Highly Accurate Bathroom Activity Recognition Using Infrared Proximity Sensors

Kévin Chapron, Patrick Lapointe, Kévin Bouchard, *Member, IEEE*, and Sébastien Gaboury, *Member, IEEE*

***Abstract*—Among elderly populations over the world, a high percentage of individuals are affected by physical or mental diseases, greatly influencing their quality of life. As it is a known fact that they wish to remain in their own home for as long as possible, solutions must be designed to detect these diseases automatically, limiting the reliance on human resources. To this end, our team developed a sensors platform based on infrared proximity sensors to accurately recognize basic bathroom activities such as going to the toilet and showering. This article is based on the body of scientific literature which establish evidences that activities relative to corporal hygiene are strongly correlated to health status and can be important signs of the development of eventual disorders. The system is built to be simple, affordable and highly reliable. Our experiments have shown that it can yield an F-Score of 96.94%. Also, the durations collected by our kit are approximately 6 seconds apart from the real ones; those results confirm the reliability of our kit.**

***Index Terms*—Activity recognition, bathroom, bathroom activity recognition, health monitoring, sensor, smart home, smart home kit.**

I. Introduction

IN OUR society, the proportion of the elderly population is constantly increasing. The United Nations revealed a report about the world population [1] stating that 12.7% of the current world population is aged over 60. In fact, this number is constantly growing, and is expected to reach 21.3% by 2050. This represents an imposing group of 2 080 million persons. Among these aging individuals, many cases of physical conditions or mental diseases appear [2]–[4], jeopardizing the quality of life of the affected populations. In the world of eldercare, this is a well-known fact that elderly people should remain in their own house, or at least retain their functional autonomy, for the longest time possible to prevent decline or enhance their quality of life [5]. This so-called aging-in-place is not without risk; without a proper medical follow-up, there is the possibility of a developing condition being missed or ignored, thus causing damage in the long term or jeopardizing the safety of the person.

Clinicians utilize a variety of tests regarding the execution of the basic and the instrumental Activities of Daily Living (*ADLs*) as a strong indicator of the level of autonomy enjoyed by a person [6], [7]. Several metrics of the execution of ADLs can help understanding unexposed conditions of the person and even lead to early diagnosis [8]. Accordingly, a lot of researchers have worked about recognizing such ADLs (e.g., cooking, bathing, eating, etc.) automatically through the implementation of ambient intelligence and distributed sensors nodes [9], [10]. The recognition of ADLs could then, in turn, help creating cognitive or physical orthosis through ambient intelligence [11]. Alternatively, activity recognition has also been used to detect anomalies in the behaviors of observed persons [12]. In fact, the daily routine is unmistakably correlated to health status, meaning that being able to monitor automatically these routines from technological tools and detecting changes in their realizations could help clinicians to detect a concealed disease or physical condition.

Corporal hygiene is a particularly good indicator of eventual disorders. Indeed, a person affected by a cognitive disorder not only will spend less and less time to take care of their own hygiene [13], [14], but also someone affected by a physical disease will take longer to perform his activities such as showering [2], [15]. Hence, the corporal hygiene is a dependable indicator for both cognitive and physical conditions. In this paper, our team proposes a sensors-based system able to recognize two daily activities about corporal hygiene : going to toilet and showering (or bathing). The kit is built with the determination to keep it affordable, easily to deploy in any housing architecture, and highly reliable and accurate. The focus is on the constant monitoring of the toilet activities to prevent further impairment by alerting healthcare providers to examine eventual problems.This kit should enable better information gathering about the bathroom activities for researchers on smart home, and thus, in turn, enable them to improve their pattern recognition or machine learning methods.

II. Related Works

The body of literature surrounding ambient intelligence and activity recognition is massive. Nevertheless, it can be narrowed down since researchers around the world are investigating a wide variety of systems for activity recognition, whereas our current

Manuscript received August 15, 2019; revised November 25, 2019; accepted December 28, 2019. Date of publication December 31, 2019; date of current version August 5, 2020. *(Corresponding author: Sebastien Gaboury.)*

The authors are with the Laboratoire d'Intelligence Ambiante pour la Reconnaissance d'Activités (LIARA), 555 Boulevard de l'Université, Chicoutimi, QC G7H 2B1, Canada (e-mail: kevin.chapron1@uqac.ca; patrick.lapointe1@uqac.ca; kevin.bouchard@uqac.ca; sebastien.gaboury@uqac.ca).

Digital Object Identifier 10.1109/JBHI.2019.2963388

research focus is on the creation of an easily-deployable kit of sensors to enhance any bathroom. To conceive the kit adequately, the type of sensors used in recognition systems were surveyed.

To begin with, smart homes almost always embed industrial hardware built directly with the equipment [16]–[18]. For example, most of those environments are using sensors in faucet to monitor the water-flowing activity or pressure mats. Even if those sensors are really interesting, they have some limitations in our situation. Sensors like that must indeed be installed with the plumbing (making it tough to install afterwards, or unacceptable considering the cost), similarly as pressure mats which need to be built with the initial flooring, also making their installation challenging in existing housing facilities. Furthermore, in the case of water-flowing sensors, even if you have the information that the faucet is open, this does not give any certainty that the person is really in the shower. In addition, installing a monitor on the toilet tank cannot provide certainty that the person used the toilet since this monitor would only provide one with the information the toilet has been flushed. Thus, if someone just throws a tissue in the toilet and then flushes the toilet, the system would wrongly deduce the person used the toilet. Nevertheless, another sensor is widely deployed in smart homes, and can be a cost-effective alternative : Passive Infrared sensors (*PIR*). In fact, these PIRs are easy to install and they provide information about motions in a coverage area. Yet, when someone is standing in front of the sensor without moving (or very slowly), it cannot detect anything, making their interpretation difficult. Also, this kind of sensor has a latency after being triggered: once its state changed, it takes a few seconds before changing it back to a wait-for-trigger state. PIRs based smart homes are often dealing with small and uncertain datasets which are challenging to exploit for accurate ADLs recognition. In opposition, some teams, such as the Grenoble Health Smart Home [19], prefer using cameras which are a rich source of information, albeit being complex to process. Nevertheless, any sensor able to record images are deemed unwelcome in bathroom environments for obvious reasons of privacy and nudity.

