



OPEN ACCESS

EDITED BY

Cristian Romagnoli,
Università telematica San Raffaele, Italy

REVIEWED BY

Farzan Majeed Noori,
University of Oslo, Norway
Alwin Poulouse,
Indian Institute of Science Education and
Research, India

*CORRESPONDENCE

Maha Abdelhaq,
✉ msabdelhaq@gmail.com
Ahmad Jalal,
✉ ahmadjalal@mail.au.edu.pk
Hui Liu,
✉ hui.liu@uni-bremen.de

RECEIVED 26 November 2023

ACCEPTED 26 January 2024

PUBLISHED 21 February 2024

CITATION

Khan D, Alonazi M, Abdelhaq M, Al Mudawi N,
Algarni A, Jalal A and Liu H (2024), Robust
human locomotion and localization activity
recognition over multisensory.
Front. Physiol. 15:1344887.
doi: 10.3389/fphys.2024.1344887

COPYRIGHT

© 2024 Khan, Alonazi, Abdelhaq, Al Mudawi,
Algarni, Jalal and Liu. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Robust human locomotion and localization activity recognition over multisensory

Danyal Khan¹, Mohammed Alonazi², Maha Abdelhaq^{3*},
Naif Al Mudawi⁴, Asaad Algarni⁵, Ahmad Jalal^{1*} and Hui Liu^{6*}

¹Department of Computer Science, Air University, Islamabad, Pakistan, ²Department of Information Systems, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia, ³Department of Information Technology, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia, ⁴Department of Computer Science, College of Computer Science and Information System, Najran University, Najran, Saudi Arabia, ⁵Department of Computer Sciences, Faculty of Computing and Information Technology, Northern Border University, Rafha, Saudi Arabia, ⁶Cognitive Systems Lab, University of Bremen, Bremen, Germany

Human activity recognition (HAR) plays a pivotal role in various domains, including healthcare, sports, robotics, and security. With the growing popularity of wearable devices, particularly Inertial Measurement Units (IMUs) and Ambient sensors, researchers and engineers have sought to take advantage of these advances to accurately and efficiently detect and classify human activities. This research paper presents an advanced methodology for human activity and localization recognition, utilizing smartphone IMU, Ambient, GPS, and Audio sensor data from two public benchmark datasets: the Opportunity dataset and the Extrasensory dataset. The Opportunity dataset was collected from 12 subjects participating in a range of daily activities, and it captures data from various body-worn and object-associated sensors. The Extrasensory dataset features data from 60 participants, including thousands of data samples from smartphone and smartwatch sensors, labeled with a wide array of human activities. Our study incorporates novel feature extraction techniques for signal, GPS, and audio sensor data. Specifically, for localization, GPS, audio, and IMU sensors are utilized, while IMU and Ambient sensors are employed for locomotion activity recognition. To achieve accurate activity classification, state-of-the-art deep learning techniques, such as convolutional neural networks (CNN) and long short-term memory (LSTM), have been explored. For indoor/outdoor activities, CNNs are applied, while LSTMs are utilized for locomotion activity recognition. The proposed system has been evaluated using the k-fold cross-validation method, achieving accuracy rates of 97% and 89% for locomotion activity over the Opportunity and Extrasensory datasets, respectively, and 96% for indoor/outdoor activity over the Extrasensory dataset. These results highlight the efficiency of our methodology in accurately detecting various human activities, showing its potential for real-world applications. Moreover, the research paper introduces a hybrid system that combines machine learning and deep learning features, enhancing activity recognition performance by leveraging the strengths of both approaches.

KEYWORDS

human activity recognition, smart IMU, GPS, deep learning, convolutional neural network, long short-term memory human activity recognition, GPS sensor, convolutional neural network (CNN)

1 Introduction

The advancement of sensing technologies (Jiang and He, 2020; Zheng et al., 2023a), notably has catalyzed progress in human activity recognition (HAR). These sensors, pivotal in health (Sobhan et al., 2021; Hussain et al., 2022; Zheng et al., 2023b) and safety monitoring (Reddy et al., 2016; Mao et al., 2022a) in smart environments (Guo et al., 2022; Jiawei et al., 2022; Liu et al., 2023g), aim to be both accurate and nonintrusive. Wearable sensors (Saboor et al., 2020; Bhelkar and Shedje, 2016; Perez and Zeadally, 2021) with their potential to capture granular movement data, have introduced new possibilities in HAR (Liu et al., 2023a). However, their challenges (Saboor et al., 2020; Liu and Schultz, 2019; Bhelkar and Shedje, 2016) concerning battery life and user acceptance underscore the importance of a balanced approach. Tools like infrared sensors (Perez and Zeadally, 2021; Liu et al., 2022a) and recent 3D data acquisition systems (Yu et al., 2023; Bruno et al., 2015) such as Microsoft Kinect (Zhao et al., 2023; Liu et al., 2022b; Shen et al., 2022) are emerging as robust alternatives, offering precision without compromising user privacy. As HAR technologies evolve, integrating wearables and non-intrusive sensors, the field is poised to offer deeper insights into human behavior (Zhang et al., 2012b; Puangragsa et al., 2022) enhancing security, health monitoring, and infrastructure management (Kamarudin, et al., 2014; Hu et al., 2022; Hassan and Gutub, 2022).

This research paper discusses the application of a Smart inertial measurement unit (IMU), global positioning system (GPS), and audio sensors, along with ambient sensors, for human activity recognition (Zheng et al., 2022; Meng et al., 2022). The combination of these sensors offers a comprehensive approach to capturing diverse aspects of human movements and actions. IMUs, which consist of accelerometers, gyroscopes, and magnetometers, provide precise motion and orientation data. In conjunction with Ambient Sensors that capture contextual information, these sensors provide insight into human activities in real-world scenarios. To achieve accurate and context-aware activity recognition, advanced signal processing techniques are used to extract relevant features from the data these sensors collect. Novel feature extraction methods have been designed for signal (Hartmann et al., 2022; Hartmann et al., 2023), GPS, and audio sensor data, enriching the system's ability to discern patterns and characteristics associated with different activities. To effectively process the information from GPS, Audio, and IMU Sensors, a Yeo-Johnson power transformation is applied for optimization. Simultaneously, IMU and Ambient features are optimized and harnessed for the identification of locomotion activities, showcasing the versatility of the proposed approach. Given the complexity and diversity of human activities, state-of-the-art deep learning techniques are employed to develop a robust and accurate HAR system (Qi et al., 2022; Wang et al., 2022; Yan et al., 2023; Ronald et al., 2021; Poulouse et al., 2022; Poulouse et al., 2019a). Convolutional neural networks (CNN) (Zhang et al., 2023; Wen et al., 2023a; Gangothri et al., 2023; Leone et al., 2022) are used for recognizing indoor/outdoor activities, while long short-term memory (LSTM) (Yao et al., 2023; Zheng, Y. et al., 2022) networks are chosen for locomotion activity recognition (Hu et al., 2023; Liu and Schultz, 2018; Liu et al., 2022c). The integration of CNN and LSTM allows the system to leverage spatial and temporal dependencies, thus

enhancing overall recognition performance. The proposed HAR system (Zhou and Zhang, 2022; Xue and Liu, 2021; Zhao et al., 2022) is evaluated using the Opportunity and Extrasensory datasets, which are well-established benchmarks in the field of localization activity recognition (Zhu et al., 2023; Qu et al., 2023a; Qu et al., 2023b; Liu et al., 2023a). The results underscore the effectiveness of the approach, achieving remarkable accuracies of 97% and 89% for locomotion activity over the Opportunity and Extrasensory datasets, respectively, and 96% for localization activity over the Extra-sensory dataset. These findings attest to the potential of Smart IMU, GPS, Audio, and Ambient Sensors in precisely identifying and classifying a range of human activities (Gioanni et al., 2016). Beyond exploring deep learning techniques, this research paper introduces a hybrid system (She et al., 2022; Liang et al., 2018; Liu et al., 2022d; Vrskova et al., 2023; Surek et al., 2023) that blends machine learning and deep learning features. By capitalizing on the strengths of both paradigms, the hybrid system further sharpens activity recognition, signaling a promising avenue for future research and development. The primary findings and contributions of this study are outlined below:

- Development of robust denoising techniques tailored for signal and Audio sensor data, enhancing activity recognition accuracy.
- Extracting novel features for detecting human localization information.
- Development of a hybrid system that combines machine learning and deep learning features to further improve activity recognition performance.
- Furthermore, a comprehensive analysis was performed on well-known benchmark datasets, which feature diverse human actions and advanced sensors.

The subsequent sections of this paper are organized as follows:

Section 2 presents a comprehensive literature review of existing methods in the field of human activity recognition. In **Section 3**, the proposed system is thoroughly discussed. The experimental setup and the results obtained from the conducted experiments are outlined in **Section 4**. In **Section 5**, we discuss the system's performance, limitations, and future directions. Finally, in **Section 6** conclusions drawn from the research are presented.

2 Literature review

Various methods exist for recognizing human activity, with some researchers utilizing RGB cameras, others employing wearable sensors, and some leveraging multimodal sensor approaches.

2.1 Visual sensor-based human locomotion recognition

A new technique for pulling out details about joints and skeletons from images was introduced in a study (Batchuluun et al., 2021). The method started by changing an original thermal image, which had 1 channel, into an image with 3 channels. This change was done to combine the images in a way that would help get

better results when pulling out information. The study used a tool called a generative adversarial network (GAN) to help extract details about joints and skeletons. Furthermore, the study tried to recognize different human actions using the information pulled out about joints and skeletons. The recognition of human actions was done by using two tools together: a CNN and LSTM. When they tested their method using their own collected data and also open data, the study found that their method worked well compared to other top methods. However, the system could not detect images that have low spatial textual information, due to which the system causes low performance. The study (Yin et al., 2021) developed a model to detect different human actions in a real-time healthcare environment. The authors utilized a multichannel LSTM. This system, built to detect actions through three-dimensional skeleton data, incorporated a unique loss function to enhance its accuracy. They used two benchmark datasets: one is NTU RGB + D and the second is TST fall detection datasets. However, the system has limitations in achieving flawless skeleton data due to a frame-level error detection approach and struggles with identifying the roots of issues related to dimensionality, which in turn impacts the overall accuracy of the system. In another study (Chen et al., 2023), the authors concentrated on recognizing actions through different video frames. Residual CNN and a second spatial attention module are utilized for the recognition of actions. The proposed system does not have integrated optical flow maps, which adversely impacts the performance of the system.

2.2 Human locomotion recognition via wearable technology

In the work conducted by Mutegeki and Han, (2020), an integrative deep learning architecture for activity recognition was introduced, utilizing a CNN-LSTM model. This approach aimed to enhance predictive accuracy for human activities derived from raw data while simultaneously reducing model complexity and negating the necessity for intricate feature engineering. The proposed CNN-LSTM network was devised to be deep in both spatial and temporal dimensions. The model manifested a 99% accuracy rate on the iSPL dataset (an internal dataset) and 92% on the publicly available UCI HAR dataset. However, the findings indicate a decline in performance when addressing complex actions, such as atomic-level activities. Additionally, as the model complexity amplified, the SoftMax loss also escalated, suggesting that the concurrent use of CNN and LSTM layers did not enhance the outcomes. Jaramillo et al. (2022) utilized a technique called Quaternion filtration by using single sensor data. In the next step, different segmentation techniques have been used to segment the data. Subsequently, features are extracted. Finally, for the classification of activities, the LSTM classifier has been utilized. We identified that the system is more computationally expensive. Hu et al. (2023), presents a system for human activity recognition is presented using IMU sensors, and the data was collected from Wearable devices. Different techniques are utilized to preprocess the data, including moving averages, sliding overlap windows, and data segmentation. For recognition of activities, five different classifiers are used including CNN, recurrent neural network, LSTM, bidirectional LSTM (BiLSTM), and gate recurrent unit. Due to a huge number of epochs, the

proposed system is very expensive in terms of time complexity. Recently, the hidden Markov model (HMM) has entered the field of vision of researchers (Liu and Schultz, 2018). Its inherently logical modeling capability of time series endows human activity recognition with a certain degree of interpretability.

2.3 Human locomotion recognition through multisensor systems

The study (Hanif et al., 2022) presents a multimodal locomotion system, utilizing the Opportunity++ and HWU-USP datasets for their study. The data was subjected to various pre-processing techniques; for image-based data, the skeleton was initially extracted, while for inertial sensors, the noise was removed followed by segmentation. Various features, including Pearson correlation, linear prediction, and cepstral coefficients, were extracted. The classification of locomotion was performed using a recursive neural network. Nonetheless, the confidence levels obtained for each extracted skeleton body point do not meet the desired standards, particularly for both ankle points. In another multimodal system, proposed (Nafea et al., 2022) data was collected using smart devices. For preprocessing the raw sensor data, different methods such as filtration, windowing, and segmentation were utilized. Multiple features were extracted, including time-based, statistical, frequency-based, and rotational features. Furthermore, various machine learning classifiers have been explored to classify both complex and basic activities, such as k nearest neighbour (k -NN), neural networks, and Naïve Bayes. However, these learning approaches tend to be susceptible to errors and often deliver suboptimal accuracy in the context of human locomotion recognition (HLR), resulting in performance that does not achieve satisfactory outcomes. In another study (Ma et al., 2023) a system was proposed to remotely monitor people, utilizing multimodal sensors to monitor activities. CNN and gated recurrent unit (GRU) were explored for recognizing different human activity patterns. Nonetheless, the suggested approach did not yield strong results due to significant losses in both the training and validation sets (M. Ronald. et al., 2021). use the iSPLInception model a deep learning architecture based on the synergistic combination of Inception modules and ResNet strategies. By refining these components, the model achieves a significant balance between depth and computational efficiency, essential for real-time processing. The researchers focused on enhancing predictive accuracy for HAR while ensuring the model's feasibility on devices with constrained computational resources. Through extensive benchmarking across diverse datasets, the iSPLInception demonstrates robustness in classifying a variety of activities. A comparison with other deep learning models such as LSTMs and CNNs confirmed its superior performance, making a notable contribution to the HAR domain. The methodology outlined by the authors provides a scalable solution that paves the way for future research in activity recognition using wearable sensor data. Poulose. et al. (2022) proposes an innovative approach to human activity recognition (HAR) using a system referred to as HIT (Human Image Threshing) machine. This system employs a smartphone camera to capture activity videos, which are then processed using a mask region-based convolutional neural network (R-CNN) for human body detection.

