



## Article

# Biosensor-Based Multimodal Deep Human Locomotion Decoding via Internet of Healthcare Things

Madiha Javeed <sup>1,\*</sup> , Maha Abdelhaq <sup>2,\*</sup>, Asaad Algarni <sup>3</sup> and Ahmad Jalal <sup>1</sup><sup>1</sup> Department of Computer Science, Air University, Islamabad 44000, Pakistan; ahmadjalal@mail.au.edu.pk<sup>2</sup> Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia<sup>3</sup> Department of Computer Sciences, Faculty of Computing and Information Technology, Northern Border University, Rafha 91911, Saudi Arabia; asaad.algarni@nbu.edu.sa

\* Correspondence: 191880@students.au.edu.pk or mjaved84@gmail.com (M.J.); msabdelhaq@pnu.edu.sa (M.A.)

**Abstract:** Multiple Internet of Healthcare Things (IoHT)-based devices have been utilized as sensing methodologies for human locomotion decoding to aid in applications related to e-healthcare. Different measurement conditions affect the daily routine monitoring, including the sensor type, wearing style, data retrieval method, and processing model. Currently, several models are present in this domain that include a variety of techniques for pre-processing, descriptor extraction, and reduction, along with the classification of data captured from multiple sensors. However, such models consisting of multiple subject-based data using different techniques may degrade the accuracy rate of locomotion decoding. Therefore, this study proposes a deep neural network model that not only applies the state-of-the-art Quaternion-based filtration technique for motion and ambient data along with background subtraction and skeleton modeling for video-based data, but also learns important descriptors from novel graph-based representations and Gaussian Markov random-field mechanisms. Due to the non-linear nature of data, these descriptors are further utilized to extract the codebook via the Gaussian mixture regression model. Furthermore, the codebook is provided to the recurrent neural network to classify the activities for the locomotion-decoding system. We show the validity of the proposed model across two publicly available data sampling strategies, namely, the HWU-USP and LARa datasets. The proposed model is significantly improved over previous systems, as it achieved 82.22% and 82.50% for the HWU-USP and LARa datasets, respectively. The proposed IoHT-based locomotion-decoding model is useful for unobtrusive human activity recognition over extended periods in e-healthcare facilities.

**Keywords:** human activity recognition; Internet of Healthcare Things; locomotion prediction; multimodal systems; recurrent neural network; RGB; wearable sensors



**Citation:** Javeed, M.; Abdelhaq, M.; Algarni, A.; Jalal, A. Biosensor-Based Multimodal Deep Human Locomotion Decoding via Internet of Healthcare Things. *Micromachines* **2023**, *14*, 2204. <https://doi.org/10.3390/mi14122204>

Academic Editors: Lei Jing,  
Yoshinori Matsumoto and  
Zhan Zhang

Received: 30 October 2023  
Revised: 28 November 2023  
Accepted: 30 November 2023  
Published: 3 December 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Recent trends in the Internet of Healthcare Things (IoHT) have boosted wearable and visual-technology-based human locomotion decoding. This boost converts the healthcare industry from cure to prevention [1–4]. Various IoHT devices are available for healthcare and research, including smart devices, inertial units, and cameras. Data from such IoHT devices have been extracted, processed, and analyzed for human locomotion decoding. For ambient assisted living, sensor-based data have been used to support and supervise people, also known as human activity recognition (HAR) [5–7]. Applications of such HAR systems include injury recognition, medical analysis, long-term or short-term care, health monitoring, and independent quality of life [8–12].

These HAR systems can use machine learning or deep learning techniques to decode the activities of daily living by extracting data from motion, ambient, or vision-based sensors [13–16]. Modern smart devices manipulate the data and thus cannot be utilized

for locomotion decoding [17–20]. Some HAR systems have less efficiency due to errors induced by the data acquisition that must be resolved using a robust filter [21–23]. Existing feature extraction methods cannot perform well for HAR systems and provide less efficient results [24–28]. Therefore, a multimodal sensor-based human-locomotion-decoding (HLD) system consisting of motion, ambient, and vision sensors is proposed in this paper. The key contributions of this research are as follows:

- An innovative multimodal system for locomotion decoding via multiple sensors fused to enhance the HAR performance [29–31];
- The effective and novel filtration of the inertial sensor data [32–34] by using a proposed state-of-the-art Quaternion-based filter;
- A novel approach to filtering the ambient-based data that includes infrared cameras and switches attached to the environment;
- Hand-crafted contemporary descriptor extraction methods [35–38] are proposed and applied to acquire related descriptors [39–42] using novel techniques;
- Efficient ambient sensor descriptor extraction based on a unique and novel graph representation;
- The proficient recognition of activities [43–46] for locomotion decoding via detection through a recurrent neural network (RNN).

Section 2 explains the sensor-based activity recognition systems presented in the literature. Next, Section 3 details the proposed locomotion-decoding system for the IoHT industry [47–51]. The experiments performed over the selected datasets using the proposed method and their results, along with a comparison of the baseline system and previous state-of-the-art models, are discussed in Section 4. The conclusion of the whole paper is presented in Section 5.

## 2. Related Work

Locomotion decoding with a combination of IoHT-based sensors can be utilized for different applications [52–55], including the execution and tagging of data, which associates the meanings of sensor data interpretations by using symbols. A single sensor is not enough to provide the semantic meaning of a situation. Therefore, multimodal sensor-based systems serve this purpose. For this resolution, multiple systems have been proposed in history to evaluate the effectiveness, completeness, and reliability of such sensor-based decoding systems.

### 2.1. Sensor-Based Locomotion Decoding

In [56], Franco et al. propose a multimodal system for locomotor activity recognition. They used RGB video and other sensors for data acquisition. Histograms of oriented gradient (HOG) descriptors and skeleton-based information were extracted from the RGB data frames to capture the most prominent body postures. For the activity classification, a voting system was defined to obtain votes from support vector machine (SVM) and random forest classifiers. However, the proposed system could not achieve higher results due to the absence of a filtration technique for the data. Another system is proposed in [57] that collects motion sensor data. Next, data are processed using a linear interpolation filter and segmentation. Features are extracted using four different extraction techniques and normalized using the z-score function. Then, features are selected via correlation and evolutionary search algorithms. Further, the class imbalance is removed using the synthetic minority over-sampling method. Features are fused, and multi-view stacking is utilized to classify humans.

### 2.2. Multimodal Locomotion Decoding

In [58], the authors propose a robust human activity recognition method. They used multimodal data based on wearable inertial and RGB-D sensors. The inertial data were pre-processed using magnitude computation and noise removal techniques, and dense HOGs were extracted from video data. Time domain features are extracted from inertial

signals, and bag-of-words encoding is utilized for video frame sequences. Furthermore, the features are fused, and K-nearest-neighbor and support vector machines are used for the human activity classification.

A long short-term memory (LSTM) network-based system is proposed in [59]. To recognize activities of daily living, the authors used a deep learning model via data acquired from real-world and synthetic environments. The sensors were attached to the wrists, ankles, and waist to detect activities, including eating and driving. Each sensor's accuracy was observed to elaborate the custom weights for each sensor fusion. This study recommended using one sensor on the upper body parts and one sensor on the lower body parts to obtain reasonable results. However, due to the restricted data used and limited weight learning in the system, the method cannot adapt to changes over time.

In [60], a system of Marfusion based on a convolutional neural network (CNN) and attention mechanism is proposed. Features are extracted from multimodal sensors and a set of CNNs is utilized for each sensor. Next, a dot-product, scaled, self-attention process is applied to give weight to each sensor. Then, CNN and attention-based modules are utilized for feature fusion with different parameters. Further, fully connected batch normalization, dropout, ReLU, and softmax layers are used for the classification via the obtainment of the probabilities for different activities. The proposed system gave an acceptable performance but experimented with limited human locomotion. Therefore, the results are not robust for real-time environments.

### 3. Materials and Methods

The proposed locomotion-decoding architecture is described in Figure 1. The input data for the proposed IoHT-based system were taken from two publicly available datasets named Logistic Activity Recognition Challenge (LARA) [61] and Heriot-Watt University/University of Sao Paulo (HWU-USP) [62], which are present in the form of time series in a time segment of size  $W$  from  $S$  sensors. Sensors of three types were used: physical signals  $\{p_i\}$ , ambient signals  $\{p_a\}$ , and visual frame sequences  $\{p_v\}$ . Algorithm 1 demonstrates the complete IoHT-based HLD system. The input  $\{p_i, p_a, p_v\}$  from the  $S$  sensors was pre-processed for each time segment of a  $W$  size. Next, the descriptors were extracted and optimized  $\{V_i^*, K_i^*, S_i^*, A_i^*\}$  for each  $W$  segment. Further, the descriptors were trained by using an RNN and tested the remaining descriptors to recognize activities  $\{A^*\}$  to decode human locomotion. All these phases of the IoHT-based HLD system are further explained in the next subsections.

---

#### Algorithm 1 HLD Algorithm

---

**Input:** physical IMU signals  $\{p_i\}$ , ambient signals  $\{p_a\}$ , visual frame sequences  $\{p_v\}$ ;

**Output:** recognized activities  $\{A^*\}$ ;

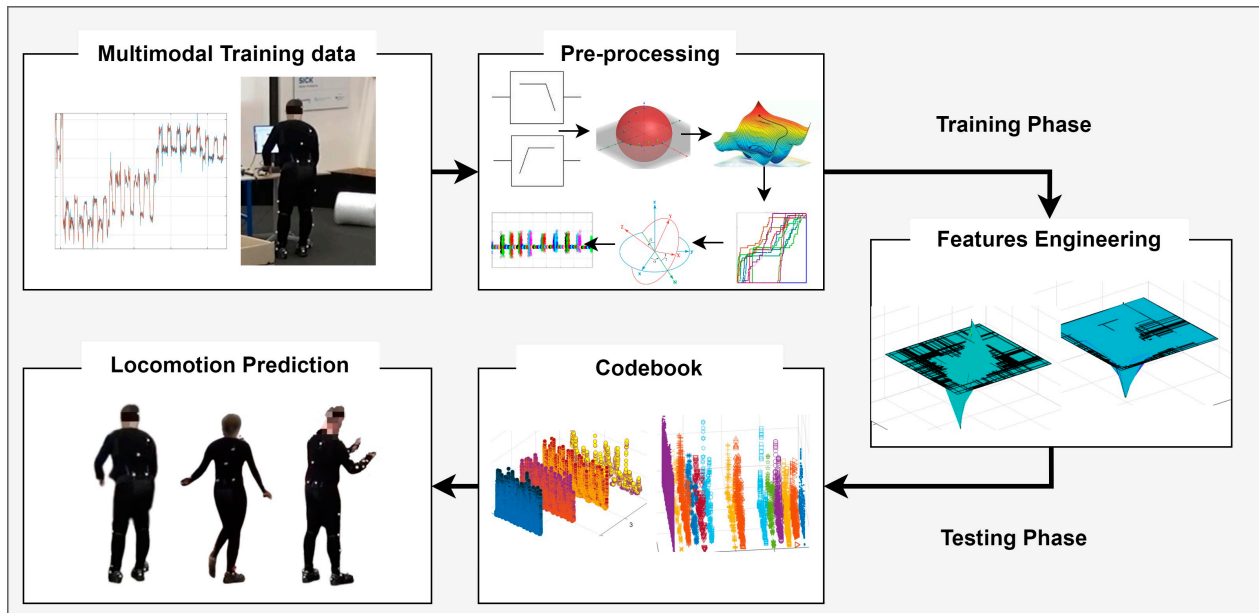
1. Pre-process  $\{p_i, p_a, p_v\}$  for each segment  $W$  in *Module I*;
  2. Extract descriptors  $\{V_i^*, K_i^*, S_i^*, A_i^*\}$  for  $W$  in *Module II*;
  3. Optimize descriptors for  $W$  in *Module III*;
  4. Train descriptors over classifier to obtain  $f(X, \theta)$ ;
  5. Test remaining descriptors to obtain  $\{\theta, \theta^*\}$ ;
  6. Recognize activities  $\{A^*\}$ ;
- 

