Original Paper

AI-Assisted Analysis of Spontaneous Affective Facial Responses for Cognitive Status Assessment in Older Adults: Exploratory Feasibility Study

Jung-Cheng Yang[1,2], PT, MSc; Ting-Ju Lai[3], MPH; Bak-Sau Yip[4], MD, PhD; Wei-Hui Wu[4], ClinPsy; Wun-Hung Chen[1], PT; Hsien-Tsung Chang[5,6,7], PhD; Pay-Shin Lin[2,8], DrPH, PT

[1]Division of Physical Therapy, Department of Physical Medicine and Rehabilitation, Hsin-Chu Branch, National Taiwan University Hospital, Hsinchu, Taiwan

[2]Master Degree Program in Health & Long-term Care Industry, College of Medicine, Chang Gung University, Taoyuan, Taiwan

[3]Department of Medical Research, Hsin-Chu Branch, National Taiwan University Hospital, Hsinchu, Taiwan

[4]Department of Neurology, Hsin-Chu Branch, National Taiwan University Hospital, Hsinchu, Taiwan

[5]Department of Artificial Intelligence, College of Intelligent Computing, Chang Gung University, Taoyuan, Taiwan

[6]Department of Computer Science and Information Engineering, College of Engineering, Chang Gung University, Taoyuan, Taiwan

[7]Center for Artificial Intelligence in Medicine, Chang Gung Memorial Hospital, Taoyuan, Taiwan

[8]School of Physical Therapy & Graduate Institute of Rehabilitation Science, College of Medicine, Chang Gung University, Taoyuan, Taiwan

Corresponding Author:
Pay-Shin Lin, DrPH, PT
Master Degree Program in Health & Long-term Care Industry
College of Medicine
Chang Gung University
No. 259, Wenhua 1st Rd., Guishan Dist.
Taoyuan, 33302
Taiwan
Phone: 886 3 2118800 ext 3393
Email: pslin@mail.cgu.edu.tw

Abstract

Background: As Taiwan enters a super-aged society, the prevalence of dementia continues to rise, highlighting the need for scalable, noninvasive approaches to identify behavioral alterations associated with cognitive decline. AI-based analysis of spontaneous affective facial responses has emerged as a potential digital phenotyping tool, but its feasibility for assessing cognitive performance remains underexplored.

Objective: This study aimed to examine the feasibility of using AI-derived spontaneous affective facial response metrics, including affective expression congruency and onset latency, to characterize differences across cognitive performance levels in older adults with healthy cognition (HC), mild cognitive impairment (MCI), or dementia.

Methods: A total of 60 participants (HC, MCI, and dementia; 20 per group) were presented with positive and negative emotion-eliciting videos and completed Mini-Mental State Examination (MMSE) recall and calculation tasks. Facial expressions were video-recorded and analyzed using a convolutional neural network–based AI system. AI-derived metrics included affective expression congruency, categorized as congruent, neutral, or incongruent according to the detected facial response, and affective expression onset latency, defined as the time from stimulus presentation to the emergence of a nonneutral facial expression detected by the AI model. Group differences were analyzed using 1-way ANOVA, Pearson chi-square or Fisher-Freeman-Halton exact tests, Bonferroni-adjusted pairwise post hoc comparisons, analysis of covariance adjusting for age and education, and exploratory receiver operating characteristic analyses.

Results: Affectively congruent responses were generally less frequent among participants with lower cognitive performance. During the 14-second positive stimulus, congruent positive expressions were observed in 65% (13/20) of HC participants, 20% (4/20) of MCI participants, and 0% (0/20) of participants with dementia. During the 48-second negative stimulus, congruent negative expressions were observed in 60% (12/20) of HC participants compared with 5% (1/20) of both MCI and dementia participants. Onset latency differed significantly across groups for the 14-second positive stimulus ($F_{2,57}$=19.63; P<.001; η^2=0.41),

XSL•FO
RenderX

52-second positive stimulus ($F_{2,57}$=4.25; P=.02; η^2=0.13), 48-second negative stimulus ($F_{2,57}$=15.61; P<.001; η^2=0.35), and MMSE-based cognitive-task condition ($F_{2,57}$=11.28; P<.001; η^2=0.28). After adjustment for age and education, the group effect remained significant for all latency outcomes ($F_{2,55}$=3.89-10.68; all P≤.03; partial η^2=0.12-0.28). Exploratory receiver operating characteristic analyses showed stronger separation in several HC vs MCI and HC vs dementia comparisons, whereas MCI vs dementia comparisons showed limited separability.

Conclusions: AI-assisted analysis of spontaneous affective facial responses revealed differences in expression congruency and onset latency across cognitive performance levels. These preliminary findings suggest that AI-derived facial behavioral metrics may serve as noninvasive adjunctive features for observational cognitive assessment. However, the findings should be interpreted cautiously because of the exploratory design, modest sample size, and absence of internal or external validation.

(JMIR Form Res 2026;10:e94422) doi: 10.2196/94422

KEYWORDS

artificial intelligence; dementia; digital phenotyping; facial expression; mild cognitive impairment

Introduction

Taiwan has entered a super-aged society, with household registration statistics from Taiwan's Ministry of the Interior showing that adults aged 65 years or older accounted for 20.06% of the population at the end of December 2025 [1]. Dementia prevalence in Taiwan has correspondingly become a major public health concern, with a recent nationwide survey estimating that approximately 7.99% of adults aged 65 years or older are affected, and prevalence rising substantially with advancing age [2]. Given that brain changes associated with Alzheimer disease may begin a decade or more before memory and other cognitive symptoms appear [3], early identification of cognitive decline represents a major public health priority.

Conventional screening approaches primarily rely on neuropsychological questionnaires or clinician-administered cognitive tests, which may be limited by administration time, language dependency, and accessibility in community-based settings. In the preclinical stage, individuals may present with subjective cognitive decline, often accompanied by depression, anxiety, or memory complaints, despite performing within normal ranges on standardized cognitive assessments [4]. Such limitations highlight the need for alternative, noninvasive behavioral indicators that may reflect early neurodegenerative change before overt cognitive impairment becomes clinically detectable. Recent reviews of AI-enabled digital phenotyping have further emphasized the potential of multimodal behavioral signals to complement conventional clinical assessment in Alzheimer disease and related cognitive disorders [5].

Beyond cognitive deficits, neurodegenerative conditions such as mild cognitive impairment (MCI) and dementia are increasingly recognized to involve alterations in emotional processing, social cognition, and affective responsiveness. Previous studies have demonstrated impairments in the perceptual recognition of emotional facial expressions in Alzheimer disease populations that may emerge even during prodromal stages such as MCI [6-8]. Similarly, reductions in emotional responsiveness have been associated with subtle global cognitive decline in older adults [9]. However, these paradigms primarily assess explicit perceptual recognition of emotional stimuli through identification tasks, rather than spontaneous affective expression.

Spontaneous affective facial responses represent observable behavioral manifestations of underlying emotional and cognitive processing that may be altered in neurodegenerative conditions. Recent advances in computer vision and AI have enabled automated quantification of such facial behavioral signals in clinical populations. Recent scoping review evidence has also highlighted the emerging role of AI-based technologies in detecting behavioral and psychological symptoms of dementia, supporting the potential value of automated behavioral assessment approaches in dementia-related care and research [10]. Automated facial expressivity analysis has been used to quantify facial expressivity in older adults with neurocognitive disorders and to examine its association with apathy symptoms [11]. AI-based facial expression analysis has been explored in individuals with Parkinson disease, major depressive disorder, and cognitive impairment to characterize alterations in emotional expressivity or affective flattening [12]. More recently, AI-assisted facial emotion analysis has been applied to examine emotion-related motor expression patterns in individuals with cognitive impairment during emotionally evocative tasks [13].

However, existing studies have primarily focused on expression intensity or recognition-related features rather than the temporal dynamics of spontaneous affective facial responses during emotionally evocative or cognitively demanding tasks. The timing and emergence of affective facial responses may reflect behavioral manifestations of emotion-cognition interaction that are not captured by conventional neuropsychological assessments [14,15].

Accordingly, this exploratory study aimed to examine the feasibility of using AI-derived measures of spontaneous affective facial responses—specifically, expression congruency and onset latency—to characterize differences across cognitive performance levels, operationally defined as healthy cognition (HC), MCI, and dementia, in older adults. By investigating temporal characteristics of spontaneous affective responses during emotionally evocative and cognitively demanding tasks, this study sought to explore the potential application of automated facial behavioral metrics as noninvasive adjunct features for observational cognitive assessment in clinical and community-based aging populations.

