Radar for assisted living in the context of Internet of Things for Health and beyond

Julien LeKernec, Francesco Fioranelli, Shufan Yang, Jordane Lorandel and Olivier Romain

August 29, 2018
Radar for assisted living in the context of Internet of Things for Health and beyond

Julien Le Kernec\textsuperscript{1,2}, \textsuperscript{*} Senior Member, IEEE, Francesco Fioranelli\textsuperscript{1}, Member, IEEE, Shufan Yang\textsuperscript{1}, Member, IEEE, Jordane Lorandel\textsuperscript{1}, Member, IEEE, Olivier Romain\textsuperscript{3}, Member, IEEE

\textsuperscript{1}Communication, Sensing and Imaging group, School of Engineering, University of Glasgow, Glasgow, UK
\textsuperscript{2}School of Information and Electronics, Univ. of Electronic Science and Technology of China, Chengdu, China
\textsuperscript{3}ETIS-ASTRE, Université Cergy-Pontoise, Cergy-Pontoise, France

Abstract—This paper discusses the place of radar for assisted living in the context of IoT for Health and beyond. First, the context of assisted living and the urgency to address the problem is described. The second part gives a literature review of existing sensing modalities for assisted living and explains why radar is an upcoming preferred modality to address this issue. The third section presents developments in machine learning that helps improve performances in classification especially with deep learning with a reflection on lessons learned from it. The fourth section introduces recent published work from our research group in the area that shows promise with multimodal sensor fusion for classification and long short-term memory applied to early stages in the radar signal processing chain. Finally, we conclude with open challenges still to be addressed in the area and open to future research directions in animal welfare.

Index Terms—Human activity classification, fall detection, ambient assisted living, inertial sensors, magnetic sensors, radar sensors, multisensory data fusion, feature selection, machine learning, micro-Doppler signatures, feature extraction.

I. INTRODUCTION

Internet of Things (IoT) in healthcare was evaluated at $60 billion and will reach $136 billion by 2021 [1]. IoT comprises intermediary components, such as devices, network connectivity, electronic systems, and software. It is networked smart electronic devices sharing information autonomously leveraging machine learning. In healthcare, this technology will facilitate managing and mining patient data and resources. Life expectancy is increasing and poses challenges for health services as it comes with medical issues (chronic illnesses, multi-morbidity) and an alarming rise in the population over 60 predicted to reach 30\% by 2050 worldwide [1-2]. This trend is not new but accelerating especially in developed countries.

In 2016, signal processing magazine had a special issue on assisted living [3–8]. It covered a range of technologies such as inertial measurement units, wearables, ambient sensors (pyroelectric infrared (PIR), vibration sensors, accelerometers, cameras, depth sensors and microphones) and radio waves with existing infrastructure (Wifi) present on site or active devices such as radar. For all sensing modalities, enhancing accuracy, lowering computational complexity, reducing power consumption, exploiting multiple domains and modalities for complementarity and robustness, are crucial in developing technology enabled self-dependent living in-home care.

II. EXISTING SENSING MODALITIES

Many systems have been proposed to tackle this problem [5,8-9] including radar sensors or a combination of these systems, whereby their information is used concurrently and fused at different levels to optimize the overall performance.

Monitoring people in their daily life poses a privacy issue; there is a correlation between the perceived privacy and richness of information collected by sensors [9]. Video provides very rich information but is perceived as intrusive; PIR sensors are not perceived as invasive but provide little information.

A review of healthcare using mobile wireless technologies shows major challenges (data acquisition, processing data locally, wireless data, quality of service over cellular network, cloud storage, security, user interface and platforms) before being feasible [11]. It also suffers from integration problems where a lot has to come together before it is practical to use and requires the lifting of technological barriers as well.

Wearable sensors despite giving good classification results [12] greater than 98\%, suffer from several major problems [13]:

\begin{itemize}
  \item require user compliance as they need to be worn or to think about it if you wake during the night to go to the washroom.
  \item easily broken if dropped, crushed while sitting or falling.
\end{itemize}

In [14], [15], entire apartments have been fitted with sensors PIR motion sensors, stove sensors, floor sensors,… and provide good density maps for activities of daily living at the macro level. However, they cannot provide a finer granularity for gait analysis change detection as well as requiring transformations in a persons living environment.

An extensive review [16] of RGB cameras, depth sensors and radar technologies for assisted living highlighting open challenges for deployment in residences or specialized homes:

\begin{itemize}
  \item For cameras, the main challenges are occlusions, working at night, dead zones in 3D, accuracy, precision, resolution and respecting privacy.
\end{itemize}
For radar systems, the presence of strong scatterers and clutter in indoor environments may generate multipath and ghost targets which is comparable to occlusion in cameras. The compliance of radar system with emission regulations limits.

