

GroundLink: A Dataset Unifying Human Body Movement and Ground Reaction Dynamics

Xingjian Han
Computer Science
Boston University
Boston, USA
xjhan@bu.edu

Deepak Kumar
Physical Therapy
Boston University
Boston, USA
kumard@bu.edu

Benjamin Senderling
Physical Therapy
Boston University
Boston, USA
bsender@bu.edu

Emily Whiting
Computer Science
Boston University
Boston, USA
whiting@bu.edu

Stanley To
Computer Science
Boston University
Boston, USA
stanleyto275@gmail.com

Jun Saito
Adobe Research
Seattle, USA
jsaito@adobe.com

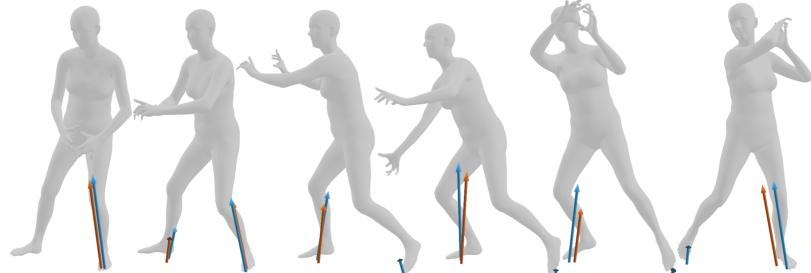
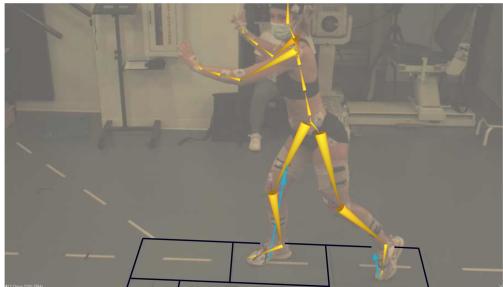


Figure 1: *GroundLink* is a dataset unifying standard kinematic motion capture with real-world physics information including ground reaction force (GRF) and center of pressure (CoP). (Left) The data is captured and synchronized at high temporal resolution in a biomechanics laboratory with mechanical force plates embedded in the ground and multiple high-speed cameras for uncompromising measurement accuracy. Our benchmark model *GroundLinkNet* trained with this data demonstrates its generalization capability by predicting GRFs and CoPs on unseen motions accurately and plausibly. (Right) Ground truth reaction forces (blue) compared to forces predicted by *GroundLinkNet* (orange) for a sample tennis motion.

ABSTRACT

The physical plausibility of human motions is vital to various applications in fields including but not limited to graphics, animation, robotics, vision, biomechanics, and sports science. While fully simulating human motions with physics is an extreme challenge, we hypothesize that we can treat this complexity as a black box in a data-driven manner if we focus on the ground contact, and have sufficient observations of physics and human activities in the real world. To prove our hypothesis, we present *GroundLink*, a unified dataset comprised of captured ground reaction force (GRF) and center of pressure (CoP) synchronized to standard kinematic motion

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SA Conference Papers '23, December 12–15, 2023, Sydney, NSW, Australia
© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0315-7/23/12...\$15.00
<https://doi.org/10.1145/3610548.3618247>

captures. GRF and CoP of *GroundLink* are not simulated but captured at high temporal resolution using force platforms embedded in the ground for uncompromising measurement accuracy. This dataset contains 368 processed motion trials ($\sim 1.59M$ recorded frames) with 19 different movements including locomotion and weight-shifting actions such as tennis swings to signify the importance of capturing physics paired with kinematics. *GroundLinkNet*, our benchmark neural network model trained with *GroundLink*, supports our hypothesis by predicting GRFs and CoPs accurately and plausibly on unseen motions from various sources. The dataset, code, and benchmark models are made public for further research on various downstream tasks leveraging the rich physics information at <https://csr.bu.edu/groundlink/>.

CCS CONCEPTS

- Computing methodologies → Motion capture.

KEYWORDS

motion capture datasets, neural network

ACM Reference Format:

Xingjian Han, Benjamin Senderling, Stanley To, Deepak Kumar, Emily Whiting, and Jun Saito. 2023. GroundLink: A Dataset Unifying Human Body Movement and Ground Reaction Dynamics. In *SIGGRAPH Asia 2023 Conference Papers (SA Conference Papers '23), December 12–15, 2023, Sydney, NSW, Australia*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3610548.3618247>

1 INTRODUCTION

Accurate and realistic representations of the human body and dynamics are important for graphics applications such as animation production and biomedical applications such as clinical movement analysis or medical device design. Understanding the underlying physics quantities and kinematics parameters, such as ground reaction force and moment, center of pressure, joint torque, and body trajectories, is critical for analyzing human movement dynamics.

In contrast to the growing demand for convincing human motions, most existing motion capture datasets focus on kinematics (e.g., [ACCAD 2001; CMU Graphics Lab 2000; Scott et al. 2020]) without physics quantities associated with them.

To address the need, we introduce *GroundLink*, a public dataset that consists of motion capture data and synchronized ground reaction forces (GRF) and center of pressure (CoP) captured from laboratory force plates. To fill the gap in the current demand for physically-plausible human movement the dataset includes subtle and slow movements such as yoga, physical therapy exercises, tennis swing motions, and walking. For example, Fig. 1 shows a noticeable weight transfer, indicating an intentional adjustment to maintain balance. This is the first public motion dataset in the graphics community that consists of laboratory-based ground reaction force measurements for various human motions.

To demonstrate the effectiveness of the kinematics data paired with physics quantities, we conduct various experiments with our benchmark deep neural network model *GroundLinkNet* trained with *GroundLink*. While its architecture is straightforward, we show that the model generalizes well in predicting GRFs and CoPs given the unseen kinematic motion sequences and body shapes. We believe this is a promising sign that a reasonable amount of kinematics data coupled with real-world physics quantities enable us to build generalizable human motion models with the understanding of physics for numerous applications in graphics, vision, and biomechanics in need of realistic digital human representation.

2 RELATED WORK

2.1 Motion Capture with Contact Dynamics

2.1.1 MoCap Dataset with Contact. Several publicly-available human motion datasets are used in the graphics community (e.g. [ACCAD 2001; CMU Graphics Lab 2000; Mahmood et al. 2019; Trumble et al. 2017b; Yin and Ying 2021]). However, motion capture datasets that contain ground truth contact information are scarce, especially datasets with continuous foot contact. Mourot et al. [2022a] and Scott et al. [2020] released valuable datasets of motion capture with synchronized pressure data that is recorded via pressure insoles and include deep neural networks to estimate pressure distributions from 3D joint positions. Yin and Pai [2003] also designed an intuitive interface for digital avatar animation with their collected foot

pressure sensor data. The datasets and tools serve specifically for animation usage and human stability analysis that contains locomotion, 24-form simplified Taichi movements and weight transferring tasks. Further, the dataset released from Luo et al. [2021] enables pose estimation from tactile signals given by pressure maps. However, pressure values captured with insoles or other types of pressure pads only contain contact information in magnitude (vertical GRF). Additionally, the computer vision community has invested effort in gathering video-based datasets focused on various body parts (e.g. hand pressure estimation from sensor data from Grady et al. [2022]) and scene interactions to address contact issues.