#### 3.1. Pre-Processing Motion and Ambient Data

A novel quaternion-based filter is proposed in this study to pre-process the physical-motion [63] and ambient data from the sensor inertial measurement units (IMUs). The signals are clarified via low- and high-pass Butterworth filters [64,65] for further processing. Next, the signals are normalized using the Euclidean distance [66,67]:

$$Norm = \sqrt{LPF_1 + LPF_2 + LPF_3} + \sqrt{HPF_1 + HPF_2 + HPF_3} \quad (1)$$

where  $LPF_1$ ,  $LPF_2$ , and  $LPF_3$  denote the filtered values for the x-, y-, and z-axes via the Butterworth filter, respectively.  $HPF_1$ ,  $HPF_2$ , and  $HPF_3$  represent the filtered values of the x-, y-, and z-axes through the Butterworth filter, respectively.



**Figure 1.** Architecture diagram of proposed IoHT-based human-locomotion-decoding system.

Then, for the accelerometer signals, gravity from a stationary activity, such as lying down, is extracted as the minimum gravity ( $g_m$ ) and average gravity ( $g_a$ ). Then, the gravitational error ( $g_e$ ) [68,69] is removed from the accelerometer signals, giving more accurate and error-free signals for further processing. Similarly, the earth's magnetic field is used to remove the magnetic errors from magnetometer signals [70,71].

After normalization, discrete wavelet transform [72] is applied to the gyroscope signals to transform them into quaternions in order to avoid the gimbal lock problem. Later, the derivative of the quaternions is considered, and gradient descent is applied to attain the minimum rate of change. Further, a local minimum [73] is selected, and the gyroscope signals are normalized using the Euler angles:

$$Axz = \text{atan2}(z, x), \quad (2)$$

$$Ayz = \text{atan2}(z, y), \quad (3)$$

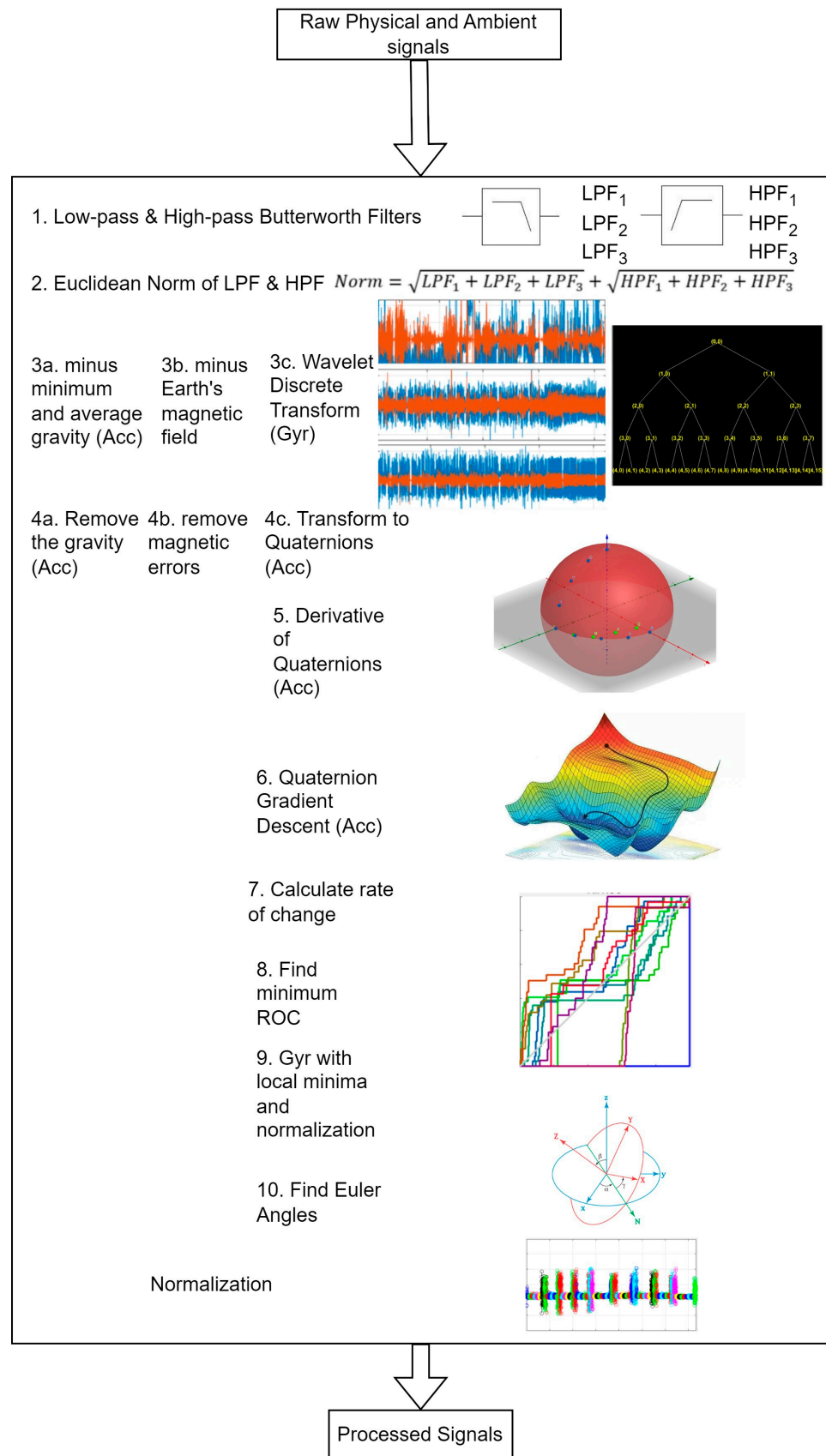
$$Axy = \text{atan2}(y, x), \quad (4)$$

where  $Axz$ ,  $Ayz$ , and  $Axy$  are the Euler angles. Lastly, all three pre-processed signals are normalized together. Figure 2 explains the pre-processing step for the physical-motion module in detail.

### 3.2. Pre-Processing Visual Data

For the pre-processing, videos from both datasets were converted into frame sequences. A delta of 50 was chosen to restrict the number of pre-processing sequences to avoid redundant data processing. Next, we retrieved a background image from both data sequences. Then, the background was removed by subtracting the background image from the original frame sequences [74,75]. The background subtraction from the original image sequence is displayed in Figure 3. Discrete wavelet transform was used over the frame sequences to reduce the noise present.



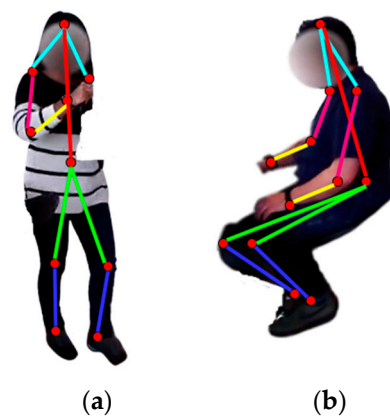


**Figure 2.** Pre-processing module proposed for physical-motion and ambient data.



**Figure 3.** (a) Before background deduction and (b) after background deduction of a frame sequence from HWU-USP dataset.

Skeleton modeling was performed through blob and centroid techniques for human detection in the frame sequences. First, the blobs were defined from the human movable parts, which was followed by taking the centroids and deciding on five types of skeleton body points—head, shoulders, elbows, wrists, torso, knees, and ankles [76]. Figure 4 shows the skeleton points extracted for drinking tea and reading a newspaper.



**Figure 4.** Skeleton point decoding from frame sequences of (a) drinking tea and (b) reading a newspaper.

### 3.3. Data Segmentation

Next, to deal with the dimensions of the datasets, this study segmented the motion and ambient pre-processed data into overlapped [77] and time-based [78] segments, whereas the vision-based data were segmented through event-based segments. For all three types of data  $\{p_i^*, p_a^*, p_v^*\}$ , Figure 5 shows the segmentation process by using manifold locomotion activities.

### 3.4. Motion Descriptor Extraction

The pre-processed and segmented motion-based data were further provided to two different techniques for the descriptor extraction, including Gaussian Markov random field (GMRF) and a novel contribution in the form of a multisynchrosqueezing transform (MSST)-based spatial-temporal graph.

GMRF can take multidimensional data, and a stochastic process becomes Gaussian when all its distributions are Gaussian-normalized [79]. Equations (5) and (6) show the expectation function ( $\tilde{\mu}_t$ ) and covariance function ( $\tilde{\Sigma}_{s,t}$ ) using  $s$  samples and  $t$  times. Figure 6 presents the results for the GMRF for a window of kinematic physical data on HWU-USP.

$$\tilde{\mu}_t = E\tilde{X}_t, \quad (5)$$

$$\widetilde{\Sigma}_{s,t} = \text{cov}(\widetilde{X}_s, \widetilde{X}_t). \quad (6)$$

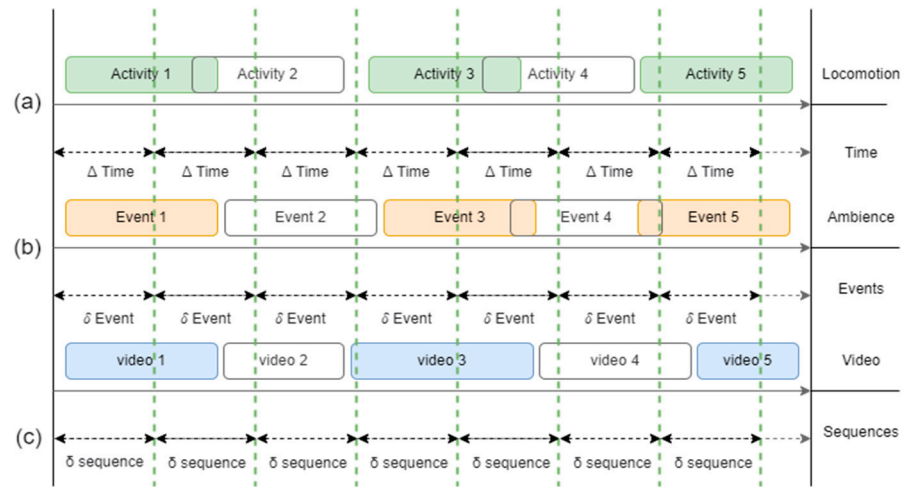


Figure 5. Segmentation for (a) motion, (b) ambient, and (c) visual data.

