Received: 26 March 2024 | Revised: 19 August 2024 | Accepted: 25 August 2024

DOI: 10.1002/dad2.70012

Alzheimer's & Dementia
Diagnosis, Assessment & Disease Monitoring

RESEARCH ARTICLE

Artificial intelligence detection of cognitive impairment in older adults during walking

Shuichi P. Obuchi[1] | **Motonaga Kojima**[2] | **Hiroyuki Suzuki**[3] | **Juan C. Garbalosa**[4] | **Keigo Imamura**[1] | **Kazushige Ihara**[5] | **Hirohiko Hirano**[6] | **Hiroyuki Sasai**[7] | **Yoshinori Fujiwara**[8] | **Hisashi Kawai**[1]

[1]Research Team for Human Care, Tokyo Metropolitan Institute for Geriatrics and Gerontology, Itabashi-ku, Tokyo, Japan

[2]Department of Physical Therapy, University of Tokyo Health Sciences, Tama-shi, Tokyo, Japan

[3]Research Team for Social Participation and Community Health, Tokyo Metropolitan Institute for Geriatrics and Gerontology, Itabashi-ku, Tokyo, Japan

[4]Department of Physical Therapy, Quinnipiac University, Hamden, Connecticut, USA

[5]Graduate School of Medicine, Hirosaki University, Hirsaki-shi, Aomori, Japan

[6]Research Team for Promoting Independence and Mental Health, Tokyo Metropolitan Institute for Geriatrics and Gerontology, Itabashi-ku, Tokyo, Japan

[7]Research Team for Promoting Independence and Mental Health, Tokyo Metropolitan Institute of Geriatrics and Gerontology, Itabashi-ku, Tokyo, Japan

[8]Tokyo Metropolitan Institute for Geriatrics and Gerontology, Itabashi-ku, Tokyo, Japan

Correspondence
Hisashi Kawai, Research Team for Human Care, Tokyo Metropolitan Institute for Geriatrics and Gerontology, 35-2 Sakae-cho, Itabashi-ku, Tokyo 173-0015, Japan.
Email: hkawai@tmig.or.jp

Funding information
Health Labour Sciences Research Grant, Grant/Award Numbers: nos. H23-Choju-Ippan-001, and 002; Ministry of Health, Labour and Welfare of Japan and JSPS KAKENHI, Grant/Award Number: grant number: 23K06919

Abstract

INTRODUCTION: To detect early cognitive impairment in community-dwelling older adults, this study explored the viability of artificial intelligence (AI)-assisted linear acceleration and angular velocity analysis during walking.

METHODS: This cross-sectional study included 879 participants without dementia (female, 60.6%; mean age, 73.5 years) from the 2011 Comprehensive Gerontology Survey. Sensors attached to the pelvis and left ankle recorded the triaxial linear acceleration and angular velocity while the participants walked at a comfortable speed. Cognitive impairment was determined using Mini-Mental State Examination scores. Deep learning models were used to discern the linear acceleration and angular velocity data of 12,302 walking strides.

RESULTS: The models' average sensitivity, specificity, and area under the curve were 0.961, 0.643, and 0.833, respectively, across 30 testing datasets.

DISCUSSION: AI-enabled gait analysis can be used to detect signs of cognitive impairment. Integrating this AI model into smartphones may help detect dementia early, facilitating better prevention.

KEYWORDS
acceleration, angular velocity, artificial intelligence, gait, screening

Highlights

- Artificial intelligence (AI)-enabled gait analysis can be used to detect the early signs of cognitive decline.
- This AI model was constructed using data from a community-dwelling cohort.
- AI-assisted linear acceleration and angular velocity analysis during gait was used.
- The model may help in early detection of dementia.

1 | BACKGROUND

The early detection of cognitive decline in older individuals is of paramount importance given that many therapies used to delay its development are most efficacious when begun in the early stages of decline.[1] Traditional screening tools for cognitive decline, such as the Mini-Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA), require skilled professionals and are time consuming to implement.[2,3] Moreover, these tools have a learning effect.[4] Given the accelerated rate at which populations are aging in developed countries, there is an urgent need to create convenient cognitive screening tools that can be used remotely and on a large scale to facilitate the early detection of cognitive decline without the need of a skilled professional.[5–10]

Walking is the strongest predictor of a decline in instrumental activities of daily living;[11] therefore, a decline in walking ability may be an early sign of a decline in cognitive function. Previous epidemiological studies demonstrated a relationship between cognitive function and walking.[12–15] However, those studies did not establish a threshold that can be used to screen for cognitive decline or the threshold validity (sensitivity and specificity). Therefore, further studies are warranted to determine whether walking can be used as a screening tool for the detection of dementia and cognitive decline.

The recent advances in the field of artificial intelligence (AI) and the widespread use of wearable devices have facilitated the identification of subtle changes in walking. Moreover, their utility in the screening of degenerative neurological diseases, including dementia, has been established.[16] Consequently, this form of gait analysis is now considered a potential screening biomarker that can be used to detect neurodegenerative diseases.[17–19] The development of digital biomarkers, along with the development of clinical biomarkers, is expected.[20,21]

At present, a variety of studies exist using either AI or machine learning algorithms in the analysis of gait as a method to detect cognitive impairment in older individuals. Zhang et al.[22] used AI-powered video camera analysis to detect early-stage dementia with an accuracy of 0.741. Using a computerized walkway and a support vector machine learning model, Ghoraani et al. were able to detect mild cognitive impairment (MCI) and Alzheimer's disease with an accuracy of 0.86.[23] Although these studies demonstrate the utility of AI in the analysis of gait to detect cognitive impairment in older participants, their use for the detection of these changes on a large scale is limited as they either require specialized equipment that is not commonly available in most clinics or cannot be used remotely. A more promising technological advance is the use of accelerometers and gyroscopes in combination with AI models to detect cognitive impairment in older adults. Jeon et al.[24] were able to differentiate patients with MCI from healthy controls, with an accuracy ranging from 0.74 to 0.73, during the performance of complex walking using a machine learning model and data acquired from accelerometers and gyroscopes. Although these studies suggest that the analysis of gait using some form of AI can distinguish between individuals with and without dementia or cognitive decline, they have certain limitations that make it unclear whether their methods can be used to screen for cognitive impairment in community-dwelling individuals on a large scale.