On the other hand, GRF and CoP are commonly understood concepts in biomechanics studies. Many motion datasets and experimental data that include kinematics, either full-body or focused on lower-extremity, along with the paired GRFs that are captured from ground force platforms or instrumented treadmill are available in the biomechanics community, which were released for gait analysis (e.g. [Fukuchi et al. 2018; Ionescu et al. 2019; Kulbacki et al. 2014; Moore et al. 2015; Shahabpoor and Pavic 2017]), balance (e.g. [dos Santos and Duarte 2016; dos Santos et al. 2017; Wang and van den Bogert 2020; Wang and van den Bogert 2020]), or running (e.g. [Fukuchi et al. 2017; Hamner and Delp 2013; Matijevich et al. 2019]). These datasets often focus on a single specific motion type for patients' recovery and healthcare purposes. Further, there are existing datasets and efforts in incorporating contact labeling tasks such as Shimada et al. [2022], where it provides scene geometry aware motion with estimated contacts guidance. We believe that including such contact labeling information in the current setup will eliminate potential grounding error introduced in the present method. Our collected dataset is unique in a way that it not only incorporates high-accuracy force data for GRF and CoP, but it also covers a variety of motion types, including subtle and balancing movements.

2.1.2 Contact Dynamics Applications. The measured data has proven to be highly valuable for foot-ground contact detection, such as the human foot keypoint dataset from Cao et al. [2018] for discrete contact labeling. Such foot contacts and contact detection have been extensively considered in character motion (e.g. [Bindiganavale and Badler 1998; Lee et al. 2002]) in order to avoid and eliminate foot artifacts to produce contact-rich character control for more realistic virtual human motion and interactions ([Holden et al. 2017, 2016; Starke et al. 2019; Zhang et al. 2018]). More than that, physics-based simulation also often incorporates contact information with contact forces modeled in the physics engines ([Adbolhosseini et al. 2019; Kwon et al. 2020; Won and Lee 2019]). Therefore, by providing more data grounded in real-world physics, such as ground reaction force, we aim to advance the realism of physical interactions of human movements with the environment.

2.2 Ground Reaction Force and Center of Pressure Estimation

Measurement of ground reaction forces and moments (GRF&M) and Center of Pressure (CoP) is common in clinical biomechanics applications. Typically, these quantities are measured using force platforms that have a high spatial and temporal resolution. In sports science, the magnitude and rates of GRFs have been used to examine

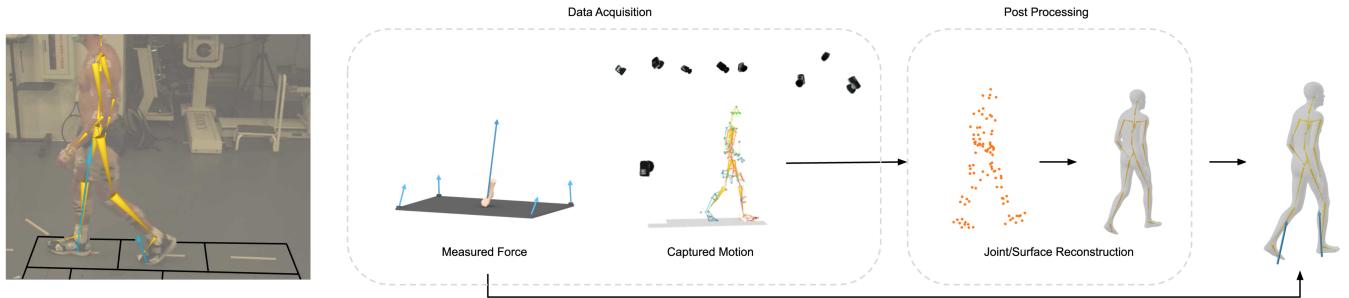


Figure 2: *GroundLink* pipeline. The Data Acquisition phase uses mechanical force platforms with high sampling frequency and an infrared motion capture system (Qualisys). Center of pressure (CoP) is obtained from the ground reaction force and moments (GRF&M). In the Post Processing phase, we annotate the MoCap raw marker data and optimize the joint angles and shape to obtain the input motion sequence parameters. Together, the resulting dataset consists of force data (GRF and CoP) synchronized with the corresponding MoCap sequences.

Table 1: Comparison of *GroundLink* with public MoCap and human pose estimation datasets.

Dataset	# participants	# motions	# frames	# markers	Human body	Force type
CMU Graphics Lab [2000]	144	2435*	3.9M+*	41	Mesh [†]	N/A
ACCAD [2001]	20*	555*	192.6K+*	82	Mesh [†]	N/A
SFU [Yin and Ying 2021]	8*	55*	109.7K+*	53	Mesh [†]	N/A
TotalCapture [Trumble et al. 2017a]	5	60	1.89M	53	Mesh [†]	N/A
Transitions [Mahmood et al. 2019]	1	110	108.7K	53	Mesh	N/A
Taiji Stability (PSU-TMM100) [Scott et al. 2020]	10**	100**	1.36M	N/A	Skeleton	Pressure (vGRF)
UnderPressure [Mourot et al. 2022a]	10	218 [‡]	2.02M [‡]	17 [§]	Skeleton	Pressure (vGRF)
GroundLink (ours)	7	368	1.59M	96	Mesh	GRF, CoP

* Online resource (accessed: May 2023).

[†] Body surface mesh provided via processing by AMASS [Mahmood et al. 2019].

** Number of participants and motion sequence obtained from Scott et al. [2020] and Collins [2023].

[‡] Number of motions given by dataset downloaded from public resource. Number of frames calculated based on 5.6 hours of data at 100 fps provided in Mourot et al. [2022a].

[§] Motions tracked via 17 inertial measurement units (IMU).

the risk of injuries in runners and other athletes ([De Bleecker et al. 2020; Van der Worp et al. 2016]). In physical rehabilitation, GRFs often provide important information about disease mechanisms and can serve as therapeutic targets ([Awad et al. 2020; Costello et al. 2021]). It is important to note that the horizontal components of GRF (i.e., anterior-posterior or medial-lateral) were found to be more important than the vertical component (vGRF) in these prior studies in people with stroke or those with knee pain. In older adults, features of CoP displacement can predict the risk of falls and are used as outcomes in fall prevention research ([Quijoux et al. 2020]).

Measurement of GRFs and CoP with high accuracy requires laboratory-based methods with 3D motion and force capture systems. This can be expensive and restrictive [Konrath et al. 2019]. Therefore, estimation of these quantities from other data (e.g., kinematics, inertial data) has been of interest in the biomechanics community with both machine learning ([Johnson et al. 2018; Mundt et al. 2023, 2020; Schwameder et al. 2019; Sharma et al. 2021; Wouda et al. 2018]) and physics-based approaches ([Bobbert et al. 1991; Jeong and Park 2020; Nedergaard et al. 2018; Skals et al. 2017; Verheul et al. 2019]) used in the past. Also, portable solutions (e.g.

wearable sensors [Pogson et al. 2020]) using pressure insoles, typically restricted to vertical GRF, (e.g. [Mourot et al. 2022a; Scott et al. 2020]) or inertial sensors (e.g. [Johnson et al. 2021; Karatsidis et al. 2017; Wouda et al. 2018]) for estimating GRFs outside the lab in people's natural environments have been attempted.

In the field of computer vision, researchers approach this problem to estimate the contact from monocular images and videos. With video-based motion and pose estimation, physics including contact is solved via optimization to create physics-aware motion (e.g. [Li et al. 2019; Shimada et al. 2021, 2020; Zell et al. 2017]). Li et al. [2019] proposed 3D motion and force estimation from videos through optimizing discrepancy between the observed and reprojected 2D joint and object endpoint positions. Brubaker et al. [2009] and Shimada et al. [2020] proposed video-based tracking for solving contact and performing physics-based contact labeling to prevent floor penetration for dynamic motions. *GRFNet* proposed by Shimada et al. [2021] estimates the GRF and compared it with the measured force provided by Shahabpoor and Pavic [2017] for gait movements. For broader body contact beyond foot-ground, Clever et al. [2021] inferred full-body structure and contact pressure using depth images related to body-mattress contact and Huang et al. [2022] utilized RICH which includes vertex-level contact labels on

the body to reconstruct 3D contact-aware human body from 2D human scene interaction.

