Non-Intrusive Load Monitoring for Multi-objects in Smart Building

Dandan Li[1], Jiangfeng Li[2], Xin Zeng[3], Vladimir Stankovic[2], Lina Stankovic[2], and Qingjiang Shi[1,4]
[1]School of Software Engineering, Tongji University, Shanghai, China
[2]Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow Scotland, G1 1XW
[3]College of Electronic and Information Engineering, Tongji University, Shanghai, China
[4]Shenzhen Research Institute of Big Data, Shenzhen, China

***Abstract*—The rapidly expansion of Internet of Things (IoT) has ignited renewed interest in energy disaggregation via non-intrusive load monitoring (NILM). Compared to the more frequent NILM approach of training one model for each appliance, this paper proposes a multi-label learning approach based on the widely cited sequence2point convolutional neural network (CNN). Using the smart meter readings collected in an office building, we demonstrate the accuracy and practicality of the proposed network compared to start-of-the-art one-to-one NILM models.**

***Index Terms*—NILM, energy disaggregation, energy efficiency.**

I. Introduction

Non-intrusive load monitoring (NILM) aims at disaggregating the total energy consumption of a building down to individual appliances using only software tools. NILM [1, 2] supports smart building automation and provides means of effective and meaningful energy feedback. Due to practical demands driven by improving energy efficiency and responding to global climate change goals, on one hand, and large-scale smart meter roll-outs, on the other hand, the research on NILM has intensified recently.

In the past, various signal processing and machine learning approaches have been used for NILM. These include hidden Markov models (HMMs) and their variants (Additive) Factorial HMM (AFHMM) with quadratic programming [3, 4, 5, 6], semi-definite programming relaxation [7], Bayesian nonparametric hidden semi-Markov models [8], etc. Another popular NILM method is based on semi-supervised learning via label propagation on graphs, that has its roots in Graph signal processing (GSP) theory [9], with promising results reported in [10, 11, 12] for various datasets. More recently, various deep learning-based architectures have been applied to NILM. Some contribution include the work of [13] that proposed a CNN architecture for sequence-to-point and sequence-to-sequence NILM learning. Furthermore, Michele et al. employed transfer learning with deep neural network model on NILM in [14]. [15] applied deep non-negative matrix factorization technique on NILM.

However, though significant progress has been made recently in improving practicality and accuracy of NILM algorithms, NILM has remained a challenging problem. As the start-of-the-art approach for NILM, HMM is not suitable when numerous loads are present because of its computational complexity; furthermore, HMM suffers from the fact that time duration of the operation of some appliance can vary significantly [16]. Additionally, HMM could be affected by unknown appliance noise as indicated in [12]. Similarly, GSP-based NILM can under-perform if the structure of the underlying graph does not capture well correlation between the collected samples [9]. Regarding deep learning approaches, training a model for each electrical appliance is the most common method currently, which leads to high complexity and high storage demands. Regarding to the current multi-label learning for appliance recognition in NILM, as indicated in [17], the multi-label learning was found to be competitive with the state-of-the-arts NILM algorithms.

In this paper, we propose a novel one-to-many CNN structure that represents the NILM problem as multi-load output using a single network. This way the complexity of training is reduced (one model is learned for the whole house instead of a separate model for each appliance). Our simulation results on a non-residential dataset with six appliances, demonstrate that the proposed multi-label network provides accurate disaggregation results.

We organized the rest of the paper as follows. After the introduction, Section II starts by describing the problem of NILM and CNN-based architectures, namely, sequence-to-point network and one-to-many network. Section II-D introduces the proposed CNN framework to address the one-to-many NILM problem. Section III follows up with detailed experiment design, experimental dataset, metrics, and results. Finally, the conclusion and future work are presented in Section IV.

