Hindawi
International Transactions on Electrical Energy Systems
Volume 2023, Article ID 9797863, 1 page
https://doi.org/10.1155/2023/9797863

Retraction

Retracted: Multimodal Opera Performance Form Based on Human-Computer Interaction Technology

International Transactions on Electrical Energy Systems

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Wu, "Multimodal Opera Performance Form Based on Human-Computer Interaction Technology," *International Transactions on Electrical Energy Systems*, vol. 2022, Article ID 4003245, 13 pages, 2022.

Hindawi
International Transactions on Electrical Energy Systems
Volume 2022, Article ID 4003245, 13 pages
https://doi.org/10.1155/2022/4003245

Research Article

Multimodal Opera Performance Form Based on Human-Computer Interaction Technology

Lin Wu

College of Music Education, Xi'an Conservatory of Music, Xi'an 710000, Shaanxi, China

Correspondence should be addressed to Lin Wu; 01074@xacom.edu.cn

Received 30 July 2022; Revised 31 August 2022; Accepted 8 September 2022; Published 8 October 2022

Academic Editor: Raghavan Dhanasekaran

"Audience Engagement (AE)" describes how a stage performance affects the audience's thoughts, provokes a bodily response, and spurs cognitive growth. With little audience involvement, theatre performing arts like opera typically have difficulty keeping audiences' attention. The brain-computer interaction (BCI) technology could be used in opera performances to alter the audience's emotional experience. Nevertheless, for such BCI systems to function, they must accurately identify their participants' present emotional states. Although difficult to evaluate, audience participation is a vital sign of how well an opera performs. Practical methodological approaches for real-time perception and comprehension of audience emotions include psychological and physiological assessments. Hence, a multimodal emotional state detection technique (MESDT) for enhancing the AE in opera performance using BCI has been proposed. Three essential steps make up a conceptual MESDT architecture. An electroencephalogram (EEG) and other biological signs from the audience are first captured. Second, the acquired signals are processed, and the BCI tries to determine the user's present psychological response. Third, an adaptive performance stimulus (APS) is triggered to enhance AE in opera performance, as determined by a rule base. To give the opera audience a high-quality viewing experience, the immersive theatre performance has been simulated. Fifty individuals have been used in the experimental assessment and performance studies. The findings demonstrated that the proposed technology had been able to accurately identify the decline in AE and that performing stimuli had a good impact on enhancing AE during an opera performance. It has been further shown that the suggested design improves the overall performance of AE by 5.8% when compared to a typical BCI design (one that uses EEG characteristics solely) for the proposed MESDT framework with BCI.

1. Introduction to Brain-Computer Interaction Technology

In the past 15 years, the domains of cognitive prosthetic systems and brain-computer interfaces (BCIs) have undergone incredible advancement. They have combined theories and techniques from electronic content and the arts with those from signal analysis, deep learning (reinforcement learning), parallel computing intelligence, cognitive science, statistics, and linear algebra [1, 2]. As a result, the multidisciplinary implications of BCIs cover a variety of fields, including recreation and the arts, sustainable population goals in the workplace, and fitness goals for individuals [3, 4]. The latter, therefore, are widely perceived with suspicion. Concepts like the reality that the artistic side cannot be studied scientifically or that art is essential for spreading scientific beliefs but should not be employed as a systematic approach to searching for scientific proof have traditionally been extended.

Lately, the multiple perspectives and novel initiatives that resulted from the cross-fertilization of many academic fields have led to the application of scientific ideas and methodologies in the artistic process and creative approaches in scientific research. Artists are amongst the forerunners in developing BCI applications, pushing the envelope for applications in realistic settings. Alvin Lucier gave the symphony for single-player performance in 1965, regarded as the first play utilizing the EEG technique, one year after the initial demonstration of an EEG-based BCI [5, 6]. Soon after, many musicians, creators, and performances appeared. A growing number of cross-disciplinary practices, including online games, interactive installations, and appearances using these interactions, have been developed. They have combined scientific and creative methodologies through new research, technological developments, and limited commercial wireless technologies since 2010.

Another essential component of characterization in human-computer interaction (HCI) is emotion detection. Embracing HCI is now possible because of the advancement of mobile, noninvasive sensing technologies like BCI. A BCI is a technology that, without external nerves and muscles, transforms signals produced by activity in the brain into commands for remote systems [7]. A research initiative in general communication led to the development of effective BCI (BCI) [8]. The study aimed to develop neurobiological devices that could recognize emotional state signs and then use that data to enable HCI. BCI research seeks to improve interactions between people and computers by better modeling physiological responses, synthesizing emotional responses and behavior, and sensing emotional states.

Emotion has a significant impact on social interaction since it includes psychotic symptoms to both external and internal stimuli and physical responses to those emotional responses in daily life. There are different levels and modalities at which emotional responses might be seen. On the one hand, ancillary inputs have been connected to the bodily nerve system and demonstrate physiological changes in emotional states. For instance, physical cues, including body language, vocal discourse, and facial emotions, might be observed [9]. On the other hand, several other factors might affect mental processes, such as coping mechanisms, including wild speculation, despair, or transferring responsibility. The study's objective is to detect emotions via a multimodal integration of external physiological inputs and brain signals like EEGs, referred to as MESDT. Dancing, singing, opera, drama, playing instruments, hypnosis, magic, mime, puppeteer, and circus arts are performing arts. This type of artwork is one in which the creators present their art to the audience live. Opera is an artistic genre (performing arts) that uses singing and music to convey a storyline. Opera singers do not utilize microphones to augment their vocals, and the orchestra plays all of the music live, unlike in a musical performance. The AE in opera performance involves brain signals like EEG and other physiological signals. So, in this work, MESDT for AE has been integrated into BCI.

In terms of model design, neural signal conditioning methods and applications have made significant advancements along with BCI developments. These BCI systems still have specific difficulties, though. On the one hand, as different noises influence emotional data, it is challenging to accurately depict emotional states using a single modality. However, specific modalities are simple to conceal and challenging to portray the actual emotional state. For example, recognizing an appropriate emotion may not always depict a person's natural emotional state since emotion can be concealed. The fundamental goal of this project is to develop multimodal emotion detection for AE in BCI systems. Building an emotional characterization model and providing a statistically accurate representation of the emotional state are two of the most significant issues in emotion detection research. A mathematical formulation of an emotional state is created by dynamic modeling so that a BCI system can classify or quantify it. Because it enables us to judge emotional states more accurately, developing an emotion model is crucial to measuring emotions.

The main contributions of this article are as follows:

(1) To design the MESDT framework for opera performance based on BCI
(2) To detect the multimodal emotional state of the audience for an opera performance by analyzing EEG and physiological signals
(3) To determine the audience's present emotional state and develop a framework that determines the decrease in AE and activates APS to enhance their engagement in the opera performance
(4) The immersive theatre performance has been simulated to give a high-quality viewing experience and perform analysis for the AE during an opera performance

The rest of the article has been organized as follows: Section 2 describes related research on BCI for enhancing AE. Section 3 gives a multimodal emotional state detection technique (MESDT) for strengthening audience engagement in opera performances using BCI which has been proposed. Simulation results and discussion are given in Section 4. Finally, the conclusion and scope for further research are shown in Section 5.

