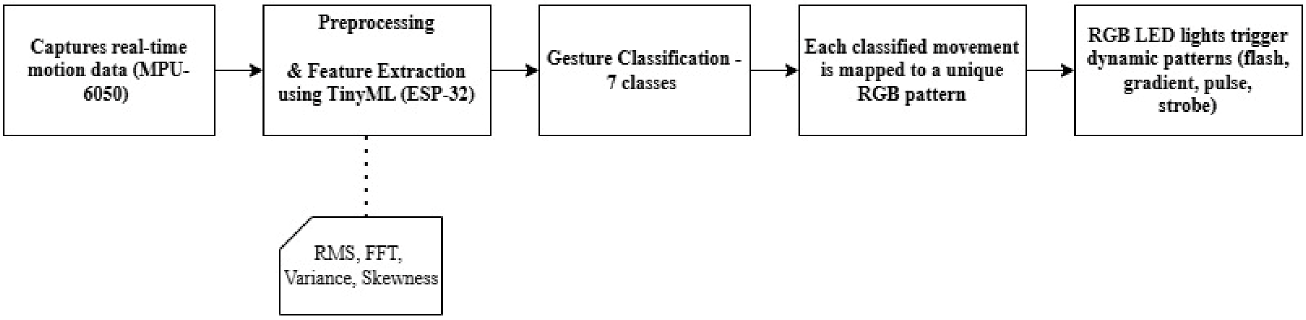
Smart Dance Shoes with Machine Learning Powered Light and Motion Synchronization

W.G.L. Harshani, E.M.R.S. Jayaweera, D.P.G.A.H. Kulathilaka, R.M.D.D. Malinda, H.K.S. Theekshan and P. Ravindra S De Silva

Department of Computer Science, University of Sri Jayewardenepura, Nugegoda, Sri Lanka

Date Received: 23-09-2025 Date Accepted: 25-12-2025



Abstract

In the era of wearable technology, integrating machine learning into performance arts opens new dimensions for user interaction and creativity. This project presents the development of Smart Dance Shoes that utilize motion sensors and machine learning algorithms to deliver real-time RGB light synchronization based on dance movements. The system is built using the ESP32-S3 microcontroller and the MPU-6050 sensor, which capture accelerometer and gyroscope data from the dancer's movements. These data inputs are processed through a machine learning model developed on Edge Impulse, which classifies different dance gestures such as jumps, spins, and steps and triggers corresponding lighting effects to enhance visual performance. The hardware is designed to be lightweight, portable, and user-friendly, making it suitable for dancers, performers, and fitness enthusiasts. Key components include RGB LED neon strips, a 3.7V LiPo battery, and Bluetooth integration for wireless customization. Testing covered unit, integration, and performance evaluations to ensure stability, low latency, and energy efficiency. Future improvements include multi-shoe synchronization, music-responsive lighting, and advanced models such as LSTM. This project demonstrates the potential of intelligent wearables in enhancing interactive and immersive experiences in performing arts.

Keywords: Wearable technology, Machine learning, Motion recognition, RGB lighting, Smart dance

**Correspondence: ravi@sjp.ac.lk*

1. Introduction

Wearable technology has emerged as one of the fastest growing fields in HCI (Human Computer Interaction) by (Carroll, 2009), offering new ways to extend human capabilities and enhance everyday experiences (Pantelopoulos & Bourbakis, 2010), (Fortino et al., 2014). From smartwatches that track health parameters to augmented reality headsets that redefine entertainment, wearables are now an integral part of modern life. In particular, motion responsive wearable devices have shown potential for creating immersive and engaging experiences in performance arts, sports, rehabilitation, and gaming (Liu et al., 2018), (Benbasat & Paradiso, 2003), (Gao et al., 2014).

In performance arts, interactivity and immersion are increasingly valued, with technology becoming a co-creative partner rather than just a background tool (Xu et al., 2016), (Dobrian & Bevilacqua, 2003). Traditional performances often rely on lighting systems that are manually synchronized with music or choreography.

While visually effective, such setups lack flexibility, personalization, and real-time responsiveness. A dancer's body movements, however, carry rich expressive information that can be captured using wearable sensors and mapped to dynamic lighting or sound effects. This coupling of motion recognition and actuation has been shown to enhance audience engagement and deepen the connection between performer and performance (Bevilacqua, Schnell & Rasamimanana, 2011), (Jensenius, 2007). Recent advances in embedded machine learning, often referred to as "TinyML," have enabled complex algorithms to run directly on resource-constrained devices such as microcontrollers (Warden & Situnayake, 2019). These developments make it feasible to integrate real-time classification and decision-making within wearable platforms, eliminating the need for external computing resources. Devices like the ESP32-S3, combined with inertial measurement units (IMUs) such as the MPU-6050, provide a powerful yet low-cost solution for motion recognition tasks (Shah & Patel, 2019), (Gálvez et al., 2019). Meanwhile, RGB LED systems offer lightweight, portable, and programmable means of delivering rich visual feedback, making them particularly suited for interactive performance environments (Tan, Lin & Chen, 2018), (Akten, 2009).

