Home Load Disaggregation using Deep Learning and Bayesian Optimization: A Case Study in Arctic Climate in Northern Norway

1st Nasrin Kianpoor
Department of Electrical Engineering
UiT – The Arctic University of Norway
Narvik, Norway
Nasrin.Kianpoor@uit.no

2nd Bjarte Hoff
Department of Electrical Engineering
UiT – The Arctic University of Norway
Narvik, Norway
Bjarte.Hoff@uit.no

3rd Trond Østrem
Department of Electrical Engineering
UiT – The Arctic University of Norway
Narvik, Norway
Trond.Ostrem@uit.no

***Abstract*—Load monitoring is an essential task in energy management systems. In this paper, an approach that relies on a long short-term memory (LSTM) model and a discrete wavelet transform (DWT) filter i s p resented t o e stimate the energy usage of flexible a ppliances. I n t he p reprocessing stage, the main features of the aggregated power signal are extracted using DWT. Deep learning methods are very sensitive to hyperparameters, and choosing optimal values can significantly improve the accuracy of the model. To optimize the performance of the LSTM model, a Bayesian optimization algorithm is used to find the optimal set of hyperparameters. The performance of the proposed approach is evaluated using real-world data collected from a residential building in northern Norway. The results show that the proposed methodology can accurately disaggregate the power consumption of different appliances, with higher accuracy compared to existing methods.**

***Index Terms*—Energy disaggregation, Bayesian optimization, deep learning, Non-intrusive load monitoring, Signal processing.**

I. Introduction

The residential sector in Europe significantly contributes to energy consumption, with buildings accounting for about 40% of the total energy use [1]. Similarly, in Norway, the household sector has a high share of the country's total energy consumption. According to the statistics provided by "Energy Facts Norway," the total energy use in households was 47.6 TWh in 2017, which was 22% of the final energy consumption. It was the third-largest energy-consuming sector after the industry and transport sectors. This statistic also shows that electricity is the most widely used type of energy in the household sector, accounting for 83% of all energy types, including electricity, biofuel, and district heating. The increasing use of electrical devices and using electricity for heating space and hot water are the reasons for the high share of electricity in the energy mix [2].

Considering the significant impact of the household sector on energy consumption, it is important to examine energy management strategies, such as load monitoring, to help households reduce their energy use and promote sustainability [3]. There are two main methods for monitoring energy consumption in residential and commercial buildings: Intrusive Load Monitoring (ILM) and Non-Intrusive Load Monitoring (NILM) [4]. In the ILM method, the meters and sensors are directly installed on each appliance or device; therefore, the ILM method provides detailed and accurate data on the usage of individual appliances. But the ILM method is expensive and time-consuming regarding the installation and maintenance of all the meters. This is why the NILM concept was proposed for the first time in the 1980s [5]. In this method, the total electricity signal consumption is analyzed to identify and estimate the usage pattern on each appliance. NILM is less expensive and easier to implement than ILM, but it may be less accurate in monitoring the energy usage of individual appliances.

Different methods have been used for NILM, such as machine learning, signal processing, optimization, and pattern recognition techniques. In the signal processing methods, Fourier transforms, wavelet transforms and time-frequency analysis are used to analyze the power consumption signal in the frequency domain. A NILM algorithm based on an improved time-frequency analysis is presented in [6]. In optimization-based methods, the problem of load disaggregation is considered an optimization problem [7]. In a study by Zoha et al. [4], a survey on NILM methods for disaggregated sensing is conducted, with a comprehensive overview of NILM structure and a review of the state-of-the-art algorithms. Another study by Hosseini et al. [8] presents a review of the NILM in the application of home energy management systems. A more recent review study of methods, challenges, and perspectives for NILM is presented by Kasemli et al. [9], where the paper provides a literature review of NILM algorithms for residential appliances.

This work was supported partly by the Project "Transformation to a Renewable and Smart Rural Power System Community (RENEW)" under Grant 310026, and partly by the Arctic Centre for Sustainable Energy (ARC), UiT-The Arctic University of Norway.