2. Related Works on BCI for Enhancing Audience Engagement

Several scholars have suggested methods for representing emotions. The six fundamental emotions are sorrow, fear, hatred, amazement, pleasure, and anger. These six basic emotions can be joined to create more complicated emotional classes [10]. Nevertheless, neither this explanation nor a device's ability to assess an emotional state from a technological standpoint can scientifically characterize the meaning of emotion. The valence-based stimulus, a two-dimensional emotion paradigm put out by Russell in 1980, has often been employed in prior research. Stimulation and valence are the two aspects the model uses to categorize emotions. The valence measurement axis's lower half denotes negative emotion, whereas its positive half denotes happy feelings [11]. The main distinction between this paradigm and the discontinuous emotion paradigm is that the multidimensional emotion model is continuous, giving it the benefit of being able to connect with the audience across a broad range and the ability to be used to describe the development of emotion.

Numerous researchers have looked at the use of biomedical signals to identify and analyze user characteristics when interacting [12]. Particularly in the context of video games, equipment employed to obtain bio and neural feedback has gained popularity. EEG devices are being used in the context of HCI because of their transparency and capacity to transparently record a user's interaction beyond his conscious and controlled activities [13]. It is possible to consider their application in domains like music listening and movie watching, now that noninvasive commercial electroencephalography (EEG) equipment is more widely accessible in the market. EEG equipment positions electrodes on specific areas of the scalp to monitor changes in electrical charge as neurons in the brain's cerebral cortex respond [14].

To understand a person's mental abilities, such as attention/engagement and calmness, the gathered signals are separated into five distinct frequency bands [15]. From the perspective of communication, BCI systems can be applied passively by observing human neural activity to identify cognitive processes. These mental processes are used as an insight into the proposal [16] by permitting users to manage a system through cognizant brain processing or by understanding the user's psychological condition as a response to the obtained stimulus. In this situation, the automatic system adaptation might be controlled by how the user's mental state is perceived [17].

The BCI system requires an approach like this, using implicit input such as "engaging" and "interest" signs to tailor each visitor's experience during a museum visit. These details about the user may be used to create a user profile, which can subsequently be used to make suggestions. Museums have recently focused on giving tailored services through their websites and on-site personalized guides and descriptions of artifacts [18]. More museums are using personalized museum interpreters to improve the tourist experience, draw in more people, and cater to the requirements of a wide range of visitors [19].

A thorough overview of the topic of customized applications in cultural heritage is provided by Pavlidis in their article [20]. With this method, the BCI can track the user's journey while touring the exhibition in real-time and provide input that can be utilized to customize their visit. Several projects have followed this strategy with effectiveness. Neuro-controlled gameplay that allows teams to manage a simulated quadcopter with their brain waves has been proposed by Tezza et al. [21]. The game also uses a BCI to gauge player interest. Recently, authors in [22, 23] demonstrated the benefits of the user experience by measuring and analyzing AE at the EEG data level during a three-dimensional simulated theatre performance.

In conclusion, no BCI approach has been applied to enhance the performance of art design from the viewer's point of view. Additionally, there are not too many virtual environment plays that let the theatre architect alter the acting signals and assess the real impact in real-time. Depending on the user's level of engagement as evaluated in real-time, an immersive opera theatre (virtual environment) has been built to produce adaptive content relevant to performing signals to increase AE under various scenarios.

3. Proposed Multimodal Emotional State Detection Technique (MESDT) for Enhancing the AE in Opera Performance Using BCI

Emotional responses are challenging to record because different forms of interference can easily influence single-modality data. Compared to the ideal single-modality correspondent system, the most delicate multimodal emotion-recognition scheme has been highly accurate, reaching 85 percent. A framework that precisely depicts the possible nature of human emotion may be built by thoroughly investigating various biological signals and their interactions.

Figure 1 depicts the movement of information in the proposed MESDT using BCI for opera performance. The devices employed in the data acquisition framework capture the modality information from cumulative neural activity and behavior modalities (such as expression from the face and motion of the eye) or different blended neurophysiological modalities of the body during interplay with other clusters or by a unique audio-visual stimulus.

Getting clean EEG readings with emotional cues is essential for signal processing. The preprocessing stage for neurophysiological signals involves denoising and eliminating artifacts from the gathered raw signals. When it comes to picture signals like gestures, unnecessary data must be removed, and data indicating psychological response are then recovered and improved. Various merging approaches can be employed at the modalities merging stage, including information merging, attribute merging, and choice merging. Emotional state output in choice-making can be done using deep learning or machine learning as the ultimate classifier for choice-making.

Multimodal techniques have produced appropriate emotional states. Suppose all that is required is emotional detection. In that case, no emotional feedback intervention is essential because the feedback will be sent through an appropriate interface (such as enhancing AE during opera performance). However, to constantly change the behavior of the audience's interaction, the participant's state information must be integrated explicitly into the closed-loop system if the activity entails AE enhancement through the initialization of performance cues.

Thus, three main steps have been used in the conceptual MESDT architecture. In the first step, the audience's EEG and other biological signals must be captured. In the second step, the acquired signals have to be processed, and the BCI determines the present psychological response of the audience. If the current response of BCI indicates reduced audience interest, an APS has been triggered to enhance AE in opera performance, which is the third step.

3.1. Acquisition of Multimodal Attributes from the Audience of Opera. In the first step, the EEG and other biological signals were acquired from the audience during the opera performance. Investigations have been carried out as part of broader research that aimed to create and test the MESDT, employing BCI to train and test the suggested system. The opera audience in this research has been obliged to visit five

