Fully-convolutional denoising auto-encoders for NILM in large non-residential buildings

Diego García-Pérez, Daniel Pérez-López, Ignacio Díaz-Blanco, Ana González-Muñiz, Manuel Domínguez-González, Abel A. Cuadrado-Vega

***Abstract*—Great concern regarding energy efficiency has led the research community to develop approaches which enhance the energy awareness by means of insightful representations. An example of intuitive energy representation is the parts-based representation provided by Non-Intrusive Load Monitoring (NILM) techniques which decompose non-measured individual loads from a single total measurement of the installation, resulting in more detailed information about how the energy is spent along the electrical system. Although there are previous works that have achieved important results on NILM, the majority of the NILM systems were only validated in residential buildings, leaving a niche for the study of energy disaggregation in non-residential buildings, which present a specific behavior. In this paper, we suggest a novel fully-convolutional denoising auto-encoder architecture (FCN-dAE) as a convenient NILM system for large non-residential buildings, and it is compared, in terms of particular aspects of large buildings, to previous denoising auto-encoder approaches (dAE) using real electrical consumption from a hospital facility. Furthermore, by means of three use cases, we show that our approach provides extra helpful funcionalities for energy management tasks in large buildings, such as meter replacement, gap filling or novelty detection.**

***Index Terms*—Energy efficiency, Building energy consumption, NILM, Energy disaggregation, Denoising auto-encoders**

I. Introduction

ENERGY disaggregation, also known as *Non-Intrusive Load Monitoring* (NILM), is the computational process of extracting individual consumptions of an electrical installation from a single total measurement. As it was shown in several studies [1]–[4], parts-based representations of the total consumption provide users with insightful feedback, which encourages them to improve their energy savings. This energy awareness makes NILM systems valuable tools that allow users to detect anomalous individual consumptions, obtain per-appliance forecasts and adjust their energy spending plans better.

In 1992, Hart [5] addressed for the first time the problem of energy disaggregation, defining each connected appliance as a *finite state machine* (FSM) whose transitions are associated with specific electrical signatures or features manually extracted from the total consumption. Hart's work led to approaches based on *Hidden Markov Models* (HMM) [6], and more specifically on *Factorial Hidden Markov Models* (FHMM) [7]–[12], in which each individual device is modeled by a set of hidden states and the total measured consumption is the sum of the demands associated with the active states per device. Although models based on hidden states achieve accurate results for "on/off" devices, they have several limitations in modeling multi-state appliances and consumptions that are dependent on human behavior. In addition to this drawback, a strong prior knowledge is needed because the feature extraction and the number of states per appliance are still manually defined.

The revival of neural networks in the last decade arose new deep convolutional architectures [13], [14] capable of automatically extracting relevant features for tasks such as computer vision [14]–[16] and speech recognition [17], [18]. These achievements encouraged several works [19], [20] to translate these novel deep neural networks (DNN) approaches into the field of energy disaggregation. Thus, these new perspectives brought novel NILM techniques that outperform the results of previous HMM-based methods using only the raw aggregated active power as input, without manually defining any feature. Subsequent DNN-based works can be grouped in those which continue the paradigm of modeling hidden states, such as previous HMM-based, by means of the use of *recurrent neural networks* (RNN) [21]–[25], and in those which address the NILM as a *denoising* problem [21], [26]–[29]. The latter ones tackle the energy disaggregation as a filtering problem where the total consumption is the noisy signal to be cleaned in order to obtain the individual loads, obtaining remarkable results using filter functions based on *denoising auto-encoders* (dAE) [30].

Most of the previous methods have been tested in residential buildings but only few of them have been applied to large non-residential buildings [31]–[33]. In recent years, interest in energy efficiency of large buildings has increased, since they constitute about 11% of the total energy consumption in developed countries [34]. Therefore, tools as NILM systems, that improve the energy awareness and help users manage the energy demand, could be convenient for developing energy management systems in large installations such as non-residential buildings [33], [35].

Applying NILM to non-residential buildings is not straightforward, since several aspects differentiate the energy disaggregation in non-residential buildings from the NILM systems in households [33], [36]. In this work, we suggest to address some of these differences of the NILM problem in large buildings by means of several architectures based on dAE. We use a similar dAE architecture as the one proposed in [19] as baseline and, then, we compare it with a dAE architecture

Diego García-Pérez, Ignacio Díaz-Blanco, Ana González-Muñiz and Abel A. Cuadrado-Vega are with the Electrical Engineering Departament, University of Oviedo, 33204 Gijón, Spain (email: diegogarcia@isa.uniovi.es).

Daniel Pérez-López and Manuel Domínguez-González are with the SUPPRESS Research Group, University of León, 24007 León, Spain.

completely made of 1D *convolutional layers* that we have called *fully-convolutional denoising auto-encoder* (FCN-dAE), using real electrical consumption data from a hospital.

In addition to mapping the total demand into individual loads, we also suggest that dAE-based NILM approaches applied to large buildings can be useful in other energy management aspects. We illustrate these extra functionalities by means of three real use cases with the data from the hospital. In the first one, we replace real meters by the NILM estimations in order to install the substituted meters in other areas of the installation, or during the maintenance of the meters. In the second one, our NILM methods fill gaps produced in the submetering data due to temporary malfunctions of the meter and, in the third use case, a method capable of finding novelties along the electrical network is presented.

The remaining sections of the article are organized as follows: in section II, the denoising NILM paradigm is formulated; in section III, the two proposed architectures are defined and the particularities of applying them to non-residential buildings are shown; in section IV, a detailed description of the data used in the experiments is presented; in section V, the evaluation metric chosen for comparing the models is defined; in section VI, the experimental set-up used in our studies is explained; and finally, in sections VII and VIII, the obtained results and the use cases are discussed.

II. NILM as a denoising problem

In addition to inferring the internal states of each load of the network, as in FHMM methods, energy disaggregation can also be tackled as a *denoising* problem using methods that are able to map a noisy total consumption $x(t)$ to a "clean" individual consumption $y_i(t)$. Note that $x(t)$ and $y_i(t)$ could be any electrical variable which gathers the total demand and the individual consumptions, such as active, reactive, apparent power, current, voltage, or a vector of all of them, but in this article, we only consider the active power as the electrical variable gathered in $x(t)$ and $y_i(t)$. Therefore, if the total aggregated active power consumption $x(t)$ is defined as the sum of all downstream loads $y_i(t)$ and the non-measured loads $e(t)$ (1), then each individual consumption of the installation $y_i(t)$ could be formulated as the total consumption minus the remaining demand (2).

$$x(t) = \sum_{i=1}^{N} y_i(t) + e(t) \tag{1}$$

$$y_i(t) = x(t) - e(t) - \sum_{l=1, l\neq i}^{N} y_l(t) \tag{2}$$

According to this, the sum of the remaining individual loads $y_l(t)$ plus the non-measured component $e(t)$ is the noise to be removed from the input $x(t)$. Therefore, energy disaggregation could be addressed by noise removal or denoising functions f_i capable of mapping a sequence of total consumption $\mathbf{x}$ into a sequence of the same length of the individual consumption $\mathbf{y}_i$.

$$f_i : \mathbf{x} \rightarrow \mathbf{y}_i \tag{3}$$

III. Denoising auto-encoder models as NILM systems in large buildings

Considering the previous paradigm of NILM as a denoising problem, in this section, we initially explain the well-known denoising auto-encoder approach (dAE) introduced in [19] and then, we present some of its particularities when it is applied to consumptions from large buildings, such as the amount of temporal context needed by the NILM models, in order to produce accurate results. Finally, we present our suggested denoising auto-encoder architecture, made completely of convolutional layers, and show why a fully-convolutional approach results in a more convenient approach for large buildings than the vanilla denoising auto-encoders.

