IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 3 July 2022, accepted 12 July 2022, date of publication 18 July 2022, date of current version 11 August 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3191685

APPLIED RESEARCH

Nontechnical Loss Detection With Duffing–Holmes Self-Synchronization Dynamic Errors and 1D CNN-Based Multilayer Classifier in a Smart Grid

CHIA-HUNG LIN[1], FENG-CHANG GU[1], JIAN-XING WU[1], AND CHAO-LIN KUO[2]
[1]Department of Electrical Engineering, National Chin-Yi University of Technology, Taichung 41170, Taiwan
[2]Department of Maritime Information and Technology, National Kaohsiung University of Science and Technology, Kaohsiung 80543, Taiwan
Corresponding authors: Chia-Hung Lin (eechl53@gmail.com) and Chao-Lin Kuo (clkuo@nkust.edu.tw)

ABSTRACT The combined Duffing–Holmes (D–H)–based quantizer and one-dimensional (1D) convolutional neural network (CNN)-based multilayer classifier were applied to perform the nontechnical loss (NTL's) (electricity fraud) feature quantification, feature extraction, and classification tasks to analyze electricity consumption data and to identify either normal or abnormal consumption patterns for NTL detection. The metering data is gathered every 15 minutes and a 3-hour screening window is used to distinguish between the normal conditions and likely NTL events or power outage events. The D–H-based quantizer in the feature quantification layer may quantify the different levels among three events for preliminary screening differences using D–H self-synchronization dynamic errors. In the feature extraction layer, two 1D convolutional-pooling processes are used to extract 1D key feature signals to enhance the distinguished levels for further classification applications. The gray relational analysis (GRA)-based multilayer network is trained as a classifier. In the classification layer, to identify electricity fraud events. The proposed method is verified and validated using simulation data and the electricity fraud attack model. The correlation coefficient and unitary averaged changed intensity index are applied in correlation analysis to discover apparent abnormality between historical consumption and metering consumption patterns within the short-time monitoring. The D-H-based quantizer and 1D CNN-based classifier then work together to accomplish classification tasks on the time-series metering data. The experimental results show that the suggested classifier model demonstrates promising performance and efficiency compared to the traditional multilayer classifier in feature extraction, training, and recall processing, and accurate classification.

INDEX TERMS Nontechnical loss, Duffing–Holmes (D-H) based quantizer, one-dimension convolutional neural network (CNN), short-time screening window, gray relational analysis (GRA).

I. INTRODUCTION

In a power grid, electricity device operations will naturally produce energy dissipation and involve large electricity operational losses, such as the winding and core losses of generators and transformers, conductor resistance losses (feeder/transmission line, dielectric losses (in cables, capacitors, and transformers), skin effect losses (electro-magnetic fields), and harmonic distortion. The above- described losses are so-called technical losses (TLs), which may be caused by current flowing through the transmission and distribution grids to different customers [1], [2]. They also have different load profiles, peak demands, and peak durations. These inherent losses can be measured and estimated depending on the regular total loads and total electricity billing charge. Nontechnical losses (NTLs), meanwhile, are caused by bypassing meter, meter reading and recording errors, meter tampering, wrong transformer ratios in the meter, electricity fraud,

The associate editor coordinating the review of this manuscript and approving it for publication was Ziang Zhang.

human errors, unmetered supplies, and embedded generation [3]–[5], which result in energy dissipation but are not accounted for in billing. Electricity fraud occupies a major proportion of NTLs, such as bypassing (illegal connections), tampering the reading of the measurement meter, and evading payment, and uses electricity from the power utility without a valid obligation to alter its measurements. NTLs are not limited to developing countries, such as India and Brazil (25% and 16% of their total electric supply, respectively) but also occur in developed countries [5], [6]. Therefore, power utilities need to handle information related to customers' energy consumption within metering and billing systems [7], [8], such as inspecting of customers with suspicious load profiles.

Distribution system operators use smart meters (SMs) to detect potential fraud users while screening the usage of load profiles for identified NTLs. (SMs) are gradually being implemented to monitor and record consumers' electrical consumption and are linked via wired or wireless secure communication systems. The customers' metering data can be transmitted to the power utility, including the usage of load profiles, illegal abstraction detection, and power outage detection, which can then be used to detect the fraud through the customers' consumption history data. The advanced metering infrastructure (AMI) [9]–[11] can collect operation information by using the upgraded SMs in state-based detection [12], such as voltages, currents, powers, and outages, to determine the electricity feedback to the power grid from the customers' premises. Both power users and grid operators can learn about frequent electricity usage and alter their consumption based on the dynamic electricity pricing. AMI can also be used to regulate electrical appliances and devices for demand responses, limit the maximum electricity consumption, and disconnect or reconnect the power supply while a fault occurs or clears. Furthermore, unauthorized use lowers the quality of the power supply and increases energy generation, resulting in increased generation costs and electricity bills for power utilities and users, respectively. Thus, with the integration of SM meterring data into the AMI infrastructure, the integration of SMs' meterring data can assist power utilities in detecting unlawful usage and electricity fraud. The meter data management system (MDMS) performs validation, estimates, and editing tasks [11], [13]–[15] to screen past recorder data as predefined reference values using detection rules to distinguish normal consumption from meter tampering or theft users [11], [16], [17]. Thus, in this study, unauthorized users can be recognized based on good matches between the pass record data and the current metering data, as well as significant changes in on-line consumption monitoring. The illegal users can then be identified. Hence, we intend to create a correlation analysis for preliminarily to screen suspicious consumptions within the short-time monitoring; and to integrate a quantizer and a classifier to automatically quantify and identify the normality and electricity fraud.

Artificial intelligence (AI) and data analysis approaches have been identified to identify clients with irregular, aberrant, and fraudulent usage in order to detect NTLs automatically. Machine learning (ML) and deep learning (DP) methods include artificial neural networks (ANNs), support vector machines (SVMs), decision trees, random forests, fuzzy inference systems (FIS), and (CNN)-based classifiers [13], [18]–[22]. The ML-based algorithms [5], [18]–[21] are trained using the gathered data and used to categorize the incoming consumers' data, whether they are yes-or-no-fraud users in a smart grid. These consumption datasets require large amounts of knowledge of the dataset parameters and distribution parameters, including the historical dataset, mean, variance, and data variability. Such as tampering with metering data and bypassing meters, which use smart technology in metered measurement, estimate these distribution parameters from sampling data with smart data.

Power assumption patterns in time-series stream data will show unusual changes in frequency and amplitude during an electricity fraud attack. Previous works [23]–[25] presented fast Fourier transform and discrete wavelet transform to divide the significant frequency-based characteristics into multiple frequency components, and then combined AI-based classifiers to automatically detect electricity fraud in a smart grid. Frequency-based features can be simply employed to analyze fraud occurrences. However, this technique requires allocating greater resolution in time with several low-pass and high-pass decomposition filters in the feature extraction operations [26]. Then, by using the chosen wavelet coefficients, such as absolute magnitudes and standard deviations of wavelet coefficients, are selected to train the supervised ML-based method for NTL detection. The selection of the sampling windows, types of wavelets, and wavelet coefficients to deal with the uncertainty signals in short-term or long-term monitoring for electricity fraud detection is a difficult task for time-varying, transit, and irregular (randomly) signal analysis. Furthermore, the supervised ML-based approaches necessitate user involvement to extract the feature patterns and label the classes in order to train the classifier [27].

The classic approaches mentioned above do not employ the quantizer to extract features and determine the appropriate feature parameters for boosting classification accuracy [20], [21], [29] based methods, in contrast to the ML, the DP-based methods, The DP-based methods, in contrast to the ML-based methods, such as CNN and a long-short-term memory (LSTM) model (CNN–LSTM), GoogLeNet and gated recurrent unit (GRU) model, wide and deep CNN, and CNN–GRU model [20], [21], [27], [28], use the multi convolutional–pooling layers and a classification layer (fully connected layer) to learn the power consumption data in order to create a classifier that can detect electricity fraud, including autonomous end-to-end feature extraction, noise removal, and classification tasks. Their models can extract associated feature patterns from customers' consumption profiles using multi convolutional-pooling methods, which eliminates the need for more human participation and improves the classification accuracy rate. These high-dimensionality features

will result in increased computational complexity, memory storage requirements, and poorer generalization ability for the high-dimensional time-series dataset [30]–[33]. To study 1D signals, the 1D CNN can perform both feature extraction and classification tasks; the 1D convolutional process, which uses linear weighted sums, and the 1D pooling process, can minimize computational complexity for feature pattern dimension [33], [34]. We plan to create a 1D CNN-based classifier to cope with short-term sequences of consumption patterns in order to automatically detect NTL events for real-time 1D signal processing.

In this study, we will apply a simplified model that It consists of a time-domain quantizer and a classifier for automatic NTL detection. The Duffing–Holmes (D–H) system [34]–[36] is used for signal preprocessing in 1D signal analysis, and it consists of a master system (MS) and a slave system (SS) to extract the differences between historical consumption and incoming metering consumption. A discrete D–H system is built as a time-domain quantizer on the basis of preliminarily quantifying the various stages of fraud activities. As a result, the suggested D–H quantizer may identify likely NTL events and indicates the discrepancies in the signal preprocessing layer based on a data-driven method [37]. Multi 1D convolutional-pooling techniques are employed in the convolutional-pooling layer processes to enrich and extract the feature signals, while also protecting essential characteristics from self-synchronization dynamic faults (SSDEs). The D-H quantizer and the convolutional-pooling method can quantify the amounts of suspected fraud activities based on daily, weekly, or monthly metering data and the measurement consumption profile. Thus, by comparing current metering consumptions to historical consumptions, fraud consumptions can be scaled. The scaled feature patterns are supplied into a gray relational analysis (GRA)-based network [38]–[40] at the classification layer to detect fraud activity for each 15-minute interval in a 1-hour period and subsequently identify possible fraud activities.

The remainder of this paper is organized as follows: Section 2 addresses the methodology for fraud attack model design, D-H-based quantizer in signal preprocessing layer, and multilayer 1D CNN design. In Sections 3 and 4, simulation results, discussion, and conclusion are given to show the efficiency of the proposed method for fraud activities screening.

