

Received February 21, 2022, accepted March 18, 2022, date of publication April 6, 2022, date of current version April 15, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3165208

Dementia Scale Score Classification Based on Daily Activities Using Multiple Sensors

AKIRA MINAMISAWA¹, SHOGO OKADA¹, KEN INOUE², AND MAMI NOGUCHI³

¹Japan Advanced Institute of Science and Technology, Nomi, Ishikawa 923-1292, Japan

²George and Shaun, Inc., Tokyo 151-0051, Japan

³Toyota Connected Corporation, Nagoya, Aichi 460-0003, Japan

Corresponding author: Shogo Okada (okada-s@jaist.ac.jp)

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant 19H01120 and Grant 19H01719, and in part by JST AIP Trilateral AI Research, Japan under Grant JPMJCR20G6.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Research Ethics Committee of the Nippon Telegraph and Telephone West Corporation.

ABSTRACT Early detection of age-related disease symptoms in older people by the use of daily activity data is one of the central challenges of home sensor systems. This paper focuses on dementia scale classification from daily activity data collected using sensors that can be deployed in actual residential environments. Activity data collected by four sensors (a door sensor, human motion sensor, location sensor, and sleep sensor) were obtained by recording 56 older adults living in common residences. We analyzed the effects of different types of sensor data, such as time spent in an individual room according to human motion sensors, location in a facility, and sleep patterns, on dementia detection. We then developed a feature extraction method related to daily activity patterns based on a clustering algorithm and analyzed its effectiveness. In the experimental evaluation, we trained binary classification models to classify dementia scale scores based on the Mini-Mental State Examination (MMSE) from these datasets. The experimental results show that a maximum accuracy of 0.871 was obtained with a linear support vector machine (SVM) model by fusing the door, location, and sleep features and by clustering activity patterns using the X-means algorithm.


INDEX TERMS Dementia, activities of daily living (ADLs), multisensor fusion, machine learning.

I. INTRODUCTION

Dementia is one of the major causes of disability and dependency among older people worldwide. It is predicted that the number of patients with dementia will reach 152 million by 2050 [1]. Dementia is a syndrome in which the ability to think, remember, and perform everyday activities is decreased [2]. A cure for dementia has not been identified, but treatments to relieve the symptoms have been developed [3]. Thus, the early detection of dementia symptoms is important to slow progression and to support patients and their families. From this perspective, some computer science researchers have started to develop computational models to identify people with dementia by using various types of activities. In these studies, researchers recorded individual behaviors and automatically predicted clinical scores based on sensor data [4], [5]. Many studies based on daily activity

have developed systems that automatically detect dementia from activity data collected unobtrusively. In recent research, activity data have been collected from an environment where a large number of sensors are installed, such as a smart home. Features related to residents' daily behaviors or activities of daily living (ADLs) that can be extracted based on activity recognition have shown effectiveness in classifying dementia. However, some difficulties are encountered in building such specific sensor networks in the residences of users.

This research aims to apply a system in actual care facilities for elderly people rather than perform an experiment in a sensor room where a large number of sensors are installed. We propose and evaluate a method to discriminate dementia scale scores obtained by a dementia screening test, namely, the Mini-Mental State Examination (MMSE), from activity data measured by installing simple (low-cost) sensors in an actual residential facility. Furthermore, by developing a classification model for each of the many independent sensors,

The associate editor coordinating the review of this manuscript and approving it for publication was Peng-Yong Kong .

we can compare the effectiveness of the measurements using each sensor.

For this purpose, we propose an effective feature extraction method by referring to a study that analyzed activity patterns during daily activities [6] to improve the dementia detection accuracy. Specifically, the main idea is to extract typical activity patterns by unsupervised dimension reduction algorithms and cluster them from a set of ADL samples observed in all participants in the experiment. A feature set composed of typical activity patterns is appropriate for comparing the activities of different people. In this paper, we refer to the extracted behavioral features based on daily activity patterns as bag-of-activity patterns (BoAPs) and verify their effectiveness in classifying dementia scale scores. The main contributions of this study are summarized as follows.

A. AN ACTIVITY SENSING DATASET COLLECTED FROM MULTIPLE SENSORS IN A REAL ENVIRONMENT

We collected an activity dataset from participants in residential facilities using four different sensors over two months to extract indoor daily activities. The dataset, which was collected from the actual living environment of elderly participants, included location and sleep information of the participants and information from the door sensor and human motion sensor in the room. Two months of data collection under these realistic conditions enabled us to model the daily behaviors of the participants. Furthermore, it is possible to compare and analyze the contribution to the estimation of information obtained from multiple aspects, such as time spent in an individual's room by motion sensors, location in a facility, and sleep patterns.

B. FEATURE EXTRACTION BASED ON EIGEN BEHAVIORS

Daily activity patterns vary among individuals, so extracting effective activity features related to classifying dementia scale scores from such diverse activity samples is still a central challenge. We propose a method to extract activity features by relating diverse activity samples to typical activity patterns that are learned using dimension reduction or clustering algorithms. We show that the activity pattern features extracted from an actual living environment are effective at classifying the differences in dementia scale scores. Collecting the activity dataset and extracting the features from these datasets were conducted in a fully autonomous manner.

C. FUSING OF VARIOUS SENSOR DATA AND COMPARING DIFFERENT SENSING TYPES

By comparing classification models of the dementia scale from different types of sensor data, we analyze the sensor features that are effective for detecting dementia. This study also addresses a novel challenge in investigating the possibility of automatic classification of dementia scale scores by integrating the features obtained from various sensors. We show that fusing these features obtained from various sensors can improve the classification accuracy of the dementia scale.

II. RELATED WORKS

A. ANALYZING DEMENTIA FROM PHYSICAL ACTIVITY DATA

The procedures for the clinical diagnosis of dementia include an interview, a physical examination, and a neurological examination to determine the presence or absence of dementia and its symptoms. Electroencephalography (EEG) is a pathologically valid indicator in the diagnosis of dementia [7], and the effectiveness of EEG-based sensing as an approach for the early detection of cognitive function has been shown [8]–[10]. Wu *et al.* [11] detected the degree of dementia based on EEG information. In recent studies, cognitive function has been detected from a variety of data, including speech, language, gait, and facial expression [12]. Jarrold *et al.* [13] detected the type of dementia from speech based on acoustic features and language. Many of the predictors of dementia levels can be interpreted as general cognitive deficits that are well-known features of dementia, such as decreased memory ability and decreased accuracy and speed of movements. For elderly health monitoring, activity performance should be sensed during daily life without burdening the user.

Daily activity performance is an important indicator of functional health. The inability to perform necessary activities during daily life is associated with increased health care utilization and the risk of developing dementia. Hodges *et al.* [14] used wearable devices (radiofrequency identification (RFID) bracelets) and RFID-tagged objects to detect indicators of cognitive impairment, such as dementia and traumatic brain injury, by monitoring individuals performing a well-defined routine task (making coffee). By extracting activity features using a motion sensor system, Hayes *et al.* [15] investigated the associations among walking speed, the amount of daily activity in the participants' residences and the level of mild cognitive impairment (MCI, high or low). To longitudinally monitor the daily activities of users in the home and to detect and evaluate health functions, it is necessary to automatically detect activities from data collected by unobtrusive sensors. Currently, with the development of ubiquitous computing, inexpensive and reliable sensors have enabled accurate pattern recognition from activity sensing data. Recently, researchers have tried to automate health assessments, including daily activities, based on sensor data. Akl *et al.* [16] acquired motion sensor data over an average period of three years and calculated several measures associated with subjects' walking speeds and general household activities to detect MCI. Dawadi *et al.* [17] investigated the relationship between the ability of a participant to complete an activity and the health assessment (dementia or cognitively healthy). In addition, Alberdi *et al.* [5] extended this work by extracting features associated with ADLs from a sensor system to estimate a health function score. Robben *et al.* [4] developed an ambient sensor monitoring system to collect sensor data in a participant's residence and proposed an algorithm for quantifying the changes in everyday behaviors.

