Grenze International Journal of Engineering and Technology, January Issue

Integrating Artificial Intelligence with Virtual Reality for Cognitive Therapy

Nuzhat Shaikh[1], Ishita Jagtap[2], Vihaa Terkar[3] and Ishita Chaudhari[4]
[1-4]Wadia College of Engineering, Pune, India
Email: nfshaikh@mescoepune.org, jishita2003@gmail.com, vihaaterkar@gmail.com, ishitachaudhari1664@gmail.com

***Abstract*— This literature review explores how Artificial Intelligence (AI), Machine Learning (ML), and immersive Virtual Reality (VR) come together to transform cognitive therapies in mental health, particularly in managing anxiety disorders and enhancing emotional regulation. It draws on from a wide range of sources, including research papers, journal articles, and systematic reviews to show the evolution of digital therapies and their current state. The reviewed studies consistently show that Virtual Reality Exposure Therapy (VRET) delivers outcomes comparable to conventional Cognitive Behavioural Therapy (CBT) for various anxiety-related conditions, such as phobias, generalized anxiety, and post-traumatic stress disorder (PTSD). A growing area of interest involves integrating biofeedback mechanisms that capture real-time physiological signals, such as Heart Rate Variability (HRV) and Electroencephalogram (EEG) data, to quantify emotional and bodily responses of a patient during therapy sessions. Although AI and ML are increasingly applied to build adaptive, data-driven therapeutic systems, current models have limitations. Traditional VRET often lacks personalization, while advanced deep learning frameworks demand large datasets and still struggle with interpretability. To overcome these challenges, this paper introduces a closed-loop, multimodal, adaptive framework that seamlessly connects physiological sensing with dynamic virtual environments, creating a more personalized and responsive therapeutic experience.**

***Index Terms*— AI-assisted Cognitive Therapy, Virtual Reality (VR), Artificial Intelligence (AI), Machine Learning (ML), Biofeedback, Cognitive Behavioral Therapy (CBT), Exposure Therapy, Multimodal Data, Anxiety Disorders, and Mental Health.**

I. INTRODUCTION

Virtual Reality (VR) has evolved significantly from its beginnings in gaming and entertainment. Today, it is transforming industries like education, retail, construction, and notably, healthcare. In medicine, virtual reality and augmented reality are being used for assistance during surgeries. This involves wearing of the goggles while performing the surgery, which allows the surgeon to see 3D projections of Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) scans, allowing a better surgical approach for the patient. Extended reality is not only being used to educate patients about their surgical procedures, but also medical students and trainee doctors for more visualised and interactive learning [1].
Mental health disorders, particularly anxiety-related conditions like phobias, Post-Traumatic Stress Disorder (PTSD), and Generalised Anxiety Disorder (GAD), represent a significant global health challenge.

Grenze ID: 01.GIJET.12.1.168_2

In 2017, according to the data published by the World Health Organisation (WHO), 38 million people in India were estimated to suffer from anxiety disorders, second highest after depression. In 2021, 359 million people worldwide had an anxiety-related disorder, making anxiety the most common mental health condition globally [2].

The established "gold standard" for psychotherapy treatment is Cognitive Behavioral Therapy (CBT), with exposure therapy being one of its fundamental components [3]. Exposure therapy is a psychological treatment that involves guiding an individual to confront their feared stimuli in a controlled environment to reduce their fear and anxiety responses. Types of exposure therapy include in vivo (real-life) exposure (directly facing feared object or situation, with gradually increasing exposure), in vitro (imagined) exposure (imagining the feared object or situation), simulation exposure (using virtual reality to simulate feared scenarios), and sensation awareness (creating harmless but feared physical sensations)[4]. However, traditional exposure methods face numerous practical and ethical limitations. As a result, virtual reality emerged as a powerful therapeutic tool, enabling delivery of exposure therapy in immersive, safe, and highly controlled simulated environments. This is known as VR-assisted CBT (VRCBT) or Virtual Reality Exposure Therapy (VRET).

Among the earliest documented therapeutic applications of VR was the treatment of Vietnam War veterans diagnosed with PTSD. Participants in these studies demonstrated substantial improvement, with reduction of PTSD symptoms by approximately 34% [5]. With recent developments, VR therapy has expanded beyond trauma treatment to address specific phobias such as arachnophobia (fear of spiders) and aerophobia (fear of flying). More recent developments have extended its use to complex psychological conditions, including obsessive-compulsive disorder (OCD) and autism spectrum disorders [3].

The current evolution of this approach involves augmenting the current VR systems with biofeedback sensors and Artificial Intelligence (AI). By integrating sensors, we can capture real-time physiological data such as HRV, EEG, and other signals, to better understand a patient's internal state instead of relying on subjective self-reports. Machine Learning (ML) algorithms can then process this data and dynamically adjust VR scenarios in real-time to create adaptive and personalised experiences [6]. The literature review will trace the evolution of these technologies. It will first examine the efficacy of VR as an enhancement to traditional CBT, followed by the role of multimodal biofeedback, concluding with the current research on how AI and ML are being applied to create truly adaptive and personalised interventions. Through this analysis, we will also see the key research gaps that motivate the present study.