II. METHODOLOGY

A. ELECTRICITY FRAUD ATTACK MODEL DESIGN

In this study, we apply an electricity fraud attack model to simulate fraud activities in a smart grid, including practical power consumption, p_t, at t timing slot, NTL ($\alpha_t + \beta_t + \gamma_t$), and TL($\delta_t$), and can be defined as follows [41]

$$P_t = \frac{1}{P_{\max}}(p_t + \alpha_t + \beta_t + \gamma_t + \delta_t), P_{max} = max(P_t) \quad (1)$$

where P_t is the total power consumption at t timing slot, $t = 0, 1/4, 1/2, \ldots, 24$ hours, which are supplied by the utility companies, and P_{max} is the maximum power demand, P_t can be normalized in per-unit value as a defined base unit quantity; parameter α_t is the reduced metering data caused by meter tampering; β_t is power loss from faulty meters in the grid; γ_t is the energy stolen by illegal bypassing at t timing slot; and δ_t is the TLs, which is produced by power dissipation, such as winding and core losses, conductor resistance losses, leakage current, and poor quality of equipment. Among these parameters, if $\alpha_t + \gamma_t > 0$, then the fraud activity exists at t timing slot; $\beta_t < 0$ means the faulty meters' report is greater than the consumers' consumptions; $\beta_t > 0$ indicates that the meters' report is less than the consumers' consumptions; TL δ_t is proportional to the P_t, as $\delta_t = \varepsilon P_t$, ε is the proportion coefficient [42]. Hence, we can use the electricity fraud attack model to simulate meter tampering, illegal meter bypassing, and faulty metering in a power grid. On the basis of daily metering data, to detect fraud activity, the consumption patterns are fed into the 1D CNN-based classifier to study automatic fraud activity screening at each 1/4-hour timing slot in a 2-hour interval (96 timing slots).

B. D-H-BASED QUANTIZER DESIGN

On the basis of second-order D-H differential equation [33]–[35], the D-H system can be reduced from a high-order to a one-order nonautonomous equation and can be expressed as

$$\dot{x} = y \quad (2)$$

$$\dot{y} = -\delta y + x - x^3 + f\sin(\omega t) \quad (3)$$

where δ is the damping parameter, f is the excitation amplitude, and ω is the excitation frequency. With the excitation amplitude, f, and frequency, ω, the D-H system can produce the dynamic chaotic phenomenon using the external influence term, "$f\sin(\omega t)$." Considering unforced condition, as excitation amplitude $f = 0$, and let state variables $x_1 = x$, and $x_2 = y$ in the D-H system, Equations (2) and (3) can be linearized as follows:

$$\dot{x}_1 = x_2 \quad (4)$$

$$\dot{x}_2 = (1 - x_1^2)x_1 - \delta x_2 \quad (5)$$

Equations (4) and (5) can also be represented in matrix form, as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 - x_1^2 & -\delta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (6)$$

where the damping parameter, $\delta \in [0, 1]$, can control the chaotic oscillations and the dynamic trajectories within a specific boundary region. The system (6) can be implemented in a simple analog electrical circuit and a digital computer [32], [40]. To screen fraud activities in a digital computer, as shown in Figure 1, we can design a SSDE system to track the SSDE between the historical consumption and incoming metering consumption patterns, including the MS and SS, with the following construction:

$$\bullet \text{MS}: \begin{bmatrix} \dot{x}_{1m} \\ \dot{x}_{2m} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 - x_{1m}^2 & -\delta \end{bmatrix} \begin{bmatrix} x_{1m} \\ x_{2m} \end{bmatrix} \quad (7)$$

$$\bullet \text{SS}: \begin{bmatrix} \dot{x}_{1s} \\ \dot{x}_{2s} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 - x_{1s}^2 & -\delta \end{bmatrix} \begin{bmatrix} x_{1s} \\ x_{2s} \end{bmatrix} \tag{8}$$

where $x_{1\text{m}}$ and $x_{2\text{m}}$ are the state variables for the MS and x_{1s}, respectively, and x_{2s} are the state variables for the SS. The variables, $x_{1\text{m}}$ and $x_{2\text{m}}$, are the historical consumptions at t timing slot, and x_{1s} and x_{2s} are incoming metering consumptions at t timing slot, respectively. The incoming metering consumptions can be produced by Equation (1). To quantify the levels of fraud activities, the error variables can be expressed as $e_1 = x_{1\text{m}} - x_{1s}$ and $e_2 = x_{2\text{m}} - x_{2s}$. Thus, the SSDE system can be expressed as follows [34]:

$$\begin{bmatrix} \dot{e}_1 \\ \dot{e}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ (1 - x_{1m}^2) - (1 - x_{1s}^2) & -\delta \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \tag{9}$$

When the MS and the SS receive the incoming metering consumptions, the dynamic errors can be produced by the self-synchronization tracking processes between the MS and SS. The damping parameter, δ, $\delta \in [0, 1]$, can control the chaotic trajectories. For digital implementation in computing applications, Equation (9) can be modified as a discrete dynamic error system with two discrete error variables, that is, e_1 and e_2, and expressed as follows [14], [34]:

$$\begin{bmatrix} \Phi_1[i] \\ \Phi_2[i] \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ (1 - (x_m[i])^2) - (1 - (x_s[i])^2) & -\delta \end{bmatrix} \begin{bmatrix} e_1[i] \\ e_2[i] \end{bmatrix} \tag{10}$$

$$\begin{bmatrix} e_1[i] \\ e_2[i] \end{bmatrix} = \begin{bmatrix} x_m[i] - x_s[i] \\ x_m[i+1] - x_s[i+1] \end{bmatrix} \tag{11}$$

where the $x_m[i] = x_{1\text{m}}[i]$ and $x_m[i+1] = x_{2\text{m}}[i]$ are the historical consumption data at ith timing slot, $i = 1, 2, 3, \ldots, N-1$, with N denoting the total number of timing slots in each screening window ($N = 12$ in this study); $x_s[i] = x_{1s}[i]$ and $x_s[i+1] = x_{2s}[i]$ are the incoming metering consumption data at ith timing slot. The chaotic variations can be quantized by two error variables, that is, $e_1[i]$ and $e_2[i]$. For the dynamic error scatter diagram, as SSDE Φ_1 versus SSDE Φ_2, the dynamic error patterns can be used to evaluate the levels of fraud activities.

In a distribution grid, we propose a D-H quantizer to quantify the NTL levels, including meter tampering, illegal meter bypassing, and faulty metering. Hence, SMs are used to install to monitor the customers' consumptions of all electric branches. The measured data can be converted into digital data with the analog-to-digital converter and microcontroller unit, and then the power consumptions are transmitted to the MDMS via wired or wireless communication system. The proposed D-H quantizer can be employed to analyze the metering consumption in an AMI. In this study, we will use correlation analysis and unified averaged changed intensity (UACI) to preliminarily estimate the differences between the historical consumption and the metering consumption patterns. As seen in Figure 2, the correlation analysis indicates the correlation coefficient (CC) will decrease ($R^2 \leq 0.98$), which indicates some differences between two consumption patterns. In addition, the UACI index can be used to evaluate the different levels as follows [43], [44]:

$$UACI = \frac{\sum_{i=1}^{N} |P_i - P_{Hi}|}{N} \times 100\%, i = 1, 2, 3, \ldots, N \tag{12}$$

where P_i is the metering consumption at ith timing slot; P_{Hi} is the historical consumption; and N is the data length of screening window. Hence, a higher $UACI$ index (the larger, the better) can preliminarily estimate possible fraud activity. In each screening window, while the $UACI$ index gradually increases ($UACI \geq 5.00\%$), the D-H quantizer will quantify the different levels with the specific screening window to further identify normal or abnormal power consumption.

C. MULTILAYER 1D CONVOLUTIONAL NEURAL NETWORK DESIGN

To deal with 1D SSDE signals from the D-H-based quantizer, the conventional 2D CNN can be modified as a 1D CNN to enhance and extract the feature signals, which can simplify the computational complexity. The 1D CNN training process is faster than 2D CNN. As seen in Figure 1, its overall configuration consists of two 1D convolutional operators and two pooling operators in the convolutional-pooling layer and GRA-based classifier in the classification layer. Each layer can be summarized as follows:

1) 1D CONVOLUTIONAL AND 1D POOLING OPERATOR

In the convolutional pooling layer, we use two 1D digital convolution operators to analyze the dynamic error signals, Φ_1 and Φ_2, for enhancing and extracting the specific different level between the historical consumption and the metering consumption patterns. The discrete Gaussian window as a feature extraction filter can slide over two dynamic error signals and then produce the feature signals. Its discrete-time convolutional operation can be expressed as $X_e[i] = \Phi_e[i] * H[j]$, $e = 1, 2$, (where $*$ is the convolution operator) [30], [31]

$$X_e[i] = \sum_{j=0}^{M-1} H[j]\Phi_e[i-j] \tag{13}$$

$$H[j] = \exp[\frac{-1}{2}(\frac{j-1}{\sigma})^2] \tag{14}$$

where $X_e[i]$ is the finite 1D convolution summation with $i = 0, 1, 2, 3, \ldots, N-1$, which is an $N+M-1$ point signal running from 0 to $N + M-2$; N is the data length of the incoming metering consumption; M is the data length of discrete Gaussian window ($N = 12$ and $M = 5$ for short-term monitoring [three hours] in this study); $H[j]$ is a discrete Gaussian-based filter to extract the feature signals from two dynamic error signals, Φ_e; and σ is the standard deviation. The filter window can cover each hour's metering consumption. In this study, the Gaussian-based filter's sliding stride is set to 1 to perform the convolutional operation. Hence, $X_e[i]$ can be obtained by summing the multiplications of the finite signal, $\Phi_e[i-j]$, and the weighted values of $H[j]$.