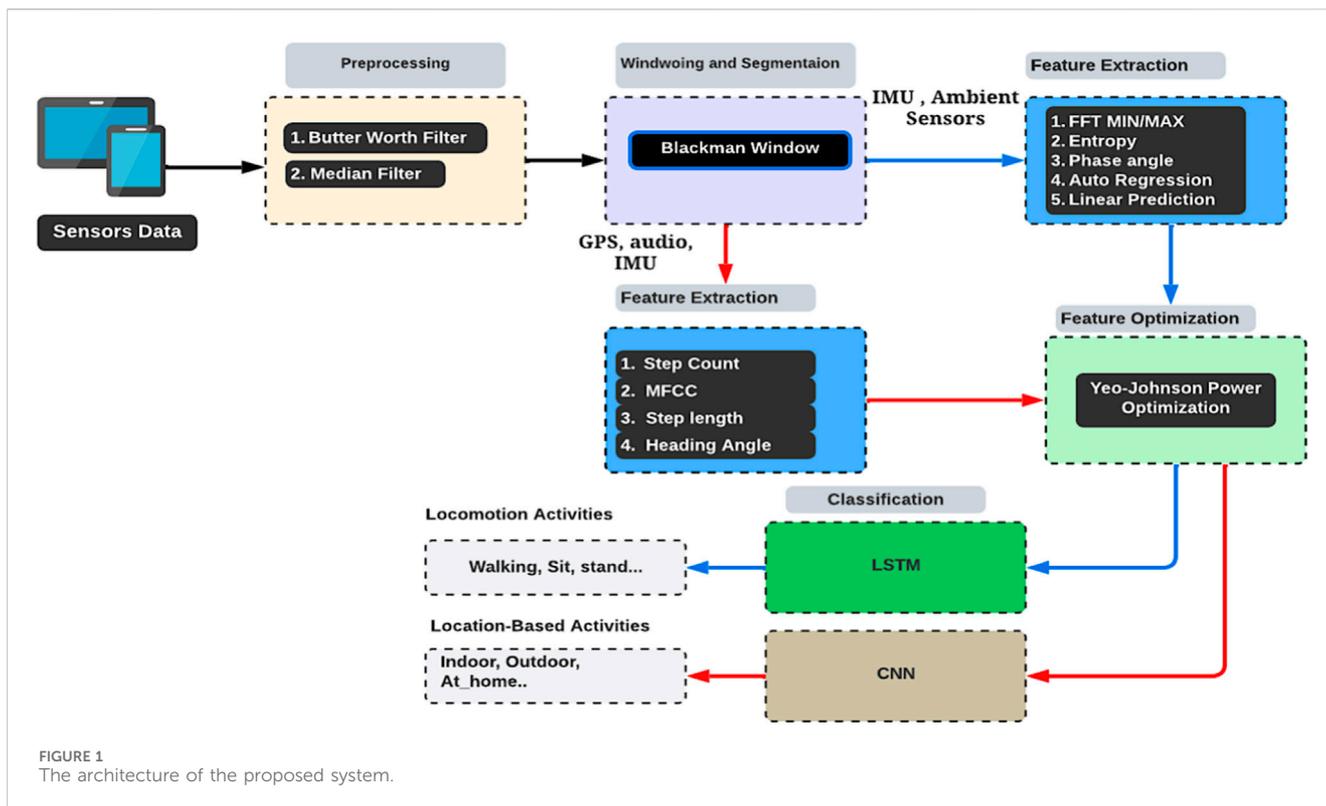


FIGURE 1 The architecture of the proposed system.

The process also includes a facial image thresholding machine (FIT) for image cropping and resizing. The core of the HIT machine’s methodology is its ability to clean and preprocess data, followed by deep feature extraction and model building for activity classification. The system is tested with various deep learning models like VGG, Inception, ResNet, and EfficientNet, achieving remarkable accuracy in classifying activities such as sitting, standing, walking, push-ups, dancing, sit-ups, running, and jumping. This approach significantly outperforms traditional sensor-based HAR systems, demonstrating the effectiveness of vision-based activity recognition using deep learning models.

3 Materials and methods

3.1 System methodology

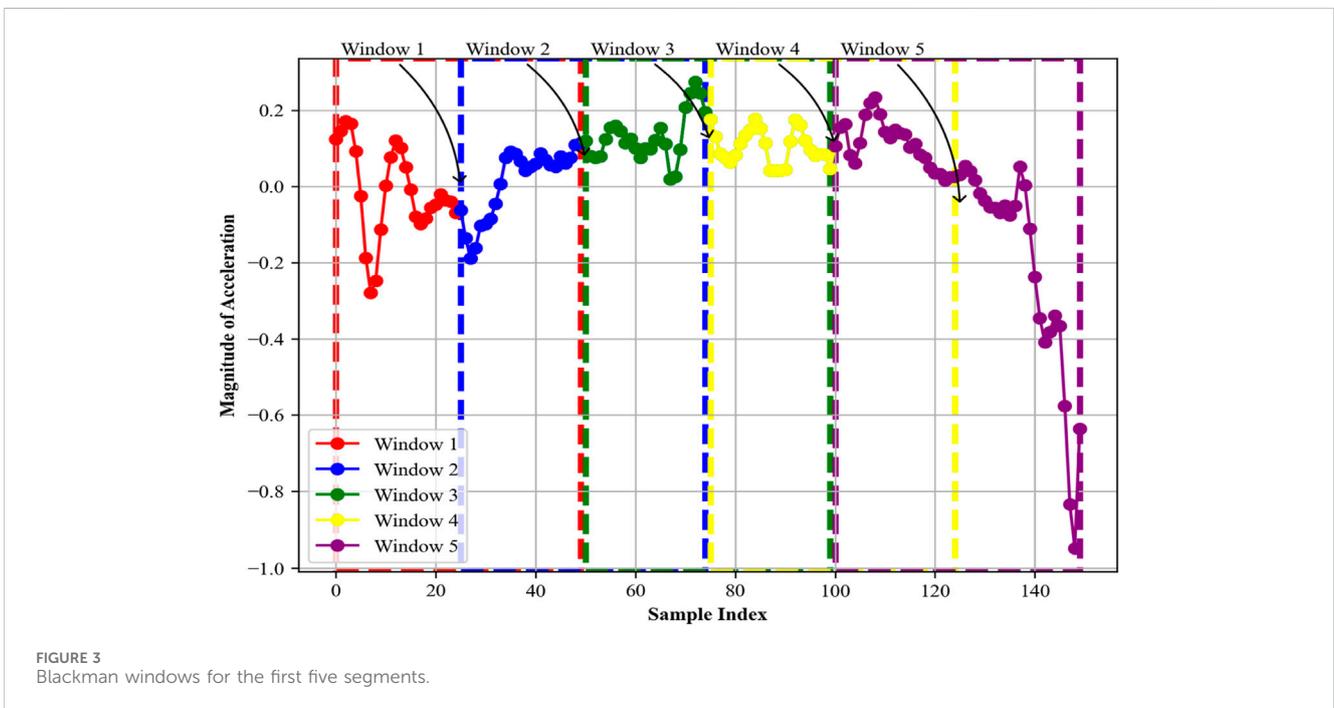
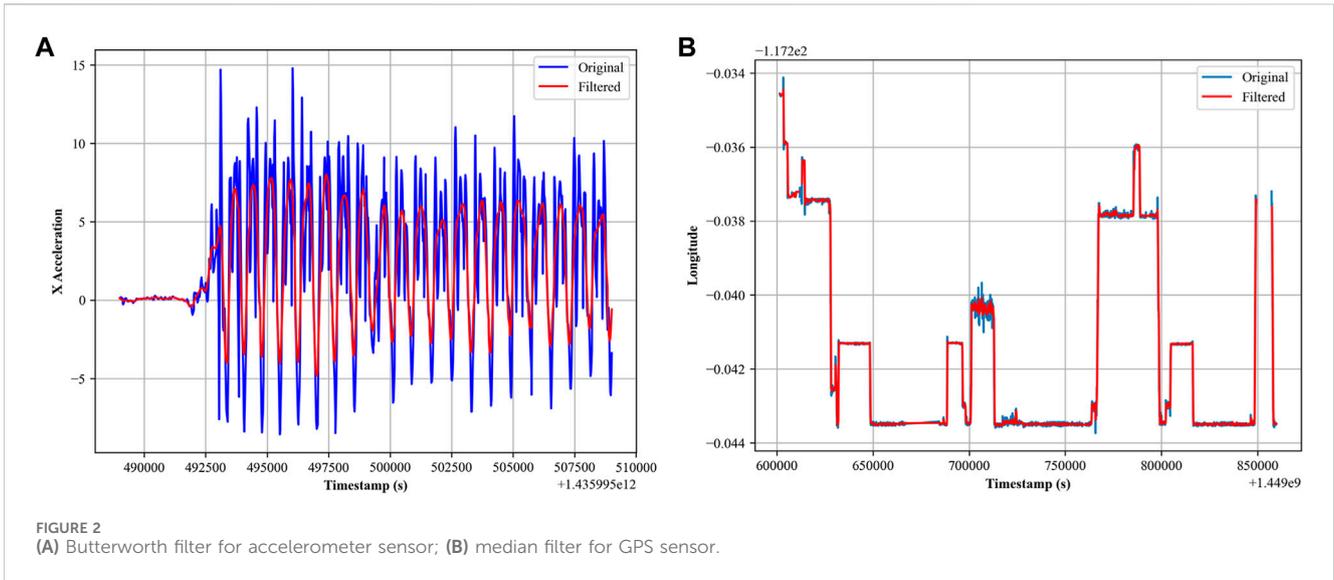
In this work, we follow a multistep approach to process and analyze data from different types of sensors (Ahmad, 2022; Zhang et al., 2022a; Latha et al., 2022). Initially, we address the issue of noise in the raw signal and use distinct filters for each sensor type. Specifically, we use a Butterworth filter for the IMU and Ambient sensors and a median filter for GPS and audio data. Next, to efficiently handle large sequence data, we utilize windowing and segmentation techniques. This allows us to break down the data into smaller segments, facilitating more effective processing. In the third step, we focus on extracting advanced features from different types of sensors. These features include statistical, phase angle, autoregressive modelling, and linear prediction features. Additionally, for the IMU and audio data, we extract various features such as step count, step length, and

Mel-frequency cepstral coefficients (MFCCs). All of these features are further optimized and combined using the Yeo-Johnson power transformation. Optimized GPS, IMU, and audio sensor features are then sent to a CNN for localization activity analysis, while the IMU and ambient sensor features are directed to an LSTM network for locomotion activity recognition (Jaiwei et al., 2022; Zhang et al., 2022b; Rustam et al., 2020). The proposed system’s architecture is visually represented in Figure 1.

3.2 Noise removal

The data was collected from raw sensors that include noise. Noise is unwanted data or irrelevant data due to many reasons during data collection. So, to handle the noise, we used 2 types of filters because of different types of sensor data. To remove noise from the IMU and ambient sensors, we used a third-order Butterworth filter (Bae et al., 2020; Liu et al., 2023f; Cömert et al., 2018; Sulistyaningsih et al., 2018) (i.e., $n = 3$) was used. The choice of this order strikes a balance between achieving a reasonable roll-off and minimizing signal distortion. The critical frequency f_c , was set to 10% of the Nyquist frequency, represented as $W_n = 0.1$. This ensures that frequencies beyond 10% of the Nyquist frequency are attenuated, providing a smooth output while preserving the essential characteristics of the input signal. The magnitude response of a Butterworth filter in the frequency domain is given by

$$|H(f)| = \frac{1}{\sqrt{1 + \left(\frac{f}{f_c}\right)^{2n}}}$$



$H(f)$ represents the magnitude of the filter response at frequency f . f_c is the critical frequency, which is the frequency at which the filter's response is $\frac{1}{\sqrt{2}}$ of its maximum (or passband) response, n denotes the order of the filter, dictating the steepness of the roll-off. Higher order results in a sharper transition between the passband and the stopband. Similarly, for the GPS and microphone sensors, we used a median filter (Altun and Barshan, 2010). To apply the median filter, we used a kernel of size 3, which essentially means that for each data point, the filter considered it and one neighboring data point on each side. The median value of these three points then replaced the original data point. Mathematically, for each component, the median of the current value and its neighbors was computed, producing the

filtered data. Mathematically, the filtered acceleration for each component can be expressed as

$$S_x = \text{median}(x[i - k], x[i - k + 1], \dots, x[i + k])$$

$$S_y = \text{median}(y[i - k], y[i - k + 1], \dots, y[i + k])$$

$$S_z = \text{median}(z[i - k], z[i - k + 1], \dots, z[i + k])$$

where S_x , S_y , and S_z are the signal.

Post the filtering process, to synthesize a unified representation of the signal component, we then employed the Pythagorean theorem:

$$\text{magnitude}_{\text{filtered}} = \sqrt{(S_x)^2 + (S_y)^2 + (S_z)^2}$$

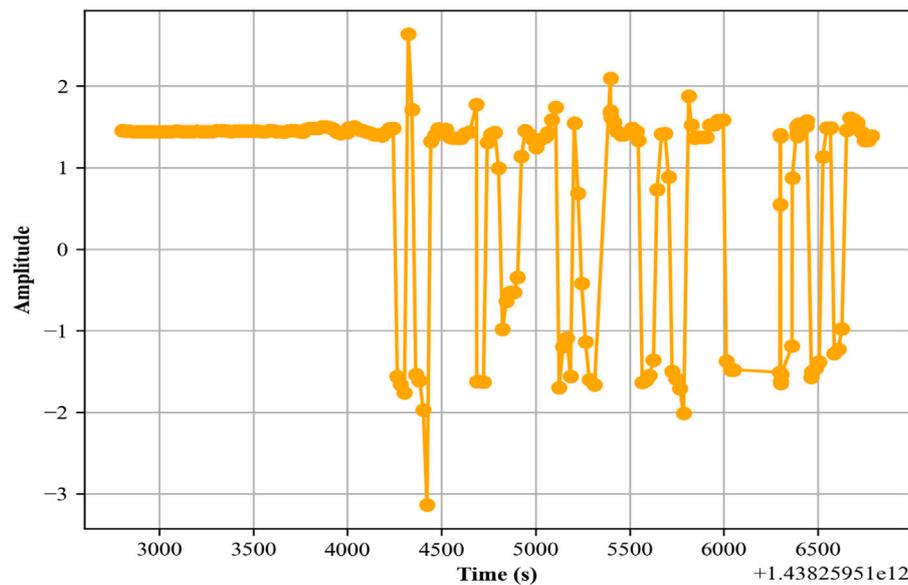


FIGURE 4
Phase angles were calculated from the accelerometer data over the Opportunity dataset.

However, it is important to note that the GPS sensor has less noise compared to other sensors, which can be seen in [Figure 2B](#).

