AddBiomechanics Dataset: Capturing the Physics of Human Motion at Scale

Keenon Werling[1], Janelle Kaneda[1], Tian Tan[1], Rishi Agarwal[1], Six Skov[1], Tom Van Wouwe[1], Scott Uhlrich[1], Nicholas Bianco[1], Carmichael Ong[1], Antoine Falisse[1], Shardul Sapkota[1], Aidan Chandra[1], Joshua Carter[2], Ezio Preatoni[2], Benjamin Fregly[3], Jennifer Hicks[1], Scott Delp[1], C. Karen Liu[1]

[1]Stanford University
{keenon,jkaneda,alantan,rishia,osskov,tvwouwe,suhlrich,nbianco,ongcf,afalisse,sapkota,aidan13,jenhicks,delp,karenliu}@stanford.edu, [2]University of Bath
{jc2369,ep314}@bath.ac.uk, [3]Rice University fregly@rice.edu

Abstract. While reconstructing human poses in 3D from inexpensive sensors has advanced significantly in recent years, quantifying the dynamics of human motion, including the muscle-generated joint torques and external forces, remains a challenge. Prior attempts to estimate physics from reconstructed human poses have been hampered by a lack of datasets with high-quality pose and force data for a variety of movements. We present the *AddBiomechanics Dataset 1.0*, which includes physically accurate human dynamics of 273 human subjects, over 70 hours of motion and force plate data, totaling more than 24 million frames. To construct this dataset, novel analytical methods were required, which are also reported here. We propose a benchmark for estimating human dynamics from motion using this dataset, and present several baseline results. The AddBiomechanics Dataset is publicly available at addbiomechanics.org/download_data.html.

Keywords: dataset, human body motion, human body physics, real to sim, benchmark

1 Introduction

The ability to accurately infer human movement *dynamics* (the motion *and* the physical forces that generated the motion) from video cameras or wearable sensors would transform the fields of computer graphics, robotics, orthopedics, biomechanics, and rehabilitation. With measurement not only of movement but also the internal and external forces driving the movement, researchers can generate more realistic animations of movement [35], design wearable robotic systems [34, 47], study injury mechanisms [3, 28], and improve treatments for mobility impairments [24, 31].

To reconstruct human movement dynamics from inexpensive sensors, we need realistic musculoskeletal models of the human body, along with accurate estimates of human pose and external moments and forces. Over the past decade,
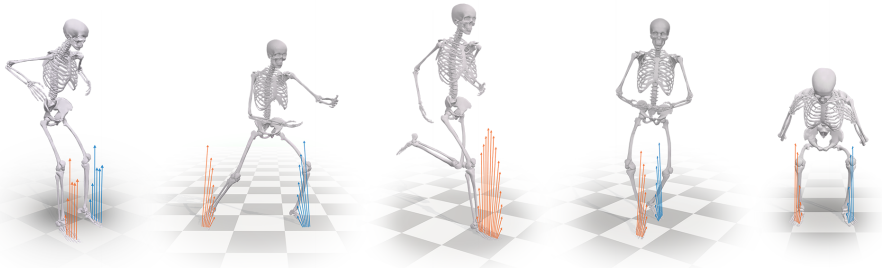
Fig. 1: Frames from recordings of a drop jump, tai chi, running, ballet, and squatting (left to right) in the AddBiomechanics Dataset. The ground reaction forces are shown with arrows on the right (orange) and left (blue) feet for representative time steps around the shown frame. The model includes 22 body segments, 37 degrees of freedom, and specialized joints for the knees, shoulders, and spine, which match the subject's mass, inertial properties, and motion capture recordings while obeying $F = ma$ with respect to the force plate recordings. Discrepancies between the reconstruction and the original sensor data are within clinically acceptable tolerances [21].

the research community has made significant progress in both generating realistic musculoskeletal models of the body [23, 56, 63, 68] and reconstructing poses from inexpensive sensors such as cameras [8, 16, 22, 32, 38, 42, 60, 66], sparse wearable sensors [26], and generative models [57]. However, the measurement of external forces outside of motion capture labs is unreliable [10, 33] and thus far the models that estimate external forces from motion alone are intractable for real time use [13, 67] or are not evaluated against measured data [17, 26, 53, 61, 62, 69, 74]. Thus their accuracy cannot be compared to the thresholds established in the literature [21]. A key bottleneck is the lack of combined motion and force datasets to train models and assess accuracy of external force estimates against real measurements for a diversity of human subjects and activities.

To address this problem, we present the AddBiomechanics Dataset, the first large scale standardized dataset of lab-based pose and force data for a variety of movements, currently with over 273 participants and 70+ hours of recorded motion. This dataset can be used to train machine learning methods to reconstruct detailed force information about human movement dynamics (*e.g.*, biological torques, ground reaction moments and forces, etc.) from easily measurable quantities (*e.g.*, motion capture from a cell phone camera). To provide a fair basis to compare future models that estimate joint torques and external moments and forces acting on the body from motion alone, we propose a set of evaluation metrics, along with a standard train/development/test split. To provide a sense for what is possible with this data, we implement both a simple model and the architecture from previous papers that have attempted data-driven ground reaction moment (GRM) and force (GRF) prediction models [19, 50], and report state-of-the-art results on the AddBiomechanics Dataset.

To construct our dataset, we present a key improvement to the AddBiomechanics tool presented by Werling *et al.* [71], which quickly and automatically provides estimates of model scales, motion, and dynamics that meet rigorous standards for analyzing human motion from the biomechanics literature [21]. In particular, the original AddBiomechanics tool is unable to process movement trials with a mix of measured and unmeasured external forces, and compute dynamics only when the external forces are available. These trial types are common since, to obtain quantities of diverse raw data recordings at a scale useful for machine learning, researchers must collect during long sessions in labs with motion capture and in-ground force plates, but instrumenting the entire floor of a lab is prohibitively expensive. Long captures of diverse data therefore contain frames where the subject stepped off of the force plates and experienced large unmeasured external ground forces acting on their body. While we still cannot use those frames in the final dataset, we present a method to use the optical motion capture on those "unobserved forces" frames to constrain the dynamics reconstruction on frames where all external forces acting on the body *are* measured, and allowing long diverse capture sessions to be optimized together. This improvement allows us to reconstruct accurate dynamics on large, diverse and previously inaccessible raw data sources, making the resulting dataset larger and more diverse.

In summary, we make the following novel contributions:

- **Releasing the largest dataset of human dynamics**, containing personalized musculoskeletal models and corresponding experimental measurements of foot-ground contact forces paired with motion capture data.
- **Improving automated human dynamics data processing tools** to enable the construction of a large and diverse dataset, by overcoming a limitation that required that all steps be on force plates. We discard timesteps with steps off of force plates (often present in long continuous sessions), but can still find dynamics on those frames where force plate data is present.
- **Proposing a standard benchmark** to evaluate methods that reconstruct human dynamics information from human motion alone. For example, some work [19,50] evaluates error only for the vertical ground reaction force, which is a strict lower bound on taking the L2-norm between predicted and measured 3D force vectors.

2   Related Work

2.1   Related Datasets

While there are several valuable, existing datasets including both pose and force-related data, they are currently small-in-scale, of insufficient accuracy, and/or in a format unsuitable for the task of predicting forces (*e.g.*, long captures with only partially observed experimental force data). For example, one of the largest available datasets of human dynamics data with pre-computed full body models is presented in UnderPressure [50], with 5.6 hours of data from 10 subjects.

