Proposal for an Early Detection Model of Cognitive Decline Based on Drawing-Process Data

1st Shun Hirai
Graduate School of Engineering
Kobe University
1-1 Rokkodai-cho, Nada, Kobe, 657-8501, Japan
hirashun@es4.eedept.kobe-u.ac.jp

2nd Keisuke Yoshida
Graduate School of Engineering
Kobe University
1-1 Rokkodai-cho, Nada, Kobe, 657-8501, Japan
yoshikei@es4.eedept.kobe-u.ac.jp

3rd Sachio Saiki
School of Data and Innovation
Kochi University of Technology
185 Miyanokuchi, Tosayamada, Kami, Kochi, 782-8502, Japan
saiki.sachio@kochi-tech.ac.jp

4th Naoki Kodama
Department of Medical Radiology
Niigata University of Health and Welfare
1398 Shimami-cho, Kita-ku, Niigata, 950–3198 Japan

5th Atsushi Satou
Faculty of Health and Medical Sciences
Aichi Shukutoku University
Katahira, Nagakute-city, Aichi Prefecture, 480-1197 Japan

6th Masahide Nakamura
Center of Mathematical and Data Science
Kobe University
1-1 Rokkodai-cho, Nada, Kobe, 657-8501, Japan
masa-n@cmds.kobe-u.ac.jp

***Abstract*—Japan is experiencing rapid population aging, and the shortage of healthcare workers has become a serious issue. In parallel, the number of people with dementia is increasing, raising the importance of early detection and robust screening methods. To address these challenges, we developed the tablet-based drawing test app EVIDENT. In this study, we propose an early screening method for cognitive decline using drawing-process data from the Clock Drawing Test (CDT) and the Cube Copying Test (CCT) collected with these apps. Specifically, we build machine learning models that use process features as explanatory variables and scoring results as target variables. We also perform clustering of the drawing-process data to reveal drawing-process patterns in participants with cognitive decline and to summarize their characteristics. Through evaluation experiments, we demonstrated the effectiveness of the proposed models and extract process features that contribute to early screening. These results suggest that our approach can serve as a basis for supporting early detection in real-world operations, including group testing, while reducing burden at the point of care. In future work, we will expand the dataset, further improve screening accuracy, and pursue applications to diagnostic support integrated with EVIDENT.**

***Index Terms*—dementia, screening, machine learning, clustering**

I. Introduction

In Japan, the rapid progression of a super-aging society has led to simultaneous increases in medical demand and shortages of healthcare workers. To address this shortage, medical DX initiatives incorporating electronic procedures and online operations are being promoted in various areas. On the other hand, with the increase in the elderly population, the number of dementia patients is also rising, making early detection and intervention while symptoms are still mild crucial. Therefore, there is a need for screening tests that can be conducted in a short time and have high reliability.

The Clock Drawing Test (CDT) [1] and the Cube Copying Test (CCT) [2] are representative drawing tests widely used in clinical practice [3]. Traditionally, it has been common to visually score the ”finished drawing.” However, in actual tests, drawing-process data such as ”the order in which the drawing was made,” ”the speed of drawing,” and ”the time taken to start drawing after thinking” can also be obtained. Combining these process information may lead to earlier detection. Additionally, for widespread and continuous regular evaluation of cognitive function, it is important to have a system that can efficiently conduct tests in groups rather than individually, and frameworks for this purpose have also been proposed.

We have previously proposed and implemented a web application called **EVIDENT** that provides healthcare workers with easily understandable visualizations of process information such as stroke count, speed, pen pressure, without relying on bulk automatic scoring using machine learning [4] [5]. Through demonstration experiments using these apps, we have collected drawing data [6] and are operating them as group tests. However, there are two challenges that remain.
P1: Development of a method to determine cognitive decline without waiting for healthcare workers' scoring
P2: Development of a method to detect cognitive decline earlier than conventional methods using drawing-process data

In this study, we aim to develop an early screening method

for cognitive decline based on the above challenges. To achieve this goal, we propose the following two approaches.

A1: Construction of a machine learning model using drawing-process data

We construct a machine learning model using drawing-process data as explanatory variables and scoring results as target variables. This allows for automatic determination of cognitive decline, reducing the burden on healthcare workers

A2: Hierarchical clustering of drawing-process data

We perform hierarchical clustering of drawing-process data to extract characteristics of participants with cognitive decline. By analyzing drawing process features that do not rely on cognitive results, we can utilize this information to construct early screening methods.

