I.J. Intelligent Systems and Applications, **2025, 6, 75-93**
Published Online on December 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijisa.2025.06.06

Non-intrusive Load Monitoring for Home Appliance Using Sequence-to-point Convolutional Neural Networks

John Oluwasegun Memud
Department of Electrical and Electronics Engineering, School of Engineering and Engineering Technology, The Federal University of Technology, Akure, Ondo State, Nigeria
E-mail: memudeee186748@futa.edu.ng
ORCID iD: https://orcid.org/0009-0004-8289-0509

Brendan Chijioke Ubochi*
Department of Electrical and Electronics Engineering, School of Engineering and Engineering Technology, The Federal University of Technology, Akure, Ondo State, Nigeria
E-mail: bcubochi@futa.edu.ng
ORCID iD: https://orcid.org/0000-0002-9967-3891
*Corresponding author

Michael Rotimi Adu
Department of Electrical and Electronics Engineering, School of Engineering and Engineering Technology, The Federal University of Technology, Akure, Ondo State, Nigeria
E-mail: mradu@futa.edu.ng
ORCID iD: https://orcid.org/0009-0007-9090-1409

Nnamdi Nwulu
Center for Cyber-Physical Food, Energy, and Water Systems, University of Johannesburg, Johannesburg, South Africa
E-mail: nnwulu@uj.ac.za
ORCID iD: https://orcid.org/0000-0003-2607-7439

Received: 10 June 2025; Revised: 21 August 2025; Accepted: 20 October 2025; Published: 08 December 2025

Abstract: Non-intrusive load monitoring (NILM) aims to estimate the operational states and power consumption of individual household appliances, providing real-time insights into energy usage for effective energy management and improved demand side response strategies. This study addresses the challenge of accurate energy disaggregation of household energy consumption data into individual appliances' consumption, an important requirement for effective energy management in smart homes. Traditional energy monitoring systems provide only aggregate data, limiting the ability to optimize energy consumption. To overcome these difficulties, this study proposes a Convolutional Neural Network (CNN)-based model for Non-Intrusive Load Monitoring (NILM) that disaggregates total energy usage into appliance-specific consumption for five key appliances: kettle, microwave, fridge, dishwasher, and washing machine. Unlike the previous approaches, our model integrates a hybrid dataset from UK-DALE and REFIT, leveraging data fusion techniques to enhance generalization. The CNN architecture employed uses five convolutional layers for effective feature extraction, capturing temporal dependencies in appliance usage patterns and thus results in an improved MAE and SAE when compared to similar published results. The preprocessing and hybridization stage involves such processes as missing data imputation, appliance state labelling, feature normalization and merging of the datasets. The developed model achieved an overall accuracy of 98.3% and an F1-score of 81.7% in seen scenarios, while in unseen environments, it attained 96.5% accuracy and an F1-score of 58.1% when tested on the UK-DALE dataset. The seen scenario refers to testing using UK-DALE House 1 and REFIT House 2 data of the validation dataset, whereas the unseen scenario involves entirely new house data not used during training and validation. It is shown that post-processing techniques reduce errors, highlighting its effectiveness, which help to enhance the model's predictive accuracy. This study contributes to the advancement of NILM technologies by combining datasets, offering a robust and scalable solution for individual appliace energy monitoring, with significant implications on energy conservation and smart home efficiency.

Index Terms: Non-intrusive Load Monitoring, Deep Learning, CNN, Energy Disaggregation, Feature Learning, Energy Conservation.

1. Introduction

Non-Intrusive Load Monitoring (NILM) involves the identification and monitoring of the energy consumption of individual appliances without requiring sensors. Thus, it provides a real-time feedback on energy use and can improve efficiency by analyzing load patterns. Unlike Intrusive Load Monitoring (ILM), which requires in-device sensors, NILM uses a single meter to collect total household consumption, offering advantages in terms of reduced installation costs and lower complexity. NILM offers low-cost and efficient solutions, which is increasingly applied in residential energy management and smart homes [1, 2]. NILM involves a number of processes that include data acquisition, event detection, feature extraction, load identification, and energy disaggregation [3]. This technology is especially important in areas with rising energy demands or limited electricity power supply and can, therefore, help to promote improved energy consumption and conservation efforts [4]. NILM also find applications in fault identification and appliance health monitoring in smart homes. Traditional load monitoring approaches provide only the total energy usage and does not provide insights into the consumption pattern of the individual appliances. Also, the absence of an accurate appliance-level consumption data makes it difficult to achieve an effective and efficient energy management [3, 5, 6].

Recent NILM studies have utilized machine learning and in particular deep learning approaches with improved outcomes. For example, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have increased the precision of appliance identification. Graph Neural Networks (GNNs), on the other hand, capture the intricate relationships among appliances, with an improved capacity to properly model household behavior [7, 8]. While sequence-to-sequence (seq2seq) CNN maps input sequences to output sequences, it struggles with challenges like the vanishing gradient problem and graphics processing unit (GPU) limitations, especially when handling long sequences. The Sequence-to-Point method addresses this problem by predicting only the midpoint of a sliding window which enhances accuracy at device state transitions [9]. Unlike the sequence-to-sequence method that smoothens the output based on averaged multiple predictions, the Sequence-to-Point improves prediction performance and eliminates edge prediction difficulties by focusing on the critical midpoint [10]. In order to improve the accuracy, larger windows are often used, since they capture more temporal patterns and transitions. A major problem with existing NILM models is that they often suffer from limited generalizability due to their reliance on single-source datasets.

This study proposes a dataset hybridization approach to improve disaggregation accuracy for both type-1 and type-2 appliances while ensuring strong generalization across unseen data. The primary objective of this research is to develop a more reliable and adaptable NILM framework by integrating multiple datasets with distinct characteristics. Unlike traditional approaches that train models on a single dataset, leading to dataset-specific biases, this study implements the fusion of UK-DALE and REFIT datasets. The fusion process perfectly aligns common electrical parameters, including active power and sampling rates, creating a comprehensive hybrid dataset for sequence-to-point convolutional neural network training. The Sequence to-Point CNN model employed in this study uses a larger window to increase accuracy and efficiency in appliance disaggregation. The sequence-to-point CNN model leverages hierarchical feature extraction through stacked 1-D convolutional layers, progressively increasing filter sizes to capture intricate temporal patterns in appliance power consumption. The selected appliances in this study include kettle, microwave, fridge, dishwasher, and washing machine. These appliances have been selected based on their availability and their consistency in the datasets used for training and validation. Additionally, a post-processing technique is designed to align predicted energy usage with ground-truth data by filtering out irrelevant activations. This is done by multiplying appliances activation status with the predicted energy. Some of the challenges include data distribution errors, dataset synchronization inconsistencies, and increased latency due to the larger window size. In order to integrate the proposed approach into existing energy management systems, we propose an NILM framework using pre-trained CNN-based models on the hybrid data-set for online appliance disaggregation.

The rest of paper is structured as follows: Section 2 reviews existing NILM algorithms; Section 3 covers problem definition, CNN-based disaggregation, post-processing, system integration, and case study details; Section 4 presents results and discussion, while Section 5 concludes with future directions.

2. Literature Review

Early research on Non-Intrusive Load Monitoring (NILM) can be traced to George Hart's foundational work in the early 1990s which introduced the Non-Intrusive Appliance Load Monitor (NALM) as a system for estimating the energy consumption of individual appliances using single-point electrical data [11].

NILM algorithms currently apply machine learning models for feature extraction and these models are broadly categorized as unsupervised and supervised machine learning. Unsupervised methods, such as Hidden Markov Models (HMMs), were commonly used in early NILM research due to their ability to operate without labeled data [4, 12, 13].

Enhanced HMMs, including Additive and Differential variants, use Maximum a Posteriori (MAP) for improved appliance disaggregation. The Additive Factorial Approximate MAP (AFAMAP), which leverages additive structures, achieves over 90% precision for appliances with distinct patterns, outperforming SMF (Switch Mode Function) algorithm, which struggles with variable assignments and generalization [4]. However, AFAMAP faces challenges with appliances sharing similar load profiles and requires iterative refinement for more complex setups. Using Factorial Hidden Markov Models (FHMM) [13], circuit-level power measurements and feature concatenation were used to recognize appliance states, achieving 90% accuracy for binary appliances and 80% for multi-state appliances. The approach outperforms traditional event-based methods and includes a real-time monitoring prototype.

Supervised learning has proven to be an effective method for NILM tasks by utilizing labeled sub-metered data to identify specific appliance patterns. Advanced deep learning models such as CNNs and Long Short-Term Memory (LSTM) networks have become popular in NILM because of their improved scalability and real-time performance [14][15]. CNNs are known to perform very well in tasks requiring feature extraction from raw signals while LSTMs have demonstrated excellent performance in multi-state appliance identification [7]. Further enhancements to these models have been achieved by combining k-means clustering with Deep Neural Network (DNN) resulting in improvements in NILM accuracy [6]. The use of LSTM networks for multi-state appliance identification, discussed in [7] achieved a 97% accuracy and a 0.91 F1 score for seen houses and effective power retrieval (MAE/MSE: 12.3/9074, 13.9/8144.5) for unseen houses. The reported model supports real-time feedback on low-cost devices like ESP32, enabling efficient energy monitoring and fault detection of appliances. Also, by combining k-means clustering with a DNN, the model achieved an accuracy of 97% for appliance class type identification and 83.9% for appliance state identification, demonstrating a strong performance for multi-state appliances [6].