UnderPressure contains joint kinematics, inertial measurement unit (IMU) data and synchronized pressure insole data collected during various types of locomotion. While the largest available, pressure insole data is of insufficient accuracy to create or evaluate 3D dynamic simulations [7, 10, 33].

There are several additional datasets, of varying size, with higher-quality force data from expensive in-floor or in-treadmill force plates, but these datasets have heterogeneous processing, incomplete observation data or formatting inconsistencies. Carter *et al.* [9] and Camargo *et al.* [6] each share nearly 20 hours of optical marker motion capture and force data for subjects performing walking, stairs, ramps, running, jumping, and other activities. Most datasets collected with in-ground force plates (*e.g.*, Camargo *et al.*) have large sections of partial force data as a subject steps onto or off of an in-ground plate, which existing analysis tools cannot handle. Additional human dynamics datasets (most under 2 hours) include video keypoints with synchronized ground force plate data [49], sparse IMU sets with instrumented treadmills [1, 41, 59], and traditional optical markers with ground force plates [27, 44, 51, 52], with GroundLink [19] representing the widest diversity of dynamics with 19 different motions. We combine these and several additional datasets into the AddBiomechanics Dataset 1.0.

2.2   Estimating Physical Forces from Motion

In computer vision and graphics, physics (either in the form of a physics engine or constraints to represent "physical intuitions") has been applied by many researchers to improve the quality of generated movements, to combat common artefacts such as foot floating or foot sliding.

One common approach is the use of fast physics-based reinforcement learning methods [2, 39, 43, 53–55, 61, 62, 73, 75–77], generally with simplified contact models (*e.g.*, "rigid box feet") to estimate joint torques necessary to achieve a measured motion. Another area of focus has been improving the quality of human motion simulations (whether synthesized *de novo* or obtained from a motion capture method) by leveraging "physical intuitions" about how humans move in the world (*e.g.*, minimizing foot sliding, disallowing foot-ground penetration, etc.). Approaches to this task have specifically targeted the problem of visual motion quality, whether in isolation [17, 26, 57, 58] or tied to a monocular motion capture pipeline [4, 5, 40, 72, 74, 79]. Both the physics based and physical intuition based methods use "physical plausibility" to describe results and evaluate against motion data alone. Quantifying the accuracy of estimated external forces and joint torques requires measured ground reaction force data with synchronized motion that isn't currently available. The ability of existing methods to reconstruct experimentally valid human dynamics remains an open question.

Combined motion and force datasets have also been used on a small scale to evaluate the models that estimate moments and forces from motion. One approach uses measured motion and force data to reconstruct dynamics with offline trajectory optimization methods and biomechanically accurate models (*i.e.*, using muscle actuators) [11, 15, 20, 29, 67]. Optimization is computationally expensive (*e.g.*, 0.001x real time [29, 67]), can only run on a few seconds of data

at a time, has errors on the order of 10% of body weight compared to measured ground reaction forces [29, 67] and depends on a task-specific movement objective (*e.g.*, minimize energy, maximize speed). A second approach leverages motion and force data to train deep learning models to predict ground reaction forces for specific tasks, such as running [1, 59], walking [41, 51, 52], sidestepping [27, 49], squatting [45], and stair climbing [44]. Although comparing the accuracy is challenging due to a lack of standard evaluation metrics, these models perform well for task-specific prediction (*e.g.*, 0.16 BW RMSE [1]; $r = 0.93$ for vertical GRF [51]). There have been few data-driven models trained on multiple activities and evaluated on out-of-distribution movements [19, 45] and their accompanying datasets are still limited (*e.g.*, less than eight subjects) with average errors exceeding 25% of body weight.

In summary, it is currently possible to predict with 10% error foot-floor contact forces for common activities like walking using offline optimization-based methods [29], but no fast, accurate, and generalized models exist.

3   AddBiomechanics Dataset

Table 1: Breakdown of subject count and hours from the raw data sources.

Source	Subjects	Total Hours	GRF Hours
Lencioni *et al.* 2019 [36]	50	0.52	0.37
Carter *et al.* 2023 [9]	50	21.60	19.80
Santos *et al.* 2017 [12]	49	4.90	4.90
Camargo *et al.* 2021 [6]	22	19.87	10.10
Tan *et al.* 2023 [65]	17	4.40	4.40
Moore *et al.* 2015 [48]	12	6.22	6.03
Falisse *et al.* 2016 [13]	11	0.49	0.10
Hamner *et al.* 2013 [18]	10	0.02	0.02
Van der Zee *et al.* 2022 [78]	10	5.46	5.31
Uhlrich *et al.* 2023 [67]	10	0.24	0.05
Tan *et al.* 2022 [64]	9	3.73	3.73
Wang *et al.* 2023 [70]	9	1.84	1.84
Han *et al.* 2023 [19]	7	1.70	0.52
Fregly *et al.* 2012 [14]	6	0.14	0.04
Li *et al.* 2021 [37]	1	0.34	0.34
Sum	**273**	**71.47**	**57.55**

Version 1.0 of the AddBiomechanics Dataset contains standardized musculoskeletal models as well as position and physics information for over 24 million frames from 70+ hours of motion for 273 participants (Table 1). Each of these frames contains optical marker locations and ground reaction moment and force measurements, along with estimated joint kinematics, estimated joint torques,

and estimated center of mass kinematics. The dataset includes nine different activity types collected from 15 publicly available raw data sources (Table 1), captured in 12 different laboratories. Each raw data source includes experimental optical motion capture and force plate data; refer to the corresponding publications for more details about the experimental data collection. All raw data was run through the same processing pipeline, presented in Section 4. For the data sources that reported demographic information, the subjects have a mean age of 30.7 ($\pm$15.8) years (range: $6 - 84$), and mean BMI of 22.8 ($\pm$3.4) (range: $11.7 - 34.4$). The majority of the subjects are male (73%), and 23% are female. A portion of the datasets did not contain reported age (9%) or biological sex (4%).
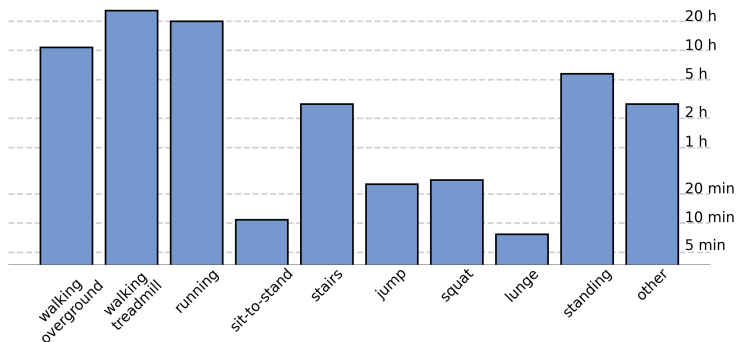
Fig. 2: Activity classification. The duration of captures in each activity class is shown on a log scale.

We manually classified the activity types in the AddBiomechanics Dataset by visualizing the kinematics of each capture. Labels were generated from the activities presented by each dataset. Sub-labels further classify each activity class; for example, there are sub-labels for overground vs. treadmill walking and running, and for walking up vs. down stairs. Any capture that could not be clearly categorized under the main activity labels is classified as "other." For example, several captures from [48] are subjects stepping onto a treadmill, and some captures contain calibration motions. Figure 2 shows the duration of each activity class within the AddBiomechanics Dataset.

