

The Design of Multimodal-based Micro-expression Recognition Training Platform for Adolescents

Tian Yu^{1,2}

¹*State Key Laboratory of Brain and Cognitive Sciences,
Institute of Psychology,
Chinese Academy of Sciences
Beijing, China*
²*Department of Psychology
University of Chinese Academy of Sciences
Beijing, China*

Liu Ye^{1,2*}

¹*State Key Laboratory of Brain and Cognitive Sciences,
Institute of Psychology,
Chinese Academy of Sciences
Beijing, China*
²*Department of Psychology
University of Chinese Academy of Sciences
Beijing, China
liuye@psych.ac.cn*

Abstract—Adolescence is a phase of emotional and social skill development, during which emotional intelligence plays a crucial role in shaping adolescents' self-concept, enhancing social interactions, and improving academic performance. Previous research demonstrated that the accurate recognition of facial expressions is an integral component of emotional intelligence. Micro-expression recognition training can enhance expression recognition ability and emotional intelligence. We designed a multimodal-based micro-expression recognition training platform by combining visual, auditory, and interactive multimodal elements to promote the micro-expression recognition ability and emotional intelligence of adolescents. The platform offers adolescents an innovative training model that incorporates more engaging, interactive, and personalized training experiences.

Keywords—multimodal, micro-expression recognition training, emotional intelligence, adolescent

I. INTRODUCTION

Emotional information in human communication is predominantly conveyed through facial expressions, making facial expression recognition critical for emotional intelligence [1]. Previous research demonstrated a strong correlation between facial expression recognition and emotional intelligence, indicating that individuals with high emotional intelligence can recognize facial expressions more rapidly and accurately than those with lower emotional intelligence. Additionally, facial expression recognition skills are improved significantly with targeted training [2]. Through a decade of behavioral research, facial expression recognition training has enhanced abilities in expression perception, emotional comprehension, emotional regulation, and the utilization of emotions to facilitate cognitive processes, thereby improving emotional intelligence [3–6].

We investigated the development and application of a multimodal micro-expression recognition training platform for adolescents considering the principles of psychological and behavioral sciences. By integrating visual, auditory, and interactive feedback modalities, the platform provides an immersive training environment tailored to the needs of adolescents. In contrast to conventional facial expression recognition training, this multimodal platform employs advanced technologies, such as multimodal emotion recognition, to offer personalized and adaptive training solutions for each adolescent user.

II. ALGORITHMS AND TECHNIQUES

A. Personalized Emotion Learning Recommendation Algorithm

The personalized recommendation provides personalized suggestions based on their preferences and historical behaviors [7]. Personalized recommendation systems can be used in education, e-commerce, and other fields [7]. The main techniques of personalized recommendation include the following.

1) Content-Based Recommendation (CBR)

Based on the users' preferences and historical behaviors, similar contents are recommended for the users. The cosine similarity is an example of calculating the similarity of two vectors A and B (1).

$$\text{similarity}(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} \quad (1)$$

where $A \cdot B$ is the dot product of the vectors, and $\|A\|$ and $\|B\|$ are the norms of the vectors

2) Collaborative Filtering Recommendation (CFR)

Based on the users' historical behaviors, the users are recommended content similar to others. The similarity between users U_i and U_j is calculated using (2).

$$\text{similarity}(U_i, U_j) = \frac{\sum_{k \in R} (r_{ik} - \bar{r}_i)(r_{jk} - \bar{r}_j)}{\sqrt{\sum_{k \in R} (r_{ik} - \bar{r}_i)^2} \sqrt{\sum_{k \in R} (r_{jk} - \bar{r}_j)^2}} \quad (2)$$

where r_{ik} is the rating of item k by user U_i , \bar{r}_i is the average rating of user U_i , and R is the rating of both user U_i and U_j

3) Hybrid Recommendation

Combining the use of content-based recommendation and collaborative filtering-based recommendation, better recommendation results can be obtained. We developed a personalized recommendation algorithm based on adolescents' emotional state during the training and the training feedback to evaluate it by monitoring their physiological signals and expression changes in real time. The difficulty and process of the training are adjusted accordingly. Combined with the results of the post-training emotional intelligence test, the algorithm generates targeted emotional guidance and social strategies based on the aforementioned equations. The algorithm enhances the adaptability and effectiveness of the training, develops

emotional intelligence in adolescents, and provides accurate and personalized guidance.

