



Research on the Mechanisms for AI-driven Students' Social-Emotional Learning

Yajin Qiu*

School of International Digital Economy, Minjiang University, Fujian, China

*3337531093@qq.com

Abstract. The integration of Artificial Intelligence (AI) into social-emotional learning (SEL) has drawn increasing attention due to its potential to enhance educational support for students' emotional development, though the application pathways, effectiveness, and influencing factors of AI in this context remain underexplored. This study aims to examine AI application approaches—including types and functions of AI, intervention efficacy, and key influencing factors—while offering theoretical and practical insights for developing AI-driven SEL tools. Through a comprehensive analysis of nearly 20 relevant studies. The findings indicate that AI supports SEL primarily through three pathways: multimodal emotion recognition, adaptive learning systems, and immersive simulation scenarios such as Augmented Reality (AR) and Virtual Reality (VR). Key influencing factors include the accuracy and cultural adaptability of algorithms, accessibility of professional teacher training, involvement of caregivers, and policy support related to ethical and equitable implementation. This study provides a solid theoretical foundation and practical guidance for future research and educational practice in AI-enhanced SEL.

Keywords: Artificial Intelligence, Social-Emotional Learning, Emotional Recognition,

1 Introduction

SEL is the process by which individuals develop self-awareness, emotional regulation, empathy, and interpersonal skills [1]. It has been widely acknowledged that SEL is crucial for both academic achievement and real-world success [2]. While traditional SEL relies on teacher guidance, AI technology has created new opportunities. For example, AI emotional computing technology employs multimodal techniques to capture students' emotions in real time. Javed developed a multimodal emotion assessment system that provides teachers with real-time student data to adjust classroom pacing accordingly [3]. Second, combined with AR-designed expressions, it reinforces emotional recognition activities tailored for children, enhancing identification accuracy [4]. Thirdly, AR and human-computer interaction technology support immersive social practices, for example, Almakky designs AR-based interactive simulations, such as scenarios mimicking social interactions, to help students with learning disabilities (LD)

© The Author(s) 2026

M. Kalogiannakis et al. (eds.), *Proceedings of the 2025 5th International Conference on Education, Language and Art (ICELA 2025)*, Advances in Social Science, Education and Humanities Research 997,

https://doi.org/10.2991/978-94-6463-988-9_8

practice empathy skills [5]. Besides, AI conversational agents can provide support for emotional development. Specifically, they can help students practice polite expressions and manage their emotions [6].

Although AI technologies offer multiple paths and approaches to current SEL educational practices similar to those mentioned above, there are still some issues and challenges at the level of research and practice. First, the accuracy of emotion recognition algorithms is inconsistent across diverse cultural and social settings [3]. This can lead to misinterpretations, for instance, in accurately identifying more subdued or reserved emotional expressions prevalent in some Eastern cultures. Second, there is a lack of teacher-technology collaboration. Third, existing AI-supported SEL tools are predominantly designed for neurotypical students, with limited applicability for special populations such as children with HFA and autism spectrum disorder (ASD).

2 Research Question

Against this backdrop, this paper seeks to address the following research questions.

(1) What types of AI technologies can effectively foster students' social-emotional competencies in educational settings?

(2) Which specific technological functions are most effective in enhancing these competencies across different implementation contexts?

(3) Can AI technologies also support the SEL of special student populations (e.g., those with LD or ASD)?

(4) What factors facilitate or hinder the effectiveness and adoption of AI-supported SCL?

3 The Type of AI Technology Applied in the SEL

3.1 Core Functions for Supporting SEL

Research indicates that AI-driven SEL tools can be architected around three core theories: emotion recognition, adaptive intervention, and scenario-based practice. The foundational theory involves emotion recognition systems, where tools such as emotion-aware chatbots serve as the primary interface. These systems utilize multimodal technology, integrating visual (e.g., facial expression recognition), auditory, and textual signals to identify a student's emotional state. The analyzed emotional data is then output to instructors, assisting them in making timely adjustments and providing targeted support to improve student well-being [7]. Similarly, Lin and Chen's AI-integrated educational platform incorporates an adaptive learning system that adjusts content difficulty based on student performance, an intelligent tutoring system (ITS) that delivers personalized feedback and Q&A, and predictive analytics to forecast academic outcomes. Together, these functionalities help alleviate negative academic emotions stemming from content mismatch or accumulated unanswered questions [8]. Furthermore, AI-driven conversational systems simulate authentic conversational scenarios embedded with emotion regulation support, thereby assisting English as a Foreign Language

(EFL) students in developing both social competencies and emotional management skills [9].

