

# Learning Based Lower Limb Joint Kinematic Estimation Using Open Source IMU Data

Benjamin Hur<sup>1</sup>, Sunin Baek<sup>1</sup>, Inseung Kang<sup>3,\*</sup>, and Daekyum Kim<sup>1,2,\*</sup>

<sup>1</sup>Korea University, School of Mechanical Engineering, Seoul, 02841, South Korea

<sup>2</sup>Korea University, School of Smart Mobility, Seoul, 02841, South Korea

<sup>3</sup>Carnegie Mellon University, Department of Mechanical Engineering, Pittsburgh, PA 15213, USA

\*corresponding author: inseung@cmu.edu, daekyum@korea.ac.kr

## ABSTRACT

This study introduces a deep learning framework for estimating lower-limb joint kinematics using inertial measurement units (IMUs). While deep learning methods avoid sensor drift, extensive calibration, and complex setup procedures, they require substantial data. To meet this demand, we leveraged an open-source dataset to develop and evaluate three training approaches. The first involved training a model exclusively on data from a single user, resulting in high accuracy for that individual only. The second approach trained a model on data from multiple users to generalize across individuals; however, demonstrated lower accuracy due to variations in gait patterns across users. The third approach added transfer learning to the second, improving estimation accuracy for new users through fine-tuning with a small portion of their data. This model overcame the limitations of the previous methods' dependency on extensive data collection, and achieved comparable performance to inverse kinematics, making it an effective solution for diverse populations. Additionally, our analysis on IMU combinations suggests that IMUs placed on the femur and calcaneus are the best for most cases. This framework not only reduces the need for extensive data collection but also enhances personalized gait analysis, enabling more efficient and accessible applications in both clinical assessments and real-world environments for broader use.

## Introduction

Lower-limb joint kinematic estimation using inertial measurement units (IMUs) holds a great promise in gait biomechanics as IMUs provide a low-cost and easy-to-use gait assessment tool for clinical populations (e.g., stroke or cerebral palsy)<sup>1,2</sup>. Furthermore, estimated joint kinematics can provide information about an individual's postural stability, which has implications to developing tools for fall detection and prevention in the aging population<sup>3-5</sup>. Given that joint kinematics can provide rich and holistic information about an individual's gait quality and stability<sup>6,7</sup>, several recent lower-limb wearable devices, such as robotic exoskeletons and exosuits, have integrated IMUs into their hardware<sup>8-10</sup>.

Previously, analytical methods were commonly used to estimate lower-limb joint kinematics based on inverse kinematics using IMUs<sup>11-13</sup>. However, these approaches require IMUs to be located on each individual limb<sup>13,14</sup>, introducing exhaustive pre-setup requirements (e.g., aligning and calibrating IMUs for each subject). This challenge scales greatly especially in the clinical population where data collection can be expensive. Additionally, IMU-based inverse kinematics suffer from long-term signal drift, primarily due to accumulated errors from time-varying biases in integrating acceleration data<sup>12,13</sup>.

Recently, data-driven approaches have been widely used to address these native challenges that analytical methods have. One major benefit of using a data-driven approach is that it can reduce the total number of required IMUs by automatically extracting the kinematic relationship between different limb segments. This strategy can also potentially mitigate IMU drift by compensating for inherent noise in IMU signals<sup>15,16</sup>. One study generated synthetic inertial data obtained from a clinical marker-based motion tracking database and used it to train a 1D convolutional neural network (1D-CNN) and a Long Short-Term Memory (LSTM) for estimating joint kinematics<sup>17</sup>. LSTM architecture has been used widely in recent literature to estimate joint kinematics in other motor tasks such as Yoga, Golf, and Swimming<sup>18</sup>. Similarly, Deep Convolutional Recurrent Neural Network (DeepConvLSTM) has been utilized in various joint kinematic estimation applications such as estimation during walking and running<sup>19</sup>.

One major bottleneck for a deep learning strategy is its requirement for a large set of data<sup>20</sup>. Developing these deep learning models for various activities generally involve a full optical motion capture(OMC) data collection with IMUs<sup>21,22</sup>. For clinical populations, this OMC data collection can be challenging due to the limited time that each patient can endure. Time-synchronizing and manually labeling of collected OMC and IMU data are also labor-intensive.

To address these challenges, leveraging open-source datasets can be a possible strategy. In domains such as computer vision and natural language processing, researchers have been leveraging open-source datasets, substantially reducing the time and

effort required for data collection and manual labeling<sup>23–25</sup>. These datasets also provide benchmarks to both objectively and quantitatively evaluate a developed algorithm. For joint kinematic estimation, there are several existing open-source datasets that consist of time-synchronized OMC and IMU data. These data cover various lower-limb motor tasks such as sit-to-stand, stand-to-walk, level-ground walking, stairs ascent and descent, ramp ascent and descent, side-step, heel raise, squat, and squat jump<sup>26–31</sup>.

However, despite the advantages of using open-source datasets in creating deep learning models, they have not been utilized to estimate lower-limb kinematics. One challenge that limits the use of these datasets in deep learning models is to accommodate variances exhibited across individuals during different tasks. For example, variations in each individual's joint lengths and gait patterns can cause the model to perform poorly when deploying a trained model to new users<sup>13</sup>. While transfer learning can potentially reduce estimation errors for new users<sup>1,18</sup>, this method is still new in the field and needs to be validated rigorously. Another challenge in the field is the lack of clear guidance in handling open-source datasets towards training a deep learning model. The lack of a unified and holistic understanding in processing and analyzing a dataset leads to inconsistent interpretations across studies. This makes it challenging to benchmark model results and limits researchers to select the proper algorithm for each application. Thus, it is evident that developing a more generalizable model and a systematic approach to handling open-source data for IMU-based kinematic estimation is critical and beneficial for the field.

In this work, we present a deep learning framework that leverages an existing open-source dataset for IMU-based joint kinematic estimation during walking. We systematically compared different models trained with individual user data and those trained without it, providing insights into how model performance varies with different training approaches. We also showed that transfer learning enables researchers to efficiently utilize open-source datasets, minimizing the need for extensive data collection for their own specific motor tasks. Lastly, through comparative analysis, we determined the optimal number of IMUs and their locations. Figure 1 outlines the general flowchart for deep learning model development using IMU and OMC data. Our framework will provide meaningful insight into helping researchers to effectively use open-source datasets in estimating joint kinematics for different target applications.