A. Vanilla denoising Auto-Encoders (dAE)

Denoising Auto-Encoders [30] were initially defined as unsupervised feature extractors capable of learning meaningful representations from high-dimensional data, even when the input data are partially occluded or corrupt. This idea fits perfectly with the denoising perspective of the energy disaggregation, since the target individual consumption is mixed with the rest of the individual loads in the total consumption and it has to be recognized by the NILM model.

The main base of dAE is an auto-encoder architecture where the input is encoded into a compact representation, typically with less dimensions than the input, called *bottleneck* from which it tries to reconstruct an output as close as posible to the input. In the specific case of electrical power demand time series, the auto-encoder is usually trained with fixed duration segments $\mathbf{x}^{(i)}$ of the input consumption, which are mapped into the bottleneck by the *encoder* e_θ and then reconstructed in the output $\hat{\mathbf{x}}^{(i)}$ by the *decoder* d_ϕ. During training, the weights θ of the encoder and the weights ϕ of the decoder are optimized by minimizing the *mean squared error*, known as *reconstruction error*, for all the training samples i:

$$\begin{aligned} \theta^*, \phi^* &= \arg\min_{\theta,\phi} \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}\big(\mathbf{x}^{(i)}, \hat{\mathbf{x}}^{(i)}\big) \\ &= \arg\min_{\theta,\phi} \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}\Big(\mathbf{x}^{(i)}, d_\phi\big(e_\theta\big(\mathbf{x}^{(i)}\big)\big)\Big) \end{aligned} \tag{4}$$

where $\mathcal{L}$ is the squared error $\mathcal{L}\big(\mathbf{x}^{(i)}, \hat{\mathbf{x}}^{(i)}\big) = ||\mathbf{x}^{(i)} - \hat{\mathbf{x}}^{(i)}||^2$ between the input and its reconstruction.

The dAEs are a special case of auto-encoders in which the network is trained to remove the noise from the input. Thus, a dAE model is trained to remove the noise from the input segment $\mathbf{x}^{(i)}$, reconstructing an estimation $\hat{\mathbf{y}}^{(i)}$ of the filtered input segment without noise $\mathbf{y}^{(i)}$. As in the case of standard auto-encoders, the weights of dAE models are optimized by minimizing the reconstruction error:

$$\begin{aligned} \theta^*, \phi^* &= \arg\min_{\theta,\phi} \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}\big(\mathbf{y}^{(i)}, \hat{\mathbf{y}}^{(i)}\big) \\ &= \arg\min_{\theta,\phi} \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}\Big(\mathbf{y}^{(i)}, d_\phi\big(e_\theta\big(\mathbf{x}^{(i)}\big)\big)\Big) \end{aligned} \tag{5}$$

Note that the reconstruction error is computed between the output of the model and the input segment without noise, which drives the model towards reconstructions different from the model input. From the denoising NILM perspective defined in (3), dAEs could learn the functions f_j which filter the total aggregated demand in order to estimate the individual consumption:

$$\hat{\mathbf{y}}_j = f_j(\mathbf{x}) = d_{\phi,j}\big(e_{\theta,j}(\mathbf{x})\big) \tag{6}$$

where f_j is a trained dAE which estimates the individual consumption under study $\mathbf{y}_j$. Thus, a complete NILM system for an installation with m individual loads would include a set of filter functions $F = \{f_1, f_2, \ldots, f_m\}$, one per individual consumption to be disaggregated.

Regarding the topology of functions f_j, we initiate our study with the architecture shown in Fig. 1a, which is a dAE architecture similar to the proposed in [19]. The encoder subnetwork maps the input segment $\mathbf{x}$ into the intermediate representation by means of a 1D convolutional layer followed by two fully-connected layers. The decoder architecture generates the $\hat{\mathbf{y}}_i$ estimated individual consumption after passing the resulting intermediate representation through two fully-connected layers and a final 1D convolutional layer with only one filter which returns the individual output consumption. All the user-defined parameters of the architecture, such as the number of units in the dense layers, number of kernels in the convolutional layers or the kernel size, were experimentally set at those which most reduce the reconstruction error after a process of trial and error.

B. Disaggregation process and the input window size selection

Once the dAEs are trained, in order to compute the estimated individual consumption $\hat{y}_i(t)$, its corresponding dAE model f_i processes a segment of the total consumption $\mathbf{x}$. These segments are the result of sliding a window of size L along the total consumption under study, with a stride M between two consecutive windows. The parameters of the windowing L and M not only determine the configuration of the architecture shown in Fig. 1a, but also affect the accuracy of the resulting disaggregation, since they define the *temporal context* used by the model to estimate the individual load. Thus, in the case of the stride, if $L > M$, there is an overlapping between windows which leads the dAE model to produce several estimations for the same time instant. This issue is typically solved in previous works [19], [27] by computing the average of all the obtained predictions. Averaging all the estimations, we noticed that disaggregation performance is not sensitive to the stride when the percent of overlapping between windows is large enough. In contrast to the stride, the size of the input windows L conditions significantly the disaggregations, since the models trained with too large input windows tend to produce poor reconstructions of the short-duration loads, whereas if the input windows are not long enough and only gather a small part of the appliance signature, the resulting model performs an inaccurate disaggregation because it lacks temporal information in the input. In addition, an increase of the window size also implies an increment in the number of trainable weights, which raises the risk of overfitting.

In residential NILM, the window size is typically set appliance-by-appliance according to its maximum active period, but in the case of non-residential buildings, where most of the individual loads remain continuously active, setting an appropriate window size is not straightforward. Therefore, an input window size invariant architecture, that prevents us from manually searching the best window size for each individual consumption to be disaggregated, would be highly desirable in the non-residential NILM approaches.

C. Fully-convolutional denoising auto-encoder (FCN-dAE)

Motivated by the problem of selecting a correct input window size, a convenient design would be a model which could reproduce long-term (low-frequency patterns) and short-term (high-frequency patterns) behaviors when its input context is long enough. In DNN literature, sequential data, such as electrical time series, were typically modeled using *Recurrent Neural Networks* (RNN) specially by means of *Long-Short Term Memory* (LSTM) [37] layers, but recent works [38] have demonstrated that long short-term behaviors could also be learned by *fully-convolutional neural network* (FCN) [39] approaches, that have proven to be more invariant to the input size than traditional convolutional neural networks (CNN) [40] in image segmentation problems. These new FCN-based approaches for time series modeling have the advantage of being more efficient in trainable weights and they present more stable gradients which makes them easier to train.

With this in mind, we suggest a *fully-convolutional dAE* which avoids the use of fully-connected layers in the conventional dAE architecture. An intuition about how a convolutional layer preserves better the temporal information than a fully-connected is that each output neuron of a fully-connected layer is connected to all the neurons of the previous layer, aggregating the input context and, therefore, vanishing the influence of short-duration input features, specially when the input size of the network is large. However, the output neurons of a convolutional layer are only connected to a small part of its input neurons known as *receptive field* [41]. This fact reduces significantly the amount of temporal information that is aggregated throughout the layer in comparison with the fully-connected layers.

