

Analyzing Activity Behavior and Movement in a Naturalistic Environment using Smart Home Techniques

Diane J. Cook, *Fellow, IEEE*, Maureen Schmitter-Edgecombe, and Prafulla Dawadi, *Member, IEEE*

Abstract—One of the many services that intelligent systems can provide is the ability to analyze the impact of different medical conditions on daily behavior. In this study we use smart home and wearable sensors to collect data while (n=84) older adults perform complex activities of daily living. We analyze the data using machine learning techniques and reveal that differences between healthy older adults and adults with Parkinson disease not only exist in their activity patterns, but that these differences can be automatically recognized. Our machine learning classifiers reach an accuracy of 0.97 with an AUC value of 0.97 in distinguishing these groups. Our permutation-based testing confirms that the sensor-based differences between these groups are statistically significant.

Index Terms—pervasive computing, machine learning, Parkinson disease, mild cognitive impairment

I. INTRODUCTION

THE population is aging – the estimated number of individuals aged 85+ will likely triple by 2050 [1]. With this changing demographic comes the urgent goal of developing cost-effective technologies that meet the needs of older adults and their caregivers [2], [3]. Specifically, we must focus attention on smart technologies and their potential to identify early indicators of cognitive and physical illness and to help individuals live independently in their own homes.

In this paper, we investigate the role that smart environments can play in observing and understanding the behavioral impact of aging and aging-related conditions including Parkinson disease (PD) and mild cognitive impairment (MCI). Researchers have argued that assessing individuals in their everyday environment will provide the most valid information about everyday functional status [4]. Currently, in the field of neuropsychology, informant-report and performance-based measures are most commonly used as proxies for real-world functioning [5]–[7]. Both methods, however, have advantages and disadvantages. For example,

although informant report measures are subject to reporter biases, they can capture information about performance across multiple unstructured environments and over an extended period of time [8]. Concomitantly, although performance-based measures typically require completion of one task at a time in a controlled artificial laboratory, they provide a more objective measure of functional capacity [9]. Data also suggests that informant-report and performance-based methods do not always correlate highly with each other [10], [11]. Clinical approaches for assessing functional status are thus limited by restricted behavior sampling, data collection in a laboratory or physician office, and lack of a “gold standard” for measuring everyday functional abilities.

Recently, attention has been directed towards using technologies to examine the quality of tasks being performed within a smart environment. Investigating this area is important, as evidenced by studies that support a relationship between daily behavior and cognitive and physical health [12], [13]. Decline in the ability to independently perform activities of daily living has been associated with placement in long-term care facilities, shorter time to conversion to dementia, and a lower quality of life for both the functionally-impaired individuals and their caregivers [14]–[16]. The maturing of sensor technologies and creation of physical smart environment testbeds provides convincing evidence that we can effectively create such smart environments and use them to aid with clinical assessment and understanding of behavioral differences between healthy older adults and older adults with cognitive and physical impairments.

We propose utilizing smart home and machine learning technologies to observe and quantify the changes in behavior that are manifested for individuals with PD and with MCI. We hypothesize that sensors placed in everyday environments can be analyzed using machine learning techniques to identify differences in the ways that normal everyday activities are performed between healthy older adults and older adults with PD and with MCI. To validate our hypothesis, we perform a study in which older adult participants perform a set of instrumental activities of daily living tasks in our smart home testbed. We apply machine learning techniques to differentiate the groups based on cognitive and physical health status. The results highlight differences in behavioral patterns between participant groups. Based on these behavioral features, our approach automates the classification of clinical health category for these individuals. This study was approved by the Institutional Review Board of Washington State University.

Date submitted: March 18, 2015. This work was supported in part by NSF Grant 1064628 and NIBIB Grant R01EB009675.

D. Cook is with the School of Electrical Engineering & Computer Science, Washington State University, Pullman, WA 99164 (djcook@wsu.edu).

P. Dawadi is with the School of Electrical Engineering & Computer Science, Washington State University, Pullman, WA 99164 (pdawadi@wsu.edu).

M. Schmitter-Edgecombe is with the Department of Psychology, Washington State University, Pullman, WA 99164 (schmitter-e@wsu.edu).

II. RELATED WORK

A smart environment such as a smart home, workplace, or car can be viewed as an environment in which artificial intelligence techniques reason about and control our physical home setting [17]. In a smart home, sensors collect data while residents perform their normal daily routines. Because sensor data can be collected in a naturalistic way without modifying an individual's behavior, it can be used as an aid for automated health assessment and to understand the behavioral impact of various health conditions.

As an example, Pavel et al. [18] observed mobility in smart homes and found evidence to support the relationship between mobility changes and cognitive decline. Lee and Dey [19] presented older adults with visualized information from an embedded sensing system to determine if this information helped them gain increased awareness of their functional abilities. In another effort, Allin and Ecker [2] used computer vision techniques to correlate stroke survivor's motion with expert functional scores on the Arm Motor Ability Test.

The ability to perform automated assessment of cognitive health has recently been given a boost because activity recognition techniques are becoming more capable of accurately identifying an individual's activities in real time from sensor data. These techniques map a sequence of readings from a variety of sensor modalities [20]–[29].

The combination of sensor data collection and activity labeling enhances the ability to study the relationship between health and behavior. First, the labeled activities themselves can be analyzed to see how completely and effectively they are performed by smart home residents. Some earlier work has measured activity correctness for scripted tasks [30], [31] and the activity performance quality is then related to the individual's cognitive and physical health. Similarly, sensing systems have been utilized to assess mental and physical health using motion and audio data [32], [33].

In this work, we enhance previous approaches to behavioral analysis by examining sensor data from smart environments as well as wearable and object sensor data to determine whether behavioral differences can be automatically detected between healthy older adults and older adults with PD and with MCI while performing activities in a smart home setting. We are also interested in determining how well machine learning algorithms can automatically categorize an individual's health condition based solely on sensor data collected in this setting.

III. STUDY DESIGN

A. Method

Participants in this study included 84 Caucasian community-dwelling older adults from a larger sample of 260 individuals. Participants were recruited through advertisements, community health and wellness fairs, physician referrals, and from past studies in our laboratory. The recruitment and data collection occurred over a period of 2 years. We first screened participants by telephone using a medical interview and the Telephone Interview of Cognitive Status (TICS). Exclusionary criteria for this study included a TICS score in the impaired range ($TICS < 26$), history of head

trauma with significant period of coma, current or recent (past year) psychoactive substance abuse, history of cerebrovascular accidents, or other known medical, neurological or psychiatric causes for cognitive difficulties (e.g., epilepsy, multiple sclerosis, significant depression) other than PD or MCI. No participant received more than 1 point for a co-morbid disorder (e.g., type II diabetes) on the Charlson comorbidity index [34]. Those who met initial screening criteria completed two testing sessions. During session 1, participants underwent a battery of neuropsychological tests. During session 2, participants completed activities at the CASAS smart home testbed located on the WSU campus. These evaluations were scheduled one week apart with each testing session lasting approximately three hours.

For this study, we analyzed two participant groups (corresponding to *Analysis 1* and *Analysis 2*). Mean demographic, medical and neuropsychological data (standard deviations in parentheses) for the participant groups can be found in Table 1. As part of Analysis 1, we compared behavior between a Healthy Older Adult (HOA) group and a PD group. A total of 25 PD participants completed both testing sessions. Individuals with PD were diagnosed by their respective neurologist in Eastern Washington; most of the neurologists were board certified with specialties in movement disorders. The majority of PD participants were prescribed anti-Parkinsonian medications ($N=23$, see Table 1) and all were tested in the ON state while on their normal medication regimen. None of the PD patients were wheelchair bound; Hoehn and Yahr Scale score of ≤ 4 [35]. We selected 50 HOA participants who also completed both testing sessions and were most closely matched by age and years of education with the PD group to comprise the HOA group.