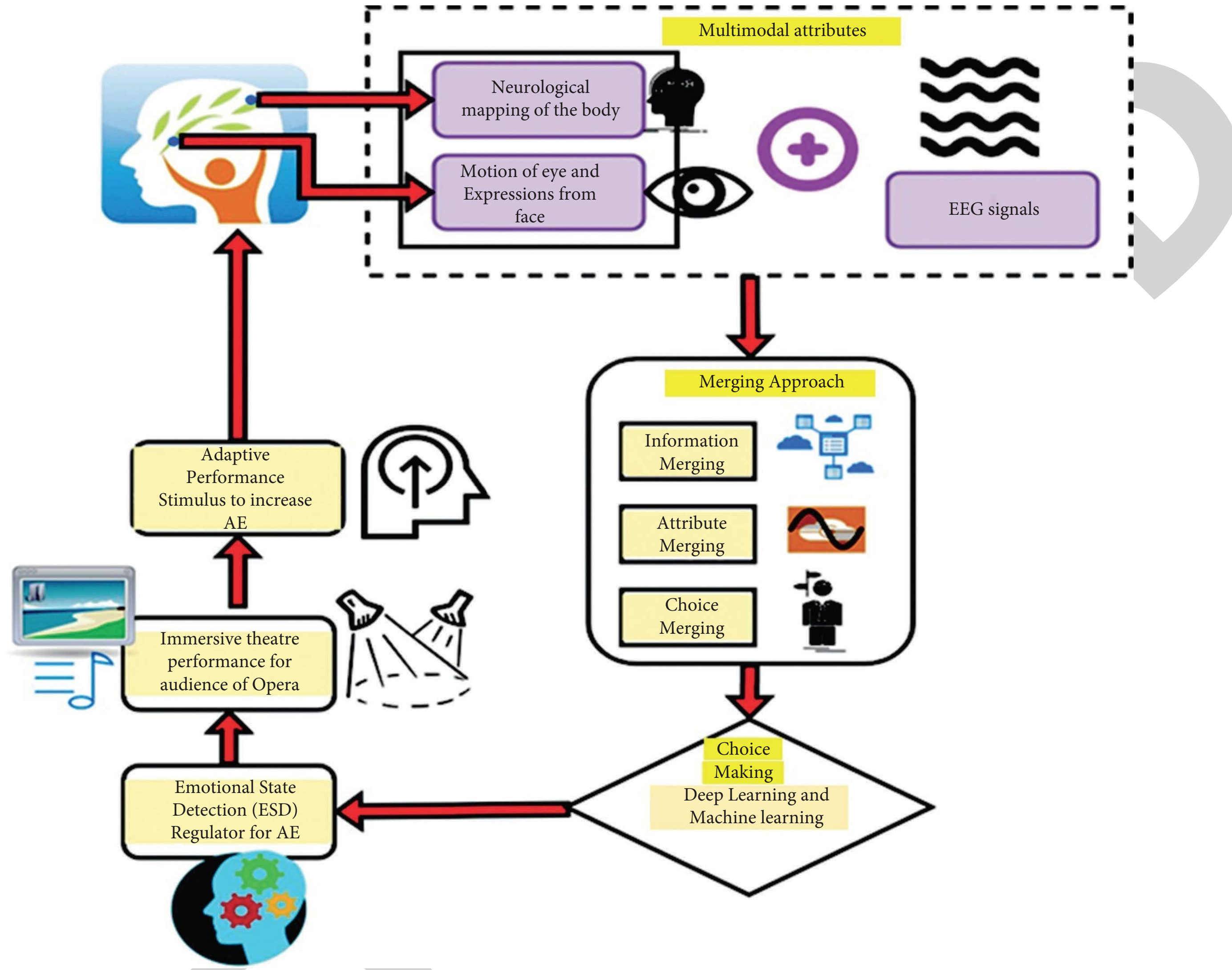

Figure 1: Movement of information in the proposed MESDT using BCI for opera performance.

opera sessions/terms in total, which have been spread over the course of four months.

3.1.1. Experiment. In the trial, there have been three different kinds of sessions referred to as "terms": one for tuning, three for training, and one for assessment. Figure 2 shows this in more detail.

Figure 2 depicts the organization of terms and rounds within the experiment. In each round, segments of opera will be played for 5 minutes. The terms are placed in order on different days (with an interval of 24 days between sessions), and within each session, rounds happen consecutively with gaps of 2 minutes in between, based on the experimental trial. The rounds that followed were used to discover courses for changing emotional states. In contrast, the tuning session has been utilized to uncover biological and psychological aspects of emotional reactions to opera performance. The framework of the training term was nearly identical. The opera was played for 5 minutes, and the music, light, and backdrop have been created to elicit two different emotional states in the audience. The first 2.5 minutes tried to produce one emotional state, while the second 2.5 minutes tried to elicit another emotional state. The training session was performed for each audience three times on different days.

One of the two different emotional states has been the focus of each round during the tuning session. A total of 10 instances of music were played throughout each tuning session to cause each audience to experience one of the two different emotional states. The identical set of two discrete emotional states utilized in the tuning round has been used as the beginning and final emotional states in each trial throughout the training sessions. The testing session evaluated the MESDT employing the BCI system while it was utilized through a live opera performance.

3.1.2. Acquiring the Multimodal Signals. The 32 EEG channels that have been placed follow the International 10/20 standard and refer to two electrodes placed at AFz and FCz. These electrodes have been used to capture EEG using a BrainAmp EEG processor (BrainProducts, Germany).

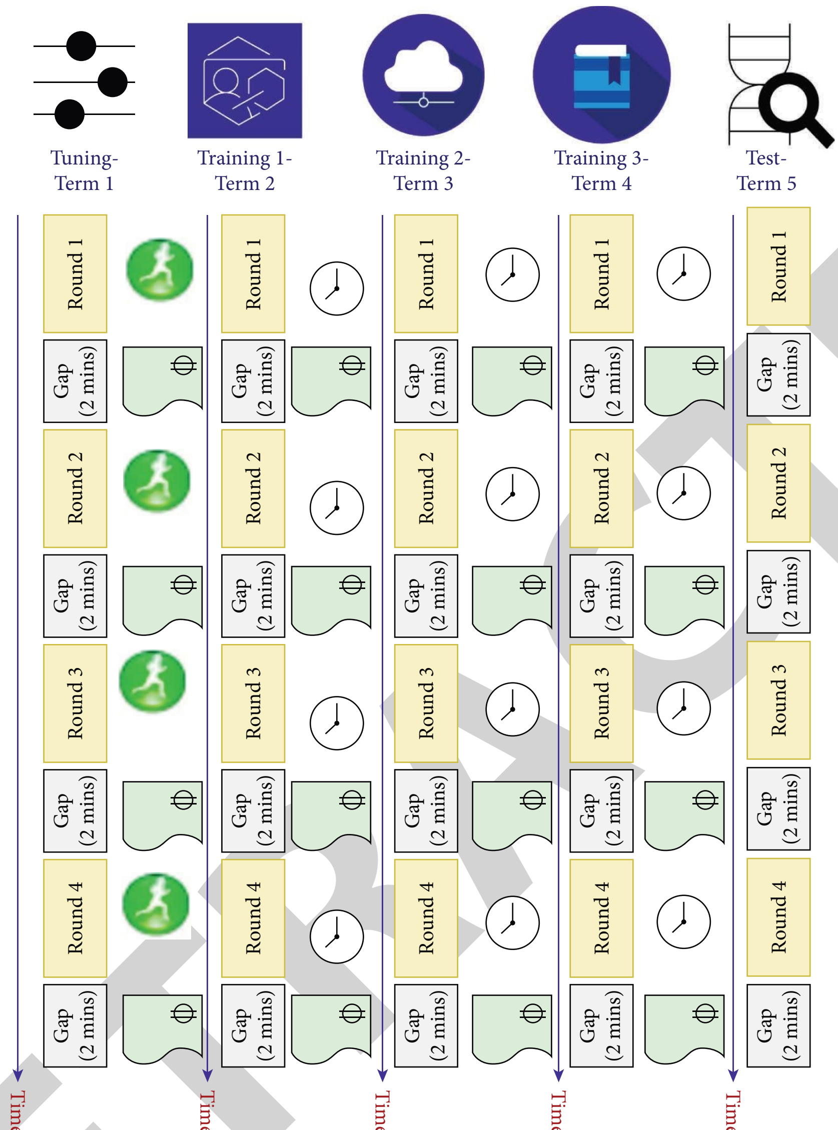

FIGURE 2: Organization of terms and rounds within the experiment.

Figure 3 shows the precise EEG channels that have been utilized.

Figure 3 depicts acquiring EEG recordings through electrodes, emotion detection, and enhancing AE through an immersive theatre experience. Due to their great relevance to eye movements and positive awakenings, two electrodes (AFz and FCz) have been utilized for assessing AE for this study. The proposed system integrated EEG into the suggested approach to identify when to use performing cues. The International 10–20 system has spatially arranged the device's 12 electrode detectors and two bipolar reference electrodes. It gives access to a laptop via Bluetooth and a USB adapter. It is comfortable and supportive enough for the audience to utilize in a makeshift theatre setting without the need for special knowledge. Once the algorithm identifies the audience's lack of interest, performance cues connected to the material (opera performance) should be summoned to recover their attention quickly.