So far, 3D GRF and CoP measurements are primarily available within the biomechanics community. The closest related datasets, Mourot et al. [2022a] and Scott et al. [2020], contain contact measurements, specifically pressure data which only denotes the vertical component of the ground reaction force. Relatedly, prediction models such as *GRFNet* from Shimada et al. [2021] provide evaluations with ground truth measurement for gait motion. Our dataset, *GroundLink* contains full-body kinematics with 3D GRF and 2D CoP measurements for a variety of human movements with a particular emphasis on subtle motions.

3 DATA ACQUISITION

3.1 Overview

GroundLink is a MoCap dataset where each motion trial has full-body kinematics with paired GRF and CoP with annotation to distinguish between left and right feet. The force data were captured at high temporal resolution using force platforms that are embedded in the floor and have high measurement accuracy. We aim to use these data in computer graphics applications to create highly accurate animation and depict the movement and interaction of humans with the environment.

We provide two different representations for human motion in our dataset as described in Sec. 3.3. The motion data is accompanied by the measured force data for the GRF and the CoP. The full pipeline of data acquisition and post-processing is illustrated in Fig. 2. Compared to other available datasets for MoCap and contact (Table. 1), such additional annotation for forces provides more comprehensive contact information for the human body, specifically for the foot segments.

3.2 Data Capture

This section describes the acquisition protocol for motion capture, GRF, and CoP data in a biomechanics lab as shown in Figs. 1, 2.

3.2.1 Participants and Motions. We invited 7 participants (4 female and 3 male, age range 18–32) and their body dimensions are shown in Table 2. Participants performed 19 movements shown in Table 3 with a focus on subtle motions. All stationary poses were held for approximately 20 sec, and most of the motions were performed with the participant making contact with the force plates. For the locomotion activities (e.g. walking), it was ensured that the acceleration and deceleration when starting and ending the movement occurred outside of the force plates. Data from the three steps that included contact with the force plates were retained for these motions. To ensure accurate measurement of GRF and CoP for both feet, each foot needs to make contact with a separate force plate during capture. Across all trials, data where only one body segment was in contact with the force plate at any instance were retained for further analyses. Annotation for left and right foot was performed during post processing given the correspondence of the force plate ID and grounded foot. Participants performed four trials of each motion and data from three trials with highest quality were processed and included in the dataset.

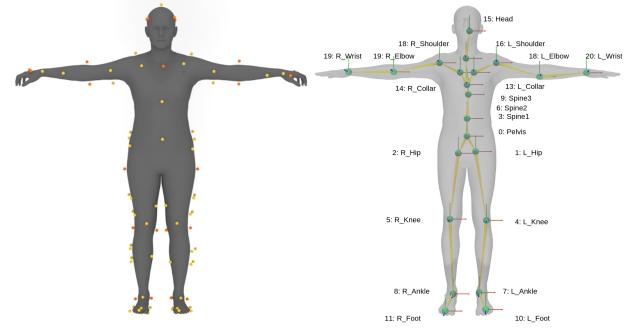


Figure 3: Marker and skeleton setup. (Left) We use a custom placement with 96 markers in total. Our setup focuses on foot measurements with 10 markers on each foot. Markers used to solve pose and shape components in MoSh++ are colored orange (26 total). (Right) Kinematics tree extracted from MoSh++ in the order of input pose parameters [Pavlakos et al. 2019] used in *GroundLinkNet*.

3.2.2 Motion Capture. To measure both the kinematics of the joints and the shape of the human body, 96 retroreflective markers (14 mm diameter) were placed on each participant. The full list of markers is illustrated in Fig. 3 and we have provided detailed marker locations in the supplement. Marker data were captured with an infrared optical motion capture system (Qualisys, Goteborg, Sweden) [Qualisys 2021] with 20 cameras at a sampling frequency of 250 Hz. A single RGB camera was used for video reference.

3.2.3 Ground Reaction Force and Moments. Force plates, as shown in Fig. 4, are specialized instruments embedded in the floor to measure the ground reaction force and center of pressure during movements. We used a total of five force plates with a sampling frequency of 2000 Hz. The GRF data were low-pass filtered at 20 Hz during post-processing. The GRF denoted as F refers to the reaction to the force that the body exerts on the ground upon contact, which has three orthogonal components that are typically denoted as F_x , F_y and F_z along the axes of the global coordinate system. The correspondence of these components with the anteroposterior (AP), mediolateral (ML), and vertical axes of the body can vary by convention used in the lab and orientation of the body relative to the global coordinate system. The sum of the measured forces is equivalent to the GRF, $F = \sum_{i=1}^4 F_i$. Further, ground reaction moments or torques, denoted as M , which represent the torques exerted by the GRF about the anterior-posterior, medial-lateral, and vertical axes, are also captured with the force plates.

3.2.4 Center of Pressure. The center of pressure is the mathematical representation of the point of application of the ground reaction force, acting on the force plate. CoPs are calculated with the ground reaction forces and moments (GRF & M) obtained from the force plates. By the definition, let $\text{CoP} = (\text{CoP}_x, \text{CoP}_y, 0)$, which is a point always on the ground plane. Recall F and M are the ground reaction force and moment, respectively. The CoP coordinates can be calculated as:

Table 2: Description of participants. We introduce diverse body shapes and sizes in *GroundLink*.

Participant ID	s001	s002	s003	s004	s005	s006	s007
Gender [†]	F	M	F	M	M	F	F
Height (m)	1.68	1.70	1.59	1.82	1.95	1.53	1.75
Mass (kg)	69.86	66.68	53.07	71.67	90.7	48.99	63.96

[†] F: Female; M: Male.

$$\text{CoP}_x = -\frac{M_y}{F_z} \quad \text{CoP}_y = \frac{M_x}{F_z} \quad (1)$$

3.3 Data Post Processing

We reconstruct the human body through different methods and representations. First, we label the raw data for all 96 markers, then we use both inverse kinematic solver embedded in Qualisys Track Manager (QTM) [Qualisys 2021] and optimization proposed from Mosh++ [Mahmood et al. 2019] to solve kinematics and additional shape representations for surface reconstruction. During the process, different subsets of the markers are used, and both representations of the data will be released.

3.3.1 Labeling and Skeleton Construction. We annotated the raw MoCap data in QTM to estimate the marker trajectories. Fully labeled data in c3d format with all 96 markers will be provided in the dataset. We also reconstruct the skeleton from 41 markers at joint positions for all the participants, with joint positions $p \in \mathbb{R}^{23 \times 3}$ for all 23 joints in the Euclidean space. This skeleton tracking step is processed with QTM with a skeleton solver and skeleton template generated from the markers mapping. This hierarchical structure can be easily compatible and supported by a variety of 3D applications. Note that this representation was only used for comparative analysis with previous work [Mourot et al. 2022a] after skeleton retargeting, it will be released but was not included in the training and evaluation for GRF and CoP prediction because of the reasons listed in Sec 4.1.

3.3.2 Joint and Surface Construction. From the labeled MoCap data, we adopted MoSh++ [Mahmood et al. 2019] to fit SMPL-X model [Pavlakos et al. 2019] and estimate both pose parameters, θ and shape parameters β to reconstruct the human body. From the 96 markers, a subset of 26 markers is selected and used to reconstruct the human body surface shape and estimate the joint parameters, which are carefully selected given by the optimized markers introduced in [Mahmood et al. 2019] for the best reconstruction results. An initial guess is given for the markers and vertices' distances to the body surface, and we use MoSh++ [Mahmood et al. 2019] to refine the correspondence. Then, we optimize the parameters with a template from SMPL-X [Pavlakos et al. 2019] to minimize the distance between the markers and the corresponding vertices on the human body mesh. Details of the subset of the markers are shown in Fig. 3.

