Using the Power Delay Profile to Accelerate the Training of Neural Network-Based Classifiers for the Identification of LOS and NLOS UWB Propagation Conditions

VALENTÍN BARRAL VALES [ID], **TOMÁS DOMÍNGUEZ-BOLAÑO** [ID],
CARLOS J. ESCUDERO, (Member, IEEE), AND JOSÉ A. GARCÍA-NAYA [ID], **(Senior Member, IEEE)**
CITIC Research Center, Department of Computer Engineering, University of A Coruña, 15071 A Coruña, Spain

Corresponding author: Valentín Barral Vales (valentin.barral@udc.es)

ABSTRACT ultra-wideband (UWB) technology enables centimeter-level localization systems based on the accurate estimation of the actual distance between transmitter and receiver, by means of the precise estimation of the signal time-of-flight. However, this is only possible when correctly detecting the first path of the incoming signal instead of a bounce or a reflection, which becomes challenging in non line-of-sight (NLOS) situations. There are many different approaches in the literature to alleviate the wrong detection of the first incoming UWB signal path. One of them considers machine learning techniques to design classifiers capable of distinguishing between line-of-sight (LOS) and NLOS propagation from available signal features. However, the performance and complexity of the obtained classifiers depend largely on the size of the input data associated to such features. Thus, features such as the channel impulse response (CIR) produce large amounts of data, yielding very complex classifiers. In this paper, we propose using a downsampled power delay profile (PDP) as an alternative feature consisting of input data much smaller than the CIR, although sufficiently representative, hence resulting in a lower computational cost while exhibiting a similar classification performance. Furthermore, another of the tasks addressed in this work is the study of the impact on the classification results of using a dataset for training where the samples of each class are not balanced from the point of view of energy. Finally, this work also studies how the classifiers based on the CIR or the PDP improve their performance when considering additional signal features such as the estimated range value or its energy level.

INDEX TERMS Channel impulse response, power delay profile, convolutional neural network, deep learning, indoor localization, non line-of-sight, ultra-wideband, ranging, received signal strength.

I. INTRODUCTION

Indoor positioning has experienced great advances in recent years, driven by an increasing number of commercial technology solutions capable of achieving positioning with centimeter accuracy. One of the radio technologies that achieves the best results in terms of precision is UWB, above others traditionally used [1], [2]. Among the

The associate editor coordinating the review of this manuscript and approving it for publication was Yeliz Karaca [ID].

commercial implementations of UWB for localization purposes, the DW1000 integrated circuit (IC) from Decawave [3] stands out for its high availability in many devices. Nowadays, this chip has become almost a de facto standard both at the enterprise-level (indeed, Decawave has more than 50 related partners [4]) and at the research-level (searching for the term "DW1000" in Google Scholar yields almost 8000 different results).

The DW1000 implements the ultra-wideband (UWB) IEEE802.15.4-2011 standard, which details the mechanisms

and protocols needed to operate with UWB in ranging tasks (distance estimation). In order to obtain signal synchronization and estimate the distance between transmitter and receiver, the DW1000 IC implements a threshold-based energy detection mechanism consisting in accumulating energy at the receiver until exceeding a certain predefined threshold value. In situations with a clear line-of-sight (LOS) between transmitter and receiver, the energy of the originally emitted signal is detected and the first path of the incoming signal is accurately detected. However, in non line-of-sight (NLOS) situations, the original signal can be attenuated or blocked completely, so that the energy exceeding the detection threshold in the receiver might not correspond to that of the first path, but that of a bounce or a delayed reflection, yielding an erroneous estimation of the signal time of flight (ToF), which later translates into an error in the range estimation much higher than that in the LOS case. Many different works in the literature try to differentiate the signals received under LOS propagation conditions with respect to those corresponding to NLOS. In general, selectively discarding the estimates coming from NLOS propagation conditions yield a significant improvement of the positioning accuracy.

There are several approaches to this LOS-NLOS classification problem, including those proposed by the authors in previous works. More specifically, in [5], [6] we consider classical machine learning techniques and shallow neural networks to attempt this classification. For this purpose, several measurement campaigns were carried out with devices equipped with the DW1000 IC to obtain a sufficient set of training samples. Later, the performance of these systems was analyzed when classifying the measurements coming from a different scenario. On that occasion, and in order to implement the classification system in computationally reduced devices, it was decided to use only received signal strength (RSS) moving averages and range estimation as the training features. In addition, to carry out the classification in real time, the idea of using the channel impulse response (CIR) samples to feed the training process was discarded due to the long time required to extract the CIR measurements corresponding to the ranging processes between the transmitter device and all the possible receiver nodes deployed in the environment and needed to generate a position estimate. However, the LOS-NLOS classification directly from the CIR samples is also useful in scenarios where either that real-time restriction is not present, or the hardware device allows for obtaining and processing the CIR faster than the DW1000.

Although in this work we focus on approaches based on machine learning, it is necessary to emphasize that there are several other parallel approaches also based on the CIR but using the samples in a different way. For example, with more heuristic and deterministic algorithms [7]–[9], or employing statistical analysis [10]. Other related works do not directly employ the CIR, but features derived from the signal such as its energy, the signal detection time instant [11]–[13], or even additional data such as sensor acceleration values [14]. Others

approach the problem exclusively from the simulation point of view [15], [16].

In addition, it is also worth mentioning that these techniques can be applied in other radio technologies [17], [18], although this work is focused on UWB.

Regarding the LOS-NLOS classification problem with UWB, there are some authors that have considered the use of the CIR samples together with machine learning or deep learning techniques to find a solution. In these cases, the process typically consists in employing all the samples from the CIR directly as training features [19], yielding classifiers with a high computational cost due to the large number of data samples provided by the CIR. A possible approach to alleviate this problem consists in employing only a subset of the CIR samples [20].

However, as in any other process where automatic learning techniques are used, in this case the initial data is the key to obtain representative results that can be exploited in final applications. Thus, both [19] and [20] used for their results the same dataset ([21]). This dataset contains 42000 samples extracted from a DW1000 module in LOS and NLOS situations captured in different indoor locations. These locations include two offices, a small apartment, a kitchen with living room, a bedroom, a small workshop, and a boiler room. The dataset is perfectly balanced, with 50% of samples from each of the classes. In addition, in order to avoid possible biases induced by the different locations, the set of samples is randomized. The dataset does not include absolute position references, but it includes the values provided by the DW1000 after a regular ranging operation, such as the range estimation itself, the CIR samples or the index within the CIR where the first path of the UWB signal was detected. This same dataset has been used in the present work, so that the results obtained with the features introduced could be compared with all the other works based on these same measurements.