Among the CNN models, the sequence-to-sequence CNN has recently been used in tasks requiring household energy disaggregation in order to address the identifiability problem. However, this method suffers from accuracy issues since its final output is the average of multiple predictions [16]. The sequence-to-point CNN model proposed in [9] to address the identifiability problem was reported to improve two standard error measures by 84% and 92%, respectively. However, a major challenge with this model is the limitation posed by the size of the sliding window. Some of the approaches used to improve the sequence-to-point CNN include integrating Bidirectional Long Short-Term Memory (Bi-LSTM) layers to enhance temporal feature extraction and by employing Bidirectional Temporal Convolutional Networks (BTCNs) for improved performance in capturing forward and backward dependencies [16,17]. Some studies have used transfer learning for NILM with considerable savings in computational cost and time [18].

Some of the current problems in NILM research are related to developing methods for improved disaggregation accuracy, reduced generalization error, and improving the reliability of real-time processing [19]. High-frequency data acquisition is capable of capturing detailed transient signatures which are important for differentiating between appliances with overlapping steady-state features. However, high frequency data are usually affected by noise and requires advanced preprocessing, such as discrete wavelet transforms, to extract meaningful patterns while reducing noise. For appliances with similar load profiles, sequence-to-point learning approaches, as discussed in [17], can enhance discrimination by using both temporal dependencies and frequency-domain transformations. Models that are trained on high-resolution datasets achieve better detection accuracies at the cost of increased computational complexity especially during data acquisition of low-power data or similar load-profile appliances [20]. During the disaggregation stage, deep neural networks often predict activations that do not correspond to the target appliance. To address this issue, some studies have used post-processing methods [19, 21], including the use of a separate deep CNN model to classify predicted appliance activations. Their classification model verified whether the predicted sequence accurately represented the target appliance.

Some of the popular NILM datasets such as UK-DALE [22], REDD [23], and REFIT [24], have diverse energy profiles that are suitable for model training and validation. Dataset hybridization in Non-Intrusive Load Monitoring (NILM) enhances disaggregation accuracy and model generalization by combining multiple datasets [25]. Methods include data fusion, which aligns common features (e.g., active power, sampling rates) from datasets like UK-DALE and REFIT [22, 24], and synthetic data generation which supplements real data for improved model training in low-data scenarios [26]. The fusion of steady state and transient features can help to enhance appliance signature performance [27].

3. Proposed Energy Disaggregation Methodology

3.1. Problem Definition

The problem involves disaggregating into individual power consumption from an aggregate household power time series represented in (1) [9]:

$$Y = \{y1, y2, \ldots, yt \ldots,\} \tag{1}$$

where 'yt' represents the aggregate power at time t. For 'sequence to point' learning, the model takes an input window $Y_{t:t+W-1}$ of size W, centered at 't', and predicts the target appliance consumption '$X_{t+\frac{W}{2}}$' at the window's midpoint as shown in (2):

$$\hat{x}_{t+\frac{W}{2}} = f_\theta\left(Y_{t:t+W-1},\right) \tag{2}$$

Where 'f_θ' is the function represented by the CNN model with parameter θ, and '$\hat{x}_{t+\frac{W}{2}}$' is the predicted power consumption of the appliance at the midpoint of the window [9].

3.2. Model Architecture

Convolutional Neural Networks (CNNs) have emerged as important architectures for image classification tasks, demonstrating exceptional efficacy in recent empirical studies. Their superiority stems from their capacity to model complex, nonlinear spatial dependencies within data structures. The sequence-to-point model, as shown in Fig. 1, uses a convolutional neural network (CNN) to capture local temporal patterns with convolutional layers followed by fully connected layers. These layers detect changes and characteristic appliance signatures within the window. Unlike sequence-to-sequence models which predict an entire output sequence, sequence-to-point models focus on a single prediction at the midpoint, simplifying the task and increasing model efficiency.

The basic CNN architecture comprises of three fundamental components: convolutional layers, pooling layers, and fully connected layers as described in [28]. The convolutional layer functions as the primary feature extraction mechanism, employing multiple convolutional filters to generate distinct feature maps. This transformation can be formally expressed as in (3):

$$z^{l}_{i,j,k} = W^{l}_{k} x^{l}_{i,j} + b^{k}_{l} \tag{3}$$

where $x^l_{i,j}$ denotes the input at spatial position i, j in the $l-th$ layer, while W^l_k and b^l_k represent the weights and bias of the $k-th$ filter in that layer, respectively.

The Rectified Linear Unit (ReLU) activation function is predominantly utilized in contemporary architectures and is mathematically defined in (4) as:

$$ReLU(x) = max(0, x) \tag{4}$$

The pooling layer is designed to lower feature maps, reducing spatial dimensions while retaining essential information. This enhances the model's shift-invariance properties. Pooling is performed using a function applied over local regions of the feature maps defined in (5) as:

$$y^l_{i,j,k} = pool(\phi^l_{m,n,k}), \forall(m,n) \in R_{i,j} \tag{5}$$

Where $R_{i,j}$ represents the local neighbourhood around (i, j). In this work, max pooling is used, which selects the maximum value within a given region. With a stride l and pooling size s, max pooling is computed in (6) as:

$$y(i) = max(x[i \times l : i \times l + s - 1]) \tag{6}$$

The fully connected layers are responsible for high-level reasoning by connecting every neuron in the previous layer to all neurons in the current layer. This ensures that the extracted features contribute to the final classification or prediction.

Training a CNN involves optimizing a loss function to improve the model's accuracy. Given a dataset of input-output pairs $\{(x^k, y^k)\}$, where x^k represents the input data and y^k is the corresponding label, the CNN parameters are denoted as θ. The loss function is computed in (7) as:

$$\Gamma = \frac{1}{N} \sum_{n=1}^{N} \ell(\theta, y^{(n)}, o) \tag{7}$$

Where O represents the CNN's output. Minimizing this loss function helps identify the optimal parameters for the network.

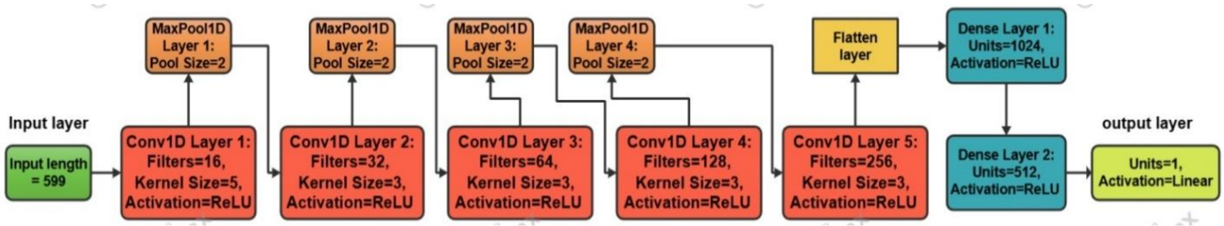


Fig.1. The architecture for sequence to point neural network

3.3. *System Integration and Scalability*

Conventional deep learning-based Non-Intrusive Load Monitoring (NILM) methodologies exhibit notable limitations regarding cross-domain generalization, appliance diversity accommodation, and robust performance in noisy environments, thus reducing their practical implementation. To address these deficiencies, we propose a sequence-to-point Convolutional Neural Network (CNN) NILM architecture designed for compatibility with existing energy monitoring infrastructure while maintaining extensibility for large datasets and diverse appliance range as shown in Fig. 2.

The scalability of the proposed NILM framework is enhanced through its ability to generalize across diverse appliance categories and adapt to larger datasets without significant performance degradation. The selected appliances represent a broad spectrum of domestic energy consumption patterns, including those with complex multi-state operations (e.g., washing machines and dishwashers) and simpler single-state appliances (e.g., kettle). As a result, other single or variable state appliances could easily be integrated into the system. The incorporation of high-performance or edge-computing can enable a real-time disaggregation even as data volume increases. These scalability features position the proposed NILM system as a viable solution for large-scale energy monitoring and management applications.

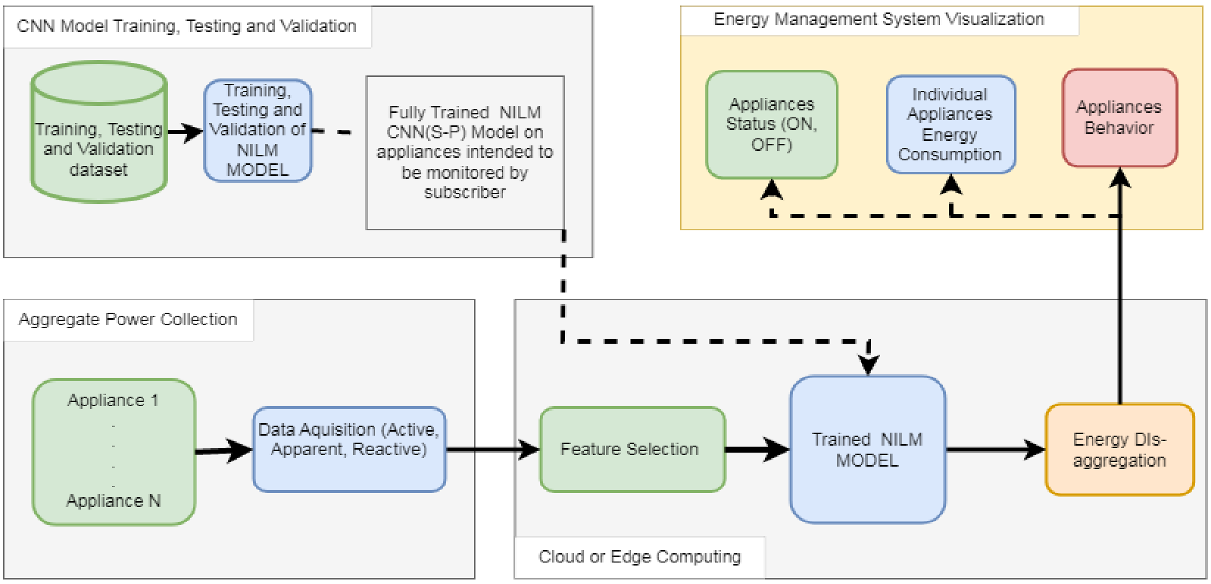


Fig.2. System integration frame work for the proposed NILM model

The proposed NILM system architecture facilitates a direct integration with contemporary smart metering systems and energy management platforms through a multi-stage implementation approach:

- **Pre-training and Generalization**: The CNN model undergoes initial training using a composite dataset integrating UK-DALE and REFIT repositories to enhance cross-environmental generalization capabilities. This pre-training methodology ensures interoperability with diverse energy monitoring infrastructures. The appliances to be monitored are individually trained, with each appliance having its own trained model.
- **Computational Implementation**: Temporal energy consumption data from smart metering devices undergoes processing via cloud-based or edge-computing infrastructure, where the pre-trained sequence-to-point CNN model executes appliance-specific disaggregation. This distributed computing paradigm minimizes computational burden on terminal devices while preserving disaggregation accuracy.

- **Extensibility Mechanisms**: The framework leverages cloud infrastructure to accommodate larger data repositories. Furthermore, the training of the model on diverse operation state appliance facilitates efficient scaling across diverse appliances with varying operational states.
- **Output Refinement and Integration**: Post-processing algorithms enhance disaggregation outputs by aligning predicted consumption patterns with verified ground-truth data. These refined results integrate seamlessly with existing energy visualization dashboards, providing actionable intelligence to both consumers and utility providers.

Through implementation of a modular and scalable architectural design, this NILM framework ensures compatibility with the existing monitoring infrastructure while supporting expanded appliance coverage and dataset growth, establishing viability for large-scale deployment scenarios.

3.4. Case Study

In this study, we employ a neural network-based approach to perform non-intrusive load monitoring (NILM) on the fusion of two widely recognized datasets, UK-DALE and REFIT, with a focus on disaggregating specific household appliances following the proposed framework of the sequence-to-point NILM algorithm as shown in Fig. 3. The goal is to evaluate the performance of the model in identifying the operational states of appliances such as the kettle, microwave, fridge, dishwasher, and washing machine.

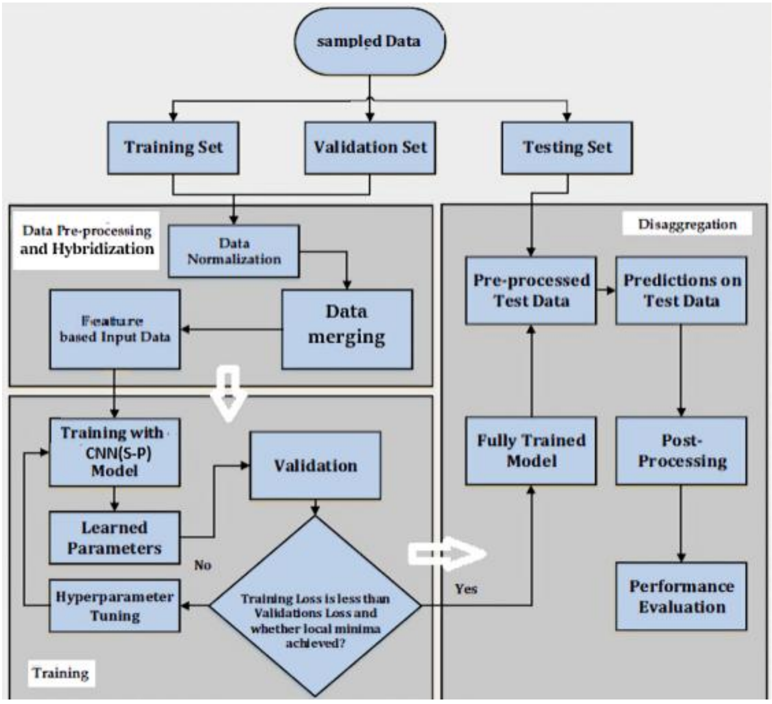


Fig.3. Proposed framework for the sequence to point NILM algorithm

A. Data Collection

This study has used two datasets (The UK-DALE and REFIT). The UK-DALE dataset provides aggregate and appliance-level power readings for five UK homes. Aggregate data is sampled at 1 Hz, while appliances like the kettle, microwave, fridge, dishwasher, and washing machine are sampled at 1/6 Hz. Only Houses 1 and 2 were selected due to their completeness and appliance data availability.

The REFIT dataset includes energy data from 20 UK homes obtained at 8 Hz sampling rate. For consistency with UK-DALE, only Houses 2, 3, 5, 9, 11, and 20 were selected, ensuring sufficient appliance usage diversity for robust disaggregation.

B. Data Pre-processing and Hybrid Dataset Construction

Data preprocessing and hybrid dataset construction entails the integration of multiple NILM datasets to enhance model performance metrics. This methodology standardizes heterogeneous data sources for optimal deep learning model training through several critical procedures. The key steps include:

- **Timestamp Standardization:** All temporal data across datasets are systematically resampled to a uniform 60-second interval, establishing consistent time-series alignment necessary for accurate event detection and analysis.

- **Reshaping and Resampling:** The UK-DALE (originally with 6-second resolution) and REFIT (originally with 8-second resolution) datasets undergo individual resampling procedures to achieve the standardized 60-second interval, with appropriate interpolation techniques applied to address intermediate values.
- **Missing Data Imputation:** Temporal gaps of less than three minutes are addressed using forward-fill imputation methodology. Additionally, appliance power consumption values are constrained within operational parameters such as 'maximum on power' as defined in Table 1, ensuring reasonable measurements.
- **Appliance State Labelling:** A uniform labelling classification is implemented utilizing predefined power-threshold criteria and temporal duration filters such as 'minimum on duration', 'minimum off duration' and 'on power threshold' as specified in Table 1, establishing consistency in appliance state identification across different datasets.
- **Dataset Partitioning:** The integrated dataset is segmented into training, validation (comprising 20% of the training data as detailed in Table 2), and testing subsets. The testing corpus includes both previously observed households from the validation set and completely new households to assess generalization capabilities.
- **Feature Normalization:** Electrical parameters undergo min-max normalization to the [0, 1] range to ensure numerical stability. Feature relevance is quantified through mutual information scores, categorized as high (>0.2), moderate (0.1–0.2), or weak (<0.1) as stated in [19].
- **Dataset Merging:** The processed UK-DALE and REFIT training datasets are merged into a unified hybrid dataset while preserving temporal dependencies and capturing diverse appliance utilization patterns.

This comprehensive pre-processing methodology produces a robust hybrid dataset that enhances NILM model adaptability across diverse residential environments.

Table 1. Parameters of the selected appliances

Appliance	Max power (watts)	On power threshold (watts)	Min. on duration (secs)	Min. off duration (secs)
Kettle	3100	2000	12	0
Fridge	3000	200	45	30
Washing m.	1000	60	400	15
Microwave	2500	20	1000	1600
Dishwasher	2500	50	1000	700

Table 2. Dataset selection for training and testing

Phase	Houses	Percentage
Training	UK-DALE House 1, REFIT Houses 2, 3, 5, 9, 11	80%
Validation	Same as training set	20%
Testing(seen)	UK-DALE House 1, REFIT House 11	from validation
Testing(unseen)	UK-DALE House 2, REFIT House 20	100%

C. Training and Hyper-parameter Optimization

The classification sub-network CNN architecture proposed contains the following layers and the hyper-parameters for each layer. It contains 13 layers, 5 Conv1-D, 4 MaxPool1-D, 1 Flatten and 3 Dense layers followed by max pooling layer after each 1-D convolutional layer. Each convolutional layer has its kernel size of which the first convolutional layer has 5 kernel size of 5 and the rest with a kernel size of 3 with each having ReLU (Rectified Linear Unit) activation. This is then followed by a flatten layer, 2 fully connected dense layers of 1024 and 512 units respectively. The last layer is an output layer with 1 unit. Key hyper-parameters, such as batch size (1024) and learning rate (10^{-4}), were selected for optimal learning, with the Adam optimizer used due to its adaptive learning features ($\beta_1 = 0.9$, $\beta_2 = 0.999$). To prevent over-fitting, early stopping was implemented, halting training when validation loss peaked. The patience level for the training iteration was set at 15 epochs and the highest number of trainings was set to 50 epochs to ensure proper appliance pattern learning. Regularization techniques, including weight decay and gradient clipping, were also applied to maintain parameter stability and reduce generalization error, ensuring effective model performance during testing.

D. Post-Processing of Predicted Appliance Activations

The post-processing phase is based on the work of [21]. It refines the output of the model by providing an array of values that range from 0 to 1, indicating the probability of an appliance being active. An appliance is considered "on" if this probability exceeds a threshold of 0.5.