Based on these approaches, we conducted evaluation experiments and demonstrated the effectiveness of the machine learning model, revealing drawing patterns of participants with cognitive decline.

In the future, we aim to improve the accuracy of early screening through further data collection and implement screening functions in EVIDENT. By conducting tests using EVIDENT, we aim to build a system that can detect cognitive decline earlier than conventional scoring methods based on drawing results, without the intervention of healthcare workers.

This study has been approved by the Ethics Committee of Niigata University of Health and Welfare (Approval No. 19323-240614).

II. Preliminaries

A. Aging Society and Dementia

In Japan, the advancement of a super-aging society has led to increased medical demand and a shortage of personnel in medical settings. According to estimates by the Ministry of Health, Labour and Welfare, there is expected to be a shortage of approximately 30,000 doctors by 2030 [7]. To compensate for these personnel constraints, efforts to improve operational efficiency through digitalization, such as electronic medical records and online medical interviews, i.e., medical DX initiatives, are being promoted. Additionally, the spread of online medical care is believed to contribute to reducing the burden of travel associated with home visits and narrowing disparities in medical access between regions.

On the other hand, **dementia** has become an important social issue in Japan due to aging. The White Paper on Aging Society by the Cabinet Office estimates that the number of elderly people with dementia will increase to approximately 5.84 million by 2040, accounting for about one in six people aged 65 and older [8]. Since early detection and diagnosis of dementia directly lead to subsequent support plans and treatment choices, continuous screening is recommended. In particular, it is important to capture changes at the stage of Mild Cognitive Impairment (MCI), which is considered a precursor to dementia, and regular cognitive function evaluation is required [9] [10].

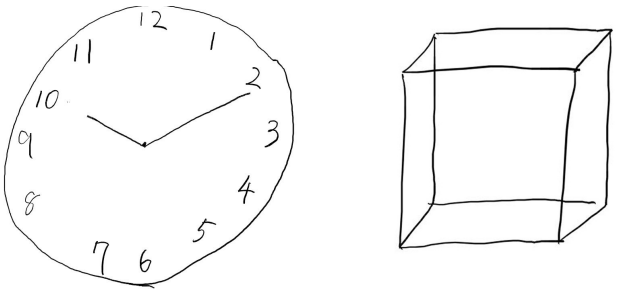

Fig. 1. Examples of drawing results (Left: CDT, Right: CCT)

B. Neuropsychological Drawing Tests

Neuropsychology is a discipline that investigates brain functions such as language, memory, attention, executive control, and visuospatial cognition, linking these to the assessment and intervention of related disorders [11]. Among its various standardized clinical tests, drawing-based tasks are particularly useful because they allow simultaneous observation of visuospatial construction, planning ability, attention, and motor control within a short time. These tests are widely applied for screening and monitoring dementia and higher brain dysfunction.

Representative examples include the **Clock Drawing Test (CDT)** and the **Cube Copying Test (CCT)**. In CDT, participants are asked to draw an analog clock showing a specified time (e.g., 10:10) on a sheet of paper [1]. In CCT, they are instructed to reproduce a given cube in perspective on paper [2]. Both tests provide insight into visuospatial and executive functions, and examples of typical drawings are shown in Fig. 1.

C. Prior Work

We developed the web application **EVIDENT** (Extraction and Visualization Interface of Drawing Execution in Neuropsychological Tests) [4] [5] for digital administration of the Clock Drawing Test (CDT) and Cube Copying Test (CCT) on tablets. The system records not only the final drawings but also detailed **drawing-process data** such as stroke count, speed, and pen pressure.

Previous studies have shown that such drawing-process data reflect subtle motor and cognitive changes related to Mild Cognitive Impairment (MCI) and dementia, for example slower drawing speed, longer pauses, or unstable pen control [12]–[14]. However, most of these studies analyzed isolated measures (e.g., speed or pause duration) and did not conduct integrated screening based on multiple process features. Moreover, prior work mainly aimed to reproduce clinician scoring, whereas practical screening in clinical and community settings requires automatic, multi-feature analysis that reduces the manual workload of healthcare workers. The current EVIDENT system supports data collection and visualization, but does not yet provide such automated early-risk screening capability.

