

Article

NURSENET: Monitoring Elderly Levels of Activity with a Piezoelectric Floor

Ludovic Minvielle * and Julien Audiffren

Centre de mathématiques et de leurs applications, CNRS, ENS Paris-Saclay, Université Paris-Saclay, 94230 Cachan, France; audiffren@cmla.ens-cachan.fr

* Correspondence: minvielle@cmla.ens-cachan.fr

Received: 2 August 2019; Accepted: 2 September 2019; Published: 6 September 2019



Abstract: Monitoring the activity of elderly individuals in nursing homes is key, as it has been shown that physical activity leads to significant health improvement. In this work, we introduce NURSENET, a system that combines an unobtrusive, affordable, and robust piezoelectric floor sensor with a convolutional neural network algorithm, which aims at measuring elderly physical activity. Our algorithm is trained using signal embedding based on atoms of a pre-learned dictionary and focuses the network's attention on step-related signals. We show that NURSENET is able to avoid the main limitation of floor sensors by recognizing relevant signals (i.e., signals produced by patients) and ignoring events related to the medical staff, offering a new tool to monitor elderly activity in nursing homes efficiently.

Keywords: deep learning; piezoelectric sensor; floor sensor; pressure sensor; activity monitoring; sparse dictionary learning

1. Introduction

1.1. General Introduction

In the recent years, the notion of frailty has received increased attention from the medical community as an early sign of significant health deterioration in the elderly population [1,2]. The authors of [1] defined frailty as a clinical syndrome in which three or more criteria are present among unintentional weight loss, self-reported exhaustion, weak grip strength, slow walking speed, and low physical activity. The early detection of this syndrome is considered important by the medical community for prevention purposes, and targeted therapy strategies have been shown to reduce the consequences of frailty significantly [3].

Nursing home residents are among the most vulnerable when it comes to frailty [4]. Indeed, people living in elderly care facilities are more likely to have disabilities than other elderly people. They also tend to be more isolated from the rest of the population and do less physical activity than the general elderly population [5]. The lack of exercise is particularly damaging, as it has been shown that physical activity reduces muscle weakness, increases mobility, improves neuronal health, limits frailty, and reduces overall risk of death [6,7]. Even short daily practice such as thirty minutes of walking is linked with significant improvement [8]. Consequently, the monitoring physical activity of the elderly in nursing homes is key, particularly as it allows focusing resources on people who are particularly vulnerable to frailty.

Recently, there has been significant interest in the development of smart systems to monitor individuals in their daily life [9]. Important aspects of such systems include the intrusiveness, the reliability, and the ability to perform different tasks. Here, intrusiveness is defined as both the capacity to be hidden from the patient and the respect of her/his privacy [10]. Some sensors

relying on technologies such as a camera [11] or microphones [12] have been shown to achieve great performance in their respective task (counting people, detecting events, etc.) at the cost of serious privacy issues. Others have used accelerometers, which are less intrusive and provide reliable signals [13,14], but wearable technology is (A) constrained by individual willingness and (B) may be rendered ineffective if individuals forget to wear them [15]. Floor sensors on the other hand resolve these issues by being unobtrusive while still providing exploitable signals. However, the main drawback of floor sensors is their *a priori* inability to distinguish between individuals. This is particularly problematic in nursing homes as most of the activity is generally generated by medical staff.

To solve this problem, in this paper, we introduce our system named the Non-invasive Unit Recognition System for the Elderly (NURSENET), which combines a piezoelectric floor sensor and a deep learning-based algorithm to recognize non-elderly activity with very high accuracy (AUC 0.94). This paper is organized as follows. Section 2 first gives the settings of our installation in the nursing home and introduces our algorithm and its training procedure. The performance of NURSENET is evaluated in Section 3 by computing its accuracy over a database containing multiple types of activity. Finally, the main advantages and drawbacks of our system are discussed in Section 4.

1.2. Related Work

1.2.1. Gait Sensing

The quantification of gait is hardly new [16–18], and a large variety of sensors has been used to record walks. Systems can be classified between wearable sensors and non-wearable, the latter being divided between image-based systems and other systems. The most popular are wearable sensors, as evidenced by the large literature on the subject [18,19]. These systems are mostly based on accelerometers, which dominate the field, but also gyroscopes, foot switches, pressure soles, or combinations of these. Accelerometers are indeed inexpensive sensors that give reliable signals and are more scalable than foot switches or soles [16].

Concerning non-wearable sensors, video-based systems seem to give good results when analyzing gait. However, considering the high intrusiveness of these solutions, they are better suited for clinical applications rather than daily monitoring. Recent works proposed systems based on WiFi sensing [20,21]. Based on the measurement of the WiFi signal after reflection on objects, many applications were proposed, but this technology seems to be in its early stages and suffers from issues such as the difficult availability of the needed metric to perform the sensing and external interferences [21]. When it comes to floor technologies, various technologies have been used. Vibration-sensor systems are very easy to install, and previous works have shown that they can be used to separate individuals with high accuracy [22,23]. However, the signal is subject to variation according to the device's location and the size of the sensing area, which may require significant additional training for each new installation. Floor-covering solutions [10], although requiring more effort in their deployment, have the advantage of being consistent since the sensing is equivalent at any location. These systems are mainly made of binary switch sensors [24,25], capacitive sensors [26–28], and pressure sensors, the latter being either piezoresistive (the material resistance is modified when put under pressure) [29] or piezoelectric (the material emits a charge or a voltage when stressed) [30,31]. When implemented, these systems have different resolutions, from only one sensor to a grid of 1 cm × 1 cm tiles. However, high resolution leads to implementation issues and higher costs. Depending on the sensor and resolution, the emitted signals can provide sufficient information for complex tasks, but can also suffer from external perturbations [10]. Our system is based on a thin piezoelectric polymer that acts as a pressure sensor and comes in bands. It is unobtrusive and gives a reliable signal while being relatively easy to install, all at a reduced cost due to precisely the ease of installation, but also the availability of its components (processing unit, sensors, and flooring) for large-scale installations and the fact that the system comes as an all-in-one solution.

1.2.2. Gait Analysis

The main families of analyses developed in previous works include: (A) the use of biomechanical features, such as single/double stance time [32]; (B) the use of descriptors derived from mathematical models [33]; and more recently, (C) the use of machine learning-based algorithms that perform end-to-end learning on gait signals [34]. Several walk detection systems rely on gait decomposition into several phases, trying to detect two to eight phases [16] depending on the sensor and the precision level needed for their application. These methods can be based on heuristic rules (e.g., thresholds [35]), signal processing techniques (e.g., Pan–Tompkins peak detection [36]), or probabilistic approaches (e.g., HMM-based algorithms [37]). However, most methods are designed to characterize gait in a controlled environment where no challenging type of signal can infer (i.e., signal that may seem like what is to be detected). Hence, when dealing with daily activities, walk detection becomes a harder task [38,39]. Another segment of step detection aims at recognizing a step signal as a whole (i.e., without considering any phase in it) in order to perform walk detection within more challenging environments or to use walk signals for patient monitoring. These methods are generally based on machine learning techniques and seem more suited to such tasks [40,41], but to the best of our knowledge, no previous work has classified activities using a floor piezoelectric sensor combined with machine learning techniques.

1.2.3. Sparse Coding and Dictionary Learning

In recent years, dictionary learning and sparse coding techniques have been successfully applied in a wide range of topics, including image classification [42,43], image restoration [44], and signal processing [45], in particular for one-dimensional signal analysis [46]. The main idea behind these representations is to learn a dictionary conjointly containing the patterns observed in the signal and sparse activations that encode the temporal or spatial locations where these patterns occur. Previous works have applied this method to walk-related problems: Poschadel et al. [47] used the accelerometer signal reconstructed from sparse coding to perform gait classification between healthy and movement impaired individuals; Zhang et al. [40] also aimed at gait classification by applying D-KSVD techniques to produce a dictionary that is both representative and discriminative. Following the same idea, we use convolutional dictionary learning to build the first layer of NURSENET in order to construct a relevant representation.

1.2.4. Signal Processing with Deep Learning

Convolutional Neural Networks (CNN) have been successfully used to detect objects in images [48–52], hence superseding previously existing techniques [53]. Among them, recent networks such as YOLO [52] and Faster R-CNN [51] have been shown to find and classify multiple objects with high accuracy in images with a small time cost. However, significantly less attention has been given to the application of CNN on one-dimensional signals, despite such signals being widely represented in medical recordings [54]. Previous works generally transformed one-dimensional signals into an image using spectrograms [55] or directly used CNN without transforming the data [56]. To the authors' knowledge, this is the first work to combine dictionary learning, transfer learning, and CNN to classify activities using the floor sensor as the sole input. As shown in Section 2.2, the combination of these three steps is revealed to be necessary due to the highly-irregular nature of the signal.

