Proceedings of CONF-MLA 2026 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2026.LD38379

Research on Personalized Learning Support System Based on AI Emotion Recognition

Dengke Wang

School of Computer Science, St. Paul University Philippines, Tu'galao, Philippines
645930567@qq.com

Abstract. With the rapid popularization of artificial intelligence (AI) technology in the field of education, personalized learning will become the core direction of digital transformation in education. Learning emotions directly affect learning efficiency, learning motivation, and learning outcomes. In traditional learning systems, only knowledge acquisition is focused on, lacking perception and intervention of learners' emotional states. AI technologies such as computer vision, facial expression recognition, speech emotion recognition, and behavior recognition are becoming increasingly mature, providing the possibility for real-time capture and learning of emotions. Building a personalized learning system that combines emotional perception, intelligent recommendation, and dynamic intervention is an important research gap in the current field of educational technology.

Keywords: personalized learning emotions, emotional perception, education, digital transformation

1. Introduction

With the deep integration of artificial intelligence technology and the education industry, the digital transformation of education is constantly advancing, and personalized learning has become an important development direction. Learning emotions directly affect learners' learning motivation, level of focus, and ultimate learning outcomes, but traditional digital learning platforms mostly focus on knowledge mastery and neglect the perception and guidance of learners' emotional states.

With the continuous maturity of computer vision and multimodal emotion recognition technology, real-time capture of learners' classroom and online learning emotions has a technological foundation. In related research, scholars such as Li have pointed out that emotional fluctuations in learning are closely related to self-regulated learning and task performance, and emotional state is an essential dimension of personalized learning that cannot be ignored; Wang et al. verified that deep learning models can effectively recognize learners' facial learning emotions, but existing systems generally suffer from a disconnect between emotion recognition and learning intervention [4]. There are also studies proposing emotion perception adaptive learning frameworks, but there is a lack of a complete and practical system solution.

In response to the above gaps, this article constructs an integrated personalized learning support system of "emotion perception intelligent recommendation dynamic intervention", which combines multimodal emotion recognition to dynamically adapt learning content and intervention strategies.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.LD38379

This study can enrich the application practice of emotional computing in educational scenarios, provide reference ideas for the design and development of intelligent learning systems, and have certain reference value for promoting the development of people-oriented digital education.

2. Literature review

In the research field of personalized learning support systems based on AI emotion recognition, system development and personalized adaptive strategy decision models are closely related. At the level of learning system development, Zhou Jin et al. [1] pointed out that traditional learning system architectures mostly focus on data collection and analysis in the cognitive dimension, lacking access modules for emotional dimension data such as learning emotions, making it difficult to perceive learners' learning emotions such as focus, confusion, and boredom. Therefore, it is necessary to build a multimodal learning emotion calculation framework that includes emotional data collection, emotional modeling and recognition, emotional visualization feedback, emotional attribution intervention, and other hierarchical architectures to provide a carrier for real-time collection, processing, and feedback of emotional data.

In terms of adaptive strategy decision-making, Yu Shuzhen et al. [2] proposed that relying solely on AI to recognize learners' emotional states cannot automatically achieve personalized learning support. Currently, emotional computing still faces the dilemma of a disconnect between technology and practice in personalized teaching applications. It is necessary to construct an adaptive strategy decision-making model that adapts to learning scenarios, in order to identify the type and intensity of emotions, combine learners' knowledge level and learning progress as model inputs, dynamically match corresponding intervention strategies such as learning content adjustment, learning prompts, rest guidance, etc., and achieve a closed loop of emotion perception to learning intervention.

Existing research has validated the feasibility of multimodal emotion recognition in education systems, but many studies have designed system development and adaptive decision models separately, resulting in insufficient coupling between emotion recognition modules and intervention strategy modules, leading to the problem of emotion data being unable to effectively drive dynamic adjustment of learning strategies. Overall, how to integrate system architecture design and construct emotion driven adaptive strategy decision-making models remains a key challenge in current research on emotion computing empowering personalized learning.

