Universidad de Alcalá
BIBLIOTECA

Document downloaded from the institutional repository of the University of Alcala: http://ebuah.uah.es/dspace/

This is a postprint version of the following published document:

Diego, L. de, Hernández, Á., Nieto, R. & Pérez, M.C. 2022, "Comparison of Neural Networks for High-Sampling Rate NILM Scenario", in 2022 IEEE International Symposium on Medical Measurements and Applications (MeMeA), 22-24 june 2022, Messina, Italy, pp. 1-6. DOI: 10.1109/MeMeA54994.2022.9856406.

Available at http://dx.doi.org/10.1109/MeMeA54994.2022.9856406

© 2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works.

(Article begins on next page)

BY NC ND

This work is licensed under a

Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Comparison of Neural Networks for High-Sampling Rate NILM Scenario

Laura de Diego-Otón[1], Álvaro Hernández[1], Rubén Nieto[2], Mª Carmen Pérez-Rubio[1]

[1] Electronic Department, University of Alcalá
Alcalá de Henares (Madrid), España
[2] Electronics Technology Department, Rey Juan Carlos University
Móstoles (Madrid), España

laura.diego@edu.uah.es; ruben.nieto@urjc.es; {alvaro.hernandez, mcarmen.perezr}@uah.es

***Abstract*— The common objective of techniques employed to identify the use of household appliances is related to energy efficiency and the reduction of energy consumption. In addition, through load monitoring it is possible to assess the degree of independence of tenants with minimal invasion of privacy and thus develop sustainable health systems capable of providing the required services remotely. Both approaches should initially deal with the load identification stage. For that purpose, this work presents three different solutions that take the events of the electrical current signal acquired at high frequency and process them for classification by using two different topologies of Artificial Neural Networks (ANN). The data of interest used as input for the ANN in the proposals are the normalized signal captured around the events, the images created by dividing that signal into sections and organizing them in a matrix, and the images coming from the Short Time Fourier Transform (STFT) of the signal around the event. The dataset BLUED is used to carry out the validation of the proposal, where some of the proposed architectures obtain an F1 score above 90% for more than fifteen devices under classification.**

Keywords*— Ambient Intelligence for Independent Living (AIIL), Non-Intrusive Load Monitoring (NILM), Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN).

I. INTRODUCTION

Non-Intrusive Load Monitoring (NILM) techniques [1] analyse changes in the electrical measurements to derive which specific household appliances are used at each instance. It is common that these aggregate consumption signals are acquired at a single point at the entrance of the house. Currently, Smart Meters (SM) are used for this issue. Therefore, since the installation of individual monitors on each appliance is no longer required, NILM might be considered a low-cost alternative. Previous works [2] have already shown that sampling frequencies in the range of kHz or even more are more suitable for this disaggregation. The main reason is that these rates allow the extraction of more significant features from the signals, which can be used to distinguish the loads connected to the mains that have lower power consumption.

Utility companies use these techniques to study energy consumption in different homes and, in addition, they are often used by their tenants to reduce that consumption or to improve energy efficiency [3]. However, from the information obtained, it is also possible to infer the behaviour patterns of people who live there, establishing a relationship between the loads and their main routines [4]. Some of these actions are called Activities of Daily Life (ADL) [5] and may be used to evaluate the quality of living. Basic ADLs are those skills required to handle basic physical needs, including ambulating, feeding, or personal hygiene, among others. Those activities that involve the use of electrical devices might be monitored in order to know if the person is able to stay in charge of their own live for as long as possible, contributing to the active ageing promotion [6]. Consequently, NILM has become an additional tool in the context of the so-called Ambient Intelligence for Independent Living (AIIL) [7]. These systems support the development of remote health systems to improve the independent living of elderly or even people with mild cognitive impairments. In addition to the low cost of NILM, a negligible intrusiveness is a determining aspect that makes these techniques more interesting for users.