3.2. Signal Processing and MESD of Audience

3.2.1. Feature Selection. Events have been divided into 5 minutes duration and nonoverlapping parts to create a MESD with high spatial and temporal resolution. During each of these 5-minute duration subtrials, average characteristics and audience reports of their perceived emotional states have been computed. An autonomous attribute selection approach based on Eigen decomposition [25] has been used to choose a subset of attributes for the classification algorithm. Attributes have been deconstructed into

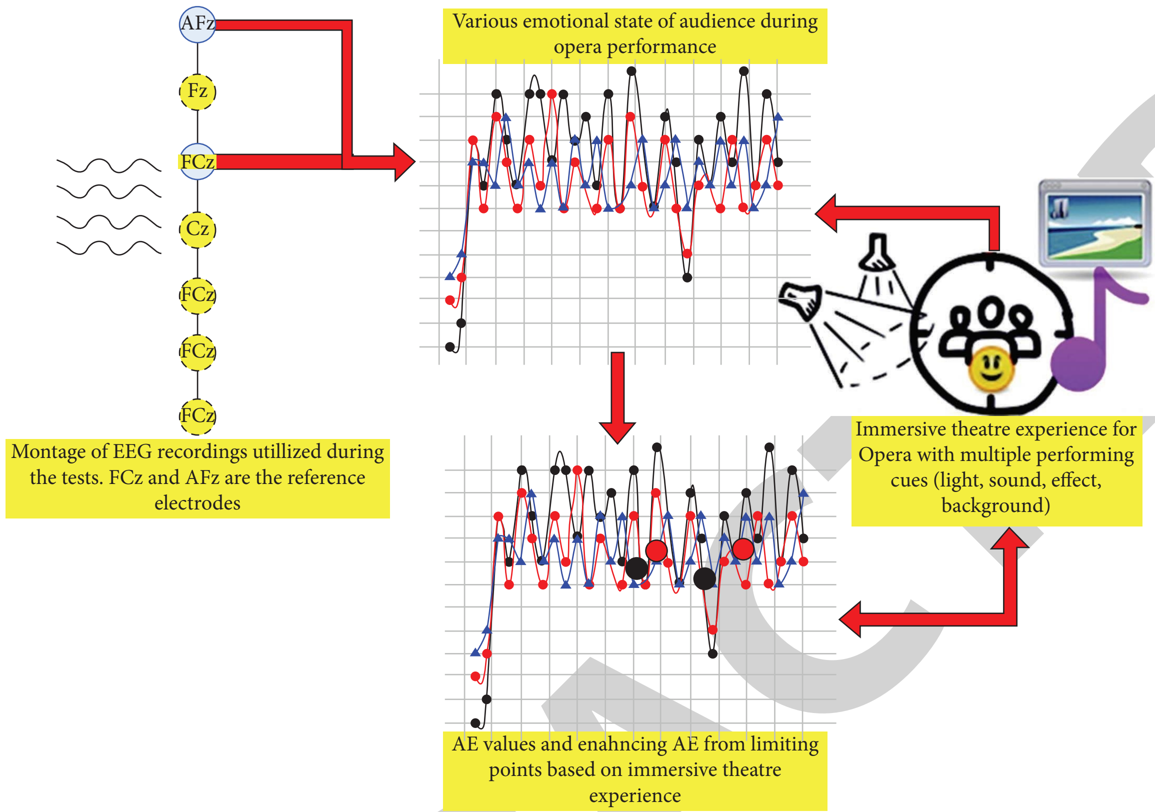

FIGURE 3: Acquiring EEG recordings through electrodes, emotion detection, and enhancing AE through immersive theatre experience (EEG montage image has been acquired from [24]).

Eigen vectors after being z-scored. The subsets of characteristics that grouped with the labels (the audience's assessments of their perceived emotional states) were then chosen using an adjusted harmonic clustering technique [26]. Collecting prospective features and membership functions have been subjected to an autonomous approach (modal clustering). The characteristics gathered into the cluster that has been found to have the membership functions were preserved. The other attributes that belonged to this cluster have been chosen as the features of interest.

3.2.2. Merging Approach. Since multimodal signals have been extracted, the obtained attributes from these signals have to be merged for effective decision-making on AE. In this work, two merging approaches have been employed. The two classifiers were given equal weights during the first method (i.e., the equal weighting method), and the Bayesian merger method [27] was utilized in the second approach. The entire dependability of each source (EEG and other neural signals) for each class is represented by Bayesian merging using the normalized matrix for each signal (eye or body signals). The merged output has been calculated as shown in the following equation:fd1

$$M_{op} = \operatorname{argmin}_{m_{ip}} \left[\prod_{cf} \operatorname{Prob}\left(M = m_{ip} | C_l = c_l\right) \right], \quad c_l \in m_1, m_1, \ldots, m_n. \tag{1}$$

The merged output has been denoted as M_{op}. m_{ip} is the input to the merger. C_l is the class identifier or label based on the classifier l. cf is the total number of classifiers. n is the total number of inputs given to the merger. The probability of the merger output has been provided by Bayes' rule as shown in the following equation:fd2

$$\operatorname{Prob}\left(M = m_{ip} | C_l = c_l\right) \propto \operatorname{Prob}(M) * \operatorname{Prob}\left(C_l = c_l | M = m_{ip}\right). \tag{2}$$

It has been identified that $\operatorname{Prob}(M)$ denotes the probability of merging has identical values for all the class identifiers. $\operatorname{Prob}(C_l = c_l | M = m_{ip})$ can be assessed from the information available for training.

3.2.3. Classification and Choice Making. The multimodal signals have been acquired, the data have been merged, and based on the unified data available from the previous sections, choice-making has to be done to predict the AE. There is a need for a boosting strategy in addition to executing the commonly used merging approach based on listing various scores among two methods through Bayesian classification. To train the AdaBoost algorithm's [28] weights, they used both classifications (facial expressions like eye movement and EEG) as subclassifiers. The following equations have been used to determine the outcomes:

$$V^{\text{boost}} = \frac{1}{\left\{1 - \text{expo}\left(\sum_{i=1}^{p} w^i v^i\right)\right\}}, \tag{3}$$

$$\text{Prediction}^{\text{boost}} = \begin{cases} \text{maximum}, \dfrac{1}{\left\{1 - \text{expo}\left(\sum_{i=1}^{p} w^i v^i\right)\right\}} > 0.5, \\ \text{average}, \dfrac{1}{\left\{1 - \text{expo}\left(\sum_{i=1}^{p} w^i v^i\right)\right\}} = 0.5, \\ \text{minimum}, \dfrac{1}{\left\{1 - \text{expo}\left(\sum_{i=1}^{p} w^i v^i\right)\right\}} < 0.5. \end{cases} \tag{4}$$

V^{boost} denotes the value of the AdaBoost algorithm. The results of the two classifiers' ratings have been combined to get the overall emotion score. $\text{Prediction}^{\text{boost}}$ denotes the predicted outcomes of the merged classifier. w^i is the weight given to the parameters of the algorithm. v^i is the outcome of i^{th} subclassifier ($v^i \in [-1, 1]$), and p denotes the total number of subclassifiers. The weight w^i ($i = 1, 2, \ldots, p$) from the training set denoted $t(a) j \in [-1, 1]$ can be the outcome of i^{th} subclassifier for the assessment sample of a.

A pipelined and unified classification has been integrated into a two-level classification method for emotion detection to decrease detection inaccuracy and false positive rate and expedite signal processing. An artificial neural network (ANN) [29] has been employed in the second layer to verify emotions. In contrast, the poor classifier's Haar-like characteristic cascades have been utilized to recognize facial objects in the first stage. Then, principal component analysis (PCA) has been used to characterize the facial feature since it retained the most power while using the most miniature primary components. Support vector machines (SVM) have been used as a classifier for each modality. The two (EEG and biological) modalities have been integrated, and the result is the average rating from both modalities. The mean value of multimodal classification has been obtained fromfd5

$$b_{\text{mean}} = \left(\frac{1}{X}\right) * \sum_{i=1}^{X} \left\{ \frac{1}{Y} \sum_{j=1}^{Y} (b(j, i)) \right\}. \tag{5}$$

b_{mean} is the mean value of multimodal classification. X denotes the number of multimodal entities from the EEG and biological signals. Y is the number of test samples considered for each modality. $b(j, i)$ denotes the outcome of the MESD system for opera performance, considering j^{th} samples from the i^{th} modality. The value $b(j, i)$ lies between 0 and 1.