During the last few years, machine learning methods, among others, support vector machines [10], decision trees [11], and deep learning has become popular in NILM. It is because of their ability to learn the features of appliance signatures. Widespread installation of smart meters in recent years makes it possible to have access to electricity consumption data of common appliances used in both residential and commercial buildings. Several recently proposed methods for the NILM problem are based on deep learning methods such as Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) [12]. A CNN-based algorithm in which the inputs and outputs are formed as data sequences are utilized for energy disaggregation in [13]. A deep learning approach based on multi-layer, feed-forward neural networks is presented in [14] to identify the common household appliances based on the total power flow. Compared to other deep learning techniques, LSTM networks often outperform because they can handle lags of unknown duration between significant events in a time series [15]. This unique feature of LSTMs gives them a competitive advantage over other deep learning methods in identifying changes in power consumption. The authors in [16] present a method based on the LSTM-RNN algorithm to address the problem of energy disaggregation, and this algorithm can be a robust solution. In another research, a combination of an adaptive ensemble filtering method and an LSTM architecture is employed to extract the power consumption of the sizable appliances from the aggregated power signal [17].

While the problem of NILM using various deep learning methods has been researched to some extent, in many researches, the effect of data preprocessing techniques on the proposed models has not been considered. But in this research, a discrete wavelet transforms filter is used to extract the features of the main signal to improve the performance of the LSTM model. Most of the research that has been conducted on NILM has treated it as a classification problem, focusing on identifying on/off events. However, in this study, the problem is being approached as a regression problem, with the goal of disaggregating the load signature of appliances from the main signal. The main contributions of this article are outlined as follows

1) A methodology for load disaggregation is proposed based on the LSTM network to identify the load profile of major appliances.
2) Discrete wavelet transform is utilized for the preprocessing stage to improve the performance of the LSTM network
3) A Bayesian optimization algorithm is implemented to find the optimal set of hyperparameters for training and testing the LSTM network.
4) A new dataset in the arctic climate of northern Norway is measured and prepared to be used for evaluating the proposed methodology.

Following the introduction section, section II briefly discusses the use case and the data that is used for the evaluation of the proposed method. Section III presents the methodology and its different stages. Experiment results are demonstrated in Section IV. Finally, Section V concludes this article.

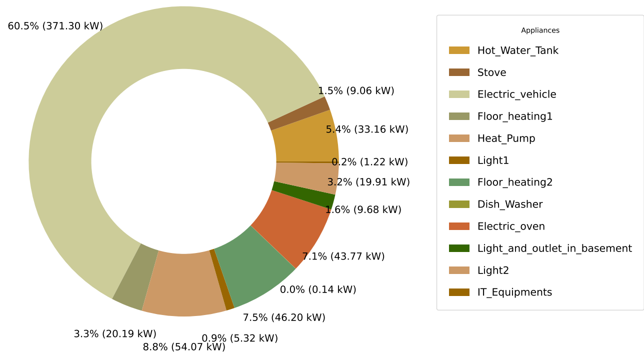

Fig. 1. The total power usage of each appliance for the entire measured period (26-09-2021 to 19-10-2021)

II. USE CASE

There are some publicly available datasets that can be used to evaluate the proposed NILM algorithms in different studies. Some of them have been widely used in different research, such as REDD [18], REFIT [19], AMPDs [20] datasets. But to the best of the author's knowledge, there is not a dataset in the arctic climate for the application of NILM. For this reason, the authors initiated data collection in the arctic climate of northern Norway, for use in the application of energy disaggregation. The data which is used in this paper are measured in a residential house located in Narvik, Norway. The data are measured from 26-09-2021 to 19-10-2021. Measurements are performed using Schneider Electric Power Tags on every outgoing circuit breaker, where the total power consumption is measured on the main circuit breaker. A Raspberry Pi is used to collect and store the measurement data in an InfluxDB database through MODBUS TCP. The data includes the measurement of the total power usage of the house, voltage, and current for different appliances including a charging station of a plug-in hybrid electric vehicle (EV), hot water tank (HWT), stove, electric oven, and heat pump. Measurements of load current, power factor, and voltage are sampled every 10 minutes. In some of the measurements, a sensor is installed for an aggregate load. For example, a sensor is used to measure the aggregate electricity consumption of light and outlet of the washing room which includes a washing machine and floor heating.

The total power usage of each appliance for the entire measured period (26-09-2021 to 19-10-2021) is calculated and shown in Fig. 1. Based on Fig. 1, EV has consumed the largest part of the total power which is 60.5 % or 371.3 (kW), therefore it is a suitable load to be considered as a flexible load, it can be scheduled to be charged during the off-peak period when the electricity demand is low.

