross various profiles. ✗ Need for an improved model architecture and the speed of convergence. ✗ Lack of implementation on online BMS.
SoH Estimation	[60]	DDPM	● A diffusion-based model for battery SoH estimation. ✓ Low prediction errors compared with RNN, LSTM, GRU, and transformer-based models. ✗ Need for a comparison with other GenAIs, e.g., GAN and VAE for the same application.

works in Table II. From the table, we can see that the VAE-based [51] and transformer-based [25] models can be utilized for battery anomaly detection. Moreover, GAN-based models, e.g., TS-DCGAN [36] and C-LSTM-WGAN-GP [66], are developed mainly for data augmentation to enhance the accuracy of SoC estimation. And TS-WGAN [65] considers both data augmentation and SoC estimation. Furthermore, the recent advancements in generative diffusion models have been applied in DDPM [60] for SoH estimation.

B. Layer 2: Individual Electric Vehicle (EV)

EV is an integrated system that combines key components like batteries in layer 1. Instead of focusing on one component as the research works for layer 1, the research in the EV layer emphasizes the integration and functionality of the EV system. GenAI has been studied for the EV layer for two aspects including charging [17], [108] and routing [26], and we present technical details below.

1) EV Charging Behaviors and Loads: The research on EV charging is important for optimizing energy usage, balancing grid demand, and improving charging efficiency. The growing adoption of EVs makes data-driven approaches ideal for related research with a growing amount of EV data generated. The approaches are typically designed for optimizing the charging parameters and analyzing parameter correlations, user travel patterns, and vehicle trajectories [109].

a) Traditional ML-based Approaches: Several ML algorithms have been applied in load forecasting and energy management for EVs, e.g., RNN [110], CNN [111], CNN-GRU [112], and LSTM with RL [12]. A common issue of the above algorithms is the demand for extensive training data (to avoid overfitting and underfitting) [113]. Such data demand cannot be satisfied by many EV service providers, facing realistic constraints, e.g., only 365 charging samples

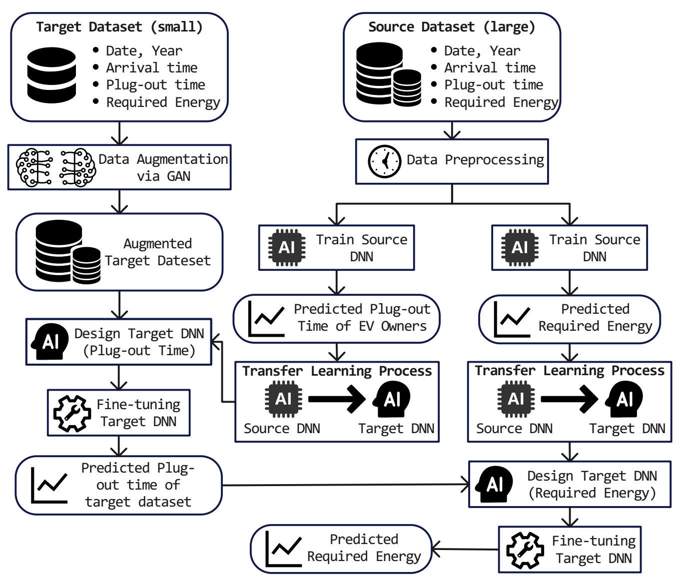

Fig. 4. The flowchart of proposed framework to address the cold-start forecasting problem in predicting the EV charging behaviors such as plug-out hour and required energy for newly committed EVs [17].

per year. This is especially true at the beginning stages of data collection, causing the famous "cold-start forecast problem" [114]. Therefore, researchers are looking for methods to generate extensive datasets with relatively small-scale data collected from real EVs, and GenAI is a promising method.

b) GAN-based Approaches: The above-mentioned cold-start problem has been addressed in [17]. The authors developed a transfer learning-based framework using a deep generative model, GAN, to address the problem of predicting EV charging behaviors. Fig. 4 shows the flowchart of the proposed framework which integrates GAN and DNN for the forecasting tasks. As seen from the figure, two source DNNs are trained on residential EV owner data to predict the plug-out time of EV owners and the required energy, respectively. The transfer learning then adapts the knowledge of the tested EV to the target DNN (plug-out time) for forecasting new EV plug-out hours. The same approach is applied to the target DNN for required energy. Also, GAN models are used to augment the target dataset. The target DNNs, with the weights from the source EV model and the GAN-generated data, are fine-tuned to predict the charging behaviors of target EVs.

Besides, the research is based on a public dataset from EA Technology, which specializes in providing asset management solutions for the owners and operators of electrical assets [115]. The available parameters of residential charging events include dates, arrival hours, plug-out hours, and required energy. Finally, the derived models can achieve significant performance gains over support vector regression (SVR) [116] by 12%, decision tree regression [117] by 57%, k-nearest neighbor regressor (KNNR) by 60%, DNN [112] by 31%, and a GAN-DNN based approach [16] by 10% [17]. For future work, the performance of proposed method could be improved by applying the clustering algorithms for efficient transfer and multi-source datasets for model generalizability.

In [118], a convolution conditional GANs with Wasserstein distance as network objective function (CW-GAN) model was developed for generating EV charging behavior parameters such as arrival time, departure time, and SoC. The study is based on a small dataset of private EVs and charging piles in three functional zones (i.e., office, business, and residential areas) [118]. The input of the model includes noise and conditional labels, and the output generates different parameters of EV charging behavior. Note that the quality of conditional labels has a significant impact on the overall performance of CW-GAN, so accurate label selection is crucial.

Besides charging behavior, GAN has also been studied for charging load. In [119], a GAN-based home electricity data generator (HEDGE) tool is introduced to semi-randomly generate synthetic daily profiles of EV and household loads as well as PV generation. The generated residential energy data spanning multiple days exhibits consistency in terms of magnitude and behavioral clusters. Several profile datasets are used for model training. Household load and solar generation are available in TC1a [120] and TC5 [121], respectively, from the customer-led network revolution (CLNR) project. The EV loads are estimated based on the general population's travel patterns dataset from the UK's National Travel Survey [122]. In the future, HEDGE will be valuable for subsequent research tasks, e.g., optimal scheduling with HEDGE generated profiles.

c) Hybrid Approach: Instead of using GAN alone, hybrid solutions have been investigated for synthetic data generation also. In [108], a VAE-GAN model is developed to generate synthetic time-series energy profiles such as EV load profiles in smart homes. The synthetic data is subsequently utilized in the Q-learning-based home energy management system (EMS) to maximize long-term profit through optimal load scheduling. The model is trained with the iHomeLab PART dataset [123], which includes power consumption profiles of five residences. The study compared the VAE-GAN with a Vanilla GAN and a GMM, employing the Kullback-Leibler (KL) divergence to assess the distance between real and synthetic data distributions. The results reveal that VAE-GAN can achieve 18% and 33% performance improvement over GAN and GMM, respectively, in generating EV load data [108]. The improvement shows that the model is able to learn various smart home data distributions (e.g., electric load, PV generation, and EV charging load), and generate realistic data samples without prior analysis before training. Nonetheless, the model relies on the quality and diversity of training data with inconsistent scalability and adaptability. Future research may incorporate diverse datasets to enhance the model's robustness and improve scalability across broader smart grid applications.

2) EV Routing: The future of urban delivery is likely to be driven by autonomous green vehicles [124]. The trend underscores the significance of efficient route planning, addressed as the electric vehicle routing problem (EVRP). In EVRP, an EV starts from a designated depot with a partial/full charge to serve customers with time restrictions. Each EV can stop at charging stations or return to the depot to recharge. The goal of EVRP is to find cost-effective routes for the EV fleet subject to battery constraints.

a) RL-based Approaches: Recently, researchers have applied supervised learning and RL to address the vehicle routing problem (VRP) amidst the growth of ML, e.g., a pointer network [125]. A significant challenge is to obtain sufficient labels for producing optimal solutions in large-scale problems like EVRP. RL is a label-free approach so it can be a viable option for addressing large-scale problems [126]. Deep reinforcement learning (DRL), as a type of RL, has been applied to solve the VRP, often utilizing the encoder-decoder architecture in neural network design [127], and achieved good performance. However, related research works often focus on basic routing issues, and overlook the complexities of EVRP with energy and charging constraints, which are unique compared to traditional VRP.

b) Transformer-based Approach: Transformer architecture with an attention mechanism has been proven to be effective in improving computational efficiency and solution quality in solving VRP [128]. Hence, the authors in [26] proposed a Transformer-based DRL method for energy minimization of EVRP. Fig. 5 shows the proposed framework where the policy network of DRL is modeled by the Transformer's encoder-decoder structure. The features of EVRP are captured by the feature embedding module and the policy gradient method is employed for policy training. In the context of an EVRP instance, the graph information including the vehicle and the

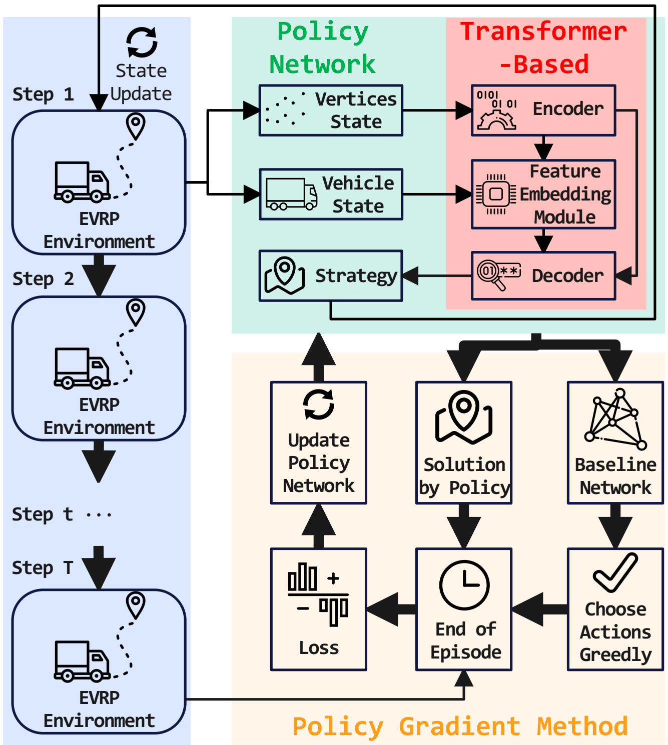

Fig. 5. Framework of DRL with Transformer for EV routing problem [26].

states of vertices undergo encoding by an encoder. This process facilitates the step-by-step construction of EV routes by the decoder, leveraging inputs from both the encoder and the feature embedding module. Subsequently, updates to the parameters of the policy network are made based on reward values derived from the policy network and the baseline network.

The EVRP instances are created according to the procedure outlined in [129]. The locations of customers and charging stations were chosen uniformly at random from a square kilometer area. Finally, the proposed transformer-based DRL method is compared with the exact algorithm [130], improved ant colony algorithm (ACO) [131], adaptive large neighborhood search (ALNS) algorithm [132], and DRL with general attention model (AM) [128] under various scenarios, e.g., different number of EVs and charging stations. Specifically for a case study with 100 EVs, the proposed method can reduce the energy consumption of EVs by 1% compared to AM, and 10% to ACO [26]. Future work includes enhancing the model's efficiency, testing with real-world data, and expanding the model to accommodate more complex scenarios, e.g., multiple vehicle types and different environmental constraints. Other GenAIs such as GAN and GDM can also be considered for navigation and route optimization application [133].

At the end of such EV layer, we would like to mention Autonomous Vehicles (AVs), which often occur with EV together as new concepts of vehicular technology. However, an AV is not necessarily an EV and it can be driven by different power sources including but not limited to electricity. Our focus in this paper is EV and some interesting discussions of utilizing GenAI in AVs are available in [134], [135] about AV-related aspects such as trustworthiness and navigation.

Overall, the GenAI applications for IoEV in layer 2 are summarized in Table III. We find that the GAN-based models, GAN-DNN [17] and CW-GAN [118], can be utilized to generate EV charging behaviors. For residential applications, the GAN-based models [108], [119] are employed for data augmentations, which is important for other applications, e.g., optimal scheduling of an EV at home [108]. Researchers also use the Transformer model [26] to solve the EVRP, which is different from traditional VRP with new EV constraints.

