Personalized Caring: Integrating EEG/Visual Analysis With ChatGPT for MCI Assistance

Simon L Liu
Shanghai High School International Division
Shanghai, China
simonleo.liu@gmail.com

***Abstract*—This paper presents an AI-driven assistance scheme for individuals with mild cognitive impairment (MCI). The scheme combines electroencephalogram (EEG) sensing to monitor brain activity with a visual sensing module that captures environmental images. The analyzed EEG and visual data are fused to trigger personalized, context-aware conversations through cloud-based AI services such as ChatGPT. Text-to-speech and voice recognition technologies ensure smooth communication, while voice style transfer simulates familiar accents to improve comfort for MCI patients. The effectiveness of the proposal is validated by a prototype and simulations that incorporate EEG/visual data analysis, and personalized AI chatting.**

Keywords*—*MCI; EEG; visual perception; AI assistance

CCS CONCEPTS

•Human-centered→ computing→ Interaction paradigms

I. Introduction

THE growing elderly demands innovative solutions, particularly for individuals with mild cognitive impairment (MCI). MCI patients often encounter difficulties related to cognition and memory, and they also require companion care resources [1]. Human-machine collaboration, particularly with advancements in AI, has gained significant attention. Machines now assist the elderly by supporting daily tasks and acting as companions [2]. AI-driven robots and technologies, such as RI-MAN [3] and Mini [4], have been introduced to support elderly individuals, providing both physical and emotional assistance. Companion robots, which help reduce loneliness and support daily activities, have shown promise in enhancing the well-being of the elderly [5]. Zhou et al. [6] demonstrated how robots can assist dementia patients with communication, while Saunders et al. [7] proposed user-personalized systems that allow elderly individuals to teach robots tasks specific to their needs.

Despite these advancements, existing solutions are still limited in providing AI-based care tailored to the user's real-time situation, especially in the context of MCI care. New solutions must predict the needs of MCI patients, recognize the surrounding environment, and interact with them effectively. Thus, integrating technologies such as environmental perception and intention recognition is essential for creating efficient MCI care services. Recent progress in AI-driven embodied intelligence, combined with new sensor technologies like EEG and visual sensing, opens up new possibilities for MCI care.

In this paper, we propose a personalized MCI assistance scheme: FV-Care, which integrates visual sensing, EEG signal detection, and AI-driven family voice interaction to offer a more interactive and tailored experience. Specifically, we introduce a novel method for analyzing EEG signals to assess cognitive states, enabling real-time monitoring of MCI patients' conditions. By recognizing these states in real-time, service robots can provide better human-robot interaction. For instance, if FV-Care detects that an MCI patient is experiencing difficulties understanding an item in front of them, the robot can provide gentle, step-by-step explanations to help the patient better comprehend the target via AI-based chatting.

II. Our Proposal

As shown in Figure 1, the proposed intelligent assistance scheme (FV-Care) consists of several key modules. The Visual Sensing Module captures images from the user's environment, enabling context-aware assistance. The Wearable EEG Signal Sensing Module monitors brain activity to track cognitive states, such as confusion or memory issues. The Contents Detection/Recognition module adopts content analysis and face recognition methods to convert the environment images into text descriptions, which are then used to guide AI chatting (e.g., ChatGPT), so that the ChatGPT could understand the user's surroundings. The EEG Data Analysis module processes brain activity to assess the user's cognitive state. Both the environmental and user's brain activity are sent to ChatGPT to adjust the conversation accordingly.

For user comfort, the system combines visual and EEG sensing into a single wearable device that pairs with a smartphone. Voice interaction is supported via earphones, with Text-to-Speech (TTS) converting ChatGPT's text responses into voice. Additionally, a voice style transfer method is adopted to modify the accent to one that the MCI patient is familiar with, helping them form an emotional connection.

