Low-Frequency Energy Disaggregation based on Active and Reactive Power Signatures

Pascal A. Schirmer and Iosif Mporas
School of Engineering and Computer Science
University of Hertfordshire
Hatfield AL10 9AB, UK
{p.schirmer,i.mporas}@herts.ac.uk

***Abstract*—Non-Intrusive Load Monitoring aims to extract the energy consumption of individual electrical appliances through disaggregation of the total power consumption as measured by a single smart meter at a household. Deep neural networks and especially Convolutional Neural Networks (CNNs) have become popular in solving the Non-Intrusive Load Monitoring problem. However, since NILM is a time series problem mostly 1-D CNNs have been utilized, thus not fully exploiting the capability of CNNs which are advantageous mostly in 2-D data such as images. Therefore, in the proposed architecture 2-D signatures of low frequency active and reactive power are utilized. The proposed architecture was evaluated on the AMPds2 dataset reporting performances up-to 96.1% in terms of estimation accuracy outperforming all previously reported approaches on the same dataset by 1.1%, in terms of absolute improvement.**

***Index Terms*—Non-Intrusive Load Monitoring (NILM), Energy Disaggregation, PQ-signatures.**

I. Introduction

Non-Intrusive Load Monitoring (NILM) aims to extract appliances' power consumption from the aggregated consumption of a building or a household [1]. NILM can be considered as a single-channel source separation problem and the NILM approaches that have been proposed can briefly be grouped into three categories. First, pattern matching (elastic matching) techniques, have been proposed in order to detect device signatures in the aggregated power consumption signal [2], [3]. Second, source separation methods, including matrix and tensor factorization as well as sparse coding, have been used in order to separate base components and activations [4], [5]. Third, data-driven approaches based on machine learning algorithms have been used, usually one per device, in order to estimate the power consumption of device(s) of interest from the aggregated signal [6], [7].

The latest advances of machine learning and the development of big datasets have led to successful deep learning based NILM methodologies. NILM architectures using Convolutional Neural Networks (CNNs) [8], [9], Long-Short-Term-Memory (LSTM) [10], [11] and Recurrent Neural Networks (RNNs) [12] have been proposed in the literature. CNN-based architectures, such as the approach proposed in [9] presenting a gate dilated CNN and the approach proposed in [12] presenting a fractional extension of [9], have reported exceptional high performances. Additionally, in [8] a CNN has been proposed for transfer learning in NILM and a high frequency concatenated CNN was proposed in [13]. Moreover, also Hidden Markov Models (HMMs) and their variants have shown good performances [6], [14], [15].

Due to the nature of the NILM problem and its time series characteristics most of the previously published CNN architectures are based on one-dimensional convolutional layers [8], [9], [12]. However, CNNs have been originally proposed for image classification, with 2-d convolutional and pooling layers performing as data-driven feature extraction engines [16], [17]. Spectrograms of 1-d signals (e.g. speech/audio) have been used to convert them to 2-d representations (images) and then been introduced to CNN models with remarkable outcomes [18], [19]. However, this is not possible in NILM as the sampling frequency is usually prohibitively low ($f_s \approx 1$ Hz) [20], [21].

In this paper we propose a two-dimensional active (P) and reactive (Q) power representation of each frame of the aggregated signal to create 2-d PQ-signatures. The proposed approach is the first low frequency NILM approach to use 2-d CNN models. The remainder of this paper is organized as follows: In Section II the introduced PQ-signature representations are described. In Section III the proposed architecture for low frequency NILM is presented. In Sections IV and V the experimental setup and the evaluation results are presented, respectively. Conclusions are provided in Section VI.

