

# CONVOLUTIONAL NEURAL NETWORK APPLIED TO THE IDENTIFICATION OF RESIDENTIAL EQUIPMENT IN NON- INTRUSIVE LOAD MONITORING SYSTEMS

Deyvison de Paiva Penha and Adriana Rosa Garcez Castro

Institute of Technology, Federal University of Para, Belém, Brazil

## **ABSTRACT**

*This paper presents the proposal of A new methodology for the identification of residential equipment in non-intrusive load monitoring systems that is based on a Convolutional Neural Network to classify equipment. The transient power signal data obtained at the time an equipment is connected in a residence is used as inputs to the system. The methodology was developed using data from a public database (REED) that presents data collected at a low frequency (1 Hz). The results obtained in the test database indicate that the proposed system is able to carry out the identification task, and presented satisfactory results when compared with the results already presented in the literature for the problem in question.*

## **KEYWORDS**

*Convolutional Neural Networks, Identification of Residential Equipment, Non-Intrusive Load Monitoring, NILM System, Energy Conservation*

## **1. INTRODUCTION**

The reduction and rationalization of electricity consumption are increasingly becoming priorities, not only for residential consumers, but also for electric power companies and government. Considering this concern, which is worldwide, research in Non-Intrusive Load Monitoring (NILM) has been emphasizing. Research in this area began in 1992 with the presentation of the work of George W. Hart [1] and since then many works have been presented, focusing on the various stages of a NILM system.

A NILM system has as main objective to measure an aggregate load of a residence through a single sensor, placed in the central meter of the residence. From the aggregate load, measured over a period of time, it is possible, through specific software, to carry out an identification of the electric equipment in operation and obtain the individual consumption thereof, in addition to obtaining the operating hours of each equipment [1]. This information can be used by residential consumers to take actions aimed at reducing and rationalizing their consumption, thus ensuring greater energy efficiency. In addition to this main functionality of the NILM systems, it is also possible to highlight: the possibility of identifying the load profile of a residence; possibility of identifying non-standard behavior of loads; possibility of detection of power failures and thefts; possibility of the use of the information of the load disaggregated by the electric power concessionaires that can promote aid to their customers in the process of identification of waste

during peak hours, thus helping to reduce consumption during these periods, offering for this incentive to consumers [2].

Considering the good results already presented by the academic community involving deep neural networks for the NILM problem, this paper presents the results obtained from the application of Convolutional Neural Networks for the problem of equipment identification. Here, unlike what we already have in the literature, a CNN network was developed to identify the type of equipment from the transient power signal data obtained at the moment an equipment is connected. The choice of the use of the transient power signal is due to the fact that each type of equipment presents different transient signal characteristics, depending on the generation mechanism, which is suitable for the development of classification systems. For the development and testing of the proposed system, the public database was used, and much used by researchers in the area, REDD (Reference Energy Disaggregation Dataset) [3]. This database has data of several equipments that were collected individually in 6 different residences at a frequency of 1 Hz. The system was developed to identify 7 equipments, these being classified as on / off loads, multilevel or variable.

## 2. NON-INTRUSIVE LOAD MONITORING SYSTEMS

The non-intrusive load monitoring aims to obtain a good approximation of the various electric devices in operation in a residence, using dedicated hardware and software [4]. The monitoring and identification of loads are performed based on the analysis of measurements of a single point of current and voltage of the aggregate load obtained through a meter outside the residence. Since each electrical equipment has its own profile of energy consumption called the electric signature, the developed algorithms try to identify such signatures in the aggregate load curve, thus indicating the periods of operation of the equipment and their respective energy consumption. The methodology of a NILM system is based on four main steps, as can be seen in Figure 1, which are the signal acquisition, event detection, characteristic extraction and equipment identification, as can be seen in Figure 1.

