Al-Ahgaff University Journal of Computer Science and Mathematics
مجلة جامعة الأحقاف لعلوم الحاسوب والرياضيات

Vol. 1, Issue. 1, Jan 2023, pp. 25~33

Collecting Complex Human Activity Dataset

Mohammed Mobark Wahdeen
Information Technology Department, Al-Ahgaff University, Mukalla, Yemen
mabomobark@yahoo.com

Article Info

Article history:

Received jan, 9, 2023
Revised jan,23, 2023
Accepted jan 25 2023

Keywords:

Complex Human Activity Dataset
Third keyword

ABSTRACT

Existing human activity datasets involve simple activities or were collected using standalone sensors. So they do not properly match the requirement to evaluate the classifiers of the complex activities that were collected using smartphone sensors. The author collected a dataset (i.e., the Complex Activity Dataset (CAD)) to solve this problem. A group of 20 subjects was selected for this task. Data was collected in the scenario where the subject prepares breakfast. The subject performs three complex activities: preparing breakfast, preparing tea, and preparing a sandwich. Those activities are categorized into two levels in the form of a hierarchy, so that the complex activities would be placed at the high level and the simple activities at the low level. CAD was collected using the accelerometer and gyroscope sensors of smartphones. This paper presents the protocol for collecting, labeling, and filtering CAD. Also, this paper evaluated the variation property of a CAD dataset and the ability to recognize its complex human activities. The result supports the variation property of the CAD dataset and presents the ability to recognize its activities with greater accuracy than other datasets.

الخلاصة

بيانات النشاط البشري المجمعة الحالية تتضمن أنشطة بسيطة أو تم جمعها باستخدام أجهزة استشعار قائمة بذاتها. لذلك فهي لا تتطابق مع متطلبات تقييم المصنفات للأنشطة المعقدة التي تم جمعها باستخدام مستشعرات الهواتف الذكية. جمع المؤلف مجموعة بيانات (CAD) لحل هذه المشكلة. تم إختيار مجموعة من ٢٠ شخصا لهذه المهمة. حيث يقوم الشخص بإعداد وجبة الإفطار. و الذي يشمل القيام بثلاثة أنشطة معقدة: تحضير الإفطار ، تحضير الشاي ، تحضير الساندويتش. يتم تصنيف هذه الأنشطة إلى مستويين في شكل تسلسل هرمي ، بحيث يتم وضع الأنشطة المعقدة على مستوى عالٍ والأنشطة البسيطة على المستوى المنخفض. تم جمع CAD باستخدام مستشعرات التسارع والجيروسكوب للهواتف الذكية. تقدم هذه الورقة بروتوكول لتجميع وتصنيف وفلترة CAD. أيضًا ، قيمت هذه الورقة خاصية التباين لمجموعة بيانات CAD والقدرة على التعرف على أنشطتها البشرية المعقدة. تدعم نتيجة هذا البحث خاصية التباين لمجموعة بيانات CAD وتوفر القدرة على إمكانية التعرف على أنشطتها بدقة أكبر من مجموعات البيانات الأخرى.

1. INTRODUCTION

Using mobile phones for Human Activities Recognition (HAR) is very helpful in observing the daily habits of the user and detecting health diseases or accidents early. In real-world situations, human activities are often performed in complex ways. Complex human activities are composite activities that occur concurrently or interleave. In those activities, the existence and variations of each activity, as well as the order and length, may vary. In this research, the composition and variations of human activity were considered as factors that impact the complexity of human activities.

Existing human activity datasets involve simple activities or were collected using standalone sensors. So they do not properly match the requirement to evaluate the classifiers of the complex activities that were collected using smartphone sensors. For example, the Opportunity [1] and UCI-HAR [2] datasets, probably the two most popular, are cases in point. The Opportunity dataset, which contains complex activities, was performed by 12 subjects. But the activities were collected using the inertial measurement unit, which contains a standalone accelerometer and gyroscope.

