



Article

Human Posture Transition-Time Detection Based upon Inertial Measurement Unit and Long Short-Term Memory Neural Networks

Chun-Ting Kuo ¹, Jun-Ji Lin ¹, Kuo-Kuang Jen ², Wei-Li Hsu ³, Fu-Cheng Wang ¹, Tsu-Chin Tsao ⁴ and Jia-Yush Yen ^{1,5,*}

¹ Department of Mechanical Engineering, National Taiwan University, Taipei 106319, Taiwan

² Missile and Rocket Research Division, National Chung Shan Institute of Science and Technology, Taoyuan 325204, Taiwan

³ School and Graduate Institute of Physical Therapy, National Taiwan University, Taipei 106319, Taiwan

⁴ Mechanical and Aerospace Engineering, Samueli School of Engineering, UCLA, Los Angeles, CA 90095, USA

⁵ Department of Mechanical Engineering, National Taiwan University of Science and Technology, Taipei 106319, Taiwan

* Correspondence: jyen@mail.ntust.edu.tw; Tel.: +886-2-2730-1125; Fax: +886-2-2737-6107

Abstract: As human–robot interaction becomes more prevalent in industrial and clinical settings, detecting changes in human posture has become increasingly crucial. While recognizing human actions has been extensively studied, the transition between different postures or movements has been largely overlooked. This study explores using two deep-learning methods, the linear Feedforward Neural Network (FNN) and Long Short-Term Memory (LSTM), to detect changes in human posture among three different movements: standing, walking, and sitting. To explore the possibility of rapid posture-change detection upon human intention, the authors introduced transition stages as distinct features for the identification. During the experiment, the subject wore an inertial measurement unit (IMU) on their right leg to measure joint parameters. The measurement data were used to train the two machine learning networks, and their performances were tested. This study also examined the effect of the sampling rates on the LSTM network. The results indicate that both methods achieved high detection accuracies. Still, the LSTM model outperformed the FNN in terms of speed and accuracy, achieving 91% and 95% accuracy for data sampled at 25 Hz and 100 Hz, respectively. Additionally, the network trained for one test subject was able to detect posture changes in other subjects, demonstrating the feasibility of personalized or generalized deep learning models for detecting human intentions. The accuracies for posture transition time and identification at a sampling rate of 100 Hz were 0.17 s and 94.44%, respectively. In summary, this study achieved some good outcomes and laid a crucial foundation for the engineering application of digital twins, exoskeletons, and human intention control.



Citation: Kuo, C.-T.; Lin, J.-J.; Jen, K.-K.; Hsu, W.-L.; Wang, F.-C.; Tsao, T.-C.; Yen, J.-Y. Human Posture Transition-Time Detection Based upon Inertial Measurement Unit and Long Short-Term Memory Neural Networks. *Biomimetics* **2023**, *8*, 471. <https://doi.org/10.3390/biomimetics8060471>

Academic Editors: Yisheng Guan, Weiwei Wan and Li He

Received: 14 August 2023

Revised: 16 September 2023

Accepted: 28 September 2023

Published: 2 October 2023

Keywords: human posture change detection; deep learning; feedforward neural network (FNN); long short-term memory (LSTM); inertial measurement unit (IMU); internal sensing; human activity recognition (HAR)



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

Currently, with the rapid development of smart-life technology, the trend of smart sensing has shifted from the field of external sensing (such as smart-home appliances) to the field of the internal sensing of the human body [1]. Human posture-change detection (HPCD) is an essential aspect of human–robot interaction research, involving collecting and analyzing signals generated by human body movements [2]. Precise and rapid prediction of human intentions can greatly enhance coordination between humans and machines, enabling collaborative robots, digital twins in the industrial field, diagnostic systems, rehabilitation devices, and even everyday-life assistance with greater ease and efficiency [3–5].

In addition to human–robot interaction, identifying human gait or posture is crucial for controlling exoskeletons. Exoskeletons typically employ two levels of control to ensure safe operation. The lower-level control utilizes feedback-based impedance or admittance control to assist joint movements [6]. However, during the process, even a small movement in the pilot’s joint may require a significant torque force from the actuator. To avoid using high feedback gain, one often selects the feedforward control technique. However, using a high feedback gain can lead to instability and result in pilot injury. Feedforward control requires prior knowledge of the motion trajectory; therefore, many commercial exoskeletons must rely on the pilot to switch between operating modes.

The focus of this study is on the application of lower-limb exoskeleton control. Optimizing posture-change detection is especially important, because if there is a delay between the exoskeleton control action and the user’s intention, the delay could cause sluggish movement and even injury to the pilot [7,8]. Therefore, it is crucial to develop methods for sensing the user’s intentions and detecting posture changes in real time to ensure safe and effective use of the exoskeleton. This future research aims to address these issues and improve the autonomous execution of exoskeleton systems.

To correctly detect human intention, there are two issues to address. The first issue is having adequate means to measure body movements. In recent decades, many researchers and commercial devices for external or internal sensing of human signals have been available, with different ways to collect body signals. Commercial external sensing systems like optical flow [9], edge contour [10], motion-history image [11,12], spatio-temporal volume [13], and the human alternative model [14] are image-based methods, and a fixed location for data gathering is required. On the other hand, some systems, such as the inertial-measurement-unit (IMU) based method [15–17], the electromyography (EMG) based method [18], and the flexible-sensor-based method, have some advantages regarding their ability to be wearable and to allow subjects to move around while gathering data. From the survey results for the hardware environment, it might be evident and appropriate to choose wearable-form tools to achieve the goals of human posture-change detection from the exoskeleton control system [19,20]. Recently, some research has focused on information from various sensors, such as the above-mentioned IMU, EMG, electrogoniometer (EGM), and so on, which have been integrated for effective action recognition. However, these also require data accumulation and analysis through moving windows and other methods [21–23]. Also, compared with other types of wearable sensing tools, the data captured by IMU is related to angle changes of the human torso, which is a more beneficial characteristic as it is not limited by individual differences of users, and its computable amount would be lower in the gait or posture-detection research field [7,15,19,20].

The second issue regarding human posture-change detection is the posture identification algorithm. A suitable algorithm that provides fast recognition and good judging accuracy is desirable; however, the highly nonlinear nature of the problem, in addition to the stringent time constraint, made it very difficult to solve with conventional methods. People have used machine-learning or statistical methods to attain the best identification results. Machine learning methods include K-nearest neighbors, mean shift, decision tree, Bayes classifier, and support vector machine (SVM) [24,25]. Statistical methods with advanced knowledge analysis of kinematics and kinesiology include the self-similarity matrix method and hidden Markov model. This research employed some of the most popular approaches, which are also used in gait detection or human activity recognition (HAR) [26–28]. Additionally, there has also been much research using neural network (NN) techniques such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and deep learning [15,29–31].

Archetti et al. and Ragni et al. compared performances between linear discriminant analysis (LDA) and random forest (RF) in predicting the intended reaching of the target with subjects wearing electromagnetic sensors [32,33]. Li et al. used action recognition, action prediction, and posture-change detection to predict the pitcher’s choice of one of the nine-square divisions by capturing and analyzing the pitcher’s RGB image and optical

flow [34]. These results mainly concern the movements of the upper limbs. However, these results have focused on identifying the movements rather than obtaining a more precise time to predict intentions utilizing artificial intelligence (AI) or statistical tools. What is more, for timeseries-type data with specific postures and particular features, these methods might not consider the time-delay risk in implementation, and the machine learning or statistic methods above might not be suitable.

The authors of this research used the Long Short-Term Memory (LSTM) method to address these problems. LSTM is a recurrent neural network designed to process and predict events with long intervals and delays. It has a unique structure that enables it to perform better than general recurrent neural networks, especially in tasks such as autonomous speech recognition [35]. LSTM is not limited to speech-recognition tasks; it can also be applied in other areas, such as predicting human decisions [36], exoskeleton control, and posture-change detection for other equipment [37–39]. When the human body performs a posture change (whether it is wearing an exoskeleton or there is human–computer interaction), there is a linkage and correlation between each of the joints of lower limbs, and there is also continuity between and memory of various movements. The deep learning method of LSTM has the memory characteristics of recurrent neural networks, and can also avoid gradient-disappearance and gradient-explosion problems during long-sequence or long-time training. The characteristics of this study also focus on the determination of both specific continuous posture-transition time and status while other machine learning literature in the same field involve posture-status detection only. What is more, the accuracy of action recognition is relatively inconsistent in the existing literature [40]; so, the LSTM network structure is more suitable, and it is also the main reason and motivation for choosing the LSTM structure in this study.

This study conducted by the authors focused on the use of LSTM in exoskeleton control and posture-change detection. To evaluate the efficiency of LSTM in comparison to other algorithms, the authors also compared its performance with the linear feedforward neural network (FNN) method, a widely applied algorithm in AI research.

The primary goal of this research was to develop an algorithm that could detect changes in human postures with high precision and in a very short time. The ultimate aim was to reduce the time it takes for the algorithm to identify a new movement, enabling the exoskeleton control system to quickly switch to the new gait or posture, hoping to avoid causing danger, such as falling, or a situation in which, for example, the lower limbs are raised at an inappropriate moment or similar due to a misjudged strategy by the controller. This study also needs to consider the time delay for software and hardware integration. In addition, another focus of this research is how to achieve this goal with the smallest variety of sensors. This research used IMU sensors as the main measurement tool for several reasons. Firstly, these sensors provide a non-invasive way of detecting changes in posture, and they can be easily integrated with existing exoskeletons. Secondly, IMU sensors offer high data rates and good measurement resolutions, which are crucial for the fast detection of posture changes. Interestingly, the research team found that very few studies have focused on using only IMU sensors for the detection of the precise time of posture changes, despite these sensors being commonly used as wearable gait sensors to distinguish different human actions. In some cases, IMU sensors were used in conjunction with other sensor types to improve detection accuracy [17,19,20,40–42].

By developing an algorithm that can quickly and automatically recognize new postures, the pilot could enjoy a more comfortable ride without noticing the switches between different operating modes. It is worth noting that the response time of the algorithm was defined as the total time taken from when the recognition label was established by the algorithm to when the amount of data was input to the algorithm.

This research involved conducting experiments with IMU sensors attached to the right leg of subjects to analyze the changes among three different movements: standing, walking, and sitting. The researchers measured the joint angular displacements, angular velocities, and angular accelerations to evaluate the performance of the algorithm in detecting changes

in human posture. The research team also compared the performance of two deep learning methods, FNN and LSTM, to assess which method could more accurately detect changes in posture and shorten the response time of the algorithm. The ultimate goal of this research was to contribute to the development of real-time human posture-change detection and exoskeleton control using simple sensor types such as IMU sensors.

