Journal of Computer Science

Original Research Paper

# Prognosis of Dementia Using Early Fusion Approach with Digital Clock Drawing and Trail-Making Tests

**Shridevi Karande and Vrushali Kulkarni**

*School of Computer Engineering and Technology, Faculty of Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune, India*

*Article history*
Received: 27-10-2023
Revised: 08-04-2024
Accepted: 23-04-2024

Corresponding Author:
Shridevi Karande
School of Computer Engineering and Technology, Faculty of Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune, India
Email: shridevi.karande@mitwpu.edu.in

**Abstract:** Dementia poses a substantial global public health challenge. Emerging evidence suggests that COVID-19's neurological impact may aggravate dementia incidences. Timely recognition and management can significantly decelerate dementia progression and enhance affected individuals' quality of life. In the realm of cognitive assessment, the Clock Drawing Test (CDT) and Trail-Making Test (TMT) stand as prominent tools. Existing research predominantly focuses on the use of these tests in isolation. As CDT checks visuospatial skills and planning, TMT focuses on processing speed and mental flexibility. Combining these allows us to better understand an individual's cognitive strengths and weaknesses. This study aims to assess whether combining the features from the digital versions of these tests as dCDT and dTMT can enhance classification accuracy and recall in dementia cases. It utilizes an early fusion technique, merging feature metrics from both these tests for dementia classification. The study includes 86 healthy control participants and 52 individuals diagnosed with dementia. The early fusion method demonstrates promising outcomes as an alternative to conventional paper-based screening methods of CDT and TMT. The model attains a prominent overall accuracy of 93%, along with 87% precision, 85% recall and 0.94 AUC. The results exhibit reasonable improvements in classification performance as against prior studies involving individual modes of dCDT and dTMT. With the increase in the dataset size, this study can be extended for the classification of dementia sub-types. The scope of the study and data collection process is reviewed and approved by an independent ethics committee.

**Keywords:** Early Fusion, Digital Clock Drawing Test, Machine Learning, Alzheimer’s Disease, Dementia, Digital Trail Making Test

## Introduction

According to estimates, the global population of people aged 60 and above is projected to surpass 2.1 billion by 2050 (UN, 2022). In India, this demographic is expected to reach 340 million by 2050 (UN, 2022). This elderly population in India has been growing rapidly at 3.9% annually, raising concerns about potential health issues, including dementia. Worldwide, there are approximately 55 million people currently suffering from dementia, with about 10 million new cases reported each year. India had 4.1 million dementia cases in 2020 and this number is anticipated to increase to 14.8 million by 2050 (Lynch, 2020). Recently, a study presented (Premraj *et al*., 2022) revealed a compelling connection between the symptoms of dementia and the neurological symptoms observed in COVID-19 patients. This finding suggests that there will be a potential increase in new dementia cases following the COVID-19 pandemic. Hence, there is an urgent requirement for easily accessible and faster screening methods to address this situation effectively.

Cognitive impairment encompasses a decline in various cognitive functions, including executive function, gnosis, language, orientation, attention, praxis, memory, visual-spatial perception and social cognition. This decline occurs across different stages:

a) Preclinical stage: This stage shows no noticeable symptoms, but there are already underlying brain changes taking place
b) Mild Cognitive Impairment (MCI): In this stage, there's a cognitive decline beyond what's expected for an individual's age and education level, yet it doesn't significantly disrupt daily activities. The observed cognitive issues or changes in brain processing are atypical for the person's age or educational background

DOI: 10.3844/jcssp.2024.898.908

c) Dementia: Prolonged cognitive impairments progress to dementia, involving a decline in cognitive skills (like memory, speech and thinking), functional abilities (such as daily activities like dressing, eating and walking), mood and behaviour. Alzheimer's Disease (AD) is a progressive, degenerative disorder affecting neurons in the brain, leading to the loss of memory, thinking, language skills and behavioural changes. It's a common form of dementia. MCI often poses challenges and can go undiagnosed

In the initial stages of impairment detection, clinical evaluations typically involve neuropsychological tests (Galvin, 2018). Among the commonly used tests is the Mini-Mental State Examination (MMSE), which assesses various cognitive domains, including language, orientation, memory, attention, motor activity and visuospatial perception. However, MMSE may not effectively detect executive dysfunction (Cullen *et al*., 2007). To address this limitation, the Montreal Cognitive Assessment (MoCA) includes tests that specifically assess executive functions. Two such sub-tests are the Clock Drawing Test (CDT) and the Trail Making Test (TMT). These tests are considered more suitable for identifying Mild to Adverse MCI (Ciesielska *et al*., 2016). CDT and TMT are frequently used as standalone assessments. The primary objective of this study is to evaluate if a combined method can better classify dementia groups from healthy controls. In this research, the group of participants with dementia and related conditions including Alzheimer's disease is collectively referred to as the dementia group.

The related work focuses on the importance of cognitive assessment tools, such as CDT and TMT, in understanding and diagnosing cognitive disorders. This section is primarily divided into comprehending the concepts of these tests, assessing their usefulness in cognitive assessment and validating them as a tool. Additionally, it explores the application of machine learning techniques to enhance prediction using these tests.

We employed two concise neuropsychological screening tests (CDT, TMT) to assess executive abilities in patients with Alzheimer's and Dementia Group (ADAG) in comparison to healthy controls. The primary aim of our research was to investigate the nature and degree of executive function impairment in the ADAG.

The Clock Drawing Test (CDT) is a widely used, simple tool for assessing cognitive impairment in individuals with dementia. It involves drawing a clock face, placing the numbers correctly and setting the hands to a specified time. The CDT test assesses various cognitive abilities, including visuospatial ability, executive function and praxis. Unlike the Mini-Mental State Examination (MMSE), which requires some educational background and may miss executive dysfunction, the CDT test is literacy agnostic, making it suitable for those with limited literacy skills (Palsetia *et al*., 2018; Kim *et al*., 2018).

