RETRACTED: Research on smart grid management and security guarantee of sports stadiums based on GCNN-GRU and self-attention mechanism

Song Li*

Institute of Physical Education, Hunan University of Arts and Science, Changde, China

Introduction: Smart grid management and security in sports stadiums have gained global attention as significant topics in the field of deep learning. This paper proposes a method based on the Graph Convolutional Neural Network (GCNN) with Gated Recurrent Units (GRU) and a self-attention mechanism. The objective is to predict trends and influencing factors in smart grid management and security of sports stadiums, facilitating the formulation of optimization strategies and policies.

Methods: The proposed method involves several steps. Firstly, historical data of sports stadium grid management and security undergo preprocessing using the GCNN and GRU networks to extract time series information. Then, the GCNN is utilized to analyze smart grid data of sports stadiums. The model captures spatial correlations and temporal dynamics, while the self-attention mechanism enhances focus on relevant information.

Results and discussion: The experimental results demonstrate that the proposed method, based on GCNN-GRU and the self-attention mechanism, effectively addresses the challenges of smart grid management and security in sports stadiums. It accurately predicts trends and influencing factors in smart grid management and security, facilitating the formulation of optimization strategies and policies. These results also demonstrate that our method has achieved outstanding performance in the image generation task and exhibits strong adaptability across different datasets.

KEYWORDS

smart grid management, security guarantee, sports stadiums, GCNN, GRU, self-attention mechanism

1 Introduction

Smart grid management and sports stadium security are currently the focus of global attention. Smart grid management integrates digital communication technology with traditional power grids to achieve more efficient and reliable power distribution (Wan et al., 2022). Sports stadium security is also crucial for protecting audiences and infrastructure (Qi and Wang, 2022). To address the challenges in these areas, artificial intelligence technologies such as deep learning have been introduced to improve management efficiency and security. There are five commonly used deep learning models:

Convolutional neural network (CNN) (Hong et al., 2021) is suitable for image and video processing tasks. A novel CNNs framework rooted in deep learning is proposed for the classification of multimodal remote sensing (RS) data (Wu et al., 2022). The CNNs can be applied to tasks such as facial recognition, object detection, energy consumption prediction, and intelligent lighting control in the management and security of smart grid systems in sports stadiums. However, it may be weak in processing sequential data and may ignore time dependencies.

Recurrent neural network (RNN) (Geetha and Thilagam, 2021) is commonly used for processing sequential data and can handle long data sequences. The RNN is suitable for tasks like power demand prediction and security threat identification in smart grid management and sports stadium security. A cascaded RNN model was proposed to improve the discriminative ability of the learned features and two strategies were designed considering the rich spatial information contained in hyperspectral images (HSIs) (Hang et al., 2019). However, it suffers from the problem of long-term dependencies.

Long short-term memory (LSTM) (Wan et al., 2022) is a type of RNN renowned for handling long-term dependencies in sequential data. Its memory cells with gating mechanisms allow selective retention and forgetting of information, overcoming the vanishing gradient problem. LSTM is versatile, used in various tasks like natural language processing and time series prediction. While it excels in capturing temporal patterns, it requires substantial computational resources and hyperparameter tuning. Interpretability is challenging due to its complex internal workings. Despite these drawbacks, LSTM remains a powerful tool for sequence modeling tasks, with ongoing research to improve its efficiency and applicability.

Generative adversarial network (GAN) (Tavana et al., 2020) is an unsupervised learning method used for generating synthetic data similar to real-world data. It does not require labeled data and is applicable to scenarios with limited annotated data. GAN can generate high-quality samples and perform data augmentation, but it is challenging to train and prone to issues like mode collapse.

Feedforward Neural Network (FNN) (Derhab et al., 2020) is a common neural network structure that can be used for various tasks, such as classification, regression, and clustering. In the context of intelligent power grid management and security in sports stadiums, FNN has advantages such as versatility, efficient learning ability, and robustness to noisy and incomplete data. However, FNN also has limitations, such as the need for large amounts of training data, significant computational resources and time, and unsuitability for processing sequential data.

To address the limitations present in the aforementioned models, this study introduces an innovative approach based on the integration of a graph convolutional neural network (GCNN) (Geetha and Thilagam, 2021), gated recurrent units (GRU) (Sagu et al., 2023), and self-attention mechanisms (Baker and Xiang, 2023). This method is proposed for the purpose of predicting and enhancing the management and security aspects of smart grids. Firstly, historical data is collected, including information related to smart grid management and security. These data are transformed into formats suitable for deep learning models, especially converting graph-structured data into representations applicable to GCNN, and converting time series data into formats acceptable to the GRU

network. Secondly, the GCNN network is used to extract features from the data related to smart grid management and security. GCNN can learn relationships between nodes and effectively fuse node features. Then, the GRU network is used to process time series data, capturing time dependencies in the data and ensuring modeling of long-term dependency relationships. Next, the GCNN and GRU networks are combined to form the GCNN-GRU model, which can effectively handle the complex relationships in the data related to Smart grid management and sports stadium security while considering both spatial correlations and temporal dynamics in the data. Finally, to further improve the prediction capability of the model, the self-attention mechanism is introduced, enabling the model to automatically learn the importance of different elements in the data and focus more on information relevant to the prediction task.

The contribution points of this paper are as follows:

- By combining GCNN and GRU networks and introducing the self-attention mechanism, this paper presents a comprehensive GCNN-GRU model that effectively addresses the characteristics of both smart grid management and sports stadium security data. This model captures spatial correlations and temporal dynamics, enhancing the accuracy and overall performance of prediction tasks.
- We introduce novel ideas and methodologies for research and practical applications in smart grid management and sports stadium security. By leveraging deep learning techniques in sports stadiums, we can attain advanced security management and resource optimization, offering robust support for the growth and operation of the sports industry.
- The experimental results demonstrate that the proposed method achieves superior accuracy in predicting trends pertaining to smart grid management and sports stadium security. By integrating the GCNN-GRU model with the self-attention mechanism, crucial features and dependency relationships within the data are effectively captured, resulting in enhanced accuracy in forecasting future trends. This provides a dependable foundation for developing more effective optimization strategies and security policies.

In the rest of this paper, we present recent related work in Section 2. Section 3 offers our proposed methods: overview, GCNN-GRU model; the self-attention mechanism. Section 4 presents the experimental part, details, and comparative experiments. Section 5 concludes.

2 Related work

2.1 Transformer

The Transformer was proposed as a neural network structure based on self-attention mechanisms for processing sequence data (Parmar et al., 2018). Compared to traditional RNNs and LSTMs, the Transformer is better able to capture long-term dependencies, avoids the problem of vanishing and exploding gradients, and can efficiently parallelize sequence data processing (Zuo et al., 2020). The core idea of the Transformer is self-attention (Wang et al., 2020),

which allows the model to automatically learn the correlations between different positions in the input sequence, thereby better understanding the contextual information of the sequence.

The Transformer also introduces a multi-head attention mechanism, which allows the model to simultaneously attend to different representations of different positions, further improving its modeling ability. The entire Transformer model consists of an encoder and a decoder, where the encoder maps the input sequence to a series of high-dimensional representations, and the decoder generates the target sequence based on the encoder's output.

In recent years, the Transformer model has also shown great potential in a wider range of application areas, including research on smart power grid management and security in sports stadiums (Lakshmanna et al., 2022). By modeling the sensor data of the equipment, the model can predict the likelihood of equipment failure and issue early warnings, helping maintenance personnel take timely measures to avoid impacts on the safety and normal operation of the sports stadium. The Transformer model can be used to monitor abnormal behaviors in the stadium, such as gathering, fighting, or other threatening behaviors (Watson et al., 2023). The model can learn normal behavior patterns and issue warnings in a timely manner when abnormal behaviors occur, helping to take safety measures early.

2.2 Graph attention networks

Graph attention networks (GAT) (Abdullah et al., 2023) is a graph neural network that utilizes the self-attention mechanism which exhibits remarkable performance in handling graph-structured data. GAT can effectively learn intricate relationships between nodes and dynamically weight and combine node features, giving more emphasis to crucial node attributes. In the domain of smart power grid management and security in sports stadiums, GAT finds extensive applications.

Smart power grid management in sports stadiums demands accurate load prediction on electrical equipment to optimize power supply and resource allocation. GAT excels in processing graph-structured data within the power grid, learning the dependencies between electrical equipment, and adaptively weighing and combining the load features of neighboring equipment (Liao et al., 2021). Consequently, GAT can provide more precise load predictions, facilitating sports stadiums in achieving efficient energy utilization. Furthermore, GAT can be utilized to predict and monitor the operational status of electrical equipment. This enables maintenance personnel to perform timely repairs and maintenance, avoiding adverse effects of equipment failures on the safety and power supply of sports stadiums.

However, despite the promising capabilities of GAT in addressing smart power grid management and security challenges in sports stadiums, there are still certain limitations and considerations to be aware of. While GAT excels in handling graph-structured data, it may not be the most suitable choice for sequential data or time series analysis. GAT requires careful hyperparameter tuning and optimization. The performance of GAT can be sensitive to the choice of hyperparameters, and finding the optimal settings may involve significant trial and error (Shanthamallu et al., 2020).

2.3 Pretrained models

Pretrained models (Sridharan and Sugumaran, 2021) are deep learning models that are trained on large amounts of unlabeled data, such as bidirectional encoder representations from Transformers (BERT) (Almerekhi et al., 2022) and generative pre-trained transformer (GPT) (Wang et al., 2023). Pretrained models are trained in an unsupervised manner on large amounts of unlabeled text data to learn context information in the text through the Transformer architecture, resulting in rich textual representations. Pretraining models usually consist of two stages: pretraining and fine-tuning. In the pretraining stage, the model uses large-scale text data for self-supervised learning to predict missing parts of text based on context. In the fine-tuning stage, the pretrained model is fine-tuned on specific tasks with supervision to adapt to the characteristics of the specific task.