II. Low frequency PQ signature representation

The proposed PQ signature is a two-dimensional representation of apparent power (S). Let $p_{agg}(t)$ and $q_{agg}(t)$ be the aggregated active and reactive power measured by a smart meter with sampling frequency f_s and $t \in \mathbb{N}$, i.e. starting at time $t = 0$. Furthermore, let p^{τ}_{agg} be the τth frame of length W, with $p^{\tau}_{agg} = [p(t_0), p(t_0+1), ..., p(t_0+W-1)]$ where $p(t)$ is the tth sample of p_{agg}. Similarly, let q^{τ}_{agg} be a frame of length W of the aggregated reactive power. A two-dimensional PQ-signature can be defined by the signals p^{τ}_{agg} and q^{τ}_{agg} as described in Eq. 1.

$$S^{\tau}_{x,y} = \sqrt{p^{\tau}_{agg}(x)^2 + q^{\tau}_{agg}(y)^2} \tag{1}$$

where $S^{\tau} \in \mathbb{R}^{(W+1)\times(W+1)}$ with $0 \leq x, y \leq W$ is the two-dimensional instantaneous apparent power representation on the PQ plane for the τth frame with $S^{\tau}_{x,0} = p^{\tau}_{agg}$ and $S^{\tau}_{0,y} = q^{\tau}_{agg}$.

In the proposed PQ representation, $S^{\tau}_{x,y}$ captures, not only the temporal change of the active and reactive power within each frame, but also the difference in the changes of p^{τ}_{agg} and q^{τ}_{agg} with respect to each other. Specifically, each PQ row, x, represents the apparent power change for active power at time x when reactive power changes from t_0 to $t_0 + W - 1$, while each PQ column, y, represents the apparent power change for reactive power at time y when active power changes from t_0 to $t_0 + W - 1$. Examples for PQ signatures of individual appliances consumption measurements (S^{τ}_{m}) as well as the PQ signatures of the corresponding aggregated signals (S^{τ}_{agg}) from the AMPpds2 dataset [22] are shown in Fig. 1.

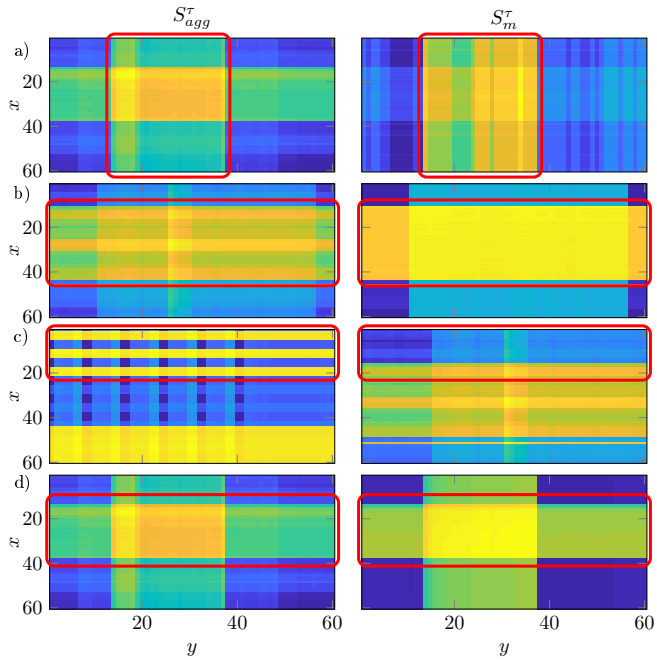

Fig. 1. Examples for aggregated (left) and appliance level (right) PQ-signatures for four different devices, namely: (a) HVAC, (b) cloth dryer, (c) dishwasher and (d) heat-pump calculated from the AMpds2 [22] dataset where x, y denotes the indices for active and reactive power respectively. Thick red lines denote PQ areas of characteristic similar patterns.

As shown in Fig. 1 PQ patterns are characteristic for each of the appliances (right column) and are visually identifiable in the aggregated PQ signatures (left column), which indicates that the proposed 2d PQ signature representations can be used in the task of energy disaggregation. In addition, the two-dimensional PQ signatures can further be transformed to the frequency domain calculating the two-dimensional discrete Fourier transform for each frame $S^{\tau}_{x,y}$, i.e.

$$\tilde{S}^{\tau}_{k,l} = \frac{1}{W^2} \sum_{x=0}^{W} \sum_{y=0}^{W} S^{\tau}_{x,y} \cdot e^{-j2\pi(\frac{k}{W}x + \frac{l}{W}y)} \tag{2}$$

where $1 \leq k < K$ and $1 \leq l < L$ being index variables and $\tilde{S}^{\tau} \in \mathbb{C}^{(W+1)\times(W+1)}$ being the complex spectrogram. Transformation of the PQ signatures to the frequency domain will represent the frequency content of active and reactive power. Examples of active and reactive aggregated power signals of two frames, (a) and (b), their PQ signatures in the time domain , (c) and (d), and their spectral magnitudes, (e) and (f) are shown in Fig. 2.