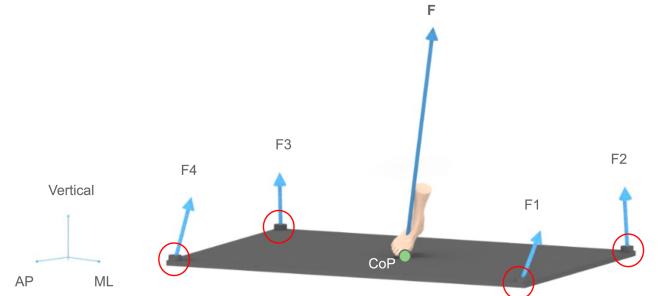


Figure 4: Force plate illustration. Each force plate has four tri-axial force sensors (red circled) embedded in the plate that measures the force between the foot and the ground in the 3 axes: x (anteroposterior, AP), y (mediolateral, ML), and z (vertical) given the illustrated body position. The vectors F_i show the reaction forces measured by the force sensors. The sum of the measured forces is equivalent to GRF, F . CoP is the point of application of F .

Table 3: Description of motions. We focus on subtle movements and weight-shifting effects of motions. During the capture, each stationary pose is maintained for at least 20 sec to capture the balancing of the participant.

Motion	Type	Duration [†] (sec)
Tree (arms down)	yoga	24.70
Tree (arms up)	yoga	24.40
Chair	yoga	23.98
Warrior I	yoga	24.34
Warrior II	yoga	25.03
Dog	yoga	19.85
Side Stretch	yoga	25.17
Lambda dance	dynamic: dancing	21.48
High leg	dynamic: ballet	13.87
Small jump	dynamic: ballet	13.43
Squat	dynamic: exercise	12.46
Jumping Jack	dynamic: exercise	13.95
Stationary hopping	dynamic: exercise	12.14
Taichi	dynamic: exercise	14.81
Swing	sports: tennis	6.68
Serving	sports: tennis	7.66
Kicking	sports: soccer	4.81
Walking	locomotion	4.18
Casual stand	idling	16.77

[†] Avg across all participants

4 EXPERIMENTS

4.1 Benchmark Neural Network Model

We prepared *GroundLinkNet* (Fig. 5), a deep neural network with a straightforward architecture trained with *GroundLink*, to validate the effectiveness of having a dataset with kinematics paired by physics quantities.

Data representation. *GroundLinkNet* primarily takes relative joint angles $\theta \in \mathbb{R}^{k \times 3}$ for k joints and pelvis trajectories $x_{\text{character}} \in$

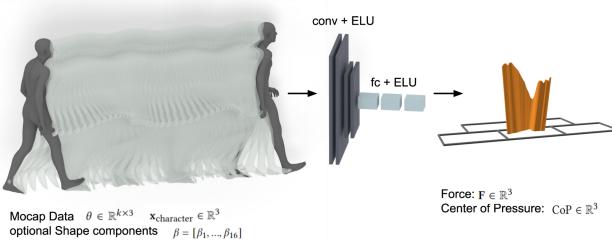


Figure 5: Overview of *GroundLinkNet*, our benchmark neural network model trained with *GroundLink* to predict the GRFs and CoPs given the SMPL [Pavlakos et al. 2019] pose and shape parameters.

\mathbb{R}^3 as inputs, with optional body shape components $\beta = [\beta_1, \dots, \beta_{16}]$, and will output the predicted GRF $F \in \mathbb{R}^3$ and $\text{CoP}_{\text{character}} \in \mathbb{R}^3$.

Using global coordinates as described in [Mourot et al. 2022b] is not ideal due to its lack of invariance of translation and rotation with respect to the ground plane. Stochastic data augmentation as proposed in [Mourot et al. 2022b] cannot generalize to the arbitrary placement of characters in the scene while increasing the training time. Hence, we represent the data with relative positions and joint orientation to avoid confusion caused by using global representation in the coordinate system. Further, due to the double-cover of raw quaternions [Pavlo et al. 2020], we follow a similar setup as [Fussell et al. 2021] to introduce our data representation.

First, we adjust the joint rotation representation. Instead of using global position, $\mathbf{x}_{k,i}$, the position for k th joint at frame i , we use the local rotation represented by rotation vectors of each joint with respect to the parent joint in the kinematics tree, as shown in Fig. 3, $\theta \in \mathbb{R}^{k \times 3}$. Then, we apply transformations to both the center of pressure and the pelvis position to avoid using any global representation in the network and transform to the character space. Given the definition of the center of pressure, since it is always located on the ground surface with $\text{CoP}_{\text{global}} = (\text{CoP}_x, \text{CoP}_y, 0)$ for both left and right foot, we use the projected pelvis as the origin for the character space. Let T be the transformation matrix of the pelvis projected on the ground plane, which is xy -plane with z -positive system, and consists of the rotation and translation of the pelvis. Then, the relative positions of the pelvis and the center of pressure are given as:

$$\mathbf{x}_{\text{character}} = T^{-1} \cdot \mathbf{x}_{\text{pelvis}} = (0, 0, \mathbf{x}_z) \quad (2)$$

$$\text{CoP}_{\text{character}} = T^{-1} \cdot \text{CoP}_{\text{global}} \quad (3)$$

Lastly, shape parameters with 16 components solved by MoSh++ [Mahmood et al. 2019] denoted as $\beta = [\beta_1, \dots, \beta_{16}]$ defined by PCA coefficients in shape space are additional input features. In the end, the input features are the concatenation of the pose θ , $\mathbf{x}_{\text{character}}$, and the optional shape β . The output features are the concatenation of the GRF F and $\text{CoP}_{\text{character}}$.

Network Architecture. Mourot et al. [2022b] proposed and structured a neural network to accommodate motion sequences with variable lengths. The architecture of the network comprises four

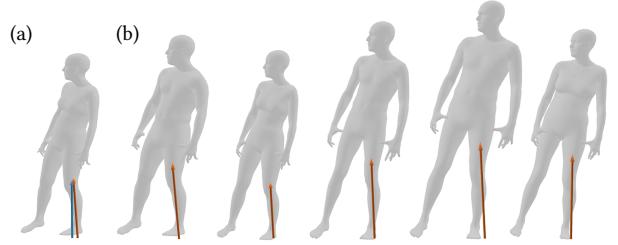


Figure 6: Shape-aware force prediction. Given the same pose from (a) and various body shapes in (b), our *GroundLinkNet* can predict plausible GRFs respecting different body sizes.

1D temporal convolutional layers, each layer with kernels spanning 7 frames. These are followed by three fully connected layers, each independently applied to every frame to maintain adaptability with variable-length sequences. After each convolutional or fully connected layer, exponential linear units (ELU) serve as the activation function, with the exception of the final layer where a softplus activation function is deployed to ensure that the relative pressure values of the sensors from different regions, namely the vertical ground reaction force components, are outputted nonnegative. In our experiments, on the other hand, it is important to note that the components that are predicted - namely the ground reaction force (excluding the vertical GRF) and the center of pressure – need not be restricted to positive values. Both the center of pressure and ground reaction force can possess negative values, reflecting the directional attributes of these vectors.