The suggested fully-convolutional architecture, shown in Fig. 1b, could mitigate this vanishing effect, replacing all fully-connected layers of the encoder by convolutional blocks which are made of two 1D convolutional layers with stride 1 followed by a 1D convolutional layer with stride 2. Convolutional layers with stride 2 reduce the input dimension along the convolutional blocks, making the bottleneck representation a coarse map which has to be transformed into a dense reconstruction of the individual consumption by the decoder. The way in which the decoder expands the intermediate coarse map is by means of 1D *transpose convolutional blocks* [42], [43], which are composed of a 1D transposed convolutional layer followed by two standard 1D convolutional layers with stride 1. The layout of the two 1D convolutional layers, before shrinking

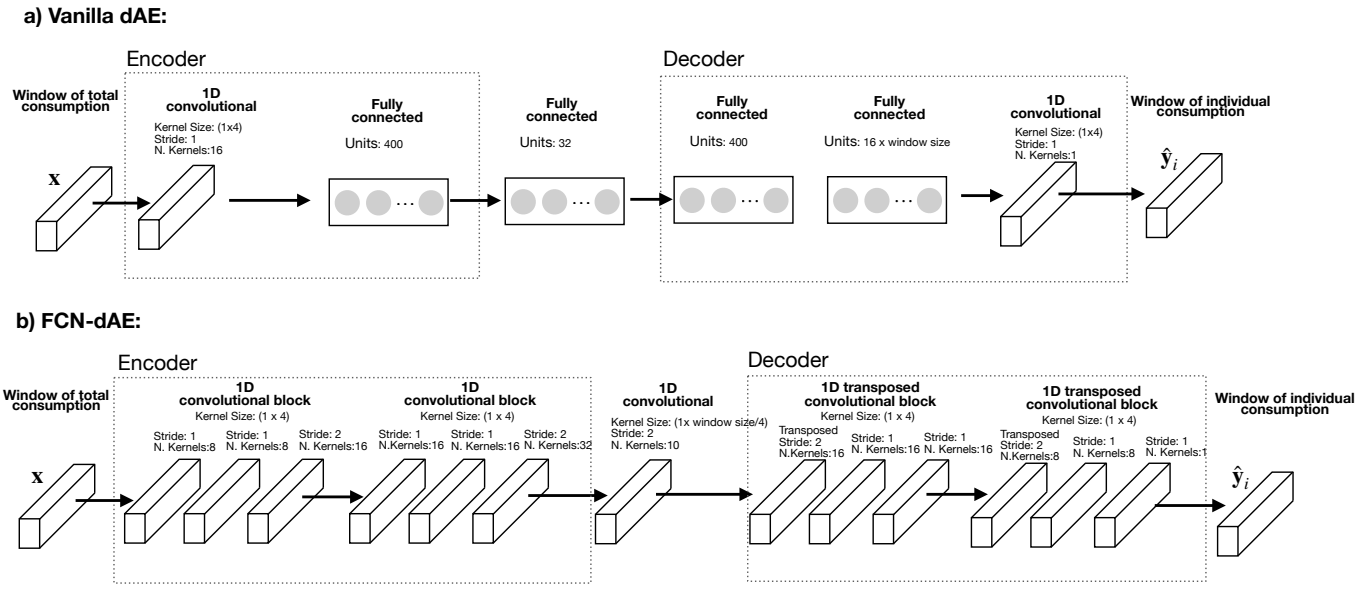


Fig. 1. The dAE-based architectures. a) Vanilla dAE architecture diagram where each cell symbolizes a layer explained by the number of units in the case of fully-connected layers; or the kernel size, stride and number of kernels in the case of convolutional layers. b) Proposed FCN-dAE architecture made entirely by 1D convolutional layers and 1D transposed convolutional layer.

the input dimensions in the case of the convolutional blocks and after expanding the input dimensions in the case of the transposed convolutional blocks, is similar to the first layers of the well-known VGG image recognition architecture [44] and it provides the FCN-dAE models with enough complexity for dealing with the problem of NILM in large buildings. Regarding the hyperparameters, such as the number and size of kernels, they were set experimentally to those values that most reduce the reconstruction error, in the same way as in the case of the vanilla dAE models.

As in the case of vanilla dAE, the FNC-dAE approach maps a window of total consumption $\mathbf{x}$ into a window of the same length of an individual consumption $\mathbf{y}_i$, so that it is necessary to train m models, one per available individual consumption, in the same way as the vanilla dAE case. Each of these models are trained with pairs of windows $\mathbf{x}$ and $\mathbf{y}_i$ and the disaggregation process from which the individual consumptions are estimated is the same as the explained in section III-B.

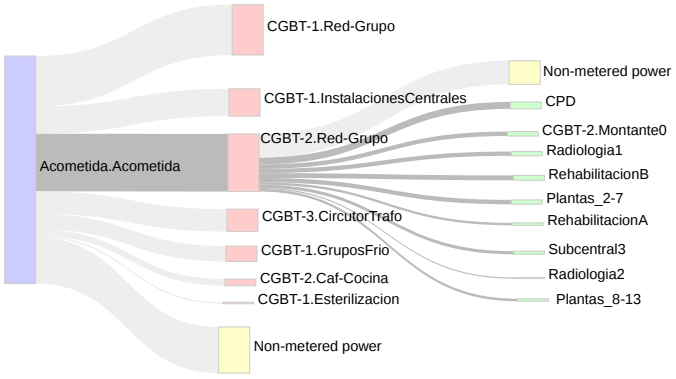


Fig. 2. Flow diagram of the hospital submetering system where nodes represent the installed meters and the width of the branches represent their mean active power. Note that the power supply meter is depicted in blue, branch meters in red and the leaf meters in green. Nodes in yellow are the non-measured electric power demand. The shaded paths in the diagram are the consumptions chosen for testing our NILM approaches.

IV. Hospital Electric Power Demand Data

The suggested models have been tested with real electric power demand data from a hospital facility, which include complete electric measurements such as current, voltage, power and power quality measurement (e.g., electrical harmonics and total harmonic distortion), gathered from 33 measuring points placed along the electrical installation, creating the submetering system shown in Fig. 2 and Table I. In the diagram, three levels of aggregation in the measurement are shown: 1) the meter connected at the supply connection; 2) seven branch meters which measure the electric demands of independent subsystems of the hospital; and 3) the leafs meters which measure downstream loads of each subsystem. Measure points were placed by the maintenance staff of the hospital so that branch meters cover independent areas of the hospital, such as *CGBT-2.Red-Grupo* which gathers all the consumptions of an entire wing of the hospital, or critical subsystems where a failure could compromise the activity of the hospital such as *CGBT-1.GruposFrio*, which measures the HVAC systems. Regarding the leaf meters, they were placed for monitoring specific downstream loads which are considered of interest by the maintenance staff because either they present large demands, or they are critical. As an example, in node *CGBT-2.Red-Grupo* several critical systems are measured, such as *CPD* (server room) or *Radiologia 1/2* (X-ray rooms).

In our experiments, we accept the leaf meters as the *individual consumptions* even though they collect several consumers, since installing one meter per device in such complex installation would be economically unfeasible. Among all the

TABLE I
DESCRIPTION OF METERS THAT CONFORM THE HOSPITAL'S SUBMETERING SYSTEM.

