Computers & Security 148 (2025) 104130

Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

An explainable unsupervised anomaly detection framework for Industrial Internet of Things

Yilixiati Abudurexiti [a], Guangjie Han [a,*], Fan Zhang [a], Li Liu [b]

[a] College of Information Science and Engineering, Hohai University, Changzhou 213200, China
[b] School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China

ARTICLE INFO

Keywords:
Anomaly detection
Unsupervised framework
Explainable
Multivariate time series

ABSTRACT

Industrial Internet of Things (IIoT) systems require effective anomaly detection techniques to ensure optimal operational efficiency. However, constructing a suitable anomaly detection framework for IIoT poses challenges due to the scarcity of labeled data. Additionally, most existing anomaly detection frameworks lack interpretability. To tackle these issues, an innovative unsupervised framework based on time series data analysis is proposed. This framework initially detects anomalous patterns in IIoT sensor data by extracting local features. An improved Time Convolutional Network (TCN) and Kolmogorov–Arnold Network (KAN) based Variational Auto-Encoder (VAE) is then constructed to capture long-term dependencies. The framework is trained in an unsupervised manner and interpreted using Explainable Artificial Intelligence (XAI) techniques. This approach offers insightful explanations regarding the importance of features, thereby facilitating informed decision-making and enhancements. Experimental results demonstrate that the framework is capable of extracting informative features and capturing long-term dependencies. This enables efficient anomaly detection in complex, dynamic industrial systems, surpassing other unsupervised methods.

1. Introduction

In the burgeoning era of Industry 4.0, the Industrial Internet of Things (IIoT) has emerged as a cornerstone of modern industrial systems, driving unprecedented levels of automation and data exchange. IIoT offers significant improvements in operational efficiency and productivity. Research indicates that the volume of IIoT-linked devices is forecasted to significantly increase in the upcoming years, reaching an estimated 29.42 billion units by 2030 (Shtayat et al., 2023). The worldwide market size for IoT is projected to escalate to $156.7 billion by the year 2025, with IIoT accounting for a large share (Al-Sarawi et al., 2020). Accompanying this growth is a surge in data production and interconnectivity. While these developments confer significant advantages, they concurrently amplify the vulnerability of IIoT systems to cyber threats (Karmakar et al., 2019). In 2015, the Ukrainian power grid was targeted by hackers using BlackEnergy malware, leading to extensive blackouts (Zhang et al., 2019a). Similarly, in 2021, a cyberattack caused the Colonial Pipeline system to cease operations for six days (Tsvetanov and Slaria, 2021). Therefore, it is essential to develop a robust anomaly detection framework to mitigate the impact of vulnerabilities on IIoT, ensuring the system continuity and reliability. Anomaly detection is characterized as the procedure of recognizing patterns that diverge from expected behaviors or norms (Fahim and Sillitti, 2019). In the context of IIoT, anomalies can occur for various reasons such as network attacks, virus intrusion, operational failures, and sensor malfunction (Fahim and Sillitti, 2019). Anomalies can manifest in different forms, such as spikes, drops, trends, oscillations, and correlations (Zolanvari et al., 2019). Fig. 1 illustrates the alterations in data from temperature sensors and accelerometers during an anomalous event within the water circulation system. This anomaly, observed within the water treatment process, results in discernible fluctuations in the sensor-recorded data.

Time series analysis is a statistical technique utilized to analyze and model data collected sequentially over time. It is commonly employed to identify patterns, trends, and relationships, enabling predictions based on them (Pang et al., 2022). This form of analysis has found widespread applications in various fields such as finance (Sezer et al., 2020), weather forecasting (Karevan and Suykens, 2020), and healthcare (Puri et al., 2022). Within the framework of IIoT, time series data, distinguished by its high dimensionality, frequency, intricacy,

* Corresponding author.
E-mail address: hanguangjie@gmail.com (G. Han).

https://doi.org/10.1016/j.cose.2024.104130
Received 28 July 2024; Received in revised form 22 August 2024; Accepted 17 September 2024
Available online 25 September 2024

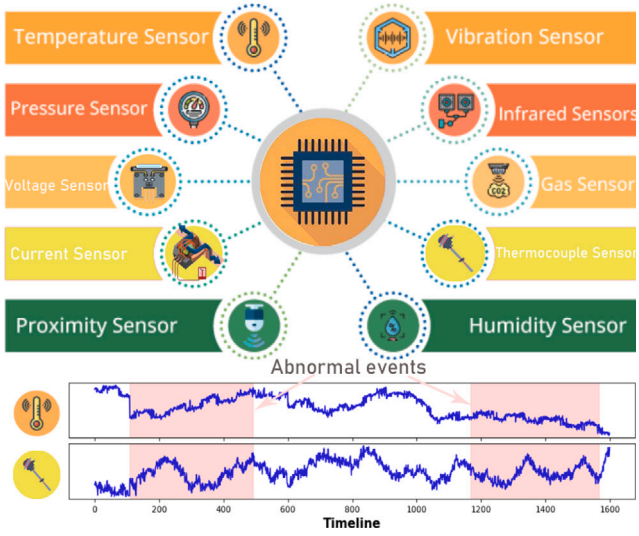


Fig. 1. Anomaly detection in IIoT.

and noise, assumes a vital part in anomaly detection. A multitude of time series analysis techniques have been put forward to address the intricacies of detecting anomalies in the IIoT. Khan et al. (2022) put forth a method for detecting anomalies that employs Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) to identify network attacks within the IIoT. In Wu et al. (2020), researchers presented a method for anomaly detection in IIoT by combining a stacked LSTM with the Gaussian Bayesian model. The authors presented a model that integrates federal learning, CNNs, LSTM, and attention mechanism to detect anomalies (Liu et al., 2021).

The reliance of these methodologies on the availability of substantial quantities of labeled data for training and validation purposes, that might be both expensive and time-intensive to gather (Nizam et al., 2022). This reliance constitutes a notable impediment, especially in situations where labeled data is limited or costly to procure. In contrast, unsupervised anomaly detection methods offer the advantage of autonomously identifying abnormalities in IIoT data, eliminating the need for labeled datasets. Su et al. (2019), present OmniAnomaly an unsupervised approach that leverages a stochastic recurrent neural network for anomaly identification through the analysis of reconstruction probabilities. A rapid and reliable method for anomaly detection, which relies on adversarial autoencoders, is presented in Audibert et al. (2020). Li et al. (2019) unveiled an unsupervised technique for multivariate anomaly detection that uses Generative Adversarial Networks (GANs) to simultaneously process all variables and discern latent correlations among them. In paper Zheng et al. (2019), the researchers design a one-class adversarial Network (OCAN) with a focus on fraud detection.

Despite the significant strides made within the domain of unsupervised anomaly detection specific to IIoT, certain obstacles continue to hinder its full potential. Two of the most prominent challenges are the efficient use of scarce labeled data and the interpretability of the employed method. Given the persistent challenges, our research presents an innovative approach to unsupervised anomaly detection in the context of IIoT. This approach aims to improve not only the detection performance but also the interpretability of the framework, thereby offering a holistic solution to the existing challenges in the domain.

Initially, our suggested approach commences with the extraction of local features using a one-dimensional CNN coupled with a Convolutional Block Attention Module (CBAM). These extracted local features are then directly fed into a Variational Auto-Encoder (VAE) that is based on the improved TCN and Kolmogorov–Arnold Networks (KAN). Following this, a dynamic scoring function is applied to calculate each data point's anomaly score. A top-k threshold function is then utilized to determine the threshold for detecting anomalies. Comparing the anomaly scores with this threshold enables effective identification of anomalies within the data points. In the final stage, SHapley Additive exPlanation (SHAP) is employed to elucidate how each feature contributes to the prediction, thereby augmenting the interpretability of the framework. The primary objective of this framework revolves around enhancing both the performance and interpretability of unsupervised anomaly detection in IIoT.

The main contributions of our proposed framework are:

(1) We present Convolutional and CBAM-enhanced Temporal Attention KAN (CCTAK), an innovative unsupervised anomaly detection framework for IIoT. This framework is designed to enhance the performance of anomaly detection by strategically extracting local features and capturing long-term dependencies. Moreover, it offers interpretability and valuable insights into the detection process.

(2) We enhance the traditional TCN by integrating a temporal attention mechanism, which is further refined with the inclusion of KAN. This unique combination within the VAE architecture bolsters the capability of capturing long-term temporal dependencies, thereby amplifying the effectiveness in identifying anomalous events.

