Islamic Azad Univer Sity Firuzkuh Branch
Journal Of Industrial Strategic Management
https://sanad.iau.ir/journal/mgmt Volume 9, Issue 1, Spring 2025, 97-113

Research Article

A Hybrid Technique for Estimating and Forecasting Household Electrical Energy Consumption Utilizing Machine Deep Learning and Fuzzy Wavelet

Abdolreza Rahmanifar[1] **Mehran Khalaj**[2] **Ali Taghizadeh Harat**[3] **Asghar Darigh**[4]

1. Department of Industrial Engineering, Pa.c., Islamic Azad University, Parand, Iran
   Email: Arahmanifar96@Piau.ac.ir
2. Department of Industrial Engineering, Pa.c., Islamic Azad University, Parand, Iran,
   Email**:** M.Khalaj@Piau.ac.ir **(Corresponding Author)**
3. Faculty MSc in Mathematical Statistics, Semnan University, Iran,
   Email: Taghizadeh@Piau.ac.ir
4. Department of Industrial Engineering, Pa.c., Islamic Azad University, Parand, Iran
   Email**:** A.Darigh@Piau.ac.ir

ABSTRACT

DOR: https://doi.org/10.71584/MGMT.2025.1207038

Article Info:

Received Date:
2025/05/16
Accepted Date:
2025/07/12

Published Online:
2025/08/30

Abstract:
This study investigates the electrical consumption of households as a microcosm of a macro society, with the individual appliances inside each home serving as the "electricity consuming units. The goal is to provide an optimal approach for addressing the issue of efficient energy usage. To accomplish this objective, it is essential to divide the total electrical consumption of the home into its component elements, which are the individual signals utilized by every appliance. Likewise, estimating the energy consumption of the appliances is a very efficient means of foreseeing how much energy each device would consume in the future and, if necessary, controlling it. In this research, a Fuzzy Wavelet- and Convolutional Network-based method is established as a way of decomposing the signals generated by individual home appliances from the overall (composite) signal. In addition, the proposed algorithm is employed in conjunction with two well-known and strong algorithms in Time-series data analysis, Long-Short Term Memory (LSTM) and Multilayer Perceptron (MLP). Hence, the proposed approach is compared to the aforementioned two renowned algorithms as well as other techniques from previous studies. The proposed neural network is trained using the Stochastic Gradient Descent (SGD) optimization approach at each stage, and the Nesterov Accelerated Gradient (NAG) optimization method is also investigated. In comparison with previous approaches, the findings demonstrate that the algorithm's prediction accuracy is greater and its error is noticeably lower. It means that the proposed algorithm is a top contender among the existing algorithms for predicting of energy consumption in residential buildings.

Keywords:
Deep Learning, Neural Network, Fuzzy Wavelet, Households

How to Cite: Rahmanifar ,A. , Khalaj ,M., Taghizadeh herat ,A.& Asghar Darigh ,A.,(2025), A Hybrid Technique for Estimating and Forecasting Household Electrical Energy Consumption Utilizing Machine Deep Learning and Fuzzy Wavelet, *Journal of Industrial Strategic Management*, 9(1): 97-113.

1. Introduction

For policymakers and allied institutions, modeling and forecasting energy consumption are crucial to a country's growth and success. When energy consumption is not taken into consideration, a power outage might occur, resulting in loss of life and economic devastation. In addition to wasting money, overestimating energy demands might result in unneeded features being implemented. Therefore, to minimize expensive errors, it is preferable to employ algorithms that forecast energy consumption with greater precision. Moreover, the models that can include non-linear energy consumption data in forecasting are preferable [1]. Research shows that Artificial Intelligence (AI) approaches are the most widely used technique for forecasting energy consumption. As a result, the metaheuristic algorithm approach is more alluring and consequential to target audiences, such as energy engineers. Since it allows for the option of developing more reliable energy applications, independent of time savings. Advantages such as fast calculation, greater affordability, simple implementation, and design by operators with minimal technical abilities are added bonuses of this technique [2, 3]. Energy is a crucial factor in almost all commercial endeavors. Most countries cannot guarantee their own safety without a steady supply of energy. Hence, the efficient generation, use, and application of energy sources in the future are of paramount importance [4].

Growing energy needs throughout the world need the creation of intelligent forecasting models and algorithms. Allocation of energy resources may be estimated and optimized utilizing economic and non-economic variables that can be derived from linear and non-linear statistical approaches, mathematics, and simulation models. Intelligent methods, including genetic algorithms, fuzzy regression, and neural networks, have been explored due to the nonlinear nature of these metrics and energy demand. In addition, nonlinear modeling and prediction employ the application of artificial neural networks [1]. While attempting to foresee future energy use, it is common practice to look at historical use patterns; these patterns in turn have connections to other elements like economy, population, climate, and so on. The widespread interest in energy modeling in recent years has focused the attention of scientists and engineers on the subject of energy generation and consumption. Several sectors of application may benefit greatly from the use of modeling in the process of establishing policies and strategies [2].

Over 40% of global power consumption and greenhouse gas emissions are attributable to buildings, according to recent studies. In truth, growing populations and higher living standards are driving forces behind the relentless increase in energy use [5, 6]. To effectively control grid loads, data on how much electricity home electrical appliances usage must be gathered. The power system stability might be threatened without understanding the energy usage of electrical devices in homes. When considering the social and cultural aspects of energy use, identifying which appliances consume the most power might assist reduce overall electricity consumption, particularly during peak usage times [7].

DSM, or "demand side management", is the practice of continuously monitoring and controlling electrical energy utilization at the end-user level. This allows planners to more effectively control and balance electrical energy generation and consumption [8]. Hart established the notion of Non-Intrusive Load Monitoring (NILM) in 1992 as a strategy to minimize electrical energy consumption [9]. With the use of a smart meter's waveform output, NILM is able to isolate and classify various consumer-side electrical loads. With this, it is no longer necessary to install smart meters for each individual part of the network, which in turn increases the network's overall cost-effectiveness and simplicity of use [10].

In this scenario, as seen in Figure 1, a cumulative meter is used for the upstream network rather than using individual meters for each device, and the waveform of each consumer is collected by decomposing the whole waveform [11]. Several researchers have attempted to address the NILM problem utilizing different techniques throughout the preceding two decades. Despite its outstanding efficiency, the machine learning (ML) approach has limitations, thus researchers often strive to eliminate computation errors by integrating it with other methods [12].

A key application of this study is the extraction of individual electrical device waveforms from the overall waveform of a residential building. By isolating and comparing these signals with standard waveforms, valuable insights can be gained about each appliance such as operational accuracy, energy consumption levels, and potential faults. Addressing any identified issues can lead to cost reductions and improved stability of the electrical grid.

