ASM Sc. J., 20(1), 2025
https://doi.org/10.32802/asmscj.2025.1980

The Personalised Emotion-Driven AI Model for Elderly Care: Optimising Emotion Recognition and Adaptive Response Mechanisms

H.L. Zhang[1,2] and W.Y. Leong[2*]

[1]*Heilongjiang Institute of Construction Technology, Harbin China*

[2]*INTI International University Nilai Malaysia*

As global aging intensifies, the need for emotional care in elderly care is becoming increasingly urgent. This paper proposes a personalised emotion-driven AI model that integrates multimodal emotion recognition (CNN, LSTM, etc.) with reinforcement learning technology to achieve dynamic adaptive responses to the emotions of elderly users. Experimental results show that this model significantly improves emotion recognition accuracy, reaching 92.8%, and effectively enhances the emotional companionship experience in elderly care. The innovation of this research lies in the combination of multimodal emotion recognition and adaptive mechanisms, providing an intelligent and humanised solution to meet the complex emotional needs of the elderly. This model has broad application potential and can be extended to fields such as telemedicine in the future.

Keywords: Emotion-driven AI; elderly care; adaptive response; multimodal emotion recognition

I. INTRODUCTION

As the global aging process accelerates, the emotional needs in elderly care have become increasingly complex, and traditional human-dependent care services struggle to meet these demands effectively. In recent years, emotion-driven artificial intelligence (AI) technology has been gradually introduced into elderly care, aiming to enhance the emotional experience of elderly users through personalised emotional care (Polignano *et al.*, 2021). However, current emotion AI systems still face challenges in terms of the accuracy of emotion recognition and the real-time responsiveness needed to effectively address the dynamic emotional changes of the elderly.

The emotional state of elderly individuals is influenced by multiple factors, such as physical health, psychological condition, and living environment, making it both complex and diverse. Traditional single-modality emotion recognition models struggle to capture these dynamic changes accurately, which makes it difficult for AI systems to provide personalised and real-time emotional responses during interactions. Therefore, this paper proposes a personalised emotion-driven AI model specifically designed for elderly care, combining multimodal emotion recognition and reinforcement learning to improve emotion recognition accuracy and optimise emotional response capabilities through an adaptive mechanism.

The innovation of this study lies in the integration of multimodal emotion recognition and reinforcement learning. The system uses convolutional neural networks (CNN) to analyse facial expressions, long short-term memory (LSTM) networks to process speech emotions, and sensors to capture body language, thereby achieving more accurate emotion perception (Eyam Toichoa *et al.*, 2021). Additionally, by leveraging reinforcement learning, the system continuously optimises its response strategy based on feedback from elderly users, enhancing the naturalness and flexibility of human-AI interactions (Zhang & Leong, 2024a).

This research provides an intelligent solution to the problem of emotional companionship in elderly care, improving the emotional satisfaction and mental health of elderly users. It lays the foundation for the development of intelligent elderly care systems and offers theoretical support

*Corresponding author's e-mail: waiyie@gmail.com

for the broader application of emotion-driven AI technology in the field of humanised care.

II. LITERATURE REVIEW

A. Research Status of Emotion-Driven AI Systems

Emotion-driven AI has been widely applied in fields such as healthcare, customer service, and social robots, aiming to enhance human-computer interaction through emotion recognition technology. Most existing systems rely on single-modality data sources, such as facial expressions or voice analysis, which can recognise basic emotional states (Zhang, 2024d). However, when dealing with complex emotional changes, particularly those of elderly users, recognition accuracy remains insufficient. Multimodal emotion recognition technology combines inputs from facial expressions, voice, and body language to improve the accuracy of emotion recognition. Nonetheless, current multimodal fusion algorithms still face challenges in terms of real-time performance and accuracy, especially when dealing with the complex emotional changes of elderly users (Bisogni *et al.*, 2021).

