Acta Informatica Pragensia
2026, Volume 15, Issue 2, 345–363
https://doi.org/10.18267/j.aip.305

PRAGUE UNIVERSITY OF ECONOMICS AND BUSINESS

Article Open Access

FearTherapy: Assessing the Impact of Therapeutic Games in Virtual Environments through Physiological State Measurements

Zoltán Balogh [1,2], **Kristián Fodor** [2], **Martin Magdin** [1,3], **Jaroslav Reichel** [1], **József Kopják** [2], **Štefan Koprda** [1], **Martin Polák** [4]

[1] Faculty of Natural Sciences and Informatics, Constantine the Philosopher University Nitra, Nitra, Slovakia
[2] Kándó Kálmán Faculty of Electrical Engineering, Óbuda University, Budapest, Hungary
[3] Faculty of Economics, University of South Bohemia in České Budějovice, České Budějovice, Czech Republic
[4] GAMETHERAPY s.r.o., Nitra, Slovakia

Corresponding author: Zoltán Balogh (zbalogh@ukf.sk)

Editorial Record

First submission received:
September 23, 2025

Revisions received:
October 10, 2025
November 19, 2025
December 22, 2025
January 27, 2026

Accepted for publication:
February 2, 2026

Academic Editor:
Zdenek Smutny
Prague University of Economics and Business, Czech Republic

This article was accepted for publication by the Academic Editor upon evaluation of the reviewers' comments.

How to cite this article:
Balogh, Z., Fodor, K., Magdin, M., Reichel, J., Kopják, J., Koprda, Š., & Polák, M. (2026). FearTherapy: Assessing the Impact of Therapeutic Games in Virtual Environments through Physiological State Measurements. *Acta Informatica Pragensia*, 15(2), 345–363. https://doi.org/10.18267/j.aip.305

Abstract

Background: Virtual reality (VR) integrated with internet of things (IoT) wearable devices offers innovative approaches to mental health interventions by enabling real-time physiological monitoring during immersive therapeutic experiences.

Objective: This study aims to evaluate the effectiveness of VR therapeutic games in identifying and measuring emotional responses through physiological signals (heart rate and galvanic skin response) and to classify these responses using machine learning.

Methods: We conduct experiments with 103 participants (aged 6–57 years) using FearTherapy, a custom VR game featuring interactions with four animals (Hermit, Bee, Wolf, Spider). Physiological data are collected using Samsung Galaxy Watch 5 for heart rate and Arduino Uno with a galvanic skin response (GSR) sensor. After preprocessing, 55 valid sessions remain for analysis. Individual baseline heart rate values are established and random forest classification with grid search optimization and 10-fold cross-validation is performed.

Results: GSR emerges as the most influential feature for classifying emotional states, followed by heart rate difference and baseline reference values. Highest emotional arousal occurs during Spider and Bee interactions. The random forest model achieves 69% cross-validation accuracy and 81% test set accuracy. The model performs well overall but encounters challenges distinguishing Bee from Hermit and Wolf from Spider, suggesting overlapping emotional states.

Conclusion: VR therapeutic games combined with IoT physiological monitoring can effectively measure and classify emotional responses. The findings support development of personalized, emotionally adaptive therapeutic interventions for anxiety and phobia treatment, emphasizing the importance of individualized baseline measurements for accurate emotional state assessment.

Index Terms

Virtual reality; Sensory network; Internet of things; Therapeutic game; Machine learning; Fear measurement; Physiological monitoring; Galvanic skin response; Heart rate.

1 INTRODUCTION

In this article, we propose a methodology for identifying problematic parts in physiological signals – galvanic skin response (GSR), heart rate (HR) and respiratory rate (RR).

The methodology is tested and validated using a system of multimodal sensors integrated into a VR game, which is intended for the treatment of anxiety and fear across multiple VR scenes. The goal of the study is to examine changes in physiological data during VR-based gameplay using a therapeutic game and wearable internet of things (IoT) devices and to explore their potential influence on emotional arousal. In line with the transactional model of stress and coping (Folkman, 2013), VR-based therapeutic games may act as controlled stressors that allow individuals to regulate and adaptively respond to stimuli that would otherwise provoke anxiety or fear responses.

The integration of therapeutic games into virtual environments has received significant attention in recent years as researchers explore their potential to improve mental health and well-being. The immersive nature of virtual environments allows for unique interactions that can enhance therapeutic outcomes, making them a promising tool for innovative mental health interventions. The use of physiological measurements, such as heart rate variability and galvanic skin response, can provide valuable insight into the effectiveness of these therapeutic games by objectively assessing users' emotional and stress responses during gameplay. Integrating these measurements will not only improve our understanding of user experiences, but also inform the development of tailored interventions that can maximize therapeutic benefits. Given the growing interest in these technologies, it is also important to examine the ethical issues associated with their use, as well as ensure protection of users' privacy and confidentiality.

Considering our previous research in the area of emotion recognition in VR applications (Magdin et al., 2021), we have concluded that most automatic emotion recognition systems are based on the analysis of data obtained from users based on static viewing of films, images or listening to various acoustic stimuli. The data thus obtained are then used in computational models without regard to the context of the data, while ignoring the basic fact (the natural behaviour of the observed subject in its natural environment), which can lead to biased interpretation of the data.

To carry out the experiment, we use commercial IoT technologies – a smartwatch for heart rate measurement, a low-cost device for GSR measurement and a VR – Meta Quest 2. Section 2 presents an analysis of the current state regarding the possibility of classifying affective states in virtual reality applications in conjunction with a brief overview of biophysical data. This innovative approach may open the door to new forms of therapy that consider different cultural and personal contexts, thereby strengthening the overall effectiveness of treatment methods. Unlike traditional approaches, VR enables a dynamic interaction with feared stimuli, potentially enhancing emotional learning through mechanisms such as extinction and habituation (Maples-Keller et al., 2017). Many other studies have investigated this topic and they served as a motivation to our implementation. Quintero et al. (2023) conducted a study on 24 participants and measured their reactions in various virtual reality environments. They reached an accuracy of 71% when classifying fear in horror games and 76% when classifying frustration, demonstrating that virtual reality is a useful tool to elicit emotional states, which are very similar in the arousal-valence emotional space (Navarro et al., 2021). The gathered data represent a larger, multimodal dataset, which can be used as the basis of a machine learning solution (Luong and Holz, 2022).

From a theoretical perspective, the present study is grounded in established models of emotional processing and exposure, particularly emotional processing theory and contemporary formulations of virtual reality exposure therapy. Rather than evaluating therapeutic efficacy or clinical outcomes, the primary contribution of this work lies in the methodological exploration of how physiological signals captured during VR exposure tasks differentiate levels of emotional arousal across controlled stimuli. Accordingly, the study is positioned as an exploratory extension of existing VR exposure paradigms, aimed at improving measurement, interpretability and signal sensitivity in immersive environments rather than making claims about treatment effectiveness or emotional change.

The structure of the paper is as follows: Section 2 describes the related work, Section 3 describes the materials and methods used, Section 4 describes the results, while Section 5 discusses the results. The last section concludes.

2 RELATED WORK

2.1 Virtual reality exposure therapy and emotional processing as a theoretical framework

Virtual reality exposure therapy (VRET) builds on classic theories of emotional processing (Foa & Kozak, 1986; Rachman, 1980), causes of fear induction or extinction (Taschereau-Dumouchel et al., 2020) and experiential learning

(Seaman et al., 2017). VRET is an effective form of exposure therapy that provides a safe, controllable and repeatable approach to analysing environmental influences that may trigger anxiety or phobic responses. As the clinical literature shows, VR exposure does not show fundamental methodological differences compared to in vivo therapy but offers greater flexibility in the gradual dosing of stimuli, reduces risk and increases the level of presence (Maples-Keller et al., 2017; Carl et al., 2020; Freeman et al., 2017). The basic principle of VRET is based on emotional processing theory (Foa & Kozak, 1986), which explains the process of gradual adaptation, reduction or even elimination of the body's reactions to the same, frequently repeated stimuli. Many meta-analyses have shown that VRET achieves comparable or even higher efficacy than traditional in vivo therapy in the treatment of various anxiety and phobic disorders (Freeman et al., 2017; Polak et al., 2022).

In the context of VR games, these mechanisms are supported by an increased sense of immersion. This can induce an increase in the intensity or change of emotional reactions, allowing them to be more accurately measured through physiological signals in real time (Zielina et al., 2022; Maples-Keller et al., 2017). Several studies report that VR-based exposure can contribute to the reduction of negative emotional responses, including fear, stress and frustration (Quintero et al., 2023; Navarro et al., 2021).

2.2 Physiological indicators of arousal: GSR, HR and HRV

Galvanic skin response, together with electrodermal activity (GSR/EDA) and heart rate (HR), are among the most reliable physiological indicators. GSR is considered one of the most sensitive ways to determine physiological changes in real time (Braithwaite et al., 2015; Boucsein, 2012) and is widely used to measure fear, stress or cognitive load. Heart rate (HR) and heart rate variability (HRV), most often obtained using photoplethysmography (PPG) or electrocardiography (ECG), provide additional information about how strongly the body can respond to stress or emotional stimuli. They therefore express the extent to which the organism is in a state of tension or, conversely, at rest (Sharma et al., 2019; Lee & Yoo, 2020; Laborde et al., 2017; Mather, 2018).

An important part of modern procedures is the use of individual baseline values. Each person has a naturally different level of HR and GSR. Without taking these differences into account, changes during the experiment could be evaluated incorrectly (Joshi et al., 2022; Laborde et al., 2017). Therefore, several authors recommend first determining the baseline measurement value, removing slow long-term fluctuations in GSR, purging the signal of noise and adjusting the recording so that it does not contain errors caused by movement or brief loss of sensor contact with the skin (Choudhury et al., 2022; Boucsein, 2012; Koprda et al., 2024).

