Xinye Wang[1]
Jing Qiu[2*]

Augmented Physics-Informed Neural Networks and Seahorse Optimization Algorithm for Music Scene Simulation and Emotion Regulation in Virtual Reality Technology

Journal of Electrical Systems

Abstract: - Virtual Reality (VR) technology has shown promise in simulating music scenes for emotion regulation. The challenge is to enhance the realism and interactivity of VR-based music experiences to create a more immersive and emotionally impactful environment for users. This manuscript proposes an augmented physics-informed neural network and seahorse optimization algorithm for music scene simulation and emotion regulation in virtual reality technology (APINN-SHO-MSS-ERVRT). Initially, the data is collected via head-mounted display (HMD) in real time basis. Afterward, the data is fed to data-adaptive Gaussian average filtering (DAGAF) based pre-processing process. In the pre-processing segment, it improves virtual presence. The pre-processing output is given to augmented physics-informed neural network (APINN)and seahorse optimization algorithm (SHO) for effectively classifying the facial expression for fear and null. Therefore, SHO is proposed to enhance weight parameter of augmented physics-informed neural network classifier, which precisely classifies the facial emotion classifier. The proposed method is implemented in MATLAB and evaluated their performance with existing methods. The performance metrics, like accuracy, precision, sensitivity, specificity, computational time and ROC is analysed to the performance of the proposed method. The proposed APINN-SHO-MSS-ERVRT methods of accuracy are provide 90% higher accuracy for fear and 97% higher accuracy for null emotion is analysed when compared with existing methods GERR-NN-AMSVR, ER-VR-EEG and MTS-VRS-CDI. Proposed APINN-SHO-MSS-ERVRT method attains 0.87%, 0.88% and 0.89%higher ROC analysed to the existing methods respectively.

Keywords: Music, Emotion Regulation, Virtual Reality, Fear, Improved Adaptive Federated Kalman Filter, Facial Expression, Seahorse Optimizer.

I. INTRODUCTION

In the contemporary era, an increasing number of individuals find themselves grappling with elevated levels of stress. The prevalence of long-term depression is on the rise, contributing to a growing societal burden [1]. The tradition of music therapy in China dates back through generations, with Chinese ancestors acknowledging the interconnectedness of music, medicine, and healing [2]. Globally, the use of music as a means of emotional regulation is widespread [3].Music therapy involves an individual listening to various music genres under a music therapist's guidance [4]. The customization of therapeutic repertoires takes into account the treatment goals, the individual's physiological and psychological state, and the specific treatment environment [5]. Unfortunately, the accessibility of music therapists is constrained by cost and availability, leaving a significant number of individuals in need of emotional regulation without adequate support [6].An automated therapist could serve as a viable alternative, addressing the gap in available resources. Furthermore, listening to music at home has become a prevalent method for individuals to alleviate negative emotions [7].

Studies indicate that emotional reactions induced by music closely resemble spontaneous emotions in terms of quality. Music has the ability to produce distinct patterns of physiological changes, such as alterations in heart rate and blood pressure, corresponding to specific emotions [8]. The research specifically concentrated on male participants experiencing four emotions joy, anger, pleasure, and sadness and achieved an impressive average recognition accuracy of 87%, surpassing previous studies [9].While the exploration of music's evocative power in generating authentic emotional states is limited, Kim and André's work contributes to the growing body of evidence supporting the notion that music can genuinely elicit emotional responses [10]. In their research, Kim and André utilized musical induction as a method to elicit emotions, recording subjects' physiological signals during the exposure to various musical pieces [11].

[1]School of Humanities and Social Sciences, Fuzhou University, Fuzhou, Fujian , China, 350108
[1]E-mail: violin_wxy@163.com
[2*]Teaching and Research Support Center, Naval Submarine Academy, Qingdao, Shandong, China, 266000
*Corresponding author e-mail: 18181093299@163.com

Subjects were encouraged to select songs associated with special memories related to the target emotions, recognizing music's role in accompanying significant life events and its connection to personal memories [12].Moreover, the solitary nature of music listening minimizes the impact of social masking artifacts, and individuals' emotional responses to music can significantly vary based on their past experiences and cultural background [13]. Studies have demonstrated that an individual's emotional reactions to music are susceptible to influence by their musical preferences.

Various studies reinforce the notion of music as a potent stimulus for evoking emotional responses, accompanied by discernible changes in the Autonomic Nervous System (ANS) [14]. Researchers have identified significant variations in physiological measures during music listening, including elements such as melody, rhythm, and continuity [15]. This aligns with Kim and André's findings and is corroborated by other research highlighting music's impact on ANS reactions [16].Notably, certain studies have delved into clinical and therapeutic contexts, uncovering the influence of music on ANS responses [17]. For instance, research has shown that music can trigger ANS reactions such as vascular constriction, heart rate fluctuations, muscle tension, and changes in skin temperature (SKT), even when subjects report reduced anxiety and increased relaxation [18]. The study revealed distinct associations between emotions and physiological changes; happiness manifested in significant alterations in respiration, sadness correlated with pronounced changes in blood pressure, heart rate, and SKT, while fear exhibited the highest change in blood flow velocity [19].These findings suggest that music doesn't merely convey easily recognizable emotions but actively generates genuine emotions in the listener [20]. However, the extent to which changes in the ANS and distinctions in musical emotions align with those observed in non-musical emotions remains unclear.

The main contribution of this manuscript includes;

- This manuscript, APINN-SHO-MSS-ERVRT is proposed. Initially input data are collected from HMD in real time basis. Afterward, the collected data are fed to pre-processing by utilizing data-adaptive Gaussian average filtering.
- Data is collected from head-mounted display (HMD) in real time basis dataset.
- In pre-processing segment, it improves the presence in VR. Following the pre-processing processes, the resulting output is inputted into the classification method.
- The proposed technique is executed and the efficiency of proposed APINsN-SHO-MSS-ERVRT for classify facial emotional expression is evaluated by several performances analysing metrics like accuracy, precision, specificity, sensitivity, computational time, and ROC.
- From the result, it concludes that the proposed approach is better compared with existing approaches like GERR-NN-AMSVR, ER-VR-EEG and MTS-VRS-CDI respectively.