2. Materials and Methods

2.1. NURSENET Hardware

In this subsection, we present the floor sensor that was used in NURSENET. The piezoelectric principle (on which the device was based) is outlined, then the unit is described in more depth.

2.1.1. The Piezoelectric Principle

The piezoelectric principle, when simplified, can be explained by the relation $d = \frac{Q}{F}$, with d being the piezoelectric coefficient, Q the charge, and F the force received by the material. A piezoelectric material emits charges when stressed (or squeezed) and on the contrary can shrink (or expand) when submitted to an electric field. Such a device can fit multiple purposes such as speakers or motors when using its deformation under an electric field, to microphones and sensors when using the electric charge emitted when submitted to a force. Our system was based on the latter use, and the piezoelectric material acted as a pressure sensor. One significant advantage of this technology is that it does not need a power supply to generate a signal.

2.1.2. The NURSENET Unit

The sensor itself was a thin (about 1 mm) polymer that comes in bands of 60 cm wide and is placed directly under the flooring (see Figure 1a). The bands were initially rolls of about 100 m that could be cut every 30 cm. Once cut, bands can be connected together without limitation in their number before the signal was recovered by the processing unit (see Figure 1b). In NURSENET, the processing unit had 8 entry channels, allowing each unit to process a large surface by connecting bands. This allows great flexibility of the system to any scale while keeping a certain ease of installation.

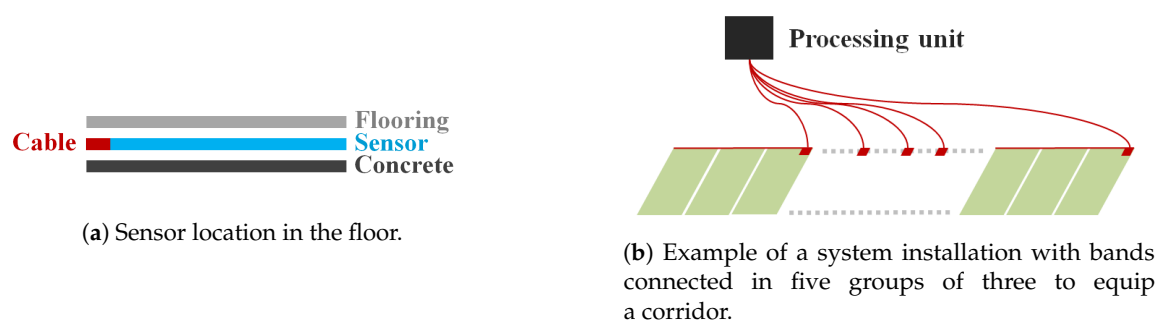


Figure 1. Pressure sensor installed in the floor.

The processing unit was designed to process the signal from end to end (from basic filtering to event detection) and to communicate with servers (sending signal recordings) and the nursing home (sending information such as activity reports or alarms). When conveyed to the unit, the signal first passed through an analog charge amplifier and was converted to numerical values before further processing. The unit was equipped with a 32-bit 500-MHz processor, accompanied with 256 MB of RAM, and 500 MB of local storage.

It should be noted that compared to the theoretical piezoelectric principle law, the sensor presented some variability in the piezoelectric coefficient d . Indeed, although d seemed to remain constant when put under humidity or high temperature (less than 40° Celsius), it could vary depending on where the impact was located on the floor. A difference of 9% between the max and min value of d was observed in controlled experiments [57]. Besides, when implemented, the system can suffer perturbations due to the way sensors, electrodes, and upper layers were installed. These elements that come with the implementation in real conditions can result in alterations to the resulting signals when compared to similar devices used in laboratory conditions [10].

2.2. The NURSENET Algorithm

In this subsection, we present our neural network NURSENET, which is used to identify and characterize activity signals, particularly between medical staff and elderly individuals. This subsection is organized as follows. First, the structure of NURSENET is introduced. Then, the data embedding used to encode the signals is described, and finally, the subnetwork used as the first layers of NURSENET is explained. Most of the strategies developed in the training process aim at circumventing our main

issues, which are the small Signal-to-Noise Ratio (SNR) (approximately 20 dB in our walking dataset) and the external alterations of the signal.

2.2.1. General Classifier

This subsection details the architecture of NURSENET and its inner workings. The main idea of NURSENET is that regardless of the type of activity recorded on the floor sensor, it is very likely to be mostly made of walks. This is why, as discussed later in this section, a significant part of the training process aimed at training the network to recognize gait and in particular the gait of the medical staff. NURSENET is a CNN whose structure is presented in Figure 2. It is a sequential network made of:

- One 1D convolutional layer with 32 filters of size 60, with a stride of 10, followed by a BatchNorm layer, with the activation function hyperbolic tangent (tanh).
- One 1D convolutional layer with 16 filters of size 1, with a stride of 1, followed by a BatchNorm layer, with the activation function tanh.
- One 1D convolutional layer with 8 filters of size 1, with a stride of 1, followed by a BatchNorm layer, with the activation function tanh.
- One 1D convolutional layer with 1 filter of size 5, with a stride of 1, with the activation function Rectified Linear Unit (ReLU) [58]. It was followed by a Maxpool layer of size 5.
- One fully-connected layer, with an output of dimension 64, with activation function ReLU.
- One fully-connected layer, with an output of dimension 16, with activation function ReLU.
- One fully-connected layer, with an output of dimension 1, with activation function sigmoid.

The output of the network is the probability for a signal to be generated from elderly activity. Note that the input of NURSENET is not the raw signal, but instead a transformation of it (see Section 2.2.2).

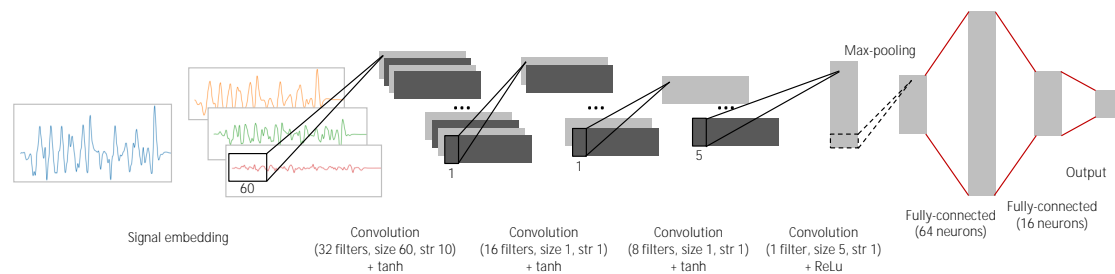


Figure 2. Architecture of the overall CNN.

Sub-networks. NURSENET can be seen as the combination of the following two sub-networks. The first one, named the Step Proposal Network (SPN), is made of the first three convolutional/BatchNorm layers. SPN is a full CNN inspired by the region proposal network, which is the first part of the Faster R-CNN algorithm [51]. At the end of this sub-network, the output is constituted of features computed on each window of the signal embedding. The second part of the network is made of a convolutional/Maxpool pair of layers, followed by three classical fully-connected layers. The role of this last part of the network is to use the previously extracted features to classify signals as *staff* or *elderly*.

Training. As mentioned before, the training of NURSENET was rather complex and involved multiple stages. This was due to the small signal-to-noise ratio, the perturbations discussed in the previous section, and the limited size of the dataset. The first phase of the training was to learn the embedding of the data. This is detailed in Section 2.2.2. Then, SPN was trained separately on a different task, in order to take advantage of the gait structure (see Section 2.2.3). The resulting weights were used to initialize the first three layers of NURSENET, and these layers were frozen during the rest of the training process to limit the number of free parameters, henceforth reducing overfitting [59].

Finally, the rest of the network was initialized using i.i.d. centered Gaussian variables with a standard deviation of 0.2. For the training phase, regularization was done using dropout [60] (with $p = 0.5$) after each fully-connected layer, except the last one. Given X_i , the embedding of each signal \hat{y}_i , the image of X_i through the network, and y_i the corresponding labels ($y = 0$ if and only if X_i is an elderly activity), the loss function \mathcal{L} is the classical binary cross entropy:

$$\mathcal{L} = \sum_i \mathbf{1}_{y_i=1} \log(\hat{y}_i) + \mathbf{1}_{y_i=0} \log(1 - \hat{y}_i). \quad (1)$$

The weights of NURSENET were optimized using backpropagation and stochastic gradient descent with a learning rate of 10^{-5} , decreasing geometrically ($\times 0.9$) every 10 epochs, and the momentum strategy proposed by Nesterov [61]. The learning was stopped using the early stopping principle [62].