C. Layer 3: Smart Grid with EV

The proliferation of EVs leads to increased stochastic power demands from the grid, accelerating grid asset deterioration and complicating power system operations. The investigation of EV charging load profiles is crucial for understanding future grid states to enable large-scale transportation electrification. However, the issues persist regarding the quantity and quality of data [14], [16], and the challenges involved in predicting EV charging loads [15], [27], [38], [136], [137] as well as the urge of understanding user experience [24].

1) Data Quantity and Quality: The digitalization and widespread deployment of charging infrastructure offers a great opportunity to gather real EV charging data, yet it is hampered by equipment failures, data collection errors, and intentional damage, resulting in missing values and outliers [138]. Given insufficient data accumulation in newly built charging facilities, the ML models can be biased with the flawed datasets [139], which are not necessary to be small-scale. Such bias and inaccuracy pose challenges to the scheduling and optimization of the grid, whether centralized or distributed. As such, one important usage of GenAI is to enhance the EV and grid datasets to improve the system performance such as load forecasting and balancing. Several GenAI algorithms, such as VAE [14] and GAN [16], [38], [140], have been explored and we present the technical details of them below.

a) VAE-based Approach: VAE's adoption is mainly for generating stochastic scenarios for EV load profiles. In [14], a VAE model is designed for such usage to capture the time-varying and dynamic nature of EV loads effectively. The paper considers five different EV load profiles, for fully battery-based and hybrid-based EVs with and without demand responses [14]. The profiles are measured at 10-minute intervals, resulting in 144 data points for each profile per day. The model uses the historical profiles as input, and accordingly generates new profiles that encapsulate the critical characteristics of the profiles. For future work, the proposed method holds the potential to be integrated with load forecasting models [27] as a data augmentation tool to address the data scarcity issue.

b) GAN-based Approaches: Similar to VAE, GAN has also been applied for data augmentation and furthermore, load forecasting. In [16], GAN is used to generate EV charging load data first, with which, load forecasting is performed. The data augmentation model is called GRU-GAN. The model uses the transactional data from various EV charging stations within a 35 kV power distribution zone spanning from July 1 to August 31 in 2019 [16]. Due to different realistic restrictions, the time-series data is incomplete. As such, the data is pre-processed to

TABLE III
SUMMARY OF GENAI FOR IOEV IN LAYER 2.
●: GENAI METHODS; ✓: PROS OF THE METHODS; ✗: CONS OF THE METHODS.

Applications	Reference	Techniques	Pros & Cons
EV Charging Behavior Data Augmentation	[17]	GAN-DNN	● A framework to predict EV charging behaviors, e.g., plug-out hour and required energy. ✓ Address the cold-start forecasting problem when limited training data is available. ✗ Need clustering algorithms and multi-source datasets for future work.
	[118]	CW-GAN	● A GAN-based model to generate charging behavior data for EVs. ✓ Learn from small samples, expanding dataset while preserving original probability distribution. ✗ Dependency on quality of conditional labels during the training.
Smart Home Data Augmentation	[108]	VAE-GAN with Q-learning	● A data generation scheme using a VAE-GAN combined with Q-learning-based home EMS. ✓ Learn various data distributions in a smart home and generate realistic samples. ✗ Reliance on quality and diversity of training data.
Residential EV Load Generation	[119]	GAN	● A GAN-based tool for semi-randomly generated data for EV load, PV, and household demand. ✓ Generate profiles that keep both magnitude of profile and behavioral consistency over time. ✗ Need other countries' datasets to extend the capability of the HEDGE tool.
EVRP	[26]	Transformer-based DRL	● A Transformer-based DRL method for solving electric vehicle routing problem (EVRP). ✓ Consider EV's energy and charging constraints. ✓ Reduced energy consumption of EV compared with exact algorithm, ACO, ALNS, and AM. ✗ Lack model testing with real-world data. ✗ Need to consider data diversity, e.g., various EV types.

simultaneously identify and handle missing values and outliers using GRU-GAN, and this process is commonly referred to as data imputation. Finally, the generated high-quality data is used for training the Mogrifier LSTM network for short-term EV load forecasting.

The comparison results show that GRU-GAN outperforms conventional imputation techniques. Compared to mean imputation [141], GRU-GAN can reduce the errors by 15%. GRU-GAN's performance improvement is even more significant when comparing with piecewise linear [142] and k-nearest neighbor (KNN) [143] imputation, achieving 43% and 19% lower errors, respectively. Such enhanced imputation accuracy is attributed to GRU-GAN's ability to capture intricate high-level data representations. The potential improvement of the solution is mainly in the forecasting stage. The forecasting of certain intervals is especially challenging, e.g., peak periods. The GenAI models, e.g., transformer-based [15], [27], may serve as viable substitutes for Mogrifier LSTM.

2) EV Charging Load Prediction: Understanding the pattern of EV load is critical for the grid scheduling. Conventionally, the load prediction is formulated as a time series forecasting problem, which can be handled by different ML algorithms such as the GenAI-based ones.

a) Traditional ML-based Approaches: There exist several load forecasting approaches driven by traditional ML, e.g., SVR [144], ANN [145], RNN [146], and LSTM [147]. These algorithms have demonstrated the potential of using ML for load forecasting. However, they either fell short in capturing the nonlinear characteristics inherent in EV load series or faced challenges in modeling long-term dependencies that are common in real-world forecasting applications [107]. In the application layer, these algorithms focus on the load patterns in time scale where spatial load distribution is overlooked. A recent work [148] developed a spatial–temporal forecasting method for load forecasting, but the method cannot model the correlation among spatial regions.

To address the aforementioned challenges, the Transformer models [15], [27] and the GAN models [38], [136], [137] have been introduced recently. The former utilizes attention mechanisms to capture the long-range dependencies and intricate patterns of the time-series data. This allows the model to focus on relevant parts of the input sequence and facilitate more accurate and robust load forecasting results. The latter helps to generate highly realistic EV charging scenarios by learning the underlying data distribution over a latent space with conditions. A wide range of plausible scenarios generated by GAN can provide a comprehensive view of potential future loads. We present the technical details as below.

b) Transformer-based Approach: The transformer-based model has been widely used for time-series forecasting because it addresses the limitations of LSTMs and effectively captures long-term dependencies [15]. The integration of a transformer-based architecture with the probabilistic forecasting technique in the temporal latent auto-encoder exhibited significantly enhanced efficacy in the context of time series forecasting endeavors [149]. Nevertheless, the vanilla transformer model exhibits significant time and memory overheads due to the quadratic computation complexity of self-attention, and imposes constraints on the maximum allowable input sequence length as a result of the accumulation of encoder and decoder layers [150]. The Informer [150] tackles these challenges by substituting the conventional self-attention computation in the standard transformer with a ProbSparse self-attention framework [150], while also introducing the self-attention distilling mechanism. However, it is limited to point forecasting which refers to the prediction of a single/specific value for a future variable or event [151].

The authors in [27] proposed a Probformer model to address the challenge of capturing long-term dependencies within charging load sequences and to facilitate the generation of probabilistic load forecasts. The model with multi-head ProbSparse self-attention [150] was an adaptation of the Transformer-based Informer framework [150]. Subsequently, the MetaProbformer [27] was developed by integrating the Probformer with a meta-learning algorithm, Reptile [152], which tackled the issue of charging stations with scarce historical data. Four datasets comprising real-world EV charging loads, sourced from various charging stations over distinct time periods, were utilized in the experiments. Each dataset details the start and end times, along with the total energy consumed (in kWh) for each charging event. While MetaProbformer is capable in load forecasting with limited historical data, it faces difficulties for new charging stations where data collection takes time. Hence, an exploration of methods for enhancing the model generalizability within a few-shot or zero-shot framework is necessary.

c) GAN-based Approach: Recently, the GAN-based techniques for charging scenario generation have begun to emerge [38], [136], [137]. For example, the authors in [136] proposed WGAN-GP to tackle spatial–temporal uncertainty in EV charging load analysis. This approach explored load dynamics and generated scenarios without uniform probability assumptions across charging stations. The undisclosed power grid structure limited data access to the EV charging load. To model the impact of these stations on the distribution network, the data from 32 charging stations in the Zhejiang region were allocated to nodes in the IEEE 33-node distribution network system [136]. The forecasting target was the EV charging load at each node in the network. Nevertheless, the comparative analysis of the WGAN-GP against contemporary GenAI approaches in the context of EV charging scenario generation remains lacking.

A Copula generative adversarial network (CopulaGAN) model combining Copula transformation with GANs was developed in [137]. Slightly different from the traditional load forecasting approaches focusing solely on EV charging load curves, the CopulaGAN captured uncertainties in EV charging sessions, including energy delivered, arrival, and departure times. Subsequently, it was used for day-ahead optimal EV scheduling. The charging session dataset used in this study was obtained from the Caltech parking lots [153]. Each site's database records the EV connection time, charging completion time, energy consumption, and vehicle departure time from the parking lot. The power market dataset including the load forecast, generation forecast, day-ahead electricity market price, and balancing energy market, was collected from the German electricity market [154]. The simulation of EV charging sessions and the prediction of day-ahead market prices were carried out for 48 hours. The outcomes of their investigation demonstrated a high degree of correspondence between the generated EV charging session data and the actual dataset, with an approximate match rate exceeding 90% [137]. This validation underscored the effectiveness of CopulaGAN in accurately capturing the inherent structure and distributional attributes of the data. However, the proposed optimal scheduling of EVs charging faces limitations due to data inadequacy and assumptions about SoC and charging rates. The comparative analysis of CopulaGAN and other GenAI methods (e.g., GAN-DNN [17] and DiffCharge [61]) shall be investigated.

3) Consumer Data and Sentiment Analysis: The transportation sector significantly contributes to global greenhouse gas (GHG) emissions [155] and impacts public health [156]. Government policies increasingly favor EVs to reduce GHG emissions. However, the analysts underutilized consumer data, particularly unstructured EV data, in decisions about charging infrastructure. Previous research employing sentiment analysis suggested prevalent negative user experiences in EV charging station reviews, yet lacked specific causal extraction [157]. Thus, the multi-label topic classification is crucial for understanding user interaction behaviors in electric mobility. Hence, the authors in [24] utilized the Transformer-based models, Bidirectional encoder representations from Transformers (BERT) [158] and XLNet [159], for the multi-label topic classification in the domain of EV charging reviews. The method proposed herein aimed to expedite the evaluation of research through automated means, utilizing extensive consumer data to assess performance and analyze regional policies. The study utilized data derived from 12,720 charging station locations across the United States, comprising 127,257 English-language consumer reviews written by 29,532 EV drivers over a four-year period from 2011 to 2015 [160]. An important future research is to enhance model interpretability through methods such as the use of rationales [161], influence functions [162], and sequence tagging approaches [163] to well understand consumers.

The GenAI applications for IoEV in layer 3 are summarized in Table IV. We observed that the GAN models were often used for data augmentation [38], [140], handling missing values and outliers of the input data [16], and generating EV charging sessions/scenarios [136], [137]. Besides GAN, VAE can be utilized for generating stochastic scenarios for EV load profiles. The Transformer-based models were usually utilized for the multi-label topic classification in the domain of EV charging reviews [24], and EV load forecasting [15], [27].

D. Layer 4: Security

Security is a critical aspect that we cannot bypass for many cyber-physical systems. In the following parts, we introduce the background of security research for electric mobility and then present related GenAI research works.

1) Background: Security's impact on electric mobility encompasses several aspects. The first one is due to the ever-increasing adoption of the ML models. For example, there has been considerable utilization of DRL algorithms within the context of EV charging schedules, aiming at acquiring optimal charging strategies for users. Such algorithms include soft actor-critic (SAC) [164], deep deterministic policy gradient (DDPG) [165], safe deep reinforcement learning (SDRL) [166], multi-agent deep reinforcement learning (MADRL) [167], proximal policy optimization (PPO) [168], etc. Same as many ML algorithms, DRL suffers from potential adversarial attacks [169]. The adversarial examples involve maliciously altered inputs to deceive the ML models, causing erroneous outputs [170]. Some methods are shown to be effective in generating adversarial data, e.g., fast gradient sign method (FGSM) [171], basic iterative method (BIM) [172], and DeepFool [173]. The ramifications of such attacks could be extensive, ranging from

TABLE IV
SUMMARY OF GENAI FOR IOEV IN LAYER 3.
●: GENAI METHODS; ✓: PROS OF THE METHODS; ✗: CONS OF THE METHODS.