Remaining manuscript is organized as follows: sector 2 depicts survey of literature, sector 3 describes proposed approach; the outcomes are proved in sector 4, and finally, the end is presented in sector 5.

II. LITERATURE SURVEY

Among the frequent research work on Music Scene and Emotion Regulation in Virtual Reality Technology based on deep learning; In this section, some of the most recent investigations were evaluated.

Ibanez et al. [21] developed an adaptive music system in virtual reality applications using a head gesture detection system to detect emotional inputs and provide them, enlightening virtual presence. The system underwent two iterations, both based on neural networks: the first iteration used hybrid convolutional neural network operating in one dimension, and the second iteration used a multi-layer perceptron. In both instances, the system successfully recognized fear by analyzing head gestures.

Suhaimi et al. [22] presented a virtual reality (VR) headgear intended to elicit four emotional categories. They used an inexpensive wearable EEG device to record brainwaves, and employed popular classifiers to compare their feasibility for this specific setup. The authors' first goal was to create a public, immersive VR database that would help with studies using virtual reality stimuli and emotion recognition. Second, they made use of a tiny, inexpensive wearable EEG headgear that was easy to affix to the scalp and gave participants unrestricted movement to view their surroundings within the immersive virtual reality stimulation.

Feng [23] introduced a virtual reality-based system for teaching music. Initially, they created a virtual piano using the hardware platform of the HTC Vive kit and the Leap Motion sensor installed on a helmet. The software platforms included Leap Motion plug-ins, Unity3D, and pertinent Steam VR plug-ins. Subsequently, they proposed and implemented a gesture recognition algorithm. To collect user gesture command data, a DCCNN, or Dual Channel Convolutional Neural Network, was employed. In order to enable DCCNN to recognize them, this network used a dual-size convolution kernel to recognize gesture commands in films and extract feature information from images. Photos with a Red-Green-Blue (RGB) color pattern were fed into the DCCNN and optical flow images after extracting spatial and temporal information.

Shen and Zhao [24] presented a model for music-based emotion recognition features, along with the construction of a Computer-Aided Design (CAD) system for musical animation. The system extracted fundamental musical features from MIDI (Musical Instrument Digital Interface) files and further derived the musical features of the music. This process involved segmenting each music section into emotional music visualization programs. By aggregating the segment features, they designed an emotional feature program capable of reflecting the musical form as a whole. Through the application of deep learning algorithms, the improved segment nodes were visualized and matched, thereby enhancing the expressive form of music emotional animation.

Zhao et al. [25] developed an innovative technique that uses inertial signals and the electroencephalogram (EEG) to identify emotions while walking. By use of end-to-end deep learning training and multi-modal fusion, emotions were recognized accurately. Wearing virtual reality head-mounted display gear, participants were able to fully experience powerful emotions while they were moving. The virtual reality environment exhibited remarkable imitation and experience qualities, significantly contributing to the elicitation and modification of emotions. Furthermore, DWT was used to convert the multi-modal information from inertial sensors and EEG into virtual emotion images. The attention-based CNN fusion model used these images as inputs.

Wang and Ko [26] presented the emotional depiction of music during the music experience. Initially, they employed Internet of Things sensors based on the principles of multi-modal technology to acknowledge the emotional impact of music. Subsequently, they utilized the Naive Bayes Classifier grounded in deep learning to classify the emotions in music compositions and users' emotional responses to the music. Lastly, they examined how consumers' emotional experiences and music's emotional representation were analysed, studying the accuracy of music emotion classification through various classification methods.

Patel [27] developed Convolutional neural networks were used to process facial features for the purpose of detecting emotions (CNN). The six main categories of emotions that were examined were: fear, disgust, anger, sadness, happiness, and surprise. These emotions were further categorized into two types: positive emotions, which entail feelings without negativity such as happiness or neutrality, and negative emotions, encompassing sensations of depression and frustration, including anger, sadness, and fear.

III. PROPOSED METHODOLOGY

The proposed method APINN-SHO-MSS-ERVRT is discussed in this section. This model composed to detect the facial expression from input data. Block diagram of APINN-SHO-MSS-ERVRT technique is presented in Figure 1. The proposed APINN-SHO-MSS-ERVRT takes the extracted data from the VR devices. This process consists of four steps: dataset preparation, pre-processing, classification, and optimization. Thus, a full explanation of all steps is provided below,

