

Article

Gait Events Prediction Using Hybrid CNN-RNN-Based Deep Learning Models through a Single Waist-Worn Wearable Sensor

Muhammad Zeeshan Arshad ¹, Ankhzaya Jamsrandorj ², Jinwook Kim ¹ and Kyung-Ryoul Mun ^{1,3,*}¹ Center for Artificial Intelligence, Korea Institute of Science and Technology, Seoul 02792, Korea² Department of Human-Computer Interface & Robotics Engineering, University of Science & Technology, Daejeon 34113, Korea³ KHU-KIST Department of Converging Science and Technology, Kyung Hee University, Seoul 02447, Korea

* Correspondence: krmoon02@kist.re.kr

Abstract: Elderly gait is a source of rich information about their physical and mental health condition. As an alternative to the multiple sensors on the lower body parts, a single sensor on the pelvis has a positional advantage and an abundance of information acquirable. This study aimed to improve the accuracy of gait event detection in the elderly using a single sensor on the waist and deep learning models. Data were gathered from elderly subjects equipped with three IMU sensors while they walked. The input taken only from the waist sensor was used to train 16 deep-learning models including a CNN, RNN, and CNN-RNN hybrid with or without the Bidirectional and Attention mechanism. The groundtruth was extracted from foot IMU sensors. A fairly high accuracy of 99.73% and 93.89% was achieved by the CNN-BiGRU-Att model at the tolerance window of ± 6 TS (± 6 ms) and ± 1 TS (± 1 ms), respectively. Advancing from the previous studies exploring gait event detection, the model demonstrated a great improvement in terms of its prediction error having an MAE of 6.239 ms and 5.24 ms for HS and TO events, respectively, at the tolerance window of ± 1 TS. The results demonstrated that the use of CNN-RNN hybrid models with Attention and Bidirectional mechanisms is promising for accurate gait event detection using a single waist sensor. The study can contribute to reducing the burden of gait detection and increase its applicability in future wearable devices that can be used for remote health monitoring (RHM) or diagnosis based thereon.

Keywords: gait event detection; deep learning; wearable sensors; hybrid models; attention



Citation: Arshad, M.Z.; Jamsrandorj, A.; Kim, J.; Mun, K.-R. Gait Events Prediction Using Hybrid CNN-RNN-Based Deep Learning Models through a Single Waist-Worn Wearable Sensor. *Sensors* **2022**, *22*, 8226. <https://doi.org/10.3390/s22218226>

Academic Editor: Marco Iosa

Received: 22 August 2022

Accepted: 7 October 2022

Published: 27 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

The gait of the elderly is abundant with health information on not only their current status but also the potential health risk they are at [1]. On top of physical conditions, even mental conditions such as cognitive impairment and dementia can be found in their gait [2–5]. In general, several medical conditions can directly or indirectly affect gait including neurologic disorders (such as Parkinson's disease, stroke, dementia, and multiple sclerosis), musculoskeletal disorders (such as sarcopenia, frailty, osteoarthritis, and lumbar spinal stenosis), cardiovascular diseases (such as arrhythmias, congestive artery disease, and orthostatic hypotension), affective disorders/psychiatric conditions (such as depression, fear of falling, and sleep disorders), infections/metabolic diseases (such as diabetes mellitus, hepatic encephalopathy, vitamin B₁₂ deficiency, and obesity), sensory abnormalities (such as hearing impairment, peripheral neuropathy, and visual impairment) and post-hospitalization or post-surgery effects [6–10]. Nonetheless, precise measurement of the gait among the elderly allows us to predict and detect their medical crisis at an early stage and to establish an active strategy to prevent unnecessary disease progression.

The challenging work of measuring human gait evolved greatly from the traditional visual observation [11] to the current methods of using threshold, peak detection, hand-crafted features, and rule-based methods [12–14]. The introduction of smaller, lighter, and cheaper sensors such as the inertial measurement unit (IMU) sensors made it possible

to break free from the laborious and costly ways of using motion capture systems and force-plates, which were limited to a strict clinical setting [15,16]. The advent of machine learning-based methods such as Hidden Markov Models (HMM) [17–19], and Support Vector Machines (SVM) [20], Deep CNN [21], and Recurrent Neural Networks (RNN) [22] eased the burden of gait measurement further and opened a new horizon of accurate gait assessments.

When measuring gait, it can be quantified through temporal characteristics in which the precise detection of the heel-strike (HS) and toe-off (TO) for each foot matters. From HS and TO detected, a gait phase gets computed. Although the existing methods have achieved fairly good performance in detecting, the use of multiple sensors, especially the placement of the sensors on the lower body parts, interferes with natural walking and limits its application to the daily life of individuals.

As an alternative to the multiple sensors on the lower body parts, a single sensor on the pelvis can be suggested for its positional advantage and the abundance of information acquirable. The pelvis might be an ideal place for any wearable sensors, for it is a common site for wearing a belt, and the site hinders common daily activities the least. Moreover, the pelvis is a valuable source of information, since a single sensor detection of the right and left feet is possible; it is linked to three of the six determinants of gait, namely pelvis rotation, pelvic tilt, and lateral displacement of the pelvis [23]. It is also aligned to the vertical midline of the body at the center of mass (COM) and essentially links the lower limb to the upper body, which enables it to transmit force between the two and control whole-body balance. Using the signals from the pelvis, activities can be recognized, and even estimating the postures such as sitting and lying is possible. The signals from the pelvis are rich with information about daily activity patterns and carry comprehensive information on gait and motion. Yet, little attempt has been made in employing the pelvis signals, for the accuracy of using them was not comparable to that of using lower body part signals [24,25].

Thus, this study aimed to explore a way of improving the accuracy of gait event detection in the elderly using a single sensor on the pelvis and deep learning models. The elderly with or without health issues were recruited and their gait was measured. Various deep learning models learned the interrelationship of the gait event information that exists in the pelvis signal and predicted gait events. Then, the prediction was compared with the groundtruth from the sensors on the feet. Suggesting a reliable way of using a single sensor on the pelvis in gait detection, the study is expected to contribute to reducing the burden of gait detection and increase its applicability in future wearable devices that can be used for remote health monitoring (RHM) or diagnosis based thereon.

The main contributions of this study are as follows:

- We use a single IMU sensor attached to the waist to accurately detect both legs' HS and TO time;
- We evaluate and compare the performance of different DL models including classical DL models, RNN models, and CNN-RNN hybrid models;
- We investigate the IMU sensor signals to find the ones that are most relevant to gait events and achieved higher accuracy than using all six axis information;
- We evaluate the best proposed model on healthy as well as patient data.

2. Methods

2.1. Data-Collection

The subjects of 169 community-dwelling elderly aged between 60 and 80 years were recruited for the study. This study was approved by the Institutional Review Board of Kyung Hee University Medical Center (IRB No. 2017-04-001). Written informed consent was obtained from all the participants before participation in the study. The subjects were divided into healthy and patient groups depending on their health conditions. The patient group included subjects with frailty ($n = 47$), cognitive impairment ($n = 8$), fall history ($n = 11$), and a combination of them ($n = 9$). Frailty was defined as the 5-item FRAIL scale

test result of 1 or higher [26] and cognitive impairment was defined as the mini-mental state examination (MMSE) [27] scores less than 24. Subjects who answered in the questionnaire that they had a history of falling within the last year and received hospital treatment for it were included in the patient group. With these criteria, 75 subjects were included in the patient group while the rest of the 94 were in the healthy group. All subjects were capable of walking without any help from others or aid from devices at the time of data collection. Table 1 summarizes the demographic details of the subjects.

