Anomaly Detection in Building Equipment Energy Consumption Using Bi-LSTM and Convolutional Block Attention Mechanism

Xin Wan[1*], Xiaoling Cai[1], Lele Dai[2]
[1]School of Architecture and Environmental Engineering, Wuxi City College of Vocational Technology
Wuxi 214000, China
[2]School of Energy and Power Engineering, Nanjing University of Science and Technology
Nanjing 210000, China
E-mail: wanxin840207@163.com, caixl123456@126.com, dyldagao@njust.edu.cn
*Corresponding author

Keyword: Bi-LSTM, CBAM, construction equipment, abnormal energy consumption

Received: August 12, 2025

With the growing complexity of building energy systems, accurate anomaly detection in equipment energy consumption has become crucial for improving operational efficiency. This study proposes an energy anomaly detection model that integrates Bidirectional Long Short-Term Memory (Bi-LSTM) networks and a Convolutional Block Attention Module (CBAM). The input features include low and high energy consumption time ratios and dynamic time warping distance, constructed into a 32-dimensional feature vector using a sliding window of 24 hours. The Bi-LSTM layer with 128 forward and 128 backward hidden units captures bidirectional temporal dependencies. CBAM refines critical feature dimensions and time steps through channel and spatial attention mechanisms. The model was trained using 70% of labeled data from the LEAD1.0 public smart meter dataset and tested on the remaining 30%. Experimental results showed that on Dataset A, the model achieved an accuracy of 0.98, a recall of 0.97, and an RMSE of 0.02. On Dataset B, it achieved 0.97, 0.98, and 0.05, respectively. Comparative analysis against baseline models, including Bi-LSTM, LSTM, and GRU, demonstrated significant improvements in both accuracy and error metrics. An ablation study confirmed the contribution of each module to model performance. Statistical validation across multiple runs showed that the improvements were consistent and robust. These findings suggest that combining Bi-LSTM with dual-attention mechanisms provides an effective solution for detecting both transient and persistent energy anomalies in dynamic building environments.

Povzetek: Študija predstavlja model Bi-LSTM s CBAM mehanizmom pozornosti za zaznavanje energetskih anomalij v stavbah, ki dosega zelo visoko natančnost in zanesljivo presega primerjalne modele.

1 Introduction

With the continuous growth of global energy requirement, energy management issues in the construction industry are receiving increasing attention. According to statistics, building Energy Consumption (EC) accounts for over 40% of the world's total EC, with Abnormal Energy Consumption (AEC) of building equipment being one of the main causes of energy waste. Therefore, how to efficiently and accurately detect the AEC of building equipment to enhance energy efficiency and lower CO2 emissions has become an important research direction for intelligent building management and energy-saving optimization. Traditional energy management mainly relies on fixed threshold monitoring and statistical analysis methods. However, due to the complex operating environment and large load fluctuations of building equipment, fixed rules are difficult to adapt to dynamically changing EC patterns, which can easily lead to misjudgments and omissions [1]. In addition, EC data of building equipment usually exhibits nonlinear and non-stationary characteristics in time series. Traditional methods have limited capabilities in feature extraction and pattern recognition, and cannot effectively address complex anomaly detection requirements [2]. Traditional threshold detection methods rely on expert experience to set fixed boundaries, which are difficult to adapt to dynamically changing operating status and are prone to false alarms due to seasonal fluctuations or equipment aging. Although statistical learning models can identify some abnormal patterns, they lack sufficient modeling of temporal dependencies and cannot capture the continuous evolution of device operating states. Currently, Long Short-Term Memory (LSTM) networks have been extensively utilized in EC prediction due to their gating mechanism's advantage in processing time-series data. However, the unidirectional network structure can only capture historical information and is difficult to synchronously analyze future trend features. Therefore, this study constructs an AEC detection model that combines Bidirectional LSTM (Bi-LSTM) and Convolutional Block Attention Mechanisms (CBAM). This method utilizes a bidirectional structure to obtain information about the time steps before and after,

enhancing the ability to model time series. Meanwhile, CBAM is introduced to optimize the feature extraction schedule and improve the model's attention to key EC patterns. The primary objective of this research is to design and validate a building equipment AEC detection model that can reliably capture both low and high EC anomalies across different operational scenarios, improving detection accuracy and reducing false alarms. The study focuses on integrating bidirectional temporal modeling with feature attention to enhance anomaly recognition performance in practical smart building environments. The innovation lies in the use of Bi-LSTM to improve time series modeling capabilities, while introducing CBAM to enhance feature selection capabilities. Different from existing anomaly detection models that predominantly use unidirectional LSTM networks or basic attention mechanisms, this study uniquely integrates the bidirectional temporal feature extraction advantage of Bi-LSTM with the channel-spatial dual attention mechanism of CBAM. This integration effectively resolves the limitations of traditional methods in capturing complex temporal dependencies and accurately extracting critical EC features simultaneously. Moreover, the method proposed herein significantly enhances generalization capability and reduces false alarms, addressing practical issues commonly encountered in diverse and dynamic building operational environments.

2 Related works

Building energy management is crucial for energy conservation, consumption reduction, and sustainable development. Accurate detection of AEC can effectively avoid energy waste and optimize equipment operation. However, EC data of building equipment have temporal and fluctuating characteristics. Traditional methods, such as fixed-threshold monitoring, statistical modeling approaches (e.g., moving average and ARIMA), and optimization-based techniques like Particle Swarm Optimization (PSO) or Genetic Algorithms (GA), are hard to precisely identify abnormal patterns in complex and dynamic building environments. In phase change heat storage system optimization, Xiao et al. designed a multi-channel phase change heat storage device with porous metal foam structure, and verified its heat transfer performance through 3D numerical simulation. By further combining the LSTM and backpropagation neural network hybrid prediction model, experiments have shown that increasing the flow rate or inlet temperature of the heat transfer medium could accelerate the melting of phase change materials. This intelligent algorithm provided a new approach for dynamic control of thermal storage systems [3]. Zheng et al. developed a dual-channel image analysis model that combines convolutional neural networks and LSTM to address the problem of partial discharge pattern recognition in power equipment. Compared with traditional machine learning and single deep learning models, its accuracy in classifying 4 kinds of discharge patterns has significantly improved. The effectiveness of the model architecture and hyperparameter design has been verified through ablation experiments [4]. In terms of predicting the lifespan of energy devices, Zhang et al. constructed a residual lifespan assessment model for supercapacitors based on a stacked Bi-LSTM recursive network. When the hidden layer was set to 2, the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) decreased to 0.0275 and 0.0241, respectively, demonstrating excellent prediction accuracy [5]. In addition, Abbasi FB et al. proposed an optimization architecture that integrates multi-condition scheduling strategies and multi-level queues for the Internet of Things workflow scheduling problem in fog computing environments. By calculating priority and LSTM load prediction to achieve dynamic task allocation, experiments have shown that the system throughput has increased by 48.6%, communication costs have been reduced by 56%, and task latency and parallel efficiency have also been significantly improved [6].

Vojdani et al. proposed a method using a meta-heuristic optimization algorithm to integrate the advantages of natural lighting into energy-saving design. This indicated that PSO and MVO algorithms exhibited better performance compared to GA and WOA in solving lighting and energy optimization issues related to improving building performance. The selection algorithm needed to consider the unique properties of the problem and the complexity of optimizing the landscape [7]. Nasouri and Delgarm proposed a simulation-oriented Pareto optimization method for building energy efficiency to reduce high EC and toxic gas emissions in the Iranian construction industry. This method combined the EnergyPlus simulation program with a multi-objective ACO algorithm. In a test on a residential building in a hot and semi-arid region of Iran, this method effectively obtained the optimal configuration of the building [8]. Jun and Fei put forth a multi-objective optimization framework based on surrogate models to obtain the optimal building design scheme to balance the conflicting goals of energy efficiency and comfort performance. The use of the NSGA-II algorithm and random forest replacement model could accurately predict EC and significantly improve building energy efficiency and comfort, providing effective guidance for building design optimization and reference for similar projects [9].

