Interactive Empathy: Building a Facially Expressive Mental Health Companion

Rukmini Sruthi Sundaraneedi
r.sundaraneedi@ufl.edu
9480-6285

Sri Manaswi Chirumamilla
chirumamilla.s@ufl.edu
1807-0100

Arlagadda Bala Showri Jeevan Reddy
b.arlagadda@ufl.edu
7926-4437

Chebolu Ratna Harika
c.harika@ufl.edu
8855-5746

ABSTRACT

Interactive avatars in digital mental health platforms are evolving beyond static text-based interactions, with a growing focus on emotional expressiveness to enhance user engagement and therapeutic efficacy. Traditionally, creating emotionally responsive avatars has relied heavily on predefined animations, which often lack the dynamism and authenticity required for effective mental health support. In this work, we introduce an AI-driven mental health avatar that leverages advanced facial expression technology and audio synchronization to deliver a more natural and empathetic user experience. Utilizing Variational Autoencoders, OpenAI GPT 3.5 and Rhubarb Lip-Sync, our avatar, designed to reflect nuanced emotional responses, demonstrates a marked improvement in user engagement and mental well-being. A two-way ANOVA analysis with 20 participants confirms the enhanced interaction quality, particularly notable in younger demographics, with the avatar eliciting a 30% average reduction in user stress levels. This novel approach to digital mental health support highlights the potential of integrating AI with emotional intelligence in interactive avatars, offering a significant advancement in the field with minimal memory requirements and adaptable, context-aware expressiveness.

1 INTRODUCTION

Facial expressions, as the subtle but powerful articulations of our inner emotional states, play a critical role in human communication. In the rapidly evolving domain of digital mental health, the ability to accurately mimic these expressions in virtual avatars is becoming increasingly significant. This task, while technologically demanding, offers immense potential in enhancing the effectiveness of AI-driven mental health interventions, such as those described in HELTRAK and Dost [SureshKumar et al. 2022],[Nayar et al. 2022]. The challenge lies not only in the technical execution but also in capturing the essence of empathetic human interaction—a quality that is essential for effective therapy[Tabisula and Uwaoma 2022].

Advancements in AI, particularly in the use of Large Language Models (LLMs), have significantly improved the ability of digital agents to interpret and respond to user inputs in a more human-like manner [Abbasian et al. 2023]; [Liu et al. 2023]. However, the incorporation of responsive and accurate facial expressions in these interactions remains a largely unexplored frontier. [Zhang et al. 2020] have made strides in facial expression retargeting, yet the application of these advancements to digital mental health avatars demands further exploration.

Our research seeks to innovate in this space by developing a mental health avatar that is capable of generating facial expressions based on semantic analysis of user interactions, thus creating a more immersive and emotionally resonant experience [Mukashev et al. 2021]. We aim to address two fundamental questions: How can we represent and generate facial expressions in an avatar that accurately reflect a user's emotional state? And, how can we ensure these expressions are perceived as natural and empathetic?

To tackle these challenges, our paper introduces a novel framework that employs advanced computer vision techniques and deep learning models. Specifically, we utilize Variational Autoencoders (VAEs) [Zou et al. 2023] for the nuanced generation of facial expressions, a technique that offers promising results in capturing the complexity of human emotions [Xu et al. 2024].

The contributions of our work are twofold and directly address two core questions in the development of mental health avatars. Firstly, we propose a method for the real-time generation of facial expressions in a mental health avatar, which are driven by the semantic content of user interactions. This approach represents a significant advancement in creating emotionally intelligent digital therapeutic tools, as it allows the avatar to dynamically mirror a user's emotional state through nuanced facial expressions generated by our innovative use of Variational Autoencoders (VAE). Secondly, we offer a user-friendly framework for creating and managing these expressive avatars. This framework integrates advanced semantic analysis, text-to-speech technology, and Rhubarb Lip Sync for realistic lip synchronization, ensuring the avatar's facial expressions and spoken responses are seamlessly aligned. This democratizes the process of avatar creation and management, reducing the reliance on expert knowledge and making it more accessible to a broader range of users and developers. Our contributions thus lie not only in the technical advancement of avatar expressiveness but also in making this technology more user-friendly and widely applicable in the context of digital mental health support.