2.3 Multimodal affect recognition and VR

Multimodal measurement of physiological responses to external stimuli integrating GSR, HR/HRV, respiratory activity or EEG has become a key research direction in the field of emotions in recent years. Many studies confirm that the combination of physiological signals in environments with high sensory stimulation, such as VR, can improve the accuracy of emotional state classification (Luong & Holz, 2022; Sorinas et al., 2020). Existing multimodal datasets (e.g., DEAP, MAHNOB-HCI, AMIGOS) show that the accuracy of emotion classification typically ranges between 60–80% when GSR and HR are present as key factors for data acquisition (Koelstra et al., 2012; Navarro et al., 2021; Yang, 2024; Al-Azani, 2025). These results are also consistent with our findings, where GSR emerges as the dominant variable in distinguishing the intensity of the emotional response. Moreover, VR provides the possibility to control the sequence of stimuli, intensity of stimulation and duration of exposure, which distinguishes it from traditional passive observation methods (Magdin et al., 2021; Paletta et al., 2020). Studies focused on measuring fear in gaming VR environments (Civitello et al., 2014; Quintero et al., 2023; Luong & Holz, 2022) demonstrate a high correlation between GSR/HR and subjectively reported fear or arousal, which supports the use of physiological signals in therapy and diagnosis.

2.4 Wearable devices and measurement reliability

Thanks to IoT-based technologies and wearables, we can now measure physiological responses not only in the laboratory but also in everyday life situations. Modern smartwatches use PPG sensors, which achieve an accuracy of around 90–95% in measuring HR under resting conditions (Charlton & Marozas, 2022); their application is not only in studies focused on researching emotional states but has also been confirmed in other works focused on, for example, stress detection (Shikha et al., 2022). Although these devices can sometimes capture distorted or unreliable data, proper signal processing and the use of multiple types of measurements significantly reduces the risk of incorrect evaluation of results (Francisti et al., 2020; Hiwaki, 2024). In a virtual reality environment, wearables are extremely useful because they do not restrict the user's movement and can continuously record data throughout the entire interaction with the VR scene.

2.5 Machine learning and deep learning approaches

Traditional algorithms, such as random forest, support vector machines or gradient boosting, are often used for emotional state classification because they tolerate dirty or partially missing data well and offer interpretability (Joshi et al., 2022; Quintero et al., 2023). However, in the last decade, there has been a significant increase in interest in deep learning models, especially convolutional neural network (CNN), long short-term memory (LSTM) and hybrid CNN-LSTM models, which can model the temporal dynamics of GSR/HR signals and extract subtle patterns in arrhythmic or noisy recordings (Faust et al., 2018; Vidyasagar et al., 2024; Liao et al., 2025; Alqudah, 2025). These methods often achieve higher accuracy but require larger datasets and better control over the quality of the measurements.

While deep learning architectures have demonstrated strong performance in large-scale affective computing datasets, their advantages are closely tied to high sample counts, dense temporal labelling and controlled signal quality. In the present study, the dataset size, heterogeneous participant population and presence of signal artifacts favour the use of classical ensemble methods over deep neural networks. Random forest models offer a favourable balance between robustness, performance and interpretability, particularly in exploratory settings where understanding feature contributions (e.g., GSR dominance) is as important as classification accuracy.

Moreover, deep learning approaches introduce additional methodological risks in small to medium-scale physiological datasets, including overfitting, reduced transparency and sensitivity to preprocessing choices. Given that the primary aim of this work is not maximal predictive performance but the characterization and differentiation of physiological responses across VR exposure scenes, an interpretable model is methodologically preferable. The selection of random forest therefore reflects a deliberate design choice aligned with the exploratory objectives of the study, rather than a limitation in methodological scope.

This review provides the methodological and empirical background for the development process described in Section 3 below. The study therefore uses random forest as a robust and well-interpretable model adapted to the size of the dataset and the heterogeneity of physiological signals. The results of the study suggest potential for future deployment of sequential deep learning architectures within VRET-based therapies.

3 RESEARCH METHODS

The integration of VR and IoT devices into emotion research raises new ethical and privacy challenges. Physiological signals, such as heart rate or skin conductance, are sensitive biometric data that can indirectly reveal the health status, stress level or emotional reactivity of participants (Afzal et al., 2024; Alneyadi et al., 2021). Therefore, it is essential to adhere to strict privacy principles, including informed consent, anonymization and secure data transmission (European Commission, 2018).

In addition to data protection from a privacy perspective, the psychological safety of participants is also crucial. Exposure experiments can cause transient negative emotions or discomfort. Current recommendations therefore emphasize the need for the presence of a qualified professional, the possibility of ending the experiment at any time and the provision of post-session consultations (Bell et al., 2022; Ling Tan et al., 2025). These principles were also implemented in our research, ensuring that technological innovation remains in line with ethical and psychological standards of human research.

Beyond compliance with formal ethical standards, physiological monitoring in immersive virtual environments raises broader interpretive and societal considerations. Signals such as GSR and heart rate are often perceived as objective indicators of internal emotional states; however, they remain indirect proxies that are sensitive to context, individual variability and situational factors. Over-interpretation of such signals risks unintended emotional labelling or attribution, particularly when physiological arousal is equated with specific affective states such as fear or anxiety.

These concerns are amplified when studies involve minors. Children and adolescents may have limited capacity to contextualize or critically interpret biometric feedback and may internalize physiological responses as indicators of personal deficiency or pathology. For this reason, the present study avoids real-time evaluative feedback, emphasizes neutral framing of physiological data and ensures post-session debriefing by a qualified professional. Importantly, physiological data are not used for individual diagnosis, profiling or adaptive difficulty modulation during gameplay.

From a privacy perspective, physiological signals constitute sensitive biometric data that can support secondary inferences beyond the original research intent, such as stress susceptibility or emotional reactivity. Although all data were anonymized and stored securely, future deployments of similar systems (especially in clinical or educational settings) should carefully consider data minimization, strict purpose limitation and explicit governance over long-term storage and reuse. Ethical acceptability should therefore be evaluated not only at the level of data protection compliance, but also in terms of how biometric information is interpreted, communicated and acted upon.

Authors confirm that all the methods were carried out in accordance with relevant guidelines and regulations. This study was approved by the Ethics Commission of Constantine the Philosopher University in Nitra, number UKF/1350/2024/191013:002. The authors confirm that all the subjects had provided appropriate informed consent, and details on how this was obtained are detailed in the manuscript.

3.1 Research setup and hypothesis formulation

The measurement of physiological functions was carried out through a sensor network that included non-invasive and commonly used IoT devices, such as bracelets or other wearables. These devices allow for long-term wearing without significant impact on daily or occupational activities. Their ability to measure and send data via wireless communication protocols such as Bluetooth or Wi-Fi ensures interoperability between different types of devices and their integration with a central control device, such as a computer. The acquisition of physiological data during the use of a therapeutic game via VR was carried out within the framework of the VEGA project. We set the following scientific hypotheses:

H1: *Using commonly available IoT devices, it is possible to non-invasively acquire physiological data of heart rate (HR) and galvanic skin resistance (GSR) so that these data are accurate and reliable.*

H2: *Using the acquired data, we will be able to identify when the respondent is in a state of calm/excitement during a therapeutic game created for VR.*

3.2 Therapeutic game in VR

To induce a change in calm and excitement in participants, we used a virtual reality game called “FearTherapy” (Gametherapy, 2023). The interactive game “FearTherapy” had four scenes, in which four animals appeared: a hermit crab, a bee, a wolf and a spider. It combined innovative game mechanics with scientific knowledge about various phobias, creating an environment that was both immersive and educational. Using specially designed elements that enhanced the feeling of playing the game, players had to solve various puzzles to get to know and interact with these animals. Individual animals also emitted sounds and visual signals.

3.3 Wearable IoT devices

We selected two IoT devices to measure physiological data, namely the Arduino Uno R2 Wi-Fi and the Samsung Galaxy Watch 5. Recent studies demonstrate that IoT devices combined with statistical methods can be used to substantiate and interpret collected physiological time series (Martinez Parrales and Téllez-Anguiano, 2022).

We used these devices to measure two physiological functions: (1) heart rate and (2) galvanic skin response. By carefully analysing the data obtained from the measurements of the above features, an accurate picture of the user's stress level was obtained, allowing the monitoring and interpretation of the user's responses to stimuli during virtual reality experience. This allowed the identification of moments of increased stress or relaxation. From a physiological and signal-processing perspective, the smartwatch estimates heart rate using photoplethysmography (PPG). PPG relies on an optical sensor (LED + photodiode) that detects pulsatile changes in blood volume in superficial vessels, typically at the wrist. The pulse rate derived from PPG can be used as a surrogate for heart rate, and pulse-rate variability can approximate heart rate variability (HRV) under suitable conditions. However, unlike clinical ECG-based monitors, PPG signals are more susceptible to motion artifacts, changes in sensor–skin contact, ambient light and local perfusion (Kwang Bok & Jae Baek, 2023).

A recent preliminary validation of the Samsung Galaxy Watch 5 against a research-grade ECG (Shimmer3) found that the watch provides an unbiased and high-quality estimate of mean heart rate at rest, with good agreement in several HRV features under stationary conditions (Rho et al., 2024). These findings are consistent with broader literature on consumer PPG wearables and support the use of the Samsung Galaxy Watch 5 for monitoring average heart rate trends, while cautioning against over-interpretation of fine-grained variability measures during movement-rich tasks.