**Table 1:** CDT TMT cognitive abilities and brain area mapping

| Assessment test | Cognitive abilities | Brain areas mapping |
|---|---|---|
| CDT Khachiyants and Kim (2012) | Planning sequencing organization | Frontal |
| | Comprehension | Temporal |
| | Ability to execute | Left parietal |
| | Construction ability spatial relation | Right parietal |
| | Visual processing | Occipital |
| TMT A Galvin (2018) | Psychomotor speed | Right-sided lesion frontal parietal |
| TMT B Galvin (2018) | Executive function | Left-sided lesion frontal parietal |

Table 1 Provides a comprehensive overview of the cognitive functions covered by both CDT and TMT. This amalgamation reveals an expanded scope of cognitive domains, emphasizing the synergistic effect of these assessments.

Notably, the CDT evaluates visuospatial abilities and executive functions, whereas the TMT delves into attention, cognitive flexibility and visual-motor coordination. The fusion of these two evaluations produces a more comprehensive and intricate portrayal of an individual's cognitive profile, thereby enriching the depth and breadth of cognitive assessment within the research context.

The functional Magnetic Resonance Imaging (fMRI) in the MCI group has revealed that specific errors in the CDT test correlate with reduced brain connectivity in certain regions (Eknoyan *et al*., 2012). Several studies have explored the sensitivity and specificity of CDT as a tool for detecting cognitive impairment. While it has shown effectiveness in screening moderate dementia cases, it may not be equally sensitive to early stages of dementia (Pinto and Peters, 2009; Ehreke *et al*., 2010).

The CDT test can be performed in two ways: (A) The Unprompted free drawing method, where participants draw the entire clock and (b) The pre-drawn method, which has two sub-variants. The pre-drawn method provides a clock face with numbers and the participants only need to draw the hands set to a specific time in one sub-variant, while in the other, they draw both the numbers and hands for a specific time (Agrell and Dehlin, 2012).

The unprompted free drawing method showed a strong correlation with both MMSE (Royall *et al*., 1998). Additionally, the free-drawn CDT presented a greater cognitive challenge and was more sensitive to detecting mild or early cognitive impairment than the incomplete-copy version, where participants copied a clock face with numbers and set the hands for a specific time

(Beber *et al*., 2016). This study consisted of 20 participants from the control group, 30 from the MCI group and 40 from the dementia group (Alzheimer's disease-AD: 20; Vascular Dementia-VD: 20).

Studies have suggested that CDT alone may be less sensitive for screening the MCI group (Powlishta *et al*. 2002) but combining it with the MMSE as the mini-clock method appears to improve sensitivity and specificity compared to using MMSE or CDT alone when it is carried out on 153 participants (Mittal *et al*., 2010; Cacho *et al*., 2010). When combined with other tests such as MMSE or Verbal Fluency Test (VFT), CDT worked well in identifying Alzheimer's Disease (AD) (Ladeira *et al*., 2009). However, its sensitivity for identifying MCI from normal controls was found to be low, even though its specificity was good.

Some studies have delved into utilizing sensor technology, such as a digital pen, to digitize the CDT test (Souillard-Mandar *et al*., 2016; Amini *et al*., 2021). These studies measured different parameters, including the number of strokes, total ink length and the time taken to draw each component. By applying various feature selection approaches to the metrics collected through the dCDT, an accuracy of 83% was achieved in classifying non-MCI, amnestic MCI (aMCI) and AD. Machine learning analysis of the digital clock drawing test approach achieved the best 2-group classification results with 10-fold cross-validation (Binaco *et al*., 2020). A newly developed digital test Geras Solutions Cognitive Test (GCST) is validated and compared with MOCA and it is found that both evaluate similar cognitive domains (Bloniecki *et al*., 2021). This study included a descriptive statistic performed considering 106 patients (SCI n = 65 MCI n = 24 dementia n = 9) and the results showed 0.91 and 0.55 in sensitivity and specificity respectively with an accuracy of 0.85. A study using visual inspection resulted in a sensitivity of 85% and a specificity of 0.75 for AD detection (Khonthapagdee *et al*., 2020).

Overall, the clock drawing test remains a valuable tool for assessing cognitive impairment, especially in the context of dementia and its combination with other tests can enhance its diagnostic capabilities.

The Trail Making Test (TMT) is another widely used neuropsychological assessment for detecting dementia and Alzheimer's disease, offering a sensitive measure of cognitive impairment. This test evaluates crucial cognitive abilities, including visual attention, executive function and mental flexibility. It comprises two parts: TMT-A and TMT-B, which assess different cognitive skills. In TMT-A, participants connect 25 circles in numerical order as quickly as possible. In contrast, TMT-B requires them to switch between connecting circles with letters and numbers in a specific order. TMT has proven valuable in assessing various cognitive and psychological impairments, such as Alzheimer's disease, schizophrenia, depression and traumatic brain injury. Research indicates that TMT performance is associated with brain regions linked to attention and executive function, like the parietal and frontal lobes (Varjacic *et al*., 2018). Several studies suggest that TMT performance is sensitive to aging (Salthouse, 2011) with TMT-B being particularly effective in detecting cognitive decline compared to TMT-A (Dahmen *et al*., 2017; Onoda *et al*., 2013; Ye *et al*., 2022). The study conducted by Linari *et al*. (2022) found that the detection performance for cognitive impairment using manual scoring of the TMT was 0.80 (Linari *et al*., 2022).

While TMT-A performance decreases with age, it is not affected by education levels (Tombaugh, 2004). Researchers have thoroughly examined the digital version of the test to explore correlations between predicted TMT scores, clinical digital test scores and traditional paper-based time-to-completion scores (Dahmen *et al*., 2017).

The TMT is a valuable screening tool, but it should be used alongside other assessments and a thorough medical history for a definitive diagnosis of dementia (Chan *et al*., 2015).

## Materials and Methods

Figure 1 illustrates the study's typical flow, starting with the identification of participant cohorts for healthy control and dementia groups. Digital variants of CDT and TMT tests were then conducted, generating the necessary digital feature set. The features obtained from various modes are combined into a unified feature vector, which is then employed to train a classifier.