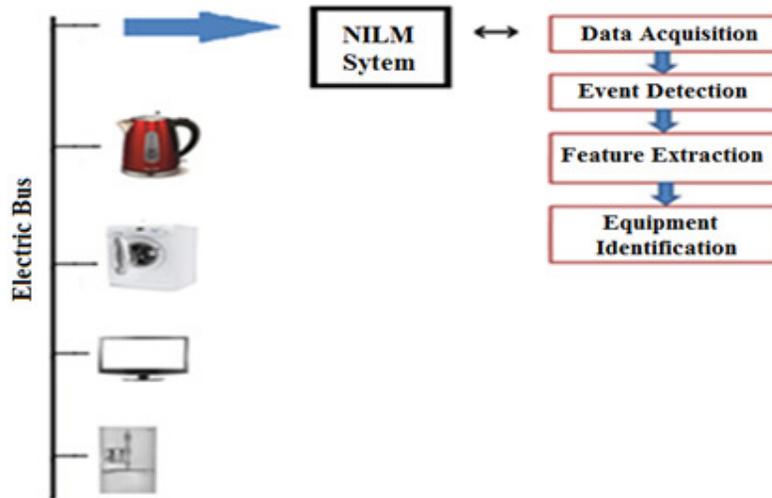


Figure 1 Residential electricity system with integrated NILM system

During the signal acquisition step, the aggregate load is measured through a single sensor on the main branch that is outside the residence. Figure 5 shows an example of the load measured over a period of 1 hour for one of the 7 equipment chosen (Refrigerator). For this stage we use the public database REDD (Reference Energy Disaggregation Data Set), being one of the most used in the

field of NILM systems research. REDD consists of data collected in six households, and contains aggregate electrical power data collected at the 1Hz frequency [3]. Table 1 shows the equipment per household that was measured in REDD.

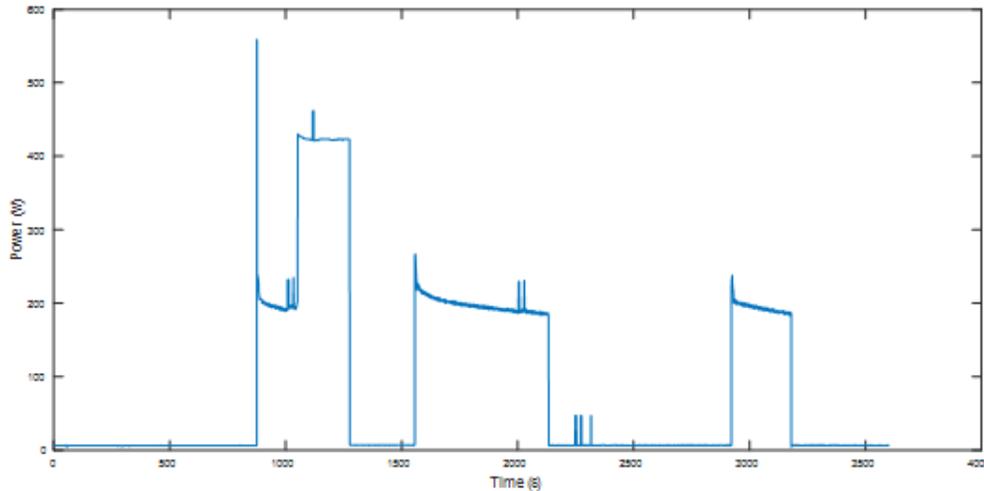


Figure 2 Load example measured over an hour

Table 1. Description of the houses and devices used in the evaluation in REDD data set [3].