The motivation behind this research is to find a method for identifying the type of electrical energy consumers from the aggregate waveform of a residential building. This enables power network planners and supervisors to implement macro-level policies for effective control and network stability, while also detecting irregular or unauthorized energy usage within the electrical grid.

1-1. Theoretical Implications

- Advancement in Non-Intrusive Load Monitoring (NILM): This research contributes to the theoretical foundation of NILM by demonstrating how aggregated electrical waveforms can be decomposed to identify individual appliance signatures, enhancing our understanding of energy disaggregation techniques.
- Signal Processing Frameworks: The study supports the development of novel signal processing models that interpret overlapping load patterns, reinforcing theories that combine time-series analysis, pattern recognition, and machine learning in electrical systems.
- Behavioral Modeling of Energy Consumers: By correlating waveform characteristics with appliance usage, the research lays theoretical groundwork for modeling consumer behavior through power consumption trends—a step toward smarter, more predictive energy analytics.
- Implications for Smart Grid Theory: Findings underscore the importance of real-time data analytics and consumer-level monitoring in maintaining grid reliability and stability, aligning with broader theoretical constructs around decentralized energy management.

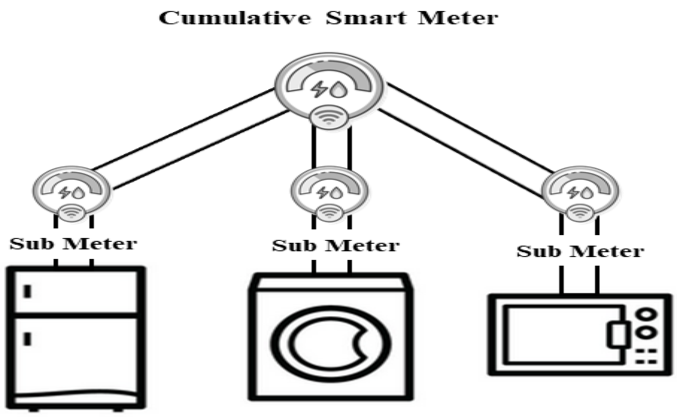

Figure 1: Cumulative meter for upstream network

1-2. Literature review

By analyzing residential sector energy consumption records and the load status decomposition method, Yin et al. [13] constructed a Gaussian mixture model with time information for probability distribution called BH-Factorial Hidden Markov Model (BH-FHMM), which was then evaluated using the REDD dataset. For a real-time, low-cost solution that can address these algorithms' challenges, Nguyen et al. [10] proposed a Non-Intrusive Appliance Load Monitoring (NIALM) method. This idea utilizes on-chip technologies to connect multiple processors simultaneously. By decomposing the complex current into independent loads and determining the switching time using a BP neural network, Wu and Lo [14] were able to identify individual loads. This strategy was used in order to modify load detection. In assisting electric energy supply companies in monitoring and analyzing household energy consumption data, Li and Dick [15] analyzed four multiple tagging algorithms to distinguish electrical equipment consumption based on a cumulative waveform. The comparison of the four algorithms allowed them to determine which one was more effective on the household dataset. For independent classifiers, Liu et al. [16] proposed a data tagging method and a non-homogeneous design and data framework connected to load decomposition. Finally, Independent loads were decomposed out of a larger group of loads using a multiple-criterion assessment based on a decision-making method. Using data on the equipment's total energy usage, Hu et al. [17] framed the load decomposition problem as an optimization problem and solved it using a genetic algorithm employing parallel supplemental computations. A multilayer artificial neural network, in addition to the previously indicated algorithm, was utilized for machine learning. Wu et al. [18] presented a technique for non-intrusive load monitoring that used a high-frequency mode to retrieve electrical data. This method is able to decompose loads automatically and in real time. Using a convolutional neural network to identify the waveforms after they have been decomposed allows for more precise load detection. To detect non-intrusive load, Wu et al. [19] proposed a multi-label classification technique using a Random Forest (RF) algorithm. These characteristics are sorted and compared based on their relative significance, and the approach is robust against load signals with no mixed signals. For the purpose of managing and decomposing load consumption from the perspective of end users, Cavdar and Faryad [20] proposed a multi-component model based on deep machine learning. The CNN-RNN model is used in conjunction with real data from residential buildings to get an estimate of the current consumption rate. Using an Artificial Neural Network and Particle Swarm Optimization (ANN-PSO) for consumer-side non-intrusive load monitoring (NILM), Lin and Hu [21] presented an Internet of Things (IoT)-based energy management system. A home system evaluation was used to assess the efficacy of this novel combination. Alotaibi [22] presented machine learning and explainable AI to predict heating and cooling loads in residential buildings. It uses data from 768 buildings and applies models like Gaussian Process Regression and Boosted Trees. The GPR-M3 model showed the highest accuracy in both heating and cooling scenarios. Results were validated using performance metrics like RMSE and PCC. The model was also tested in Ecotect software for energy simulation. Azim et al. [23] presented new artificial neural networks to predict energy use in Tabriz homes based on resident behavior. Key influencing factors include number of walls, housing direction, family size, and occupation. Regression analysis helped select input variables for the ANN model. Seasonal variations were considered in the prediction. The model achieved accurate forecasts using real consumption data. Khodadadi et al. [24] presented a hybrid deep learning model combining three CNNs and a DNN using a voting mechanism. It was trained on the WiDS Datathon dataset for residential energy prediction. The ensemble model outperformed traditional ML methods like Random Forest and Linear Regression. Results showed high accuracy and robustness against new data. The approach is scalable for future smart building applications. Neshat et al. [25] presented an adaptive ensemble learning models (Bagging, Stacking, Voting) with evolutionary hyperparameter tuning. It uses sensor data (temperature, humidity, lighting) from a smart building in Belgium. The proposed model outperformed 15 other ML models in accuracy and error reduction. It demonstrated strong performance in predicting appliance-level energy use. The method supports real-time energy management in smart homes. Alam [26] presented a stacked deep learning model combining CNN, LSTM, and DNN for energy demand forecasting. It emphasizes feature engineering and normalization for improved accuracy. The model was trained on real-world datasets and achieved high performance metrics. It supports short- and long-term forecasting for residential buildings. The framework aims to aid energy conservation and smart grid planning.

2. Methodology and Material

2-1. Energy Consumption Decomposition

2-1-1. Why Machine learning and Fuzzy wavelet

Machine learning algorithms can uncover hidden and nonlinear patterns in energy consumption data something traditional methods often miss. These models analyze past building behavior under various conditions (temperature, time of day, day of the week, etc.) to provide accurate predictions. ML models can be updated and refined with new data, meaning they improve over time. Energy consumption is often a time-varying and complex signal. Wavelet transforms are highly effective at decomposing these signals into various frequency components. Fuzzy logic allows for reliable decisions even with uncertainty or measurement errors (like consumption fluctuations). Combining wavelets for feature extraction with fuzzy logic for decision-making enhances the model’s accuracy in real-world conditions.