On the other hand, the UCI-HAR dataset contains inertial data collected from 30 subjects who performed a set of common daily activities while carrying a smartphone. It provides data collected from the smartphone's accelerometer and gyroscope. But the subjects performed simple activities such as walking (straight, upstairs, downstairs), sitting, standing, and lying down. The author collected CAD to solve this problem. The following sections present the protocol for collecting, labeling, and filtering CAD. Also they show the experiments to evaluate the variation property of CAD dataset and the ability to recognition its complex human activities.

2. METHODOLOGY

CAD was collected using an Android (Samsung SM-G935F) that is commonly utilized [3]. It was attached to the upper arm and forearm of the subject. This smartphone contains a tri-axial accelerometer and a gyroscope derived from the findings of [4]. We want to see the effect of using the smartphone gyroscope sensor in addition to the accelerometer to smooth the data and increase recognition accuracy. Smartphone sensors collected data about 30 subjects who perfomed complex activities.

In addition to the requirement that the human activities should be collected using mobile phone sensors, the collected activitiy should be complex. There are several factors that make recognising complex human activities a challenge. In this research, the composition and variations of human activity were examined as factors that impact the complexity of the human activity recognition. For example, the complex activity comprises more than one activity that might be performed in changing order such as in an interleave or parallel manner (composition property). However, the activities should be in particular structures and sequences to be recognized by current recognition methods [5]. Also, several factors can affect the performance of the activity such as physical body differences or the environmental state in which the activity is performed. Hence, the same activity may be performed differently by different subjects (variation property).

In our own dataset, the complex activities are organized into levels to reflect the composition property of human activities whereby the high level consists of complex activities such as making sandwiches or preparing tea. Meanwhile, the low level contains meaningful, elementary (basic) movements of a person's body parts to perform the complex activity, for instance, stretching an arm or raising a leg. Table 1 shows the levels of activities in our own dataset.

Table 1 Hierarchal labelling of complex activity

High level activities	Medium level activities	Low level activities
Preparing Tea	Get boiling water	Shoulder Extension
	Add tea	Shoulder Internal Rotation
	Mix the tea	Wrist rotation

2.1. Data collection setup

The selected smartphone to conduct our experiment was an Android (Samsung SM-G935F) that is commonly utilized [3]. Table 2 shows the specifications of the smartphone used. It was attached to the upper arm of the subject. The smartphone contains a tri-axial accelerometer and a gyroscope derived from the findings of [4]. The sensors record timestamp motion data at the "fastest" sampling rate which can reach a maximum of 80 Hz [6]. The selected sampling rate for acquiring the body movement is contained within frequency components below 20 Hz as recommended by [3]. It is equipped with a SensorDataCollector program for collecting subject data and for storing it in a log file at SD card.

Table 2 Specifications of the smartphone used

Device	Smartphone
Brand	Samsung SM-G935F
CPU	Exynos 8890 8 Cores (Octa-Core)
ROM Memory	64GB
RAM Memory	4GB
Operating System	Android v6.0.1 (Marshmallow)

Device	Smartphone
Accelerometer And Gyroscope Make Model Power Rang Resolution	Sample rate is set to fastest which can reach a maximum of 80 Hz. STM. K6DS3TR. 0.2500 mA. 8.0 g. 0.002394 m/s^2.
Mobile Network Type Battery Capacity	HSUPA (High-Speed Uplink Packet Access). 3600mAh.

2.2. Data collection protocol

The experiment was carried out to obtain the HAR datasets. A group of 20 subjects were selected for this task based on the findings of [7]. Data of the subjects is presented in Table 3. Each subject was instructed to follow a protocol of activities while carrying the selected smartphone in his upper arm to infer overall body motion. Data was collected in the scenario where the subject prepares breakfast. This scenario has been used extensively in other works in literature [8]. The subject performs three complex activities: preparing breakfast, preparing tea, and preparing a sandwich. Those activities are categorized into two levels in the form of a hierarchy so that the complex activities would be placed in the high level and the simple activities in the low level [9]. The details of the low level activities in each complex activity are as follows:

A. Preparing:
Lying down on the deckchair.
Getting up.
Retrieving bread, cheese, cup, tea, sugar, plate, spoon, and knife from the cupboard and putting them on the cooking table.
B. Preparing Tea (Pre. Tea):
Getting a cup of water from the water boiler machine.
Adding the tea and sugar.
Mixing the tea.
Putting the cup on the dining table.
C. Preparing Sandwich (Pre. Sandwich):
Making bread and cheese sandwich at the cooking table.
Heating it in the microwave.
Putting it on the dining table.