II. Search Strategy and Selection Criteria

A systematic literature search was conducted across IEEE Xplore, PubMed, Scopus, Web of Science, and PsycINFO to identify studies examining the integration of AI, ML, and immersive technologies like VR or AR with cognitive or behavioral therapies for mental health. The search focused on studies published between 2020 and 2025. Keywords were selected around four themes: AI/ML techniques, immersive technologies, therapeutic approaches, and mental health applications. A Combination of various terms like "Artificial Intelligence", "Machine Learning", "Deep Learning", "Virtual Reality", "Augmented Reality", "Cognitive Behavioral Therapy", and "Anxiety" was used in the search.

Conference papers, peer-reviewed articles, systematic reviews and meta-analyses having a main focus on integrating AI/ML with AR/VR for mental health related conditions were selected. Approximately 130 records were yielded through the initial search. After filtering these according to their relevance, they were categorised by AI technique, type of biofeedback, target mental health disorder, or study the design to facilitate a structured analysis of the findings.

III. Literature Review

Although CBT has long been established as an effective treatment approach, its real-world implementation and accessibility face multiple challenges. While digital interventions have attempted to overcome some of these barriers, issues related to effectiveness, scalability, and user acceptance continue to persist. Traditional in-person therapy is limited by a shortage of trained professionals, high treatment costs, and geographical constraints that make access uneven. In addition, social stigma remains a significant barrier with many individuals avoiding seeking help due to concerns about being judged or labeled within clinical environments [7], [8]. Exposure-based therapies, despite their proven success, present unique practical difficulties. In vivo exposure (real-world confrontation with fears) can be logistically difficult, costly (in cases involving fear of flying) and occasionally unsafe or ethically questionable, especially when involving traumatic recreations. Conversely, in vitro exposure (imagined scenarios) depends heavily on a patient's ability to visualize fear-inducing situations, a skill that not all possess. These constraints often result in avoidance behaviors and high dropout rates [3], [8]. Earlier digital formats such as Computer-Based CBT (CCBT) and Internet-Based CBT (ICBT) enhanced accessibility but suffered from low interactivity, poor engagement, and high attrition levels. Similarly, many existing VR-based therapies remain static, relying on uniform virtual environments that fail to account for changes in an individual's emotional or physiological state. The absence of real-time adaptive feedback often leads to experiences that are either insufficiently stimulating or overly intense, reducing overall therapeutic effectiveness and patient retention [7].

A. VR-based Therapy

VR technology provides the ideal platform for exposure therapy by creating immersive environments that are safe, highly controllable, repeatable, and private. The discreet nature of VR therapy can also help reduce the stigma that often prevents individuals from seeking traditional care [6],[9].

This paper [8] reviews VRET as a treatment for mental illnesses, especially for patients who feel shame and stigma towards seeking help. It highlights advantages of VRET over traditional therapy, including high controllability, safety, repeatability, privacy and the ability to create multi-sensory experiences than purely imaginal therapy, which is more beneficial for people who have trouble visualising. The review summarises studies demonstrating the effectiveness of VRET in treating phobias (acrophobia, claustrophobia), relieving depression symptoms, teaching real-world skills to children with autism, and successfully treating PTSD in military veterans. The authors conclude that despite current limitations (like hardware costs), VRET has significant potential to make mental healthcare more accessible after mass production of VR equipment to meet the lack of psychotherapists.

This systematic review and meta-analysis [3] of 11 studies evaluates the effectiveness of VRCBT for anxiety disorders.

The quantitative analysis of 6 randomised controlled trials, with a total of 626 participants (aged 18-65) across all studies, found that VRCBT is significantly more effective than no treatment at reducing both anxiety and depression. The therapeutic effect of VRCBT was also found to be statistically similar to standard CBT, with no significant difference in patient withdrawal rates. Follow-up studies suggested that the effects of VRCBT were maintained at 6 and 12-month follow-ups. The authors conclude that VRCBT is a viable and effective treatment for anxiety disorders with comparable outcomes to the current gold-standard therapy.

This literature review [10] discusses how CBT, the "gold standard" treatment for GAD, can be enhanced with VR. The authors note that GAD is one of the most difficult anxiety disorders to treat using traditional CBT interventions due to its chronic nature and high relapse rates. Even though the outcomes of in vivo exposure and VR exposure are similar, they propose the use of VR as a therapeutic tool, primarily for scenarios that are difficult or impossible to replicate in real life. The paper positions VR not as a more effective tool than standard therapy, but as a more practical and versatile one. It also provides a comprehensive outline of current CBT protocol for GAD and concludes that augmentation of VR with this protocol would produce an efficient and effective treatment method for GAD as well as other anxiety disorders.

