ISSN 2616-5775 Vol. 9, Issue 5: 36-47, DOI: 10.25236/AJCIS.2026.090505

EEG- and Language Generation-Based Emotional Arousal Assistance System for Alzheimer's Disease Patients

Zimo Xu

YK Pao School Songjiang Campus, Shanghai, China
s21309@ykpaoschool.cn

Abstract: *Patients with Alzheimer's disease (AD) frequently present both cognitive decline and emotional dysregulation. Pharmacological treatment may reduce symptom progression, but side effects and long-term cost remain concerns, while traditional non-pharmacological interventions are often limited in personalization and objective evaluation. To address this gap, this paper proposes an emotional-arousal assistance system that combines electroencephalography (EEG) and large language models. The system performs real-time emotion recognition from EEG, generates simplified personalized poetry from autobiographical memory cues, and delivers coordinated multimodal stimulation with soothing music. Experiments on 2,132 EEG samples with six deep-learning models show that LSTM performs best (Accuracy 91.56%, Precision 92.9%, Recall 91.51%, F1-score 91.42%). Pilot testing indicates improved emotional arousal experience. The study provides a practical path toward personalized and quantifiable non-pharmacological intervention for AD.*

Keywords: *Alzheimer's disease; electroencephalography (EEG); emotion classification; deep learning; poetry generation; non-pharmacological treatment*

1. Introduction

1.1 Background and Motivation

Alzheimer's disease (AD) is one of the most common neurodegenerative disorders and is characterized by memory impairment, cognitive decline, and emotional disturbance. With accelerating population aging, the number of patients continues to increase, making AD a major global public-health challeng[1]. Figure 1 shows the trend.

Current pharmacological treatment can delay symptoms to some extent, but it cannot reverse disease progression and may bring side effects and financial burden[2][3]. Non-pharmacological approaches (e.g., music therapy) are generally safer, but they often lack personalization, objective efficacy assessment, and timely feedback adjustment[4]. Therefore, a closed-loop framework integrating real-time monitoring, emotion recognition, content generation, and intervention feedback is highly valuable.In addition, recent medical AI literature indicates that large language models can provide clinically relevant language support and personalization capability, which motivates their integration into non-pharmacological AD intervention workflows[7].

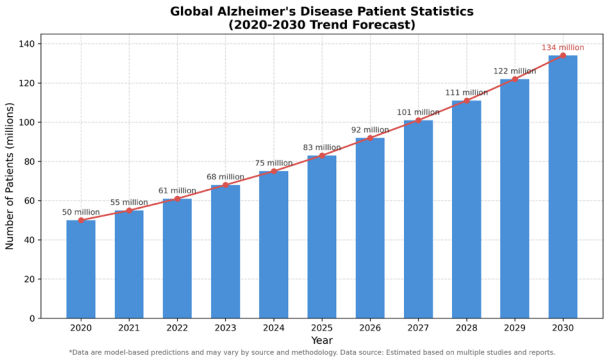

Fig. 1. Global prevalence of Alzheimer's disease (2020--2030 forecast).

ISSN 2616-5775 Vol. 9, Issue 5: 36-47, DOI: 10.25236/AJCIS.2026.090505

1.2 Objectives and Contributions

The goal of this work is to build a personalized non-pharmacological intervention system for AD and realize a closed loop between emotion recognition and content intervention. Specifically, we: (1) compare multiple deep-learning models and identify the optimal EEG emotion-classification model; (2) generate simplified poetry aligned with patient experiences using an LLM; (3) combine poetry and music to provide multimodal emotional stimulation; and (4) implement and preliminarily validate an end-to-end prototype.

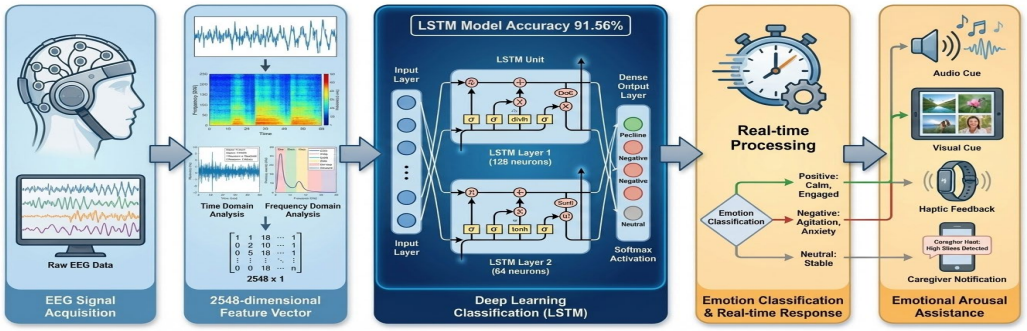

Fig. 2. Overview of the proposed framework.