D. Challenges to Address

As mentioned in Section II-C, we are collecting drawing test data using EVIDENT, including detailed drawing-process data. At present, EVIDENT mainly visualizes this information for healthcare workers after the test. To make it useful as an *early* screening tool in practice, especially in group testing where expert time is limited, two specific challenges arise.

P1: Development of a method to determine cognitive decline without waiting for healthcare workers' scoring. Currently, healthcare workers visually score the drawing results each time data is collected, which is time-consuming and burdensome. Addressing P1 means providing an indication of possible decline *immediately after the test*, before manual scoring is completed. In other words, P1 targets temporal early detection—shortening the delay between test administration and feedback.

P2: Development of a method to detect cognitive decline earlier than conventional methods using drawing-process data. In current drawing tests, signs of cognitive decline are mainly judged from the final shape and layout of the drawing. However, drawing-process data include stroke count, speed, and pen pressure, which may start to change even when the completed drawing still appears normal. Addressing P2 means capturing such subtle process-level changes and using them as an early clinical indicator, corresponding to clinical early detection.

These two challenges define the direction of this study.

III. Proposed Method

A. Objectives and Approaches

The objective of this study is to address the challenges mentioned in Section II-D by utilizing data collected using EVIDENT to develop an early screening method for cognitive decline. In this study, we focus on two aspects of early detection: temporal early detection, which involves determining cognitive decline without waiting for healthcare workers' scoring, and clinical early detection, which involves identifying features that cannot be determined from drawing results to detect cognitive decline earlier than conventional methods. The key ideas are the construction of a machine learning model using drawing-process data and hierarchical clustering of drawing-process data.

To achieve the above objectives, we propose the following two approaches.

A1：Construction of a machine learning model using drawing-process data

We construct a machine learning model using drawing-process data as explanatory variables and scoring results as target variables. We perform learning and evaluation using various machine learning models and select the model that achieves high accuracy in screening for cognitive decline.

A2：Hierarchical clustering of drawing-process data

We perform hierarchical clustering of drawing-process data to extract features. Through hierarchical clustering, we aim to clarify drawing process patterns that cannot be determined from drawing results and consider early screening methods for cognitive decline.

The ultimate goal is to establish an early screening method based on these approaches, enabling early detection of dementia simply by conducting tests using EVIDENT.

B. A1: Construction of a machine learning model using drawing-process data

We construct a machine learning model for screening dementia using drawing-process data as explanatory variables and scoring results as target variables. Specifically, we proceed with the following steps.

First, we score the drawing-process data of CDT and CCT collected using EVIDENT, as described in Section II-C, based on specific scoring items. Based on the scoring results, we assign binary labels indicating the presence or absence of cognitive decline. Additionally, we preprocess the drawing-process data and convert it into a format suitable for machine learning models. Specifically, we extract and quantify features such as stroke count, speed, and pen pressure. Next, we use the quantified drawing-process data as explanatory variables and the labels assigned above as target variables to train various machine learning models. We then evaluate the performance of the constructed models and select the one that achieves high accuracy in screening for cognitive decline.

This approach is expected to enable automatic determination of cognitive decline using drawing-process data, thereby reducing the burden on healthcare workers.

C. A2: Hierarchical clustering of drawing-process data

We perform hierarchical clustering of drawing-process data to extract features that lead to early screening for dementia. Similar to Section III-B, we assign labels based on scoring results, quantify the drawing-process data, and perform hierarchical clustering. We extract clusters that include data with labels indicating cognitive decline and analyze the characteristics of the drawing-process data belonging to those clusters.

This approach is believed to clarify drawing process patterns that cannot be determined from drawing results, enabling the construction of methods for earlier detection of cognitive decline.

IV. Evaluation Experiments

To evaluate the effectiveness of the proposed methods described in Chapter III, we collected data using EVIDENT and conducted the following evaluation experiments.

A. Data Collection using EVIDENT

Data collection using EVIDENT has been ongoing since 2021, and as of August 2025, we have collected and accumulated 86 datasets. Below, we present the details of the tests and information about the collected data.