(3) We implement a dynamic scoring function coupled with a top-k threshold strategy to enhance anomaly detection efficiency. The framework leverages Explainable Artificial Intelligence (XAI) techniques to clarify the complexities involved in making determinations, consequently boosting the interpretability and reliability of the framework.

The rest of this paper unfolds as follows. Section 2 offers an overview of the current studies on anomaly detection in time series data. The suggested framework is unveiled in Section 3. Section 4 delves into the specifics of our experimental design and deliberates on the outcomes. The paper wraps up with Section 5, providing a recap of the discoveries and sketching out possible avenues for subsequent research.

2. Related work

An array of methods aimed at detecting anomalies within time series data has emerged, encompassing statistical approaches, supervised models for anomaly detection, and unsupervised methods (Cook et al., 2020). This section provides a concise overview of current techniques aimed at identifying anomalies within time series data.

2.1. Statistical methods

Statistical methodologies employ a variety of techniques, including hypothesis testing, mean, and standard deviation, to identify anomalies (Cook et al., 2020). The three prevalent statistical methods for anomaly detection are control charts (Albazli et al., 2020), autoregressive integrated moving average models (ARIMA) (Kadri et al., 2016), and seasonal and trend decomposition using Loess (STL) (Wen et al., 2019).

Control charts serve as graphical instruments that track process quality over time by comparing observed data with predefined control limits (Albazli et al., 2020).

ARIMA models the data as a linear combination of past values, errors, and differencing terms. ARIMA can be utilized in identifying anomalies through the calculation of discrepancies from observed to forecasted values (Kadri et al., 2016).

STL decomposes time series data into three separate components: seasonal, trend, and remainder (Wen et al., 2019).

Although statistical methods are straightforward and easy to implement, they do demand high data quality and have limited model generalization capabilities. As a result, they may not be effective in identifying novel or complex anomalies.

2.2. Supervised anomaly detection methods

Supervised learning techniques have garnered interest in the realm of time series anomaly identification, attributed to their ability to learn from annotated instances. Scholars have utilized a range of supervised strategies for identifying anomalies in IIoT systems, such as Support Vector Machines (SVM) (Vos et al., 2022), random forests (Kopp et al., 2020), and LSTM (Zhou et al., 2021). These methods are trained on labeled data, where anomalies are explicitly identified. Subsequently, unseen samples can be categorized as typical or anomalous based on discerned patterns and attributes (Das et al., 2022).

Supervised learning methods, when applied to anomaly detection in IIoT systems, face several challenges. Firstly, they necessitate a substantial volume of annotated data, which may be difficult to procure. Secondly, they may struggle to adapt to new or unseen situations, particularly when data distributions change or anomalies are infrequent. Thirdly, they can suffer from overfitting or underfitting due to complex models and low-quality data. Consequently, it becomes crucial to assess the resilience of these models by employing appropriate metrics and standards (Das et al., 2022).

2.3. Semi-supervised anomaly detection methods

Semi-supervised anomaly detection methods are particularly effective in scenarios where labeled data is scarce. These methods typically leverage a large amount of unlabeled data alongside a small set of labeled examples to train models that can identify anomalies by learning the underlying patterns of normal behavior. Researchers have explored various semi-supervised strategies for anomaly detection in IIoT systems, such as generative model (Abdel-Basset et al., 2023), graph-based method (Zhou et al., 2023), self-clustering method (Ma et al., 2024) and reconstruction based method (Kumaran Santhosh et al., 2022). These approaches share the common feature of combining a small amount of labeled data with a larger pool of unlabeled data to improve the model's ability to identify anomalies. By leveraging both types of data, these methods enhance the generalization capacity of the models, particularly in situations where labeled data is limited. This hybrid strategy allows for the effective learning of normal behavior patterns, making it possible to detect deviations indicative of anomalies.

However, semi-supervised anomaly detection methods also present certain challenges. The effectiveness of these methods heavily depends on the quality of the labeled and unlabeled data, as well as handling imbalanced datasets, and ensuring scalability remain.

2.4. Unsupervised anomaly detection models

Unsupervised anomaly detection seeks to detect anomalous patterns in time series data without relying on prior knowledge or labeled examples (Alghanmi et al., 2022). A multitude of unsupervised techniques for anomaly detection have been suggested in the context of IIoT environments. One common approach is to employ deep learning techniques for measuring the deviation of new data from an established representation of normal data (Alghanmi et al., 2022). Models such as the deep convolutional autoencoding memory network (CAE-M) (Zhang et al., 2023) and the anomaly transformer (AT) model (Xu et al., 2022) have been proposed for IIoT scenarios.

Although these methods are advantageous in that they do not require labeled data, they still face some obstacles. Often, these methods operate under the belief that standard data points are situated in areas of high density, while anomalies are found in regions of low density, a concept that may not always be valid. They are sensitive to hyperparameters and do not utilize available label information, potentially leading to suboptimal performance. Furthermore, they generally offer less interpretability than statistical methods. Therefore, it is crucial to confront these obstacles in order to propel the domain of unsupervised anomaly detection within IIoT forward.

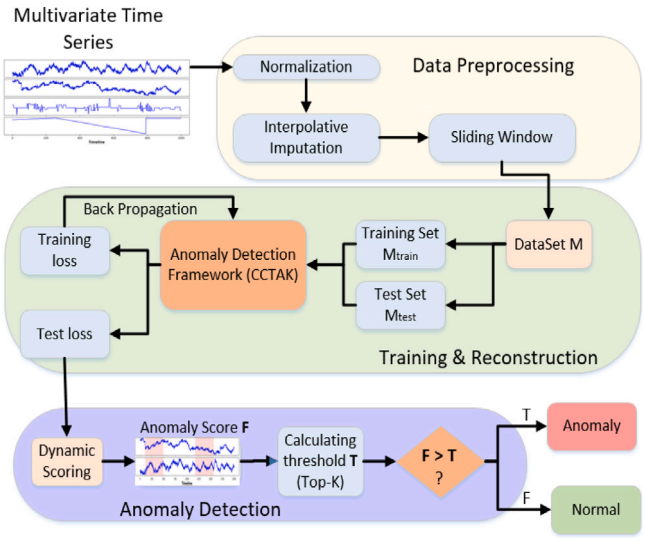


Fig. 2. Flowchart of anomaly detection in IIoT.

3. Proposed framework

3.1. Problem statement

Consider the time series $F = \left(d^1, d^2, ..., d^n\right)^T = \left(d_1, d_2, d_3, \ldots, d_t\right) \in \mathbb{R}^{n\times t}$, where n denotes the count of features, while t signifies the magnitude of timestamps. At a specific timestep t in the time series data, the input is represented as $d_{t=}\left(d_t^1, d_t^2, \ldots, d_t^n\right) \in \mathbb{R}^n$. The categorization of the input series as either univariate or multivariate sequences is contingent upon the quantity of variables present at each timestamp. The task of IIoT anomaly detection addresses two primary issues: (1) detecting all anomaly points, and (2) classifying the anomalies within a given target sequence.

By utilizing sliding windows, the task of detecting anomalies can be converted into a classification or clustering task (Xu et al., 2022). The sliding window can split the entire sequence into consecutive short sequences to obtain a data set $D = \left(w_1, w_2, \ldots, w_{t-L+1}\right) = \left(d_1, \ldots, d_L\right)$, $\left(d_2, \ldots, d_{l+1}\right), \ldots, \left(d_{t-l+1}, \ldots, d_t\right) \in \mathbb{R}^{L\times(t-L+1)}$, where L is the sliding window length. If d_i contains any anomaly, then the original sequence is considered as an anomalous sequence. Consequently, the aim of identifying anomalies within the time series data F is redefined as the task of assigning labels to each input vector in D.

3.2. Anomaly detection with unsupervised framework

The procedure for detecting anomalies, depicted in Fig. 2, initiates with the normalization, interpolation, and sliding window segmentation of data from diverse IIoT sensors. The processed data is divided into subsets for training and testing. The training subset is fed into the CCTAK framework for learning and updating parameters. Upon completion of training, the framework undergoes evaluation on the test set to determine the loss during testing. Anomaly scores are then calculated using a dynamic scoring function based on the test loss. A threshold is determined using the top-k threshold calculation method, and anomalies are detected by contrasting the anomaly scores with this threshold.