House	Device Categories
1	Electronics, Lighting, Refrigerator, Disposal, Dishwasher, Furnace, Washer Dryer, Smoke Alarms, Bathroom GFI, Kitchen Outlets, Microwave
2	Lighting, Refrigerator, Dishwasher, Washer Dryer, Bathroom GFI, Kitchen Outlets, Oven, Microwave, Electric Heat, Stove
3	Electronics, Lighting, Refrigerator, Disposal, Dishwasher, Furnace, Washer Dryer, Bathroom GFI, Kitchen Outlets, Microwave, Electric Heat, Outdoor Outlets
4	Lighting, Dishwasher, Furnace, Washer Dryer, Smoke Alarms, Bathroom GFI, Kitchen Outlets, Stove, Disposal, Air Conditioning
5	Lighting, Refrigerator, Disposal, Dishwasher, Washer Dryer, Kitchen Outlets, Microwave, Stove
6	Lights, refrigerator, crazy washer, heater, clothes dryer, bathroom equipment, cooking utensils, cooker, electronic, air conditioning.

Still in the first stage, 7 electric appliances were chosen for the development of the planned system that was based on a convolutional neural network for the identification of the equipment. These were as follows: a microwave, oven, stove, a dishwasher, an air conditioning, a washer/dryer and a refrigerator. The chosen equipment can be regarded as comprising the machines that consume most energy in a household. According to Batra [5], priority should be given to identifying the equipment that uses most energy in the dwellings because these appliances have the most significant features in the aggregate load and thus other appliances that consume less can be regarded as the only noisy items in the total aggregate load.

In the event detection stage, the on / off moments of equipment in a residence (event) are detected from the aggregate signal. In order to detect abrupt changes in the signal, a methodology was

used based on an analysis window that scans the whole measured aggregate load, and it is possible to identify the occurrence of an event when the difference between the final average and the initial mean (left margin mean) of the window reach a predetermined threshold value, as can be seen in Figure 2. For each detected event, the first twelve transient samples were separated to form the training database of the system. The choice of the number of samples to be used as input to the system was based on the evaluation, for all equipment, of the number of samples sufficient to characterize a complete transient.

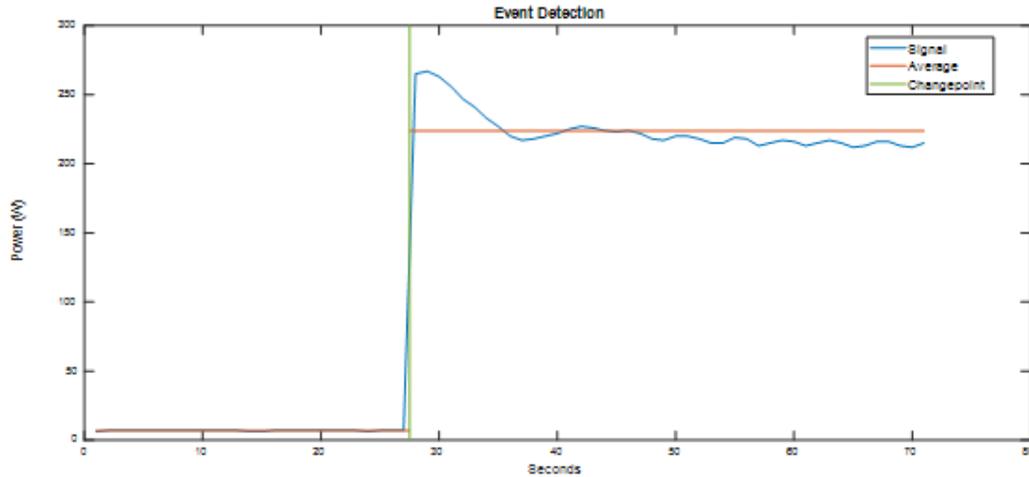


Figure 3 Event Detection Through Windowing

With the detected events, the third stage, of characteristic extraction or electric signatures, takes place. Electrical signatures represent a set of characteristics of voltage, current or power for a given equipment, and can be divided into macroscopic and microscopic. The macroscopic (low frequency) characteristics must be obtained from a sample period of up to one sample per cycle (1 Hz), which is the focus of this work. In the fourth and last step, from the characteristics / signatures extracted, we have the identification of each equipment for each detected event. Methods for identifying equipment used in NILM systems may be of the supervised or unsupervised type.

