energies

Article

Energy Theft Detection Model Based on VAE-GAN for Imbalanced Dataset

Youngghyu Sun, Jiyoung Lee, Soohyun Kim, Joonho Seon, Seongwoo Lee, Chanuk Kyeong and Jinyoung Kim *

Department of Electronic Convergence Engineering, Kwangwoon University, Seoul 01897, Republic of Korea

* Correspondence: jinyoung@kw.ac.kr; Tel.: +82-02-940-5567

Abstract: Energy theft causes a lot of economic losses every year. In the practical environment of energy theft detection, it is required to solve imbalanced data problem where normal user data are significantly larger than energy theft data. In this paper, a variational autoencoder-generative adversarial network (VAE-GAN)-based energy theft-detection model is proposed to overcome the imbalanced data problem. In the proposed model, the VAE-GAN generates synthetic energy theft data with the features of real energy theft data for augmenting the energy theft dataset. The obtained balanced dataset is applied to train a detector which is designed as one-dimensional convolutional neural network. The proposed model is simulated on the practical dataset for comparing with various generative models to evaluate their performance. From simulation results, it is confirmed that the proposed model outperforms the other existing models. Additionally, it is shown that the proposed model is also very useful in the environments of extreme data imbalance for a wide variety of applications by analyzing the performance of detector according to the balance rate.

Keywords: energy theft; imbalanced dataset; data augmentation; variational autoencoder; generative adversarial network

Citation: Sun, Y.; Lee, J.; Kim, S.; Seon, J.; Lee, S.; Kyeong, C.; Kim, J. Energy Theft Detection Model Based on VAE-GAN for Imbalanced Dataset. *Energies* **2023**, *16*, 1109. https://doi.org/10.3390/en16031109

Academic Editor: Fernando Morgado-Dias

Received: 12 December 2022
Revised: 9 January 2023
Accepted: 16 January 2023
Published: 19 January 2023

1. Introduction

Every year, energy theft causes globally enormous economic losses of electric utilities [1]. The energy theft can be defined as an illegal usage of energy service with dishonest intentions such as meter tampering, bypassing, direct tapping from feeders, etc. [2]. By implementing smart meters and advanced metering infrastructure (AMI), the electric utilities may prevent traditional energy theft [3]. However, intelligent energy thefts can manipulate smart meters by advanced techniques [4].

To detect the intelligent energy thefts, many studies have actively been conducted in the past decades. The energy theft detection methods can be divided into four types: hardware-based, grid analysis-based, game theory-based and machine learning-based methods [3]. Hardware-based methods have been treated with developing devices for energy theft detection. The modified current meter was designed by three modules to detect malicious consumers [5]. In addition, a device for anomaly detection was proposed that compares the deviation of real-time and estimated values [6]. Grid analysis-based methods have been usually focused on detecting abnormal consumers by monitoring the current flow and line voltage of a power grid. The method formulated by power flow analysis was presented to effectively detect energy thieves [7]. A state estimation-based method with weighted least-squares was presented for estimating the voltage of consumers and comparing it to the measured voltage [8]. Game theory-based methods have been dealt with the interactive decision-making process among energy thieves, utilities and consumers. The energy theft detection problem was conceptualized as a game between energy thieves and utilities [9], where the Nash equilibrium was attempted to be found over several assumptions. To detect energy theft meters, a Stackelberg game-theoretic model based on Benford's law was proposed [10]. Due to limitations of machine learning models, machine learning-based methods have recently been focused on deep learning techniques. A novel convolutional neural network architecture was proposed to handle the energy theft detection problem [11]. Using a clustering algorithm that learns patterns from unlabeled data, the correlation between meter data and outliers was determined [12]. An energy theft detection scheme based on a semi-supervised learning algorithm was presented for the imbalanced labeled data [13].

1.1. Related Works

An imbalanced data problem is major issue of the deep learning-based energy theft methods dealing with a practical environment [14–18]. In the practical environment, it has been considerably simpler to obtain the energy consumption data of users than the energy theft data. The biased dataset can be formed due to the tiny number of energy theft events. The imbalanced data problem may cause performance degradation of deep learning models. In general, the approaches for designing the feature cost-efficiency model have been applied to alleviate the imbalanced data problem, but occasionally the approaches for balancing the datasets have also been employed [14–19].

In the approaches of balancing the datasets, synthetic data are generated by augmenting energy theft data as much as energy consumption data. Although simple methods such as false data injection [15] can be applied for generating the synthetic data, the synthetic data generated by the simple methods may be unsuitable for the practical environment. In the practical environment, the synthetic data with realistic features is needed to improve the degraded performance of deep learning models by imbalanced dataset. The synthetic data with realistic features can be augmented by sampling methods or generative models [16,17].

Recently, synthetic minority oversampling technique (SMOTE), which is sampling method, is a well-known data augmentation method for imbalanced dataset problem [20]. SMOTE generates synthetic data of minority class using K-nearest neighbor algorithm. However, SMOTE may lead to the overfitting problem and generate other class data due to sampling without consideration of other classes.

Though the generative models have been actively employed in various fields for data augmentation, the generative models have been rarely applied to research area of energy theft detection with imbalance data problem. The generative models can generate synthetic data with diversity. The degraded performance of deep learning models can be enhanced by training the balanced dataset with the synthetic data generated from the generative models. However, due to addition of generative models in the augmenting process, the complexity of energy theft detection model is increased. In this study, it is attempted to apply a generative model to an energy theft detection model with imbalance dataset and to confirm the performance improvement.

In deep learning-based methods, a variety of models have been utilized as detectors, including convolutional neural network (CNN), since CNN can handle large amount of data, automate feature extraction and perform well enough in terms of classification [3]. By designing the cost efficiency deep learning model, the performance of the energy theft detection model can be enhanced. In this study, it is focused on the effect of balancing datasets rather than on designing the cost efficiency detector.