To identify the appliances, the changes that their own operation generates in the electrical signals (events) should be recognized and classified. This is known as an event-based approach, and it may be based on different methods, such as Artificial Neural Networks (ANN) [8], probabilistic methods [9], or Principal Component Analysis (PCA) [10], among others. ANN-based approaches have already been successfully implemented for this purpose. For example, recurrent neural networks (RNN) are commonly used in energy disaggregation, especially with sequential data, due to their ability to obtain the relationship over time [11]. On the other hand, Convolutional Neural Networks (CNNs) are able to extract the most relevant features from an image by applying different filters [12] [13]. Furthermore, the methods applied to this process might use two learning alternatives, supervised or unsupervised, whether or not the ground-truth data is available [14].

This work presents three architectures focused on identifying the loads by using the temporal windows around an event detected on the incoming electrical signals. All of them have a first stage, where the input samples are adapted to the following processing. Afterwards, a second stage particularly deals with the load classification by means of ANNs. The first architecture is based on RNN, and the other two on CNN, but using different types of features from the samples: either time domain or frequency domain. These architectures have been verified using real data obtained from the public Building-Level fUlly labelled Electricity Disaggregation (BLUED) database [15]. This database contains electrical energy measurements at both low and high sampling rates, as well as the ground-truth labels for the on/off events. The experimental classification results show an F1-score around 90% for sixteen or seventeen appliances (or groups of appliances). The main contribution of this work is the comparative evaluation of the household load disaggregation by these three ANN-based architectures, considering the high-frequency electric current as signal of interest. The outline of the manuscript is as follows: Section II describes the two main stages of the proposed recurrent architecture and presents its performance, whereas Section III deals with the CNN architectures; Section IV compares the results from the three approaches, as well as their computational complexity; and, finally, conclusions and future works are discussed in Section V.

II. Proposed Recurrent Architecture

The first proposed architecture is based on the one presented for a low sampling frequency in [16], but, in this case, the analysis will be carried out at a high sampling frequency. It consists of a first pre-processing stage, in which the input samples of the electric current are prepared, and a second stage for load classification, where these samples are processed by an RNN. The details for each part are presented below.

A. Pre-Processing Stage

Information about the effects produced by the usage of the different household appliances is extracted from the electrical current obtained from the BLUED database [15]. The electrical signal has been sampled at a frequency of 12 kHz; however, in order to reduce the amount of data to be processed and adapt to the operation of commercial Analog Front-Ends (AFE), this frequency is reduced to 4 kHz by taking one out of every three samples. Afterwards, the signal is normalized by adjusting the values with respect to a common scale. The range is between -1 and +1 since the cells of the architecture use activation functions that saturate at those limits. Consequently, unless they are avoided, vanishing gradient problems might appear.

The inputs of the classification stage are temporal windows of the normalized signal captured around on/off events from the different electrical appliances. The location of the events is carried out by using the ground-truth provided by the database. The length of these temporal windows is 4096 samples. An illustrative example is provided in Fig. 1.

The classification algorithms require large amounts of data for a proper training. As there may not be enough samples for certain devices, some techniques involving the introduction of noise have been used to increase the size of the dataset. This step will only affect those classes that have a lower number of detected events. The total set of samples has been divided into three independent subsets (training, testing and validation) to address the learning process. The 50% of the available data corresponds to the training subset and the rest is divided equally into testing and validation, each one with 25%.

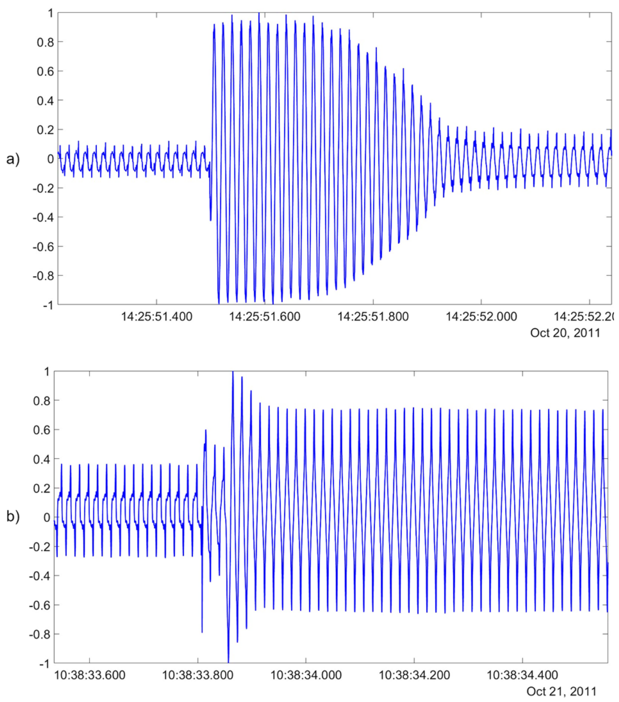

Fig. 1. Temporal window of: a) a fridge on event; and b) a garage door on event from the BLUED database.