This paper [11] describes a proposed study to develop and validate a VRET tool for patients with Obsessive-Compulsive Disorder (OCD) by using a virtual house environment designed to trigger specific OCD symptoms across three dimensions: "contamination/cleaning", "symmetry/ordering", and "fear-of-harm/checking". The study plans to involve 40 OCD patients (aged 18-55) to participate in seven weekly VR sessions combined with standard CBT to confront triggering situations while preventing compulsive responses. Data collection after each session will include questionnaires for anxiety (STAI-6), compulsions (Visual Analogue Scale), simulator sickness (SSQ), clinician-rated symptom severity (Y-BOCS), and behavioral measures within the VR environment. Proposed VR equipment is the HTC Vive with built-in trackers for head and hand movements. The goal of the proposed study is to validate VRET as a therapeutic tool and show a significant decrease in OCD symptoms compared to patients receiving only standard treatment.

B. Biofeedback and Physiological Monitoring

The integration of biofeedback sensors (eg.: Heart Rate Variability (HRV), Electrodermal Activity (EDA), Electroencephalogram (EEG)) enables real-time, objective monitoring of a patient's emotional and physiological responses.This way, we can move beyond reliance on subjective self-reports, and provide rich, multimodal data streams reflecting a patient's authentic internal state [6], [12].

Each year, in the United States of America alone, 50 million surgeries are performed. Most of these surgeries are performed on middle-aged or old-aged adults. After surgeries, patients often go through pain and anxiety. The most common way to treat these symptoms is by prescribing opioids. However, it is accompanied by addiction risks and the risk of becoming a long-term user. To address this issue, the paper [13] proposes a virtual reality experience which analyses your heart rate, depending on which it changes what you see and gets you relaxed. The VR environment basically involved creating a virtual beach integrated with HRV biofeedback to help manage these symptoms. As compared to control groups, this VR experience method showed significant decrease in both surgical pain (29.6%) and anxiety (31.2%), and it was popularly accepted amongst the older adults undergoing Total Knee Arthroplasty (TKA).

The main idea that this paper [12] proposes revolves around combining real-time physiological analysis with VR to have personalisations made to the system. Instead of just looking at the heart rate, the system uses it to transform the VR environment. It helps in creating a dynamic feedback loop which can actively help manage a patient's emotional and cognitive states and facilitate their regulation. It consists of a framework using real-time physiological sensing like heart rate, EEG, HRV, EDA and respiratory rhythms. Abstract cognitive intentions are converted into immersive virtual experiences, enhancing user autonomy and self-regulation. The broader idea is based on real-time biofeedback in which different bio-signals like heart rate and EEG are used to generate dynamic visual/auditory feedback within an immersive VR performance, to enhance digital therapies for conditions like anxiety, PTSD, Attention Deficit Hyperactivity Disorder(ADHD), and depression.

This paper [6] presents an adaptive VRET system for treating phobias by integrating a Meta Quest 3 VR headset with real-time biofeedback using heart rate and Galvanic Skin Response (GSR) sensors. Based on the measured stress levels of the user, the system dynamically adjusts the intensity of the virtual environment, reducing the intensity if anxiety levels become high. The proposed study suggested 20 participants to test the system's effectiveness for fear of heights and spiders.

The study in this paper [14] is based around real-time biofeedback to examine impact on audience perception, immersion, and narrative understanding. The main types of biofeedback augmentation ways explored in the paper are based on heart rate and EEG to analyse how they affect the virtual reality performance for the audience. Dynamic visual and auditory feedback are generated using these biosignals within an immersive VR performance. It uses the performer's heart rate (using an HR monitor) and the audience's brainwaves (using an EEG headband). The VR system is named "Scary Puppet", where the audience watches a witch character, whose movements are controlled by a real performer. The study showed that the performer's heart rate visual feedback showed more concentration, while the audience's EEG visual biofeedback showed improved narrative understanding. This shows that it made the audience more engaged with the performance and the story.

The paper [15] proposes an intervention that combines CBT and Heart Rate Variability Biofeedback (HRV-BF) to reduce

stress, implemented across three different immersive technologies: virtual, augmented, and mixed reality (AR, VR, MR) environments. A volumetric video is used to create a virtual therapist in 3-D style, who will then deliver evidence based CBT, and the HRV-BF will reduce stress by engaging with the body's nervous system. The goal of the project is to evaluate and compare the effectiveness of these stress management tools across the different immersive platforms.

C. AI/ML for Emotion or Disorder Detection

AI and ML play the central role in transforming the raw biofeedback data into meaningful, adaptive actions. The motivation is to use ML algorithms to analyse real-time biofeedback and dynamically adjust the difficulty/intensity of the VR environment. This real-time adaptability helps in maximising the progress while minimising the risk of overwhelming distress [16].