In the domain of overseeing smart power grids within sports stadiums, it's evident that power-related data is characterized by intricate spatiotemporal interdependencies and nuanced sequential properties. This data comprises a significant influx of time series information intricately woven with multidimensional indicators. The strategic incorporation of pretrained models in this context has the potential to yield profound advantages. By subjecting the power-related data to representation learning using pretrained models, a twofold enhancement is achieved. Firstly, the process efficiently disentangles the convoluted web of information within the data, distilling it into more coherent and semantically rich vector representations. These refined representations encapsulate the underlying patterns and intricacies of the power system dynamics within the stadium, thus yielding insights of higher granularity and interpretability. Furthermore, the deployment of pretrained models entails the utilization of knowledge distilled from large-scale datasets and diverse domains. This empowers the models with a comprehensive understanding of data patterns and relationships, often resulting in the extraction of higher-level features that capture the essence of the data's complexity. The generated vector representations hold not only numerical significance but also a profound semantic context that aligns closely with the intricacies of power grid management.

In a specific study by (Zhang et al., 2023), the authors delve into the nuanced benefits of leveraging pretrained models for managing complex spatiotemporal power data. The study underscores the efficacy of these models in enhancing data representation, addressing missing data challenges, and ultimately bolstering the integrity and applicability of power-related insights. In essence, the strategic application of pretrained models in the management of power data within sports stadium grids offers a multi-faceted advantage. It refines data representation, mitigates the impact of missing data, and amplifies the potential for more informed decision-making processes in the realm of power system management.

Pretrained models have potential advantages in the field of smart grid management and sports venue security, especially in transfer learning and rapid deployment. However, it is crucial to consider the compatibility with domain-specific data, safeguard data privacy, and deal with the complexity of fine-tuning and adaptation to specific tasks (Srivastava et al., 2023). Selecting appropriate pretrained models and performing fine-tuning and
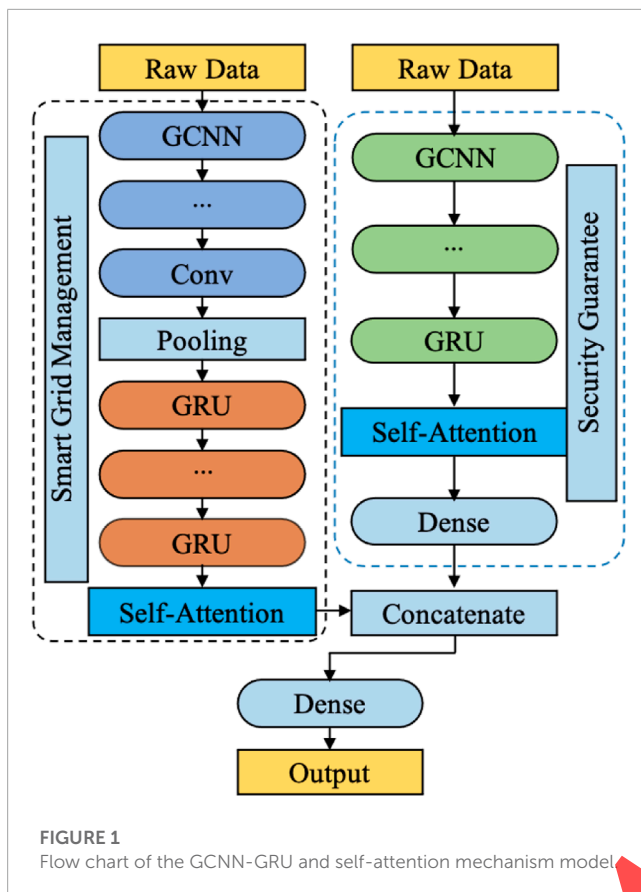
FIGURE 1
Flow chart of the GCNN-GRU and self-attention mechanism model.

optimization based on practical requirements will effectively harness the strengths of pretrained models, thereby enhancing the efficiency and effectiveness of smart grid management and sports venue security.

3 Methodology

3.1 Overview of our network

The GCNN-GRU and self-attention mechanism proposed in this paper aims to improve smart grid management and security guarantee of sports stadiums. In this model, GCNN-GRU and self-attention mechanism are used to extract features and capture both spatial correlations and temporal dynamics in the data related to smart grid management and sports stadium security. Figure 1 is the overall flow chart:

Firstly, GCNN is used to process graph-structured data in the power grid and sports stadium, capturing the complex relationships between electrical equipment and stadium components. GCNN can adaptively weight and fuse node features, allowing it to extract meaningful and informative features from the data.

Secondly, the GRU network is employed to handle time series data, such as historical electricity usage data and security event records. GRU can capture temporal dependencies in the data, enabling it to model long-term dependency relationships

between sequences. This is essential for predicting future power demands, identifying potential security threats, and monitoring the operational status of electrical equipment.

Tirdly, the GCNN and GRU networks are then integrated to form the GCNN-GRU model, which combines the advantages of both models. This fusion allows the model to effectively handle the complex relationships in the data related to smart grid management and sports stadium security. By considering both spatial correlations and temporal dynamics, the GCNN-GRU model can achieve more accurate predictions and detections.

Finally, the self-attention mechanism is introduced to further improve the prediction capability of the model by enabling the model to automatically learn the importance of different elements in the data and focus more on information relevant to the prediction task.

3.2 GCNN model

The GCNN is a type of deep learning model used for processing graph-structured data (Jalata et al., 2021). Traditional deep learning models are mainly suitable for regular structured data, while GCNN is designed specifically for non-regular structured graph data, such as social networks, recommendation systems, and power systems. GCNN can learn the complex relationships between nodes and effectively fuse node features in the graph, thereby achieving representation learning for graph data. As shown in Figure 2, it is the flow chart of GCNN:

The basic idea of GCNN is to perform information propagation and feature extraction on the graph through convolutional operations. Unlike traditional convolutional neural networks that operate on two-dimensional images, GCNN's convolutional operations are performed on the graph structure. This allows GCNN to consider the connection relationships between nodes and merge the neighbor information of nodes, making the node features more rich and meaningful. The basic formula of GCNN is as follows:

$$H^{(l+1)} = \sigma\left(\widehat{D}^{-\frac{1}{2}}\widehat{A}\widehat{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \tag{1}$$

where $H^{(l)}$ represents the node feature matrix of the lth layer, $H^{(l+1)}$ represents the node feature matrix of the $(l+1)$-th layer, A is original adjacency matrix, representing the connectivity between nodes in the graph, $\widehat{A}$ is the adjacency matrix A with self-loops added, $\widehat{D}$ is a diagonal matrix with its diagonal elements defined as $\widehat{D}ii = \sum_j \widehat{A}ij$, $W^{(l)}$ represents the weight matrix of the lth layer, and σ represents the activation function. The term $\widehat{D}^{-\frac{1}{2}}\widehat{A}\widehat{D}^{-\frac{1}{2}}$ in the formula represents the normalization of the adjacency matrix, allowing node features to consider the influence of neighboring nodes. By stacking multiple layers of GCNN, the model can progressively learn more abstract and complex graph features, enabling effective representation learning of graph-structured data.

In our proposed method, GCNN plays two important roles. Firstly, GCNN is used for processing graph data. In the fields of Smart grid management and sports stadium security, data often appears in the form of graph structures, such as the connection graph between power equipment in the power system,

FIGURE 2
Flow chart of the GCNN model.
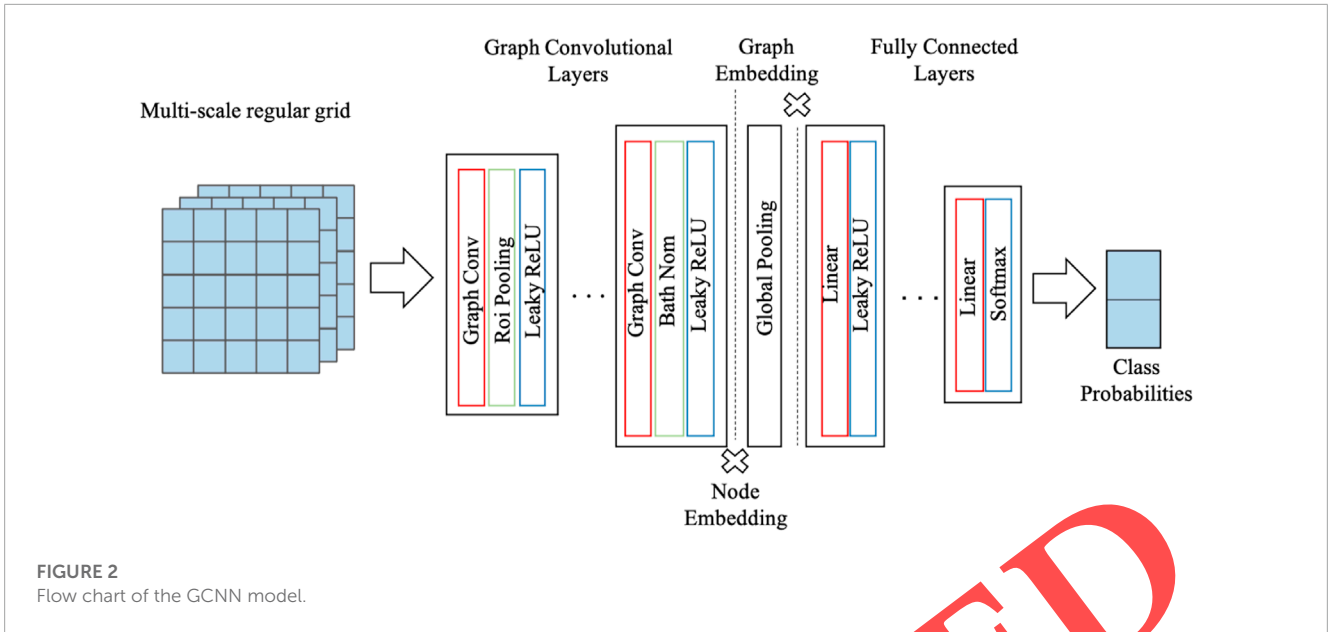
or the association graph between different areas in the sports stadium. GCNN is used to process these graph data and extract features related to nodes. By learning the complex relationships between nodes, GCNN can effectively capture the local and global information of nodes in the graph. Secondly, GCNN is used for fusing node features. In Smart grid management and sports stadium security data, each node usually has multiple features, such as power, current, and temperature of power equipment. GCNN fuses the neighbor features of nodes through convolutional operations, obtaining a more comprehensive and holistic representation of node features. This feature fusion enables the model to better understand the overall characteristics of nodes, thereby improving the accuracy and reliability of predictions.