The drawing tests were conducted using an iPad and Apple Pencil. A paper-like film was applied to the iPad, and a

TABLE I
FEATURES IN DRAWING-PROCESS DATA

Item	Features
Speed	Minimum, Median, Maximum, Mean, Variance
Pen Pressure	Minimum, Median, Maximum, Mean, Variance
Time (seconds)	Total Drawing Time, Average Pen-Off Time, Total Pen-Off Time, Variance of Pen-Off Time, Ratio of Pen-Off Time to Total Drawing Time
Stroke Count	Stroke Count

TABLE II
AGE DISTRIBUTION OF PARTICIPANTS IN GROUP TESTS USING EVIDENT

Age Group	Number of Participants
30s	1
40s	0
50s	4
60s	15
70s	39
80s	20
90s	7

pencil-like cover was attached to the Apple Pencil to create an environment similar to that of drawing tests on paper. Hereafter, we refer to them as "screen" and "pen," respectively. Before conducting the tests, a tutorial was provided as a preliminary exercise, where participants were asked to write simple shapes and characters.

The features of the collected data are shown in Table I. Speed is calculated by dividing the distance drawn from one coordinate to another on the screen over a certain time interval by the time interval. Pen pressure is measured sequentially when the pen is in contact with the screen. The values are normalized within a range where the application can measure, with a minimum value of 0.0005 and a maximum value of 0.9995. Total drawing time refers to the time from when the pen touches the screen and starts drawing to when the drawing is completed and the pen is lifted off the screen for each drawing test. Pen-off time indicates the time when the pen is lifted off the screen between the end of one stroke and the start of the next stroke. These features were selected based on prior research [13] [14].

The age distribution of the participants is shown in Table II.

B. Evaluation Experiment for A1

1) Preparation: We scored 86 sets of CDT and CCT drawing data collected via EVIDENT. For CDT, we adapted Royall et al.'s 15-point CLOX method [15], modifying the target time to 10:10 and adding items such as the presence of tic marks. CCT was scored on a 10-point scale following Yorimitsu et al.'s framework [16], with cutoffs of 10 and 7 points, respectively. These thresholds served as analytical criteria rather than clinical diagnoses. Samples scoring above the cutoff were labeled as "normal," and those at or below were labeled as "cognitive decline," treated as the positive class. As a result, 22.1% of CDT and 33.7% of CCT samples were labeled positive.

From each session, we extracted quantitative process features—drawing speed, pen pressure, total drawing time, pen-off time, and stroke count (Table I). Identifiers and geometric size/position features (e.g., width, aspect ratio) were excluded to avoid bias. Missing values were imputed by the median, and zero-variance features were removed. Z-score normalization was applied to scale-sensitive models (Logistic Regression, SVM with RBF kernel, KNN), while tree-based models (RandomForest, GradientBoosting, LightGBM, CatBoost) used only imputation and variance filtering. All preprocessing steps were unified in each model's `Pipeline` for consistency.

We trained seven binary classifiers: Logistic Regression, SVM_RBF, KNN, RandomForest, GradientBoosting, LightGBM, and CatBoost. Class imbalance was mitigated using class-weighting where applicable. Model performance was evaluated with stratified 5-fold cross-validation at the subject level. In each fold, 80% of the subjects were used for training and 20% for validation while maintaining the class ratio. For each model, we report the mean $\pm$ standard deviation across folds for Accuracy, Precision, Recall, F1-score, ROC-AUC, PR-AUC (average precision), and the Brier score. ROC-AUC reflects the separability of the positive and negative classes, PR-AUC emphasizes performance under class imbalance, and the Brier score measures the calibration quality of predicted probabilities (lower is better). This evaluation setup enabled us to assess both the discriminative performance and probability calibration of each classifier.

2) Results: Table III summarizes the stratified 5-fold cross-validation performance. Because of space constraints, we omit the full CCT metrics table.

3) Discussion: In CDT, KNN and CatBoost showed high discriminative performance, while RandomForest provided the most reliable probability calibration. In CCT, tree-based ensemble models such as GradientBoosting and RandomForest consistently achieved strong results. Overall, cognitive decline could be classified with high accuracy using only drawing-process features.

There are drawing data that were not correctly classified, we analyzed the characteristics of the data included in FP (false positives) and FN (false negatives). Examples of data that were not correctly classified in CDT are shown in Fig. 2.

The respective scoring results were 12/15 points for CDT and 8/15 points for CCT. A similar analysis was conducted for CCT, and in both FP and FN cases, the scores were close to the cutoff value. From this result, it can be understood that data with scoring results near the cutoff value may not be correctly classified. In the future, we will explicitly handle uncertainty near the cutoff and aim to reduce misclassification from both operational and learning perspectives.