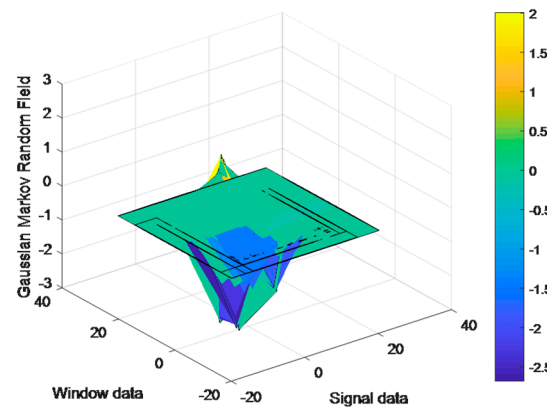


Figure 6. Results of GMRF application on HWU-USP dataset.

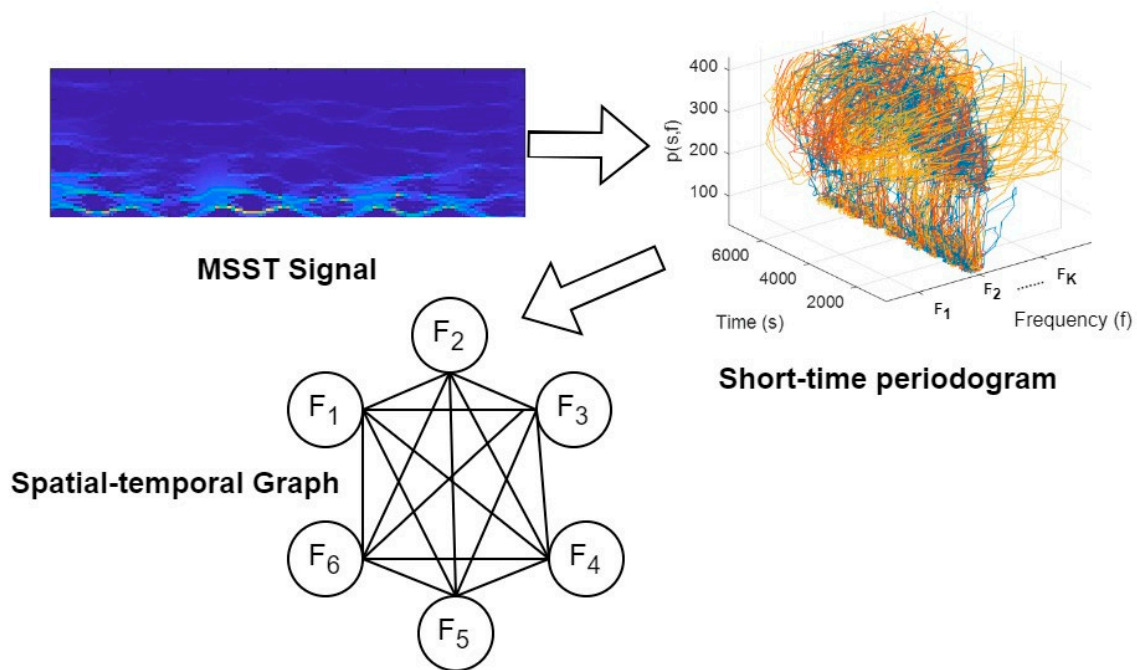
MSST represents multiple synchrosqueezing transforms iteratively [80] and is calculated as follows:

$$Ts^{[M]}(t, \gamma) = \int_{-\infty}^{+\infty} Ts^{M-1}(t, \gamma) \delta(\gamma - \hat{\omega}(t, \omega)) d\omega, \quad (7)$$

where  $M$  gives the iteration number  $\leq 2$  and  $Ts^{[M]}(t, \gamma)$  is the spread time–frequency coefficient. The short-time periodogram is further calculated as follows:

$$p(s, f) = \frac{1}{T} \left| Y(s, f) \right|^2 \quad (8)$$

where  $p(s, f)$  is the result of frequency ( $f$ ) and time ( $s$ ).  $T$  shows the window length. Further, the spatial–temporal graph was constructed using six nodes or frequencies. Figure 7 shows the novel spatial–temporal graph extracted from MSST for a random static pattern.



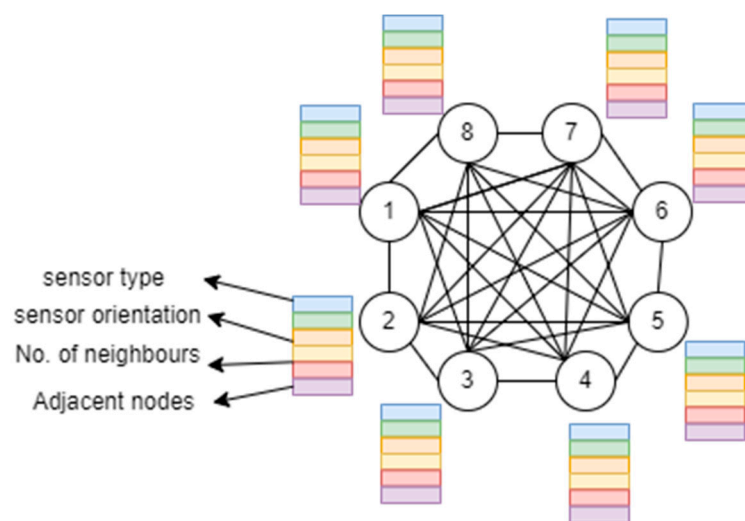
**Figure 7.** Process of constructing novel spatial-temporal graph from MSST.

### 3.5. Ambient Descriptor Extraction

A graph-based representation has been proposed as a novel descriptor extraction for ambient sensor pre-processing [81]. For each sensor attached to the ambient, a graph ( $R$ ) is produced using a descriptor matrix ( $M$ ) and adjacency matrix ( $K$ ) given by the following:

$$R = (M, K) \quad (9)$$

where  $M$  is the descriptor matrix consisting of the sensor type, number of neighbors, and sensor orientation.  $K$  contains the number of adjacent sensors for each node and the names of neighboring sensors. Figure 8 presents the details of the proposed graph-based ambient descriptors.

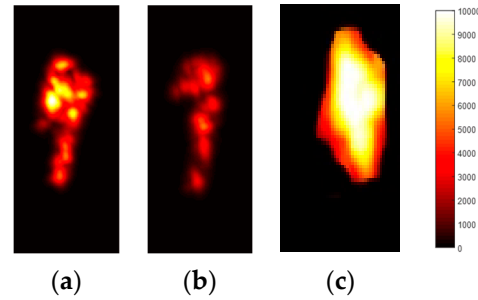


**Figure 8.** Proposed novel graph-based ambient feature extraction.

### 3.6. Vision Descriptor Extraction

In thermal descriptors, the movement from one frame to another is captured in the form of thermal maps. More movement is described using higher heat values in yellow, and less movement is shown using red or black [82]. In Equation (10),  $x$  represents a one-dimensional vector comprising the extracted values,  $i$  represents the index value, and  $R$  denotes the RGB value. Figure 9 presents the heat map for the full-body frame sequence.

$$TM(x) = \sum_{i=0}^k \ln R(i). \quad (10)$$

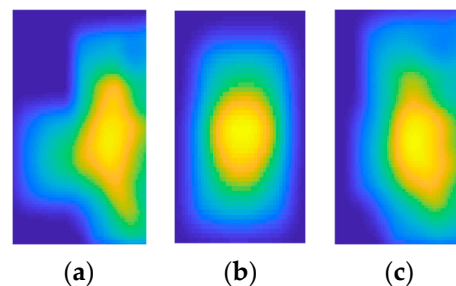


**Figure 9.** Thermal heat map extracted for activities including (a) drinking tea, (b) opening a drawer, and (c) reading a newspaper.

The full-body descriptor extraction method for visual data utilized is called the saliency map (SM) approach. It is computationally expensive to process an entire frame simultaneously; therefore, the SM approach suggests sequentially looking at or fixating on the salient locations of a frame. The fixated region is analyzed, and then attention is redirected to other salient regions using saccade movements requiring more focus [83]. The SM approach is a successful and biologically plausible technique for modeling visual attention. The generalized Gaussian distribution shown in Equation (11) is used to model each of these:

$$P(f_i) = \frac{\theta_i}{2\sigma_i\gamma(\theta_i^{-1})} \exp\left(-\left|\frac{f_i}{\sigma_i}\right|^{\theta_i}\right), \quad (11)$$

where  $\theta_i > 0$  is the shape parameter,  $\sigma_i > 0$  provides the scale parameter, and  $\gamma$  gives the gamma function. Figure 10 presents the results of SMs applied over a full-body frame sequence.



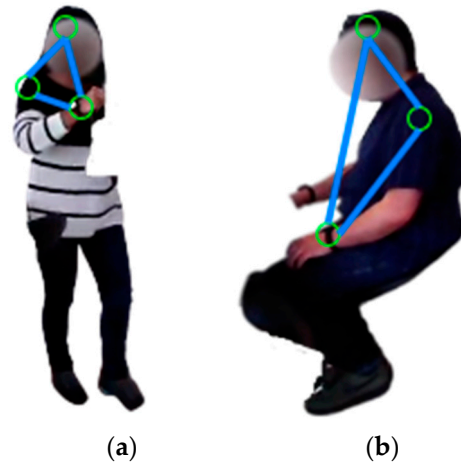
**Figure 10.** Results of saliency maps applied over full-body frame sequences for (a) drinking tea, (b) opening a drawer, and (c) reading a newspaper.

