energies

MDPI

Article

Efficient One-Class False Data Detector Based on Deep SVDD for Smart Grids

Hany Habbak [1,†], Mohamed Mahmoud [2,*,†], Mostafa M. Fouda [3,4,*,†], Maazen Alsabaan [5,†], Ahmed Mattar [1], Gouda I. Salama [1,†] and Khaled Metwally [1,†]

[1] Department of Computer Engineering and AI, Military Technical College, Cairo 11766, Egypt; c-helshall@tntech.edu (H.H.); a.mattar@ieee.org (A.M.); gisalama@mtc.edu.eg (G.I.S.); k.metwally@mtc.edu.eg (K.M.)
[2] Department of Electrical and Computer Engineering, Tennessee Technological University, Cookeville, TN 38505, USA
[3] Department of Electrical and Computer Engineering, College of Science and Engineering, Idaho State University, Pocatello, ID 83209, USA
[4] Center for Advanced Energy Studies (CAES), Idaho Falls, ID 83401, USA
[5] Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11451, Saudi Arabia; malsabaan@ksu.edu.sa
* Correspondence: mmahmoud@tntech.edu (M.M.); mfouda@ieee.org (M.M.F.)
† These authors contributed equally to this work.

Abstract: In the smart grid, malicious consumers can hack their smart meters to report false power consumption readings to steal electricity. Developing a machine-learning based detector for identifying these readings is a challenge due to the unavailability of malicious datasets. Most of the existing works in the literature assume attacks to compute malicious data. These detectors are trained to identify these attacks, but they cannot identify new attacks, which creates a vulnerability. Very few papers in the literature tried to address this problem by investigating anomaly detectors trained solely on benign data, but they suffer from these limitations: (1) low detection accuracy and high false alarm; (2) the need for knowledge on the malicious data to compute good detection thresholds; and (3) they cannot capture the temporal correlations of the readings and do not address the class overlapping issue caused by some deceptive attacks. To address these limitations, this paper presents a deep support vector data description (*DSVDD*) based unsupervised detector for false data in smart grid. Time-series readings are transformed into images, and the detector is exclusively trained on benign images. Our experimental results demonstrate the superior performance of our detectors compared to existing approaches in the literature. Specifically, our proposed *DSVDD*-based schemes have exhibited improvements of 0.5% to 3% in terms of recall and 3% to 9% in terms of the Area Under the Curve (*AUC*) when compared to existing state-of-the-art detectors.

Keywords: false data detection; electricity theft; smart meters; automatic metering infrastructure; smart power grid; deep-SVDD; one-class classification

Citation: Habbak, H.; Mahmoud, M.; Fouda, M.M.; Alsabaan, M.; Mattar, A.; Salama, G.I.; Metwally, K. Efficient One-Class False Data Detector Based on Deep SVDD for Smart Grids. *Energies* **2023**, *16*, 7069. https://doi.org/10.3390/en16207069

Academic Editors: Zoya Pourmirza, Mustafa A. Mustafa and Roberto Metere

Received: 15 September 2023
Revised: 4 October 2023
Accepted: 9 October 2023
Published: 12 October 2023

1. Introduction

The concept of smart grids (SGs) aims to revolutionize the energy sector by facilitating intelligent monitoring, control, and management of power generation and distribution. With the advancement of smart metering technologies, an enormous amount of energy consumption data is being generated and collected at unprecedented scales. These data can be used for smart management of energy generation, load forecasting and billing using dynamic tariffs as well. However, this vast volume of data has also brought significant challenges, particularly in the detection and prevention of electricity theft, a growing concern for energy providers worldwide [1].

The idea is that malicious consumers may tamper with their smart meters (SMs) to report false power consumption readings to reduce their electrical bills. The false data not

only leads to substantial financial losses for the electrical utility companies, but they also pose a significant threat to the performance and stability of the *SGs* because they can cause suboptimal energy management decisions. Traditional detection methods have struggled to keep pace with the sophisticated techniques employed by perpetrators, necessitating more advanced and intelligent solutions [2,3]. The current advancement in artificial intelligence (*AI*) and big data technologies and the possibility of seamless integration of them have opened up new avenues for remarkable advancements in electricity theft detection (*ETD*) within the *SG* environment. The marriage of *AI* techniques with big data technologies enables the efficient processing, storage, and analysis of massive volumes of energy-consumption data [4]. This powerful synergy empowers utilities with real-time insights to combat theft with unparalleled effectiveness and precision [5].

Smart meters provide an abundance of energy consumption data, encouraging researchers to introduce machine learning (*ML*) models for the detection of electricity theft [6,7]. These *ML*-based detectors encompass both supervised classifiers and anomaly detectors, aiming to accurately identify instances of suspicious electricity usage patterns [8]. Supervised classifiers utilize both benign and malicious energy consumption profiles of customers during training. However, they have a limitation, heavily relying on the availability of both benign and malicious samples in customers' energy consumption data for effective training. This condition poses challenges, particularly in case of new attacks. Consequently, in such cases, supervised classifiers are not practical choices, as their detection capability is confined to the attacks they have been trained on (seen attacks) only [9], and their detection accuracy significantly degrades when new attacks are launched.

In contrast, unsupervised (or anomaly) detectors exclusively use benign data for training to learn legitimate customer consumption patterns [10]. During testing, these detectors identify false power consumption readings by detecting deviations from the learned benign power consumption patterns. The key advantage of unsupervised detectors over the supervised ones lies in their ability to detect new attacks because unlike supervised learning detectors that are trained on the false data of specific attacks, the unsupervised learning detectors are trained only on benign data and thus any deviations from the benign patterns are indicators for false data [11,12]. However, most of the existing unsupervised anomaly detectors suffer from low detection accuracy and high false alarm. Specifically, they suffer from several limitations that result in low detection accuracy. These include: (1) shallow-architecture based unsupervised detectors struggle with capturing temporal correlations of the power consumption readings; (2) other detectors need knowledge on the malicious data to compute detection threshold that can achieve high detection accuracy; and (3) unsupervised detectors struggle in addressing the class overlapping issue posed by some deceptive attacks. To address these limitations, this paper presents a deep support vector data description (*DSVDD*) based unsupervised detector for false data in smart grid.

Initially, time-series *SM* readings are transformed into images (RGB and grayscale). The model is then exclusively trained on benign images. *DSVDD* has emerged as a promising one-class deep learning (*DL*) approach [13]. It trains a deep neural network while also optimizing a hypersphere encapsulating the data in the output space. This unique approach allows *DSVDD* to extract shared factors of variation from the benign data, ultimately utilizing these benign samples to construct a hypersphere feature space that serves as an effective detection threshold. Because *DSVDD* leverages *DL*, it can learn the inherent characteristics of legitimate energy consumption patterns and accurately identifies deviations caused by fraudulent activities.

Additionally, we investigate how to address the challenge of overlapping data and empower electrical utilities to combat electricity theft with greater precision, efficiency, and scalability. To achieve this, we employ *DSVDD*, which effectively reduces the span of benign data within the feature space, addressing the issue of class overlapping. We also utilize the likelihood metric to fit the probability distribution of benign data within the feature space. By addressing the issue of the overlap of benign and malicious data, caused by deceptive attacks, and developing an accurate detector that is not trained on specific

attacks, our goal is to contribute to the development of robust and secure *SG* systems capable of effectively mitigating the adverse impacts of electricity theft.

The primary contributions of this paper can be outlined as follows:

- We create and propose utilizing an image-based dataset for our detector. To achieve this, we leverage the Irish Smart Energy Trial (*ISET*) dataset [14], which was initially captured in a time series format and stored as a .csv file. By converting the *SM* readings into image representations, both RGB and grayscale, we aim to enhance the effectiveness of electricity theft detection. Additionally, we can take advantage of the significant advancements in the field of computer vision and *DL*, which have demonstrated remarkable progress when applied to image-based datasets.
- Given the availability of only benign data for power consumption readings, our objective is to build a robust one-class classifier that can accurately classify data, even in the presence of overlapping classes. To achieve this, we first investigate a One Class *DSVDD* (*OC-DSVDD*)-based scheme to reduce the span of benign data within the feature space, addressing the issue of class overlapping. Then, we investigate using *DSVDD* with one-class support vector machine (*OC-SVM*) to leverage the decision boundary created by *OC-SVM* for detecting malicious data and providing a decision. Concurrently, we utilize the reduced span characteristics generated by *DSVDD* to address the issue of class overlap. Finally, we investigate using *DSVDD* with Gaussian Mixture Model (*GMM*) to fit the probability distribution after reducing the span of benign data within the feature space using the likelihood metric. The likelihood, derived from *GMM*, serves as an anomaly score, allowing for the identification of malicious samples as they exhibit lower likelihood values.
- Extensive experiments have been conducted to evaluate our detector and compare it to the state-of-the-art detectors. The results demonstrate that the performance of our *DSVDD* with *GMM* is superior in terms of accuracy, precision, recall, F1 score, and Area Under the Curve *(AUC)*.

The subsequent sections of this article are organized as follows. Section 2 offers an inclusive overview of various *ETD* techniques, encompassing both supervised and unsupervised models. Section 3 delves into data preparation, including energy consumption profiles, along with the conversion of the dataset from its time series format into RGB and grayscale images. Moving on, Section 4 furnishes essential preliminaries, including key concepts such as autoencoders, one-class classification, *OC-SVM*, *SVDD*, *DSVDD*, and *GMM*. In Section 5, the architecture of the proposed *DSVDD*-based schemes is presented. Subsequently, Section 6 showcases the results of the conducted experiments. Lastly, Section 7 offers a conclusive summary to round off this article.

2. Related Work

Due to the abundance of energy consumption data provided by *SMs*, researchers have been leveraging *ML* to detect instances of electricity theft. The vast majority of the existing works focused on supervised classifiers. Due to the unavailability of malicious data needed to train the classifiers, some electricity theft attacks are assumed. The classifiers are trained to identify these attacks and they may fail to identify new attacks. Very few works have investigated unsupervised classifiers but these works either do not achieve good detection accuracy or they need knowledge on the attacks to tune the model to improve it accuracy. In this section, we initiate our survey with supervised classifiers, including both shallow and deep learning-based models. Subsequently, we explore the few works that investigated unsupervised classifiers. Finally, we conclude by providing an overview of the identified limitations and existing research gaps within the literature survey.

2.1. Supervised False Data Detector

Murthy et al. [15] conducted a study using data mining techniques to investigate non-technical losses in the power distribution system. Their model consists of two stages; the first stage employs Fuzzy C-Means clustering to group consumers with similar con-

sumption profiles, and the second stage applies a finely-tuned Naïve Bayes classification technique to identify potential fraudsters. Data mining based approaches use simple techniques that usually do not lead to accurate classification comparing to *ML*-based approaches that can make more accurate decisions because they learn the consumption patterns of benign and malicious data.

For shallow classifiers, P. Jokar et al. [16] presented a consumption pattern-based energy theft detector *(CPBETD)*. The proposed scheme utilizes the foreseeability of normal and malicious customer behaviors to detect energy theft. It combines a multiclass SVM, silhouette plots for identifying different dataset distributions, and distribution transformer meters to detect non-technical losses at the transformer level. R. Wu et al. [17] propose a combined approach using Adaptive Boosting algorithm (AdaBoost) and SVM to find anomalous consumers in an imbalanced dataset. They also employ General Regression Neural Network (GRNN) to estimate electricity theft intervals for abnormal consumers. The method outperforms conventional approaches and achieves better performance on imbalanced datasets.

A hybrid convolutional neural network-random forest (CNN-RF) model for automatic energy theft detection is introduced in [18,19]. The CNN learns features from *SM* data, and the RF classifies the data using these features. The given results show that CNN-RF model outperforms benchmark models, including SVM, RF, gradient boosting decision tree (GBDT), and logistic regression (LR), when trained using the same datasets.

A top-down classifier utilizing decision trees (DT) and SVM to detect electricity theft by malicious consumers is presented in [20]. It operates effectively across the power network, considering various features and achieving improved accuracy when SVM is combined with DT. Yan et al. [21] propose an electricity theft detector based on XGBoost. The detector outperforms various *ML*-based models, including the SVM, the backpropagation neural network, extreme learning machine (ELM), Deep ELM, the k-nearest Neighbors (KNN) algorithm, LR, DT, RF, the Naive Bayes classifier, and AdaBoost, achieving good performance even with an imbalanced training set. A comprehensive summary of all supervised shallow classifiers for electricity detection is presented in Table 1.

For *DL*-based classifiers, Nabil et al. [22] introduced *DL*-based detectors to combat Cyberattacks targeting electricity theft in AMI networks within SGs. They developed both consumer-specific and generalized detectors utilizing deep feed-forward (FF) and recurrent neural networks (RNN). In consumer-specific classifiers, one *DL* model is trained and used for each customer while a generalized detector is trained on the data of all consumers and can be used for all of them. The experimental results indicate that *DL*-based detectors outperform shallow *ML* approaches. Additionally, the work in [23] develops an RNN-based detector leveraging consumers' electricity consumption time series and employing a gated recurrent unit (GRU) for enhanced detection performance with effective hyperparameter fine-tuning.