1.2. Main Contributions

In this paper, an energy theft detection model based on a generative adversarial network combining a variational autoencoder is proposed to solve the imbalanced data problem over a practical environment. For an efficient detection of energy theft, handling of imbalanced datasets is known to be a crucial point. The augmentation of synthetic data with realistic features is one of the most promising solutions to tackle the imbalance problems of the dataset. From this viewpoint, among various machine learning algorithms, generative models have been working well by efficiently relieving the data imbalance. Therefore, the representative two models have been chosen and combined to enhance detection performance of energy theft. One is a generative adversarial network (GAN), which generates samples as similar as possible to the data in training dataset, and the other is a variational autoencoder (VAE) which generates new types of samples with features in the training dataset. The GAN is known to be unstable during the learning phase while VAE is more stable than GAN. By combining GAN and VAE, a generative model with high quality and stability can be produced. The synthetic data is generated from a variational autoencoder-generative adversarial network (VAE-GAN) to mitigate the imbalanced data problem in the proposed energy theft model. Then, a detector, which is designed as one dimensional (1D-CNN), is trained by both real and synthetic data for detecting energy theft. The 1D-CNN is employed for dealing with energy consumption data and energy theft data, which are one-dimensional time-series data.

For clarity, the main contributions of this paper are summarized as follows:

- The GAN combined with VAE is proposed for energy theft detection in a practical environment. VAE generates data with diversity and appears to be stable in its learning process. Meanwhile, GAN generates data with fidelity, which turns out to be unstable.
- To ensure that VAE-GAN can be effective, the proposed energy theft detection model is evaluated in terms of data generation and classification over different data augmentation schemes.
- The performance of the proposed model is evaluated according to the balance rate, so as to look into the imbalanced data problem.

The remainder of this paper is organized as follows. Section 2 presents the proposed energy theft detection model based on the VAE-GAN and metrics for evaluating the performance of the proposed model. In Section 3, the simulation settings and results are reported. Here, the performance of the proposed energy theft detection model is analyzed and discussed. Finally, concluding remarks are given in Section 4.

2. System Model

In this section, the VAE-GAN and the proposed energy theft detection model are illustrated for imbalanced data and high-dimensional data problems. The metrics are described for evaluating the performance of the proposed model.

2.1. VAE-GAN

In this study, VAE-GAN is employed to alleviate the imbalanced data problem. GAN produces samples with fidelity and is unstable during the learning process while VAE produces samples with diversity and is relatively stable during the learning process. By combining a couple of generative models, it can contain the benefits of each model. VAE-GAN may generate samples with high fidelity and diversity as well as be stable in the learning process.

The VAE is usually composed of an encoder and decoder [21]. The encoder maps the input to a latent vector while the decoder reconstructs the latent vector into an approximated input. The encoder and decoder can be represented as follows:

$$z \sim Enc(x) = q_{\phi}(z|x), \tag{1}$$

$$\hat{x} \sim Dec(z) = p_{\theta}(x|z), \tag{2}$$

where x, z, and $\hat{x}$ denote the input, latent vector and approximated input, respectively, and ϕ and θ are parameters of the encoder and decoder models. The term $q_{\phi}(z|x)$ is the approximation of the true posterior $p_{\theta}(x|z)$. The loss function of VAE is represented as the sum of the reconstruction error and a prior regularization term.

$$J_{VAE} = J_{recon} + J_{prior}, \text{with} \tag{3}$$

$$J_{recon} = -\mathbb{E}_{q_{\phi}(z|x)}[\log p_{\theta}(x|z)], \tag{4}$$

$$J_{prior} = D_{KL}(q_{\phi}(x|z) \parallel p_{\theta}(z)), \tag{5}$$

where D_{KL} and $p_\theta(z)$ denote the Kullback-Leibler divergence and the prior distribution of z.

Typically, the GAN model consists of a generator and discriminator [21]. The generator maps the latent vector to data space while the discriminator allocates probability v and probability $1 - v$. The main objective of GAN is to find the discriminator distinguishing between true and generated data and simultaneously the generator adapting to the true data distribution. The loss function of GAN is represented by the binary cross entropy with respect to the generator and discriminator.

$$v = Dis(u) \in [0,1], u = Gen(w), \tag{6}$$

$$J_{GAN} = \log(Dis(u)) + \log(1 - Dis(Gen(w))), \tag{7}$$

where u denotes an actual sample and w is random variable with probability density function $p(w)$.

Although it is possible to obtain synthetic data from random noise without density functions using GAN, it may be advantageous to attain new samples from the generator with specific distributions under imbalance data. Then, a generative model is designed by assigning the decoder of VAE as the generator of GAN. Figure 1 depicts the structure of the VAE-GAN.

