Efficient Deep Learning Techniques for Water Disaggregation

Nikolaos Virtsionis Gkalinikis *
School of Informatics
Aristotle University of Thessaloniki
Thessaloniki 54124, Greece
nvirtsion@csd.auth.gr

Dimitris Vrakas
School of Informatics
Aristotle University of Thessaloniki
Thessaloniki 54124, Greece
dvrakas@csd.auth.gr

***Abstract*—The goal of water disaggregation is to specify the consumption of the individual water fixtures using only the aggregate consumption measurements of a single house meter. This task belongs to the family of blind source separation methods, where neural networks are proven to excel. Despite that recent breakthroughs in Internet of Things (IoT) led to the development of meters able to measure consumption in frequency higher than 1 point per month or day, the use of deep learning in this domain is infeasible due to the lack of publicly available data sets. To overcome these challenges a set of deep learning techniques is introduced to achieve great results in cases where limited data is available. In addition, the method of domain knowledge transfer is explored in an effort to take advantage of models that were previously trained in energy data as feature extractors to address the water separation problem.**

***Index Terms*—water disaggregation, non intrusive load monitoring, artificial neural networks, knowledge transfer, transfer learning**

I. Introduction

The concept of blind source separation [1] refers to the decomposition of a signal down to its individual sources, without any more information but the actual signal. The term disaggregation is mainly used in the domain of non-intrusive load monitoring (NILM) [2], where the given signal is the total power consumption of an installation and the targets are the individual consumptions of the electrical appliances. NILM constitutes an important part of intelligent home energy management systems (HEMS) [3], providing numerous insights from how users operate their electrical appliances to possible hardware functionality inconsistencies. Eventually, the energy management could lead to the optimization of the electricity consumption alongside the decrease of household bills.

Water disaggregation could also be viewed from a similar stand point, as a blind source separation task. Instead of electricity data, water consumptions are processed, with the total consumption signal being decomposed to the individual water fixtures estimates. Successful disaggregation in residential installations could provide tools to properly monitor and manage the available potable water resources. Hence, the water demand could be controlled and possible waste of water could be avoided. The latter is of great interest, given the fact that in the recent years both the available resources and the weather conditions are affected by climate change [4]. Furthermore, an effective design of water disaggregation system in residential installations could also provide more insights on user behavior and habits. That could facilitate a more proactive approach to water demand management especially in geographical areas or systems, where limited resources are available [5], [6].

The rise of IoT towards cheap and easy to deploy solutions resulted in an infiltration of IoT monitoring systems to every day life activities. Given the fact that many types of services are being produced from house [7] to health systems, the time has come for efficient water consumption management to be also introduced. Thus, water management could be a vital part of the smart home ecosystems, which are currently being deployed to residential and small commercial sector installations [8].

NILM solutions are usually applied on low frequency measurements from 1 sec up to a few minutes, even though electrical data can be measured in a great range of frequencies from sub Hz to MHz. The main reasons are the following. Firstly, most of the commercial smart meters draw measurements in low frequencies. Secondly, in these low frequencies there is a good trade off between system performance and cost. As a result, the majority of the publicly available data sets have been recorded at low frequency [9].

On the other hand, in the sector of water disaggregation the developed market solutions usually make use of measurements that are gathered in monthly and/or daily periods. This is either due to the lack of proper smart meter hardware and/or the fact that there is not a clear direction of how and why higher frequency data is of use. It is obvious that in such low data resolution no useful information could be extracted regarding the habits of the residents in day-to-day basis and water disaggregation is intractable. Although, there are research works where very high frequency data (up to kHz) coming from prototype smart meters are used the design of such a practical system could be very costly. Current approaches tend to use data for water disaggregation in a range of frequencies similar to NILM, considering both the performance and the overall application costs.

II. Contributions

The contribution of the current paper to the field research of water disaggregation could be summarized in three key points. To begin with, as a proof of concept by adjusting a set of lightweight yet powerful deep learning architectures and techniques to solve the current task. Deep learning is the current go-to approach in blind source separation tasks, but it is not used as much for water disaggregation, due to the lack of data. In addition, with the introduction of the concept of domain transfer learning in water disaggregation field of study. Transfer learning is a technique commonly used to Computer Vision in situations where limited data is available. Lastly, with a set of baseline results for further research usage.