Building on these developments, this paper presents the Smart Dance Shoes, an interactive system that integrates an ESP32-S3 microcontroller with an MPU-6050 motion sensor and RGB LED strips. A machine learning model deployed on the embedded device classifies dance movements in real time, triggering corresponding lighting effects. This enables performers to achieve synchronization between movement and stage visuals without requiring extensive manual setup or external operators. The objectives of this research can be understood through three interconnected directions. First, the system aims to develop a real-time motion recognition framework that seamlessly synchronizes RGB lighting effects with human movement using the ESP32-S3 microcontroller and the MPU-6050 inertial measurement unit. By capturing acceleration and gyroscopic data from the dancer's body in real time, the system is designed to recognize subtle shifts in movement and translate them into dynamic lighting responses. This objective emphasizes responsiveness and reliability, ensuring that the visual effects remain tightly coupled with the physical performance without noticeable delay or misalignment.
Second, the project focuses on implementing machine learning algorithms capable of classifying a variety of dance movements and mapping them to appropriate lighting patterns. Rather than relying on pre-programmed or static effects, the system leverages embedded models to interpret motion sequences and make context-aware decisions directly on the wearable platform. This not only enhances the adaptability of the system across different performers and choreographies but also allows for a higher level of personalization and creative expression. By enabling real-time classification on a resource-constrained device, the research contributes to the growing field of TinyML and its applications in interactive art and performance.

Finally, the research seeks to ensure that the overall system remains lightweight, efficient, and practical for real-world use in performance environments. This involves designing a compact and unobtrusive hardware setup that can be easily integrated into wearable form factors, while also

*Correspondence: ravi@sjp.ac.lk

optimizing the embedded algorithms for low latency and power consumption. Beyond technical efficiency, this objective also highlights usability: the system should require minimal setup, reduce reliance on external operators, and allow performers to intuitively engage with the technology without disrupting the natural flow of their performance.

Together, these objectives establish a foundation for wearable technologies that not only extend technical capabilities but also enrich the expressive possibilities of performance art, offering a pathway toward more immersive, interactive, and human-centered experiences.

2. Literature Review

The literature on wearable and interactive technologies reveals a rich and diverse set of research directions, yet it also highlights the need for integrated, real-time solutions tailored to performance contexts. Broadly, prior work can be categorized into three major areas: wearable devices, machine learning for motion recognition, and interactive performance systems.

The first body of research focuses on wearable devices, which have evolved from early health-monitoring sensors to multifunctional platforms capable of supporting entertainment, rehabilitation, and artistic applications. This work emphasizes portability, ergonomics, and continuous data collection, laying the foundation for real-time human–computer interaction.

A second strand of literature examines machine learning methods for motion recognition. This area has seen rapid growth with the availability of inertial measurement units (IMUs) and embedded hardware. Early rule-based approaches have gradually been replaced by data driven learning models capable of classifying complex gestures and movement patterns. Recent advancements in TinyML further extend this capability by enabling efficient on-device inference, reducing dependency on external computing resources.

The third research domain explores interactive performance systems, where technology is not merely a supporting tool but an active collaborator in artistic expression. Studies in this area demonstrate how motion, sound, and visuals can be coupled in real time to enhance audience immersion.

However, most implementations either rely on external computing setups or lack full integration of motion recognition with dynamic visual feedback, leaving room for systems that are lightweight, embedded, and performer centered.

Considered collectively, these three strands of literature provide a foundation for this research while also exposing its novelty. This work highlights how wearable technologies have matured, how machine learning has advanced the recognition of movement, and how interactive performance systems have demonstrated artistic potential. Yet, the intersection of these areas where portable, embedded, and intelligent wearables drive real-time visual interaction—remains underdeveloped.
This study positions itself within this intersection, aiming to bridge the gap by proposing a fully integrated system that combines motion recognition with synchronized RGB lighting in the context of live dance performance.

2.1 *Wearable Devices*

Wearables are increasingly used in healthcare (Preece et al., 2019), fitness (Khan et al., 2010), and entertainment applications (Zheng et al., 2018). Accelerometer and gyroscope sensors are widely adopted due to their ability to capture fine-grained motion data. Early systems primarily targeted fitness monitoring, focusing on step counting or gait analysis (Bao & Intille, 2004), (Ravi et al., 2005). Later, more complex tasks such as gesture recognition and fall detection were explored (Anguita et al., 2012), (Lara & Labrador, 2013).