2.2.2. Data Embedding

The first step of our classification algorithm was the data embedding, which consisted successively of preprocessing the data and encoding the resulting signal using convolution with atoms of a pre-learned dictionary. The preprocessing phase was made of a classical tool that aimed at denoising and cleaning the data. However, our approach of data embedding is rather new: while we used convolutional sparse coding to learn the custom dictionary, a standard approach since sparsity has been shown to produce more relevant atoms (see [63] and the references therein), the actual embedding of the data was done with (non-sparse) convolution.

The reasoning behind this approach is as follows. It has been shown that the first layers of a CNN tend to learn general feature extractors: for example, in image processing, the first layers usually exhibit features similar to Gabor filters and color blobs [64]. The aim of our data embedding phase was to replace the first layers of the CNN with some already trained filters, here the atoms of the dictionary. Since the training was done on a small dataset, reducing the number of layers in the network may improve the results. Hence, in order to mimic the behavior of the first layers of the CNN, regular convolution was used instead of convolutional sparse coding. All data embedding details were given below.

Preprocessing. It has been shown that preprocessing improves the performance of CNN networks [65], particularly when dealing with a small labeled dataset. The preprocessing of data was done as follows. Let $(C_t^k)_{t=1}^T$ denote the signal produced by the channel k ($1 \leq k \leq K = 5$), where $t = 1, \dots, T$ represent the data points recorded with a frequency of 100 Hz.

1. Each channel $(C_t^k)_{t=1}^T$ was filtered with a low-pass Butterworth filter with a 10-Hz cutoff frequency, fifth order, and zero lag. This step aimed to limit the amount of electronic noise present in the signal, as the 10-Hz cutoff frequency is the reference in gait-related signals [66].
2. The linear trend of each channel was removed using a least squares model.
3. Each channel whose signal maximum amplitude was small was then set to zero, as the channel was assumed to only account for noise.
4. The resulting signal \mathbf{s} was obtained as the sum of all the channels:

$$\mathbf{s} \doteq (C_t)_{t=1}^T = \left(\sum_{k=1}^K C_t^k \right)_{t=1}^T$$

Figure 3 shows an example of the resulting signal.

Signal embedding. Data augmentation is a key part of training complex neural networks, particularly when only small labeled datasets are available [67]. However, these techniques are difficult to apply in our setting, as it is unclear that commonly-used transformations would preserve the nature and structure of the data. For instance, a re-scaled step signal may be confused with a fall signal, as the amplitude of the applied force is the main difference between the two signals (see [68]). In our case,

we are dealing with a small dataset, which is composed of complex signals, thus making end-to-end learning a difficult task. To circumvent this difficulty, the data were first transformed using convolution with atoms learned from Convolutional Dictionary Learning (CDL) in order to extract key elements and features of the signal. This first phase was intended to improve the quality of the CNN input to balance out the limited amount of available data.

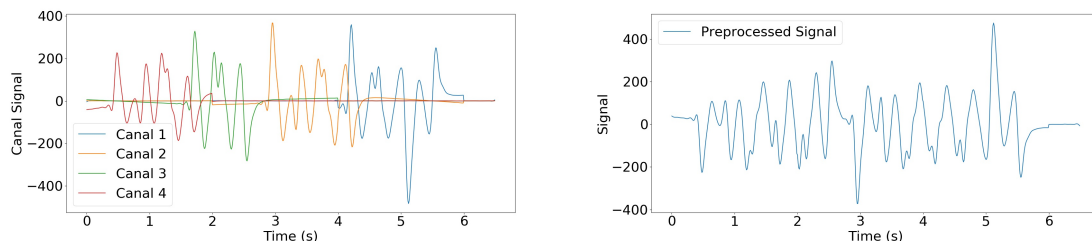


Figure 3. Example of raw (left) and preprocessed (right) signals resulting from medical staff walking on the sensor. In this case, no channel was set to zero during Step 3.

Convolutional dictionary learning. CDL is defined as the learning of a set of atoms to best represent data (e.g., images or one-dimensional signals) while insuring a sparse representation. Given a vector \mathbf{s} that contains data to be represented, the objective is to find atoms \mathbf{d}_m and activation signals \mathbf{x}_m such that the reconstruction is accurate, i.e., $\mathbf{s} \approx \sum_m \mathbf{x}_m * \mathbf{d}_m$. Solving CDL generally uses the following objective function, referred to as convolutional basis pursuit denoising [46]:

$$\arg \min_{\mathbf{x}_m, \mathbf{d}_m} \frac{1}{2} \left\| \sum_m \mathbf{x}_m * \mathbf{d}_m - \mathbf{s} \right\|_2^2 + \lambda \sum_m \|\mathbf{x}_m\|_1 \quad (2)$$

Previous works have shown that many walk alterations can be detected and quantified by studying the characteristics and variety of steps [69]. Following this idea, CDL has been successfully used to quantify and evaluate gait [47]. However, it is worth noting that CDL is generally used on signals derived from accelerometers, which have been proven to be a robust and reliable way to record walks. In our setting, the recording was of lower quality, and nearly-identical steps, such as the ones produced by the walking of young healthy individual [70], were recorded with significant variations between them (see Figure 3). Besides, a large number of signals in the dataset contained elements that were not related to steps (e.g., pushing a wheelchair or a cart). Finding a suitable dictionary was therefore significantly more challenging, and this is why atoms were created by selecting signals that were almost entirely walk related.

Learning step atoms. As previously said, in order to learn atoms relevant to our step description process, we used a CDL approach on a subset of the walk signals, to guide the learning towards step-related atoms. Equation (2) shows that when learning a dictionary, we learn at the same time the dictionary and sparse representations of input signals. The standard procedure of dictionary learning is to alternate between a sparse coding step (i.e., updating the sparse representation of the data according to the current dictionary) and a dictionary update (i.e., updating the dictionary according to the current sparse representation). This means that (2) is solved successively with \mathbf{d}_m fixed (sparse coding) or \mathbf{x}_m fixed (dictionary update). We used the method of Garcia-Cardona and Wohlberg [71], which is derived from the Alternating Direction Method of Multipliers (ADMM). Proposed by Boyd et al. [72], ADMM solves problems of the form:

$$\arg \min_{x,y} f(x) + g(y) \quad \text{such that} \quad Ax + By = c \quad (3)$$

By using the augmented Lagrangian multipliers. Adapted to sparse coding and dictionary update with a specific coupling between the two steps (named consensus), this method proved to be efficient among the existing ones [46]. We used the implementation of Wohlberg [73] of the

algorithm, available in a Python package named SPORCO (SParse Optimization Research COde). The number of atoms m was considered a hyperparameter of our CDL problem, and we chose it using cross-validation. Our experiments showed that $m = 3$ achieved the best trade-off between the quality of the reconstruction of the training set and the one of the validation set, while maintaining the sparsity of the representation. Using less atoms led to either poor reconstruction or non-sparse representation of the training set, whereas using more atoms significantly decreased the reconstruction quality of the validation set. These three atoms were 0.7 s long, which corresponded to an upper bound of the duration of a step in the walking of a healthy young individual [74]. They were obtained with $\lambda = 5$ and 2000 iterations during the optimization process. Figure 4 shows the resulting dictionary.

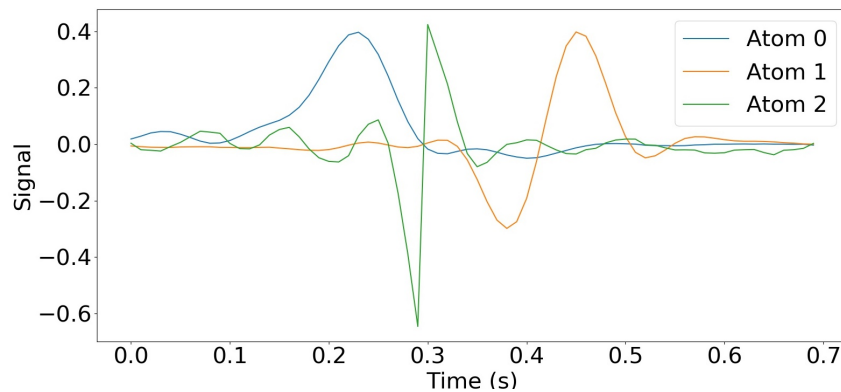


Figure 4. Dictionary learned using Equation (2). The amplitude of the atoms is small compared to the signal because they are normalized.

Convolutional dictionary embedding. As discussed above, while we used a sparsity constraint on the dictionary learning process, the actual embedding of the input data \mathbf{s} into \mathbf{S}^d was done with standard convolution:

$$\mathbf{S}^d \doteq \left(\mathbf{s} * d_m \right)_{1 \leq m \leq 3} \quad (4)$$

Resulting in a three-channel encoding of the signal. Figure 5 shows an example of the data encoding process.