This paper makes three main contributions: (1) a unified intervention framework combining EEG monitoring and generative content; (2) experimental validation that LSTM is effective for this task; and (3) a working prototype covering the full pipeline from signal acquisition to intervention presentation, providing an engineering basis for future clinical validation. Figure 2 shows the framework.

2. Related Work

2.1 Emotion Recognition Techniques

Emotion recognition is a core topic in affective computing. Common signal sources include facial expression, speech, ECG, and EEG, each with different trade-offs in accuracy, robustness, and deployability.

1) Facial-Expression-Based Emotion Recognition: Facial-expression methods infer emotion from visual features, and the FACS system provides a standardized representation[8]. Deep-learning approaches, including CNN and attention-based models, have improved performance on public datasets[9][10]. However, these methods remain sensitive to illumination, occlusion, head pose, and inter-individual variation. Because facial expressions can also be intentionally controlled, this modality is limited in medical scenarios requiring objective and continuous monitoring.

2) Other Physiological Emotion Recognition Methods: Speech-based methods rely on prosodic and spectral features but are strongly affected by language and individual differences[11]. ECG-based methods are objective but vulnerable to motion and environmental noise[12]. GSR reflects arousal intensity but has limited capability for emotion-valence discrimination[13]. In contrast, EEG directly captures neural activity, offers high temporal resolution, is difficult to fake, and supports both valence and arousal analysis, making it more suitable for the real-time closed-loop setting in this study. This choice is further supported by prior EEG-emotion studies reporting consistent discriminative patterns across affective states and strong feasibility of deep-learning-based recognition in practical settings[5][6].

2.2 Deep-Learning-Based EEG Emotion Recognition

Deep learning has significantly improved EEG emotion recognition. Different architectures provide different trade-offs among local feature extraction, temporal modeling, and parameter efficiency.

1) Convolutional Neural Network (CNN): Convolutional neural networks (CNNs) are effective at extracting local patterns. A typical convolution is:

$$yy_{i,j}=f\left(\sum_{m}\sum_{n} w_{m,n}\cdot x_{i+m,j+n}+b\right) \quad (1)$$

where x is the input, w is the kernel, b is the bias, and f is the activation function. CNNs can effectively

ISSN 2616-5775 Vol. 9, Issue 5: 36-47, DOI: 10.25236/AJCIS.2026.090505

extract local time-frequency features for EEG and have shown promising results on public datasets[15]. Their main limitation is weaker long-range temporal modeling, and pooling may lose temporal detail.

2) RNN and Its Variants: Recurrent neural networks (RNNs) model sequence dependency through hidden-state transitions:

$$h_t = f(W_{hh}h_{t-1} + W_{xh}x_t + b_h) \tag{2}$$

$$y_t = W_{hy}h_t + b_y \tag{3}$$

Standard RNNs suffer from vanishing/exploding gradients, so practical studies more commonly use LSTM and GRU.

Long short-term memory (LSTM) enhances long-term memory using gates and a cell state:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{4}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{5}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{6}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{7}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{8}$$

$$h_t = o_t \odot \tanh(C_t) \tag{9}$$

LSTM is generally stable in EEG sequence tasks, but it has higher parameter count and training cost.

Gated recurrent unit (GRU) is a lightweight variant of LSTM[14], with core computations:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \tag{10}$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \tag{11}$$

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \odot h_{t-1}, x_t] + b_h) \tag{12}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{13}$$

GRU has fewer parameters and faster training, which is advantageous in resource-constrained settings, though it can be slightly weaker than LSTM on complex long-term dependencies.

*3) Multilayer Perceptron (MLP):*The multilayer perceptron (MLP) is a common feed-forward baseline model with:

$$h^{(l)} = f(W^{(l)}h^{(l-1)} + b^{(l)}) \tag{14}$$

MLP is simple and efficient to train, but it does not explicitly model temporal dependency and is typically used as a baseline. Existing evidence indicates that LSTM/GRU are generally more suitable for EEG emotion-recognition tasks, which motivates our model-selection design.

3. Methodology

3.1 Overall System Architecture

The proposed emotional-arousal assistance system adopts an end-to-end modular architecture with five components: EEG acquisition, preprocessing and feature extraction, emotion classification, personalized poetry generation, and multimodal presentation. The workflow follows a ``perception-cognition-generation-presentation'' pipeline, forming a closed loop from physiological input to intervention output. Figure 3 shows module interactions and data flow.

The operational sequence is as follows: multi-channel EEG is sampled at 250 Hz and streamed in real time; signals are filtered, artifact-corrected, and baseline-adjusted to produce a 2,548-dimensional feature vector; an LSTM model predicts positive/negative/neutral states; DeepSeek then generates simplified personalized poetry from autobiographical cues; finally, text and music are delivered jointly as multimodal stimulation. The design emphasizes (1) real-time and accurate inference (about 15 ms per inference, 91.56% accuracy), (2) personalization, (3) multimodal synergy, and (4) non-invasive safety.