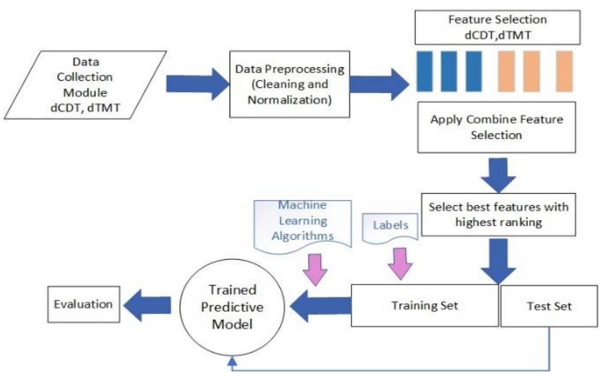


**Fig. 1:** Flow of the research

**Table 2:** Participants demographics

| Metrics | | Healthy controls (86) | AD and dementia (52) |
|---|---|---|---|
| Age | Mean | 74.98 | 72.81 |
| | Std deviation | 06.35 | 06.25 |
| Gender | Female | 46 | 32 |
| | Male | 40 | 20 |
| Language | English | 51 | 18 |
| | Hindi | 26 | 25 |
| | Marathi | 9 | 9 |

DOI: 10.3844/jcssp.2024.898.908

### *Participants*

The study involved a total of 138 participants, comprising 52 patients in the dementia group (32 females, 20 males) and 86 healthy controls (46 females, 40 males). Detailed participant demographics are presented in Table 2. The assessments of dementia patients took place at two Jagruti rehabilitation centres in Pune, specifically in the In-Patient Department (IPD). A healthy control sample was selected by conducting an MMSE screening test, where a score > = 24 is considered healthy. Before commencing the live data collection from patients, we sought approval from an independent ethics committee to address any ethical concerns. The committee was provided with details regarding the data collection process and a sample e-consent Form used for study participation was submitted for their review.

### *Data Collection Modules*

To support this study, we developed two digital versions of the assessment modules, namely dCDT and dTMT, as a cross-platform Android and IoS App. The application is optimized for both finger touch and stylus-pen interactions, allowing for convenient data collection using either method.

Recent findings (Rosselli *et al*., 2022) have shed light on the significant impact of a subject's preferred language, culture, ethnicity and country of origin on cognitive test performance. employing a generalized model may inadvertently introduce biases that skew the outcomes of the assessments. Hence, to meet Indian linguistic requirements, data collection modules (dCDT and dTMT) were localized to support Hindi and Marathi languages, in addition to English.

The dCDT module utilizes touch events to capture essential drawing metrics, such as path length, the number of paths, drawing speed, paths drawn within the clock's circle and spacing between independent paths. These metrics help assess visuospatial abilities. To determine a path length, the SkiaSharp graphics library is employed to measure the pixels traversed during the touch-down action (Mukherjee, 2019). Figure 2 depicts the flow of dCDT feature extraction.

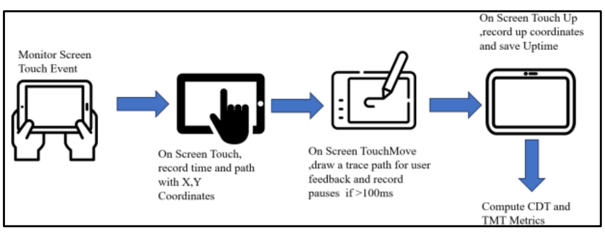


**Fig. 2:** Flow diagram for dCDT Feature extraction

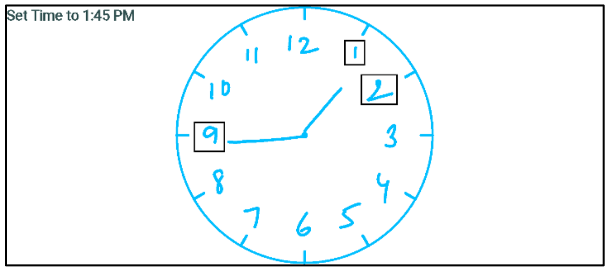


**Fig. 3:** Illustration of CDT test with identified bounding rectangles

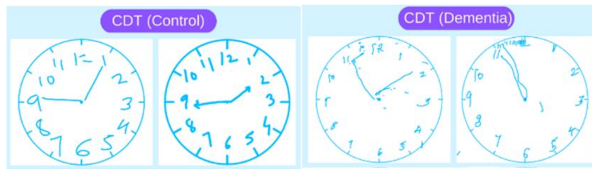


**Fig. 4:** CDT Test samples for; (a) Class 0; (b) Class 1

To identify the clock's hands, straight paths drawn from the circle's center to the clock's circumference are used.

The length of these paths distinguishes between minute and hour hands and their positions are estimated based on the outer end point's coordinates relative to the bounding rectangle of the numbers 1 and 2, as illustrated in Fig. 3. The bounding rectangles of the numbers are identified using the AWS extract (OCR) service. These techniques collectively enable the dCDT module to comprehensively analyze drawing-related data and visuospatial capabilities.

In this scenario, "lift duration" refers to the time between two path-drawing actions, indicating how much time the user spends contemplating their next move. This might be related to the user's memory recall process, as they mentally visualize the clock and its positions. On the other hand, "drawing speed" reflects the user's motor abilities. After each test completion, screen images are captured to facilitate CDT scoring. These CDT images are used to validate the CDT Score which is calculated using the method (Shua-Haim *et al*., 1996). The primary metrics captured at the end of the CDT test are utilized to compute secondary metrics such as Average durations and drawing speeds. Refer to Fig. 3 for the flow of dCDT feature extraction.

Figure 4a-b shows the CDT test samples of the control group and the dementia group respectively.

The dTMT module, similar to the dCDT module, calculates pause durations and lift durations based on touch-up, touch-down and touch-move events. It also verifies if the drawn path intersects with any of the bubbles during the touch move event. When a bubble is touched, it is recorded as a hit and later compared with an expected bubble order list to determine whether it is an error or a correct hit. To prevent missing circle connections erroneously, a tolerance buffer of 5-point coordinates is applied around each circle. Figure 5 depicts the flow of dTMT feature extraction.