This article is structured as follows. In section 3 there is a brief presentation of the related work in the field of water disaggregation. Additionally, there is a short presentation of the key deep learning techniques that are used in the domain of energy disaggregation. Section 4, contains an introduction to the basic types of layers used in deep learning. Furthermore the concept of transfer learning in machine learning is described. Section 5, presents a description of the followed experiment methodology and the datasets that were used alongside with a thorough presentation of the proposed architectures. The produced results are presented in section 6. In the final section, conclusions and future work suggestions are discussed.

III. Related Work

The increasing deployment of smart meters in cities across the world constitutes water disaggregation as a new interesting research area. Pressure-based sensors have been designed for installation on water fixtures to help identify activity and estimate the corresponding consumption for individual household devices [10]. By utilizing both occupancy sensors and whole house water flow meter data [11], categorization of the aggregated consumption at the fine-grained device level could be achieved.

Although such methods are capable of high accuracy results, they depend on high-sampling-rate sensing data (as high as 1 Khz) to capture the characteristic open/close signatures of devices. A HMM (Hidden Markov Model) based approach was developed in [12] for separating low-sampling-rate (1/900 Hz) data, while [13], [14] proposed a hybrid combination of HMM and DTW (Dynamic Time Warping) to automate the categorisation of residential water end use events and estimate the consumption of each device. However, a HMM based structure inherently restricts its ability to infer consumption for devices that are active in parallel.

There is a lack of models designed for disaggregating low-sampling-rate water consumption. The existing HMM based method analyses the activities with interval based consumption; however, it has limited ability to estimate the consumption for parallel devices due to its inherent serial structure.

From the machine learning perspective two approaches are commonly used to tackle NILM and disaggregation; regression and multi-label classification. In regression approach, the power consumption of a single appliance is estimated [15]–[17]. Thus, one model per device is created. In multi-label classification approach, the model identifies operating states of various devices. Hence, one model learns to disaggregate a set of devices. Recently published research showed that a multi-label approach achieves descent results [18]–[20].

IV. Introduction to Neural Networks

The recent advances in hardware solutions and especially the breakthroughs in Graphics processing units (GPUs) technology provided the necessary computational power to design and train neural network architectures. Neural networks constitute a natural solution in machine learning problems due to the fact that they produce state of the art results in a wide range of problems, even though researchers do not accurately now why.

A neural network is made of a number of neural layers. The basic types of these layers are; the fully connected, the convolutional and the recurrent [21]. In this article we evaluate architectures with a combination of these layers.

A. Neural Networks in a nutshell

Essentially, an artificial neural network (ANN) is a directed graph, where the nodes are artificial neurons and the edges allow information from one neuron to pass to another neuron (or the same neuron in a future time step). Neurons are typically arranged into layers such that each neuron in layer l connects to every neuron in layer l + 1. Connections are weighted and it is through modification of these weights that ANNs learn. ANNs have an input layer and an output layer. Any layers in between are called hidden layers. Each artificial neuron calculates a weighted sum of its inputs, adds a bias and passes this sum through an activation function. Multiple nonlinear hidden layers can be used to re-represent the input data (hopefully by learning a hierarchy of feature detectors).

Neural network architectures learn by example. In order to train a neural network in a supervised learning manner [22] the network's output given a specific input is compared with the ground truth and the error is calculated. Then, the weights are modified in the direction which should reduce the error with an optimization algorithm based on gradient descent [23]. Depending of the type of neural network the optimization algorithm may differ, but the basic principals are the same.

B. Transfer Learning

In order to reduce the computational costs a popular technique on is knowledge transfer known as also transfer learning [24]. This method was firstly applied on Computer Vision problems with very good results [25]. It should be noted that transfer learning was also used in NILM research with some success in [26], [27] in order to reduce computational resources.

The general idea of transfer learning is to use the features extracted from training on one domain to a completely different one. After fine tuning the parameters of the last layers of the network the model is ready to be used on the new domain. This framework usually applies to problems, where the domain data is limited and/or to applications where low computational cost is the main concern. Usually, transfer learning produce good results faster than the complete retraining of the network.