In summary, many scholars have researched AEC testing of building equipment and achieved certain results. Although these methods have improved the accuracy of AEC detection to some extent, there are still issues such as high misjudgment rates and insufficient extraction of time-dependent features. Therefore, the paper constructs a building equipment AEC detection model grounded on Bi-LSTM and CBAM, aiming to increase the accuracy and robustness of AEC recognition, and provide an efficient and reliable solution for intelligent building energy management. The summary of related works is in Table 1.

Table 1: Related works.

Research	Method	Research Content	Key Performance Metrics	Reference
Xiao et al.	LSTM-BP Neural Network	Phase change heat storage system performance prediction	Faster melting with higher flow/inlet temp; validated via 3D simulation	[3]
Zheng et al.	CNN + LSTM	Partial discharge pattern recognition in power equipment	Improved classification accuracy on 4 discharge types	[4]
Zhang et al.	Stacked Bi-LSTM RNN	Remaining useful life prediction for supercapacitors	RMSE = 0.0275; MAE = 0.0241	[5]
Abbasi et al.	LSTM + Priority Queue Scheduling	Fault-tolerant task scheduling in IoT fog/cloud environments	Throughput ↑48.6%, Communication cost ↓56%, latency ↓	[6]
Vojdani et al.	PSO, MVO, GA, WOA	Optimization of daylight and EC in building design	PSO and MVO achieved better convergence and performance	[7]
Nasouri & Delgarm	Multi-objective ACO + Energy Plus	Energy efficiency and emission reduction optimization in residential buildings	Optimal building configuration found; case study in hot/semi-arid climate	[8]
Jun & Fei	NSGA-II + Random Forest Surrogate Model	Trade-off optimization between building energy efficiency and comfort performance	Improved prediction accuracy; better comfort-energy balance	[9]

3 Methods

3.1 AEC testing model for construction equipment

The proposed approach distinctively combines Bi-LSTM with a CBAM, forming a novel hybrid deep learning model. Although Bi-LSTM is a widely adopted choice for sequence data tasks, recent advances in time series anomaly detection have shown that transformer-based models offer significant advantages, especially in modeling long-range dependencies. Transformer uses a self-attention mechanism, allowing the model to focus on important time steps in long sequences, solving the vanishing gradient problem common in RNN-based architectures. Temporal Convolutional Networks (TCNN) are also effective in capturing long-range dependencies using causal convolutions. However, unlike Bi-LSTM, they do not explicitly model bidirectional dependencies, which may limit their ability to capture complex temporal dynamics in anomaly detection tasks. Despite these alternatives, Bi-LSTM remains a strong choice due to its simplicity, interpretability, and efficiency in training large datasets. While Transformer models provide cutting-edge performance, they also come with higher computational costs and longer training times, especially when processing large-scale data. Given the trade-offs between model complexity and computational feasibility, Bi-LSTM is selected for this study for its balanced performance, especially in scenarios where computational resources are constrained and real-time deployment is essential. Future work may explore hybrid models that combine Bi-LSTM and Transformer-based architectures to enhance anomaly detection performance while maintaining scalability. This integration not only enhances the comprehensive understanding of temporal context through bidirectional analysis but also employs channel and spatial attention simultaneously, improving feature extraction precision. The development of building intelligence has made EC monitoring and management of building equipment an important research direction for improving energy utilization efficiency. However, due to the complex operating environment and significant load fluctuations, AEC detection still faces many challenges. Therefore, this study proposes an AEC detection model for building equipment and first establishes an AEC feature mathematical model, whose model structure is shown in Figure 1.

In Figure 1, the model mainly contains three parts: Low EC Time Ratio (LECTR), High EC Time Quantity (HECTQ), and DTW Distance Characteristics (DTWDC). LECTR refers to the proportion of time during a certain time window when the operating EC of building equipment is below a set threshold. This feature is used to measure whether the device has been in a low EC state for a long time, to determine whether the device has abnormal standby or shutdown phenomena. Its expression is given by equation (1).

$$n(i)=\begin{cases}1, E_{\min}<E(i)<E_{\max}\\ 0, Other\end{cases} \tag{1}$$

In equation (1), $E_{\min}$ and $E_{\max}$ are the lower and upper limits of EC data, and EC data within this range are considered low EC data. $E(i)$ is the EC time series. The definition of ELECR is given by equation (2).

$$X=\frac{1}{N}\sum_{i=0}^{N-1}\mathrm{sgn}(n(i)) \tag{2}$$

In equation (2), X represents ELECTR, which is the proportion of time during the entire time period that is in a low EC state. N is the total duration of the time series. HECTQ refers to the total amount of time during a certain period when the EC of building equipment is higher than the normal operating range. This indicator is used to determine whether the device has been in an abnormally high EC state for a long time, as shown in equation (3).

$$T=\sum_{i=0}^{N-1}\mathrm{sgn}(E(i)-E_{\max}) \tag{3}$$

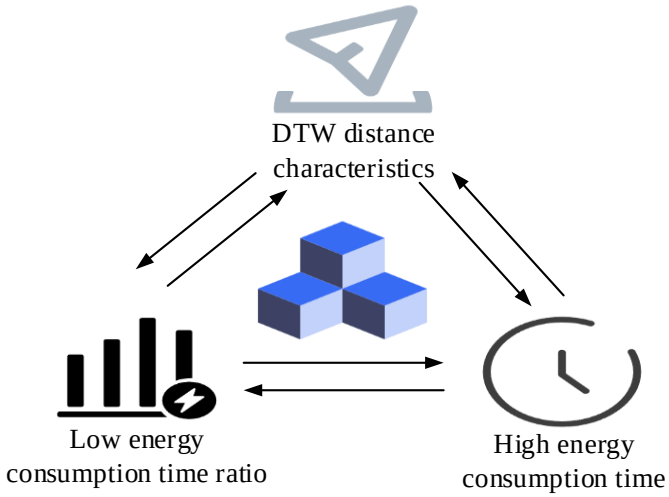


Figure 1: Structure of AEC feature mathematical model.

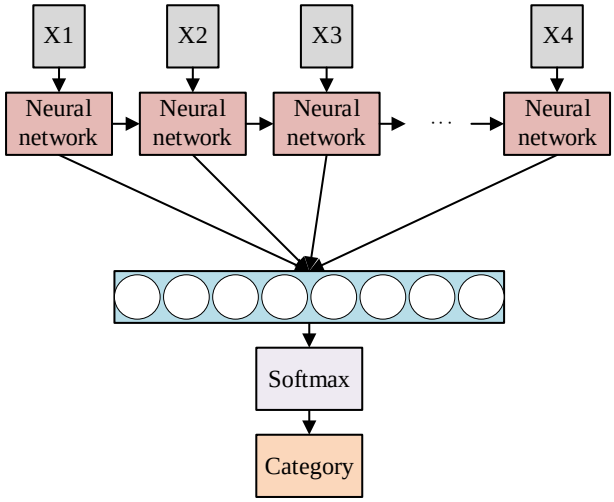


Figure 2: AEC classification model for building equipment.