Next, some researchers specialized their work in the recognition of activities done in the bathroom. Forgary *et al.* [20] used an oriented microphone taped on water pipes, while Chen *et al.* [21] used a basic microphone placed in the bathroom. This last method is also used by Debes *et al.* [22]; this sensor is able to detect whenever water is flowing through the pipes, whether it is the faucet or the flush. This achieves great results to detect events, but it cannot retrieve the duration of the activity performed by the inhabitant. Moreover, this kind of device must be placed directly on the main water circuit, which is generally hidden behind walls, making it difficult to access. The general microphone method is very easy to install, but it is also very noise sensitive. Indeed, if the person receives a phone call, or talks while being in the bathroom, it may disturb the device, making the recognition inaccurate. Moreover, for the same reasons as the cameras, having a microphone constantly recording sounds emitted can be rather privacy-invasive, making the inhabitant uncomfortable with the technology. A last work realized by Tapia *et al.* [23] used sensors similar as the well-known electromagnetic contacts. Those are very simple sensors easily installable and extremely accurate. Yet, as they only change their state when someone activates it (e.g.: by flushing the toilet), they can record the exact time of their activation but are unable to measure the duration of the performed activity. In our case, since the observations must be performed over the time taken by the inhabitant to do an activity, such devices are inadequate.

Finally, as all sensors described above lack in robustness, accuracy, or in their acceptance, the team opted for a type of sensor that has not been used for this purpose in the literature : an Infrared Proximity Sensor (*IRPS*).[1] Indeed, microphone are very noise-sensitive or not easily placed, cameras are out of the scope because of our environment and sensors like electromagnetic contacts do not measure the time taken by the inhabitant to perform his activities, it results in the fact that no already adopted sensors can be used in our case. An IRPS is based on Infrared waves and is constantly measuring distance via an integrated combination of Position Sensitive Detector (*PSD*), Infrared Emitting Diode (*IRED*), and signal processing circuit within a range of 150 cm. When something is interfere with the light emitted by the sensor's diode, it is able to compute the distance between the object and itself. This is the reason why this sensor has been selected instead of the others seen above. Moreover, a *PIR* would only activates when something is moving in front of it, while the *IRPS* is constantly measuring a distance. These sensors are straightforward to install and can compete with the industrial hardware found in the well-established smart homes infrastructures including ours (water-flowing sensor, microphones). Finally, privacy is preserved with this sensor since it does not record anything else than distances and timestamps (compared to cameras, microphones). All things considered, this sensor is the best for our problematic.

After this review of related works, our team decided to focus on building a cheap, non-invasive and really easy to install sensors kit for the bathroom activities that would also be easily reproducible. Moreover, it may also become very helpful for a better pattern recognition, since much of the body of literature relies on simple sensors that provide very scarce and unreliable information on those activities. Therefore, there is a fundamental limit on the improvement of such methods without improving the basic data collection.

To sum up, our work has been designed to improve the future works related to smart homes by adding the information provided by our sensor. Our team adopt the philosophy of sharing findings and improving reproducibility in smart home researches. Hence, the system demonstrated in this paper is totally accessible in our GitHub Repository.[2]

III. A New Portable Bathroom Activity Recognition System

The whole system to recognize activities has been designed to be very easy to install. In the next subsections, the physical and networking architecture, and the activity recognition system are introduced. As a side note, our team wishes the system to be

[1]Sparkfun, 2006. GP2Y0A02YK0F. [Online]. Available: https://www.sparkfun.com/datasheets/Sensors/Infrared/gp2y0a02yk_e.pdf.

[2]Available upon acceptance of the article.

(a) Sensors kit (b) Sensor placed in the bath (c) Sensor placed on the toilet

Fig. 1. Physical devices kit and positioning.

used as widely at it can be. In that regards, every schematics, 3D designs and programs to recreate it are provided on our GitHub repository.[2]

A. Physical Architecture

Our architecture is based on many Raspberry Pi devices.[3] There were chosen since the Raspberry Pi is an open-source and reputable platform. Also, this organization proposes a wide range of board from powerful devices to much smaller ones, thus creating ideal units to work with in our case. Moreover, they are very easily programmable because of they exploit a complete Operating System (*OS*) and are able to operate multiple wireless communication protocols (Wi-Fi, Bluetooth, Bluetooth Low Energy (BLE)). Being able to switch from a protocol to another is a very useful asset in the context of Internet of Things (*IoT*) devices that our system fall into. Besides, as the IRPS returns analog values and our nano-computers only takes digital ones, we had to add another chip, to transform the raw values of the sensor into readable data for our controller. To ease the development of our module, we selected the ADS1115[4] since it can transfer data through the I^2C bus. The schematics to reproduce each module is shown in Fig. 2. Both of these modules are encapsulated in a case our team designed, and can be observed in Fig. 1a (more specifically, the shower module in Fig. 1b and the toilet module in Fig. 1c). Also, since we opted to place them directly in front of the toilet/shower to continuously monitor the distance from the user, one of them had to be water-resistant (shower can cause many droplets to infiltrate the module, and thus lead to a broken sensor).

In terms of price (USD), one of our module is rather cheap since it only costs 10 $ for the Raspberry Pi Zero W, 13.95 $ for the inner sensor and 14.95 $ for the analog to digital converter (8 $ for the production version). So, one of our module is affordable at a price of 38.90 $. To have our whole system, it must be considered to buy a central unit, such as a Raspberry Pi 3 B+ which price is 35.00 $.

[3]Raspberry Pi Foundation. Raspberry Pi 3 Model B. [Online]. Available: https://www.raspberrypi.org/products/raspberry-pi-zero-w/.
[4]Texas Instruments, "ADS111x," pp. 1–53, 2018

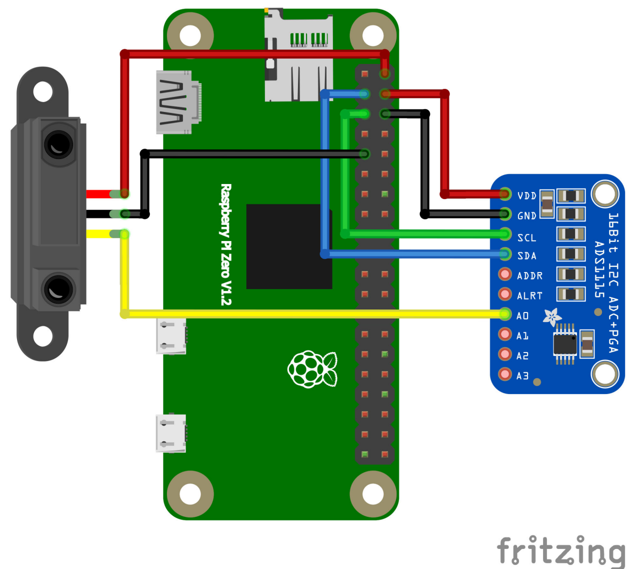

Fig. 2. Schema of a module.