The AddBiomechanics Dataset contains a variety of speeds and individual trial lengths (Figure 3). Averaged over each trial, the majority of absolute speeds fall between about 0.0 and 5.0 m/s. Trials were split into segments of at most 2,000 frames as a result of computational limits in the processing pipeline. The majority of remaining trials contain less than 1,000 frames. Overall, 53.4% of the GRF in the dataset is single support, followed by 21.2% for double support and 25.4% for flight phase.
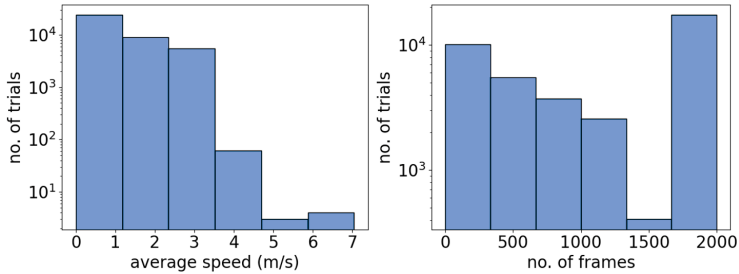
Fig. 3: Distributions of speed and capture length in the AddBiomechanics Dataset. The magnitude of speed averaged per trial (**left**). Trial lengths as number of frames (**right**). Original trials longer than 2000 frames were split up into multiple trials segments.

3.1 Personalized Subject Models

We provide individual subject skeletons as scaled and mass-optimized versions of the Rajagopal Full Body Model [56], which is the standard in human biomechanics for maximum biomechanical accuracy in modelling the physics of the body. Recent work by Keller *et al.* [30] means the Rajagopal model can also be transformed to a variant of the SMPL Model, ensuring that the AddBiomechanics Dataset can be used with existing SMPL-based datasets (*e.g.*, [25, 46]) and methods (*e.g.*, [16, 17, 26, 53, 57, 61, 62, 69, 74]).

3.2 Evaluating Dataset Quality

It is best practice in biomechanics to evaluate dynamic reconstructions of human movement from raw sensor data (marker locations, force plate data) using the discrepancy between the raw measurements and what the dynamic model implies the measurements should be [21]. These standard measurements are available in Table 2. There are thresholds for clinical grade dynamics suggested by Hicks *et al.* [21], also reported in the table. Many human modelling experts do not reach this bar in practice [71], so passing these thresholds at scale is a significant achievement. 12.1 hours (21.2%) of the data is classified as clinical-grade human dynamics data by this measure, an unprecedentedly large dataset to reach that quality bar. For overall distributions of residuals, see Figure 4.

4 Raw Data Processing and Aggregation

Raw data sources contain optical marker traces, and time synced force plate data. It requires a substantial optimization problem to reconstruct human dynamics. The backbone of our approach is to manually review our raw data sources frame-by-frame to determine frames where steps off of force plates occurred, and then run the annotated raw data through the pipeline presented in Werling *et al.* [71], with the modifications described below to handle steps off of force plates.

Measurement	Mean ± Std-dev	Hicks Threshold for Clinical Use [21]
Marker Error	2.17 ± 2.44 (cm)	N/A
Linear Residual	0.046 ± 0.027 (BW)	0.05 BW
Angular Residual	0.11 ± 0.02 (BW*h)	0.1 BW*h

Table 2: Evaluating dataset quality. "Marker Error" is measured in centimeters RMS to evaluate the quality of pose reconstruction. "Linear Residual" is force discrepancy between model and raw data, normalized across subjects as body-weights (BW) RMS. "Angular Residual" is torque dynamic discrepancy at the model's root, normalized across subjects as body-weights * height (BW*h) RMS. The Hicks Thresholds [21] for clinical model accuracy are presented with each metric.
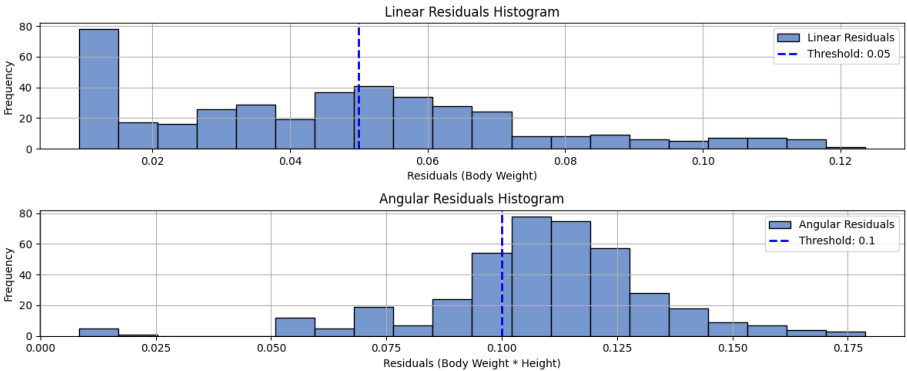
Fig. 4: Evaluating dataset quality. See Table 2 for descriptions of quantities.

In collecting large amounts of motion data with overground force plates, it is common for subjects to step on uninstrumented ground and have unmeasured forces acting on them several times during a single capture. These frames of data are not useful for learning about human dynamics, so should be thrown away, but that must be done after we have solved for dynamics on the remaining frames. If we were to simply delete the frames, and solve for skeleton scaling and motion dynamics separately on each range of frames with continuously observed forces, we could get quite unrealistic motion, since there are no useful constraints on the start and end conditions of each short group of frames with observed forces, and so the optimizer can cheat with unrealistic initial and terminal velocities. Optimizing for dynamics with the motion from frames with unobserved force data helps prevent that from happening.

In order to construct this dataset, we needed to be able to solve for accurate human dynamics while ignoring physical consistency on frames with steps off of force plates, and yet not allowing any discontinuous jumps in human kinematics when they step off of or back onto force plates. This section describes the relevant background, and our algorithms to handle this problem.
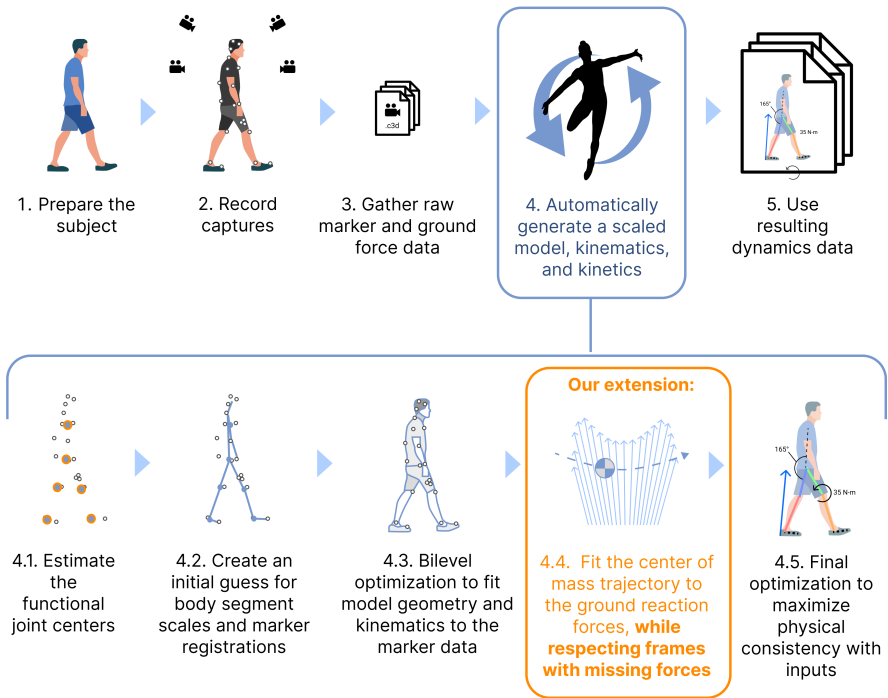
Fig. 5: Adapted with permission from Werling *et al.* [71], this figure shows the processing pipeline we used to turn raw sensor data into accurate human dynamics, and shows where the method described below swaps into the original pipeline from [71]

These methods are enhancements to the methods originally presented in Werling *et al.* [71] These algorithms are already merged into the upstream Werling *et al.* [71] tool, and are available in the open source codebase.