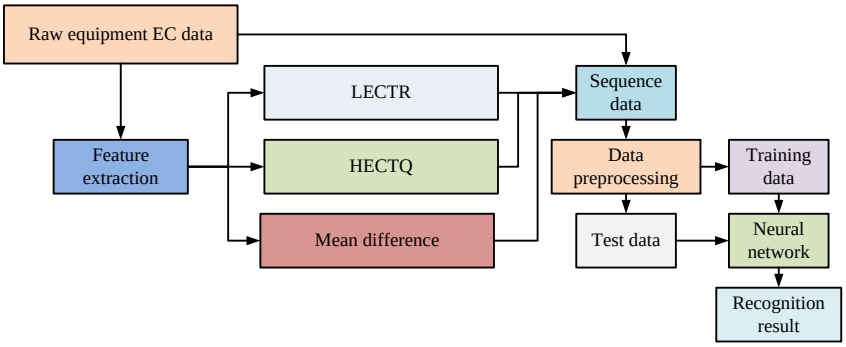


Figure 3: Overall structure of AEC detection model for building equipment.

In equation (3), T is HECTQ, which represents the total number of hours the equipment is in a high-EC state throughout the entire time series. The dynamic time bending distance feature is utilized to measure the discrepancies between the current EC mode and the historical normal EC mode. A historical EC data sequence during normal operation is selected as the reference sequence, and then the EC data of the current time window is selected as the test sequence, and its distance is calculated using the formula shown in equation (4).

$$DTW(X,Y) = \min \sum_{(i,j)\in W} d(x_i, y_j) \quad (4)$$

In equation (4), $d(x_i, y_j)$ is the Euclidean distance between sequences X and Y, and W is the alignment path. The small distance indicates that the current EC mode is similar to the normal mode and the device is

operating normally. A large distance indicates a significant change in the current EC mode, possibly due to AEC [10-12]. Then, a neural network is employed to solve the mathematical model. The structure of the AEC classification model for building equipment based on neural networks is shown in Figure 2.

In Figure 2, the entire model includes an Input Layer (IL), a Neural Network Layer (NNL), a Fully Connected Layer (FCL), and an Output Layer (OL). The IL receives time-series data, where each time step represents the EC information of the device at different time points. The NNL is used to extract long-term dependency features from time series. Each neural network unit receives the hidden state from the previous time step and fuses it with the present input to generate a novel hidden state, which is gradually passed on to the next time step [13-14]. Finally, the output of the NNL is passed to the FCL for feature transformation and mapping to a fixed length vector. Then, the Softmax activation function is utilized to calculate the class probability, and the classification result is generated in the OL. The general framework of the AEC testing model for building equipment is displayed in Figure 3.

In Figure 3, firstly, the system obtains raw equipment EC data and performs feature extraction, including key features such as ELECTR and HECTQ. These features are used to characterize the EC status of the device. The extracted features are constructed into 32-dimensional Time Series Data (TSD) and enter the data preprocessing stage. At this stage, the data are segmented into training and testing data for subsequent model training and validation. The preprocessed training data are used to train a neural network classifier. This classifier is capable of learning patterns and trends in TSD and capturing temporal dependencies through NNLs to improve the accuracy of anomaly classification. The test data are used to evaluate the generalization ability of the model, and the final output of the model is the recognition result, that is, whether there are abnormalities in the EC status of the equipment.

3.2 AEC detection model for building equipment based on Bi-LSTM

For the established AEC detection model of building equipment, Bi-LSTM is employed as the core neural network due to its superior ability in sequential feature extraction. The proposed architecture is designed in four key stages: input preprocessing, Bi-LSTM temporal encoding, CBAM-based feature refinement, and final classification.

IL receives a 32-dimensional time-series feature vector at each time step, including LECTR, HECTQ, DTW distance, and derived statistical indicators. A sliding time window of length T (e.g., 24 hours) forms the model input with shape (T×32). The Bi-LSTM Layer consists of a forward and a backward LSTM, each with 128 hidden units. The forward LSTM captures past-to-present dependencies, while the backward LSTM captures future-to-present dependencies. Their outputs are concatenated, producing a feature map of size (T×256), which encodes complete bidirectional temporal patterns. CBAM Attention Module is applied to the Bi-LSTM output. The CBAM refines features through sequential channel attention and spatial (temporal-step) attention. Channel attention emphasizes the most informative feature dimensions (e.g., key energy indicators), while spatial attention highlights critical time steps where AEC occurs. This results in an attention-weighted feature representation of size (T×256). FCL and Softmax Output: The attention-enhanced feature map is globally pooled and passed through an FCL to produce a fixed-length 128-dimensional vector, which is finally classified via Softmax into normal or AEC categories.

In the AEC testing task of building equipment, the EC patterns of the equipment often have complex temporal dependencies. Traditional LSTM can only process data unidirectionally from the past to the future, and may overlook certain important temporal features. Bi-LSTM can simultaneously utilize historical and future information to improve the recognition ability of AEC and perfect the classification performance. The structure of the AEC detection model for building equipment based on Bi-LSTM is shown in Figure 4.

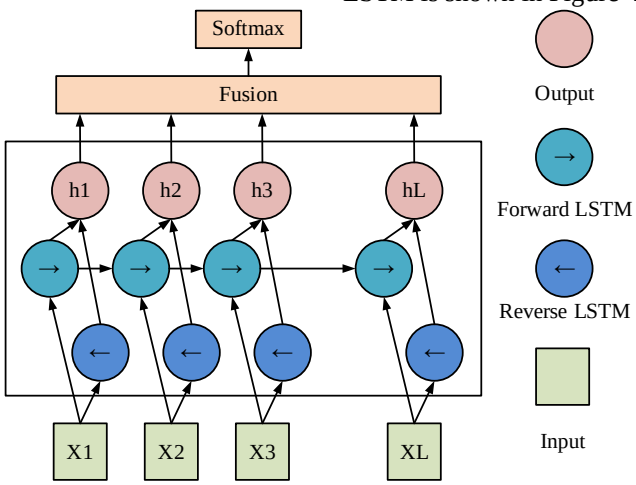


Figure 4: AEC detection model for building equipment based on Bi-LSTM.

In Figure 4, the model includes the IL, Bi-LSTM layer, fusion layer, FCL, and OL. The IL receives time-series data, with each time step representing the EC information of the device at different time points. The Bi-LSTM layer has two LSTMs, namely the forward LSTM and backward LSTM. The forward LSTM processes sequence data from left to right, while the backward LSTM processes inversely. This bidirectional structure can simultaneously utilize past and future information, improving the effectiveness of time series modeling [15]. The outputs of both LSTMs are integrated in the fusion layer to enable the model to have a more comprehensive understanding of the features in the time series. The fused result is passed to the FCL for feature mapping and converted into a fixed length vector. Then, the Softmax is taken to calculate the class probability, and finally, the classification result is generated in the OL. Since not all features or time steps contribute equally to anomaly detection, an attention mechanism is required to guide the model toward the most informative patterns. In EC time series, different features have varying sensitivity to anomalies, and abnormal patterns often appear only in specific time intervals. CBAM is particularly suitable because it provides a dual-attention mechanism: channel attention emphasizes the most discriminative energy features, while spatial attention focuses on critical time steps where anomalies occur. Compared to simpler attention mechanisms such as Squeeze-and-Excitation (SE) or single-head temporal attention, CBAM can simultaneously refine feature selection across channels and temporal dimensions, which is essential for capturing subtle and sporadic AEC patterns in complex building environments. Its structure is shown in Figure 5.