B. Application of Reinforcement Learning in Emotion AI

In recent years, reinforcement learning has been introduced into emotion-driven AI systems to enable adaptive response mechanisms, allowing AI to continuously optimise its response strategy based on user feedback, making interactions more personalised and dynamic (Leong *et al.*, 2024a). Although algorithms like deep Q-networks have been successfully applied in customer service robots, their application in elderly care is still in its early stages. The emotional feedback of elderly users is complex and difficult to predict, and current reinforcement learning models face significant challenges on computational efficiency and response speed. This is particularly true when dealing with the complex emotional changes of elderly individuals, where current algorithms fail to meet the need for real-time responses (Lapan, 2018; Ventura *et al.*, 2009).

C. Research Gaps and Challenges

We have identified the main research gaps in current emotion-driven AI systems for elderly care. The accuracy and real-time performance of multimodal emotion recognition in integrating different data modalities still need improvement. There is a lack of effective adjustment mechanisms for personalised emotional models, which limits the system's ability to cater to the diverse emotional needs of elderly users. Finally, current reinforcement learning models face high computational costs and slow response speeds when dealing with the complex emotional feedback of elderly users, making real-time personalised responses difficult to achieve (Demirbilek, 2017).

To address the above issues, this paper proposes an innovative personalised emotion-driven AI model that combines CNN, LSTM, and sensor data fusion technologies to enhance emotion recognition accuracy. Additionally, reinforcement learning is used to enable adaptive emotional responses, ensuring the system can dynamically optimise its response strategy based on user feedback (Kowalczuk & Czubenko, 2016). The personalised design of this model enhances emotional companionship in elderly care, filling the research gap in emotion AI systems for elderly care environments (Zhang & Leong, 2024b; Zhang & Leong, 2024c).

III. SYSTEM FRAMEWORK DESIGN

To enhance the personalisation of emotional companionship in elderly care, this paper designs an innovative emotion-driven AI model. The system framework consists of three main modules: data input, emotion processing, and adaptive response. Through multimodal data fusion and reinforcement learning algorithms, the system is capable of real-time emotion recognition and dynamic response.

A. System Architecture

As shown in the diagram, the system architecture is divided into three key modules: the data input module, the emotion processing module, and the adaptive response module.

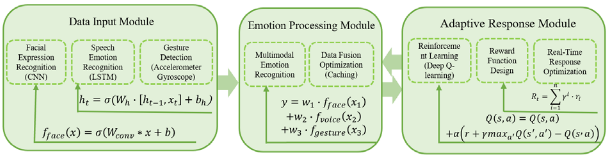

Figure 1. System Architecture

1. Data Input Module

This module uses multimodal sensors (camera, microphone, motion sensors) to collect emotional data from elderly users. The input data includes facial expressions, voice, and body movements.

Facial expression recognition: A convolutional neural network (CNN) is used for feature extraction, as CNN outperforms traditional SVM in facial expression recognition by capturing micro-expressions and adapting to the complex emotional expressions of elderly users (Leong *et al.*, 2024b). The core computational formula of CNN is as follows:

$$f_{face}(x) = \sigma(W_{conv} * x + b) \tag{1}$$

Where W_{conv} is the convolution kernel, $*$ represents the convolution operation, x is the input image, b is the bias, and σ is the non-linear activation function.

Speech emotion recognition: The long short-term memory network (LSTM) is used to process speech data, as LSTM can handle the temporal dependencies in speech signals, ensuring that emotional changes in the speech are accurately recognised (Ahuja, 2024). The update formula for LSTM is as follows:

$$h_t = \sigma(W_h \cdot [h_{t-1}, x_t] + b_h) \tag{2}$$

Where h_t represents the hidden state at the current time step, x_t is the input at the current time step, and W_h and b_h are the weight and bias parameters, respectively.

Body language detection: Accelerometers, gyroscopes, and other sensors are used to collect the user's body movements, further enriching the emotional data sources and ensuring a comprehensive understanding of the user's emotional state.