This systematic review [16] analyses 141 studies to explore the integration of ML with AR and VR for cognitive therapies. The review answers six research questions concerning efficacy, personalisation, and technological innovation within these intersecting fields. It concludes that ML-driven interventions offer significant advancements in the same, especially for PTSD, anxiety disorders, and phobias. Predicting treatment success with up to 77% accuracy, these therapies offer durable long-term benefits, maintained for up to six months, and lead to improved patient adherence due to their immersive nature, resulting in lower dropout rates (83% vs. 68%).

This review [7] examines how AI is being incorporated into CBT. The review structures the outcomes based on different stages of the treatment (pre-treatment, therapeutic process, post-treatment) by assessing the details of AI applications, emphasizing the role of chatbots and VR. The review also highlights how AI, especially Large Language Models (LLMs), can be used to make CBT more efficient and personalised. It concludes with how AI should be used to support and enhance, and not replace, human therapists.

This comprehensive literature review [17] systematically categorizes existing research on anxiety detection via ML models based on data modality. The literature review classifies studies into five primary groups: clinical test data, textual data, diverse signals (biological, neurological, audio/video), background information, and psychological features. The review findings confirm that machine learning algorithms effectively detect anxiety from a broad spectrum of features, including physiological signals, brain activity, textual data, speech, facial expressions, and demographic/lifestyle information. Multimodal methodologies, which integrate diverse data sources (e.g., physiological signals with wearable-derived behavioral data), demonstrate considerable potential for improving detection accuracy. Despite notable progress, existing research still faces considerable challenges. Many studies rely on culturally or demographically narrow datasets, lack longitudinal validation, and often overlook comorbidities that accompany mental disorders. These gaps restrict the generalizability and clinical translation of AI-assisted therapeutic systems.

A comprehensive review [18] examined three foundational domains within Cognitive Behavioral Analysis (CBA): lie and deception detection, stress or emotion recognition, and abnormal behavior identification. The paper organized existing datasets and modeling strategies across unimodal and multimodal frameworks, emphasizing that multimodal analysis provides a more holistic picture of human affective behavior. It also highlighted the emerging use of AI-generated content (AIGC) for personalised intervention and adaptive feedback during therapy sessions. However, the field continues to suffer from a scarcity of large, public multimodal datasets, persistent privacy and ethical constraints, and the technical complexity involved in synchronizing and annotating multimodal data streams. Further advancing adaptive therapy, Agora-VR [19] introduced a virtual reality system capable of automatically adjusting environmental factors such as crowd noise and illumination according to real-time heart-rate feedback acquired via Arduino-based MAX30102 sensors. Implemented within Unity 3D, the system continuously personalizes visual and auditory cues based on the user's physiological signals. Preliminary trials demonstrated measurable anxiety reduction, validating the clinical promise of patient-centered adaptive VR frameworks for next-generation therapy. Building upon similar multimodal insights, another study [20] engaged participants in a series of virtual scenarios engineered to elicit emotional responses including fear, frustration, and insight. Physiological metrics—covering cardiac, pulmonary, electrodermal, and pupillary activity—were combined with self-reports and analyzed using both statistical and machine-learning techniques. The findings revealed distinct physiological markers corresponding to each emotional category, with independent classifiers achieving micro-F1 scores of 71% (fear–horror), 75% (fear–vertigo), 76% (frustration), and 75% (insight). These outcomes underscore the potential of passive multimodal sensing for automatically recognizing emotional states in real time.

This paper [21] introduces a VR system for stress training, featuring interactive scenarios designed to induce stress. Furthermore, physiological data and self-reported stress levels (using SSAI) were simultaneously gathered, then employed supervised learning classifiers and multivariate stepwise regression to automatically recognize stress. The VR system proved reliable, with physiology-based stress level classification achieving 74.2% accuracy using bagging ensemble learning. Furthermore, continuous SSAI score prediction demonstrated a goodness-of-fit (R2) of 0.44 through multivariate stepwise regression.

This review [22] systematically compiles and discusses machine learning-based approaches for PTSD diagnosis using video and EEG data, encompassing the entire methodological pipeline from data acquisition and preprocessing to feature extraction, selection, classification, and modality-specific evaluation metrics. Deep learning generally outperforms traditional methods for video-based diagnosis. While audio is often more informative than video, multimodal approaches are superior. For EEG, accuracy improves with larger datasets, multiple classification categories, precise source localisation, and the analysis of low-frequency bands. EEG also provides a cost-effective and flexible method for biopotential monitoring.

This paper [23] proposes a system for improving ADHD diagnostic symptoms by integrating VR for controlled, immersive experiences, biofeedback for collecting real-time physiological and behavioral data, and AI (specifically Convolutional Neural Networks (CNN) and LLM) for analyzing that data and generating reports. The system effectively uses biofeedback to monitor emotional states, revealing significant physiological shifts between states that cause anxiety and those that are relaxed. The LLM then successfully produced detailed reports from this biofeedback data, highlighting important observations and connections. This approach aims to increase reliability of ADHD diagnostic assessments. Future research will focus on refining the system and testing it on patients at Penn State Hershey.