3.3 Windowing and segmentation

To window and segment large sequence data for efficient processing, we turned to the Blackman window ([Kwapisz et al., 2011](#)) windows technique to modulate the signal. Windows plays an important role in this phase. By applying a Blackman window to the signals during segmentation, we smooth the abrupt beginnings and endings of segments, thereby reducing spectral leakage, a phenomenon where energy from one frequency leaks into another, potentially obscuring important features. This ensures that the Fourier transform of the windowed signal provides a more faithful representation of its frequency content. Furthermore, in human activity recognition, activities can span varying durations and might be best represented by capturing their essence within specific windows ([Poulose et al., 2019b](#)). The Blackman window, with its inherent properties, ensures that each segmented frame is appropriately weighted, reducing discontinuities at the boundaries. This results in improved frequency domain representations, enabling more accurate feature extraction, and consequently more precise activity recognition. Mathematics of the Blackman Window is

$$W_n = 0.42 - 0.5 \cos\left(\frac{2\pi n}{N-1}\right) + 0.8 \cos\left(\frac{4\pi n}{N-1}\right)$$

where $W(n)$ is the window function. N is the total number of points in the window, and n ranges from 0 to $N-1$. For our specific implementation, we used a 50-sample window to represent 5 s ([He and Jin, 2008](#); [Hao, 2021](#); [Liu et al., 2021](#); [Hatamikia et al., 2014](#)) of

activity with 25% overlap. After generating the Blackman window values based on the formula, we multiplied each point in our data segments with its corresponding Blackman window value. To bring clarity to our process, we visualized the results through distinct line plots, with each of the five windows represented in a unique color in [Figure 3](#), and [Algorithm 1](#) shows the working of the Blackman windowing technique.

```

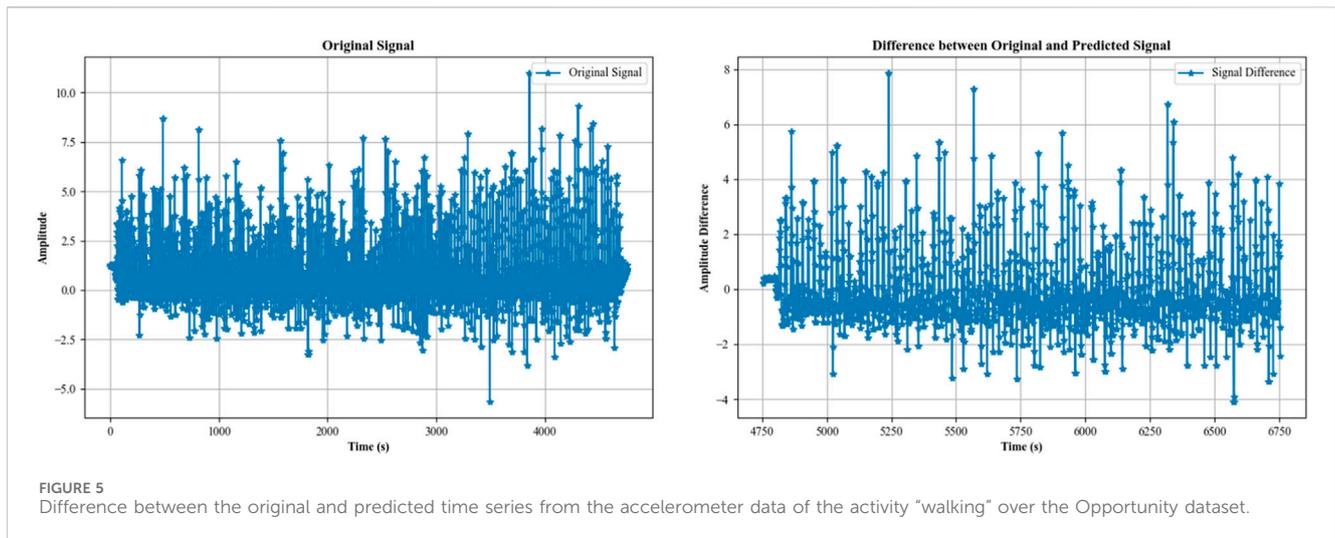
Input: Time-series data array  $D$ 
        Window size  $N$ 
Output: List  $F$  containing feature vectors for
        each segment
Method: Create a Blackman window  $W$  of size  $N$ 
        Initialize an empty list  $F$  to store feature vectors
        for each segment
For  $i = 0$  to length of  $D - N$  with a step size of  $N$ :
    Extract a segment  $S$  from  $D [ i : i + N ]$ 
    Multiply  $S$  with  $W$  element-wise to get  $S_{windowed}$ 
    Compute features  $f$  from  $S_{windowed}$  Append  $f$  to  $F$ .
return list  $F$  containing feature vectors for
        each segment

```

Algorithm 1. Blackman Windowing and Segmentation

3.4 Feature extraction for locomotion activity

Another essential step in this research is the extraction of features, ensuring that the model effectively recognizes data patterns. We derived unique features for various sensor types. For both IMU and Ambient sensors, we extracted features such as phase angle, linear predictions, FFT Max/Min, Shannon entropy, skewness, kurtosis, and autoregressive analysis.



3.4.1 Phase angle

Phase angles hold significance in signal analysis, particularly in the field of human activity recognition. Phase angles capture the temporal alignment and synchronization of cyclic movements, helping in the extraction of valuable information from complex signals (Zhang, 2012; Liu et al., 2022a). These angles provide insight into the relative timing of movements in different dimensions, enabling the identification of specific activities and patterns. Mathematically, the phase angle between two signals A and B can be calculated using the arctangent function, which takes into account the ratio of their spectral components in the frequency domain. For accelerometer data, the phase angle between the x and y components (ϕ_{xy}), x and z components (ϕ_{xz}), and y and z components (ϕ_{yz}), can be computed as

$$\phi_{xy} = \arctan\left(\frac{\text{FFT}(Ay)}{\text{FFT}(Ax)}\right)$$

$$\phi_{xz} = \arctan\left(\frac{\text{FFT}(Az)}{\text{FFT}(Ax)}\right)$$

$$\phi_{yz} = \arctan\left(\frac{\text{FFT}(Az)}{\text{FFT}(Ay)}\right)$$

where FFTA_x , FFTA_y , and FFTA_z represent the fast Fourier transforms of the x , y , and z components of the sensor data, respectively. Figure 4 exemplifies the phase angles calculated in XY graphically.

3.4.2 Auto regressive model

Autoregressive (AR) modeling (Li et al., 2020; Xu et al., 2016; Gil-Martin et al., 2020a) is a powerful technique in signal analysis, particularly for human activity recognition. It involves predicting a data point in a time series based on previous data points and capturing temporal dependencies and patterns. This is especially useful in recognizing periodic or rhythmic activities, as the model captures the repeating patterns inherent in activities like walking, running, or cycling. By comparing the predicted and actual values, deviations can be detected, helping to identify anomalies or changes in activity patterns (Bennasar et al., 2022; Liu et al., 2021). For

example, variations in step lengths, gait irregularities, or sudden changes in motion can be indicative of different activities or health conditions (Wen et al., 2023). We used an AR model to model the time series data for the walking activity opportunity dataset. In an AR model, the value at time t is predicted as a linear combination of the p previous values. For an AR model of order p , the value X_t at time t is modeled as

$$X_t = c + \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t$$

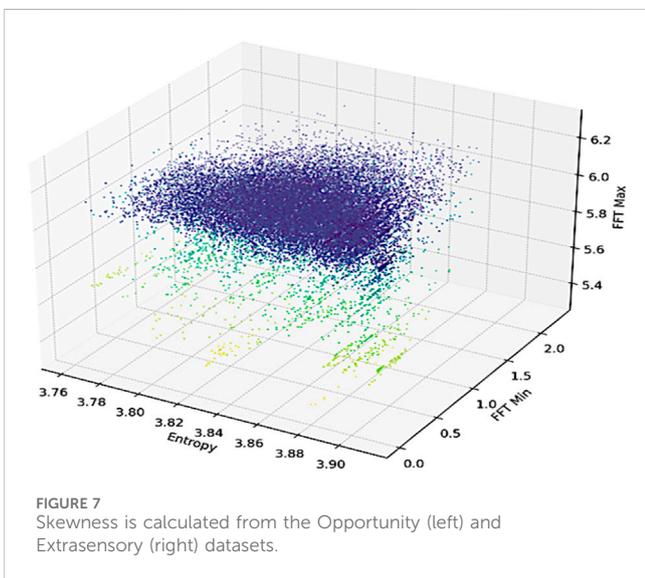
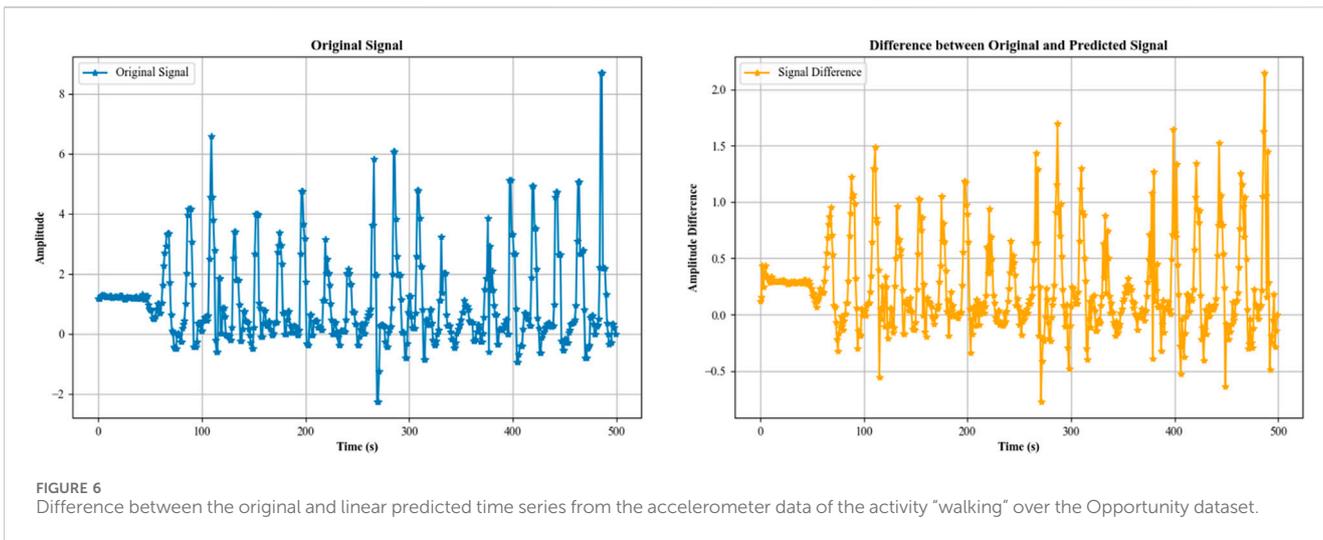
where: X_t , is the value at time t , c is a constant, ϕ_i are the parameters of the model and ε_t is the white noise. After fitting the AR model to the data, we used the model to make predictions for future points. The prediction step is based on the AR model equation. For each future point X_t , the predicted value is calculated as

$$\widehat{X}_t = c + \sum_{i=1}^p \phi_i X_{t-i}$$

The difference between the actual AR model and the prediction step is that the actual AR model includes a noise term ε_t while the prediction step does not. The noise term represents uncertainty and random fluctuations that cannot be predicted by the AR model. Thus, it is not included in the prediction step. Finally, we plotted the difference between original and predicted time series data in Figure 5.

3.4.3 Linear prediction for signal

After calculating the autoregression, we then calculated the linear prediction. Linear prediction is a powerful method employed in signal analysis for uncovering meaningful patterns and trends in data. This approach is particularly useful when dealing with time-series data, such as movement patterns. This concept finds the relationship between current and previous data points; linear prediction enables us to forecast how the signal might evolve over time. This predictive capability enables the identification of distinctive movement patterns and characteristics that are indicative of specific activities. We preprocess the accelerometer data to ensure its quality and reliability. We then apply linear prediction techniques



to model the temporal patterns of each activity. This involves training linear models that predict future data points based on a history of previous observations. The optimization of model coefficients is carried out to minimize prediction errors, resulting in predictive models that capture the underlying motion dynamics. For a time, series x_t , linear prediction estimates x_t as a weighted sum of p previous values $x_{t-1}, x_{t-2}, \dots, a_p x_{t-p}$:

$$x_t = c + a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_p x_{t-p}$$

where c is a constant term and a_1, a_2, \dots, a_p are the coefficients of the linear model. These coefficients are determined through optimization methods that minimize the prediction error. Figure 6 portrays the linear prediction for walking activity.

3.4.4 Fast fourier transformation (FFT) min/max and entropy

We first calculated FFT (Javeed et al., 2021; Li et al., 2018), a mathematical algorithm that unveils the frequency-domain

representation of a time-domain signal. By applying FFT to sensor data, it becomes possible to uncover the underlying frequency components inherent in various human activities. Peaks and patterns in the resulting frequency spectrum can be associated with specific motions or actions, offering crucial insights into the dynamic nature of movements (Liu, 2021). We calculated the minimum and maximum components from the FFT spectrum. It can be calculated as

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt$$

where $X(f)$, is the frequency-domain representation, $x(t)$, is the time-domain signal, f is the frequency, and j is the imaginary unit. Furthermore, we extracted the Shannon entropy feature. In the context of signal analysis for human activity recognition, Shannon entropy (Khairy, 2022) can reveal the complexity and diversity of frequency components in the signal. Higher entropy values suggest a broader range of frequencies and more varied motion patterns. Mathematically, it can be computed as

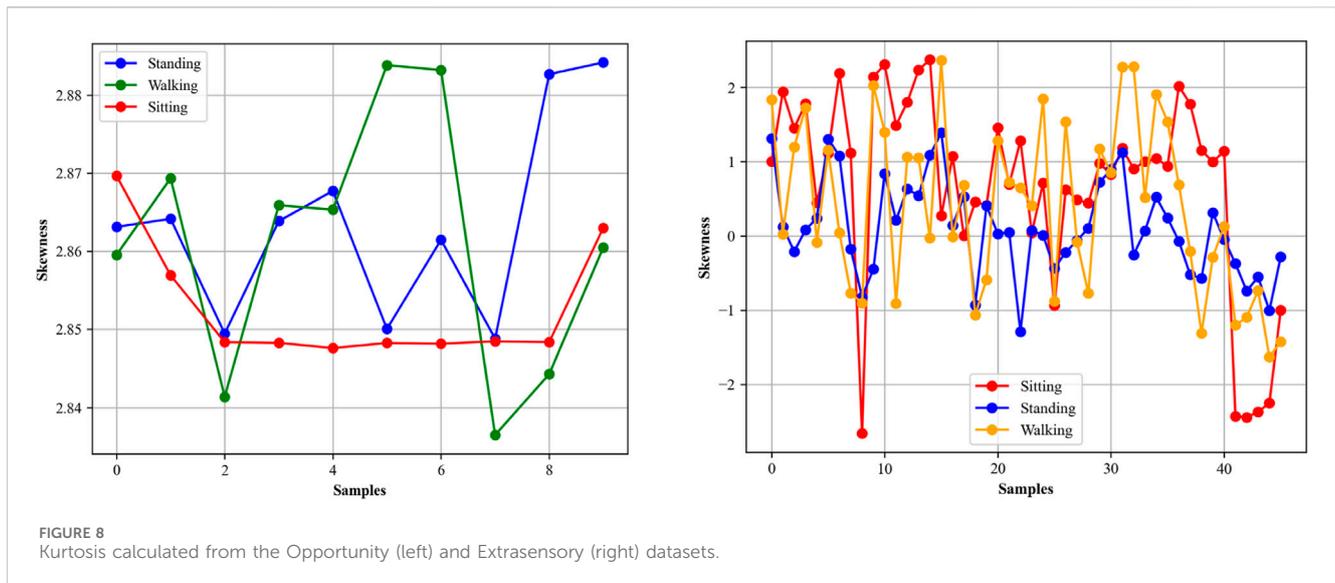
$$H = -\sum_{i=1}^N p(f_i) \log_2 p(f_i)$$

where N is the number of frequency bins, f_i is the i th frequency bin and $p(f_i)$ is the probability of occurrence of f_i in the signal’s frequency distribution. These features are demonstrated in Supplementary Figure S1.