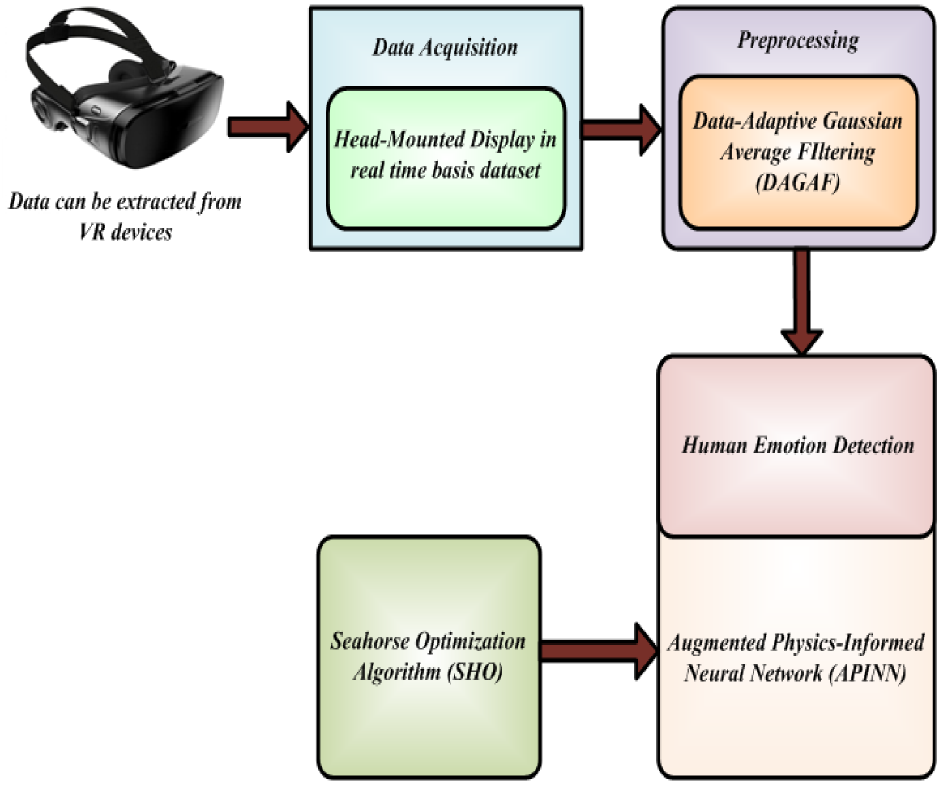

Figure 1: Block diagram of proposed APINN-SHO-MSS-ERVRT technique

A. Data Acquisition

The information are collected from head-mounted display in real time basis dataset. A system for recognizing gestures has been developed to capture emotional cues from players using a HMD. Additionally, an adaptive music generator has been implemented to tailor the game experience based on these emotional inputs. The latter system serves as an emotional regulator, modifying the soundtrack to enhance or diminish a detected emotion. Recognizing dynamic gestures, characterized through a succession of motions made over time at a particular spatial position, poses a particular challenge. The time-dependent data stream generated by the HMD proves valuable in predicting outcomes in this domain, where a continuous flow of data is present.

B. Pre-processing using data-adaptive Gaussian average filtering(DAGAF)

DAGAF is used in this part to pre-process the input data. To enhance virtual presence in virtual reality applications, it is necessary to recognize emotional inputs and feed them into an adaptive music system [28]. Because head-mounted displays (HMDs), which are increasingly used in virtual reality (VR), may provide a very rich stream of information regarding a player's position, posture, and rate of acceleration when moving, a room-scale VR environment was chosen for this study technique. This filtering approach is moving average process performed to data, weights of average filter are created using a Gaussian window. Gaussian function with a standard deviation on its distribution is given in equation (1).

$$G_a(\omega) = \rho \cdot (2\vartheta)^{1/2} \cdot b^{-(\rho\omega)^2/2} \tag{1}$$

Where ρ is the analogous continuous-time Fourier transform (CTFT).In the discrete time domain, the Gaussian function becomes G_a assuming sample interval ρ equal to 1, ϑ is integer which is followed by Fourier transform shown in equation (2).

$$l[N] = l[-N] = b^{-\frac{1}{2}\alpha^2 < 0.05} \tag{2}$$

Where, the Gaussian distribution's standard deviation and the parameter α are inversely related. Making values at end point (N) lesser than 5% of maximal window value yields parameter α in this research. The maximum may be seen at b, when $l[0]$ has a value of 1 and is shown in the following equation (3).

$$L(\omega) = \frac{N}{\alpha}(2\vartheta)^{1/2} \cdot b^{-\frac{1}{2}\left(\frac{N}{\alpha}\omega\right)^2} \tag{3}$$

The spectrum obtained from above equation has a bell-shaped centre at $L(\omega)$ can be considered as a low pass filter when given the reasonable values N and α. DAGAF average filters the spatial data using a normalized discrete truncated Gaussian window, provided in order to preserve the data during the filtering process and given in equation (4).

$$l_M[n] = \frac{l[n]}{\sum_{w=-N}^{N} L[w]} \tag{4}$$

Where the above equation defines Gaussian average filter and it is in fact low pass filter. Where, is then assuming Gaussian average filter $l_M[n]$ has been found and that the missing temporal h to be analyzed is represented as $L[w]$. The moving-average technique that follows can be used to get the instantaneous mean in DAGAF and shown in equation (5).

$$n_i[h] = \sum_{n=-N}^{N} l_M[n] \cdot r[n+h], \qquad for \quad 0 \le h \le H-1 \tag{5}$$

A system for recognition that gathers data from a virtual world and identifies pertinent feelings. Initially, the gesture recognition component of the system processes the input it receives from the head-mounted display used to play the game. The processed data in one-decomposition iteration is indicated as $l_M[n]$, while the data to be analysed is represented by $g_i[h]$. This method involves first extending the data by a "reflection" extension, and then reflecting the expanded segment in a filtering process with regard to a fixed value. Then the contemporaneous mean in DAGAF is then obtained by extending the data using the following equation (6).

$$g_i[h] = \sum_{n=-N}^{N} l_M[n] \cdot r_b[g+h], \qquad for \quad 0 \le h \le H-1 \tag{6}$$

Due to the symmetric structure of the Gaussian average filter $l_M[n]$, the instantaneous mean calculation in equation is essentially the convolution total of $r_b[\cdot]$. To identify the relevant emotion in the traffic data will be recovered in the data domain by directly multiplying the spectrum of $r_b[\cdot]$ and the data bank of $l_M[\cdot]$.Using the above equation the input data is pre-processed imbalance ratio through duplicating instances in minority class DAGAF filtering method. After the data pre-processed is then further given for classification process in the following section.

C. Classification Using Augmented Physics-Informed Neural Networks (APINNs)

This sector discusses the use of APINN for facial cue recognition in the identification of basic human emotions. Several techniques have gained significant prominence among these, playing a crucial role in various virtual reality (VR) research endeavours [29]. They have proven instrumental in diverse applications, including the detection of faces and facial expressions, tracking human body movements, visualizing data, and recognizing speech. Neural network optimization should abide by data and physical rules as prescribed by a partial differential equation (PDE) in order to approximate its solution is the driving force behind the Physics-Informed Neural Network (PINN). This entails reducing the training loss, which is represented as follows and is made up of a residual loss and a boundary loss:

$$G_S(\theta) = \frac{1}{m_a} \sum_{j=1}^{m_a} \left| V_\theta(Z_{a,j}) - R(Z_{a,j})^2 \right| + \frac{1}{m_r} \sum_{j=1}^{m_r} \left| LV_\theta(Z_{r,j}) - F(Z_{r,j}) \right|^2 \tag{7}$$

Let m_a represent the boundary training points, m_r represent the residual training points, and The PINN model V_θ. approximates the ground truth PDE solution. In the first term, In the first semester, PINN studies boundary conditions; in the second term, it studies the physical laws that PDEs explain.