For Analysis 2, we decomposed the PD group into two subgroups, PD participants with no indication of mild cognitive impairment (PDNOMCI, $n=16$) and participants who had PD and met criteria for MCI (PDMCI, $n=9$). MCI was diagnosed by two neuropsychologists following review of the testing data, interview data, and medical records when available. Diagnosis of MCI required the following (a) self- or informant-report of decline in cognitive abilities; (b) objective evidence of impairment in one or more cognitive domains represented by scores >1.5 standard deviations below age-based norms or in relation to estimated pre-morbid abilities; (c) not demented, and (d) generally preserved independence in functional abilities as confirmed by self-report and informant-report on an IADL questionnaire. Most of the PDMCI participants met criteria for amnesic MCI ($n=8$), and both single-domain ($n=3$) and multi-domain ($n=6$) MCI are represented in this sample. In order to maintain a uniform class distribution, we selected 18 of the HOA group that were most closely matched to the PDNOMCI group in age and education. In addition, we selected a group of 9 participants who did not have PD but did have MCI; the majority of these individuals met criteria for amnesic MCI ($n=8$) and single-domain MCI ($n=6$). The PD groups had difficulty with processing speed and verbal fluency tests, while the MCI groups differed significantly from the HOA and PDNOMCI groups on verbal learning and memory measures.

Table 1. Mean demographic and neuropsychological data (with standard deviations) for Analysis 1 and 2 participant groups. HOA=healthy older adult, PD=Parkinson's disease, MCI = mild cognitive impairment, WTAR = Wechsler Test of Adult Reading (a measure of premorbid verbal ability), TICS = Telephone Interview of Cognitive Status. ^aMedical records unavailable for one PDNOMCI and one PDMI participant. ^bA measure of manipulative dexterity.

Variable or Test	Groups (Analysis 1)		Groups (Analysis 2)			
	HOA (n=50)	PD (n=25)	HOA (n = 18)	PDNOMCI (n = 16)	PDMCI (n = 9)	MCI (n = 9)
Age	68.60 (9.87)	69.08 (10.12)	70.00 (9.76)	71.50 (10.10)	64.78 (9.15)	68.67 (6.46)
Education	16.66 (2.47)	16.00 (2.94)	16.89 (2.22)	15.88 (3.20)	16.22 (2.59)	17.33 (1.80)
Gender (% female)	72%	52%	72%	50%	44%	67%
WTAR standard score	115.08 (7.45)	112.88 (11.43)	114.88 (7.27)	116.00 (11.14)	107.67 (10.49)	111.22 (10.88)
TICS total score	35.41 (2.44)	34.20 (3.35)	35.29 (2.31)	34.87 (3.79)	33.00 (2.06)	32.71 (1.60)
Years since PD diagnosis	-	6.35	-	4.85	9.02	-
PD Medications ^a	-	70%	-	67%	75%	-
Grooved Pegboard ^b						
Dominant hand	87.94 (24.14)	141.45 (70.83)	82.94 (17.14)	120.71 (58.51)	177.75 (79.54)	95.56 (21.28)
Non-dominant hand	95.65 (24.32)	129.68 (60.59)	98.15 (45.31)	118.92 (33.64)	148.14 (91.10)	94.00 (41.29)

Table 2. Nine activities performed by study participants.

#	Name	Description
1	<i>Water Plants</i>	Retrieve a watering can from the kitchen closet and water all of the plants in the apartment.
2	<i>Medication Management</i>	Retrieve medicine containers and dispenser. Fill the dispenser according to directions.
3	<i>Wash Countertop</i>	Retrieve a rag and cleaning supplies from the kitchen closet, clean all kitchen countertops.
4	<i>Sweep and Dust</i>	Sweep the kitchen, dust the dining living rooms using supplies from the kitchen closet.
5	<i>Cook</i>	Cook a cup of soup, following the package directions and using the kitchen microwave.
6	<i>Wash Hands</i>	Wash hands using the hand soap next to the kitchen sink. Dry hands using a paper towel.
7	<i>TUG Test</i>	Sit in a chair at the end of a hallway that is 3 meters long. Stand up, walk to the end of the hallway, turn around, walk back to the chair, and sit down.
8	<i>TUG Test with Name Generation</i>	Repeat the TUG test while simultaneously completing a semantic fluency task (i.e., generating girls' names).
9	<i>Day Out Task</i>	Perform a complex set of activities as efficiently as possible to prepare for a day out.

perform a sequence of nine activities. Instructions were given before each activity. Once the activity started, no further instructions were provided unless the participant explicitly asked for assistance.

The activities are summarized in Table 2. The first six activities represent instrumental activities of daily living (IADLs) commonly assessed by IADL questionnaires [36] as well as by performance-based measures of everyday competency [37]. Successful completion of IADLs requires intact cognitive abilities such as memory and executive functions. Researchers have shown that declining ability to perform IADLs is related to decline in cognitive abilities [38].

The next two activities represent versions of the Timed Up and Go (TUG) test. The TUG is used to assess a person's mobility and requires both static and dynamic balance. The TUG is used frequently with the older adult population to assess cognitive and motor functions because it is easy to administer and complete. It is valuable for our study because of the reliability that has been demonstrated in using TUG to detect mobility changes in subjects with PD [39].

The ability to multi-task, or perform concurrent tasks by interleaving, has been said to be at the core of competency in everyday life [40]. Therefore, we designed a naturalistic "Day Out Task" (DOT) that participants complete by interweaving subtasks. Participants were told to imagine that they were planning for a day out, which would include meeting a friend at a museum at 10am and later traveling to the friend's house for dinner. The eight subtasks that need to be completed to prepare for the day out are explained and participants are told to multi-task and perform steps in any order to complete the preparation as efficiently as possible. Participants are also provided with a brief description of each subtask that they can reference. The DOT subtasks are summarized in Table 3.

During the experiment, each participant was familiarized with the smart home testbed. The participant was then asked to

Table 3. Subtasks performed as part of the Day Out Task.

#	Name	Description
S1	<i>Magazine</i>	Choose a magazine from the coffee table to read on the bus ride.
S2	<i>Heating pad</i>	Microwave (3 minutes) a heating pad located in the cabinet to take on the bus.
S3	<i>Medication</i>	Right before leaving, take motion sickness medicine from kitchen cabinet.
S4	<i>Bus map</i>	Plan a bus route using a provided map, determine the trip time and calculate when to leave to catch the bus.
S5	<i>Change</i>	Gather correct change for the bus.
S6	<i>Recipe</i>	Find a recipe for spaghetti, collect ingredients to make the sauce.
S7	<i>Picnic basket</i>	Pack all of the items in a picnic basket located in the closet.
S8	<i>Exit</i>	When all the preparations are made, take the picnic basket to the front door.

B. Smart Home Testbed

Study activities were performed in our smart home testbed. A smart home environment can be defined as one that acquires and applies knowledge about its residents and their physical surroundings in order to improve their experience in that setting [17]. Our behavioral study was performed using the CASAS smart home testbed [41]. CASAS components include sensors, actuators, and software algorithms which communicate via publish/subscribe instant messaging. All of the CASAS smart homes can store information locally or securely transmit them to the cloud to be stored in an SQL database. The CASAS smart home architecture has been used to gather behavior data in 50 smart homes including the on-campus apartment that we use for the study described here.

Our on-campus smart home testbed contains a living room, dining room, and kitchen on the first floor and two bedrooms, an office, and a bathroom on the second floor. The apartment is instrumented with infrared sensors on the ceiling which detect motion inside their field of view, magnetic door sensors on cabinets and doors which detect door openings and closings, ambient light and temperature sensors which report significant light/temperature changes, and vibration sensors on selected items including the dustpan, broom, duster, oatmeal container, watering can, bowl of noodles, hand soap dispenser, dish soap dispenser, medicine dispenser, and picnic basket which report when the item has been moved. Each of the ambient sensors sends a text message when there is a significant change in the sensed value. We refer to these messages as “events” which are collected and stored by CASAS. All activities were performed on the first floor of the apartment with sensors located as shown in Fig. 1.