V. Experiments

All the experiments were designed using Torch-NILM [28] and were executed on the same machine with a Titan Xp GPU. Torch-NILM is the first Pytorch oriented framework that is used to build and run NILM experiments easily. It contains numerous build in models alongside known pre-processing methods popular in NILM problems. To use the framework minor adjustments had to be made. The complete code is provided in the official Torch-NILM repository github.com/Virtsionis/torch-nilm.

The Sliding Window approach 1 proposed by [29] was used for all the experiments alongside data standardization pre-processing method. In the Sliding Window method the time series is divided into chunks with predefined length. Given a chunk the model predicts the last point of it. This method is suitable for online disaggregation in the sense that the model estimates the last point of the window. For all the models under evaluation the input window was set to 100 points.

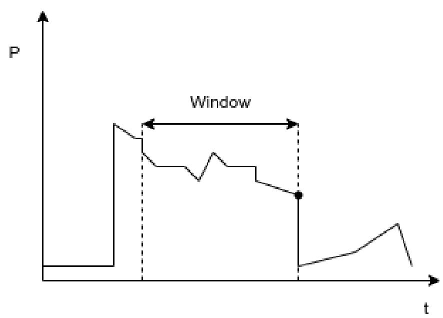

Fig. 1. Sliding window approach.

A. Methodology

The experiments in this article are divided in two categories. The first one involves training and inference on a dataset with measurements from pure water end uses. Hence the ability of the models to successfully learn from water data is measured. In the second category a transfer learning schema was executed in order to explore the transfer-ability between the two domains of NILM and water disaggregation.

The idea of transfer learning is that a pre-trained model can be used as a feature extraction for a different task. At first the models are trained on the electrical measurements of washing machine appliance data coming from a popular dataset in NILM. Then, fine-tuning on the last layers of the network was applied on the target device and inference was performed. It should be noted that the fine-tuning and the inference were executed on water data.

B. Neural Network Topologies

In the current work three known NILM topologies were used; NFED [30], SAED [15] and WGRU [29]. Alongside those architectures a simpler and lighter recurrent architecture called SimpleGru was developed. The NFED and WGRU models produce state-of-the-art results on NILM tasks, whereas SAED shows strong generalization capabilities with small number of parameters and fast training and inference speeds. The architectures WGRU, SAED and SimpleGru are based on Bidirectional GRU layers, an efficient version of the LSTM network [31]. WGRU has two of those layers in series whereas the other two have one. Between the layers dropout is used [32].

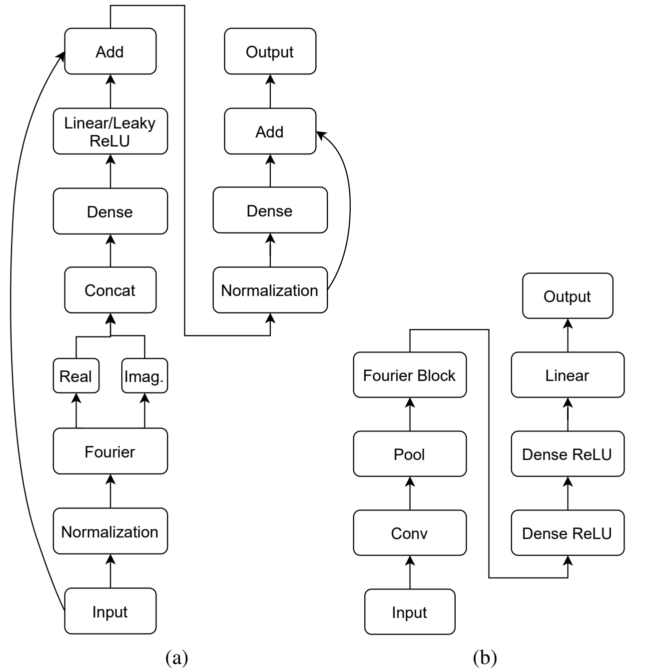

Fig. 2. The NFED neural network. (**a**) Fourier block; (**b**) NFED architecture.

NFED consists mostly of fully connected and normalization layers along with residual connections. The basic component is called the fourier block which contains a fourier transformation at the top.