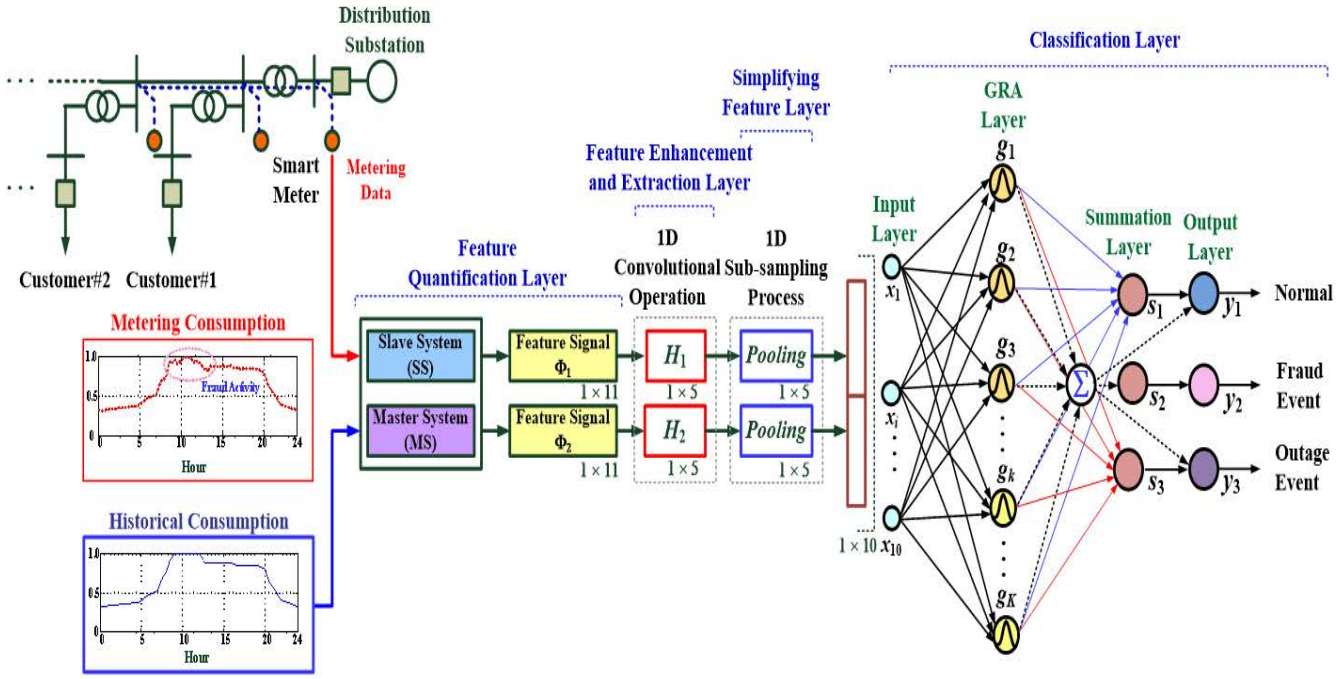


FIGURE 1. Configuration of the proposed D-H based quantizer and 1D CNN-based classifier for automatic fraud activity screening.

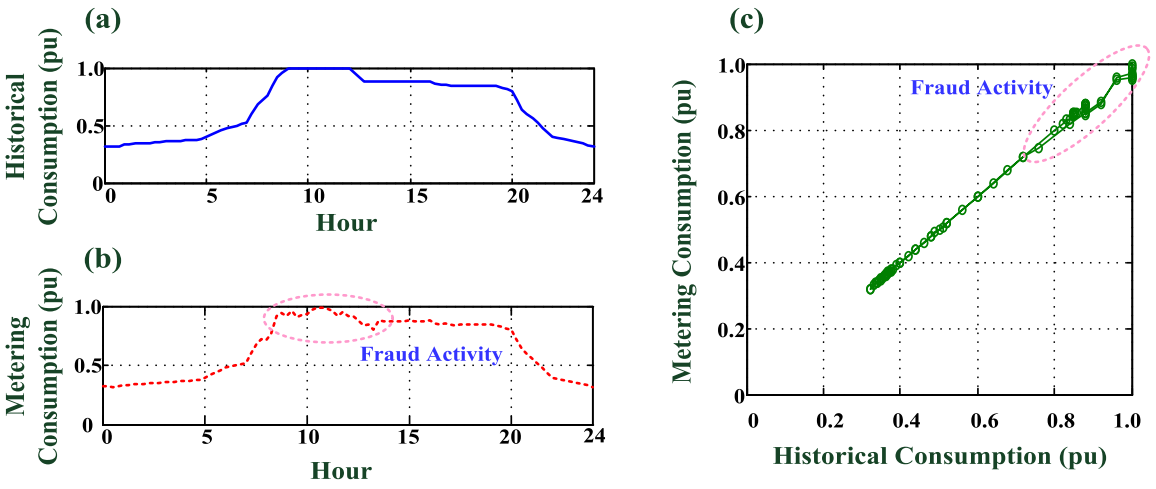


FIGURE 2. Historical consumption and metering consumption patterns and their correlation analysis. (a) the historical consumption pattern, (b) metering consumption pattern involving fraud activity, (c) correlation analysis with metering consumption versus historical consumption pattern.

In the pooling layer, a 1D downsampling process is used to reduce the dimension of the feature pattern and expressed as follows:

$$x_e[i] = X_e[3i], I = 1, 2, 3, \ldots, n' \tag{15}$$

$$n' = \frac{N + M - 2}{3} \tag{16}$$

where $X_e[i]$, $e = 1, 2$, are the pooling features obtained with the sliding stride $= 3$ from the feature signals, X_e ($n' = 5$ in this study). Then, two pooling feature signals are combined as an input pattern, that is, $X = [x_1[i]|x_2[i]] = [x_1, x_2, x_3, \ldots, x_{10}]$ in vector form, and is fed to a GRA-based classification network to further screen fraud activities.

2) GRA-BASED CLASSIFICATION NETWORK

As seen in Figure 1, a fully connected GRA-based network [38]–[40], [45] is applied as a classifier, consisting of input layer (10 nodes), GRA layer, summation layer (four nodes), and output layer (three nodes). In the GRA layer, the Gaussian functions are used to measure the similarity between a reference feature signal (testing feature signal) and comparative feature signals (training feature signals),

as represented by $x_0 = [x_1(0), x_2(0), x_3(0), \ldots, x_i(0), \ldots, x_{10}(0)]$ and $x_k = [x_1(k), x_2(k), x_3(k), \ldots, x_i(k), \ldots, x_{10}(k)]$, $k = 1, 2, 3, \ldots, K$, respectively. The Gaussian functions can produce gray grades, $g(k)$, and can be defined as [35]–[37]

$$g(k) = \exp(-\frac{1}{2}(\frac{ED(k)^2}{\sigma^2})^2), k = 1, 2, 3, \ldots, K \quad (17)$$

where $ED(k)$ is the Euclidean distance (ED); K is the number of training feature signals; and σ is the standard deviation, which can be represented as [20], [21], [23], [24]

$$ED(k) = \sqrt{\sum_{i=1}^{10}(\Delta d_i(k))^2}, d_i(k) = x_i(0) - x_i(k), \\ i = 1, 2, 3, \ldots, 10, k = 1, 2, 3, \ldots, K, \quad (18)$$

The standard deviation can be automatically adjusted by

$$\sigma^2 = (\Delta d_{\max} - \Delta d_{\min})^2, \begin{cases} \Delta d_{\min} = \min\limits_{\forall i \forall k}(\Delta d_i(k)) \\ \Delta d_{\max} = \max\limits_{\forall i \forall k}(\Delta d_i(k)) \end{cases}, \\ (\Delta d_{\max} - \Delta d_{\min}) \neq 0 \quad (19)$$

where parameter, $d_i(k)$, is the difference between a reference feature signal, x_0, and comparative feature signals, x_k; $\Delta d_{\max}$ and $\Delta d_{\min}$ are the maximum and minimum deviation values, respectively; and K training data are created by feature signals and are used to establish the GRA-based connected network between the input layer and the summation layer. Then, the classifier's outputs, $y_j, j = 1, 2, 3, \ldots, m$ ($m = 3$ in this study) can be normalized by

$$y_j = \frac{\sum\limits_{k=1}^{K} w_{kj} g(k)}{\sum\limits_{k=1}^{K} g(k)} = \sum_{k=1}^{K} w_{kj} s_j, s_j = \frac{g(k)}{\sum\limits_{k=1}^{K} g(k)}, \quad (20)$$

$$Y_j = \begin{cases} 1, y_j \geq 0.50 \\ 0, y_j < 0.50 \end{cases} \quad (21)$$

where W_{KJ} is the connected weighted values between the GRA layer and summation layer, which can be established by the output training data, as denoted by binary values, 1 or 0. The output weighted values in vector can be labeled as [normal, fraud event, outage event] = [0 / 1, 0 / 1, 0 / 1], including (1) normal: [1, 0, 0]; (2) fraud event: [0, 1, 0]; and (3) outage event: [0, 0, 1]. The final output as Equation (11) can be determined by the threshold value of 0.5, which is set to confirm the "*possible event* (value 1)" or "*event absent* (value 0)."

III. EXPERIMENTAL RESULTS

The proposed screening NTL classifier was implemented on a tablet PC using a high-level graphical programming language in the LabVIEW and MATLAB software (NI™, Austin, TX, USA) and on a graphics processing unit (GPU, NVIDIA®GeForce®RTX™2080 Ti, 1755 MHz, 11 GB GDDR6). In this study, we used four processes to screen possible NTL events (as seen flowchart in Figure 3).

Step 1) Abnormal events such as fraud activity events and power outage events are simulated with Equation (1);

Step 2) The differences between the historical consumption and the metering consumption patterns are estimated preliminarily with the correlation analysis, as shown in Equation (12), and then the different levels are quantified with the D-H-based quantizer as Equations (10) and (11);

Step 3) The feature signals are enhanced and extracted with the 1D convolutional processes, as Equations (13) and (14);

Step 4) The feature signals are identified using the GRA-based classifier, including normal condition, fraud activity events, and power outage events.

In situational simulations, we could simulate the normal condition, fraud activities, and power outage events; hence, with three events taken into consideration, the changes of power consumptions, $\Delta P = p_t - P_t$, could be set: (1) 0.00 $\leq \Delta P <$ 0.05 (pu) for normal condition (Nor); (2) 0.05 $\leq \Delta P <$ 0.20 (pu) for likely fraud and serious fraud activities (NTL); and (3) 0.50 $\leq \Delta P <$ 0.90 (pu) for power outages (POut). Then, after feature quantification and extraction processes, as seen in Figure 4, we could produce the feature signals, X_1 and X_2, by using the D-H quantizer and two 1D convolutional processes for the established training dataset to train the proposed classifier, including 6 training patterns for "*Nor*," 16 training patterns for "*NTL*," and 18 training patterns for "*POut*," which could be distinguishably applied to the three identified classes. Table 1 shows the related data of the proposed method combining the D-H quantizer and 1D CNN-based multilayer classifier, including its layer functions, manners, and feature patterns (dimension of 1D feature signal). The feasibility study could be validated as described in detail in the subsequent sections.