A deep recurrent vector embedding model is proposed by Takiddin et al. [24] to detect electricity theft cyber-attacks, using vector embedding to represent energy consumption profiles as real-numbered vectors and capturing patterns in reported readings. Meanwhile, Ismail et al. [25] examine electricity theft detection in renewable energy-based distributed generation units, employing a hybrid *C-RNN DL* architecture. Various cyber-attack functions are introduced to manipulate benign power readings to compute malicious dataset. The experimental results demonstrate that the *C-RNN* model exhibits superior detection performance compared to other *DL*-based models.

The proposed energy detection scheme in [26] combines CNN and LSTM. It utilizes a new algorithm for preparing data prior to handle missing instances and addresses class imbalance through synthetic data generation. The experimental results show good accuracy in classifying both normal and malicious data. Furthermore, in [27], a Hybrid Deep Neural Network (HDNN) that amalgamates CNN, GRU, and particle swarm optimization (PSO) for electricity theft detection is introduced. The preprocessing phase refines the data, the CNN contributes to dimensionality reduction, and the GRU-PSO

mechanism distinguishes between benign and malicious power consumption readings. When benchmarked against techniques such as LR, SVM, LSTM, and GRU, the proposed HDNN exhibits superior performance in energy theft detection, effectively addressing class imbalance concerns.

Table 1. Summary of Data Mining and Shallow Classifier for Electricity Theft Detection.

Technique	Ref.	Description	Dataset
Naïve Bayes	[15]	data mining techniques used for *ETD*, employing Fuzzy c-Means clustering to group end users and Naïve Bayes classification to identify potential fraudsters.	APSPDCL
multiclass SVM	[16]	*CPBETD* employs M-SVM, silhouette plots, and distribution transformer meters for NTL detection.	ISET
AdaBoost-SVM	[17]	An ensemble strategy utilizing AdaBoost and SVM detects anomalies in imbalanced user data. GRNN is applied to estimate electricity theft intervals for unusual consumers.	Tangshan City Dataset
CNN-RF	[18,19]	Introducing the CNN-RF model, where CNN captures *SM* data features, and RF detects the theft.	SEAI & LCL
DT-SVM	[20]	An efficient two-step scheme using DT and SVM classifiers detects intentional electricity theft by malicious users across the power network.	OpenEI
XGBoost	[21]	An *ETD* scheme based on XGBoost for AMI.	ISET

In [28], a hybrid *DL* model is presented for electricity theft detection. The model combines both GRU and CNN to distinguish between benign and malicious electricity consumption patterns. The GRU layers extract temporal patterns, while the CNN retrieves optimal abstract patterns from the dataset. Another hybrid improved wide and deep CNN method, proposed in [29], addresses electricty theft detection using an imbalanced real dataset. The method introduces a channel dimensional adaptive attention module along with dilated convolutions, and utilizes focal loss to tackle the data imbalance problem.

Emadaleslami et al. [30] propose a two-stage *DL* model for detection of electricity theft in AMI. In the first stage, they develop several CNN models based on the predictability of normal and malicious patterns. These models predict theft patterns using the available load profile of fraudulent customers, addressing the data shortages of malicious customers. Unlike previous methods that used synthetic minority over-sampling (SMOTE) analysis to handle data imbalance, the CNN models efficiently identify a wide range of theft patterns for all customers. In the second stage, a Deep Neural Network (DNN) model is utilized to distinguish between normal and malicious customers. It presents an improved methodology for implementing electricity theft detection using CNN model. The implementation employs two different approaches for data processing, aiming to determine the most suitable approach for the proposed model. Table 2 offers a summary of supervised *DL* classifiers utilized for electricity theft detection.

Table 2. Summary of Supervised DL-based Classifier for Electricity Theft Detection.

Technique	Ref.	Description	Dataset
FF-RNN	[22]	New *DL* detectors combat electricity theft cyber-attacks in AMI networks. Customized and generalized, using deep FF-RNN.	ISET
GRU-RNN	[23]	RNN-based detector uses customer consumption data with GRU-RNN for improved detection, and fine-tuning hyperparameters.	ISET
Vector Embedding	[24]	A deep model is proposed for *ETD*, utilizing real-numbered vectors for representing consumption profiles and capturing reading patterns.	SGCC & ISET
C-RNN	[25]	Investigating theft detection in renewable energy-based units using *C-RNN DL* with introduced cyber-attack functions for power reading manipulation.	Created Realistic Synthetic Data
CNN-LSTM	[26]	Integrating CNN and LSTM for Effective *ETD*. Innovative Data Pre-processing and Synthetic Data Tackling Class Imbalance.	SGCC
CNN-GRU-PSO HDNN	[27]	Introducing CNN-GRU-PSO HDNN for *ETD*. Including data Pre-processing, dimensionality reduction with CNN, and honesty-fraud classification via GRU-PSO.	SGCC
GRU-CNN	[28]	The model integrates GRU and CNN components to differentiate between benign and malicious electricity consumption patterns. GRU layers capture temporal patterns, while CNN identifies optimal abstract patterns from the dataset.	SGCC
Wide & Deep CNN	[29]	Hybrid improved wide and deep CNN for imbalanced *ETD*. Adaptive attention, dilated convolutions, and focal loss for effective results.	SGCC
CNN-DNN	[30]	Two-stage *DL* model detects energy fraud in AMI. Stage 1: CNN predicts theft patterns without SMOTE. Stage 2: DNN distinguishes normal from suspicious customers.	ISET

2.2. Unsupervised False Data Detector

The work in [16] trains supervised and unsupervised electricity theft detectors and compares between them. For the unsupervised detector, benign power consumption readings are used to train an *OC-SVM* classifier. The experimental results show that supervised classifiers significantly outperform the unsupervised one. Another study done by Krishna et al. [31] shows that cross-validation techniques confirm the effectiveness of first-order differentiation, rendering the data weakly stationary. This finding supports the use of the ARIMA model as a better choice for capturing consumption behavior and forecasting future behaviors.

In [32,33] a Principal Component Analysis (PCA) based electricity theft detection approach is proposed. It uses PCA to transform a high-dimensional dataset into a lower-dimensional one and then calculate an anomaly score to compare with a predefined threshold value. The scheme is tested with real data under different attack scenarios. Yeckle et al. [34] used outliers detection technique to detect electricity theft in AMI. Preprocessing with k-means clustering reduced measurement samples, improving *AUC* performance. Influenced Outlierness (INFLO) and Relative Density-based Outlier Score (RDOS) are used to classify data.

A combined unsupervised learning approach is presented for electricity theft detection and loss estimation in [35]. It uses three anomaly measurement indexes (mean, fluctuation,

and trend) to detect various anomalies. Through the application of two unsupervised learning techniques (sample-to-subsamples decomposition algorithm and clustering algorithm), typical ranges of index values are derived from historical electricity consumption data, facilitating the identification of malicious power consumption samples. The combination of these anomaly measurement indices enables the identification of electricity thieves based on the majority voting.

Takiddin et al. [36–40] utilized multiple deep autoencoder anomaly detectors to detect electricity theft. The results indicate that deep architectures outperform shallow detectors in terms of detection performance and the recurrent LSTM-based architectures could further enhance the detection performance compared to static fully connected detectors. This work assumes the existence of malicious dataset and uses it to optimize the detection threshold of the proposed model. The summary of the unsupervised classifiers for electricity theft detection can be found in Table 3.

Table 3. Summary of Unsupervised Classifier for Electricity Theft Detection.

Technique	Ref.	Description	Dataset
OC-SVM	[16]	Using *OC-SVM* for *ETD* in AMI systems. Relies on benign data to identify anomalies as potential theft.	ISET
ARIMA	[31]	Using ARIMA model to predict consumption behavior and evaluate its performance against an attack involving modified *SM* readings for electricity theft.	ISET
PCA	[32,33]	PCA-based scheme for *ETD*. Transforms data, calculates anomaly scores against threshold. Tested on real data with diverse attacks.	ISET
Outliers Detection	[34]	*AMI-ETD* using outlier detection algorithms. K-means preprocessing improves *AUC*. INFLO and RDOS notably effective in theft detection.	ISET
Mean Index Fluctuation Index Trend Index	[35]	A combined unsupervised approach for *ETD* using anomaly indexes. Leveraging sample-to-subsamples decomposition and clustering, historical data defines index ranges for fraudulent load identification. Detects thieves through predominant fraudulent load samples.	ISET
Multiple Autoencoders	[36–40]	Several autoencoders are Utilized to detect electricity theft. Showed deep architectures outperform shallower ones, especially LSTM-based models.	ISET

2.3. Limitations and Research Gaps

Supervised classifiers face a major limitation as they require the availability of both benign and malicious customers' energy consumption samples for training. In power consumption readings, only benign data is publicly available, and no malicious data is accessible. Because of this limitation, in the literature, a set of electricity theft attacks are introduced and used to compute malicious data. Because the supervised classifiers are trained only on these attacks, they may not be able to detect new attacks [41,42]. Consequently, supervised classifiers are a practical choice only when the malicious data is known and they may fail in case of launching new attacks, and thus, the use of unsupervised classifiers trained only on benign data becomes necessary.

The existing unsupervised detectors suffer from several limitations that result in low detection accuracy of the false power consumption readings. These include: (1) shallow-architecture based unsupervised detectors struggle with capturing temporal correlations of the power consumption readings; (2) other detectors need knowledge on the malicious data to compute detection threshold that can achieve high detection accuracy; and (3) the existing unsupervised detectors struggle in addressing the class overlapping issue posed by some

deceptive attacks. Consequently, there is a pressing need for unsupervised detectors that exclusively rely on benign energy consumption profiles, providing significantly enhanced detection performance and effectively overcoming these serious limitations.

3. Data Preparation

In this section, we present the electricity consumption data that has been employed for the training and testing of the investigated detectors. Anomaly detectors are exclusively trained on benign dataset, and subsequently tested on both benign and malicious datasets, while supervised classifiers are trained and tested using both benign and malicious datasets. The dataset utilized in our study comprises real electricity consumption samples obtained from the ISET [14]. In addition to this benign data, we have generated malicious samples using six different attack functions introduced in [16,43] and widely used in the literature. The malicious dataset is used only for evaluation purpose and it is not needed at all to compute the detector.

The ISET dataset encompasses data obtained from *SMs* installed in approximately 3600 residential units. These meters record electricity consumption every 30 min over a span of 536 days. This results in approximately 25,728 readings per customer, providing an ample amount of data for training and testing our electricity theft detectors. Figure 1a illustrates a sample electricity consumption pattern throughout a day for a benign customer sourced from the ISET dataset.

3.1. Generating Malicious Data

Dishonest users manipulate their meters to report inaccurate power consumption readings in an effort to reduce their electricity expenses. The malicious customers employ diverse attack methods to manipulate the integrity of consumption readings, with the intention of reducing their electricity bills by causing a discrepancy between the reported consumption $K_c(d,t)$ and the actual consumption $E_c(d,t)$, where, $E_c(d,t)$ denote the electricity consumption value for customer c on a particular day d and time t. These values collectively form the entries of matrix E_c. In the case of an honest customer, the reported energy consumption by their *SMs*, $K_c(d,t)$, adheres to the condition $K_c(d,t) = E_c(d,t)$. Consequently, matrices E_c and K_c are identical. In order to construct the malicious dataset, we utilize the electricity theft attacks outlined in [16,36]. The attacks we adopted to compute the malicious data can be classified intopartial reduction, selective by-pass, and price-based load control attacks.

In the partial reduction attacks, an attack function denoted as $g_1(E_c(d,t))$ reduces the actual electricity consumption reading by a fixed random fraction, $\delta = \text{rand}(0.1, 0.8)$. This captures both low-level and high-level attacks across all samples. As a consequence, the reported electricity consumption reading $K_c(d,t)$ is altered accordingly as shown in Equation (1).

$$g_1(E_c(d,t)) = \delta E_c(d,t) \tag{1}$$

The attack function $g_2(E_c(d,t))$ applies a dynamic random fraction, $\Delta(d,t) = \text{rand}(0.1, 0.8)$, to multiply each reading of the electricity consumption data as follows:

$$g_2(E_c(d,t)) = \Delta(d,t) E_c(d,t) \tag{2}$$

In the second category of attacks, known as selective by-pass attacks, the reported energy consumption reading is set to zero during a specific time interval $[t_i(d), t_f(d)]$. Outside of this interval, the reported electricity consumption readings reflect the actual consumption levels.

$$g_3(E_c(d,t)) = \begin{cases} 0 & \forall t \in [t_i(d), t_f(d)] \\ E_c(d,t) & \forall t \notin [t_i(d), t_f(d)] \end{cases} \tag{3}$$

The interval is determined by an initial time, $t_i(d)$, which is randomly selected from the range of 0 to $(23-4)$. The length of the interval, $t_l(d)$, is also randomly selected from the range of 4 to 24. The final time, $t_f(d)$, is calculated as $t_i(d) - t_l(d)$. This range encompasses low-level attacks with a minimum off-time of 4 h and high-level attacks with a maximum off-time of 24 h.