Each activity lasted a minute and was repeated twice for each subject. The duration of the entire experiment was around 15 minutes per person excluding the setting up of the sensors and the repetition of the protocol. The collected human activities were designed to closely represent the natural world in both the style and time of action classes executed. The subject is free to perform the sequence of activities, so we will get activities with wide range of variations. Also, there was no time limitation on the execution of each task, so some tasks took naturally longer than others. Table 4 shows the proportion of classes in our dataset.

Table 3 Data of the subjects

Sex	Male (11)+female (9)
Age	24-49 years
Average length of upper arm	29.28 cm

Average length of Forearm	23.11 cm

Table 4 Proportions of our dataset classes

Class	Instances	Proportion
Preparing	1383	19%
Preparing Tea	2300	31%
Preparing Sandwich	3648	50%
Total	7331	100%

2.3. Data Labeling

Once data was collected from the experiment, the log files were filled in order to generate the HAR datasets. Firstly, smartphone and video signals were synchronized manually by specifying the start and end of the basic movements and complex activities. All the labels of the experiments were collected in a file (labels file) which was used as one of the inputs for the dataset generation process.

2.4. Signal filtering

Most of the time, raw sensor signals from the accelerometer and the gyroscope are noise, so they should be preprocessed by a series of filters. We used the following filters to utilize the best performance of each sensor after carefully examining the sensor's dynamic models:

- Low pass filter that only allows signals with lower frequencies than certain cutoff frequencies. It was used to extract the low frequency of the accelerometer. The break frequency of low-pass filter was chosen at 2.5 rad/s.
- High pass filter that only allows signals with higher frequencies than certain cutoff frequencies. It was used to extract the high frequency of gyroscope. The break frequency of high pass filter was selected at 3.3 rad/s.

2.5. Time window size

The labeled and preprocessed signals are segmented into time window samples. Every window supposedly has an associated activity. We used fixed-width sliding windows and assigned them into activity label with 50% overlap between windows. The overlap avoids any missing activity data that begins during the time window and continues into the next one when splitting the data into segments. We evaluated the set of sizes {1, 2, 4, and 8} seconds following the recommendations of [10]. We chose a one second time window to segment our data. This decision is based on preliminary experiments which showed that using one second time window yields the best recognition accuracy.

3. PERFORMANCE EVALUATION

In addition, this research evaluated the variation property of CAD dataset and the ability to recognition its complex human activities. For this purpose, the following two experiments were conducted:

3.1. Experiment I –Evaluating the variation property of CAD

The experiments were conducted to check the variation property of the collected human activities in our own dataset. Different subjects were chosen to perform three selected tasks namely boiling water, adding tea, and mixing the tea. These tasks were chosen because each task is represented by basic movements of different arm joints as shown in Table 5. The F-test measure of analysis of variance (ANOVA) was used to investigate the effect of the variations of those task and subject factors in the wrist velocity. The two factors were tested and verified statistically by 20 x 3 (subject x task) ANOVA analysis at probability levels ($p<0,05$).

Table 5 Hierarchical labelling of complex activity

High level activities	Medium level activities	Low level activities
Preparing Tea	Get boiling water	Shoulder Extension

Add tea	Shoulder Internal Rotation
Mix the tea	Wrist rotation

The result of the F-test for one task performed by four subjects is presented in Table 6. Meanwhile, the result of the F-test for the three tasks performed by the same subject is shown in Table 7. Table 6 presents the result of the F-test for the (get boiling water) task performed by four subjects. It shows that the variations between the subjects (3.82E+09) who performed the same task are greater than the variations inside the task (2.93E+08). Also, it displays that the P -value (1.74E-8) is more than 0,05 showing that there is no significant difference in the four subjects when performing the same task.

