Unobtrusive Activity Recognition of Elderly People Living Alone Using Anonymous Binary Sensors and DCNN

Article *in* IEEE Journal of Biomedical and Health Informatics · May 2018

DOI: 10.1109/JBHI.2018.2833618

6 authors, including:

Munkhjargal Gochoo
National Taipei University of Technology

Tan-Hsu Tan
National Taipei University of Technology

Shing-Hong Liu
Chaoyang University of Technology

Fady Shibata Alnajjar
United Arab Emirates University

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2833618, IEEE Journal of Biomedical and Health Informatics

Unobtrusive Activity Recognition of Elderly People Living Alone Using Anonymous Binary Sensors and DCNN

Munkhjargal Gochoo, Tan-Hsu Tan, *Senior Member, IEEE*, Shing-Hong Liu, Fu-Rong Jean, Fady S. Alnajjar, and Shih-Chia Huang, *Senior Member, IEEE*

***Abstract*—Elderly population (over the age of 60) is predicted to be 1.2 billion by 2025. Most of the elderly people would like to stay alone in their own house due to the high eldercare cost and privacy invasion. Unobtrusive activity recognition is the most preferred solution for monitoring daily activities of the elderly people living alone rather than the camera and wearable devices based systems. Thus, we propose an unobtrusive activity recognition classifier using Deep Convolutional Neural Network (DCNN) and anonymous binary sensors that are PIR motion sensors and door sensors. We employed Aruba annotated open dataset that aquired in a smart home while a voluntary single elderly woman was living inside for eight months. First, ten basic daily activities, namely: Eating, Bed_to_Toilet, Relax, Meal_Preparation, Sleeping, Work, Housekeeping, Wash_Dishes, Enter_Home, and Leave_Home are segmented with different sliding window sizes, and then converted into binary activity images. Next, the activity images are employed as the ground truth for the proposed DCNN model. The 10-fold cross-validation evaluation results indicated that our proposed DCNN model outperfors the existing models with F_1-score of 0.79 and 0.951 for all 10 activities and eight activities (excluding Leave_Home and Wash_Dishes), respectively.**

***Index Terms*— unobtrusive, device-free, deep learning, activity recognition, elder care**

I. Introduction

In the last few decades, a single independent lifestyle among the elderly people is trending worldwide [1]–[6]. Study predicts that the number of the elderly people who are 60 and above will reach 1.2 billion by 2025 [7]. However, many of them cannot continue their independent life due to the physical or mental issues as they age.

Activity recognition of elderly people is important for it can help elderly people to maintain their independent life by recommending a healthy lifestyle based on the monitored activities' information and detecting early symptoms of any abnormalities. Basically, monitoring systems are divided into three main categories that are (1) camera-based [8]–[11]; (2) wearable devices based [12]–[26]; (3) and anonymous binary sensors based [7], [27]–[30], [46]–[48]. Anonymous binary sensors are the sensors that generate only two outputs, that is, passive infrared (PIR) sensors, magnetic door sensors, etc. Among these three types of monitoring systems, the camera-based systems are the best in terms of accuracy; however, the camera-based monitoring systems are not preferred systems because of their privacy-invasiveness. Next, the wearable device-based systems don't have the privacy concern; however, such systems may not be practical when employed for the long-term activity recognition applications because of its natural flaws such as the easy loss of the wearable device, maintenance burden, short battery life, and discomfort of wearing it [31]. Thus, the most preferred solution for the real-life long-term monitoring is the unobtrusive (device-free and non-privacy invasive) system. However, the existing anonymous binary sensor-based systems [46]–[48] have employed the conventional machine learning models and their highest F_1-score for the same Aruba open dataset is 0.75.

The main goal of this work is to propose an activity recognition model for the elderly people living alone employing Deep Convolutional Neural Network (DCNN) and an fully annotated Aruba open dataset [32]. Only the anonymous binary sensors are employed for the collecting Aruba dataset. First, ten daily activities are segmented and sliding windowed from the dataset, then converted into activity images. DCNN is

This study was partially supported by Ministry of Science and Technology of the Republic of China (Taiwan) under the Contract No. MOST 106-2218-E-027-017 and MOST 107-2218-E-027-011. This is an extended version of the paper that was presented at the International Conference on Electrical and Computing Technologies and Applications 2017 [55].

(*Corresponding author*: *Shih-Chia Huang*)

M. G. is with Department of Electronic Engineering, National Taipei University of Technology, Taipei 10608, Taiwan, R.O.C., and School of Information and Communication Technology, Mongolian University of Science and Technology, Ulaanbaatar 13341, Mongolia (e-mail: g.munkhjargal@must.edu.mn).

T.-H. T. is with Department of Electrical Engineering, National Taipei University of Technology, Taipei 10608 Taiwan, R.O.C. (e-mail: thtan@ntut.edu.tw).

S.-H. L. is with Department of Information Engineering, Chaoyang University of Technology, Taichung 413, Taiwan, R.O.C. (e-mail: shliu@cyut.edu.tw).

F.-R. J. is with Department of Electrical Engineering, National Taipei University of Technology, Taipei 10608 Taiwan, R.O.C. (e-mail: frjean@ntut.edu.tw).

F.-S. A. is with College of Information Technology (CIT), The United Arab Emirates University, Al-Ain P.O. Box 15551, United Arab Emirates (e-mail: fady.alnajjar@uaeu.ac.ae)

S.-C. H is with Department of Electronic Engineering, National Taipei University of Technology, Taipei 10608, Taiwan, R.O.C. (e-mail: schuang@ntut.edu.tw).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2833618, IEEE Journal of Biomedical and Health Informatics

employed to extract spatial information (intra-sensor patterns) from the activity images to distinguish them. The 10-fold cross-validation results indicated that our proposed model outperformed the existing works with F_1-scores of 0.951 and 0.79 for, respectively, 8 and 10 activities. The contributions of the study is summarized as follows:

- We propose an unobtrusive DCNN activity recognition model for the elderly people living alone;
- We propose an activity image that is a 2D representation of the binary sensor logs;
- To the best of our knowledge, our proposed DCNN classifier yielded the highest recognition rate among the existing systems that employed the same dataset.

II. Related Works

Recently, several activity recognition methods using convolutional neural networks (CNN) and camera video [10], [33], [34], wearable devices [35]–[38], and binary sensors [39] have been proposed.

Gowda *et al.* [10] provide a CNN model that recognizes human activities from the videos. They have employed two datasets that have 51 actions with more than 100 videos for each action. The proposed method yielded an accuracy of 80.48% and 91.21% for two datasets. Wang *et al*. [33] present a Recurrent CNN (RNN) model that detects daily activities from egocentric videos. They employed a dataset consisting of 18 activities of daily living (ADL) that are collected by 20 different persons. The accuracy of the model was 53%. Adhikari *et al*. [34] have implemented an indoor fall detection model using Kinect cameras and CNN. The overall accuracy was 74%. Papakostas *et al*. [11] have developed a CNN classifier for indoor activity recognition with camera videos. Its accuracy for sitting, standing, and walking was, respectively, 96.41%, 79.18%, and 98.94%.

Ordóñez and Roggen [35] have proposed activity recognition classifier using DCNN and long short-term memory (LSTM) and two open datasets collected by seven inertial measurement units and 12 triaxial accelerometer wearable sensors. They have classified 27 hand gestures (opening door, washing dishes, cleaning table, and etc.), five movements (standing, walking, sitting, lying, and nulling) using a combination of DCNN and LSTM after converting the sensory data into sensor signal graphs. Simulation results showed that F_1 score of 0.93 and 0.958 was achieved with the two datasets. Similarly, Ha *et al.* [26] have implemented a CNN activity recognition model employing open datasets collected by tri-axial accelerometers and gyroscopes. The model yielded an accuracy of 98.29% and 97.92% with two datasets. Matsui *et al*. [36] have proposed a CNN model for outdoor activity recognition using an accelerometer, a magnetometer, and a gyroscope. They have recorded six activities (taking a train, riding a car/bus, riding a bicycle, walking, running, and keeping still) of nine people for 43 hours. The model's average accuracy was 88.2%. Wagner *et al.* [37] have used CNN to extract features from the spectrogram of the wearable inertial sensors. Then, the features were fed into a Support Vector Machine (SVM) to detect four activities (standing, walking, walking downstairs, and walking upstairs). It achieved an overall accuracy of 99.75%. However all these studies employed a camera or wearbale devices for the single person activity recognition.