Similarly, the neural architecture used in this work is very close to that presented in [19], [20]. In this way we guarantee that the differences observed between the different proposals are due exclusively to the set of features selected, since both the training data and the algorithms are the same.

In this article we present three main contributions. First, we perform a comparison in the classification of LOS-NLOS measurements using a downsampled version of the PDP versus using the complete CIR or only a part of it. The idea of this experiment was to check if similar results could be obtained while reducing the input size of the neural network. Section II details the experiment and the results obtained.

The second contribution is to analyze the impact of using the CIR to classify when both classes have a very different level of energy in their samples. Section IV describes the problem and shows the classification results after applying a normalization and noise removal process on the CIR samples.

Finally, the third contribution is to show the performance of a classification system based on the CIR samples when two additional features are added: the range value itself and the

energy value of the samples. Section V shows the experiments performed and the results obtained.

II. USING THE PDP AS TRAINING SET

This section introduces the use of the PDP samples instead of those from the CIR to perform the training of a multilayer convolutional neural network capable of classifying the UWB measurements captured under LOS or NLOS propagation conditions. The samples employed are those corresponding to the dataset available in [21], originally created by the authors of the paper [20]. This dataset includes 42000 UWB measurements obtained in up to 7 different indoor locations, and it has the same number of measurements for both of the LOS and NLOS classes. Notice that, although obtaining the CIR is typically a hardware-dependant task requiring low-level software development, computing the PDP from the CIR is a generic operation that does not depend on the hardware.

Section II-A details the processing of the CIR samples to obtain the PDP, while Section II-B shows the details of the considered convolutional neural network (CNN) and the selection mechanism for the training and test sets.

A. TRAINING FEATURES

The DW1000 stores, for each measurement, 992 samples for the 16 MHz mean pulse repetition frequency (PRF), or 1016 samples for the case of 64 MHz mean PRF [22]. For the dataset considered in [21], a 64 MHz mean PRF was set, hence, for each measurement, we have 1016 valid CIR samples. The DW1000 provides the real and imaginary values for the CIR samples. However, the considered dataset [21] only provides the modulus of the samples, which is defined for a given measurement as

$$|h[n]|, \quad n \in \mathbb{Z} \text{ and } 0 \leq n < 1016. \tag{1}$$

The selected training features derived from the CIR and considered in this work are defined as described below.

1) ALL THE CIR SAMPLES (AS IN [19])

The idea considered in [19] consists of using all the CIR samples as training features in the same way as Eq. (1).

2) FIRST 152 SAMPLES OF THE CIR AFTER THE FIRST PATH (AS IN [20])

As noted in [20], most of the information about the propagation characteristics is contained in the 152 samples of the CIR starting at the one corresponding to the first path. Thus, if the first path starts at index n_0, then the training feature is

$$|h[n]|, \quad n \in \mathbb{Z} \land n_0 \leq n < n_0 + 152 \tag{2}$$

where we assume that $n_0 \geq 0$ and $n_0 + 152 < 1016$.

For each measurement, the index of the first path is also stored by the DW1000 in an internal register which can be accessed later. Although this value could be used to extract the most representative set of samples from the CIR, in our case we decided to use a different algorithm. The reason is

that the DW1000 uses a threshold algorithm to decide which is the first sample of the first path. In an LOS scenario this method works well, but in NLOS scenarios it may sometimes not detect that first path correctly if its energy level is low. That is why, since in our study we do not have the immediacy requirements that the algorithm implemented in the chip has, we decided to perform a CIR processing to try to detect more accurately that first significant sample.

For that matter, we employed a similar method to the one proposed in [23] to calculate the so-called rise time feature. This value is defined as follows:

$$t_{rise} = t_L - t_H \tag{3}$$

where t_H corresponds to the instant when the signal exceeds a high level based on the maximum value of the signal, and t_L corresponds to a later instant when the signal goes below another threshold dependent on the noise level. Mathematically:

$$t_L = \max\{t : |r(t)| \geq \alpha\sigma_n\} \tag{4}$$
$$t_H = \min\{t : |r(t)| \geq \beta r_{max}\} \tag{5}$$

being $r(t)$ the received signal, σ_n the standard deviation of the thermal noise, $r_{max} = \max\{|r(t)|\}$, and α and β two constants selected empirically. Thus, similarly to [23], but in discrete time, we obtain n_H as

$$n_H = \min\{n : |h[n]| \geq \beta h_{max}\} \tag{6}$$

where $h_{max} = \max\{|h[n]|\}$ is the maximum amplitude of the CIR, and $\beta \in [0, 1]$ is a parameter used to select the threshold to detect the first path based on the maximum value of $h[n]$. The value of β is chosen empirically and in our case we used $\beta = 0.4$, manually inspecting the results to check that the first path was accurately detected. It should be also noted that the CIRs have a small rise time n_r from its actual start to the detected index n_H. Therefore, in order to include this rise time in the selected CIR samples, n_0 is obtained as

$$n_0 = n_H - n_r. \tag{7}$$

In our case, we use a value $n_r = 3$, which was also chosen empirically by inspecting the results.

Fig. 1 shows an example of the CIR for one measurement where the two estimates of the beginning of the first path have been marked: the one returned by the DW1000 and the one generated by our algorithm. It can be seen how in this case the estimate included in the measurements does not include a previous sample that is clearly above the noise level.

3) DOWNSAMPLED PDP (PROPOSAL)

Although most of the CIR information is contained in 152 samples (see Eq. (2)), we can reduce this number even further while retaining a similar performance, hence speeding up the training time of the network. This is feasible because: *1)* closely spaced CIR samples use to be highly correlated, and *2)* even though the CIR may exhibit rapid variations, which reduce the correlation of closely spaced samples in
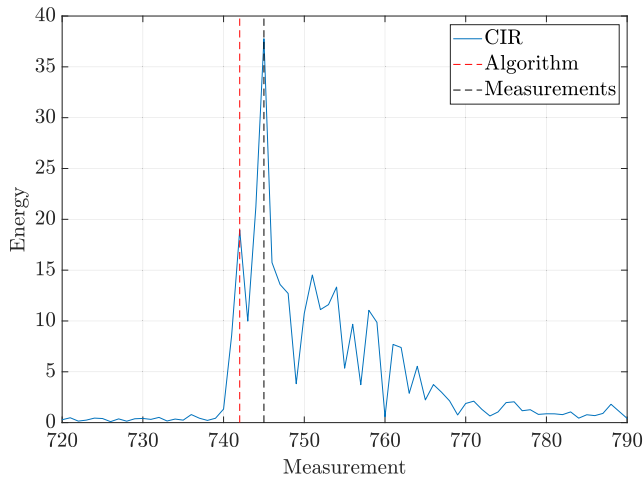
FIGURE 1. First path detection using the value given by the DW1000 and usign a different offline detection algorithm.

