

Multimodal monitoring of Parkinson's and Alzheimer's patients using the ICT4LIFE platform

Federico Alvarez¹, Mirela Popa², Nicholas Vretos³, Alberto Belmonte-Hernández¹, Stelios Asteriadis², Vassilis Solachidis³, Triana Mariscal¹, Dario Dotti², Petros Daras³.

1 Universidad Politécnica de Madrid, 2 Universiteit Maastricht, 3 Centre for Research and Technology Hellas

fag@gatv.ssr.upm.es, mirela.popa@maastrichtuniversity.nl, vretos@iti.gr, abh@gatv.ssr.upm.es, stelios.asteriadis@maastrichtuniversity.nl, vsol@iti.gr, tmc@gatv.ssr.upm.es, dario.dotti@maastrichtuniversity.nl, daras@iti.gr

Abstract

The analysis of multimodal data collected by innovative imaging sensors, Internet of Things (IoT) devices and user interactions, can provide smart and automatic distant monitoring of patients and reveal valuable insights for early detection and/or prevention of events related to their health situation.

In this paper, we present a platform called ICT4LIFE which starting from low-level data capturing and performing multimodal fusion to extract relevant features, can perform high-level reasoning to provide relevant data on monitoring and evolution of the patient, and trigger proper actions for improving the quality of life of the patient.

1. Introduction

Almost 10 million Europeans live with Parkinson, Alzheimer and other dementia related diseases today. As a result of ageing, the number of people affected by one of these conditions is forecasted to double by 2030 [8]. Coping with these diseases in society results in major health and financial challenges. Parkinson's and Alzheimer's patients want to live in their own homes for as long as possible. Nevertheless, the symptoms they face make their daily life hard to both managing their own care and living independently.

As the population ages, it is important to develop unobtrusive, affordable, computationally inexpensive, off-the-shelf smart home technologies for use in elder care, particularly for elderly people experiencing health problems.

To reach this goal, ICT4Life platform deliver a series of innovative services to patients affected by Parkinson's, Alzheimer's and other dementia related diseases and also to health professionals and formal and informal caregivers. All solutions are developed following a user-centered methodology and tested in real life scenarios including patients' home, healthcare centres and rehabilitation ones.

The proposed technology is focused on the analysis of multimodal data from visual sensors such as RGB-D or 360° cameras, combined with wearable devices, mobile,

binary sensors and the user digital interaction with the system. All these informative cues can provide valuable data to both, early detect and prevent patient events related to their health situation. Moreover, this information can provide a smart and automatic patient monitoring useful for caregivers or physicians.

The remainder of the paper is organized as follows: in Section 2 a summary of the relevant prior art is outlined. In Section 3 we describe the system's overview including the sensing technologies that are used. Section 4 presents one of the main modules of the solution, the abnormal behaviour analysis including testing and results obtained in real-life scenarios. Finally, Section 5 concludes the paper.

2. Related work

Advances in everyday technologies, such as smartphones and wearable devices have allowed for new information acquisition and thus for higher semantic processing capabilities as well as the incorporation of biomedical sensors for general purposes. As a consequence, a wide perspective for ad-hoc & full time health monitoring applications using ICT has been created [1].

Numerous works in vision-based monitoring of indoor human behaviour have been proposed [2][3], using different types of modalities (i.e., cameras, inertial sensors, RGB-D sensors and others) and system architectures. The joint analysis of the sensory inputs leads to an improved understanding of the context (environment), enhancing the understanding of an event through different informative cues.

One common type of human behaviour modelling applications is anomalies detection, and the main challenge is the unclear definition of abnormalities, as they are context dependent and can appear very rarely in the training set of a machine learning system. There are several attempts in the literature, especially in the context of smart homes [3] or elderly assisted living [4] trying to cope with these issues.

One approach towards learning behaviour patterns is based on trajectory analysis and has been investigated in [5]. In this paper the authors aim at capturing the overall

spatial arrangement of local motion displacement vectors, as well as in [6] where the authors propose the use of hierarchical clustering, as a proposition to the analysis of motion trajectories.