The architecture of the proposed CCTAK framework, as shown in Fig. 3, begins by extracting local features from IIoT time series data using a pair of one-dimensional CNN convolutional blocks. A CBAM attention layer is then applied to capture complex patterns. The extracted local features are then input into an improved TCN and KAN based VAE model for compression and reconstruction, which enables the capture

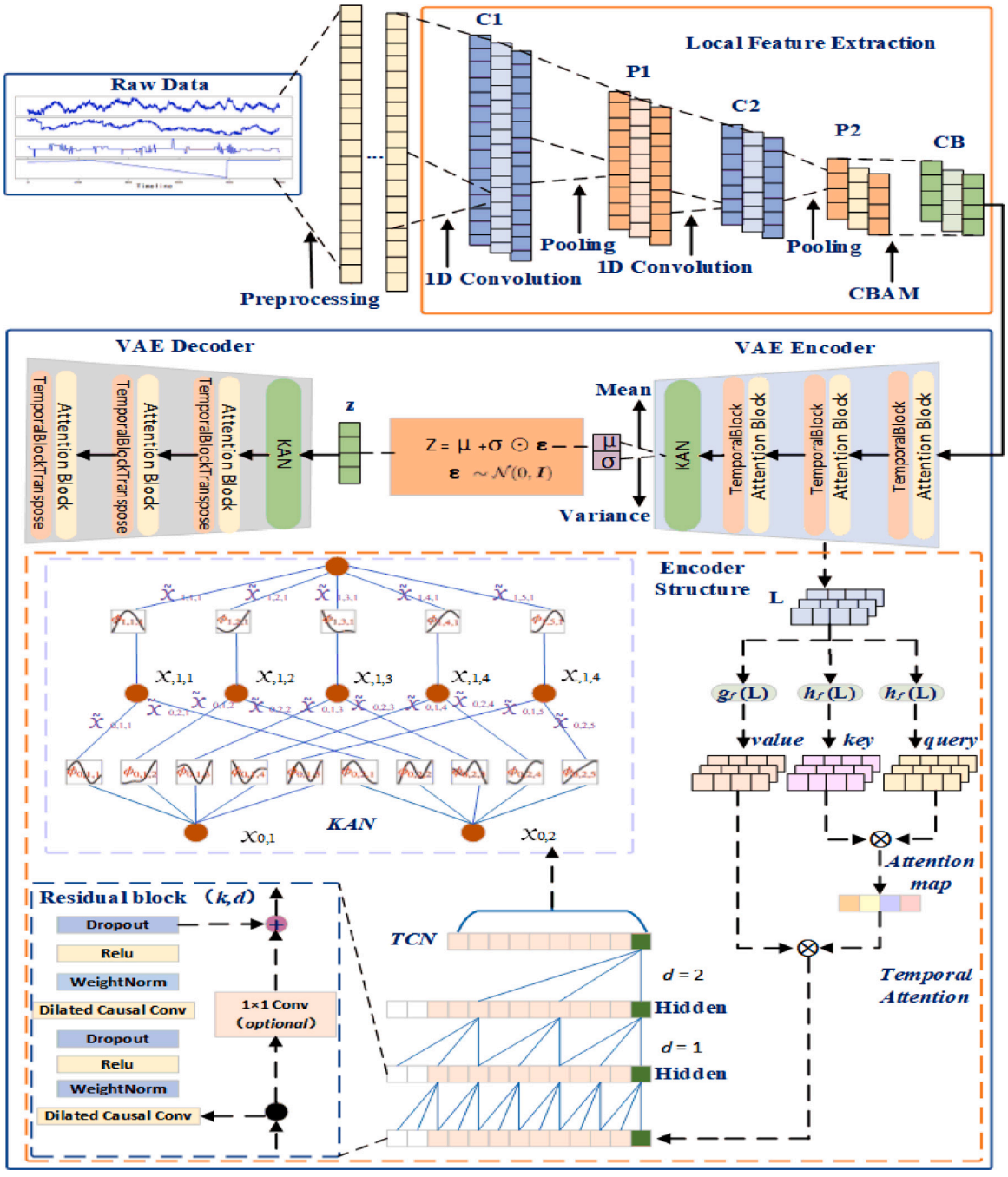


Fig. 3. Overall architecture of the proposed CCTAK.

of long-term dependencies. The Gaussian-D scoring function is used to compute the anomaly score, and a top-k threshold function is applied to determine the threshold value. Through the assessment of the anomaly score against the threshold value, the data points are distinguished as either normal or anomalous. The framework comprises four main components: a feature extraction component (i.e. convolutional blocks and CBAM), a long-term dependency capture component (i.e. improved TCN and KAN based VAE), and components for computing anomaly scores and thresholding (i.e. Gaussian-D and top-k thresholding functions). A comprehensive elucidation of each component is provided below:

(1) Convolutional block: This framework processes time series data using two convolutional blocks. Each consists of a one-dimensional (1D) convolutional layer, a pooling layer, and a ReLU activation function. The formula for processing time series data using the 1D convolution is as follows:

$$y_t = \sum_{i=1}^{k} x_{t+i-1} w_i + b, \tag{1}$$

where y_t denotes the output at time step t, x_{t+i-1} is the input at time step $t+i-1$, w_i represents the weight of the convolutional kernel at position i, b is the bias term, k refers to the size of the kernel.

The output of two 1D convolution blocks is given by:

$$h = act\,(MaxPool(Conv(act(MaxPool(Conv(\mathrm{x})))))), \tag{2}$$

where x is the inputted time series data, Conv is the 1D convolution operation, $MaxPool$ is the maximum pooling operation, and act is the ReLU activation function.

(2) CBAM (Woo et al., 2018): It is a self-attention mechanism that enhances the framework's focus on critical information. It encompasses two sub-modules, namely the channel attention module, which generates vectors and computes channel weights, and the spatial attention module, which generates a matrix and computes spatial location weights. The output generated by the CBAM attention layer is:

$$X_{cbam} = Sigmoid\left(Conv\left(AvgPool\left(X_{cnn}\right) + MaxPool\left(X_{cnn}\right)\right)\right) \odot X_{cnn}, \tag{3}$$

where X_{cnn} is the feature extracted from the 1D convolutional block, $\odot$ denotes Hadamard product, $AvgPool$ is the average pooling operation, and $Sigmoid$ is the sigmoid activation function.

(3) The improved TCN and KAN based VAE: It is a VAE-based method for time series data compression and reconstruction. It uses enhanced TCN and KAN to capture long-term dependencies and dynamics of the data. The VAE is designed to convert high-dimensional data into a compressed latent space, subsequently generating novel data samples. It is composed of two primary components: an encoder and a decoder. The encoder's role is to transform the input into the mean and variance of a concealed variable, while the decoder's function is to rebuild the input from this variable. The VAE's objective function is inclusive of the reconstruction error and the KL divergence. These measure the disparity between the input and the reconstructed data, and the deviation between the posterior and prior distributions of the latent variables, respectively. The mathematical representation of the objective function is as follows:

$$L_{vae=}E_{q(z|x)}\left[\log p(x|z)\right] - D_{KL}\left(q(z|x) \parallel p(z)\right), \tag{4}$$

where $q(z|x)$ is the posterior distribution of the hidden variable, $p(x|z)$ is the conditional distribution of the reconstructed data, and $p(z)$ is the prior distribution of the hidden variable, which is typically presumed to follow a standard normal distribution.

TCN is a deep neural network that employs causal convolution, which can effectively handle time series data of varying lengths while preserving the temporal order consistency. TCN consists of multiple residual blocks, each of which includes two dilated convolution layers, a residual connection, a chopping function, and a $ReLU$ activation function.

The TCN in our framework consists of three temporalblocks, the output of the TCN can be expressed as:

$$X_{tcn} = TemporalBlock_3\left(TemporalBlock_2\left(TemporalBlock_1\left(X\right)\right)\right), \tag{5}$$

where each $TemporalBlock$ can be represented by the following formula

$$X_{tb} = (X + LayerNorm(ReLU(Conv1D(LayerNorm(ReLU((Conv1D(X; k, d))); k, d))))), \tag{6}$$

where $Conv1D$ is one dimensional dilated convolution, $ReLU$ is the activation function, $LayerNorm$ denotes normalization, k stands for the number of the convolution kernel, and d is indicative of the dilation factor.

In addition to the traditional TCN, we introduce a temporal attention mechanism. This mechanism is designed to accentuate the most significant temporal features in time series data. It applies attention weights to the output of each TCN residual block, which enables the network to concentrate on the most relevant time steps. This enhancement makes the VAE more effective for time series anomaly detection.

