This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2020.2985907, IEEE Journal of Biomedical and Health Informatics

Prediction of mild cognitive impairment using movement complexity

Taha Khan and Peter Jacobs

***Abstract*— *Objective:* Aimless movement or wandering may be a symptom of mild cognitive impairment (MCI) that arises as a consequence of confusion and forgetfulness. This paper presents a support vector machine (SVM) framework based on movement analysis for the prediction of the onset and progression of MCI. *Methods:* Movement data of 22 subjects with MCI, and 22 other healthy subjects, living independently in smart homes were collected for ten years using motion sensors. Features were extracted from the sensor data using movement metrics, including cyclomatic complexity, detrended fluctuation analysis, fractal index, entropy, and room transitions. Two different SVM classification algorithms were trained using the features, first to predict the progression of MCI in the post-transition period, and second to predict the onset of MCI in the pre-transition phase. *Results:* The two SVMs were able to detect the onset six months earlier than the clinical diagnosis. The model accuracy in classifying MCI increased monotonically from the onset month and reached maximum (81%) at the 11th post-transition month. The features of cyclomatic complexity contributed significantly to the prediction results. *Conclusion:* Findings support the use of movement complexity measures and machine learning for monitoring cognitive behavior in an independent living environment.**

***Index Terms*—Cyclomatic complexity, mild cognitive impairment, movement analysis, support vector machines.**

I. Introduction

Dementia is a degenerative neurological disorder that affects millions of people worldwide [1]. The symptoms include memory and concentration loss, change in behavior, disorientation, and confusion, typically when confronted with unfamiliar surroundings. Diagnosing dementia is difficult in the early stages since the symptoms resemble the natural signs of aging. Studies suggest that age is the most critical risk factor for dementia onset and is ubiquitous in approximately 5%–7% of people over the age of 60 [2].

A prominent symptom of dementia is the change in movement patterns that arise as a consequence of forgetfulness and confusion [3]. Previous literature suggests that changes in movement start years before the onset of dementia [4], suggesting that monitoring the movement could be useful to enable early diagnosis. Prior research indicates that people with mild cognitive impairment (MCI) show greater variability in movement patterns than those with severe impairment [5, 6]. MCI is an intermediate state between normal cognition and dementia. According to Birkett [7], severe changes in movement patterns occur during MCI that go away as dementia gets worse due to immobility that occurs with dementia severity.

Currently, no pharmacological treatment has been reported to cure dementia permanently [8]. With an increasing number of demented people, the management accounts for approximately 640 billion US dollars of direct and indirect costs annually [9]. On the other hand, the management imposes a high burden on caregivers. In a survey, 27% of caregivers reported feelings of burden and frustration [10]. The early detection of variability in movement patterns in MCI may allow early treatment that potentially can slow down dementia onset, hence improving the quality of life of the people with dementia (PWD) and delaying the transition to costly healthcare facilities.

Studies have been reported to quantify movement patterns to identify a decline in cognitive abilities indicative of dementia [3]. In order to correlate between movement and cognitive performance, two wandering dimensions, frequency, and temporal distribution, were estimated. The frequency was measured as the rate of wandering episodes per hour. Temporal distribution was measured as the time spent on wandering, the variability of locomotion in wandering, and room transitions. In one study, Hayes et al. [6] found that variability in walking time was higher among seven elderly seniors having MCI compared to seven other elderly seniors not having MCI. In another study, Makimoto et al. [11] employed RFID tagging systems to record two indicators of wandering locomotion, the percentage of hours with locomotion, and the median distance traveled per hour. They found that the percentage of hours with locomotion was negatively correlated with cognitive functions. In another study, Algase et al. [12] videotaped 44 wanderers in a long term setting. They measured the rate and duration of wandering episodes and found that the Mini-Mental State Examination (MMSE) scores were negatively correlated with the overall wandering duration and not significantly correlated with the rate-related parameters [11].

This manuscript was submitted for review on 21st October 2019. The work was supported by the National Institute of Health, USA, grants NIH R01 AG024059, P30 AG024978, P30 AG008017.

Taha Khan is with the department of Intelligent Systems at Halmstad University, Sweden (email: taha.khan@hh.se).

Peter Jacobs is an Associate Professor and director of the Artificial Intelligence for Medical Systems (AIMS) lab within the Department of Biomedical Engineering, Oregon Health and Science University, OR, USA (email: jacobsp@ohsu.edu).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2020.2985907, IEEE Journal of Biomedical and Health Informatics

Nams et al. [13] reported that the fractal dimension (a measure of randomness) of the movement paths of 14 elderly residents recorded over a month was significantly higher in individuals with lower MMSE scores (range: 0-30) indicating cognitive impairment. In a later study of 28 subjects [14] recorded over a month, the findings of Nams et al. [13] were replicated, and that the randomness of the paths taken by the elderly residents reliably differentiated PWD from those without the disorder.

Detrended fluctuation analysis (DFA) is a method to quantify self-similarity in movement [15]. The utilization of DFA for gait assessment relies on the fact that movement patterns are not random but exhibit dependencies such that the future variation in movement depends on the past variation. These dependencies are expressed in terms of long-range correlations within a time series that can be computed by DFA using a scaling exponent called α. The parameter α estimates the fractal scaling properties of the time series and can be used to describe the complexity of the process [16]. Hausdorff et al. [15] performed DFA on stride-interval time series to compare 17 subjects with Huntington's disease and 32 healthy subjects. They reported that the values of α were lower for patients with Huntington's disease compared to those for the healthy subjects (Huntington's disease: 0.60 ± 0.24; controls: 0.88 ± 0.17; $P < 0.005$). Moreover, they reported that α was linearly correlated with the degree of functional impairment in Huntington's disease ($r = 0.78$, $P < .0005$). It was concluded that abnormal alterations in the fractal properties of gait dynamics were associated with degradation in the central nervous system.

Room transitions were used as a measure of activity in sensor-based monitoring systems installed at Bristol and Deptford flats [17]. A room transition was defined as a trigger of a motion sensor in a room followed by a trigger in an adjacent room. It was reported that the number of transitions occurring within a defined time frame is an essential measure of wandering in Dementia. However, this metric was not robust to the floor plan variations, i.e., a bigger home may entail a higher number of rooms and vice versa. To quantify movement complexity in dementia, we present a robust approach known as cyclomatic complexity [18]. Cyclomatic complexity is a metric used to indicate the complexity of the control flow of a software program, which is analogous to the complexity of the movement flow of a subject in a home having a number of rooms. The advantage is that the measures of cyclomatic complexity are robust to floor plans, i.e., the value of complexity does not change for the movement flow of a subject in a home having a different number of rooms. We combined the measures of cyclomatic complexity with fractal index, the number of room transitions, movement entropy, and scaling exponent α and used these features to train a support vector machine (SVM) classifier to predict MCI.

II. Methods

A. Data acquisition

The data acquisition was conducted by the Oregon Center for Aging and Technology (ORCATECH) [19]. ORCATECH deployed a platform of ubiquitous sensors in the homes of older adults. In order to assess cognitive decline associated with neurodegenerative disorders such as dementia, these subjects were monitored for several years starting from January 2006 until 2016. All of these subjects were recruited from the Portland Oregon metropolitan area, including older adults living in local senior centers and retirement communities. The subjects provided written informed consent and agreed to be continuously monitored through a platform of sensors deployed in their homes. The subjects who were recruited met the following inclusion criteria: 1) 60 years and older man or woman; 2) living independently; 3) cognitively healthy, i.e., subjects with MMSE-score > 24 and clinical dementia rating (CDR) ≤0.5, and 4) of average health for age with no chronic symptoms.