Our work is producing comparable results meanwhile offering a more general and adaptable framework which is able to cover a wider range of user patterns and behavioural analysis.

3. System overview

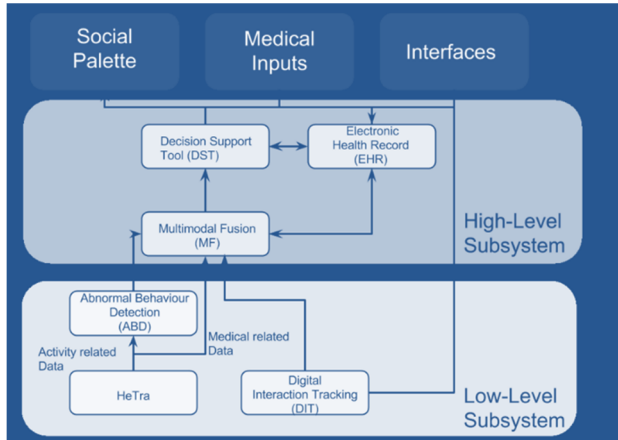


Figure 1: ICT4Life Platform Architecture and Main modules

The proposed ICT4Life system aims to create a platform able to provide a wide range of services to Parkinson's and Alzheimer's patients, caregivers and health professionals in order to assist them in the continuous care that all individuals need when dealing with chronic diseases.

Additionally, this platform should fulfil requirements in terms of flexibility, modularity, scalability, usability and accessibility.

In order to guarantee these requirements, a three-layer architecture has been developed as described in Figure 1. This architecture is mainly composed of a Services Subsystem, a High-Level Subsystem and a Low-Level Subsystem. Within the High-level subsystem, there are functionalities that ingest data from the Low-Level subsystem modules, manual inputs and digital events, in order to fuse this information with health-related variables (historical data, activity analysis, health records and input from professionals), to support decision-making on patient management and build a dynamic repository. Finally, within the Low-level subsystem, there are core functionalities that retrieve information from sensors to be sent to the High-level subsystem and end-user services.

In the top part of the figure we can find different elements which are also providing input to the rest of the modules, which are not always related to automated monitoring of the patient. Interfaces are related to user interfaces where it is possible to provide additional medical information from the users or caregivers. Medical inputs are information which is providing relevant data to the Electronic Health Record and supporting the rules in the Decision Support Tool. Social Palette is a platform allowing communication among patients, where information can be extracted.

3.1. Sensing technologies

In order to perform the patients monitoring, different sensing technologies are used (providing visual, motion and depth information) and their outputs analysed and