4.1 Processing Raw Sensor Data

Biomechanical motion capture labs output optical marker traces over time, and time-synced ground reaction force measurements (which include force, locations, and torque for each contact) from force plates in the ground or in treadmills. It is possible to run inverse kinematics on the locations of the optical markers over time to produce an accurate reconstruction of motion by picking joint angles q and bone scales s to minimize the distance between virtual markers $f(q, s)$ and observed markers o.

$$\min_{q} ||f(q, s) - o|| \tag{1}$$

It is also possible, given a skeleton model, joint accelerations and observations of external forces f_t from the force plates, to run standard inverse dynamics to

compute joint torques.

$$\boldsymbol{\tau}_t = \boldsymbol{M}(\ddot{\boldsymbol{q}}_t) - \boldsymbol{C}_t - \boldsymbol{J}^T \boldsymbol{f}_t \tag{2}$$

We adopt standard notation that $\boldsymbol{M}$ is the mass matrix, $\boldsymbol{C}_t$ is the vector of joint torques from coriolis forces and gravity, $\boldsymbol{J}$ is the jacobian of the contacts, and $\boldsymbol{f}_t$ is the contact forces.

Here $\ddot{\boldsymbol{q}}_t$ is estimated by central differencing the joint angles, with the equation given below.

$$\ddot{\tilde{\boldsymbol{q}}}_t = \frac{\tilde{\boldsymbol{q}}_{t-1} - 2\tilde{\boldsymbol{q}}_t + \tilde{\boldsymbol{q}}_{t+1}}{\Delta t^2} \tag{3}$$

It is worth noting that noise in motion capture systems that estimate our $\boldsymbol{q}_t$ values tends to have very high frequency content, because each frame's errors are approximately independently and identically distributed at 100fps. This leads to a signal to noise ratio that is 4 orders of magnitude worse after finite differencing. We use a Butterworth low pass filter at 30Hz to process the finite differenced estimates, which does a good job attenuating the noise in our estimates of $\ddot{\boldsymbol{q}}_t$ but makes them only trustworthy in their lower frequency components.

Returning to Equation 2, the problem now arises that in order to be physically possible, whichever degrees of freedom correspond to the free translation and rotation of the skeleton in the world (generally the first 6, 3 for translation and 3 for rotation) in $\boldsymbol{\tau}$ must be 0. This constraint (*i.e.,* $\boldsymbol{\tau}^{[0:6]} = 0$) reflects the fact that all external forces are applied by the force plate, and not by a giant invisible robot arm at the root segment. However, there is no guarantee in the standard inverse dynamics equations that the first 6 entries of $\boldsymbol{\tau}_t$ will be 0. To the extent they are not zero, they have typically been called "residual forces" in the literature.

The goal in producing realistic digital twins of subjects is to find subject bone scales $\boldsymbol{s}$, positions over time $\boldsymbol{q}_t$, and bone masses and inertial properties $\boldsymbol{m}$ such that the residual forces are minimized when we solve inverse dynamics. Leaving out some less important variables for clarity, Werling *et al.* [71] optimize the following non-convex problem to solve for $\boldsymbol{s}$, $\boldsymbol{q}_t$, and $\boldsymbol{m}$:

$$\min_{\boldsymbol{q}_t, \boldsymbol{s}, \boldsymbol{m}} \sum_t \underbrace{||f(\boldsymbol{q}_t, \boldsymbol{s}) - \boldsymbol{o}_t||}_{\text{Inverse Kinematics}} + \underbrace{||\boldsymbol{\tau}_t^{[0:6]}||}_{\text{Inverse Dynamics}} \tag{4}$$

The challenge with this approach is that it requires a good initial guess for the values of $\boldsymbol{q}_t$, $\boldsymbol{s}$, $\boldsymbol{m}$ because the optimization problem is non-convex.

In the original method, these initial guesses for $\boldsymbol{q}_t$ and $\boldsymbol{s}$ are constructed through a series of linear and convex optimization problems with marker data. The $\boldsymbol{q}_t$ initialization is then adjusted so the center of mass trajectory implied by $\boldsymbol{q}_t$ is consistent with the trajectory dictated by the ground reaction forces, defined by the differential equation:

$$\ddot{\boldsymbol{z}} = \frac{\boldsymbol{f}}{m} - \boldsymbol{g} \tag{5}$$

where $\ddot{z} \in \mathbb{R}^3$ is center of mass acceleration, $\boldsymbol{f}$ is the ground reaction force vector, m is the system mass, and $\boldsymbol{g}$ is gravitational acceleration.

4.2 Fitting Continuous Dynamics While Ignoring Frames with Missing Force Data

When all the ground reaction forces are known from experimental data, then the entire trajectory $\boldsymbol{z}_t$ is linearly determined by the inverse of mass μ, $\boldsymbol{z}_1$ and $\dot{\boldsymbol{z}}_1$. However, given a set of frames where the external forces are no longer observed, $\boldsymbol{U} = \{\boldsymbol{u}_1 \in \mathbb{N}, \boldsymbol{u}_2 \in \mathbb{N}, \ldots\}$, then this simple linear relationship ceases to hold. The center of mass trajectory $\boldsymbol{z}_t$ is now also dependent on the total external forces on those frames, divided by the unknown subject mass, then integrated over time, which makes the relationship between μ, $\boldsymbol{z}_1$, $\dot{\boldsymbol{z}}_1$, and all $\boldsymbol{f}_{\boldsymbol{u}_i}$ for $\boldsymbol{u}_i \in \boldsymbol{U}$ quadratic.

To restore linearity in the presence of unknown forces, we improve the original Werling *et al.* method by explicitly solving for the center of mass acceleration on each unobserved frame: $\ddot{z}_{\boldsymbol{u}_i}$ for all $\boldsymbol{u}_i \in \boldsymbol{U}$. We define an extended vector $\boldsymbol{\zeta}$ that contains our unknowns:

$$\boldsymbol{\zeta} = \begin{bmatrix} \boldsymbol{z}_1 & \dot{\boldsymbol{z}}_1 & \mu & \ddot{z}_{\boldsymbol{u}_1} & \ddot{z}_{\boldsymbol{u}_2} & \ldots \end{bmatrix} \in \mathbb{R}^{7+3|U|} \tag{6}$$