Table 1. Demographic information of the subjects.

| Characteristic | All Subjects $n = 169$ | Healthy Subjects $n = 94$ | Patients $n = 75$ |
|-----------------------------------|-------------------------------|------------------------------|---------------------------------|
| Age (years), Mean \pm SD(Range) | 74.89 \pm 5.08 (60–87) | 74.66 \pm 4.75 (64–87) | 75.17 \pm 5.48 (60–87) |
| Height (cm), Mean \pm SD(Range) | 159.67 \pm 7.23 (141.9–171) | 160.5 \pm 7.01 (141.9–171) | 155.65 \pm 7.55 (151.3–170.4) |
| Weight (kg), Mean \pm SD(Range) | 61.1 \pm 9.34 (42.3–91) | 62.01 \pm 9.69 (42.3–91) | 59.92 \pm 8.79 (42.5–80) |
| Gender | | | |
| - Male n (%) | 68 (40.24%) | 42 (44.68%) | 26 (34.67%) |
| - Female n (%) | 101 (59.77%) | 52 (55.32%) | 49 (65.34%) |
| Temporal gait pars. | | | |
| - DLS1 (s), Mean (SD) | 0.096 (0.027) | 0.093 (0.025) | 0.101 (0.028) |
| - SLS_R (s), Mean (SD) | 0.408 (0.027) | 0.405 (0.024) | 0.41 (0.031) |
| - DLS2 (s), Mean (SD) | 0.101 (0.026) | 0.095 (0.021) | 0.108 (0.029) |
| - SLS_L (s), Mean (SD) | 0.402 (0.025) | 0.401 (0.023) | 0.404 (0.028) |
| - STEP (s), Mean (SD) | 0.504 (0.045) | 0.498 (0.042) | 0.511 (0.049) |
| - STANCE (s), Mean (SD) | 0.605 (0.062) | 0.593 (0.054) | 0.62 (0.069) |
| - STRIDE (s), Mean (SD) | 1.007 (0.08) | 0.994 (0.071) | 1.023 (0.089) |

Three commercial IMU sensors were used for this study: one on the pelvis and two on the feet (Xsens MVN, Enschede, The Netherlands). An IMU sensor was attached to a belt that the subjects wore around their waist, while the other two sensors were tied around the feet, one for each foot, as depicted in Figure 1. Wearing the sensors, the subjects were asked to walk a 10-m path three times at their preferred speed and faster than their usual speed, which made each subject walk the path a total of six times. The three translational and three rotational inertial data were collected at a sampling rate of 100 Hz and passed through a 0.5–6 Hz band-pass filter to remove high-frequency noise.

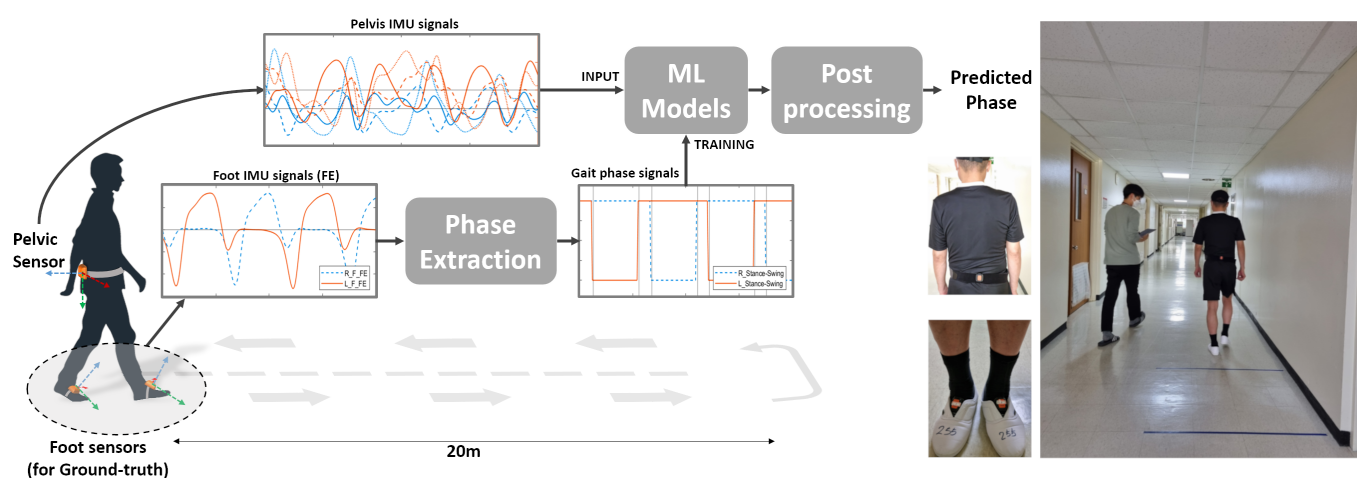


Figure 1. Overview of the proposed phase detection method and the experiment used for data acquisition. Images on the right show the actual sensors and the experiment being carried out.

Although foot switches or foot pressure insoles are generally considered the gold standard among wearable sensors, the IMU sensors can allow detailed kinematic information of the gait [28]. Furthermore, since gait event detection through angular velocity from the IMU sensor placed on the foot has been demonstrated to be as accurate as foot switches in estimating times of IC and EC for normal and abnormal gait patterns [29], this study uses this method and algorithm to obtain the ground truth gait events.

Figure 2a,b show the acceleration and angular velocity signals from the pelvis of a healthy subject. The acceleration signals included anteroposterior (AP), mediolateral (ML), vertical (V), while the angular velocity signals included tilt (TIL), obliquity (obl), and rotation (rot). To detect the actual HS and TO, the angular velocity signals for the flexion and extension of the right and left foot in the sagittal plane were used (Figure 2c). The toe-off events were detected as inverted high-amplitude peaks marked with squares, and the heel-strike events were detected as the zero-crossing before the inverted low-amplitude peaks marked with triangles in the same figure [29]. Using the four events of HS and TO of each foot, the groundtruth data was generated (Figure 2d). The period of the gait cycle with the foot on the ground (HS to TO) was called the stance phase, while that with the foot in the air (TO to HS) was called the swing phase [30]. Figure 2d shows the right foot phase signal represented as a dashed red line and the left foot signals as a solid black line. A value of 1 was assigned for the stance phase and -1 for the swing phase.

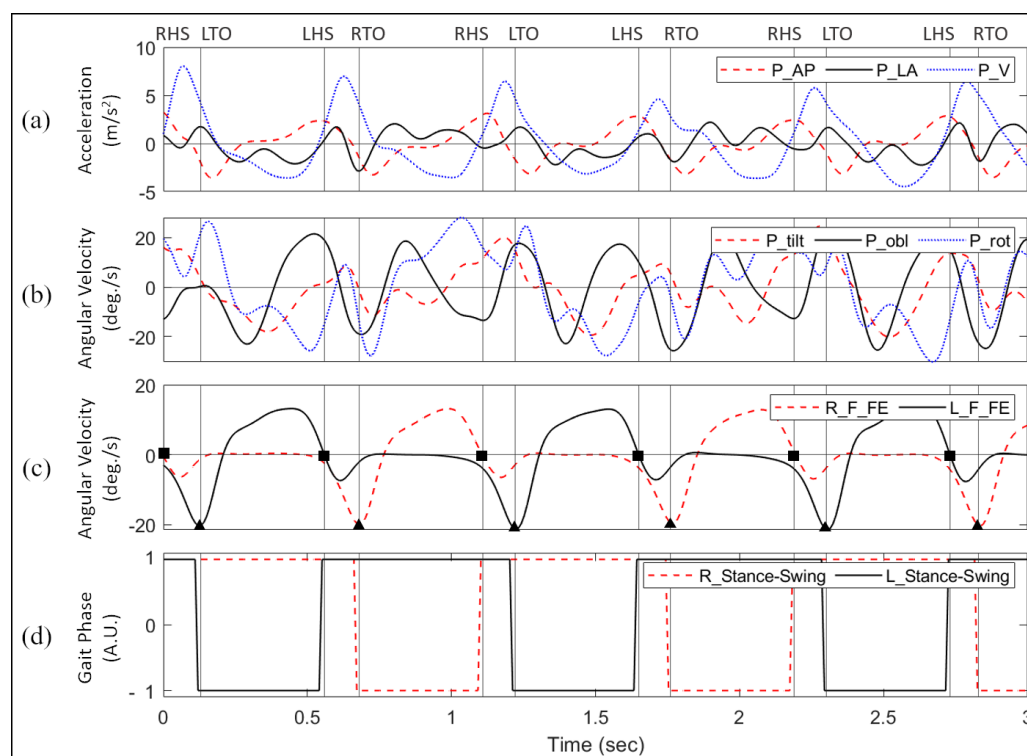


Figure 2. Input signals consist of (a) acceleration signals; (b) angular-velocity signals from the pelvis. (c) Groundtruth is extracted by identifying events from the angular-velocity signals from the feet. (d) The stance and swing phase signals were generated for the right and left foot as two continuous groundtruth signals for the regression-based models.