Meter	Sample rate	No. samples	Description
Acometida.Acometida	1 min	519278	Meter connected to the supply connection
CGBT-1.Red-Grupo	1 min	519278	All the consumptions of a building
CGBT-1.InstalacionesCentrales	1 min	519278	Meter connected to principal installations
CGBT-3.CircuitorTrafo	1 min	519278	Unknown
CGBT-1.GruposFrio	1 min	519278	Loads from HVAC system
CGBT-2.Caf-Cocina	1 min	519278	All the consumptions of kitchen facilities
CGBT-1.Esterilizacion	1 min	519278	Loads from sterilization facilities
CGBT-2.Red-Grupo	1 min	519278	Consumptions from entire wing of the hospital
CGBT-2.Montante0	1 min	507353	All consumptions from lifts
Radiologia1	1 min	507305	Consumptions from X-ray room 1
Radiologia2	1 min	507354	Consumptions from X-ray room 2
RehabilitacionA	1 min	507355	Consumptions from rehab facilities A
RehabilitacionB	1 min	507307	Consumptions from rehab facilities A
Subcentral3	1 min	507353	Unknown
CPD	1 min	507353	Consumptions of the server anda data center
Plantas_2-7	1 min	507353	Consumptions of the hospitalization floor from 2 to 7
Plantas_8-13	1 min	507355	Consumptions of the hospitalization floor from 8 to 13

measuring points, we decided to focus our NILM experiments on the *active power* of the branch meter *CGBT-2.Red-Grupo* and its downstream meters, because it aggregates all the loads from a wing of the main building of the hospital, collecting a rich variety of individual loads such as diagnostic equipment, loads from inpatient hospitalization floors or the demand from the servers of the data center, as it is shown in Table I. All the downstream nodes of the line *CGBT-2.Red-Grupo* were measured during one year with a sample rate of 1 min, which results in 525600 samples per node. However, as it is shown in Table I, some nodes have missing samples due to temporary malfunctions of the meter or maintenance works.

V. EVALUATION METRIC

Metrics used in the literature to evaluate the performance of NILM methods are commonly divided into: 1) classification metrics, which measure the accuracy of NILM methods in determining whether the state of an individual consumption is "on" or "off" and 2) disaggregation metrics, which measure how well the whole signal of an individual consumption is estimated by the algorithms.

In our case, classification metrics are inappropriate since our individual consumptions aggregate several downstream loads, so they are continuously active and we cannot make a binary classification. For this reason, our models should be evaluated by means of regression metrics which assess the deviation between the estimated individual loads and the real demand gathered by the meter. Although the standard method to evaluate regression problems in most domains is the *root mean squared error* (RMSE), some previous works on NILM [27], [45], [46] suggest that when the RMSE is used as the evaluation metric, the comparison between individual disaggregations is not coherent, since the total energy of each consumer has to be taken into account for a fair comparison. Thus, metrics normalized by the total power consumed of the load under study are commonly adopted, such as *Normalized error in assigned power* (NEP), which is computed dividing the sum of the absolute values of the difference between estimated power $\hat{y}_i(t)$ and real consumed power $y_i(t)$ of an individual consumption i by its actual total energy consumed during the whole test sequence of duration T:

$$NEP_i = \frac{\sum_{t=1}^{T} \mid \hat{y}_i(t) - y_i(t) \mid}{\sum_{t=1}^{T} y_i(t)} \tag{7}$$

From now on, we shall adopt NEP as the evaluation metric to compare the performance of the proposed NILM approaches.

VI. EXPERIMENTAL SET-UP

Deep learning (DL) models are well known for their stunning results, but also for being difficult to train due to the huge quantity of parameters to be configured, such as the network topology, the optimization algorithm or the layer-specific parameters. In our case, all the models used in the following sections are based on the topologies described in Fig. 1 and were trained with all the available active power demand data (total demand and individual loads) from the node *CGBT-2.Red-Grupo* of the hospital, except one month of demand, which was reserved as the test set. Hence, we approximately use 90% of the available data for training and 10% for testing. Windows of the total active power consumption and the individual load under study with the desired size are randomly taken from the training set and they are used as the training examples in the optimization. Once the training windows are selected, the total consumption windows (inputs of the models) are normalized in order to speed up the optimization process, subtracting the mean and dividing by the standard deviation of the whole signal. By contrast, the individual consumptions windows (outputs of the models) are scaled between $[0, 1]$, dividing by the maximum value of electrical power demand of the node to be disaggregated. After data normalization, each disaggregation model is trained by means of the *Adam optimizer* [47] in order to minimize the mean squared error between the output and the ground truth node consumption during an indicated number of training epochs which is set to 100, using batches of size 100.

Regarding the disaggregation process, we split the segment of the total consumption under study into windows with the same size of the model and a stride between windows fixed to 5 minutes. Finally, both the training and the disaggregation processes were performed in a Linux-based operating system with an *NVIDIA GeForce GTX 1060 6GB* GPU which significantly reduces the computational time needed to optimize and test the proposed models.

VII. RESULTS

A. The effect of the window size

In order to explore how the input window size affects the disaggregation accuracy, both dAE and FCN-dAE approaches were implemented with input windows of different sizes: 32, 120, 1440 minutes. Input window sizes larger than 1440 minutes would make the number of dAE's trainable weights explode and therefore problems as overfitting or extremely long training times may appear. For this reason, we decided to keep the architectures as minimal as possible and set the maximum input window size of the models to 1440 minutes.

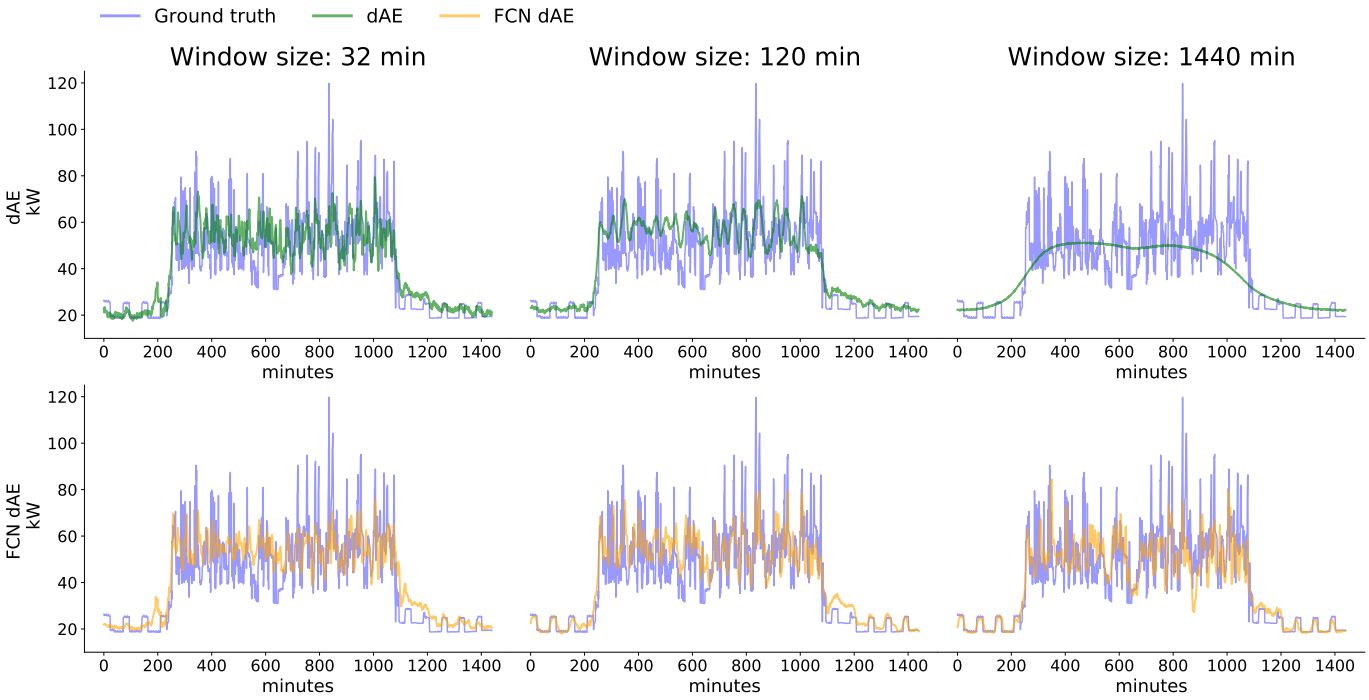


Fig. 3. Disaggregation results of node *Radiologia1* obtained after applying the two proposed architectures, trained with input window sizes of 30, 120 and 1440 minutes, to a daily snapshot of the total consumption.