3.2 Core Technologies for Supporting SEL

Core algorithmic technologies can be broadly categorized into SEL-specific technologies, AI agent technologies, AR technologies, and AI-driven educational system technologies (See Table 1.)

Table 1. The type of AI technology applied in the SEL

Technology Category	Core Algorithm / Model / Interaction Logic	SEL Application Value
SEL-Specific Technologies	Multimodal affective computing (facial + vocal + textual); HCI optimization (multi-input); custom CNN (FER-2013, 7 emotions, ~70% accuracy); AR social simulation & emotional agents	Early intervention [7]; reduce technophobia & enhance acceptance [5]; support emotional awareness in autism [10]; enhance empathy & authentic interaction [11]
AI Agent Technologies	Multimodal weighted fusion (facial/vocal/textual/physiological signals); deep learning-based hierarchical feature fusion; hidden emotion detection	Real-time emotional data for personalized instruction; reduce frustration, increase engagement & retention; support students with learning difficulties such as autism [12]
AR	Real-time AR expression capture & virtual avatar matching; sensory adaptation (lighting/color); real-world scenario modeling (e.g., supermarket, classroom)	Improve emotion recognition accuracy[13]; reduce sensory overload [5]; support transfer of social skills to real life [14]
AI-Driven Education Systems	Generative AI content customization (GPT-4 + MidJourney); real-time feedback (star rewards + voice encouragement + gentle error correction with demonstration)	Meet individualized SEL needs; improve emotion recognition and learning outcomes; practical reference for generative AI in education [4]

Among these, SEL-specific technologies rely on affective computing systems tailored to the characteristics of special populations. These systems aim to address the low accuracy of single-modality emotion recognition and the poor adaptability of traditional tools. For example, affective computing systems primarily use multimodal technologies, leveraging cameras to capture facial muscle movements and natural language processing (NLP) to analyze textual semantics and vocal intonation, thereby generating emotional analysis reports. These systems assist educators in identifying students' negative emotions, uncovering underlying affective states, and providing evidence to support the application of SEL technologies in educational settings [7].

Secondly, for students with LD, adaptive learning systems often incorporate Human-Computer Interaction (HCI) optimization strategies, such as text amplification, voice descriptions, and multiple input methods (voice, touch, gesture), combined with AR social stories and situational simulations. These not only enhance empathy but also effectively alleviate technophobia and improve technology acceptance [5].

Moreover, for children with ASD, custom convolutional neural networks (CNNs) trained on the FER-2013 dataset are used to recognize seven types of emotions with approximately 70% accuracy. Through features like an “emotion diary,” these applications help autistic children independently practice expression recognition and develop emotional awareness [10]. Additionally, by using emotional pedagogical agents and feedback interactions, affective computing systems guide students to express their inner feelings. This approach overcomes the limitation of mere emotion “recognition” and enhances the authenticity and depth of interaction [11].

First, AI agent technology employs multimodal weighted fusion algorithms to integrate facial expressions, vocal tone, text semantics, and physiological signals like EEG. Using deep learning, it detects hidden emotions where outward behavior contradicts inner feelings, providing teachers with real-time emotional data to support personalized intervention, enhance classroom engagement, and improve knowledge retention—delivering particular benefit to students with learning difficulties such as autism. [12]. Second, AR technology constructs highly immersive social scenarios (e.g., supermarket checkouts, classroom interactions) to support emotion recognition training and social rehearsal for children with autism. By using real-time AR expression capture and virtual avatar matching, along with sensory adaptations such as lighting adjustment and color modification, AR improves recognition accuracy and promotes the transfer of social skills to real-world situations [13][14]. Meanwhile, AR tools can also reduce sensory overload and technology anxiety among LD students through customizable interfaces and progressive task design, thereby ensuring more stable and accessible usage [5]. Last, AI-driven educational systems utilize generative AI models (e.g., GPT-4 for generating social stories and MidJourney for creating high-saturation visual elements) to achieve content customization that aligns with the “low-stimulus, high-recognition” preferences of children with HFA. Combined with real-time positive feedback mechanisms like star rewards and voice encouragement, as well as gentle error correction with demonstrations, these systems meet individualized SEL needs while offering practical references for the application of generative AI in education [4].