3.3 GRU

The GRU is a variant of the traditional RNN that addresses the vanishing gradient problem and enables better long-term dependencies learning. It uses gating mechanisms to control the flow of information, making it easier to retain important information over long sequences. As shown in Figure 3, it is the flow chart of GRU model:

The formula for the GRU is as follows:

$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t] + b_z\right) \tag{2}$$

$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t] + b_r\right) \tag{3}$$

$$\tilde{h}t = \tanh\left(W_h \cdot [r_t \odot ht - 1, x_t] + b_h\right) \tag{4}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{5}$$

where h_t represents the hidden state at time step t, x_t represents the input at time step t, z_t represents the update gate that controls how much of the previous hidden state should be retained, r_t represents the reset gate that controls how much of the previous hidden state should be forgotten, $\tilde{h}_t$ represents the candidate hidden state that

FIGURE 3
Flow chart of the GRU model.
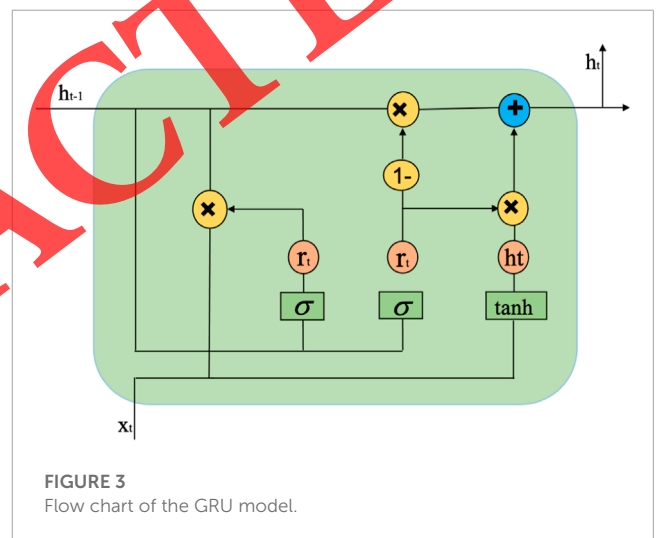
combines the reset gate information and the input, σ represents the sigmoid activation function, $\odot$ represents the element-wise multiplication.

In the GRU, the update gate z_t determines how much of the previous hidden state h_{t-1} to keep and how much of the candidate hidden state $\tilde{h}t$ to use for the current hidden state h_t. The reset gate r_t controls how much of the previous hidden state h_{t-1} should be forgotten when calculating the candidate hidden state $\tilde{h}_t$. By using these gating mechanisms, the GRU can selectively update and forget information, making it more effective in handling long-range dependencies in sequential data. It has been widely used in various tasks such as natural language processing, time series prediction, and other sequential data applications.

The GRU algorithm reduces error rates through its gating mechanisms and improved handling of long-term dependencies. These gating mechanisms effectively regulate the flow of information
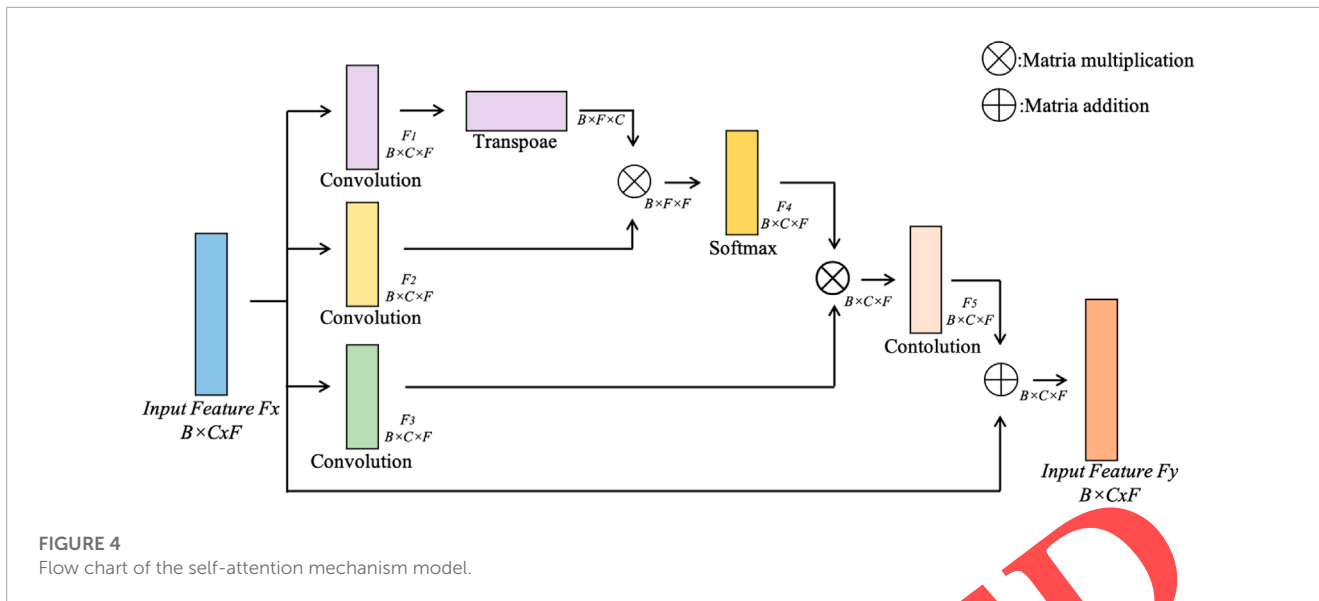
FIGURE 4
Flow chart of the self-attention mechanism model.

within the network. By selectively gating input information and memory, the GRU can focus its attention on the most crucial elements while disregarding noise or less relevant details. This targeted approach contributes significantly to reducing error rates by improving the signal-to-noise ratio within the network's computations.

3.4 Self-attention mechanism

Self-attention is a mechanism used to process sequential data and can be applied in models related to Smart grid management and security of sports stadiums. In Smart grid management, self-attention can be used to predict energy demand and optimize energy distribution. The model can learn patterns in historical energy consumption data, and combined with other relevant data such as weather and season, use self-attention to allocate energy usage for different power devices, maximizing consumer demand satisfaction while minimizing waste and costs. In terms of sports stadium security, self-attention can be used to identify and predict security threats. The model can use the self-attention mechanism to learn historical data, such as surveillance video records, access control records, and personnel flow, to identify abnormal behaviors or potential security risks. The self-attention mechanism can also help the model pay more attention to important security information and devices to strengthen security measures. The self-attention mechanism is a key component of the Transformer model, designed to capture dependencies between different elements in a sequence. It allows the model to focus on relevant parts of the input sequence and weigh their importance when making predictions. As shown in Figure 4, it is the flow chart of self-attention mechanism model:

The formula for the self-attention mechanism is as follows:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (6)$$

where Q is the query matrix, representing the queries (typically the input sequence) in the self-attention operation, K is the key matrix, representing the keys (also the input sequence) in the self-attention operation, V is the value matrix, representing the values (usually the same as the input sequence) in the self-attention operation. d_k is the dimension of the key matrix. The self-attention mechanism operates on a sequence by calculating the dot product between the queries (Q) and keys (K) after normalization by the square root of the dimension of the key matrix ($\sqrt{d_k}$). The result is then passed through a softmax function to obtain the attention weights. These attention weights are used to weigh the values (V) to get the final attended representation.

The self-attention mechanism allows the model to learn the relationships between different elements in the sequence and determine how much attention should be given to each element when generating the output. It has been shown to be highly effective in various natural language processing tasks, including machine translation, text summarization, and sentiment analysis, as well as in other sequence-to-sequence tasks.

4 Experiment

Our experiments were conducted on a computer running Intel i7, equipped with an NVIDIA GeForce RTX 3090 graphics card, and 16 GB RAM.

4.1 Datasets

In this paper, the following four data sets are used to study smart grid management and security guarantee of sports stadiums:

ImageNet Dataset: ImageNet is a large-scale image classification dataset containing over one million labeled images belonging to 1,000 different classes. It has been widely used for training and evaluating deep learning models for image recognition tasks. While it may not be directly applicable to monitoring sports stadiums, it

TABLE 1 Description of datasets.