XSL•FO
RenderX

Methods

Study Setting and Participants

This study was conducted at a medical center in Taiwan. Participants were recruited and study data were collected between August 2024 and April 2025. Participants were recruited from 2 sources. Participants in the HC group were recruited through institutional review board–approved printed study flyers posted on hospital bulletin boards and through word-of-mouth referral by friends and colleagues. Community-dwelling individuals who reported no prior clinical diagnosis of dementia were invited to undergo Mini-Mental State Examination (MMSE) screening to determine eligibility for the HC group. Participants in the MCI and dementia groups were recruited from neurology outpatient clinics after routine clinical evaluation conducted by board-certified neurologists. For these clinically referred participants, classification and referral were determined independently of the AI-based behavioral analysis and prior to study enrollment. Clinical classification was based on neurological assessment, Clinical Dementia Rating (CDR) staging, and review of computed tomography and/or magnetic resonance imaging findings performed as part of standard clinical care. For the purpose of behavioral comparison in this exploratory feasibility study, participants were stratified into cognitive performance levels operationally defined as HC, MCI, and dementia. A total of 60 participants were enrolled in the study and categorized into 3 cognitive performance strata, with 20 participants in each group. Because this was an exploratory feasibility study, the sample size was determined pragmatically based on recruitment feasibility, with a target of 20 participants per cognitive performance group.

Inclusion and Exclusion Criteria

Cognitive grouping in this study was operationally defined using screening instruments and clinical staging measures, including the MMSE and CDR, for the purpose of behavioral group-level comparison in exploratory analysis, rather than to establish new clinical diagnoses.

Participants were eligible for inclusion if they were aged between 50 and 75 years, were alert and able to follow instructions during testing procedures, and were able to complete the audiovisual stimulus procedures. For the HC group, participants were required to have an MMSE score of 26 or higher, consistent with commonly used screening thresholds for normal cognitive performance in older adults [16]. For the MCI group, participants were required to have a CDR score of 0.5, representing MCI for the purpose of this study [17]. For the dementia group, participants were required to have a CDR score of 1.0 or higher and to have been evaluated by neurologists [17].

Participants were excluded if they had a diagnosis of Parkinson disease, a history of stroke with associated facial paralysis, an inability to complete the assessments due to severe cognitive impairment, altered consciousness, or language deficits, or use of psychiatric medications that could substantially affect facial expression.

Study Procedure

The study protocol consisted of cognitive assessment, demographic data collection, and facial expression testing under emotionally evocative and cognitively demanding conditions.

Global cognitive performance was assessed using the MMSE administered by a licensed clinical psychologist. Screening-based cognitive performance strata were subsequently used for group-level behavioral comparison in the analysis of spontaneous affective facial responses. Participants also completed a demographic and health questionnaire including gender, age, education level, marital status, BMI, exercise habits, social engagement, and history of chronic medical conditions.

Participants were then seated in front of a computer screen in a controlled testing environment. Prior to emotional stimulus presentation, a 5-second neutral baseline recording was obtained while participants were instructed to maintain a relaxed facial expression. This baseline segment was used to establish an individualized reference for subsequent AI-based detection of nonneutral affective facial responses during task conditions. Before stimulus presentation, participants' ability to perceive the audiovisual stimuli was checked in the testing setting. When participants or family members reported presbyopia, reduced visual acuity, or hearing difficulty, participants were asked to wear their usual corrective glasses if available, verbal instructions were provided at an increased volume, and the video playback volume was adjusted as needed to support adequate perception of the stimuli during testing.

Emotional video stimuli were selected through a pilot screening procedure conducted prior to study implementation. Candidate clips were evaluated by adult volunteers across a wide age range for perceived emotional category (positive or negative) and engagement. Based on pilot feedback and consultation with a clinical psychologist, 2 positive emotional video stimuli of differing durations, approximately 14 seconds and 52 seconds, were selected to examine whether stimulus length influenced the emergence of observable affective facial responses, particularly in cognitively impaired individuals who may demonstrate delayed emotional engagement.

Following baseline recording, participants were presented with the 2 positive emotional video clips in sequence. A cognitively demanding task condition was subsequently introduced between positive and negative emotional stimuli. Specifically, participants completed MMSE subtests involving short-term memory recall and calculation ability, which served as a structured transitional condition intended to reduce potential affective carryover between emotionally evocative tasks. Finally, a 48-second negative emotional video stimulus depicting an affectively salient real-life narrative was presented to elicit sadness-related responses. After the negative stimulus, a brief positive video clip was presented as a mood-recovery procedure to help participants return to a more positive emotional state. Data from this recovery clip were not included in the primary analysis.

During both the emotional stimulation and cognitive task conditions, synchronized video recording was performed to capture facial expressions for subsequent AI-based analysis of

XSL•FO
RenderX

affective expression onset latency and expression congruency. Facial expressions recorded during MMSE subtests were not used to evaluate cognitive task performance but rather to quantify spontaneous affective expressivity under cognitive load conditions.

Ethical Considerations

This study was approved by the Institutional Review Board of National Taiwan University Hospital, Hsinchu Branch (approval number 202405102RINA). Written informed consent was obtained from all participants or their legally authorized representatives in accordance with the Declaration of Helsinki. When consent was provided by a legally authorized representative, participant assent was sought whenever possible before study procedures. Participants who declined participation or showed unwillingness, discomfort, or distress during the study procedures were not enrolled or were withdrawn from testing. All study data were deidentified before analysis. Video recordings and AI-derived data were stored on a secure hospital server with access restricted to authorized research personnel. No participant compensation was provided.

Facial Expression Capture and Processing

Overview

Facial expression data were acquired at a standard frame rate of 30 frames per second, enabling the capture of subtle facial movements for temporal affective expression analysis. The AI-based analysis pipeline consisted of automated facial landmark detection and deep learning–based facial affective expression classification. All video data were processed using the same predefined pipeline, and AI-derived outputs were used for group-level behavioral analysis rather than clinical diagnosis.

Preprocessing and Feature Extraction

Facial landmark detection was performed using the MediaPipe Face Mesh module developed by Google, which extracts 468 3D facial landmarks. For compatibility with conventional facial landmark–based alignment approaches, 68 landmarks corresponding to the widely used dlib 68-point facial landmark model were retained for facial alignment and pose estimation [18].

To correct potential head movements during recording, 3D pose estimates were calculated based on key landmarks (eg, nose tip, chin, eye corners, and mouth corners) and converted to Euler angles (pitch, yaw, and roll). If any angle exceeded a predefined threshold, affine transformations were applied to realign the face to a frontal position.

The aligned facial images were then converted to grayscale, resized to 48×48 pixels, normalized to a 0-1 scale, and transformed into tensor format for convolutional neural network (CNN) input.

The CNN-based model classified facial expressions into 1 of 7 affective categories (angry, disgusted, fearful, happy, neutral, sad, or surprised), with probabilities calculated via a softmax layer. The category with the highest probability was assigned as the dominant affective expression for each frame. If the classification confidence was below 0.6, the frame was labeled as "uncertain" to reduce potential misclassification of ambiguous expressions. Frames labeled as uncertain were excluded from the nonneutral frame count used for affective responses onset detection.

For the purpose of behavioral analysis, the 7 predicted affective categories were subsequently grouped into 3 emotional expression classes: positive, negative, and neutral. This grouping was used to account for the low frequency of certain affective categories in the sample and to improve the interpretability of group-level analyses. In the primary analysis, happy and surprised expressions were classified as positive affective responses; angry, disgusted, fearful, and sad expressions were classified as negative affective responses; and neutral expressions were retained as a separate category. Because surprise is considered a valence-ambiguous expression whose interpretation depends on contextual appraisal [19,20], the classification of surprised expressions as positive affective responses was explicitly defined as context-specific and based on the positive video stimuli, which were designed to evoke amusement, enjoyment, and pleasant surprise. This classification was not intended to imply that surprise is universally positive.

Establishing a Neutral Baseline

To objectively determine the onset of emotional expressions, an individualized neutral baseline was established for each participant. During the initial 5-second neutral gaze period, CNN-derived feature vectors from all frames were averaged to generate a participant-specific neutral reference vector.

Subsequent affective responses were determined based on feature-level deviations from this individualized neutral reference rather than absolute emotion classification, allowing detection of subject-specific changes in facial expressivity relative to their baseline state.

Affective Response Detection Algorithm

During the emotional video viewing and cognitive-task conditions, CNN-based feature vectors and classification outputs were extracted for each frame. A sliding window of the most recent 30 frames was used to evaluate expression consistency.

If more than 50% of the frames within the window were classified as nonneutral, and the L1 distance from the neutral reference vector exceeded 0.8 for more than 2 consecutive seconds, equivalent to 60 frames, this period was identified as a significant affective facial response. Onset latency was defined as the time from stimulus or task onset to the beginning of the first 2-second period that met the detection criteria.

This baseline-referenced detection approach was designed to minimize interindividual variability in resting facial expressivity and improve sensitivity to task-related affective changes.

Data Export

The system exported timestamped affective response durations and corresponding expression categories for each stimulus condition in CSV format. These AI-derived metrics were used to quantify spontaneous emotional expressivity and response latency for between-group comparisons.