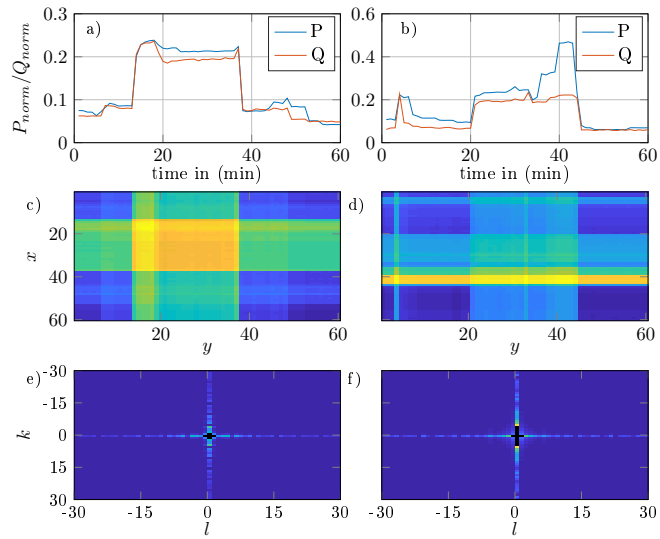

Fig. 2. Two frames of p_{agg} and q_{agg} (a-b) and their corresponding PQ singatures in (c-d) the time and the (e-f) frequency domain where x,y denotes the time-domain indices of $S^{\tau}_{x,y}$ and k,l denotes the frequency-domain indices.

As shown in Fig. 2 two different frames of aggregated active and reactive power result into characteristic representations of the corresponding PQ-signatures both in the time and in the frequency domain.

III. Proposed PQ-Signature based NILM

Considering a set of $M-1$ known devices each consuming power p_m with $1 \leq m \leq M$, the aggregated power p_{agg} measured by the sensor will be:

$$p_{agg} = f(p_1, ..., p_{M-1}, g) = \sum_{m=1}^{M-1} p_m + g = \sum_{m=1}^{M} p_m \tag{3}$$

where $g = p_M$ is a 'ghost' power consumption (noise) consumed by one or more unknown devices and $f(\cdot)$ is the aggregation function. In NILM the goal is to find precise estimations $\hat{p}_m, \hat{g}$ of the power consumption of each device m using an estimation method $f^{-1}(\cdot)$ with minimal estimation error and $\hat{p}_M = \hat{g}$, i.e.

$$\begin{aligned} \hat{P} &= \{\hat{p}_1, \hat{p}_2, ..., \hat{p}_{M-1}, \hat{g}\} = f^{-1}(p_{agg}) \\ s.t. \quad & \underset{f^{-1}}{\operatorname{argmin}} \{(p_{agg} - \sum_{m=1}^{M} \hat{p}_m)^2\} \end{aligned} \tag{4}$$

In the proposed approach the disaggregation function f^{-1} is using as input the two-dimensional PQ-signatures in the time and in the frequency domain (spectral magnitude), namely S^{τ} and $|\tilde{S}^{\tau}|$ as described in Eq. 5.

$$\hat{P} = \{\hat{p}_1, \hat{p}_2, ..., \hat{p}_{M-1}, \hat{g}\} = f^{-1}(S^{\tau}, |\tilde{S}^{\tau}|) \tag{5}$$