FIGURE 2. $|h[n]|^2$ and $p[m]$ from a given measurement with $n_0 = 743$.

some cases, we assume that both the amplitude and the overall shape of the CIR play a more important role in our case.

To do this, we consider the PDP, which gives us the average power of the channel impulse response as a function of the delay. In our case, we can obtain the PDP as the squared modulus of the CIR, i.e., $|h[n]|^2$. Therefore, to reduce the input size of the neural network, we propose as a training feature a downsampled version of the PDP with a given factor L. The proposed feature considers the average of the PDP over equispaced intervals, defined as follows:

$$p[m] = \frac{1}{L} \sum_{n=0}^{L-1} \left| h\big[n_0 + mL + n\big] \right|^2, \quad 0 \le m < \lfloor 152/L \rfloor \quad (8)$$
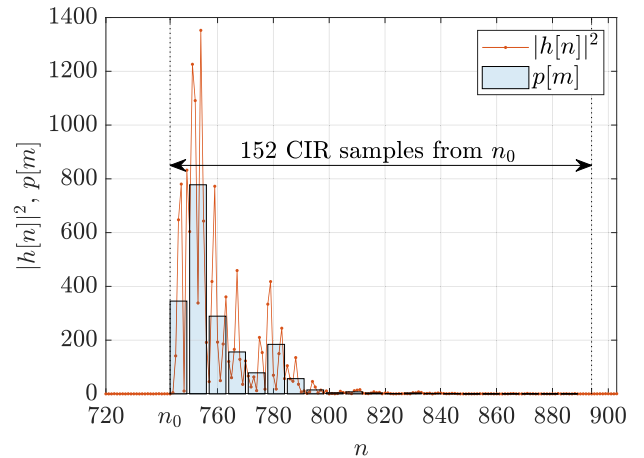
To select the L value with the best result, a comparison was made between different values. Details of this comparison are shown in Section IV, since the results of this comparison are closely related to the energy normalization process described in that section. Note that the number of samples in this proposal depends on the factor L, and this value corresponds to $\lfloor 152/L \rfloor$. Thus, for example, for $L = 5$ the number of samples is 30.

Fig. 2 shows, for an exemplary measurement, the PDP, $|h[n]|^2$, and the corresponding downsampled PDP, $p[m]$, for $L = 7$.

B. NEURAL NETWORK TOPOLOGY AND PARAMETERS
A neural network architecture similar to the one proposed in [20] was generated to evaluate the impact of using the PDP instead of the CIR in the training. Based on the description proposed in [20], a CNN was built with the architecture shown in Fig. 3. Note that, although not shown in Fig. 3, several batch normalization layers were used between each layer to speed up and to facilitate the training. Note also that the design of the neural network is not an objective of this paper, instead we try to replicate the structure used in the other works in order to obtain comparable results.

As shown in Fig. 3, the neural network consists of several convolutional layers. The idea of each of these layers is to extract the relevant features from a section of the input signal (in this case the CIR or PDP samples) and provide them to the next layer where the process is repeated. In our case, the activation function of each neuron in these layers was a rectified linear unit (ReLU). This function is defined as:

$$f(x) = max(0, x) \quad (9)$$

Another layer type used in the network is the *Reshape* layer, whose mission is to transform the shape of a matrix into another different shape. It is used only to adapt the output of one layer to the expected input of the next one.

A spatial reduction layer is typically placed in a CNN after several convolutional layers. In this work, as in the reference architecture described in [20], a *Max Polling* layer was considered to output the maximum input value depending on the size of the window selected.

Finally, a *Dense* layer of multiple neurons fully connected is added to the stack. Its mission is to perform the actual classification using the features extracted by the previous convolutional layers.

The network implementation was done using Tensorflow version 2.2.0 [24], a free, open source Python library capable of running on different operating systems. A PC running Windows 10 LTSC on an AMD Ryzen 5 3600 processor, 16 GB of RAM, and an NVIDIA GeForce GTX 1660 Ti GPU was used for testing. The implementation was configured to use the GPU, so the times presented in the results section are those obtained with it. In an attempt to minimize the possibility of overfitting, an iterative mechanism was designed consisting of the following steps:

1) Of all the samples, 60% were assigned to the training set, 20% to the validation set, and the remaining 20% to the test set, ensuring that all sets contained the same amount of samples from each class (LOS and NLOS).
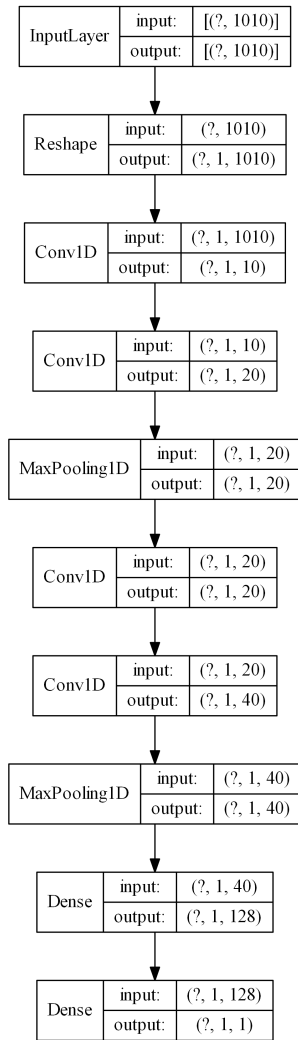2) The training phase was performed using the training set. 20 epochs and a batch size of 64 samples were considered.

FIGURE 3. Network architecture using all the CIR samples. Question marks indicate the input size that depends on the training set size.

FIGURE 4. Evolution of training and validation loss by epoch, using the PDP with $L = 5$.

FIGURE 5. F1-score results using all the CIR measurements (labeled as "cir"), the first 152 samples of the CIR after the detection of the first path (labeled as "cir[152]"), and the PDP with $L = 5$ (labeled as "pdp($L = 5$)"). 10 repetitions. Numbers next to the boxes correspond to the mean value.

3) After the training was completed, the time spent was saved and the network was tested using the test set. The results were also recorded.