B. Load Classification Stage

To classify the load profile, a recurrent topology based on Long Short-Term Memory (LSTM) cells is proposed. The structure, shown in Fig. 2, is similar to the first version defined in [16]. The configuration of the network is described below:

- Input Layer: The input size coincides with the dimension of the temporal windows.
- LSTM 1 Layer: The number of hidden units corresponds to 1/20 the input sequence length. The activation function is the hyperbolic tangent function (tanh).
- LSTM 2 Layer: With respect to the previous LSTM layer, the number of hidden units is halved to reduce the treated information.
- Dense Layer: The number of units corresponds to the number of devices involved in the classification process. The activation function is the *SoftMax* (normalized exponential function) [17].

Firstly, the inputs are processed by two stacked LSTM layers. The stacked LSTMs were introduced by Graves, et al. in their application of LSTMs for speech recognition [18]. These layers extract information by computing the hyperbolic tangent function. Additionally, they must be defined so that the first *LSTM* layer provides a sequence output rather than a single-value output to the *LSTM* layer below. For the multiclass classification it is used a *SoftMax* activation function in the last dense layer.

C. Experimental Results

The evaluation of the classification performance has been done by using the electricity consumption data from the BLUED database [15]. In this assessment, not only the appliances that may be of interest to estimate or predict routines within a home have been considered, but also the rest

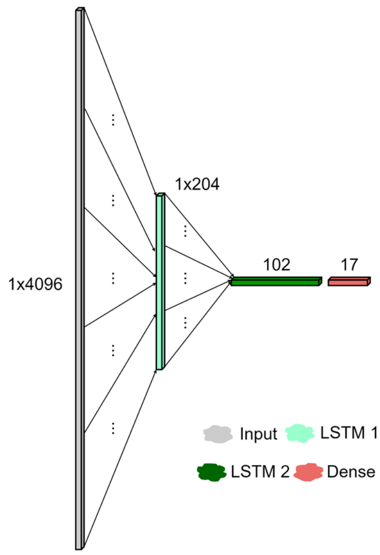

Fig. 2. Structure of the proposed recurrent neural network.

of the electrical devices included in the dataset. The total number of electronic devices that have been considered is seventeen. Regarding the devices that can be observed in Table I, it is worth mentioning that labels 00 and 01, which are the total number of lights in the house and the distribution circuits that the house contains, respectively, involve the events of different electrical appliances. Although there are some appliances that have better characteristics to recognise daily routines, being able to identify all of them allows to discard the less informative ones. Concretely, the main characteristics that make a device a good candidate are that it must be operated manually and used frequently. For instance, the usage of the TV (label 11) provides a good indicator of anomalies in the behaviour such as sleep disorder [6].

Before presenting the obtained results, the training of the model is carried out by using the Adam algorithm [19], with an initial learning rate of 0.0001, which will be varied during the process using a decay rate of 10^{-6}. The duration of the learning process is set by the maximum number of epochs. This maximum number defined for this process is 200, however, this time may be shortened if the performance obtained in the validation does not improve. Furthermore, the number of batches is 100; since the division of samples per batch may not be exact, the samples are shuffled avoiding discarding the same samples at each iteration.

In order to consider the possible limitations of the model, some variations have been made. Firstly, it must be considered that there is an imbalance in the number of samples for each device, since not all of them are used with the same frequency. To avoid that this disproportion could affect the algorithm's performance, the model will provide different weights during the learning process, in such a way that a higher value is established for the minority classes, whereas the weight is reduced for the majority ones. The other variation implies discarding the samples from label 01. The main reason is that this class includes the events of the housing distribution circuits that may contain events from other classes already present in the classification. This could mislead the classification results.