In addition, two wearable sensors collected continual information while participants performed activities. An Android smart phone was on the upper dominant arm, affixed by a strap. Additionally, a wearable sensor was attached to the ankle on the dominant side. Both of these devices collected data 30 times a second that contains 3-axis accelerometer, gyroscope, and magnetometer readings.

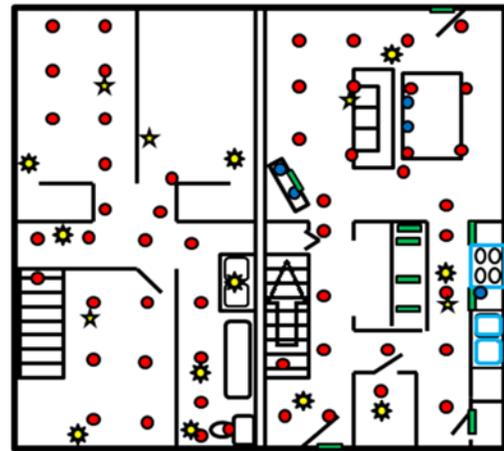


Fig. 1. CASAS smart home testbed. The floorplan of the testbed first floor is shown on the top. Motion sensors are indicated in the floorplan by red circles (example motion sensors are shown lower left), light sensors by eight-pointed stars, door sensors by green rectangles (example door sensors are shown lower right), and temperature sensors by five-pointed stars. The testbed also contains vibration sensors placed on individual items throughout the home.

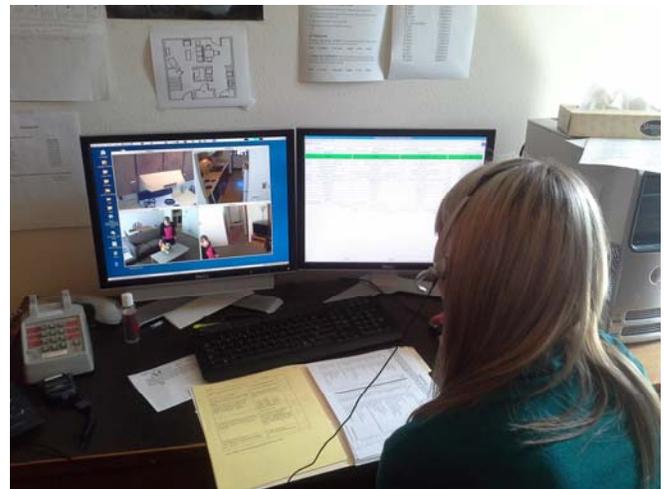


Fig. 2. An experimenter observes a participant performing activities via web cameras and logs information using the RAT.

During the study, an experimenter delivered instructions to participants and answered questions through an intercom while observing the participants’ activities from an upstairs room via a web camera (see Fig. 2), using a Wizard of Oz experiment design. The experimenter used our Real-Time Annotation (RAT) system [42] to log information as it occurred, including the beginning and ending of each activity step. The sensor readings (sensor events) are recorded by CASAS and are formatted based on the date, time, sensor identifier, sensor value, and the RAT annotation, if available.

C. Feature Extraction

The goal of this study is to analyze sensor data to determine if behavioral differences exist between healthy

older adults and older adults with PD based on activity performance. Our machine learning algorithms will analysis features of the raw sensor data, summarized in Table 4.

Table 4. Feature vector used to describe an activity.

Category	Feature	Description
Ambient sensor	<i>Duration</i>	Length of the activity in time
	<i>NumSensorEvents</i>	Number of sensor readings that were generated during activity
	<i>NumSensors</i>	Number of different ambient sensors (total) that generated readings
	<i>NumMSensors</i>	Number of motion sensors that generated readings
	<i>NumDSensors</i>	Number of door sensors that generated readings
	<i>NumISensors</i>	Number of item (vibration) sensors that generated readings
	<i>NumIrrSensorEvents</i>	Number of sensor readings from irrelevant sensors for this activity
Wearable sensor (defined for each wearable sensor feature s and each possible sensor value $s_i \in S$)	<i>Max</i>	Maximum value for this feature
	<i>Min</i>	Minimum value for this feature
	<i>Sum</i>	Sum of the values, $\sum_{i=1}^N s_i$
	<i>Mean</i>	Mean of the values, $\bar{s} = Sum(S) / N$
	<i>Median</i>	Median of the values
	<i>Standard Deviation</i>	Standard deviation, $\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (s_i - \mu)^2}$
	<i>MeanAbs Dev</i>	Mean absolute deviation, $\frac{1}{N} \sum_{i=1}^N s_i - \bar{s} $
	<i>MedianAbsDev</i>	Median absolute deviation, $\frac{1}{N} \sum_{i=1}^N s_i - Median(S) $
	<i>Coeff Variation</i>	Coefficient of variation, σ / μ
	<i>Skewness</i>	Asymmetry in the value distribution, $\frac{1}{N} \sum_{i=1}^N (s_i - \mu)^3 / (\frac{1}{N} \sum_{i=1}^N (s_i - \mu)^2)^{\frac{3}{2}}$
	<i>Kurtosis</i>	The shape of the value distribution, $(\frac{1}{N} \sum_{i=1}^N (s_i - \mu)^4 / (\frac{1}{N} \sum_{i=1}^N (s_i - \mu)^2)^3) - 3$
	<i>Signal Energy</i>	Signal energy, $\sum_{i=1}^N s_i^2$
	<i>LogSignal Energy</i>	$\sum_{i=1}^N \log(s_i^2)$
	<i>Power</i>	Average energy, $\frac{1}{N} \sum_{i=1}^N s_i^2$
	<i>SMA</i>	Signal magnitude area for accelerometer/gyroscope axes, $\sum_{i=1}^N (x_i + y_i + z_i)$
	<i>AutoCorrelation</i>	Amount of correlation between the values at times t and $t+1$, $\sum_{i=1}^{N-1} (s_i - \bar{s})(s_{i+1} - \bar{s}) / \sum_{i=1}^N (s_i - \bar{s})^2$
<i>Axis Correlation</i>	Correlation between each pair of accelerometer / gyroscope axes S and V , $\sum_{i=1}^N (s_i - \bar{s})(v_i - \bar{v}) / \sqrt{\sum_{i=1}^N (s_i - \bar{s})^2 \sum_{i=1}^N (v_i - \bar{v})^2}$	
Day Out Task	<i>Magazine Count</i>	Binary value indicating whether magazine subtask was performed
	<i>HeatingPadCount</i>	Binary value indicating whether heating pad subtask was performed
	<i>MedCount</i>	Binary value indicating whether medicine subtask was performed
	<i>BusMap Count</i>	Binary value indicating whether bus map subtask was performed
	<i>Change Count</i>	Binary value indicating whether change subtask was performed
	<i>RecipeCount</i>	Binary value indicating whether recipe subtask was performed
	<i>PicnicBasketCount</i>	Binary value indicating whether picnic basket subtask was performed
	<i>ExitCount</i>	Binary value indicating whether exit subtask was performed
	<i>PValue</i>	Number of different subtasks were interrupted at least once
	<i>IValue</i>	Maximum number of subtasks that were performed in parallel
	<i>Task1</i>	Subtask performed first
	<i>Task2</i>	Subtask performed second
	<i>Task3</i>	Subtask performed third
Participant	<i>Age</i>	Participant age
Activity	<i>ActivityNum</i>	Activity number

The data features that we extracted can be grouped into five categories: 1) ambient sensor features, 2) wearable sensor features, 3) day out task (DOT) features (for activity 9), 4) participant features, and 5) activity features. The feature vector is summarized in Table 4.