We focused on the subjects who were the single occupants of their homes and either remained cognitively intact (CIN) or transitioned to MCI without bouncing back to cognitive intactness. The in-home assessments of these subjects were carried out at baseline and then during annual visits by clinicians who administered physical and neurological examinations using MMSE and CDR. CDR served as the ground truth and was used to determine if a subject was cognitively intact or impaired. A CDR score of '0' indicates cognitive intactness. A CDR score of '0.5' represents MCI. The conversion to cognitive impairment is a gradual process, and cognitive status is in flux just before the onset [20]. Hence, in case of a subject scoring '0' on a baseline visit and '0.5' on a subsequent visit the following year, the amount of time between these two visits was labeled *pre-transition*, i.e., this amount of time neither belonged to CIN nor MCI. On the other hand, the amount of time before the baseline visit was labeled as CIN. Likewise, the amount of time following the subsequent visit was labeled as MCI.

The data collection for this study was initially started in 2006 and continued until 2016. Protocols were developed, and pilot tested by ORCATECH in 2006 [19]. The enrollment of participants began in 2007. Within the period between 2006 and 2016, 205 enrolled subjects were single occupants of their homes. Thirty four subjects out of these 205 subjects developed MCI over time. The rest of the 171 subjects remained CIN. Some subjects, including 106 CIN and 12 MCI, moved to a new home. The rest of the 87 subjects, including 65 CIN and 22 MCI, did not change their residence throughout this period.

The subjects' homes ranged from a simple one-bedroom apartment to houses with as many as five bedrooms, a laundry room, a garage, etc. The details of the sensor platform and its installation in these homes were described previously [21]. An example of sensor deployment is shown in Fig. 1. In order to detect movement, wireless IR motion sensors (MS16A, The Home Automation Store, Carrolton, TX, USA) were strategically installed in rooms that were frequently visited by a subject such as a bedroom, bathroom, hallways, kitchen, living room, etc. A sensor activation records the timestamp and location of the subject. Another set of wireless magnetic contact sensors (DS10A, The Home Automation Store, Carrolton, TX, USA) was placed on the refrigerator, and the doors to track the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2020.2985907, IEEE Journal of Biomedical and Health Informatics

subject's absence from home and also outsider's visits. Both of these sensors MS16A and DS10A are generally used in setups for home security.

When motion occurs near the sensor, the motion sensors send a wireless signal to a data acquisition computer placed in the subject's home. Similarly, the contact sensors send event-based codes to the computer when the doors are opened or closed. The contact sensors placed on the exit doors were also used to track visitors in the home and any absences from the home. Moreover, the subjects were given weekly questionnaires to report visitors during the week, days spent outside the home, and changes in medication or health conditions. The recorded data in the computer was transferred and stored in a SQL database server through a broadband connection.

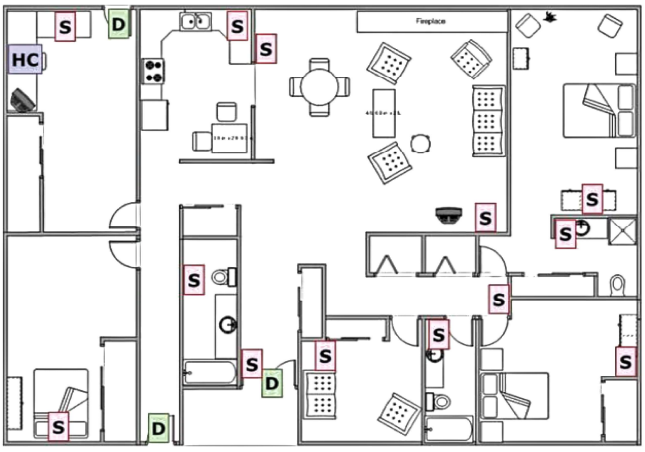


Fig. 1. An example of floorplan and sensor deployment [21]. Red boxes (S) are locations of infrared motion sensors (MS16A), green boxes (D) are magnetic contact sensors (DS10A), and the purple box (HC) is the location of the home computer.

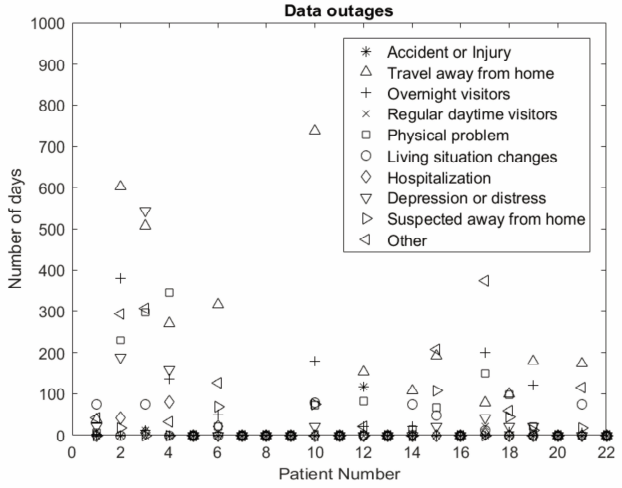


Fig. 2. Data outages. A cumulative number of days spent outside the home by a resident or time spent by outsiders at the resident's home is shown.

The data collected from various sensors consisting of timestamps and location coordinates were fused to generate the movement time series representing a subject's movement at home. The segments of this time series representing absences of the subject from home or an outsider's visit to the subject's home were discarded.

Data outages of 22 MCI study participants are shown in Fig. 2. The figure shows cumulative days that were spent by the participant outside home due to an accident or injury, physical problems, hospitalization or visits to a hospital emergency room, depression, or traveling, or if they are suspected to be away from home, i.e., when motion sensors did not trigger over a long period of time. Also, data were discarded if there was a change in the living situation or overnight visits of a spouse. 'Other' outages include a sensor or other equipment malfunction. Since the data were collected throughout the years for a given subject, the accumulated data provide a longitudinal profile of changes in the movement patterns of the subject that can be used to develop predictive models of MCI.

B. Features

In order to detect MCI, the complexity of the movement time series was quantified using a large initial set of features. The feature set, defined further below, included the cyclomatic complexity features, the total number of room transitions, the entropy, the scaling exponent α, and the fractal index of movement of a subject across rooms in a home.

Cyclomatic complexity is used as a metric to indicate the complexity of the control flow of a software program [18]. It is a quantitative measure of the linearly independent paths throughout the program. The nodes of the control-flow graph of this program correspond to the commands in that program. A directed edge connects two nodes in case if the second command is executed immediately after the first command. A group of commands that are executed within a specified period constitutes one connected component. Mathematically, the cyclomatic complexity is defined as

$$M = E - N + 2P \quad (1)$$

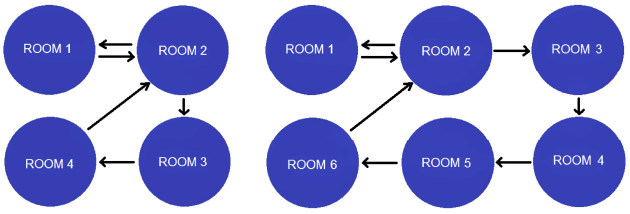
Where E is the number of edges in the control flow graph. N is the number of nodes in the graph. P is the number of connected components in the graph. Fig. 3 shows a control flow graph for the execution of a series of movements in a home from one room to another. These movements were executed in a sequence in a way that movement to room 2 was executed immediately after movement to room 1, followed by the movements to rooms 3, 4, 2, and 1, respectively. The graph consists of 4 nodes, 5 edges, and 1 connected component. The value of M for this control graph is $5 - 4 + 2(1) = 3$.