DOI: 10.3844/jcssp.2024.898.908

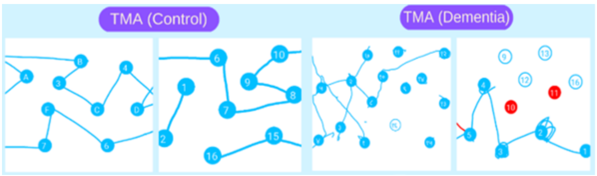


**Fig. 5:** TMT-A test samples for; (a) Class 0; (b) Class 1

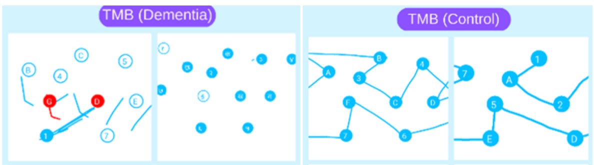


**Fig. 6:** TMT-B test samples for; (a) Class 1; (b) Class 0

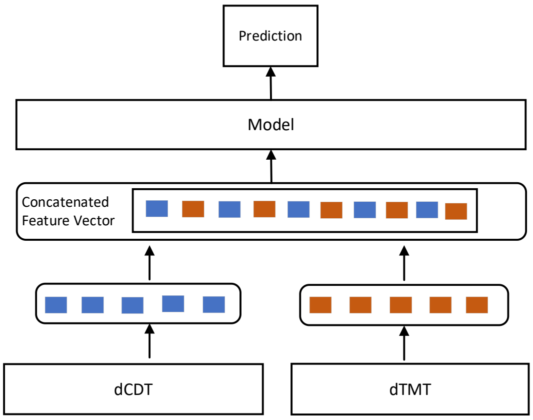


**Fig. 7:** Early fusion method for dCDT and dTMT metrics

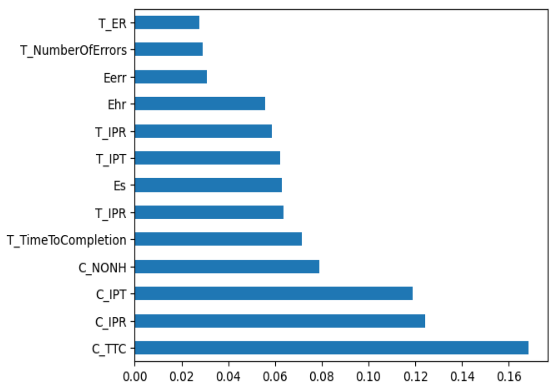


**Fig. 8:** Top-ranked features from the extra tree classifier selection method

Since dTMT is designed for self-administration, when a user makes an erroneous connection, the application visually indicates the error by changing the color of the incorrect circle and the connecting path from blue to red. This immediate feedback informs the user about the mistake committed, as depicted in Fig. 6a-b.

Trail-making tests are traditionally conducted using pen and pencil. The digital version of this test provides additional insights into abilities that may not be easily tracked by humans during the paper-based test. Trail-making tests, despite their apparent simplicity, are considered to be reflective of a wide array of cognitive functions. These functions encompass attention, visual scanning, order shifting, psychomotor speed, mental flexibility, adaptability in altering plans of action and the ability to maintain two separate streams of thought.

Modifications were made to adapt the TMT-part A and B tests for tablet screen size. In part A, the test involved 16 circles, numbered from 1-16. Part B consisted of 14 circles, labeled with numbers 1-7 and letters A-G (Fig. 6a-b). TMT -B snapshots Fig. 7a-b.

The dCDT and dTMT modules produce initial digital features based on the participant's test performance. From these primary features, secondary or derived features are generated. These features are classified into three types: Information processing, execution ability and visuospatial ability (specifically, timing features, motion/motor features and spatial features). Apart from the time to completion metric obtained in the traditional paper approach, the digital tests also capture additional information, including lifts, pauses, path length drawn and the number of correct and incorrect bubble connections (hits). Duration greater than 100 ms is counted as a pause. Refer glossary section for features definition and formula.

### *Early Fusion Method*

The early fusion method is commonly applied to raw or pre-processed data obtained from various sources or modalities. To ensure a seamless fusion process, it's crucial to extract data features beforehand, especially if the data sources have varying sampling rates across the modalities. Neglecting this step can significantly complicate the fusion process. This approach, also known as input-level fusion or data-level fusion, assumes that multiple data sources are independent of each other conditionally.

In the early fusion method, feature vectors are randomly concatenated to form a comprehensive list of feature sets, as illustrated in Fig. 8. The integration of various tests can significantly broaden the feature set accessible to machine learning algorithms, thereby equipping them with the capability to detect intricate patterns that might otherwise remain concealed when relying on a single test. After fusing the input data features, a feature selection step follows, where the master list of features serves as the input.

Statistical tests can be employed to identify the attributes that exhibit the most robust correlation with the output variable. The scikit-learn package offers the select K-best class, which may be utilized with various statistical tests to choose a certain number of features. To obtain the feature importance of each feature in your dataset, you can utilize the feature importance attribute of the model. Feature importance provides a numerical score for each feature in our dataset, indicating the degree of importance or relevance of the item to the output variable. The feature significance is a built-in class that is included in tree-based classifiers. In this case, we will utilize the extra tree classifier to extract the top 10 features from the dataset. The result of the extra tree classifier with the top 10 features is Fig. 8.

The alternative feature selection approaches, such as ANOVA and information gain, yielded similar results to the additional tree classifier selection method. These features are namely Total Information Processing Time (T_IPT), Executive Speed (ES) and Executive hit rate (Ehr) from the TMT test, along with features such as Information Processing Rate (C_IPR), Information Processing Time (C_IPT) and Time to Completion (C_TTC) from the CDT test, hold significant importance in classifying the dementia group. This finding underscores the significance of logical thinking, memory recall and motor skills as crucial indicators for dementia group classification.

Additionally, noteworthy indicators that surfaced during the analysis were time-to-test completion from dCDT and executive hit rate from dTMT (part-A) emerging as key indicators for their respective tests.

In this study, the collected data consisted of labelled information for two classes: Healthy control (class 0) and dementia (class 1). Given this classification scenario, the adoption of linear classifier algorithms was deemed appropriate for model construction. Considering the data distribution, we choose linear classifiers. The classifiers considered for this study are namely Logistic Regression (LR), Gaussian Naïve Bayes (GNB) and Support Vector Machine (SVM). To ensure optimal performance, the top-ranking features obtained from the feature selection step were used as inputs for these classifiers.

Classification analysis involved 138 participants, comprising 86 healthy controls and 52 from the dementia group. Utilizing Python and scikit-learn (Abraham *et al.*, 2014), we built LR, GNB and SVM, classifier models.

## Results and Discussion

The objective of this study was to explore the potential of building a classifier by combining features from multiple neuropsychological assessment tests, specifically CDT and TMT, through an early fusion approach. When these tests are used independently, they only address specific aspects of cognitive domains, such as visuospatial abilities in the case of dCDT and executive and psychomotor speed abilities in the case of dTMT.

Figure 9, the Average Time between Letters (ATBL) and Average Time between Numbers (ATBN) displays a right-skewed distribution. Class 1 participants tend to take more time to progress to subsequent letters or numbers compared to class 0. Additionally, these participants took more time to reach letters than numbers. Within class 0, there are 3-4 outliers (indicated by a cross) located on the class boundary lines. Class 1 exhibits lower values due to some participants withdrawing from the tests, citing their complexity as the reason for withdrawal.

Table 3 presents a comparison of three classifiers LR, GNB and SVM on a binary classification problem, considering accuracy, recall and precision as evaluation metrics. While Accuracy measures the proportion of correctly classified instances by a model, recall (sensitivity or true positive rate), provides a quantitative evaluation of the ratio of true positive instances as against total positive instances within the dataset. Finally, precision determines the proportion of true positive instances relative to the total number of instances classified as positive by the model.

The classifiers were assessed individually and using the combined feature list for both CDT and TMT features. Based on the evaluation metric, the top-performing classifier varied. For accuracy, the fusion method of LR and SVM achieved the highest score above 0.85, while GNB obtained the next highest recall score (0.83) after LR. Regarding precision, LR's fusion method outperformed others with a score of 0.93, with SVM's fusion method at 0.85. Interestingly, the GNB model exhibited relatively low precision scores across all three modes, suggesting a greater susceptibility to false positives.

Figure 10 displays a clustered bar chart presenting the LR classifier's performance across various test modes. In the chart, the single mode refers to the LR classifier trained on individual feature sets of CDT or TMT, while the fusion method denotes the LR classifier trained on a combined feature set of CDT and TMT. Notably, the LR classifier trained in the fusion method, utilizing combined features, achieved the highest performance concerning accuracy (0.93), recall (0.85) and precision (0.87). In comparison, the LR classifier trained on the Single mode of CDT attained the second-best performance in the analysis. A small number of cases from class 1 in the dataset were initially classified as class 0 when utilizing a single TMT mode. However, when the results from the combined mode were taken into account, these cases were appropriately classified as class 1. Utilizing a combination approach significantly enhanced the sensitivity and specificity of the system.

**Table 3:** Performance comparison of different classifiers

| Classifier | Accuracy CDT mode | Accuracy TMT mode | Accuracy Fusion method | Recall CDT Mode | Recall TMT mode | Recall Fusion method | Precision CDT mode | Precision TMT mode | Precision Fusion method |
|---|---|---|---|---|---|---|---|---|---|
| LR | 0.85 | 0.71 | 0.93 | 0.69 | 0.67 | 0.85 | 0.8 | 0.56 | 0.87 |
| GNB | 0.88 | 0.73 | 0.71 | 0.67 | 0.33 | 0.83 | 0.89 | 0.71 | 0.05 |
| SVM | 0.85 | 0.73 | 0.85 | 0.58 | 0.67 | 0.58 | 0.88 | 0.59 | 0.88 |

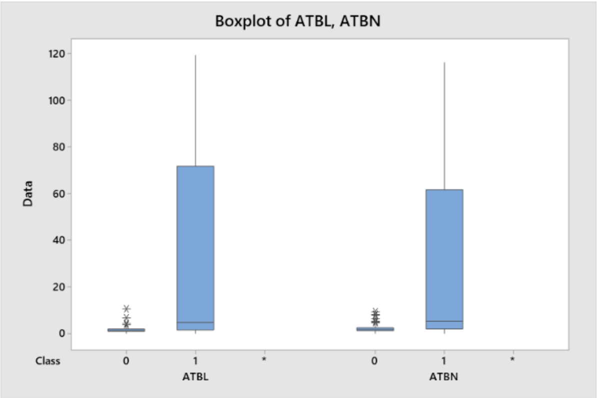


**Fig. 9:** TMT-B boxplots for ATBL, ATBN

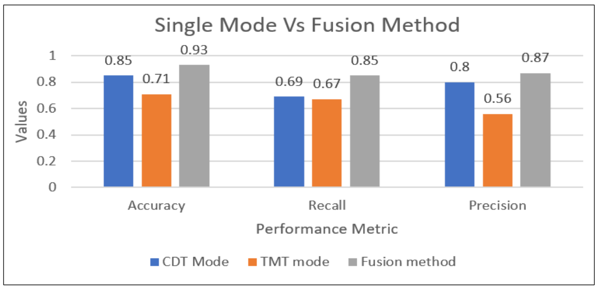


**Fig. 10:** Performance comparison of LR for single mode and fusion method

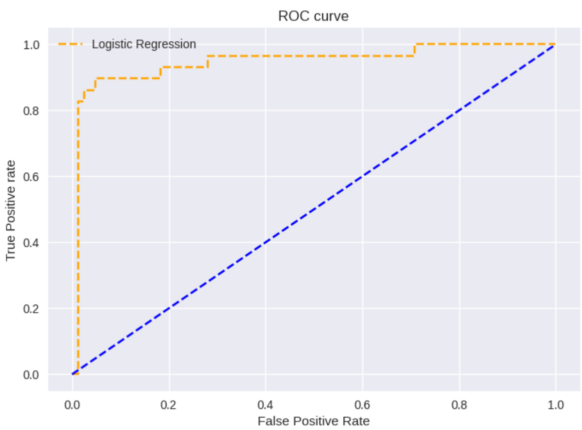


**Fig. 11:** ROC curve

In machine learning, AUC (area under the receiver operating characteristic curve) is a widely used metric to evaluate the performance of a binary classification model. AUC ranges between 0 and 1, where a higher AUC value indicates better model performance in separating the positive and negative samples. Figure 11 depicts the AUC of LR classifiers with an AUC score of 0.94. This means the model is better at ranking predictions, with a lower false positive rate and a higher true positive rate. A higher Area Under the Curve (AUC) value is typically considered more favorable, although the exact interpretation of a score is dependent upon the specific context and the associated cost of misclassification. In our specific scenario, when the consequences of a false negative (failing to detect a disease) are quite serious it may be desirable to select a higher threshold for identifying individuals as positive, even if this results in a little decrease in the AUC score.

Prior studies by Binaco *et al*. (2020); Dahmen *et al*., (2017) though effective focused on employing digital tests like dCDT and dTMT in isolation, leading to a limited evaluation of cognitive domains. In contrast, the present study employing a fusion method demonstrates an enhanced model accuracy of 93% as against earlier investigations (Binaco *et al*., 2020), which achieved a classifier accuracy of 85% for binary classification of MCI and non-MCI groups. Furthermore, the fusion method exhibited a notable improvement with an AUC score of 0.94 as compared to an AUC score of 0.65 in an earlier study (Dahmen *et al*., 2017) utilizing digital features of the TMT and an AUC score of 81.3 reported in another study (Amini *et al*., 2021) using digital features from CDT. The class 1 group, as defined in this study, encompasses a spectrum of dementia cases, including those classified as MCI.

The overall results indicate that the proposed early fusion method enables the classification of the dementia group using logistic regression with an accuracy of 93%, precision of 87% and recall of 85%. The study highlights the potential of machine learning through the early fusion method in enhancing the accuracy and objective interpretation of multi-domain cognitive assessments for dementia and related conditions. The limitation of this study is the relatively small sample size and focusing on a specific population nevertheless, the relatively small sample size calls for further studies with larger samples encompassing different cities in India to validate the generalizability of the findings.

DOI: 10.3844/jcssp.2024.898.908

## Conclusion

The study emphasizes the importance of employing the early fusion method with the digital Clock Drawing Test (dCDT) and Trail Making Test (dTMT) to gain valuable insights into cognitive function efficiently. Utilizing digital versions enhances accessibility and ensures cost-effective, accurate data collection, making it a highly advantageous tool for initial screening.

The combined assessment can provide clinicians with a more detailed cognitive profile, facilitating informed clinical decision-making. Clinicians can tailor interventions based on the specific cognitive domains affected. Also, a cognitive profile derived from the combined assessment can aid in designing personalized treatment plans for individuals with cognitive impairments.

Consequently, continued research efforts, involving a larger and more diverse cohort, are vital to strengthen the credibility and generalizability of the fusion method's outcomes. Through such extensive validation, the fusion approach using digital CDT and TMT assessments could become an indispensable tool in the field of cognitive health assessment ultimately leading to improved healthcare strategies and interventions tailored to individual cognitive needs.

## Acknowledgment

The authors would like to express special gratitude to the Centre head and support staff of Jagruti Rehabilitation Centre, Pune for facilitating the conduction of this study at their Centre. Their cooperation and support are highly appreciated.

## Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Author's Contributions

**Shridevi Karande:** Designed and acquired data, analyzed and interpretation of data and drafted the article.

**Vrushali Kulkarni:** Reviewed the article critically for significant intellectual content and gave final approval of the version to be submitted.

## Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and that no ethical issues are involved.

This study on human subjects was approved (approval no. RPIEC091221) by the Royal Pune independent ethics committee (DCGI REG NO: ECR/45/Indt/MH/2013/RR-19) and the opt-out consent process was granted.

## References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*, 00014. https://doi.org/10.3389/fninf.2014.00014

Agrell, B., & Dehlin, O. (2012). The clock-drawing test. *1998. Age and Ageing, 41 Suppl 3, iii41–iii45*. https://doi.org/10.1093/ageing/afs149

Amini, S., Zhang, L., Hao, B., Gupta, A., Song, M., Karjadi, C., Lin, H., Kolachalama, V. B., Au, R., & Paschalidis, I. C. (2021). An Artificial Intelligence-Assisted Method for Dementia Detection Using Images from the Clock Drawing Test. *Journal of Alzheimer's Disease*, *83*(2), 581-589. https://doi.org/10.3233/jad-210299

Beber, B. C., Kochhann, R., Matias, B., & Chaves, M. L. F. (2016). The Clock Drawing Test: Performance differences between the free-drawn and incomplete-copy versions in patients with MCI and dementia. *Dementia and Neuropsychologia*, *10*(3), 227-231. https://doi.org/10.1590/S1980-5764-2016DN1003009

Binaco, R., Calzaretto, N., Epifano, J., McGuire, S., Umer, M., Emrani, S., Wasserman, V., Libon, D. J., & Polikar, R. (2020). Machine Learning Analysis of Digital Clock Drawing Test Performance for Differential Classification of Mild Cognitive Impairment Subtypes Versus Alzheimer's Disease. *Journal of the International Neuropsychological Society, 26*(7), 690-700. https://doi.org/10.1017/s1355617720000144

Bloniecki, V., Hagman, G., Ryden, M., & Kivipelto, M. (2021). Digital Screening for Cognitive Impairment-A Proof of Concept Study. *The Journal of Prevention of Alzheimer's Disease*, *8*, 127-134. https://doi.org/10.14283/jpad.2021.2

Chan, E., MacPherson, S. E., Robinson, G., Turner, M., Lecce, F., Shallice, T., & Cipolotti, L. (2015). Limitations of the Trail Making Test Part-B in Assessing Frontal Executive Dysfunction. *Journal of the International Neuropsychological Society*, *21*(2), 169-174. https://doi.org/10.1017/s135561771500003x

Ciesielska, N., Sokołowski, R., Mazur, E., Podhorecka, M., Polak-Szabela, A., & Kędziora-Kornatowska, K. (2016). Is the Montreal Cognitive Assessment (MoCA) test better suited than the Mini-Mental State Examination (MMSE) in Mild Cognitive Impairment (MCI) detection among people aged over 60? Meta-analysis. *Psychiatria Polska*, *50*(5), 1039-1052. https://doi.org/10.12740/pp/45368

Cullen, B., O'Neill, B., Evans, J. J., Coen, R. F., & Lawlor, B. A. (2007). A review of screening tests for cognitive impairment. *Journal of Neurology, Neurosurgery and Psychiatry*, *78*(8), 790-799. https://doi.org/10.1136/jnnp.2006.095414

Dahmen, J., Cook, D., Fellows, R., & Schmitter-Edgecombe, M. (2017). An analysis of a digital variant of the Trail Making Test using machine learning techniques. *Technology and Health Care*, *25*(2), 251-264. https://doi.org/10.3233/thc-161274

Royall, D. R., Cordes, J. A., & Polk, M. (1998). CLOX: An executive clock drawing task. *Journal of Neurology, Neurosurgery and Psychiatry*, *64*(5), 588-594. https://doi.org/10.1136/jnnp.64.5.588

Ehreke, L., Luppa, M., König, H.-H., & Riedel-Heller, S. G. (2010). Is the Clock Drawing Test a screening tool for the diagnosis of mild cognitive impairment? A systematic review. *International Psychogeriatrics*, *22*(1), 56-63. https://doi.org/10.1017/s1041610209990676

Eknoyan, D., Hurley, R. A., & Taber, K. H. (2012). The Clock Drawing Task: Common Errors and Functional Neuroanatomy. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *24*(3), 260-265. https://doi.org/10.1176/appi.neuropsych.12070180

Galvin, J. E. (2018). Using Informant and Performance Screening Methods to Detect Mild Cognitive Impairment and Dementia. *Current Geriatrics Reports*, *7*(1), 19-25. https://doi.org/10.1007/s13670-018-0236-2

Linari, I., Juantorena, G. E., Ibáñez, A., Petroni, A., & Kamienkowski, J. E. (2022). Unveiling Trail Making Test: Visual and manual trajectories indexing multiple executive processes. *Scientific Reports*, *12*(1), 14265. https://doi.org/10.1038/s41598-022-16431-9

Cacho, J., Benito-León, J., García-García, R., Fernández-Calvo, B., Vicente-Villardón, J. L., & Mitchell, A. J. (2010). Does the Combination of the MMSE and Clock Drawing Test (Mini-Clock) Improve the Detection of Mild Alzheimer's Disease and Mild Cognitive Impairment?. *Journal of Alzheimer's Disease*, *22*(3), 889-896. https://doi.org/10.3233/jad-2010-101182

Khachiyants, N., & Y. Kim, K. (2012). Mini-mental status exam mapping to the corresponding brain areas in dementia. *Applied Technologies and Innovations*, *7*(2), 55-58. https://doi.org/10.15208/ati.2012.7

Kim, S., Jahng, S., Yu, K.-H., Lee, B.-C., & Kang, Y. (2018). Usefulness of the Clock Drawing Test as a Cognitive Screening Instrument for Mild Cognitive Impairment and Mild Dementia: An Evaluation Using Three Scoring Systems. *Dementia and Neurocognitive Disorders*, *17*(3), 100. https://doi.org/10.12779/dnd.2018.17.3.100

Ladeira, R. B., Diniz, B. S., Nunes, P. V., & Forlenza, O. V. (2009). Combining Cognitive Screening Tests for the Evaluation of Mild Cognitive Impairment in the Elderly. *Clinics*, *64*(10), 967-973. https://doi.org/10.1590/s1807-59322009001000006

Lynch, C. (2020). World Alzheimer Report 2019: Attitudes to dementia, a global survey: Public health: Engaging people in ADRD research. Alzheimer's & Dementia, 16, e038255. https://doi.org/10.1002/alz.038255

Mittal, C., Gorthi, S. P., & Rohatgi, S. (2010). Early Cognitive Impairment: Role of Clock Drawing Test. *Medical Journal Armed Forces India*, *66*(1), 25-28. https://doi.org/10.1016/s0377-1237(10)80087-5

Mukherjee, S. (2019). *Learn Microsoft Visual Studio App Center* (1st Ed.,). Apress Berkeley, CA. https://doi.org/10.1007/978-1-4842-4382-4

Onoda, H., Nabika., A., Takayoshi, N., Ishihara Mitaki, S., Yamaguchi, O., Shiwaku, & Yamaguchi. (2013). Validation of a new mass screening tool for cognitive impairment: Cognitive Assessment for Dementia, iPad version. *Clinical Interventions in Aging*, *2013*, 353-360. https://doi.org/10.2147/cia.s42342

Palsetia, D., Rao, G. P., Tiwari, S. C., Lodha, P., & De Sousa, A. (2018). The Clock Drawing Test versus Mini-mental Status Examination as a Screening Tool for Dementia: A Clinical Comparison. *Indian Journal of Psychological Medicine*, *40*(1), 1-10. https://doi.org/10.4103/ijpsym.ijpsym_244_17

Pinto, E., & Peters, R. (2009). Literature Review of the Clock Drawing Test as a Tool for Cognitive Screening. *Dementia and Geriatric Cognitive Disorders*, *27*(3), 201-213. https://doi.org/10.1159/000203344

Powlishta, K. K., Von Dras, D. D., Stanford, A., Carr, D. B., Tsering, C., Miller, J. P., & Morris, J. C. (2002). The clock drawing test is a poor screen for very mild dementia. *Neurology*, *59*(6), 898-903. https://doi.org/10.1212/WNL.59.6.898

Premraj, L., Kannapadi, N. V., Briggs, J., Seal, S. M., Battaglini, D., Fanning, J., ... & Cho, S. M. (2022). Mid and long-term neurological and neuropsychiatric manifestations of post-COVID-19 syndrome: A meta-analysis. *Journal of the Neurological Sciences*, *434*, 120162. https://doi.org/10.1016/j.jns.2022.120162