Finally, our system integrates a central unit (Raspberry Pi 3), powering a wireless Access Point (*AP*) using Wi-Fi technology to which both developed modules automatically connect to, allowing them to communicate their data through the network.

B. Networking Architecture

Our whole network architecture is based on messages of maximum 1 024 bytes, sent through WebSockets over Wi-Fi. This choice was supported by our need of real-time data transfers. Indeed, basic unilateral methods such as pushing data to the server or pulling data from the sensors increased the global latency of the system. A websocket being a bilateral channel, it ensures that both the sensor and the server can always communicate in real-time.

To ensure a minimal security level and to prevent most of the attacks regarding the sniffing of our data on the network, an AES-128 algorithm is used to encrypt every message. The first 2 bytes are reserved for the data length, followed by the

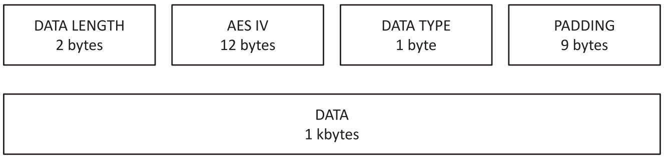

Fig. 3. Full network packet.

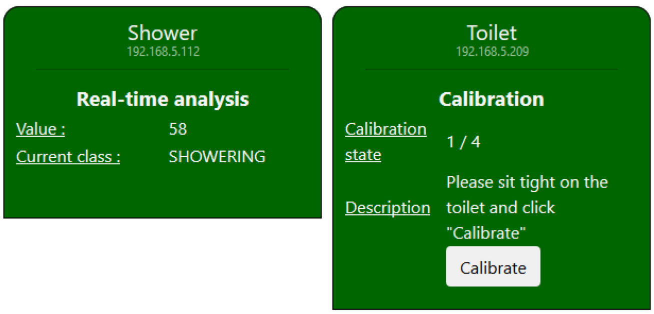

Fig. 4. User interface from a remote device.

12-bytes of the AES IV. Then, the 15th byte is allocated to the data type. Our system is able to manage 3 types of messages: the basic message to transfer data, an identification message and a calibration message. The next bytes from the 16th to the 24th one are padding, resulting in a full data of 1 000 bytes formatted using JSON Standard, encrypted using AES algorithm. The whole packet is summed up in the following Fig. 3.

Now that the packets are defined, the procedure is going to be examined. First, whenever a module is connected to the network, it sends an authentication message to the central unit, to inform it of its presence. This message contains its name, its address and all the calibration steps required to enable the sensor. It is indeed not enabled while its calibration values are not set.

The next step is for the central unit to calibrate the modules via their remote interface. This phase has been designed to be used on a mobile device so that a user can easily move while realizing the different actions asked by the modules. For example, the toilet module has to take values related to the distance of the user while he is sitting on the seat: the user has to perform the action and press the right calibration button while being in place. Once every step of the modules' calibration have been validated, it starts sending real-time values. An example of the interface available via the central unit is shown in Fig. 4.

C. Calibration Phase

Once the authentication of every module has been done, the calibration phase starts. The shower module is easier to calibrate since it simply must remain empty for a minute after the user has pressed the "Calibrate" button. In fact, during this minute, the module collects data and extracts the average value γ and the standard deviation σ_γ of the gathered data. γ represents the maximum value that the actual sensor can go to (capped at 150 cm by the sensor's capacity) before meeting a wall and σ_γ the variability of this value within a complete minute (*null* if γ has been capped by the sensor).

The toilet module require a little more complex calibration phase, since it involves users to perform actions. In a preliminary phase, the participants were asked to execute 4 actions: ρ_1 represents the action of sitting straight on the toilet, ρ_2 the action of sitting and leaning on the front of the toilet, ρ_3 the action of standing close to the toilet and ρ_4 the action of standing far from the toilet (close enough to be able to urinate). However, our experiment proved us that recognizing the standing position could often lead to inaccurate measurements. This is caused by the triggering of the module whenever a person is passing in front of the sensor which created several false positives (this issue is explained in Sec. VI). In consequence, ρ_3 and ρ_4 are no longer taken into consideration, and our algorithm only recognize activities involving sitting on the toilet.

D. Activity Recognition System

Each module has an independent recognition system, specifically designed for each task; this is also why the modules have different calibration phases previously explained. By realizing this locally, if the network ever fails for any reason, every module would keep recognizing activities until the network comes back. Those recognized activities could then be fetched by the central unit (maximum of a whole day of data).

1) Activity Recognition for the Shower: Before explaining the algorithm, let us define several variables: α is the lower bound that the sensor can measure (20 cm), β is the higher bound, defined by the minimum between 150 cm (max value the sensor can measure) and γ (calibration value).

The activity recognition system for the shower analysis is performed using a sliding window of $T = 20$ seconds using an overlapping frame of 75%, $T_\theta = 20 \times \frac{75}{100}$ seconds (those values have been found experimentally and are discussed in details in Sec. V-B2). The first step of the algorithm is to check if the mean of the current window is in the range $[\alpha, \beta - \sigma_\gamma]$. Then, every window with an average value in the specified range are stored in an array. The accumulation stops whenever 5 consecutive windows have an average value outside the range. Thus, the resulting array contains the whole potential activity.

However, to decrease our false positive rate, it is necessary to differentiate the noise (values outside the range) from the real values. To do so, an average value is computed on the raw data of every accumulated windows and again compared to $[\alpha, \beta - \sigma_\gamma]$. If the global mean is outside the range, it means our potential activity is actually mostly noise, and is ignored going forward.

The first step serves to extract candidate activities while the second one can be seen as a filter to remove failures ("activity" recognized that contains more noise than real data). For example, if some data happens to be in the range every 4 windows, it would still be a potential activity. Yet, as the window has 4 times more noise than actual data, the second step would remove it. So, those two steps extract a real activity realized into the shower.

Last but not least, it is possible to have a single real activity split into two close recognized activities. So, to finalize the event, if the last time recorded in the last event is less than 60 seconds apart from the first time recorded in the current event, those two events are merged into one (See Algorithm 1).

Algorithm 1: Activity Recognition for the Shower.