C. Evaluation Experiment for A2

1) Preparation: Similar to Section IV-B1, we used the 86 datasets of CDT and CCT drawing-process data to assign

TABLE III
STRATIFIED 5-FOLD CROSS-VALIDATION PERFORMANCE ON CDT (MEAN±STD ACROSS FOLDS)

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC	PR-AUC	Brier
CatBoost	0.851±0.029	0.890±0.028	0.925±0.054	0.906±0.022	0.875±0.068	0.959±0.031	0.132±0.026
GradientBoosting	0.839±0.062	0.868±0.060	0.941±0.033	0.902±0.037	0.868±0.086	0.949±0.044	0.147±0.061
KNN	0.863±0.050	0.878±0.031	0.956±0.040	0.915±0.032	0.889±0.097	0.952±0.052	0.098±0.035
LightGBM	0.839±0.050	0.910±0.026	0.880±0.072	0.893±0.040	0.862±0.071	0.949±0.034	0.125±0.047
LogisticRegression	0.816±0.029	0.925±0.048	0.837±0.063	0.876±0.020	0.855±0.075	0.943±0.048	0.126±0.016
RandomForest	0.840±0.044	0.867±0.040	0.941±0.033	0.902±0.027	0.914±0.058	0.974±0.022	0.100±0.021
SVM_RBF	0.792±0.081	0.890±0.040	0.835±0.103	0.859±0.064	0.824±0.090	0.943±0.034	0.124±0.027

TABLE IV
STRATIFIED 5-FOLD CROSS-VALIDATION PERFORMANCE ON CCT (MEAN±STD ACROSS FOLDS)

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC	PR-AUC	Brier
CatBoost	0.814±0.128	0.878±0.101	0.847±0.135	0.858±0.094	0.887±0.074	0.946±0.033	0.156±0.100
GradientBoosting	0.850±0.048	0.918±0.080	0.864±0.125	0.881±0.049	0.919±0.055	0.955±0.035	0.148±0.054
KNN	0.826±0.037	0.810±0.028	0.965±0.048	0.880±0.026	0.824±0.010	0.865±0.048	0.138±0.018
LightGBM	0.825±0.084	0.887±0.107	0.862±0.096	0.868±0.058	0.929±0.055	0.970±0.023	0.133±0.052
LogisticRegression	0.849±0.098	0.932±0.098	0.842±0.146	0.877±0.090	0.923±0.071	0.965±0.033	0.115±0.043
RandomForest	0.838±0.045	0.870±0.054	0.898±0.108	0.879±0.039	0.906±0.050	0.947±0.039	0.107±0.026
SVM_RBF	0.792±0.082	0.859±0.055	0.827±0.156	0.834±0.083	0.851±0.065	0.932±0.026	0.152±0.031

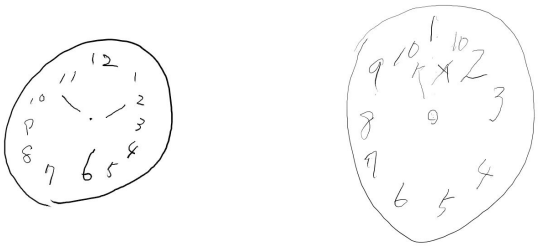

Fig. 2. Examples of misclassified cases in CDT binary classification model (Left: FP, Right: FN)

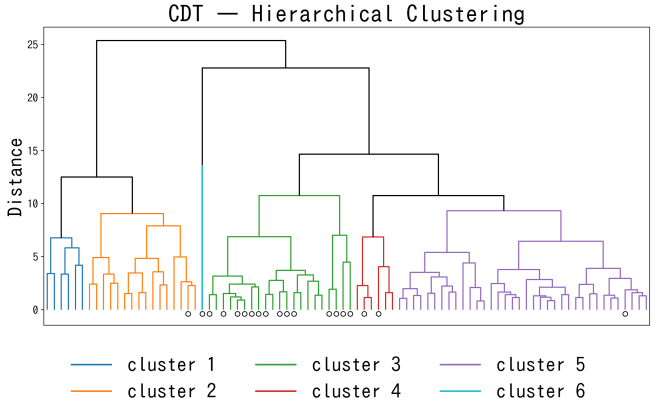

Fig. 3. Hierarchical clustering of CDT

labels indicating the presence or absence of cognitive decline based on scoring results. Additionally, we preprocessed the drawing-process data by quantifying and normalizing it. The features used were the same as those in Table I. Next, we performed hierarchical clustering based on the created features. From the clustering results, we determined the optimal number of clusters and extracted clusters that included data labeled as indicating the presence or absence of cognitive decline. Subsequently, we analyzed the characteristics of the drawing-process data belonging to each cluster and extracted features to capture signs of cognitive decline.