II. Methods

A. Problem Formulation

NILM is a technique to estimate the electric consumption of each appliance while only the main meter is monitored. The main meter power reading y_τ can be expressed as:

$$y_\tau = \sum_{i=1}^{N} x_\tau^{(i)} + e, \tag{1}$$

where $\tau = 1, ..., T$, and $i = 1, ..., N$. T and N denote the number of time windows and appliances, respectively. $x_\tau^{(i)}$ is the electric consumption of appliance i in time-window τ, and

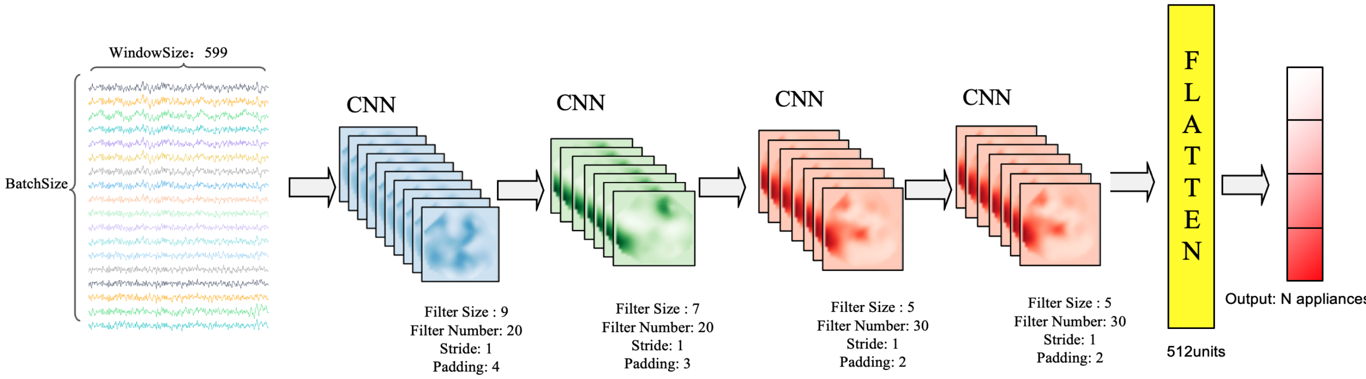

Fig. 1. The structure of the model used for multi-objects disaggregation.

e represents the noise recordings that includes unknown loads and measurement noise. The task of NILM is to estimate $x_\tau^{(i)}$ given y_τ for each time window τ.

B. Sequence-to-point (seq2point) CNN architecture

In seq2point CNN architecture [13], the input raw data is split into short time windows (sequences), each represented by its middle sample (point) [18]. The sequence-to-point model can be expressed as $x_t^{(i)} = f(y_\tau) + \epsilon$, in which y_τ is the input sequence, ϵ is the bias and $x_t^{(i)}$ the output of model, the true consumption of appliance i at time t which is the middle sample of window τ. $f(\cdot)$ is a mapping function of y_τ to $x_t^{(i)}$ for all appliances. The output of the model is decided based on maximising aposteriori probability of the model and can be formulated as:

$$\max \ p(x_t^{(i)} \mid y_\tau, \theta), \tag{2}$$

where θ denotes the overall network parameters, as described in the next section.

C. One-to-many model

Most recent NILM works focus on one-to-one model [5, 12, 14], where one model is trained for each appliance to maximize the probability of predicting a particular type of appliance, i.e.,

$$\max \ p\left(\mathbf{x}^{(i)} \mid \mathcal{Y}_\tau, \theta\right) = \max \prod_{t=1}^{T} p_t \ (x_t^{(i)} \mid y_\tau, \ \theta), \tag{3}$$

where $\mathcal{Y}_\tau$ is a set of T time windows, of fixed length, of the smart meter readings.

In our work, we adopt the one-to-many (one model for many appliances) structure to improve the efficiency of disaggregation. The model considers all N appliances at once and maximises:

$$\begin{aligned} &\max \ p(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(N)} \mid \mathcal{Y}_\tau, \ \theta) \\ &= \max \prod_{t=1}^{T} p_t(x_t^{(1)}, x_t^{(2)}, ..., x_t^{(N)} \mid y_\tau, \ \theta), \end{aligned} \tag{4}$$

where $p(x^{(1)}, x^{(2)}, ..., x^{(N)} \mid \mathcal{Y}_\tau, \theta)$ is a joint distribution probability function of all N appliances $\forall t \in [1, ..., T]$. $p_t(x_t^{(1)}, x_t^{(2)}, ..., x_t^{(N)} \mid y_\tau, \ \theta)$ is a joint distribution probability function of all N appliances at particular time t, $\mathcal{Y}_\tau$ is a set of T time windows of main data, and $\mathbf{x}^{(i)}, i \in [1, .., N]$ is a set of T predicted samples (one for each time window y_{tau} for appliance i.