Since XPINN's interface losses might not always guarantee continuity across different sub-PINNs, bigger mistakes could be seen close to the interface. This challenge arises from the difficulty in accurately enforcing residual continuity conditions for partial differential equations (PDEs) that involve higher-order derivatives. The presence of these higher-order derivatives makes it challenging to maintain precise continuity conditions. As a result, making sure that first-order derivatives continue amongst various sub-PINNs is proposed as a resolution to this issue:

$$D_b\left(\theta^j,\theta^i\right)=\frac{1}{m_{I,ji}}\sum_{k=1}^{m_{I,ji}}\sum_{n=1}^{q}\left|\frac{\partial V_{\theta^{ij}}\left(Z_{I,K}^{ji}\right)}{\partial Z_n}-\frac{\partial V_{\theta^{J}}\left(Z_{I,K}^{ji}\right)}{\partial Z_n}\right| \tag{8}$$

Where q represent the problem dimension. While θ^j represent the parameters for subdomain j and $Z_{I,K}^{ji}$ represent the kth interface points.

When the enhanced PINN is parameterized by θ, its output is expressed as;

$$V_\theta(Z)=\sum_{j=1}^{N}(D(Z))_j\, F_j(g(Z)) \tag{9}$$

Where $(D(Z))_j$ represent the jth entry of $D(Z)$ and θ represent the collection of all parameters in g, D and F_j. Both g and F_j are trainable in APINN, while D can be either trainable or fixed. If D represent the trainable model of APINN. The F_j, g, D sub-modules are all nonlinear fully-connected neural networks, making the APINN model nonlinear. Additionally, the APINN is a universal approximator.

Returning to the APINN paradigm, we can identify the function class of APINN as

$$\text{APINN}=\left\{F \mid 3F_1,...,F_N, g\in \text{NN}, D\in G, s.t., F=\sum_{j=1}^{N}D(Z)_j\, F_j\,(g\,(Z))\right. \tag{10}$$

Here, NN represents all neural networks, and the function class of the gating network is denoted as G. Since multilayer neural networks are a subset of the model APINN and are by nature universal approximators, APINN must likewise be a universal approximator.

There are three possible approaches to constructing the APINN model. The first and simplest option involves excluding parameter sharing in APINN, resulting in the model becoming:

$$V_\theta(\delta)=\sum_{j=1}^{N}(D(Z))_j\, F_j(Z) \tag{11}$$

The goal of each sub-PINN's parameter sharing is to increase parameter efficiency. By using the shared network as the identity mapping and applying the aforementioned equation F_j networks. APINN stands out as a widely employed computing system for gesture recognition, particularly with devices capable of extracting substantial data, including position and rotation information. Given each sub-PINN's contribution to the target function as a whole, it makes intuitive sense that the functions that each sub-PINN learns should be somewhat similar. Our model explicitly utilizes this intuition by include network sharing, resulting in enhanced parameter efficiency. Facial expressions related to human emotions, such as fear and neutrality, are predicted using APINNN with the equation mentioned above. Notably, APINN lacks an inherent optimization structure for refining weight parameters to achieve more accurate predictions of facial expressions for human emotions. To address this, an optimization algorithm is employed, as detailed in the following section.

D. Stepwise process for Sea-Horse Optimization Algorithm

In this section, optimization using SHO is discussed. Here the proposed method is utilized to classify the facial expression of human emotion. Drawing inspiration from the way seahorses navigate, hunt, and breed in the ocean, the Sea-Horse Optimizer (SHO) is developed [30]. The SHO algorithm incorporates the principles of exploration and exploitation, mirroring the social behaviour of seahorses in their movement and search for prey. The algorithm's design is influenced by the phases of seahorse breeding, with the final phase initiated once the two components have concluded. A detailed modelling of the SHO method is presented as follows:

Step 1: Initialization

Initialize population of SHO weight parameter values of APINN. It expressed in equation (12)

$$Zs=\begin{bmatrix} Z_{1,j} & Z_{1,Dim-1} & \cdots & Z_{1,H} \\ Z_{2,j} & Z_{2,Dim-1} & \cdots & Z_{2,H} \\ \cdots & \cdots & \cdots & \cdots \\ Z_{n,j} & Z_{n,Dim-1} & \cdots & Z_{n,H} \end{bmatrix} \tag{12}$$

Here Zs denotes the SHO population matrix, n is denoted as the size of population, H is the dimension of variables, correspondingly.

Step 2: Random Generation

Following initialization, input fitness function developed randomness via SHO method

Step 3: Fitness Function

An initialization value, result is random solution. Assessment of fitness values utilizes outcomes of weight parameter optimization δ . It expressed in equation (13),

$$fitness\ \ function = Optimizing\ \ [\delta] \tag{13}$$

Step 4: Seahorse Movement Behaviour

The movement pattern of seahorses serves as a reference for the normal distribution. Two examples are provided to help achieve a balance between exploration and exploitation, each having a boundary point set at 0.

Case 1: Exploration

In order to expand the local solution zone, the agent continuously modifies the rotation angle as it spirals towards the $Xelite$. Mathematically, exploration (σ) can be expressed as follows:

$$\sigma = \left(\frac{\Gamma(1+\lambda) \times \sin\left(\frac{\pi\lambda}{2}\right)}{\Gamma\left(\frac{1+\lambda}{2}\right) \times \lambda \times 2\left(\frac{\lambda-1}{2}\right)} \right) \tag{14}$$

Where λ represent the random value between 0 and 2.

Case 2: Exploitation

The seahorse uses a Brownian motion to improve its traversal in the presence of ocean waves by replicating the motion length of another seahorse. This process can be expressed as follows:

$$\beta_j = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{Z^2}{2}\right) \tag{15}$$

Where β_j represent the Brownian motion's random walk coefficient.

Step 5: Seahorse Foraging Behaviour

When seahorses search for food, two potential outcomes exist: failure and success. The success condition is established when $r_2 > 0.1$, indicating that the seahorse travels more quickly than its prey. On the other hand, if the answer varies, a failure condition happens. The conditions for success and failure in seahorses' quest for food can be expressed as follows:

$$\alpha = \left(1 - \frac{s}{S}\right)^{\frac{2s}{S}} \tag{16}$$

Where S represent the maximum number of iteration.