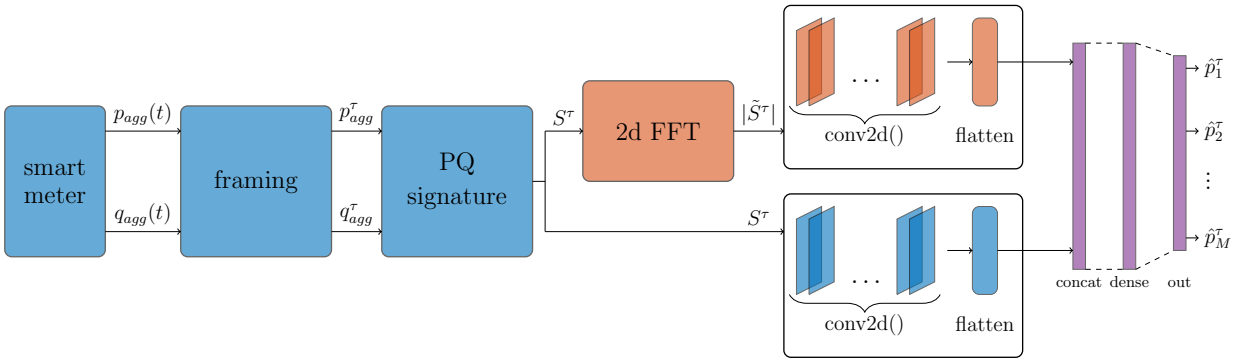

Fig. 3. Block diagram of the proposed NILM architecture, for the mth device, using PQ signatures in the time- and frequency-domain based on CNN regression. conv2d(x,y) denotes two-dimensional convolutional layers with x being the number of filters and y being the size of the convolution in both dimensions. Blue denotes time domain, red frequency domain and purple time/frequency domain operations.

where $|\tilde{S}^{\tau}|$ is the magnitude of the two-dimensional frequency representation. As Eq. 5 cannot be solved analytically [23], numerical approaches must be considered. Therefore, the two-dimensional PQ signatures are selected as input for the CNNs as part of a learning-based approach. The block diagram of the proposed architecture is illustrated in Fig. 3.

As illustrated in Fig. 3 the proposed architecture consists of four steps namely the data acquisition of the aggregated signals $p_{agg}(t)$ and $q_{agg}(t)$ from one smart-meter, segmentation into time frames p^{τ}_{agg} and q^{τ}_{agg}, calculation of PQ signatures (one PQ signature per frame) in the time and the frequency domain (S^{τ} and $|\tilde{S}^{\tau}|$) and CNN regression providing a numerical estimation, $\hat{P} = \{\hat{p}_1, \hat{p}_2, ..., \hat{p}_M\}$, of the power consumption of each of the M target devices.

IV. Experimental Setup

The NLIM architecture utilizing PQ signature representations presented in Section III was evaluated using the datasets, experimental protocols and CNN regression algorithms described below.

A. Datasets and Experimental Protocols

There are several different datasets for NILM [22], [24]–[26] with varying characteristics such as sampling frequency, appliances, duration and measured features. To evaluate the proposed low-frequency NILM methodology the AMPds2 [22] dataset was used, a low-frequency dataset with sampling frequency of 1 sample per minute and a monitoring duration of 2 years. Specifically, the dataset was chosen as it contains active and reactive power measurements from 20 different devices as well as the aggregated power and current signals. Next to using all 20 loads, five out of the 20 loads (deferrable loads [6], [9]), namely the HVAC system (AC), the Heat Pump (HP), the Wall Oven (WO), the Cloth Dryer (CD) and the DishWasher (DW), were chosen for disaggregation on a subset of appliances similarly to [9].

Regarding the experimental setup the protocol followed in [9] was adopted in this study for the purpose of direct comparison with the state-of-the-art approaches reporting results on the AMPds2 dataset. Specifically, training was conducted using randomly 90% of the data and testing on the remaining 10% using active and reactive power measurements normalized in the amplitude range (0,1). In detail, three different experimental protocols were created with respect to the PQ signature representations described in Section II. The three experimental protocols, namely PQ signatures (#1) in time (#2) in frequency and (#3) in time-and-frequency as well as the corresponding size of each PQ frame, are tabulated in Table I.

TABLE I
Three evaluated experimental protocols including their features and dimensionality D.

Protocol	Features	D
#1	PQ-signature time ($S^{\tau}_{x,y}$)	31×31
#2	PQ-signature frequency ($\|\tilde{S}^{\tau}_{x,y}\|$)	31×31
#3	PQ-signature time/frequency ($[S^{\tau}_{x,y}, \|\tilde{S}^{\tau}_{x,y}\|]$)	31×62

The length of the frames was selected equal to W=30 (30 minutes) after empirical optimization on a bootstrap training dataset utilizing two months of the AMPds2 dataset, thus resulting to the $W + 1$ PQ signature sizes tabulated in the last column of Table I.