The temporalBlock after the introduction of Temporal Attention mechanism can be expressed as:

$$X_{tb} = X + LayerNorm(ReLU(Conv1D(a(LayerNorm(ReLU(Conv1D(a(X); k, d))))))), \tag{7}$$

The attention weight 'a' can be derived as the equation given below:

$$a(X) = softmax\left(\frac{(W_qX)(W_kX)^T}{\sqrt{d_k}}\right)(W_vX), \tag{8}$$

where W_q, W_k and W_v are the linear transformation matrices used to compute the query, key and value, respectively, and d_k is the dimension of key.

In addition, the VAE encoder is improved with the incorporation of Kolmogorov–Arnold Networks (KAN) (Liu et al., 2024). KANs represent an innovative neural network structure, drawing inspiration from the Kolmogorov–Arnold theorem. Contrasting with conventional neural networks that utilize static activation functions, KANs incorporate adaptable activation functions at the network edge. The learnable activation functions in KAN provide the model with additional flexibility. The purpose of integrating KAN into the VAE encoder is to further enhance the framework's capacity to encapsulate intricate temporal patterns within the data. The computation of KAN can be articulated through the subsequent equation:

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{2n+1} \psi_i \left(\sum_{j=1}^{n} \phi_{ij}(x_j)\right), \tag{9}$$

where $f : [0,1]^n \rightarrow \mathbb{R}$ is a continuous function, $x_1, x_2, \ldots, x_n \in [0,1]$, $\phi_i : [0,1] \rightarrow \mathbb{R}$ and $\psi_i : \mathbb{R} \rightarrow \mathbb{R}$ are some appropriate continuous functions.

The output of the TCN-KAN-VAE structure can be expressed as:

$$z = TCN - KAN\left(X_{cbam}\right), \tag{10}$$

$$X_{rec} = TransposeConv(TCN - KAN(z)), \tag{11}$$

where z is the sampling of the hidden variable, and $Transpo- seConv$ is the transposed convolution operation.

(4) Selection of scoring and threshold functions: A scoring function is employed to assess the difference between the input and output data of the VAE, serving as a crucial component in anomaly detection. A threshold function is also defined to determine if the score surpasses a certain value. For these purposes, we select a dynamic Gaussian scoring function (Gauss-D) and a top-k threshold function. The Gauss-D scoring function assumes that the anomaly scores adhere to a Gaussian distribution characterized by a time-dependent mean and standard deviation. These parameters are estimated using an Exponential Moving Average (EMA) method, which gives more weight to recent observations. The Gauss-D score is then calculated as the negative log-likelihood of the reconstruction error under the Gaussian distribution

$$S_t = -\log\left(\frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(e_t-\mu_t)^2}{2\sigma_t^2}}\right), \tag{12}$$

where e_t denotes the reconstruction error at the instance t, while μ_t and σ_t stand for the mean and standard deviation of the reconstruction error at the same instance t, respectively. A comprehensive sketch of the suggested anomaly detection framework is provided in Algorithm 1.

The EMA method updates μ_t and σ_t as follows:

$$\mu_t = \alpha e_t + (1-\alpha)\mu_{t-1}, \tag{13}$$

$$\sigma_t = \sqrt{\beta(e_t-\mu_t)^2 + (1-\beta)\sigma_{t-1}^2}, \tag{14}$$

where α and β are smoothing parameters that control the decay rate of the EMA. The advantage of the Gauss-D scoring function is that it can adapt to the changing statistics of the reconstruction error over time.

The top-k threshold function operates on the principle that anomalies are infrequent occurrences appearing in the tail end of the score distribution. The top-k threshold function selects a threshold value that corresponds to the kth highest score in a sliding window of size w. The window size w determines how sensitive the anomaly detection is to recent changes in the score distribution. The top-k threshold function is defined as:

$$T_t = \max\left\{S_i | i \in [t-w+1, t], S_i \in \text{top } k \text{ scores}\right\}, \tag{15}$$

The advantage of the top-k threshold function lies in its ability to autonomously modify the threshold value in accordance with the attributes of the data. Should the anomaly score S_t for a data point at time

Table 1
Details of the datasets.

Dataset name	Industrial domain	Number of sensors	Average train size	Average test size	Average number of events	Average Anomalies
SKAB	Water circulation	8	9401	35 600	34	36.70%
SWAT	Water treatment	51	473 400	414 569	35	4.65%
DAMADICS	Sugar manufacturing	32	507 600	217 802	17	1.48%
MSL	Spacecraft	55	2160	2731	35	12.02%

Algorithm 1 Anomaly detection applying the CCTAK

Input: IIoT time series data TS_{raw}
Output: Anomalies
Data-Pre-Processing:
1: Process the missing values in raw data with linear interpolation and backward fill:
$TS_m = Libf(TS_{raw})$
2: Standardized time series data TS_m:
$TS_{norm} = Normalization(TS_m)$
3: Apply sliding window to segment data:
$TS_w = SW(TS_{norm})$
4: Split the dataset randomly into a training set TS_{train} and a test set TS_{test}.
Model Training:
5: Construct CCTAK and initialize parameters with θ
6: **for** $i = 1$ to $epoch$ **do**
7: Compute the model output: $H_{train} = \text{CCTAK}(TS_{train})$
8: Calculate reconstruction loss: $L_{train} = \text{MSELoss}(H_{train}, Y_{train})$
9: Update the model parameters θ
10: **end for**
Anomaly Detection:
11: Compute the model output: $H_{test} = \text{CCTAK}(TS_{test})$
12: Calculate reconstruction loss: $L_{test} = \text{MSELoss}\ (H_{test}, Y_{test},)$
13: Calculate anomaly scores using the Gauss-D scoring function: $Anomaly\ Score = Gauss - D(L_{test})$
14: Calculate the threshold using the top-k function:
$T_t = top - k(Anomaly\ Score)$
15: **if** $Anomaly\ Score >= T_t$ **then**
16: Anomaly data
17: **else**
18: Normal data
19: **end if**

t surpass the threshold T_t, it is flagged as an anomaly. A comprehensive sketch of the suggested anomaly detection framework is provided in Algorithm 1.

3.3. Explanation of the framework

XAI techniques are crucial in enhancing the transparency and interpretability of machine learning models, especially in critical applications such as IIoT systems. In our proposed framework, XAI methods, particularly SHAP, are employed to analyze and visualize the contribution of each feature to the model's decisions. The SHAP values were computed for each feature, providing a clear ranking of feature importance. Based on the SHAP value analysis, we selected the top N most important features and removed less significant ones, which were then used to refine the model. This not only allows stakeholders to understand the rationale behind the anomaly detection results but also facilitates more informed decision-making. XAI thus plays a dual role in our framework by both improving model accuracy through feature selection and ensuring that the model's outputs are understandable and actionable.

SHAP interprets black-box models using Shapley values, a concept from cooperative game theory that quantifies the contribution of each player to the total payoff of a coalition. In our context, each player is a feature in the time series data, and the payoff is the anomaly score computed by our framework. SHAP assigns an importance score to each feature, which reflects its impact on the anomaly score when included or excluded from the coalition. The Shapley value of feature i is calculated as:

$$\phi_i = \sum_{S \subseteq \{x_1,\ldots,x_n\} \setminus \{x_i\}} \frac{|S|!(n-|S|-1)!}{n!} [f_x(S \cup \{x_i\}) - f_x(S)], \tag{16}$$

where $f_x(S)$ is the predicted value given the set of features S, where n denotes the total quantity of features. We employ the KernelExplainer to compute SHAP values, which estimates Shapley values by sampling coalitions from a weighted kernel function.

4. Experimental verification and analysis

4.1. Description of datasets

To evaluate the effectiveness of the anomaly detection approach we utilize four distinct IIoT anomaly datasets. The particulars of these datasets are outlined in Table 1.

(1) Skoltech Anomaly Benchmark (SKAB) (Katser and Kozitsin, 2020): This dataset emulates anomalies within a water circulation system, offering a realistic environment for the testing and evaluation of anomaly detection algorithms.

(2) Secure Water Treatment (SWaT) (Goh et al., 2016): This dataset serves as a significant asset for cybersecurity research within water treatment systems. It offers data derived from a testbed, which replicates a miniaturized model of an industrial water treatment facility.

(3) Development and Application of Methods for Actuator Diagnosis in Industrial Control Systems (DAMADICS) (Damadics, 2020): Originating from the Cukrownia Lublin SA sugar factory in Poland, this dataset is designed for the diagnosis of industrial actuator failures, with a particular focus on control valves.

(4) Mars Science Laboratory Rover (MSL) (Hundman et al., 2018): This dataset offers an extensive array of observational data from the Curiosity rover's mission on Mars, including geological and atmospheric measurements, as well as high-resolution images.