In summary, this research is distinguished by its attributes and substantial contributions:

- It stands out in its precise identification of specific postures, relying solely on IMU data to capture the intricate dynamics within the human body, obviating the need for external sensors;
- A central focus of this study revolves around the meticulous prediction of transitions between different postures. This pioneering innovation holds the promise of facilitating seamless and secure transitions, whether in the context of exoskeleton utilization or the synchronization of digital twins in future applications;
- The development of a tailor-made deep learning framework designed for the detection and prediction of human intentions opens vast possibilities in the domains of digital twins, exoskeleton technology, and human intention control.

2. Materials and Methods

2.1. The Experimental Structure

The main objective of this study was to develop a systematic method for accurately detecting changes in human intentions and predicting the transition time using inertial measurement units (IMUs) and artificial intelligence technology. To achieve this objective, the researchers divided the experimental design into three parts.

The first part involved designing a protocol for clinical trials with IMUs. This protocol would specify the number and placement of IMUs on the human body to collect data on the movements and changes in posture. The second part of this study focused on data preparation and treatment to build a better model using artificial intelligence technology. The researchers collected data from the IMUs and used various techniques to preprocess the data, such as filtering, smoothing, and normalization. They then used machine learning algorithms to build a predictive model to accurately detect changes in human intentions. In the third part of this study, the researchers used the deep learning model to predict the transition time of human intentions. This involved analyzing the data collected from the IMUs to detect changes in posture or movements and predicting the time it would take for a human to transition from one intention to another.

2.2. The IMU Equipment

The IMU equipment used in this study is NGIMU from x-io Technologies, Bristol, UK. Each NGIMU contains a 3-axial gyroscope, 3-axial accelerometer, and 3-axial magnetometer, as shown in Figure 1a. The sensor's static roll, pitch, and yaw accuracy are less than 2°, and the maximum sampling rate is 400 Hz [43–45]. The system provided instant sagittal, transverse, and coronal plane angles when our subject wore the sensors on his right leg. The NGIMU would also be used with the iSen 3.0 analysis system from the STT system company, San Sebastián, Spain.



Figure 1. (a) The IMU equipment: NGIMU; (b) the front and back view of the subject wearing IMU sensors; (c) the setting of IMU sensors displayed in the iSen 3.0 analysis system.

2.3. Participants

A convenience sample of 30 healthy subjects was recruited from July 2021 to December 2022, including 16 male and 14 female subjects. None of the subjects had current or previous neurological or orthopedic pathology of the right leg. This study was approved by National Taiwan University Hospital (NTUH) Research Ethics Committee (NTUH IRB approval number: 202209051RINB), and it also was approved by [ClinicalTrials.gov](https://www.clinicaltrials.gov) (NCT number: NCT05650255); All recruited subjects gave their written informed consent for study participation.

Table 1 lists detailed information on the subjects. The subjects' ages ranged between 20 to 32 years old and the ratio of male to female was 1.14, which shows that this study was focused on the young adult stage and was able to reduce the impact of the imbalance between males and females effectively.

Table 1. The information on the 30 healthy voluntary subjects.

	Parameters
Number of subject's sex	Male: 16; Female: 14
The ratio of male to female	1.14
	Mean \pm SD
The mean of subjects' age (years)	26 \pm 5.92 (Range: 22–50)
The mean of subjects' height (cm)	165.57 \pm 8.94 (Range: 142–179)
The mean of subjects' weight (kg)	61.07 \pm 12.34 (Range: 39–95)
Lengths of right thigh of subjects (cm)	41.43 \pm 5.43
Lengths of right calf of subjects (cm)	40.37 \pm 4.35
Lengths of right foot of subjects (cm)	24.53 \pm 1.94

2.4. The Experimental Protocol for Clinical Trials with IMU Equipment

A detailed protocol was designed for this experiment to establish a smooth and accurate measurement of each subject. First, subjects were asked to wear four IMU sensors on their right leg, as shown in Figure 1b. The locations for the four sensors were on the sacrum, right thigh, right calf, and right foot. Second, the IMU sensors and iSen 3.0 analysis program were turned on to ensure successful signal capturing, as shown in Figure 1c. This system allows a selection of sampling rates, such as 25, 100, 200, or 400 Hz. This study used 25 Hz and 100 Hz. Afterward, the subjects were required to stand straight for a while to establish the initial state. Next, subjects were asked to perform a series of "stand-walk-stand-sit" cycles for five to six minutes. "Stand" in this scenario meant that the subject stood straight in his initial state for 15 s, "Walk" meant that the subject walked around the room for 15 to 20 s at their usual pace, and "Sit" meant that the subject sat on a chair for 15 s. After completing the cycle, the data were ready for analysis. The skeleton

pictures of the right foot of the subjects were also displayed on the iSen 3.0 system when the subject performed various movements, as shown in Figure 2.

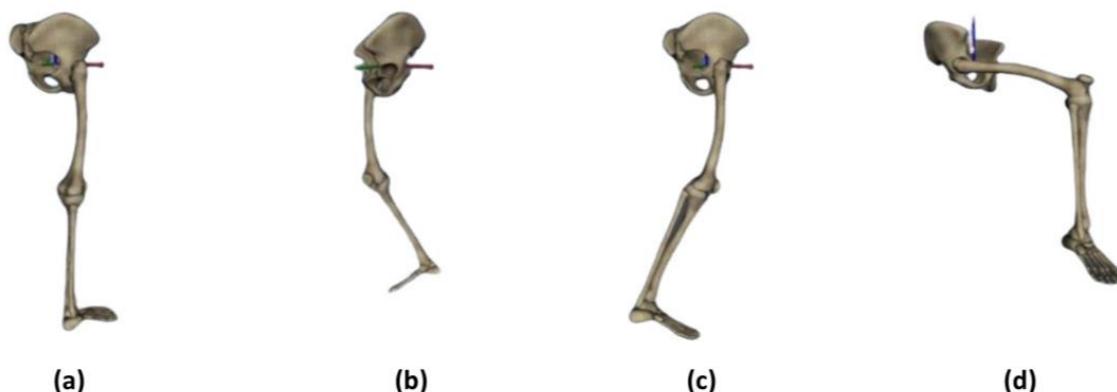


Figure 2. Skeleton pictures of the right foot of the subject were displayed on the iSen 3.0 system when the subject performed different human actions: (a) standing, (b) walking 01, (c) walking 02, (d) sitting.

2.5. Data Processing

This research developed the necessary deep learning programs using Python to treat and extract the features from the IMU data. The program was written in Python, and PyCharm Community Edition was used as the integrated development environment (IDE). The data processing used an Intel® Core (TM) i7-11700F processor (2.50 GHz) running Win10. As depicted in the previous section, four IMU sensors were placed on the sacrum, right thigh, right calf, and right foot during the experiments to obtain the flexion and extension angles of the subject in his sagittal plane. The iSen 3.0 analysis system translated the collected data above into the joint angles. The system also computed the first and second derivatives to obtain the angular velocity and angular acceleration for the use of more potential features. Also, the validity and repeatability of these parameters from IMU sensors and the iSen 3.0 system were verified using a motion-capture system [46]. These variables comprised the training input signals during the training process.

2.6. Posture-Change Detection Algorithm

In the experiment-procedure step, the IMU sensors worn by the subjects provided the main features of the subjects' sagittal plane movement from the pelvis, the right hip, the right knee, and the right ankle: angular displacement, angular velocity, and angular acceleration, which were the joint parameters relative to the pelvis. Thus, there are a total of 12 data signals in each set. These signals were used as the input for the deep learning network structure for the training process. The network output is then the seven labels of human actions: standing, walking, sitting down, sitting, standing up, starting walking, and stopping walking. Also, for the motivation and application of the generalized deep learning model, the data from Subject 1 with a total duration of 3630 s was chosen to be the training dataset in this study.

This research considered various network components and combinations to identify the most suitable algorithm model for the best intention-identification performance. A critical criterion for the network was that it needed to allow sequential input of the time-domain data so that the system could detect the motion change as the pilot moved around. To achieve this criterion, the authors proposed two choices for network structures: a linear Feedforward Neural Network (FNN) structure and a Long Short-Term Memory (LSTM) network structure. The FNN contained three fully connected layers and a SoftMax layer, as shown in Figure 3a. The network took a matrix with the 2 dimensions of $n \times 12$ as an input signal because the data sequence contains three features of the four joints and n was the test time of about 5–6 min. Also, the output of the AI network obtained a matrix with the

2 dimensions of $n \times 7$. The fully connected layers were each composed of seven nodes. The Softmax layer also had seven nodes for the seven labels/classifications of human motions. The outcome of the fully connected layers was then processed by a Softmax layer, which rescaled the outputs to fall within the range of [0, 1]. The Softmax function was designed with the constraint that the sum of the output values must equal 1, thereby producing a probability distribution over the possible output labels. Consequently, this criterion ensures that the output corresponds to a single output label with a probability distribution that conforms to the nature of the classification task.

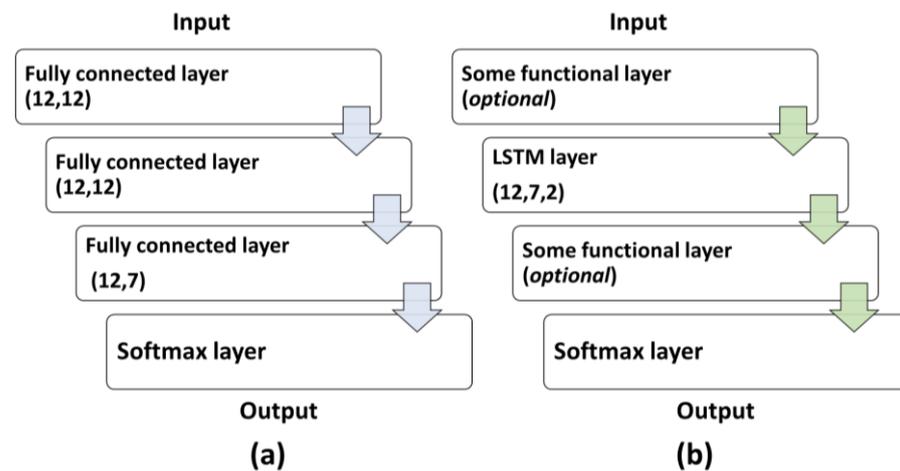


Figure 3. (a) The structure of the FNN deep learning system; (b) the general structure of the LSTM deep learning system.

To further examine the possibility of improving performance, this research also tried a 2-layer LSTM structure for comparison (Figure 3b). The reason for adopting the 2-layer LSTM is that it could minimize effort for application on a mobile device with limited computing power. There is also the consideration that the LSTM networks for two layers could decrease the complexity for a regular PC to process.