## Methods

### Open-Source Data

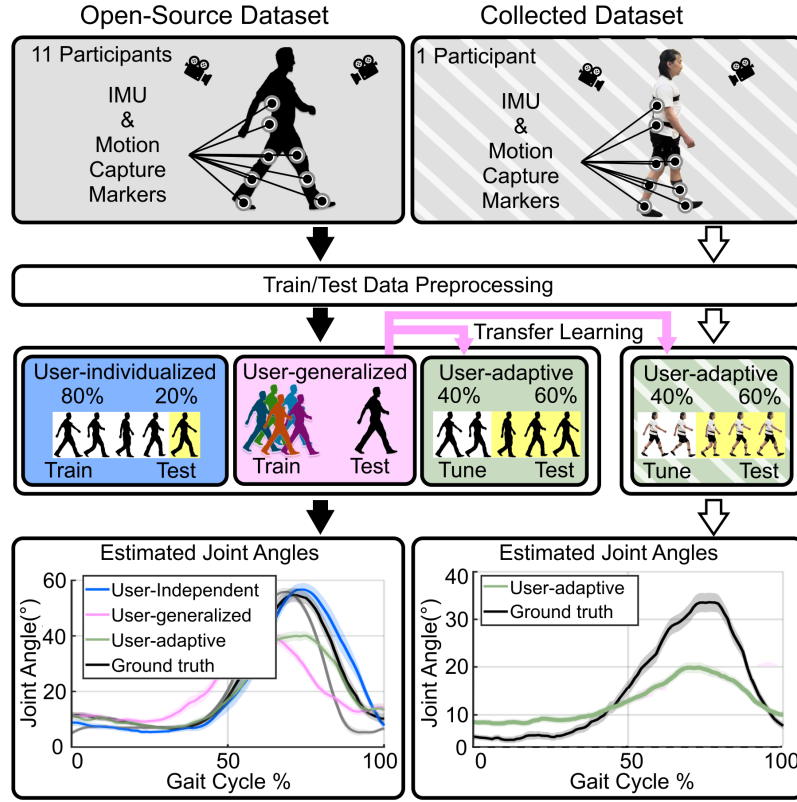
For this study, we used an open-source gait data that included IMU(MTw Awinda, Xsens North America Inc., Culver City, CA, USA) and OMC(Motion Analysis Corporation, Santa Rosa, CA, USA) from 11 healthy individuals (9 males and 2 females)<sup>27</sup>. In this data, subjects had an average age of  $27.9 \pm 6.7$  years and a body mass index of  $23.7 \pm 2.4$  kg/m<sup>2</sup>. This dataset was specifically chosen for its comprehensive representation of lower-limb movements using both IMUs and OMC. The dataset consisted of the following protocol: 1) 5-second calibration pose, 2) 10 minutes of 5-meter straight walking with a 180° turn, 3) additional 5-second calibration pose, and 4) 10 minutes of lower-extremity movements. Each participant was equipped with 8 IMUs located on the torso, pelvis, tibia, femur, and calcaneus to record accelerometer, gyroscope, and magnetometer signals at a sampling rate of 100 Hz. However, for participants(denoted as subjects<sup>27</sup>) 6 and 10, data were sampled at 40 Hz. All OMC data were consistently recorded at 100 Hz sampling rate. We excluded data from participants 5, 9, 10, and 11 due to data corruption, including missing IMU data and synchronization issues between the IMU and OMC. For participant 6, who had a reduced sampling rate on the IMUs, we upsampled the accelerometer, gyroscope, and magnetometer data provided by the IMU using linear interpolation to match the OMC.

### Data Preprocessing

Accurate time synchronization of IMU and OMC data is crucial for precise kinematics analysis. However, open-source datasets often lack time synchronization and have varying lengths between OMC and IMU data. As for the dataset we used in this study, IMU data extended beyond the OMC data. To address this issue, we processed the data to ensure accurate alignment between IMU and OMC data. First, we derived quaternion representations from the rotation matrices of both the IMU data and the OMC marker data. Next, we applied a sliding window in time to the quaternion representation of the IMU data, using a window size matching that of the OMC marker quaternion data. To align the windowed IMU data with the OMC quaternion data, we calculated both the quaternion shortest angle and Euclidean distances between the datasets. Interestingly, the results from both methods were nearly identical. We opted to use the quaternion shortest angle as it is the mathematically correct way to measure distances in orientation space described by quaternions. The starting time point of the window with the smallest quaternion shortest angle was chosen as the synchronization point.

Additionally, we confirmed that time synchronization was successfully completed by comparing the yaw angle data of the torso IMU to the walking model generated in OpenSim. We focused on key moments during turning (start, halfway, and end) and made adjustments as necessary to ensure precise synchronization.

The open-source dataset provides continuous recordings of lower-limb movements, in which the participants perform repeated sequences of straight walking, turning, and resuming straight walking. Since our focus was on kinematics during



**Figure 1.** General flowchart for deep learning model development using IMU and OMC data. The workflow begins with the preprocessing of both open-source and collected datasets. The open-source dataset is then employed to train CNN/LSTM models aimed at estimating joint kinematics through both user-individualized and user-generalized methods. Following this, the user-generalized model serves as the foundation for the transfer learning approach (user-adaptive method), which is utilized to evaluate the model’s adaptiveness. Additionally, the collected dataset is used to train LSTM models specifically for user-adaptive methods.

straight walking, we extracted only the straight walking segments from the dataset and excluded the corner sections. This approach aligns with general clinical gait assessments, such as the 10-meter walk test, which involve only straight walking<sup>32</sup>. Additionally, the direction and method of turning varied among participants, with some changing direction without stopping while others stopped to rotate in place, introducing variance. To focus on consistent gait patterns, we excluded the rotation and turning segments.

We first analyzed the yaw angle data of the torso-mounted IMU at each time point. To calculate the yaw angle at time =  $T$ , we added up the yaw angle differences at each timestep, starting from time = 0. Specifically, we calculated the change in yaw angle from time = 0 to time = 0.01, then from time = 0.01 to time = 0.02, and so on, up to the change from time =  $T-1$  to time =  $T$ . By summing all these yaw angle differences, we obtained the total yaw angle change from time = 0 to time =  $T$ . This angle was used to analyze the walking direction. We then found segments of straight walking by selecting periods where the yaw angle remained within the range of  $n \times \pi \pm 30^\circ$  (where  $n$  is an integer), inferring minimal lateral rotation. These segments were then categorized and analyzed as straight walking data.

To ensure consistency across individuals in the dataset and minimize potential biases from variations in IMU placement, we normalized the IMU data before inputting it into our deep learning models. This step involved converting the IMU rotation matrix data into Euler angles to standardize orientation representation. We then normalized the accelerations, the angular velocities, and the Euler angles, to have zero mean and unit variance. Since all participants in the dataset performed standardized walking tests with consistent IMU placements, we assessed that normalizing the Euler angles would effectively capture the changes in orientation for each axis within the predefined range of motion observed during the tests. By standardizing these angles, we aimed to ensure the model could focus on relative orientation changes rather than absolute values, which is particularly useful for identifying movement patterns. This process enables more accurate comparisons across subjects and enhances the deep learning model’s ability to generalize from the training data. The entire preprocessing workflow was

conducted using Matlab R2023b(The MathWorks Inc., Natick, MA, USA).