Applications	Reference	Techniques	Pros & Cons
EV Charging Load Profile Data Augmentation	[14]	VAE	● A framework for generating and enhancing EV load profiles. ✓ Ensure consistency in power consumption between generated and original profiles over time. ✓ Capture temporal correlations, probability distributions, and volatility of original load profiles. ✗ Challenge in balancing reconstruction loss and KL divergence during training.
	[16]	GRU-GAN with Mogrifier LSTM	● A load data generation model for missing values and outliers of the input data. ✓ Generate high-quality data closely resembling real data. ✓ Outperform traditional mean [141], piecewise linear [142], and KNN [143] imputation methods ✗ Sub-optimal forecasting performance during peak and plateau periods.
Analysis of EV Consumer	[24]	BERT, XLNet	● Transformer-based models for analyzing EV charging reviews. ✓ Outperform traditional LSTM and CNN models in terms of accuracy and F1 scores. ✗ Lack of interpretability.
EV Load Forecasting	[27]	MetaProbformer	● A Transformer-based model for EV charging load forecasting. ✓ Perform well in point and probabilistic forecasting for the load at EV charging station. ✓ Competitive performance in probabilistic forecasting for both short-term and long-term tasks. ✓ Adaptable to seen and unseen scenarios. ✗ Need an extension to multivariate forecasting for future work. ✗ Need to improve model's generalization capabilities within a few-shot/zero-shot framework.
Charging Scenarios Generation	[136]	WGAN-GP	● A GAN-based model for generating EV charging scenarios at charging stations. ✓ High degree of spatial correlation similarity in terms of SSIM and FSIM indexes. ✗ Lack of comparative analysis for different GenAI methods in EV charging scenario generation.
	[137]	CopulaGAN	● A GAN-based model for capturing and modelling the uncertainties in EV charging sessions. ✓ Generate EV charging sessions data which closely align with the real dataset. ✗ Real conditions not fully captured with simplified assumptions of SoC and charging rate. ✗ Need a comparison study with other GenAI approaches for scenarios generation.

increasing charging expenses and fluctuations in grid loads to jeopardizing the stability of power grids [174].

Besides adversarial attacks, security is also a concern in power grids with EV penetration. Specifically, the integration of power systems with electrified transportation networks poses new challenges concerning the reliability and resilience of charging infrastructure [175]. According to [176], there is an increase in the occurrence of power system attacks targeting customer satisfaction levels in charging services. Deep learning algorithms such as deep belief network (DBN) [177] and feed-forward neural network (FNN) [178] have been used for detecting cyber-physical attacks in power systems because of their advanced feature extraction capabilities. The algorithms achieve high detection rates, yet they overlook the extraction of the important spatial relationships inherent in the data, as they disregard the topological grid attributes [179].

The connection between EVs and the grid is largely controlled by the EV's supply equipment, that is responsible for managing and maintaining the charging operations. It also facilitates communication among cloud services, payment providers, EVs, battery management systems, and other relevant entities to enable efficient and intelligent charging [180]. However, the connection is often associated with potential vulnerabilities such as DoS attacks, FDIAs, and spoofing [181]. A popular connection standard is CAN, which is a bus protocol for communication within vehicles. CAN has many advantages such as reduced wiring expenses, minimal weight, and simplified design [182]. However, CAN has security vulnerabilities also, including inadequate authentication mechanisms, susceptibility to multiple attack vectors, and the absence of encryption technologies [183]. The security vulnerabilities have been studied with several methods such as local outlier factor (LOF), one-class support vector machines (OCSVM), and principal component analysis (PCA) [184]. Nevertheless, these methods often are sensitive to noise, incapable to handle high-dimensional data, and fail to handle the intricate dynamics of systems.

2) Adversarial Attacks: To investigate adversarial attacks against DRL in the EV charging process, a GAN-based approach namely RL-AdvGAN was introduced in [39]. RL-AdvGAN could support the adversary to leverage the stolen data to engage in behavior cloning, thereby constructing an adversarial policy network to mimic the user's policy. In this research, the EV charging environment was set up using real-world data from California independent system operator (ISO) [185]. The dataset spanned a duration of approximately 2 years with data recorded hourly, consisting of the electricity prices and the grid load information. The paper assumed that the commuting behavior of EV users follows the normal distributions with parameters of arrival time, departure time, and battery SoC. Four types of DRL algorithms, i.e., deep Q-networks (DQN) [186], DDPG [187], PPO [188], and SAC [189] were tested under the adversarial attacks generated by FGSM [190] and RL-AdvGAN [39]. According to the results, FGSM attack is effective, and the algorithms, i.e., DQN, DDPG, PPO, and SAC, only managed to maintain a normal SoC range of the battery in 6%, 0%, 28%, and 8%, of the time, respectively [39]. The proposed RL-AdvGAN attack [39] was even more effective, where all the tested DRL algorithms failed to maintain the normal range of SoC all the time during the attack. As such the proposed RL-AdvGAN posed a significant risk to the security and stability of the battery system. Future work includes further exploring the threat and damage of GenAI methods for adversarial attacks with different network architectures and developing effective attack detection methods.

In [191], the authors considered both the generation and detection of adversarial attacks against vehicle-to-microgrid (V2M) systems and proposed GAN-based models. The research modeled a system where the adversaries sought to manipulate the ML classifier at the network edge, causing it to incorrectly classify the energy requests received from microgrid users, e.g. EVs' charging/discharging requests. The FGSM and conditional generative adversarial network (CGAN) models were introduced to serve as attackers in generating adversarial instances. On the adversarial detection side, a GAN-based adversarial training framework was introduced to create adversarial training instances. The instances are used to train SVM classifiers to detect these adversarial attacks. The simulation was set up using the iHomeLab RAPT dataset [192] consisting of electrical power consumption and generation of 5 Swiss households, and the supplementary dataset with power generation from the batteries and wind farm from their previous work [193]. The results indicate that their proposed method outperforms the density-based spatial clustering of applications with noise (DBSCAN) algorithm, leading to an improvement in the adversarial detection rate ranging from 13.2% to 25.6% [191]. For future work, the investigation of the impact of limited resources on the adversarial detection rate is necessary since the classifier is designed to be deployed at the network edge.

3) Detection of Attacks: The authors in [45] extended the study of potential attacks [39] and developed a graph autoencoder (GAE) model for detecting the FDIAs in power systems. The model leverages the correlations between power system data (e.g., active and reactive power measurements) and transportation data (e.g., hourly traffic volume) to enhance charging satisfaction. In the event of FDIAs, the malicious entities can manipulate power measurements to simulate additive, deductive, and camouflage attacks. The level of satisfaction for customers can diminish due to insufficient power supply at the charging stations, e.g., blocked charging requests or extended charging times. The proposed method was tested via simulation of the Texas power grid consisting of 2,000 buses and 360 active charging stations. The load data was sourced from electric reliability council of Texas (ERCOT) which manages the distribution of electric power to more than 27 million customers in Texas [194]. Compared to the benchmark graph CNN model [179], the proposed GAE model improved the detection accuracy by about 15% in various FDIAs scenarios on vulnerable nodes when 30% of data was under attack. The results show that the proposed GAE model outperforms the state-of-the-art detector in various attacks. For example, compared to the benchmark graph CNN model [179], the proposed GAE model can improve the detection accuracy by around 14.4% to 14.8% in the various FDIAs scenarios e.g, additive attacks, deductive attacks, and combined attacks [45]. Future research may focus on real-time model updates and decision-making in dynamic operational contexts of power and transportation networks.

Based on the above introduced detection model, the researchers in [47] developed an interpretable anomaly detection system, referred to as RX-ADS. The system was designed to identify intrusions within the CAN protocol communications of the EVs with active charging connections. Besides, the ResNet Autoencoder was employed to learn normal behavior from data and detect anomalies based on reconstruction errors. The performance of the system was investigated with two publicly available datasets of EV CAN protocol, including offset ratio and time interval based intrusion detection system (OTIDS) [195] and Car-Hacking [196]. The results showed that the system outperformed a GAN-based intrusion detection system [197] by 4% under DoS attacks. Nevertheless, the proposed method requires a substantial dataset that accurately reflects the system's typical normal behavior. If new behaviors arise that deviate from the established norms, the model shall be updated. Otherwise, such deviations might be incorrectly classified as anomalies, leading to an increase in false positives.

The same research group of [47] developed a ResNet AE-based approach for unsupervised physical anomaly detection with high resilience in [198], specifically in EV charging stations without labeled data. The experiments were conducted using data under normal and physical attack scenarios from the Idaho National Laboratory's EV charging station system testbed. The proposed ResNet Autoencoder based approach was compared with two benchmark algorithms: LOF and OCSVM. The results in terms of F1 score show that the proposed method is able to improve the detection performances of LOF algorithm by around 20.1% and OCSVM algorithm by about 3.7% [198]. For further work, enhancing the proposed anomaly detection framework can be achieved by integrating cyber security related scenarios pertinent to EV charging systems.

In addition to detecting FDIAs [45], [199], the authors in [46] proposed a ResNet AE-based anomaly detection framework which consists of cyber anomaly detection (Cy-ADS) and physical anomaly detection (Phy-ADS) for the cyber and physical data streams, respectively. The former has the capability to monitor and analyze packet data in real-time for the detection of cyber anomalous behaviors within EV charging stations. Meanwhile, the latter is designed to capture and analyze real-time physical sensor data (e.g., voltage, current, power, and thermal measurements) to identify physical anomalous behavior. Their results show that the proposed Cy-ADS for detection of cyber attack and Phy-ADS for physical anomaly detection outperforms the LSTM AE method by about 18% and 15% respectively, VAE approach by around 3% and 5% respectively [46]. However, the proposed method is subject to the quality of the training data and model retraining is needed if system behaviors change over time.

The GenAI research for EV and IoEV security in layer 4 is summarized in Table V. As seen from the table, AE-based and GAN-based models are commonly found in recent research for security. Among them, GAN-based models are mainly used for generating and identifying adversarial attacks, e.g., RL-AdvGAN [39] and CGAN [191]. Whereas the AE-based models [45]–[47], [198] are good at detecting FDIAs, DoS,

TABLE V
SUMMARY OF GENAI FOR IOEV IN LAYER 4.
●: GENAI METHODS; ✓: PROS OF THE METHODS; ✗: CONS OF THE METHODS.

Applications	Reference	Techniques	Pros & Cons
Adversarial Attacks	[39]	RL-AdvGAN	● A GAN-based model for generating adversarial attacks against DRL algorithms. ✓ Effectively reduce the performance of common DRL algorithms for optimal EV charging. ✓ Higher security threat than FGSM [171]. ✗ Need to improve training stability with model refinement. ✗ Need to propose methods to prevent adversarial attacks.
	[191]	GAN and CGAN	● A GAN-based detection framework for identifying the adversarial attacks. ✓ Higher adversarial detection rate compared with traditional DBSCAN algorithm [191]. ✗ Incompatible with network edge services due to resource constraints.
False Data Injection Attacks	[45]	Graph Autoencoder	● An AE-based detection scheme for identifying FDIAs. ✓ Performance improvement of 15-25% compared with SVM, FNN, CNN, and LSTM [45]. ✗ Need to update and deploy the offline-trained model to real-time applications for future work.
Cyber and Physical Attacks	[46]	ResNet Autoencoder (AE)	● An AE-based anomaly detection framework for identifying both cyber and physical attacks. ✓ Capable of detecting both simple and complex cyber-physical attack scenarios. ✓ Low training and inference time, making it suitable for real-time applications. ✗ Sensitive to the setting of threshold values.
DoS and Fuzzy Attacks	[47]	ResNet Autoencoder (AE)	● An AE-based method for detecting intrusions in CAN communication protocol for EV charging. ✓ Competitive results compared with the GAN-based approach [197]. ✗ Require a large amount of data that reflects the normal behavior of the system. ✗ Need model update and retraining for new behaviors in the future.
Physical Attacks	[198]	ResNet Autoencoder (AE)	● An AE-based approach for physical anomaly detection. ✓ Outperform two benchmark algorithms including LOF and OCSVM [198]. ✗ Not capable of detecting cyber attacks.

fuzzy, and physical attacks, and serve as viable alternatives to traditional methods such as LOF and OCSVM.