A. Visual Sensing Module and Image to Environment Description

1) Cloud-based AI for Environment Description: The Visual Sensing Module captures environmental images in front of the user and transmits them to the cloud via the smartphone connecting the cloud. Advanced AI models, such as OpenAI's GPT-4 Vision model [8], process these images and generate textual descriptions.

2) Smartphone-based Face Recognition: We adopt a lightweight face recognition model running on the smartphone [9]. Users can create a familiar face library, allowing the smartphone to recognize them locally, ensuring privacy. The

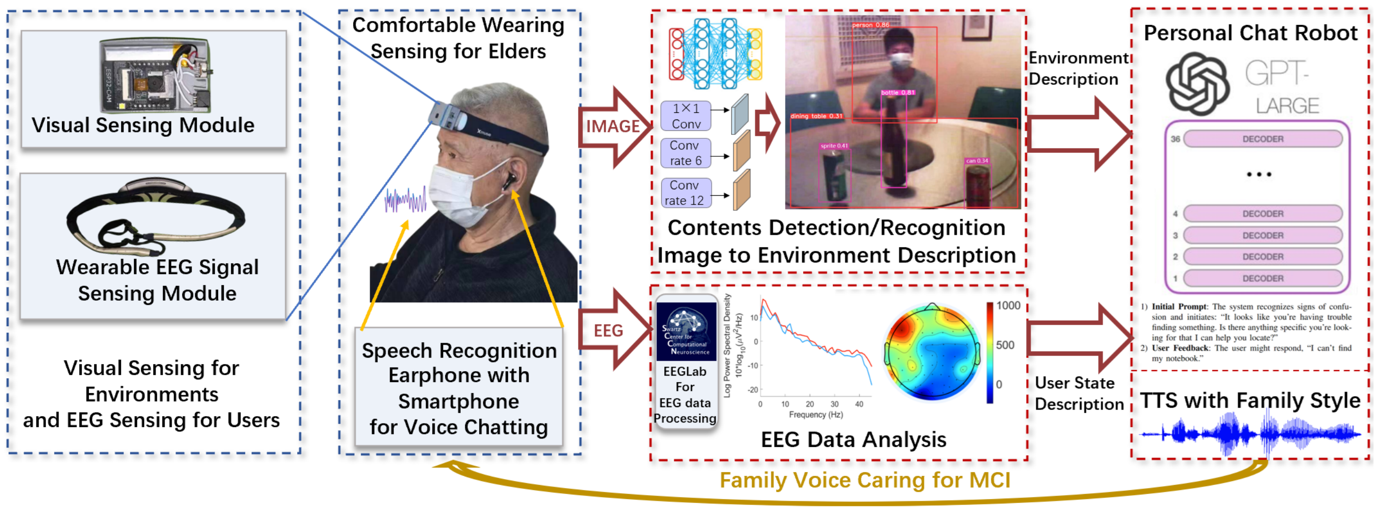


Fig. 1. The Scheme of Personalized Caring for MCI Caring via AI Chatting

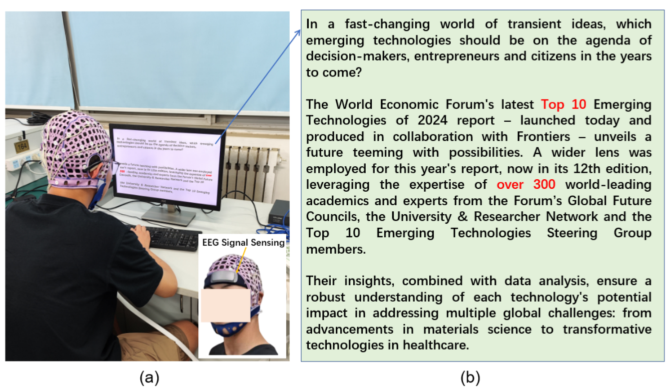


Fig. 2. Visual evoked potentials EEG experiments for cognitive activity analysis

environmental images are captured and analyzed by the cloud and smartphone, and the results fused to control personalized AI-based conversations, thereby enhancing interaction with MCI patients.