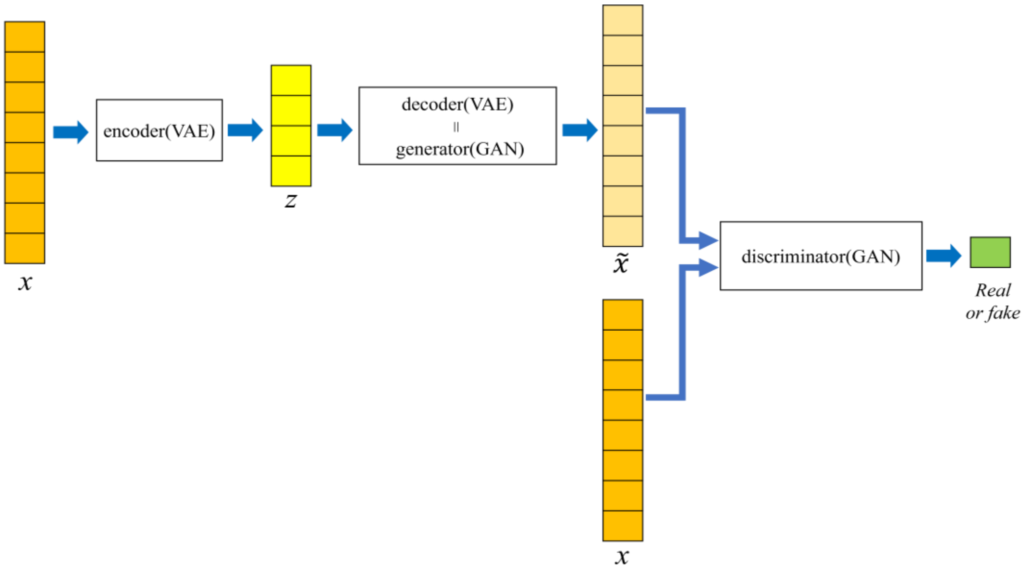

Figure 1. Structure of VAE-GAN.

The loss function of VAE-GAN can be represented as follows [21].

$$J_{VAE-GAN} = J_{prior} + J_{Dis_l} + J_{GAN}, \quad \text{with} \tag{8}$$

$$J_{Dis_l} = -\mathbb{E}_{q(z|x)}[\log p(Dis_l(x)|z)], \tag{9}$$

where $Dis_l(x)$ represents a Gaussian observation model with mean $Dis_l(\tilde{x})$ and identity covariance.

2.2. Energy Theft Detection Model Based on VAE-GAN

A classification model based on VAE-GAN is proposed to detect energy theft in this paper. In practical environments, the energy consumption data of users are collected much more than the energy theft data. The performance of classifiers trained on this imbalanced dataset may be degraded. In order to enhance the performance of classifiers, the balanced dataset is necessary. The balanced dataset can be formed by generating artificial data similar to authentic data. The VAE-GAN is employed to produce the artificial data in this study. The proposed VAE-GAN approach can combine enhanced discrimination capability of the discriminator and higher fidelity of mapping of the encoder by jointly training the encoder, generator and discriminator. From the improved encoder, generator and discriminator capabilities, the proposed VAE-GAN approach can achieve superior performance by generating data with higher diversity by VAE and higher fidelity by GAN compared with other algorithms. And 1D-CNN is utilized as a classifier for detecting energy theft. The architecture of the proposed energy theft detection model is described in Figure 2. The proposed model consists of four steps. The VAE-GAN network is learned by the real energy theft data for generating the synthetic energy theft data. Then, the synthetic energy theft data is generated by the learned VAE-GAN network. And then, the energy theft detector, which is designed as 1D-CNN network, is learned by using the dataset comprised of energy consumption data from real users, real energy theft and synthetic data. Finally, the learned detector is deployed to detect energy theft in the dataset composed of real energy consumption and energy theft data for checking the performance of the detector.

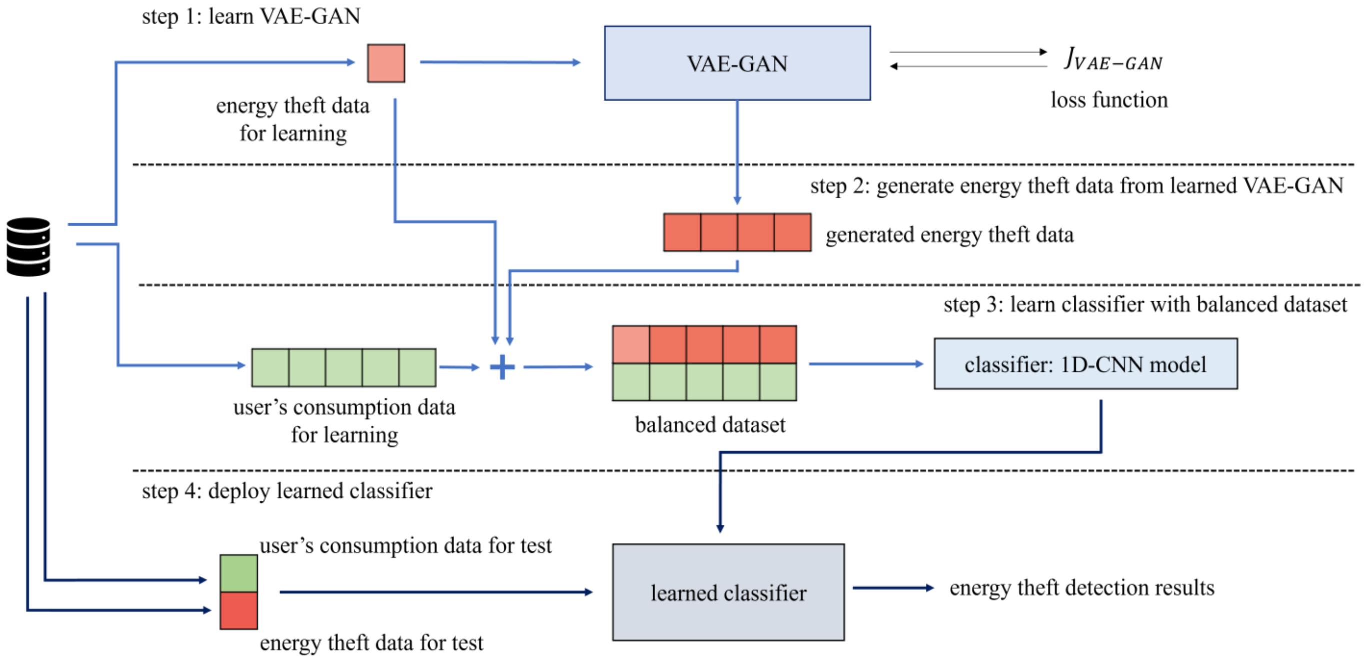
step 1: learn VAE-GAN
VAE-GAN
$J_{VAE-GAN}$
loss function
energy theft data for learning
step 2: generate energy theft data from learned VAE-GAN
generated energy theft data
step 3: learn classifier with balanced dataset
classifier: 1D-CNN model
user's consumption data for learning
balanced dataset
step 4: deploy learned classifier
user's consumption data for test
learned classifier
energy theft detection results
energy theft data for test

Figure 2. Architecture of proposed energy theft detection model.