3. System development method

This study independently designed and developed a four layer architecture personalized learning support system, which includes a lightweight multimodal emotion recognition model and an adaptive learning strategy decision-making model as the core innovation points of the entire article. The specific model principles and formulas are as follows.

The multimodal emotion fusion recognition model abandons the traditional single face recognition mode and integrates three modalities: facial visual features, learning behavior features, and text emotion features. It adopts a late stage weighted fusion strategy to solve the problem of low recognition accuracy in complex classroom environments.

Single modal feature output formula, let the emotional confidence vector output by the single modal model be: $F=[f_1,f_2,f_3,f_4,f_5,f_6]$, Among them, $f_1 \sim f_6$ correspond to confidence levels for six types of learning emotions: focus, pleasure, confusion, fatigue, anxiety, and exhaustion, with a range of values [0,1].

DOI: 10.54254/2755-2721/2026.LD38379

Visual modality output: F_v (lightweight EfficientNet model), behavioral modality output: F_b (learning time, accuracy, and operation frequency temporal statistics), text modality output: F_t (text emotion pre training model). Multi modal weighted fusion formula: $F_{final}=\alpha F_v+\beta F_b+\gamma F_t$, constraint condition: $\alpha+\beta+\gamma=1$. In this study, the optimal values obtained through experiments were: $\alpha=0.6,\beta=0.25,\gamma=0.15$. Parameter definition: Visual expressions are the core of emotions based on the highest weight, learning behavior serves as auxiliary verification, text interactive emotions serve as supplementary correction, and are in line with the characteristics of real learning scenarios.

Emotion classification judgment formula: $Label=argmax(F_{final})$, take the emotion label corresponding to the highest final confidence level as the student's real-time learning emotional state.

Personalized adaptive strategy decision-making model: Construct a dual dimensional quantitative scoring model of emotion index and cognitive ability index to achieve dynamic quantitative adjustment of learning strategies and break away from traditional fixed rule intervention.

Quantitative formula for cognitive ability:

$$S_c=\omega_1 A_{acc}+\omega_2(1-T_{time}/T_{std}) \tag{1}$$

A_{acc}: Correct answer rate for knowledge points, T_{time}: The duration of students answering questions; T_{std}: Standard answer time; The larger the S_c, the higher the level of cognitive mastery. Quantitative formula for emotional stability:Se=ffocus+fpleasure-fconfusion-ffatigue-fanxiety-ffatigue, the higher Se, the better the student's learning emotional state. The decision formula for learning adaptation intensity is $D_{level}=k_1 S_c+k_2 S_e$. Based on the Dlevel value range, it automatically divides the learning difficulty level into three levels: basic, improvement, and expansion, achieving personalized and accurate adaptation.

4. Dataset construction and data sample

To improve the adaptability of emotion recognition models in educational learning scenarios and mitigate the domain discrepancy between general-purpose emotion datasets and real classroom contexts, this study constructs data resources by combining public standard datasets with self-built classroom-scenario datasets. The public datasets mainly include the FER-2013 facial expression dataset, the IEMOCAP multimodal emotion dataset, and the GoEmotions text emotion dataset, which are used to provide foundational samples for visual, multimodal, and textual emotion recognition. The self-built dataset is obtained by collecting data offline from primary and secondary school classrooms and online self-directed learning scenarios; after de-identification and cleaning, an emotion dataset dedicated to educational scenarios is constructed. In terms of data organization, this study uniformly sets six fields: sample ID, visual features, behavioral features, textual features, emotion labels, and confidence, to support the training and validation of multimodal emotion recognition models. Specific data examples are shown in Table 1.