TABLE II
DISAGGREGATION PERFORMANCE OF ALL INDIVIDUAL NODES MEASURED BY MEANS OF NEP METRIC. EACH INDIVIDUAL NODE WAS ESTIMATED BY THE DAE AND FCN-DAE ARCHITECTURES WITH THE FOLLOWING INPUT WINDOW SIZES: 32, 120 AND 1440 MINUTES.

Architecture	**dAE**			**FCN-dAE**		
Window size	32	120	1440	32	120	1440
CGBT-2.Montante0	0.2977	0.1523	0.5326	0.3013	0.1582	**0.0922**
CPD	0.2676	0.1089	0.0776	0.2643	0.1088	**0.0770**
Plantas_2-7	0.3268	0.1966	**0.0813**	0.3267	0.1810	0.0822
Plantas_8-13	0.4920	0.3940	0.3531	0.5158	0.4046	**0.1467**
Radiologia1	0.3509	0.1808	0.2476	0.3465	0.1927	**0.1709**
Radiologia2	0.4602	**0.2491**	0.3094	0.4421	0.3198	0.2796
RehabilitacionA	0.3110	0.1771	0.1265	0.2982	0.1697	**0.0963**
RehabilitacionB	0.2835	0.1218	0.4821	0.2881	0.1397	**0.0962**
Subcentral3	0.3562	0.2390	0.1377	0.3604	0.2375	**0.1191**

In addition, limiting the input temporal context of the models to daily windows is consistent with the nature of the problem of NILM in large buildings due to the strong daily patterns observed in their energy demand signals.

The two architectures under analysis were trained to filter each individual node from the total consumption, resulting in 9 different models per architecture and window size, so that the total number of models in the experiment adds up to 54. Table II shows the evaluation of the experiment in terms of the NEP metrics computed on the disaggregated individual consumptions resulting from applying each proposed model to the test portion of total consumption not used in the training. In the case of FCN-dAE, the higher the window size, the more accurate the disaggregation, so that FCN-dAE models with long input windows outperform all the models except in two individual consumptions, where a dAE model beat the FCN-dAE topologies by a small margin. By contrast, a window size increase on dAE architectures may not result in an improvement of the accuracy of the disaggregations, since in some individual consumptions such as *CGBT-2.Montante0*, when the window size is enlarged from 120 to 1440, the accuracy of the outcome decreases sharply.

For a qualitative analysis of the experiment, all the obtained disaggregations for the individual node *Radiologia1* for a day are depicted in Fig. 3, showing how the short-term behaviors of real consumption cannot be reproduced by the dAE approach trained with long windows, while the similar FCN model achieves accurate disaggregations on both long-term and short-term consumptions. This suggests that the drop in the performance of dAE trained with long windows is closely related to the reconstruction of short-term behaviors.

In order to provide insight into the error analysis, we suggest to analyze the residuals between the real and the disaggregated consumptions in the frequency domain by computing their *power spectral density* (PSD), using the Welch method which maps the residuals into the frequency spectrum, showing which frequency bands concentrate more error. This representation in the frequency domain provides us with a suitable knowledge about both the short-term and long-term errors.

The PSD of the residuals obtained from the estimated individual loads of the node *Radiologia1* are shown in Fig. 4, where it can be seen that two principal harmonics emerge with daily and 75 minutes periods, respectively. These highlighted frequencies are strongly related to the consumption *Radiologia1*, since, as it is shown in Fig. 3, it presents a common daily pattern and a periodic load during the night with a duration of approximately 75 minutes. Although enlarging the window size reduces the error density in the daily harmonic for all the approaches, the error in the harmonic corresponding to the 75

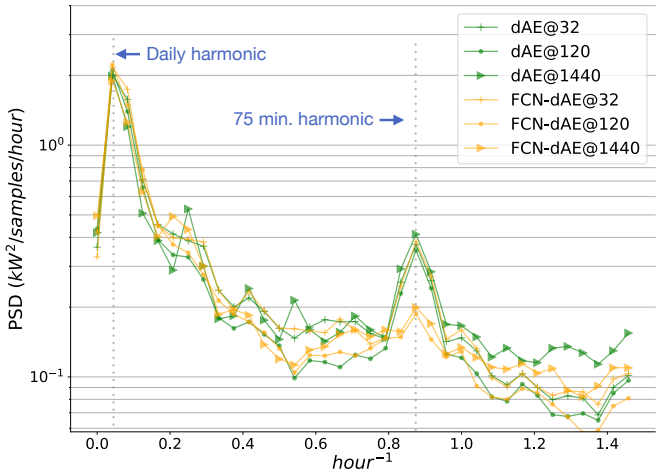


Fig. 4. Power spectral density of the residuals calculated from the disaggregations resulting from applying the dAE and FCN-dAE architectures, trained with input window sizes of 30, 120 and 1440 minutes, to the individual node *Radiologia1*. The highlighted harmonics correspond to the errors made by the models at daily (freq. $= 0.041\,\mathrm{hour}^{-1}$) and at 75 minutes (freq. $= 0.8\,\mathrm{hour}^{-1}$) periods.

minutes signature rises in the case of the dAE models, while FCN-dAE models are less sensitive to it, keeping errors low for all window sizes except the case of 32 minutes input window, where the error at this harmonic remains high due to the lack of temporal input context. Therefore, FCN-dAEs, when the input window size is large enough, provide disaggregations that are not only more accurate in general terms, but also are more robust at reproducing both long-term and short-term patterns in individual nodes.

VIII. Use cases in large buildings

In this section, we explain how to use dAE-based NILM models in other energy demand management tasks in large buildings, such as the organization and maintenance of the submetering data acquisition, filling gaps in the gathered data or detecting novelties in the individual consumptions which could be related to faults in the installations.

A. Replacement of real meters

The most obvious application is the replacement of real meters by their corresponding estimated disaggregations. Thus, the replaced meters could be installed in other non-measured areas of the installation, increasing the knowledge of the system without any additional cost. This functionality is also useful when a meter needs maintenance, since it can be replaced by the NILM system, avoiding the loss of measures in the process.

B. Gap filling

A common issue in large submetering installations is the occasional loss of data due to temporary malfunctions of meters, network traffic or routine maintenance works on the devices. Suggested NILM systems provide accurate estimations of downstream individual consumptions by means of the learned functions f_i. We suggest the use of functions f_i for accurately filling gaps in the individual meters measurements $\mathbf{y}_i$ since they only need the input consumption $\mathbf{x}$ to compute their corresponding estimations $\hat{\mathbf{y}}_\mathbf{i}$, as it is shown in Eq. 6. Therefore, the elements of sequence $\mathbf{y}_i$ corresponding to the gap could be directly replaced by their corresponding elements of sequence $\hat{\mathbf{y}}_i$. However, when the failure affects all the nodes, including the total consumption $\mathbf{x}$, the input windows of the models have missing elements. A solution for these gaps is to fill them with zeros, but the individual node estimations might still be compromised, since the denoising models lack temporal information in the input.