2.3. Performance Metrics

The performance of the proposed model can be by the following two points: data generation and data classification. In the case of data generation, fidelity and diversity are considered for validating generative models. The fidelity means the quality of generated data from generative models, i.e., how similar the generated data is to real data. The diversity means a variety of generated data from generative models. In other words, it is how close the distribution of the generated data is to the distribution of real data. Considering the fidelity and diversity, the Inception Score (*IS*) and Fréchet Inception Distance (*FID*) are used for evaluating generative models [22]. The *IS* is a method to measure the fidelity of outputs of a generative model. For including meaningful features, data should have a conditional distribution of low entropy. Moreover, a marginal distribution should have high entropy to generate various data from the model. Then, the *IS* is derived as the following equation:

$$IS = exp(\mathbb{E}_{x \sim p_G}(D_{KL}(p(y|x) \parallel p(y))), \tag{10}$$

where x denotes the generated data from a generative model G and y denotes the labels obtained from a pretrained classifier applied x; $p(y)$ and $p(y|x)$ denote a distribution of y

and a conditional distribution of y given x, respectively; $D_{KL}(p \parallel q)$ is the KL-divergence between a distribution p and q. The high score of *IS* indicates both high diversity in data and that the data are meaningful.

The *FID* evaluates the fidelity and diversity of outputs of generative models using the Fréchet distance which is a distance between the real data distribution and the generated data distribution. For the multivariate normal distributions case, the Fréchet distance can be defined by the following:

$$d(D_1, D_2) = \parallel \mu_1 - \mu_2 \parallel^2 + Tr\left(\Sigma_1 + \Sigma_2 - 2\sqrt{\Sigma_1 \Sigma_2}\right), \tag{11}$$

where D_1 and D_2 are marginal distributions of X and Y, respectively; μ_i and Σ_i are mean and covariance matrix of D_i; $Tr(\cdot)$ is trace of a matrix which means the sum of the elements on the main diagonal. In (11), the former term represents the distance between the centers of the two distributions and the latter term describes a metric on the space of all covariance matrices. The *FID* is calculated by the Fréchet distance between the generated data distribution and real data distribution using a feature extractor. Here, it is assumed that the extracted features have a multivariate normal distribution. The *FID* is described as the following equation:

$$FID = \parallel \mu_R - \mu_G \parallel^2 + Tr\left(\Sigma_R + \Sigma_G - 2\sqrt{\Sigma_R \Sigma_G}\right), \tag{12}$$

where μ_R and Σ_R are the mean and covariance matrix of the real data distribution; μ_G and Σ_G are the mean and covariance matrix of the generated data distribution. In this paper, the *IS* and *FID* are employed to evaluate the data augmentation performance of the energy theft detection model based on VAE-GAN.

The energy theft detection problem can be handled as a classification problem. Therefore, the performance metrics of classification problems can be utilized to evaluate the energy theft detection model. In this paper, four performance metrics—*positive predictive value (PPV)*, *true positive rate (TPR)*, *F1-score*, and *Matthews correlation coefficient (MCC)*—are used to evaluate the proposed model [23–25]. The metrics are described in the following equations:

$$PPV = \frac{TP}{TP + FP}, \tag{13}$$

$$TPR = \frac{TP}{TP + FN}, \tag{14}$$

$$F1 - score = \frac{2TP}{2TP + FP + FN}, \tag{15}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}, \tag{16}$$

where *TP* denotes true positive which means that the result of classifying actual energy thieves as energy thieves; *FP* denotes false positive which means that the result of classifying actual normal users as energy thieves; *FN* denotes false negative which means that the result of classifying actual energy thieves as normal users; and *TN* denotes true negative which means that the result of classifying actual normal users as normal users. *PPV* is the ratio of actual energy thieves among the results of classifying as energy thieves. *TPR* is the ratio of results of classifying as energy thieves among actual energy thieves. *F1-score* is a harmonic mean of *PPV* and *TPR*. *MCC*, which is called phi coefficient, is the performance metric used where the dataset is imbalanced. The value of *F1-score* is represented between 0 and 1 while the value of *MCC* is represented between −1 and 1. However, the closer the value of two metrics approaches one, the better the classification performance.

For evaluating the performance depending on the balance ratio of dataset, the balance rate (BR) is defined as follow:

$$BR = \frac{augmented\ abnormal\ data}{normal\ data}. \tag{17}$$

The data augmentation scheme is not applied to learn the detection model at $BR = 0$, while the synthetic data are only used to learn the detection model at $BR = 1$. The results of performance metrics are shown and discussed in the next section.

3. Simulation Results

In this section, the environment is described for simulating the proposed model. Furthermore, the simulation results are shown and analyzed to evaluate the performance of the proposed energy theft model.

3.1. Simulation Environment

3.1.1. Dataset

The energy consumption dataset, which has been released by SGCC (State Grid Corporation of China), is used to simulate over-realistic environment [11]. The dataset consists of 38,757 normal users and 3615 energy thieves within 1035 days (from January 2014 to October 2016). The data collected by meters includes the daily energy consumption histories of residential users. In addition, it is confirmed that normal users and energy thieves have different consumption patterns [17]. It has been well known that the data of abnormal(thief) users tend to fluctuate more drastically in amplitude compared with those of normal users. The fluctuation pattern of energy consumption data is used for classification. The energy thieves constitute roughly 9% of the released dataset. The composition of the released dataset demonstrates that the imbalanced dataset is manifested in actual world. The dataset contains some error values. Therefore, data pre-processing is described in following subsection for addressing this issue.