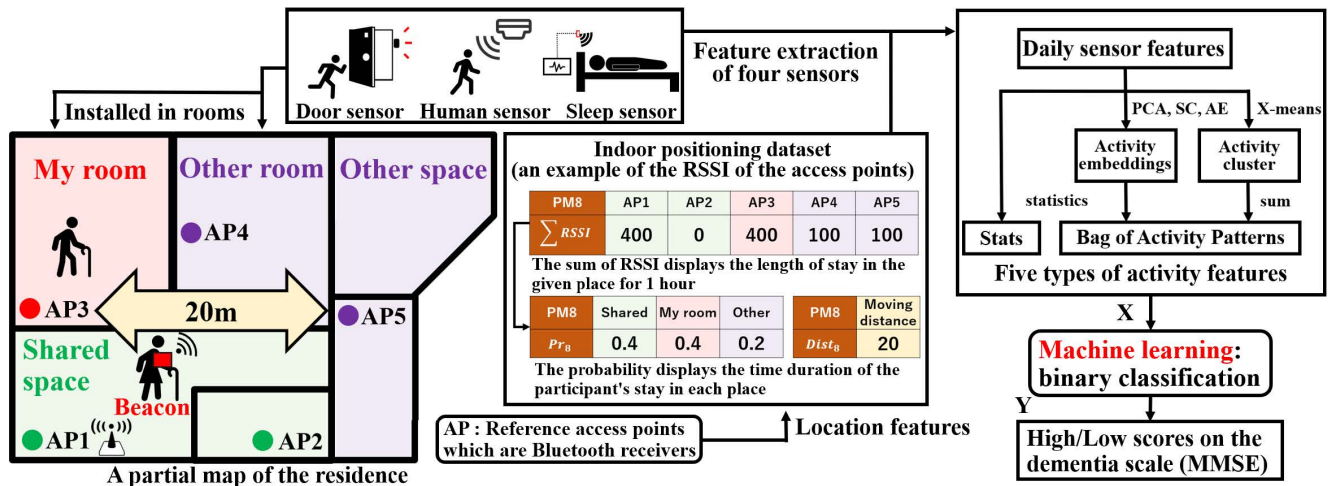


FIGURE 1. Overview of the dementia scale score classification.

Many studies on the estimation of cognitive functions based on ADLs show that large-scale sensing systems such as smart homes enable activity recognition or feature extraction corresponding to daily living behaviors. There are several challenges in implementing such activity sensing environments in actual residential facilities for elderly individuals. Prediction models depend on the different floor plans of the residential facilities and the measured environment.

On the other hand, [18] reported that the frequency of social activity is associated with dementia, and [19] showed that both social interaction and intellectual stimulation may be relevant to preserving mental function in elderly individuals. From these findings, [20] proposed a method to extract daily activity information based on the time spent in each room from indoor positioning information using mobile beacons and evaluated the classification accuracy of dementia scale scores.

B. DETECTING PHYSIOLOGICAL DISEASE FROM SLEEP ACTIVITY

Dementia causes sleep disorders, such as day and night reversal, difficulty falling asleep, and midway awakening problems [21], [22]. On the basis of more than 9 years of research, Hahn *et al.* [23] reported that changes in sleep patterns cause dementia and that respiratory rate, heart rate, and body movement are the main physiological parameters that indicate sleep quality. Previous research applied the differences in sleep behavior estimated from the bedroom and the participant's position to detect dementia and has shown that sleep duration and sleep patterns differ according to dementia scores [5], [17].

Monitoring sleep behavior in depth is known to be effective at estimating health status. EEG is a typical sensing method for identifying specific patterns in sleep, and many studies have been conducted on methods of automatically detecting sleep patterns, such as the automatic classification of

sleep stages [24], [25]. Paradiso *et al.* [26] collected activity and respiration levels during sleep using noninvasive microbiosensors embedded in objects such as clothing items to identify patients with mood disorders. Nam *et al.* [27] used a pressure sensor to obtain behavioral data during sleep, such as body movement and heart rate, to estimate health status. From the viewpoint of achieving a collection of sleep behaviors without burdening the user, such as by wearing a sensor, contact-free bed sensors can now provide high monitoring accuracy [28]. Therefore, we attempt to estimate the dementia scale score from the sleep state acquired by a bed sensor and the biological signals during sleep rather than to perform sleep activity recognition based on an individual's location information. Sleep sensor information enables a comparative analysis of the contribution to the estimation of dementia from information obtained from different sensors, such as location information and information from human motion sensors.

C. POSITION OF THIS RESEARCH

The main difference between this paper and many previous ones analyzing health assessments based on activity data is that we analyze data sensing methods to introduce a dementia estimation system into users' actual living environments. In this study, we used a sensing method that can be easily implemented in actual living environments to collect sleep information and location information, including outings and movement information. Different types of sensors have different advantages and disadvantages in terms of price, reliability, and user acceptability for sensing in the home. Since the type of sensors that can be installed depends on the living environment, we investigate whether the behavioral sensor data acquired by door and human motion sensors in addition to indoor positioning are effective at detecting dementia. Furthermore, we aim to extract the differences in individual activity patterns from sensor data collected in an easy-to-install

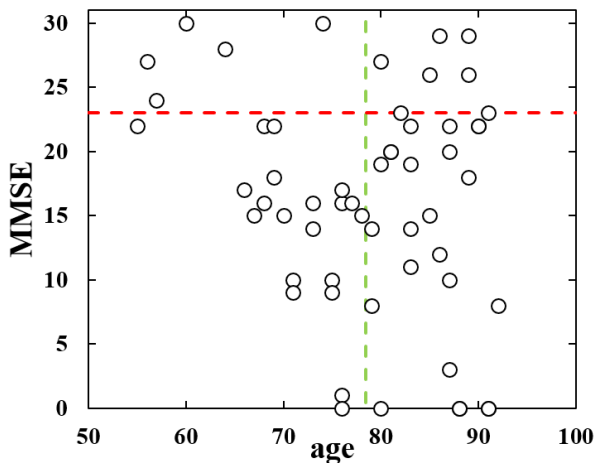


FIGURE 2. Distribution of the MMSE scores (y-axis) with respect to age (x-axis). The horizontal line represents 23 points, which is the threshold for dementia suspicion, and the vertical line represents the mean age.

way and automatically extract features of ADLs that are known to be effective in estimating the dementia scale score. While many previous studies have shown that features of daily activities extracted by sensing systems contribute to the estimation of dementia, they have not focused on comparing the contribution of features obtained from multiple sensors or on fusing features extracted from different types of sensors. The purpose of this study is to determine the best method for collecting activity data that is effective for the classification of dementia scale scores from multiple types of sensors and to evaluate whether the accuracy of the model can be improved by fusing different types of sensor information.

III. RESEARCH OVERVIEW AND DATASET

A. OVERVIEW OF DEMENTIA SCALE CLASSIFICATION

Fig 1 shows an overview of the dementia scale score classification. The procedure for developing a classification model for the dementia scale from a behavioral dataset collected by a variety of sensors is shown here.

1. Extract features for dementia detection from the data from each sensor (IV-A).
2. Extract features based on typical activity patterns extracted by using dimension reduction and clustering algorithms (IV-B).
3. Develop a model to estimate the dementia scale score by machine learning using the activity feature set (V).

B. DEMENTIA SCALE TEST

We recorded the indoor daily activity datasets in two nursing residential facilities in Japan. We recruited 56 Japanese participants. Daily life behavior data were collected using the four types of sensors during a two-month period. The Research Ethics Committee of the Nippon Telegraph and Telephone West Corporation reviewed and approved the collection of data and the corresponding research using this dataset. The dataset, excluding personal information (age,

TABLE 1. Statistics of the participants and the number of participants in the high- and low-scoring groups on a dementia scale.

	No. Participants	MMSE [High/Low]
Door	54	[10/44]
Human	55	[9/46]
Location	47	[8/39]
Sleep	44	[9/35]
All	37	[6/31]
Total	56	[10/46]
Sex [F/M]	[32/24]	
Average age	78.43(\pm 9.63)	

sex, name, and audio) that could be used by a third party to identify and discriminate against the participants, was shared only among the coauthors of this study for the purpose of academic research. Only the following age and sex information was collected and shared: the average age of the participants was 78.43 (\pm 9.63) years old, and the participants included 24 male participants and 32 female participants. Written informed consent was obtained from all the participants or from a capable family member before the following data were collected.

The participants initially completed the MMSE [29], a dementia screening test. In the MMSE, if the score is below 23 points, the likelihood of dementia is high, and a score below 27 points is defined as suspicious for MCI [30]. Fig 2 plots the scores against the ages of the participants. In this dataset, because of the number of participants suspected of having dementia, we classified participants into two groups, namely, dementia suspicion or no dementia suspicion, instead of three groups (dementia, MCI, or other). The task of estimating the dementia scale score is defined in this study as the two-class classification of high- and low-scoring groups on a dementia scale based on the threshold of 23 points on the MMSE.¹ The information of the experimental participant's high/low-score groups collected by each sensor is given in Table 1.