## 2.1. Previous work on approach NILM systems

In [6] the authors point out the main supervised techniques to solve NILM problems, such as Artificial Neural Networks (ANN), Supporting Vector Machines (SVM), Naive Bayes Classifier and K-Nearest Neighbor (KNN). Recently the researchers have turned their attention to the use of Deep Neural Networks to the problem of equipment identification. In [7] the authors apply 3 types of deep neural networks, a recurrent neural network based on Long Short Term Memory Units, a self-encoder neural network and a convolutional neural network, to predict the start and end time of an event of an equipment, as well as to predict the average demand of each device. In [8] the author sought to make an analysis of the various methods of deep learning to improve the performance of a NILM system. In [9], the authors used convolutional neural networks for the task of load disaggregation, promoting the individual identification of equipment loads based on the time series of the aggregate load. In [10], it is shown that CNN networks can also be used in the NILM context for equipment classification based on the VI path of an equipment.

This work differs from the other works by the fact that it possesses a single variable as input (transient power signal), while several authors already mentioned as [6] use current harmonics, current waveform, active and reactive power. In the context of the deep neural networks cited in

[7-10], this study performed better than the evaluation metrics used. This is due to the difficulty of other methods in classifying multi-state appliances, such as the dishwasher and the washing machine. In addition, the CNN already presented in the literature need to transform the transient signal of each equipment in an image (spectrogram), to extract the characteristics of the image through the intensity of the colors and finally to make the classification, while in our approach we use directness the power signal, causing our CNN to interpret these values as being the color intensity in an image.

## 2.2. Evaluating NILM Algorithms

In order to evaluate the performance of the proposed system, some evaluation metrics have been used that are generally used to evaluate equipment identification systems in the context of NILM systems:

**Confusion Matrix:** Allows an effective measure of the classification model, presenting the number of correct classifications versus classifications predicted for each class, on a set of examples [9]. The main diagonal presents for each class the correct classification number and the percentage that this number represents within the complete number of data of the class.

**Accuracy:** presents the percentage of positive and negative samples correctly classified on the sum of positive and negative samples.

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

Being True positive (TP), the number of times an equipment is correctly classified as ON; True Negative (TN), the number of times an equipment is correctly classified as OFF; False Positive (FP) The number of times an equipment is incorrectly classified as ON and False Negative (FN) is the number of times an equipment is incorrectly classified as OFF.

**Sensitivity:** percentage of positive samples correctly classified on the total of positive samples.

$$\text{Sens} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{TP}}{\textit{Positive}} \quad (2)$$

**Precision:** percentage of positive samples correctly classified on the total of samples classified as positive.

$$\text{Prec} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

**F-score:** It is a weighted average of precision and sensitivity

$$\text{F - score} = \frac{2 \times (\text{Prec} \times \text{Sens})}{(\text{Prec} + \text{Sens})} \quad (4)$$

## 3. CONVOLUTIONAL NEURAL NETWORK

A convolutional neural network (CNN) can be considered as a variant of the neural network Perceptron of Multiple Layers (MLP). Instead of using fully-connected hidden layers, such as

MLP, the architecture of a CNN is based on the alternation of convolution layers - the layer that names the network; and pooling layers. Each layer will have a set of filters, also known as kernel, that will be responsible for extracting local features from an input. With this, we can create several convolution and pooling maps, containing several specific characteristics like borders, colour intensity, contours and shapes. Each feature map will have a shared set of weights, which decreases the computational complexity of the network [11]. Finally, we have the layer responsible for the classification process, which have the fully connected layer, which connects all the neurons of the layer before it to the output neurons, as shown in Figure 3.