In Figure 5, CBAM includes Channel and Spatial attention mechanisms (i.e., CAM and SAM). In the CAM section, CBAM learns the features of various channels and weights them, allowing the model to concentrate more on information rich channels. In the SAM section, CBAM focuses on selecting key features from feature images, guiding the network to focus on these key regions and extract relevant feature information. This dual attention mechanism that combines channels and space enables CBAM to refine feature extraction while further improving the efficiency and accuracy of network processing of image data [16]. The formula of CBAM is equation (5).

$$F' = M_c(F) \otimes F \tag{5}$$

In equation (5), F means the Feature Map (Fmap) obtained through the convolution operation of the last layer. $\otimes$ is the multiplication symbol among features. M_c represents CAM. F' is a novel Fmap processed through CAM. The expression of the Fmap after SAM processing is shown in equation (6).

$$F'' = M_s(F') \otimes F' \tag{6}$$

In equation (6), M_s denotes SAM. F'' is the attention weight value gained by processing F' through SAM. The final F'' is weighted and summed with the original Fmap, and then output as a new Fmap. The final model structure is shown in Figure 6.

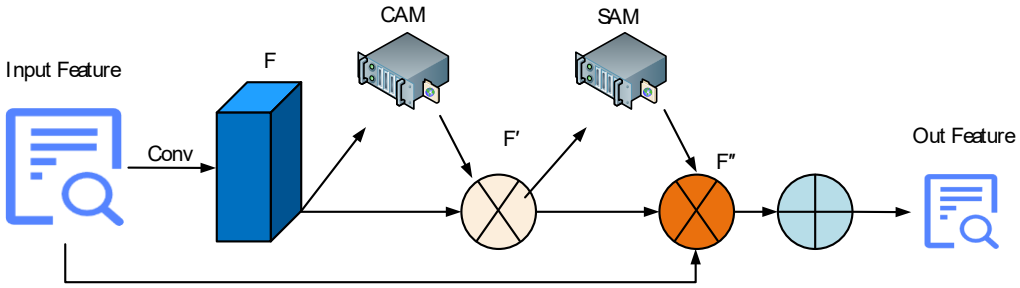


Figure 5:CBAM structure.

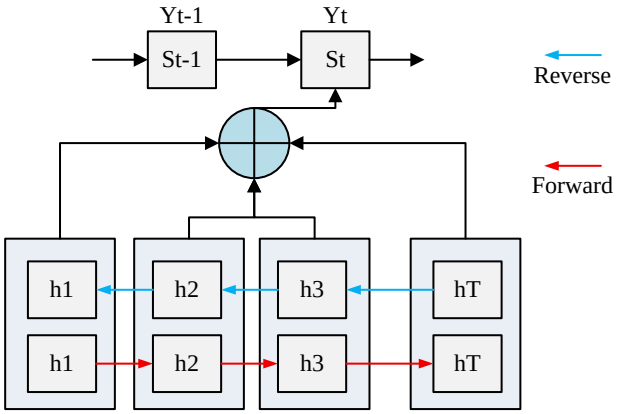


Figure 6:AEC detection model for building equipment based on Bi-LSTM-CBM.

In Figure 6, the model operates in four main stages. The IL receives a time-series feature matrix of size (T×32), where T is the time window length. The Bi-LSTM layer encodes bidirectional temporal dependencies, producing a (T×256) feature representation. The CBAM module applies channel and spatial attention to focus on critical energy features and abnormal time steps. The output is aggregated through an FCL and Softmax to generate the anomaly detection result. This detailed architecture ensures the model can accurately identify AEC patterns in complex building environments. The Bi-LSTM layer is used to process input TSD. The hidden states of the Bi-LSTM at each moment are output by the forward and backward LSTM, respectively. These hidden states contain characteristic information of the time series. To perfect the model's attention to key time points, an attention mechanism is selected to calculate the importance weights of different time steps, and then obtain a global representation through weighted summation. This process can highlight the most important time step features for classification tasks. Next, the fused representation is passed to the classification layer, where the Softmax function calculates the class probability and ultimately generates the classification result for AEC detection.

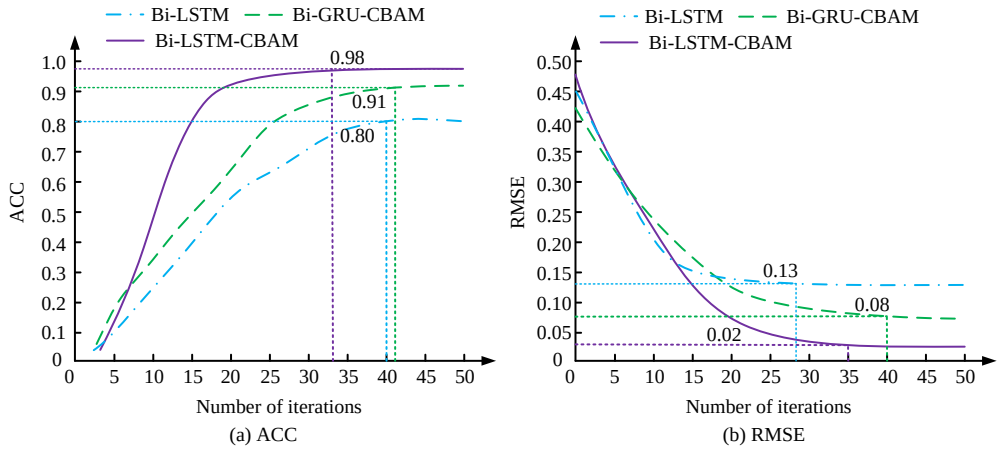
To ensure generalization and prevent overfitting, several regularization strategies are employed during model training. A dropout layer is inserted after the Bi-LSTM and before the FCL, with a dropout rate set to 0.3, which is empirically determined to balance regularization strength and convergence stability. Additionally, L2 weight regularization is applied to all trainable parameters, with a decay coefficient of 1e-5. Early stopping is also used during training, with a patience setting of 10 epochs, monitoring validation loss to halt training once the model’s performance plateaus. This strategy ensures that the model does not continue fitting on noise or overemphasize infrequent patterns in the training set. Furthermore, the training set is shuffled at each epoch to minimize sequential bias, and batch normalization is incorporated after attention modules to stabilize feature distributions across mini-batches. These techniques jointly reduce the risk of overfitting despite the large training data volume. Experimental results confirm that the model maintains robust generalization across both test datasets and unseen time windows.