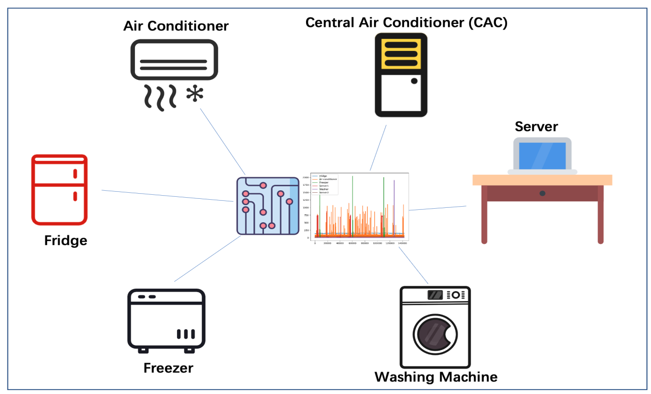

Fig. 2. A schematic of the smart meters deployed in the non-residential building room, which consist of a main electricity meter, and six appliance electricity meters.

D. One-to-many framework

Rather than estimating one appliance each time, we propose an efficient CNN architecture, which establishes the one-to-many model to estimate all appliances in the household. Firstly, we fragment the continuous recorded main power consumption into T time-windows of length of l samples each, and the input to the network is a batch of length-l time-series signals; thus, the network input is a $BatchSize * l * 1$ tensor.

As illustrated in Fig. 1, the designed CNN architecture comprises four convolution layers to capture time-dependent information in the receptive field. The particular filter size, stride, and padding used in our experiments are indicated in Fig. 1. The output of the convolution layer is then:

$$y_{ct} = \sum_{\forall k} y_{t+k} w_c + b_c, \tag{5}$$

where ct, a subscript of y, refers to c-th layer and t-th samples. t is the sample number in the input data, k is the length of the receptive field; w_c and b_c (bias) are the c-th layer parameters to be learned.

The fully connected layers follow the convolution layers, and can be formulated as:

$$y_{lt} = w_l y_t + b_l, \tag{6}$$

where y_{lt} represents the t-th output sample of the l-th fully connected layer. w_l and b_l represents the weight and bias parameters to be learned.

The loss function is set as:

$$L(x|f(y)) = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}(x_t^{(i)} - f^{(i)}(y_\tau))^2, \tag{7}$$

where $x_t^{(i)}$ denotes the true consumption of appliance i at time t, $f^{(i)}(y_\tau)$ represents the estimated consumption of appliance i at time t, and time t is the midpoint of time window τ.

III. Experiments

A. Data Sets

We evaluate the proposed method using our own recorded data. The results are then compared to those from literature that are based on two open access datasets, namely REDD and UK-DALE.

1) Own recorded IoT data: In order to evaluate the proposed network in a non-residential building, we deployed a set of smart electricity meters in an office building. There are six appliances in the office, namely, freezer, fridge, washer, air conditioner, central air conditioner (CAC), and server. Each of the appliances is monitored by a smart meter to provide ground truth values. The main meter is used as well. This data are sampled every second, and the dataset covers a period of 3 months. Fig. 2 is a schematic of the smart meters deployed in the non-residential building room. Fig. 3 shows an example of the true electric consumption of all appliances. Other public non-residential electrical datasets such as Commercial Building Energy Dataset (COMBED), and building-level office environment dataset (BLOND), *etc* described in [19], will be tested as part of the future work.

2) REDD data: Reference Energy Disaggregation Data set (REDD) contains electricity usage data of 6 American households. Each house has 2 mains meters, as well as individual monitoring meters 10 to 25 in each house. The data is sampled at a frequency of about once a second for a mains and once every three seconds for the circuits [20]. Appliances in REDD data set are similar to our own IoT data.