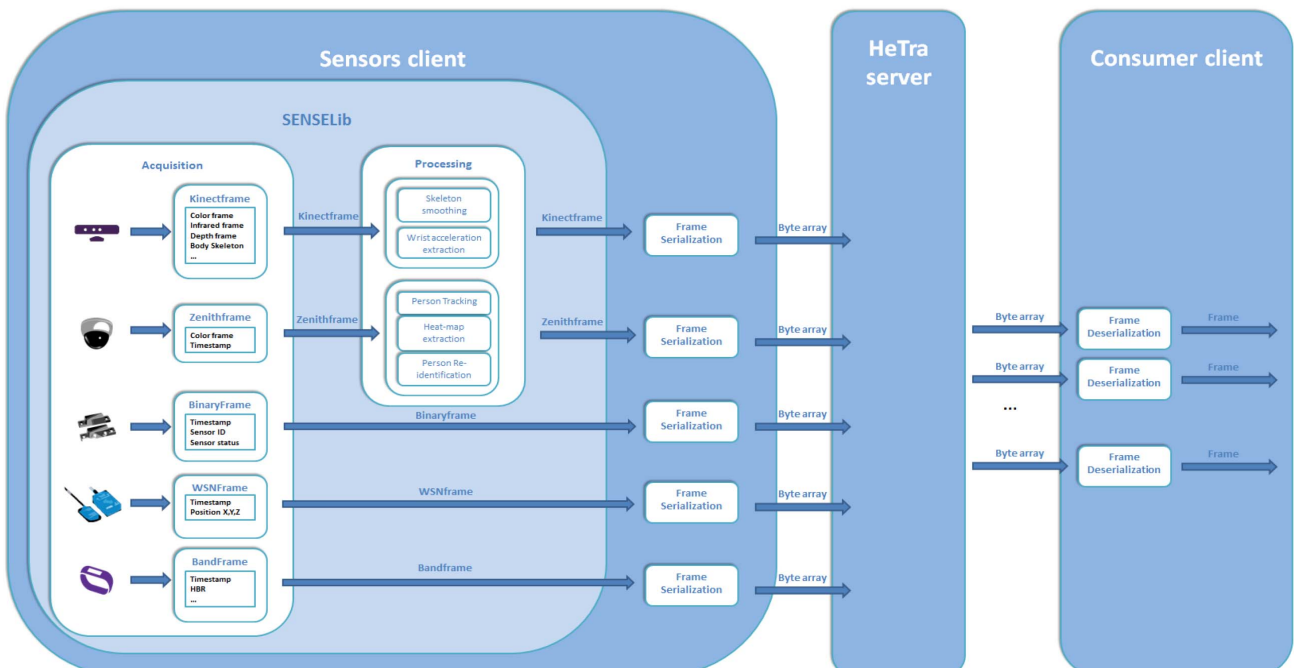


Figure 2: HeTra subsystem architecture overview

combined to produce relevant data which can provide information on the patient health status. The different devices have been chosen according to the scenarios which were selected for the patient monitoring: a daily centre (where patients meet others to socialise and perform leisure activities), rehabilitation centre (where patients do guided rehabilitation exercises) and home environment (patients home).

- Multi-sensory band, this device is the most important of the system and it is present in all the scenarios. It can provide useful data regarding the health of the patient such as blood pressure, skin temperature and other health related information. In addition, it also allows the acquisition of motion data including sensing technologies such as accelerometer, gyroscope or magnetometer.

- Binary sensor, this technology is intended to provide information regarding the behaviour of the patient, while being inside his/her house. It is related to different functionalities, such as detecting when the patient is leaving the house, or in which room the patient is located at a specific moment.

- RGB-D sensor (Kinect) adds new sources of data in the infrared domain that allows the extraction of the skeleton of a patient. Information expressed in these skeletons allow a deep motion analysis to monitor evolution of the patients during their rehabilitation exercises.

- Zenith camera, this kind of camera provides a panoramic view of the scene (360°) allowing monitoring big indoor areas such as healthcare daily centres. Moreover, due to their placement, results of human tracking algorithms are improved since occlusions are less present with this kind of devices.

- Wireless Sensor Network (WSN), this component consists of different anchors placed normally at the roof monitoring the radio signal from patient’s wearables for patient indoor location and tracking, in a non-intrusive manner.

3.2. Data Collection

An acquisition system named HeTra (Health Tracking) has been implemented in order to gather data from all mentioned devices. HeTra consists of a server and two types of clients, the HeTra sensors client and the HeTra consumer client. The HeTra server is connected to both clients, requesting data from the former and delivering them to the latter. The system architecture is quite flexible, allowing the deployment of the server and the clients on the same PC or on different PCs.

The developed tool allows to organise and synchronize the data collected as well as to safely store the information to fulfil security requirements. HeTra system is illustrated in Fig. 2

4. ICT4LIFE platform use for Abnormal Behaviour detection

Due to the constrained length of the paper, we will focus on one example of the use of the ICT4LIFE platform: detecting abnormal behaviour. The Abnormal Behaviour Detection (ABD) module (see Fig. 3) focuses on analysing sensing information to identify behaviour patterns and inform any interested party (health professionals, formal/informal caregivers) about changes in the patient’s situation. Exploiting intelligent capabilities of an indoors activity recognition methodology, employing ambient and camera sensors, this core module provides the ICT4Life platform with the capability of assessing on the real deviations from the expected daily conduct of the person (e.g. the senior shows signs of apathy over a long time interval) and any (previously parameterized) human abnormal behaviour (e.g. for Parkinson’s patients, it needs to be controlled if the patient gets blocked or has fallen to help the patient or warn the caregivers).