Home appliance energy consumption data may be decomposed by receiving a cumulative signal from the building's energy system and then separating that signal into data relevant to each appliance. As a mathematical problem, it may be stated as:

$$\mathrm{P(t)} = p_1(t) + p_2(t) + , \ldots , + p_n(t), \tag{1}$$

In the preceding equation, the variable $p_i(t)$ represents the energy used by the device over time. The goal is to decompose the overall consumption pattern (denoted by P(t)) of a household into its component parts $p_i(t)$. Hence, the total energy consumption of all machines is determined. Moreover, depending on the power consumption of each device, a customized power management scheme may be put into place.

There are two methods for calculating estimated electrical consumption. I) An average consumption estimates for each consumer's (home's) power usage; II) Micro-electricity consumption modeling to predict individual users' power usage (each electrical device in the house). Decomposing energy in this context might involve two distinct phases: the first would involve reestablishing the device's power consumption patterns upon startup. The second is how much energy the equipment used on average between the beginning and end of the task [27].

Datasets, like those of home appliance power usage, that exhibit substantial fluctuation over time are, in general, more difficult to analyze and assess in time series. In light of this, it stands to reason that sub-time series derived from the signals of the main or total time series would exhibit less noticeable variations and be simpler to analyze than the main time series itself. Hence, the wavelet approach is used to decompose the ground truth into sub-datasets and get it ready for the algorithm's analysis and training phases. Figure 2 depicts the method by which the house's overall signal is converted to the signals used by the various electrical devices located inside the home.

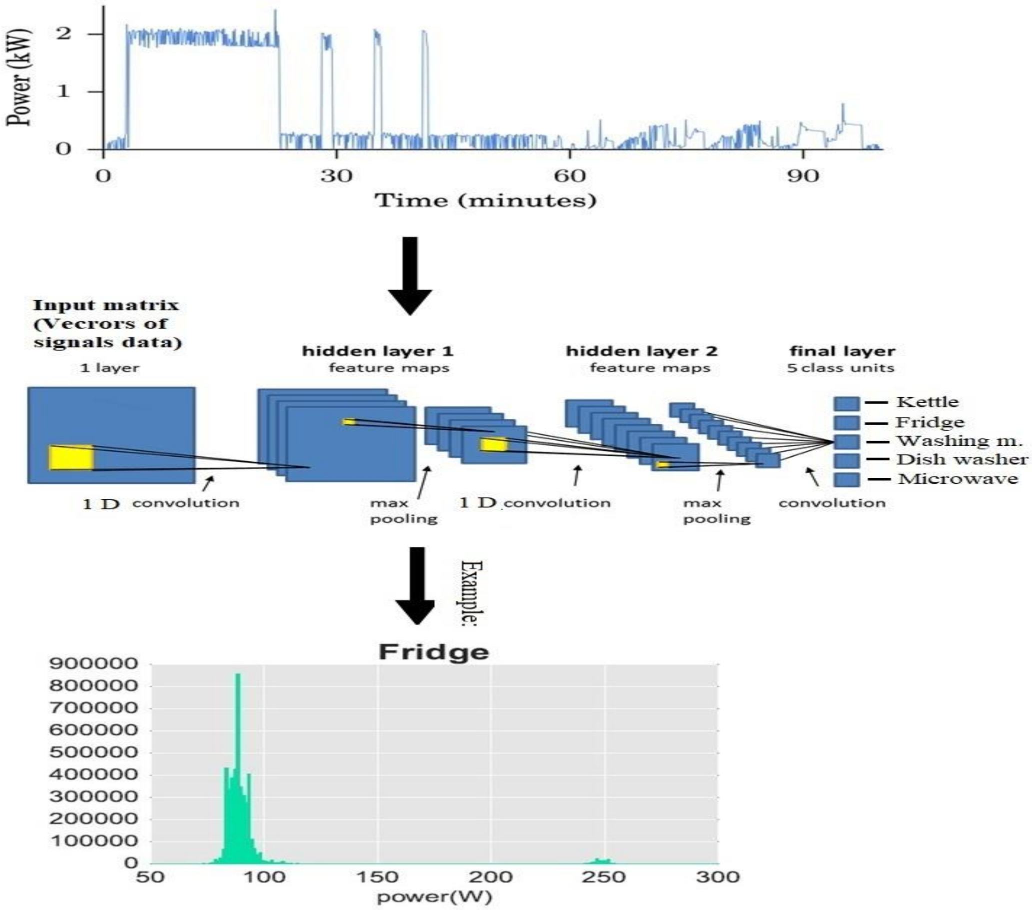

Figure 2. Converting the house's main signal to the signals used by individual home appliances.

Wavelet basically breaks the primary signals into two subgroups, one containing low frequencies and the other with high frequencies. After a series of wavelet transforms, the original signal is divided into "partial" and "approximate" sub-parts. The wave's basic tendency may be found in the approximate section, while the severe fluctuations can be found in the partial part. Daubechies wavelet methodology has been utilized for signal decomposition [28].

Figure 3 depicts the signal-breaking process. The data is first divided into its constituent parts, or "partial signals," or "D_1," and "approximate values," or "A_1." The high-frequency and low-frequency wavelets for this part are generated by further decomposing the D_1wavelet into D1-High and D1-Low wavelets. Moreover, wavelet A is decomposed into its constituent sub-wavelets, A_3, D_3, and D_2. Thereafter, the sub-wavelets (labeled A_3, D_3, D_2, $\mathrm{D}_1 - \mathrm{Low}$, and $\mathrm{D}_1 - \mathrm{High}$) are entered into the algorithm's convolution layer of a neural network.

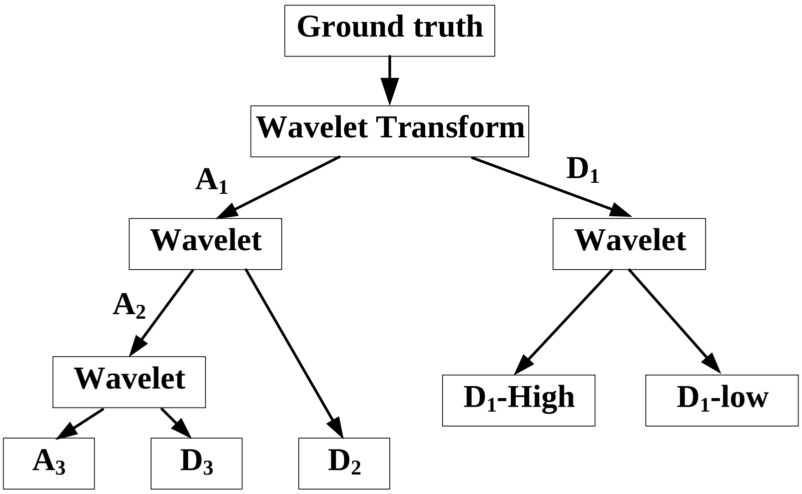

Figure 3. Sub-signal decomposition refers to the procedure of taking ground truth and breaking it down further

2-2. The primary neural networks for estimating:

In this study, the authors' custom-built estimation network was employed for all estimation-required components. The layers of this network are what make it a neural network. However, convolutional networks are two-dimensional (2D). The 2D nature of each convolutional layer is what makes these networks ideal for processing images. Unfortunately, two-dimensional convolutional layers cannot be used for time series analysis. For this reason, the convolutional network's layers have been designed in a 1-dimensional (1D) structure that is well-suited to time-series data. The TensorFlow and Keras libraries include the necessary functions for creating 1D convolutional layer.