3) UK-DALE data: UK Domestic Appliance-Level Electricity (UK-DALE) data set [21] contains main meter readings of current and voltage of 16 kHz collected from three UK homes and power data of individual devices collected every six seconds. The low frequency data are often used in neural network models for NILM.

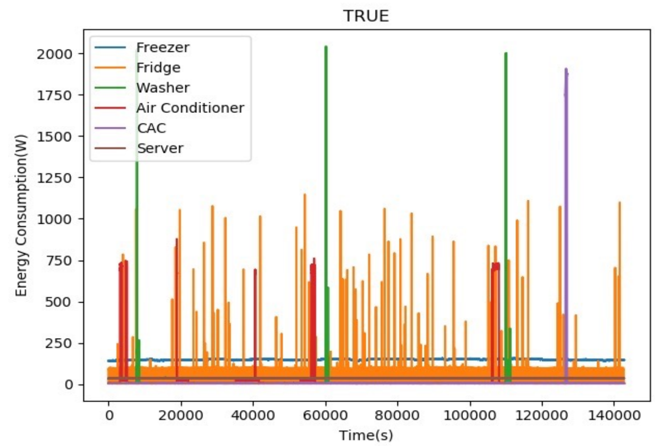

Fig. 3. The actual power consumption curve of each electrical appliances.

B. Metrics

As in the previous work on NILM [14], three metrics are used to measure the proposed multi-label disaggregation model. Usually, mean normalized disaggregation error (MNDE) is used as a metric as defined:

$$MNDE = \frac{\sum_{n,t}(x_{n,t} - \hat{x}_{n,t})^2}{\sum_{n,t} x_{n,t}^2}, \tag{8}$$

where $n = 1, ..., N$ denotes the appliance number, and t is the time index. $x_{n,t}$ and $\hat{x}_{n,t}$ means the true consumption and the estimated consumption of appliance n at time t, respectively.

Normalized signal aggregate error (SAE) is deployed to evaluate the difference between the total electric consumption and estimated electric consumption:

$$SAE = \sum_n \frac{|\hat{r_n} - r_n|}{r_n}. \tag{9}$$

Energy per day (NDE) is a metric to measure the prediction error on daily energy consumption:

$$EpD = \frac{1}{D}\sum_{n=1}^{D}|\hat{e} - e|, \tag{10}$$

where e and $\hat{e}$ represent the true energy consumption in a day, and D denotes the total number of days in the test data.

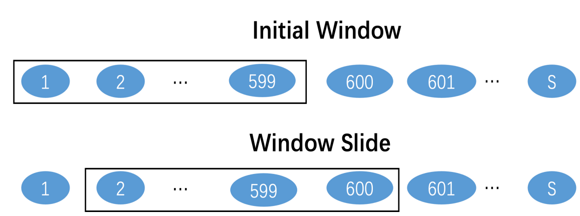

Fig. 4. The process of data generation.

TABLE I
METRICS OF MNDE, SAE AND EPD ON EACH APPLIANCE.

Metric	Method	DataSet	Fridge	Washing machine
MNDE	AFHMM[3]	REDD	0.99	84.53
	Proposed	IoT	0.90	0.99
SAE	AFHMM [3]	REDD	0.84	0.99
	Seq2seq [22]	REDD	0.24	0.11
	Seq2point[13]	REDD	0.06	0.18
	AFHMM[3]	UK-DALE	0.98	8.28
	Seq2seq [22]	UK-DALE	0.50	13.83
	Seq2point[13]	UK-DALE	0.37	0.45
	Proposed	IoT	0.77	0.73
EpD[kW]	AFHMM[3]	REDD	1.50	0.08
	AFHMM[3]	UK-DALE	0.90	0.32
	Proposed	IoT	0.40	0.20

C. Settings for training neural networks

Our method adopts one (model) to-many (appliances) structure, in particular, one model for six appliances. Based on the architecture shown in Fig. 1, a fixed-length window is adopted for the input data. The window length is set to 599, which was selected based on experiments. Each window data is generated by moving a sliding time window one sample a time. Fig. 4 shows the process of sliding window, where S is the length of the dataset (power reading samples). The first 70% of the dataset are selected as training samples, 20% testing samples, and 10% validation samples. Pytorch is used as the machine learning framework. The hyper-parameters are as follows: batchsize is 1024, max epoch is 100, Adam is used as the optimiser algorithm, learning rate is 0.001, and dropout is used as well to reduce overfitting.