B. Multimodal Emotion Recognition

Multimodal emotion recognition is a precise analysis tool that combines multiple modal data such as facial expression, speech, and gesture [8]. Given the complexity and variability of emotional expressions in adolescents, multimodal emotion recognition technology offers a holistic approach to capturing their emotional states. This integration minimizes the limitations and errors associated with single-modal recognition, providing a more nuanced understanding of adolescents' emotional states and potential underlying emotions. This capability is valuable in the context of micro-expression recognition training platforms.

The key modalities of multimodal emotion recognition technology include facial expression analysis, speech emotion analysis, and gesture analysis [8]. Facial expression and gesture analysis require computer vision techniques to extract features from pictures or videos and perform expression analysis and emotion classification. Speech emotion analysis includes the characterization of intonation, pitch, speech rate, and speech intensity [8]. Multimodal sentiment analysis is used to accurately categorize emotions such as happiness, anger, sadness, and other basic emotions using machine learning or deep learning classifiers [9]. Multimodal fusion improves the accuracy of emotion recognition and supports personalized education and emotional intelligence development [10].

III. SYSTEM CONSTRUCTION

A. Platform Design

Unlike traditional micro-expression recognition training tools, the developed platform integrates oral and visual signals to realize real-time monitoring and in-depth analysis of adolescents' emotional states in micro-expression recognition training. The difficulty of training is adjusted based on the analysis results. For example, when users perform well with high emotions in emotion recognition tasks, the platform increases the complexity of the task and the difficulty in the users' micro-expression recognition ability. Conversely, when users encounter difficulties and become discouraged with a particular task, the platform adjusts the task difficulty for improvement (Fig. 1). At the same time, the training platform tracks the user's performance, analyzes progress, and provides adaptive feedback according to the individual's ability. This progress tracking is for single training sessions and the long-term evaluation of the training effect to ensure continuous improvement. In addition, the platform also integrates emotional intelligence testing tools to recognize, manage, and express emotions. Based on the results, the platform customizes personalized emotion management strategies to help adolescents understand and cope with emotions. In addition, data security and privacy protection are also considered in the platform design.

All collected data is encrypted and anonymized to ensure compliance with data protection regulations. The user interface is designed to be intuitive and user-friendly, allowing schools and counselors to access sentiment analysis results and engage effectively with adolescents. The platform's scalability and integration ensure upgraded responses and technological advancements for user needs, while seamlessly interfacing with educational technology.

This multimodal micro-expression recognition training platform serves as a training tool for micro-expression recognition and an emotional support system for adolescents with its advanced architecture and intelligent features. By enhancing adolescents' emotional intelligence and social skills through targeted training, the platform simultaneously provides educators and mental health professionals with valuable data and insights, facilitating collaborative efforts to support adolescents' mental health and personal development.

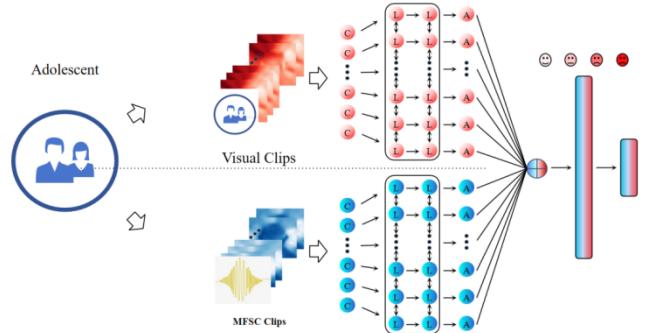


Fig. 1. Construction diagram of multimodal micro-expression recognition training platform.

B. Multimodal Micro-expression Recognition Training

The micro-expression recognition training process is based on Ekman's micro-expression recognition training tool (METT, <http://www.mettonline.com>). The tool includes benchmarks, training, practice, review, and improvement measures to help adolescents effectively enhance emotional intelligence through systematic training and practice to enhance emotional intelligence. This process helps participants improve their emotional awareness, management ability, empathy, and social skills by combining theory and practice. It verifies the enhancement effect through post-test to promote the continuous growth of individuals in emotional intelligence (Fig. 2).