Step 6: Seahorse Breeding Behaviour

During the breeding period, seahorses are categorized into two gender groups: male and female, each comprising an equal composition of 50%.

$$\begin{cases} Father = Z_{sort}\left(1 : \frac{pop}{2}\right) \\ Mother = Z_{sort}\left(\frac{pop}{2} + 1 : pop\right) \end{cases} \tag{17}$$

Where Z_{sort} represent the ascending order of fitness value. $Father$ and $Mother$ were chosen randomly. In the SHO algorithm, each pair produces one child.

$$Z_j = (1 - r_3)\ Z_{Mother} + r\ Z_{Father} \tag{18}$$

Where r_3 represent the random number and male as well as female randomly chosen members represented by Z_{Father} and Z_{Mother} respectively.

Step 7: Termination

Check the termination criteria, if it met the condition means optimal solution is obtained, otherwise repeat the process. The weight parameter is δ from APINN is optimized with SHO, for effectively for classify the facial emotional expression as fear and null. Figure 2 illustrate the flowchart of SHO.

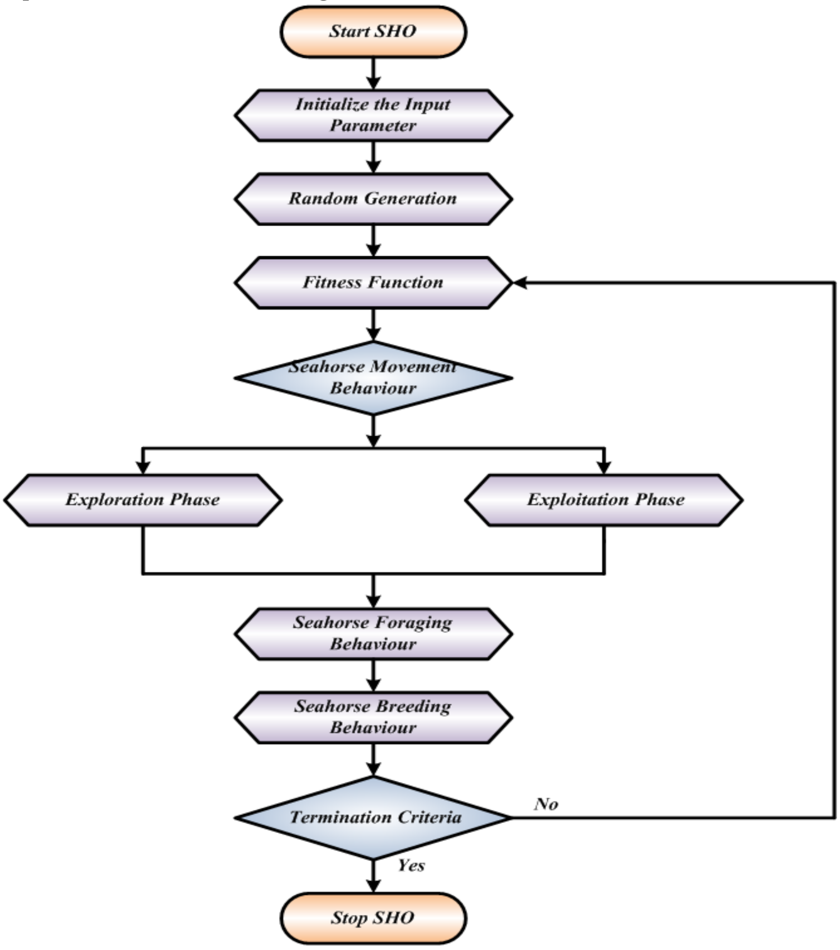

Figure 2: Flowchart of SHO

IV. RESULT AND DISCUSSION

This section discusses the experimental results using the proposed approach. The proposed technique simulated utilizing python based numerous performance measures comprising accuracy, precision, sensitivity, specificity, computational time and ROC. Obtained results of proposed APINN-SHO-MSS-ERVRT technique are analysed with existing approaches such as GERR-NN-AMSVR [21], ER-VR-EEG [22] and MTS-VRS-CDI [23] respectively.

A. Performance measures

Performance of proposed approach articulated utilizing the accuracy, ROC, precision, specificity and sensitivity performance measures. For that, following confusion matrix is necessary.

- TP: “Instances that are actually positive then categorized as Null”.
- FP: “Instances that are actually negative but categorized as Null”.
- FN: “Instances that are actually positive but categorized as Fear”.
- TN: “Instances that are actually negative then categorized as Fear”.

1) Accuracy

It is accurate to predict emotion rather than their status rate. Accuracy serves as a proxy for the overall correctness of the information gathered. It represents the ratio of true positives, true negatives to all occurrences in dataset. It is given in equation (19)

$$Accuracy = \frac{Tp + Tn}{Tp + Tn + Fp + Fn} \tag{19}$$

Here, true positive refers to properly anticipated emotion as null. False positive, or incorrectly diagnosed as null, is known as FP. TN stands for true negative, or accurately estimated emotion as fear. False negative, or incorrectly projected emotion as fear, is FN.

2) Precision

The degree of accuracy of the data that was obtained is measured by precision. It is proportion of "true positives" to all "positive instances," where "true positives" refers to the number of correctly returned results. The precision is computed using the following equation (20),

$$Precision = \frac{Tp}{Tp + Fp} \tag{20}$$

3) Specificity

The percentage of true negatives that the method correctly identifies is called specificity. It is determined by equation (21),

$$Specificity = \frac{Tn}{Tn + Fp} \tag{21}$$

4) ROC

ROC can be stated as ratio among changes in single variable relative to corresponding change in another, graphically; rate of change represents slope of line. It is given in equation (22)

$$ROC = 0.5 \times \left(\frac{TP}{TP + FN} + \frac{TN}{TN + TP} \right) \tag{22}$$

B. Performance analysis

Figure 3 to 8 portrays model result of APINN-SHO-MSS-ERVRT method. Then, the proposed APINN-SHO-MSS-ERVRT technique analyzed with existing method likes GERR-NN-AMSVR, ER-VR-EEG and MTS-VRS-CDI models respectively.