The orientation descriptor technique is the first descriptor extraction technique for the skeleton body points. Five skeleton body points are used to make triangles and obtain angles from them. The tangent angle in Equation (12) is measured between the three sides of each triangle [84]:

$$\tan \theta = \frac{u \cdot v}{|u||v|}, \quad (12)$$



where  $u \cdot v$  is the dot product of vectors  $u$  and  $v$  that are any two sides of a triangle. Figure 11 demonstrates the examples of triangles formed by combining two human skeleton body points in some activities, such as drinking tea and reading a newspaper.

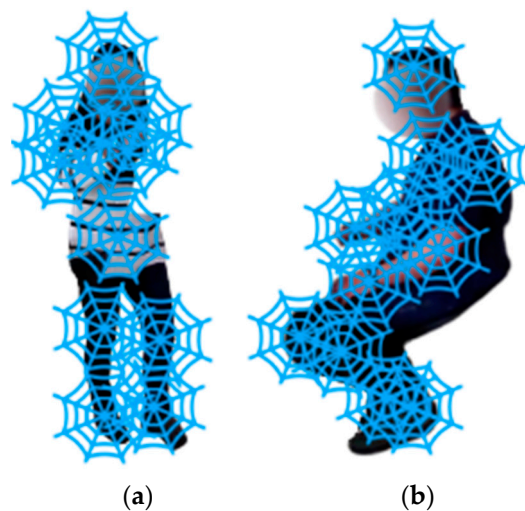


**Figure 11.** The triangular shape is formed by combining human skeleton body points for (a) drinking tea and (b) reading a newspaper.

The second descriptor extraction technique used for the skeleton body points is the spider local image feature (SLIF) technique. A spiderweb representation emulates the skeleton body point nodes as web intersection points in a frame sequence [85,86]. The position of each node  $(n, m)$  is denoted by a set of two-dimensional coordinates, as follows:

$$x_{n,m} = \left( \frac{m \cdot \cos\left(\frac{2\pi n}{N}\right)}{M}, \frac{m \cdot \sin\left(\frac{2\pi n}{N}\right)}{M} \right), \quad (13)$$

where the first and second terms represent the horizontal and vertical coordinates, respectively. For a set of previously defined skeleton body points, the SLIFs are extracted by selectively extracting pixel information from around the neighborhood of each point and applying a spiderweb over the point. Figure 12 shows a spiderweb applied over two sample frame sequences.



**Figure 12.** Spiderweb applied for (a) drinking tea and (b) reading a newspaper over HWU-USP dataset.

### 3.7. Codebook Generation

A Gaussian mixture model (GMM) codebook is used to encode the descriptors extracted from previous subsections. An expectation maximization (EM) algorithm is used in the GMM to present complex descriptors. This algorithm approximates the parameter set ( $\Theta$ ) and aids in calculating the maximum likelihood through an initial parameter set ( $\Theta_1$ ), and then continuously applies the  $E$  and  $M$  steps. Then, it produces  $\{\Theta_1, \Theta_2, \dots, \Theta_m, \dots\}$  and both  $E$  and  $M$  steps as follows:

$$\gamma^m(z_k^j | x_j, \Theta^m) = \frac{\omega_k^m f(x_j | \mu_k^m, \Sigma_k^m)}{\sum_{i=1}^K \omega_i^m f(x_j | \mu_i^m, \Sigma_i^m)}, \quad (14)$$

$$\sum_k^{m-1} = \frac{\sum_{j=1}^N \gamma^m(z_k^j | x_j, \Theta^m) (x_j - \mu_k^{m+1})(x_j - \mu_k^{m+1})^T}{\sum_{j=1}^N \gamma^m(z_k^j | x_j, \Theta^m)}. \quad (15)$$

where  $\gamma^m(z_k^j | x_j, \Theta^m)$  gives the probability of the  $j$ th example and the  $k$ th Gaussian at the  $m$ th iteration with weights ( $\omega_k^m$ ), means ( $\mu_k^m$ ), and covariance ( $\Sigma_k^m$ ) values. Similarly, a single generalized signal is extracted from the set of descriptors given using Gaussian mixture regression (GMR). Henceforth, a smooth signal via regression can be taken out by coding the temporal signal features [87] through a mixture of Gaussians. Each vector of the signals' GMM is taken as the input (xI) and output (xO) using GMR via this method.

### 3.8. Locomotion Decoding

A simple feedforward neural network poorly handles the sequence of data. It never forms a cycle between two hidden layers, and information always flows in one direction, never going back. It comprises an input layer, a hidden layer, and an output layer. An RNN [88] also contains these three layers, but it focuses on considering the current state along with the previous state in the form of output from the previously hidden layer via memory. Thus, the current state and previous state are used to produce output for the next time step, as shown in Figure 13. An activation function is also used to calculate the current state; we used  $\tanh$  as the activation function. Due to the input pattern change, the RNN performs better by incorporating backpropagation.

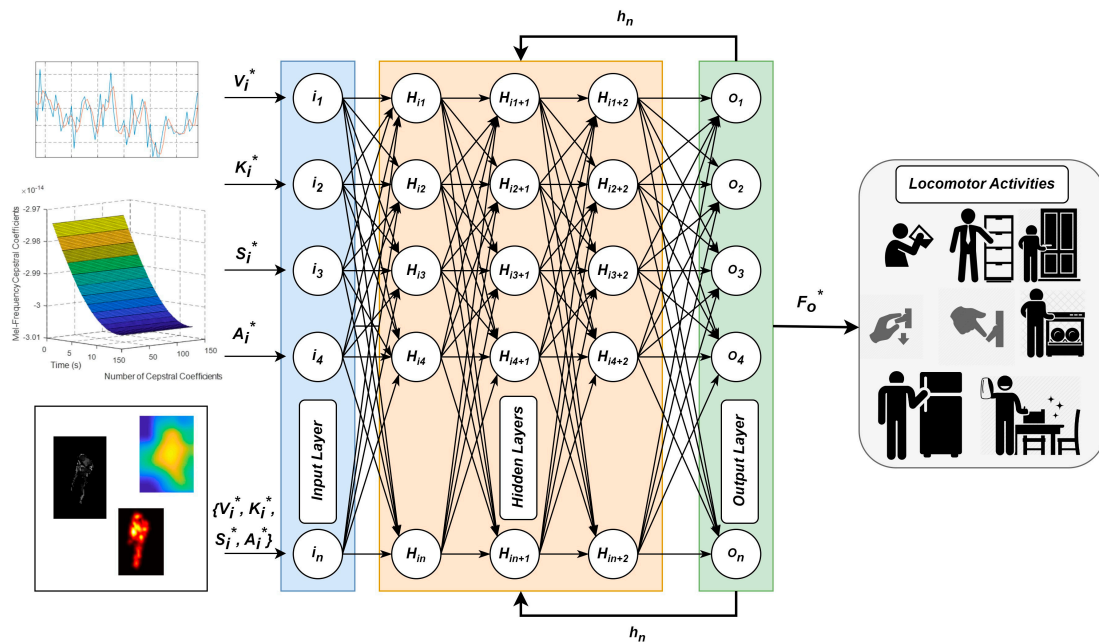


Figure 13. RNN incorporated into the proposed IoHT-based HLD system.

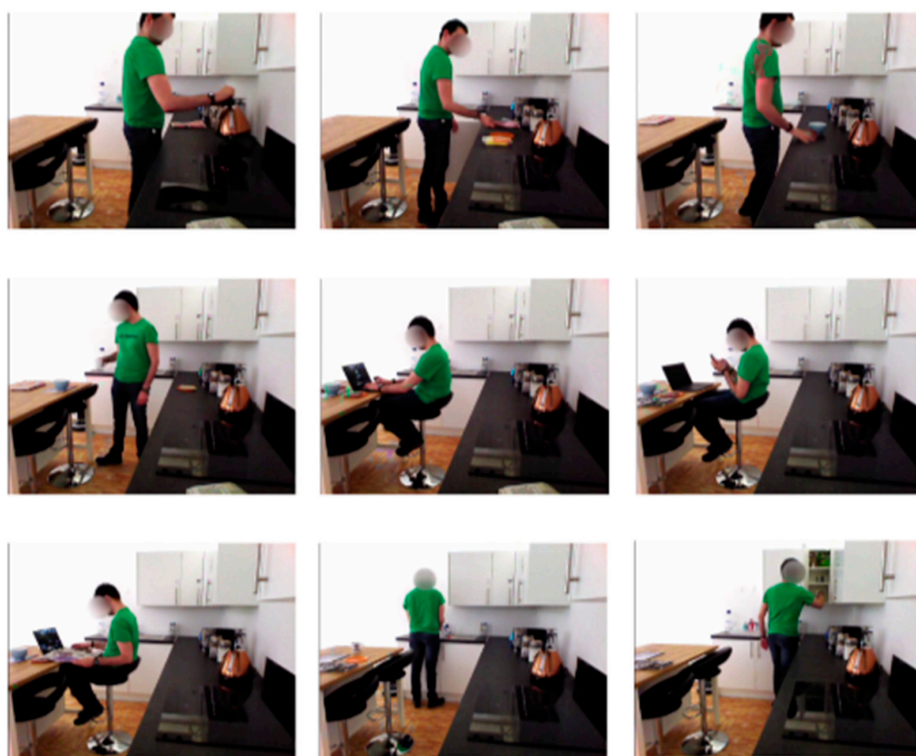
#### 4. Performance Evaluation

In order to evaluate the IoHT-based HLD system, the following datasets and evaluation criteria were used.