4.2. Data preprocessing

The framework begins by addressing missing data values using linear interpolation and backward padding. It then employs a sliding window technique to segment the data into subsequences. This is followed by data normalization and cropping to condense the data and reduce noise.

The data preprocessing method is adopted to segment the time series data into windows of length WS that overlapped by a step size of SL. We set the window size $WS = 100$ and step size $SL = 1$. After that, the training data is scaled between [0,1] by performing maximum–minimum normalization, and then the test data is cropped in the range [−4,5] to prevent the impact of oversized data on the anomaly scores. Linear interpolation and backward fill are utilized to handle the missing values of the data before model training.

Table 2
Comparison of anomaly detection methods.

Method	SKAB		SWaT		MSL		DAMADIC	
	Ave-Fc_1	AU–ROC scores	Ave-Fc_1	AU–ROC scores	Ave-Fc_1	AU–ROC scores	Ave-Fc_1	AU–ROC scores
LSTM-VAE	0.5439	0.5002	0.4456	0.7570	0.3910	0.5986	0.5999	0.9146
OmniAnomaly	0.5482	0.5058	0.4948	0.7984	0.4120	0.6523	0.1425	0.6804
OCAN	0.5459	0.5108	0.4829	0.7918	0.3009	0.5803	0.2532	0.6848
MSCRED	0.5526	0.5270	0.4315	0.7069	0.3944	0.6436	0.2906	0.6866
TCN-AE	0.5488	0.5002	0.4732	0.7596	0.4354	0.6184	0.5989	0.8109
BeatGAN	0.5437	0.4941	0.4777	0.7889	0.4531	0.6517	0.4531	0.7960
USAD	0.5648	0.5323	0.4986	0.7926	0.4481	0.6406	0.4602	0.8028
InterFusion	0.5964	0.5427	0.5109	0.8013	0.4816	0.6715	**0.6128**	**0.9280**
CCTAK (Ours)	**0.6136**	**0.5509**	**0.5316**	**0.8191**	**0.5053**	**0.6913**	0.6034	0.9109

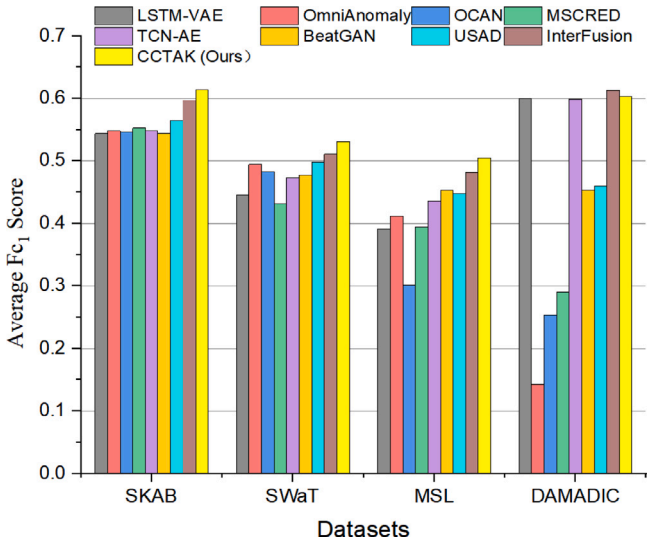


Fig. 4. Anomaly detection results.

Table 3
Impact of different scoring functions on model performance.

Dataset	Model	Normalized Error	Gauss-S	Gauss-D	Gauss-D-k
SKAB	LSTM-VAE	0.4412	0.4480	0.5439	0.5433
	OmniAnomaly	0.4425	0.4521	0.5482	0.5481
	OCAN	0.4369	0.4460	0.5459	0.5325
	MSCRED	0.3926	0.4028	0.5526	0.5502
	TCN-AE	0.4260	0.4390	0.5488	0.5515
	BeatGAN	0.4480	0.4396	0.5437	0.5391
	USAD	0.4216	0.4319	0.5648	0.5571
	InterFusion	0.4537	0.4251	0.5964	0.5812
	CCTAK (Ours)	**0.4669**	**0.4753**	**0.6136**	**0.5986**
SWAT	LSTM-VAE	0.2207	0.0965	0.4456	0.4444
	OmniAnomaly	0.2278	0.1648	0.4948	0.5118
	OCAN	0.2188	0.1547	0.4829	0.4775
	MSCRED	0.1893	0.2142	0.4315	0.3769
	TCN-AE	0.2132	0.2124	0.4732	0.4353
	BeatGAN	0.2574	0.1878	0.4777	0.4760
	USAD	0.2651	0.2490	0.4986	0.4925
	InterFusion	0.2741	0.2284	0.5109	0.5098
	CCTAK (Ours)	**0.2809**	**0.2353**	**0.5316**	**0.5223**

4.3. Evaluation metrics of anomaly detection

In practical scenarios, attaining a high F-score for the selected threshold is of paramount importance (Xu et al., 2018). Numerous studies have employed various F-scores to verify the efficacy of anomaly detection methods, with the point-wise F-score (F_1) being the most prevalent. However, in real-world applications, our emphasis is on the detection of unusual events characterized as a series of linked anomalous time instances, instead of single time points (Garg et al., 2022).

The point-adjusted F-Score (Fpa_1), proposed in Xu et al. (2018), has a drawback in that it may yield high anomaly scores even if numerous anomalous events remain undetected (Garg et al., 2022).

A novel metric for anomaly detection in time series data, namely the composite F_1 score Fc_1, is introduced in Garg et al. (2022). Unlike traditional F_1 scores, Fc_1 is designed to reward high time-wise precision, It serves as a more accurate evaluation metric in scenarios where anomalies are characterized by a series of linked time instances, rather than isolated points. Fc_1 is calculated as the harmonic mean of time-wise precision. Consequently, in this paper, we select Fc_1 and AU-ROC scores as our model evaluation metrics. This dual-metric approach provides a comprehensive evaluation of the model's performance offering a more holistic assessment of the model's performance in detecting complex anomalies in time series data. The results are displayed in Table 2 and Fig. 4.

4.4. Selection of comparison algorithm

The efficacy of the proposed algorithm is showcased through a comparative analysis with various unsupervised detection methods, as outlined below:

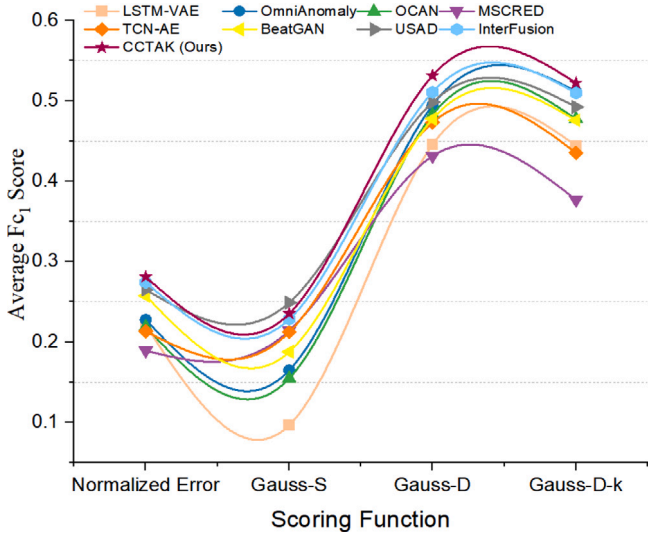


Fig. 5. Impact of different scoring functions on model performance based on top-k threshold function with SWaT dataset.

(1) LSTM-VAE (Chen et al., 2021): This approach depicts the process of data generation from latent representations to observable data, utilizing variational methods during the training phase.

(2) OmniAnomaly (Su et al., 2019): A method that leverages a stochastic recurrent neural network to model the temporal dependencies and uncertainties of multivariate time series.

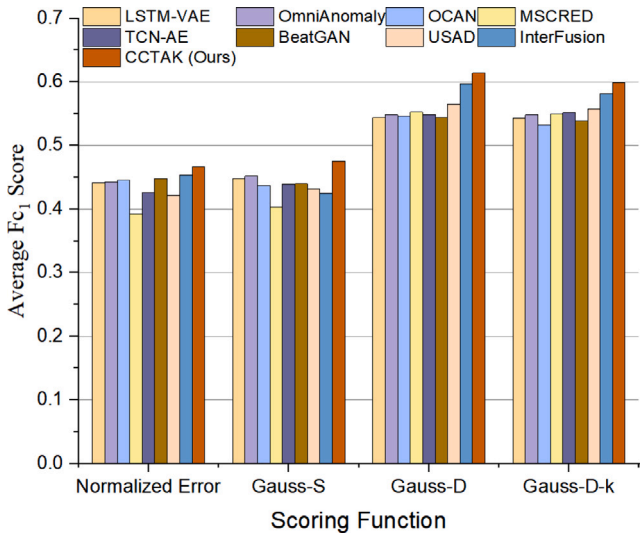


Fig. 6. Impact of different scoring functions on model performance based on top-k threshold function with SKAB dataset.