3.4.5 Skewness

Skewness and kurtosis (Wang et al., 2021; Ramanujam et al., 2019; AlZubi et al., 2014) are statistical measures that describe the shape and characteristics of a distribution. Skewness quantifies the extent and direction of the skew in the data. A negative skew indicates that the left tail is longer, while a positive skew indicates a longer right tail. The mathematical equation for skewness (Yu et al., 2021; Zhang et al., 2021; Qi et al., 2022; Zheng et al., 2023c) is

$$\text{Skewness} = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{(N - 1).s^3}$$



where x_i are the individual sample points, \bar{x} is the sample mean, s is the standard deviation (Liu et al., 2021; J, X. et al., 2022; Mao et al., 2022b; Guo et al., 2022; Xu et al., 2022b), and N is the number of samples. Figure 7 shows skewness for different activities over both datasets.

3.4.6 kurtosis

Kurtosis (Lu et al., 2023; Liu et al., 2023b; Liu et al., 2023c; Liu et al., 2023d) on the other hand, measures the tailedness of the distribution. Higher kurtosis indicates a more extreme result, meaning that more of the variance is the result of infrequent extreme deviations, as opposed to frequent modestly sized deviations (Miao et al., 2023; Di et al., 2023; Ahmad et al., 2020; Liu et al., 2023e). The mathematical equation for kurtosis is

$$\text{Kurtosis} = \frac{\sum_{i=1}^N (x_i - \bar{x})^4}{(N - 1) \cdot s^4} - 3$$

Both skewness and kurtosis provide valuable information on the nature of variability in a set of numbers and are especially useful in the field of Human Activity Recognition (HAR) to distinguish between different types of activity. Skewness could provide clues about the symmetry of the user's motion, and kurtosis could indicate the extremity of the user's activities. We extracted kurtosis for different types of activities in Figure 8.

3.5 Feature extraction for location-based activity

For localization activity, we extracted separate features. These features include step count detection, step length calculation (Gu et al., 2017; Kang et al., 2018), and MFCCs. We describe each feature in detail one by one.

3.5.1 Step count detection

In indoor localization and activity recognition, the step count (Sevinç et al., 2020; Gu et al., 2019) emerges as an important metric with multifaceted applications. It serves as a fundamental parameter

for activity profiling, aiding in the differentiation of various human movements such as walking, running, or standing. Key features like step count and heading angle are integral to the development of robust and precise indoor localization systems, especially in environments where GPS signals are weak or entirely absent (Zhang and Jiang, 2021; Xu et al., 2023).

The step count was determined using accelerometer data (Pham et al., 2021) collected from the waist of the subject during walking activity. First, we combined the raw acceleration data along the x , y , and z axes to form a composite magnitude signal. This signal was then shifted to ensure that all values were positive. The mean of the shifted signal was calculated, and peaks that exceeded this mean were identified and counted as steps in Supplementary Figure S2. The magnitude of the acceleration A was calculated as

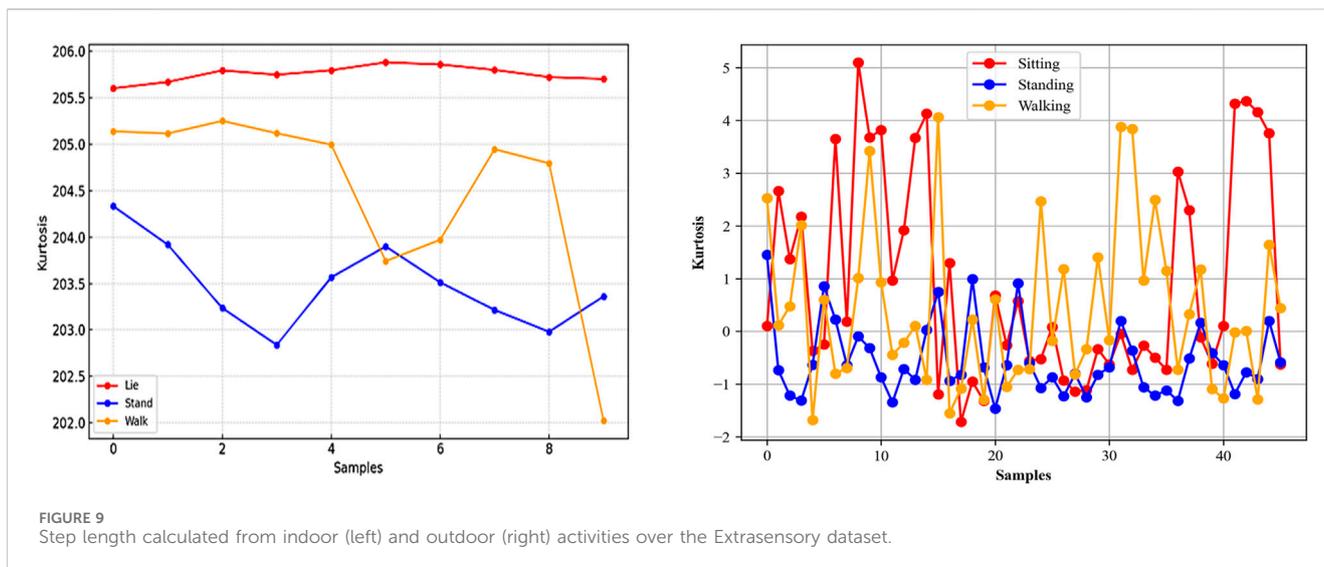
$$A = \sqrt{x^2 + y^2 + z^2}$$

3.5.2 Step length estimation

Step length, or stride length (Ahn and Yu, 2007; Hu et al., 2020) is important in the domain of indoor localization (Yoon and Kim, 2022) and human activity tracking. This metric essentially quantifies the distance covered in a single stride and serves as an essential parameter for accurately estimating an individual's location within a confined space. We utilized valley points in the position-time curve to estimate the stride length. Valley points in the position-time curve typically represent instances where the same foot hits the ground in successive strides. The curve itself is derived from double-integrating the acceleration data. This method is particularly beneficial in indoor settings, where GPS data may be unreliable or unavailable. Mathematically, the first step involves calculating the velocity V by integrating the acceleration A :

$$V = \int A dt$$

Following this, the position P is calculated by integrating the velocity:



$$P = \int V dt$$

We then identified valley points in this position-time curve. These points are local minima in the curve and represent the moments where a complete stride has occurred, that is, the same foot has hit the ground twice. The time difference between successive valley points is calculated as

$$t_{\text{valley}_n} - t_{\text{valley}_{n-1}}$$

This time difference Δt , when multiplied by a constant or average speed, gives the stride length for that particular step. In Figure 9, step lengths calculated for indoor and outdoor activities can be seen intuitively.

3.5.3 Heading angles

The calculation of the heading angle (Javeed and Jalal, 2023; Azmat et al., 2022) is an important component in indoor localization (Jiang et al., 2023), as it provides the orientation or directional information of an individual in relation to Earth's magnetic North. This orientation data is particularly for accurate path tracking and route reconstruction within indoor environments, where GPS signals are often weak. In our study, the heading angle, was calculated using magnetometer data, which measures the Earth's magnetic field components along the x , y , and z -axes. Given that the magnetometer can capture the Earth's magnetic field, it serves as a reliable sensor for determining orientation. To compute the heading angle, we employed the arctangent function on the y and x components of the magnetic field as per the following equation:

$$\theta = \arctan 2 \left(\frac{M_y}{M_x} \right)$$

The resulting angle θ was calculated in radians and later converted to degrees for easier interpretation and application. This angle gives us an understanding of the individual's orientation at any given point in time, significantly enhancing indoor localization systems. Supplementary Figure S3 displays the

heading angles for indoor and outdoor activities over the Extrasensory dataset.

3.5.4 Mel-frequency cepstral coefficients (MFCCs)

The Mel-frequency cepstral coefficients (MFCCs) (González et al., 2015; Hou et al., 2023) are widely used in various applications. They serve as a compact representation of an audio signal, capturing essential characteristics while ignoring less informative variations. In the context of location recognition, MFCCs can help distinguish between different types of environments based on ambient noise or specific sound patterns. For instance, an indoor location might exhibit different MFCC patterns compared to an outdoor location due to the presence of echoes, HVAC noise, or human activity. MFCCs are computed through a series of transformations. We already segmented the audio data in section B. Each segment is passed through an FFT to obtain its frequency spectrum. Then we applied a set of Mel-filters to the frequency spectrum to capture the human perception of pitch. The logarithm of the energies of these Mel-frequencies is then taken, and a discrete cosine transform (DCT) is applied to the log energies. The resulting coefficients are the MFCCs. The equation for the k^{th} MFCC (c_k) can be summarized as

$$c_k = \sum_{n=1}^N \log(MF_n) \cos \left(\frac{\pi k(n-0.5)}{N} \right)$$

where MF_n is the Mel-filtered energy of the n^{th} frequency bin, and N is the total number of Mel-filters. The MFCCs calculated for indoor and outdoor activities over the Extrasensory dataset can be seen in Supplementary Figure S4.

3.6 Feature optimization using Yeo-Johnson power transformation

The Yeo-Johnson transformation (Xu et al., 2023) is a power transformation technique aimed at making the data more closely

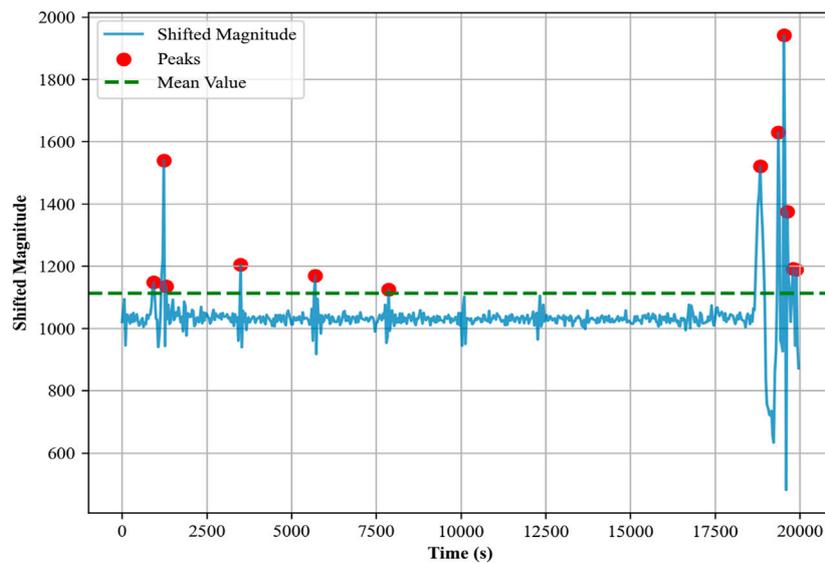


FIGURE 10
(A) original feature vector; (B) optimized feature vector over the Extrasensory dataset.

follow a Gaussian distribution, thereby optimizing its characteristics for further analysis. The transformation is similar to the Box-Cox transformation, but it has the advantage of handling negative numbers as well. We started by extracting various features from the time-series data. Each of the features serves as a column in our feature matrix. To apply the Yeo-Johnson transformation, we used the **PowerTransformer** class from the **sklearn.preprocessing** package, which internally fits the optimal λ for each feature based on the likelihood maximization. After fitting, the transformation was applied to each feature vector to create an optimized feature set. Here it is important to note that after optimization, we got two feature vectors, one for localization activities and the second for locomotion activities. We plotted two feature vectors the original *versus* optimized for Walking, Sitting, and Lying activities using only a few features including (Alazeb et al., 2023), FFT-Min/Max, Shannon entropy, and Kurtosis over the Extrasensory dataset in Figure 10. The transformation is defined as

$$y'_i = \begin{cases} \frac{[(y_i + 1)^\gamma - 1]}{\gamma}, & \text{if } y_i \geq 0 \text{ and } \gamma \neq 0, \\ \log(y_i + 1), & \text{if } y_i \geq 0 \text{ and } \gamma = 0, \\ \frac{-[(-y_i + 1)^\gamma - 1]}{\gamma}, & \text{if } y_i < 0 \text{ and } \gamma \neq 2, \\ -\log(y_i + 1), & \text{if } y_i < 0 \text{ and } \gamma = 2. \end{cases}$$

Here, y_i is the i^{th} observation, y'_i is the transformed value, and λ is the transformation parameter. It is important to mention that the Yeo-Johnson transformation is often determined by optimizing a likelihood function to find the best λ that minimizes the deviation from normality. The objective function for this is usually the negative log-likelihood, given by

$$L(\gamma) = -\log\left(\prod_{i=1}^n \int \left(\frac{y'_i}{\gamma}\right)\right)$$

where $\int \left(\frac{y'_i}{\gamma}\right)$ is the probability density function of the transformed data.

3.7 Feature evaluation analysis and comparisons

In order to validate and evaluate the robustness of the proposed feature set, we compare the extracted features in this study with other latest existing state-of-the-art methods. Initially, we categorize all features into different sets, from the latest SOTA systems (Bennasar et al., 2022; Tian et al., 2019; Muaaz et al., 2023). The features are partitioned into 4 sets. Each set is subjected to model training and validation. Our observations indicate that our proposed feature set outperforms other sets in performance, thereby validating the effectiveness, robustness, and novelty of our proposed features in enhancing model performance. The details of feature sets and their description are given in Table 1.

4 Experimental setup and datasets

4.1 Experimental setup

The research experiments were carried out on a laptop equipped with an Intel Core i5-8500U processor running at 3.10 GHz, 16.0 GB of RAM, and the Mac operating system. The Jupyter Notebook was utilized as the primary programming environment. We conducted comprehensive experiments using two widely recognized benchmark datasets, Opportunity and Extrasensory. The Opportunity dataset, a renowned benchmark in the field, captures data from various sensors. Another dataset used in our research is the Extrasensory dataset. With its rich sensory data, it offers an extensive range of

TABLE 1 Comparison of the proposed feature extraction with the latest SOTA.