We define a linear system with matrix $\boldsymbol{A} \in \mathbb{R}^{3(T+|U|) \times 7+3|U|}$ and offset $\boldsymbol{b} \in \mathbb{R}^{3(T+|U|)}$ that maps the vector $\boldsymbol{\zeta}$ onto $\boldsymbol{Z} \in \mathbb{R}^{3T}$, a vector of concatenated center of mass position vectors over time, and $\ddot{\boldsymbol{Z}}_U$ the concatenated $\ddot{z}_{\boldsymbol{u}_i}$'s, scaled by a regularization term α:

$$\boldsymbol{A}\boldsymbol{\zeta} + \boldsymbol{b} = \begin{bmatrix} \boldsymbol{Z} & \alpha\ddot{\boldsymbol{Z}}_U \end{bmatrix} = \begin{bmatrix} \boldsymbol{z}_1 & \boldsymbol{z}_2 & \ldots & \boldsymbol{z}_T & \alpha\ddot{z}_{\boldsymbol{u}_1} & \ldots & \alpha\ddot{z}_{\boldsymbol{u}_{|U|}} \end{bmatrix} \in \mathbb{R}^{3(T+|U|)} \tag{7}$$

We can construct $\boldsymbol{A}$ and $\boldsymbol{b}$ using a semi-explicit Euler integration scheme of the center of mass trajectory to relate the unknowns $\boldsymbol{\zeta}$ to the center of mass positions, $\boldsymbol{Z}$. We proceed in blocks:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{\boldsymbol{Z},3} & \boldsymbol{A}_{\boldsymbol{Z},\ddot{\boldsymbol{Z}}_U} \\ 0 & \alpha I \end{bmatrix} \tag{8}$$

Here, $\boldsymbol{A}_{\boldsymbol{Z},3}$ represents the contributions from μ, $\boldsymbol{z}_1$ and $\dot{\boldsymbol{z}}_1$ to the trajectory, similar to the original adjustment in Werling *et al.* [71]. Then, $\boldsymbol{A}_{\boldsymbol{Z},\ddot{\boldsymbol{Z}}_U}$ is the matrix block relating our acceleration on time steps where there are unobserved external forces is the simple time integration of the change in velocity generated by our $\ddot{z}_{\boldsymbol{u}_i}$ for each unobserved time step i, given as follows:

$$\boldsymbol{A}_{\boldsymbol{Z},\ddot{\boldsymbol{Z}}_U}^{[:,i]} = \begin{bmatrix} 0 & \ldots & 0 & \Delta_t^2 I & 2\Delta_t^2 I & \ldots & (T - \boldsymbol{u}_i)\Delta_t^2 I \end{bmatrix} \tag{9}$$

The vector $\boldsymbol{b}$ is as expressed in the original method, now concatenated with a 0 vector in $\mathcal{R}^{3|U|}$, so that $\boldsymbol{b} \in \mathcal{R}^{3T+3|U|}$.

Using Equation 7, we apply the pseudoinverse of $\boldsymbol{A}$ to obtain a least-squares best fit of the initial conditions and mass of the system, $\hat{\boldsymbol{\zeta}} = \boldsymbol{A}^{\dagger}(\boldsymbol{\mathcal{Z}} - \boldsymbol{b})$, along with all the α regularized center of mass accelerations on time steps with unobserved external forces $\boldsymbol{u}_i \in \boldsymbol{U}$ and can continue with the original adjustment process [71]. This adjusted initialization then serves as an initialization for the final problem in Equation 4, which we adapt slightly as Equation 10 by only including the inverse dynamics loss on frames with observed forces.

$$\min_{\boldsymbol{q}_t, \boldsymbol{s}, \boldsymbol{m}} \sum_t \underbrace{||f(\boldsymbol{q}_t, \boldsymbol{s}) - \boldsymbol{o}_t||}_{\text{Inverse Kinematics}} + \underbrace{\mathbb{I}(t \notin \boldsymbol{U})||\boldsymbol{\tau}_t^{[0:6]}||}_{\text{Observed Inverse Dynamics}} \tag{10}$$

This new algorithm vastly expands the available data with which to construct the dataset by enabling the platform released by Werling $et\ al.$ [71] to process motion data with observed external forces on only some time steps.

5   Benchmark for Estimating Dynamics from Motion

The key evaluation task of interest enabled by the AddBiomechanics Dataset is estimating physical forces from an observed movement. In this section, we describe in more detail the evaluation task, provide standard metrics relevant to Computer Science and Biomechanics to assess the accuracy of new models, and provide some initial baseline models and their performance on the evaluation metrics.

5.1   Task Description

The task assumes an "off the shelf" motion capture model ($e.g.$, [8, 16, 26, 66]) produces a noisy time series of joint angles over time, $\tilde{\boldsymbol{q}}_t$ for $t \in [0, T]$. We also assume that a scaled human body model is available for the subject – modern motion capture algorithms provide this, $e.g.$ [16]. In processing the AddBiomechanics Dataset and in the baselines discussed below, we use the model of Rajagopal $et\ al.$ [56]; future models to solve the task could employ this or other skeletal models.

To construct a full physical simulation of our subject, we now need to solve inverse dynamics (Equation 2) using estimated joint accelerations $\ddot{\boldsymbol{q}}_t$, and either contact forces and moments acting at the origin of the foot $\boldsymbol{f}_t$ or joint torques $\boldsymbol{\tau}_t$. The simplest way to achieve this is using central differencing (Equation 3) and filtering on our input pose estimates $\tilde{\boldsymbol{q}}_t$ to estimate $\ddot{\boldsymbol{q}}_t$ and using a model to estimate contact forces and moments $\boldsymbol{f}_t$, then solving Equation 2 for compute joint torques $\boldsymbol{\tau}_t$. However, this is only one of many possible approaches to estimating these quantities.

To summarize, the inputs are $\tilde{\boldsymbol{q}}_t$ for $t \in [0, T]$ (a time series of joint angles) and a body model for the subject with correctly scaled bones. And outputs are all three of the following which are consistent with Equation 2:

1. $\ddot{\boldsymbol{q}}_t$ for $t \in [0, T]$: A time series of joint accelerations.
2. $\boldsymbol{f}_t[: 3]$ for $t \in [0, T]$: A time series of external moments acting at the origin of the foot.
3. $\boldsymbol{f}_t[3 :]$ for $t \in [0, T]$: A time series of external forces.
4. $\boldsymbol{\tau}_t$ for $t \in [0, T]$: A time series of joint torques.

5.2 Proposed Metrics

Previous work has used a diversity metrics, measuring different aspects of the problem (*e.g.*, r-value for vertical GRF [13], RMSE for vertical GRF [19,50]). We propose a superset of previous metrics. We also include ground reaction moment, which is often neglected in the literature despite contributing significantly to joint torque estimates.

Two clear metrics to evaluate the accuracy of models are the average L2-norm error for our ground reaction force and moment estimate $\hat{\boldsymbol{f}}_t$ and average element-wise absolute error for the joint torque estimate $\hat{\boldsymbol{\tau}}_t$. Because estimates for $\ddot{\boldsymbol{q}}_t$ are not experimentally observed, it does not make sense to include RMSE on $\ddot{\boldsymbol{q}}_t$ as a metric. However, we can partially evaluate the quality of $\ddot{\boldsymbol{q}}_t$ by checking the RMSE of $\hat{\ddot{\boldsymbol{z}}}_t$, the center of mass acceleration estimated by $\ddot{\boldsymbol{q}}_t$.