In Table 4, there are eight features that extract information about the ambient sensors. These include the time spent on the activity and the number of sensor readings that were generated during the activity. We also include the number of different sensors (motion, door, item, and total) that generated readings. Unlike the wearable sensors, the ambient sensors only generate readings if they sense an event (e.g., motion in the sensor’s viewing area, door opening or closing, or a significant change in temperature or light). The fact that a particular sensor generated an event during an activity indicates that the participant spent time in that part of the apartment and possibly manipulated doors or items in that region.

The last two ambient sensor features refer to “irrelevant sensors”. There are areas in the apartment in which an activity is typically performed, as well as items and doors that are typically manipulated. If sensors that are not typical for an activity generate readings, this may be an indication that the participant is wandering or not performing the activity correctly. We used the entire set of participants to generate a list of irrelevant sensors for each activity, by noting the number of participants that triggered each sensor for a particular activity. We then use z scores to determine which sensors are outliers for each activity and add these to the irrelevant sensor list for the activity.

The next category of sensor features is extracted from wearable sensors, which continuously collect data throughout the activities. These are standard signal processing features, which are calculated based on values collected during a particular activity. These features complement the ambient sensor features because they provide data corresponding to movement patterns, in contrast with the location and object manipulation insights provided by ambient sensors.

The third category of sensor features is extracted only from the DOT activity. We note that because the activity is complex, some participants do not complete all of the subtasks. The experimenter uses the RAT to note which subtasks are performed, and we encode these as binary values in the feature vector. The parallelism value (pvalue) and interweave value (ivalue) provide an indication of the number of activities that are interrupted, interwoven, and/or performed in parallel. Finally, the last three features in this category indicate which subtask is started first, second, and third. Some sequences of activities will be more efficient than others, so these values give an indication of the strategy that is used to complete the DOT efficiently.

Finally, the last two categories provide information about the participant (participant age) and the activity (activity number). The feature vector contains 352 feature values for each activity. In contrast, a single participant’s feature vector combined over all activities contains 3,152 features.

IV. DATA VISUALIZATION

First, we examine some of the data features using data visualization strategies. The box plots in Figure 4 show the duration of a selection of tasks for the HOA and PD population groups. The box shows the range of durations for the group with a horizontal line at the median point, whiskers at the smallest and largest values still within 1.5 interquartile range, and individual stars indicating outliers.

We only show visualization highlights here. For many of the activities and features there is not a large difference between each subpopulation. The largest duration differences between the groups occur with Cook (which is longer than the other activities) and TUG performed while listing items belonging to a given category (i.e., girls’ names). We note that the PD participants take longer to perform these activities and the range of times for this group is larger.

The second type of plot we include is a probability density function of the task duration. The population mean is indicated with the smooth curve. As shown in these plots, the differences between groups are not extremely pronounced.

We can tell from looking at the graphs that distinguishing behavior between the different population groups cannot be effectively based on any single activity or sensor feature. Instead, we will use machine learning algorithms that can distinguish the groups using more complex learned functions.

V. ANALYSIS RESULTS

Our goal for this study is to determine whether differences exist in the way that individuals with PD and with MCI perform complex daily activities in comparison with HOA, and whether these differences can be detected by sensors using machine learning algorithms. We therefore pose a number of hypotheses that we will evaluate using our study data:

- **H1:** Machine learning algorithms can use sensor data to distinguish HOA participants from PD participants (*Analysis 1*).
- **H2:** Machine learning algorithms can use sensor data to distinguish participants in the categories HOA, MCI, PDNOMCI, and PDMCI (*Analysis 2*).
- **H3:** The *choice of sensor* (e.g., ambient, wearable) will have a noticeable impact on the ability to differentiate different participant groups.
- **H4:** The ability to differentiate between different participant groups will vary depending upon the *activity* that is being monitored and analyzed.
- **H5:** *Dimensionality reduction* techniques can improve classifier performance, particular when the data dimensionality is high and the number of data points is relatively small.
- **H6:** The relationship between our sensor-based learned patterns and the actual participant groups will be *statistically significant*.

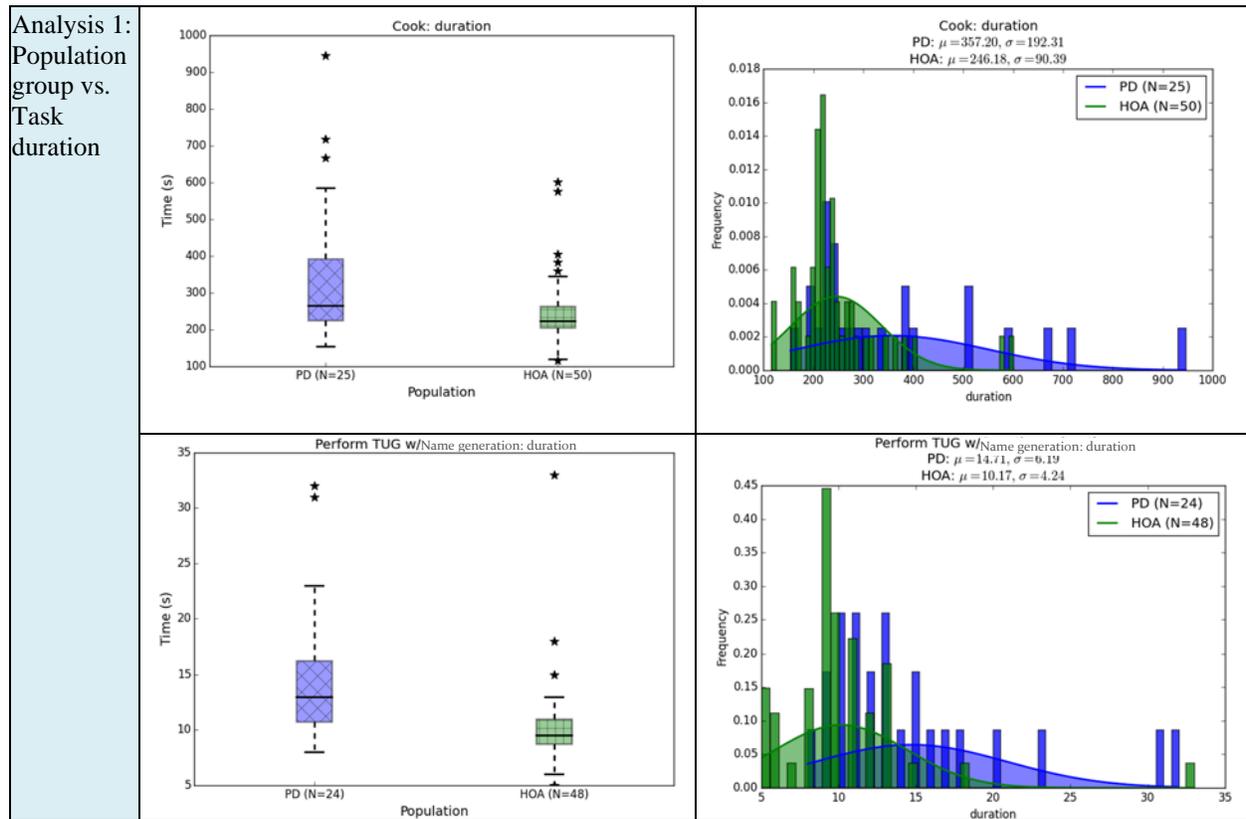


Fig. 3. Summary of Analysis 1 groups: box plots (left) and pdf histogram plots (right) for the cook and TUG with name generation. PD=Parkinson’s Disease population group, HOA=Healthy Older Adults.

HI: Automated classification of HOA and PD participants

We first want to determine whether participant group category can be automatically classified. We see from Figure 3 that the groups may not be easily distinguished based on a single feature. However, machine learning algorithms generate class descriptions and class boundaries as a complex function of all available features and therefore may be able to find behavioral differences between the groups.