TABLE III
NEP metrics obtained after the disaggregation of three test total consumptions with gaps of size: 20, 40 and 60 minutes. Moreover, the disaggregation without any gap is included as a baseline.

Gap size	dae@32	dae@120	dae@1440	fcn@32	fcn@120	fcn@1440
0 min	0.901	0.177	0.126	0.878	0.169	0.096
20 min	0.937	0.177	0.127	0.877	0.168	0.096
40 min	0.954	0.182	0.131	0.884	0.169	0.098
60 min	0.977	0.190	0.135	0.890	0.171	0.099

In order to determine which models among the proposed ones are more robust in this scenario, we suggest an experiment where all the trained models for a specific individual node, such as *RehabilitacionA*, are tested on three different versions of the test segment of total consumption where gaps of sizes 10, 40 and 60 minutes have been respectively introduced. In all cases, a total of 100 gaps have been placed randomly along the month of total consumption in order to evaluate the gap sensitivity regardless of when the missing data is introduced. Table III contains the resulting NEP metrics of the experiment and, in addition, we included the NEP obtained from the disaggregation of the consumption without any gap, as a baseline. As expected, when the gap size in the input is augmented, the accuracy of the disaggregations for all the models decreases because of the loss of temporal context in the input. Although this drop in the performance is generalized, the disaggregations produced by FCN models remain more stable since their internal activations only depend on the convolutional receptive field of the input.

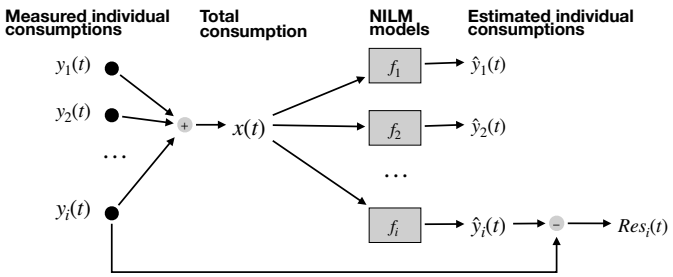


Fig. 5. Explanatory diagram of the mechanism for obtaining the residuals from the real measured individual consumptions and the estimated individual consumptions obtained.

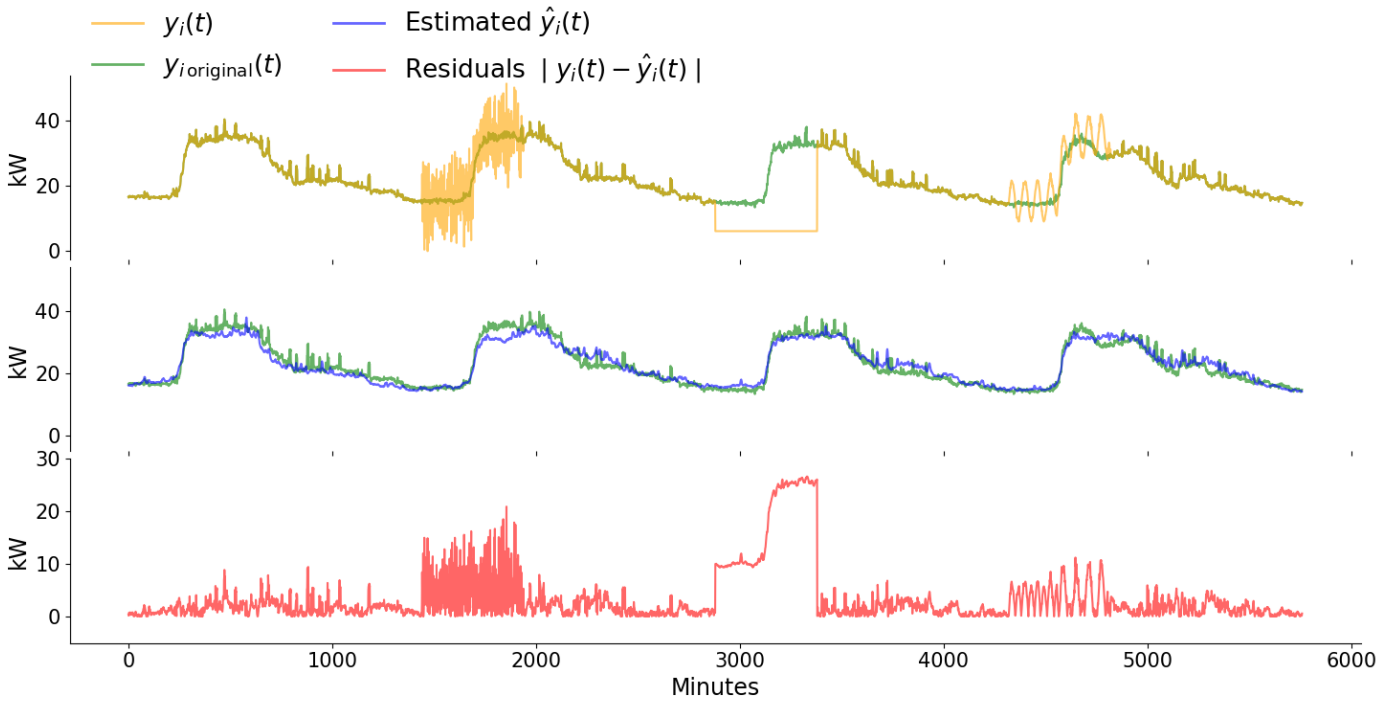


Fig. 6. Novelties detected in individual node *RehabilitacionA* using FCN-dAE with a window size of 1440 minutes as a disaggregator. The first row illustrates the modified individual consumption $y_i(t) = y_{i\,\text{original}}(t) + \text{noise}$ along with the original individual load $y_{i\,\text{original}}(t)$; in the second row, the resulting disaggregation $\hat{y}_i(t)$ is compared to the original individual load $y_{i\,\text{original}}(t)$; and in the last row, the residuals between the estimated individual consumption $\hat{y}(t)$ and the modified load $y_i(t)$ are shown.

C. Novelty detection

Although our approaches provide accurate results in total consumptions similar to those used in training, when an individual node presents abnormal consumptions, dAE-based disaggregators provide estimations that ignore the anomalies and produce estimations similar to the individual consumption used in the training. Hence, in the case that the real measurement is available, the difference or *residual* between the estimated and the measured individual consumption might be used as an indicator of novelty in the individual nodes. This idea is graphically explained in Fig. 5 where the residuals $Res_i(t)$ for a specific individual node i are the absolute value of the difference between the estimated and the real individual consumption:

$$Res_i(t) = \mid y_i(t) - \hat{y}_i(t) \mid \tag{8}$$

where $y_i(t)$ is the real consumption and $\hat{y}_i(t)$ is the disaggregated individual consumption. The correlation between the value of Res_i and the novelty degree in a individual node can be evaluated by applying the previously trained FCN-dAE model with a window size of 1440 minutes to a 4-day segment of a modified total consumption where the individual node $RehabilitacionA$ has been altered by the following novelties: 1) adding a Gaussian noise with a standard deviation of $6\,kW$; 2) setting a constant value of $6\,kW$; and 3) adding a sine wave with an amplitude of $6\,kW$ and a frequency of 1 hour. We choose an amplitude of $6\,kW$ for the introduced novelties because it represents a variation of 13% with respect to the maximum demand registered in the individual node (about $45\,kW$), and we consider that it is comparable to a new device connection or a downstream failure. These modifications are also added to the total input consumption simulating a real situation where the general meter captures the novelty as well. The first row of Fig. 6 shows the modified consumptions together with the normal consumptions of $RehabilitacionA$. Note that in this first experiment, the added signals have a duration of 500 minutes and are placed in the last three days, leaving the first day intact in order to compare the disaggregation with and without novelties. The second row of Fig. 6 shows that the estimated individual loads follow the same distribution as the individual load without novelties regardless of the introduced signals. This invariance to the modifications on the individual load makes Res_i increase where the individual node has been modified, as it can be seen in the third row of the figure.

