Domain Adaptation for Fear of Heights Classification in a VR Environment Based on EEG and ECG

Andrea Apicella[1] · Pasquale Arpaia[1] · Simone Barbato[2] · Giovanni D'Errico[4] · Giovanna Mastrati[1] · Nicola Moccaldi[1,5] · Ersilia Vallefuoco[1] · Selina Christin Wriessnegger[3]

Abstract

Three levels of fear of heights were detected in subjects with different severities of acrophobia, based on the electroencephalographic (EEG) and electrocardiographic (ECG) signals. The study aims to demonstrate the feasibility of a data-fusion-based method for real-time assessment of the fear of heights intensity to integrate into adaptive Virtual Reality Exposure Therapy for acrophobia. The generalization performance of classification tasks on fear states is improved by exploiting both trait-based clustering and Domain Adaptation methods. Participants were gradually exposed to increasing height levels through a Virtual Reality (VR) scenario representing a canyon. The initial severity of fear of heights, the level of distress at each height, and the anxiety level before and after the exposure were assessed through the Acrophobia Questionnaire, the Subjective Unit of Distress, and the State and Trait Anxiety Inventory, respectively. The Simulator Sickness Questionnaire was administered to exclude possible motion sickness interference in the experiment. The EEG and ECG signals were acquired through a 32-channel headset and 1 Lead ECG derivation during the exposure to the eliciting VR scenario. Four classifiers (i.e. Support Vector Machines, Deep Neural Networks, Random Forests, and k-Nearest Neighbors) were adopted in the experimental environment. Preliminary tests were performed in a within-subject experiment, achieving the best classification accuracy of $87.1\% \pm 7.8\%$ with a Deep Neural Network. As the cross-subject approach is concerned, three strategies, namely Domain Adaptation (DA), data fusion (combining EEG with ECG), and participant clustering (based on the acrophobia severity), were evaluated. DA resulted in the most effective strategies by determining an improvement of more than 20 % in classification accuracy. Random Forest performed the best classification accuracy for the severe acrophobia cluster with a mean of 63.6% and a standard deviation of 13.4% over three classes by exploiting Stratified Normalization.

Keywords Electroencephalography (EEG) · Virtual Reality (VR) · fear of heights · domain adaptation

1 Introduction

Specific phobias are anxiety disorders characterized by a marked and persistent fear of a specific object or situation, which leads to significant distress and avoidance behavior Association et al. (2015). The fear is usually out of proportion to any actual danger posed by the situation. Specific phobias can have a significant impact on quality of life, thus compromising daily functioning (Choy et al., 2007; Magee et al., 1996; Essau et al., 2000). Frequently, specific phobias occur before the onset of other mental disorders and can be considered early indicators of vulnerability to psychopathology (Wardenaar et al., 2017).

Both psychological and pharmacologic treatments are used for specific phobias. The main psychological therapies are cognitive-behavioral therapy (CBT) and exposure therapy (Wolitzky-Taylor et al., 2008). CBT is a talking therapy aimed to support the patient in changing his negative thoughts and beliefs about phobic stimuli. During exposure therapy, individuals are gradually exposed to phobic stimuli in a safe and controlled environment, with the help of a therapist. This therapy can help individuals to overcome their

✉ Andrea Apicella
andrea.apicella@unina.it

Giovanna Mastrati
giovanna.mastrati@unina.it

[1] Department of Electrical Engineering and Information Technology, University of Naples Federico II, Naples, Italy

[2] Idego Digital Psychology, Rome, Italy

[3] Institute of Neural Engineering, Graz University of Technology, Graz, Austria

[4] Department of Applied Science and Technology, Politecnico di Torino, Turin, Italy

[5] Department of Engineering for Innovation University of Salento, Lecce, Italy

avoidance behaviors and develop a sense of control over the fear (Thoma et al., 2015). This exposure can take the form of either imagined scenarios (imaginal modality), or real-life situations (in-vivo modality) (Hodges et al., 1995). In particular, *systematic desensitization* combines exposure with relaxation techniques (Marks & Gelder, 1965). On the other hand, *graded exposure* entails gradually exposing the individual to the phobic source in a carefully controlled environment without the utilization of relaxation techniques (Boehnlein et al., 2020).

In recent years, Virtual Reality (VR) has been increasingly used to implement exposure therapy (Powers & Emmelkamp, 2008). By means of Virtual Reality Exposure Therapy (VRET), the therapist can control the intensity and duration of the exposure, according to the patient's reaction. VRET has several advantages over traditional exposure therapy, including greater control over the exposure, the ability to create personalized scenarios, and the opportunity for patients to practice coping skills in a safe environment. It has been shown to be effective in a variety of clinical settings and is increasingly being used as a therapeutic tool (Botella et al., 2017; Parsons & Rizzo, 2008).

Traditional VRET treatments require operator intervention to modulate phobic stimuli. In recent years, adaptive solutions have been proposed by combining VR and biosignals monitoring (Arpaia et al., 2022; Apicella et al., 2022; Choo and A. May, 2014; Kosunen et al., 2016). Heart rate variability (HRV), skin conductance, eye movements, and electroencephalographic (EEG) signals are typical biosignals used in this context. In particular, the portability and wearability advancements of EEG devices have made them compatible with VR headsets (Arpaia et al., 2022), enabling their simultaneous use in therapeutic interventions (Andersen et al., 2023). Notably, EEG signals exhibit information richness and high temporal resolution (Liu et al., 2021; Uchitel et al., 2021; Yoshimura et al., 2017) resulting successfully employed in real-time brain-computer interface (BCI) applications (Arpaia et al., 2021a, b). Furthermore, also the ECG signals offer distinctive advantages such as ease of integration with VR setups, providing valuable information on autonomic nervous system responses, and offering insights into emotional states with high temporal resolution (Arpaia et al., 2023; Bornas et al., 2006).

Five subtypes of specific phobias are defined in the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (APA, 2013): animal, natural environment, blood/injection/injury, situational, and other (Lépine et al., 2005; Straube et al., 2006). Acrophobia is the most prevalent situational-specific phobia and is characterized by an excessive and irrational fear of heights or elevated places, leading to avoidance of such situations (Menzies, 1997).

Machine Learning (ML) allows to automatically identify the severity of fear of heights by analyzing the EEG signals

(Wang et al., 2021). However, real-time adaptivity requires the assessment of the current mental condition of the user (by focusing on his/her *state* (Steyer et al., 1992)) instead of the diagnosis of acrophobia disease severity (a feature of his/her *trait*). At the same time, the information on the acrophobia trait can be exploited to improve the state classification by clustering the participants according to their acrophobia severity and so handling the EEG bias related to specific traits. Moreover, EEG signals suffer from high nonstationarity (Shen & Lin, 2019), leading toward a significant statistical difference between signals acquired from different subjects or from the same subject at different times. ML literature is addressing this issue by adopting methods used for the Dataset Shift problem (Apicella et al., 2023). In a nutshell, starting from the assumption that data acquired in different conditions belong to different distributions (*domains*), these methods project the data in new feature spaces where these differences are relieved. In particular, Domain Adaptation (DA) approaches applied to EEG data are reporting interesting results in the literature, also in challenging problems such as emotion recognition (Lan et al., 2018). However, less attention is given to the application of DA strategies to phobias analysis and classification.

In a prior research investigation, a pipeline for EEG analysis was suggested to categorize fear of heights into three distinct levels of intensity (namely, low, medium, and high) (Apicella et al., 2023). In this work, three strategies were implemented to improve the cross-subject and cross-session generalization capability of the aforementioned processing pipeline (i) *clustering* of the experimental sample according to the acrophobia severity level, (ii) *data fusion*, and (iii) *domain adaptation* methods. In Section 2, related works are presented and discussed. The experimental sample, the instrumentation, the protocol, the psychometric and signal processing tools are described in Section 3. This marks the initial stride towards the development of a BCI adaptive system designed for VR therapy aimed at addressing the fear of heights.

2 Background

In (Wang et al., 2021), the EEG was employed to distinguish different groups of subjects according to their severity of acrophobia. The Acrophobia Questionnaire (AQ) (Cohen, 1977) was employed for an initial assessment of the experimental sample. The experiment ("Richie's Plank Experience on Steam") consisted of the exposure of the participants to a fear-inducing virtual environment reproducing a wooden plank hanging at a height of about 160 m. After the exposure, subjects reported their feelings through the Subjective Unit of Distress (SUD) (Keptner et al., 2021). The AQ and SUD scores were employed to divide all the subjects into three

groups, that are *not fear of heights*, *slight fear of heights*, and *severe fear of heights*. However, the SUD scale, typically used to assess the current state of the subject, in (Wang et al., 2021) was utilized in conjunction with the AQ scale to measure the subject's trait. EEG data were processed following the Harvard automated processing pipeline, decomposed in nine-layer with Wavelet Packet Decomposition (WPD), and functional connectivity features were evaluated between each pair of channels. The resulting features are then fed to Convolutional Neural Networks (CNNs) achieving an accuracy of 98.46 ± 0.42 %. However, the study was based exclusively on the acrophobia trait without considering the fear state. Moreover, only the EEG signal was exploited. In (Tychkov et al., 2023), a spectral analysis was conducted to identify markers for anxious-phobic disorders, specifically acrophobia. The analysis was performed under various conditions, including rest, while wearing a VR headset, and during exposure to a stressful stimulus (a height in a VR environment through the "New City" scene). Power spectrum in the following frequency ranges: (0.5-4.0 Hz), (4.0-8.0 Hz), (8.0-12.0 Hz), and (12.0-35.0 Hz) was explored on 12 male subjects. Results demonstrate that the exposure to height causes an increase in the β power spectrum by 2-3 times with respect to the state of rest and regardless of the intensity of the fear manifestation.