To enhance energy accuracy and reduce false activations such as spike in the proposed energy disaggregation process, the following steps are implemented in the post-processing algorithm:

- De-normalization of Aggregate Data: The aggregate data was de-normalized to restore its original scale.
- Calculation of Ground Truth and Prediction Status: The ground truth and prediction status were computed using ‘minimum on duration’, ‘minimum off duration’ and on power threshold as specified in Table 1, establishing consistency in appliance state identification across different datasets.
- Batch Adjustment: Adjustments were made because the ‘test_provider’ function defaults to producing complete batches.
- Elimination of Irrelevant Activations: Irrelevant activations were removed by multiplying the computed status by the corresponding energy values.

This structured post-processing ensures that predicted energy consumption closely aligns with actual usage, enhancing the model's performance [19].

E. Performance Evaluation Metrics

The performance of the proposed model was evaluated for identifying the activation status of appliances and estimating their power consumption. Several metrics were utilized for this evaluation.

a. Activation State Metrics

For an ‘i-th’ appliance at time ‘t’ having the actual activation state denoted as $a_i(t)$ with $\hat{a}_i(t)$ representing the model's prediction of this state. Where $\hat{a}_i(t)$= 1 if the appliance is predicted to be on and $\hat{a}_i(t)$=0 if the appliance is predicted to be off. The following definitions apply:

- True Positive (TP): Correctly predicting the appliance as on.
- True Negative (TN): Correctly predicting the appliance as off.
- False Positive (FP): Incorrectly predicting the appliance as on when it is off.
- False Negative (FN): Incorrectly predicting the appliance as off when it is on.

Using these definitions, we can compute the following metrics:

- Accuracy: This represents the proportion of correctly predicted instances among the total predictions.

$$Acc = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{P+N}} \tag{8}$$

- Precision: This measures the proportion of true positive predictions to all predicted positives

$$Pre = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{9}$$

- Recall: This indicates the proportion of actual positives that were correctly identified.

$$Rec = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{10}$$

- F1 Score: The harmonic means of precision and recall, useful for imbalanced datasets used to train and test the model.

$$F1 = \frac{2 \times \mathrm{precision} \times \mathrm{recall}}{\mathrm{precision} + \mathrm{recall}} \tag{11}$$

- Matthews Correlation Coefficient (MCC): A measure of the quality of binary classifications.

$$MCC = \frac{TN \times TP - FN \times FP}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{12}$$

b. Power Estimation Metrics

To evaluate the accuracy of power consumption estimates, the following metrics were employed:

- Mean Absolute Error (MAE): The average absolute difference between predicted and actual values.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|x_i - x_i\right| \tag{13}$$

- Normalized Disaggregation Error (NDE): Measures normalized error between predicted and actual power usage

$$NDE = \frac{\sum_{i,t}(x_{i,t} - x_{i,t})^2}{\sum_{i,t} x_{i,t}^2} \tag{14}$$

- Normalized Signal Aggregate Error (SAE): Compares actual total energy (r) to predicted total energy ($\hat{r}$):

$$SAE = \frac{\left|\hat{r} - r\right|}{r} \tag{15}$$

- Relative error in total energy: Measures the relative error between the total energy (E) to predicted total energy ($\hat{E}$)

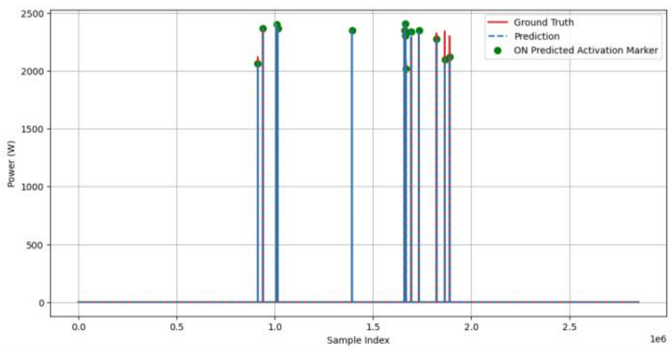
$$TRE = \frac{\hat{E} - E}{Max\left(E, \hat{E}\right)} \tag{16}$$

4. Results and Discussion

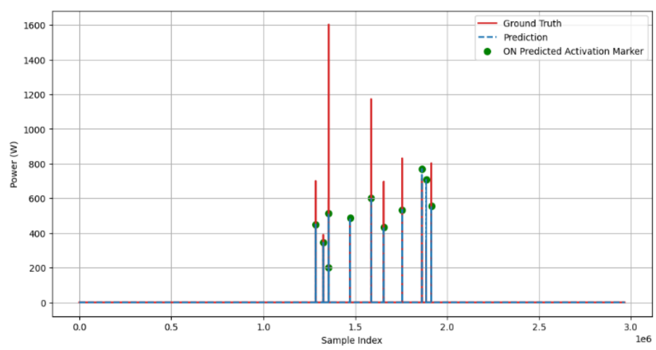
4.1. Testing in Seen Scenario (Seen Data from UK-DALE House-1 and REFIT House-11)

A. Results with the UK-DALE Dataset (House 1)

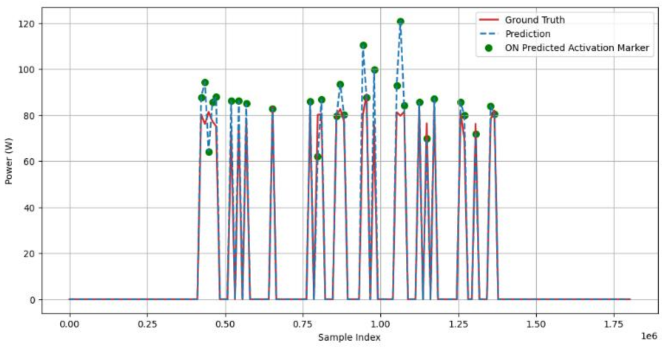
Testing the sequence-to-point CNN model on seen data from House 1 in the UK-DALE dataset highlights the significant impact of the data fusion. The Seen scenario refers to part of the test data from the selected houses used for model validation. The model was evaluated on individual appliances, including the kettle, microwave, fridge, dishwasher, and washing machine, with the goal of predicting clean, disaggregated signals.



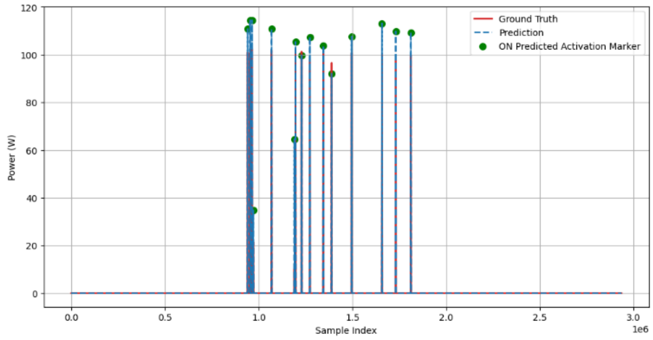
(a)



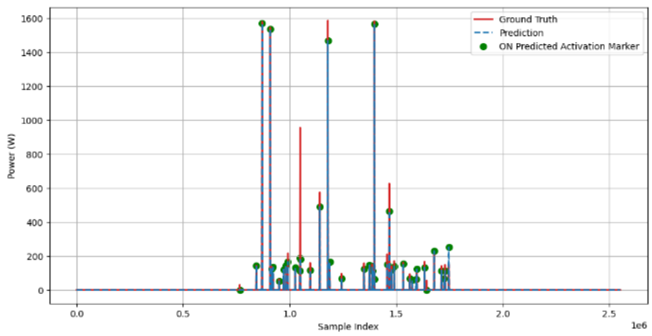
(b)



(c)



(d)



(e)

Fig.4. Disaggregation results of selected appliances on seen scenario from the UK-DALE dataset with post-processing (a) Kettle; (b) Microwave; (c) Fridge; (d) Dishwasher; (e) Washing machine

Performance metrics revealed exceptional classification and prediction accuracies. Accuracy values ranged from 0.929 for fridges to 0.999 for kettles and dishwashers, substantiating the model's reliability. The Matthews Correlation Coefficient (MCC) demonstrated significant variability, spanning from 0.387 for microwaves to 0.981 for Dishwashers. Similarly, F1 scores exhibited comparable variations, ranging from 0.371 to 0.981.

Table 3. Performance evaluation of the proposed algorithm in a seen scenario based on the UK-DALE datasets (House 1)

Appliance	Acc	MCC	F1	Without post-processing				With post-processing			
				MAE	NDE	SAE	TRE	MAE	NDE	SAE	TRE
Kettle	0.999	0.842	0.842	3.189	0.226	0.041	0.711	3.189	0.297	0.041	0.04
Microwave	0.992	0.387	0.371	7.164	0.861	0.775	0.736	7.14	0.863	0.779	-0.779
Fridge	0.929	0.856	0.916	8.588	0.154	0.014	0.148	8.668	0.159	0.02	-0.02
Dishwasher	0.999	0.981	0.981	0.754	0.017	0.001	0.008	0.754	0.017	0.002	-0.002
Washing M.	0.997	0.973	0.974	3.357	0.031	0.06	-0.047	3.438	0.032	0.065	-0.065
Overall	0.983	0.808	0.817	4.21	0.258	0.178	0.311	4.238	0.274	0.181	-0.165

Post-processing had an impact on MAE for appliances like microwave which reduced from 7.16 to 7.14. While there were no changes with kettle and dishwasher, it is slightly increased in the case of the fridge (8.588 Wh to 8.668 Wh) and washing machine, indicating inefficiencies in handling variable load profiles. While NDE and SAE improved for the kettle and dishwasher, the microwave retained high errors due to its sporadic, short-duration consumption pattern. TRE improved significantly for the kettle (0.711 to 0.04). Type-2 appliances such as washing machines and microwaves exhibited high errors due to their complex operational states, making them challenging to model with simple post-processing techniques. Fig. 4 shows the visual disaggregation results of the selected appliances after post-processing. Table 3 summarizes the performance of the model on seen data from House 1, with and without post-processing.