A. FEATURE EXTRACTION AND MULTILAYER CLASSIFIER ESTABLISHMENT

Considering the 40 possible events, including 6 *Nors*, 16 *NTL*s, and 18 *POuts*, the proposed D-H-based quantizer with the damping parameter, δ =0.10 ($\delta \in [0.0, 1.0]$), was used to track the differences between the historical consumptions and the metering consumptions and then produce the SSDE signals, Φ_1 and Φ_2 (1×11), in the feature quantification layer. In the feature enhancement and extraction layer, we could use two 1D convolutional operators (with moving stride = 1) to deal with the SSDE signals, which could obtain the enhanced and extracted two feature signals from those, as seen from the feature signals, X_1 and X_2 (1×15), for *Nor*, *NTL*, and *POut* in Figure 4. Subsequently, the two pooling processes (with moving stride = 3) were used to downsample the two feature signals, x_1 and x_2 (1×5), and then two feature signals were combined into an input pattern $[x_1 | x_2]_{1\times10}$ for further classification tasks.

According to the 40 training patterns (40 input-output paired training dataset), we could establish a GRA-based network for performing the classification tasks, including 10 nodes in the input layer, 40 Gaussian function nodes in the GRA layer, 4 nodes in the summation layer, and 3 nodes in

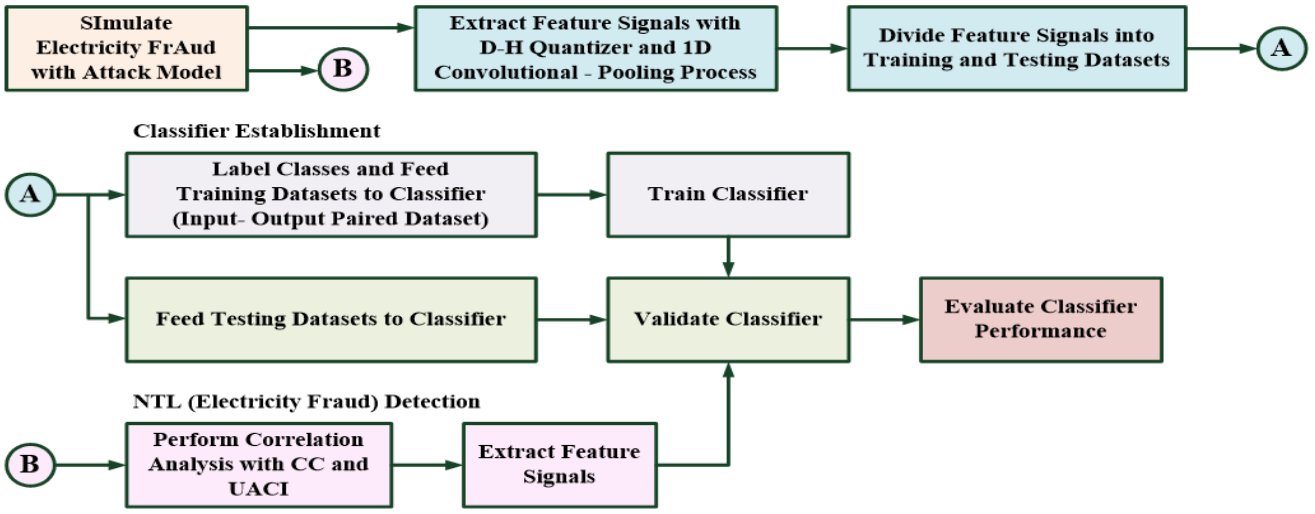


FIGURE 3. Flowchart of classifier establishment and automatic NTL (electricity fraud) detection.

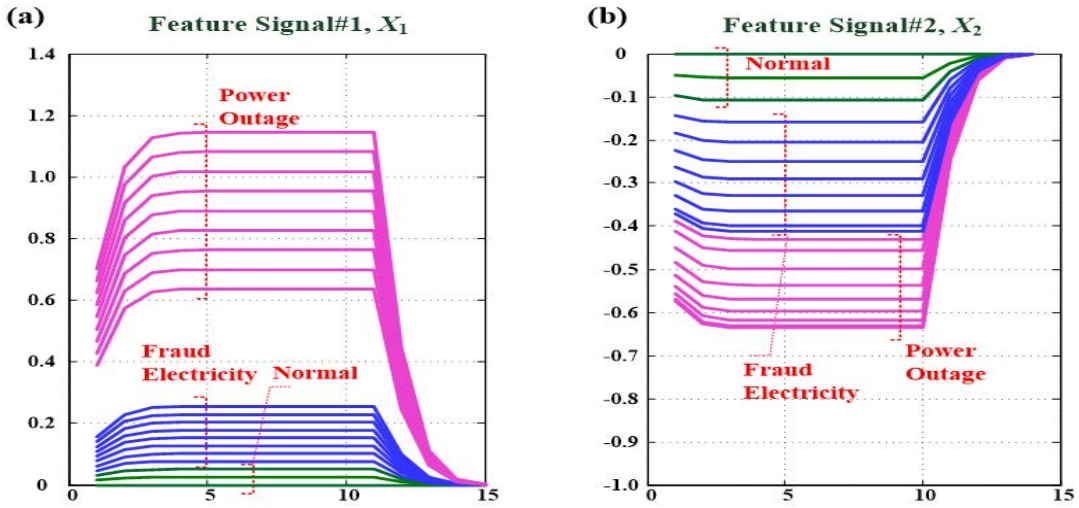


FIGURE 4. Two feature signals, X_1 (feature signal #1) and X_2 (feature signal #2), for identifying normal condition, electricity fraud, and power outage. (a) feature signal #1 extraction with convolution operator H_1, (b) feature signal #2 extraction with convolution operator H_2.

TABLE 1. Related data of D-H based quantizer and 1D CNN based multilayer classifier.

Layer Function	Manner	1D Feature Signal
Feature Quantification Layer	D-H based Quantizer with the Damping Parameter, $\delta = 0.10$	Φ_1 and Φ_2 (1×11)
Feature Enhancement and Extraction Layer	2 1D Convolutional Processes (stride = 1)	X_1 and X_2 (1×15)
Simplifying Feature Layer	2 1D Pooling Processes (stride = 3)	x_1 and x_2 (1×5)
Classification Layer	Multi-Layer Classifier: 10 input nodes, 40 pattern nodes, 4 summation nodes, 3 output nodes	Input Pattern Feeding into GRA-based Network: $[x_1 \mid x_2]$ (1×10)
	Learning Algorithm: GRA Method	

the output layer, as seen in Table 1. In the learning stage, the GRA-based classifier could perform mathematical operations by using Equations (17)–(19) to construct the classification network and make the final decision to identify the possible class by using Equations (20) and (21). Its pattern recognition scheme could deal with the incoming feature signals without the optimization algorithms and iterative computations to adjust the network's parameter. The network parameter in the GRA layer was automatically determined by using the maximum and minimum deviation values between a reference

feature signal and comparative feature signals, as in Equation (19). In the recalling stage, it had a 100% accuracy in identifying the possible classes.

B. CORRELATION ANALYSIS WITH LINEAR REGRESSION AND UACI

In the real word, fraud activity and power outage events might occur at any time. For example, as seen from the power consumption pattern in Figure 5(a), with two durations of fraud activities between 7:30 and 11:15 a.m. and between 4:15 and 6:45 p.m. The MDMS could perform the validation and estimation tasks for preliminarily screening abnormal events. Correlation analysis with linear regression method was used to measure the linear relationship between the metering consumptions and the historical consumptions within the short-time or the long-time monitoring as follows:

- Long-time monitoring: The timing slot was set as 24 Hours (as seen in Figure 5(b)). When any fraud activity occurred, 96 slots of metering data on metering consumptions versus historical consumptions were noted, and CC would decrease, reaching less than 0.98;
- Short-time monitoring: The screening window in Figure 6(a). shows that the timing slot was set to 3 hours (8:15 to 11:15 a.m.), Each screening window had 12 slots of metering data (as shown in Figure 6(b)), and the CC was 0.78, which was less than 0.98, as seen in Figure 6(c).

Short-time or long-time monitoring verified the use of correlation analysis with linear regression to preliminarily measure the linear relationship between the metering consumption and the historical consumption. The correlation study revealed that the CC was 0.78, indicating that there was a significant. There is a difference between the metered consumption and the historical consumption in this time slot. Equation (12) was also used to quantify the different values of the UACI [44], [45], which were greater than 5.00% for identified abnormal event occurrences and 6.38% and 9.38% for long-time and short-time monitoring, respectively. Hence, in this study, we chose short-time monitoring for screened fraud activities and power outages. When CC was less than 0.98 and the UACI index was greater than 5.00%, the D–H-based quantizer and 1D convolutional operators were used to extract the two feature signals for identifying the possible classes, either "*NTL*" or "*Pout*," as seen in the SSDE signals and 1D feature signals in Figures 6 (d) and 6 (e), respectively. As a result, correlation analysis revealed that the CPU time for estimating CC and UACI indexes for preliminary screening of aberrant events was less than 0.10 s in each short-time screening window. This finding confirms that the proposed method is a promising automatic detection approach for real-time applications.

C. NTL AND POWER OUTAGE DETECTION

We simulated case study #1 with serious fraud activities from 2:15 to 5:15 p.m. on the first day and case study #2 with power outages and fraud activities from 9:00 to 10:45 a.m. on the second day, as presented by the power consumption patterns in Figures 7(a) and 7(b), respectively. With the metering sequence data of the power consumption patterns, the correlation analysis with UACI index could be used to preliminarily identify the abnormal events, as seen in the UACI indexes versus power consumption changes, ΔP (%), for case study #1 and case study #2 ($R^2 = 0.96$) in Figure 8(a). The UACI indexes were sensitive to preliminary identification of the Nor, NTL, and POut events (20 experimental tests) for short-time monitoring. Hence, while the UACI index was greater than 5%, the D-H based quantizer quantified the difference levels between the metering consumptions and the historical consumptions. With two 1D convolutional processes, two feature signals could be extracted and obtained for each screening window, as seen in Figures 7(c) and 7(d). As seen in Figure 8, the comprehensive feature signal X (as seen in Figure 8(b)) was positively correlated with power consumption changes ($R^2 = 0.94$). The comprehensive feature signal X was between 0.50 and 1.0, thus indicating that the power consumption changes from 5% to 12% and signifying that the power consumptions involved possible fraud activities. In addition, the comprehensive feature signals greater than 2.0 indicated the ΔP (%) >50% for "*POut*" event or "*POut* involved *NTL*." Hence, two feature signals (X_1 and X_2) could be promising feature patterns for separating the normal condition from abnormal events.