Training. Since both direction and magnitude are significant to reconstruct GRF and CoP, we use mean squared error (MSE) to measure the difference between the predicted and target values. The loss function is then defined as:

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N \| \mathbf{F}_i - \hat{\mathbf{F}}_i \|^2 \quad (4)$$

where $\hat{\mathbf{F}}$ is the predicted GRF by the neural network.

For evaluation, we use s007's data as a testing set. When performing cross-validation, a different participant serves as a testing subject. With the remaining served as training data, we split it with a ratio of 7 : 3 for the training and validation set.

4.2 Evaluations

Testing with GroundLink. At the stage of inference, our model predicts the GRF and CoP components F_x, F_y, F_z and $\text{CoP}_x, \text{CoP}_y, \text{CoP}_z$, respectively, from the inputs of pelvis position, joint rotations, and shape components. To evaluate the robustness of the models, we perform cross-validation for all the participants (error report in Table 5) and provide a report in Fig. 9, which contains a comparison of ground truth and predicted GRF and CoP for three example motions: taichi, side stretching, and hopping movements. These examples not only demonstrate weight shifting between feet but also shows our model's performance on more dynamic movement such as jumping

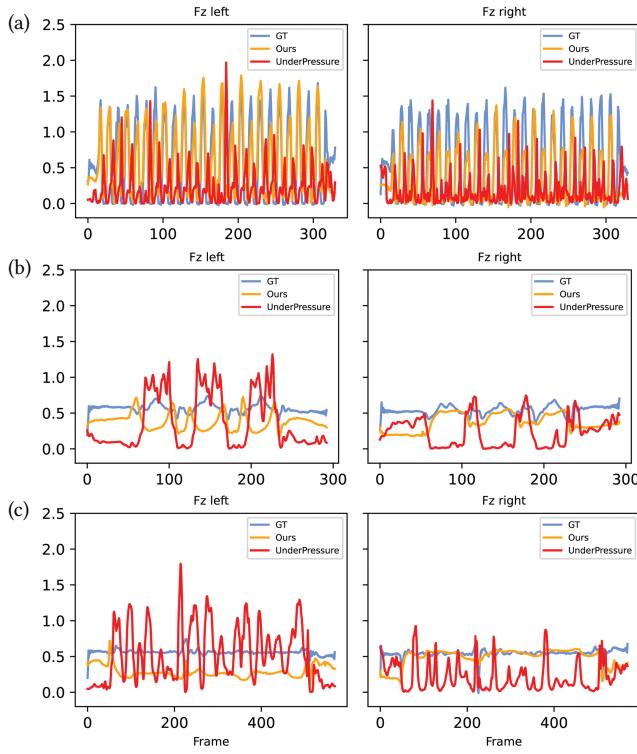


Figure 7: Comparison against UnderPressure [Mourot et al. 2022a] for motions: (a) jumping jacks (b) squatting (c) chair. Horizontal axis: frame # (downsampled by 1/10); vertical axis: force normalized by weight (test subject weighs 90.7kg). Our model yields improved predictions for the vertical component (vGRF), even though our model handles a more challenging task of predicting the full 3D GRF and CoP.

and stationary hopping. Further, we perform a comparative analysis with vGRF prediction provided by UnderPressure [Mourot et al. 2022a] and present the results in Fig. 7 and Table 4.

Testing with kinematics-only data. Many existing datasets have optimized pose and shape components through [Mahmood et al. 2019], such as OSU ACCAD [ACCAD 2001]. While they do not contain any force data associated, we can still evaluate the performance of the model visually. Among all the testing data that are available, it shows that our model predicts convincing GRFs for stationary poses such as martial arts punching and casual standing. We note that while our CoP predictions generally follow the foot positions, they can incorrectly offset from the feet, as observed by the arrows coming off from the feet in visualizations as shown in Fig. 8(b). We believe this is due to the CoP representation relative to the pelvis. We show more results from different datasets in our supplementary video.

Testing with various body shapes. Body shape and body weight are factors that can affect the ground reaction force. Among all the components, vertical ground reaction force (vGRF) is the most sensitive to the body shape. With the shape parameters incorporated into the training setup, we evaluate the correspondence of the

Table 4: vGRF Prediction error comparison between GroundLinkNet and UnderPressure. Values are mean square error (MSE) with the target measured values normalized by weight (test subject weighs 90.7kg).

Motion	UnderPressure [†]	Ours [†]
Chair	0.94/0.19	0.08/0.02
Tree (arm up)	0.54/0.11	0.49/0.03
Tree (arm down)	0.78/0.10	0.14/0.03
Warrior I	0.15/0.16	0.02/0.06
Warrior II	0.09/0.09	0.02/0.04
Dog	0.69/0.13	0.04/0.07
Side Stretch	0.31/0.46	0.06/0.04
Lambada dance	0.27/0.40	0.06/0.13
Squat	0.16/0.14	0.06/0.04
Jumping Jack	0.60/0.50	0.08/0.08
Stationary Hopping	0.80/0.27	0.23/0.03
Taichi	0.19/0.18	0.06/0.09
Soccerkick	0.49/0.41	0.45/0.18
Total	0.44/0.23	0.18/0.07

[†] The error values are calculated errors for the left and right foot, normalized by body weight.

Table 5: Mean square error (MSE) with target measured values

Motion	F_x [†]	F_y [†]	F_z [†]	CoP_x [‡]	CoP_y [‡]
Chair	0.12/0.13	1.85/1.59	25.62/20.54	1.46/6.56	2.61/3.54
Tree (arm up)	0.05/0.03	1.34/0.45	149.8/14.55	2.09/8.72	1.56/2.37
Tree (arm down)	0.05/0.02	1.30/0.41	130.8/17.08	3.88/14.68	0.70/3.28
Warrior I	0.13/0.14	1.96/1.78	22.80/66.49	1.19/31.90	4.63/14.83
Warrior II	0.19/0.24	0.96/1.37	15.25/48.12	1.23/20.46	1.98/9.07
Dog	0.65/0.76	1.80/2.54	55.34/86.38	12.1/10.82	13.46/48.52
Side Stretch	0.22/0.26	3.22/4.09	43.36/51.04	9.26/49.63	4.26/20.39
Lambada Dance	0.50/0.59	3.30/4.08	70.37/93.94	6.79/11.90	3.52/7.47
Squat	0.14/0.15	1.27/1.30	24.36/28.07	1.71/4.27	1.71/2.49
Jumping Jack	2.82/3.20	5.89/4.98	61.68/74.59	3.21/3.54	2.45/3.79
Hopping*	2.36/0.09	1.85/0.49	221.5/29.34	5.0/14.74	2.56/11.88
Taichi	0.12/0.15	1.98/2.06	41.41/54.54	7.95/9.00	5.70/7.89
Tennis: Swing	0.28/0.22	0.88/1.83	19.24/30.28	4.20/2.58	6.32/7.26
Tennis: Serve	0.23/0.29	1.64/1.17	42.10/42.17	2.27/5.76	2.72/24.09
Soccerkick	5.05/3.79	2.88/2.02	178.9/118.5	14.02/25.56	11.21/39.41
Casual Stand	0.07/0.07	0.55/0.92	46.69/52.71	3.36/11.19	2.03/8.37

* refers to Stationary Hopping.

[†] The calculated error for left/right foot across all test subjects ($\times 10^{-3}$), given by force normalized by weight.

[‡] Error for left/right foot across all test subject ($\times 10^{-3}$) in m

human body shape with both the ground reaction force and the center of pressure. The analysis of our predicted GRF for people with varying body shapes, as shown in Fig. 6, reveals that individuals with larger physiques tend to exhibit higher vGRF (F_z) as expected.