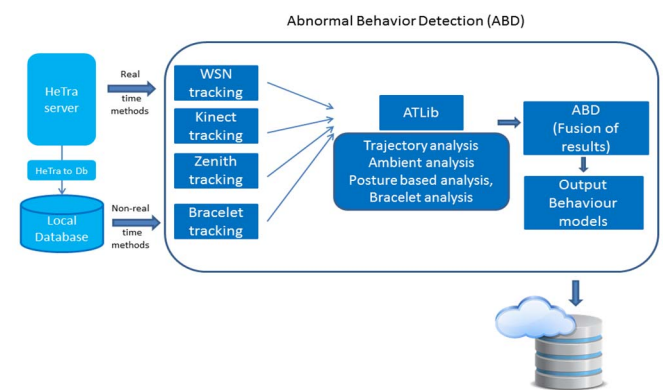


Figure 3: Abnormal Behaviour Detection architecture overview

4.1. Multimodal tracking

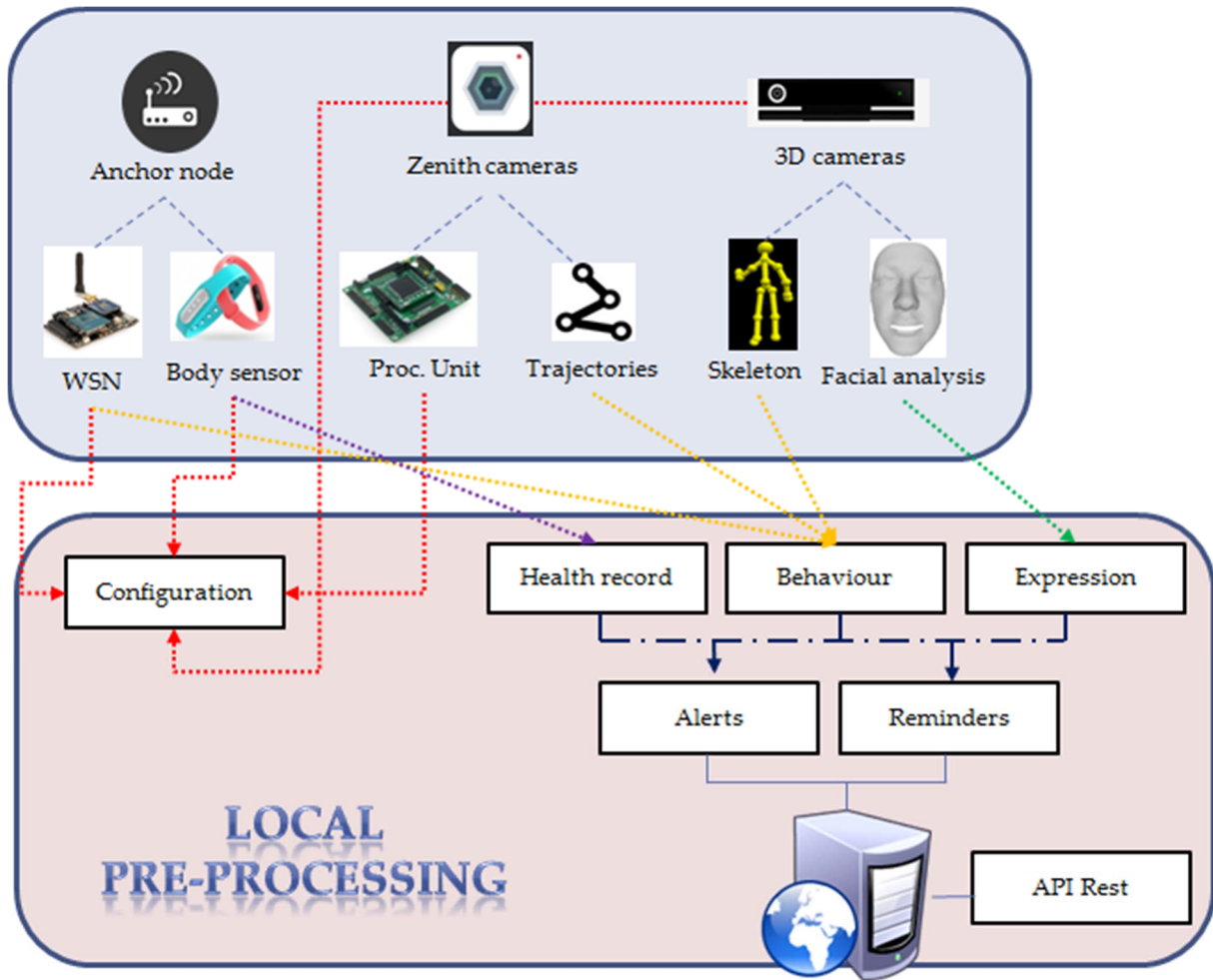


Figure 4: Multimodal tracking architecture overview

Multimodal tracking (see Fig. 4) is a process which provides input to the abnormal behaviour detection module. This is useful to perform the patient identification (especially relevant in situations where different patients are present), and the activity recognition. Once these processes are complete, they facilitate the detection of abnormal behaviour patterns by using high-level reasoning and taking into account the usual motion patterns of the patient.

4.1.1 Patient Identification.

The main limitations of the explored vision-based tracking techniques (either for RGB or RGB-D devices) are due to occlusions or strong lighting changes among others.

Specifically, RGB-D sensors, WSN and multi-sensory bands (bracelets) are considered for solving the aforementioned problem and link patients and RGB-D information. RGB-D sensors, in addition to image

information, are able to provide data from 1 up to 6 persons, that can be simultaneously tracked.

Moreover, a WSN is able to obtain the physical address of the bracelet (in particular, the MAC address). These data can be employed to perform inference of the patient's behaviour in the scene and can be used as an identification parameter.

Finally, in addition to health-related sensors, the bracelets are equipped with inertial sensors, such as accelerometer, gyroscope and magnetometer allowing for the estimation of acceleration, angular speed and relative/absolute position orientation.