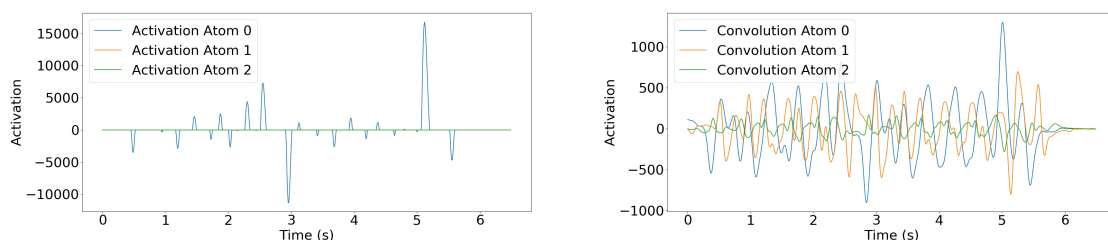


Figure 5. (Left) Convolutional sparse decomposition of the signal of Figure 3. Note that for this individual, only Atom 0 is used. (Right) Convolution of the same signal with each atom. For this signal, only Atom 0 is used in the sparse coding solution, but this differs in other signals. Moreover, it is interesting to note that since Atom 0 is mildly similar to a Gaussian kernel, the result of the convolution of the signal with Atom 0 is close to a smoothed version of the original signal.

2.2.3. Pre Training Weights with Step Detection

This subsection discusses the idea behind the first sub-network of NURSENET and presents its dedicated training process.

As discussed before, this sub-network is inspired by the Region Proposal Network (RPN). The idea behind RPN is to help the main network of Faster R-CNN [51] focus only on interesting parts of an image. To this end, a sliding window was passed on a feature map of the input image. On this sliding window, boxes of different scales and ratios (in total K) were used as region proposals, and these regions were evaluated through two layers. The first one output the probability of a box to be an object and the other, one gave corrected coordinates of the box. This method allowed reducing the number of false examples that may mislead the training process and greatly accelerated the item search. In our case, since we were dealing with one-dimensional signals (hence significantly smaller than images), we were not especially seeking high speed execution. However in this approach, steps were considered as a key element to distinguish between medical staff and elderly activities. Therefore, by training a network to identify and characterize steps, the idea was to ensure that most of the information that was passed to the second sub-network of NURSENET was relevant. Moreover, as this step recognition part brought a separate training, it added a constraint to the overall model, hence reducing the chances of overfitting (see [75] for a discussion related to constraints and overfitting).

The architecture of the CNN used to train the first sub-network of NURSENET is presented in Figure 6. This CNN was made of:

- One 1D convolutional layer with 32 filters of size 60, with a stride of 10, followed by a BatchNorm layer, with the activation function hyperbolic tangent (tanh).
- One 1D convolutional layer with 16 filters of size 1, with a stride of 1, followed by a BatchNorm layer, with the activation function tanh.
- One 1D convolutional layer with 8 filters of size 1, with a stride of 1, followed by a BatchNorm layer, with the activation function tanh.
- One 1D convolutional layer with 3 filters of size 1, with a stride of 1, with the activation function sigmoid.

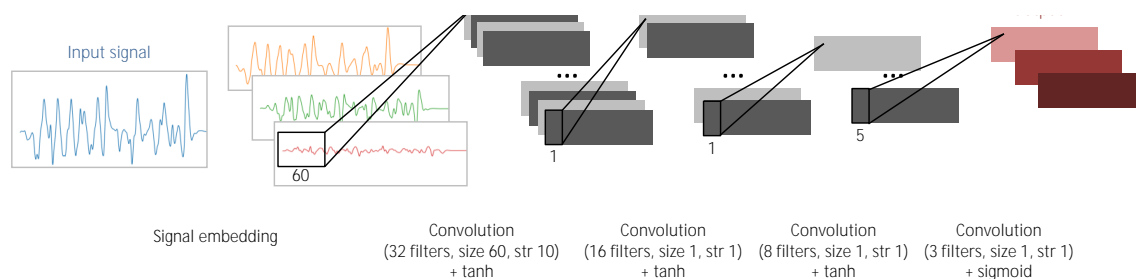


Figure 6. Architecture of the first part of the CNN, the SPN. Note that the last layer, which gives the output, is not present in the general NURSENET algorithm (Figure 2).

Training set. The training of this CNN was done as follows. For this task, we selected a subset of signals that only included easily-identifiable walks of the medical staff. Each of these signals was manually segmented by an expert using external sensor information, such as the video recordings of the equipped area. More precisely, boxes (denoted \mathbf{b}_i) were manually labeled for each step of each embedded signal. Figure 7 gives an example of step delimiting boxes. The purpose of the SPN is to produce boxes $\hat{\mathbf{b}}$ with a large Intersection over Union (IoU) score:

$$\text{IoU}(\hat{\mathbf{b}}) \doteq \max_j \frac{|\mathbf{b}_j \cap \hat{\mathbf{b}}|}{|\mathbf{b}_j \cup \hat{\mathbf{b}}|} \quad (5)$$

i.e., boxes that are nearly equal to at least one of the ground-truth boxes. More precisely, the output of the network is a matrix $W \in \mathbb{R}^{T \times K}$, with T being the signal length and K the number of different box sizes. Here, W_t^k is interpreted as the probability that the box b_t^k starting at time t and of size 20, 30, or 40 (for resp. $k = 1, 2$, or 3) has a large IOU score. Those durations are chosen according to typical step durations of young healthy individuals. In line with [51], we define positive instances of boxes, i.e., b_t^k such that $\text{IOU}(b_t^k) > \sqrt{0.7}$, and negative instances, i.e., b_t^k such that $\text{IOU}(b_t^k) < \sqrt{0.3}$. It should be noted that in the original paper, authors used 0.7 (resp. 0.3) for the upper threshold (resp. the lower threshold). However, in their case, IOU was computed between 2D boxes, instead of 1D in our case. All other boxes were considered neutral boxes and did not participate in the training process. The loss function \mathcal{L} is then defined as:

$$\mathcal{L} = \sum_t \sum_{k \in [1,2,3]} \mathbf{1}_{\text{IOU}(b_t^k) > \sqrt{0.7}} \log(W_t^k) + \mathbf{1}_{\text{IOU}(b_t^k) < \sqrt{0.3}} \log(1 - W_t^k). \quad (6)$$

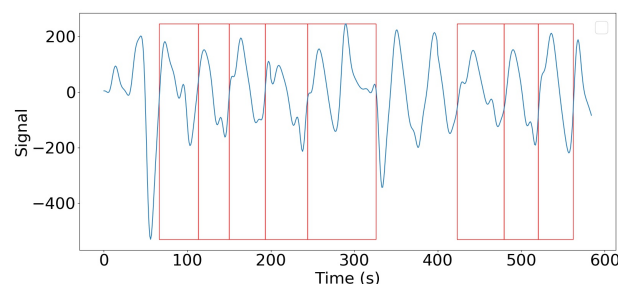


Figure 7. Example of the boxes' delimiting steps in a walk signal. Only parts of the signal that can be undoubtedly connected to a step are labeled.

Training process. The weights of each layer were initialized using i.i.d. centered Gaussian variables with a standard deviation of 0.2. Regularization was done using dropout (with $p = 0.5$) after each BatchNorm layer. Equation (6) was optimized using stochastic gradient descent with a learning rate of 10^{-5} , decreasing geometrically ($\times 0.9$) every 10 epochs, and the Nesterov momentum strategy. The learning was stopped using the early stopping principle.

Remark. After its training, SPN achieved good, but lower than the state-of-the-art performance: for an IoU threshold of 0.7 (resp 0.9), it achieved an Average Precision (AP) of 85.2% (resp. 55.3%).

2.3. Data Collection

The objective of NURSENET was to provide an unobtrusive tool to monitor physical activity of elderly individuals in a nursing home. To this end, a partner retirement house was chosen to collect the signals using the following setting. A corridor and a nearby common room were equipped with the system (Figure 8), and signals were recorded and labeled (event types and step boxes for the signals concerned) by an expert using two cameras. The areas, a corridor and a common room, were chosen for their accessibility (i.e., ease of installation of the system and event labeling) and the great frequency of people passing through. Indeed, the corridor is surrounded by the dining room, a caregiver office, and another corridor to another aisle of the nursing home, making it an important crossing point. Data were labeled according to the type of event (walking, rolling a cart, pushing a chair, etc.), the status of the person (patient or caregiver), and the number of persons on the floor.