3.2.4. Measuring AE from Acquired Signals. According to considerable research, EEG waves can reveal information about a participant's emotional state, including enthusiasm, mindfulness, enjoyment, and aggravation. The psychological conditions of engagement, attentiveness, and working mode and the impression of user-machine mistakes are all detectable by EEG measurements as shown in the following equation:fd6

$$\text{EEG}_{AE} = \frac{\beta}{(\alpha + \theta)}. \tag{6}$$

Depending on the magnitude of the alpha α, theta θ, and beta β waves, the formula above is frequently used to assess AE in opera performance. The EEG determined engagement is a proportion estimate and is unitless.

Commercial devices are now considerably more practical because they are readily accessible and inexpensive wireless EEG headsets. The headset poses no threat to the audience, can be used in any location, and does not need special training or experience. Additionally, EEG has been utilized to accurately classify different cognitive activities. The proposed MESDT applies to an immersive theatre performance of opera, which creates an embodied narrative agent that uses movements and variable speech intensity to enhance AE when it notices a drop in interest. Although the algorithm has been able to recognize significant reductions in AE and restore it, it is not clear how distinct changes in behavioral signals influence audiences. The noninvasive, inert BCI strategy is the backbone of the proposed MESDT. It uses it to track AE and improve user experience by designing an experience that reacts to the participant's psychological process regulated by multimodal signals.

The audience of an opera put on an EEG headset and sensors, which record the multimodal signals throughout the tests. The degrees of AE were then determined by analyzing the collected multimodal data. The proposed MESDT detects declines in AE and begins executing stimuli to bring values back. Significant performance elements and cues have been realized through immersive theatre in a virtual world. The simulated performance's output has been rendered and shown on an immersive 3D circumferential display.

To eliminate noise and neuromuscular artifacts (such as gaze, eye blinking, and brain activity), the proposed framework filters the signal using standard mode rejections, electronic spike screening at 45 Hz and 55 Hz, and other methods. Fast Fourier transforms (FFT) are used to divide frequencies between 0.5 and 45 Hz into α, θ, and β waveforms, and the equipment sweeps at a rate of 130 Hz. The headgear has been used to measure the EEG values for different frequencies α, θ, and β.

Figure 4 depicts the enhancement of AE for opera performance based on MESDT. It is a graph between the limiting values of AE and the time duration of opera performance. The curves include multimodal signals consisting of mean AE value, actual AE value, and smoothed AE value. It consists of the following two points: points to indicate that the AE has dropped below the limiting value and points at which APS has been triggered to enhance the AE.

According to Figure 4, the AE values $AE(j, i)$ have been calculated by taking the mean value of the information from two electrodes (AFz and FCz) using the proposed MESDT algorithm for AE in the opera mentioned above. Using the

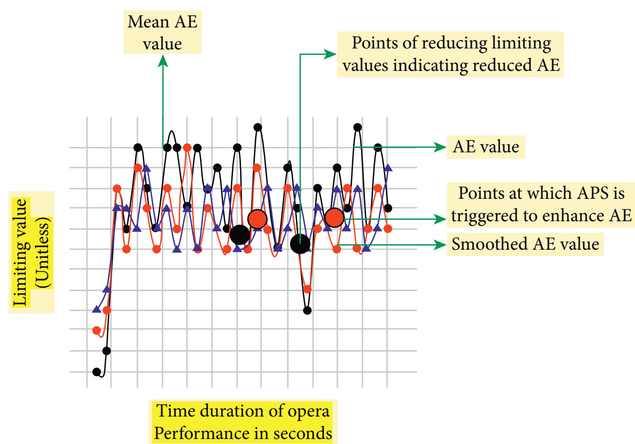

Figure 4: Enhancement of AE for opera performance-based MESDT.

moving mean filter, the AE value $AE(j,i)$ has been smoothed as shown in the following equation:fd7

$$AE_{\text{smoothed}}(n) = \frac{1}{T}\sum_{j=n-T+1}^{n} AE(j,i). \tag{7}$$

$AE_{\text{smoothed}}(n)$ is the smoothed AE value corresponding to n^{th} the frame. $AE(j,i)$ is the audience engagement for j^{th} sample from the i^{th} modality in a multimodal BCI. T is the time frame or period. This is selected via pretesting to determine optimum durations that guarantee there is enough data to generate reasonable estimates for opera performance.

3.3. Triggering APS to Enhance AE in Opera Performance. Because the multimodal signals obtained using the BCI approach mentioned above are audience-dependent, it is challenging to gauge the level of participation of different audiences during an opera performance. The performance levels of AE have been classified based on $AE_{\text{smoothed}}(n)$ values shown:fd8

$$\text{Levels of AE} = \begin{cases} \frac{1}{T}\sum_{j=n-T+1}^{n} AE(j,i) = 1 \in \text{pleasant AE}, \\ \frac{1}{T}\sum_{j=n-T+1}^{n} AE(j,i) = 0 \in \text{consisent AE}, \\ \frac{1}{T}\sum_{j=n-T+1}^{n} AE(j,i) = -1 \in \text{reduced AE}. \end{cases} \tag{8}$$

For each audience member, two thresholds have been created to distinguish between the following three unique levels of AE: pleasant AE (value of 1), consistent AE (value of 0), and reduced AE (value of −1) from the AE curve $AE_{\text{smoothed}}(n)$ given in Figure 4. $AE(j,i)$ is the audience engagement for j^{th} sample from the i^{th} modality in a multimodal BCI. T is the time frame or period. The limiting values of AE are denoted as $LV(n)$ and are given as shown in the following equation:

$$LV(n) = -1, \text{when} AE_{\text{smoothed}}(n) - AE_{\text{smoothed}}(n-1) < 0 \& AE_{\text{smoothed}} < Avg(AE). \tag{9}$$

$LV(n) = -1$ indicates the threshold for reduced AE. n is the frame number during the current period. $n-1$ is the frame number for the previous period. $Avg(AE)$ is the mean value of AE for the entire period of signal acquisition. The point $LV(n) = -1$ has been indicated by the circled dot in Figure 4. This point is the reference at which APS has to be triggered to enhance AE.

Likewise, another limiting value, known as the "AE enhancement threshold," is established to assess if the behavioral signals effectively reengaged the audience and how drastically they could alter the AE values. The limiting value for AE enhancement has been shown in the following equation:fd10

$$LV(n) = +1, \text{when} AE_{\text{smoothed}}(n) - AE_{\text{smoothed}}(n-1) > 0 \& AE_{\text{smoothed}} > Avg(AE). \tag{10}$$

The delivering signal has been found to successfully enhance AE if in $LV(n)$, the output is 1. Due to a delay from the smoothed AE value (AE_{smoothed}), it has been set at 5 minutes (based on pretest findings) as the reaction time.

Through carefully examining many biological signals and their interactions, the suggested MESDT framework—which accurately portrays the potential nature of human emotionality—has been developed. The proposed system included EEG in the recommended strategy to determine when to perform cues. Without specialized understanding, it is cozy and supportive enough for an audience to use in a makeshift theatre environment. Opera performance cues should be called upon as soon as the proposed framework detects the audience's disinterest to regain their attention swiftly.