Microcontrollers such as Arduino and ESP32 have been central to wearable device prototyping. The ESP32-S3 stands out because of its dual-core architecture, low power consumption, and integrated wireless capabilities (Espressif Systems, 2021). Similarly, the MPU-6050 sensor has been widely

*Correspondence: ravi@sjp.ac.lk

adopted in research projects due to its combination of accelerometer and gyroscope in a compact form factor (InvenSense, 2013).

2.2 ***Machine Learning for Motion Recognition***

Machine learning has played a pivotal role in analyzing motion data. Traditional methods, such as decision trees and support vector machines, were initially applied to activity recognition (Anguita et al., 2012). However, these methods struggled with complex sequential patterns. Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) models improved classification accuracy by capturing temporal dependencies (Ordoñez & Roggen, 2016), (Hammerla et al., 2016), (Zeng et al., 2014). More recently, attention-based models and Transformers have been proposed for handling sequential motion data more effectively (Vaswani et al., 2017), (Chen et al., 2021).

2.3 ***Interactive Performance Systems***

The integration of wearable devices into performance art has gained momentum. Benbasat & Paradiso, 2003 explored expressive footwear that responded to dance movements. Choi et al., 2016 developed smart wearable systems for interactive dance performances. While these systems showcased creativity, they often lacked portability or required external computing devices. Jiang & Yin, 2015 proposed deep learning-based recognition but relied on smartphones for processing. Our work addresses this gap by embedding ML models directly into the shoe hardware, enabling real-time and portable performance augmentation.

Although wearable motion recognition has seen significant advances in recent years, several limitations remain that restrict its broader adoption in performance contexts. One of the key challenges lies in portability. Many existing systems depend heavily on external devices, such as smartphones or computers, to handle computation and classification tasks (Jiang & Yin, 2015). While this approach provides access to more processing power, it undermines real-time usability on stage, where performers require systems that are compact, self-contained, and capable of operating without tethered devices. Another critical issue concerns responsiveness. For applications in dance and other live performances, synchronization between body movements, music, and lighting must occur with minimal latency. However, many current solutions introduce noticeable delays between movement detection and system response, which diminishes the immersive quality of performances and reduces the sense of natural interaction (Choi et al., 2016). Addressing this issue requires systems that can process sensor data and trigger outputs in real time, even under the constraints of embedded hardware.

Equally important is the question of usability. Existing solutions are often not tailored to the specific needs of dancers and performers, who require wearable systems that are lightweight, unobtrusive, and intuitive to use. Complex setup procedures, bulky hardware, or non-ergonomic designs can interfere with a performer's freedom of movement and distract from the creative experience. Thus, designing systems that prioritize comfort, ease of use, and seamless integration into performance practices remains an underexplored area. Finally, there is a clear lack of integration in current approaches. While research has demonstrated wearable motion recognition and, separately, dynamic lighting systems, relatively few projects have successfully combined these elements into a single embedded platform. The absence of integrated solutions means that performers often rely on fragmented setups, which limit the expressiveness and creative possibilities of interactive stage technologies (Benbasat & Paradiso, 2003).

This research seeks to address these gaps by proposing a compact, embedded system that unifies machine learning–based motion recognition with synchronized RGB lighting. By emphasizing portability, real-time responsiveness, performer-centered usability, and holistic integration, the project aims to advance the state of wearable technologies in performance arts and contribute to more immersive and expressive interactive experiences.

*Correspondence: ravi@sjp.ac.lk

3. Research Methodology

The methodology for this research was designed to establish a complete pipeline for the development and deployment of a wearable motion-responsive system. It begins with the design of the hardware platform, where an ESP32-S3 microcontroller is integrated with an MPU-6050 motion sensor and RGB LED strips to form a lightweight and portable solution. Once the hardware foundation is in place, motion data was systematically collected from professional dance performances, ensuring a balanced dataset that captured both subtle and dynamic movements across multiple dance styles.

The data collected was then preprocessed and used to train a machine learning model, with an emphasis on optimizing both accuracy and efficiency for embedded deployment. A convolutional neural network (CNN) was trained using the Edge Impulse platform, which provided a streamlined environment for data handling, feature extraction, model training, and quantization. To ensure real-time usability, the trained model was deployed directly on the ESP32-S3 microcontroller, allowing movement recognition to be performed locally without reliance on external servers.

Finally, the system was fully integrated by mapping the classified dance movements to the corresponding lighting effects on the RGB LED strips. This integration also ensured real-time synchronization between performance and stage visuals but also enabled customization through a Bluetooth-enabled web interface, allowing performers to adjust lighting patterns according to their artistic preferences. Overall, the methodology combines hardware design, data collection, machine learning, and system integration into a unified workflow aimed at delivering practical, responsive, and performer-centered wearable technology.

