# Exploring Time Series Imaging for Load Disaggregation

Hafsa Bousbiat
University of Klagenfurt, Austria
hafsa.bousbiat@aau.at

Christoph Klemenjak
University of Klagenfurt, Austria
klemenjak@ieee.org

Wilfried Elmenreich
University of Klagenfurt, Austria
wilfried.elmenreich@aau.at

## ABSTRACT

In this paper, we investigate the benefits of time-series imaging in load disaggregation, as we augment the wide-spread sequence–to–sequence approach by a key element: an imaging block. The approach presented in this paper converts an input sequence to an image, which in turn serves as input to a modified version of a common Denoising Autoencoder architecture used in load disaggregation. Based on these input images, the Autoencoder estimates the power consumption of a particular appliance. The main contribution presented in this paper is a comparison study between three common imaging techniques: Gramian Angular Fields, Markov Transition Fields, and Recurrence Plots. Further, we assess the performance of our augmented networks by a comparison with two benchmarking implementations, one based on Markov Models and the other one being a common Denoising Autoencoder. The outcome of our study reveals that in 19 of 24 cases, the considered augmentation techniques provide improved performance over the baseline implementation. Further, the findings presented in this paper indicate that the Gramian Angular Field could be better suited, though the Recurrence Plot was observed to be a viable alternative in some cases.

## CCS CONCEPTS

• **Computing methodologies** → *Machine learning*; *Image processing*; *Image representations*; • **Hardware** → *Smart grid*.

## KEYWORDS

time-series imaging, load disaggregation, NILM, neural networks

**ACM Reference Format:**
Hafsa Bousbiat, Christoph Klemenjak, and Wilfried Elmenreich. 2020. Exploring Time Series Imaging for Load Disaggregation. In *The 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '20), November 18–20, 2020, Virtual Event, Japan.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3408308.3427975

## 1 INTRODUCTION

Non-Intrusive Load Monitoring (NILM), or load disaggregation, describes the problem of identifying electrical appliances within a time series, which is typically provided by an energy meter [9]. Ever since NILM was introduced, data analysis techniques have been considered for adoption to improve estimation performance. Recently, data augmentation from 1D to 2D representation has received increased interest.

Many load disaggregation techniques perform sequence–to–sequence mapping in the sense that the technique learns to reconstruct the time series of a particular household appliance on the basis of one-dimensional input, a time series of a smart meter. Prominent examples are autoencoders as in [13] and, more recently, sequence–to–sequence estimation [21]. While it was demonstrated that such techniques can provide decent performance under certain conditions [18], a recent study finds that further research is required to close the performance gap to existing load disaggregation algorithms [8]. Inspired by the success of time-series imaging in related research fields [7, 10], we suspect that looking beyond one-dimensional NILM could bear the potential to close this performance gap. In particular, we propose to augment the existing sequence–to–sequence approach by a key element: *image transforms.*

In this paper, we explore the use of three time-series imaging techniques for low-frequency load disaggregation (i.e. sampling intervals in the range of 1 s to 60 s). We present an experimental setup consisting of an image transform stage followed by a denoising autoencoder, as we strive to evaluate if augmenting the sequence–to–sequence learning approach with image transforms results in enhanced performance. Further, we aim to identify the most promising type of time-series imaging.

The remainder of the paper is organized as follows: Section 2 discusses related work with regard to the application of image transforms in NILM. Section 3 introduces the image transforms considered and the experimental setup applied in our investigations. Section 4 presents the outcome of our evaluation. Section 5 concludes the paper. Auxiliary material as well as the implementation of our setup can be obtained from our online repository[1].

## 2 RELATED WORK

Image representations of time series have received some attention in NILM. The authors of [2] represent the plots of the current-voltage trajectory as binary images from which they extract contours by using the marching squares algorithm. Once the contours are extracted, they calculate their elliptic Fourier descriptors which serve as input to a classifier. In a follow-up investigation, the authors examine the automated detection of unknown appliances in the household using the same VI (Voltage–Current) binary images [3]. Those VI images are mapped to a new feature space learned by Siamese neural networks. To improve load disaggregation based on VI features, a weighted recurrence plot was proposed in [6], where the values of the matrix are no more binary but rather in a specified interval. Thus, this approach aims to reduce information loss.