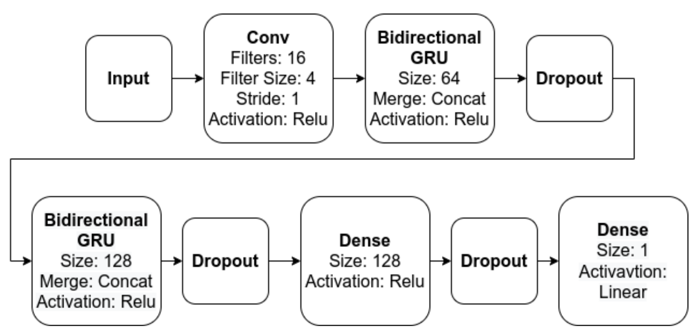

Fig. 3. Architecture of WGRU.

Comparing to WGRU, SAED is many times faster and smaller in size. In addition, the SAED model contains an Attention layer before the GRU. The intuition is that the attention mechanism after the convolution aids the model to focus on the most important features of the input sequences.

A summary of the architecture size and parameters number is depicted on table I.

TABLE I
PROPERTIES OF THE TESTED ARCHITECTURES. NUMBER OF PARAMETERS, SIZE OF THE MODEL AND TRAINING SPEED (GPU TIME).

Architecture	**Parameters**	**Size (MB)**	**Train (it/s)**
NFED	900 K	3.6	30.92
SAED	59.9 K	0.240	37.51
SimpleGru	39.9 K	0.160	39.41
WGRU	698 K	2.794	16.94

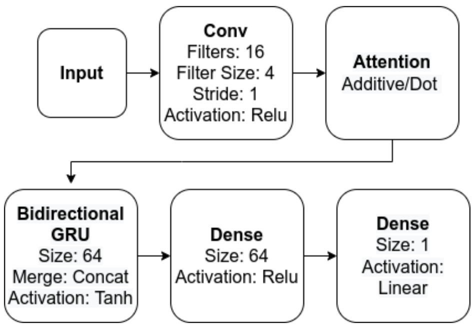

Fig. 4. Architecture of SAED.

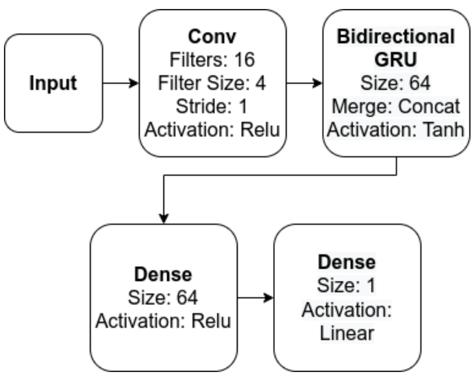

Fig. 5. Architecture of SimpleGru.

C. Data

To train and conduct the experiments, water and electricity data with granularity 1-10 seconds is necessary. As the electricity data, UKDALE [33] was used. This dataset contains appliance measurements in 6 seconds and total consumption measurements in 1 second time period.

In general, not many datasets for water disaggregation are publicly available. In this work data from WEUSEDTO dataset was utilized [34], [35] which contains measurements in 1 second sampling period from a single resindent appartement in Naples. The dataset refers to 1 year of monitoring between 2019 and 2020 (March to November 2019 and July to October 2020). The water data was down sampled to 6 seconds to speed up the training and also match the electricity data that was used for transfer learning.

The dataset contains ground truth measurements for 7 kind of water appliances. In this work data only for Washbasin, Bidet, Kitchen faucet, Shower and Washing Machine were used. Due to the lack of good aggregate measurements, the aggregate water consumption was artificially composed by adding the separate water end uses. From the final dataset 10 months were used for training and the rest were used for testing.

D. Metrics

In order to evaluate and compare the models, some well known NILM metrics were calculated; F1 score, Relative Error in Total Volume (REVol) and Mean Absolute Error (MAE). The ability of model to detect on/off states is evaluated with F1 score. As seen in eq. 3, F1 score is computed as the harmonic mean of Precision and Recall, presented in eq. 4 and 5. Precision measures the ratio of the actual true positives (TP) versus the total predicted positives. In addition, Recall is the percentage of TP versus the actual positives.