Price-based load control attacks can be launched when there are varying electricity tariffs throughout the day. In one way to launch these attacks, an attack function is employed to report a constant consumption value throughout the entire day as shown in Equation (4). Here, Avg represents the expected average consumption value, and $E_c(d)$ represents the power consumption readings of day d.

$$g_4(E_c(d,t)) = Avg[E_c(d)] \tag{4}$$

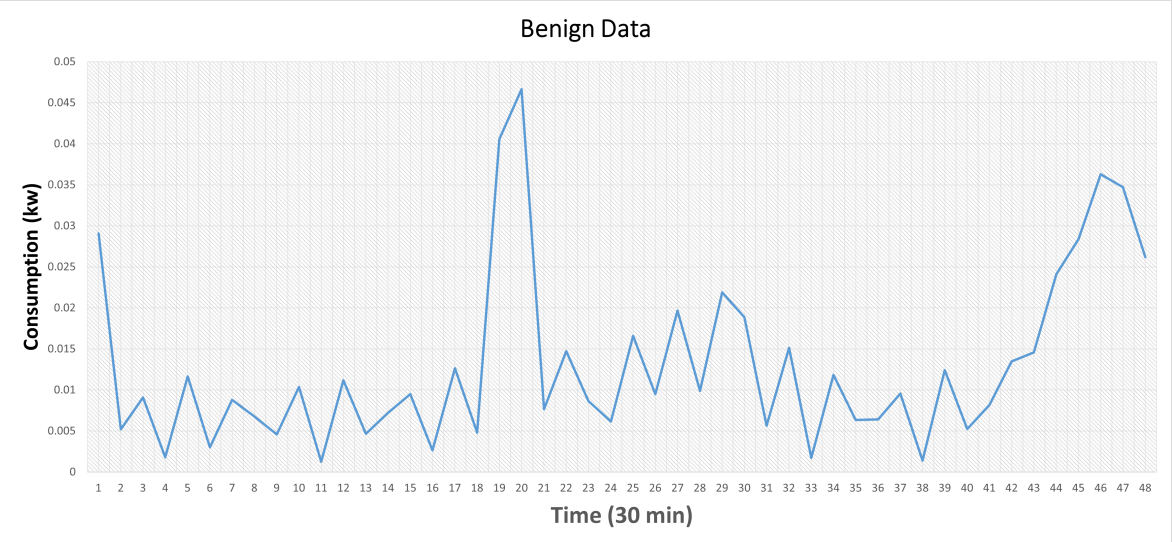
To make the attack stealth and avoid the easy detection of constant consumption values reported throughout the day, a random and time-varying fraction $\Delta(d,t) = \text{rand}(0.1, 0.8)$ is applied as shown in Equation (5).

$$g_5(E_c(d,t)) = \Delta(d,t)Avg[E_c(d)] \tag{5}$$

The final attack function reports high values of energy consumption readings during the time intervals of low electricity tariff, and vice versa (6).

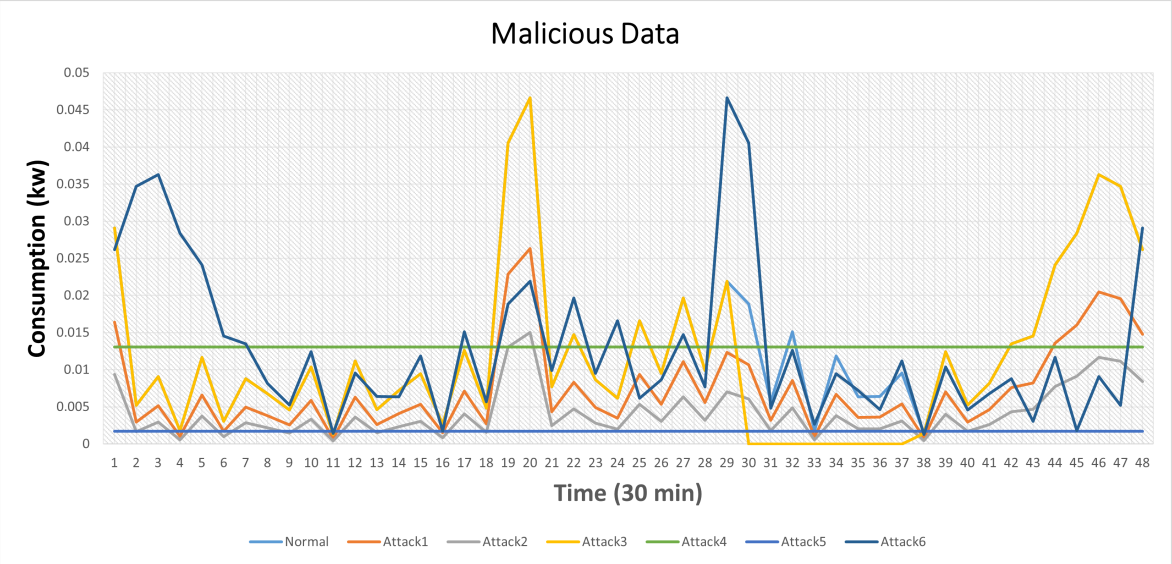
$$g_6(E_c(d,t)) = E_c(d, T-t+1) \tag{6}$$

For each customer, we apply the previously-explained six attack functions to their benign consumption profile matrix, E_c. As a result, each customer is left with six malicious matrices. Figure 1b showcases examples of malicious energy consumption patterns created using the six attack functions, using the benign energy consumption pattern depicted in Figure 1a.



(**a**)

Figure 1. *Cont.*



(**b**)

Figure 1. A sample of benign and malicious power consumption readings. (**a**) A benign power consumption reading. (**b**) Malicious power consumption readings computed by the six attacks.

It can be seem from Figure 1 that some attacks possess a deceptive nature. For instance, attack 3 exhibits high similarity to the benign data, except for a specific period where it provides zero readings. Attack 4 employs an averaging technique, resulting in a constant value throughout the entire day, leading to a good overlap between benign and malicious data. Similarly, Attack 6 reports elevated electricity consumption during periods of low electricity tariffs and vice versa, exacerbating the issue of overlapping between the two types of data.

3.2. Transforming Consumption Readings into Images

To address the challenge of data overlapping, which occurs in specific attacks such as 3, 5, and 6 where benign and malicious data overlap, we propose employing an image-based dataset. Leveraging recent advancements in computer vision and deep learning, we convert the time-series *SM* readings from the ISET dataset into image representations [44,45]. To facilitate this conversion, we initially transformed the CSV file format into 48 columns, each representing one day of data for each user from a total of 536 days. These days were divided into 17 months per user, creating data segments of 30 rows and 48 columns for each month. Subsequently, using Matlab version 2023, we converted these monthly datasets into image representations. Both RGB and grayscale images were generated for both benign and malicious data. During the training phase, we exclusively utilized benign images, while during the testing phase, we employed both benign and malicious images. The resulting images were resized to 32 $\times$ 32 pixels for further analysis. The provided MATLAB code reads data from the CSV file, converts it into RGB and grayscale images, and saves them to a designated folder. It defines the dimensions for each image, extracts the image data from the CSV file, and resizes the images to 32 $\times$ 32 pixels, as demonstrated in Algorithms 1 and 2.

Algorithm 1: Convert CSV Data to Grayscale Images

Data: Data from 'input CSV file'
Result: Grayscale images saved in specific folder
Parameter: $subimage\_height = 30$
Parameter: $subimage\_width = 48$
for *i in range(2 to size(data)–30, with step subimage_height)* **do**
    Extract subimage data *subimage_data* from *data*;
    Convert *subimage_data* to grayscale image *img*;
    Resize *img* to 32×32 pixels, resulting in *img_data_grayscale_resized*;
    Construct file name $file\_name$ as 'image_' concatenated with $((i - 2) / subimage\_height) + 1$;
    Save *img_data_grayscale_resized* as a PNG file in folder 'grayscale' with $file\_name$;

Algorithm 2: Convert CSV Data to RGB Images

Data: Data from 'full_data_6_normalized.csv'
Result: Rescaled RGB images saved in RGB folder
Parameter: $subimage\_height = 30$
Parameter: $subimage\_width = 48$
Input : RGB image *img_data_rgb* with dimensions $subimage\_height \times subimage\_width \times 3$
Output : Rescaled RGB image *img_data_rgb_resized* with dimensions $32 \times 32 \times 3$
for *i in range(2 to size(data, 1) - 6, with step subimage_height)* **do**
    Extract subimage data *subimage_data* from *data*;
    Reshape *subimage_data* into *img_data_r*, *img_data_g*, and *img_data_b* with dimensions $subimage\_height \times (subimage\_width/3)$;
    Scale *img_data_r*, *img_data_g*, and *img_data_b* to the range [0, 255];
    Create *img_data_rgb* by concatenating *img_data_r*, *img_data_g*, and *img_data_b*;
    Resize *img_data_rgb* to $32 \times 32 \times 3$, resulting in *img_data_rgb_resized*;
    Construct file name $file\_name$ as 'image_' concatenated with $((i - 2) / subimage\_height) + 1$;
    Save *img_data_rgb_resized* as a PNG file in folder 'RGB' with $file\_name$;

Figures 2 and 3 illustrate samples of the converted images from each class. Specifically, Figures 2a and 3a depict the benign images in RGB and grayscale formats, respectively. Meanwhile, Figures 2b–g and 3b–g illustrate the six types of attacks in both RGB and grayscale images. Notably, attacks 4 and 5 produce distinctive images in both RGB and grayscale formats compared to the other attacks and the benign image. This distinction arises because these two attacks entirely replace the benign sample with a new sample containing the average power consumption throughout the day. This indicates that the malicious sample significantly differs from the benign data, making it easier for the detector to identify them.

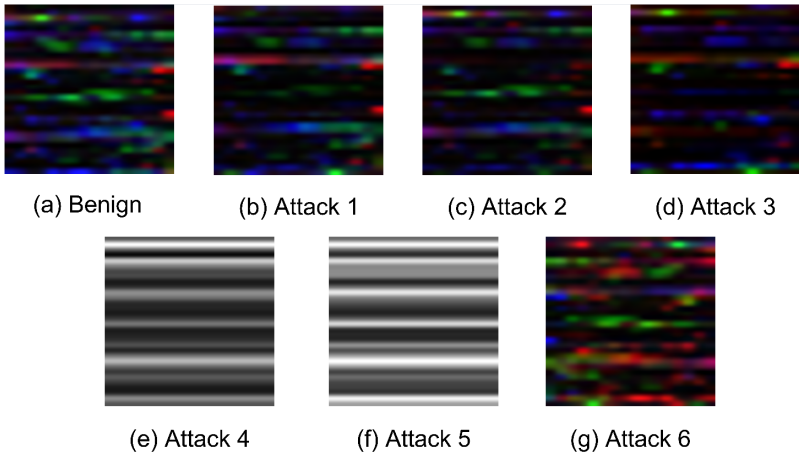


Figure 2. Visualization of Images Generated in the Dataset Conversion Process (RGB).

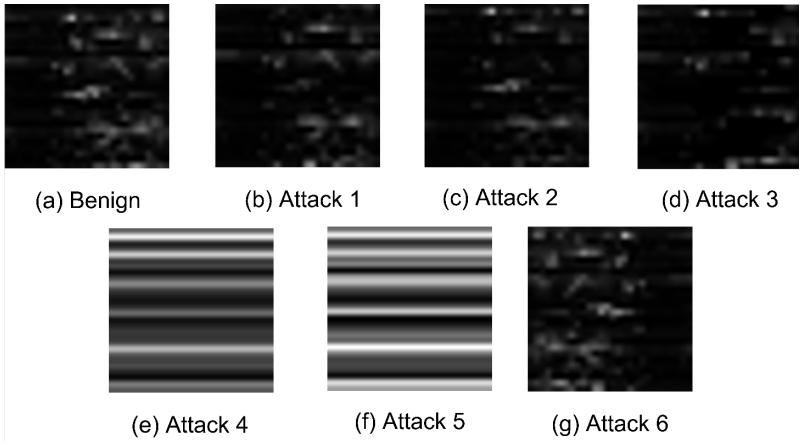


Figure 3. Visualization of Images Generated in the Dataset Conversion Process (grayscale).

On the other hand, the images of attack 3 in Figures 2d and 3d indicate that the detection of this attack relies on the duration of zero-time interval $[t_i(d), t_f(d)]$ introduced by the malicious consumer. As this interval increases, so does the probability of detection by the detector because the image becomes far from the benign image. Furthermore, the images of attack 6 in Figures 2g and 3g indicate that they are far from the benign images because of the flipping done by the attack and thus by detecting the correlations within the data, the attack can be detected.

4. Preliminaries

In this section, we provide an overview of the essential concepts utilized in our solution, including Autoencoders, one-class classification, *OC-SVM*, *SVDD*, *DSVDD*, and *GMM*.

4.1. Autoencoders

Autoencoders are a type of artificial neural network architecture that learns compression and reconstruction of input data [46]. They consist of an encoder network and a decoder network, which are typically symmetrical in structure as illustrated in Figure 4. The encoder network takes the input data and maps it to a lower-dimensional latent space representation [47]. This latent representation serves as a compressed encoding of the input data, capturing its most essential features. The decoder network then aims to reconstruct the original input from the latent representation [48].