B. Wearable EEG Signal Sensing and Analysis

EEG is widely used for analyzing brain activity. Accurate decoding of EEG signals is essential for real-time identification of user brain states. The raw EEG signals are preprocessed to remove noise and artifacts, which includes filtering, downsampling, and baseline correction. Then Independent Component Analysis (ICA) is applied to separate neural signals from artifacts like eye movements [10]. After ICA, the independent components are analyzed in the frequency domain. Power Spectral Density (PSD) is computed for key brainwave bands: delta (δ: 0.5-4Hz), theta (θ: 4-8Hz), alpha (α: 8-14Hz), beta (β: 14-30Hz), and gamma (γ: 30-45Hz). The power of each frequency band provides insights into cognitive and emotional states, as different bands relate to specific neural activities.

1) Cognitive Analysis with EEG: The FV-Care scheme uses the EEG sensing module of Ref. [11], which records data from four EEG channels (AF7, AF8, TP9, TP10) at 256 Hz. Based on [12], we hypothesized that EEG patterns of healthy individuals can be similar to those of MCI patients under similar conditions. Therefore, 32 healthy volunteers participated in our study.

Cognitive state is often associated with specific EEG frequency band, particularly α and θ waves, which are linked to various cognitive processes. For instance, α waves are commonly associated with states of relaxed alertness and working memory, while θ waves are often observed during focused attention and deep concentration. To investigate these cognitive states, we need to set up a specially designed experiment to collect EEG data and extract features of different brain activity patterns.

First, we collected resting-state EEG data from the 32 volunteers, asking them to stay relaxed. The resting-state EEG will serve as a baseline for further analysis. Then, the volunteers are invited to complete a carefully designed experiment. During the experiment, participants were shown a news video (Figure 3), which was chosen to engage their cognitive processes. After watching the video, participants were asked to complete a related questionnaire with specific numbers from the news content. EEG data collected during the experiments were considered valid if the participant successfully answered the questions.

We then PSD of valid samples for the α and θ bands to assess relative power changes in cognitive state. According to our analysis, both α and θ waves are highly related to cognitive state, specifically, high θ activity always indicate complex cognitive tasks. Based on this, we then defined a simple method to calculate the Cognitive Load Index (CLI) as follows:

$$CLI = \frac{P_\theta}{P_\alpha} \tag{1}$$

where P_θ and P_α represent the average power in the θ and α of the four bands (AF7, AF8, TP9, TP10).

Using the CLI, we can calculate different indices for users in different brain states. For the experiments we completed, only valid participant data were considered, the mean value of the resting-state was:

$$CLI_r = 2.45$$

while the mean value of the cognitive-task was:

$$CLI_c = 3.03$$

Based on the CLI statistics, the ratio between the CLI in the cognitive task and the resting state could be: $\frac{3.03}{2.45} \approx 1.24$. Although the results are derived from normal individuals, the relationship between the resting and cognitive states may also be applicable to MCI patients [12]. In our FV-Care framework, each MCI user will be recorded their resting-state CLI to derive (CLI_r), then continuously calculate the real-time CLI relative to the resting-state (CLI_r), we can assess whether the user is in a recognized cognitive state. For instance, if a participant's real-time CLI exceeds 1.24 times their resting-state CLI_r, they are classified as being in a cognitive state. Although the current method for calculating the CLI is relatively simple, it provides a foundational approach to distinguish between a user's resting and cognitive states. In the future, this method can be further refined and adjusted based on the actual number of users and data.

C. "$4W$" Control Prompts for ChatGPT

ChatGPT's responses can be variable, making it essential to control its output in the FV-Care scheme for MCI patient assistance. The system uses structured prompts in the format *'role: system, content: descriptions"* to ensure responses are relevant to the user's environment and cognitive state [13]. while, the *descriptions* parameter should be used for to the chatting conditions.