On the other hand, MAE (measured in mili litres) and REVol (dimensionless) evaluate the capability of the models to estimate the actual water consumption of the device. MAE and REVol are given in equations 1 and 2, where V' is the predicted total volume, V is the true value of total volume, T is the length of the predicted sequence, yt' and yt the estimated and the true water flow values at moment t correspondingly.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1 = 2\frac{Precision * Recall}{Precision + Recall} \tag{3}$$

$$REVol = \frac{|V' - V|}{max(V', V)} \tag{4}$$

$$MAE = \frac{1}{T}\sum |y'_t - y_t| \tag{5}$$

VI. RESULTS

The models were compared on two experiments. At first, in a schema where training and inference were executed on the same water dataset. Next, on a knowledge transfer scenario, where the models were pre-trained on Washing Machine end uses of house 1 of UKDALE and then they were fine tuned with water data. In tables II and III the metrics F1, MAE and REVol are presented. An example of the output of the models for the second experiment is depicted in figures 6 and 7.

For the first experiment it is obvious that the WGRU is the clear winner regarding the F1 measure and the MAE, for 3 to 5 appliances and next is the SimpleGru with two wins. The greatest differences occurred on Washing Machine and Shower appliances. In terms of REVol the SAED shows the best performance with 4 wins and then follows the NFED.

The transfer learning technique that was used in the second experiment affects the performance of certain models. Specifically, SimpleGru performs better in terms of F1 score for the appliances Washing Machine, Shower and Kitchen Faucet with up to 7% boost. On the other hand, SAED has a drop in performance for Washing Machine and Shower and a raise of 13% for the Kitchen Faucet. NFED and WGRU do not show differences in performance between the two experiments. In terms of MAE and REVol the results among the experiments are very close.

Considering the results and the architecture properties showed in tables I, II and III, SimpleGru with knowledge transfer technique could be the best choice for the current data because it combines good performance with low computational cost.

TABLE II
PERFORMANCE COMPARISON FOR EXPERIMENT 1.

Appliance	Architecture	F1	REVol	MAE
Bidet	NFED	0.3	0.3	0.062
	SAED	0.29	**0.06**	0.063
	SimpleGru	**0.34**	0.31	**0.059**
	WGRU	0.33	0.31	0.057
KitchenFaucet	NFED	0.22	0.34	0.17
	SAED	0.38	**0.18**	0.13
	SimpleGru	**0.41**	0.29	**0.11**
	WGRU	0.25	0.42	0.12
Shower	NFED	0.76	0.16	0.117
	SAED	0.77	**0.13**	0.083
	SimpleGru	0.72	0.31	0.078
	WGRU	**0.79**	0.14	**0.073**
Washbashin	NFED	0.52	0.126	0.11
	SAED	0.48	**0.078**	0.12
	SimpleGru	0.54	0.119	0.11
	WGRU	**0.56**	0.126	**0.09**
Washing Machine	NFED	0.66	0.268	0.071
	SAED	0.66	0.194	0.061
	SimpleGru	0.74	0.149	0.048
	WGRU	**0.79**	**0.097**	**0.047**

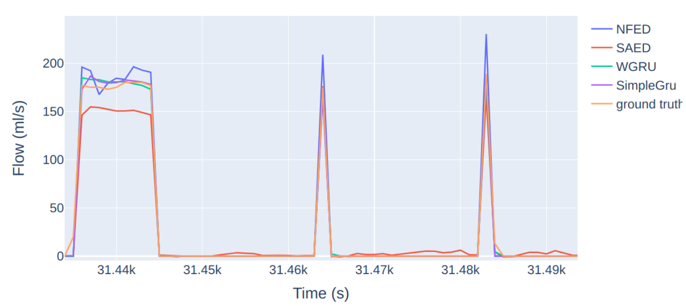

Fig. 6. Model output versus ground truth for Washing Machine, for Experiment 2.

VII. CONCLUSIONS AND PROPOSALS FOR FUTURE WORK

The produced results in the current work highlight a numerous of key points regarding the problem of water disaggregation. To start with, the similarities between water disaggregation and NILM emphasize the fact that the first can be formulated as a blind source type of task. Secondly, even in situations with limited data, deep learning techniques can be applied to water disaggregation by producing descent results. Furthermore, the use of transfer learning concept across different disaggregation problems shows potential. The fact that the models performed the same or even better in some cases of the second experiment indicates that pre-trained models from the NILM domain could be easily adapted and used as feature extractors on water disaggregation problems.