Several studies [40]–[43] have used CASAS open dataset for recognizing individual or joint activities by two persons inside the house. The dataset is collected by anonymous binary sensors (PIR sensors, door sensors, water flow sensors, electricity usage sensor etc.) when two people are performing simulated activities. All studies employed HMM based classification algorithms for recognizing activities like hanging up clothes, moving furniture, reading magazine, sweeping floor, playing checkers and so on. However, these models developed for activity recognition of multiple persons.

Krishnan and Cook [44] and Cook *et al.* [45] have developed activity recognition models for the single resident using three datasets that collected in different testbeds (Bosch 1, Bosch 2, and Bosch 3) with 32 binary sensors (motion and door sensors) for about a period of half-year. All three datasets include 11 daily activities; seven of them are the same as the dataset employed in our work. They employed 3-fold cross-validation technique, and the experimental results illustrate that the highest average F_1-score for one of the three datasets was 0.61 by SVM model.

There are few studies [46]–[48] have employed the same Aruba dataset that we employed in this study for the elderly activity recognition in a smart home. Fahad *et al*. [46] have proposed an Activity Recognition approach by Clustering-based Classification (AR-CbC) that employed evidence theoretic k-nearest neighbors method. F_1-score and accuracy of the classifier were, 0.75 and 91.4%, respectively. Yala *et al*. [47], [48] have proposed three feature extraction methods combined with SVM or KNN classification models. The highest F_1-score and accuracy were 53.78 and 69.09%.

III. Proposed Methods

Fig. 1 illustrates a framework for (a) preprocessing of Aruba open dataset, (b) training, and (c) evaluation of DCNN activity recognition classifiers for the elderly people living alone. We employed CNN due to its powerful capability for extracting the features from the proposed activity image which represents the binary sensory data. In other words, CNN can extract intra-sensor patterns from the activity images to distinguish different activities.

A. Aruba Testbed and Resident

The open dataset is collected in Aruba testbed by monitoring a single elderly woman for 8 months. This testbed is one of over 40 smart home testbeds in CASAS project [49], [50], is shown in Fig. 2. This project was conducted under the ethical approval from their institutional review board.

Aruba testbed is a two bedrooms' house with a backyard, a garage, an office, and a kitchen. Totally 31 wireless motion sensors are installed around the house. Moreover, four temperature sensors and four door sensors are installed in the testbed. Though, the temperature sensors were not depicted in Fig. 2; because we only utilized the data obtained from the door sensors and the motion sensors in this work.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2833618, IEEE Journal of Biomedical and Health Informatics

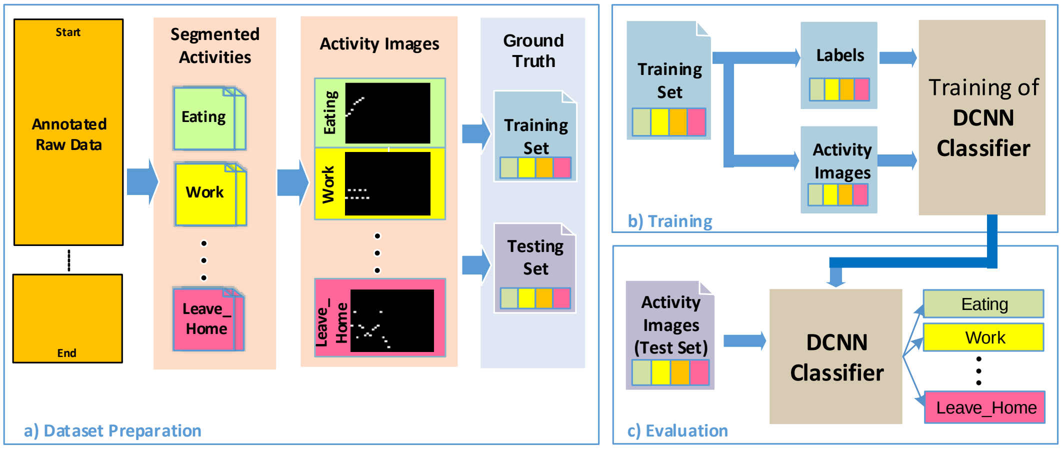


Fig. 1. Framework of preprocessing of the dataset, and training and evaluation of DCNN classifier.

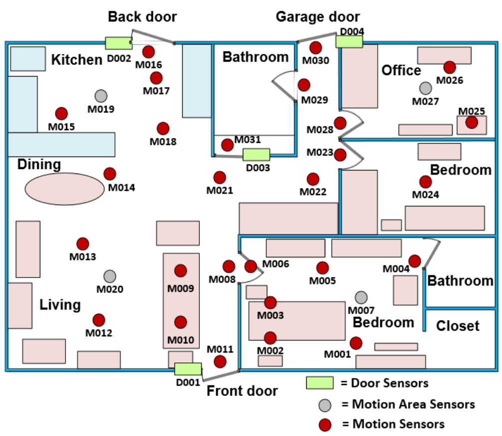


Fig. 2. Aruba testbed layout showing the positions of the door and PIR motion sensors.

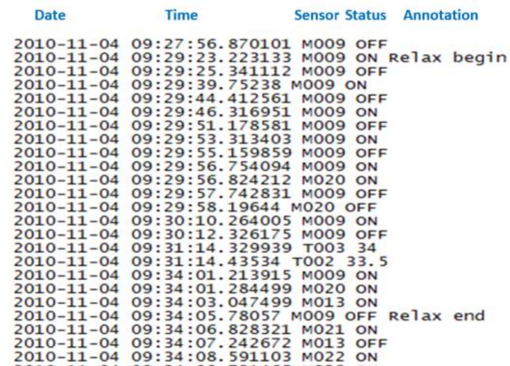
B. Binary Sensors

The binary sensors [51] employed in Aruba testbed are connected to the server through the mesh wireless network using ZigBee wireless module. The motion sensors are installed on the ceilings and the door sensors are installed on the door frames. Supposedly, the positions of the binary sensors are well selected so that the most visited/important places can be detected.

In Fig. 2, motion sensors are illustrated with the red and grey circles. The grey ones are called Motion Area Sensors; because they have a wider field of views which can detect the motion within the room. The red ones called Motion Sensors can only detect the motion under them. Similarly, door sensors are represented by green rectangles, they are the magnetic sensors that detect the opening and closing of the doors, and they can be operated on either main power or battery.

C. Dataset and Events

The dataset is a .txt file which comprised of 1719558 contiguous events recorded in 219 days during 2010-2011. The dataset has 6462 annotations of 10 different activities. Fig. 3 represents samples of Aruba raw dataset which include an annotation of Relax activity. As shown in the figure, all activity annotations have the beginning and the end, thus we can realize that this particular Relax activity is started at 9:29:23AM and finished at 9:34:05AM on Nov. 04, 2010.



```
Date        Time                Sensor Status   Annotation
2010-11-04 09:27:56.870101 M009 OFF
2010-11-04 09:29:23.223133 M009 ON Relax begin
2010-11-04 09:29:25.341112 M009 OFF
2010-11-04 09:29:39.75238 M009 ON
2010-11-04 09:29:44.412561 M009 OFF
2010-11-04 09:29:46.316951 M009 ON
2010-11-04 09:29:51.178581 M009 OFF
2010-11-04 09:29:53.313403 M009 ON
2010-11-04 09:29:55.159859 M009 OFF
2010-11-04 09:29:56.754094 M009 ON
2010-11-04 09:29:56.824212 M020 ON
2010-11-04 09:29:57.742831 M009 OFF
2010-11-04 09:29:58.19644 M020 OFF
2010-11-04 09:30:10.264005 M009 ON
2010-11-04 09:30:12.326175 M009 OFF
2010-11-04 09:31:14.329939 T003 34
2010-11-04 09:31:14.43534 T002 33.5
2010-11-04 09:34:01.213915 M009 ON
2010-11-04 09:34:01.284499 M020 ON
2010-11-04 09:34:03.047499 M013 ON
2010-11-04 09:34:05.78057 M009 OFF Relax end
2010-11-04 09:34:06.828321 M021 ON
2010-11-04 09:34:07.242672 M013 OFF
2010-11-04 09:34:08.591103 M022 ON
```

Fig. 3. Samples of Aruba raw dataset.

Events indicate any detected motion or no motion by motion sensors and detected closing or opening of the door by the door sensors. Events are received in the server by wireless mesh network and then chronologically logged in the storage. The event logs include date, time, sensor type, status, and annotation. There are 24 motion and temperature sensor events in Fig. 3.

Motion Sensors and Motion Area Sensors send "ON" status message once there is any motion in their field of view, then they send "OFF" status message when there is no motion. Door sensors send "Open" or "Close" status to the server when they sense opening or closing of the door.

D. Technical Specification

In this study, we employed a PC with i7-7700 CPU @ 3.6 GHz and a graphical processor unit (GPU) GTX 1080. GPU speeds up the computation of the DCNN classifier.