##### 4.1. Dataset Descriptions

Several publicly available datasets are present for human locomotion decoding via activity recognition. However, they can be different in terms of the number of subjects, number of activities performed, environmental setup, number of sensors, type of sensors, and sampling rates. In the proposed IoHT-based HLD system, we used two publicly available datasets, HWU-USP [62] and LARa [61], captured in diverse environmental setups and three different sensor modalities to make the system more robust. A 10-fold cross-validation technique was utilized to evaluate the proposed system. The following sections give details on the datasets mentioned above:

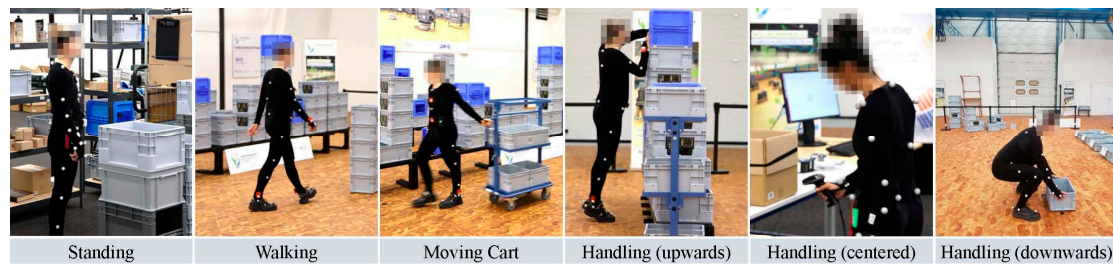
**HWU-USP:** A dataset recorded in a “living-lab” was selected for this study. It contains recordings from binary switches, PIR sensors, RGBD cameras installed over a robot, and IMU devices. The camera color is VGA  $640 \times 480$  at 25 fps. A total of 16 participants performed nine activities with 144 instances with an average length of 48 s [62]. The participants were voluntary and healthy with neither functional nor visual impairments. The dataset contains activities of daily living with either periodical patterns or long-term dependencies and, hence, it is different from other multimodal environments. A variety of activities have been performed, such as *making a cup of tea*, *making a sandwich*, *making a bowl of cereal*, *setting the table*, *using a laptop*, *using a phone*, *reading a newspaper*, *cleaning the dishes*, and *tidying the kitchen*. Figure 14 represents the sample of activities performed by one of the participants in the HWU-USP dataset;



**Figure 14.** Activities performed by subject in HWU-USP dataset [61].

**LARa:** This dataset consists of an OmoCap system, a VICON system of 38 infrared cameras, three sets of IMU devices, and 30 recordings of 2 min for each of the 14 subjects. A wide range of participants were selected, including both male and female, ranging in age from 22 to 59 years, weighing from 48 to 100 lbs, left- and right-handed, and with heights from 163 to 185 cm. The dataset was recorded in a total of seven sessions of 758 min

of recording. Acceleration sensors recorded the locomotion at a rate of 100 Hz [61]. The dataset is unbalanced regarding the annotations due to the complex process. The dataset is based on the activities performed in a logistics-based context. An expert trained the subjects in advance to recordings. A total of eight activities were recorded for each subject, including standing, walking, carting, handling (upwards), handling (centered), handling (downwards), synchronization, and none. Figure 15 gives a few sample frame sequences from the dataset.



**Figure 15.** Sample frame sequences from the LARa dataset [62].

#### 4.2. Experiment 1: Evaluation Protocol

Evaluation metrics can be used to evaluate the performance of the chosen deep learning classifier, including the accuracy, precision, and F1-score [89]. Table 1 shows the evaluation metrics derived from the experimental results. In our study, these metrics were chosen where the accuracy was the ratio between the decoded samples and the total number of samples. The three metrics can be defined as follows:

$$Acc = \frac{TP + TN}{TP + FN + FP + TN}, \quad (16)$$

$$rec = \frac{TP}{TP + FN}, \quad (17)$$

$$F1 - score = \frac{2 \times (rec \cdot pre)}{(rec + pre)}, \quad (18)$$

where  $TP, TN$  are the true-positive and true-negative values,  $FP, FN$  give the false-positive and false-negative values, and  $pre$  is the precision, which can be calculated as follows:

$$pre = \frac{TP}{TP + FP} \quad (19)$$

**Table 1.** Comparative analysis of proposed IoHT-based HLD system with other deep learning approaches using accuracy, recall, precision, and F1-score for the two benchmarked datasets.

Performance	Proposed System with First Novelty	Proposed System with Second Novelty	Proposed System with Both Novelties	CNN	LSTM
<b>HWU-USP</b>					
Accuracy	78.89%	80.00%	<b>82.22%</b>	72.22%	70.00%
Recall	0.79	0.80	<b>0.82</b>	0.72	0.70
Precision	0.79	0.81	<b>0.83</b>	0.73	0.71
F1-Score	0.79	0.81	<b>0.82</b>	0.72	0.70
<b>LARa</b>					
Accuracy	80.00%	77.50%	<b>82.50%</b>	78.75%	76.25%
Recall	0.80	0.77	<b>0.82</b>	0.78	0.76
Precision	0.80	0.78	<b>0.83</b>	0.79	0.76
F1-Score	0.80	0.77	<b>0.82</b>	0.79	0.76

#### 4.3. Experiment 2: Comparison with Baseline HLD Systems

In the first experiment, we tested to highlight the importance of novel techniques introduced in this system. The first novelty is the motion and ambient data filtration technique that can handle sensor signal-based errors, biasness, and drift. The second novelty is ambient and motion descriptor extraction through a graph-based approach that helps extract robust descriptors related to the data type. The comparative results for the proposed IoHT-based HLD system with the first novelty, second novelty, and both together are given in Table 1, along with a comparison of the same system classification through the CNN [90] and LSTM [91].

We used the scikit-learn library to train all three classifiers. We set the learning rate for the CNN to 0.001, and the maximum epoch number was 200. The input layer contained the descriptors extracted. Then, we proposed three convolution layers with the ReLU activation function. Next, the pooling layer was utilized after each convolution layer. A flattened layer was also used to flatten the shape of the layers. Further, a fully connected layer with two hidden layers and a softmax layer were also used to test the trained data through output. For the LSTM, we used the architecture proposed in [92], where an input layer, a few LSTM-based temporal models, a flattened layer, and a fully connected network were used to recognize the ADL. Table 2 shows the confidence levels of extracted skeleton body-points compared to the ground truth values over HWU-USP and LARa datasets.

**Table 2.** Confidence levels for skeleton body points over HWU-USP and LARa datasets.

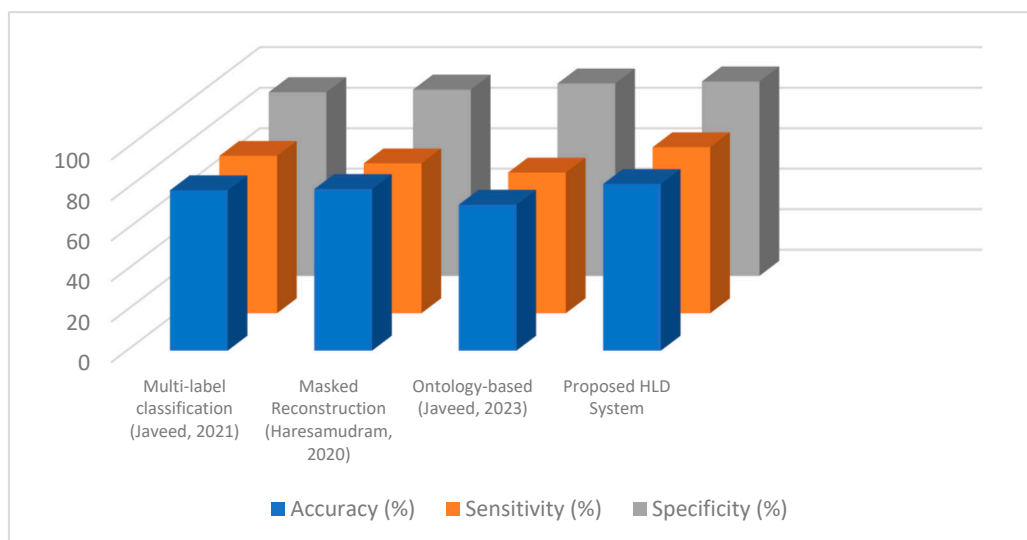
Skeleton Body Points	Confidence Levels for HWU-USP	Confidence Levels for LARa
Head	0.95	0.94
Shoulders	0.92	0.90
Elbows	0.88	0.89
Wrists	0.91	0.90
Torso	0.85	0.88
Knees	0.89	0.92
Ankles	0.95	0.94
<b>Mean Confidence</b>	<b>0.90</b>	<b>0.91</b>

#### 4.4. Experiment 3: Comparison with Other Works Utilizing Filtration and Descriptors

This section will focus on comparing the two novelties with the existing techniques by comparing them with the proposed HLD system. Figure 16 compares the accuracies of the proposed HLD mechanism and other existing techniques [93–95] that also used data filtration along with feature extraction. In [93], the authors utilized a combination of IMU, mechanomyography, and electromyography sensors and filtered them using median, band-pass, and moving-average filters to remove noise. Next, they made windows of 5 s each from the data and applied different techniques for the feature extraction, including peak-to-peak, abrupt changes, skewness, and mean frequency. Further, to reduce the features' vector dimension, they propose a multi-layer sequential forward selection method followed by classification via the random forest.

Haresamudram et al. present a self-supervised technique called masked reconstruction for HAR in [94]. They used small-labeled datasets and filtered data using transformer encoders. Then, they trained the network using different features and transfer learning mechanisms. In [95], a similar method to filter the data from motion, ambient, and vision-based sensors is proposed. The authors extracted features such as dynamic time warping, hidden Markov random fields, Mel-frequency cepstral coefficients, a gray-level co-variance matrix, and geodesic distance. Further, these features were optimized using a genetic algorithm and the system-recognized activities via a hidden Markov model-based classifier. As can be observed in Figure 16, the proposed HLD system with two novelties outperformed the existing works in terms of accuracy, sensitivity, and specificity.





**Figure 16.** Comparison of previous works [93–95] with proposed HLD systems over the two novelties proposed.

#### 4.5. Experiment 4: Comparisons with Existing Works

This section gives a comparison of our proposed IoHT-based HLD method with other previous state-of-the-art systems. We compared the proposed HLD system with methodologies that have hand-crafted descriptor extraction techniques, multiple datasets, machine learning, and applied deep learning techniques. Table 3 summarizes the comparison of the proposed system with other systems based on the classifiers, descriptor domain, modality, and accuracy achieved.

**Table 3.** Comparative analysis of proposed IoHT-based HLD system in terms of accuracy with existing work in the literature.

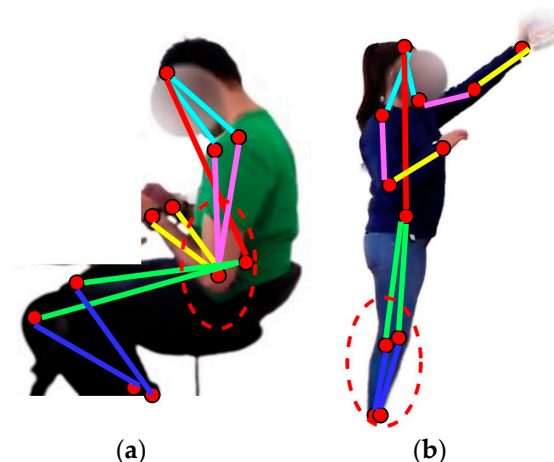
Ref.	Classifier	Descriptor Domain	Modality	Accuracy
[96]	Random Forest	Time-based	Multiple	81.00
[97]	CNN-LSTM	Deep-learning-based	Multiple	75.00
[98]	HMM	Machine learning	Single	78.33
[99]	Multi-Layer Perceptron	Frequency and time	Single	74.20
[100]	Multi-Layer Perceptron	Entropy	Multiple	75.50
[101]	Markov Chain	Multi-features	Multiple	74.94
[102]	Recurrent Neural Network	Convolutional	Multiple	82.00
[103]	Recurrent Neural Network	Raw	Single	80.43
<b>Proposed</b>	<b>Recurrent Neural Network</b>	<b>Energy, Graph, Frequency, and Time</b>	<b>Multiple</b>	<b>82.36</b>

The comparison between multiple human activity recognition models is explained in the table. It focuses on the classifiers used to recognize these activities. The descriptors extracted for classification are also presented. Different models acquired either single- or multiple-sensor-based raw data. Single-sensor-based means that the data were acquired from one sensor type. In contrast, multimodal-sensor-based means that the data were gathered from multiple sensor types. The accuracies of each system compared are given in the table.

