Results in Engineering 27 (2025) 105931

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Research paper

Bioengineering an AI-augmented platform for remote mental health interventions

Ju Yu Wu [a,b], Yu-Jie Chen [a], Yen-Feng Lin [c], Congo Tak Shing Ching [b,d,f,g,h], Hui-Min David Wang [b,d], Yi-Chung Chen [e], Yang-Chou Juan [e], Lun-De Liao [a,b,*]

[a] *Institute of Biomedical Engineering and Nanomedicine, National Health Research Institutes, 35 Keyan Road, Zhunan Town, Miaoli County 350, Taiwan*
[b] *Doctoral Program in Tissue Engineering and Regenerative Medicine, National Chung Hsing University, 145, Xingda Road, South District, Taichung City 402, Taiwan*
[c] *Center for Neuropsychiatric Research, National Health Research Institutes, 35, Keyan Road, Zhunan Town, Miaoli County 350, Taiwan*
[d] *Graduate Institute of Biomedical Engineering, National Chung Hsing University, 145, Xingda Road, South District, Taichung City 402, Taiwan*
[e] *Department of Computer Science and Engineering, National Chung Hsing University, 45, Xingda Road, South District, Taichung City 402, Taiwan*
[f] *Department of Electrical Engineering, National Chi Nan University*
[g] *International Doctoral Program in Agriculture, National Chung Hsing University*
[h] *Advanced Plant and Food Crop Biotechnology Center, National Chung Hsing University*

A R T I C L E I N F O

Keywords:
Teletherapy
Mental health therapy
Eye movement desensitization and reprocessing
Artificial intelligence

A B S T R A C T

This study developed and tested a new teletherapy system that uses artificial intelligence (AI) to improve remote psychological treatment and reduce therapists' workload. The system includes a case management platform and a video-based therapy platform with support for remote Eye Movement Desensitization and Reprocessing (EMDR). After each session, the system uses AI tools such as emotion recognition, speech-to-text, and a large language model to automatically create a session summary and report, including emotional changes and a transcript. Real-time eye movement tracking and distance monitoring help standardize remote EMDR procedures. The case management system improves efficiency by helping schedule sessions and reminding clients (rated 4.25±0.83 out of 5). In clinical trials, AI-generated summaries reduced report writing time by over 50 %. Less experienced therapists found the emotion recognition tool helpful for self-reflection (4.3 ± 0.64), while experienced therapists believed the EMDR tools improved their remote consultation skills (4.2 ± 0.75). User feedback shows the system is easy to use, with therapists rating it 4.83 and clients 4.33 for ease of use. Overall, the system improves access to mental health services, increases workflow efficiency, and supports professional development in digital therapy.

1. Introduction

Mental health concerns have become increasingly prominent in modern society, especially since the COVID-19 pandemic. Uncertainty can induce mental health issues, such as stress, anxiety, and depression [1]. In 2019, a cross-sectional study revealed that approximately one-fifth of children and youths globally were affected by mental disorders [2]. During the COVID-19 pandemic, the overall prevalence rates of anxiety and depression were 33 % and 28 %, respectively. The prevalence of anxiety and depression was highest among patients with pre-existing medical conditions and those with COVID-19, at 56 % and 55 %, respectively. China, Italy, Turkey, Spain, and Iran showed higher prevalence rates of anxiety and depression [3]. After the pandemic, 30 % to 40 % of people in the United States, Europe, and China with post-COVID-19 syndrome had a greater likelihood of experiencing depression and anxiety, sleep problems, and posttraumatic stress disorder (PTSD) [4,5]. This situation has resulted in a steady rise in the demand for psychological counseling services [6].

The need for psychotherapists is escalating, and educating qualified psychological practitioners requires rigorous professional training and extensive practical experience. This situation increases psychotherapists' workload and leads to longer working hours and limited resources [7,8]. This situation is further exacerbated by the uneven geographical distribution of mental health professionals, who are predominantly concentrated in urban areas, leaving rural populations with limited access to psychological services [9,10]. Psychotherapy delivered remotely

* Corresponding author at: Institute of Biomedical Engineering and Nanomedicine, National Health Research Institutes, 35 Keyan Road, Zhunan, Miaoli County, 35053, Taiwan.
E-mail address: ldliao@nhri.edu.tw (L.-D. Liao).

https://doi.org/10.1016/j.rineng.2025.105931
Received 11 April 2025; Received in revised form 10 June 2025; Accepted 22 June 2025
Available online 27 June 2025

via a well-supported platform might solve these problems.

Many studies have shown that video- and audio-based teletherapy can reduce depression, anxiety, and other mental health issues and that their effectiveness is similar to that of in-person services [11–13]. Furthermore, remote cognitive behavioral therapy (CBT) and remote eye movement desensitization and reprocessing (EMDR) can alleviate PTSD, stress, and depression [6,14–20]. EMDR can help people process stress-related traumatic memories by using eye movements or other bilateral stimulation during the treatment session [21,22]. Most of the literature on telepsychotherapy (especially EMDR) is focused on the period after the epidemic. The pre-COVID-19 literature primarily addresses telepsychological health promotion, such as studies by Backhaus, who published a systematic review showing that videoconferencing technology, which allows audio and video information to be shared concurrently across geographical distances, offers an alternative that may improve access to treatment [23]. However, no remote EMDR apps were recommended for use by a client before the pandemic [24]. Since the pandemic, online EMDR therapy has become widely popular, and guidelines that provide specific recommendations for safety, treatment considerations, and client selection have been published by many EMDR organizations [25–27].

An effective remote delivery system for psychological interventions using artificial intelligence (AI) technologies might be a viable solution to these problems. Teletherapy offers significant advantages for clients by eliminating travel time to healthcare facilities and providing greater scheduling flexibility. This increased convenience, combined with the reduced stigma associated with seeking mental health services remotely, increases individuals' willingness to pursue professional psychological assistance [28,29]. Furthermore, teletherapy transcends geographical boundaries, making quality mental health services accessible to populations in remote areas [30].

However, the digital nature of remote therapy introduces certain limitations, such as digital exclusion (e.g., a lack of digital equipment or skill to access online services), difficulty engaging people, and difficulty recognizing nonverbal cues. Specifically, because therapists are restricted to observing clients' emotional states through computer screens, they may overlook comprehensive emotional cues that would be apparent in face-to-face interactions. This constraint raises concerns among practitioners about potentially compromised therapeutic effectiveness [31–33]. The development of AI-assisted tools that can enhance therapists' ability to detect and interpret emotional states in virtual settings could address these limitations while preserving the benefits of remote service delivery.

Preliminary platforms, such as EMDR Remote, already provide remote eye movement command functions. However, no complete system integrates AI emotion recognition, summary generation, and eye tracking into a single platform, and clinical empirical evaluation is lacking. This study presents the development of an AI-enhanced remote mental health service platform that features innovative functionalities designed to optimize therapeutic workflows to reduce the workloads of psychotherapists, as well as easy access to overcome digital exclusion and incorporate clients' nonverbal cues. The system incorporates multiple AI technologies, such as deep learning and computer vision for facial expression analysis, natural language processing to organize session content, and precise eye-tracking technology to enable remote EMDR therapy. These technological integrations address key challenges in the delivery of remote psychological services.

The scope of this study is focused on the technological architecture and an evaluation of the initial usability and technical feasibility. The subsequent sections of this paper are structured as follows. First, we present a detailed examination of the system architecture and performance metrics. Next, we present a comprehensive analysis of the user experience and an efficacy-validated test based on sandbox testing to evaluate the results. Finally, we present real-world clinical trials to validate the efficacy of this study.

2. Materials and methods

2.1. System architecture

This study presents an integrated remote psychotherapy system that combines AI technology with teletherapy and remote EMDR therapy capabilities. The system architecture consists of two primary components, as illustrated in Fig. 1: a case management system and a remote psychotherapy consultation platform. The case management system can automatically remind clients of appointments. Furthermore, after the video treatment session, an AI-generated report can be downloaded that includes session transcripts, summaries, and an emotional analysis of the client. The remote psychotherapy consultation platform includes two embedded functions: video calling and an EMDR therapy controller. AI models process a video image of the client and full audio to generate the components of the report. Additionally, eye-tracking technology and a real-time distance detector ensure the accuracy and effectiveness of remote EMDR therapy. For better performance, the entire system is deployed using Amazon Web Services (AWS).

2.1.1. Case management system

The case management system was developed as a comprehensive platform for psychotherapists to manage client information and consultation appointments. It includes a secure registration process exclusively for qualified psychotherapists. Upon submission of a registration application, the system generates an automated notification while administrators verify the applicant's professional credentials. Following credential verification, the system provides the psychotherapist with secure login credentials, including an access URL, a username, and a temporary password.

The system architecture incorporates role-based access control that grants psychotherapists comprehensive case management capabilities. Upon authentication, psychotherapists can create and manage client profiles. The system automatically generates client account activation credentials, which are securely transmitted to clients via email. The case overview interface enables psychotherapists to access client demographics, including name, gender, age, and consultation history. Detailed client records are accessible through individual profile pages, where psychotherapists can modify information and access system-generated consultation reports.

The appointment management module facilitates flexible scheduling of consultation sessions. The system features an automated reminder system that operates at three intervals: one month, one week, and one day before scheduled appointments. All scheduling information is synchronized with the system calendar. Psychotherapists can access the teletherapy interface directly on the consultation day through the daily schedule dashboard.

The client interface features a simplified design with essential functionalities. Clients can view their consultation schedule, access the remote psychotherapy platform, and review their consultation history, including session timestamps and therapist information. The system includes password management capabilities for psychotherapists and clients to maintain account security. This architectural design prioritizes professional service quality and user accessibility and effectively supports psychotherapists in case management and the delivery of teletherapy services.

2.1.2. Remote psychotherapy system

The remote psychotherapy consultation platform integrates two core functionalities, video consultation and EMDR therapy, which are enhanced by AI technologies to optimize remote therapeutic outcomes. During video consultations, the integrated AI-powered facial emotion recognition system processes clients' facial expressions and provides psychotherapists with quantitative emotional data for a more precise assessment of clients' psychological states.