Rosselli, M., Uribe, I. V., Ahne, E., & Shihadeh, L. (2022). Culture, Ethnicity and Level of Education in Alzheimer's Disease. *Neurotherapeutics*, *19*(1), 26-54. https://doi.org/10.1007/s13311-022-01193-z

Khonthapagdee, S., Lownoppakul, S., Chotchoey, N., Wiwatwattana, N., Mongkolluksamee, S., & Sa-ing, V. (2020). Alzheimer Screening using Drawing Test Scores. *2020 17th International Conference on Electrical Engineering Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 101-104. https://doi.org/10.1109/ecti-con49241.2020.9158244

Salthouse, T. A. (2011). What cognitive abilities are involved in trail-making performance? *Intelligence*, *39*(4), 222-232. https://doi.org/10.1016/j.intell.2011.03.001

Shua-Haim, J., MD, G. K., & MD, J. G. (1996). A Simple Scoring System for Clock Drawing in Patients with Alzheimer's Disease. *Journal of the American Geriatrics Society*, *44*(3), 335-335. https://doi.org/10.1111/j.1532-5415.1996.tb00931.x

Souillard-Mandar, W., Davis, R., Rudin, C., Au, R., Libon, D. J., Swenson, R., Price, C. C., Lamar, M., & Penney, D. L. (2016). Learning classification models of cognitive conditions from subtle behaviors in the digital Clock Drawing Test. *Machine Learning*, *102*(3), 393-441. https://doi.org/10.1007/s10994-015-5529-5

Tombaugh, T. (2004). Trail Making Test A and B: Normative data stratified by age and education. *Archives of Clinical Neuropsychology*, *19*(2), 203-214. https://doi.org/10.1016/s0887-6177(03)00039-8

Varjacic, A., Mantini, D., Demeyere, N., & Gillebert, C. R. (2018). Neural signatures of Trail Making Test performance: Evidence from lesion-mapping and neuroimaging studies. *Neuropsychologia*, *115*, 78-87. https://doi.org/10.1016/j.neuropsychologia.2018.03.031

UN. (2022). *Department of Economic and Social Affairs Population Division*. United Nations. https://population.un.org/wpp/

Ye, S., Sun, K., Huynh, D., Phi. H. Q., Ko, B., Huang, B., & Ghomi, R. H. (2022). A Computerized Cognitive Test Battery for Detection of Dementia and Mild Cognitive Impairment: Instrument Validation Study. *JMIR Aging*, *5*(2), e36825. https://doi.org/10.2196/36825

**Glossary**

| Sr. No. | Metric | Definition | Formula |
|---|---|---|---|
| 1 | Total pause Duration (TPD) | Duration greater than 100 ms is counted as a pause and its time is added to the total pause duration | $TPD = \sum_{i=0}^{n} PD(i)$ |
| 2 | Total lift Duration (TLD) | It is the time interval for which the user is not touching the screen (i.e., lifting the stylus) and its time is added to the total lift duration | $TPD = \sum_{i=1}^{m} PD(i)$ |
| 3 | Number of Lifts (NOL) | The number of times the user lifts his/her finger after the test has started, is measured as the number of lifts | NOL = Count of number of lifts |
| 4 | Time To Completion (TTC) | Total time taken for completion of the test. C_TTC denote time to completion for CDT | TTC = End time-start time |
| 5 | Hit Count/No. of Hits (NOH) | This is the number of bubbles connected by the user during TMT-A & B tests. Based on the expected order of bubble connections, the tests categorize the hit as correct or incorrect hit | NOH = No. of correct and incorrect hits |
| 6 | Number of Errors (NOE) | This is the number of hits which were categorized as incorrect/error hit. In the case of TMT-A & B tests, incorrect order of bubbles is counted as an error hit. | NOE = No. of incorrect hits |
| 7 | Executive Speed (Es) | This is the speed at which the user draws a path on the screen and this metric is applicable to both | Es = Path length/time taken draw the path |

**DOI: 10.3844/jcssp.2024.898.908**

**Glossary:** Continue

| | | | |
|---|---|---|---|
| | | dCDT and dTMT tests | |
| 8 | Visuospatial Ability (VA) | This ability is determined only in the case of the CDT test | $f(x) = \vee(\forall y(Csym)y) \wedge (\forall z(Chs(z), Cmhp(z), Chhp(z)))$ Where: Y = Presence of numbers from1-12 z = Presence of minute and hour hands |
| 9 | Csym | Symmetry followed while writing numbers in the clock contour | observed manually from the drawn clock. It will have values as True or False |
| 10 | Chs | Correct hand size | Minute and hour hand sizes. It will have values from 0-2. 0- incorrect min and hour hand sizes.1-either min or hour hand size is correct 2-both min and hour hand sizes are correct |
| 11 | Cmhp | Correct minute hand position | It will have values as True or false |
| 12 | Chhp | Correct hour hand position | It will have values as True or false |
| 13 | Executive Ability (EA) | This ability indicates the participant's overall executive i.e., decision-making skills and gross Psycho motors skills | $EA = fn(Es, Ehr, Eer)$ Where: |
| 14 | Executive hit rate (Ehr) and Executive error rate (Eer) | Ehr-no. of correct hits Eer-no. of correct hits | $Ehr = \frac{Es}{NOH}$ $Eer = \frac{NOE}{Total\ Hits}$ |
| 15 | Information Processing Time (IPT) | This refers to the processing time required to convert information into action while drawing. It serves as an indicator of the time spent thinking or recalling information from memory. It encompasses the total duration during which a person is not actively drawing. C_IPT denote Information Processing Time in case of CDT. T_IPT denote information processing time in case of TMT | $IPT = TPD + TLD$ |
| 16 | Information Processing Rate (IPR) | This is the processing time for the total number of hits. In the case of the CDT test, the number of hits is the total numbers drawn inside the circle and in TMT, it is the number of bubbles connected including errors and the number of hits. C_IPR denote information processing rate in case of CDT | $IPR = \frac{NOH}{IPT}$ |