Integrating these IoT devices with VR systems requires seamless connectivity and data synchronization. Bluetooth and Wi-Fi technologies are commonly used for wireless data transfer, which can be integrated into various applications, ensuring accurate capture of physiological data and its transmission to the VR system in real time.

3.4 VR therapeutic game scenario

At the beginning of the game, players found themselves in a room, specifically in a library, where individual animals gradually began to appear. At each launch, there was also an introductory instruction that lasted 66 seconds, which explained to the game participants what they could expect in the interactive therapeutic game. During this time, the participants were in a standing position and were listening to the instructions with both controllers in the hands and the devices were already measuring their physiological functions. The game was controlled very simply, just by looking at given things, such as an object which they wanted to move, lift or interact with. At the end of the game, there was an opportunity to go through breathing exercises, which we did not implement with the participants.

The goal of the game was not only to overcome the fear of various animals, but also to master all the tasks in a safe environment. Thanks to the various animal sounds and the presence of a narrator, the player could perceive the dynamics of the environment and the seriousness of the situation, while safely moving forward at their own pace and overcoming their fear.

Figure 1. *A scene in FearTherapy VR game.*

FearTherapy (Figure 1) thus uniquely combines the potential of virtual reality with therapeutic goals, providing an innovative and effective tool for treating fears of individual animals such as hermit crabs, bees, wolves and spiders. When designing the game, special attention was paid to how the gameplay elements would serve therapeutic purposes while providing an experience that players would understand and remember.

3.5 Demographics and procedure

All participants experienced the four scenes in a fixed order (Hermit → Bee → Wolf → Spider). This fixed within-subject sequence could induce order and carryover effects (e.g., novelty, anticipatory arousal, fatigue, habituation). Consequently, differences between animals may have been partially confounded by their position in the sequence. To mitigate this retrospectively, we baseline-corrected HR/GSR within each participant. This step did not eliminate the limitation but bound its likely impact.

In this experiment, we tested and collected physiological data from 103 respondents (aged 6–57 years), representing a diverse sample across children, adolescents and adults. Participants were elementary school, high school and university students as well as teachers. Inclusion criteria required that all individuals had normal or corrected-to-normal vision and no known neurological or psychiatric conditions. Informed consent was obtained from adult participants or legal guardians of minors. All the collected data were fully anonymized to ensure participant privacy. During the experiment, participants engaged in a therapeutic virtual reality (VR) game where they were exposed to four animal stimuli (hermit, bee, wolf and spider) in a fixed order across all sessions.

3.6 Sensory network design

The goal of creating a sensor network was to design and create a system that was able to collect physiological data through a smart device. Today's smart devices, such as smartwatches, bracelets and rings, can collect physiological data about the user in real time thanks to biometric sensors. The collected data were processed and then sent to a server. The data were sent to the server via Wi-Fi or 4G/5G data connection. The server provided the means to register participants, assign devices and configure application parameters for data collection. The received data were stored in a central database. On site, the purpose of the research was to capture the measured physiological data of the user. Subsequently, these data were analysed and visualized in a web application. Through an application programming interface (API), we obtained data from the server. The tools to achieve this goal were a Samsung Galaxy Watch 5 smartwatch and an Arduino microcontroller with a GSR sensor. The development environment included Android Studio for smartwatches on the Wear OS platform (Samsung Galaxy Watch 5) and for smartphones with the Android operating system, as well as Arduino IDE for microcontrollers (Arduino UNO II Wi-Fi – GSR).

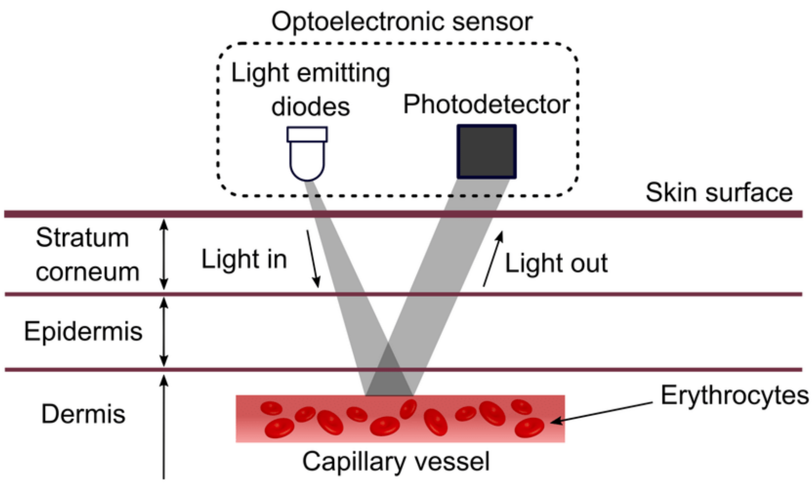

Figure 2. *Principle of operation of PPG sensors.*

3.6.1 Smartwatches

A smartwatch is a type of digital or electronic watch that can perform advanced functions beyond simply displaying the time. These functions may include displaying notifications from a connected smartphone, tracking fitness or exercise or providing additional features such as mobile payments or voice assistants. Smartwatches are often designed to be worn on the wrist and may have touchscreens or physical buttons for interacting with the watch functions (Davie and Hilber, 2016). Several IoT devices are available for non-invasive heart rate monitoring. Wearables such as smartwatches and fitness trackers use photoplethysmography (PPG) sensors to measure changes in blood volume in the skin, which provides an estimate of heart rate (Figure 2). Brands such as Fossil, Samsung and Withings offer products with built-in heart rate monitors that are user-friendly and accurate (Charlton and Marozas, 2022).

3.6.2 Galvanic skin response monitoring

Galvanic skin response (GSR) measures the electrical conductivity of the skin, which changes in response to sweat gland activity. Changes in GSR are confirmed by feedback on the subject's experienced emotions, such as stress or excitement (Christian et al., 2009). Wearable devices such as the Empatica E4 bracelet and the Grove GSR sensor for Arduino allow monitoring galvanic skin response. The Grove GSR sensor connects to two fingers on one hand and is suitable for creating projects related to emotions. It can monitor sleep quality or the impact of gaming. These devices are equipped with sensors that ensure accurate and reliable data collection.

Figure 3. *GSR sensor module.*

By monitoring GSR (Figure 3) in conjunction with VR gameplay, developers and therapists can gain insight into a player's emotional state and tailor the gaming experience or therapeutic intervention accordingly. GSR can respond to changes in stress caused by game difficulty and is significantly correlated with negative gaming events such as frustration rather than positive events such as success.

3.6.3 Virtual Reality

Virtual reality (VR) is at the forefront of technological innovation in gaming, offering gamers a unique experience. By utilizing advanced computer graphics, audio and interactive elements, it transports players into virtual worlds, creating a sense of presence and engagement that goes beyond traditional gaming media. Understanding players' physiological responses during VR gameplay is essential for several reasons. These data allow game developers to refine the gaming experience and ensure that it elicits intense emotional and physiological responses. Consider a horror VR game that aims to increase the player's heart rate and induce fear; monitoring physiological responses helps developers assess whether these goals have been achieved. Virtual reality has become an effective tool in the treatment of various phobias because it provides a controlled and safe environment for exposure therapy. According to Maples-Keller et al. (2017), virtual reality provides tools to enhance traditional exposure methods in the treatment of phobias. Since exposure therapy in virtual reality has not shown significant methodological differences compared to "in vivo" therapy, VR has great potential to be more effective and at the same time easier and cheaper to implement and conduct such sessions compared to traditional methods, which require more time, preparation and other resources.

Figure 4 visualizes the data collection process. The data from the IoT devices (GSR and HR) were sent to the server via a POST API request. The therapeutic game also sent details from the game to the server, mainly the name of the animal that the user interacted with. The server processed the data and stored it in the relevant database tables. The therapeutic game displayed the GSR and HR rates in the top right corner of the screen; therefore, the server sent the most recent GSR and HR values to the therapeutic game via a POST API request. After the experiment, the data in the database were processed and synchronized by using Structured Query Language (SQL) and exported into a comma-separated values (CSV) file (Figure 5). The data also contained metadata regarding the sessions – *session_se* (start of the session) and *session_clear* (end of the session).

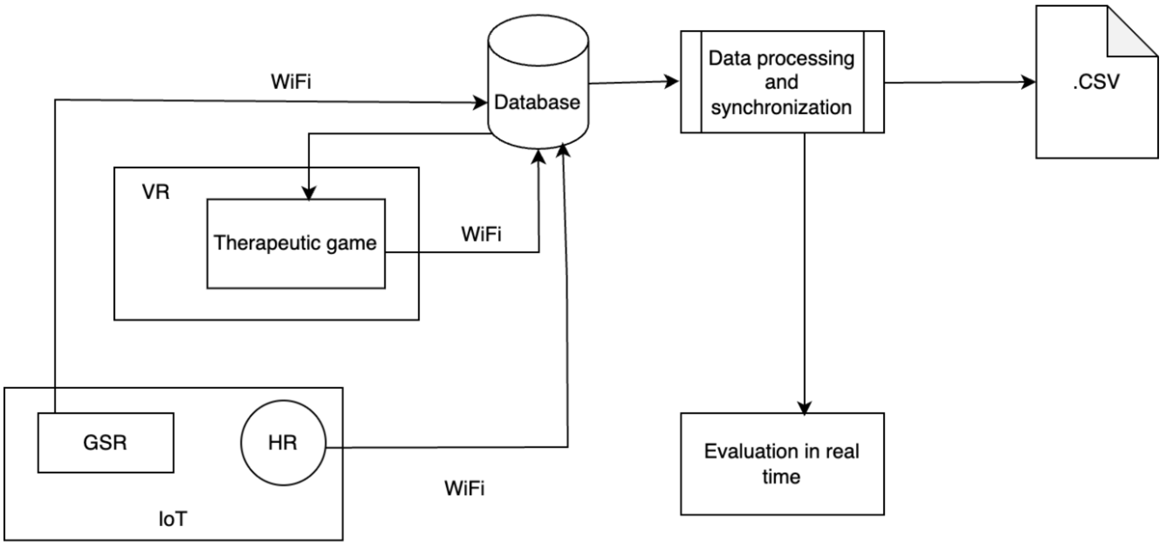

Figure 4. *Data engagement and acquisition model.*

id	animal	date_time	heart_rate	gsr	delay	change	start	session	end	session_se	session_clear	calculated	hr_ref	hr_dif
0	Other	2023-10-05 08:41:48	90.0	125.0	1.0	0.0	1.0	2.0	0.0	2.0	2	0	88.535714	1.464286
1	Other	2023-10-05 08:41:49	89.0	125.0	1.0	0.0	0.0	2.0	0.0	2.0	2	0	88.535714	0.464286
2	Other	2023-10-05 08:41:50	87.0	126.0	1.0	0.0	0.0	2.0	0.0	2.0	2	0	88.535714	-1.535714
3	Other	2023-10-05 08:41:51	85.0	124.5	1.0	0.0	0.0	2.0	0.0	2.0	2	0	88.535714	-3.535714
4	Other	2023-10-05 08:41:52	84.0	124.0	1.0	0.0	0.0	2.0	0.0	2.0	2	0	88.535714	-4.535714