In the deep learning field, the splitting ratio of the training and testing data has conventionally been used as 8:2 [26]. Consequently, the dataset is empirically divided into training, validation and test datasets with the proportion of 8:1:1 based on trial and error.

3.1.2. Data Pre-Processing

The data pre-processing stage consists of two steps which are the initialization and normalization step. Initially, missing values in dataset are calculated by a pre-processing scheme. And then, the compensated data is normalized by a normalization scheme.

Missing values are often included in energy consumption data by various causes such as unreliable measurement data, the error of smart meters, etc. [27]. In order to recover the error values, the interpolation method is used following the below equation:

$$f(x_i) = \begin{cases} \frac{x_{i-1}+x_{i+1}}{2}, & x_i \in \mathrm{NaN}, \quad x_{i-1}\ and\ x_{i+1} \notin \mathrm{NaN} \\ 0, & x_i \in \mathrm{NaN}, \quad x_{i-1}\ or\ x_{i+1} \notin \mathrm{NaN} \\ x_i, & x_i \notin \mathrm{NaN}, \end{cases}, \tag{18}$$

where x_i is the value in energy consumption data and NaN denotes a set of null value.

Moreover, after handling the missing values, the data are normalized to apply on the neural networks by the MIN-MAX scaler. Due to the nature of deep learning model, when the data have a high range of values then the performance of the deep learning model can be degraded. It can help to the training procedure that the values of the dataset are allocated to the same scale. The values of the energy consumption data are transformed into the range of 0 and 1 by normalization. The MIN-MAX scaler is calculated as follows:

$$f(x_i) = \frac{x_i - min(\boldsymbol{x})}{max(\boldsymbol{x}) - min(\boldsymbol{x})}, \tag{19}$$

where max(x) is the maximum value in x, and min(x) is the minimum value in x.

3.1.3. Hyper-Parameters

Hyper-parameters of training models are presented in Table 1. The early stopping [28] is applied during the training process. The hyper-parameters are optimized by trial and errors.

Table 1. Hyperparameters of training models.

	VAE-GAN	1D-CNN
batch size	100	100
epoch	10,000	1000
learning rate	0.0004	0.002
optimizer	Adam	Nadam

3.1.4. Structure of VAE-GAN Network

The structure of VAE-GAN network is summarized in Table 2. The encoder consists of three convolutional (Conv) layers with a rectified linear unit (ReLU) activation function for mapping input to a latent vector. The generator is composed of three transposed convolutional layers with ReLU activation function and one transposed convolutional (ConvT) layer with sigmoid activation function to reconstruct the input from the latent vector (i.e., it generates a new sample from the latent vector). The discriminator is comprised of three fully connected (FC) layers with leaky ReLU activation function and one fully connected layer with sigmoid activation function for distinguishing real data and fake data. It has been widely known that for the GAN-type approaches, the bounded activation functions such as ReLU and LeakyReLU, have achieved superior performance compared with other activation functions by accelerating the training leading to fast convergence [29]. Therefore, the ReLU and the LeakyReLU were employed in designing both the generator and discriminator.

Table 2. Structure of VAE-GAN.

	Layer Type (with Activation Function)
encoder	(Conv + ReLU) × 3
generator (decoder)	(ConvT + ReLU) × 3 ConvT + Sigmoid
discriminator	(FC + LeakyReLU) × 3 FC + Sigmoid

3.1.5. Structure of Detector

The specification of detector is detailed in Table 3. The 1D-CNN layers are used for dealing with energy consumption data which are one-dimensional data type. The models designed with 1D-CNN have relatively lower computational complexity and shallower architectures than other deep learning models [30]. Then, the models may be fast for training due to low computational requirements. Scaled exponential linear unit (SELU) is selected as activation function of each layer, except of output layer. SELU relieves the vanishing gradient problem of deep neural networks. Furthermore, SELU learns faster and better than other activation functions [31]. The activation function of output layer is used as sigmoid function for binary classification.

Table 3. Specification of detector.

Layer	Type (with Activation Function)	# of Neurons	Size of Kernel	Stride	# of Parameters
1	Conv1D + SELU	256	7	1	2048
2	AvgPool	-	4	3	-
3	Conv1D + SELU	128	7	1	229,504
4	AvgPool	-	4	3	-
5	Conv1D + SELU	64	7	1	57,408
6	AvgPool	-	4	3	-
7	Conv1D + SELU	32	7	1	14,368
8	AvgPool	-	4	3	-
9	GlobalAvgPool	-	-	-	-
10	Linear + SELU	512	-	-	16,896
11	Linear + SELU	256	-	-	131,328
12	Linear + SELU	128	-	-	32,896
13	Linear + SELU	64	-	-	8256
14	Linear + SELU	32	-	-	2080
15	Output (Sigmoid)	1	-	-	33

3.2. Performance Analysis

The proposed model was compared with three other models and the baseline model to evaluate the performance. The data augmentation schemes used in comparison models are VAE, GAN and SMOTE. By comparing three generative model-based models with baseline model, the influence of applying generative model can be analyzed in energy theft detection field. The effect of combining VAE and GAN in the proposed model can be confirmed by comparing with the models applied as VAE and GAN, respectively. By comparing the proposed model with the model applied SMOTE, which is the sampling method, the applicability of generative model and the ability of the proposed model to alleviate the overfitting problem can be confirmed. The *IS* and *FID* are displayed for validating the performance in point of data generation in Table 4.

Table 4. *IS* and *FID* of proposed and conventional models.