ISSN 2616-5775 Vol. 9, Issue 5: 36-47, DOI: 10.25236/AJCIS.2026.090505

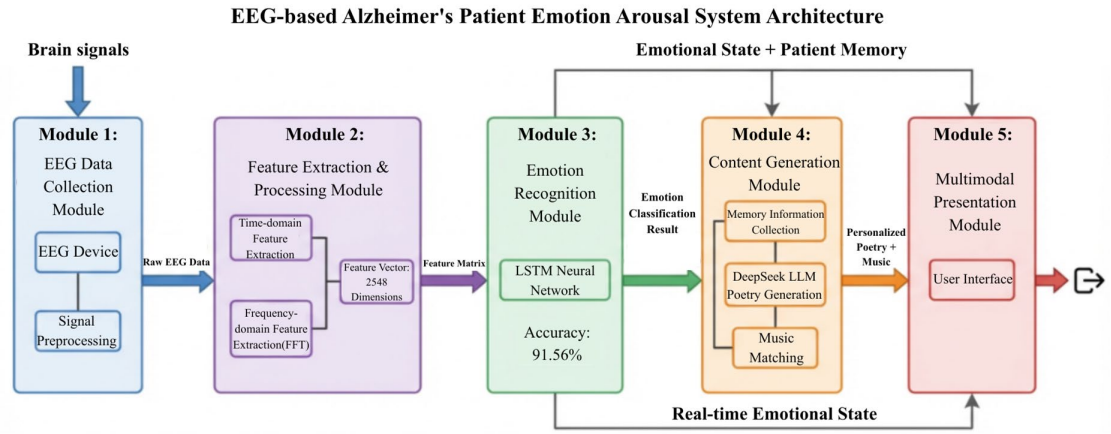

Fig. 3. Overall system architecture

3.2 EEG Signal Acquisition

1) Electrodes and Chip Selection: The acquisition setup uses Ag/AgCl wet electrodes and the ADS1299 chip. Ag/AgCl electrodes can reliably capture weak EEG signals (typically 10--100 μV), with low contact impedance (usually < 5kΩ) and high signal-to-noise ratio.

Electrodes follow the international 10-20 system. The 14 channels are Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, and T8. The reference electrode is placed at the right earlobe or right mastoid, and ground is placed at Fpz.

ADS1299 is a 24-bit, 8-channel synchronous-sampling chip that supports 250 Hz sampling, programmable gain (1--24), low-noise input (about 1 μV RMS), and lead-off detection, making it suitable for portable biosignal acquisition.

2) EEG Acquisition Procedure: The acquisition procedure includes four steps. First, clean the scalp and apply conductive gel to reduce electrode-skin impedance and improve signal quality.

Second, place electrodes according to the 10-20 system and check impedance. Each channel should remain below 5 kΩ; otherwise, the site is cleaned again and gel is reapplied.

Third, configure acquisition parameters: 250 Hz sampling rate, gain 24, 0.5--45 Hz band-pass filtering, and lead-off monitoring.

Finally, start 14-channel continuous acquisition with real-time transmission and storage. Waveforms are displayed for online quality control and quick handling of issues such as loose electrodes or EMG interference.

3) Selection Rationale: This configuration is selected for four reasons: high signal quality, device reliability, controllable cost, and portability. Its main limitation is wet-electrode dependence on conductive gel, which increases preparation effort and may reduce long-term wearing comfort. Future work will explore dry or semi-dry electrodes for better home-use experience.

EEG Emotion Classification

1) Dataset: The EEG emotion dataset is the public Kaggle dataset “EEG Brainwave Dataset: Feeling Emotions"[18]. It contains recordings from two healthy participants (one male, one female; ages 20 and 23) under emotion-induction video stimuli. Videos are drawn from DEAP and each clip lasts one minute. EEG is collected with the Emotiv EPOC+ wearable device (14 channels: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4) at 128 Hz.

The dataset contains 2,132 EEG records. Each record corresponds to a 3-second window and is labeled as POSITIVE (708, 33.2%), NEGATIVE (708, 33.2%), or NEUTRAL (716, 33.6%). The class distribution is highly balanced.

After preprocessing and feature extraction, each EEG sample is represented as a 2,548-dimensional vector. Table 1 lists partial feature values for one POSITIVE sample.

ISSN 2616-5775 Vol. 9, Issue 5: 36-47, DOI: 10.25236/AJCIS.2026.090505

Table 1 Example EEG feature vector (partial features).

Feature Type	Feature Name	Value
Time-domain	AF3 channel mean	4523.7
Time-domain	AF3 channel standard deviation	1842.3
Frequency-domain	AF3 alpha-band power	2.847
Frequency-domain	AF3 beta-band power	1.563
Spatial-domain	AF3-F7 correlation coefficient	0.742
Information-theoretic	AF3 sample entropy	1.234

The data are randomly split into 70% training and 30% testing sets while preserving class distribution. The split is: training 1,493 (POSITIVE 496, NEGATIVE 496, NEUTRAL 501) and testing 639 (POSITIVE 212, NEGATIVE 212, NEUTRAL 215). A fixed seed (42) is used for reproducibility.