Cyclomatic complexity is analogous to the complexity of the movement flow of a subject in a home having some rooms where a room corresponds to a node (N) in the movement flow. An edge (E) represents the transition of a subject from one room to another. One trip of a subject between the rooms without a considerable time break corresponds to one connected component (P). The cyclomatic complexity for such a trip (i.e. P=1) can be given as $M = E - N + 2(1)$.

An advantage of using M to quantify movement complexity is that the value of M is robust to the change in the number of rooms in different homes. For instance, consider two homes are having four (Fig. 3a) and six rooms (Fig. 3b), respectively. The movement flow of a subject in home-1 making a round-trip between the rooms produces an M value of 3, i.e., M = 5-4+2(1)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2020.2985907, IEEE Journal of Biomedical and Health Informatics

= 3. Likewise, the movement flow of a subject in home-2, making a round-trip between the rooms produces the same M value, i.e., M = 7-6+2(1) =3.



a. Home-1 with 4 rooms b. Home-2 with 6 rooms

Fig. 3. Control flow graphs. Two homes with a different number of rooms are compared. In both a and b, the person moved similarly through the home, from the NW to NE to SE to SW back to NE back to NW. Notice that the cyclomatic complexity of the movement flow in both these homes is the same, i.e., M=3.

As we explored how cyclomatic complexity related to movement, we considered different statistics of *M* that were computed using the movement time series. These measures include the mean of M (*MM*), the standard deviation of M (*SM*), the coefficient of variance of M (*CM*), and total M (*TM*) given in (2), (3), (4) and (5), respectively.

$$MM = \sum_{i=1}^{n} M_i / n \tag{2}$$

$$SM = \sqrt{\sum_{i=1}^{n} (M_i - \bar{M})^2 / n - 1} \tag{3}$$

$$CM = (SM/MM) \times 100 \tag{4}$$

$$TM = \sum_{i=1}^{n} E_i - \sum_{i=1}^{n} N_i + 2P \tag{5}$$

Where $M_i = E_i - N_i + 2(1)$ is the M value for the i^{th} trip of a subject across the rooms without a break of 5 minutes in between the trip and *n* is the total number of trips in a day. A break longer than 5 minutes means that the starting point of the next trip is the room where the subject stayed in that room for longer than 5 minutes. Note that exits from the home are not considered a part of the cyclomatic complexity. Rather a flow graph would only start once the person has entered the home.

Other than the cyclomatic complexity features, the sum of room transitions (*RT*), the entropy (*E*), and the fractal index (*FI*) of the movement of a subject across rooms in a day was computed. The *RT* was computed using (6).

$$RT = \sum_{i=1}^{n} i \tag{6}$$

Where *i* is a subject's visit to a room and *n* is the total number of visits in different rooms in a day. The entropy *E* can be computed using (7).

$$E = -\sum p_i \times log_2(p_i) \tag{7}$$

Where *i* in the equation is a room in the subject's home, and *p* is the probability of the subject's visit to that room in a day. The entropy would be larger if the rooms in the home would have an equal probability of visits.

FI is a measure of the self-similarity of a signal that is assumed to be a realization of fractional Brownian motion (or random motion). It is a ratio that provides a complexity index that is indicative of variance in detail of the signal relative to the change in the measurement scale. *FI* can be computed in different ways [13, 14, 22-25]. In this study, it was computed using the method described by Istas and Lang [25]. According to their method, first, the fractional Gaussian noise is computed using the movement time-series based on the assumption that this time-series is the realization of the fractional Brownian motion given in (8).

$$Y(n) = X(n+1) - X(n) \quad for\ n = 1..N-1 \tag{8}$$

Where *X* is the movement time-series for a day, and *(N-1)* is the total length of this time-series. In the next step, the cumulative sum of *Y* is computed using (9).

$$C(n) = Y(n) + \sum_{i=1}^{n-1} Y(i) \quad for\ n = 1..N-1 \tag{9}$$

Now, *C (n)* is filtered using two tapped delay-line filters. The filter functions are given in (10) and (11) respectively

$$y_1(n) = C(n) - 2C(n-1) + C(n-2) \quad for\ n = 1..N \tag{10}$$

$$y_2(n) = C(n) - 2C(n-2) + C(n-4) \quad for\ n = 1..N \tag{11}$$

In the next step, the means of the element-wise squares of signals y_1 and y_2 are computed using (12) and (13), respectively.

$$S_1 = \sum_{n=1}^{N} y_1(n)^2 \Big/ N \tag{12}$$

$$S_2 = \sum_{n=1}^{N} y_2(n)^2 \Big/ N \tag{13}$$

Finally, *FI* is computed using (14)

$$FI = 0.5 \times log_2 \frac{S2}{S1} \tag{14}$$

Long-range correlations in the movement time-series were calculated using DFA [26]. In order to compute the DFA, first, the movement time-series $A_{1...N}$ of length *N* was integrated using (15).

$$y(k) = \sum_{i=1}^{k} [A_i - A_{avg}] \quad where\ A_{avg} = \sum_{i=1}^{k} A_i / k \quad for\ k = 1..N \tag{15}$$

This integrated time-series was divided equally into the boxes of length *n*. In this paper, the value of *n* was set to 3 hours (60 seconds × 60 minutes × 3 hours = 10800 samples) for the 24 hours (86400 samples) movement time-series. For each box of length *n*, a least-squares line was fit to the data representing the trend in that box. The y-coordinate of the fitted line in each box was represented by $y_n(k)$. Next, the segment of the integrated time-series in each box was detrended by subtracting between *y(k)* and the local trend $y_n(k)$. Finally, the fluctuation *F(n)* of the time series was calculated using (16).

$$F(n) = \sqrt{\sum_{k=1}^{N} (y_k - y_n(k))^2 / N} \quad for\ n = n, 2n, 3n..N/n \tag{16}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2020.2985907, IEEE Journal of Biomedical and Health Informatics

Typically, an integrated time series is self-similar if fluctuation $F(n)$ increases with the increasing value of n. In order to determine the power-law scaling between the two variables, a log-log plot was created between $F(n)$ and n. The slope of the fitted line on this log-log plot was computed, which is known as the scaling exponent α (or self-similarity measure).

In order to investigate the rise and fall of the feature values throughout a month, slopes and intercepts were computed for each feature by fitting a line over the feature values across 30 days. Consistency and variation in the feature values were estimated using the mean, standard deviation, and coefficient of variance of each feature across 30 days. All 40 features are listed in Table 1.

Since there were more CIN than MCI participants, we randomly selected 22 of the CIN participants from the total 65

TABLE I
THE FEATURE SET. THE PREFIX REPRESENTS THE FEATURE VARIANT, AND THE SUFFIX REPRESENTS THE FEATURE NAME

Features	*Cyclomatic complexity M*				*Other features*			
Values estimated over a month	Mean (*MM*)	Standard deviation (*SM*)	Total (*TM*)	Variance coefficient (*CM*)	Entropy (*E*)	Room transitions (*RT*)	Fractal index (*FI*)	DFA self similarity (α)
Slope S:	S_{MM}	S_{SM}	S_{TM}	S_{CM}	S_E	S_{RT}	S_{FI}	S_α
Slope Intercept I:	I_{MM}	I_{SM}	I_{TM}	I_{CM}	I_E	I_{RT}	I_{FI}	I_α
Mean μ:	μ_{MM}	μ_{SM}	μ_{TM}	μ_{CM}	μ_E	μ_{RT}	μ_{FI}	μ_α
Standard deviation σ:	σ_{MM}	σ_{SM}	σ_{TM}	σ_{CM}	σ_E	σ_{RT}	σ_{FI}	σ_α
Variance coefficient cv:	cv_{MM}	cv_{SM}	cv_{TM}	cv_{CM}	cv_E	cv_{RT}	cv_{FI}	cv_α

participants and used these along with the 22 MCI subjects to perform the classification. Features were extracted from the CIN subjects using the data collected during the last 12 months of their residence in ORCATECH homes. The dimension of feature space for the CIN group was 40 features × 264 samples (22 subjects × 12 months).