We initially compare the HOA and PD groups. To do this, we first input all of the features for each activity as a separate data point to a supervised classifier. Given 75 participants and 9 activities, we have 675 data points labeled “Individual Activities”. Next, we combine the features for all activities into a single data point, so there is one data point for each participant, or 75 data points labeled “Combined Activities”. We measure classification accuracy and area under the ROC curve (AUC). These measures are based on 10-fold cross validation. The AUC values provide additional insights when the class distribution is not perfectly balanced. If behavioral differences exist, the classification performance will be better than that of a random classifier (0.5 for the two-class case).

Different machine learning strategies offer distinct advantages in terms of the complexity of the learned concept, the generalization power, the ability to deal with noisy data, and the ability to process discrete or continuous feature values. We experiment with the following types of classifiers.

- Decision tree classifier (DT). DT selects features that maximally reduce the uncertainty or entropy of the dataset. The learned tree can be expressed as human-readable rules, which allows us to interpret the findings that help differentiate the participant groups.

- Naïve Bayes classifier (NBC). NBC selects a class label that has the greatest probability based on the observed data, using Bayes rule and assuming conditional independence.

- Random forest (RF). RF generates a set of 100 decision trees, each created from a random selection of data points.

- Support vector machine (SVM). SVM identifies a function that maximizes the distance to data points in both classes. Cases which involve more than two classes are modeled as a combination of one-versus-all binary classifiers.

- Adaptive Boosting. AdaBoost combines multiple base classifiers by weighting each one according to its accuracy. We combine boosting with a decision tree (Ada/DT) and with a random forest (Ada/RF).

Table 5. Performance of classifiers on Analysis 1.

Classifier	Individual Activities		Combined Activities	
	Acc	AUC	Acc	AUC
Decision Tree	0.78	0.77	0.55	0.56
Naïve Bayes	0.51	0.62	0.73	0.75
Random Forest	0.74	0.77	0.68	0.69
SVM	0.71	0.57	0.61	0.55
Ada / DT	0.79	0.82	0.59	0.56
Ada / RF	0.74	0.75	0.70	0.70

Table 5 summarizes the performance results of the classifiers for the HOA and PD participants (Analysis 1). As the results show, all classifiers do a better job than random classification ($p < .01$ for each classifier), indicating that there are observable differences in behavior between participants in the two groups. On the other hand, the sensitivity of Ada/DT for the individual activities is .90 and the specificity is .57. For the combined experiment the Ada/DT sensitivity is .70 and the specificity is .36. This indicates that in both cases, the healthy adult class is easier to recognize, possibly as a result of having more data points in this category.

Decision trees provide human-interpretable rules by which classifications are generated. The tree orders features in a greedy search method according to their ability to decrease the entropy on the dataset, or conversely increase the ability to classify the data. Based on this, we note that the features which most clearly separated the two population groups were specific wearable sensor features (power, sma, autocorrelation, skewness, and median), a few ambient sensors (number of irrelevant sensors, magazinecount, and duration), and the participant's age. Including the participant's age raises some interesting questions because this is not a property that is strictly observable from the smart home sensors. As Figure 4 shows, age does have a slight impact on activity performance (in this case, on activity duration). To determine the quantifiable impact of this feature on performance, we remove it from consideration and retest the Ada/DT algorithm, which was the best performing classifier. When age is removed, the performance of the Ada/DT algorithm does drop to an accuracy of 0.70 and an AUC of 0.70. The groups are matched in age, so we conclude that age does play a factor in combination with sensor-observed features. However, the smart home sensors combined with machine learning can still do a reasonable job of separating the two groups without including the age variable.

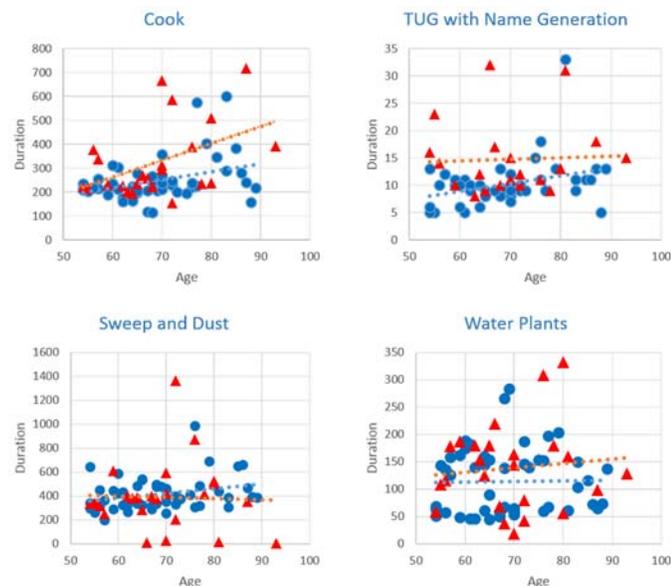


Fig. 4. Scatter plots of the cook, TUG with name generation, sweep and dust, and water plants activity plotting duration as a function of age with trend lines for Healthy Older Adult participants (blue circles) and Parkinson's Disease participants (red triangles).

H2: Automated classification of HOA, MCI, PDNOMCI, and PDMCI participants

In the first experiment, we observe that the PD group is a complex category. Some of the PD participants also have MCI and others are cognitively healthy. The cognitive health of the participant can impact the way the activity is performed, as can the motor symptoms (e.g., resting tremor, slow movement, rigidity, postural instability) that are often observed with Parkinson patients. As a result, we introduce Analysis 2 in which we compare groups of PD without MCI (PDNOMCI), age and education-matched healthy older adults (HOA), PD with MCI (PDMCI), and participants who do not have PD but do have MCI (MCI). We hypothesize that these four categories will be easier in general to distinguish than the HOA and PD groups. However, the classifiers may have more difficulty with Analysis 2 because each group has fewer data points than for Analysis 1. Machine learning classifiers need a large number of data points in order to learn such complex concepts based on a large number of features.

Table 6. Performance of classifiers on Analysis 2.

Classifier	Individual Activities		Combined Activities	
	Acc	AUC	Acc	AUC
Decision Tree	0.81	0.88	0.27	0.61
Naïve Bayes	0.37	0.59	0.33	0.69
Random Forest	0.56	0.76	0.27	0.48
SVM	0.44	0.63	0.25	0.45
Ada / DT	0.85	0.96	0.32	0.64
Ada / RF	0.50	0.71	0.24	0.49

The results of Analysis 2 are shown in Table 6. Accuracies for Analysis 2 have a wider range than for Analysis 1 but also have higher overall accuracies, as high as 0.85 accuracy and 0.96 AUC using Adaboost with decision trees. The accuracy is better than a random classifier with statistically significant results when using separate data points for each activity and when using combined activities data points, both with the Ada / DT classifier ($p < .01$). These results provide evidence that there are behavior differences between these four participant groups. Using Ada/DT, the sensitivity is higher for the healthy adult category (.90, as opposed to .79 for PD, .89 for MCI, and .82 for PDMCI). The specificity for all categories using Ada/DT is fairly high (.90 for HOA, .92 for PD, .97 for MCI, and .97 for PDMCI).

As with Analysis 1, several features offer discriminative power. However, in this case the most-used features are the wearable sma, the number of unique sensors and total number of sensor events, the number of irrelevant sensors, the activity duration, and the participant's age. The features that appear in the Analysis 2 model but not the Analysis 1 model focus more on the amount of time and movement for each activity, which may highlight differences between MCI participants and non-MCI participants. Prior direct observation work in our smart home testbed has shown that, in comparison to HOA, individuals with MCI require greater time to complete complex activities and demonstrate more incomplete and inaccurate activity completion [43].

Although the classifiers vary in performance, a one-way ANOVA indicates that the differences are not statistically significant for either classification experiment. As a result, we stick with the highest-overall-performing classifier, adaptive boosting with decision trees (Ada/DT), for the remainder of the experiments we discuss in this paper.