Table II compares the results of the performed tests. It is necessary to mention that, although there are more metrics, the accuracy and F1-score will be used to measure the classification performance. In this way, whereas accuracy corresponds to the ratio of correct predictions, the F1 Score is formed by the combination of precision and recall values, giving them equal importance. In addition, the F1-score is often more useful than accuracy, especially when there is an uneven class distribution [20]. As can be observed, the different variations made from the main structure present a similar result, although, on one hand, discarding the circuits seems to improve the metrics except for accuracy. Furthermore, adding class weights behaves the other way around.

III. Proposed Convolutional Architectures

As in the previous section, the proposed convolutional architectures consist of two stages: first, input samples are prepared and then, through their processing, the devices corresponding to these samples are identified. Hereinafter, the differences and similarities with respect to the previous architecture are discussed.

TABLE I. Selected devices for the proposal evaluation.

Label	Devices
00	Lamps and lights
01	Circuits
02	Garage door
03	Kitchen aid chopper
04	Fridge
05	A/V Living room
06	Computer 1
07	Laptop 1
08	DVD player basement
09	Air compressor
10	LCD Monitor 1
11	TV basement
12	Printer
13	Hair Dryer
14	Iron
15	Empty living room socket
16	Monitor 2

TABLE II. Experimental results for the proposed recurrent architecture.

Configuration	Accuracy	Precision	Recall	F1-score
Baseline	99.75 %	95.24 %	97.41 %	96.11 %
Adding class weights	99.76 %	94.97 %	97.54 %	95.97 %
Discarding label 01	99.67 %	96.29 %	97.67 %	96.93 %

A. Pre-Processing Stage

The topology used in this case is convolutional. This type of neural network usually works with images, therefore, temporal windows used by the LSTM neural network should be transformed into images. Regarding the creation of images, two different approaches have been considered. The first one consists of splitting the temporal window into 64 sections with a length of 64 samples, with the purpose of obtaining images with 64×64 square dimensions. The sections were entered into the rows of a matrix, which will be finally transformed into a grayscale image. This procedure was presented in [21] and Fig. 3 shows an example of the resulting images.

The second method includes information about the spectral content of the signal. The process used to obtain the images is the Short Time Fourier Transform (STFT) of the temporal windows [22] [23]. The outcomes are matrices made up of complex numbers. In order not to lose information when transforming them into images, the values should be divided. Two different forms of division have been considered: distinguishing between magnitude and phase, or between the real and imaginary parts. The resulting values are normalized and expressed in grayscale to finally create the input images such as those shown in Fig. 4.

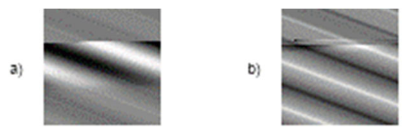

Fig. 3. Time sequence imagen of: a) a fridge on event; and b) a garage door on event.

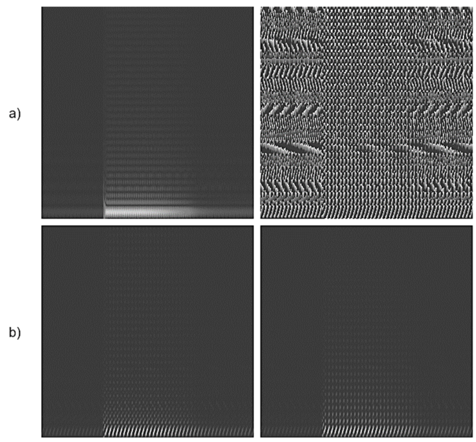

Fig. 4. STFT images of: a) the magnitude (left) and phase (right); and b) the real (left) and imaginary (right) parts for a fridge on event.

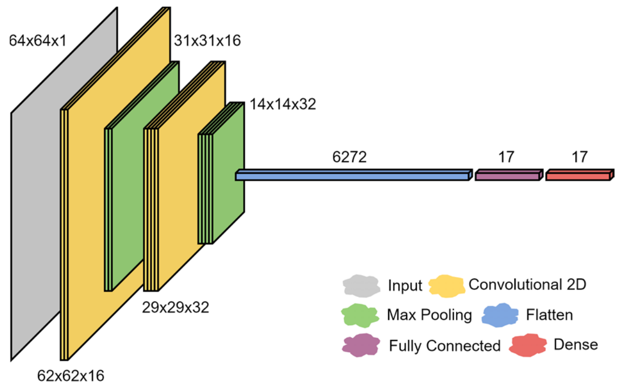

Fig. 5. Structure of the proposed convolution neural network that uses time sequence images as inputs.