Table 1. Data examples

Sample ID	Learning behavior characteristics	Text interactive content	Identify emotional labels	confidence level
001	Stuck answer sheet, repeated page review, dwell time of 90 seconds	I don't understand this question, I don't know how to start	perplexed	0.92

DOI: 10.54254/2755-2721/2026.LD38379

Table 1. (continued)

002	Quick answer, 95% accuracy, proactive practice	This type of question is very simple, I have already mastered it	Pleasure/Focus	0.94
003	Idle page, reduced mouse operation, learning time exceeding 40 minutes	I don't want to learn anymore, I feel very tired	Fatigue/fatigue	0.89
004	Frequent page switching and sharp drop in answer accuracy	Always making mistakes, I feel like I can't learn	Anxiety	0.90

5. Research methods and experimental design

5.1. Research design

This study develops an overall experimental design centered on validating the effectiveness of an AI emotion recognition-based personalized learning support system, adopting an empirical research method based on controlled experiments to ensure the scientific rigor and objectivity of the results. The study follows a closed-loop technical route consisting of theoretical analysis, data construction, model optimization, system development, empirical validation, and summary and iteration. In the preliminary stage, a literature review is conducted to synthesize the current research landscape, identify existing research gaps and the innovations of this paper, and determine the core technical scheme for a multimodal emotion recognition model and personalized adaptive intervention. Two parallel classes in Grade 8 of a middle school with no significant differences in learning conditions are selected as the experimental participants, totaling 92 students, who are divided into an experimental group and a control group, with 46 students in each group. The two groups are consistent in baseline levels of prior academic achievement and learning ability, thus meeting the criteria for a controlled experiment. This study sets whether the AI emotion recognition and adaptive learning module is deployed as the independent variable, and takes emotion recognition accuracy, students' academic performance, learning emotional state, and learning satisfaction as the core dependent variables. Meanwhile, variables such as teaching resources, teaching progress, instructor, learning duration, and equipment environment are held constant to avoid interference from irrelevant variables on the experimental results.

5.2. Experimental procedure

The experiment lasted for 8 weeks and was carried out in an orderly manner according to standardized procedures throughout. In the early stage of the experiment, the researchers completed the screening of experimental samples, uniformly conducted a pre-test of students' academic performance, simultaneously distributed and collected the Adolescent Academic Emotions Scale and the Personalized Learning Satisfaction Scale, screened valid questionnaires, and completed data organization and entry, thereby ensuring baseline equivalence between the two groups. In the data preprocessing stage, public datasets and the self-built learning emotion dataset were integrated; data denoising, de-identification, cleaning, and annotation were completed; and dataset partitioning was carried out to provide data support for model training. In the system and model preparation stage, a layered system architecture was built; core modules for emotion collection, intelligent recognition, personalized adaptation, and data visualization were developed; a multimodal weighted fusion algorithm was designed; model training, parameter tuning, and functional debugging were completed; and stable system operation was ensured through stress testing and privacy optimization.

DOI: 10.54254/2755-2721/2026.LD38379

In the formal experiment stage, the experimental group used the AI emotion recognition personalized learning system developed in this study, while the control group used a traditional personalized learning system without emotion perception. The teaching process was followed throughout, system operation status and student learning behavior data were recorded, and various problems during the experiment were identified and resolved in a timely manner. After the experiment, post-tests of academic performance, learning emotions, and learning satisfaction were uniformly organized, and all-dimensional experimental data collection was completed.

5.3. Experimental evaluation and result analysis

This study establishes experimental evaluation metrics from two dimensions: technical performance and teaching effectiveness. At the technical level, multimodal emotion recognition accuracy and model inference latency are used as the core metrics. At the teaching level, students' academic performance, positive and negative learning emotions, and learning satisfaction are used as the evaluation bases. After the experimental data are collected, all valid data are cleaned, outliers are removed, and standardization is performed. SPSS 26.0 statistical software is used to conduct independent samples t-tests and paired samples t-tests to quantitatively analyze differences in various data between the experimental group and the control group before and after the experiment. The expected experimental results are as follows: the accuracy of the multimodal recognition model in this study is no lower than 88%, and the inference latency is no more than 200 ms, with performance superior to that of traditional unimodal models. Meanwhile, students in the experimental group show significantly improved academic performance, a marked reduction in negative learning emotions such as anxiety and burnout, and significantly higher learning satisfaction than the control group (Table 2). Finally, by integrating model operation data, system usage data, and empirical statistical results, the application effectiveness of the system is comprehensively evaluated, its advantages and existing shortcomings are analyzed, and the core research conclusions are summarized, thereby providing solid data support for the research outcomes of the project.