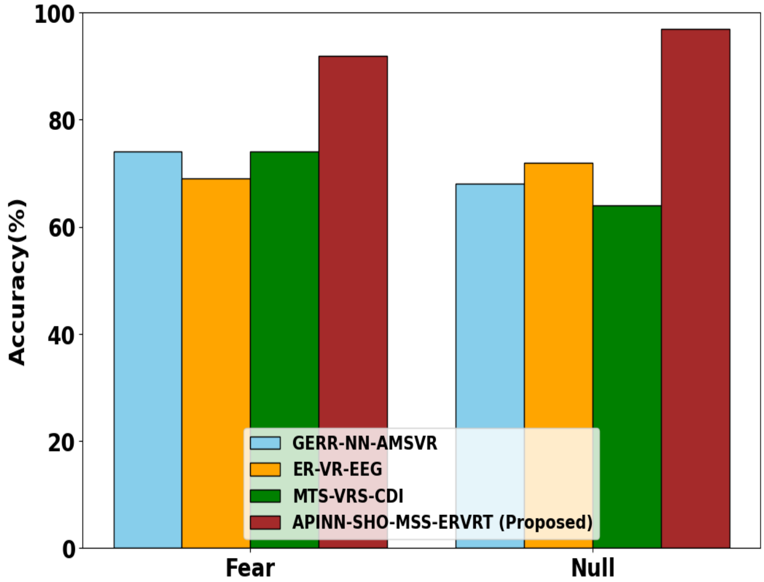

Figure 3: Performance analysis of accuracy

Figure 3 displays performance analysis of accuracy. Here, proposed APINN-SHO-MSS-ERVRT method attains 90% higher accuracy for fear: 97% higher accuracy for null is analysed with existing technique likes GERR-NN-AMSVR, ER-VR-EEG and MTS-VRS-CDI methods respectively.

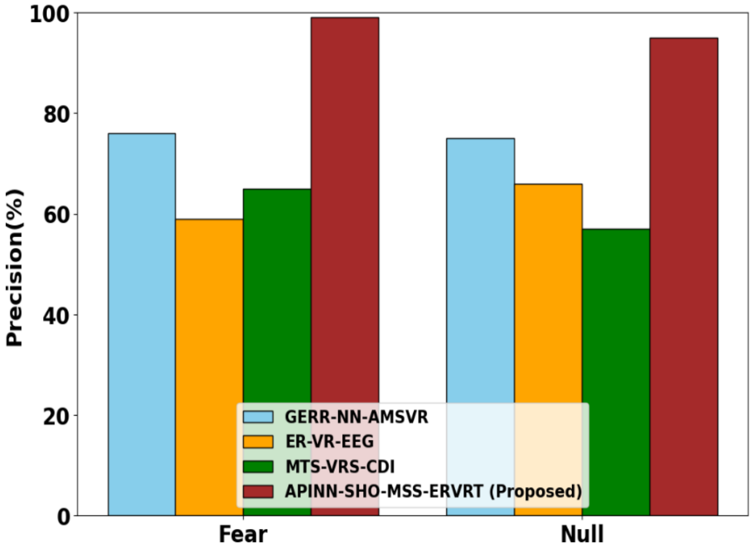

Figure 4: Performance analysis of precision

Figure 4 displays performance analysis of Precision. Here, proposed APINN-SHO-MSS-ERVRT method attains 99% higher precision for fear: 96% higher precision for null is analysed with existing technique likes GERR-NN-AMSVR, ER-VR-EEG and MTS-VRS-CDI methods respectively.

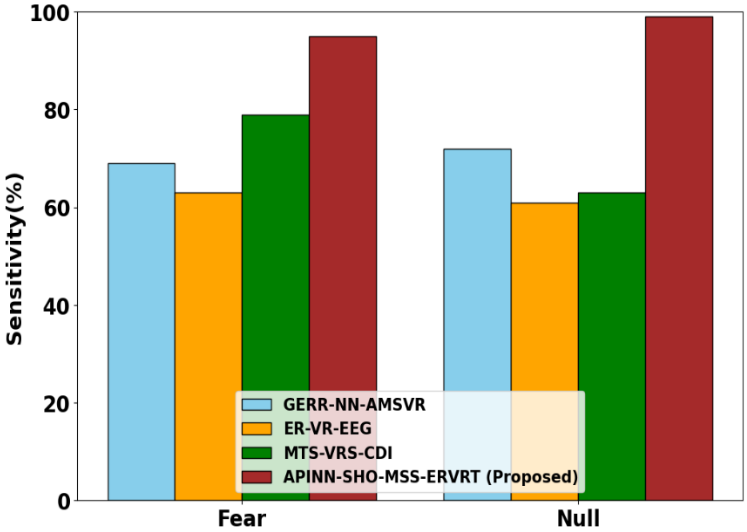

Figure 5: Performance analysis of sensitivity

Figure 5 displays performance analysis of sensitivity. Here, proposed APINN-SHO-MSS-ERVRT method attains 96% higher sensitivity for fear: 96% higher sensitivity for null emotion is analysed with existing technique likes GERR-NN-AMSVR, ER-VR-EEG and MTS-VRS-CDI methods respectively.

Figure 6 displays performance analysis of specificity. Here, proposed APINN-SHO-MSS-ERVRT method attains 97% higher specificity for fear: 96% higher specificity for null emotion is analysed with existing technique likes GERR-NN-AMSVR, ER-VR-EEG and MTS-VRS-CDI methods respectively.

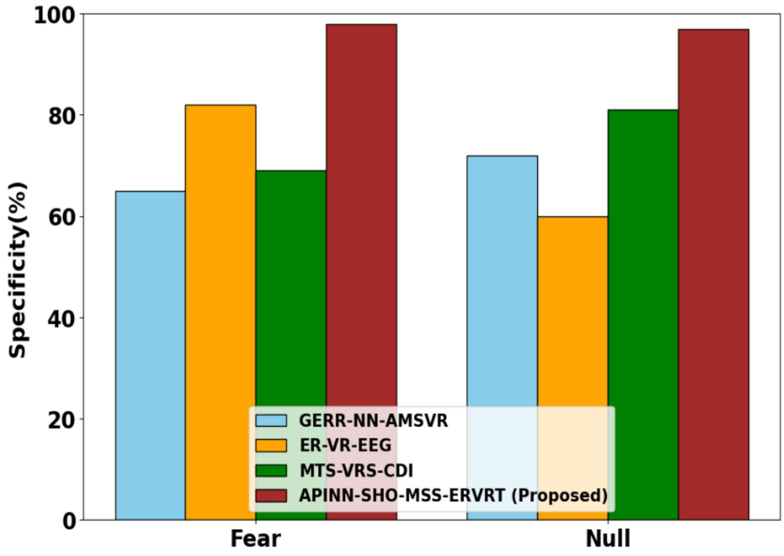

Figure 6: Performance analysis of specificity

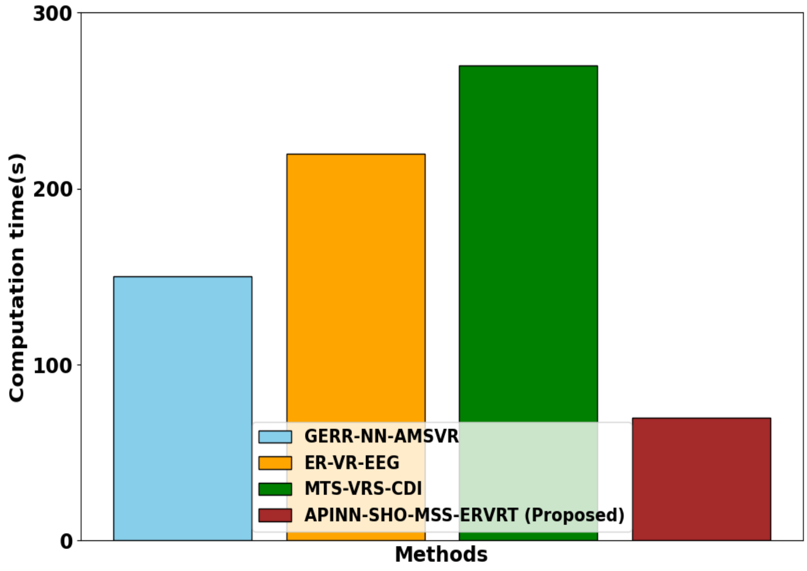

Figure 7: Performance analysis of computation time

Figure 7 displays performance analysis of computational time. Here, proposed APINN-SHO-MSS-ERVRT method attains 70s lower computational time for fear and null emotion is analysed with existing technique likes GERR-NN-AMSVR, ER-VR-EEG and MTS-VRS-CDI methods respectively.

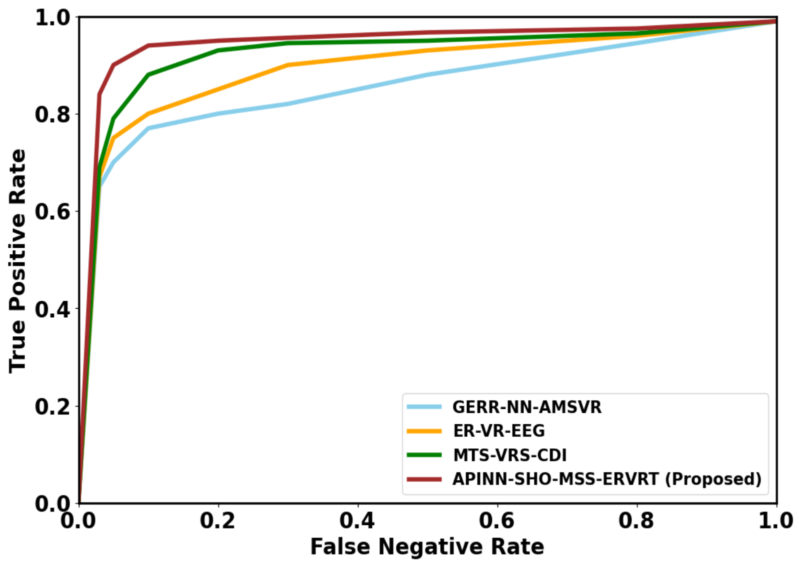

Figure 8: Performance analysis of ROC

Figure 8 portrays ROC analysis. The proposed APINN-SHO-MSS-ERVRT technique then provides a higher ROC in 0.87%, 0.88% and 0.89% than the existing GERR-NN-AMSVR, ER-VR-EEG and MTS-VRS-CDI method.

V. CONCLUSION

In present study, an augmented physics-informed neural networks and seahorse optimization algorithm for music scene simulation and emotion regulation in virtual reality technology (APINN-SHO-MSS-ERVRT) is successfully executed. The proposed APINN-SHO-MSS-ERVRT method is executed in python utilizing head-mounted display (HMD) in real time basis dataset. This paper mainly proposes a method based on the APINN to better achieve for classify facial expression for emotion as fear and null. The results demonstrate the capability of APINN-SHO-MSS-ERVRT in providing an immersive and emotionally impactful VR environment for music scene simulation and emotion regulation. This research not only contributes to the advancement of VR technology but also offers a promising avenue for the development of more realistic and emotionally engaging virtual experiences in the field of music and entertainment. Performance of proposed APINN-SHO-MSS-ERVRT approach contains 97% high accuracy is analysed with existing methods likes GERR-NN-AMSVR, ER-VR-EEG and MTS-VRS-CDI method respectively.

REFERENCE

[1] Colombo, D., Díaz-García, A., Fernandez-Álvarez, J., &Botella, C. (2021). Virtual reality for the enhancement of emotion regulation. *Clinical Psychology & Psychotherapy, 28*(3), 519-537.

[2] Van Kerrebroeck, B., Caruso, G., &Maes, P. J. (2021).A methodological framework for assessing social presence in music interactions in virtual reality.*Frontiers in Psychology*, *12,* 663725.

[3] Benlamine, M. S., Dufresne, A., Beauchamp, M. H., &Frasson, C. (2021). BARGAIN: behavioral affective rule-based games adaptation interface–towards emotionally intelligent games: application on a virtual reality environment for socio-moral development. *User Modeling and User-Adapted Interaction, 31*, 287-321.

[4] Kako, N., Waugh, C. E., & McRae, K. (2023). The future of immersive mood induction in affective science: Using virtual reality to test effects of mood context on task performance. *Affective Science, 4*(3), 570-579.

[5] Brivio, E., Serino, S., Negro Cousa, E., Zini, A., Riva, G., & De Leo, G. (2021). Virtual reality and 360 panorama technology: a media comparison to study changes in sense of presence, anxiety, and positive emotions. *Virtual Reality, 25*, 303-311.

[6] Akdere, M., Jiang, Y., & Lobo, F. D. (2022). Evaluation and assessment of virtual reality-based simulated training: exploring the human–technology frontier. *European Journal of Training and Development, 46*(5/6), 434-449.

[7] Agres, K. R., Schaefer, R. S., Volk, A., van Hooren, S., Holzapfel, A., Dalla Bella, S., ... & Magee, W. L. (2021). Music, computing, and health: a roadmap for the current and future roles of music technology for health care and well-being. *Music & Science*, *4*, 2059204321997709.

[8] Ren, H., Du, Y., Feng, X., Pu, J., & Xiang, X. (2021). Mitigating psychological trauma on adult burn patients based on virtual reality technology of smart medical treatment. *Journal of healthcare engineering, 2021*.

[9] Zhang, Y. (2021). VR technology to the adjustment of piano playing mentality. *Mathematical Problems in Engineering, 2021,* 1-12.

[10] Jiang, A., Gong, Y., Yao, X., Foing, B., Allen, R., Westland, S., ...& Zhu, Y. (2023). Short-term virtual reality simulation of the effects of space station colour and microgravity and lunar gravity on cognitive task performance and emotion.*Building and Environment*, *227*, 109789.

[11] Theodorou, A., Spano, G., Bratman, G. N., Monneron, K., Sanesi, G., Carrus, G., ...&Panno, A. (2023). Emotion regulation and virtual nature: cognitive reappraisal as an individual-level moderator for impacts on subjective vitality. *Scientific Reports, 13*(1), 5028.

[12] Drigas, A., Mitsea, E., &Skianis, C. (2022). Subliminal Training Techniques for Cognitive, Emotional and Behavioral Balance.The Role of Emerging Technologies.*Technium Soc. Sci. J.*, *33*, 164.

[13] Dirin, A., Nieminen, M., Laine, T. H., Nieminen, L., &Ghalebani, L. (2023). Emotional contagion in collaborative virtual reality learning experiences: an esports approach. *Education and Information Technologies, 28*(11), 15317-15363.

[14] Ruan, Y. (2022). Application of immersive virtual reality interactive technology in art design teaching. *Computational Intelligence and Neuroscience, 2022*.

[15] Li, H., Dong, W., Wang, Z., Chen, N., Wu, J., Wang, G., & Jiang, T. (2021). Effect of a virtual reality-based restorative environment on the emotional and cognitive recovery of individuals with mild-to-moderate anxiety and depression.*International Journal of Environmental Research and Public Health*, *18*(17), 9053.

[16] Jiang, D., Jin, D., Zhuang, J., Tan, D., Chen, D., & Liang, Y. (2021). A computational model of emotion based on audio-visual stimuli understanding and personalized regulation with concurrency. *Concurrency and Computation: Practice and Experience*, *33*(17), e6269.

[17] Holt, S. (2023). Virtual reality, augmented reality and mixed reality: For astronaut mental health; and space tourism, education and outreach. *ActaAstronautica, 203*, 436-446.

[18] Baía Reis, A., &Ashmore, M. (2022).From video streaming to virtual reality worlds: an academic, reflective, and creative study on live theatre and performance in the metaverse. *International Journal of Performance Arts and Digital Media*, *18*(1), 7-28.

[19] Shahab, M., Taheri, A., Mokhtari, M., Shariati, A., Heidari, R., Meghdari, A., &Alemi, M. (2022). Utilizing social virtual reality robot (V2R) for music education to children with high-functioning autism. *Education and Information Technologies,* 1-25.

[20] Xiang, Y. (2022). Computer analysis and automatic recognition technology of music emotion. *Mathematical Problems in Engineering, 2022*, 1-9.

[21] Ibáñez, M. L., Miranda, M., Alvarez, N., &Peinado, F. (2021). Using gestural emotions recognised through a neural network as input for an adaptive music system in virtual reality. *Entertainment Computing, 38*, 100404.

[22] Suhaimi, N. S., Mountstephens, J., &Teo, J. (2022). A dataset for emotion recognition using virtual reality and EEG (DER-VREEG): emotional state classification using low-cost wearable VR-EEG headsets. *Big Data and Cognitive Computing, 6*(1), 16.

[23] Feng, Y. (2023). Design and research of music teaching system based on virtual reality system in the context of education informatization. *Plos one, 18*(10), e0285331.

[24] Shen, D., & Zhao, W. (2024). Research on the Construction of Music Animation CAD System Based on Music Form and Emotion Recognition.

[25] Zhao, Y., Guo, M., Chen, X., Sun, J., &Qiu, J. (2023). Attention-Based CNN Fusion Model for Emotion Recognition During Walking Using Discrete Wavelet Transform on EEG and Inertial Signals. *Big Data Mining and Analytics, 7*(1), 188-204.

[26] Wang, C., &Ko, Y. C. (2023). Emotional representation of music in multi-source data by the Internet of Things and deep learning. *The Journal of Supercomputing*, *79*(1), 349-366.

[27] Patel, L. (2022). Music Therapy-Based Emotion Regulation Using Convolutional Neural Network.*In Applications of Machine Learning and Artificial Intelligence in Education* (pp. 73-96).IGI Global.

[28] Lin, Y. D., Tan, Y. K., &Tian, B. (2022). A novel approach for decomposition of biomedical signals in different applications based on data-adaptive Gaussian average filtering.*Biomedical Signal Processing and Control, 71*, 103104.

[29] Hu, Z., Jagtap, A. D., Karniadakis, G. E., & Kawaguchi, K. (2023). Augmented Physics-Informed Neural Networks (APINNs): A gating network-based soft domain decomposition methodology. *Engineering Applications of Artificial Intelligence*, *126*, 107183.

[30] Aribowo, W. (2023). A Novel Improved Sea-Horse Optimizer for Tuning Parameter Power System Stabilizer. *Journal of Robotics and Control (JRC), 4*(1), 12-22.