Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
2326-9865

Energy Theft Detection and Preventive Measures for IoT Using Machine Learning

Mali H. Alameady[#1], Salah Albermany [*2], Loay E. George [#3]
[#1,*2]University of Kufa, Najaf, Iraq
[#3] Technology and Communication, (UoITC), Baghdad, Iraq
maali.alameedi@uokufa.edu.iq, salah.albermany@uokufa.edu.iq, loayedwar57@uoitc.edu.iq

Issue: *Special Issue on Mathematical Computation in Combinatorics and Graph Theory in Mathematical Statistician and Engineering Applications*

Article Info
Page Number: 155 - 168
Publication Issue:
Vol 71 No. 3s3 (2022)

Article History
Article Received: *30 April 2022*
Revised: *22 May 2022*
Accepted: *25 June 2022*
Publication: *02 August 2022*

Abstract

Electricity theft is a big challenge for utilities. The smart grid infrastructure collects and sends out a huge amount of data, Using this data, algorithms for machine learning and deep learning may be able to reliably detect electrical thieves. Automatic power theft detection was carried out using CNNs, and convolution neural networks. CNN with some layers to extract features map by convolution layer and is classified by a softmax layer. State Grid Corporation of China (SGCC) dataset is used, and it handles missing values by linear interpolation, removes an empty record from a dataset, and orders the dataset by date. Then, the slope, average, and moment for each month are determined and entered into the CNN model. Finally, the attained computed accuracy reached 100, which is encouraging compared to previously published classification systems

Keywords: Deep Learning, Machine Learning, Internet of Things, and energy theft.

1. Introduction

The power systems are under high pressure to provide a stable and reliable electricity supply. New consumption models (smart plug-in vehicles and smart houses) are predicted to increase power demand by 30 % by 2035, according to the US Department of Energy [1,2]. The nation's primary source of energy is electricity. However, electricity theft is a problem that makes it difficult to maintain a reliable supply of electric power [3].

Recent advancements in communication technology, such as the Internet of Things (IoT), have greatly advanced the practice of environmental sensing. IoT technologies have the potential to gather, quantify, and comprehend the environment around them, enabling modernizations that enhance the living quality [4]. Using Machine Learning, computers and smart devices may learn from their knowledge and human-generated data. An IoT solution may also be characterized as the capacity of a smart device to change or automate the environment or behavior based on knowledge. ML techniques have been used in classification, regression, and density estimation tasks. We can gain a lot from IoT's enormous volumes of data. As a result, machine learning (ML) may play an

important role in designing intelligent IoT systems and delivering smart IoT services. The Internet of Things (IoT) generates large amounts of data from various physical things.

On the other hand, physical devices generate enormous amounts of data that must be examined in real-time to provide relevant insights. Researchers have studied many techniques of merging big data analytical methodologies with IoT architecture to gain insights from this data. Unlike traditional analytical approaches, big data can be efficiently analyzed using machine learning and deep learning (ML and DL) without human intervention [5-7].

The authors in [8] proposed a detection technique based on the Convolutional Neural Networks (CNN) to increase the accuracy of ETD by artificial feature extraction. Also, a combination of a long short-term memory (LSTM) architecture and CNN has been developed in [9] by using a synthetic minority over-sampling technique (SMOTE). This Non-Technical Loss (NTL) may cause significant problems in the power system because of the Electricity Theft Detection (EDT) algorithms. H. M. Rouzbahani and etc. (2020). A solution has been proposed for using an Ensemble Deep Convolutional Neural Network (EDCNN) algorithm for ETD in smart grids [10]. Bhat et al. investigated three deep learning methods for detecting electricity theft: convolutional neural networks (CNNs), long short-term memories (LSTMs), Loaded autoencoders, and recurrent neural networks. Synthetic data were used to investigate the detector operation. However, the performance of detectors with shallow designs could not be reliably evaluated. Due to the detectors being tested with shallow designs [1], this is what happened. An extensive and detailed CNN model, ETD for usage in SGs, was provided by Zheng et al. (2018). When it comes to detecting power theft, most currently used approaches are inaccurate because they rely on 1-D data from electricity use and fail to disrupt the periodicity of electricity use [12]. Ibrahim Noor and others (2021), When it comes to determining which configuration of the sequential model (SM) is best for categorizing and identifying cases of electricity theft, this study relies on experimentation. Combining the first layer of 128 nodes with the second layer of 64 nodes has resulted in an ideal performance level. The accuracy was 0.92 [13]. Shuan Li, etc. (2019) used a convolutional neural network (CNN) first and foremost; it was intended to understand the characteristics between various hours of the day and different days by using the operations of convolution and down-sampling on the massive and variable data that was collected from smart meters. Dropout layers and the back-propagation method are used to delay over-fitting and alter network parameters during training [14]. Hussain F. & etc. (2020) presented a study of the security requirements, attack vectors, and existing security solutions for IoT networks. Then, we illuminate the holes in these security solutions that need ML and DL techniques. In-depth, we also explore the current ML and DL techniques for solving various IoT network security issues. Finally, based on the comprehensive evaluation of the previously released solutions, we suggest potential future research approaches for ML- and DL-based IoT security [15]. Hu W. & etc. (2020) proposed recognizing electricity-theft behavior via multi-source data. Non-technical loss (NTL) and temperature in the transformer area are used to examine user behavior and energy use information. Analytical studies reveal many fascinating patterns: Electricity thieves, for example, are more prone to utilize electricity than the average