5 CONCLUSION

In this paper, we introduce *GroundLink*, a new motion capture data with corresponding GRFs and CoPs, which are difficult to obtain due to the restricted resources of the lab and spaces. From the motion and contact force data, we present *GroundLinkNet*, a benchmark model trained from *GroundLink* to predict GRFs and

CoPs from the kinematics and shapes of the human body. With the possibility of estimating the contact force from pure kinematics of human characters, we hope our data and model can serve numerous downstream applications of digital humans with physical realism.

Discussion and Future Work. As a benchmark model, we noticed some prediction inconsistencies delivered by *GroundLinkNet*. First, the model is currently set to use the projected pelvis as the character space origin, which results in a noticeable offset in predicting CoP, as illustrated in Fig. 8(b). To remedy this, a more appropriate frame-of-reference, such as projected feet for referring left and right GRF and CoP, can enhance prediction accuracy. Second, *GroundLink* only contains walking trials with a maximum of three-step strides to keep the participants on the limited region of the force plates. Because of this, we observe that our benchmark model struggles to predict GRFs for different types of locomotions. Third, we also noticed the model being too sensitive to small changes in the pose for upper body, e.g. hand rotations, affecting the output. These are signs that there is not enough data (such as isolated upper body movements) to deal with such out-of-distribution test scenarios. While we showed that *GroundLinkNet* can generalize to unseen scenarios to some extent, it is also clear that the data scarcity issue must be addressed. We believe there are open research opportunities in (a) deriving a better machine learning scheme, including data representation, to understand physics from a limited amount of data, and (b) simulating physics to obtain more data, while also leveraging the captured physics quantities simultaneously.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their valuable comments. The authors also thank the participants for their time and commitment to the study. This project was partially supported by the Alfred P. Sloan Foundation: Sloan Research Fellowship.

REFERENCES

- OSU ACCAD. 2001. ACCAD MoCap System and Data. <https://accad.osu.edu/research/motion-lab/mocap-system-and-data> Accessed: 2023-05-01.
- Farzad Adbolhosseini, Hung Yu Ling, Zhaoming Xie, Xue Bin Peng, and Michiel van de Panne. 2019. On Learning Symmetric Locomotion. In *Proc. ACM SIGGRAPH Motion, Interaction, and Games (MIG 2019)*.
- Louis N. Awad, Michael D. Lewek, Trisha M. Kesar, Jason R. Franz, and Mark G. Bowden. 2020. These legs were made for propulsion: advancing the diagnosis and treatment of post-stroke propulsion deficits. *Journal of NeuroEngineering and Rehabilitation* 17, 1 (2020), 139. <https://doi.org/10.1186/s12984-020-00747-6>
- Rama Bindiganavale and Norman I. Badler. 1998. Motion Abstraction and Mapping with Spatial Constraints. In *Modelling and Motion Capture Techniques for Virtual Environments*, Nadia Magnenat-Thalmann and Daniel Thalmann (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 70–82.
- Maarten F. Bobbeld, Henk C. Schamhardt, and Benno M. Nigg. 1991. Calculation of vertical ground reaction force estimates during running from positional data. *Journal of Biomechanics* 24, 12 (1991), 1095–1105. [https://doi.org/10.1016/0021-9290\(91\)90002-5](https://doi.org/10.1016/0021-9290(91)90002-5)
- Marcus A. Brubaker, Leonid Sigal, and David J. Fleet. 2009. Estimating contact dynamics. In *2009 IEEE 12th International Conference on Computer Vision*. 2389–2396. <https://doi.org/10.1109/ICCV.2009.5459407>
- Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2018. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. In *arXiv preprint arXiv:1812.08008*.
- Henry M. Clever, Patrick Grady, Greg Turk, and Charles C. Kemp. 2021. Body Pressure – Inferring Body Pose and Contact Pressure from a Depth Image. [arXiv:2105.09936 \[cs.CV\]](https://doi.org/10.4236/ojs.20211112100936)
- CMU CMU Graphics Lab. 2000. CMU Graphics Lab Motion Capture Database. <http://mocap.cs.cmu.edu/>
- Robert T. Collins. 2023. A Light-Weight Contrastive Approach for Aligning Human Pose Sequences. *arXiv:2303.04244 [cs.CV]*
- K.E. Costello, D.T. Felson, T. Neogi, N.A. Segal, C.E. Lewis, K.D. Gross, M.C. Nevitt, C.L. Lewis, and D. Kumar. 2021. Ground reaction force patterns in knees with and without radiographic osteoarthritis and pain: descriptive analyses of a large cohort (the Multicenter Osteoarthritis Study). *Osteoarthritis and Cartilage* 29, 8 (2021), 1138–1146. <https://doi.org/10.1016/j.joca.2021.03.009>
- C De Bleeker, S Vermeulen, C De Blaiser, T Willems, R De Ridder, and P Roosen. 2020. Relationship Between Jump-Landing Kinematics and Lower Extremity Overuse Injuries in Physically Active Populations: A Systematic Review and Meta-Analysis. *Sports Medicine* 50, 8 (2020), 1515–1532. <https://doi.org/10.1007/s40279-020-01296-7> PMID: 32514700.
- Damiana A. dos Santos and Marcos Duarte. 2016. A Public Data set of Human Balance Evaluations. *PeerJ Life and Environment* (2016). <https://doi.org/10.7717/peerj.2648>
- Damiana A. dos Santos, Claudiene A. Fukuchi, Reginaldo K. Fukuchi, and Marcos Duarte. 2017. A Data set with Kinematic and Ground Reaction Forces of Human Balance. *PeerJ Life and Environment* (2017). <https://doi.org/10.7717/peerj.3626>
- Claudiene A. Fukuchi, Reginaldo K. Fukuchi, and Marcos Duarte. 2018. A public dataset of overground and treadmill walking kinematics and kinetics in healthy individuals. *PeerJ Life and Environment* (2018). <https://doi.org/10.7717/peerj.4640>
- Reginaldo K. Fukuchi, Claudiene A. Fukuchi, and Marcos Duarte. 2017. A public dataset of running biomechanics and the effects of running speed on lower extremity kinematics and kinetics. *PeerJ Life and Environment* (2017). <https://doi.org/10.7717/peerj.3298>
- Levi Fussell, Kevin Bergamin, and Daniel Holden. 2021. SuperTrack: Motion Tracking for Physically Simulated Characters Using Supervised Learning. *ACM Trans. Graph.* 40, 6, Article 197 (dec 2021), 13 pages. <https://doi.org/10.1145/3478513.3480527>
- Patrick Grady, Chengcheng Tang, Samarth Brahmbhatt, Christopher D. Twigg, Chengde Wan, James Hays, and Charles C. Kemp. 2022. PressureVision: Estimating Hand Pressure from a Single RGB Image. *arXiv:2203.10385 [cs.CV]*
- Samuel R. Hamner and Scott L. Delp. 2013. Muscle contributions to fore-aft and vertical body mass center accelerations over a range of running speeds. *Journal of Biomechanics* 46, 4 (2013), 780–787. <https://doi.org/10.1016/j.jbiomech.2012.11.024>
- Daniel Holden, Taku Komura, and Jun Saito. 2017. Phase-Functioned Neural Networks for Character Control. *ACM Trans. Graph.* 36, 4, Article 42 (jul 2017), 13 pages. <https://doi.org/10.1145/3072959.3073663>
- Daniel Holden, Jun Saito, and Taku Komura. 2016. A Deep Learning Framework for Character Motion Synthesis and Editing. *ACM Trans. Graph.* 35, 4, Article 138 (jul 2016), 11 pages. <https://doi.org/10.1145/2897824.2925975>
- Chun-Hao P. Huang, Hongwei Yi, Markus Höschle, Matvey Safroshkin, Tsvetelina Alexiadis, Senya Polikovsky, Daniel Scharstein, and Michael J. Black. 2022. Capturing and Inferring Dense Full-Body Human-Scene Contact. In *Proceedings IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*. 13274–13285.
- Catalin Ionescu, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu. 2019. A multimodal dataset of human gait at different walking speeds established on injury-free adult participants. *Scientific Data* 6, 1 (jul 2019), 111. <https://doi.org/10.1038/s41597-019-0124-4>
- Hyunoo Jeong and Sukyung Park. 2020. Estimation of the ground reaction forces from a single video camera based on the spring-like center of mass dynamics of human walking. *Journal of Biomechanics* 113 (12 2020), 110074. <https://doi.org/10.1016/j.jbiomech.2020.110074>
- William R. Johnson, Ajmal Mian, Cyril J. Donnelly, David Lloyd, and Jacqueline Alderson. 2018. Predicting athlete ground reaction forces and moments from motion capture. *Scientific Data* 56, 10 (oct 2018), 1781–1792. <https://doi.org/10.1007/s11517-018-1802-7>
- William R. Johnson, Ajmal Mian, Mark A. Robinson, Jasper Verheul, David G. Lloyd, and Jacqueline A. Alderson. 2021. Multidimensional Ground Reaction Forces and Moments From Wearable Sensor Accelerations via Deep Learning. *IEEE Transactions on Biomedical Engineering* 68, 1 (2021), 289–297. <https://doi.org/10.1109/TBME.2020.3006158>
- Angelos Karatsidis, Giovanni Bellusci, H. Martin Schepers, Mark De Zee, Michael S. Andersen, and Peter H. Veltink. 2017. Estimation of Ground Reaction Forces and Moments During Gait Using Only Inertial Motion Capture. *Sensors* 17, 1 (2017). <https://doi.org/10.3390/s17010075>
- Jason Konrath, Angelos Karatsidis, Martin Schepers, G. Bellusci, Mark de Zee, and Michael Andersen. 2019. Estimation of the Knee Adduction Moment and Joint Contact Force during Daily Living Activities Using Inertial Motion Capture. *Sensors* 19 (04 2019), 1681. <https://doi.org/10.3390/s19071681>
- Marek Kulbacki, Jakub Segeñ, and Jerzy Nowacki. 2014. 4GAIT: Synchronized MoCap, Video, GRF and EMG Datasets: Acquisition, Management and Applications. 555–564. https://doi.org/10.1007/978-3-319-05458-2_57
- Taesoo Kwon, Yoonsang Lee, and Michiel Van De Panne. 2020. Fast and Flexible Multilegged Locomotion Using Learned Centroidal Dynamics. *ACM Trans. Graph.* 39, 4, Article 46 (aug 2020), 24 pages. <https://doi.org/10.1145/3386569.3392432>
- Jehee Lee, Jinxiang Chai, Paul S. A. Reitsma, Jessica K. Hodgins, and Nancy S. Pollard. 2002. Interactive Control of Avatars Animated with Human Motion Data. *ACM Trans. Graph.* 21, 3 (jul 2002), 491–500. <https://doi.org/10.1145/566654.566607>