This process was repeated N times and, finally, the results were averaged. For this experiment a value of $N = 10$ was set. The entire software implementation created for these experiments is publicly available [25].

As an illustration of the training process, and following the approach in [26], Fig. 4 shows the evolution of the loss function with respect to the epoch index for both the training and validation sets. This example corresponds to one of the repetitions using the PDP.

III. RESULTS CONSIDERING PDP

The first experiment consisted on comparing the effect of considering the PDP samples versus employing all the samples from the CIR [19] or only the first 152 samples after the detection of the first path [20]. The CNN was trained
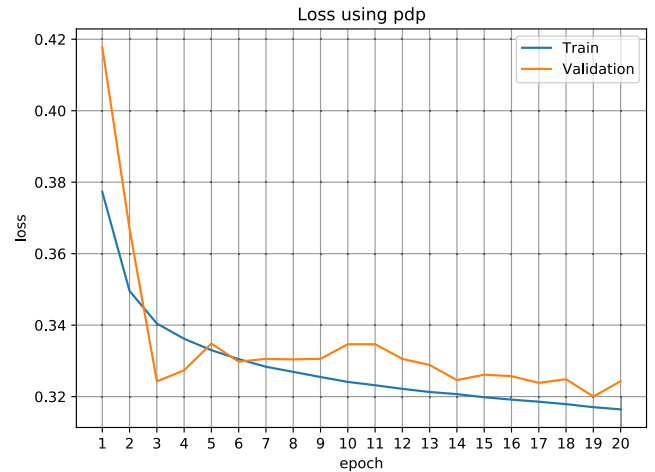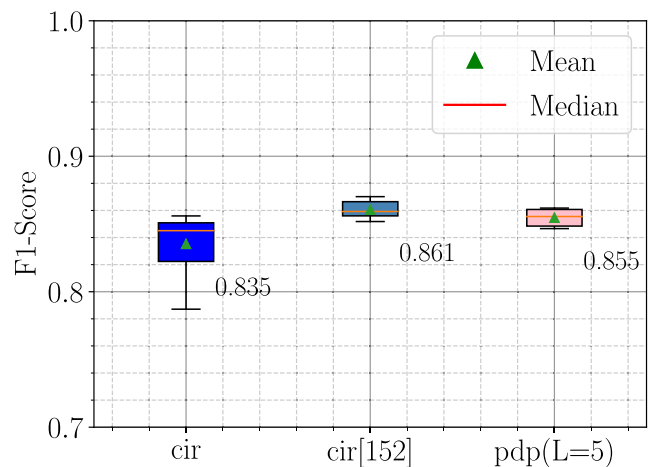
using each of these sets separately. Fig. 5 shows the results for the three sets of features seen in Section II-A (i.e. all CIR Samples, CIR; First 152 samples, CIR[152]; and Downsampled PDP with $L = 5$) after repeating the training and test processes 10 times for each of the different sets of features. The F1-score is a common metric in the literature, by relating in a single value the indicators of precision and recall, and is defined as

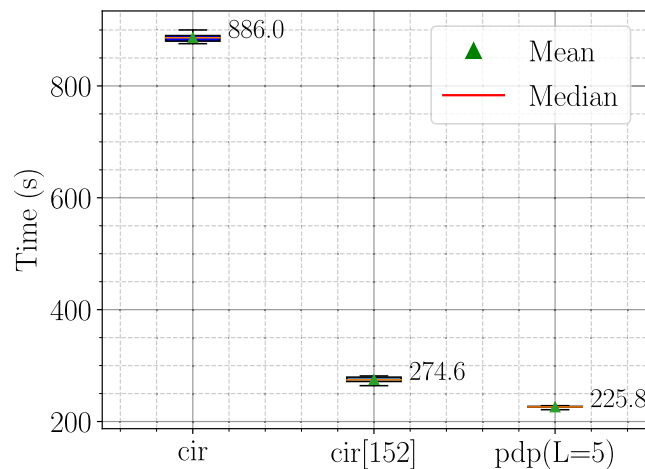$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (10)$$

where precision $= t_p/p$ and recall $= t_p/f$, with t_p being the set of true positives, p being the union of the true positives and the false positives, and f the union of the true positives and the false negatives.

For this comparison a value of $L = 5$ was used to generate the PDP. The choice of this value is detailed in Section IV. With regard to the results shown in Fig. 5, it can be seen

that, despite obtaining close values in all three cases, using all the CIR samples as training features yields the worst and most variable results. The reason for this lies in the fact that most of the CIR samples are just noise, and the actual CIR information is contained only in the samples following the arrival signal. This set of samples with significant information is however very reduced in comparison to the total number of CIR samples captured. Thus the remaining samples (which do not provide any useful information) cause the deterioration of the classifier performance. This can be easily verified by comparing the value obtained with that of the case where only the first 152 samples from the CIR are used from the first path. In the latter case, a higher and more precise (with lower variance) F1-score value is shown in Fig. 5. Consequently, the training result is more predictable, no matter how the samples are distributed in the training and test sets. When considering only the first 152 CIR samples the results in Fig. 5 are quite similar to those obtained in [20], in which the obtained F1-score value was 87.6 and falls within the range obtained in the present experiment (see Fig. 5). This also indicates that the CNN considered reproduces the one used in [20] obtaining similar results.

Regarding the use of the PDP, the results shown in Fig. 5 are on par with those obtained using only the first 152 CIR samples. However, it is important to remember that in this experiment the PDP only contains 30 samples, which will have a positive influence on the computational cost and therefore on the time needed to train and run the network.

Fig. 6 shows the training times averaged over the different executions of the experiment. The time corresponding to the PDP case also includes the time necessary to extract the PDP from the CIR. This time, in the same machine where the simulations were carried out, was 9.807 s. Fig. 6 shows that, in spite of this extra time, the final result is the lowest of the three variants. Obviously, training with all the samples from
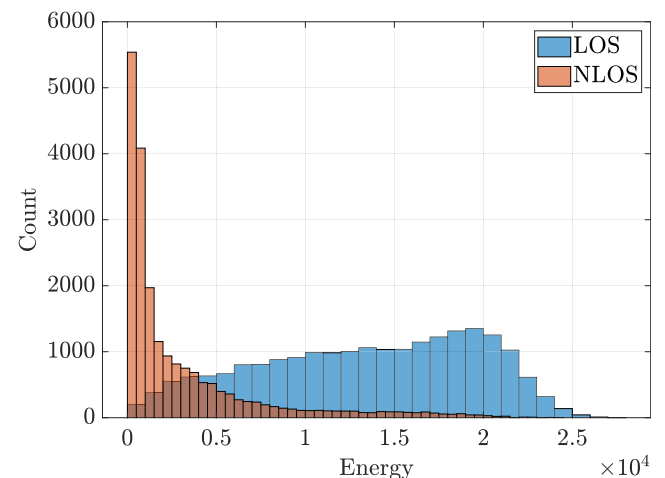
the CIR is the most expensive approach, whereas doing it with only 152 samples reduces the time considerably. Training using only the 30 samples of the PDP is the fastest approach and, since it does not worsen the classifier performance with respect to the other two approaches, it is the best alternative in this case.