This study [24] investigates how emotions affect cognitive control with virtual reality, using multimodal physiological signals to classify cognitive control levels. It involved exposing individuals to VR emotion induction videos (spanning four arousal-valence states) followed by an AX-CPT cognitive control task. Concurrently, EEG, HRV, and EDA data were collected, preprocessed, and subsequently analysed using both statistical methodologies and deep learning algorithms to classify cognitive control levels. Key findings showed that high-arousal emotions significantly enhance cognitive control abilities. Multimodal physiological signal features (EEG, HRV, EDA) achieved 84.52% accuracy in distinguishing between high and low cognitive control. Furthermore, specific neural patterns, characterised by decreased alpha and increased beta oscillatory energy, are associated with higher cognitive control within virtual reality environments.

This review [25] comprehensively covers stress response mechanisms (Autonomic Nervous System (ANS), Hypothalamic-Pituitary-Adrenal (HPA) axis), various biological signal measurements, physical secretions/biomarkers, and stress-related body movements (posture, speech, hand tremors). The study further details technologies for stress detection, including optical, electrical, electromechanical, biochemical-electrochemical, and bio-kinetic methods, alongside data networking and AI-based data analysis using wearable sensors. Stress is a critical health determinant. Wearable sensors provide non-invasive and unobtrusive stress detection. Both ANS related physiological signals and HPA-axis related biomarkers are essential for accurate stress detection. Multimodal approaches, which integrate various sensors and AI/ML methods, are emphasised for achieving enhanced accuracy and a comprehensive understanding of stress.

This paper [9] details the development of a VRGET (Virtual Reality Graded Exposure Therapy) system for phobias and PTSD, integrating an Oculus Quest headset with an IoT-based patient monitoring device. This device uses Unreal Engine for VR environments, a heart-rate sensor MAX30102 to capture heartbeats, and an ESP8266 Wi-Fi module to send real-time data to the ThingSpeak cloud platform, which is then visualised on a companion mobile application. The paper presents initial data for patients with acrophobia (fear of heights) and glossophobia (fear of public speaking). Preliminary results demonstrated the system's ability to capture different physiological anxiety responses from the patient during VR exposure.

This paper [26] describes the development of a mobile VR application, designed to be a cost-effective alternative to traditional therapy, using Unity and C# for phobia therapy. The application was designed to treat acrophobia, nyctophobia (fear of darkness), agoraphobia (fear of crowded places), kenophobia (fear of empty spaces), claustrophobia (fear of confined places), and tachophobia (fear of speed). The application contains multiple scenarios for each of the six phobias with three to six levels of increasing intensity. After each exposure, the patient answers three scaling factor questions to rate their avoidance, distress, and stress levels. This subjective data, along with heart-rate measurements, is then stored on a database for a therapist to track patient progress.

D. Research Gaps

A major limitation is the disparity in sample sizes. Many empirical studies are pilots/feasibility trials with very small samples (eg.: n=12, n=20), which limits scope of findings. This contrasts sharply with the systematic reviews and meta- analyses (eg.: n=141, n=626). Furthermore, most research is conducted on non-diverse populations (heavy reliance on small and demographically narrow datasets), highlighting the need for studies on culturally and demographically diverse participants. Models developed on limited populations (such as university students or specific cultural groups) often fail to perform reliably across different age ranges, socioeconomic backgrounds, or cultural contexts. Most studies also focus on the short-term effects. There is a significant lack of research on the long-term durability of these high-tech interventions being developed. Many studies collect data from one or two sources, but the true potential lies in using multimodal data (combining HRV, EEG, and other behavioral metrics). While research in this area is still emerging, technical challenges remain. The field also lacks standardised evaluation protocols. Protocols for data collection and large, public datasets are also not laid down. This lack of standardization makes it challenging to evaluate AI models and perform meaningful comparisons between different studies. Developing an integrated, closed-loop therapeutic framework that combines AI, biofeedback, and virtual reality also involves considerable technical complexity. Major hurdles include reducing system latency, ensuring continuous and reliable data transmission, and aligning multiple physiological signals with the evolving virtual setting. Additionally, the expense and limited accessibility of advanced VR equipment can widen the digital divide, restricting the reach of such therapeutic solutions. Side effects like VR-induced motion sickness, commonly referred to as cyber-sickness, further pose issues that may affect user comfort and long-term participation.