Source	Features	Accuracy (%)
Bennasar et al. (2022)	Mean, standard deviation, root mean square, autocorrelation, permutation entropy, etc.	76
Tian et al. (2019)	Mean, variance, skewness, kurtosis, signal magnitude area, minimum/maximum, interquartile range, etc.	83
Muaaz et al. (2023)	Mean, variance, skewness, kurtosis, entropy, total energy, slope, etc.	88
Proposed	Skewness, kurtosis, phase angle, linear prediction, auto-regression etc.	96

human locomotion and localization activity. The time-series data was partitioned into approximately equal-length segments for the purpose of cross-validation (Xu et al., 2022a). In each of the k iterations, $k - 1$ segments were designated for training, and the remaining segment was set aside for testing. This procedure was repeated k times, guaranteeing that each segment served as a test set once, the rest being used as training sets. Importantly, we maintained a strict separation between training and test sets in every iteration, preventing any overlap or data sharing between them.

4.2 Dataset description

In the subsequent subsection, we provide comprehensive and detailed descriptions of each dataset used in our study. Each dataset is thoroughly introduced, highlighting its unique characteristics, data sources, and collection methods.

4.2.1 The opportunity dataset

The Opportunity dataset (Lukowicz et al., 2010) stands as a key benchmark in the domain of human activity recognition. It was collected from 12 subjects participating in various daily activities, ensuring a diverse representation. The dataset captures data from different sensors, such as accelerometers, gyroscopes, and magnetometers, strategically positioned on the participants' bodies and on certain daily-use objects. These sensors record data during both dynamic and static human activities. The dynamic activities include actions like walking, jogging, and opening doors, while the static activities encompass states like standing, sitting, and lying down. Additionally, there are more complex activities, like making coffee or preparing a sandwich, which involve interactions with objects and the environment. In total, the Opportunity dataset covers 17 different activities, ranging from basic locomotion tasks to more intricate, multi-step actions. These activities were recorded in diverse scenarios, both scripted and unscripted, to ensure a comprehensive representation of real-world conditions.

4.2.2 The extrasensory dataset

The Extrasensory dataset (Vaizman et al., 2017) is a robust collection of data sourced from 60 distinct participants, each uniquely identified by a UUID. These participants contributed thousands of data samples. While the majority of these samples were recorded at consistent 1-min intervals, there are instances where time gaps exist. Each data sample encompasses measurements derived from a variety of sensors present in the

participants' personal smartphones and a provided smartwatch. Furthermore, a large portion of these data points come furnished with context labels, as self-reported by the individual participants. In terms of device usage, the dataset includes data from 34 iPhone users and 26 Android users. The gender distribution is fairly balanced, with 34 females and 26 males. Sensors integrated into the dataset include accelerometer, gyroscope, magnetometer (He and Jin, 2008), GPS, audio, compass, and smartwatch sensors. The dataset provides a variety of human activities, including indoor, outdoor, transportation, and locomotion.

5 Results and analysis

In this section, we performed different experiments for the proposed system. The system is evaluated using different matrices, including confusion matrix, precision, recall, F1 score and receiver operating characteristic (ROC) curve. The detailed discussion and analysis are described below.

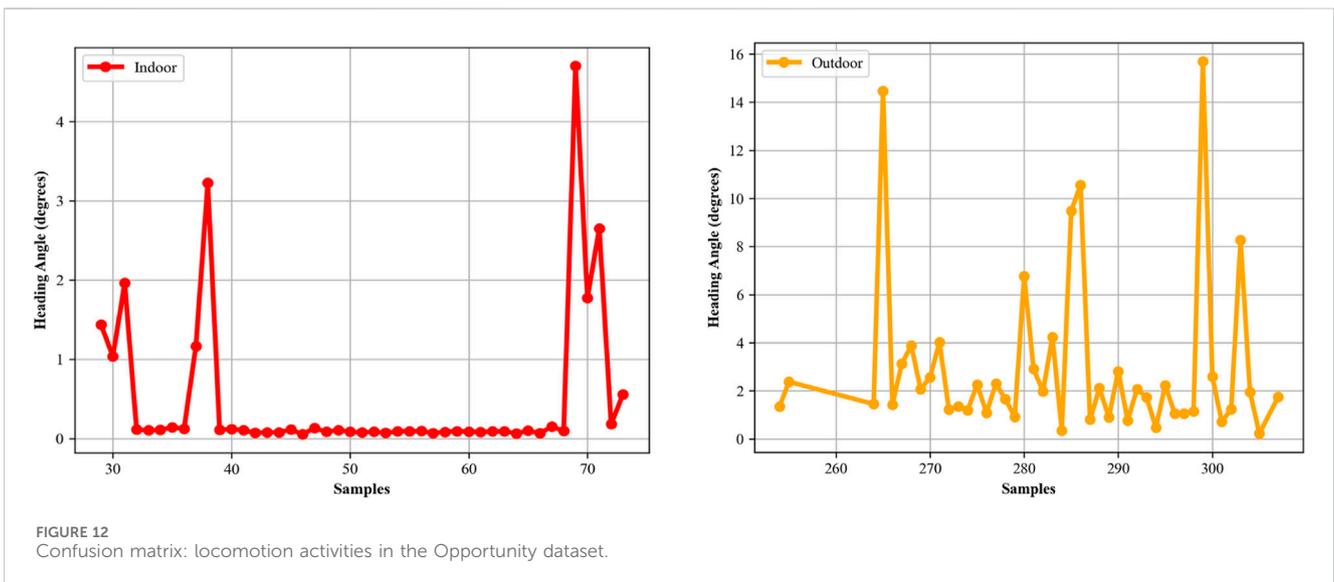
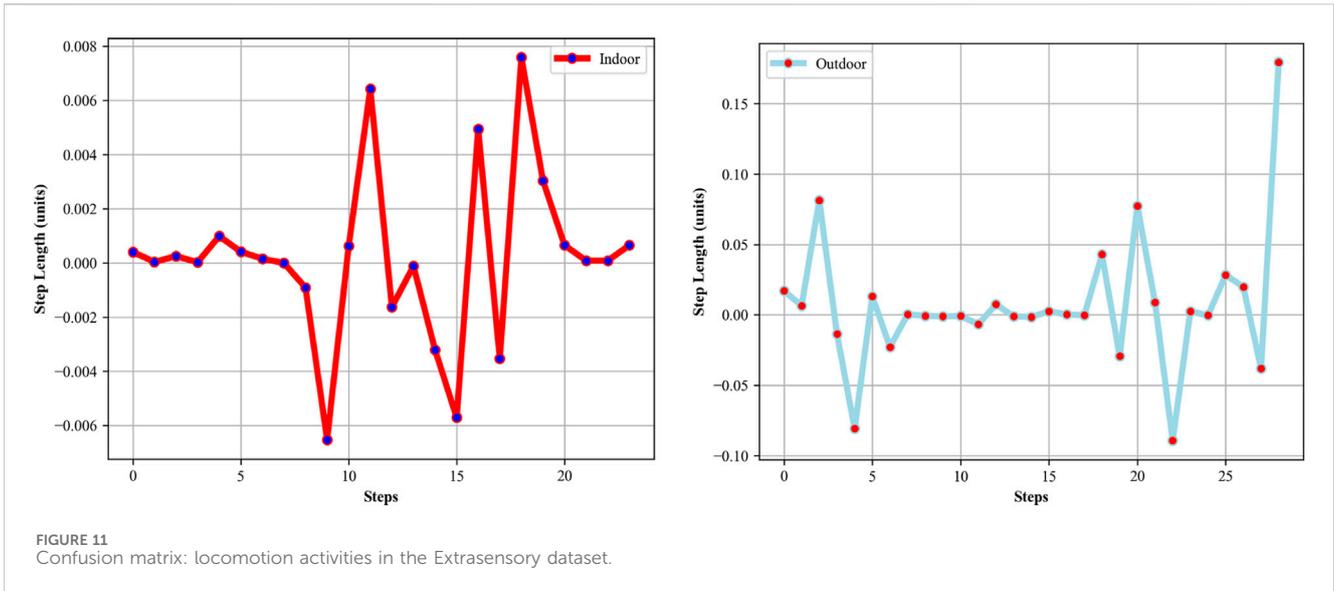
5.1 Confusion matrices for locomotion activities

We assessed our system's performance for locomotion activities across both datasets. Impressively, the system achieved a 97% accuracy rate on the Opportunity dataset and 89% on the Extrasensory dataset. Figures 11, 12 present the confusion matrices for both datasets.

The system shows high performance in identifying stationary activities. The system shows excellent performance in correctly identifying Sitting and Standing, with accuracies of 97% and 99%, respectively. This suggests that the system is highly effective in recognizing stationary activities. This capability is particularly applicable in contexts like workplace ergonomics or patient monitoring, where it is important to track the amount of time spent sitting or standing.

5.1.1 Moderate performance in dynamic activities

The performance in recognizing Walking is moderate, with an accuracy of 67%. The system seems to confuse Walking with Lying Down in some cases, which might be due to similar sensor patterns during slow walking or transitional movements. This indicates a potential area for improvement, especially in applications like fitness tracking or elderly care, where accurate recognition of walking is crucial.



5.1.2 Strong recognition of lying down

The system accurately identifies Lying Down in 96% of the cases, indicating its effectiveness in distinguishing this activity from others. This could be particularly relevant in healthcare applications, like patient monitoring systems, where detecting prolonged periods of lying down is important.

5.1.3 Near-perfect recognition of all activities in the opportunity dataset

The system shows near-perfect accuracy in recognizing all four activities: Standing, Walking, lying, and sitting, with accuracies of 100%, 98%, 96%, and 95% respectively. This high level of accuracy is significant for applications that require precise activity recognition, such as in advanced assistive technologies or smart home environments.

5.1.4 Applicability across diverse scenarios

Given the high accuracy in all activities, this system can be confidently applied to diverse real-world scenarios, from fitness tracking to elderly care, where accurate activity recognition is crucial. The system's ability to distinguish between similar activities (like lying and sitting) demonstrates its sophistication and reliability.

5.1.5 General observations

The higher overall mean accuracy in the Opportunity dataset (97.25%) compared to the Extrasensory dataset (89.75%) could be attributed to differences in sensor quality, data collection protocols, or the inherent nature of the activities in each dataset. The system's performance on the Opportunity dataset suggests its potential effectiveness in environments with structured activities, while the Extrasensory dataset results indicate the need for refinement in more complex or less structured environments.

TABLE 2 Precision, recall, and F1 score: locomotion activities in the Extrasensory dataset.

Classes	Precision	Recall	F1 score
Sitting	0.97	0.98	0.97
Walking	1.00	0.33	0.50
Lying Down	0.96	0.93	0.95
Standing	0.88	1.00	1.00
Macro-average	0.95	0.81	0.85

TABLE 3 Precision, recall and F1 score: locomotion activities in the Opportunity dataset.

Classes	Precision	Recall	F1 score
Standing	1.00	1.00	1.00
Walking	1.00	0.95	0.99
Lie	0.99	0.98	0.99
Sitting	1.00	1.00	1.00
Macro-average	0.99	0.98	0.99

5.2 Precision, recall, and F1 score values for locomotion activities

We continued to investigate in more depth the evaluation of our system using precision, recall, and F1 score. Across both datasets, the system demonstrated strong performance in all of these metrics. Tables 2 and 3 showcase the system's performance.

The performance evaluation of our system on the Extrasensory and Opportunity datasets, as reflected in Tables 4 and 5, highlights its strengths and areas for improvement in activity recognition. In the Extrasensory dataset, the system exhibits high precision across all activities, particularly for 'Sitting' and 'Lying Down', with scores of 0.97 and 0.96, respectively. This indicates a strong capability to correctly identify these activities when they occur. However, there is a notable discrepancy in the recall for 'Walking', at only 0.33, despite a perfect precision score. This suggests that while the system is accurate when it detects walking, it frequently misses walking instances. The overall macro-average scores of 0.95 for precision and 0.81 for recall, with an F1 score of 0.85, reflect competent performance but highlight the need for improvements in consistently recognizing walking activities. In contrast, the system's performance on the Opportunity dataset is exemplary, achieving near-perfect scores across all activities. Precision and recall are both 1.00 for 'Standing', 'Walking', and 'Sitting', with 'Lie' closely following at 0.99 for both metrics. This exceptional performance, encapsulated in macro-average scores of 0.99 for both precision and recall, and an F1 score of 0.99, demonstrates the system's high efficacy in structured environments with clear activity definitions.

5.3 Receiver operating characteristic curves for locomotion activities

To further investigate the system and stability of the system, we evaluated the system using the roc curve. The receiver operating

TABLE 4 Precision, recall, and F1 Score: localization activities in the Extrasensory dataset.

Classes	Precision	Recall	F1 score
Indoors	1.00	1.00	1.00
At School	0.88	1.00	0.93
Location Home	1.00	1.00	1.00
Location Workplace	1.00	1.00	1.00
Outside	1.00	0.93	0.96
Macro-average	0.97	0.98	0.97

TABLE 5 Comparisons of the proposed system with other systems.

Methods	Opportunity	Extrasensory
Javeed, M. et al. (2023)	0.88	-
Vanijkachorn and Visytsak (2021)	0.88	-
Han, J. (2019)	0.87	-
Gil-Martin, et al. (2020b)	0.67	-
Gioanni et al. (2016)	0.74	-
Vaizman et al. (2018)	-	0.83
Asim et al. (2020)	-	0.87
Abduallah et al. (2022)	-	0.87
Proposed system mean accuracy	0.97	0.96

characteristic (ROC) curve is a graphical representation used to evaluate the performance of a classifier across various decision thresholds. An important aspect of the ROC curve is the area under the curve (AUC). The AUC provides a single-number summary of the classifier's performance. A value of 1 indicates perfect classification, while a value of 0.5 suggests that the classifier's performance is no better than random guessing. In Supplementary Figures S5, S6, the ROC curves for both datasets can be observed clearly.