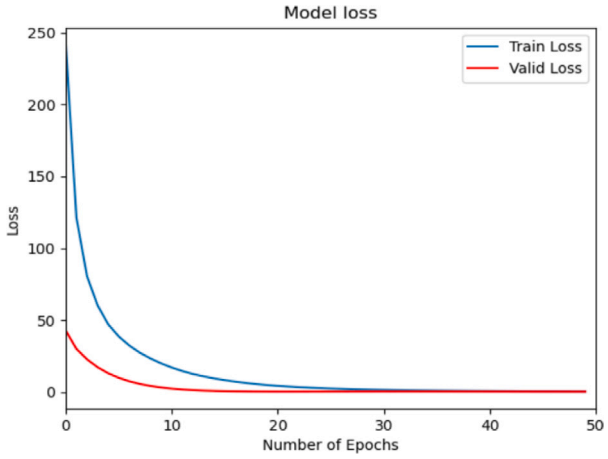


Fig. 7. Training and validation losses.

(3) OCAN (Zheng et al., 2019): It is designed to discern inherent data patterns, it utilizes adversarial learning for distinguishing normal from anomalous behavior, thus offering a robust fraud detection mechanism.

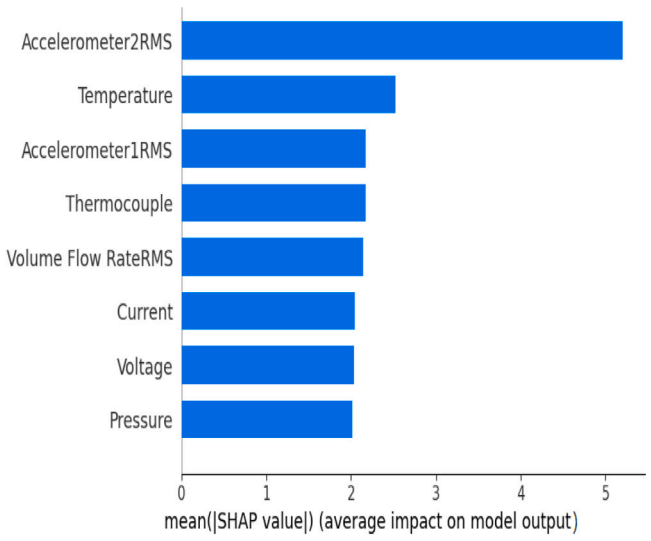
(4) MSCRED (Zhang et al., 2019b): It is designed to acquire the capability of reconstructing signature matrices. These matrices serve as representations of the cross-correlation relationships between channels.

(5) TCN-AE (Bai et al., 2018): Constructed upon the TCN model in our implementation, the encoder consists of a sequence of TCN residual blocks, while the decoder employs transpose convolutions in place of convolutions within TCN residual blocks.

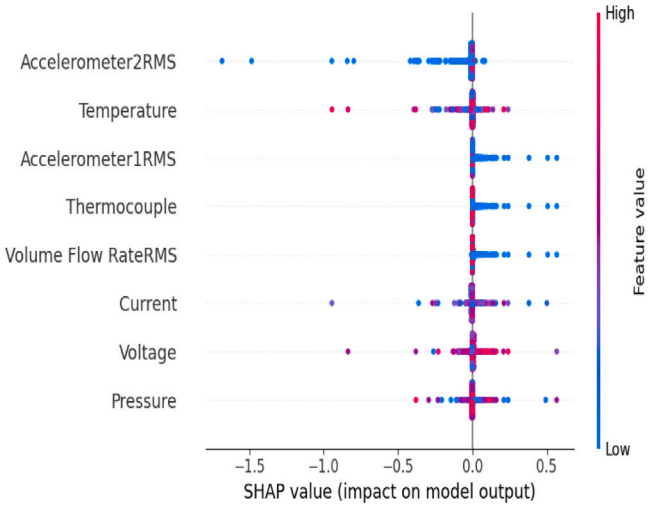
(6) BeatGAN (Zhou et al., 2019) : A model for anomaly detection in time series data, utilizing generative adversarial networks (GANs) within a framework based on reconstruction. By incorporating adversarial regularization, BeatGAN exhibits robustness in its performance.

(7) USAD (Audibert et al., 2020): A reconstruction-based method that uses an autoencoder architecture. It is constructed to learn and understand normal patterns, subsequently identifying anomalies as divergences from this established normal behavior.

(8) InterFusion (Li et al., 2021): This approach employs VAE to grasp the temporal and spatial interdependencies in multivariate time series, utilizing reconstruction discrepancies for anomaly detection.



(a) Mean SHAP Value for Each Feature



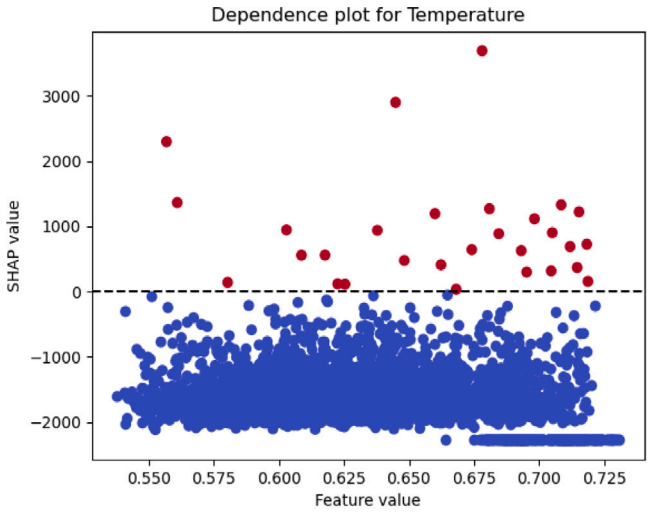
(b) SHAP Values for Individual Predictions

Fig. 8. Feature impact on model output using SHAP.

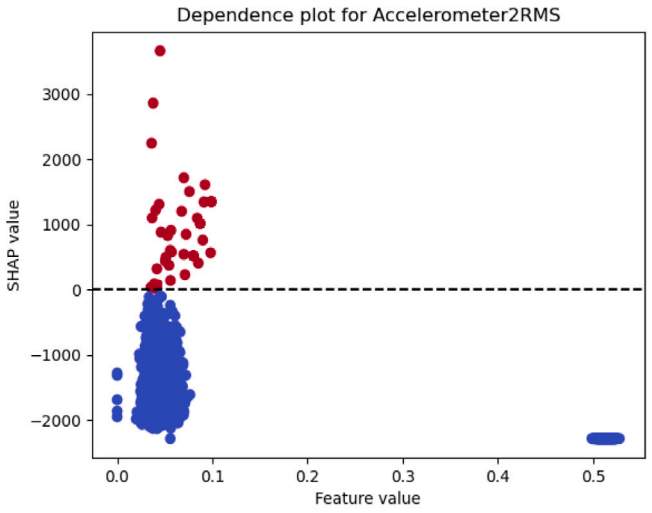
4.5. Analysis of anomaly detection results

In order to affirm the stability of our findings, we carried out five distinct tests for each algorithm. In the experiments, we evaluate our framework with eight other unsupervised anomaly detection algorithms. None of these algorithms employ SHAP values or any other XAI techniques to analyze feature importance or perform feature selection and model optimization. Our experimental results clearly indicate that the utilization of XAI techniques, specifically SHAP, significantly enhances both the interpretability and performance of our anomaly detection framework. This approach led to the creation of new composite features, further improving the model's ability to detect anomalies. The optimized framework, which was retrained with these selected features and subjected to hyperparameter tuning, demonstrated superior performance on key metrics Fc_1 and AU-ROC score. Additionally, when compared to traditional unsupervised anomaly detection models that lack XAI capabilities, our framework not only achieved higher accuracy but also provided transparent explanations for its decisions, making it a more reliable and interpretable solution for industrial applications.

Additionally, as observed in Table 2 and Fig. 4, the InterFusion algorithm demonstrates superior performance over CCTAK on the



(a) Dependence Plot for Temperature



(b) Dependence Plot for Accelerometer

Fig. 9. Dependence plot of SHAP values.