C. ACTIVITY DATASET FROM FOUR DIFFERENT SENSORS

In this section, we describe how to collect the activity dataset. Although the two residential nursing facilities hosting the participants are very different, they have individual rooms with a bed and a toilet and a shared space for free communication among the residents. To collect the indoor location data, we used a Bluetooth Low Energy (BLE) beacon, Biblle 3, which was produced by George and Shaun Co., Ltd. This Bluetooth beacon has a small radio transmitter (6 \times 0.6 \times 2.2 cm, 9.07 g) that sends signals within a radius of 10-30 m (interior spaces). These beacons are cost-effective and can be installed with minimal effort. We installed reference access points (APs), which are the receivers for the

¹The MMSE is a cognitive function test, but the test results do not always match the results of the diagnosis of dementia. Therefore, the proposed classification model that identifies and estimates high/low MMSE scores (the value of one indicator that investigates dementia) does not estimate the actual diagnostic results of dementia.

Bluetooth signal, in the participants' rooms and the shared spaces of the residential facilities. We estimated the position of a participant with a beacon using the received signal strength indication (RSSI) of the Bluetooth signal and the location coordinates of the APs. Additionally, we installed three types of single sensors (door sensors, human motion sensors, and sleep sensors) in the participants' private rooms to collect information on their daily activities using a sensing method different from indoor positioning.

The door sensor (DS1A-A01WH) and human motion sensor (HM92-01WHC) were produced by Nissha Co., Ltd. These sensors were installed on the entrance doors and ceilings of private rooms to detect the entry/exit and presence/absence, respectively, of the participants in the private rooms. The sensors do not need to be replaced because they use solar cells, and they record the opening/closing of the door and presence/absence of the participant as binary values and wirelessly transmit them. Although the interval at which data are collected is irregular, the door sensor periodically transmits its current state every 20-30 minutes, even outside of the open/close timing. In the case of the human motion sensor, the sensor sends the detected status at a maximum interval of once every 5 seconds.

The sleep data of the participants were acquired by using a noncontact vital sensor (MS-106) produced by Mio Corporation. A bed sensor with a 24-GHz microwave Doppler sensor ($13.35 \times 12.12 \times 2.6$ cm, 246 g) was installed on the ceiling or at the top of the wall in private rooms. Microwaves were used to detect the left and right sides of the sensor at $32^\circ/58^\circ$ angles within a range of 1.4-3 m, and two patterns of biological signals were obtained for each of the sensing results from the two directions. Specifically, the heart rate, respiratory rate, body movement level, and whether the participant was in his or her bed were estimated using the sensing results.

The number of participants and recording days were different for each sensor. The participant-level averages for the location, door sensor, human motion sensor, and sleep sensor data were 31.3, 42.4, 42.5, and 33.4 days, respectively.

IV. ACTIVITY FEATURE EXTRACTION

Fig 3 provides an overview of the feature extraction method for the participants' dementia scale score categories from the time series data acquired from each sensor; the steps are listed as follows:

- 1) Represent the activity feature vector of one sample as the activity of one day and prepare an activity sample set containing the samples corresponding to all the dates.
- 2) Discover the typical activity patterns from the activity samples by using clustering and dimension reduction algorithms.
- 3) Extract the activity feature vector of an activity sample using typical activity patterns.
- 4) Transform the activity samples observed from each participant into one sample for each participant.

Daily activity data generally have a specific structure and regularity. When daily activity patterns are compared, health conditions can also be compared [31]. Based on these findings, we represent an activity feature vector for a unit of one day.

Extracting activity types from behavioral data is important for estimating health status. It has been reported that changes in activity patterns correspond to changes in health status [32]. Riboni *et al.* [33] focused on detecting abnormal behaviors of MCI patients from sensor data and built a system for early detection. Extracting patterns of activities, such as those proposed by Dawadi *et al.* [34], is based on an activity recognition algorithm that requires the collection of participants' activity events by many sensors. In this study, we applied dimension reduction and unsupervised clustering algorithms to obtain features that represent daily activity patterns from automatically extracted features based on data from easily installed sensors.

A. DAILY SENSOR FEATURES

First, to acquire the behavior trend, we extracted activity features from time-series sensor data segmented by days. In previous sensing systems, specific activity features associated with ADLs were extracted (e.g., time spent on toilet needs at night, daily sleep duration) [4], [5]. In this study, we designed features that are less dependent on differences among sensor types or living environments, considering their application to actual living environments. Table 2 shows the hourly activity features that can be automatically extracted from each sensor. The daily activity sample x was obtained by concatenating the sensor features calculated for each hour. This approach for extracting features based on hourly sensor values is not limited to the four types of sensors used in this study and can be applied to other types of sensor data that are not used in this study.

- 1) **Door and human motion sensor features:** The door and human motion sensors output a value of 1 when they detect an action. We extracted feature values of 24 dimensions by calculating the total count for which the event (door movement or human detection) was observed within each hour.
- 2) **Indoor positioning features:** The indoor location data obtained using the mobile beacons captured the amount of time that the participants spent in each room daily. We calculated these durations using the location data [20].
 - a) When an AP receives a signal from a beacon (ID of the participant), a data sample is added to the database. The data are composed of the following attributes: (1) ID, M of the AP; (2) ID, P of the beacon; and (3) RSSI, the RSSI of the signal. Each AP M has a coordinate of position $Pos_M = (X_{ax}, Y_{ax})$ in the residential facility.
 - b) Every participant lives in one of the two residential facilities. Therefore, we need to extract common

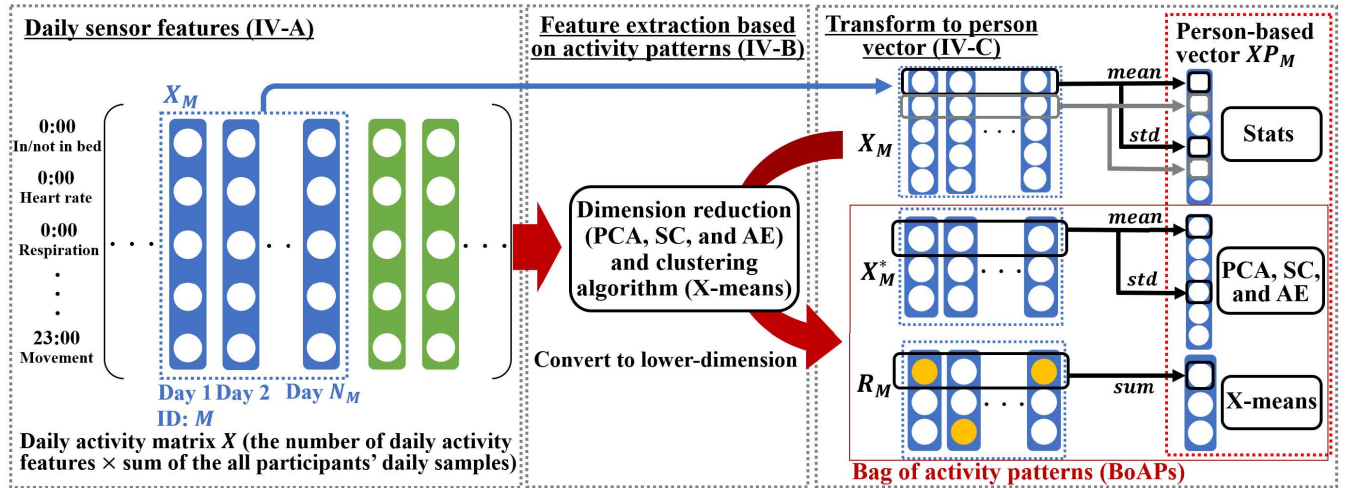


FIGURE 3. Overview of the feature extraction method.

activity features that are independent of the residential facility. Common features in both residential facilities are the shared space and private rooms of the participants. We classify the positions, Pos , into three types: a shared space (shared living and recreation rooms) ($C = 1$), a participant’s own room ($C = 2$), and other places, including other participants’ rooms ($C = 3$) in the residential facility (an example is shown in Figure 1).

- c) The location where participant P stays from time T to time $T + 1$ is estimated using the RSSIs. Let the RSSI vectors of all the Bluetooth signals detected by an AP in location $C \in \{1, 2, 3\}$ be $RSSI_{T,C} = \{rssi_{C,1}, \dots, rssi_{C,N_{T,C}}\}$, where $N_{T,C}$ corresponds to the number of times the signals were received by the APs in location C from time T to time $T + 1$. $rssi$ [dBm] is a negative value ($RSSI > -80$), and we normalize the RSSI by $w_{C,n} = 80 + rssi_{C,n}$.
- d) The probability that participant N stays in location C from time T to time $T + 1$ is $Pr_{T,C} = \sum_n^{N_{T,C}} w_{C,n} / Z$, where $Z = \sum_C \sum_n w_{C,n}$.

