Estimating Attentional States of Older Adults for Remote Caregiving

Thomas M. Kwok[1], Yuanze Huang[2], Clara Colombatto [2], Yue Hu[1]

Abstract— **Remote caregiving robots have the potential to support older adults aging in place, but caregivers often lack awareness of users' attentional states during remote interactions. Existing attentiveness estimation approaches have primarily been developed using younger populations and may not generalize well to older adults due to age-related changes in facial appearance and behavior. This pilot study investigates facial indicators of attentional engagement in older adults performing a Sustained Attention to Response Task (SART). Four older adults (69.5 ± 5.1 years) completed the task while facial videos were recorded and analyzed using OpenFace 2.0. Facial Action Units (AUs) were extracted and examined using mixed-effects statistical models. Results identified several facial behaviors associated with attentiveness, including increased occurrence and intensity of eye blinks (AU45), lip parting (AU25), lip corner pulling (AU12), and chin raising (AU17), while AU14 (dimpler) was negatively associated with attentive states. These findings provide preliminary evidence that attentional engagement in older adults can be inferred from age-specific facial behaviors and offer a foundation for developing age-aware attentiveness estimation systems for remote caregiving and human–robot interaction.**

I. INTRODUCTION

As populations age, the demand for caregiving support continues to rise. At the time of writing, individuals aged 60 or older will account for nearly 22% of the global population by 2050 [1]. While many older adults prefer to age in place, age-related functional decline and limited access to care can make independent living challenging [2], [3].

Recent advances in assistive robotics [4], [5] show promise in delivering remote caregiving support. Such robots enable caregivers to assist older adults from a distance with physical and verbal support for daily tasks. However, existing teleoperation frameworks mainly focus on the user's physical state and surrounding environments. Comparatively little work has been done to assess users' cognitive states, such as attentional engagement, which are essential for caregivers to ensure that users understand caregiving instructions and perform tasks safely.

Attentional engagement is a fundamental cognitive process that enables individuals to maintain focus, monitor ongoing events, and detect errors during everyday activities [6], [7]. Although older adults can maintain sustained attention for extended periods of time [8], aging is associated with changes in response behavior and information processing [9]. This may affect their performance in daily activities. For caregivers operating remotely, determining whether an older adult is attentive and ready to receive assistance remains challenging, particularly when camera feedback captures mostly the task rather than the older adult, and when the complexity of multimodal sensing interfaces further limits observation [4]. As a result, important cues to attentional engagement may be overlooked, leading to less effective robotic assistance.

Equipping assistive robots with attentiveness estimation capabilities could identify disengagement and notify caregivers to adapt their support accordingly. However, most existing attentiveness estimation methods in psychology and computer vision have been developed and validated using younger adult populations [10], [11], [12], [13], [14]. Aging is associated with changes in facial morphology, musculature [15], gaze characteristics [16], scleral coloration [17], and motor behavior [18] that influence how attentional states are manifest in their appearance. These age-related changes may reduce the reliability and generalizability of conventional attentional cues. Consequently, attentiveness estimation systems may produce biased assessments when applied to older adults.

To address this gap, there is a need to systematically investigate how attentional states manifest in older adults and to identify age-specific facial cues. In this work, we present a pilot study examining attentional facial behaviors in older adults during a Sustained Attention to Response Task (SART) [19]. Using facial Action Units [20] extracted with OpenFace 2.0 [21], we analyze how attentional states are expressed through facial activity and identify candidate cues for age-aware attentiveness estimation. By bridging cognitive psychology and human–robot interaction, this work contributes toward the development of socially and cognitively aware assistive robots for remote caregiving.

II. METHODOLOGY

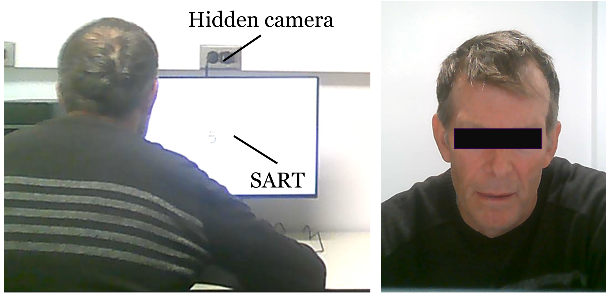
A. Experiment Task

We conducted an experiment using a hidden camera to capture the facial expressions of older adults during the Sustained Attention to Response Task (SART). The setup is shown in Fig. 1. In this task, participants were presented with a rapid stream of single-digit numbers displayed sequentially on a computer screen. They were instructed to respond as quickly as possible to each digit by pressing the spacebar, except for a designated target digit ("3"), for which they were required to withhold their response.