E. Studies Across Multiple Layers

Besides studies on individual layer, some studies were carried out on multiple IoEV layers, e.g., [40], [61]. Understanding user behavior [200], coupling characteristics [201] among user behavior, road networks [202], and EVs are crucial for accurate demand prediction, but it remains challenging due to various factors such as time and SoC. Nevertheless, research on precise mathematical models for charging and discharging strategies is lacking due to the complexity of influencing factors.

a) Generation of EV Charging Scenarios: In [61], a diffusion model namely DiffCharge was developed to generate EV charging scenarios. The generated scenarios could be divided into battery-level (e.g., charging current in Ampere) and station-level (e.g., charging load in kW). The traditional ML e.g., GMM could be utilized to estimate the daily EV charging load profiles, but it faced a challenge in accurately capturing temporal dynamics across various time-series EV charging data. On the other hand, DiffCharge as one of the diffusion-based approaches is capable of deriving the challenging uncertainties associated with charging and producing a range of charging load profiles characterized by realistic and unique temporal features. On top of DDPM [52], the DiffCharge framework consists of LSTM, broadcast, multi-head self-attention, and 1D-CNN. DiffCharge was trained by using the ACN-Data [153] dataset which comprises real-world charging data of individual EVs in California. The data attributes including connection time, done charging time, kWh delivered, and charging current in Ampere were considered for training the model to generate EV charging curves. The daily charging load profile could be aggregated and extracted from the ACN-Data dataset. Then, it was integrated with arrival/departure time and actual scheduled energy for training the model to generate the EV load profile at the station. The proposed method was compared to the baseline models: GMM [203], VAEGAN [204], and TimeGAN [205]. Their results showed that the proposed DiffCharge could generate realistic charging curves and it outperformed the baseline models in terms of marginal score, discriminative score, and tail score. For example, the marginal score of DiffCharge was improved by 91% from GMM, by 35% from VAEGAN, and by 32% from TimeGAN [61]. However, the control capability of the DiffCharge is restricted, thereby limiting the model's ability to generate tailored charging scenarios for varying conditions such as initial SoC, types of batteries, and station congestion.

b) Generation of Regional Electric Vehicle (EV) Charging Demand: Similar to [61], [40] explored the EV charging demand from perspectives encompassing both battery and station levels. However, a notable distinction of [40] lies in the emphasis placed on the spatial–temporal distribution of EV charging demand in the region. The authors in [40] addressed the challenges of predicting spatial–temporal EV charging demand at both battery and station levels by proposing a deep learning framework that consists of methodologies such as GAIL, PPO, and XGBoost. The paper classified strategies concerning user charging and discharging into three categories: driving policies, travel target mileage policies, and charging duration selection policies. Following this categorization, the research leveraged GAIL to obtain insights into these delineated policies i.e., GAIL was used as a strategy learning model. Then, employing the PPO, the strategy learning model underwent

TABLE VI
SUMMARY OF GENAI FOR IOEV ACROSS MULTIPLE LAYERS.
●: GENAI METHODS; ✓: PROS OF THE METHODS; ✗: CONS OF THE METHODS.

Applications	Reference	Techniques	Pros & Cons
Scenarios Generation	[61]	DiffCharge	● A diffusion-based model for generating EV charging scenarios for both battery-level (Layer 1) and station-level (Layer 3). ✓ Outperform GMM [203], VAEGAN [204], and TimeGAN [205] in charging scenarios generation. ✗ Limited control over diverse conditions, e.g., initial SoC, battery types, and station congestion.
SoC and Load Forecasting	[40]	GAIL, PPO, and XGBoost [40]	● A GAN-based model, GAIL, as a strategy learning model assisting DRL algorithm. ✓ Good SoC prediction with low MAE and RMSE values. ✗ Overlook predictions at the individual charging station level.

optimization utilizing an SoC forecasted through the XGBoost algorithm. The data utilized in this study were acquired from the Shanghai New Energy Electric Vehicle Monitoring Center [206], pertaining to a cohort of 1,000 EVs subjected to testing over the course of one month. The data attributes included speed, acceleration, SoC, temperature, longitudes, and latitudes. The data points were sampled every 10 seconds. The output variables were the 24-hour SoC predictions for individual vehicles and the forecast of regional spatial–temporal charging demand. Four categories of vehicle's SoC, encompassing logistics vehicles, taxis, buses, and private cars, underwent predictive analysis. The assessment criteria exhibited a variability spanning approximately 1.66% to 3.15% for MAE and 2.32% to 4.44% for RMSE [40]. Future research will refine the findings of this study. Considering additional factors such as road conditions and user demographics will enhance the accuracy of EV SoC predictions.

Table VI summarizes GenAI implementations for IoEV applications across multiple layers. We discovered that a few papers considered both layer 1 and layer 3 applications. For example, DiffCharge [61] is able to generate EV charging scenarios for both battery level and EV charging station level. In contrast, in [40], SoC prediction and regional charging load forecasting were completed by a framework with GAIL.

F. Discussion

We have presented technical details of GenAI's usage in IoEV in different layers in various aspects and we present our discussions about GenAI's performance below.

1) Key Features for GenAI's Competitiveness: We may notice that GenAI is not the first ML technology but manages to outperform traditional ML algorithms for many discussed applications and tasks. We summarize and highlight the key features that drive GenAI's competitiveness in IoEV.

a) Data Generation and Robustness: Data scarcity has been among the toughest challenges for traditional ML, e.g., in layers 2 and 3 for predicting supply and demand. GenAI algorithms such as GAN and VAE help generate realistic synthetic data by learning the underlying distribution of the limited raw data. Such GenAI algorithms demonstrate a high level of noise control during data generation and this enhances the stability of anomaly detection, predictive modeling, etc.

b) Advanced Pattern Recognition: Compared to traditional ML, GenAI exhibits improved pattern recognition performance, e.g., for modeling and prediction tasks. This is critical for IoEV-related applications that involve complex (e.g., high-dimensional and nonlinear) patterns and dependencies. Two representative techniques are GDM and transformer. The former captures inherent distribution characteristics and models complex relationships to understand system dynamics. The latter extracts multi-scale features and exploits long-term dependencies well with its self-attention mechanism.

With these strengths, GenAI becomes a versatile and competitive solution, and we foresee GenAI advancements and its increased usage in IoEV applications in the future.

2) Selection of GenAI Algorithms: GenAI has been used for various IoEV applications and the optimal GenAI performance in part depends on the choice of GenAI algorithms. Each GenAI algorithm owns unique characteristics that can influence the algorithm's performance in solving different problems. For example, GAN has been shown to be a popular technology for data augmentation by generating high-fidelity synthetic data (e.g., charging behavior). When robustness is a key concern, GAN becomes less competitive due to its ineffectiveness in capturing data diversity well and mode collapse issue. VAE is capable of probabilistic data generation and particularly useful in generating stochastic scenarios. However, its reliance on the Gaussian latent space sacrifices data details sometimes. The diffusion model outperforms GAN for producing high-resolution outputs by iteratively refining noisy data, but its demand for computing resources is significant. Overall, we urge a comprehensive evaluation of different GenAI algorithms and the selection of suitable algorithms to meet specific requirements and constraints in different tasks.

IV. TECHNICAL REVIEWS: DATASET

Data is highly important for GenAI for model training, system customization, performance improvement, and so on. In this section, we provide a summary of the available public dataset in the domain of GenAI-based electric mobility applications. We summarize the datasets in Table VII and describe them in detail as below.

Layer 1: In the battery layer, the dataset from NSMC-EV [74] and the unlabeled dataset on multiple faults and failure scenarios [76] are used by GRU-VAE [51] and BERTtery [25], respectively, for the anomaly detection. NSMC-EV serves as China's national big data platform for EVs, offering extensive real-time online data on EVs utilized in public transportation.

TABLE VII
SUMMARY OF DATASET USED IN GENAI FOR IOEV AT DIFFERENT LAYERS.

Applications	Dataset	Properties	Dataset Size	Algorithms
		Layer 1		
Anomaly Detection	NSMC-EV [74]	13-dimensional time series, e.g., vehicle speed, charging state, insulation resistance, and SoC.	NSMC-EV platform for over three million EVs	GRU-VAE [51]
	Faults and Failure [76]	Multiple scenarios, e.g., short circuit and thermal runaway. Time series of voltage, current, etc.	316 NCM battery cells [25]	BERTtery [25]
SoC Estimation	EV [88]	Time series, e.g., vehicle speed, voltage, cell temperature, motor controller voltage, and SoC.	Driving data from five identical vehicles over a year [88]	TS-WGAN [65]
	Li-ion Battery [89]	Time series of recorded variables, e.g., cell voltage, current, battery temperature, and ampere-hours.	Various drive cycles, including US06, UDDS, and LA92.	TS-WGAN [65]
SoH Estimation	LFP Battery [105]	Rated capacity, number of cells, charging current, discharging current, cut-off voltage, etc.	4 cells; cycles: 1062, 1266, 1114, and 1047. [60]	DDPM [60]
		Layer 2		
EV Charging Behaviors	EA Technology [115]	Residential charging events, e.g., date time, arrival hour, plug-out hour, and required energy	Charging behaviors of over 200 participants in 2014 and 2015.	GAN [17]
Smart Home	iHomeLab PART [123]	Energy consumption of households and specific appliances as well as PV generation.	Five houses in the Lucerne region, Switzerland.	VAE-GAN [108], GANs [191]
Residential EV Load	CLNR TC1a [120]	UK electricity customers' electricity consumption measured by British Gas's smart meter;	Up to 8,000 customers for the year 2011.	GAN [119]
	CLNR TC5 [121]	TC5 [121]: including customers' energy use and solar PV performance.	Part of the project involving over 12,000 consumers.	GAN [119]
		Layer 3		
EV Load Forecasting	City of Boulder [207]	Charging load records, e.g., address, arrival and departure time, type of the plug, data, and energy.	4 years; 25 public charging stations in Boulder.	Transformer [15]
Scenarios Generation	ACN Data [153]	Time of EV connection, charging time, amount of energy received by EV, and time of leaving.	Over 30,000 charging sessions; growing daily [153].	CopulaGAN [137]; DiffCharge [61]
	Belgian Elia Group [208]	PV and load power in the microgrid.	2 months of 15-minute interval data.	GAN [38]
		Layer 4		
Adversarial Attacks	California OASIS [185]	Electricity prices and the grid load information from OASIS site.	Continuously updating [185]	RL-AdvGAN [39]
FDIAs	ERCOT's Data [194]	Load profiles from the grid.	Continuously updating [194].	GAE [45]
DoS and Fuzzy Attacks	OTIDS dataset [195]	States of DoS attacks, fuzzy attacks, impersonation attacks, and attack-free conditions.	Attacks: 657K DoS, 592K fuzzy, and 995K impersonation.	RX-ADS [47]
	Car-Hacking [196]	Attack types, e.g., DoS, fuzzy, drive gear spoofing, and revolutions per minute (RPM) gauge spoofing.	Attacks: 3.6M DoS, 3.8M fuzzy, 4.4M spoofing drive gear, etc.	RX-ADS [47]
		Multiple Layers		
SoC and Load Forecasting	Shanghai New Energy [206]	Data sampled for speed, acceleration, SoC, temperature, longitudes, and latitudes.	Continuously updating [206]	GAIL, PPO, XGBoost [40]

The dataset used for EV's SoC estimation can be found in [88], [89], which provides the battery's voltage, current, temperature, SoC, etc. EV dataset from [88] consists of driving data of 5 identical vehicles over a year where each vehicle's data for charging and discharging events is sampled at 10-second intervals. An LPF battery dataset from [105] is used for EV's SoH estimation. It consists of 4 cells with a number of charge-discharge cycles over a thousand times for each cell.

Layer 2: For the EV layer in Table VII, data from EA Technology [115] consists of residential EV charging events with charging behaviors of over 200 participants observed from February 2014 to November 2015. EA Technology offers specialized asset management solutions for electrical asset owners and operators worldwide [115]. This data is used by [17] for the study of EV charging behaviors. The iHomeLab PART dataset [123] includes the energy consumption of households and specific appliances as well as PV generation. The data are collected from 5 houses in the Lucerne region, Switzerland, and recorded over 1.5 to 3.5 years. It is used by [108] for data augmentation in the smart home applications, and by [191] for the study of adversarial attacks. CLNR's dataset TC1a [120] and TC5 [121] consist of UK electricity customers' overall electricity consumption and customers' energy use and solar PV performance, respectively. These datasets are used in GAN [119] for generating residential EV load. CLNR is a project supported by the Ofgem's Low Carbon Network Fund, which aimed to facilitate UK's low carbon energy sector [209].