IV. EFFECT OF THE DIFFERENT ENERGY LEVELS OF EACH CLASS ON THE RESULT

In order to choose a suitable L value when generating the PDP, classification tests were carried out with different values of this parameter. The results (Fig. 7) did not show significant differences, despite the fact that the final number of samples went from 30 ($L = 5$) to only 3 ($L = 40$). This counter-intuitive result indicated that some other factor outside the simple relationship between the CIR samples was being overlapped. Thus, a comparison was made by analyzing the energy level of all the samples for each of the two classes, LOS and NLOS. The results can be seen in Fig. 8. Clearly,
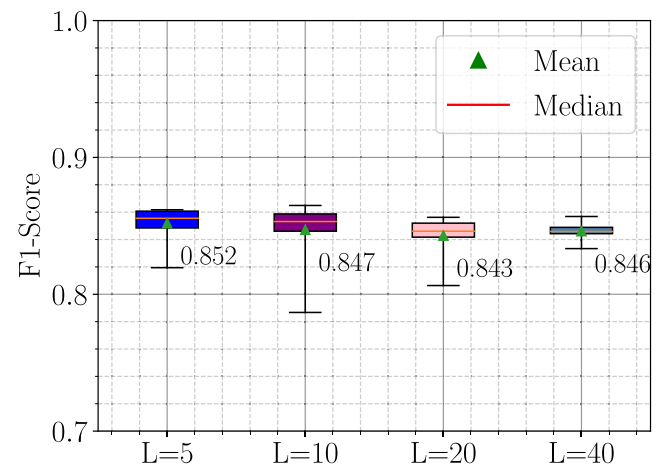
FIGURE 7. F1-score results using PDP and different *L* values. No energy normalization. Numbers next to the boxes correspond to the mean value.

FIGURE 6. Training time using all the CIR samples, the first 152 CIR samples after the detection of the first path, and the PDP with *L* = 5. 20 epochs and 32000 samples. Numbers next to the boxes correspond to the mean value.

FIGURE 8. Measurements energy distribution in both classes.

the energy distribution is very different in both classes, which makes the classifier use this underlying information as a main feature when classifying, obtaining an overfitting for this dataset. Probably, this distribution is a consequence of the way in which the NLOS samples were obtained, maybe with an obstacle in between (i.e. a wall).

In order to be able to observe the results without this effect, we decided to perform an energy normalization to the CIR. However, note that normalizing the CIR energy has the unwanted effect of changing the noise power, which could be also used by the classifier. To avoid this problem, we decided to also remove the noise samples, leaving only those samples with a value above a certain threshold.

It should be noted that, for this particular dataset, the detection value of the first path stored in the measurements, as well as the one calculated by us, is always above the 700-th sample, so we can consider that the first samples of the CIR always correspond to noise. Thus, we considered the first 500 samples of the CIR to obtain the noise power estimation. Using 500 samples allows us to obtain a good estimation of the noise power and ensures that we are only taking noise samples and not samples where the actual CIR is present. Thus, for each CIR measurement $|h[n]|$ we estimate the noise power as

$$\sigma_n^2 = \frac{1}{500} \sum_{n=0}^{500-1} |h[n]|^2. \tag{11}$$

Then, for each measurement $|h[n]|$, we calculated the normalized and noise-free CIR as

$$|\tilde{h}[n]| = \begin{cases} |h[n]|/\mathcal{E}_h & \text{if } n \geq n_0 \wedge n \leq n_L \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

where $\mathcal{E}_h$ is the energy of the CIR, calculated without the effect of the noise estimated in Eq. (11), as

$$\mathcal{E}_h = \sum_{n=500}^{1016-1} \left(|h[n]|^2 - \sigma_n^2 \right) \tag{13}$$

n_0 in Eq. (12) is the index of the first path as defined in Eq. (6), and n_L in Eq. (12) is the instant when the signal exceeds the threshold of the noise, defined in a similar way as t_L in Eq. (4):

$$n_L = \max\{n_0 \leq n < 1016 : |h[n]| \geq \alpha\sigma_n\} \tag{14}$$

where α is a parameter used to select the threshold to detect the last index n where the CIR power is still significantly higher than the noise power. In our case we empirically set $\alpha = 6$ by manually inspecting the results to check that the noise samples were successfully detected and removed.

Fig. 9 shows the classification results with PDP and different L values after energy normalization. In this occasion we observed that the F1-score values decrease as the number of input features decrease. Thus, with the value of $L = 5$ (which gives 30 samples) the best result is obtained, whereas with $L = 40$ (only 3 samples) the worst result is obtained.

Regarding the training time, Fig. 10 shows the results per execution for the different values of L. As expected the time
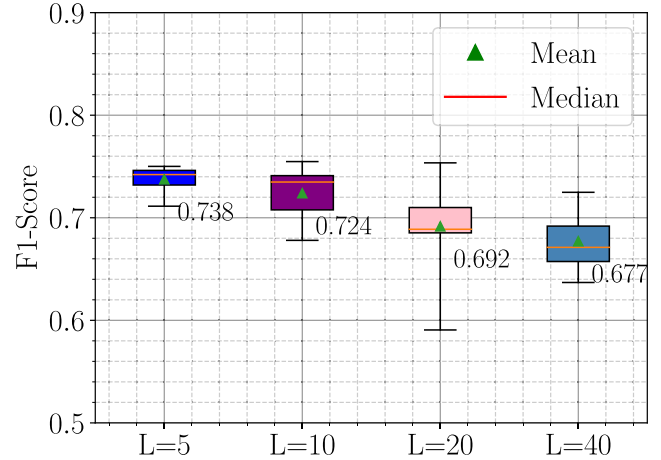
FIGURE 9. F1-score results using PDP and different *L* values. Energy normalized. Numbers next to the boxes correspond to the mean value.
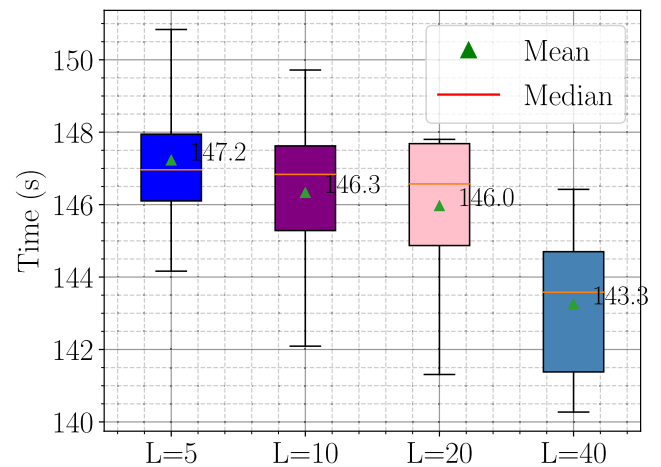
FIGURE 10. Training time using PDP and different *L* values. Energy normalized. Numbers next to the boxes correspond to the mean value.