Features for the MCI participants were extracted during (1) the *pre-transition period* and (2) during the *post-transition period*. The features extracted during the pre-transition period were computed using the data collected during the 12 months that were labeled as *pre-transition* by the clinicians (i.e., before the diagnosis of MCI). Twelve months was chosen because people who are diagnosed with dementia are typically clinically diagnosed with CIN a year prior to the diagnosis. This feature set had the dimensions of 40 features × 217 samples. Similarly, the features extracted during the *post-transition period* were computed using the data collected during the 18 months after the participant transitioned to MCI. Importantly, the data collected after the 18th month were discarded since this analysis was primarily focused on using movement changes to do early detection of MCI transitions. The time period of 18 months was chosen because according to [27], the average time to conversion from MCI to Alzheimer's disease is 18 months. The feature set for this group had the dimensions of 40 features × 178 samples.

It is important to note that the pre- and post-MCI data sets were not entirely complete. Firstly, this was because data were discarded due to visitors and study participants moving away from home (Fig. 2). Secondly, some of the subjects were diagnosed with MCI close to the end of the data collection period in their home; hence, for these subjects, the entire 18 months post-transition data could not be gathered. For this group, 178 total samples were available for model training out of a maximum possible size of 396 samples (22 subjects × 18 post-transition months = 396). Likewise, some other subjects were diagnosed with MCI early during their stay in the ORCATECH homes. For these subjects, we could not collect complete data for 12 pre-transition months. For this group, 217 total samples were available for training out of a maximum possible of 264 samples (22 subjects × 12 pre-transition months = 264). However, in the case of CIN, a complete dataset was available for model training i.e. 264 total samples (22 subjects × 12 months = 264). These three sets of 217 pre-transition, 178 post-transition, and 264 CIN samples were used for feature extraction and analysis. The three feature sets were normalized between 0 and 1 and passed through a feature selection process before they were used for training an SVM.

C. Determining relevant features

Relevant features were identified using three feature selection methods for two different classification algorithms. The first classification algorithm was used to predict *post-transition MCI* vs. *CIN*. The second classifier algorithm was used to predict *pre-transition MCI* vs. *CIN*.

We compared three different feature selection methods, including 1) Kruskal-Wallis (KW) test [28], 2) Lasso (least absolute shrinkage and selection operator) [29], and 3) logistic regression (LR) test [30].

The KW test performs a nonparametric one-way analysis of variance between two or more groups [28]. The test computes the statistics by ranking the data from smallest to largest across all the groups. It compares the median of the ranks in each group to test the null hypothesis that independent samples come from continuous distributions that are identical and have equal medians. The test returns the p-value of the null hypothesis.

Lasso minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients β being less than a constant in a linear regression setting [29]. Because of this constraint, some coefficients β shrink to zero, which means that the features associated with those coefficients are eliminated. The features producing $\beta \geq 0$ may contribute towards predicting the response [29].

LR finds the optimal equation that best predicts the value of the response variable Y for feature X. It uses the natural log of the odds $Y/1-Y$ in regression equation $ln(Y/1-Y)=a+bX$. In order to find the best-fitting equation, the slope (b) and intercept (a) are computed using the maximum likelihood function [30]. The statistical significance (p-value) of this regression equation is computed using the chi-square method [31].

We used a leave-two-subjects out cross-validation (L2-CV) to avoid over-fitting and bias by assuring that testing was always done on a study participant using a model that was not

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2020.2985907, IEEE Journal of Biomedical and Health Informatics

trained on this study participant. Two randomly selected subjects were held out of the training set. The left-out data were used for testing, and the remaining data were used for training. A subject once selected for testing is not selected in other iterations for testing. In this way, data from the same individual were never used simultaneously in model training and testing thereby preventing overfitting.

When using the KW test, the average p-value was computed using L2-CV on the training set to select the features that were able to distinguish between the two groups with statistical significance ($p<0.05$). In the case of Lasso, the mean absolute value of the β coefficients was computed using L2-CV on the training set to select the features. Likewise, when using LR, the average p-value of the regression equation was computed using L2-CV on the training set to select the features that were able to discriminate between the two groups with statistical significance ($p<0.05$). The selected features were used for training the SVM classifiers.

D. Classification

SVM was chosen due to its ability to implement flexible decision boundaries that can generate optimal training results [32]. SVM works by constructing hyperplanes to separate classes by mapping predictors (support vectors) onto high-dimensional feature space that enables non-linear classification boundaries. These non-linear classifications and mapping of predictors to an outcome class are performed by the SVM using a kernel function. The SVM tends to find the most significant distance between the hyperplanes to facilitate a kernel function that leads to a quadratic programming optimization problem. This optimization problem can be solved using the sequential minimum optimization algorithm [32]. Since the data distribution was non-parametric, the radial basis kernel function [33] and sequential minimum optimization algorithm were utilized for training the SVM. The scaling factor sigma in the radial basis function was kept as 1. Hyperparameters were tuned using a grid search to optimize accuracy during cross-validation.

Using this SVM configuration, two different classification algorithms were built. *Classification algorithm 1* was trained only on the *CIN* and the *post-transition MCI* data. This algorithm was trained to classify between *post-transition MCI* and *CIN* using selected features and then tested to predict the onset and progression of MCI in the pre- and post-transition periods. We expected that *Classification algorithm 1* would be effective at predicting the progression of MCI after diagnosis of MCI. *Classification algorithm 2* was trained to classify the *pre-transition MCI* data from the *CIN* data. We expected that *Classification algorithm 2* would perform better in classifying the onset of MCI prior to diagnosis.

Classification algorithm 1 – Predicting progression of MCI: An SVM was trained using a minimal set of features that produced the highest accuracy in discriminating between the MCI and CIN samples. This set of features was selected iteratively using the three feature selection algorithms described above.

For both the KW and LR feature selection, the features producing the lowest p-values and the highest classification rate were used. We started by using a threshold p-value of 1.00, which means that, in the first selection round, all the features producing p-values less than the threshold were selected for training. In each subsequent round, the threshold p-value was divided by 10. Features with p-values more than the threshold were taken off from the feature set. The remaining features were selected for training. This procedure was repeated until only one feature was left for training the SVM. The classification rate was computed for each round, and the set of features producing the highest rate was selected for predicting MCI.

For the Lasso feature selection, we started by keeping a threshold β-value of 0, which means that all the features were utilized for training the SVM in the first round. Next, the value of β was increased by 0.1. The features producing the mean absolute value of β less than the threshold were removed. The remaining features were used for training. This procedure was repeated until only one feature was left for training the SVM. Classification accuracy was computed for each set of features, and the features set producing the highest accuracy was selected to predict the progression of MCI.