$\alpha \leftarrow 20$
$\beta \leftarrow min(150, \gamma)$

$array \leftarrow$ new detection array

for all do
  Wait for $T - T_\theta$ **new data**
  $window \leftarrow T_\theta$ previously obtained data + $(T - T_\theta)$ last collected data
  $M \leftarrow mean(window)$
  $cpt \leftarrow 0$
  $Ind_1, Ind_2 \leftarrow 0$
  if $\alpha < M < \beta$ **then**
    $Add \leftarrow true$
    $cpt \leftarrow 0$
    $Ind_1 \leftarrow$ index of $window$
 else
    if Add **then**
      $cpt \leftarrow cpt + 1$
  if Add and cpt=5 **then**
      $Ind_2 \leftarrow$ index of $window$

      Add items from $i = Ind_1$ to $i = Ind_2$
        into $array$
      $Add \leftarrow false$
      $cpt \leftarrow 0$

$data \leftarrow lastof(array)$
$\overline{array} \leftarrow mean(data)$
if $not(\alpha < \overline{array} < \beta)$ **then**
  delete *data*from *array*

Combine every recognition in *array* with less than 60 seconds of differences

2) Activity Recognition for the Toilet: Before explaining the algorithm, let recall the previously defined variables ρ_1 and ρ_2 being the distance between the module and the user respectively performing the activities "sitting straight on the toilet" and "sitting/leaning on the front of the seat". Obviously, as the sensor cannot correctly detect any distance less than 20 cm, ρ_1 has to be higher than this value (if not, it is set to 20 cm). Likewise to the previous algorithm, this recognition also uses a sliding window, but with a value of $T = 30$ seconds and an overlapping time of 80%, $T_\sigma = 30 \times \frac{80}{100}$ seconds (as for the previous algorithm, best-fitting values have been examined more deeply in Sec. V-B2).

For this algorithm, the first step is exactly as the previous one. However, the second step is different since it includes a binary conversion of the first step resulting array. This step will extract potential activities with exact start/end time from the first step resulting array. For example, when the inhabitant use his toilet, the real duration is desired. Once he stops doing the activity, the data before and after the activity must be excluded. In fact, mean values and standard deviations would be the best metrics to exclude them, but they are very affected by highly disparate values. Besides, since noise values are outside the box (greater than ρ_2), it drastically change the mean value of the whole data set. So, those metrics are not good enough in our case. Yet, in order to cut through those problems, a binary conversion is applied; if the mean value of a window is in the range $[\alpha, \beta]$, a 1 is added otherwise it is a 0. Then, first and last 0s are deleted so that only the full activity remains. Hence, the exact times the activity started and ended are obtained (since it must be more precise than a shower, which is much longer). Now, even if the outer noise is removed, inner noise could still be part of the data. For this reason, the mean λ and standard deviation σ_λ of the combined windows are computed. They are respectively compared to α and $\alpha \times 1.75$. This value has been found experimentally, by adjusting it to have the best results possible. If $\lambda < \alpha$ or if $\sigma_\lambda > \alpha \times 1.75$ (or if there is only one value left), the potential recognition is removed.

Those operations left us a series of recognized activities for the toilet. The merging process is the same explained for the previous algorithm, except for the duration which is 30 seconds instead of the 60 seconds of the shower case.

As for the previous algorithm, the Algorithm 2 summarizes it.

IV. Experiments

Our experiments have been approved by the Ethics committee of the UQAC (*Université du Québec à Chicoutimi*) with the file number *602.636.01*. In the next sections, participants and procedure are described.

A. Participants

For our experiments, 8 participants have been recruited, distributed as 5 males and 3 females aged from 22 to 29 years old. All of them were healthy adults without any motor or cerebral issues. In total, more than 5 million seconds of data have been recorded during this experiment (59 days of data for toilet and shower combined). Moreover, even if a majority of them is very familiar with technologies, some of them have no prior knowledge in this specific field, adding to the diversity of the group taking part into this set of experiments. Furthermore, participants were ask to give their permission to diffuse their data online: the resulting dataset is available on our GitHub repository linked in Sec. III.

B. Procedure

Once a participant has accepted to take part in our experiments, he had to meet us, in order to receive the instructions and the kit of modules (See Fig. 1a). Then, he had to take it home, install and calibrate every module by himself (an installation guide was provided for the participant to refer to in case the verbal explanations were not sufficient). The calibration had to be done by the participant since it was impractical for the team to send an installer for every settings. Additionally, every bathroom is different and dimensions can vary a lot. Next, the participants

Algorithm 2: Activity Recognition for the Oilet.

$\alpha \leftarrow max(20, \rho_1)$
$\beta \leftarrow \rho_2$

$array \leftarrow$ new detection array

for all do
  Wait for $T - T_\theta$ **new data**
  $window \leftarrow T_\theta$ previously obtained data + $(T - T_\theta)$ last collected data
  $M \leftarrow mean(window)$
  $cpt \leftarrow 0$
  $Ind_1, Ind_2 \leftarrow 0$
  if $\alpha < M < \beta$ **then**
    $Add \leftarrow true$
    $cpt \leftarrow 0$
    $Ind_1 \leftarrow$ index of $window$
 else
  if Add **then**
    $cpt \leftarrow cpt + 1$
  if Add and cpt=5 **then**
    $Ind_2 \leftarrow$ index of $window$

    Add items from $i = Ind_1$ to $i = Ind_2$ into $array$
    $Add \leftarrow false$
    $cpt \leftarrow 0$

 $data \leftarrow lastof(array)$
 $\lambda \leftarrow mean(data)$
 $\sigma_\lambda \leftarrow stdev(data)$
 if $\lambda < \alpha$ or $\sigma_\lambda \geqslant \alpha \times 1.75$ **then**
   delete *data* from *array*

 Combine every recognition in *array* with less than 30 seconds of differences

are instruted to not change their habits due to the influenced of the presence of the modules. Obviously, the module may slightly affect their behavior in the beginning, but unless it is something major, it should not affect the results significantly. This phase was planned for a duration of 10 days (± 2 days). During this phase, the two modules placed as shown in Fig. 1b and Fig. 1c are constantly recording their data with a timestamp. Besides, they also run the activity recognition algorithm in real-time (explained in details in Sec. III) based on the incoming data to recognize the ongoing activity. On top of that, it was asked to every participant to log whenever they start doing a listed activity and the time it lasted, so that the logs could be compared with the data recorded by our modules. At the end of the experiment, participants were asked to fill in two forms, respectively about the easiness of installation and the acceptability of such system. To finish with this experiment, they had to give back the whole system with their notes and forms.