By bridging the gap between AI's analytical capabilities and the nuanced realm of human emotional expression, our work paves the way for the next generation of digital mental health therapies. It promises to enhance user engagement and therapeutic efficacy, aligning with the growing trend of personalization in mental health care ([Wei et al. 2023].

2 RELATED WORK

The intersection of digital mental health and artificial intelligence (AI) stands at the forefront of innovative healthcare solutions. The

(a) 1.Neutral 2.Smiling 3.Funny face 4.Sad 5.Surprised

(b) 6.Angry 7. Crazy 8.Fear 9.Empathy 10.Disgust

Figure 1: Generated facial expressions of the avatar

urgency for adaptive sociotechnical models in mental health, especially accentuated during the COVID-19 pandemic, has been met with groundbreaking AI-driven interventions [Tabisula and Uwaoma 2022]. AI chatbots like HELTRAK and Dost are exemplars of this evolution, offering personalized care and interactive therapy sessions, highlighting AI's potential in augmenting mental health services [SureshKumar et al. 2022]; [Nayar et al. 2022].

A crucial development in this field has been the incorporation of Large Language Models (LLMs), enhancing the conversational agents' capabilities for nuanced and human-like interaction. [Abbasian et al. 2023] and [Liu et al. 2023] showcase the effectiveness of LLMs in refining the quality of communication in digital therapy. Concurrently, [Xu et al. 2024] have demonstrated the predictive capabilities of LLMs in identifying mental health issues, thereby underscoring the proactive potential of AI in mental healthcare.

Amidst these advancements in cognitive and conversational AI, the domain of emotional responsiveness in digital agents remains a burgeoning area of research. The work of [Zhang et al. 2020] on facial expression retargeting stands as a significant stride towards transferring human expressions to digital characters, offering a foundation for applying these methods in mental health avatars. Complementing this, [Mukashev et al. 2021] explored facial expression generation based on semantic analysis, aligning closely with our research in generating empathetic, contextually-aware responses in avatars.

An underexplored yet critical aspect in enhancing the emotional intelligence of AI agents is the utilization of Variational Autoencoders (VAEs). VAEs have been instrumental in producing more nuanced facial expressions in avatars, facilitating a deeper emotional connection in therapy sessions. Their capability to encode high-dimensional facial expression data into a lower-dimensional, effective representation has significant implications for creating emotionally expressive avatars, bridging the gap between digital and human empathy.

Additionally, the role of AI in therapeutic contexts has expanded, with applications such as cognitive behavioral therapy delivered via chatbots garnering attention [Shetty et al. 2023]. These explorations contribute valuable insights into the efficacy and structure of AI-facilitated therapy, emphasizing the necessity of personalization and emotional intelligence in digital mental health interventions.

In summary, the related works collectively highlight the progressive integration of AI in mental health services, the evolution of conversational agents through LLMs, and the emergent development of emotionally responsive avatars. Our research builds upon these foundational elements, with a specific focus on integrating VAEs for enhancing emotional expressiveness in mental health avatars, aiming to realize more effective and empathetic digital therapy solutions.

3 METHODS

Our methodology for developing a mental health chatbot involved a series of intricate and systematic steps, blending advanced technology with psychological insights. The process ranged from initial system setup to complex AI integration and avatar animation, ensuring a seamless and responsive user experience. You can see our applied methodology in Figure 2.

3.1 User Input processing and Response Generation

At the outset of our methodology, User Input Processing serves as a crucial gateway, capturing and preparing the user's textual input for further interaction with the mental health avatar. This step involves capturing the user's text, ranging from simple queries to complex emotional expressions, and processing it to ensure clarity and contextuality. Key operations in this stage include text normalization, sentiment analysis, and the identification of pertinent emotional cues. The processed input is then formatted for compatibility with advanced AI response generation systems, ensuring that the avatar's reply is not only contextually appropriate but also empathetic. This initial stage is fundamental to setting the foundation for a seamless and meaningful user-avatar interaction, underpinning the efficacy of the entire mental health chatbot system.