person, especially in extreme temperatures [16]. Finardi P. & etc. (2020) Propose a self-attention mechanism model to solve the detection of electric theft in the State Grid Corporation of China's unbalanced real-world dataset. Convolutions of kernel size 1 concatenated with dilated convolutions provide a multi-head self-attention mechanism. To help the network learn how to deal with missing values, we have added a binary input channel (Binary Mask) to the system. AUC of 0.926 is achieved by our model [17]. Rouzbahani H. M. (2020) It has been proposed that the Ensemble Deep Convolutional Neural Network (EDCNN) algorithm be used for ETD in smart grids. The model's first layer employs a random under bagging strategy to deal with the unbalanced input. After that, DCNNs (Deep Convolutional Neural Networks). In the last section, there is an incorporated voting mechanism. Area Under the Curve (AUC) is used to evaluate the accuracy, precision, recall, and f1-score [18].

The following sections of this work are structured as follows. Overviews of IoT are presented in section 2. Section 3 describes Energy Theft Detection. Section 4 describes deep learning. Section 5 describes a proposed method. The results are discussed in Section 6, which may be seen below. In the last part of the paper, which is Section 8, we will conclude.

2. Overviews of IoT

Near-field communication (NFC), RFID and GPS sensors, weather monitors, and emergency sirens are just a few examples of the many context-aware products and technologies that make up this new generation of connected gadgets. These Internet of Things (IoT) devices are constantly collecting, processing, and exchanging data. Many systems frequently monitor, connect, and interact with real-time communication. " In addition, these IoT devices record essential data such as sound, light intensity, temperature readings, power consumption, mechanical motions, a chemical response to impact, biological changes, and geo-location. They are all part of the Internet of Things. It is also possible to use IoT devices to link machines, connect machines and connect machines to humans [19].

3. Energy Theft Detection

Technical and non-technical losses are the two types of energy waste in power distribution networks. Copper losses, dielectric losses, and induction with radiation losses are power dissipation in electrical system components. On the other hand, non-technical (NT) losses describe energy losses incurred by utility companies that are not directly related to the electricity grid [20]. Tapping the power cables and tampering with meters are the most common methods of NT losses. Electric utilities suffer significant financial harm due to this widespread theft of electricity. Non-technical losses cost utilities an estimated 96 billion dollars a year [21].

Three types of energy theft exist physical, cybernetic, and digital theft. Physical assaults include tampering with the meter, disconnecting the meter, and reversing or bypassing the meter to lower recorded energy usage. Cyber-attacks involve hacking into computerized energy metering systems or remotely intercepting communication lines to alter or delete energy data. Finally, any assault on

the meter data that might compromise its integrity is considered a data breach. For example, a big load appliance may be removed from the data collection, or the energy data may be slashed in half or zeroed out entirely. As a result of these attacks, power utilities confront considerable hurdles.