Table 6 F-test for one task performed by four subjects

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	1.15E+10	3	3.82E+09	13.05485	1.74E-8	2.606929
Within Groups	1.29E+12	4396	2.93E+08			
Total	1.3E+12	4399				

Table 7 presents the result of the F-test for the three tasks performed by the same subject. It shows that the variations between the tasks (2.07E+12) are greater than the variations inside the same task (1.21E+08). Also, it displays that the P-value (0) is less than 0,05 showing that there is a significant difference in the three tasks when performed by the same subject.

Table 7 F-test for three tasks performed by the same subject

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	4.15E+12	2	2.07E+12	17127.18	0	2.998456
Within Groups	3.99E+11	3297	1.21E+08			
Total	4.55E+12	3299				

The result supports the variation property of the collected human activities in our own dataset. There are variations in the performed task more than the variations in the subject.

3.2. Experiment II –Evaluating the ability to recognize the complex activities of CAD

In addition, experiments were conducted to check how far we could recognize the complex human activities collected in our datasets. Our recognition system, the Complex Activity Recognizer through Wrist Velocity (CARWV) [14], was used to recognize the complex human activities that were collected using the accelerometer and gyroscope of smartphones in CAD. The result was compared, to check the ability to recognize them, with the Oppurnity dataset [1] that was collected using the Inertial Measurement Unit that contains a standalone accelerometer and gyroscope.

3.2.1 Opportunity dataset

We used the Opportunity dataset to test our system's capability (CARWV) in generalizing the recognized complex human activities that were collected using standalone sensors. The Opportunity dataset was collected from four subjects who performed 17 different Activities of Daily Life (ADLs). In our experiments, we chose the scenario that consists of four high level activities i.e. Early morning moving, Coffee time, Sandwich time, and Cleanup. These high level activities and their low level activities are further described below.

A. Early morning moving:

Getting up

Opening the door

Closing the door

Walking

B. Coffee time:

Sipping

Opening drawer

Closing drawer

Reaching for an item

C. Sandwich time:

Slicing

Opening the fridge

Closing the fridge

Reaching for an item

D. Cleanup:

Opening dish washer

Closing dish washer

Reaching for an item

Moving item

Releasing item

Wiping

The Opportunity dataset includes a high number of instances of different gestures recorded by a high number of on-body, environmental and object-attached sensors at a sampling frequency of 30Hz. In our experiments, we chose the Inertial Measurement Unit that contains standalone accelerometer and gyroscope to evaluate our system. It was placed at the Right Upper Arm (RUA) of the subjects.

3.2.2 Experiment setup

To evaluate our system (CARWV), a 5-fold cross validation method was used. In 5-fold cross validations, the dataset is randomly divided into 5 groups (folds) of equal sizes. Each time, one fold is taken as a testing set whilst the remaining is used for training our system. This process is repeated 5 times before arriving at the final performance by taking the average of test errors that resulted from each step. The K folds cross validation system incurs less computational cost compared to other validation systems.

The recognition performance of our system was measured by two performance metrics: accuracy and F1 measure. We used the recognition accuracy metric to measure the performance of our system because it is a popular measure in the literature of human activity recognition ([11], [12], and [9]). But recognition accuracy is affected by imbalanced classes in the dataset, so we also used the F1 measure that is independent of the class distribution and measures the effect of false negatives and false positives. These two metrics (i.e. accuracy and F1 measure) had been used in previous works ([5], [13], and [9]) which makes the comparison easier. The F1 measure is the mean of precision and recall metrics. The evaluation process was simulated using MATLAB R2018b on a notebook computer with Intel i7-7700K CPU and 8GM RAM.

3.2.3 Experiment result

Firstly, this experiment tests the capabilities of our system to recognize complex human activities in our own dataset. For this purpose, the 5-fold cross validation method was used. The results are shown in Table 8 which shows the confusion matrix of complex activities in our own dataset when applying our system using the 5-fold cross validation. The total recognition accuracy of applying the proposed system to recognize complex activities in our own dataset is 86.2 percent and all classes obtained more than 55 percent.