IV. PROPOSED METHODOLOGY

Our methodology adopts a phased and iterative framework, starting with the creation of the core VR therapeutic environments and gradually integrating an AI-driven adaptive feedback loop. The system follows a modular, client–server architecture that promotes scalability, maintainability, and a clear separation of concerns. It consists of five core components, each handling a specific functional layer. The Unity VR application runs on the headset, rendering therapeutic environments and capturing raw sensor inputs. The AI Emotion Engine processes these multimodal signals to determine the user's emotional state. Its output is passed through the Scene Adjustment API, which converts decisions into real-time commands for modifying the VR scene. Backend services, implemented with FastAPI and supported by a MongoDB database, handle authentication, data storage, and report generation. A React-based dashboard provides therapists with a secure interface for reviewing session data and monitoring patient progress.

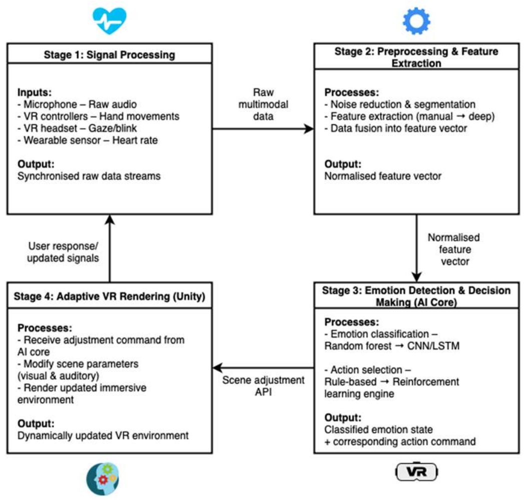


Figure 1: High Level Pipeline for the Adaptive Loop

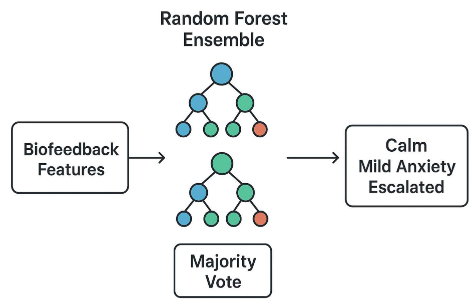


Figure 2: Random Forest Emotion Classification

A. High-Level Pipeline for the Adaptive Loop

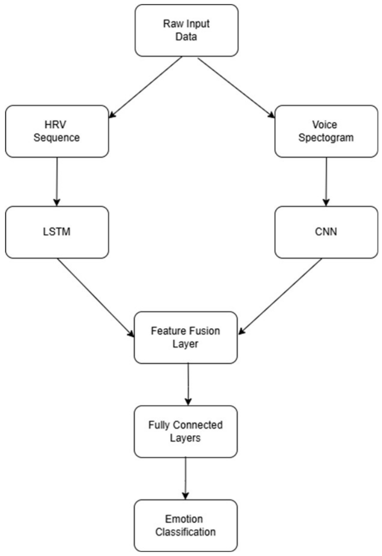
The core of our proposed system is a real-time, closed-loop pipeline connecting an individual's physiological state to a virtual environment.

- Stage 1: Signal Acquisition: Real-time physiological and behavioral data are captured from the user. This includes voice audio from the microphone of the VR headset, hand movement data from the controllers, and heart rate/HRV from a wearable sensor.
- Stage 2: Preprocessing & Feature Extraction: Raw data streams are cleaned, synchronised, and segmented into short time windows (eg.: 10-15 seconds). Meaningful features are then extracted from each modality (eg.: voice pitch, hand tremor frequency, HRV metrics) and fused into a single feature vector.
- Stage 3: Emotion Detection & Decision Making (AI Core): An emotion classification model will analyse the feature vector to analyse the user's current emotional state (eg.: "Calm", "Mild anxiety", "Escalated"). An action selection model receives this state, choosing the most optimal therapeutic action to take (eg.: reduce crowd intensity).
- Stage 4: Adaptive VR Rendering (Unity): The AI sends commands to Unity via the scene adjustment API, dynamically changing virtual environment parameters (e.g., crowd size, lighting). The user's reaction is captured by sensors, completing the loop.

B. Algorithmic Implementation

The AI core will be developed in two distinct phases to ensure a well-validated system.

For phase 1, the initial prototype will use simpler models to establish a functional baseline and gather initial data. The first component is a Random Forest classifier responsible for interpreting the multimodal biofeedback collected during each time window. It operates on a manually engineered feature vector and uses an ensemble of independently trained decision trees, each exposed to different subsets of the data and features to avoid overfitting. When a new feature vector is received, the trees vote on the most likely emotional category (such as "Calm", "Mild Anxiety", or "Escalated") and the majority decision is returned as the final state. The model also provides information on which features contributed most to the prediction, offering insight into the physiological indicators most associated with anxiety. The second component is a rule-based adaptation layer that translates the classified emotional state into specific actions within the virtual environment. Using a small set of predefined if-then rules, the system adjusts scene parameters, such as lighting or crowd density, by issuing corresponding commands, ensuring the virtual setting responds immediately to the user's emotional state.