*BuildSys '20, November 18–20, 2020, Virtual Event, Japan*

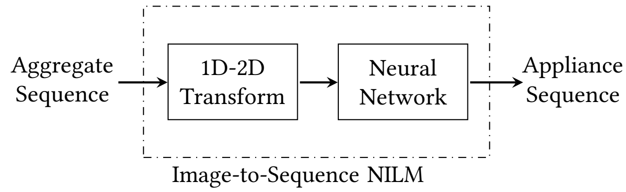
ACM ISBN 978-1-4503-8061-4/20/11…$15.00
https://doi.org/10.1145/3408308.3427975

[1] https://github.com/BHafsa/image-nilm

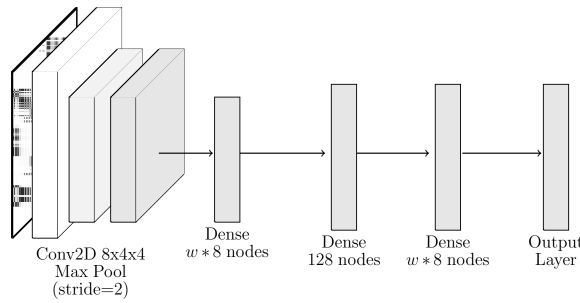
In [20], the authors proposed the use of data sampled at a high rate to formulate sequences of the current waveform as a one-channel image that serves as an input of a CNN to classify appliance's consumption. To overcome the vanishing gradient problem, a residual model was proposed in [11] to classify appliance's states using the Gramian Angular Field (GAF) representation of high-frequency power measurements.

In [15], the authors proposed to encode low-frequency power measurements as images using GAF. Once the time series is converted, its image representation is fed to a pre-trained network, a VGG16 in this case. This step is followed by a classifier that detects appliances states (ON/OFF). Though the experiments presented in this study are promising, they are limited to a single appliance, a fridge. The authors of [16] also adopted the GAF representation to train CNNs on the encoded images and perform load disaggregation. However, in the latter approach the target consumption was represented in polar coordinates.

From a general point of view, NILM scholarship has considered image representations. However, some promising imaging techniques, such as MTF [19], have not been considered at all. Furthermore, we claim that scholars have not unlocked the full potential of time series consisting of power readings in this context, as the majority of energy meters provides data in this particular form. We identify the need for a comparison study of image transforms for this kind of time series. The main motivation of this contribution is to close this gap by identifying the most promising imaging technique in the context of load disaggregation for sampling intervals in the range of 1 s to 60 s.




(a) Processing pipeline




(b) Architecture of the neural network

**Figure 1: The experimental setup in this work**

## 3 TIME SERIES IMAGING FOR NILM

A considerable number of load disaggregation techniques perform a sequence–to–sequence mapping between an input sequence (aggregate power readings) and an output sequence that consists of power readings associated with a particular electrical appliance [13, 21]. In our experimental setup, we augment this pipeline with a 1D-2D transform: a time series imaging block. Fig. 1a shows the big picture of our approach. The neural network takes as input the image representation of a sequence of power values obtained by a sliding window. The role of the neural network is to learn the mapping between the image representation and the corresponding appliance sequence. In Fig. 1b, we illustrate the architecture of the network embedded in our implementation. We adapted the Denoising Autoencoder (DAE) presented in [12], where our adaptation consisted of changing the first layer from a 1D convolutional layer into a 2D convolutional layer. This way, we minimize the effect of the network on the disaggregation results, as we strive to measure the influence of the 2D representation in our investigation. The output layer gives a sequence of values representing the power consumption of the target appliance.

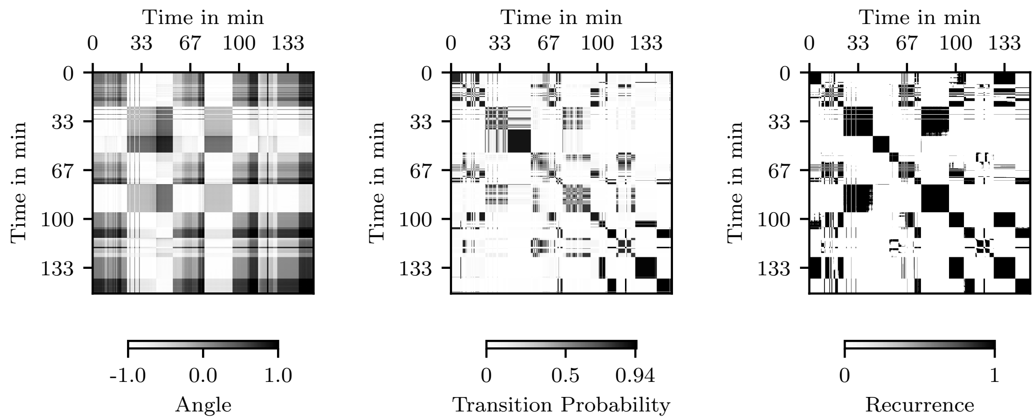
In the course of experimental evaluations, we aim to identify the most promising imaging technique for a disaggregation pipeline as depicted in Fig. 1a. We incorporate three imaging techniques provided by [5]:

The *Gramian Angular Field [19]* is a square image representation of the input sequence. It preserves temporal dependencies between different instants of the sequence, as the time increases with the position (from top-left to bottom-right). This approach comprises two main techniques: Gramian Angular Summation Field (GASF) and Gramian Angular Difference Field (GADF). In this work, we consider GASF only. To construct this representation, the input data is first scaled using a min-max normalization and then represented in the polar coordinate system. The final elements of the 2D representation consist of the angular cosine function of the sum between each two points which help to identify temporal correlations.

A *Markov Transition Field [19]* generates images that consist of transition probabilities. A Markov Transition Field (MTF) identifies the $Q$ quantile bins from the input sequence X (based on the magnitude) and assigns each element of the sequence $x_i$ to a quantile $q_j$. Each element of the Markov matrix represents the transition probability from quantile $q_i$ to quantile $q_j$. The final MTF representation is constructed as a matrix whose elements $M_{ij}$ represent the probability of transition from the quantile of the value $x_i$ to the quantile of the value $x_j$.

The *Recurrence Plot [4]* is a binary representation of a similarity measure between the instants of the input sequence X. In our work, we consider the *Euclidean distance*. To construct the image representation, the distance between each two instants is calculated. Afterwards, a pre-defined threshold is subtracted from these values which are than transformed to binary values with the Heaviside function.

We provide three sample images created from 150 min of smart meter data in Fig. 2 to get a basic understanding of how images obtained from GASF, MTF and RP differ. All three transforms generate symmetric images where the value of a pixel gives insights on the similarity between two instants.




(a) Gramian Angular Field (b) Markov Transition Field (c) Recurrence Plot

**Figure 2: The output of common time-series imaging techniques applied to 150 minutes of smart meter data**

## 4 EVALUATION

We considered data from three different energy consumption datasets in this study: REFIT [17], SynD [14], and UK-DALE [13]. For one household per dataset, we selected four common household appliances: dishwasher, fridge, microwave and washing machine. We extracted a subset of 105 days, which we split into 90 days to serve as training set and the remaining 15 days to evaluate our algorithms on. We applied a sampling interval of 10 s throughout our investigation.

The experimental setup was designed to be fully compatible to NILMTK [1], a NILM toolkit for reproducible experiments. We incorporate implementations of GAF, MTF, and RP provided by the PyTS package [5]. In order to get a broader understanding of the performance of our network, we benchmark the outcome of our experiments to the Factorial Hidden Markov Model (FHMM) and the Denoising Autoencoder (DAE) approaches provided by NILMTK. We trained DAE as well as our augmented network for 20 epochs each. To be precise, we trained one network for every image transform and every appliance.

In accordance with previous studies [12], we set the input time window to 20 min for fridges and microwaves, whereas we selected a time window of 30 min for dishwashers and washing machines.

We employ two metrics to measure how well the disaggregated data resembles the actual power consumption. The Mean Average Error (MAE) for an appliance i is defined as

$$\mathrm{MAE}^{(i)} = \frac{1}{N} \cdot \sum_{t=0}^{N-1} |\hat{y}_t^{(i)} - y_t^{(i)}| \tag{1}$$

where, $y_t$ is the actual power consumption, $\hat{y}_t$ is the estimated power consumption, and $N$ represents the number of samples. As absolute measure, MAE is reported in Watts. In addition to MAE, we report a normalized metric, the *Normalized Disaggregation Error (NDE)*:

$$\mathrm{NDE}^{(i)} = \sqrt{\frac{\sum_{t=0}^{N-1} (\hat{y}_t^{(i)} - y_t^{(i)})^2}{\sum_{t=0}^{N-1} (y_t^{(i)})^2}} \tag{2}$$

### 4.1 Results & Discussion

Table 1 summarizes the outcome of our experiments on houses of three different energy datasets: house 1 of SynD, house 6 of REFIT, and house 1 of UK-DALE. We observe that the baseline DAE as well as our augmented networks clearly outperform the FHMM in every single case. More interestingly, the outcome of our investigation shows that in 19 of 24 test runs, the networks augmented by time-series imaging provide better disaggregation performance than the baseline DAE, which differs from the augmented networks in one layer only. We measure considerable improvements in the case of the fridges in SynD and UK-DALE for all three image transforms. Exceptions to this strong trend, however, can be identified. The dishwasher in SynD and UK-DALE shows better performance in case of the conventional DAE for both metrics as well as the microwave in REFIT with regard to the NDE metric. We suspect this is due to differences in terms of optimal hyper-parameters between DAE and our augmented network, as we used the same settings for both approaches.