H3: Analysis of Sensor Choices

In this paper we are focusing on activity performance differences between individuals with PD and healthy older adults. In the literature, the choice of sensor platform varies greatly for activity modeling and analysis [20]–[24]. Different types of sensors offer advantages in terms of the types of activities they can easily track [44]. They also differ in terms of cost, placement challenges, usability, and information granularity.

We are interested in determining which sensor platform is most effective at detecting differences between our population groups while they perform complex activities. In this case, we consider two classes of sensors: ambient and wearable. We postulate that differences between the HOA and PDNOMCI groups will be more easily detected by the wearable sensors. To test this, we repeat our supervised classification task using all sensors, just the ambient sensors, and just the wearable sensors. We also perform this experiment with just the HOA and MCI groups (*Analysis 3*), where we predict that the ambient sensors will better detect the differences. We also include the HOA and PD classification for comparison. Our experiments have observed that PD participants may have differences in body movement patterns from HOA participants. On the other hand, we have found that MCI participants exhibit differences from HOA participants with respect to task accuracy and duration and their patterns of movement around the home [32].

The results from this experiment, summarized in Table 7, are interesting and somewhat mixed. Overall, we see that distinguishing the HOA group from the MCI is simpler than distinguishing HOA from PD or PDNOMCI. This is intuitive, because as the direct observation data indicates, the MCI participants tend to exhibit more errors in the activities (e.g., missed steps, substituting objects), more off task activities (e.g., wandering) and take more time than the HOA group.

Table 7. Performance of all, ambient, and wearable sensors. HOA/PD=Healthy Older Adults and older adults with Parkinson’s Disease, HOA=Healthy Older Adults, PD=Parkinson’s Disease, MCI=Mild Cognitive Impairment, PDNOMCI=Parkinson’s Disease no MCI.

	Sensor Choice	HOA/PD		HOA/PDNOMCI		HOA/MCI	
		Acc	AUC	Acc	AUC	Acc	AUC
Individual	All	0.79	0.82	0.79	0.88	0.87	0.96
	Ambient	0.70	0.68	0.64	0.67	0.87	0.93
	Wearable	0.78	0.80	0.38	0.41	0.96	0.99
Combined	All	0.59	0.56	0.53	0.48	0.68	0.79
	Ambient	0.65	0.61	0.53	0.56	0.63	0.61
	Wearable	0.65	0.66	0.75	0.80	0.49	0.44

In contrast, the differences between the HOA and PDNOMCI participants are subtler. Nevertheless, HOA/PDNOMCI classification is better than random even when a subset of sensors is utilized. The wearable sensors perform best when each participant is considered as a single data point (the combined data), which is consistent with the theory that many of the differences will be related to body movement patterns rather than the method of performing activities. On the other hand, ambient sensors performed best for HOA/MCI classification in the combined case. When every activity is treated as a separate data point, the wearable sensors do separate the classes, better than when integrating ambient sensors and with close to perfect accuracy. When all of the data for a single participant is combined, accuracy drops due to the smaller number of data points. This provides evidence that body movement patterns differ not only between HOA and PDNOMCI participants but also between HOA and MCI participants, although the ambient sensors generally boost the overall recognition accuracy. This is consistent with recent literature which suggests that changes in gait speed may precede diagnosis of mild cognitive impairment [45].

H4: Analysis of Individual Activities

In the same way that different sensor platforms highlight particular types of differences between individuals and groups, so we hypothesize that activities will highlight to a greater or lesser extent the differences in activity performance between individuals with Parkinson disease and healthy older adults. To validate this hypothesis, we look at the sensor data for each activity individually to see how well our machine learning algorithm can classifier individuals as HOA or PD (Analysis 1) or HOA, MCI, PDNOMCI, or PDMCI (Analysis 2).

Table 8. Individual activity performance for Analysis 1 and 2. HOA=Healthy Older Adults, PD=Parkinson’s Disease, MCI=Mild Cognitive Impairment, PDNOMCI=Parkinson’s Disease no MCI, PDMCI=Parkinson’s Disease with Mild Cognitive Impairment.

Activity	HOA/PD		HOA/MCI/PDNOMCI/PDMCI	
	Acc	AUC	Acc	AUC
Water plants	0.65	0.57	0.27	0.45
Medication	0.60	0.54	0.29	0.54
Wash countertop	0.68	0.65	0.35	0.54
Sweep / dust	0.67	0.67	0.35	0.54
Cook	0.63	0.57	0.29	0.57
Hand wash	0.71	0.58	0.35	0.58
TUG	0.63	0.60	0.29	0.51
TUG with names	0.57	0.55	0.42	0.56
Day Out Task	0.64	0.67	0.37	0.59

The results are shown in Table 8. The classifiers perform better when utilizing information from all activities than when focusing on any one activity. This is consistent with current diagnostic approaches which rely on multiple pieces of data, and also supports the idea of performing long-term continual monitoring of activity behavior in order to predict, detect, and monitor the progression of disease. This is the role that smart

home technologies can play in analyzing, diagnosing, and assisting with PD and other age-related changes and disorders.

A one-way ANOVA indicates that the difference between the groups is not significant. However, the difference in performance between using all activities and using any one single activity is significant ($p < .05$). Looking at the activities, we see that the more complex tasks such as TUG with name generation and the day out task provide good discriminating power for HOA/MCI/PDNOMCI/PDMCI based on AUC values. This is not the case for the HOA/PD classification task, however, for which analysis no single activity appears to provide the most insights on population group patterns.

H5: Dimensionality Reduction

At this point, we observe that the datasets have very large feature vectors. The individual activities datasets have 352 features and the combined dataset contains 3,152 features. Learning concepts in high-dimensional spaces can sometimes present a challenge for machine learning algorithms, referred to as the “curse of dimensionality”. Dimensionality reduction techniques such as principal component analysis (PCA) can be applied to the data before classification to map the original feature space onto a lower-dimensional space in a way that maximizes the variance of the data in the new space.

We apply PCA to each of the datasets, which reduces the individual activities datasets to 32 features and the combined datasets to a range of 51-53 features. In the HOA/PD analysis, reducing the feature vector has a dramatic impact as shown in Table 9, improving accuracy to 0.75 and AUC to 0.72 ($p < .05$) for the combined activities. Reducing the dimensionality does not consistently improve performance for the other cases, however, and for some of the datasets the performance actually drops. This is consistent with findings across the field of machine learning, which indicate that the impact of dimensionality reduction is highly dependent on the nature of the data and the selected classification algorithm [46].

We can also go the opposite direction. Dimensionality reduction can be valuable when the feature vector is large and the number of data points is fairly small. However, in some cases the feature vector is not large enough. In particular, because the features are specified a priori, they may not fully and accurately describe the nature of the data. For this data, then, we experiment with adding features to the vector. We do this by applying a simple k means clustering algorithm with $k=2$ clusters. With this additional feature the performance improves for one of the smaller (individual activities) datasets. In addition, we can address the data sparsity problem by performing random resampling of the data. As expected, this has the greatest effect on HOA/PDNOMCI, which has very few data points for one of the classes, as well as for Analysis 2, which has the greatest number of classes and smallest number of data points for each class.

H6: Permutation Test

Finally, we perform a permutation-based p -value test to determine if there is a statistically-relevant relationship between the sensor data and participant group labels. This technique for evaluating classifiers was introduced by Ojala and Garriga [47] and is used to address situations such as ours

where the data is characterized by a large number of features and a relatively small number of data points, so it may be difficult to determine whether the classifier results are trustworthy and generalizable to new data points (in our case, people who did not participate in the study).