XSL•FO
RenderX

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics version 25. Demographic and clinical characteristics of the 3 cognitive performance groups (HC, MCI, and dementia) were summarized as mean (SD) for continuous variables and as frequencies and percentages (n, %) for categorical variables. Between-group differences in demographic and clinical characteristics were examined using 1-way ANOVA for continuous variables and Pearson chi-square tests or Fisher-Freeman-Halton exact tests, as appropriate, for categorical variables.

AI-derived affective expression onset latency was analyzed as a continuous outcome. Group differences in onset latency across cognitive performance groups were examined using 1-way ANOVA, followed by Bonferroni-adjusted post hoc pairwise comparisons. The significance threshold for pairwise comparisons was set at $P<.0167$. Because age and education level differed across cognitive performance groups, additional analysis of covariance (ANCOVA) models were performed for each latency outcome, with cognitive performance group entered as the fixed factor and age and education level entered as covariates. Education level was coded as college education or higher vs below college education. Partial η^2 was reported as an estimate of effect size for ANCOVA models.

Using the 3 affective response classes defined above, positive affective responses were considered congruent for positive video conditions. For the negative video condition and the MMSE-based cognitive-task condition, negative affective responses were considered congruent. In the MMSE-based cognitive-task condition, negative affective facial expressions were operationally classified as task-congruent responses because such expressions may reflect task-related cognitive effort, frustration, uncertainty, or mental strain during recall and calculation demands. This operational definition was informed by prior studies suggesting that facial behavioral signals and facial action units may be associated with cognitive load, task difficulty, listening effort, confusion, and frustration during cognitively demanding or difficult task conditions [21,22]. This definition was not intended to imply that negative expressions exclusively or directly represent cognitive effort.

AI-derived affective expression congruency was analyzed as a categorical outcome by comparing response distributions across cognitive performance groups using Pearson chi-square tests or Fisher-Freeman-Halton exact tests, as appropriate. The MMSE-based cognitive-task condition included short-term memory recall and serial subtraction. Frames classified as nonneutral expressions were further used for latency estimation to quantify spontaneous affective expressivity under cognitive load.

In trials where no nonneutral affective facial expression was detected during a given video stimulus condition, affective onset latency was assigned as n+1 seconds, where n represents the duration of the corresponding video stimulus. This coding approach was used to indicate that no affective response was detected within the observation window and that the true onset latency, if present, exceeded the stimulus duration. Because n+1 represents the smallest possible extension beyond the observation window, this approach was intended to preserve the ordinal relationship between responders and nonresponders while minimizing artificial variance inflation compared with assigning larger arbitrary latency values. For the MMSE-based cognitive-task condition, participants with no AI-detected affective response throughout the task were assigned a maximum latency value of 90 seconds, corresponding to the predefined maximum observation window for the cognitive-task condition. These observations were conceptually treated as right-censored responses for descriptive and group-level latency analyses.

Because this was an exploratory feasibility study with a modest sample size, receiver operating characteristic (ROC) curve analyses were conducted as hypothesis-generating analyses to examine the preliminary discriminative ability of AI-derived affective expression onset latency for distinguishing between cognitive performance groups across emotional video and MMSE-based cognitive-task conditions. No internal cross-validation procedure, such as leave-one-out cross-validation or k-fold cross-validation, was performed. Therefore, the area under the curve (AUC) values and Youden index–derived cutoff values should be interpreted as preliminary estimates derived from the present sample rather than validated classification thresholds.

All statistical tests were 2-tailed. A P value of <.05 was considered statistically significant, except for Bonferroni-corrected pairwise post hoc comparisons, as described above.

Results

Participant Characteristics

A total of 60 participants were enrolled in this exploratory feasibility study, with 20 individuals in each of 3 cognitive performance strata: HC, MCI, and dementia. There were no significant differences in gender distribution among the 3 groups ($\chi^2_2=0.94$; n=60; $P=.63$; Cramér $V=0.13$). Age differed significantly across cognitive performance groups ($F_{2,57}=13.62$; $P<.001$; $\eta^2=0.32$). Post hoc comparisons indicated that participants in the HC group were significantly younger than those in the MCI and dementia groups, whereas no significant difference in age was observed between the MCI and dementia groups.

BMI also differed significantly across groups ($F_{2,57}=4.77$; $P=.01$; $\eta^2=0.14$), as did the proportion of participants with a college education or higher ($\chi^2_2=13.65$; n=60; $P=.001$; Cramér $V=0.48$). Social participation differed significantly across cognitive performance groups ($\chi^2_2=10.10$; n=60; $P=.006$; Cramér $V=0.41$).

Significant between-group differences were also observed for several exercise-related characteristics, including regular exercise, exercise frequency greater than 150 minutes per week, exercise intensity, and walking as an exercise type (Table 1). No significant between-group differences were observed for current smoking, alcohol consumption, insomnia, or retired status.

XSL•FO
RenderX

Table 1. Demographic characteristics by cognitive performance group.a

Variable	HC[b] (n=20)	MCI[c] (n=20)	Dementia (n=20)	Test statistic (*df*)	*P* value	Effect size
Female, n (%)	12 (60)	9 (45)	11 (55)	0.94 (2)[d]	.63	0.13
Age (years), mean (SD)	61.5 (7.71)	69.1 (4.99)	70.0 (4.89)	13.62 (2,57)[e]	<.001	0.32
BMI, mean (SD)	25.07 (4.06)	25.16 (3.40)	22.21 (2.71)	4.77 (2,57)[e]	.01	0.14
College education or higher, n (%)	12 (60)	7 (35)	1 (5)	13.65 (2)[d]	.001	0.48
Current smoking, n (%)	3 (15)	2 (10)	1 (5)	1.11 (2)[d]	.86[f]	0.14
Alcohol consumption, n (%)	3 (15)	4 (20)	3 (15)	0.24 (2)[d]	>.99[f]	0.06
Insomnia, n/N (%)	7/20 (35)	11/20 (55)	6/19 (31.6)	2.62 (2)[d]	.27	0.21
Retired status, n (%)	13 (65)	15 (75)	19 (95)	5.50 (2)[d]	.06[f]	0.30
Social participation, n (%)	16 (80)	11 (55)	6 (30)	10.10 (2)[d]	.006	0.41
Regular exercise, n (%)	14 (70)	20 (100)	13 (65)	8.45 (2)[d]	.008[f]	0.38
Exercise frequency >150 min/week, n/N (%)	11/14 (78.6)	7/20 (35.0)	1/13 (7.7)	14.49 (2)[d]	.001	0.56
Exercise intensity, Borg RPE[g] 10-13, n/N (%)	12/14 (85.7)	8/20 (40.0)	5/13 (38.5)	8.48 (2)[d]	.01	0.43
Exercise type: walking, n/N (%)	8/14 (57.1)	19/20 (95.0)	12/13 (92.3)	9.46 (2)[d]	.02[f]	0.45

[a]Between-group differences in continuous variables were examined using 1-way ANOVA. Categorical variables were analyzed using Pearson chi-square tests or Fisher-Freeman-Halton exact tests, as appropriate. Effect sizes are reported as η^2 for 1-way ANOVA and Cramér *V* for categorical comparisons. For categorical comparisons with sparse expected cell counts, statistical significance was determined using Fisher-Freeman-Halton exact *P* values, while Pearson chi-square statistics and Cramér V are reported as the corresponding test statistic and effect size. Exercise frequency, exercise intensity, and exercise type were calculated among participants who reported regular exercise (healthy cognition: n=14; mild cognitive impairment: n=20; dementia: n=13). Insomnia data were available for 19 participants in the dementia group.

[b]HC: healthy cognition.

[c]MCI: mild cognitive impairment.

[d]Chi-square test (*df*).

[e]*F* test (*df*).

[f]Fisher-Freeman-Halton exact *P* value was used because expected cell counts were sparse.

[g]RPE: rating of perceived exertion.

AI-Derived Affective Expression Congruency Across Cognitive Performance Strata

AI-derived affective expression distributions differed significantly across cognitive performance strata under both emotional video and cognitive-task conditions. Across groups, affectively congruent facial expressions were generally less frequent and neutral responses more frequent among participants with lower cognitive performance.

During the first positive video stimulus (14 seconds), affectively congruent positive expressions were observed in 65% (13/20) of participants in the HC group, compared with 20% (4/20) in the MCI group and 0% (0/20) in the dementia group (χ^2_4=28.48; exact *P*<.001; Cramér *V*=0.49). During the second positive video stimulus (52 seconds), affectively congruent positive expressions were observed in 75% (15/20) of HC participants, 45% (9/20) of MCI participants, and 20% (4/20) of participants with dementia (χ^2_4=15.53; exact *P*=.001; Cramér *V*=0.36).