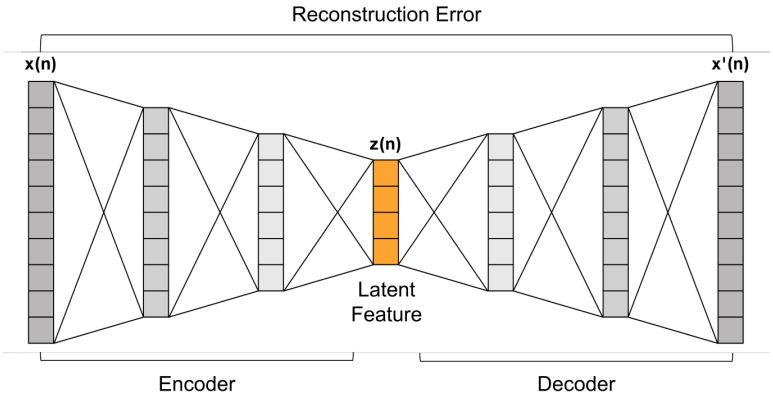


Figure 4. Structure of an Autoencoder.

During the training process, autoencoders aim to reduce the gap between the original input and its reconstruction version [49]. This is achieved through an optimization algorithm such as gradient descent. By minimizing the reconstruction error, the autoencoder learns a compressed representation that captures the most salient information in the data. This capability makes autoencoders useful for dimensionality reduction tasks, where high-dimensional data can be efficiently represented in a lower-dimensional space [50].

Autoencoders can also be used for anomaly detection. During training, they learn the reconstruction of normal or expected patterns in the data [51]. When presented with anomalous data during testing, the reconstruction error tends to be higher, indicating the presence of anomalies. This makes autoencoders valuable in various applications including fraud detection and fault diagnosis [52].

In addition to their utility in unsupervised learning tasks, autoencoders have found applications in supervised learning as well. They can be trained on unlabeled data to learn meaningful features and then optimized for a particular supervised task, such as image classification or sentiment analysis. This process, known as pre-training, enables the model to leverage the learned representations and potentially improve performance on the supervised task [53,54].

Furthermore, autoencoders have been extended with variations such as variational autoencoders (VAEs) and denoising autoencoders [55]. VAEs introduce a probabilistic interpretation to the latent space, allowing for generating new data samples [36,38]. Denoising autoencoders are trained to reconstruct clean versions of input data corrupted by noise, which helps in learning robust representations and denoising capabilities.

In summary, autoencoders are powerful neural network architectures that learn to compress and reconstruct input data. Their versatility and ability to capture meaningful features make them valuable in tasks such as dimensionality reduction, anomaly detection, pre-training for supervised learning, and generative modeling.

4.2. One-Class Classification

One-class classification, also known as one-class learning or outlier detection, is an *ML* approach that focuses on training a model to classify instances belonging to a single class. Unlike traditional classification, where multiple classes are considered, one-class classification aims to distinguish normal (or inlier) instances from anomalies (or outliers) [56]. In one-class classification, the training data consists only of examples from the target class, representing the normal behavior or characteristics of the data. The goal is to build a model that can accurately identify and generalize the patterns and properties of the target class, enabling it to identify cases that substantially differ from the learnt normal behavior [57,58].

One-class classifiers, such as *OC-SVM*, attempt to define a decision boundary or construct a representation of the target class in the feature space. During the testing or

inference phase, the model assigns new instances either as part of the target class (inliers) or as outliers based on their proximity to the learned representation or decision boundary [59].

One-class classification finds applications in various domains, such as fraud detection, network intrusion detection, anomaly detection in industrial systems, and outlier detection in healthcare or finance. It is particularly useful in scenarios where obtaining labeled instances of outliers or anomalies is difficult or costly. By focusing on learning the characteristics of the target class alone, one-class classification provides a valuable tool for identifying unusual or potentially harmful instances in real-world data [60].

4.3. OC-SVM

OC-SVM is a specific approach within the family of SVM that is used for one-class classification or outlier detection tasks. *OC-SVM* is designed to learn a boundary or decision function that encapsulates the normal data instances, aiming to separate them from outliers or anomalies Figure 5. The training process of *OC-SVM* involves constructing a hypersphere or a hyperplane in a high-dimensional feature space. This boundary is positioned to enclose as many normal instances as possible while maintaining a maximal distance from the origin or center of the feature space. By doing so, *OC-SVM* effectively captures the support of the normal class, encompassing the majority of normal instances within the hypersphere or hyperplane [61]. During the testing or inference phase, *OC-SVM* assigns new instances as either normal or anomalous based on their proximity to the learned boundary. Instances lying within the boundary are classified as normal, while those outside are classified as outliers [62,63].

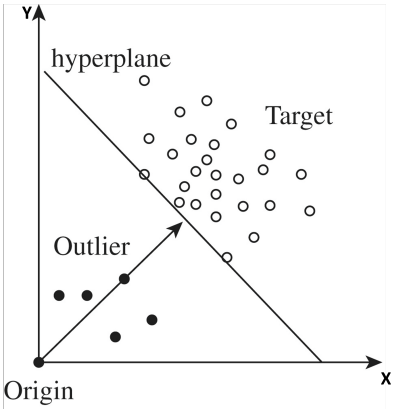


Figure 5. One Class Support Vector Machine (OC-SVM).

OC-SVM is particularly useful in situations where the training data only contains examples from the normal class, making it suitable for one-class classification problems where labeled instances of outliers are scarce or unavailable. It relies on the assumption that the normal class occupies a relatively small region in the feature space, allowing the model to identify instances that significantly deviate from this region. Applications of *OC-SVM* include fraud detection, intrusion detection in network security, anomaly detection in industrial systems, and outlier detection in various domains. It offers an effective approach for identifying and flagging unusual or potentially harmful instances that do not conform to the learned patterns of the normal class.

4.4. SVDD

SVDD is an *ML* algorithm that belongs to the family of SVM but it is specifically designed for one-class classification or anomaly detection. *SVDD* aims to construct a hypersphere or a hyperellipsoid in the feature space that encapsulates most of the training data, which represents the target class or typical instances [64]. In *SVDD*, the objective is to find the center and radius of the hypersphere or hyperellipsoid that minimizes the volume or surface area while containing the training instances. This is achieved by solving

an optimization problem that involves maximizing the margin or distance between the center and the data instances, subject to a constraint that all instances should lie within or on the boundary of the hypersphere or hyperellipsoid as shown in Figure 6. During testing or inference, *SVDD* classifies new instances as either normal or anomalies based on their proximity to the learned boundary. Instances that fall within the boundary are classified as normal, while those lying outside are classified as anomalies.

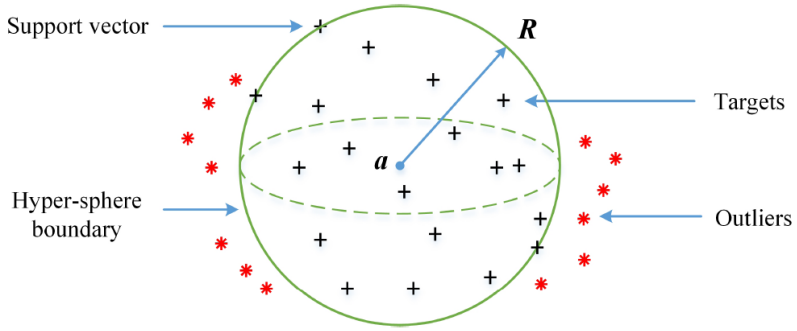


Figure 6. Support Vector Data Description (SVDD).

SVDD is a powerful technique for one-class classification because it can effectively learn the characteristics of the target class and capture the distribution of normal instances in the feature space. It is particularly useful when only instances from the normal class are available for training, making it suitable for scenarios where labeled instances of anomalies or outliers are scarce or difficult to obtain. Applications of *SVDD* include anomaly detection in various domains such as intrusion detection, fraud detection, medical diagnosis, and quality control. By learning a compact representation of the normal instances, *SVDD* provides a robust method for identifying and isolating instances that exhibit substantial deviations from the learnt patterns of the target class.

4.5. DSVDD

The utilization of DNNs offers a new and innovative method for extracting discriminative features directly from raw data. These features, obtained through DNNs, can be defined as some input space $\chi \subseteq \mathbb{R}^s$ and some output space $\beta \subseteq \mathbb{R}^k$. Let $\phi(\cdot; \mathcal{W}) : \chi \longrightarrow \beta$ be a neural network with weights $\mathcal{W} = \{W^1, \cdots, W^L\}$, where W^l corresponds to the weight of hidden layer l. To effectively train the network parameters $\mathcal{W}$ while simultaneously minimizing the volume of the *SVDD* hypersphere, the objective function of the *OC-DSVDD* can be formulated as follows:

$$min \frac{1}{n} \sum_{i=1}^{n} \|\phi(x_i; \mathcal{W}) - a\|^2 + \frac{\lambda}{2} \sum_{l=1}^{L} \|\mathcal{W}^l\|_F^2 \tag{7}$$

where a denotes the center of the sphere, and $\|\cdot\|_F$ is the Frobenius norm. The first term of Equation (7) computes the quadratic loss based on the distances to the sphere center. The second term represents a weight decay regularizer of $\mathcal{W}$ with $\lambda > 0$ introduced as a hyperparameter [13,64].

Equation (7) demonstrates that in the context of *OC-DSVDD*, the characterization of the sphere solely requires the center a. On the other hand, the contraction of the sphere is accomplished by taking the mean value of the distances from each feature to a. It is important to note that *OC-DSVDD* strictly encloses every sample from the training set within the sphere, without allowing any tolerance for outliers. To address this limitation and introduce a more flexible approach, a variant of *DSVDD* with a soft boundary is proposed, outlined as follows:

$$minR^2 + \frac{1}{vn}\sum_{i=1}^{n} Max\{0, \|\phi(x_i; \mathcal{W}) - a\|^2 - R^2\} + \frac{\lambda}{2}\sum_{l=1}^{L} \|\mathcal{W}^l\|_F^2 \quad (8)$$

In contrast to the previous formulation Equation (7), the soft-boundary *DSVDD* incorporates both the center a and the radius R to characterize the sphere. The presence of a penalty term in Equation (8), where $v \in (0, 1]$ manages the trade-off between the volume of the sphere and the extent of violations of the boundary. In other words, it allows for the possibility of certain points being mapped outside the sphere, introducing a level of flexibility in the model. Flexibility refers to the ability of the method to accommodate the case of some benign points being mapped outside the sphere. This ensures that all points inside the sphere are indeed benign while sacrificing a few benign points outside the sphere to prevent any malicious data from being classified as benign.

4.6. Gaussian Mixture Model

GMM is a probabilistic model that represents a dataset as a mixture of multiple Gaussian distributions as shown in Figure 7. It is a popular technique used for unsupervised learning tasks such as clustering and density estimation. In a *GMM*, each Gaussian component represents a cluster or mode in the data distribution [65]. The *GMM* assumes that the recorded data points arise from a combination of gaussian distributions, where each component is associated with a weight indicating its contribution to the overall distribution. The model's goal is to estimate the parameters of the Gaussian components (mean, covariance, and weight) that best fit the data [66].

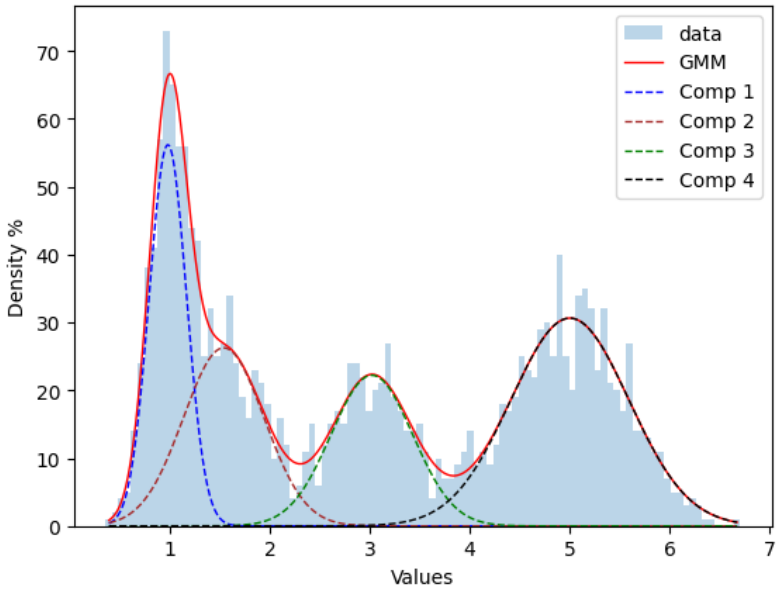


Figure 7. Gaussian Mixture Model (GMM).

The training process of a *GMM* involves an iterative algorithm, such as the Expectation-Maximization (EM) algorithm that has two steps, named expectation step (E-step) and maximization step (M-step). In E-step, the algorithm estimates the posterior probabilities of each data point belonging to each Gaussian component. These probabilities represent the soft assignments of data points to different clusters. In the M-step, the algorithm updates the parameters of the Gaussian components based on the weighted data points.