### Data Collection for Model Adaptability Verification

We collected additional data to evaluate our user-adaptive model, which applies transfer learning to unseen data using a pre-trained model and an independent dataset. Our dataset includes IMU and OMC data from three individuals performing a 5-minute session of 3 meters straight walking and 180° turns at a self-selected pace. Before walking, the participants were told to do a 5-second calibration pose to perform marker-to-segment calibration via Opensim static marker position calibration. Eight IMUs (Xsens DOT, Xsens North America Inc., Culver City, CA, USA) were placed at the same locations and their coordinate axes were manually aligned according to the same local frame orientations described in the open-source dataset. Each IMU's orientation data, originally expressed in the IMU global reference frame, were rotated into the local laboratory reference frame (OMC frame). Specifically, we first converted the IMU's Euler angles into a rotation matrix. Then we multiplied this rotation matrix to the transformation matrix, which is defined by how the IMU global frame is oriented relative to the local laboratory reference frame. By multiplying these matrices, we obtained a new rotation matrix that represents the IMU's orientation in the local laboratory reference frame (OMC frame). The OMC data (OptiTrack, NaturalPoint Inc. Corvallis, OR, USA) were sampled at 100 Hz. The accelerometer, gyroscope, and magnetometer signals were recorded at a sampling rate of 60 Hz and were interpolated over time to match the sampling rate of 100Hz. For this collected dataset, we performed preprocessing steps of extracting straight walking and normalizing IMU data.

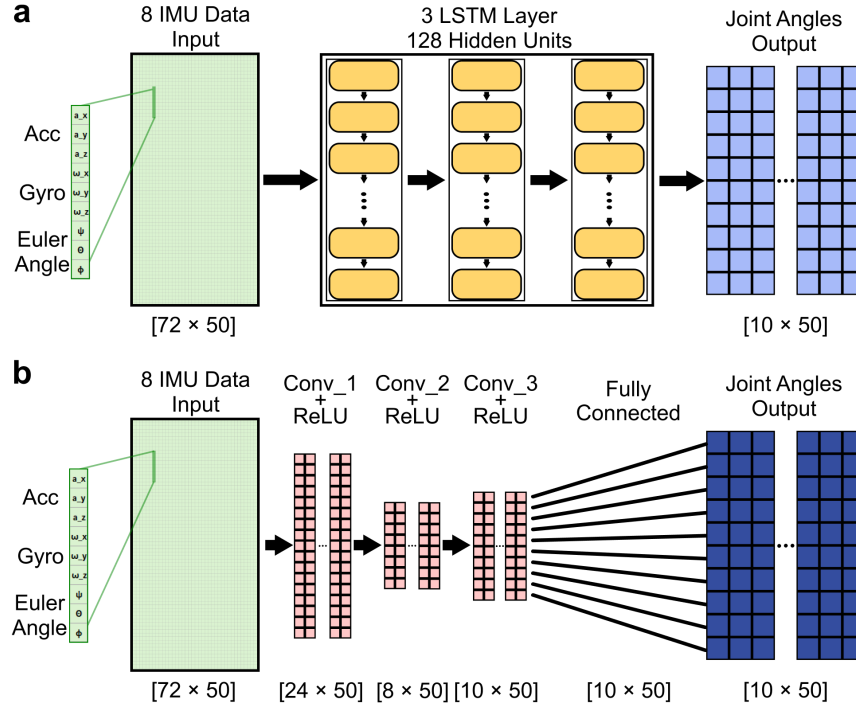
### Deep Learning

In developing our model for estimating lower-limb kinematics during gait, we used Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTMs) networks, which are widely used and are proven to be effective in deep learning applications for joint kinematic estimation<sup>26,33</sup>. CNN are renowned for their ability to extract features through convolutional filters, making them adept at identifying spatial hierarchies in data. This capability is particularly beneficial for analyzing the complex patterns present in kinematic data<sup>34</sup>. LSTM excel at capturing temporal dependencies by using information from previous time steps to enhance current predictions, which makes them especially suited for the sequential nature of gait data<sup>35</sup>. To optimize these models, we employed the Adam optimizer, with a set learning rate of 0.001. The determination of the best epoch for each model was carried out by incrementing the epoch count in steps of 200, starting from 500 and capping at 3,000 epochs. Early stopping was applied if the validation loss did not decrease at the subsequent epoch increment. This approach allowed us to methodically identify the epoch at which the model achieved the lowest root mean square error (RMSE) on validation data, ensuring optimal model performance without overfitting.

In constructing our models to estimate lower-limb kinematics, we processed the data from 8 IMUs to have 72 variables for each time step as a deep learning input, encompassing x, y, z measurements of acceleration, gyroscope, and Euler angles. Figure 2 illustrates the schematic diagram for both the LSTM and CNN models. For the LSTM model, a structure with hidden dimensions of 128 and three layers were used to create the model. For the CNN model, a sequence of 1D convolution layers was used. The first layer consisted of a 1D convolution with 24 channels and a kernel size of 5, selected to capture the essential features without excessive computational load. The second and third layer used 1D convolutions with 8 and 10 channels respectively, maintaining the same kernel size to ensure consistency in feature extraction across layers. Each convolution step used the ReLU activation function to introduce non-linearity and enhance model learning capabilities. Additionally, sliding window technique was applied to the input data. Window size was set to 50 (0.5 seconds). This window size ensures that the input data does not contain data of a whole gait, so that the model is able to estimate joint kinematics without requiring every part of data from a gait. The sliding stride length was set to 25 (0.25 seconds) so that each data contributes to the model consistently.

We evaluated three different methods of selecting training data to develop deep learning models that estimate lower-limb kinematics: the 'user-individualized method', the 'user-generalized method', and the 'user-adaptive method'. Each approach was designed to tackle the inherent challenges of kinematic estimation from different perspectives.

- *User-individualized method(UI)*: In this approach, the model is trained using the first 80% of each participant's data, with the remaining 20% reserved for testing. Additionally, 20% of the training data is set aside as a validation set. This method focuses on optimizing the model's performance for known participants. However, this method may be limited in its ability to generalize to new data(participants).
- *User-generalized method(UG)*: Unlike the user-individualized method, this approach trains the model using data from a specific group of participants (1, 2, 3, 4, and 6), validates it with data from another participant (7), and subsequently tests it with data from a final participant (8). This method aims to assess the model's ability to generalize to new, unseen individuals.



**Figure 2.** Schematic diagram for **a** LSTM and **b** CNN models. Total of eight IMUs, positioned on the torso, pelvis, bilateral femurs, tibias, and calcanei, are used as model inputs. The model outputs are the bilateral sagittal, frontal, and transverse hip, sagittal knee, and sagittal ankle joint angles.

- *User-adaptive method(UA)*: This method combines the strengths of the user-individualized and user-generalized methods while minimizing their weaknesses to enhance generalizability without sacrificing performance. It starts with a model pre-trained using the user-generalized method and then fine-tunes it with data from a target participant. By adjusting the train-test split ratios and employing a detailed epoch selection process (from 1 up to a maximum of 15 epochs), this approach aims to achieve both the generalization of the user-generalized method and the specificity of the user-individualized method.

### IMU Combination

CNN and LSTM were trained using different combinations of IMUs to cover all possible configurations of the pelvis, both tibias, femurs, and calcanei. This approach allowed us to assess both the individual contributions of each IMU and their combined effects on model performance. We tested this using all three methods—UI, UG, and UA approaches—to develop a total of 378 LSTM models, with 126 models created for each method.