In response to the negative emotional stimulus (48 seconds), 60% (12/20) of participants in the HC group displayed affectively congruent negative facial expressions, whereas only 5% (1/20) of both the MCI and dementia groups did so (χ^2_4=26.29; exact *P*<.001; Cramér *V*=0.47). A similar pattern was observed during the MMSE-based cognitive-task condition, in which affectively congruent negative facial expressions, operationally defined as task-congruent responses, were observed in 50% (10/20) of HC participants, compared with 15% (3/20) of MCI participants and 10% (2/20) of participants with dementia (χ^2_4=15.35; *P*=.004; Cramér *V*=0.36).

These findings indicate that AI-derived patterns of affective facial responses varied across cognitive performance levels, with neutral responses generally more prevalent among participants with lower cognitive performance (Table 2).

XSL•FO
RenderX

Table 2. AI-derived affective expression congruency by cognitive performance level across emotional and cognitive-task conditions.[a]

Stimulus condition and affective expression category	HC[b] (n=20), n (%)	MCI[c] (n=20), n (%)	Dementia (n=20), n (%)	Chi-square (*df*)	*P* value	Effect size
Positive video, 14 seconds				28.48 (4)	<.001	0.49
Congruent	13 (65)	4 (20)	0 (0)			
Neutral	5 (25)	16 (80)	20 (100)			
Incongruent	2 (10)	0 (0)	0 (0)			
Positive video, 52 seconds				15.53 (4)	.001	0.36
Congruent	15 (75)	9 (45)	4 (20)			
Neutral	4 (20)	11 (55)	16 (80)			
Incongruent	1 (5)	0 (0)	0 (0)			
Negative video, 48 seconds				26.29 (4)	<.001	0.47
Congruent	12 (60)	1 (5)	1 (5)			
Neutral	6 (30)	17 (85)	19 (95)			
Incongruent	2 (10)	2 (10)	0 (0)			
MMSE[d]-based cognitive-task condition				15.35 (4)	.004	0.36
Congruent	10 (50)	3 (15)	2 (10)			
Neutral	3 (15)	11 (55)	14 (70)			
Incongruent	7 (35)	6 (30)	4 (20)			

[a]Happy and surprised expressions were classified as positive responses; angry, disgusted, fearful, and sad expressions were classified as negative responses; and neutral expressions were retained separately. Positive responses were considered congruent for positive videos, whereas negative responses were considered congruent for the negative video and MMSE-based cognitive-task condition. The MMSE-based cognitive-task condition included MMSE memory recall and calculation subtests. Overall between-group differences in response distribution were assessed using Pearson chi-square tests or Fisher-Freeman-Halton exact tests, as appropriate. Effect sizes are reported as Cramér *V*. For comparisons with sparse expected cell counts, statistical significance was determined using the Fisher-Freeman-Halton exact *P* values, while Pearson chi-square statistic and Cramér *V* are reported as the corresponding test statistic and effect size.

[b]HC: healthy cognition.

[c]MCI: mild cognitive impairment.

[d]MMSE: Mini-Mental State Examination.

Post Hoc Comparisons of AI-Derived Affective Expression Congruency Between Groups

Bonferroni-corrected post hoc comparisons using a significance threshold *P*<.0167 showed significant between-group differences in selected testing conditions.

Compared with the MCI group, participants in the HC group showed significantly different affective expression distributions during the 14-second positive video stimulus and the 48-second negative video stimulus, whereas no significant differences were observed during the 52-second positive video stimulus, or MMSE-based cognitive-task condition. Similarly, the HC group showed significantly different affective expression distributions compared with the dementia group across all 4 conditions. In contrast, no statistically significant differences were observed between the MCI and dementia groups across any testing condition (Table 3).

XSL•FO
RenderX

Table 3. Pairwise post hoc comparisons of AI-derived affective expression congruency across cognitive performance groups.[a]

Condition	Congruent HC[b], n (%)	Congruent MCI[c], n (%)	Congruent dementia, n (%)	HC vs MCI, *P* value	HC vs dementia, *P* value	MCI vs dementia, *P* value
Positive video, 14 seconds	13 (65)	4 (20)	0 (0)	.002[d]	<.001[d]	.11
Positive video, 52 seconds	15 (75)	9 (45)	4 (20)	.048	<.001[d]	.18
Negative video, 48 seconds	12 (60)	1 (5)	1 (5)	<.001[d]	<.001[d]	.74
MMSE[e]-based cognitive task	10 (50)	3 (15)	2 (10)	.02	.001[d]	.65

[a]Positive responses were considered congruent for positive videos, whereas negative responses were considered congruent for the negative video and MMSE-based cognitive-task condition. Pairwise comparisons were based on the full distribution of congruent, neutral, and incongruent responses using Fisher-Freeman-Halton exact tests. Unadjusted exact *P* values are reported and were evaluated against a Bonferroni-corrected significance threshold of $P<.0167$ for the 3 pairwise comparisons within each condition.

[b]HC: healthy cognition.

[c]MCI: mild cognitive impairment.

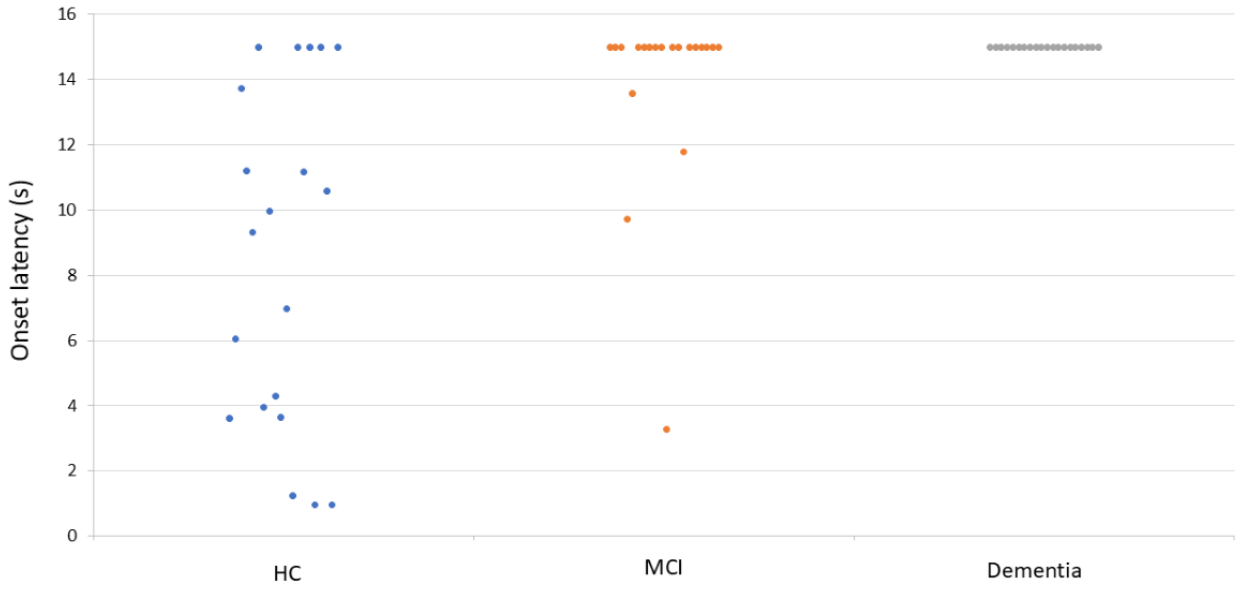
[d]Statistically significant at the Bonferroni-corrected threshold.

[e]MMSE: Mini-Mental State Examination.

AI-Derived Affective Expression Onset Latency Across Cognitive Performance Strata

AI-derived affective expression onset latency differed significantly across cognitive performance groups in all testing conditions, including the 14-second positive video stimulus ($F_{2,57}=19.63$; $P<.001$; $\eta^2=0.41$), 52-second positive video stimulus ($F_{2,57}=4.25$; $P=.02$; $\eta^2=0.13$), 48-second negative video stimulus ($F_{2,57}=15.61$; $P<.001$; $\eta^2=0.35$), and MMSE-based cognitive-task condition ($F_{2,57}=11.28$; $P<.001$; $\eta^2=0.28$). Overall, onset latency showed a consistent pattern of longer latency in the MCI and dementia groups than in the HC group across the emotional video conditions and the MMSE-based cognitive-task condition. Bonferroni-adjusted post hoc comparisons primarily indicated shorter onset latency in the HC group than in the MCI and dementia groups, whereas differences between the MCI and dementia groups were generally not statistically significant. The distributions of individual onset latency values are illustrated in Figures 1-4.

Figure 1. Distribution of AI-derived affective expression onset latency during the 14-second positive video stimulus. Individual data points represent participant-level onset latency values. HC: healthy cognition; MCI: mild cognitive impairment.