DAMADICS dataset. This performance advantage can be attributed to InterFusion's VAE architecture, which is particularly effective in detecting actuator anomaly patterns prevalent in this dataset. The focus on capturing subtle temporal patterns allows InterFusion to achieve higher accuracy in anomaly detection under these specific conditions, which may require further tuning in the CCTAK framework to achieve similar results.

Table 3, Figs. 5 and 6, illustrate the impact of the scoring function and the algorithm on model performance, using top-k as the threshold function and Fc_1 as the performance metric. These visualizations underscore the significant influence of the scoring function on the performance of the method. Dynamic scoring functions, such as Gauss-D and Gauss-D-K, demonstrate superior performance over static scoring functions like Normalized Error and Gauss-S. This superiority stems from their adaptability to both normal and abnormal testing situations, which makes them particularly effective in accurately assessing a variety of scenarios. Fig. 7 depicts the training and validation losses, which decrease with each epoch iteration, thus indicating rapid convergence of the framework.

To elucidate the influence of individual data features on the anomaly detection output, we employ the KernelExplainer function. This function measures the impact of each feature on anomaly detection within the SKAB dataset. Figs. 8 and 9 offer visual interpretations of the framework. Plot a in Fig. 8 succinctly summarizes the average impact of each feature on the framework output. The features are sorted by their mean absolute SHAPley value. The plot clearly shows that "Accelerometer2MS" is the most influential feature in contributing to the framework prediction. Plot b provides a comprehensive overview of feature importance and impacts. Every marker on the graph corresponds to a Shapley value, tied to a feature and an instance. The color signifies the feature's value, ranging from low to high. The ordering of features is determined by the aggregate magnitude of their SHAPley values across all instances.

Fig. 9 visualizes the effect of a specific feature on the prediction made by the framework. Each point on the plot represents an instance in the dataset. The color of the points represents the sign of the SHAP value, with blue indicating a negative impact and red indicating a positive impact. This plot allows us to observe how changes in the feature value influence the prediction.

5. Conclusion and future work

This study introduces a novel unsupervised anomaly detection framework for IIoT. This framework leverages convolutional blocks and CBAM to extract local features and employs an enhanced TCN and KAN-based VAE to capture the long-term dependencies of time series data. Then, the dynamic scoring function and the top-k threshold function are utilized to compute the anomaly score and threshold, respectively. Anomalies are identified by comparing these scores and thresholds. The efficacy of the framework is verified using four datasets. In addition, the XAI method is utilized to graphically depict how features affect the outcome. The data from our experiments suggest that the framework we propose excels in comparison to other unsupervised anomaly detection algorithms.

While the proposed anomaly detection framework is specifically designed for IIoT systems, its underlying principles are broadly applicable to other domains. For example, in information technology (IT) network environments, where the detection of network traffic anomalies is critical, this framework could be adapted with minimal modification to identify abnormal patterns effectively. However, extending the framework to different domains may present challenges, such as domain-specific feature extraction, the need for additional tuning of model parameters, and ensuring that the explainability remains consistent across different contexts. Future work could explore these challenges in greater depth and explore semi-supervised anomaly detection methods to enhance the effectiveness of anomaly detection with limited labeled data, offering a broader validation of the framework across multiple application areas.

CRediT authorship contribution statement

Yilixiati Abudurexiti: Writing – original draft, Methodology, Conceptualization. **Guangjie Han:** Supervision. **Fan Zhang:** Writing – review & editing. **Li Liu:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant Number B240201135, in part by the National Natural Science Foundation of China under Grant Number 62402162, in part by the China Postdoctoral Science Foundation under Grant Number 2024M750732.

References

Abdel-Basset, M., Moustafa, N., Hawash, H., 2023. Privacy-preserved generative network for trustworthy anomaly detection in smart grids: A federated semisupervised approach. IEEE Trans. Ind. Inform. 19 (1), 995–1005.

Al-Sarawi, S., Anbar, M., Abdullah, R., Al Hawari, A.B., 2020. Internet of things market analysis forecasts 2020–2030. In: 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability. WorldS4, London, UK, pp. 449–453.

Albazli, Ahmad O., Aslam, Muhammad, Dobbah, Saeed A., 2020. A control chart for exponentially distributed characteristics using modified multiple dependent state sampling. Math. Probl. Eng. 5682587. 26pages.

Alghanmi, N., Alotaibi, R., Buhari, S.M., 2022. Machine learning approaches for anomaly detection in IoT: An overview and future research directions. Wirel. Pers. Commun. 122, 2309–2324.

Audibert, J., Michiardi, P., Guyard, F., Marti, S., Zuluaga, M.A., 2020. USAD: Unsupervised anomaly detection on multivariate time series. In: Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining. pp. 3395–3404.

Bai, S., Kolter, J.Z., Koltun, V., 2018. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv:1803.01271.

Chen, R.-Q., Shi, G.-H., Zhao, W.-L., Liang, C.-H., 2021. A joint model for IT operation series prediction and anomaly detection. Neurocomputing 448, 130–139.

Cook, A.A., Mısırlı, G., Fan, Z., 2020. Anomaly detection for IoT time series data: A survey. IEEE Internet Things J. 7 (7), 6481–6494.

2020. Damadics benchmark website. [Online]. Available: http://diag.mchtr.pw.edu.pl/damadics/.

Das, T., Shukla, R.M., Sengupta, S., 2022. What could possibly go wrong?: Identification of current challenges and prospective opportunities for anomaly detection in internet of things. IEEE Netw..

Fahim, M., Sillitti, A., 2019. Anomaly detection, analysis and prediction techniques in IoT environment: A systematic literature review. IEEE Access 7, 81664–81681.

Garg, A., Zhang, W., Samaran, J., Savitha, R., Foo, C.-S., 2022. An evaluation of anomaly detection and diagnosis in multivariate time series data. IEEE Trans. Neural Netw. Learn. Syst. 33 (6), 2508–2517.

Goh, J., Adepu, S., Junejo, K.N., Mathur, A., 2016. A dataset to support research in the design of secure water treatment systems. In: Proc. Int. Conf. Crit. Inf. Infrastruct. Secur. Springer, Cham, Switzerland, pp. 88–99.

Hundman, K., Constantinou, V., Laporte, C., Colwell, I., Soderstrom, T., 2018. Detecting spacecraft anomalies using LSTMs and nonparametric dynamic thresholding. In: Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining. pp. 387–395.

Kadri, Farid, Harrou, Fouzi, Chaabane, Sondès, Sun, Ying, Tahon, Christian, 2016. Seasonal ARMA-based SPC charts for anomaly detection: Application to emergency department systems. Neurocomputing 173 (Part 3), 2102–2114.

Karevan, Zahra, Suykens, Johan A.K., 2020. Transductive LSTM for time series data prediction: An application to weather forecasting. Neural Netw. 125, 1–9.

Karmakar, A., Dey, N., Baral, T., Chowdhury, M., Rehan, M., 2019. Industrial internet of things: A review. In: 2019 International Conference on Opto-Electronics and Applied Optics. Optronix, Kolkata, India, pp. 1–6.

Katser, I.D., Kozitsin, V.O., 2020. Skoltech anomaly benchmark (skab). [Online]. Available: https://www.kaggle.com/dsv/1693952.

Khan, I.A., Moustafa, N., Pi, D., Sallam, K.M., Zomaya, A.Y., Li, B., 2022. A new explainable deep learning framework for cyber threat discovery in industrial IoT networks. IEEE Internet Things J. 9 (13), 11604–11613.

Kopp, Martin, Pevný, Tomáš, Holeňa, Martin, 2020. Anomaly explanation with random forests. Expert Syst. Appl. 149, 113187.

Kumaran Santhosh, K., Dogra, D.P., Roy, P.P., Mitra, A., 2022. Vehicular trajectory classification and traffic anomaly detection in videos using a hybrid CNN-VAE architecture. IEEE Trans. Intell. Transp. Syst. 23 (8), 11891–11902.

Li, D., Chen, D., Jin, B., Shi, L., Goh, J., Ng, S.-K., 2019. MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks. In: Proc. Int. Conf. Artif. Neural Netw. Springer, Cham, Switzerland, pp. 703–716.

Li, Zhihan, Zhao, Youjian, Han, Jiaqi, Su, Ya, Jiao, Rui, Wen, Xidao, Pei, Dan, Multivariate time series anomaly detection and interpretation using hierarchical inter-metric and temporal embedding. In: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, KDD'2021. pp. 3220–3230, 2021.