### Inverse Kinematics

For the ground truth values of joint angles, we used OMC-based inverse kinematics calculated through OpenSim 4.4. The open-source dataset provided scaled OpenSim models with calibrated markers. Using these models, we performed inverse kinematics in OpenSim with the OMC data. For the dataset we collected, we created an OpenSim model based on the base model from the open-source dataset. This model included a physiological skeletal structure with 22 segments and 43 degrees of freedom. We calibrated the marker positions using a static pose calibration. As with the open-source dataset, we ran OpenSim inverse kinematics to solve for the joint angles.

For the baseline when evaluating our deep learning models, we used IMU-based inverse kinematics. We followed the steps outlined in the original study that provided the dataset<sup>27</sup>. The calibration pose data included in the open-source dataset was used to determine the relative orientation of the IMU to each body segment. When solving the inverse kinematics, we utilized the setup weights provided in the dataset, which reduced the influence of the IMUs on the shanks and feet by lowering the relative weights of those IMUs that were closer to the ground. We did not perform IMU-based inverse kinematics for our collected dataset, as this dataset was specifically gathered to evaluate the performance of the user-adaptive method on unseen data.



## Results Analysis

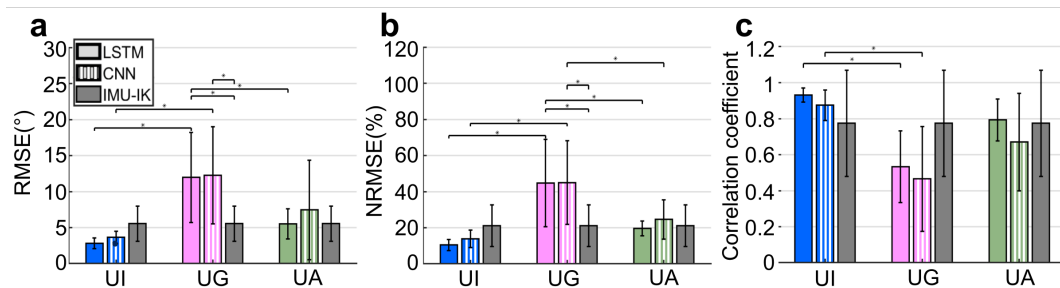
To evaluate the performance of the UI, UG, and UA, we segmented the ground truth and estimated joint kinematics into gait cycles, defined by heel strike events. Based on a previous study on joint angles during walking<sup>36</sup>, the hip flexion peak typically occurs at approximately 85% of the gait cycle. To align our analysis with this benchmark, we performed peak detection on the sagittal hip angle data and then shifted it by 15%. Additionally, during cornering, some participants walked through the turn while others stopped and rotated in place. This variability, combined with the exclusion of corner sections, resulted in some gait cycles being incomplete or partial, slightly altering the averaged gait pattern. However, as the primary focus of this study is on estimation performance rather than the precise gait cycle representation, we included all gaits in the result analysis. We used the last 20% of data from one specific participant, which is a data segment included in the test data for all six models (three methods applied to two architectures), to analyze the estimation results.

To compare the estimation performances of the deep learning models, we calculated the root mean square error (RMSE) between the joint angles estimated by the models and those calculated from OpenSim OMC-based inverse kinematics. We also computed the normalized root mean square error (NRMSE) because each joint's range of motion differs. This is important because the same RMSE value may indicate a larger error percentage for a joint with a smaller range of motion compared to one with a larger range, leading to inconsistent assessments of estimation accuracy. To calculate NRMSE, we divided the RMSE by the range of motion for each joint. Additionally, we calculated the Pearson correlation coefficient for each gait cycle and joint. This was done to examine how closely the patterns of the estimation results align with the ground truth throughout a gait cycle. To further analyze the performance across different methods and architectures, we conducted a two-way ANOVA test. For pairwise comparisons, we used the Bonferroni post-hoc correction with a significance level of  $\alpha = 0.05$ . All statistical analyses were performed using MATLAB R2023b.

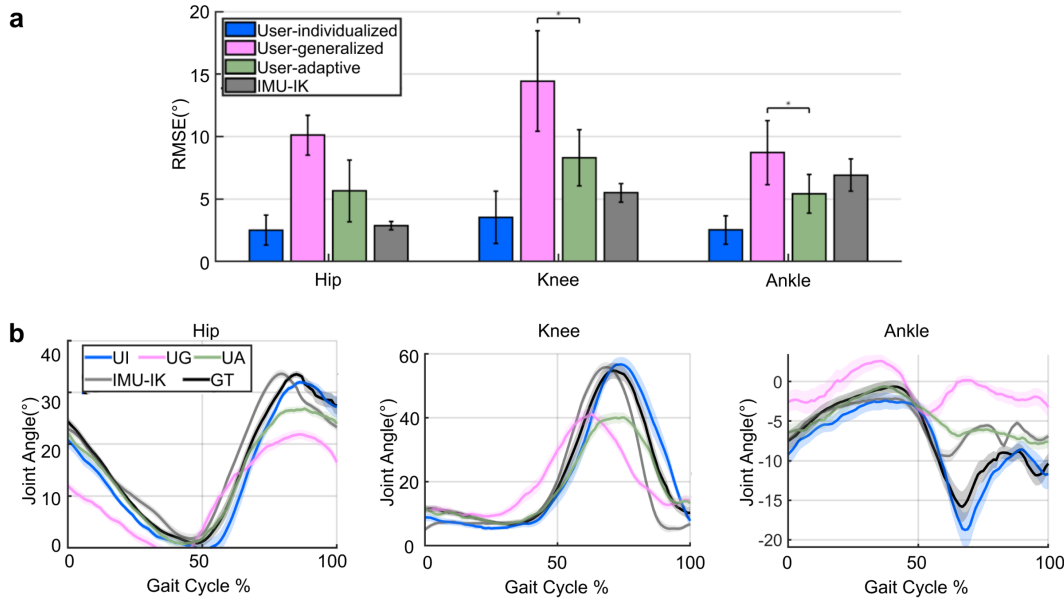
## Results

### Overall model performance comparison

The UI outperformed IMU-based inverse kinematics in RMSE values, NRMSE values, and Pearson correlation coefficient for both LSTM and CNN models. Specifically, the LSTM model showed 49.20% lower average RMSE, 50.65% lower average NRMSE, and 20.13% higher correlation coefficient compared to IMU-based inverse kinematics. The UI demonstrated the smallest standard deviation across joints. Compared to RMSE values of IMU-based inverse kinematics, the UG showed higher error, with RMSE values 115.87% and 121.50% larger for the LSTM and CNN models, respectively. It showed 117.60% and 112.83% larger NRMSE values, along with 33.33% and 42.84% lower Pearson correlation coefficients. We tested whether the performance degradation of the UG could be overcome through transfer learning. The average RMSE and NRMSE for the UA LSTM model were slightly lower than those of IMU-based inverse kinematics (0.4% and 7% respectively). The Pearson correlation coefficient was 0.024% higher for the UA than for the IMU-based inverse kinematics. Although the average RMSE and NRMSE values were similar between the UA and IMU-based inverse kinematics, the UA shows substantially smaller standard deviation across joints in NRMSE and correlation coefficient compared to those of IMU-based inverse kinematics (65.03% and 60.73% smaller, respectively). This highlights that the UA enhances the consistency of model performance across different joints. Across all methods, the LSTM model outperformed the CNN model. Figure 3 presents the average RMSE, NRMSE, and Pearson correlation coefficient values for joint angles estimated by each method compared to OpenSim inverse kinematics, with error bars representing standard deviations and statistical significance indicated.