III. METHODOLOGY

In this paper, an approach is proposed to address the problem of residential load desegregation on a new dataset collected in the arctic climate of northern Norway. By accurately disaggregating the power consumption of different appliances, the overall energy consumption of buildings can be reduced. The suggested solution can be extended to other regions and can provide a valuable tool for building energy management systems. A combination of a discrete wavelet transform (DWT) filter and long short-term memory (LSTM) model is utilized to estimate the energy usage of flexible appliances from the aggregated power. The methodology has different stages, first stage is data collection which is explained in II. Then in the preprocessing stage, in addition to data cleaning and normalization, a DWT filter is applied to the total power signal to extract the main features of the signal. Deep learning methods are very sensitive to hyperparameters, therefore a Bayesian optimization algorithm is implemented to find the optimal set of hyperparameters for the LSTM network. The detail of all the above-mentioned stages and the structure of an LSTM network is explained in the following subsections.

A. Pre-processing

1) Data normalization and cleaning: Data cleaning is the first step before training any machine learning algorithm. This can improve data quality and consistency. Data cleaning includes various steps such as detecting outliers data, interpolating or imputing missing values, finding missing values, standardization, or normalization. In the data set used in this research, some missing values were found, which are replaced by the median of the data. The preprocessing stage also involves normalizing the data, which helps to bring all the features of the data to a similar scale. This can improve the performance of the LSTM network, as it can reduce the impact of features with larger values dominating the model's predictions. The data are normalized as follows

$$x_{new} = \frac{x - \bar{x}}{\sigma} \tag{1}$$

where x_{new}, $\bar{x}$, x, and σ are the normalized value, mean of real values, real value, and the standard deviation of the true values, respectively.

2) Signal denoising: One of the important steps in the preprocessing stage is signal denoising, where unwanted noises are removed from the signal while the important features of the signal are preserved. Therefore, removing noise from the time series can lead to an increase in modeling accuracy. There are different methods for signal denoising such as low-pass filtering, moving average, wavelet denoising, and principle component analysis (PCA). In this paper, the discrete wavelet transform (DWT) method is chosen to remove the noise of the time series. DWT is a powerful tool for signal smoothing, with features that make it superior to other methods. For example, specifying a threshold value allows the identification and removal of noise at certain frequencies, while preserving the important characteristics of the signal [21].

B. Bayesian optimization

The performance of deep learning models is highly dependent on their hyperparameters. A set of configuration parameters used during the training and testing of a deep learning model are called hyperparameters. Therefore, it is very crucial to select the optimal hyperparameters to train a deep learning model. For example, in an LSTM model, the number of hidden layers, batch size, validation split, etc. are considered hyperparameters. There is not one set of hyperparameters that is suitable for all the models, it varies from model to model according to the problem statement [22]. There are different methods for hyperparameter tuning such as random search, grid search, genetic algorithm, and Bayesian optimization. One of the most common hyperparameter tunings is hand tuning which is based on trial and error, finding the optimal set of hyperparameters using trial and error is a difficult and time-consuming method. The other method is grid search which searches between all the defined boundaries to find the optimum solution, it is not scalable for higher dimensions, and it takes a long time to search among all the searching ranges. Bayesian optimization has shown a fast and acceptable performance in the literature. The algorithm of this method does not check all the points within the defined boundaries one by one, rather a probabilistic model is used to suggest the next set of parameters that are close to the region that is likely to contain the optimal solution. In other words, it does not check the outlier points [23]. Therefore it takes a shorter time to use Bayesian optimization rather than grid search or random search methods. Bayesian optimization is an optimization method that is commonly used for hyperparameter tuning in machine learning. In this method, the objective function, $f(x)$ is a black box or unknown function. The basic idea of Bayesian optimization is to model the black-box function using a probabilistic model, which is usually considered a Gaussian function. In the procedure of Bayesian optimization, posterior information of $f(x)$ is updated using the prior information, to find the global optimal point of the function. More details about Bayesian optimization can be found in [24].