Various function layers were added before and after the LSTM layer to improve identification accuracy. First of all, the settings for dropout layers reduced the chance of overfitting. The probability of forcing the input to be zeroed (p values) could be set to either 0.2 or 0.8. Secondly, the settings for batch-normalization layers could help avoid unbalanced weight growth. Thirdly, adding ReLU layers could help eliminate the negative terms.

Table 2 lists the ablation study of the LSTM layer groups tested in our design, in which the number in each group stands for the order in which it was placed in the deep learning network. The ratio between the training group and validation group in the training dataset during all network training was 7:3 in this study. In addition to the various LSTM structures, this research also included FNN to compare with the LSTM performances.

Table 2. The ablation study of the LSTM layer groups in this study.

	Group 1	Group 2	Group 3	Group 4	Group 5
Dropout layer ($p = 0.2$)	1				
Dropout layer ($p = 0.4$)		3	1		
Dropout layer ($p = 0.5$)			5		
Dropout layer ($p = 0.8$)	6			2	4
LSTM layer	2	2	2	1	1
Fully connect layer 1	3		3		2
Fully connect layer 2	7	4	6	3	5
Batch Normalization layer	4	1			3
Leakly ReLu layer	5		4	4	6
Softmax layer	8	5	7	5	7

2.7. Performance Comparison and Transition-Time Analysis

Once the candidate structures had been chosen, the next step was to determine which models would yield better results. In this stage, training loss is a good indicator for gauging the performance of the deep learning structures. The classification functions, such as the sigmoid and Softmax functions or mean square deviation can still impose a high computational load. Alternatively, a relatively straightforward and efficient calculation can be achieved using a cross-entropy classification function. Therefore, this study adopts cross-entropy as the training-loss function. The formula for multiple-sample cross-entropy calculations is as follows:

$$H(p, q) = - \sum_{i=1}^m \sum_{j=1}^n p(x_j) \cdot \ln q(x_j) \tag{1}$$

where p and q were the true label distribution and the predicted label distribution, respectively, and m and n were the sample size and the number of classes.

Once an effective AI model was selected and trained, it was utilized as the primary algorithm. When input signals for a given case were prepared for the AI model, it generated a series of predictive labels. These outputs were subsequently transmitted to the MATLAB environment (MATLAB 2019b) for post-processing and data analysis to determine the intention-transition time. Figure 4 illustrates the complete operational flowchart of this study, encompassing both the experimental and analytical components.

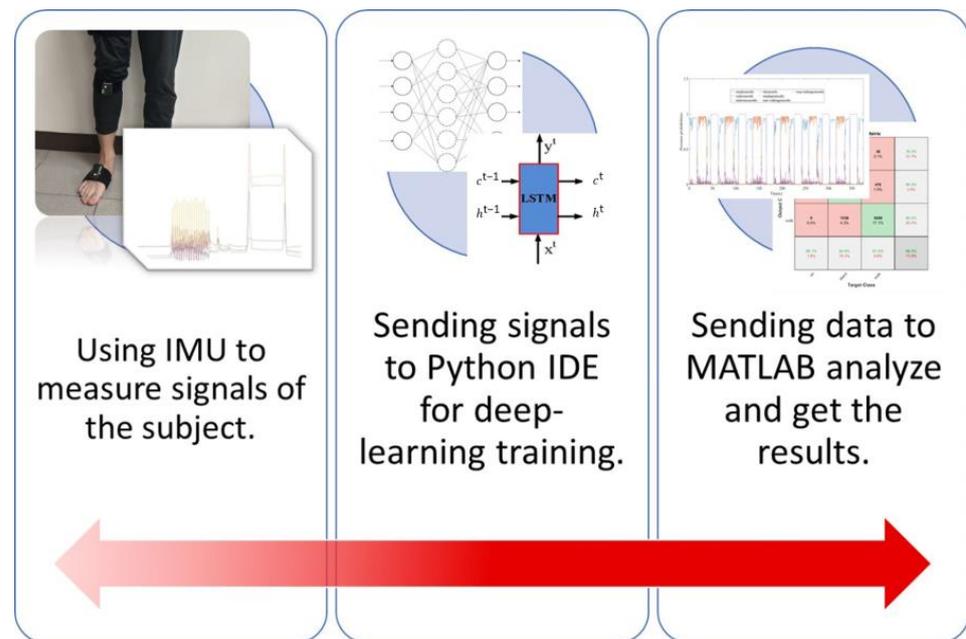


Figure 4. The complete operating flowchart in this study.

The MATLAB environment also determined how fast the network could detect human intention change. The performance was defined as the proportion of the time required for the system to identify a change. The accuracy was calculated with the following equation.

$$\text{Accuracy for intention transition time prediction} = 1 - \frac{\sum \Delta T_{stand/walk/sit} + \sum \Delta T_{start-walk/sit} + \sum \Delta T_{end-walk/sit}}{\text{Total experiment protocol time}} \times 100\% \tag{2}$$

Again, $\Delta T_{stand/walk/sit}$, $\Delta T_{start-walk/sit}$, and $\Delta T_{end-walk/sit}$ were the absolute value of the time difference between the true label situation, such as standing, walking, sitting down, sitting, standing up, starting walking, and stopping walking, and the predicted label situation, respectively.

3. Results

3.1. The IMU Signal Characteristics

Figure 5 shows the IMU recorded data signals from the subject's pelvis, hip on the right side, knee on the right leg, and ankle on the right leg. One can clearly observe specific signal patterns corresponding to different movements of the subject. The IMU signals' distinctive nature is suitable for human posture-change detection.

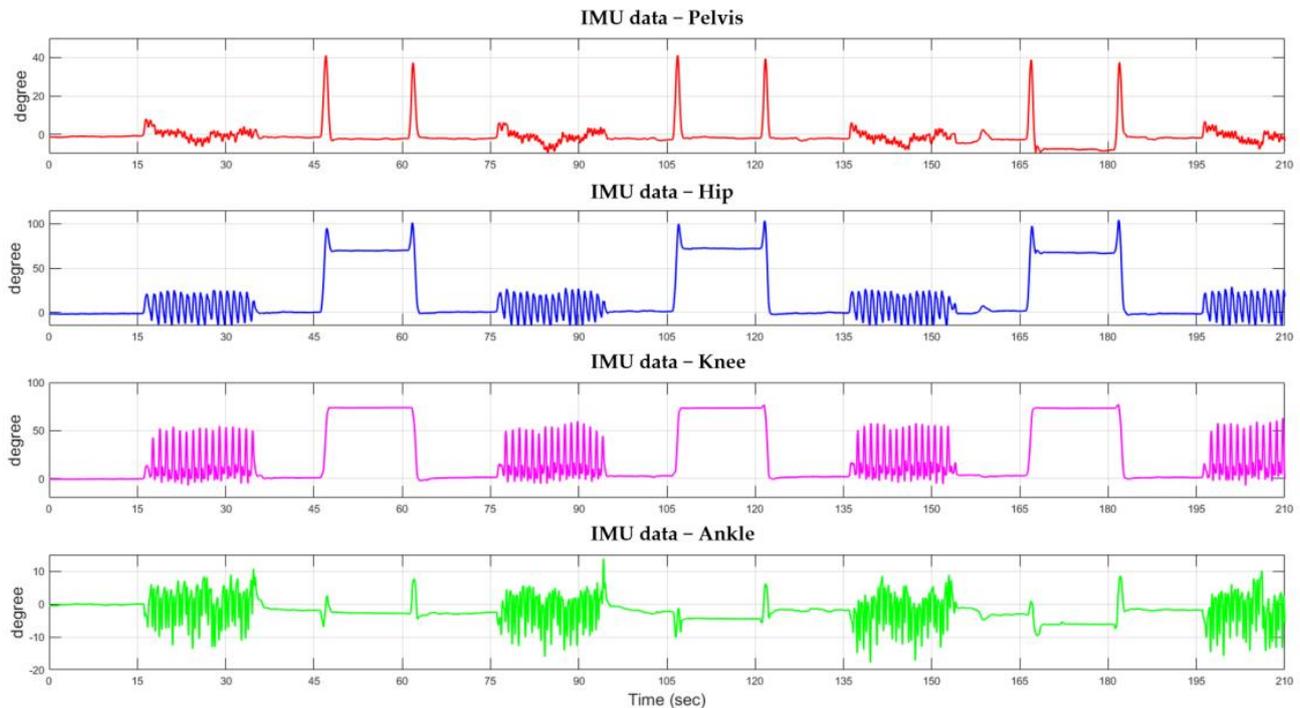


Figure 5. The specific tendency of IMU signals changed when the subject moved.

In this study, through the IMU sensor and iSen 3.0 software system, the IMU input signal is made up of the angle of joints of the lower limbs, such as the hip, knee, and ankle joints, during motions. Table 3 shows the results of Pearson's correlation coefficient of the input parameters from IMU sensors with each other, analyzed by SPSS statistic software. The main results indicated that pelvis parameters would be correlated with the parameters of the hip and knee. Also, the parameters of the hip would be correlated with those of the knee, and the parameters of the ankle are only connected with themselves. Meanwhile, it can also be seen obviously from the content of this table that the joints are affected by each other, and the higher the body part is, the more joints it involves. This would also be consistent with general logical thinking.

3.2. Data and the Design of Test Posture-Change Sequence

This research used two sets of data from Subject 1, for which sampling was both 25 Hz and 100 Hz, with a total duration of 3630 s to train the proposed deep learning detection networks.

As shown in Figure 6, the signal responses from the different movements were quite distinctive. In the process of labeling, the author used motion-capture technology in the form of a Vicon device to ensure the validity and objectivity of labeling by testing the timestamp from the knee-angle data of IMU sensors, whose sampling rate was 100 Hz, and the Vicon 3D capture system (whose sample rate is 120 Hz) when the subject stood, started to walk, and stopped walking (the timestamp starts from 0 s). When the human body posture started to change, one author himself recorded the frame of Vicon at the moment from which the timestamp can be computed. At the same time, the angle data from IMU sensors was also changed, which one author also recorded and labeled himself. It can be found that the transition time from the two types of device data when posture changes is almost the same and manual labeling for supervised learning was possible, as shown in Table 4 and Figure 7. So, the later experiment would involve three types of motion: standing, walking, and sitting; however, to enable transition detection, we decided to separate the movements into seven events to enable posture-transition detection. The labeling was carried out manually. The seven labels include: standing, walking, sitting down (“sd”), sitting, standing up (“su”), starting walking (“stw”), and stopping walking (“spw”). The sample size for each label above after processing is 157,181, 87,073, 6006, 94,580, 6106, 6055, and 6051, which are considered imbalanced data; however, they can be applied properly, such as by adding appropriate functional layers to the subsequent network architecture. From another perspective, the data of those small sample sizes can serve as outliers, which stand for posture-change time detection, which is the thing that this study focuses on.