B. Results with the REFIT Dataset (House 11)

Table 4. compares appliance-level performance metrics of each appliance. The appliances demonstrated significant variability with the kettle and microwave achieving high accuracy of 0.996 while the fridge presented the most substantial analytical challenge, recording the lowest accuracy at 0.687. The dishwasher and washing machine exhibited moderate performance, with accuracies of 0.992 and 0.959, respectively.

Aggregate performance metrics indicated an average appliance accuracy of 0.926, accompanied by an average F1 score of 0.234 and a Matthews Correlation Coefficient (MCC) of 0.240. The kettle demonstrated moderate predictive reliability with an F1 score of 0.507 and an MCC of 0.536. The microwave exhibited comparatively lower predictive capabilities.

Post-processing had mixed effects on error metrics across different appliances. While the MAE for most appliances remained relatively stable, it reduced for the fridge (27.318 Wh to 26.997 Wh) and washing machine (9.161Wh to 8.865Wh), suggesting that post-processing did reduce absolute errors for these appliances. NDE and SAE improvements were marginal for most cases with the fridge showing a substantial increase in NDE (1.051 to 2.089), indicating a potential over-correction. The microwave and dishwasher retained high errors highlighting challenges in disaggregating appliances with sporadic or multi-state consumption patterns. TRE improvements were inconsistent, with the fridge showing a notable increase (0.368 to 0.895) while other appliances like the dishwasher and microwave exhibited reduction in the TRE values.

Quantitative improvements were seen, where the overall MAE decreased from 10.791 W to 10.666 W, NDE transitioned from 0.916 to 1.178, SAE reduced from 2.450 to 2.413, and TRE shifted from 0.320 to -0.122. These modifications underscored post-processing's pivotal role in refining energy prediction accuracy and reliability.

Table 4. Performance evaluation of the proposed algorithm in a seen scenario based on the REFIT datasets (House 11)

Appliance	Acc	MCC	F1	Without post-processing				With post-processing			
				MAE	NDE	SAE	TRE	MAE	NDE	SAE	TRE
Kettle	0.996	0.536	0.507	9.445	0.511	0.48	0.559	9.445	0.75	0.48	-0.48
Microwave	0.996	0.18	0.179	2.244	1.155	0.309	0.923	2.235	1.162	0.315	-0.315
Fridge	0.687	-0.003	0.011	27.318	1.051	8.618	0.368	26.997	2.089	8.491	0.895
Dishwasher	0.992	0.354	0.339	3.789	0.773	0.744	-0.505	3.786	0.778	0.745	-0.745
Washing M.	0.959	0.131	0.132	9.161	1.091	0.099	0.253	8.865	1.11	0.035	0.034
Overall	0.926	0.24	0.234	10.791	0.916	2.45	0.32	10.666	1.178	2.413	-0.122

In summary, post-processing emerges as a critical optimization strategy for the sequence-to-point CNN model in appliance energy disaggregation. The analysis definitively illustrates its significance in improving classification accuracy, minimizing predictive errors, and generating more reliable energy consumption predictions across diverse appliance types.

4.2. Testing in Unseen Scenario (Unseen Data from UK-DALE House-2 and REFIT House-20)

A. Results with the UK-DALE Dataset (House 2)

The sequence-to-point CNN model's ability to generalize was assessed with data that it had not encountered during training. The Unseen scenario refers to test data from the selected houses used for testing which was completely not used during training stage of the model. For this, we used the complete dataset from House-2 in UK-DALE, ensuring that the test data captured all target appliances (kettle, microwave, dishwasher, fridge, and washing machine). Table 5 provides the performance metrics which revealed substantial predictive accuracy variations across appliances, reflecting the influence of appliance-specific characteristics and noise levels.

Post-processing led to reductions in key error metrics, indicating improvements in NILM performance for specific appliances. The most notable decrease was observed in the total relative error (TRE) for all appliances, particularly for the microwave (0.783 to -0.796). These reductions suggest that post-processing corrected some overestimations, although at the cost of introducing negative TRE values which may indicate underestimation issues. Additionally, decreases in MAE were recorded for the microwave (5.999 Wh to 5.995 Wh) and dishwasher (31.158 Wh to 31.141 Wh), reflecting marginal improvements in absolute error reduction. However, the overall improvements were limited, as NDE and SAE values remained nearly unchanged, highlighting the challenge of effectively filtering errors for complex appliances.

The Table 5 compares appliance-level performance metrics with and without post-processing. These results reveal notable performance enhancements due to post-processing and some variability compared to the seen scenario, largely due to House-2's specific appliance usage patterns and noise levels. However, the overall findings confirm that the model effectively estimates power consumption of specific appliances in both familiar and entirely new household settings.

Table 5. Performance evaluation of the proposed algorithm in Unseen Scenario based on UK-DALE dataset (House 2)

Appliance	Acc	MCC	F1	Without post-processing				With post-processing			
				MAE	NDE	SAE	TRE	MAE	NDE	SAE	TRE
Kettle	0.999	0.93	0.93	7.047	0.125	0.261	0.439	7.047	0.136	0.261	-0.261
Microwave	0.994	0.384	0.353	5.999	0.89	0.795	0.783	5.995	0.892	0.796	-0.796
Fridge	0.877	0.755	0.851	12.412	0.296	0.147	-0.063	12.52	0.308	0.153	-0.153
Dishwasher	0.968	0.416	0.405	31.158	0.877	0.853	-0.834	31.141	0.878	0.854	-0.854
Washing M.	0.988	0.434	0.367	6.236	0.838	0.848	-0.808	6.262	0.841	0.853	-0.853
Overall	0.965	0.584	0.581	12.97	0.605	0.581	0.079	12.993	0.611	0.583	-0.583

B. Results with the REFIT Dataset (House 20)

The testing results on unseen REFIT houses using a sequence-to-point CNN model underscore the considerable impact of post-processing on appliance energy disaggregation performance. In this evaluation, data from previously unobserved homes was utilized to assess the model's ability to generalize effectively. Table 6 compares appliance-level performance metrics with and without post-processing for individual appliances including the kettle, microwave, dishwasher, fridge, and washing machine.

In Table 6, performance metrics revealed variability across appliances. The dishwasher and kettle demonstrated exceptional predictive capabilities, with the kettle achieving 0.998 accuracy, 0.794 Matthews Correlation Coefficient, and 0.789 F1 score. Conversely, the fridge displayed moderate performance with a 0.347 MCC and 0.617 F1 score. Aggregate metrics indicated an average accuracy of 0.929, and F1 score of 0.513. The initial mean absolute error (MAE) stood at 9.305 Wh, with a normalized disaggregation error (NDE) of 0.658.

Post-processing helped to reduce the key error metrics, particularly in total relative error (TRE) for multiple appliances, suggesting improved accuracy in specific cases. The microwave experienced a significant decrease in TRE, dropping from 0.801 to -0.853, indicating an overcorrection that may have resulted in underestimation. Similarly, the dishwasher's TRE declined from -0.039 to -0.124, and the washing machine saw a further reduction from -0.729 to -0.838, reinforcing the pattern of TRE corrections introducing negative values. Additionally, the fridge exhibited a slight improvement in SAE, decreasing from 0.049 to 0.038, suggesting an enhancement in scaling accuracy. While these reductions indicate some benefits of post-processing, the overall improvements were marginal. These results highlight the impact of post-processing in achieving accurate and appliance-level energy disaggregation.

Table 6. Performance evaluation of the proposed algorithm in Unseen Scenario based on REFIT dataset (House 20)

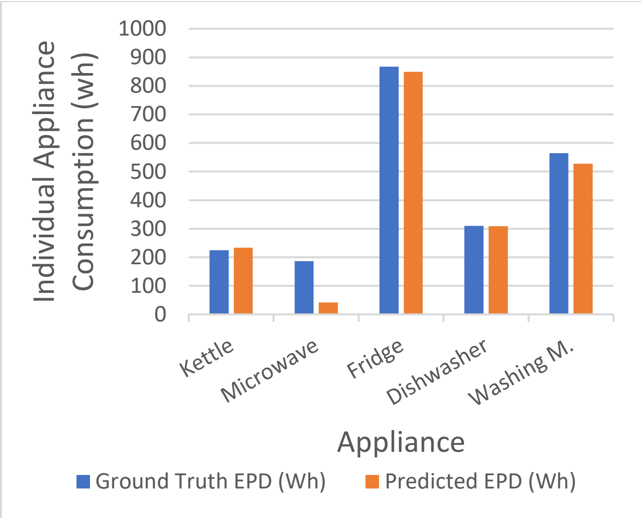
Appliance	Acc	MCC	F1	Without post-processing				With post-processing			
				MAE	NDE	SAE	TRE	MAE	NDE	SAE	TRE
Kettle	0.998	0.794	0.789	4.761	0.305	0.137	0.736	4.761	0.374	0.137	0.12
Microwave	0.993	0.146	0.142	5.558	0.988	0.848	0.801	5.537	0.99	0.853	-0.853
Fridge	0.683	0.347	0.617	27.731	0.852	0.049	0.29	27.721	0.859	0.038	0.036
Dishwasher	0.994	0.823	0.815	4.065	0.256	0.122	-0.039	4.059	0.256	0.124	-0.124
Washing M.	0.976	0.279	0.2	6.41	0.888	0.833	-0.729	6.414	0.916	0.838	-0.838
Overall	0.929	0.478	0.513	9.305	0.658	0.398	0.212	9.298	0.679	0.398	-0.332

The unseen REFIT test results clearly demonstrate the effectiveness of post-processing in improving the model's performance across all tested appliances. Post-processing enhances energy accuracy and reduces key error metrics, although it may have resulted in underestimation of the energy in some cases. The model benefits greatly from post-processing by eliminating irrelevant activations which lead to better disaggregation of energy consumption. While post-processing reduced some discrepancies, its tendency to introduce new errors in TRE suggests the need for a more refined approach that better preserves correct activations while minimizing misclassifications. This shows that the sequence-to-point CNN model, coupled with post-processing, is robust enough to generalize to unseen environments, offering reliable and accurate appliance-level energy predictions in real-world scenarios.