**Figure 3.** **a** Average RMSE values for joint angles estimated by each method compared to OpenSim inverse kinematics of the OMC data. **b** Average NRMSE values for joint angles estimated by each method compared to OpenSim inverse kinematics of the OMC data. **c** Average Pearson correlation coefficients for joint angles estimated by each method compared to OpenSim inverse kinematics of the OMC data. Error bars represent  $\pm 1$  standard deviation. Asterisks indicate statistical significance ( $p < 0.05$ ).



**Figure 4.** **a** Average RMSE values over gaits for joint angles estimated by LSTM models and IMU-based inverse kinematics compared to OpenSim inverse kinematics of the OMC data. Error bars represent  $\pm 1$  standard deviation. Asterisks indicate statistical significance ( $p < 0.05$ ). **b** Average estimated sagittal joint angles for the hip, knee, and ankle during the gait cycle, calculated using the LSTM architecture. IMU-IK is the IMU-based inverse kinematics, and GT is the ground truth from mocap-based inverse kinematics. Gaits with durations less than 70% or greater than 130% of the average gait time were excluded, resulting in 32 gaits used to generate this graph. The bands represent the standard error of the mean.

	UI				UG				UA			
	Hip	Knee	Ankle		Hip	Knee	Ankle		Hip	Knee	Ankle	
C	4.14	4.54	2.58	+	14.88	17.03	8.28	+	14.01	16.91	7.61	+
F	3.83	5.82	3.18	↑	14.90	17.05	8.30	↑	14.03	16.98	7.64	↑
P	6.31	7.97	4.00		14.90	17.06	8.30		14.02	17.00	7.64	
T	4.01	4.94	2.87		14.29	16.71	8.00		13.43	16.29	7.33	
F_C	3.34	4.12	2.57		13.07	16.03	8.17		11.81	15.63	6.99	
P_C	4.01	4.70	2.67		14.89	17.06	8.29		14.07	16.95	7.64	
P_F	3.70	5.28	3.00		14.93	17.06	8.30		14.00	16.99	7.63	
P_T	3.84	4.92	2.79		13.86	15.92	8.09		12.42	16.10	7.05	
T_C	3.64	4.15	2.45		14.40	16.33	7.98		13.02	15.99	7.22	
T_F	3.54	4.64	2.70		12.95	15.50	7.93		12.21	15.35	7.40	
P_F_C	3.09	4.26	2.50		13.43	15.82	8.22		11.59	14.22	7.00	
P_T_C	3.45	4.12	2.54		13.52	15.79	8.11		12.12	15.65	7.15	
P_T_F	3.77	4.66	2.92		13.29	16.34	8.28		11.24	15.32	7.00	
T_F_C	3.28	4.04	2.56	-	13.10	14.93	7.97	-	11.50	14.27	6.69	-

C: Calcaneus  
F: Femur  
P: Pelvis  
T: Tibia

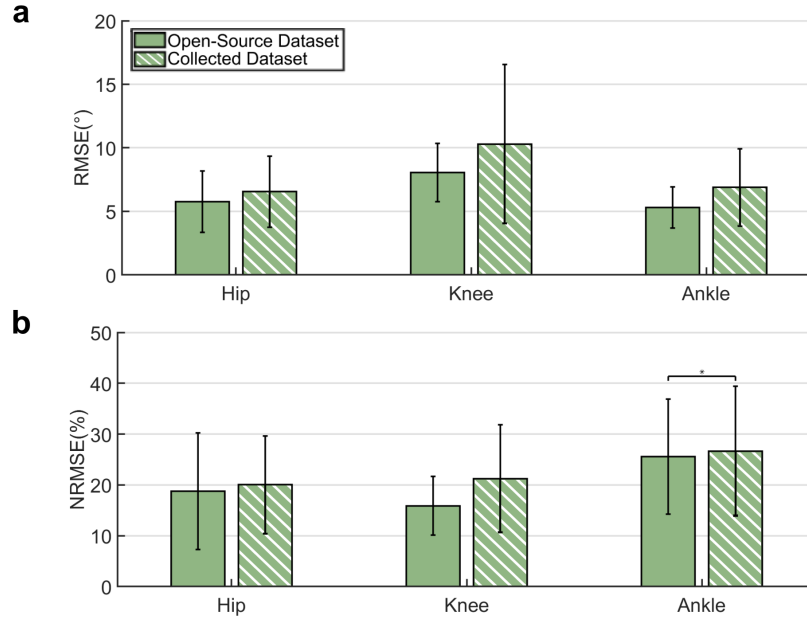
F\_C: Femur & Calcaneus  
P\_C: Pelvis & Calcaneus  
P\_F: Pelvis & Femur  
P\_T: Pelvis & Tibia  
T\_C: Tibia & Calcaneus  
T\_F: Tibia & Femur

P\_T\_F: Pelvis & Tibia & Femur  
T\_F\_C: Tibia & Femur & Calcaneus  
P\_F\_C: Pelvis & Femur & Calcaneus  
P\_T\_C: Pelvis & Tibia & Calcaneus

**Figure 5.** RMSE values for different IMU combinations. Darker colors in each column indicate higher RMSE values.

#### Error for each joint

Figure 4-a presents the RMSE values for each joint and method. Figure 4-b presents the average sagittal hip, knee, and ankle joint angles over gait cycles. The RMSE difference between the UI and the UG was large but was not statistically significant. However, for the UA, sagittal knee and ankle angle estimation results showed a statistical significant reduction ( $p = 0.0044$  and



**Figure 6.** Comparison of joint angle estimation errors for each dataset. **a** RMSE values for each joint angle. **b** NRMSE values for each joint angle. Error bars represent  $\pm 1$  standard deviation.

$p = 9.71 \times 10^{-7}$ , respectively) in RMSE values compared to results of the UG.

### IMU Combinations

In our analysis on various IMU combinations, we calculated the RMSE values for each combination, thereby identifying the best combinations for each joint. We performed this procedure for all three methods (UI, UG, and UA). Figure 5 shows results from the IMU combinations for sagittal hip, knee, and ankle joint angle. Across all methods, IMU combinations including femur and calcaneus showed the lowest RMSE values for sagittal hip and knee joint angle estimation. For the sagittal ankle angle estimation, IMU combinations that included the tibia also demonstrated strong performance, with average RMSE values of  $2.69^\circ$  for the UI,  $8.05^\circ$  for the UG, and  $7.12^\circ$  for the UA. This indicates that IMUs on the hip and feet provide the most information for estimating lower-limb joint kinematics during gait. IMUs on the shank also provide useful data, whereas IMUs on the pelvis do not yield beneficial information under the study conditions.