Layer 3: This layer is about the interaction between EVs and power grid. The City of Boulder Open Data Hub [207] is used by a Transformer-based model for EV load forecasting. It provides EV charging load records spanning about 4 years collected from 25 public charging stations in Boulder, Colorado where the charging stations are equipped with 22 kW-rated connectors. Datasets from [208] and [153] are employed for scenario generations. The Belgian grid dataset from Elia Group [208] includes PV and load in the microgrid with about two months of data. ACN data [153] include the time of EV connection, done charging time, amount of energy received by EV, and time of EV leaving the parking lot. This dataset contains charging sessions at Caltech parking lots in California, and is continuously growing every day [153].

Layer 4: For the security layer, California ISO OASIS site [185] provides the continuously updating dataset for electricity prices and the grid load information where OASIS offers real-time data related to the ISO transmission system and its market. Data from [185] is used by RL-AdvGAN [39] for the study of adversarial attacks. The load profiles from the grid can be found in ERCOT [194]. However, it requires an IP address from the U.S. to access data from ERCOT. This dataset is used by GAE [45] for the investigation of FDIAs. OTIDS dataset [195] and Car-Hacking dataset [196] are employed in [47] for the detection of DoS, fuzzy, and impersonation attacks.

Multiple Layers: Moreover, datasets can be used in multiple layers. For example, the Shanghai New Energy EV Monitoring Center [206] is a data sharing and cooperation platform with extensive operational data on new energy vehicles in Shanghai, and provides the EV's data such as speed, acceleration, battery status, and location. The data is continuously updated and used in [40] for the prediction of EV's SoC and load.

V. Future Directions

GenAI has demonstrated its potential in the IoEV, and there are still areas to be improved. This section outlines several directions to unlock new opportunities in the IoEV ecosystem.

A. Improving Existing Solutions

While GenAI models have been developed for electric mobility, they are not perfect.

1) Up-to-date Models: One challenge is to make sure models are updated and data plays an important role. Data patterns may evolve over time due to the dynamic nature of IoEV environments. This requires the models to have continuous learning capability where new information should be adapted without forgetting previous knowledge. Integrating continual learning into the GenAI system can keep it remain effective over the lifespan of an EV.

2) Hallucination: Another challenge could be solving the hallucination issues in GenAI models. The hallucination refers to ML generating outputs that are plausible but incorrect. For future work, developing methods that can detect and mitigate hallucinations in GenAI is very important, especially for high-stakes applications such as EV routing and EV's battery management system. The hybrid systems combining GenAI with the traditional rule-based approach or the new architectures that can verify the correctness of generated content/data could be developed to avoid hallucination issues.

3) Transfer Learning: Additionally, models, though are accurate for certain applications, may suffer from performance drops when the applications are different, even slightly. Transfer learning could be used to apply knowledge gained from one domain (e.g., load forecasting in Europe) to another (e.g., load forecasting in Asia), which reduces training time, lowers computational costs, and improves results with smaller datasets.

4) Integration of EVs as Distributed Energy Resources: Furthermore, combining traditional grid-to-vehicle (G2V) and advanced V2G techniques allows EVs act as mobile energy storage units and enables bidirectional energy flow between the vehicle and the grid. The G2V techniques view EVs as the energy consumers, while V2G techniques provide EVs with the opportunity to deliver power back to the grid. This bidirectional energy flow capability transforms EVs into essential elements of distributed energy storage systems, provides the benefits to the grid such as helping for load balance, mitigating peak demand, enhancing grid resilience and stability. For future work, developing GenAI-based optimization algorithms which can dynamically manage energy flow between EVs and the grid, as well as energy transactions between EVs, in coordination with renewable energy sources, is crucial for future of smart grids and sustainable energy solutions.

B. New Methodologies

Besides enhancing existing models, new technologies can also be explored.

1) LLMs: LLMs are often discussed together with GenAI, both representing the latest ML technology advancements. LLMs were originally developed for language-centric applications and language has not been the focus of EV related applications. With the ever-increasing interactions between EV systems and users, there is a demand to incorporate NLP into the technology stack to facilitate the interaction and contribute to the enhancement of IoEV. For example, LLMs such as GPT-4 and LLaMA can offer enhanced interactions between EV users and charging infrastructure. The EV users can specify their charging preferences in natural language, such as "*I need to charge my car fully by 7 AM, tomorrow morning*". Then, the LLMs-based system can translate this into an optimized charging schedule. Such interaction simplifies the process and offers a more accessible and personalized experience for the EV users, not necessarily from a technical background of IoEV. LLMs can also be used to interpret users' historical queries, largely language-based, to predict user demand and optimize real-time responses. Overall, the integration of LLMs in IoEV charging systems shall be investigated to improve user experience and facilitate the adoption of electric mobility.

2) Federated Learning: Privacy concerns, systems' robustness, and decentralized systems are critical in IoEV applications. Federated learning can address security issues by enabling EVs to learn collaboratively without sharing sensitive data. Future work could study how federated learning can be effectively implemented in large-scale IoEV to promote the system's scalability and security.

3) Hybrid Models: When one model is insufficient to perform well, the hybrid models can be considered and the models are not limited to GenAI. For example, combining GenAI with DRL may improve the models' performance in complex IoEV applications. Several existing research efforts are GRU and VAE [51] for battery anomaly detection, LSTM and GAN for SoC estimation [66], DNN and GAN for EV charging behaviors prediction [17], etc. Future research may explore cross-layer optimizations with hybrid models, e.g., advanced BMS informs routing decisions and grid interactions. This helps create an integrated and efficient system.

4) Embodied AI and AGI: Looking toward future advancements, the embodied AI and AGI are expected to play transformative roles in IoEV. The former combines cognitive processing with physical interaction to enhance adaptability and real-world decision-making, e.g., navigating complex environments and responding to traffic conditions. The latter aims at achieving broad cognitive abilities similar to human intelligence. It can potentially revolutionize IoEV, e.g., make decisions in unpredictable environments without relying on pre-programmed instructions.

Together, new methodologies and technologies could drive further innovation in IoEV in various aspects such as sustainability and safety.

VI. Conclusions

In this survey, we explored the applications of GenAI in the IoEV from various perspectives. We grouped relevant applications across four different layers: EV's battery layer, individual EV layer, the grid layer, and security layer. We concluded that these applications primarily utilize GenAI's capabilities in data feature extraction, enhancement, and generation. The characteristics of GenAI make it ideal for data augmentation. Hence, research works in the first three layers all used GenAI for this purpose. At layer 1, GenAIs including GAN, GDM, VAE, and Transformer were introduced for anomaly detection, as well as SoC and SoH estimations of the EV's battery. At layer 2, GAN and Transformer were discussed for EV charging behaviors and loads as well as the optimal EV routing problem. At layer 3, GAN, VAE, and Transformer are the main techniques currently employed for EV charging load forecasting/charging scenarios generation, and LLMs for the analysis of EV charging reviews. At layer 4, GAN and AE were often applied for the detection and generation of attacks that may threaten the systems across layers 1 to 3. Subsequently, we summarized the publicly available dataset and the possible further research directions for GenAI's applications in IoEV. In conclusion, this survey highlights the essential role of GenAI in IoEV and underscores the urgent need for further exploration of its applications.

References

[1] S. Becker, "Are US EV sales a disaster or a booming segment? the answer may be both," BBC: https://www.bbc.com/worklife/article/20240228-are-us-ev-sales-a-disaster-or-a-booming-segment-the-answer-may-be-both, 29 Feb. 2024, accessed: 20 Mar 2024.

[2] Editor OilPrice.com, "China's ev growth set to explode in 2024," Yahoo!finance: https://finance.yahoo.com/news/china-ev-growth-set-explode-150000597.html, 21 Feb. 2024, accessed: 03 Apr 2024.

[3] VANTAGE Market Research, "Electric vehicle market - global industry assessment & forecast," VANTAGE Market Research's report: https://www.vantagemarketresearch.com/industry-report/electric-vehicle-ev-market-1853, Feb. 2024, accessed: 03 Apr 2024.

[4] S. Zhang, S. Zhang, L. K. Yeung *et al.*, "Urban internet of electric vehicle parking system for vehicle-to-grid scheduling: Formulation and distributed algorithm," *IEEE Transactions on Vehicular Technology*, 2023.

[5] O. Gul M. Khan, F. Elghitani, A. Youssef *et al.*, "Real-time congestion-aware charging station assignment model for evs," *IEEE Internet of Things Journal*, vol. 11, no. 7, pp. 11 723–11 736, 2024.

[6] H. Wang, S. Chen, M. Li *et al.*, "Demand-driven charging strategy based distributed routing optimization under traffic restrictions in internet of electric vehicles," *IEEE Internet of Things Journal*, pp. 1–1, 2024.

[7] S. Lee and D.-H. Choi, "Three-stage deep reinforcement learning for privacy-and safety-aware smart electric vehicle charging station scheduling and volt/var control," *IEEE Internet of Things Journal*, vol. 11, no. 5, pp. 8578–8589, 2024.

[8] S. Liu, C. Wu, J. Huang *et al.*, "Blockchain-based interpretable electric vehicle battery life prediction in iov," *IEEE Internet of Things Journal*, vol. 11, no. 4, pp. 7214–7227, 2024.

[9] M. Naik, A. P. Singh, N. R. Pradhan *et al.*, "Tokengreen: A versatile nft framework for peer-to-peer energy trading and asset ownership of electric vehicles," *IEEE Internet of Things Journal*, vol. 11, no. 8, pp. 13 636–13 646, 2024.

[10] A. Mohammad, R. Zamora, and T. T. Lie, "Transactive energy management of pv-based ev integrated parking lots," *IEEE Systems Journal*, vol. 15, no. 4, pp. 5674–5682, 2021.

[11] J. Liu, G. Lin, S. Huang *et al.*, "Optimal ev charging scheduling by considering the limited number of chargers," *IEEE Transactions on Transportation Electrification*, vol. 7, no. 3, pp. 1112–1122, 2021.

[12] Y. Li, S. He, Y. Li *et al.*, "Probabilistic charging power forecast of evcs: Reinforcement learning assisted deep learning approach," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 1, pp. 344–357, 2023.

[13] X. Diao, L. Jiang, T. Gao *et al.*, "Research on electric vehicle charging safety warning based on a-lstm algorithm," *IEEE Access*, vol. 11, pp. 55 081–55 093, 2023.

[14] Z. Pan, J. Wang, W. Liao *et al.*, "Data-driven ev load profiles generation using a variational auto-encoder," *Energies*, vol. 12, no. 5, 2019.

[15] S. Koohfar, W. Woldemariam, and A. Kumar, "Prediction of electric vehicles charging demand: A transformer-based deep learning approach," *Sustainability*, vol. 15, no. 3, 2023.

[16] X. Shen, H. Zhao, Y. Xiang *et al.*, "Short-term electric vehicles charging load forecasting based on deep learning in low-quality data environments," *Electric Power Systems Research*, vol. 212, p. 108247, 2022.

[17] A. Forootani, M. Rastegar, and H. Zareipour, "Transfer learning-based framework enhanced by deep generative model for cold-start forecasting of residential ev charging behavior," *IEEE Transactions on Intelligent Vehicles*, pp. 1–9, 2023.

[18] BioLogic, "Battery states: State of charge (SoC), state of health (SoH). electrochemistry basics series." https://www.biologic.net/topics/battery-states-state-of-charge-soc-state-of-health-soh/, 06 Oct 2023, accessed: 05 May 2024.

[19] A. Vaswani, N. Shazeer, N. Parmar *et al.*, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[20] F. Liu, B. Ding, C. Guan *et al.*, "Demystify adult learning: A social network and large language model assisted approach," in *2024 IEEE Annual Congress on Artificial Intelligence of Things (AIoT)*, 2024, pp. 207–212.

[21] Z. Kwan, W. Zhang, Z. Wang *et al.*, "Nutrition estimation for dietary management: A transformer approach with depth sensing," *arXiv:2406.01938*, 2024.

[22] A. Radford, J. W. Kim, T. Xu *et al.*, "Robust speech recognition via large-scale weak supervision," *arXiv:2212.04356*, 2022.

[23] W. Kim, B. Son, and I. Kim, "ViLT: Vision-and-language transformer without convolution or region supervision," *arXiv:2102.03334*, 2021.

[24] S. Ha, D. J. Marchetto, S. Dharur *et al.*, "Topic classification of electric vehicle consumer experiences with transformer-based deep learning," *Patterns*, vol. 2, no. 2, p. 100195, 2021.

[25] J. Zhao, X. Feng, J. Wang *et al.*, "Battery fault diagnosis and failure prognosis for electric vehicles using spatio-temporal transformer networks," *Applied Energy*, vol. 352, p. 121949, 2023.