4. Machine Learning

The machine learning algorithms may be divided into four groups: supervised, unsupervised, semi-supervised, and reinforcement learning algorithms. The breakdown of each of these categories can be found below [22-24]:

- **Supervised Learning**: This is called supervised learning when specific goals are set to reach certain inputs. For this type of learning, the data is first labelled and then used to train (having inputs and desired outputs).
- **Unsupervised Learning**: In unsupervised learning, the environment only gives the learner inputs, not goals. It does not need data to be labelled and can look for similarities in data that are not labelled and put them into different groups.
- **Semi-supervised Learning**: In the first two types, either none of the observations in the dataset have labels or all of the observations have labels. In between these two types of learning is semi-supervised learning.
- **Reinforcement Learning**: In Reinforcement Learning (RL), no specific results are arranged, and the agent learns from feedback after interacting with the environment. It does some tasks and then decides what to do based on its reward.

Deep Learning (DL), an artificial neural network (ANN)-derived machine learning approach Neurons (variables) are connected via weighted connections in the neural network (considered parameters). A learning approach, either supervised or unsupervised, is used to obtain the required outputs. An iterative process is used to adjust the weights between each pair of neurons in a supervised or unsupervised learning strategy, followed by learning. Therefore, while discussing deep learning, we use the term "deep" to refer to the number of layers in the network [25, 26]. DL is well-known for its ability to process, learn from, and analyze massive amounts of unlabelled, uncategorized, or unsupervised data in a distributed manner. A layered learning and feature representation model is generated based on how the human brain learns [27].

5. Proposed Method

The proposed CNN model is designed with many layers to classify the dataset. As a result, the dataset has many problems, such as missing values, which handles by pre-processing stage, feature extraction stage, and a classification stage. These stages explain as follows:

A. Pre-processing Stage:

For each recorded represented day reading $(R(i))$ that has missing values, it is handled as the followings:

1. Order the dataset by date.
2. Remove all empty records from the SGCC dataset. Find 5 empty records, where empty record numbers are $[421.686.3030.37832.41000]$.
3. Find the closest $M1$ non-empty readings at the before side $\{(Rb(1), Rb(2), \dots Rb(M1)\}$ whose days numbers are $\{posb(1), posb(2). \dots . posb(M1)\}$, where $M1$ is all cells non-empty before empty reading.

4. Determine the centroid value of the closest non-empty nearest neighbor and its position:

$$C_b = \frac{1}{M1}\sum_{i=1}^{M1} R_b(i). \qquad P_b = \frac{1}{M1}\sum_{i=1}^{M1} Pos_b(i) \tag{1}$$

4. Find the closest $M2$ non-empty readings at the after side $\{(Ra(1), Ra(2), \dots Ra(M2)\}$ whose days numbers are {*posa(1), posa(2),…,* $posa(M2)$}, where $M2$ is all cells non-empty after empty reading.
5. Determine the centroid value of the closest non-empty nearest neighbor and its position:

$$C_a = \frac{1}{M2}\sum_{i=1}^{M2} R_a(i), \qquad P_a = \frac{1}{M2}\sum_{i=1}^{M2} Pos_a(i) \tag{2}$$

6. Apply linear interpolation the assess the missing value $R(i)$ using the centroids points:

$$R(i) = \frac{C_a - C_b}{P_a - P_b}(i - P_b) + C_b \tag{3}$$

7. Delete records that contain more than ten empty cells.

B. Feature Extraction Stage

After hands missing values, we extract features for each reading $(R(i))$.

1. Compute the size of the recorded, where $[n\ N] = size(R)$;
2. Divided each record into 30 days, where $Pre = floor(N/30)$.
3. $T = 1{:}\ N$. Where $T(1..N)$ is the day's numbers.
4. Compute Slope is the gradient, and Yo is the average of readings, as follows:

For $i = 1\ to\ n$

Set F as empty

For $j = 1\ to\ N - Pre\ increasing\ by$ 30

initialize the values are:

$SmR = SmT = SmRT = SmT2 = 0.$

$SmR = SmR\ +\ sum(R(i.\ j{:}j + Pre - 1))$

$SmT = SmT\ +\ T(1.j{:}j + Pre - 1)$

$SmRT = SmRT\ +\ R(i.j{:}j + Pre - 1).*\ T(1.j{:}j + Pre - 1)$

$$SmT2 = SmT2\ +\ T(1.j{:}j + Pre - 1).*\ T(1.j{:}j + Pre - 1)$$

The value of $Slope$ is:

$$Slope = (SmRT * N - SmR * SmT)/(SmT2 * N - SmT * SmT)$$

The value of Yo is:

$$Yo = (SmT2.SmR - SmRT * SmT) / (SmT2 * N - SmT * SmT)$$

Set the Feature vector as:

$$F = [SmR/N \;\; Slope \;\; Yo]$$

End for j

Store Feature vector $f(i,:) = F$

End for i

C. Classification Stage

This stage includes many layers are Convolutional Layer, $Relu$ layer, pooling layer, and $softmax$ layer.