For case study #1, following the 10 screening windows, Table 2 showed the overall screening results as "*NTL event*" from 2:15 to 7:00 p.m. Equation (1) was used to simulate the electricity fraud attacks, and then the metering consumptions could be obtained for further analysis and classification. To feed the feature signals, as seen in Figure 7(c), the proposed CRA-based classifier could identify the possible events accurately. Performing the screening tasks for each screening window took an average CPU time of < 0.10 s, thus offering promising results for identifying the electricity fraud events from screening windows #1 to #8. This finding confirms that the proposed screening method is feasible for fraudulent activity detection in short-time monitoring applications.

For case study #2, a fault is assumed to have occurred at a feeder between a power system and the customers' ownership demarcation point in the grid. The protection relays tripped the feeder's circuit breakers to isolate the fault section at 9:00 a.m. According to the restoration scheme, the downstream transformer, loads, and distributed generation (DG) units (wind generators and photovoltaic arrays) are first isolated. About 60% of the downstream loads might be picked up by 90% of the capacity of the neighboring supportive transformers because of the peak load. During a power outage event, key loads and partial loads could be transferred to the supportive transformers and also tripped to prevent islanding operation [14], [46]. After a fault section is located, a fault is isolated, and the system is kept energized by supportive transformers, then the DG units offer additional service for the overall loads. At 10:45 a.m., the customers could obtain the scheduled supply consumptions, as seen in Figure 7(b).

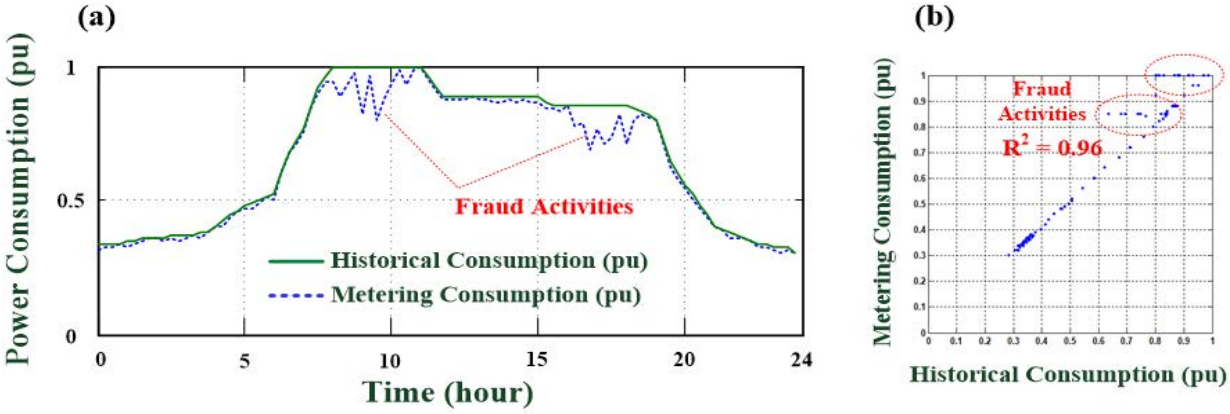


FIGURE 5. Power consumption pattern involving fraud activities (red dashed line). (a) power consumption patterns for historical consumptions and metering consumptions during one workday, (b) correlation analysis ($R^2 = 0.96$) with metering consumptions versus historical consumptions.

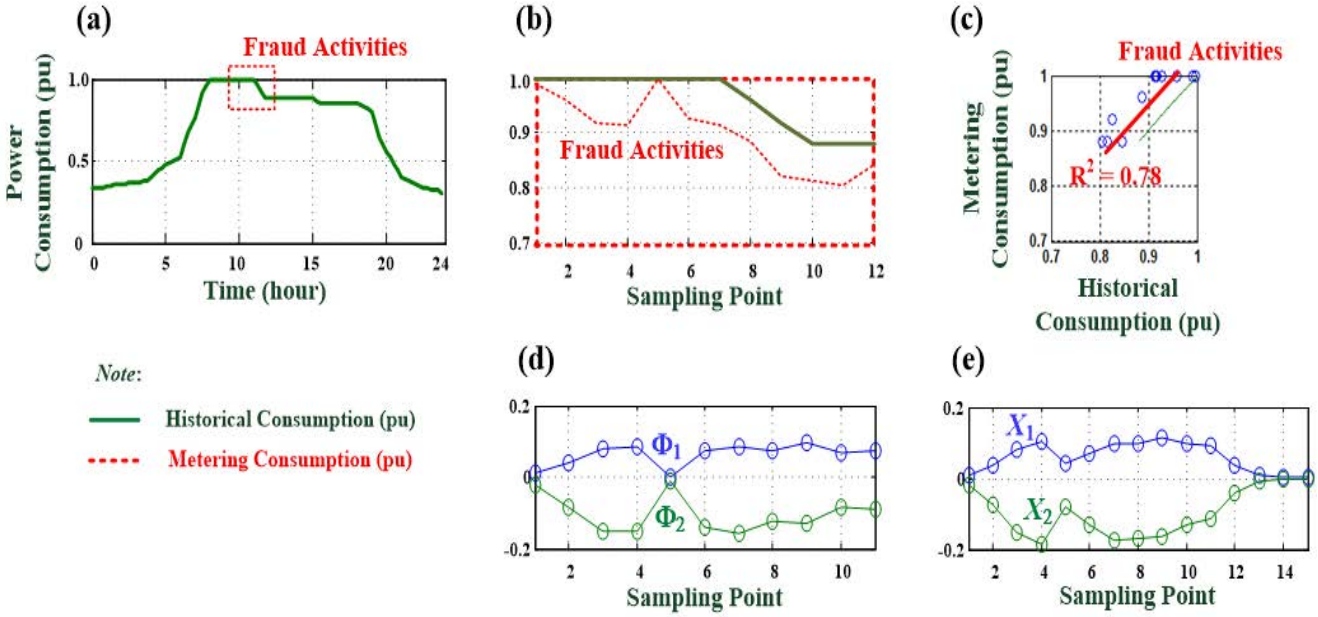


FIGURE 6. Short-time monitoring for fraud activity detection. (a) screening window (3 hours duration) for fraud activity detection, (b) power consumption pattern for fraud activity, (c) correlation analysis for fraud activity ($R^2 = 0.78$), (d) SSDE signal (Φ_1 and Φ_2) extraction with D-H based quantizer, (e) 1D feature signal (X_1 and X_2) extraction with 1D convolutional operator.

In the above-described scenarios, we also simulated a power outage involving fraudulent activity by using Equation (1). During a power outage, the proposed screening method could identify the power outage event (windows #1 to #4) that involves fraudulent activities, as seen from the screening results in Table 3. Two simulation tests for the two scenarios confirmed that the proposed method is a promising approach for NTL and POut detection for AMI monitoring in real-time applications.

D. DISCUSSIONS

A comparison of performance comparison with other approaches revealed that the proposed screening method performed well and produced promising findings, as shown in Table 4. In previous studies [47], [48], SVM was widely applied in customer usage pattern classification. Based on classification-based detection, the SVM technique had an accuracy of roughly 60%–70% [47], [48], which indicates that the regular power consumption was identified as illegal consumption. As a result, the SVM with FIS [49] was also built based on historical data to improve screening accuracy, which may boost the accuracy level. The average accuracy could be higher than 70%. However, in order to train the SVM-based classifier, the effects of fraud activities in a smart grid must be examined by comparing the power consumption with or without considering NTL. As a result, the

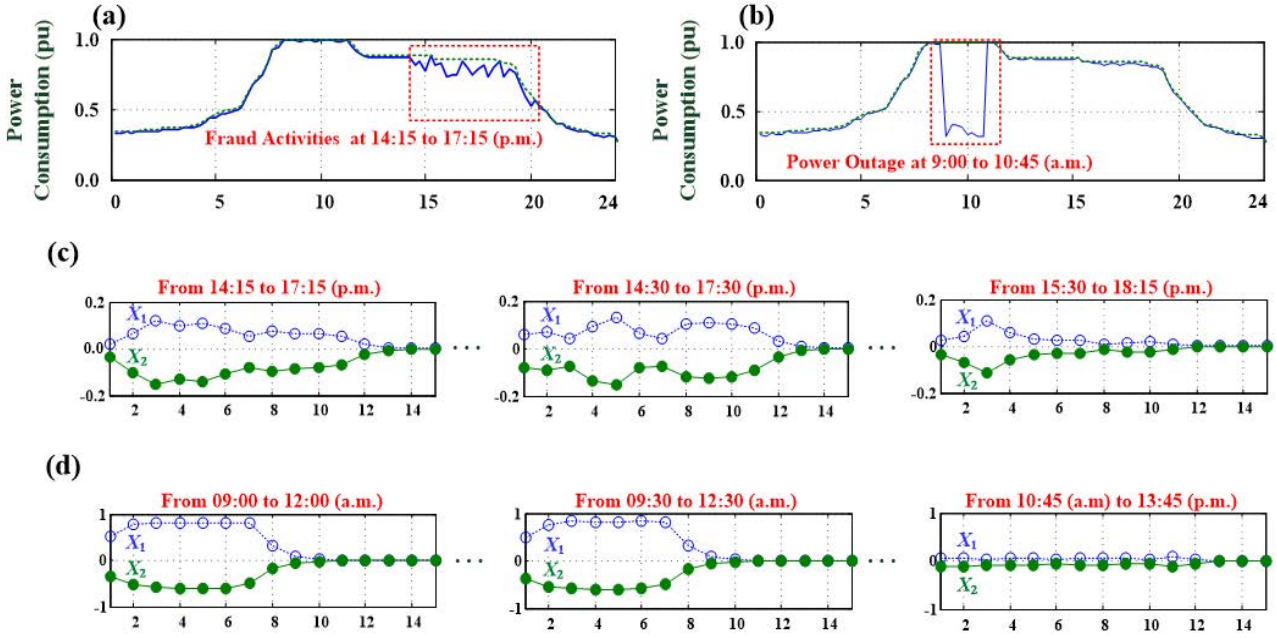
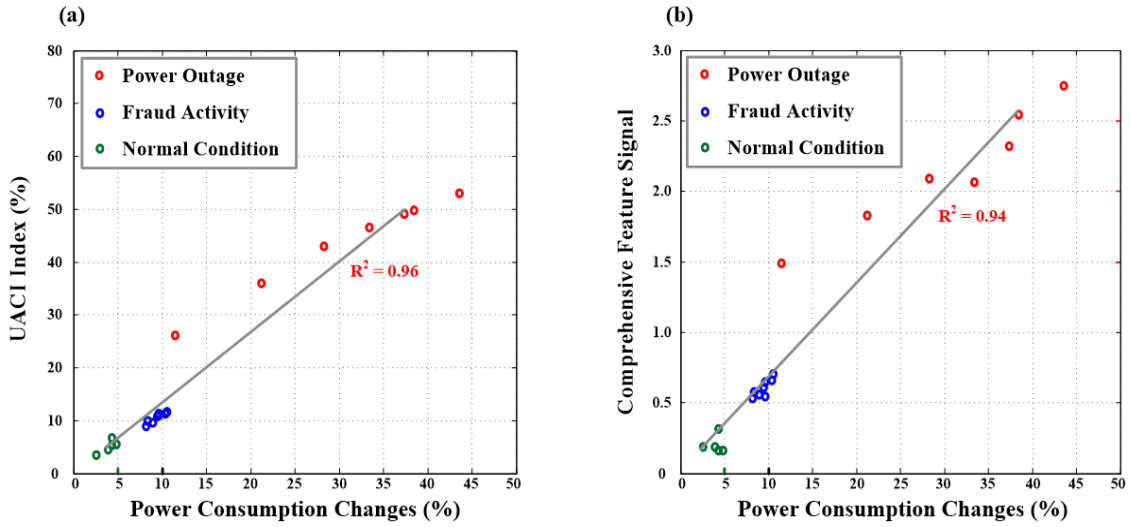