[26] M. Tang, W. Zhuang, B. Li *et al.*, "Energy-optimal routing for electric vehicles using deep reinforcement learning with transformer," *Applied Energy*, vol. 350, p. 121711, 2023.

[27] X. Huang, D. Wu, and B. Boulet, "Metaprobformer for charging load probabilistic forecasting of electric vehicle charging stations," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 10, pp. 10 445–10 455, 2023.

[28] I. Goodfellow, J. Pouget-Abadie, M. Mirza *et al.*, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.

[29] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv:1511.06434*, 2015.

[30] T. Karras, T. Aila, S. Laine *et al.*, "Progressive growing of gans for improved quality, stability, and variation," *arXiv:1710.10196*, 2017.

[31] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *International conference on machine learning*. PMLR, 2017, pp. 214–223.

[32] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 4401–4410.

[33] S. Tang, Q. Yang, D. Gündüz *et al.*, "Evolving semantic communication with generative model," *arXiv:2403.20237*, 2024.

[34] T. Han, J. Tang, Q. Yang *et al.*, "Generative model based highly efficient semantic communication approach for image transmission," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5.

[35] C.-H. Lee, Z. Liu, L. Wu *et al.*, "Maskgan: Towards diverse and interactive facial image manipulation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 5549–5558.

[36] C. Hu, F. Cheng, Y. Zhao *et al.*, "State of charge estimation for lithium-ion batteries based on data augmentation with generative adversarial network," *Journal of Energy Storage*, vol. 80, p. 110004, 2024.

[37] W. Gang, L. Wu, and G. Xuan, "A load forecasting method of electric vehicles charging station group based on GAN-RF model," in *2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2)*, 2021, pp. 3076–3079.

[38] W. Shi, D. Zhang, X. Han *et al.*, "Coordinated operation of active distribution network, networked microgrids, and electric vehicle: A multi-agent ppo optimization method," *CSEE Journal of Power and Energy Systems*, pp. 1–12, 2023.

[39] Z. Peng, Q. Yang, D. Li *et al.*, "Adversarial attacks on deep reinforcement learning applications in electric vehicle charging scheduling: A dual-stage attack framework," *SSRN*, 2023.

[40] H. Li, D. Liang, Y. Zhou *et al.*, "Spatiotemporal charging demand models for electric vehicles considering user strategies," *Frontiers in Energy Research*, vol. 10, 2023.

[41] J. Yoon, J. Jordon, and M. van der Schaar, "GAIN: Missing data imputation using generative adversarial nets," in *Proceedings of the 35th International Conference on Machine Learning*, vol. 80, Jul 2018, pp. 5689–5698.

[42] D. Pathak, P. Krahenbuhl, J. Donahue *et al.*, "Context encoders: Feature learning by inpainting," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2536–2544.

[43] Z. Liu, R. A. Yeh, X. Tang *et al.*, "Video frame synthesis using deep voxel flow," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 4463–4471.

[44] T. Chen, S. Kornblith, K. Swersky *et al.*, "Big self-supervised models are strong semi-supervised learners," *Advances in neural information processing systems*, vol. 33, pp. 22 243–22 255, 2020.

[45] S. R. Fahim, R. Atat, C. Kececi *et al.*, "Graph autoencoder-based power attacks detection for resilient electrified transportation systems," *IEEE Transactions on Transportation Electrification*, pp. 1–1, 2024.

[46] H. S. Mavikumbure, V. Cobilean, C. S. Wickramasinghe *et al.*, "Cyphy ads: Cyber physical anomaly detection framework for ev charging systems," *IEEE Transactions on Transportation Electrification*, pp. 1–1, 2024.

[47] C. S. Wickramasinghe, D. L. Marino, H. S. Mavikumbure *et al.*, "Rx-ads: Interpretable anomaly detection using adversarial ml for electric vehicle can data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 12, pp. 14 051–14 063, 2023.

[48] R. Yeh, Z. Liu, D. B. Goldman *et al.*, "Semantic facial expression editing using autoencoded flow," *arXiv:1611.09961*, 2016.

[49] J. Walker, C. Doersch, A. Gupta *et al.*, "An uncertain future: Forecasting from static images using variational autoencoders," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VII 14*. Springer, 2016, pp. 835–851.

[50] L. Pan, X. Chen, Z. Cai *et al.*, "Variational relational point completion network," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 8524–8533.

[51] C. Sun, Z. He, H. Lin *et al.*, "Anomaly detection of power battery pack using gated recurrent units based variational autoencoder," *Applied Soft Computing*, vol. 132, p. 109903, 2023.

[52] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in neural information processing systems*, vol. 33, pp. 6840–6851, 2020.

[53] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," *arXiv:2010.02502*, 2020.

[54] R. Rombach, A. Blattmann, D. Lorenz *et al.*, "High-resolution image synthesis with latent diffusion models," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 10 684–10 695.

[55] L. Zhang, A. Rao, and M. Agrawala, "Adding conditional control to text-to-image diffusion models," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 3836–3847.

[56] H. Du, R. Zhang, Y. Liu *et al.*, "Enhancing deep reinforcement learning: A tutorial on generative diffusion models in network optimization," *IEEE Communications Surveys & Tutorials*, pp. 1–1, 2024.

[57] J. Wang, H. Du, D. Niyato *et al.*, "A unified framework for guiding generative ai with wireless perception in resource constrained mobile edge networks," *IEEE Transactions on Mobile Computing*, pp. 1–17, 2024.

[58] ——, "Generative artificial intelligence assisted wireless sensing: Human flow detection in practical communication environments," *IEEE Journal on Selected Areas in Communications*, pp. 1–1, 2024.

[59] ——, "Generative ai for integrated sensing and communication: Insights from the physical layer perspective," *IEEE Wireless Communications*, pp. 1–10, 2024.

[60] C. Luo, Z. Zhang, S. Zhu *et al.*, "State-of-health prediction of lithium-ion batteries based on diffusion model with transfer learning," *Energies*, vol. 16, no. 9, 2023.

[61] S. Li, H. Xiong, and Y. Chen, "Diffcharge: Generating ev charging scenarios via a denoising diffusion model," *IEEE Transactions on Smart Grid*, pp. 1–1, 2024.

[62] Z. W. Boyang Chen and R. Zhao, "From fiction to fact: the growing role of generative ai in business and finance," *Journal of Chinese Economic and Business Studies*, vol. 21, no. 4, pp. 471–496, 2023.

[63] K. Moulaei, A. Yadegari, M. Baharestani *et al.*, "Generative artificial intelligence in healthcare: A scoping review on benefits, challenges and applications," *International Journal of Medical Informatics*, vol. 188, p. 105474, 2024.

[64] O. Noroozi, S. Soleimani, M. Farrokhnia *et al.*, "Generative ai in education: Pedagogical, theoretical, and methodological perspectives." *International Journal of Technology in Education*, vol. 7, no. 3, pp. 373–385, 2024.

[65] X. Gu, K. See, Y. Liu *et al.*, "A time-series wasserstein gan method for state-of-charge estimation of lithium-ion batteries," *Journal of Power Sources*, vol. 581, p. 233472, 2023.

[66] X. Qiu, S. Wang, and K. Chen, "A conditional generative adversarial network-based synthetic data augmentation technique for battery state-of-charge estimation," *Applied Soft Computing*, vol. 142, p. 110281, 2023.

[67] MoviTHERM, "Mitigating EV battery fires with infrared technology," https://movitherm.com/2024/03/21/blog/mitigating-ev-battery-fires-with-infrared-technology/, 21 Mar 2024, accessed: 28 Apr 2024.

[68] A. Naha, A. Khandelwal, S. Agarwal *et al.*, "Internal short circuit detection in li-ion batteries using supervised machine learning," *Scientific reports*, vol. 10, no. 1, p. 1301, 2020.

[69] J. Xie, L. Zhang, T. Yao *et al.*, "Quantitative diagnosis of internal short circuit for cylindrical li-ion batteries based on multiclass relevance vector machine," *Journal of Energy Storage*, vol. 32, p. 101957, 2020.

[70] Y. Jia, J. Li, C. Yuan *et al.*, "Data-driven safety risk prediction of lithium-ion battery," *Advanced Energy Materials*, vol. 11, no. 18, p. 2003868, 2021.

[71] J. Hong, Z. Wang, and Y. Yao, "Fault prognosis of battery system based on accurate voltage abnormity prognosis using long short-term memory neural networks," *Applied Energy*, vol. 251, p. 113381, 2019.

[72] J. Hong, Z. Wang, W. Chen *et al.*, "Synchronous multi-parameter prediction of battery systems on electric vehicles using long short-term memory networks," *Applied Energy*, vol. 254, p. 113648, 2019.

[73] H. Zhao, Y. Wang, J. Duan *et al.*, "Multivariate time-series anomaly detection via graph attention network," in *2020 IEEE International Conference on Data Mining (ICDM)*, 2020, pp. 841–850.

[74] D. Li, Z. Zhang, P. Liu *et al.*, "Battery fault diagnosis for electric vehicles based on voltage abnormality by combining the long short-term memory neural network and the equivalent circuit model," *IEEE Transactions on Power Electronics*, vol. 36, no. 2, pp. 1303–1315, 2021.

[75] N. Smol, "The risk of lithium-ion batteries – thermal runaway in EV's," Fire Isolator News: https://fireisolator.com/the-risk-of-lithium-ion-batteries-thermal-runaway-in-evs/, 28 Jun 2023, accessed: 29 Apr 2024.

[76] J. Zhao, H. Ling, J. Wang *et al.*, "Data-driven prediction of battery failure for electric vehicles," *Iscience*, vol. 25, no. 4, 2022.

[77] BMS Control Systems Limited, "Exploring BMS state of charge (SOC): Monitoring battery health," BMS Controls Articles: https://bmscontrols.co.uk/blog/exploring-bms-state-of-charge-soc-monitoring-battery-health/, accessed: 28 Apr 2024.

[78] K. S. Mawonou, A. Eddahech, D. Dumur *et al.*, "State-of-health estimators coupled to a random forest approach for lithium-ion battery aging factor ranking," *Journal of Power Sources*, vol. 484, p. 229154, 2021.

[79] J. Wang, R. Zhao, Q.-A. Huang *et al.*, "High-efficient prediction of state of health for lithium-ion battery based on ac impedance feature tuned with gaussian process regression," *Journal of Power Sources*, vol. 561, p. 232737, 2023.

[80] Z. Zhao, Y. Zhang, Y. Yang *et al.*, "Load forecasting via grey model-least squares support vector machine model and spatial-temporal distribution of electric consumption intensity," *Energy*, vol. 255, p. 124468, 2022.

[81] X. Fan, W. Zhang, C. Zhang *et al.*, "Soc estimation of li-ion battery using convolutional neural network with u-net architecture," *Energy*, vol. 256, p. 124612, 2022.

[82] F. Yang, X. Song, F. Xu *et al.*, "State-of-charge estimation of lithium-ion batteries via long short-term memory network," *IEEE Access*, vol. 7, pp. 53 792–53 799, 2019.

[83] X. Zhang, Y. Qin, C. Yuen *et al.*, "Time-series regeneration with convolutional recurrent generative adversarial network for remaining useful life estimation," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 10, pp. 6820–6831, 2021.

[84] H. Ni, L. Szpruch, M. Sabate-Vidales *et al.*, "Sig-wasserstein gans for time series generation," in *Proceedings of the Second ACM International Conference on AI in Finance*, ser. ICAIF '21. New York, NY, USA: Association for Computing Machinery, 2022.

[85] F. Naaz, J. Channegowda, M. Lakshminarayanan *et al.*, "Solving limited data challenges in battery parameter estimators by using generative adversarial networks," in *2021 IEEE PES/IAS PowerAfrica*, 2021, pp. 1–3.

[86] Y. Tian, Q. Dong, J. Tian *et al.*, "Capacity estimation of lithium-ion batteries based on optimized charging voltage section and virtual sample generation," *Applied Energy*, vol. 332, p. 120516, 2023.

[87] Q. Wang, X. Zhou, C. Wang *et al.*, "Wgan-based synthetic minority over-sampling technique: Improving semantic fine-grained classification for lung nodules in ct images," *IEEE Access*, vol. 7, pp. 18 450–18 463, 2019.

[88] X. Gu, K. See, Y. Wang *et al.*, "The sliding window and shap theory—an improved system with a long short-term memory network model for state of charge prediction in electric vehicle application," *Energies*, vol. 14, no. 12, 2021.