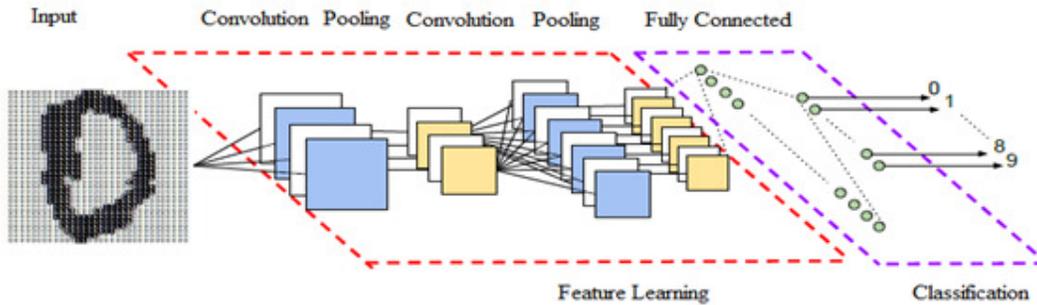


Figure 4 Illustration of the architecture of a CNN [11]

For this approach, which is focused on the classification of equipment through the behavior of its power transients, an architecture based on three layers of convolution followed by pooling was used. Between each convolution and pooling layer normalization is applied in the filter sets (batches), which serves to accelerate network formation and reduce sensitivity for initialization. In addition, we used the non-linear activation function (ReLU) which is simply the identity function for positive values. After the 3 layers of convolution and pooling a fully connected layer is used, followed by the Softmax function. This architecture, derived from a reduction in the convolutional network GoogLeNet [12] (that has five layers of convolution always followed by a pooling), is represented in Figure 4, containing specifications such as: the number of filters in each layer, the size of the stride and the configuration of the output layer.

The convolution layer consists of neurons that are responsible for extracting different sub-region resources from the input images [13]. These areas are derived from the filters used in this layer, being able to extract specific characteristics of the input. In this layer we specify the amount of filters, their sizes, in addition to the stride, which defines the size of the neighbourhood that each layer's neuron will process. [11]

The Pooling layer follows the convolutional layer reducing the number of connections to the following layers, being Max-Pooling in our work. A Max-Pooling layer returns the maximum values obtained in its filters. This layer does not perform any learning, but reduces the number of parameters to be learned in the following layers. [13,11]

The fully connected layer connects all the neurons of the anterior layer with the output neurons, which represent the classes to be classified. This layer combines all the characteristics (local information) learned in previous layers, sweeping the input to identify the highest standards. For our classification problem, it will combine the characteristics of the transients to classify the equipment. At the output of the classification layer, the Softmax activation function is applied which is responsible for performing the multi-class classification (for example: object recognition). [13,11]