In total, 225 single-digit stimuli were presented, with each digit from 1 through 9 appearing 25 times. Each digit

[1] Department of Mechanical and Mechatronics Engineering, University of Waterloo, Canada, thomasm.kwok@uwaterloo.ca, yue.hu@uwaterloo.ca, [2] School of Psychology, University of Waterloo, Canada, yuanze.huang@uwaterloo.ca, clara.colombatto@uwaterloo.ca

This work was supported by the University of Waterloo's Network for Aging Research (NAR) Catalyst Grant.



(a) (b)

Fig. 1. (a) Experiment setup showing a participant performing the Sustained Attention to Response Task (SART) while being recorded by a hidden camera. (b) Example video frame captured from the hidden camera, showing the participant's facial expressions during the task.

was displayed for 250 ms and followed by a mask for 900 ms, resulting in a 1150-ms interval between successive digit onsets. Across the task, the digit "3" appeared on 25 trials, distributed throughout the sequence in a quasi-random manner. This design established a strong prepotent response tendency, making failures to inhibit responses to the target digit indicative of disengagement, while correct inhibitions reflected sustained attentional engagement.

B. Data Processing

As for recording processing, facial features were extracted offline using OpenFace 2.0, with facial action units (AUs) as indicators of attentional engagement.

C. Statistical Analyses

We analyzed the relationships between facial Action Units (AUs) and attentional conditions. For binary AUs, indicating whether an AU was present or absent, we used logistic mixed-effects models. For continuous AU intensity measures, we used linear mixed-effects models. Each AU was examined separately to account for variability across participants and frames. To compare multiple AUs, adjusted p-values were reported using false discovery rate (FDR) in Table I.

TABLE I

SIGNIFICANT ACTION UNITS ASSOCIATED WITH ATTENTIONAL STATE.

AU*	Estimate	Statistic	p-value#
AU12$_c$ (Lip Corner Puller)	4.83	OR = 125.28	< 0.001
AU25$_c$ (Lips Part)	4.44	OR = 84.66	< 0.001
AU45$_c$ (Blink)	4.09	OR = 59.88	< 0.001
AU17$_c$ (Chin Raiser)	1.29	OR = 3.64	< 0.01
AU14$_c$ (Dimpler)	-7.34	OR = 0.001	< 0.001
AU45$_r$ (Blink)	0.07	$t = 4.98$	< 0.001
AU26$_r$ (Jaw Drop)	0.09	$t = 4.32$	< 0.001
AU25$_r$ (Lips Part)	0.06	$t = 3.81$	< 0.01
AU15$_r$ (Lip Corner Depressor)	0.02	$t = 3.63$	< 0.01
AU01$_r$ (Inner Brow Raiser)	0.04	$t = 3.50$	< 0.01
AU17$_r$ (Chin Raiser)	0.05	$t = 3.26$	< 0.01
AU09$_r$ (Nose wrinkler)	0.008	$t = 2.84$	< 0.05
AU20$_r$ (Lip Stretcher)	0.02	$t = 2.60$	< 0.05
AU23$_r$ (Lip Tightener)	0.02	$t = 2.52$	< 0.05

* AU$_c$ denotes Action Unit occurrence, and AU$_r$ denotes Action Unit intensity, analyzed with logistic and linear mixed models, respectively.
# p-values were adjusted using false discovery rate (FDR).

D. Participants

A pilot sample of 4 older adults (69.5 ± 5.1 years old, 2 males, 4 Caucasian) was recruited. Participants met inclusion criteria, including normal vision with/without glasses, no attention-related disorders and no motor impairments affecting facial or body movements. All participants provided informed consent before participation and were fully debriefed afterward. The protocols were reviewed and received ethics clearance through a University of Waterloo Research Ethics Board (REB-46826).

III. Results

The results in Table I demonstrate that attentional states in older adults are associated with distinct facial behaviors, primarily involving eye blinks and lower-face movements.