4 Results

4.1 Performance analysis of AEC detection model for building equipment

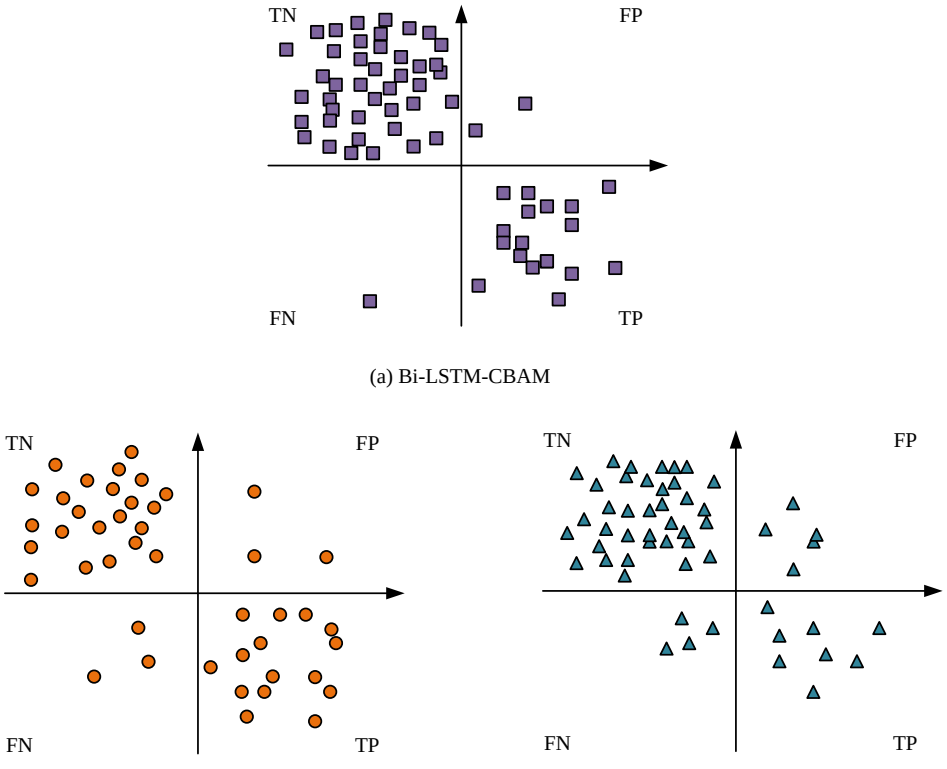
The hardware configuration used in this experiment is Intel Core i7-13490KF CPU, NVIDIA GeForce GTX4070Ti GPU, 8GB VRAM, and 16GB RAM. The dataset adopts the LEAD1.0 public dataset, which contains annual TSD of 1413 smart meters. The data of each smart meter records the electricity consumption for a period of one year. The data are collected on an hourly basis, covering EC patterns of various building types. The dataset provides detailed anomaly labeling for EC data, including point anomalies and sequence anomalies. Point anomaly refers to the abnormal EC value at a single time point. Sequence anomaly refers to abnormal EC patterns within a continuous period of time. This study utilizes two datasets, Dataset A and Dataset B, both derived from the LEAD1.0 public smart meter database. LEAD1.0 contains annual EC records from 1,413 smart meters, covering various building types such as office buildings, residential apartments, and commercial facilities. The original sampling frequency is hourly, resulting in 8,760 time points per meter per year. Prior to model training, a structured data preprocessing pipeline is applied to the raw time-series EC data. First, missing values within hourly readings are imputed using linear interpolation to preserve temporal continuity. Outliers that exceeded three standard deviations from the mean within a local window of seven days are flagged and removed, based on robust statistical profiling. All time-series features, including EC values and derived metrics such as ELECTR and HECTQ, are min-max normalized to the range of zero to one. This ensures consistent feature scaling across different building types and meter devices. For segmentation, a sliding time window approach is used, with a window size of 24 hours and a stride of 1 hour, generating overlapping time slices to preserve local temporal dynamics. Additionally, data from each smart meter are individually normalized to retain device-specific consumption patterns, avoiding distributional shifts caused by large-scale aggregation. No synthetic data or resampling techniques are applied to preserve the integrity of real-world usage trends. This preprocessing strategy enables stable convergence of the Bi-LSTM-CBAM model and ensures that temporal patterns are not distorted by data quality issues.

The models are implemented using Python 3.10 and PyTorch 1.13. Training is conducted on an NVIDIA RTX 4070Ti GPU. The optimization algorithm used is Adam with a learning rate of 0.001, and the batch size is set to 64. A dropout rate of 0.3 is applied to the Bi-LSTM and FCL to prevent overfitting. The training process utilizes the categorical cross-entropy loss function, and each model is trained for up to 100 epochs with an early stopping mechanism (patience = 10) based on validation loss. No data augmentation is applied due to the temporal integrity requirement of EC sequences. All time series are normalized using min-max scaling before training. In defining abnormal consumption, the LECTR threshold is set at 20% of the device's average operating range, while the HECTQ threshold is fixed at +30% above the upper bound of normal load behavior. These thresholds are empirically determined based on the statistical distribution of historical data and validated by domain experts. This study selects Bidirectional Gated Recurrent Unit (Bi-GRU) based on attention mechanism and Bi-LSTM as comparative models, as shown in Figure 7.



(a) ACC (b) RMSE

Figure 7:ACC and RMSE for various models.



(a) Bi-LSTM-CBAM

(b) Bi-GRU-CBAM (c) Bi-LSTM

Figure 8:Analysis of classification performance of various models.

Table 2: Comprehensive performance analysis of the model.

Dataset	Model	Accuracy	Recall	F1 score	RMSE	Convergence rate
Dataset A	Bi-LSTM-CBAM	0.98	0.97	0.97	0.02	31
	Bi-GRU-CBAM	0.91	0.89	0.90	0.08	48
	Bi-LSTM	0.80	0.76	0.78	0.13	50
Dataset B	Bi-LSTM-CBAM	0.97	0.98	0.99	0.05	32
	Bi-GRU-CBAM	0.86	0.91	0.95	0.12	46
	Bi-LSTM	0.82	0.74	0.88	0.17	57

Figures7 (a) and (b) show the variation of accuracy and RMSE of different models when training with the number of iterations. In Figure 7 (a), the convergence speed of Bi-LSTM-CBM is significantly faster than that

of Bi-LSTM and Bi-GRU-CBM, and the final accuracy is the highest, reaching 0.98. In contrast, Bi-GRU-CBM reaches 0.91 after 40 iterations, while the final accuracy of Bi-LSTM is only 0.80. In Figure 7 (b), the RMSE of Bi-LSTM-CBM decreases the fastest and eventually converges to 0.02, much lower than the 0.13 of Bi-LSTM and 0.08 of Bi-GRU-CBM. This is because CBAM enhances the ability of Bi-LSTM to focus on effective features, reduces the interference of irrelevant information, and makes the model more efficient in error optimization, resulting in a much lower RMSE compared to other models. This indicates that Bi-LSTM-CBM performs the best in AEC detection tasks, not only with higher accuracy but also significantly reducing prediction errors. The results of analyzing the classification performance of these models are shown in Figure 8.

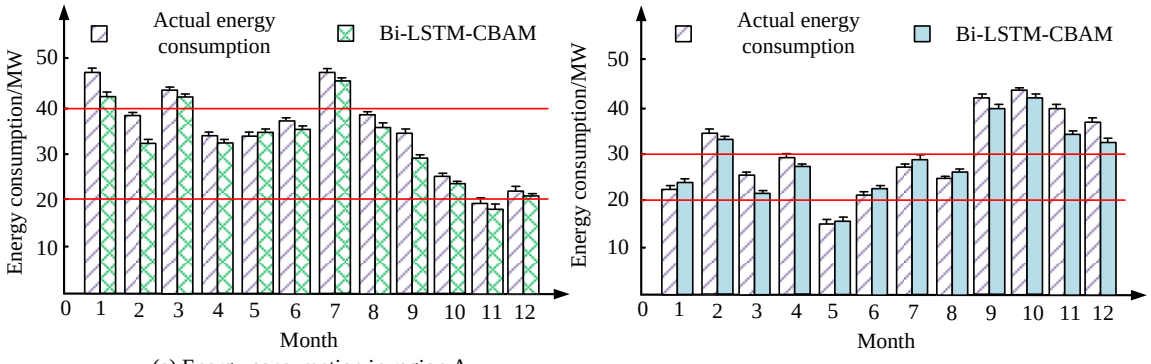
In Figure 8, TN represents the model's determination of normal EC, but in reality, it is normal EC. FP indicates that the model determines normal EC, but in reality, it is AEC. FN indicates that the model has determined AEC, but in reality, the EC is normal. TP indicates that the model is judged as AEC in reality. Figures 8 (a) to (c) show the classification performance of Bi-LSTM-CBM, Bi-GRU-CBM, and Bi-LSTM. The classification performance of Bi-LSTM-CBM is the best, with highly concentrated TN and TP samples and fewer FP and FN samples, indicating that the model can accurately identify normal and abnormal EC with a low misjudgment rate. The introduction of CBAM increases the model's focus on key features, rises classification accuracy, and generalization ability. Bi-GRU-CBM also introduces CBAM, but compared to Bi-LSTM-CBM, the number of FN and FP samples increases significantly, especially the FN samples are more scattered, indicating that the model is more likely to ignore abnormal samples, leading to a decrease in the recall rate. The ability of GRU to capture long-term dependency information is slightly inferior to LSTM, causing a slight decrease in the stability of abnormal EC detection. Due to the lack of CBAM, Bi-LSTM has the worst classification performance, with the highest number of FP and FN samples and the least concentrated distribution of TP samples. This indicates that the proposed model has excellent performance. Table 2 analyzes the comprehensive performance of the model.