## 5. Discussion

Although human locomotion decoding was achieved successfully using the proposed IoHT-based HLD system, this study also has a few limitations. The skeleton body points extracted can be obstructed in different human postures, which can cause limitations for accurate locomotion decoding. A couple of examples are highlighted in Figure 17 using red

ellipses. The proposed filtration technique and descriptor extraction methodologies have to be assessed using some systems and datasets to verify the results. There is still a need to test this novel HLD system over different settings and datasets to validate the outcomes.



**Figure 17.** Samples of obstruction caused by human postures in activities over HWU-USP dataset: (a) using a phone and (b) taking out a bowl.

## 6. Conclusions

This article proposes a deep-learning-based human-locomotion-decoding system via novel filtration techniques and two innovative descriptor extraction mechanisms. The study compared two novelties of the proposed system using an RNN, a CNN, and LSTM. The RNN outperformed the other two deep learners concerning the accuracy of the IoHT-based HLD system. We have also shown that all the compared classifiers performed acceptably over the HWU-USP and LARa datasets. By comparing the three classifiers and other previous state-of-the-art methodologies, we conclude that the proposed IoHT-based HLD architecture enhances the accuracy rates for human locomotion decoding. Therefore, the proposed system has many applications in human activity decoding and can be scaled for more practical solutions in smart homes, ambient assisted living, and care-based facilities. In the future, we can compare and improve the results of the current study using different settings, datasets, and deep learning techniques.

**Author Contributions:** Conceptualization: M.J., M.A. and A.A.; methodology: M.A. and M.J.; software: M.A. and A.J.; validation: M.J. and M.A.; formal analysis: A.J. and A.A.; resources: M.J., A.A., M.A. and A.J.; writing—review and editing: M.J., M.A. and A.J.; funding acquisition: M.J., M.A., A.A. and A.J. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2023R97), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

**Data Availability Statement:** Data are contained within the article.

**Acknowledgments:** Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2023R97), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ramanujam, E.; Perumal, T.; Padmvavathi, S. Human Activity Recognition with Smartphone and Wearable Sensors Using Deep Learning Techniques: A Review. *IEEE Sens. J.* **2021**, *21*, 13029–13040. [\[CrossRef\]](#)
2. Ouyed, O.; Allili, M.S. Group-of-features relevance in multinomial kernel logistic regression and application to human interaction recognition. *Expert Syst. Appl.* **2020**, *148*, 113247. [\[CrossRef\]](#)
3. Abid Hasan, S.M.; Ko, K. Depth edge detection by image-based smoothing and morphological operations. *J. Comput. Des. Eng.* **2016**, *3*, 191–197. [\[CrossRef\]](#)

4. Batool, M.; Jalal, A.; Kim, K. Telemonitoring of daily activity using Accelerometer and Gyroscope in smart home environments. *J. Electr. Eng. Technol.* **2020**, *15*, 2801–2809. [\[CrossRef\]](#)
5. Javeed, M.; Mudawi, N.A.; Alabdualah, B.I.; Jalal, A.; Kim, W. A Multimodal IoT-Based Locomotion Classification System Using Features Engineering and Recursive Neural Network. *Sensors* **2023**, *23*, 4716. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Shen, X.; Du, S.-C.; Sun, Y.-N.; Sun, P.Z.H.; Law, R.; Wu, E.Q. Advance Scheduling for Chronic Care Under Online or Offline Revisit Uncertainty. *IEEE Trans. Autom. Sci. Eng.* **2023**, 1–14. [\[CrossRef\]](#)
7. Wang, N.; Chen, J.; Chen, W.; Shi, Z.; Yang, H.; Liu, P.; Wei, X.; Dong, X.; Wang, C.; Mao, L.; et al. The effectiveness of case management for cancer patients: An umbrella review. *BMC Health Serv. Res.* **2022**, *22*, 1247. [\[CrossRef\]](#)
8. Hu, S.; Chen, W.; Hu, H.; Huang, W.; Chen, J.; Hu, J. Coaching to develop leadership for healthcare managers: A mixed-method systematic review protocol. *Syst. Rev.* **2022**, *11*, 67. [\[CrossRef\]](#)
9. Azmat, U.; Ahmad, J. Smartphone inertial sensors for human locomotion activity recognition based on template matching and codebook generation. In Proceedings of the IEEE International Conference on Communication Technologies, Rawalpindi, Pakistan, 21–22 September 2021.
10. Lv, Z.; Chen, D.; Feng, H.; Zhu, H.; Lv, H. Digital Twins in Unmanned Aerial Vehicles for Rapid Medical Resource Delivery in Epidemics. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 25106–25114. [\[CrossRef\]](#) [\[PubMed\]](#)
11. İnce, Ö.F.; Ince, I.F.; Yildirim, M.E.; Park, J.S.; Song, J.K.; Yoon, B.W. Human activity recognition with analysis of angles between skeletal joints using a RGB-depth sensor. *ETRI J.* **2020**, *42*, 78–89. [\[CrossRef\]](#)
12. Cheng, B.; Zhu, D.; Zhao, S.; Chen, J. Situation-Aware IoT Service Coordination Using the Event-Driven SOA Paradigm. *IEEE Trans. Netw. Serv. Manag.* **2016**, *13*, 349–361. [\[CrossRef\]](#)
13. Sun, Y.; Xu, C.; Li, G.; Xu, W.; Kong, J.; Jiang, D.; Tao, B.; Chen, D. Intelligent human computer interaction based on non-redundant EMG signal. *Alex. Eng. J.* **2020**, *59*, 1149–1157. [\[CrossRef\]](#)
14. Muneeb, M.; Rustam, H.; Ahmad, J. Automate Appliances via Gestures Recognition for Elderly Living Assistance. In Proceedings of the 2023 4th International Conference on Advancements in Computational Sciences (ICACS), Lahore, Pakistan, 20–22 February 2023; pp. 1–6. [\[CrossRef\]](#)
15. Nguyen, N.; Bui, D.; Tran, X. A novel hardware architecture for human detection using HOG-SVM co-optimization. In Proceedings of the APCCAS, Bangkok, Thailand, 11–14 November 2019. [\[CrossRef\]](#)
16. Nadeem, A.; Jalal, A.; Kim, K. Automatic human posture estimation for sport activity recognition with robust body parts detection and entropy markov model. *Multimed. Tools Appl.* **2021**, *80*, 21465–21498. [\[CrossRef\]](#)
17. Zank, M.; Nescher, T.; Kunz, A. Tracking human locomotion by relative positional feet tracking. In Proceedings of the IEEE Virtual Reality (VR), Arles, France, 23–27 March 2015. [\[CrossRef\]](#)
18. Jalal, A.; Mahmood, M. Students' behavior mining in e-learning environment using cognitive processes with information technologies. *Educ. Inf. Technol.* **2019**, *24*, 2797–2821. [\[CrossRef\]](#)
19. Batool, M.; Jalal, A.; Kim, K. Sensors Technologies for Human Activity Analysis Based on SVM Optimized by PSO Algorithm. In Proceedings of the 2019 International Conference on Applied and Engineering Mathematics (ICAEM), Taxila, Pakistan, 27–29 August 2019; pp. 145–150. [\[CrossRef\]](#)
20. Prati, A.; Shan, C.; Wang, K.I.-K. Sensors, vision and networks: From video surveillance to activity recognition and health monitoring. *J. Ambient Intell. Smart Environ.* **2019**, *11*, 5–22. [\[CrossRef\]](#)
21. Wang, Y.; Xu, N.; Liu, A.-A.; Li, W.; Zhang, Y. High-Order Interaction Learning for Image Captioning. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 4417–4430. [\[CrossRef\]](#)
22. Zhang, C.; Xiao, P.; Zhao, Z.-T.; Liu, Z.; Yu, J.; Hu, X.-Y.; Chu, H.-B.; Xu, J.-J.; Liu, M.-Y.; Zou, Q.; et al. A Wearable Localized Surface Plasmons Antenna Sensor for Communication and Sweat Sensing. *IEEE Sens. J.* **2023**, *23*, 11591–11599. [\[CrossRef\]](#)
23. Lin, Q.; Xiongbo, G.; Zhang, W.; Cai, L.; Yang, R.; Chen, H.; Cai, K. A Novel Approach of Surface Texture Mapping for Cone-beam Computed Tomography in Image-guided Surgical Navigation. *IEEE J. Biomed. Health Inform.* **2023**, 1–10. [\[CrossRef\]](#)
24. Hu, Z.; Ren, L.; Wei, G.; Qian, Z.; Liang, W.; Chen, W.; Lu, X.; Ren, L.; Wang, K. Energy Flow and Functional Behavior of Individual Muscles at Different Speeds During Human Walking. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2023**, *31*, 294–303. [\[CrossRef\]](#)
25. Zhang, R.; Li, L.; Zhang, Q.; Zhang, J.; Xu, L.; Zhang, B.; Wang, B. Differential Feature Awareness Network within Antagonistic Learning for Infrared-Visible Object Detection. *IEEE Trans. Circuits Syst. Video Technol.* **2023**. [\[CrossRef\]](#)
26. Mahmood, M.; Ahmad, J.; Kim, K. WHITE STAG model: Wise human interaction tracking and estimation (WHITE) using spatio-temporal and angular-geometric (STAG) descriptors. *Multimed. Tools Appl.* **2020**, *79*, 6919–6950. [\[CrossRef\]](#)
27. Zheng, M.; Zhi, K.; Zeng, J.; Tian, C.; You, L. A hybrid CNN for image denoising. *J. Artif. Intell. Technol.* **2022**, *2*, 93–99. [\[CrossRef\]](#)
28. Gao, Z.; Pan, X.; Shao, J.; Jiang, X.; Su, Z.; Jin, K.; Ye, J. Automatic interpretation and clinical evaluation for fundus fluorescein angiography images of diabetic retinopathy patients by deep learning. *Br. J. Ophthalmol.* **2022**, *107*, 1852–1858. [\[CrossRef\]](#)
29. Wang, W.; Qi, F.; Wipf, D.P.; Cai, C.; Yu, T.; Li, Y.; Zhang, Y.; Yu, Z.; Wu, W. Sparse Bayesian Learning for End-to-End EEG Decoding. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 15632–15649. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Lu, S.; Liu, S.; Hou, P.; Yang, B.; Liu, M.; Yin, L.; Zheng, W. Soft Tissue Feature Tracking Based on Deep Matching Network. *Comput. Model. Eng. Sci.* **2023**, *136*, 363–379. [\[CrossRef\]](#)
31. Sreenu, G.; Saleem Durai, M.A. Intelligent video surveillance: A review through deep learning techniques for crowd analysis. *J. Big Data* **2019**, *6*, 48. [\[CrossRef\]](#)