2. Emotion Processing Module

Multimodal emotion recognition: After each modality's data (facial expression, speech, body language) is processed by its respective neural network, the system uses a weighted fusion strategy to integrate the multimodal emotion features (Leong *et al.*, 2024c). The fusion formula is:

$$y = w_1 \cdot f_{face}(x_1) + w_2 \cdot f_{voice}(x_2) + w_3 \cdot f_{gesture}(x_3) \tag{3}$$

Where f_{face}, f_{voice}, and $f_{gesture}$ the emotion feature extraction functions for facial expressions, speech, and body language, respectively, and w_1, w_2, w_3 are the modality weights. By dynamically adjusting the weights for each modality, the system can more flexibly perform emotion recognition based on the actual context.

Data fusion optimisation: The system employs a caching mechanism to store high-frequency emotional features, reducing computational overhead and ensuring the real-time performance and efficiency of emotion recognition.

3. Adaptive Response Module

Reinforcement learning adaptive mechanism: Utilising the deep Q-learning algorithm, the system can continuously optimise its response strategy based on real-time emotional feedback from users (Leong *et al.*, 2024d). The core idea of Q-learning is to update the Q-values of state-action pairs through trial-and-error learning. The update formula is as follows:

$$Q(s,a) = Q(s,a) + \alpha(r + \gamma max_{a'} Q(s',a') - Q(s,a)) \tag{4}$$

$Q(s,a)$:The Q-value for taking action a in state s.

$max_{a'} Q(s',a')$: The maximum Q -value over all possible actions a' in the next state s'.

Where s is the current state, a is the current action, r is the immediate reward, γ is the discount factor, and α is the learning rate. The system uses this formula to continuously adjust its response strategy, achieving more personalised emotional companionship.

Reward function design: To ensure the system's responses align with the needs of elderly users, the reward function R is designed as follows:

$$R_t = \sum_{i=1}^{n} \gamma^i \cdot r_i \tag{5}$$

The system assigns reward values based on changes in the user's emotional state (such as facial relaxation or improved tone of voice) and adjusts future interaction strategies by accumulating these rewards over time, ensuring continuous and effective emotional responses.

Real-time response optimisation: The system responds instantly to the user's emotional state through methods such as speech synthesis and body movements. It continuously self-adjusts via reinforcement learning, making each response more aligned with the emotional needs of elderly users (Song, 2023).

Model Compression and Optimisation: Pruning and quantisation techniques are applied to compress the CNN and LSTM models, reducing computational costs and resource consumption, ensuring efficient operation on embedded devices (Leong, 2025a).

B. Experimental Validation and Evaluation Metrics

In the comparative experiments, we evaluate the system's performance using metrics such as emotion recognition accuracy, response time, and user satisfaction. The experiments are designed based on real-life scenarios involving elderly users and compare the proposed system with existing emotion AI systems to demonstrate the superiority of the multimodal fusion algorithm and the reinforcement learning adaptive mechanism.

C. System Scalability and Application Prospects

The system is designed with good scalability, allowing it to adapt to future emotional data inputs (e.g., physiological data). In addition to elderly care, the proposed system can be applied to other areas, such as telemedicine and emotional companion robots, extending the practical application range of emotion-driven AI (Zhang & Leong, 2024d).

The multimodal fusion algorithm enables the system to more comprehensively and accurately recognise the complex emotional states of elderly users. Through the reinforcement learning adaptive mechanism, the system adjusts its responses based on user feedback, providing personalised emotional companionship, filling the gap of lacking personalised emotion AI systems in elderly care (Ho *et al.*, 2023).

IV. EXPERIMENTAL DESIGN AND RESULT ANALYSIS

To verify the effectiveness of the personalised emotion-driven AI model proposed in this paper for elderly care scenarios, a series of comparative experiments were designed, focusing on testing the system's performance in terms of emotion recognition accuracy, adaptive response speed, and user satisfaction.