The complex activity of Preparing Sandwich consists of three simple activities (i.e. Making bread and cheese sandwich at the cooking table, Heating it in the microwave, Putting it on the dining table) in which each one involves a number of basic arm movements. It obtained the highest recognition accuracy (93 percent) with little confusion with other classes. The misclassification of Preparing Sandwich activity with other activities (42 percent with Preparing Tea and 20 percent with Preparing activity) may be because it is a dominant class that shares similar abduction and adduction basic arm motions to reach items such as the bread, cheese, tea, or cupboard.

The other complex activity in our dataset is (Preparing) which consists of three simple activities (i.e. Lying down on the deckchair, Getting up and Retrieving bread, cheese, cup, tea, sugar, plate, spoon, and knife from the cupboard, and Putting them on the cooking table). Each simple activity involves a set of basic arm movement. The CARWV recognized it with a 73 percent accuracy with 20 percent confusion for the Preparing Sandwich activity. The lowest percentage of recognition accuracy was for the Preparing Tea activity with 55 percent and 42 percent confusion for the Preparing Sandwich activity. The complex activity (Preparing Tea) consists of four simple activities with basic arm movements (i.e. Getting a cup of water from the water boiler machine, Adding the tea and sugar, Mixing the tea, and Putting the cup on the dining table). The results show the ability of our system to recognize complex human activities with relatively high recognition accuracy (92-55 percent) in our own dataset.

Table 8 Confusion matrix of our dataset classes

Classes of Complex Activity dataset	Cross validation (Folds=5, Total accurate=86.2%)		
	Preparing	Preparing Tea	Preparing Sandwich
Preparing	**73**	7	20
Preparing Tea	3	**55**	42
Preparing Sandwich	2	5	**93**

Table 9 Confusion matrix of Opportunity classes

Classes of Opportunity	Cross validation (Folds=5, Total accurate =87%)			
	Coffee time	Cleanup	Sandwich time	Early morning moving
Coffee time	78	0	0	22
Cleanup	0	58	2	40
Sandwich time	0	2	48	50

Early morning moving	2	2	4	92

Secondly, Table 9 shows the confusion matrix of classes in the Opportunity dataset when the 5-fold cross validation system was applied. The total recognition accuracy of applying the proposed system to recognize complex activities in the Opportunity dataset is 87 percent compared to the total recognition accuracy in our own dataset. It was noticed that the total recognition accuracy of our system in the Opportunity dataset (87 percent) is better than that in our own dataset (86.2 percent). This might be in part due to the use of standalone sensors to collect the Opportunity dataset instead of using smartphone sensors as how it was carried out in our own dataset. Dernbach et al. (2012) noted that the capabilities of standalone sensors are better than the ones used in smartphones for acquiring data.

All classes obtained more than 48 percent. The Early Morning Moving complex activity obtained the highest recognition accuracy with 92 percent. It consists of four simple activities (i.e. Getting up, Opening the door, Closing the door, and Walking) in which each one involves a number of basic arm movements. The misclassification of the Early Morning Moving activity with other activities (50 percent with Sandwich time, 40 percent with Cleanup, and 22 percent with Coffee time activity) may be because it is a dominant class that shares similar flexion and extension basic arm movements when the arm swings during walking, opening doors, cutting bread, wiping, and moving hand near mouth to sip coffee.

The next activity was Coffee time which obtained 78 percent with a 22 percent confusion with Early Morning Moving activity. It consists of four simple activities (i.e. Sipping, Opening drawer, Closing drawer, Reaching for an item). Each simple activity involves a set of basic arm movements. Following that is the complex activity of Cleanup which consists of six simple activities with their basic arm movements (i.e. Opening dish washer, Closing dish washer, Reaching for an item, Moving item, Releasing item, and Wiping). It received a 58 percent accuracy with a 40 percent confusion with Early Morning Moving activity. The lowest percentage was 48 percent for Sandwich time activity with 50 percent confusion with Early Morning Moving activity with 44 percent difference between its accuracy and the one of the best performing class. This activity consists of four simple activities (i.e. Slicing, Opening the fridge, Closing the fridge, and Reaching for an item) in which each one involves a number of basic arm movements. The results of this experiment show the ability of our system in recognizing complex human activities with relatively high recognition accuracy on the two datasets.