Based on the aforementioned features, several comparison approaches for sensor association and identification have been proposed, which we combine depending on the result sought:

- Approach 1: to extract bracelet axis accelerations from inertial sensors (assuming that the patient is wearing the bracelet on the hand). Afterwards, perform an initial comparison of these data with the Kinect (i.e linear acceleration from x,y,z coordinates for

the corresponding skeleton joints) to establish the similarity degree.

- Approach 2: to extract yaw from bracelet sensors as well as orientation from Kinect (x,y coordinates to obtain movement patterns). Additionally, orientation from the WSN ($x,y,0$) is attached to estimate movement patterns. Finally, by comparing the similarity of the movement vectors, it is possible to associate the corresponding measurements.
- Approach 3: from WSN, to extract the patient's position ($x,y,0$) and compare it with the estimated position from Kinect ($x,y,0$ whenever joints close to the floor are chosen).

4.1.2 Activity recognition

The activity recognition module employs as input information to classify patients activity, their indoor movement (trajectories), namely the location (x,y) of the person across time. The trajectory of a person is estimated by using Kinect and Zenith cameras and WSN sensors. As the Kinect camera provides the coordinates of 25 joints of the patient's body, by using the coordinates of the head or spine-base joint, the person trajectory is obtained. By employing a Zenith camera that has a 360 degree top view of the room, the 2D trajectories of people moving in a room can also be obtained. Finally, the trajectory can be extracted from the WSN sensors, which estimate the user's location from the signal strength of the patient's wearable.

Examples of activities which can be detected using trajectory information are Daily motion and Sign of Apathy for which the user daily motion patterns are learned using the movement trajectories and then, every new trajectory is classified as normal or abnormal.

Basic user interaction with the environment (e.g. door/drawer opening/close) can be detected using binary sensors. Simple activities can be detected using the output of the binary sensors (e.g. number of visits to the bathroom) or by fusing them with the patient's trajectory for assessing Dream quality (Night Motion). Further events that can be detected are Patient Disoriented/Confused/Wandering, Patient (Disoriented and) leaving the house, which are especially relevant for Alzheimer's patients.

Besides patient location detection by the adoption of WSN, patient's wearable can be used in order to identify the patient's movement. By employing the accelerometer, gyroscope and magnetometer sensors, the user's movement can be estimated. By combining this information with the patient 3D skeleton coordinates acquired by the Kinect, fall, freezing, festination and loss of balance can be detected, which are common activities in the case of Parkinson's patients.

Finally, the proposed platform can track and evaluate the patient's movement during their rehabilitation exercises, record their performance and assess the patient's evolution.

4.2. High level fusion and probabilistic models to infer behaviour

The Multimodal Fusion (MF) module is responsible for fusing the retrieved signals from several Low-Level subsystem modules such as ABD and health monitoring (HM) with the medical information of the patient available in the Electronic Health Record (EHR), while its output is sent to the module responsible for taking decisions regarding the patient's health status and to the database where the patient-specific profile is stored. The fusion process of the input information takes into account the reliability of each source of information, using a weighted probabilistic model, where each modality is weighted according to its contribution in the fusion scheme. The probabilistic models for apathy are initially based on expert knowledge, provided by professionals and are later on personalized for each patient, by taking into account their daily routine.

$$P_i(\text{apathy}) = w_a P_{\text{ABD}}(\text{apathy}) + w_h P_{\text{HM}}(\text{apathy}) + w_e P_{\text{EHR}}(\text{apathy}) \quad (1)$$

An example of the fusion model proposed for computing the probability of a patient, having the id i , to be in an apathy state is shown in (1). Different modalities are taken into consideration, while their confidence level is provided by their associated probability measure. The weights associated with each modality (e.g. w_a , w_h , w_e) are based on knowledge or learnt during training. Furthermore, this model is also able to cope with missing or not-reliable information, by discarding that modality in the fusion process.

Provided that enough training data is available, a better fusion model, such as a Bayesian Network (BN) model can be used. The BN model is trained and used to infer the high-level information regarding the patient's health status based on observable variables provided by the Low-Level subsystem modules (such as heart rate, level of activity or time spent interacting with the digital platform). In case the training data is not sufficient to train the BN model, we can set the conditional probabilities using expert knowledge derived from the project end-users.

After obtaining the health status of the patient, it is used to update the EHR of the patient and sent to the Decision Support Tool (DST) for formulating recommendations.

5. Testing and results of patient behaviour analysis

An adaptive system has been developed and tested for monitoring an indoor environment, able to learn repeated patterns of activities and to detect non-expected behaviours (abnormalities), based on trajectory information. We recorded a dataset by tracking people for twenty-four days. We employ different feature descriptors, both spatial features, such as Occupancy histograms (OH), motion features and a fused spatio-temporal descriptor. On top of the raw features we use a Sparse AutoEncoder (SAE) algorithm [5] which is useful at learning an optimized representation in an unsupervised way, leading to a new and compact feature representation. As a classification model, we employ the Logistic Regression classifier.

In our experiments, we used 700 sample trajectories, from which 30% were labeled as abnormal and the as rest normal behaviours. The obtained accuracy for discriminating between normal and abnormal behaviours was 98.4%, while also the precision (98.7%) and the recall (98.3%) were really high, showing that our proposed methodology was successful and reliable.

Furthermore, the SAE representation is beneficial, improving the performance of the AHOT (Adaptive Histogram of Oriented Tracklets) descriptor by 2%. The Occupancy Histogram descriptor achieves a good accuracy of 97.4%, as it is able to capture the correlation between activity patterns and the specific spatial regions where they are usually performed.

Examples of detected normal and abnormal behaviour patterns are presented in Figure 5. Fig. 5a depicts the most common behaviour pattern in an office (in green), as people are most of the time sitting at their desk, working, while in Fig. 5b an abnormal behaviour is shown (in red), as no one is sitting at their desk and the tracks are distributed all over the room and near the door region.

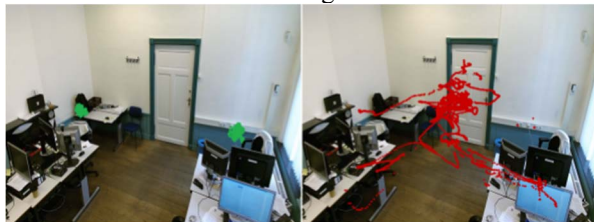


Figure 5 Examples of (a) normal behaviour (b) abnormal behaviour.

6. Conclusions

In this paper a complete ICT-based system to support people with cognitive impairments in their daily activities has been presented. In this article we show how regular sensor technology, involving RGB-D, Zenith camera, wearables, and other IoT devices, is transforming the way context and human behaviour is reasoned about and we

highlight its applicability at improving patients' quality of life.

The Multimodal Fusion techniques presented in this work guarantee the modularity of the entire system by allowing its operation even in cases where not all presented sensors are available.

The architecture for integrating multimodal sensing technologies and the fusion of sensory data, are offering new ways of inferring patient behaviour, being applicable to different situations where the patient might be at risk. Furthermore, the methodology for discriminating between normal and abnormal behaviour patterns was introduced, achieving a good performance which is expected to play an important role in signalling potential changes in the physical, mental or emotional state of the patient.

References

- [1] Pawar, P., Jones, V., van Beijnum, B.-J. F., Hermens, H. (2012) "A framework for the comparison of mobile patient monitoring systems", *Journal of Biomedical Informatics*, vol. 45 no. 3, pp. 544-556, 2012.
- [2] Kasteren, T. v., Englebienne, G., and Krse, B. (2010). "Activity recognition using semi-markov models on real world smart home datasets", in the *Journal of Ambient Intelligent Smart Environments*, nr. 2, pp. 311-325.
- [3] Nef, T., Urwyler, P., Buchler, M., Tarnanas, I., Stucki, R., Cazzoli, D., Muri, R., and Mosimann, U. (2015). "Evaluation of Three State-of-the-Art Classifiers for Recognition of Activities of Daily Living from Smart Home Ambient Data", *Sensors*, vol. 15, no. 5, pp. 11725-11740.
- [4] Hoque, E., Dickerson, R. F., Preum, S. M., Hanson, M., Barth, A., and Stankovic, J. A. (2015). "Holmes: A comprehensive anomaly detection system for daily in-home activities", In the *11th IEEE Int. Conf. on Distributed Computing in Sensor Systems*.
- [5] Efros, A., Berg, A., Mori, G., and Malik, J. (2003). "Recognizing action at a distance". In the *IEEE Int. Conf. on Computer Vision*, pp. 726-733.
- [6] Jiang, F., Wu, Y., and Katsaggelos, A. K. (2009). "A dynamic hierarchical clustering method for trajectory based unusual video event detection." In the *IEEE Transactions on Image Processing*, vol. 18, no. 4, pp. 907-913.
- [7] Masci, J., Meier, U., Ciresan, D., and Schmidhuber, J. (2011) "Stacked convolutional auto-encoders for hierarchical feature extraction" In the *21th International Conference on Artificial Neural Networks (ICAN'11)*, volume 1, pp. 52-59.
- [8] Knickman JR, Snell EK. The 2030 Problem: Caring for Aging Baby Boomers. *Health Services Research*. 2002;37(4):849-884