The system incorporates advanced voice processing capabilities and

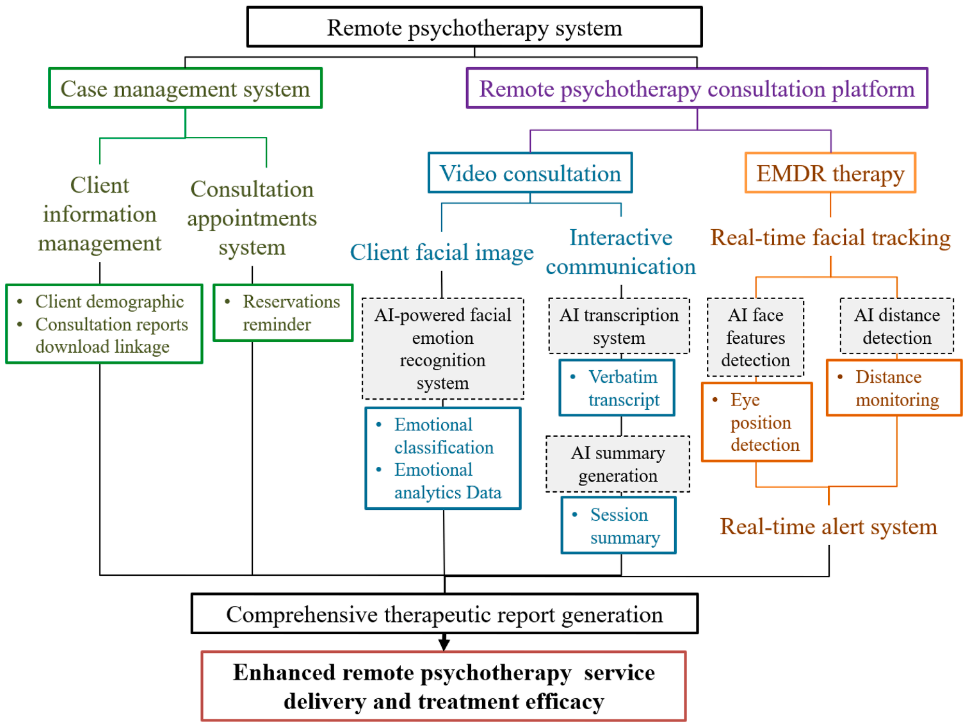


Fig. 1. System architecture of the integrated teletherapy platform. The system consists of two primary components: a case management system (left) and a teletherapy consultation platform (right). The case management system handles client information and appointment scheduling, while the teletherapy platform integrates video consultation and EMDR therapy functionalities. The platform incorporates multiple AI-powered features, including facial emotion recognition, speech transcription, and real-time monitoring systems (shown in dashed boxes). All session data are automatically compiled into a comprehensive therapeutic report, which enhances the delivery and efficacy of remote psychotherapy services.

automatically records client-therapist interactions during sessions. AI algorithms process these audio recordings to generate verbatim transcripts and session summaries, which provide therapists with comprehensive documentation of the therapeutic process.

The EMDR therapy module uses webcam technology to monitor clients' facial features, eye positions, and screen distance in real time. The system tracks eye movement compliance with the EMDR guidance stimulus and assesses movement effectiveness. When deviations are detected, the system generates immediate alerts that enable therapists and clients to make necessary adjustments to maintain the fidelity and effectiveness of treatment.

Following each session, the system automatically compiles the collected data into a comprehensive therapeutic report. This report includes client demographics, session duration, emotional analysis percentages, access to verbatim transcripts, session summaries, and EMDR eye-tracking compliance metrics. This integrated reporting system facilitates both the monitoring of treatment progress and informed therapeutic planning.

2.2. Technical implementation

2.2.1. Components of system architecture

The system was developed using a decoupled architecture that separates the frontend and backend components. The frontend interface was implemented using modern web technologies, including JavaScript, HTML, and CSS, to provide a responsive and interactive user experience. The backend services were developed using the Django framework in Python [34], which offers robust server-side functionality and database management. The system architecture can be categorized into three main components.

First, the frontend components handle real-time processing tasks, particularly those that require immediate feedback during therapy sessions. These include eye position detection and distance monitoring, which are implemented using the JavaScript version of the MediaPipe library [35]. This architecture decision ensures minimal latency for critical real-time features that are essential for effective EMDR therapy.

Second, the backend services manage postsession processing tasks using Python-based implementations. These include emotion analysis, speech-to-text conversion, session summary generation, and automated report creation. This approach allows for comprehensive data processing without impacting real-time performance during therapy sessions.

Third, the system uses RESTful APIs for communication between frontend and backend components, which enables efficient data exchange and system integration. The Django backend provides structured API endpoints for various functionalities, including user authentication, session management, and data processing services.

This architectural design ensures optimal performance for real-time interactions and complex postsession analysis while maintaining system scalability and maintainability. The entire system is deployed on Amazon Web Services (AWS) [36] and leverages its cloud infrastructure for reliable service delivery and scalability.

2.2.2. Security implementation and compliance

The system's security architecture leverages AWS cloud infrastructure to ensure data protection and regulatory compliance [36]. The security implementation involves multiple layers of protection and focuses mainly on safeguarding sensitive client information and data from therapeutic sessions.

Access control is implemented through AWS identity and access management (IAM), which manages user authentication and authorization. The system employs role-based access control (RBAC) to assign specific permissions to different user types, including administrators, psychotherapists, and clients. Multifactor authentication (MFA) is mandatory for all psychotherapist accounts to enhance access security.

The system uses AWS Key Management Service (KMS) to manage encryption keys for data encryption. Transport Layer Security (TLS 1.3) encryption secures all data transmission between clients and servers. All stored data, including client records and session recordings, are encrypted at rest using AES-256 encryption standards.

The video consultation module implements end-to-end encryption for all therapeutic sessions to ensure that video and audio streams remain secure and private. Session recordings are immediately encrypted upon capture and stored in Amazon S3 with server-side encryption enabled. Access to these recordings is strictly controlled through time-limited, secure URLs.

Database security is maintained through Amazon RDS encryption, automated backup systems, and point-in-time recovery capabilities. The AWS Web Application Firewall (WAF) enhances the system's network security and protects against common web exploits and attacks. The virtual private cloud (VPC) configuration ensures network isolation and protection.

To maintain compliance with healthcare regulations, the system implements comprehensive audit logging and monitoring through AWS CloudTrail and CloudWatch. All system access and data modifications are logged and monitored for security analysis and compliance reporting. Regular security assessments and penetration testing are conducted to identify and address potential vulnerabilities.

2.2.3. Facial expression analysis for emotion recognition

2.2.3.1. Dataset. This study validated the emotion recognition model using the Taiwan Corpora of Chinese Emotions and relevant psychophysiological data. The dataset includes seven basic emotional categories: neutral, happiness, sadness, anger, disgust, fear, and surprise (Fig. 2(A)) [37].

2.2.3.2. Data distribution. The dataset was randomly partitioned using an 80–20 split ratio, which resulted in 977 images for the training set and 244 for the validation set. This distribution maintained a relative balance across emotional categories to facilitate stable model training.

2.2.3.3. Model architecture. The study employed the Swin Transformer as the foundational architecture with a transfer learning strategy. Swin Transformer is a vision transformer-based deep learning model with a hierarchical window-based self-attention mechanism that enables effective processing of local and global image features [38]. Through transfer learning, we initialized the model with pretrained weights and fine-tuned it specifically for Chinese facial emotion recognition tasks to enhance performance on the target objective.

2.2.3.4. Data preprocessing and augmentation. To ensure data consistency and stability, all input images underwent standardization procedures, including dimension adjustment and pixel value normalization. Data augmentation techniques, such as random flipping and rotation methods, were implemented to enhance the generalizability of the model.

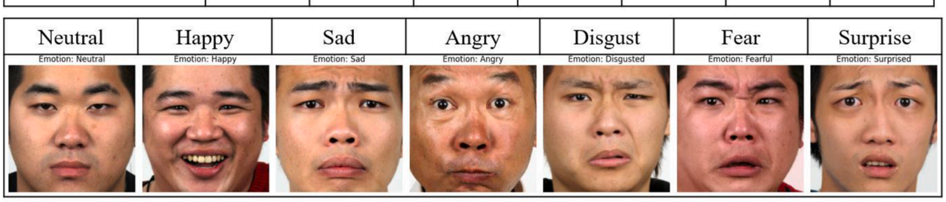
2.2.3.5. Model interpretability. To elucidate the model's decision-making process, we implemented attention mechanism visualization techniques. This approach revealed the regions of focus during emotion recognition by the Swin Transformer and facilitated understanding of the model's operational mechanisms and validation of its performance.

2.2.4. Speech recognition and transcription

This study evaluated the Chinese transcription capabilities of three mainstream cloud-based speech-to-text services. The assessment utilized a 6-minute counseling session video-to-test model of Google Cloud [39], Microsoft Azure [40], and AWS platforms [41]. The evaluation metric was the recognition accuracy rate (RAR) and the characteristic error rate (CER) instead of the word error rate (WER) because Chinese word

(A) Distribution of emotion categories in Taiwan Corpora of Chinese Emotions and representative samples

	Neutral	Happy	Sad	Angry	Disgust	Fear	Surprise
Training set	55	181	157	159	158	121	146
Validation dataset	14	45	39	40	40	30	36



(B) Eye-tracking measurement configuration: distance and position parameters

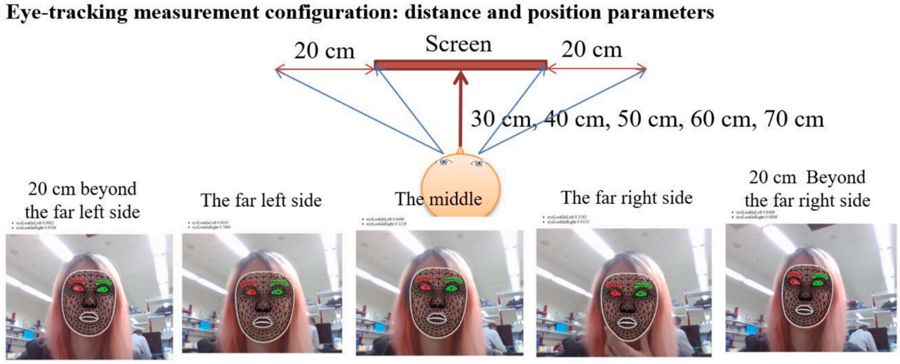


Fig. 2. Dataset overview and real-time eye-tracking experimental setup. (A) Dataset characteristics and representative samples from the Taiwan Corpora of Chinese Emotions. The top panel shows the distribution of emotion categories across the training ($n = 977$) and validation ($n = 244$) datasets. The bottom panel displays representative facial expressions for each of the seven basic emotions: neutral, happy, sad, angry, disgust, fear, and surprise. (B) Schematic representation of the eye-tracking measurement configuration. The setup illustrates the spatial parameters for system validation, including five viewing distances (30 cm, 40 cm, 50 cm, 60 cm, 70 cm) and five horizontal gaze positions (20 cm beyond left, far left, middle, far right, and 20 cm beyond right of screen). The bottom row shows facial landmark detection results across different head positions, with visible left (red) and right (green) eye detection markers.

boundaries are nondiscrete. The system chose the best audio-to-text model for operational deployment.

2.2.5. Automated summary generation

The automated summarization capabilities were evaluated using 25 counseling dialogs from "Helping Skills: Facilitating Exploration, Insight, and Action (3rd Edition)". The evaluation process comprised two phases: initial assessment and optimization. ChatGPT-3.5 [42] and Claude-3.5 Sonnet [43] generated summaries during the initial assessment using basic prompts. A licensed psychologist evaluated these summaries in relation to four criteria (completeness, terminology accuracy, structural clarity, and practical utility), each of which was scored on a 5-point scale. In the optimization phase, enhanced prompts were generated using psychologist-modified summaries and original dialogs.