TABLE I
CONFUSION MATRIX FOR THE TOILET MODULE

Expected \ Observed	Detected	Not Detected
Detected	40,508	3,161
Not Detected	866	2,694,045

V. RESULTS

In this section, data from the experiments described has been analyzed. The results were extracted from the main classification metrics used in the literature. Those metrics are presented in Sec. V-A. The results are then presented in the remainder of the section. Also, it is very important to note that we did not use a machine learning approach. Indeed, even if most works use such algorithms, it is very correlated to their sensor types (e.g. cameras). Basing our approach on a simple algorithm also avoids the need of a required training, resulting in a real plug-and-play solution.

A. Classification Performance Metrics

The problem presented is a classification task and whenever the performance of such a work needs to be evaluated, metrics must be representative. To this end, since the amount of data is large, basic ones such as simple accuracy care not sufficient to properly indicate the performance qualities of our system [24], [25]. It is especially important to use a metric that would not equivocating with the high number of True Negatives (*TN*) confronted to the much smaller number of True Positives (*TP*) combined with False Positives (*FP*) and False Negatives (*FN*). In our case, the *TN* represent 98.56 % of the combined results. Therefore, a classifier putting all the samples in the negative class would get a very high accuracy. In this respect and to ensure the quality of our work, the metrics chosen are Precision (*Precision*), Recall (*Recall*) and F-Score (*F*).

B. Results Obtained

The results are divided among the following five categories:

- Accuracy in activity detection
- Full test of sliding windows parameters for both shower and toilet
- Precision in recorded times for activities
- Easiness for the installation of the system
- Acceptability of such a system

1) Results About Activity Recognition: As it has been explained in Sec. IV-B above, participants in our experiment were given tables to write whenever they use their toilet/shower. The two confusion matrices about the toilet module and the shower one are given in the Tables I and III. Those values match the number of seconds for each category (*TN* means no-activity). Now that the confusion matrix are exposed, extraction of the classification performance metrics shown above is possible. Those metrics can be observed in the Table III below.

TABLE II
CONFUSION MATRIX FOR THE SHOWER MODULE

Expected \ Observed	**Detected**	**Not Detected**
Detected	$27,522$	772
Not Detected	0	$2,375,246$

TABLE III
CLASSIFICATION METRICS EXTRACTED

Metric	**Toilet**	**Shower**
Precision	0.9276	0.9727
Recall	0.9791	1.0000
F-Score	0.9526	0.9862

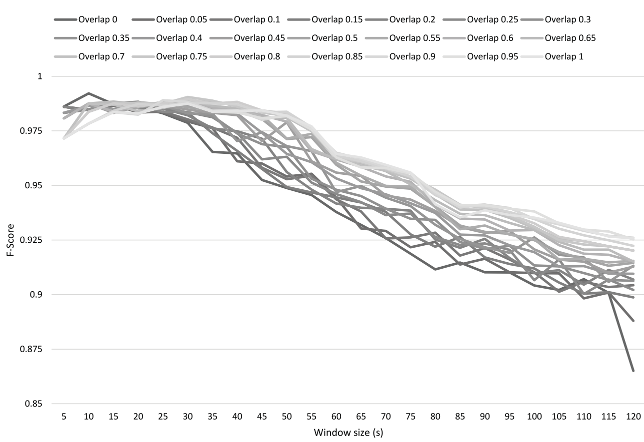

Fig. 5. F-Score for the toilet analysis in a single house.

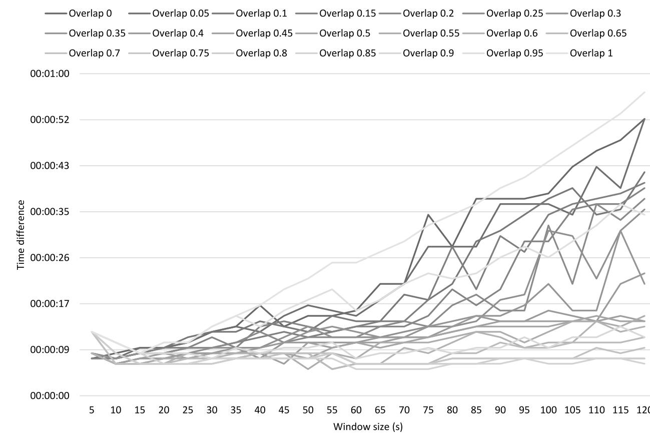

Fig. 6. Time differences for the toilet analysis in a single house.

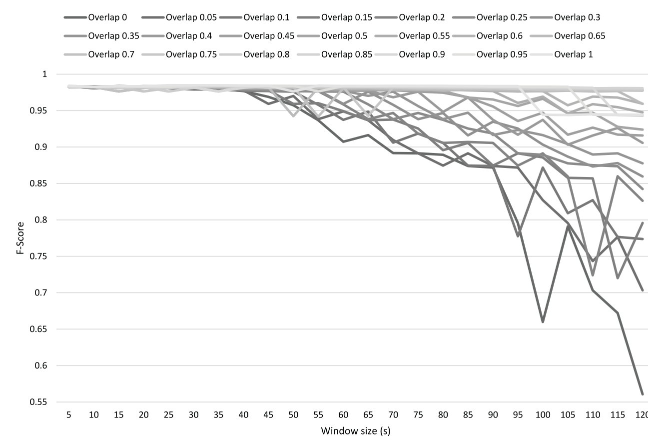

Fig. 7. F-Score for the shower analysis in a single house.

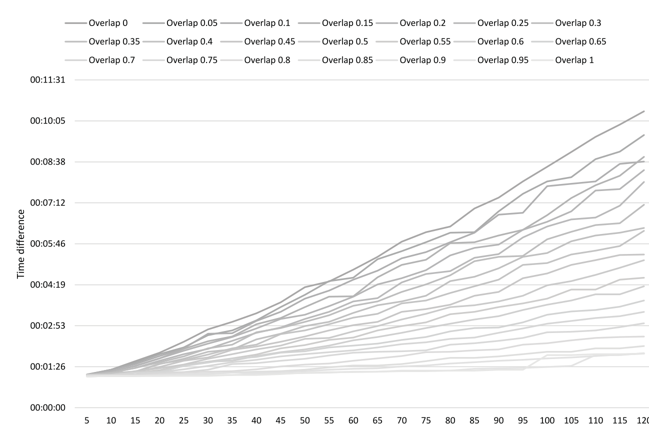

Fig. 8. Time differences for the shower analysis in a single house.