[89] C. Vidal, P. Kollmeyer, E. Chemali *et al.*, "Li-ion battery state of charge estimation using long short-term memory recurrent neural network with transfer learning," in *2019 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2019, pp. 1–6.

[90] TOMO Battery, "Lg inr18650hg2," Quick Overview: https://www.tomobattery.com/product/lg-inr18650hg2/, Apr 2024, accessed: 27 Apr 2024.

[91] Monolithic Power Systems, Inc. , "The importance of state of charge (SOC) and state of health (SOH) in battery management systems," MPS Article: https://www.monolithicpower.com/en/the-importance-of-state-of-charge-and-state-of-health-in-battery-management-systems, accessed: 28 Apr 2024.

[92] X. Lai, C. Deng, J. Li *et al.*, "Rapid sorting and regrouping of retired lithium-ion battery modules for echelon utilization based on partial charging curves," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1246–1254, 2021.

[93] W. Li, D. Cao, D. Jöst *et al.*, "Parameter sensitivity analysis of electrochemical model-based battery management systems for lithium-ion batteries," *Applied Energy*, vol. 269, p. 115104, 2020.

[94] S. Amir, M. Gulzar, M. O. Tarar *et al.*, "Dynamic equivalent circuit model to estimate state-of-health of lithium-ion batteries," *IEEE Access*, vol. 10, pp. 18 279–18 288, 2022.

[95] H.-T. Lin, T.-J. Liang, and S.-M. Chen, "Estimation of battery state of health using probabilistic neural network," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 2, pp. 679–685, 2013.

[96] Z. Deng, L. Yang, Y. Cai *et al.*, "Online available capacity prediction and state of charge estimation based on advanced data-driven algorithms for lithium iron phosphate battery," *Energy*, vol. 112, pp. 469–480, 2016.

[97] C. Luo, Z. Zhang, D. Qiao *et al.*, "Life prediction under charging process of lithium-ion batteries based on automl," *Energies*, vol. 15, no. 13, 2022.

[98] X. Li, Z. Wang, and J. Yan, "Prognostic health condition for lithium battery using the partial incremental capacity and gaussian process regression," *Journal of Power Sources*, vol. 421, pp. 56–67, 2019.

[99] S. Sohn, H.-E. Byun, and J. H. Lee, "Cnn-based online diagnosis of knee-point in li-ion battery capacity fade curve," *IFAC-PapersOnLine*, vol. 55, no. 7, pp. 181–185, 2022, 13th IFAC Symposium on Dynamics and Control of Process Systems, including Biosystems DYCOPS 2022.

[100] Y. Harada, L. Lin, and M. Fukui, "An rnn approach for lithium-ion battery internal impedance estimation," in *2020 IEEE REGION 10 CONFERENCE (TENCON)*. IEEE, 2020, pp. 1145–1146.

[101] M. Singh, S. Bansal, Vandana *et al.*, "A Genetic Algorithm and RNN-LSTM Model for Remaining Battery Capacity Prediction," *Journal of Computing and Information Science in Engineering*, vol. 22, no. 4, p. 041009, 02 2022.

[102] R. Rouhi Ardeshiri and C. Ma, "Multivariate gated recurrent unit for battery remaining useful life prediction: A deep learning approach," *International Journal of Energy Research*, vol. 45, no. 11, pp. 16 633–16 648, 2021.

[103] Y. Zhao, Z. Wei, Q. Yan *et al.*, "Practical battery health monitoring using uncertainty-aware bayesian neural network," *arXiv:2404.14444*, 2024.

[104] L. Yang, Z. Zhang, Y. Song *et al.*, "Diffusion models: A comprehensive survey of methods and applications," *ACM Computing Surveys*, vol. 56, no. 4, pp. 1–39, 2023.

[105] K. A. Severson, P. M. Attia, N. Jin *et al.*, "Data-driven prediction of battery cycle life before capacity degradation," *Nature Energy*, vol. 4, no. 5, pp. 383–391, 2019.

[106] D. Chen, W. Hong, and X. Zhou, "Transformer network for remaining useful life prediction of lithium-ion batteries," *IEEE Access*, vol. 10, pp. 19 621–19 628, 2022.

[107] S. Li, X. Jin, Y. Xuan *et al.*, "Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting," *Advances in neural information processing systems*, vol. 32, 2019.

[108] M. Razghandi, H. Zhou, M. Erol-Kantarci *et al.*, "Smart home energy management: Vae-gan synthetic dataset generator and q-learning," *IEEE Transactions on Smart Grid*, vol. 15, no. 2, pp. 1562–1573, 2024.

[109] Z. Chen, Z. Zhang, J. Zhao *et al.*, "An analysis of the charging characteristics of electric vehicles based on measured data and its application," *IEEE Access*, vol. 6, pp. 24 475–24 487, 2018.

[110] H. Jahangir, H. Tayarani, A. Ahmadian *et al.*, "Charging demand of plug-in electric vehicles: Forecasting travel behavior based on a novel rough artificial neural network approach," *Journal of cleaner production*, vol. 229, pp. 1029–1044, 2019.

[111] X. Zhang, K. W. Chan, H. Li *et al.*, "Deep-learning-based probabilistic forecasting of electric vehicle charging load with a novel queuing model," *IEEE transactions on cybernetics*, vol. 51, no. 6, pp. 3157–3170, 2020.

[112] A. J. Jahromi, M. Mohammadi, S. Afrasiabi *et al.*, "Probability density function forecasting of residential electric vehicles charging profile," *Applied Energy*, vol. 323, p. 119616, 2022.

[113] A. Forootani, M. Rastegar, and A. Sami, "Short-term individual residential load forecasting using an enhanced machine learning-based approach based on a feature engineering framework: A comparative study with deep learning methods," *Electric Power Systems Research*, vol. 210, p. 108119, 2022.

[114] M. Gilanifar and M. Parvania, "Clustered multi-node learning of electric vehicle charging flexibility," *Applied energy*, vol. 282, p. 116125, 2021.

[115] EA Technology, "My electric avenue data download," MPS Article: https://eatechnology.com/resources/projects/my-electric-avenue-data-download/, 04 May 2017, accessed: 02 May 2024.

[116] Y.-W. Chung, B. Khaki, T. Li *et al.*, "Ensemble machine learning-based algorithm for electric vehicle user behavior prediction," *Applied Energy*, vol. 254, p. 113732, 2019.

[117] S. Shahriar, A.-R. Al-Ali, A. H. Osman *et al.*, "Prediction of ev charging behavior using machine learning," *Ieee Access*, vol. 9, pp. 111 576–111 586, 2021.

[118] S. Zhang, Z. Ji, J. Zhang *et al.*, "Multi-dimensional data generation method of electric vehicle charging behaviors based on improved generative adversarial network," in *2022 IEEE / IAS Industrial and Commercial Power System Asia (I & CPS Asia)*, 2022, pp. 1827–1832.

[119] F. Charbonnier, T. Morstyn, and M. McCulloch, "Home electricity data generator (hedge): An open-access tool for the generation of electric vehicle, residential demand, and pv generation profiles," *MethodsX*, vol. 12, p. 102618, 2024.

[120] Customer-Led Network Revolution, "Dataset (tc1a): Basic profiling of domestic smart meter customers," http://www.networkrevolution.co.uk/project-library/dataset-tc1a-basic-profiling-domestic-smart-meter-customers/, 28 Aug 2014, accessed: 16 Jul 2024.

[121] ——, "Dataset (tc5): Enhanced profiling of domestic customers with solar photovoltaics (pv)," http://www.networkrevolution.co.uk/project-library/dataset-tc5-enhanced-profiling-solar-photovoltaic-pv-users/, 29 Aug 2014, accessed: 16 Jul 2024.

[122] Department for Transport (2021), "National travel survey, 2002-2020," http://doi.org/10.5255/UKDA-SN-5340-11, 25 October 2021, accessed: 17 Jul 2024.

[123] P. Huber, M. Ott, M. Friedli *et al.*, "Residential power traces for five houses: the ihomelab rapt dataset," *Data*, vol. 5, no. 1, p. 17, 2020.

[124] Y. Yingfei, Z. Mengze, L. Zeyu *et al.*, "Green logistics performance and infrastructure on service trade and environment-measuring firm's performance and service quality," *Journal of King Saud University - Science*, vol. 34, no. 1, p. 101683, 2022.

[125] O. Vinyals, M. Fortunato, and N. Jaitly, "Pointer networks," *Advances in neural information processing systems*, vol. 28, 2015.

[126] T. Barrett, W. Clements, J. Foerster *et al.*, "Exploratory combinatorial optimization with reinforcement learning," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, 2020, pp. 3243–3250.

[127] W. Pan and S. Q. Liu, "Deep reinforcement learning for the dynamic and uncertain vehicle routing problem," *Applied Intelligence*, vol. 53, no. 1, pp. 405–422, 2023.

[128] W. Kool, H. Van Hoof, and M. Welling, "Attention, learn to solve routing problems!" *arXiv:1803.08475*, 2018.

[129] S. Erdoğan and E. Miller-Hooks, "A green vehicle routing problem," *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 1, pp. 100–114, 2012, select Papers from the 19th International Symposium on Transportation and Traffic Theory.

[130] Gurobi Optimization, LLC, "Computer Software: Gurobi Optimizer (Version 9.5)," https://www.gurobi.com, 2022.

[131] S. Zhang, Y. Gajpal, S. Appadoo *et al.*, "Electric vehicle routing problem with recharging stations for minimizing energy consumption," *International journal of production economics*, vol. 203, pp. 404–413, 2018.

[132] D. Goeke and M. Schneider, "Routing a mixed fleet of electric and conventional vehicles," *European Journal of Operational Research*, vol. 245, no. 1, pp. 81–99, 2015.

[133] R. Zhang, K. Xiong, H. Du *et al.*, "Generative ai-enabled vehicular networks: Fundamentals, framework, and case study," *IEEE Network*, vol. 38, no. 4, pp. 259–267, 2024.

[134] M. Andreoni, W. T. Lunardi, G. Lawton *et al.*, "Enhancing autonomous system security and resilience with generative ai: A comprehensive survey," *IEEE Access*, vol. 12, pp. 109 470–109 493, 2024.

[135] H. Lin, Y. Liu, S. Li *et al.*, "How generative adversarial networks promote the development of intelligent transportation systems: A survey," *IEEE/CAA Journal of Automatica Sinica*, vol. 10, no. 9, pp. 1781–1796, 2023.

[136] N. Huang, Q. He, J. Qi *et al.*, "Multinodes interval electric vehicle day-ahead charging load forecasting based on joint adversarial generation," *International Journal of Electrical Power & Energy Systems*, vol. 143, p. 108404, 2022.

[137] M. A. Gharibi, H. Nafisi, H. Askarian-abyaneh *et al.*, "Deep learning framework for day-ahead optimal charging scheduling of electric vehicles in parking lot," *Applied Energy*, vol. 349, p. 121614, 2023.

[138] Y. Luo, X. Cai, Y. Zhang *et al.*, "Multivariate time series imputation with generative adversarial networks," *Advances in neural information processing systems*, vol. 31, 2018.

[139] M. Gilanifar and M. Parvania, "Clustered multi-node learning of electric vehicle charging flexibility," *Applied energy*, vol. 282, p. 116125, 2021.

[140] R. Buechler, E. Balogun, A. Majumdar *et al.*, "Evgen: Adversarial networks for learning electric vehicle charging loads and hidden representations," *arXiv:2108.03762*, 2021.

[141] M. Jamshidian and M. Mata, "2 - advances in analysis of mean and covariance structure when data are incomplete," in *Handbook of Latent Variable and Related Models*, 2007, pp. 21–44.

[142] Tobin A. Driscoll and Richard J. Braun, "5.2. piecewise linear interpolation — fundamentals of numerical computation," https://tobydriscoll.net/fnc-julia/localapprox/pwlin.html, 2022, accessed: 19 May 2024.

[143] S. Zhang, "Nearest neighbor selection for iteratively knn imputation," *Journal of Systems and Software*, vol. 85, no. 11, pp. 2541–2552, 2012.

[144] Q. Sun, J. Liu, X. Rong *et al.*, "Charging load forecasting of electric vehicle charging station based on support vector regression," in *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*. IEEE, 2016, pp. 1777–1781.

[145] T. S. Mahmoud, D. Habibi, M. Y. Hassan *et al.*, "Modelling self-optimised short term load forecasting for medium voltage loads using tunning fuzzy systems and artificial neural networks," *Energy Conversion and Management*, vol. 106, pp. 1396–1408, 2015.