This study aimed to determine whether AI analysis of acceleration and angular velocity gait patterns using readily available inertial motion sensors can be used to detect cognitive impairment in a cohort of community-dwelling older adults. This method, if proven accurate, could serve as an alternative to commonly used tests to assess cognitive decline. Furthermore, it may facilitate objective assessments without the learning effects associated with traditional testing methods.

2 | METHODS

2.1 | Participants

A total of 879 community-dwelling older adults who had participated in the 2011 Comprehensive Gerontology Survey (Otassha Study) were included in the present study. The Otassha Study is a comprehensive gerontological survey that is conducted once a year to assess physical, oral, psychological, cognitive, and social functions. In addition, medical conditions, including a diagnosis of dementia, were identified via interviews. The interviews included questions regarding the medical history of hypertension, stroke, heart disease, and diabetes but did not include questions about the history of medications, history of depression, and neuromuscular abnormalities. Cognitive impairment was defined as a score < 24 on the MMSE (positive case); a score ≥ 24 was defined as an absence of cognitive impairment (negative case).

RESEARCH IN CONTEXT

1. **Systematic review**: Previous studies have demonstrated a relationship between cognitive function and walking and have identified thresholds for cognitive decline screening. However, they did not assess the validity of those thresholds. To detect early cognitive decline in community-dwelling older adults, we explored the viability of artificial intelligence (AI)-assisted linear acceleration and angular velocity analysis during walking. Cognitive decline was determined using Mini-Mental State Examination scores. Deep learning models were used to discern the linear acceleration and angular velocity data of 12,302 walking strides.
2. **Interpretation**: AI-enabled gait analysis can be used to detect early signs of cognitive decline. If this AI model is integrated into smartphones, it could represent a significant step toward early preventive measures for dementia.
3. **Future directions**: Given that the degenerative changes responsible for cognitive decline develop relatively early and require early management, larger studies using AI models capable of predicting cognitive decline early are needed.

2.2 | Measurement of linear acceleration and angular velocity during walking

From a static standing position, the participants were asked to walk along a 16-m gait path at their comfortable walking speed twice while triaxial sensors attached to the pelvis and left ankle recorded the linear acceleration and angular velocity of the lower leg and pelvis at a sampling rate of 1000 Hz (Figure 1). The triaxial sensors comprised a general-purpose three-dimensional accelerometer and gyroscope (MVP-RF8-BC, MicroStone Corp.). The initial and last three steps of each walking trial were excluded from further data analysis.

The linear acceleration was measured as the linear acceleration in a Cartesian coordinate system on the sensor. The angular velocity was defined as the vector representing the rotational velocity around each axis of the Cartesian coordinates on the sensor. The measured data were not subjected to any pre-processing such as filtering. Each walking trial consisted of approximately seven strides. A dataset was created by extracting the linear acceleration and angular velocity data for each stride.

3 | DATA ANALYSIS

3.1 | Data structure

A dataset was created by collecting 12,302 strides of linear acceleration and angular velocity from 879 participants. Each sample comprised 12 features, including the triaxial linear acceleration and angular velocity of the pelvis and lower leg during walking. In addition, the stride times were recorded and added to the 12 features, yielding a dataset comprising 13 features (six linear accelerations and six angular velocities with one time feature).

3.2 | Data preparation

The time required to complete each stride was recorded. In addition, the average linear acceleration and angular velocity along and about each axis were determined for each stride. These values were used as representative values for the stride. The averages of the linear accelerations and angular velocities over each walking stride from the larger original dataset and stride times were used subsequently to create a new dataset for AI analysis. The feature set comprised 13 variables, including one stride time, as well as three linear acceleration and three angular velocity measurements for the pelvis and lower legs. Training and testing datasets were created by randomly splitting the dataset in an 8:2 ratio. The training and testing sets comprised 9838 and 2464 samples, respectively. The allocation of samples from individual participants was performed with care to ensure that there was no overlap between the training and testing datasets in terms of participant data.

3.3 | Training dataset

The training set comprised 9236 negative samples from 660 participants classified as negative cases and 602 positive samples from 43 participants classified as positive cases. However, random sampling was performed to extract 602 samples from all 9236 negative samples such that the number of negative and positive samples was equal because the imbalance between the number of negative and positive samples in the training dataset can negatively influence the learning of models. A training dataset comprising 1204 samples was created by combining 602 randomly extracted negative samples and 602 positive samples. The random sampling of the negative cases was repeated 29 times to obtain 30 different training datasets comprising 1204 samples. Additional data augmentation was not performed during the process.

We attempted to prevent the degradation of model accuracy due to data imbalance by using several methods, including oversampling with the SMOTE library. However, we found that the most accurate model was produced using a method similar to bootstrapping. According to Blagus and Lusa,[25] oversampling can lead to overly optimistic estimates, while undersampling can avoid such optimistic estimates in machine learning methods that use cross-validation similar to our method (bootstrapping). Thus, we adopted undersampling based on the validity supported by their findings. In this approach, we created 30 training datasets by randomly sampling negative cases from a large dataset to match the number of positive cases. As discussing which sampling method is more useful is beyond the scope of this study, we have omitted detailed comparisons of the sampling methods.