2) Results: The results of hierarchical clustering for CDT are shown in Fig. 3. The circular marks indicate data labeled as cognitive decline. A similar analysis was conducted for CCT, which is omitted here for brevity.

Based on these results, we set the number of clusters to 6 for CDT and 5 for CCT. To extract the characteristics of each cluster, we calculated the mean of the features for each cluster and created stacked bar graphs. This allows us to determine which clusters have higher or lower values for each feature. As an example, the results of creating this graph for speed-related features in CDT are shown in Figure 4. Similar analyses were conducted for other features and for CCT.

3) Discussion: From Fig. 3, it can be seen that data labeled as cognitive decline are biased towards clusters 3 and 6 in CDT. In CCT, it was found that the data were generally biased towards two of the five clusters. Therefore, we focused on these data and analyzed Fig. 4 and other stacked bar graphs, summarizing the characteristics of the clusters, as shown in Table IV.

From the tables, it can be seen that the focused clusters in both CDT and CCT have characteristics such as low speed and pen pressure, long drawing time and pen-off time, and high total stroke count. Drawing result data for CDT and CCT that belong to these characteristics but do not have a cognitive decline label are shown in Figure 5.

The scoring results for these drawing data are 15/15 points for CDT and 9/10 points for CCT. Since there are no signs

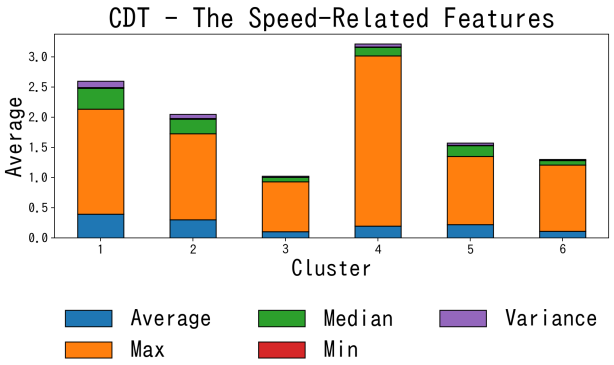

Fig. 4. Cluster means of speed-related features in CDT

TABLE V
TRENDS IN FOCUSED CLUSTERS

Item	Characteristics
Speed	Low mean, minimum, median, and variance High maximum value
Pen Pressure	Low mean, maximum, median, and variance
Time	Long total drawing time, total pen-off time, high variance in pen-off time, and high pen-off ratio High average pen-off time
Stroke Count	High total stroke count

of cognitive decline in the drawing results, it is considered that early screening may be possible using the drawing-process data for such subjects. However, since it has not been confirmed whether cognitive function is actually declining, it is necessary to examine in detail the correspondence with clinical cognitive function evaluations such as brain imaging tests.

In addition, there are several data points labeled as cognitive decline outside the clusters mentioned above, and it cannot be said that screening can be completely done using only the characteristics of these drawing processes. Therefore, by using it in conjunction with the machine learning models shown in A1, it is expected to build a more accurate screening method.

Specifically, first, temporal early screening is performed using the method in A1, and for subjects who are not judged to have cognitive decline by the method in A1, further screening is conducted using the method in A2. This approach is expected to capture more signs of cognitive decline.

V. CONCLUSION AND FUTURE WORK

In this study, we developed an early screening method based on the drawing-process data of CDT and CCT collected through EVIDENT. Two complementary approaches were examined—machine learning and hierarchical clustering—and their effectiveness was verified through experiments. In A1 (machine learning), binary classification using features such as size, speed, pen pressure, time, and stroke count achieved high accuracy across multiple models, with CatBoost showing consistently strong results. In A2 (hierarchical clustering), clusters with concentrated cognitive decline labels were identified,

Fig. 5. Examples of drawing data without cognitive decline labels (Left: CDT Cluster 3, Right: CCT Cluster 2)

and influential feature combinations were visualized. These findings suggest that the proposed methods can reduce the workload of healthcare professionals and support early detection of cognitive decline, including in large-scale screenings.