In Table 2, on Dataset A, the accuracy of Bi-LSTM-CBM reaches 0.98, the recall rate is 0.97, the F1 score is 0.97, the RMSE is 0.02, and the convergence rate is 31. On Dataset B, its accuracy and recall are 0.97 and 0.98, F1 score and RMSE are 0.99 and 0.05, and the convergence rate is 32. On Dataset A, the accuracy, recall, and F1 score of Bi-GRU-CBM are 0.91, 0.89, and 0.90, with an RMSE of 0.08 and a convergence rate of 48. On Dataset B, the recall, accuracy, and F1 score of Bi-GRU-CBM are 0.91, 0.86, and 0.95, with an RMSE of 0.12 and a convergence rate of 46. Therefore, the proposed Bi-GRU-CBM detection model has excellent performance.

4.2 Simulation analysis of AEC detection model for construction equipment

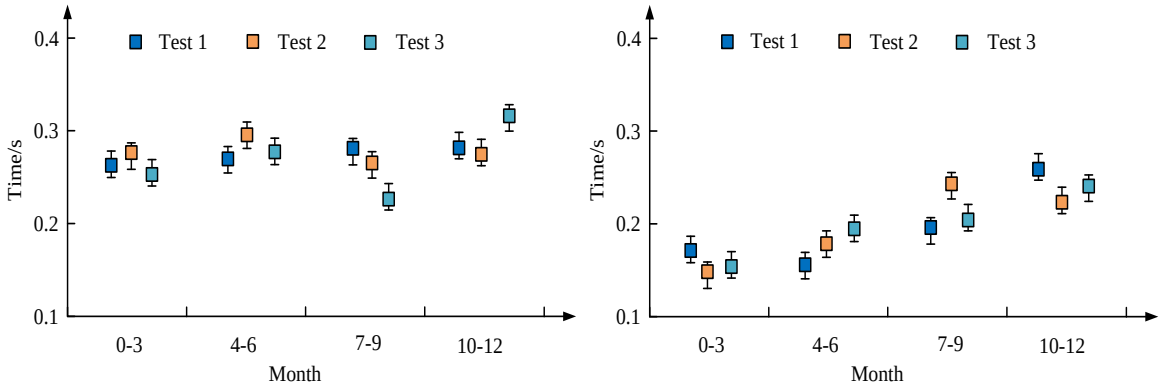
In the anomaly detection simulation, the threshold for anomaly classification (indicated by the red line) is statistically determined based on historical EC data. Specifically, the LECTR threshold is set at the 20th percentile of the device's normal operating range, while the HECTQ threshold is chosen as the 80th percentile of consumption values exceeding the average baseline. These thresholds are derived from the statistical distribution of the energy data, ensuring that the thresholds are contextually relevant to the specific building characteristics and usage patterns. This method provides an objective and data-driven approach to anomaly classification, enhancing the replicability and validity of the results. Additionally, sensitivity analysis is conducted to ensure robustness across varying operational conditions, confirming that the thresholds can generalize to different building types and energy profiles.

The Bi-LSTM-CBAM model performs anomaly classification on a per-time-step basis to capture the temporal dynamics of EC. Each time step is evaluated independently for potential anomalies, with a sliding window of 24 hours used to assess local patterns. The model then assigns an anomaly score to each time-step based on the predicted EC and the established thresholds. For temporal localization, anomalies are flagged at the time-step level, ensuring precise identification of the exact moment when the abnormal consumption occurs. This approach allows the model to detect anomalies in real-time and provides accurate localization of energy usage issues, which is essential for dynamic building management systems. To assess early detection capability, the model is evaluated by introducing an early detection metric: the Time-to-Detection (TTD). TTD is defined as the number of time-steps from the onset of an anomaly until it is detected by the model. In the evaluation, the model demonstrates the ability to detect anomalies within 2-3 hours of their occurrence, achieving early detection for a significant portion of the anomalies across both datasets. To further validate the performance of the model, simulation analysis is conducted using real EC data from two regions for a period of one year, as shown in Figure 9.



(a) Energy consumption in region A (b) Energy consumption in region B

Figure 9: Results of AEC detection model for regions A and B using the model.



(a) Time comparison in region A (b) Time comparison in region B

Figure 10: Analysis of recognition efficiency of the model in different quarters.

In Figure 9, the red lines represent the highest and lowest EC thresholds. Figures 9 (a) and (b) show the AEC detection results for regions A and B, respectively. In Figure 9 (a), the overall EC in region A fluctuates considerably, with actual EC in January, March, and July significantly exceeding the predicted baseline. A two-sample t-test conducted on these months confirms that the mean EC values differ significantly from normal consumption levels (p<0.01), indicating strong evidence of AEC. Conversely, from October to December, EC remains consistently below the baseline, corresponding to a statistically significant drop in energy usage (p<0.05). These findings demonstrate that the Bi-LSTM-CBAM model accurately captures abnormal consumption trends and seasonal deviations in region A. In Figure 9 (b), region B exhibits distinct abnormal peaks in February, May, and from September to November. These deviations are also statistically significant compared to baseline consumption patterns (p<0.01), suggesting the presence of sustained AEC events likely caused by extended high-load operations, climatic factors, or management inefficiencies. Despite the structural and seasonal differences between regions A and B, the Bi-LSTM-CBAM model maintains stable detection performance, with over a 95% confidence level in identifying true anomalies across both locations. This supports the model's robustness and generalization capability in practical multi-regional building environments. Figure 10 analyzes the predictive efficiency of each model.

Figures 10 (a) and (b) illustrate the temporal inference trends of the Bi-LSTM-CBAM model under different seasonal conditions in regions A and B. In Figure 10 (a), the average inference time during January to March remains relatively low, ranging from 0.20 to 0.25 seconds. The time increases slightly from April to June, reaching approximately 0.30 seconds, and rises further from July to September. Statistical analysis using one-way analysis of variance indicates that the differences among the three seasonal phases in region A are statistically significant (p<0.05), suggesting that seasonal variations in EC patterns can influence the model's execution time. In Figure 10 (b), region B demonstrates generally shorter inference times. From January to March, the time stays between 0.15 and 0.20 seconds. While a gradual increase is observed from April to June, it remains under 0.25 seconds. From July to September, inference time increases further, and from October to December, it continues to rise with higher volatility. An independent samples t-test comparing overall inference times between regions A and B confirms a statistically significant difference (p<0.01), with region B showing faster average inference performance. These findings indicate that the Bi-LSTM-

CBAM model maintains high computational efficiency across both regions. The faster inference observed in region B may result from simpler energy usage patterns, fewer anomalies, or lower sequence complexity during testing. Table 3 conducts ablation experiments on the model.