Figure 3 illustrates how the input–output data pairs were generated for training the one-step-ahead prediction, where the input x refers to the pelvis IMU data in the sliding window and the output y refers to the right and left phase signal values for the timestep next to the sliding window. The first pair started with the input of x_1, x_2, \dots, x_w , where w is the window length and with the output of y_{w+1} . The window was then shifted by one timestep. Hence for each pair at timestep t , the input was $x_{t-w}, x_{t-w+1}, \dots, x_{t-1}$ with

the output of y_t . For the input data of timestep n , a total of $n - w$ input-output pairs were generated.

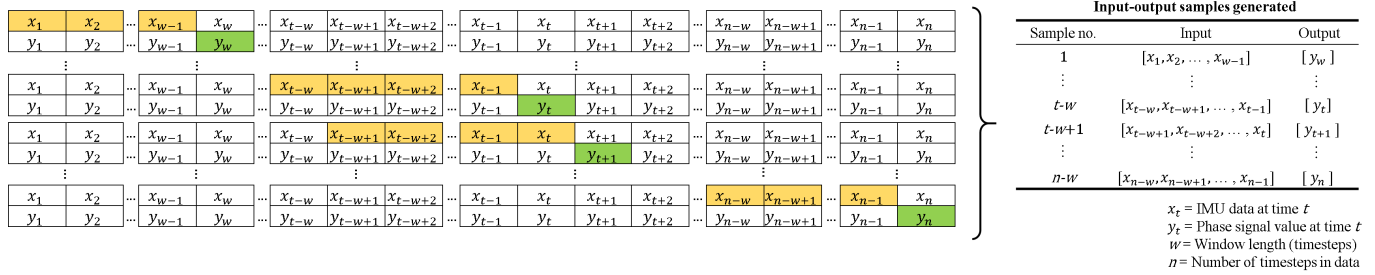


Figure 3. The training data is prepared as input–output pairs where the input consist of previous values (yellow) of the pelvis IMU signals in a moving window and the output is the next value (green) of groundtruth phase signal after the window.

2.2. Deep Learning Models

A CNN consists of a convolutional layer, an activation function, and a pooling layer, which can be defined as

$$a_{i,j} = f \left(\sum_{m=1}^M \sum_{n=1}^N w_{m,n} \cdot x_{i+m,j+n} + b \right) \quad (1)$$

where $a_{i,j}$ is the respective activation, f is the non-linear activation function, $W_{m,n}$ is the weight matrix of convolution kernel $m \times n$, $x_{i+m,j+n}$ is the upper neuron activation connected to the neuron (i, j) , and b is the bias term. The pooling layer is used to reduce network parameters and simplify the operations.

Among the RNNs, LSTMs and GRUs have become the most commonly used sequential models. LSTMs were originally proposed to overcome short-term dependency in classical RNN models [31]. Each LSTM unit is controlled by three gates, a forget gate, an input gate, and an output gate. For each time step t , the LSTM layer takes as input x^t , previous cell state c^{t-1} and previous output h^{t-1} , all real-valued vectors, and computes the new cell state, c^t , by:

$$\begin{aligned} \hat{x}_t &= [x_t; h_{t-1}] \\ i_t &= \sigma(W_i \hat{x}_t + b_i) \\ f_t &= \sigma(W_f \hat{x}_t + b_f) \\ c_t &= f_t \cdot c_{t-1} + y_t \cdot \tanh(W_c \hat{x}_t + b_c) \end{aligned}$$

where i_t and f_t are called the *input* and *forget* gates. Finally, the output h_t is generated by passing c_t through a tanh and multiplying with an output gate, o_t .

$$\begin{aligned} o_t &= \sigma(W_o \hat{x}_t + b_o) \\ h_t &= o_t \cdot \tanh(c_t) \end{aligned}$$

GRUs were introduced to reduce the computational burden in LSTM by integrating the forget gate and input gate into an update gate [32]. The mathematical expressions of GRU cell are as follows:

$$\begin{aligned} \hat{x}_t &= [x_t; h_{t-1}] \\ z_t &= \sigma(W_z \hat{x}_t + b_z) \\ r_t &= \sigma(W_r \hat{x}_t + b_r) \\ \tilde{h}_t &= \tanh(W_h r_t \cdot \hat{x}_t) \end{aligned}$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tanh(W_h r_t \cdot \tilde{h}_t)$$

Self-attention mechanism [33] enables the information from all the hidden states corresponding to the whole sequence input in RNN by learning the weights for each hidden state through the following equations where z_t is the update gate and r_t is the reset gate.

$$a_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})}$$

where

$$e_{ij} = a(s_{i-1}, h_j)$$

$$c_i = \sum_{j=1}^{T_x} a_{ij} h_j$$

A total of 16 deep learning models were trained and tested. The models included classical, convolutional neural network (CNN), Recurrent Neural Networks (RNN) models with or without the Bidirectional or Attention mechanisms, along with some hybrid ones that combined CNN with RNN models with or without the Bidirectional or Attention mechanisms. The classical models included Convolutional Neural Networks (CNN) and Multi-Layer Perceptron (MLP) model, while the RNN models included Long Short-term Memory (LSTM), Gated Recurrent Unit (GRU), Bidirectional LSTM (BiLSTM), Bidirectional GRU (BiGRU), Stacked LSTM, Stacked GRU, Stacked LSTM with the Attention mechanism (stacked-LSTM-Att), and stacked-GRU with the Attention mechanism (stacked-GRU-Att). The hybrid models included CNN combined with LSTM (CNN-LSTM), CNN combined with GRU (CNN-GRU), CNN combined with Bidirectional LSTM (CNN-BiLSTM), CNN combined with Bidirectional GRU (CNN-BiGRU), CNN combined with Bidirectional LSTM and the Attention mechanism (CNN-BiLSTM-Att), and CNN combined with Bidirectional GRU and the Attention mechanism (CNN-BiGRU-Att).

For all models, the raw data for training consisted of six IMU signals from the pelvis as inputs and the stance and swing phase signals from both feet as outputs. The models used the generated input and output data samples and made a one-step-ahead prediction. The hyper-parameters for all models were optimized to obtain the best accuracy for each. A Dense layer followed by a Linear activation function (AF) was used for the final output.

The architecture adopted for the deep learning models is presented in Figure 4. Figure 4a shows the architecture for MLP. The model included a Dense layer of 100 neurons for each of the six input signals and a concatenate layer for combining them before connecting to a Dense layer. As for the CNN model, the input was first reshaped to the dimensions [samples, timesteps, features] for making it compatible with the one-dimensional convolutional layer, Conv1D (Figure 4b). The Conv1D layer was followed by a one-dimensional pooling layer, MaxPool1D. After flattening, the outputs went into a Dense layer of 50 neurons, and the last Dense layer connected this sequence to the final output after a Linear AF. As for the RNN models with a single or stacked LSTM or GRU with or without attention, the single-layer vanilla LSTM and GRU networks had 100 hidden units and a Dense layer and a Linear AF followed (Figure 4c). The stacked LSTM and GRU included two layers stacked together with 100 hidden units for each. The first layer fed its hidden state to the second one, which was used for the output prediction.

For the hybrid models combining CNN and RNN models, the input with dimensions [samples, window-length, features] was reshaped by splitting the window-length dimension into two segments. With a window length of 80 timesteps, the input dimensions [samples, 80, 6] were transformed to [samples, 2, 40, 6]. The reshaped input was fed to Conv1D. MaxPooling1D followed the convolutional layer, which was then flattened at the end (Figure 4d). The time-distributed layer that wrapped the convolutional blocks enabled applying the same instance of blocks to all the temporal slices of the input [34]. The output

from this CNN went into the single layer of LSTM or GRU with 600 hidden units. The Dense layer followed with Linear AF as in other models.