The hierarchical structure of activities and result of F-measure of Anova in our own dataset shows the sufficiency of our dataset in representing the required two factors for evaluating the complex activities which are the ability to recognize complex, with variation human activities which were collected using mobile phone sensors. The experiments show also the ability to recognize its activities with accuracy more than other datasets.

4. CONCLUSION

The protocol for collecting, labeling, and filtering CAD is presented in this work. It also assessed the CAD dataset's variation property and its ability to recognize complex human activities. The outcome validates the CAD dataset's variation property and demonstrates the capacity to recognize its activities with more accuracy than other datasets. As a result, it might be used to evaluate the classifiers of complex activities recorded using smartphone sensors.

ACKNOWLEDGEMENTS

The author thanks Alahgaff University for giving him this opportunity to present and publish his research.

REFERENCES

[1] Chavarriaga, R., Sagha, H., Calatroni, A., Digumarti, S. T., Tröster, G., Millán, J. D. R., & Roggen, D. "The Opportunity challenge: A benchmark database for on-body sensor-based activity recognition", Pattern Recognition Letters, 34(15), 2033-2042. 2013.

[2] Ortiz, J. L. R. *Smartphone-based human activity recognition*. Springer. 2015.

[3] Shoaib, M., Bosch, S., Incel, O. D., Scholten, H., & Havinga, P. J. "A survey of online activity recognition using mobile phones", *Sensors*, 15(1), 2059-2085. 2015.

[4] Wu, W., Dasgupta, S., Ramirez, E. E., Peterson, C., & Norman, G. J. "Classification accuracies of physical activities using smartphone motion sensors", *Journal of medical Internet research*, 14(5), e130. 2012.

[5] Liu, L., Peng, Y., Wang, S., Liu, M., & Huang, Z. "Complex activity recognition using time series pattern dictionary learned from ubiquitous sensors", *Information Sciences*, 340, 41-57. 2016.

[6] Dernbach, S., Das, B., Krishnan, N. C., Thomas, B. L., & Cook, D. J. Simple and complex activity recognition through smart phones. *Intelligent Environments (IE), 2012 8th International Conference*. June. IEEE. 214-221. 2012.

[7] Su, X., Tong, H., & Ji, P. "Activity recognition with smartphone sensors", *Tsinghua Science and Technology*, 19(3), 235-249. 2014.

[8] Lukowicz, P., Pirkl, G., Bannach, D., Wagner, F., Calatroni, A., Förster, K., ... & Doppler, J. Recording a complex, multi modal activity data set for context recognition. *Architecture of Computing Systems (ARCS), 2010 23rd International Conference*. February. VDE. 1-6. 2010.
[9] Saguna, S., Zaslavsky, A., & Chakraborty, D. "Complex activity recognition using context-driven activity theory and activity signatures", *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(6), 32. 2013.
[10] Bao, L., & Intille, S. S. Activity recognition from user-annotated acceleration data. *International Conference on Pervasive Computing*. April. Berlin, Heidelberg: Springer. 1-17. 2004.
[11] Zhan, Y., & Kuroda, T. "Wearable sensor-based human activity recognition from environmental background sounds", *Journal of Ambient Intelligence and Humanized Computing*, 5(1), 77-89. 2014.
[12] Liu, L., Peng, Y., Liu, M., & Huang, Z. "Sensor-based human activity recognition system with a multilayered model using time series shapelets", *Knowledge-Based Systems*, 90, 138-152. 2015.
[13] Vaka, P., Shen, F., Chandrashekar, M., & Lee, Y. PEMAR: A pervasive middleware for activity recognition with smart phones. *Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference*. March .IEEE. 409-414. 2015.
[14] M. M. Wahdeen, "Task oriented feature extraction for complex human activity recognition," Ph.D. dissertation, UTM, Malaysia, 2022.