In Table 3, Bi-LSTM-CBM performs the best, with an accuracy of 0.98, a recall of 0.97, an RMSE of 0.02, an MAE of 0.015, an AUC of 0.99, an F1 score of 0.975, and the fastest convergence rate, requiring only 35 rounds of training convergence. This model combines the bidirectional temporal feature extraction ability of Bi-LSTM and the attention enhancement effect of CBAM, making it superior to other models in all indicators. The accuracy of Bi-LSTM without CBAM decreases to 0.91, the recall rate decreases to 0.89, the F1 score is 0.9, and the RMSE and MAE increase to 0.08 and 0.06, indicating that CBAM plays an important role in improving classification accuracy and reducing errors. Although the LSTM-CBAM model uses CBAM, its performance is lower than Bi-LSTM-CBM due to the lack of a Bi-LSTM structure, with an accuracy of only 0.86, an F1 score of 0.84, and RMSE and MAE increasing to 0.1 and 0.075, respectively. This indicates that the Bi-LSTM structure has stronger feature capture capability when processing TSD. In the experiment, the combination of Bi-LSTM and CBAM significantly improves the classification performance. Adding CBAM reduces RMSE from 0.08 to 0.02 and improves F1 score from 0.90 to 0.97, confirming that CBAM effectively guides the model to focus on the most critical energy features and time steps. This justifies the relevance of CBAM to anomaly detection tasks in building EC, as it enhances sensitivity to sparse and transient abnormal patterns that traditional recurrent networks may overlook. The Bi-LSTM-CBM scheme exhibits optimal performance in accuracy, convergence rate, and error control, demonstrating the effectiveness of the Bi-LSTM structure and CBAM. The baseline model is selected and its performance is analyzed. The results are shown in Table 4.

Table 4 presents the performance comparison between the proposed Bi-LSTM-CBAM model and three baseline models, including Bi-LSTM, LSTM, and GRU, on both Dataset A and Dataset B. This demonstrates the superiority of the proposed approach in classification accuracy and error minimization. On Dataset A, the Bi-LSTM-CBAM model achieves an accuracy of 0.98 and an F1 score of 0.97. The RMSE of Bi-LSTM-CBAM is as low as 0.02, indicating that its predictions align closely with actual EC patterns, while the RMSE values of LSTM and GRU are 0.16 and 0.15, respectively. This shows that the baseline models are more prone to prediction errors. On Dataset B, which contains more complex and highly fluctuating EC sequences, the proposed model maintains a high accuracy of 0.97 and an F1 score of 0.99. The RMSE of Bi-LSTM-CBAM on Dataset B is 0.05, which is considerably lower than 0.17 for Bi-LSTM and nearly three times lower than 0.19 for LSTM, highlighting its robust ability to capture transient and long-term anomaly patterns. These results indicate that integrating bidirectional temporal modeling with dual attention enables the proposed model to identify AEC more effectively, reducing false alarms and missed detections compared with classical RNN-based methods.

Table 3: Analysis of ablation experiments.

Model	Accuracy	Recall	F1 score	RMSE	MAE	AUC	Training time/s	Convergence rate
Bi-LSTM-CBAM	0.98	0.97	0.97	0.02	0.015	0.99	120	35
Bi-LSTM	0.91	0.89	0.90	0.08	0.06	0.94	110	40
LSTM-CBAM	0.86	0.82	0.84	0.10	0.075	0.90	105	45
LSTM	0.80	0.76	0.78	0.13	0.095	0.87	100	50

Table 4: Baseline model comparison table.

Dataset	Model	Accuracy	Recall	F1 Score	RMSE
Dataset A	Bi-LSTM-CBAM	0.98	0.97	0.97	0.02
	Bi-LSTM	0.80	0.76	0.78	0.13
	LSTM	0.75	0.71	0.73	0.16
	GRU	0.77	0.73	0.75	0.15
Dataset B	Bi-LSTM-CBAM	0.97	0.98	0.99	0.05
	Bi-LSTM	0.82	0.74	0.88	0.17
	LSTM	0.78	0.7	0.74	0.19
	GRU	0.80	0.72	0.76	0.18

5 Discussion

The experimental findings validate that the Bi-LSTM-CBAM model exhibits strong performance in detecting AEC patterns across various building types. Its superior accuracy, lower RMSE, and faster convergence rate compared to traditional recurrent models can be attributed to two primary design advantages. The bidirectional structure of Bi-LSTM captures historical and future time dependencies. The CBAM module enhances the model's focus on information energy characteristics and critical time steps. Compared with LSTM-BP or CNN-LSTM, which focus on one-way sequence learning or rely only on convolutional feature extraction, the Bi-LSTM-CBAM architecture introduces more comprehensive temporal modeling capabilities and more sophisticated attention-based feature selection. This enables the model to effectively capture both transient and persistent anomalies in EC, particularly under conditions with complex load fluctuations or irregular usage patterns. The combination of dual-directional sequence encoding and dual-domain attention provides a more adaptive and sensitive detection framework. Despite these advantages, the model has certain limitations. Its performance may degrade in scenarios with highly imbalanced datasets or scarce anomaly labels, where the learning process tends to be biased toward dominant normal patterns. Additionally, the architecture introduces increased computational complexity due to the recurrent and attention components, making real-time deployment in low-resource environments challenging. To address these issues, future research can explore lightweight variants of the model, integrate data augmentation techniques to enrich rare anomaly samples, or incorporate transformer-based encoders to improve the modeling of long-range dependencies with higher efficiency. Furthermore, semi-supervised or self-supervised learning frameworks may help the model generalize better in scenarios with limited labeled anomalies.

6 Conclusion

This study proposed a Bi-LSTM-CBAM-based model for detecting AEC patterns in building equipment using time-series data. By incorporating bidirectional temporal modeling and dual-attention mechanisms, the model demonstrated competitive performance in identifying both transient and sustained anomalies across two public datasets. Key features such as LECTR and HECTQ and dynamic time warping distances were extracted and leveraged for accurate classification. Despite these results, several limitations remain. The current approach assumes clean and complete input data, which may not reflect real-world conditions where sensor noise, data loss, or inconsistent reporting intervals are common. The model’s robustness under noisy or missing data is not explicitly evaluated and should be addressed in future work. Furthermore, while inference latency is measured, full-system deployment constraints, such as edge device memory limits and streaming throughput, are not considered. These factors may impact the model’s viability for real-time building energy management. Finally, the current binary classification framework limits interpretability in practical applications. Future research will investigate multi-class extensions for fault type identification, as well as techniques such as semi-supervised learning, data augmentation, and transformer-based encoders to improve generalization, especially in low-resource or dynamic environments. Despite the excellent performance, the model exhibits certain limitations when encountering scarce or highly imbalanced anomaly data. In such scenarios, rare or short-duration abnormal consumption patterns are underrepresented during training, leading to biased learning toward normal behaviors. As a result, the model may suffer from reduced recall, increased false-negative rates, and limited generalization ability in real-world applications where anomalies are infrequent but critical. This sensitivity to data scarcity suggests the need for enhanced anomaly synthesis techniques, data augmentation, or semi-supervised learning frameworks to improve robustness under imbalanced conditions. Future research can be expanded to comprehensive monitoring of multiple energy types to enhance the comprehensiveness of energy management, and can be combined with multi-source data fusion technology to improve the adaptability of the model in complex environments. Going forward, Transformer-based architectures provide a promising direction to further enhance the model's ability to capture long-range dependencies and fine-grained temporal patterns in EC sequences. Unlike RNNs, Transformers rely on self-attention mechanisms that allow parallel processing and dynamic weighting of all time steps, making them particularly effective for modeling sparse, non-stationary, or irregular anomaly patterns over extended periods. Integrating a Transformer encoder or building a hybrid framework that combines Bi-LSTM for local sequence modeling with a Transformer layer for global context extraction can significantly improve the model's interpretability, scalability, and performance in complex operating environments. Furthermore, multi-head attention mechanisms within Transformers can facilitate nuanced learning of interactions among multiple energy indicators, enabling more robust detection of subtle consumption anomalies often missed by traditional recurrent models.