GMMs have several advantages, including their flexibility in modeling complex data distributions, their ability to capture different modes of data, and the ability to provide probabilistic assignments rather than hard clustering. However, they may be sensitive to

the choice of the number of Gaussian components and are prone to overfitting if the model is overly complex.

The *GMM* can be used for various tasks. In clustering, the *GMM* assigns each data point to the most likely component, allowing for the identification of clusters in the data. *GMMs* can also be used for density estimation, where they can estimate the probability density function of the data. This makes *GMMs* useful for tasks such as outlier detection, anomaly detection, and data generation.

5. Proposed Scheme

In this section, we present the design of our image-based anomaly detectors, which have been developed to identify electricity theft attacks in smart grid AMI as shown in Figure 8.

Our approach introduces new architectures that leverage *DSVDD* techniques. These architectures are specifically designed to improve the overall performance of detecting electricity theft attacks in multiple aspects. Firstly, our detectors should accurately identify instances of electricity theft by relying solely on benign data. This ensures that the detectors are effective in flagging abnormal consumption patterns associated with theft. Secondly, our detectors should be able to identify new types of attacks that have not been trained on before. This adaptability allows for robust detection capability even in the presence of evolving attack methods. Lastly, our detectors should address the challenge of data overlapping between benign and malicious samples. This issue arises due to deceptive attacks that attempt to make malicious data resemble benign patterns. By incorporating advanced techniques, our detectors can effectively differentiate between the two types of data, overcoming the challenge of data overlapping.

Deep learning is a subfield within representation learning that leverages model architectures featuring multiple processing layers. These layers work together to acquire data representations characterized by multiple levels of abstraction. This characteristic enables the encoding of a diverse range of features within a compact and distributed framework. Deep neural networks, particularly those with multiple layers, excel at learning hierarchical representations of data. This capability is especially advantageous for handling data with inherent hierarchical structures, such as images or text.

In the context of *DSVDD*, we present a pioneering approach to unsupervised anomaly detection [13]. *DSVDD* is designed to uncover the shared underlying patterns within a data distribution. This is achieved through the training of a neural network, which is tasked with fitting the network outputs within a hypersphere of minimal volume. This innovative approach harnesses the power of deep learning to distill and represent complex data distributions efficiently.

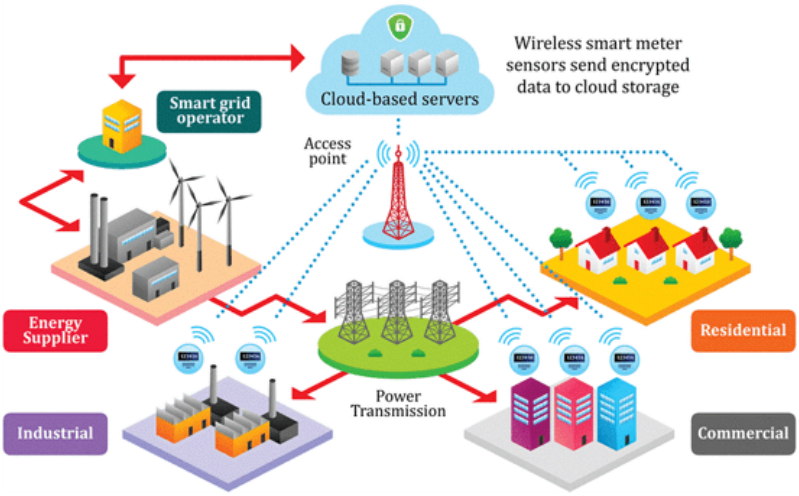


Figure 8. Smart Grid Network Digram [67].

5.1. OC-DSVDD

After transforming the ISET dataset from its original time series format, stored as a .csv file, into both RGB and grayscale images, we proceeded to employ a 2D autoencoder to build the *OC-DSVDD*. This *OC-DSVDD* utilizes the latent features extracted by the autoencoder to construct a hypersphere within the feature space. This hypersphere is characterized by a center, denoted as a, and a radius, denoted as R. The primary objective is to encapsulate the majority of the training data, which corresponds to the benign images, as elaborated in Section 4.4. Figure 9 visually outlines the architecture of the proposed scheme incorporating *OC-DSVDD*.

Within the broader context of *OC-DSVDD*, we introduce an innovative approach to unsupervised anomaly detection. *DSVDD's* core mission is to unveil shared underlying patterns within a given data distribution. This is effectively accomplished by training a neural network, tasked with fitting the network outputs within a hypersphere of minimal volume. This pioneering approach effectively leverages the capabilities of deep learning to efficiently distill and represent complex data distributions.

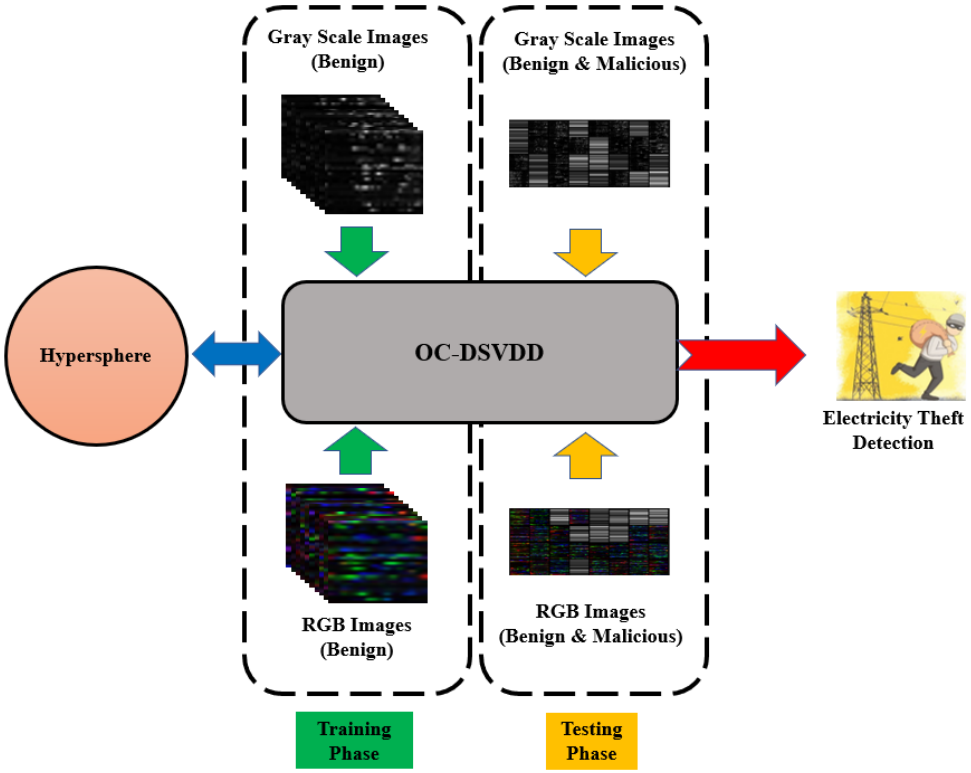


Figure 9. Our proposed Scheme OC-SVDD.

5.2. DSVDD with OC-SVM

In this section, we integrate *OC-SVM* with *DSVDD* to leverage its capability of making decisions based on the decision boundary obtained during the training process, rather than relying on the threshold created by *OC-DSVDD*. Furthermore, we propose that integrating *OC-SVM* with *DSVDD* can enhance the results and eliminate the need for using a threshold to do classification. In the case of *DSVDD*, the threshold is calculated solely based on benign data using the center a and radius R, without considering the malicious data, as done in [36,38] that uses the autoencoder alone. Figure 10 depicts the structure of the proposed scheme employing DSVDD-OCSVM.

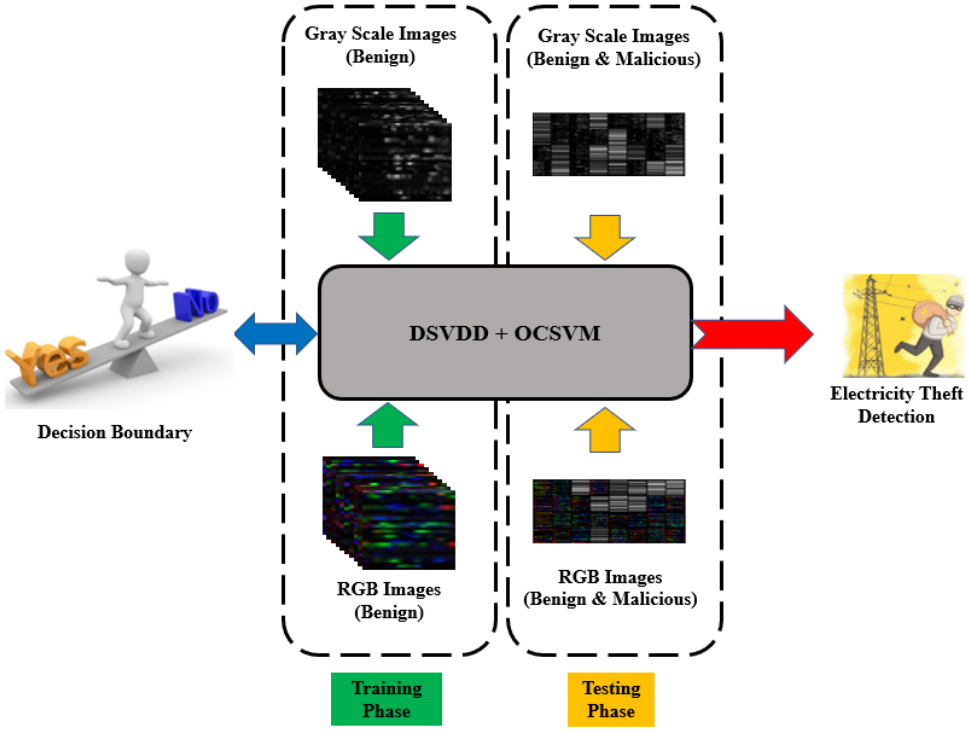


Figure 10. Our proposed Scheme DSVDD with OC-SVM.

5.3. DSVDD with GMM

GMM possesses the capability of estimating the probability density function of data and exhibits flexibility in modeling complex data distributions. It can capture different modes of data and provides probabilistic assignments, making it useful for tasks such as outlier detection and anomaly detection. To determine the maximum likelihood, the EM algorithm is employed, updating parameters iteratively through the expectation step and the maximization step as described in Section 4.6. Note that the maximum likelihood indicates the highest probability that a point is benign, while minimum probability suggests that the point is an outlier. When *GMM* is used in anomaly detection, it can generate probability densities for the samples after training. Consequently, the likelihood serves as an anomaly score for detecting anomalous samples, as abnormal ones exhibit lower likelihoods.

GMM is a good option for better precision and recall in the context of anomaly detection due to its ability to handle the class overlap problem, which can be effectively mitigated by utilizing *GMM* after *DSVDD*. The class overlap problem arises when normal and anomalous data instances share similar characteristics, making it challenging to accurately distinguish between them. *DSVDD* is a powerful technique for defining a decision boundary around normal data, aiming to encapsulate it within a hypersphere. However, when there is significant class overlap, *DSVDD* alone struggles to precisely identify and separate anomalous instances.

By incorporating *GMM* after *DSVDD* as illustrated in Figure 11, the performance in terms of precision and recall can be improved. *GMM* has the ability to model complex data distributions and capture different modes, which allows it to better handle situations where there is an overlap between normal and anomalous data. *GMM* can provide a more nuanced understanding of the underlying data distribution and assign probabilistic scores to individual instances. This combination of *DSVDD* and *GMM* leverages the strengths of both methods, where the *DSVDD's* strength is in its ability for defining the initial decision boundary and *GMM's* strength is for refining anomaly detection by considering the probability densities of the samples. By utilizing *GMM* after *DSVDD*, the class overlap problem can be mitigated, leading to improved precision and recall in electricity theft detection.

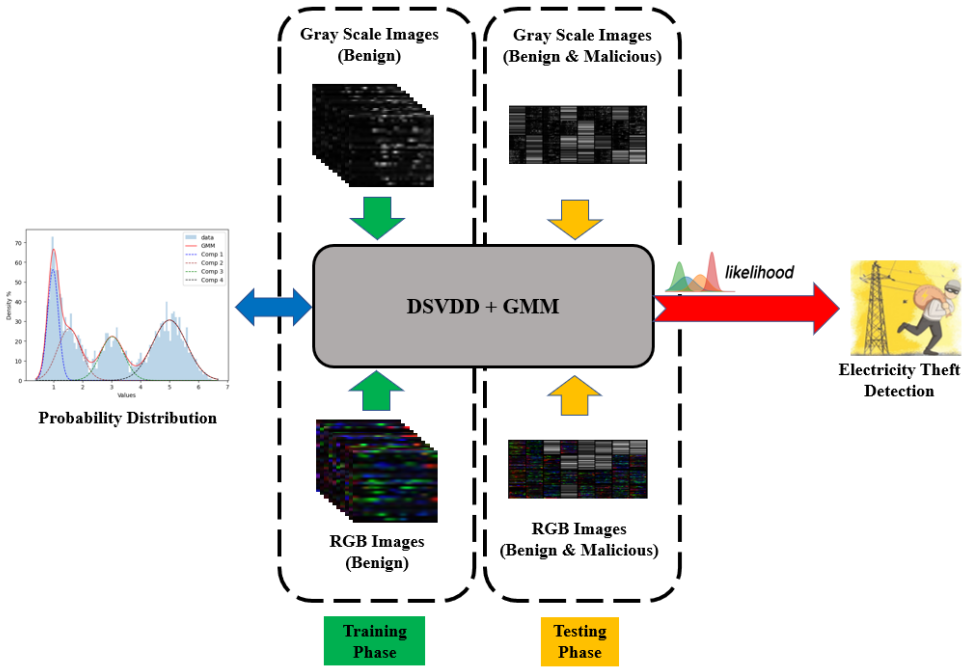


Figure 11. Our proposed Scheme DSVDD with GMM.