Figure 4 depicts the neural network architecture utilized in the final step of the algorithm and the estimation of power consumption, as well as in the signal decomposition stage for each piece of equipment. The performance of the created network is compared to that of various neural networks already present in the academic literature. That is, the wavelet decomposes the original signal into smaller signals, and then the data from those signals is evaluated using Long-Short Term Memory (LSTM) and Multilayer Perceptron (MLP) networks in conjunction with the proposed convolution network algorithm.

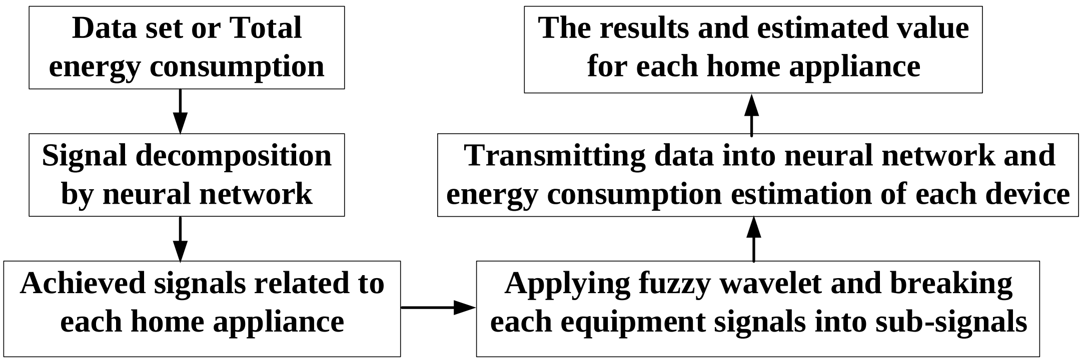

Figure 4. Decomposing a dataset into individual signals for each machine that uses energy.

As shown in Table 1, the developed convolutional network algorithm has the following structure.

Table 1. The layers of the proposed neural network

Dense Layers (hidden units = window length, activation= linear)
Dense Layers (hidden units = window length*5, activation = ReLU)
1D Convolutional Layer (filter num. = 3, filter size = 4, stride = 1, activation = ReLU)
Dense Layers (hidden units = window length*4, activation = ReLU)
Dense Layers (hidden units = window length*4, activation = ReLU)
Dense Layers (hidden units = window length*4, activation = ReLU)
Dense Layers (hidden units = window length*4, activation = ReLU)
1D Convolutional Layer (filter num. = 16, filter size = 4, stride = 1,activation = ReLU)
1D Convolutional Layer (filter num. = 32, filter size = 4, stride = 1, activation = ReLU)
1D Convolutional Layer (filter num. = 64, filter size = 4, stride = 1, activation = ReLU)
SpatialDropout1D(rate, **kwargs)
1D Convolutional Layer (filter num. = 128, filter size = 4, stride = 1, activation = ReLU)
1D Convolutional Layer (filter num. = 128, filter size = 4, stride = 1, activation = ReLU)
SpatialDropout1D(rate, **kwargs)
1D Convolutional Layer (filter num. = 256, filter size = 4, stride = 1, activation = ReLU)
1D Convolutional Layer (filter num. = 256, filter size = 4, stride = 1, activation = ReLU)
1D Convolutional Layer (filter num. = 512, filter size = 4, stride = 1, activation = ReLU)
1D Convolutional Layer (filter num. = 512, filter size = 4, stride = 1, activation = ReLU)
1D Convolutional Layer (filter num. = 1024, filter size = 4, stride = 1, activation = ReLU)
1D Convolutional Layer (filter num. = 1024, filter size = 4, stride = 1, activation = ReLU)

2-3. FWCNN Formulation

After the individual components of the fuzzy wavelet convolutional neural network (FWCNN) technique have been outlined, the underlying mathematical equation may be formulated as follows:

$$x_i^{(l+1)} = (\omega_d)_i^{(l+1)}(y_d)^{(l)} + (\omega_f)_i^{(l+1)}(y_f)^{(l)} + b_i^{(l+1)}, \tag{2}$$

Time-series data computation is described by the equation (2), in which y_d and y_f refer to the output part of the deep representation, respectively. Weights of ω_d and ω_f also denote the output of the fuzzy representation. Then, the nonlinear function modifies the reaction result in the reaction layer. As such, the following expression defines the value predicted at period $t - th$.

$$\hat{y} = g\left(x_i^{(l+1)}\right) = \frac{1-e^{-2x_i^{(l+1)}}}{1-e^{-2x_i^{(l+1)}}} \tag{3}$$

Where $x_i^{(l+1)}$ refers to the combined result of both fuzzy and neural representations. In addition, the Hyperbolic Tangent ($Tanh$) of the activation function is defined as g. $Tanh$ ensures that the output values range between -1 and1. That is, close to the normalized input values.

By minimizing the mean squared error between the predicted and actual values, the FWCNN model may be trained to predict energy consumption from input series data.