Figure 5. *Sample data from CSV file.*

4 SOLUTIONS AND RESULTS

Participants wore **wearable devices (smartwatches)** that continuously recorded their **heart rate (HR)** and **galvanic skin response (GSR)** in real time. The primary objective was to analyse physiological variations in response to each stimulus and evaluate the level of emotional arousal induced by the VR experience. In our previous work, we have already analysed the accuracy of the sensors used (Francisti et al., 2023; Koprda et al., 2024)

Certain technical factors had to be addressed during data acquisition and processing to ensure the reliability of the dataset. These included **variability in signal quality** caused by **incorrect positioning of the trimmer**, which led to occasional sensor signal losses. This issue was analysed in detail and discussed in Koprda et al. (2024). To optimize data quality, systematic **data cleaning and filtering procedures** were applied before the analysis.

4.1 Data processing and cleaning

Several methodological challenges were addressed to ensure consistency and reliability of the dataset for further analysis:

- Interpolation of missing values – In cases where data gaps were detected, missing values were interpolated based on the transition between the last recorded data point before the gap and the first available data point afterwards. For instance, if HR was recorded as 88 bpm at second 1 and 90 bpm at second 3, but second 2 was missing, its value was estimated as the mean (89 bpm). This approach minimized the impact of short-term sensor failures on the dataset.
- Handling duplicate records – If multiple HR values were recorded within the same second, their values were averaged to maintain consistency.
- Identification and removal of invalid records – In some instances, sensors (particularly the GSR sensor) continued recording data beyond the active experimental phases or during non-relevant periods. Since the VR application transmitted real-time information regarding the displayed stimulus, these extraneous records were identified and removed to prevent distortions in the dataset.

- Ensuring measurement completeness – To be considered valid, each measurement session had to include all four animal stimuli in the predefined sequence. Sessions with missing stimuli were excluded from the analysis.

After applying these preprocessing steps, a total of 55 valid measurement sessions remained for further analysis.

4.2 Calculation of reference HR and HR variability

To quantify changes in heart rate during the experiment, an individual reference HR (*hr_ref*) was established for each participant. The *hr_ref* value was computed as the average HR recorded during the first 66 seconds of the session, a phase where no stimuli were presented and physiological responses were expected to be stable.

The momentary HR variability was then computed as follows: $hr\_dif = heart\_rate - hr\_ref$

This ***hr_dif*** variable represented the participant's **heart rate deviation** relative to their baseline and served as a **key metric** for assessing physiological responses to the VR stimuli.

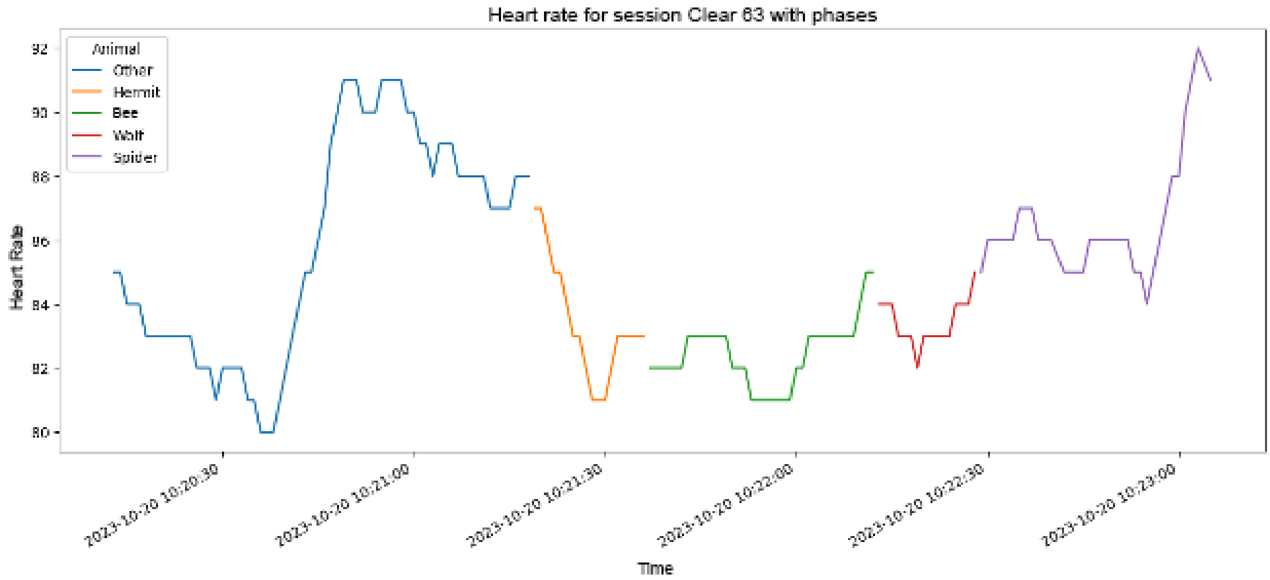

Figure 6. *Heart rate visualization for session with ID 63.*

Figure 6 shows the times-series analysis of the heart rate during a single session separated by the phases of the VR game. In this session, the user's heart rate started to rise a few seconds after the introduction had started, possibly after the narrator in the game had said that the user would see some animals that are considered scary. Later, we only saw a rise in heart rate after the spider had appeared on screen. For the other three animals, we considered the user calm; however, the user did not reach the lowest measured heart rate during the baseline measurement (80 bpm) while any of the animals were on screen. The trends in other sessions are similar. We saw noticeable spikes in heart rate during interactions with the spider, indicating heightened emotional or physical responses. The physiological differentiation across stimuli is in line with psychological findings that link arousal to threat perception and attentional bias, especially in the context of specific phobias (e.g., arachnophobia), as discussed in connection with fear-learning paradigms (Foa and Kozak, 1986; Liu et al., 2013; Taschereau-Dumouchel et al., 2020). More stable heart rate patterns were seen during interactions with the hermit. GSR spikes coincide with heart rate spikes for the spider, reinforcing the hypothesis of increased arousal or stress. For the hermit, the GSR remained relatively steady, suggesting a more relaxed state. Moreover, the heightened response to the spider may have reflected both evolutionary preparedness and cognitive schema activation, mechanisms described in contemporary fear-conditioning models.

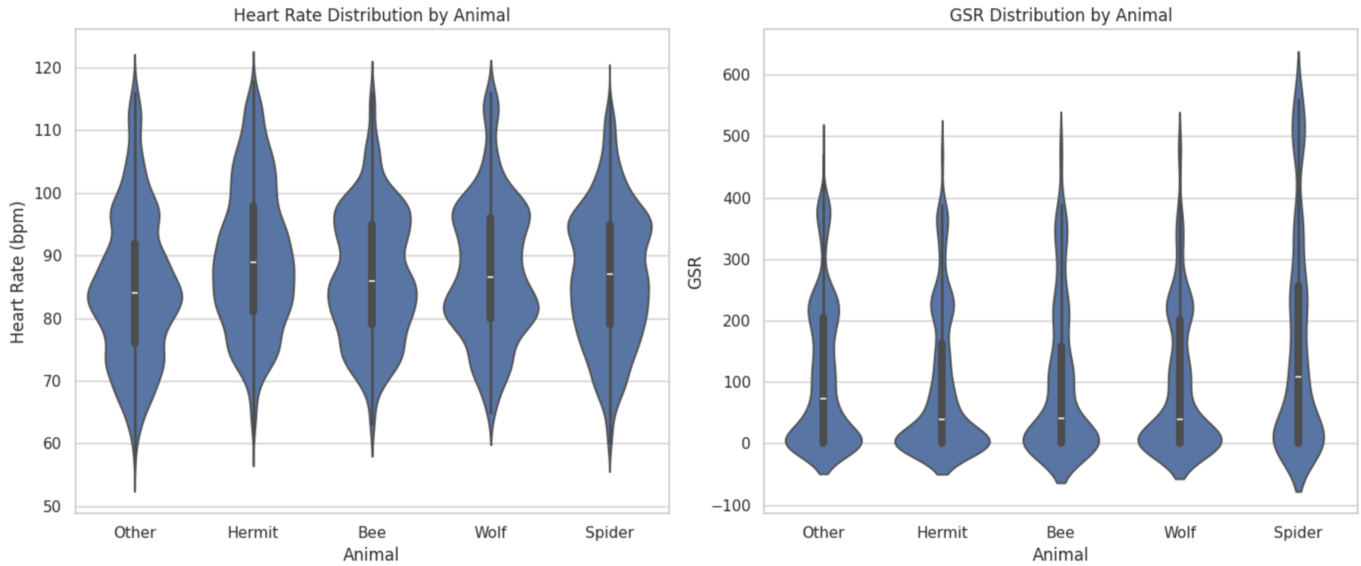

Figure 7. *Heart rate and GSR distributions per animal.*

Figure 7 visualizes the distribution of the measured bio signals (HR and GSR) across the entire dataset. The heart rate varied noticeably across different animal interactions. The highest variability in heart rate was observed with the spider, possibly indicating a stronger emotional response. Interactions with the hermit showed comparatively lower heart rates and less variability. GSR, which measures emotional arousal, was notably higher and more variable for interactions with the spider. Lower GSR values were observed for the hermit, suggesting a calmer response.