The results were obtained by comparing the power spectrum ranges computed under the aforementioned different conditions. The study demonstrates that it is possible to discriminate between the three conditions (rest, while wearing a VR headset, and at a certain height in a VR environment) based on the EEG signal. However, this study does not consider different levels of fear or the acrophobia traits of participants. Furthermore, the analysis is limited to EEG signals and no machine-learning strategies are exploited.

In (Aspiotis et al., 2022), the EEG and ECG signals were monitored to explore their relationship with height-related stress. The participant was exposed to a fear-inducing virtual environment, the "Richie's Plank Experience on Steam". Subjects reported their perceived stress during the exposure through the Perceived Stress Scale (PSS). The scale assesses the stress felt in the last month but it is here employed as a measure of state. The experimental sample made of 16 subjects was divided in two groups according to Hearth Rate variation with respect to a baseline. As EEG signal is concerned, a fourth-order Butterworth band-pass filter was applied to filter between 0.4 and 48 Hz. Artifact rejection was performed by using the Artifact Subspace Reconstruction (ASR) and the Independent Component Analysis (ICA). The Power Spectral Density (PSD) of each frequency band at each electrode was calculated using the Welch method. Each frequency band was averaged across the electrodes for each

cortex of the brain, leading to the calculation of the average absolute power for each brain region. The PSS score was found to be correlated with the increase in the frontal, parietal, occipital and, temporal beta powers and the increase in the parietal, temporal, and occipital gamma powers. Furthermore, statistically relevant differences emerged between the EEG biomarkers of the two groups identified on the basis of the Hearth Rate variation. Also in this case, the study does not consider different levels of fear neither the acrophobia traits of participants. Furthermore, no machine learning strategy is adopted.

Differently from the previous studies, in (Bălan et al., 2020) different fear levels are considered in a classification problem based on biosignals. The participants' severity of acrophobia was preliminarily assessed by the Visual Height Intolerance Questionnaire (VHIQ) (Huppert et al., 2017). Forty-four subjects were involved in VR-based experimental activities. Subjects were required to rate the fear through the SUD scale. EEG data were processed through a bandpass Butterworth filter. The signal was then averaged and log-normalization was applied. Finally, the alpha, beta and theta frequency powers were extracted. A range of ML and deep learning classifiers were used and evaluated in both user-dependent and user-independent scenarios. The fear levels were categorized by the authors into two and four classes, forming two distinct classification problems. In particular, a 4-class (relax, low, medium, and high) fear level classification was conducted starting from the Galvanic Skin Response, Hearth Rate, and EEG signals. The best performances of 79.12% and of 52.75% of accuracy were achieved in the user-dependent setting through an SVM and of 89.50% and of 42.50% in the user-independent setting through a Deep Neural Network (DNN) on two and four classes, respectively. Data fusion and machine learning strategies were employed to evaluate the participants' level of fear. However, a notable drawback lies in the absence of employing leave-n-subject-out strategies. Indeed, the Authors adopted a 10-fold cross-validation strategy, randomly distributing participants' data across these 10 folds during the validation process. This choice limits the extent to which the claimed accuracy results can be generalized. Similarly to this last approach, the current work aims to recognize different levels of fear of heights. Differently, the innovative contribution with respect to the discussed literature is twofold: (i) incorporating the trait of acrophobia into a classification problem regarding the fear of height state and (ii) the use of Domain Adaptation to handle the non-stationarity of EEG signals. The proposed approaches are evaluated by means of leave-one-out-subject strategies.

3 Materials and Methods

3.1 Psychometric Tools

Participants were required to fill in some scales and questionnaires for an initial screening of the sample and for pre- and post-observation of the anxiety level. The psychometric tools employed are (i) the Acrophobia Questionnaire (AQ) (Cohen, 1977) before the experiments; (ii) the State-Trait Anxiety Inventory (STAI) - Y1 before and after the experiments, (iii) the Simulator Sickness Questionnaire (SSQ) (Kennedy et al., 1993) at the end of the activity, and (iv) the Subjective Unit of Distress (SUD) (McCabe, 2015) after each run. The AQ and SUD scores were used to evaluate the severity of fear of heights.

The AQ self-report is made of two sub-scales which evaluate respectively the anxiety and avoidance levels associated with 20 height-relevant situations. Each sub-scale is made of 20 items.

The SUD was used to observe the reaction of the subjects to the different heights during the ongoing exposure session. This visual analogue scale is a widely employed tool for individuals to self-assess and report their levels of anxiety, restlessness, stress, or other unpleasant emotions experienced during exposure therapy. The participant can express an opinion about the current level of anxiety on a Likert scale from 0 (no distress) to 100 (extreme distress). In this research, a Likert scale ranging from zero to ten for assessing levels of distress was used. During exposure therapy, SUD is primarily utilized to construct fear hierarchies, organizing triggering stimuli based on their intensity levels. SUD ratings are commonly used to assess the initial level of fear experienced by the participants as well.

The STAI-Y1 was used to evaluate the efficacy of the fear eliciting stimulus through a comparison of the state anxiety levels acquired before and after the experiment. The STAI is a widely used psychological assessment tool designed to measure two distinct types of anxiety: state anxiety (Y1) and trait anxiety (Y2). State anxiety refers to the temporary or situational anxiety that individuals experience in response to a specific event or situation. Trait anxiety refers to the more stable and enduring aspect of anxiety that is characteristic of an individual's personality. The STAI is composed of 20 items for measuring state anxiety and 20 items for measuring trait anxiety on a 4-point Likert scale. Finally, the SSQ was employed to exclude motion sickness induced by the VR headset. It was designed to measure the intensity of three main symptom clusters associated with simulator sickness: (i) nausea, (ii) oculomotor (eye-related) symptoms, and (iii) disorientation symptoms. The SSQ consists of 16 items, and users are asked to rate their experience of each symptom on a scale from 0 (no symptom) to 3 (severe symptom). The total score is calculated by summing the ratings across all items, with the score indicating the level of simulator sickness: negligible (< 5), minimal ($5 - 10$), significant ($10 - 15$), and concerning ($15 - 20$) symptoms.

3.2 Participants

A sample of 20 healthy subjects (age 26.3 ± 7.9; 8 males and 12 females) took part in the experimental activity. The experimental protocol was approved by the ethical committee of the University of Naples Federico II. Participants were recruited within the university context. Thirty-five subjects voluntarily joined a call shared on social networks. Then they were divided into clusters according to their acrophobia level based on AQ and SUD scores. Finally, twenty participants were included in the study to manage the trade-off between sample size and cluster balance. Subjects had never participated in experiments involving emotion-related stimulation in a VR environment. The benefits and risks of the experimental procedure were clearly explained and the participants were instructed on the purpose of the experiment. Prior written informed consent to participate was provided by all the subjects.

3.3 Hardware

The EEG signals were acquired through the *LiveAmp* amplifier from Brain Products (Liveamp, 2022) (Fig. 1). The system is wearable and ultra-light, it is equipped with 32 gel-based active electrodes placed according to the International 10/20 Positioning System. The electrodes are placed on the scalp by means of caps of different sizes. Specifically, the actiCAP is provided with impedance conversion circuitry and visual feedback on the electrode-scalp impedance.

The LiveAmp comes with an ADC that has a resolution of 24 bits. The EEG signal can be recorded at three different sampling rates, namely 1000, 500, and 250 Sa/s. According to the sample rate, data are filtered through the amplifier's built-in third-order low-pass filter with a cut-off frequency of 262, 131, or 65 Hz, respectively. The amplifier is also provided with a built-in 3-axis acceleration sensor. Data can be wireless transmitted and/or internally stored on a micro memory card. The BrainVision Recorder software guides through the entire hardware setup. Specifically, it enables the ability to alternate between different channel configurations, verify the electrical resistance between the electrodes and the surface of the scalp, and view the EEG data in real time. The LiveAmp LSL connector app allows to connect the amplifier to LabStreamingLayer (LSL) thus enabling the unified and synchronized collection of data streams from different sources. Besides the EEG signal, it is possible to record other physiological signals thanks to the 8 AUX inputs provided by the *sensor & trigger extension* (Fig. 1b) connected with the LiveAmp (Fig. 1a), Moreover, the *bipolar-to-auxiliary*
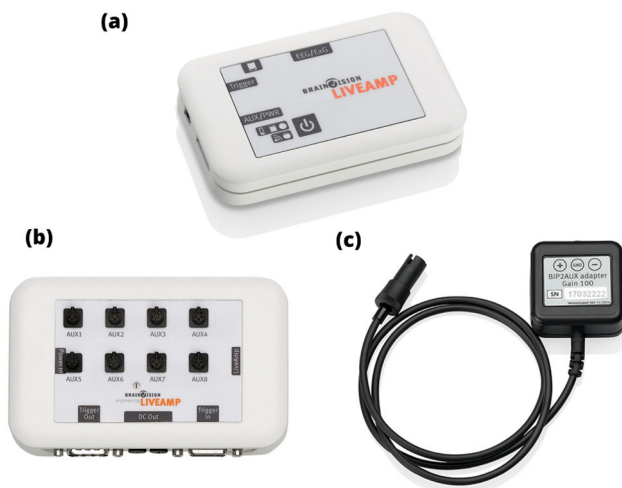
Fig. 1 Brain Vision LiveAmp (a) , Sensor & Trigger Extension (b), and Bipolar-to-auxiliary adapter (c)

(BIP2AUX) adapter (Fig. 1c) was connected to the STE in order to measure a bipolar recording of the ECG signal. The BIP2AUX adapter is an analog differential DC amplifier that optimizes the input signal by improving the quality of the acquired signal. The potential difference between the right arm and left arm was measured with the I Lead ECG.