For AU occurrence, the linear mixed model revealed several facial actions that significantly differentiated attentional states after multiple comparison correction. In particular, AU12 (lip corner puller), AU25 (lips part), and AU45 (blink) showed strong positive associations with participants' attentiveness, with large effect sizes (odds ratios = 125.28, 84.66 and 59.88 respectively, $p < 0.001$), indicating that these movements were substantially more likely to occur during attentive periods. Similarly, AU17 (chin raiser) was positively associated with attentiveness (odds ratios = 3.64, $p < 0.01$), suggesting subtle lower-face engagement during focused states. In contrast, AU14 (dimpler) demonstrated a significant negative relationship, with a very small odds ratio (i.e., 0.001, $p < 0.001$), indicating that this facial configuration was less frequent during attentive states.

Complementing the occurrence analysis, linear mixed modeling of AU intensities revealed a broader set of facial dynamics associated with attentiveness. Similar to AU occurrence, AU45 (blink intensity)($t = 4.98, p < 0.001$) emerged as one of the strongest predictors. As for mouth-related activity, increased intensities were observed for AU26 (jaw drop) ($t = 4.32, p < 0.001$), AU25 (lips part) ($t = 3.81, p < 0.01$), AU15 (lip corner depressor) ($t = 3.63, p < 0.01$), AU20 (lip stretcher) ($t = 2.60, p < 0.05$), and AU23 (lip tightener) ($t = 2.52, p < 0.05$). Besides, we observed other significant effects included AU01 (inner brow raiser) ($t = 3.50, p < 0.01$), AU17 (chin raiser) ($t = 3.26, p < 0.01$), and AU09 (nose wrinkler) ($t = 2.84, p < 0.05$), suggesting coordinated engagement across upper and lower facial regions.

IV. Discussion and Future Work

The results of this study provide preliminary evidence that attentional states in older adults can be inferred from AUs. Notably, attentiveness was more consistently reflected in specific AUs, such as blinks and lower-face movements.

From a human–robot interaction perspective, these findings have important implications for the design of socially assistive robotic systems. Incorporating attentiveness estimation into teleoperated or autonomous robots could enable more responsive and adaptive interactions, allowing systems to adjust their behavior based on the user's cognitive engagement. For example, a robot could slow down instructions, repeat information, or prompt re-engagement when signs of inattention are detected. Such capabilities are particularly relevant for remote caregiving scenarios, where caregivers must rely on limited sensory input and may otherwise miss subtle cues of disengagement. By grounding attentional inference in age-aware facial features, this work contributes toward the development of robots that are not only functionally effective but also socially and cognitively attuned to older users.

Despite these contributions, the present study has several limitations that point to directions for future research. The most immediate limitation is the small sample size, which constrains the generalizability of the findings. Expanding the dataset to include a larger cohort of older adults will be essential for validating the findings and ensuring robustness across demographic factors such as age range, ethnicity, and cognitive ability. In addition, while this work focuses exclusively on AUs, attentiveness is inherently multimodal. Future work should integrate complementary signals such as gaze directions and head movement to develop a more comprehensive representation of attentional engagement. Another important direction involves moving from offline analysis toward real-time attentiveness estimation. Embedding the proposed approach within a teleoperation framework would enable continuous monitoring of user engagement and adaptive robot control strategies during interaction.

Finally, understanding how cognitive states manifest in older adults is critical for developing assistive technologies that support not only task completion but also well-being, independence, and dignity. In this way, the present work represents an initial step toward a more holistic, human-centered approach to socially assistive robotics for aging populations.

References

[1] World Health Organization, "Ageing and health," 2024, accessed: Feb. 7, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

[2] National Research Council Canada, "Aging in Place: Technology and Innovation – Program plan." 2024, accessed: Mar. 27, 2025. [Online]. Available: https://nrc.canada.ca/en/research-development/research-collaboration/programs/aging-place-technology-innovation-proposed-program-plan

[3] Canadian Institute for Health Information , "1 in 9 new long-term care residents potentially could have been cared for at home," 2020, accessed: Mar. 27, 2025. [Online]. Available: https://www.cihi.ca/en/1-in-9-new-long-term-care-residents-potentially-could-have-been-cared-for-at-home

[4] T. M. Kwok, J. Li, and Y. Hu, "Leveraging GCN-based Action Recognition for Teleoperation in Daily Activity Assistance," Apr. 2025.