### Transfer learning using our IMU data

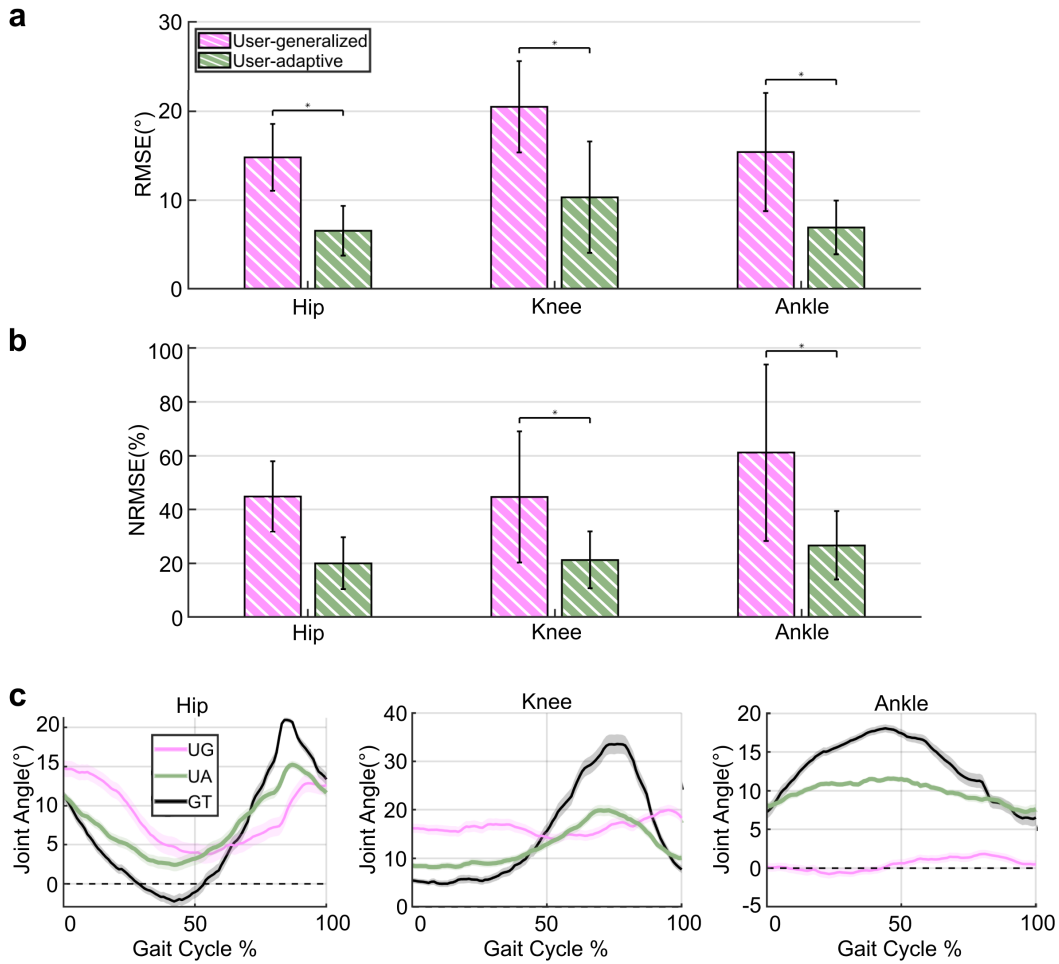
We evaluated the performance of our UA by applying it to both the open-source dataset and the collected dataset. We compared the results from train/validation/test ratio of 20/20/60 because it was the best split ratio for UA when applied to the open-source dataset. For the open-source dataset, we found an RMSE of  $6.38 \pm 1.47^\circ$  and for the collected dataset,  $7.91 \pm 2.07^\circ$ . Applying transfer learning with the collected dataset substantially improved the model's estimation accuracy, bringing it to a level comparable to that achieved with the open-source dataset. This indicates that the UA is effective not only for the original open-source dataset, but also for the newly collected dataset. Figure 6 presents the RMSE and NRMSE values of each joint angle for each dataset. Figure 7 presents the comparison of joint angle estimation errors for UG and UA along with estimated joint angle results.

## Discussion

We presented benchmarks for deep learning models designed to estimate lower-limb joint kinematics, using an open-source IMU dataset. Three different methods were implemented: UI, UG, and UA. These methods were applied to two different deep learning architectures, CNN and LSTM networks. Our results demonstrated that applying transfer learning to a generalized pre-trained model improves estimation performance by 53.86%, achieving accuracy comparable to that of IMU-based inverse kinematics. Additionally, our comparative analysis of IMU sensor placements revealed that positioning the sensors on the femur and calcaneus provides the most accurate estimation of lower-limb kinematics during gait.

While estimation performances of UA were on par with previous studies<sup>17,26,27,37</sup>, there is still potential for improvement through the integration of additional data. Joint kinematic estimation performance of models developed by UI and UA showed





**Figure 7.** Comparison of joint angle estimation errors for UG and UA. **a** RMSE values for each joint angle. **b** NRMSE values for each joint angle. Error bars represent  $\pm 1$  standard deviation. Asterisks indicate statistical significance ( $p < 0.05$ ). **c** Average estimated joint angles during the gait cycle for each method. GT is the ground truth from mocap-based inverse kinematics. Gaits with durations less than 70% or greater than 130% of the average gait time were excluded, resulting in 119 gaits to generate this graph. The bands represent the standard error of the mean.

RMSE values of  $2.81 \pm 0.75^\circ$  and  $5.52 \pm 2.10^\circ$ , respectively, which falls in the range of previous studies which reported RMSE values under  $5^\circ$ <sup>17,26,27,37</sup>. However, models developed using UG showed higher errors, with RMSE values exceeding  $10^\circ$ . The UG model's ability to generalize is constrained by the size and diversity of its training dataset, requiring a larger and more diverse dataset to fully perform. Given these limitations, UG is insufficient for generalizing to new users, making the application of transfer learning with a small amount of target user data(UA) highly valuable. Since UG served as the base model for the UA, performance improvements of the UG model could also enhance performance of the UA model<sup>38</sup>. As model generalizability typically improves with larger training datasets, we expect that incorporating additional data into the UG training process will enhance its performance<sup>39-41</sup>. One way to achieve this is by including other open-source datasets<sup>42</sup> and processing them based on the guidelines provided in our paper. Another approach is to generate synthetic IMU data to further increase the available data pool<sup>39</sup>.

The UA enhanced the performance of joint kinematic estimation for participants whose data were not used for training. This demonstrated its potential for effective model personalization. In previous research, transfer learning showed an improvement in estimation accuracy of approximately 10 to  $15^\circ$  for joint kinematic estimation of unseen activities<sup>18</sup>. While the prior study focused on unseen activities, our work applied transfer learning to unseen users. Despite this difference, we observed similar performance improvements, further confirming that transfer learning is effective for joint kinematic estimation in various contexts. Moreover, the enhanced estimation accuracy of the UA applied to both the open-source and the collected datasets suggests that transfer learning is effective not only for the unseen participants within the same dataset, but also for those from

newly collected datasets using a different IMU set<sup>18</sup>. This demonstrated that a personalized joint kinematic estimation model can be constructed by developing a generalized pre-trained model using open-source datasets from prior studies and applying transfer learning with a small portion of a novel individual's data. Because only a minimal amount of new data is necessary, this approach significantly reduced both the time and cost associated with data collection and model training<sup>43,44</sup>.