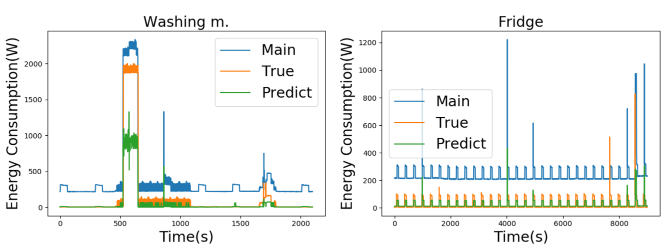

Fig. 5. Example disaggregation results on IoT data.

D. Results

The proposed one-to-many method is applied to the disaggregation problem using our data set. Fig. 5 shows an example of the disaggregation results using our IoT dataset. The left is an example for washing machine, and the right is an example for fridge which is on non stop. Table. I gives the results of the three metrics for estimating two appliances using different methods on the REDD data, UK-DALE data, and our own IoT data. For the convenience of comparison, these two appliances are of the same type in the three datasets.

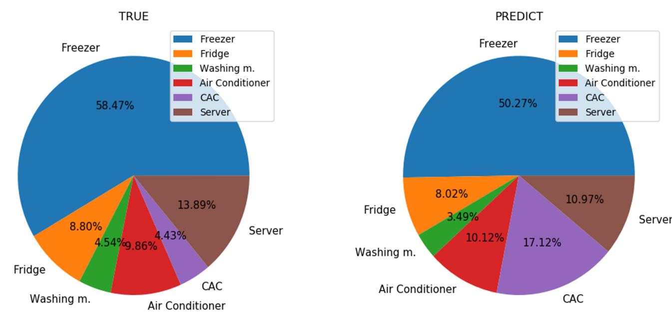

Fig. 6. The proportion of electricity consumed by each electrical appliance in the IoT dataset over the whole 3 months.

Compared to AFHMM, our method has better MNDE, SAE, and EpD results. Compared to the one-to-one seq2seq model and the one-to-one seq2point model, the results of our method are comparable based MNDE, SAE, and EpD metrics. The results suggest that while we have improved the effectiveness and practicality by disaggregating all appliances at once using a simple one-to-many NILM model, there is no significant decrease in accuracy.

The pie chart of disaggregation on all six appliances are shown in Fig. 6. The subgraph on the left is the true energy consumption of all the six appliances, and the right subgraph is the disaggregation result of the energy usage. The exact proportions of the six appliances are 58.6%, 8.8%, 4.54%, 9.86%, 4.43%, and 13.89%, respectively. The estimated proportions of the usage of the six appliances are 50.27%, 8.02%, 3.49%, 10.12%, 17.12%, and 10.97%. This chart shows that the one-to-many NILM method can provide a reasonable strategy for disaggregation, and the main disaggregation error comes from the appliance CAC represented in purple. The reason is that CAC data are more sparse as it is rarely used, thus the amount of training data available is insufficient. In the future, we will further improve the accuracy of disaggregation by few-shot learning method.

IV. Conclusion

Non-intrusive load monitoring (NILM) is vital for planning energy demand and developing energy conservation and energy efficiency tools.

In order to improve the practicality of NILM, we propose a one-to-many model for multi-label disaggregation. Experimental results on a non-residential dataset that contains 6 appliances, using MNDE, SAE, EpD metrics show that the proposed disaggregation method can achieve acceptable results using a single model, this way reducing the training complexity and storage demands. Further work will be perfomed in two areas, namely improving the accuracy of disaggregation for rarely-used appliances via few-shot learning, and testing the proposed method on other public datasets such as COMBED, BLOND *etc.*

Acknowledgment

This work was supported in part by the National Key Research and Development Program, China (Grant no. 2017YFE0119300), and Shanghai Sailing Program (Grant No. 19YF1451500). It was also partially supported by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 734331.