4. Results and Discussion for the Proposed MESDT for Opera Performance

The third of Wagner's four operas that make up "The Ring of the Nibelung," the classical opera "Siegfried," has been selected as one of the experimental performances. In a five-minute scenario [30], Siegfried longed for Fafner to emerge and join him in battle after receiving the ring and changing

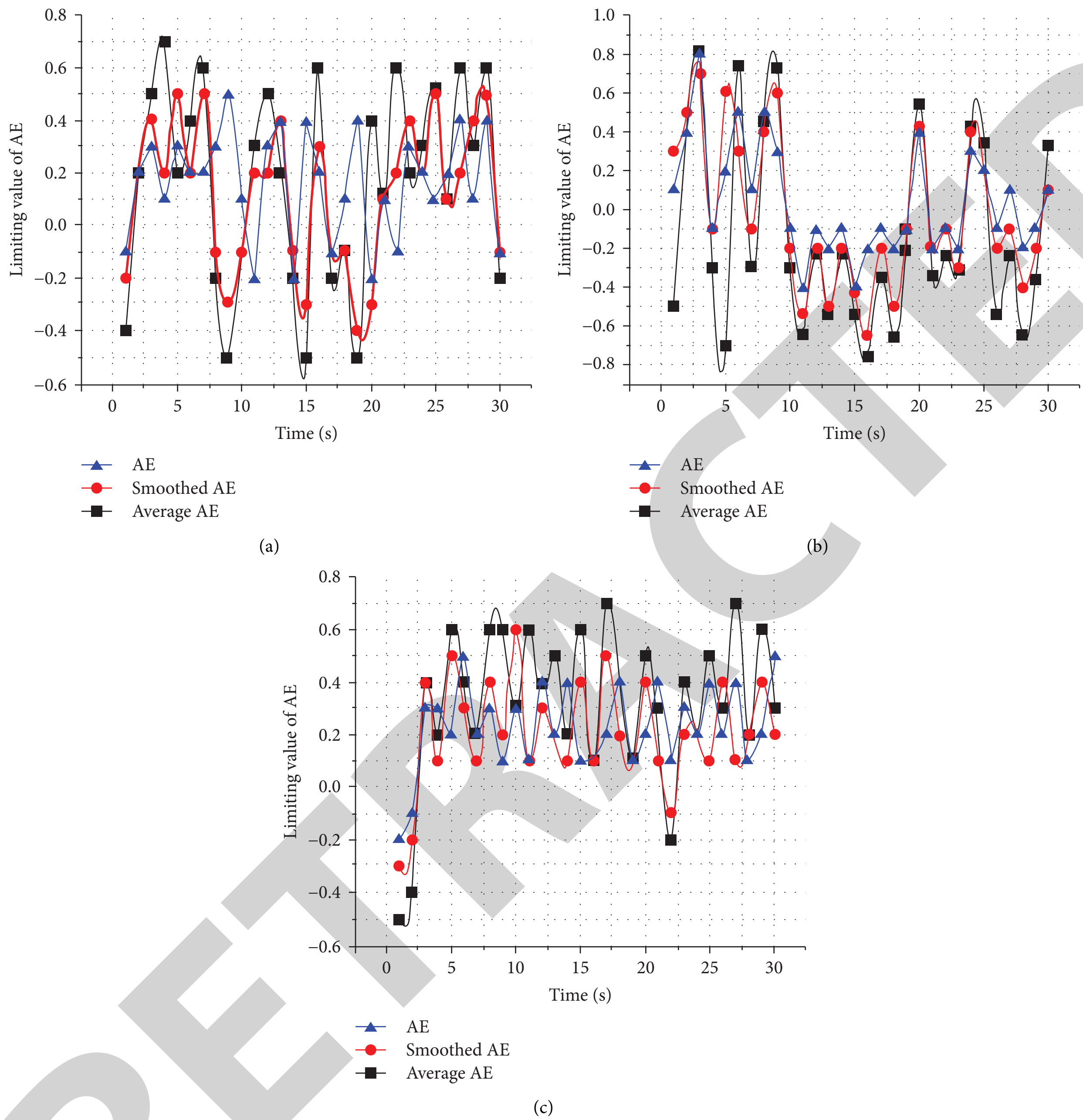

Figure 5: (a) No performing cues analysis. (b) Single performing cues analysis. (c) Multiple performing cue analysis for the proposed MESDT (with the classical opera "Siegfried").

into a monster. The action has been divided into the following three main parts: the actor who plays Siegfried's bugle while attempting to call the monster, the monster emerging from its tunnel and appearing in the shadowy jungle at the rear of the theatre, and Siegfried engaging the monster and stabbing it in the chest. In each step, pairs of single and multiple cues have been produced. One stimulus lasted for 20 seconds and was unrelated to the performance's substance. However, it also extracted emotional aspects and linked them to the show's illumination, acoustics, and special effects. The immersive theatre performance has been replicated for the proposed MESDT to provide the opera audience with a premium watching experience. In the experimental evaluation and performance research, 50 audiences have been included.

The different performing cues like no performing cues, single performing cues, and multiple performing cues are expressed in Figures 5(a)–5(c), respectively. The AE value, average AE value, and smoothed AE values are used to analyze the impact of BCI. The different opera performances with timing functions are analyzed, AE_{smoothed} is produced by removing noise from the actual AE value, and the average of the samples $Avg(AE)$ is used to analyze the opera performance. The α, θ, and β waveforms can be identified, and

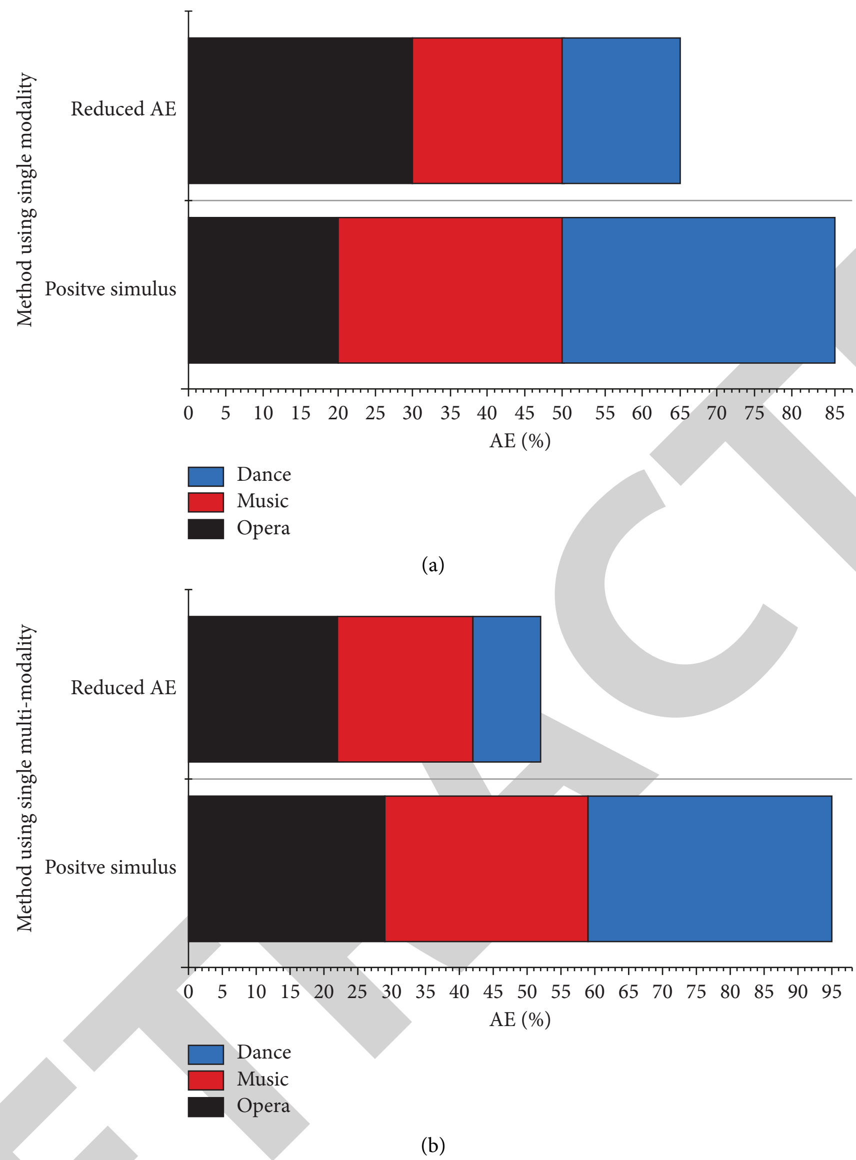

FIGURE 6: (a) AE evaluation analysis for a single modality using dance, music, and opera. (b) AE evaluation analysis for multimodality using dance, music, and opera.

the values of AE show the threshold limit. It has also been demonstrated that using the proposed MESDT framework with multimodal cues enhances the overall performance of AE by 5.8% when compared to a standard BCI design (one that employs only EEG features or physiological features).