XSL•FO
RenderX

Figure 2. Distribution of AI-derived affective expression onset latency during the 52-second positive video stimulus. Individual data points represent participant-level onset latency values. HC: healthy cognition; MCI: mild cognitive impairment.

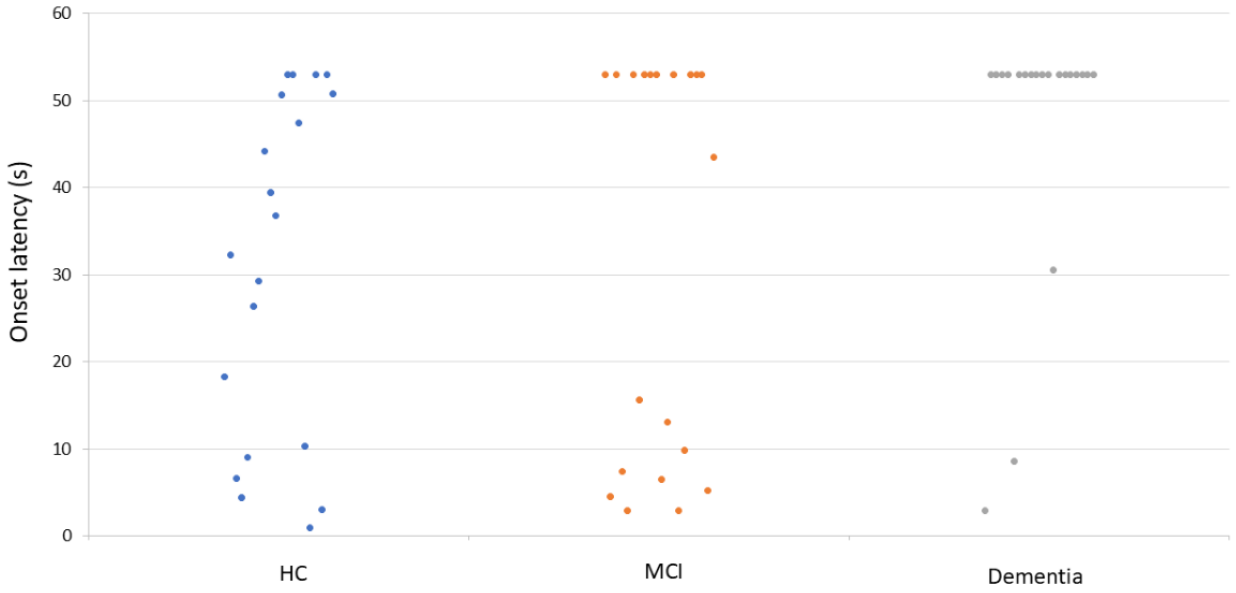
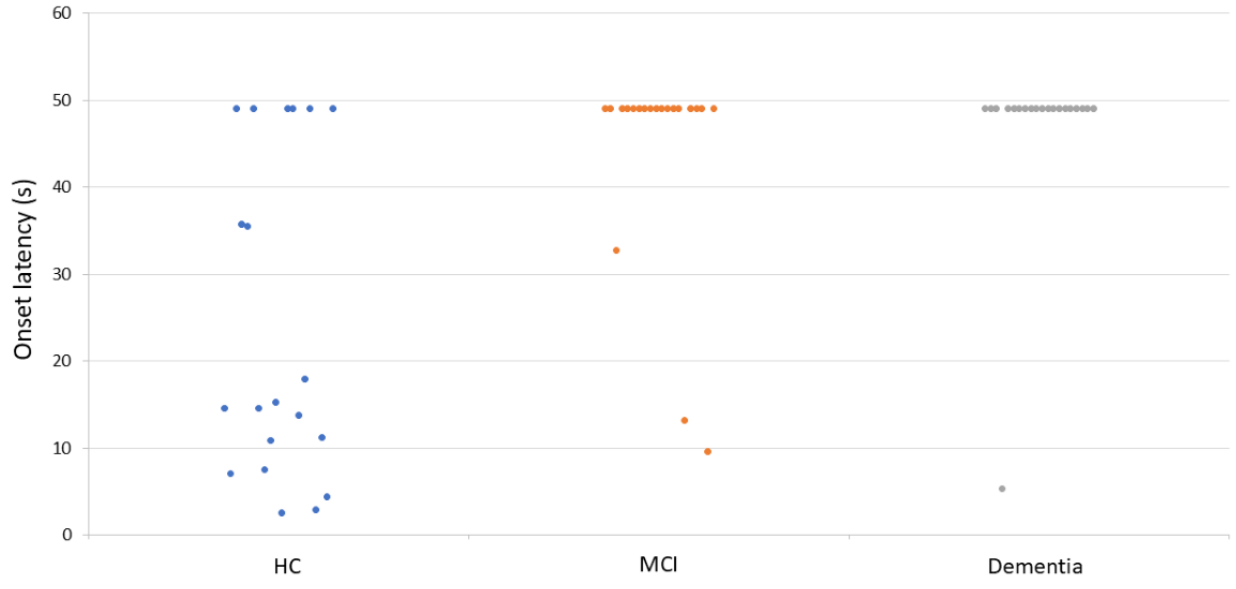
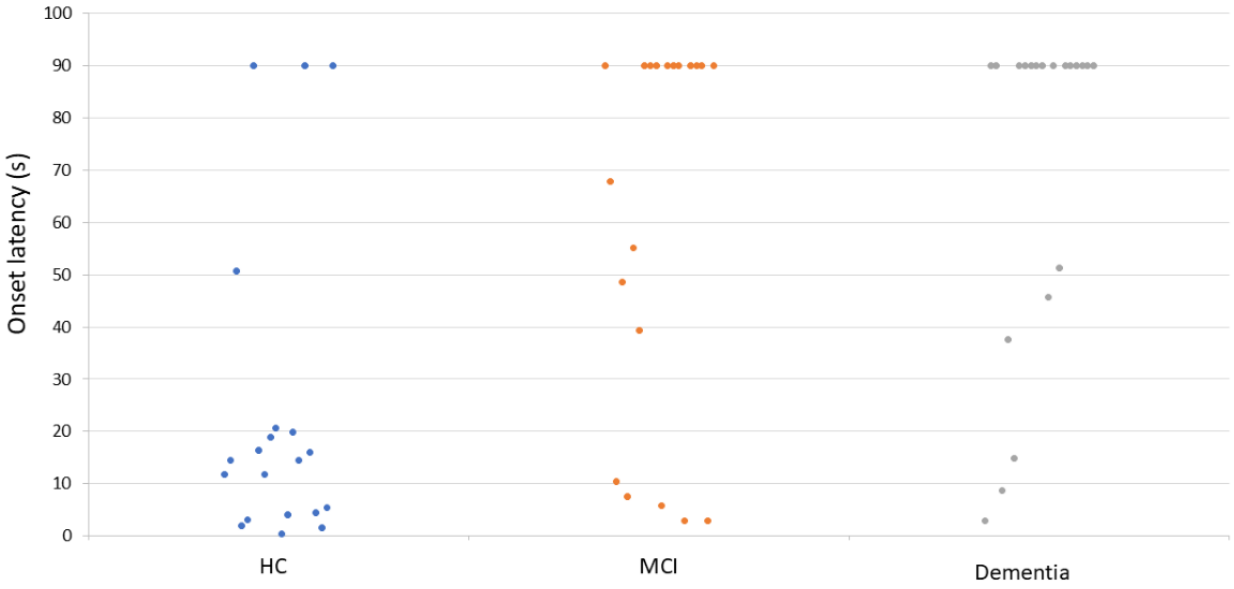


Figure 3. Distribution of AI-derived affective expression onset latency during the 48-second negative video stimulus. Individual data points represent participant-level onset latency values. HC: healthy cognition; MCI: mild cognitive impairment.



XSL•FO
RenderX

Figure 4. Distribution of AI-derived affective expression onset latency during the Mini-Mental State Examination–based cognitive-task condition. Individual data points represent participant-level onset latency values. HC: healthy cognition; MCI: mild cognitive impairment.



Covariate-Adjusted Analysis of Affective Expression Onset Latency

Because age and education level differed across cognitive performance groups, additional ANCOVA models were performed to examine whether group differences in AI-derived affective expression onset latency remained significant after adjustment for these variables. Cognitive performance group remained significantly associated with onset latency across all 4 testing conditions after adjustment for age and education level ($F_{2,55}$=3.89-10.68; all P≤.03; partial η^2=0.12-0.28; Table 4). In contrast, neither age nor education level showed significant covariate effects in any ANCOVA model. These findings suggest that the observed group differences in affective expression onset latency were not fully explained by baseline differences in age or education level.