The exposure to the VR environment was achieved by the *Meta Quest 2* (Meta quest 2, 2022), produced by Meta Platforms (Fig. 2). The Quest 2 runs an Android-based operative system, it can be used as a standalone device or can be connected through a wired connection to a PC with the VR software running on it. The headset is provided with a fast-switch LCD display with a per-eye resolution of $1832 * 1920$ and an adjustable refresh rate of 60, 72, or 90 Hz. Loud and left/right positional audio are available thanks to the built-in speakers thus offering the users a more immersive experience. A motion tracker with 6 degrees of freedom (DOF) guarantees precise tracking of the user's head and body movements. Finally, the Meta Quest 2 can be easily integrated in EEG applications due to its small dimensions and weight, $(224 * 450\ mm, 503\ g)$.

Fig. 2 Meta Quest 2

3.4 VR App

The *AKRON* application[1], developed by IDEGO (Idego - Digital Psychology, 2022), was employed in this study for the treatment of fear of heights. The scenario represents a canyon in a rocky desert: the river, the rocks, and the barren nature dominate the landscape, Fig. 3. The app allows to gradually expose the user to different height levels arranged in ascending order. Some elements such as the wall steepness, the river, and the wooden platform were exploited to enhance the eliciting power of the scenario (Azimisefat et al., 2022). The rock steepness contributes to increasing the sense of "dangerousness". The river provides the user with a spatial reference point and enhances the overall sense of depth. Finally, the wooden platform was built with spaces between the planks to provide the user with reference points downwards, and also with the aim of increasing the perception of height and depth. The wooden lift allows people to go up to 4 distinct floors: ground floor, first floor, second floor, and third floor. On each floor, the platform rises approximately 15 m in height. The lift is equipped with protective barriers present on each side with the aim of letting the user feel safe while the platform rises. The only exception is when the platform stops, in that situation the frontal barriers are no longer present in order to leave the user in greater eye contact with the sensation of empty space. The application was created using the Unity game engine (version 2019.4.16) on the Android platform. The application was created for Android using the Unity game engine (version 2019.4.16), programmed in C# and utilized the OpenGLES3 graphics library, along with an IL2CPP-type backend configuration. The app has been optimized through a ASTC compression system to use it on low-end VR viewers such as the Meta Quest 2. The refresh rate was set at 72 Hz.

3.5 Experimental Protocol

The activities were carried out at the Institute of Neural Engineering (BCI Lab) at the Graz University of Technology. Each subject concluded one session composed of three trials, on the same day (Fig. 4). A preliminary phase preceded the experimental activities. First, participants were carefully instructed on the purpose of the experiment after the researchers set up the EEG device and the VR headset. The use of wet electrodes filled with conductive gel guaranteed electrode-scalp impedance lower than $25k\Omega$ and the quality of the EEG signal was visually inspected. Once the EEG configuration was completed, participants were asked to wear the VR headset which had been previously calibrated by the

---

[1] Video of the application: https://shorturl.at/dgluI.

Fig. 3 Eliciting VR scenario

operator. The EEG signals were again visually inspected to ensure no perturbations occurred (Fig. 5).

The individual needs to position themselves inside a VR setting reproducing a canyon where they will be standing in a wooden elevator. This elevator enables the user to progressively ascend to greater heights with each attempt. Before the first run, the subject stands at the ground level on the river bank and is asked to look around in order to become familiar with the environment. At the ground level, the baseline is recorded and the subject answers the SUD for the first time.

A 5 s visual countdown informs the subject of the start of each run. After, the platform starts rising and in a few seconds the subject reaches the desired level. The participant is required to stand on the platform at a certain height for 90 s. The run is followed by a 60 s relaxation phase at the

ground level and then the user is asked to answer the SUD again. The duration of the session is approximately 15 min. The overall experimental process is presented in Fig. 4.

3.6 EEG Signal Preprocessing and Feature Extraction

The pre-processing stage of the raw EEG data was realized using Matlab v. R2022a. A digital filtering was applied in order to filter out the power line noise and extract the frequency bands of interest. Specifically, a 50 Hz notch IIR filter and a 4th order Butterworth band pass IIR filter, with cutoff frequencies between 0.5 and 48.5 Hz were built and applied to the row data. A robust artifact removal was then required to clean the EEG data from the movement artifacts occurred during the experiments. Artifact Subspace Reconstruction
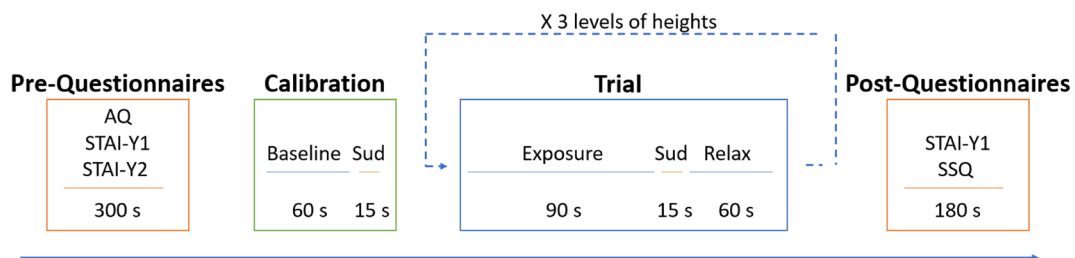

Fig. 4 Structure of the experimental session , including one-time pre/post-questionnaires (AQ = Acrophobia Questionnaire, STAI = State-Trait Anxiety Inventory, SSQ = Simulator Sickness Question-naire). After the initial assessment and the subsequent exposure stages, the Subjective Unit of Distress (SUD) is conducted

Fig. 5 EEG data acquisition

(ASR) and Independent Component Analysis (ICA) (Arpaia et al., 2022) were employed for removing artifacts from the EEG signal using the EEGLAB Matlab toolbox, version 2019 (Radüntz et al., 2015). For the ASR, a cutoff parameter of 15 was employed in order to retain the brain component of the signal while effectively removing artifacts (Chang et al., 2018). For the ICA, all the components identified as eye, muscle, hearth, line noise, and channel noise artifacts with a percentage greater than 95% were removed. Additionally, subjects' head movements in the x, y, and z dimensions recorded through the accelerometer sensors were used to further clean the EEG data from movement-related artifacts. Independent components (ICs) previously computed by ICA were band-pass filtered with a 4th order Butterworth band pass IIR filter between [1 − 10] Hz. The accelerometer data were filtered in the same way. The frequency band was cho-

sen in order to retain artifacts related to head movement while removing low-frequency drifts (Daly et al., 2013). Following, the Pearson's correlation coefficient between ICs and accelerometer signals is computed to identify ICs that were most likely related to head movement. IC components that resulted highly correlated (Pearson's correlation coefficient greater than two standard deviations above the mean correlation calculated between the IC and the accelerometer signals) with at least one of the accelerometer signals, were removed from the EEG signal (Table 6).

Next, signals were segmented into 10s time windows. Typical time windows for EEG signal processing range in the interval [1-10] s (Apicella et al., 2022). Furthermore, findings from a functional magnetic resonance imaging (fMRI) study (Caseras et al., 2010) indicate that brain activation increases compared to normal values within the timeframe of approximately 1 to 10 s. In (Knopf & Pössel, 2009), it is stated that in phobic subjects, heart rate (HR) begins to increase after 7 s from the start of stimulation. Additionally, it has been demonstrated (Shaffer and J. P. Ginsberg, 2017) that the heart rate variability (HRV), that we employ as an ECG feature, is more reliable when considering measurements of 10 s (Ultra Short-Term, UST). Therefore, a 10 s time window was adopted to take into account the temporal trend of both EEG and ECG signals. The Fast Fourier Transform (FFT) was then applied to the signals and the PSD was computed in the following frequency bands: *delta* [0.5, 4] *theta* [4, 8] Hz, *alpha* [8, 13] Hz, *low-beta* [13, 21] Hz, *high-beta* [21, 28] Hz, *low-gamma* [28, 38] Hz, and *high-gamma* [38, 48.5] Hz.