Classification accuracy (CA) for each iteration i of the L2-CV was computed using (17).

$$CA(i) = (TP + TN)/(TP + TN + FP + FN) \qquad (17)$$

Where *TP* or true positives are the MCI samples correctly identified as MCI. *TN* or true negatives are the CIN samples correctly identified as CIN. *FP* or false positives are the CIN samples incorrectly identified as MCI. *FN* or false negatives are the MCI samples incorrectly identified as CIN. Sensitivity or true positive rate (TPR) and specificity or true negative rate (TNR) refer to the model's ability to correctly classify MCI and CIN groups, respectively. In addition to CA, classification performance was interpreted using the area under the ROC curves (AUC) computed using different thresholds of sensitivity and specificity. The average *CA* was computed using (18).

$$Avg\_CA(n) = \left(\sum_{i=1}^{I}\sum_{j=1}^{J} CA_{i,j}\right)/(I \times J) \quad \begin{matrix} for \\ n = 1..N \end{matrix} \qquad (18)$$

Where I is the total number of CIN subjects, J is the total number of MCI subjects, and N is the total number of sets of features evaluated by the feature selection process. The average CA produced using each of the feature selection algorithms is shown in Fig. 4. Further analysis was done using box-plots by comparing the median ranks of the significant features producing the highest CA (Fig. 5b).

Once the optimal features were selected, the SVM was trained on the CIN subjects and the post-transition MCI subjects. The performance of the SVM was evaluated on both pre-transition and post-transition months from the MCI subjects to observe how well the classifier performed, leading up to the transition to MCI diagnosis. The training and the testing sets were balanced such that the number of MCI and CIN samples was equal in the training and testing sets. Classification

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2020.2985907, IEEE Journal of Biomedical and Health Informatics

accuracy was computed for CIN and MCI subjects separately using (17). In the case of MCI, the accuracy was computed for each month between 12 months pre-transition and 12 months post-transition period, to identify the onset of progression. Section III-A presents the results.

Classification algorithm 2 – Prediction of the MCI onset: This classifier was designed the same as for algorithm 2, with the only exception being that pre-transition MCI data was used to train the classifier against CIN data. We expected that by training on pre-transition data, we might get a better prediction of the onset of MCI than when post-transition MCI data was used.

Data was trained using L2-CV on an equal number of pre-transition MCI vs. CIN subjects. The algorithm was then evaluated on both the pre-transition MCI data as well as the post-transition data. The training and the testing sets were balanced such that the number of pre-transition and the CIN samples was equal in the training and the testing sets. In the case of the pre-transition MCI group, the accuracy was computed for each month between the 12th pre-transition month and the 12th post-transition month. Section III-B presents the results.

III. Results

A. Prediction of the MCI progression:

The results from the KW test suggests that the TM features I_{TM}, μ_{TM} and σ_{TM}, the entropy features σ_E and cv_E, the FI features I_{FI} and μ_{FI} and the DFA features I_α, μ_α, σ_α and cv_α were able to distinguish between MCI and CIN groups with statistical significance ($p<0.05$). Likewise, the β values indicate that the TM features σ_{TM} and cv_{TM}, the entropy feature cv_E, the FI features I_{FI}, μ_{FI} and cv_{FI} and the DFA features μ_α and cv_α may contribute in predicting the MCI and CIN groups correctly. Similarly, the results from the LR test suggests that several cyclomatic complexity features (σ_{MM}, cv_{MM}, I_{SM}, μ_{SM}, σ_{SM}, I_{TM}, μ_{TM}, σ_{TM} and σ_{CM}), entropy features (σ_E and cv_E), FI features (I_{FI} and μ_{FI}) and the DFA features (I_α, μ_α, σ_α and cv_α) were able to distinguish between MCI and CIN groups with statistical significance ($p<0.05$). Importantly, none of the room transition features were able to produce p-value <0.05 or $\beta > 1$ in any of the three tests.

Fig. 4 shows the average classification rates produced by each feature selection algorithm. Lasso produced the highest accuracy of 79% using six features, including σ_{TM}, cv_{TM}, cv_E, I_{FI}, $cv_{FI,}$ and cv_α. The sensitivity, specificity, and AUC were 91.4%, 82.7%, and 76.1%, respectively. The β coefficients and box-plots for each of these features are shown in Fig. 5. The feature cv_E produced the highest value of β, followed by cv_{FI}, σ_{TM}, cv_{TM}, $I_{FI,}$ and cv_α. The boxplots of cv_E, σ_{TM}, cv_{TM}, $I_{FI,}$ and cv_α revealed that the median and the mean rank of the CIN group was higher than the median and the mean rank of the MCI group with statistical significance (p-value<0.05). On the other hand, the medians of cv_{FI} were not significantly different.

The testing of the samples of CIN subjects on this SVM model produced a mean classification accuracy of 81%, with only one subject having samples classified with an accuracy

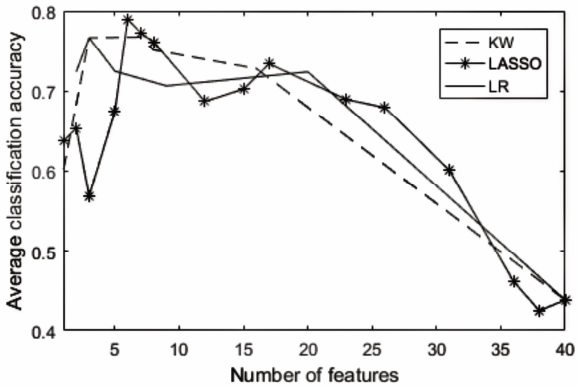
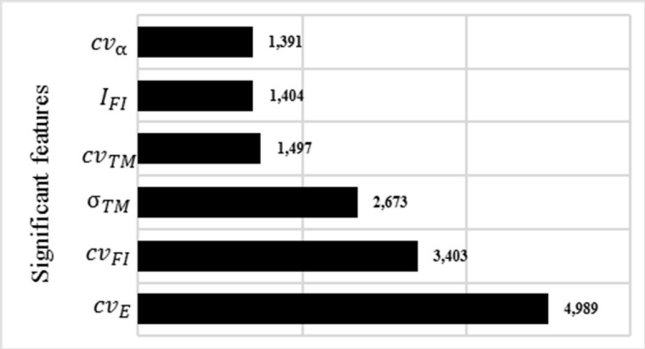
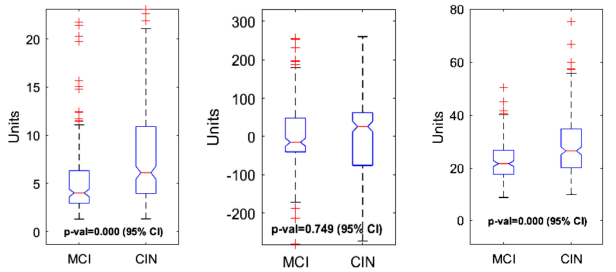


Fig. 4. A comparison between the feature selection algorithms used for classification between the CIN and MCI samples. The leave-two-subjects-out cross-validation on SVM produced the highest classification accuracy (79%) using six features (σ_{TM}, cv_{TM}, cv_E, I_{FI}, $cv_{FI,}$ and cv_α) each of which produced the mean absolute LASSO-β value ≥ 1.3.