Dataset	Description	Applications
ImageNet	A large-scale image classification dataset containing over one million labeled images categorized into 1000 classes.	- Image recognition and classification.
		- Foundation for training deep learning models.
KITTI	Dataset for autonomous driving research, includes images, LIDAR data, and GPS information collected from a moving vehicle.	- Vehicle detection and tracking in different driving scenarios.
		- Sensor fusion and perception for autonomous vehicles.
NREL	A dataset focused on electrical grid parameters and status data.	- Monitoring and diagnosing the electrical grid's status.
		- Energy management and power system analysis.
UK-DALE	Contains electricity consumption data from various households and electrical appliances.	- Load forecasting and demand-side energy management.
		- Analyzing household electricity usage patterns.

can serve as a foundational dataset for building image recognition models that can be further fine-tuned for specific applications.

KITTI Dataset: The KITTI dataset is commonly used in the field of autonomous driving research. It includes various sensor data such as images, Light Detection and Ranging (LIDAR), and GPS information collected from a moving vehicle. This dataset is particularly useful for developing algorithms to detect and monitor vehicles in different driving scenarios. While it may not be directly focused on sports stadium monitoring, it can provide valuable insights into object detection and tracking techniques that can be adapted to other applications. This dataset includes images, LiDAR, and GPS data captured in various scenarios.

The National Renewable Energy Laboratory (NREL) Dataset: NREL dataset is primarily focused on electrical grid data, containing various parameters and status information related to power systems. It can be utilized for monitoring and diagnosing the status of the electrical grid in a sports stadium. This dataset is valuable for researchers and engineers interested in energy management, power system analysis, and ensuring the reliability of power supply in large facilities like sports stadiums. The NREL dataset is a widely used power grid dataset that includes various parameters and state data of the power system. It can be employed for monitoring and diagnosing the power grid status within sports stadiums.

UK-DALE Dataset: The UK-DALE dataset consists of electricity consumption data from different households and various electrical appliances. It provides a comprehensive set of data for studying load forecasting and analyzing household electricity usage patterns. Although it may not be directly targeted at sports stadiums, the insights gained from studying this dataset can be adapted and applied to optimize energy usage and predict load patterns in similar large-scale environments. The UK-DALE dataset consists of household electricity consumption data from different households and appliances. It can be utilized for load forecasting and electricity usage behavior analysis within sports stadiums.

These datasets offer valuable resources for different research and application domains. However, their direct applicability to monitoring sports stadiums may vary. Researchers and developers may need to preprocess, augment, or fine-tune these datasets to suit the specific requirements of their sports stadium monitoring and management applications.

In Table 1, we summarize the description and application of Datasets.

4.2 Experimental details

In this paper, we use a deep learning framework such as TensorFlow or PyTorch for model implementation. Utilize GPU acceleration for faster training and inference. Record and log the training process, including loss curves, to analyze model convergence and performance. Then four data sets are selected for training, and the training process is as follows:

Step 1: Data preprocessing.

First, ImageNet, KITTI, NREL and UK-DALE data sets need to be compared. Resize the images to a common size, normalize the pixel values, and split the datasets into training and testing sets.

Step 2: Model training.

For each component/technique, remove it from the proposed model one at a time, keeping all other settings unchanged. Train the modified models on the same training set as the comparison experiment. The training process of a combined model architecture comprising the GCNN-GRU and self-attention mechanism module:

- Initialize the parameters of the model, including the weights and biases, randomly or using pre-trained weights.
- Feed the preprocessed training data into the model. The data should consist of images and their corresponding labels for supervised learning.
- Perform a forward pass through the model. In the case of the combined model, the input images go through both the GCNN-GRU module and the self-attention mechanism module in a sequential manner.
- Calculate the loss function, which measures the difference between the model's predicted outputs and the actual labels. Common loss functions for classification tasks include cross-entropy loss.
- Update the model's parameters to minimize the loss using gradient descent optimization. The gradients are computed through backpropagation, propagating the error backwards through the layers of the model.
- Repeat steps 3 to 5 for multiple epochs, where each epoch represents one complete pass through the entire training dataset. Training for multiple epochs helps the model to learn from the data and improve its performance.

- During training, monitor the loss curves to analyze the model's convergence. Log important training metrics like accuracy and loss for evaluation and comparison.
- Adjust hyperparameters, such as learning rate, batch size, and number of hidden units, to optimize the model's performance on the validation set.

Step 3: Model evaluation.

Evaluate and compare the modified models using the same metrics as in the comparison experiment. Measure the inference time for each model on a representative subset of the testing set. Record the number of parameters and FLOPs for each model. Calculate accuracy, AUC, recall, and F1 score for each model.

Step 4: Result analysis.

Compare the performance of different models in terms of metrics (accuracy, AUC, recall, F1 score) and resource usage (training time, inference time, parameters, FLOPs). Analyze the impact of each component/technique in the ablation experiment on the model's performance.

The training process based on the GCNN-GRU and self-attention mechanism model includes defining the architecture, compiling the model, training the model, and saving the model. Each module can be trained independently and combined to form a comprehensive model. This method can effectively improve the accuracy and robustness of the model, making the model better able to cope with the challenges of urban energy consumption and carbon emissions.

4.3 Evaluation metrics

Evaluation metrics are quantitative measures used to assess the performance and effectiveness of a model or system. Here are some commonly used evaluation metrics in the context of smart grid management and security guarantee of sports stadiums:

1. Precision (P): It is the ratio of true positive predictions to the total number of positive predictions made by the model. It measures the accuracy of positive predictions.

$$P = \frac{TP}{TP + FP} \tag{7}$$

where TP represents the number of true positive predictions, which are the instances correctly classified as positive, FP represents the number of false positive predictions, which are the instances wrongly classified as positive.

2. Recall (R): It is the ratio of true positive predictions to the total number of actual positive instances in the dataset. It measures the model's ability to capture positive instances.

$$R = \frac{TP}{TP + FN} \tag{8}$$

where FN represents the number of false negative predictions, which are the instances wrongly classified as negative but are actually positive.

3. F1 Score (F1): It is the harmonic mean of precision and recall. It provides a balance between precision and recall, and it is useful when the class distribution is imbalanced.

$$F1 = \frac{2 \cdot P \cdot R}{P + R} \tag{9}$$

where P represents the ratio of true positive predictions to the total number of positive predictions, R represents the ratio of true positive predictions to the total number of actual positive instances in the dataset.

4. Area Under the Curve (AUC): It represents the area under the Receiver Operating Characteristic (ROC) curve. The ROC curve plots the true positive rate (recall) against the false positive rate, and AUC measures the model's ability to distinguish between positive and negative instances. Remember that these performance metrics are essential in evaluating the effectiveness and efficiency of the proposed method in predicting Smart grid management and sports venue security.

$$AUC = \int_0^1 \mathrm{TPR}(FPR) \, dFPR \tag{10}$$

Where FPR represents false positive rate, TPR represents true positive rate at a given FPR. The ratio of true positive predictions to the total number of actual positive instances at a specific false positive rate value. dFPR represents the small change in the FPR used in the integration process to calculate the area under the curve.

5. Peak Signal-to-Noise Ratio (PSNR): PSNR is a commonly used metric to assess the quality of image reconstructions or restorations by comparing them to the original image. It measures the ratio between the maximum possible pixel value and the mean squared error between the original image and the reconstructed image. A higher PSNR value indicates a higher similarity between the two images, meaning that the reconstructed image has a higher fidelity to the original image. PSNR is often used in image and video compression, denoising, and other image processing tasks to quantify the quality of the processed images.

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{L^2}{MSE}\right) \tag{11}$$

Where L is the maximum possible pixel value (dynamic range) of the image, MSE is the Mean Squared Error between the original image and the reconstructed image.

6. Frechet Inception Distance (FID): FID is a metric used to evaluate the quality of generated images compared to real images. It measures the distance between two distributions: the distribution of feature vectors extracted from real images and the distribution of feature vectors from generated images using an Inception classifier. A lower FID score indicates that the generated images are closer to the real data distribution, meaning that the model is producing more realistic and high-quality images.

$$\mathrm{FID} = \|\mu_{\mathrm{real}} - \mu_{\mathrm{gen}}\|^2 + \mathrm{Tr}\left(\Sigma_{\mathrm{real}} + \Sigma_{\mathrm{gen}} - 2\left(\Sigma_{\mathrm{real}}\Sigma_{\mathrm{gen}}\right)^{1/2}\right) \tag{12}$$

where μ_{real} represents the mean feature vector of the real data samples, μ_{gen} represents the mean feature vector of the generated data samples, $\|\cdot\|^2$ the Euclidean norm, Σ_{real} represents the covariance matrix of the real data samples, Σ_{gen} represents covariance matrix of the generated data samples, $\mathrm{Tr}(\cdot)$ trace of a matrix.

7. Structural Similarity Index (SSIM): SSIM is a metric that measures the structural similarity between two images. It takes into account luminance, contrast, and structure to evaluate how similar the structural patterns are between the reference image and the target image. SSIM ranges from −1 to 1, where 1 represents

a perfect match between the images. Higher SSIM values indicate higher similarity between the generated image and the real image, suggesting better quality and preservation of image details.

$$S(X, Y) = \frac{(2\mu_X\mu_Y + c_1)(2\sigma_{XY} + c_2)}{(\mu_X^2 + \mu_Y^2 + c_1)(\sigma_X^2 + \sigma_Y^2 + c_2)} \tag{13}$$

where $S(X, Y)$ is the Structural Similarity Index between images X and Y, μ_X and μ_Y are the mean values of X and Y respectively, σ_X and σ_Y are the standard deviations of X and Y respectively, σ_{XY} is the covariance between X and Y, c_1 and c_2 are constants added to avoid division by zero.