In both cases, the total set of samples has been used, including those samples artificially created to compensate for the class imbalance. For the classification stage, the samples have been divided in the same way into three different subsets.

B. *Load Classification Stage*

The algorithms proposed for load disaggregation are CNNs. Since two different pre-processing have been defined, there will be two versions, although both will be related. The first structure, shown in Fig. 5 [21], uses the input images built from time-domain information. The layers that form the neural network are:

- Input Layer: The input size corresponds to images size.
- Convolutional 2D Layer: The dimensionality of the output space is 16, with a kernel size of 3×3.
- Max Pooling Layer: The size of the max pooling window is 2×2.
- Convolutional 2D Layer: The dimensionality of the output space is 32, with a kernel size of 3×3.
- Max Pooling Layer: The size of the max pooling window is 2×2.
- Flatten Layer.
- Fully connected Layer: The number of units corresponds to the number of devices used in the classification process. The activation function is the *ReLU* (Rectified Linear Unit) [24].
- Dense Layer: The number of units corresponds to the number of devices used in the classification process. The activation function is the *SoftMax*.

The images are introduced into the neural network to be processed by two *Convolutional 2D* layers, with which the main characteristics of the events are extracted. It contains more than one *Convolutional 2D* layer to increase the ability to extract more complex features than those that could be learned by using a single layer. *Max Pooling* layers have been included after convolutional layers in order to reduce the size of the filtered images and, consequently, those of the subsequent layers. In addition, these layers allow only the most relevant features to prevail. Once the necessary information is available, the *Flatten* layer resizes the output dimensions so that this information can be entered into the Fully Connected layer. The *Fully Connected* layer is responsible for obtaining the probability that each characteristic extracted in the previous layers may correspond to a specific class. The final classification is carried out by the last layer, defined by the *SoftMax* activation function. The dimensions of each layer are due to the image size.

On the other hand, as can be observed in Fig. 6, the neural network used in the processing of spectral images is made up of two equal parts, whose structure is based on the previous algorithm. However, since the size of the input images is much larger than that used in the other analysis, the size of the structure is enlarged. The layers forming the neural network are:

- Input Layer: The input size corresponds to the images size.
- Convolutional 2D Layer: The dimensionality of the output space is 16, with a kernel size of 3×3.
- Max Pooling Layer: The size of the max pooling window is 2×2.
- Convolutional 2D Layer: The dimensionality of the output space is 32, with a kernel size of 3×3.
- Max Pooling Layer: The size of the max pooling window is 2×2.
- Convolutional 2D Layer: The dimensionality of the output space is 64, with a kernel size of 3×3.
- Max Pooling Layer: The size of the max pooling window is 2×2.
- Concatenate Layer.
- Flatten Layer.
- Fully connected Layer: Reduce the number of units with respect to the previous output layer. The activation function is the *ReLU*.
- Dense Layer: The number of units corresponds to the number of devices used in the classification process. The activation function is the *SoftMax*.

The main difference is that, after the extraction phase performed by three *Convolutional 2D* and three *Max Pooling* layers to minimize the size of the filtered images, a *Concatenate* layer is used to connect both parts so that the processed information of each pair of images is used jointly in the final classification stage.