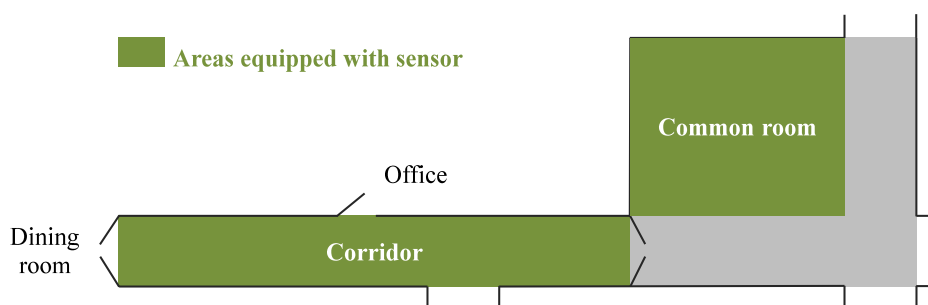
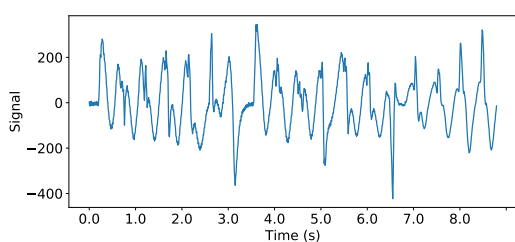


Figure 8. System installation in the nursing home.

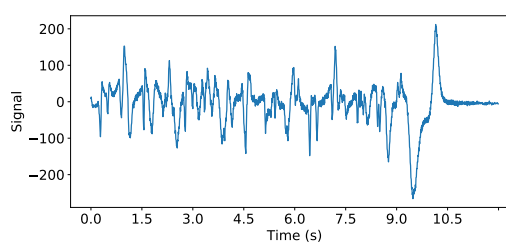
The final database contained 93 labeled signals with a labeling resolution of 1 s for event labels (e.g walk, rolling a cart, etc.) and 10 ms for the steps of the concerned signals. Table 1 shows the number of events of each main class. It should be noted that the majority of signals are from caregivers. Consequently, it is crucial to identify them as accurately as possible in order not to alter a potential elderly activity monitoring system. Figure 9 shows two examples of signals recorded in the nursing home.

Table 1. Number of instances in the dataset. Signals have an average duration of 10 s. Note that in the *walking, >1 person* category, all the walkers were either all elderly or all medical staff.

Event	Number of Instances
Walking, 1 person, staff	42
Walking, 1 person, elderly	16
Walking, >1 person	11
Wheelchair	9
Wheelchair pushed by another	5
Walking with a cart	5
Other events	5



(a) Walking of a staff member (SNR \approx 23 dB).



(b) Walking of an elderly person (SNR \approx 18 dB).

Figure 9. Example of signal of a young healthy person from the staff (a) and an elderly (b). The first observation is that the signal generated by the elderly footsteps has a significantly smaller amplitude, leading to a lower signal-to-noise ratio. We recognize some similarities between these signals as step patterns, as can be seen in (b). However as a whole, the (b) signal seems less repeatable than (a). This is probably due to the fact that the walking of an elderly can be less regular and less pronounced than a young healthy person whose feet leave and hit the floor at a regular pace.

3. Results

In this section, we thoroughly evaluate the performance and behavior of NURSENET on multiple types of signals. First, the accuracy of NURSENET is studied over each subgroup of our database, and we take a closer look at misclassified signals. Then, a complete ablation analysis is done to show the relative improvement added by each part of the algorithm, and our approach is compared to another robust off-the-shelf classification algorithm. The algorithm was tested and run with a Python

implementation on a laptop with a 2.4-GHz Intel Core i7 processor. The training of the several stages (i.e., CDL, SPN, and NURSENET) took a few hours. However, the complete inference process was very fast (nearly instantaneous for a 10-s signal).

3.1. Performance Evaluation

The algorithm was evaluated by performing a stratified split of the dataset into training and testing sets (70 and 30%), resulting in similar distributions of data in both sets. The hyperparameters of CDL, SPN, and NURSENET were chosen using cross-validation on the training set. The separation was done using stratified random sampling along the three labels *single walk*, where one individual was walking on the sensor, *multiple walks*, where strictly more than one individual was walking at the same time on the sensor, and *others*, where one non-walk-related event occurred (e.g., pushing a wheelchair or a cart). The signals were labelled as *staff* if all the persons involved were parts of the medical staff and *elderly* otherwise. The demographics are detailed in Table 2.

Table 2. Quantity and nature of the signals contained in our dataset. Data were split into training and test using stratified random sampling along the classes single walk, multiple walks and other.

Label	Staff	Elderly
Single Walk	40	18
Multiple Walks	8	3
Other	7	17

NURSENET achieved a global AUC of 0.94(± 0.02) on the test set for the classification between staff and non-staff activities. This high performance was evenly distributed on all the labels, as shown by Figure 10. Additionally, it should be noted that the accuracy of NURSENET was similar for both the test set and the entire dataset, which highlights the generalizability of our approach. Table 3 presents two examples of classification obtained by NURSENET for different thresholds, and Figure 11 shows examples of misclassified signals.

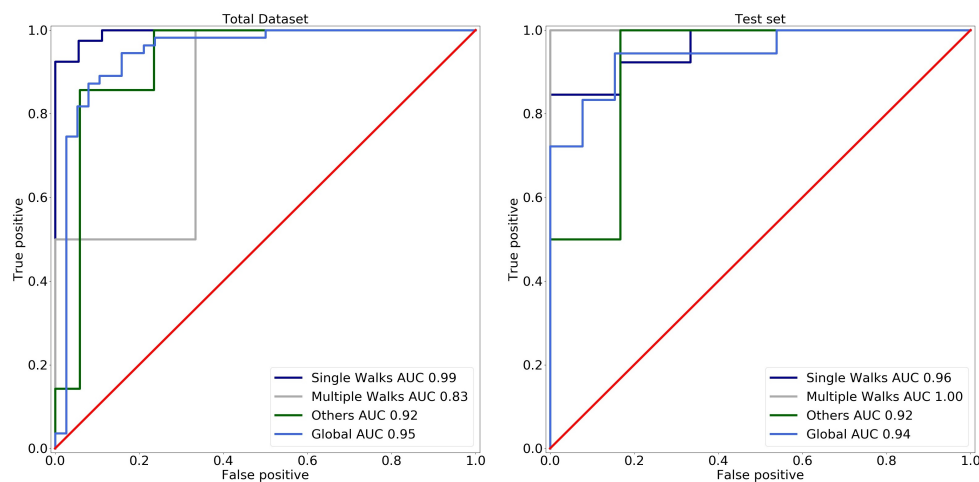


Figure 10. ROC obtained when classifying between medical staff and elderly individuals for different activities (single individual walk, multiple individuals walks, other) using NURSENET on (left) the entire dataset and (right) the test set. Performances are similar in both cases.

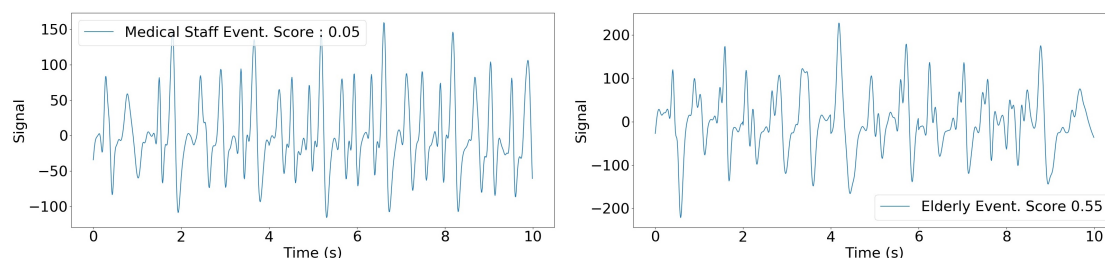


Figure 11. Examples of signals misclassified by NURSENET. **(Left)** Nurse pushing a wheelchair. Note that the signal is irregular. **(Right)** Elderly walking. Contrary to the previous example (Figure 9), this signal presents higher periodicity and amplitude, and the recorded individual is healthier.

Table 3. Confusion matrix for different thresholds τ , i.e., activity is classified as staff if the network produces an output greater than or equal to τ . Note that the staff activities that are confused by the networks are essentially the nurse pushing medical material such as wheelchairs or carts on the sensor.