In our work, *Recall* is the most important metric since it involves activities that actually happened but has not been recognized. As observed in the Table III, the shower has a perfect *Recall* whereas the toilet has an excellent one of 97.91

2) Results About Best-Fitting Parameters: In a real-time case of activity recognition, the choice of window size and overlap is very important and could very well be determinant in the performance of any chosen algorithm. As time spent during a shower or a toilet-related activity is strongly dependant on the person observed, it is difficult to estimate with a satisfying precision. In order to try to have the best-fitting parameters, a systematic search of the window size/overlap was performed. To do so, the algorithms were ran on each of the houses of the experiment (3 houses). Actually, the window sizes varied from 5 seconds to 120 seconds (step of 5 seconds) and the overlaps from 0% to 100% (step of 5%). Also, since duration vary a lot between each house, it has been performed on each dataset individually. An example of those analysis for the toilet can be observed in Fig. 5 *(F-Score)* and in Fig. 6 *(time differences)*. The equivalent example for the shower is in Fig. 7 *(F-Score)* and in Fig. 8 *(time differences)*. For each graph, each curve represents a single overlap going from no-overlap (0%) in darker lines to full-overlap (100%) in light lines.

After the analysis of all graphs (several are missing in the paper, but are on the GitHub repository), it was determined that the best-fitting parameters are : a window size of 30 seconds with an overlapping time of 80% for the toilet and a window size of 20 seconds with an overlapping time of 75% for the shower.

3) Results About Activity Duration: In this paper, a great recognition rate is not enough to evaluate the good performance of our system. Indeed, we had to ensure that each activity recognized by our system matches the duration written by our

TABLE IV
TIME DIFFERENCES (IN SECONDS) AND ITS REPRESENTATION IN PERCENTAGE BETWEEN ACTIVITIES RECOGNIZED AND GROUND TRUTH GIVEN BY THE PARTICIPANTS USING PARAMETERS DEFINED IN V-B2

Activity	**House 1**		**House 2**		**House 3**	
Going to the toilet	6 ± 22	*3.90 %*	86 ± 261	*62.32 %*	20 ± 47	*22.22 %*
Showering	32 ± 58	*6.48 %*	70 ± 213	*13.51 %*	310 ± 595	*56.26 %*

participants on the observation sheet. Besides, the participants had to write (in the most precise way they could) the duration of their activities realized with the toilet or with the shower. Given that this creates our ground truth, we could easily compare recognized times with the written ones.

To confirm the system performance, the loss of precision was extracted (generally a few seconds) by using Equation 1, with x_A being the recorded time and x_B being the time written by the participant.

$$Time\_Diff = |x_A - x_B| \quad (1)$$

The time difference has been computed for each experiment (in seconds), for the toilet and the shower using the values presented in the Sec. V-B2. Every measured loss are given in the Table IV below. It can be easily observed that the loss of precision is mostly related to the ability of the participants to write times in a rigorous manner. Also, the Algorithms 1, 2 previously described had a big trouble for House 3 since both of the inhabitants from this environment took showers at less than a minute of interval, thus merging two recognized activities that should not be merged, causing the high gap.

Everything considered, the average difference represents 137 ± 288 seconds for the shower and 46 ± 110 seconds for the toilet.However, it is highly affected by two factors : the rigor taken by the participant to annotate the activities and the global window size/overlap chosen in V-B2. In fact, one of the authors took part in the experiment (House 1) to set a ground truth. His results are incredibly close to the real duration since it only differs of 6 seconds for the toilet and 32 seconds for the shower. Those results are very good considering that the duration written by the participant often includes the journey taken to go to the shower/toilet. Indeed, to improve those results, a stopwatch should has been provided to every participant, to automatically compute the activities' duration. The analysis performed leads to conclude that the system is sufficiently accurate.

4) Results About the Installation of the System: Our experiment has been a great opportunity to receive feedback about our system installation (since we want to create a very easy to install kit) and the acceptability of such a system. This second part will be described in the next section whereas the first one is overviewed here. As mentioned in the Sec. IV-B, we gave two forms to the participants to fill in. On those forms, four attributes were examined by the participant; they were asked to note each criteria in the range [1,5] following the rules below. The first criteria (*A*) represents their knowledge in technology with 5 being a technology expert and 1 a novice. The second one (*B*) evaluates the easiness of the installation (*B*) with 5 being a very easy installation and 1 being a complex one. The third criteria (*C*) represents the time taken to install the modules with 5 being a long time and 1 being very short. Lastly, the fourth one (*D*) is the complexity of calibration (*D*) with 5 being hard and 1 being easy. The results can be observed in the Table V. Also, participants indicated that the installation of our system in their house took them an average of 11 minutes.

TABLE V
SCORES ACCORDING TO INSTALLATION OF THE SYSTEM (V-B4) AND TO THE ACCEPTABILITY OF THE SYSTEM (V-B5)

About the system	**A**	**B**	**C**	**D**	**E**
Installation	4.2857	4.5714	2.0000	1.2857	*N/A*
Acceptability	2.0000	1.6250	1.7500	1.3750	1.8750

By analyzing these scores, we can easily conclude that our system is indeed very easy to install (*B, D*), and takes just a few minutes to do so (*C*).

5) Results About the Acceptability of Such a System: As explained in the previous section, the acceptability of our system was an important element to analyze since it is placed in the bathroom, where they feel the most vulnerable. This is the reason why we wanted to know how the participant felt about each module (toilet, shower) at the beginning of the experiment, and at its end. In the Table V, we can observe the marks given by the participants about their feeling relative to our modules. They were asked to note each criteria from 1 to 5, 1 being the lowest disturbance and 5 being the highest one. The first one (*A*) represents the disturbance of the user while using the toilet module at the beginning of the experiment by contrast to (*B*) which is the one at the end of the experiment. Criterias *C* and *D* are exactly the same question as *A* and *B*, but for the shower module. The last one (*E*) describes the number of participants who felt observed during the experiment. Our results clearly demonstrates that even if some people felt disturbed or observed at first, people get used to it very fast. In fact, participants who felt observed at first did not feel any more observed at the end: the sensation had dwindled with the passing time. Moreover, some of the participants found the global experimentation annoying because of the numerous sheets to file every time they needed to use their bathroom, statistics that could influence the data from *A* and *B*.