Table 4. Covariate-adjusted analyses of AI-derived affective expression onset latency.[a]

Outcome variable	HC[b], adjusted mean (SE)	MCI[c], adjusted mean (SE)	Dementia, adjusted mean (SE)	*F* test (*df*)	*P* value	Partial η^2
Positive video 1 latency, 14 seconds	8.97 (0.91)	13.89 (0.80)	14.69 (0.86)	10.20 (2,55)	<.001	0.27
Positive video 2 latency, 52 seconds	29.09 (5.14)	32.90 (4.51)	48.34 (4.87)	3.89 (2,55)	.03	0.12
Negative video latency, 48 seconds	23.88 (3.66)	45.03 (3.21)	46.73 (3.47)	10.68 (2,55)	<.001	0.28
MMSE[d]-based cognitive-task latency	22.10 (8.66)	61.48 (7.61)	73.32 (8.21)	8.25 (2,55)	.001	0.23

[a]Analysis of covariance models included cognitive performance group as the fixed factor and age and education level as covariates. Education level was entered as a binary covariate coded as college education or higher vs below college education. Partial η^2 is reported as the effect size estimate for the group effect.

[b]HC: healthy cognition.

[c]MCI: mild cognitive impairment.

[d]MMSE: Mini-Mental State Examination.

XSL•FO
RenderX

ROC Curve Analysis of AI-Derived Affective Expression Onset Latency

In exploratory ROC analyses comparing HC and MCI performance levels, AI-derived affective facial expression onset latency showed variable separation across testing conditions. The 14-second positive video stimulus demonstrated an AUC of 0.803 (cutoff=11.48 seconds; sensitivity=90%; specificity=70%), while the 48-second negative video stimulus yielded a comparable AUC of 0.783 (cutoff=42.38 seconds; sensitivity=85%; specificity=70%). The MMSE-based cognitive-task condition showed an AUC of 0.741 (cutoff=29.99 seconds; sensitivity=75%; specificity=80%), whereas the 52-second positive video stimulus demonstrated limited separability (AUC=0.560).

In comparisons between HC and dementia performance levels, higher AUC values were observed across several task conditions. The 14-second positive stimulus yielded an AUC of 0.875 (cutoff=14.36 seconds; sensitivity=100%; specificity=75%), while the 52-second positive video stimulus and the 48-second negative video stimulus demonstrated AUC values of 0.800 and 0.815, respectively. The MMSE-based cognitive-task condition yielded an AUC of 0.823 (cutoff=29.14 seconds; sensitivity=85%; specificity=80%).

In comparisons between MCI and dementia performance levels, latency-based separation was comparatively limited, with AUC values ranging from 0.546 to 0.678 across testing conditions. Although sensitivity estimates were relatively high for several thresholds, specificity values remained low, suggesting substantial overlap in affective response latency between these groups.

Overall, AI-derived onset latency demonstrated variable exploratory classification performance across cognitive performance levels, with greater separation observed between cognitively healthy participants and those with lower cognitive performance. These latency-based behavioral signals may reflect group-level variability associated with cognitive performance; however, they should not be interpreted as diagnostic indicators and may require integration with additional behavioral or clinical measures in future studies (Table 5).

Table 5. Exploratory receiver operating characteristic analyses of AI-derived affective expression onset latency.

Comparison and condition	AUC[a]	Youden index	Cutoff (seconds)[b]	Sensitivity[c]	Specificity[c]
HC[d] vs MCI[e]					
Positive video 1 (14 s)	0.803	0.60	11.48	0.90	0.70
Positive video 2 (52 s)	0.560	0.30	51.89	0.50	0.80
Negative video (48 s)	0.783	0.55	42.38	0.85	0.70
MMSE[f]-based cognitive-task condition	0.741	0.55	29.99	0.75	0.80
HC vs dementia					
Positive video 1 (14 s)	0.875	0.75	14.36	1.00	0.75
Positive video 2 (52 s)	0.800	0.65	51.89	0.85	0.80
Negative video (48 s)	0.815	0.65	42.38	0.95	0.70
MMSE-based cognitive-task condition	0.823	0.65	29.14	0.85	0.80
MCI vs dementia					
Positive video 1 (14 s)	0.600	0.20	14.29	1.00	0.20
Positive video 2 (52 s)	0.678	0.35	48.22	0.85	0.50
Negative video (48 s)	0.546	0.10	48.86	0.95	0.15
MMSE-based cognitive-task condition	0.578	0.15	8.09	0.95	0.20

[a]AUC: area under the curve.

[b]Cutoff values indicate AI-derived affective expression onset latency in seconds.

[c]Sensitivity and specificity are presented as proportions.

[d]HC: healthy cognition.

[e]MCI: mild cognitive impairment.

[f]MMSE: Mini-Mental State Examination.

Discussion

Principal Findings

This exploratory feasibility study examined whether AI-derived measures of spontaneous affective facial responses, including affective expression congruency and onset latency, varied across cognitive performance levels in older adults. The main findings were that participants with lower cognitive performance demonstrated lower affective expression congruency and longer onset latency during both emotionally evocative video viewing and MMSE-based cognitive-task conditions. These group

XSL•FO
RenderX

differences remained significant for onset latency after adjustment for age and education level. Exploratory ROC analyses further suggested that latency-based features showed greater separation between cognitively healthy participants and those with lower cognitive performance than between the MCI and dementia groups. Together, these findings suggest that AI-derived facial behavioral metrics may capture observable variation in spontaneous affective engagement across cognitive performance levels, although they should be interpreted as preliminary and nondiagnostic markers.

Comparison With Prior Work

Alterations in emotional processing and affective responsiveness have been increasingly recognized in individuals with cognitive impairment. Prior studies have reported impairments in the perceptual recognition of emotional facial expressions in individuals with MCI and Alzheimer dementia, particularly for negative emotions such as sadness or fear [9,23]. More recent evidence also supports impairments in facial emotion recognition among older adults with MCI and broader alterations in social cognition across the MCI-dementia continuum [7,8]. However, these paradigms primarily assess explicit perceptual recognition of emotional stimuli rather than spontaneous affective expression during emotionally evocative or cognitively demanding conditions. These findings suggest that reduced congruency of observable facial expressions in cognitively impaired individuals may reflect alterations in spontaneous affective engagement rather than perceptual emotion recognition ability alone. Such observable variation in affective expression may be associated with reduced behavioral responsiveness or attenuated affective expressivity, phenomena that have previously been described in individuals with MCI and dementia and linked to neuropsychiatric features such as apathy [6].

The relevance of automated facial expressivity analysis in neurocognitive disorders is also supported by a previous exploratory study that used computer vision–based facial analysis to quantify facial expressivity in older adults with neurocognitive disorders and examined its association with apathy symptoms [11]. That study suggested that reduced facial expressivity may provide objective behavioral information related to emotional blunting and apathy in neurocognitive populations. In line with this work, this study further extends automated facial behavior analysis by focusing on spontaneous affective response congruency and onset latency during emotionally evocative and cognitively demanding conditions.

Prolonged onset latency of spontaneous affective facial responses may be interpreted as a behavioral manifestation of altered emotional expressivity rather than as evidence of impaired perceptual recognition of emotional stimuli. Previous studies have reported reduced emotional facial responding and variation in affective motor expression in individuals with neurodegenerative disorders, which may contribute to delayed or attenuated observable emotional responses under emotionally evocative conditions [15,24]. Similarly, the increased prevalence of neutral facial expressions observed during cognitively demanding tasks may reflect reduced behavioral expression of affective engagement under cognitive load conditions. The limited separation observed between MCI and dementia performance levels across task conditions may further suggest that observable variation in spontaneous affective expression is present across multiple levels of reduced cognitive performance but does not clearly distinguish adjacent impairment levels.

In the cognitive-task condition, negative facial expressions were interpreted as task-congruent only in an operational sense. Recall and calculation tasks may elicit observable negative affective responses related to cognitive effort, frustration, uncertainty, or perceived task difficulty. This interpretation is consistent with recent evidence suggesting that facial behavioral signals and facial action units may be informative for cognitive load estimation and that automatically detected facial expressions may reflect difficulty-related emotional responses under challenging task conditions [21,22]. Therefore, reduced negative expressivity during this condition may indicate attenuated observable affective engagement under cognitive load rather than a direct absence of cognitive effort.

Digital Phenotyping and Clinical Implications

These findings may also be interpreted within the emerging framework of digital phenotyping and digital biomarkers. Recent work has emphasized that digital biomarkers in Alzheimer disease and related cognitive disorders may include objective and quantifiable physiological or behavioral data collected through digital technologies, including wearable sensors, eye tracking, speech analysis, and other behavioral signal streams [25]. In addition, recent scoping evidence has highlighted the growing use of AI-based technologies for the early detection of behavioral and psychological symptoms of dementia, supporting the clinical relevance of AI-assisted behavioral monitoring in dementia care settings [10]. Such approaches may complement conventional episodic clinical assessments by enabling more objective and potentially longitudinal monitoring of subtle cognitive, behavioral, and functional changes. In parallel, recent studies on contactless facial expression assessment have highlighted automated facial expression analysis as a promising approach for monitoring cognitive and emotional changes in older adults [26]. Within this context, this study extends prior work by examining not only the presence or category of facial affective expressions but also their temporal characteristics, particularly onset latency, during emotionally evocative and cognitively demanding conditions. These findings support the potential relevance of AI-derived facial behavioral metrics as adjunctive, noninvasive features within future multimodal digital phenotyping frameworks.