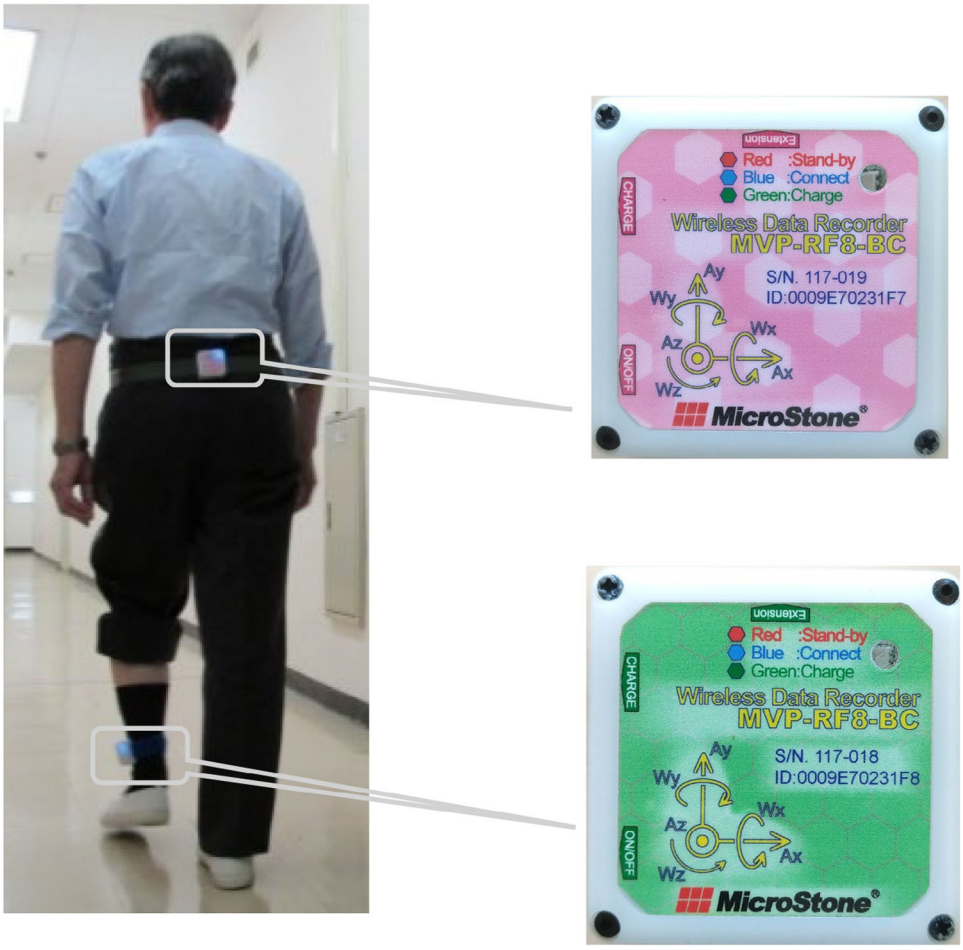

FIGURE 1 Attachment positions of the triaxial sensors

3.4 | Testing dataset

The testing dataset comprised 2310 samples from 165 participants classified as negative cases and 154 samples from 11 participants classified as positive cases. Random sampling of the available negative samples was performed as described in the preceding section as a large imbalance between negative and positive cases could artificially inflate the accuracy of the AI model and hinder the evaluation of the model. In total, 154 samples were extracted from the 2310 negative samples in the original testing dataset during each random sampling and combined with the 154 positive samples in the original testing dataset. Thus, a smaller but more balanced testing dataset comprising 308 samples was created. Random sampling was performed 30 times to yield 30 new testing datasets that differed from each other. These sets were used to validate the classification model.

3.5 | Deep learning model

The model comprised the channel attention mechanism, multilayer perceptron, and SoftMax (Figure 2A). The 13 feature datasets (13 channels) were the input, and the probability of being classified into a class (presence or absence of cognitive impairment) was the output. A fully connected layer was placed before the SoftMax layer. Figure 2B depicts the internal structure of the sublayer, comprising a linear layer, GELU activation function, and Batch Norm 1D layer. The internal

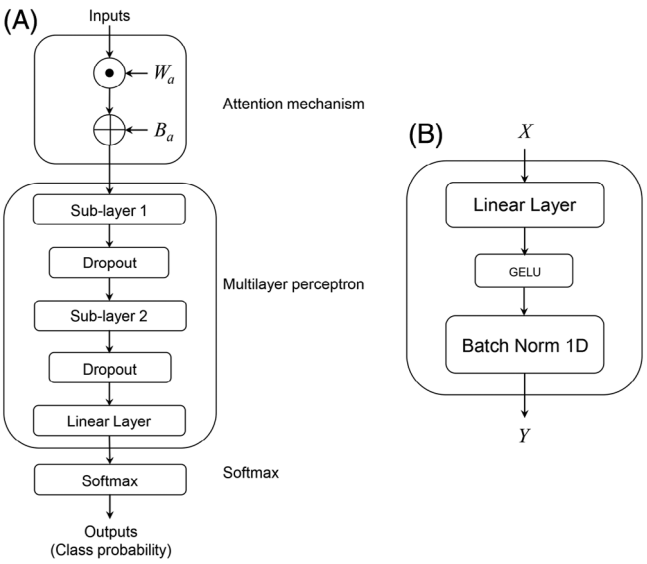

FIGURE 2 Deep learning model used in this study. A, Structure of the model and (B) the internal structure of sublayers 1 and 2

TABLE 1 Parameters of the deep learning model.

Parameter	Setting
Optimizer	AdamW
Learning rate	1.E-03
Betas	(0.9, 0.999)
Eps	1.E-08
Weight decay	0.01
Max epoch	1000
Dropout	0.2
Label smoothing	0.2

structures of the two sublayers in the multilayer perceptron were identical. Sublayers 1 and 2 consisted of 8 and 4 units, respectively. The liner layer consisted of 2 units.

Python (version 3.7.16) and PyTorch (version 1.10.2) were used to test the proposed model. The PyTorch seeds were set at 100 during the experiment. AdamW was used as the optimizer during training, and the learning rate was set to 1e-3. The remaining parameters were maintained at their default values using PyTorch. The number of epochs was set to 1000 during training, the model dropout was set to 0.2, and label smoothing was set to 0.2 as a regularization method to suppress overfitting. To address the concern about overfitting and the choice of parameters, we conducted experiments using grid search with varying values for label smoothing and dropout rate. Specifically, we tested the following three conditions: [0,0], [0.1,0.1], and [0.2,0.2]. Table 1 presents the details of the learning parameters.