2.2.6. Eye movement detection and analysis

In this study, a two-stage experiment was designed to evaluate the performance of the eye movement detection system. In the first stage, MediaPipe [35] was used to locate facial features at five distances (30 cm, 40 cm, 50 cm, 60 cm, 70 cm) and five sight directions (20 cm beyond the far left side of the screen, the far left of the screen, the middle, the far right of the screen, 20 cm beyond the far right side of the screen), and the confidence score of the position of both eyes was recorded (shown in Fig. 2(B)). The ratio value was calculated by Eq. (1) and drawn into a box-and-whisker plot to establish the criterion for eye position. In the second phase of the experiment, the eye movement detection module was combined with the EMDR module, and group tests were conducted at distances of 25, 40, and 75 cm to evaluate the accuracy of the eye movement tracking module.

$$\text{Ratio} = \ln \frac{\textit{confidence score of LEFT eye}}{\textit{confidence score of RIGHT eye}} \tag{1}$$

2.2.7. Real-time distance monitoring

This study used MediaPipe to locate the two iris positions. Based on the average human iris distance of 11.7 mm [44], the distance between the user and the camera was calculated by Eq. (2). The experiment was conducted at a distance of 30 cm to 80 cm, and the actual distance and predicted distance of the system were recorded every 10 cm. Each distance was measured 10 times and averaged, and its correlation was analyzed.

$$\text{Distance} = \frac{\textit{camera focal length} \times \textit{actual iris distance}}{\textit{detected iris distance} \times 10} \tag{2}$$

2.3. Experimental design for the user experience evaluation

2.3.1. Participant recruitment and demographics

This study recruited 24 subjects for system evaluation, 16 of whom participated in system usability testing and 8 of whom participated in actual consultation testing (Table 1). All subjects had basic computer operating skills and a stable internet connection. In the system usability testing group, 62.5 % (10 people) of the subjects were women, and 37.5 % (6 people) were men. In terms of professional background, 43.8 % (7 people) of the subjects were licensed psychologists, 25 % (4 people) had completed internships, and 31.3 % (5 people) were other types of professionals with program backgrounds. Notably, among the subjects in this group, 50 % had EMDR certification, and 50 % did not have EMDR certification (8 people each). The distribution of clinical experience was relatively scattered, with 31.3 % (5 people) of subjects having no experience and 25 % (4 people) having >10 years of experience. The age distribution was dominated by subjects aged 31–40 years (37.5 %), followed by those aged 20–30 years and over 50 years (31.3 % and 25 %, respectively).

Most of the patients in the consultation test group were female (87.5 %). The members of this group were all psychologists, of whom 50 % were licensed, and 50 % had completed internships (4 each). Seventy-five percent of the members did not have EMDR certification, and subjects had <1 year or 1–5 years of clinical experience (37.5 % each). With respect to the age distribution, 75 % of subjects were in the 20–30 age range, while the remaining 25 % were aged 41–50 years. The background distributions of the two groups of subjects had specific differences that helped them evaluate the system performance from different perspectives. Recruitment was conducted through the Taiwan Eye Movement Desensitization & Reprocessing Association and related professional societies.

This study was approved by the institutional review board (IRB) of the National Health Research Institutes in Taiwan (IRB number: EC1130508-E), and all research procedures strictly complied with research ethics regulations. All subjects signed informed consent forms, and the experimental data were processed anonymously and stored securely. The audio and video of the consultation process were used only for subsequent analysis and will be destroyed after the research is completed.

Table 1
Demographic and professional characteristics practitioners ($n = 24$).

Characteristic	In total ($n = 24$)	In system usability test group ($n = 16$)	In actual consultation test group ($n = 8$)
Gender, n (%)			
Female	17 (70.8)	10 (62.5)	7 (87.5)
Male	7 (29.2)	6 (37.5)	1 (12.5)
Professional status, n (%)			
Licensed psychologist	11 (45.8)	7 (43.8)	4 (50.0)
Post-internship psychologist	8 (33.3)	4 (25.0)	4 (50.0)
Other professional	5 (20.9)	5 (31.3)	0 (0.0)
EMDR certification, n (%)			
Yes	10 (41.7)	8 (50.0)	2 (25.0)
No	14 (58.3)	8 (50.0)	6 (75.0)
Clinical experience (years), n (%)			
0	5 (20.8)	5 (31.3)	0 (0.0)
0–1	4 (16.7)	1 (6.3)	3 (37.5)
1–5	6 (25.0)	3 (18.8)	3 (37.5)
5–10	3 (12.5)	3 (18.8)	0 (0.0)
>10	6 (25.0)	4 (25.0)	2 (25.0)
Age range (years), n (%)			
20–30	11 (45.8)	5 (31.3)	6 (75.0)
31–40	6 (25.0)	6 (37.5)	0 (0.0)
41–50	3 (12.5)	1 (6.3)	2 (25.0)
> 50	4 (16.7)	4 (25.0)	0 (0.0)

2.3.2. System usability testing

The system usability test group adopted a structured experimental protocol. The experimental process was divided into three stages: preparation, execution, and closing. In the first stage, an experimental facilitator, a psychologist, and a simulated patient interacted synchronously through Google Meet. The screen was shared and recorded. The psychologist and the patient received training in think-aloud skills and were familiar with the experimental task requirements.

During the second phase, the psychologist completed eight system operation tasks, including system registration, account login, platform browsing, case account creation, consultation appointment, consultation room entry, consultation execution with the paired patient, and report download. The patient performed three tasks: system login, platform browsing, and entry to the consultation room at the same time as the paired psychologist. There were three different consultation scripts: general counseling, EMDR desensitization of an emotional memory of shame in the eighth session with self-soothing skills, and EMDR desensitization of an emotional memory of fear in the fifth session with self-soothing skills. The researchers used structured observation methods to record the following indicators: task completion time, completion status, unfinished projects, usage confusion types, operation error descriptions, and solutions.

In the final stage, all participants completed the designated questionnaires and a summative interview to obtain comprehensive data on the user experience. This experimental design was used to effectively evaluate usability and identify potential opportunities for improvement in the user interface.

2.3.3. *Clinical application testing*

A simplified experimental protocol was adopted for an actual consultation test group to evaluate the application of the system in clinical practice situations. The experimental design had three main stages: pre-preparation, execution, and interviews. In the first phase, the researcher provided the psychotherapist and the patient with preset account credentials and guided the participants into the remote psychotherapy consultation platform of this study. The focus of the execution phase was to conduct a video consultation and assess the AI-assisted report. During the case closing stage, the AI-assisted reports generated by the system, including the accuracy, clinical relevance, and timeliness of the report content, were evaluated.

The appraisals were collected systematically through structured questionnaires and in-depth interviews that focused on the practicality, workload impact, and professional development potential of the AI report generation function. This evaluation design was used to verify the feasibility of the system in practical situations and to provide a specific empirical basis for developing remote psychotherapy consultation services. This step has significant reference value for future applications to clinical practice.

2.3.4. *Evaluation metrics*

A quantitative and qualitative approach was used for the systematic assessment. The quantitative evaluation used a five-point Likert scale (1–5 points) to evaluate system usability and calculated the mean and standard deviation to understand the overall user evaluation. The qualitative assessment focused on clinical application testing in relation to three aspects: (1) measuring the difference in the time required for psychologists to write reports before and after using the system to evaluate the actual improvement in work efficiency; (2) conducting in-depth interviews to understand the system's impact on reducing; and (3) exploring the system's impact on improving the professional capabilities of young psychologists. This multifaceted evaluation framework reflects the system's performance and value in practical applications.

3. Results

3.1. *System interface implementation*

The system interface adopts an intuitive design approach and features distinct interfaces for psychotherapists and clients. As shown in Fig. 3(A) and Fig. 3(B), the psychotherapist's main interface consists of a functional menu on the upper page, including a case overview, new case registration, schedule management, and daily appointments. The lower panel displays the corresponding content on the basis of menu selection. The client interface (Fig. 3(C)) shows only the consultation schedule list and password management. The interface uses a warm color scheme to enhance the user experience and reduce visual fatigue.

Fig. 3(D) illustrates the remote psychotherapy interface, which implements a split-screen design. The primary video area occupies the central screen position and displays real-time video feeds of the therapist and client. The bottom control panel for the psychotherapist contains essential functions, including audio controls, video settings, and session control buttons for EMDR therapy. In contrast, the client interface includes only a video call controller and an exit function.

The EMDR treatment interface ensures accuracy and immediate

Fig. 3. System interface components and functionality design for the AI-integrated teletherapy platform. (A) The user authentication interface presents a secure login portal where practitioners and clients can access the system using their credentials. (B) The psychotherapist portal interface includes comprehensive functionality for case management, such as new client registration, appointment scheduling, and detailed case information management. (C) A streamlined portal interface is implemented for client access with essential features for viewing consultation schedules and video sessions. (D) The video consultation interface is designed with distinct layouts tailored for both psychotherapists and clients. The psychotherapist's view incorporates session management controls while maintaining simplified functionality for clients. (E) The remote EMDR therapy module features separate interfaces for practitioners and clients. The psychotherapist's view includes comprehensive controls for managing sight and auditory bilateral stimulation parameters, whereas the client's view focuses on presenting clear visual guidance. (F) The automated report interface synthesizes session data through emotional analysis visualizations, consultation summaries, and complete session transcripts to provide comprehensive documentation of therapeutic interactions.

feedback (Fig. 3(E)). The upper left corner of the interface for clients and psychotherapists displays warning messages, whereas the central area shows the EMDR guidance stimulus trajectory. The system displays immediate alerts when anomalies are detected. Only the psychotherapist can decide when to use EMDR therapy and the controller for EMDR guidance, such as the size, speed, color, sound, and track style. The interface features a dark theme to minimize visual distractions and enhance treatment focus.

The automated report interface, shown in Fig. 3(F), employs a modular design to clearly present the data and the results of the analysis. The report page appears only in the psychotherapist's interface and is organized into distinct sections, including basic information, emotional analysis charts, conversation summaries, and verbatim transcripts. Each module features clear headings for efficient information review and interpretation. The report is available for download in docx format to facilitate documentation and editing capabilities. The supplement video demonstrates the main interface of the psychotherapist and the client. The video and the EMDR therapy with an AI assistant are also shown in the video, as well as how to download an AI-generated report.

3.2. AI system performance analysis

3.2.1. Emotion recognition accuracy

This study established an emotion recognition model based on the Swin Transformer architecture, which demonstrated robust recognition performance. During the training process, as illustrated in Fig. 4(A), the loss function curves exhibited a consistent downward trend for both the training (blue line) and validation (red line) sets and converged in the later stages of training. The accuracy curves showed gradual improvement throughout the training process and reached a stable plateau.

The confusion matrix results, which are presented in Fig. 4(B), revealed varying recognition accuracies across different emotional categories. The normalized confusion matrix shows that the model achieved excellent performance in distinguishing between most emotion classes, with particularly strong recognition capabilities for "neutral" and "happy" emotional states, as evidenced by the high diagonal values and minimal off-diagonal confusion.