1. Convolutional layer: it computes the feature map as follows:

$$\mathrm{OFea(x.y.f)} = \mathrm{AF}\left(\sum_{\mathrm{v=0}}^{\mathrm{c}}\sum_{\mathrm{i=0}}^{\mathrm{k}}\sum_{\mathrm{j=0}}^{\mathrm{k}}\mathrm{A(x+i.y+j.v)}\times\mathrm{W(i.j.v.f)}+b(\mathrm{f})\right) \quad (24)$$

Where $i.j\ the$ index of filter, v number of channels, f number of filters, AF is the activation function.

The size of the output volume is set by three hyper-parameters: the depth is a number of filters, stride (S slide the filter), and zero-padding (P zeros around the border). Calculating the number of neurons is as follows:

$$H = W = \frac{n - f + 2 \times P}{S} + 1 \quad (25)$$

Where f size of filter and n size of the input image.

$$p = \frac{f - 1}{2} \quad (26)$$

2. $ReLU$: is Rectified Linear Unit. The $ReLU$ activation function performs depending on the zero thresholds.

$$f(x) = \max(0, x) \quad (27)$$

3. Pooling layers: it reduces the number of parameters. Pooling can be different: Max, mini, and average pooling. Max pooling takes the maximum element from the feature map, Mini takes the minimum element from the feature map, and average takes the average element from the feature map.

$$D^{(L)} = D^{(L-1)} \quad (28)$$

$$H^{(L)} = \frac{H^{(L-1)} - F^{(L)}}{S^{(L)}} + 1 \quad (29)$$

$$W^{(L)} = \frac{W^{(L-1)} - F^{(L)}}{S^{(L)}} + 1 \quad (30)$$

Where D is the depth of filters; H & W is the high and width of images, L is a layer, *L-1* is the previous layer, $F^{(L)}$ size of filters, and $S^{(L)}$ Stride.

4. Fully Connected Layer: is ANN, its input is 1D-array, where Flattening is converted data into a vector, all neurons in the layer have full connections to all nodes in the previous layer.
5. $Softmax$ layer: is used probabilities associated with many classes, where probabilities summation equal one. Compute $softmax$ layer as follow:

$$Softmax\ (y_i) = \frac{e^{y_i}}{\sum_j e^{y_j}} \quad (31)$$

6. Classification layer: it takes input value from the $Softmax$ layer and assigns into one class using the cross-entropy function.

The procedures of the proposed CNN model are as follows:

- Divide the dataset into 80% $training$. 10% $validation$. and 10% $testing$.
- DCNN layers are the input layer, Convolution Layer, Batch Normalization, $Relu$, Maximum Pooling, Fully Connected, Dropout Layer, $Soft_{max}$. Classification Layers.
- Convolution Layer with size $(3x3)$, and the number of filters is 200.
- Batch Normalization Layerelu Layer
- Maximum Pooling Layer with the size is 2; a stride is 2
- Convolution Layer with size $3 * 3$. and number of filters is 200
- $Relu$ Layer
- Maximum Pooling Layer with the size is 2, and a stride is 2
- Convolution Layer with size $3 * 3$, and number of filters is 200
- $Relu$ Layer
- Maximum Pooling Layer with the size is 2, and a stride is 2
- The fully Connected Layer is 200
- The dropout Layer is 0.5
- The fully Connected Layer is 2
- $Softmax$ Layer
- Classification Layer
- Define Training Parameters are ADAM training method,
- Mini Batch size equals 256
- Max epochs equal 10
- Initial Learn Rate equals 0.0003

- Train Network
- Extract the class labels from the test data by classification layer
- Measures predicated accuracy

6. Results Discussion (Measures)

The proposed algorithm has been tested on an SGCC real data set collected by a smart meter and found some measures [24]:

- Accuracy (Ac) is the rate of correctly classified samples to overall samples.

$$\text{Ac} = \frac{TP + TN}{TP + FP + FN + TN} \quad (33)$$

- Precision (P) is the rate of true positive samples to predicted positive samples.