References

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

[2] W. Kong, Z. Y. Dong, B. Wang, J. Zhao, and J. Huang, "A practical solution for non-intrusive type ii load monitoring based on deep learning and post-processing," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 148–160, 2019.

[3] J. Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial hmms with application to energy disaggregation," in *Artificial intelligence and statistics*, pp. 1472–1482, PMLR, 2012.

[4] M. Zhong, N. Goddard, and C. Sutton, "Signal aggregate constraints in additive factorial hmms, with application to energy disaggregation," *Advances in Neural Information Processing Systems*, vol. 27, pp. 3590–3598, 2014.

[5] F. Paradiso, F. Paganelli, D. Giuli, and S. Capobianco, "Context-based energy disaggregation in smart homes," *Future Internet*, vol. 8, no. 1, p. 4, 2016.

[6] M. A. Mengistu, A. A. Girmay, C. Camarda, A. Acquaviva, and E. Patti, "A cloud-based on-line disaggregation algorithm for home appliance loads," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3430–3439, 2018.

[7] K. Shaloudegi, A. György, C. Szepesvári, and W. Xu, "Sdp relaxation with randomized rounding for energy disaggregation," *arXiv preprint arXiv:1610.09491*, 2016.

[8] Johnson, J. Matthew, Willsky, and S. Alan, "Bayesian nonparametric hidden semi-markov models.," *Journal of Machine Learning Research*, 2013.

[9] A. Ortega, P. Frossard, J. Kovacevic, J. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[10] K. He, L. Stankovic, J. Liao, and V. Stankovic, "Non-intrusive load disaggregation using graph signal processing," *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 1739–1747, 2016.

[11] K. He, D. Jakovetic, B. Zhao, V. Stankovic, L. Stankovic, and S. Cheng, "A generic optimisation-based approach for improving non-intrusive load monitoring," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6472–6480, 2019.

[12] B. Zhao, L. Stankovic, and V. Stankovic, "On a training-less solution for non-intrusive appliance load monitoring using graph signal processing," *IEEE Access*, vol. 4, pp. 1784–1799, 2016.

[13] C. Zhang, M. Zhong, Z. Wang, N. Goddard, and C. Sutton, "Sequence-to-point learning with neural networks for non-intrusive load monitoring," vol. 32, no. 1, 2018.

[14] M. D'Incecco, S. Squartini, and M. Zhong, "Transfer learning for non-intrusive load monitoring," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1419–1429, 2019.

[15] J. R. Hershey, J. L. Roux, and F. Weninger, "Deep unfolding: Model-based inspiration of novel deep architectures," *arXiv preprint arXiv:1409.2574*, 2014.

[16] C. Nalmpantis and D. Vrakas, "Machine learning approaches for non-intrusive load monitoring: from qualitative to quantitative comparation," *Artificial Intelligence Review*, vol. 52, no. 1, pp. 217–243, 2019.

[17] A. Faustine and L. Pereira, "Multi-label learning for appliance recognition in nilm using fryze-current decomposition and convolutional neural network," *Energies*, vol. 13, no. 16, p. 4154, 2020.

[18] M. Yang, X. Li, and Y. Liu, "Sequence to point learning based on an attention neural network for nonintrusive load decomposition," *Electronics*, vol. 10, no. 14, p. 1657, 2021.

[19] L. Pereira and N. Nunes, "Performance evaluation in non-intrusive load monitoring: Datasets, metrics, and tools—a review," *Wiley Interdisciplinary Reviews: data mining and knowledge discovery*, vol. 8, no. 6, p. e1265, 2018.

[20] 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, pp. 59–62, 2011.

[21] J. Kelly and W. Knottenbelt, "The uk-dale dataset, domestic appliance-level electricity demand and whole-house demand from five uk homes," *Scientific data*, vol. 2, no. 1, pp. 1–14, 2015.

[22] J. Kelly and W. Knottenbelt, "Neural nilm: Deep neural networks applied to energy disaggregation," pp. 55–64, 2015.