The model's classification performance across all seven emotion categories is further validated in the ROC curve analysis in Fig. 4(C). The area under the curve (AUC) values demonstrate strong discriminative ability: Happy achieved the highest AUC of 0.989, followed by Neutral (0.973), Angry (0.953), Surprise (0.949), and Sad (0.943). Disgust and Fear had relatively lower but acceptable AUC values of 0.927 and 0.863,

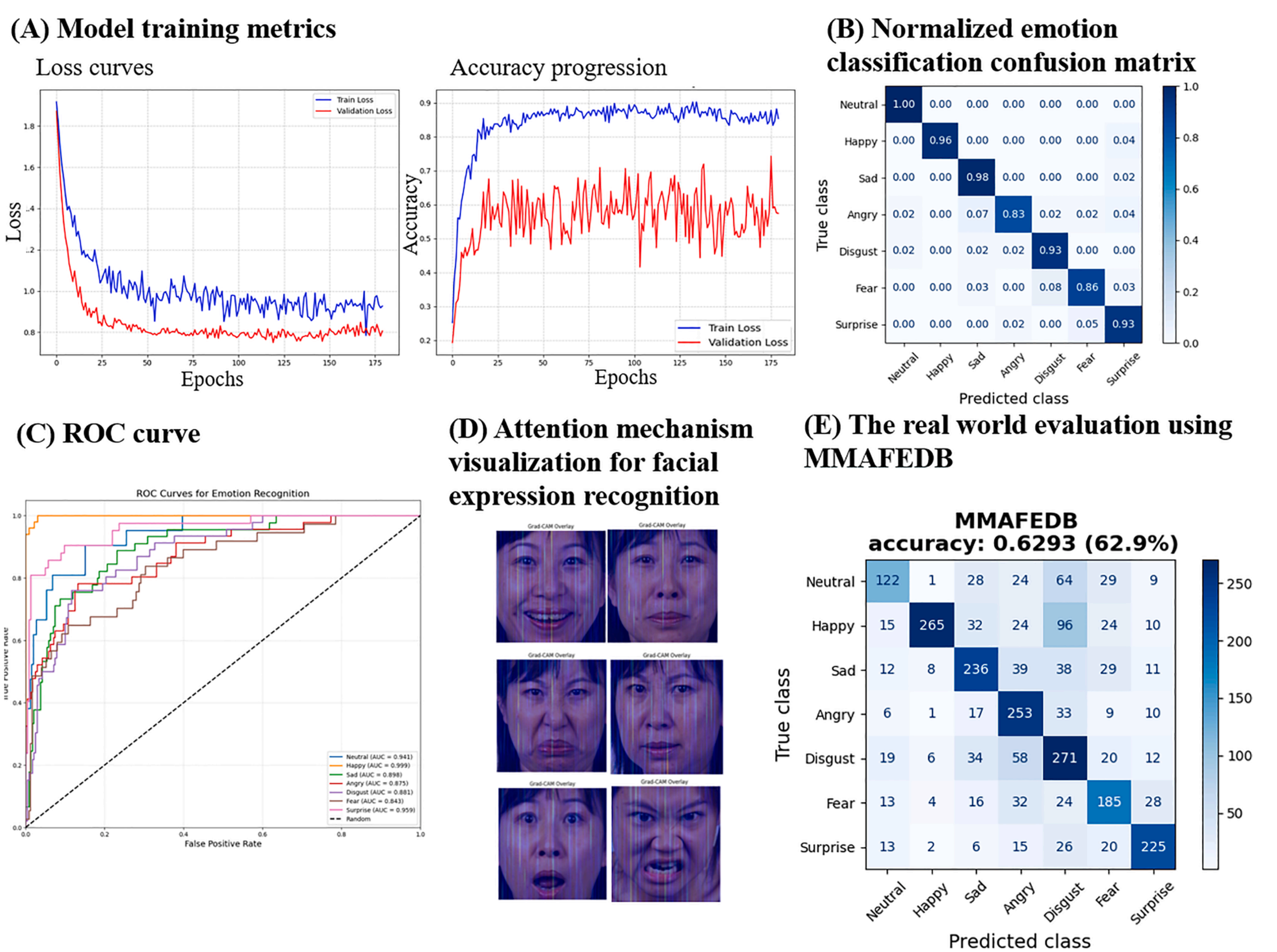


Fig. 4. Performance evaluation and visualization of the emotion recognition model based on the Swin Transformer architecture. (A) Loss curves and accuracy progression over training epochs. (B) The confusion matrix provides detailed insight into the model's classification capabilities across seven basic emotions. (C) ROC curve for multiclass emotion recognition performance evaluation. The diagonal dashed line represents random classifier performance, whereas curves closer to the top-left corner indicate better classification performance. (D) A series of attention mechanism visualizations using Grad-CAM overlays reveal how different facial expressions are processed and interpreted by the model during the recognition process. (E) The real-world evaluation using MMAFEDB shows the confusion matrix results on the real-world MMAFEDB dataset, with an overall accuracy of 0.6293 (62.9 %).

respectively. All curves positioned well above the diagonal reference line indicate superior performance compared with random classification.

In terms of overall performance metrics, the model had a test accuracy of 0.78 and an F1 score of 0.831, indicating an effective balance between precision and recall. The test loss value of 0.89 further confirmed the model's stability. Fig. 4(D) shows the model's visualization results, which present recognition examples across six emotional expressions. These results intuitively demonstrate the model's performance in practical application scenarios and substantiate the feasibility and practical utility of the proposed Swin transformer-based emotion recognition model for real-world applications.

To evaluate real-world applicability, the model was further tested on the MMAFEDB dataset, a multicultural facial expression dataset, as shown in Fig. 4(E). The real-world evaluation revealed a decreased overall accuracy of 62.9 %, reflecting the challenges of cross-dataset generalization. The confusion matrix analysis revealed distinct prediction patterns: Neutral (72 % of actual cases) and Happy (62 % of actual cases) emotions were underpredicted by the model, whereas Angry (135 % of actual cases) and Disgust (131 % of actual cases) emotions were overpredicted by the model. However, the predictions of the model were balanced for Sad (99 %), Fear (105 %), and Surprise (99 %), demonstrating the selective robustness of the model across different emotional categories.

3.2.2. Speech recognition performance

Fig. 5(A) compares the RAR and CER of the three cloud services. Google Cloud's Speech-to-Text API demonstrated an RAR of 67.3 % (CER of 32.7 %), Microsoft Azure's Azure Speech to Text achieved an RAR of 84.6 % (CER of 15.4 %), and Amazon Transcribe Service in AWS exhibited superior performance at 92.6 % (CER of 7.4 %). AWS was subsequently implemented because of its transcription accuracy, cost-effectiveness, and seamless integration with existing AWS infrastructure.

In operational deployment, the audio-to-text model of AWS incorporates noise reduction and echo cancellation algorithms to maintain transcription quality during remote sessions. The service demonstrated proficiency in general transcription and domain-specific terminology recognition and established a reliable foundation for subsequent AI analysis.

3.2.3. Effectiveness of summary generation

The initial evaluation revealed that Claude-3.5 Sonnet outperformed ChatGPT-3.5 by 18.2/20 and 16.8/20, respectively. After optimized prompts were implemented, performance improved to 19.5/20 for the prompt from Claude-3.5 Sonnet and the prompt from 18.6/20 for ChatGPT-3.5 (Fig. 5(B)). The final system implemented the Claude-3.5 Sonnet model, and the prompt was also generated from Claude-3.5 Sonnet. The prompt framework included intervention analysis of the

(A) Speech recognition accuracy comparison across different platform

Cloud Service	RAR	CER
Google Cloud	67.3%	32.7%
Microsoft Azure	84.6%	15.4%
AWS	92.6%	7.4%

(B) Summary generation quality assessment

	Initial assessment		Optimization Prompts	
	ChatGPT-3.5	Claude-3.5 Sonnet	ChatGPT-3.5	Claude-3.5 Sonnet
Completeness	18.0	19.2	19.0	19.8
Professional terminology accuracy	15.5	17.5	18.4	19.2
Structural clarity	16.2	18.4	18.7	19.6
Practical utility	17.5	17.7	18.3	19.4
Average	16.8	18.2	18.6	19.5

(C) Eye position detection results using AI model

(D) Correlation analysis between actual and measured viewing distances

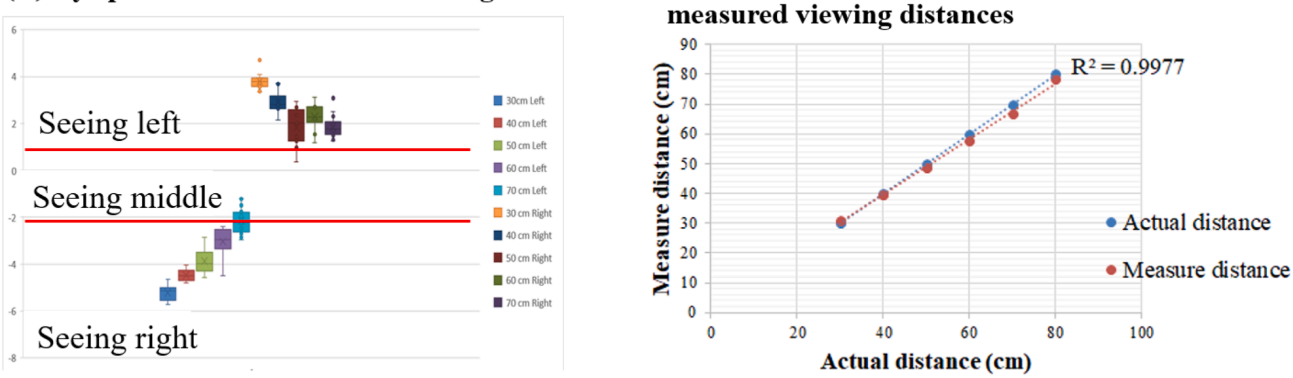


Fig. 5. Performance analysis of speech recognition, summary generation, and eye-tracking components. (A) Performance analysis of speech recognition in three different cloud services. AWS achieves a superior recognition accuracy rate (RAR) of 92.6 % and the lowest character error rate (CER) of 7.4 %, significantly outperforming Google Cloud and Microsoft Azure in Chinese language processing. (B) Quality assessment results of summary generation capabilities comparing ChatGPT-3.5 and Claude-3.5 Sonnet. (C) The eye position detection results obtained via the AI model show the system's ability to track gaze direction across different viewing distances. (D) There is a strong correlation between the actual and measured viewing distances with a high coefficient of determination ($R^2 = 0.9977$).

psychologist, client state assessment, and therapeutic interaction patterns. The system adopted the prompts and the LLM model from Claude-3.5 Sonnet. Testing confirmed the system's ability to identify and reject nontherapeutic conversations to ensure domain-specific application accuracy.

3.2.4. Precise eye movement detection

The results of the first phase of the experiment are shown in Fig. 5(C). The second phase of the experiment revealed that at an ideal distance of 35 cm, the system achieved 100 % detection accuracy. However, when the distance between the user and the webcam was shortened to 25 cm or extended to 75 cm, the accuracy decreased to 46.2 % and 48.6 %, respectively. The detection effect may be affected by distances that are too long for the size of the subject's eye.

3.2.5. Distance monitoring accuracy

As shown in Fig. 5(D), the predicted distance of the system is highly linearly correlated with the actual distance ($R^2 = 0.9977$), and the error between the expected value and the exact value is within ±2 cm. The average mistake at 30 cm is 1.2 cm. As the distance increases to 80 cm, the error increases slightly to 1.8 cm, which remains within the acceptable range. This accuracy is sufficient to support the distance monitoring requirements of the EMDR therapy module and to ensure that the client maintains an appropriate distance from the screen during treatment. The system demonstrates stable and reliable distance detection performance within the recommended treatment distance range of 35–75 cm.