$$P = \frac{TP}{TP + FP} \quad (34)$$

- Recall (R) is the rate of true positive samples to total positive samples.

$$R = \frac{TP}{TP + FN} \quad (35)$$

- F-measure (F) is the average of the precision and the recall.

$$F = \frac{2 * P * R}{P + R} \quad (36)$$

- False Negative Rate (FNR) is the rate of false-negative samples to total positive samples.

$$FNR = \frac{FN}{TP + FN} \quad (37)$$

- False Positive Rate (FPR) is the rate of false-positive samples to predicted positive samples.

$$FPR = 1 - precision(P) \quad (38)$$

Where TP is the true positive, FP is the false positive, TN is the true negative, and FN is the false negative.

7. SGCC Dataset and Results

This section explains dataset properties, where electricity theft data; the study was based on a collection of real electricity usage data from consumers that the State Grid Corporation of China made available (http://www.sgcc.com.cn/). In Table (1), the metadata about the dataset is shown. The dataset is two years and ten months. Smart meters or other sensors at the user end are usually used to get information about how much electricity is used. The data are then sent to a central location through a data communication network. In this case, there is a chance that the smart meter will not work, that the sensor will not work, or that the data transmission and storage server will go wrong. Therefore, there will always be missing or wrong information in the datasets about how much electricity is used. We found a lot of missing data in this set. If those missing instances are just thrown away, the size of the dataset goes down a lot, making it hard to do reliable analysis. Therefore, to avoid downsizing the dataset.

Table (1) Metadata information of the electricity theft dataset.

Description	Value
Time window of data collection	1 January 2014–31 October 2016
Total number of consumers	42 372
Number of normal users	38 757 approx. 91.5%
Number of aberrant users or electricity thieves	3 615 approx. 8.55%
Missing data cases approx.	25.6393%

The dataset is divided into a training ratio of 80%, a Testing ratio of 10%, and a Validation ratio of 10%. Tables (2-5) show the affected learning rate, min-batch, training method, and epoch parameters respectively on the training dataset. Figure (1) shows the confusion matrix of model classification. Figure (2) shows the training process and loss rate for the proposed classification model. Finally, table (6) show the measures of the training dataset.

Table (2): Effect of Learning Rate Parameter on the training Dataset.

Learning Rate	0.0003	0.3	0.04	0.0001	0.00000003
Accuracy	100	99.96	99.7	100	91.9

Table (3): Effect min-batch parameter on the training dataset

Mini-batch	64	128	256	512	1024
Accuracy	97.9	100	100	100	100

Table (4): Effect of training method parameter on the training dataset.

Training method	SGDA	RMSPROP	ADAM
Accuracy	99.92	100	100

Table (5): Effect epoch parameter on the training dataset.

Epoch	1	2	10	20	30	50
Accuracy	99.8	100	100	100	99.96	99.96

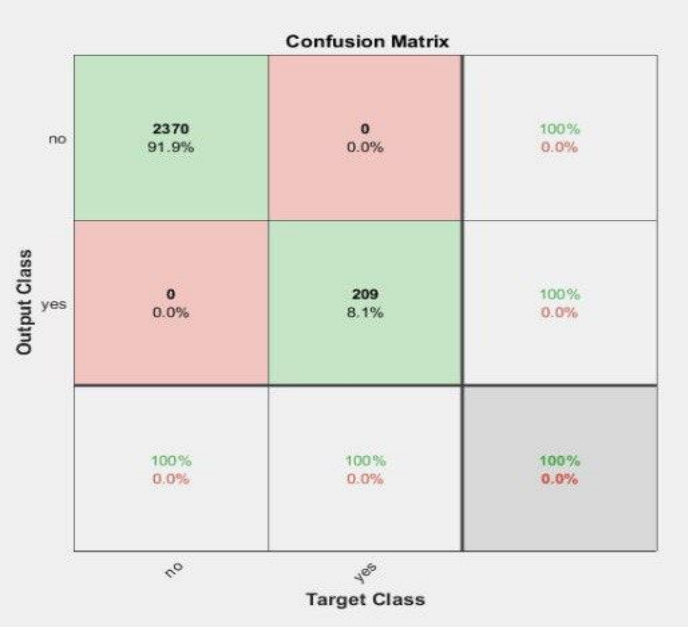

Figure (1): confusion matrix.