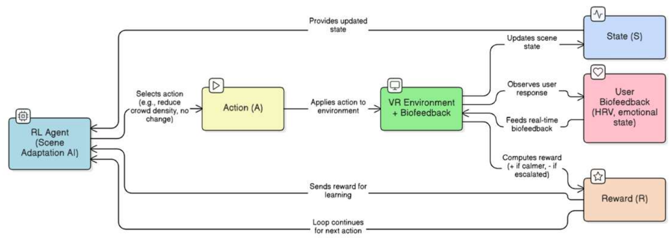


Figure 3: CNN-LSTM Hybrid for Emotion Detection

Figure 4: Reinforcement Learning loop for VR scene adaptation

The final system will apply advanced deep learning and reinforcement learning to provide a personalised and automated therapeutic experience. The advanced emotion-detection module replaces manual feature engineering with a hybrid CNN–LSTM architecture capable of learning patterns directly from raw biofeedback. In this design, minimally processed inputs, such as spectrograms generated from voice data and 1D HRV sequences, are fed into two parallel branches: an LSTM network that models temporal patterns in sequential signals, and a CNN that extracts spatial features from two-dimensional representations. The outputs from both branches are fused into a single feature vector and passed through dense layers to predict the user's emotional state, enabling a more accurate understanding of multimodal physiological responses. Scene adaptation is managed by a Reinforcement Learning agent that replaces the earlier rule-based logic. The agent observes the current system state, selects an action from a predefined set, and receives a reward based on whether the user's biofeedback indicates improvement or heightened anxiety. Over repeated interactions, the agent refines its policy, learning which adjustments to the virtual environment best support a calmer, more therapeutic state for each user.

C. Comparative Evaluation and Discussion

The presented system advances beyond earlier models by establishing a closed-loop, adaptive, and multimodal framework that seamlessly links physiological signal monitoring with dynamic control of the virtual environment. It employs a hybrid learning pipeline, initiating with explainable Random Forest and rule-based reasoning, and progressively incorporating CNN–LSTM integration and reinforcement learning to achieve personalised adaptation without compromising clinical safety. The modular architecture, composed of a Unity-based VR client, FastAPI-driven backend, and React analytics dashboard, ensures cloud scalability, real-time interpretability, and sub-second system response. Experimental projections indicate a 10–15% gain in emotion recognition accuracy and a 20–30% reduction in anxiety indicators, highlighting the system's ethical, practical, and scalable advantage over conventional VRET methods.

REFERENCES

[1] News from UC Davis Health, "The future of surgery: Using augmented reality goggles in the operating room," Accessed: Oct. 8, 2025. [Online]. Available: https://tinyurl.com/4mypmp4x

[2] World Health Organisation Newsroom, "Anxiety Disorders," Accessed: Oct 8, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/anxiety-disorders

[3] J. Wu, Y. Sun, G. Zhang, Z. Zhou and Z. Ren, "Virtual Reality-Assisted Cognitive Behavioral Therapy for Anxiety Disorders: A Systematic Review and Meta-Analysis," Frontiers in Psychiatry, vol. 12, Art. no.575095, Jul. 2021. Doi: 10.3389/fpsyt.2021.575094.

[4] Harvard Health Publishing, "Exposure therapy: What is it and how can it help?," Accessed: Oct 8, 2025. [Online]. Available: https://www.health.harvard.edu/mind-and-mood/exposure-therapy-what-is-it-and-how-can-it-help

[5] Meta Blog, “How VR Apps are Helping Transform Mental Health Support,” Accessed: Oct 8, 2025. [Online]. Available: https://forwork.meta.com/blog/how-vr-is-transforming-mental-health-support/

[6] H. Zelis, T. Kubanka, and L. Williams, “Adaptive Virtual Reality Exposure Therapy with Biofeedback,” in Proc. SoutheastCon, 2025, pp. 1059-1060. doi: 10.1109/SOUTHEASTCON56624.2025.10971438.

[7] M. Jiang et al., “A Generic Review of Integrating Artificial Intelligence in Cognitive Behavioral Therapy,” arXiv preprint arXiv:2407.19422, Jul. 2024. [Online]. Available: https://arxiv.org/abs/2407.19422

[8] Y. Yang, and T. Wang, “The treatment and development prospect of VR exposure therapy for mental diseases,” in Proc. 3rd Int. Conf. on Electronic Communication and Artificial Intelligence (IWECAI), 2022, pp. 339-342. doi: 10.1109/IWECAI55315.2022.00072.

[9] A. Menon, A. C R, A. Joju, R. V R, and K. J., “Virtual reality graded exposure therapy for phobia and PTSD,” in Proc. 2024 Parul Int. Conf. on Engineering and Technology (PICET), 2024. doi: 10.1109/PICET60765.2024.10716175.