6. Experimental Work

The detectors being investigated are trained and tested using the PyTorch API. The electricity theft detectors, including benchmarks, are first trained offline at the electricity utility company. Following that, the electricity company performs real-time online detection to identify malicious samples.

6.1. Evaluation Matrices

The number of correctly identified malicious samples is represented by true positive (TP), while the number of correctly identified benign samples is represented by true negative (TN). On the other hand, a false positive FP refers to the number of benign samples wrongly identified as malicious, and a false negative FN represents the number of malicious samples wrongly identified as benign. To evaluate the performance of the detectors under investigation, we employ multiple evaluation metrics, including accuracy, precision, recall, F1-score, and AUC of the Receiver Operating Characteristics (ROC) curve. These evaluation metrics provide comprehensive insights into the performance of classification models, allowing for a more thorough assessment of their effectiveness in differentiating between positive and negative instances. In this subsection, we define these metrics.

Accuracy assesses the model's overall prediction correctness, calculated as the ratio of correctly classified instances to the total number of instances. The accuracy is computed using Equation (9).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

Precision measures the quality of positive predictions, specifically, the proportion of correctly identified affirmative cases (TPs) among all instances identified as positive. The precision is computed using Equation (10). It helps determine the model's ability to avoid false positives.

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

Recall, referred to as true positive rate or sensitivity, quantifies the model's capability to correctly detect malicious consumption. The recall is computed using Equation (11). Recall provides insight into the model's ability to avoid false negatives.

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

The F1-score is a metric that consolidates precision and recall into one value. It signifies the harmonic mean of precision and recall, offering a balanced assessment of the model's performance. The F1-score is calculated using Equation (12).

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{12}$$

ROC-curve plots the true positive rate (recall) against the false positive rate at different classification thresholds. *AUC*-curve summarizes the performance of the classifier across all possible thresholds. A higher *AUC* value indicates better overall performance, as the model achieves higher true positive rates while maintaining lower false positive rates.

6.2. Benchmark Detectors

We evaluate the performance of our proposed *DSVDD*-based detectors by conducting a comparative analysis against existing supervised and unsupervised detectors. The supervised detectors undergo training and testing using both benign and malicious samples. These encompass shallow classifiers such as Naïve Bayes [15] and multiclass SVM [16], as well as deep classifiers like FF-RNN [22] and CNN-LSTM [26]. Conversely, the unsupervised detectors are trained exclusively on benign samples and subsequently tested on datasets comprising both benign and malicious instances. This category includes shallow models such as *OC-SVM* [16] and ARIMA [31], alongside a variety of deep autoencoders [36–40].

However, static classifiers like SVM, Naïve Bayes, and feed forward-based detectors lack the ability to capture the time-series nature of the dataset or handle the overlap between malicious and benign data. The dynamic ARIMA model can capture temporal dependencies but it has a shallow architecture that fails to detect data overlap adequately. Moreover, the autoencoders introduced in [36–40] assume the existence of malicious data and use it to optimize the threshold of the detector. This dependency on specific malicious data limits the detector's practicality in detecting new attacks, particularly deceptive attacks that cause data overlap.

6.3. Experimental Results and Discussion

Table 4 provides a summary of the performance of the proposed and benchmark detectors using the ISET dataset. The performance metrics are computed using entirely unseen data, i.e., test dataset. It is important to note that this dataset is distinct from the data used for selecting hyperparameters and constructing the *ROC* curves, which were derived from the validation dataset.

The experimental results presented in Table 4 indicate that *OC-DSVDD* demonstrates significant enhancement in performance compared to the best results achieved in the literature. Specifically, achieving a 1.5% increase in recall, 4% in precision, 3.5% in accuracy, 2.5% in F1-score, and an impressive 8% boost in *AUC* compared to the *LSTM-AEA* model, which represents the best performing model in the literature. When we integrate *DSVDD* with *OC-SVM*, further improvements are observed, with a 0.5% and 6% increase in recall and precision respectively, 3.5% in accuracy, 2% in F1-score, and 3% in *AUC*. Additionally, the *DSVDD-GMM* combination results in remarkable enhancements of 3% in recall, 5.5% in precision, 4.5% in accuracy, 4.5% in F1-score, and a substantial 9% improvement in *AUC*.

Table 4. Performance Evaluation Summary.

Detector/Metric	Rec	PR	ACC	F1	AUC
OC-DSVDD	**96**	**97**	**98**	**96**	**98**
DSVDD+OC-SVM	**94.5**	**99**	**98**	**95.5**	**93**
DSVDD+GMM	**97**	**98.5**	**99**	**97.5**	**99**
Benchmark Supervised Classifier					
Naïve Bayes [15]	73	73	77.5	73	70
Multiclass-SVM [16]	91	90	91.5	90.5	89
FF-RNN [22]	90	89	89.5	89.5	88
CNN-LSTM [26]	90.5	89.5	90	90	89
Benchmark Unsupervised Detectors					
OC-SVM [16]	90	89	90.5	89.5	87
ARIMA [31]	86	86	87	86	87
LSTM-AEA [36]	94	93	94.5	93.5	90

Notably, when compared to benchmark detectors, *DSVDD-GMM* achieves the most impressive results, surpassing both supervised and unsupervised models. This demonstrates that the deep *SVDD* and probabilistic attributes of *DSVDD-GMM* contribute to its superior performance compared to shallow and deep classifiers, as well as other models investigated in our study. Figures 12 and 13 illustrate the *AUC-ROC* curves for our proposed detectors in both RGB and grayscale formats, using varying training dataset sizes of 70%, 80%, and 90%, respectively.

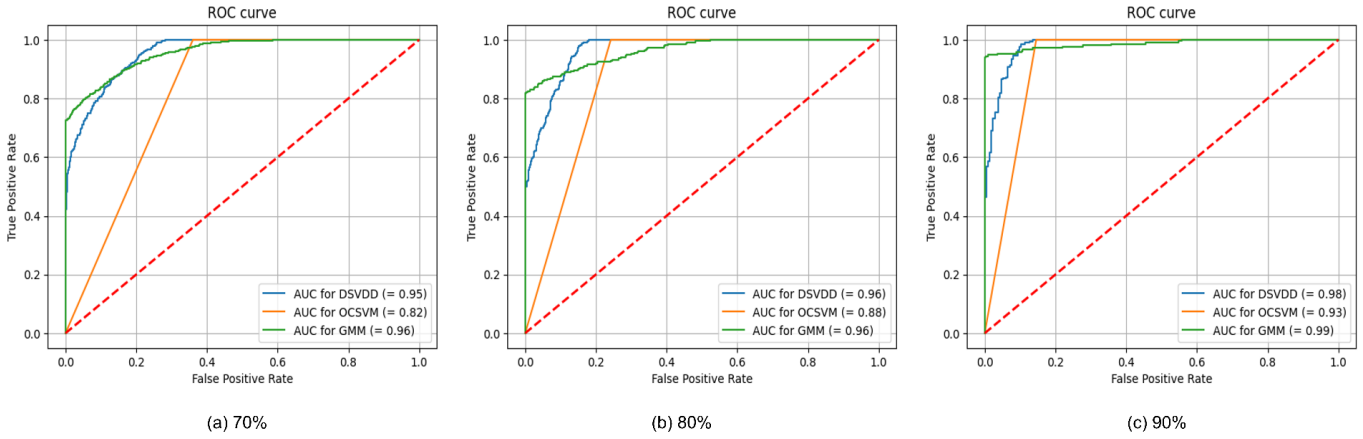


Figure 12. AUC-ROC Curves for the Proposed Schemes Utilizing RGB Images.

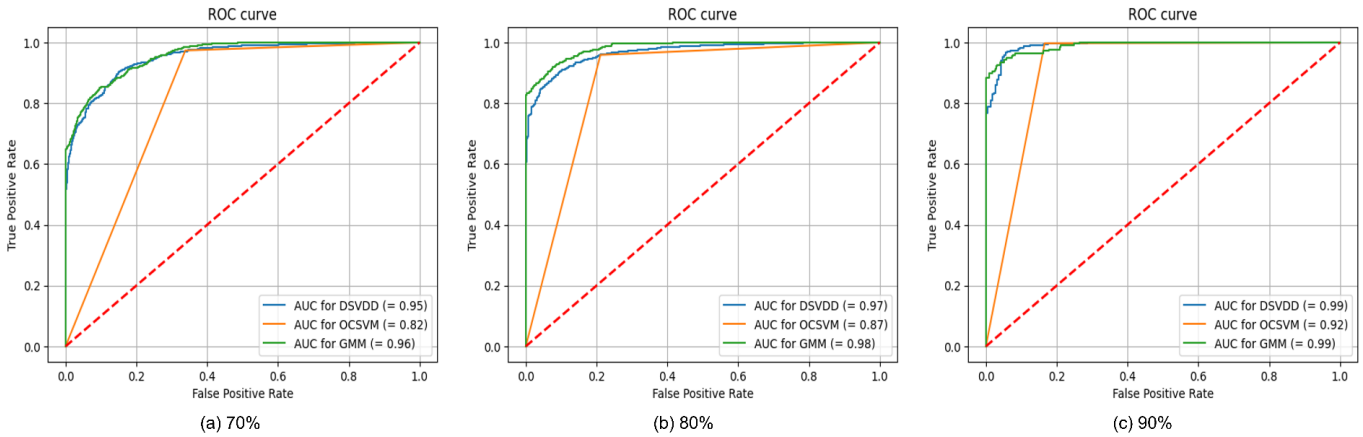


Figure 13. AUC-ROC Curves for the Proposed Schemes Utilizing Grayscale Images.

To assess the influence of varying training dataset sizes on the performance of the detectors, we partitioned the dataset into different ratios, including 70:30, 80:20, and 90:10, where the first number represents the size of the training dataset and the second number represents the size of the test dataset. The detectors' performance results with using these three different ratios are given in Tables 5–7. As evident from the tables, increasing the size of the training data results in improvement across several metrics, including recall, precision, accuracy, F1-Score, and *AUC*.

Table 5. The performance of *OC-DSVDD* at different data sizes.

Training Data Size	ACC	PR	REC	F1	AUC	Image Type
70%	93	86	87	86	95	RGB
80%	97	95	92	93	96	
90%	98	97	95	96	98	
70%	90	78	88	82	95	Gray Scale
80%	90	79	92	83	97	
90%	96	90	96	93	99	

Table 6. The performance of DSVDD + OC-SVM at different data sizes.

Training Data Size	ACC	PR	REC	F1	AUC	Image Type
70%	95	96.5	82	87.5	82	RGB
80%	96	98	88	92	88	
90%	98	99	92.5	95.5	93	
70%	93	87.5	81.5	84	82	Gray Scale
80%	94	89.5	87.5	86.5	87	
90%	98	98	91.5	94.5	92	

Table 7. The performance of DSVDD + GMM at different data sizes.

Training Data Size	ACC	PR	REC	F1	AUC	Image Type
70%	96	97.5	90	90.5	96	RGB
80%	97	98	91	93.5	96	
90%	99	98.5	97	97.5	99	
70%	90	95	87.5	88	96	Gray Scale
80%	97	98	89	92.5	96	
90%	98	98	96	96	99	

In addition, we conducted experiments with varying numbers of epochs to assess their impact on the performance of the proposed detectors. Through this exploration, we observed that increasing the number of epochs led to improved data compactness and overall performance enhancement. These findings are illustrated in Figure 14, where the PCA method was used for visualization. From Figure 14, it is noticeable that the red dots, representing benign samples, exhibit increased compactness as the number of epochs rises. The maximum level of compactness is observed in Figure 14d, where the number of epochs reaches 1000, resulting in the minimum hypersphere radius (R). While the blue dots, which represent the malicious samples, become outside the hypersphere radius (R) and can be easily identified as outliers. This enhanced compactness is reflected in the evaluation metrics, thus addressing the issue of overlapping between benign and malicious samples.