increases as the number of input features increase, however, it can be seen that the differences are small with the dataset used. In view of these data, it is clear that with the energy of the CIR normalized and eliminating the noisy samples, the best results will always be obtained with the smallest value of L, that is, with the highest number of samples. Thus, for the comparison of Section II we decided to select the value of $L = 5$, since this was the smallest L value among all the analyzed ones.

V. RESULTS CONSIDERING ADDITIONAL FEATURES

In the second experiment, the objective was to test the performance of the classifier by adding a number of extra features in addition to the PDP or the two CIR variants. Based on a previous work ([27]), where these properties demonstrated to contain enough information to make a good classification when using shallow neural networks, two of the chosen features were the range value returned by the DW1000 chip and the RSS estimation. However, on this occasion and since we
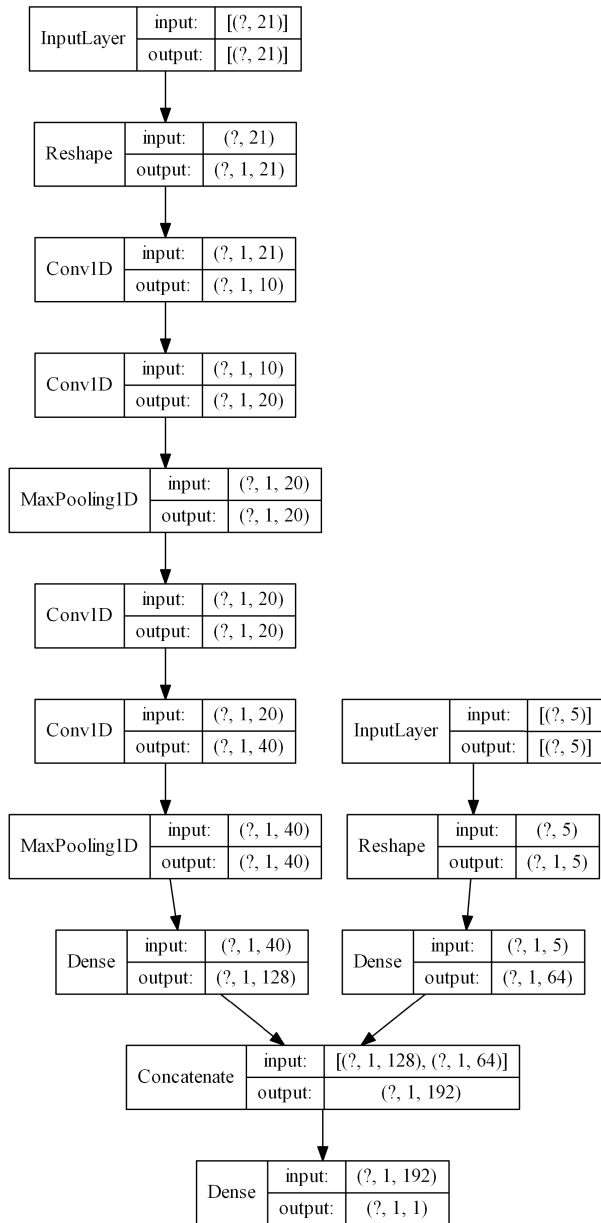
FIGURE 11. Network architecture using PDP and extra features. Question marks indicate the input size that depends on the training set size.

FIGURE 12. F1-score results using additional features. Energy normalized. 10 repetitions using extra features. Numbers next to the boxes correspond to the mean value.

or temporally related (as is the case with the CIR or the PDP samples), it is not the best solution for features that do not have these properties. This is why the new neural network was designed with two different branches: a first one composed of several convolutional layers to treat the CIR/PDP samples (similar to Section II-B), and a different branch composed of a full connected layer whose input corresponds to the two extra features. The final network design is shown in Fig. 11. To carry out the training and test of this network, each sample of the training set was separated into two parts, each one corresponding to one of the two branches of the network. Obviously, the LOS/NLOS class value was common to both.

B. RESULTS

Fig. 12 shows the results obtained in terms of the F1-score with the three different variants: the whole CIR plus the extra features, the first 152 CIR samples after detecting the first path plus the extra features, and the PDP plus the extra features. Note that the F1-score values for the three variants have improved substantially from the results of the first experiment seen in Fig. 5. Thus, when considering all the CIR samples, the F1-score value improves from 0.835 to 0.883; when 152 samples are employed, the improvement goes from 0.861 to 0.889; and finally, an improvement from 0.855 to 0.886 is observed for the PDP case. Therefore, adding the extra features allows for increasing the classifier performance. On the other hand, as in the first experiment (Fig. 5), almost the same values are obtained for the variants using the PDP and the 152 CIR samples since both have similar information, although the PDP requires only 30 samples (with $L = 5$).

Regarding the training time, in this second experiment all the values increase in all the three cases because a more complex network with a greater number of parameters is being employed. Such time increase is also related to the number of features handled, hence the faster variant (using the PDP) increases by only 61.9 s with respect to the version that

have the samples from the CIR, we decided to replace the RSS estimate with the energy extracted from these samples. Thus, a new experiment was carried out in which the following features were added to those already analyzed previously:

- Range: The distance estimation obtained by the DW1000.
- Energy: The energy of the CIR, i.e., the zero-th order moment of the PDP, as shown in Eq. (13).