Regarding future work the following proposals can be made. Firstly, a more extensive investigation of the knowledge transfer technique with different scenarios of available data could be insightful. Next, more pre-processing techniques similar to the sliding window approach could be tested in order to decide which one fits better the current problem. Finally, more neural network architectures and state-of-the-art models from NILM could be used.

TABLE III
PERFORMANCE COMPARISON FOR EXPERIMENT 2.

Appliance	Architecture	F1	REVol	MAE
Bidet	NFED	0.32	1.14	0.114
	SAED	0.29	**0.23**	0.072
	SimpleGru	0.32	0.337	0.063
	WGRU	**0.33**	0.31	**0.057**
KitchenFaucet	NFED	0.22	0.336	0.168
	SAED	**0.43**	0.234	**0.101**
	SimpleGru	**0.43**	0.29	0.106
	WGRU	0.42	**0.146**	0.116
Shower	NFED	0.78	0.18	0.118
	SAED	0.75	**0.11**	0.10
	SimpleGru	0.77	0.31	0.075
	WGRU	**0.79**	0.14	**0.073**
Washbashin	NFED	0.52	0.126	0.104
	SAED	0.47	0.135	0.124
	SimpleGru	0.54	**0.119**	0.115
	WGRU	**0.56**	0.126	**0.097**
Washing Machine	NFED	0.66	0.268	0.071
	SAED	0.58	0.154	0.085
	SimpleGru	0.77	0.149	**0.041**
	WGRU	0.79	**0.097**	0.047

Fig. 7. Model output versus ground truth for Shower, for Experiment 2.

Acknowledgment

This research was carried out as part of the project "iWet: Intelligent IoT System for Quantitative and Qualitative Measurements for water distribution networks" (Project code: KMP6-0072250) under the framework of the Action "Investment Plans of Innovation" of the Operational Program "Central Macedonia 2014 2020", that is co-funded by the European Regional Development Fund and Greece.

We gratefully acknowledge the support of the NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

References

[1] M. Pal, R. Roy, J. Basu, and M. S. Bepari, "Blind source separation: A review and analysis," in *2013 International Conference Oriental COCOSDA held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE)*, 2013, pp. 1–5.

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

[3] B. Naghibi and S. Deilami, "Non-intrusive load monitoring and supplementary techniques for home energy management," in *2014 Australasian Universities Power Engineering Conference (AUPEC)*. IEEE, 2014, pp. 1–5.

[4] J. Laitinen, R. Antikainen, J. J. Hukka, and T. S. Katko, "Water supply and sanitation in a green economy society: The case of finland," *Public Works Management & Policy*, vol. 25, no. 1, pp. 33–50, 2020. [Online]. Available: https://doi.org/10.1177/1087724X19847211

[5] Y. Shan, L. Yang, K. Perren, and Y. Zhang, "Household water consumption: Insight from a survey in greece and poland," *Procedia Engineering*, vol. 119, pp. 1409–1418, 12 2015.

[6] M. Salehi, "Global water shortage and potable water safety; today's concern and tomorrow's crisis," *Environment International*, vol. 158, p. 106936, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0160412021005614

[7] B. Mahapatra and A. Nayyar, "Home energy management system (hems): concept, architecture, infrastructure, challenges and energy management schemes," *Energy Systems*, pp. 1–27, 2019.

[8] K. C. Armel, A. Gupta, G. Shrimali, and A. Albert, "Is disaggregation the holy grail of energy efficiency? the case of electricity," *Energy Policy*, vol. 52, pp. 213–234, 2013.

[9] H. K. Iqbal, F. H. Malik, A. Muhammad, M. A. Qureshi, M. N. Abbasi, and A. R. Chishti, "A critical review of state-of-the-art non-intrusive load monitoring datasets," *Electric Power Systems Research*, vol. 192, p. 106921, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779620307197

[10] J. E. Froehlich, E. Larson, T. Campbell, C. Haggerty, J. Fogarty, and S. N. Patel, "Hydrosense: Infrastructure-mediated single-point sensing of whole-home water activity," in *Proceedings of the 11th International Conference on Ubiquitous Computing*, ser. UbiComp '09. New York, NY, USA: ACM, 2009, pp. 235–244. [Online]. Available: http://doi.acm.org/10.1145/1620545.1620581

[11] V. Srinivasan, J. Stankovic, and K. Whitehouse, "Watersense: water flow disaggregation using motion sensors," 11 2011, pp. 19–24.