B. CNN Regression Model

For the regression stage a CNN model-based approach was evaluated. Specifically, for the CNN the architecture from [8] was adapted using one branch for time domain PQ features and one for frequency domain PQ features as illustrated in Fig. 3. The architecture of each of the two branches was set to conv2d(30,10); conv2d(30,8); conv2d(40,6); conv2d(50,5); conv2d(50,5), as proposed in [8]. The size of dense layer was equal to 1024 nodes and the learning hyper-parameters are tabulated in Table II.

V. Experimental Results

The architecture presented in Section III was evaluated according to the experimental setup described in Section IV. The performance was evaluated in terms of estimation accuracy (E_{ACC}) considering device operation on state level with a double counting for errors as proposed in [24], i.e.

TABLE II
HYPER-PARAMETERS OF THE CNN MODEL AND PARAMETERS OF THE ADAM SOLVER.

Parameter	Protocols #1/#2	Protocol #3
Input size	31×31	31×62
Batch size	1000	1000
Epochs	50	50
Learning rate	0.001	0.001
Beta-1	0.9	0.9
Beta-2	0.999	0.999
Epsilon	1e-8	1e-8

$$E_{ACC} = 1 - \frac{\sum_{t=1}^{T}\sum_{m=1}^{M}|\hat{p}_m^\tau - p_m^\tau|}{2\sum_{t=1}^{T}\sum_{m=1}^{M}|p_m^\tau|} \quad (6)$$

where $\hat{p}_m^\tau$ is the estimated power of the mth device, T is the number of disaggregated frames and M is the number of devices. Removing the summation over M the disaggregation accuracy for the mth device can be written as in Eq. 7.

$$E_{ACC}^m = 1 - \frac{\sum_{t=1}^{T}|\hat{p}_m^\tau - p_m^\tau|}{2\sum_{t=1}^{T}|p_m^\tau|} \quad (7)$$

Additionally, a pattern matching approach based on DTW was evaluated, similarly as in [2], comparing each two-dimensional PQ signature against a set of reference PQ signatures created from the training data. The experimental results for the deferrable loads in terms of E_{ACC} for the CNN and the DTW based approaches are tabulated in Table III.

TABLE III
ENERGY DISAGGREGATION RESULTS FOR DEFERRABLE LOADS IN TERMS OF E_{ACC} ('AVG' ROW) AND E_{ACC}^m ('APP' ROWS) FOR THREE PROPOSED PROTOCOLS FOR CNN AND DTW RESPECTIVELY.

App	CNN			DTW		
	#1	#2	#3	#1	#2	#3
DW	49.6%	52.9%	63.6%	37.5%	19.7%	38.9%
AC	93.2%	92.3%	93.2%	91.8%	91.4%	91.6%
HP	97.9%	92.4%	97.8%	96.7%	93.8%	96.7%
WO	76.7%	54.5%	56.0%	61.7%	51.8%	62.7%
CD	96.1%	80.1%	96.4%	89.4%	84.7%	89.6%
AVG	95.2%	87.3%	**95.8%**	91.5%	88.3%	**92.4%**

As shown in Table III the CNN based regression model outperforms the DTW based approach for all three experimental protocols with protocol #3 presenting the highest overall performance for both CNN (95.8%) and DTW (92.4%). Also, time domain PQ-signatures (protocol #1) outperform frequency-based ones (protocol #2) for both CNN and DTW. This is probably due to the very low sampling frequency of 1/60 Hz, resulting in loss of frequency content.

To further evaluate the proposed NILM approach, comparison with the highest reported accuracies [9], [12] found in the literature is presented in Table IV, in terms of power disaggregation (OUT: P) and current disaggregation (OUT: I) for all loads (ALL) and deferrable loads (DEF).