C. The LSTM model

An LSTM model is an extended version of a Recurrent neural network that is widely used for time-series predictions. An LSTM unit is comprised of several gates and a cell state. In Fig.2 the LSTM unit structure is presented. The cell state works as a memory for the LSTM unit. The gates can add or remove information from the cell state. They are different neural networks that are trained to decide which information should be deleted or kept in the cell state. An LSTM unit has three gates including an input gate, a forget gate, and an output gate. The sigmoid activation function (σ) is used for each gate to constrain values between 0 and 1 while the $tanh$ activation function is used for cell state and input gate to compress the input values to a range between -1 and 1. The forget gate removes information that is less important or no longer needed by the LSTM unit. The output of the forget gate is a vector with values between 0 and 1, where 0 implies

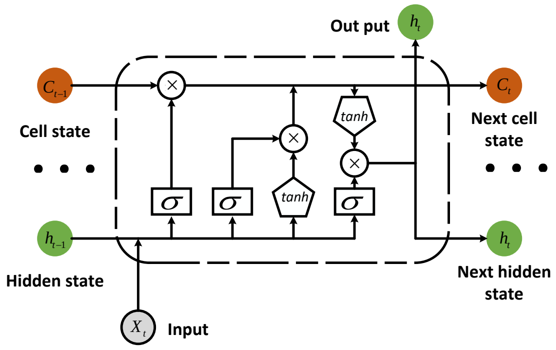

Fig. 2. The structure of the LSTM unit.

TABLE I
THE SEARCH BOUNDARIES FOR HYPERPARAMETERS OPTIMIZATION

Hyperparameters	Initial value	Final value	Type
Window_size	6	360	int
Number of Unit 1	3	150	int
Number of Unit 2	3	150	int
Drop_out 1	0	0.5	float
Drop_out 2	0	0.5	float
Batch_size	16	1024	int
Learning_rate	0.0001	0.1	float
Decay	1e-6	1e-2	float
L2_reg	1e-7	1e-4	float
Validation_split	0.1	0.3	float

deleting information and 1 means keeping information. The task of the input gate is to add new information to the cell state. The output gate takes useful information from the input and from the state cell. It sends them as output for the current cell and as hidden state for the next cell. More information about the architecture and equations for the LSTM cell can be found in [15].

IV. Experimental results

A. Experimental setup

Electricity usage of different appliances in a case study in northern Norway is utilized for evaluating the proposed model for load disaggregation application. More detail about the dataset is explained in section II. The proposed method is implemented in “Google Colab” using Pandas, Numpy, TensorFlow, Keras, and scikit-learn libraries.

The proposed approach is implemented to disaggregate the electricity consumption of four different appliances which have a high share of total power including hybrid electric vehicle (EV), hot water tank (HWT), electric heater (EH), and floor heating/washing machine/light (FH/WM/Lght) (the sensor measured the aggregate power consumption of light and outlet of washing room which includes washing machine and floor heating).

Results are evaluated based on three different metrics, the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Normalised Mean Square Error (NRMSE) which are commonly used to evaluate the energy disaggregation problem. NRMSE is the ratio between RMSE and the range of the true value ($y_{\max} - y_{\min}$), its value should be between 0 and 1, where smaller values indicate better performance of the model. The mathematical relations of the metrics are as follows

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2}$$

$$RMSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{3}$$

$$NRMSE = \frac{RMSE}{y_{\max} - y_{\min}} \tag{4}$$

where y_i, $\hat{y}_i$, $y_{\max}$, and $y_{\min}$ in the above formulas are true value, estimated value, maximum true value, and the minimum true value, respectively.

In the following subsections, the details of the experimental setup for different steps are presented separately.

1) Denoising the data using discrete wavelet transform: In this stage, the total power is smoothed out using the DWT method described in III-A2, and afterward, the extracted feature is used as one of the inputs of the LSTM network to train the model. An open-source wavelet transforms software for Python called ”PyWavelets” [25] is used for filtering out the aggregated power signal using the DWT method.

2) Bayesian Optimization: In this article, Bayesian optimization is utilized for finding the optimal parameters for training and testing the LSTM model. For this reason, a Python package for Bayesian optimization is used to find the optimal solution, more detail about the package can be found in [26]. The Bayesian algorithm estimates a set of LSTM network parameters, which include window size, number of units, dropout value, batch size, validation split, learning rate, decay, and kernel regularizer. The search range for different parameters is summarized in Table I.