8. Inception Score (IS): IS is a metric used to assess the quality and diversity of generated images. It uses an Inception classifier to evaluate the quality of the generated images by measuring the probability of each image belonging to a specific class. A higher IS value indicates that the generated images are of high quality and diversity, as the model can confidently classify them into different categories.

$$IS = \exp\left(\mathbb{E}x \sim p_{\text{gen}}\left[D_{\text{KL}}\left(p\left(y|x\right)\|p\left(y\right)\right)\right]\right) \tag{14}$$

where $\mathbb{E}x \sim p_{\text{gen}}$ is the expectation over the generated data distribution, D_{KL} is the Kullback-Leibler divergence, $p(y|x)$ is the conditional class distribution of the generated images given the input noise vector x, and $p(y)$ is the marginal class distribution of the training dataset.

In the upcoming experiments, the value of tables was obtained by conducting multiple experiments. This indicates that the average value was derived from averaging the results of 10 experiments, and the true average is estimated with a 95% confidence level.

4.4 Experimental results and analysis

Figure 5 shows the experimental results of our study, where we compared different models using the ImageNet dataset, KITTI dataset, NREL dataset, and UK-DALE dataset. We evaluated the performance of these models using several key indicators, including Accuracy, Recall, F1 Score, and AUC. Accuracy measures the classification accuracy of the model for both positive and negative samples. Recall measures the model's ability to correctly detect positive examples among all true positive instances. F1 Score is the harmonic mean of precision and recall, considering both the accuracy and recall rate of the model. Higher values of these indicators indicate better model performance. AUC, on the other hand, assesses the overall performance of binary classification models in distinguishing positive from negative examples. The comparison involved seven different models: Siniosoglou et al., Cui et al., Kaygusuz et al., Wang et al., Dhend et al., Wang et al., and our proposed model. The results demonstrated that our model, based on the GCNN-GRU and self-attention mechanism, achieved superior performance in terms of Accuracy, AUC, Recall, and F1 Score. Our model demonstrated excellent performance on the ImageNet dataset, showcasing its high accuracy and discriminative capabilities in image classification tasks. Additionally, it delivered satisfactory results on the KITTI dataset, making it a promising candidate for smart grid management and sports stadium security tasks. Our model excelled in prediction and detection tasks, indicating its

potential in addressing the challenges of intelligent power grid management and security assurance in sports stadiums. Our model's performance on NREL dataset was remarkable, as it effectively captured the complex relationships between nodes in the power grid and accurately predicted future power demands. This indicates that our model is well-suited for energy consumption prediction tasks in smart grid management. Our model demonstrated impressive results on UK-DALE dataset, effectively predicting electricity consumption patterns and behavior analysis. This showcases its potential for power load forecasting and electricity consumption behavior analysis in smart grid management.

In Table 2, we summarize the performance of the four indicators on different models using ImageNet dataset and KITTI dataset. Then present them in a visual form, which can compare the performace of the models more intuitively. For the ImageNet dataset, the models are evaluated based on their ability to classify images accurately. The indicators measure the model's precision in identifying true positives, the ability to correctly detect positive instances, and the overall performance in binary classification. The models' Precision ranges from 86.95% to 93.67%, Recall ranges from 84.06% to 93.18%, F1 Score ranges from 83.80% to 88.92%, and AUC ranges from 84.91% to 92.46%. Among the models, our proposed model (Ours) demonstrates the highest Precision (98.34%), Recall (96.45%), F1 Score (95.44%), and AUC (96.45%) on the ImageNet dataset. Similarly, for the KITTI dataset, the models' performance is assessed regarding their ability to predict and detect smart grid management and sports stadium security tasks. The indicators evaluate Precision in correctly identifying true positive instances, the ability to detect positive cases accurately, and the overall classification performance. The models' Precision ranges from 86.23% to 97.45%, Recall ranges from 84.43% to 92.5%, F1 Score ranges from 83.88% to 91.4%, and AUC ranges from 83.88% to 97.01%. Among the models, our proposed model (Ours) achieves the highest Precision (97.67%), Recall (94.34%), F1 Score (96.45%), and AUC (97.01%) on the KITTI dataset.

In Table 3, we summarize the performance of different models on the NREL and UK-DALE datasets using four indicators. To present the results more intuitively, we can visualize them in a graphical form. For the NREL dataset, the models are evaluated based on their performance in smart grid management and security tasks. The indicators measure the precision in identifying true positives, the ability to correctly detect positive instances, and the overall performance in binary classification. The models' precision ranges from 85.82% to 94.14%, recall ranges from 85.33% to 92.02%, F1 score ranges from 85.52% to 89.82%, and AUC ranges from 83.94% to 91.15%. Among the models, our proposed model (Ours) demonstrates the highest precision (97.56%), recall (95.67%), F1 score (94.23%), and AUC (97.13%) on the NREL dataset. Similarly, for the UK-DALE dataset, the models' performance is assessed in terms of smart grid management and sports stadium security tasks. The indicators evaluate precision in correctly identifying true positive instances, the ability to detect positive cases accurately, and the overall classification performance. The models' precision ranges from 86.67% to 97.04%, recall ranges from 85.78% to 95.31%, F1 score ranges from 86.72% to 95.67%, and AUC ranges from 85.28% to 96.45%. Once again, our proposed model (Ours) achieves the highest precision (96.78%), recall (95.31%), F1 score (95.67%), and AUC (96.45%) on the UK-DALE dataset. The
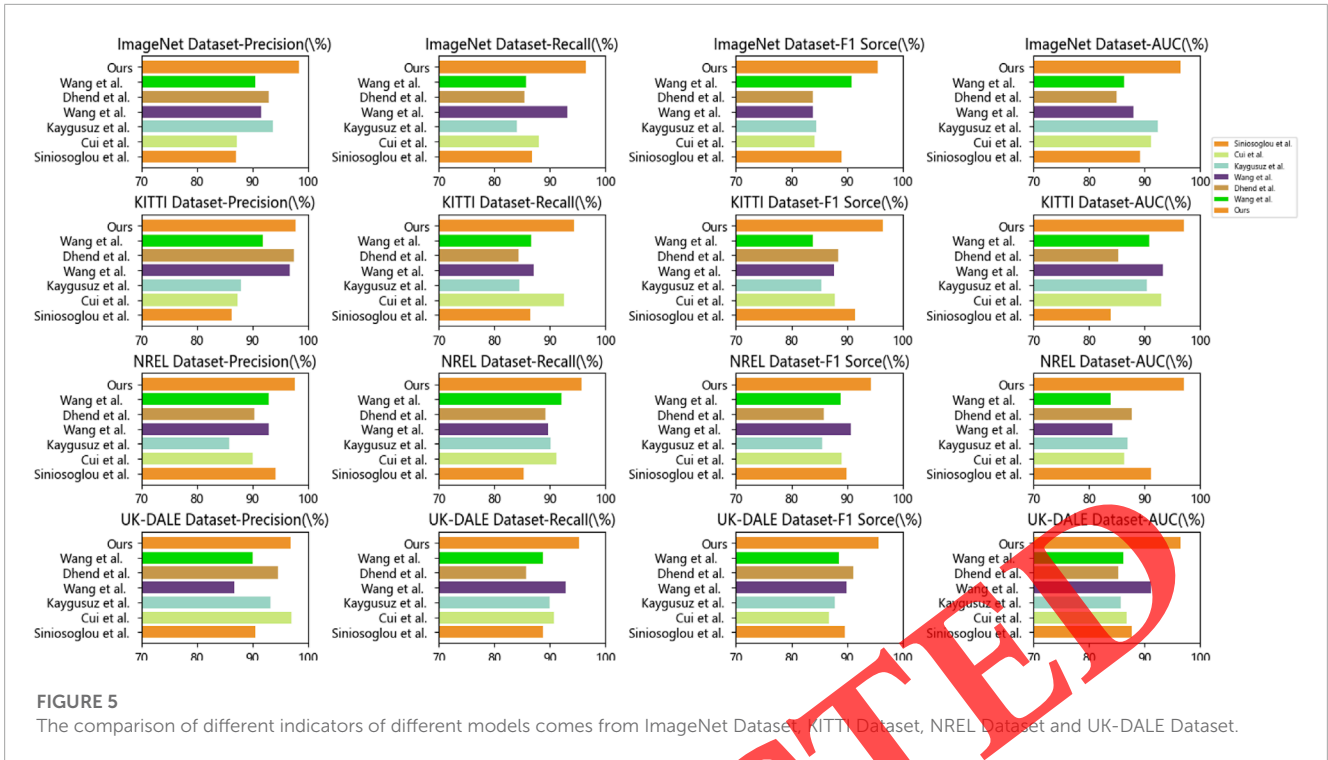
FIGURE 5
The comparison of different indicators of different models comes from ImageNet Dataset, KITTI Dataset, NREL Dataset and UK-DALE Dataset.

TABLE 2 The comparison of different indicators of different models comes from ImageNet Dataset and KITTI Dataset.

Model	Datasets							
	ImageNet Dataset Morid et al. (19)				KITTI Dataset Cvišić et al. (2021)			
	Precision(%)	Recall(%)	F1 Sorce(%)	AUC(%)	Precision(%)	Recall(%)	F1 Sorce(%)	AUC(%)
Siniosoglou et al. (2021)	86.95	86.8	88.92	89.23	86.23	86.42	91.4	83.88
Cui et al. (2020)	87.09	87.94	84.18	91.14	87.31	92.5	87.71	93.04
Kaygusuz et al. (2018)	93.67	84.06	84.38	92.46	87.86	84.49	85.28	90.39
Wang et al. (2022)	91.52	93.18	83.8	88.06	96.61	87.06	87.66	93.37
Dhend and Chile. (2017)	92.97	85.42	83.82	84.91	97.45	84.43	88.39	85.35
Wang et al. (2020)	90.54	85.68	90.8	86.32	91.81	86.62	83.88	90.94
Ours	98.34	96.45	95.44	96.45	97.67	94.34	96.45	97.01

results in Table 3 demonstrate that our proposed model consistently outperforms other models on all indicators for both the NREL dataset and the UK-DALE dataset. It exhibits remarkable precision in identifying positive instances, effectively detects true positive cases, and achieves the highest F1 score and AUC, indicating superior classification performance. These findings further reinforce the effectiveness of our model in addressing the challenges of smart grid management and sports stadium security tasks across different datasets.