The single modality and multimodality evaluation of the AE are analyzed and plotted in Figures 6(a) and 6(b), respectively, for 50 participants. Opera, dance, and music are the different components of the performance. The different α, θ, and β waveforms are monitored and fetched from the BCI signals. The different levels of the waveforms, like reduced AE level and positive stimulus level, are considered for the outcomes. The multimodal classification function b_{mean} is used to analyze and classify the different performances like dance, music, and opera activities. The actual EEG_{AE} is considered as the base for the simulation analysis. AE during opera performance has increased through appropriately triggering APS and considering multimodal feature extraction. The positive stimulus of AE with the proposed MESDT has an improved value of 95% compared to a single modality (value of 85%).

The stimulus and classification accuracy for the proposed MESDT for varying numbers of audiences are represented in Figures 7(a) and 7(b). The EEG signal from the BCI is fetched and EEG_{AE} is computed to find the opera performance and differentiate the opera results from dance and music AE_{smoothed}. The different values $.V^{\text{boost}}$ and $\text{Prediction}^{\text{boost}}$ are computed using the different samples, which contain EEG, physiological, and multimodal features to enhance the classification accuracy. The smoothed AE_{smoothed} values and the limiting AE values directly affect the accuracy of the EEG signals.

The different EEG samples and dance, music, and opera performances are considered for this analysis, and the

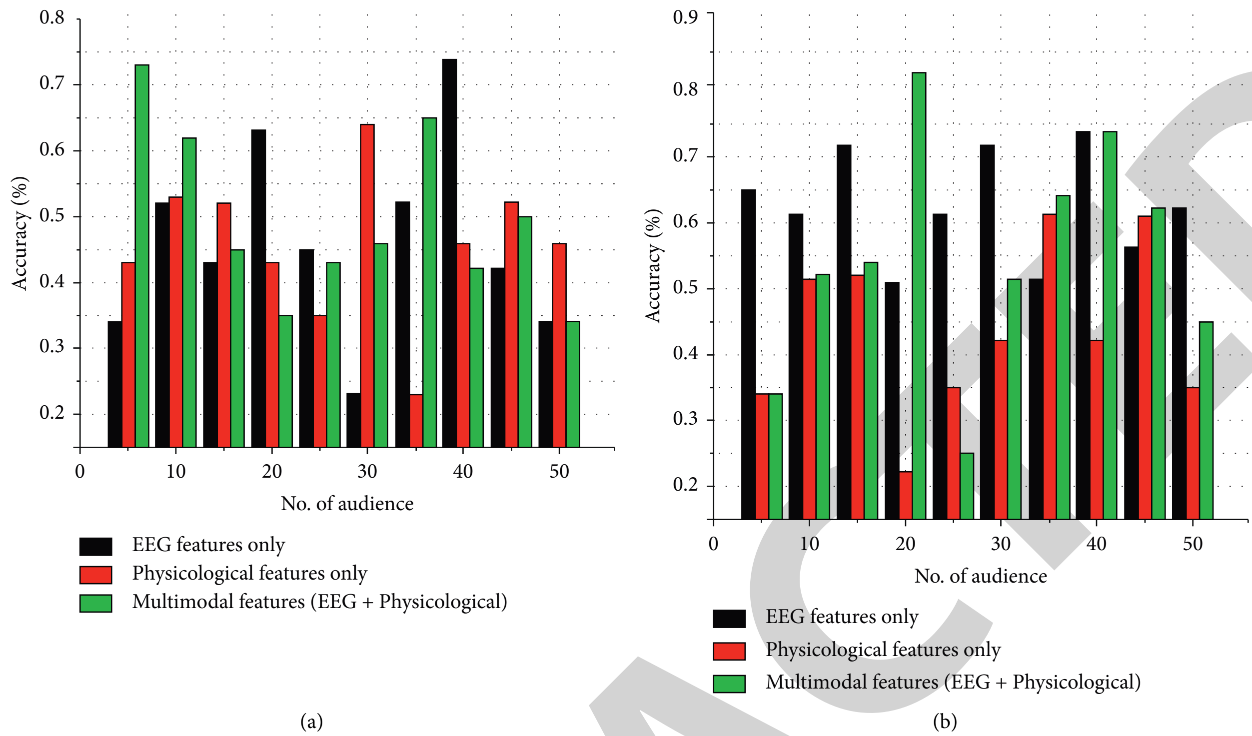

Figure 7: (a) Stimulus and (b) classification accuracy outcomes for the proposed MESDT.

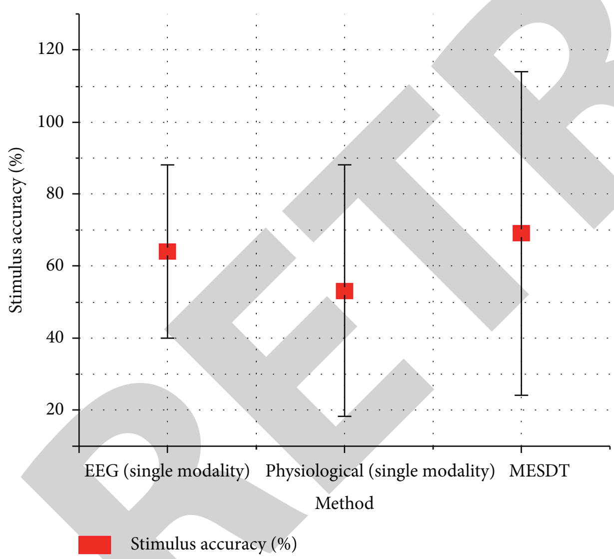

Figure 8: Stimulus accuracy outcomes.

outcomes are shown in Figure 8. The stimulus accuracy is computed using the different samples containing only EEG, physiological, and multimodal models with EEG and physiological features. The samples from the BCI interface are fetched EEG_{AE}, and the smoothed samples AE_{smoothed} are used to find the accuracy. The mean multimodal value b_{mean} ensures the highest accuracy in stimulus using the proposed MESDT system.

The MESDT system ensures the highest accuracy with multimodal (EEG and other physiological) signals. The EEG waveforms are fetched from the BCI, the AE smoothed AE_{smoothed}, and the average $Avg(AE)$ ensures the highest arousal and classification accuracy. The proposed MESDT framework with multimodal cues enhances the overall performance of AE by 5.8% when compared to a standard single-modal BCI design.