3.3 Product Applications

AI technology facilitates the development of various SEL products by integrating specific scenarios with SEL needs (see Table 2).

First of all, robot/agent products address limitations such as “single-scenario use and delayed feedback” through multimodal emotion recognition and scenario-based social intervention [1]. Furthermore, the DOCO robot exemplifies this by providing bullying risk intervention and self-regulation recommendations based on emotional analysis [15]. Then, on the aspect of immersive learning. VR creates immersive environments for conflict mediation [12]. Meanwhile, AR products simulate real-life social situations,

allowing children with autism to practice social skills in a controlled setting [14]. Furthermore, AR-based object scanning and real-time facial expression analysis can enhance user engagement [13]. Besides, mobile products emphasize "fragmented training and long-term emotion recording" through lightweight personal SEL tools. Systems like EmoPal effectively improve users' emotion recognition abilities via emotion diary logging and tiered emotion-matching games [10]. Last, educational platforms have been widely adopted by both teachers and students. For students, these platforms can achieve scalable SEL through academic-emotional integration and instructional synergy, using algorithms that monitoring their emotional states and adjust task difficulty in real time[8]. Furthermore, customized systems designed for children with HFA significantly improve emotion recognition via structured training [4]. For teachers, AI-powered emotion reporting systems provide valuable insights, enabling them to optimize instructional strategies and promote the scaling of SEL across diverse classrooms [11].

Table 2. AI drives the classification of social-emotional learning products

Category	Key Features
Robot/agent	Multimodal emotion recognition [1]; Bullying intervention [15]
Immersive technology	VR conflict mediation [12]; AR social scenario simulation [14]; object scanning, real-time facial expression analysis [13]
Mobile products	Emotion diary, tiered matching game, lightweight training [10]
Educational platforms	Academic-emotion integration [8]; Autism-specific training [4]; Emotion reports [11]

4 The Function of AI in the Virtual Learning Environments

4.1 Expression Recognition Function

In virtual learning environments, Facial expression recognition (FER) technology identifies learners' emotions by analyzing facial expressions and converts this information into quantifiable data, providing a basis for instructors to adjust their teaching pace and implement interventions[3]. Furthermore, continuous monitoring through multimodal models helps identify negative emotions like anxiety both during and after class, allowing teachers to pinpoint high-risk students and offer proactive support for their SEL development [7].

The application of FER effectively extends to specialized educational needs. In AR environments, for example, systems capture the expressions of both children and caregivers. A key feature of these systems is a real-time feedback mechanism that compares a user's expression with standardized emotional templates—a design particularly beneficial for helping autistic children form fundamental "expression-emotion" associations. By providing immediate visual feedback through computer vision, these tools reinforce emotional understanding via imitation and practice [13].

Technologically, facial expression recognition systems often leverage tools like the Google Cloud Vision API and custom Convolutional Neural Network (CNN) models trained on specific datasets (e.g., FER-2013) to enable automated emotion classification [10]. To enhance reliability, some frameworks integrate multimodal data such as electroencephalography (EEG) to provide complementary physiological insights and significantly improves the accuracy of emotion assessment [11].

Together, these technologies establish a robust multidimensional SEL framework, supporting applications from real-time classroom intervention to long-term emotional documentation for both general and specialized student populations.

4.2 Emotion Recognition, Analysis, and Feedback Function

In virtual learning environments, emotion recognition, analysis, and feedback technologies are integrated through multimodal data to accurately identify emotions and provide targeted support for SEL.