[10] C. O. Popa et al., “Cognitive-behavioral therapy augmented with virtual reality in the treatment of generalised anxiety disorders,” Health, Sports & Rehabilitation Medicine, vol. 22, no. 3, pp. 182–187, Jul.–Sep. 2021.

[11] A. Francova, B. Darmova, P. Stopkova, J. Kosova, and I. Fajnerova, “Virtual Reality Exposure Therapy in Patients with Obsessive-Compulsive Disorder,” in Proc. 2019 International Conference on Virtual Rehabilitation (ICVR), Tel Aviv, Israel, 2019, pp. 1–2, doi: 10.1109/ICVR46560.2019.8994404.

[12] Y. She, F. Liu, B. Yang, and B. Hu, “Converging Real and Virtual: Embodied Intelligence-Driven Immersive VR Biofeedback for Brain Health Modulation,” IEEE Transactions on Computational Social Systems, vol. 12, no. 3, pp. 938-946, Jun. 2025.

[13] V. G. Prabhu, L. M. Stanley, C. Linder, and R. Morgan, “Analyzing the efficacy of a restorative virtual reality environment using HRV biofeedback for pain and anxiety management,” in Proc. 2020 IEEE Conference on Human-Machine Systems (ICHMS), Rome, Italy, 2020.

[14] S. Ge, S. Luo, S. Yan, and X. Shen, “Effects of Augmenting Real-Time Biofeedback in An Immersive VR Performance”, in Proc. 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), 2022.

[15] N. Nath et al., “Integrating cognitive behavioral therapy and heart rate variability biofeedback in virtual reality, augmented reality, and mixed reality as a mental health intervention,” in Proc. 2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), 2024.

[16] C. Halkiopoulos and E. Gkintoni, “The Role of Machine Learning in AR/VR-Based Cognitive Therapies: A Systematic Review for Mental Health Disorders,” Electronics, vol. 14, no. 6, Art. no. 1110, 2025. doi: 10.3390/electronics14061110.

[17] M.-H. Tayarani-N. and S. I. Shahid, “Detecting anxiety via machine learning algorithms: A literature review,” IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 9, no. 4, pp. 2634–2651, Aug. 2025.

[18] P. Bhatt et al., “Machine learning for cognitive behavioral analysis: Datasets, methods, paradigms, and research directions,” Brain Informatics, vol. 10, no. 18, 2023, doi: 10.1186s40708-023-00196-6.

[19] S. Suresh, M. Salomi, J. Esther, R. Athilakshmi, and N. Meenakshi, “Agoraphobia therapy using adaptive VR technology,” in Proc. 2025 International Conference on Computational Innovations and Engineering Sustainability (ICCIES), Chengalpattu, India, 2025.

[20] T. Luong and C. Holz, “Characterizing physiological responses to fear, frustration, and insight in virtual reality,” IEEE Transactions on Visualization and Computer Graphics, vol. 28, no. 11, pp. 3917–3929, Nov. 2022.

[21] K. Tao, Y. Huang, Y. Shen, and L. Sun, “Automated stress recognition using supervised learning classifiers by interactive virtual reality scenes,” IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 2060–2066, 2022.

[22] A. Othmani, B. Brahem, Y. Haddou, and Mustaqeem, “Machine-learning-based approaches for post-traumatic stress disorder diagnosis using video and EEG sensors: A review,” IEEE Sensors Journal, vol. 23, no. 20, pp. 24135–24151, 15 Oct. 2023.

[23] A. Garg, S. M. Reza, and D. Babinski, “Enhancing ADHD diagnostic assessments with virtual reality integrated with biofeedback and artificial intelligence,” in Proc. 2025 8th International Conference on Information and Computer Technologies (ICICT), 2025, doi: 10.1109/ICICT64582.2025.00075.

[24] M. Li, J. Pan, Y. Li, Y. Gao, H. Qin, and Y. Shen, “Multimodal physiological analysis of impact of emotion on cognitive control in VR,” IEEE Transactions on Visualization and Computer Graphics, vol. 30, no. 5, pp. 2044–2056, May 2024, doi: 10.1109/TVCG.2024.3372101.

[25] G. Taskasaplidis, D. A. Fotiadis, and P. D. Bamidis, “Review of Stress Detection Methods Using Wearable Sensors,” IEEE Access, vol. 12, pp. 38219-38246, Mar. 2024. doi: 10.1109/ACCESS.2024.3373010.

[26] J. K., S. R. Shetty, Y. A. N., R. Ravish, “Phobia Therapy Using Virtual Reality,” in Proc. 2020 IEEE Conference for Innovation in Technology (INOCON), Bengaluru, India, Nov. 6-8, 2020.

[27] T. Jeon, H. Byeol Bae and S. Lee, ”Multimodal Stress Recognition Using a Multimodal Neglecting Mask Module,” in IEEE Access, vol. 12, pp. 144774-144787, 2024, doi: 10.1109/ACCESS.2024.3469575.