We calculate $Pr_{T,C}$ as the probability of the participant staying in a location at each time T from 0-23. $Pr_{T,C}$ is calculated for each day for all the participants. In addition to the percentage of the day spent in each location, we calculate the maximum distance moved at each time. $Dist_T$ is calculated as the maximum distance [m] between the AP where the signal was detected at each hour and the AP of the participant’s own room. The dimension of the location activity feature space is 96.

- 3) **Sleep sensor features:** For each hour, we calculate the average of the eight sensor output values for heart rate, respiratory rate, and body movement levels and the average of the binary value of whether the patient is in bed based on the sensor’s estimate. The eight biological signals to be acquired are listed as follows:

TABLE 2. Extracted sensor features.

Sensor type	Feature	Details
Door	Number of detections	Total number of detected door openings
Human	Number of detections	Total number of detections by the human sensor
Location	Probability of staying	Time spent in shared areas, own room, and other places.
	Moving distance	Maximum distance between the received access point and the participant’s room.
Sleep	In/not in bed	Mean value of bed status
	Heart rate	Mean heart rate
	Respiration	Mean value of respiratory rate
	Body movement	Mean body movement level

heart rate, “heart rate level”, “left heart rate level”, and “right heart rate level”; respiration, “respiration level”, “left respiration level”, and “right respiration level”; and body movement, “left body movement level” and “right body movement level”. The dimension of the sleep feature space is 216.

B. FEATURE EXTRACTION BASED ON ACTIVITY PATTERNS

From the sensor features (daily activity samples) calculated in Section IV-A, we extract features based on activity patterns that are expected to be effective at classifying the dementia scale score. The activity data matrix (the number of all the participants’ daily samples (n) \times the number (d) of daily activity features) was composed by using all the participants’ daily activity samples (data samples) (x_i), so the matrix is represented as $X = [x_1, \dots, x_i, \dots, x_n] \in \mathbb{R}^{d \times n}$. Using unsupervised dimension reduction and unsupervised clustering algorithms of the activity data matrix, we obtain a set of typical daily activity patterns. Principal component analysis (PCA), sparse coding (SC) [35], and an autoencoder (AE) [36] are used for dimension reduction. Additionally, since the

number of activity clusters is unknown, we use X-means [37] as a clustering algorithm.

1) PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA is commonly used for dimension reduction with a transformation matrix $\mathbf{A} \in \mathbb{R}^{d \times d_p}$ by projecting each input data sample x_i in dataset X into a lower-dimensional data sample $x^* \in \mathbb{R}^{d_p}$. Let the lower-dimensional dataset be X^* ; then, the transformation is written as $x^* = \mathbf{A}^T x$ and so $X^* = \mathbf{A}^T X$. In the setting of dimension reduction for activity analysis, X is set as the activity data matrix. In PCA, A is optimized so that the variance of projected data samples x^* is maximized. It is well known that the optimized A can be obtained as an eigenvector ev of the covariance matrix of the input data set (X), so \mathbf{A} is represented as $\mathbf{A} = [ev_1, \dots, ev_{d_p}]$, where ev_1 denotes the first principal component, which is the eigenvector corresponding to the largest eigenvalue. Therefore, \mathbf{A} is composed of eigenvectors corresponding to the largest d_p eigenvalues. We regard these first d_p principal components as typical activity patterns in all activity samples.

2) SPARSE CODING (SC)

The objective of the SC algorithm is to decompose the activity data matrix ($X \in \mathbb{R}^{d \times n}$) into d_s basic elements (dictionary) ($\mathbf{D} \in \mathbb{R}^{d \times d_s}$, $d_s < d$) and sparse coefficient arrays ($\alpha = [\alpha_1, \dots, \alpha_n] \in \mathbb{R}^{d_s \times n}$). In SC, we regard the d_s basic elements (dictionaries) as the typical activity patterns in the activity dataset. The optimization problem in SC is expressed as follows:

$$\arg \min_{\mathbf{D} \in \mathcal{C}, \alpha \in \mathbb{R}^{d_s \times n}} \sum_{i=1}^N \|x_i - \mathbf{D}\alpha_i\|^2 + \lambda \|\alpha_i\|_1, \quad (1)$$

where

$$\mathcal{C} \triangleq \{\mathbf{D} \in \mathbb{R}^{d \times d_s} : d_j^T d_j \leq 1, \forall j = 1, \dots, d_s\} \quad (2)$$

In the optimization setting, the coefficient α of the sparse decomposition is obtained by minimizing the error: $\|x_i - \mathbf{D}\alpha_i\|^2$ with ℓ_1 penalties: $\|\alpha_i\|_1$. λ is a model parameter that controls the trade-off between sparsity and minimization error.

To solve this optimization problem, we applied an alternative optimization technique, referred to as online dictionary learning [35], which optimizes with respect to each of the two variables \mathbf{D} and α when the other variable is fixed.

3) AUTOENCODER (AE)

The AE is used to transform the data into a lower-dimensional space of feature vectors by an unsupervised learning neural network. The objective of the AE is essentially to learn the encoding matrix from the original input data ($x \in \mathbb{R}^d$) to low-dimensional space (\mathbb{R}^{d_a} , $d_a < d$) as close as possible to its original input data. The optimization problem in the AE is expressed as follows:

$$\arg \min_{\mathbf{W} \in \mathbb{R}^{d \times d_a}} \sum_{i=1}^N \|x_i - \sigma(\mathbf{W}^T \sigma(\mathbf{W}x_i + b_1) + b_2)\|^2, \quad (3)$$

where $\sigma(\cdot)$ is the activation function in the neural network. We applied the rectified linear unit function as $\sigma(\cdot)$, and $b_1 \in \mathbb{R}^{d_a}$ and $b_2 \in \mathbb{R}^d$ are bias vectors. The \mathbf{W} , b_1 , b_2 are optimized using the backpropagation algorithm based on the adaptive moment estimation (Adam) optimizer [38] in this study.

4) X-MEANS

Dimension reduction is a qualitatively easier representation to clearly separate the data clusters. To identify typical activity patterns, clustering algorithms can be used to assign the activity samples into K clusters. As a feature representation method, we used the X-means algorithm [37], which is an extension of the K-means algorithm with the ability to choose the number of clusters (K) based on information criteria. First, the optimization problem in both K-means and X-means is set as follows:

$$\arg \min_{\mathbf{R} \in \mathbb{R}^{N \times d_k}, \mu \in \mathbb{R}^{d \times d_k}} \sum_{i=1}^N \sum_{k=1}^{d_k} r_{i,k} \|x_i - \mu_k\|^2, \quad (4)$$

$$r_{i,k} = \begin{cases} 1, & (k = \arg \min_j \|x_i - \mu_j\|^2) \\ 0, & (k \neq \arg \min_j \|x_i - \mu_j\|^2). \end{cases} \quad (5)$$

μ_k are d -dimensional centroid vectors of the k th cluster. $r_{i,k}$ is the binary indicator variable. If the i th sample is assigned to the k th cluster, $r_{i,k} = 1$; otherwise, $r_{i,k} = 0$. To solve this optimization problem, an alternative optimization technique is applied to optimize with respect to each of the two variables $r_{i,k}$ and μ when the other variable is fixed. In X-means [37], the optimal number of clusters X is determined by recursively running 2-means until the Bayesian information criterion (BIC) stops improving.

5) HYPERPARAMETER SETTING

For dimension reduction by PCA, we set the cumulative contribution rate $th_c = 0.9$ as the threshold. For SC, we set the basis of the dictionary $d_s = 100$ and $\lambda = 1$. For the AE, we set the dimension of the middle layer d_a to 20, the number of epochs to 100, and the minibatch size to 64.

C. TRANSFORMATION TO PARTICIPANT-BASED FEATURE VECTOR

To train the classification models of MMSE by using a data set that consists of a set of samples per participant, we need to convert the daily activity dataset X_M of person ID M , which is composed of feature vectors per day, into the person-based activity dataset XP with a feature vector per participant (Fig 3).

We calculate the statistics of the daily activity features. Statistical features (called ‘‘Stats’’), which are directly calculated from the daily sensor features, comprise a comparative method with the proposed feature extraction methods based on typical activity patterns.

Consider the daily activity matrix (sample set) of a participant whose ID is M represented as X_M , composed of a daily

activity sample $\{x_1, x_2, \dots, x_{N_M}\}$, where N_M is the number of measurement days of the participant. We describe how to convert $X_M \in \mathbb{R}^{N_M \times d}$ to the person-based feature vector $XP_M \in \mathbb{R}^d$ with five types of features: (i) statistical features, and features extracted based on (ii) PCA, (iii) SC, (iv) AE, and (v) X-means in Section IV-B.