3.1 *Hardware Design*

The hardware configuration for the proposed system comprised an ESP32-S3 microcontroller (Espressif Systems, 2021), an MPU-6050 motion sensor (InvenSense, 2013), and programmable RGB LED strips, all powered by a rechargeable Lithium Polymer (LiPo) battery. Each component was selected carefully to ensure efficiency, compactness, and scalability of the system.

The ESP32-S3 was chosen as the primary processing unit due to its superior performance compared to traditional microcontrollers such as Arduino. It features a dual-core Xtensa LX7 processor with a higher clock speed, integrated WiFi, and Bluetooth Low Energy (BLE) support, which significantly enhances wireless communication and connectivity. This allows seamless integration with external devices and cloud-based platforms, making it suitable for real-time applications. Additionally, the ESP32-S3 supports hardware acceleration for machine learning (ML) tasks, which is advantageous for deploying lightweight motion recognition models directly on the device without relying on external servers.

The MPU-6050 motion sensor was selected because of its compact design and capability to capture six degrees of freedom (6-DoF) motion data, combining a 3-axis accelerometer and a 3-axis gyroscope into a single module. This provides precise measurement of both linear acceleration and angular velocity, which are critical for accurate movement tracking and classification.
Its compatibility with the I2C communication protocol also simplifies the integration process with the ESP32-S3, reducing wiring complexity and ensuring efficient data transfer. For the visual output, addressable RGB LED strips were employed, offering high brightness, programmability, and dynamic color rendering. These LEDs were used to translate the detected motion patterns into synchronized lighting effects, providing an interactive and immersive user experience.

A LiPo battery was utilized as the primary power source due to its high energy density, lightweight design, and rechargeability, making the system portable and suitable for long duration use. Voltage regulation modules were included to ensure stable and safe power delivery to the microcontroller, sensor, and LED strips.

*Correspondence: ravi@sjp.ac.lk

Overall, this hardware design provides a balance between computational capability, motionsensing accuracy, energy efficiency, and portability, forming a robust foundation for the realtime motion recognition and interactive lighting system.

3.2 *Data Collection*

Motion data were systematically collected from a professional dancer performing a range of movements to capture the dynamic variations of human activity. The performed movements included spins, jumps, and step sequences of varying intensities and speeds. These activities were deliberately chosen to represent both rapid and gradual motion transitions, ensuring that the dataset captured a broad spectrum of real-world movements. The data collection process focused on capturing six degrees of freedom (6-DoF) information from the MPU-6050 sensor, consisting of three-axis accelerometer readings (linear acceleration) and three-axis gyroscope readings (angular velocity). To ensure robustness and prevent bias, the dataset was balanced across all activity classes. Each class contained a comparable number of samples, thereby avoiding the over-representation of any single activity. This balanced approach improves the performance and generalizability of the trained models when applied to unseen data (Bao & Intille, 2004). Data were collected at a uniform sampling frequency to maintain temporal consistency, and preprocessing steps such as normalization and noise filtering were applied to enhance signal quality.

The collected motion data were labeled according to predefined activity categories that correspond to distinct dance or movement actions. Table 1 presents the activity labels and their descriptions.

Table 1: Activity Labels for Collected Motion Data

Label	**Description**
hinchipinchi	A signature movement style used by the professional dancer, characterized by rhythmic and expressive gestures.
jump	Vertical motion involving a sudden lift of the body from the ground, with variations in intensity and landing style.
left-right	Side-to-side stepping or body movement, capturing lateral shifts in position.
round-r	Rotational spin to the right side, capturing angular velocity and circular motion patterns.
stand	Static posture representing minimal movement, primarily used as a baseline reference class.
up-down	Vertical oscillatory movement of the body, involving repeated raising and lowering actions.
walk	Forward locomotion with alternating left and right steps, representing a natural walking gait.

By systematically labeling and balancing the dataset across these categories, the collected data formed a reliable foundation for training and evaluating machine learning models, ensuring robust classification across varying motion types and intensities.

3.3 *Data Preprocessing and Model Training*

The raw motion data collected from the MPU-6050 sensor were initially preprocessed to ensure quality and consistency prior to model training. Preprocessing involved two key steps: normalization and segmentation. Normalization was applied to scale the accelerometer and gyroscope readings into a consistent range, thereby reducing the influence of sensor drift and ensuring uniform feature distribution across samples. Segmentation was performed by dividing the continuous time-series data

*Correspondence: ravi@sjp.ac.lk

into fixed-length windows, each representing a single instance of an activity. This sliding window approach enabled the preservation of temporal dependencies while avoiding overlap between different motion classes.