3.3. User experience analysis

3.3.1. System usability assessment

The usability testing was conducted with psychologists ($n = 8$) and clients ($n = 8$) by analyzing task completion time, success rates, and error types, which are shown in Table 2. For psychologists, all basic functionality tests achieved 100 % completion rates except for the consultation execution phase (75 %). Most tasks averaged a completion time of <2 min. Due to different rates of typing, the most extended task was system registration ($M = 3.77$ min, SD=0.92). However, only 12.5 % of users completed account login without errors, primarily because the system instructions were displayed too briefly, which caused 87.5 % of users to miss crucial operational guidance. Other errors mainly occurred as a result of typing mistakes during personal information input. During the consultation, 75 % of users completed tasks without support. Issues stemmed from network restrictions, insufficient bandwidth, computer malfunctions, and unexplained webcam errors. All psychologists successfully downloaded the post-session consultation reports.

Client-side testing demonstrated 100 % success rates for account login and session room entry. Login errors resulted primarily from password input mistakes and forgotten passwords, leading to an average execution time of 1.21 min (SD=1.19). Consultation room entry averaged only 0.04 min (SD=0.01), indicating an intuitive interface design. The phase-matched psychologists had a 75 % success rate in consultation execution due to shared platform connection requirements. System performance issues focused primarily on the stability of the network connection, suggesting a need for improved adaptation to the network environment and enhanced user guidance.

3.3.2. User satisfaction analysis

Statistical analysis of the user satisfaction metrics (Table 3) revealed distinct patterns across multiple system components. The graphical user interface demonstrated high levels of satisfaction ($M = 4.25$, SD=0.66) and favorable overall user experience ratings ($M = 4.00$, SD=0.50). The EMDR module exhibited comparable performance metrics ($M = 4.00$, SD=0.87). The technical parameters, especially system stability ($M = 3.50$, SD=1.00) and audio-visual transmission quality ($M = 3.5$, SD=1.22), presented lower mean scores that were attributed primarily to infrastructure limitations.

AI-driven functionalities were stratified on the basis of clinicians' experience. Summary content usability ($M = 3.75$, SD=0.83) and emotion analysis tools ($M = 3.63$, SD=0.48) demonstrated an inverse correlation with years of clinical practice. Compared with less experienced psychotherapists ($M = 4.25$, SD=0.83 and $M = 4.00$, SD=0.00, respectively), experienced practitioners (>10 years) consistently assigned lower ratings to both summary content ($M = 3.25$, SD=0.43) and emotion analysis capabilities ($M = 3.25$, SD=0.43).

The qualitative investigation revealed that veteran clinicians prioritized comprehensive case documentation that included therapeutic processes and intervention strategies. While automated summaries may not fully capture these nuanced elements, they adequately fulfill institutional requirements for documentation. Similarly, experienced practitioners preferred to use clinical judgment in emotional assessment, particularly when considering culturally specific emotional expression patterns in Asian populations. Nevertheless, the emotional assessment tool demonstrated utility in longitudinal monitoring and case analysis. The EMDR module evaluation yielded predominantly positive outcomes, with a single outlier rating (2 points) related to treatment parameter customization. Integrating AI-powered eye-movement tracking and distance monitoring capabilities appeared to enhance therapeutic efficacy.

Client-side evaluation metrics exhibited consistently positive trends, particularly with regard to interface satisfaction ($M = 4.63$, SD=0.48). Despite moderate ratings for system stability ($M = 3.88$, SD=0.93) and audio-visual quality ($M = 3.63$, SD=1.22), overall satisfaction remained

Table 2
System usability testing results by user type.

	Time to complete (min)		Task difficulty rating	Successful, n (%)	Perform with error, n (%)	Need assistance, n (%)	Error could not be solved, n (%)
	Mean (SD)	Range	Mean (SD)				
Psychologist tasks ($n = 8$)							
System registration	3.77 (0.92)	1.85–4.90	4.76 (0.40)	8 (100.0)	2 (25.0)	2 (25.0)	0 (0.0)
Account login	0.92 (0.33)	0.35–1.37	4.88 (0.33)	8 (100.0)	7 (87.5)	7 (87.5)	0 (0.0)
Platform navigation	1.32 (0.45)	0.68–2.03	5.00 (0.00)	8 (100.0)	1 (12.5)	0 (0.0)	0 (0.0)
Case account creation	2.18 (0.48)	1.33–2.83	5.00 (0.00)	8 (100.0)	1 (12.5)	1 (12.5)	0 (0.0)
Consultation scheduling	1.38 (0.37)	1.03–2.05	4.88 (0.33)	8 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)
Session room entry	0.48 (0.37)	0.22–1.12	4.88 (0.33)	8 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)
Consultation execution	–	–	4.50 (0.50)	6 (75.0)	2 (25.0)	2 (25)	2 (25.0)
Report download	–	–	4.75 (0.43)	8 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)
Client tasks ($n = 8$)							
Account login	1.21 (1.19)	0.38–3.55	4.25 (1.39)	8 (100.0)	2 (25.0)	1 (12.5)	0 (0.0)
Session room entry	0.04 (0.01)	0.03–0.05	4.25 (1.40)	8 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)
Consultation execution			4.5 (0.50)	6 (75.0)	2 (25.0)	2 (25.0)	2 (25.0)

Table 3
User satisfaction ratings.

	Mean (SD)	Distribution				
		5 (%)	4 (%)	3 (%)	2 (%)	1 (%)
Psychologist satisfactions (*n* = 8)						
User interface satisfaction	4.25 (0.66)	3 (37.5)	4 (50.0)	1 (12.5)	0 (0.0)	0 (0.0)
System stability	3.5 (1.00)	1 (12.5)	4 (50.0)	1 (12.5)	2 (25.0)	0 (0.0)
Audiovisual transmission quality	3.5 (1.22)	2 (25.0)	3 (37.5)	0 (0.0)	3 (37.5)	0 (0.0)
Overall system user experience	4 (0.50)	1 (12.5)	6 (75.0)	1 (12.5)	0 (0.0)	0 (0.0)
Summary content usability						
Clinical experience in total	3.75 (0.83)					
Clinical experience >10 (*n* = 4)	3.25 (0.43)	0 (0.0)	1 (25.0)	3 (75.0)	0 (0.0)	0 (0.0)
Clinical experience <10 (*n* = 4)	4.25 (0.83)	2 (50.0)	1 (25.0)	1 (25.0)	0 (0.0)	0 (0.0)
Emotional assessment tool availability						
Clinical Experience in total	3.63 (0.48)					
Clinical Experience >10 (*n* = 4)	3.25 (0.43)	1 (25.0)	3 (75.0)	0 (0.0)	0 (0.0)	0 (0.0)
Clinical Experience <10 (*n* = 4)	4 (0.00)	4 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
EMDR therapy tool availability						
Clinical experience in total	4 (0.87)					
Clinical experience >10 (*n* = 4)	3.5 (0.87)	0 (0.0)	3 (75.0)	1 (25.0)	0 (0.0)	0 (0.0)
Clinical experience <10 (*n* = 4)	4.5 (0.50)	2 (50)	2 (50)	0 (0.0)	0 (0.0)	0 (0.0)
Client satisfactions (*n* = 8)						
User interface satisfaction	4.63 (0.48)	5 (62.5)	3 (37.5)	0 (0.0)	0 (0.0)	0 (0.0)
System stability	3.88 (0.93)	2 (25.0)	4 (50.0)	1 (12.5)	1 (12.5)	0 (0.0)
Audiovisual transmission quality	3.63 (1.22)	2 (25.0)	3 (37.5)	2 (25.0)	0 (0.0)	1 (12.5)
Overall system user experience	4.38 (0.70)	4 (50.0)	3 (37.5)	1 (12.5)	0 (0.0)	0 (0.0)

above threshold levels. The observed difference between client and therapist ratings may be attributed to varying operational requirements and user sophistication levels.

The operational complexity assessment (scale: 1=extremely complicated, 5=extremely simple, in Table 2) indicated high usability among psychotherapists, particularly in navigation and account management functions (M = 5.00, SD=0.00). While client ratings were slightly lower because of the use of activation procedures and network variables, they maintained satisfactory levels.

The semantic analysis of the qualitative data in Fig. 6(A) highlights recurring descriptors, including "intuitive," "user-friendly," and "innovative." Psychotherapists emphasized "high-fidelity" characteristics, reflecting professional requirements. Clients (Fig. 6(B)), noted "complex" aspects that suggested opportunities for interface optimization.

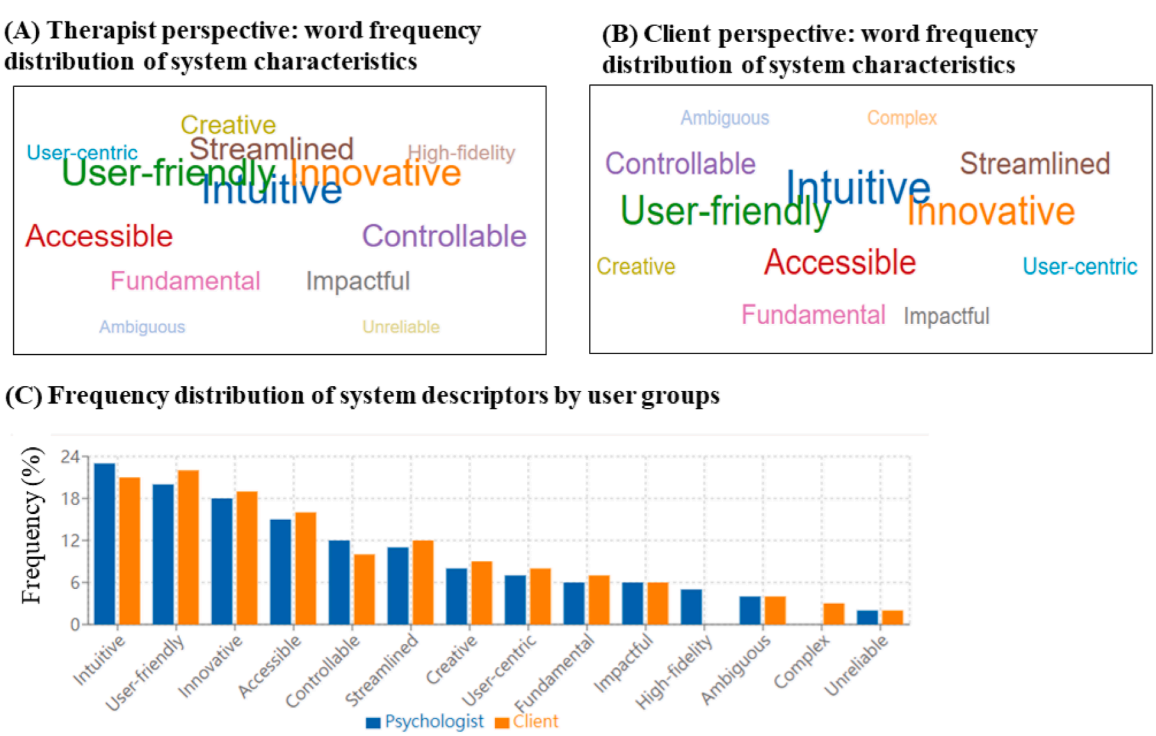


Fig. 6. Word cloud visualization of system characteristics based on user feedback analysis. (A) The therapist's perspective shows similar prominence of "intuitive" and "user-friendly" but additionally highlights professional considerations through terms such as "high fidelity" and "streamlined." The size of each word corresponds to its frequency in user feedback, with larger text indicating more frequent mentions. (B) The client perspective, where terms such as "intuitive" and "user friendly" emerge as dominant descriptors, reflects the system's accessibility to end-users. The prominence of "accessible" and "innovative" further emphasizes the system's perceived ease of use and modern approach. (C) Frequency distribution of system descriptors by user groups. Comparative analysis between psychologists (n = 8) and clients (n = 8) reveals that the primary descriptors are "intuitive", "user friendly", and "innovative" (>18 % each). Professional-specific concerns focused on system fidelity, while clients noted complexity issues. Both groups reported minimal negative descriptors (<4 %).