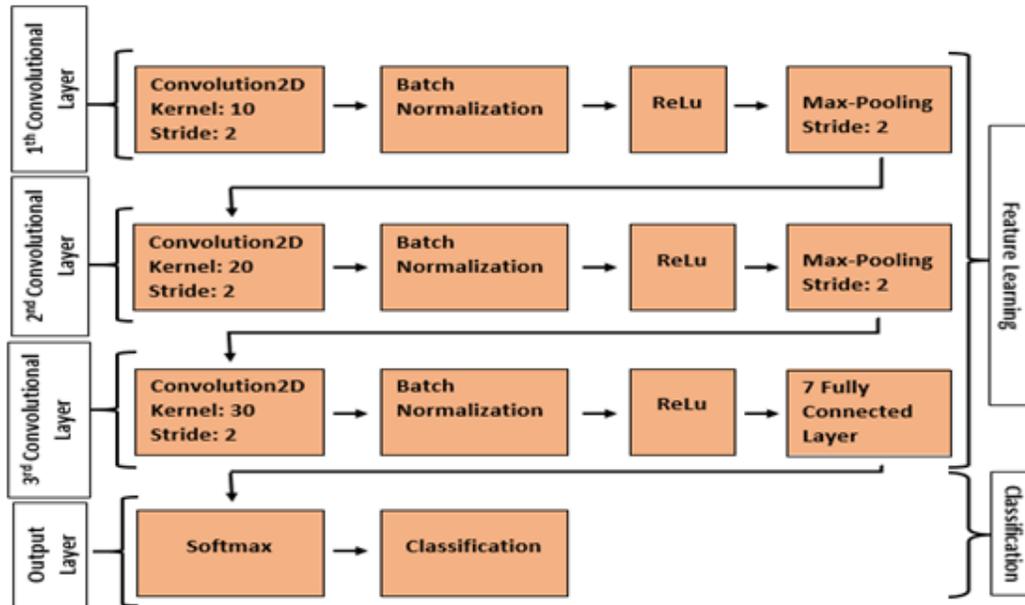


Figure 5 CNN architecture developed for the proposal

### 3.1. CNN Training

The database for developing the identification system had 448 patterns, when all the 7 appliances were taken into account. Each pattern has 7 transient samples for a particular appliance, thus forming a bidimensional matrix (7x448). The data were divided into training, validation and test categories, comprising approximately 60%, 20% and 20% respectively for the total number of patterns. Table 2 shows the arrangement of the data in greater detail.

Table 2. Data Organization.

N°	Equipment	Trai	Valid.	Test	Total
1	Refrigerator	60	14	15	89
2	Microwave	65	14	15	94
3	Stove	67	14	14	95
4	Oven	60	14	13	87
5	Dishwasher	61	12	15	88
6	Air conditioning	55	12	13	80
7	Washer / Dryer	80	15	18	113
#	Overall	448	95	103	646

The approach involves the direct use of 7 samples of power supply transient signals of the appliances as an entry to the CNN, without the need for the application of signal processing to images such as spectrogram [14], or binary images [15]. All that was necessary to achieve this was to re-size the entry of the training matrix to 4D, and thus take on the dimensions of 1x7x1x448, and in this way the CNN can interpret the data as a numerical 4-D matrix (an agglomeration of colored images). While the first three dimensions refer to height, width and channels, the last dimension must index the individual images, or rather, index the transients.

## 4. RESULTS

Table 3 shows the results obtained for the test data, after the training of the projected CNN network. The results are given in the form of metrics: sensitivity (Sens), precision (Prec) and F-score (F). The 3 assessment metrics used in this study can assist us in measuring the performance of the CNN from another perspective. Thus, for example, there are the Oven, Air-Conditioner and Washing-machine which were classified in a precise way, since they had a low amount of FP. However, they did not have the same level of performance for sensitivity, which measures the capacity of the system to predict correctly in the cases that really have it (TP). For this reason, the F-score is used to harmonize the two assessment metrics already mentioned and make a better comparison between the appliances by means of the F-score metric. Hence, it can be seen from the analysis in Column F that the Air-conditioner and Oven had a score above 90%, which demonstrates that the model shown had an excellent performance.

Table 3. Results for Test Data

N	Equipment	Sens.	Prec.	F
1	Refrigerator	0.8667	0.6842	0.7647
2	Microwave	0.9333	0.7000	0.8000
3	Stove	0.6429	0.7500	0.6923
4	Oven	0.8462	1.0000	0.9167
5	Dishwasher	0.6667	0.7692	0.7143
6	Air conditioning	0.9231	1.0000	0.9600
7	Washer / Dryer	0.8889	1.0000	0.9412
#	Overall	0.8239	0.8433	0.8270

Table 4, in turn, shows the results obtained in the training, testing and validation simulations where the metrics used were accuracy and the F-score. On the basis of these results, it can be noted that although we are confronted with a complex classificatory problem, involving multi-stage types of equipment, the CNN on average, had a general rate of accuracy of 82.43% and an F-score of 82.46%, which are very promising results.