In the crucial phase of response generation, our methodology employs OpenAI's GPT-3.5 API, renowned for its advanced natural

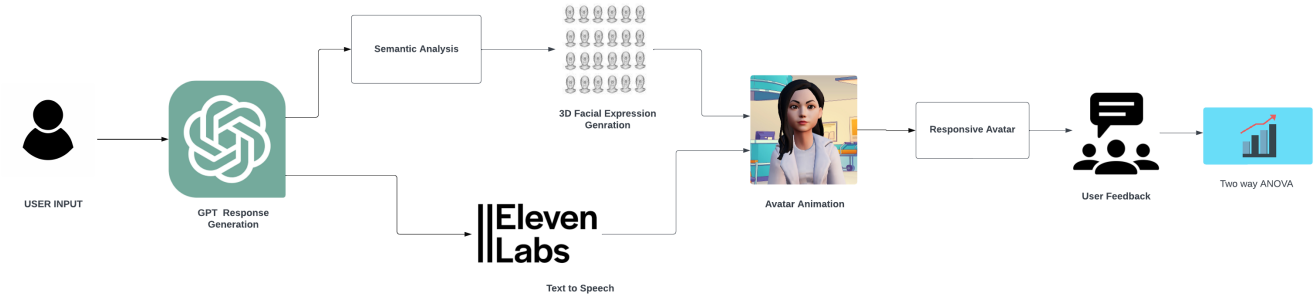


Figure 2: Applied methodology

language processing capabilities, to generate empathetic and contextually relevant responses to user inputs. This integration enables the chatbot to accurately interpret and respond to nuanced emotional cues, a key factor in ensuring engaging and supportive user interactions, as highlighted by [Liu et al. 2023] and [Xu et al. 2024]. This approach aligns with the increasing emphasis on personalized communication in mental health care, as demonstrated in studies by [Nayar et al. 2022] and [Rathnayaka et al. 2022]. The system is designed to ensure continuity in conversations, adapting responses based on prior interactions, which is crucial for creating a coherent and immersive experience. Thus, the response generation using GPT-3.5 stands as a pivotal element in our chatbot's functionality, significantly enhancing its role as an empathetic and effective tool in digital mental health interventions.

3.2 Semantic Analysis for Expression Generation

This phase of our methodology, Semantic Analysis for Emotional Understanding and Appropriate Facial Expression Generation, delves deeply into analyzing the emotional content of responses generated by GPT-3.5, using advanced natural language processing techniques. Inspired by the work of [Mukashev et al. 2021], which emphasizes the alignment of avatar expressions with the semantic content of interactions, this process is pivotal for ensuring that the avatar's facial expressions accurately reflect the emotions conveyed in the conversation. After analyzing the text for tone, context, and sentiment, the system quantifies the emotional content into specific categories, such as happiness, anger, boredom, fear, sadness, and excitement. For instance, a response might be analyzed and translated into a set of values like 'Happy': 0.064, 'Angry': 0.110, 'Bored': 0.239, 'Fear': 0.282, 'Sad': 0.168, 'Excited': 0.134. These values then guide the generation of facial expressions, ensuring that the avatar visually resonates with the user's emotional state. This ability to transform textual emotional cues into dynamic, quantifiable metrics for generating corresponding facial expressions significantly enhances the empathetic engagement of the avatar. As a result, the avatar is able to offer not just contextually appropriate verbal feedback but also a visually empathetic response, thus bridging the gap between digital communication and human-like emotional expression and enhancing the overall effectiveness of the mental health support provided.