32. Xu, H.; Pan, Y.; Li, J.; Nie, L.; Xu, X. Activity recognition method for home-based elderly care service based on random forest and activity similarity. *IEEE Access* **2019**, *7*, 16217–16225. [\[CrossRef\]](#)
33. Beddiar, D.R.; Nini, B.; Sabokrou, M.; Hadid, A. Vision-based human activity recognition: A survey. *Multimed. Tools Appl.* **2020**, *79*, 30509–30555. [\[CrossRef\]](#)
34. Hu, X.; Kuang, Q.; Cai, Q.; Xue, Y.; Zhou, W.; Li, Y. A Coherent Pattern Mining Algorithm Based on All Contiguous Column Biclust. *J. Artif. Intell. Technol.* **2022**, *2*, 80–92. [\[CrossRef\]](#)
35. Quaid, M.A.K.; Ahmad, J. Wearable sensors based human behavioral pattern recognition using statistical features and reweighted genetic algorithm. *Multimed. Tools Appl.* **2020**, *79*, 6061–6083. [\[CrossRef\]](#)
36. Ahmad, F. Deep image retrieval using artificial neural network interpolation and indexing based on similarity measurement. *CAAI Trans. Intell. Technol.* **2022**, *7*, 200–218. [\[CrossRef\]](#)
37. Zhang, J.; Ye, G.; Tu, Z.; Qin, Y.; Qin, Q.; Zhang, J.; Liu, J. A spatial attentive and temporal dilated (SATD) GCN for skeleton-based action recognition. *CAAI Trans. Intell. Technol.* **2022**, *7*, 46–55. [\[CrossRef\]](#)
38. Lu, S.; Yang, J.; Yang, B.; Yin, Z.; Liu, M.; Yin, L.; Zheng, W. Analysis and Design of Surgical Instrument Localization Algorithm. *Comput. Model. Eng. Sci.* **2023**, *137*, 669–685. [\[CrossRef\]](#)
39. Zhang, J.; Zhu, C.; Zheng, L.; Xu, K. ROSEFusion: Random optimization for online dense reconstruction under fast camera motion. *ACM Trans. Graph.* **2021**, *40*, 1–17. [\[CrossRef\]](#)
40. Meng, J.; Li, Y.; Liang, H.; Ma, Y. Single-image Dehazing based on two-stream convolutional neural network. *J. Artif. Intell. Technol.* **2022**, *2*, 100–110. [\[CrossRef\]](#)
41. Ma, K.; Li, Z.; Liu, P.; Yang, J.; Geng, Y.; Yang, B.; Guan, X. Reliability-Constrained Throughput Optimization of Industrial Wireless Sensor Networks With Energy Harvesting Relay. *IEEE Internet Things J.* **2021**, *8*, 13343–13354. [\[CrossRef\]](#)
42. Zhuang, Y.; Jiang, N.; Xu, Y.; Xiangjie, K.; Kong, X. Progressive Distributed and Parallel Similarity Retrieval of Large CT Image Sequences in Mobile Telemedicine Networks. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 6458350. [\[CrossRef\]](#)
43. Miao, Y.; Wang, X.; Wang, S.; Li, R. Adaptive Switching Control Based on Dynamic Zero-Moment Point for Versatile Hip Exoskeleton Under Hybrid Locomotion. *IEEE Trans. Ind. Electron.* **2023**, *70*, 11443–11452. [\[CrossRef\]](#)
44. He, B.; Lu, Q.; Lang, J.; Yu, H.; Peng, C.; Bing, P.; Li, S.; Zhou, Q.; Liang, Y.; Tian, G. A New Method for CTC Images Recognition Based on Machine Learning. *Front. Bioeng. Biotechnol.* **2020**, *8*, 897. [\[CrossRef\]](#)
45. Li, Z.; Kong, Y.; Jiang, C. A Transfer Double Deep Q Network Based DDoS Detection Method for Internet of Vehicles. *IEEE Trans. Veh. Technol.* **2023**, *72*, 5317–5331. [\[CrossRef\]](#)
46. Hassan, F.S.; Gutub, A. Improving data hiding within colour images using hue component of HSV colour space. *CAAI Trans. Intell. Technol.* **2022**, *7*, 56–68. [\[CrossRef\]](#)
47. Zheng, W.; Xun, Y.; Wu, X.; Deng, Z.; Chen, X.; Sui, Y. A Comparative Study of Class Rebalancing Methods for Security Bug Report Classification. *IEEE Trans. Reliab.* **2021**, *70*, 1658–1670. [\[CrossRef\]](#)
48. Zheng, C.; An, Y.; Wang, Z.; Wu, H.; Qin, X.; Eynard, B.; Zhang, Y. Hybrid offline programming method for robotic welding systems. *Robot. Comput.-Integr. Manuf.* **2022**, *73*, 102238. [\[CrossRef\]](#)
49. Zhang, X.; Wang, Y.; Yang, M.; Geng, G. Toward Concurrent Video Multicast Orchestration for Caching-Assisted Mobile Networks. *IEEE Trans. Veh. Technol.* **2021**, *70*, 13205–13220. [\[CrossRef\]](#)
50. Qi, M.; Cui, S.; Chang, X.; Xu, Y.; Meng, H.; Wang, Y.; Yin, T. Multi-region Nonuniform Brightness Correction Algorithm Based on L-Channel Gamma Transform. *Secur. Commun. Netw.* **2022**, *2022*, 2675950. [\[CrossRef\]](#)
51. Zhao, W.; Lun, R.; Espy, D.D.; Reinthal, M.A. Rule based real time motion assessment for rehabilitation exercises. In Proceedings of the IEEE Symposium Computational Intelligence in Healthcare and E-Health, Orlando, FL, USA, 9–12 December 2014. [\[CrossRef\]](#)
52. Hao, S.; Jiali, P.; Xiaomin, Z.; Xiaoqin, W.; Lina, L.; Xin, Q.; Qin, L. Group identity modulates bidding behavior in repeated lottery contest: Neural signatures from event-related potentials and electroencephalography oscillations. *Front. Neurosci.* **2023**, *17*, 1184601. [\[CrossRef\]](#)
53. Barnachon, M.; Bouakaz, S.; Boufama, B.; Guillou, E. Ongoing human action recognition with motion capture. *Pattern Recognit.* **2014**, *47*, 238–247. [\[CrossRef\]](#)
54. Lu, S.; Yang, B.; Xiao, Y.; Liu, S.; Liu, M.; Yin, L.; Zheng, W. Iterative reconstruction of low-dose CT based on differential sparse. *Biomed. Signal Process. Control* **2023**, *79*, 104204. [\[CrossRef\]](#)
55. Ordóñez, F.; Roggen, D. Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition. *Sensors* **2016**, *16*, 115. [\[CrossRef\]](#)
56. Franco, A.; Magnani, A.; Maio, D. A multimodal approach for human activity recognition based on skeleton and RGB data. *Pattern Recognit. Lett.* **2020**, *131*, 293–299. [\[CrossRef\]](#)
57. Nweke, H.F.; Teh, Y.W.; Mujtaba, G.; Alo, U.R.; Al-Garadi, M.A. Multi-sensor fusion based on multiple classifier systems for human activity identification. *Hum. Cent. Comput. Inf. Sci.* **2019**, *9*, 34. [\[CrossRef\]](#)
58. Ehatisham-Ul-Haq, M.; Javed, A.; Azam, M.A.; Malik, H.M.A.; Irtaza, A.; Lee, I.H.; Mahmood, M.T. Robust Human Activity Recognition Using Multimodal Feature-Level Fusion. *IEEE Access* **2019**, *7*, 60736–60751. [\[CrossRef\]](#)
59. Chung, S.; Lim, J.; Noh, K.J.; Kim, G.; Jeong, H. Sensor Data Acquisition and Multimodal Sensor Fusion for Human Activity Recognition Using Deep Learning. *Sensors* **2019**, *19*, 1716. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Zhao, Y.; Guo, S.; Chen, Z.; Shen, Q.; Meng, Z.; Xu, H. Marfusion: An Attention-Based Multimodal Fusion Model for Human Activity Recognition in Real-World Scenarios. *Appl. Sci.* **2022**, *12*, 5408. [\[CrossRef\]](#)