A comparison among the performance of the networks enhanced by image transforms techniques (MTF, GAF, and RP) reveals that in 15 of 24 cases, the network augmented by GAF provides superior performance compared to RP and MTF. We hypothesize that the cosine function employed by the GAF representation may be better in representing recurrent patterns in time series of household appliances and hence, boost the performance of the neural network attached to it. It should be noted, however, that the network utilizing RP scores best in 8 of 24 cases, which indicates that this image transform can provide an alternative to GAF in certain cases and should not be neglected.

A limitation of data augmentation with time-series imaging lies in increased demand of computational power and memory requirements compared to sequence–to–sequence approaches. While we record enhanced performance as a result of the use of time series imaging for the vast majority of test runs, it was observed that this performance gain comes at the cost of considerably increased training time. A promising but yet untested approach to overcome

**Table 1: Comparing the disaggregation performance of time series imaging techniques on three selected datasets**

| | | Mean Absolute Error (MAE) in Watts | | | | | Normalized Disaggregation Error (NDE) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Appliance | FHMM | DAE | DAE+MTF | DAE+GAF | DAE+RP | FHMM | DAE | DAE+MTF | DAE+GAF | DAE+RP |
| SynD 1 | fridge | 8.6 | 22.2 | 2.0 | **1.46** | 1.6 | 0.64 | 0.80 | **0.15** | 0.19 | 0.14 |
| | dish washer | 107.0 | **4.9** | 12.7 | 8.0 | 8.8 | 1.20 | **0.17** | 0.38 | 0.30 | 0.25 |
| | washing machine | 66.3 | 11.7 | 9.4 | 10.3 | **6.9** | 1.06 | 0.35 | 0.28 | 0.28 | **0.23** |
| | microwave | 56.3 | 2.3 | 1.4 | **1.1** | 1.4 | 4.99 | 0.99 | 0.88 | **0.74** | 0.85 |
| REFIT 6 | fridge | 32.7 | 27.0 | 25.4 | **24.2** | 24.3 | 1.22 | 0.88 | 0.86 | **0.84** | 0.85 |
| | dish washer | 58.6 | 12.6 | 22.8 | **12.2** | 20.6 | 2.23 | 0.85 | 0.99 | **0.81** | 0.99 |
| | washing machine | 227.4 | 8.8 | 9.2 | 7.4 | **6.4** | 2.33 | 0.96 | 0.85 | 0.74 | **0.69** |
| | microwave | 381.9 | 7.6 | 8.9 | **5.8** | 9.9 | 5.21 | **0.71** | 0.99 | 0.79 | 0.95 |
| UK-DALE 1 | fridge | 60.0 | 46.3 | 32.8 | **26.9** | 30.4 | 1.17 | 0.80 | 0.66 | **0.58** | 0.63 |
| | dish washer | 76.9 | **6.3** | 20.3 | 10.8 | 11.2 | 2.30 | **0.49** | 0.80 | 0.62 | 0.55 |
| | washing machine | 233.5 | 14.5 | 36.2 | 13.9 | **11.9** | 1.61 | 0.38 | 0.71 | 0.36 | **0.33** |
| | microwave | 71.4 | 9.5 | 11.5 | **7.5** | 9.9 | 2.51 | 0.76 | 0.99 | **0.64** | 0.96 |

this drawback could be to disable the image transform element in the disaggregation pipeline for sub-sequences that contain minor changes or turn out to be sequences of constants.

## 5 CONCLUSIONS

Inspired by the success of data augmentation techniques in related work, we propose to augment the traditional sequence–to–sequence approach in NILM by a key element: time series imaging. We combine a modified version of a common sequence–to–sequence technique, a Denoising Autoencoder, with an image transform element to investigate the performance gain provided by this augmentation. We find that in 19 of 24 test runs, this augmentation results in increased disaggregation performance. Further, we observe that the Gramian Angular Field outperformed other imaging techniques in the vast majority of experiments, though there are indications that Recurrence Plots could serve as a viable alternative in some cases.