The reconstruction error can be achieved as follows:

$$L(\theta) = \|y_t - \hat{y}_t\|_2^2 \tag{4}$$

In whichy_t, $\hat{y}_t$ and θ describe observed value, predicted value, and all the learnable parameters in the FWCNN model, respectively. Before making any predictions, it is essential to develop the FWCNN model by establishing initial values for its parameters and then adjusting it to perfection. The convergence of the neural network to a desirable minimum may be aided by greater preparation. Parameterization of the FWCNN model involves both the FN and CNN. The CNN's weights are initially randomized for convenience, based on the uniform distribution principle.

$$(\omega_d)_i^{(l)} \sim U\left[-\frac{1}{\sqrt{n^{(l-1)}}}, \frac{1}{\sqrt{n^{(l-1)}}}\right] \quad (5)$$

Where $n^{(l-1)}$ is the number of $(1-l)-th$ layer nodes on node i in the $l-th$ layer. From the start, each node has an orientation b of 0. In both the FN and CNN sections, the number of nodes required for the pooling layer is present in the final layer. Weights between layers, as well as the mean value μi and variance $\sigma i\ ^2$of the membership function, are all parameters that must be initialized in the FN portion of the FWCNN. The weight between the "Fuzzification," layers and "Operation" layers is adjusted to 1. The value of μi is set by a statistical approach, and σi2 may be calculated from the mean value. Tuning settings for the FWCNN model may be adjusted in a task-oriented manner once its components have been correctly set up. The FWCNN model is trained using back-propagation and the Adam algorithm to ensure that the parameters are properly adjusted. The Adam algorithm excels in non-convex maximum optimization and is therefore well-suited to problems involving big datasets and high-dimensional spaces. The procedure for updating parameters is outlined below:

The gradient $b_t$of the parameters is calculated concerning the equation (6):

$$b_t = \frac{\partial L(\theta)}{\partial \theta^{(l)}} = \sum_n \frac{\partial L(\theta)}{\partial y_i^{(l)}} \cdot \frac{\partial y_i^{(l)}}{\partial x_i^{(l)}} \cdot \frac{\partial x_i^{(l)}}{\partial \theta^{(l)}} \quad (6)$$

Where $L(\theta)$ refers to the reconstruction error defined in equation 7, and θ reflects the FWCNN model's general parameter adjustment. The activation function and the neuron's output $y_i^{(l)}$ are used to derive the last two components in equation (6), the first of which is the back-propagation term.

Estimates for the first and second moments of orientation may be found in equations (7) and (8), respectively.

$$m_t = \beta_1 . m_{t-1} + (1-\beta_1) b_t \quad (7)$$

$$v_t = \beta_2 . v_{t-1} + (1-\beta_2) b_t^2 \quad (8)$$

In which β_1 and β_2 reflect exponential decline (decay) rates for the first and second instant estimates, respectively. Vector m_{t-1} is the first moment, while v_t represents the second raw moment vector. Generally, m_0 and v_0 are defined as having an initial value of 0. Also, $\beta_1, \beta_2 \in [0,1)$.

With the use of equations (7) and (8), it determines a revised estimate for the elapsed time since the first instant and a new estimate for the time since the first moment (second initial moment estimate). Updates to the parameters are calculated using equations (9) and (10),

$$\widehat{m}_t = m_t / (1-\beta_1^t) \quad (9)$$

$$\hat{v}_t = v_t / (1-\beta_2^t) \quad (10)$$

In equations (8) and (9), m_t and v_t represent the first torque vector and the second raw torque vector, respectively. Updates to the parameters are calculated using equation 11.

$$\theta_t = \theta_{t-1} - \alpha . \widehat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon\ , \quad (11)$$

Where parameter values at time $t\ -\ 1$ are denoted by θ_{t-1}, the learning rate is given by α, and the constant of 8-10 is designated by ε. The FWCNN model has the following sets of parameters: θ = {W, b, μ, σ}.

2-4. Assessment of algorithms' results

In most cases, the algorithm's performance may be measured via different metrics. There are a variety of functions that may be employed to determine either the algorithm's error or its accuracy, with the decision ultimately falling to the algorithm architect and the condition of the problem under consideration. Several approaches are explored in this investigation, as shown below. The following metrics are used to assess the precision and recall of classified

data, as well as the overall accuracy of the neural network's classification [29]. Within the realm of information retrieval, the practical criteria of accuracy and recall determine how well the documents retrieved by the system meet the user's requirements. The following are some definitions of these metrics:
[Accuracy=Related document number after retrieval / the number of documents retrieved in total]
[Recall=Related document number after retrieval / the number of documents retrieved in total in the database]

Table 2. A table for determining the algorithm's accuracy: the confusion matrix

		Model-assigned class	
		Positive	**Negative**
Actual class	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

According to Table 2, there are some descriptions as follows [24]:
True Positive (TP): it identifies the correctly classified positive samples.
True Negative (TN): it identifies the correctly classified negative samples.
False Positive (FP): it identifies samples that are being falsely classified as positive.
False Negative (FN): it identifies samples that are being falsely classified as negative.
[Recall (Sensitivity): The number of system true positive samples/ Total number of true positive samples=TP/(TP+FN)]
[Precision: The number of system true positive samples/ Total number of predicted positive samples=TP/(TP+FP)]
[Accuracy: (TP + TN)/(TP + TN + FP + FN)].

2-5. The algorithm error calculation

The algorithm error has been determined in two different ways in this study, and both metrics are equivalent in their application. For the sake of facilitating comparisons between the proposed algorithm and other studies, both approaches are used. This is due to the fact that research papers use different approaches. That is to say, some authors have used one approach and others have utilized another one. Both approaches are presented as equations (12) and (13).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_t - y_t)^2}{N}} \tag{12}$$

$$MAE = \frac{1}{T}\sum_{t=1}^{T}|\hat{y}_t - y_t| \tag{13}$$

2-6. The utilized dataset

This dataset is a standard one, and it was taken from reference [27]. Several recent high-quality studies have utilized this dataset to train and test AI and ML models. Included in this dataset are details about the electrical use of five different London homes, each of which has been assigned a unique "Household" identifier. Simply put, the household-related data in this dataset is organized as Household 1, Household 2, Household 3, Household 4, and Household 5. This dataset contains records that individually indicate the entire amount of power used by a single residence's worth of equipment, including but not limited to TVs, washing machines, refrigerators, and more. The data in this dataset was gathered over the course of a year, making it a time-series dataset. A 24-hour time series plot of all available data is shown in Figure 5 [27].

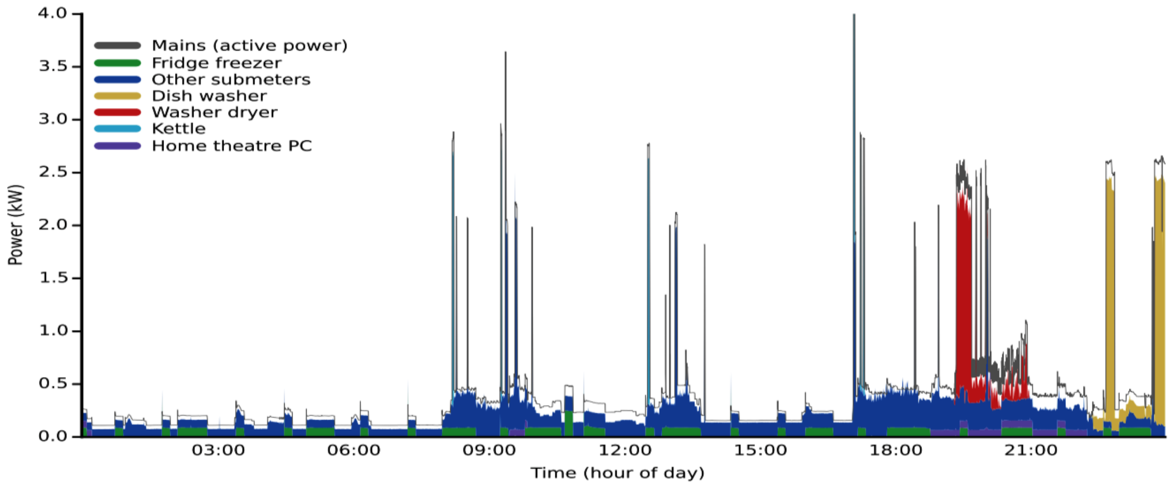

Figure 5. Data from the DLAE dataset depicting the overall electrical appliance consumption over 24 hours [27]