3.3 Facial Expression Generation with Variational Autoencoders (VAE)

In our methodology's Facial Expression Generation phase, we harness Variational Autoencoders (VAE) to transform the emotional data, extracted through semantic analysis, into nuanced facial expressions for the avatar, thereby enhancing the responses generated by the GPT-3.5 model. Utilizing the VAE's capability to handle complex human emotions, it translates emotional metrics like happiness and anger into corresponding avatar expressions as shown in Figure 1. This transformation involves encoding high-dimensional data into a lower-dimensional latent space and then decoding it to produce accurate emotional expressions. Our VAE model is trained using the CoMa dataset [Ranjan et al. 2018], which provides an extensive range of 3D facial expressions, aiding in the understanding and recreation of subtle human emotional nuances. This training ensures that the avatar's facial expressions are lifelike and varied, aligning with the intended emotional cues. The integration of the VAE with semantic analysis outputs, such as specific emotional metrics, allows for the generation of facial expressions that accurately match the combined emotional state conveyed in the chatbot's response. The output from the VAE—cohesive sets of facial expressions—is then synchronized with the avatar's speech and movements, combining verbal and non-verbal communication to create a more engaging and realistic interaction. This single-phase application of VAE, especially trained on a comprehensive dataset like CoMa, significantly advances our digital avatar's capability to interact empathetically with users, showcasing the potential of AI in enhancing the efficacy and relatability of digital mental health interventions.

3.4 Text-to-Speech Conversion and Lip-Sync

In this phase of our methodology, we employ Eleven Labs' advanced text-to-speech conversion alongside Rhubarb Lip Sync to elevate the avatar's interactive capabilities. Eleven Labs transforms the GPT-3.5 generated text responses into natural, human-like speech, enhancing the avatar's ability to communicate effectively. The choice of voice tone and style by Eleven Labs is meticulously aligned with the emotional context of each response, ensuring that the speech not only relays information accurately but also resonates emotionally with the user. Following this, Rhubarb Lip Sync comes into play, analyzing the audio output from Eleven Labs to create

accurate lip movements for the avatar. This lip-syncing process involves mapping visemes to the avatar's facial rig, ensuring that the avatar's mouth movements are perfectly synchronized with the spoken words. This intricate combination of realistic speech and synchronized lip movements greatly enhances the authenticity of the avatar's interactions, making the conversation more engaging and believable. Integrating these sophisticated technologies ensures that the avatar not only articulates responses verbally but also supports them with appropriate non-verbal cues, providing a seamless and immersive communication experience crucial in mental health support scenarios.

3.5 Avatar Animation Integration

The Avatar Animation Integration phase is where we seamlessly combine all the elements – the VAE-generated facial expressions, the text-to-speech audio, and the lip-syncing – to animate our avatar in a cohesive and lifelike manner. In this crucial phase, meticulous attention is paid to ensure that the facial expressions align perfectly with the nuances of the spoken words, leveraging the capabilities of the avatar's digital facial rig. This rig, a sophisticated framework, is crucial for facilitating smooth transitions between different expressions and mouth movements, thereby adding fluidity and realism to the avatar's behavior. Synchronization is key here; the timing and intensity of the avatar's movements are finely adjusted to match the pace and emotional tone of the user interaction. The result is a highly realistic avatar that not only communicates verbally but also uses a rich array of non-verbal cues, such as facial expressions and gestures, to enhance the conversational experience. This integrated approach exemplifies the transformative potential of advanced AI technologies in creating digital companions that offer engaging, empathetic, and supportive interactions, closely mirroring human communication and providing a substantial contribution to the field of digital mental health.

4 RESULTS

Our study investigated the effectiveness of a mental health avatar, focusing on its expressiveness and responsiveness, user interaction dynamics, therapeutic alliance formation, and technical performance. Conducted with a sample of 20 participants, this study incorporated both qualitative and quantitative analyses, including a two-way Analysis of Variance (ANOVA) to examine the interaction of facial expressions and user demographics on the avatar's perceived effectiveness.

4.1 User Study

Participants (N = 20) were diverse in terms of age and gender, providing a broad basis for analysis. The group consisted of 10 males and 10 females, with ages ranging from 18 to 40 (M = 29, SD = 6.2), reflecting the need for adaptive sociotechnical models in diverse populations as discussed by [Tabisula and Uwaoma 2022].