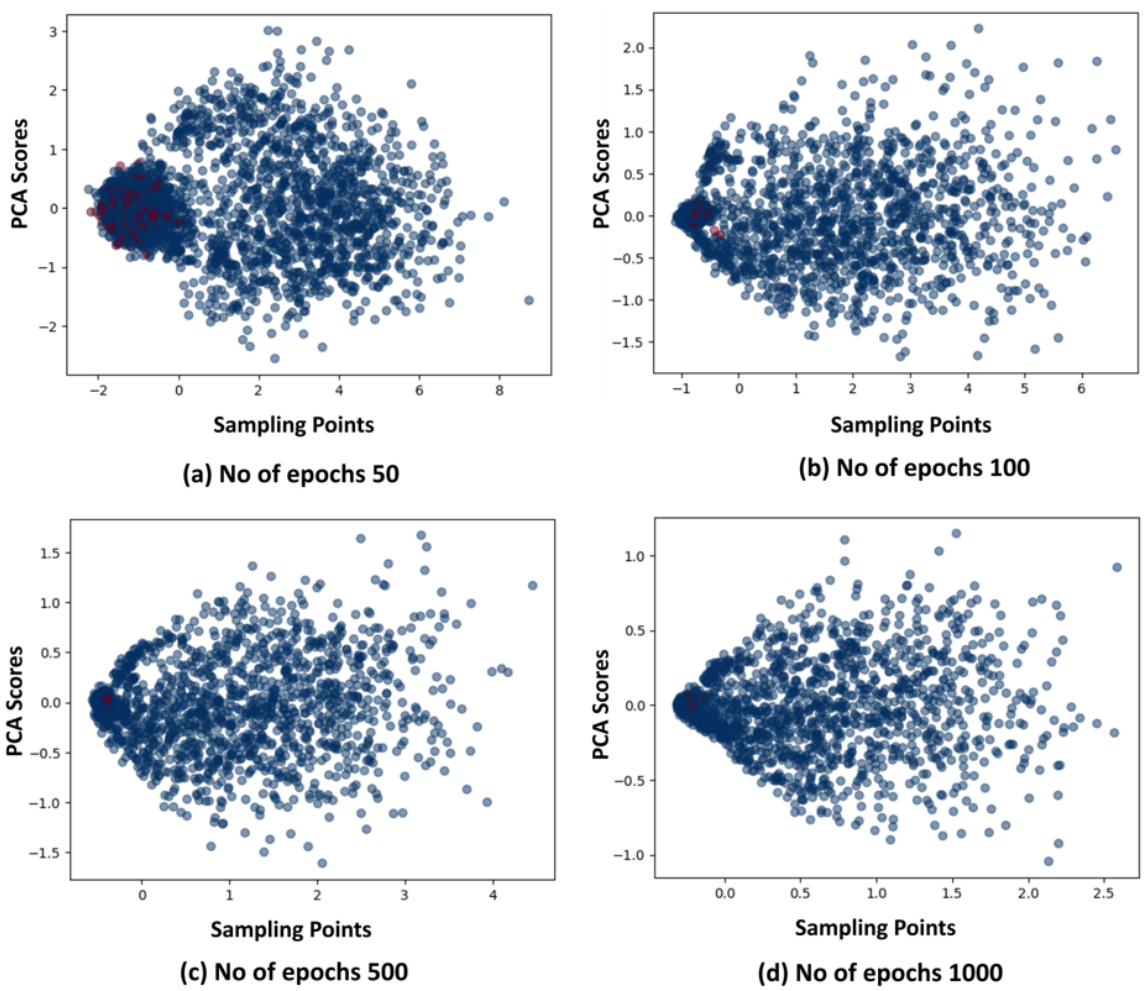


Figure 14. The Impact of the Number of Epochs on the Efficiency of DSVDD.

7. Conclusions and Future Work

In conclusion, we have proposed an efficient one-class false data detector based on DSVDD for smart grids. Our work can address two main limitations including the unavailability of malicious data and the inability to detect new (unseen) attacks. Our approach enhances data representation by converting smart meter time-series readings into images, which serve as the input to our machine learning model. Additionally, it does not only train a DNN but also optimizes a hypersphere that encapsulates the data within the output space. A distinctive feature of our proposal is the integration of both *OCSVM* and *GMM* into a unified class classifier, greatly enhancing the overall effectiveness of our approach. To address the problem of data overlapping, we have introduced a solution that effectively manages these cases, significantly improving the accuracy of our approach. Additionally, we have eliminated the reliance on thresholds for the detection of false data because finding optimal thresholds without knowledge on the malicious data is a challenge. This enhancement greatly boosts the reliability and flexibility of our approach, particularly in boundary decision-making and likelihood estimation. To assess the performance of our detector, we have conducted extensive experiments. The results have demonstrated better performance across a range of evaluation metrics, including accuracy, *AUC*, precision, recall, and F1-score, comparing to the existing proposals in the literature. Specifically, comparing to the current cutting-edge detectors, the results have showed enhancements of 1–3% in terms of recall and 3–9% in terms of *AUC*. However, although our classifier is performing well on the consumption patterns it learned during the training, it may fail when the consumption pattern changes and in this case new retraining process is needed which

require high computation resources. To address this limitation, in our future work, we will integrate our classifier with a reinforcement learning approach. This integration will empower our classifier to efficiently adapt and detect new power consumption patterns.

Author Contributions: Conceptualization, H.H., M.M., M.A. and K.M.; methodology, M.M.F., G.I.S., M.A. and A.M.; software, H.H., M.M. and M.A.; validation, H.H., M.M. and K.M.; formal analysis, H.H., M.M., G.I.S. and K.M.; investigation, H.H., M.M., M.A. and K.M.; resources, H.H., M.M., M.M.F., M.A. and A.M.; data creation, H.H., M.M. and M.A.; writing—original draft preparation, H.H.; writing—review and editing, M.M. and M.A.; visualization, H.H., M.A., A.M. and K.M.; supervision, M.M., M.M.F. and G.I.S.; project administration, H.H., M.M. and K.M.; funding acquisition, M.M., M.M.F. and M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Researchers Supporting Project number (RSPD2023R636), King Saud University, Riyadh, Saudi Arabia, and the Center for Advanced Energy Studies (CAES), USA.

Institutional Review Board Statement: An Institutional Review Board Statement and approval is not needed.

Informed Consent Statement: The study deos not involve humans.

Data Availability Statement: All links to the data used in the study are cited in the text.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DSVDD	Deep Support Vector Data Description
SG	Smart Grid
SM	Smart Meter
AI	Artificial Intelligence
ETD	Electricity Theft Detection
ML	Machine Learning
DL	Deep Learning
ISET	Irish Smart Energy Trial
OC-DSVDD	One-Class Deep Support Vector Data Description
SVM	Support Vector Machine
OC-SVM	One-Class Support Vector Machine
GMM	Gaussian Mixture Model
CPBETD	Consumption Pattern-Based Energy Theft Detector
CNN	Convolutional Neural Network
RF	Random Forest
AdaBoost	Adaptive Boosting algorithm
GRNN	General Regression Neural Network
DT	Decision Tree
GBDT	Gradient Boosting Decision Tree
LR	Logistic Regression
ELM	Extreme Learning Machine
KNN	k-nearest Neighbors
FF	Feed-Forward
RNN	Recurrent Neural Networks
GRU	Gated Recurrent Unit
HDNN	Hybrid Deep Neural Network
PSO	Particle Swarm Optimization
SMOTE	Synthetic Minority Over-sampling
DNN	Deep Neural Network
PCA	Principal Component Analysis
INFLO	Influenced Outlierness
RDOS	Relative Density-based Outlier Score
RGB	Red, Green and Blue

VAE	Variational Autoencoder
TP	True Positive
TN	True Negative
FN	False Negative
FP	False Positive
AUC	Area Under the Curve
ROC	Receiver Operating Characteristics

References

1. Habbak, H.; Mahmoud, M.; Metwally, K.; Fouda, M.M.; Ibrahem, M.I. Load Forecasting Techniques and Their Applications in Smart Grids. *Energies* **2023**, *16*, 1480. [CrossRef]
2. Abdulaal, M.J.; Mahmoud, M.M.E.A.; Bello, S.A.; Khalid, J.; Aljohani, A.J.; Milyani, A.H.; Abusorrah, A.M.; Ibrahem, M.I. Privacy-Preserving Detection of Power Theft in Smart Grid Change and Transmit (CAT) Advanced Metering Infrastructure. *IEEE Access* **2023**, *11*, 68569–68587. [CrossRef]
3. Habbak, H.; Baza, M.; Mahmoud, M.M.E.A.; Metwally, K.; Mattar, A.; Salama, G.I. Privacy-Preserving Charging Coordination Scheme for Smart Power Grids Using a Blockchain. *Energies* **2022**, *15*, 8996. [CrossRef]
4. Habbak, H.; Metwally, K.; Mattar, A.M. Securing Big Data: A Survey on Security Solutions. In Proceedings of the 2022 13th International Conference on Electrical Engineering (ICEENG), Cairo, Egypt, 29–31 March 2022; pp. 145–149. [CrossRef]
5. Saad, M.H.; Serageldin, A.; Salama, G.I. Android spyware disease and medication. In Proceedings of the 2015 Second International Conference on Information Security and Cyber Forensics (InfoSec), Cape Town, South Africa, 15–17 November 2015; pp. 118–125. [CrossRef]
6. El-Toukhy, A.T.; Badr, M.M.; Mahmoud, M.M.E.A.; Srivastava, G.; Fouda, M.M.; Alsabaan, M. Electricity Theft Detection Using Deep Reinforcement Learning in Smart Power Grids. *IEEE Access* **2023**, *11*, 59558–59574. [CrossRef]
7. Takiddin, A.; Ismail, M.; Zafar, U.; Serpedin, E. Robust Electricity Theft Detection Against Data Poisoning Attacks in Smart Grids. *IEEE Trans. Smart Grid* **2021**, *12*, 2675–2684. [CrossRef]
8. Punmiya, R.; Choe, S. Energy Theft Detection Using Gradient Boosting Theft Detector With Feature Engineering-Based Preprocessing. *IEEE Trans. Smart Grid* **2019**, *10*, 2326–2329. [CrossRef]
9. Badr, M.M.; Mahmoud, M.M.E.A.; Abdulaal, M.; Aljohani, A.J.; Alsolami, F.; Balamsh, A. A Novel Evasion Attack Against Global Electricity Theft Detectors and a Countermeasure. *IEEE Internet Things J.* **2023**, *10*, 11038–11053. [CrossRef]
10. Ren, H.; Xu, B.; Wang, Y.; Yi, C.; Huang, C.; Kou, X.; Xing, T.; Yang, M.; Tong, J.; Zhang, Q. Time-Series Anomaly Detection Service at Microsoft. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD'19, New York, NY, USA, 4–8 August 2019; pp. 3009–3017. [CrossRef]
11. Nassif, A.B.; Talib, M.A.; Nasir, Q.; Dakalbab, F.M. Machine Learning for Anomaly Detection: A Systematic Review. *IEEE Access* **2021**, *9*, 78658–78700. [CrossRef]
12. Wang, R.; Nie, K.; Wang, T.; Yang, Y.; Long, B. Deep Learning for Anomaly Detection. In Proceedings of the 13th International Conference on Web Search and Data Mining, WSDM'20, New York, NY, USA,3–7 February 2020; pp. 894–896. [CrossRef]
13. Ruff, L.; Vandermeulen, R.; Goernitz, N.; Deecke, L.; Siddiqui, S.A.; Binder, A.; Müller, E.; Kloft, M. Deep One-Class Classification. *PMLR* **2018**, *80*, 4393–4402.
14. *Smart Metering Project—Electricity Customer Behaviour Trial*, 1st ed.; [dataset]; Irish Social Science Data Archive; Commission for Energy Regulation (CER): Dublin, Ireland, 2012.
15. Murthy, T.S.; Gopalan, N.; Ramachandran, V. A Naive Bayes Classifier for Detecting Unusual Customer Consumption Profiles in Power Distribution Systems—APSPDCL. In Proceedings of the 2019 Third International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 10–11 January 2019; pp. 673–678. [CrossRef]
16. Jokar, P.; Arianpoo, N.; Leung, V.C.M. Electricity Theft Detection in AMI Using Customers' Consumption Patterns. *IEEE Trans. Smart Grid* **2016**, *7*, 216–226. [CrossRef]
17. Wu, R.; Wang, L.; Hu, T. AdaBoost-SVM for Electrical Theft Detection and GRNN for Stealing Time Periods Identification. In Proceedings of the IECON 2018—44th Annual Conference of the IEEE Industrial Electronics Society, Washington, DC, USA, 21–23 October 2018; pp. 3073–3078. [CrossRef]
18. Li, S.; Han, Y.; Yao, X.; Song, Y.; Wang, J.; Zhao, Q. Electricity theft detection in power grids with deep learning and random forests. *J. Electr. Comput. Eng.* **2019**, *2019*, 4136874. [CrossRef]
19. Harshini, C.; Deepthi, G.; Reddy, G.A.; Laxmi, G.V.; Rajasree, G. ELECTRICITY THEFT DETECTION IN POWER GRIDS WITH DEEP LEARNING AND RANDOM FORESTS. *Int. J. Manag. Res. Rev.* **2023**, *13*, 1–10.
20. Jindal, A.; Dua, A.; Kaur, K.; Singh, M.; Kumar, N.; Mishra, S. Decision Tree and SVM-Based Data Analytics for Theft Detection in Smart Grid. *IEEE Trans. Ind. Inform.* **2016**, *12*, 1005–1016. [CrossRef]
21. Yan, Z.; Wen, H. Electricity Theft Detection Base on Extreme Gradient Boosting in AMI. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–9. [CrossRef]
22. Nabil, M.; Ismail, M.; Mahmoud, M.; Shahin, M.; Qaraqe, K.; Serpedin, E. Deep Learning-Based Detection of Electricity Theft Cyber-Attacks in Smart Grid AMI Networks. In *Deep Learning Applications for Cyber Security*; Alazab, M., Tang, M., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 73–102. [CrossRef]