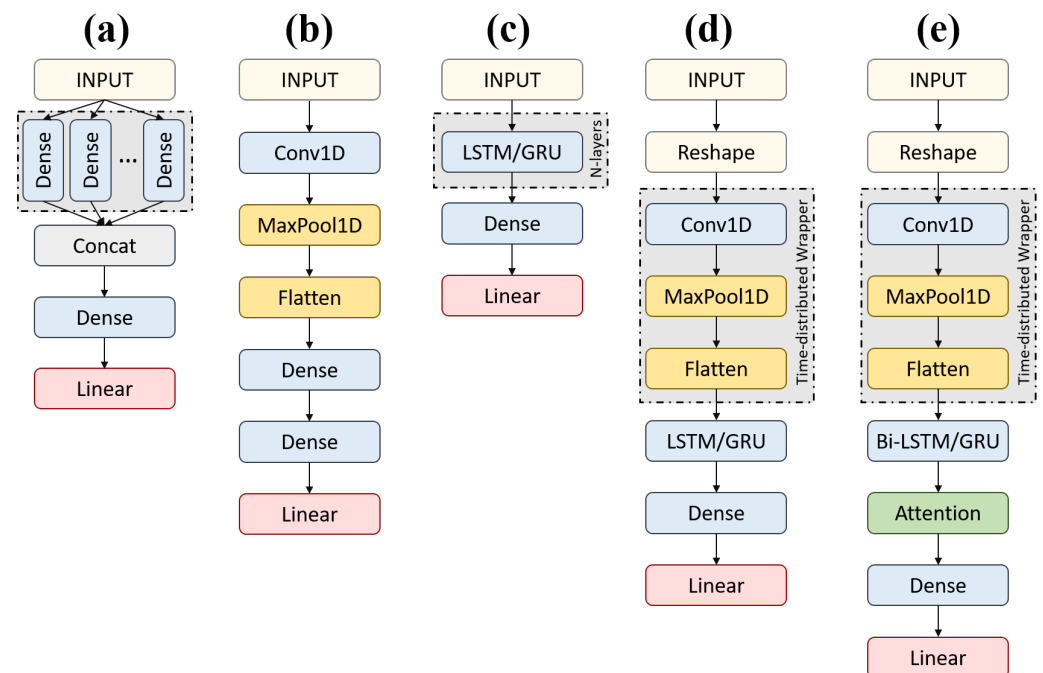


Figure 4. Brief architectural details for models, (a) MLP; (b) CNN; (c) LSTM and GRU-based models; (d) CNN-RNN hybrid models; (e) CNN-RNN hybrid models with bidirectional and Attention mechanisms.

As for the models with the Bidirectional and Attention mechanisms, the Bidirectional mechanism provided the forward and backward sequence of the input to the two different RNN layers, allowing the network to make use of the past and future context for each point in the input to make predictions [35]. The Self-attention mechanism enabled the model to give more weightage to a specific part of the input sequence [36]. In this context, the Attention allowed the models to find temporal dependencies in certain periods in the sliding window instead of relying on all to make more accurate predictions. Figure 4e shows the CNN-RNN hybrid with the Bidirectional mechanism incorporated into the RNN, and the Attention mechanism that is connected to the output of RNN networks. For the final output, the Attention layer was followed by the Dense layer and Linear AF as the others.

All models were trained using the Adam optimizer and mean squared error as the loss metric. An early stopping criterion was used to retrieve the best model by minimizing the validation loss with the patience of 10 epochs.

The parameters and data dimensions for each layer of the CNN-RNN hybrid model with Attention have been given in Figure 5.

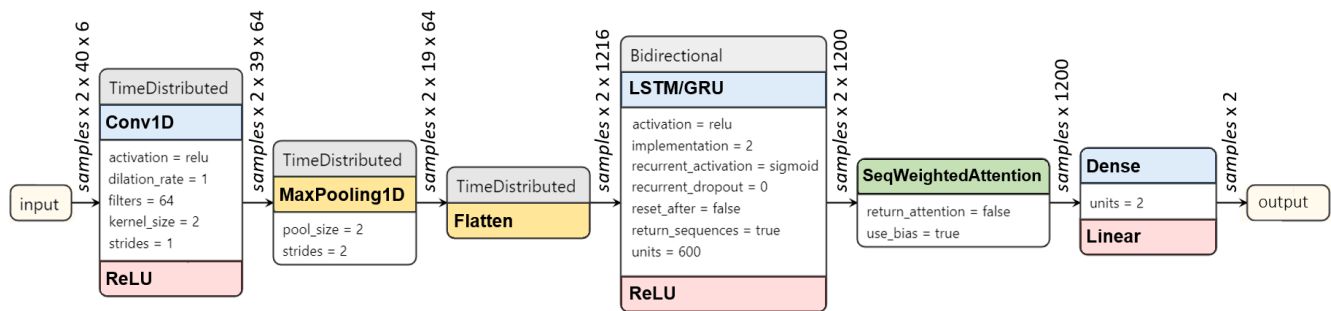


Figure 5. The parameters and data dimensions for each layer of CNN-GRU/LSTM Attention model.

2.3. Output Post-Processing

Post-processing of the output signals was performed to remove the noisy spikes in the raw outputs and improve the prediction accuracy (Figure 6). As the first step, all transitions across the zero-line were extracted from the output signals. Then, these transitions were filtered to remove the disturbances, which could be mistaken as phase transitions. Since the output signal is essentially a pulse train, the valid phase transitions were identified by distinguishing between noisy spikes and real pulses. For a pulse to be identified as valid, three conditions were used: (1) the maximum value of the pulse must be higher than 0.5, (2) the mean value of the pulse must be higher than 0.6, and (3) the pulse-width must be higher than three timesteps. The processed output matched the groundtruth better (Figure 6b).

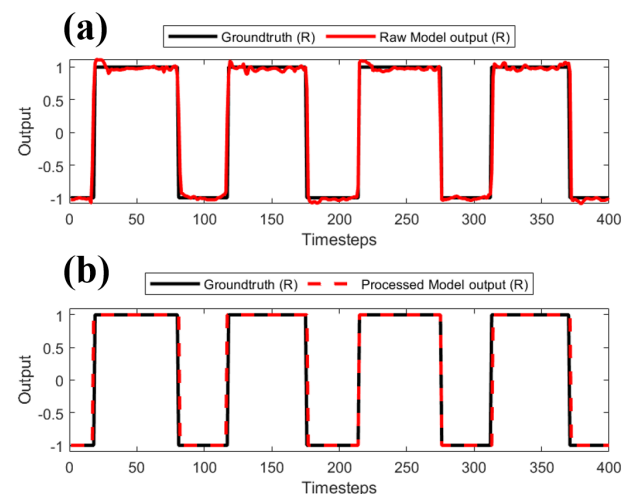


Figure 6. Groundtruth compared to (a) pre-processed output signal and (b) post-processed output signal.

2.4. Accuracy Measurement

The accuracy was measured by different tolerance windows to verify the precision of the detection. An event was defined as successfully detected when the output transition occurred within W , the tolerance window of size (Figure 7). The performance of each model was measured with the tolerance windows ± 1 TS, ± 2 TS, \dots , ± 6 TS where 1 TS refers to one timestep of 1 ms. The overall accuracy was defined as the percentage of correctly detected events to the total number of events.

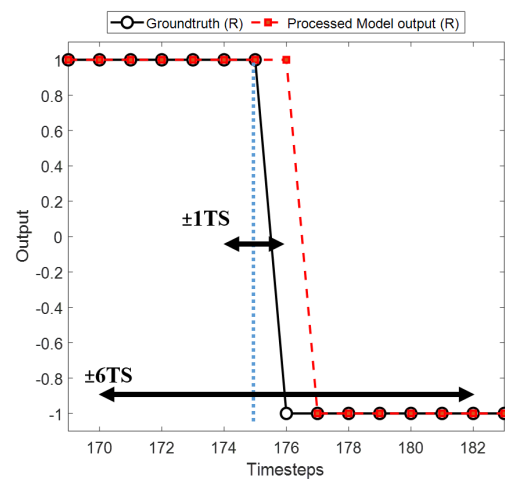


Figure 7. The accuracy is measured with different precision levels termed as tolerance windows ± 1 TS, ± 2 TS, ..., ± 6 TS, where 1 TS is a single timestep of 1 ms measured from the groundtruth event (blue dotted line).

To investigate further, the accuracy of the models including the data from the patient group was investigated. The models were trained and tested with the healthy group, trained with the healthy group but tested with the patient group, and trained and tested with both groups.

3. Results

Table 2 summarizes the average accuracy of the models trained and tested with the healthy group by different sizes of tolerance windows. CNN-BiGRU-Att model using the tolerance window of ± 6 TS achieved the highest accuracy of 99.73. All hybrid models with the Bidirectional mechanism with or without the Attention mechanism demonstrated an accuracy comparable to the accuracy of the best-performing CNN-BiGRU-Att.