To summarize, we propose four evaluation metrics, corresponding to the four outputs of the models, where ground truth values from the dataset are denoted $\ddot{\boldsymbol{z}}_t$, $\boldsymbol{f}_t$, and $\boldsymbol{\tau}_t$ (see 3), and T is the number of time steps and N is the number of degrees of freedom in the skeleton:

1. Center of Mass Acceleration: $\sum_t ||\ddot{\boldsymbol{z}}_t - \hat{\ddot{\boldsymbol{z}}}_t||_2 / T$
2. Ground Reaction Moment (GRM at foot): $\sum_t ||\boldsymbol{f}_t[: 3] - \hat{\boldsymbol{f}}_t[: 3]||_2 / T$
3. Ground Reaction Force (GRF): $\sum_t ||\boldsymbol{f}_t[3 :] - \hat{\boldsymbol{f}}_t[3 :]||_2 / T$
4. Joint Torques: $\sum_t ||\boldsymbol{\tau}_t - \hat{\boldsymbol{\tau}}_t||_1 / NT$

5.3 Preliminary Baseline Experiments

To set a baseline for future work to improve upon, we tried several trivial or previously reported models on the data. All of these models predict $\boldsymbol{f}$ given $\tilde{\boldsymbol{q}}$, and then use Equation 2 to compute $\boldsymbol{\tau}_t$. We evaluate these baselines on the held-out test set (Table 3). Brief details of each baseline method are reported below.

Analytical Baseline: To evaluate performance without machine learning or optimization, we construct a simple baseline method. We estimate $\ddot{\boldsymbol{q}}_t$ using Equation 3. Then we estimate center of mass acceleration $\ddot{\boldsymbol{z}}_t$ using $\ddot{\boldsymbol{q}}_t$. The total external force acting on the body is estimated using $F = ma$. Then we divide that force up between the feet to create ground force $\boldsymbol{f}_t[3 :]$, splitting the force evenly during double support. We set ground moment $\boldsymbol{f}_t[: 3] = 0$.

Multi-Layer Perceptron Baseline: We train a simple two-layer MLP with a hidden state size of 512 to predict $\boldsymbol{f}_t$ from $\boldsymbol{q}_t$. We use as input features a window of recent $\boldsymbol{q}_t$, our estimates of $\ddot{\boldsymbol{q}}_t$ from Equation 3, and joint center locations

Method	CoM Acc.	Joint Torque	GRM (at foot)	GRF
Analytical	4.47 m/s^2	2.77 Nm/kg	6.15 Nm/kg	2.79 N/kg
MLP Baseline	1.41 m/s^2	**1.33 Nm/kg**	**0.03 Nm/kg**	1.29 N/kg
GroundLinkNet on AddB	**1.40 m/s^2**	1.34 Nm/kg	0.07 Nm/kg	**1.17 N/kg**
UnderPressure [50]	N/A	N/A	N/A	> 5.74 N/kg
GroundLinkNet [19]	N/A	N/A	N/A	> 3.06 N/kg
Trajectory Optimization [29]	N/A	N/A	N/A	∼ 1 N/kg

Table 3: Above the double-line is accuracy of simple baseline methods and the model from [19] retrained on the AddBiomechanics dataset, with errors presented as means over the test set. Below the double-line are errors self-reported by previous work. These are not perfectly comparable to our baseline evaluations (or each other), as they each evaluate on their own different test sets, and use their own metrics. Where precise conversion to N/kg errors is not possible (for example, authors only report error on the y-component, or only report r-value), approximate numbers or lower bounds are included to give the reader a general sense of approximate relative performance.

in the pelvis reference frame. We then use our prediction for f_t to compute $\ddot{z}_t = \sum f_t / m$, and adjust the root-translation coordinates of $\ddot{q}_t$ to reflect our new $\ddot{z}_t$.

Previous Work - GroundLinkNet: Two previous papers [19, 50] have tackled the problem of predicting f_t (though not $\ddot{q}_t$ or τ_t) using a similar convolution-based architecture, which we refer to as "GroundLinkNet" (where the two models differ we use [19]). For fair comparison, we perform the same procedure as our MLP baseline to predict all three of f_t, $\ddot{q}_t$, and τ_t from just f_t: the inputs, outputs, and post-processing equations are all identical, the only difference is the model architecture is more complex and larger than the MLP. The model is trained and tested on the same data.

6   Conclusion and Future Work

We present a large standardized dataset of high quality human dynamics. This dataset is currently biased towards walking and running, and we plan to release a follow-up dataset with a wider diversity of tasks.

We lay the groundwork for using this data to train models that reconstruct accurate human dynamics information from simple motion capture. We are excited to see the integration of continued progress in inexpensive motion capture systems with methods to infer physics from motion.

We also anticipate other exciting uses of the dataset. It could be possible to use estimates of foot-ground contact and forces derived from this data to remove artefacts from motion capture, like foot sliding. Another exciting application area for the data is "real to sim", developing more accurate human body simulators by learning better foot-ground contact colliders from data. This data could also be a useful additional input to motion models, and could have applications in physical motion auto-encoders.

References

1. Alcantara, R.S., Edwards, W.B., Millet, G.Y., Grabowski, A.M.: Predicting continuous ground reaction forces from accelerometers during uphill and downhill running: a recurrent neural network solution. PeerJ **10**, e12752 (2022)
2. Bergamin, K., Clavet, S., Holden, D., Forbes, J.R.: Drecon: data-driven responsive control of physics-based characters. ACM Transactions On Graphics (TOG) **38**(6), 1–11 (2019)
3. Boden, B.P., Sheehan, F.T., Torg, J.S., Hewett, T.E.: Non-contact acl injuries: mechanisms and risk factors. The Journal of the American Academy of Orthopaedic Surgeons **18**(9), 520 (2010)
4. Brubaker, M.A., Fleet, D.J., Hertzmann, A.: Physics-based person tracking using simplified lower-body dynamics. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–8. IEEE (2007)
5. Brubaker, M.A., Sigal, L., Fleet, D.J.: Estimating contact dynamics. In: 2009 IEEE 12th International Conference on Computer Vision. pp. 2389–2396. IEEE (2009)
6. Camargo, J., Ramanathan, A., Flanagan, W., Young, A.: 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)
7. Camomilla, V., Cereatti, A., Cutti, A.G., Fantozzi, S., Stagni, R., Vannozzi, G.: Methodological factors affecting joint moments estimation in clinical gait analysis: a systematic review. Biomedical engineering online **16**(1), 1–27 (2017)
8. Cao, Z., Simon, T., Wei, S.E., Sheikh, Y.: Realtime multi-person 2d pose estimation using part affinity fields. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7291–7299 (2017)
9. Carter, J., Chen, X., Cazzola, D., Trewartha, G., Preatoni, E.: Full body kinematics and ground reaction forces of fifty heterogeneous runners completing treadmill running at various speeds and gradients. University of Bath Research Data Archive (2023)
10. DeBerardinis, J., Dufek, J.S., Trabia, M.B., Lidstone, D.E.: Assessing the validity of pressure-measuring insoles in quantifying gait variables. Journal of rehabilitation and assistive technologies engineering **5**, 2055668317752088 (2018)
11. Dembia, C.L., Bianco, N.A., Falisse, A., Hicks, J.L., Delp, S.L.: Opensim moco: Musculoskeletal optimal control. PLOS Computational Biology **16**(12), e1008493 (2020)
12. Dos Santos, D.A., Fukuchi, C.A., Fukuchi, R.K., Duarte, M.: A data set with kinematic and ground reaction forces of human balance. PeerJ **5**, e3626 (2017)
13. Falisse, A., Van Rossom, S., Jonkers, I., De Groote, F.: Emg-driven optimal estimation of subject-specific hill model muscle–tendon parameters of the knee joint actuators. IEEE Transactions on Biomedical Engineering **64**(9), 2253–2262 (2016)
14. Fregly, B.J., Besier, T.F., Lloyd, D.G., Delp, S.L., Banks, S.A., Pandy, M.G., D'lima, D.D.: Grand challenge competition to predict in vivo knee loads. Journal of orthopaedic research **30**(4), 503–513 (2012)
15. Gärtner, E., Andriluka, M., Xu, H., Sminchisescu, C.: Trajectory optimization for physics-based reconstruction of 3d human pose from monocular video. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 13106–13115 (2022)
16. Goel, S., Pavlakos, G., Rajasegaran, J., Kanazawa, A., Malik, J.: Humans in 4d: Reconstructing and tracking humans with transformers. arXiv preprint arXiv:2305.20091 (2023)