		Staff			Elderly		
		Single Walk	Multiple Walks	Other	Single Walk	Multiple Walks	Other
$\tau = 0.2$							
Test	Classified Staff	13	3	1	2	0	0
	Classified Elderly	0	0	1	4	1	6
All	Classified Staff	40	7	5	4	1	1
	Classified Elderly	0	1	2	14	2	16
$\tau = 0.4$							
Test	Classified Staff	12	2	1	1	0	0
	Classified Elderly	1	1	1	5	1	6
All	Classified Staff	39	3	3	1	0	1
	Classified Elderly	1	5	4	17	3	16

3.2. Ablation Analysis

This subsection aims to perform an ablation analysis on NURSENET and compare it to non-deep learning-based approaches. To achieve this result, a version of NURSENET was trained without the SPN pre-training part. In other words, the network was trained from scratch on the final classification task. Another version of NURSENET was trained, in which there was neither dictionary embedding nor SPN pre-training, i.e., it used directly the raw signal as an input, and the whole CNN was trained on the classification task. Finally, NURSENET was compared to a Random Forest (RF) algorithm, which is known to provide good results. The RF features were 10-s padded signals decomposed using dictionary embedding and the atoms learned in Section 2.2.2. Results are presented in Figure 12.

It is important to notice that while all the previously-described algorithms achieved good performance on the entire dataset (which includes the training set), their AUCs significantly dropped on the test set, hence suggesting significant overfitting. This was particularly true for RF, whose AUC was only $0.72(\pm 0.04)$ on the test set, highlighting the advantage of a CNN-based approach in our setting. Additionally, NURSENET significantly outperformed both its simplified versions, which illustrates the advantages of our training method. Finally, it should be noted that NURSENET without embedding, which is a CNN directly applied to the raw signal, achieved the surprisingly high AUC of $0.84(\pm 0.03)$.

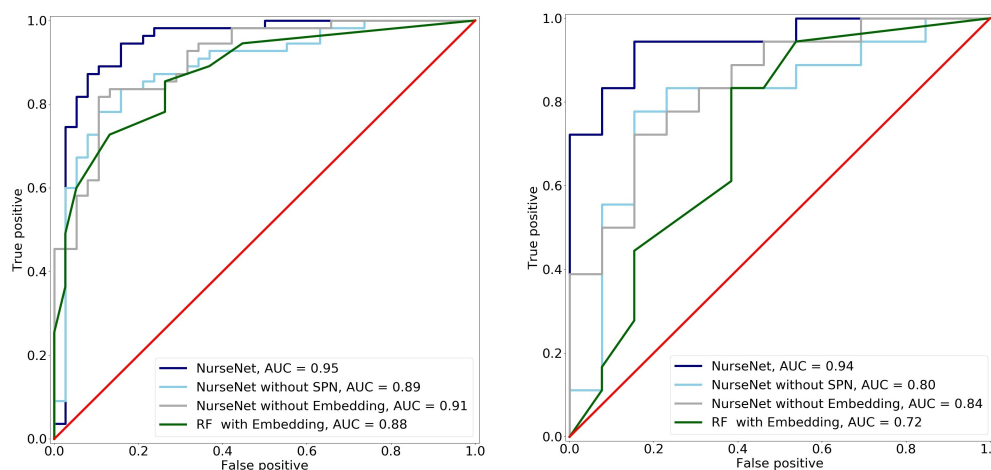


Figure 12. ROC obtained when classifying staff/elderly activities using the different variants of NURSENET and RF on (left) the entire dataset and (right) the test set. The corresponding AUCs are displayed in the figure legend. Note that NURSENET outperformed all the other algorithms in both cases. It also generalized well, as it had consistent performances between the test set and the entire dataset, contrary to the others.

4. Discussion

Classification across all labels. It is interesting to note that NURSENET performed well across all labels described in Table 2. This might seem counter-intuitive at first as the algorithm was pretrained toward detecting steps and some signals contained non-gait-related activities (e.g., pushing a cart). A possible explanation comes from the fact that all signals contained some elements of walks. In the example where a nurse was pushing a cart, while not obvious to the untrained eye, it was possible that some characteristics of the medical staff walk were encoded in the signal. Therefore, the features learned by the SPN could be relevant to correctly assess the signals.

Multiple training. Interestingly, the CNN without the dictionary encoding, nor SPN pre-trainings performed significantly better than the CNN with the sole dictionary encoding. This may be explained as follows. Both the dictionary encoding and the SPN are two complementary stages of step detection: the dictionary extracted features linked to step characteristics, and the SPN was trained to use these features to segment the signal into steps. On the one hand, when the CNN was used with only dictionary embedding, the training task became more complex as the network had to learn how to use these features to recognize steps while trying to classify these signals. On the other hand, the regular CNN algorithm could use any non-step-related features to perform the classification. Consequently, its objective function was significantly less constrained and had a larger number of reasonable local minima, making it easier to train. However, it should be noted that when properly guided toward a step-related solution (NURSENET), the CNN appeared to perform even better, which emphasized the interest of our approach.

NURSENET vs. shallow. On this classification task, the random forest model gave reasonable performance on the test set (AUC 0.72). However, it was significantly lower than its overall performance (AUC 0.88), which indicates that it was overfitting. We observed similar behavior for most off-the-shelf algorithms (e.g., SVM). This may be explained by the complex nature of the signal: while NURSENET (as well as other deep network algorithms) can learn relevant features to describe the signal, shallow algorithms tackle a significantly harder task. It is the authors belief that with properly-designed features, RF may achieve performance similar to CNN. However, due to the very good AUC obtained by NURSENET, we chose not to push further into the manual feature engineering, as it is always possible to use features learned from NURSENET to improve the results of other algorithms.

In-depth classification. As seen in Section 3, NURSENET was able to separate medical staff and elderly activities with very high accuracy (AUC 0.94). Future work will focus on separating different categories of activity as each one of them entails different indicators regarding elderly activity monitoring. However, this task requires a significantly larger dataset, and therefore, the installation of NURSENET on multiple locations.

Signal Quality. Serra et al. [10] used a similar sensor to NURSENET and showed examples of footstep recordings that proved to be very reliable in their interpretability. For instance, different phases of a footstep could be seen: heel, toe, weight transfer, and foot removal. Moreover, these footsteps were recorded multiple times with different persons, showing that the generated signal was repeatable, illustrating it with a superimposition of several footsteps of one person (see Figure 13). However, in our application, the output was of significantly lower quality. There were multiple factors that may account for this difference. The authors of [10] recorded their signals in a controlled environment: a single band was directly linked to an acquisition unit dedicated to laboratory prototyping. These laboratory conditions lead to high-quality output signals, and events such as walks could be easily decomposed. However, in our setting and due to scale and cost issues, the signal was collected on multiple large surfaces and sent through a processing unit designed to perform multiple tasks including the amplification of the signal, which may alter the quality of the data. Moreover, in addition to the previously-acknowledged limitations (i.e., non-homogeneous piezoelectric coefficient across space), the system appeared to be sensitive to humidity. Indeed, in this setting, the signal was collected through crimps hooked to the sensor and linked to a printed circuit. This connection was vulnerable to humidity that may be in the surfaces in contact with the system (e.g., adjacent walls), which may even lead to corrosion. Besides, this humidity may enter in contact with the edges of the sensor, thus creating an additional resistance between the two sides of the device. Depending on the characteristics of the humidity, these two phenomena may significantly alter the output of the unit. Future improvement of the sensor would aim to reduce these effects.

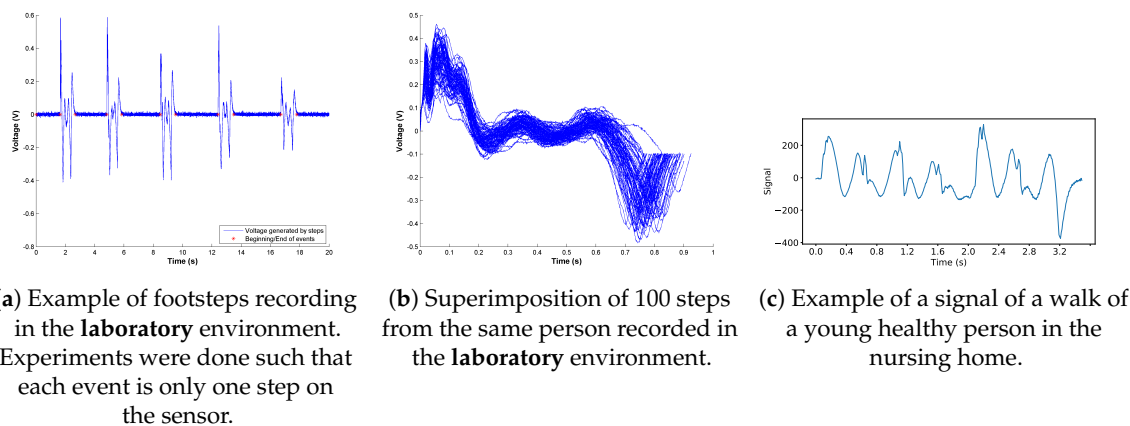


Figure 13. When considering signals recorded in laboratory conditions (a,b), we notice a certain regularity and good repeatability of events that are not of the same quality when looking at signals taken in the nursing home (c). Figure (a,b) were taken with the author's agreement.