VI. DISCUSSION

In this paper, our teams conducted experiments and analyzed results for three characteristics related to the software of the developed modules (recognition, best-fitting parameters and time precision) and two related to their hardware/appreciation (installation, acceptability). In the Sec. V-B2, the best window sizes and overlapping times have been computed, resulting in 30 seconds and 80% for the toilet and 20 seconds and 75% for the shower. With those settings, it can be concluded that the recognition rate is excellent (96.94% of F-Score) for more than 8 participants over 30 days, suggesting that our system is highly accurate in real-world usage context. However, the system must not only recognizes correctly, but also takes the duration of an activity as precisely as possible. Also, times recorded by the modules were compared to times written by participants: they are really substantially identical with a mean difference of just a few seconds, representing approximately 6 seconds of difference for the toilet and 34 seconds for the shower, in the house of the most rigorous participant. Since it must take into account that the participants had not a proper chronometer, the duration written in the given sheets were approximate.Also, it has been seen in our dataset that each results could be improved by personalizing window size and overlapping time per house. Indeed, since people have different habits regarding their bathroom-related activities, the time according to it is drastically different.

Moreover, an interesting thing to be concerned about is the reaction of the participants about the installation of such a system and how they felt about it. On the forms provided to the participants, they estimated that our system was very easy to install ($4.57/5$) even if they were not all comfortable with technology (25% were not accustomed to this kind of mechanics). In addition to this, the acceptability of our system were evaluated by our participants in terms of disturbance at start and at the end of the experiment. Indeed, even if 37.5% felt disturbed at start, every affected participant get familiarized with it and did not feel any more troubled at the end of the experiment. A good inference about it could be that those sensors are not conventional in the everyday-life of people, but they get used to it very fast.

Furthermore, since we want to make this research domain more accessible, we deliver every schematics and datasets on our GitHub Repository previously linked in Sec. III. The reader may also be interested into learning about our first version of this system which lead to failing experiments. In the same vision as previously explained, this previous failure is explained in this paper to ensure anyone that would want to reproduce our system would avoid reproducing those mistakes. First, the Raspberry Pi hardware is very good, but the SD cards provided with the development kits are now of a cheaper brand than before. As this kind of storage is very sensitive to the number of write/read interactions, and most of all, to a hard-reboot, those cards were not reliable. Also, a subsequent trouble happened to us; one of the house had a corrupted database because of the faulty storage. So, we strongly recommend to use high quality SD cards with a well-known manufacturer to avoid any of those troubles; this was our solution and it worked perfectly. Secondly, another concern that have been fed back to us was the distance between the power source and the system in the bathroom. Indeed, we used the USB cables furnished with our kits, which were short. Actually, we would encourage people to use longer cables in order to completely power the modules more than 5 meters away of the shower.

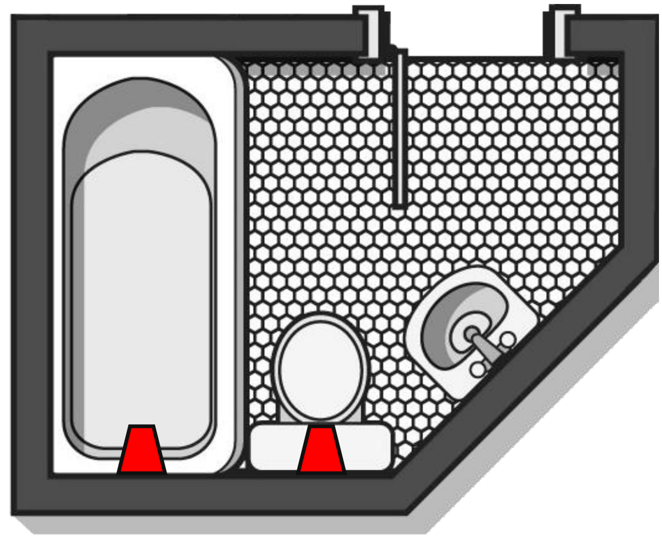

Fig. 9. Schematics of a typical bathroom (red boxes being our modules, not representative size.).

On top of that, participants of our experimentation gave us feedback to improve the system. They were actually thinking that the system would be calibrating automatically, like most measuring system. In fact, the module placed in the shower could be easily calibrated without the need of the user to press "Calibrate," since there is only one phase to it. However, the one installed on the toilet having multiple classes to classify, its calibration is harder. Yet, we discovered, by looking at the raw data of our experiments, that every recognition was done by persons using the toilet by sitting on it. The recognition of people standing in front of it (possibly a man urinating) was affected by the passage of the inhabitants in front of the sensors. In most cases, this passage was to access the shower in a typical bathroom schematized in Fig. 9: the toilet has to be triggered before accessing the shower. This specific situation causes the algorithm to fail a lot, which is why we decided to ignore the recognition dealing with the standing position. In addition to our decision, most of the people targeted by our experiment are more affected by actions requiring standing and sitting transitions than by just standing. Accordingly to our reasoning and the ground facts, we encourage people to only monitor activities requiring sitting on the toilet.

VII. Conclusion

In this work, we proposed a first system built to host many sensors in a smart-home context. This approach introduced a unique and very easy to install system, composed of two modules placed in the bathroom of lambda residents, in order to recognize and quantify basic activities such as going to the toilet and showering. The technologies used in it are different from the commonly used sensors (Sec. II) and have resulted in outstanding performance. This allows people to monitor their bathroom-related activities, in terms of duration and quantity. Besides, those activities have been highly correlated to the health status of a person [13], [14]. Also, since our system induce excellent results, it should be a good assumption to say that it could be used to monitor behavioral changes related to those activities, in short-term, mid-term and long-term experiments. Moreover, every people who participated in our experiment found it very easy to install and to live with, making it very compliant with daily life. This paper aimed to develop a new system to continuously observe those activities, which is an achieved goal. Now, as everything we did on this project is public via our GitHub Repository previously given[2] and as there is multiple existing studies relying on those activities to monitor health status [26], our technology could be very useful in a real-world context. Also, researchers could get a simulation of a smart bathroom by using our kit; easing the access to smart home context to the research community.

VII. Future Works

As the proposed system is built to host more sensors than presented ones, it would be great to add more in it. Besides, as authors already worked on a smart wristband being able to recognize daily activities [27], it would be very interesting to couple those two researches: indeed, being able to correlate activities recognized by the toilet (being sit) with activities recognized by the wristband (sitting down, standing up) would be a great asset to ensure that the user really did detected activities. Also, as mentioned in Sec. VI, the calibration phase could be improved by avoiding user interaction. In fact, it should be possible to take a full day to calibrate itself. Finally, this work has been tested on young population since it is a proof of concept. However, it is designed to be used by either an elderly population or a medium-aged population suffering from heavy diseases (needing more supervision).