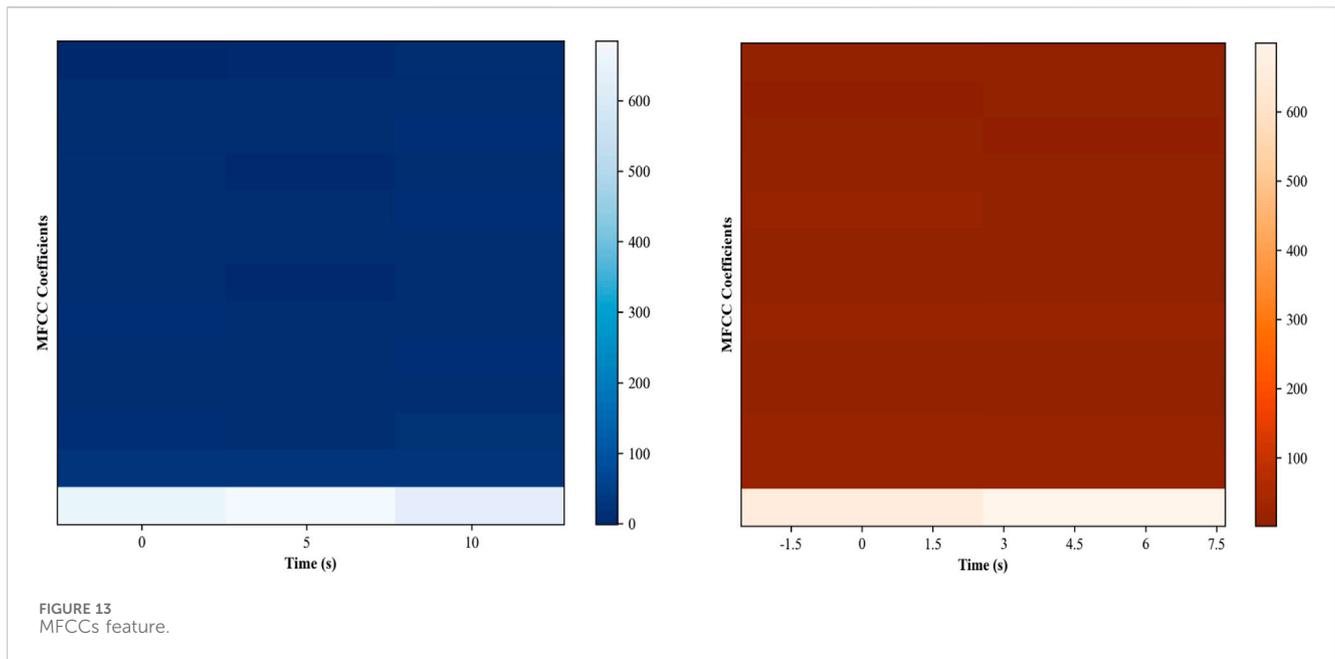
5.3.1 The Opportunity dataset

Standing (AUC = 1.00): The model's perfect score in identifying standing activities underscores its precision in environments such as elderly care, where detecting prolonged stationary periods is crucial for monitoring wellbeing and preventing health risks.

Walking (AUC = 0.98): The high AUC value for walking reflects the model's strong capability in accurately tracking walking movements, which is particularly beneficial for applications in fitness tracking and urban navigation systems.

Lying (AUC = 0.99): This near-perfect score indicates the model's effectiveness in recognizing lying down postures, an essential feature for patient monitoring in healthcare settings, especially for bedridden individuals.

Sitting (AUC = 1.00): The model's flawless detection of sitting activities is critical to workplace ergonomics and sedentary lifestyle analysis, aiding in developing interventions for prolonged inactivity.



5.3.2 The extrasensory dataset

Sitting (AUC = 0.97): The high AUC for sitting activities demonstrates the system's reliability in identifying sedentary behaviors, which is vital in office settings for promoting active work habits.

Lying Down (AUC = 0.96): This score reflects the model's adeptness in detecting lying down positions, applicable in sleep analysis and residential healthcare monitoring.

Walking (AUC = 0.67): The lower AUC in this category suggests challenges in distinguishing walking from other movements in complex environments, pointing to potential areas for improvement in applications requiring precise motion tracking.

Standing (AUC = 0.99): The high accuracy in identifying standing positions is crucial in varied contexts such as retail analytics and customer behavior studies, where understanding patterns of movement and pause can enhance service strategies.

5.4 Confusion matrix for localization activities

We conducted experiments to recognize localization activities. These experiments were conducted using the extrasensory dataset, which offers a variety of human localization activities. Initially, we generated a confusion matrix, followed by an assessment of the system's performance using precision, recall, and the F1 score. Moreover, to assess the system's stability and effectiveness, we employed the ROC curve. Each experiment is thoroughly discussed, and the resulting outcomes are presented below.

In this experiment, we evaluate the system for localization activities. We observed a good accuracy rate of 96%. The confusion matrix is given in [Figure 13](#).

The confusion matrix for the Extrasensory dataset's localization activities provides valuable insights into the system's capability to

accurately distinguish between different environmental contexts. **Indoors (accuracy = 100%):** The perfect accuracy in identifying indoor activities showcases the model's precision in environments like homes, offices, or malls. This is crucial for applications such as smart home automation, where accurate indoor localization can trigger context-specific actions like adjusting lighting or temperature. **At School (accuracy = 90%):** The high accuracy in recognizing activities at school is significant for educational applications, such as attendance tracking or student activity monitoring. The confusion with other locations, although minimal, suggests room for improvement in differentiating between similar educational and other public environments. **Location Home (accuracy = 100%):** Flawless detection of activities at home points to the model's effectiveness in residential settings, crucial for applications like security systems or elder care monitoring, where distinguishing home activities is essential for providing personalized and situational services. **Outside (accuracy = 100%):** The model's ability to perfectly identify outdoor activities is vital for location-based services, such as navigation aids and outdoor fitness tracking. This can enhance user experiences in applications that rely on outdoor localization. **Location Workplace (accuracy = 94%):** The high accuracy in identifying workplace activities is important for enterprise solutions, like workforce management and safety compliance monitoring. The slight confusion with other locations highlights the need for further refinement to distinguish workplace activities from similar environments with greater accuracy.

5.5 Precision, recall, and F1 score values for localization activities

To check the performance of the system for localization activities, we calculated the precision, recall, and F1 score. In [Table 4](#).

5.6 Receiver operating characteristic curve for localization activities

We plotted the ROC curve of localization activities to ensure that the proposed system is well trained, accurate, and stable. The system showed very impressive results in recognizing location-based activities. The ROC curve can be examined in [Supplementary Figure S7](#).

5.7 Detailed performance analysis

In this subsection, we delve deeper into the performance metrics across different datasets and activities, shedding light on the trade-offs between accuracy and computational efficiency.

5.7.1 Locomotion activities

For the Opportunity dataset, as the number of iterations increased from 5 to 50, the accuracy improved from 74.76% to 97.14%, while the computation time increased from 2.53 s to 19.49 s. For the Extrasensory dataset's locomotion activity, accuracy improved from 61.76% at 5 iterations to 89.14% at 50 iterations, with computation time increasing from 1.49 s to 14.49 s. It is evident that the model trained on the Opportunity dataset achieved a relatively higher accuracy with more iterations compared to the Extrasensory dataset. However, the computational cost was also higher for the Opportunity dataset. The time complexity and efficiency plot can be seen in [Supplementary Figure S8](#).

5.7.2 Localization activities

For the Extrasensory dataset's localization activity, the accuracy improved from 85.76% at 10 iterations to 95.61% at 50 iterations. The corresponding computation time rose from 1.93 s to 11.90 s. The model's accuracy for localization showed a steady improvement with increased iterations, and the computational cost was relatively consistent, indicating efficient model performance. The time complexity and efficiency plot for the localization activities can be seen in [Supplementary Figure S9](#).

5.8 Comparison between locomotion and localization activities

For the Extrasensory dataset, the model's performance for localization activities was consistently higher in terms of accuracy compared to locomotion activities, across the same number of iterations. However, the computation time for localization was slightly longer, indicating a trade-off between accuracy and computational efficiency.

Finally, our system was subject to a comparative analysis against other existing systems, revealing that our model excels in terms of accuracy. [Table 5](#) provides a comprehensive overview of the comparisons between our system and state-of-the-art techniques.

6 Discussion

A system for recognizing human locomotion and location-related activities is introduced in this work. This system utilizes advanced noise filtering techniques, signal segmentation methods, feature extraction processes, and hybrid models to effectively identify both locomotion and localization activities. It is designed to be versatile and can find applications in various real-world scenarios such as sports, healthcare, security, location recognition, and many more real-world applications. Our system, while advanced, has certain limitations. The sensors we utilized, especially GPS and IMU, have inherent challenges; GPS may not always be accurate indoors or amidst tall urban structures, and IMUs can drift over time. Second, our reliance on the Opportunity and Extrasensory datasets, although reputable, does not capture all human activity nuances, as evidenced by the challenge of recognizing walking activities. Additionally, our experiments were conducted on a specific laptop configuration. When transitioning to real-world wearable devices, differing computational capabilities might influence the system's performance. Moving forward, we plan to enhance our system by exploring advanced sensor fusion techniques, allowing for more robust data integration from various sensors. We also recognize the need to diversify our datasets and will seek collaborations to gather more varied and balanced human activity data. Importantly, to ensure our system's efficacy on wearable devices, we will explore optimizations customized to devices with varied computational constraints, ensuring our HAR system remains efficient and real-time in practical scenarios.

7 Conclusion

This study introduces a new and resilient system designed to identify human locomotion and localization activities effectively. The system was developed with a focus on utilizing advanced techniques such as sensor data filtering, windowing, and segmentation, along with innovative methods for feature extraction. It is important to mention that our primary emphasis was on recognizing localization activities, for which we employed robust feature extraction techniques, including step count, step length, and heading angle. In addition to manual feature extraction, we introduced a hybrid model that harnesses both machine learning and deep learning approaches to enhance accuracy in recognition tasks. As a result, the presented system demonstrates precise and efficient recognition of both locomotion and localization activities.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: The Opportunity Dataset: <https://archive.ics.uci.edu/dataset/226/opportunity+activity+recognition>; The ExtraSensory Dataset: <http://extrasensory.ucsd.edu/>.

Author contributions

DK: Conceptualization, Software, Writing—original draft, Writing—review and editing. MoA: Funding acquisition, Methodology, Validation, Writing—review and editing. MaA: Funding acquisition, Resources, Software, Writing—review and editing. NA: Conceptualization, Formal Analysis, Funding acquisition, Resources, Validation, Writing—original draft, Writing—review and editing. AA: Conceptualization, Funding acquisition, Resources, Writing—review and editing. AJ: Formal Analysis, Funding acquisition, Resources, Writing—original draft, Writing—review and editing. HL: Formal Analysis, Funding acquisition, Resources, Software, Writing—original draft, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The APC was funded by the Open Access Initiative of the University of Bremen and the DFG via SuUB Bremen. Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2023R97), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2024/R/1445). The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA for funding this research work through the project number “NBU-FFR-2024-231-02”. The authors are thankful to the Deanship of Scientific Research at Najran University for funding this work under the Research Group Funding program grant code (NU/RG/SERC/12/40).

References

- Abdullah, M., Lejarza, F., Cahail, S., Claudel, C., and Thomaz, E. (2022). “HAR-GCNN: deep graph CNNs for human activity recognition from highly unlabeled mobile sensor data,” in Proceedings of the 2022 IEEE International Conference on Pervasive Computing and Communications Workshops and Other Affiliated Events, 335–340.
- Ahmad, F. (2022). Deep image retrieval using artificial neural network interpolation and indexing based on similarity measurement. *CAAI Trans. Intelligent Technol.* 7 (2), 200–218. doi:10.1049/cit2.12083
- Ahmad, J., Batool, M., and Kim, K. (2020). Stochastic recognition of physical activity and healthcare using tri-axial inertial wearable sensors. *Appl. Sci.* 10, 7122. doi:10.3390/app10207122
- Ahn, H.-S., and Yu, W. (2007). “Indoor mobile robot and pedestrian localization techniques,” in Proceedings of the 2007 International Conference on Control, Seoul, Korea, 2350–2354. Automation and Systems.
- Alazeb, A., Azmat, U., Al Mudawi, N., Alshahrani, A., Alotaibi, S. S., Almujally, N. A., et al. (2023). Intelligent localization and deep human activity recognition through IoT devices. *Sensors* 23, 7363. doi:10.3390/s23177363
- Altun, K., and Barshan, B. (2010). “Human activity recognition using inertial/magnetic sensor units,” in Proceedings of the 1st International Workshop on Human Behavior Understanding, Istanbul, Turkey, 38–51.
- AlZubi, H. S., Gerrard-Longworth, S., Al-Nuaimy, W., Goulermas, Y., and Preece, S. (2014). “Human activity classification using a single accelerometer,” in Proceedings of the 2014 14th UK Workshop on Computational Intelligence, Bradford, United Kingdom (UKCI), 1–6.
- Asim, Y., Azam, M. A., Ehatisham-ul-Haq, M., Naeem, U., and Khalid, A. (2020). Context-aware human activity recognition (CAHAR) in-the-Wild using smartphone accelerometer. *IEEE Sensors J.* 20, 4361–4371. doi:10.1109/jsen.2020.2964278
- Azmat, U., Ghadi, Y. Y., Shloul, T. a., Alsuhibany, S. A., Ahmad, J., and Park, J. (2022). Smartphone sensor-based human locomotion surveillance system using multilayer perceptron. *Appl. Sci.* 12, 2550. doi:10.3390/app12052550
- Bae, T. W., Lee, S. H., and Kwon, K. K. (2020). An adaptive median filter based on sampling rate for R-peak detection and major-arrhythmia analysis. *Sensors* 20, 6144. doi:10.3390/s20216144
- Batchuluun, G., Kang, J. K., Nguyen, D. T., Pham, T. D., Arsalan, M., and Park, K. R. (2021). Action recognition from thermal videos using joint and skeleton information. *IEEE Access* 9, 11716–11733. doi:10.1109/access.2021.3051375
- Bennasar, M., Price, B. A., Gooch, D., Bandara, A. K., and Nuseibeh, B. (2022). Significant features for human activity recognition using tri-axial accelerometers. *Sensors* 22, 7482. doi:10.3390/s22197482
- Bhelkar, V., and Shedje, D. K. (2016). “Different types of wearable sensors and health monitoring systems: a survey,” in Proceedings of the 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology, Bangalore, India (iCATccT), 43–48.
- Bruno, B., Mastrogiovanni, F., and Sgorbissa, A. (2015). Wearable inertial sensors: applications, challenges, and public test benches. *IEEE Robotics Automation Mag.* 22, 116–124. doi:10.1109/mra.2015.2448279
- Chen, B., Meng, F., Tang, H., and Tong, G. (2023). Two-level attention module based on spurious-3D residual networks for human action recognition. *Sensors* 23, 1707. doi:10.3390/s23031707
- Cömert, Z., Boopathi, A. M., Velappan, S., Yang, Z., and Kocamaz, A. F. (2018). “The influences of different window functions and lengths on image-based time-frequency features of fetal heart rate signals,” in Proceedings of the 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, 1–4.
- Di, Y., Li, R., Tian, H., Guo, J., Shi, B., Wang, Z., et al. (2023). A maneuvering target tracking based on fastLMM-extended Viterbi algorithm. *Neural Comput. Appl.* doi:10.1007/s00521-023-09039-1
- Gangothri, B. N., Satamraju, K. P., and Malarkodi, B. (2023). “Sensor-based ambient healthcare architecture using blockchain and internet of things,” in Proceedings of the 2023 10th International Conference on Signal Processing and Integrated Networks (SPIN), India (Noida), 543–546.