Following preprocessing, feature extraction and classification were conducted using the Edge Impulse Studio platform (Edge Impulse, 2023). Edge Impulse provides an integrated environment for embedded machine learning development, allowing efficient handling of sensor data pipelines, automated feature generation, and real-time model performance evaluation. Both time-domain features (e.g., mean, variance, signal magnitude area) and frequency-domain features (e.g., spectral energy, dominant frequency components) were considered to capture the dynamics of human motion.

A Convolutional Neural Network (CNN) architecture was employed for classification due to its ability to automatically learn spatial and temporal patterns in sensor signals. CNNs have shown superior performance in modeling complex human activity data compared to traditional machine learning techniques (Hammerla et al., 2016). The network was trained using labeled motion segments, with optimization techniques such as learning rate scheduling, dropout regularization, and data augmentation applied to prevent overfitting and enhance generalization.

Furthermore, the model was optimized specifically for deployment on resource-constrained embedded systems. Quantization and pruning techniques were applied to reduce memory footprint and computational overhead while maintaining classification accuracy. The final trained model was validated on held-out test data to evaluate performance metrics, including accuracy, precision, recall, and F1-score, ensuring its robustness under real-time operational conditions (Chen et al., 2021).This workflow established an efficient end-to-end pipeline, from raw motion data to an optimized CNN model, enabling accurate activity recognition on the ESP32-S3 microcontroller in real-world scenarios.

3.4 *System Integration*

The trained Convolutional Neural Network (CNN) model was exported from the training environment and deployed on the ESP32-S3 microcontroller using the Edge Impulse SDK. Movements were mapped to specific lighting effects, such as color transitions for spins and flashing lights for rapid footwork. A Bluetooth-enabled web application provided customization options for performers. This process involved model quantization to ensure compatibility with the limited memory and computational capacity of the microcontroller. The integration workflow began with preprocessing the motion data from the MPU-6050 sensor to match the input format expected by the trained CNN model. Once deployed, the model performed real-time inference directly on the device, reducing the dependency on external servers and ensuring low-latency responses. Movements recognized by the CNN were mapped to specific lighting effects. For instance, spins were represented by smooth color transitions to convey fluid motion, while rapid footwork triggered flashing or strobe-like patterns to accentuate high-energy movements. This mapping was designed not only for aesthetic appeal but also to provide visual feedback to the performer and audience in real time.

To enhance system usability, a Bluetooth-enabled web application was developed. This application allowed performers and choreographers to customize lighting effects without requiring hardware-level programming. Users could adjust parameters such as color palettes, brightness, transition speed, and effect-trigger mapping, thereby enabling creative freedom and adaptability to different performances or dance styles. The web application’s interface was designed with simplicity in mind, ensuring that performers without technical expertise could personalize the system to their preferences.

4. Results

The experimental procedure was divided into multiple stages to ensure a comprehensive evaluation of the system’s functionality, accuracy, and real-world applicability. Testing was carried out both in controlled laboratory conditions and in live performance environments.

*Correspondence: ravi@sjp.ac.lk

Unit testing was performed to validate the functionality of each individual hardware and software component. The ESP32-S3 microcontroller was tested for reliable power management and program execution. The MPU-6050 sensor was evaluated for its ability to accurately capture accelerometer and gyroscope data across different movement intensities and orientations. LED strips were examined for uniform brightness, smooth transition rendering, and response to different command signals. At the software level, test scripts were used to confirm data preprocessing steps, sensor calibration, and CNN model inference outputs. These unit tests ensured that each module could function independently before integration into the larger system.

After successful unit validation, integration testing focused on the interactions between the microcontroller, motion sensor, and lighting system. The primary objective was to verify seamless communication and synchronization across modules. The sensor data pipeline was monitored to ensure that raw motion signals were correctly preprocessed and fed into the CNN without data loss or corruption. Latency was measured as the time between the physical movement and the activation of the corresponding lighting effect. Repeated trials indicated system delays remained within the acceptable threshold of less than 150 milliseconds, thereby ensuring that the light effects appeared instantaneous to both performers and audience members. Additionally, stress testing was conducted under continuous operation to ensure system stability during long-duration performances. Performance testing was conducted to assess the accuracy and robustness of the motion recognition system. A labeled test dataset comprising various dance movements was collected and used to evaluate the deployed CNN model. Metrics such as accuracy, precision, recall, confusion matrices, and F1-scores were computed to quantify recognition performance (Anguita et al., 2012).

The results highlighted the strengths of the model in identifying large, distinct movements such as spins, while also indicating areas for improvement in differentiating subtle or overlapping motions. Cross-validation was employed to confirm the generalizability of the trained model across multiple performers with varying body types and movement styles. To further assess real-world usability, performance tests were conducted under different lighting and stage conditions, ensuring that external environmental factors did not interfere with system responsiveness.

