A Practical Solution Based on Convolutional Neural Network for Non-Intrusive Load Monitoring

Moradzadeh, Arash ; Mohammadi-ivatloo, Behnam; Abapour, Mehdi; Anvari-Moghaddam, Amjad; Gholami Farkoush, Saeid ; Rhee, Sang-Bong

Published in:
Journal of Ambient Intelligence and Humanized Computing

DOI (link to publication from Publisher):
10.1007/s12652-020-02720-6

Publication date:
2021

Document Version
Accepted author manuscript, peer reviewed version

Citation for published version (APA):
Moradzadeh, A., Mohammadi-ivatloo, B., Abapour, M., Anvari-Moghaddam, A., Gholami Farkoush, S., & Rhee, S.-B. (2021). A Practical Solution Based on Convolutional Neural Network for Non-Intrusive Load Monitoring. *Journal of Ambient Intelligence and Humanized Computing*, *12*(10), 9775-9789. https://doi.org/10.1007/s12652-020-02720-6

A Practical Solution Based on Convolutional Neural Network for Non-Intrusive Load Monitoring

Arash Moradzadeh[1], Behnam Mohammadi-Ivatloo[1,2,*], Mehdi Abapour[1], Amjad Anvari-Moghaddam[3,1], Saeid Gholami Farkoush[4,*], Sang-Bong Rhee[4,*]

ABSTRACT-In recent years, the introduction of practical and useful solutions to solve the non-intrusive load monitoring (NILM) as one of the sub-sectors of energy management has posed many challenges. In this paper, an effective and applicable solution based on deep learning called convolutional neural network (CNN) is employed for this purpose. The proposed method with the layer-to-layer structure and extraction of features in the power consumption (PC) curves of each household appliances will be able to detect and distinguish the type of electrical appliances (EAs). Likewise, the load disaggregation for the total home PC will be based on identifying the PC patterns of each EA. To do this, experimental evaluation of reference energy data disaggregation dataset (REDD) related to real-world data and measurement at low frequency is used. The PC curves of each EA are used as input data for training and testing the network. After initial training and testing by the PC data of EAs, the total PC of building obtained from the smart meter are used as input for each network in order to load disaggregation. The trained networks prove to be able to disaggregate the total PC for REDD houses 1, 2, 3, and 4 with a 96.17% mean accuracy. The presented results show the precision and efficiency of the suggested technique for solving NILM problems compared to other used methods.

Keywords- Residential load, non-intrusive load monitoring, deep learning, convolutional neural network

---

Arash Moradzadeh
arash.moradzadeh@tabrizu.ac.ir

Behnam Mohammadi-Ivatloo
mohammadi@ieee.org

Mehdi Abapour
mehdi.abapour@gmail.com

Amjad Anvari-Moghaddam
aam@et.aau.dk

Saeid Gholami Farkoush
saeid_gholami@ynu.ac.kr

Sang-Bong Rhee
rrsd@yu.ac.kr

Abbreviations

[1] Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran
[2] Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam
[3] Department of Energy Technology, Alborg University, Alborg, Denmark
[4] Department of Electrical Engineering, Yeungnam University, Gyeongsan, South Korea

ILM	Intrusive load monitoring	LP	Label power-set
NILM	Non-intrusive load monitoring	ML-kNN	Multilabel k-nearest neighbor
EA	Electrical appliance	LSTM	Long short-term memory
PC	Power consumption	CNN	Convolutional neural network
SVM	Support vector machine	REDD	Reference energy disaggregation data set
DTL	Decision tree learner	ReLU	Rectified linear unit

1. Introduction

Human daily life is influenced by many issues, such as promoting energy efficiency, finding economic and environmental solutions such as carbon reduction and effective management of energy supply and demand. This has created a great interest in implementation measures in supply, management, and energy saving in the field of science and technology (Morais and Castro 2019; Roy et al. 2020). Today, residential and commercial buildings consume nearly 60% of the world's electricity. Therefore, saving energy consumption of buildings can have a significant impact on reducing overall energy demand. Important management of an energy system requires information about the characteristics of loads or appliances (Moradzadeh et al. 2020a, c). Nowadays, load disaggregation is regarded as an important program for the effective monitoring and management of an electric power system. Load disaggregation requires a monitoring system to sense the significant parameters of the circuit and be able to identify and label existing activities (Mengistu et al. 2019; Quek et al. 2020). Intrusive load monitoring (ILM) and non-intrusive load monitoring (NILM) are two types of home electrical appliances (EAs) monitoring. In the ILM model, a sub-meter is appended to each EA between the socket and the appliance to achieve the consumption data. A large number of sensors required by this method and lack of attention to consumer privacy have made this method expensive and inopportune. In NILM, the energy consumption data collected on smart electrical meters are analyzed through the power consumption (PC) patterns of EAs without the need for recording sensors in the exclusive energy consumption of home appliances (Egarter et al. 2016). With this technology, the next generation of smart electrical meters can give customers an insight into how much energy they spend on their EAs or particular tasks. Accordingly, smart electricity meters are being installed in many parts of the world and based on previous research predicted that by the end of 2020, about 72% of European houses would have smart electricity meters (Devlin and Hayes 2019; Zeinal-Kheiri et al. 2020). Smart electricity meters record the total energy consumption of each customer in the building, and research has shown that using the data recorded by these meters can accurately identify PC patterns of home appliances and their energy consumption at different times (Kong et al. 2020).

Therefore, taking any technologies to improve NILM problems, which can detect the signature of personal appliance signals by reading the total PC, is of great importance in the scientific and industrial fields. As be observed in Fig. 1, an NILM system is built on four main steps. Identifying PC patterns of EAs and PC disaggregation of appliances that work simultaneously in a house is the principle of the NILM system (Morais and Castro 2019). The first step in NILM is accessing collected energy measurements at sufficient rate to detect specific load patterns. The sampling rate can be done in both high and low frequency. Since high-frequency sampling requires sophisticated hardware that results in additional costs in the monitoring process, low-frequency sampling is used instead. A low sampling rate of 1 Hz is typically utilized for NILM because it can be obtained by smart meters without modification. In the second step, behavioral data detection algorithms are used to identify the start and end of an energy consumption event. Every residential EA has its own unique consumption pattern, referred to as a home appliance signature. Detecting the PC of any EA by the collected load requires to detect the PC patterns of that appliance. The third stage begins with identifying PC patterns of EAs. Research shows that transient and steady-state features are two main class examples of household EA signatures used to recignize PC patterns of residential EAs. Steady-state patterns are intransitive for reconstructing the PC of each EA over time. Transient state patterns are short-term oscillations in the current or voltage before it becomes a steady state. In the transient state, due to the possibility of a low overlap of signatures based on the short duration of transient state, these state features are used to recognize PC patterns of household EAs. Selecting a suitable method or algorithm to extract the features and recognizing the PC patterns of each EA from the total data recorded in the smart electricity meter is the most important choice in this operation. After identifying the PC patterns of each EA, in the fourth step, it is time to identify the EA types. The better the feature learning and pattern recognition processes are, the more successful the identification operation will be (Liu et al. 2019; Morais and Castro 2019; Zeinal-Kheiri et al. 2020).

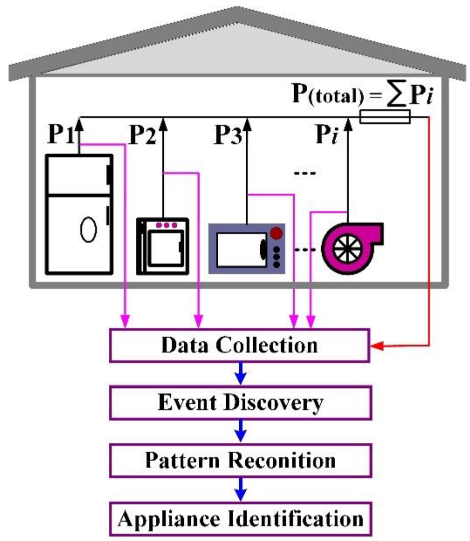


Fig. 1. Main steps of the NILM system methodology

A wide range of researches have been done to improve the accuracy and effectiveness of NILM methods.

Hart et al. first presented the NILM concept in 1986 (Hart et al. 1989), and expanded it in the seminal paper by Hart in 1992 (Hart 1992). Subsequently, increasing the smart electricity meters in recent years has had a huge impact on solving NILM problems (Klemenjak and Goldsborough 2016). Some clustering techniques for NILM analysis and identification of certain two-state (on/off) EAs are presented in (Wang and Zheng 2012) with high accuracy coefficient.

Some of the above approaches have some problems in identifying complex appliances that have several different performance modes. On the other hand, they do not perform well in recognizing multiple appliances simultaneously (Kelly and Knottenbelt 2012). Clustering techniques were used to resolve these deficiencies and identify the patterns of household EAs in (Beckel et al. 2014; Gajowniczek and Zabkowski 2015). In some simillir studies, different optimization-based approaches are proposed for load disaggregation (Piga et al. 2016; Bhotto et al. 2017; Zeinal-Kheiri et al. 2020). In some others, different techniques, based on the hidden Markov model, are used to solve the load disaggregation problems (Egarter et al. 2015; Ferrández-Pastor et al. 2017).

In (Gonçalves et al. 2011; Kim et al. 2011; Parson et al. 2014; Liu et al. 2019; Moradzadeh et al. 2020b) unsupervised techniques have been employed to solve the NILM problems. The technique used in (Liu et al. 2019) is based on a fuzzy algorithm called the entropy index of competitive aggregation constraints. In some recent studies, signal processing techniques have been proposed for load disaggregation problems (He et al. 2018). A Cepstrum-smoothing-based approach is described in (Kong et al. 2015) to solve NILM problems. This method is described to deal effectively with the concurrent on/off events of multiple EAs. In (Bhotto et al. 2017; Wittmann et al. 2018) NILM models are based on mixed-integer linear programming and integer programming, respectively.

Some applications of deep learning and machine learning focus on the activity classification of the appliances in question but do not consider their energy consumption estimates. For example, in (Basu et al. 2015; Tabatabaei et al. 2017; Moradzadeh et al. 2020d) different approaches of machine learning such as binary relevance, support vector machine (SVM), decision tree learner (DTL), label power-set (LP), and multilabel k-nearest neighbor (ML-kNN) have been proposed to detect all EAs in the system under study to advance NILM objectives. In (Mauch and Yang 2016), a long short-term memory (LSTM) network as a deep learning applications is selected to classify the types of EAs into a set. Deep learning applications called convolutional neural network (CNN) and auto-encoder have been employed in (Sirojan et al. 2018) to predict the energy consumption of each EA from the aggregated smart meter data. Also, to estimate the PC of EAs in consumer sites, three different hybrid structures incorporating recurrent network structures and CNNs have been used in (Kong et al. 2020). In the other work for this purpose, a 6-layer CNN has been adopted (Kelly and Knottenbelt 2015). A CNN technique for identifying multi-state appliances is presented in (Zhang et al. 2018), in which low-frequency power measurements are utilized. In (Gaur and Majumdar 2018) a deep learning based technique called transform learning is applied for solving the NILM issues. In (Singhal et al. 2019), improving the NILM models based on deep learning algorithms is represented, in which transform learning and deep dictionary learning algorithms are adopted.

The used method should be based on low-frequency data compatible with smart electricity meters. It should be appropriate for the identification of a large number of home EAs, and the proposed method should have a good identification performance for a large number of non-target appliances that are unknown to us.

In this paper, CNN, as one of the powerful structures of deep learning, is used for accurate disaggregation of power consumption curves of residential EAs and solve the NILM problem. Here, to increase operational accuracy, we propose a solution based on a deep CNN architecture and low-frequency PC measurements. The proposed approach has some key properties compared to other machine learning and deep learning-based solutions. This method, in addition to the extraction of features and inherent patterns of data and strong capability to avoid over-fitting, perfectly fits with the practical paradigm of this issue as an innovative processing solution. It is capable of effectively training even the low data of electrical PC of EAs and does not require trial and error tests to adjust the coefficients and parameters of the network during the training phase.

In deep learning applications, the structure of the network and how to select the input data for training and test the network are the most important criteria in the performance of any network. In this paper, the CNN is designed in such a way that in the shortest time and with the least complexity compared to the other methods mentioned in the literature review, recognize the behavior pattern of the least data recorded from the smart meter of the whole home. In addition, the contribution of the paper is technically expressed as follows and comparison with other valuable studies in Table 1:

- CNN is employed as a framework for solutions that are fully consistent with the practical model.
- Network training based on home appliance data is such that it requires minimum prerequisites and presuppositions for unseen households.
- The proposed technique provides a pattern of behavior and detailed information about the consumption of each appliance for a predetermined period of several hours.
- Input samples are selected in such a way that high-dimensional samples are divided into multiple samples so that instead of increasing the dimension of each sample, the number of samples from one class is high. With this behavior, the problems and errors related to the overfitting problems in the fully-connected layers are eliminated to ensure the performance of the method, even if a large number of home appliances are non-target appliances about which we have no information.
- The CNN method has been layered in such a way that based on the extraction of features related to the behavior of each EA, no direct subnet measurement data is required to identify a new unseen household.
- Selecting the optimal number of the CNN layers, kernel size, and fully interconnected layers, and optimizing network parameters is one of the most important factors in achieving the highest values of the performance in this paper that the other papers suffer from.

TABLE 1. Contribution of this paper in comparison to other valuable studies

Ref.	**Technique**	**Number of evaluated appliance**	**Description**	**Limitations**
(Hosseini et al. 2019)	Markovian and Recurrent Neural Network (RNN)	5 appliance	- On and off mode is considered for each EA. - The network is trained with suitable data. - For each EA, the transient power signal was selected as the input and output vector to train the CNN.	- Predictions are time-dependent and time intervals are effective in identifying the PC behavior of EAs. - It is not possible to predict unseen households. - Natural noise is effective in identifying and disaggregating the total PC of the household. - One type of dataset is utilized. - In all used datasets, only some of the special EAs in the household have been evaluated.
(D'Incecco et al. 2020)	Transfer Learning	5 appliance	- Convert the PC curves to images to clearly show the changes and behavior in them. - Method testing on different datasets. - View features in the PC curves of EAs. - Utilized technique is fully consistent with the practical model.	- Predictions are time-dependent and time intervals are effective in identifying the PC behavior of EAs. - It is not possible to predict unseen households. - Natural noise is effective in identifying and disaggregating the total PC of the household. - The proposed method requires a large amount of memory to process data and convert it to images.

				- In all used datasets, only some of the special EAs in the household have been evaluated.
(Morais and Castro 2019)	ANNs	7 appliance	- For each EA, the transient power signal was selected as the input and output vector to train the network. - Method testing on different datasets. - On and off mode is considered for each EA. - The network is trained with suitable data.	- It is not possible to predict unseen households. - Natural noise is effective in identifying and disaggregating the total PC of the household. - In all used datasets, only some of the special EAs in the household have been evaluated. - Predictions are time-dependent and time intervals are effective in identifying the PC behavior of EAs. - ANNs suffer from the problems of overfitting in the face of big data.
(Moradzadeh et al. 2020b)	PCA	13 appliance	- Behavioral patterns and features are clearly seen in the PC curves of household EAs. - Operates without any dependence on input data and extracts their features. - Utilized technique is fully consistent with the practical model. - For each EA, the transient power signal was selected as the input and output vector to PCA. - Evaluates the behavior of all EAs in the household.	- It is not possible to predict unseen households. - Natural noise is effective in identifying and disaggregating the total PC of the household. - One type of dataset is utilized and various datasets can be employed for future work.
(Welikala et al. 2019)	PBN	15 appliance	- For each EA, the transient power signal obtained directly in the smart meter was selected as the input and output vector. - The priori probabilities of individual appliances at the current time instant are calculated using a developed fuzzy system. - Evaluates the behavior of all EAs in the household. - The most important information about each EA based on its consumption behavior trend is considered for the employed NILM method to find the best probable currently turned ON EA combination.	- One type of dataset is utilized. - Natural noise is effective in identifying and disaggregating the total PC of the household. - Predictions are time-dependent and time intervals are effective in identifying the PC behavior of EAs. - It is not possible to predict unseen households. - One type of dataset is utilized and various datasets can be employed for future work. - The method used requires complex mathematical calculations that may be difficult to predict in the face of big data from PC curves.
This paper	CNN	15 appliance	- Utilized technique is fully consistent with the practical model. - The network is trained in such a way that without time dependence, only based on the behavior pattern of each appliance can disaggregate the total PC load of the household. - Selecting the optimal number of layers and network parameters to extract the most and useful features from the consumption curves of each EAs. - Ability to predict data related to unseen households. - Ineffectiveness of natural noises in PC curves of EAs. - Trained networks are tested and validated with new and unseen data related to PC curves of appliances. - For each EA, the transient power signal was selected as the input and output vector to train the CNN. - Evaluates the behavior of all EAs in the household.	- The extracted behavioral pattern features of each EA are not available and visible. - One type of dataset is utilized and various datasets can be employed for future work.

To implement the suggested technique, the low-frequency data of the PC obtaining at the meter corresponding to reference energy disaggregation data set (REDD) (Kolter and Johnson 2011) is utilized. This dataset includes detailed PC information from several homes in the real world.

The rest of this paper is structured as follows: Section II outlines the methodology for CNN. Section III describes the studied case. The CNN design and how to extract features from the PC curves are described in Section IV. Section V depicts the results. The comparisons of the methods used to solve the NILM problems are presented in section VI. Finally, the conclusions are presented in Section VII.

2. CONVOLUTIONAL NEURAL NETWORK (CNN)

CNN, as a multistep neural network, has made great achievements in the classification of large-scale images and extracting data features in recent years (Lecun et al. 2015; Moradzadeh and Pourhossein 2019a). As seen in Fig. 2, CNN consists of 3 layers; one or more convolutional layers for extracting the features, pooling layers for compiling these features, fully connected layers performed to forecast the feature vector in the output of the network, and finally, a Softmax function to classify the features in the last layer of the fully connected layers (Moradzadeh and Pourhossein 2019b; Kong et al. 2020). A nonlinear activation function called rectified linear unit (ReLU) is utilized in the convolutional step as follows (Anthimopoulos et al. 2016; Han et al. 2019):

$$C_r^n = ReLU\left(\sum_m v_{r-1}^m * w_r^n + b_r^n\right) \quad (1)$$

where C_r^n is the n^{th} filter output in convolutional layer r, v_{r-1}^m is the m^{th} output of the last layer $r-1$, $*$ indicates the convolution and ω_r^n indicates the n^{th} filter kernel of the running layer r, andb_r^n shows the bias (Han et al. 2019). In each convolution layer, after identifying the features, the pooling layers act to compile the significant and most extracted features from each filter and moving it to the next convolutional layer as input (Anthimopoulos et al. 2016; Moradzadeh and Pourhossein 2019a). The Max pooling that takes the most substantial part of the input is depicted as (Han et al. 2019):

$$h_{r+1}^n = \max\ C_r^n(t), \quad (\mathrm{i}-1)\mathrm{l}+1 \le \mathrm{t} \le \mathrm{il} \quad (2)$$

where C_r^n shows the feature map, $h_{r+1}^n$is the pooling layer result, l is the local area length for pooling, i is the number of input features to the pooling layer, and t is the number of collected features by the Max pooling. After extracting and pooling the features, fully connected layers are formed to forecast the feature vector in the ultimate output of the network (Han et al. 2019). In the final step, a Softmax function is used to do the classification as follow) Dong et al. 2016):

$$O_j = \begin{bmatrix} P(y=1)|x;\theta \\ P(y=2)|x;\theta \\ \dots \\ P(y=k)|x;\theta \end{bmatrix} = \frac{1}{\sum_{j=1}^{k} \exp(\theta^j x)} \begin{bmatrix} \exp(\theta^1 x) \\ \exp(\theta^1 x) \\ \dots \\ \exp(\theta^k x) \end{bmatrix} \quad (3)$$

where, k is the number of categories and $\Theta^j x$ represents the classification layer factors (Moradzadeh and Pourhossein 2019b).

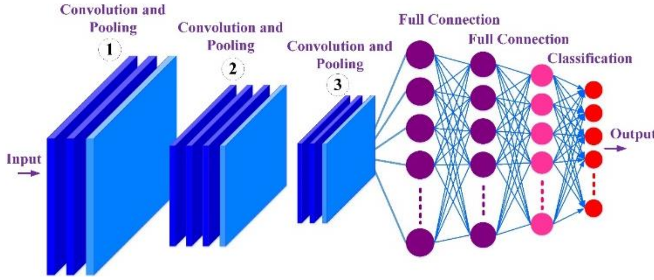


Fig. 2. Layer-to-layer structure of a Convolutional Neural Network

3. CASE STUDY

In this paper, the REDD, which includes aggregated and disaggregated EA data for six households in Massachusetts, USA is used to test the NILM applications (Kolter and Johnson 2011). The collection uses a smart power meter to record data. The measurements are related to 2 weeks with the 1s sampling period and some household appliances with the 3s sampling period (Kolter and Johnson 2011). To apply the proposed method, data from house 1, house 2, house 3, and house 4 of the REDD dataset were selected. Table 2 represents the types of household EAs and their target number for train CNN.

4. CNN DESIGN

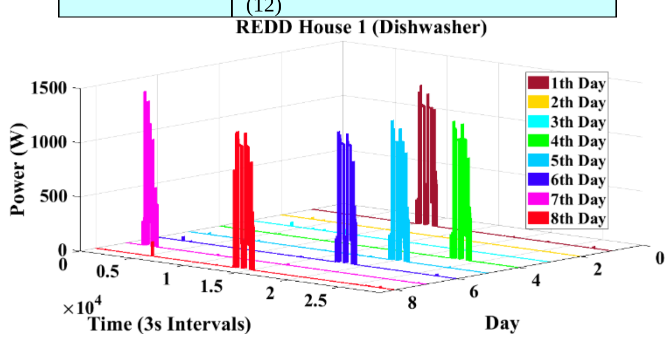
The design and structure of networks play an important role in its efficiency. The optimal selection of the number of network layers and the size of the convolution kernel is obtained by separate experiments. In optimizing one parameter, the other parameters are kept constant and initial experiments are started to narrow the search range to select the optimal value of the desired parameter. The selection of the most optimal structure and architecture is the most important criterion of the network used so that it can extract even the smallest features in the behavioral pattern of PC of household EAs. The CNN used in this paper consists of an input layer, three convolution layers, three max pooling layers, a fully connected layer, and a classifier layer. The structure of each layer is described as follows:

4.1. INPUT LAYER

The input layer reads data and generates output to the first convolution layer. In this paper, PC related to eight days of home EAs was considered as input. The one - day PC was assumed a curve, and totally, eight one-day PC curves from each EA were forming the input layer. Fig. 3 shows samples of PC curves for the four types of EAs from REDD houses that were considered as inputs.

TABLE 2. Names of the EAs considered from REDD houses

REDD Houses	Appliances (Target)
House 1	Wall Oven (1), refrigerator (2), kitchen outlets (4), dishwasher (3), lighting (5), washer dryer (6), microwave (7), stove (10), bathroom ground fault interrupter (GFI) (8), electric heat (9), different (11)
House 2	Kitchen outlets (1), dishwasher (2), lighting (3), washer dryer (4), microwave (5), refrigerator (6), garb (7), stove (8), different (9)
House 3	Electronics (1), bathroom GFI (2), refrigerator (3), kitchen outlets (4), dishwasher (5), furnace (6), washer dryer (7), microwave (8), smoke alarms (9), lighting (10), garb (11), unknown (12), different (13)
House 4	Lighting (1), furnace (2), kitchen outlets (3), washer dryer (4), stove (5), air conditioning (6), smoke alarms (7), dishwasher (8), bathroom GFI (9), disposal (10), unknown (11), different (12)



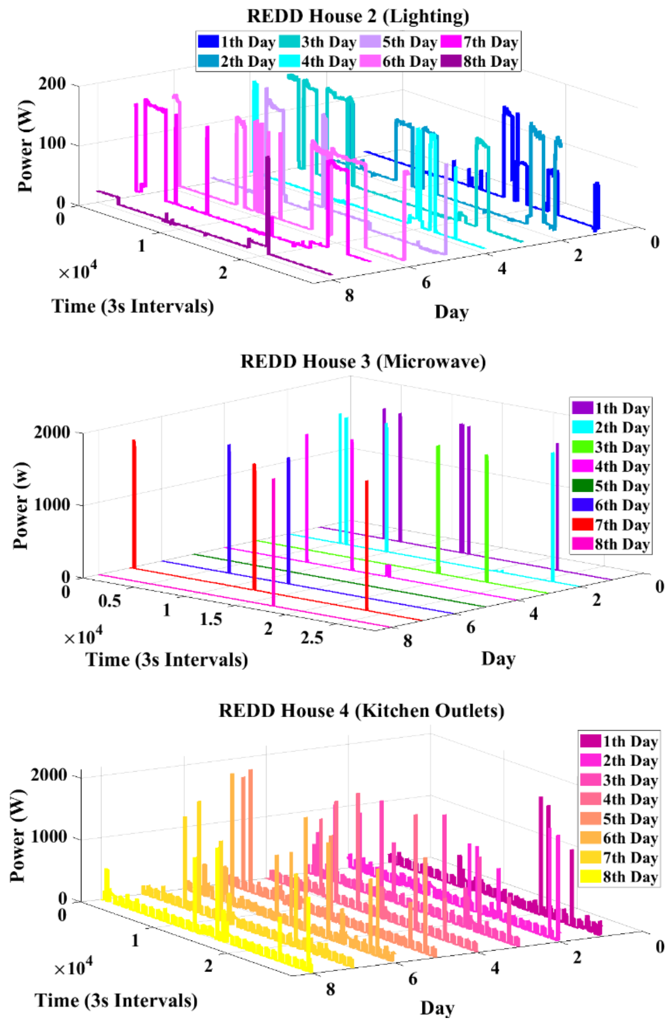


Fig. 3. Examples of PC curves for the two types of EAs from two houses

4.2. CONVOLUTION LAYER

Convolution layers are made up of kernels as filters that are responsible for the extraction of features. During the forward propagation process, the features extracted from each layer acts as the next layer input and are convolves with each convolutional kernel in the next layer to produce convolved feature maps (a sample of convolution process presented in Fig. 4) (Patterson and Gibson 2017). The size of the kernel in each convolution layer affects the performance of feature extraction directly. The small size of the kernel leads to information loss and its large size will increase the computational cost rapidly (Dong et al. 2016; Peng et al. 2019). In this study and in the first, second, and third convolution layers, 5, 16, and 20 filters were used, respectively. The filter size of the first convolutional layer is 4×4, and this size for the second and third layers is 3×3. Stride is set to 1. In this paper, the conventional size of the kernels relative to the size of the inputs was used to prevent the data missing and redundant computations in feature extraction. It should be noted that the approving and selecting this size of kernels was performed after optimization of their and obtaining the most appropriate values of network performance. In each convolution layer, a ReLU activation function has been used. It has been proven that ReLU speeds up the training process and avoids over-fitting better than other activation functions.

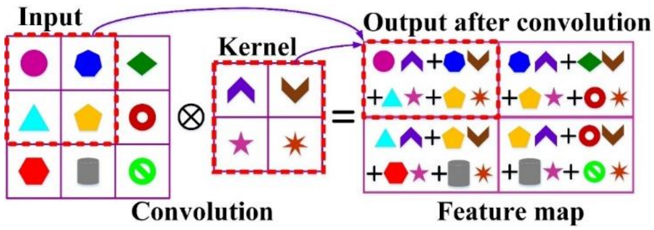


Fig. 4. Sample of convolution process: a 2×2 kernel convolves with a 3×3 input feature map to produce a 2×2 feature map

4.3. POOLING LAYER

In each convolution layer, the pooling layer is used to pool the features with the largest value and reduce the dimensions of the extracted features. Average pooling and max pooling are two common types of pooling (Patterson and Gibson 2017). Average pooling takes the mean of input features, which may lose some useful data. Max pooling prevents the loss of features or information that may be lost in average pooling by taking the greatest value of the features (Bagheri et al. 2018; Peng et al. 2019). The max pooling has been used in this design. Fig. 5 illustrates max pooling. Each max pooling layer has a window dimension of 2×2 and stride is set to 2.

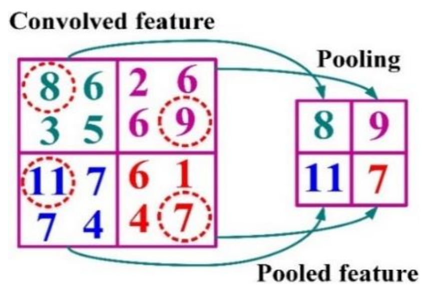


Fig. 5. Illustration of a 2×2 max pooling window

4.4. FULLY-CONNECTED LAYER

At the end of the CNN, there are a number of fully connected layers. Here, only one fully-connected layer is used. To complete the CNN structure, a Softmax layer is added to the fully-connected layer. The Softmax layer performs the classification and prediction of the feature vector at the final output of the network. Fig. 6 shows the flowchart of CNN design and use.

The PCs of residential EAs are used as inputs to the CNN input layer. The dataset was randomly divided into two sections: training and testing datasets. The type of appliances was considered as target data.

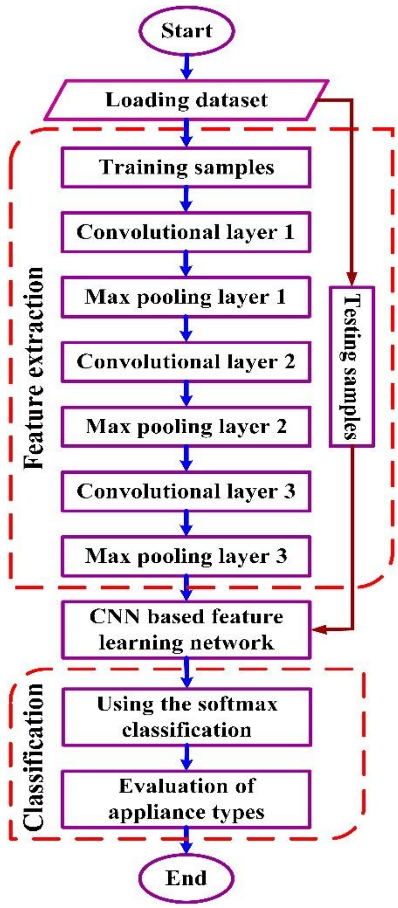


Fig. 6. Flowchart of CNN design and use

In each dataset, 65% of data was considered as training data and 35% was used to test the network. The training dataset was used to regulate the parameters of the network via comparing output with the target. During the process of CNN training, network outputs are calculated using the network's current weights and network inputs. Network error is calculated using the outputs and sample labels. A backpropagation algorithm is used to calculate the error derivative to the network weight. Finally, the weight update method is used to update the weights. By extracting the features of PC curves, their original nature and their consumption patterns can be achieved.

5. EXPERIMENTAL RESULTS

The designed network is trained with training data. The test data selected from the input data is used to test the network training. Fig. 7 shows the network detection and classification results for the training and testing data on the PC of the REDD house 1 EAs. Another network with the same design characteristics of the previous network is trained using data from the REDD house 2. Fig. 8 shows the network detection and classification results for the training and testing data on the PC of the REDD house 2 EAs.

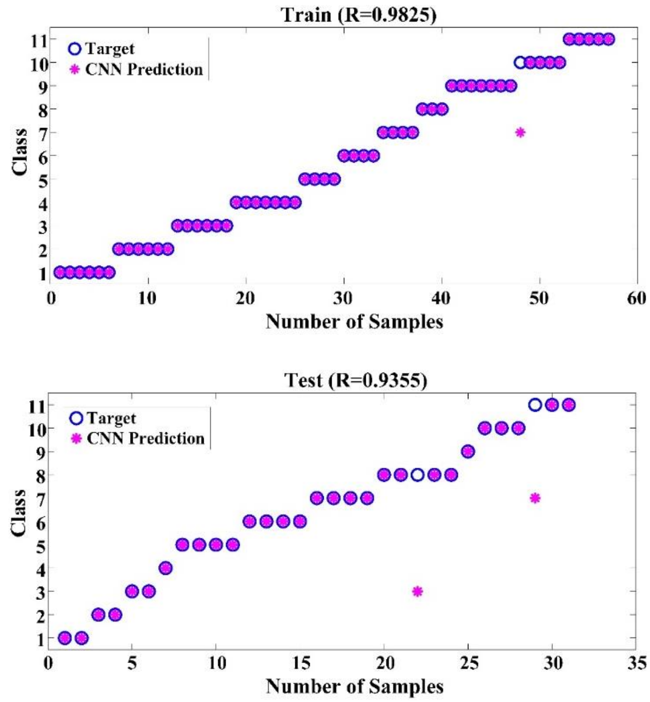


Fig. 7. CNN classification for the training and testing data of the REDD house 1 EAs

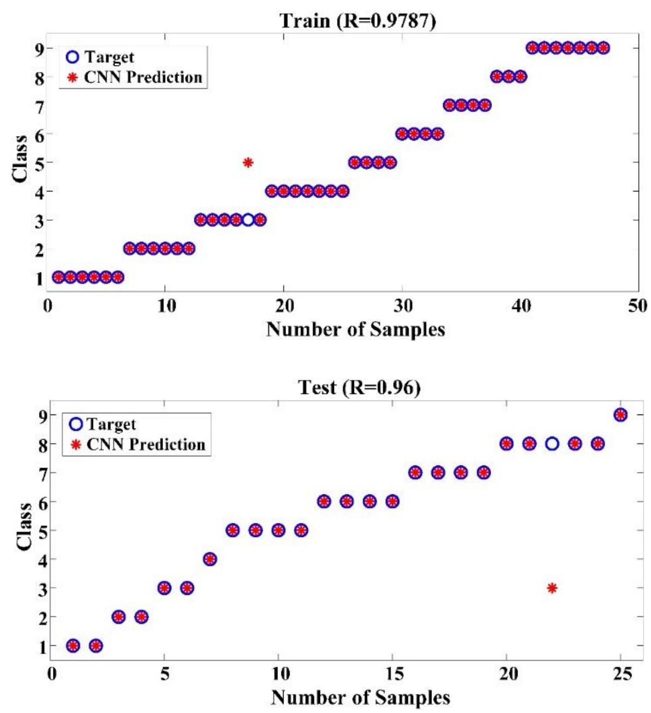


Fig. 8. CNN classification for the training and testing data of the REDD house 2 EAs

High correlation between target data and CNN prediction, and classification of household EAs type based on their PC patterns in REDD house 1 with a 98.25% accuracy for training and 93.55% for testing the network were shown in Fig. 7. Due to Fig. 8, this classification was done with 97.87% accuracy for training and 96% accuracy for testing the network in REDD house 2. Other networks are trained with the same specifications introduced for previous network designs and using data related to REDD house 3 and REDD house 4. The results of network detection and classification for training and testing data of PC for home EAs of REDD house 3 and REDD house 4 are presented in Figs. 9 and 10, respectively.

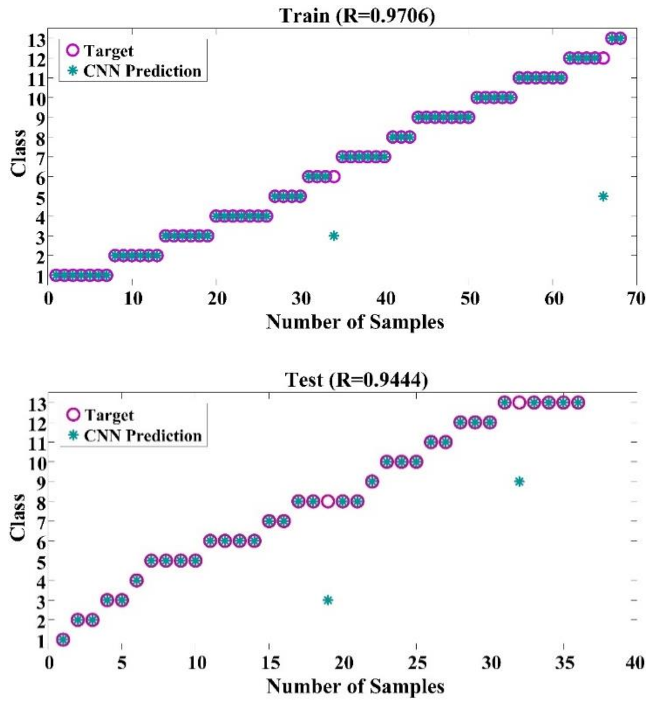


Fig. 9. CNN classification for the training and testing data of the REDD house 3 EAs

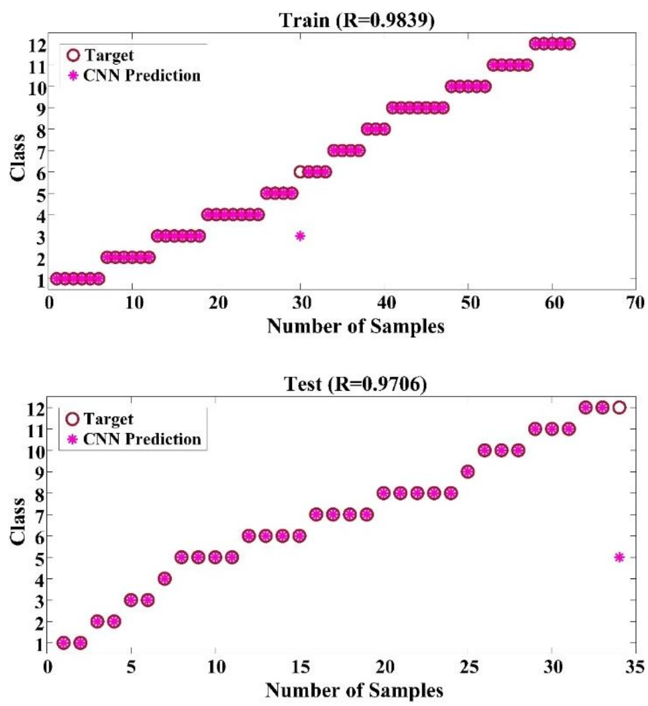


Fig. 10. CNN classification for the training and testing data of the REDD house 4 EAs

As with the above Figures, also a good correlation between the actual data and the CNN predict has been shown in these Figures. The EA types classification was done in REDD house 3 with a 97.06% accuracy for training and 94.44% for testing the network. This classification in the REDD house 4 was done with 98.39% accuracy for training and 97.06% for testing the network. It can be seen that the designed networks were able to do this with considerable accuracy. Trained networks containing extracted features of data (electrical PC curves) are saved. Given that these networks have the necessary training, they can be used to identify new data that was not among the input data. Now to test the saved networks to identify the types of EAs using their PC patterns, we consider some new examples of their PC as inputs to the networks. To do this, 15 samples of PC of each EA from each home were considered for six hours. The network testing to identify new data and its results must be evaluated via the metrics accuracy. Accuracy measure is calculated as follows (Morais and Castro 2019):

$$Accuracy = \frac{TI}{TI + FI} \tag{4}$$

where TI and FI demonstrates the number of samples that were truly and falsely identified, respectively.

Table 3 represents the results taken for each network for the testing data of all REDD houses 1, 2, 3, and 4. Figs. 11 and 12 show the confusion matrix for the network for the testing data of REDD house 1 and REDD house 2, respectively. Figs. 13 and 14 show the confusion matrix for the network for the testing data of REDD houses 3 and 4, respectively.

Table 3. Results for each network for the testing data of all REDD house 1, 2, 3, and 4

Class	**Accuracy (%)**			
Appliance	REDD House 1	REDD House 2	REDD House 3	REDD House 4
Dishwasher	100	100	100	100
Kitchen outlets	100	100	100	100
Lighting	100	93.3	100	100
Washer dryer	100	100	93.3	100
Stove	100	100	100	100
Refrigerator	93.3	100	100	-
Microwave	93.3	100	100	-
Bathroom GFI	100	-	93.3	100
Furnace	-	-	100	100
Electronics	-	-	100	100
Garb	-	100	93.3	-
Wall oven	100	-	-	-
Smoke alarm	-	-	-	86.6
Air conditioning	-	-	-	100
Electric heat	100	-	-	-
Unknown	-	-	100	93.3
Different	86.6	93.3	86.7	86.6
General	97.6	98.6	97.45	97.25

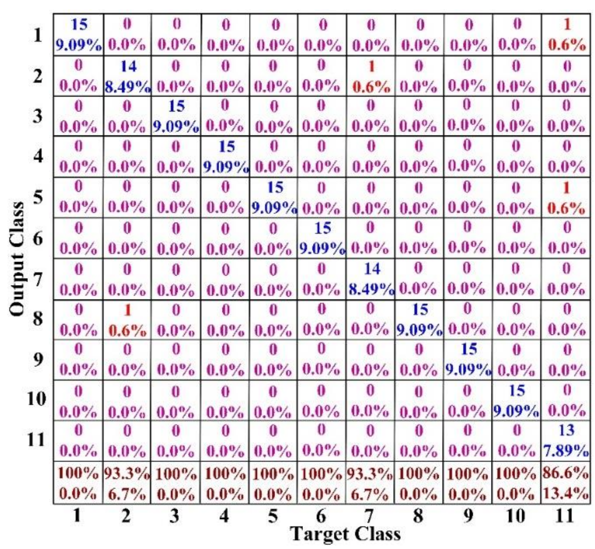


Fig. 11. REDD house 1 confusion matrix

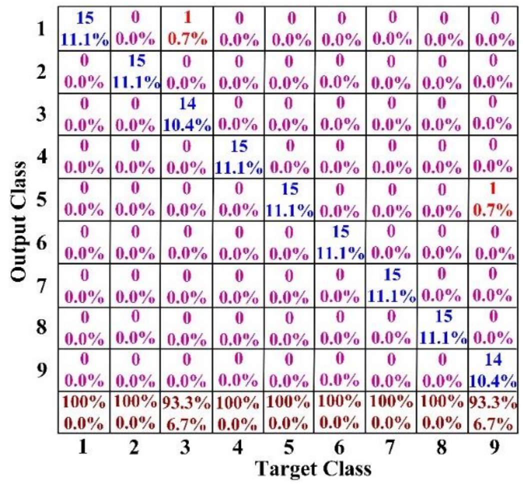


Fig. 12. REDD house 2 confusion matrix

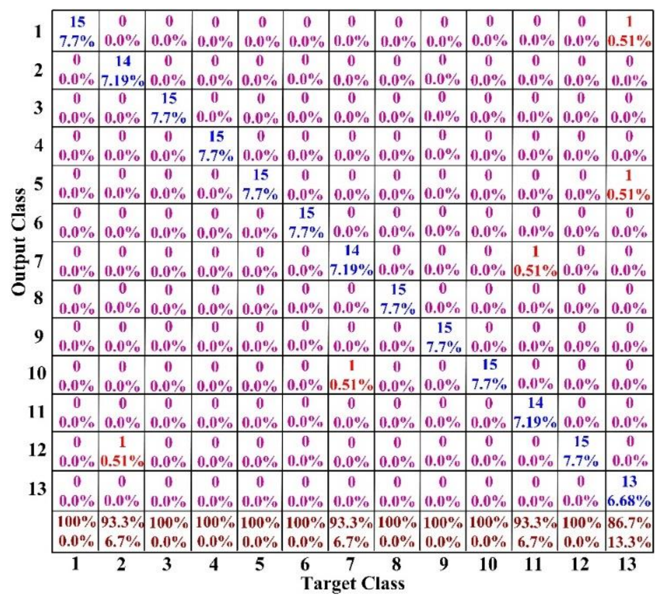


Fig. 13. REDD house 3 confusion matrix

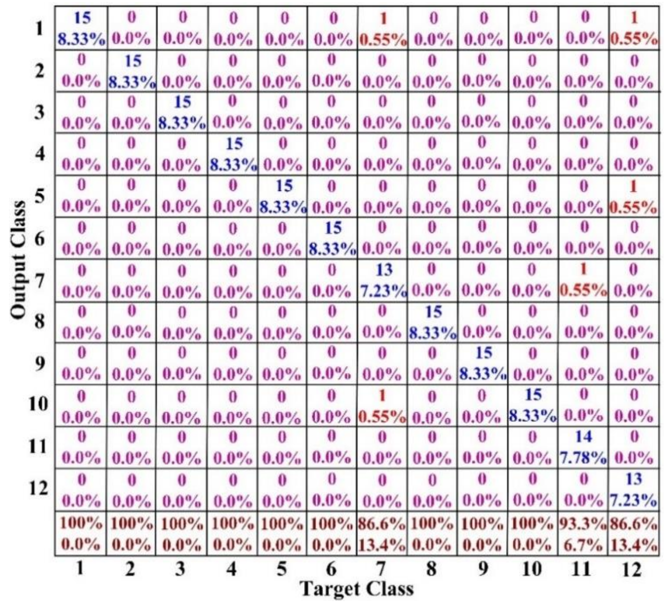


Fig. 14. REDD house 4 confusion matrix

Due to the Figs. 11 and 12, it can be seen that the trained network was prosperous in the identification proceeding of the EAs with 97/6% accuracy for REDD house1 and 98/6% accuracy for REDD house 2. In a total of 165 samples for REDD house 1, the network did not correctly identify 4 cases, and in a total of 135 samples for REDD house 2, the network did not correctly identify 2 cases.

According to Figs. 13 and 14, it can be seen that the trained network was prosperous in the identification proceeding of the EAs with 97.45% accuracy for REDD house 3 and 97.25% accuracy for REDD house 4. In a total of 195 samples for REDD house 3, the network did not correctly identify 5 cases, and in a total of 180 samples for REDD house 4, the network did not correctly identify 5 cases.

Results show the high effectiveness and precision of the suggested technique in detecting patterns of PC of household EAs. The main objective in solving the NILM problems is the identification and disaggregation of the total home PC at different times. Knowing which appliances were most consumed during peak hours of consumption or other hours of the day can contribute to the excellent planning of energy management. To do this, samples of total home PC were considered as inputs to each saved network. Eighty-five samples (each of which corresponding to two hours) were selected from the PC of the entire house to be disaggregated based on PC patterns of each EA. Table 4 presents the results for each saved network related to the total PC data for load disaggregation of all REDD houses 1, 2, 3, and 4.

Based on the results in Tables 4, it can be concluded that if CNN is properly designed, it will be able to easily identify and classify test data or any new data with high accuracy. Accordingly, the saved networks in this paper were able to easily disaggregate and classify the total home PC samples from each dataset using prior training. The use of transient state signals from the PC of any EA as input, reducing the computational cost of measuring data, and the need for no complicated calculations in the detection operation are the advantages of the proposed method over other traditional and common methods.

Table 4. The results of identifying the PC of entire houses using saved networks

House	Number of Samples	TI	FI	Accuracy (%)
REDD House 1	85	82	3	96.47
REDD House 2	85	80	5	94.11
REDD House 3	85	82	3	96.47
REDD House 4	85	83	2	97.64
General	340	327	13	96.17

6. COMPARISONS OF METHODS

Demonstrating the effectiveness of the proposed method is achieved by comparing the results with other methods used in this regard. The average accuracy obtained for all the EAs of all four houses studied in this paper is compared with the results obtained in other papers. All of the comparison methods have been applied to the REDD dataset. Given that in each of the works done in this regard, division of the database for training, testing and validation operations is not the same, the direct comparison should be done with caution. Table 5 shows the comparison of the results of the different methods with the proposed method in this paper. Using power signals as network inputs, which reduces computational costs, extracting the features of PC data in several steps to identify and classify PC patterns of EAs, and the use of low frequency (1 Hz) signals to implement the proposed method through the use of low-cost meters for all similar real-world data are the advantages of the proposed method over other methods.

Table 5. Comparison of the Accuracy of different methods using the REDD dataset and the number of categorized appliance types

Appliance Identification Method	Remarks	Accuracy
Proposed Method	Using all appliances from REDD houses 1, 2, 3, and 4	96.17%
AANNs (Morais and Castro 2019)	Using 7 appliances selected from the REDD	95.4%
PCA (Moradzadeh et al. 2020b)	Using all appliances from REDD houses 1, 2, and 3	94.68%
PBN (Welikala et al. 2019)	Using all appliances from REDD	85.5%
Basic NILM (Dinesh et al. 2016)	Using all appliances from REDD	79.7%
Bayesian Classifier (Zeifman 2012)	Using 9 appliances selected from the REDD	83%
Viterbi algorithm (Zeifman 2012)	Using 9 appliances selected from the REDD	88.1%
Supervised DT (Liao et al. 2014)	Using 9 appliances selected from the REDD	76.4%
Additive FHMM (Kolter and Jaakkola 2012)	Using 7 appliances selected from the REDD	71.3%
OLDA (Hart et al. 1989)	Using 7 appliances selected from the REDD	84%

7. CONCLUSION

In this paper, an application of Deep Learning based on feature extraction for the effective performance of the NILM was presented. This approach that was able to identify types of residential electrical appliances (EAs) by extracting and identifying the power consumption (PC) patterns is called Convolutional Neural Network (CNN). The extraction of sensible features and recognition of the electrical PC patterns of each EA in the load disaggregation gives consumers the information they need to be aware of their PC in any time period. Low-frequency sampled data from the Reference Energy Data Disaggregation Dataset (REDD) were selected to test the suggested procedure through various performance criteria. PC curves received from each residential EA at diverce times were employed as input data for CNN. By implementing the presented procedure to the input data and training the networks, the PC patterns were saved as a toolbox. Subsequently, these saved networks were used to disaggregate the samples of whole residential PC. By applying the proposed solution, the PC of household EAs was predicted from the total PC of the house at diverse time horizons. Saved networks showed that they were able to disaggregate the total PC with a 96.17% mean accuracy for the used dataset. It should be noted that the suggested solution could be utilized for all measured data in the real world.

REFERENCES

Anthimopoulos M, Christodoulidis S, Ebner L, et al (2016) Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. IEEE Transactions on Medical Imaging 35:1207–1216. https://doi.org/10.1109/TMI.2016.2535865

Bagheri A, Gu IYH, Bollen MHJ, Balouji E (2018) A Robust Transform-Domain Deep Convolutional Network for Voltage Dip Classification. IEEE Transactions on Power Delivery 33:2794–2802. https://doi.org/10.1109/TPWRD.2018.2854677

Basu K, Debusschere V, Bacha S, et al (2015) Nonintrusive load monitoring: A temporal multilabel classification approach. IEEE Transactions on Industrial Informatics 11:262–270. https://doi.org/10.1109/TII.2014.2361288

Beckel C, Sadamori L, Staake T, Santini S (2014) Revealing household characteristics from smart meter data. Energy 78:397–410. https://doi.org/10.1016/j.energy.2014.10.025

Bhotto MZA, Makonin S, Bajić I V. (2017) Load Disaggregation Based on Aided Linear Integer Programming. IEEE Transactions on Circuits and Systems II: Express Briefs 64:792–796. https://doi.org/10.1109/TCSII.2016.2603479

D'Incecco M, Squartini S, Zhong M (2020) Transfer Learning for Non-Intrusive Load Monitoring. IEEE Transactions on Smart Grid 11:1419–1429. https://doi.org/10.1109/TSG.2019.2938068

Devlin MA, Hayes BP (2019) Non-Intrusive Load Monitoring and Classification of Activities of Daily Living Using Residential Smart Meter Data. IEEE Transactions on Consumer Electronics 65:339–348. https://doi.org/10.1109/TCE.2019.2918922

Dinesh C, Nettasinghe BW, Godaliyadda RI, et al (2016) Residential Appliance Identification Based on Spectral Information of Low Frequency Smart Meter Measurements. IEEE Transactions on Smart Grid 7:2781–2792. https://doi.org/10.1109/TSG.2015.2484258

Dong C, Loy CC, He K, Tang X (2016) Image Super-Resolution Using Deep Convolutional Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence 38:295–307. https://doi.org/10.1109/TPAMI.2015.2439281

Egarter D, Bhuvana VP, Elmenreich W (2015) PALDi: Online load disaggregation via particle filtering. IEEE Transactions on Instrumentation and Measurement 64:467–477. https://doi.org/10.1109/TIM.2014.2344373

Egarter D, Monacchi A, Khatib T, Elmenreich W (2016) Integration of legacy appliances into home energy management systems. Journal of Ambient Intelligence and Humanized Computing 7:171–185. https://doi.org/10.1007/s12652-015-0312-9

Ferrández-Pastor FJ, Mora-Mora H, Sánchez-Romero JL, et al (2017) Interpreting human activity from electrical consumption data using reconfigurable hardware and hidden Markov models. Journal of Ambient Intelligence and Humanized Computing 8:469–483. https://doi.org/10.1007/s12652-016-0431-y

Gajowniczek K, Zabkowski T (2015) Data mining techniques for detecting household characteristics based on smart meter data. Energies 8:7407–7427. https://doi.org/10.3390/en8077407

Gaur M, Majumdar A (2018) Disaggregating transform learning for non-intrusive load monitoring. IEEE Access 6:46256–46265. https://doi.org/10.1109/ACCESS.2018.2850707

Gonçalves H, Ocneanu A, Bergés M, Fan RH (2011) Unsupervised disaggregation of appliances using aggregated consumption data. Environmental Engineering

Han T, Liu C, Yang W, Jiang D (2019) A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults. Knowledge-Based Systems 165:474–487. https://doi.org/10.1016/j.knosys.2018.12.019

Hart GW (1992) Nonintrusive Appliance Load Monitoring. Proceedings of the IEEE 80:1870–1891. https://doi.org/10.1109/5.192069

Hart GW, Kern EJC, Schweppe FC (1989) Non-intrusive appliance monitor apparatus. Google Patents 13

He K, Stankovic L, Liao J, Stankovic V (2018) Non-Intrusive Load Disaggregation Using Graph Signal Processing.

IEEE Transactions on Smart Grid 9:1739–1747. https://doi.org/10.1109/TSG.2016.2598872

Hosseini S, Henao N, Kelouwani S, et al (2019) A Study on Markovian and Deep Learning Based Architectures for Household Appliance-level Load Modeling and Recognition. In: 2019 IEEE 28th International Symposium on Industrial Electronics (ISIE). IEEE, pp 35–40

Kelly J, Knottenbelt W (2012) Disaggregating Multi-State Appliances from Smart Meter Data. In: Imperial College Energy and Performance Colloquium

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

Kim H, Marwah M, Arlitt M, et al (2011) Unsupervised disaggregation of low frequency power measurements. In: Proceedings of the 11th SIAM International Conference on Data Mining, SDM 2011. pp 747–758

Klemenjak C, Goldsborough P (2016) Non-intrusive load monitoring: A review and outlook. In: Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft fur Informatik (GI). pp 2199–2210

Kolter JZ, Jaakkola T (2012) Approximate inference in additive factorial HMMs with application to energy disaggregation. In: Journal of Machine Learning Research. pp 1472–1482

Kolter JZ, Johnson MJ (2011) REDD : A Public Data Set for Energy Disaggregation Research. SustKDD workshop xxxxx:1–6

Kong S, Kim Y, Ko R, Joo SK (2015) Home appliance load disaggregation using cepstrum-smoothing-based method. IEEE Transactions on Consumer Electronics 61:24–30. https://doi.org/10.1109/TCE.2015.7064107

Kong W, Dong ZY, Wang B, et al (2020) A practical solution for non-intrusive type II load monitoring based on deep learning and post-processing. IEEE Transactions on Smart Grid 11:148–160. https://doi.org/10.1109/TSG.2019.2918330

Lecun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444

Liao J, Elafoudi G, Stankovic L, Stankovic V (2014) Power Disaggregation for Low-sampling Rate Data. NILM Work 2014

Liu Q, Kamoto KM, Liu X, et al (2019) Low-Complexity Non-Intrusive Load Monitoring Using Unsupervised Learning and Generalized Appliance Models. IEEE Transactions on Consumer Electronics 65:28–37. https://doi.org/10.1109/TCE.2019.2891160

Mauch L, Yang B (2016) A new approach for supervised power disaggregation by using a deep recurrent LSTM network. In: 2015 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2015. pp 63–67

Mengistu MA, Girmay AA, Camarda C, et al (2019) A Cloud-Based On-Line Disaggregation Algorithm for Home Appliance Loads. IEEE Transactions on Smart Grid 10:3430–3439. https://doi.org/10.1109/TSG.2018.2826844

Moradzadeh A, Mansour-Saatloo A, Mohammadi-Ivatloo B, Anvari-Moghaddam A (2020a) Performance evaluation of two machine learning techniques in heating and cooling loads forecasting of residential buildings. Applied Sciences (Switzerland) 10:3829. https://doi.org/10.3390/app10113829

Moradzadeh A, Pourhossein K (2019a) Location of Disk Space Variations in Transformer Winding using Convolutional Neural Networks. In: 2019 54th International Universities Power Engineering Conference, UPEC 2019 - Proceedings. IEEE, pp 1–5

Moradzadeh A, Pourhossein K (2019b) Short Circuit Location in Transformer Winding Using Deep Learning of Its Frequency Responses. In: Proceedings 2019 International Aegean Conference on Electrical Machines and Power Electronics, ACEMP 2019 and 2019 International Conference on Optimization of Electrical and Electronic Equipment, OPTIM 2019. IEEE, pp 268–273

Moradzadeh A, Sadeghian O, Pourhossein K, et al (2020b) Improving residential load disaggregation for sustainable development of energy via principal component analysis. Sustainability (Switzerland) 12:3158. https://doi.org/10.3390/SU12083158

Moradzadeh A, Zakeri S, Shoaran M, et al (2020c) Short-term load forecasting of microgrid via hybrid support vector regression and long short-term memory algorithms. Sustainability (Switzerland) 12:7076. https://doi.org/10.3390/su12177076

Moradzadeh A, Zeinal-Kheiri S, Mohammadi-Ivatloo B, et al (2020d) Support Vector Machine-Assisted Improvement Residential Load Disaggregation. In: 2020 28th Iranian Conference on Electrical Engineering (icee). pp 1–6

Morais LR, Castro ARG (2019) Competitive Autoassociative Neural Networks for Electrical Appliance Identification for Non-Intrusive Load Monitoring. IEEE Access 7:111746–111755. https://doi.org/10.1109/access.2019.2934019

Parson O, Ghosh S, Weal M, Rogers A (2014) An unsupervised training method for non-intrusive appliance load monitoring. Artificial Intelligence 217:1–19. https://doi.org/10.1016/j.artint.2014.07.010

Patterson J, Gibson A (2017) Deep learning: A practitioner's approach

Peng X, Yang F, Wang G, et al (2019) A Convolutional neural network-based deep learning methodology for recognition of partial discharge patterns from high-voltage cables. IEEE Transactions on Power Delivery 34:1460–1469. https://doi.org/10.1109/TPWRD.2019.2906086

Piga D, Cominola A, Giuliani M, et al (2016) Sparse Optimization for Automated Energy End Use Disaggregation. IEEE Transactions on Control Systems Technology 24:1044–1051. https://doi.org/10.1109/TCST.2015.2476777

Quek YT, Woo WL, Logenthiran T (2020) Load Disaggregation Using One-Directional Convolutional Stacked Long Short-Term Memory Recurrent Neural Network. IEEE Systems Journal 14:1395–1404. https://doi.org/10.1109/JSYST.2019.2919668

Roy SS, Samui P, Nagtode I, et al (2020) Forecasting heating and cooling loads of buildings: a comparative performance analysis. Journal of Ambient Intelligence and Humanized Computing 11:1253–1264. https://doi.org/10.1007/s12652-019-01317-y

Singhal V, Maggu J, Majumdar A (2019) Simultaneous Detection of Multiple Appliances from Smart-Meter Measurements via Multi-Label Consistent Deep Dictionary Learning and Deep Transform Learning. IEEE Transactions on Smart Grid 10:2969–2978. https://doi.org/10.1109/TSG.2018.2815763

Sirojan T, Phung BT, Ambikairajah E (2018) Deep Neural Network Based Energy Disaggregation. In: 2018 6th IEEE International Conference on Smart Energy Grid Engineering, SEGE 2018. pp 73–77

Tabatabaei SM, Dick S, Xu W (2017) Toward Non-Intrusive Load Monitoring via Multi-Label Classification. IEEE Transactions on Smart Grid 8:26–40. https://doi.org/10.1109/TSG.2016.2584581

Wang Z, Zheng G (2012) Residential appliances identification and monitoring by a nonintrusive method. IEEE Transactions on Smart Grid 3:80–92. https://doi.org/10.1109/TSG.2011.2163950

Welikala S, Dinesh C, Ekanayake MPB, et al (2019) Incorporating Appliance Usage Patterns for Non-Intrusive Load

Monitoring and Load Forecasting. IEEE Transactions on Smart Grid 10:448–461. https://doi.org/10.1109/TSG.2017.2743760

Wittmann FM, Lopez JC, Rider MJ (2018) Nonintrusive Load Monitoring Algorithm Using Mixed-Integer Linear Programming. IEEE Transactions on Consumer Electronics 64:180–187. https://doi.org/10.1109/TCE.2018.2843292

Zeifman M (2012) Disaggregation of home energy display data using probabilistic approach. IEEE Transactions on Consumer Electronics 58:23–31. https://doi.org/10.1109/TCE.2012.6170051

Zeinal-Kheiri S, Shotorbani AM, Mohammadi-Ivatloo B (2020) Residential Load Disaggregation Considering State Transitions. IEEE Transactions on Industrial Informatics 16:743–753. https://doi.org/10.1109/TII.2019.2925323

Zhang C, Zhong M, Wang Z, et al (2018) Sequence-to-point learning with neural networks for non-intrusive load monitoring. In: 32nd AAAI Conference on Artificial Intelligence, AAAI 2018. pp 2604–2611