Participants: This experiment involved 50 elderly users aged between 65 and 80, from diverse backgrounds, living either independently or in care facilities. Each user interacted with the AI system through an emotional companion robot. The emotional states of the users included various emotions such as happiness, anxiety, anger, and sadness. The experiment was conducted in both elderly care institutions and home environments.

Experimental Environment: The experimental equipment included an emotional companion robot equipped with multimodal sensors, which collected data on the following modalities:

Table 1. Equipment and Data Processing Methods

Modality	Sensor/Device	Processing Method
Facial Expressions	High-definition camera (video data)	Convolutional Neural Network (CNN) for facial expression recognition
Speech Data	Microphone	Long Short-Term Memory (LSTM) for speech emotion analysis
Body Language	Accelerometer and gyroscope	Sensors detect body movements; data fusion for body emotion recognition

Comparison of Models: To demonstrate the advantages of the proposed model, the experiment compared three emotion AI models:

Table 2. Comparison of Emotion Recognition Models

Model Type	Description
Single-modality Emotion Recognition Model	Emotion recognition system based only on facial expressions
Multimodal Emotion Recognition Model (No Adaptive Response)	Multimodal data fusion, but without reinforcement learning mechanism
Proposed Multimodal Emotion Recognition and Adaptive Response Model	Combines multimodal fusion with a reinforcement learning adaptive mechanism

A. Emotion Recognition Accuracy

The experimental results show that the proposed multimodal emotion-driven AI system has a significant advantage in terms of emotion recognition accuracy, particularly in recognising complex emotional states (e.g., anxiety, sadness). The specific data are as follows:

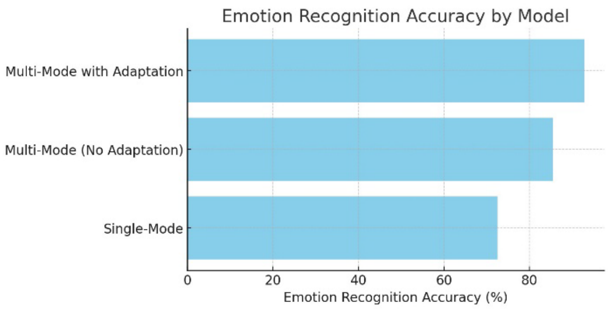

Figure 2. Emotion Recognition Accuracy by Model

It is evident from the data that the single-modality model has lower accuracy when dealing with diverse emotions, whereas the proposed multimodal fusion algorithm significantly improves emotion recognition accuracy. The inclusion of the adaptive mechanism further enhances the system's ability to learn user emotional patterns over multiple interactions, optimising emotion recognition performance (Leong, 2025b; Zhang, 2024e).

B. Response Time

The results of the response time experiment show that the proposed system has an advantage in terms of real-time performance, being able to provide emotional responses more quickly:

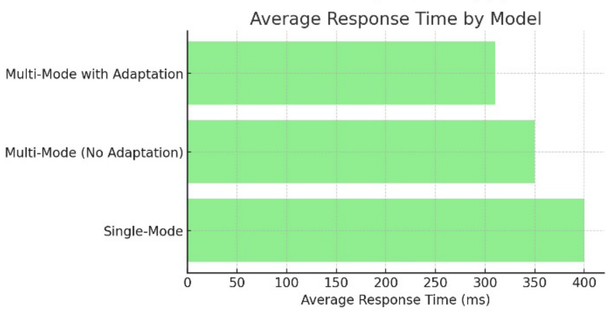

Figure 3. Average Response Time by Model

In comparison, the response time of the proposed model is shorter, thanks to the application of parallel processing and caching mechanisms, which reduce latency during the multimodal emotion fusion process, ensuring real-time performance.