Table 4. Performance Results

Simulation	Acc.	F
Training	0.8795	0.8785
Validation	0.7684	0.7685
Test	0.8252	0.8270
Geral	0.8243	0.8246

Figure 5 shows the confusion matrix obtained for the test data which thus allows a broader view of the performance of our algorithm, as well as providing a detailed account of the results obtained in Table 3. The 6 appliances are defined as follows: Refrigerator (1), Microwave (2), Stove (3), Oven (4), Dishwasher (5), Air Conditioning (6) and washer/dryer (7). Each matrix column represents the categories of appliances predicted by the CNN, while the lines represent the real categories. The number of checks for each class can be found on the main diagonal of the

matrix. Thus, it can be inferred that the appliances that have FN values, had a reduction of sensitivity, such as the Stove and Dishwasher. The Refrigerator had the worst rate of precision owing to the fact that this appliance had had a high FP value, with 6 FP values and 13 TP values. However, the Washer-Dryer and Oven did not have any FP values, and attained a 100% precision rate.

1	13 12.6%	1 1.0%	0 0.0%	0 0.0%	4 3.9%	1 1.0%	0 0.0%	68.4% 31.6%
2	0 0.0%	14 13.6%	4 3.9%	1 1.0%	1 1.0%	0 0.0%	0 0.0%	70.0% 30.0%
3	0 0.0%	0 0.0%	9 8.7%	1 1.0%	0 0.0%	0 0.0%	2 1.9%	75.0% 25.0%
4	0 0.0%	0 0.0%	0 0.0%	11 10.7%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
5	2 1.9%	0 0.0%	1 1.0%	0 0.0%	10 9.7%	0 0.0%	0 0.0%	76.9% 23.1%
6	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	12 11.7%	0 0.0%	100% 0.0%
7	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	16 15.5%	100% 0.0%
	86.7% 13.3%	93.3% 6.7%	64.3% 35.7%	84.6% 15.4%	66.7% 33.3%	92.3% 7.7%	88.9% 11.1%	82.5% 17.5%
	1	2	3	4	5	6	7	

Figure 6 Confusion Matrix for Test Data

#### 4.1. COMPARISON WITH STATE OF THE ART

In this section, we compare our results with some state-of-the-art NALM algorithms, proposed for low sampling rates and active power measurements. Table 5 presents the results of some systems already developed to identify equipment in NILM systems using as input the power transient measurements for low frequency. A direct comparison of results should be carried out with caution since for all the presented systems one has the database used for different training, test and validation and equipment and number of equipment also identified different.

Table 5. Comparison between systems presented in the literature

Authors	Technique	N° of Appliance categories	Sens	Prec	F	Acc
This Study	CNN	7	0.82	0.84	0.82	0.82
Kelly [7]	Autoencoder	5	0.80	0.58	0.55	0.91
Kelly [7]	LSTM	5	0.69	0.39	0.39	0.68
WONG [16]	PDT	6	0.77	0.76	0.73	-----
Zhao [17]	GSP	8	0.51	0.89	0.64	0.77

[7] Uses long short-term memory; [16] Uses Particle-based Distribution Truncation (PDT) and [17] Uses Graph Signal Processing (GSP).

## 5. CONCLUSIONS

In this article, we describe how to apply CNNs to the recognition of technical equipment in an innovative manner. From the results obtained, the efficiency of the proposed system is clearly evident, where a weighted average of precision and sensitivity was obtained that was higher than 75%; and with an average degree of accuracy of 82%. The results obtained can be regarded as satisfactory when compared with the results of the identification systems already shown in the literature and also when account is taken of the complexity of the system put forward which was designed to identify loads in a multilevel or variable state.

One point that should be stressed with regard to the direct use of the power supply transient signal as an entry to the identification system, is that it speeds up the system. This means that it is a system that can achieve good results in classification by using data where the measured power is of a low frequency. This is beneficial since the use of low frequencies is common in available low-cost measuring devices which are currently being used for the development of NILM systems.