As mentioned, from the records before the first animal (hermit), we created the baseline (reference) heart rate (*hr_ref*). The heart rate for each user was also averaged during each animal. These two parameters were used to calculate the deviation from the reference heart rate.

Table 1. *Heart rate difference for each animal.*

Animal	***hr_dif***
Hermit	0.9756
Other	0.0000
Spider	1.3859
Bee	1.6346
Wolf	0.8151

Importantly, all the participants experienced the VR scenes in a fixed sequence (Hermit → Bee → Wolf → Spider). As a result, the observed physiological differences may reflect not only stimulus-specific effects, but also order-related influences such as novelty, anticipatory arousal, habituation or cumulative fatigue. While baseline normalization was applied to reduce inter-individual variability and partially bound these effects, it does not fully disentangle stimulus identity from presentation order. Consequently, the reported differences should be interpreted as scene-dependent physiological responses under a fixed exposure sequence rather than as definitive evidence of animal-specific fear hierarchies.

Table 1 displays the heart rate difference for each animal. Based on this table, we see that the users' heart rate increased moderately while interacting with the spider and the bee, while decreasing slightly when interacting with the hermit and the wolf. The wolf and the hermit showed moderate variability, but lower means compared to the spider and the bee.

To examine whether the physiological differences between the four VR animal scenes were statistically meaningful, we conducted repeated-measures ANOVA analyses on the baseline-corrected heart rate (*hr_mean*) and galvanic skin response (*gsr_mean*). Each participant contributed one mean value per scene, enabling a within-subject comparison across conditions.

4.3 Heart rate (HR)

The repeated-measures ANOVA revealed that the mean heart rate did not differ significantly across the four animals: $F(4, 216) = 1.92$, $p = 0.108$, $\eta^2 = 0.034$. Bonferroni-corrected post-hoc comparisons showed no significant pairwise differences (all $p > 0.26$). Although the HR tended to increase during the bee and spider scenes (Table 1), the effect size was small, indicating that the heart rate changes were modest and showed considerable inter-individual variability. These findings suggest that HR alone was not a strong discriminator of emotional arousal across the VR stimuli.

The small effect size ($\eta^2 = 0.034$) indicates that the heart rate differences across scenes accounted for only a minor proportion of the within-subject variance, suggesting limited practical relevance despite observable trends.

4.4 Galvanic skin response (GSR)

In contrast, the GSR differed significantly between the animals, $F(4, 192) = 18.20$, $p < 0.0001$, $\eta^2 = 0.275$, indicating a large effect size. Post-hoc tests showed that the spider scene elicited reliably higher GSR than the bee, hermit and other (all corrected $p \leq 0.015$), reflecting substantially elevated sympathetic activation. No significant differences were found between the hermit, wolf or bee after correction. These results confirmed that GSR was highly sensitive to the emotional intensity of the scenes, particularly for fear-relevant stimuli such as spiders.

Taken together, the statistical analyses highlight an important distinction between significance and effect magnitude. While the heart rate differences did not reach statistical significance and exhibited small effect sizes, the galvanic skin response demonstrated both statistical robustness and large practical effects. This contrast reinforces the interpretation of GSR as a more sensitive indicator of arousal in short-term VR exposure contexts, while heart rate appears more susceptible to inter-individual variability and contextual modulation.

Next, we examined whether the measured physiological signals and engineered variables (e.g., session) could be used to classify the animal with which the user interacted. To prepare the data for machine learning, we first carried out a feature importance analysis to find out which of the features were the most important.

According to Figure 8, GSR was the most important feature for classifying animal interactions. This suggests that emotional arousal varied significantly depending on the animal. Heart rate difference was the second most influential feature, indicating a meaningful change in the users' HR levels. Heart rate and baseline reference heart rate were also identified as significant features.

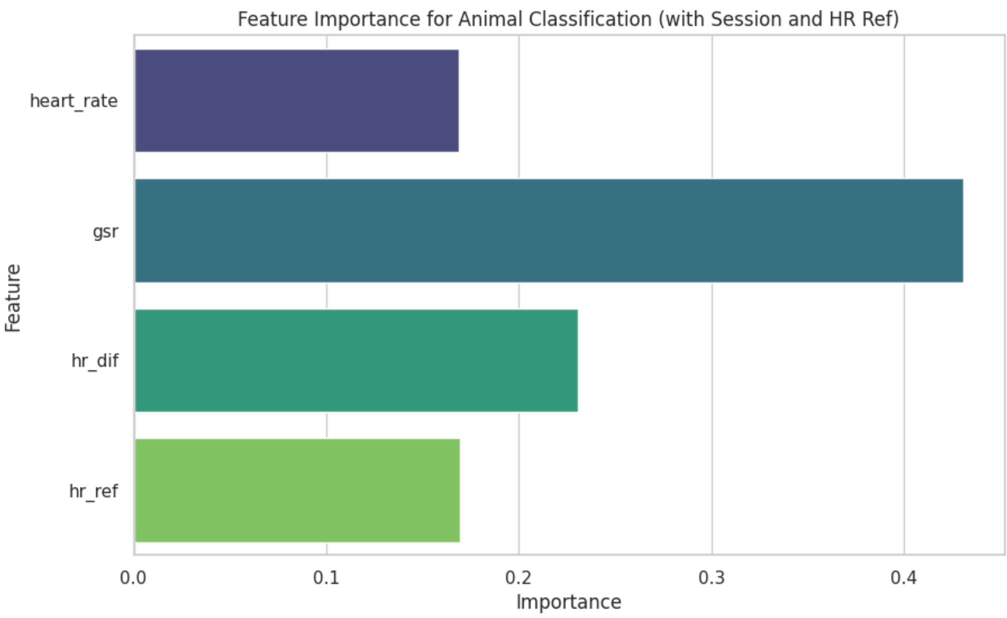

Figure 8. *Feature importance for animal classification.*

Before starting the machine learning, we encoded the categorical and target variables, standardized the features and split the data into training and test sets (0.7 to 0.3), while also using random state to shuffle the data. The classifier we used was random forest (RF). Random forest is well-suited for capturing complex non-linear interactions between physiological features (heart rate, GSR, *hr_dif*) and categorical variables (session, animal). RF generally provides high accuracy and robust performance compared to simpler models such as decision trees or logistic regression. One of the key advantages of random forest is its ability to estimate feature importance, helping us understand which physiological features (e.g., GSR, *heart_rate, hr_dif*) are most influential in predicting animal interactions. This is particularly useful for interpreting emotional responses and understanding the impact of personalized baselines (e.g., *hr_ref*).

We used grid search to find the best combination of hyperparameters (*n_estimators, max_depth, min_samples_split, min_samples_leaf, bootstrap*) for the random forest together with 10-fold cross-validation. Best cross-validation accuracy was reported at 69%.

5 DISCUSSION

The findings presented in this section should be interpreted within an exploratory and measurement-focused framework. While differences in physiological arousal were observed across the VR scenes, these responses reflect short-term autonomic activation rather than sustained emotional regulation or therapeutic change. Elevated GSR or heart rate deviations indicate momentary arousal associated with stimulus exposure, but they do not in themselves constitute evidence of fear reduction, habituation or clinical improvement. Consequently, the results are best understood as demonstrating the sensitivity of physiological monitoring within VR exposure contexts, rather than as direct indicators of therapeutic impact.

Table 2. *Classification report for random forest.*

Animal	Precision	Recall	F1-score	Support
Bee	0.82	0.65	0.72	333
Hermit	0.80	0.83	0.82	903
Spider	0.84	0.90	0.87	721
Wolf	0.81	0.83	0.82	413
Other	0.79	0.67	0.72	1009
Accuracy	-	-	0.81	3379
Macro average	0.81	0.78	0.79	3379
Weighted average	0.81	0.81	0.81	3379

Table 2 shows the classification report of the random forest classifier for all the classes. Classes included the 4 animals and the “other” class, which represented the introduction section of the game, while no animal was visible on screen. The reported accuracy (69%) was the average accuracy across all 10 folds during cross-validation. It represented the generalization performance of the model on unseen data. It considered the mean accuracy of all folds, which might be lower if certain folds contained more challenging data points (e.g., more ambiguous samples or imbalanced classes).

The classification report yielded an 81% accuracy. This was calculated on the test set after selecting the best model. The test set may have had a slightly different distribution than the validation sets used during cross-validation. The model might have generalized better on the test set, leading to a higher accuracy compared to the cross-validation average. The classification report also included macro average, which is the arithmetic mean of precision, recall and F1-score for each class. This was useful in our case as all the classes were equally important. Another variable in the classification report was the weighted average, which weighted the score of each class by its support (number of true instances).