	VAE-GAN	VAE	GAN	SMOTE
IS	**1.13**	1.01	1.08	1.00
FID	**13.99**	14.13	14.21	14.33

It is shown that the *IS* of the VAE-GAN model is the highest and the *FID* is the lowest in Table 4. The higher the *IS* and the lower the *FID*, the better the generation performance of the model. Therefore, it can be demonstrated that the generation performance of the proposed model is relatively better than the other models. The *IS* and *FID* of the VAE model is lower than the GAN model. It may be analyzed that the VAE model generates more diverse data than the GAN model while the GAN model generates better quality data than the VAE model. It can be found the reason in the difference of learning objective. VAE aims to learn the distribution of original data. Consequently, the various data can be generated from the learned distribution. In contrast, the learning objective of GAN is to create data like original data. Accordingly, generated data from GAN have superior quality than diversity. It is shown in simulation results that data quality and diversity may be improved compared with VAE and GAN by combining VAE and GAN. In the instance of SMOTE, the *IS* is lowest and the *FID* is highest among them. It can be said that data with low quality and diversity is generated due to overfitting.

In Table 5, it is shown the simulation results in terms of data classification varying data augmentation schemes. The confusion matrices of classification result on the proposed model and conventional models are depicted in Figure 3.

Table 5. Classification performance metrics on proposed and conventional models.

	VAE-GAN	VAE	GAN	SMOTE	Baseline
PPV	**0.925**	0.754	0.733	0.783	0.661
TPR	**0.909**	0.803	0.693	0.729	0.65
F1-score	**0.905**	0.76	0.622	0.677	0.567
MCC	**0.834**	0.62	0.446	0.517	0.352

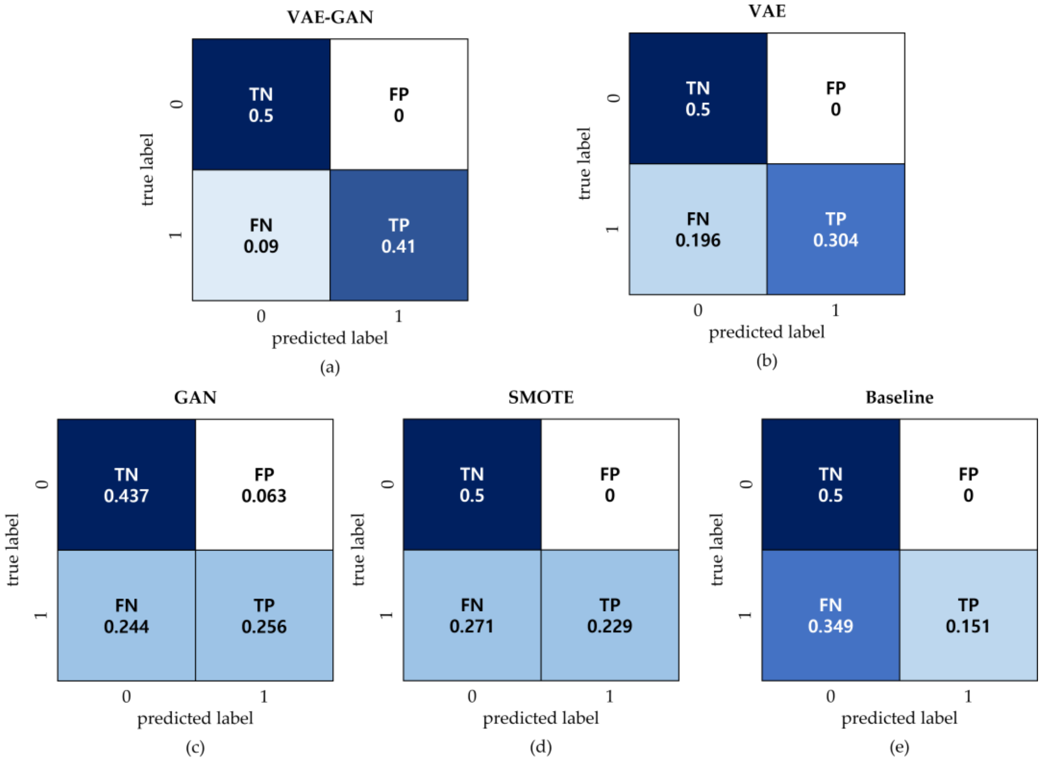

Figure 3. Confusion matrices of classification result on (**a**) VAE-GAN; (**b**) VAE; (**c**) GAN; (**d**) SMOTE; (**e**) Baseline.

From Table 5 and Figure 3, it can be confirmed that all performance metrics of the proposed model are higher than those of other schemes. *PPV* is related to the number of correctly classified energy thieves among all predicted energy thieves. *TPR* is related to the accurately classified energy thieves among all actual energy thieves. It is indicated that the energy thieves are effectively classified by the proposed model compared to other models. *F1-score* and *MCC* are related to the number of overall accurately classified data. Therefore, the classification performance is improved by applying the VAE-GAN.

According to Tables 4 and 5, the proposed model was very advantageous in detecting theft data compared with other models. It has been generally known that performance of deep learning algorithms depends on how component blocks such as encoder, generator and discriminator are efficiently and optimally designed. The proposed VAE-GAN approach combined enhanced discrimination capability of the discriminator and higher fidelity of mapping of the encoder by jointly training the encoder, generator and discriminator. From the improved encoder, generator and discriminator capabilities, the proposed VAE-GAN approach achieved superior performance by generating data with higher diversity by VAE and higher fidelity by GAN compared with other algorithms. The performance of the energy theft detection model can be improved by utilizing the balanced dataset by the synthetic data from VAE-GAN in the training process.

Furthermore, the simulation results are shown on balance rate in Figure 4. The VAE-GAN is not applied to the proposed model at *BR* = 0, while the generation data are only used to learn the model based on the VAE-GAN at *BR* = 1. In Figure 4, all performance metrics have an increasing trend depending on *BR*. The tendency can be described that the new types of data with similar features to energy theft data augmented by VAE-GAN. The more detector is learned by new types of data, the better its performance. Then, it can be confirmed that the proposed model is more effective in extreme imbalance situations than other situations from Figure 4.