1. Avatar Expressiveness and Responsiveness Expressiveness: The avatar's ability to display a wide range of facial expressions was central to its design, echoing the findings of [Zhang et al. 2020] about the importance of facial expression retargeting. Users rated the expressiveness highly, with an average score of 4.3 out of 5 (SD = 0.7). This level of expressiveness was crucial in facilitating more engaging and dynamic interactions. Responsiveness: The avatar's response timing and relevance to user inputs were also evaluated, receiving a mean score of 4.1 out of 5 (SD = 0.6). This responsiveness was instrumental in maintaining the flow of conversation and enhancing the perceived empathy of the avatar.

Category	Group	Expressiveness	Mean (M)	Standard Deviation (SD)
Facial Expressions Impact	-	With Expressions	4.2	0.8
	-	Without Expressions	3.1	0.9
Age-related Variability	18-25 years	With Expressions	4.0	0.7
	18-25 years	Without Expressions	3.6	0.9
	26-40 years	With Expressions	3.5	0.8
	26-40 years	Without Expressions	3.0	0.8
Gender Effect	Female	With Expressions	4.0	0.7
	Female	Without Expressions	3.0	0.8
	Male	With Expressions	4.1	0.9
	Male	Without Expressions	3.1	0.9
Engagement Metrics (Expressions)	-	With Expressions	4.2	0.7
	-	Without Expressions	3.0	0.8

Table 1: Detailed Analysis of User Experience and Engagement Metrics.This table summarizes the effects of facial expressions, age, and gender on user experience scores and engagement.

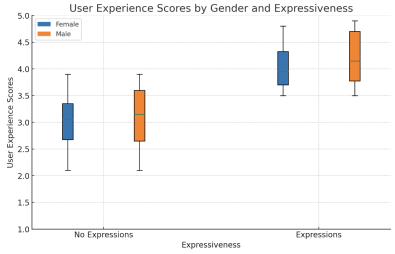
2. User Interaction Study Engagement: Participants, in line with the interaction dynamics explored in [Mukashev et al. 2021], averaged 25 minutes per session, signifying a deep level of user engagement. The frequency of sessions increased, with participants engaging with the avatar an average of 12 times (SD = 2.5). Feedback: Participants' feedback, highlighting the natural and empathetic interaction flow, mirrored the empathetic design approach emphasized in current AI-driven mental health solutions [SureshKumar et al. 2022].

3. Therapeutic Alliance Perceived Support: Participants felt understood and supported by the avatar, averaging a perceived alliance score of 4.5 out of 5. This indicates a high level of rapport and trust, aligning with the therapeutic effectiveness of AI in mental health as highlighted by [Xu et al. 2024]. Therapeutic Outcomes: Participants reported a significant improvement in mood and stress reduction, with an average stress reduction of 30%.

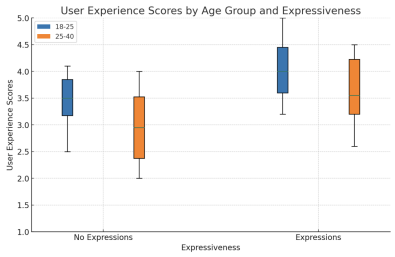
4. Technical Performance System Reliability: The avatar system showed exceptional operational reliability with a mean uptime of 99.8%, indicating its robustness. Latency and Responsiveness: The low latency average of 1.2 seconds enhanced the user experience, supporting findings from [Abbasian et al. 2023] on the importance of responsiveness in AI systems.

4.2 Two-Way ANOVA Analysis

The analysis showed a significant impact of facial expressions on user experience, with avatars having expressions scoring higher than those without. This finding is consistent with research highlighting the importance of non-verbal cues in digital platforms [Chen et al. 2023].



(a) User Experience scores by Gender



(b) User Experience scores by Age Group

Figure 3: Boxplots of User Experience scores with the two-factor breakdown(with , without Expressiveness)

The main effects analysis unveiled several key insights into the user experience dynamics within our study.