5. Conclusion, Limitations, and Scope for Further Research

A multimodal emotional state detection technique (MESDT) has been suggested to improve the AE in opera performance utilizing BCI. A hypothetical MESDT architecture is composed of three crucial components. First, audience members' electroencephalograms (EEG) and other biological indicators are recorded. The BCI then processes the collected signals and attempts to ascertain the user's current psychological reaction. Third, an adaptive performance stimulus (APS), as specified by a limiting value, is activated to improve AE in opera performance. The immersive theatre performance has been replicated to give the opera audience a premium watching experience. In the experimental evaluation and performance research, 50 people were included. The results showed that the APS had a positive effect on boosting AE during opera performance and that the suggested system had effectively identified AE decline. With multimodal (EEG and other physiological) inputs, the MESDT system guarantees maximum accuracy.

The poor signal-to-noise ratio of the EEG and other physiological inputs utilized for the customized emotional state identification system is a shortcoming of the methodology described in this work. The fact that this study only had a small number of participants (only 50) might also be a possible flaw. The relatively low number of audiences nevertheless offers a sufficiently reliable test case for the strategy because this study aims to develop a method for customized emotional state detection rather than a general solution that would work for everyone. Future research will examine the method's appropriateness for detecting emotional states in more individuals.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] M. D. Cirillo, A. Affonsêca Netto, G. M. Toffolo, and A. M. F. L. Miranda de Sa, "Development of a brain computer interface based on steady-state visual evoked potential with multiple intermittent photo stimulation," *Biomedical Signal Processing and Control*, vol. 57, 2020.

[2] F. Lotte, L. Bougrain, A. Cichocki et al., "A review of classification algorithms for EEG-based brain-computer interfaces: a 10-year update," *Journal of Neural Engineering*, vol. 15, no. 3, 2018.

[3] A. E. Hramov, V. A. Maksimenko, and A. N. Pisarchik, "Physical principles of brain-computer interfaces and their applications for rehabilitation, robotics, and control of human brain states," *Physics Reports*, vol. 918, pp. 1–133, 2021.

[4] M. A. Khan, R. Das, H. K. Iversen, and S. Puthusserypady, "Review on motor imagery-based BCI systems for upper limb post-stroke neurorehabilitation: from designing to application," *Computers in Biology and Medicine*, vol. 123, 2020.

[5] E. P. Torres, E. A. Torres, M. Hernández-Álvarez, and S. G. Yoo, "EEG-based BCI emotion recognition: a survey," *Sensors*, vol. 20, no. 18, p. 5083, 2020.

[6] R. Abiri, S. Borhani, E. W. Sellers, Y. Jiang, and X. Zhao, "A comprehensive review of EEG-based brain–computer interface paradigms," *Journal of Neural Engineering*, vol. 16, no. 1, 2019.

[7] I. Lazarou, S. Nikolopoulos, P. C. Petrantonakis, I. Kompatsiaris, and M. Tsolaki, "EEG-based brain-computer interfaces for communication and rehabilitation of people with motor impairment: a novel Approach of the 21st century," *Frontiers in Human Neuroscience*, vol. 12, p. 14, 2018.

[8] N. Jamil, A. N. Belkacem, S. Ouhbi, and C. Guger, *Cognitive and Affective Brain-Computer Interfaces for Improving Learning Strategies and Enhancing Student Capabilities: A Systematic Literature Review*, IEEE Access, Piscataway, NJ, USA, 2021.

[9] Y. W. Shen and Y. P. Lin, "Challenge for affective brain-computer interfaces: non-stationary spatio-spectral EEG oscillations of emotional responses," *Frontiers in Human Neuroscience*, vol. 13, p. 366, 2019.

[10] M. B. Akçay and K. Oğuz, "Speech emotion recognition: emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers," *Speech Communication*, vol. 116, pp. 56–76, 2020.

[11] A. Evmenenko and D. S. Teixeira, "The circumplex model of affect in physical activity contexts: a systematic review," *International Journal of Sport and Exercise Psychology*, vol. 20, no. 1, pp. 168–201, 2022.

[12] B. Steichen and B. Fu, "Cognitive style and information visualization—modeling users through eye gaze data," *Frontiers of Computer Science*, vol. 2, 2020.

[13] C. Daudén Roquet, C. Sas, and D. Potts, "Exploring anima: a brain-computer interface for the peripheral materialization of mindfulness states during mandala coloring," *Human-Computer Interaction*, vol. 41, 2021.

[14] C. Pernet, M. I. Garrido, A. Gramfort et al., "Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research," *Nature Neuroscience*, vol. 23, no. 12, pp. 1473–1483, 2020.

[15] K. Tsiakas, M. Abujelala, and F. Makedon, "Task engagement as personalization feedback for socially-assistive robots and cognitive training," *Technologies*, vol. 6, no. 2, p. 49, 2018.

[16] T. Shi, L. Ren, and W. Cui, "Feature extraction of brain–computer interface electroencephalogram based on motor imagery," *IEEE Sensors Journal*, vol. 20, no. 20, pp. 11787–11794, 2020.

[17] B. Almarri, S. Rajasekaran, and C. H. Huang, "Automatic subject-specific spatiotemporal feature selection for subject-independent affective BCI," *PLoS One*, vol. 16, no. 8, Article ID e0253383, 2021.

[18] K. Ryding, J. Spence, A. S. Løvlie, and S. Benford, "Interpersonalizing intimate museum experiences," *International Journal of Human-Computer Interaction*, vol. 37, no. 12, pp. 1151–1172, 2021.

[19] A. Thanou, E. E. Tsiropoulou, and S. Papavassiliou, "Quality of experience under a prospect theoretic perspective: a cultural heritage space use case," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 1, pp. 135–148, 2019.

[20] G. Pavlidis, "Recommender systems, cultural heritage applications, and the way forward," *Journal of Cultural Heritage*, vol. 35, pp. 183–196, 2019.

[21] D. Tezza, D. Caprio, B. Pinto, I. Mantilla, and M. Andujar, "An analysis of engagement levels while playing brain-controlled games," in *International Conference on Human-Computer Interaction*, pp. 361–372, Springer, Berlin, Germany, 2020, July.

[22] M. Prpa and P. Pasquier, "Brain-computer interfaces in contemporary art: a state of the art and taxonomy," *Brain art*, vol. 65, 2019.

[23] I. Daly, D. Williams, A. Malik et al., "Personalised, multimodal, affective state detection for hybrid brain-computer music interfacing," *IEEE Transactions on Affective Computing*, vol. 11, no. 1, pp. 111–124, 2020.

[24] R. Behroozmand and S. Sangtian, "Neural bases of sensorimotor adaptation in the vocal motor system," *Experimental Brain Research*, vol. 236, no. 7, pp. 1881–1895, 2018.

[25] S. S. Shekhawat, H. Sharma, S. Kumar, A. Nayyar, and B. Qureshi, "bSSA: binary salp swarm algorithm with hybrid data transformation for feature selection," *IEEE Access*, vol. 9, pp. 14867–14882, 2021.

[26] X. Li, Z. Wu, Z. Zhao, F. Ding, and D. He, "A mixed data clustering algorithm with noise-filtered distribution centroid and iterative weight adjustment strategy," *Information Sciences*, vol. 577, pp. 697–721, 2021.

[27] H. Huo, J. He, and X. Guan, "A Bayesian fusion method for composite damage identification using Lamb wave," *Structural Health Monitoring*, vol. 1475921720945000, 2020.
[28] Y. Cao, Q. G. Miao, J. C. Liu, and L. Gao, "Advance and prospects of AdaBoost algorithm," *Acta Automatica Sinica*, vol. 39, no. 6, pp. 745–758, 2014.
[29] P. P. Mini, T. Thomas, and R. Gopikakumari, "EEG based direct speech BCI system using a fusion of SMRT and MFCC/LPCC features with ANN classifier," *Biomedical Signal Processing and Control*, vol. 68, 2021.
[30] https://www.youtube.com/watch?v=5xRNL8IQm7I.