The frequency analysis of system descriptive terms (Fig. 6(C)) revealed three predominant characteristics across both user groups: "intuitive" (psychologists: 23 %, clients: 21 %), "user-friendly" (psychologists: 20 %, clients: 22 %), and "innovative" (psychologists: 18 %, clients: 19 %). Notably, "high-fidelity" appeared exclusively in psychologists' descriptions (5 %), whereas "complex" was unique to client feedback (3 %), indicating distinct perceptions between the user groups. Negative descriptors such as "ambiguous" (4 % for both groups) and "unreliable" (2 % for both groups) consistently presented low frequencies across both populations and were primarily related to the technical constraints of the system.

3.3.3. Perceived workflow optimization

The study demonstrated significant benefits in workflow optimization through system implementation (Table 4). The evaluation involved two groups: a group that involved consultation with scripts ($n = 11$), which represented the sandbox text after brief exposure, and a group that involved consultation in the real world, which reflected actual usage assessment.

With respect to the efficiency of appointment management, the script counseling group provided notably high ratings ($M = 4.82$, SD=0.39), with >80 % of users rating it above 4 points. Although the real-world counseling group provided slightly lower scores ($M = 4.25$, SD=0.83), it maintained overall positive evaluations. The automated appointment reminder feature received favorable ratings from both groups (script group: $M = 4.55$, SD=0.66; real-world group: $M = 4.25$, SD=0.66), indicating an effective reduction in administrative burden.

Qualitative interviews revealed various organizational structures that affected the utility of the system. Some institutions employ dedicated staff for appointment management and client reminders, but many counselors work across multiple facilities. Consequently, counselors often maintain separate calendars to coordinate appointments. Notably, some practitioners viewed missed appointments as therapeutic discussion points and questioned the need for reminder systems.

The session summary of the system received strong approval. The script counseling group unanimously scored it above 4 points ($M = 4.45$, SD=0.5). In contrast, 87.5 % of the real-world group provided positive evaluations ($M = 4.25$, SD=0.66), confirming that it saved >50 % of report-writing hours. However, the emotional analysis component received moderate ratings, averaging approximately 3.5 in both groups (script group: $M = 3.55$, SD=0.5; real-world group: $M = 3.38$, SD=0.48). This lower rating may be attributed to the current single-mode emotional model and suggests room for improvement.

The interview findings highlighted limitations in automated report generation due to varying institutional requirements. Some facilities demand detailed documentation, others prefer bullet-point summaries, and some require handwritten reports. Consequently, the system's capability to generate reports is a supplementary tool rather than a complete solution. Practitioners noted its utility for generating basic reports but acknowledged limitations in automatically producing therapeutic processes and strategies. Senior practitioners often develop customized templates for EMDR therapy documentation, which primarily involves recording clients' sensation scores and negative emotional imagery. This circumstance underscores the current constraints in the automation of counseling reports.

The analysis of practitioner feedback revealed significant insights related to emotional analysis. Senior psychologists emphasized that emotional assessment remains a fundamental professional competency. Consequently, practitioners consistently prioritize their professional judgment when discrepancies arise between clinicians' observations and AI-generated emotional model reports.

Furthermore, the system's current presentation of emotional analysis through percentage-based metrics presents limitations in clinical utility. Practitioners noted that the absence of temporal markers to indicate specific therapeutic moments of emotional fluctuation impeded their ability to develop targeted intervention strategies. However, some clinicians identified a valuable application in comparing variations in emotional percentages across multiple sessions for individual clients, suggesting potential utility in tracking therapeutic effectiveness longitudinally.

These findings suggest that while automated emotional analysis tools may serve as supplementary resources for assessment, they currently function better as complementary rather than primary diagnostic instruments in therapeutic settings.

In terms of overall effectiveness, the system demonstrated exceptional performance in reducing administrative paperwork, with both groups maintaining scores above 4 points (script group: $M = 4.45$, SD=0.89; real-world group: $M = 4.38$, SD=0.48). The system also

Table 4
Comparative analysis of system implementation effectiveness between script-based and real-world consultation groups.

	Consultation with scripts ($n = 11$)					Consultation in real world ($n = 8$)			
	Mean (SD)	Distribution, n(%)				Mean (SD)	Distribution, n(%)		
		5 (%)	4 (%)	3 (%)	2 (%)		5 (%)	4 (%)	3 (%)
Case management system implementation									
The system demonstrates enhanced efficiency in managing session reservations.	4.82 (0.39)	7 (63.6)	2 (18.2)	2 (18.2)	0 (0.0)	4.25 (0.83)	3 (37.5)	2 (28.6)	3 (37.5)
The automated appointment reminder reduces administrative workload.	4.55 (0.66)	7 (63.6)	3 (27.2)	1 (0.9)	0 (0.0)	4.25 (0.66)	3 (37.5)	4 (50.0)	1 (12.5)
I think overall, the system can reduce my workload	4.45 (0.78)	7 (63.6)	2 (18.2)	2 (18.2)	0 (0.0)	4(0.87)	3 (37.5)	2 (28.6)	3 (37.5)
Therapeutic report generation									
The session summary facilitates more effective post-consultation report preparation.	4.45 (0.5)	5 (45.5)	6 (54.5)	0 (0.0)	0 (0.0)	4.25 (0.66)	3 (37.5)	4 (50.0)	1 (12.5)
The emotional analytics component provides enhanced insights into client emotional patterns.	3.55 (0.5)	0 (0.0)	6 (54.5)	5 (45.5)	0 (0.0)	3.38 (0.48)	0 (0.0)	3 (37.5)	5 (62.5)
I think overall, the system can reduce my workload	4.45 (0.87)	7 (63.6)	2 (18.2)	2 (18.2)	0 (0.0)	4.38(0.7)	4 (50.0)	3 (37.5)	1 (12.5)
In overall									
The system exhibits effectiveness in reducing administrative documentation.	4.45 (0.89)	7 (63.6)	3 (27.2)	0 (0.0)	1 (12.5)	4.38 (0.48)	3 (37.5)	5 (62.5)	0 (0.0)
The system enables more optimal allocation of work schedules.	4.36 (0.98)	7 (63.6)	2 (18.2)	1 (0.9)	1 (0.9)	4.25 (0.66)	3 (37.5)	4 (50.0)	1 (12.5)
The system demonstrates improvement in work productivity efficiency.	4.45 (0.78)	7 (63.6)	2 (18.2)	0 (0.0)	0 (0.0)	4.25 (0.43)	2 (28.6)	6 (75.0)	0 (0.0)
The system contributes to enhanced quality of work deliverables.	3.91(1)	4(36.4)	3 (27.2)	3 (27.2)	1 (0.9)	4(0.71)	2 (28.6)	4 (50.0)	2 (28.6)

received positive evaluations for work schedule optimization (script group: $M = 4.36$, SD=0.98; real-world group: $M = 4.25$, SD=0.66). Both groups reported favorable assessments of work productivity efficiency (script group: $M = 4.45$, SD=0.78; real-world group: $M = 4.25$, SD=0.43) that indicated significant improvements in operational efficiency.

Regarding work quality enhancement, both groups provided similar ratings (script group: $M = 3.91$, SD=1; real-world group: $M = 4$, SD=0.71). Practitioners emphasized that the quality of counseling work manifests primarily during therapeutic dialog and psychological treatment sessions. They noted that factors such as the therapist's mental state and continued professional development have direct impacts on work quality. Consequently, while the system successfully reduces workload, it does not directly influence the qualitative aspects of therapeutic practice.

These data collectively indicate positive system benefits in optimizing workflows, mainly administrative management, and enhancing efficiency. However, certain functionalities (such as emotional analysis) and practical application scenarios require further refinement. Disparities between the evaluations of the script and real-world counseling groups highlight potential challenges and the need for adaptation in the context of practical implementation.

3.3.4. Professional development benefits

Research findings on professional development benefits revealed significant disparities in the evaluation of the system between less experienced (<5 years, $n = 10$) and more experienced (≥5 years, $n = 9$) practitioners (Table 5).

With respect to enhanced therapeutic self-reflection, less experienced practitioners demonstrated greater appreciation ($M = 4.3$, SD=0.64), with 90 % providing ratings above 4. While more experienced practitioners presented lower ratings ($M = 3.89$, SD=0.99), 66.6 % maintained scores above 4. In terms of the development of professional confidence, the evaluations of less experienced practitioners ($M = 3.8$, SD=0.6) were substantially higher than those of experienced practitioners ($M = 3.33$, SD=1.25), with 70 % of less experienced practitioners providing ratings above 4, in contrast with more varied ratings among experienced practitioners. The most pronounced disparity emerged in the enhancement of remote consultation capabilities: less experienced practitioners provided notably positive evaluations ($M = 4.2$, SD=0.75), with 80 % providing ratings above 4 points, whereas experienced practitioners showed lower ratings ($M = 3.33$, SD=1.25), with only 44.4 % exceeding 4 points.

The qualitative interview data revealed divergent perspectives on approaches to professional development. Experienced practitioners emphasized traditional enhancement methods, including continuous education, supervision, and detailed analysis of clients' emotional responses and therapeutic interactions. Conversely, less experienced practitioners valued the system's capability for emotional modeling to facilitate self-reflection and identify previously unnoticed emotional reactions, which informed subsequent therapeutic sessions. However, some experienced practitioners expressed concern that significant discrepancies between AI-generated emotional models and practitioners' perceptions might impact therapeutic confidence.

The quantitative and qualitative data suggest that the system's support for professional development provides more substantial benefits for less experienced practitioners. In contrast, experienced practitioners derive comparatively lower developmental value from the features of the system.

4. Discussion

4.1. Technological innovation

This study presents a significant advancement in AI-assisted remote psychotherapy through the development of a validated system. This system's comprehensive integration of AI technologies, including emotion recognition, speech transcription, and automated report generation, represents a novel approach to digital mental healthcare delivery. Clinical validation demonstrates that the system reduces report writing time by approximately 50 %, allowing practitioners to allocate more time to the development of therapeutic strategies and professional advancement. This system is the first comprehensive remote psychotherapy assistance platform that does not require the download of a specific application or extra equipment.

This system's innovative approach to remote EMDR therapy addresses several critical challenges in the implementation of remote psychotherapy. Through real-time eye movement tracking and distance monitoring, the system overcomes the common limitations of remote EMDR therapy, such as ineffective bilateral stimulation and suboptimal eye movement patterns [45,46]. This technological solution benefits novice practitioners by providing standardized guidance for the delivery of remote EMDR therapy.