3. Results and discussion

3-1. The proposed neural network results in signal decomposition

It is at this stage that the results of the convolutional neural network technique used to decompose the appliance data from the whole signal are shown. At this point, what follow is the results of processing the data from each appliance through the convolutional neural network algorithm, which decomposed the individual appliance signals from the whole. The accuracy and error of the algorithm for each electrical appliance are included in the findings. Values for Accuracy, Precision, and Recall regarding the signal separation of electrical appliances are shown in Table 3. These values are associated with the assessment of the convolution network's output in decomposing the produced data as belonging to a certain appliance.

Table 3. Error and accuracy in signal decomposition for each electrical device

Device name	Precision	Recall	Accuracy	MAE
Dishwasher	0.88%	0.99%	0.97%	16%
Kettle	1.00%	0.96%	0.99%	14%
Fridge	0.84%	0.82%	0.86%	16%
Microwave	0.96%	0.87%	0.99%	5%
Washing m.	0.51%	0.99%	0.78%	35%

3-2. Training neural networks to estimate energy consumption

The proposed neural network's training has extensively used the Stochastic Gradient Descent (SGD) optimization approach, and the Nesterov Accelerated Gradient (NAG) optimization method has been tested as a secondary option for the new algorithm at the end of the process. Table 4 lists the outcomes achieved by the proposed method while using both optimization techniques.

Table 4. Comparing the output of the proposed algorithm using SGD and NAG optimization methods

Home appliances	Accuracy		Loss		MAE	
	NAG	SGD	NAG	SGD	NAG	SGD
Dishwasher	0.96	0.97	7.6	7.3	9.8	9.4
Kettle	0.99	1.0	3.9	3.9	5.2	5.3
Fridge	0.90	0.89	2.9	3.1	12.8	12.0
Microwave	0.95	0.99	6.6	6.3	11.1	10.11
Washing m.	0.94	0.97	8.7	8.8	16.9	14.0
Average	**0.948%**	**0.964%**	**5.94%**	**5.88%**	**11.16%**	**10.162%**

Estimates generated by various algorithms and the proposed technique are graphically shown in Figure 6 to 8. In each graph, the black lines indicate the ground truth, while the colored lines represent the algorithms' estimates. In fact, the black lines in the figures typically represent the test data. Estimates for the test data have been made using the proposed algorithm, which has been trained using the training data. The test compares the estimated results to the ground truth in order to evaluate the algorithms' performance in a future simulation.

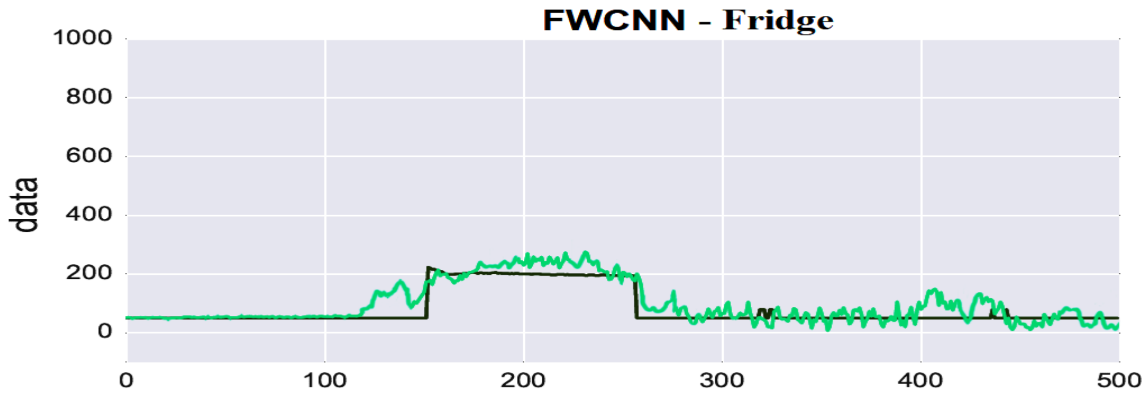

Figure 6. FWCNN estimation for Fridge Consumption.

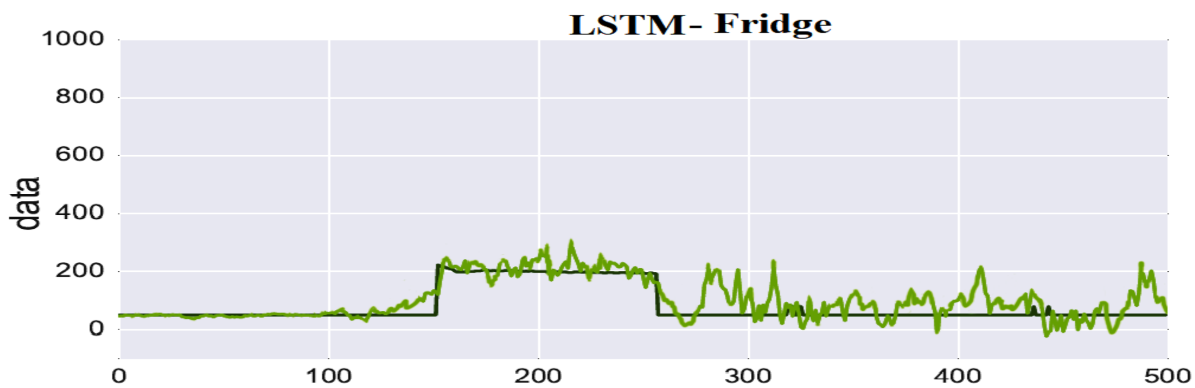

Figure 7. LSTM estimation for Fridge Consumption.