61. Niemann, F.; Reining, C.; Rueda, F.M.; Nair, N.R.; Steffens, J.A.; Fink, G.A.; Hompel, M.T. LARa: Creating a Dataset for Human Activity Recognition in Logistics Using Semantic Attributes. *Sensors* **2020**, *20*, 4083. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Ranieri, C.M.; MacLeod, S.; Dragone, M.; Vargas, P.A.; Romero, R.A.F. Activity Recognition for Ambient Assisted Living with Videos, Inertial Units and Ambient Sensors. *Sensors* **2021**, *21*, 768. [\[CrossRef\]](#)
63. Bersch, S.D.; Azzi, D.; Khusainov, R.; Achumba, I.E.; Ries, J. Sensor data acquisition and processing parameters for human activity classification. *Sensors* **2014**, *14*, 4239–4270. [\[CrossRef\]](#)
64. Huang, H.; Liu, L.; Wang, J.; Zhou, Y.; Hu, H.; Ye, X.; Liu, G.; Xu, Z.; Xu, H.; Yang, W.; et al. Aggregation caused quenching to aggregation induced emission transformation: A precise tuning based on BN-doped polycyclic aromatic hydrocarbons toward subcellular organelle specific imaging. *Chem. Sci.* **2022**, *13*, 3129–3139. [\[CrossRef\]](#)
65. Schrader, L.; Vargas Toro, A.; Konietzny, S.; Rüping, S.; Schäpers, B.; Steinböck, M.; Krewer, C.; Müller, F.; Güttler, J.; Bock, T. Advanced sensing and human activity recognition in early intervention and rehabilitation of elderly people. *Popul. Ageing* **2020**, *13*, 139–165. [\[CrossRef\]](#)
66. Lee, M.; Kim, S.B. Sensor-Based Open-Set Human Activity Recognition Using Representation Learning with Mixup Triplets. *IEEE Access* **2022**, *10*, 119333–119344. [\[CrossRef\]](#)
67. Patro, S.G.K.; Mishra, B.K.; Panda, S.K.; Kumar, R.; Long, H.V.; Taniar, D.; Priyadarshini, I. A Hybrid Action-Related K-Nearest Neighbour (HAR-KNN) Approach for Recommendation Systems. *IEEE Access* **2020**, *8*, 90978–90991. [\[CrossRef\]](#)
68. Li, J.; Tian, L.; Wang, H.; An, Y.; Wang, K.; Yu, L. Segmentation and recognition of basic and transitional activities for continuous physical human activity. *IEEE Access* **2019**, *7*, 42565–42576. [\[CrossRef\]](#)
69. Chen, D.; Wang, Q.; Li, Y.; Li, Y.; Zhou, H.; Fan, Y. A general linear free energy relationship for predicting partition coefficients of neutral organic compounds. *Chemosphere* **2020**, *247*, 125869. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Hou, X.; Zhang, L.; Su, Y.; Gao, G.; Liu, Y.; Na, Z.; Xu, Q.; Ding, T.; Xiao, L.; Li, L.; et al. A space crawling robotic bio-paw (SCRBP) enabled by triboelectric sensors for surface identification. *Nano Energy* **2023**, *105*, 108013. [\[CrossRef\]](#)
71. Hou, X.; Xin, L.; Fu, Y.; Na, Z.; Gao, G.; Liu, Y.; Xu, Q.; Zhao, P.; Yan, G.; Su, Y.; et al. A self-powered biomimetic mouse whisker sensor (BMWS) aiming at terrestrial and space objects perception. *Nano Energy* **2023**, *118*, 109034. [\[CrossRef\]](#)
72. Mi, W.; Xia, Y.; Bian, Y. Meta-analysis of the association between aldose reductase gene (CA)n microsatellite variants and risk of diabetic retinopathy. *Exp. Ther. Med.* **2019**, *18*, 4499–4509. [\[CrossRef\]](#) [\[PubMed\]](#)
73. Ye, X.; Wang, J.; Qiu, W.; Chen, Y.; Shen, L. Excessive gliosis after vitrectomy for the highly myopic macular hole: A Spectral Domain Optical Coherence Tomography Study. *Retina* **2023**, *43*, 200–208. [\[CrossRef\]](#) [\[PubMed\]](#)
74. Chen, C.; Liu, S. Detection and Segmentation of Occluded Vehicles Based on Skeleton Features. In Proceedings of the 2012 Second International Conference on Instrumentation, Measurement, Computer, Communication and Control, Harbin, China, 8–10 December 2012; pp. 1055–1059. [\[CrossRef\]](#)
75. Chen, C.; Jafari, R.; Kehtarnavaz, N. A survey of depth and inertial sensor fusion for human action recognition. *Multimed. Tools Appl.* **2017**, *76*, 4405–4425. [\[CrossRef\]](#)
76. Amir, N.; Ahmad, J.; Kibum, K. Human Actions Tracking and Recognition Based on Body Parts Detection via Artificial Neural Network. In Proceedings of the 2020 3rd International Conference on Advancements in Computational Sciences (ICACS), Lahore, Pakistan, 17–19 February 2020.
77. Zhou, B.; Wang, C.; Huan, Z.; Li, Z.; Chen, Y.; Gao, G.; Li, H.; Dong, C.; Liang, J. A Novel Segmentation Scheme with Multi-Probability Threshold for Human Activity Recognition Using Wearable Sensors. *Sensors* **2022**, *22*, 7446. [\[CrossRef\]](#)
78. Yao, Q.-Y.; Fu, M.-L.; Zhao, Q.; Zheng, X.-M.; Tang, K.; Cao, L.-M. Image-based visualization of stents in mechanical thrombectomy for acute ischemic stroke: Preliminary findings from a series of cases. *World J. Clin. Cases* **2023**, *11*, 5047–5055. [\[CrossRef\]](#)
79. Su, W.; Ni, J.; Hu, X.; Fridrich, J. Image Steganography With Symmetric Embedding Using Gaussian Markov Random Field Model. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *31*, 1001–1015. [\[CrossRef\]](#)
80. Li, X.; Zhao, H.; Yu, L.; Chen, H.; Deng, W.; Deng, W. Feature Extraction Using Parameterized Multisynchrosqueezing Transform. *IEEE Sens. J.* **2022**, *22*, 14263–14272. [\[CrossRef\]](#)
81. Jin, K.; Gao, Z.; Jiang, X.; Wang, Y.; Ma, X.; Li, Y.; Ye, J. MSHF: A Multi-Source Heterogeneous Fundus (MSHF) Dataset for Image Quality Assessment. *Sci. Data* **2023**, *10*, 286. [\[CrossRef\]](#) [\[PubMed\]](#)
82. Amir, N.; Ahmad, J.; Kim, K. Accurate Physical Activity Recognition using Multidimensional Features and Markov Model for Smart Health Fitness. *Symmetry* **2020**, *12*, 1766.
83. Kanan, C.; Cottrell, G. Robust Classification of Objects, Faces, and Flowers Using Natural Image Statistics. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 2472–2479. [\[CrossRef\]](#)
84. Arbain, N.A.; Azmi, M.S.; Muda, A.K.D.; Radzid, A.R.; Tahir, A. A Review of Triangle Geometry Features in Object Recognition. In Proceedings of the 2019 9th Symposium on Computer Applications & Industrial Electronics (ISCAIE), Kota Kinabalu, Malaysia, 27–28 April 2019; pp. 254–258. [\[CrossRef\]](#)
85. Fausto, F.; Cuevas, E.; Gonzales, A. A New Descriptor for Image Matching Based on Bionic Principles. *Pattern Anal. Appl.* **2017**, *20*, 1245–1259. [\[CrossRef\]](#)
86. Yu, Y.; Yang, J.P.; Shiu, C.-S.; Simoni, J.M.; Xiao, S.; Chen, W.-T.; Rao, D.; Wang, M. Psychometric testing of the Chinese version of the Medical Outcomes Study Social Support Survey among people living with HIV/AIDS in China. *Appl. Nurs. Res.* **2015**, *28*, 328–333. [\[CrossRef\]](#) [\[PubMed\]](#)



87. Ali, H.H.; Moftah, H.M.; Youssif, A.A.A. Depth-based human activity recognition: A comparative perspective study on feature extraction. *Future Comput. Inform. J.* **2018**, *3*, 51–67. [[CrossRef](#)]
88. Nguyen, H.-C.; Nguyen, T.-H.; Scherer, R.; Le, V.-H. Deep Learning for Human Activity Recognition on 3D Human Skeleton: Survey and Comparative Study. *Sensors* **2023**, *23*, 5121. [[CrossRef](#)]
89. Singh, S.P.; Sharma, M.K.; Lay-Ekuakille, A.; Gangwar, D.; Gupta, S. Deep ConvLSTM With Self-Attention for Human Activity Decoding Using Wearable Sensors. *IEEE Sens. J.* **2021**, *21*, 8575–8582. [[CrossRef](#)]
90. Farag, M.M. Matched Filter Interpretation of CNN Classifiers with Application to HAR. *Sensors* **2022**, *22*, 8060. [[CrossRef](#)] [[PubMed](#)]
91. Husni, N.L.; Sari, P.A.R.; Handayani, A.S.; Dewi, T.; Seno, S.A.H.; Caesarendra, W.; Glowacz, A.; Oprzędkiewicz, K.; Sułowicz, M. Real-Time Littering Activity Monitoring Based on Image Classification Method. *Smart Cities* **2021**, *4*, 1496–1518. [[CrossRef](#)]
92. Khatun, M.A.; Abu Yousuf, M.; Ahmed, S.; Uddin, Z.; Alyami, S.A.; Al-Ashhab, S.; Akhdar, H.F.; Khan, A.; Azad, A.; Moni, M.A. Deep CNN-LSTM With Self-Attention Model for Human Activity Recognition Using Wearable Sensor. *IEEE J. Transl. Eng. Health Med.* **2022**, *10*, 2700316. [[CrossRef](#)]
93. Javeed, M.; Jalal, A.; Kim, K. 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 (IBCAST), Islamabad, Pakistan, 12–16 January 2021; pp. 512–517. [[CrossRef](#)]
94. Haresamudram, H.; Beedu, A.; Agrawal, V.; Grady, P.L.; Essa, I. Masked Reconstruction Based Self-Supervision for Human Activity Recognition. In Proceedings of the 24th annual International Symposium on Wearable Computers, Cancun, Mexico, 12–16 September 2020.
95. Javeed, M.; Mudawi, N.A.; Alazeb, A.; Alotaibi, S.S.; Almujaal, N.A.; Jalal, A. Deep Ontology-Based Human Locomotor Activity Recognition System via Multisensory Devices. *IEEE Access* **2023**, *11*, 105466–105478. [[CrossRef](#)]
96. Cosoli, G.; Antognoli, L.; Scalise, L. Wearable Electrocardiography for Physical Activity Monitoring: Definition of Validation Protocol and Automatic Classification. *Biosensors* **2023**, *13*, 154. [[CrossRef](#)]
97. Ehatisham-ul-Haq, M.; Murtaza, F.; Azam, M.A.; Amin, Y. Daily Living Activity Recognition In-The-Wild: Modeling and Inferring Activity-Aware Human Contexts. *Electronics* **2022**, *11*, 226. [[CrossRef](#)]
98. Jalal, A.; Kamal, S.; Kim, D. A depth video sensor-based life-logging human activity recognition system for elderly care in smart indoor environments. *Sensors* **2014**, *14*, 11735–11759.
99. Aşuroğlu, T. Complex Human Activity Recognition Using a Local Weighted Approach. *IEEE Access* **2022**, *10*, 101207–101219. [[CrossRef](#)]
100. Azmat, U.; Ahmad, J.; Madiha, J. Multi-sensors Fused IoT-based Home Surveillance via Bag of Visual and Motion Features. In Proceedings of the 2023 International Conference on Communication, Computing and Digital Systems (C-CODE), Islamabad, Pakistan, 17–18 May 2023; pp. 1–6. [[CrossRef](#)]
101. Ahmad, J.; Kim, Y.-H.; Kim, Y.-J.; Kamal, S.; Kim, D. Robust human activity recognition from depth video using spatiotemporal multi-fused features. *Pattern Recognit.* **2017**, *61*, 295–308.
102. Boukhechba, M.; Cai, L.; Wu, C.; Barnes, L.E. ActiPPG: Using deep neural networks for activity recognition from wrist-worn photoplethysmography (PPG) sensors. *Smart Health* **2019**, *14*, 100082. [[CrossRef](#)]
103. Sánchez-Caballero, A.; Fuentes-Jiménez, D.; Losada-Gutiérrez, C. Real-time human action recognition using raw depth video-based recurrent neural networks. *Multimed. Tools Appl.* **2023**, *82*, 16213–16235. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.