17. Guo, H., Ji, Q.: Physics-augmented autoencoder for 3d skeleton-based gait recognition. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 19627–19638 (2023)

18. Hamner, S.R., Delp, S.L.: Muscle contributions to fore-aft and vertical body mass center accelerations over a range of running speeds. Journal of biomechanics **46**(4), 780–787 (2013)

19. Han, X., Senderling, B., To, S., Kumar, D., Whiting, E., Saito, J.: Groundlink: A dataset unifying human body movement and ground reaction dynamics. arXiv preprint arXiv:2310.03930 (2023)

20. HARAGUCHI, N., HASE, K.: Prediction of ground reaction forces and moments and joint kinematics and kinetics by inertial measurement units using 3d forward dynamics model. Journal of Biomechanical Science and Engineering pp. 23–00130 (2023)

21. Hicks, J.L., Uchida, T.K., Seth, A., Rajagopal, A., Delp, S.L.: Is my model good enough? best practices for verification and validation of musculoskeletal models and simulations of movement. Journal of biomechanical engineering **137**(2), 020905 (2015)

22. Holmquist, K., Wandt, B.: Diffpose: Multi-hypothesis human pose estimation using diffusion models. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 15977–15987 (2023)

23. Holzbaur, K.R., Murray, W.M., Delp, S.L.: A model of the upper extremity for simulating musculoskeletal surgery and analyzing neuromuscular control. Annals of biomedical engineering **33**, 829–840 (2005)

24. Hunt, M.A., Simic, M., Hinman, R.S., Bennell, K.L., Wrigley, T.V.: Feasibility of a gait retraining strategy for reducing knee joint loading: increased trunk lean guided by real-time biofeedback. Journal of biomechanics **44**(5), 943–947 (2011)

25. Ionescu, C., Papava, D., Olaru, V., Sminchisescu, C.: Human3. 6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. IEEE transactions on pattern analysis and machine intelligence **36**(7), 1325–1339 (2013)

26. Jiang, Y., Won, J., Ye, Y., Liu, C.K.: Drop: Dynamics responses from human motion prior and projective dynamics. arXiv preprint arXiv:2309.13742 (2023)

27. Johnson, W.R., Alderson, J., Lloyd, D., Mian, A.: Predicting athlete ground reaction forces and moments from spatio-temporal driven cnn models. IEEE Transactions on Biomedical Engineering **66**(3), 689–694 (2018)

28. Kalkhoven, J.T., Watsford, M.L., Coutts, A.J., Edwards, W.B., Impellizzeri, F.M.: Training load and injury: causal pathways and future directions. Sports Medicine **51**, 1137–1150 (2021)

29. Karatsidis, A., Jung, M., Schepers, H.M., Bellusci, G., de Zee, M., Veltink, P.H., Andersen, M.S.: Musculoskeletal model-based inverse dynamic analysis under ambulatory conditions using inertial motion capture. Medical engineering & physics **65**, 68–77 (2019)

30. Keller, M., Werling, K., Shin, S., Delp, S., Pujades, S., Liu, C.K., Black, M.J.: From skin to skeleton: Towards biomechanically accurate 3d digital humans. Transaction on Graphics **42** (Dec 2023)

31. Killen, B.A., Falisse, A., De Groote, F., Jonkers, I.: In silico-enhanced treatment and rehabilitation planning for patients with musculoskeletal disorders: Can musculoskeletal modelling and dynamic simulations really impact current clinical practice? Applied Sciences **10**(20), 7255 (2020)

32. Kocabas, M., Athanasiou, N., Black, M.J.: Vibe: Video inference for human body pose and shape estimation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 5253–5263 (2020)
33. Lambrich, J., Hagen, M., Schwiertz, G., Muehlbauer, T.: Concurrent validity and test–retest reliability of pressure-detecting insoles for static and dynamic movements in healthy young adults. Sensors **23**(10), 4913 (2023)
34. Lee, D., McLain, B.J., Kang, I., Young, A.J.: Biomechanical comparison of assistance strategies using a bilateral robotic knee exoskeleton. IEEE Transactions on Biomedical Engineering **68**(9), 2870–2879 (2021)
35. Lee, S., Park, M., Lee, K., Lee, J.: Scalable muscle-actuated human simulation and control. ACM Transactions On Graphics (TOG) **38**(4), 1–13 (2019)
36. Lencioni, T., Carpinella, I., Rabuffetti, M., Marzegan, A., Ferrarin, M.: Human kinematic, kinetic and emg data during different walking and stair ascending and descending tasks. Scientific data **6**(1), 309 (2019)
37. Li, G., Shourijeh, M.S., Ao, D., Patten, C., Fregly, B.J.: How well do commonly used co-contraction indices approximate lower limb joint stiffness trends during gait for individuals post-stroke? Frontiers in Bioengineering and Biotechnology **8**, 588908 (2021)
38. Li, W., Liu, H., Tang, H., Wang, P., Van Gool, L.: Mhformer: Multi-hypothesis transformer for 3d human pose estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 13147–13156 (2022)
39. Li, Z., Cheng, X., Peng, X.B., Abbeel, P., Levine, S., Berseth, G., Sreenath, K.: Reinforcement learning for robust parameterized locomotion control of bipedal robots. In: 2021 IEEE International Conference on Robotics and Automation (ICRA). pp. 2811–2817. IEEE (2021)
40. Li, Z., Sedlar, J., Carpentier, J., Laptev, I., Mansard, N., Sivic, J.: Estimating 3d motion and forces of person-object interactions from monocular video. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8640–8649 (2019)
41. Lim, H., Kim, B., Park, S.: Prediction of lower limb kinetics and kinematics during walking by a single imu on the lower back using machine learning. Sensors **20**(1), 130 (2019)
42. Lin, K., Wang, L., Liu, Z.: End-to-end human pose and mesh reconstruction with transformers. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 1954–1963 (2021)
43. Ling, H.Y., Zinno, F., Cheng, G., Van De Panne, M.: Character controllers using motion vaes. ACM Transactions on Graphics (TOG) **39**(4), 40–1 (2020)
44. Liu, D., He, M., Hou, M., Ma, Y.: Deep learning based ground reaction force estimation for stair walking using kinematic data. Measurement **198**, 111344 (2022)
45. Louis, N., Corso, J.J., Templin, T.N., Eliason, T.D., Nicolella, D.P.: Learning to estimate external forces of human motion in video. In: Proceedings of the 30th ACM International Conference on Multimedia. pp. 3540–3548 (2022)
46. Mahmood, N., Ghorbani, N., Troje, N.F., Pons-Moll, G., Black, M.J.: AMASS: Archive of motion capture as surface shapes. In: International Conference on Computer Vision. pp. 5442–5451 (Oct 2019)
47. Molinaro, D.D., Kang, I., Camargo, J., Young, A.J.: Biological hip torque estimation using a robotic hip exoskeleton. In: 2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob). pp. 791–796. IEEE (2020)
48. Moore, J.K., Hnat, S.K., van den Bogert, A.J.: An elaborate data set on human gait and the effect of mechanical perturbations. PeerJ **3**, e918 (2015)