First, the integration of multimodal data is crucial for accurate emotion identification. Jave's research demonstrated that AI-driven systems analyze vocal tone, interaction patterns, learning behaviors, and other multidimensional data to determine emotional states with high accuracy [3]. This multimodal approach overcomes the limitations of single-mode identification, captures subtle emotional shifts, and facilitates personalized SEL interventions through continuous behavioral monitoring.

Second, behavioral analytics are increasingly being utilized in this field. For instance, some research found that examining students' typing rhythms and interaction habits can reveal underlying emotional states such as academic frustration and worry [7]. Such analysis offers educators valuable insights into emotional patterns, enabling instructional adjustments. Furthermore, Zong and Yang's research findings emphasized that AI-driven emotion analysis can identify student anxiety and motivation levels, generate visualized emotional analytics, and provide classroom-level insights to support targeted group-based SEL activities [9]. Moreover, AI-powered tools, such as the Amazon Alexa-Speak Speaking Assessment System [16] and the pre-trained RoBERTa model [15], enable real-time feedback in specific learning contexts.

In summary, these technologies work together to enable multimodal recognition, in-depth analysis, and timely feedback, forming an end-to-end "Recognition–Analysis–Intervention" loop that effectively supports SEL implementation in digital learning settings.

4.3 Intelligent Dialogue and Understanding Function

Studies have explored the underlying technologies and mechanisms of intelligent dialogue systems.

Firstly, research by Saifuddin found that NLP-enabled systems achieve human-like interactions and act as “social-emotional practice partners,” offering multi-dimensional support. Their study showed that chatbots with fault-tolerant dialogue mechanisms can accurately interpret student expressions and deliver friendly interactions [17]. Equipped with appealing personalities, these chatbots help reduce hesitation and encourage open communication of emotions. Furthermore, by integrating multimodal emotion recognition—including facial, vocal, and movement analysis—along with interaction histories, these systems adapt to context, communicate naturally, and provide emotional support without teacher intervention, making students feel understood and accepted [12]. Moreover, conversational agents are designed to support SEL through interactive scenarios such as "Listening Quiz," "Chit Chat," and "Good Table Manners Made Easy," which target competencies like self-awareness, relationship skills, and responsible decision-making. However, analysis reveals that commercial skills remain limited by shallow interactions, lack of contingent feedback, and minimal adaptability, restricting their efficacy in meaningful SEL development [6].

In summary, these technologies work together to establish a foundation for "daring to express," provide emotional support so students feel "understood," and deliver practical "knowledge-based skills," collectively forming a comprehensive SEL support system in virtual environments.

4.4 Personalized Learning Support Function

Personalized learning support relies on AI technology to dynamically adapt to students' emotions and learning characteristics, thereby providing tailored assistance for diverse learners with varying needs and traits.

Firstly, adaptive learning systems can analyze students' emotional states and learning performance to dynamically adjust task difficulty. For instance, when anxiety is detected, the system reduces task difficulty to alleviate frustration. Conversely, when students show confidence, it increases the challenge appropriately to stimulate creativity. Moreover, such systems offer personalized tasks and real-time feedback, promoting the synergistic development of self-regulation and learning motivation [8].

Secondly, in the English learning context, AI-enhanced SEL frameworks modulate learning content and pace based on the student's emotional state, thereby improving concentration and sustained engagement. These systems also support the development of social-emotional skills through customized interventions. For example, they may guide speaking practice when low anxiety is detected or use AI to simulate social scenarios to help manage higher anxiety, effectively integrating SEL into language learning [9]. Furthermore, for children with ASD, AR training scenarios can be designed to dynamically adjust task difficulty based on facial expression recognition. This enables personalized learning content that aligns with the children's perceptual characteristics, avoids sensory overload, and gradually improves their emotion recognition ability

through a structured process [13]. Similarly, VR technology can simulate real-world social environments like supermarkets or campuses, customized for children with ASD. These immersive environments provide a safe space to practice social skills, helping to reduce anxiety, enhance motivation, improve learning efficiency, and promote the generalization of skills to real-life situations.[14].