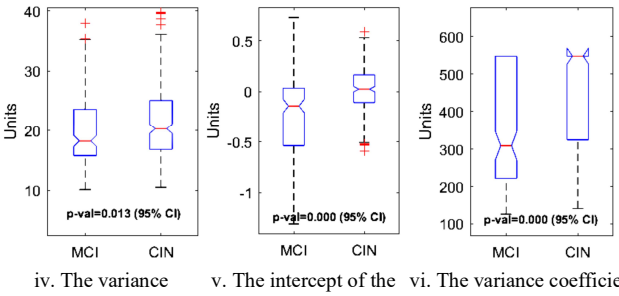
a. Lasso-β coefficients of selected features



i. The variance coefficient of entropy (cv_E) over a month

ii. The variance coefficient of the fractal index (cv_{FI}) over a month

iii. The standard deviation of TM (σ_{TM}) over a month



iv. The variance coefficient of TM (cv_{TM}) over a month

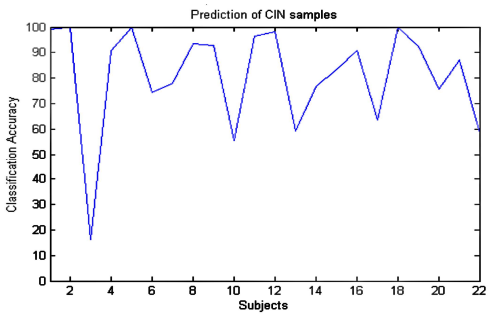
v. The intercept of the slope of the fractal index (I_{FI}) over a month

vi. The variance coefficient of scaling exponent (cv_α) over a month

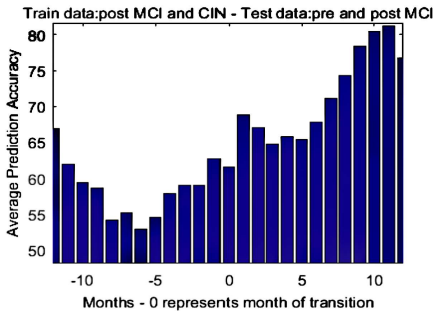
b. Box-plot comparison of selected features.

Fig. 5. Selected features for classification between MCI and CIN groups.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2020.2985907, IEEE Journal of Biomedical and Health Informatics



a. Classification of the CIN dataset. The samples were classified with an average accuracy of 81%.



b. Classification of the MCI dataset. An increasing trend between the sixth pre-transition month and the eleventh post-transition month was observed. The peak accuracy (81.2%) was at the eleventh post-transition month.

Fig. 6. Classification algorithm 1's results of prediction of MCI progression.

below 50% (Fig. 6a). On the other hand, the testing of the samples of the MCI subjects produced the highest classification accuracy of 81.2% at the eleventh post-transition month (Fig. 6b). The lowest accuracy of 52.8% was at the sixth pre-transition month. An increasing trend between the sixth pre-transition and eleventh post-transition months was observed that indicates that the symptoms start to elevate six months before the actual diagnosis and continue to increase up to the eleventh month. Besides, the backward progression of classification accuracies before the sixth pre-transition month is indicative of the subjects being cognitively intact.

B. Prediction of the MCI onset

Results suggest that some cyclomatic complexity, entropy, FI, and DFA features were able to discriminate between pre-transition and CIN groups with statistical significance ($p<0.05$). However, none of the room transition features were able to produce $p<0.05$ or $\beta > 1$ in any of the tests.

Fig. 7 presents a comparison between the feature selection algorithms used in the classification of the pre-transition group. The KW test produced the highest classification accuracy of 75% using three features, including σ_{TM}, μ_{FI}, and cv_α. The sensitivity, specificity, and AUC were 72.7%, 77.3%, and 75%, respectively. Box-plots were produced for each of these features to compare between group-medians and mean-ranks (Fig. 8b). For all the three features, the medians and mean-ranks of the CIN group were significantly higher ($p<0.05$) than the medians and mean-ranks of the pre-transition group. In the case of σ_{TM}, the median and the mean rank of the CIN were 26.57 (± 34.8 – 20.1) and 279.49, as compared to the median and mean rank of the pre-transition group that was 20.06 (± 26.33 – 14.51) and 194.17 respectively. For μ_{FI}, the median and the mean rank of the CIN were 0.02 (± 0.18 – -0.12) and 271.76, as compared to the median and the mean rank of the pre-transition group that was -0.15 (± 0.106 – -0.49) and 203.57 respectively. Lastly, for cv_α, the median and the mean rank of the CIN were 547.72 ± (569.25 – 324.7) and 276.02, as compared to the median and the mean rank of the pre-transition group, which were 380.94 ± (547.7 – 209.35) and 198.38 respectively.

Importantly, the median value of σ_{TM} for the pre-transition group (20.06 units; Fig. 8b.i) was approximately equal to the median value of the MCI group (21.74 units; Fig. 5b.iii), suggesting that σ_{TM} is predictive of the MCI progression during the pre-transition period. Likewise, the median value of cv_α for the pre-transition group (380.94 units; Fig. 8.b.iii) was closer to the median value of cv_α for the MCI group (349 units; Fig. 5.b.vi), suggesting that cv_α is another strong feature that is predictive of the MCI progression during the pre-transition period.

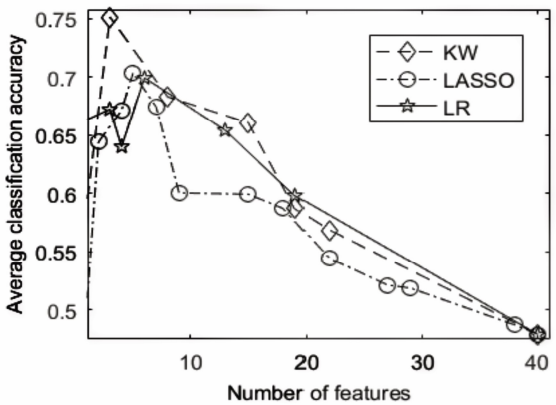
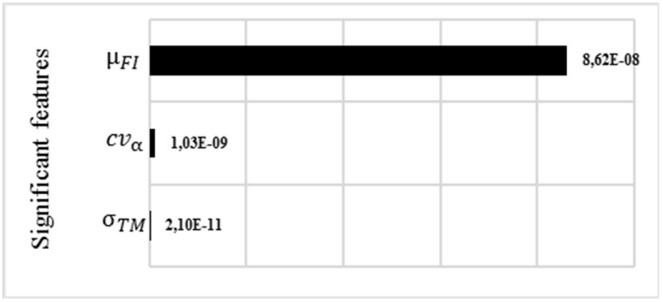
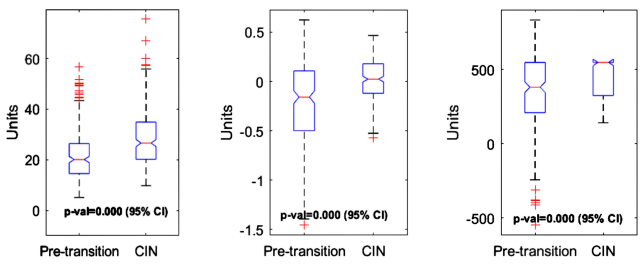


Fig. 7. A comparison between the feature selection algorithms for classification between the CIN and pre-transition groups. The leave-two-subjects-out cross-validation on SVM produced the highest classification accuracy (75.1%) using three features (σ_{TM}, μ_{FI}, and cv_α) producing the KW p-value < 0.00001.



a. KW p-values of selected features



i. The standard deviation of TM (σ_{TM}) over a month.

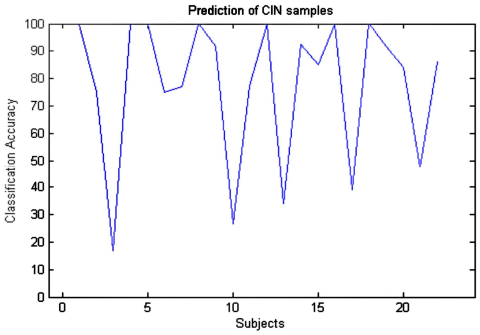
ii. Mean of the fractal index (μ_{FI}) over a month.

iii. The variance coefficient of scaling exponent (cv_α) over a month.