In this analysis, we keep the original collection of data points and their class labels. We train a classifier on the data (as shown in Tables 5 and 6). Next, we keep the data points but randomly shuffle the class labels and repeat the training and testing process, making sure that each participant is not assigned more than one unique class label. The original set of data points is used as is the original class distribution, but the labels are randomly permuted. The permutation test measures how likely the observed classifier would be obtained by chance and the resulting p -value represents the fraction of random data sets under a null hypothesis where the classifier performed at least as well as in the original data.

Table 9. Result of applying data preprocessing techniques: 1) original, 2) PCA reduction, 3) addition of a cluster identifier, and 4) random resampling. HOA=Healthy Older Adults, PD=Parkinson’s Disease, MCI=Mild Cognitive Impairment, PDNOMCI=Parkinson’s Disease no MCI, PDMCI=Parkinson’s Disease with Mild Cognitive Impairment.

	App- roach	HOA/PD		HOA/ MCI/ PDNO- MCI/ PDMCI		HOA/ PDNO- MCI		HOA/ MCI	
		Acc	AUC	Acc	AUC	Acc	AUC	Acc	AUC
Indi- vidual	1	0.79	0.82	0.84	0.96	0.79	0.88	0.87	0.96
	2	0.69	0.62	0.40	0.62	0.60	0.64	0.43	0.57
	3	0.80	0.84	0.84	0.96	0.78	0.81	0.85	0.96
	4	0.77	0.79	0.86	0.97	0.91	0.97	0.86	0.96
Com- bined	1	0.59	0.56	0.32	0.64	0.53	0.48	0.68	0.79
	2	0.75	0.72	0.24	0.44	0.36	0.25	0.35	0.45
	3	0.64	0.63	0.29	0.64	0.50	0.50	0.41	0.65
	4	0.67	0.69	0.35	0.67	0.97	0.97	0.35	0.60

Table 10 presents the results from the permutation test. In this case we focus on the combined activity set. This combines features from all nine activities into one feature vector, so there is one data point per participant. This is the most challenging case for our classifier because there is a high-dimensional feature vector and relatively few data points from which to learn the concept.

We used the original classifier method presented in Section 5.1. Given the original classification results for this combined dataset, there may be a concern whether the classifier is identifying a true difference between the groups or if the results are due to chance. However, the statistically significant result for the null hypothesis indicates that there does exist a relationship between sensor-based activity performance and participant groups. The results suggest that the machine learning classifier applied to our sensor data is significant, even if the data is sparse with high dimensionality.

Table 10. Results of permutation test for Analysis 1 and 2.

Analysis	Accuracy	p value
1	0.59	$p < .0001$
2	0.32	$p < .0007$

VI. DISCUSSION

The study described in this paper does face some limitations. The current study relies on experimenters entering activity information into RAT. For this information to be used in a home setting activities would need to be recognized automatically, as described in the literature [48]. The type of cross-sectional study described here also presumes that individuals will perform activities in a fairly uniform manner. Use of this technology in a home setting will require that activity recognition and analysis be sensitive to individual activity differences that are not linked with PD or MCI. The technology will also need to be enhanced to operate in settings with multiple residents and interrupted activities.

Another potential limitation is user acceptance of the technology. All of the collected information in this study was encrypted and password protected. User perception of privacy as well as the actual security of home-based systems is an ongoing issue that will need to be addressed for the technology to be widely accepted and used.

Finally, we point out that some of the features that we include in the analysis are not directly computable from the features. In particular, the age of the smart home resident would need to be supplied by the individual. Other features such as the irrelevant sensors and the activity duration would rely on automated activity recognition and segmentation. While such tools are available, they were not used for this particular study.

VII. CONCLUSIONS

In this paper, we investigated whether the impact of PD and MCI could be sensed and identified using sensor data and machine learning algorithms. Our results indicate that smart homes, wearable devices, and ubiquitous computing technologies can be useful for monitoring activity behavior and analyzing the data to pinpoint differences between healthy older adults and older adults with PD or MCI. The technologies can be used to perform in-home health monitoring as well as early detection of functional changes associated with PD and MCI. The technologies can also assist with treatment validation by providing an ecologically-valid setting in which residents are monitored in their own homes while performing their normal routines. In future work, we will investigate methods of automatically evaluating the quality of the performed activities and comparing this across population groups. We also plan to increase our study to include more participants and population groups to better understand using our sensor-based machine learning methods.

ACKNOWLEDGEMENTS

We want to thank Jennifer Walker, Kaci Johnson and Kylee McWilliams for their help with overseeing data collection, and Gina Sprint for her help generating visualization graphs.

REFERENCES

- [1] V. Lesser, M. Atighetchi, B. Benyo, B. Horling, A. Raja, R. Vincent, T. Wagner, X. Ping, and S. X. Q. Zhang, "The intelligent home testbed," in *Proceedings of the Autonomy Control Software Workshop*, 1999.
- [2] A. Essen, "The two facets of electronic care surveillance: an exploration of the reviews of older people who live with monitoring devices," *Soc. Sci. Med.*, vol. 67, pp. 128–136, 2008.
- [3] J. Powell, L. Gunn, P. Lowe, B. Sheehan, F. Griffiths, and A. Clarke, "New networked technologies and carers of people with dementia: An interview study," *Ageing Soc.*, vol. 30, pp. 1073–1088, 2010.
- [4] T. D. Marcotte, J. C. Scott, R. Kamat, and R. K. Heaton, "Neuropsychology and the prediction of everyday functioning," in *Neuropsychology of Everyday Functioning*, 2010, pp. 5–38.
- [5] L. S. Miller, C. L. Brown, M. B. Mitchell, and G. M. Williamson, "Activities of daily living are associated with older adult cognitive status: Caregiver versus self-reports," *J. Appl. Gerontol.*, vol. 32, pp. 3–30, 2013.
- [6] M. Mitchell, L. S. Miller, J. L. Woodard, A. Davey, P. Martin, M. Burgess, and L. W. Poon, "Regression-based estimates of observed functional status in centenarians," *Gerontologist*, vol. 51, pp. 179–189, 2010.
- [7] R. S. M. Tsang, K. Diamond, L. Mowszowski, S. J. G. Lewis, and S. L. Naismith, "Using informant reports to detect cognitive decline in mild cognitive impairment," *Int. Psychogeriatrics*, vol. 24, pp. 967–973, 2012.
- [8] S. A. M. Sikkes, E. S. M. de Lange - de Klerk, Y. A. L. Pijnenburg, P. Scheltens, and B. M. J. Uitdehag, "A systematic review of instrumental activities of daily living scales in dementia: Room for improvement," *J. Neurol. Neurosurg. Psychiatry*, vol. 80, pp. 7–12, 2009.
- [9] D. J. Moore, B. W. Palmer, T. L. Patterson, and D. V. Jeste, "A review of performance-based measures of functional living skills," *J. Psychiatr. Res.*, vol. 41, pp. 97–118, 2007.
- [10] D. Loewenstein and A. Acevedo, "The relationship between instrumental activities of daily living and neuropsychological performance," in *Neuropsychology of Everyday Functioning*, T. D. Marcotte and I. Grant, Eds. New York: The Guilford Press, 2010, pp. 93–112.
- [11] M. Schmitter-Edgecombe, C. Parsey, and D. J. Cook, "Cognitive correlates of functional performance in older adults: Comparison of self-report, direct observation and performance-based measures," *J. Int. Neuropsychol. Soc.*, vol. 17, no. 5, pp. 853–864, 2011.
- [12] A. K. Coppin, A. Shumway-Cook, J. S. Saczynski, K. V. Patel, A. Ble, L. Ferrucci, and J. M. Guralnik, "Association of executive function and performance of dual-task physical tests among older adults: Analyses from the InChianti study," *Age Ageing*, vol. 35, no. 6, pp. 619–624, 2006.
- [13] M. Schmitter-Edgecombe, C. M. Parsey, and R. Lamb, "Development and psychometric properties of the instrumental activities of daily living – compensation scale (IADL-C)," *Neuropsychology*.
- [14] T. Luck, M. Lupp, M. C. Angermeyer, A. Villringer, H. H. König, and S. G. Riedel-Heller, "Impact of impairment in instrumental activities of daily living and mild cognitive impairment on time to incident dementia: results of the Leipzig Longitudinal Study of the Aged," *Psychol. Med.*, vol. 41, no. 5, pp. 1087–1097, 2011.
- [15] D. Marson and K. Hebert, "Functional assessment," in *Geriatric Neuropsychology Assessment and Intervention*, 2006, pp. 158–189.