Table 2. 8-week teaching control experiment

indicator	control group	experimental group	difference
Post test average score of academic performance	62.41±8.36	69.73±7.92	$p<0.05$,significantly improve
Score on the Learning Positive Emotions Questionnaire	2.87±0.52	3.42±0.48	$p<0.05$
Negative emotions (anxiety, frustration)	3.21±0.64	2.63±0.57	significant decline
System satisfaction	—	78.6% of students believe that emotional intervention is helpful for learning	—

6. Feasibility analysis

This study is based on academic emotion theory, affective computing theory, and adaptive learning theory, with a mature theoretical framework and ample literature support; multimodal fusion, lightweight model optimization, and reinforcement learning-based adaptive strategies are all mainstream research directions, with no theoretical barriers. At the technical level, lightweight EfficientNet, multimodal late fusion, and temporal statistical methods are all open-source and

DOI: 10.54254/2755-2721/2026.LD38379

mature technologies. Model derivation, parameter tuning, and module design have been completed, enabling millisecond-level real-time recognition. Deployment is possible on ordinary computers and tablets, and local processing and de-identification mechanisms ensure privacy and security. In terms of data, public datasets are open-source and abundant, while self-collected classroom data are authentic and comprehensive, with standardized annotation and sufficient samples, which can meet the needs of model training and validation. At the practical level, the partner schools have well-developed smart teaching environments. The system fits primary and secondary school teaching scenarios, and the intervention strategies conform to educational principles. It can be directly applied to online self-directed learning and offline smart classrooms, and has strong applicability and promotion value.

7. Conclusion

This paper takes a personalized learning support system based on AI emotion recognition as its research object, reviews theories related to smart education, AI emotion recognition, and personalized learning, analyzes the current application status and prominent problems of existing intelligent learning systems, and completes the system architecture, functional module design, and logical verification. The research shows that most current intelligent learning systems focus on learning behavior analysis and resource recommendation, lack an emotion perception component, and adopt relatively homogeneous personalized adaptation methods. To address this, this study constructs an integrated system framework comprising emotion recognition, data analysis, resource recommendation, intelligent intervention, and effect feedback. It dynamically allocates learning resources and emotion regulation strategies by integrating multidimensional data such as learning progress and question-answering performance, thereby compensating for the homogeneous recommendation and insufficient emotional intervention of traditional systems. The empirical results show that personalized support based on AI emotion recognition is feasible, can improve the learning experience, and promotes the transformation of learning services from behavior-level personalization to human-centered intelligent services.

However, the experimental scenarios of this study are limited to online self-directed learning and do not cover offline classrooms or group collaboration. The sample coverage is limited, and there is a lack of long-term, large-scale pilots. The accuracy of emotion recognition and its cross-grade adaptability still need to be tested. Future research can expand to diverse teaching scenarios and samples, introduce multimodal fusion recognition and iteratively optimize algorithms, carry out long-term teaching pilots, improve the closed loop of emotional intervention, and form a scalable smart education solution.

References

[1] Zhou, J., Ye, J., & Li, C. (2021). Multimodal learning emotion computing: Motivation, framework, and suggestions. *Research on Electronic Education*.

[2] Yu, S., Shan, J., & Yan, H. (2024). Emotional computing empowers personalized teaching: Logical framework, problem deconstruction, and diverse reshaping. *Modern Distance Education*.

[3] He, J., Liu, Y., & He, Z. (2018). Research progress on multimodal emotion recognition. *Computer Application Research, 35*(11), 3201-3205.

[4] Artif. Intell. (2024, November 28). Implementation of deep reinforcement learning models for emotion detection and personalization of learning in hybrid educational environments. *Artif. Intell., 7*.

[5] Siddiqui, M. F., & Shah, S. A. (2023). A systematic review of multimodal emotion-aware adaptive learning systems. *Computers & Education, 194*, 104689.