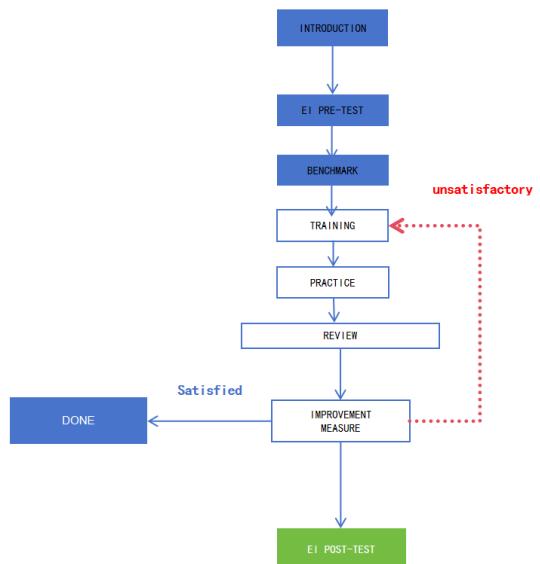


Fig. 2. Flowchart of micro-expression recognition training.

The developed platform encompasses a variety of modules to offer interactive personalized service. It features training modules that extend beyond basic instruction to include additional emotion-related components. The

personalized learning module delivers customized guidance derived from training analysis results and enhances practical skills through the simulation of social scenarios. This module recommends tailored learning resources according to user needs, and the learning resource module. The community support module provides an extensive array of learning materials and a platform for communication.

Although research on micro-expression recognition training and its impact on adolescents' emotional intelligence is limited, such training helps individuals better understand emotional states in real-life scenarios, thereby enabling appropriate responses in interpersonal interactions. Individuals trained for micro-expression exhibit superior performance in emotion recognition, emotion understanding, emotion regulation, and social communication [11,12]. Additionally, increased emotional competence is crucial for improving social communication and interpersonal relationships [5,12–15]. Micro-expression recognition ability helps healthcare professionals better understand the emotional changes of patients, thus improving the efficiency of communication with patients [12]. In business, micro-expression recognition ability helps understand the intention of the other party for effective negotiation strategies [15]. Using systematic multimodal training in micro-expression recognition, adolescents can enhance their sensitivity to subtle emotional signals and improve their adaptability to complex social situations.

C. Data Acquisition

The training platform uses CASME and CASME2, the multimodal emotion recognition platform to obtain samples and emotion expressions. These two datasets cover pictures and videos of facial expressions from people of different ages, and emotion types, providing training materials for the model. CASME and CASMEII are micro-expression databases developed by the Institute of Psychology of the Chinese Academy of Sciences [16,17]. The CASME database contains high temporal-resolution facial video and picture data from Chinese participants in emotion-evoking tasks, which provides a high ecological sanitization of micro-expressions for the study of the micro-expressions [16]. CASME2 extends the dataset to include more samples, finer annotations, and newer technical approaches to enhance the representativeness and diversity of the data [17]. The introduction of the CASME and CASME2 datasets adds the video data of facial expressions in Chinese culture to the developed model, allowing the model to better understand and recognize emotional expressions.

D. Personalized Micro-expression Recognition Training and Emotional Education

The personalized micro-expression recognition training tool enhances the micro-expression recognition ability and emotional intelligence of adolescents using an intelligent technology framework. The tool adopts advanced user-personalized modeling technology, analyzes the adolescents' emotional state, behavioral habits, and psychological characteristics, and constructs a personalized learning profile. Then, the tool provides data for the intelligent recommendation engine which integrates emotion recognition, collaborative filtering, content filtering, and deep learning technologies to monitor and accurately interpret emotional states. Emotion recognition technology captures subtle changes in facial expressions to recognize emotional states. Collaborative filtering technology uses

group emotional responses to recommend content that triggers similar emotional experiences, while content filtering technology recommends emotional management strategies and social skill training based on personalized interests and emotional responses.

The tool adjusts the training difficulty and progress by monitoring the emotional state and feedback results during the training process and recommends targeted emotional guidance and social strategies in conjunction with the results of the emotional intelligence test for the development of youth emotional intelligence. The tool improves the personalization and effectiveness of micro-expression recognition training and provides a scientific and systematic solution for the enhancement of adolescents' emotional intelligence.