A. NEURAL NETWORK TOPOLOGY AND PARAMETERS

For this second experiment, we decided to create a new differentiated neural network to handle the new features. Although a CNN is useful for treating information spatially
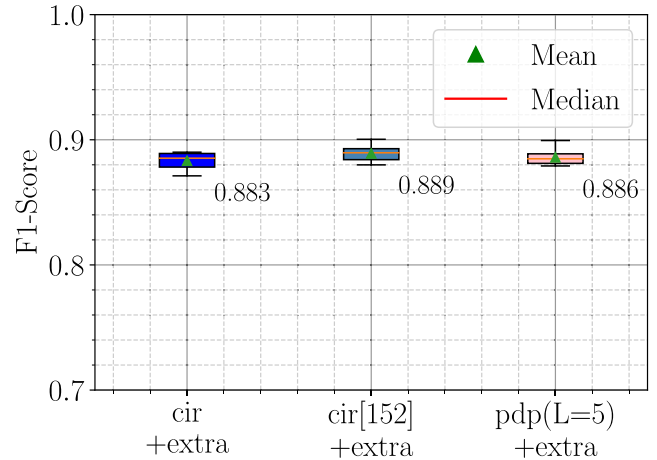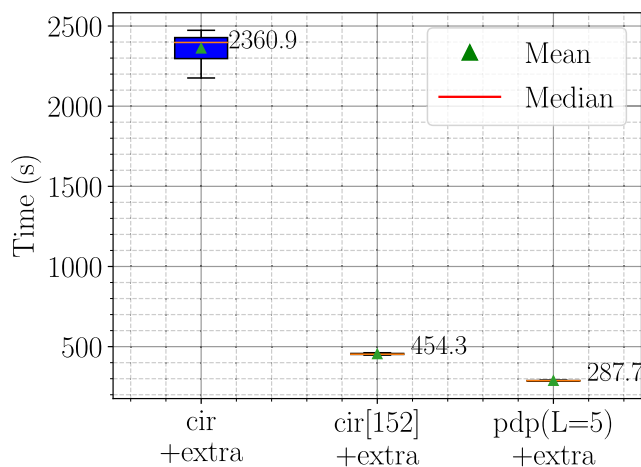
FIGURE 13. Training time using additional features. 20 epochs and 32000 samples using extra features. Numbers next to the boxes correspond to the mean value.

does not use the extra features. The variant that uses the 152 CIR samples suffers a slighter increase, with a final difference of 179, 7 s. Finally, the slower variant employing all the CIR samples is the one that suffers the largest increment with 1474.9 s more than the version without the extra features. In view of this data, we can conclude that additional training features, such as range and energy enables for improving the classifier performance at the expense of a increase of the training time, which depends on the total number of features considered.

VI. CONCLUSION

A series of experiments to test two hypotheses related to the classification of UWB measurements into LOS and NLOS according to its propagation conditions was carried out using deep learning mechanisms. On the one hand, one of the questions to be analysed was the impact of considering a downsampled version of the PDP as a source of information in the learning process instead of directly using the CIR samples. For this purpose, an experiment was carried out in which the same CNN was trained with three variants of the same set of samples: using all the CIR samples, considering only a subset of these containing most of the information, and employing the PDP samples. The results obtained showed that the use of a downsampled PDP offers similar or superior classifier performance than the other alternatives, but requiring less training time.

Another point of interest in this work was to analyze the impact that an unbalanced energy distribution in the classes has on the final classification result. We analyzed the problem showing the values before and after a process of energy normalization and noise removal, checking substantial variations in the final results of the F1-score. This shows us that for the particular case of LOS-NLOS classification of UWB measurements using the CIR or its derivatives, it is not only sufficient to have a balanced dataset in terms of the number of measurements of each class, but it is also necessary that the

energy level is similar in both classes. Otherwise the energy level itself is sufficient to make a satisfactory classification that can mask the real contribution of the relationship between the CIR samples.

Another question to which this work was intended to answer is whether the performance in the classification using the CIR or the PDP samples could be improved by adding to the training set other additional data derived from the CIR, such as the range or the energy. The results obtained in this case showed an improvement in the performance with respect to the first experiment at the cost of an increase in the computational cost.

Although the considered dataset was useful to compare the results with other works based on these data, the conclusions obtained cannot be generalized without performing new experiments considering a different set of measurements. As demonstrated in the present work, the measurements of the considered dataset can be classified at some degree only by their differences with respect to their energy level, beyond the possible relationship between the different paths observed in the CIR. Thus, it becomes mandatory for a future work to obtain a more balanced dataset in which the range and variety of measurements (for example, with long distance LOS samples), allows for assessing the contribution of each information source to the final classification result.

REFERENCES

[1] A. Alarifi, A. Al-Salman, M. Alsaleh, A. Alnafessah, S. Al-Hadhrami, M. Al-Ammar, and H. Al-Khalifa, "Ultra wideband indoor positioning technologies: Analysis and recent advances," *Sensors*, vol. 16, no. 5, p. 707, May 2016.

[2] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2568–2599, 3rd Quart., 2019.

[3] Decawave. *Dw1000 Datasheet*. Accessed: Nov. 23, 2020. [Online]. Available: https://www.decawave.com/wp-content/uploads/2020/04/DW1000_Datasheet.pdf

[4] Decawave. *Decawave List of Partners*. Accessed: Nov. 23, 2020. [Online]. Available: https://www.decawave.com/partners/

[5] V. Barral, C. J. Escudero, J. A. García-Naya, and R. Maneiro-Catoira, "NLOS identification and mitigation using low-cost UWB devices," *Sensors*, vol. 19, no. 16, p. 3464, Aug. 2019.

[6] V. Barral, C. J. Escudero, J. A. García-Naya, and P. Suárez-Casal, "Environmental cross-validation of NLOS machine learning classification/mitigation with low-cost UWB positioning systems," *Sensors*, vol. 19, no. 24, p. 5438, Dec. 2019.

[7] Decawave. *Dw1000 Metrics for Estimation of Non Line of Sight Operating Conditions*. Accessed: Nov. 23, 2020. [Online]. Available: https://www.decawave.com/sites/default/files/aps006_part_3_dw1000_diagnostics_for_nlos_channels.pdf

[8] S. K. Meghani, M. Asif, F. Awin, and K. Tepe, "Empirical based ranging error mitigation in IR-UWB: A fuzzy approach," *IEEE Access*, vol. 7, pp. 33686–33697, 2019.

[9] K. Yu, K. Wen, Y. Li, S. Zhang, and K. Zhang, "A novel NLOS mitigation algorithm for UWB localization in harsh indoor environments," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 686–699, Jan. 2019.

