

RESEARCH ARTICLE

Body-Worn Sensors for Recognizing Physical Sports Activities in Exergaming via Deep Learning Model

MIR MUSHHOOD AFSAR¹, SHIZZA SAQIB¹, MOHAMMAD ALADFAJ²,
MOHAMMED HAMAD ALATIYYAH³, KHALED ALNOWAISER⁴,
HANAN ALJUAID⁵, AHMAD JALAL¹, AND JEONGMIN PARK⁶

¹Department of Computer Science, Air University, Islamabad 44000, Pakistan

²Department of Natural Engineering, College of Science, King Saud University, Riyadh 11421, Saudi Arabia

³Department of Computer Science, College of Sciences and Humanities in Aflaj, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

⁴Department of Computer Engineering, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

⁵Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh 11671, Saudi Arabia

⁶Department of Computer Engineering, Tech University of Korea, Siheung-si, Gyeonggi-do 15073, South Korea

Corresponding author: Khaled Alnowaiser (k.alnowaiser@psau.edu.sa)

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN (ICT Challenge and Advanced Network of HRD) program (IITP-2022-RS-2022-00156326) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation). The authors acknowledge Princess Nourah bint Abdulrahman University Researchers supporting Project number (PNURSP2023R54), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

ABSTRACT Obesity and laziness are some of the common issues in the majority of the youth today. This has led to the development of a proposed exergaming solution where users can play first-person physical games. This research study not only proposes a solution for physical fitness in the form of a game using wearable sensors but also proposes a multi-purpose system that provides different applications when trained for the domain-specific dataset. Critical tasks of gesture recognition and depiction in virtual reality can be applied to many applications in the domains of crime detection, fitness, healthcare, online learning, and sports. In particular, the proposed system enables a user to perform, detect, and depict different gestures in the virtual reality game. First, the system pre-processes input data by applying a median filter to overcome the anomalies. Then, features are extracted through a convolutional neural network, power spectral density, skewness, and kurtosis methods. Further, the system optimizes different features by using the grey wolf optimization. Lastly, the feature set which is optimized is fed to a recurrent neural network for classification. When Compared to the traditional methods, the suggested system gives better results while being easier to use. The IMSporting behaviors (IMSB) dataset includes badminton and other physical activities, the WISDM dataset includes common locomotor motions, and the ERICA dataset which includes a variety of exercises, were used in the experimentation. According to experimental findings, the suggested approach outperformed current methods, which showed detection accuracies of 85.01%, 88.46%, and 93.18% over the IMSB, WISDM, and ERICA datasets, respectively.

INDEX TERMS Convolution neural network, exergaming, grey wolf optimization, recurrent neural network, virtual reality, wearable sensors.

I. INTRODUCTION

According to global statistics, the gaming sector is now the one that is expanding the quickest globally. The gaming

The associate editor coordinating the review of this manuscript and approving it for publication was Alberto Cano¹.

industry, which was estimated to be worth “\$179.7 billion” in 2021, is predicted to increase at a “CAGR of 8.94%” from 2022 to 2027 and reach a value of \$339.95 billion. [1]. Playing games using a mouse and keyboard is old-fashioned, and the need for a new methodology for playing games is essential to further expand this billion-dollar industry.

Apart from that, studies have revealed that people with more screen time are prone to become obese, lazy, sleepless, and tired [2]. Generally, sleeplessness can cause anxiety, mental disorders, and anger issues in adults. Most of these adults have no physical activities, and they are addicted to gaming, which causes a lot of rash effects on their behavior and their bodies [3]. Researchers have been trying to find new ways to make games less harmful and more useful in areas like health, education, sports, and the military. This is because games are used by a lot of people and can have negative effects on them.

Morton Heilig created the first virtual system in 1957 and gave rise to the concept of virtual reality (VR). The proposed gadget was known as sensorama, but later, in 1987, researcher Jaron Lanier came up with the phrase “virtual reality” [4]. VR headsets are now too expensive for the general public to use. Oculus Go, for instance, has the highest pricing at \$545. The Oculus headgear may be purchased for as low as \$249 [5]. There are some systems similar to Meta Quest 2, HTC Vive Pro 2, Sony PlayStation VR, and Wii Nintendo that provides similar and amazing experiences to their users, but they are quite costly. Apart from that these VR devices have long wires attached to computers which restrict users from moving freely in the space around them. There is no gesture recognition system inside these systems. They can only be used for a single purpose at a time. Powerful computing machines are required for generating VR views, and VR illusions and the controllers for playing games only work after buttons on the controllers are pressed. Some VR headsets have cameras that continuously detect the controller movements, and, in this way, they performed an action. In such a system it is common to witness action delay, speed issues, and gesture accuracy issues. Apart from these issues, there is also a configuration issue that is common in almost all VR headsets i.e., login, signup, and connectivity-related issues. Hence, a novel, more efficient, wireless, cost-effective, and the sensors-based wearable system is suggested in order to make a difference. It will help youngsters having obesity and other related issues by encouraging the use of physical exercise in them.

- 1) The proposed approach is multi-purpose and not only limited to gaming, but it can also be used in other domains such as fitness, robotics, drones, sports, and e-learning.
- 2) It connects VR and human physical health through playing games in an indoor environment.
- 3) The system will also enhance the trend of old-fashioned games by introducing a sensors-based wearable device that can control gaming objects precisely through accelerometer data generated from human body gestures.
- 4) The system makes a unique virtual reality experience by using inertial sensors, and it is both affordable and user-friendly.
- 5) The system removes the wires and the need for powerful computers. Instead, the wireless approach is

designed for playing exergames in a virtual reality headset using wearable sensors.

- 6) To anticipate the values, an accurate recurrent neural network (RNN) classifier is utilized.

The proposed approach uses 6 DOF (degrees of freedom) inertial sensors to measure acceleration. This data will be transferred via a transmitter to the computer, where a pre-trained deep-learning model will test via an RNN classifier. Finally, the gaming behavior will be predicted, and the appropriate interface will be loaded into the game. The sensor data controls the avatar gestures and sends the real-time gesture data to a pre-trained model. The motions and activity recognition may be shown over the personal computer (PC) at the same time as the user watches in their VR headset.

Numerous head-mounted displays for VR (HMD-VR) exergames are being utilized widely for rehabilitation and to aid the recovery of patients [6]. In addition, VR games have several uses in the medical field, such as preventative health and well-being along with medical evaluations in clinical treatments [7]. VR gaming has a wide range of educational uses as well. The use of games in a classroom increases students’ engagement and motivation, whereas first-person VR game usage will keep students active and healthy while having fun [8]. There has been a resurgence in the research related to delivering therapy using VR gaming systems. VR games offer potential in the treatment of numerous ailments, including post-stroke, Parkinson’s disease, and others [9].

The following are our system’s significant contributions:

- Our suggested system provides a cost-effective solution to very serious issues like childhood obesity and other health-related problems.
- Similar systems exist in the realm of VR, but they are exceedingly expensive and out of the reach of the common man. As a result, the suggested method creates a product that allows individuals to keep in shape while playing games in an indoor environment.
- Our hardware-based system is being developed and tested. Therefore, it is a reliable solution for physical health and other applications.
- To make our system efficient and affordable, we are incorporating a straightforward and inexpensive VR device.
- The system outperformed other already made approaches in terms of accuracy rates.

The remaining article is structured as follows: Section II looks at similar research in the area of VR sports action detection using sensors and cameras. The suggested technique is thoroughly explained in Section III. The many datasets that were utilized to verify the effectiveness of the suggested strategy and the outcomes of those tests are described in Section IV. The paper’s concluding part and goals for the future are included in Section V.

II. RELATED WORKS

For deep learning and machine learning-based systems employing a range of sensors, including inertial measurement units (IMU), cameras, and other fused sensors, several methodologies have been proposed by numerous academics. This section reviews the research on camera-based and wearable sensors-based systems.

A. VIRTUAL REALITY EXERGAMING WITH WEARABLE SENSORS

Virtual reality games and sensors have been utilized in several applications in recent years. I. Paraskevopoulos and E. Tsekles suggested a system in [10] that incorporated motion capture technology that was more affordable, flexible, and off-the-shelf with video games that were specially designed to meet the needs of Parkinson's Disease (PD) rehabilitation. However, they used larger controllers for the game, and D. Fitzgerald et al. developed a computer game for VR in order to lead an athlete through several advised rehabilitation activities [11].

In an effort to enhance physical performance while avoiding or treating musculoskeletal disorders, certified professionals have prescribed training programs to athletes. With the use of serious games and virtual environments, Mondragón Bernal et al. developed and assessed a system for teaching power distribution operation. Building information modelling from a "115 kV" substation was utilised to create a scenario with high technical details suited for professional training in the VR simulator. [12].

Immersive 3D virtual worlds and serious games, or video games meant to be educational, are both growing in popularity. Serious games have just lately been tested for healthcare education. Following a review of educational philosophies highlighting the importance of serious games and virtual simulations as teaching aids, Ma et al. examined various instances of early teaching models and evaluated procedures in their study [13]. They further made recommendations for how to assess their worth in a learning environment.

VR technologies are gaining popularity as a way to model, evaluate, and improve the assembly process. Abidi et al. discussed the development of a "haptic virtual reality platform" for virtual assembly planning, execution, and evaluation. The technology enables real-time handling and interaction with virtual components. To examine the advantages and disadvantages of combining haptics with physical-based modeling, the system consists of several software programs including Open Haptics, PhysX, and OpenGL/GLUT libraries [14].

B. VIRTUAL REALITY EXERGAMING WITH CAMERA

Researchers have used a variety of approaches while employing camera-oriented VR systems. To put handicapped persons at ease, Gerling et al. [15] devised a system that employed the "Kinect v2 depth camera" to evaluate the movement of wheelchair and created two Unity VR games. The study's findings were highly encouraging since an immersive

VR experience for persons with disabilities proved to be a wonderful experience for them.

Stomp Joy was a camera based VR game which was specific to one task which was rehabilitation after lower limbs stroke developed by Xu et al. [16]. Sangmin et al. [17] created a VR game in Unity for the A-Camera and A-Visor. They displayed cutting-edge head-mounted virtual reality controllers that enthusiasts could easily construct for themselves using corrugated cardboard, an Arduino, and sensors.

Another study expands the use of VR in manufacturing by incorporating ideas and research from training simulations into the evaluation of assembly training efficacy and training transfer [18]. A research was carried out by Abidi et al. to evaluate and contrast the virtual assembly training method used for the first Saudi Arabian vehicle prototype. Three learning contexts were examined in this study: conventional engineering, computer-aided design environments, and immersive VR. Random assignments were made to the various training contexts for 15 university students [19].

Industrial design, planning, and prototyping are more successful and economical when done in VR. The study conducted by Abdulrahman M. Al-Ahmari and colleagues and reported in this paper was primarily concerned with creating a "virtual manufacturing assembly simulation system" that tackles the limitations of VR settings. Using a virtual environment, the proposed system builds an interactive workbench for looking at different assembly options and teaching how to put them together [20].

Dissimilar from the above systems in literature, our framework proposed wireless body-worn sensors for controlling 3D game objects, a deep learning-based approach for recognizing sports behaviors, and activity recognition for an indoor gaming activity that is used to predict the label of the considered game activity.

III. OUR APPROACH

This section elaborates on the proposed architecture for active monitoring of the sports-related activities of humans and their conversion into IMU data for its recognition in gaming activity. Such recognition can be very helpful for demonstrating complex sports behaviors in VR. It is also helpful for artificial intelligence-based gaming objects to recognize a set of sports behavior.

Fig. 1 shows a description of the overall system. According to the figure, wearable sensors generate accelerometer data for particular gaming activity, which is performed by humans using the body-worn sensors-based device. The publicly available benchmarked datasets were used to evaluate the proposed system. The data was pre-processed, and the corresponding features were extracted. A well-known approach of grey wolf optimization (GWO) was efficiently used for feature optimization over the extracted features. RNN classifier is further applied for the classification of these optimized features to recognize the gestures performed by humans. Lastly, the predicted gesture was depicted in the

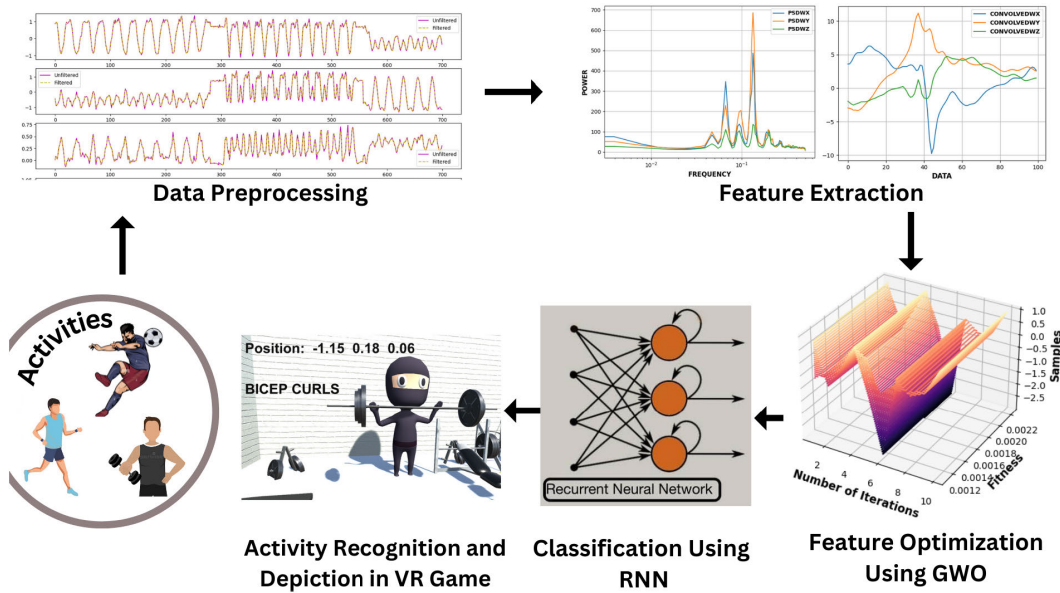


FIGURE 1. Overall flow of the architecture of system.

VR game and the user can play the game as first person in the VR world.

A. DATA PRE-PROCESSING

The ERICA dataset contained 3-axes accelerometer data, which is obtained by integrating mpu6050 device with Arduino. Each value in the data portrayed a certain position of a body part where the body-worn device is attached in 3D space. Hence, each value of the dataset is equally important in the proposed approach, but the data may contain irrelevancy, irregularity, inconsistencies, and repetition that can affect the proposed model and generate false predictions [21]. It is called the noise in data that needs to be calibrated before it is fed to the classifier. To reduce this noise, the data was divided into frames, which improve the quality of data and ensure the employment of signal enhancement for data filtering to identify undesired features. It also helps in avoiding irrelevancy, irregularity, inconsistencies, and repetition issues [22]. Then, a 3rd order median filter has been used to cancel the significant noise artifact. However, the median filter is a “non-linear filter”, which removes “speckle noise” from a given signal. It provides the median of the signal in a required size and outperformed the “low pass filter” because of reducing noise with keeping the original signal. The median filter is calculated as:

$$Med(X) = \begin{cases} x \left[\frac{n}{2} \right] & \text{if } n \text{ is even} \\ \frac{X[n - 12] + X[n + 12]}{2} & \text{if } n \text{ is odd} \end{cases} \quad (1)$$

where n is the count of values and X is the “ordered set of values” in the dataset [23]. ERICA is a lightweight dataset containing sensor data from three different gym exercises. Every data value is important, and a small proportion of data

requires pre-processing. For this purpose, median filter was applied. Fig. 2 shows the result of the median filter applied over ERICA dataset in the form of filtered and unfiltered signals. The dotted line displays the filtered wave, and the solid line shows the unfiltered data.

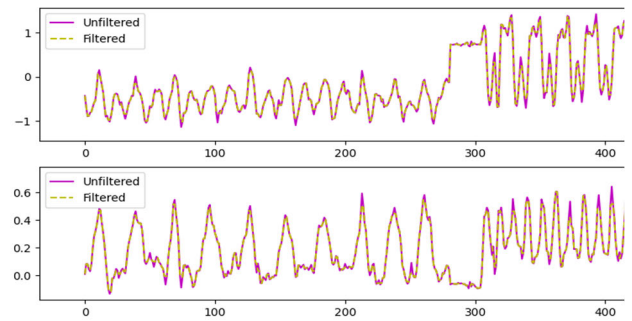


FIGURE 2. Filtered and Unfiltered 3-axes accelerometer signal of motion sensor from ERICA dataset.

B. FEATURES EXTRACTION

After pre-processing, the data is further subjected to features extraction methods to collect unique features from the data. These features were then passed to the features optimization module for further processing. We have utilized four different feature extraction methods including power spectral density (PSD), skewness, kurtosis, and convolutional neural network (CNN).

1) POWER SPECTRAL DENSITY (PSD)

PSD determines the “power of a signal as a function of frequency” by using the signal’s per unit frequency [26]. Watts per hertz (W/Hz) is a typical unit of measurement for PSD. Fast Fourier Transform (FFT) has it’s function to

produce the Discrete Fourier Transform $X(w_i)$ of a signal, where w_i gives the frequency point. Following is the equation to calculate PSD:

$$P(w_i) = \frac{1}{N} |X(w_i)|^2 \tag{2}$$

The average power P_x can be explained as $S(f)df$, where the function $S(f)$ is used to express the power of each minimal limit unit's frequency component, it will be referred to as the PSD [27].

$$\begin{aligned} \bar{P}_x &= \lim_{t_0 \rightarrow \infty} \frac{1}{2t_0} \int_{-\infty}^{\infty} E[|X_{t_0}(f)|] df \\ &= \int_{-\infty}^{\infty} \lim_{t_0 \rightarrow \infty} \frac{E[|X_{t_0}(f)|]}{2t_0} df \end{aligned} \tag{3}$$

$$S(f) = \lim_{t_0 \rightarrow \infty} \frac{E[|X_{t_0}(f)|]}{2t_0} \tag{4}$$

PSD shows the energy of fluctuations relative to frequency. In other words, it demonstrates the frequencies whose certain variants are strong and those frequencies that are weak. PSD is applied over three columns of the dataset, which contains accelerometer data in the x, y, and z-axis. This data is collected concerning time domain. When PSD is applied over the ERICA dataset, unique features are extracted from the frequency domain. Fig. 3 elucidates the results, which show the signals' power vs frequency. This helped in explaining the distribution of data between multiple frequency domains.

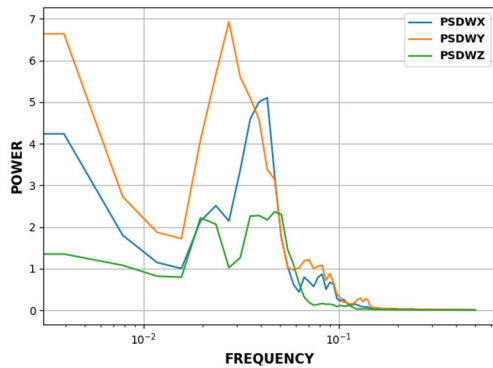


FIGURE 3. Features Extraction using PSD over ERICA Dataset.

2) SKEWNESS

Skewness can be defined as a slight deviation from the “normal distribution or symmetrical bell curve” in a collection of data. There are various conditions. If the curve is seen inclined towards right or left, data has skewness. Skewness can have zero, negative, positive, or undefined values [28]. Skewness can be calculated as:

$$Skewness = \frac{\sum_{i=1}^N (X_i - X)^3}{(N - 1) * \sigma^3} \tag{5}$$

where N is the total sample count in data, X_i is the value of the samples, X gives the mean, and σ shows the standard deviation. Skewness scores are calculated between -3 and $+3$.

If a value of skewness in a distribution is more than 1 or lesser than -1 , it is said to be strongly skewed, if it is between 0.5 and 1 or -0.5 and -1 , it is said to be mildly skewed. Additionally, the distribution is said to be very symmetrical if the value of skewness is between -0.5 and 0.5 [29]. The skewness of the “ERICA dataset” is shown in Fig. 4.

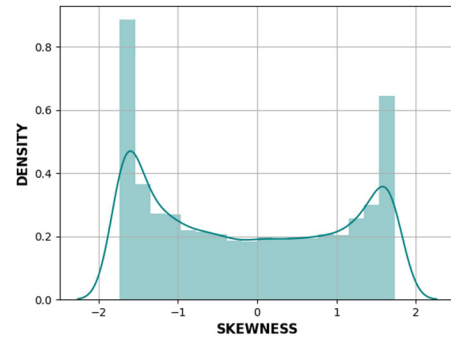


FIGURE 4. Features Extraction using Skewness over ERICA Dataset.

3) KURTOSIS

The final feature extraction method used in our proposed system was kurtosis. It can be termed as the cumulative weight of a distribution's tails in relation to its middle point. A set of essentially normal data may be visualized using a histogram to reveal a bell-shaped peak with the majority of data falling within three standard deviations (plus or minus) of the average [30]. The equation below is a mathematical formula for kurtosis:

$$Kurtosis = \frac{\sum_{i=1}^N \frac{(X_i - X)^4}{N}}{\sigma^4} \tag{6}$$

where N is the total sample count in data, X_i is the value of the samples, X denotes mean and σ denotes standard deviation. A metric in the statistics field is called kurtosis, which expresses the measure of divergence of a distribution's tail from those of a normal distribution. Kurtosis, thus, tells if a particular distribution's tails include greater values. [31]. Fig. 5 displays the features selected from ERICA dataset via the kurtosis-based features extraction method.

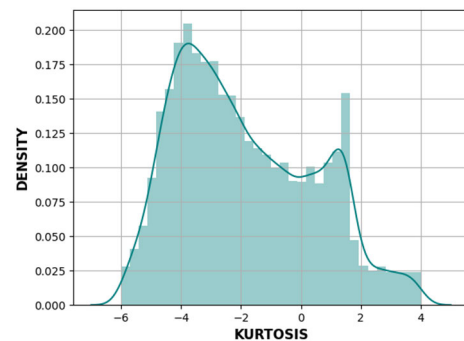


FIGURE 5. Features Extraction using Kurtosis over ERICA Dataset.

4) CONVOLUTIONAL NEURAL NETWORK

To extract features, we applied CNN over the filtered data that collected the features, while a dissimilar neural network classified the features. The feature extraction network uses input data. Three layers make up a neural network; input, output, and hidden layer. The neurons in CNN are similar to the neurons of human body. The way they take the input, analyze it and send the response to the body is similar. Data arrays are accepted as input by the input layer. CNN's may have several hidden layers that employ mathematics to extract characteristics from the provided data. Several instances of this include convolution, pooling, corrected linear units, and fully connected layers. Formally, the following formulas were used to extract key features map via one-dimensional convolution operation:

$$\begin{aligned}
 a_j^{(l+1)}(\tau) &= \sigma(b_j^l + \sum_{f=1}^{F^l} K_{jf}^l(\tau) * a_f^l(\tau)) \\
 &= \sigma(b_j^l + \sum_{f=1}^{F^l} [\sum_{p=1}^{p^l} K_{jf}^l(p) * a_f^l(\tau - p)])
 \end{aligned}
 \tag{7}$$

where $a_j^l(\tau)$ denotes the “feature map j in layer l ”, σ is a “non-linear function”, F^l gives the “number of feature maps in layer l ”, K_{jf}^l displays the “kernel convolved over feature map f in layer l ” to form the “feature map j in layer $(l+1)$ ”, p^l is the “length of kernels in layer l ” and b_j^l provides a “bias vector” [25]. The datasets have some activities and when we pass the dataset columns to CNN based feature extraction method, we will get a unique feature. The features extracted of CNN on the ERICA dataset are demonstrated by three different colors in Fig. 6.

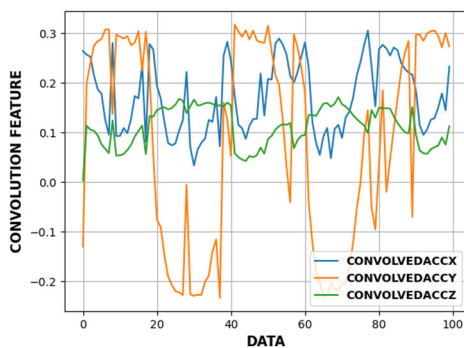


FIGURE 6. Features Extraction using CNN over ERICA Dataset.

The algorithm for data preprocessing and features extraction methods described above is shown in Algorithm 1:

C. FEATURE OPTIMIZATION

The dimensionality reduction method was employed next over multiple datasets to divide and reduce the vector size to make them more manageable groups. A necessary step in the “feature selection process” of a predictive model is to make the feature array smaller and use only the features that are important in certain cases. Fewer input variables

Algorithm 1 “Data Preprocessing and Feature Extraction”

Input: Dataset

Output: Filtered dataset and extracted features

```

For i from 0 to length of dataset
  For j from 0 to length of dataset[i]
    Set dataset[i][j] to median(dataset[j])
For i from 0 to length of dataset
  For j from 0 to length of dataset[i]
    Apply 1 by 1 convolutional layer
    Extract and add result feature map to extracted features
For i from 0 to length of data
  For j from 0 to length of data[i]
    Compute Skewness of data
    Extract and add result feature map to extracted features
For i from 0 to length of dataset
  For j from 0 to length of data[i]
    Compute Kurtosis of data
    Extract and add result feature map to extracted features
For i from 0 to length of dataset
  For j from 0 to length of data[i]
    Compute Power Spectral Density of data
    Extract and add result feature map to extracted features
    
```

might improve the model’s performance while simultaneously minimizing the computational expense of modeling. Modern advanced feature selection techniques choose a subset of essential features utilizing the strength of optimization algorithms to improve classification outcomes [32].

Numerous controlling factors were used by the majority of optimization algorithms including the genetic algorithm. They must be tuned for improved performance. The optimization step of the proposed system uses the GWO technique, which is a novel meta-heuristic optimization technique. Its guiding premise is to model cooperative hunting behavior similar to that of grey wolves in the wild. Compared to other techniques, GWO has a unique model structure. The goal of the GWO is to use population interaction to locate the best areas of the complicated search space [33]. The pack finds its prey by changing the positions of the individual agents with respect to the prey location as follows:

$$X(t + 1) = X_p(t) + A.D \tag{8}$$

where X_p is the prey position, X is the grey wolf position t is the iteration, the dot operator shows vector entry-wise multiplication, and D is defined as:

$$D = |C.X_p(t) - X(t)| \tag{9}$$

where coefficient vectors (A and C) are computed as follows:

$$A = 2a.r_1 - a \tag{10}$$

$$C = 2r_2 \tag{11}$$

where “ r_1 and r_2 ” are random vectors with ranges $[0, 1]$ and a is a linear function of the number of exploration and

exploitation repetitions. All wolves have the same value for a . According to these calculations, a wolf can modify its location in the search area around its prey at any random time. The entire pack engages in hunting based on information provided by the beta, alpha, and delta wolves, who are aware of the whereabouts of the prey, as stated in the following:

$$X(t + 1) = \frac{X_1 + X_2 + X_3}{3} \tag{12}$$

where X_1, X_2, X_3 are calculated as follows:

$$X_1 = |X_\alpha - A_1 \cdot D_\alpha| \tag{13}$$

$$X_2 = |X_\beta - A_2 \cdot D_\beta| \tag{14}$$

$$X_3 = |X_\gamma - A_3 \cdot D_\gamma| \tag{15}$$

where X_1, X_2, X_3 are best results and $D_\alpha, D_\beta, D_\gamma$ are calculated as [34]:

$$D_\alpha = |C_1 \cdot X_\alpha - X| \tag{16}$$

$$D_\beta = |C_2 \cdot X_\beta - X| \tag{17}$$

$$D_\gamma = |C_3 \cdot X_\gamma - X| \tag{18}$$

Fig. 7, Fig. 8, and Fig. 9 display the visualization of the fitness value or best solution with the number of iterations by applying GWO over the ERICA, IMSporting Behaviors (IMSB), and WISDM datasets, respectively.

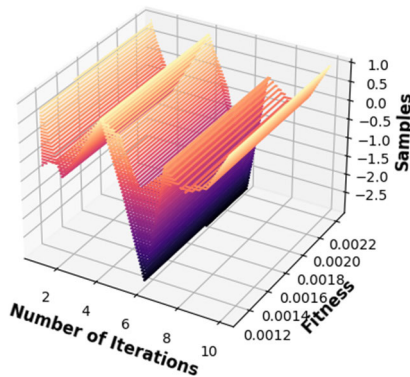


FIGURE 7. Grey Wolf Optimization Visualization over ERICA Dataset.

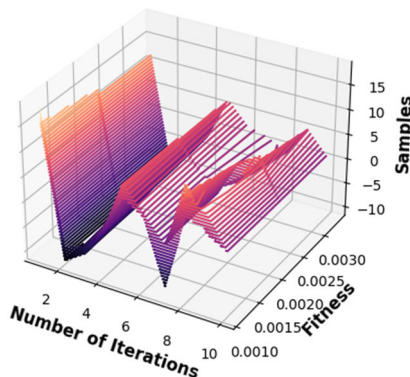


FIGURE 8. Grey Wolf Optimization Visualization over IMSB Dataset.

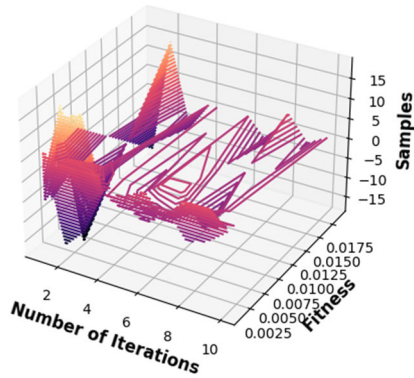


FIGURE 9. Grey Wolf Optimization Visualization over WISDM Dataset.

D. RECURRENT NEURAL NETWORK

The classification of interactions has been carried out by a classifier named RNN and it is the last phase of the proposed system. The RNN is a fast, robust, and one of the most reliable neural networks currently available due to its unique feature called internal memory [35]. Fig. 10 shows a visual representation of the RNN for the ERICA dataset, which has a 5 hidden layers of LSTM, one input layer of LSTM and an output-dense layer.

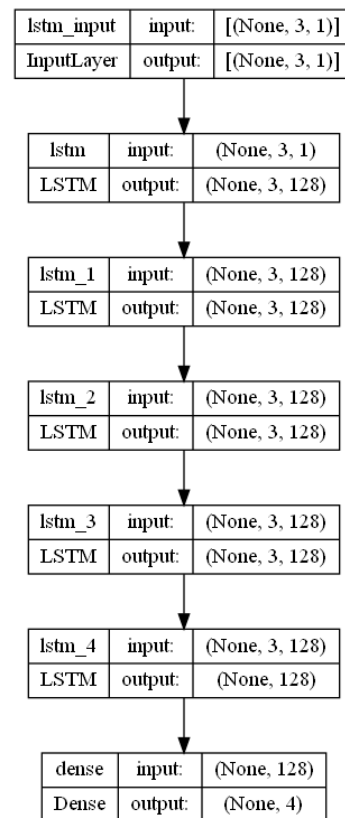


FIGURE 10. Recurrent Neural Network Visualization over ERICA Dataset.

$x(t)$ is used as input at any time step t in RNN. One-hot vector x_1 , for instance, may correspond to a word in a text. $H(t)$ serves as the network's "memory" and represents a

concealed state at time t . The hidden state of the previous time step and the current input is used to determine $h(t)$. The RNN has connections for hidden input, hidden-to-hidden recurrent connections, and hidden-to-output connections, all of which were parameterized by a “weight matrix U ,” “weight matrix W ,” and “weight matrix V ,” respectively. Over time, all of these “weights (U, V, W)” were shared. By $o(t)$, the network output is shown. [36]. The following set of equations can be used to model the RNN forward pass:

$$a^{(t)} = b + Wh^{(t-1)} + Ux^{(t)} \quad (19)$$

$$o^{(t)} = c + Vh^{(t)} \quad (20)$$

$$h^{(t)} = \tanh(a^{(t)}) \quad (21)$$

$$\hat{y}^{(t)} = \text{softmax}(o^{(t)}) \quad (22)$$

The equations shown above are an illustration of a recurrent network that converts an “input sequence” into an identically lengthened “output sequence”. The sum of the losses across all the time steps would thus be the overall loss for a particular series of x and y values. We suppose that the vector of probabilities over the output was obtained by using the “outputs $o(t)$ ” in the softmax function [37]. We also suppose that, given the current input, the loss L is the “negative log-likelihood” of the genuine goal $y(t)$. The algorithm for feature optimization and classification by RNN is shown below in Algorithm 2:

Algorithm 2 Grey Wolf Optimization and Classification by RNN

Input: Dataset with extracted features

Output: Optimized features and classification result

Initialize the grey wolf population $X_i, i = 1, n$

Initialize a, A and C

Find each search agent fitness

X_α = “Best Search Agent”

X_β = “Second Best Search Agent”

X_γ = “Third Best Search Agent”

while $t <$ iterations maximum number **do**

for each search agent **do**

 Randomly initialize r_1 and r_2

 Update current search agent position

 Update a, A and C

 Find all search agent fitness

 Update X_α, X_β and X_γ

$t = t + 1$

Return X_α

Initialize X to dataset[features]

Initialize Y to dataset[classes]

Split X and Y to train and test data by 80% and 20% ratio

Set RNN to sequential mode

Add an “LSTM” layer with 128 units, and `return_sequence = True`

Add an “LSTM” layer with 128 units, and `return_sequence = True`

For each i which is [True, True, False]

 Add LSTM layer with `return_sequence = i`

 Add Dense layer with softmax activation function and units = `outputclasses.length`

 Compile RNN with adam optimizer and

`categorical_crossentropy` loss function

 Run RNN with 100 epochs and batch size = 8

IV. EXPERIMENTAL SETUP AND RESULTS

The experimental results part discusses about the benchmark datasets which we have used in our study, and experimental setup for the proposed system, statistical evaluation, results by implementing the proposed architecture, and comparison of this work with other body-worn systems.

A. DATASET DESCRIPTION AND METRICS

High-intensity interval training (HIIT) can be delivered through exergaming in virtual reality, which combines exercise and enjoyment. We have used three different datasets in this study i.e., IMSB [38], [39], WISDM [40], and ERICA [41].

The first dataset used is IMSB created by Intelligent Media Centre, Air University, Islamabad. There are six different sports-related interactions in the IMSB dataset including table tennis, football, cycling, badminton, basketball, skipping. Three tri-axial accelerometers were attached to the knee, wrist, and lower neck regions of the subjects. The dataset contained motion data from participants performing six different sports as mentioned above. The dataset also contained 120 data sequences with varying exercise time period from the 40s to 60s. A total of 20 subjects were engaged in repetitive behaviors. Fig. 11 shows the plots of raw data from three accelerometers in x, y , and z coordinates for *basketball* and *badminton* behavior.

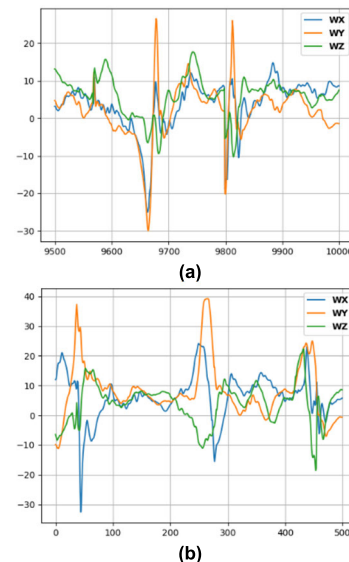


FIGURE 11. Raw data plots of three accelerometers showing (a) basketball and (b) badminton behavior from IMSB dataset.

The WISDM is the next dataset that was used. The activities included in the dataset are running, sitting, standing, going up and down stairs, and so on. There are 1,098,207 total samples in this dataset, including 424,400 walking samples, 342,177 jogging samples, 122,869 upstairs samples, 100,427 downstairs samples, 59,939 sitting samples, and 48,395 standing samples. Fig. 12 represents the *walking* and *jogging* behavior over the WISDM dataset.

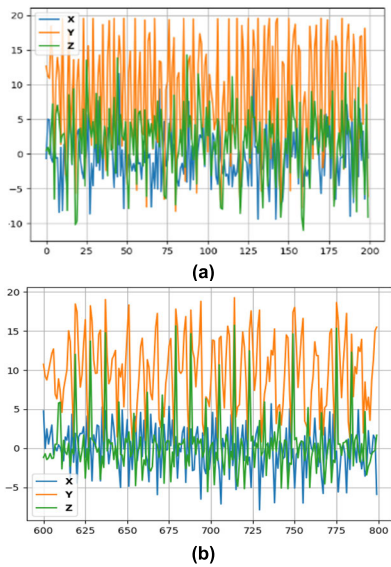


FIGURE 12. Raw data plot of one accelerometer displaying (a) walking and (b) jogging behavior from WISDM dataset.

The third dataset utilized was ERICA, which was created for the automated tracking and analysis of exercise activities at the individual level. This dataset was acquired as part of the development of a low-cost, pervasive digital personal training system that combines affordable IoT sensors linked to dumbbells with personal wireless ear-worn devices (earables) to enable fine-grained tracking of a person’s free-weight exercise training. Total of 324 samples from three separate free-weight workouts carried out by 27 subjects are included in this dataset. The activities performed in a dataset are *biceps curls*, *lateral raises*, and *triceps extensions*. Fig. 13 shows

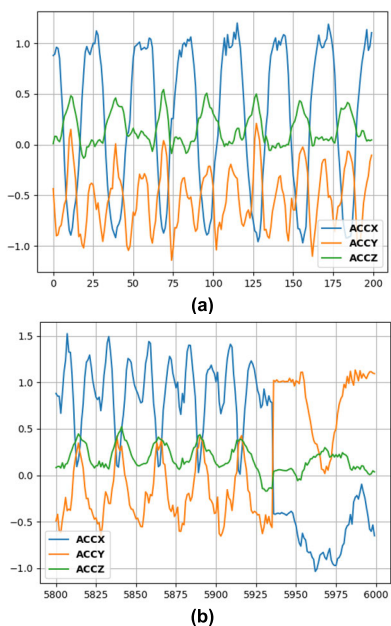


FIGURE 13. Raw data plot over one accelerometer showing (a) Biceps curls and (b) lateral raises behavior from ERICA dataset.

the *biceps curls* and *lateral raises* behavior of the ERICA dataset.

B. EXPERIMENTAL SETUP

This section gives a brief description of the implementation of our proposed system. A 3D VR game will be made in Unity3d which when playing is visible to the person by a screen attached to the VR headset and on the PC as well through screen casting method [42].

A 6 DOF sensor MPU6050 is utilized for capturing the motion data of the human body during exergaming. Arduino nano read the sensor data and send it via nrf24l01 to the receiving point, which is computer, using serial communication. Further, the User Datagram Protocol (UDP) communication was used to establish a secure connection using specific IP based path to convey data from host to destination and in our system, it helps in establishing a secure channel connection between computer and VR game to convey the predicted results from the proposed model to the game [43]. The results are presented in the form of confusion matrices and a precision-recall table. On a Windows 10 computer running the Unity3D and Python programming languages, with 16 GB of RAM, and a Core i7-7500U CPU running at 2.70 GHz, all processing and experimentation were carried out. nRF24L01 and MPU6050 on an Arduino Nano were utilized to create body-worn. Finally, using the IMSB, WISDM, and ERICA datasets, the suggested system’s performance is compared to the precision of other already made systems.

C. STATISTICAL EVALUATION

We will go over the experimental findings of the suggested model using the publically accessible IMSB, WISDM, and ERICA datasets in this part. We will also be comparing the results with other state-of-the-art methods.

1) IMSporting BEHAVIORS DATASET

Regarding the IMSB dataset, confusion matrices are used to demonstrate interaction recognition for different dataset types. A “confusion matrix” measures the effectiveness of a classifier on the basis of “true positives, false positives, true negatives, and false negatives”. [44]. The amount of true positives reveals the correctly detected classes represented on the matrix diagonal. Table 1 demonstrates the confusion matrix over the IMSB dataset.

TABLE 1. Confusion matrix over imsporting behaviors dataset.

Classes	BDN	BTB	Cg	FB	Sg	TT
BDN	0.84	0.02	0.00	0.14	0.00	0.00
BTB	0.01	0.97	0.00	0.02	0.00	0.00
Cg	0.00	0.00	1.00	0.00	0.00	0.00
FB	0.16	0.09	0.00	0.68	0.00	0.07
Sg	0.02	0.01	0.00	0.07	0.88	0.02
TT	0.02	0.00	0.00	0.05	0.00	0.93
Mean Accuracy = 85.01%						

*BDN = Badminton; BTB = Basketball; Cg = Cycling; FB = Football; Sg = Skipping; TT = Table Tennis

The confusion matrix in Table 1 shows that a few interaction classes having related activity types are confused with each other. The mean accuracy achieved by applying the classifier is 85.01%. The recall, precision, and F1-score for different classes over the IMSB dataset are shown in Table 2.

TABLE 2. Classification results over importing behaviors dataset.

Interaction Classes	Precision	Recall	F1 Score
Badminton	0.82	0.94	0.88
Basketball	0.98	0.97	0.98
Cycling	1.00	1.00	1.00
Football	0.89	0.80	0.84
Skipping	0.99	0.95	0.97
Table Tennis	0.97	1.00	0.99
Mean	0.94	0.94	0.94

Hence, an accurate system was developed, which was able to recognize each game with high precision [45]. The results of the gaming interface using the IMSB dataset are presented in Fig. 14.

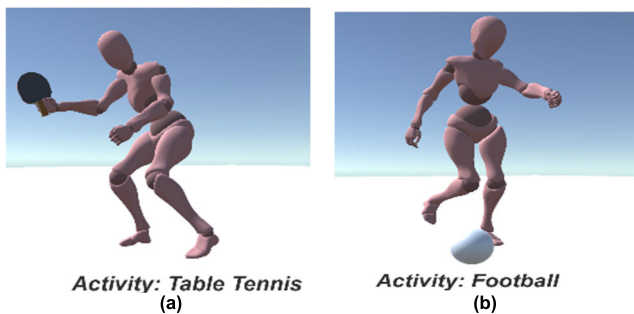


FIGURE 14. Gaming Interface for predicted gaming activities (a) Table tennis and (b) Football using IMSB dataset.

Table 3 represents a comparison of classifier results when compared to the other state-of-the-art methods. Many systems were developed that are similar to the proposed method. A comparison of the systems developed over the IMSB dataset has been displayed in Table 3. The results from other

TABLE 3. Importing behaviors dataset accuracy (%) comparison with state of the art methods.

Method	Recognition Accuracy (%)
MLP [49]	75.90
Classification using LSVM [48]	80.0
Artificial Neural Networks [46]	82.83
Random Forest [47]	83.42
Proposed	85.01

systems using different classifiers over the IMSB dataset are compared with the proposed method. Authors have used artificial neural networks algorithm along with features extraction methods and got an accuracy of 82.83% [46]. Another system utilized a random forest algorithm and achieved an accuracy of 83.42% [47]. One of the other proposed systems used for classification through LSTM with multi-fused features extraction achieved an accuracy of 80% [48]. Lastly, a system using the multi-layer perceptron (MLP) achieved an accuracy of 75.90% [49]. The system proposed in this paper has achieved an accuracy of 85.01% that outperformed all the previously proposed systems.

2) WISDM DATASET

For WISDM dataset, the results from RNN classifier over the optimized features produced the confusion matrix that is shown in table 4. It is clear that the result of interaction classes was efficient and acceptable. A small amount of data from some interaction classes was confused i.e., achieving a mean accuracy of 88.46%.

TABLE 4. Confusion matrix of WISDM dataset.

Classes	WLK	Jg	Up	Dw	Stg	Std
WLK	0.82	0.13	0.02	0.03	0.00	0.00
Jg	0.05	0.88	0.03	0.04	0.00	0.00
Up	0.10	0.31	0.48	0.10	0.00	0.01
Dw	0.09	0.35	0.13	0.42	0.00	0.01
Stg	0.00	0.01	0.00	0.00	0.99	0.00
Std	0.00	0.03	0.02	0.00	0.00	0.95
Mean Accuracy = 88.46%						

*WLK = Walking; Jg = Jogging; Up = Upstairs; Dw = Downstairs; Stg = Sitting; Std = Standing

The ratio of “correct positive predictions” to the “total positives” is precision while the recall is the “true positive rate”, and it is the ratio of “correct positive” to the “total predicted positives”. The average of precision and recall is the F1 score [50]. The precision, recall, and F1-score for classes of the dataset are given in Table 5.

TABLE 5. Classification results over WISDM dataset.

Interaction Classes	Precision	Recall	F1 Score
Walking	0.87	0.82	0.84
Jogging	0.76	0.88	0.82
Upstairs	0.62	0.48	0.54
Downstairs	0.54	0.42	0.47
Sitting	0.99	0.99	0.99
Standing	0.93	0.95	0.94
Mean	0.78	0.76	0.77

An accurate system was developed that can recognize each game with high precision. The results of the gaming interface using the WISDM dataset are shown in Fig. 15.

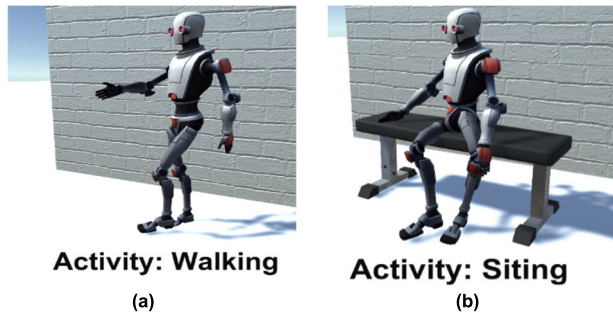


FIGURE 15. Gaming Interface for showing predicted gaming activities (a) Waling and (b) Sitting using WISDM dataset.

Table 6 represents a comparison of RNN results over the WISDM dataset with other state-of-the-art models. The results were compared with the conventional system and compared with the proposed method. Authors in [51] have used a reweighted genetic algorithm and achieved an accuracy of 87.75%. Another proposed system in [52] utilized MLP and achieved a 75.09% accuracy rate. Another system applied classification through CNN and achieved 75.90% accuracy [53]. The Hoeffding tree algorithm has achieved an accuracy rate of 75.54% in another proposed method [53]. Lastly, a system utilized support vector machines and achieved 82.77% accuracy [38]. The proposed system in this paper has outperformed these systems by achieving an accuracy rate of 88.46%.

TABLE 6. Wisdm dataset accuracy (%) comparison with state of the art methods.

Methods	Recognition Accuracy (%)
MLP [52]	75.09
Hoeffding tree [53]	75.54
CNN [53]	75.90
SVM [38]	82.77
reweighted genetic algorithm (rGA) [51]	87.75
Proposed	88.46

3) ERICA DATASET

Due to its light nature, many classifiers generate efficient results over the ERICA dataset. The accuracy acquired by the proposed system via RNN and optimized features from ERICA dataset is 93.18%. The confusion matrix over ERICA dataset is shown in Table 7. It is shown that the results of interaction classes have achieved a mean accuracy rate of 93.18%. Despite the complications in activities, the results show that few activities are confused with other activities.

TABLE 7. Confusion matrix over erica dataset.

Classes	BC	LR	TE
BC	0.87	0.11	0.02
LR	0.09	0.89	0.02
TE	0.02	0.03	0.95
Mean Accuracy = 93.18%			

* BC = Bicep Curls; LR= Lateral Raise; TE = Triceps Extension

TABLE 8. Classification results over erica dataset.

Interaction Classes	Precision	Recall	F1 Score
Biceps curls	0.89	0.87	0.88
Lateral Raise	0.86	0.89	0.88
Triceps extension	0.96	0.95	0.95
Mean	0.90	0.90	0.90



FIGURE 16. Gaming Interface for showing predicted gaming activities of (a) Biceps curls and (b) Lateral raises using ERICA dataset.

TABLE 9. Erica dataset accuracy (%) comparison with state of the art methods.

Method	Recognition Accuracy (%)
Random Forest [41]	70.0
Decision trees [54]	81.7
Artificial neural network [55]	84.0
Proposed	93.18

The precision, recall, and F1-score for activities recognized are given in Table 8.

The outcomes from the gaming interface over ERICA dataset are shown in Fig. 16. The interface also illustrates the gaming object’s gesture position that will give the information regarding gaming label prediction.

Table 9 represents a comparison of RNN results over the ERICA dataset with other state-of-the-art methodologies. According to the table, Radhakrishnan et al. used the ERICA dataset for their experiment and achieved 70.0% accuracy using a random forest classifier [41]. An accuracy rate of 81.7% was attained for identifying gym workouts while monitoring the leg muscles using a pressure-sensing system [54].

The accuracy of the Kalman filter is 84.0% in a filter-based sensor fusion activity recognition system [55].

V. CONCLUSION

The proposed system effectively implemented a VR first-person game with an accurate deep learning-based gesture recognition system. Recently, it solved a major problem of obesity caused by a lack of physical activity particularly in the young generation. Three datasets were utilized for experimenting with the proposed approach, which are IMSB, WISDM, and ERICA datasets. First, the dataset used is pre-processed by applying a third-order median filter. Next, features were extracted by four well-known techniques called power spectral density, CNN-based features extraction, skewness, and kurtosis. Then, the datasets were further reduced through grey wolf optimization to get the optimized features. Further, the gestures were classified by applying the RNN classifier. After gesture prediction, the hardware is implemented using Arduino and motion sensors. Furthermore, the hardware and software components are created and combined by serial communication. Extensive experiments have been performed over the three datasets and demonstrated the effectiveness and efficiency of the system by achieving remarkable results and superior performance. It also outperformed the recognition accuracy of conventional state-of-the-art systems.

As for limitations, the sensor must be calibrated otherwise, it will generate wrong results. The sensor must be placed in the right place to get the desired gesture result. The dataset was generated by taking data on exercises from healthy and young persons, and if you try Bodyworn on disabled persons, you may get the wrong result. Activity labels might appear after a few seconds of delay if you performed multiple activities together in less than 2 seconds. A battery power of 5V is required for the sensor to work normally.

By including new features and playable games, we want to increase the effectiveness of the suggested system in the future. Additionally, we want to create a jacket with body-worn sensors. In the future, we also hope to increase the system's precision and provide consumers a better user interface so they may play and take pleasure in a virtual gaming experience.

REFERENCES

- [1] Z. Belyaeva, A. Petrosyan, and S. M. R. Shams, "Stakeholder data analysis in the video gaming industry: Implications for regional development," *EuroMed J. Bus.*, vol. 17, no. 3, pp. 333–349, Aug. 2022, doi: 10.1108/EMJB-10-2021-0150.
- [2] E. Zamani, "Effect of addiction to computer games on physical and mental health of female and male students of guidance school in city of Isfahan," *Addiction Health*, vol. 1, no. 2, pp. 98–104, 2009.
- [3] S. A. Bowman, "Television-viewing characteristics of adults: Correlations to eating practices and overweight and health status," *Preventing Chronic Disease*, vol. 3, pp. 1–11, Jan. 2006.
- [4] B. Poetker, "The very real history of virtual reality (+A look ahead)," *Vis. Coupled Airborne Syst.*, USA, Tech. Rep. re:2019, 2019.
- [5] O. Ouyed and M. A. Said, "Group-of-features relevance in multinomial kernel logistic regression and application to human interaction recognition," *Exp. Syst. Appl.*, vol. 148, Jun. 2020, Art. no. 113247.
- [6] G. Tao, B. Garrett, T. Taverner, E. Cordingley, and C. Sun, "Immersive virtual reality health games: A narrative review of game design," *J. NeuroEng. Rehabil.*, vol. 18, no. 1, pp. 1–21, Dec. 2021.
- [7] L. Jones, "VR and...better healthcare," *Eng. Technol.*, vol. 11, no. 3, pp. 36–39, Apr. 2016.
- [8] P. Zikas, S. Kateros, and P. Margarita, "Mixed reality serious games and gamification for smart education," in *Proc. Eur. Conf. Games Based Learn.*, 2016, pp. 805–812.
- [9] M. Yates, A. Kelemen, and C. S. Lanyi, "Virtual reality gaming in the rehabilitation of the upper extremities post-stroke," *Brain Injury*, vol. 30, no. 7, pp. 855–863, 2016.
- [10] I. Paraskevopoulos and E. Tsekleves, "Use of gaming sensors and customised exergames for parkinson's disease rehabilitation: A proposed virtual reality framework," in *Proc. 5th Int. Conf. Games Virtual Worlds Serious Appl. (VS-GAMES)*, Sep. 2013, pp. 1–5.
- [11] D. Fitzgerald, J. Foody, D. Kelly, T. Ward, C. Markham, J. McDonald, and B. Caulfield, "Development of a wearable motion capture suit and virtual reality biofeedback system for the instruction and analysis of sports rehabilitation exercises," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 4870–4874.
- [12] I. F. Mondragón Bernal, N. E. Lozano-Ramírez, J. M. Puerto Cortés, S. Valdivia, R. Muñoz, J. Aragón, R. García, and G. Hernández, "An immersive virtual reality training game for power substations evaluated in terms of usability and engagement," *Appl. Sci.*, vol. 12, no. 2, p. 711, Jan. 2022.
- [13] M. Ma, L. Jain, and P. Anderson, "Healthcare training enhancement through virtual reality and serious games," *Virtual, Augmented Reality Serious Games Healthcare*, vol. 68, pp. 9–27, Jan. 2014.
- [14] M. H. Abidi, A. Ahmad, S. Darmoul, and A. M. Al-Ahmari, "Haptics assisted virtual assembly," *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 100–105, 2015.
- [15] K. Gerling, P. Dickinson, K. Hicks, L. Mason, A. L. Simeone, and K. Spiel, "Virtual reality games for people using wheelchairs," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, Apr. 2020, pp. 1–11.
- [16] Y. Xu, M. Tong, W.-K. Ming, Y. Lin, W. Mai, W. Huang, and Z. Chen, "A depth camera-based, task-specific virtual reality rehabilitation game for patients with stroke: Pilot usability study," *JMIR Serious Games*, vol. 9, no. 1, Mar. 2021, Art. no. e20916.
- [17] S. Park, H. Aan, J. Jo, H. Kim, S. Han, J. Kim, P. Yoon, and K. Kim, "A-visor and A-camera: Arduino-based cardboard head-mounted controllers for VR games," in *Proc. IEEE Conf. Virtual Reality 3D User Int. Abstr. Workshops (VRW)*, Atlanta, GA, USA, Mar. 2021, pp. 434–435.
- [18] M. H. Abidi, A. Al-Ahmari, A. Ahmad, W. Ameen, and H. Alkhalefah, "Assessment of virtual reality-based manufacturing assembly training system," *Int. J. Adv. Manuf. Technol.*, vol. 105, no. 9, pp. 3743–3759, Dec. 2019.
- [19] M. H. Abidi, A. M. E. Tamimi, A. M. A. Ahmari, and E. S. A. Nasr, "Assessment and comparison of immersive virtual assembly training system," *Int. J. Rapid Manuf.*, vol. 3, no. 4, p. 266, 2013.
- [20] A. Al-Ahmari, M. Abidi, A. Ahmad, and S. Darmoul, "Development of a virtual manufacturing assembly simulation system," *Adv. Mech. Eng.*, vol. 8, no. 3, pp. 1–17, 2016.
- [21] S. Tang, S. Yuan, and Y. Zhu, "Data preprocessing techniques in convolutional neural network based on fault diagnosis towards rotating machinery," *IEEE Access*, vol. 8, pp. 149487–149496, 2020.
- [22] A. Famili, W.-M. Shen, R. Weber, and E. Simoudis, "Data preprocessing and intelligent data analysis," *Intell. Data Anal.*, vol. 1, no. 1, pp. 3–23, 1997.
- [23] S. Tripathy and T. Swarnkar, "Performance observation of mammograms using an improved dynamic window based adaptive median filter," *J. Discrete Math. Sci. Cryptogr.*, vol. 23, no. 1, pp. 167–175, Jan. 2020.
- [24] D. Varshni, K. Thakral, L. Agarwal, R. Nijhawani, and A. Mittal, "Pneumonia detection using CNN based feature extraction," in *Proc. IEEE Int. Conf. Electr., Comput. Commun. Technol. (ICECCT)*, Feb. 2019, pp. 1–7.
- [25] X. Shao, C.-S. Kim, and P. Sontakke, "Accurate deep model for electricity consumption forecasting using multi-channel and multi-scale feature fusion CNN-LSTM," *Energies*, vol. 13, no. 8, p. 1881, Apr. 2020.
- [26] H. Nassuna, O. S. Eyobu, J.-H. Kim, and D. Lee, "Feature selection based on variance distribution of power spectral density for driving behavior recognition," in *Proc. 14th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2019, pp. 335–338.

- [27] J. Chen, J. Wang, and J. M. W. Brownjohn, "Power spectral-density model for pedestrian walking load," *J. Structural Eng.*, vol. 145, no. 2, Feb. 2019, Art. no. 04018239.
- [28] A. Koduru, H. B. Valiveti, and A. K. Budati, "Feature extraction algorithms to improve the speech emotion recognition rate," *Int. J. Speech Technol.*, vol. 23, no. 1, pp. 45–55, Mar. 2020.
- [29] D. Trafimow, T. Wang, and C. Wang, "From a sampling precision perspective, skewness is a friend and not an enemy!" *Educ. Psychol. Meas.*, vol. 79, no. 1, pp. 129–150, Feb. 2019.
- [30] A. K. Alanazi, "Application of neural network and time-domain feature extraction techniques for determining, volumetric percentages and the type of two phase flow regimes independent of scale layer thickness," *Appl. Sci.*, vol. 12, no. 3, p. 1336, 2022.
- [31] M. Altaf, T. Akram, M. A. Khan, M. Iqbal, M. M. I. Ch, and C.-H. Hsu, "A new statistical features based approach for bearing fault diagnosis using vibration signals," *Sensors*, vol. 22, no. 5, p. 2012, Mar. 2022.
- [32] T. M. Ghazal and N. Taleb, "Feature optimization and identification of ovarian cancer using internet of medical things," *Exp. Syst.*, vol. 39, no. 9, p. e12987, Nov. 2022.
- [33] K. Asghari, M. Masdari, F. S. Gharehchopogh, and R. Saneifard, "A chaotic and hybrid gray wolf-whale algorithm for solving continuous optimization problems," *Prog. Artif. Intell.*, vol. 10, no. 3, pp. 349–374, Sep. 2021.
- [34] E. Krishna, V. Institute of Technology, T. Arunkumar, and V. Institute of Technology, "Hybrid particle swarm and gray wolf optimization algorithm for IoT intrusion detection system," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 4, pp. 66–76, Aug. 2021.
- [35] H. Singh Gill, O. I. Khalaf, Y. Alotaibi, S. Alghamdi, and F. Alassery, "Multi-model CNN-RNN-LSTM based fruit recognition and classification," *Intell. Autom. Soft Comput.*, vol. 33, no. 1, pp. 637–650, 2022.
- [36] J. A. Nasir, O. S. Khan, and I. Varlamis, "Fake news detection: A hybrid CNN-RNN based deep learning approach," *Int. J. Inf. Manag. Data Insights*, vol. 1, no. 1, Apr. 2021, Art. no. 100007.
- [37] J. Yan, L. Jin, Z. Yuan, and Z. Liu, "RNN for receding horizon control of redundant robot manipulators," *IEEE Trans. Ind. Electron.*, vol. 69, no. 2, pp. 1608–1619, Feb. 2022.
- [38] A. Jalal, M. A. K. Quaid, and A. S. Hasan, "Wearable sensor-based human behavior understanding and recognition in daily life for smart environments," in *Proc. Int. Conf. Frontiers Inf. Technol. (FIT)*, Dec. 2018, pp. 105–110, doi: 10.1109/FIT.2018.00026.
- [39] A. Jalal, M. A. K. Quaid, and M. A. Sidduqi, "A triaxial acceleration-based human motion detection for ambient smart home system," in *Proc. 16th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Jan. 2019, pp. 353–358.
- [40] S. K. Challa, A. Kumar, and V. B. Semwal, "A multibranch CNN-BILSTM model for human activity recognition using wearable sensor data," *Vis. Comput.*, vol. 38, pp. 4095–4109, Aug. 2021.
- [41] M. Radhakrishnan, D. Rathnayake, O. K. Han, I. Hwang, and A. Misra, "ERICA: Enabling real-time mistake detection & corrective feedback for free-weights exercises," in *Proc. 18th Conf. Embedded Networked Sensor Syst.* New York, NY, USA: Association for Computing Machinery, 2020, pp. 558–571.
- [42] F. N. Nezami, M. A. Wächter, N. Maleki, P. Spaniol, L. M. Kuhne, A. Haas, J. M. Pingel, L. Tiemann, F. Nienhaus, L. Keller, S. U. König, P. König, and G. Pipa, "Westdrive X LoopAR: An open-access virtual reality project in unity for evaluating user interaction methods during takeover requests," *Sensors*, vol. 21, no. 5, p. 1879, Mar. 2021.
- [43] Y. Peng, J. Chung, Q. Cao, and Y. Cai, "Design of a virtual home for special needs children to learn life skills," in *When VR Serious Games Meet Special Needs Education*. Berlin, Germany: Springer, 2021, pp. 31–61.
- [44] A. Luque, A. Carrasco, A. Martin, and A. De Las Heras, "The impact of class imbalance in classification performance metrics based on the binary confusion matrix," *Pattern Recognit.*, vol. 91, pp. 216–231, Oct. 2019.
- [45] N. Tatbul, T. J. Lee, S. Zdonik, M. Alam, and J. Gottschlich, "Precision and recall for time series," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 1–11.
- [46] S. B. Tahir, A. Jalal, and K. Kim, "Wearable inertial sensors for daily activity analysis based on Adam optimization and the maximum entropy Markov model," *Entropy*, vol. 22, no. 5, p. 579, 2020.
- [47] S. B. U. D. Tahir, A. Jalal, and M. Batool, "Wearable sensors for activity analysis using SMO-based random forest over smart home and sports datasets," in *Proc. 3rd Int. Conf. Advancements Comput. Sci. (ICACS)*, Feb. 2020, pp. 1–6.
- [48] S. B. U. D. Tahir, A. Jalal, and K. Kim, "Daily life log recognition based on automatic features for health care physical exercise via IMU sensors," in *Proc. Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Jan. 2021, pp. 494–499.
- [49] S. F. Idowu, A. Selamat, and R. Ibrahim, "Performance evaluation of classifiers on activity recognition for disasters mitigation using smartphone sensing," *Jurnal Teknologi*, vol. 77, no. 13, Nov. 2015.
- [50] K. Takahashi, K. Yamamoto, A. Kuchiba, and T. Koyama, "Confidence interval for micro-averaged F_1 and macro-averaged F_1 scores," *Int. J. Speech Technol.*, vol. 52, no. 5, pp. 4961–4972, Mar. 2022.
- [51] M. Batool, A. Jalal, and K. Kim, "Telemonitoring of daily activity using accelerometer and gyroscope in smart home environments," *J. Electr. Eng. Technol.*, vol. 15, no. 6, pp. 2801–2809, 2020.
- [52] W. Ahmad, B. M. Kazmi, and H. Ali, "Human activity recognition using multi-head CNN followed by LSTM," in *Proc. 15th Int. Conf. Emerg. Technol. (ICET)*, Dec. 2019, pp. 1–6.
- [53] K. H. Walse, R. V. Dharaskar, and V. M. Thakare, "Performance evaluation of classifiers on WISDM dataset for human activity recognition," in *Proc. 2nd Int. Conf. Inf. Commun. Technol. Competitive Strategies*, Mar. 2016, pp. 1–7.
- [54] B. Zhou, M. Sundholm, J. Cheng, H. Cruz, and P. Lukowicz, "Never skip leg day: A novel wearable approach to monitoring gym leg exercises," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Mar. 2016, pp. 1–9.
- [55] T. A. Shloul, M. Javeed, M. Gochoo, S. A. Alsubhany, Y. Y. Ghadi, A. Jalal, and J. Park, "Student's health exercise recognition tool for E-learning education," *Intell. Autom. Soft Comput.*, vol. 35, no. 1, pp. 149–161, 2023.



MIR MUSHHOOD AFSAR received the B.S. degree in computer science from Air University, Islamabad. He is currently a Research Assistant with the Intelligent Media Centre. His research interests include machine learning, deep learning, camera and sensor-based gesture recognition, and virtual reality.



SHIZZA SAQIB received the bachelor's degree in computer science from Air University, Islamabad. She is currently a Research Assistant with the Intelligent Media Center. Her research interests include machine learning, deep learning, image processing, and virtual reality.

MOHAMMAD ALADFAJ is currently with the Department of Natural Engineering, College of Science, King Saud University, Saudi Arabia.



MOHAMMAD HAMAD ALATIYYAH is an Assistant Professor of computer science with the Computer Science Department, Prince Sattam Bin Abdulaziz University, Saudi Arabia. His research interests include the recommender systems and computer vision, such as group recommender systems, travel recommender systems, and drone vision.



KHALED ALNOWAISER received the Ph.D. degree in computer science from Glasgow University, Scotland. He is an Assistant Professor with the Computer Engineering Department, Prince Sattam Bin Abdulaziz University, Saudi Arabia. His research interests include computer vision, optimization techniques, and performance enhancement.



AHMAD JALAL received the Ph.D. degree from the Department of Biomedical Engineering, Kyung Hee University, South Korea. He is currently an Associate Professor with the Department of Computer Science and Engineering, Air University, Pakistan. Currently, he is a Postdoctoral Research Fellowship with POSTECH. His research interests include multimedia contents and artificial intelligence.

HANAN ALJUAID received the B.S. degree from KAU University and the M.S. and Ph.D. degrees in computer science from UTM University, in 2014. She is currently with the Computer Sciences Department, College of Computer and Information Sciences, Princess Nourah Bint Abdul Rahman University (PNU), Saudi Arabia. Much of her work has been on improving the understanding, design, and performance of pattern recognition, mainly through the application of data mining and machine learning. She has given numerous invited talks and tutorials. She has published numerous articles in pattern recognition, the IoT, and data science. Her research interests include computer vision and NLP.



JEONGMIN PARK received the Ph.D. degree from the College of Information and Communication Engineering, Sungkyunkwan University, in 2009. He was a Senior Researcher at the Electronics and Telecommunications Research Institute (ETRI) and a Research Professor at Sungkyunkwan University, South Korea. In 2014, he joined with the Tech University of Korea, South Korea, where he is an Associate Professor with the Department of Computer Engineering.

His research interests include high-reliable autonomic computing mechanism and human-oriented interaction systems.

...