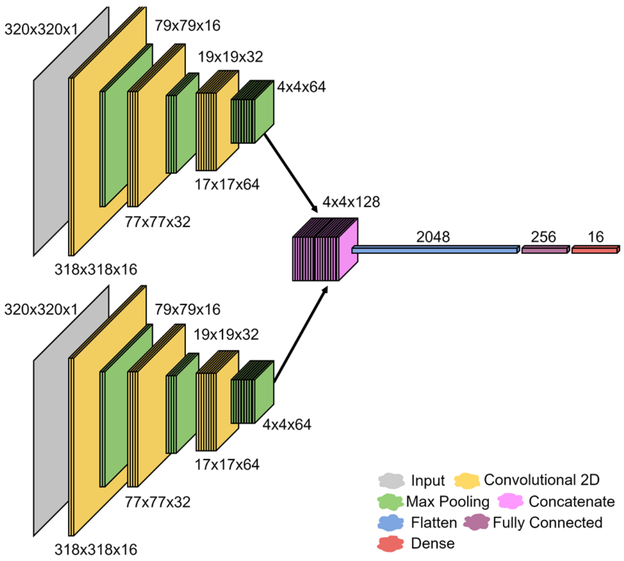

Fig. 6. Structure of the proposed convolution neural network that uses STFT images as inputs.

C. *Experimental Results*

For classification, the involved electrical devices are the same as before for the recurrent proposal; however, the number of appliances involved can be sixteen or seventeen, since the distribution circuits are excluded because of their negative influence on the results. In addition, the hyperparameters used in this topology are the same as those used with the recurrent architecture, also including the option of adding class weights.

Table III presents the comparison between the results from the proposed convolutional architectures. The metrics of the architecture that uses the images resulting from the STFT process are worse than those obtained with the images extracted directly from the temporal windows, especially for the F1-score.

IV. DISCUSSION

After detailing the classification performance for all the proposals described previously, it is also worth considering the computational complexity for every proposal. This aspect is associated with the number of learning parameters, provided in Fig. 7, which define the internal configuration of the model. They depend on the number of layers, as well as on their dimensions. The architecture with the highest number of learning parameters is the one based on RNN, whereas the parameters of convolutional structures roughly correspond to 1/7 of that number. The difference between each CNN is mainly due to the size of the input images.

For further comparison of the classification performance, Fig. 8 shows the experimental metrics of the three architectures, including the distinction of the two ways of representing the information obtained from the STFT of the images. It shows that the recurrent method gives better results in all metrics, compared to the convolutional approaches, although the one that uses time sequence images obtains similar results. It can also be observed, as discussed above, that the worst results are those achieved by the architecture that uses the images resulting from the STFT process.

TABLE III. EXPERIMENTAL RESULTS FOR THE PROPOSED CONVOLUTIONAL ARCHITECTURES.

Architectures	Accuracy	Precision	Recall	F1-score
Convolutional Time Sequence Images	99.70 %	92.17 %	91.78 %	91.91 %
Convolutional STFT Images (Magnitude and Phase)	97.17 %	79.93 %	66.39 %	70.34 %
Convolutional STFT Images (Real and Imaginary Parts)	96.91 %	73.03 %	67.07 %	68.45 %

Finally, a comparison with other previous works is considered here. The use of a LSTM-RNN model in [11] obtains an average F1-score of 85.76 % for its application on the UK-DALE dataset and a 90.48 % on the REDD dataset, classifying up to fifteen appliances. In that work the information used for disaggregation is low-frequency signals. Alternatively, some convolutional approaches are presented in [12] and [13]. In [12] voltage and current changes from the V-I trajectories are analysed and the obtained results are 78.16 % and 66.01 % in F1-score, respectively, for eleven appliances from the two subsets of data contained in the PLAID database. On the other hand, in [13] the spectral information of the signal is included to observe its influence, besides the changes in time domain. The combination of both sources of information achieves an F1-score of 99.8 %, using the same database as the one used in this work and distinguishing the on/off events of 34 different types of devices.

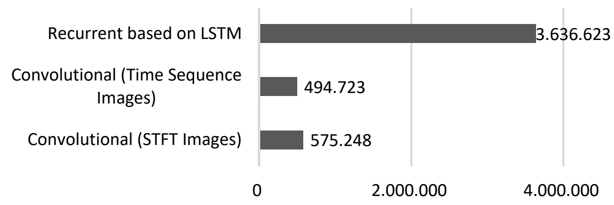

Fig. 7. Number of learning parameters for the three architectures.