3.6 | Statistical analysis

Differences in demographic characteristics between the negative and positive groups were assessed using the *t* test and chi-squared test. The dimensions of the features were reduced to two using principal component analysis (PCA), and a scatter plot was drawn for each dataset to determine whether the characteristics of the training and testing datasets were identical. In addition, the accuracy, F1 score, sensitivity, specificity, false negative and positive rates, and positive and negative predictive values were calculated for each testing dataset. The means, standard deviations, and ranges of these measures were calculated subsequently. The positive predictability of each sample was determined using the final output of the SoftMax function in the AI model. The areas under the curve (AUCs) of each testing dataset were calculated. The overall receiver operating characteristic (ROC) curve was created using 2464 samples in the testing dataset. All statistical analyses were performed using SPSS version 27 (IBM Corporation).

4 | RESULTS

Of the participants analyzed, 346 were men and 533 were women. The analysis included 825 healthy individuals (negative case) and 54 individuals (positive case) with cognitive impairment. The demographic characteristics of the participants in the positive and negative groups are shown in Tables 2 and 3. Participants were significantly older in the positive group than in the negative group, and normal walking speed, maximum walking speed, and MMSE scores were significantly higher in the negative group than in the positive group. The proportion of women was significantly higher in the negative group than in the positive group. The prevalence of hypertension was higher in the positive group than in the negative group.

The PCA plots of the training and testing datasets were similar (Figure 3). Using grid search, we experimented with the hyperparameters of label smoothing and dropout at values of [0,0], [0.1,0.1], and [0.2,0.2] (Figure 4). We observed excessive overfitting in the [0,0] condition (no label smoothing and no dropout). The training accuracy was significantly higher than the testing accuracy, indicating that the model was not generalizing well to unseen data. In the [0.1,0.1] condition, there was still noticeable overfitting, but to a lesser extent compared to the [0,0] condition. The gap between training and testing accuracy persisted, suggesting that the regularization was insufficient. In the [0.2,0.2] condition, the overfitting was considerably reduced. The training and testing accuracies were more aligned, indicating better generalization. Therefore, we adopted this setting for our final model. Table 4 presents the basic statistics of the accuracy measurements for the 30 testing datasets. The mean accuracy, F1 score, sensitivity, and specificity of the 30 testing datasets were 0.802, 0.831, 0.961, and 0.643, respectively. The mean AUC was 0.833. Figure 5 presents the ROC curve for the total of 2464 testing dataset samples.

5 | DISCUSSION

This study underscores the efficacy of using linear acceleration and angular velocity data acquired during one comfortable walking stride to distinguish community-dwelling older adults with and without cognitive impairment. Notably, a high average sensitivity of 0.961 was observed, indicating that most individuals with cognitive impairment were correctly identified. However, the study's specificity, at 0.643, was low, indicating a moderate false positive rate. Nevertheless, the positive predictive value, at 0.736, was moderate, and the negative predictive value, at 0.941, was particularly robust, underscoring the reliability and use of this AI model as an initial screening tool.

Ruengchaijatuporn et al. used the AI interpretations of the clock-drawing test to differentiate between individuals with and without MCI and reported F1 scores of 0.74 to 0.84.[26] The model developed in the present study yielded comparable results, with F1 scores of 0.736 to 0.928. The model used herein may be a more suitable initial screening tool for cognitive impairment than the method proposed by Ruengchaijatuporn et al.[26] as it is simple, and the data can be collected non-intrusively. Our method, if it can be deployed using a cell phone or other wearable technology (watch), does not require the patient to make changes in their daily routine. They do not have to visit a clinic but can be monitored remotely.

The significance of the model developed herein lies in its accuracy and applicability. Unlike the models developed in previous studies,

TABLE 2 Demographic characteristics of the study participants.

	Negative ($n = 825$)		Positive ($n = 54$)		Overall ($N = 879$)	
	Mean	SD	Mean	SD	Mean	SD
Age, years	73.3	4.97	76.7[†]	4.65	73.5	5.02
Height, cm	155.6	8.39	156.2	8.07	155.6	8.37
Weight, kg	55.5	10.22	54.9	9.41	55.4	10.17
Grip strength, kg	25.6	7.88	24.8	8.16	25.5	7.90
Normal walking speed, m/s	1.4[a]	0.22	1.2	0.29	1.4	0.23
Maximum walking speed, m/s	1.9[a]	0.33	1.6	0.39	1.8	0.34
MMSE score	28.1[a]	1.60	21.1	2.64	27.7	2.39

Abbreviations: MMSE, Mini-Mental State Examination; SD, standard deviation.
[a]Significantly higher than the other group (paired *t* test, $P < 0.05$).

TABLE 3 Sex distribution and the prevalence of selected chronic diseases among the participants.

	Negative ($n = 825$)		Positive ($n = 54$)		Overall ($N = 879$)	
	Number of participants	Percentage of participants	Number of participants	Percentage of participants	Number of participants	Percentage of participants
Women	513	62.2[a]	20	37.0	533	60.6
Chronic disease						
Hypertension	371	45.0	28	51.9[a]	399	45.4
Stroke	40	4.8	6	11.1	46	5.2
Heart disease	133	16.1	8	14.8	141	16.1
Diabetes	94	11.4	9	16.7	103	11.7

[a]Significantly higher than the other group (chi-squared test, $P < 0.05$).