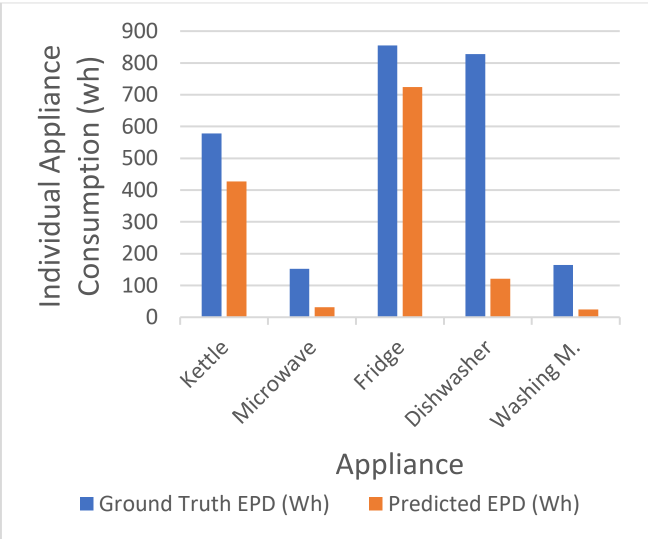
4.3. Energy Contributions by Selected Appliances

Evaluating the energy contributions of key appliances like the kettle, microwave, fridge, dishwasher, and washing machine provides critical insights into the performance of disaggregation algorithms in both seen and unseen cases. Beyond individual appliance analysis, it is essential to assess overall energy contributions within the NILM system to gauge the accuracy of the algorithm in estimating composite appliance power consumption.

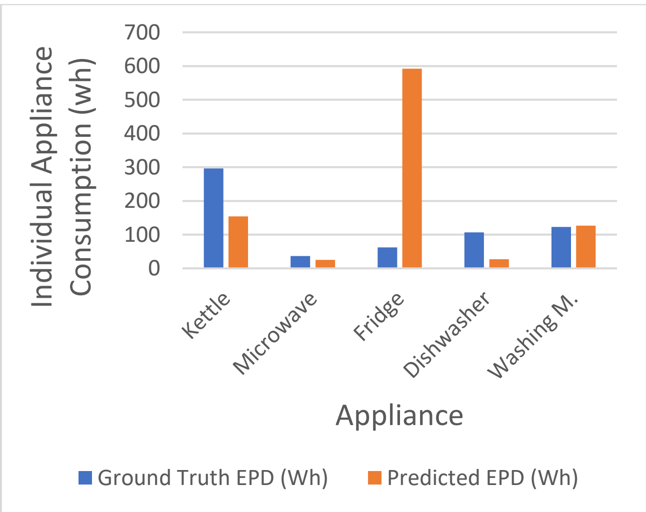
Fig. 4 illustrates appliance energy contributions in the UK-DALE and REFIT datasets, showing that predicted consumption is generally lower than the actual consumption due to irrelevant activations being filtered out. Notably, larger errors occur with type-2 appliances (dishwasher, washing machine) due to their complex, variable operational states, which DNN models struggle to capture accurately. In the UK-DALE House 1 seen scenario, appliance-specific predictions showed varied accuracy: kettle at 4% relative error (233.4 Wh predicted vs. 224.1 Wh actual), microwave at -77.9% relative error (41.3 Wh predicted vs. 187.0 Wh actual), fridge at -0.2%, dishwasher at -6.5%, and washing machine at -16.5%. The unseen UK-DALE House 2 scenario revealed further predictive variability. REFIT dataset analysis confirmed these findings, with House 11 showing lower relative errors for kettle and microwave, while House 20 exposed more substantial predictive challenges.



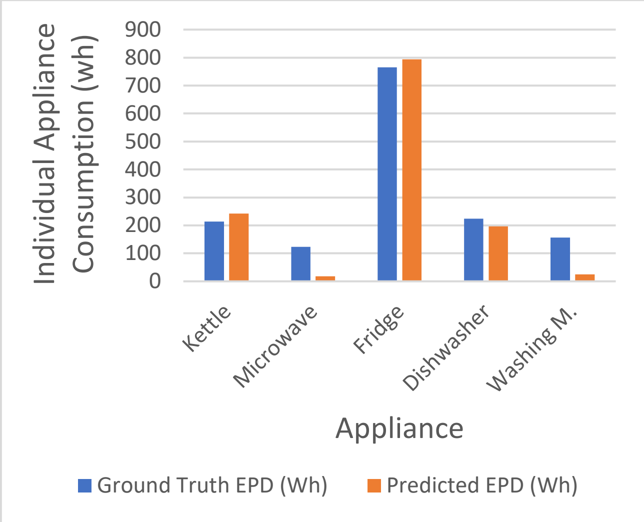
(a)



(b)



(c)



(d)

Fig.5. Energy Contributions by individual appliance from (a) seen scenario (UK-DALE Houses 1), (b) Unseen scenario (UK-DALE Houses 2), (c) seen scenario (REFIT HOUSE 11), (d) Unseen scenario (REFIT Houses 20)

These results highlight challenges in achieving precise NILM predictions, especially for appliances with multiple operating states.

Table 7. Details of energy contributions by selected appliances in UKDALE and REFIT

Dataset	Total Ground Truth EPD (Wh)	Total Predicted EPD (Wh)	Overall Relative Error (%)
UK-DALE Houses 1	2151.53	1960.277	-16.5
UK-DALE Houses 2	2578.96	1327.38	-58.3
REFIT Houses 11	624.93	925.92	-12.2
REFIT Houses 20	1482.59	1276.52	-33.2

Table 7 demonstrates the algorithm's performance on the UK-DALE and REFIT datasets. For UK-DALE, Houses 1 and 2 had overall relative errors of -16.5% and -58.3%, while REFIT Houses 11 and 20 showed errors of -12.2% and -33.2%. Though the algorithm reasonably estimates power consumption, these findings highlight the need to further algorithmic optimization to enhance predictive accuracy across diverse household environment.

4.4. Performance Comparison with State-of-the-Art Disaggregation Algorithms

A. Performance Evaluation of Disaggregation Algorithms in the Seen Scenario (UK-DALE House 1).

In evaluating energy disaggregation models, the proposed sequence-to-point (S-P) CNN model shows exceptional performance across key metrics including F1 Score, Precision, Recall, Accuracy, Relative Error, and Mean Absolute Error (MAE). This model significantly outperforms traditional methods such as Combinatorial Optimization, Factorial HMM, Autoencoder, Rectangles, and LSTM across all major metrics in Non-Intrusive Load Monitoring (NILM), as supported by previous research [3]. The disaggregation results comparing this work with existing models for the seen scenario are shown in Fig. 6.

The proposed sequence-to-point CNN model demonstrated exceptional performance across multiple metrics. F1 scores reached 0.842 for kettle, 0.916 for fridge, averaging 0.817 across appliances—significantly outperforming the LSTM model's 0.39. Precision values of 0.818 for kettle, 0.983 for dishwasher, and 0.988 for washing machine yielded an average of 0.853. Recall scores included 0.868 for kettle and 0.961 for washing machine, with a 0.797 average. Accuracy approached perfection at 1.0 for kettle and dishwasher, averaging 0.983 across appliances.

Relative error analysis showed minimal deviations from 0.04 to -0.779, while Mean Absolute Error averaged just 4.238 watts, substantially lower than the LSTM model's 68 watts. Overall, the CNN (S-P) model establishes itself as a reliable approach for energy disaggregation in NILM, outperforming other models in every metric and enhancing accuracy and energy-saving potential.