For the deep learning computation, we employed Keras which is an open source high-level neural networks Application Program Interface (API) written in Python programming language. Keras is capable of running on top of other deep learning APIs, i.e., TensorFlow, CNTK, or Theano [52].

E. Preprocessing

The preprocessing part consists of four steps i.e. (1) an activity segmentation step, (2) a sliding window step, (3) an activity image step, and (4) 10-fold training set and test set

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2833618, IEEE Journal of Biomedical and Health Informatics

Activities \ Sensors	M001	M002	M003	M004	M005	M006	M007	M008	M009	M010	M011	M012	M013	M014	M015	M016	M017
Bed_to_Toilet	0	0	0	53.5	16.8	0	29.7	0	0	0	0	0	0	0	0	0	0
Eating	0	0	0.02	0.04	0.01	0.09	0.07	0.10	0.60	0.74	0.02	0.23	1.58	74.6	2.18	0.41	1.12
Enter_Home	0	0	0	0	0	0.05	0	0	0	0	0	0	0	0.69	0	0.10	0
Housekeeping	1.74	0.15	1.21	2.63	4.70	2.93	10.8	3.93	5.72	2.96	0.67	2.37	5.03	6.69	3.62	0.82	1.97
Leave_Home	0	0	0	0	0	0.20	0	0.45	0	0	0	0.60	0.60	2.58	0	0.05	0
Meal_Preparation	0.01	0	0.01	0.05	0.06	0.11	0.13	0.26	1.05	0.40	0.04	0.38	0.94	3.86	23.3	5.81	9.63
Relax	0.04	0.02	0.03	0.06	0.08	0.11	0.13	0.20	64.9	8.29	0.03	1.95	8.05	0.96	0.22	0.15	0.16
Sleeping	0.37	19.5	58.2	0.29	0.46	0.88	5.71	0.32	0.20	0.16	0.16	0.16	0.22	0.25	0.15	0.20	0.22
Work	0.01	0	0	0.01	0.01	0	0.03	0	0	0	0	0	0.06	0.02	0.02	0	0
Wash_Dishes	0	0	0.02	0.01	0.01	0.07	0.04	0.11	0.21	0.01	0	0.08	0.45	3.48	30.8	3.69	8.53
Involved Activities	5	3	6	8	8	8	8	7	6	6	5	7	8	9	7	8	6

Activities \ Sensors	M018	M019	M020	M021	M022	M023	M024	M025	M026	M027	M028	M029	M030	M031	D001	D002	D003	D004	Involved Sensors
Bed_to_Toilet	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
Eating	3.54	3.45	4.90	0.97	0.58	0.43	3.91	0	0.06	0.03	0.11	0.14	0.02	0	0	0.01	0	0.02	29
Enter_Home	0.10	0.10	0	0.99	4.10	0.99	0.20	0	0	0	1.97	9.9	38.5	0	0	0.20	0	42.1	14
Housekeeping	4.37	5.90	13.7	4.76	2.56	1.20	5.48	0.24	0.98	0.60	0.67	0.61	0.10	0.71	0.14	0.02	0	0.02	34
Leave_Home	0.45	0.20	1.39	1.79	2.98	0.50	0	0.10	0.20	0.10	1.29	18.8	25.3	0	0	0.25	0	42.1	20
Meal_Preparation	10.8	34.0	2.58	1.34	0.86	0.56	2.79	0.09	0.14	0.09	0.12	0.19	0.18	0.04	0	0.03	0	0.13	33
Relax	0.51	0.42	5.19	0.58	0.45	0.40	5.54	0.14	0.65	0.27	0.14	0.15	0.10	0.01	0	0	0	0.07	34
Sleeping	0.26	0.22	0.42	0.35	0.40	0.79	8.48	0.16	0.16	0.16	0.21	0.34	0.24	0.15	0.03	0.03	0	0.15	34
Work	0.02	0.01	0.09	0.10	0.41	0.12	4.02	3.71	64.7	25.4	1.07	0.11	0.06	0	0	0	0	0.09	21
Wash_Dishes	10.2	37.3	1.71	0.63	0.39	0.06	1.03	0.16	0.28	0.19	0.24	0.20	0.08	0	0	0	0	0.03	28
Involved Activities	9	9	8	9	9	9	8	7	8	8	9	9	9	4	4	7	0	9	

0% … 100%

Fig. 4. Distribution map of involved sensors for each activity, occurrence of the sensors represented by light green meaning no involvement to red meaning 100%.

division step that prepare the ground truth for the classifiers from Aruba annotated raw dataset.

1) *Involved Sensors*

Fig. 4 shows the distribution map of involved sensors for each activity as a heat map, the occurrence of the sensors represented by light green (meaning no involvement) to red (meaning 100% involvement). From the heat map, we can realize that D003 was not involved in any activities, and M002 and M031, and D001 were involved only in 3 to 4 activities. On the other hand, M014, M018, M019, M021-M023, M028-M030, and D004 were involved in nine activities. Only three sensors were involved in Bed_to_Toilet; however, 34 sensors were involved in Housekeeping, Relax, and Sleeping, and 14 to 33 sensors were involved in the rest of the activities. However, all activities, except Housekeeping, have only 1 to 4 sensors that are involved in most of the time, and most of the other sensors' involvement is nearly negligible. We can also observe that Enter_Home and Leave_Home have a similar distribution over all sensors, and so as Meal_Preparation and Wash_Dishes.

2) *Activity segmentation*

Firstly, totally 6413 annotated activities were segmented as .csv files from the dataset. Temperature sensor events and 49 overlapped activities were ignored during the segmentation. Fig. 5 shows (a) the number of samples, (b) the average number of events, and (c) the maximum number of events for each activity.

Meal_Preparation and Relax have the highest numbers in terms of number of samples and the maximum number of events. However, Housekeeping has only 32 samples, but they have the highest average number of events and highest maximum number of events which are 32 and 2341, respectively. Thus, we can realize that the resident does not regularly do housekeeping; however, Housekeeping lasts longer than any other activities. Wash_Dishes has the similar features as Housekeeping does. Contrarily, Bed_to_Toilet, Enter_Home, and Leave_Home occur quite regularly than Housekeeping and Wash_Dishes; however, their average number of events are 8.4, 4.7, and 4.5, respectively, which makes them the shortest three activities among the others.

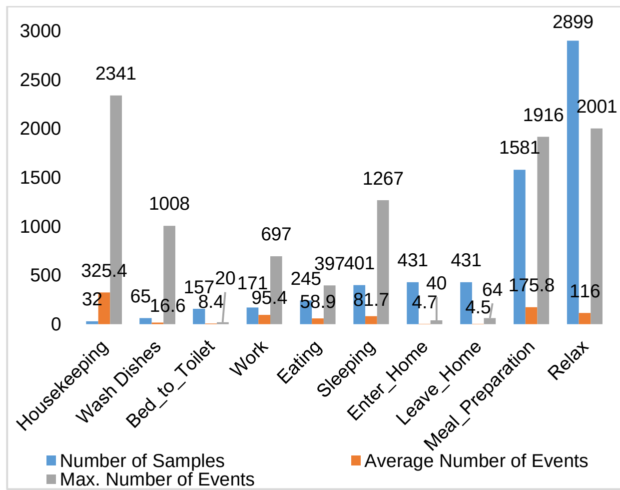


Fig. 5. Bar graph showing number of samples, average number of events, and the maximum number of events for each activity.

Moreover, we can say that the resident spent longer time on Work than Eating during the experimental period.

3) *Sliding Window*

The sliding window step divides the segmented activities into several separate .csv files. The number of separate .cv files depends on the number of events *N* of the activity, the sliding window length *L*, and the sliding step *S*-- the longer *N* and the shorter *L* and *S* the higher the number of fragmented samples. We have evaluated various values for *L* and *S* to find the best recognition rate for the classifiers and we chose *L*=250 and *S*=50 for this study. For example, if *N*=350, then there will be 3 .csv files with 250 events; however, the first, the second, and the third files will include 1-250th, 50-300th, and 100-350th events of the original file. On the other hand, if *N*<*L* there will be no divisions. Thus, short activities like Leave_Home and Bed_to_Toilet would not be sliding windowed.