These solutions cater to a wide range of learners, from students to children with special needs, offering broad adaptability and specialized support, and contributing to an integrated system of personalized SEL in virtual learning environments.

5 The Effectiveness of AI on Supporting SEL of Special People

This study focuses on two special populations—ASD/HFA and LD—to examine the value of AI-driven targeted interventions.

For children with ASD/HFA, a multimodal emotion recognition system integrating facial expressions, speech, and physiological signals was developed, along with adaptive behavioral modeling algorithms to dynamically interpret atypical emotional expressions and generate socially appropriate responses [12]. This supports the adaptation of children’s emotional states and customization of learning content through adaptive learning systems, helping address challenges in interpreting emotional signals in educational contexts [11]. Furthermore, AI-powered technologies, such as AR and VR, are leveraged to create adaptive virtual scenarios and gamified content that simulate real-life social situations. These tools, exemplified by systems like EMooly that generate personalized social stories using GPT-4, progressively increase in complexity based on individual profiles. This approach reduces social anxiety and promotes active social engagement by enhancing the ability to interpret emotions in dynamic social contexts [13][14].

In summary, the effectiveness of AI interventions for SEL in special populations lies in its ability to combine “technology accurately adapted to special physiological and cognitive characteristics” with “scenarios to reduce the pressure of practice.”

6 The Factors During the Progress of AI-Supported Social-Emotional Learning

The efficacy of AI-supported SEL is concurrently constrained by a constellation of internal technological limitations and external environmental factors.

6.1 Internal Factors

On the internal front, the foundational technology faces significant challenges across three core areas. First, in terms of accuracy, emotion recognition algorithms exhibit substantial errors when processing complex emotional states (e.g., embarrassment) and cross-cultural expressions [10]. This issue is exacerbated by insufficient multimodal

data fusion and ongoing debates surrounding the universality of emotional expression rooted in Basic Emotion Theory [18]. Research indicates that non-Western cultural samples often exhibit higher misclassification rates, necessitating model optimization with localized data [11]. Second, in terms of technical stability issues, such as hardware malfunctions and software latency can critically disrupt interventions. This is particularly detrimental for sensitive populations like children with ASD, as it erodes their trust and engagement [13][14].

Finally, and perhaps most critically, the prevalent "one-size-fits-all" design paradigm fails to adapt to vast individual differences in sensory preferences and cognitive profiles. This limitation severely curtails the intervention's effectiveness and underscores a pressing need for robust personalization strategies [13][14].

6.2 External Factors

Externally, the transition from technical feasibility to practical effectiveness is non-trivial, hindered by several barriers. A primary obstacle is the lack of sustained user engagement, often stemming from shallow interaction designs. This is exemplified by the finding that 40% of SEL-oriented conversational agents employ simplistic question-and-answer dialogues with content that fails to align with user interests [6]. This problem is further compounded by AI-generated content that lacks personalized feedback, leading to higher user churn rates [3].

Another determining factor is the critical role of human facilitators, specifically teachers and caregivers. Their ability to operationalize these tools and accurately interpret the generated data is paramount for transferring learned SEL skills to real-world contexts. However, a widespread lack of specialized training severely hampers this potential [5]. Empirical studies indicate a strong positive correlation between teacher competency in using AI tools and student outcomes, whereas misinterpretation of data by caregivers can lead to misalignment between home and school interventions [13][1].

7 Conclusion

The integration of AI-enhanced SEL frameworks represents a groundbreaking shift in the educational landscape. As this study demonstrates, AI not only enhances but fundamentally reconfigures SEL delivery through technologically and pedagogically innovative approaches. At a technological level, it facilitates precise, context-aware interventions by leveraging multimodal data and adaptive algorithms to provide real-time, personalized feedback within immersive scenarios. On a practical level, AI fosters a human-machine collaborative framework, undertaking data analysis and content delivery to free educators to focus on emotional guidance and real-world skill application. Nevertheless, the implementation of AI-supported SEL faces persistent challenges related to individual adaptability, ethical compliance, and equitable access, which currently constrain its scalability. Looking ahead, the continued evolution of AI-assisted SEL will depend on the creation of culturally responsive algorithms and evidence-based pedagogical frameworks to fully support the holistic development of every student.