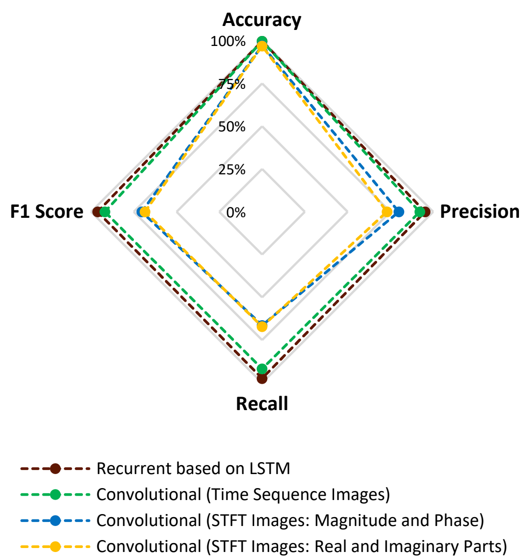

Fig. 8. Comparison of the performance of the architectures with different inputs.

V. CONCLUSIONS

In this work, three different architectures for the load classification in NILM have been proposed. Their common feature is that all the three are supervised event-based solutions, using the electrical current sampled at high frequency as the source of information. RNN and CNN have been compared, and, in the case of CNN, training data have been considered in the time and frequency domains. The experimental results obtained from the BLUED dataset indicates the proposed architectures achieve a suitable classification performance, similar to those provided by previous works. The computational complexity of the three architectures has also been taken into account, since it might influence the future implementation.

Future works will deal with the design of an unsupervised architecture able to classify the appliances without having any a priori information about in the existing categories in the dataset. These techniques might allow the proposal to become closer to a real situation, where the useful life of household appliances is limited, and they are often replaced.

ACKNOWLEDGMENT

This work has been possible thanks to the Comunidad de Madrid (RACC project, CM/JIN/2021-016), and the Spanish Ministry of Science, Innovation and Universities (PoM project, PID2019-105470RA-C33).

REFERENCES

[1] J. Revuelta Herrero, Á. Lozano Murciego, A. López Barriuso, D. Hernández de la Iglesia, G. Villarrubia González, J. M. Corbacho Rodríguez, and R. Carreira, “Non Intrusive Load Monitoring (NILM): A State of the Art,” in *Proc. 15th International Conference on Practical Applications of Agents and Multi-Agent Systems*, pp. 125-138, 2018, doi: 10.1007/978-3-319-61578-3_12.

[2] A. Siddiqui and A. Sibal, “Energy Disaggregation in Smart Home Appliances: A Deep Learning Approach,” *Energies, Elsevier*, 2020, in press ⟨hal-02954362⟩.

[3] J. He, Z. Zhang, L. Zhu, Z. Zhu, J. Liu and K. Gai, "An Efficient and Accurate Nonintrusive Load Monitoring Scheme for Power Consumption," in *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 9054-9063, 2019, doi: 10.1109/JIOT.2019.2926815.

[4] A. Ruano, A. Hernandez, J. Ureña, M. Ruano, and J. Garcia, “NILM Techniques for Intelligent Home Energy Management and Ambient Assisted Living: A Review,” *Energies*, vol. 12(11), pp. 2203, 2019, doi: 10.3390/en12112203.

[5] C. M. Giebel, C. Sutcliffe and D.Challis, “Activity of daily Living and Quality of Life across different Satages of Dementia: A UK study,” *Aging & Mental Health*, vol. 19, no. 1, pp. 63-71, 2015, doi: 10.1080/13607863.2014.915920.

[6] Bousbiat, Hafsa et al. “Ageing Safely in the Digital Era: A New Unobtrusive Activity Monitoring Framework Leveraging on Daily Interactions with Hand-Operated Appliances.” *Sensors*, vol. 22(4), 1322, 2022, doi:10.3390/s22041322.

[7] T. Kleinberger, M. Becker, E. Ras, A. Holzinger and P. Müller, “Ambient Intelligence in Assisted Living: Enable Elderly People to Handle Future Interfaces,” in *Proc. 12th International Conference on Universal Access in Human-Computer Interaction*, pp. 104-111, 2007, doi: 10.1007/978-3-540-73281-5_11.

[8] S. Ranjan Kabat, D. Jayashree, K.G.S. Venkatesan, and H. Budaraju Venkata, “Non-intrusive load monitoring technique using deep neural networks for energy disaggregation”, *Materials Today: Proceedings*, 2021, doi: 10.1016/j.matpr.2021.01.192.