The strong correlations between $Res_i(t)$ and the deviations from the normal behavior in individual loads lead us to use the residuals as decision values in a normal/novelty binary classification. The final decision of the binary classification is made comparing the residue value $Res_i(t)$ with a user-defined threshold θ_i over which the samples are classified as positive class or novel consumption.

$$C_i(t) = \begin{cases} 0, & \text{if } \mid y_i(t) - \hat{y}_i(t) \mid < \theta_i \\ 1, & \text{otherwise} \end{cases} \tag{9}$$

Note that θ_i has to be defined for each individual load i independently, since the range of values of the residuals for each node could differ. In order to evaluate the performance of C_i for different values of θ_i, Fig. 7 shows the precision-recall curves obtained from applying all the proposed models

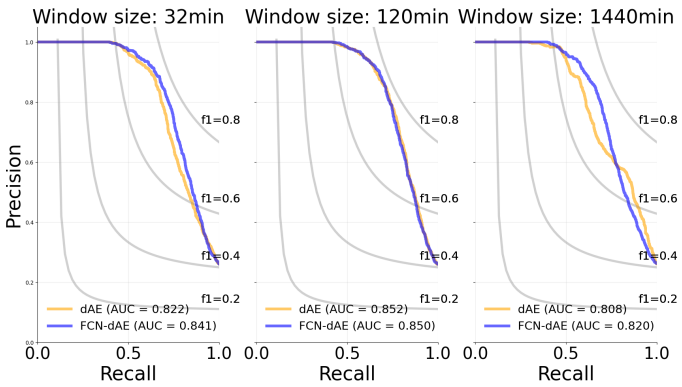


Fig. 7. Precision-recall curves for residuals computed from all the proposed models ordered by window size. Grey lines indicate the precision-recall curve for a model with the specified F1-score.

in section VII-A to the modified total consumptions plotted in Fig. 6.

Precision-recall curves indicate that all the tested models reach an accuracy higher than 0.7 for some value of θ, using the F1-score [48] as a binary classification metric. Moreover, FCN-dAE models have slightly higher areas under the curve (AUC) than dAE models and therefore they provide a more accurate classification. Although this difference in AUC is minimal and it could be related to the node under study, precision-recall curves show that the classifications based on FCN-dAE models are more stable to the window sizes variations. This result is aligned with all the findings of this article, reaffirming the robustness of the FCN-dAE architectures against the traditional dAE models.

IX. Conclusions

In this paper, two approaches based on deep neural networks a) denoising auto-encoder (dAE) and b) fully-convolutional denoising auto-encoder (FCN-dAE) were compared as NILM systems using real data from a hospital facility. Several models were used for studying the performance of disaggregation in both architectures taking into account each individual node and several input window sizes. FCN-dAE models with a large input window size outperformed vanilla dAE for the majority of the individual consumptions. In addition, after analyzing the disaggregation of a particular node in terms of frequency, FCN-dAE models have proved to be more accurate disaggregating short and long-term consumptions.

Moreover, use cases show the usefulness of the suggested NILM systems for the energy management in large buildings. In particular, we tackled the gap filling and novelty detection problems in individual consumptions. In both cases, FCN models have demonstrated more robustness, being more invariant to gaps in the total input consumptions than the dAE approaches and detecting abnormal consumptions in the individual nodes with a remarkable precision.

Acknowledgment

The authors would like to thank financial support from the Principado de Asturias government through the predoctoral grant "Severo Ochoa" and to the Hospital Universitario de León for the cession of data and knowledge about the facility used in the experiments.

References

[1] S. Ahmadi-Karvigh, B. Becerik-Gerber, and L. Soibelman, "A framework for allocating personalized appliance-level disaggregated electricity consumption to daily activities," *Energy and Buildings*, vol. 111, pp. 337–350, 2016.

[2] K. Ehrhardt-Martinez, K. A. Donnelly, S. Laitner *et al.*, "Advanced metering initiatives and residential feedback programs: a meta-review for household electricity-saving opportunities," in *American Council for an Energy-Efficient Economy Washington, DC*, 2010.

[3] W. Gans, A. Alberini, and A. Longo, "Smart meter devices and the effect of feedback on residential electricity consumption: Evidence from a natural experiment in northern ireland," *Energy Economics*, vol. 36, pp. 729–743, 2013.

[4] J. Carroll, S. Lyons, and E. Denny, "Reducing household electricity demand through smart metering: The role of improved information about energy saving," *Energy Economics*, vol. 45, pp. 234–243, 2014.

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

[6] N. J. Yadwadkar, C. Bhattacharyya, K. Gopinath, T. Niranjan, and S. Susarla, "Discovery of application workloads from network file traces." in *FAST*, 2010, pp. 183–196.

[7] H. Kim, M. Marwah, M. Arlitt, G. Lyon, and J. Han, "Unsupervised disaggregation of low frequency power measurements," in *Proceedings of the 2011 SIAM International Conference on Data Mining*. SIAM, 2011, pp. 747–758.

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

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

[10] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "An unsupervised training method for non-intrusive appliance load monitoring," *Artificial Intelligence*, vol. 217, pp. 1 – 19, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0004370214001003

[11] Y. Li, Z. Peng, J. Huang, Z. Zhang, and J. H. Son, "Energy disaggregation via hierarchical factorial hmm," in *Proceedings of the 2nd International Workshop on Non-Intrusive Load Monitoring*, 2014.

[12] S. Makonin, F. Popowich, I. V. Bajić, B. Gill, and L. Bartram, "Exploiting hmm sparsity to perform online real-time nonintrusive load monitoring," *IEEE Transactions on smart grid*, vol. 7, no. 6, pp. 2575–2585, 2015.

[13] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.

[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[15] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1915–1929, 2012.

[16] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[17] G. Hinton, L. Deng, D. Yu, G. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, B. Kingsbury *et al.*, "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal processing magazine*, vol. 29, 2012.

[18] T. Mikolov, A. Deoras, D. Povey, L. Burget, and J. Černockỳ, "Strategies for training large scale neural network language models," in *2011 IEEE Workshop on Automatic Speech Recognition & Understanding*. IEEE, 2011, pp. 196–201.

[19] D. Kelly, "Disaggregation of domestic smart meter energy data," Ph.D. dissertation, Imperial College London, 2016.

[20] P. P. M. do Nascimento, "Applications of Deep Learning Techniques on NILM," PhD Thesis, Universidade Federal do Rio de Janeiro, 2016.

[21] J. Kelly and W. Knottenbelt, "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*, 2015, pp. 55–64.

[22] A. M. Sudoso and V. Piccialli, "Non-intrusive load monitoring with an attention-based deep neural network," *arXiv preprint arXiv:1912.00759*, 2019.

[23] M. Kaselimi, N. Doulamis, A. Voulodimos, E. Protopapadakis, and A. Doulamis, "Context aware energy disaggregation using adaptive bidirectional LSTM models," *IEEE Transactions on Smart Grid*, 2020, publisher: IEEE.