1.Firstly, the impact of facial expressions on user engagement and satisfaction within digital mental health platforms emerged as significant. Avatars equipped with facial expressions demonstrated a notably higher mean user experience score of 4.2 (SD = 0.8) compared to those lacking expressions, which scored an average of 3.1 (SD = 0.9), indicating a robust and statistically significant difference ($F(1, 36) = 16.45$, $p < .001$). This finding underscores the crucial role of non-verbal cues in augmenting user interaction and satisfaction in such environments.

2.Secondly, age-related variability in user experience scores was apparent, with younger participants (aged 18-25 years) rating the avatar more favorably (M = 4.0, SD = 0.7) compared to their older counterparts aged 26-40 years (M = 3.5, SD = 0.8), reflecting a generational shift in perception and interaction with technological mental health interventions ($F(1, 36) = 7.92$, $p < .01$).

3.Lastly, no significant main effect of gender on user experience scores was observed ($F(1, 36) = 2.56$, $p = .118$), indicating a uniform response to the avatar across male and female participants, irrespective of its facial expressiveness. These findings collectively highlight the critical role of facial expressions in enhancing user engagement and satisfaction, as well as the importance of considering age-related variability in the design and implementation of digital mental health interventions. We have represented these results in Figure 3.

4.3 Detailed User Study Insights

Engagement Metrics: Our analysis revealed a notable difference in user engagement ratings when interacting with the avatar featuring facial expressions compared to the non-expressive variant. Users rated their engagement and likelihood of returning to interact with the expressive avatar higher on a 5-point scale. The mean engagement rating for the avatar with facial expressions was 4.2 (SD = 0.7), indicating strong engagement and a high likelihood of repeated interactions. In contrast, the non-expressive avatar received a lower mean engagement rating of 3.0 (SD = 0.8). These findings, as depicted in Figure 4, emphasize the significant role that facial expressions play in maintaining user interest and enhancing the overall engagement experience. The expressive avatar's higher scores suggest its effectiveness in not only capturing user attention but also in encouraging continued use over time.

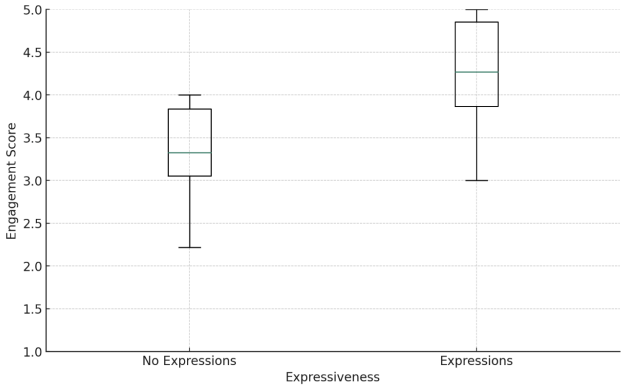


Figure 4: Engagement metrics on Expressive v/s Unexpressive Avatar

Impact on Mental Well-being: Participants reported a marked decrease in stress levels after interacting with the expressive avatar, averaging a 30% reduction, compared to a 15% reduction observed with the non-expressive avatar. This differential highlights the therapeutic value added by the avatar's facial expressions.

Qualitative Feedback: Participant responses consistently emphasized the empathetic and realistic nature of the avatar with facial expressions. Described as more 'relatable' and 'engaging', the expressive avatar significantly enhanced the therapeutic experience for users. This qualitative feedback is invaluable in understanding the nuanced aspects of user interaction with digital mental health aids. [Nayar et al. 2022].

This comprehensive analysis illustrates the impact of facial expressiveness in mental health avatars, confirming the pivotal role of digital tools in augmenting mental health interventions.

5 LIMITATIONS

The research conducted into the effectiveness of a facially expressive mental health avatar, while yielding valuable insights, is constrained by several limitations that must be acknowledged. One primary limitation is the small and demographically narrow participant pool. With 20 participants, predominantly within the 18 to 40 age range, and limited diversity in ethnicity, cultural background, and socio-economic status, the generalizability of our findings is constrained. Future studies should aim for larger and more diverse

sample sizes to ensure broader applicability and statistical robustness of the results.