After sliding window (*L*=250 and *S*=50) step, 6413 original samples are fragmented into 9281 samples. The highest numbers of samples were 2992 and 4104 for Meal_Preparation and Relax, respectively. To reduce the unbalanced sample size, we randomly chose 1000 samples from the activities with over

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2833618, IEEE Journal of Biomedical and Health Informatics

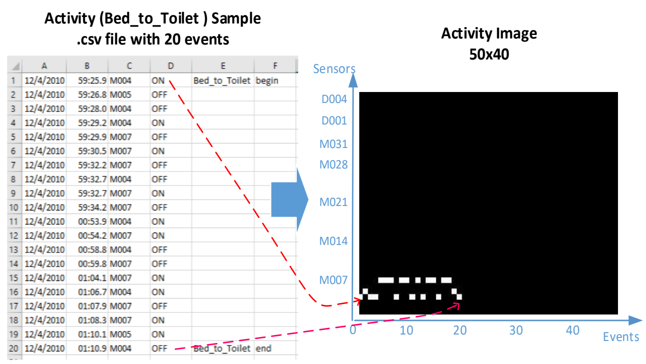


Fig. 6. Converted activity image from the segmented .csv file.

1000 samples. Thus, after the random selection, 9281 samples were reduced to 4185.

4) *Activity Image*

The activity image is a 2D visual representation of the binary sensory data collected during a particular activity made by the resident in the testbed. The activity image is a binary image with a black background and white pixels corresponding to each "ON" and "OFF" signals of the motion sensors and "OPEN" and "CLOSE" signals of the door sensors, respectively. Fig. 6 represents a sample 50×40 activity image that is converted from the segmented Bed_to_Toilet .csv file with 20 events.

Two dimensions of the activity image are x-axis (length) and y-axis (height) which represent, respectively, temporal and intra-sensor patterns of the activities. The length is equal to the sliding window length L; on the contrary, the height H is a subject to choose and it poses the intra-sensor patterns information of the activities. We observed that choosing different y coordinates for the binary sensors changes the recognition rate a lot, because it would change the intra-sensor patterns which is the main feature of the activity images to distinguish them using 2D DCNN model. For example, the recognition rate increases when we increase the difference between the y-coordinates of sensors in two different rooms while keeping the same difference between two subsequent sensors in the same room.

Table 1 indicates the y coordinates of the binary sensors that we chose in this study, because this setting gave the highest recognition rate among the different settings we tried. The motion sensors and door sensors are assigned with y coordinates alphabetically between, respectively, 20 to 100 and 110 to 140. We did leave some spaces on the top and bottom of the image for the sake of better look by human eyes. As mentioned previously, we chose closer y coordinates for closely located sensors so that the difference between y coordinates of adjacent two motion sensors is one, and farther y coordinates for the distant sensors so that the difference between y coordinates of two motion sensors in different rooms is more than six. On the contrary, the difference between y coordinates of two adjacent door sensors is ten. Since the maximum y value is 140 for D004, we chose H=150, thus the activity image size is 250×150 pixels. Fig. 7 illustrates three random samples with size of 50×40 pixels for each activity. For the sake of saving space, we show 50×40 activity images rather than 250×150 activity images, because 250×150 activity images will look 5

TABLE I.
Y-COORDINATES OF THE BINARY SENSORS.

No.	Sensor	y-coordinates	No.	Sensor	y-coordinates
1	M001	20	19	M019	54
2	M002	21	20	M020	40
3	M003	22	21	M021	60
4	M004	23	22	M022	61
5	M005	24	23	M023	71
6	M006	25	24	M024	72
7	M007	26	25	M025	80
8	M008	27	26	M026	81
9	M009	35	27	M027	82
10	M010	36	28	M028	83
11	M011	28	29	M029	90
12	M012	37	30	M030	91
13	M013	38	32	M031	100
14	M014	39	33	D001	110
15	M015	50	34	D002	120
16	M016	51	35	D003	130
17	M017	52	36	D004	140
18	M018	53			

and 3.75 times bigger in x-axis and y-axis, respectively, otherwise they have similar patterns.

5) *10-fold Training Sets and Test Sets*

We employed 10-fold cross-validation method [53] to evaluate the classifiers. Thus we divided the activity images into 10 different folds where each fold is consisting of a training set and a test set which are 9/10 and 1/10 of all activity images, respectively. Thus, each test set has different images compared to the other test sets; however, each training set shares 8/10 of all images with any of the other training sets.

F. *DCNN classifier*

Fig. 8 represents the architecture of DCNN classifier we proposed in this study which consists of three convolutional layers followed by pooling layers as a feature extraction part, and neural networks with three fully connected layer (FCL)s as a classification part. We have changed the size of the kernel (feature filter) for tuning DCNN classifier. We employed 2×2 pooling layers, thus the output size is twice the smaller than the input size.

Outputs of the last max-pooling layer are flattened and fed to the neurons of the first FCL. All neurons in the consecutive two FCLs are linked to each other. In the end, the neurons in the last layer are linked to 10 outputs.

Generated feature maps are formulated as follows [54]:

$$y_j = \frac{1}{1+\exp(b_j+\sum_i k_{ij} * x_i)} \quad (1)$$

where the convolutional operator is denoted by $*$, k_{ij} is the convolutional filter (kernel), y_j is the generated output feature map, x_i is the i-th input map, b_j is a bias.

G. *Training of DCNN Classifiers*

The training process of the DCNN models is illustrated in Fig. 1 (b), where the labels and the activity images of the training set are employed for the training. Several DCNN classifiers with different parameters are trained for the comparison. We only changed parameters of convolutional

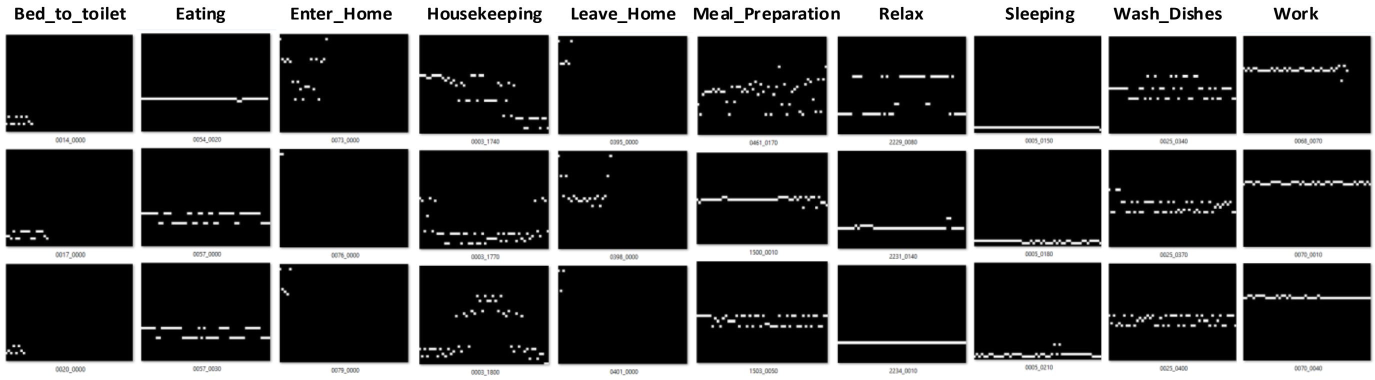


Fig. 7. Sample activity images for all activities.

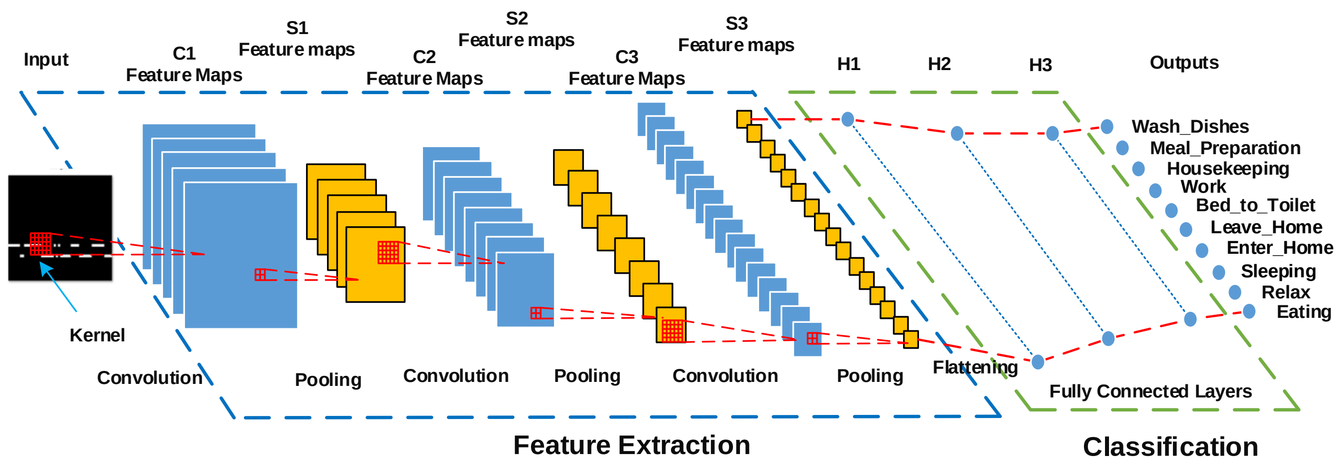


Fig. 8. The architecture of the proposed DCNN classifier.

layers; however, parameters of FCLs were constant with 254, 128, and 32 neurons for three layers.