(i) XP with statistical features: To represent the representative feature of a person, we calculated the mean vector $mean(X_M)$ of X_M and standard deviation vector $std(X_M)$ of X_M along the dimension of a day. The person-based feature vector XP_M is represented by concatenating $mean(X_M)$ and $std(X_M)$ as $XP_M = [mean(X_M)^T, std(X_M)^T]^T$.

(ii) XP extracted based on PCA: X_M is converted to lower-dimensional data $X_M^* = A^T X_M$ by projection for dimension reduction with PCA. The person-based feature vector XP_M is represented by concatenating $mean(X_M^*)$ and $std(X_M^*)$, $XP_M = [mean(X_M^*)^T, std(X_M^*)^T]^T$.

(iii) XP extracted based on SC: After training the SC model, all activity samples X are decomposed into d_s basic elements (\mathbf{D}) and sparse coefficient arrays ($\alpha \in \mathbb{R}^{d_s \times N}$). Coefficient values $\alpha_{k,n}$ in α , which are almost zero, are implemented to denote the weight of the k th basic element to represent the n th activity sample.

The corresponding $\alpha_M^* \in \mathbb{R}^{d_s \times N_M}$ is calculated by encoding the data set X_M as a sparse combination of the basic elements \mathbf{D} , so we set $X_M^* = \alpha_M^*$. The person-based feature vector XP_M is calculated as $XP_M = [mean(X_M^*)^T, std(X_M^*)^T]^T$ in the same manner as PCA.

(iv) XP extracted based on AE: After the AE model is trained, the activity matrix X_M is encoded into X_M^* as follows:

$$X_M^* = \sigma(\mathbf{W}X_M + B_1),$$

where

$$B_1 = [b_1, \dots, b_1] \in \mathbb{R}^{d_a \times N_M}. \quad (6)$$

The person-based feature vector XP_M is calculated as $XP_M = [mean(X_M^*)^T, std(X_M^*)^T]^T$ in the same manner as PCA.

(v) XP extracted based on X-means: After the X-means model is trained, the indicator variable $r_{i,k}$ is used to convert X_M to XP_M . Let the indicator matrix, which has $r_{i,k}$ as each value, be $R \in \mathbb{R}^{d_a \times N}$. The indicator matrix for activity samples of person ID M can be written by $R_M \in \mathbb{R}^{d_a \times N_M}$. XP_M is calculated as $XP_M = sum(R_M)$.

In this paper, we refer to the extracted activity features based on the daily activity patterns by PCA, SC, AE and X-means as BoAP features. In total, five feature sets were acquired by extracting the Stats features and four BoAP features for each participant.

V. EXPERIMENT

We perform a two-class classification task on the high/low MMSE scores of the experimental participants and evaluate the classification accuracy in two different experiments. First, to analyze the effectiveness of each feature group, we train classification models with five sets of features for each type of sensor and report the classification accuracies of models with

each feature group (Experiment 1). In the first experiment, we used all samples obtained from each sensor.

Second, we compare the feature groups of the different sensors and analyze the effectiveness of fusing features from multiple sensors (Experiment 2). For this purpose, we used only daily activity samples, which consist of data observed from all sensors. Samples for which sensor data are missing are removed from the dataset in Experiment 2.

Third, we compare the accuracy of the classification models based on daily activity samples and classification models using person-based feature vectors to clarify the effectiveness of the proposed BoAP features (person-based features) extracted for each participant in Experiment 3.

A. EXPERIMENTAL SETTING

Leave-one-person-out cross-validation is conducted to evaluate the performance. The balanced accuracy (mean accuracy for both classes) is selected as the evaluation criterion for the classification because the numbers of samples are imbalanced. The majority baseline when all samples are classified into one majority class is set to 50%.

We explain the machine learning models for the classification of the MMSE score as follows.

1) PERSON-BASED BoAP APPROACH

As classification models to be trained with the proposed BoAP features (person-based features), we use a linear support vector machine (SVM), a logistic regression (LR) model, and a random forest (RF) model. Because the sample size is equal to the number of participants in our approach (using person-based features) and the sample size is small, we do not use a nonlinear classifier such as a deep neural network (DNN) or a long short-term memory (LSTM) network [39], which would require many training samples, in Experiments 1 and 2. We normalize the data such that each feature has a mean of zero and a standard deviation of one.

The parameters of the SVM are optimized using a nested cross-validation scheme, with C parameter values selected from [0.01, 0.1, 1, and 10]. The parameters of the RF are optimized similarly using a nested cross-validation scheme, with the numbers of trees per forest selected from [10, 100, and 200]. The number of random samples per tree is set as the square root of the size of the training sample set. For the feature sets, independent-sample t-tests between the activity features and the high/low scores on the dementia scale are performed for feature selection. Variables with significant differences based on $p < 0.1$ are selected for the training process.

2) DAY-SAMPLE-BASED APPROACH

Our proposed method for classifying each person's dementia scale score is an approach that effectively extracts a person's activity features from a daily activity sample and then estimates the MMSE from the daily activities of each person. However, such a person-based classification method

TABLE 3. Experiment 1: Balanced accuracy of the model for each sensor feature set. Row 2 denotes the method of feature extraction. The features (excluding Stats) are the proposed method’s BoAPs, which represent the difference in activity patterns.

	Door					Human					Location					Sleep				
	Stats	BoAPs				Stats	BoAPs				Stats	BoAPs				Stats	BoAPs			
		PCA	SC	AE	X-means		PCA	SC	AE	X-means		PCA	SC	AE	X-means		PCA	SC	AE	X-means
SVM	0.502	0.580	0.647	0.513	0.612	0.475	0.675	0.670	0.525	0.595	0.511	0.524	0.646	0.473	0.694	0.511	0.524	0.521	0.486	0.631
LR	0.459	0.525	0.516	0.592	0.560	0.602	0.725	0.614	0.625	0.586	0.596	0.559	0.559	0.473	0.595	0.510	0.534	0.534	0.497	0.552
RF	0.447	0.524	0.481	0.524	0.612	0.659	0.725	0.689	0.514	0.550	0.585	0.612	0.510	0.433	0.609	0.581	0.513	0.513	0.386	0.511
Mean	0.469	0.543	0.548	0.543	0.595	0.579	0.708	0.658	0.555	0.577	0.564	0.565	0.572	0.460	0.632	0.534	0.524	0.523	0.456	0.565

TABLE 4. Experiment 2: Comparison of the classification accuracy of each sensor feature set.

	D	H	L	S	D+H	D+L	D+S	H+L	H+S	L+S	D+H+L	D+H+S	D+L+S	H+L+S	ALL
Stats	0.570	0.435	0.637	0.685	0.586	0.605	0.602	0.653	0.685	0.801	0.605	0.570	0.602	0.801	0.602
PCA	0.503	0.489	0.753	0.669	0.718	0.653	0.669	0.753	0.669	0.852	0.653	0.669	0.669	0.669	0.669
SC	0.586	0.554	0.538	0.737	0.702	0.685	0.551	0.522	0.618	0.538	0.570	0.702	0.470	0.570	0.704
AE	0.720	0.409	0.669	0.637	0.522	0.737	0.527	0.737	0.605	0.704	0.637	0.589	0.785	0.704	0.540
X-means	0.573	0.516	0.723	0.723	0.589	0.823	0.589	0.790	0.672	0.772	0.772	0.508	0.871	0.868	0.626
Mean	0.590	0.481	0.664	0.690	0.623	0.701	0.588	0.691	0.650	0.733	0.647	0.608	0.679	0.722	0.628

cannot provide sufficient samples for classification training because the number of samples is equal to the number of participants.

In Experiment 3, we analyze the effectiveness of our approach using person-based features by comparison with classification models trained based on daily activity samples (see the left-hand side of Fig 3). In the approach using the daily samples, daily activity samples observed from a participant are assigned to the MMSE label of that participant. In this approach, all daily samples (total N=1034) are available as training data; thus, for the approach using the daily samples, we use DNN and LSTM classifiers in addition to SVM, LR, and RF models.

For the DNN, we use a fully connected feedforward neural network consisting of two hidden layers, with dropout after each layer. We set the number of units in each hidden layer to 100, the number of epochs to 100, the minibatch size to 64, and the dropout rate to 0.5. We use the Adam optimizer and set the learning rate to 0.001. For the LSTM classifier, we stack a sigmoid layer on top of the LSTM neuron to accomplish two-class classification. We set the number of hidden dimensions equal to 20 and apply dropout before the fully connected layer. In the LSTM case, the sensor features are learned as time-series features with a length of 24 ($N_D \times 24$) corresponding to the feature vectors extracted for each hour (24 hours in total). The other network settings are the same as for the DNN. In the inference phase of the day-based approach, a class label (dementia or no dementia) is output by the trained model for each day sample; therefore, we need to convert the label set (inferred class) for each day sample into labels for the participants. We use a majority voting method for this purpose, meaning that the class label that is assigned for the greater number of days is taken as the final dementia scale classification result for the participant. The features used for the training process are selected using independent-sample t-tests, similar to the approach used in the person-based method, except for the LSTM classifier.