- Zongmian Li, Jiri Sedlar, Justin Carpentier, Ivan Laptev, Nicolas Mansard, and Jose Sivic. 2019. Estimating 3D Motion and Forces of Person-Object Interactions from Monocular Video. In *Computer Vision and Pattern Recognition (CVPR)*.
- Yiyue Luo, Yunzhu Li, Foshey Michael, Shou Wan, Pratyusha Sharma, Tomas Palacios, Antonio Torralba, and Wojciech Matusik. 2021. Intelligent Carpet: Inferring 3D Human Pose from Tactile Signals. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 11250–11260.
- Naureen Mahmood, Nima Ghorbani, Nikolaus F. Troje, Gerard Pons-Moll, and Michael J. Black. 2019. AMASS: Archive of Motion Capture as Surface Shapes. In *International Conference on Computer Vision*. 5442–5451.
- Emily S. Matijevich, Lauren M. Branscombe, Leon R. Scott, and Karl E. Zelik. 2019. Ground reaction force metrics are not strongly correlated with tibial bone load when running across speeds and slopes: Implications for science, sport and wearable tech. *PLoS ONE* (2019). <https://doi.org/10.1371/journal.pone.0210000>
- Jason K. Moore, Sandra K. Hnat, and Antonie J. van den Bogert. 2015. An elaborate data set on human gait and the effect of mechanical perturbations. *PeerJ Life and Environment* (2015). <https://doi.org/10.7717/peerj.918>
- Lucas Mourot, Ludovic Hoyet, François Le Clerc, and Pierre Hellier. 2022a. UnderPressure: Deep Learning for Foot Contact Detection, Ground Reaction Force Estimation and Footskate Cleanup. arXiv:2208.04598 [cs.GR]
- Lucas Mourot, Ludovic Hoyet, François Le Clerc, and Pierre Hellier. 2022b. UnderPressure: Deep Learning of Foot Contacts and Forces for Character Animation. (2022). <https://doi.org/10.48550/arXiv.2208.04598> arXiv:2208.04598 To be published in proceedings of Symposium on Computer Animation (SCA).
- Marion Mundt, Zachery Born, Molly Goldacre, and Jacqueline Alderson. 2023. Estimating Ground Reaction Forces from Two-Dimensional Pose Data: A Biomechanics-Based Comparison of AlphaPose, BlazePose, and OpenPose. *Sensors* 23, 1 (2023). <https://doi.org/10.3390/s23010078>
- Marion Mundt, Arnd Koeppe, Sina David, Franz Bamer, Wolfgang Potthast, and Bernd Markert. 2020. Prediction of ground reaction force and joint moments based on optical motion capture data during gait. *Medical Engineering & Physics* 86 (2020), 29–34. <https://doi.org/10.1016/j.medengphy.2020.10.001>
- Niels J. Nedergaard, Jasper Verheul, Barry Drust, Terence Etchells, Paulo Lisboa, Mark A. Robinson, and Jos Vanrenterghem. 2018. The feasibility of predicting ground reaction forces during running from a trunk accelerometry driven mass-spring-damper model. *PeerJ* 6 (dec 2018), e6105. <https://doi.org/10.7717/peerj.6105>
- Georgios Pavlakos, Vasileios Choutas, Nima Ghorbani, Timo Bolkart, Ahmed A. A. Osman, Dimitrios Tzionas, and Michael J. Black. 2019. Expressive Body Capture: 3D Hands, Face, and Body from a Single Image. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. 10975–10985.
- Dario Pavillo, Christoph Feichtenhofer, Michael Auli, and David Grangier. 2020. Modeling Human Motion with Quaternion-Based Neural Networks. *Int. J. Comput. Vision* 128, 4 (apr 2020), 855–872. <https://doi.org/10.1007/s11263-019-01245-6>
- Mark Pogson, Jasper Verheul, Mark A. Robinson, Jos Vanrenterghem, and Paulo Lisboa. 2020. A neural network method to predict task- and step-specific ground reaction force magnitudes from trunk accelerations during running activities. *Medical Engineering & Physics* 78 (2020), 82–89. <https://doi.org/10.1016/j.medengphy.2020.02.002>
- Qualisys. 2021. Qualisys Motion Capture System. <https://www.qualisys.com/> Accessed: 2022-05-30.
- Flavien Quijoux, Aliénor Vienne-Jumeau, François Bertin-Hugault, Philippe Zawieja, Marie Lefèvre, Pierre-Paul Vidal, and Damien Ricard. 2020. Center of pressure displacement characteristics differentiate fall risk in older people: A systematic review with meta-analysis. *Ageing Research Reviews* 62 (2020), 101117. <https://doi.org/10.1016/j.arr.2020.101117> PMID: 32565327.
- Hermann Schwameder, Josef Kröll, Julia Lindorfer, and Julian Fritz. 2019. Estimation of ground reaction forces from kinematics in ski-jumping imitation jumps.
- Jesse Scott, Bharadwaj Ravichandran, Christopher Funk, Robert T. Collins, and Yanxi Liu. 2020. From Image to Stability: Learning Dynamics from Human Pose. In *Computer Vision – ECCV 2020*, Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm (Eds.). Springer International Publishing, Cham, 536–554. https://doi.org/10.1007/978-3-030-58592-1_32
- Erfan Shahabpoor and Aleksandar Pavic. 2017. Measurement of Walking Ground Reactions in Real-Life Environments: A Systematic Review of Techniques and Technologies. *Sensors (Basel)* 17, 9 (Sep 2017), 2085. <https://doi.org/10.3390/s17092085> PMID: PMC5620730.
- Dharmendra Sharma, Pavel Davidson, Philipp Müller, and Robert Piché. 2021. Indirect Estimation of Vertical Ground Reaction Force from a Body-Mounted INS/GPS Using Machine Learning. *Sensors* 21 (02 2021), 1553. <https://doi.org/10.3390/s21041553>
- Soshi Shimada, Vladislav Golyanik, Zhi Li, Patrick Pérez, Weipeng Xu, and Christian Theobalt. 2022. HULC: 3D HUman Motion Capture with Pose Manifold SampLing and Dense Contact Guidance. In *European Conference on Computer Vision (ECCV)*. 516–533.
- Soshi Shimada, Vladislav Golyanik, Weipeng Xu, Patrick Pérez, and Christian Theobalt. 2021. Neural Monocular 3D Human Motion Capture with Physical Awareness. *ACM Transactions on Graphics* 40, 4, Article 83 (aug 2021).
- Soshi Shimada, Vladislav Golyanik, Weipeng Xu, and Christian Theobalt. 2020. PhysCap: Physically Plausible Monocular 3D Motion Capture in Real Time. *ACM Trans. Graph.* 39, 6, Article 235 (nov 2020), 16 pages. <https://doi.org/10.1145/3414685.3417877>
- Sebastian Skals, Moon Jung, Michael Damsgaard, and Michael Andersen. 2017. Prediction of ground reaction forces and moments during sports-related movements. *Multibody System Dynamics* 39 (03 2017). <https://doi.org/10.1007/s11044-016-9537-4>
- Sebastian Starke, He Zhang, Taku Komura, and Jun Saito. 2019. Neural State Machine for Character-Scene Interactions. *ACM Trans. Graph.* 38, 6, Article 209 (nov 2019), 14 pages. <https://doi.org/10.1145/3355089.3356505>
- Matt Trumble, Andrew Gilbert, Charles Malleson, Adrian Hilton, and John Collomosse. 2017a. Total Capture: 3D Human Pose Estimation Fusing Video and Inertial Sensors. In *2017 British Machine Vision Conference (BMVC)*.
- Matt Trumble, Andrew Gilbert, Charles Malleson, Adrian Hilton, and John Collomosse. 2017b. Total Capture: 3D Human Pose Estimation Fusing Video and Inertial Sensors. In *2017 British Machine Vision Conference (BMVC)*.
- H Van der Worp, JW Vrielink, and SW Bredeweg. 2016. Do runners who suffer injuries have higher vertical ground reaction forces than those who remain injury-free? A systematic review and meta-analysis. *British Journal of Sports Medicine* 50, 8 (2016), 450–457. <https://doi.org/10.1136/bjsports-2015-094924> PMID: 26729857.
- Jasper Verheul, Warren Gregson, Paulo Lisboa, Jos Vanrenterghem, and Mark A. Robinson. 2019. Whole-body biomechanical load in running-based sports: The validity of estimating ground reaction forces from segmental accelerations. *Journal of Science and Medicine in Sport* 22, 6 (2019), 716–722. <https://doi.org/10.1016/j.jssams.2018.12.007>
- Huawei Wang and Antonie J. van den Bogert. 2020. Identification of the human postural control system through stochastic trajectory optimization. *Journal of Neuroscience Methods* 334 (2020), 108580. <https://doi.org/10.1016/j.jneumeth.2020.108580>
- Huawei Wang and Antonie J. van den Bogert. 2020. Standing Balance Experiment with Long Duration Random Pulses Perturbation. <https://api.semanticscholar.org/> CorpusID:244988575
- Jungdam Won and J. Lee. 2019. Learning body shape variation in physics-based characters. *ACM Transactions on Graphics* 38 (11 2019), 1–12. <https://doi.org/10.1145/3355089.3356499>
- Frank Wouda, Matteo Giuberti, G. Bellusci, Erik Maartens, Jasper Reenalda, Bert-Jan BEIJNUM, and Peter Veltink. 2018. Estimation of Vertical Ground Reaction Forces and Sagittal Knee Kinematics During Running Using Three Inertial Sensors. *Frontiers in Physiology* 9 (03 2018). <https://doi.org/10.3389/fphys.2018.00218>
- KangKang Yin and Dinesh K. Pai. 2003. FootSee: an interactive animation system. In *Symposium on Computer Animation*. <https://api.semanticscholar.org/> CorpusID: 14380898
- KangKang Yin and Goh Jing Ying. 2021. SFU Motion Capture Database. <https://mocap.cs.sfu.ca/> Accessed: 2023-05-01.
- Petrissa Zell, Bastian Wandt, and Bodo Rosenhahn. 2017. Joint 3D Human Motion Capture and Physical Analysis from Monocular Videos. 17–26. <https://doi.org/10.1109/CVPRW.2017.9>
- He Zhang, Sebastian Starke, Taku Komura, and Jun Saito. 2018. Mode-Adaptive Neural Networks for Quadruped Motion Control. *ACM Trans. Graph.* 37, 4, Article 145 (jul 2018), 11 pages. <https://doi.org/10.1145/3197517.3201366>