Acknowledgments

The authors are thankful to Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2023R97), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphys.2024.1344887/full#supplementary-material>

- Gil-Martín, M., Sánchez-Hernández, M., and San-Segundo, R. (2020a). Human activity recognition based on deep learning techniques. *Proceedings* 42, 15. doi:10.3390/ecs-a-6-06539
- Gil-Martín, M., San-Segundo, R., Fernández-Martínez, F., and de Córdoba, R. (2020b). Human activity recognition adapted to the type of movement. *Comput. Electr. Eng.* 88, 106822. doi:10.1016/j.compeleceng.2020.106822
- Gioanni, L., Dartigues-Pallex, C., Lavirotte, S., and Tigli, J.-Y. (2016). "Opportunistic human activity recognition: a study on opportunity dataset," in Proceedings of MOBIQUITOUS 2016, New York, NY (ACM), 290–291.
- González, S., Sedano, J., Villar, J. R., Corchado, E., Herrero, Á., and Baroque, B. (2015). Features and models for human activity recognition. *Neurocomputing* 167, 52–60. doi:10.1016/j.neucom.2015.01.082
- Gu, F., Khoshelham, K., Shang, J., Yu, F., and Wei, Z. (2017). Robust and accurate smartphone-based step counting for indoor localization. *IEEE Sensors J.* 17, 3453–3460. doi:10.1109/jsen.2017.2685999
- Gu, F., Khoshelham, K., Yu, C., and Shang, J. (2019). Accurate step length estimation for pedestrian dead reckoning localization using stacked autoencoders. *IEEE Trans. Instrum. Meas.* 68, 2705–2713. doi:10.1109/tim.2018.2871808
- Guo, L., Cheng, S., Liu, J., Wang, Y., Cai, Y., and Hong, X. (2022). Does social perception data express the spatio-temporal pattern of perceived urban noise? A case study based on 3,137 noise complaints in Fuzhou, China. *Appl. Acoust.* 201, 109129. doi:10.1016/j.apacoust.2022.109129
- Han, J. (2019). GraphConvLSTM: spatiotemporal learning for activity recognition with wearable sensors. *Proc. GLOBECOM*. doi:10.1109/GLOBECOM38437.2019.9013934
- Hanif, M. A., Akram, T., Shahzad, A., Khan, M. A., Tariq, U., Choi, J.-I., et al. (2022). Smart devices based multisensory approach for complex human activity recognition. *Comput. Mater. Continua* 70, 3221–3234. doi:10.32604/cmc.2022.019815
- Hao, T. (2021). "Human activity recognition based on WaveNet," in Proceedings of the 2021 IEEE 7th World Forum on Internet of Things, New Orleans, LA, USA (WF-IoT), 824–829.
- Hartmann, Y., Liu, H., Lahrberg, S., and Schultz, T. (2022). "Interpretable high-level features for human activity recognition," in *Biosignals*, 40–49.
- Hartmann, Y., Liu, H., and Schultz, T. (2023). "High-level features for human activity recognition and modeling," in International Joint Conference on Biomedical Engineering Systems and Technologies, Springer Nature Switzerland (Cham), 141–163.
- Hassan, F. S., and Gutub, A. (2022). Improving data hiding within colour images using hue component of HSV colour space. *CAAI Trans. Intelligent Technol.* 7 (1), 56–68. doi:10.1049/cit.2.12053
- Hatamikia, S., Maghooli, K., and Nasrabad, A. M. (2014). The emotion recognition system based on autoregressive model and sequential forward feature selection of electroencephalogram signals. *J. Med. Signals Sensors* 4, 194–201. doi:10.4103/2228-7477.137777
- He, Z.-Y., and Jin, L.-W. (2008). "Activity recognition from acceleration data using AR model representation and SVM," in Proceedings of the 2008 International Conference on Machine Learning and Cybernetics, China (Kunming), 2245–2250.
- Hou, X., Zhang, L., Su, Y., Gao, G., Liu, Y., Na, Z., et al. (2023). A space crawling robotic bio-paw (SCRBP) enabled by triboelectric sensors for surface identification. *Nano Energy* 105, 108013. doi:10.1016/j.nanoen.2022.108013
- Hu, G., Zhang, W., Wan, H., and Li, X. (2020). Improving the heading accuracy in indoor pedestrian navigation based on a decision tree and kalman filter. *Sensors* 20, 1578. doi:10.3390/s20061578
- Hu, X., Kuang, Q., Cai, Q., Xue, Y., Zhou, W., and Li, Y. (2022). A coherent pattern mining algorithm based on all contiguous column bicluster. *J. Artif. Intell. Technol.* 2 (3), 80–92. doi:10.37965/jait.2022.0105
- Hu, Z., Ren, L., Wei, G., Qian, Z., Liang, W., Chen, W., et al. (2023). Energy flow and functional behavior of individual muscles at different speeds during human walking. *IEEE Trans. Neural Syst. Rehabilitation Eng.* 31, 294–303. doi:10.1109/TNSRE.2022.3221986
- Hussain, I., Tahir, S., Humayun, M., Almfareh, M. F., Jhanjhi, N. Z., and Qamar, F. (2022). "Health monitoring system using internet of things (IoT) sensing for elderly people," in Proceedings of the 2022 14th International Conference on Mathematics, Actuarial Science, Karachi, Pakistan (Computer Science and Statistics (MACS)), 1–5.
- Jiawei, X., Park, S. H., Zhang, X., and Hu, J. (2022). The improvement of road driving safety guided by visual inattentive blindness. *IEEE Trans. Intelligent Transp. Syst.* 23 (6), 4972–4981. doi:10.1109/TITS.2020.3044927
- Jaramillo, I. E., Jeong, J. G., Lopez, P. R., Lee, C.-H., Kang, D.-Y., Ha, T.-J., et al. (2022). Real-time human activity recognition with IMU and encoder sensors in wearable exoskeleton robot via deep learning networks. *Sensors* 22, 9690. doi:10.3390/s22249690
- Javeed, M., and Jalal, A. (2023). "Deep activity recognition based on patterns discovery for healthcare monitoring," in Proceedings of the 2023 4th International Conference on Advancements in Computational Sciences, Lahore, Pakistan (ICACS), 1–6. doi:10.1109/ICACS55311.2023.10089764
- Javeed, M., Jalal, A., and Kim, K. (2021). "Wearable sensors based exertion recognition using statistical features and random forest for physical healthcare monitoring," in Proceedings of the 2021 International Bhurban Conference on Applied Sciences and Technologies, Islamabad, Pakistan (IBCAST), 512–517.
- Javeed, M., Mudawi, N. A., Alazeb, A., Alotaibi, S. S., Almujally, N. A., and Ahmad, J. (2023). Deep ontology-based human locomotor activity recognition system via multisensory devices. *IEEE Access* 11, 105466–105478. doi:10.1109/access.2023.3317893
- Jiang, H., Chen, S., Xiao, Z., Hu, J., Liu, J., and Dustdar, S. (2023). Pa-count: passenger counting in vehicles using wi-fi signals. *IEEE Trans. Mob. Comput.*, 1–14. doi:10.1109/TMC.2023.3263229
- Jiang, Y., and He, X. (2020). Overview of applications of the sensor technologies for construction machinery. *IEEE Access* 8, 110324–110335. doi:10.1109/access.2020.3001968
- Jiawei, X., Sicheng, P., Poly, S., Seok, P., and Kun, G. (2022). Human-Factors-in-Driving-Loop: driver identification and verification via a deep learning approach using psychological behavioral data. *IEEE Trans. Intelligent Transp. Syst. (IEEE-TITS)*. doi:10.1109/TITS.2022.3225782
- Kamarudin, K., Mamduh, S. M., Shakaff, A. Y. M., and Zakaria, A. (2014). Performance analysis of the Microsoft Kinect sensor for 2D simultaneous localization and mapping (SLAM) techniques. *Sensors* 14, 23365–23387. doi:10.3390/s141223365
- Kang, X., Huang, B., and Qi, G. (2018). A novel walking detection and step counting algorithm using unconstrained smartphones. *Sensors* 18, 297. doi:10.3390/s18010297
- Khairy, H. (2022). "Statistical features versus deep learning representation for suspicious human activity recognition," in Proceedings of the 2022 4th Novel Intelligent and Leading Emerging Sciences Conference, Giza, Egypt (NILES), 126–130.
- Kwapisz, J. R., Weiss, G. M., and Moore, S. A. (2011). "Activity recognition using cell phone accelerometers," in Proceedings of the 17th Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, 74–82.
- Latha, L., Musammil, M., and Sazwan, F. A. (2022). "Human action recognition using deep learning methods (CNN-LSTM) without sensors," in Proceedings of the 2022 1st International Conference on Computational Science and Technology (Chennai, India: ICCST), 447–451.
- Leone, A., Rescio, G., Diraco, G., Manni, A., Siciliano, P., and Caroppo, A. (2022). Ambient and wearable sensor technologies for energy expenditure quantification of ageing adults. *Sensors* 22, 4893. doi:10.3390/s22134893
- Li, F., Shirahama, K., Nisar, M. A., Köping, L., and Grzegorzek, M. (2018). Comparison of feature learning methods for human activity recognition using wearable sensors. *Sensors* 18, 679. doi:10.3390/s18020679
- Li, R., Li, H., and Shi, W. (2020). Human activity recognition based on LPA. *Multimedia Tools Appl.* 79, 31069–31086. doi:10.1007/s11042-020-09150-8
- Liang, X., Huang, Z., Yang, S., and Qiu, L. (2018). Device-free motion and trajectory detection via RFID. *ACM Trans. Embed. Comput. Syst.* 17 (4), 1–27. doi:10.1145/3230644
- Liu, A., Zhai, Y., Xu, N., Nie, W., Li, W., and Zhang, Y. (2022c). Region-aware image captioning via interaction learning. *IEEE Trans. Circuits Syst. Video Technol.* 32 (6), 3685–3696. doi:10.1109/tcsvt.2021.3107035
- Liu, C., Wu, T., Li, Z., Ma, T., and Huang, J. (2022d). Robust online tensor completion for IoT streaming data recovery. *IEEE Trans. Neural Netw. Learn. Syst.* 34, 10178–10192. doi:10.1109/TNNLS.2022.3165076
- Liu, H. (2021). *Biosignal processing and activity modeling for multimodal human activity recognition*. Doctor thesis (Universität Bremen).
- Liu, H., Hartmann, Y., and Schultz, T. (2021). "Motion units: generalized sequence modeling of human activities for sensor-based activity recognition," in Proceedings 29th European Signal Processing conference, 680–688.
- Liu, H., Hartmann, Y., and Schultz, T. (2022a). "A practical wearable sensor-based human activity recognition research pipeline," in *Healthinf*, 847–856.
- Liu, H., and Schultz, T. (2018). "ASK: a framework for data acquisition and activity recognition," in *Biosignals*, 262–268.
- Liu, H., and Schultz, T. (2019). "A wearable real-time human activity recognition system using biosensors integrated into a knee bandage," in Proceedings BIODEVICES, 47–55.
- Liu, H., and Schultz, T. (2022). "How long are various types of daily activities? Statistical analysis of a multimodal Wearable sensor-based human activity dataset," in Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2022), 680–688.
- Liu, H., Xu, Y., and Chen, F. (2023a). Sketch2Photo: synthesizing photo-realistic images from sketches via global contexts. *Eng. Appl. Artif. Intell.* 117, 105608. doi:10.1016/j.engappai.2022.105608
- Liu, H., Xu, Y., and Chen, F. (2023d). Sketch2Photo: synthesizing photo-realistic images from sketches via global contexts. *Eng. Appl. Artif. Intell.* 117, 105608. doi:10.1016/j.engappai.2022.105608
- Liu, H., Xue, T., and Schultz, T. (2023c). "On a real real-time wearable human activity recognition system," in Proceedings of the WHC, 711–720.
- Liu, H., Yuan, H., Hou, J., Hamzaoui, R., and Gao, W. (2022b). PUFA-GAN: a frequency-aware generative adversarial network for 3D point cloud upsampling. *IEEE Trans. Image Process.* 31, 7389–7402. doi:10.1109/TIP.2022.3222918