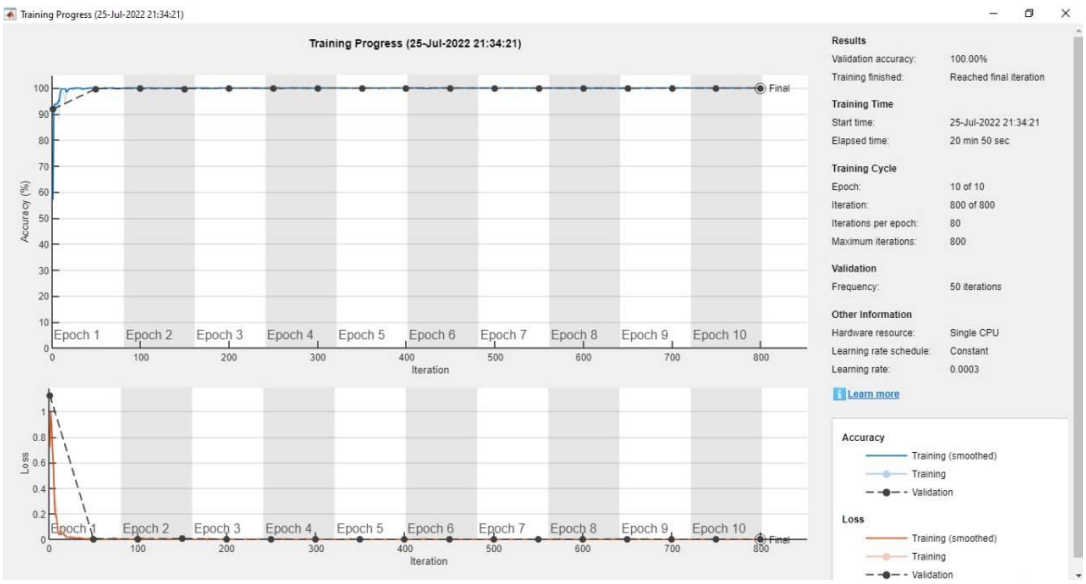

Figure (2): training process and loss rate.

Table (6): Measures of the training dataset.

Parameters	**Training method =ADAM; Mini¬ Batch=256; Epochs=10; Learn Rate=3e-4**
Accuracy	100
Precision	100
Recall	100
F-measure	100
False Negative Rate	0

False Positive Rate	0

8. Comparison with previous work

This section explains the comparison with other works as shown in table (7).

Table (7): Summary of models, and performance measures.

Method	Year	Accuracy
Proposed DCNN model	2022	100
self-attention mechanism model [29]	2020	0.926
convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) [30]	2019	89
combined convolutional neural networks (CNNs) [31]	2019	92.67
hybrid DL model [32]	2022	88
CNN-GRU-PSO [33]	2020	87
CNN [34]	2019	93
Hybrid 2DCNN and BiLSTM [35]	2022	97
Ensemble Deep Convolutional Neural Network (EDCNN) algorithm.[36]	2020	99
autoencoder-bidirectional gated recurrent unit (AE-BiGRU) model [37]	2022	91.1
Autoregressive Integrated moving average (ARIMA). In the second stage, the distributed random forest (DRF) [38]	2022	98

9. Conclusion

According to this study, supervised learning approaches are superior to unsupervised learning methods since labeled data is available for training models. Furthermore, because these models are developed using large datasets and sophisticated computers, they have great power in handling electricity consumption data. This work used DCNN to classify the energy theft dataset, where the system runs with high accuracy equals 100; compared with previous works.

References

[1] H. Karimipour and V. Dinavahi, “Parallel domain decomposition based distributed state estimation for large-scale power systems,” IEEE Transactions on Industry Applications, vol. 2015, 2015.

[2] H. M. Ruzbahani and H. Karimipour, “Optimal incentive-based demand response management of smart households,” in Conference Record - Industrial and Commercial Power Systems Technical Conference, 2018, vol. 2018-May, pp. 1–7.

[3] Hung-Po Chao and Stephen Peck. 1996. A market mechanism for electric power transmission. Journal of regulatory economics 10, 1 (1996), 25–59.

[4] A. V. Dastjerdi and R. Buyya, "Fog computing: Helping the Internet of Things realize its potential," Computer, vol. 49, no. 8, pp. 112-116, 2016.