FIGURE 7. Case studies for fraud activity and power outage and their feature signals. (a) Power consumption pattern with fraud activities from 14:15 to 17:15 (p.m.), (b) Power consumption pattern with fraud activities and power outage from 9:00 to 10:45 (a.m.), (c) Feature signals of fraud activities for each screening window, (d) Feature signals of power outages and fraud activities for each screening window.



Note: Comprehensive Feature Signal: $X = \sqrt{\sum_{i=1}^{15} (X_1[i])^2 + (X_2[i])^2}$, $i = 1, 2, 3, \ldots, 15$

FIGURE 8. Correlation analysis with linear regression for (a) UACI indexes versus power consumption changes ($R^2 = 0.96$), (b) comprehensive feature signal versus power consumption changes ($R^2 = 0.94$).

model required a huge number of training datasets collected from SMs to train the classifier. to continuously model the intended tasks. The ML-based methods required more intervention to select feature patterns (manual selection) and to feed new training datasets to train the classifier. Over time, it became challenging to maintain the classifier's performance and modify its capabilities in real-time applications. The FIS also requires the establishment of a fuzzy membership matrix and inference methods to identify the possible classes. Furthermore, SVM required a learning algorithm, such as gradient descent, least mean square, or genetic algorithms (GA) [50], to adjust the network parameters through iteration computations, which were used to find the optimal parameters to achieve the convergent condition, with the screening accuracy being improvable. Such an ML method lacked the feature extraction function in data preprocessing.

TABLE 2. Screening Results for Case Study#1 with fraud activities detection with 10 timing-slot testing.

Timing Slot for Screening		Abnormal Event (average ΔP)	correlation analysis		Screening Results
			CC	UACI (%)	
1	14 : 15 ~ 17 : 15 (p.m.)	Fraud Activity (average ΔP = 8.38%)	0.04	10.03	[0, 1, 0] : NTL
2	14 : 30 ~ 17 : 30 (p.m.)	Fraud Activity (average ΔP = 9.58%)	0.03	11.02	[0, 1, 0] : NTL
3	14 : 45 ~ 17 : 45 (p.m.)	Fraud Activity (average ΔP = 9.43%)	0.49	10.77	[0, 1, 0] : NTL
4	15 : 00 ~ 18 : 00 (p.m.)	Fraud Activity (average ΔP = 10.37%)	0.02	11.29	[0, 1, 0] : NTL
5	15 : 15 ~ 18 : 15 (p.m.)	Fraud Activity (average ΔP = 8.21%)	0.03	8.94	[0, 1, 0] : NTL
6	15 : 30 ~ 18 : 30 (p.m.)	Fraud Activity (average ΔP = 8.99%)	0.02	9.61	[0, 1, 0] : NTL
7	15 : 45 ~ 18 : 45 (p.m.)	Fraud Activity (average ΔP = 9.62%)	0.03	11.33	[0, 1, 0] : NTL
8	16 : 00 ~ 19 : 00 (p.m.)	Fraud Activity (average ΔP = 10.55%)	0.19	11.60	[0, 1, 0] : NTL
9	16 : 15 ~ 19 : 15 (p.m.)	Normal (average ΔP = 4.32%)	0.61	6.69	[1, 0, 0] : Nor
10	16 : 30 ~ 19 : 30 (p.m.)	Normal (average ΔP = 2.53%)	0.66	3.50	[1, 0, 0] : Nor

TABLE 3. Screening results for case study#2 with power outages and fraud activities detection with 10 timing-slot testing.

Timing Slot for Screening		Abnormal Event (average ΔP)	correlation analysis		Screening Results
			CC	UACI (%)	
1	9 : 00 ~ 12 : 00 (a.m.)	Power Outage (average ΔP = 43.54%)	0.44	52.96	[0, 0, 1] : POut
2	9 : 15 ~ 12 : 15 (a.m.)	Power Outage (average ΔP = 38.44%)	0.47	49.86	[0, 0, 1] : POut
3	9 : 30 ~ 12 : 30 (a.m.)	Power Outage (average ΔP = 37.78%)	0.77	49.18	[0, 0, 1] : POut
4	9 : 45 ~ 12 : 45 (a.m.)	Power Outage (average ΔP = 33.41%)	0.78	46.62	[0, 1, 0] : NTL
5	10 : 00 (a.m) ~ 13 : 00 (p.m.)	Power Outage (average ΔP = 28.29%)	0.77	42.92	[0, 1, 0] : NTL
6	10 : 15 (a.m) ~ 13 : 15 (p.m.)	Power Outage (average ΔP = 21.24%)	0.74	35.94	[0, 1, 0] : NTL
7	10 : 30 (a.m) ~ 13 : 30 (p.m.)	Power Outage (average ΔP = 11.48%)	0.29	26.15	[0, 1, 0] : NTL
8	10 : 45 (a.m) ~ 13 : 45 (p.m.)	Normal (average ΔP = 4.84%)	0.77	5.49	[1, 0, 0] : Nor
9	11 : 00 (a.m) ~ 14 : 00 (p.m.)	Normal (average ΔP = 4.36%)	0.69	5.29	[1, 0, 0] : Nor
10	11 : 15 (a.m) ~ 14 : 15 (p.m.)	Normal (average ΔP = 3.93%)	0.69	4.55	[1, 0, 0] : Nor

TABLE 4. Comparison of performances between the proposed multilayer classifier and other traditional classifiers.

Method / Task	D-H based Quantizer + 1D CNN based Classifier	D-H based Quantizer + CRA based Classifier	D-H based Quantizer + SVM [47-48] Based Classifier
Feature Extraction	SSDE and 1D Convolutional Process	SSDE Process	Historical Consumption Data
Classifier Configuration	Multilayer Network	Multilayer Network	Multilayer Network
Feature Parameter Reduction	Pooling Process	—	—
Training Dataset	Yes, Key Feature Signals Requirement	Yes, Key Feature Signals Requirement	Yes, Major
Learning Algorithm	GRA Manner	Traditional GRA Manner	Gradient Descent Algorithm Least Mean Square Algorithm
Network Parameter Assignment	Yes, Minor	Yes, Minor	Yes, Major
Network Parameter Adjustment	No	No	Yes
Iteration Computation	No	No	Yes
Convergent Condition	No	No	Yes, $< 10^{-2}$
Precision (%)	97.50 (TP: 39, FP: 1)	95.00 (TP: 38, FP: 2)	92.50 (TP: 37, FP: 3)
Recall (%)	95.12 (TP: 39, FN: 2)	92.68 (TP: 38, FN: 3)	88.10 (TP: 37, FN: 5)
F1 Score	0.9630 (TP: 39, FP: 1, FN: 2)	0.9383 (TP: 38, FP: 2, FN: 3)	0.9024 (TP: 37, FP: 3, FN: 5)
Accuracy (%)	94.00 (TP: 39, FP: 1, TN: 8, FN: 2)	90.00 (TP: 38, FP: 2, TN: 7, FN: 3)	84.00 (TP: 37, FP: 3, TN: 5, FN: 5)

Note: (1) $Recall\ (\%) = (\frac{TP}{TP + FN}) \times 100\,\%$, (2) $Precision\,(\%) = (\frac{TP}{TP + FP}) \times 100\%$,

(3) $Accuracy\ (\%) = (\frac{TP + TN}{TP + FN + TN + FP}) \times 100\,\%$, (4) $F1\ Score = \frac{2TP}{2TP + FP + FN}$.

where *TP* and *TN* are the true positive and true negative, respectively; *FP* and *FN* are the false positive and false negative, respectively [34, 38].

As shown in Table 5, DL-based methods such as wide and deep CNN, CNN–GRU, and CNN–LSTM [22], [27], [28], [50], [51], based on different datasets (for example, Smart Grid Corporation of China [21] or State Grid Corporation of China [22]), had also created a classifier to automate the feature extraction and classification processes for metering data classification, which could achieve greater than 85% classification accuracy and had higher indexes like Precision

TABLE 5. Comparison of performances between the proposed classifier and DL-based classifiers.