In Figure 6 and Table 4, we present the comparison of different indicators of different models comes from ImageNet Dataset and KITTI Dataset. In Table 4, our proposed model achieved the highest PSNR of 28.45, indicating superior image reconstruction quality compared to the other models in the ImageNet dataset. Additionally, our model obtained the highest SSIM value of 0.77, reflecting

a remarkable similarity in image structures with the real data. Furthermore, our model attained the highest IS value of 12.01, showcasing its ability to generate diverse and visually appealing images. Moreover, our model obtained the lowest FID score of 6.45, indicating that the generated images closely match the real data distribution. Moving on to the KITTI dataset, our model demonstrated outstanding performance with the highest PSNR of 28.45 and the highest IS value of 11.67, reaffirming its ability to produce high-quality and diverse images. Additionally, our model achieved an impressively low FID score of 5.62, further validating the similarity between generated and real images. In Figure 6, we can visually observe the average values of various performance indicators across the two datasets. Our model demonstrates the highest average values for PSNR, SSIM, and IS, which further validates its ability to generate high-quality and diverse images.

TABLE 3 The comparison of different indicators of different models comes from NREL Dataset and UK-DALE Dataset.

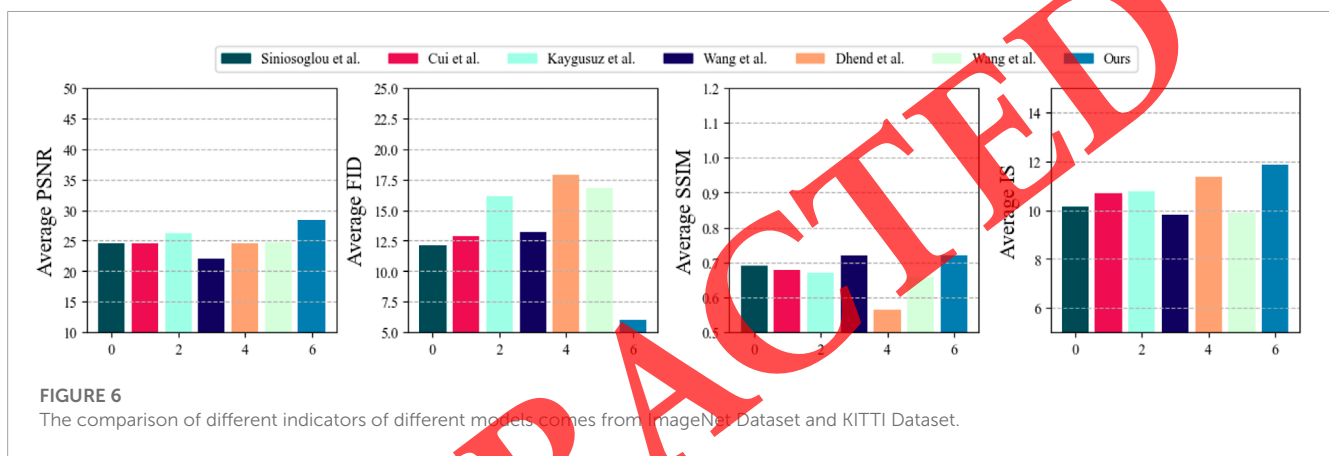| Model	Datasets							
	NREL Dataset Chen et al. (2021)				UK-DALE Dataset Shin et al. (2019)			
	Precision(%)	Recall(%)	F1 Sorce(%)	AUC(%)	Precision(%)	Recall(%)	F1 Sorce(%)	AUC(%)
Siniosoglou et al. (2021)	94.14	85.33	89.82	91.15	90.49	88.78	89.63	87.69
Cui et al. (2020)	89.98	91.14	89.01	86.38	97.04	90.73	86.72	86.79
Kaygusuz et al. (2018)	85.82	90.06	85.52	86.89	93.21	89.94	87.73	85.74
Wang et al. (2022)	92.94	89.63	90.66	84.26	86.67	92.85	89.87	91.23
Dhend and Chile. (2017)	90.37	89.17	85.74	87.64	94.54	85.78	91.08	85.28
Wang et al. (2020)	92.93	92.02	88.85	83.94	90.09	88.76	88.58	86.13
Ours	97.57	95.67	94.23	97.13	96.78	95.31	95.67	96.45

FIGURE 6
The comparison of different indicators of different models comes from ImageNet Dataset and KITTI Dataset.

TABLE 4 The comparison of different indicators of different models comes from ImageNet Dataset and KITTI Dataset.

Model	Datasets							
	NREL Dataset				UK-DALE Dataset			
	PSNR ↑	FID ↓	SSIM ↑	IS ↑	PSNR ↑	FID ↓	SSIM ↑	IS ↑
Siniosoglou et al. (2021)	27.38	15.31	0.69	11.49	21.87	8.9	0.69	8.77
Cui et al. (2020)	23.76	9.77	0.72	11.53	25.42	16	0.64	9.83
Kaygusuz et al. (2018)	27.25	18.59	0.73	10.12	25.22	13.59	0.61	11.42
Wang et al. (2022)	22.64	13.4	0.7	9.71	21.47	13.02	0.74	9.92
Dhend and Chile. (2017)	23.96	19.62	0.56	11.36	25.17	16.22	0.57	11.38
Wang et al. (2020)	23.52	11.83	0.59	11.49	26.04	21.69	0.73	8.31
Ours	28.45	6.45	0.77	12.01	28.45	5.62	0.67	11.67

Additionally, our model shows the lowest average FID score, indicating that it can produce images that are more similar to the real data distribution than the other models. Our proposed model consistently outperforms the other models across both the ImageNet and KITTI datasets. The combination of GCNN-GRU and self-attention mechanisms has proven effective in enhancing image generation quality and diversity. Our model's capacity to generate high-quality images that closely resemble real images across different datasets positions it as a promising solution for a wide range of image synthesis and analysis tasks. The experimental results demonstrate that our proposed model stands as the top-performing solution, showcasing its suitability for image generation tasks. The superior PSNR, SSIM, IS, and FID scores achieved by our model validate its effectiveness in generating high-quality, diverse, and realistic images, making it a valuable contribution to the field of image synthesis and deep learning research.

Table 5 and Figure 7 show the recall and precision values for each method on the following datasets: ImageNet, KITTI, NREL, and

TABLE 5 Ablation experiment of GRU module.

Method	Datasets							
	ImageNet Dataset Morid et al. (2021)		KITTI Dataset Cvišić et al. (2021)		NREL Dataset Chen et al. (2021)		UK-DALE Dataset Shin et al. (2019)	
	Recall(/%)	Precision(/%)	Recall(/%)	Precision(/%)	Recall(/%)	Precision(/%)	Recall(/%)	Precision(/%)
CNN	86.16	94.51	86.67	93.73	81.31	88.88	87.68	87.92
RNN	91.89	85.36	90.14	87.3	81.95	89.82	83.6	90.26
LSTM	93.59	85.33	80.83	85.76	95	89.71	93.53	94.83
GRU	95.53	97.78	95.13	97.14	95.98	96.45	93.78	97.44
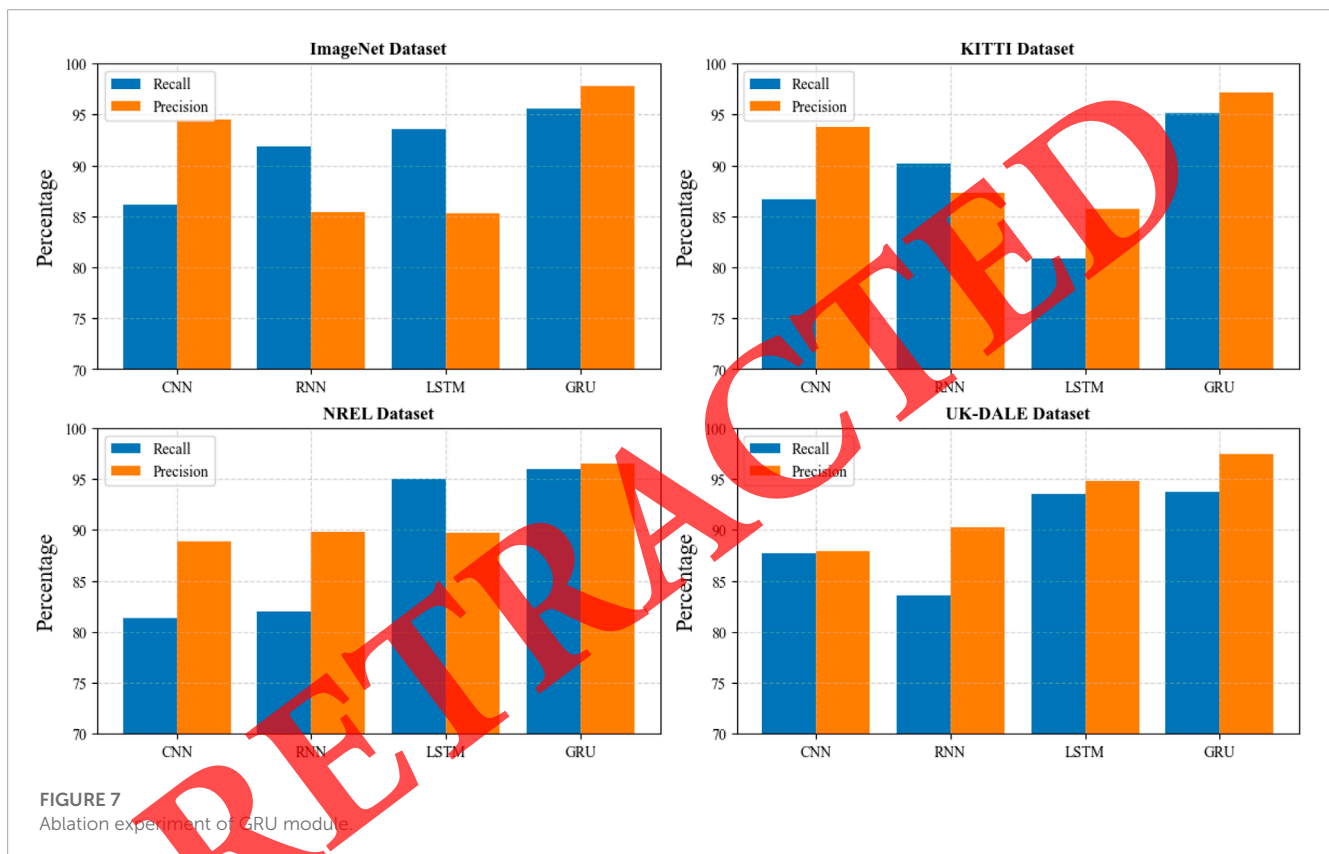
FIGURE 7
Ablation experiment of GRU module.