3.7 ECG Signal Preprocessing and Feature Extraction

The ECG signal processing was developed in Matlab v. R2021b. The Pan Tompkins algorithm (Pan & Tompkins, 1985; Sedghamiz, 2014) was applied to the raw ECG signal for the pre-processing phase. It consists of a series of filtering operations and of a final detection of the QRS complexes
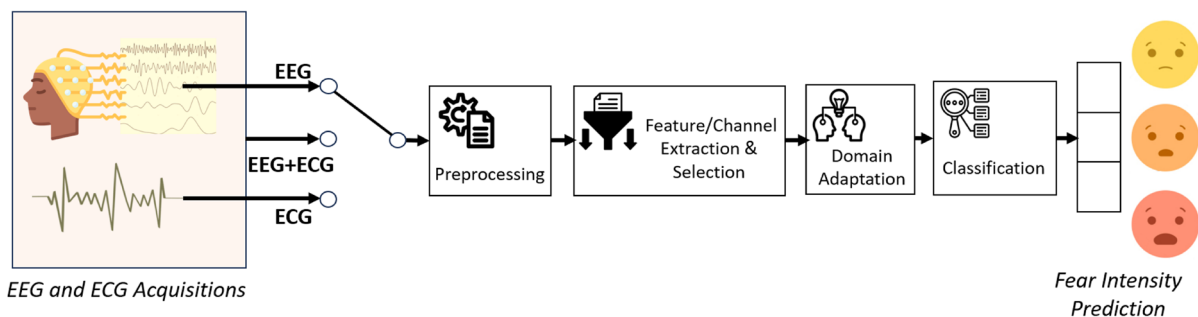

EEG and ECG Acquisitions

Fear Intensity Prediction

Fig. 6 Pipeline of the classification process involving EEG, ECG, or a combination of both. After the data has been preprocessed and segmented into epochs, suitable features are extracted, along with a channel selection strategy. Subsequently, a Domain Adaptation algorithm is applied, followed by a classification step that outputs the level of experienced phobic intensity

corresponding to the ventricular depolarization. Specifically, the algorithm expects the following steps:

- a digital filtering between ([5 − 15] Hz) carried out through a bandpass Butterworth filter. This first stage allows to attenuate the 50 Hz, muscle artifact, and baseline wander.
- a derivative filter that allows to highlight the QRS complexes;
- a signal amplitude squaring to enhance the high frequencies;
- a moving window integration with a window length of 30 samples to obtain information on the slope of the R wave;
- an adaptive thresholding with a decision rule algorithm that allows to distinguish ECG signal peak from noise peak and also allows to discriminate T-waves.

Once detected the NN interval, defined as the Normal-to-Normal interval obtained from the signal without abnormal beats, several features related to HRV were extracted (Table 1): linear statistical and geometric features, and non-linear features (T. F. of The European Society of Cardiology et al., 1996; Shaffer and J. P. Ginsberg, 2017).

3.8 Domain Adaptation and Classification

Due to the intrinsic non-stationary of the EEG signal, DA strategies have been explored in this study. In particular, we focus on Unsupervised DA strategies, which start from the hypothesis that unlabeled data of the target domain (i.e. EEG acquisitions belonging to the target subject/session) are available during the training. Unsupervised DA resulted particularly suitable because of the availability of target data during the training of the model, together with data belonging to the source domain(s). Unsupervised DA strategies can be categorized into two main families: feature-based approaches, where a proper feature transformation is induced and applied to the data before the classification training stage, and end-to-end approaches, where the most suitable feature space is learned together with the classification model. This last family is particularly suitable for ML methods based on Deep Neural Networks (DNNs), since they allow to build complex functional architectures able to extract features and pursue ML tasks at the same time. In particular, methods based on adversarial learning (HassanPour Zonoozi & Seydi, 2022) are gaining success in several applications. In this work, we explore how Subspace Alignment (SA, (Fernando et al., 2013)) feature method impacts classical ML methods (for instance, RF, SVM, and kNN) and how Domain-Adversarial Neural Networks (DANN, (Ganin et al., 2016)) impacts on DNNs in a Fear classification task. In a nutshell, SA searches for a linear transformation able to align the source and target spaces finding the best linear transformation of the source points projected in a PCA space. On the other side, DANN learns a DNN feature space considering the discrepancy between the source and the target domain with the aim of generating a common representation space such that data belonging to different domains are indistinguishable for an ad-hoc domain discriminator. To this aim, a DNN model able to project the data in a feature space able to maximize both the class prediction performances and the domain classification loss is trained. To verify if DA's methods are responsible for the resulting improvement, an ablation study was made to evaluate each classifier's performance with and without the DA approach for the proposed Fear classification task. These methods search for common features where the source and target distributions result aligned. The classification is then carried out in this encoded feature space. SA and DANN methods provided by the Python package ADAPT

Table 1 Features extracted from the ECG

Linear and Statistical	Linear Geometric	Non Linear
Root Mean Square of successive differences of NN intervals (RMSNN)	Triangular Index (TI)	SD1 and SD2
Standard Deviation of the NN Interval	Triangular Interpolation of the NN Interval Histogram (TINN)	SD1/SD2 ratio
Standard Deviation of successive differences of NN intervals (SDSD)		Tone
Mean value of NN intervals		
Median value of NN intervals		
Difference between the maximum and minimum value of NN intervals		
Median Absolute Deviation (MAD) of the intervals NN		
pNN50		
pNN20		

Table 2 Classifiers, optimized hyperparameters, and variation ranges

Classifier	Hyperparameter	Variation Range
Random Forest (RF)	Depth (D)	[3, 5, 7, 10]
	N° of Trees (NT)	[100, 200, 400, 600, 800, 1000]
Support Vector Machine (SVM)	BoxConstraint (BC)	{0.1, 1.0, 10.0, 100.0}
	KernelFunction (KF)	{rbf (Vert et al., 2004)}
	KernelScale (KS)	log-scaled in the range [1e-3,1e3]
k-Nearest Neighbour (k-NN)	N° Neighbors (NN)	{1,3,5,7,9,11}
	Weight	{uniform, inverse of distance (Coomans & Massart, 1982)}
	Nearest Neighbors Computation	{brute, ball tree (Omohundro, 1989), kd tree (Bentley, 1975)}
Deep Neural Network (DNN)	Activation Function (AF)	{ReLU(Nair & Hinton, 2010), sigmoid (Hecht-Nielsen, 1992), tanh (Hecht-Nielsen, 1992), VAF(Apicella et al., 2019)}
	N° of Hidden Layers	[1, 3]
	N° of Neurons for each Hidden Layer	[32, 1024]
DANN (used only for DNN case)	Activation Function (AF)	{ReLU (Nair & Hinton, 2010), sigmoid (Hecht-Nielsen, 1992), tanh (Hecht-Nielsen, 1992), VAF (Apicella et al., 2019)}
	N° of Hidden Layers for Domain Classifier	[1, 3]
	N° of Neurons for Domain Classifier	[32,1024] with step of 16
	N° of Hidden Layers for Feature Extractor	[1, 2]
	N° of Neurons for Feature Extractor	[32, 1024] with step of 16

were employed. Furthermore, experiments with and without Stratified Normalization (SN) were employed to reduce inter-participant variability (Fdez et al., 2021). Indeed, proper normalization strategies applied to the data can heavily affect the classification performance in EEG data classification, in some cases outperforming classical DA strategies (Apicella et al., 2023). k-Nearest Neighbors (k-NN, (Bishop & Nasrabadi, 2006)), Random Forests (RFs, (Ho, 1995)), Support Vector Machines (SVMs, (Cortes and V. Vapnik, 1995)),

and DNNs with fully-connected layers (LeCun et al., 2015) were the employed classifiers. The objective was to categorize fear of heights into three distinct levels of intensity (i.e., low, medium, and high). A Leave-One-Subject-Out (LOSO) Cross-Validation (CV) was employed in a cross-subject setting. Since we enrolled 20 subjects each of them providing 123 samples, each cross-validation round adopted $(20 - 1) \times 123$ points for the training and 123 samples for the test set. Instead, in the within-subject classification a strati-

Table 3 Cross-subject classification accuracy (mean and standard deviation) in % of fear of heights on a 3-level intensity scale for the whole experimental sample

Bio-signal	DA method	Classifier			
		RF	SVM	kNN	DNN
ECG	None	36.2 ± 10.2	35.0 ± 11.4	31.6 ± 7.0	35.1 ± 4.9
	SA	31.8 ± 7.1	33.0 ± 8.6	33.6 ± 7.4	–
	DANN	–	–	–	34.8 ± 11.8
	SN	43.0 ± 15.7	40.6 ± 11.0	41.1 ± 13.3	43.5 ± 15.8
EEG	None	43.0 ± 10.8	40.1 ± 9.5	38.4 ± 8.3	37.1 ± 7.7
	SA	40.8 ± 10.3	45.4 ± 18.6	46.3 ± 18	–
	DANN	–	–	–	46.0 ± 13.9
	SN	$\mathbf{56.5 \pm 11.1}$	54.0 ± 19.7	42.7 ± 15.9	55.1 ± 15.1
EEG & ECG	None	36.6 ± 5.6	40.1 ± 9.5	38.9 ± 10	37.5 ± 6.9
	SA	42.0 ± 7.5	46.2 ± 16.2	43.8 ± 14.7	–
	DANN	–	–	–	46.8 ± 13.7
	SN	56.4 ± 15.7	54.4 ± 18.4	42.0 ± 13	54.4 ± 14.6

fied 5-fold CV was employed, resulting in $123 \times 4/5$ points composing the training set and the remaining points composing the test set for each subject. Given the limited sample size, simple ML architectures, such as neural networks with max 3 layers, was chosen deliberately to prevent overfitting and to maintain a balance between model complexity and the available data's capacity to generalize. Indeed, while employing a more complex neural network with a larger number of parameters might offer a higher capacity to capture intricate relationships within the data, it could also result in overfitting, where the model memorizes the noise present in the limited dataset rather than learning meaningful patterns. For all the classifiers, the hyperparameters used during the CV procedure are reported in Table 2. The classification was conducted considering: (i) the whole experimental sample made of 20 subjects, and (ii) the three clusters of subject with different severity of fear of heights, separately. A stratified 5-fold CV was employed in the within-subject classification.