Table 2. Accuracy results for all models.

| Model | ± 1 TS | ± 2 TS | ± 3 TS | ± 4 TS | ± 5 TS | ± 6 TS |
|------------------|------------|------------|------------|------------|------------|------------|
| CNN-BiGRU-Att | 93.89 | 98.29 | 99.02 | 99.46 | 99.64 | 99.73 |
| CNN-BiLSTM | 93.68 | 98.28 | 99.14 | 99.48 | 99.68 | 99.76 |
| CNN-BiLSTM-Att | 93.52 | 98.21 | 99.02 | 99.46 | 99.64 | 99.70 |
| CNN-BiGRU | 93.27 | 98.19 | 99.10 | 99.41 | 99.56 | 99.73 |
| stacked-LSTM-Att | 92.71 | 97.60 | 98.91 | 99.31 | 99.49 | 99.59 |
| CNN-GRU | 92.33 | 97.82 | 98.93 | 99.40 | 99.57 | 99.65 |
| CNN-LSTM | 91.77 | 97.89 | 98.94 | 99.44 | 99.61 | 99.76 |
| stacked-GRU-Att | 91.67 | 97.54 | 98.68 | 99.30 | 99.59 | 99.65 |
| BiGRU | 89.99 | 97.08 | 98.76 | 99.35 | 99.55 | 99.64 |
| stacked-GRU | 88.86 | 96.51 | 98.53 | 99.24 | 99.51 | 99.60 |
| stacked-LSTM | 88.04 | 96.64 | 98.53 | 99.18 | 99.47 | 99.59 |
| BiLSTM | 86.54 | 95.98 | 98.17 | 99.02 | 99.35 | 99.51 |
| GRU | 78.98 | 92.70 | 96.72 | 98.32 | 99.13 | 99.47 |
| LSTM | 76.83 | 91.61 | 96.21 | 98.18 | 98.89 | 99.24 |
| MLP | 69.55 | 86.57 | 92.88 | 95.73 | 97.05 | 97.53 |
| CNN | 68.58 | 87.23 | 94.25 | 96.76 | 98.01 | 98.56 |

The accuracy increased as the sizes of the tolerance window increased. With a wider tolerance window, the detection rate increased while the precision decreased. The best accuracy for each model was achieved with the tolerance window of ± 6 TS. Since all models gave little deviation from the tolerance window greater than ± 3 TS, the comparisons

between the models were made with the accuracy at the tolerance window of ± 1 TS, which means the most precise detection.

Figure 8 shows the accuracy plots for all events (Figure 8a), HS events (Figure 8b), and TO events (Figure 8c). The nine key models were chosen as the best-performing ones for each type. The events for the right and left foot were averaged for all three plots. In all three events, the accuracy increased as the tolerance window increased. In most models, higher accuracies were achieved with TO events compared to HS events.

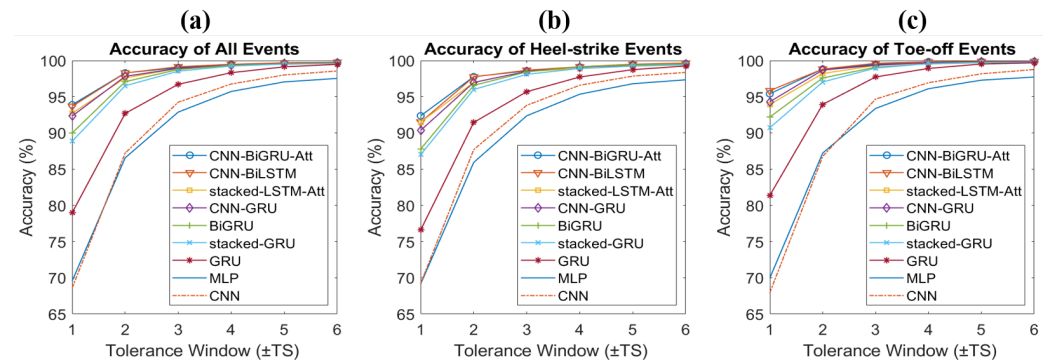


Figure 8. Accuracy plots for key models for (a) all events, (b) HS events, and (c) TO events.

To observe the differences in the performance of the events from the right and left foot, the absolute errors in timesteps between the predicted and the groundtruth events were computed. Figure 9 shows the event prediction error of CNN-BiGRU-Att. The mean absolute error (MAE) for all events was 5.77 ms, whereby those for HS and TO events were 6.239 ms and 5.24 ms, respectively. The TO events showed fewer errors than the HS events. No significant difference was found between using the right foot events and the left foot events.

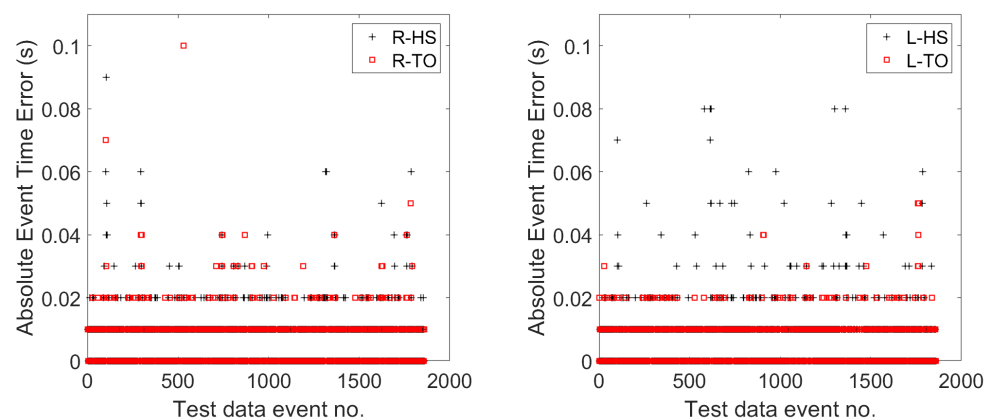


Figure 9. Event prediction errors for right and left, heel-strike, and toe-off events.

An ablation study with subsets of the six input signals was performed to investigate the prediction accuracy by the number of input signals. Table 3 shows the list of the best-performing input combinations for the CNN-BiGRU-Att model. Using either ML or ROT and using both of them increased the accuracy. When two input signals were used, the combination of AP and ML and that of AP and ROT gave an accuracy of about 88% at the tolerance window of ± 1 TS. The highest accuracy of over 94% was achieved by using four input signals AP, ML, V, and ROT. When all six input signals were used, an accuracy of 93.89% was achieved.

Table 3. Results for ablation study using subsets of input signals.

| No. of Input Signals | Input | ± 1 TS | ± 2 TS | ± 3 TS | ± 4 TS | ± 5 TS | ± 6 TS |
|----------------------|----------------------------|------------|------------|------------|------------|------------|------------|
| 1 | [AP] | 20.57 | 27.75 | 32.12 | 34.65 | 36.26 | 37.52 |
| | [ML] | 60.12 | 79.12 | 88.20 | 92.57 | 94.80 | 96.14 |
| | [V] | 15.37 | 21.60 | 25.12 | 27.47 | 28.83 | 29.74 |
| | [TIL] | 24.52 | 33.52 | 38.65 | 41.89 | 44.03 | 45.42 |
| | [OBL] | 15.52 | 24.05 | 31.13 | 36.47 | 40.53 | 43.18 |
| | [ROT] | 58.00 | 75.96 | 84.42 | 88.45 | 90.72 | 92.09 |
| 2 | [AP, ML] | 88.16 | 97.12 | 98.66 | 99.14 | 99.45 | 99.65 |
| | [AP, ROT] | 87.23 | 95.45 | 97.61 | 98.66 | 99.27 | 99.45 |
| 3 | [AP, ML, TIL] | 90.15 | 97.28 | 98.68 | 99.21 | 99.60 | 99.70 |
| | [AP, ML, ROT] | 90.28 | 97.34 | 98.86 | 99.18 | 99.50 | 99.68 |
| 4 | [AP, ML, V, TIL] | 93.40 | 98.25 | 98.95 | 99.30 | 99.60 | 99.68 |
| | [AP, ML, V, ROT] | 94.11 | 98.48 | 99.13 | 99.46 | 99.68 | 99.79 |
| 5 | [AP, ML, V, TIL, OBL] | 87.97 | 96.57 | 98.31 | 99.14 | 99.49 | 99.62 |
| | [AP, ML, V, TIL, ROT] | 93.89 | 98.04 | 98.97 | 99.53 | 99.73 | 99.84 |
| | [AP, ML, V, OBL, ROT] | 93.99 | 98.28 | 98.94 | 99.38 | 99.68 | 99.81 |
| 6 | [AP, ML, V, TIL, OBL, ROT] | 93.89 | 98.29 | 99.02 | 99.46 | 99.64 | 99.73 |

Table 4 shows the accuracy of the CNN-BiGRU-Att model trained and tested with the same or different subject groups. When all six input signals were used, the models trained and tested with the healthy group exhibited an accuracy of 93.89% at the tolerance window of ± 1 TS. When trained with the healthy group but tested with the patient group, an accuracy of 63.10% was achieved. When trained and tested with both groups, the models achieved an accuracy of 93.63%. When four input signals of AP, ML, V, and ROT were used, the models trained and tested with the healthy group achieved the highest accuracy of 94.11% at the tolerance window of ± 1 TS, which was higher than that of using all six signals. When the model was trained with both groups, using all six inputs achieved higher accuracy at the tolerance window of ± 1 TS and ± 2 TS. To observe the accuracies of the HS and TO events for these results, accuracy plots are given in Figure 10. As observed earlier, the TO events were identified with higher accuracy than the HS events.