However, the current dataset (86 samples each for CDT and CCT) remains limited, and additional data collection is essential to improve screening accuracy. Future work includes conducting brain imaging and other clinical assessments to validate the medical relevance of the detected patterns. Although this study focused on data analysis, EVIDENT does not yet incorporate automated screening functions. We plan to expand group testing using EVIDENT and implement real-time screening capabilities within the application. This will ultimately enable an app that facilitates both temporal and clinical early detection of dementia.

ACKNOWLEDGMENT

This research was partially supported by JSPS KAKENHI Grant Numbers JP25H01167, JP25K02946, JP25K24389, JP24K02765, JP24K02774, JP23K17006, JP23K28091, JP23K28383.

REFERENCES

[1] A. Berit and D. Ove, "The clock-drawing test," *Age and Ageing*, vol. 27, no. 3, pp. 399–403, 1998.

[2] M. S, O. A, M. S, O. K, S. T, K. I, and S. E, "Clinical examination on reliability and validity of cube copying test (cct) scoring method," *Japanese Journal of Comprehensive Rehabilitation Science*, pp. 5–102, 2014.

[3] N. Masami, N. Masahiro, I. Eiji, O. Kanami, K. Masako, H. Akiko, M. Yasuyo, and W. Shinnichi, "MMSE and scoring of the clock drawing test increase the accuracy of diagnosis of dementia," *Igaku Kensa (Japanese Journal of Medical Technology)*, vol. 68, no. 3, pp. 424–429, 2019.

[4] S. Ryukichi, S. Sachio, N. Masahide, K. Naoki, and S. Atsushi, "Implementing visualization interface evident of drawing execution in neuropsychological tests," IEICE, Tech. Rep. 232 SC2020-31, 2020.

[5] ——, "Developing a platform for the digitalization of cognitive screening tests based on drawing process," IEICE, Tech. Rep. 416, 2022.

[6] Y. Keisuke, H. Shun, S. Sachio, S. Atsushi, K. Naoki, and N. Masahide, "Conducting a mass testing experiment using the drawing test app evident and exploratory data analysis," IEICE, Tech. Rep. 245 SC2024-40, 2024.

[7] J. N. Association, "2019 survey on hospital nursing," [Online]. Available: https://www.nurse.or.jp/up_pdf/20200330151534_f.pdf, 2019, accessed: Sep. 20, 2025.

[8] C. O. G. of Japan, "White paper on aging society 2024," [Online]. Available: https://www8.cao.go.jp/kourei/whitepaper/w-2024/html/zenbun/s1_2_2.html, 2024, accessed: Sep. 20, 2025.

[9] A. Takashi, "Mild cognitive impairment (MCI)," *Cognitive Neuroscience*, vol. 11, no. 3+4, pp. 252–257, 2009.
[10] J. S. of Neurology, "Clinical practice guideline for dementia 2017," [Online]. Available: https://www.neurology-jp.org/guidelinem/nintisyo_2017.html, 2017, accessed: Sep. 20, 2025.
[11] K. Tagawa, *Handbook of Neuropsychological Assessment*. Nishimura Co., Ltd., 2004.
[12] D. Catherine, A. Franchesca, A. Shawna, D. Randall, P. Dana, L. D. J, and P. C. C, "Cognitive correlates of digital clock drawing metrics in older adults with and without mild cognitive impairment," *Journal of Alzheimer's Disease*, vol. 75, no. 1, pp. 73–83, 2020.
[13] S.-M. William, P. Dana, G. Sebastian, H. Daniel, R. Roni, G. John, L. D. J, S. Randall, A. Rhoda, and D. Randall, "DCTclock: Clinically-interpretable and automated artificial intelligence analysis of drawing behavior for capturing cognition," *Frontiers in Digital Health*, vol. 3, 2021.
[14] K. Masahiro *et al.*, "Automated early detection of alzheimer's disease by capturing impairments in multiple cognitive domains with multiple drawing tasks," *Journal of Alzheimer's Disease*, vol. 88, no. 3, pp. 1075–1089, 2022.
[15] R. D. R, C. J. A, and P. Marsha, "Clox: An executive clock drawing task," *Journal of Neurology Neurosurgery and Psychiatry*, vol. 64, no. 5, pp. 588–594, 1998.
[16] Y. Miyuki, T. Kenshin, W. Yasuko, and Y. Ryoji, "Development of fixed quantitative scoring method of cube copying test—from the drawings by the inpatients of brain surgery in our hospital—," *Higher Brain Function Research*, vol. 33, no. 1, pp. 12–19, 2013.