Following laboratory validation, real-world testing was conducted during live dance performances to evaluate the system's practical effectiveness. This phase measured not only technical performance but also human-centered factors such as responsiveness, usability, and dancer satisfaction. The system successfully provided synchronized lighting feedback, enhancing the visual impact of performances and improving audience engagement. Feedback from dancers indicated that the system did not interfere with movement or comfort, as the hardware components were lightweight and unobtrusive. Optimization techniques were applied to address challenges observed during live testing. Model compression methods, such as pruning and quantization, were employed to reduce inference time without sacrificing recognition accuracy (Han et al., 2016). Additionally, power management strategies such as dynamic frequency scaling and sensor sleep modes were implemented to extend operational time during extended rehearsals and performances. Overall, the real-world trials demonstrated that the system was reliable, adaptable, and capable of delivering meaningful enhancements to live dance experiences.

4.1 *Model Overview*

The neural network model was developed and trained using the Edge Impulse Studio platform. It was implemented as a fully connected neural network (dense layers) designed for the classification of motion-related activities, including HINCHIPINCHI, Jump, Left-Right, Round-R, Stand, Up-Down, and Walk. The input to the model consisted of 63 extracted features derived from motion sensor data, while the output layer contained seven classes corresponding to the defined activities.

*Correspondence: ravi@sjp.ac.lk

4.2 ***Neural Network Architecture***

The model architecture is summarized as follows:

• **Input Layer:** 63 features, processed from raw sensor data through feature extraction (e.g., spectral analysis).

• **Hidden Layers:**

– Dense layer with 20 neurons, activation function f(x) = max(0, x) (ReLU).

– Dense layer with 10 neurons (ReLU).

• **Output Layer:** 7 neurons with a softmax activation function:

$$\hat{y_i} = \frac{e^{z_i}}{\sum_{j=1}^{7} e^{z_j}}, \quad i = 1,2,\dots,7$$

where z_i is the input to the i-th output neuron and y_i represents the predicted probability for class i.

4.3 ***Training Configuration***

The training was performed on a CPU included 40 Training cycles (epochs) with a Learning rate of 0.0005. It uses the Categorical cross-entropy as the Loss function and Adam (without the learned optimizer option) as the Optimizer.

The categorical cross-entropy loss is defined as:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log\left(\widehat{y_{i,c}}\right)$$

where N is the number of samples, C is the number of classes, $y_{i,c}$ is the ground truth label, and $y_{i,c}$ is the predicted probability.

4.4 ***Training Results***

The trained model achieved an overall validation accuracy of 98.5% with a cross-entropy loss of 0.06. The confusion matrix (Table 2) illustrates class-wise performance. Most classes were correctly classified with very high accuracy, with Left-Right and Walk achieving perfect classification (100%), while Up-Down and HINCHIPINCHI exceeded 97%. Minor misclassifications were observed between Round-R and HINCHIPINCHI (4.1%) and between Up-Down and Jump (2.1%).

4.5 ***Evaluation Metrics***

The performance of the model was further evaluated using precision, recall, F1-score, and the area under the ROC curve (AUC). The definitions of the metrics are given below: (TP – True Positive, TN - True Negative, FP - False Positive, FN - False Negative)

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP},$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

*Correspondence: ravi@sjp.ac.lk

Table 2: Confusion Matrix of the Validation set (in %)

Class	HIN	Jump	Left-Right	Round-R	Stand	Up-Down	Walk
HINCHIPINCHI	97.6	0.0	0.0	2.4	0.0	0.0	0.0
Jump	0.0	97.3	0.0	0.0	0.0	2.7	0.0
Left-Right	0.0	0.0	100	0.0	0.0	0.0	0.0
Round-R	4.1	0.0	0.0	95.9	0.0	0.0	0.0
Stand	0.0	0.0	0.0	0.0	100	0.0	0.0
Up-Down	0.0	2.1	0.0	0.0	0.0	97.9	0.0
Walk	0.0	0.0	0.0	0.0	0.0	0.0	100

Table 3 presents the results achieved. The AUC reached 1.0, indicating excellent class separability. Weighted averages for precision, recall, and F1-score all reached 0.98, confirming that the model performs consistently across all activity classes.

Table 3: Confusion Matrix of the Validation set (in %)

Metric	Value
Accuracy	98.5%
Loss	0.06
Area under ROC curve (AUC)	1.00
Weighted Precision	0.98
Weighted Recall	0.98
Weighted F1 Score	0.98

4.6 Strengths of the Model

The trained model exhibited robust performance across multiple evaluation metrics. It achieved high classification accuracy for all motion categories, indicating reliable recognition of diverse movement patterns. The model's low validation loss demonstrated strong convergence during training, ensuring stability and generalization. Moreover, excellent class separability was observed, with an area under the curve (AUC) of 1.0, highlighting its ability to distinguish between different motion types effectively. Performance was well-balanced across all classes, minimizing any bias toward dominant categories and ensuring consistent results across the dataset. These results validate the robustness of the proposed approach and its suitability for deployment in real-time embedded systems.