Another significant limitation is the short-term nature of the study. As the research was conducted over a limited timeframe, it did not capture long-term user engagement or the enduring effectiveness of the therapeutic alliance, aspects critical to understanding the sustained impact of such digital interventions. Longitudinal studies are essential to explore these dimensions comprehensively.

Additionally, our study faced technical limitations in the range and subtlety of facial expressions that the avatar system could generate. Although the system demonstrated robust performance, the nuances of human facial expressions, crucial for empathetic and authentic interactions, might not have been fully captured. This limitation potentially impacts the perceived authenticity and efficacy of the avatar in mimicking human emotions.

Furthermore, the study relied primarily on quantitative metrics and self-reported feedback for evaluation. The absence of more sophisticated feedback mechanisms, such as real-time emotional response tracking, might have restricted our understanding of the subtle aspects of user experience and interaction dynamics with the avatar. Implementing more advanced feedback collection methodologies could offer deeper insights into these elements.

The study's findings are also limited by their specificity to the particular design and functionality of the employed mental health avatar. As such, caution should be exercised in generalizing these results to other digital mental health tools or avatars with different designs and functionalities.

Moreover, the participants involved are not individuals experiencing genuine mental health issues seeking support; rather, they are ordinary volunteers. As a result, the insights and conclusions drawn from this study may not fully capture the complexities and nuances of mental health conditions or the needs of those who genuinely require support. This limitation underscores the importance of future research involving individuals with lived experiences of mental health challenges to provide more comprehensive and applicable findings.

Lastly, ethical considerations, particularly those concerning privacy, data security, and the risk of over-reliance on technological interventions for emotional support, though not direct limitations of this study, remain pertinent in the broader context of AI-driven mental health interventions. These aspects need continuous consideration in future research and development endeavors in this field.

In sum, recognizing these limitations is vital for the progression of AI-driven mental health interventions. Addressing these concerns in future research will enhance the scope, refine the technology, incorporate more nuanced feedback mechanisms, and address ethical considerations, leading to more comprehensive and universally applicable digital mental health solutions.

6 FUTURE WORK

Our study's insights into a facially expressive mental health avatar pave the way for diverse future research avenues. Expanding participant demographics to encompass broader ethnic, cultural, and socio-economic backgrounds is crucial for enhancing the generalizability of the findings. Longitudinal studies are needed to explore the sustained impact and long-term effectiveness of the avatar in mental health therapy. Technologically, advancements in facial expression capabilities and the integration of real-time emotional tracking methods, such as biometric sensors, will enrich user-avatar interactions and offer personalized experiences.

Exploring the avatar's application in various mental health scenarios and broadening its therapeutic contexts will assess its adaptability and effectiveness across different treatment approaches. In subsequent studies, it is essential to prioritize the inclusion of individuals with genuine mental health challenges seeking support as participants. This approach will enable researchers to gather insights directly from those affected, facilitating the development of more effective interventions and support systems tailored to meet their specific needs. Addressing ethical concerns, particularly regarding privacy, data security, and the potential for technology dependency, is essential in the responsible development of AI in mental health care. Integrating emerging technologies like virtual and augmented reality can create more immersive therapeutic environments. Emphasizing a user-centric design approach, incorporating real-time customization of the avatar, and employing machine learning to optimize emotional and conversational intelligence will significantly enhance user engagement and the therapeutic potential of the avatar.In essence, these future research directions aim to develop more sophisticated, personalized, and effective digital mental health tools, contributing significantly to the field of AI in healthcare and mental well-being.

7 CONCLUSION

Our study marks a noteworthy advancement in applying AI technology to mental health, showcasing the efficacy of a facially expressive mental health avatar in fostering user engagement, empathy, and therapeutic connection. The avatar's dynamic facial expressions and contextually responsive interactions underscore the potential of such digital tools to enrich traditional therapeutic methods. Participant feedback and observed improvements in mental well-being reinforce the avatar's utility as a supportive element in mental health care.