- Liu, X., Shi, T., Zhou, G., Liu, M., Yin, Z., Yin, L., et al. (2023g). Emotion classification for short texts: an improved multi-label method. *Humanit. Soc. Sci. Commun.* 10 (1), 306. doi:10.1057/s41599-023-01816-6
- Liu, X., Wang, S., Lu, S., Yin, Z., Li, X., Yin, L., et al. (2023f). Adapting feature selection algorithms for the classification of Chinese texts. *Systems* 11 (9), 483. doi:10.3390/systems11090483
- Liu, X., Zhou, G., Kong, M., Yin, Z., Li, X., Yin, L., et al. (2023e). Developing multi-labelled corpus of twitter short texts: a semi-automatic method. *Systems* 11 (8), 390. doi:10.3390/systems11080390
- Liu, Z., Wen, C., Su, Z., Liu, S., Sun, J., Kong, W., et al. (2023b). Emotion-semantic-aware dual contrastive learning for epistemic emotion identification of learner-generated reviews in MOOCs. *IEEE Trans. Neural Netw. Learn. Syst.*, 1–14. doi:10.1109/TNNLS.2023.3294636
- Lu, S., Ding, Y., Liu, M., Yin, Z., Yin, L., and Zheng, W. (2023). Multiscale feature extraction and fusion of image and text in VQA. *Int. J. Comput. Intell. Syst.* 16 (1), 54. doi:10.1007/s44196-023-00233-6
- Lukowicz, P., Pirkil, G., Bannach, D., Wagner, F., Calatroni, A., Förster, K., et al. (2010). "Recording a complex, multi-modal activity data set for context recognition," in Proceedings, 161–166.
- Ma, X., Dong, Z., Quan, W., Dong, Y., and Tan, Y. (2023). Real-time assessment of asphalt pavement moduli and traffic loads using monitoring data from built-in sensors: optimal sensor placement and identification algorithm. *Mech. Syst. Signal Process.* 187, 109930. doi:10.1016/j.ymssp.2022.109930
- Mao, Y., Sun, R., Wang, J., Cheng, Q., Kiong, L. C., and Ochieng, W. Y. (2022b). New time-differenced carrier phase approach to GNSS/INS integration. *GPS Solutions* 26 (4), 122. doi:10.1007/s10291-022-01314-3
- Mao, Y., Zhu, Y., Tang, Z., and Chen, Z. (2022a). A novel airspace planning algorithm for cooperative target localization. *Electronics* 11 (18), 2950. doi:10.3390/electronics11182950
- Meng, J., Li, Y., Liang, H., and Ma, Y. (2022). Single-image dehazing based on two-stream convolutional neural network. *J. Artif. Intell. Technol.* 2 (3), 100–110. doi:10.37965/jait.2022.0110
- Miao, Y., Wang, X., Wang, S., and Li, R. (2023). Adaptive switching control based on dynamic zero-moment point for versatile hip exoskeleton under hybrid locomotion. *IEEE Trans. Industrial Electron.* 70 (11), 11443–11452. doi:10.1109/TIE.2022.3229343
- Muaz, M., Waqar, S., and Pätzold, M. (2023). Orientation-independent human activity recognition using complementary radio frequency sensing. *Sensors* 23, 5810. doi:10.3390/s23135810
- Mutegeki, R., and Han, D. S. (2020). "A CNN-LSTM approach to human activity recognition," in Proceedings of the International Conference on Artificial Intelligence and Information Communication (ICAIIIC), 362–366.
- Nafea, O., Abdul, W., and Muhammad, G. (2022). Multi-sensor human activity recognition using CNN and GRU. *Int. J. Multimedia Inf. Retr.* 11, 135–147. doi:10.1007/s13735-022-00234-9
- Perez, A. J., and Zeadally, S. (2021). Recent advances in wearable sensing technologies. *Sensors* 21, 6828. doi:10.3390/s21106828
- Pham, V. T., Nguyen, T. K., Nguyen, D. A., Dang, N. D., Huynh, H. T., and Tran, D.-T. (2021). "Adaptive step length estimation support indoor positioning system using low-cost inertial measurement units," in Proceedings of the 2020 IEEE Eighth International Conference on Communications and Electronics, Phu Quoc Island, Vietnam (ICCE), 271–275.
- Poulose, A., Eyobu, O. S., and Han, D. S. (2019a). An indoor position-estimation algorithm using smartphone IMU sensor data. *IEEE Access* 7, 11165–11177. doi:10.1109/ACCESS.2019.2891942
- Poulose, A., Kim, J. H., Han, D. S., and Medina, J. (2022). HIT HAR: human image thresholding machine for human activity recognition using deep learning models. *Intell. Neurosci.* 2022, 1808990. doi:10.1155/2022/1808990
- Poulose, A., Senouci, B., and Han, D. S. (2019b). Performance analysis of sensor fusion techniques for heading estimation using smartphone sensors. *IEEE Sensors J.*
- Puangragsa, U., Setakornnukul, J., Dankulchai, P., and Phasukkit, P. (2022). 3D Kinect camera scheme with time-series deep-learning algorithms for classification and prediction of lung tumor motility. *Sensors* 22, 2918. doi:10.3390/s22082918
- Qi, M., Cui, S., Chang, X., Xu, Y., Meng, H., Wang, Y., et al. (2022). Multi-region nonuniform brightness correction algorithm based on L-channel gamma transform. *Security and communication networks*, 2022. doi:10.1155/2022/2675950
- Qu, J., Mao, B., Li, Z., Xu, Y., Zhou, K., Cao, X., et al. (2023a). Recent progress in advanced tactile sensing technologies for soft grippers. *Adv. Funct. Mater.* 33 (41), 2306249. doi:10.1002/adfm.202306249
- Qu, J., Yuan, Q., Li, Z., Wang, Z., Xu, F., Fan, Q., et al. (2023b). All-in-One strain-triboelectric sensors based on environment-friendly ionic hydrogel for wearable sensing and underwater soft robotic grasping. *Nano Energy* 111, 108387. doi:10.1016/j.nanoen.2023.108387
- Ramanujam, E., Padmavathi, S., Dharshani, G., and Madhumita, M. R. R. (2019). "Evaluation of feature extraction and recognition for human activity using smartphone based accelerometer data," in Proceedings of the 2019 11th International Conference on Advanced Computing, Chennai, India (ICoAC), 86–89.
- Reddy, N. S., Saketh, M. S., and Dhar, S. (2016). "Review of sensor technology for mine safety monitoring systems: a holistic approach," in Proceedings of the 2016 IEEE First International Conference on Control, Measurement and Instrumentation, Kolkata, India (CMI), 429–434.
- Ronald, M., Poulose, A., and Han, D. S. (2021). iSPLInception: an inception-ResNet deep learning architecture for human activity recognition. *IEEE Access* 9, 68985–69001. doi:10.1109/ACCESS.2021.3078184
- Rustam, F., Reshi, A. A., Ashraf, I., Mehmood, A., Ullah, S., Khan, D. M., et al. (2020). Sensor-based human activity recognition using deep stacked multilayered perceptron model. *IEEE Access* 8, 218898–218910. doi:10.1109/access.2020.3041822
- Saboor, A., Kask, T., Kuusik, A., Alam, M. M., Le Moulec, Y., Niazi, I. K., et al. (2020). Latest research trends in gait analysis using wearable sensors and machine learning: a systematic review. *IEEE Access* 8, 167830–167864. doi:10.1109/access.2020.3022818
- Sevinç, H., Ayvaz, U., Özlem, K., Elmoughni, H., Tunçay Atalay, A., Atalay, Ö., et al. (2020). "Step length estimation using sensor fusion," in Proceedings of the 2020 IEEE International Conference on Flexible and Printable Sensors and Systems (FLEPS), Manchester, United Kingdom, August 16, 2020.
- She, Q., Hu, R., Xu, J., Liu, M., Xu, K., and Huang, H. (2022). Learning high-DOF reaching-and-grasping via dynamic representation of gripper-object interaction. *ACM Trans. Graph.* 41 (4), 1–14. doi:10.1145/3528223.3530091
- Shen, X., Jiang, H., Liu, D., Yang, K., Deng, F., Lui, J. C. S., et al. (2022). PupilRec: leveraging pupil morphology for recommending on smartphones. *IEEE Internet Things J.* 9 (17), 15538–15553. doi:10.1109/JIOT.2022.3181607
- Sobhan, S., Islam, S., Valero, M., Shahriar, H., and Ahamed, S. I. (2021). "Data analysis methods for health monitoring sensors: a survey," in Proceedings of the 2021 IEEE 45th Annual Computers, Software, and Applications Conference, Madrid, Spain (COMPSAC), 669–676.
- Sulistyaningsih, Putranto, P., Daud, P., and Desvasari, W. (2018). "Fast fourier transform (FFT) data sampling using hamming and blackman method for radar," in Proceedings of the 2018 International Conference on Electrical Engineering and Computer Science, Pangkal, Indonesia (ICECOS), 183–188.
- Surek, G. A. S., Seman, L. O., Stefenon, S. F., Mariani, V. C., and Coelho, L. d.S. (2023). Video-based human activity recognition using deep learning approaches. *Sensors* 23, 6384. doi:10.3390/s23146384
- Tian, Y., Wang, X., Chen, L., and Liu, Z. (2019). Wearable sensor-based human activity recognition via two-layer diversity-enhanced multiclassifier recognition method. *Sensors* 19 (9), 2039. doi:10.3390/s19092039
- Vaizman, Y., Ellis, K., and Lanckriet, G. (2017). Recognizing detailed human context in the wild from smartphones and smartwatches. *IEEE Pervasive Comput.* 16 (4), 62–74. doi:10.1109/mprv.2017.3971131
- Vaizman, Y., Weibel, N., and Lanckriet, G. (2018). Context recognition in-the-wild: unified model for multi-modal sensors and multi-label classification. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 1–22. doi:10.1145/3161192
- Vanijkachorn, P., and Visutak, P. (2021). A deep convolutional LSTM for ADLs classification of the elderly. *Proc. ICDABI*, 2021. doi:10.1109/ICDABI53623.2021.9655856
- Vrskova, R., Kamenecy, P., Hudec, R., and Sykora, P. (2023). A new deep-learning method for human activity recognition. *Sensors* 23, 2816. doi:10.3390/s23052816
- Wang, F., Wang, H., Zhou, X., and Fu, R. (2022). A driving fatigue feature detection method based on multifractal theory. *IEEE Sensors J.* 22 (19), 19046–19059. doi:10.1109/jsen.2022.3201015
- Wang, L., Sun, J., Pan, T., Ye, Y., He, W., and Yang, K. (2021). "Personalized human activity recognition using hypergraph learning with fusion features," in Proceedings of the 2021 IEEE 4th International Conference on Electronics Technology, Chengdu, China (ICET), 1251–1255.
- Wen, C., Huang, Y., and Davidson, T. N. (2023b). Efficient transceiver design for MIMO dual-function radar-communication systems. *IEEE Trans. Signal Process.* 71, 1786–1801. doi:10.1109/tsp.2023.3275274
- Wen, C., Huang, Y., Zheng, L., Liu, W., and Davidson, T. N. (2023a). Transmit waveform design for dual-function radar-communication systems via hybrid linear-nonlinear precoding. *IEEE Trans. Signal Process.* 71, 2130–2145. doi:10.1109/tsp.2023.3278858
- Xu, H., Liu, J., Hu, H., and Zhang, Y. (2016). Wearable sensor-based human activity recognition method with multi-features extracted from hilbert-huang transform. *Sensors* 16, 2048. doi:10.3390/s16122048
- Xu, J., Guo, K., and Sun, P. Z. H. (2022b). Driving performance under violations of traffic rules: novice vs. Experienced drivers. *IEEE Trans. Intelligent Veh.* 7, 908–917. doi:10.1109/TIV.2022.3200592
- Xu, J., Guo, K., Zhang, X., and Sun, P. Z. H. (2023). Left gaze bias between lht and rht: a recommendation strategy to mitigate human errors in left- and right-hand driving. *IEEE Trans. Intelligent Veh.* 8, 4406–4417. doi:10.1109/TIV.2023.3298481
- Xu, J., Zhang, X., Park, S. H., and Guo, K. (2022a). The alleviation of perceptual blindness during driving in urban areas guided by saccades recommendation. *IEEE Trans. Intelligent Transp. Syst.* 23, 16386–16396. doi:10.1109/TITS.2022.3149994
- Xue, T., and Liu, H. (2021). "Hidden Markov Model and its application in human activity recognition and fall detection: a review," in Proceedings of the International Conference in Communications, 863–869. Signal Processing, and Systems.

- Yan, L., Shi, Y., Wei, M., and Wu, Y. (2023). Multi-feature fusing local directional ternary pattern for facial expressions signal recognition based on video communication system. *Alexandria Eng. J.* 63, 307–320. doi:10.1016/j.aej.2022.08.003
- Yao, Y., Shu, F., Li, Z., Cheng, X., and Wu, L. (2023). Secure transmission scheme based on joint radar and communication in mobile vehicular networks. *IEEE Trans. Intelligent Transp. Syst.* 24, 10027–10037. doi:10.1109/tits.2023.3271452
- Yin, J., Han, J., Xie, R., Wang, C., Duan, X., Rong, Y., et al. (2021). MC-LSTM: real-time 3D human action detection system for intelligent healthcare applications. *IEEE Trans. Biomed. Circuits Syst.* 15, 259–269. doi:10.1109/TBCAS.2021.3064841
- Yoon, J., and Kim, S. (2022). Practical and accurate indoor localization system using deep learning. *Sensors* 22, 6764. doi:10.3390/s22186764
- Yu, J., Lu, L., Chen, Y., Zhu, Y., and Kong, L. (2021). An indirect eavesdropping attack of keystrokes on touch screen through acoustic sensing. *IEEE Trans. Mob. Comput.* 20 (2), 337–351. doi:10.1109/TMC.2019.2947468
- Yu, J., Xian, S., Zhang, Z., Hou, X., He, J., Mu, J., et al. (2023). Synergistic piezoelectricity enhanced BaTiO₃/polyacrylonitrile elastomer-based highly sensitive pressure sensor for intelligent sensing and posture recognition applications. *Nano Res.* 16 (4), 5490–5502. doi:10.1007/s12274-022-5084-x
- Zhang, J., Tang, Y., Wang, H., and Xu, K. (2022a). ASRO-DIO: active subspace random optimization based depth inertial odometry. *IEEE Trans. Robotics* 39, 1496–1508. doi:10.1109/TRO.2022.3208503
- Zhang, J., Ye, G., Tu, Z., Qin, Y., Qin, Q., Zhang, J., et al. (2022b). A spatial attentive and temporal dilated (SATD) GCN for skeleton-based action recognition. *CAAI Trans. Intelligent Technol.* 7 (1), 46–55. doi:10.1049/cit2.12012
- Zhang, J., Zhu, C., Zheng, L., and Xu, K. (2021). ROSEFusion: random optimization for online dense reconstruction under fast camera motion. *ACM Trans. Graph.* 40 (4), 1–17. doi:10.1145/3450626.3459676
- Zhang, X., Huang, D., Li, H., Zhang, Y., Xia, Y., and Liu, J. (2023). *Self-training maximum classifier discrepancy for EEG emotion recognition*. CAAI Transactions on Intelligent Technology.
- Zhang, X., and Jiang, S. (2021). “Application of fourier transform and Butterworth filter in signal denoising,” in Proceedings of the 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP), China (Xi’an), 1277–1281.
- Zhang, Z. (2012). Microsoft Kinect sensor and its effect. *IEEE Multimed.* 19, 4–10. doi:10.1109/mmul.2012.24
- Zhao, K., Jia, Z., Jia, F., and Shao, H. (2023). Multi-scale integrated deep self-attention network for predicting remaining useful life of aero-engine. *Eng. Appl. Artif. Intell.* 120, 105860. doi:10.1016/j.engappai.2023.105860
- Zhao, Z., Xu, G., Zhang, N., and Zhang, Q. (2022). Performance analysis of the hybrid satellite-terrestrial relay network with opportunistic scheduling over generalized fading channels. *IEEE Trans. Veh. Technol.* 71 (3), 2914–2924. doi:10.1109/tvt.2021.3139885
- Zheng, C., An, Y., Wang, Z., Qin, X., Eynard, B., Bricogne, M., et al. (2023c). Knowledge-based engineering approach for defining robotic manufacturing system architectures. *Int. J. Prod. Res.* 61 (5), 1436–1454. doi:10.1080/00207543.2022.2037025
- Zheng, M., Zhi, K., Zeng, J., Tian, C., and You, L. (2022). A hybrid CNN for image denoising. *J. Artif. Intell. Technol.* 2 (3), 93–99. doi:10.37965/jait.2022.0101
- Zheng, W., Liu, M., Liu, C., Wang, D., and Li, K. (2023b). Recent advances in sensor technology for healthcare and biomedical applications (volume II). *Sensors* 23, 5949. doi:10.3390/s23135949
- Zheng, W., Yang, Y., Liu, C., and Zhou, W. (2023a). Recent advancements in sensor technologies for healthcare and biomedical applications. *Sensors* 23, 3218. doi:10.3390/s23063218
- Zheng, Y., Lv, X., Qian, L., and Liu, X. (2022). An optimal BP neural network track prediction method based on a GA-ACO hybrid algorithm. *J. Mar. Sci. Eng.* 10 (10), 1399. doi:10.3390/jmse10101399
- Zhou, X., and Zhang, L. (2022). SA-FPN: an effective feature pyramid network for crowded human detection. *Appl. Intell.* 52 (11), 12556–12568. doi:10.1007/s10489-021-03121-8
- Zhu, T., Ding, H., Wang, C., Liu, Y., Xiao, S., Yang, G., et al. (2023). Parameters calibration of the GISSMO failure model for SUS301L-MT. *Chin. J. Mech. Eng.* 36 (1), 20. doi:10.1186/s10033-023-00844-2