- [16] Y. Ouchi, K. Akanuma, M. Meguro, M. Kasai, H. Ishii, and K. Meguro, "Impaired instrumental activities of daily living affect conversion from mild cognitive impairment to dementia: the Osaka-Tajiri Project," *Psychogeriatrics*, vol. 12, no. 1, pp. 34–42, 2012.
- [17] D. J. Cook and S. K. Das, *Smart Environments: Technologies, Protocols, and Applications*. Wiley, 2005.
- [18] M. Pavel, A. Adami, M. Morris, J. Lundell, T. L. Hayes, H. Jimison, and J. A. Kaye, "Mobility assessment using event-related responses," in *Transdisciplinary Conference on Distributed Diagnosis and Home Healthcare*, 2006, pp. 71–74.
- [19] M. L. Lee and A. K. Dey, "Embedded assessment of aging adults: A concept validation with stake holders," in *International Conference on Pervasive Computing Technologies for Healthcare*, 2010, pp. 22–25.
- [20] N. Krishnan and D. J. Cook, "Activity recognition on streaming sensor data," *Pervasive Mob. Comput.*, vol. 10, pp. 138–154, 2014.
- [21] P. Palmes, H. K. Pung, T. Gu, W. Xue, and S. Chen, "Object relevance weight pattern mining for activity recognition and segmentation," *Pervasive Mob. Comput.*, vol. 6, no. 1, pp. 43–57, 2010.
- [22] J. R. Smith, K. P. Fishkin, B. Jiang, A. Mamishev, M. Philipose, A. D. Rea, S. Roy, and K. Sundara-Rajan, "RFID-based techniques for human-activity detection," *Commun. ACM*, vol. 48, no. 9, pp. 39–44, 2005.
- [23] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Comput. Surv.*, vol. 46, no. 3, pp. 107–140, 2014.
- [24] O. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Commun. Surv. Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.
- [25] D. Cook, "Learning setting-generalized activity models for smart spaces," *IEEE Intell. Syst.*, vol. 27, no. 1, pp. 32–38, 2012.
- [26] M. Philipose, K. P. Fishkin, M. Perkowitz, D. J. Patterson, D. Hahnel, D. Fox, and H. Kautz, "Inferring activities from interactions with objects," *IEEE Pervasive Comput.*, vol. 3, no. 4, pp. 50–57, 2004.
- [27] J. K. Aggarwal and M. S. Ryoo, "Human activity analysis: A review," *ACM Comput. Surv.*, vol. 43, no. 3, pp. 1–47, 2011.
- [28] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu, "Sensor-based activity recognition," *IEEE Trans. Syst. Man, Cybern. Part C Appl. Rev.*, vol. 42, no. 6, pp. 790–808, 2012.
- [29] A. Reiss, D. Stricker, and G. Hendeby, "Towards robust activity recognition for everyday life: Methods and evaluation," in *Pervasive Computing Technologies for Healthcare*, 2013, pp. 25–32.
- [30] D. J. Cook, M. Schmitter-Edgecombe, and G. Singla, "Assessing the quality of activities in a smart environment," *Methods Inf. Med.*, vol. 48, no. 5, pp. 480–485, 2009.
- [31] M. Hodges, N. Kirsch, M. Newman, and M. Pollack, "Automatic assessment of cognitive impairment through electronic observation of object usage," in *International Conference on Pervasive Computing*, 2010, pp. 192–209.
- [32] P. Dawadi, D. J. Cook, and M. Schmitter-Edgecombe, "Automated cognitive health assessment using smart home monitoring of complex tasks," *IEEE Trans. Syst. Man, Cybern. Part B*, vol. 43, no. 6, pp. 1302–1313, 2013.
- [33] M. Rabbi, S. Ali, T. Choudhury, and E. Berke, "Passive and in-situ assessment of mental and physical well-being using mobile sensors," in *ACM International Conference on Ubiquitous Computing*, 2011, pp. 385–394.
- [34] M. E. Charlson, P. Pompei, K. L. Ales, and C. R. MacKenzie, "A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation," *J. Chronic Dis.*, vol. 40, no. 5, pp. 373–383, 1987.
- [35] M. Hoehn and M. Yahr, "Parkinsonism: Onset, progression and mortality," *Neurology*, vol. 17, no. 5, pp. 427–442, 1967.
- [36] B. Reisburg, S. Finkel, J. Overall, N. Schmidt-Gollas, S. Kanowski, H. Lehfeld, F. Hulla, S. G. Sclan, H.-U. Wilms, K. Heininger, I. Hindmarch, M. Stemmler, L. Poon, A. Kluger, C. Cooler, M. Bergener, L. Hugonot-Diener, P. H. Robert, and H. Erzigkeit, "The Alzheimer's disease activities of daily living international scale (ADL-IS)," *Int. Psychogeriatrics*, vol. 13, no. 2, pp. 163–181, 2001.
- [37] M. Diel, M. Marsiske, A. Horgas, A. Rosenberg, J. Saczynski, and S. Willis, "The revised observed tasks of daily living: A performance-based assessment of everyday problem solving in older adults," *J. Appl. Gerontechnology*, vol. 24, pp. 211–230, 2005.
- [38] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician," *J. Psychiatr. Res.*, vol. 12, no. 4, pp. 189–198, 1975.
- [39] K. J. Brusse, S. Zimdars, K. R. Zalewski, and T. M. Steffen, "Testing functional performance in people with Parkinson disease," *Phys. Ther.*, vol. 85, no. 2, pp. 134–141, 2005.
- [40] P. W. Burgess, "No TitleStrategy application disorder: The role of the frontal lobes in multitasking," *Psychol. Res.*, vol. 63, pp. 279–288, 2000.
- [41] D. J. Cook, A. Crandall, B. Thomas, and N. Krishnan, "CASAS: A smart home in a box," *IEEE Comput.*, vol. 46, no. 7, pp. 62–69, 2012.
- [42] K. Feuz and D. J. Cook, "Real-time annotation tool (RAT)," in *AAAI Workshop on Activity Context-Aware System Architectures*, 2013.
- [43] M. Schmitter-Edgecombe, C. McAlister, and A. Weakley, "Naturalistic assessment of everyday functioning in individuals with mild cognitive impairment: The day out task," *Neuropsychology*, vol. 26, pp. 631–641, 2012.
- [44] Y. Sahaf, N. Krishnan, and D. J. Cook, "Defining the complexity of an activity," in *AAAI Workshop on Activity Context Representation: Techniques and Languages*, 2011.
- [45] T. Burracchio, H. H. Dodge, D. Howieson, D. Wasserman, and J. Kaye, "The trajectory of gait speed preceding mild cognitive impairment," *Arch. Neurol.*, vol. 67, no. 8, pp. 980–986, 2010.
- [46] A. Sarveniazi, "An actual survey of dimensionality reduction," *Am. J. Comput. Math.*, vol. 4, pp. 55–72, 2014.
- [47] M. Ojala and C. G. Garriga, "Permutation tests for studying classifier performance," *J. Mach. Learn. Res.*, vol. 11, pp. 1833–1863, 2010.
- [48] P. Dawadi, D. J. Cook, and M. Schmitter-Edgecombe, "Longitudinal functional assessment of older adults using smart home sensor data," *IEEE J. Biomed. Heal. Informatics*, 2015.