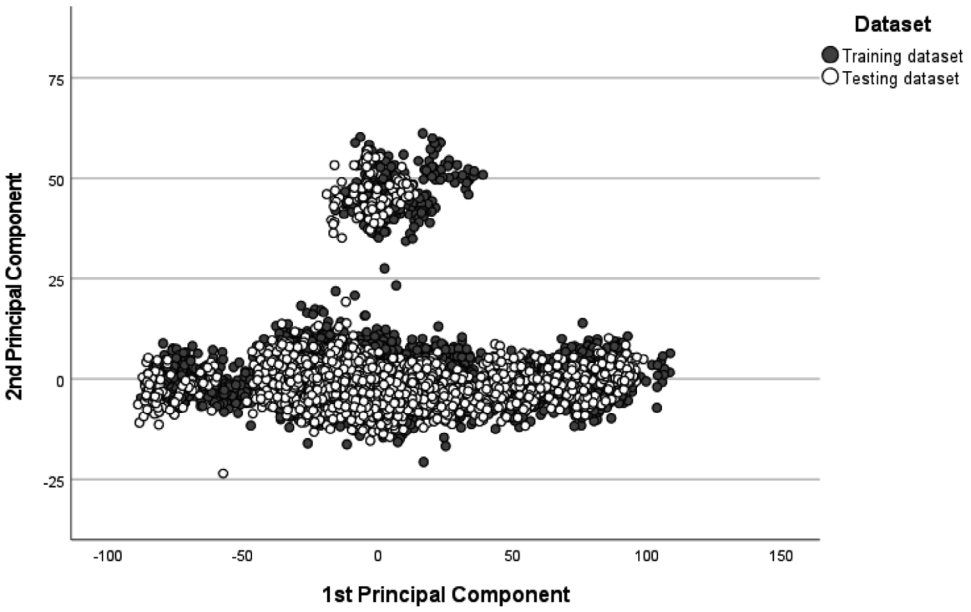

FIGURE 3 Scattergram of the principal component analysis of the training and testing datasets

the AI model developed herein was constructed using data from a community-dwelling cohort. AI models rely heavily on training data. Therefore, they should be constructed using datasets that are faithful to the classification target.[27] Secondary gait disturbances reportedly become more pronounced as the duration since the dementia diagnosis increases.[28] Thus, gait differences in these individuals were more pronounced than those in the community-dwelling cohort data. Significant differences in the datasets can enhance the accuracy of the

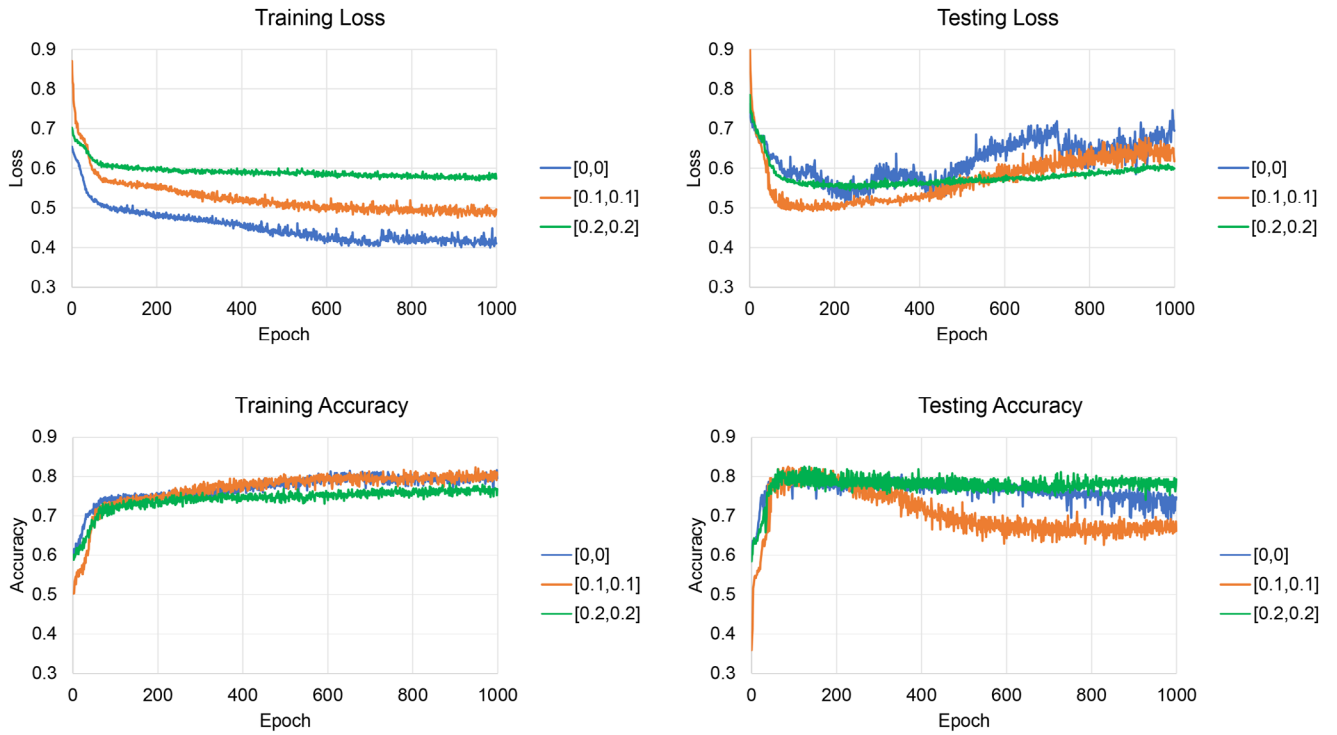

FIGURE 4 Results of the experiments using grid search with varying values for label smoothing and dropout rate

TABLE 4 Statistical performance metrics of the artificial intelligence model across 30 testing datasets.

	Mean	Standard deviation	Minimum	Maximum	Count
Accuracy	0.802	0.063	0.656	0.925	30
F1 score	0.831	0.046	0.736	0.928	30
Sensitivity	0.961	0.000	0.961	0.961	30
Specificity	0.643	0.126	0.351	0.890	30
False negative rate	0.039	0.000	0.039	0.039	30
False positive rate	0.357	0.126	0.110	0.649	30
Positive predictive value	0.736	0.072	0.597	0.897	30
Negative predictive value	0.941	0.012	0.900	0.958	30
AUC	0.833	0.072	0.705	0.976	30

Abbreviation: AUC, area under curve.