2) Model Training： We compare six deep-learning models for EEG emotion classification: LSTM, GRU, CNN-LSTM, MLP, MLP with Dropout, and pure CNN. All models are trained under unified settings for fair comparison.

Training is conducted on a workstation with an NVIDIA GeForce RTX 3090 GPU (24 GB VRAM). The software stack includes Python 3.8, PyTorch 1.12, CUDA 11.3, and MNE-Python 0.24 for EEG preprocessing.

Unified hyperparameters are used: batch size 32, up to 100 epochs with early stopping (patience 10), initial learning rate 0.001, Adam optimizer, and cross-entropy loss for multi-class classification. Evaluation metrics include Accuracy, Precision, Recall, and F1-score.

For LSTM, the architecture is: input layer with 2,548 features; first bidirectional LSTM layer with 128 hidden units; second bidirectional LSTM layer with 64 hidden units; dropout (0.5); fully connected layer mapping 128 features (64x2) to three classes; and Softmax output.

The approximate parameter counts are: LSTM 1.58M, GRU 1.19M, CNN-LSTM 2.07M, and MLP 3.28M.

During training, validation loss and accuracy are computed each epoch (20% of the training set as validation). Early stopping stores the best checkpoint by validation loss. Final performance is reported on the held-out test set.

3.3 Personalized Poetry Generation

1) Large Language Model Principle: This work uses the DeepSeek LLM for personalized poetry generation. LLMs are Transformer-based pretrained models that learn language knowledge and reasoning ability through large-scale self-supervised training[16].

The core is self-attention:

$$\text{Attention}(Q,K,V)=\text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{15}$$

where Q, K, and V are Query, Key, and Value matrices obtained via linear projection of the input sequence:

$$Q = XW_Q, \quad K = XW_K, \quad V = XW_V \tag{16}$$

d_k is the key dimension, and division by $\sqrt{d_k}$ stabilizes gradients for Softmax.

DeepSeek uses a Transformer-decoder architecture with multiple self-attention and feed-forward layers (e.g., 67 layers). With 67B parameters pretrained on trillions of tokens, it provides strong language understanding and generation. In autoregressive generation, the next-token distribution is:

$$P(x_t|x_1,x_2,\ldots,x_{t-1})=\text{softmax}(W_o h_t) \tag{17}$$

where $h_t$is the hidden state at step t, and $W_o$is the output projection matrix.

To balance generation quality and diversity, the model uses Top-p (nucleus) sampling[17]. Tokens are sorted by probability, a nucleus is formed up to cumulative probability p (e.g., 0.9), and sampling is performed within that set instead of greedy decoding.

2) Prompt Design: Prompt design is central to controlled LLM generation. We use a two-part structure:

ISSN 2616-5775 Vol. 9, Issue 5: 36-47, DOI: 10.25236/AJCIS.2026.090505

System Prompt and User Input.

System Prompt defines role and task constraints:

You are a professional poetry writer specialized in therapeutic poems for Alzheimer's disease patients. Your task is to create concise, warm, and positive poems from each patient's life stories to help evoke positive memory and stabilize emotions.

Writing constraints:

1. Simplified language: use common words and simple syntax; keep each line short and easy to read.

2. Emotional orientation: maintain a warm, positive, and calm tone; avoid negative wording.

3. Personalization: include specific experiences, places, people, and events.

4. Memory activation: use concrete scene details and sensory cues (visual/auditory/olfactory/tactile).

5. Moderate length: keep the poem concise so patients can complete reading and comprehension.

User Input contains patient-specific information collected via structured questions:

Family memory: recall a warm moment with family (e.g., holiday reunion, family trip, or daily happiness).

Life experience: describe a memorable trip, including places, scenes, and feelings.

Interests: favorite activities or hobbies (e.g., gardening, chess, opera, calligraphy).

Emotional memory: who influenced your life most? Which period do you miss most?

Current emotion: current state from EEG classification (positive/negative/neutral).

An **example prompt** is shown below:

System Prompt: (as above)

User Input:

Patient name: Ms. Li

Age: 72

Family memory: Every Spring Festival, the whole family makes dumplings together while grandchildren play nearby.

Life experience: She once visited West Lake with her husband, took a boat ride, and enjoyed lotus flowers.

Interests: She likes gardening, especially jasmine and roses, and enjoys morning watering.

Emotional memory: Her husband was the most important person in her life, and these memories remain warm.

Current emotion: Negative (EEG result)

Please generate a poem that helps Ms. Li recall positive memories and stabilize her mood.

Example generated poem:

Spring Festival comes, we make dumplings,

Laughter fills the room with children around.