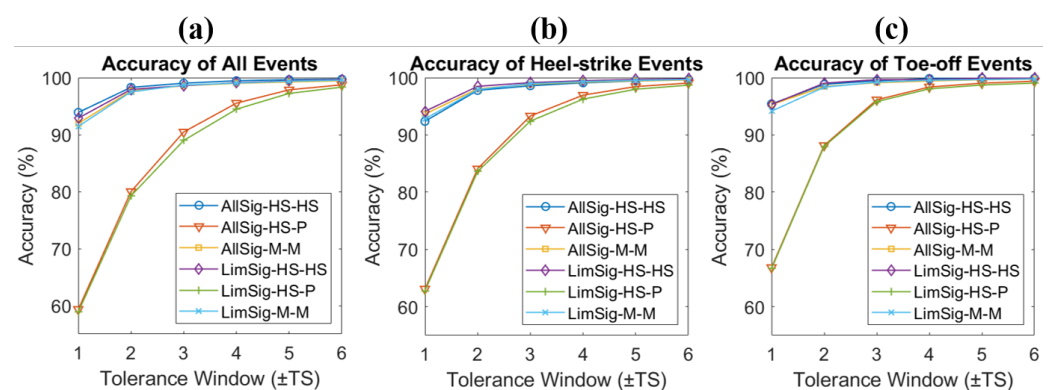


Figure 10. Accuracy plots for (a) all events, (b) heel-strike events, and (c) toe-off events when working with patient data. The plots include results from models trained with all 6 pelvis signals (AllSig) and with limited 4 signals (LimSig). The training and testing include healthy-trained, patient-tested (HS-P), healthy-trained, healthy-tested (HS-HS), and mixed-trained and mixed-tested (M-M).

Table 4. Accuracy results when patient data is used.

| Input Signals | Training | Testing | ± 1 TS | ± 2 TS | ± 3 TS | ± 4 TS | ± 5 TS | ± 6 TS |
|----------------------------|----------|---------|------------|------------|------------|------------|------------|------------|
| [AP, ML, V, TIL, OBL, ROT] | HS | HS | 93.89 | 98.29 | 99.02 | 99.46 | 99.64 | 99.73 |
| | HS | P | 63.10 | 84.09 | 93.30 | 96.94 | 98.44 | 99.05 |
| | Mixed | Mixed | 93.63 | 98.04 | 98.85 | 99.21 | 99.44 | 99.59 |
| [AP, ML, V, ROT] | HS | HS | 94.11 | 98.48 | 99.13 | 99.46 | 99.68 | 99.79 |
| | HS | P | 62.78 | 83.63 | 92.40 | 96.25 | 97.97 | 98.69 |
| | Mixed | Mixed | 92.80 | 97.91 | 98.95 | 99.30 | 99.52 | 99.66 |

4. Discussion

This study aimed to explore a way of improving the accuracy of gait event detection among the elderly using a single sensor on the pelvis and deep learning technology. A total of 16 models were trained and tested to predict the gait events of the elderly, and their prediction was compared with the groundtruth acquired from the feet. The CNN-BiGRU-Att model achieved the highest accuracy of 99.73% at the tolerance window of ± 6 TS, which was an accuracy comparable to that of using multiple sensors on the lower body parts. An ablation study demonstrating using four signal sets of AP, ML, V, and ROT achieved an accuracy of over 94% at the tolerance window of ± 1 TS. The study pioneered the utilization of deep learning technology in predicting gait events using the data from the pelvis and suggested a reliable way of using a single sensor on the pelvis in gait event detection among the elderly. The findings are expected to contribute to reducing the burden of gait measurement and increase the potential of various future technologies being incorporated with the suggested method.

The use of CNN together with RNN models improved the prediction accuracy since CNN first extracted effective features and then the following RNN models could process these features sequentially. The added Bidirectional mechanism took into account both forward and backward temporal perspectives, which improved the accuracy further. Adding the Attention mechanisms to the CNN-BiGRU increased the accuracy even more since the Attention mechanisms could focus on the timesteps that were more relevant for the output prediction. Being unable to utilize the temporal nature of the information in the sliding window, MLP and CNN were not able to predict the phase transitions in the output signal accurately until ± 4 TS, although their accuracy improved as the tolerance window sizes grew bigger.

Many previous studies have looked at the use of multiple sensors to analyze gait events. Lin et al. [22] used LSTM-based regression model using five IMUs, two on the thighs, two on the shins, and one on the left shoe. They achieved the mean error (ME) of 2 ms for HS; however, large errors for TO were reported with the ME of -18 ms. In [21], Hannink et al. used a deep CNN-based network with input from two inertial sensors placed on the feet and predicted different gait parameters including the HS and TO events as output, for which they reported errors of ± 70 ms and ± 120 ms, respectively. Sarshar et al. [37], who also used RNN to train two IMU sensors attached to the shanks, reported an accuracy of 0.9977 for both HS and TO events; however, they did not compute the error in prediction delays. Utilizing rule-based algorithms, Fadillioglu et al. [38] presented an automated gait event detection method using a gyroscope attached to the right shank and reported an MAE of 7 ms and 19 ms for the HS and TO events, respectively. More recently, Yu et al. [39] also used a single sensor on the right foot with an LSTM-HMM hybrid model for gait event detection, reporting the accuracy only without mentioning the delays. They reported accuracy of 0.9679 and 0.9846 for the HS and TO events. But since both of these works used only a single sensor on the right side, they were unable to get events for the left side. In comparison, only a single sensor on a comparatively less accurate position of the waist was used in the proposed method but a much superior overall performance was shown. The studies using multiple sensors not only hinder the natural gait but are also far from practical usage of these methods in daily living. If a single sensor is attached to one side of

the limb, the extraction of gait events from the other side is not possible, hence for complete gait event information two sensors are required for any location on the lower limb except for the waist. Even though the waist is a less accurate position as compared to the lower limb, the proposed method in the current study has still managed to achieve more accurate gait event detection using a single sensor.

Compared with previous attempts that detected gait events from a single sensor on the waist, the proposed CNN-BiGRU-Att model achieved far advanced accuracy. According to a recent survey of gait event detection methods using an IMU sensor mounted on the waist, Gonzalez et al. [40] used a rule-based method to achieve the lowest MAE of 15 ms and 9 ms for the HS and TO events, respectively. McCamley et al. [41] proposed a Gaussian CWT-based gait event estimation method using a single inertial sensor on the waist, reporting an MAE of 19 ms and 32 ms for HS and TO, respectively. Soaz et al. [42] also used a rule-based method with a single waist sensor to assess the gait of the elderly for their experiment and reported an error of 20 ms for HS. Apart from relatively higher errors than the proposed method, these studies also did not have enough subjects to show the generalizability of their method.

The motion capture system, though considered the gold standard for gait measurement, requires the installation of expensive equipment. Furthermore, it must be used indoors in a limited space. On the other hand, gait event detection techniques based on electromyography signals cannot give better performance than IMU-based sensors due to sensor location sensitivity, low intra-operator repeatability, low inter-operator reproducibility, and higher inter-subject variability [43]. For example, in [44], Morbidoni et al. used multilayer perceptron (MLP) architectures for ground walking EMG signals and reported an MAE of 21.6 ms and 38.1 ms for HS and TO, respectively. Similarly, Nazmi et al. [45] also used the artificial neural network for the walking task and reported an MAE of 16 ± 18 ms and 21 ± 18 ms for HS and TO, respectively.