Based on the Swin Transformer architecture, the emotion recognition component demonstrates high accuracy in the detection of neutral and happy emotional states with a higher rate of detection that found in 2013 [47]. However, cultural considerations emerged as a significant factor that affected system performance. The study revealed that traditional Taiwanese patterns of emotional expression, which are characterized by emotional suppression and the cultural tendency to mask negative emotions, challenged the accuracy of the AI model [48]. For example, the system frequently misclassified self-deprecating smiles, which are commonly used when difficulties are discussed, as expressions of happiness [49,50]. Despite these limitations, the emotion recognition feature is a valuable tool for therapeutic reflection and enables therapists to compare their subjective assessments with objective data. However, multimodel emotion recognition could be a solution that combines tone of voice, hand posture, and text content [51].

The integration of AWS speech recognition services, which achieved a 92.6 % accuracy rate, establishes a reliable foundation for session documentation and analysis. This high performance in speech recognition combined with other AI-powered features suggests that the system successfully addresses key operational challenges in the delivery of remote psychotherapy while maintaining clinical efficacy.

These findings collectively indicate that while AI-assisted teletherapy systems can significantly enhance operational efficiency,

Table 5
Comparison of system benefits on professional development between experienced and less experienced practitioners.

	Clinical experience ≥ 5 years, $n = 9$						Clinical experience < 5 years, $n = 10$			
	Mean (SD)	Distribution, n(%)					Mean (SD)	Distribution, n(%)		
		5 (%)	4 (%)	3 (%)	2 (%)	1 (%)		5 (%)	4 (%)	3 (%)
The system can enhanced self-reflection on therapeutic consultation processes.	3.89 (0.99)	3 (33.3)	3 (33.3)	2 (22.2)	1 (11.1)	0 (0.0)	4.3(0.64)	4 (40.0)	5 (50.0)	1 (1.0)
The system can accelerate development of professional confidence.	3.33 (1.25)	2 (22.2)	2 (22.2)	3 (33.3)	1 (11.1)	1 (11.1)	3.8(0.6)	1 (1.0)	6 (60.0)	3 (30.0)
The system can enhance professional capabilities in remote consultation	3.33 (1.25)	2 (22.2)	2 (22.2)	3 (33.3)	1 (11.1)	1 (11.1)	4.2(0.75)	4 (40.0)	4 (40.0)	2 (20.0)

cultural nuances and patterns of emotional expression remain important considerations for future development. The system's ability to reduce administrative burden as well as the burden of purchasing extra material while providing objective data for therapeutic reflection represents a meaningful step forward in the evolution of mental healthcare delivery.

4.2. Clinical applications and impact

The remote psychotherapy system significantly optimized psychological counseling workflows. Statistical analysis revealed high satisfaction ratings for appointment management efficiency in both the script consultation group ($M = 4.82$) and the real-world consultation group ($M = 4.25$). Notably, the automated appointment reminder function received unanimous positive feedback from both groups, with practitioners reporting a 70 % reduction in time spent on appointment management and client reminders. This substantial decrease in administrative workload represents a significant improvement in operational efficiency.

The system's session summary functionality also received strong approval from both user groups. The script consultation group unanimously rated this feature above 4 points ($M = 4.45$), whereas the real-world consultation group maintained positive evaluations with 87.5 % approval ($M = 4.25$) despite varying institutional requirements. These results suggest that the summary feature is robustly adaptable across clinical settings.

However, the study revealed significant disparities in the impact of the system based on practitioners' levels of experience. Less experienced psychotherapists ($<$ 5 years) reported substantial benefits to therapeutic self-reflection ($M = 4.3$) and remote consultation capabilities ($M = 4.2$). In contrast, more experienced psychotherapists ($\geq$ 5 years) provided lower ratings for these features ($M = 3.89$ and $M = 3.33$, respectively). This disparity suggests that the features of the system that support professional development may be particularly beneficial for early-career practitioners but may offer less value to more experienced clinicians.

These findings indicate that while the system successfully reduces administrative burden across all user groups, its impact on professional development varies significantly based on practitioners' experience. This difference in benefits to users highlights the importance of considering experience levels when developing and implementing AI-assisted therapeutic tools.

While the current study is focused primarily on technical feasibility, system integration, and the user experience, we recognize the importance of addressing broader ethical and clinical implications. The system is explicitly designed to support, not replace, professional judgment. Even though AI-generated summaries and emotion analyses are helpful, final decisions must remain in the hands of human clinicians, and this should be emphasized.

4.3. Feature comparison with commercial teletherapy systems

To further contextualize the contributions of our system, a feature-level comparison with three widely used commercial teletherapy platforms, EMDR Remote, SimplePractice, and TheraNest, was conducted. While these platforms offer robust case management, scheduling, and billing tools, integrated AI features and comprehensive EMDR-specific support are unavailable.

Only EMDR Remote has visual bilateral stimulation (BLS), but it does not support real-time eye tracking, emotion recognition, or automated clinical documentation. In contrast, our system incorporates the Amazon Transcribe service to transcribe speech to text, facial emotion analysis, session summarization using large language models (Claude 3.5), and automated report generation—all of which are absent in commercial counterparts.

Additionally, the remote psychotherapy system in this study supports whole browser-based operation with no installation needed, which enhances accessibility for therapists and clients. Unlike existing systems, it offers a built-in eye-tracking module with distance alerts to maintain EMDR treatment fidelity during virtual sessions.

Furthermore, our system has undergone formal user experience testing and IRB-approved clinical validation, which have not been reported by the other platforms (Table 6). These distinctions position our platform not only as a technically innovative solution but also as one with demonstrated clinical feasibility and readiness for real-world deployment.

4.4. Study limitations and future directions

The current study has several limitations. The emotion analysis model operated in a single-mode framework that potentially limited its ability to capture complex emotional variations. Presenting emotional data as percentages proved less beneficial for reviewing individual session dynamics, which suggests that alternative visualization methods are needed [52]. While automated report generation successfully reduced the documentation workload, institutional variations in documentation requirements necessitate manual customization. Nevertheless, the system had a positive effect on therapists' psychological burden and facilitated case recall.

Network stability emerged as a critical factor that affected audiovisual transmission quality. Connection failures were observed in cases of low network speeds or specific institutional intranets. These findings suggest the potential benefit of implementing adaptive resolution adjustment mechanisms that can automatically modulate audio and video quality based on network conditions.

Future research should focus on the following. First, the emotional model should be expanded to reflect the dynamic nature of therapeutic emotional processes. Second, report generation should be enhanced to accommodate diverse institutional requirements. Third, system adaptability and variable network environments should be improved. Future versions of the system will incorporate adaptive streaming techniques and latency-tolerant mechanisms to ensure stability across low-bandwidth environments.

Additionally, we plan to conduct controlled network simulations using tools such as tc to simulate packet delay and loss and evaluate the impacts of bandwidth, jitter, and packet loss on emotion recognition accuracy and real-time eye-tracking performance. The development of a comprehensive user training process, particularly for new users, is also recommended to optimize system utilization and enhance remote counseling capabilities. Our current prototype is hosted on AWS Free Tier to support operational scalability. While this setup has been sufficient for small-scale research use, the estimated costs for clinical deployment are approximately \$0.18–0.25 USD per user-hour. Future implementations will adopt serverless architectures using services such as AWS Lambda and S3 to increase cost efficiency and minimize idle-time overhead.

While the system demonstrated significant improvements in administrative efficiency and workflow optimization, its direct impact on therapeutic quality remains limited. Future developments should prioritize integrating technological and professional therapeutic practices to enhance service quality. This integration should focus on maintaining the balance between technological efficiency and therapeutic effectiveness. Finally, future trials will involve larger and more diverse participant cohorts to test generalizability across clinical contexts and cultures. The aim of these efforts is to ensure that the system remains technically resilient, clinically relevant, and economically sustainable, as it scales across different healthcare environments.

The findings highlight the importance of addressing technical and practical limitations while focusing on the system's primary goal of supporting therapeutic processes. Future research should explore more sophisticated approaches to emotion recognition, customizable reporting systems, and adaptive network solutions to increase the system's utility across diverse clinical settings.

Table 6
Feature comparison with commercial teletherapy systems.

Feature category	Feature	Our study	RemotEMDR	SimplePractice	TheraNest
Platform overview	–	A platform that integrates basic case management system, remote EMDR, AI emotion recognition and automatic generation of summaries and transcripts	Video-based EMDR platform with visual bilateral stimulation	Clinic Management System, which integrates booking, billing, video, and recording functions	A team-oriented platform for psychologists and institutions
EMDR features	Remote EMDR	V	V	X	X
	Remote EMDR Support	V: eye tracking, visual BLS, real-time distance feedback	X	X	X
AI enhancements	Facial emotion recognition	V: With emotion trend analysis	X	X	X
	AI-generated summary and transcripts	V	X	X	X
	Auto-generated report	V: Downloadable DOCX report	X	Partial (manual)	Partial
	EMDR treatment fidelity during virtual therapy	V: Instant eye movement tracking, distance tracking with alert.	X	X	X
Client management	Case management system	V	Basic profile only	V	V
	Scheduling and reminders	V	V	V	V
Use threshold	Installation requirement	X	V	X	X
Security and deployment	Cloud security architecture	V	Unclear	Unclear	Unclear
Validation and feedback	User experience testing	V	X	X	X
	Clinical validation (IRB-backed)	V	No public data	No public data	No public data

5. Conclusion

This study presents the first remote psychotherapy platform to successfully develop an AI-integrated teletherapy system that demonstrates technological innovation and clinical value across multiple dimensions. The system effectively reduces therapists' administrative burden through automated case management, session documentation, and report generation capabilities. It reduces user report writing time by an average of 50 %, allowing therapists to dedicate more time to the development of therapeutic strategies and professional growth. Notably, in remote EMDR therapy, the system's real-time eye movement tracking and distance monitoring features successfully overcome the technical limitations of remote treatment and provide a reliable solution for the delivery of standardized remote EMDR therapy. The accuracy of emotion recognition is 78.3 %. The error of distance monitoring is <2 cm, and the eye movement detection accuracy is 100 % when the client is at the proper distance. The research findings indicate different impacts across therapists' experience levels, with particular benefits for professional development for less experienced practitioners. The user experience exam indicates that the platform is easy to access. However, several limitations and directions for future development were identified, including improving cultural sensitivity in emotion analysis models, optimizing network stability, and enhancing the system's adaptability to diverse institutional requirements for documentation. Nevertheless, the system substantially improves mental health service accessibility and operational efficiency and provides valuable insights for the future development of digital mental healthcare. This research demonstrates the potential of AI technology applications in mental health and establishes a crucial foundation for the digital transformation of teletherapy services.

Ethics approval

Ethics approval is not required.

CRediT authorship contribution statement

Ju Yu Wu: Methodology, Data curation, Software, Conceptualization, Project administration. **Yu-Jie Chen:** Supervision. **Yen-Feng Lin:** Supervision. **Congo Tak Shing Ching:** Supervision. **Hui-Min David Wang:** Supervision. **Yi-Chung Chen:** Supervision. **Yang-Chou Juan:** Software. **Lun-De Liao:** Writing – original draft, Visualization, Validation, Writing – review & editing, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported in part by the National Science and Technology Council of Taiwan under grant numbers NSTC 114-2221-E-400-001-MY3, NSTC 113–2221-E-400–003, NSTC 113-2218-E-007-026, 112-2221-E-005-042 and 111-2221-E-005-018.; by the National Health Research Institutes of Taiwan under grant numbers BN-114-PP-15, BN-114-GP-02, BN-114-GP-03 and BN-114-GP-05; by the Ministry of Economic Affairs of Taiwan under the grant numbers 114-EC-17-A-22-1906 and 114-EC-17-A22-1905; by the Metal Industries Research & Development Centre of Taiwan under the grant number K00069651K and the Southern Taiwan Science Park Bureau, Ministry of Science and Technology under the grant number 113CH-1-03. Ju-Yu Wu carried out this research with funding support in part by the Doctoral Program in Tissue Engineering and Regenerative Medicine of National Chung Hsing University and the National Health Research Institutes.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rineng.2025.105931.