You roll the dough, I fold each piece,

A lively New Year, warm and bright.

By West Lake, lotus fragrance drifts,

The boat sways gently over ripples.

Hand in hand beside clear water,

Warmth still rests in my heart.

ISSN 2616-5775 Vol. 9, Issue 5: 36-47, DOI: 10.25236/AJCIS.2026.090505

Morning watering in the garden,

Jasmine blooms and roses glow.

I remember your smiling face,

And happiness returns inside.

The poem uses simple language and short lines, integrating Ms. Li's family memory (Spring Festival dumplings), life experience (West Lake trip), hobbies (gardening), and emotional memory (time with her husband), with a positive tone to support emotional stabilization and memory recall.

3) Rationale for Collecting Personalized Informatio: The rationale includes three aspects. First, neuroscience indicates episodic memory is highly affected in AD, but specific cues (place, person, emotion) can reactivate residual traces[21]. Personal life stories therefore provide effective prompts for emotional connection and recall.

Second, psychology studies show autobiographical memory is important for identity and emotion regulation; recalling positive experiences can improve mood and reduce depressive or anxious symptoms[22].

Third, personalized medicine emphasizes tailoring interventions to individual differences[23]. Since AD patients have diverse life and cultural backgrounds, personalized content can better maximize engagement and therapeutic relevance.

4. Experimental results

4.1 EEG Signal Feature Analysis

1) EEG Features Under Different Emotional States: We analyzed EEG characteristics under three emotional states (POSITIVE, NEGATIVE, and NEUTRAL). Figure 4 presents raw AF3-channel waveforms (left frontal area) within a 3-second window.

The waveform patterns differ clearly across classes. Under POSITIVE emotion, EEG amplitude is generally larger with stronger fluctuations, indicating more active neural dynamics. Frequency-domain analysis shows increased Beta (13--30 Hz) and Gamma (30--45 Hz) power, consistent with findings linking positive affect and high-frequency activity[19].

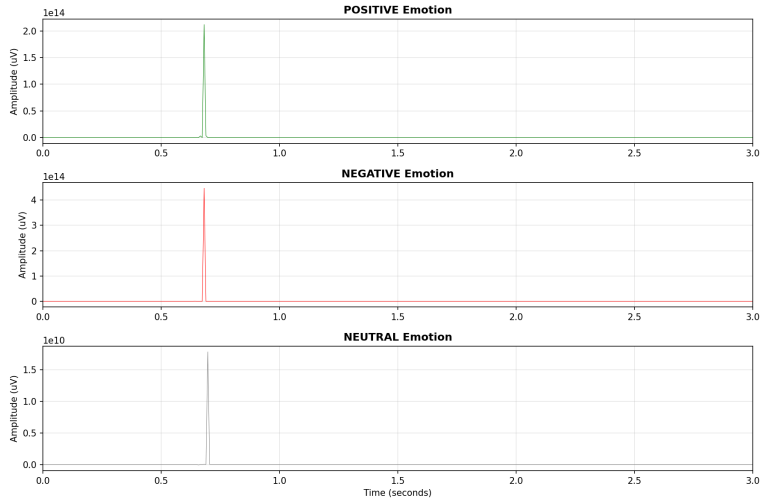

Fig. 4. EEG waveform comparison across three emotional states (AF3 channel)

Under NEGATIVE emotion, EEG amplitude is relatively smaller and waveforms are smoother, while Alpha (8--13 Hz) power is enhanced, especially in right frontal regions. This aligns with frontal-asymmetry theory, which associates left frontal activity with approach/positive affect and right frontal activity with withdrawal/negative affect[20].

Under NEUTRAL emotion, amplitude and spectral characteristics lie between POSITIVE and

ISSN 2616-5775 Vol. 9, Issue 5: 36-47, DOI: 10.25236/AJCIS.2026.090505

NEGATIVE, with more balanced band power.

2) Statistical Analysis of Frequency-Domain Features: Table 2 reports average band power ($\mu V^2/Hz$) under the three emotional states. Values are averaged across channels and show significant differences (ANOVA, $p<0.001$).

Table 2 Comparison of EEG band power across emotional states

Band	POSITIVE	NEGATIVE	NEUTRAL
Delta (0.5-4 Hz)	3.24 ± 0.87	3.18 ± 0.92	3.21 ± 0.85
Theta (4-8 Hz)	2.67 ± 0.64	2.81 ± 0.71	2.73 ± 0.66
Alpha (8-13 Hz)	2.15 ± 0.58	2.84 ± 0.73	2.47 ± 0.62
Beta (13-30 Hz)	1.93 ± 0.52	1.54 ± 0.41	1.71 ± 0.47
Gamma (30-45 Hz)	0.87 ± 0.23	0.62 ± 0.18	0.74 ± 0.20

Beta and Gamma power are highest in POSITIVE and lowest in NEGATIVE states, making them key discriminative features. Alpha power is highest in NEGATIVE and lowest in POSITIVE, supporting frontal-asymmetry findings. Delta and Theta differences are relatively smaller and contribute less to classification.