As shown in Table IV the proposed approach based on 2-d PQ based CNNs outperforms both previously reported

TABLE IV
PERFORMANCE IN TERMS OF E_{ACC} FOR ALL AND DEFERRABLE LOADS FOR TWO DIFFERENT OUTPUT SIGNALS. ACCURACIES DENOTED WITH '*' ARE USING APPARENT POWER AND CURRENT ADDITIONALLY TO ACTIVE AND REACTIVE POWER, THUS HAVE ADDITIONAL INFORMATION COMPARED TO THE PROPOSED NILM METHOD.

Loads	CNN	DTW	[9]	[12]
ALL (OUT: P)	**89.6%**	83.4%	87.5%	88.9%
DEF (OUT: P)	**95.8%**	92.4%	93.9%	94.7%
ALL (OUT: I)	**91.4%**	84.7%	90.2%*	90.8%*
DEF (OUT: I)	**96.1%**	92.6%	95.0%*	92.7%*

approaches [9], [12], which are based on 1-d CNN models. The maximum absolute performance improvement was 2.1% (ALL (Out: P)) when comparing to [9], while the maximum absolute improvement was 1.1% and 1.2% for the setup utilizing deferrable loads and current as output signal. The performances of DTW are 2-6% worse compared to the CNN based approaches, however DTW comes with the advantage of not having to train a model [2]. Finally, except the improvements in performance the proposed approach is advantageous in terms of learning rate as it converged within 50 epochs, while the approach in [9] needed 300 (deferrable loads) or 500 (all loads) epochs to converge. The two convergence curvatures are illustrated in Fig. 4.

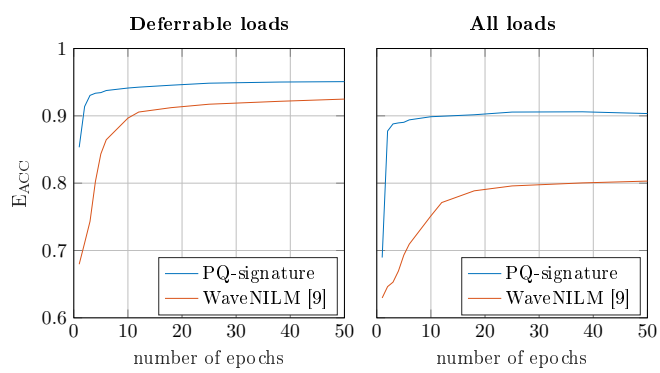

Fig. 4. Convergence behaviour for the proposed PQ-signatures (OUT: I) as well as for the waveNILM (OUT: I) approach presented in [9] for the first 50 epochs of training. For convergence of "waveNILM" beyond 50 epochs the interested reader is referred to [9].

VI. CONCLUSION

We presented a two-dimensional representation of energy consumption using active and reactive power measurements, which was used as input to a convolutional neural network model for regression. Using the proposed two-dimensional representations in the time and in the frequency domain, after 2-d discrete Fourier transform, significantly improved the non-intrusive load monitoring performance when directly compared the top performing approaches reported in the literature for low frequency energy consumption signals. The proposed architecture was evaluated on the AMPds2 dataset reporting performances up-to 96.1% in terms of estimation accuracy outperforming all previously reported approaches on the same dataset by 1.1%, in terms of absolute improvement.

REFERENCES

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

[2] Pascal A. Schirmer, Iosif Mporas, and Michael Paraskevas, "Energy disaggregation using elastic matching algorithms," *Entropy*, vol. 22, no. 1, pp. 71, 2020.

[3] Kanghang He, Lina Stankovic, Jing Liao, and Vladimir Stankovic, "Non-intrusive load disaggregation using graph signal processing," *IEEE Transactions on Smart Grid*, p. 1, 2016.

[4] Shikha Singh and Angshul Majumdar, "Deep sparse coding for non-intrusive load monitoring," *IEEE Transactions on Smart Grid*, p. 1, 2017.

[5] Shikha Singh and Angshul Majumdar, "Analysis co-sparse coding for energy disaggregation," *IEEE Transactions on Smart Grid*, p. 1, 2017.

[6] Stephen Makonin, Fred Popowich, Ivan V. Bajic, Bob Gill, and Lyn Bartram, "Exploiting hmm sparsity to perform online real-time nonintrusive load monitoring," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2575–2585, 2016.