Literature	Method	Precision (%)	Recall (%)	Accuracy (%)	F1 Score	AUC
[20]	CNN - LSTM	90.00	87.00	89.00	—	—
[21]	Time Least Square Generative Adversarial Network (TLSGAN)	—	—	—	—	ROC-AUC: 0.96 PR-AUC: 0.97
[22]	CNN - LSTM	92.00	96.00	89.00	0.9400	—
[27]	Wide and Deep CNN	—	—	—	—	0.7815
[28]	CNN - GRU	82.20	83.25	85.00	0.8400	0.85
	CNN-GRU–MRFO (Manta Ray Foraging Optimization)	91.11	93.00	91.10	0.8900	0.91
[51]	CNN - LSTM	—	—	96.90	—	—
Proposed Method	D-H based Quantizer + 1D CNN based Classifier	97.50	95.12	94.00	0.9630	—

(%), Recall (%), and F1 score to evaluate the classifier's performance (as shown in Table 4). The DL-based approaches had a more complicated scheme, with many convolutional-pooling layers and a fully linked network to set up the various multilayer classifiers, which could automatically extract feature patterns without operator intervention and improve the classification accuracy. Their models, on the other hand, took a longer time to train the classifier using the adaptive moment estimation method (ADAM) or a backpropagation algorithm. However, once trained, their models could rapidly produce results and also required large volumes of training datasets to retrain the classifiers to maintain the intended purposes.

In the feature quantification layer, the D–H based quantizer and the standard CRA approach [38]–[40] were coupled. This model was also a multilayer network that could estimate the difference in consumption patterns and be used to distinguish nor from NTL or POut occurrence, which could be employed to separate Nor from NTL or POut events. It required modest network parameter assignment without the use of iterative computations for making adjustments. This model included a feature extraction function that allowed it to overcome the challenges of adjustable mechanism construction and network parameter optimization. We randomly constructed testing patterns and feature signals for untrained consumption patterns, including 10 Nors, 20 NTLs, and 20 Pouts, to validate the performance of the D–H-based quantizer + SVM-based classifier, the D–H-based quantizer + CRA-based classifier, and the proposed multilayer classifier. Table 4 shows the performance comparisons and experimental findings for three multilayer classifiers using the same untrained dataset. For screening NTL and POut events, the approach achieved a recall of 95.12% (true positive (TP): 39, FN: 2)), a precision of 97.50% (TP: 39, FP: 1) for screening NTL and POut events, an accuracy of 94% (TP: 39, FP: 1, TN: 8, FN: 2), as well as an F1 score of 0.9630 (TP: 39, FN: 2). As shown in Table 4, the suggested multilayer of technique performed better than existing traditional multilayer classifiers. The suggested multilayer technique method has better performance than existing traditional multilayer classifiers. Some advantages of the proposed multilayer classifiers are shown below.

- The difference levels could be quantified by discrete SSDE equation in feature quantification preprocessing) layer with simple mathematic operations;
- The feature signal could enhance the characteristics for distinguished levels to separate the Nor events from the NTL or POut events;
- The training process was completed using simple mathematical operations that did not require sophisticated iteration computations or optimization algorithm requirements.

IV. CONCLUSION

In this study, we developed the D-H-based quantizer and 1D CNN-based multilayer classifier for electricity fraud and power outage screening in a smart grid. In the AMI monitoring system, AMI enables bidirectional communication metering electricity consumptions between SMs and utility companies, which could be processed in real time and sends responses for demand management. In energy supplier-side management, distribution system operators need to collect metering data at each meter, which are critical to analyzing the customers' load patterns and to understanding customers' behaviors. Hence, fault detection, isolation, and restoration verification could be performed during outage detection. Utility companies aim to supply stable energy for customers and protect customers from electricity fraud. Thus, the classification-based detection approach is widely used to compare the historical consumption and metering consumption as observed through short-time or long-time monitoring. This process consists of metering data acquisition, data preprocessing, feature extraction, classifier training and validation, and classification. Hence, the combined D-H based quantizer and 1D CNN-based multilayer classifier was

used to identify abnormal consumption patterns from the metering data based on a testing dataset that contained the training dataset of Nor, NTL, and POut classes. Experimental results indicated that the proposed method obtained promising results, with higher precision (97.50%), recall (95.12), F1 score (0.9630), and accuracy (94.00%) than other traditional classifiers in identifying possible fraud events. In addition, the proposed classifier's algorithm can be implemented easily in a programmable microprocessor or an embedded system by using high-level programming language. The proposed multilayer classifier could also be integrated into existing AMI monitoring systems for use in real-time analysis and classification applications.

ABBREVIATIONS

NTL	Nontechnical Loss (Electricity Fraud)
TL	Technical Loss
Nor	Normal Condition
POut	Power Outage
D-H	Duffing – Holmes
MS	Master System
SS	Slave System
CNN	Convolutional Neural Network
SSDE	Self-synchronization Dynamic Error
GRA	Gray Relational Analysis
CC	Correlation Coefficient
UACI	Unified Averaged Changed Intensity
SM	Smart Meter
AMI	Advanced Metering Infrastructure
MDMS	Meter Data Management System
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
DG	Distributed Generation
ANN	Artificial Neural Network
SVM	Support Vector Machine
FIS	Fuzzy Inference System
CNN-LSTM	CNN and a long short-term memory (LSTM)
GRU	Gated Recurrent Unit
CNN-GRU	CNN and GRU
ED	Euclidean Distance
GPU	Graphics Processing Unit
GA	Genetic Algorithms
ADAM	Adaptive Moment Estimation Method
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
TLSGAN	Time Least Square Generative Adversarial Network
AUC	Area Under Curve
PR-AUC	Precision Recall - Area Under Curve
ROC-AUC	Receiver Operating Characteristic – Area Under Curve

REFERENCES

[1] M. U. Cantillo, R. González, J. J. Mares, and C. Q. Monroy, "Intelligent system for non-technical losses management in residential users of the electricity sector," *Ingeniería e Investigación*, vol. 38, no. 2, pp. 52–60, May 2018.
[2] S. Sahoo, D. Nikovski, T. Muso, and K. Tsuru, "Electricity theft detection using smart meter data," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT)*, Feb. 2015, pp. 1–5.
[3] A. Y. Kharal, H. A. Khalid, A. Gastli, and J. M. Guerrero, "A novel features-based multivariate Gaussian distribution method for the fraudulent consumers detection in the power utilities of developing countries," *IEEE Access*, vol. 9, pp. 81057–81067, 2021.
[4] A. A. Esmael, H. H. da Silva, T. Ji, and R. da Silva Torres, "Non-technical loss detection in power grid using information retrieval approaches: A comparative study," *IEEE Access*, vol. 9, pp. 40635–40648, 2021.
[5] C. Muniz, K. Figueiredo, M. Vellasco, G. Chavez, and M. Pacheco, "Irregularity detection on low tension electric installations by neural network ensembles," in *Proc. Int. Joint Conf. Neural Netw.*, Jun. 2009, pp. 2176–2182.
[6] P. Jokar, N. Arianpoo, and V. C. M. Leung, "Electricity theft detection in AMI using customers' consumption patterns," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 216–226, Jan. 2016.
[7] A. H. Nizar, Z. Y. Dong, M. Jalaluddin, and M. J. Raffles, "Load profiling method in detecting non-technical loss activities in a power utility," in *Proc. IEEE Int. Power Energy Conf.*, Nov. 2006, pp. 82–87.
[8] A.-H. Mohsenian-Rad, V. W. S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid," *IEEE Trans. Smart Grid*, vol. 1, no. 3, pp. 320–331, Dec. 2010.
[9] R. Jiang, R. Lu, Y. Wang, J. Luo, C. Shen, and X. Shen, "Energy-theft detection issues for advanced metering infrastructure in smart grid," *Tsinghua Sci. Technol.*, vol. 19, no. 2, pp. 105–120, Apr. 2014.
[10] M. Wen, R. Lu, J. Lei, H. Li, X. Liang, and X. Shen, "SESA: An efficient searchable encryption scheme for auction in emerging smart grid marketing," *Secur. Commun. Netw.*, vol. 7, no. 1, pp. 234–244, 2014.
[11] A. P. S. Meliopoulos, G. Cokkinides, R. Huang, E. Farantatos, S. Choi, Y. Lee, and X. Yu, "Smart grid technologies for autonomous operation and control," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 1–10, Mar. 2011.
[12] R. Balaambikha and P. G. Scholar, "ENHDF: Intrusion detection system for distribution networks in smart grid," *Adv. Natural Appl. Sci.*, vol. 10, no. 8, pp. 138–147, 2016.
[13] F. Li, W. Qiao, H. Sun, H. Wan, and P. Zhang, "Smart transmission grid: Vision and framework," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 168–177, Sep. 2010.
[14] S.-J. Chen, T.-S. Zhan, C.-H. Huang, J.-L. Chen, and C.-H. Lin, "Nontechnical loss and outage detection using fractional-order self-synchronization error-based fuzzy Petri nets in micro-distribution systems," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 411–420, Jan. 2015.
[15] S.-C. Huang, Y.-L. Lo, and C.-N. Lu, "Non-technical loss detection using state estimation and analysis of variance," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2959–2966, Aug. 2013.
[16] S. S. S. R. Depuru, L. Wang, V. Devabhaktuni, and P. Nelapati, "A hybrid neural network model and encoding technique for enhanced classification of energy consumption data," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2011, pp. 1–8.
[17] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, and S. Zonouz, "A multi-sensor energy theft detection framework for advanced metering infrastructures," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1319–1330, Jul. 2013.
[18] B. C. Costa, B. L. A. Alberto, A. M. Portela, M. W, and E. O. Eler, "Fraud detection in electric power distribution networks using an ann-based knowledge-discovery process," *Int. J. Artif. Intell. Appl.*, vol. 4, no. 6, pp. 17–23, Nov. 2013.

[19] P. Glauner, J. A. Meira, P. Valtchev, R. State, and F. Bettinger, "The challenge of non-technical loss detection using artificial intelligence: A survey," *Int. J. Comput. Intell. Syst.*, vol. 10, pp. 760–775, Feb. 2017.

[20] M. Hasan, R. N. Toma, A. A. Nahid, M. Islam, and J. M. Kim, "Electricity theft detection in smart grid systems: A CNN-LSTM based approach," *Energies*, vol. 12, no. 17, pp. 1–18, 2019.

[21] F. Shehzad, N. Javaid, A. Almogren, A. Ahmed, S. M. Gulfam, and A. Radwan, "A robust hybrid deep learning model for detection of non-technical losses to secure smart grids," *IEEE Access*, vol. 9, pp. 128663–128678, 2021.

[22] C. Tian, J. Ma, C. Zhang, and P. Zhan, "A Deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network," *Energies*, vol. 11, no. 12, pp. 1–13, 2018.

[23] R. Jiang, H. Tagaris, A. Lachsz, and M. Jeffrey, "Wavelet based feature extraction and multiple classifiers for electricity fraud detection," in *Proc. IEEE/PES Transmiss. Distrib. Conf. Exhib.*, Oct. 2002, pp. 2251–2256.