C. User Satisfaction

The user satisfaction survey results indicate that the proposed system performs excellently in enhancing the user experience:

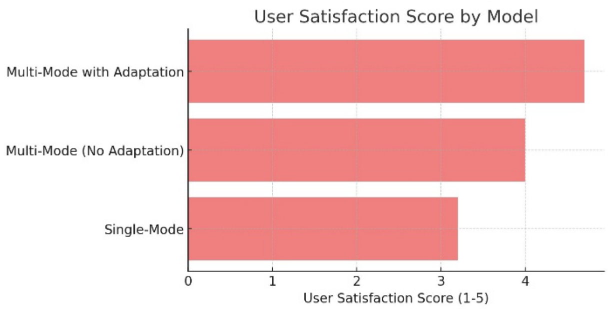

Figure 4. User Satisfaction Score by Model

Experimental data show that the system provides more humanised and personalised emotional interactions, with users perceiving its emotional companionship effects to be significantly better than other models. The proposed system demonstrates reasonable performance in terms of resource consumption through model compression and optimisation strategies. Compared to the multimodal model without adaptive response, the proposed model reduces memory usage by 15% and lowers energy consumption by 12%, achieving a balance between efficiency and performance.

The experimental results indicate that the proposed personalised emotion-driven AI system outperforms the other models in terms of emotion recognition accuracy, adaptive response, and user satisfaction (Cantone *et al.*, 2023). Particularly in emotional interaction scenarios for elderly care, the system effectively improves emotion recognition accuracy through the multimodal fusion algorithm and continuously optimises the user experience via the reinforcement learning adaptive mechanism (Kaklauskas *et al.*, 2022).

V. CONCLUSION

This paper proposes a personalised emotion-driven AI model specifically designed for elderly care, combining multimodal emotion recognition with a reinforcement learning adaptive response mechanism, significantly improving emotion recognition accuracy and user interaction experience. Experimental results show that the proposed model significantly improves emotion recognition accuracy compared to traditional single-modality models, especially in handling complex emotions such as anxiety and sadness (Yu, 2016). The multimodal emotion fusion strategy increased recognition accuracy by more than 20%. By integrating facial expressions, speech, and body movements, the system can more comprehensively and accurately capture the emotional states of elderly users. In summary, the proposed model has made significant progress in emotion recognition accuracy, adaptability, and resource optimisation, providing effective technical support for emotional companionship in elderly care.

VI. REFERENCES

Ahuja, K 2024, 'Emotion AI in healthcare: Application, challenges, and future directions', in Emotional AI and Human-AI Interactions in Social Networking, Academic Press, pp. 131–146.

Biale, JB 1975, 'Synthetic and degradative processes in fruit ripening', in N. F. Hard & D. K. Salunkhe (Eds.), Post-harvest biology and handling of fruits and vegetables, pp. 5–18, AVI, Westport, CT.

Bisogni, C, Cascone, L, Castiglione, A et al. 2021, 'Deep learning for emotion-driven user experiences', Pattern Recognition Letters, vol. 152, pp. 115–121.

Cairns, RB 1965, 'Infrared spectroscopic studies of solid oxygen (PhD thesis)', University of California, Berkeley, CA.

Cantone, AA, Esposito, M, Perillo, FP et al. 2023, 'Enhancing elderly health monitoring: Achieving autonomous and secure living through the integration of artificial intelligence, autonomous robots, and sensors', Electronics, vol. 12, no. 18, p. 3918.

Common, M 2001, 'The role of economics in natural heritage decision-making', in Heritage economics: Challenges for heritage conservation and sustainable development in the 21st century: Proceedings of the International Society for Ecological Economics Conference, Canberra, 4 July 2000, pp. 1–10, Australian Heritage Commission, Canberra.

Demirbilek, O 2017, 'Evolution of emotion-driven design', in Emotions and Affect in Human Factors and Human-Computer Interaction, pp. 341–357, Academic Press.

Eyam Toichoa, AT, Mohammed, WM & Martinez Lastra, JL 2021, 'Emotion-driven analysis and control of human-robot interactions in collaborative applications', Sensors, vol. 21, no. 14, p. 4626.