3) Experimental Setup Photo: Figure 5 shows the EEG acquisition preparation process. Researchers prepare the Emotiv EPOC+ wearable EEG headset, adjust electrode placement for stable contact, and verify channel impedance. Experiments are conducted in a quiet indoor environment to ensure signal quality.

Fig. 5. EEG signal acquisition scene

4.2 EEG Emotion Classification Performance

1) Model Performance Comparison: We compared six deep-learning models for EEG emotion classification. Table 3 summarizes test-set Accuracy, Precision, Recall, and F1-score.

Table 3 Performance comparison of six deep-learning models.

Model	Accuracy	Precision	Recall	F1-score
LSTM	0.9156	0.9290	0.9151	0.9142
GRU	0.8983	0.8957	0.8962	0.8878
CNN-LSTM	0.7172	0.7750	0.7145	0.6549
MLP	0.6953	0.7522	0.6892	0.6158
MLP (Dropout)	0.6188	0.6035	0.6171	0.6055
CNN	0.4344	0.6996	0.4298	0.3291

LSTM achieved the best results on all metrics (Accuracy 91.56%, Precision 92.9%, Recall 91.51%, F1-score 91.42%). GRU ranked second at 89.83% accuracy, only 1.73 percentage points lower than LSTM, indicating it is a feasible lightweight alternative. CNN-LSTM reached 71.72%, likely limited by added complexity on already extracted feature vectors. MLP achieved 69.53%, while MLP with Dropout decreased to 61.88%, suggesting limited temporal modeling ability. Pure CNN performed worst (43.44%), indicating that a convolution-only structure is not suitable for this feature-vector setting.

ISSN 2616-5775 Vol. 9, Issue 5: 36-47, DOI: 10.25236/AJCIS.2026.090505

1) Training-Curve Analysis: Figure 6 shows training/validation loss and accuracy curves for representative models (LSTM, GRU, and MLP).

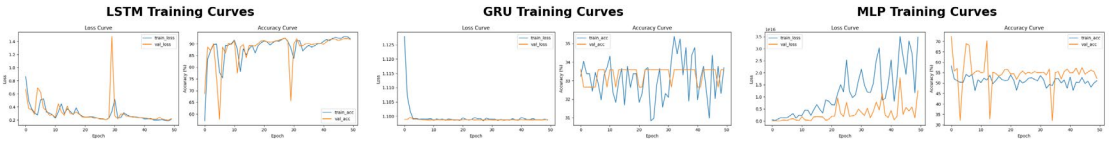

Fig. 6. Training-curve comparison of LSTM, GRU, and MLP

LSTM showed the most stable optimization behavior. Training and validation losses decreased smoothly and converged around epoch 35, with no evident overfitting pattern.

GRU followed a similar trend and converged slightly earlier (around epoch 30), likely due to its lower parameter count. However, final performance remained slightly below LSTM, indicating LSTM's stronger long-dependency modeling in this task.

MLP displayed less stable validation behavior and lower final accuracy than recurrent models, highlighting its limited ability to capture temporal dependencies in EEG signals.

2) Model Selection and Performance Summary: Based on all experiments, LSTM is selected as the core emotion-classification model due to best overall metrics, stable convergence, and good generalization. Its gate mechanism and memory structure effectively capture long-term EEG dependencies while retaining robustness to noise.

GRU remains a practical lightweight alternative in resource-constrained scenarios (e.g., mobile or embedded deployment). Future work can explore compression and distillation to transfer LSTM knowledge into smaller GRU models.

4.3 System Website and Visualization

1) Website Function Overview: We implemented a web-based interactive prototype for patients and caregivers. The system covers EEG emotion monitoring, personalized poetry generation, and multimodal presentation. It follows a front-end/back-end separated architecture: React for responsive UI, Flask for EEG processing/model inference/content generation, and SQLite for patient profiles and historical records.

Main functions include account/login management for patient-caregiver roles, patient-profile editing with autobiographical inputs, real-time 14-channel EEG display with contact-quality monitoring, LSTM-based emotion-state visualization, personalized poetry generation with manual review options, synchronized text-to-speech plus background music playback, and historical session logging with export and trend analysis.

2) Interface Screenshots and Functional Description: Figure 7 presents the homepage, designed with a clear and elder-friendly visual style. It provides three entries: ``Start Session" for EEG/emotion workflow, ``Patient Management" for profile maintenance, and ``History" for record review and statistics.