H. Evaluation

The evaluation process is illustrated in Fig. 1 (c) where the DCNN classifier predicts the labels for the corresponding activity image. The recognition rate of the DCNN models highly depends on the model architecture, model parameters, and number of training samples, etc.

The performance of the models is measured by seven validation metrics: recall, precision, F_1-score, specificity, accuracy, error, and latency. Latency is used to evaluate the computational cost that is the time spent time during classification of an activity image.

Let's call correctly labeled "positive" images as TP (true positives), correctly labeled "negative" images as TN (true negatives), incorrectly labeled "positive" images as FP (false positives), and incorrectly labeled "negative" images as FN (false negatives). Then recall, precision, specificity, F_1-score, accuracy, and error are calculated as follows:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4}$$

$$F_1 = 2 \cdot \frac{Precision \cdot Recall}{Precesion + Recall} \tag{5}$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \cdot 100 \tag{6}$$

$$Error = 100 - Accuracy \tag{7}$$

IV. EXPERIMENTAL RESULTS

A. Recognition of 10 Activities

Table 2 represents 10-fold cross-validation results of a DCNN model with parameters 32C5×5−S2×2−64C5×5 −S2×2−128C5×5−S2×2 for all 10 activities. Each row of the table represents the averaged results of 10 different folds. Normalized confusion matrix shows how well the classifier performs on each activity, and mean values of 10 rows of seven measures are represented in the bottom of the table. The model recognizes Bed_to_Toilet activity with recall of 1.000, and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2833618, IEEE Journal of Biomedical and Health Informatics

TABLE II.
10-FOLD CROSS-VALIDATION RESULTS FOR DCNN CLASSIFIERS (10 ACTIVITIES) *

Image (250x150) 32C5×5−S2×2−64C5×5−S2×2−128C5×5−S2×2	Bed_to_Toilet	Eating	Enter_Home	Housekeeping	Leave_Home	Meal_Preparation	Relax	Sleeping	Wash_Dishes	Work	precision	recall	specificity	F_1-score	accuracy	error	latency [ms]
Bed_to_Toilet	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.994	1.000	0.9999	0.997	99.99	0.01	<10
Eating	0.000	0.895	0.000	0.004	0.000	0.083	0.018	0.000	0.000	0.000	0.981	0.896	0.9996	0.926	99.75	0.25	<10
Enter_Home	0.002	0.000	0.691	0.002	0.305	0.000	0.000	0.000	0.000	0.000	0.660	0.691	0.987	0.649	97.75	2.25	<10
Housekeeping	0.000	0.000	0.002	0.725	0.000	0.093	0.162	0.018	0.000	0.000	0.844	0.725	0.994	0.765	98.39	1.61	<10
Leave_Home	0.000	0.000	0.369	0.002	0.627	0.002	0.000	0.000	0.000	0.000	0.721	0.627	0.990	0.628	97.76	2.24	<10
Meal_Preparation	0.000	0.001	0.000	0.008	0.000	0.981	0.009	0.000	0.002	0.000	0.925	0.981	0.951	0.952	96.23	3.77	<10
Relax	0.000	0.000	0.000	0.005	0.000	0.014	0.976	0.002	0.000	0.003	0.973	0.976	0.982	0.974	98.00	2.00	<10
Sleeping	0.000	0.000	0.001	0.001	0.001	0.000	0.014	0.983	0.000	0.000	0.973	0.983	0.998	0.977	99.75	0.25	<10
Wash_Dishes	0.000	0.000	0.000	0.000	0.000	0.960	0.000	0.000	0.040	0.000	0.157	0.041	0.999	0.062	97.92	2.08	<10
Work	0.000	0.000	0.000	0.003	0.000	0.003	0.017	0.000	0.000	0.977	0.967	0.977	0.999	0.969	99.85	0.15	<10
Mean Value (all)											**0.819**	**0.790**	**0.990**	**0.790**	**98.54**	**1.46**	**<10**

*"C" and "S" denote convolutional layer and subsampling (maxpooling) layer, respectively. The architecture is described as "{the number of output maps}C{map size}-S{pooling size}.

precision, specificity, and F_1-score are nearly 1.000, and the accuracy is 99.99%. Then, Eating, Meal_Preparation, Relax, Sleeping, and Work are the next highest recognized activities after Bed_to_Toilet, and their F_1-score were 0.926, 0.952, 0.974, 0.977, and 0.969. Enter_Home, Housekeeping, and Leave_Home activities which were not well recognized as their normalized TP were 0.691, 0.725, 0.625. The worst classified activity was Wash_Dishes in which its precision, recall, and F_1-score were 0.157, 0.41, and 0.62, respectively. In addition, 96%, 8.3%, and 9.3% of, respectively, Wash_Dishes, Eating, and Housekeeping samples were misclassified as Meal_Preparation. Moreover, 30.4% and 36.9% of, respectively, Enter_Home and Leave_Home samples were mistaken for each other. Classification latency is less than 10 ms for all activities. The overall mean value of precision, recall, specificity, F_1-score, and accuracy were 0.819, 0.79, 0.99, 0.79, and 98.54%, respectively.

1) Kernel Size

Fig. 9 illustrates true positives graph of all activities for three different classifiers with different kernel sizes, true positives are interpreted by percent. Orange, yellow, and green lines are the results of 32C**5×5**−S2×2− 64C**4×4**−S2×2−128C**3×3**−S2×2, the same classifier in Table 2, and 32C**5×5**−S2×2−64C**7×7**−−S2×2−128C**9×9**−S2×2. We can see that the results of three classifiers with different kernel sizes are quite close to each other except for Enter_Home and Leave_Home activities where the difference is 25.2% and 9.9%, respectively. Performance of constant kernel size is average, decreasing kernel size was the best for Leave_Home; however, increasing kernel size was the best for Enter_Home.

2) Model Accuracy Graph

Fig. 10 illustrates a model accuracy graph during the training of DCNN classifier with 9th fold training and test set, and the other graphs for the rest of the folds are similar to it. In the graph, training accuracy increases as the epoch number increases; however, testing accuracy converges around 90% as the epoch number increases.

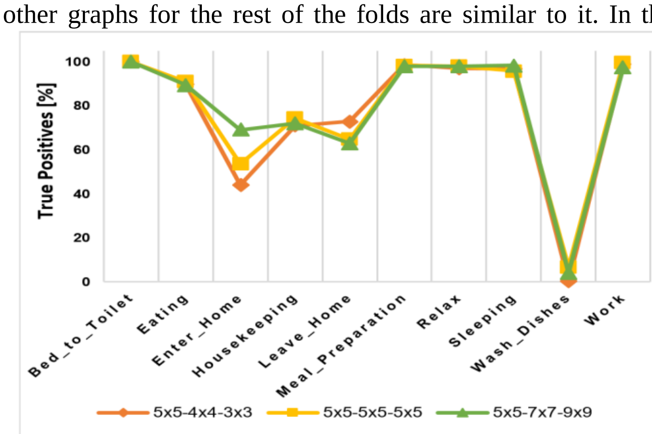


Fig. 9. True positives graph with (orange) decreasing, (yellow) constant, and (green) increasing kernel sizes.