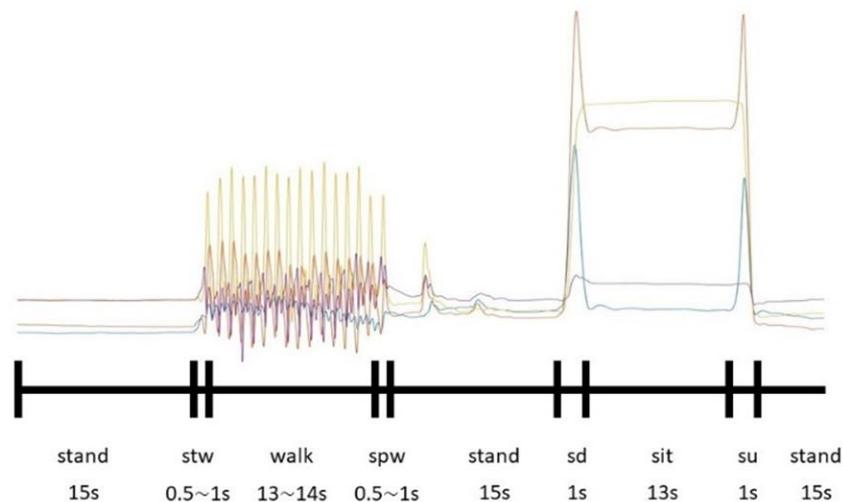


Figure 6. The specific tendency of IMU signals changed when a subject moved. The curves with different colors represented angles with different joints.

Table 4. The time label from the Vicon device and IMU sensors (the timestamp of start to walk and stop to walk). The experiment was performed in triplicate for the same subject.

	The Time Label from IMU Sensors (Unit: s) (Mean ± S.D.)	The Time Label from the Vicon Device (Unit: s) (Mean ± S.D.)
Start to walk	7.4150 ± 0.473	7.65826 ± 0.709
Stop to walk	11.9350 ± 0.544	11.16389 ± 0.359

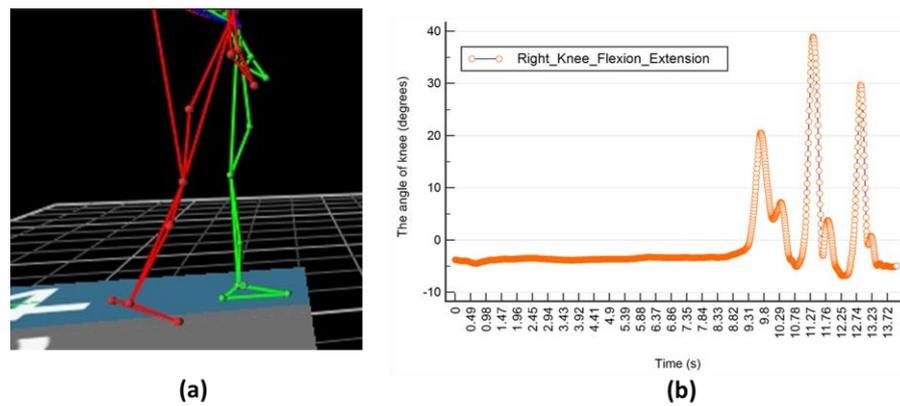


Figure 7. Schematic diagram of kinematic parameters. (a) obtained by Vicon device; (b) obtained by IMU sensors.

3.3. The Comparison of Deep Learning Structures between FNN and LSTM Structures

In this section, two primary deep learning structures, FNN and LSTM, were examined. The first dataset comprised data with sampling rates provided by the iSen 3.0 system, with a sampling rate of 25 Hz, and this study included testing on three subjects (Subject 01 to Subject 03). The LSTM structure was selected from Group 5 in Table 2.

Table 5 shows the computing time of the two deep learning structures. There were no significant differences in the computer time for MATLAB analysis, regardless of whether the structure type was FNN or LSTM (about 3–4 s, accounting for 50% of the total operating time, which included the running time of the AI model for intention-transition-time probability prediction with Python IDE and the running time with MATLAB analysis). However, for the LSTM structure, the training time in Python IDE was 3.6 times that of the FNN structure. The results of the two different structures are shown in Figures 8 and 9, and the accuracy of intention-transition time was compared.

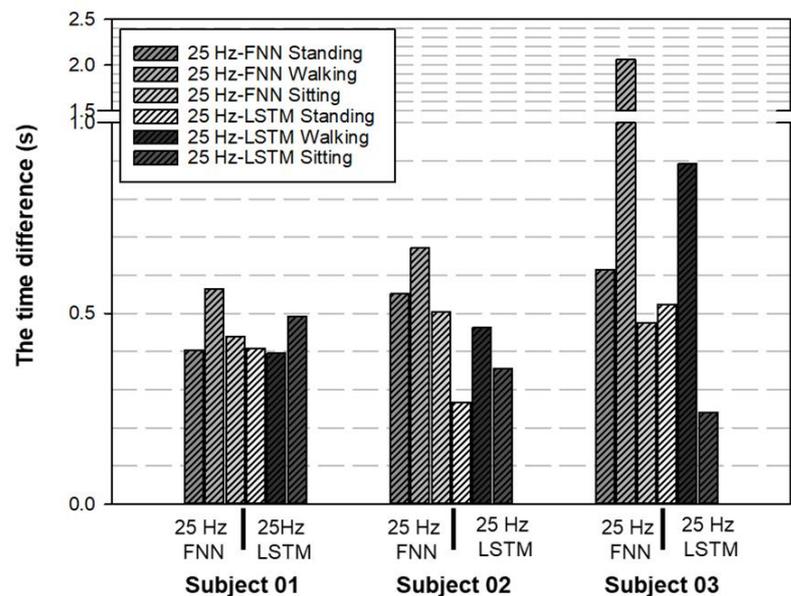


Figure 8. The performance of the time difference (ΔT_{stand} , ΔT_{walk} , and ΔT_{sit}) between FNN and LSTM structure. The stability of the LSTM structure seemed better than that of the FNN for the distributions of the time difference in the case of subject 02 and subject 03.

Table 5. The basic time information with two deep learning structures.

Training Dataset	Sample Rate (Unit: Hz)		Experiment Time (Unit: s)		
	25		3630		
	Deep Learning Type	Sample Rate (Unit: Hz)	Experiment Time (Unit: s)	Training Time in Python IDE (Unit: mins)	MATLAB Analysis Time (Unit: s)
Subject 01	FNN	25	375	24.60	3.804
Subject 02			315		3.888
Subject 03			315		3.789
Subject 01	LSTM	25	375	88.80	3.731
Subject 02			315		3.856
Subject 03			315		3.905

About 50% of the total operating time, which included the running time of the AI model with Python IDE and the running time with MATLAB analysis.

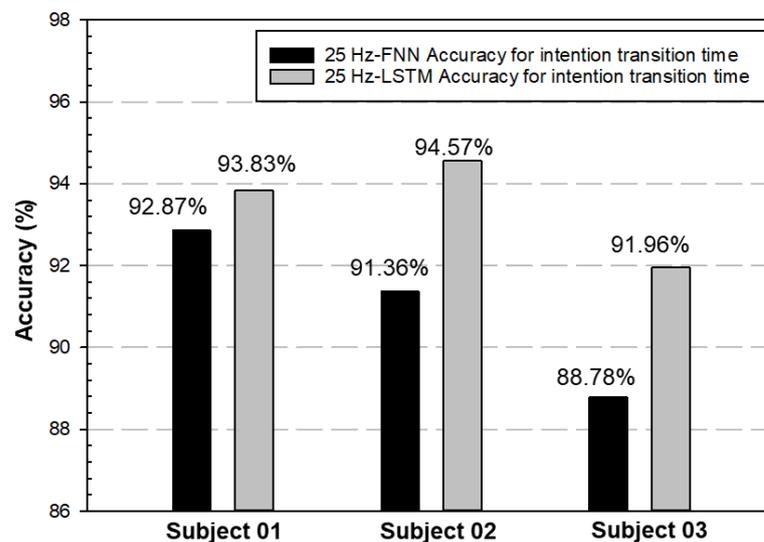


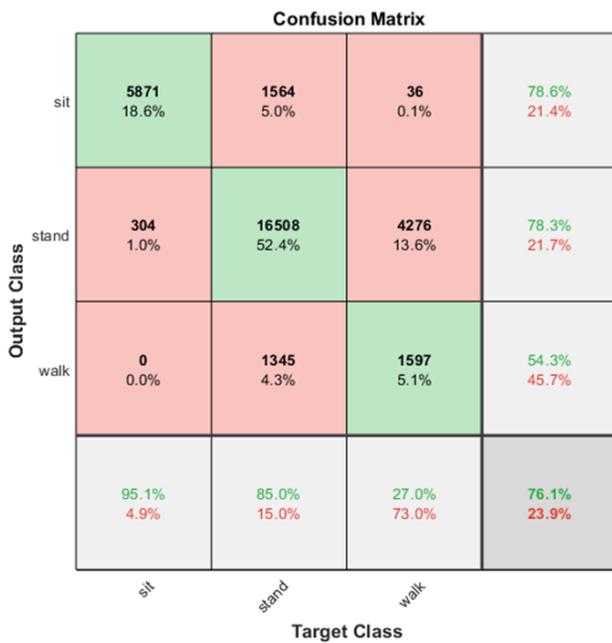
Figure 9. Both methods achieved better results for the measurement of intention-transition time, which ranged from 88% to 94%.

From Figure 9, both methods achieved better results for the accuracy of the intention-transition time, which ranged from 88% to 94%; however, the stability of the LSTM structure was better than the FNN, for the distributions of the identification time.

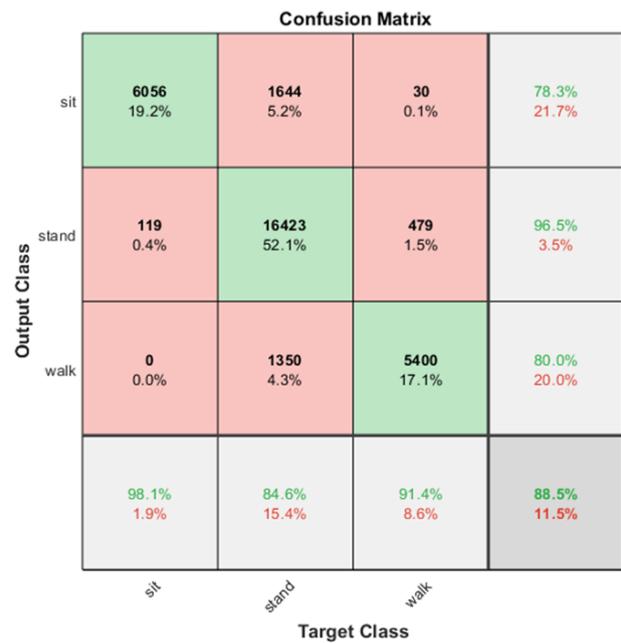
On the other hand, to obtain more precise information, the angle data of IMU sensors with a 100 Hz sample rate are used. In this criterion, the suitable one is still the LSTM structure and its performance is also better for posture-identification accuracy, as shown in Figure 10. As a result, comparing the FNN and LSTM, the latter is faster and more accurate in detecting the motion-switching time. As a result, the LSTM structure with a 100 Hz sampling rate was chosen in this study for later analysis.