The entire pipeline related to the process of signal acquisition, preprocessing, feature extraction, Domain Adaptation, and classification is illustrated in Fig. 6.

4 Results

4.1 Psychometric Analysis

Relying on the AQ and SUD scores, subjects were grouped in 3 clusters according to the severity of fear of heights (Wang et al., 2021). Individuals were placed in the *slight-acrophobia* cluster when reported both AQ scores for anxiety and avoidance below 20 and 6, respectively, along with a SUD score below 2. Subjects simultaneously reporting AQ scores in ranges $(20, 40)$, $(6, 12)$ and a SUD score in range $(3, 4)$ were grouped together in the *mild-acrophobia* clus-

ter. Subjects simultaneously reporting AQ scores > 40 and > 12 and a SUD score > 5 were grouped together in the *severe-acrophobia* cluster. As a result, 7 participants were classified with slight acrophobia, 8 with mild acrophobia, and 5 with severe acrophobia. To evaluate whether there was a change in reported levels of fear, following the VR exposure, the Wilcoxon signed-rank test was used to analyze the pre- and post- scores of STAI-Y1. Statistical analysis was performed using R Software (version 4.1.1) and a p-value < 0.05 indicated statistically significant differences. By comparing the results from the STAI-Y1 acquired before and after the experiments, a significant difference emerged ($p = 0.037$, $V = 44$). STAI scores acquired after the VR exposure were higher than the scores acquired before, thus proving the effectiveness of the fear induction. From the analysis of the SSQ, an average score of 7.74 ± 3.51 established that subjects suffered from minimal motion sickness symptoms during the activity.

4.2 Classification

For the cross-subject analysis, the classification performance were computed by combining the following binary criteria: i) exploiting (or not) DA methods, and ii) exploiting only EEG data (or both EEG and ECG). In Table 3, accuracy was computed on the entire experimental sample. The best result was achieved with the combination of SN and RF applied to the EEG signal. The merge of EEG and ECG allows a compatible mean accuracy with a larger standard deviation. In Tables 4, 5, and 6, the accuracy values are reported for the severe, mild, and slight-acrophobia clusters, respectively. For all the clusters the combination of SN and RF allowed the best accuracy. Considering the type of signal, the ECG was significant for clusters of mild and slight acrophobia.

Table 4 Cross-subject classification accuracy (mean and standard deviation) in % of fear of heights on a 3-level intensity scale considering the severe acrophobia cluster

Bio-signal	DA method	Classifier			
		RF	SVM	kNN	DNN
ECG	None	32.0 ± 3.2	38.8 ± 5.5	43.0 ± 4.8	36.3 ± 10.1
	SA	32.7 ± 2.2	37.8 ± 3.9	38.8 ± 7.0	–
	DANN	–	–	–	34.5 ± 4.1
	SN	50.9 ± 4.4	29.6 ± 4.8	33.3 ± 4.3	29.7 ± 9.4
EEG	None	49.6 ± 21.8	49.1 ± 11.1	47.2 ± 12.8	56.9 ± 14.6
	SA	39.3 ± 6.0	56.3 ± 17	57.5 ± 13.8	–
	DANN	–	–	–	47.2 ± 5.2
	SN	$\mathbf{63.6 \pm 13.4}$	55.7 ± 9.6	53.3 ± 7.5	58.1 ± 7.0
EEG & ECG	None	49.1 ± 17.8	52.7 ± 5.2	46.1 ± 5.9	55.8 ± 18.2
	SA	44.8 ± 6.4	60.0 ± 13.8	52.7 ± 9.8	–
	DANN	–	–	–	55.1 ± 12.6
	SN	63.6 ± 15	54.5 ± 12.7	48.5 ± 5.4	55.1 ± 12.4

Table 5 Cross-subject classification accuracy (mean and standard deviation) in % of fear of heights on a 3-level intensity scale considering the mild acrophobia cluster

Bio-signal	DA method	Classifier			
		RF	SVM	kNN	DNN
ECG	None	40.2 ± 8.9	36.3 ± 10.7	33.7 ± 6.3	40.9 ± 10.2
	SA	33.7 ± 6.3	34.8 ± 4.7	35.2 ± 7.8	–
	DANN	–	–	–	30.3 ± 11.4
	SN	50.3 ± 15.3	37.8 ± 11.4	45.0 ± 15	48.5 ± 14.2
EEG	None	41.3 ± 14.6	36.3 ± 4.2	34.8 ± 6.9	35.2 ± 5.0
	SA	31.0 ± 5.8	34.8 ± 3.0	38.2 ± 15.8	–
	DANN	–	–	–	44.7 ± 13.2
	SN	50.7 ± 15.2	42.0 ± 7.3	34.0 ± 6.5	46.6 ± 16.2
EEG & ECG	None	39.0 ± 12.3	33.0 ± 0.0	34.1 ± 4.7	37.1 ± 10.2
	SA	36.0 ± 7.3	34.8 ± 3.7	36.0 ± 5.6	–
	DANN	–	–	–	48.9 ± 16.7
	SN	**52.6 ± 11.8**	44.3 ± 13.7	34.1 ± 8.8	52.2 ± 12.2

Results of the within-subject classification are reported in Table 7. DNN was the best classifier and the use of the ECG in addition to the EEG reduced the standard deviation.

5 Discussion

In this study, three distinct strategies, namely DA, data fusion, and participant clustering (based on the severity of fear of heights), were employed to enhance the classification accuracy. The effectiveness of fear induction was confirmed by the statistical significant difference between pre- and post-STAI state. The use of VR technology did not cause motion sickness to the participants, as confirmed by the SSQ results. Thus, the reported SUD and STAI scores were not affected from external factors and were only related to the felt fear of heights. The achieved classification accuracy is remarkable given that the chance level in a three-class problem is

33 %. This result is even more significant because a LOSO strategy was employed to ensure the generalizability of the results. However, it hard to compare the results with previous studies due to different scientific goals and experimental conditions such as the number of classes, the bio-signals considered, the hardware for signal acquisition, and the sample size. Despite the small sample sizes, DA strategies exhibited an increase of more than 20 % in accuracy (Table 3). In particular, SN resulted in the most effective DA method in all the experiments. The successful application of SN was demonstrated in literature with a cross-subject generalization scenario in an emotion recognition task (Fdez et al., 2021). In essence, SN operates by normalizing features to mitigate inter-participant variability, aiming to preserve only the pertinent emotion-related information within the data. The significance of normalization in EEG classification has been highlighted in (Apicella et al., 2023), showcasing that employing a suitable normalization procedure can

Table 6 Cross-subject classification accuracy (mean and standard deviation) in % of fear of heights on a 3-level intensity scale considering the slight acrophobia cluster

Bio-signal	DA method	Classifier			
		RF	SVM	kNN	DNN
ECG	None	32.0 ± 4.8	37.2 ± 8.5	33.7 ± 7.6	29.4 ± 9.5
	SA	34.2 ± 8.2	30.7 ± 8.9	26.4 ± 3.1	–
	DANN	–	–	–	27.7 ± 7.3
	SN	47.6 ± 12.1	40.7 ± 5.0	39.4 ± 10.4	42.0 ± 12.3
EEG	None	45.8 ± 11.1	38.5 ± 4.7	37.2 ± 7.5	41.1 ± 10.2
	SA	41.9 ± 6.1	45.8 ± 10.7	43.3 ± 10.3	–
	DANN	–	–	–	43.7 ± 14.5
	SN	57.1 ± 11.2	51.0 ± 11.8	37.2 ± 3.8	51.1 ± 11.3
EEG & ECG	None	45.4 ± 11.3	35.0 ± 4.2	41.6 ± 11	41.1 ± 11.0
	SA	44.1 ± 5.3	54.1 ± 13.4	45.4 ± 11.3	–
	DANN	–	–	–	47.6 ± 9.2
	SN	**58.0 ± 11.7**	52.0 ± 12.6	39.8 ± 6.1	54.1 ± 14.8

Table 7 Within-subject classification accuracy (mean and standard deviation) in % of fear of heights on a 3-level intensity scale

Biosignal	Classifier			
	RF	SVM	kNN	DNN
ECG	57.9 ± 13.7	52.9 ± 9.5	51.5 ± 8.9	48.9 ± 7.6
EEG	85.1 ± 14.1	84.5 ± 11	82.1 ± 13.3	**87.5±9.3**
EEG & ECG	86.4 ± 14	85.3 ± 10.9	84.0 ± 11.4	**87.1±7.8**

yield results comparable or even superior to more intricate DA techniques. Consequently, based on experimental evaluations, SN emerged as particularly well-suited for the specific task and dataset involved.