## ACKNOWLEDGMENTS

This work was supported by the Doctoral College DECIDE. The authors thank Andreas Reinhardt for providing insightful comments.

## REFERENCES


[1] Nipun Batra, Jack Kelly, Oliver Parson, Haimonti Dutta, William Knottenbelt, Alex Rogers, Amarjeet Singh, and Mani Srivastava. 2014. NILMTK: an open source toolkit for non-intrusive load monitoring. In *Proceedings of the 5th international conference on Future energy systems*. 265–276.

[2] Leen De Baets, Chris Develder, Tom Dhaene, and Dirk Deschrijver. 2017. Automated classification of appliances using elliptical fourier descriptors. In *2017 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 153–158.

[3] Leen De Baets, Chris Develder, Tom Dhaene, and Dirk Deschrijver. 2019. Detection of unidentified appliances in non-intrusive load monitoring using siamese neural networks. *International Journal of Electrical Power & Energy Systems* 104 (2019), 645–653.

[4] JP Eckmann, S Oliffson Kamphorst, D Ruelle, et al. 1995. Recurrence plots of dynamical systems. *World Scientific Series on Nonlinear Science Series A* 16 (1995), 441–446.

[5] Johann Faouzi and Hicham Janati. 2020. Pyts: A python package for time series classification. *Journal of Machine Learning Research* 21, 46 (2020), 1–6.

[6] Anthony Faustine and Lucas Pereira. 2020. Improved appliance classification in non-intrusive load monitoring using weighted wecurrence graph and convolutional neural networks. *Energies* 13, 13 (2020), 3374.

[7] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. 2019. Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery* 33, 4 (2019), 917–963.

[8] R Gopinath, Mukesh Kumar, C Prakash Chandra Joshua, and Kota Srinivas. 2020. Energy management using non-intrusive load monitoring techniques–State-of-the-art and future research directions. *Sustainable Cities and Society* (2020), 102411.

[9] George William Hart. 1992. Non-intrusive appliance load monitoring. *Proc. IEEE* 80, 12 (1992), 1870–1891.

[10] Nima Hatami, Yann Gavet, and Johan Debayle. 2018. Classification of time-series images using deep convolutional neural networks. In *Tenth international conference on machine vision (ICMV 2017)*, Vol. 10696. International Society for Optics and Photonics, 106960Y.

[11] Dongning Jia, Xianqing Huang, Zehua Du, Ruixue Li, and Kexin Li. 2020. Identification of electrical equipment based on two-dimensional time series characteristics of power. *MS&E* 768, 6 (2020), 062019.

[12] Jack Kelly and William Knottenbelt. 2015. Neural nilm: Deep neural networks applied to energy disaggregation. In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*. 55–64.

[13] Jack Kelly and William Knottenbelt. 2015. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Scientific data* 2, 1 (2015), 1–14.

[14] Christoph Klemenjak, Christoph Kovatsch, Manuel Herold, and Wilfried Elmenreich. 2020. A synthetic energy dataset for non-intrusive load monitoring in households. *Scientific Data* 7, 1 (2020), 1–17.

[15] Lamprini Kyrkou, Christoforos Nalmpantis, and Dimitris Vrakas. 2019. Imaging time-series for NILM. In *International Conference on Engineering Applications of Neural Networks*. Springer, 188–196.

[16] Mohammad Mottahedi and Somayeh Asadi. 2016. Non-intrusive load monitoring using imaging time series and convolutional neural networks.

[17] David Murray, Lina Stankovic, and Vladimir Stankovic. 2017. An electrical load measurements dataset of United Kingdom households from a two-year longitudinal study. *Scientific data* 4, 1 (2017), 1–12.

[18] Andreas Reinhardt and Christoph Klemenjak. 2020. How does load disaggregation performance depend on data characteristics? insights from a benchmarking study. In *Proceedings of the Eleventh ACM International Conference on Future Energy Systems*. 167–177.

[19] Zhiguang Wang and Tim Oates. 2015. Encoding time series as images for visual inspection and classification using tiled convolutional neural networks. In *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*.

[20] Dongsheng Yang, Xiaoting Gao, Liang Kong, Yongheng Pang, and Bowen Zhou. 2020. An event-driven convolutional neural architecture for non-intrusive load monitoring of residential appliance. *IEEE Transactions on Consumer Electronics* 66, 2 (2020), 173–182.

[21] Chaoyun Zhang, Mingjun Zhong, Zongzuo Wang, Nigel Goddard, and Charles Sutton. 2018. Sequence-to-Point learning with neural networks for non-intrusive load monitoring. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*.