We further showed the best IMU placements for estimating lower-limb joint kinematics, allowing for reduction in the number of IMUs needed. Consistent with previous study, our findings confirm that femur and calcaneus are the most critical IMU placements, whereas the pelvis is the least important<sup>14</sup>. However, among the combinations of two IMUs, the combination of pelvis and tibia showed better accuracy compared to combinations of the pelvis with either the calcaneus or the femur. This suggests that the synergy between IMUs has a more effect on estimation accuracy than relying solely on the precision of individual IMUs. This result differs from the previous study<sup>14</sup>, where feet and thigh IMU combinations provided the best accuracy during their inter-subject examination. In contrast, our findings highlight that combinations including the tibia offered the lowest RMSE values for generalized models. This can also be seen in the best combinations for specific joints. For sagittal ankle estimation, combinations including the tibia and calcaneus provided the highest accuracy, while for sagittal knee joint estimation, the femur and tibia or calcaneus combinations showed the best results. These findings emphasize that IMUs positioned near joints or able to capture the full range of motion are the most effective, reinforcing that optimal IMU synergy is key to improving estimation accuracy. By determining these best IMU combinations, we can offer reference for reducing the number of IMUs required for joint kinematic estimations. This reduction considerably lowers the computational load for deep learning models, enhancing the feasibility of real-time application by allowing faster processing and reduced latency<sup>43,44</sup>.

The results for UA applied with different train/test split ratio shows the possibility of optimal amount of training data for transfer learning. Previous studies typically used 20 to 25% of the target data for transfer learning<sup>18,45,46</sup>. Results from the UA applied to the open-source dataset confirm these findings, showing that model accuracy does not considerably improve once the transfer rate exceeds 20%. However, when applied to the collected dataset, the UA achieved maximum accuracy when the transfer rate exceeded 35%. One possible explanation for this discrepancy is that the model in our collected dataset-based UA transferred data from a new dataset, rather than from within the same dataset. Although we maintained fundamental aspects of the data collection protocol, such as IMU placements and tasks performed, variations in sensor modules and environmental conditions may have necessitated additional data for effective model adaptation through transfer learning.

Several limitations in this study should be addressed in future work. One key limitation is that the UA was tested on only a few participants. While this provided an initial proof of concept, further testing across a broader range of datasets is necessary to gain a more comprehensive understanding of its performance. Gait patterns can vary considerably between individuals, especially in different contexts such as age, gender, or health status. Testing the model on diverse datasets that reflect a wide range of gait types, including those of participants with abnormal walking patterns or varying speeds, will help determine how effectively the UA generalizes across different population groups. Moreover, evaluating the estimation performance for clinical populations with distinct walking patterns, such as those with gait impairments or using assistive devices, could provide valuable insights into its clinical applicability. Identifying how these differences influence the accuracy of the UA model is critical for refining its performance for clinical settings. Additionally, expanding the dataset types to include various environments, such as indoor versus outdoor walking or different types of terrains, would be valuable in understanding the adaptability of the model in real-world scenarios. Such diversity is essential to better evaluate the optimal transfer rate for different types of gaits and ensure the model's robustness in handling a wide array of motion data. Another limitation is the training of the base model, which relied on a limited set of participants. The pre-trained base model used in this study, referred to as the UG, was developed with data from a relatively small group of individuals, resulting in lower estimation accuracy. Transfer learning models are highly influenced by the quality and diversity of the data used in pre-training, as it sets the foundation for how well the model adapts to new data<sup>47</sup>. Expanding the base model with more diverse training data is crucial for improving the accuracy of the UA. Incorporating a larger pool of participants representing different demographics, such as age groups, physical conditions, or movement styles, will enable the base model to capture a broader spectrum of gait variations. Additionally, including data from various types of walking activities (e.g., brisk walking, slow walking, walking with assistive devices) in the base model can help ensure the transfer learning process is more effective when applied to target datasets with different characteristics. This approach would likely lead to higher overall accuracy and better generalizability when adapting the model to new datasets in future studies. Furthermore, although our method is readily applicable when open-source data includes both IMU and OMC, it faces a limitation if a dataset only provides IMU data, since validation would require OMC. Additionally, the UA requires a small amount of marker-based OMC data for transfer learning, indicating that continued dependence on OMC systems remains a limitation. To address this limitation, future research could consider using alternative methods of extracting kinematic data, such as OpenCap<sup>48</sup> which can utilize smartphone cameras, to reduce or eliminate the need for traditional OMC systems. This strategy would not only broaden accessibility but also enhance the practicality of the system in various real-world scenarios.

This work provides guidelines for utilizing transfer learning with LSTM and CNN models for lower-limb joint kinematic estimation, offering a practical solution for adapting models to individual users and varied conditions. This method substantially

enhances the applicability of open-source IMU datasets for personalized gait analysis, fostering opportunities for broader adoption in both clinical and commercial settings.

## Data availability

The datasets and codes that support the findings of this study are available on our GitHub repository at <https://github.com/mintlabkorea/Learning-Based-Lower-Limb-Joint-Kinematic-Estimation-Using-Open-Source-IMU-Data>.

## References

1. Nuckols, R. W., et al. Design and evaluation of an independent 4-week, exosuit-assisted, post-stroke community walking program. *Annals of the New York Academy of Sciences*, 2023.
2. Revi, D. A., et al. Indirect measurement of anterior-posterior ground reaction forces using a minimal set of wearable inertial sensors: From healthy to hemiparetic walking. *Journal of NeuroEngineering and Rehabilitation*, 17(1), 1–13, 2020.
3. McManus, K., et al. Development of data-driven metrics for balance impairment and fall risk assessment in older adults. *IEEE Transactions on Biomedical Engineering*, 69(7), 2324–2332, 2022.
4. Liu, J., Zhang, X., and Lockhart, T. E. Fall risk assessments based on postural and dynamic stability using inertial measurement unit. *Safety and Health at Work*, 3(3), 192–198, 2012.
5. Subramaniam, S., Faisal, A. I., and Deen, M. J. Wearable sensor systems for fall risk assessment: A review. *Frontiers in Digital Health*, 4, 921506, 2022.
6. Topham, L.K., et al. Gait identification using limb joint movement and deep machine learning *IEEE Access*, 10, pp.100113-100127, 2022.
7. Sabo, A., et al. Estimating parkinsonism severity in natural gait videos of older adults with dementia *IEEE journal of biomedical and health informatics*, 26(5), pp.2288-2298. 2022
8. Kang, I., Kunapuli, P., and Young, A. J. Real-time neural network-based gait phase estimation using a robotic hip exoskeleton. *IEEE Transactions on Medical Robotics and Bionics*, 2(1), 28–37, 2019.
9. Kang, I., Hsu, H., and Young, A. The effect of hip assistance levels on human energetic cost using robotic hip exoskeletons. *IEEE Robotics and Automation Letters*, 4(2), 430–437, 2019.
10. Kang, I., et al. Real-time gait phase estimation for robotic hip exoskeleton control during multimodal locomotion. *IEEE Robotics and Automation Letters*, 6(2), 3491–3497, 2021.
11. O'Reilly, M., et al. Wearable inertial sensor systems for lower limb exercise detection and evaluation: A systematic review. *Sports Medicine*, 48, 1221–1246, 2018.
12. Picerno, P. 25 years of lower limb joint kinematics by using inertial and magnetic sensors: A review of methodological approaches. *Gait & Posture*, 51, 239–246, 2017.
13. Hafer, J. F., et al. Challenges and advances in the use of wearable sensors for lower extremity biomechanics. *Journal of Biomechanics*, 2023.
14. Moghadam, S. M., Yeung, T., and Choisne, J. The Effect of IMU Sensor Location, Number of Features, and Window Size on a Random Forest Model's Accuracy in Predicting Joint Kinematics and Kinetics During Gait. *IEEE Sensors Journal*, 2023.
15. Chen, H., et al. Improving inertial sensor by reducing errors using deep learning methodology. *NAECON 2018-IEEE National Aerospace and Electronics Conference*, IEEE, 2018.
16. Brossard, M., Bonnabel, S., and Barrau, A. Denoising IMU gyroscopes with deep learning for open-loop attitude estimation. *IEEE Robotics and Automation Letters*, 5(3), 4796–4803, 2020.
17. Rapp, E., et al. Estimation of kinematics from inertial measurement units using a combined deep learning and optimization framework. *Journal of Biomechanics*, 116, 110229, 2021.
18. Li, J., et al. 3D knee and hip angle estimation with reduced wearable IMUs via transfer learning during yoga, golf, swimming, badminton, and dance. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2024.
19. Hernandez, V., et al. Lower body kinematics estimation from wearable sensors for walking and running: A deep learning approach. *Gait & Posture*, 83, 185–193, 2021.