[5] C. W. Tsai, C.-F. Lai, M.-C. Chiang, and L. T. Yang, "Data mining for Internet of Things: A survey," IEEE Communications Surveys and Tutorials, vol. 16, no. 1, pp. 77-97, 2014.

[6] Ahmed E., Yaqoob I., Hashem, I. A. T. Khan, I. Ahmed, A. I. A. Imran, M., & Vasilakos A. V., "The role of big data analytics in Internet of Things," Computer Networks, vol. 129, pp. 459-471, 2017.

[7] S. G. Leem, I. C. Yoo, and D. Yook, "Multitask Learning of Deep Neural Network-Based Keyword Spotting for IoT Devices", IEEE Transactions on Consumer Electronics, vol. 65, no. 2, pp. 188-194, 2019.

[8] 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 Transactions on Industrial Informatics, vol. 14, no. 4, pp. 1606–1615, Apr. 2018.

[9] Md. N. Hasan, R. N. Toma, A.-A. Nahid, M. M. M. Islam, and J.-M. Kim, “Electricity Theft Detection in Smart Grid Systems: A CNN-LSTM Based Approach,” Energies, vol. 12, no. 17, p. 3310, Aug. 2019.

[10] H. M. Rouzbahani, H. Karimipour and L. Lei, "An Ensemble Deep Convolutional Neural Network Model for Electricity Theft Detection in Smart Grids," 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Toronto, ON, 2020, pp. 3637-3642.

[11] R. R. Bhat, R. D. Trevizan, R. Sengupta, X. Li, A. Bretas, “Identifying Nontechnical Power Loss via Spatial and Temporal Deep Learning,” 15th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 272-279, 2016, doi: 10.1109/ICMLA.2016.0052.

[12] Z. Zheng, Y. Yang, X. Niu, H. Dai, Y. Zhou, “Wide and Deep Convolutional Neural Networks for Electricity-Theft Detection to Secure Smart Grids,” in IEEE Transactions on Industrial Informatics, vol. 14, no. 4, pp. 1606-1615, April 2018, doi: 10.1109/TII.2017.2785963.

[13] Ibrahim Noor, Sufyan Al-Janabi, and Belal Al-Khateeb. "Electricity-Theft Detection in Smart Grid Based on Deep Learning." Bulletin of Electrical Engineering and Informatics 10.4 (2021): 2285-2292.

[14] Shuan Li, Yinghua Han, Xu Yao, Song Yingchen, Jinkuan Wang, and Qiang Zhao, Electricity Theft Detection in Power Grids with Deep Learning and Random Forests, Journal of Electrical and Computer Engineering Volume 2019, Article ID 4136874, 12 pages.

[15] Hussain F., Hussain R. Hassan, S. A. & Hossain E.," Machine learning in IoT security: Current solutions and future challenges". IEEE Communications Surveys & Tutorials, 22(3), 1686-1721, 2020.

[16] Hu W., Yang, Y. Wang, J. Huang, X. & Cheng Z.," Understanding electricity-theft behavior via multi-source data". In Proceedings of The Web Conference 2020 (pp. 2264-2274), 2020, April.

[17] Finardi P., Campiotti I., Plensack G., de Souza R. D., Nogueira R., Pinheiro, G. & Lotufo R.," Electricity theft detection with self-attention", arXiv preprint arXiv:2002.06219, 2020.

[18] Rouzbahani H. M., Karimipour H., & Lei L. ". An ensemble deep convolutional neural network model for electricity theft detection in smart grids". In 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 3637-3642). IEEE, 2020.

[19] K. Zhao and L. Ge, "A Survey on the Internet of Things Security," in 2013 Ninth International Conference on Computational Intelligence and Security, 14-15 Dec. 2013 2013, pp. 663-667, doi: 10.1109/CIS.2013.145.

[20] Tanveer Ahmad. Non-technical loss analysis and prevention using smart meters. Renewable and Sustainable Energy Reviews, 72:573–589, 2017.

[21] Sonal Jain, Kushan A Choksi, and Naran M Pindoriya. Rule-based classification of energy theft and anomalies in consumers load demand profile. IET Smart Grid, 2(4):612–624, 2019.

[22] A. B. YILMAZ, Y. S. TASPINAR, and M. Koklu, “Classification of Malicious Android Applications Using Naive Bayes and Support Vector Machine Algorithms”, Int J Intell Syst Appl Eng, vol. 10, no. 2, pp. 269–274, May 2022.