AI models; however, the present study demonstrated that the accuracy of community-based datasets was similar to or even greater than that of models developed with traditional case–control datasets. To our knowledge, this is one of the few studies using community-based research to effectively assess cognitive impairment using AI, indicating the potential applicability of our AI model in real-world social settings.[29]

Individuals with cognitive impairment are often underdiagnosed, as demonstrated by comprehensive surveys in our region, which revealed that many individuals with MCI or dementia remain undiagnosed.[30] Thus, by the time they undergo a cognitive function test, they often present with dementia that has progressed to a point at which interventions are no longer effective.[31] Consequently, an increasing need exists to develop a method to screen for cognitive impairment in a healthy state, which can be administered conveniently, thereby facilitating the implementation of appropriate interventions on a timely basis.[32] Decentralized clinical trials (DCTs)[32,33] are under way to address these issues. DCTs actively enroll individuals in a research project at a healthy stage and monitor their cognitive function regularly without them having to attend specialized medical institutions. Recommendations for further screening tests, such as MMSE, MoCA, or computed tomography scans, are made if cognitive impairment is detected. Multilayered screening can help detect early dementia symptoms and facilitate more efficient medical resource use.[34] The streamlined method used in the present study for daily cognitive function monitoring may be especially beneficial for screening purposes. Moreover, it can be implemented as part of a multilayer screening procedure. Furthermore, this AI model can be integrated into

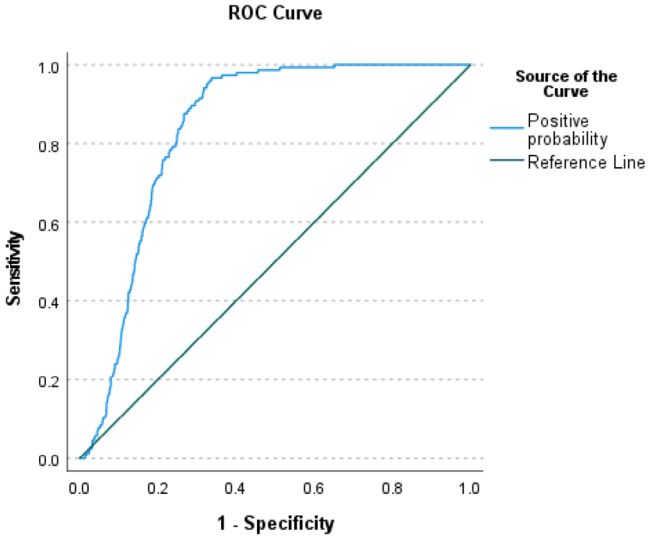

FIGURE 5 Receiver operating characteristic (ROC) curve for the testing dataset samples

smartphone applications as the sensors used herein (accelerometers and gyroscopes) are standard smartphone components. Achieving this goal will open new avenues for screening applications, thus highlighting the greater social impact of our study.

The classification of multichannel time-series signals was developed by exploring various AI model architectures. Typically, longer short-term memory-based models[35] are preferred as they can handle sequential data; however, these models are prone to overfitting. Convolutional neural network-based models are often implemented after converting the data into images via wavelet analysis or similar techniques.[36] However, these AI models could not detect cognitive impairment with sufficient accuracy in our preliminary analysis. Therefore, a multilayer perceptron-based model enhanced with an attention mechanism[37,38] was designed using the global average pooling method to achieve high discrimination accuracy with fewer computational resources. The present study's findings are promising, demonstrating moderate classification accuracy with an average F1 score of 0.780.

Although our AI model exhibited sufficient accuracy for screening, its accuracy can be improved further by adding the variance of each data feature rather than using only their average values as inputs. Abnormalities in gait rhythm correlate with cognitive function decline.[39,40] The decline in gait speed is more closely related to gait rhythm rather than stride length.[41–43] Moreover, compared to those in the highest quartile, the participants in the lowest quartile of gait rhythms were 1.89 and 1.66 times more likely to develop cognitive impairment and dementia, respectively.[43] This rhythmic abnormality also resulted in a decreased gait rhythm and an increased variability. Compared to low or moderate variability, high variability in gait was associated with a 12-fold higher risk of developing MCI in a 4-year prospective study.[19] Gait rate and variability were related to dementia onset in the study by Darweesh et al.[44] Increased variability in walking rhythm is also associated with white matter hyperintensity.[45] These studies' findings underscore the potential for improving the predictive capabilities of the AI model by adding measures of variability.

The present study has certain limitations. The primary concern is the limited number of individuals with cognitive impairment included in the study. A bootstrap method was used to statistically estimate the range of sensitivity and specificity to address this limitation. However, using a large sample of individuals with cognitive impairment is preferred for creating a robust AI model. The high sensitivity compared to the specificity obtained in this study may also be because positive samples were used repeatedly in the training and testing datasets. Currently, guidelines regarding the sample number required in a dataset for classification by AI have not been established; however, 54 positive cases is considered small. In such cases, learning may not converge, or overfitting may occur. Fortunately, an accurate model was created in the present study despite these limitations. Nevertheless, increasing the case number and verifying the validity would be advantageous. The prevalence of cognitive impairment among community-dwelling older adults was 6% in the present study, suggesting that ≈ 10 times the number of participants in the cohort must be enrolled to obtain enough cases for constructing the AI model. The diagnosis of positive cases was based on MMSE scores alone. A more detailed diagnosis may be necessary. We could not examine the detailed demographics of participants, such as the history of medications, depression, and neuromuscular abnormalities, in this study. In addition, low MMSE scores can be attributed not only to cognitive impairment but also to inattention, drowsiness due to poor sleeping quality, and uncooperative disposition, thereby limiting the prediction of future dementia using only one MMSE test. In this study, we used 10-m steady-state walking data to develop our model; however, the effects of cognitive impairment may be more pronounced during the initiation or turning phases of walking. We believe further research is warranted to analyze various phases of walking for improving the model's accuracy.

Second, the present study used cross-sectional data to construct the AI model. Although the AI model could identify individuals with cognitive impairment, it could not predict future cognitive decline. The current AI model has the potential to be used as a screening tool for detecting cognitive impairment in older adults. To establish the model's utility as a predictive tool, future studies should use a prospective design. It would be preferable for an AI model to predict cognitive decline years in advance to facilitate earlier interventions, given that the degenerative changes responsible for cognitive decline develop relatively early.[46] Additionally, we did not perform hyperparameter tuning as we achieved sufficient accuracy with the initial settings in this study. However, hyperparameter tuning could potentially lead to slight improvements in model accuracy. Further research exploring this aspect is warranted to optimize the predictive model's performance. Because our study aims to explore the potential sensor-based application for issuing precautionary alerts before clinical visits, we limited our model to variables that can be obtained using sensors; thus, we did not account for any covariates or comorbidities. However, future studies should consider the covariates that can affect walking.