The technological challenges are greater for radar technology, but the fact there are no judicial issues regarding rights to image and plain images are not recorded, thus respecting privacy, facilitating acceptance of end users and investors. For these reasons, the radar sensing modality is an interesting research trend however still underutilized in specialized homes.

Radar is attractive due to reliability, low power emissions for indoor use (similar to WiFi), safety, which brings it at the frontier of indoor monitoring modalities rivaling video cameras and wearable devices for health. Radar can be used for fall detection, gait analysis and activities of daily living (ADL) to provide supplemental information to detect early signs of deteriorating physical/cognitive health. It would allow greater healthcare coverage, better quality of provision through 24/7 monitoring of the elderly well-being while respecting privacy.

Furthermore, the elderly may suffer from reduced cognitive capabilities and memory loss. To enable assistive technologies to help them to deal with ADL and monitoring their condition, a system requiring no intervention from their part is more suited. Existing radar systems can be used to monitor activities [12,16–20], but it could create a paradigm shift in health monitoring and moving from reactive technologies to preventive. If they are made smart enough to learn the daily activity pattern of an end user, and identify deviations/anomalies linked to declining health, they could foresee the occurrence of possible critical events (e.g. falls, strokes).

Radar will enable prompt emergency responses following critical events (reactive), continuous in-home health monitoring for medical professionals to improve diagnostics and develop precision medicine for individuals (predictive). It would also enable persuasive feedback to individuals to advise/influence behaviours for safer and better practice, when variations in their routine are identified (prevention & assistance).

III. MACHINE LEARNING PERSPECTIVE

Machine learning is becoming an integral part of technology development given the advantages it provides, radar system applications are also leveraging machine learning for enhanced performances and accuracy in activity classification.

Generally, to classify activities, radar micro-Doppler (mD) signatures are used as a base. The relative motion of limbs and head with respect to the torso generates unique signatures in the time-frequency domain of the radar returns. Different activities create uniquely identifiable features in mD signatures used for classification. A comprehensive coverage of the subject can be found in [21–22]. Spectrograms are then processed to extract features [23] followed by different classifiers [24–25].

Here is a non-exhaustive list of machine learning techniques for classification: Fisher Discriminant Analysis [26–27], K-nearest neighbors [28–29], Naive Bayes [30], Ensembles (e.g. Bagging [31]) and Support Vector Machine (SVM) [32].

A review classifiers for activity classification [16] advise to use multiple sensors to enhance classification accuracy by covering multiple aspect angles and combat occlusions. Another way to improve accuracy is to fuse data and select the most salient features [12,18,33–34]. Many classifiers are used in activity classification of which SVM is the most common [35]. The choice of classifier is important, but choosing the most salient features has a greater impact on accuracy than the classifier [36]. There is a wealth of contributions trying to extract features and classify activities from mD signatures [16].

Beyond machine learning lies deep learning thanks to advances in computational power (GPUs). Feature extraction is an expert-knowledge based task. Deep learning techniques however can figure out relevant features for classification, sparse representations and time-dependencies through several layers of neurons with activation functions e.g. recognize faces with convolutional neural networks [37]. Another class of deep learning algorithms used for speech recognition are Recurrent Neural Networks (RNN) with Gated Recurrent Units [38], [39] and Long Short-Term Memory (LSTM) [40].

A general belief is that deep learning requires “Big Data” to be effective; but small datasets also produce good results [41–42] via data augmentation and transfer learning.

Figure 1 summarizes the research on activity classification using deep learning for enhanced accuracy [43–57] yielding precisions from 80 to almost 100%.

It is hard to assess the different performances since all the deep learning algorithms are ad hoc and the size and the nature of datasets vary. Because the intra class variance for similar activities is smaller than for different ones, therefore advertised accuracy varies in meaning. Deep learning is already showing better performance than expertly pre-trained models [58].

The problem of multiple people in the field of view [47, 54, 56, 59] is rarely studied as they mostly consider only one person. In [16, 19, 24, 25, 60–63], the multi-static radar approach is utilized for classification from spectrograms using feature extraction. The difficulty in research with multi-static radar is the synchronization requirement between radar units and they are not commercially available. Aspect angle dependence in classification is rarely discussed although it has a large effect on accuracy [36, 53, 60]; most studies adopt actions happening in the radial direction of the radar.

Generally, the activities are looked at different activity snapshots and not in a continuum like in [51] for wearables.
The lessons from the literature are that CNN can recognize elaborate features from signals/images for particular snapshots at a given time where RNN of which LSTM is the leading technique takes into consideration time dependencies between snapshots. [51] shows a combination of CNN and LSTM for wearables and [49] presents a multimodal CNN multi-stream in parallel with LSTM with fusion; showing new ways to think about classifying data as a continuum.

Great efforts should go on preparing datasets, neural network architectures, training/optimizing to avoid overfitting, bias and ensure the model generalizes the activities to recognize unseen data or people accurately even with small datasets.