3.4. Effect of the Sampling Rates

To examine the effect of the different sampling rates further, the authors also studied the performance of the LSTM network with sampling rates of both 25 and 100. Three subjects (Subject 01 to Subject 03) were chosen for the test again. Figure 11 shows that the performance of the time difference (ΔT_{stand} , ΔT_{walk} , and ΔT_{sit}) with a sampling rate of 100 Hz was better than that of 25 Hz. In other words, the time difference was smaller with a 100 Hz sampling rate. Figure 12 shows that a higher sampling rate also leads to higher accuracy in both the transition detection and motion classification, 91% to 95% and 37% to 88%, respectively. Thus, the selected LSTM structure in Section 3.4 with 100 Hz sampling was confirmed again and chosen for later studies.

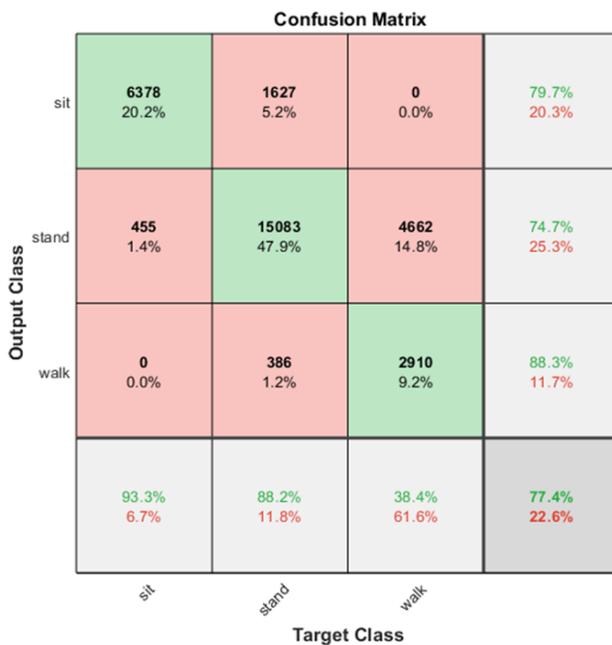


Subject 01 – FNN structure

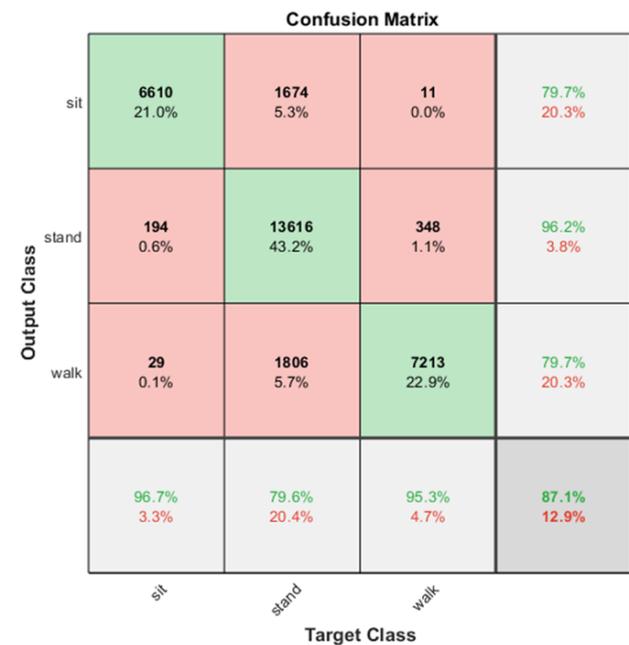


Subject 01 – LSTM structure

(a)



Subject 02 – FNN structure



Subject 02 – LSTM structure

(b)

Figure 10. Cont.

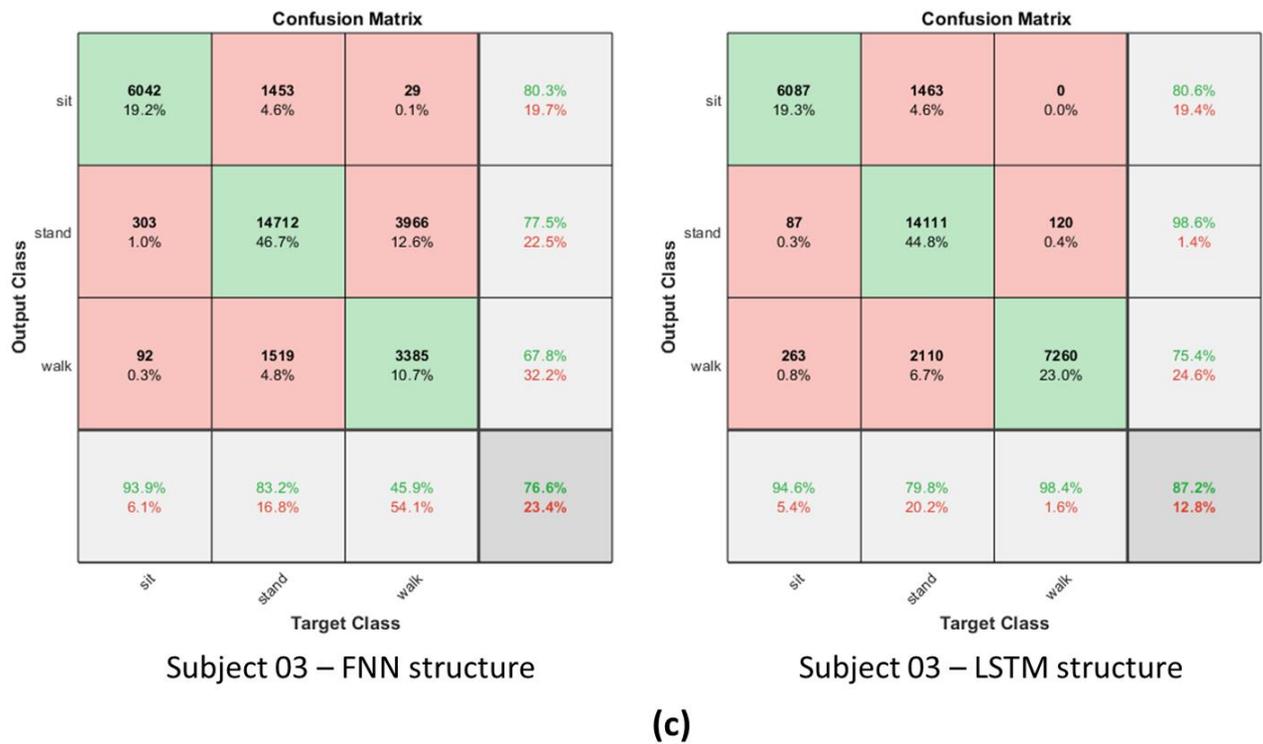


Figure 10. The confusion matrix results in the FNN structure and LSTM structure. (a) Subject 01; (b) Subject 02; (c) Subject 03. (Sampling rate is 100 Hz).

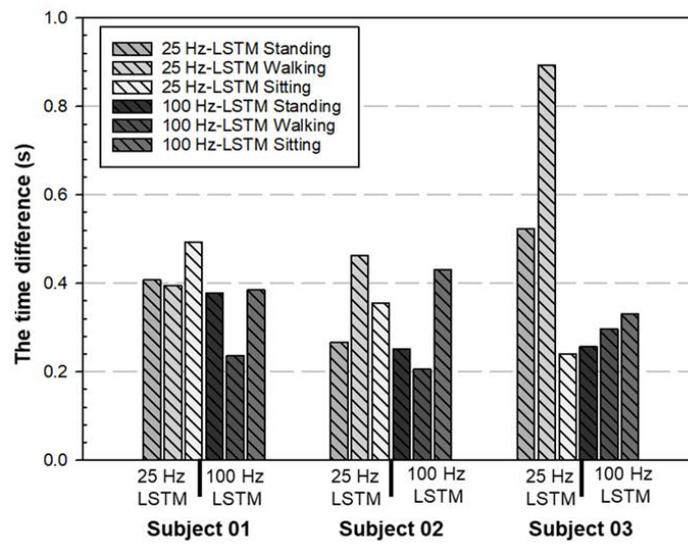


Figure 11. The performance of the time difference (ΔT_{stand} , ΔT_{walk} , and ΔT_{sit}) between different sample rates with LSTM structure. The time difference was decreased when the sample rate was 100 Hz.

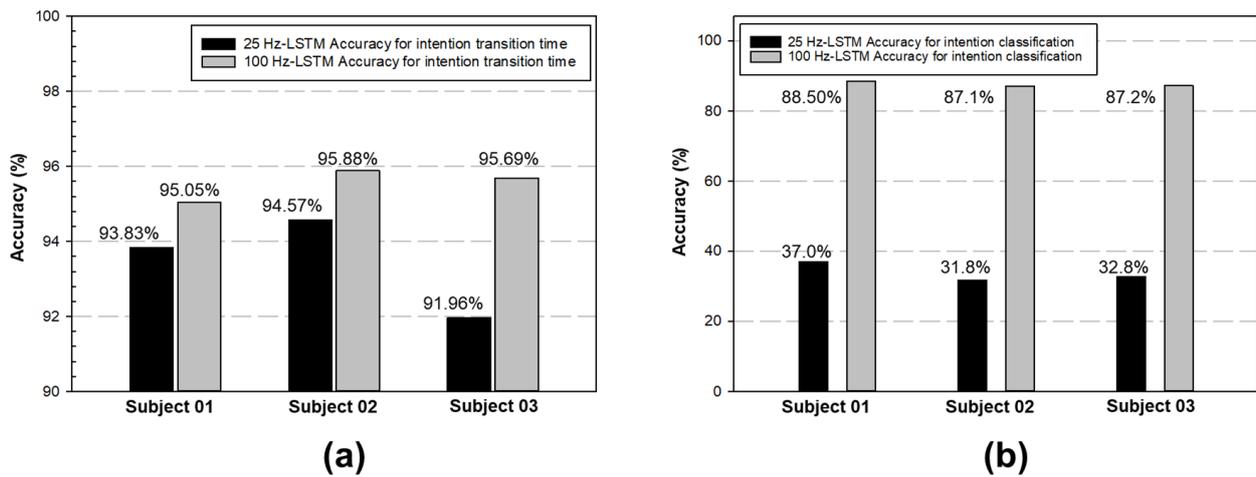


Figure 12. (a) The outcome, with both sample rates achieving good results for the measurement of intention-transition time, which range from 91% to 95%.; (b) The stability of the LSTM structure with a 100 sample rate seemed better for the accuracy of intention classification.