5. Discussion and Conclusion

The experimental results demonstrate that the proposed system can effectively classify seven distinct dance related movements using a lightweight fully connected neural network deployed on the ESP32-S3. The accuracy of 92% indicates that the selected features and model architecture are well suited for real-time motion recognition on resource-constrained hardware. Compared to existing wearable-based motion recognition studies, which often rely on smartphones or cloud computing, the presented system emphasizes on-device computation, thereby reducing latency and ensuring independence from external networks.

A key observation is that most misclassifications occurred between visually and kinematically similar movements, such as Round-R and HINCHIPINCHI. This suggests that the extracted statistical features may not fully capture fine-grained rotational patterns, indicating potential benefits of more advanced temporal models such as LSTMs or Transformers. Additionally, variations in movement

*Correspondence: ravi@sjp.ac.lk

speed and intensity introduced subtle inconsistencies in the sensor readings, which may have affected classification robustness. The hardware configuration, particularly the integration of ESP32-S3 with the MPU-6050, proved to be reliable for continuous data acquisition and processing. However, the size and placement of the module could affect user comfort during extended dance sessions. While the RGB LED visualization enriched the performance experience, synchronization between music and motion remains an area for improvement. Overall, the discussion highlights both the strengths—lightweight architecture, real-time inference, and system portability—and the limitations—dataset diversity, feature sensitivity, and hardware ergonomics—of the current implementation. This research presented the design and implementation of a wearable, sensor-based system for real-time dance motion recognition and visualization. By combining an ESP32-S3 microcontroller, an MPU-6050 motion sensor, and a compact RGB LED interface, the system successfully classified seven different movements with high accuracy while operating entirely on-device. The work demonstrates the feasibility of deploying neural networks trained in Edge Impulse Studio on embedded hardware for interactive performance applications.

The contributions of this study are threefold: (i) the development of a compact and portable hardware module for motion sensing, (ii) the creation of a balanced dataset of professional dance movements and its processing pipeline, and (iii) the deployment of a neural network that achieves reliable classification under real time constraints. The system advances the intersection of wearable technology, performing arts, and embedded machine learning, showcasing a novel approach to augmenting dance performances. Future directions include expanding the dataset with multiple dancers and genres, adopting advanced sequential models for improved recognition accuracy, and refining hardware ergonomics for long-duration comfort. Furthermore, integrating audio-driven synchronization and multi-shoe communication could enable enhanced interactive performances. The outcomes of this study set the foundation for broader applications in sports training, rehabilitation, and entertainment technologies.

6. Future Work

Future research can explore several avenues to further enhance the motion recognition and light-synchronization system. One promising direction is the adoption of advanced models such as LSTM or Transformer architecture, which can improve sequential motion recognition (Vaswani et al., 2017). Expanding the dataset to include multiple dancers and diverse dance genres would enhance model generalization and robustness. Hardware redesign could focus on creating more compact and ergonomic modules to improve user comfort during performances. Additionally, implementing multi-shoe synchronization would enable coordinated effects for group performances. Integrating music beat recognition could allow the system to synchronize lighting effects with audio, creating more immersive experiences (Essid et al., 2009). Finally, leveraging cloud-based updates and edge computing could support continuous learning and model improvements over time (Satyanarayanan, 2017).

References

Akten, M. 2009. Visualisation of sound and music for live performance. In: *Proceedings of the Conference on New Interfaces for Musical Expression*. Pittsburgh, USA, pp. 26–29.

Anguita, D. et al. 2012. A public domain dataset for human activity recognition using smartphones. In: *Proceedings of ESANN*. Bruges, Belgium.

Bao, L. & Intille, S.S. 2004. Activity recognition from user-annotated acceleration data. *Pervasive Computing*, pp. 1–17.

Benbasat, A.Y. & Paradiso, J.A. 2003. Design and implementation of expressive footwear for dance and entertainment. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, New York.

*Correspondence: ravi@sjp.ac.lk

Bevilacqua, F., Schnell, N. & Rasamimanana, N. 2011. Gestural control of sound synthesis. In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. Oslo, Norway.

Carroll, J.M. 2009. Human–computer interaction (HCI). In: *Encyclopedia of Human-Computer Interaction*. The Interaction Design Foundation, Aarhus, pp. 21–62.