[9] J. A. Mueller and J. W. Kimball, "Accurate Energy Use Estimation for Nonintrusive Load Monitoring in Systems of Known Devices," in *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2797-2808, 2018, doi: 10.1109/TSG.2016.2620120.

[10] J. M. Alcalá, J. Ureña, Á. Hernández and D. Gualda, "Sustainable Homecare Monitoring System by Sensing Electricity Data," *in IEEE Sensors Journal*, vol. 17, no. 23, pp. 7741-7749, 2017, doi: 10.1109/JSEN.2017.2713645.

[11] J. Kim, T. Le and H. Kim, “Nonintrusive Load Monitoring Based on Advanced Deep Learning and Novel Signature” in *Computational Intelligence and Neuroscience*, pp. 1-22, 2017, doi: 10.1155/2017/4216281.

[12] D. Baptista, S. Mostafa, L. Pereira, L. Sousa, and F.Morgado-Dias, “Implementation strategy of convolution neural networks on field programmable gate arrays for appliance classification using the voltage and current (V-I) trajectory”, *Energies*, vol. 11(9), pp. 2460, 2018.

[13] F. Ciancetta, G. Bucci, E. Fiorucci, S. Mari and A. Fioravanti, "A New Convolutional Neural Network-Based System for NILM Applications," in *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-12, 2021, doi: 10.1109/TIM.2020.3035193.

[14] R. Sathya and A. Abraham, “Comparison of supervised and unsupervised learning algorithms for pattern classification,” *International Journal of Advanced Research in Artificial Intelligence*, vol. 2(2), pp. 34-38, 2013, doi: 10.14569/IJARAI.2013.020206.

[15] K. Anderson, A. F. Ocneanu, D. Benítez, D. Carlson, A. Rowe and M. Bergés, “Blued: a fully labelled public dataset for event-based nonintrusive load monitoring research,” in *Proc. 2nd KDD Workshop on Data Mining Applications in Sustainability*, pp. 1-5, 2012.

[16] L. de Diego-Otón, D. Fuentes-Jimenez, Á. Hernández and R. Nieto, "Recurrent LSTM Architecture for Appliance Identification in Non-Intrusive Load Monitoring," in *2021 IEEE International Instrumentation and Measurement Technology Conference*, pp. 1-6, 2021, doi: 10.1109/I2MTC50364.2021.9460046.

[17] A. Martins and R. F. Astudillo. “From softmax to sparsemax: A sparsemodel of attention and multi-label classification,” in *Proc. 33rd International conference on Machine Learning*, vol. 48, pp. 1614-1623, 2016.

[18] A. Graves, A.-r. Mohamed and G. Hinton, “Speech recognition with Deep Recurrent Neural Networks,” in *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 6645-6649, 2013.

[19] D. Kingma and J. Ba, “Adam: a method for stochastic optimization,” in *Proc. 3th International Conference on Learning Representations*, pp. 1-15, 2015.

[20] H. He and Y. Ma, "Foundations of Imbalanced Learning," in *Imbalanced Learning: Foundations, Algorithms, and Applications*, IEEE, pp.13-41, 2013, doi: 10.1002/9781118646106.ch2.

[21] L. de Diego-Otón, Á. Hernández, R. Nieto and C. Pérez-Rubio "Evaluación de una Arquitectura CNN para la Identificación de Cargas en NILM," in *28th Seminario Anual de Automática, Electrónica Industrial e Instrumentación*, pp. 116-121, 2021.

[22] W. Yuegang, J. Shao, X. Hongtao, "Non-stationary signals processing based on STFT," in *8th International Conference on Electronic Measurement and Instruments*, pp. 3-301–3-304, 2007, doi: 10.1109/ICEMI.2007.4350914.

[23] S. Zhang, D. Yu, S. Sheng, "A discrete STFT processor for real-time spectrum analysis," in *APCCAS 2006 – 2006 IEEE Asia Pacific Conference on Circuits and Systems*, pp. 1943–1946, 2006, doi: 10.1109/APCCAS.2006.342241.

[24] A.F. Agarap, “Deep learning using rectified linear units (ReLU)”, 2018, ArXiv abs/1803.08375.