3.5. The Comparison of Different LSTM Structures' Performance and Decisions of Deep Learning Models

After the deep learning type was selected, the next step was to find out which LSTM structure was better for human posture-change detection. As mentioned in Section 2, this study compared five LSTM candidate structures for ablation studies. The loss values at the 150th epoch with the cross-entropy method, the value after validation, and the value at the prediction stage were all recorded to determine which structure achieved the best performance and the outcome of the ablation studies.

As shown in Table 6, the loss of Group 5 achieved the minimum training loss, validation loss, and average prediction loss of 0.3112, 0.3141, and 0.3995, respectively. As a result, it was appropriate to choose the network of Group 5 as the identification network. The parameters for the selected structure are also listed in Table 6. Figure 13 shows the tendencies of training loss for the five LSTM groups in the ablation studies. From the outcome of Figure 13, it can be found that if Group 1 was the baseline group and the tendency of the training loss was the performance index for ablation studies, the performance of Group 2, Group 3, and Group 4 was also not better than that of Group 1 and the performance of Group 5 was better than that of Group 1.

Table 6. The performance comparisons between different LSTM group conditions.

Parameters	Learning Rate: 0.001		Weight Decay: 0.00001		Loss Function: Cross-Entropy Use One GPU Device
	Epochs: 150		Batch Size: 10		
Training Dataset	The Number of Features: 12			The Number of Labels: 7	
	The Number of Samples: 217,880 Samples from One Person. (3630 s, Whose Sample Rate Was 100 Hz)				
Group Types	Training Stage		Prediction Stage	Training Time in Python IDE (mins)	
	Training Loss	Validation Loss	Average Prediction Loss		
Group 1	0.3082	0.3101	0.4109	206	
Group 2	1.2410	1.2409	0.6357	229	
Group 3	1.7446	1.9525	0.4520	241	
Group 4	0.3980	0.3982	0.4362	244	
Group 5	0.3112	0.3141	0.3995	249	

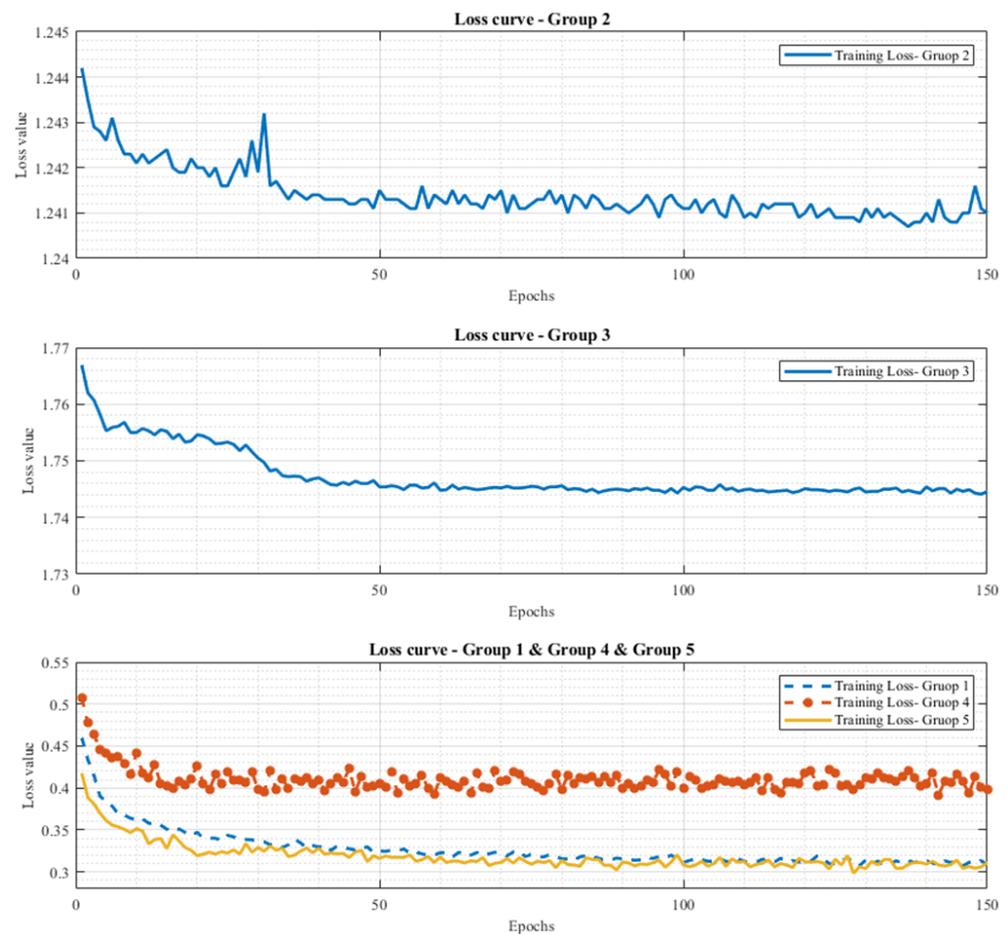


Figure 13. The tendencies of training loss for five LSTM group conditions in the ablation studies.

3.6. The Identification Performance of Human Posture-Transition Time among Different Subjects

Once the main LSTM algorithm was determined, one could test the network's identification ability on the 30 subjects. Again, the deep learning type was an LSTM network, and the training dataset was from only the data of Subject 1. An additional test was to see whether one could use the same pre-trained network for all the subjects. The sampling rate of all tests was 100 Hz.

Figure 14 shows the schematic diagram of the test results. The input to the network was the total input variables, as mentioned in Section 2. The output results represented the probability of the identified action (movement) at each time instance. The action with the highest probability value would be considered the recognized human movement at that time. Recall that the primary purpose of this study was to achieve real-time human posture-change identification and to study the possibility of using only the IMU sensor signals to detect human intention when a posture changes. Therefore, this study paid particular focus to the time required for the network to identify the transition. In particular, this study also tried to identify the transition between consecutive postures.

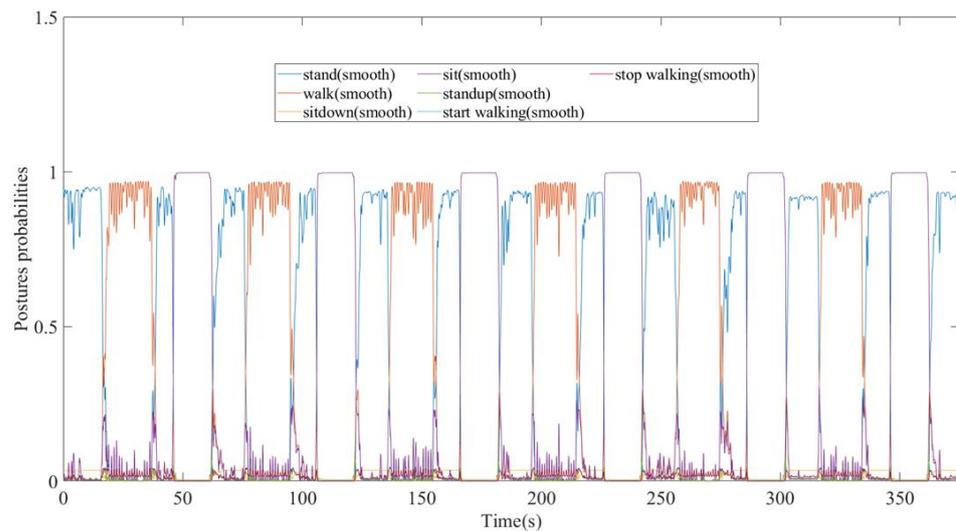


Figure 14. The schematic diagram of the output results from the LSTM model, which represent the probability of each specific intention action vis-à-vis time.

Figure 15 shows the statistical analysis of the variation in the identification time. The figure shows the statistics of the thirty subjects on the identification-time differences of the seven postures: “standing”, “walking”, “sitting”, “stop standing and start walking”, “stop standing and starting sitting”, and “stop sitting and start standing” ($\Delta T_{stand/walk/sit}$, $\Delta T_{start-walk/sit}$ and $\Delta T_{end-sit}$). One notices that the identification time for the postures was all less than 0.5 s (0.4255 s, 0.39 s, 0.429 s, 0.479 s, 0.405 s, and 0.3188 s, respectively). The figure also shows a significant difference in the statistics under the t-test if the analysis of the posture was based on “stop walking and start standing”. One also observes significant differences in the t-test result for all postures when the comparison was based on “stop sitting and start standing” except for “stop standing/walking and start walking/standing”. The result indicated that the transition identification of “stop sitting and start standing” was close to the ground truth. The detailed results are listed in Table 7.

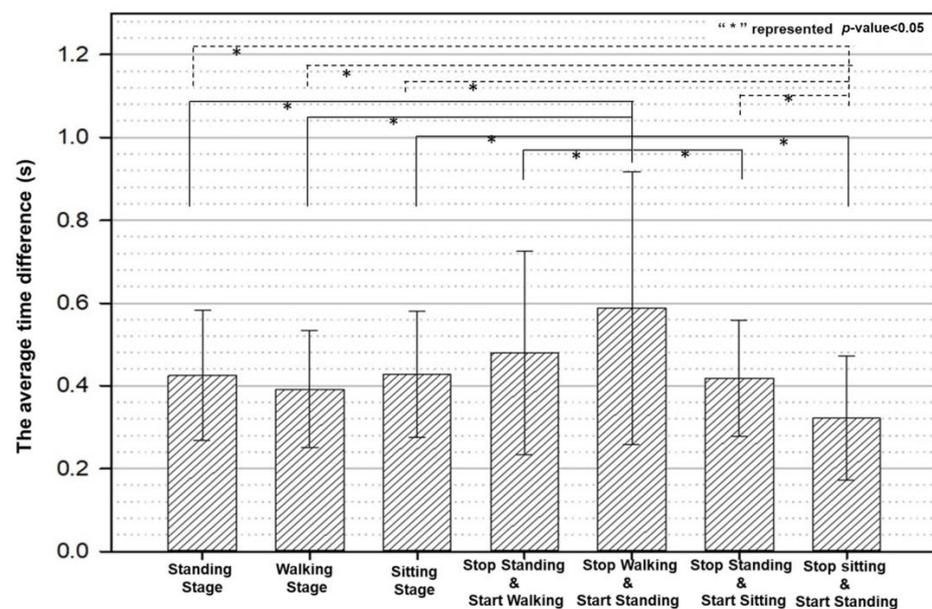


Figure 15. The time difference outcome among 30 different subjects in this study. (Average value \pm standard deviation).