B. EXPERIMENT 1: CLASSIFICATION WITH DAILY ACTIVITIES COLLECTED BY EACH SENSOR

To discuss the effectiveness of the feature groups, we trained models with five sets of features for each type of sensor. Table 3 shows the classification results of these models. For the door, human motion, location, and sleep sensors, the highest model accuracies (0.647, 0.725, 0.694, and 0.631, respectively) outperform the classification accuracy of the baseline (0.5). This result shows that each obtained sensor data point has a certain effectiveness in classifying dementia scale scores. Comparing the mean accuracy of each machine learning model, the X-means algorithm is the most effective for door sensors, indoor positioning, and sleep sensors (0.595, 0.632, and 0.565, respectively), while PCA is especially effective for human motion sensors (0.708).

The highest accuracy among all the sensors was obtained with the features extracted based on activity patterns. Each model improves the accuracy by 0.066-0.183 points compared to the same machine learning model in Stats. This result indicates that BoAPs are more effective features for dementia scale score classification than Stats are, which are extracted directly from the sensor features. Among the machine learning models, the highest accuracy for all sensors, except the human motion sensor, was achieved by the linear SVM, so only the linear SVM was used in the following experiments.

C. EXPERIMENT 2: FUSING THE MULTISENSOR FEATURES

In this section, we compare the feature groups of the different sensors and analyze the effectiveness of fusing features from multiple sensors. Although different participants provided data for each installed sensor, 37 participants agreed to have their data recorded by all the sensors. Because the measurement dates differed for each sensor, we extracted only the daily activity sample collected by all four sensors for each participant. The total number of days of measurement data acquired from all experimental participants was 1034.

TABLE 5. Comparison of the highest classification accuracy between the single-sensor scenario and the multisensor scenario. Column 4 shows the number of improved model in the multisensor scenario.

	Single-sensor	Multisensor	Improved
Stats	0.685	0.801	4/11
PCA	0.753	0.852	2/11
SC	0.737	0.704	2/11
AE	0.720	0.785	5/11
X-means	0.723	0.871	7/11

TABLE 6. Confusion matrix for the highest accuracy in Experiment 2: Estimation results of the classification model for each participant's high/low MMSE scores.

		Estimated	
		High	Low
Actual	High	6	0
	Low	8	23

We evaluated the balanced accuracy of the four types of sensor features: door features (D), human motion features (H), location features (L), and sleep features (S). The fused features of each sensor produced 15 features. Five linear SVM models, which are based on the activity pattern forms in each case, are built for evaluation.

Table 4 shows the classification accuracy for the fusion of each feature set. In Columns 2-5, when we compare the accuracy of each sensor by using the common measurement data, the features extracted from the location and sleep sensors (the mean accuracy values are 0.664 and 0.690, respectively) are more effective than those from the door and human motion sensors (the mean accuracy values are 0.590 and 0.481, respectively).

Columns 6-16 show that the model can be improved by the fusion of features obtained from multiple sensors. Row 2 shows that among the 11 multisensor patterns, the classification accuracy of four sensor combinations (D+H, H+L, L+S, and H+L+S) in Stats is better than that of any of the single sensors selected for fusion. Row 3 shows that the classification accuracy of the PCA model is higher for the combination of two sensors (D+H, L+S). Similarly, rows 4-6 show that the model improves in two cases for SC (D+H, D+L), five cases for AE (D+L, H+L, L+S, D+L+S, H+L+S), and seven cases for X-means (D+H, D+L, H+L, L+S, D+H+L, D+L+S, H+L+S). Table 5 summarizes the highest accuracy of each single-sensor and multisensor case in Table 4 and the number of sensor combinations whose models were improved by sensor fusion. These results indicate that the model can be improved by the fusion of features obtained from multiple sensors. The highest accuracy of 0.871 is obtained by fusing the door, location, and sleep features and clustering the activity patterns by the X-means algorithm. Table 6 shows the confusion matrix; 8 of the 31 participants in the dementia group were incorrectly identified as being healthy. The highest mean accuracy of 0.773 in Experiment 2 is obtained for the combination of the location and sleep sensors. Table 7 shows the classification accuracy of each machine learning model for the combination of the location and sleep sensors, and the highest accuracy is obtained with the features extracted based on activity patterns when using any machine learning model.

TABLE 7. Comparison of machine learning models for the feature combination that yielded the highest mean accuracy in Experiment 2.

	Stats	PCA	SC	AE	X-means
SVM	0.801	0.852	0.538	0.704	0.772
LR	0.718	0.769	0.602	0.554	0.677
RF	0.702	0.677	0.685	0.487	0.750
Mean	0.740	0.766	0.608	0.582	0.733

TABLE 8. Experiment 3: Comparison of classification accuracy of models trained on day-based and person-based samples.

	Day-based	Person-based
SVM	0.634	0.852
LR	0.618	0.769
RF	0.720	0.750
DNN	0.720	-
LSTM	0.667	-

D. EXPERIMENT 3: CLASSIFICATION PER PERSON VS. DAY SAMPLES

We compare the proposed person-based BoAP method with the day-sample-based (day-based) method to clarify the effectiveness of BoAP for the feature representation of each participant. For day-based classification, the activity sample set (N=1034) can be used to train the classification models, so the number of samples meets the minimum requirement for nonlinear algorithms, including DNNs. This section compares the accuracy of the day-based and person-based classification methods with the fusion of indoor positioning and sleep features, which was the sensor combination with the highest mean accuracy in Experiment 2. In day-based classification, we evaluate the classification accuracy of models trained on daily samples using DNN and LSTM classifiers in addition to SVM, LR, and RF classifiers. Each classification model is trained on the daily sensor features and outputs classification results for the participants for all days.

Table 8 compares the classification accuracy results of the day-based and person-based methods for each machine learning model. In day-based classification, the RF and DNN models obtain the highest accuracy of 0.720. The results of the person-based method show the highest accuracy for each machine learning model in Table 7, and the accuracy of all machine learning models is higher with the person-based method than with the day-based method. These results indicate the importance of classifying dementia based on typical activity features for each person rather than using each day as a sample.

The reason why the person-based method yields better accuracy even though the number of samples is small is discussed as follows. The hypothesis behind the sample-based method is that all daily samples observed from a participant with dementia characterize dementia because all daily samples are annotated as dementia samples in this case. However, this hypothesis might be too strong when modeling the specific behavioral features of dementia because it is unclear what and when specific activity patterns of participants with dementia are observed, and not all daily samples (daily behavior) always display the characteristics of dementia. This finding is consistent with the result (Table 7) that the sample-based method is less effective for the dementia

TABLE 9. Ablation results for the feature groups.

		Stats		PCA	
All feature		0.801		0.852	
Removed feature group		Accuracy	Difference	Accuracy	Difference
Location	Probability of staying	0.669	+0.132	0.621	+0.231
	Moving distance	0.801	± 0	0.489	+0.363
Sleep	In/not in bed	0.785	+0.016	0.769	+0.083
	Heart rate	0.718	+0.083	0.702	+0.150
	Respiratory rate	0.785	+0.016	0.685	+0.167
	Body movement	0.769	+0.032	0.734	+0.118

TABLE 10. Balanced classification accuracy for different measurement times.

	Door	Human	Location	Sleep
Time	0.612	0.595	0.694	0.631
0:00-3:00	0.464	0.591	0.569	0.546
3:00-6:00	0.463	0.505	0.277	0.595
6:00-9:00	0.535	0.509	0.519	0.348
9:00-12:00	0.569	0.500	0.518	0.492
12:00-15:00	0.435	0.586	0.479	0.578
15:00-18:00	0.469	0.480	0.492	0.437
18:00-21:00	0.376	0.509	0.657	0.662
21:00-0:00	0.473	0.364	0.599	0.638

classification task than the person-based method. From the experiments reported in this section, we conclude that the person-based method of learning effective feature representations from daily samples of participants with dementia is effective for dementia classification.

E. FEATURE ANALYSIS

To discuss the feature groups that contribute to classification accuracy, we conduct feature ablation tests while fusing indoor positioning with sleep features, which is the highest mean accuracy sensor combination in Experiment 2. Table 9 shows the classification accuracy when one feature group is removed from all the features used in the case of Stats and PCA, where the accuracy is high when all feature groups are used. The accuracies of the linear SVM models, excluding the features related to the probability of staying in each area and the moving distance, are particularly low. This result indicates that the indoor location information of the participants is important for accurate classification. Rows 6-9 show that the accuracy of the model decreases when the sleep features related to heart rate, respiration, and body movement are excluded, as well as the state of being in bed. This result indicates that biological information during sleep also has the potential to improve the model accuracy.