[12] E. Kermany, H. Mazzawi, D. Baras, Y. Naveh, and H. Michaelis, "Analysis of advanced meter infrastructure data of water consumption in apartment buildings," in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '13. New York, NY, USA: Association for Computing Machinery, 2013, p. 1159–1167. [Online]. Available: https://doi.org/10.1145/2487575.2488193

[13] K. Nguyen, H. Zhang, and R. Stewart, "Development of an intelligent model to categorise residential water end use events," *Journal of Hydro-environment Research*, vol. 7, 09 2013.

[14] K. Nguyen, R. Stewart, and H. Zhang, "An intelligent pattern recognition model to automate the categorisation of residential water end-use events," *Environmental Modelling and Software*, vol. 47, pp. 108–127, 09 2013.

[15] N. Virtsionis-Gkalinikis, C. Nalmpantis, and D. Vrakas, "Saed: self-attentive energy disaggregation," *Machine Learning*, pp. 1–20, 2021.

[16] C. Athanasiadis, D. Doukas, T. Papadopoulos, and A. Chrysopoulos, "A scalable real-time non-intrusive load monitoring system for the estimation of household appliance power consumption," *Energies*, vol. 14, no. 3, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/3/767

[17] C. L. Athanasiadis, D. I. Doukas, T. A. Papadopoulos, and G. A. Barzegkar-Ntovom, "Real-time non-intrusive load monitoring: A machine-learning approach for home appliance identification," in *2021 IEEE Madrid PowerTech*. IEEE, 2021, pp. 1–6.

[18] S. M. Tabatabaei, S. Dick, and W. Xu, "Toward non-intrusive load monitoring via multi-label classification," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, pp. 26–40, 2016.

[19] C. Nalmpantis and D. Vrakas, "On time series representations for multi-label nilm," *NEURAL COMPUTING & APPLICATIONS*, 2020.

[20] C. L. Athanasiadis, T. A. Papadopoulos, and D. I. Doukas, "Real-time non-intrusive load monitoring: A light-weight and scalable approach," *Energy and Buildings*, vol. 253, p. 111523, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378778821008070

[21] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, 04 2014.

[22] S. Trivedi, B. K. Dhaliwal, and G. Singh, "A review paper on a comparative study of supervised learning approaches," in *2021 International Conference on Computing Sciences (ICCS)*, 2021, pp. 95–100.

[23] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.

[24] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. PP, pp. 1–34, 07 2020.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 06 2016, pp. 770–778.

[26] 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, Mar. 2020.

[27] S. Houidi, D. Fourer, F. Auger, H. B. A. Sethom, and L. Miègeville, "Comparative evaluation of non-intrusive load monitoring methods using relevant features and transfer learning," *Energies*, vol. 14, no. 9, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/9/2726

[28] N. Virtsionis Gkalinikis, C. Nalmpantis, and D. Vrakas, "Torch-nilm: An effective deep learning toolkit for non-intrusive load monitoring in pytorch," *Energies*, vol. 15, no. 7, 2022. [Online]. Available: https://www.mdpi.com/1996-1073/15/7/2647

[29] O. Krystalakos, C. Nalmpantis, and D. Vrakas, "Sliding window approach for online energy disaggregation using artificial neural networks," in *Proceedings of the 10th Hellenic Conference on Artificial Intelligence*, 2018, pp. 1–6.

[30] C. Nalmpantis, N. Virtsionis Gkalinikis, and D. Vrakas, "Neural fourier energy disaggregation," *Sensors*, vol. 22, no. 2, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/2/473

[31] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," in *NIPS 2014 Workshop on Deep Learning, December 2014*, 2014.

[32] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[33] K. Jack and K. William, "The uk-dale dataset domestic appliance-level electricity demand and whole-house demand from five uk homes," *Sci. Data*, vol. 2, p. 150007, 2015.

[34] A. Di Mauro, A. Di Nardo, G. Santonastaso, and S. Venticinque, "An iot system for monitoring and data collection of residential water end-use consumption," 08 2019.

[35] ——, "Development of an iot system for the generation of a database of residential water end-use consumption time series," *Environmental Sciences Proceedings*, vol. 2, p. 20, 08 2020.