Table 7. The statistical outcome information from the LSTM model between 30 different drivers in this study.

Software Operating Time	Mean ± SD
Python algorithm with LSTM model for intention-transition-time probability prediction (s)	6.76667 ± 1.30472 (80.37% of operating time)
Analysis with MATLAB for intention-transition-time prediction (s)	1.65217 ± 0.71406 (19.63% of operating time)
Statistic results of the time difference when human posture changed	Mean ± SD
In the standing stage (s)	0.42557 ± 0.16049
In the walking stage (s)	0.39019 ± 0.14458
In the sitting stage (s)	0.42941 ± 0.15568
Stop standing and start walking (s)	0.47986 ± 0.25106
Stop walking and start standing (s)	0.59732 ± 0.33405
Stop standing and start sitting (s)	0.40483 ± 0.12591
Stop sitting and start standing (s)	0.31880 ± 0.15251
Accuracy for posture-change detection	Mean ± SD
Accuracy for prediction of intention classification (%)	85.87 ± 7.49
Accuracy for prediction of intention-transition time (%)	94.44 ± 1.70

In addition to the transition-time values, Table 7 also shows the identification accuracy. The definition of identification accuracy is in Section 2. This research also utilized the confusion matrix for all drivers (participants) to judge the identification accuracy. From the accuracy results in Table 7, the network achieved 94.44% accuracy for transition time and 85.87% accuracy for posture-change detection.

4. Discussion

Instead of just identifying the posture, this research focuses on how fast a network can detect a posture transition, a function essential to human compliance robotics. The main goal is to catch the instant when humans intend to change gait or posture using IMU sensor signals. In contrast to similar studies, this research strongly emphasizes pinpointing the moment when a person's intention shifts, with the aim of facilitating the adequate switching of control laws in exoskeletons. What is more, these pieces of information on human intention cannot be obtained directly from the exoskeleton motor.

This study identified standing, walking, and sitting as the three main scenarios for exoskeleton initial activities. We also stress the importance of accurately detecting the instant when a human initializes a posture change. Initial observation of the experimental results reveals that the changes in the flexion and extension angles were not immediate when the subjects switched between standing and walking. The response of the general sagittal plane flexion and extension angle (Figure 6) allows for an easy distinction between "sd"—sitting down—and "su"—standing up. Here, "sd" means that the subject changed posture from standing to sitting, while "su" means the subject stood up from a sitting posture. Both movements take about one second and show quite distinctive characteristics compared to "sitting still." In particular, this research separated the transition phase as separate postures: "stw"—starting walking—and "spw"—stopping walking—for precise identification of the moment when posture changes. Here, "stw" means the subject changed his posture from standing to walking, while "spw" means the subject slowed down from the walking state. Both movements took about half a second to one second for all the subjects. For the same reason, we have added the labels "sitting down", "standing up", "start walking", and "stop walking". It is worth pointing out that these movements also correspond to very distinctive joint-torque characteristics.

This research compared FNN and LSTM as candidate networks and defined the difference between the actual posture-switching time and the identified switching time as the detection time. FNN required about 0.6 to 1.9 s to detect a transition into walking, while LSTM required about 0.4 to 0.8 s, as shown in Figure 8, which used the same case data with a 25 Hz sampling rate. As a result, LSTM could detect the transition moment much faster. The fastest it could detect a transition is 0.24 to 0.49 s in detecting a sitting down movement. Also, LSTM was a recurrent neural network (RNN) suitable for exoskeleton controller applications.

There was a substantial improvement in the detection accuracy when the sampling rate of the angle information of the IMU sensors increased from 25 to 100 Hz (Figure 12). As the dataset from different test subjects increased, we saw that the network trained for one subject could detect posture changes in all subjects equally fast, whether they be male or female. The average detection time was 0.4 s with a p -value of 0.05 (Figure 15 and Table 7). Although the response time was not instantaneous, it was fast enough for practical purposes. The test result showed that the system proposed in this study could effectively reflect the pilots' initial intention to change posture and that training a preliminary controller network that already performed well for general use was possible.

Compared to previous research in this field, it has been demonstrated that linear discriminant analysis (LDA) and random forest (RF) can effectively predict a subject's intention to initiate movement. Notably, the average prediction transition time was approximately 31×10^{-4} seconds and 11×10^{-5} seconds, respectively [31,33]. It is important to note that these studies predominantly focused on the upper limbs and arms only, characterized by relatively simple joint freedom and kinematic parameters.

On the other hand, a separate study indicated that the prediction times for transitions varied from 277 ms to 488 ms when considering walking as the reference state [43]. However, it is worth mentioning that achieving such results necessitated the integration of various sensors, including IMU, EMG, and other types of sensors. In contrast, this study achieved an average prediction time of approximately 400 ms for lower limb action transitions. Remarkably, this was accomplished using just one type of sensor (IMU) in combination with an effective LSTM algorithm. This outcome indicates that comparable results to previous research can be obtained with more streamlined sensor configurations.

Furthermore, based on the results of ablation studies involving various LSTM deep learning structures, it was observed that placing the Dropout layer before the LSTM layer with a p value of 0.4 did not outperform Group 1 (in the case of Group 3). Additionally, when the Batch Normalization layer and Dropout layer were situated before or after the LSTM layer, their performance still did not surpass that of Group 1 (in the case of Group 2) if we consider Group 1 as the baseline group and the trend of training loss as the performance metric for ablation studies. It becomes evident that the primary functions of the Batch Normalization layer and Dropout layer, which include regularization, overfitting prevention, and managing weight imbalances in the network architecture, apply to situations involving imbalanced data. However, the sequence in which these layers are added can significantly influence the final performance index and target-prediction performance. This information would be very useful for the modification of the LSTM network architecture.

To sum up, this study built up a deep learning algorithm and structure using only IMU sensors that started with some initial posture or intention prediction, which proved that AI technology could be further applied to the more complicated field of human posture change or exoskeleton control by human-intention detection.

5. Limitations

The proposed method showed a good outcome in measuring or detecting human intention in this study, but it could be worth further exploring. First, inertial measurement units (IMUs) were used for the current study. Other wearable sensors could be combined with this, such as electromyography (EMG), torque sensors, or the application of sensor-fusion technology with a more modern AI algorithm [47], and the relationship between

the unitary duration of activities and the duration of all activity switches for kinematic modeling in the HAR research field [48] might be included in the next phase of research to obtain more meaningful data. Second, the present research only considered the sagittal plane joint angles of the subjects' right feet. It would be interesting also to consider the influence of the angles in the transverse and the coronal plane. Third, the postures studied were relatively simple; different postures or gaits, actual implementation problems, and combination with an exoskeleton could be considered in future studies to match real-life scenarios and engineering applications.

6. Conclusions

This research proposed two machine learning network structures, a linear Feedforward Neural Network (FNN) and a Long Short-Term Memory (LSTM), to detect the switching of human postures, and three everyday initial human postures were used, including standing, walking, and sitting. The system used only IMU sensors to enable portability. To speed up posture-change detection, the authors introduced the transition stages as distinct features for identification. Comparing the detection results showed that the LSTM model with a 100 Hz sampling rate achieved the best results for posture-change detection time and posture-classification accuracy, about 94.44% and 85.87%, respectively. Even though LSTM took longer to train, it does not affect the implementation requirement.

Further investigation of the sampling-rate effect showed that a faster sampling rate (four times faster) improved the accuracy of posture-change detection from an average of about 91% to 95%. In addition to the network design, this research also showed that using a network trained for one subject was adequate for use on the other subjects. This result suggests the feasibility of selecting personalized or generalized deep learning models in related applications for the field of detecting human posture changes.

Author Contributions: Conceptualization, C.-T.K.; Formal analysis, W.-L.H.; Funding acquisition, J.-Y.Y.; Methodology, C.-T.K.; Resources, W.-L.H. and J.-Y.Y.; Software, C.-T.K.; Supervision, W.-L.H. and J.-Y.Y.; Validation, W.-L.H.; Writing—original draft, J.-J.L.; Writing—review and editing, K.-K.J., F.-C.W., T.-C.T. and J.-Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the National Science and Technology Council, R.O.C. under Project No. 108-2221-E-011 -166 -MY3.

Institutional Review Board Statement: This study was approved by National Taiwan University Hospital (NTUH) Research Ethics Committee (NTUH IRB approval number: 202209051RINB), and it was also approved by [ClinicalTrials.gov](https://clinicaltrials.gov) (NCT number: NCT05650255); all recruited subjects gave their written informed consent for study participation.

Data Availability Statement: Data are available on request due to ethical restrictions. The data presented in this study are available on request from the corresponding author. The data are not publicly available due to regulations of the Research Ethics Committee.

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

References

1. Lara, O.D.; Labrador, M.A. A Survey on Human Activity Recognition using Wearable Sensors. *IEEE Commun. Surv. Tutor.* **2013**, *15*, 1192–1209. [[CrossRef](#)]
2. Preatoni, E.; Nodari, S.; Lopomo, N.F. Supervised Machine Learning Applied to Wearable Sensor Data Can Accurately Classify Functional Fitness Exercises Within a Continuous Workout. *Front. Bioeng. Biotechnol.* **2020**, *8*, 664. [[CrossRef](#)]
3. Dallel, M.; Havard, V.; Dupuis, Y.; Baudry, D. Digital twin of an industrial workstation: A novel method of an auto-labeled data generator using virtual reality for human action recognition in the context of human–robot collaboration. *Eng. Appl. Artif. Intell.* **2023**, *118*, 105655. [[CrossRef](#)]
4. Cangelosi, A.; Invitto, S. Human-Robot Interaction and Neuroprosthetics: A review of new technologies. *IEEE Consum. Electron. Mag.* **2017**, *6*, 24–33. [[CrossRef](#)]
5. Zhang, L.; Liu, G.; Han, B.; Wang, Z.; Zhang, T. sEMG Based Human Motion Intention Recognition. *J. Robot.* **2019**, *2019*, 3679174. [[CrossRef](#)]