The data are divided into three categories, training, validation, and testing. The amount of data for validation is one of the parameters that the Bayesian algorithm finds in the range of 0.1 to 0.3 percent of the data. The amount of data for testing varies based on the type of appliances. For example, one day of data is considered to test the model for the electric heater and hot water tank, but for electric vehicles and FH/WM/Lght, approximately two and a half days of data is used to test the model to include more samples of the device in the operating mode. The rest of the data is used to train the model.

The features and characteristics of different appliances are not the same, that is why the electricity consumption pattern of household appliances is different. It depends on factors such as power rating, usage pattern, and efficiency. Therefore, a set of optimal hyperparameters can not be used for all the appliances. For this reason, the algorithm of Bayesian optimization is implemented for all the case studies separately. The results show different configurations for different appliances. The optimal solution for LSTM network configuration for different appliances using the Bayesian algorithm is presented in Table

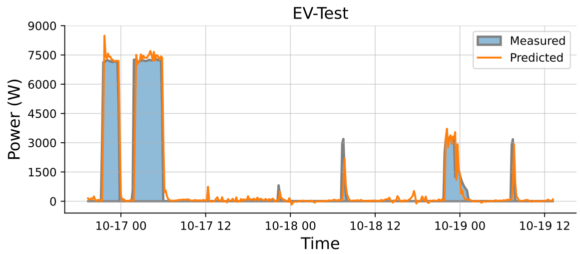

Fig. 3. Comparison of ground truth data with the proposed method for the Electric vehicle.

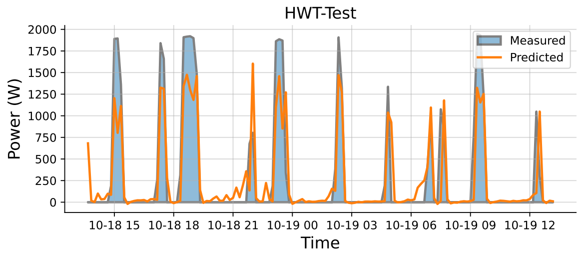

Fig. 4. Comparison of the proposed method with ground truth for the hot water tank.

II.

B. *Results and discussion*

Based on the configuration mentioned in Table II, the proposed algorithm is implemented for the above-mentioned appliances. The signature identification of appliances are shown in Figs. 3 - 6. These figures depict the performance of the proposed approach compared to ground truth data. As observed, the proposed method can estimate complicated energy patterns.

To demonstrate the effectiveness of the proposed method compared to other techniques, it is compared with widely used state-of-the-art approaches in the application of NILM, such as linear regression, decision tree regression, and also with a standalone LSTM model. The results for DWT-LSTM and its comparison with other methods based on different metrics (MAE, RMSE, NRMSE) are presented in Table III and Fig. 7. Based on the results summarized in Table III, the stand-alone LSTM model performs better than LR and DTR methods in estimating the power consumption of the considered case studies. The superiority of the LSTM model to compared ones can be because of its ability to capture long-term dependencies in time series data, it can recognize patterns in the power consumption data of specific appliances. The DWT-LSTM outperforms the stand-alone LSTM model. It means that filtering the main signal before feeding it to the LSTM model leads to improving the performance of the model by removing noise and extracting relevant features.

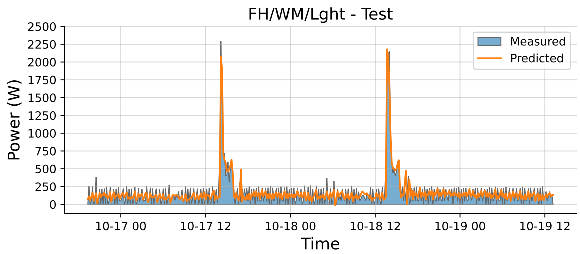

Fig. 5. Comparison of the proposed method with ground truth for Floor heating cable/Washing Machine /Light.

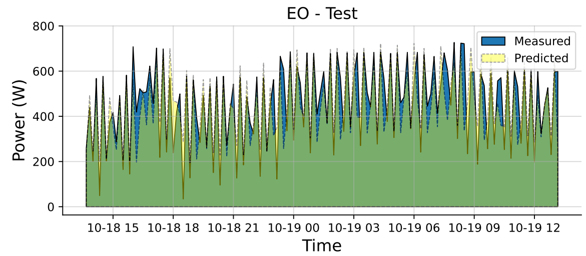

Fig. 6. Comparison of ground truth data with the proposed method for Electric heater.