The study found that the TO events show fewer errors than the HS events. The occurrence of extrema in the pelvis signals around these events can serve as a possible explanation. As presented in Figure 2, the TO events were more aligned with the peaks in pelvic signals such as V, TIL, and ROT, which were followed by the TO events. These characteristics may have contributed to the better performance of the TO events compared with the HS events.

As for the single input signals used, ML and ROT demonstrated an outstanding performance since they are rich in information on the right and left. The signals AP and V are the clearest but void of right and left information, so when the prediction was made based on only either of them, they exhibited the lowest performance. Combining AP with either ML or ROT might have incorporated the movement information on both the forward and backward with that on the right and left resulting in fairly improved accuracy. The overall accuracy of using AP, ML, V, and ROT was higher than the accuracy of using all six signals, probably because of the bigger variation found in TILT and OBL among the elderly [46]. AP, ML, and V being acceleration signals may have contributed as well, since the accelerometer output is generally less prone to sensor location errors compared to a gyroscope.

When the model was trained and tested with the same and different groups, the models trained with the healthy group but tested with the patient group exhibited an accuracy of 63.10 at ± 1 TS with all the six input signals used. Trained with the healthy group, the model was not familiar with the variations and abnormalities that the patient group had but its accuracy improved, as the tolerance window sizes grew bigger. The inclusion of TILT and OBL in the healthy group did not significantly improve the accuracy of the model, whereas the inclusion of them in the patient group improved the accuracy, probably because of the variation in all signals being greater in the patient group. It would be advisable to consider using all six signals for patients whose signals demonstrate a large variation, even for the signals considered relatively clear.

To investigate how the Attention weight varied when there were all six input signals and when they were limited to four; the value of the average attention weight for each timestep was plotted using the stacked-LSTM-Att model (Figure 11). When the model used all six input signals, it gave much greater attention to the double-limb-support (DLS) phase between the HS and TO events than to the single-limb-support (SLS) phase. When four signals were used, slightly more attention was paid to the SLS phase with its peak attention in the DLS phase being lower than that of using all six signals.

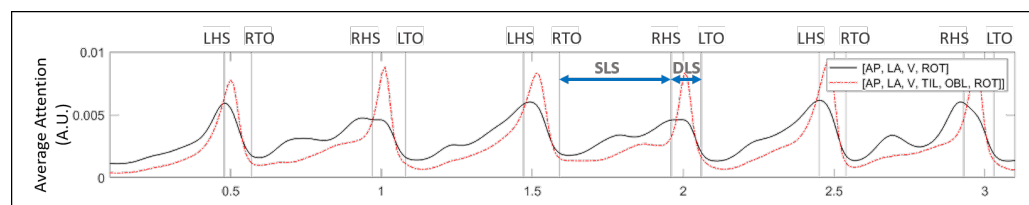


Figure 11. Average Attention plotted for the stacked-LSTM model using all 6 input pelvis signals and limited 4 input signals. The models put more attention to DLS phase between right HS and left TO as compared to the SLS between right TO and right HS.

One of the limitations of this study is that the proposed method could not be evaluated on other physical impairments such as osteoarthritis and skeletal deformities, and neurological diseases such as hemiplegia, Parkinson's disease, Huntington's disease, and Alzheimer's disease. Furthermore, the gait data has not been acquired through the long-term or continuous monitoring of subjects in their natural environment and everyday lives. Therefore, our future work will include gait event detection for real-world walking in non-conventional environments and under unconstrained and uncontrolled conditions. Furthermore, we will focus more on the real time implementations of these methods to support gait patients through exoskeleton devices.

5. Conclusions

The study proposed deep learning-based gait detection as a novel and reliable way of using a single sensor on the pelvis in detecting the gait of the elderly. A total of 16 models including the CNN, RNN, and CNN-RNN hybrid with or without the Bidirectional and Attention mechanism were trained and tested and a fairly high accuracy of 99.73% and 93.89% was achieved by the CNN-BiGRU-Att model at the tolerance window of ± 6 TS and ± 1 TS, respectively. Advancing from the previous studies exploring gait event detection, the model showed a great improvement in terms of its prediction error having an MAE of 6.239 ms and 5.24 ms for HS and TO events, respectively. For healthy subjects, using all three acceleration signals with ROT as input signals exhibited better performance compared to using all six signals; meanwhile, the performance of using all six signals was better for the patients. Suggesting that reliable gait detection is possible from a single sensor on the pelvis, the study is expected to contribute to lowering the burden of gait detection and expand the applicability of gait detection in future wearable devices.

Author Contributions: Conceptualization, M.Z.A. and K.-R.M.; methodology, M.Z.A. and K.-R.M.; software, M.Z.A.; validation, K.-R.M.; formal analysis, M.Z.A. and A.J.; investigation, M.Z.A., A.J. and K.-R.M.; resources, K.-R.M.; data curation, M.Z.A. and A.J.; writing—original draft preparation, M.Z.A.; writing—review and editing, M.Z.A. and K.-R.M.; visualization, M.Z.A.; supervision, J.K. and K.-R.M.; project administration, J.K. and K.-R.M.; funding acquisition, J.K. and K.-R.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (Project No. 1711139131) and by Athletes' training/matches data management and AI-based performance enhancement solution technology Development Project (No.1375027374).

Institutional Review Board Statement: This study was approved by the Institutional Review Board of Kyung Hee University Medical Center (IRB No. 2017-04-001).

Informed Consent Statement: Written informed consent was obtained from all the participants before participation in the study.

Data Availability Statement: The data used for this study cannot be shared publicly, so supporting data is not available.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|------|-------------------------------|
| IMU | Inertial measurement unit |
| HMM | Hidden Markov Models |
| SVM | Support Vector Machines |
| CNN | Convolutional Neural Network |
| RNN | Recurrent Neural Network |
| HS | Heel-strike |
| TO | Toe-off |
| AP | Aanteroposterior |
| ML | Mediolateral |
| V | Vertical |
| TIL | Tilt |
| OBL | Obliquity |
| ROT | Rotation |
| COM | Center of mass |
| RHM | Remote health monitoring |
| MMSE | Mini-mental state examination |
| MLP | Multi-Layer Perceptron |
| LSTM | Long Short-term Memory |
| GRU | Gated Recurrent Unit |
| Bi | Bidirectional |
| Att | Attention |
| AF | Linear activation function |
| MAE | Mean absolute error |
| ME | Mean error |
| DLS | Double limb support |
| SLS | Single limb support |

References

1. Studenski, S.; Perera, S.; Patel, K.; Rosano, C.; Faulkner, K.; Inzitari, M.; Brach, J.; Chandler, J.; Cawthon, P.; Connor, E.B.; et al. Gait speed and survival in older adults. *JAMA* **2011**, *305*, 50–58. [\[CrossRef\]](#)
2. Arshad, M.Z.; Jung, D.; Park, M.; Shin, H.; Kim, J.; Mun, K.R. Gait-based Frailty Assessment using Image Representation of IMU Signals and Deep CNN. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Online, 1–5 November 2021; pp. 1874–1879.
3. Jung, D.; Kim, J.; Kim, M.; Won, C.W.; Mun, K.R. Classifying the Risk of Cognitive Impairment Using Sequential Gait Characteristics and Long Short-Term Memory Networks. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 4029–4040. [\[CrossRef\]](#)
4. Verghese, J.; Robbins, M.; Holtzer, R.; Zimmerman, M.; Wang, C.; Xue, X.; Lipton, R.B. Gait dysfunction in mild cognitive impairment syndromes. *J. Am. Geriatr. Soc.* **2008**, *56*, 1244–1251. [\[CrossRef\]](#)
5. Mielke, M.M.; Roberts, R.O.; Savica, R.; Cha, R.; Drubach, D.I.; Christianson, T.; Pankratz, V.S.; Geda, Y.E.; Machulda, M.M.; Ivnik, R.J.; et al. Assessing the temporal relationship between cognition and gait: Slow gait predicts cognitive decline in the Mayo Clinic Study of Aging. *J. Gerontol. Ser. Biomed. Sci. Med. Sci.* **2013**, *68*, 929–937. [\[CrossRef\]](#)
6. Salzman, B. Gait and balance disorders in older adults. *Am. Fam. Phys.* **2010**, *82*, 61–68.
7. Alexander, N.B.; Goldberg, A. Gait disorders: Search for multiple causes. *Clevel. Clin. J. Med.* **2005**, *72*, 586–589. [\[CrossRef\]](#)
8. Moylan, K.C.; Binder, E.F. Falls in older adults: Risk assessment, management and prevention. *Am. J. Med.* **2007**, *120*, 493. [\[CrossRef\]](#)