We propose to construct descriptions in the "$4W$" prompt format, which consists of five parameters: *Where*, *When*, *Who*, and *What*. The *Where* and *When* are derived from the user's smartphone GPS and time data. The *Who* is obtained from the smartphone's face recognition. The *What* is generated by a cloud AI model to describe the user's environment.

Based on EEG analysis, FV-Care can identify whether the user is in a resting or cognitive state. Once the cognitive state is detected, the FV-Care system creates a tailored prompt for ChatGPT and automatically triggers a chat. For instance, based on the visual and EEG data from Figure 4 (a), the prompt might be: *role: system, content: I am in a room at ddmm, N/S, dddmm, E/W, current time is year-month-day-time, somebody is here in front of me, with some objects... and I am in a cognitive state*. This structured prompt ensures that ChatGPT's responses are contextually relevant and personalized for the user.

D. Chatting Improvement with Family Voice

MCI patients often need assistance from voice assistants, similar to family support. In our FV-Care scheme, we utilize cloud-based models for speech-to-text and text-to-speech conversions, using OpenAI models for efficient text generation [14] and speech recognition [15]. To address privacy concerns,

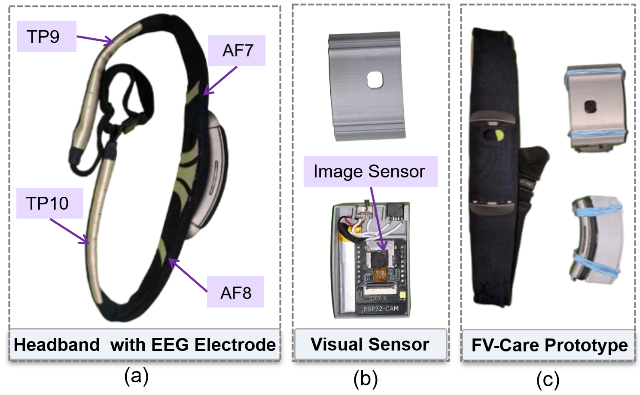


Fig. 3. (a). EEG module, (b).Visual sensing module, (c).Wearable FV-Care prototype

we deploy a voice style transfer model [16] on the user's Android device, ensuring secure voice data handling. The family member's voiceprint is locally collected and used for personalizing the voice style. ChatGPT's text responses are converted to speech via the cloud-based text-to-speech model, then processed through the smartphone's voice style transfer model, ensuring secure and efficient communication for MCI patients.

III. Simulation for Our Proposals

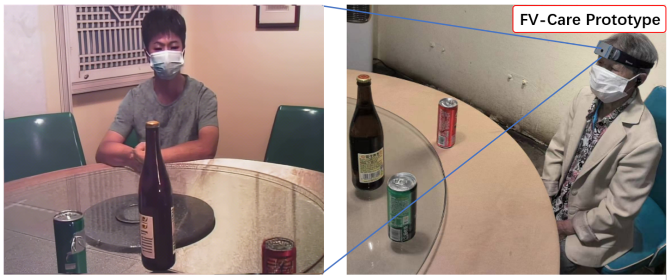
A. Participant Selection

We conducted participant selection following ethical standards, screening participants for cognitive function using the Montreal Cognitive Assessment (MoCA). The MoCA evaluates cognitive domains such as attention, memory, and executive functions, and is highly effective for identifying MCI. A total of 30 volunteers were screened, with a mean age of 66.15 years (±8.12 years). The sample was predominantly female (63.3%, n=19). The mean MoCA score was 21.23 (±6.16). For further experiments, we selected participants with MoCA scores between 20 and 23, as they were able to effectively engage in communication and follow instructions.