IV. Recent Results from the Communications, Sensing and Imaging Group at University of Glasgow

Now we have gone around the state of the art and the context, it is time to present some results from our recent studies.

A. Multisensor approach for remote health monitoring of older people [12], [18]

The experimental setup (Figure 2) shows the placement of the radar and the wearable sensor. The activities (walk, walk while carrying an object, sit down, stand up, pick up an object, crouch to tie shoe laces, drink, answer the phone, frontal fall, check under a piece of furniture) were measured with 9 volunteers giving 270 samples in total. 177 features were extracted using the inertial sensor in time and frequency domains and 28 features from spectrograms using radar.

The classifiers were quadratic kernel SVM and 10-NN trained using 10-folds randomly (9 for training and 1 for testing). The results in Table I are the result of the average of 10 folds. Notice, radar underperforms compared to wearables.

Table I: Classification accuracy using a single sensor [12]

<table>
<thead>
<tr>
<th>Classification Accuracy (%)</th>
<th>SVM</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accelerometer</td>
<td>85.2</td>
<td>79.6</td>
</tr>
<tr>
<td>Gyroscope</td>
<td>84.1</td>
<td>79.6</td>
</tr>
<tr>
<td>Magnetometer</td>
<td>80.4</td>
<td>69.6</td>
</tr>
<tr>
<td>Inertial</td>
<td>89.3</td>
<td>85.2</td>
</tr>
<tr>
<td>Radar</td>
<td>77.9</td>
<td>70.7</td>
</tr>
</tbody>
</table>

Feature selection (Fscore [64], ReliefF [65], SFS [64]) on single modality significantly increases classification accuracy.

Table II: Improvements with feature selection methods for IMU and radar in terms of accuracy and number of features

<table>
<thead>
<tr>
<th>Method</th>
<th>IMU  Accuracy(%)</th>
<th>IMU  Features #</th>
<th>Radar  Accuracy(%)</th>
<th>Radar  Features #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fscore SVM</td>
<td>90.7</td>
<td>73</td>
<td>78.8</td>
<td>17</td>
</tr>
<tr>
<td>Fscore KNN</td>
<td>88.2</td>
<td>76</td>
<td>74.1</td>
<td>17</td>
</tr>
<tr>
<td>ReliefF SVM</td>
<td>91.1</td>
<td>164</td>
<td>74</td>
<td>20</td>
</tr>
</tbody>
</table>

Table II shows 5-9% improvement for both sensors. So “less is more”: having more features does not improve accuracy but using/identifying the salient features does.

To increase accuracy, it is interesting to explore the benefit of using multimodal fusion at various levels (signal, feature and decision) [66–68] (Table III).

Table III: classification accuracy improvement with fusion

<table>
<thead>
<tr>
<th>Fusion method</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature level</td>
<td>97.4</td>
</tr>
<tr>
<td>Decision level logP [69]</td>
<td>96.7</td>
</tr>
<tr>
<td>Decision level fuzzy logic [66]</td>
<td>94.8</td>
</tr>
<tr>
<td>Decision level voting [12]</td>
<td>97.8</td>
</tr>
</tbody>
</table>

Figure 3: sensitivity and specificity for the fall action using various classifications methods [12]

Figure 3 shows an improvement overall without affecting fall specificity (reaching 100% with voting) by applying suitable feature selection and fusion.

B. Activity Classification Using Raw Range and I & Q Radar Data With Long Short Term Memory Layers [70]

LSTM are used to classify directly from raw data and range maps for binary classification every 0.5 s of action recorded. 5 subjects contributed, actions were recorded continuously for 60s giving 19 recordings (10 ‘walk’ and 9 ‘sitting & standing’).

For both, 2,280 samples were obtained by dividing the recording in 0.5s snapshots and the data is presented to the LSTM as described in Table IV and Figure 4.

Table IV: LSTM data parameters for classification

<table>
<thead>
<tr>
<th>LSTM data</th>
<th>Number of time samples</th>
<th>Number of features</th>
</tr>
</thead>
<tbody>
<tr>
<td>I&amp;Q</td>
<td>64000</td>
<td>2</td>
</tr>
<tr>
<td>Range profiles</td>
<td>500</td>
<td>35</td>
</tr>
</tbody>
</table>

Figure IV: left) Range profiles for walking (60s) and sitting & standing movements (60s). right) illustrate I & Q data for walking movement. The time patterns that exist in the signals are exploited through the LSTM layers. [70]
The samples were shuffled in a stratified manner (80% for training and 20% for testing) under the 5-fold scheme. Table V shows that I&Q data yields better accuracy than range profiles and LSTM are able to process backscattered data as time series. The drawback is the time for an estimation is 2s for raw data and 2ms for range profiles for 0.5s of a data continuum.

### REFERENCES