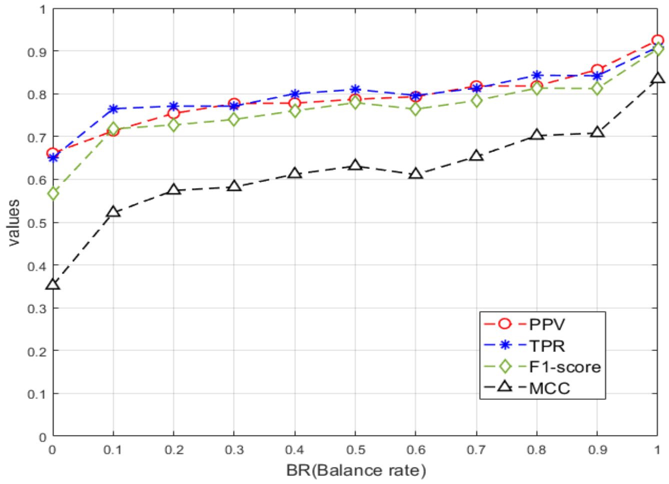

Figure 4. Performance metrics values over BR.

In Table 6, the number of weights of the employed generative models is described.

Table 6. Model complexity for various generative models.

Generative Models	Model Complexity (the Number of Weights)
VAE	134,165
GAN	1,248,642
VAE-GAN	5,689,813

In this paper, the model complexity of the employ generative models is analyzed by the number of weights. instead of time complexity [32]. The model complexity of the generative models can be represented as $O(|W|\log(|W|))$, where $O(\cdot)$ denotes big O notation; $|W|$ is the cardinality of total number of weights. It was confirmed that the complexity is increased by combining a couple of generative models.

4. Conclusions and Discussions

In this paper, an energy theft detection model based on VAE-GAN was proposed to overcome the imbalanced data problem and simulated under the practical dataset environment. It was demonstrated from simulation results that the proposed model outperformed various existing generative models in terms of both data generation and classification. The generated data by VAE-GAN showed relatively high quality and diversity compared to the other schemes. This was achieved by slightly improving the data generation performance metrics while considerably doing the classification performance metrics. Moreover, it was demonstrated that the proposed model is helpful in applying to extreme environments by investigating the performance of a detector depending on the balance rate. The proposed model will be applied to anomaly detection problems with an imbalanced dataset such as fault diagnosis of electrical devices and machines.

Despite the result that the imbalanced data problem was effectively alleviated by generating data with fidelity and diversity from VAE-GAN, the complexity of the proposed model was increased due to employing GAN combined with VAE in the augmentation process. The deep learning-based energy theft methods are relatively complicated due to the learning process and structure of detector. The employment of VAE-GAN with high complexity will inevitably increase its complexity. Consequently, the deep learning model with low complexity should be developed in the augmentation process.

The deep learning-based model like VAE-GAN may be interpreted as a black box. In the augmentation process, it is actually challenging to figure out which features of the original data are most reflected into the generated synthetic data. The identification of the features in the augmentation process can help generate more meaningful synthetic data.

From the aforementioned limitations of this study, future research will concentrate on a data augmentation method with low complexity producing data with fidelity and diversity. In addition, research will be conducted to design an efficient detector to minimize the increased complexity for solving imbalanced data problem.

Author Contributions: Conceptualization, Y.S. and J.L.; methodology Y.S. and J.L.; software, J.S.; validation, S.L. and C.K.; investigation, S.L. and C.K.; writing—original draft preparation, Y.S. and J.K.; writing—review and editing, S.K. and J.K.; visualization, S.K. and J.S.; supervision, J.K.; project administration, J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Publicly available datasets were analyzed in this study. These datasets can be found here: http://www.sgcc.com.cn/ (accessed on 23 February 2022).

Acknowledgments: This work supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2022R1F1A1064414); in part by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2021-0-00892-002, Research on advanced physical layer technologies of LEO (low-earth orbit) satellite communication systems for ubiquitous intelligence in space); and in part by Kwangwoon University in 2022.