[7] Shirantha Welikala, Chinthaka Dinesh, Mervyn Parakrama B. Ekanayake, Roshan Indika Godaliyadda, and Janaka Ekanayake, "Incorporating appliance usage patterns for non-intrusive load monitoring and load forecasting," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 448–461, 2019.

[8] Michele DrIncecco, Stefano Squartini, and Mingjun Zhong, "Transfer learning for non-intrusive load monitoring," *IEEE Transactions on Smart Grid*, p. 1, 2019.

[9] Alon Harell, Stephen Makonin, and Ivan V. Bajic, "Wavenilm: A causal neural network for power disaggregation from the complex power signal," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2019, pp. 8335–8339, IEEE.

[10] Lukas Mauch and Bin Yang, "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): 14-16 Dec. 2015*, Piscataway, NJ and Piscataway, NJ, 2015, pp. 63–67, IEEE.

[11] Maria Kaselimi, Nikolaos Doulamis, Anastasios Doulamis, Athanasios Voulodimos, and Eftychios Protopapadakis, "Bayesian-optimized bidirectional lstm regression model for non-intrusive load monitoring," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2019, pp. 2747–2751, IEEE.

[12] Pascal A. Schirmer and Iosif Mporas, "Energy disaggregation using fractional calculus," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 5/4/2020 - 5/8/2020, pp. 3257–3261, IEEE.

[13] Qian Wu and Fei Wang, "Concatenate convolutional neural networks for non-intrusive load monitoring across complex background," *Energies*, vol. 12, no. 8, pp. 1572, 2019.

[14] J. Zico Kolter and Tommi Jaakkola, "Approximate inference in additive factorial hmms with application to energy disaggregation," in *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*, Neil D. Lawrence and Mark Girolami, Eds., La Palma, Canary Islands, 2012, vol. 22 of *Proceedings of Machine Learning Research*, pp. 1472–1482, PMLR.

[15] Matthew J. Johnson and Alan S. Willsky, "Bayesian nonparametric hidden semi-markov models," *The Journal of Machine Learning Research*, vol. 14, no. 1, pp. 673–701, 2013.

[16] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks," 2012, pp. 1097–1105.

[17] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," *IEEE transactions on neural networks*, vol. 8, no. 1, pp. 98–113, 1997.

[18] Ossama Abdel-Hamid, Abdel-rahman Mohamed, Hui Jiang, Li Deng, Gerald Penn, and Dong Yu, "Convolutional neural networks for speech recognition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 10, pp. 1533–1545, 2014.

[19] Ossama Abdel-Hamid, Abdel-rahman Mohamed, Hui Jiang, and Gerald Penn, "Applying convolutional neural networks concepts to hybrid nn-hmm model for speech recognition," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012*, Piscataway, NJ, 3/25/2012 - 3/30/2012, pp. 4277–4280, IEEE.

[20] Pascal A. Schirmer and Iosif Mporas, "Energy disaggregation from low sampling frequency measurements using multi-layer zero crossing rate," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 5/4/2020 - 5/8/2020, pp. 3777–3781, IEEE.

[21] Pascal A. Schirmer and Iosif Mporas, "Improving energy disaggregation performance using appliance-driven sampling rates," in *2019 27th European Signal Processing Conference (EUSIPCO)*. 2019, pp. 1–5, IEEE.

[22] Stephen Makonin, Bradley Ellert, Ivan V. Bajić, and Fred 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.

[23] Pascal A. Schirmer, Iosif Mporas, and Akbar Sheikh-Akbari, "Energy disaggregation using two-stage fusion of binary device detectors," *Energies*, vol. 13, no. 9, pp. 2148, 2020.

[24] J. Zico Kolter and Matthew J. Johnson, Eds., *REDD: A Public Data Set for Energy Disaggregation Research*, 2011.

[25] Jack Kelly and William Knottenbelt, "The uk-dale dataset, domestic appliance-level electricity demand and whole-house demand from five uk homes," *Scientific data*, vol. 2, pp. 150007, 2015.

[26] David Murray, Lina Stankovic, and Vladimir Stankovic, "An electrical load measurements dataset of united kingdom households from a two-year longitudinal study," *Scientific data*, vol. 4, pp. 160122, 2017.