Next, to analyze the effective measurement time for classification, we compare the model accuracy when using only the features of each time period in the X-means algorithm of Experiment 1. Table 10 shows the linear SVM model accuracy when using only three hours of the sensor features extracted for each hour in Section IV-A. The classification accuracy of each time for the location and sleep sensors shows that an accurate model can be obtained when the features after 18:00 are applied. This result suggests that monitoring

activity at night is effective for estimating the dementia scale score category.

VI. DISCUSSION

In this study, we compare multiple sensors from various aspects, with a focus on introducing sensor networks for estimating the dementia scale score in an actual living environment.

Indoor positioning can provide an approximate location of the participants at each time. More accurate indoor positioning algorithms using the RSSIs are surveyed in [33]. In general, we need to develop a large-scale reference point set to estimate accurate positions. However, we aimed to be as independent of the collection environment as possible and applied a simple method to extract features. In this study, location information based on indoor positioning contributed the most to classification, but participants needed to carry a beacon at all times during sensing. Because the sensor system is expected to be applied in living facilities with different floor plans, some challenges arise in extracting features that are unified under different environments.

Although door and human motion sensors are relatively inexpensive and easy to install and can collect movement information for each participant, the accuracy of estimating the dementia scale in this experiment was lower than the accuracies of the indoor positioning and sleep sensors. Similar to indoor positioning, some difficulties are encountered in realizing an estimation model that can be adapted to different living facilities. Furthermore, sensors that measure activity data may also detect the behaviors of individuals other than the subject. Because the sensors are not associated with the participants themselves, the approach is available only for those participants who live alone or in a private room that utilizes door and human sensors.

In this experiment, sleep data, including biological information recorded by bed sensors, were effective for classifying dementia scale scores. The sleep data acquired by the noncontact sensor are the least affected by the monitoring environment among the sensors used in this experiment, so it is relatively easy to continue collecting data over a long time in a stable sensing environment. Collecting sleep data from elderly individuals can be utilized for various nursing care support systems, such as the analysis of sleep stages and sleep quality [40], [41]. In future works, to identify patterns according to dementia scale scores, we will focus on building more elaborate models using sleep sensor data by extracting features that represent the differences in detailed sleep patterns. The sensing method may be effective in support systems, including dementia detection.

In this study, we created two classes based on MMSE scores, but we did not estimate the results of doctors' diagnoses, and the experimental participants with extremely low MMSE scores included bedridden persons. Brodaty *et al.* [42] showed that MCI symptoms do not necessarily progress to dementia and that recovery can be expected. Early detection of MCI is much more important

than estimating MMSE scores in support systems for elderly people. Because of the small number of participants with MCI in this dataset (refer to Fig 2), we did not attempt to address the classification of MCI. Based on this study, future studies should focus on long-term data collection and tracking of changes in the MMSE for participants with MCI using a sensing system that can be implemented in actual living environments.

Moreover, it is difficult to understand what specific behavioral characteristics are related to the feature values extracted based on activity patterns. When assuming the development of a support system for medical professionals to analyze dementia scale scores, it is desirable to use interpretable feature values, such as the time asleep and amount of movement. Detailed analyses of the features extracted using dimensionality reduction algorithms such as PCA and AEs are provided in [43], [44]. Future work should focus on the visualization of features based on activity patterns that can be applied to estimate dementia accurately.

VII. CONCLUSION

In this study, we collected activity sensor datasets by installing multiple sensors in participants' residential facilities and built a dementia scale score estimation model. We also proposed a method for extracting features based on activity patterns to automatically acquire differences in daily activity types, which are expected to be effective at detecting dementia from sensor data.

In Experiment 1, the accuracy of all the sensors was higher than the baseline, thus showing that the four sensing methods used in this experiment were effective at estimating the dementia scale. Experiment 2 showed that indoor positioning was particularly effective for collecting behavioral data, as it could capture general activities all day, and that activity sensing using sleep sensors was particularly effective at recording sleep rhythms and detailed sleep states. Experiment 3 showed, through comparison with the day-based method, that the person-based method learns effective feature representations for dementia classification from participants' sample data. Throughout the three experiments, the activity-pattern-based features were particularly effective. The extracted activity vectors represent differences in daily activity patterns and can be regarded as effective features for classifying dementia scale scores.

We aimed to provide insight into the types of sensors and activity features that may be applied for dementia classification. Future tasks will include the development of sensing methods that can be adapted to different living facilities, the construction of models for estimating dementia levels, and the analysis of activity pattern features that are effective for improving the classification accuracy. However, this work has demonstrated the possibility that easily installed and unobtrusively collected in-home behavior data can estimate the dementia scale score. Further model improvement can be expected by collecting larger datasets for location and sleep information. We conclude that our results suggest the

possibility of analyzing the activity patterns specific to dementia and realizing an early dementia detection system by feasible data collection in a real-world environment.

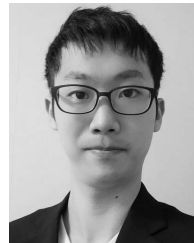
ACKNOWLEDGMENT

The authors would like to thank the cooperation of the business design division at Nippon Telegraph and Telephone West Corporation.

REFERENCES

- [1] World Health Organization. (2020). *Fact Sheets: Dementia*. [Online]. Available: <http://www.who.int/en/news-room/fact-sheets/detail/dementia>
- [2] M. Prince, A. Comas-Herrera, M. Knapp, M. Guerchet, and M. Karagiannidou, "World Alzheimer report 2016: Improving healthcare for people living with dementia: Coverage, quality and costs now and in the future," *Alzheimer's Disease Int.*, London, U.K., Tech. Rep., 2016.
- [3] K. G. Yiannopoulou and S. G. Papageorgiou, "Current and future treatments for Alzheimer's disease," *Therapeutic Adv. Neurol. Disorders*, vol. 6, no. 1, pp. 19–33, 2013.
- [4] S. Robben, G. Englebienne, and B. Kröse, "Delta features from ambient sensor data are good predictors of change in functional health," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 4, pp. 986–993, 2017.
- [5] A. Alberdi, A. Weakley, M. Schmitter-Edgecombe, D. J. Cook, A. Aztiria, A. Basarab, and M. Barrenechea, "Smart home-based prediction of multidomain symptoms related to Alzheimer's disease," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 6, pp. 1720–1731, 2018.
- [6] N. Eagle and A. Pentland, "Eigenbehaviors: Identifying structure in routine," *Behav. Ecol. Sociobiol.*, vol. 63, no. 7, pp. 1057–1066, 2009.
- [7] N. K. Al-Qazzaz, S. H. B. M. Ali, S. A. Ahmad, K. Chellappan, M. S. Islam, and J. Escudero, "Role of EEG as biomarker in the early detection and classification of dementia," *Sci. World J.*, vol. 2014, pp. 1–16, 2014.
- [8] E. Q. Wu, D. Hu, P.-Y. Deng, Z. Tang, Y. Cao, W.-M. Zhang, L.-M. Zhu, and H. Ren, "Nonparametric Bayesian prior inducing deep network for automatic detection of cognitive status," *IEEE Trans. Cybern.*, vol. 51, no. 11, pp. 5483–5496, 2021.
- [9] E. Q. Wu, P. Xiong, Z.-R. Tang, G.-J. Li, A. Song, and L.-M. Zhu, "Detecting dynamic behavior of brain fatigue through 3-D-CNN-LSTM," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 52, no. 1, pp. 90–100, 2022.
- [10] E. Q. Wu, L.-M. Zhu, G.-J. Li, H.-J. Li, Z. Tang, R. Hu, and G.-R. Zhou, "Nonparametric hierarchical hidden semi-Markov model for brain fatigue behavior detection of pilots during flight," *IEEE Trans. Intell. Transp. Syst.*, early access, 2021.
- [11] E. Q. Wu, X.-Y. Peng, S.-D. Chen, X.-Y. Zhao, and Z.-R. Tang, "Detecting Alzheimer's dementia degree," *IEEE Trans. Cognit. Develop. Syst.*, vol. 14, no. 1, pp. 116–125, 2022.
- [12] C. P. Guruge, S. Oviatt, P. D. Haghighi, and E. Pritchard, "Advances in multimodal behavioral analytics for early dementia diagnosis: A review," in *Proc. Int. Conf. Multimodal Interact. (ICMI)*, 2021, pp. 328–340.
- [13] W. Jarrold, B. Peintner, D. Wilkins, D. Vergryi, C. Richey, M. L. Gorno-Tempini, and J. Ogar, "Aided diagnosis of dementia type through computer-based analysis of spontaneous speech," in *Proc. Workshop Comput. Linguistics Clin. Psychol., From Linguistic Signal Clin. Reality*, 2014, pp. 27–37.
- [14] M. R. Hodges, N. L. Kirsch, M. W. Newman, and M. E. Pollack, "Automatic assessment of cognitive impairment through electronic observation of object usage," in *Proc. Int. Conf. Pervasive Comput.* Berlin, Germany: Springer, 2010, pp. 192–209.
- [15] T. L. Hayes, F. Abendroth, A. Adami, M. Pavel, T. A. Zitzelberger, and J. A. Kaye, "Unobtrusive assessment of activity patterns associated with mild cognitive impairment," *Alzheimer's Dementia*, vol. 4, no. 6, pp. 395–405, 2008.
- [16] 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.
- [17] P. N. Dawadi, D. J. Cook, and M. Schmitter-Edgecombe, "Automated cognitive health assessment from smart home-based behavior data," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 4, pp. 1188–1194, 2016.
- [18] S. Duggan, T. Blackman, A. Martyr, and P. Van Schaik, "The impact of early dementia on outdoor life: A 'shrinking world?'" *Dementia*, vol. 7, no. 2, pp. 191–204, 2008.