The best hyperparameters were: *{'bootstrap': True, 'max_depth': 20, 'min_samples_leaf': 1, 'min_samples_split': 5, 'n_estimators': 300}*. With *bootstrap* sampling enabled (set to True), each decision tree was trained on a random subset of the data with replacement. It introduced variability, helping the model generalize better and reduce overfitting. Random subsets allowed trees to learn different patterns, improving overall accuracy. It is particularly useful when there are some noisy or outlier data points.

Max_depth was set to 20. Depth 20 was sufficient to capture complex interactions between physiological features (e.g., heart rate, GSR) and categorical variables (e.g., session). Too deep (e.g., None) would have led to overfitting by memorizing the training data. Too shallow would have led to underfitting, failing to capture important patterns. In this context, a depth of 20 allowed the model to capture non-linear interactions without overfitting.

Min_samples_leaf was set to 1. This allowed the trees to grow fully and capture detailed interactions between features. It is beneficial in datasets with complex patterns or when some classes have fewer samples (e.g., bee and wolf). This setting maximized model flexibility, ensuring that the model did not prematurely stop splitting when important patterns were present.

The best configuration for *min_samples_split* was 5. It prevented the model from creating overly complex branches, which would have led to overfitting. A split requirement of 5 ensured that nodes were split only when there was a meaningful pattern to learn. Too low (e.g., 2) would have led to overfitting, while too high would have resulted in underfitting.

300 *n-estimators* were found as the best configuration. More trees provided better averaging, reducing variance and increasing stability. This was particularly beneficial when dealing with noisy physiological data, ensuring robust predictions. 300 trees were enough to achieve a good balance between accuracy and computation time. More trees (e.g., 500) would have yielded diminishing returns in accuracy but would have significantly increased computation time. Fewer trees (e.g., 100) would have increased variance and reduced accuracy.

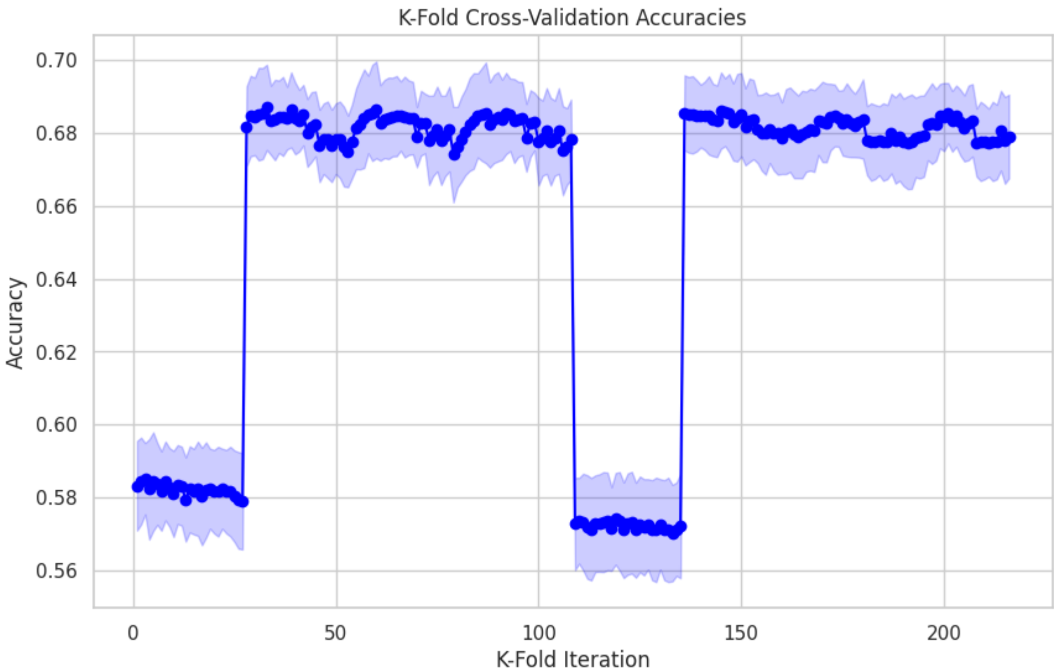

Figure 9. *K-fold cross-validation accuracies.*

Figure 9 describes the cross-validation accuracies during the 10-fold validation. The accuracy generally oscillated between 0.60 and 0.70, with noticeable peaks and drops. There were two significant drops in the accuracy around iterations 100 and 150. These drops indicate that the model struggled on certain validation folds, likely due to challenging data distributions in those specific splits. The figure also contains the standard deviation in accuracy for each fold. High variance (wider bands) indicates more variability and sensitivity to the specific training/validation split. Lower variance (narrower bands) suggests more consistent performance. Despite the drops, the model maintains relatively stable accuracy around 0.68 for most iterations. This suggests that the model generalizes well across different folds, except for a few challenging splits (which might have been caused by noise in physiological responses, class imbalance or session-specific patterns).

Figure 10 shows the normalized confusion matrix. The values are presented as percentages, representing the proportion of correct and incorrect predictions for each class. The most significant misclassifications occurred between the bee and hermit and between the bee and spider: possibly due to moderate arousal levels. The wolf and the spider showed similarity in fear or excitement responses. As suggested before, there was a low arousal reaction to the hermit; therefore, there was difficulty in distinguishing neutral or calm states.

The model performed well overall, with "other" being the easiest class to identify, likely due to its neutral emotional signature. The most challenging classes were the bee and the wolf, indicating overlapping emotional states or physiological responses with other animals.

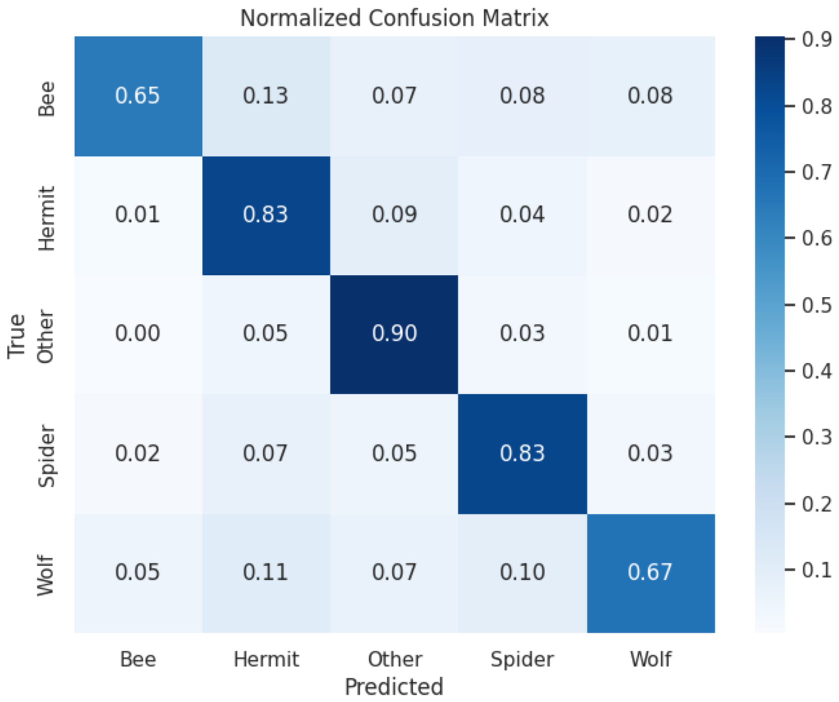

Figure 10. *Normalized confusion matrix.*

Accordingly, the machine learning results should be interpreted as evidence of physiological separability between VR scenes rather than as optimized emotion recognition performance. More complex models may yield higher accuracy in future work, but their deployment should be guided by dataset scale, interpretability requirements and ethical constraints, particularly in studies involving vulnerable populations.

5.1 Limitations

This study had several limitations that qualify interpretation and generalizability.

- Fixed-order stimulus presentation is a primary limitation of this study. Since all the participants encountered the same sequence of animals, order and carryover effects may have influenced physiological responses. Later scenes may have been affected by anticipatory arousal, habituation to the VR environment or cumulative emotional load. Although within-subject baseline correction was applied, this approach cannot fully isolate stimulus-specific effects. Accordingly, the findings should be interpreted as exploratory indicators of physiological differentiation across scenes rather than conclusive evidence of stimulus-specific emotional intensity. Future studies should employ counterbalanced or randomized presentation orders (e.g., Latin-square designs) to establish stronger internal validity.
- Time-series dependence and prior ML framing. The per-second tabular approach implicitly assumed independence and allowed potential leakage when random splits mixed windows from the same participant across train/test; we remedied this via windowing and removed identifier predictors (e.g., session). The reported performance should be interpreted as conservative and feasibility-focused.
- Sensor and preprocessing constraints. Signal interruptions, missing samples and device idiosyncrasies (e.g., GSR dropouts) required interpolation and cleaning that may have smoothed short-lived responses. Although we standardized these steps and excluded incomplete sessions, residual artifacts may remain.
- Sample and context. Participants were primarily students and teachers in a controlled VR setting; responses may differ in clinical populations or real-world exposure contexts, limiting external validity.

- Scope of inference. The findings are framed as physiological differentiation across scenes in a fixed order, not as proof of animal-specific fear signatures. Confirmatory work should employ counterbalancing/Latin-square designs, expand modalities (e.g., HRV-specific features) and consider sequence-aware models (e.g., HMM/TCN/LSTM).

6 CONCLUSION

This study demonstrates the effectiveness of virtual reality therapeutic games in eliciting and measuring emotional states using physiological signals such as heart rate and galvanic skin response. By integrating IoT wearable devices and advanced machine learning models, we successfully classified emotional responses induced by interactions with various virtual animals. The random forest classifier proved effective in capturing complex non-linear patterns in physiological data, achieving an average accuracy of 69% during cross-validation and 81% on the test set.