The deep learning algorithms enhance the system's ability to understand adolescents' complex emotional characteristics, making the recommendations accurate and personalized. In the design of scenario-based learning, the emotional performance of adolescents is considered in different scenarios, such as school, family, or social occasions. Scenario-compatible solutions are provided for the cultivation of emotional intelligence. The adjustment of dynamic learning paths ensures continuity and challenge, and at the same time, develops the youth's emotional intelligence. Emotional intelligence cultivation is important. Using the developed emotion recognition and management training strategies, the adolescents can understand and express their emotions better, and recognize and respond to other's emotions. The user feedback mechanism enables the system to self-learn and adjust according to the youth's direct feedback, optimizing the recommendation strategies to further enhance the learning experience.

IV. RESULTS

A questionnaire survey was conducted involving 320 adolescents using the micro-expression recognition test and the emotional intelligence inventory (EIS) to assess each participant's competence in the two domains. The data analysis result revealed a significant positive correlation between micro-expression recognition ability and emotional intelligence. Specifically, adolescents who scored higher on the test exhibited higher levels of emotional intelligence. Adolescents who accurately recognized subtle emotional signals of others demonstrated heightened emotional sensitivity, empathy, emotional regulation, and social skills (Fig. 5).

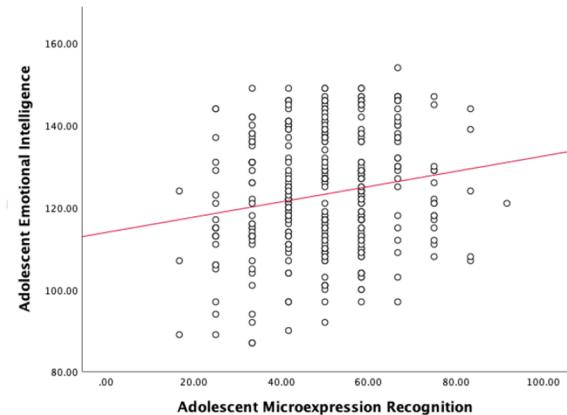


Fig. 3. Linear regression analysis result of micro-expression recognition ability and adolescents' emotional intelligence.

In the other experiment, 42 adolescents were randomly selected from the first experiment and were assigned to either the experimental or control group. The participants in the experimental group engaged in an online micro-expression recognition training program, while those in the control group watched videos related to facial expressions and emotions. The emotional intelligence pre-test was conducted before training, and a post-test was conducted following a three-week training. An obvious improvement was observed in the post-test scores compared to the pre-test scores of the experimental group, while no substantial differences were observed in the control group. Thus, the intervention effectively improved the participants' emotional intelligence and provided information on the mechanisms of emotional intelligence (Fig. 4).

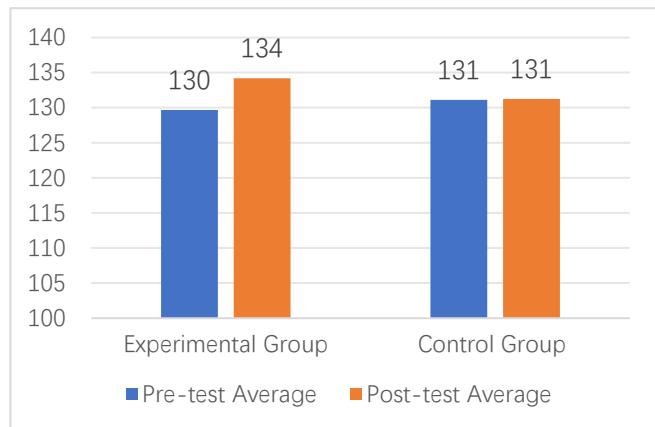


Fig. 4. Comparison of pre-test and post-test average of emotional intelligence total scores between the experimental and control groups.

V. CONCLUSION

We designed and developed a multimodal micro-expression recognition training platform for adolescents to enhance micro-expression recognition and emotional intelligence. The correlation between micro-expression recognition and emotional intelligence confirmed a significant relationship between these constructs. The platform incorporated visual, auditory, and interactive multimodal elements for immersive and dynamic training, thereby enhancing the engagement and effectiveness of the learning process. By personalizing the training experience, the platform delivers targeted emotional guidance and social strategies and optimizes the training outcomes for each individual. Adolescents who participated in the study demonstrated significant improvements in the accuracy of micro-expression recognition, as well as emotional intelligence. These results validate the efficacy of the multimodal training approach and highlight the platform's potential for practical application.