49. Morris, C., Mundt, M., Goldacre, M., Weber, J., Mian, A., Alderson, J.: Predicting 3d ground reaction force from 2d video via neural networks in sidestepping tasks. ISBS Proceedings Archive **39**(1), 300 (2021)

50. Mourot, L., Hoyet, L., Clerc, F.L., Hellier, P.: Underpressure: Deep learning for foot contact detection, ground reaction force estimation and footskate cleanup. In: Computer Graphics Forum. vol. 41, pp. 195–206. Wiley Online Library (2022)

51. Mundt, M., Koeppe, A., David, S., Bamer, F., Potthast, W., Markert, B.: Prediction of ground reaction force and joint moments based on optical motion capture data during gait. Medical Engineering & Physics **86**, 29–34 (2020)

52. Oh, S.E., Choi, A., Mun, J.H.: Prediction of ground reaction forces during gait based on kinematics and a neural network model. Journal of biomechanics **46**(14), 2372–2380 (2013)

53. Peng, X.B., Abbeel, P., Levine, S., Van de Panne, M.: Deepmimic: Example-guided deep reinforcement learning of physics-based character skills. ACM Transactions On Graphics (TOG) **37**(4), 1–14 (2018)

54. Peng, X.B., Guo, Y., Halper, L., Levine, S., Fidler, S.: Ase: Large-scale reusable adversarial skill embeddings for physically simulated characters. ACM Transactions On Graphics (TOG) **41**(4), 1–17 (2022)

55. Peng, X.B., Ma, Z., Abbeel, P., Levine, S., Kanazawa, A.: Amp: Adversarial motion priors for stylized physics-based character control. ACM Transactions on Graphics (ToG) **40**(4), 1–20 (2021)

56. Rajagopal, A., Dembia, C.L., DeMers, M.S., Delp, D.D., Hicks, J.L., Delp, S.L.: Full-body musculoskeletal model for muscle-driven simulation of human gait. IEEE transactions on biomedical engineering **63**(10), 2068–2079 (2016)

57. Rempe, D., Birdal, T., Hertzmann, A., Yang, J., Sridhar, S., Guibas, L.J.: Humor: 3d human motion model for robust pose estimation. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 11488–11499 (2021)

58. Rempe, D., Guibas, L.J., Hertzmann, A., Russell, B., Villegas, R., Yang, J.: Contact and human dynamics from monocular video. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part V 16. pp. 71–87. Springer (2020)

59. Scheltinga, B.L., Kok, J.N., Buurke, J.H., Reenalda, J.: Estimating 3d ground reaction forces in running using three inertial measurement units. Frontiers in Sports and Active Living **5**, 1176466 (2023)

60. Shi, M., Starke, S., Ye, Y., Komura, T., Won, J.: Phasemp: Robust 3d pose estimation via phase-conditioned human motion prior. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 14725–14737 (2023)

61. Shimada, S., Golyanik, V., Xu, W., Pérez, P., Theobalt, C.: Neural monocular 3d human motion capture with physical awareness. ACM Transactions on Graphics (ToG) **40**(4), 1–15 (2021)

62. Shimada, S., Golyanik, V., Xu, W., Theobalt, C.: Physcap: Physically plausible monocular 3d motion capture in real time. ACM Transactions on Graphics (ToG) **39**(6), 1–16 (2020)

63. Silva, M.P., Ambrósio, J.A.: Sensitivity of the results produced by the inverse dynamic analysis of a human stride to perturbed input data. Gait & posture **19**(1), 35–49 (2004)

64. Tan, T., Strout, Z.A., Cheung, R.T., Shull, P.B.: Strike index estimation using a convolutional neural network with a single, shoe-mounted inertial sensor. Journal of Biomechanics **139**, 111145 (2022)

65. Tan, T., Wang, D., Shull, P.B., Halilaj, E.: Imu and smartphone camera fusion for knee adduction and knee flexion moment estimation during walking. IEEE Transactions on Industrial Informatics **19**(2), 1445–1455 (2023)

66. Toshev, A., Szegedy, C.: Deeppose: Human pose estimation via deep neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1653–1660 (2014)

67. Uhlrich, S.D., Falisse, A., Kidziński, Ł., Muccini, J., Ko, M., Chaudhari, A.S., Hicks, J.L., Delp, S.L.: Opencap: Human movement dynamics from smartphone videos. PLoS computational biology **19**(10), e1011462 (2023)

68. Vasavada, A.N., Li, S., Delp, S.L.: Influence of muscle morphometry and moment arms on the moment-generating capacity of human neck muscles. Spine **23**(4), 412–422 (1998)

69. Vondrak, M., Sigal, L., Jenkins, O.C.: Dynamical simulation priors for human motion tracking. IEEE transactions on pattern analysis and machine intelligence **35**(1), 52–65 (2012)

70. Wang, H., Basu, A., Durandau, G., Sartori, M.: A wearable real-time kinetic measurement sensor setup for human locomotion. Wearable Technologies **4**, e11 (2023). https://doi.org/10.1017/wtc.2023.7

71. Werling, K., Bianco, N.A., Raitor, M., Stingel, J., Hicks, J.L., Collins, S.H., Delp, S.L., Liu, C.K.: Addbiomechanics: Automating model scaling, inverse kinematics, and inverse dynamics from human motion data through sequential optimization. bioRxiv pp. 2023–06 (2023)

72. Xu, J., Yu, Z., Ni, B., Yang, J., Yang, X., Zhang, W.: Deep kinematics analysis for monocular 3d human pose estimation. In: Proceedings of the IEEE/CVF Conference on computer vision and Pattern recognition. pp. 899–908 (2020)

73. Yang, C., Yuan, K., Zhu, Q., Yu, W., Li, Z.: Multi-expert learning of adaptive legged locomotion. Science Robotics **5**(49), eabb2174 (2020)

74. Yang, G., Yang, S., Zhang, J.Z., Manchester, Z., Ramanan, D.: Ppr: Physically plausible reconstruction from monocular videos. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 3914–3924 (2023)

75. Yi, X., Zhou, Y., Habermann, M., Shimada, S., Golyanik, V., Theobalt, C., Xu, F.: Physical inertial poser (pip): Physics-aware real-time human motion tracking from sparse inertial sensors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 13167–13178 (2022)

76. Yuan, Y., Song, J., Iqbal, U., Vahdat, A., Kautz, J.: Physdiff: Physics-guided human motion diffusion model. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 16010–16021 (2023)

77. Yuan, Y., Wei, S.E., Simon, T., Kitani, K., Saragih, J.: Simpoe: Simulated character control for 3d human pose estimation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 7159–7169 (2021)

78. van der Zee, T.J., Mundinger, E.M., Kuo, A.D.: A biomechanics dataset of healthy human walking at various speeds, step lengths and step widths. Scientific data **9**(1), 704 (2022)

79. Zell, P., Wandt, B., Rosenhahn, B.: Joint 3d human motion capture and physical analysis from monocular videos. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 17–26 (2017)