9. Alexander, N.B. Gait disorders in older adults. *J. Am. Geriatr. Soc.* **1996**, *44*, 434–451. [\[CrossRef\]](#)
10. Sudarsky, L. Gait disorders: Prevalence, morbidity, and etiology. *Adv. Neurol.* **2001**, *87*, 111–117.
11. Krebs, D.E.; Edelman, J.E.; Fishman, S. Reliability of observational kinematic gait analysis. *Phys. Ther.* **1985**, *65*, 1027–1033. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Kim, M.; Lee, D. Development of an IMU-based foot-ground contact detection (FGCD) algorithm. *Ergonomics* **2017**, *60*, 384–403. [\[CrossRef\]](#)
13. Oudre, L.; Barrois-Müller, R.; Moreau, T.; Truong, C.; Vienne-Jumeau, A.; Ricard, D.; Vayatis, N.; Vidal, P.P. Template-based step detection with inertial measurement units. *Sensors* **2018**, *18*, 4033. [\[CrossRef\]](#)
14. Lee, H.K.; Hwang, S.J.; Cho, S.P.; Lee, D.R.; You, S.H.; Lee, K.J.; Kim, Y.H.; Choi, H.S. Novel algorithm for the hemiplegic gait evaluation using a single 3-axis accelerometer. In Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 3–6 September 2009; pp. 3964–3966.
15. Barker, S.; Craik, R.; Freedman, W.; Herrmann, N.; Hillstrom, H. Accuracy, reliability, and validity of a spatiotemporal gait analysis system. *Med Eng. Phys.* **2006**, *28*, 460–467. [\[CrossRef\]](#)
16. Winter, D.A. *Biomechanics and Motor Control of Human Movement*; John Wiley & Sons: Hoboken, NJ, USA, 2009.
17. Mannini, A.; Sabatini, A.M. Gait phase detection and discrimination between walking–jogging activities using hidden Markov models applied to foot motion data from a gyroscope. *Gait Posture* **2012**, *36*, 657–661. [\[CrossRef\]](#)
18. Taborri, J.; Rossi, S.; Palermo, E.; Patanè, F.; Cappa, P. A novel HMM distributed classifier for the detection of gait phases by means of a wearable inertial sensor network. *Sensors* **2014**, *14*, 16212–16234. [\[CrossRef\]](#)
19. Bae, J.; Tomizuka, M. Gait phase analysis based on a Hidden Markov Model. *Mechatronics* **2011**, *21*, 961–970. [\[CrossRef\]](#)
20. Mannini, A.; Trojaniello, D.; Cereatti, A.; Sabatini, A.M. A machine learning framework for gait classification using inertial sensors: Application to elderly, post-stroke and huntington’s disease patients. *Sensors* **2016**, *16*, 134. [\[CrossRef\]](#)
21. Hannink, J.; Kautz, T.; Pasluosta, C.F.; Gaßmann, K.G.; Klucken, J.; Eskofier, B.M. Sensor-based gait parameter extraction with deep convolutional neural networks. *IEEE J. Biomed. Health Inform.* **2016**, *21*, 85–93. [\[CrossRef\]](#)
22. Lin, P.H.; Shih, C.L.; Wong, D.P.; Chou, P.H. Gait Parameters Analysis Based on Leg-and-shoe-mounted IMU and Deep Learning. In Proceedings of the 2021 International Symposium on VLSI Design, Automation and Test (VLSI-DAT), Hsinchu, Taiwan, 19–21 April 2021; pp. 1–4.
23. Inman, V.T.; Eberhart, H.D. The major determinants in normal and pathological gait. *JBS* **1953**, *35*, 543–558.
24. Zijlstra, W.; Hof, A.L. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait Posture* **2003**, *18*, 1–10. [\[CrossRef\]](#)
25. De Ridder, R.; Lebleu, J.; Willems, T.; De Blaiser, C.; Detrembleur, C.; Roosen, P. Concurrent validity of a commercial wireless trunk triaxial accelerometer system for gait analysis. *J. Sport Rehabil.* **2019**, *28*. [\[CrossRef\]](#)
26. Morley, J.E.; Malmstrom, T.; Miller, D. A simple frailty questionnaire (FRAIL) predicts outcomes in middle aged African Americans. *J. Nutr. Health Aging* **2012**, *16*, 601–608. [\[CrossRef\]](#)
27. Folstein, M.F.; Folstein, S.E.; McHugh, P.R. “Mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **1975**, *12*, 189–198. [\[CrossRef\]](#)
28. Taborri, J.; Palermo, E.; Rossi, S.; Cappa, P. Gait partitioning methods: A systematic review. *Sensors* **2016**, *16*, 66. [\[CrossRef\]](#)
29. Jasiewicz, J.M.; Allum, J.H.; Middleton, J.W.; Barriskill, A.; Condie, P.; Purcell, B.; Li, R.C.T. Gait event detection using linear accelerometers or angular velocity transducers in able-bodied and spinal-cord injured individuals. *Gait Posture* **2006**, *24*, 502–509. [\[CrossRef\]](#)
30. Whittle, M.W. *Gait Analysis: An Introduction*; Butterworth-Heinemann: Oxford, UK, 2014.
31. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [\[CrossRef\]](#)
32. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
33. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *arXiv* **2014**, arXiv:1409.0473.
34. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. Tensorflow: A system for large-scale machine learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), Savannah, GA, USA, 2–4 November 2016; pp. 265–283.
35. Graves, A. Supervised sequence labelling. In *Supervised Sequence Labelling with Recurrent Neural Networks*; Springer: Cham, Switzerland, 2012; pp. 5–13.
36. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 5998–6008.
37. Sarshar, M.; Polturi, S.; Schega, L. Gait phase estimation by using LSTM in IMU-based gait analysis—Proof of concept. *Sensors* **2021**, *21*, 5749. [\[CrossRef\]](#)
38. Fadillioglu, C.; Stetter, B.J.; Ringhof, S.; Krafft, F.C.; Sell, S.; Stein, T. Automated gait event detection for a variety of locomotion tasks using a novel gyroscope-based algorithm. *Gait Posture* **2020**, *81*, 102–108. [\[CrossRef\]](#)
39. Yu, Z.; Zhao, J.; Zhou, X.; Liu, K.; Yan, Y. Gait Phase Detection Based on a Foot-Mounted Inertial Sensor Using Long Short-Term Memory Enhanced by Hidden Markov Model. In Proceedings of the 2021 26th International Conference on Automation and Computing (ICAC), Portsmouth, UK, 2–4 September 2021; pp. 1–5.

40. González, R.C.; López, A.M.; Rodríguez-Uría, J.; Alvarez, D.; Alvarez, J.C. Real-time gait event detection for normal subjects from lower trunk accelerations. *Gait Posture* **2010**, *31*, 322–325. [[CrossRef](#)] [[PubMed](#)]
41. McCamley, J.; Donati, M.; Grimpampi, E.; Mazza, C. An enhanced estimate of initial contact and final contact instants of time using lower trunk inertial sensor data. *Gait Posture* **2012**, *36*, 316–318. [[CrossRef](#)]
42. Soaz, C.; Diepold, K. Step detection and parameterization for gait assessment using a single waist-worn accelerometer. *IEEE Trans. Biomed. Eng.* **2015**, *63*, 933–942. [[CrossRef](#)] [[PubMed](#)]
43. Agostini, V.; Ghislieri, M.; Rosati, S.; Balestra, G.; Knaflitz, M. Surface electromyography applied to gait analysis: How to improve its impact in clinics? *Front. Neurol.* **2020**, *11*, 994. [[CrossRef](#)]
44. Morbidoni, C.; Cucchiarelli, A.; Fioretti, S.; Di Nardo, F. A deep learning approach to EMG-based classification of gait phases during level ground walking. *Electronics* **2019**, *8*, 894. [[CrossRef](#)]
45. Nazmi, N.; Rahman, M.A.A.; Yamamoto, S.I.; Ahmad, S.A. Walking gait event detection based on electromyography signals using artificial neural network. *Biomed. Signal Process. Control* **2019**, *47*, 334–343. [[CrossRef](#)]
46. Perry, J.; Burnfield, J.M. *Gait Analysis—Normal and Pathological Function*, 2nd ed.; Slack: San Francisco, CA, USA, 2010.