23. Nabil, M.; Mahmoud, M.; Ismail, M.; Serpedin, E. Deep Recurrent Electricity Theft Detection in AMI Networks with Evolutionary Hyper-Parameter Tuning. In Proceedings of the 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Atlanta, GA, USA, 14–17 July 2019; pp. 1002–1008. . [CrossRef]
24. Takiddin, A.; Ismail, M.; Nabil, M.; Mahmoud, M.M.E.A.; Serpedin, E. Detecting Electricity Theft Cyber-Attacks in AMI Networks Using Deep Vector Embeddings. *IEEE Syst. J.* **2021**, *15*, 4189–4198. [CrossRef]
25. Ismail, M.; Shaaban, M.F.; Naidu, M.; Serpedin, E. Deep Learning Detection of Electricity Theft Cyber-Attacks in Renewable Distributed Generation. *IEEE Trans. Smart Grid* **2020**, *11*, 3428–3437. [CrossRef]
26. Hasan, M.N.; Toma, R.N.; Nahid, A.A.; Islam, M.M.M.; Kim, J.M. Electricity Theft Detection in Smart Grid Systems: A CNN-LSTM Based Approach. *Energies* **2019**, *12*, 3310. [CrossRef]
27. Ullah, A.; Javaid, N.; Samuel, O.; Imran, M.; Shoaib, M. CNN and GRU based Deep Neural Network for Electricity Theft Detection to Secure Smart Grid. In Proceedings of the 2020 International Wireless Communications and Mobile Computing (IWCMC), Limassol, Cyprus, 15–19 June 2020; pp. 1598–1602. [CrossRef]
28. Khattak, A.; Bukhsh, R.; Aslam, S.; Yafoz, A.; Alghushairy, O.; Alsini, R. A Hybrid Deep Learning-Based Model for Detection of Electricity Losses Using Big Data in Power Systems. *Sustainability* **2022**, *14*, 13627. [CrossRef]
29. Xia, R.; Gao, Y.; Zhu, Y.; Gu, D.; Wang, J. An attention-based wide and deep CNN with dilated convolutions for detecting electricity theft considering imbalanced data. *Electr. Power Syst. Res.* **2023**, *214*, 108886. [CrossRef]
30. Emadaleslami, M.; Haghifam, M.R.; Zangiabadi, M. A two stage approach to electricity theft detection in AMI using deep learning. *Int. J. Electr. Power Energy Syst.* **2023**, *150*, 109088. [CrossRef]
31. Badrinath Krishna, V.; Iyer, R.K.; Sanders, W.H. ARIMA-Based Modeling and Validation of Consumption Readings in Power Grids. In *Critical Information Infrastructures Security, Proceedings of the 10th International Conference, CRITIS 2015, Berlin, Germany, 5–7 October 2015*; Rome, E., Theocharidou, M., Wolthusen, S., Eds.; Springer: Cham, Switzerland, 2016; pp. 199–210.
32. Singh, S.K.; Bose, R.; Joshi, A. PCA based electricity theft detection in advanced metering infrastructure. In Proceedings of the 2017 7th International Conference on Power Systems (ICPS), Pune, India, 21–23 December 2017; pp. 441–445. [CrossRef]
33. Singh, S.K.; Bose, R.; Joshi, A. Energy theft detection in advanced metering infrastructure. In Proceedings of the 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), Singapore, 5–8 February 2018; pp. 529–534. [CrossRef]
34. Yeckle, J.; Tang, B. Detection of Electricity Theft in Customer Consumption Using Outlier Detection Algorithms. In Proceedings of the 2018 1st International Conference on Data Intelligence and Security (ICDIS), South Padre Island, TX, USA, 8–10 April 2018; pp. 135–140. [CrossRef]
35. Xu, L.; Shao, Z.; Chen, F. A combined unsupervised learning approach for electricity theft detection and loss estimation. *IET Energy Syst. Integr.* **2023**, *5*, 213–227. [CrossRef]
36. Takiddin, A.; Ismail, M.; Zafar, U.; Serpedin, E. Deep Autoencoder-Based Anomaly Detection of Electricity Theft Cyberattacks in Smart Grids. *IEEE Syst. J.* **2022**, *16*, 4106–4117. [CrossRef]
37. Takiddin, A.; Ismail, M.; Zafar, U.; Serpedin, E. Variational Auto-encoder-based Detection of Electricity Stealth Cyber-attacks in AMI Networks. In Proceedings of the 2020 28th European Signal Processing Conference (EUSIPCO), Amsterdam, The Netherlands, 18–21 January 2021; pp. 1590–1594. [CrossRef]
38. Takiddin, A.; Ismail, M.; Zafar, U.; Serpedin, E. Deep Autoencoder-based Detection of Electricity Stealth Cyberattacks in AMI Networks. In Proceedings of the 2021 International Symposium on Signals, Circuits and Systems (ISSCS), Iasi, Romania, 15–16 July 2021; pp. 1–6. [CrossRef]
39. Takiddin, A.; Ismail, M.; Serpedin, E. Detection of Electricity Theft False Data Injection Attacks in Smart Grids. In Proceedings of the 2022 30th European Signal Processing Conference (EUSIPCO), Belgrade, Serbia, 29 August–2 September 2022; pp. 1541–1545. [CrossRef]
40. Takiddin, A.; Ismail, M.; Serpedin, E. Robust Data-Driven Detection of Electricity Theft Adversarial Evasion Attacks in Smart Grids. *IEEE Trans. Smart Grid* **2023**, *14*, 663–676. [CrossRef]
41. Khan, Z.A.; Adil, M.; Javaid, N.; Saqib, M.N.; Shafiq, M.; Choi, J.G. Electricity Theft Detection Using Supervised Learning Techniques on Smart Meter Data. *Sustainability* **2020**, *12*, 8023. [CrossRef]
42. Hussain, S.; Mustafa, M.W.; Jumani, T.A.; Baloch, S.K.; Alotaibi, H.; Khan, I.; Khan, A. A novel feature engineered-CatBoost-based supervised machine learning framework for electricity theft detection. *Energy Rep.* **2021**, *7*, 4425–4436. [CrossRef]
43. Abdelrahman, A.A.; Dahshan, H.; Salama, G.I. Enhancing the Actual Throughput of the AES Algorithm on the Pascal GPU Architecture. In Proceedings of the 2018 3rd International Conference on System Reliability and Safety (ICSRS), Barcelona, Spain, 23–25 November 2018; pp. 97–103. [CrossRef]
44. Hussain, F.; Abbas, S.G.; Husnain, M.; Fayyaz, U.U.; Shahzad, F.; Shah, G.A. IoT DoS and DDoS Attack Detection using ResNet. In Proceedings of the 2020 IEEE 23rd International Multitopic Conference (INMIC), Bahawalpur, Pakistan, 5–7 November 2020; pp. 1–6. [CrossRef]
45. Gopi, A.P.; Gowthami, M.; Srujana, T.; Gnana Padmini, S.; Durga Malleswari, M. Classification of Denial-of-Service Attacks in IoT Networks Using AlexNet. In *Human-Centric Smart Computing*; Bhattacharyya, S., Banerjee, J.S., Köppen, M., Eds.; Springer: Singapore, 2023; pp. 349–357.
46. Zhang, G.; Liu, Y.; Jin, X. A survey of autoencoder-based recommender systems. *Front. Comput. Sci.* **2020**, *14*, 430–450. [CrossRef]
47. Tschannen, M.; Bachem, O.; Lucic, M. Recent advances in autoencoder-based representation learning. *arXiv* **2018**, arXiv:1812.05069.

48. Bank, D.; Koenigstein, N.; Giryes, R. Autoencoders. *arXiv* **2020**, arXiv:2003.05991.
49. Zhai, J.; Zhang, S.; Chen, J.; He, Q. Autoencoder and Its Various Variants. In Proceedings of the 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 7–10 October 2018; pp. 415–419. [CrossRef]
50. Chen, Z.; Yeo, C.K.; Lee, B.S.; Lau, C.T. Autoencoder-based network anomaly detection. In Proceedings of the 2018 Wireless Telecommunications Symposium (WTS), Phoenix, AZ, USA, 17–20 April 2018; pp. 1–5. [CrossRef]
51. Chen, D.; Zhang, R. Building Multimodal Knowledge Bases with Multimodal Computational Sequences and Generative Adversarial Networks. *IEEE Trans. Multimed.* **2023**, 1–14. [CrossRef]
52. Li, Y.; Liu, L.; Deng, S.; Qin, H.; El-Yacoubi, M.A.; Zhou, G. Memory-Augmented Autoencoder based Continuous Authentication on Smartphones with Conditional Transformer GANs. *IEEE Trans. Mob. Comput.* **2023**, 1–16. [CrossRef]
53. Furtney, I.; Bradley, R.; Kabuka, M.R. Patient Graph Deep Learning to Predict Breast Cancer Molecular Subtype. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2023**, 3117–3127. [CrossRef] [PubMed]
54. Romanelli, F.; Martinelli, F. Synthetic Sensor Data Generation Exploiting Deep Learning Techniques and Multi-Modal Information. *IEEE Sens. Lett.* **2023**, *7*, 1–4. [CrossRef]
55. Chiang, H.T.; Hsieh, Y.Y.; Fu, S.W.; Hung, K.H.; Tsao, Y.; Chien, S.Y. Noise Reduction in ECG Signals Using Fully Convolutional Denoising Autoencoders. *IEEE Access* **2019**, *7*, 60806–60813. [CrossRef]
56. Perera, P.; Patel, V.M. Learning Deep Features for One-Class Classification. *IEEE Trans. Image Process.* **2019**, *28*, 5450–5463. [CrossRef]
57. Goyal, S.; Raghunathan, A.; Jain, M.; Simhadri, H.V.; Jain, P. DROCC: Deep Robust One-Class Classification. In Proceedings of the 37th International Conference on Machine Learning, Virtual Event, 13–18 July 2020; Volume 119, pp. 3711–3721.
58. Khan, S.S.; Madden, M.G. One-class classification: Taxonomy of study and review of techniques. *Knowl. Eng. Rev.* **2014**, *29*, 345–374. [CrossRef]
59. Khan, S.S.; Madden, M.G. A Survey of Recent Trends in One Class Classification. In *Artificial Intelligence and Cognitive Science, Proceedings of the 20th Irish Conference, AICS 2009, Dublin, Ireland, 19–21 August 2009*; Coyle, L., Freyne, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 188–197.
60. Sohn, K.; Li, C.L.; Yoon, J.; Jin, M.; Pfister, T. Learning and evaluating representations for deep one-class classification. *arXiv* **2020**, arXiv:2011.02578.
61. Zhu, F.; Yang, J.; Gao, C.; Xu, S.; Ye, N.; Yin, T. A weighted one-class support vector machine. *Neurocomputing* **2016**, *189*, 1–10. [CrossRef]
62. Amer, M.; Goldstein, M.; Abdennadher, S. Enhancing One-Class Support Vector Machines for Unsupervised Anomaly Detection. In Proceedings of the ACM SIGKDD Workshop on Outlier Detection and Description, ODD'13, New York, NY, USA, 11 August 2013; pp. 8–15. [CrossRef]
63. Shin, H.J.; Eom, D.H.; Kim, S.S. One-class support vector machines—An application in machine fault detection and classification. *Comput. Ind. Eng.* **2005**, *48*, 395–408. [CrossRef]
64. Liu, C.; Gryllias, K. A deep support vector data description method for anomaly detection in helicopters. *Phm Soc. Eur. Conf.* **2021**, *6*, 9. [CrossRef]
65. Zong, B.; Song, Q.; Min, M.R.; Cheng, W.; Lumezanu, C.; Cho, D.; Chen, H. Deep Autoencoding Gaussian Mixture Model for Unsupervised Anomaly Detection. In Proceedings of the International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.
66. McLachlan, G.J.; Rathnayake, S. On the number of components in a Gaussian mixture model. *WIREs Data Min. Knowl. Discov.* **2014**, *4*, 341–355. [CrossRef]
67. Yip, S.C.; Wong, K.; Hew, W.P.; Gan, M.T.; Phan, R.C.W.; Tan, S.W. Detection of energy theft and defective smart meters in smart grids using linear regression. *Int. J. Electr. Power Energy Syst.* **2017**, *91*, 230–240. [CrossRef]