V. Conclusion

The problem of residential load disaggregation considering a case study in the arctic climate of northern Norway is addressed in this paper. A combination of a discrete wavelet transform filter with an LSTM model is used to disaggregate the signature of four different appliances from the total power. For the optimal performance of the LSTM model, a Bayesian optimization algorithm is implemented to find the best set of hyperparameters for each appliance. Experimental results on the real-world data show the superiority of the proposed approach compared to other techniques.

References

[1] European commission, “In focus: Energy efficiency in buildings,” online, accessed 3/13/2023, Feb 2020. [Online]. Available: https://commission.europa.eu/news/focus-energy-efficiency-buildings-2020-02-17_en/

[2] Energy facts Norway, “ENERGY USE BY SECTOR,” online, accessed 3/13/2023. [Online]. Available: https://energifaktanorge.no/en/norsk-energibruk/energibruken-i-ulike-sektorer/

[3] N. Kianpoor, B. Hoff, and T. Østrem, “Load modeling from smart meter data using neural network methods,” in *2021 22nd IEEE International Conference on Industrial Technology (ICIT)*, vol. 1. IEEE, 2021, pp. 611–616.

[4] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, “Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey,” *Sensors*, vol. 12, no. 12, pp. 16 838–16 866, 2012.

[5] G. W. Hart, “Nonintrusive appliance load monitoring,” *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.

[6] Y.-H. Lin and M.-S. Tsai, “Development of an improved time–frequency analysis-based nonintrusive load monitor for load demand identification,” *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 6, pp. 1470–1483, 2013.

[7] E. Azizi, A. M. Shotorbani, M.-T. Hamidi-Beheshti, B. Mohammadi-Ivatloo, and S. Bolouki, “Residential household non-intrusive load monitoring via smart event-based optimization,” *IEEE Transactions on Consumer Electronics*, vol. 66, no. 3, pp. 233–241, 2020.

[8] S. S. Hosseini, K. Agbossou, S. Kelouwani, and A. Cardenas, “Non-intrusive load monitoring through home energy management systems: A comprehensive review,” *Renewable and Sustainable Energy Reviews*, vol. 79, pp. 1266–1274, 2017.

TABLE II
THE OPTIMAL LSTM NETWORK CONFIGURATION FOR DIFFERENT APPLIANCES USING BAYESIAN ALGORITHM

Appliances	**LSTM Hyperparameters**									Window_size
	Unit 1	Unit 2	Dropout 1	Dropout 2	Batch Size	Validation_split	Learning_rate	Decay	Kernel_regularizer	
Electric Vehicle	12	67	0.04	0.07	910	0.14	0.05	0.008	2.5e-06	46
Hot Water Tank	12	125	0.04	0.43	383	0.11	0.05	0.004	6.7e-05	258
FH/WM/Lght	72	57	0.45	0.18	123	0.15	0.02	0.004	8.2e-05	316
Electric Heater	13	7	0.43	0.09	48	0.17	0.04	0.006	1.1e-05	77

TABLE III
PERFORMANCE EVALUATION OF THE PROPOSED ALGORITHM COMPARED TO OTHER TECHNIQUES USING DIFFERENT METRICS.

	EV			HWT			FH/WM/Lght			EH		
	MAE (W)	RMSE (W)	NRMSE	MAE (W)	RMSE (W)	NRMSE	MAE (W)	RMSE (W)	NRMSE	MAE (W)	RMSE (W)	NRMSE
DWT-LSTM	211.60	631.80	0.09	162.04	312.82	0.16	74.48	98.99	0.04	90.10	119.60	0.17
LSTM	238.73	643.77	0.09	181.40	332.80	0.17	82.04	130.25	0.05	104.26	134.75	0.19
DTR	320.44	758.9	0.11	492.6	823.6	0.42	171.28	328.03	0.14	213.72	258.46	0.37
LR	679.87	933.72	0.12	402.05	539.77	0.27	131.37	252	0.1	199.78	233.17	0.33

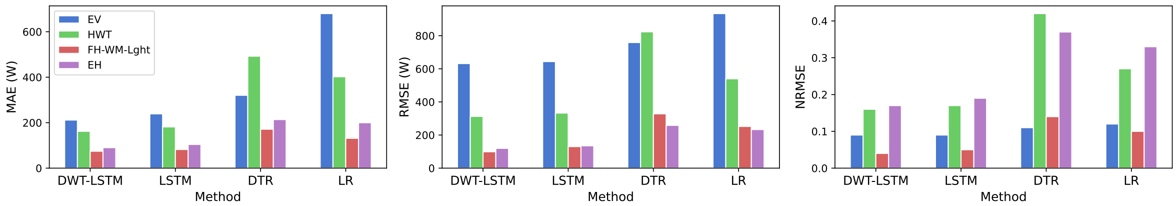

Fig. 7. Performance evaluation of the proposed algorithm compared to other techniques using different metrics.