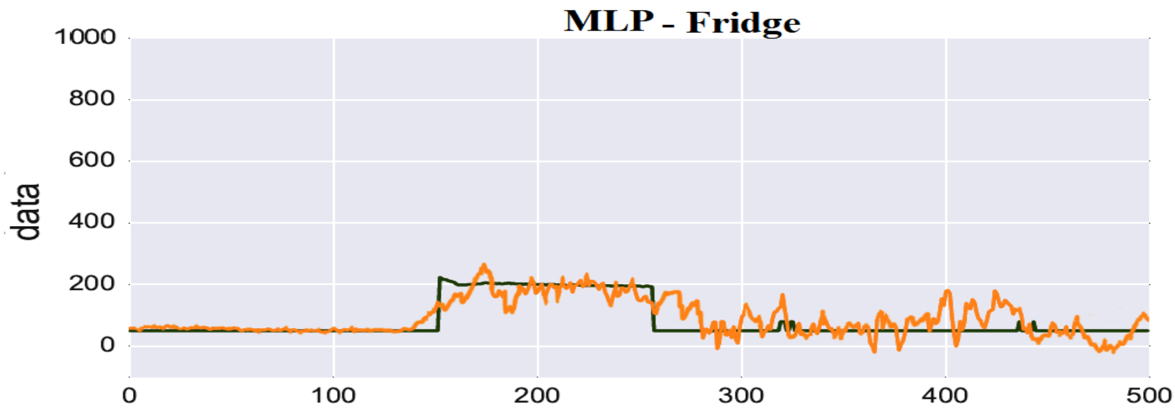

Figure 8. MLP estimation for Fridge Consumption

3-3. Evaluation of the proposed algorithm compared to existing algorithms

This section compares the main algorithm described in this thesis to other algorithms used to estimate energy consumption in order to establish the effectiveness of the proposed main algorithm. Also, a comparison is made with other methodologies using a similar dataset. Table 5 illustrates the results of this comparison.

Table 5. Comparison between the proposed algorithm and other algorithms

Algorithm	dataset	Accuracy (%)	MAE (%)	Ref.
The proposed FWCNN	Uk-dale	**96.4**	**10.162**	[*]
CNN	Uk-dale	-	14.86	[30]
U-NET	Uk-dale	-	11.174	[30]
ARIMA	Uk-dale	-	15.5670	[31]
SVR	Uk-dale	-	10.6512	[31]
Persis.	Uk-dale	-	13.6995	[31]
Multi-Step Short-Term Hybrid Deep Learning	Uk-dale	-	10.1582	[31]
SVM	Uk-dale	79.35	-	[32]
unsupervised data clustering and frequent pattern mining analysis	Uk-dale	89.58	-	[32]
seq2seq	Uk-dale	-	17.999	[33]
seq2point	Uk-dale	-	15.472	[33]
Markov model (AFHMM)	Uk-dale	-	82.79	[33]
seq2seq(Kelly)	Uk-dale	-	93.488	[27]
VDOCNN	Uk-dale	85.84	-	[34]
Xception	Uk-dale	85.84	-	[34]
Concatenate-DenseNet121	Uk-dale	90.25	-	[34]
he Neuro-Fuzzy Hybridization-	Uk-dale	86.49	-	[35]
Autoencoder	Uk-dale	94.00	-	[36]
FFNN	Uk-dale	95.385	-	[37]
LPH	Uk-dale	98.51	-	[38]
CNN–LSTM	Uk-dale	97.995	-	[39]

4. Conclusion

Table 5 indicates that various algorithms have produced varying outcomes, with CNN-based algorithms showing the most improvement in recent years. Nevertheless, the fact that the algorithms are not being compared on identical datasets makes the comparisons slightly challenging. For instance, the number of samples used in the publication discussed in the CNN-LSTM algorithm differs from the number of samples used in this research and other papers. The result of the algorithm may be viewed to provide insight into how well it performs when classifying relevant data since; in general, all algorithms utilize relatively similar datasets. This means that the proposed algorithm is a top contender among the existing algorithms.

5. Challenges and future studies

The absence of a concrete criterion for evaluating the presented algorithms is only one of the numerous obstacles in the way of these tests. While it is important to know how many samples were used for training and testing an algorithm, in some research studies, this information is left vague. Owing to the significance of this factor in comparing algorithms, it is almost impossible to identify with precision which algorithm is superior to others. This provides a foundation for future research into the means of addressing these difficulties and into the means of various combinations to improve this algorithm. Furthermore, improved approaches for optimizing the descending gradient may be obtained to increase the algorithm's accuracy by incorporating contemporary optimization techniques including SGD and NAG.

Future NILM studies can benefit from integrating IoT data such as temperature, humidity, and occupancy sensors to enhance appliance detection accuracy. Employing deep learning models like RNNs or transformers can improve time-series analysis of consumption behavior. Researchers should focus on distinguishing appliances with similar usage profiles using higher-level features such as operational sounds or specific frequency patterns. Real-time NILM systems could enable early detection of anomalies and faults. Developing region-specific algorithms tailored to cultural and climatic energy-use patterns can improve adaptability. Lastly, behavioral research should explore how detailed consumption feedback affects user habits and energy efficiency.

NILM research faces several limitations, including difficulty in distinguishing between appliances with similar energy signatures such as refrigerators and air conditioners and vulnerability to signal and environmental noise, which can reduce algorithm accuracy. Scalability is another concern, as models effective in one building type may not perform equally well in different settings. The lack of diverse, standardized datasets restricts the replicability and comparison of results. Privacy concerns also emerge when collecting detailed consumption data, particularly in real-time scenarios. Furthermore, accurately separating multiple devices operating simultaneously presents a challenge due to signal overlap. Finally, the computational cost and complexity of advanced models, like deep learning, may limit their practical deployment in low-resource devices.

Funding

There is no funding support.

Declaration of Competing Interest

The author has no conflicts of interest to declare that are relevant to the content of this article.

ACKNOWLEDGEMENTS

We would like to express our gratitude to the anonymous reviewers for their valuable comments, which greatly contributed to improving our work.

References

Y.Liu, S.Yu, Y.Zhu, D.Wang, J.Liu, Modeling, planning ,application and management of energy systems for isolated areas: A review, Renewable and Sustainable Energy Reviews, 82 (2018), 460-470 .

Council, W. E. (2013). World energy scenarios composing energy futures to 2050. PSI, London .

M.Sehatpour, A.Kazemi, H.Sehatpour, Evaluation of alternative fuels for light-duty vehicles in Iran using a multi-criteria approach, Renewable and Sustainable Energy Reviews, 72(2017) 295-31.

B.Salvi, K.Subramanian, N.Panwar, Alternative fuels for transportation vehicles: a technical review, Renewable and Sustainable Energy Reviews, 25(2013), 404-419 .

T. Chaudhuri, Y.C. Soh, H. Li, L. Xie, A feedforward neural network based indoor-climate control framework for thermal comfort and energy saving in buildings, Applied Energy, 248 (2019), 44-53.