Individual identification. While NURSENET achieved very high accuracy for classifying the elderly and medical staff, it is not able at the moment to recognize individuals. This is due to the fact that the dataset used only included a handful of signals for each individual. However, it is the authors' belief that with significantly more data, NURSENET can be improved to distinguish between individuals, albeit at a lower accuracy than the staff/elderly separation. Future work will aim at improving NURSENET in that regard. It should also be noted that despite this limitation, it was already possible to assign the activity to a person with reasonable accuracy using external information, using for instance the location of the system (e.g., if the system is installed near or inside a person's room).

Author Contributions: Conceptualization, L.M. and J.A.; methodology, J.A.; validation, L.M. and J.A.; investigation, L.M. and J.A.; data curation, L.M.; writing, review and editing, L.M. and J.A.; supervision, J.A.

Funding: Part of this work was funded by the Industrial Data Analytics And Machine Learning chairs of ENSParis-Saclay. Ludovic Minvielle is partially supported by Tarkett.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

NURSENET	Non-invasive Unit Recognition System for the Elderly Network
SPN	Step Proposal Network
RPN	Region Proposal Network
CNN	Convolutional Neural Network
CDL	Convolutional Dictionary Learning
RF	Random Forest

References

1. Fried, L.P.; Tangen, C.M.; Walston, J.; Newman, A.B.; Hirsch, C.; Gottdiener, J.; Seeman, T.; Tracy, R.; Kop, W.J.; Burke, G.; et al. Frailty in older adults: Evidence for a phenotype. *J. Gerontol. Ser. A Biol. Sci. Med. Sci.* **2001**, *56*, M146–M157. [[CrossRef](#)] [[PubMed](#)]
2. Morley, J.E.; Vellas, B.; Van Kan, G.A.; Anker, S.D.; Bauer, J.M.; Bernabei, R.; Cesari, M.; Chumlea, W.; Doehner, W.; Evans, J.; et al. Frailty consensus: A call to action. *J. Am. Med. Dir. Assoc.* **2013**, *14*, 392–397. [[CrossRef](#)] [[PubMed](#)]
3. Vellas, B. Implementing frailty screening, assessment, and sustained intervention: The experience of the G erontop ole. *J. Nutr. Health Aging* **2015**, *19*, 673–680. [[CrossRef](#)] [[PubMed](#)]
4. Fried, T.R.; Mor, V. Frailty and hospitalization of long-term stay nursing home residents. *J. Am. Geriatr. Soc.* **1997**, *45*, 265–269. [[CrossRef](#)] [[PubMed](#)]
5. Avorn, J.; Langer, E. Induced disability in nursing home patients: A controlled trial. *J. Am. Geriatr. Soc.* **1982**, *30*, 397–400. [[CrossRef](#)] [[PubMed](#)]
6. Fiatarone, M.A.; O’Neill, E.F.; Ryan, N.D.; Clements, K.M.; Solares, G.R.; Nelson, M.E.; Roberts, S.B.; Kehayias, J.J.; Lipsitz, L.A.; Evans, W.J. Exercise training and nutritional supplementation for physical frailty in very elderly people. *N. Engl. J. Med.* **1994**, *330*, 1769–1775. [[CrossRef](#)] [[PubMed](#)]
7. McPhee, J.S.; French, D.P.; Jackson, D.; Nazroo, J.; Pendleton, N.; Degens, H. Physical activity in older age: Perspectives for healthy ageing and frailty. *Biogerontology* **2016**, *17*, 567–580. [[CrossRef](#)]
8. Blair, S.N.; Kohl, H.W.; Gordon, N.F.; Paffenbarger, R.S., Jr. How much physical activity is good for health? *Annu. Rev. Public Health* **1992**, *13*, 99–126. [[CrossRef](#)]
9. Mukhopadhyay, S.C. Wearable sensors for human activity monitoring: A review. *IEEE Sens. J.* **2014**, *15*, 1321–1330. [[CrossRef](#)]
10. Serra, R.; Knittel, D.; Di Croce, P.; Peres, R. Activity recognition with smart polymer floor sensor: Application to human footprint recognition. *IEEE Sens. J.* **2016**, *16*, 5757–5775. [[CrossRef](#)]
11. Cucchiara, R.; Prati, A.; Vezzani, R.; Emilia, R. A multi-camera vision system for fall detection and alarm generation. *Expert Syst.* **2007**, *24*, 334–345. [[CrossRef](#)]
12. Zhuang, X.; Huang, J.; Potamianos, G.; Hasegawa-Johnson, M. Acoustic fall detection using Gaussian mixture models and GMM supervectors. In Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 19–24 April 2009; pp. 69–72.
13. Jonghee, H.; Hyo Sun, J.; Beom Suk, J.; Kwang Suk, P. Gait detection from three dimensional acceleration signals of ankles for the patients with Parkinson’s disease. In Proceedings of the IEEE The International Special Topic Conference on Information Technology in Biomedicine, Ioannina, Epirus, Greece, 26 October 2006.
14. Bourke, A.; O’Brien, J.; Lyons, G. Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait Posture* **2007**, *26*, 194–199. [[CrossRef](#)] [[PubMed](#)]
15. Mahoney, E.L.; Mahoney, D.F. Acceptance of wearable technology by people with Alzheimer’s disease: Issues and accommodations. *Am. J. Alzheimer’s Dis. Other Dementias* **2010**, *25*, 527–531. [[CrossRef](#)] [[PubMed](#)]

16. Taborri, J.; Palermo, E.; Rossi, S.; Cappa, P. Gait partitioning methods: A systematic review. *Sensors* **2016**, *16*, 66. [[CrossRef](#)] [[PubMed](#)]
17. Muro-De-La-Herran, A.; Garcia-Zapirain, B.; Mendez-Zorrilla, A. Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications. *Sensors* **2014**, *14*, 3362–3394. [[CrossRef](#)] [[PubMed](#)]
18. Ao, B.; Wang, Y.; Liu, H.; Li, D.; Song, L.; Li, J. Context Impacts in Accelerometer-Based Walk Detection and Step Counting. *Sensors* **2018**, *18*, 3604. [[CrossRef](#)]
19. Derawi, M.O. Accelerometer-based gait analysis, a survey. In Proceedings of the Nor Informasjonssikkerhetskoneranse NISK 2010, Gjøvik, Norway, 23–24 November 2010.
20. Wang, W.; Liu, A.X.; Shahzad, M. Gait recognition using wifi signals. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany, 12–16 September 2016; pp. 363–373.
21. Jiang, H.; Cai, C.; Ma, X.; Yang, Y.; Liu, J. Smart home based on WiFi sensing: A survey. *IEEE Access* **2018**, *6*, 13317–13325. [[CrossRef](#)]
22. Pan, S.; Yu, T.; Mirshekari, M.; Fagert, J.; Bonde, A.; Mengshoel, O.J.; Noh, H.Y.; Zhang, P. Footprintid: Indoor pedestrian identification through ambient structural vibration sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2017**, *1*, 89. [[CrossRef](#)]
23. Fagert, J.; Mirshekari, M.; Pan, S.; Zhang, P.; Noh, H.Y. Characterizing left-right gait balance using footstep-induced structural vibrations. In *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2017*; International Society for Optics and Photonics: Bellingham, WA, USA, 2017; Volume 10168, p. 1016819.
24. Yun, J.; Woo, W.; Ryu, J. User identification using user's walking pattern over the ubiFloorII. In *Computational Intelligence and Security*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 949–956.
25. Suutala, J.; Fujinami, K.; Röning, J. Gaussian process person identifier based on simple floor sensors. In *Smart Sensing and Context*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 55–68.
26. Rimminen, H.; Lindström, J.; Sepponen, R. Positioning accuracy and multi-target separation with a human tracking system using near field imaging. *Int. J. Smart Sens. Intell. Syst.* **2009**, *2*, 156–175. [[CrossRef](#)]
27. Henry, R.; Matti, L.; Raimo, S. Human tracking using near field imaging. In Proceedings of the 2008 Second International Conference on Pervasive Computing Technologies for Healthcare, Tampere, Finland, 30 January–1 February 2008; pp. 148–151.
28. Valtonen, M.; Maentausta, J.; Vanhala, J. Tiletrack: Capacitive human tracking using floor tiles. In Proceedings of the 2009 IEEE International Conference on Pervasive Computing And Communications, Galveston, TX, USA, 9–13 March 2009; pp. 1–10.
29. Tanaka, O.; Ryu, T.; Hayashida, A.; Moshnyaga, V.G.; Hashimoto, K. A smart carpet design for monitoring people with dementia. In *Progress in Systems Engineering*; Springer: Cham, Switzerland, 2015; pp. 653–659.
30. Suutala, J.; Röning, J. Methods for person identification on a pressure-sensitive floor: Experiments with multiple classifiers and reject option. *Inf. Fusion* **2008**, *9*, 21–40. [[CrossRef](#)]
31. Rodríguez, R.V.; Lewis, R.P.; Mason, J.S.; Evans, N.W. Footstep recognition for a smart home environment. *Int. J. Smart Home* **2008**, *2*, 95–110.
32. Sacco, I.d.C.N.; Amadio, A.C. A study of biomechanical parameters in gait analysis and sensitive cronaxie of diabetic neuropathic patients. *Clin. Biomech.* **2000**, *15*, 196–202. [[CrossRef](#)]
33. Dockstader, S.L.; Berg, M.J.; Tekalp, A.M. Stochastic kinematic modeling and feature extraction for gait analysis. *IEEE Trans. Image Process.* **2003**, *12*, 962–976. [[CrossRef](#)] [[PubMed](#)]
34. Zhang, C.; Liu, W.; Ma, H.; Fu, H. Siamese neural network based gait recognition for human identification. In Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 20–25 March 2016; pp. 2832–2836.
35. González, R.C.; López, A.M.; Rodríguez-Uría, J.; Álvarez, D.; Alvarez, J.C. Real-time gait event detection for normal subjects from lower trunk accelerations. *Gait Posture* **2010**, *31*, 322–325. [[CrossRef](#)] [[PubMed](#)]
36. Ying, H.; Silex, C.; Schnitzer, A.; Leonhardt, S.; Schiek, M. Automatic step detection in the accelerometer signal. In Proceedings of the 4th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2007), Aachen, Germany, 26–28 March 2007; pp. 80–85.
37. Bae, J.; Tomizuka, M. Gait phase analysis based on a Hidden Markov Model. *Mechatronics* **2011**, *21*, 961–970. [[CrossRef](#)]