Finally, the AI model developed in the present study was constructed to classify individuals with an MMSE score < 24, a threshold commonly used to detect dementia. Thus, participants in a transitional state, known as those with MCI, were included in the healthy group.

These participants may have exhibited a gait pattern similar to that observed in individuals with cognitive impairment despite possessing MMSE scores > 24. Owing to our model's relatively low specificity, the number of false positives is not insignificant. A web-based cognitive test, Compbased-CAT, has been designed to detect and predict MCI[47,48] in such participants. This secondary screening tool can help distinguish the presence of MCI in patients who were initially identified as false positives.

In summary, the AI analysis of the acceleration and angular velocity of one stride during comfortable walking used in the present study could reasonably identify cognitive impairment in community-dwelling older adults. Fostering and motivating further research in this field with other datasets is warranted.

ACKNOWLEDGMENTS

We wish to extend our heartfelt appreciation to Mr. Tomoketsu Senri for his invaluable guidance in navigating the intricate landscape of AI model selection and fine-tuning learning parameters. His expertise has been instrumental in our journey, and we are profoundly grateful for his contributions. This work was supported by a Health Labour Sciences Research Grant (nos. H23-Choju-Ippan-001 and 002) from the Ministry of Health, Labour and Welfare of Japan and JSPS KAKENHI (grant number: 23K06919).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest. Author disclosures are available in the supporting information.

CONSENT STATEMENT

This study upholds the principles of diversity, equity, and inclusion (DEI) and adheres to the Declaration of Helsinki guidelines for ethical research. The research protocol was approved by the ethics committee of the affiliated institution (approval number: R22-081). Written informed consent was obtained from all participants.

ORCID

Hisashi Kawai https://orcid.org/0000-0003-3015-6041

REFERENCES

1. Morley JE. Mild cognitive impairment-a treatable condition. *J Am Med Dir Assoc*. 2014;15:1-5.
2. Mansbach WE, Mace RA, Tanner MA, Simmons SP. A path to early diagnosis of mild cognitive impairment and dementia: validity and reliability of the myMemCheck® self-administered screening tool. *Fam Pract*. 2020;37:561-567.
3. Tsai PH, Liu JL, Lin KN, et al. Development and validation of a dementia screening tool for primary care in Taiwan: brain health test. *PLoS One*. 2018;13:e0196214.
4. Whelan R, Barbey FM, Cominetti MR, Gillan CM, Rosicka AM. Developments in scalable strategies for detecting early markers of cognitive decline. *Transl Psychiatry*. 2022;12:473.
5. Makino K, Lee S, Bae S, et al. Development and validation of new screening tool for predicting dementia risk in community-dwelling older Japanese adults. *J Transl Med*. 2021;19:448.
6. Navickaite E, Saltvedt I, Lydersen S, et al. Diagnostic accuracy of the Clock Drawing Test in screening for early post-stroke neurocognitive disorder: the Nor-COAST study. *BMC Neurol*. 2024;24:22.
7. Karimi L, Mahboub-Ahari A, Jahangiry L, Sadeghi-Bazargani H, Farahbakhsh M. A systematic review and meta-analysis of studies on screening for mild cognitive impairment in primary healthcare. *BMC Psychiatry*. 2022;22:97.
8. Siqueira GSA, Hagemann PMS, Coelho DS, Santos FHD, Bertolucci PHF. Can MoCA and MMSE be interchangeable cognitive screening tools? a systematic review. *Gerontologist*. 2019;59:e743-e763.
9. De Roeck EE, De Deyn PP, Dierckx E, Engelborghs S. Brief cognitive screening instruments for early detection of Alzheimer's disease: a systematic review. *Alzheimers Res Ther*. 2019;11:21.
10. Wang H, Sheng L, Xu S, et al. Develop a diagnostic tool for dementia using machine learning and non-imaging features. *Front Aging Neurosci*. 2022;14:945274.
11. Masugi Y, Kawai H, Ejiri M, et al. Early strong predictors of decline in instrumental activities of daily living in community-dwelling older Japanese people. *PLoS One*. 2022;17:e0266614.
12. Verghese J, Wang C, Bennett DA, Lipton RB, Katz MJ, Ayers E. Motoric cognitive risk syndrome and predictors of transition to dementia: a multicenter study. *Alzheimers Dement*. 2019;15:870-877.
13. Jayakody O, Breslin M, Ayers E, et al. Relative trajectories of gait and cognitive decline in aging. *J Gerontol A Biol Sci Med Sci*. 2022;77:1230-1238.
14. Byun S, Han JW, Kim TH, et al. Gait variability can predict the risk of cognitive decline in cognitively normal older people. *Dement Geriatr Cogn Disord*. 2018;45:251-261.
15. Grande G, Triolo F, Nuara A, Welmer AK, Fratiglioni L, Vetrano DL. Measuring gait speed to better identify prodromal dementia. *Exp Gerontol*. 2019;124:110625.
16. Jing Y, Qin P, Fan X, et al. Deep learning-assisted gait parameter assessment for neurodegenerative diseases: model development and validation. *J Med Internet Res*. 2023;25:e46427.
17. Cicirelli G, Impedovo D, Dentamaro V, Marani R, Pirlo G, D'Orazio TR. Human gait analysis in neurodegenerative diseases: a review. *IEEE J Biomed Health Inform*. 2022;26:229-242.
18. Hansson O. Biomarkers for neurodegenerative diseases. *Nat Med*. 2021;27:954-963.
19. Mc Ardle R, Galna B, Donaghy P, Thomas A, Rochester L. Do Alzheimer's and Lewy body disease have discrete pathological signatures of gait?. *Alzheimers Dement*. 2019;15:1367-1377.
20. Zhou H, Park C, Shahbazi M, et al. Digital biomarkers of cognitive frailty: the value of detailed gait assessment beyond gait speed. *Gerontology*. 2022;68:224-233.
21. Chan LLY, Choi TCM, Lord SR, Brodie MA. Development and large-scale validation of the Watch Walk wrist-worn digital gait biomarkers. *Sci Rep*. 2022;12:16211.
22. Zhang Z, Wang L, Lee C. Recent Advances in Artificial Intelligence Sensors. *Adv Sens Res*. 2023:2:2200072.
23. Ghoraani B, Boettcher LN, Hssayeni MD, Rosenfeld A, Tolea MI, Galvin JE. Detection of mild cognitive impairment and Alzheimer's Disease using dual-task gait assessments and machine learning. *Biomed Signal Process Control*. 2021;64:102249.
24. Jeon Y, Ho TKK, Kang J, et al. Machine learning–based detection model of early Alzheimer's disease using wearable device and gait assessment. *Alzheimers Dement*. 2021;17(Suppl. 5):e057563.
25. Blagus R, Lusa L. Joint use of over-and under-sampling techniques and cross-validation for the development and assessment of prediction models. *BMC Bioinformatics*. 2015;16:363.
26. Ruengchaijatuporn N, Chatnuntawech I, Teerapittayanon S, et al. An explainable self-attention deep neural network for detecting mild cognitive impairment using multi-input digital drawing tasks. *Alzheimers Res Ther*. 2022;14:111.