Our results highlight the importance of personalized baselines, as emotional responses varied significantly across sessions. The study identified GSR as the most influential feature in classifying emotional states, followed by heart rate and session-specific variations. Notably, the highest arousal levels were observed during interactions with the spider and the bee, while the hermit elicited calmer responses. Misclassifications were mainly observed between the bee and the hermit, as well as between the wolf and the spider, suggesting overlapping emotional states that warrant further investigation.

This research contributes to the growing field of VR-based mental health therapies by offering insights into designing emotionally adaptive therapeutic games. It underscores the need for personalized emotional baselines and highlights the potential of VR environments to simulate realistic emotional experiences safely and controllably. This work supports the application of cognitive-behavioural principles in virtual environments, particularly through the lens of exposure therapy, emotional processing theory (Foa and Kozak, 1986; Francisti et al., 2023) and fear extinction models. Future research should explore more advanced machine learning models, incorporate additional physiological signals and investigate session-specific emotional patterns to enhance the accuracy and effectiveness of VR therapeutic interventions.

These findings pave the way for the development of personalized VR-based systems that can adapt content based on users' physiological responses. Importantly, the present results should be viewed as supporting the feasibility of physiological monitoring and scene differentiation within VR exposure settings, rather than as evidence of therapeutic efficacy. Future work should therefore include controlled clinical studies, counterbalanced designs and longitudinal assessments to evaluate whether such systems lead to meaningful emotional regulation or therapeutic outcomes.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding: This work was supported by the Scientific Grant Agency of the Ministry of Education of the Slovak Republic (ME SR) and of Slovak Academy of Sciences (SAS) under the contract No. VEGA 1/0385/23 and VEGA 1/0492/26.

Conflict of Interests: Author Martin Polák works at the GAMETHERAPY company, which developed the FearTherapy VR game used for this research. The FearTherapy VR game is offered as a product by the company. The other authors declare no conflict of interest.

Author Contributions: Z.B.: Conceptualization, Investigation, Writing – original draft, Funding acquisition, Writing – review and editing, Supervision, Project administration. K.F.: Investigation, Methodology, Software, Analysis of the results, Writing – original draft, Writing – review and editing, Validation. M.M.: Investigation. J.R.: Visualization, Data Curation. J.K.: Methodology, Formal Analysis. Š.K.: Formal Analysis. M.P.: Methodology, Writing – review and editing, Validation.

Institutional Review Board Statement: Authors confirm that all the methods were carried out in accordance with relevant guidelines, and ethical approval was obtained from the Ethical Committee of the Constantine the Philosopher University in Nitra (chairman Prof. Dr. M. Bauerová). Number of Ethics Committee Approval: UKF/1350/2024/191013:002. The Ethics Committee stated and confirmed that the research does not contradict any ethical rules and confirmed that all respondents were informed about the use of their data.

Informed Consent Statement: All participants were initially informed about the objectives of the research. All methods were performed in accordance with the relevant guidelines and regulations. Informed consent was obtained from all participants and/or their legal guardians. A certified psychologist was present during the sessions to provide post-session support.

Statement on the Use of Artificial Intelligence Tools: The authors declare that they didn't use artificial intelligence tools for text or other media generation in this article.

Data Availability: The data is available in a public repository: https://doi.org/10.57760/sciencedb.29414. Alternatively, the data is also available in a public repository at GitHub: https://github.com/ChrisFodor333/feartherapy.

REFERENCES

Afzal, S., Khan, H. A., Piran, M. J., & Lee, J. W. (2024). A Comprehensive Survey on Affective Computing: Challenges, Trends, Applications, and Future Directions. *IEEE Access*, 12, 96150–96168. https://doi.org/10.1109/ACCESS.2024.3422480

Al-Azani, S., & El-Alfy, E.-S. M. (2025). A review and critical analysis of multimodal datasets for emotional AI. *Artificial Intelligence Review*, 58(10), 334. https://doi.org/10.1007/s10462-025-11271-1

Alneyadi, M., Drissi, N., Almeqbaali, M., & Ouhbi, S. (2021). Biofeedback-Based Connected Mental Health Interventions for Anxiety: Systematic Literature Review. *JMIR MHealth and UHealth*, 9(4), e26038. https://doi.org/10.2196/26038

Alqudah, A. M., & Moussavi, Z. (2025). A Review of Deep Learning for Biomedical Signals: Current Applications, Advancements, Future Prospects, Interpretation, and Challenges. *Computers, Materials and Continua*, 83(3), 3753–3841. https://doi.org/10.32604/cmc.2025.063643

Balan, O., Moldoveanu, A., Petrescu, L., Moise, G., Cristea, S., Petrescu, C., Moldoveanu, F., & Leordeanu, M. (2019). Sensors system methodology for artefacts identification in Virtual Reality games. In *2019 International Symposium on Advanced Electrical and Communication Technologies, ISAECT 2019*. IEEE. https://doi.org/10.1109/ISAECT47714.2019.9069719

Beitle, J. (2021). *Video Games and Virtual Reality as a Relaxation Therapy: An EEG Study*. Master thesis. Northeastern University.

Bell, I. H., Nicholas, J., Alvarez-Jimenez, M., Thompson, A., & Valmaggia, L. (2020). Virtual reality as a clinical tool in mental health research and practice. *Dialogues in Clinical Neuroscience*, 22(2), 169–177. https://doi.org/10.31887/DCNS.2020.22.2/lvalmaggia

Boucsein, W. (2012). Applications of Electrodermal Recording. In W. Boucsein (Ed.), *Electrodermal Activity* (pp. 259–523). Springer US. https://doi.org/10.1007/978-1-4614-1126-0_3

Braithwaite, J. J., Derrick, D., Watson, G., Jones, R., & Rowe, M. (2015). *A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments*. Technical Report, 2nd Version. University of Birmingham.

Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., Asmundson, G. J. G., Carlbring, P., & Powers, M. B. (2019). Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. *Journal of Anxiety Disorders*, 61, 27–36. https://doi.org/10.1016/j.janxdis.2018.08.003

Charlton, P. H., & Marozas, V. (2022). Wearable photoplethysmography devices. In J. Allen & P. Kyriacou (Eds.), *Photoplethysmography* (pp. 401–439). Academic Press. https://doi.org/10.1016/B978-0-12-823374-0.00011-6

Choudhury, A. D., Banerjee, R., Kimbahune, S., & Pal, A. (2022). Chapter 3 – Sensor signal analytics. In A. D. Choudhury, R. Banerjee, S. Kimbahune, & A. Pal (Eds.), *New Frontiers of Cardiovascular Screening Using Unobtrusive Sensors, AI, and IoT* (pp. 61–89). Academic Press. https://doi.org/10.1016/B978-0-12-824499-9.00003-9

Christian, P., Ebert, E. & Beikirch, H. (2009). Physiological Sensing for Affective Computing. In T. Tao Jianhua and Tan (Ed.), *Affective Information Processing* (pp. 293–310). Springer. https://doi.org/10.1007/978-1-84800-306-4_16

Civitello, D., Finn, D., Flood, M., Salievski, E., Schwarz, M., & Storck, Z. (2014). How do physiological responses such as respiratory frequency, heart rate, and galvanic skin response (GSR) change under emotional stress? https://api.semanticscholar.org/CorpusID:220621974

Davie, N., & Hilber, T. (2016). Smartwatches as a learning tool: A survey of student attitudes. In *12th International Conference Mobile Learning 2016*, (pp. 117–122). IADIS.

Faust, O., Hagiwara, Y., Hong, T. J., Lih, O. S., & Acharya, U. R. (2018). Deep learning for healthcare applications based on physiological signals: A review. *Computer Methods and Programs in Biomedicine*, 161, 1–13. https://doi.org/10.1016/j.cmpb.2018.04.005

Feng, J. (2023). Research on physiological signal analysis based on clinical statistics and machine learning. In G. Royle & S. M. Lipkin (Eds.), *Second International Conference on Biological Engineering and Medical Science (ICBioMed 2022)* (126112G). SPIE. https://doi.org/10.1117/12.2669358

Foa, E., & Kozak, M. J. (1986). Emotional Processing of Fear. Exposure to Corrective Information. *Psychological Bulletin*, 99, 20–35. https://doi.org/10.1037//0033-2909.99.1.20

Folkman, S. (2013). Stress: Appraisal and Coping. In M. D. Gellman & J. R. Turner (Eds.), *Encyclopedia of Behavioral Medicine,* (pp. 1913–1915). Springer. https://doi.org/10.1007/978-1-4419-1005-9_215

Francisti, J., Balogh, Z., Reichel, J., Benko, Ľ., Fodor, K., & Turčáni, M. (2023). Identification of heart rate change during the teaching process. *Scientific Reports*, 13, 16674. https://doi.org/10.1038/s41598-023-43763-x

Francisti, J., Balogh, Z., Reichel, J., Magdin, M., Koprda, Š., & Molnár, G. (2020). Application Experiences Using IoT Devices in Education. *Applied Sciences,* 10(20), 7286. https://doi.org/10.3390/app10207286

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., & Slater, M. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological medicine*, 47(14), 2393–2400. https://doi.org/10.1017/S003329171700040X

Gametherapy. (2023). *FearTherapy on Meta.* https://www.meta.com/experiences/feartherapy/6831911683537521/

Hiwaki, O. (2024). Whole-Head Noninvasive Brain Signal Measurement System with High Temporal and Spatial Resolution Using Static Magnetic Field Bias to the Brain. *Bioengineering*, 11(9). https://doi.org/10.3390/bioengineering11090917