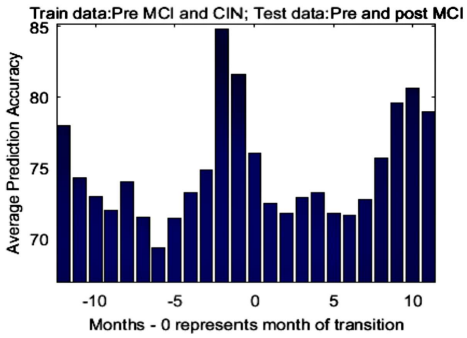
b. Box-plot comparison of selected features.

Fig. 8. Selected features for classification between pre-transition MCI and CIN groups.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2020.2985907, IEEE Journal of Biomedical and Health Informatics



a. Prediction of the CIN dataset. The samples were classified with an average accuracy of 77%.



b. Prediction of the MCI onset. An increasing trend between the sixth and second pre-transition months was observed.

Fig. 9. Classification algorithm 2 results for predicting the onset of MCI.

The testing of the CIN group on this SVM model, trained using the selected features, produced a mean classification accuracy of 77.23%. The testing of the pre-transition group produced the highest classification accuracy of 84.8% at two months before the transition diagnosis. The lowest accuracy of 69.3% was at the sixth pre-transition month.

An increasing trend in the accuracy between the sixth and second pre-transition months was observed in Fig. 9b. The elevation in the accuracy from the sixth pre-transition month validates the results of the first classification experiment that movement complexity begins to elevate from the sixth pre-transition month of MCI (Fig. 6b). The pre-transition MCI and the CIN groups were most separable at the second pre-transition month. Moreover, the selected features were able to classify the MCI group with a maximum accuracy of 80.6% at the tenth post-transition month. Importantly, two of the features, $\sigma_{TM,}$ and $cv_{\alpha,}$ contributed the most in identifying the onset and progression of MCI in both the classification experiments. These results confirm that the selected features in both classification algorithms were the excellent representatives of movement patterns as they related to MCI.

IV. Conclusions

To the best of our knowledge, there has not yet been a study that demonstrates how movement complexity can be used to predict MCI onset and progression. We demonstrated how movement complexity features could be used as features in a machine-learning algorithm to predict both the onset of MCI and the progression of MCI depending on the types of data on which you train the classifier.

The results of this study suggested that features based on cyclomatic complexity, which are robust to normalizing between different residential floor plans, may be used by a machine-learning algorithm to help discriminated between CIN and MCI. Other movement features such as room transition features were not useful in helping to classify the difference between MCI and CIN group. Features extracted from the data indicated that the variability of the total cyclomatic complexity (σ_{TM}) and the variability of the DFA self-similarity measure (σ_{α} and cv_{α}) contributed most significantly in predicting MCI.

Analysis of the features indicated that the CIN participants showed greater movement variability than the MCI participants, which is in line with Nams et al. [13], who showed that movement variability estimated using fractal dimension was positively correlated with MMSE scores for subjects with MMSE scores >24. Also, the results were in line with Hausdorff et al. [15], who showed that α was lower for Huntington's patients compared to the healthy subjects, indicating that movement variability declines as cognition declines.

We compared our study with a recent study on dementia that used a simulation of sensor activations to represent abnormal activity patterns of seniors living in smart homes [34]. The study simulated abnormal behavior by inserting random sensor activations in the night-time activity sequence of residents to represent sleep-disorders and night-time wandering. Deep learning algorithms, including graph convolutional networks, long short-term memory, and convolutional neural networks when trained separately on raw sensor activations, i.e., 'ON' and 'OFF' states recorded over 10 seconds, classified normal and abnormal patterns with an AUC of 66%, 57%, and 59% respectively. The study [34] explained the reason for the poor model performance by stating that they trained their models on data collected in small time windows of 10 seconds. The short time-slices were unable to encode temporal information required to characterize the gradual change in activity patterns representing cognitive decline.

The study in [34] emphasizes the importance of using feature variants representing temporal characteristics such as slopes, intercepts, consistency, and variation of feature values over longer durations, such as the month duration used in the current manuscript, to predict cognitive decline. Importantly, the aggregation of feature values over a longer duration assimilates varying movement patterns of individuals subjective to their daily routine. However, a drawback of estimating temporal characteristics of features over long time windows is the reduced size of training data that restricts the choice of the learning algorithm. For instance, the size of our training sets was 264 and 178 samples for predicting the onset and progression of MCI, respectively, which is sufficient to train an SVM but not sufficient to train a deep neural network learning algorithm due to the risk of overfitting for training on a small dataset (n<200) [35]. We plan to improve the model's ability to predict cognitive decline by collecting a big dataset of activity patterns for training recurrent neural networks that take

temporal information into account. Future models could potentially be improved by the fusion of multi-modal assessment of brain images such as MRI and PET scans that have shown promise in predicting MCI [36 and 37].

In conclusion, the proposed framework based on movement analysis was able to predict the MCI onset months before the diagnosis and was able to track the progression between the onset and advanced stages. The model was able to classify MCI with a maximum accuracy of 81.2% in the eleventh post-transition month. In the future, activity recognition using pervasive sensor computing in smart homes will be incorporated with movement analysis to improve the model.

V. References

[1] R. Sousa et al., "Contribution of chronic diseases to disability in elderly people in countries with low and middle incomes: a 10/66 Dementia Research Group population-based survey", *The Lancet*, vol. 374, no. 9704, pp. 1821-1830, 2009. Available: 10.1016/s0140-6736(09)61829-8.

[2] M. Prince, R. Bryce, E. Albanese, A. Wimo, W. Ribeiro, and C. Ferri, "The global prevalence of dementia: A systematic review and meta-analysis," *Alzheimer's & Dementia*, vol. 9, no. 1, pp. 63-75.e2, 2013. Available: 10.1016/j.jalz.2012.11.007.

[3] N.K. Vuong, C. Syin, and T.L. Chiew. *mHealth Sensors, Techniques, and Applications for Managing Wandering Behavior of People with Dementia: A Review* (5th ed) Springer; 2015. p. 11-42.

[4] I. Ramakers et al., "Symptoms of Preclinical Dementia in General Practice up to Five Years before Dementia Diagnosis," *Dementia and Geriatric Cognitive Disorders*, vol. 24, no. 4, pp. 300-306, 2007. Available: 10.1159/000107594.

[5] M. Pavel, L. H. Tamara, A. Andre, J. Holly, and K. Jeffrey. "Unobtrusive assessment of mobility," in *Conf Proc. IEEE Eng Med Biol Soc* 2006. p. 6277-80.

[6] T.L. Hayes, M. Pavel, and J. Kaye. "An unobtrusive in-home monitoring system for detection of key motor changes preceding cognitive decline." in *26th Conf Proc. IEEE Eng Med Biol Soc* 2004. p. 2480–3.

[7] D.P. Birkett. *Psychiatry in the nursing home.* Routledge; 2013.

[8] H. Nygaard, "Consumption of psychotropic drugs in elderly persons registered at the home nursing center in Bergen, Norway," *International Journal of Geriatric Psychiatry*, vol. 7, no. 1, pp. 53-58, 1992. Available: 10.1002/gps.930070109.

[9] A. Wimo, B. Winblad, and L. Jönsson, "The worldwide societal costs of dementia: Estimates for 2009", *Alzheimer's & Dementia*, vol. 6, no. 2, pp. 98-103, 2010. Available: 10.1016/j.jalz.2010.01.010.