As far the data fusion is concerned, in almost all the cases, the obtained results show that ECG alone leads to worse results with respect to its use with EEG. However, the combined use of ECG and EEG had a marginal impact on the classification accuracy.

Clustering allowed an increment of accuracy, especially for the severe acrophobia cluster (more than 10 %). The impact of clustering was minor on the slight acrophobia cluster and negligible on the mild acrophobia cluster. Neurocorrelates of Acrophobia are well documented in the literature and also localized at the cortical level, therefore they are also detectable via EEG. In particular, studies based on the use of magnetic resonance imaging have focused on the analysis of acrophobic subjects at rest compared with control groups (Hang et al., 2022; Guo et al., 2023). To date, in our knowledge, there are no specific studies in the literature on the interaction between trait and state in acrophobic subjects. The results of the present study could be interpreted as the effect of a more accentuated sensitivity of the neuro correlates related to the acrophobic trait when persons suffering from severe acrophobia are exposed to phobic stimuli. Regarding the observed discrepancy in classification accuracy between severity levels of acrophobia, the severity of acrophobia appears to impact the intensity of the fear responses. Indeed, severe acrophobia tends to evoke more overt and distinguishable reactions, which could translate into more readily identifiable patterns in the EEG signals. These distinct patterns can help machine learning algorithms to more accurately recognize and classify instances of severe acrophobia compared to moderate cases. The achieved levels of accuracy in the medium and slight acrophobia clusters may be influenced by additional factors, including the influence exerted on EEG patterns by the vestibular system. This system processes spatial information input, impacting EEG signal modulation (Ibitoye et al., 2023; Ehinger et al., 2014). However, the fear system can affect vestibular system operation (Neumann et al., 2023). It can be hypothesized that, in the case of mild acrophobia severity, the vestibular circuit expresses new EEG patterns that are more difficult to discriminate. Similarly, the fear system also expresses itself at levels that are not particularly intense and therefore less discrim-

inable. Instead, when the phobic state does not perturb the vestibular system, EEG activity could be more detectable. Therefore, under a slight acrophobia condition, EEG patterns due to the vestibular system can enable the classifier to achieve higher accuracy than under a mild acrophobia condition. In contrast, when the phobic condition reaches high intensities, EEG patterns attributable to the fear system are more easily discriminable by the classifier.

Finally, the Random Forest model achieved the best accuracy of 63.6 % ± 13.4 in classifying the severe acrophobia cluster. The observed superior performance of the RF could stem from several factors, including the dataset used and its size. This is particularly evident in scenarios where the amount of available data is relatively small. Therefore, simpler models can better handle limited amounts of data without succumbing to overfitting (Bejani & Ghatee, 2021) or struggling with the vast parameter space typical of DNNs. Moreover, in cases where the patterns within the data are simpler to discern, traditional algorithms might benefit from handcrafted or domain-specific feature engineering tailored to EEG signal characteristics (such as PSD, adopted in the current study), leading superior performance as they can efficiently capture and model these patterns without requiring the depth and complexity inherent in neural networks.

6 Conclusions and Future Works

In the present study, three levels of fear of heights (namely, low, medium, and high) were detected in subjects with different severities of acrophobia starting from the EEG and ECG signals. The generalization performance of classification tasks on fear states is improved by exploiting both trait-based clustering and Domain Adaptation methods. A VR scenario representing a canyon was exploited to expose 20 healthy participants to increasing height levels. Subjects were asked to fill in some psychometric tools to assess their initial severity of fear of heights, the level of distress at each height, the anxiety level before and after the exposure, and motion sickness induced by VR through the AQ, the SUD, the STAI, and the SSQ, respectively. The EEG and ECG signals were acquired through a 32-channel headset and I Lead ECG derivation during the entire experimental activity. Three distinct strategies, namely Domain Adaptation (DA), data fusion (combining EEG with ECG), and participant clus-

tering (based on the severity of acrophobia), were employed to enhance the accuracy in the cross-subject classification. DA strategies exhibited the highest impact on the increment of accuracy, followed by clustering (especially for the severe acrophobia cluster). Data fusion had a marginal impact on the classification accuracy. Regarding the observed discrepancy in classification accuracy between severity levels of acrophobia, this result can be interpreted as support for the hypothesis that the severity of acrophobia (trait) impacts the intensity of the fear responses (state).

The study demonstrated the feasibility of a data-fusion-based method for real-time assessment of the fear of heights intensity to integrate into adaptive Virtual Reality Exposure Therapy for acrophobia. In future works the impact of specific EEG-band and further EEG features on the classification accuracy will be investigated to enhance fear of heights classification, aiming to improve both classification accuracy and cross-subject generalization.

Author Contributions Andrea Apicella, Pasquale Arpaia, Giovanni D'Errico, Giovanna Mastrati, Nicola Moccaldi, Ersilia Vallefuoco, and Selina Christin Wriessenegger collaborated equally in developing the research framework, designing experiments, and interpreting results. They collectively contributed to the drafting and critical review of the manuscript. Simone Barbato developed the VR application.

Availability of data and material The data that support the findings of this study are available from the corresponding author upon request.

Declarations

Competing interests The authors declare that they have no conflict of interest. The authors have no affiliation with or involvement in any organization or entity with a direct or indirect financial interest or non-financial interest in the subject matter discussed in the manuscript.

References

Andersen, N. J., Schwartzman, D., Martinez, C., Cormier, G., & Drapeau, M. (2023) Virtual reality interventions for the treatment of anxiety disorders: A scoping review *Journal of Behavior Therapy and Experimental Psychiatry*, 101851.

APA, A. P. A. (2013). Diagnostic and statistical manual of mental disorders. textitThe American Psychiatric Association.

Apicella, A., Arpaia, P., Giugliano, S., Mastrati, G., & Moccaldi, N. (2022). High-wearable eeg-based transducer for engagement detection in pediatric rehabilitation. *Brain-Computer Interfaces, 9*(3), 129–139.

Apicella, A., Barbato, S., Chacón, L. A. B., D'Errico, G., De Paolis, L. T., Maffei, L., Massaro, P., Mastrati, G., Moccaldi, N., Pollastro, A., et al. (2023). Electroencephalography correlates of fear of heights in a virtual reality environment. *Acta IMEKO, 12*(2), 1–7.

Apicella, A., Isgrò, F., Pollastro, A., & Prevete, R. (2023). On the effects of data normalization for domain adaptation on eeg data. *Engineering Applications of Artificial Intelligence, 123*, 106205.

Apicella, A., Isgrò, F., & Prevete, R. (2019). A simple and efficient architecture for trainable activation functions. *Neurocomputing, 370*, 1–15.

Arpaia, P., Barbato, S., D'Errico, G., Mastrati, G., Moccaldi, N., Robbio, R., & Wriessenegger, S. C. (2023). Hrv-based detection of fear of heights in a vr environment. In: International Conference on Extended Reality. Springer, pp. 500–513.

Arpaia, P., Callegaro, L., Cultrera, A., Esposito, A., & Ortolano, M. (2021). Metrological characterization of a low-cost electroencephalograph for wearable neural interfaces in industry 4.0 applications. In: 2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0&IoT). IEEE, pp. 1–5.

Arpaia, P., Coyle, D., D'Errico, G., De Benedetto, E., De Paolis, L. T., du Bois, N., Grassini, S., Mastrati, G., Moccaldi, N., Vallefuoco, E. (2022). Virtual reality enhances eeg-based neurofeedback for emotional self-regulation. In: International Conference on Extended Reality. Springer, pp. 420–431.

Arpaia, P., De Bendetto, E., Esposito, A., Natalizio, A., Parvis, M., & Pesola, M. (2022). Comparing artifact removal techniques for daily-life electroencephalography with few channels. In: 2022 IEEE International Symposium on Medical Measurements and Applications (MeMeA). IEEE, pp. 1–6.

Arpaia, P., Callegaro, L., Cultrera, A., Esposito, A., & Ortolano, M. (2021). Metrological characterization of consumer-grade equipment for wearable brain-computer interfaces and extended reality. *IEEE Transactions on Instrumentation and Measurement, 71*, 1–9.

Aspiotis, V., Miltiadous, A., Kalafatakis, K., Tzimourta, K. D., Giannakeas, N., Tsipouras, M. G., Peschos, D., Glavas, E., & Tzallas, A. T. (2022). Assessing electroencephalography as a stress indicator: A vr high-altitude scenario monitored through eeg and ecg. *Sensors, 22*(15), 5792.

Association, A. P., et al.,(2015). Diagnostic and statistical manual of mental disorders fifth edition, vol. 656. https://doi.org/10.1176/appi.books.9780890425596.19.

Azimisefat, P., de Jongh, A., Rajabi, S., Kanske, P., & Jamshidi, F. (2022). Efficacy of virtual reality exposure therapy and eye movement desensitization and reprocessing therapy on symptoms of acrophobia and anxiety sensitivity in adolescent girls: A randomized controlled trial. *Frontiers in Psychology, 13*, 919148.