Joshi, S., P, S., Mohanram, N. J., & K, A. (2022). Data Mining for Human Emotions Classification Based on Skin Conductance Response and Heart Rate – A Survey. In *2022 International Conference on Computer Communication and Informatics (ICCCI)*, (pp. 1–4). IEEE. https://doi.org/10.1109/ICCCI54379.2022.9740917

Kim, K. B., & Baek, H. J. (2023). Photoplethysmography in Wearable Devices: A comprehensive review of technological advances, current challenges, and future directions. *Electronics*, 12(13), 2923. https://doi.org/10.3390/electronics12132923

Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., & Patras, I. (2012). DEAP: A Database for Emotion Analysis: Using Physiological Signals. *IEEE Transactions on Affective Computing*, 3(1), 18–31. https://doi.org/10.1109/T-AFFC.2011.15

Koprda, Š., Magdin, M., Šášik, M., Balogh, Z., Fodor, K., & Molnár, G. (2024). Identification of obtaining incorrect values from the GSR sensor during skin resistance measurements. In *2024 7th International Conference on Information and Computer Technologies (ICICT),* (pp. 509–513). IEEE. https://doi.org/10.1109/ICICT62343.2024.00089

Laborde, S., Mosley, E., & Thayer, J. F. (2017). Heart Rate Variability and Cardiac Vagal Tone in Psychophysiological Research - Recommendations for Experiment Planning, Data Analysis, and Data Reporting. *Frontiers in psychology*, 8, 213. https://doi.org/10.3389/fpsyg.2017.00213

Lee, J. E., & Yoo, S. K. (2020). Recognition of negative emotion using long short-term memory with bio-signal feature compression. *Sensors*, 20(2), 573. https://doi.org/10.3390/s20020573

Liao, Y., Gao, Y., Wang, F., Zhang, L., Xu, Z., & Wu, Y. (2025). Emotion recognition with multiple physiological parameters based on ensemble learning. *Scientific reports*, 15(1), 19869. https://doi.org/10.1038/s41598-025-96616-0

Liu, J., Wei, W., Kuang, H., Zhao, F., & Tsien, J. (2013). Changes in Heart Rate Variability Are Associated with Expression of Short-Term and Long-Term Contextual and Cued Fear Memories. *PloS One*, 8, e63590. https://doi.org/10.1371/journal.pone.0063590

Luong, T., & Holz, C. (2022). Characterizing Physiological Responses to Fear, Frustration, and Insight in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 28(11), 3917–3927. https://doi.org/10.1109/TVCG.2022.3203113

Magdin, M., Balogh, Z., Reichel, J., Francisti, J., Koprda, Š., & György, M. (2021). Automatic detection and classification of emotional states in virtual reality and standard environments (LCD): comparing valence and arousal of induced emotions. *Virtual Reality*, 25(4), 1029–1041. https://doi.org/10.1007/s10055-021-00506-5

Maples-Keller, J., Yasinski, C., Manjin, N., & Rothbaum, B.O. (2017). Virtual Reality-Enhanced Extinction of Phobias and Post-Traumatic Stress. *Neurotherapeutics*, 14, 554–563. https://doi.org/10.1007/s13311-017-0534-y

Martinez Parrales, R., & Téllez-Anguiano, A. (2022). Vibration-based Fault Detection System with IoT Capabilities for a Conveyor Machine. *Acta Polytechnica Hungaric*a, 19, 7–24. https://doi.org/10.12700/APH.19.9.2022.9.1

Mather, M., & Thayer, J. F. (2018). How heart rate variability affects emotion regulation brain networks. *Current Opinion in Behavioral Sciences*, 19, 98–104. https://doi.org/10.1016/j.cobeha.2017.12.017

Navarro, D., Sundstedt, V., & Garro, V. (2021). Biofeedback methods in entertainment video games. *Proceedings of the ACM on Human-Computer Interaction*, 5(CHI PLAY), 1–32. https://doi.org/10.1145/3474695

Paletta, L., Dini, A., & Pszeida, M. (2020). Emotion Measurement from Attention Analysis on Imagery in Virtual Reality. *Advances in Intelligent Systems and Computing*, 952, 13–20. https://doi.org/10.1007/978-3-030-20441-9_2

Philippe-Antoine Dubé Daniel Imbeau, D. D. I. A., & Leone, M. (2015). Prediction of work metabolism from heart rate measurements in forest work: some practical methodological issues. *Ergonomics*, 58(12), 2040–2056. https://doi.org/10.1080/00140139.2015.1044920

Polak, M., Tanzer, N., & Carlbring, P. (2022). PROTOCOL: Effects of virtual reality exposure therapy versus in vivo exposure in treating social anxiety disorder in adults: A systematic review and meta-analysis. *Campbell Systematic Reviews*, 18(3), e1259. https://doi.org/10.1002/CL2.1259

Quintero, L., Fors, U., & Papapetrou, P. (2023). Personalized Feature Importance Ranking for Affect Recognition From Behavioral and Physiological Data. *IEEE Transactions on Games*, 17(3), 594–603. https://doi.org/10.1109/TG.2023.3263070

Rachman, S. (1980). Emotional processing. *Behaviour Research and Therapy*, 18(1), 51–60. https://doi.org/10.1016/0005-7967(80)90069-8

Rho, Gianluca & Di Rienzo, Francesco & Marinai, Carlotta & Giannetti, Francesca & Arcarisi, Lucia & Righetti, Francesca & Vallati, Carlo & Carbonaro, Nicola & Tognetti, Alessandro & Greco, Alberto & Bufano, Pasquale & Zanoletti, Michele & Laurino, Marco. (2024). Preliminary Assessment of the Samsung Galaxy Watch 5 Accuracy for the Monitoring of Heart Rate and Heart Rate Variability Parameters. In *MEDICON'23 and CMBEBIH'23*, (pp. 22–30). Springer. https://doi.org/10.1007/978-3-031-49062-0_3

Seaman, J., Brown, M., & Quay, J. (2017). The Evolution of Experiential Learning Theory: Tracing Lines of Research in the JEE. *Journal of Experiential Education*, 40(4), NP1–NP21. https://doi.org/10.1177/1053825916689268

Sharma, K., Castellini, C., Van Den Broek, E. L., Albu-Schaeffer, A., & Schwenker, F. (2019). A dataset of continuous affect annotations and physiological signals for emotion analysis. *Scientific Data*, 6(1), 196. https://doi.org/10.1038/s41597-019-0209-0

Shikha, Aryan, Arya, L., & Sethia, D. (2022). HRV and GSR as Viable Physiological Markers for Mental Health Recognition. In *2022 14th International Conference on COMmunication Systems & NETworkS*, (pp. 37–42). IEEE. https://doi.org/10.1109/COMSNETS53615.2022.9668439

Singh, N. K., Kumar, A., Singh, S., Kumar, V., Kumar, D., & Singh, S. (2024). Bio-Signal Processing. In *Futuristic Trends in Biotechnology Volume 3 Book 11* (pp. 75–92). Iterative International Publishers. https://doi.org/10.58532/V3BJBT11P2CH1

Sorinas, J., Ferrández, J. M., & Fernandez, E. (2020). Brain and Body Emotional Responses: Multimodal Approximation for Valence Classification. *Sensors*, 20(1), 313. https://doi.org/10.3390/s20010313

Tan, Y. L., Chang, V. Y. X., Ang, W. H. D., Ang, W. W., & Lau, Y. (2024). Virtual reality exposure therapy for social anxiety disorders: a meta-analysis and meta-regression of randomized controlled trials. *Anxiety, Stress, & Coping*, 38(2), 141–160. https://doi.org/10.1080/10615806.2024.2392195

Taschereau-Dumouchel, V., Kawato, M., & Lau, H. (2020). Multivoxel pattern analysis reveals dissociations between subjective fear and its physiological correlates. *Molecular Psychiatry*, 25(10), 2342–2354. https://doi.org/10.1038/s41380-019-0520-3

Valenti, R., & Gevers, T. (2012). Accurate eye center location through invariant isocentric patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(9), 1785–1798. https://doi.org/10.1109/TPAMI.2011.251

Vidyasagar, K. E. C., Kumar, K. R., Sai, G. N. K. A., Ruchita, M., & Saikia, M. J. (2024). Signal to Image Conversion and Convolutional Neural Networks for Physiological Signal Processing: A Review. *IEEE Access*, 12, 66726–66764. https://doi.org/10.1109/ACCESS.2024.3399114

Wang, Y., Qian, K., Nelson, J., Yagi, H., Kishi, A., Morita, K., & Yamamoto, Y. (2020). Can Affective Computing Better the Mental Status of the Electronic Games Player? A Perspective. In *LifeTech 2020 - 2020 IEEE 2nd Global Conference on Life Sciences and Technologies*, 366–367. https://doi.org/10.1109/LIFETECH48969.2020.1570620404

Wu, Y., Krishnan, S., & Ghoraani, B. (2022). Computational Methods for Physiological Signal Processing and Data Analysis. *Computational and Mathematical Methods in Medicine*, 2022(1), 9861801. https://doi.org/10.1155/2022/9861801

Yang, P., Liu, N., Liu, X., Shu, Y., Ji, W., Ren, Z., Sheng, J., Yu, M., Yi, R., Zhang, D., & Liu, Y.-J. (2024). A Multimodal Dataset for Mixed Emotion Recognition. *Scientific Data*, 11(1), 847. https://doi.org/10.1038/s41597-024-03676-4

Zielina, M., Šmahaj, J., Raudenská, J., & Javůrková, A. (2022). Využívání a vytváření terapeutických her ve virtuální realitě a model hráč/hra/terapie. *Československá Psychologie*, 66(3), 332–348. https://doi.org/10.51561/cspsych.66.3.332

Acta Informatica Pragensia is published by the Prague University of Economics and Business, Czech Republic | eISSN: 1805-4951