Liu, Ziming, Wang, Yixuan, Vaidya, Sachin, Ruehle, Fabian, Halverson, James, Soljačić, Marin, Hou, Thomas Y., Tegmark, Max, 2024. KAN: Kolmogorov–Arnold networks. arXiv preprint arXiv:2404.19756.

Liu, Y., et al., 2021. Deep anomaly detection for time series data in industrial IoT: A communication-efficient on-device federated learning approach. IEEE Internet Things J. 8 (8), 6348–6358.

Ma, X., Keung, J., He, P., Xiao, Y., Yu, X., Li, Y., 2024. A semisupervised approach for industrial anomaly detection via self-adaptive clustering. IEEE Trans. Ind. Inform. 20 (2), 1687–1697.

Nizam, H., Zafar, S., Lv, Z., Wang, F., Hu, X., 2022. Real-time deep anomaly detection framework for multivariate time series data in industrial IoT. IEEE Sens. J. 22 (23), 22836–22849.

Pang, Guansong, Shen, Chunhua, Cao, Longbing, Van Den Hengel, Anton, 2022. Deep learning for anomaly detection: A review. ACM Comput. Surv. 54 (2), 1–38, Research Collection School Of Computing and Information Systems.

Puri, C., Kooijman, G., Vanrumste, B., Luca, S., 2022. Forecasting time series data in Healthcare with Gaussian processes and dynamic time warping based subset selection. IEEE J. Biomed. Health Inf. 26 (12), 6126–6137.

Sezer, Omer Berat, Gudelek, Mehmet Ugur, Ozbayoglu, Ahmet Murat, 2020. Financial time series data forecasting with deep learning : A systematic literature review: 2005–2019. Appl. Soft Comput. 90, 106181.

Shtayat, M.M., Hasan, M.K., Sulaiman, R., Islam, S., Khan, A.U.R., 2023. An explainable ensemble deep learning approach for intrusion detection in industrial internet of things. In: IEEE Access. Vol. 11, pp. 115047–115061.

Su, Y., Zhao, Y., Niu, C., Liu, R., Sun, W., Pei, D., 2019. Robust anomaly detection for multivariate time series through stochastic recurrent neural network. In: Proc. 25th ACMSIGKDD Int. Conf. Knowl. Discovery Data Mining. pp. 2828–2837.

Tsvetanov, Tsvetan, Slaria, Srishti, 2021. The effect of the Colonial Pipeline shutdown on gasoline prices. Econom. Lett. 209, 110122.

Vos, Kilian, Peng, Zhongxiao, Jenkins, Christopher, Shahriar, Md Rifat, Borghesani, Pietro, Wang, Wenyi, 2022. Vibration-based anomaly detection using LSTM/SVM approaches. Mech. Syst. Signal Process. 169, 108752.

Wen, Qingsong, Gao, Jingkun, Song, Xiaomin, Sun, Liang, Xu, Huan, Zhu, Shenghuo, 2019. RobustSTL: A Robust Seasonal-Trend Decomposition Algorithm for Long Time Series Data. AAAI.

Woo, S., Park, J., Lee, J.Y., Kweon, I.S., 2018. CBAM: Convolutional block attention module. In: Proceedings of the European Conference on Computer Vision. ECCV, pp. 3–19.

Wu, D., Jiang, Z., Xie, X., Wei, X., Yu, W., Li, R., 2020. LSTM learning with Bayesian and Gaussian processing for anomaly detection in industrial IoT. IEEE Trans. Ind. Inform. 16 (8), 5244–5253.

Xu, Jiehui, Wu, Haixu, Wang, Jianmin, Long, Mingsheng, 2022. Anomaly Transformer: Time Series Data Anomaly Detection with Association Discrepancy. ICLR.

Xu, H., et al., 2018. Unsupervised anomaly detection via variational auto-encoder for seasonal KPIs in web applications. In: Proceedings of the World Wide Web Conference. pp. 187–196.

Zhang, Y., Chen, Y., Wang, J., Pan, Z., 2023. Unsupervised deep anomaly detection for multi-sensor time series data signals. IEEE Trans. Knowl. Data Eng. 35 (2), 2118–2132.

Zhang, F., Kodituwakku, H.A.D.E., Hines, J.W., Coble, J., 2019a. Multilayer data-driven cyber-attack detection system for industrial con- trol systems based on network, system, and process data. IEEE Trans. Ind. Inform. 15 (7), 4362–4369.

Zhang, C., et al., 2019b. A deep neural network for unsupervised anomaly detection and diagnosis in multivariate time series data. Proc. AAAI Conf. Artif. Intell. 33, 1409–1416.

Zheng, P., Yuan, S., Wu, X., Li, J., Lu, A., 2019. One-class adversarial nets for fraud detection. In: Proc. AAAI Conf. Artif. Intell. Vol. 33, pp. 1286–1293.

Zhou, X., Hu, Y., Liang, W., Ma, J., Jin, Q., 2021. Variational LSTM enhanced anomaly detection for industrial big data. IEEE Trans. Ind. Inform. 17 (5), 3469–3477.

Zhou, S., Huang, X., Liu, N., Zhou, H., Chung, F.-L., Huang, L.-K., 2023. Improving generalizability of graph anomaly detection models via data augmentation. IEEE Trans. Knowl. Data Eng. 35 (12), 12721–12735.

Zhou, B., Liu, S., Hooi, B., Cheng, X., Ye, J., 2019. BeatGAN: Anomalous rhythm detection using adversarially generated time series data. In: Proceedings of the 28th International Joint Conference on Artificial Intelligence Menlo Park. AAAI Press, CA, USA, pp. 4433–4439.

Zolanvari, M., Teixeira, M.A., Gupta, L., Khan, K.M., Jain, R., 2019. Machine learning-based network vulnerability analysis of industrial internet of things. IEEE Internet Things J. 6 (4), 6822–6834.

Yilixiati Abudurexiti received his M.S. degree in Computer Applications from Xinjiang University, Urumqi, China, in 2012. He is currently pursuing the Ph.D. degree with the College of Information Science and Engineering, Hohai University, Changzhou, China. He is currently a lecturer at the College of Information Engineering, Urumqi Vocational University, China. His main research interests include machine learning based anomaly detection for industrial IoT and intelligent health monitoring for industrial equipment, as well as radiation source identification.

Guangjie Han (Fellow, IEEE) is currently a Professor with the Department of Information Science and Engineering, Hohai University, Changzhou, China. He received his Ph.D. degree from Northeastern University, Shenyang, China, in 2004. In February 2008, he finished his work as a Postdoctoral Researcher with the Department of Computer Science, Chonnam National University, Gwangju, Korea. From October 2010 to October 2011, he was a Visiting Research Scholar with Osaka University, Suita, Japan. From January 2017 to February 2017, he was a Visiting Professor with City University of Hong Kong, China. From July 2017 to July 2020, he was a Distinguished Professor with Dalian University of Technology, China. His current research interests include Internet of Things, Industrial Internet, Machine Learning and Artificial Intelligence, Mobile Computing, Security and Privacy. Dr. Han has over 500 peer-reviewed journal and conference papers, in addition to 160 granted and pending patents. Currently, his H-index is 63 and i10-index is 265 in Google Citation (Google Scholar). The total citation count of his papers raises above 14500+ times. Dr. Han is a Fellow of the UK Institution of Engineering and Technology (FIET). He has served on the Editorial Boards of up to 10 international journals, including the IEEE TII, IEEE TCCN, IEEE Systems, etc. He has guest-edited several special issues in IEEE Journals and Magazines, including the IEEE JSAC, IEEE Communications, IEEE Wireless Communications, Computer Networks, etc. Dr. Han has also served as chair of organizing and technical committees in many international conferences. He has been awarded 2020 IEEE Systems Journal Annual Best Paper Award and the 2017–2019 IEEE ACCESS Outstanding Associate Editor Award.

Fan Zhang received the B.S. degree in Internet of Things engineering from Hohai University, Changzhou, China, in 2019, where she is currently pursuing the Ph.D. degree at the College of Information Science and Engineering. Her current research interests include mobile edge computing and machine learning.

Li Liu received the Ph.D. degree in Internet of Things Technology from Hohai University, Nanjing, China, in 2019. He was a Senior Research Assistant in Department of Electrical Engineering in City University of Hong Kong from October 2017 to February 2018. He is currently a Lecturer with the School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi, China. He has published over 20 papers in related international conferences and journals, including IEEE TMC, IEEE TII, IEEE IOTJ, IEEE Communications Magazine, etc. His research interests include wireless sensor networks, industrial internet of things, and machine learning.