Bălan, O., Moise, G., Moldoveanu, A., Leordeanu, M., & Moldoveanu, F. (2020). An investigation of various machine and deep learning techniques applied in automatic fear level detection and acrophobia virtual therapy. *Sensors, 20*(2), 496.

Bejani, M. M. & Ghatee, M. (2021). A systematic review on overfitting control in shallow and deep neural networks. *Artificial Intelligence Review*, 1–48.

Bentley, J. L. (1975). Multidimensional binary search trees used for associative searching. *Communications of the ACM, 18*(9), 509–517.

Bishop, C. M., & Nasrabadi, N. M. (2006). Pattern recognition and machine learning. Springer, 4(4).

Boehnlein, J., Altegoer, L., Muck, N. K., Roesmann, K., Redlich, R., Dannlowski, U., & Leehr, E. J. (2020). Factors influencing the success of exposure therapy for specific phobia: A systematic review. *Neuroscience & Biobehavioral Reviews, 108*, 796–820.

Bornas, X., Llabres, J., Noguera, M., López, A. M., Gelabert, J. M., & Vila, I. (2006). Fear induced complexity loss in the electrocardiogram of flight phobics: a multiscale entropy analysis. *Biological Psychology, 73*(3), 272–279.

Botella, C., Fernández-Álvarez, J., Guillén, V., García-Palacios, A., & Baños, R. (2017). Recent progress in virtual reality exposure therapy for phobias: a systematic review. *Current Psychiatry Reports, 19*, 1–13.

Caseras, X., Mataix-Cols, D., Trasovares, M. V., López-Solà, M., Ortriz, H., Pujol, J., Soriano-Mas, C., Giampietro, V., Brammer, M. J., & Torrubia, R. (2010). Dynamics of brain responses to phobic-related stimulation in specific phobia subtypes. *European Journal of Neuroscience, 32*(8), 1414–1422.

Chang, C.-Y., Hsu, S.-H., Pion-Tonachini, L., & Jung, T.-P. (2018). Evaluation of artifact subspace reconstruction for automatic eeg artifact removal. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, pp. 1242–1245.

Choo, A., May, A. Virtual., & mindfulness meditation: (2014). Virtual reality and electroencephalography for health gamification. In. IEEE Games Media Entertainment. *IEEE, 2014*, 1–3.

Choy, Y., Fyer, A. J., & Lipsitz, J. D. (2007). Treatment of specific phobia in adults. *Clinical Psychology Review, 27*(3), 266–286.

Cohen, D. C. (1977). Comparison of self-report and overt-behavioral procedures for assessing acrophobia. *Behavior Therapy, 8*(1), 17–23.

Coomans, D., & Massart, D. L. (1982). Alternative k-nearest neighbour rules in supervised pattern recognition: Part 1 k-nearest neighbour classification by using alternative voting rules. *Analytica Chimica Acta, 136*, 15–27.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*, 273–297.

Daly, I., Billinger, M., Scherer, R., & Müller-Putz, G. (2013). On the automated removal of artifacts related to head movement from the eeg. *IEEE Transactions on neural systems and rehabilitation engineering, 21*(3), 427–434.

Ehinger, B. V., Fischer, P., Gert, A. L., Kaufhold, L., Weber, F., Pipa, G., & König, P. (2014). Kinesthetic and vestibular information modulate alpha activity during spatial navigation: a mobile eeg study. *Frontiers in Human Neuroscience, 8*, 71.

Essau, C. A., Conradt, J., & Petermann, F. (2000). Frequency, comorbidity, and psychosocial impairment of specific phobia in adolescents. *Journal of Clinical Child Psychology, 29*(2), 221–231.

Fdez, J., Guttenberg, N., Witkowski, O., & Pasquali, A. (2021). Cross-subject eeg-based emotion recognition through neural networks with stratified normalization. *Frontiers in Neuroscience, 15*, 626277.

Fernando, B., Habrard, A., Sebban, M., & Tuytelaars, T. (2013). Unsupervised visual domain adaptation using subspace alignment. In: Proceedings of the IEEE international conference on computer vision, pp. 2960–2967.

Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., & Lempitsky, V. (2016). Domain-adversarial training of neural networks. *The journal of Machine Learning Research, 17*(1), 2096–2030.

Guo, M., Zhong, Y., Xu, J., Zhang, G., Xu, A., Kong, J., Wang, Q., Hang, Y., Xie, Y., Wu, Z., et al. (2023). Altered brain function in patients with acrophobia: A voxel-wise degree centrality analysis. *Journal of Psychiatric Research, 164*, 59–65.

Hang, Y., Zhong, Y., Zhang, G., Wu, Z., Kong, J., Wang, Q., Ding, H., Wu, H., Yuan, S., Wang, C., et al. (2022). Altered spontaneous neural activity in frontal and visual regions in patients with acrophobia. *Journal of Affective Disorders, 303*, 340–345.

HassanPour Zonoozi, M., & Seydi, V. (2022). A survey on adversarial domain adaptation. *Neural Processing Letters*, 1–41.

Hecht-Nielsen, R. (1992). Theory of the backpropagation neural network. In: Neural networks for perception. Elsevier, pp. 65–93.

Ho, T. K. (1995). Random decision forests. In: Proceedings of 3rd international conference on document analysis and recognition, vol. 1. IEEE, pp. 278–282.

Hodges, L. F., Kooper, R., Meyer, T. C., Rothbaum, B. O., Opdyke, D., Graaff, J. J. D., Williford, J. S., & North, M. M. (1995). Virtual environments for treating the fear of heights. *IEEE computer, 28*(7), 27–34.

Huppert, D., Grill, E., & Brandt, T. (2017). A new questionnaire for estimating the severity of visual height intolerance and acrophobia by a metric interval scale. *Frontiers in Neurology, 8*, 211.

Ibitoye, R. T., Castro, P., Ellmers, T. J., Kaski, D. N., & Bronstein, A. M. (2023). Vestibular loss disrupts visual reactivity in the alpha eeg rhythm. *NeuroImage Clinical, 39*, 103469.

Idego - Digital Psychology. https://www.idego.it/. Accessed: 2022-12-06.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology, 3*(3), 203–220.

Keptner, K. M., Fitzgibbon, C., & O'Sullivan, J. (2021). Effectiveness of anxiety reduction interventions on test anxiety: a comparison of four techniques incorporating sensory modulation. *British Journal of Occupational Therapy, 84*(5), 289–297.

Knopf, K., & Pössel, P. (2009). Individual response differences in spider phobia: comparing phobic and non-phobic women of different reactivity levels. *Anxiety, Stress, & Coping, 22*(1), 39–55.

Kosunen,, I. Salminen,, M. Järvelä, S., Ruonala, A., Ravaja, N., Jacucci, G. (2016). Relaworld: neuroadaptive and immersive virtual reality meditation system. In: Proceedings of the 21st International Conference on Intelligent User Interfaces, pp. 208–217.

Lan, Z., Sourina, O., Wang, L., Scherer, R., & Müller-Putz, G. R. (2018). Domain adaptation techniques for eeg-based emotion recognition: A comparative study on two public datasets. *IEEE Transactions on Cognitive and Developmental Systems, 11*(1), 85–94.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444.

Lépine, J.-P., Gasquet, I., Kovess, V., Arbabzadeh-Bouchez, S., Negre-Pages, L., Nachbaur, G., & Gaudin, A. (2005). Prevalence and

comorbidity of psychiatric disorders in the french general population. *L'encephale, 31*(2), 182–194.

Liu, Z., Shore, J., Wang, M., Yuan, F., Buss, A., & Zhao, X. (2021). A systematic review on hybrid eeg/fnirs in brain-computer interface. *Biomedical Signal Processing and Control, 68*, 102595.

Liveamp. https://www.brainproducts.com/solutions/liveamp/. Accessed: 2022-11-17.

Magee, W. J., Eaton, W. W., Wittchen, H.-U., McGonagle, K. A., & Kessler, R. C. (1996). Agoraphobia, simple phobia, and social phobia in the national comorbidity survey. *Archives of General Psychiatry, 53*(2), 159–168.

Marks, I., & Gelder, M. (1965). A controlled retrospective study of behaviour therapy in phobic patients. *The British Journal of Psychiatry, 111*(476), 561–573.

McCabe, R. E. (2015). Subjective units of distress scale. *J Phobias Psychol Irrational Fear, 18*, 361.

Menzies, R. (1997). Height phobia. Phobias-A Handbook of Theory, Research and Treatment. Chichester, West Sussex, England: John Wiley, pp. 139–151.

Meta quest 2. https://www.meta.com/at/en/quest/products/quest-2/. Accessed: 2022-11-17.

Nair, V., & Hinton, G. E., (2010). Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th international conference on machine learning (ICML-10), pp. 807–814.

Neumann, N., Fullana, M. A., Radua, J., Brandt, T., Dieterich, M., & Lotze, M. (2023). Common neural correlates of vestibular stimulation and fear learning: an fmri meta-analysis. *Journal of Neurology, 270*(4), 1843–1856.

T. F. of The European Society of Cardiology et al. (1996). The north american society of pacing and electrophysiology (membership of the task force listed in the appendix), heart rate variability standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, 17, 354–381.

Omohundro, S. M. (1989). *Five balltree construction algorithms*. International Computer Science Institute Berkeley.