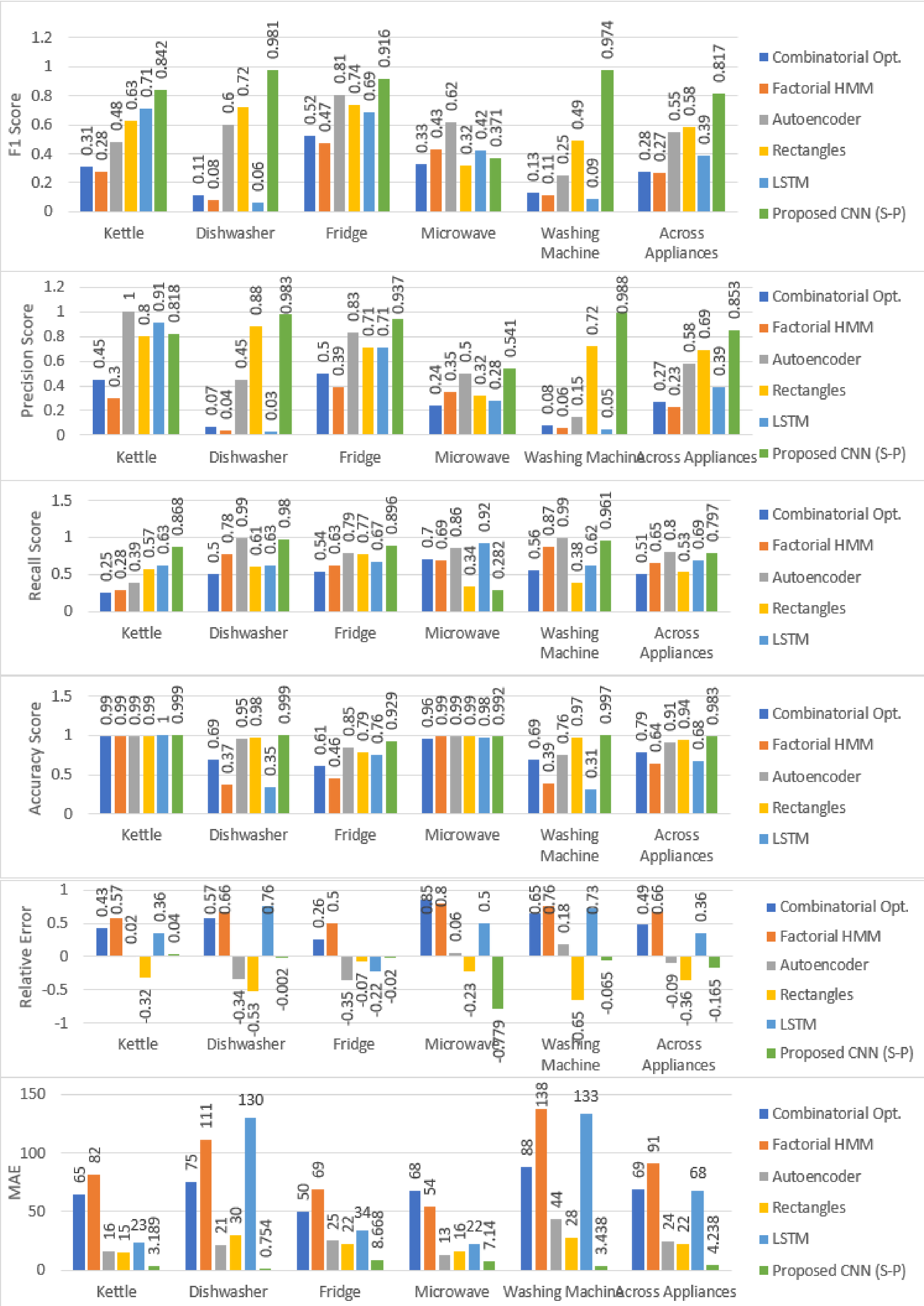


Fig.6. Disaggregation performance comparing this work with existing models in the seen scenario (UK-DALE House 1)

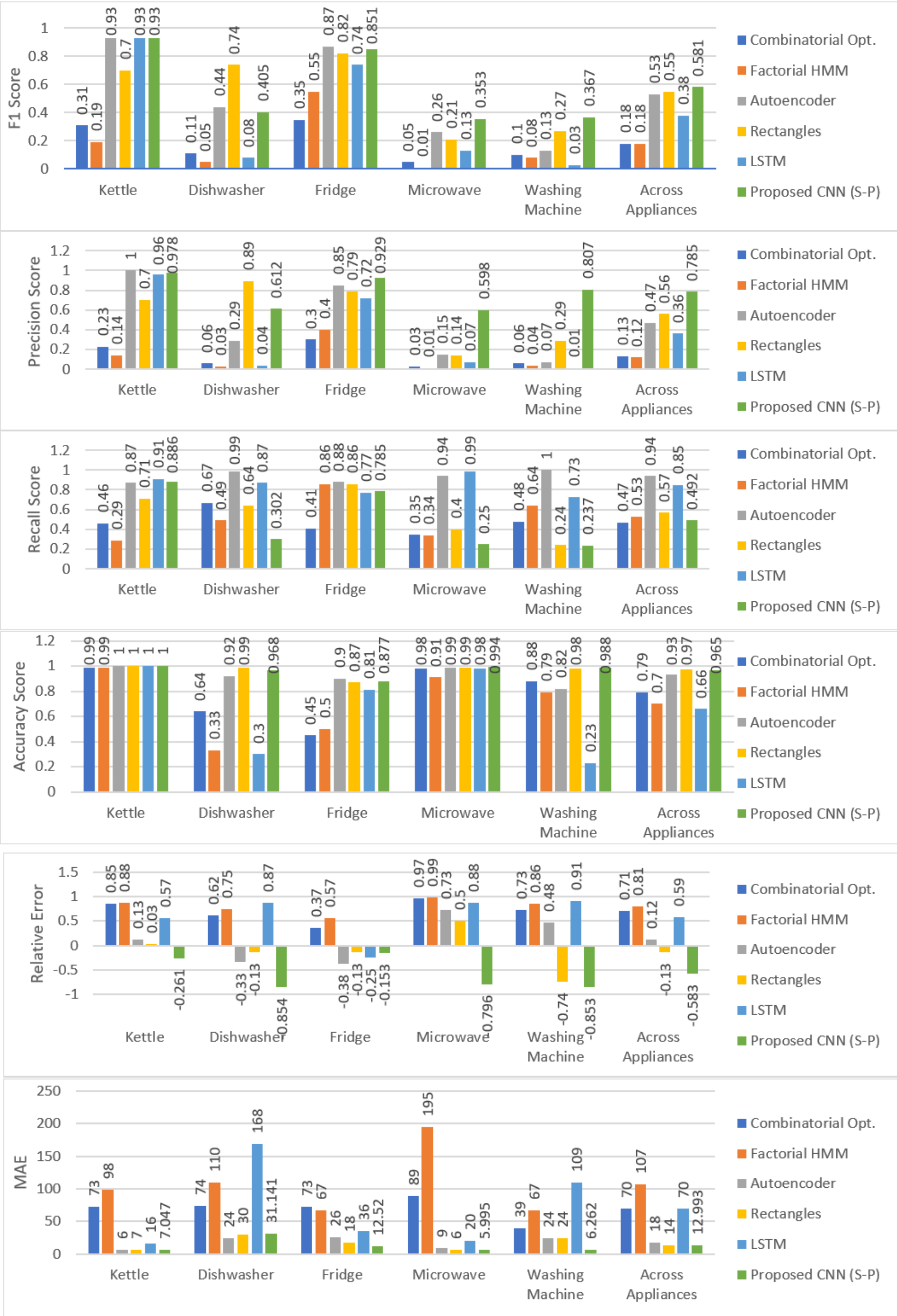


Fig.7. Disaggregation performance comparing this work with existing models in the Unseen scenario (UK-DALE House 2)

B. Performance Evaluation of Disaggregation Algorithms in the Unseen Scenario (UK-DALE House 2).

The evaluation shows that the proposed sequence-to-point CNN model outperforms traditional disaggregation models (e.g., Combinatorial Optimization, Factorial HMM, Autoencoder, Rectangles, and LSTM) across all major metrics in Non-Intrusive Load Monitoring (NILM), as supported by previous research [3]. The disaggregation results comparing this work with existing models for the Unseen scenario are shown in Fig. 7.

The sequence-to-point CNN model demonstrated sophisticated performance across appliances, with F1 Scores ranging from 0.93 for kettle to 0.851 for fridge, averaging 0.581—substantially exceeding traditional models' 0.18 to 0.55. Precision metrics achieved near-perfect scores for kettle (0.978) and fridge (0.929), averaging 0.785 across appliances. Recall showed variability, averaging 0.492, while overall accuracy reached 0.965—a significant improvement over traditional approaches' 0.3 to 0.79 range. Relative error analysis revealed minimal deviation at -0.583, and Mean Absolute Error averaged 12.993 watts. The model's superior performance across metrics establishes its effectiveness in energy disaggregation technologies, marking a substantial advancement in non-intrusive load monitoring methodologies.

5. Conclusions

The work has demonstrated the use of the sequence-to-point (S-P) CNN model for NILM. The model shows a strong potential in achieving load monitoring and energy disaggregation with high accuracy and error minimization that is enabled by advanced feature extraction techniques. The proposed algorithm demonstrates high accuracy in appliance-level energy disaggregation, especially in seen scenarios, with post-processing significantly improving performance metrics like MAE and NDE. In the UK-DALE and REFIT datasets (seen scenario), the model achieves overall accuracies of 98.3% and 92.6%, respectively, while maintaining adaptability in unseen scenarios with an accuracy of 96.5%. Due to the large window size, the latency of the disaggregation is slightly increased for a real time system. The developed model has an improved appliance identification accuracy compared to previous models and can perform the energy disaggregation in real-time on edge devices like the Raspberry Pi or high-performance computing, making it scalable and suitable for NILM applications.

