

# Low-Cost Automatic Ambient Assisted Living System

Hossein Malekmohamadi, Armaghan Moemeni, Ahmet Orun and Jayendra Kumar Purohit

School of Computer Science and Informatics

De Montfort University, The Gateway, Leicester, LE1 9BH, UK

Email: {hossein.malekmohamadi, amoemeni, aorun, jay.purohit@dmu.ac.uk}

**Abstract**—The recent increase in ageing population in countries around the world has brought a lot of attention toward research and development of ambient assisted living (AAL) systems. These systems should be inexpensive to be installed in elderly homes, protecting their privacy and more importantly being non-invasive and smart. In this paper, we introduce an inexpensive system that utilises off-the-shelf sensor to grab RGB-D data. This data is then fed into different learning algorithms for classification different activity types. We achieve a very good success rate (99.9%) for human activity recognition (HAR) with the help of light-weighted and fast random forests (RF).

**Index Terms**—Assisted Living, Kinect, Naïve Bayes, Multi Layer Perceptron, Random Forest, Human activity recognition.

## I. INTRODUCTION

Many countries around the world are facing ageing population. This fact inevitably necessitates an innovative development of more efficient low-cost ambient assisted living (AAL) monitoring system with the use of optimised infrastructure. Essentially, an AAL system comprises a number of subsystems requiring a multi-disciplinary approach in terms of research, design, development, integration and deployment. The motivation behind an AAL system is to provide elderly or disabled people with affordable chronic health care monitoring facilities in their own homes and thereby promote well-being and independence. To realise this vision, the research community is currently actively working in various related fields such as sensor technologies (wearable, environmental, physiological, audio and video), activity identification and analysis of behavioural patterns for long-term predictive health-assessment analytics. Identifying human behaviour requires appropriate sensors of which there are of three main types: wearable, distributed environmental and vision-based.

Abnormal activity detection and classification (e.g. fall detection) is a subject of interest in research communities including computer vision. The ultimate goal of AAL technologies is to have a robust, inexpensive, non-invasive and accurate system [1], [2]. One of the earliest studies; pre-Kinect era; [3] involves a 3D video camera for human activity recognition (HAR) which is used to classify three activities. The Microsoft Kinect sensor is an inexpensive ( 100 GBP) RGB-D video camera of relatively recent invention (2010). The depth information can be used to construct a skeleton model of a human subject in a real-time fashion. Most significantly, facial and body features are stripped away and can be hidden from

view or storage. The Kinect sensor in the context of HAR can be truly regarded as a game changing device.

In [4], a hierarchical two-layer maximum entropy Markov model is introduced to infer activities from the Kinect sensor data. It assumes an activity comprises a set of sub-activities. Tests were performed on four people carrying out twelve activities. A different method of detecting activities is presented in [5] where a k-means clustering algorithm forms clusters from features, which are then processed by hidden Markov models (HMMs). Reference [6] considers activity detection in terms of interactions between a sub-activities sequence and objects, by modelling data from the Kinect RGB-D video and using a structured support vector machine (SSVM). The approach in [7] is to identify significant sub-activities using support vector machine (SVM). Using HMM, each high-level activity is then represented as a sequence of known sub-activities. A useful literature review on semantic HAR can be found in [8], [9]. Postural control assessment where physical function and fall risk can be assessed using Kinects 3D joints data [10], [11]. Fall risk identification by measuring gait parameters is another area of earnest research. Further readings can be found in [9], [12], [13], [14]. Some of these studies employ Bayesian classification probabilistic modelling for estimating gait parameters.

In this paper, a low-cost system for HAR is presented. This system utilises Microsoft Kinect V2 sensor for data acquisition and learning algorithms: naïve Bayes (NB), multi layer perceptron (MLP), random forest (RF). For this purpose, details of experimental set-up to record RGB-D data are presented in Section II. Classification results are in Section III while Section IV concludes this paper.

## II. EXPERIMENTAL SET-UP

The system architecture comprises some main component sub-systems as shown in Figure 1:

- Microsoft Kinect sensor interfaced to a local processing hub for RGB-D processing of depth images for the purpose of activity monitoring. The Kinect safeguards the privacy of people by capturing body skeleton pose data only. No video images need be captured or saved.
- Raspberry Pi 2, a low-cost single board computer which performs as an IoT hub for interfacing environmental monitoring and body sensing devices. The small size and

low cost enable multiple such devices to be installed as needed in a practical AAL environment.

- Microsoft Azure, an open and flexible leading cloud infrastructure offering integrated services such as computing, database, networking, storage, web services, analytics and IoT.

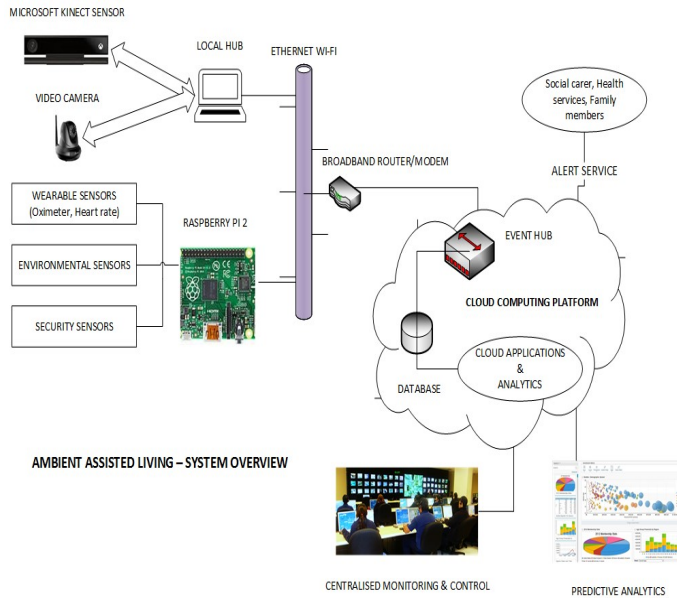


Figure 1. The proposed low-cost AAL system architecture

The Microsoft Kinect v2 grabs a high-definition (1920×1080) RGB map and a depth map (512×424) at 30fps, and a microphone array to provide audio data. The depth camera comprises an infra-red (IR) emitter and a detector that provides depth information in millimetres. The IR emitter projects into the 3D scene a dot pattern which is reflected and received by the IR detector. The depth sensor has a physical range from 0.5m to 8m. However, reliable body detection range is from 0.5m to 4.5m (0.8m to 3.5m is recommended). The depth angle of vision is 60 degrees vertical and 70 degrees horizontal. The sensor API is provided by Kinect for Windows 2.0 which can be downloaded from the Microsoft development website<sup>1</sup>. The API allows up to six people to be tracked (30fps) as whole skeletons, each with 25 body joints. Seated skeletons are tracked with the upper 10 body joints only. Each joint has an orientation and 3D position in space.

One useful feature of the Kinect is to capture IR images which exclude ambient light. These can be very useful for machine learning approaches in computer vision processing applications. In this paper, to satisfy the privacy requirements, only the skeleton joints data is used for activity monitoring purposes. Specifically, the RGB colour image frames are not used. Body frame data is captured once every second. Pre-processing of the raw captured frame data produces a total of

four datasets as follows: Raw data as captured by the Kinect sensor, normalised data with the spine body joint at the origin of a 3D coordinate system, normalised data with the three axis normalised as well, and normalised data with only the  $y$ -axis (representing height) normalised.

The Kinect Windows Application is used to create an activity dataset which is then later utilised in the training of machine learning algorithms for activity classification purposes. Fifteen human volunteers were recruited to help with activity tests, and were provided with a participant information sheet and signed consent obtained from each of them to satisfy the ethics requirements. Each test subject performed, in turn, thirteen structured activities in front of the Kinect. Body joints positional data was recorded for each activity lasting 30 seconds. Data capture rate was 1 frame per second. Pre-processing of the raw dataset is programmed in the software application, and this produces a total of 4 datasets as follows: Raw data as captured by the Kinect sensor, normalised data with the spine body joint at the origin of a 3D coordinate system, normalised data with the three axis normalised as well, and normalised data with only the  $y$ -axis (representing height) normalised. In this paper, the simulated environment for activity recording is a simplistic representation of real-world scenarios, necessitated by the constraints in available resources and project needs. However, the approach is judged to be adequate for demonstrating the basic methodology and analysis of results. Crucially, a real-world home environment would present greater challenges, such as: multiple activity locations, distance limitations, presence of multiple persons, variable viewing angles and body view obstructions. In this paper, 13 activity classes have been described which are: 1) standing, 2) sitting, 3) lying down in sleep position, 4) drinking from a mug, 5) eating with knife and fork, 6) using a mobile phone, 7) brushing teeth, 8) vacuuming, 9) ironing, 10) washing hands in a bowl, 11) brushing hair, 12) exercise foot cycle and 13) floor fall position. A sample of activities is shown in Figure 2.

A total of four datasets are created from the captured body frames as follows: 1) Raw data as captured by the Kinect sensor, 2) normalised data with the spine body joint set at the origin of a 3D coordinate system, 3) normalised data with the three axis normalised (-1000mm to 1000mm) as well, and 4) normalised data with only the  $y$ -axis (representing height) normalised (-1000mm to 1000mm). These datasets are labelled as:  $Raw_{data}$ ,  $NO_{data}$ ,  $NA_{data}$  and  $NAY_{data}$ . Each of the four activity datasets is made up of one row per activity per person. With 15 participants each performing 13 activities for a duration of 30 seconds, at one frame-per-second, gives a total of 5850 rows per dataset. Each row represent one frame of body joints pose data. Each joint position in 3D space is represented by  $x$ ,  $y$  and  $z$  co-ordinates. With 25 joints per frame, this gives 75 columns, each column in the dataset is considered as a feature for supervised machine learning purposes. A sample skeleton single frame datasets is shown in Figure 3.

<sup>1</sup><https://www.microsoft.com/en-gb/download/details.aspx?id=44561>

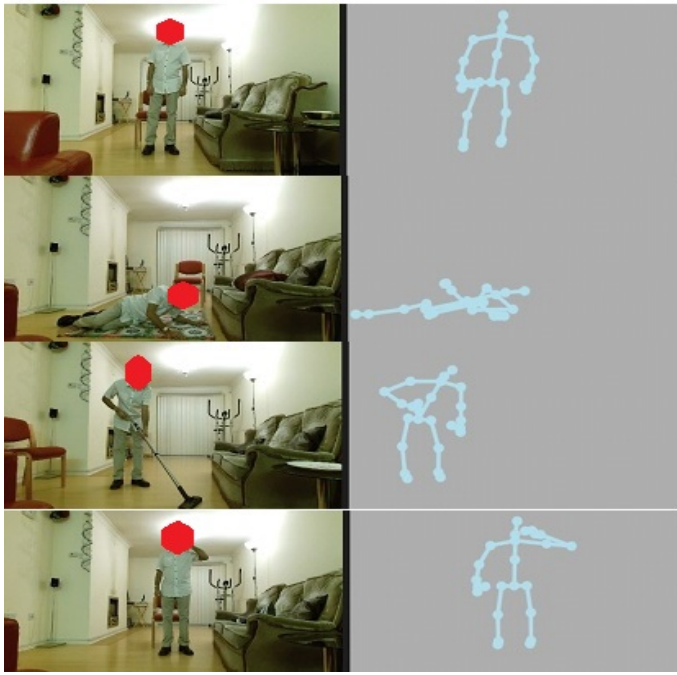


Figure 2. Sample activity used in this paper: standing, floor fall position, vacuuming and brushing teeth are shown.

Joint	Raw data			NO data			NA data			NAY data		
	X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z
0 SpineBase	-327	28	2021	0	0	0	0	0	0	0	0	0
1 SpineMid	-332	387	2014	-5	359	-7	-15	427	-14	-7	427	-10
2 Neck	-336	729	1994	-9	701	-27	-28	833	-55	-13	833	-39
3 Head	-312	870	1992	15	842	-29	48	1000	-59	18	1000	-42
4 ShoulderLeft	-493	595	1949	-166	567	-72	-528	674	-148	-242	674	-105
5 ElbowLeft	-641	470	1894	-314	442	-127	-1000	525	-262	-458	525	-185
6 WristLeft	-620	697	1836	-293	669	-185	-933	795	-382	-427	795	-270
7 HandLeft	-607	763	1834	-280	735	-187	-891	873	-386	-408	873	-272
8 ShoulderRight	-172	595	2018	155	567	-3	493	674	-6	185	674	-4
9 ElbowRight	-104	351	2050	223	323	29	708	384	1000	265	384	35
10 WristRight	-48	166	1996	279	138	-25	886	164	-51	332	164	-36
11 HandRight	-36	92	1962	291	64	-59	924	77	-121	346	77	-86
12 HipLeft	-390	34	1967	-63	6	-54	-200	8	-111	-91	8	-78
13 KneeLeft	-414	-399	1915	-87	-427	-106	-277	-623	-219	-127	-623	-154
14 AnkleLeft	-429	-657	1861	-102	-685	-160	-324	-1000	-330	-148	-1000	-233
15 FootLeft	-356	-647	1702	-29	-675	-319	-92	-985	-659	-42	-985	-465
16 HipRight	-251	21	1991	76	-7	-30	242	-10	-61	91	-10	-43
17 KneeRight	-217	-363	1981	110	-391	-40	350	-570	-82	131	-570	-58
18 AnkleRight	-184	-567	1682	143	-595	-339	454	-868	-700	170	-868	-494
19 FootRight	-158	-603	1537	169	-631	-484	537	-921	-999	201	-921	-706
20 SpineShoulder	-335	646	2001	-8	618	-20	-25	734	-41	-11	734	-29
21 HandTipLeft	-592	842	1850	-265	814	-171	-843	967	-353	-386	967	-249
22 ThumbLeft	-555	740	1815	-228	712	-206	-726	846	-425	-332	846	-300
23 HandTipRight	-12	6	1923	315	-22	-98	1000	-32	-202	375	-32	-143
24 ThumbsRight	-33	68	1898	294	40	-123	934	48	-254	350	48	-179

Figure 3. Sample skeleton single frame datasets.

### III. CLASSIFICATION

In this section, we present the results and analysis for naïve Bayes, multi layer perceptron and random forest for training classification models for human activity recognition. Multi layer perceptron is a feed forward artificial neural network model that maps sets of input data onto a set of output. Multi layer perceptron consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one.

Multi layer perceptron utilises a supervised learning technique called back-propagation for training. Two key advantages of multi layer perceptron are generalisation and fault tolerance [15]. Naïve Bayes is another classification technique which assumes that the value of a particular feature is independent of the value of any other feature, given the class variable. For some types of probability models, these classifiers can be trained very efficiently in a supervised learning setting. One advantage of naïve Bayes is that it only requires a small number of training data to estimate the parameters necessary for classification. Random forest is an ensemble classifier that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes of the individual trees. Unlike decision trees, random forest tackles over-fitting of training set.

Before presenting the results, the performance metrics used for evaluating performance are defined as:

- True Positive (TP) : The activity class has been correctly identified.
- True Negative (TN) : The activity class has been correctly rejected.
- False Positive (FP) : The activity class has been incorrectly identified.
- False Negative (FN) : The activity class has been incorrectly rejected.
- Precision (positive predictive value) : The correctly identified proportion in predicted activities which is:  
Precision = TP / (TP + FP)
- Recall (sensitivity) : The correctly identified proportion in actual activities which is:  
Recall = TP / (TP + FN)
- Overall Accuracy : The correctly identified proportion of the total activities.
- F1 Score : Weighted average of precision and recall is calculated as:  
F1 = 2 (Precision × Recall) / (Precision + Recall)

All the four activity datasets (one raw and three normalised) were experimented with. The results presented here are from one dataset  $NO_{data}$  which gave the best results. Each dataset comprises 5850 samples, randomly divided as 50% for training and 50% for testing. Table I compares the performance metrics of recall, precision and F measure for each activity class. On all three measures, the RF model performs better with the precision of 99.9%. It is fair to say that activity a9 (ironing) negatively impacts the performance. But overall, the RF appears to be the more preferred approach, especially when the model will have to be enhanced to cope with more complex sequence of activities, as would be expected in a real living environment. RF is trained with 5-fold cross validation and it has 11 trees. The precision measure of multi layer perceptron is the second best in this paper (98.7%). Compared to random forests, multi layer perceptron performance is less for activities: a1 (standing), a4 (drinking from a mug), a6 (using a mobile phone), a7 (brushing teeth), a9 (ironing), a10 (washing hands in a bowl) and a11 (brushing hair).

Table I  
 NAÏVE BAYES (NB), MULTI LAYER PERCEPTRON (MLP) AND RANDOM FOREST (RF) TRAINED MODELS PERFORMANCE COMPARISON: SHOWN RESULTS ARE FOR TESTING DATASETS (50% OF OVERALL DATA).

Class	Recall			Precision			F1 Score		
	NB	MLP	RF	NB	MLP	RF	NB	MLP	RF
a1: standing	0.991	1	1	0.996	0.996	1	0.993	0.998	1
a10: washing hands in a bowl	0.981	1	1	0.963	0.991	1	0.972	0.995	1
a11: brushing hair	0.482	0.991	1	0.566	0.978	1	0.521	0.984	1
a12: exercise foot cycle	0.965	1	1	0.909	1	1	0.936	1	1
a13: floor fall position	0.908	1	1	0.857	1	1	0.882	1	1
a2: sitting	0.898	1	1	0.818	1	1	0.856	1	1
a3: lying down in sleep position	0.85	1	1	0.903	1	1	0.876	1	1
a4: drinking from a mug	0.834	0.995	0.986	0.503	0.981	1	0.627	0.988	0.993
a5: eating with knife and fork	0.764	1	1	0.847	1	1	0.803	1	1
a6: using a mobile phone	0.924	0.924	1	0.846	0.987	1	0.884	0.954	1
a7: brushing teeth	0.425	1	1	0.696	0.922	1	0.527	0.96	1
a8: vacuuming	0.966	0.92	1	0.996	1	1	0.981	0.958	1
a9: ironing	0.786	1	1	0.99	0.976	0.986	0.876	0.988	0.993
Average	0.83	0.986	0.999	0.841	0.987	0.999	0.828	0.986	0.999

However, multi layer perceptron advantages like extracting meaning from complicated and imprecise data especially in HAR research make them very popular in this field. MLP has 38 nodes and activation function is *tanh*. Finally, the naïve Bayes lacks performance noticeably in: a4 (drinking from a mug), a7 (brushing teeth) and a11 (brushing hair). Drinking activity with naïve Bayes is often misclassified as brushing hair or brushing teeth. Brushing teeth with naïve Bayes is also misclassified as brushing hair and drinking from mug. Lastly, brushing hair with naïve Bayes is misclassified as exercise foot cycle, drinking from a mug and brushing teeth. This is an evidence that naïve Bayes cannot deal with noisy and incomplete data properly.

In another experiment, principal component analysis (PCA) is applied to the feature set. PCA is a statistical procedure that uses an orthogonal transformation to convert a set of possibly correlated features into a set linearly uncorrelated principal components. The number of principal components is less than or equal to the smaller of the number of original variables or the number of observations. This transformation is defined in such a way that the first principal component has the largest possible variance, and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. In this paper, the number of feature set for training algorithms is reduced to 10 principal components. The first 5 major components are (Variables are shown in Figure 3):

- $P_1 = -0.169X_{14} - 0.168X_{15} - 0.165X_{13} - 0.165X_{19} - 0.164X_{18}$ .  
It takes X co-ordinates of Left Ankle, Left Foot, Left Knee, Right Foot and Right Ankle into consideration.
- $P_2 = -0.195Z_1 - 0.192Z_{12} - 0.191Z_0 - 0.188Z_{10} - 0.188Z_{16}$ .  
It takes Z co-ordinates of Mid Spine, Left Hip, Base Spine, Right Wrist and Right Hip into consideration.
- $P_3 = 0.254X_4 + 0.25X_2 + 0.248X_5 + 0.246X_{20} +$

$0.243X_3$ .

It considers x co-ordinates Left shoulder, Neck, Left Elbow, Shoulder Spine and Head.

- $P_4 = -0.287Y_{19} - 0.279Y_{18} - 0.276Y_{15} - 0.258Y_{14} + 0.244Z_{19}$ .  
It considers y co-ordinates of Right Foot, Right Ankle, Left Foot, Left Ankle and z co-ordinate of Right Foot.
- $P_5 = -0.318X_{22} + 0.294Y_{24} - 0.283X_{23} - 0.279X_7 + 0.264Y_{11}$ .  
It considers x co-ordinates of Left Thumb, Right Hand tip and Left Hand. It also takes y co-ordinates of Right Thumb and Right Hand into consideration.

This is partly because the type of the actions discussed in this paper implies the importance of these parameters. For example, y co-ordinates play less important role because of the human pose of these activities. The above-mentioned classifiers are trained and tested considering PCA. The results are shown in Table II. On all three measures, the RF model performs better with the precision of 99.4%. The precision measure of multi layer perceptron is the second best in this paper (91%). Compared to random forests, multi layer perceptron under-performance is noticeable for brushing teeth and drinking from a mug.

The naïve Bayes performance is much less (80.1%); compared to random forest and multi layer perceptron; in activities: brushing teeth, eating with knife and fork, drinking from a mug, sitting and brushing hair. Misclassification of similar activities like brushing teeth, drinking and brushing hair proves the need for better data acquisition techniques to tackle noisy or incomplete data. All the classifiers mentioned in this paper are very fast in training, robust and easy to deploy in single board computers like Raspberry Pi 2. The processing power and memory requirements for these classifiers are not remarkable hence they can be implemented with inexpensive hardware for HAR purposes.

Table II  
 NAÏVE BAYES (NB), MULTI LAYER PERCEPTRON (MLP) AND RANDOM FOREST (RF) TRAINED MODELS PERFORMANCE COMPARISON FOR REDUCED FEATURE SETS WITH PCA: SHOWN RESULTS ARE FOR TESTING DATASETS (50% OF OVERALL DATA).

Class	Recall			Precision			F1 Score		
	NB	MLP	RF	NB	MLP	RF	NB	MLP	RF
a1: standing	0.973	1	1	0.846	0.991	0.996	0.905	0.996	0.998
a10: washing hands in a bowl	0.730	0.976	0.981	0.875	0.976	1	0.796	0.976	0.990
a11: brushing hair	0.545	0.703	0.982	0.654	0.821	0.995	0.595	0.757	0.989
a12: exercise foot cycle	0.882	0.921	0.982	0.939	0.925	1	0.910	0.923	0.991
a13: floor fall position	0.908	0.979	0.996	0.867	0.932	1	0.887	0.955	0.998
a2: sitting	0.898	0.961	1	0.679	0.895	1	0.773	0.927	1
a3: lying down in sleep position	0.863	0.975	1	0.900	0.947	1	0.881	0.961	1
a4: drinking from a mug	0.820	0.763	0.991	0.479	0.805	0.995	0.605	0.783	0.993
a5: eating with knife and fork	0.564	0.851	1	0.671	0.892	0.970	0.613	0.871	0.985
a6: using a mobile phone	0.924	0.924	1	0.936	0.973	0.996	0.930	0.948	0.998
a7: brushing teeth	0.248	0.881	0.996	0.683	0.724	0.987	0.364	0.794	0.991
a8: vacuuming	0.929	0.828	1	0.936	1	1	0.932	0.906	1
a9: ironing	0.915	1	0.996	0.863	0.922	0.984	0.888	0.959	0.990
Average	0.789	0.907	0.994	0.801	0.910	0.994	0.781	0.906	0.994

After training and validating the classifier, it is possible for MATLAB to generate a packaged code module which can be embedded to an external application. In a live HAR system, this would then be used to infer the user activity type by reading Kinect skeletal data stream in real-time. In this paper and for cloud deployment, three software applications (a) the Kinect simulator for live activity data generation, (b) RPi2 simulator for environmental sensors data generation and (c) cloud storage data simulator have been developed. They provide simulated real-time test data on demand and generate event data at an accelerated rate.

#### IV. CONCLUSIONS

This paper demonstrates the feasibility of a low-cost end-to-end functioning AAL system. Scalability, data storage and secure communications with IoT type devices are necessary in AAL system which are addressed in this paper. Moreover, data acquisition and classification results of this paper show the reliability and accuracy of the overall system. Future research include investigating other camera types, applying new feature sets based on image/video and using other learning algorithms, e.g. deep learning.

#### REFERENCES

- [1] D. N. Monekosso, F. Florez-Revuelta, and P. Remagnino, "Guest editorial special issue on ambient-assisted living: Sensors, methods, and applications," *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 5, pp. 545–549, 2015.
- [2] D. Monekosso, F. Florez-Revuelta, and P. Remagnino, "Ambient assisted living [guest editors' introduction]," *IEEE Intelligent Systems*, vol. 30, no. 4, pp. 2–6, 2015.
- [3] B. Jansen, F. Temmermans, and R. Deklerck, "3d human pose recognition for home monitoring of elderly," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE, 2007, pp. 4049–4051.
- [4] J. Synnott, C. Nugent, and P. Jeffers, "Simulation of smart home activity datasets," *Sensors*, vol. 15, no. 6, pp. 14 162–14 179, 2015.
- [5] T. Liu, Y. Song, Y. Gu, and A. Li, "Human action recognition based on depth images from microsoft kinect," in *Intelligent Systems (GCIS), 2013 Fourth Global Congress on*. IEEE, 2013, pp. 200–204.
- [6] H. S. Koppula, R. Gupta, and A. Saxena, "Learning human activities and object affordances from rgb-d videos," *The International Journal of Robotics Research*, vol. 32, no. 8, pp. 951–970, 2013.

- [7] P. Cottone, G. Maida, and M. Morana, "User activity recognition via kinect in an ambient intelligence scenario," *IERI Procedia*, vol. 7, pp. 49–54, 2014.
- [8] M. Ziaefard and R. Bergevin, "Semantic human activity recognition: a literature review," *Pattern Recognition*, vol. 48, no. 8, pp. 2329–2345, 2015.
- [9] E. E. Stone and M. Skubic, "Unobtrusive, continuous, in-home gait measurement using the microsoft kinect," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 10, pp. 2925–2932, 2013.
- [10] R. A. Clark, Y.-H. Pua, K. Fortin, C. Ritchie, K. E. Webster, L. Denehy, and A. L. Bryant, "Validity of the microsoft kinect for assessment of postural control," *Gait & posture*, vol. 36, no. 3, pp. 372–377, 2012.
- [11] H. P. Shum, E. S. Ho, Y. Jiang, and S. Takagi, "Real-time posture reconstruction for microsoft kinect," *IEEE transactions on cybernetics*, vol. 43, no. 5, pp. 1357–1369, 2013.
- [12] A. Prochazka, O. Vysata, M. Valis, O. Tupa, M. Schatz, and V. Marik, "Bayesian classification and analysis of gait disorders using image and depth sensors of microsoft kinect," *Digital Signal Processing*, vol. 47, pp. 169–177, 2015.
- [13] A. Procházka, O. Vyšata, M. Vališ, O. Ťupa, M. Schätz, and V. Mařík, "Use of the image and depth sensors of the microsoft kinect for the detection of gait disorders," *Neural Computing and Applications*, vol. 26, no. 7, pp. 1621–1629, 2015.
- [14] X. Xu and R. W. McGorry, "The validity of the first and second generation microsoft kinect for identifying joint center locations during static postures," *Applied ergonomics*, vol. 49, pp. 47–54, 2015.
- [15] S. S. Haykin, *Neural networks and learning machines*. Pearson Upper Saddle River, NJ, USA., 2009, vol. 3.