38. Brajdic, A.; Harle, R. Walk detection and step counting on unconstrained smartphones. In Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Zurich, Switzerland, 8–12 September 2013; pp. 225–234.
39. Barralon, P.; Vuillerme, N.; Noury, N. Walk detection with a kinematic sensor: Frequency and wavelet comparison. In Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, USA, 30 August–3 September 2006; pp. 1711–1714.
40. Zhang, Y.; Ogunbona, P.O.; Li, W.; Munro, B.; Wallace, G.G. Pathological gait detection of parkinson's disease using sparse representation. In Proceedings of the 2013 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Hobart, Australia, 26–28 November 2013; pp. 1–8.
41. Li, M.; Tian, S.; Sun, L.; Chen, X. Gait Analysis for Post-Stroke Hemiparetic Patient by Multi-Features Fusion Method. *Sensors* **2019**, *19*, 1737. [[CrossRef](#)] [[PubMed](#)]
42. Mairal, J.; Ponce, J.; Sapiro, G.; Zisserman, A.; Bach, F.R. Supervised dictionary learning. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2009; pp. 1033–1040.
43. Huang, K.; Aviyente, S. Sparse representation for signal classification. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2007; pp. 609–616.
44. Aharon, M.; Elad, M.; Bruckstein, A. *rmk*-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.* **2006**, *54*, 4311–4322. [[CrossRef](#)]
45. Mairal, J.; Bach, F.; Ponce, J.; Sapiro, G. Online learning for matrix factorization and sparse coding. *J. Mach. Learn. Res.* **2010**, *11*, 19–60.
46. Garcia-Cardona, C.; Wohlberg, B. Convolutional dictionary learning: A comparative review and new algorithms. *IEEE Trans. Comput. Imaging* **2018**, *4*, 366–381. [[CrossRef](#)]
47. Poschadel, N.; Moghaddamnia, S.; Alcaraz, J.C.; Steinbach, M.; Peissig, J. A dictionary learning based approach for gait classification. In Proceedings of the 2017 22nd International Conference on Digital Signal Processing (DSP), London, UK, 23–25 August 2017; pp. 1–4.
48. Ouyang, W.; Wang, X.; Zeng, X.; Qiu, S.; Luo, P.; Tian, Y.; Li, H.; Yang, S.; Wang, Z.; Loy, C.C.; et al. Deepid-net: Deformable deep convolutional neural networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 2403–2412.
49. Cai, Z.; Fan, Q.; Feris, R.S.; Vasconcelos, N. A unified multi-scale deep convolutional neural network for fast object detection. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; pp. 354–370.
50. Li, G.; Yu, Y. Deep contrast learning for salient object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 478–487.
51. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2015; pp. 91–99.
52. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
53. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436. [[CrossRef](#)]
54. Suder, K.; Drepper, F.R.; Schiek, M.; Abel, H.H. One-dimensional, nonlinear determinism characterizes heart rate pattern during paced respiration. *Am. J. Physiol. Heart Circ. Physiol.* **1998**, *275*, H1092–H1102. [[CrossRef](#)]
55. Rubin, J.; Parvaneh, S.; Rahman, A.; Conroy, B.; Babaeizadeh, S. Densely connected convolutional networks and signal quality analysis to detect atrial fibrillation using short single-lead ECG recordings. In Proceedings of the 2017 Computing in Cardiology (CinC), Rennes, France, 24–27 September 2017; pp. 1–4.
56. Kiranyaz, S.; Ince, T.; Gabbouj, M. Real-time patient-specific ECG classification by 1-D convolutional neural networks. *IEEE Trans. Biomed. Eng.* **2015**, *63*, 664–675. [[CrossRef](#)] [[PubMed](#)]
57. Serra, R. Développement et Caractérisation d'un Système de sol Piézoélectrique Intelligent. Application à la Détection des Chutes. Ph.D. Thesis, Université de Strasbourg, Strasbourg, France, 2017.
58. Nair, V.; Hinton, G.E. Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, 21–24 June 2010; pp. 807–814.
59. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2014; pp. 3320–3328.

60. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
61. Nesterov, Y. A method for unconstrained convex minimization problem with the rate of convergence $O(1/k^2)$. *Doklady AN USSR* **1983**, *269*, 543–547.
62. Prechelt, L. Early stopping-but when? In *Neural Networks: Tricks of the Trade*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 55–69.
63. Rubinstein, R.; Zibulevsky, M.; Elad, M. Double sparsity: Learning sparse dictionaries for sparse signal approximation. *IEEE Trans. Signal Process.* **2009**, *58*, 1553–1564. [[CrossRef](#)]
64. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2012; pp. 1097–1105.
65. Pal, K.K.; Sudeep, K. Preprocessing for image classification by convolutional neural networks. In Proceedings of the 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 20–21 May 2016; pp. 1778–1781.
66. Antonsson, E.K.; Mann, R.W. The frequency content of gait. *J. Biomech.* **1985**, *18*, 39–47. [[CrossRef](#)]
67. Um, T.T.; Pfister, F.M.J.; Pichler, D.; Endo, S.; Lang, M.; Hirche, S.; Fietzek, U.; Kulić, D. Data augmentation of wearable sensor data for Parkinson’s disease monitoring using convolutional neural networks. *arXiv* **2017**, arXiv:1706.00527.
68. Minvielle, L.; Atiq, M.; Serra, R.; Mougeot, M.; Vayatis, N. Fall detection using smart floor sensor and supervised learning. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Seogwipo, Korea, 11–15 July 2017; pp. 3445–3448.
69. Robinson, J.L.; Smidt, G.L. Quantitative gait evaluation in the clinic. *Phys. Ther.* **1981**, *61*, 351–353. [[CrossRef](#)]
70. Gabell, A.; Nayak, U. The effect of age on variability in gait. *J. Gerontol.* **1984**, *39*, 662–666. [[CrossRef](#)]
71. Garcia-Cardona, C.; Wohlberg, B. Subproblem coupling in convolutional dictionary learning. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 1697–1701.
72. Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **2011**, *3*, 1–122. [[CrossRef](#)]
73. Wohlberg, B. SPORCO: A Python package for standard and convolutional sparse representations. In Proceedings of the 15th Python in Science Conference, Austin, TX, USA, 11–17 July 2016; pp. 1–8.
74. Holden, M.K.; Gill, K.M.; Magliozzi, M.R.; Nathan, J.; Piehl-Baker, L. Clinical gait assessment in the neurologically impaired: Reliability and meaningfulness. *Phys. Ther.* **1984**, *64*, 35–40. [[CrossRef](#)] [[PubMed](#)]
75. Rifkin, R.M. Everything Old Is New Again: A Fresh Look at Historical Approaches in Machine Learning. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2002.