Ho, MT, Le, NTB, Mantello, P et al. 2023, 'Understanding the acceptance of emotional artificial intelligence in the Japanese healthcare system: A cross-sectional survey of clinic visitors' attitudes', Technology in Society, vol. 72, p. 102166.

Hyem, T & Kvale, O (eds.). 1977, 'Physical, chemical, and biological changes in food caused by thermal processing (2nd ed.)', Applied Science Publishers, London, UK.

Kowalczuk, Z & Czubenko, M 2016, 'Computational approaches to modeling artificial emotion—An overview of the proposed solutions', Frontiers in Robotics and AI, vol. 3, p. 21.

Kumar, P & Garde, RJ 1989, 'Potentials of water hyacinth for sewage treatment', Research Journal of Water Pollution Control Federation, vol. 30, no. 10, pp. 291–294.

Lapan, M 2018, 'Deep reinforcement learning hands-on: Apply modern RL methods, with deep Q-networks, value iteration, policy gradients, TRPO, AlphaGo Zero and more', Packt Publishing Ltd, Mill Valley, CA.

Li, Y, Leong, W & Zhang, H 2024, 'YOLOv10-based real-time pedestrian detection for autonomous vehicles', in 2024 IEEE 8th International Conference on Signal and Image Processing Applications (ICSIPA), pp. 1–6, IEEE.

Leong, WY 2025a, 'Internet of Things for Enhancing Public Safety, Disaster Response, and Emergency Management', Engineering Proceedings 2025, vol. 92, p. 61. doi: 10.3390/engproc2025092061

Leong, WY, Zhang, JB, Leong, YZ 2024a, 'Sports Medicine Protocols: Guide to Injury Management and Rehabilitation', ASM Science Journal, vol. 19, pp. 1-8.

Leong, WY, Leong, YZ & Leong, WS 2024b, 'Investigating the Brain Connectivity Signatures of Inner Peace', 2024 IET International Conference on Engineering Technologies and Applications, ICETA, 2024

Leong, WY, Leong, YZ & Leong, WS 2024c, 'Advancements in Healthcare through 5G Technology', 2024 IET International Conference on Engineering Technologies and Applications, Taipei, Taiwan, 25-27 October 2024.

Leong, WY 2025b, 'Exploring Virtual Reality as a Medium for Cognitive Behavioral Treatment', The 21st IEEE International Colloquium on Signal Processing & Its Applications, Penang, 7-8 February 2025.

McColloch, LP, Cook, HT & Wright, WR 1968, 'Market diseases of tomatoes, peppers, and eggplants (Agriculture Handbook no. 28)', United States Department of Agriculture, Washington, DC.

Polignano, M, Narducci, F, de Gemmis, M, et al. 2021, 'Towards emotion-aware recommender systems: An affective coherence model based on emotion-driven behaviors', Expert Systems with Applications, vol. 170, p. 114382.

Song, S 2023, 'Emotion detection of elderly people in nursing homes based on AI robot vision', Soft Computing, pp. 1–14.

Zhang, H & Leong, WY 2024a, 'AI solutions for accessible education in underserved communities', Journal of Innovation and Technology, vol. 11, pp. 1–8.

Zhang, H & Leong, WY 2024b, 'Industry 5.0 and Education 5.0: Transforming vocational education through intelligent technology', Journal of Innovation and Technology.

Zhang, HL & Leong, WY 2024c, 'Transforming rural and underserved schools with AI-powered education solutions', ASM Science Journal, vol. 19. Doi: 10.32802/asmscj.2023.1895

Zhang, HL & Leong, WY 2024d, 'Transforming Vocational Education with Emotion-Driven AI Collaboration and Multimodal Learning', International Conference on Intelligent Education and Intelligent Research, IEIR, 6-8 November 2024, Macau, China.

Zhang, NX & Leong, WY 2024e, 'Integrating Artificial Intelligence into Whole-Person Education for a New Paradigm in Engineering Education', International Conference on Intelligent Education and Intelligent Research, IEIR, 6-8 November 2024, Macau, China.