[23] Alpaydin, E., " Introduction to machine learning". MIT Press, 2009.

[24] Géron A., "Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems. " O'Reilly Media, Inc.", 2017.

[25] Kaelbling L. P., Littman M. L., & Moore A. W., "Reinforcement learning: A survey", Journal of artificial intelligence research, 4, 237-285, 1996.

[26] Yang L., Chen Y., Li X. Y., Xiao C., Li M., & Liu Y., “Tagoram: Real-time tracking of mobile RFID tags to high precision using COTS devices,” ACM international conference on Mobile computing and networking, vol. 1, pp. 237–248, 2014.

[27] Lane N. D., Georgiev P., & Qendro L., “DeepEar: robust smartphone audio sensing in unconstrained acoustic environments using deep learning,” ACM International Conference on Pervasive and Ubiquitous Computing, vol. 1, pp. 283–294, 2015.

[28] T. Wng, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, “Deep Learning for Wireless Physical Layer: Opportunities and Challenges,” IEEE China Communication, vol. 14, pp. 92–111, October 2017.

[29] Amato G., Carrara F., Falchi F., Gennaro C., Meghini C., Vairo C., “Deep learning for decentralized parking lot occupancy detection”, Expert System, 72, 327–334, 2017.

[30] Finardi, P., Campiotti, I., Plensack, G., de Souza, R. D., Nogueira, R., Pinheiro, G., & Lotufo, R.. "Electricity theft detection with self-attention". arXiv preprint arXiv:2002.06219., 2020.

[31] Hasan, M. N., Toma, R. N., Nahid, A. A., Islam, M. M., & Kim, J. M.," Electricity theft detection in smart grid systems": A CNN-LSTM based approach. Energies, 12(17), 3310., 2019.

[32] Yao, D., Wen, M., Liang, X., Fu, Z., Zhang, K., & Yang, B., "Energy theft detection with energy privacy preservation in the smart grid", IEEE Internet of Things Journal, 6(5), 7659-7669., (2019).

[33] Ullah, A., Javaid, N., Asif, M., Javed, M. U., & Yahaya, A. S.," AlexNet, AdaBoost and Artificial Bee Colony Based Hybrid Model for Electricity Theft Detection in Smart Grids". IEEE Access, 10, 18681-18694., 2022.

[34] Ullah, A.; Javaid, N.; Samuel, O.; Imran, M.; Shoaib, M. CNN and GRU based deep neural network for electricity theft detection to secure smart grid. In Proceedings of the 2020 International Wireless Communications and Mobile Computing (IWCMC), Limassol, Cyprus, 15–19 June 2020; pp. 1598–1602.

[35] Yao, D.; Wen, M.; Liang, X.; Fu, Z.; Zhang, K.; Yang, B. Energy theft detection with energy privacy preservation in the smart grid. IEEE Internet Things J. 2019, 6, 7659–7669.

[36] Asif, M., Nazeer, O., Javaid, N., Alkhammash, E. H., & Hadjouni, M. (2022). Data Augmentation Using BiWGAN, Feature Extraction and Classification by Hybrid 2DCNN and BiLSTM to Detect Non-Technical Losses in Smart Grids. IEEE Access, 10, 27467-27483.

[37] Rouzbahani, H. M., Karimipour, H., & Lei, L. (2020, October). An ensemble deep convolutional neural network model for electricity theft detection in smart grids. In 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 3637-3642). IEEE.

[38] Garg, D. K. . (2022). Understanding the Purpose of Object Detection, Models to Detect Objects, Application Use and Benefits. International Journal on Future Revolution in Computer Science &Amp; Communication Engineering, 8(2), 01–04. https://doi.org/10.17762/ijfrcsce.v8i2.2066

[39] Javaid, N., Qasim, U., Yahaya, A. S., Alkhammash, E. H., & Hadjouni, M. (2022). Non-technical losses detection using autoencoder and bidirectional gated recurrent unit to secure smart grids. IEEE Access.

[40] Badawi, S. A., Guessoum, D., Elbadawi, I., & Albadawi, A. (2022). A Novel Time-Series Transformation and Machine-Learning-Based Method for NTL Fraud Detection in Utility Companies. Mathematics, 10(11), 1878.