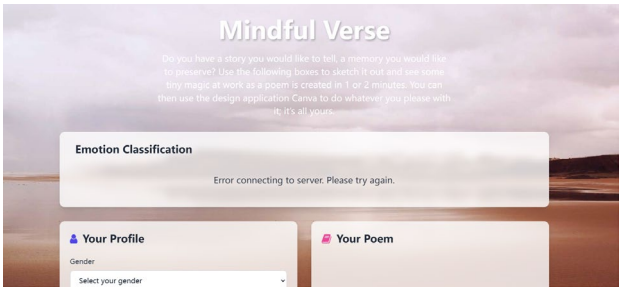

Fig. 7. System homepage

Figure 8 shows the core interaction page, integrating EEG monitoring, emotion classification, and poetry presentation.

ISSN 2616-5775 Vol. 9, Issue 5: 36-47, DOI: 10.25236/AJCIS.2026.090505

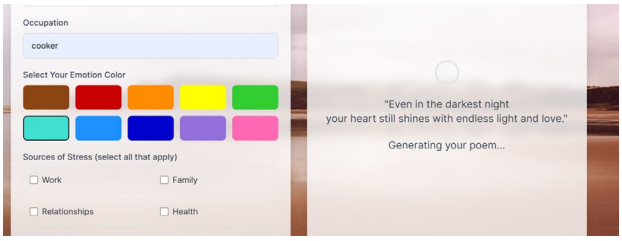

Fig. 8. Core interaction page of the system

The layout is organized into three areas. The left panel displays real-time 14-channel EEG waveforms (y-axis: potential in μV ; x-axis: time in seconds), with channel-wise impedance indicators (green/yellow/red for good/medium/poor contact). The waveform panel supports rolling display and adjustable time/amplitude scale.

The center panel shows emotion-monitoring results. A primary indicator displays current state (POSITIVE/NEGATIVE/NEUTRAL), followed by class-confidence bars and a short-term trend plot for recent emotional dynamics.

The right panel presents personalized poetry in large readable text, with controls for playback, pause/replay, save, and regenerate. A feedback area captures patient/caregiver comments on recalled memory and mood, supporting future personalization updates.

3) Prompt Visualization: To improve transparency and user engagement, the interface provides prompt visualization. When caregivers click ``View Generation Process,'' a dialog shows the complete prompt used for poem generation, including System Prompt constraints and User Input (patient profile + current emotion). This improves interpretability and trust in AI-generated content.

Caregivers can also update patient information (e.g., new life stories or revised interests), and the system synchronizes these updates into User Input in real time. This supports a human-in-the-loop design and improves controllability and safety.

5. Conclusion

5.1 Key Findings

This study developed an EEG- and language-generation-based emotional-arousal assistance system and established a full “monitor-recognize-generate-present” pipeline for practical non-pharmacological AD intervention.

1) Technical Outcomes: For emotion recognition, the six-model benchmark shows LSTM as best (Accuracy 91.56%, Precision 92.9%, F1-score 91.42%), outperforming GRU and CNN-LSTM. Inference is about 15 ms per sample and end-to-end response is around 500 ms, satisfying online intervention requirements. In content generation, the system produces simplified and positive personalized poetry; in pilot tests, 90% of participants reported good autobiographical alignment and 80% rated readability and emotional expression positively.

2) Clinical Application Value: For patients and caregivers, the system provides objective emotion assessment, personalized stimulation content, and continuous support in home settings. In pilot observations, 100% of participants reported emotional arousal from combined poetry-music stimulation, and 70% reported stronger effects than single-modality stimulation. For healthcare practice, the study validates the feasibility of a ``monitor-feedback-intervene'' closed-loop paradigm and provides a transferable idea for personalized non-pharmacological intervention in other neuropsychiatric conditions.

3) Societal Impact: With population aging, AD care demand continues to increase[24]. This system has the potential to provide scalable low-cost home assistance, reduce caregiving burden, and improve patient quality of life, especially in regions with limited medical resources.

5.2 Limitations

Despite positive findings, the study still has technical and application-level limitations.

1) Technical Limitations: First, dataset scale and population representativeness are limited: current data come mainly from a small number of healthy young participants, and model generalization to real

ISSN 2616-5775 Vol. 9, Issue 5: 36-47, DOI: 10.25236/AJCIS.2026.090505

AD patients remains to be verified. Second, wet electrodes have limited long-term comfort, which may affect compliance. Third, LLM outputs are stochastic and still require more stable automatic quality assessment and filtering.

2) Application Limitations: First, clinical evidence is still limited: only small-scale proof-of-concept testing has been completed, without large randomized controlled trials in real patient populations. Second, long-term efficacy and optimal intervention frequency remain unclear. Third, the system handles sensitive data (personal memories and EEG), requiring continuous strengthening of privacy protection, regulatory compliance, and ethical control of generated content.

5.3 Future Work

Future work will focus on data expansion, hardware optimization, generation-quality control, and clinical validation to move from proof-of-concept toward clinical usability.

1) Technical Improvements: We plan to expand patient datasets and apply transfer learning/domain adaptation to improve model generalization; adopt dry or semi-dry electrodes to improve comfort; and build automated content-assessment pipelines to reduce manual review workload and stabilize generation quality.