Y. Himeur, A. Alsalemi, F. Bensaali, A. Amira, Robust event-based non-intrusive appliance recognition using multi-scale wavelet packet tree and ensemble bagging tree, Applied Energy, 267 (2020) 114877.

A. Wójcik, R. Lukaszewski, R. Kowalik, W. Winiecki, Nonintrusive Appliance Load Monitoring: An Overview, Laboratory Test Results and Research Directions, *Sensors*, *19* (2019) 3621.

R. Usman, P. Mirzania, S.W. Alnaser, P. Hart, C. Long, Systematic Review of Demand-Side Management Strategies in Power Systems of Developed and Developing Countries. *Energies*, 15 (2022) 7858.

G.W. Hart, Nonintrusive appliances load monitoring, Proceedings of the IEEE, 80(1992) 1870-1891.

T.K. Nguyen, E. Dekneuvel, G.Jacquemod, Development of a real-time non-intrusive appliance load monitoring system: an application level model, Power Energy System 90(2017) 168–180.

C. Puente, R.Palacios, Y.González-Arechavala, E.F. Sánchez-Úbeda,. Non-Intrusive Load Monitoring (NILM) for Energy Disaggregation Using Soft Computing Techniques. Energies, 13(2020) 3117.

X. Wu, Y. Gao, D. Jiao, Multi-label classification based on random forest algorithm for non-intrusive load monitoring system, Processes, 7 (2019) 337.

B. Yin, Z. Li, J. Xu, Non-intrusive load monitoring algorithm based on household electricity use habits, Neural Computer & Application, 34 (2022) 15273–15291.

S. Wu, K.L. Lo, Non-intrusive monitoring algorithm for resident loads with similar electrical characteristic, Processes 8 (2020) 1385.

D. Li, S. Dick, Non-intrusive load monitoring using multi-label classification methods. Electr. Eng. 103(2021) 607–619.

Y. Liu, Q. Shi, Y. Wang, X. Zhao, S. Gao, X. Huang, An Enhanced Ensemble Approach for Non-Intrusive Energy Use Monitoring Based on Multidimensional Heterogeneity, *Sensors 21* (202) 7750.

Y. Hu, Y. Lin, C. Lin, Artificial intelligence accelerated in parallel computing and applied to nonintrusive appliance load monitoring for residential demand-side management in a smart grid: A comparative study, Appl. Sci., 10 (2020).

X. Wu, D. Jiao, Y. Du, Automatic implementation of a self-adaption non-intrusive load monitoring method based on the convolutional neural network, Processes 8 (2020) 704.

X. Wu, Y. Gao, D. Jiao, Multi-label classification based on random forest algorithm for non-intrusive load monitoring system, Processes 7 (2019) 337.

İ.H. Çavdar, V. Faryad, New design of a supervised energy disaggregation model based on the deep neural network for a smart grid, Energies 12 (2019) 1217.

Y. Lin, Y. Hu, Electrical energy management based on a hybrid artificial neural network-particle swarm optimization-integrated two-stage non-intrusive load monitoring process in smart homes, Processes 6 (2018) 236.

S. Alotaibi, “Advancing energy performance efficiency in residential buildings for sustainable design: Integrating machine learning and optimized explainable AI (AIX),” Int. J. Energy Res., Article ID 6130634, 2024.

Z. Azim, H. Pooyafar, B. A. Naieri, D. P. Khiabani, and A. M. Bazargan, “Prediction of energy consumption for residential buildings based on the behavior of its residents using artificial intelligence neural network method in Tabriz city,” Strategic Soc. Energy Stud., vol. 11, no. 1, pp. 1–20, 2024.

M. Khodadadi, L. Riazi, and S. Yazdani, “A novel ensemble deep learning model for building energy consumption forecast,” Int. J. Eng., vol. 37, no. 6, pp. 1067–1075, 2024.

M. Neshat, M. Thilakaratne, M. El-Abd, S. Mirjalili, A. H. Gandomi, and J. Boland, “Smart buildings energy consumption forecasting using adaptive evolutionary ensemble learning models,” arXiv preprint, 2025. [Online]. Available: https://doi.org/10.48550/arXiv.2506.11864

S. Alam, “Deep learning applications for residential energy demand forecasting,” AI, IoT and the Fourth Industrial Revolution Review, vol. 14, no. 2, pp. 27–38, 2024.

J. Kelly, W. Knottenbelt, The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes, Scientific data, 2 (2015) 1-14.

B. Zhao, H. Song, Fuzzy Shannon wavelet finite element methodology of coupled heat transfer analysis for clearance leakage flow of single screw compressor. Engineering with Computers, (2021) 1-11.

A. Kulkarni, D.Chong, F. A. Batarseh, Foundations of data imbalance and solutions for a data democracy, Data Democracy, 5 (2020) 83-106.

A. Faustine, L. Pereira, H. Bousbiat, S. Kulkarni, UNet-NILM: A deep neural network for multi-tasks appliances state detection and power estimation in NILM, In Proceedings of the 5th International Workshop on Non-Intrusive Load Monitoring, (2020) 84-88.

K. Yan, X. Wang, Y. Du, N. Jin, H. Huang, H. Zhou,. Multi-step short-term power consumption forecasting with a hybrid deep learning strategy. Energies, 11 (2018) 3089.

S. Singh A. Yassine, Big data mining of energy time series for behavioral analytics and energy consumption forecasting, Energies, 11 (2018) 452.

C. Zhang, M. Zhong, Z. Wang, N. Goddard, C. Sutton, Sequence-to-point learning with neural networks for non-intrusive load monitoring. In Proceedings of the AAAI Conference on Artificial Intelligence (2018) Vol. 32, No. 1.

H. Kim, S. Lim,. Temporal Patternization of Power Signatures for Appliance Classification in NILM. Energies, 14 (2021) 2931.

Y.C. Hu, Y. H. Lin, H.L. Gururaj, Partitional Clustering-Hybridized Neuro-Fuzzy Classification Evolved through Parallel Evolutionary Computing and Applied to Energy Decomposition for Demand-Side Management in a Smart Home, Processes, 9 (2021) 1539.

J. Kelly, W. Knottenbelt, Neural NILM: Deep neural networks applied to energy disaggregation, In Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments (2015) 55–64.

P. Franco, J.M. Martínez, Y.C. Kim, M.A. Ahmed, IoT based approach for load monitoring and activity recognition in smart homes, IEEE Access, 9 (2021) 45325-45339.

Y. Himeur, A. Alsalemi, F. Bensaali, A. Amira, Smart non-intrusive appliance identification using a novel local power histogramming descriptor with an improved k-nearest neighbors classifier, Sustainable Cities and Society, 67 (2021) 102764.

X. Zhou, J. Feng, Y. Li, Non-intrusive load decomposition based on CNN–LSTM hybrid deep learning model. Energy Reports, 7 (2021) 5762-5771.