[5] H. Lyu, G. Yang, H. Zhou, X. Huang, H. Yang, and Z. Pang, "Teleoperation of Collaborative Robot for Remote Dementia Care in Home Environments," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 8, pp. 1–10, 2020.

[6] I. H. Robertson and P. M. Dockree, "Aging and Attention," in *Cognitive Changes and the Aging Brain*, K. M. Heilman and S. E. Nadeau, Eds. Cambridge: Cambridge University Press, 2019, pp. 140–152.

[7] E. Commodari and M. Guarnera, "Attention and aging," *Aging Clinical and Experimental Research*, vol. 20, no. 6, pp. 578–584, Dec. 2008.

[8] J. Veríssimo, P. Verhaeghen, N. Goldman, M. Weinstein, and M. T. Ullman, "Evidence that ageing yields improvements as well as declines across attention and executive functions," *Nature Human Behaviour*, vol. 6, no. 1, pp. 97–110, Jan. 2022.

[9] A. Vallesi, V. Tronelli, F. Lomi, and R. Pezzetta, "Age differences in sustained attention tasks: A meta-analysis," *Psychonomic Bulletin & Review*, vol. 28, no. 6, p. 1755, Mar. 2021.

[10] C. Colombatto and B. J. Scholl, "Attending to attention: Reverse correlation reveals subtle cues to attentiveness in others' faces," *Psychonomic Bulletin & Review*, vol. 32, no. 6, pp. 3176–3184, 2025.

[11] H.-R. Chen, "Assessment of Learners' Attention to E-Learning by Monitoring Facial Expressions for Computer Network Courses," *Journal of Educational Computing Research*, vol. 47, no. 4, pp. 371–385, Dec. 2012.

[12] J. Bhanushali, A. S. John, and M. Muniyadi, "Attention Score: Objective Measure of Attentiveness in Immersive Omnidirectional Videos," in *2022 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, Dec. 2022, pp. 163–170, iSSN: 2771-7453.

[13] M. Elbawab and R. Henriques, "Machine Learning applied to student attentiveness detection: Using emotional and non-emotional measures," *Education and Information Technologies*, vol. 28, no. 12, pp. 15 717–15 737, Dec. 2023.

[14] S. Zhao, G. Bury, A. Milne, and M. Chait, "Pupillometry as an Objective Measure of Sustained Attention in Young and Older Listeners," *Trends in Hearing*, vol. 23, p. 2331216519887815, Jan. 2019.

[15] M. Fölster, U. Hess, and K. Werheid, "Facial age affects emotional expression decoding," *Frontiers in Psychology*, vol. 5, Feb. 2014.

[16] S. A. Grainger and J. D. Henry, "Gaze patterns to emotional faces throughout the adult lifespan," *Psychology and Aging*, vol. 35, no. 7, pp. 981–992, 2020.

[17] R. Russell, J. R. Sweda, A. Porcheron, and E. Mauger, "Sclera color changes with age and is a cue for perceiving age, health, and beauty," *Psychology and Aging*, vol. 29, no. 3, pp. 626–635, 2014.

[18] L. Zapparoli, M. Mariano, and E. Paulesu, "How the motor system copes with aging: a quantitative meta-analysis of the effect of aging on motor function control," *Communications Biology*, vol. 5, no. 1, pp. 1–15, Jan. 2022.

[19] I. H. Robertson, T. Manly, J. Andrade, B. T. Baddeley, and J. Yiend, "'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects," *Neuropsychologia*, vol. 35, no. 6, pp. 747–758, May 1997. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0028393297000158

[20] T. Baltrusaitis, M. Mahmoud, and P. Robinson, "Cross-dataset learning and person-specific normalisation for automatic Action Unit detection," in *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*. Ljubljana: IEEE, May 2015, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/7284869/

[21] T. Baltrusaitis, A. Zadeh, Y. C. Lim, and L.-P. Morency, "OpenFace 2.0: Facial Behavior Analysis Toolkit," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. Xi'an: IEEE, May 2018, pp. 59–66. [Online]. Available: https://ieeexplore.ieee.org/document/8373812/