2) Domain-Specific LLM Training: Beyond general-purpose LLMs, we will explore training and fine-tuning smaller models for AD-specific use cases to improve controllability, inference efficiency, and privacy security while reducing dependence on external APIs.

3) Expansion from Web to Software and VR: The product will expand from Web to desktop and mobile platforms to improve interaction smoothness and offline capability. We will also explore immersive VR scene reconstruction to strengthen memory activation and emotional arousal effects.

4) Large-Scale Clinical Trials: Clinical evaluation will proceed in phases: first, a pilot study with 20--30 participants; then, randomized controlled trials with at least 100 participants. Key outcomes include safety, acceptability, and long-term efficacy. These steps are expected to support translation from technical validation to clinical deployment.

References

[1] World Health Organization. (2022). Dementia: Key facts. Retrieved from https://www.who.int/news-room/fact-sheets/detail/dementia
[2] Smith, J., et al. (2024). Efficacy and safety of monoclonal antibodies targeting amyloid-beta in Alzheimer's disease. Journal of Neurology, 271(3), 123-145.
[3] Johnson, A., et al. (2024). Pharmacological interventions for cognitive symptoms in Alzheimer's disease. Cochrane Database of Systematic Reviews, 2024(4), CD001234.
[4] García, M., et al. (2023). Music therapy as non-pharmacological treatment in Alzheimer's disease. Healthcare, 11(15), 1493.
[5] Alarcão, S. M., & Fonseca, M. J. (2019). Emotions recognition using EEG signals: A survey. IEEE Transactions on Affective Computing, 10(3), 374-393.
[6] Zhang, T., Wang, X., Xu, X., & Chen, C. L. (2024). Deep learning for electroencephalography emotion recognition. Scientific Reports, 14, Article 17863.
[7] Thirunavukarasu, A. J., et al. (2023). Large language models in medicine. Nature Medicine, 29(8), 1930-1940.
[8] Ekman, P., & Friesen, W. V. (1978). Facial action coding system. Palo Alto: Consulting Psychologists Press.
[9] Li, Y., Zeng, J., Shan, S., & Chen, X. (2022). Occlusion aware facial expression recognition using CNN with attention mechanism. IEEE Transactions on Image Processing, 28(5), 2439-2450.
[10] Zhang, F., Zhang, T., Mao, Q., & Xu, C. (2023). Multi-scale spatial-temporal interaction network for facial expression recognition. IEEE Transactions on Affective Computing, 14(2), 1103-1116.
[11] Akçay, M. B., & Oğuz, K. (2020). Speech emotion recognition. Speech Communication, 116, 56-76.
[12] Can, Y. S., Arnrich, B., & Ersoy, C. (2019). Stress detection in daily life scenarios. Journal of Biomedical Informatics, 92, 103139.
[13] Guo, H. W., et al. (2016). Heart rate variability signal features for emotion recognition. 2016 IEEE 16th International Conference on Bioinformatics and Bioengineering, 274-277.
[14] Cho, K., et al. (2014). Learning phrase representations using RNN encoder-decoder. arXiv preprint arXiv:1406.1078.

ISSN 2616-5775 Vol. 9, Issue 5: 36-47, DOI: 10.25236/AJCIS.2026.090505

[15] Alhagry, S., Fahmy, A. A., & El-Khoribi, R. A. (2017). Emotion recognition based on EEG using LSTM. International Journal of Advanced Computer Science and Applications, 8(10), 355-358.
[16] Brown, T., et al. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877-1901.
[17] Holtzman, A., et al. (2019). The curious case of neural text degeneration. arXiv preprint arXiv:1904.09751.
[18] Bird, J. J., et al. (2019). Mental emotional sentiment classification with an EEG-based brain-machine interface. Proceedings of theInternational Conference on Digital Image and Signal Processing (DISP'19).
[19] Davidson, R. J. (2003). Affective neuroscience and psychophysiology. Psychophysiology, 40(5), 655-665.
[20] Harmon-Jones, E., & Allen, J. J. (2004). The role of affect in the mere exposure effect. Personality and Social Psychology Bulletin, 23(8), 889-898.
[21] Eldaief, M. C., et al. (2011). Transcranial magnetic stimulation modulates the brain's intrinsic activity. Proceedings of the National Academy of Sciences, 108(52), 21229-21234.
[22] Williams, J. M. G., et al. (2007). Autobiographical memory specificity and emotional disorder. Psychological Bulletin, 133(1), 122.
[23] Hamburg, M. A., & Collins, F. S. (2010). The path to personalized medicine. New England Journal of Medicine, 363(4), 301-304.
[24] GBD 2019 Dementia Forecasting Collaborators. (2022). Estimation of the global prevalence of dementia in 2019. The Lancet Public Health, 7(2), e105-e125.