[9] M. Kaselimi, E. Protopapadakis, A. Voulodimos, N. Doulamis, and A. Doulamis, “Towards trustworthy energy disaggregation: A review of challenges, methods, and perspectives for non-intrusive load monitoring,” *Sensors*, vol. 22, no. 15, p. 5872, 2022.

[10] H. Altrabalsi, V. Stankovic, J. Liao, and L. Stankovic, “Low-complexity energy disaggregation using appliance load modelling,” *Aims Energy*, vol. 4, no. 1, pp. 884–905, 2016.

[11] J. Liao, G. Elafoudi, L. Stankovic, and V. Stankovic, “Non-intrusive appliance load monitoring using low-resolution smart meter data,” in *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2014, pp. 535–540.

[12] H. Çimen, N. Çetinkaya, J. C. Vasquez, and J. M. Guerrero, “A microgrid energy management system based on non-intrusive load monitoring via multitask learning,” *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 977–987, 2020.

[13] M. Kaselimi, E. Protopapadakis, A. Voulodimos, N. Doulamis, and A. Doulamis, “Multi-channel recurrent convolutional neural networks for energy disaggregation,” *IEEE Access*, vol. 7, pp. 81 047–81 056, 2019.

[14] M. A. Devlin and B. P. Hayes, “Non-intrusive load monitoring and classification of activities of daily living using residential smart meter data,” *IEEE transactions on consumer electronics*, vol. 65, no. 3, pp. 339–348, 2019.

[15] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[16] J. Kim, T.-T.-H. Le, H. Kim *et al.*, “Nonintrusive load monitoring based on advanced deep learning and novel signature,” *Computational intelligence and neuroscience*, vol. 2017, 2017.

[17] N. Kianpoor, B. Hoff, and T. Østrem, “Deep adaptive ensemble filter for non-intrusive residential load monitoring,” *Sensors*, vol. 23, no. 4, p. 1992, 2023.

[18] J. Z. Kolter and M. J. Johnson, “Redd: A public data set for energy disaggregation research,” in *Workshop on data mining applications in sustainability (SIGKDD), San Diego, CA*, vol. 25, no. Citeseer. Citeseer, 2011, pp. 59–62.

[19] D. Murray, L. Stankovic, and V. Stankovic, “An electrical load measurements dataset of united kingdom households from a two-year longitudinal study,” *Scientific data*, vol. 4, no. 1, pp. 1–12, 2017.

[20] S. Makonin, B. Ellert, I. V. Bajić, and F. Popowich, “Electricity, water, and natural gas consumption of a residential house in canada from 2012 to 2014,” *Scientific data*, vol. 3, no. 1, pp. 1–12, 2016.

[21] S. Mallat, *A wavelet tour of signal processing*. Elsevier, 1999.

[22] A. H. Victoria and G. Maragatham, “Automatic tuning of hyperparameters using bayesian optimization,” *Evolving Systems*, vol. 12, pp. 217–223, 2021.

[23] J. Snoek, H. Larochelle, and R. P. Adams, “Practical bayesian optimization of machine learning algorithms,” *Advances in neural information processing systems*, vol. 25, 2012.

[24] P. I. Frazier, “A tutorial on bayesian optimization,” *arXiv preprint arXiv:1807.02811*, 2018.

[25] G. Lee, R. Gommers, F. Waselewski, K. Wohlfahrt, and A. O’Leary, “Pywavelets: A python package for wavelet analysis,” *Journal of Open Source Software*, vol. 4, no. 36, p. 1237, 2019.

[26] F. Nogueira, “Bayesian Optimization: Open source constrained global optimization tool for Python,” 2014–. [Online]. Available: https://github.com/fmfn/BayesianOptimization