[10] I. Güvenç, C.-C. Chong, F. Watanabe, and H. Inamura, "NLOS identification and weighted least-squares localization for UWB systems using multipath channel statistics," *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 1, pp. 1–14, Dec. 2007.

[11] B. Song, S.-L. Li, M. Tan, and Q.-H. Ren, "A fast imbalanced binary classification approach to NLOS identification in UWB positioning," *Math. Problems Eng.*, vol. 2018, pp. 1–8, Dec. 2018.

[12] Z. Xiao, H. Wen, A. Markham, N. Trigoni, P. Blunsom, and J. Frolik, "Non-line-of-sight identification and mitigation using received signal strength," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1689–1702, Mar. 2015.

[13] Z. Zeng, S. Liu, and L. Wang, "UWB NLOS identification with feature combination selection based on genetic algorithm," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2019, pp. 1–5.

[14] Z. Zeng, S. Liu, and L. Wang, "UWB/IMU integration approach with NLOS identification and mitigation," in *Proc. 52nd Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2018, pp. 1–6.

[15] M. J. Kuhn, J. Turnmire, M. R. Mahfouz, and A. E. Fathy, "Adaptive leading-edge detection in UWB indoor localization," in *Proc. IEEE Radio Wireless Symp. (RWS)*, Jan. 2010, pp. 268–271.

[16] J. Fan and A. S. Awan, "Non-Line-of-Sight identification based on unsupervised machine learning in ultra wideband systems," *IEEE Access*, vol. 7, pp. 32464–32471, 2019.

[17] J.-S. Choi, W.-H. Lee, J.-H. Lee, J.-H. Lee, and S.-C. Kim, "Deep learning based NLOS identification with commodity WLAN devices," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3295–3303, Apr. 2018.

[18] J. Li, Y. Chang, T. Zeng, and Y. Xiong, "Channel correlation based identification of LOS and NLOS in 3D massive MIMO systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–6.

[19] C. Jiang, J. Shen, S. Chen, Y. Chen, D. Liu, and Y. Bo, "UWB NLOS/LOS classification using deep learning method," *IEEE Commun. Lett.*, vol. 24, no. 10, pp. 2226–2230, Oct. 2020.

[20] K. Bregar and M. Mohorcic, "Improving indoor localization using convolutional neural networks on computationally restricted devices," *IEEE Access*, vol. 6, pp. 17429–17441, 2018.

[21] K. Bregar. *Ewine-Project/UWB-Localization Source Code Repository*. Accessed: Nov. 23, 2020. [Online]. Available: https://github.com/ewine-project/UWB-localization

[22] Decawave. *Dw1000 Manual*. Accessed: Nov. 23, 2020. [Online]. Available: https://www.decawave.com/dw1000/usermanual/

[23] S. Marano, W. Gifford, H. Wymeersch, and M. Win, "NLOS identification and mitigation for localization based on UWB experimental data," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 7, pp. 1026–1035, Sep. 2010.

[24] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, and S. Ghemawat. (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*. [Online]. Available: https://www.tensorflow.org/

[25] V. Barral. (2020). *Source Code Repository of: Using the Power Delay Profile to Accelerate the Training of Neural Network-Based Classifiers for the Identification of Los and NLOS UWB Propagation Conditions*. [Online]. Available: https://github.com/valentinbarral/uwb_los_classification_pdp_dl

[26] A. Barua, Z.-Y. Zhang, F. Al-Turjman, and X. Yang, "Cognitive intelligence for monitoring fractured post-surgery ankle activity using channel information," *IEEE Access*, vol. 8, pp. 112113–112129, 2020.

[27] V. Barral, C. J. Escudero, and J. A. Garcia-Naya, "NLOS classification based on RSS and ranging statistics obtained from low-cost UWB devices," in *Proc. 27th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2019, pp. 1–5.

VALENTÍN BARRAL VALES received the degree in computer engineer from the University of A Coruña. He currently works as an Associate Researcher with the University of A Coruña. His research interest includes indoor localization using radio technologies, a field in which he has published several works, since 2012. In recent years his research has focused on ultra wideband technology and its use for location purposes in complex indoor environments. His Ph.D. thesis titled: *Ultra Wideband Location in Scenarios Without Clear Line Of Sight*.
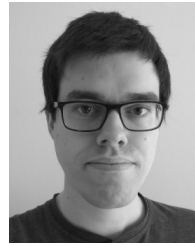
TOMÁS DOMÍNGUEZ-BOLAÑO received the B.S. degree in computer engineering and the Ph.D. degree in computer engineering (with the distinction Doctor with European Mention) from the University of A Coruña, A Coruña, Spain, in 2014 and 2018, respectively. Since 2014, he has been with the Group of Electronics Technology and Communications. He was a Visiting Scholar with Tongji University, Shanghai, China, in 2018. He is the author of more than 15 articles in peer-reviewed international journals and conferences. He was awarded with a predoctoral grant and two research-stay grants. His research interests include channel measurements, parameter estimation, and modeling and experimental evaluation of wireless communication systems.

CARLOS J. ESCUDERO (Member, IEEE) received the Ph.D. degree in computer science from the University of A Coruña (UDC), Spain, in 1998. Since 1992, he has been a Faculty Member with the Department of Computer Engineering, UDC, where he was an Associate Professor, in 1998. From March 2000 to February 2005, he was a Secretary of the Department, the Vice-Dean of the Faculty of Informatics, from March 2009 to April 2013, the Head of the Department of Electronics and Systems, from April 2013 to January 2016, and the Deputy Rector of the University of A Coruña for Information and Communication Technologies, since 2016. His research interests include applications of signal processing in communications, wireless sensor networks, and indoor location systems. Proof of this are some of the merits alleged in the curriculum: publications in journals and books most prestigious, many national and international collaborations with renowned, dozens of papers / publications in national and international conferences, highlighting the most significant in the curriculum, participation in 25 competitive research projects in which he has been the principal investigator on seven, responsible for 20 contracts with companies, author of two patents, founding member of two companies associated with the University.

JOSÉ A. GARCÍA-NAYA (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in computer engineering from the University of A Coruña (UDC), Spain, in 2005 and 2011, respectively. He is currently with the Group of Electronics Technology and Communications, UDC, where he is also an Associate Professor, and with the CITIC Research Center. He is the coauthor of more than 90 peer-reviewed articles in journals and conferences. He is also a member of the research team in more than 40 research projects funded by public organizations and private companies. His research interests include signal processing and experimental evaluation of wireless communication systems, including indoor location, time-modulated arrays, and high-mobility transportation.

• • •