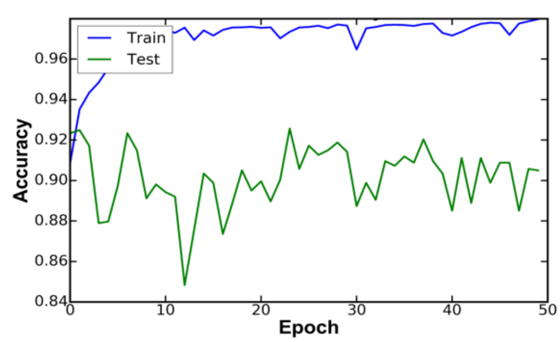


Fig. 10. Model Accuracy vs. Epoch Number graph.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2833618, IEEE Journal of Biomedical and Health Informatics

TABLE III.
10-FOLD CROSS-VALIDATION RESULTS FOR DCNN CLASSIFIERS (8 ACTIVITIES) *

Image (250x150) 32C5×5−S2×2−64C5×5−S2×2−128C5×5−S2×2	Bed_to_Toilet	Eating	Enter_Home	Housekeeping	Meal_Preparation	Relax	Sleeping	Work	precision	recall	specificity	F_1-score	accuracy	error	latency [ms]
Bed_to_Toilet	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	100.00	0.00	<10
Eating	0.000	0.925	0.000	0.004	0.060	0.011	0.000	0.000	0.982	0.925	1.000	0.946	99.80	0.20	<10
Enter_Home	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.993	1.000	1.000	0.997	99.98	0.02	<10
Housekeeping	0.000	0.004	0.000	0.756	0.131	0.107	0.002	0.000	0.840	0.756	0.994	0.784	98.46	1.54	<10
Meal_Preparation	0.000	0.001	0.000	0.002	0.993	0.004	0.000	0.000	0.969	0.993	0.978	0.981	98.40	1.6	<10
Relax	0.000	0.000	0.000	0.008	0.016	0.966	0.006	0.004	0.982	0.966	0.988	0.973	97.88	2.12	<10
Sleeping	0.000	0.000	0.004	0.022	0.000	0.013	0.961	0.000	0.963	0.961	0.997	0.961	99.53	0.47	<10
Work	0.000	0.000	0.000	0.000	0.000	0.007	0.000	0.993	0.957	0.993	0.998	0.969	99.80	0.2	<10
Mean Value (all)									**0.961**	**0.949**	**0.994**	**0.951**	**99.23**	**0.77**	**<10**

*"C" and "S" denote convolutional layer and subsampling (maxpooling) layer, respectively. The architecture is described as "{the number of output maps}C{map size}-S{pooling size}.

TABLE IV.
COMPARISON OF CONVENTIONAL MODELS AND DCNN MODEL

No.	Model	Study	F_1-score	Accuracy [%]
1	AR-CbC	[46]	0.7500	91.40
2	MkRENN	[47], [48]	0.5378	46.01
3	SVM	[47], [48]	0.4738	69.09
4	DCNN (10 activites)	Proposed	0.7900	98.54
4	DCNN (8 activites)	Proposed	0.9510	99.23

B. Recognition of 8 Activities

Table 3 represents the 10-fold cross-validation results of the same DCNN model in Table 2; however, the classifier is trained with a dataset that consists of 8 activities excluding Leave_Home and Wash_Dishes. From the table, we can see that the classifier performances are all improved than the classifier in Table 2. The classifier recognizes eight activities with more than 0.946 of F_1-score except Housekeeping (F_1-score=0.784). About 23.8% of Housekeeping samples were mistaken as Meal_Preparation and Relax. The overall mean value of precision, recall, specificity, F_1-score, and accuracy were 0.961, 0.949, 0.994, 0.951, and 99.23%, respectively.

C. Conventional Models vs. DCNN Model

Table 4 illustrates the comparison of the existing classification models and our proposed DCNN models. We can see that both the DCNN models that trained with 10 and 8 activites outperform the existing models. As mentioned in Section II, the study [44] has employed three datasets (Bosch 1-3) that have the same seven activities as the dataset (Aruba) used in this study. Graphs in Fig. 11 illustrate the F_1-scores of our proposed DCNN model on Aruba dataset and the best result performed by SVM model on Bosch datasets. DCNN model excels the conventional machine learning models on all activities except Enter_Home. The average F_1-score of DCNN model on Aruba dataset was 0.872, on the contrary, the average F_1-score of SVM model on Bosch 1, Bosch 2, and Bosch 3 datasets were, respectively, around 0.53, 0.569, and 0.51.

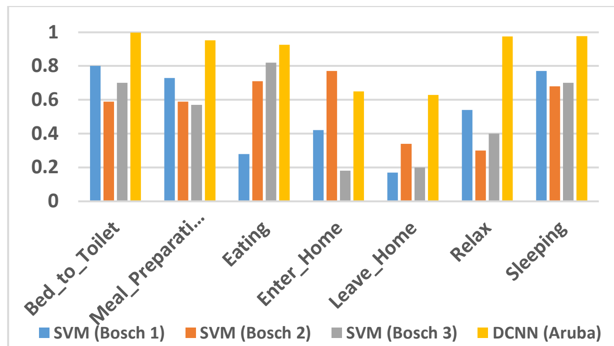


Fig. 11. Bar graph illustrating the F_1-scores of DCNN model and SVM model on different datasets.

Note that, the study [44] represents the F_1-scores of individual activities on the bar graph without the values, thus we could not calculate the exact average F_1-scores for the seven activities.

V. DISCUSSION

Proposing the activity image enabled us to exploit the powerful capabilities of 2D DCNN for extracting intra-sensor patterns of binary sensory data on different activities. Observing Fig. 11 where our dataset and previous work's datasets are acquired with the same unobtrusive binary sensors, but in a different smart house with a different single resident living inside. The result shows that the proposed DCNN model outperforms the conventional models when these seven activities are considered. Moreover, our proposed model outperforms the existing activity recognition classifiers that employed the same Aruba dataset.

In general, our model is very bad at recognizing Wash_Dishes and recall them as Meal_Preparation; because Wash_Dishes and Meal_Preparation are quite similar activities in terms of the location where both occur in the kitchen and their distribution of involved sensors are almost the same. Similarly, Enter_Home and Leave_Home have the same distribution of involved sensors, and thus the classifiers mix them with each other.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2833618, IEEE Journal of Biomedical and Health Informatics

Wash_Dishes and Meal_Preparation can be distinguished if the testbed is equipped with other binary sensors such as a water flow sensor, electricity/gas usage sensor. Because in Wash_Dishes activity residents mostly only use water; on the other hand, in the Meal_Preparation activity water and electricity/gas are used for cooking.

Moreover, Wash_Dishes, Enter_Home, and Leave_Home can be recognized without the help of extra binary sensors by combining Long Short-Term Memory (LSTM) to DCNN, since Wash_Dishes mostly occurs after Meal_preparation–Eating and Enter_Home should occur after Leave_Home.

VI. Conclusions and Future Works

In this study, an unobtrusive activity recognition DCNN model for the elderly people living alone is proposed. Aruba open dataset is compiled for the training and evaluating the proposed classifiers.

Experimental results show that, F_1-score of the best DCNN classifier among the simulated classifiers for all 10 activities and eight activities (excluding Leave_Home and Wash_Dishes) were 0.79 and 0.951, respectively. Therefore, we conclude the proposed unobtrusive activity recognition model is a useful tool for the daily life activities monitoring application of the elderly people living alone.

Our next step is to enhance the classifier by integrating LSTM model with DCNN. Moreover, instead of using open dataset only, we will apply and enhance the proposed model on the real-life long-term activity monitoring application.

VII. References

[1] M. Gochoo, "Device-Free Non-Privacy Invasive Activity Monitoring of Elderly People in A Smart House," Ph.D. Dissertation, Department of Electrical Engineering, National Taipei University of Technology, 2017.

[2] S. C. Mukhopadhyay, *Next Generation Sensors and Systems*, vol. 16. Springer, 2015.

[3] S. Häfner, J. Baumert, R. T. Emeny, M. E. Lacruz, M. Bidlingmaier, M. Reincke, H. Kuenzel, R. Holle, R. Rupprecht, and K. H. Ladwig, "To Live Alone and To Be Depressed, An Alarming Combination for The Renin–Angiotensin–Aldosterone-System (RAAS)," *Psychoneuroendocrinology*, vol. 37, no. 2, pp. 230–237, Feb. 2012.

[4] K. Kinsella, J. Beard, and R. Suzman, "Can Populations Age Better, Not Just Live Longer?," *Gener. - J. Am. Soc. Aging*, vol. 37, no. 1, pp. 19–27, 2013.

[5] K. Raworth, *State of the World 2013: Is Sustainability Still Possible?*, Washington, D.C.: Island Press/Center for Resource Economics, 2013. pp. 1-26.

[6] U. S. C. Euromonitor International, "Living Alone Statistics," 2015. [Online]. Available: http://www.statisticbrain.com/living-alone-statistics/.

[7] A. Akl, B. Taati, and A. Mihailidis, "Autonomous Unobtrusive Detection of Mild Cognitive Impairment in Older Adults," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 5, pp. 1383–1394, 2015.

[8] D. Chen, A. J. Bharucha, and H. D. Wactlar, "Intelligent Video Monitoring To Improve Safety of Older Persons," in *Proc. 29th IEEE Int. Conf. Eng. in Med. and Biol. Soc.*, 2007, pp. 3814–3817.

[9] C. Di Huang, C. Y. Wang, and J. C. Wang, "Human Action Recognition System for Elderly and Children Care Using Three Stream ConvNet," in *Proc. Int. Conf. on Orange Technologies*, 2016, no. 4, pp. 5–9.