[24] T. Wang, T. Ji, and M. Li, "A new approach for supervised power disaggregation by using a denoising autoencoder and recurrent lstm network," in *2019 IEEE 12th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*. IEEE, 2019, pp. 507–512.

[25] J. Kim, T. Le, and H. Kim, "Nonintrusive load monitoring based on advanced deep learning and novel signature." *Computational intelligence and neuroscience*, vol. 2017, pp. 4 216 281–4 216 281, 2017.

[26] W. He and Y. Chai, "An empirical study on energy disaggregation via deep learning," *Advances in Intelligent Systems Research*, vol. 133, 2016.

[27] R. Bonfigli, A. Felicetti, E. Principi, M. Fagiani, S. Squartini, and F. Piazza, "Denoising autoencoders for non-intrusive load monitoring: Improvements and comparative evaluation," *Energy and Buildings*, vol. 158, pp. 1461 – 1474, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378778817314457

[28] K. S. Barsim and B. Yang, "On the feasibility of generic deep disaggregation for single-load extraction," *CoRR*, vol. abs/1802.02139, 2018. [Online]. Available: http://arxiv.org/abs/1802.02139

[29] M. Valenti, R. Bonfigli, E. Principi, and S. Squartini, "Exploiting the reactive power in deep neural models for non-intrusive load monitoring," in *2018 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2018, pp. 1–8.

[30] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the 25th international conference on Machine learning*. ACM, 2008, pp. 1096–1103.

[31] S. Henriet, U. Şimşekli, B. Fuentes, and G. Richard, "A generative model for non-intrusive load monitoring in commercial buildings," *Energy and Buildings*, vol. 177, pp. 268–278, 2018.

[32] L. K. Norford and S. B. Leeb, "Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms," *Energy and Buildings*, vol. 24, no. 1, pp. 51 – 64, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0378778895009582

[33] N. Batra, O. Parson, M. Berges, A. Singh, and A. Rogers, "A comparison of non-intrusive load monitoring methods for commercial and residential buildings," *arXiv*, pp. arXiv–1408, 2014.

[34] L. Pérez-Lombard, J. Ortiz, and C. Pout, "A review on buildings energy consumption information," *Energy and buildings*, vol. 40, no. 3, pp. 394–398, 2008.

[35] Y. Liu and M. Chen, "A review of nonintrusive load monitoring and its application in commercial building," in *The 4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent*. IEEE, 2014, pp. 623–629.

[36] L. K. Norford and S. B. Leeb, "Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms," *Energy and Buildings*, vol. 24, no. 1, pp. 51–64, 1996.

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

[38] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv preprint arXiv:1803.01271*, 2018.

[39] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[40] S. Gupta, R. Girshick, P. Arbeláez, and J. Malik, "Learning rich features from rgb-d images for object detection and segmentation," in *European conference on computer vision*. Springer, 2014, pp. 345–360.

[41] A. Araujo, W. Norris, and J. Sim, "Computing receptive fields of convolutional neural networks," *Distill*, 2019, https://distill.pub/2019/computing-receptive-fields.

[42] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," *arXiv preprint arXiv:1603.07285*, 2016.

[43] A. Odena, V. Dumoulin, and C. Olah, "Deconvolution and checkerboard artifacts," *Distill*, 2016. [Online]. Available: http://distill.pub/2016/deconv-checkerboard

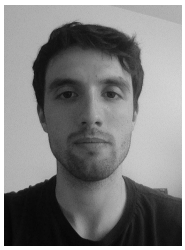
[44] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[45] A. Faustine, N. H. Mvungi, S. Kaijage, and K. Michael, "A Survey on Non-Intrusive Load Monitoring Methodies and Techniques for Energy Disaggregation Problem," *arXiv:1703.00785 [cs]*, Mar. 2017, arXiv: 1703.00785. [Online]. Available: http://arxiv.org/abs/1703.00785

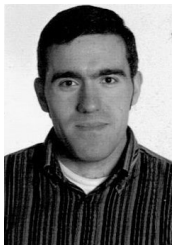
[46] N. Batra, J. Kelly, O. Parson, H. Dutta, W. Knottenbelt, A. Rogers, A. Singh, and M. Srivastava, "Nilmtk: an open source toolkit for non-intrusive load monitoring," in *Proceedings of the 5th international conference on Future energy systems*. ACM, 2014, pp. 265–276.

[47] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[48] N. Chinchor and B. M. Sundheim, "Muc-5 evaluation metrics," in *Fifth Message Understanding Conference (MUC-5): Proceedings of a Conference Held in Baltimore, Maryland, August 25-27, 1993*, 1993.



DIEGO GARCÍA PÉREZ recieved the Bachelor's degree in automation and industrial electronic engineering from the University of León in 2015 and the Master's degree in automation engineering and industrial informatics from the University of Oviedo in 2017. Currently, he is pursuing the Ph.D. degree in the University of Oviedo and his research focuses on the application of deep learning and visual analytics techniques to energy data in order to improve the energy efficiency in large buildings.



DANIEL PÉREZ LÓPEZ received the M.Eng. degree in electrical engineering from University of Oviedo in 2007, later he completed the PhD degree related to visual analytics in 2015. From 2006 to 2017, he worked as research assistant in the Department of Electrical, Electronics, Computer, and Systems Engineering in University of Oviedo. In 2018, he joined the University of León as postdoctoral researcher and currently he is Assistant Professor. His main research interests include information visualization, process monitoring and machine learning techniques applied to industrial environments.

IGNACIO DÍAZ BLANCO is an associate professor of Electrical Engineering Department at the University of Oviedo since 2004. He received a M.Eng. in 1995 and his Ph.D. in industrial engineering in 2000 from the University of Oviedo. His main research interests are the application of data visualization and intelligent data analysis algorithms to industrial problems. He has led several R&D projects financed by the Spanish Government and the European Union and published his research in indexed journals, as well as numerous publications in international conferences. Professor Díaz is member of the IEEE since 1997.

ANA GONZÁLEZ MUÑIZ received the Bachelor's and Master's degrees in automation and industrial electronic engineering in 2016 and 2018, respectively, both from Gijón Polytechnic School of Engineering (University of Oviedo, Spain), where she is currently pursuing the Ph.D. degree. Her research focuses on the application of deep learning techniques for process analysis and monitoring.

MANUEL DOMÍNGUEZ GONZÁLEZ was born in 1956. He holds a M.S. degree in industrial engineering and received his Ph.D. from the Universidad de Oviedo, Gijón, Spain, in 2003. He is a Professor of Automatic Control in the Department of Electrical and Systems Engineering, at the School of Engineering of the Universidad de León, León Spain. He is author of more than 100 publications in international conferences and journals, as well as several book chapters. His main research interests are industrial remote laboratories, the development of innovative tools for education in automatic control, remote monitoring, automation and machine learning applications in industrial automation and cybersecurity. He is a member of IFAC and IEEE, he is also member of the Spanish Committee on Automatic control.

ABEL ALBERTO CUADRADO VEGA received the M.S. and Ph.D. degrees in electronic and control engineering from the University of Oviedo, Gijon, Spain, in 1998 and 2003, respectively. He was a Visiting Researcher at Betriebsforschungsinstitut, Düsseldorf, Germany, in 2005. He is currently an Associate Professor with the Department of Electrical, Electronic, Computer, and Systems Engineering, University of Oviedo, teaching control theory and automation. His research interests include supervision of complex industrial processes, especially those related to electric power generation and the steel industry. Dr. Cuadrado Vega received the University of Oviedo Outstanding Ph.D. Thesis Award.