SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2024, 21(4), 468.
https://doi.org/10.62617/mcb468

Article

Application of human-computer interaction technology integrating biomimetic vision system in animation design with a biomechanical perspective

Jing Han

School of Art and Design, Yantai Institute of Science and Technology, Yantai 265600, China; 13081635526@163.com

CITATION

Han J. Application of human-computer interaction technology integrating biomimetic vision system in animation design with a biomechanical perspective. Molecular & Cellular Biomechanics. 2024; 21(4): 468. https://doi.org/10.62617/mcb468

ARTICLE INFO

Received: 8 October 2024
Accepted: 23 October 2024
Available online: 25 December 2024

Abstract: The combination of Human-Computer Interaction (HCI) technology with biomimetic vision systems has transformational potential in animation design, particularly by incorporating biomechanical principles to create immersive and interactive experiences. Traditional animation approaches frequently lack sensitivity to real-time human motions, which can restrict engagement and realism. This study addresses this constraint by creating a framework that uses Virtual Reality (VR) and Augmented Reality (AR) to generate dynamic settings that include a variety of human activities, informed by biomechanical analysis. A biomimetic vision system is used to record these motions with wearable sensors, allowing for precise monitoring of user activity while considering biomechanical factors such as joint angles, force distribution, and movement patterns. The recorded data is preprocessed using Z-score normalization methods and extracted using Principal Component Analysis (PCA). This study proposed an Egyptian Vulture optimized Adjustable Long Short-Term Memory Network (EVO-ALSTM) technique for motion classification, specifically tailored to recognize biomechanical characteristics of human movements. Results demonstrate a significant improvement in precision (93%), F1-score (91%), accuracy (95%), and recall (90%) for the motion recognition system, highlighting the effectiveness of biomechanical insights in enhancing animation design. The findings indicate that integrating real-time biomechanical data into the animation process leads to more engaging and realistic user experiences. This study not only advances the subject of HCI but also provides the framework for future investigations into sophisticated animation technologies that use biomimetic and biomechanical systems.

Keywords: animation; motion recognition; biomimetic vision system; human activities; biomechanics; Human-Computer Interaction (HCI)

1. Introduction

Human-computer interaction (HCI) technology plays an important function in animation design since it allows natural interaction between people and computers. It permits animators to create and manipulate complicated visual elements through user-friendly interfaces and interactive systems [1]. When applied in the layout of animation, the approach is additionally efficient, inventive, and accessible to designers, enabling them to convey their ideas into consequence more accurately. It reduces the need for lots of manual intervention and narrows down the creative process. It presents new opportunities for launching innovative and complex animation projects [2].

A biomimetic vision system is an artificial visible processing generation that mimics biological vision, mainly human or animal vision structures. These systems are designed to capture, process, and interpret visual information further to how residing organisms work [3]. In animation design, incorporating a biomimetic vision system allows for more realistic animations because the system can observe and interpret natural movements and visual cues, ensuring accurate person and surroundings rendering. It also adapts to varying lighting and environmental situations to reflect how human vision works in diverse settings [4]. The era can help designers obtain better ranges of elements and visual accuracy in their work.

The integration of the biomimetic vision framework into animation enhances the realism of the visual factors. By simulating the way human beings understand light intensity, color, and movement, this era permits animators to create reasonable scenes and characters [5]. It results in extra immersive and convincing animators, where the movers interact with the surrounding experiences naturally, enriching the viewer’s experience. Realism can help animations evoke stronger emotional responses from the target audience [6]. Moreover, it lets in for extra attractive storytelling through making the visuals seem towards real life.

HCI technology, when mixed with biomimetic vision, can massively enhance the interactive design of animations. Animators and architects can engage with the system more intuitively through the use of gestures, eye tracking, or voice instructions to control the animation technique [7]. The hand-free approach allows for smoother workflows and extra engaging user experiences, making it less complicated to create dynamic animations that reply to real-time inputs. Additionally, it fosters an extra immersive improvement where designers can, at once, affect results with minimum. It ensures greater innovative flexibility and reduces repetitive guide responsibilities.

In 3D animation and visible outcomes, the biomimetic vision system can simulate complicated visual phenomena inclusive of reflections, shadows, and textures with higher accuracy [8]. These structures analyze the scene as a human eye track, supporting animators to generate more specified and visually attractive environments. The technology can also enhance motion capture using imitation of how human eyes follow movement, resulting in additional natural and fluid character movements. Thus, biomimetic vision makes it easy to grab and record subtle information, including changes in texture or light, to produce near-perfect digital worlds [9]. It contributes a dimension of professional excellence to tasks, generally making them even more pleasant across the board.

For animation designers and manufacturers, the combination of HCI technology and biomimetic vision systems offers significant benefits. The time and effort exerted in the production of quality animation are also cut down, in that the device optimizes the critical animated values mainly based on real-world recorders [10]. Furthermore, this era also allows designers to explicitly express more visible styles and concepts while enabling realistic and interactive designs. It additionally makes collaboration simpler, as a couple of designers can work seamlessly with the device. It leads to quicker manufacturing cycles and complements the overall quality [11].

To enhance animation design, the integration of the HCI era and biomimetic and prescient systems can be applied to human activity. By leveraging actual time motion tracking and visual recognition, those technologies can analyze and interpret human moves with excessive precision. Biomimetic vision systems mimic the way human movements, bearing in mind more accurate classification of various activities together [12]. The approach is particularly useful in fields like animation, digital reality, and video game development, wherein sensible character moves are crucial. **Figure 1** represents the general framework of virtual interaction and human activities.

Figure 1. The general framework of virtual interaction and human activities.

HCI generation additionally complements the process by supplying intuitive interplay methods, including gesture control and body movement monitoring, permitting designers to seamlessly manipulate animations that accurately replicate human activities. Additionally, device mastering algorithms can be integrated to continuously improve the accuracy of interest type by getting to know big information about human motion [13]. The effects include smoother and more fluid animations that reflect human movements; this integration streamlines workflows complements realism, and opens new opportunities for interactive actual-time animation improvement.

The paper aims to improve animation design by proposing an Egyptian vulture-optimized adjustable long short-term memory network (EVO-ALSTM) for human activity classification.

Key contributions

- The data was gathered from 30 participants using wearable sensors, which included accelerometers and gyroscopes.
- Z-score normalization was used for preprocessing and PCA was used to extract the complex features from the preprocessed data.
- EVO-ALSTM is proposed to classify human activity.

The remaining parts of this paper: Part 2 represents the related work, a methodology that includes dataset, preprocessing, and feature extraction, and the proposed method was described in Part 3. Part 4 presents the result and discussion. Part 5 covered the paper's conclusion.

2. Related work

To enhance the natural relationship between humans and machines, [14] created a detailed blueprint for a robot with a humanoid head that possessed human-like feelings and activities. It also evaluated how well human behavior and emotional expression were simulated by the motor and sensory control system. The field of biomimetic humanoid robots benefited from these efforts. Experimental data from the survey indicated that participants expressed emotions, and participants replicated actions.

A didactic, graphical tool that presented the most recent applications of biomimicry in medicine [15] suggested a webpage was created to show 2D artwork and visuals (animations). Animation was a tried-and-true method of educating the public about health-related topics. All respondents agreed that biomimicry could provide useful responses for medical design. Investigation showed that for public outreach, visual motions could successfully communicate complex ideas.

A presentation on 2D animation using artificial intelligence and biomechanics modeling (2D-AI-BM) was presented in [16]. Deep neural networks (DNN) for movement predictions and development based on biopsychological principles were employed in the process to better resemble real human motions. Research contrasting that approach with conventional animation approaches has demonstrated that it reduced the production period while indeed generating realistic movements for 2D characters. It presented numerical results proving that the implementation of the 2D-AI-BM model enhances an accuracy rate.

A flexible simulated tactile approach that made use of the stick-slip sensing model in [17] offered a general approach to identify a failure and quantify the surface properties of an object through slippage. The system comprised a read-out system in the form of tips of hands, a display unit, and an artificial intelligence component. Based on the stick-slip sensing approach, the system had a high identification rate for slippage monitoring. The multipurpose system was also demonstrated for interactive gaming, robotic hand deception, and identifying objects, allowing for extensive and prospective interactions between humans and machines.

A new architecture that employed motion data recorded [18] by human webcams. Because of that technique, which used a lot of movement information in the real-time recreation of such animation as animals moving, designers could be in a position to design those characters with more accurate movements, which could depict real-world settings. Moreover, users' actions were tied to virtual reality, making the whole action more realistic and exciting.

A model of random forests [19] was to process and create animation data, and from the collected animation data, the knowledge that could guide the development of animation was extracted. Based on the design goal and execution approach of the animated information processing and development platform, the features and categories of the random forest model were separated. The findings from the experiment showed that the platform for developing and processing 3D animation data was both practical and efficient.

Created several techniques to expedite and reduce the expense of that procedure [20] developing a mobile robot that could follow the actor and record the scene while keeping them where he needs to be in the frame. By feeding the recorded video through a range of deep learning algorithms, the team could then determine the actor's 3D position. It could be used to animate the required 3D model; therefore, there is no need to use several cameras and a mobcap suit to capture the movements of the actor.

A customized DNN to constantly and accurately detect external haptic stimuli [21] suggested a new method of data enhancement process was pioneered by identifying the hexagonal structure of the sensor, which has six-fold rotational symmetry and possesses mirror images. The generated pseudo data could enhance the generalization performance of the DNN model by adding the obtained training data. The sensor proved its effectiveness and the feasibility of the proposed data augmentation technique and provided a good generalization of five touch modes and potential for further development to improve human-robot interaction.

Four methods for incorporating deep learning models and Kinect camera-based animated manufacturing systems with natural human movement [22] were examined. The selection of each approach was contingent upon environmental circumstances and accuracy. The initial solution made use of a Kinect camera. A camera and a calibration algorithm were employed in the second technique. The third option made use of a deep learning framework. A deep learning model was employed when combined with Kinect in the fourth method. Comparing the recommended method's experiments to previous approaches, it was found that the fourth method, which combines a Kinect and a deep learning model, produced the greatest results.

A complex HCI program to point at the problem of communication impairment between groups of individuals with hearing disabilities and without such problems was designed [23]. The advancement of artificial intelligence has made it quite possible by the hard work of listening and people without disabilities to communicate as per their desire. Using near-wearable technology and utilizing backpropagation (BP) neural network models to classify gestures, the proposed system was able to successfully close the communication gap between the impaired and the non-impaired individuals.

A revolutionary biomimetic bidirectional cooperation perceiving system (BBCPS) [24] suggested the gaze function in human eyes served as inspiration for their creation of a simple yet versatile BBCPS. The results of the simulation demonstrate that it reduced operational energy consumption and enhanced braking effectiveness. Furthermore, a system for initial position calibration was established to ensure that the BBCPD state matches the control strategy that followed. The methodology allowed for the certification and modification of the camera pose and servo motors' zero-position. The gaze error was fewer than three pixels across in real testing, confirming the control performance of the BBCPS.

Neurological control systems [25] enhance significantly the realism of the simulation of human movement. They have underscored some of the challenges in relocating head pose and facial emotions from the pictures and clinical movies into muscle-actuated modeling of facial and head and neck. A complicated biomechanical system was also involved in generating locomotion-based animations with biomechanical plausibility. It extended the use of imitation, physics-based humanoid simulators, and modeling in graphic design and vision by showcasing the adaptability of the face and body controlled by muscles.

A unique learning-based method for biomechanically modeling the face-head-neck complicated by importing facial emotions and head motions from images and videos was described in [26] and suggests training a DNN to take in the face actions coding system (FACS) with action units (AUs) and produce appropriate facial muscle and mouth motion signals for the biomechanics model by using the FACS as a substitute for representing emotion distance. Experiments including the projection of different facial emotions and head poses from films onto the face-head-neck model demonstrated the model's efficacy.

A simple yet efficient method that used deep learning (DL) [27] to create a basic 3D animation of numerous people moving in 2D. Despite recent considerable advances in 3D human posture calculation, multi-person determination of pose was rather a challenging topic, and most previous works were yet limited to single-person estimation of poses. Using the publicly available dataset, the proposed system performed comparably to prior state-of-the-art 3D multi-person pose approximation approaches and surpassed previous competitive human pose tracking devices by a significant margin.

The efficacy and precision of human annotators, whether employing video, data, or both for annotating events across four human activity recognition (HAR) tasks [28] observed that annotators were more accurate in classifying kinds of events when employing video alone on all four tasks and more effective while using data alone on three of the four assignments. The annotations of event boundaries based on data alone were more accurate. The experimental findings discovered that the data and video collected for HAR task annotations had multiple functions and that the functions might vary across the HAR tasks.

A novel system for deep learning based on signals from movement to identify human activities and address these limitations and difficulties using deep learning techniques [29] Utilized convolutional neural networks (CNNs) and laboratory metrics, the methodology was effectively studied and obtained better accuracy in comparison to machine learning techniques. The research's innovative approach was to improve classification accuracy while executing tasks more quickly and with a lower mistake rate. It also introduced a new technique that uses CNN with Adam's optimization technique to detect human involvement in the dataset.

Problem statement

The current animation design using HCI with a biomimetic vision system brings massive challenges, such as capturing and responding to actual time human motion. Traditional animation techniques regularly fail to offer the level of sensitivity and interactivity required for attractive and practical user experiences. Existing frameworks often struggle to appropriately interpret complicated moves, leading to a disconnection between user actions and animated responses. There is also a lack of standardized protocols for integrating HCI and biomimetic systems, complicating the development of cohesive frameworks that make certain compatibility and effectiveness. The high cost related to the advanced technology implementation can restrict accessibility for big adoption and innovation within the field. These challenges collectively hinder the potential of animation design to create trust immersive and interactive investigations that resonate with users. The proposed method deals with the limitations of cutting-edge animation techniques with the aid of growing a framework that utilizes VR and AR to create dynamic environments that contain numerous human activities.

3. Methodology

The data consists of analysis of the relevant data where Z-score normalization is needed to standardize the distribution of the data. To reduce dimensionality, and retain the most important variance, Principal Component Analysis (PCA) is used. The present method EVO-ALSTM, combines the benefits of Egyptian Vulture optimized adjustable Long Short-Term Memory networks prediction power as well as optimizing the learning process adaptively. The purpose of the coupling is to focus on improving the performance of the model concerning complex data sets these overall procedures are shown in **Figure 2**.

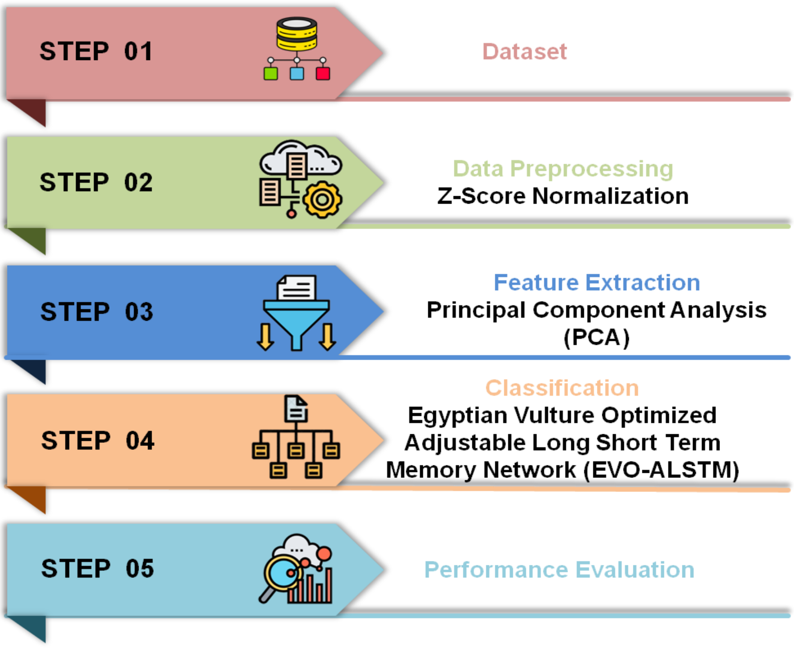

Figure 2. Methodology's overall procedure.

3.1. Dataset

The dataset was gathered from Kaggle (https://www.kaggle.com/datasets/ziya07/human-motion-dataset-for-animation-design/data). The dataset comprises human activity data acquired using attached motion sensors; accelerometers and gyroscopes, which were embedded into a biomimetic vision sensor system on thirty subjects (15-males, 15-females). This system intraocular visualizes dynamic activities as a human eye would increase the precision in such activities. Each of the subjects performed a set of defined motions, namely walking, jumping, waving arms, and performing sports-mimicking actions. The activities were recorded in an enclosed area as most of the activities were captured and implemented in a 3D world so a 3D vision system was used along with augmented reality (AR). Before data gathering, the sensors and the imaging system were set up to enhance their functional efficiency and accuracy for the measurements of motion and body dynamics. Each session per participant lasted 30 min with the directive to act as freely as possible while the recordings were taken. The data obtained, which was further enhanced by the imaging system, was collected and classified in an orderly fashion along the lines of subject identification number and type of activity to facilitate in-depth tracking of human movements for other purposes.

3.2. Z score normalization

The process of normalizing data involves scaling or mapping abnormal data to standard data. The Z-score approach, a numerical data type, is used in the model and its normalized values in the dataset. A common statistical method for standardizing and normalizing numerical features in a dataset is the Z-score method. It calculates the Z-score for every fact point by subtracting the implied and dividing, utilizing the standard deviation of the dataset. The normalization statistics can have a mean of 0 and a popular deviation of 1, improving the model's overall performance in the movement class. It facilitates minimizing the influence of outliers and ensures that different capabilities contribute similarly to the type system. The formula for Z-score normalization is represented in Equation (1).

$$z = \frac{(x_i - \mu)}{\sigma} \tag{1}$$

3.3. Extraction of feature

PCA utilized to remove features to maximize records variability and then convert it directly into a space with a low number of dimensions. It is a powerful method used for feature extraction in movement category responsibilities related to human activity datasets. By reducing the dimensionality of the statistics, PCA identifies the most sizeable capabilities that account for the variance in human actions, facilitating improved model performance. The method eliminates the noise and redundant data, making it easier to classify critical patterns and movements. This algorithm can function extra efficaciously and correctly, enhancing the recognition of numerous human activities and it contributes to extracting effective and reliable movement classification structures. An ortho basis set that is identical is the resultant vector set. Since the fundamental elements are the vectors that form part of the balanced interaction matrix, each of them is orthogonal. In mathematics, if l samples are taken from a dataset and the class label is not considered, then every measurement is n-dimensional. Assume $w_1, w_2, \dots\dots, w_l \in \Re^m$ that the following steps for PCA calculation.

Determine the mean vector μ in n observations by Equation (2):

$$\mu = \frac{1}{l}\sum_{j=1}^{l} w_j \tag{2}$$

Obtain the expected coefficients matrix T for the acquired data by Equation (3):

$$T = \frac{1}{l}\sum_{j=1}^{l}(w_j - \mu)(w_j - \mu)^s \tag{3}$$

Compute the appropriate equations and$T$ as an eigenvalue where$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l \geq 0$.

Using the l starting parameters, produce the l necessary components by Equation (4):

$$\begin{gathered} z_1 = b_{11}w_1 + b_{12}w_2 + \cdots + b_{1l}w_l \\ z_2 = b_{21}w_1 + b_{22}w_2 + \cdots + b_{2l}w_l \\ \cdots \\ z_l = b_{l1}w_1 + b_{l2}w_2 + \cdots + b_{ll}w_l \end{gathered} \tag{4}$$

Attempts to explain the first variation in the data set as much as possible, and z_2 attempts to explain the remaining variance, etc. A few instances of bigger eigenvalues usually control the rest in the most valuable data sets that are represented in Equation (5).

$$\gamma_l = \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_n}{\lambda_1 + \lambda_2 + \cdots + \lambda_n + \cdots + \lambda_l} \tag{5}$$

whereλ_l represents the percentage preserved in the data forms. Since the produced principal components via PCA extracted features constitute the variability in the data, they ought to be retained.

3.4. Egyptian vulture optimized adjustable long short-term memory network (EVO-ALSTM)

Hybrid approach combining the evolutional optimization approach of the Egyptian Vulture optimized (EVO) with adjustable LSTM (ALSTM) algorithms. The optimization mimics the foraging characteristic of Egyptian vultures to fine-tune hyperparameters, including learning rate, amount of hidden units, and dropout rate within the ALSTM, which is designed to regulate its memory and neglect gates dynamically. This hybridization approach improves the ALSTM's capacity to evolve to varying temporal dependencies in movement classification responsibilities. The integration ensures faster convergence and better motion sequence reputation performance by leveraging each efficient optimization and flexible memory modification in the LSTM. It achieves improved accuracy and efficiency in movement classification compared to standard LSTM networks.

3.4.1. ALSTM

ALSTM is a variant of the LSTM network designed for motion classification tasks, wherein the model dynamically adjusts its internal memory gates based on various temporal dependencies in movement sequences. This adaptability allows the ALSTM to selectively consider or forget information at special time steps, enhancing in classifying complex and time-sensitive movement patterns. There are four layers in the ALSTM prediction model: input, hidden, output, and EVO. The prediction model is optimized by the usage of EVO. Based on the RNN design approach and consider the properties. $C = \{W_1, W_2, \ldots., W_r\}$ denotes the entire collection, it is separated into two subsets: a training data set called $W_{train} = \{W_1, W_2, \ldots.., W_o\}$ and a test data set called$W_{test} = \{W_{o+1}, W_{o+2}, \ldots \ldots., W_r\}$. Each of those samplings is represented by the symbol $W'_{train} = \{W'_1, W'_2, \ldots., W'_o\}$, which includes n attributes $W_j = \{w_{j,1}, w_{j,1}, \ldots.., w_{j,m}\} o < ro, r \in M$. Utilizing the Z-score normalization technique (mean 0, variance 1), standardize the training set as expressed in Equation (6):

$$W_s^{'} = \left(W_i - \frac{\sum_{s=1}^{o} W_s}{o}\right) / \sqrt{\sum_{s=1}^{o} (W_i - \sum_{s=1}^{o} W_s\,/o)^2/o} \tag{6}$$

In the above instance Ws' is the standardized amount for the property at times$W_s^{'} \in [-1,1]$ and $s \in [1, o]$, data are split to adjust to the hidden layer's input properties. Equations (7)–(9) denote the data set separation appears as follows when the separation length is set to$K$:

$$C = \{C_1, C_2, \dots, C_{o/K-1}\} \tag{7}$$

$$C_i = \{W_i^{'}, W_{i+1}^{'}, \dots \dots, W_{i+K-1}^{'}\} \quad 1 \le i \le K \le o \tag{8}$$

$$W_i^{'} = [w_{i,1}^{'}, \dots., w_{i,i}^{'}, \dots.., w_{i,m}^{'}] j \in [1, m] \tag{9}$$

The amount of $w_{i,j}^{'}$ at a given time is entered as an algorithm in the hidden layer, which is composed of K identical LSTM units interconnected at the times before and after. The hidden layer's output in Equation (10) is as follows:

$$(g_s, d_s) = LSTM([W_s^{'}, g_{s-1}], d_{s-1}, X \tag{10}$$

where the current state and output of the preceding LSTM unit have the following values: d_{s-1} andg_{s-1}. $LSTM(*)$ is the LSTM unit's forward computation. Equation (11) represents the final output of the LSTM network after it has passed through a fully connected layer, which is the ALSTM's output.

$$Z_s = \sigma(X^{zd} g_s + a^{zd}) \tag{11}$$

The training method uses the mean square error, which is defined in Equation (12), as the loss function.

$$loss = \sum_{j=1}^{K \times N} (o_j - z_{real})^2 / (K \times N) \tag{12}$$

where K and N stand for the isomorphic LSTM unit and the total amount of input samples, correspondingly. The numerical value of the LSTM model's training output is indicated byo_j. The sample's real value is denoted byz_{real}. The network algorithm's layer counts, training intervals, and hidden layer neuron count are all simultaneously adjusted to the optimization.

3.4.2. EVO

EVO was a meta-heuristic method initially developed to solve complicated arrangement problems. It is aroused by the Egyptian vulture's behavior to obtain a solution. This avian creature's cunning behavior is converted into an algorithm that can handle challenging optimization issues. The system has been modified to incorporate the detailed EVOA processes. It optimizes the selection of key features, improving the accuracy and efficiency of ALSTM in recognizing complicated human motions. By integrating this bio-inspired optimization technique, the ALSTM can better capture temporal dependencies and diffused variations in movement information. This method leads to greater specific and dependable motion types, especially in dynamic environments. **Figure 3** represents the process of EVO, algorithm 1 represents the EVO-ALSTM algorithm, and the following are the steps in the procedure:

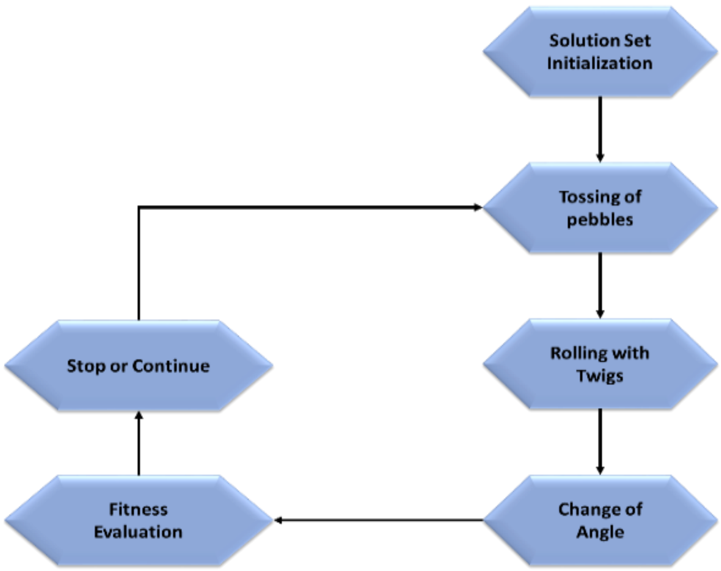

Figure 3. Process of EVO.

Algorithm 1 EVO-ALSTM algorithm

```
Step 1: def initialize_population():
   population = []
   for i in range(pop_size):
      individual = random_hyperparameters()
population.append(individual)
   return population
Step 2: def fitness (individual, train_data, val_data):
lstm_model = build_lstm(individual)
lstm_model.train(train_data)
   performance = evaluate_model (lstm_model, val_data)
   return performance
Step 3: def egyptian_vulture_optimization(population, train_data, val_data):
   for generation in range (max_generations):
      for vulture in population:
fitness_score = fitness (vulture, train_data, val_data)
update_best_solution(fitness_score)
      population = evolve_population(population)
      return best_solution
Step 4: def adjustable_lstm (input_data, hyperparameters):
lstm_layer = LSTM (hyperparameters['hidden_units'], return_sequences=True)
dropout_layer = Dropout(hyperparameters['dropout'])
adjusted_output = adjust_memory_for_sequence(lstm_layer, input_data)
   return dropout_layer(adjusted_output)
Step 5: def evo_alstm_motion_classification (train_data, val_data):
   population = initialize_population ()
best_hyperparameters = egyptian_vulture_optimization (population, train_data, val_data)
lstm_model = build_lstm (best_hyperparameters)
Step 6: hyperparameters
lstm_model.train (train_data)
motion_classification_results = lstm_model.classify (test_data)
return motion_classification_results
```

Step 1: The initialization of the solution set of strings contains changeable representations of the parameters. One possible solution state is represented by a string with a set of parameters.

Step 2: Conditions are verified, limitations are superimposed, and representative variables are refined.

Step 3: Stones are thrown at predetermined or random locations.

Step 4: Either a portion of the string or the complete one is picked for the Rolling of the Twigs performance.

Step 5: The strategy of changing the angle is used to reverse a specific portion of the solution.

Step 6: Fitness is assessed.

Step 7: Checking the stopping criterion is necessary.

4. Result and discussion

In this section, the confusion matrix for classification is calculated, the performance of the proposed method based on the factors including walking, jumping, arm waving, and sports action is evaluated, and the effectiveness of the proposed method EVO-ALSTM with the conventional technique ALSTM is compared based on the metrics (accuracy, precision, recall, and F1-score).

4.1. Experimental setup

An Intel i7-7500U CPU running at 2.70GHz with 8 GB of RAM and Mat lab R2014a was used to simulate the proposed method.

4.2. Confusion matrix

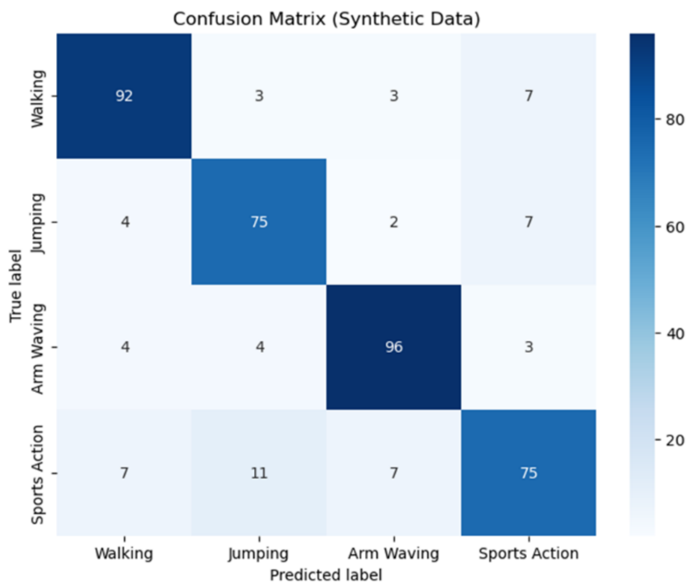

Figure 4. Output of confusion matrix with 30 participants.

To evaluate the performance of a classification method, it compares the prediction classification with the actual results by organizing outcomes into four classes, including true positive (TP), false positive (FP), true negative (TN), and false negative (FN). **Figure 4** represents the output of the confusion matrix with 30 participants. In **Figure 4**, the diagonal values represent accurate detections, with excessive accuracy for walking (92), jumping (75), arm waving (96), and sports action (75). Misclassifications are seen, which include walking expected as sports action (7 instances) and jumping misclassified as sports action (7 instances). Darker cells imply better accuracy, at the same time as lighter cells reflect fewer correct classifications. Overall, the model successfully classifies all instances without errors.

4.3. Motion detection accuracy

The motion detection accuracy that estimates how well the system features the ability to cope with the concerns of identification of movement within an environment or in a video frame movement tracking. It is expressed as the number of correctly identified motion events over the number of existed motion events. The accuracy level of various human activities classified using EVO-ALSTM was evaluated, which is collected using sensors. **Table 1** represents the motion detection accuracy of human activity.

Table 1. Motion detection accuracy of human activity.

Human activities	Motion detection accuracy (%)
Walking	94.7%
Jumping	92.3%
Arm waving	95%
Sports action	93.4%

Figure 5 shows the wearable sensors and a biomimetic vision system used to record various human activities indicating reliable movement detection accuracy. Among the different activities, 94.7 % accuracy was recorded in walking while arm waving was slightly higher at 95%. Complementing this, jumping and sports-related activities recorded an accuracy of 92.3 % and 93.4 % respectively. This underlines the efficient performance of the system in the execution of any other dynamic activities. According to the findings, the proposed method achieved a high accuracy level in classifying arm-waving movement.

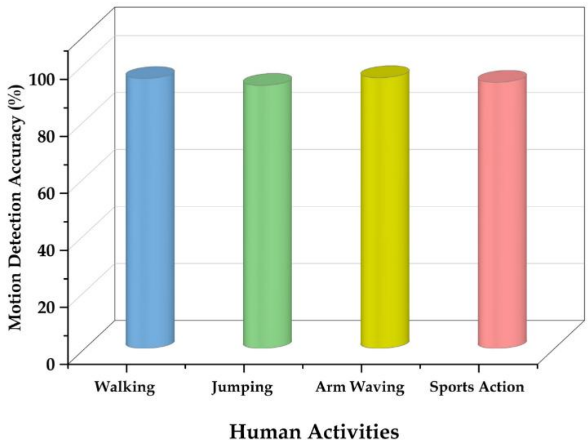

Figure 5. Motion detection accuracy of human activity.

4.4. Comparison

4.4.1. Accuracy

The model's accuracy is calculated by dividing the total number of true positive (TP) and true negative (TN) predictions by the total number of forecasts. This gives an overall performance level of the model with consideration of all the classes. It consists of the proportion of accurately identified cases to all occurrences of human activity. **Table 2** and **Figure 6** display the performance of accuracy.

Table 2. Values of four metrics.

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
ALSTM	88%	83%	86%	82%
EVO-ALSTM	95%	93%	90%	91%

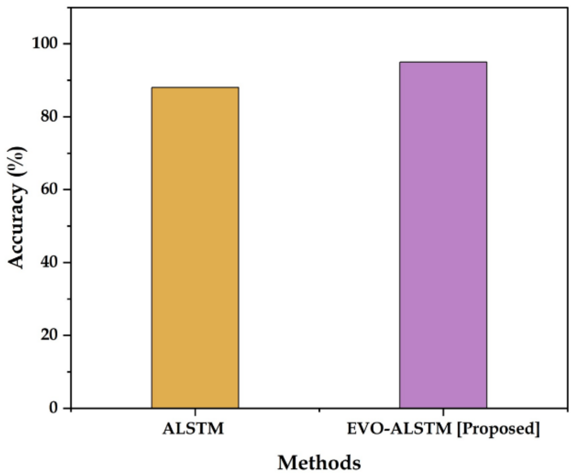

Figure 6. Performance of accuracy.

Figure 6 compares the accuracy of two methods: ALSTM and EVO-ALSTM (Proposed). EVO-ALSTM demonstrates improved accuracy achieved 95%, while ALSTM shows slightly lower performance around 88%. This indicates that the proposed EVO-ALSTM method outperforms the standard ALSTM in terms of accuracy.

4.4.2. Precision

Precision focuses on the number of true positive predictions when compared to the total amount of accurate forecasts that the model produced. This indicates that based on the model's overall number of accurate predictions during the period under consideration, how many of them are real predicted positive cases and how better the model in making non-positive predictions. It measures how the model accurately classifies a motion. It indicates how many positive instances are positive. **Figure 7** and **Table 2** show the precision evaluation.

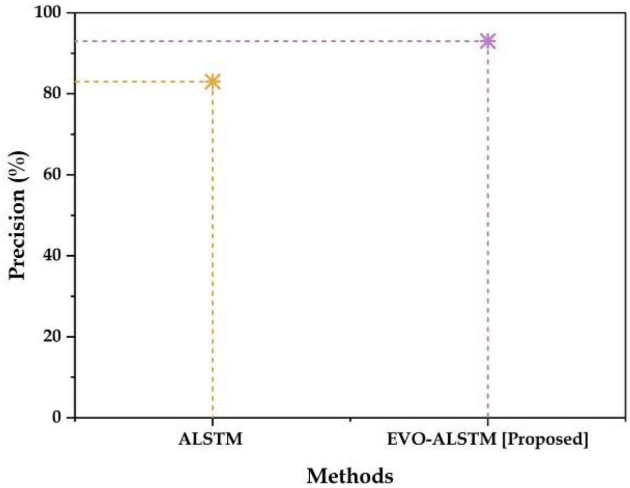

Figure 7. Precision performances.

Figure 7 depicts an assessment of the precision of ALSTM and the EVO-ALSTM model. The precision of the EVO-ALSTM technique is remarkably better 93% which confirms its capability to locate pertinent instances with minimal errors. The ALSTM achieves an accuracy level of about 83%, which infers a relatively higher false positive ratio.

4.4.3. Recall

Sensitivity also referred to as recall is the measure of the ratio of the number of true positives as compared to the number of positive cases in the population. This shows the capability of the model to recognize all the relevant elements of a certain class. Recall measures the classifier's capacity to find every relevant occurrence of a given motion by calculating the proportion of the true positive to the actual positive. **Table 2** and the recall performance is shown in **Figure 8**.

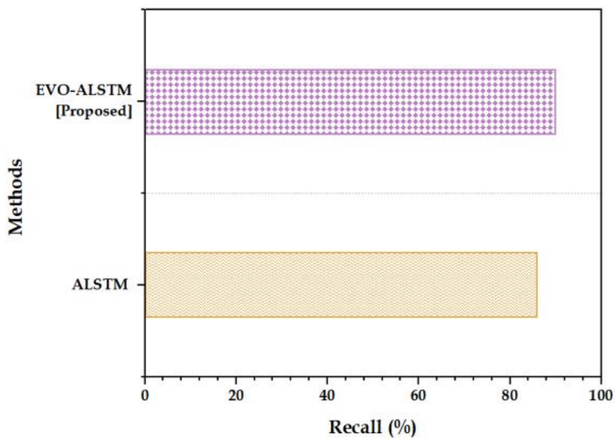

Figure 8. Performance of recall.

In **Figure 8**, the recall percentage in the two methods ALSTM and EVO-ALSTM is presented with the proposed EVO-ALSTM with very high recall 90% proving its effectiveness in instances recovery task. The recall performance of ALSTM is less than that, at about 86%, suggesting that this method loses more relevant information when compared to EVO-ALSTM.

4.4.4. F1-score

The F1 score is a metric that takes into account both precision and recall by averaging their numerical scores. It is especially positive in cases of irregular data distribution where both false positive and false negative are important. It is the balance between precision and recall, ensuring that the model performs well in both classifying and accurately assigning motion. **Table 2** and the analysis of the F1-score are shown in **Figure 9**.

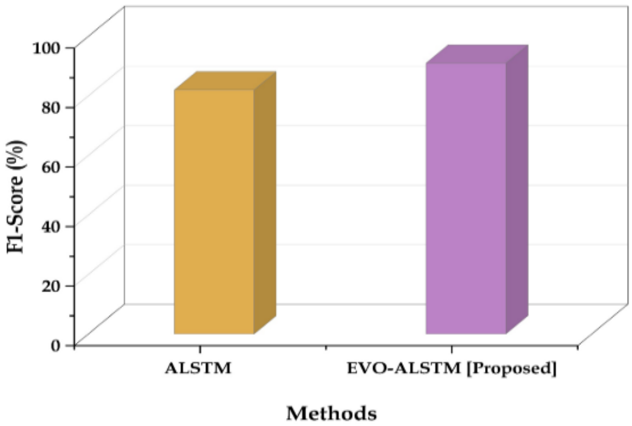

Figure 9. Evaluation of F1-score.

Figure 9 shows the F1 score for both the ALSTM and EVO-ALSTM approaches. The efficiency of the suggested EVO-ALSTM in various recovery tasks is demonstrated by its extremely high F1 score of 91%. However, ALSTM's F1 score performance is lower than that, at around 82%, indicating that this approach loses more pertinent data than EVO-ALSTM.

4.5. Discussion

The standard method EVO, while effective for hyperparameters optimization, could struggle with local minima and convergence speed, particularly in high dimensional search spaces. Additionally, it is computationally expensive, as it requires multiple evaluations of the fitness function, a process that can be time-consuming in large data sets. Moreover, EVO lacks the dynamic adjustment that is necessary for real-time motion classification because it is designed to deal only with parameters rather than model structures. Compared to it, ALSTM has a flexible memory mechanism with hyperparameter sensitivity, and the model cannot adapt to the variety of motion patterns correctly without tuning hyperparameters. Further, ALSTM can end up in an overfitting situation, particularly when few data are used in training and its efficiency tends to drop sharply when input information is noisy. The EVO-ALSTM models are assisted by the evolutionary algorithms for adaptive optimization, which improves the ability of the model to optimize and escape out of local minima reducing the overall convergence times. This allows for improved generalization and accuracy when dealing with complex and nonlinear data distributions. Furthermore, due to its flexible structure, the architecture can be customized to fit the specific attributes of the dataset, thereby enhancing performance. The proposed EVO-ALSTM method addressed these limitations by integrating the strengths of both techniques: it leverages the efficient exploration that EVO affords to dynamically configure the ALSTM architecture, ensuring that the model not only learns effectively but also manages memory in real-time. This hybridization results in enhanced accuracy and robustness in motion classification tasks, effectively managing the weakness of the individual methods and enhancing overall performance.

5. Conclusion

In this paper, an Egyptian Vulture optimized Adjustable Long Short-Term Memory Network (EVO-ALSTM) was introduced for motion classification. The dataset was gathered from 30 participants. The wearable sensors, like accelerometers and gyroscopes, were used to collect human activity, including walking, jumping, arm waving, and sports actions, from the participants. Preprocessed the data by using Z score normalization and extracted the complex features by using PCA. The proposed EVO-ALSTM method was used as a classification to identify the motions. As a result, the four human activities measured for the motion detection accuracy of the proposed method showed that arm waving (95%) has a high detection motion accuracy level. The proposed EVO-ALSTM method was compared with the standard method (ALSTM) based on the metrics, including accuracy (95%), precision (93%), recall (90%), and F1-score (91%). According to the findings, the proposed method has superior performance than other methods to classify human activity and it helps to enhance the animation design.

Limitation and future scope

The framework's dependence on wearable sensors may restrict the range of motion statistics, and the computational complexity of the EVO-ALSTM approach will be resource-intensive for real-time processing. Additionally, generalization to various person environments and motions may require further optimization. This investigation opens avenues for similar studies into integrating biomimetic imaginative and visual structures with greater superior device mastering models for even greater accuracy in motion popularity. Future research may want to explore expanding the variety of human sports recorded and integrating this technology into various packages like gaming, digital learning, and interactive storytelling. Enhancing actual time feedback capabilities in VR and AR environments could also increase immersion.

Ethical approval: Not applicable.

Conflict of interest: The author declares no conflict of interest.

References

1. Wee, C., Yap, K.M. and Lim, W.N., 2021. Haptic interfaces for virtual reality: Challenges and research directions. IEEE Access, 9, pp.112145-112162.

2. Wang, G., Zheng, C., Fu, Y., Zhu, K., Lai, F., Zhang, L., Li, M., Wu, X., Ren, M., Zheng, Y. and Lian, B., 2024, July. KiPneu: Designing a Constructive Pneumatic Platform for Biomimicry Learning in STEAM Education. In Proceedings of the 2024 ACM Designing Interactive Systems Conference (pp. 441-458).
3. Han, L., Afzal, N., Wang, Z., Wang, Z., Jin, T., Guo, S., Gong, H. and Wang, D., 2024. Ambient haptics: bilateral interaction among humans, machines, and virtual/real environments in the pervasive computing era. CCF Transactions on Pervasive Computing and Interaction, pp.1-33.
4. Sauda, E., Karduni, A. and Lanclos, D., 2024. Architecture in the Age of Human-Computer Interaction. Taylor & Francis.
5. CHEN, Y., WU, X., ZHANG, J., LIU, Y. and LI, H., 2024. A Type of Human-Computer Collaborative.
6. Coban, M. and Coştu, B., 2023. Integration of biomimicry into science education: Biomimicry teaching approach. Journal of Biological Education, 57(1), pp.145-169.
7. Soliman, M.E. and Bo, S., 2023. An innovative multifunctional biomimetic adaptive building envelope based on a novel integrated methodology of merging biological mechanisms. Journal of Building Engineering, 76, p.106995.
8. Jović, B.S. and Mitić, A.D., 2020. Exploration of nature-based biomimetic approach in landscape architectural design: a parametric study of candelabra model design. Visual Computing for Industry, Biomedicine, and Art, 3, pp.1-11.
9. Li, S., 2021. Ancient Architecture Animation Design Method of 3D Technology and Its Application. In Journal of Physics: Conference Series (Vol. 2037, No. 1).
10. Pezent, E., Macklin, A., Yau, J.M., Colonnese, N. and O'Malley, M.K., 2023. Multisensory Pseudo-Haptics for Rendering Manual Interactions with Virtual Objects. Advanced Intelligent Systems, 5(5), p.2200303.
11. Gallerani, M., Vazzoler, G., De Novi, G., Razzoli, R., Berselli, G. and Ottensmeyer, M.P., 2023. Integrated design and prototyping of a robotic eye system for ocular and craniofacial trauma simulators. International Journal on Interactive Design and Manufacturing (IJIDeM), 17(6), pp.3103-3116.
12. Shreyas, D.G., Raksha, S. and Prasad, B.G., 2020. Implementation of an anomalous human activity recognition system. SN Computer Science, 1(3), p.168.
13. Mahbub, U. and Ahad, M.A.R., 2022. Advances in human action, activity, and gesture recognition. Pattern Recognition Letters, 155, pp.186-190.
14. Minh Trieu, N. and Thinh, N.T., 2024. Advanced Design and Implementation of a Biomimetic Humanoid Robotic Head Based on Vietnamese Anthropometry. Biomimetics, 9(9), p.554.
15. McMahon, M. and Erolin, C., 2024. Biomimicry–medical design concepts inspired by nature. Journal of Visual Communication in Medicine, 47(1), pp.27-38.
16. Dai, F. and Li, Z., 2024. Research on 2D Animation Simulation Based on Artificial Intelligence and Biomechanical Modeling. EAI Endorsed Transactions on Pervasive Health and Technology, 10.
17. Li, Y., Zhao, M., Yan, Y., He, L., Wang, Y., Xiong, Z., Wang, S., Bai, Y., Sun, F., Lu, Q. and Wang, Y., 2022. Multifunctional biomimetic tactile system via a stick-slip sensing strategy for human-machine interactions. NPJ Flexible electronics, 6(1), p.46.
18. Reddy, E.V., Manideep, K. and Agarwal, O., 2024. Machine Learning-Based Human Movements Mimicking System for Animation and Virtual Reality. Asian Journal of Research in Computer Science, 17(7), pp.84-94.
19. Gao, Q., 2022. Design and Implementation of a 3D Animation Data Processing Development Platform Based on Artificial Intelligence. Computational Intelligence and Neuroscience, 2022(1), p.1518331.
20. Vignesh, T., 2021, March. Pipeline for Development of 3-dimensional motion animation using 2-dimensional video. In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 991-995). IEEE.
21. Kong, D., Yang, G., Pang, G., Ye, Z., Lv, H., Yu, Z., Wang, F., Wang, X.V., Xu, K. and Yang, H., 2022. Bioinspired Co-Design of Tactile Sensor and Deep Learning Algorithm for Human–human-robot interaction. Advanced Intelligent Systems, 4(6), p.2200050.
22. Kim, S.J., Lee, Y.J. and Park, G.M., 2021. Real-Time Joint Animation Production and Expression System using Deep Learning Model and Kinect Camera. Journal of Broadcast Engineering, 26(3), pp.269-282.
23. Fu, Q., Fu, J., Zhang, S., Li, X., Guo, J. and Guo, S., 2021. Design of intelligent human-computer interaction system for hard of hearing and non-disabled people. IEEE Sensors Journal, 21(20), pp.23471-23479.
24. Qin, X., Xia, X., Ge, Z., Liu, Y. and Yue, P., 2024. The Design and Control of a Biomimetic Binocular Cooperative Perception System Inspired by the Eye Gaze Mechanism. Biomimetics, 9(2), p.69.

25. Lu, W., 2024. Learning-Based, Muscle-Actuated Biomechanical Human Animation: Bipedal Locomotion Control and Facial Expression Transfer (Doctoral dissertation, UCLA).
26. Zeng, X.S., Dwarakanath, S., Lu, W., Nakada, M. and Terzopoulos, D., 2021. Neuromuscular Control of the Face-Head-Neck Biomechanical Complex with Learning-Based Expression Transfer from Images and Videos. In Advances in Visual Computing: 16th International Symposium, ISVC 2021, Virtual Event, October 4-6, 2021, Proceedings, Part I (pp. 116-127). Springer International Publishing.
27. Kumarapu, L. and Mukherjee, P., 2021. Animepose: Multi-person 3d pose estimation and animation. Pattern Recognition Letters, 147, pp.16-24.
28. Jones, M., Byun, C., Johnson, N. and Seppi, K., 2023. Understanding the Roles of Video and Sensor Data in the Annotation of Human Activities. International Journal of Human-Computer Interaction, 39(18), pp.3634-3648.
29. Yadav, R.K., Arockiam, D. and Bhaskar Semwal, V., 2024. Motion Signal-based Recognition of Human Activity from Video Stream Dataset Using Deep Learning Approach. Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science), 17(3), pp.77-91.