[10] S. N. Gowda, "Human Activity Recognition Using Combinatorial Deep Belief Networks," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 1589–1594.

[11] M. Papakostas, T. Giannakopoulos, F. Makedon, and V. Karkaletsis, "Short-Term Recognition of Human Activities Using Convolutional Neural Networks," in *Proc. 12th Int. Conf. on Signal Image Technology and Internet-Based Systems*, 2017, pp. 302–307.

[12] S. Abbate, M. Avvenuti, and J. Light, "MIMS: A Minimally Invasive Monitoring Sensor Platform," *IEEE Sens. J.*, vol. 12, no. 3, pp. 677–684, 2012.

[13] N. K. Vuong, S. Chan, C. T. Lau, S. Y. W. Chan, P. L. K. Yap, and A. S. H. Chen, "Preliminary Results of Using Inertial Sensors to Detect Dementia-Related Wandering Patterns," in *Proc. 37th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2015, pp. 3703–3706.

[14] A. Kumar, C. T. Lau, S. Chan, M. Ma, and W. D. Kearns, "A Unified Grid-based Wandering Pattern Detection Algorithm," in *Proc. 38th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2016, pp. 5401–5404.

[15] K.-J. Kim, M. M. Hassan, S. Na, and E.-N. Huh, "Dementia Wandering Detection and Activity Recognition Algorithm Using Tri-Axial Accelerometer Sensors," in *Proc. 4th Int. Conf. on Ubiquitous Information Technologies & Applications*, 2009, pp. 1–5.

[16] K. Makimoto, M. Yamakawa, N. Ashida, Y. Kang, and K.-R. Shin, "Japan-Korea Joint Project on Monitoring People with Dementia," in *Proc. 11th World Congress on the Internet and Med.*, 2006, pp. 1–5.

[17] W. D. Kearns, D. Algase, D. H. Moore, and S. Ahmed, "Ultra wideband radio: A Novel Method for Measuring Wandering in Persons with Dementia," *Gerontechnology*, vol. 7, no. 1, pp. 48–57, 2008.

[18] C. Price, "Monitoring People with Dementia -- Controlling or Liberating?," *Qual. Ageing*, vol. 8, no. 3, pp. 41–44, Apr. 2007.

[19] J. Cohen-Mansfield, P. Werner, W. J. Culpepper, M. Wolfson, and E. Bickel, "Assessment of Ambulatory Behavior in Nursing Home Residents Who Pace or Wander: A Comparison of Four Commercially Available Devices," *Dement. Geriatr. Cogn. Disord.*, vol. 8, no. 6, pp. 359–365, Oct. 1997.

[20] D. De Venuto, V. F. Annese, and G. Mezzina, "Remote Neuro-Cognitive Impairment Sensing based on P300 Spatio-Temporal Monitoring," *IEEE Sens. J.*, vol. 16, no. 99, pp. 8348–8356, 2016.

[21] K. Ohnishi, A. Kanehira, A. Kanezaki, and T. Harada, "Recognizing Activities of Daily Living with a Wrist-Mounted Camera," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2016, pp. 3103–3111.

[22] T. Zebin, P. J. Scully, and K. B. Ozanyan, "Human Activity Recognition with Inertial Sensors Using a Deep Learning Approach," in *Proc. of IEEE Sensors 2016*, 2016, no. 1, pp. 1–3.

[23] M. Panwar, S. R. Dyuthi, K. C. Prakash, D. Biswas, A. Acharyya, K. Maharatna, A. Gautam, and G. R. Naik, "CNN Based Approach for Activity Recognition Using a Wrist- Worn Accelerometer," in *Proc. 39th Ann. Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, 2017, pp. 2438–2441.

[24] J. Li, R. Wu, J. Zhao, and Y. Ma, "Convolutional Neural NetWorks (CNN) for Indoor Human Activity Recognition Using Ubisense System," in *Proc. 29th Chinese Control And Decision Conference*, 2017, pp. 2068–2072.

[25] S.-M. Lee, S. M. Yoon, and H. Cho, "Human Activity Recognition From Accelerometer Data Using Convolutional Neural Network," in *Proc. IEEE Int. Conf. on Big Data and Smart Computing*, 2017, pp. 131–134.

[26] S. Ha, J.-M. Yun, and S. Choi, "Multi-modal Convolutional Neural Networks for Activity Recognition," in *Proc. IEEE Int. Conf. on Systems, Man, and Cybernetics*, 2015, pp. 3017–3022.

[27] H. H. Dodge, N. C. Mattek, D. Austin, T. L. Hayes, and J. A. Kaye, "In-Home Walking Speeds and Variability Trajectories Associated with Mild Cognitive Impairment," *Neurology*, vol. 78, no. 24, pp. 1946–1952, 2012.

[28] P. N. Dawadi, D. J. Cook, and M. Schmitter-Edgecombe, "Automated Cognitive Health Assessment Using Smart Home Monitoring of Complex Tasks," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 43, no. 6, pp. 1302–1313, Nov. 2013.

[29] B. Das, D. J. Cook, N. C. Krishnan, and M. Schmitter-Edgecombe, "One-Class Classification-Based Real-Time Activity Error Detection in Smart Homes," *IEEE J. Sel. Top. Signal Process.*, vol. 10, no. 5, pp. 914–923, 2016.

[30] J. Petersen, D. Austin, N. Mattek, and J. Kaye, "Time Out-of-Home and Cognitive, Physical, and Emotional Wellbeing of Older Adults: A Longitudinal Mixed Effects Model," *PLoS One*, vol. 10, no. 10, pp. 1–16, 2015.

[31] "Wearables Have A Dirty Little Secret: 50% of Users Lose Interest - TechRepublic." [Online]. Available: http://www.techrepublic.com/article/wearables-have-a-dirty-little-secret-most-people-lose-interest/. [Accessed: 16-Dec-2016].

[32] D. J. Cook, "Learning Setting-Generalized Activity Models for Smart Spaces," *IEEE Intell. Syst.*, vol. 27, no. 1, pp. 32–38, 2012.

[33] M. Wang, C. Luo, B. Ni, J. Yuan, J. Wang, and S. Yan, "First-Person Daily Activity Recognition with Manipulated Object Proposals and Non-linear Feature Fusion," *IEEE Trans. Circuits Syst. Video Technol.*, vol.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2833618, IEEE Journal of Biomedical and Health Informatics

8215, no. c, pp. 1–21, 2017.

[34] K. Adhikari, H. Bouchachia, and H. Nait-Charif, "Activity Recognition for Indoor Fall Detection Using Convolutional Neural Network," in *Proc. 15th Int. Conf. on Machine Vision Applications*, 2017, pp. 81–84.

[35] F. J. Ordóñez and D. Roggen, "Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition," *Sensors*, vol. 16, no. 1, 2016.

[36] S. Matsui, N. Inoue, Y. Akagi, G. Nagino, and K. Shinoda, "User Adaptation of Convolutional Neural Network for Human Activity Recognition," in *Proc. 25th European Signal Processing Conference*, 2017, pp. 783–787.

[37] D. Wagner, K. Kalischewski, Jorg Velten, and A. Kummert, "Activity Recognition Using Inertial Sensors and A 2-D Convolutional Neural Network," in *Proc. 10th Int. Workshop on Multidimensional Systems*, 2017, pp. 1-6.

[38] Y. Chen and Y. Xue, "A Deep Learning Approach to Human Activity Recognition Based on Single Accelerometer," in *Proc. IEEE Int. Conf. on Systems, Man, and Cybernetics*, 2015, pp. 1488–1492.

[39] G. Liu, J. Liang, G. Lan, Q. Hao, and M. Chen, "Convolution Neutral Network Enhanced Binary Sensor Network for Human Activity Recognition," in *Proc. IEEE Sensors*, 2017, pp. 1–3.

[40] G. Singla, D. J. Cook, and M. Schmitter-Edgecombe, "Recognizing Independent and Joint Activities Among Multiple Residents in Smart Environments," *J. Ambient Intell. Humaniz. Computer*, vol. 1, no. 1, pp. 57–63, 2010.

[41] R. Chen and Y. Tong, "A Two-Stage Method for Solving Multi-Resident Activity Recognition in Smart Environments," *Entropy*, vol. 16, no. 4, pp. 2184–2203, 2014.

[42] C. Yi-Ting, K. C. Hsu, C. H. Lu, F. Li-Chen, and H. Jane Yung-Jen, "Interaction Models for Multiple-Resident Activity Recognition in a Smart home," in *Proc. Intelligent Robots and Systems*, 2010, pp. 3753–3758.