Acknowledgment

We would like to thank our main financial sponsors: the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Foundation for Innovation (CFI). Moreover, authors would like to acknowledge every participant in our experiment.

References

[1] D. Population Division, "World population ageing," *United Nations*, pp. 1–124, 2017.

[2] Z. Milanovic, S. Pantelic, N. Trajkovic, G. Sporis, R. Kostic, and N. James, "Age-related decrease in physical activity and functional fitness among elderly men and women," *Clin. Intervent. Aging*, vol. 8, pp. 549–556, 2013.

[3] E. Perry *et al.*, "Nicotinic receptor subtypes in human brain ageing, Alzheimer and Lewy body diseases.," *Eur. j. Pharmacol.*, vol. 393, no. 1–3, pp. 215–22, 2000.

[4] S. A. Paluska and T. L. Schwenk, "Physical activity and mental health," *Sports Med.*, vol. 29, pp. 167–180, 2000.

[5] M. Nakhodaeezadeh, M. A. Jafarabadi, H. Allahverdipour, H. Matlabi, and F. R. Dehkordi, "Home environment relation with quality of life older people," *J. Housing Elderly*, vol. 31, no. 3, pp. 272–285, 2017.

[6] E. C. Nelson, T. Verhagen, and M. L. Noordzij, "Health empowerment through activity trackers: An empirical smart wristband study," *Comput. Human Behav.*, vol. 62, pp. 364–374, 2016.

[7] J. B. Bussmann, W. L. Martens, J. H. Tulen, F. C. Schasfoort, H. J. van den Berg-Emons, and H. J. Stam, "Measuring daily behavior using ambulatory accelerometry: The activity monitor.," *Behav. Res. Methods, Instrum., Comput : A J Psychonomic Soc., Inc*, vol. 33, no. 3, pp. 349–356, 2001.

[8] L. Walsh, B. Greene, A. Burns, and C. Ní Scanaill, "Ambient assessment of daily activity and gait velocity," in *Proc. 5th Int. Conf. Pervasive Comput. Technol Healthcare (PervasiveHealth) Workshops*, 2012, pp. 418–425.

[9] S. Giroux *et al.*, "Cognitive assistance to meal preparation: Design, implementation, and assessment in a living lab," in *Proc. AAAI Spring Symp. Ser.*, 2015, pp. 01–25.

[10] B. Bouchard, K. Bouchard, S. Gaboury, and A. Bouzouane, "Smart senior residences: how to improve services efficiency and quality by exploiting big data and ambient intelligence," *Big Data Manage. Internet Things Improved Health Syst.*, pp. 65–85, 2018.

[11] R. Li, B. Lu, and K. D. McDonald-Maier, "Cognitive assisted living ambient system: A survey," *Digit. Commun. Netw.*, vol. 1, no. 4, pp. 229–252, 2015.

[12] G. Sprint, D. J. Cook, R. Fritz, and M. Schmitter-Edgecombe, "Using smart homes to detect and analyze health events," *Computer*, vol. 49, no. 11, pp. 29–37, 2016.

[13] A. M. Fonseca and E. Soares, "The care taker's discourse about taking care of the elderly with alzheimer's disease," *Northeast Netw. Nursing J.*, vol. 9, no. 3, 2008.

[14] M. Folstein, J. C. Anthony, I. Parhad, B. Duffy, and E. M. Gruenberg, "The meaning of cognitive impairment in the elderly," *J. Amer. Geriatrics Soc.*, vol. 33, no. 4, pp. 228–235, 1985.

[15] B. M. Van Gelder, M. A. Tijhuis, S. Kalmijn, S. Giampaoli, A. Nissinen, and D. Kromhout, "Physical activity in relation to cognitive decline in elderly men: The fine study," *Neurology*, vol. 63, no. 12, pp. 2316–2321, 2004.

[16] K. Bouchard, B. Bouchard, and A. Bouzouane, "Smart homes: Practical guidelines," Opportunistic Networking: Vehicular, D2D and Cognitive Radio Networks, CRC press, Taylor & Francis, pp. 206–234, Apr. 2017.

[17] J. A. Kientz, S. N. Patel, B. Jones, E. Price, E. D. Mynatt, and G. D. Abowd, "The georgia tech aware home," in *Proc. 26th Annu. CHI Conf. Extended Abstr. Human Factors Comput. Syst.*, 2008, p. 3675.

[18] D. J. Cook, A. S. Crandall, B. L. Thomas, and N. C. Krishnan, "CASAS: A smart home in a box," vol. 100, no. 2, pp. 130–134, 2012.

[19] P. Lago *et al.*, "The contextact @ A4H real-life dataset of daily-living activities activity recognition using model checking," 2017.

[20] J. Fogarty, C. Au, and S. E. Hudson, "Sensing from the basement: A feasibility study of unobtrusive and low-cost home activity recognition," in *Proc. 19th Annu. ACM Symp. User Interface Softw. Technol.*, 2006, pp. 91–100.

[21] J. Chen, J. Zhang, A. H. Kam, and L. Shue, "An automatic acoustic bathroom monitoring system," in *Proc. - IEEE Int. Symp. Circuits Syst.*, 2005, pp. 1750–1753.

[22] C. Debes, A. Merentitis, S. Sukhanov, M. Niessen, N. Frangiadakis, and A. Bauer, "Monitoring activities of daily living in smart homes: Understanding human behavior," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 81–94, Mar. 2016.

[23] E. M. Tapia, S. S. Intille, and K. Larson, "Activity recognition in the home using simple and ubiquitous sensors," in *Pervasive Comput.* (A. Ferscha and F. Mattern, eds.), (Berlin, Heidelberg), Springer Berlin Heidelberg, pp. 158–175, 2004.

[24] A. Ben-David, "A lot of randomness is hiding in accuracy," *Eng. Appl. Artif. Intell.*, vol. 20, no. 7, pp. 875–885, 2007.

[25] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond accuracy, F-Score and ROC: A family of discriminant measures for performance evaluation," *Adv. Artif. Intell.*, vol. 4304, pp. 1015–1021, 2006.

[26] M. D. Leon, K. B. Rajan, and D. A. Evans, "Decline in cognitive function and risk of elder self-neglect," vol. 58, no. 12, pp. 2292–2299, 2011.

[27] K. Chapron, V. Plantevin, F. Thullier, K. Bouchard, E. Duchesne, and S. Gaboury, "A more efficient transportable and scalable system for real-time activities and exercises recognition," *Sensors*, vol. 18, no. 1, 2018, Art. no. 268.