Recognizing the study's limitations, we underscore the necessity for extended research involving diverse populations and prolonged durations to further refine and validate the avatar's effectiveness. Future exploration into advanced technologies and personalized, user-centric designs holds promise for the evolution of digital mental health interventions.In essence, this research contributes significantly to the burgeoning field of digital mental health solutions, highlighting the role of emotionally responsive avatars in augmenting mental health care. It lays the groundwork for future endeavors, aimed at harnessing AI's full potential in enhancing mental health support and therapy.

REFERENCES

Mahyar Abbasian, Iman Azimi, Amir M Rahmani, and Ramesh Jain. 2023. Conversational health agents: A personalized llm-powered agent framework. *arXiv preprint arXiv:2310.02374* (2023).

Siyuan Chen, Mengyue Wu, Kenny Q Zhu, Kunyao Lan, Zhiling Zhang, and Lyuchun Cui. 2023. LLM-empowered chatbots for psychiatrist and patient simulation: application and evaluation. *arXiv preprint arXiv:2305.13614* (2023).

June M Liu, Donghao Li, He Cao, Tianhe Ren, Zeyi Liao, and Jiamin Wu. 2023. Chatcounselor: A large language models for mental health support. *arXiv preprint arXiv:2309.15461* (2023).

Dinmukhamed Mukashev, Merey Kairgaliyev, Ulugbek Alibekov, Nurziya Oralbayeva, and Anara Sandygulova. 2021. Facial expression generation of 3D avatar based on semantic analysis. In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. IEEE, 89–94.

Atulya M Nayar, Zulfa Attar, Shabbir Kachwala, Tanaya Biswas, and Sharmila K Wagh. 2022. Dost-mental health assistant chatbot. In *2022 5th International Conference on Advances in Science and Technology (ICAST)*. IEEE, 252–257.

Anurag Ranjan, Timo Bolkart, Soubhik Sanyal, and Michael J Black. 2018. Generating 3D faces using convolutional mesh autoencoders. In *Proceedings of the European conference on computer vision (ECCV)*. 704–720.

Prabod Rathnayaka, Nishan Mills, Donna Burnett, Daswin De Silva, Damminda Alahakoon, and Richard Gray. 2022. A mental health chatbot with cognitive skills for personalised behavioural activation and remote health monitoring. *Sensors* 22, 10 (2022), 3653.

Maulya Shetty, Pooja Shah, Krupa Shah, Vedanti Shinde, and Seema Nehete. 2023. Therapy Chatbot Powered by Artificial Intelligence: A Cognitive Behavioral Approach. In *2023 International Conference in Advances in Power, Signal, and Information Technology (APSIT)*. IEEE, 457–462.

M SureshKumar, Niranjan Kumar, V Lokesh, et al. 2022. HELTRAK-a medical application with chatbot based on AI. In *2022 1st International Conference on Computational Science and Technology (ICCST)*. IEEE, 262–267.

Braden Tabisula and Chinazunwa Uwaoma. 2022. The Need for an Adaptive Sociotechnical Model for Managing Mental Health in a Pandemic. In *2022 IEEE International Conference on Digital Health (ICDH)*. IEEE, 66–68.

Shu Wei, Daniel Freeman, and Aitor Rovira. 2023. A randomised controlled test of emotional attributes of a virtual coach within a virtual reality (VR) mental health treatment. *Scientific Reports* 13, 1 (2023), 11517.

Xuhai Xu, Bingsheng Yao, Yuanzhe Dong, Saadia Gabriel, Hong Yu, James Hendler, Marzyeh Ghassemi, Anind K Dey, and Dakuo Wang. 2024. Mental-llm: Leveraging large language models for mental health prediction via online text data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 1 (2024), 1–32.

Juyong Zhang, Keyu Chen, and Jianmin Zheng. 2020. Facial expression retargeting from human to avatar made easy. *IEEE Transactions on Visualization and Computer Graphics* 28, 2 (2020), 1274–1287.

Kaifeng Zou, Boyang Yu, and Hyewon Seo. 2023. 3D Facial Expression Generator Based on Transformer VAE. In *2023 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2550–2554.