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

References

1. Shah, A.L.; Mesbah, W.; Al-Awami, A.T. An Algorithm for Accurate Detection and Correction of Technical and Nontechnical Losses Using Smart Metering. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 8809–8820. [CrossRef]
2. Nagi, J.; Yap, K.S.; Tiong, S.K.; Ahmed, S.K.; Mohamad, M. Nontechnical Loss Detection for Metered Customers in Power Utility Using Support Vector Machines. *IEEE Trans. Power Del.* **2010**, *25*, 1162–1171. [CrossRef]
3. Yan, Z.; Wen, H. Performance Analysis of Electricity Theft Detection for The Smart Grid: An Overview. *IEEE Trans. Instrum. Meas.* **2021**, *71*, 1–28. [CrossRef]
4. Korba, A.A.; Karabadji, N.E.I. Smart Grid Energy Fraud Detection Using SVM. In Proceedings of the 2019 International Conference on Networking Advanced Systems, Annaba, Algeria, 26–27 June 2019.
5. Henriques, H.O.; Barbero, A.P.L.; Ribeiro, R.M.; Fortes, M.Z.; Zanco, W.; Xavier, O.S.; Amorim, R.M. Development of Adapted Ammeter for Fraud Detection in Low-voltage Installations. *Measurement* **2014**, *56*, 1–7. [CrossRef]
6. Engelbrecht, J.; Hancke, G.P.; Osifeko, M.O. Design and Implementation of An Electrical Tamper Detection System. In Proceedings of the 45th Annual Conference on the IEEE Industrial Electronics Society, Lisbon, Portugal, 14–17 October 2019.
7. Kim, J.Y.; Hwang, Y.M.; Sun, Y.G.; Sim, I.; Kim, D.I.; Wang, X. Detection for Non-technical Loss by Smart Energy Theft with Intermediate Monitor Meter in Smart Grid. *IEEE Access* **2019**, *7*, 129043–129053. [CrossRef]
8. Weckx, S.; Gonzalez, C.; Tant, J.; De Rybel, T.; Driesen, J. Parameter Identification of Unknown Radial Grids for Theft Detection. In Proceedings of the 2012 3rd IEEE PES Innovative Smart Grid Technologies Europe, Berlin, Germany, 14–17 October 2012.
9. Cardenas, A.A.; Amin, S.; Schwartz, G.; Dong, R.; Sastry, S. A Game Theory Model for Electricity Theft Detection and Privacy-aware Control in AMI Systems. In Proceedings of the 50th Annual Allerton Conferences, Monticello, IL, USA, 1–5 October 2012.
10. Wei, L.; Sundararajan, A.; Sarwat, A.I.; Biswas, S.; Ibrahim, E. A Distributed Intelligent Framework for Electricity Theft Detection Using Benford's Law and Stackelberg game. In Proceedings of the 2017 Resilience Week, Wilmington, DE, USA, 18-22 September 2017.
11. Zheng, Z.; Yang, Y.; Niu, X.; Dai, H.-N.; Zhou, Y. Wide and Deep Convolutional Neural Networks for Electricity-theft Detection to Secure Smart Grids. *IEEE Trans. Ind. Informat.* **2018**, *14*, 1606–1615. [CrossRef]
12. Zheng, K.; Chen, Q.; Wang, Y.; Kang, C.; Xia, Q. A Novel Combined Data-driven Approach for Electricity Theft Detection. *IEEE Trans. Ind. Informat.* **2019**, *15*, 1809–1819. [CrossRef]
13. Wang, X.; Yang, I.; Ahn, S.-H. Sample Efficient Home Power Anomaly Detection in Real Time Using Semi-supervised Learning. *IEEE Access* **2019**, *7*, 139712–139725. [CrossRef]
14. Aldegheishem, A.; Anwar, M.; Javaid, N.; Alrajeh, N.; Shafiq, M.; Ahmed, H. Towards Sustainable Energy Efficiency With Intelligent Electricity Theft Detection in Smart Grids Emphasising Enhanced Neural Networks. *IEEE Access* **2021**, *9*, 25036–25061. [CrossRef]

15. Lee, J.; Sun, Y.G.; Sim, I.; Kim, S.H.; Kim, D.I.; Kim, J.Y. Non-Technical Loss Detection Using Deep Reinforcement Learning for Feature Cost Efficiency and Imbalanced Dataset. *IEEE Access* **2022**, *10*, 27084–27095. [CrossRef]
16. Aslam, Z.; Javaid, N.; Ahmad, A.; Ahmed, A.; Gulfam, S.M. A Combined Deep Learning and Ensemble Learning Methodology to Avoid Electricity Theft in Smart Grids. *Energies* **2020**, *13*, 5599. [CrossRef]
17. 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]
18. Yan, Z.; Wen, H. Electricity Theft Detection Base on Extreme Gradient Boosting in AMI. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–9. [CrossRef]
19. 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]
20. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kengelmeyer, W.P. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* **2022**, *16*, 321–357. [CrossRef]
21. Larsen, A.B.L.; Sonderby, S.K.; Larochelle, H.; Winther, O. Autoencoding Beyond Pixels Using A Learned Similarity Metric. In Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016.
22. Benny, Y.; Galanti, T.; Benaim, S.; Wolf, L. Evaluation Metrics for Conditional Image Generation. *Int. J. Comput. Vis.* **2021**, *129*, 1712–1731. [CrossRef]
23. Xie, X.; Jiang, W.; Guo, J. Research on Rockburst Prediction Classification Based on GA-XGB Model. *IEEE Access* **2021**, *9*, 83993–84020. [CrossRef]
24. Chicco, D.; Starovoitov, V.; Jurman, G. The Benefits of The Matthews Correlation Coefficient (MCC) over The Diagnostic Odds Ratio (DOR) in Binary Classification Assessment. *IEEE Access* **2021**, *9*, 47112–47124. [CrossRef]
25. Albahar, M.A. Skin Lesion Classification Using Convolutional Neural Network with Novel Regularizer. *IEEE Access* **2019**, *7*, 38306–38313. [CrossRef]
26. Joseph, V.R. Optimal Ratio for Data Splitting. *Stat. Anal. Data Min.* **2022**, *15*, 531–538. [CrossRef]
27. Genes, C.; Esnaola, I.S.; Perlaza, M.; Ochoa, L.F.; Coca, D. Recovering Missing Data via Matrix Completion in Electricity Distribution Systems. In Proceedings of the IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications, Edinburgh, UK, 3–6 July 2016.
28. Raskutti, G.; Wainwright, M.J.; Yu, B. Early Stopping for Non-parametric Regression: An Optimal Data-dependent Stopping Rule. In Proceedings of the 49th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, USA, 28–30 September 2011.
29. Radford, A.; Metz, L.; Chintala, S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. In Proceedings of the 6th International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.
30. Kiranyaz, S.; Avci, O.; Abdeljaber, O.; Ince, T.; Gabbouj, M.; Inman, D.J. 1D Convolutional Neural Networks and Applications: A Survey. *Mech. Syst. Signal Process.* **2021**, *151*, 107398. [CrossRef]
31. Klambauer, G.; Unterthiner, T.; Mayr, A.; Hochreiter, S. Self-normalizing Neural Networks. In Proceedings of the 31st International Conference on Neural Information Processing System, Long Beach, CA, USA, 4–9 December 2017.
32. Sontag, E.D. VC Dimension of Neural Networks. *NATO ASI Ser. F Comput. Syst. Sci.* **1998**, *168*, 69–96.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.