REFERENCES

- [1] Salovey, P., Caruso, D., & Mayer, J. D. (2004). Emotional intelligence in practice. *Positive psychology in practice*, 447-463.
- [2] Wezowski, K., & Penton-Voak, I. (2023). Associations between workplace emotional intelligence and micro expression recognition. *Acta Psychologica*, 240, 104046.
- [3] Yu, E. H., Choi, E. J., Lee, S. Y., Im, S. J., Yune, S. J., & Baek, S. Y. (2016). Effects of micro-and subtle-expression reading skill training in medical students: a randomized trial. *Patient education and counseling*, 99(10), 1670-1675.
- [4] Zhang, X., Chen, L., Zhong, Z., Sui, H., & Shen, X. (2018). The effects of the micro-expression training on empathy in patients with schizophrenia. In *Man-Machine-Environment System Engineering: Proceedings of the 17th International Conference on MMESE 17* (pp. 189-194). Springer Singapore.
- [5] Cannavo, M., Zomparelli, W., Carta, M. G., Romano, F., & La Torre, G. (2021). Effectiveness of a training on the recognition of subtle facial emotions in health and social workers. *Riv Psichiatr*, 56, 334-339.
- [6] Qu, F., Yan, S., Liang, J., & Wang, J. (2019). Effect of short-term micro-expression training on the micro-expression recognition performance of preschool children. In *Cognitive Systems and Signal Processing: 4th International Conference, ICCSIP 2018, Beijing, China, November 29-December 1, 2018, Revised Selected Papers, Part I* 4 (pp. 54-62). Springer Singapore.
- [7] Raghavendra, C. K., Srikanthiah, K. C., & Venugopal, K. R. (2018). Personalized recommendation systems (pres): a comprehensive study and research issues. *International Journal of Modern Education and Computer Science*, 10(10), 11.
- [8] Kessous, L., Castellano, G., & Caridakis, G. (2010). Multimodal emotion recognition in speech-based interaction using facial expression, body gesture and acoustic analysis. *Journal on Multimodal User Interfaces*, 3, 33-48.
- [9] Ezzamel, K., & Mahersia, H. (2023). Emotion recognition from unimodal to multimodal analysis: A review. *Information Fusion*, 99, 101847.
- [10] Geetha, A. V., Mala, T., Priyanka, D., & Uma, E. (2024). Multimodal Emotion Recognition with deep learning: advancements, challenges, and future directions. *Information Fusion*, 105, 102218.
- [11] Matsumoto, D., & Hwang, H. S. (2011). Evidence for training the ability to read microexpressions of emotion. *Motivation and emotion*, 35, 181-191.
- [12] Koc, E., & Boz, H. (2020). Development of hospitality and tourism employees' emotional intelligence through developing their emotion recognition abilities. *Journal of Hospitality Marketing & Management*, 29(2), 121-138.
- [13] Higginbotham, A. (2013). Deception Is Futile. *Wired*, 21(2), 90-97.
- [14] Ning, J., & Wang, R. (2020, November). Application of psychological analysis of micro-expression recognition in teaching evaluation. In *International Conference on Education Studies: Experience and Innovation (ICESEI 2020)* (pp. 71-77). Atlantis Press.
- [15] Bratianu, C., & Iordache, S. (2013). Knowledge dynamics analysis in negotiations. *Electronic Journal of Knowledge Management*, 11(1), pp 18-29.
- [16] Yan, W. J., Wu, Q., Liu, Y. J., Wang, S. J., & Fu, X. (2013, April). CASME database: A dataset of spontaneous micro-expressions collected from neutralized faces. In *2013 10th IEEE international conference and workshops on automatic face and gesture recognition (FG)* (pp. 1-7). IEEE.
- [17] Yan, W. J., Li, X., Wang, S. J., Zhao, G., Liu, Y. J., Chen, Y. H., & Fu, X. (2014). CASME II: An improved spontaneous micro-expression database and the baseline evaluation. *PloS one*, 9(1), e86041.