Reference

1. Sethi, Surbhi Seema, and Kanishk Jain. "AI technologies for social emotional learning: recent research and future directions." *Journal of Research in Innovative Teaching & Learning* 17.2 (2024): 213-225.
2. Durlak, Joseph A., et al. "The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions." *Child development* 82.1 (2011): 405-432.
3. Javed, Sunbal, et al. "How AI can detect Emotional cues in Students, Improving Virtual Learning Environments by providing Personalized support and Enhancing Social-emotional Learning." *Review of Applied Management and Social Sciences* 8.2 (2025): 665-682.
4. Tang, Yilin, et al. "Emoeden: Applying generative artificial intelligence to emotional learning for children with high-function autism." *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024.
5. Almakky, Hana Ghazy. "Fostering Empathy and Social-Emotional Learning in Students with Learning Disabilities through HCI and Augmented Reality." *International Journal of Interactive Mobile Technologies* 19.5 (2025).
6. Fu, Yue, et al. "Social Emotional Learning with Conversational Agents: Reviewing Current Designs and Probing Parents' Ideas for Future Ones." *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6.2 (2022): 1-23.
7. Shifa, Salman, et al. "Emotional Intelligence Meets Artificial Intelligence Affective Computing and Student Well-Being Online Classrooms." *The Critical Review of Social Sciences Studies* 3.3 (2025): 271-289.
8. Lin, Haozhuo, and Qiu Chen. "Artificial intelligence (AI)-integrated educational applications and college students' creativity and academic emotions: students and teachers' perceptions and attitudes." *BMC psychology* 12.1 (2024): 487.
9. Zong, Yue, and Lei Yang. "How AI-enhanced social-emotional learning framework transforms EFL students' engagement and emotional well-being." *European Journal of Education* 60.1 (2025): e12925.
10. La Fauci De Leo, Andrea, et al. "A New AI Framework to Support Social-Emotional Skills and Emotion Awareness in Children with Autism Spectrum Disorder." *Computers* 14.7 (2025): 292.
11. Vistorte, Angel Olider Rojas, et al. "Integrating artificial intelligence to assess emotions in learning environments: a systematic literature review." *Frontiers in psychology* 15 (2024): 1387089.
12. Annapareddy, Venkata Narasareddy, et al. "Emotional Intelligence in Artificial Agents: Leveraging Deep Multimodal Big Data for Contextual Social Interaction and Adaptive Behavioral Modelling." *Jai Kiran Reddy, Emotional Intelligence in Artificial Agents: Leveraging Deep Multimodal Big Data for Contextual Social Interaction and Adaptive Behavioral Modelling (April 14, 2025)* (2025).
13. Lyu, Yue, et al. "EMooly: supporting autistic children in collaborative social-emotional learning with caregiver participation through interactive AI-infused and AR activities." *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8.4 (2024): 1-36.
14. Adako, Oyeyemi Patricia, Oluwafemi Clement Adeusi, and Peter Adeniyi Alaba. "Revolutionizing autism education: Harnessing AI for tailored skill development in social, emotional, and independent learning domains." *Journal of Computational and Cognitive Engineering* 3.4 (2024): 348-359.

15. Alves-Noreña, Alba Catherine, et al. "Technology and Emotions: AI-Driven Software Prototyping for the Analysis of Emotional States and Early Detection of Risky Behaviors in University Students." *Education Sciences* 15.3 (2025): 350.
16. Tajik, Aliakbar. "Integrating AI-Driven Emotional Intelligence in Language Learning Platforms to Improve English Speaking Skills through Real-Time Adaptive Feedback." (2025).
17. Saifuddin, Tasneem. "Chatbots on Emotional and Social Skills; Are they really improving Students Learning?." *Pakistan Research Journal of Social Sciences* 4.2 (2025).
18. Stark, Luke, and Jesse Hoey. "The ethics of emotion in artificial intelligence systems." *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 2021.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