Data availability

Data will be made available on request.

References

[1] S.M.A. Shah, D. Mohammad, M.F.H. Qureshi, M.Z. Abbas, S. Aleem, Prevalence, psychological responses and associated correlates of depression, anxiety and stress in a global population, during the coronavirus disease (COVID-19) pandemic, Community Ment. Health J. 57 (1) (2021) 101–110.
[2] C. Kieling, et al., Worldwide prevalence and disability from mental disorders across childhood and adolescence: evidence from the global burden of disease study, JAMA Psychiatry 81 (4) (2024) 347–356.
[3] M. Luo, L. Guo, M. Yu, W. Jiang, H. Wang, The psychological and mental impact of coronavirus disease 2019 (COVID-19) on medical staff and general public–a systematic review and meta-analysis, Psychiatry Res. 291 (2020) 113190.
[4] A. Nalbandian, et al., Post-acute COVID-19 syndrome, Nat. Med. 27 (4) (2021) 601–615.
[5] N. Seighali, et al., The global prevalence of depression, anxiety, and sleep disorder among patients coping with post COVID-19 syndrome (long COVID): a systematic review and meta-analysis, BMC Psychiatry 24 (1) (2024) 105.
[6] J.-Y. Wu, et al., Digital transformation of mental health therapy by integrating digitalized cognitive behavioral therapy and eye movement desensitization and reprocessing, Med. Biol. Eng. Comput. (2024) 1–16.
[7] Y. Zhang, H. Ma, Y. Wang, X. Du, D. Chi, Burnout and its associated factors in psychotherapists, Curr. Psychol. (2024) 1–11.
[8] Y. Kotera, R. Maxwell-Jones, A.-M. Edwards, N. Knutton, Burnout in professional psychotherapists: relationships with self-compassion, work–life balance, and telepressure, Int. J. Env. Res. Public Health 18 (10) (2021) 5308.
[9] L. Vilela e Souza, F. Scorsolini-Comin, Counselling psychology practice in Latin America, Couns. Psychother. Res. 22 (4) (2022) 839–843.
[10] A. Bloch-Atefi, E. Day, T. Snell, G. O`Neill, A snapshot of the counselling and psychotherapy workforce in Australia in 2020: underutilised and poorly remunerated, yet highly qualified and desperately needed, Psychother. Couns. J. Aust. 9 (2) (2021).
[11] A.K. Giovanetti, S.E. Punt, E.-L. Nelson, S.S. Ilardi, Teletherapy versus in-person psychotherapy for depression: a meta-analysis of randomized controlled trials, Telemed. e-Health 28 (8) (2022) 1077–1089.
[12] K.E. Venturo-Conerly, O.M. Fitzpatrick, R.L. Horn, A.M. Ugueto, J.R. Weisz, Effectiveness of youth psychotherapy delivered remotely: a meta-analysis, Am. Psychol. 77 (1) (2022) 71.
[13] J.C. Markowitz, et al., Psychotherapy at a distance, Am. J. Psychiatry 178 (3) (2021) 240–246.
[14] S. Zandieh, et al., Therapist-guided remote versus in-person cognitive behavioural therapy: a systematic review and meta-analysis of randomized controlled trials, CMAJ 196 (10) (2024) E327–E340.
[15] V.Trenoska Basile, T. Newton-John, B.M. Wootton, Remote cognitive-behavioral therapy for generalized anxiety disorder: a preliminary meta-analysis, J. Clin. Psychol. 78 (12) (2022) 2381–2395.
[16] G. Efron, B.M. Wootton, Remote cognitive behavioral therapy for panic disorder: a meta-analysis, J. Anxiety Disord. 79 (2021) 102385.
[17] S. Russo, F. Fiani, C. Napoli, Remote eye movement desensitization and reprocessing treatment of long-COVID-and Post-COVID-related traumatic disorders: an innovative approach, Brain Sci. 14 (12) (2024) 1212.
[18] I.W. McGowan, N. Fisher, J. Havens, S. Proudlock, An evaluation of eye movement desensitization and reprocessing therapy delivered remotely during the Covid–19 pandemic, BMC Psychiatry 21 (2021) 1–8.
[19] N. Goga, et al., An efficient system for eye movement desensitization and reprocessing (EMDR) therapy: a pilot study, in: Healthcare, 10, MDPI, 2022, p. 133.
[20] A.A. Drewes, J.W. Seymour, Integrative psychotherapy with children. Handbook of Psychotherapy Integration, 2019, pp. 341–356.
[21] F. Shapiro, Adaptive information processing and case conceptualization, in: Presentation at the Annual Conference of the EMDR-Europe Association, Brussels, Belgium, 2005.
[22] F. Shapiro, EMDR 12 years after its introduction: past and future research, J. Clin. Psychol. 58 (1) (2002) 1–22.
[23] A. Backhaus, et al., Videoconferencing psychotherapy: a systematic review, Psychol. Serv. 9 (2) (2012) 111.
[24] S.A. Marotta-Walters, K. Jain, J. DiNardo, P. Kaur, S. Kaligounder, A review of mobile applications for facilitating EMDR treatment of complex trauma and its comorbidities, J. EMDR Pract. Res. 12 (1) (2018) 2–15.
[25] S.K. Kaptan, Z.M. Kaya, A. Akan, Addressing mental health need after COVID-19: a systematic review of remote EMDR therapy studies as an emerging option, Front. Psychiatry 14 (2024) 1336569.
[26] N. Fisher, Using EMDR therapy to treat clients remotely, J. EMDR Pract. Res. 15 (1) (2021).
[27] L.S. Bozzoli, Eye Movement Desensitization Reprocessing Treatment Protocol for Telehealth, Walden University, 2025.
[28] A. Thompson-de Benoit, U. Kramer, Work with emotions in remote psychotherapy in the time of Covid-19: a clinical experience. How the COVID-19 Pandemic Transformed the Mental Health Landscape, Routledge, 2023, pp. 190–198.
[29] V.L. Forman-Hoffman, J.C. Kvedar, K. Ranta, Potential benefits of remote continuous care for depression, Int. J. Digit. Health 1 (1) (2021) 15.
[30] M. Eisenstadt, S. Liverpool, E. Infanti, R.M. Ciuvat, C. Carlsson, Mobile apps that promote emotion regulation, positive mental health, and well-being in the general population: systematic review and meta-analysis, JMIR Ment. Health 8 (11) (2021) e31170.
[31] M. Blair, L. Tweedlie, H. Minnis, I. Cronin, F. Turner, Online therapy with families-what can families tell us about how to do this well? a qualitative study assessing families' experience of remote Dyadic Developmental Psychotherapy compared to face-to-face therapy, Plos One 19 (4) (2024) e0301640.
[32] R.Mc Kenny, E. Galloghly, C.M. Porter, F.R. Burbach, Living in a zoom world': survey mapping how COVID-19 is changing family therapy practice in the UK, J. Fam. Ther. 43 (2) (2021) 272–294.
[33] K. Ahlström, C. Von Below, D. Forsström, A. Werbart, Therapeutic encounters at the onset of the COVID-19 pandemic: psychodynamic therapists' experiences of transition to remote psychotherapy, Psychoanal. Psychother. 36 (3) (2022) 256–274.
[34] J. Forcier, P. Bissex, W.J. Chun, Python Web Development With Django, Addison-Wesley Professional, 2008.
[35] C. Lugaresi et al., "Mediapipe: a framework for building perception pipelines," *arXiv preprint arXiv:1906.08172,* 2019.
[36] S. Mathew, J. Varia, Overview of amazon web services, Amaz. Whitepapers 105 (1) (2014) 22.
[37] C.-C. Chen, S.-L. Cho, R.-Y. Tseng, Taiwan corpora of Chinese emotions and relevant psychophysiological data—behavioral evaluation norm for facial expressions of professional performer, Chin. J. Psychol. (2013).
[38] Z. Liu, et al., Swin transformer: hierarchical vision transformer using shifted windows, in: Proceedings of the IEEE/CVF International Conference On Computer Vision, 2021, pp. 10012–10022.
[39] P. Yellamma, P.R. Varun, N.C.N.L. Narayana, Y. Chowdary, P. Manikanth, K.H. G. Sai, Automatic and multilingual speech recognition and translation by using Google Cloud API, in: 2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI), IEEE, 2024, pp. 566–571.
[40] K. Miao, O. Biermann, Z. Miao, S. Leung, J. Wang, K. Gai, Integrated parallel system for audio conferencing voice transcription and speaker identification, in: 2020 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), IEEE, 2020, pp. 1–8.
[41] M. Łajszczak et al., "BASE TTS: lessons from building a billion-parameter text-to-speech model on 100K hours of data," *arXiv preprint arXiv:2402.08093,* 2024.
[42] *ChatGPT.* (2024).
[43] *Claude 3.5 Sonnet.* (2024).
[44] A. Rosebrock, "Find distance from camera to object/marker using Python and OpenCV," *línea].[Consultado: 10 de Octubre de 2019]. Disponible en:* https://www.pyimagesearch.com/2015/01/19/find-distancecamera-objectmarker-using-python-opencv*,* 2015.
[45] F. Fiani, S. Russo, C. Napoli, An advanced solution based on machine learning for remote emdr therapy, Technologies 11 (6) (2023) 172.
[46] S. Rollins, M. Cohen-Peck, Guidelines for virtual EMDR therapy, in: A Report Of The Virtual Training And Therapy Task Group. Presented to the EMDRIA Board of Directors, 2020.
[47] M. Bie, H. Xu, Y. Gao, K. Song, X. Che, Swin-FER: swin Transformer for facial expression recognition, Appl. Sci. (2076-3417) 14 (14) (2024).
[48] G. Benitez-Garcia, T. Nakamura, M. Kaneko, Multicultural facial expression recognition based on differences of western-caucasian and east-asian facial expressions of emotions, IEICE Trans. Inf. Syst. 101 (5) (2018) 1317–1324.
[49] L.-C. Chen, "A socio-pragmatic analysis of Taiwanese and Polish humor: casual conversations and television variety shows," 2016.
[50] Y.-C. Chiang, H.-S. Sheng, W.-C. Wu, L.-L. Yen, Clusters of humor expression and related factors among junior high school students in Taiwan, Bull. Educ. Psychol. 46 (2) (2014) 289–310.
[51] S. Zhang, Y. Yang, C. Chen, X. Zhang, Q. Leng, X. Zhao, Deep learning-based multimodal emotion recognition from audio, visual, and text modalities: a systematic review of recent advancements and future prospects, Expert Syst. Appl. 237 (2024) 121692.
[52] R.A. Calix, S.A. Mallepudi, B. Chen, G.M. Knapp, Emotion recognition in text for 3-D facial expression rendering, IEEE Trans. Multimed. 12 (6) (2010) 544–551.