Chen, Y. et al. 2021. Deep learning for sensor-based human activity recognition: A survey. *Pattern Recognition Letters*, 147, pp. 1–11.

Choi, J. et al. 2016. Smart wearable systems for interactive dance performance. *Multimedia Tools and Applications*, 75(22), pp. 14773–14788.

Dobrian, C. & Bevilacqua, F. 2003. Using motion capture to control interactive music systems. In: *Proceedings of the Conference on New Interfaces for Musical Expression*. Montreal, Canada.

Edge Impulse. 2023. *Edge Impulse Documentation*. Available at: https://docs.edgeimpulse.com (Accessed 2023).

Espressif Systems. 2021. *ESP32-S3 Datasheet*. Espressif Systems Inc.

Essid, S. et al. 2009. Music information retrieval for music and motion synchronization. *IEEE Transactions on Audio, Speech, and Language Processing*, 17(4), pp. 580–589.

Fortino, G., Parisi, D., Pirrone, V. & di Fatta, G. 2014. BodyCloud: A SaaS approach for community body sensor networks. *Future Generation Computer Systems*, 35, pp. 62–79.

Gao, W. et al. 2014. Wearable physiological monitoring systems: Current progress and future challenges. *IEEE Reviews in Biomedical Engineering*, 7, pp. 147–156.

Gálvez, J.C., Medina, C., Koutsou, A.D., López-Nava, I.H. & De la Cuerda, R.C. 2019. Comparing IMU-based motion capture systems for gait analysis. *Sensors*, 19(3), 552.

Hammerla, N.Y. et al. 2016. Deep, convolutional, and recurrent models for human activity recognition using wearables. In: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1533–1540.

Han, S. et al. 2016. Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding. In: *Proceedings of the International Conference on Learning Representations (ICLR)*.

InvenSense. 2013. *MPU-6050 6-axis MotionTracking Device Datasheet*. InvenSense Inc.

Jensenius, A.R. 2007. *Action–sound: Developing methods and tools to study music-related body movement*. PhD thesis, University of Oslo, Norway.

Jiang, W. & Yin, Z. 2015. Human activity recognition using wearable sensors by deep convolutional neural networks. In: *Proceedings of ACM Multimedia*. ACM, New York.

Khan, A. et al. 2010. A review on human activity recognition using wearable sensors. *IEEE Sensors Journal*, 10(6), pp. 1167–1179.

Lara, O.D. & Labrador, M.A. 2013. A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys & Tutorials*, 15(3), pp. 1192–1209.

Liu, Y. et al. 2018. Wearable computing for motion analysis: Current status and future challenges. *ACM Computing Surveys*, 51(6), pp. 1–36.

Ordoñez, F. & Roggen, D. 2016. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1), pp. 1–25.

Pantelopoulos, A. & Bourbakis, N.G. 2010. A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 40(1), pp. 1–12.

Preece, J. et al. 2019. Wearable technologies for human activity recognition: A review. *IEEE Sensors Journal*, 19(19), pp. 9047–9060.

Ravi, N. et al. 2005. Activity recognition from accelerometer data. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI Press, pp. 1541–1546.

Satyanarayanan, M. 2017. The emergence of edge computing. *Computer*, 50(1), pp. 30–39.

*Correspondence: ravi@sjp.ac.lk

Shah, S. & Patel, R. 2019. Performance evaluation of ESP32 module for IoT applications. *International Journal of Advanced Research in Computer Science*, 10(5), pp. 7–10.

Tan, J., Lin, C-H. & Chen, B. 2018. LightDance: A system for wearable lighting design and control in dance performance. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, New York, pp. 1–12.

Turaga, P. et al. 2008. Machine recognition of human activities: A survey. *IEEE Transactions on Circuits and Systems*, 18(11), pp. 1473–1488.

Vaswani, A. et al. 2017. Attention is all you need. In: *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 5998–6008.

Warden, P. & Situnayake, D. 2019. *TinyML: Machine learning with TensorFlow Lite on Arduino and ultra-low-power microcontrollers*. O'Reilly Media, Sebastopol.

Xu, W., Woodruff, A., Wu, M. et al. 2016. Dancing in data: Exploring movement-based interaction in digital performance. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, New York, pp. 2760–2771.

Yao, S. et al. 2017. DeepSense: A unified deep learning framework for time-series mobile sensing data processing. In: *Proceedings of the World Wide Web Conference (WWW)*. ACM.

Zeng, M. et al. 2014. Convolutional neural networks for human activity recognition using mobile sensors. In: *Proceedings of Mobile Computing*. ACM.

Zheng, Y. et al. 2018. Wearable devices for human activity recognition: A survey. *Journal of Biomedical Informatics*, 81, pp. 44–56.

*Correspondence: ravi@sjp.ac.lk