Exploratory Classification Performance

Exploratory ROC analyses indicated that AI-derived onset latency of spontaneous affective facial responses demonstrated variable classification performance across cognitive performance levels. Greater separation was observed between cognitively healthy participants and those with lower cognitive performance, whereas separation between the MCI and dementia groups was limited. Although sensitivity estimates were relatively high in several emotionally evocative and cognitively demanding task conditions, specificity remained limited, suggesting substantial overlap in affective response timing between adjacent cognitive performance levels.

XSL•FO
RenderX

These findings suggest that temporal characteristics of spontaneous affective facial responses may reflect group-level behavioral variability but may have limited separability when used in isolation. Behavioral timing metrics alone may therefore provide only partial information regarding affective response variability, underscoring the potential need to integrate these features with additional behavioral or clinical measures in future multimodal observational frameworks [25].

Limitations

This study has several limitations. First, the modest sample size may limit generalizability, and the cross-sectional design does not permit examination of longitudinal changes in behavioral response timing.

Second, although the ANCOVA models adjusted for age and education level, other baseline differences, including lifestyle-related factors such as physical activity and social participation, were not included as covariates. Previous epidemiological studies and recent dementia prevention evidence have reported associations between lifestyle-related factors, including physical activity and social participation, and cognitive health in older adults [27,28]. Therefore, variation in AI-derived affective facial expression onset latency may partly reflect residual confounding related to baseline lifestyle differences rather than cognitive performance alone.

Third, baseline affective state, apathy, depressive symptoms, and other preexisting mood states were not assessed using standardized instruments such as the Positive and Negative Affect Schedule [29]. Given prior evidence linking reduced facial expressivity with apathy-related features in neurocognitive disorders [11], spontaneous facial expressivity and AI-derived onset latency may have been influenced by participants' preexisting emotional states rather than task-evoked affective responses alone. The interpretation of negative facial expressions during the MMSE-based cognitive-task condition should also be considered exploratory. Although these expressions were operationally treated as task-congruent behavioral responses, they may reflect a mixture of cognitive effort, frustration, uncertainty, or affective discomfort rather than cognitive effort alone [21,22].

Fourth, cognitive grouping in this study was based on screening-level cognitive assessment for the HC group and clinical evaluation, CDR staging, and neuroimaging review for clinically referred participants in the MCI and dementia groups as part of standard clinical care. However, comprehensive neuropsychological profiling, biomarker-supported classification, or dementia subtype–specific analyses were not systematically performed. Therefore, the dementia group was analyzed as a single clinically defined category, and subtype heterogeneity, including Alzheimer disease, dementia with Lewy bodies, or vascular cognitive impairment, may have influenced facial expressivity and AI-derived onset latency.

Fifth, emotional stimuli were restricted to 2 emotional categories, positive and negative, which may not fully represent the broader spectrum of affective responsiveness across cognitive performance levels. The relatively short video durations may also have constrained emotional elicitation, particularly among participants with slower spontaneous affective engagement. In addition, all participants received the same fixed sequence of stimuli, including positive video stimuli, MMSE-based cognitive-task condition, and a negative video stimulus; therefore, potential order effects, habituation, fatigue, or affective carryover between testing conditions cannot be excluded. Although practical steps were taken during testing to support audiovisual stimulus perception, formal standardized visual and hearing assessments were not performed. Therefore, unmeasured sensory impairment may have influenced stimulus perception and AI-derived onset latency independently of cognitive performance. Cultural and linguistic mismatches between certain stimuli and participants may have further influenced emotional engagement and observable facial responses. These methodological considerations are consistent with recent reviews emphasizing the need for standardized data collection, improved interpretability, and multimodal integration in AI-based facial emotion analysis for dementia-related applications [30].

Finally, the ROC analyses were exploratory and were not internally validated using leave-one-out or k-fold cross-validation. Therefore, the AUC values and Youden index–derived cutoff values may be sample-dependent and should be interpreted as preliminary estimates rather than validated classification thresholds.

Future Directions

Future research should consider methodological refinements to improve the interpretability, ecological validity, and generalizability of AI-assisted facial expression analysis across varying levels of cognitive performance. Expanding the range of emotional stimuli beyond positive and negative affective categories may provide a more comprehensive characterization of emotion-cognition interactions across cognitive performance levels. The delayed affective onset observed in participants with lower cognitive performance suggests that brief stimuli may be insufficient to elicit measurable spontaneous facial responses in some individuals; therefore, future protocols may benefit from longer, culturally and linguistically appropriate, or more emotionally salient stimuli. In addition, longitudinal investigations in larger and more diverse cohorts will be necessary to examine intraindividual changes in affective facial responsiveness over time and to determine whether AI-derived facial behavioral metrics can contribute to multimodal monitoring frameworks for cognitive health.

Conclusion

In this exploratory feasibility study, AI-assisted facial expression analysis revealed observable variation in the timing and congruency of spontaneous affective facial responses across cognitive performance levels. Participants with lower cognitive performance demonstrated delayed onset of observable affective facial expressions and a higher prevalence of neutral or incongruent responses during both emotionally evocative video viewing and cognitively demanding task conditions.

These findings suggest that temporal and categorical features of facial expressions, particularly onset latency and expression congruency, may reflect group-level variation in spontaneous

XSL•FO
RenderX

affective engagement across cognitive performance levels. Latency-based measures showed preliminary potential for distinguishing cognitively healthy participants from those with lower cognitive performance, although separation between adjacent impairment levels remained limited.

AI-derived quantification of affective facial responsiveness may complement existing behavioral and clinical observations by providing an additional noninvasive observational measure in both clinical and community-based aging contexts. Further studies involving larger cohorts, longitudinal designs, and internal and external validation procedures are required to examine the stability, generalizability, and clinical relevance of these behavioral features over time.

Acknowledgments

We gratefully acknowledge the staff of the Department of Medical Research at National Taiwan University Hospital, Hsin-Chu Branch, for their valuable assistance with statistical analysis. We also sincerely thank our colleagues for their support in participant recruitment and study coordination. Finally, we are deeply grateful to all participants and their families for their voluntary contribution to this research.

Funding

This research received no external research funding. Institutional support from National Taiwan University Hospital was limited to nonresearch-related assistance, including language editing services and possible partial publication grants upon manuscript acceptance.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request. Deidentified data will be shared in accordance with institutional and ethical guidelines.

Generative AI Use

ChatGPT (OpenAI) was used to support language editing, grammar and spelling checks, and formatting during manuscript revision. It was also used to assist in compositing, stylizing, and editing two author-provided source images, including a photograph of the study setup and a backend facial-landmark analysis image, to create the TOC/Feature image. The final image is a substantially modified, cartoon-like conceptual illustration that does not depict a recognizable individual and is not presented as a quantitative study result. All scientific content, analyses, interpretations, and references were reviewed and verified by the authors, who take full responsibility for the final manuscript.

Authors' Contributions

JCY contributed to conceptualization, methodology, investigation, data curation, and writing–original draft. TJL performed formal analysis, data curation, visualization, and writing–review & editing. BSY was responsible for investigation, resources, methodology, supervision, and writing–review & editing. He also contributed to clinical evaluation, diagnostic classification, neurological assessment, neuroimaging review of computed tomography findings, and case referral for study recruitment. WHW contributed to investigation, resources, project administration, and writing–review & editing. WHC was responsible for data curation, resources, project administration, and writing–review & editing. HTC contributed to conceptualization, methodology, software, validation, supervision, and writing–review & editing. PSL contributed to conceptualization, methodology, supervision, and writing–review & editing.