- [19] H.-X. Wang, A. Karp, B. Winblad, and L. Fratiglioni, "Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: A longitudinal study from the kungsholmen project," *Amer. J. Epidemiol.*, vol. 155, no. 12, pp. 1081–1087, 2002.
- [20] S. Okada, K. Inoue, T. Imai, M. Noguchi, and K. Kuwamura, "Dementia scale classification based on ubiquitous daily activity and interaction sensing," in *Proc. Int. Conf. Affect. Comput. Intell. Interact. (ACII)*, 2019, pp. 192–198.
- [21] B. Guarnieri et al., "Prevalence of sleep disturbances in mild cognitive impairment and dementing disorders: A multicenter Italian clinical cross-sectional study on 431 patients," *Dementia Geriatric Cognit. Disorders*, vol. 33, no. 1, pp. 50–58, 2012.
- [22] L. Peter-Derex, P. Yammine, H. Bastuji, and B. Croisile, "Sleep and Alzheimer's disease," *Sleep Med. Rev.*, vol. 19, pp. 29–38, 2015.
- [23] E. A. Hahn, H.-X. Wang, R. Andel, and L. Fratiglioni, "A change in sleep pattern may predict Alzheimer disease," *Amer. J. Geriatric Psychiatry*, vol. 22, no. 11, pp. 1262–1271, 2014.
- [24] K. A. I. Aboalayon, M. Faezipour, W. S. Almuhammadi, and S. Moslehpour, "Sleep stage classification using EEG signal analysis: A comprehensive survey and new investigation," *Entropy*, vol. 18, no. 9, p. 272, 2016.
- [25] J. Zhou, G. Wang, J. Liu, D. Wu, W. Xu, Z. Wang, J. Ye, M. Xia, Y. Hu, and Y. Tian, "Automatic sleep stage classification with single channel EEG signal based on two-layer stacked ensemble model," *IEEE Access*, vol. 8, pp. 57283–57297, 2020.
- [26] R. Paradiso, T. Faetti, and S. Werner, "Wearable monitoring systems for psychological and physiological state assessment in a naturalistic environment," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2011, pp. 2250–2253.
- [27] X. Nam, Y. Kim, and J. Lee, "Sleep monitoring based on a tri-axial accelerometer and a pressure sensor," *Sensors*, vol. 16, no. 5, p. 750, 2016.
- [28] A. Tal, Z. Shinar, D. Shaki, S. Codish, and A. Goldbart, "Validation of contact-free sleep monitoring device with comparison to polysomnography," *J. Clin. Sleep Med.*, vol. 13, no. 3, pp. 517–522, 2017.
- [29] 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. Psychiatric Res.*, vol. 12, no. 3, pp. 189–198, 1975.
- [30] M. Sugishita, Y. Koshizuka, S. Sudou, and E. A. Sugishita, "The validity and reliability of the Japanese version of the mini-mental state examination (MMSE-J) with the original procedure of the attention and calculation task (2001) (in Japanese)," *Jpn. J. Cognit. Neurosci.*, vol. 20, no. 2, pp. 91–110, 2018.
- [31] J. Merilahti, P. Viramo, and I. Korhonen, "Wearable monitoring of physical functioning and disability changes, circadian rhythms and sleep patterns in nursing home residents," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 3, pp. 856–864, 2016.
- [32] G. Sprint, D. J. Cook, R. Fritz, and M. Schmitter-Edgecombe, "Using smart homes to detect and analyze health events," *Computer*, vol. 49, no. 11, pp. 29–37, 2016.
- [33] D. Riboni, C. Bettini, G. Civitarese, Z. H. Janjua, and R. Helaoui, "Fine-grained recognition of abnormal behaviors for early detection of mild cognitive impairment," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, 2015, pp. 149–154.
- [34] P. N. Dawadi, D. J. Cook, and M. Schmitter-Edgecombe, "Modeling patterns of activities using activity curves," *Pervas. Mobile Comput.*, vol. 28, pp. 51–68, 2016.
- [35] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2009, pp. 689–696.
- [36] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [37] P. Dan and A. W. Moore, "X-means: Extending k-means with efficient estimation of the number of clusters," in *Proc. Int. Conf. Mach. Learn. (ICML)*, vol. 1, 2000, pp. 727–734.
- [38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015.
- [39] A. Graves, "Long short-term memory," in *Supervised Sequence Labelling With Recurrent Neural Networks*. Berlin, Germany: Springer, 2012, pp. 37–45.
- [40] R. Yi, M. Enayati, J. M. Keller, M. Popescu, and M. Skubic, "Non-invasive in-home sleep stage classification using a ballistocardiography bed sensor," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, 2019, pp. 1–4.
- [41] R. Sadeghi, T. Banerjee, J. C. Hughes, and L. W. Lawhorne, "Sleep quality prediction in caregivers using physiological signals," *Comput. Biol. Med.*, vol. 110, pp. 276–288, 2019.
- [42] H. Brodaty, M. Heffernan, N. A. Kochan, B. Draper, J. N. Trollor, S. Reppermund, M. J. Slavin, and P. S. Sachdev, "Mild cognitive impairment in a community sample: The Sydney memory and ageing study," *Alzheimer's Dementia*, vol. 9, no. 3, pp. 310–317, 2013.
- [43] M. M. López, J. Ramírez, J. M. Górriz, I. Álvarez, D. Salas-Gonzalez, F. Segovia, and R. Chaves, "SVM-based CAD system for early detection of the Alzheimer's disease using kernel PCA and LDA," *Neurosci. Lett.*, vol. 464, no. 3, pp. 233–238, 2009.
- [44] E. Q. Wu, G.-R. Zhou, L.-M. Zhu, C.-F. Wei, H. Ren, and R. S. F. Sheng, "Rotated sphere Haar wavelet and deep contractive auto-encoder network with fuzzy Gaussian SVM for pilot's pupil center detection," *IEEE Trans. Cybern.*, vol. 51, no. 1, pp. 332–345, 2021.



AKIRA MINAMISAWA received the B.S. degree in electrical engineering from the Tokyo University of Science, in 2019, and the M.S. degree in information science from the Japan Advanced Institute of Science and Technology, in 2021. His research interests include social signal processing, machine learning, and deep learning.



SHOGO OKADA received the Ph.D. degree from the Tokyo Institute of Technology, Japan, in 2008. He directs the Social Signal and Interaction Group, Japan Advanced Institute of Science and Technology (JAIST), Japan, where he is currently an Associate Professor. In 2008 and 2011, he joined Kyoto University, Tokyo Institute of Technology, as an Assistant Professor. He joined the IDIAP Research Institute, Switzerland, as a Visiting Faculty Member, in 2014. His research interests include social signal processing, human dynamics, multimodal interaction, and machine learning. He is a member of the ACM.



KEN INOUE received the B.S. degree from Sophia University, Japan, in 2004, and the M.S. degree from the Tokyo Institute of Technology, Japan, in 2006. He is currently the Chief Executive Officer of George and Shaun, Inc.



MAMI NOGUCHI received the B.S. degree from the Department of Civil Engineering, Kyoto University, in 2004, and the M.S. degree from the Department of Civil and Earth Resources Engineering, Kyoto University, in 2006. She is currently the Group Manager of Toyota Connected Corporation, Japan.