20. Bansal, A., Sharma, R., and Kathuria, M. A systematic review on data scarcity problem in deep learning: solution and applications. *ACM Computing Surveys (CSUR)*, 54(10s), 1–29, 2022.
21. Lee, C. J., and Lee, J. K. Inertial motion capture-based wearable systems for estimation of joint kinetics: A systematic review. *Sensors*, 22(7), 2507, 2022.
22. Weygers, I., et al. Inertial sensor-based lower limb joint kinematics: A methodological systematic review. *Sensors*, 20(3), 673, 2020.
23. Tsipras, D., et al. From ImageNet to image classification: Contextualizing progress on benchmarks. *International Conference on Machine Learning*, PMLR, 2020.
24. Lin, T. Y., et al. Microsoft COCO: Common objects in context. *Computer Vision–ECCV 2014: 13th European Conference*, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13, Springer International Publishing, 2014.
25. Wang, A., et al. GLUE: A multi-task benchmark and analysis platform for natural language understanding. *arXiv preprint arXiv:1804.07461*, 2018.
26. Moghadam, S. M., Yeung, T., and Choisine, J. A comparison of machine learning models' accuracy in predicting lower-limb joints' kinematics, kinetics, and muscle forces from wearable sensors. *Scientific Reports*, 13(1), 5046, 2023.
27. Al Borno, M., et al. OpenSense: An open-source toolbox for inertial-measurement-unit-based measurement of lower extremity kinematics over long durations. *Journal of NeuroEngineering and Rehabilitation*, 19(1), 1–11, 2022.
28. Dimitrov, H., Bull, A. M., and Farina, D. High-density EMG, IMU, kinetic, and kinematic open-source data for comprehensive locomotion activities. *Scientific Data*, 10(1), 789, 2023.
29. Lavikainen, J., et al. Open-source software library for real-time inertial measurement unit data-based inverse kinematics using OpenSim. *PeerJ*, 11, e15097, 2023.
30. Camargo, J., et al. A comprehensive, open-source dataset of lower limb biomechanics in multiple conditions of stairs, ramps, and level-ground ambulation and transitions. *Journal of Biomechanics*, 119, 110320, 2021.
31. Topham, L.K., et al. A diverse and multi-modal gait dataset of indoor and outdoor walks acquired using multiple cameras and sensors. *Scientific Data*, 10(1), p.320, 2023.
32. Scivoletto, G., et al. Validity and reliability of the 10-m walk test and the 6-min walk test in spinal cord injury patients. *Spinal cord*, 49(6), 736–740, 2011.
33. Robert-Lachaine, X., et al. Accuracy and repeatability of single-pose calibration of inertial measurement units for whole-body motion analysis. *Gait & Posture*, 54, 80–86, 2017.
34. Alzubaidi, L., et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, 1–74, 2021.
35. Graves, A. Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*, 2013.
36. Williams, Gavin, et al. The effect of walking speed on lower limb angular velocity. *Archives of Physical Medicine and Rehabilitation*, 99.10 (2018): e12-e13.
37. Gholami, M., et al. Lower body kinematics monitoring in running using fabric-based wearable sensors and deep convolutional neural networks. *Sensors*, 19(23), 5325, 2019.
38. Entezari, R., et al. The role of pre-training data in transfer learning. *arXiv preprint arXiv:2302.13602*, 2023.
39. Sharifi Renani, M., et al. The use of synthetic IMU signals in the training of deep learning models significantly improves the accuracy of joint kinematic predictions. *Sensors*, 21(17), 5876, 2021.
40. Hestness, J., et al. Deep learning scaling is predictable, empirically. *arXiv preprint arXiv:1712.00409*, 2017.
41. Sun, C., et al. Revisiting unreasonable effectiveness of data in deep learning era. *Proceedings of the IEEE International Conference on Computer Vision*, 2017.
42. Scherpereel, K., et al. A human lower-limb biomechanics and wearable sensors dataset during cyclic and non-cyclic activities. *Scientific Data*, 10(1), 924, 2023.
43. Justus, D., et al. Predicting the computational cost of deep learning models. *2018 IEEE International Conference on Big Data (Big Data)*, IEEE, 2018.
44. Bai, L., et al. Dnnabacus: Toward accurate computational cost prediction for deep neural networks. *arXiv preprint arXiv:2205.12095*, 2022.

45. Ameri, A., et al. A deep transfer learning approach to reducing the effect of electrode shift in EMG pattern recognition-based control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(2), 370–379, 2019.
46. Kang, P., et al. Wrist-worn hand gesture recognition while walking via transfer learning. *IEEE Journal of Biomedical and Health Informatics*, 26(3), 952–961, 2021.
47. Pan, S. J., and Yang, Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359, 2009.
48. Uhlrich, Scott D., et al. OpenCap: Human movement dynamics from smartphone videos. *PLoS computational biology*, 19.10, 2023.

## Acknowledgements

This work was supported by the Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (Ministry of Trade, Industry and Energy, MOTIE) in 2024 through the International Joint Technology Development Program for Industrial Technology International Cooperation (Project No. R2417081, Project Name: Development of a real-time integrated control SW platform for autonomous driving/collaborative robots using generative AI).

## Author contributions statement

All authors contributed to the conceptualization of the study. B.H. contributed to the development of data processing algorithms and deep learning models. B.H and S.B. conducted the experiments. B.H, S.B, I.K, and D.K. analyzed and interpreted the data. D.K. contributed to funding acquisition. All authors prepared and revised the manuscript.

## Competing interests

The authors declare no competing interests.

## Supplementary information

## Additional information