27. Nguyen TV, Dakka MA, Diakiw SM, et al. A novel decentralized federated learning approach to train on globally distributed, poor quality, and protected private medical data. *Sci Rep*. 2022;12:8888.
28. Stephen R, Barbera M, Peters R, et al. Development of the first WHO guidelines for risk reduction of cognitive decline and dementia: lessons learned and future directions. *Front Neurol*. 2021;12:763573.
29. Gamble LD, Matthews FE, Jones IR, et al. Characteristics of people living with undiagnosed dementia: findings from the CFAS Wales study. *BMC Geriatr*. 2022;22:409.
30. Ura C, Okamura T, Inagaki H, et al. Characteristics of detected and undetected dementia among community-dwelling older people in Metropolitan Tokyo. *Geriatr Gerontol Int*. 2020;20:564-570.
31. Walter E, Pinquart M. How effective are dementia caregiver interventions? An updated comprehensive meta-analysis. *Gerontologist*. 2020;60:609-619.
32. De Brouwer W, Patel CJ, Manrai AK, Rodriguez-Chavez IR, Shah NR. Empowering clinical research in a decentralized world. *NPJ Digit Med*. 2021;4:102.
33. Goodson N, Wicks P, Morgan J, Hashem L, Callinan S, Reites J. Opportunities and counterintuitive challenges for decentralized clinical trials to broaden participant inclusion. *NPJ Digit Med*. 2022;5:58.
34. Contador J, Vargas-Martinez AM, Sanchez-Valle R, Trapero-Bertran M, Llado A. Cost-effectiveness of Alzheimer's disease CSF biomarkers and amyloid-PET in early-onset cognitive impairment diagnosis. *Eur Arch Psychiatry Clin Neurosci*. 2023;273:243-252.
35. Cinar A, Tuncer SA. Classification of normal sinus rhythm, abnormal arrhythmia and congestive heart failure ECG signals using LSTM and hybrid CNN-SVM deep neural networks. *Comput Methods Biomech Biomed Engin*. 2021;24:203-214.
36. Cheng J, Zou Q, Zhao Y. ECG signal classification based on deep CNN and BiLSTM. *BMC Med Inform Decis Mak*. 2021;21:365.
37. Qiao M, Liang Y, Tavares A, Shi X. Multilayer perceptron network optimization for chaotic time series modeling. *Entropy*. 2023:25:973.
38. Raghavendra P, Charitha C, Begum KG, Prasath VBS. Deep learning-based skin lesion multi-class classification with global average pooling improvement. *J Digit Imaging*. 2023;36:2227-2248.
39. Verghese J, Wang C, Lipton RB, Holtzer R, Xue X. Quantitative gait dysfunction and risk of cognitive decline and dementia. *J Neurol Neurosurg Psychiatry*. 2007;78:929-935.
40. Tuena C, Maestri S, Serino S, Pedroli E, Stramba-Badiale M, Riva G. Alzheimer's disease neuroimaging initiative. Prognostic relevance of gait-related cognitive functions for dementia conversion in amnestic mild cognitive impairment. *BMC Geriatr*. 2023;23:462.
41. Morris R, Lord S, Lawson RA, et al. Gait rather than cognition predicts decline in specific cognitive domains in early Parkinson's disease. *J Gerontol A Biol Sci Med Sci*. 2017;72:1656-1662.
42. Hausdorff JM. Gait variability: methods, modeling and meaning. *J Neuroeng Rehabil*. 2005;2:19.
43. Quan M, Xun P, Chen C, et al. Walking pace and the risk of cognitive decline and dementia in elderly populations: a meta-analysis of prospective cohort studies. *J Gerontol A Biol Sci Med Sci*. 2017;72:266-270.
44. Darweesh SKL, Licher S, Wolters FJ, Koudstaal PJ, Ikram MK, Ikram MA. Quantitative gait, cognitive decline, and incident dementia: the Rotterdam study. *Alzheimers Dement*. 2019;15:1264-1273.
45. Sakurai R, Inagaki H, Tokumaru AM, et al. Differences in the association between white matter hyperintensities and gait performance among older adults with and without cognitive impairment. *Geriatr Gerontol Int*. 2021;21:313-320.
46. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7:280-292.
47. Takahashi J, Kawai H, Suzuki H, et al. Development and validity of the computer-based cognitive assessment tool for intervention in community-dwelling older individuals. *Geriatr Gerontol Int*. 2020;20:171-175.
48. Takahashi J, Kawai H, Ejiri M, et al. Predicting the incidence of mild cognitive impairment with a computer-based cognitive assessment tool in community-dwelling older adults: the Otassha study. *PLoS One*. 2024;19:e0297433.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Obuchi SP, Kojima M, Suzuki H, et al. Artificial intelligence detection of cognitive impairment in older adults during walking. *Alzheimer's Dement*. 2024;16:e70012. https://doi.org/10.1002/dad2.70012