References

[1] Zoha, A.; Gluhak, A.; Imran, M.A.; Rajasegarar, S. (2012). Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey. Sensors, 12, 16838–16866. https://dx.doi.org/10.3390/s121216838

[2] Zeifman, M.; Roth, K. (2011). Nonintrusive Appliance Load Monitoring: Review and Outlook. IEEE Trans. Consum. Electron., 57, 76–84. https://dx.doi.org/10.1109/TCE.2011.5735484

[3] Kelly, J.; Knottenbelt, W. (2015a). 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 (pp. 55-64).

[4] Kolter, J. Z.; Jaakkola, T. (2012). Approximate inference in additive factorial HMMs with application to energy disaggregation. In Artificial Intelligence and Statistics, 1472-1482. Retrieved from http://proceedings.mlr.press/v22/zico12.html

[5] He, W.; Chai, Y. (2016). An Empirical Study on Energy Disaggregation via Deep Learning. In 2nd International Conference on Artificial Intelligence and Industrial Engineering, 338–342. https://www.atlantis-press.com/proceedings/aiie-16/25866379

[6] Zhang, Y.; Yin, B.; Cong, Y.; Du, Z. (2020). Multi-state Household Appliance Identification Based on Convolutional Neural Networks and Clustering. Energies, 13 (4), 792. https://dx.doi.org/10.3390/en13040792

[7] Aghera, R.; Chilana, S.; Garg, V.; Reddy, R. (2021). A Deep Learning Technique Using Low Sampling Rate for Residential Non-Intrusive Load Monitoring. arXiv preprint arXiv:2111.05120. https://doi.org/10.48550/arXiv.2111.05120

[8] Jiao, X.; Chen, G.; Liu, J. (2023). A non-intrusive load monitoring model based on graph neural networks. In 2023 IEEE 2nd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA) (pp. 245–250). IEEE. https://doi.org/10.1109/EEBDA56825.2023.10090820

[9] Song, J.; Wang, H.; Du, M.; Peng, L.; Zhang, S.; Xu, G. (2021). Non-Intrusive Load Identification Method Based on Improved Long Short-Term Memory Network. Energies, 14 (3), 684. https://doi.org/10.3390/en14030684

[10] Zhang, C.; Zhong, M.; Wang, Z.; Goddard, N.; Sutton, C. (2018). Sequence-to-point learning with neural networks for non-intrusive load monitoring. In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), 32(1) https://doi.org/10.1609/aaai.v32i1.11873

[11] Hart, G. W. (1992). Non-intrusive appliance load monitoring. Proceedings of the IEEE, 80 (12), 1870-1891.

[12] Bonfigli, R.; Principi, E.; Fagiani, M.; Severini, M.; Squartini, S.; Piazza, F. (2017). Non-Intrusive Load Monitoring by Using Active and Reactive Power in Additive Factorial Hidden Markov Models. Applied Energy, 208, 1590–1607. https://doi.org/10.1016/j.apenergy.2017.08.203

[13] Zoha, A.; Gluhak, A.; Nati, M.; Imran, M.A. (2013). Low-Power Appliance Monitoring Using Factorial Hidden Markov Models. In Proceedings of the 2013 IEEE 8th International Conference on Intelligent Sensors, Sensor Networks and Information Processing: Sensing the Future, ISSNIP 2013, Melbourne, Australia, 2–5 April 2013; Volume 1, pp. 527–532. https://dx.doi.org/10.1109/ISSNIP.2013.6529845

[14] Lin, G. Y.; Lee, S. C.; Hsu, J. Y. J.; Jih, W. R. (2010). Applying Power Meters for Appliance Recognition on the Electric Panel. In Proceedings of the 2010 5th IEEE Conference on Industrial Electronics and Applications (ICIEA 2010) (pp. 2254–2259). https://doi.org/10.1109/ICIEA.2010.5515385

[15] Mauch, L.; Yang, B. (2015). A New Approach for Supervised Power Disaggregation by Using a Deep Recurrent LSTM Network. In Proceedings of the 2015 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2015, Orlando, FL, USA, 14–16 December 2015; pp. 63–67. https://dx.doi.org/10.1109/GlobalSIP.2015.7418157

[16] Varanasi, L. N. S.; Karri, S. P. K. (2024). Enhancing non-intrusive load monitoring with channel attention guided bi-directional temporal convolutional network for sequence-to-point learning. Electric Power Systems Research, 228, 110088.

[17] Xiong, C.; Cai, Z.; Liu, S.; Luo, J.; Tu, G. (2023). An improved sequence-to-point learning for non-intrusive load monitoring based on Discrete Wavelet Transform. IEEE Transactions on Instrumentation and Measurement, 72, 1-16.

[18] D'Incecco, M.; Squartini, S.; Zhong, M. (2020). Transfer Learning for Non-Intrusive Load Monitoring. IEEE Trans. Smart Grid, 11 (2), 1419–1429. https://dx.doi.org/10.1109/TSG.2019.2938068

[19] Rafiq, H.; Shi, X.; Zhang, H.; Li, H.; Ochani, M. K. (2020). A deep recurrent neural network for non-intrusive load monitoring based on multi-feature input space and post-processing. Energies, 13 (9), 2195. https://doi.org/10.3390/en13092195

[20] Le, T.-T.-H.; Kim, H. (2018). Non-intrusive load monitoring based on novel transient signal in household appliances with low sampling rate. Energies, 11(12), 3409. https://doi.org/10.3390/en11123409

[21] Manca, M. M.; Massidda, L. (2022). Deep learning based non-intrusive load monitoring with low resolution data from smart meters. Communications in Applied and Industrial Mathematics, 13(1), 39-56. https://doi.org/10.2478/caim-2022-0004

[22] Kelly, J.; Knottenbelt, W. (2015b). The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. Scientific Data, 2 (1), 150007. https://doi.org/10.1038/sdata.2015.7

[23] Johnson, M. J.; Kolter, J. Z. (2011). REDD: A public data set for energy disaggregation research. In Workshop on Data Mining Applications in Sustainability (SIGKDD) 25: 59–62. https://www.cs.cmu.edu/afs/cs.cmu.edu/Web/People/zkolter/pubs/kolter-kddsust11.pdf

[24] Murray, D.; Stankovic, L. (2016). REFIT: Electrical load measurements (Cleaned) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.5063428

[25] Pereira, L.; Nunes, N. (2018). Performance evaluation in non-intrusive load monitoring: Datasets, metrics, and tools—A review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(6), e1265. https://doi.org/10.1002/widm.1265

[26] Klemenjak, C.; Makonin, S.; Elmenreich, W. (2020). A synthetic energy dataset for non-intrusive load monitoring in households. Scientific Data, 7(1), 1-12. https://doi.org/10.1038/s41597-020-0434-6

[27] Reddy, R.; Garg, V.; Pudi, V. (2020). A feature fusion technique for improved non-intrusive load monitoring. Energy Informatics, 3 (1), 9. https://doi.org/10.1186/s42162-020-00112-w

[28] Laouali, I.; Ruano, A.; Ruano, M. D. G.; Bennani, S.D.; Fadili, H. E. (2022). Non-intrusive load monitoring of household devices using a hybrid deep learning model through convex hull-based data selection. Energies, 15 (3), 1215. https://doi.org/10.3390/en15031215

Authors' Profiles

Mr. John Oluwasegun Memud obtained the B.Eng. degree in Electrical and Electronics Engineering from the Federal University of Technology, Akure, Nigeria, in 2024. He is currently a Graduate Researcher in the Department of Electrical and Electronics Engineering at the Federal University of Technology, Akure. His research interests include Machine learning, Artificial Intelligence, Natural language processing, cyber security and communication systems.

Dr. Brendan Ubochi obtained the B.Eng. (2008), M.Sc. (2011) and Ph.D. (2018) degrees from The Federal University of Technology. Owerri, Nigeria, The University of Manchester, Manchester, United Kingdom and Swansea University, Swansea, United Kingdom, respectively. Since 2021, he has been a Senior Lecturer in the Department of Electrical and Electronics Engineering at the Federal University of Technology, Akure, focusing on modelling and characterization of nano-electronic devices, power electronics, and embedded systems.

Dr. Michael Rotimi Adu is currently an Associate professor with the department of Electrical and Electronics Engineering, Federal University of Technology, Akure, Nigeria. He is also the incumbent Ag. Head of Department. His research interest is in the area of Power Systems Engineering, Renewable Energy and Electrical machines.

Prof. Nnamdi Nwulu (Senior Member, IEEE) is currently a Full Professor with the Department of Electrical and Electronic Engineering Science, University of Johannesburg, and the Director of the Centre for Cyber Physical Food, Energy and Water Systems (CCP-FEWS). His research interests include the application of digital technologies, mathematical optimization techniques, and machine learning algorithms in food, energy, and water systems. He is a Professional Engineer registered with the Engineering Council of South Africa (ECSA), a Senior Member of the South African Institute of Electrical Engineers (SMSAIEE), and a Y-Rated Researcher by the National Research Foundation of South Africa. He is the Editor-in-Chief of the Journal of Digital Food Energy and Water Systems (JDFEWS) and an Associate Editor of IET Renewable Power Generation (IET-RPG) and African Journal of Science, Technology, Innovation and Development (AJSTID).

How to cite this paper: John Oluwasegun Memud, Brendan Chijioke Ubochi, Michael Rotimi Adu, Nnamdi Nwulu, "Non-intrusive Load Monitoring for Home Appliance Using Sequence-to-point Convolutional Neural Networks", International Journal of Intelligent Systems and Applications(IJISA), Vol.17, No.6, pp.75-93, 2025. DOI:10.5815/ijisa.2025.06.06