[10] Profiling the dementia family carer in Singapore. 2019. [Online]. Available: http://alz.org.sg/wp-content/uploads/2017/04/Research-Profiling-Dementia-Family-Carer-SG.pdf. [Accessed: 26- Feb- 2019].

[11] K. Makimoto, Eun Ah Lee, Y. Kang, M. Yamakawa, N. Ashida, and Kyung Rim Shin, "Temporal Patterns of Movements in Institutionalized Elderly With Dementia During 12 Consecutive Days of Observation in Seoul, Korea", *American Journal of Alzheimer's Disease & Other Dementias*, vol. 23, no. 2, pp. 200-206, 2008. Available: 10.1177/1533317507312625.

[12] D. Algase, C. Antonakos, E. Beattie, C. Beel-Bates, and L. Yao, "New Parameters for Daytime Wandering," *Research in Gerontological Nursing*, vol. 2, no. 1, pp. 58-68, 2009. Available: 10.3928/19404921-20090101-02.

[13] V. Nams, J. Fozard, and W. Kearns, "Tortuosity in Movement Paths are Related to Cognitive Impairment," *Methods of Information in Medicine*, vol. 49, no. 06, pp. 592-598, 2010. Available: 10.3414/me09-01-0079.

[14] W. Kearns, C. Fozard, V. Nams, and J. Craighead, "Wireless tele surveillance system for detecting dementia," *Gerontechnology*, vol. 10, no. 2, 2011. Available: 10.4017/gt.2011.10.2.004.00.

[15] J. Hausdorff et al., "Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease," *Journal of Applied Physiology*, vol. 82, no. 1, pp. 262-269, 1997. Available: 10.1152/jappl.1997.82.1.262.

[16] A. Goldberger, L. Amaral, J. Hausdorff, P. Ivanov, C. Peng, and H. Stanley, "Fractal dynamics in physiology: Alterations with disease and aging," *Proceedings of the National Academy of Sciences*, vol. 99, no. 1, pp. 2466-2472, 2002. Available: 10.1073/pnas.012579499.

[17] T. Adlam, B. E. Carey-Smith, N. Evans, R. Orpwood, J. Boger, and Mihailidis. *Implementing Monitoring and Technological Interventions in Smart Homes for People with Dementia-Case Studies. BMI Book 3*. 2009:159-82.

[18] T. McCabe, "A Complexity Measure," *IEEE Transactions on Software Engineering*, vol. 2, no. 4, pp. 308-320, 1976. Available: 10.1109/tse.1976.233837.

[19] T. Hayes, "The ORCATECH Living Lab: From smart homes to smart communities," *Gerontechnology*, vol. 9, no. 2, 2010. Available: 10.4017/gt.2010.09.02.146.00.

[20] B. Kelley and R. Petersen, "Alzheimer's Disease and Mild Cognitive Impairment," *Neurologic Clinics*, vol. 25, no. 3, pp. 577-609, 2007. Available: 10.1016/j.ncl.2007.03.008.

[21] J. Kaye et al., "Intelligent Systems for Assessing Aging Changes: Home-Based, Unobtrusive, and Continuous Assessment of Aging," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 66, no. 1, pp. i180-i190, 2011. Available: 10.1093/geronb/gbq095.

[22] V. Nams, "Improving Accuracy and Precision in Estimating Fractal Dimension of Animal movement paths," *Acta Biotheoretica*, vol. 54, no. 1, pp. 1-11, 2006. Available: 10.1007/s10441-006-5954-8.

[23] P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 910-917, 1992. Available: 10.1109/18.119751.

[24] J.M. Bardet, G. Lang, G. Oppenheim, A. Philippe, S. Stoev, and M. S. Taqqu. *Semi-parametric estimation of the long-range dependence parameter: a survey***.** Birkhauser; 2003. p. 557-77.

[25] J. Istas and G. Lang, "Quadratic variations and estimation of the local Hölder index of a Gaussian process," *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, vol. 33, no. 4, pp. 407-436, 1997. Available: 10.1016/s0246-0203(97)80099-4.

[26] C. Peng, S. Buldyrev, A. Goldberger, S. Havlin, M. Simons, and H. Stanley, "Finite-size effects on long-range correlations: Implications for analyzing DNA sequences," *Physical Review E*, vol. 47, no. 5, pp. 3730-3733, 1993. Available: 10.1103/physreve.47.3730.

[27] P. Julayanont, Parunyou, B. Mélanie, C. Howard, P. Natalie, and S.N. Ziad. "Montreal Cognitive Assessment Memory Index Score (MoCA-MIS) as a Predictor of Conversion from Mild Cognitive Impairment to Alzheimer's Disease." *Journal of the American Geriatrics Society*, vol. 62, no. 4, pp. 679-684, 2014. Available: 10.1111/jgs.12742.

[28] E. Theodorsson-Norheim, "Kruskal-Wallis test: a BASIC computer program to perform nonparametric one-way analysis of variance and multiple comparisons on ranks of several independent samples," *Computer Methods and Programs in Biomedicine*, vol. 23, no. 1, pp. 57-62, 1986. Available: 10.1016/0169-2607(86)90081-7.

[29] A. Belloni, V. Chernozhukov, and L. Wang, "Pivotal estimation via square-root Lasso in nonparametric regression," *The Annals of Statistics*, vol. 42, no. 2, pp. 757-788, 2014. Available: 10.1214/14-aos1204.

[30] R. Zakharov, and P. Dupont. "Ensemble logistic regression for feature selection." *In PRIB* 2011. p. 133-44.

[31] D. Hosmer, T. Hosmer, S. Le Cessie, and S. Lemeshow, "A comparison of goodness-of-fit tests for the logistic regression model," *Statistics in Medicine*, vol. 16, no. 9, pp. 965-980, 1997. Available: 10.1002/(sici)1097-0258(19970515)16:9<965::aid-sim509>3.0.co;2-o.

[32] B. Schölkopf, J. Platt, J. Shawe-Taylor, A. Smola and R. Williamson, "Estimating the Support of a High-Dimensional Distribution," *Neural Computation*, vol. 13, no. 7, pp. 1443-1471, 2001. Available: 10.1162/089976601750264965.

[33] J. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293-300, 1999. Available: 10.1023/a:1018628609742.

[34] D. Arifoglu, N.C. Hammadi, and B. Abdelhamed, "Detecting indicators of cognitive impairment via Graph Convolutional Networks." *Engineering Applications of Artificial Intelligence*, vol 89, pp. 103401, 2020. Available: 10.1016/j.engappai.2019.103401.

[35] T. Eslami, and S. Fahad, "Auto-ASD-network: a technique based on deep learning and support vector machines for diagnosing autism spectrum disorder using fMRI data." *In Proc. 10th ACM Int. Conf. Bioinformatics, Computational Biology and Health Informatics*, 2019. p. 646-651.

[36] T. Zhou, L. Mingxia, T. Kim-Han, and S. Dinggang, "Latent representation learning for Alzheimer's disease diagnosis with incomplete multi-modality neuroimaging and genetic data." *IEEE transactions on medical imaging* vol. 38, no. 10, pp. 2411-2422, 2019. Available: 10.1109/TMI.2019.2913158

[37] Y. Zhang, Z. Han, C. Xiaobo, L. Mingxia, Z. Xiaofeng, L. Seong-Whan, and S. Dinggang, "Strength and similarity guided group-level brain functional network construction for MCI diagnosis." *Pattern Recognition* vol. 88, pp. 421-430, 2019. Available: 10.1016/j.patcog.2018.12.001.