While the current model focuses on binary classification (normal vs. abnormal) for AEC detection, there is potential for expanding the model to a multivariate classification framework, such as fault type identification or categorization of different EC patterns. Identifying specific fault types (e.g., electrical faults, HVAC system malfunctions, etc.) can provide valuable insights for proactive building management, enabling more targeted interventions and reducing downtime. However, transitioning to multivariate classification introduces new challenges, including the need for additional labeled data, the complexity of managing multiple categories, and the risk of class imbalance in fault type distribution. This will be an area of future work. The goal is to explore advanced multi-class classification techniques, such as hierarchical classification models or multi-output models, to improve

diagnostic capabilities and extend model applicability to real-world fault detection systems.

Author contribution

X.W. provided the concept, designed the experiment and wrote the manuscript; X.C. revised manuscript and validated the experiment, L.D. analyzed the data and revised the manuscript. All authors reviewed the manuscript and approved the submission.

Fundings

The research is supported by: Young and Middle-aged Academic Leaders of Qinglan Project of Wuxi City Vocational and Technical College in 2023 (NO:070300200208).

References

[1] Yuexiang Wang, Mingzhi Cui, Bing Xie, Quan Li, Xu Li, Youbin Wu, Ranhong Xie, and Jiangfeng Guo. Tight sandstone reservoir classification based on 1DCNN-BLSTM with conventional logging data. Acta Geophysica, 73(3):2373-2389, 2025.https://doi.org/10.1007/s11600-024-01506-0

[2] Ömer Harun Özkaynak, and Gönül Tuğrul İçemer. Determining the bilge water waste risk and management in the gulf of antalya by the monte carlo method. Journal of the Air & Waste Management Association, 71(12):1545-1554, 2021.https://doi.org/10.1080/10962247.2021.1972055

[3] Tian Xiao, Zhengguang Liu, Liu Lu, Hongcheng Han, Xinyu Huang, Xinyi Song, Xiaohu Yang, and Xiangzhao Meng. LSTM-BP neural network analysis on solid-liquid phase change in a multi-channel thermal storage tank. Engineering Analysis with Boundary Elements, 146(12):226-240, 2023.https://doi.org/10.1016/j.enganabound.2022.10.014

[4] Qinghe Zheng, Ruoyu Wang, Xinyu Tian, Zhiguo Yu, Hongjun Wang, Abdussalam Elhanashi, and Sergio Saponara. A real-time transformer discharge pattern recognition method based on CNN-LSTM driven by few-shot learning. Electric Power Systems Research, 219(6):1-12, 2023.https://doi.org/10.1016/j.epsr.2023.109241

[5] Chunli Liu, Yang Zhang, Jianrui Sun, Zhenhua Cui, and Kai Wang. Stacked bidirectional LSTM RNN to evaluate the remaining useful life of supercapacitor. International Journal of Energy Research, 46(3):3034-3043, 2022.https://doi.org/10.1002/er.7360

[6] Felor Beikzadeh Abbasi, Ali Rezaee, Sahar Adabi, and Ali Movaghar. Fault-tolerant scheduling of graph-based loads on fog/cloud environments with multi-level queues and LSTM-based workload prediction. Computer Networks, 235(10):1.1-1.16, 2023.https://doi.org/10.1016/j.comnet.2023.109964

[7] Bahereh Vojdani, Morteza Rahbar, Mohammadreza Fazeli, Mohammad Hakimazari, and Holly W. Samuelson. Comparative study of optimization methods for building energy consumption and daylighting performance. Energy and Buildings, 323(2):141-152, 2024.https://doi.org/10.1016/j.enbuild.2024.114753

[8]Masoud Nasouri, and Navid Delgarm. Efficiency-based pareto optimization of building energy consumption and thermal comfort: a case study of a residential building in Bushehr, Iran. Journal of Thermal Science, 33(3):1037-1054, 2024.https://doi.org/10.1007/s11630-023-1933-5

[9] Hu Jun, and Hu Fei. Research on multi-objective optimization of building energy efficiency based on energy consumption and thermal comfort. Building Services Engineering Research & Technology, 45(4):391-411, 2024.https://doi.org/10.1177/01436244241240066

[10] Min-Der Lin, Ping-Yu Liu, Chi-Wei Huang, and Yu-Hao Lin. The application of strategy based on LSTM for the short-term prediction of PM_(2.5) in city. The Science of the Total Environment, 906(1):167892.1-167892.10, 2024. https://doi.org/10.1016/j.scitotenv.2023.167892

[11] Luigi Martirano, Alessandro Ruvio, Matteo Manganelli, Federico Lettina, Andrea Venditti, and Gianluca Zori. High efficiency lighting systems with advanced controls. IEEE Transactions on Industry Applications, 57(4):3406-3515, 2021.https://doi.org/10.1109/TIA.2021.3075185

[12] Yi Yan, Fan Wang, Chenlu Tian, Wei Xue, Longlong Lin, Wenpeng Cao, and Linfei Qiao. Transfer learning and source domain restructuring-based BiLSTM approach for building energy consumption prediction. International Journal of Green Energy, 22(3):536-550, 2025.https://doi.org/10.1080/15435075.2024.2421328

[13] Anuruddha Jayasuriya, Matthew J. Bandelt, and Matthew P. Adams. Stochastic mesoscopic modeling of concrete systems containing recycled concrete aggregates using monte Carlo methods. ACI Materials Journal, 119(2):3-18, 2022.https://doi.org/10.14359/51734483

[14] Chenyang Cai, Zechang Wei, Chunxiang Ding, Bianjing Sun, Wenbo Chen, Christoph Gerhard, Evgeny Nimerovsky, Yu Fu, and Kai Zhang. Dynamically tunable all-weather daytime cellulose aerogel radiative supercooler for energy-saving building. Nano Letters, 22(10):4106-4114, 2022.https://doi.org/10.1021/acs.nanolett.2c00844

[15] Chunbo Zhang, Mingming Hu, Benjamin Sprecher, Xining Yang, Xiaoyang Zhong, Chen Li, and Arnold Tukker. Recycling potential in building energy renovation: a prospective study of the Dutch residential building stock up to 2050. Journal of Cleaner Production, 301(7):12683-12700, 2021.https://doi.org/10.1016/j.jclepro.2021.126835

[16] Iná Maia, Lukas Kranzl, and Andreas Müller. New step-by-step retrofitting model for delivering optimum timing. Applied Energy, 290(5):1164-1189,

2021.https://doi.org/10.1016/j.apenergy.2021.116714