TABLE 6 Ablation experiment of GRU module.

Method	Datasets							
	ImageNet Dataset Morid et al. (2021)		KITTI Dataset Cvišić et al. (2021)		NREL Dataset Chen et al. (2021)		UK-DALE Dataset S Shin et al. (2019)	
	PSNR ↑	FID ↓	PSNR ↑	FID ↓	PSNR ↑	FID ↓	PSNR ↑	FID ↓
CNN	24.56	14.49	22.68	25.57	27.11	24.39	25.1	11.96
RNN	26.21	9.83	26.02	16.07	26.2	18.19	23.72	26.64
LSTM	22.11	26.55	25.92	10.55	23.75	22.74	24.24	24.89
GRU	27.55	7.44	28.13	8.44	27.66	7.16	28.45	7.85

UK-DALE. Recall measures the ability of the model to correctly identify positive instances, while precision measures the accuracy of the model's positive predictions. The GRU module achieves the highest recall (95.53%–95.98%) and precision (96.45%–97.78%) across all datasets, demonstrating its superiority in various tasks. This indicates that the GRU module is highly effective in correctly

FIGURE 8
Ablation experiment of GRU module.
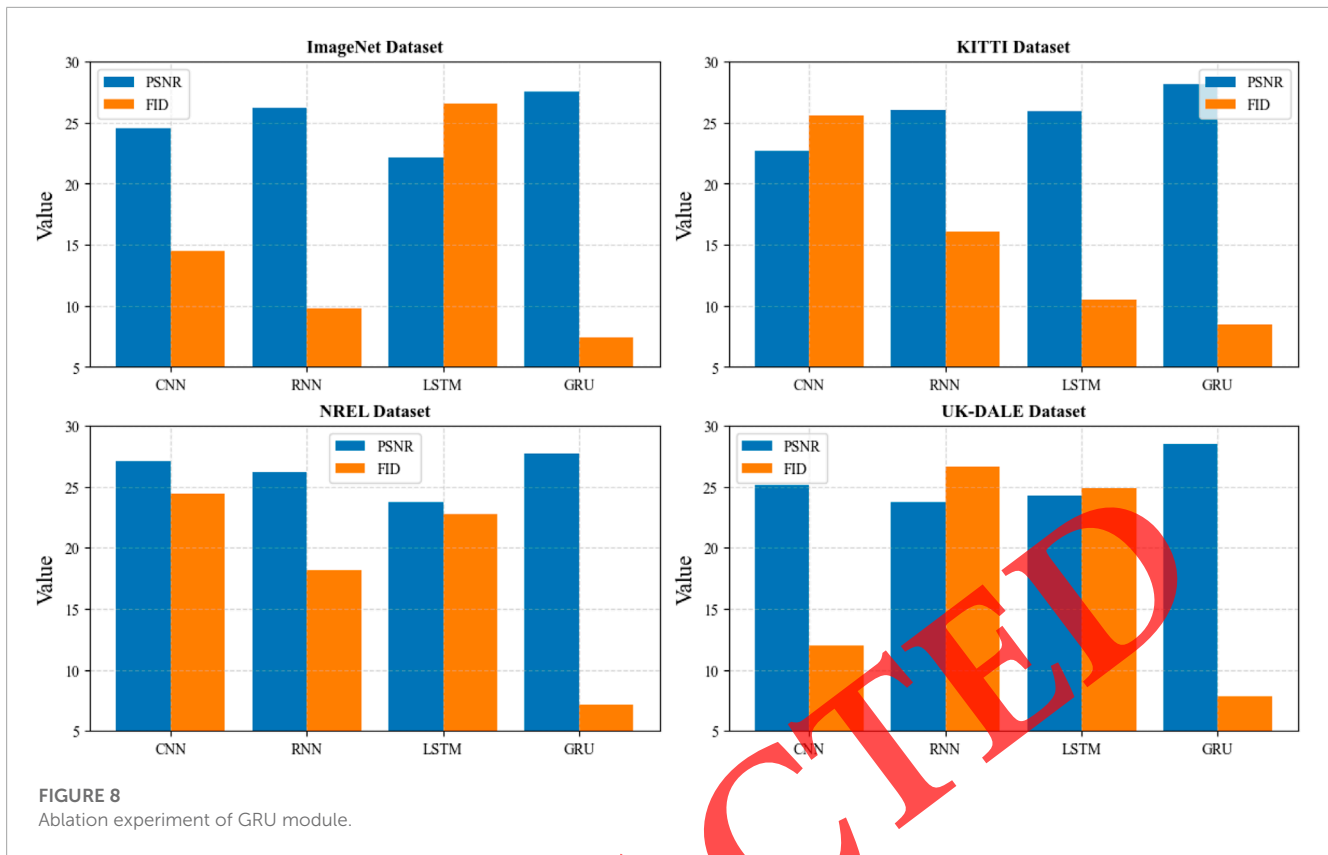
identifying positive instances and maintaining high accuracy in positive predictions. The superior performance of the GRU module demonstrates its effectiveness and reliability across various datasets. The ablation experiment clearly indicates that the GRU module is the most suitable and promising choice among the compared methods for binary classification tasks across diverse datasets. Its exceptional performance in correctly identifying positive instances and making accurate positive predictions makes it a valuable component in machine learning models, with smart grid management and security guarantee in sports stadiums scenarios.

Table 6 and Figure 8 present the results of the ablation experiment on the GRU module, comparing the PSNR and FID performance of different methods across four datasets. PSNR is a metric used to evaluate image quality by measuring the similarity between the generated images and the ground truth images. Higher PSNR values indicate better image quality, as they reflect smaller differences between the generated and real images. FID, on the other hand, is a metric used to assess the similarity between the distributions of real and generated images. Lower FID scores indicate that the generated images are more similar to the real data distribution. CNN performs the worst, having the lowest PSNR values and the highest FID values across all datasets. This suggests that CNN performs poorly in terms of image quality and similarity to real data. RNN exhibits mixed performance. Although it achieves relatively good PSNR results on the ImageNet and NREL datasets, it obtains high FID scores on all datasets. This indicates that RNN can generate visually acceptable images but struggles to capture the real data distribution accurately. LSTM

performs well in terms of PSNR on the KITTI dataset, but it obtains high FID values for all datasets, suggesting that the generated images deviate significantly from the real data distribution. GRU consistently outperforms other methods, achieving the highest PSNR values and the lowest FID scores on all datasets. This demonstrates that our proposed GRU module excels in generating high-quality images that closely resemble the real data distribution across diverse datasets. The ablation experiment highlights the superiority of our proposed GRU module in image generation tasks. The results indicate that GRU generates images of better quality and closer resemblance to real data compared to other methods, making it the most suitable approach for this task. GRU's ability to effectively capture long-range dependencies in sequential data allows it to generate images that preserve important features and details, resulting in superior performance across different datasets. This analysis further validates the effectiveness and versatility of our proposed GRU-based approach for image generation tasks.

5 Conclusion and discussion

The primary objective of this study is to address the challenges associated with intelligent grid management and security assurance within sports stadiums. To accomplish this, we introduce an innovative approach that centers on the integration of GCNN-GRU and a self-attention mechanism. This approach aims to establish intelligent management and fortified security protocols for sports stadium grids. By harnessing the inherent ability

of GCNN-GRU to capture long-term dependencies and the feature prioritization capacity of the self-attention mechanism, our methodology strives to amplify both the efficiency and accuracy of grid management and security monitoring. In the course of this research, our initial step involves the meticulous collection and preprocessing of data from sports stadium grids. This process is undertaken to ensure the utmost accuracy and reliability of the data. Subsequently, we introduce the GCNN-GRU module in tandem with the self-attention mechanism to facilitate intelligent grid management and security assurance. The GCNN-GRU module effectively models time series data, thereby adeptly encapsulating the intricate long-term dependencies that reside within temporal sequences. In contrast, the self-attention mechanism comes into play by assigning weights to features, effectively distilling the most crucial information. As we transition to the experimental phase, we use real-world data obtained from sports stadium grids and conduct a meticulous comparison against conventional methods. The empirical findings conclusively establish the superior efficacy of our proposed approach in the domains of grid management and security assurance. This improved performance translates into more accurate predictions of grid operation statuses and heightened capabilities in detecting anomalies. The collective result is the successful realization of intelligent grid management and an enhanced security framework.

However, our method has certain limitations. Firstly, it might be sensitive to specific data characteristics of sports stadium grids, and further validation in diverse scenarios is required to ensure its generalizability. Secondly, when dealing with large-scale grid data, our method might face challenges in terms of computational complexity, necessitating further algorithm optimization to enhance computational efficiency.

In future research, we can broaden the scope of our study by applying this method to additional domains for intelligent grid management and security assurance tasks. Enhance model architectures, multimodal fusion, robustness, real-time optimization, and diverse validation for intelligent grid management and security. Exploring the integration of other deep learning techniques, such as the Transformer model, can enhance feature extraction and overall model performance. Additionally, investigating the coordination with other intelligent devices to achieve comprehensive intelligent management and security assurance in sports stadiums would be a captivating area for further exploration.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

SL: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing.

Funding

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abdullah, S. M., Periyasamy, M., Kamaludeen, N. A., Towfek, S., Marappan, R., Kidambi Raju, S., et al. (2023). Optimizing traffic flow in smart cities: soft gru-based recurrent neural networks for enhanced congestion prediction using deep learning. *Sustainability* 15, 5949. doi:10.3390/su15075949

Almerekhi, H., Kwak, H., Salminen, J., and Jansen, B. J. (2022). Provoke: toxicity trigger detection in conversations from the top 100 subreddits. *Data Inf. Manag.* 6, 100019. doi:10.1016/j.dim.2022.100019

Baker, S., and Xiang, W. (2023). Artificial intelligence of things for smarter healthcare: A survey of advancements, challenges, and opportunities. *IEEE Commun. Surv. Tutorials* 25, 1261–1293. doi:10.1109/comst.2023.3256323

Chen, Z., Yu, H., Luo, L., Wu, L., Zheng, Q., Wu, Z., et al. (2021). Rapid and accurate modeling of pv modules based on extreme learning machine and large datasets of iv curves. *Appl. Energy* 292, 116929. doi:10.1016/j.apenergy.2021.116929

Cui, L., Qu, Y., Gao, L., Xie, G., and Yu, S. (2020). Detecting false data attacks using machine learning techniques in smart grid: A survey. *J. Netw. Comput. Appl.* 170, 102808. doi:10.1016/j.jnca.2020.102808

Cvišić, I., Marković, I., and Petrović, I. (2021). "Recalibrating the kitti dataset camera setup for improved odometry accuracy," in 2021 European Conference on Mobile Robots (ECMR), Bonn, Germany, 31 August 2021 - 03 September 2021 (IEEE), 1–16.

Derhab, A., Aldweesh, A., Emam, A. Z., and Khan, F. A. (2020). Intrusion detection system for internet of things based on temporal convolution neural network and efficient feature engineering. *Wirel. Commun. Mob. Comput.* 2020, 1–16. doi:10.1155/2020/6689134

Dhend, M. H., and Chile, R. H. (2017). Fault diagnosis of smart grid distribution system by using smart sensors and symlet wavelet function. *J. Electron. Test.* 33, 329–338. doi:10.1007/s10836-017-5658-9

Geetha, R., and Thilagam, T. (2021). A review on the effectiveness of machine learning and deep learning algorithms for cyber security. *Archives Comput. Methods Eng.* 28, 2861–2879. doi:10.1007/s11831-020-09478-2

Hang, R., Liu, Q., Hong, D., and Ghamisi, P. (2019). Cascaded recurrent neural networks for hyperspectral image classification. *IEEE Trans. Geoscience Remote Sens.* 57, 5384–5394. doi:10.1109/TGRS.2019.2899129

Hong, D., Gao, L., Yokoya, N., Yao, J., Chanussot, J., Du, Q., et al. (2021). More diverse means better: multimodal deep learning meets remote-sensing imagery classification. *IEEE Trans. Geoscience Remote Sens.* 59, 4340–4354. doi:10.1109/TGRS.2020.3016820

Jalata, I. K., Truong, T. D., Allen, J. L., Seo, H. S., and Luu, K. (2021). Movement analysis for neurological and musculoskeletal disorders using graph convolutional neural network. *Future Internet* 13, 194. doi:10.3390/fi13080194

Kaygusuz, C., Babun, L., Aksu, H., and Uluagac, A. S. (2018). "Detection of compromised smart grid devices with machine learning and convolution techniques," in 2018 IEEE International Conference on Communications (ICC), Kansas City, MO, USA, 20-24 May 2018 (IEEE), 1–16. doi:10.1109/ICC.2018.8423022

Lakshmanna, K., Kaluri, R., Gundluru, N., Alzamil, Z. S., Rajput, D. S., Khan, A. A., et al. (2022). A review on deep learning techniques for iot data. *Electronics* 11, 1604. doi:10.3390/electronics11101604

Liao, W., Bak-Jensen, B., Pillai, J. R., Wang, Y., and Wang, Y. (2021). A review of graph neural networks and their applications in power systems. *J. Mod. Power Syst. Clean Energy* 10, 345–360. doi:10.35833/mpce.2021.000058

Morid, M. A., Borjali, A., and Del Fiol, G. (2021). A scoping review of transfer learning research on medical image analysis using imagenet. *Comput. Biol. Med.* 128, 104115. doi:10.1016/j.compbiomed.2020.104115

Parmar, N., Vaswani, A., Uszkoreit, J., Kaiser, L., Shazeer, N., Ku, A., et al. (2018). "Image transformer," in *International conference on machine learning* (PMLR), 4055–4064.

Qi, W., and Wang, Z. (2022). "Intelligent system construction of gymnasium based on internet of things resource sharing technology," in 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), Dalian, China, 14-16 April 2022 (IEEE), 1328–1332.

Sagu, A., Gill, N. S., Gulia, P., Singh, P. K., and Hong, W. C. (2023). Design of metaheuristic optimization algorithms for deep learning model for secure iot environment. *Sustainability* 15, 2204. doi:10.3390/su15032204

Shanthamallu, U. S., Thiagarajan, J. J., and Spanias, A. (2020). "A regularized attention mechanism for graph attention networks," in ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 04-08 May 2020 (IEEE), 3372–3376.

Shin, C., Lee, E., Han, J., Yim, J., Rhee, W., and Lee, H. (2019). The enertalk dataset, 15 hz electricity consumption data from 22 houses in korea. *Sci. data* 6, 193. doi:10.1038/s41597-019-0212-5

Siniosoglou, I., Radoglou-Grammatikis, P., Efstathopoulos, G., Fouliras, P., and Sarigiannidis, P. (2021). A unified deep learning anomaly detection and classification approach for smart grid environments. *IEEE Trans. Netw. Serv. Manag.* 18, 1137–1151. doi:10.1109/tnsm.2021.3078381

Sridharan, N. V., and Sugumaran, V. (2021). Convolutional neural network based automatic detection of visible faults in a photovoltaic module. *Energy Sources, Part A Recovery, Util. Environ. Eff.* 2021, 1–16. doi:10.1080/15567036.2021.1905753

Srivastava, S., Paul, B., and Gupta, D. (2023). Study of word embeddings for enhanced cyber security named entity recognition. *Procedia Comput. Sci.* 218, 449–460. doi:10.1016/j.procs.2023.01.027

Tavana, M., Hajipour, V., and Oveisi, S. (2020). Iot-based enterprise resource planning: challenges, open issues, applications, architecture, and future research directions. *Internet Things* 11, 100262. doi:10.1016/j.iot.2020.100262

Wan, B., Xu, C., Mahapatra, R. P., and Selvaraj, P. (2022). Understanding the cyber-physical system in international stadiums for security in the network from cyber-attacks and adversaries using ai. *Wirel. Personal. Commun.* 127, 1207–1224. doi:10.1007/s11277-021-08573-2

Wang, D., He, Y., Ma, Y., Wu, H., and Ni, G. (2023). The era of artificial intelligence: talking about the potential application value of chatgpt/gpt-4 in foot and ankle surgery. *J. Foot Ankle Surg.* doi:10.1053/j.jfas.2023.07.002

Wang, T., Liu, W., Cabrera, L. V., Wang, P., Wei, X., and Zang, T. (2022). A novel fault diagnosis method of smart grids based on memory spiking neural p systems considering measurement tampering attacks. *Inf. Sci.* 596, 520–536. doi:10.1016/j.ins.2022.03.013

Wang, W., Wei, F., Dong, L., Bao, H., Yang, N., and Zhou, M. (2020). Minilm:deep self-attention distillation for task-agnostic compression of pre-trained transformers. *Adv. Neural Inf. Process. Syst.* 33, 5776–5788. doi:10.48550/arXiv.2002.10957

Watson, E., Viana, T., and Zhang, S. (2023). Augmented behavioral annotation tools, with application to multimodal datasets and models: A systematic review. *AI* 4, 128–171. doi:10.3390/ai4010007

Wu, X., Hong, D., and Chanussot, J. (2022). Convolutional neural networks for multimodal remote sensing data classification. *IEEE Trans. Geoscience Remote Sens.* 60, 1–10. doi:10.1109/TGRS.2021.3124913

Zhang, R., Yao, W., Shi, Z., Ai, X., Tang, Y., and Wen, J. (2023). Identification and screening of key traffic violations: based on the perspective of expressing driver's accident risk. *IEEE Trans. Power Syst.* 2023, 1–18. doi:10.1080/17457300.2023.2245804

Zuo, S., Jiang, H., Li, Z., Zhao, T., and Zha, H. (2020). "Transformer hawkes process," in *International conference on machine learning (PMLR)* 119, 11692–11702. doi:10.48550/arXiv.2002.09291