6. Lee, C.-Y.; Lan, S.C.; Lin, J.-J.; Lin, Y.-T.; Chiang, P.-S.; Hsu, W.-L.; Jen, K.-K.; Huang, A.Y.S.; Yen, J.-Y. Realization of Natural Human Motion on a 3D Biped Robot For Studying the Exoskeleton Effective. *J. Med. Biol. Eng.* **2021**, *41*, 856–869. [[CrossRef](#)]
7. Lee, J.; Kwon, K.; Yeo, W.-H. Recent advances in wearable exoskeletons for human strength augmentation. *Flex. Print. Electron.* **2022**, *7*, 023002. [[CrossRef](#)]
8. Ding, M.; Nagashima, M.; Cho, S.G.; Takamatsu, J.; Ogasawara, T. Control of Walking Assist Exoskeleton with Time-delay Based on the Prediction of Plantar Force. *IEEE Access* **2020**, *8*, 138642–138651. [[CrossRef](#)]
9. Wang, H.; Schmid, C. In Action Recognition with Improved Trajectories. In Proceedings of the 2013 IEEE International Conference on Computer Vision, Las Vegas, NV, USA, 1–8 December 2013; Volume 2013, pp. 3551–3558.
10. Herath, S.; Harandi, M.; Porikli, F. Going deeper into action recognition: A survey. *Image Vis. Comput.* **2017**, *60*, 4–21. [[CrossRef](#)]
11. Younsi, M.; Diaf, M.; Siarry, P. Comparative study of orthogonal moments for human postures recognition. *Eng. Appl. Artif. Intell.* **2023**, *120*, 105855. [[CrossRef](#)]
12. Bobick, A.F.; Davis, J.W. The recognition of human movement using temporal templates. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 257–267. [[CrossRef](#)]
13. Gorelick, L.; Basri, R.; Blank, M.; Shechtman, E.; Irani, M. Actions as Space-Time Shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *29*, 2247–2253. [[CrossRef](#)]
14. Yan, S.; Xiong, Y.; Lin, D. In Spatial temporal graph convolutional networks for skeleton-based action recognition. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, Louisiana, LA, USA, 2–7 February 2018.
15. Khodabandelou, G.; Moon, H.; Amirat, Y.; Mohammed, S. A fuzzy convolutional attention-based GRU network for human activity recognition. *Eng. Appl. Artif. Intell.* **2023**, *118*, 105702. [[CrossRef](#)]
16. Papi, E.; Bo, Y.N.; McGregor, A.H. A flexible wearable sensor for knee flexion assessment during gait. *Gait Posture* **2018**, *62*, 480–483. [[CrossRef](#)] [[PubMed](#)]
17. Taheri, O.; Salarieh, H.; Alasty, A. Human leg motion tracking by fusing imus and rgb camera data using extended kalman filter. *arXiv* **2011**, arXiv:2011.00574.
18. Ito, T.; Ayusawa, K.; Yoshida, E.; Kobayashi, H. Evaluation of active wearable assistive devices with human posture reproduction using a humanoid robot. *Adv. Robot.* **2018**, *32*, 635–645. [[CrossRef](#)]
19. Zhu, L.; Wang, Z.; Ning, Z.; Zhang, Y.; Liu, Y.; Cao, W.; Wu, X.; Chen, C. A Novel Motion Intention Recognition Approach for Soft Exoskeleton via IMU. *Electronics* **2020**, *9*, 2176. [[CrossRef](#)]
20. Bruinsma, J.; Carloni, R. IMU-Based Deep Neural Networks: Prediction of Locomotor and Transition Intentions of an Osseointegrated Transfemoral Amputee. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2021**, *29*, 1079–1088. [[CrossRef](#)] [[PubMed](#)]
21. Liu, H.; Xue, T.; Schultz, T. On a Real Real-Time Wearable Human Activity Recognition System. In Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies—WHC, Lisbon, Portugal, 16–18 February 2023; Volume 5.
22. Manivasagam, K.; Yang, L. Evaluation of a New Simplified Inertial Sensor Method against Electrogoniometer for Measuring Wrist Motion in Occupational Studies. *Sensors* **2022**, *22*, 1690. [[CrossRef](#)]
23. Liu, H.; Schultz, T. Biosignal Processing and Activity Modeling for Multimodal Human Activity Recognition. Ph.D. Thesis, Universität Bremen, Bremen, Germany, 2021.
24. Harris, E.J.; Khoo, I.-H.; Demircan, E. A Survey of Human Gait-Based Artificial Intelligence Applications. *Front. Robot. AI* **2022**, *8*, 749274. [[CrossRef](#)]
25. Kececi, A.; Yildirak, A.; Ozyazici, K.; Ayluctarhan, G.; Agbulut, O.; Zincir, I. Implementation of machine learning algorithms for gait recognition. *Eng. Sci. Technol. Int. J.* **2020**, *23*, 931–937. [[CrossRef](#)]
26. Hartmann, Y.; Liu, H.; Schultz, T. High-Level Features for Human Activity Recognition and Modeling. In *Biomedical Engineering Systems and Technologies*; Roque, A.C.A., Gracanin, D., Lorenz, R., Tsanas, A., Bier, N., Fred, A., Gamboa, H., Eds.; Springer Nature Switzerland: Cham, Switzerland, 2023; pp. 141–163.
27. Rodrigues, J.; Liu, H.; Folgado, D.; Belo, D.; Schultz, T.; Gamboa, H. Feature-Based Information Retrieval of Multimodal Biosignals with a Self-Similarity Matrix: Focus on Automatic Segmentation. *Biosensors* **2022**, *12*, 1182. [[CrossRef](#)] [[PubMed](#)]
28. Liu, H.; Hartmann, Y.; Schultz, T. Motion Units: Generalized Sequence Modeling of Human Activities for Sensor-Based Activity Recognition. In Proceedings of the 29th European Signal Processing Conference (EUSIPCO), Dublin, Ireland, 23–27 August 2021; pp. 1506–1510.
29. Kong, Y.; Fu, Y. Human Action Recognition and Prediction: A Survey. *Int. J. Comput. Vis.* **2022**, *130*, 1366–1401. [[CrossRef](#)]
30. Wang, J.; Chen, Y.; Hao, S.; Peng, X.; Hu, L. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognit. Lett.* **2019**, *119*, 3–11. [[CrossRef](#)]
31. Chen, K.; Zhang, D.; Yao, L.; Guo, B.; Yu, Z.; Liu, Y. Deep Learning for Sensor-based Human Activity Recognition: Overview, Challenges, and Opportunities. *ACM Comput. Surv.* **2021**, *54*, 77. [[CrossRef](#)]
32. Archetti, L.; Ragni, F.; Saint-Bauzel, L.; Roby-Brami, A.; Amici, C. Inclusive Human Intention Prediction with Wearable Sensors: Machine Learning Techniques for the Reaching Task Use Case. *Eng. Proc.* **2020**, *2*, 13.
33. Ragni, F.; Archetti, L.; Roby-Brami, A.; Amici, C.; Saint-Bauzel, L. Intention Prediction and Human Health Condition Detection in Reaching Tasks with Machine Learning Techniques. *Sensors* **2021**, *21*, 5253. [[CrossRef](#)]
34. Li, S.; Zhang, L.; Diao, X. Deep-Learning-Based Human Intention Prediction Using RGB Images and Optical Flow. *J. Intell. Robot. Syst.* **2020**, *97*, 95–107. [[CrossRef](#)]

35. Lindemann, B.; Müller, T.; Vietz, H.; Jazdi, N.; Weyrich, M. A survey on long short-term memory networks for time series prediction. *Procedia CIRP* **2021**, *99*, 650–655. [[CrossRef](#)]
36. Lin, B.; Bouneffouf, D.; Cecchi, G. In Predicting Human Decision Making with LSTM. In Proceedings of the 2022 International Joint Conference on Neural Networks (IJCNN), Padua, Italy, 18–23 July 2022; pp. 1–8.
37. He, J.; Guo, Z.; Shao, Z.; Zhao, J.; Dan, G. An LSTM-Based Prediction Method for Lower Limb Intention Perception by Integrative Analysis of Kinect Visual Signal. *J. Healthc. Eng.* **2020**, *2020*, 8024789. [[CrossRef](#)]
38. Ren, B.; Zhang, Z.; Zhang, C.; Chen, S. Motion Trajectories Prediction of Lower Limb Exoskeleton Based on Long Short-Term Memory (LSTM) Networks. *Actuators* **2022**, *11*, 73. [[CrossRef](#)]
39. Kang, K.-Y.; Lee, S.-G.; Kang, H.; Kim, J.-G.; Tack, G.-R.; Choi, J.-S. A Pilot Study of the Efficiency of LSTM-Based Motion Classification Algorithms Using a Single Accelerometer. *Appl. Sci.* **2022**, *12*, 7243. [[CrossRef](#)]
40. Peng, Y.; Gan, C.; Zhang, Z. A survey of feature extraction methods in human action recognition. *Comput. Appl. Software* **2022**, *39*, 8.
41. Liu, Y.X.; Wang, R.; Gutierrez-Farewik, E.M. A Muscle Synergy-Inspired Method of Detecting Human Movement Intentions Based on Wearable Sensor Fusion. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2021**, *29*, 1089–1098. [[CrossRef](#)]
42. Yin, P.; Yang, L.; Yang, M. In Research on Recognition of Human Motion State Based on Force and Motion Sensor Fusion. In Proceedings of the 2022 IEEE 2nd International Conference on Power, Electronics and Computer Applications (ICPECA), Shenyang, China, 21–23 January 2022; pp. 238–242.
43. Tang, H.-Y.; Tan, S.-H.; Su, T.-Y.; Chiang, C.-J.; Chen, H.-H. Upper Body Posture Recognition Using Inertial Sensors and Recurrent Neural Networks. *Appl. Sci.* **2021**, *11*, 12101. [[CrossRef](#)]
44. Cui, J.; Li, Z. Prediction of Upper Limb Action Intention Based on Long Short-Term Memory Neural Network. *Electronics* **2022**, *11*, 1320. [[CrossRef](#)]
45. Gibb, K.; Yovchev, N.; Aubertin, C.; Greenwood, K.; Redpath, S.; Ibey, A.A.M.; Chan, A.D.C.; Green, J.R.; Langlois, R.G. Developing an Instrumentation Package to Measure Noise and Vibration in Neonatal Patient Transport. In Proceedings of the 2023 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Jeju, Republic of Korea, 14–16 June 2023; pp. 1–6. [[CrossRef](#)]
46. Piche, E.; Guilbot, M.; Chorin, F.; Guerin, O.; Zory, R.; Gerus, P. Validity and repeatability of a new inertial measurement unit system for gait analysis on kinematic parameters: Comparison with an optoelectronic system. *Measurement* **2022**, *198*, 111442. [[CrossRef](#)]
47. Qiu, S.; Zhao, H.; Jiang, N.; Wang, Z.; Liu, L.; An, Y.; Zhao, H.; Miao, X.; Liu, R.; Fortino, G. Multi-sensor information fusion based on machine learning for real applications in human activity recognition: State-of-the-art and research challenges. *Inf. Fusion* **2022**, *80*, 241–265. [[CrossRef](#)]
48. Liu, H.; Schultz, T. 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)—HEALTHINF, Virtual Event, 9–11 February 2022.

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.