All authors meet all 4 authorship criteria established by the International Committee of Medical Journal Editors (ICMJE), including substantial contributions to the conception or design of the work or the acquisition, analysis, or interpretation of data; drafting the manuscript or critically revising it for important intellectual content; final approval of the submitted version; and agreement to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

References

1. Analysis of household registration statistics for December 2025. Ministry of the Interior. 2026. URL: https://www.moi.gov.tw/News_Content.aspx?n=9&s=335870&sms=9009 [accessed 2026-09-13]
2. Ma KZM, Hsu CC. Prevalence, subtypes, and associated factors of dementia in Taiwan: a nationwide survey of 11,298 community‐dwelling adults aged 65 years or older. Alzheimers Dement. 2025;21(S6):e102326. [doi: 10.1002/alz70860_102326] [Medline: 41434825]
3. Rafii MS, Aisen PS. Detection and treatment of Alzheimer's disease in its preclinical stage. Nat Aging. 2023;3(5):520-531. [FREE Full text] [doi: 10.1038/s43587-023-00410-4] [Medline: 37202518]

XSL•FO
RenderX

4. Mohanty M, Kumar P. Multi-component interventions in older adults having subjective cognitive decline (SCD)-a review article. Geriatrics (Basel). 2022;8(1):4. [FREE Full text] [doi: 10.3390/geriatrics8010004] [Medline: 36648909]
5. Kim SJ, Kim MJ, Kim JS, Jang J, Choi W, Lee HJ, et al. AI-enabled digital phenotyping for Alzheimer's disease: a review of multimodal sensor integration and symptom trajectories. Alzheimers Res Ther. 2026;18(1):111. [FREE Full text] [doi: 10.1186/s13195-026-02013-8] [Medline: 41913245]
6. Spoletini I, Marra C, Di Iulio F, Gianni W, Sancesario G, Giubilei F, et al. Facial emotion recognition deficit in amnestic mild cognitive impairment and Alzheimer disease. Am J Geriatr Psychiatry. 2008;16(5):389-398. [doi: 10.1097/JGP.0b013e318165dbce] [Medline: 18403572]
7. Ju EY, Kim CY, Choi BY, Ryoo SW, Min JY, Min KB. Deficits of facial emotion recognition in elderly individuals with mild cognitive impairment. Dement Geriatr Cogn Disord. 2024;53(6):321-328. [doi: 10.1159/000540364] [Medline: 39043155]
8. Shi P, Chapman H, Liu L, Rodgers F, Shaw J, Livingston G, et al. Social cognition in mild cognitive impairment and dementia: a systematic review and meta-analysis. Alzheimers Dement. 2025;21(3):e70076. [doi: 10.1002/alz.70076] [Medline: 40145360]
9. Virtanen M, Singh-Manoux A, Batty GD, Ebmeier KP, Jokela M, Harmer CJ, et al. The level of cognitive function and recognition of emotions in older adults. PLoS One. 2017;12(10):e0185513. [FREE Full text] [doi: 10.1371/journal.pone.0185513] [Medline: 28977015]
10. Fernandes S, Rosselet Amoussou J, Gomes da Rocha C, Perruchoud E, von Gunten A, Mabire C, et al. Using artificial intelligence-based technologies for the early detection of behavioral and psychological symptoms of dementia: scoping review. JMIR Aging. 2025;8:e76074. [FREE Full text] [doi: 10.2196/76074] [Medline: 41118658]
11. Zeghari R, König A, Guerchouche R, Sharma G, Joshi J, Fabre R, et al. Correlations between facial expressivity and apathy in elderly people with neurocognitive disorders: exploratory study. JMIR Form Res. 2021;5(3):e24727. [FREE Full text] [doi: 10.2196/24727] [Medline: 33787499]
12. Jiang Z, Seyedi S, Haque RU, Pongos AL, Vickers KL, Manzanares CM, et al. Automated analysis of facial emotions in subjects with cognitive impairment. PLoS One. 2022;17(1):e0262527. [FREE Full text] [doi: 10.1371/journal.pone.0262527] [Medline: 35061824]
13. Bergamasco L, Coletta A, Olmo G, Cermelli A, Rubino E, Rainero I. AI-based facial emotion analysis for early and differential diagnosis of dementia. Bioengineering (Basel). 2025;12(10):1082. [FREE Full text] [doi: 10.3390/bioengineering12101082] [Medline: 41155081]
14. Kumfor F, Irish M, Hodges JR, Piguet O. Discrete neural correlates for the recognition of negative emotions: insights from frontotemporal dementia. PLoS One. 2013;8(6):e67457. [FREE Full text] [doi: 10.1371/journal.pone.0067457] [Medline: 23805313]
15. Sturm VE, Ascher EA, Miller BL, Levenson RW. Diminished self-conscious emotional responding in frontotemporal lobar degeneration patients. Emotion. 2008;8(6):861-869. [FREE Full text] [doi: 10.1037/a0013765] [Medline: 19102597]
16. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
17. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology. 1993;43(11):2412-2414. [doi: 10.1212/wnl.43.11.2412-a] [Medline: 8232972]
18. Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees. 2014. Presented at: 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2014 June 24-27:1867-1874; Columbus, OH, USA. [doi: 10.1109/cvpr.2014.241]
19. Neta M, Davis FC, Whalen PJ. Valence resolution of ambiguous facial expressions using an emotional oddball task. Emotion. 2011;11(6):1425-1433. [FREE Full text] [doi: 10.1037/a0022993] [Medline: 21707167]
20. Neta M, Kim MJ. Surprise as an emotion: a response to Ortony. Perspect Psychol Sci. 2023;18(4):854-862. [FREE Full text] [doi: 10.1177/17456916221132789] [Medline: 36356055]
21. Oppelt MP, Foltyn A, Deuschel J, Lang NR, Holzer N, Eskofier BM, et al. ADABase: a multimodal dataset for cognitive load estimation. Sensors (Basel). 2023;23(1):340. [FREE Full text] [doi: 10.3390/s23010340] [Medline: 36616939]
22. Venkitakrishnan S, Wu YH. Facial expressions as an index of listening difficulty and emotional response. Semin Hear. 2023;44(2):166-187. [FREE Full text] [doi: 10.1055/s-0043-1766104] [Medline: 37122878]
23. Elferink MW, van Tilborg I, Kessels RPC. Perception of emotions in mild cognitive impairment and Alzheimer's dementia: does intensity matter? Transl Neurosci. 2015;6(1):139-149. [FREE Full text] [doi: 10.1515/tnsci-2015-0013] [Medline: 28123797]
24. Marshall CR, Hardy CJD, Russell LL, Clark CN, Bond RL, Dick KM, et al. Motor signatures of emotional reactivity in frontotemporal dementia. Sci Rep. 2018;8(1):1030. [FREE Full text] [doi: 10.1038/s41598-018-19528-2] [Medline: 29348485]
25. Qi W, Zhu X, Wang B, Shi Y, Dong C, Shen S, et al. Alzheimer's disease digital biomarkers multidimensional landscape and AI model scoping review. NPJ Digit Med. 2025;8(1):366. [FREE Full text] [doi: 10.1038/s41746-025-01640-z] [Medline: 40523935]

XSL•FO
RenderX

26. Jiang D, Yan L, Mayrand F. Emotion expressions and cognitive impairments in the elderly: review of the contactless detection approach. Front Digit Health. 2024;6:1335289. [FREE Full text] [doi: 10.3389/fdgth.2024.1335289] [Medline: 39040877]
27. Blondell SJ, Hammersley-Mather R, Veerman JL. Does physical activity prevent cognitive decline and dementia?: a systematic review and meta-analysis of longitudinal studies. BMC Public Health. 2014;14:510. [FREE Full text] [doi: 10.1186/1471-2458-14-510] [Medline: 24885250]
28. Kuiper JS, Zuidersma M, Oude Voshaar RC, Zuidema SU, van den Heuvel ER, Stolk RP, et al. Social relationships and risk of dementia: a systematic review and meta-analysis of longitudinal cohort studies. Ageing Res Rev. 2015;22:39-57. [doi: 10.1016/j.arr.2015.04.006] [Medline: 25956016]
29. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol. 1988;54(6):1063-1070. [doi: 10.1037//0022-3514.54.6.1063] [Medline: 3397865]
30. Gerbaudo-González N, Condorí-Fernández N, Catalá A, Gandoy-Crego M. AI-based emotion recognition in dementia through facial expression: a scoping review. Ageing Res Rev. 2026;114:102975. [FREE Full text] [doi: 10.1016/j.arr.2025.102975] [Medline: 41318099]

Abbreviations

ANCOVA: analysis of covariance
AUC: area under the curve
CDR: Clinical Dementia Rating
CNN: convolutional neural network
HC: healthy cognition
MCI: mild cognitive impairment
MMSE: Mini-Mental State Examination
ROC: receiver operating characteristic

Edited by L MacNeill; submitted 17.Mar.2026; peer-reviewed by C Oh; comments to author 28.May.2026; revised version received 21.Sep.2026; accepted 09.Oct.2026; published 02.Nov.2026

Please cite as:
Yang J-C, Lai T-J, Yip B-S, Wu W-H, Chen W-H, Chang H-T, Lin P-S
AI-Assisted Analysis of Spontaneous Affective Facial Responses for Cognitive Status Assessment in Older Adults: Exploratory Feasibility Study
JMIR Form Res 2026;10:e94422
URL: https://formative.jmir.org/2026/1/e94422
doi: 10.2196/94422
PMID:

©Jung-Cheng Yang, Ting-Ju Lai, Bak-Sau Yip, Wei-Hui Wu, Wun-Hung Chen, Hsien-Tsung Chang, Pay-Shin Lin. Originally published in JMIR Formative Research (https://formative.jmir.org), 02.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX