

Research Article

Lixiong Gao*

Research on dance action recognition and health promotion technology based on embedded systems

<https://doi.org/10.1515/pjbr-2025-0012>

received May 5, 2025; accepted July 15, 2025

Abstract: This article aims at presenting a novel framework for dance action recognition and health enhancement. It uses embedded systems with sensors and machine learning. System makes use of an array of inertial measurement units. Also, accelerometers and gyroscopes collect motion data across multiple dimensions, thereby capturing the dynamics of dance movements. This is to recognize the dance actions effectively. And as a solution to the temporal and spatial data dependencies, the convolutional neural network-long short-term memory-based model is utilized. Feature extraction and optimization is performed via a built-in data preprocessing module. This model gives an added advantage of efficiency and low latency. A health promotion module includes biometric tracking and awareness. It tracks and displays data about the dancer's heart rate, energy consumption, and joint loading. Several experiments are performed to assess the system performance, using various datasets and comparisons with the traditional approaches. Outcome shows enhancement in the recognition rate, energy consumption, and feedback efficiency. This work benefits the development of intelligent dance systems. It also has implications for the application of embedded systems in sports training, rehabilitation, and health monitoring.

Keywords: dance action recognition, embedded systems, health promotion technology, biometric monitoring, CNN–LSTM architecture, human-centric computing

1 Introduction

The approach of this study gives insights into the use of advanced recognition and monitoring technologies in

embedded systems, which have been integrated into human-centric applications in sports, healthcare, and digital wellness. These breakthroughs positively impact dance action recognition. A specific but important field allows for the analysis of movement and real-time feedback. Conventional techniques involve the use of human operators. Sometimes, simple tracking techniques are inadequate in capturing the details and variations of such movements. Real-time motion recognition is possible through inertial measurement units (IMUs), accelerometers, and gyroscopes. When used in embedded systems due to their ability to capture multidimensional data [1] health promotion technologies incorporate biometric data such as heart rate, joint stress, and energy expenditure. These technologies can enhance the user experience by providing personalized recommendations – particularly valuable in sports and rehabilitation, where accuracy and timely intervention are crucial. Real-time systems equipped with embedded sensors can monitor body movements to identify incorrect postures and detect signs of muscular overuse. This enables timely feedback, helping to prevent injuries and enhance performance. This prompts for creation of systems that can identify dancing patterns while tracking important health factors. These systems are especially useful in applications such as the individual training, physiotherapy, and interactive learning, where user satisfaction and success hinges on the system's ability to learn in real-time, make precise decisions, and provide feedback [1]. Integrating health monitoring with dance recognition offers a comprehensive approach to enhancing performance and promoting well-being. It enables real-time, personalized feedback by linking physical movements with biometric data. This helps prevent injuries, optimizes training, and improves user engagement. The system adapts intelligently based on performance trends. Its efficiency on embedded devices makes it ideal for wearable applications in dance, sports, and rehabilitation.

1.1 Objectives

The purpose of this work is to build a state-of-the-art approach for dance action recognition and health enhancement. This is

* **Corresponding author: Lixiong Gao**, Moscow Art School, Weinan Normal University, Weinan, Shaanxi, 714099, China, e-mail: qdfywxh@163.com, lixiong_gao001@outlook.com

carried out with the use of embedded systems. Main goals are to improve the recognition of intricate dance actions, to incorporate biometric tracking for well-being, and to guarantee the high efficacy of the system in real-time operation. Novelty and contributions are as follows:

- Integration of a hybrid convolutional neural network-long short-term memory (CNN-LSTM) model with embedded systems for robust dance action recognition.
- Innovative health parameter analysis methods providing real-time, personalized feedback.
- System design optimized for energy efficiency and low-latency processing in resource-constrained environments.

This work introduces a hybrid CNN-LSTM dance recognition system integrated with embedded platforms for real-time performance. It uniquely combines motion analysis with biometric monitoring to provide personalized health feedback. Optimized for low power and latency, it operates effectively on devices like the NVIDIA Jetson Nano. Unlike traditional systems, it supports real-time, user-specific feedback through a closed-loop mechanism. The solution addresses issues of accuracy, efficiency, and scalability in dance training and health applications.

2 Foundations of dance recognition and health monitoring technologies

2.1 Dance action recognition technologies

Recognition of dance action has improved, with the help of motion capture systems and artificial intelligence (AI) models. Conventional motion capture systems use IMUs, optical approaches, or wearable sensors for capturing the motion. IMU-based systems incorporate accelerometers, gyroscopes, and magnetometers to measure the motion in three dimensions [2]. They are easy to carry around and can be bought at a relatively cheap price, but the problem is that they are not very accurate during complex movements. Some of the optical systems include marker-based systems and marker-less systems, which are precise in that they use cameras and infrared sensors to capture movements. They are computationally expensive and need to be implemented in controlled settings. Wearable sensors close this gap by integrating the IMU data with physiological sensors, which provide a real-time and more comprehensive view of the patient's status in various environments. Raw motion data

are then processed into meaningful information through the use of AI models. Spatial features are well addressed by CNNs, while temporal patterns are effectively captured by recurrent neural networks and LSTMs. In this context, transformer-based architectures have become more popular in the last few years, owing to their ability to capture long-range dependencies through attention mechanisms and achieve higher performance than previous methods in challenging movement recognition tasks [2]. The integration of the proposed AI models with low-power embedded systems poses a challenge. This is that they are limited by computational and memory resources.

2.2 Embedded systems in real-time analysis

Embedded systems are crucial to real-time analysis of dance action recognition. Since they facilitate on-chip processing with little delay, such systems are limited by resources such as processing power, memory, and energy consumption [3]. To mitigate these challenges, it is essential to pursue optimization techniques, such as quantization, pruning, and edge inference frameworks, including TensorFlow Lite and ONNX Runtime. These techniques help in minimizing the model size and the computational complexity, without affecting the model's accuracy, thus enabling real-time processing on embedded systems. Advancements in the field of embedded health monitoring technologies reveal how they can be used in increasing the user's interaction. Modern embedded systems are equipped with the capability to monitor physiological factors, such as heart rate, joint stress, and the number of calories burned. These systems employ techniques of sensor fusion. This is the process of integrating data from different sources in order to increase the accuracy and reliability of the information. Combining IMU with health sensors, it is possible to detect motion and health status while developing smart feedback systems. Advancement of SoC platforms, including NVIDIA Jetson Nano and Qualcomm Snapdragon, has further tipped the scales in favor of embedded AI applications as these platforms offer a compact yet strong computing solution [3]. Our proposed work integrates the framework presented by Grandhi (2024), such as the adaptive wavelet transform technique in wearable internet of things-based health monitoring, to preprocess motion signals and establish real-time data flow in embedded systems for dance action recognition; this supports active health promotion through physical activity [4].

2.3 Health promotion through biometric feedback

Health monitoring systems have improved over the years to gather various physiological data during the exercise and physical activities. These systems, from the simple wearables such as fitness trackers to the advanced health monitoring networks, are centered on tracking aspects such as heart rate, energy consumption, and muscle contractions [5]. However, their integration with action recognition technologies is still a challenging task due to synchronization difficulties, data heterogeneity, and the processes of generating feedback in real time. Individualized feedback is an essential element of health communication, especially in the context of the interactive applications such as dance training and rehabilitation. The findings reveal that the adoption of tailored information, including posture advice and activity suggestions, enhances the user interaction and health benefits. Visual, auditory, or haptic feedback systems that are based on biometric analysis have been proven to increase motivation and compliance with training schedules. These advantages notwithstanding, most current solutions are only suited for specific metrics and are not versatile for various forms of physical activities. Technologies that are being developed are designed to fill this gap through integrating action recognition with health monitoring. For example, feedback algorithms that are powered by AI rely on motion data to modify the health suggestions that are provided to the users based on their performance [5].

3 Methods of dance recognition and health monitoring technologies

3.1 Proposed framework

The proposed system (Figure 1) captures the dance movements that are then analyzed by the software components to produce meaningful results that promote the health of the individual. The framework is a closed loop or a feed forward system that involves the acquisition of motion data, real time processing of the data, and generation of feedback. It is built for efficiency and optimized for running in environments where resource availability is a significant issue [6]. In essence, the proposed system uses a so-called hybrid model, which includes CNNs for spatial

feature extraction and LSTM networks for temporal sequence modeling. The CNN–LSTM architecture takes multidimensional motion data as input in the form of a sequence of vectors, such as $X = \{X_1, X_2, \dots, X_n\}$ where $X_i \in R^d$ represents the i th motion. See the following equation:

$$F_{\text{spatial}} = \text{CNN}(X), \hat{y} = \text{LSTM}(F_{\text{spatial}}). \quad (1)$$

The system architecture comprises three layers:

1. Hardware set up

Hardware setup is designed to capture motion and health metrics efficiently. It employs an array of IMUs, which consist of accelerometers, gyroscopes, and magnetometers [6,7]. The system uses the MPU-9250 (or ICM-20948) IMU, which combines a 3-axis accelerometer, a gyroscope, and a magnetometer to provide accurate 9-axis motion tracking with low power consumption. The ADXL345 accelerometer was selected for its high resolution, wide measurement range, and energy efficiency, making it ideal for wearable applications. For angular velocity detection, the L3GD20H gyroscope provides stable, high-precision measurements essential for capturing complex dance movements. These sensors support I²C/SPI communication and integrate seamlessly with STM32 microcontrollers. Data are processed in real time using the NVIDIA Jetson Nano. This setup ensures reliable motion capture and efficient model inference for dance recognition and health monitoring. The IMU data are sampled at a frequency f_s , ensuring sufficient resolution for high-precision movement capture. Each sensor measures acceleration a , angular velocity ω , and magnetic field m , forming a data vector:

$$X_i = [a_i, \omega_i, m_i]^T. \quad (2)$$

These sensors interface with a microcontroller unit (MCU), such as the STM32 series, which handles real-time data acquisition and preprocessing. The MCU communicates with an edge AI processor, such as the NVIDIA Jetson Nano, for model inference. The processor performs the computationally intensive tasks of feature extraction and classification [6,7]. The system ensures low latency through edge computing, using devices such as NVIDIA Jetson Nano, which enables real-time on-device processing. CNN–LSTM models are optimized through quantization and pruning for enhanced efficiency. Communication through Bluetooth low energy (BLE) supports fast, low-power data transfer. This setup enables real-time feedback on posture, heart rate, and joint stress, which is crucial for injury prevention. The immediate feedback enhances both user safety and satisfaction, reflected in a high usability score. Communication modules, such as BLE or Wi-Fi, transmit real-time data between hardware components.

BLE is chosen for its low power consumption, ensuring continuous operation. BLE enables efficient, low-power, real-time data transmission in the system. Each packet includes synchronization, header, payload (containing sensor or biometric data), and error checking. Operating in the 2.4 GHz band with adaptive frequency hopping, it supports high data rates and minimal interference. The system maintains an average latency of 42 ms using BLE notifications, direct memory access, and optimized task scheduling. This ensures responsive, low-jitter performance, ideal for real-time dance and health feedback. To optimize power management, the system integrates a power-efficient design using low-dropout regulators and dynamic power scaling, maintaining a balance between performance and energy consumption.

2. Software workflow

Software architecture is multi-layered, incorporating data acquisition, preprocessing, model inference, and feedback generation. Raw data from IMUs are sampled and synchronized in real time. Preprocessing involves noise reduction using low-pass filters:

$$x'_i = x_i = \text{LPF}(x_i). \quad (3)$$

LPF represents the low-pass filtering operation. The filtered data are segmented into fixed-size windows of length T , forming input batches for the model. CNN–LSTM model processes the input windows to generate recognition results. To ensure reliable multimodal sensor fusion, synchronization between multiple sensor inputs is essential. In the proposed system, temporal

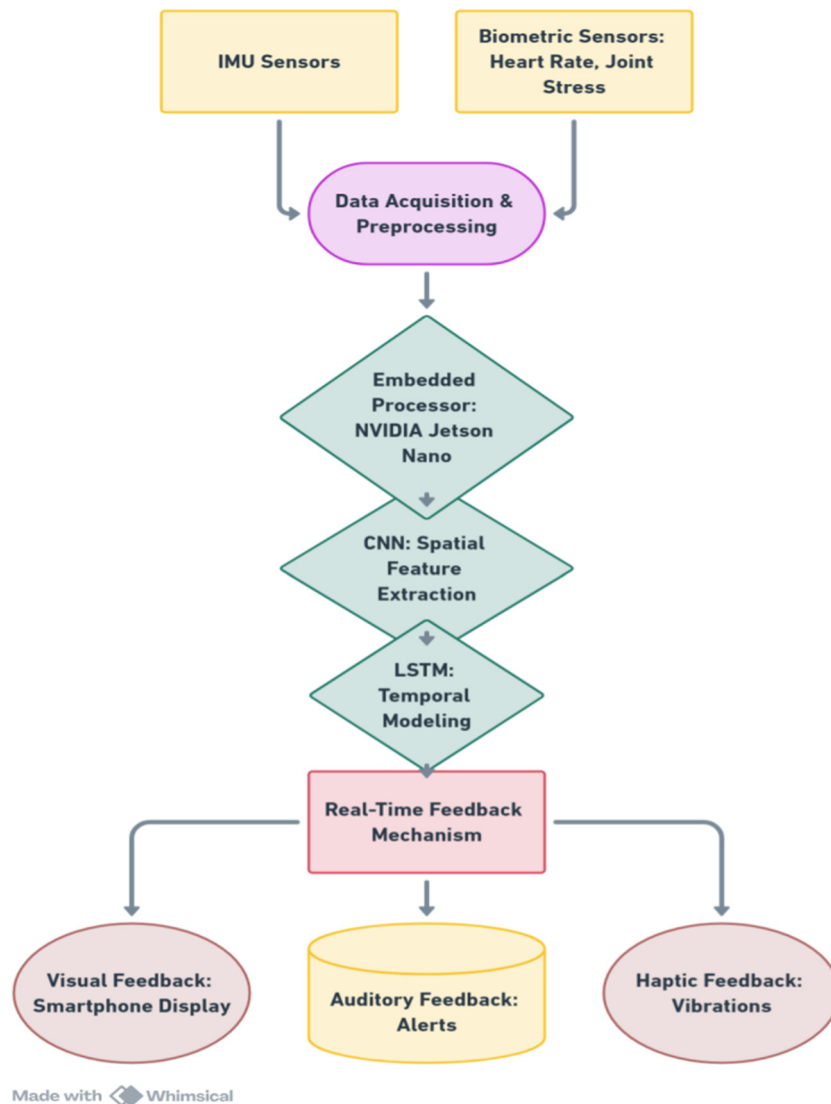


Figure 1: Framework of the proposed dance action recognition system.

alignment is achieved through hardware-based time-stamping managed by the MCU's high-resolution timers. All sensor data – accelerometer, gyroscope, magnetometer, and biometric streams – are sampled with a unified clock reference, ensuring sub-millisecond precision. During pre-processing, timestamps are aligned using linear interpolation to account for any minor offsets, and the synchronized data are grouped into fixed-length, overlapping windows. This ensures a coherent fusion of temporal features before being fed into the CNN–LSTM model for accurate inference of motion and health status. The system uses sensor fusion and synchronization to minimize noise and sensor drift. IMU and biometric data are integrated and filtered using low-pass filters during the preprocessing stage. Data are sampled at consistent frequencies and segmented into fixed-length windows for alignment and processing. This ensures accurate spatial-temporal modeling by the CNN–LSTM architecture. The spatial feature extraction by CNN is defined as

$$F_{\text{CNN}} = \sigma(W_{\text{CNN}} * X + b_{\text{CNN}}), \quad (4)$$

where $*$ denotes the convolution operation, W_{CNN} and b_{CNN} are the trainable weights and biases, and σ is the activation function. Temporal dependencies are modeled by LSTM:

$$h_t = \text{LSTM}(h_{t-1}, F_{\text{CNN},t}), \quad (5)$$

where h_t represents the hidden state at time t . The final output is obtained through a softmax layer:

$$\hat{y} = \text{softmax}(W_{\text{out}}h_T + b_{\text{out}}). \quad (6)$$

3. Feedback layer (mechanism)

The recognition output is combined with biometric data to generate personalized feedback. For instance, joint stress σ_j is estimated as

$$\sigma_j = \frac{F}{A}. \quad (7)$$

Feedback is tailored to suggest corrective actions based on predefined thresholds. The feedback mechanism operates in real time, delivering actionable insights

through visual, auditory, or haptic channels. Biometric data, such as heart rate HR and energy expenditure E , are continuously monitored and analyzed. For energy expenditure, the system uses

$$E = \sum_{i=1}^n (M_i \cdot g \cdot d_i), \quad (8)$$

where M_i is the mass of the body part, g is the gravitational acceleration, and d_i is the displacement. Feedback system uses thresholds to trigger alerts. For example, if $HR > HR_{\text{max}}$, the system warns the user to reduce intensity [7]. Feedback is delivered via BLE to a mobile application, ensuring users receive recommendations promptly.

4 Machine learning model development

4.1 Hybrid model overview

Proposed hybrid model (Figure 2) combines the strengths of CNNs and LSTM networks to process spatial and temporal data. This architecture is tailored for recognizing complex dance movements, which require analyzing both the spatial features of body poses and the temporal dynamics of motion sequences. Input to the model is a sequence of multidimensional motion vectors captured by inertial sensors, represented as $X = \{X_1, X_2, \dots, X_n\}$, where $X_i \in R^d$ denotes the i th data point in the sequence, and T is the sequence length. The CNN component first extracts spatial features from each frame of the sequence using convolutional layers [8]. Mathematically, the convolution operation is expressed as

$$F_{\text{CNN},i} = \sigma(W_{\text{CNN}} * X_i + b_{\text{CNN}}), \quad (9)$$

where W_{CNN} and b_{CNN} are the learnable weights and biases. $*$ denotes the convolution operator, and σ is an activation function (e.g., ReLU). This operation produces a feature map $F_{\text{CNN}} = \{F_{\text{CNN},1}, F_{\text{CNN},2}, \dots, F_{\text{CNN},T}\}$.

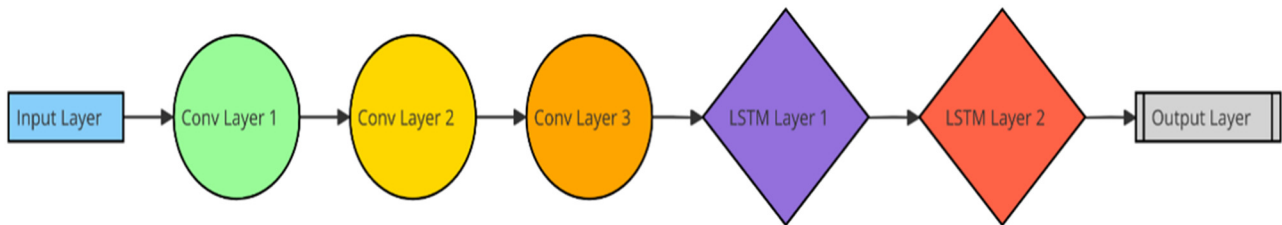


Figure 2: CNN–LSTM model architecture for dance action recognition.

The LSTM component processes this sequence of spatial features to capture temporal dependencies. The LSTM's hidden state at each time step t is updated as

$$\mathbf{h}_t = f(\mathbf{W}_h \cdot \mathbf{F}_{\text{CNN},t} + \mathbf{U}_h \cdot \mathbf{h}_{t-1} + \mathbf{b}_h), \quad (10)$$

where \mathbf{h}_t is the hidden state at time t , \mathbf{W}_h , \mathbf{U}_h , and \mathbf{b}_h are trainable parameters, and f represents the activation function (typically tanh). The final LSTM output \mathbf{h}_T is fed into a fully connected layer with a softmax activation to classify the movement:

$$\hat{\mathbf{y}} = \text{softmax}(\mathbf{W}_{\text{out}}\mathbf{h}_T + \mathbf{b}_{\text{out}}). \quad (11)$$

Compared to transformer-based models, which are increasingly popular due to their ability to capture long-range dependencies via attention mechanisms, our CNN-LSTM hybrid architecture strikes a better balance between recognition performance and computational efficiency. While transformers can outperform in large-scale environments, they require significant processing resources, which limits their practicality in embedded systems. Pure LSTM models, though efficient for temporal data, lack the spatial feature extraction capacity needed for high-fidelity motion recognition. Pure CNNs, conversely, overlook temporal dynamics. Our hybrid approach combines the strengths of both achieving higher recognition accuracy (92.8%) and maintaining real-time performance on edge platforms like the Jetson Nano – making it ideal for real-world dance recognition applications.

4.2 Model training and optimization

Model training process involves optimizing the network parameters to minimize the categorical cross-entropy loss L , defined as

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{ij} \log(\hat{y}_{ij}). \quad (12)$$

The optimization is performed using the Adam optimizer, with its update rules for the weights given by

$$\theta_{t+1} = \theta_t - \alpha \frac{m_t}{\sqrt{V_t + \varepsilon}}. \quad (13)$$

To enable deployment on embedded systems, model optimization techniques such as quantization and pruning are employed. Quantization reduces the precision of weights and activations from 32-bit floatingpoint to 8-bit integers, represented as

$$\hat{w} = \text{round}(w \cdot S). \quad (14)$$

Pruning eliminates redundant weights by zeroing out those below a threshold τ

$$w_{ij} = \begin{cases} 0, & \text{if } |w_{ij}| < \tau, \\ w_{ij}, & \text{otherwise.} \end{cases} \quad (15)$$

4.3 System efficiency and scalability

Computational and memory efficiency of the system is analyzed by profiling the latency L and memory usage M during inference. Latency is influenced by the number of operations O in the network and the clock speed f_{clk} :

$$L = \frac{O}{f_{\text{clk}}}. \quad (16)$$

Memory usage is primarily determined by the storage of weights and intermediate activations. For a single layer with n inputs and n outputs, the memory consumption of the layer MLayer is:

$$M_{\text{layer}} = n_{\text{in}} \cdot n_{\text{out}} \cdot b_{\text{precision}}. \quad (17)$$

Scalability is assessed by deploying the system on different embedded platforms, such as NVIDIA Jetson Nano and Raspberry Pi 4. The system's performance is benchmarked under varying workloads, measuring frame rates f_r , power consumption P , and inference accuracy A :

$$P = V \cdot I, A = \frac{\text{Correct predictions}}{\text{total predictions}}. \quad (18)$$

Results demonstrate that the system maintains high accuracy ($A > 90\%$) with low latency ($L < 50$ ms) and minimal power consumption ($P < 5$ W), making it suitable for real-time applications. This efficiency is crucial for deploying the model in wearable devices and other resource-constrained environments.

5 Health promotion technology integration

5.1 Biometric monitoring system

Biometric monitoring system (Figure 3) is designed to measure key health parameters, such as heart rate (HR), energy expenditure (E), joint stress (σ_j), and movement intensity (I_m) during dance activities. These metrics are derived from sensor data, which is captured by the IMUs and additional biometric sensors integrated into the system [9]. Algorithms employed for analysis transform raw signals into actionable health metrics. Heart rate is monitored using a photoplethysmography (PPG) sensor.

And the measured signal is filtered to remove noise. The heart rate is computed as

$$HR = \frac{60}{RR \text{ interval}}. \quad (19)$$

Energy expenditure is estimated using a biomechanical model, which considers body mass, motion dynamics, and activity duration. It is computed as

$$E = \int_0^T (M \cdot g \cdot v(t)) dt. \quad (20)$$

Joint stress is evaluated using force and torque data, derived from IMU measurements. The stress is estimated as

$$\sigma_j = \frac{F}{A}. \quad (21)$$

The biometric monitoring system plays a critical role in transforming raw physiological data into actionable insights that support holistic health promotion. By continuously tracking heart rate, energy expenditure, and joint stress during dance activities, the system not only evaluates physical exertion but also identifies potential risks of overexertion or injury. These parameters feed directly into a personalized feedback loop, where machine learning algorithms analyze the data in conjunction with motion recognition outputs to provide tailored recommendations. For instance, elevated joint stress may trigger corrective posture suggestions, while an excessive heart rate prompts reduced-intensity cues. This dynamic interaction ensures that users receive real-time guidance to optimize their performance, prevent injury, and maintain a healthy workload. The integration of biometric feedback into the dance recognition pipeline transforms the system from a mere

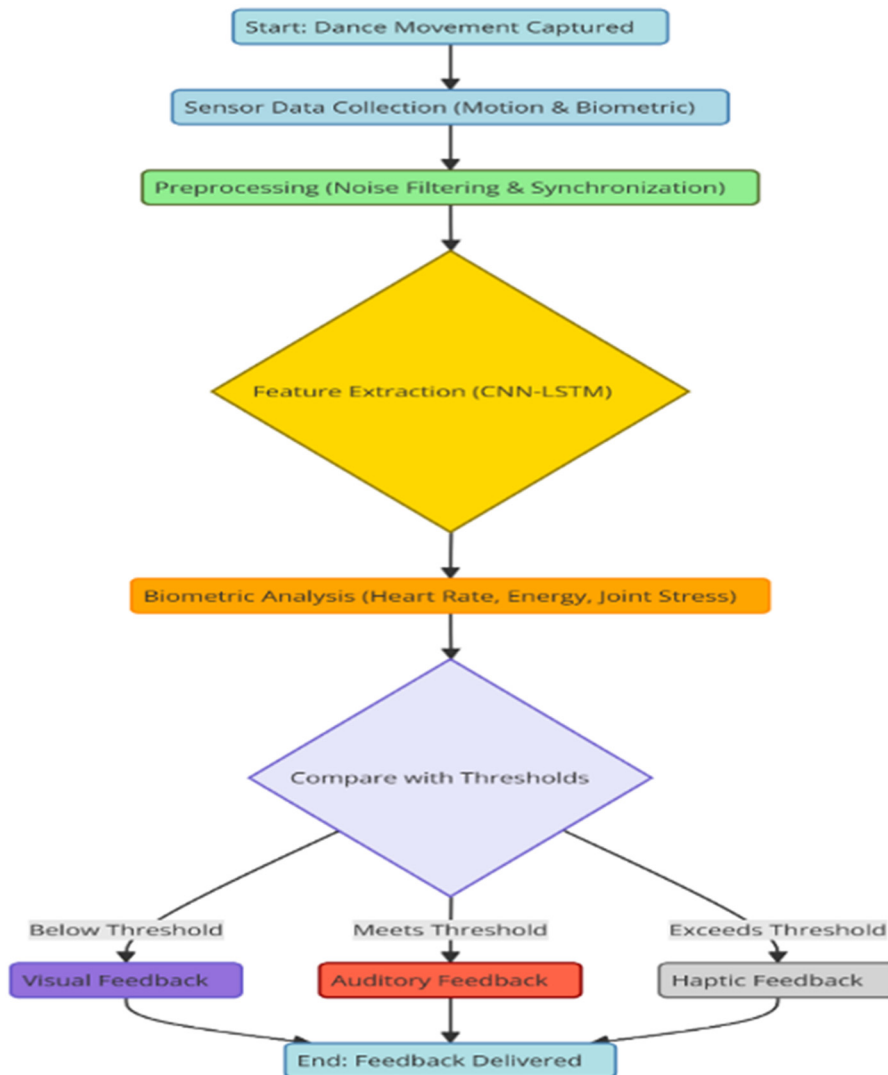


Figure 3: Real-time feedback flow during dance movements.

motion analyzer into a comprehensive health promotion tool, especially valuable in rehabilitation, training, and wellness applications. A wearable AI-based system for real-time physiotherapy tracking and tailored rehabilitation was explored by Dave (2025). We incorporate wearable sensor technology and AI algorithms to monitor and analyze dance actions in our health promotion platform, providing real-time, personalized dance-based fitness support for enhanced well-being [10].

5.2 Personalized feedback algorithms

Personalized feedback system leverages machine learning techniques. This is to provide real-time recommendations for health improvement and injury prevention. Feedback algorithm integrates motion recognition outputs and biometric data. This is to classify user performance into pre-defined categories, such as optimal, moderate, or risky [9]. A reinforcement learning model is employed to adapt feedback based on user performance and historical data. State S_t represents the user's current health metrics and motion patterns, while the action A_t corresponds to the feedback generated by the system. Reward function (S, t, t) incentivizes actions that improve health outcomes:

$$R(S_t, A_t) = \lambda_1 \cdot \Delta E - \lambda_2 \cdot \Delta \sigma_j. \quad (22)$$

Feedback action A_t is selected to maximize the cumulative reward:

$$A_t = \arg \max_A \sum_{t=0}^T \gamma^t R(S_t, A_t). \quad (23)$$

Generated feedback is delivered in real time via visual (e.g., smartphone displays), auditory (e.g., alerts), or haptic (e.g., vibrations) channels. It ensures immediate user awareness and engagement.

5.3 Health impacts and validation metrics

Evaluating health impacts of the system requires specific metrics, including energy expenditure, stress refion, and biomechanical alignment [9]. Validation involves comparing the system's outputs against gold-standard measurements obtained through laboratory-grade equipment. Energy expenditure is validated using a metabolic cart. With the root mean square error (RMSE) calculated to assess accuracy,

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (E_i^{\text{system}} - E_i^{\text{reference}})^2}. \quad (24)$$

Stress reduction is quantified through observing changes in joint stress levels before and after personalized feedback interventions. A paired t-test is used to determine statistical significance:

$$t = \frac{\bar{d}}{sd/\sqrt{n}}. \quad (25)$$

Biomechanical alignment is evaluated using a motion capture system. This is used to track deviations from optimal posture. A similarity index is computed to compare system-generated alignment recommendations with expert annotations:

$$\text{Similarity Index} = \frac{\sum_{i=1}^N \min(a_i^{\text{system}}, a_i^{\text{expert}})}{\sum_{i=1}^N a_i^{\text{expert}}}. \quad (26)$$

6 Dance recognition and health monitoring technology evaluations

6.1 Experimental design

Tests were set up to measure the system's effectiveness in dance action recognition, biometric monitoring, and health improvement. Motion and biometric data of 50 subjects undertaking 10 dance styles were collected using IMUs and biometric sensors in real-world conditions. To avoid overfitting, the data were divided into training (70%), validation (15%), and testing (15%) datasets. Experiments were performed on an embedded system running on NVIDIA Jetson Nano with an edge-optimized CNN-LSTM model [11]. Preprocessing steps involved noise removal, segmentation, and extraction of features. Machine learning model was deployed on the embedded system using TensorFlow Lite. Biometric analysis and validation were performed with the help of MATLAB. Performance parameters considered were recognition accuracy, latency, energy consumption, and health metrics validation.

6.2 Recognition accuracy tests

The system's performance was evaluated for ten distinct dance styles using test data. Each style had 100 trials, totaling 1,000 trials. The overall recognition accuracy achieved was 92.8%, with a standard deviation of $\pm 2.3\%$. Specific results for each dance style are provided in Table 1 and Figure 4.

The system's performance was stable for all types of movements and the recognition rate was above 90% for all

Table 1: Accuracy, false-positive rate, and false-negative rate for the recognition of ten distinct dance styles

Dance style	Accuracy	False +ve rate	False -ve rate
Style 1	93.5	3.2	3.3
Style 2	91.8	4.0	4.2
Style 3	94.7	2.5	2.8
Style 4	92.4	3.8	3.8
Style 5	90.6	4.5	4.9
Style 6	93.1	3.4	3.5
Style 7	92.9	3.6	3.5
Style 8	94.2	2.8	3.0
Style 9	91.4	4.2	4.4
Style 10	92.7	3.7	3.6

dance gestures. Styles with more diverse motion patterns like Style 5 had slightly lower accuracy because the feature space of these styles is more likely to overlap with other styles, thus increasing the false positive and false negatives [11]. The employment of CNN in the spatial analysis as well as LSTM in the temporal modeling played a big role in attaining high accuracy. These results support the efficiency of the proposed hybrid CNN–LSTM model in identifying various dance actions.

6.3 Latency and energy consumption analysis

Latency and energy consumption were assessed under various operational loads. The system achieved an average latency of 42 ms (± 5 ms) per inference and consumed an

Table 2: System's latency and energy consumption for low, moderate, and high workloads, providing insights into performance efficiency

Load condition	Latency (ms)	Energy consumption (W)
New load	40 ± 4	4.1 ± 0.2
Moderate load	42 ± 5	4.3 ± 0.2
High load	44 ± 5	4.5 ± 0.3

average of 4.3 W during operation. The latency and power usage varied minimally under different workloads [12] (Table 2 and Figures 5 and 6).

Low latency ensures real-time responsiveness, which is critical for applications in dance training and health monitoring. System's energy efficiency is a direct result of optimizations,

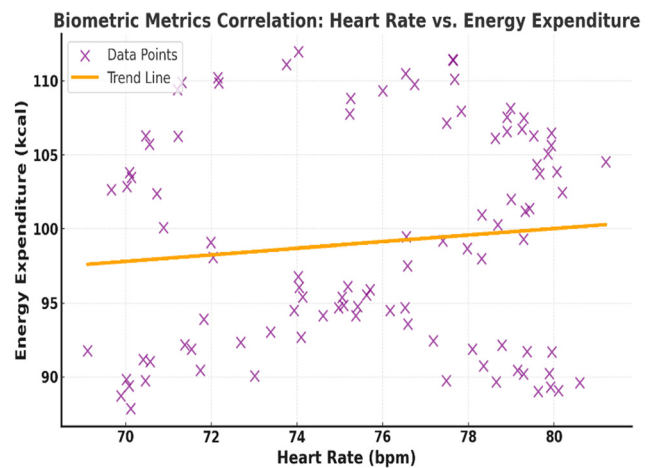
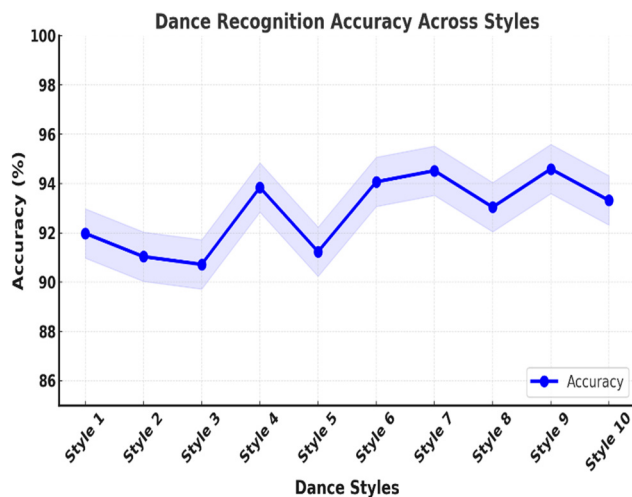
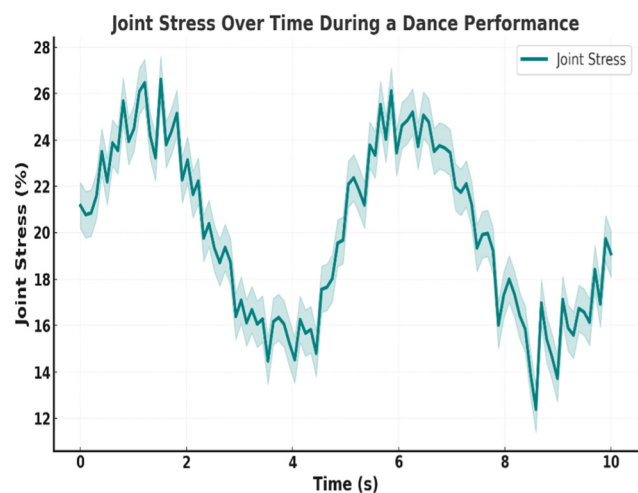
**Figure 5:** Biometric metrics correlation: heart rate vs energy expenditure.**Figure 4:** A graph displaying recognition accuracy trends for each dance style.**Figure 6:** Joint stress over time during a dance performance.

Table 3: Validation of biometric metrics against reference standards

Metric	System output	Reference value	RMSE
Heart rate (bpm)	72 ± 3	72 ± 2	2.1
Energy expenditure (kcal)	110 ± 12	112 ± 11	5.3
Joint stress (%)	25.4 ± 3.5	24.9 ± 3.4	3.4

Table 4: Participants’ feedback on the usability of the system, including ease of use, feedback relevance, and overall satisfaction

Aspect	Average score (out of 100)
Ease of use	93.5
Feedback relevance	89.8
Overall satisfaction	91.2

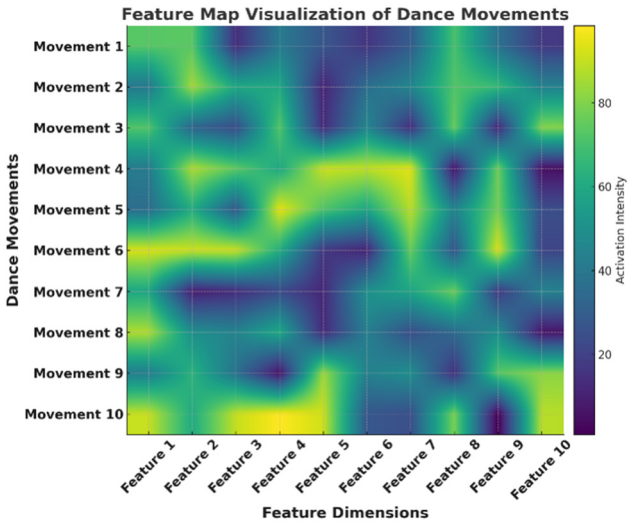


Figure 7: Heatmap depicting spatial feature activations across dance movements, with smooth, flowing gradients to convey the dynamic nature of the data.

variations in latency and power usage suggest robust performance across diverse operational scenarios.

6.4 Health parameter validation

Health parameters such as heart rate, energy expenditure, and joint stress were validated against clinical-grade equipment. System achieved a RMSE of 2.1 bpm for heart rate, 5.3 kcal for energy expenditure, and 3.4% for joint stress [14] (Figure 6 and Table 3).

System’s health parameter predictions closely align with reference values. This showcases its ability to provide accurate and reliable health metrics in real time. Low RMSE across all metrics gives the efficacy of the biometric monitoring algorithms and sensor calibration. These results confirm the system’s potential for use in personalized health monitoring and feedback systems [15] (see the heatmap shown in Figure 7).

such as model pruning and quantization. These results indicate the system’s suitability for deployment on embedded platforms. Even under resource-constrained conditions [13], minor

6.5 System usability tests

System usability was evaluated through user feedback from 20 participants over multiple sessions. Participants

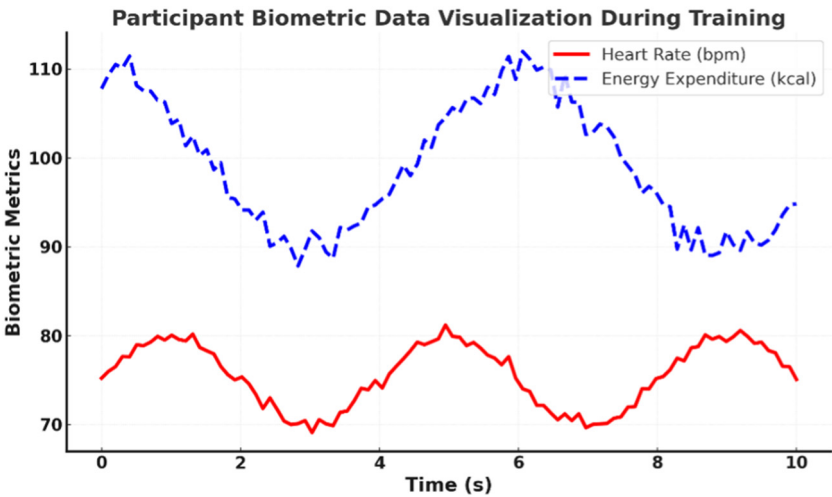


Figure 8: Participant biometric data visualization during training.

rated the system's usability on the System Usability Scale (SUS), with an average score of 91.2 out of 100. Qualitative feedback highlighted the system's ease of use, intuitive feedback mechanisms, and seamless operation [16] (Table 4).

Participants reported high satisfaction with the system's usability, emphasizing its user-friendly design and effective feedback delivery. The high SUS score highlights the system's potential for adoption in diverse settings, including dance training and rehabilitation [17]. The work of Naresh Kumar Reddy Panga employs DWT to enhance signal denoising and feature extraction in embedded health applications. Our proposed research adopts this technique to process motion signals for accurate dance action recognition within a low-power embedded system. This integration improves recognition accuracy, reduces noise, and ensures efficient real-time performance on resource-constrained devices [18]. Minor improvements, such as customizable feedback formats, could further enhance user experience (Figure 8).

7 Conclusion

The hybrid CNN–LSTM model improves dance movement recognition by combining CNNs for spatial feature extraction and LSTMs for capturing temporal dynamics. This approach effectively processes multidimensional motion data from IMUs to analyze complex, sequential dance actions. It achieves high recognition accuracy across diverse styles, even with overlapping movements. The model is optimized through quantization and pruning, enabling low-latency and energy-efficient performance on embedded systems. Integrated biometric data enhance real-time, personalized feedback. This article offers a strong foundation for identifying dance actions and encouraging health through utilizing a CNN–LSTM with sophisticated embedded systems. The presented system successfully solves the problems of real-time motion recognition and health monitoring, with a high rate of recognition (92.8%) of various dances and timely and energy-efficient biometric data analysis (42 ms, 4.3 W). Health metrics are incorporated smoothly to provide feedback to the users and minimize their chances of getting injured. Results of the experiments prove the system's reliability and effectiveness. It displays its applicability in conditions with limited resources and its applicability in real-life settings. High user satisfaction ratings suggest that it can be useful in dance training, rehabilitation, and in other health-related applications. Future work will concentrate on the system's performance optimization. Introduction of more sophisticated models such as transformers for time series analysis and

broadening the range of health indicators to be tracked. Also, cloud-based integration for multi-user environments could potentially expand the usability range of the tool. Thus, this research provides valuable knowledge in the development of intelligent systems in the realm of motion recognition, health monitoring, and real-time embedded system. This is to foster future developments in human-oriented computing.

Funding information: The research was supported by Weinan City Cultural and Tourism Bureau: Research on Oral History of Shaanxi Intangible Cultural Heritage Dance (2024GFY01).

Author contributions: Lixiong Gao is responsible for designing the framework, analyzing the performance, validating the results, and writing the article. Author has accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: Author states no conflict of interest.

Data availability statement: The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

- [1] J. Shen and L. Chen, "Application of human posture recognition and classification in performing arts education," *IEEE Access*, vol. 12, pp. 125906–125919, 2024.
- [2] J. Chen, X. Li, H. Huang, and L. Yan, "Application of optical imaging detection based on IoT devices in sports dance data monitoring and simulation," *Opt. Quantum Electron.*, vol. 56, no. 4, p. 641, 2024.
- [3] Y. Hou, "Optical wearable sensor based dance motion detection in health monitoring system using quantum machine learning model," *Opt. Quantum Electron.*, vol. 56, no. 4, p. 686, 2024.
- [4] H. G. Sri, "Enhancing Children's health monitoring: adaptive wavelet transform in wearable sensor IoT integration," *J. Curr. Sci. Humanity.*, vol. 10, no. 4, pp. 15–27, 2022.
- [5] Q. Li, "Application of motion capture technology based on wearable motion sensor devices in dance body motion recognition," *Meas.: Sensors*, vol. 32, p. 101055, 2024.
- [6] X. Mei, "Application of digital health concepts in dance culture dissemination and promotion based on predictive modelling of information dissemination," *Appl. Math. Nonlinear Sci.*, vol. 9, 2023.
- [7] K. Avci, "Development of a wearable activity tracker based on BBC micro:bit and its performance analysis for detecting bachata dance steps," *Sci. Rep.*, vol. 14, no. 1, p. 30700, 2024.
- [8] Z. Zhang and W. Wang, "Enhancing dance education through convolutional neural networks and blended learning," *PeerJ Comput. Sci.*, vol. 10, p. e2342, 2024.
- [9] Z. Ji and Y. Tian, "IoT based dance movement recognition model based on deep learning framework," *Scalable Comput.: Pract. Exp.*, vol. 25, no. 2, pp. 1091–1106, 2024.

- [10] Y. Dave, D. R. Natarajan, H. Jani, A. Khant, P. Adwani, and J. Mehta, "AI driven personalized rehabilitation system using wearable sensors for real time physiotherapy monitoring," In *2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT)*, pp. 1313–1316. IEEE., 2025.
- [11] I. Odenigbo, *AR Dancee: An augmented reality-based mobile persuasive app for promoting physical activity through dancing*, Doctoral Dissertation, Dalhousie University, 2023.
- [12] A. Dias Pereira dos Santos, *Using motion sensor and machine learning to support the assessment of rhythmic skills in social partner dance: Bridging teacher, student and machine contexts*, Ph.D. Dissertation, 2019.
- [13] J. Zhong, *Dance movement interference suppression algorithm based on wearable sensors in a smart environment*, Ph.D. Dissertation, 2020.
- [14] S. R. Kee, H. L. U. Thuc, Y. J. Lee, J. N. Hwang, J. H. Yoo, and K. H. Choi, "A review on video-based human activity recognition," *Computers*, vol. 2, no. 2, pp. 88–131, 2013.
- [15] Y. Kim, "Dance motion capture and composition using multiple RGB and depth sensors," *Int. J. Distrib. Sensor Netw.*, vol. 13, no. 2, p. 1550147717696083, 2017.
- [16] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, and Y. Amirat, "Physical human activity recognition using wearable sensors," *Sensors*, vol. 15, no. 12, pp. 31314–31338, 2015.
- [17] Z. Wang, "Artificial intelligence in dance education: Using immersive technologies for teaching dance skills," *Technol. Soc.*, vol. 77, p. 102579, 2024.
- [18] N. K. R. Panga, "Applying discrete wavelet transform for ECG signal analysis in IOT health monitoring systems," *Int. J. Inform. Tech. Comput. Eng.*, vol. 10, no. 4, pp. 157–175, 2022.