[24] S. I. Kim, H. S. Kim, Y. J. Joo, and J. H. Kim, "Power usage pattern and consumption separation method by load devices based on remote metering system's load profile data," in *Proc. Int. Conf. Control, Autom. Syst.*, Oct. 2011, pp. 1669–1671.

[25] N. Fabian, G. Figueroa, and C.-C. Chu, "NTL detection in electric distribution systems using the maximal overlap discrete wavelet-packet transform and random undersampling boosting," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 7171–7180, Nov. 2018.

[26] C. Leon, F. Biscarri, I. Monedero, J. I. Guerrero, J. Biscarri, and R. Millan, "Variability and trend-based generalized rule induction model to NTL detection in power companies," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 1798–1807, Nov. 2011.

[27] Z. Zheng, Y. Yang, X. Niu, H.-N. Dai, and Y. Zhou, "Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids," *IEEE Trans. Ind. Informat.*, vol. 14, no. 4, pp. 1606–1615, Apr. 2018.

[28] N. Ayub, K. Aurangzeb, M. Awais, and U. Ali, "Electricity theft detection using CNN-GRU and manta ray foraging optimization algorithm," in *Proc. IEEE 23rd Int. Multitopic Conf. (INMIC)*, Nov. 2020, pp. 1–6.

[29] N. Javaid, H. Gul, S. Baig, F. Shehzad, C. Xia, L. Guan, and T. Sultana, "Using GANCNN and ERNET for detection of non technical losses to secure smart grids," *IEEE Access*, vol. 9, pp. 98679–98700, 2021.

[30] A. Ullah, S. M. Anwar, M. Bilal, and R. M. Mehmood, "Classification of arrhythmia by using deep learning with 2-D ECG spectral image representation," *Remote Sens.*, vol. 12, no. 10, pp. 1–14, 2020.

[31] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications: A survey," *Mech. Syst. Signal Process.*, vol. 151, pp. 1–21, Apr. 2021.

[32] Y. W. F. Yang, Y. Liu, X. Zha, and S. Yuan, "A comparison of 1-D and 2-D deep convolutional neural networks in ECG classification," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* Honolulu, HI, USA, 2018, pp. 1–4.

[33] C.-H. Lin, J.-X. Wu, C.-D. Kan, P.-Y. Chen, and W.-L. Chen, "Arteriovenous shunt stenosis assessment based on empirical mode decomposition and 1D convolutional neural network: Clinical trial stage," *Biomed. Signal Process. Control*, vol. 66, Apr. 2021, Art. no. 102461.

[34] P.-Y. Chen, Z.-L. Sun, J.-X. Wu, C. C. Pai, C.-M. Li, C.-H. Lin, and N.-S. Pai, "Photoplethysmography analysis with Duffing–Holmes self-synchronization dynamic errors and 1D CNN-based classifier for upper extremity vascular disease screening," *Processes*, vol. 9, no. 11, pp. 1–17, 2021.

[35] E. Tamaševičiute, A. Tamaševičius, G. Mykolaitis, and S. Bumeliene, "Analogue electrical circuit for simulation of the Duffing–Holmes equation," *Nonlinear Anal., Model. Control*, vol. 13, no. 2, pp. 241–252, Apr. 2008.

[36] Y.-P. Kuo, C.-L. Kuo, C.-H. Lin, Y.-R. Pu, and S.-M. Liang, "Terminal fuzzy sliding mode control for the Duffing–Holmes system," in *Proc. Int. Conf. Fluid Power Mechatronics*, Aug. 2011, pp. 473–476.

[37] C.-H. Lo and N. Ansari, "CONSUMER: A novel hybrid intrusion detection system for distribution networks in smart grid," *IEEE Trans. Emerg. Topics Comput.*, vol. 1, no. 1, pp. 33–44, Jun. 2013.

[38] C.-H. Lin, J.-X. Wu, C.-M. Li, P.-Y. Chen, N.-S. Pai, and Y.-C. Kuo, "Enhancement of chest X-ray images to improve screening accuracy rate using iterated function system and multilayer fractional-order machine learning classifier," *IEEE Photon. J.*, vol. 12, no. 4, pp. 1–19, Jul. 2020.

[39] Y. Wua, F. Zhou, and J. Kong, "Innovative design approach for product design based on TRIZ, AD, fuzzy and Grey relational analysis," *Comput. Ind. Eng.*, vol. 140, pp. 1–12, 2020.

[40] F. Sarrafa and S. H. Nejadb, "Improving performance evaluation based on balanced scorecard with grey relational analysis and data envelopment analysis approaches: Case study in water and wastewater companies," *Eval. Program Planning*, vol. 79, pp. 1–12, Apr. 2020.

[41] J. Y. Kim, Y. M. Hwang, Y. G. Sun, I. Sim, D. I. Kim, and X. Wang, "Detection for non-technical loss by smart energy theft with intermediate monitor meter in smart grid," *IEEE Access*, vol. 7, pp. 129043–129053, 2019.

[42] S.-C. Yip, W.-N. Tan, C. Tan, M.-T. Gan, and K. Wong, "An anomaly detection framework for identifying energy theft and defective meters in smart grids," *Int. J. Electr. Power Energy Syst.*, vol. 101, pp. 189–203, Oct. 2018.

[43] A. N. K. Telem, C. M. Segning, G. Kenne, and H. B. Fotsin, "A simple and robust gray image encryption scheme using chaotic logistic map and artificial neural network," *Adv. Multimedia*, vol. 2014, pp. 1–13, Dec. 2014.

[44] M. Dridi, M. A. Hajjaji, B. Bouallegue, and A. Mtibaa, "Cryptography of medical images based on a combination between chaotic and neural network," *IET Image Process.*, vol. 10, no. 11, pp. 830–839, 2016.

[45] K. C. Chang and M. F. Yeh, "Grey relational analysis based approach for data clustering," *IEE Proc. Vis., Image Signal Process.*, vol. 152, no. 2, pp. 165–172, Apr. 2005.

[46] A. P. S. Meliopoulos, "Smart grid technologies for autonomous operation and control," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 1–10, Mar. 2011.

[47] R. Jiang, R. Lu, Y. Wang, J. Luo, C. Shen, and X. Shen, "Energy-theft detection issues for advanced metering infrastructure in smart grid," *Tsinghua Sci. Technol.*, vol. 19, no. 2, pp. 105–120, Apr. 2014.

[48] S. S. S. R. Depuru, L. Wang, and V. Devabhaktuni, "Support vector machine based data classification for detection of electricity theft," in *Proc. IEEE/PES Power Syst. Conf. Expo.*, Mar. 2011, pp. 1–8.

[49] J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and F. Nagi, "Improving SVM-based nontechnical loss detection in power utility using the fuzzy inference system," *IEEE Trans. Power Del.*, vol. 26, no. 2, pp. 1284–1285, Apr. 2011.

[50] J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and A. M. Mohammad, "Detection of abnormalities and electricity theft using genetic support vector machines," in *Proc. TENCON IEEE Region 10 Conf.*, Nov. 2008, pp. 1–6.

[51] R. R. Bhat, R. D. Trevizan, R. Sengupta, X. Li, and A. Bretas, "Identifying nontechnical power loss via spatial and temporal deep learning," in *Proc. 15th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2016, pp. 272–279.

CHIA-HUNG LIN was born in Kaohsiung, Taiwan, in 1974. He received the B.S. degree in electrical engineering from the Tatung Institute of Technology, Taipei, Taiwan, in 1998, and the M.S. and Ph.D. degrees in electrical engineering from the National Sun Yat-sen University, Kaohsiung, in 2000 and 2004, respectively.

He was a Professor with the Department of Electrical Engineering, Kao-Yuan University, Kaohsiung, from 2004 to 2017. Since 2018, he has been with the National Chin-Yi University of Technology, Taichung, Taiwan, where he is currently a Professor with the Department of Electrical Engineering. His research interests include neural network computing and its applications in power system and biomedical engineering, biomedical signal and image processing, healthcare, hemodynamic analysis, and pattern recognition.

FENG-CHANG GU was born in Taichung, Taiwan, in August 1985. He received the B.S. and M.S. degrees from the National Chin-Yi University of Technology, Taichung, in 2007 and 2009, respectively, and the Ph.D. degree from the National Taiwan University of Science and Technology, in 2013.

Since 2018, he has been with the Department of Electrical Engineering, National Chin-Yi University of Technology, where he has been an Associate Professor, since 2021. His research interests include partial discharge detection, fault diagnosis, neural network computing and its applications, and pattern recognition.

JIAN-XING WU was born in 1985. He received the B.S. and M.S. degrees in electrical engineering from the Southern Taiwan University of Science and Technology, Tainan, Taiwan, in 2007 and 2009, respectively, and the Ph.D. degree in biomedical engineering from the National Cheng Kung University, Tainan, in 2014.

He was a Postdoctoral Research Fellow of the X-Ray and IR Imaging Group, National Synchrotron Radiation Research Center, Hsinchu, Taiwan, from 2014 to 2017. He was also a Postdoctoral Research Fellow of the Department of Niche Biomedical LLC, California NanoSystems Institute, UCLA, Los Angeles, USA, from 2017 to 2018. Since 2019, he has been with the National Chin-Yi University of Technology, Taichung, Taiwan, where he is currently an Assistant Professor with the Department of Electrical Engineering. His research interests include artificial intelligence applications in electrical engineering and biomedical engineering, biomedical signal processing, medical ultrasound, medical device design, and X-ray microscopy.

CHAO-LIN KUO received the B.S. degree from the Department of Automatic Control Engineering, Feng Chia University, Taichung, Taiwan, in 1998, the M.S. degree from the Institute of Biomedical Engineering, National Cheng Kung University, Tainan, Taiwan, in 2000, and the Ph.D. degree from the Department of Electrical Engineering, National Cheng Kung University, in 2006.

He was an Associate Professor at the Institute of Maritime Information and Technology, National Kaohsiung Marine University, Kaohsiung, Taiwan, from 2011 to 2017. Since 2017, he has been with the National Kaohsiung University of Science and Technology, Kaohsiung, where he is currently a Professor with the Department of Maritime Information and Technology. Since 2018, he has also been the Chief of the Department of Maritime Information and Technology. His current research interests include artificial intelligence applications in electrical engineering and ocean engineering, intelligent control systems, fuzzy systems, and embedded systems and its applications.