[146] G. Lai, W.-C. Chang, Y. Yang *et al.*, "Modeling long-and short-term temporal patterns with deep neural networks," in *The 41st international ACM SIGIR conference on research & development in information retrieval*, 2018, pp. 95–104.

[147] X. Huang, D. Wu, and B. Boulet, "Ensemble learning for charging load forecasting of electric vehicle charging stations," in *2020 IEEE Electric Power and Energy Conference (EPEC)*. IEEE, 2020, pp. 1–5.

[148] M. Li and M. Lenzen, "How many electric vehicles can the current australian electricity grid support?" *International Journal of Electrical Power & Energy Systems*, vol. 117, p. 105586, 2020.

[149] N. Nguyen and B. Quanz, "Temporal latent auto-encoder: A method for probabilistic multivariate time series forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 10, pp. 9117–9125, May 2021.

[150] H. Zhou, S. Zhang, J. Peng *et al.*, "Informer: Beyond efficient transformer for long sequence time-series forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 12, pp. 11 106–11 115, 2021.

[151] I. Svetunkov, "Forecasting and analytics with the augmented dynamic adaptive model (adam)," https://openforecast.org/adam/typesOfForecasts.html, 08 May 2024, accessed: 20 May 2024.

[152] A. Nichol, "On first-order meta-learning algorithms," *arXiv:1803.02999*, 2018.

[153] Z. J. Lee, T. Li, and S. H. Low, "Acn-data: Analysis and applications of an open ev charging dataset," in *Proceedings of the Tenth ACM International Conference on Future Energy Systems*, ser. e-Energy '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 139–149.

[154] Bundesnetzagentur, "Smard platform," https://www.smard.de/en, 2024, accessed: 29 May 2024.

[155] United States Environmental Protection Agency, "Inventory of u.s. greenhouse gas emissions and sinks: 1990-2016," https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2016, 2018, accessed: 14 Jun 2024.

[156] N. R. Council, G. Affairs, B. on Science *et al.*, *Hidden costs of energy: unpriced consequences of energy production and use*. National Academies Press, 2010.

[157] O. I. Asensio, K. Alvarez, A. Dror *et al.*, "Real-time data from mobile platforms to evaluate sustainable transportation infrastructure," *Nature Sustainability*, vol. 3, no. 6, pp. 463–471, 2020.

[158] J. Devlin, M.-W. Chang, K. Lee *et al.*, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv:1810.04805*, 2018.

[159] Z. Yang, Z. Dai, Y. Yang *et al.*, "Xlnet: Generalized autoregressive pretraining for language understanding," *Advances in neural information processing systems*, vol. 32, 2019.

[160] K. Alvarez, "Evaluating electric vehicle user mobility data using neural network-based language models," in *Proceedings of the Transportation Research Board (TRB) annual meeting*, 2019.

[161] O. F. Zaidan, J. Eisner, and C. Piatko, "Machine learning with annotator rationales to reduce annotation cost," in *Proceedings of the NIPS* 2008 workshop on cost sensitive learning*, 2008, pp. 260–267.

[162] S. Serrano and N. A. Smith, "Is attention interpretable?" *arXiv:1906.03731*, 2019.

[163] A. T. Nguyen, B. C. Wallace, J. J. Li *et al.*, "Aggregating and predicting sequence labels from crowd annotations," in *Proceedings of the conference. Association for Computational Linguistics. Meeting*, vol. 2017. NIH Public Access, 2017, p. 299.

[164] J. Qian, Y. Jiang, X. Liu *et al.*, "Federated reinforcement learning for electric vehicles charging control on distribution networks," *IEEE Internet of Things Journal*, vol. 11, no. 3, pp. 5511–5525, 2024.

[165] F. Zhang, Q. Yang, and D. An, "Cddpg: A deep-reinforcement-learning-based approach for electric vehicle charging control," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3075–3087, 2021.

[166] H. Li, Z. Wan, and H. He, "Constrained EV charging scheduling based on safe deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2427–2439, 2020.

[167] J. Zhang, L. Che, and M. Shahidehpour, "Distributed training and distributed execution-based Stackelberg multi-agent reinforcement learning for EV charging scheduling," *IEEE Transactions on Smart Grid*, vol. 14, no. 6, pp. 4976–4979, 2023.

[168] J. Jin, S. Mao, and Y. Xu, "Optimal priority rule-enhanced deep reinforcement learning for charging scheduling in an electric vehicle battery swapping station," *IEEE Transactions on Smart Grid*, vol. 14, no. 6, pp. 4581–4593, 2023.

[169] J. Wei, L. Yao, and Q. Meng, "Self-adaptive logit balancing for deep neural network robustness: Defence and detection of adversarial attacks," *Neurocomputing*, vol. 531, pp. 180–194, 2023.

[170] A. Chakraborty, M. Alam, V. Dey *et al.*, "A survey on adversarial attacks and defences," *CAAI Transactions on Intelligence Technology*, vol. 6, no. 1, pp. 25–45, 2021.

[171] S. Huang, N. Papernot, I. Goodfellow *et al.*, "Adversarial attacks on neural network policies," *arXiv:1702.02284*, 2017.

[172] A. Kurakin, I. J. Goodfellow, and S. Bengio, "Adversarial examples in the physical world," in *Artificial intelligence safety and security*. Chapman and Hall/CRC, 2018, pp. 99–112.

[173] S.-M. Moosavi-Dezfooli, A. Fawzi, and P. Frossard, "Deepfool: a simple and accurate method to fool deep neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2574–2582.

[174] P. Wang, Y. Li, S. Shekhar *et al.*, "Adversarial attacks on reinforcement learning based energy management systems of extended range electric delivery vehicles," *arXiv:2006.00817*, 2020.

[175] W. Gan, M. Shahidehpour, J. Guo *et al.*, "A tri-level planning approach to resilient expansion and hardening of coupled power distribution and transportation systems," *IEEE Transactions on Power Systems*, vol. 37, no. 2, pp. 1495–1507, 2021.

[176] K.-D. Lu and Z.-G. Wu, "Multi-objective false data injection attacks of cyber–physical power systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 9, pp. 3924–3928, 2022.

[177] Y. Li and Y. Wang, "False data injection attacks with incomplete network topology information in smart grid," *IEEE Access*, vol. 7, pp. 3656–3664, 2018.

[178] D. Xue, X. Jing, and H. Liu, "Detection of false data injection attacks in smart grid utilizing elm-based ocon framework," *IEEE Access*, vol. 7, pp. 31 762–31 773, 2019.

[179] O. Boyaci, A. Umunnakwe, A. Sahu *et al.*, "Graph neural networks based detection of stealth false data injection attacks in smart grids," *IEEE Systems Journal*, vol. 16, no. 2, pp. 2946–2957, 2021.

[180] J. Johnson, T. Berg, B. Anderson *et al.*, "Review of electric vehicle charger cybersecurity vulnerabilities, potential impacts, and defenses," *Energies*, vol. 15, no. 11, 2022.

[181] J. Johnson, B. Anderson, B. Wright *et al.*, "Cybersecurity for electric vehicle charging infrastructure," Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), Tech. Rep., 2022.

[182] H. Lee, S. H. Jeong, and H. K. Kim, "Otids: A novel intrusion detection system for in-vehicle network by using remote frame," in *2017 15th Annual Conference on Privacy, Security and Trust (PST)*. IEEE, 2017, pp. 57–5709.

[183] M. E. Verma, M. D. Iannacone, R. A. Bridges *et al.*, "Road: The real ornl automotive dynamometer controller area network intrusion detection dataset (with a comprehensive can ids dataset survey & guide)," *ArXiv*, vol. abs/2012.14600, 2020.

[184] D. L. Marino, C. S. Wickramasinghe, K. Amarasinghe *et al.*, "Cyber and physical anomaly detection in smart-grids," in *2019 Resilience Week (RWS)*, vol. 1. IEEE, 2019, pp. 187–193.

[185] California ISO, "Open access same-time information system (oasis)," http://oasis.caiso.com/mrioasis/logon.do, 2024, accessed: 08 May 2024.

[186] V. Mnih, K. Kavukcuoglu, D. Silver *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[187] T. Lillicrap, "Continuous control with deep reinforcement learning," *arXiv:1509.02971*, 2015.

[188] J. Schulman, F. Wolski, P. Dhariwal *et al.*, "Proximal policy optimization algorithms," *arXiv:1707.06347*, 2017.

[189] T. Haarnoja, A. Zhou, P. Abbeel *et al.*, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *International conference on machine learning*. PMLR, 2018, pp. 1861–1870.

[190] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," *arXiv:1412.6572*, 2014.

[191] A. Omara and B. Kantarci, "Generative adversarial networks to secure vehicle-to-microgrid services," in *2023 IEEE Virtual Conference on Communications (VCC)*, 2023, pp. 151–156.

[192] P. Huber, M. Ott, M. Friedli *et al.*, "Residential power traces for five houses: the ihomelab rapt dataset," https://doi.org/10.5281/zenodo.3581895, 18 Dec 2019, accessed: 08 Jun 2024.

[193] A. Omara and B. Kantarci, "On the impact of data integrity attacks on vehicle-to-microgrid services," in *2021 IEEE 26th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*. IEEE, 2021, pp. 1–7.

[194] Electric Reliability Council of Texas (ERCOT), "Backcasted (actual) load profiles - historical," https://www.ercot.com/mktinfo/loadprofile/alp, 2024, accessed: 12 Jun 2024.

[195] Hacking and Countermeasures Research Lab (HCRL), "Can dataset for intrusion detection (otids)," https://ocslab.hksecurity.net/Dataset/CAN-intrusion-dataset, 2017, accessed: 23 Jul 2024.

[196] ——, "Car-hacking dataset for the intrusion detection," https://ocslab.hksecurity.net/Datasets/car-hacking-dataset, 2018, accessed: 23 Jul 2024.

[197] E. Seo, H. M. Song, and H. K. Kim, "Gids: Gan based intrusion detection system for in-vehicle network," in *2018 16th annual conference on privacy, security and trust (PST)*. IEEE, 2018, pp. 1–6.

[198] H. S. Mavikumbure, V. Cobilean, C. S. Wickramasinghe *et al.*, "Physical anomaly detection in ev charging stations: Physics-based vs resnet ae," in *2023 IEEE 32nd International Symposium on Industrial Electronics (ISIE)*, 2023, pp. 1–7.

[199] Z. Liu, Y. Li, Q. Wang *et al.*, "Tscw-gan based fdias defense for state-of-charge estimation of battery energy storage systems in smart distribution networks," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 4, pp. 5048–5059, 2024.

[200] M. Li, M. Lenzen, F. Keck *et al.*, "Gis-based probabilistic modeling of bev charging load for australia," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3525–3534, 2018.

[201] Y. Liu, W. Liu, S. Gao *et al.*, "Fast charging demand forecasting based on the intelligent sensing system of dynamic vehicle under evs-traffic-distribution coupling," *Energy Reports*, vol. 8, pp. 1218–1226, 2022.

[202] M. B. Arias, M. Kim, and S. Bae, "Prediction of electric vehicle charging-power demand in realistic urban traffic networks," *Applied energy*, vol. 195, pp. 738–753, 2017.

[203] S. Powell, G. V. Cezar, E. Apostolaki-Iosifidou *et al.*, "Large-scale scenarios of electric vehicle charging with a data-driven model of control," *Energy*, vol. 248, p. 123592, 2022.

[204] D. A. R. de Jesús, P. Mandal, T. Senjyu *et al.*, "Unsupervised hybrid deep generative models for photovoltaic synthetic data generation," in *2021 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2021, pp. 1–5.

[205] J. Yoon, D. Jarrett, and M. Van der Schaar, "Time-series generative adversarial networks," *Advances in neural information processing systems*, vol. 32, 2019.

[206] Shanghai New Energy Electric Vehicle Monitoring Center, "The public data open platform for new energy vehicles," https://labs.shevdc.com/home, 2024, accessed: 13 Jun 2024.

[207] City of Boulder government, "City of boulder open data," https://open-data.bouldercolorado.gov/, 2021, accessed: 27 May 2024.

[208] Elia Group, "Transparency on grid data," http://www.elia.be/en/grid-data, 2024, accessed: 10 Jun 2024.

[209] Customer-Led Network Revolution (CLNR), "What was driving the project?" http://www.networkrevolution.co.uk/the-project/what-were-the-drivers/, 2024, accessed: 21 Jul 2024.