B. Experimental Equipment and Workflow

We designed an FV-Care prototype with a 4-channel EEG sensor, a mini-camera module capable of capturing 640×480 images, and a rechargeable power supply. The prototype is shown in Figure 3. The workflow of the FV-Care scheme is as follows:

1) EEG Monitoring: The system continuously monitors the user's brain activity using EEG sensing. Based on EEG data, the system identifies whether the user is in cognitive state.
2) Visual Sensing Activation: Once a cognitive state is detected, the visual sensing module is triggered to capture environmental images.
3) Image Analysis: The captured images are analyzed using the cloud-based GPT-4 Vision model and smartphone-based familiar face recognition.



(a) (b)

Fig. 4. (a) FV-Care prototype sensed image, (b) Volunteer with FV-Care prototype

4) Prompt Construction: The environmental description and the user's brain state are combined to form "$4W$" prompts.
5) Conversation Initialization: The prompts are sent to ChatGPT to initiate an interactive conversation with the user.
6) TTS Conversion with Familiar Style: The text generated from the conversation is processed by the cloud-based TTS model for speech synthesis, then transmitted to the user via the smartphone based familiar voice style transfer model.

C. Simulation with Case Study: Assisting in Recognizing the Difference between Drinks

In this section, we demonstrate the functionality of FV-Care through a case study. The simulation setup, shown in Figure 4, involved a round table with three beverages (alcoholic drink, Coke, and Sprite). Two participants were involved: one wearing the FV-Care prototype and the other sitting across from them. This setup was designed to test the system's ability to help users differentiate between the beverages and make decisions.

In case of Figure 4, the volunteer is tasked with recognizing three beverages placed in front of her: a can of Coke, a can of Sprite, and a bottle of beer. The person named Simon who is familiar with the volunteer will ask her to choose a beverage. Using the working flow presented in previous Section, the "$4W$" prompt used to initiate ChatGPT for the environment (Figure 4) could be: *role: system, content: "I am in a room with GPS ddmm,N/S,dddmm,E/W, current time year-month-day-time, Simon is here, in front of me, a man, indoors, sitting, wearing a face mask, table etc., and the user is in cognition state"*. The FV-Care scheme then provides relevant information on each drink, including taste, alcohol content, and flavor profiles, and initiates a conversation to assist the user in making an informed decision.

1) **System:** *I see you're looking at the three drinks in front of you. Would you like some help understanding what they are?*
2) **User:** *Yes, I'm trying to figure out what these drinks are.*
3) **System:** *Let me explain. The can with the red label is coke. It's sweet and contains caffeine. The can with the green label is sprite, a lemon-lime flavored soft drink. The bottle is beer, which has a mild bitterness and contains alcohol.*
4) **User:** *Alcohol? I'm not sure I want that... What about the others?*
5) **System:** *The coke and sprite are both soft drinks. Coke has a richer, caramel-like flavor and more sweetness. Sprite is lighter, with a refreshing lemon-lime taste and no caffeine.*
6) **User:** *I think I prefer something without caffeine. Sprite sounds good.*
7) **System:** *That's a great choice! The green one is light and refreshing, perfect if you're looking for something without caffeine or alcohol. By the way, there's someone nearby—his name is Simon. He's a regular here and always enjoys a cold sprite, just like you. Have you met him before?*
8) **User:** *Simon? I think I've seen him around. He usually sits near the window, right?*
9) **System:** *Exactly! That's Simon. He often comes in for a quiet drink and some time to relax. It looks like you've made a choice he'd approve of. Enjoy your drink!*

D. Subjective Experience and Usability Evaluation

To assess the comfort and usability of our FV-Care prototype, we invited over 20 participants with MoCA scores ranging from 18 to 25. Participants wore the FV-Care prototype during the experiment, which offered real-time navigation and environmental interaction functionalities. A subjective evaluation form was used to gather feedback on various aspects. The evaluation results were highly positive. On a 1-5 scale, the average score for wearing comfort was 4.7, indicating that most participants found the prototype comfortable to wear. The ease of use was rated 4.8 on average, suggesting that the system was easy to operate. Most participants found the system intuitive, comfortable, and beneficial, and expressed a desire for continued use.

IV. Future Work and Conclusion

This paper presents FV-Care, an intelligent assistance scheme for individuals with MCI, integrating EEG signal monitoring, visual recognition, and cloud-based dialogue for real-time, context-aware support. Future work will focus on training domain-specific dialogue models for MCI care, and developing adaptive EEG analysis methods to improve the detection of MCI-related neural activities. We demonstrated the effectiveness of the FV-Care prototype, which uses EEG-based cognitive and memory state detection along with ChatGPT control via "$4W$" prompts. While the current implementation serves as a proof-of-concept, further developments will include mobile integration, improved signal processing, and custom hardware for real-world applications.

References

[1] F. H. Frech *et al.*, "Economic impact of progression from mild cognitive impairment to alzheimer disease in the united states," *Journal of Prevention of Alzheimer's Disease*, vol. 11, no. 6, pp. 983–991, Nov 2024.

[2] S. Li *et al.*, "Proactive human–robot collaboration: Mutual-cognitive, predictable, and self-organising perspectives," *Robotics and Computer-Integrated Manufacturing*, vol. 81, p. 102510, Jan 2023.

[3] M. Toshiharu *et al.*, "Development of the tactile sensor system of a human-interactive robot 'ri-man'," *IEEE Transactions on Robotics*, vol. 24, no. 2, pp. 505–512, Apr 2008.

[4] M. A. Salichs *et al.*, "Mini: A new social robot for the elderly," *International Journal of Social Robotics*, vol. 12, no. 6, pp. 1231–1249, Dec 2020.

[5] J. M. R. . J. K. Clara Berridge 1, Yuanjin Zhou 2, "Companion robots to mitigate loneliness among older adults: Perceptions of benefit and possible deception," *Front. Psychol*, vol. 14, Feb 2023.

[6] D. Zhou, E. I. Barakova, P. An, and M. Rauterberg, "Assistant robot enhances the perceived communication quality of people with dementia: A proof of concept," *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 3, 2022.

[7] J. Saunders, D. S. Syrdal, K. L. Koay, N. Burke, and K. Dautenhahn, ""teach me–show me"—end-user personalization of a smart home and companion robot," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 1, 2016.

[8] Jiaye, Fok, and GPT4, "Gpt4v-image-captioner," https://github.com/jiayev/GPT4V-Image-Captioner, 2023, [Accessed 11-Nov-2024].

[9] S. Panchal, "Face recognition with facenet on android," https://github.com/shubham0204/FaceRecognition_With_FaceNet_Android, 2022, [Accessed 24-Nov-2024].

[10] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," *NIPS'95: Proceedings of the 9th International Conference on Neural Information Processing Systems*, 1995.

[11] Muse by Interaxon Inc., "Muse s: The brain sensing headband," *Muse Official Website*, 2024, available at: https://choosemuse.com/products/muse-s-gen-2 [Accessed 24-Aug-2024].

[12] J. Schumacher, J.-P. Taylor, C. A. Hamilton, M. Firbank, R. A. Cromarty, P. C. Donaghy, G. Roberts, L. Allan, J. Lloyd, R. Durcan *et al.*, "Quantitative eeg as a biomarker in mild cognitive impairment with lewy bodies," *Alzheimer's Research & Therapy*, vol. 12, no. 1, pp. 1–12, 2020.

[13] OpenAI, "Openai api documentation," https://platform.openai.com/docs, [Accessed 29-Nov-2024], 2024.

[14] OpenAI., "Openai text-to-speech guide," https://platform.openai.com/, [Accessed 29-Nov-2024], 2024.

[15] OpenAI, "Openai speech-to-text guide," https://platform.openai.com/, [Accessed 29-Nov-2024], 2024.

[16] RVC-Boss, "Vioce-tranfer," https://github.com/RVC-Boss/GPT-SoVITS , 2024, [Accessed 11-Nov-2024].