[43] C.-H. Lu and Y.-T. Chiang, "Interaction-Feature Enhanced Multiuser Model Learning for A Home Environment Using Ambient Sensors," *Int. J. Intell. Syst.*, vol. 29, no. 11, pp. 1015–1046, Nov. 2014.

[44] N. C. Krishnan and D. J. Cook, "Activity Rrecognition on Streaming Sensor Data," *Pervasive Mob. Comput.*, vol. 10, no. Part B, pp. 138–154, 2014.

[45] D. J. Cook, N. C. Krishnan, and P. Rashidi, "Activity Discovery and Activity Recognition: A New Partnership," *IEEE Trans. on Cybern.*, vol. 43, no. 3, pp. 820–828, 2013.

[46] L. G. Fahad, S. F. Tahir, and M. Rajarajan, "Activity Recognition in Smart Homes Using Clustering Based Classification," in *Proc. 22nd Int. Conf. Pattern Recognit.*, 2014, pp. 1348–1353.

[47] N. Yala, B. Fergani, and A. Fleury, "Feature extraction for human activity recognition on streaming data," in *Proc. Int. Symp. Innov. Intell. Syst. Appl.*, 2015, pp. 1–6.

[48] N. Yala, B. Fergani, and A. Fleury, "Towards improving feature extraction and classification for activity recognition on streaming data," *J. Ambient Intell. Humaniz. Comput.*, vol. 8, no. 2, pp. 177–189, 2017.

[49] D. J. Cook, A. S. Crandall, B. L. Thomas, and N. C. Krishnan, "CASAS: A Smart Home in a Box," *Computer*, vol. 46, no. 7, pp. 62–69, Jul. 2013.

[50] T. H. Tan, M. Gochoo, F. R. Jean, S. C. Huang, and S. Y. Kuo, "Front-Door Event Classification Algorithm for Elderly People Living Alone in Smart House Using Wireless Binary Sensors," *IEEE Access*, vol. 5, pp. 10734–10743, 2017.

[51] B. Minor and D. J. Cook, "Forecasting Occurrences of Activities," *Pervasive Mob. Comput.*, vol. 38, pp. 77–91, Jul. 2017.

[52] "Keras Documentation." [Online]. Available: https://keras.io/. [Accessed: 14-Nov-2017].

[53] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in *Proc. Int. Joint Conf. on Articial Intelligence*, 1995, vol. 5, pp. 1–7.

[54] M. Gochoo, T.-H. Tan, V. Velusamy, S.-H. Liu, D. Bayanduuren, and S.-C. Huang, "Device-Free Non-Privacy Invasive Classification of Elderly Travel Patterns in A Smart House Using PIR Sensors and DCNN," *IEEE Sens. J.*, vol. 18, no. 1, pp. 390–400, 2018.

[55] M. Gochoo, T.-H. Tan, S.-H. Liu, S.-C. Huang, and F. S. Alnajjar, "DCNN-Based Elderly Activity Recognition Using Binary Sensors," in *Proc. Int. Conf. on Electrical and Computing Technologies and Applications*, 2017, pp. 1–5.

Munkhjargal Gochoo was born in Ulaanbaatar City, Mongolia, in 1984. He received B.S. and M.S. degrees in Electronics Engineering from the Mongolian University of Science and Technology in 2004 and 2005, respectively. He stayed at Department of Electronics, Mongolian University of Science and Technology as a Lecturer during 2005-2011. In 2017, he got his Ph.D. degree from National Taipei University of Technology in electrical engineering field. Currently, he is a Research Assistant Professor at Department of Electronic Engineering, National Taipei University of Technology, Taipei, Taiwan. Dr. Gochoo is a member of the IEEE and his main research interests are telecare, eldercare, IoT, wearable systems, machine learning, image processing, and deep learning classification algorithms.

Tan-Hsu Tan received B.S. degree from National Taiwan Institute of Technology and M.S. degree from National Tsing Hua University, in 1983 and 1988, respectively, both in Electrical Engineering, and Ph.D. in Electronics Engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1998. Since 1988, he has been with Department of Electrical Engineering at the National Taipei University of Technology, Taipei, Taiwan, where he is currently a Full Professor. His research interests are wireless communications, telecare, machine learning, and optimization algorithms.

Shing-Hong Liu received B.S. degree in electronic engineering from Feng-Jia University, Taichung, Taiwan, R.O.C., in 1990, M.S. degree in biomedical engineering from the National Cheng Kung University, Tainan, Taiwan, R.O.C., in 1992, and Ph.D. degree from the Department of Electrical and Control Engineering, National Chiao-Tung University, Hsinchu, Taiwan, R.O.C., in 2002. August 1994, he has been a Lecturer in Department of Biomedical Engineering, Yuanpei University, Hsinchu, Taiwan, and was an Associate Professor from 2002 to 2008. Now, he is a Professor in Department of Computer Science and Information Engineering, Chaoyang University of Technology. His current research interests are digital signal processing, fuzzy control and designing biomedical instrument.

Fu-Rong Jean received the B.S. degree from the National Taiwan University of Science and Technology, Taipei, Taiwan, in 1985 and the M.S. and Ph.D. degrees in electrical engineering from the National Tsing Hua University, Hsinchu, Taiwan, in 1987 and 1995, respectively. From 1987 to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2833618, IEEE Journal of Biomedical and Health Informatics

1995, he was a lecturer with the Department of Electrical Engineering, National Taipei Institute of Technology, Taipei. Since 1995, he has been an Associate Professor with the Department of Electrical Engineering, National Taipei University of Technology, Taipei. His main research interests include speech enhancement, speech and audio coding, and distributed speech recognition.

Fady S. Alnajjar received his MSc in Artificial Intelligence and his PhD in System Design Engineering at the University of Fukui, Japan in 2007 and 2010, respectively. In 2010, Dr. Fady worked as a research scientist at Brain Science Institute (BSI), RIKEN, Japan. He conducted neuro-robotics study to understand the underlying mechanisms for embodied cognition and mind. From 2012, he started an interest in exploring the neural mechanisms of motor learning, adaptation, and recovery after brain injury from the sensory- and muscle-synergies perspectives. His research target is to propose an advance neuro-rehabilitation application for patients with brain injuries. Since 2016, he joined a three-year joint research group, including Nagoya University-Japan, and the University of Michigan-USA, to develop a bio-inspired prosthetic arm. Currently Dr. fady is an Assistant Professor in the department of Computer Science and Software Eng-(CIT) in United Arab Emirates University (UAEU), and he is the founder of Intelligent Robot Interactive laboratory (IRI Lab).

Shih-Chia Huang (M’09-SM’14) is a Full Professor with the Department of Electronic Engineering at National Taipei University of Technology, Taiwan, and an International Adjunct Professor with the Faculty of Business and Information Technology at the University of Ontario Institute of Technology, Canada. He has been named a senior member of the Institute of Electrical and Electronic Engineers (IEEE). He is currently the Chair of the IEEE Taipei Section Broadcast Technology Society, and was a Review Panel Member of the Small Business Innovation Research (SBIR) program for the Department of Economic Development of Taipei City and New Taipei City, respectively. Professor Huang has published more than 80 journal and conference papers and holds more than 60 patents in the United States, Europe, Taiwan, and China. Dr. Huang received B.S. and M.S. degrees from National Taiwan Normal University and National Chiao Tung University, respectively. In 2009, Dr. Huang received a doctorate degree in Electrical Engineering from National Taiwan University, Taiwan. He was presented with the Kwoh-Ting Li Young Researcher Award in 2011 by the Taipei Chapter of the Association for Computing Machinery, the 5th National Industrial Innovation Award in 2017 by the Ministry of Economic Affairs, Taiwan, as well as the Dr. Shechtman Young Researcher Award in 2012 by National Taipei University of Technology. Professor Huang was the recipient of an Outstanding Research Award from National Taipei University of Technology in 2014 and 2017, and the College of Electrical Engineering and Computer Science, National Taipei University of Technology in 2014-2016. In addition, he has been an associate editor of the Journal of Artificial Intelligence and a guest editor of the Information Systems Frontiers and the International Journal of Web Services Research. He is also the Services and Applications Track Chair of the IEEE CloudCom 2016-2017 conference, the Applications Track Chair of the IEEE BigData Congress in 2015, General Chair of the 2015-2016 IEEE BigData Taipei Satellite Session, and the Deep learning, Ubiquitous and Toy Computing Minitrack Chair of the 2017-2018 Hawaii International Conference on System Sciences. His research interests include intelligent multimedia systems, image processing and video coding, video surveillance systems, cloud computing and big data analytics, artificial intelligence, and mobile applications and systems.