Pan, J., & Tompkins, W. J. (1985). A real-time qrs detection algorithm. *IEEE Transactions on Biomedical Engineering, 3*, 230–236.

Parsons, T. D., & Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(3), 250–261.

Powers, M. B., & Emmelkamp, P. M. (2008). Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of Anxiety Disorders, 22*(3), 561–569.

Radüntz, T., Scouten, J., Hochmuth, O., & Meffert, B. (2015). Eeg artifact elimination by extraction of ica-component features using image processing algorithms. *Journal Of Neuroscience Methods, 243*, 84–93.

Sedghamiz, H. (2014). Matlab implementation of pan tompkins ecg qrs detector. *Code Available at the File Exchange Site of MathWorks*.

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. Frontiers in public health, 258.

Shen, Y.-W., & Lin, Y.-P. (2019). Challenge for affective brain-computer interfaces: Non-stationary spatio-spectral eeg oscillations of emotional responses. *Frontiers in Human Neuroscience, 13*, 366.

Steyer, R., Ferring, D., & Schmitt, M. J. (1992). States and traits in psychological assessment. *European Journal of Psychological Assessment*.

Straube, T., Mentzel, H.-J., & Miltner, W. H. (2006). Neural mechanisms of automatic and direct processing of phobogenic stimuli in specific phobia. *Biological psychiatry, 59*(2), 162–170.

Thoma, N., Pilecki, B., & McKay, D. (2015). Contemporary cognitive behavior therapy: A review of theory, history, and evidence. *Psychodynamic Psychiatry, 43*(3), 423–461.

Tychkov, A. Y., Chernyshov, D., Bofonova, N., Churakov, P., Yuldashev, Z., Gorbunov, V., Zolotarev, R., & Alimuradov, A. (2023). Spectral analysis of the eeg of subjects with anxious-phobic disorders in a virtual reality environment. *Biomedical Engineering, 56*, 429–434.

Uchitel, J., Vidal-Rosas, E. E., Cooper, R. J., & Zhao, H. (2021). Wearable, integrated eeg-fnirs technologies: A review. *Sensors, 21*(18), 6106.

Vert, J.-P., Tsuda, K., & Schölkopf, B. (2004). A primer on kernel methods. *Kernel Methods in Computational Biology, 47*, 35–70.

Wang, Q., Wang, H., Hu, F., Hua, C., & Wang, D. (2021). Using convolutional neural networks to decode eeg-based functional brain network with different severity of acrophobia. *Journal of Neural Engineering, 18*(1), 016007.

Wardenaar, K. J., Lim, C. C., Al-Hamzawi, A. O., Alonso, J., Andrade, L. H., C. d. Benjet, B. Bunting, G. De Girolamo, K. Demyttenaere, S. E. Florescu, et al. (2017). The cross-national epidemiology of specific phobia in the world mental health surveys. *Psychological Medicine, 47*(10), 1744–1760.

Wolitzky-Taylor, K. B., Horowitz, J. D., Powers, M. B., & Telch, M. J. (2008). Psychological approaches in the treatment of specific phobias: A meta-analysis. *Clinical Psychology Review, 28*(6), 1021–1037.

Yoshimura, N., Koga, O., Katsui, Y., Ogata, Y., Kambara, H., & Koike, Y. (2017). Decoding of emotional responses to user-unfriendly computer interfaces via electroencephalography signals. *Acta IMEKO, 6*(2), 93–98.

Andrea Apicella received the M.Sc. degree in Computer Science and the Ph.D. degree in Mathematics and Computer Science from University of Naples Federico II, Italy, in 2014 and 2019, respectively. He is currently assistant professor in the Department of Information Technology and Electrical Engineering of University of Naples Federico II, and he is a research associate of the ARHeMLab (Augmented Reality for Health Monitoring Laboratory), and AIPA Lab (Laboratory of Artificial Intelligence, Privacy & Applications) sited in Naples. The current research topics of Andrea Apicella cover design, testing and development of Trustworthy AI and eXplainable Artificial Intelligence (XAI) approaches for explaining the AI system's decisions, together with EEG signal processing for Emotion Recognition and Attention Monitoring using Artificial Intelligence methods.

Pasquale Arpaia (Senior Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from the University of Naples Federico II (Unina), Naples, Italy, in 1987 and 1992, respectively. He was an Associate Professor with the University of Sannio, Benevento, Italy. He is currently a Full Professor of instrumentation and measurements with Unina and a Team Leader with CERN, Geneva, Switzerland. He is also the Head of the Interdepartmental Center for Research on Management and Innovation of Health (CIRMIS), the Head of the Hi-Tech Academic FabLab, Department of Electrical Engineering and Information Technology (DIETI), Unina, and the Chairperson of the Stage Project of Unina. He is also the Head of the Instrumentation and Measurement for Particle Accelerators Laboratory (IMPALab) and the Augmented Reality for Health Monitoring Laboratory (ARHeMlab). His current research interests include augmented reality, brain-computer interfaces, cyber-security, digital instrumentation and measurement techniques, evolutionary diagnostics, and distributed measurement systems.

Simone Barbato is the co-founder of IDEGO Digital Psychology. He is a Clinical Psychologist registered with the Order of Psychologists of Lazio. He serves as a contract lecturer at the University of Cassino and Southern Lazio and at Niccoló Cusano University. He is the author of the book "Windows on the Future: Foundations of Digital Psychology," published by Alpes Italia.

Giovanni D'Errico received the M.S. degree in Computer Science Engineering in University of Salento. He is currently a PhD student at the Polytechnic University of Turin (doctoral school in Metrology). His current research interests include the study of Passive Brain Computer Interfaces (BCI) and eXtended Reality (XR) solutions to foster the EEG-based measurement of mental states. His past research interests include the study of XR applications and gamification in the context of smart cities, cultural heritage smart fruition and neurocognitive rehabilitation. He is currently part of AVR Lab (University of Salento) and ARHeMLab (University of Naples Federico II).

Giovanna Mastrati received the M.S. Degree (cum laude) in Biomedical Engineering in 2018, followed by a Ph.D. in Information and Communication Technology for Health in 2024, both from the University of Naples Federico II. Throughout her doctoral studies, she has been a researcher at the ARHeMLab (Laboratory of Augmented Reality for Health Monitoring) of DIETI Excellence Department, University of Naples Federico II, Naples. Her research interests centered around biomedical instrumentation and measurement, with a particular emphasis on Brain-Computer Interfaces (BCIs). She specialized in electroencephalographic (EEG) data acquisition and processing for the assessment of Emotions and Engagement, representing her primary areas of investigation.

Nicola Moccaldi (Member, IEEE) received the M.A. degree in communication science from the University of Salerno, Fisciano, Italy, in 1999, and the M.Sc. degree in electronic engineering and the Ph.D. degree in information technologies and electrical engineering from the University of Naples Federico II (UNINA), Napoli, Italy, in 2018 and 2021, respectively. He is currently an assistant professor (RTDB) and the Scientific Coordinator of Passive BCI Research Line for the ARHeMLab of DIETI Excellence Department (Laboratory of Augmented Reality for Health Monitoring) - Federico II University of Naples. Dr. Moccaldi is a member of the IEEE Instrumentation and Measurement Society and the BCI Society.

Ersilia Vallefuoco received the M.Sc. degree in biomedical engineering and the Ph.D. degree in information technology and electrical engineering from the University of Naples Federico II, in 2016 and 2020, respectively. She is currently a Research Fellow with the Augmented Reality for Health Monitoring Laboratory (ARHeMLab), DIETI Excellence Department, University of Naples Federico II. She was a speaker at national and international conferences. She is the author and coauthor of several publications. Her main research interests include healthcare, technologies for people with disabilities, serious games, and extended reality applications for healthcare.

Selina Christin Wriessnegger is Associate Professor and Deputy Head at the Institute of Neural Engineering (BCI-Lab), Graz University of Technology, Austria. From 2001 to 2005 she was PhD student at the Max-Planck-Institute for Human Cognitive and Brain Sciences and received her PhD from the Ludwig-Maximilians University (Munich). During that time, she spent one year in Rome as research assistant at IRCCS (Fondazione Santa Lucia), Laboratory for Human Psychophysiology. From 2005 to 2008 she was University Assistant at the Karl-Franzens-University Graz, Section Neuropsychology. From 2009 until May 2016 she was senior researcher at the Institute of Neural Engineering (BCI-Lab). From 2016 until 2020 she was Assistant Professor at the Institute of Neural Engineering and her habilitation was in 2020 at TUG. In 2017 she was visiting professor at SISSA (Scuola Internazionale Superiore di Studi Avanzati), Trieste and in 2019 she was guest professor at the University of Padova. She is Associate Editor of Frontiers in Neuroscience, special section Brain-Computer Interfaces and Topic Editor "Women in Brain-Computer Interfaces and Beyond". Since 2011 she is Co-organizer of the International Graz BCI Conference, since 2022 member of the IEEE MetroXRAINE steering committee and board member of the The Styrian Brain Research Initiative (INGE St). She has authored more than 120 peer reviewed publications which were cited more than 2500 times. Her research interests are, neural correlates of covert actions, novel applications of BCIs for healthy users, affective computing, passive BCIs, neuroadaptive systems, VR/AR applications in learning and rehabilitation.