## REFERENCES

- [1] HART, George William. Nonintrusive appliance load monitoring. Proceedings of the IEEE, v. 80, n. 12, p. 1870-1891, 1992.
- [2] FIGUEIREDO, Marisa. Contributions to Electrical Energy Disaggregation in a Smart Home. 2014. Tese de Doutorado. APA. Disponível em: <www:http://hdl.handle.net/10316/24256>. Acessado em: novembro de 2017.
- [3] KOLTER, J. Zico; JOHNSON, Matthew J. REDD: A public data set for energy disaggregation research. In: Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA. 2011. p. 59-62.
- [4] KATO, Takekazu et al. Appliance Recognition from Electric Current Signals for Information-Energy Integrated Network in Home Environments. ICOST, v. 9, p. 150-157, 2009.
- [5] BATRA, Nipun et al. A comparison of non-intrusive load monitoring methods for commercial and residential buildings. arXiv preprint arXiv:1408.6595, 2014.
- [6] WONG, Yung Fei et al. Recent approaches to non-intrusive load monitoring techniques in residential settings. In: Computational Intelligence Applications In Smart Grid (CIASG), 2013 IEEE Symposium on. IEEE, 2013. p. 73-79.
- [7] KELLY, Jack; KNOTTENBELT, William. 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. ACM, 2015. p. 55-64.
- [8] DO NASCIMENTO, Pedro Paulo Marques. Applications of Deep Learning Techniques on NILM. 2016. Tese de Doutorado. Universidade Federal do Rio de Janeiro.
- [9] Wan He and Ying Chai. An Empirical Study on Energy Disaggregation via Deep Learning, in Advances in Intelligent Systems Research, volume 133, 2nd International Conference on Artificial Intelligence and Industrial Engineering (AIIE2016), pp338-341, 2016
- [10] DE BAETS, Leen et al. Appliance classification using VI trajectories and convolutional neural networks. Energy and Buildings, v. 158, p. 32-36, 2018.

- [11] VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: Proceedings of the XXIX Conference on Graphics, Patterns and Images. 2016. p. 1-4.
- [12] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in CVPR 2015, 2015.
- [13] HIJAZI, Samer; KUMAR, Rishi; ROWEN, Chris. Using convolutional neural networks for image recognition. Tech. Rep., 2015. [Online]. Available: <http://ip.cadence.com/uploads/901/cnn-wp-pdf>.
- [14] ABDEL-HAMID, Ossama et al. Convolutional neural networks for speech recognition. *IEEE/ACM Transactions on audio, speech, and language processing*, v. 22, n. 10, p. 1533-1545, 2014.
- [15] Atabay, H.A.: Binary shape classification using convolutional neural networks. *IIOAB J.* 7(5), 332–336 (2016)
- [16] WONG, Yung Fei; DRUMMOND, T.; ŞEKERCIOĞLU, Y. A. Real-time load disaggregation algorithm using particle-based distribution truncation with state occupancy model. *Electronics Letters*, v. 50, n. 9, p. 697-699, 2014.
- [17] ZHAO, Bochao; STANKOVIC, Lina; STANKOVIC, Vladimir. On a training-less solution for non-intrusive appliance load monitoring using graph signal processing. *IEEE Access*, v. 4, p. 1784-1799, 2016.

## **AUTHORS**

B. Sc. Deyvison de Paiva Penha is a Master Student in the Electrical Engineering Graduate Program in Federal University of Pará. He received his bachelor degree in 2009 at the Federal University of Pará.

Prof. Dr. Adriana Rosa Garcez Castro has a Master's degree in Electrical Engineering from the Federal University of Pará in 1995 and a PhD in Electrical Engineering from the Faculty of Engineering of the University of Porto in 2004. She is currently a Professor at the Federal University of Pará. His areas of interest are: Control of Electronic Processes and Computational Intelligence applied to Energy Systems.