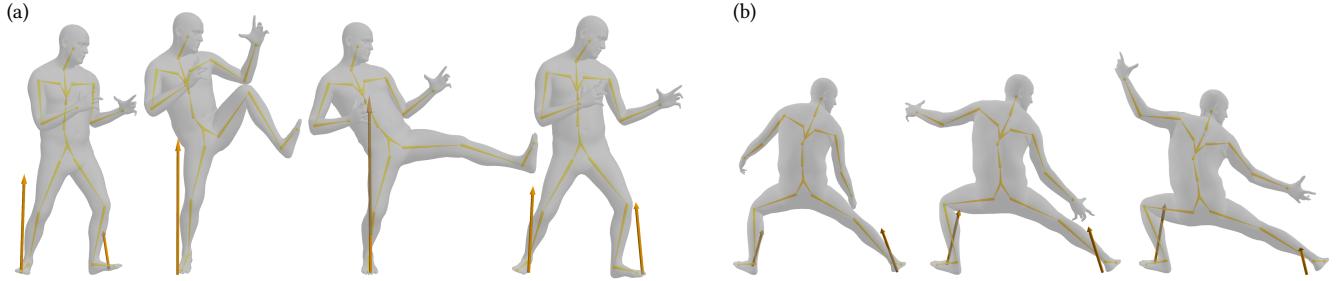


Figure 8: Testing result for OSU ACCAD: (a) push kicking motion and (b) Victory martial arts. Preprocessed MoCap data is obtained from AMASS to retrieve the pose components for input parameters for the model. Even in the absence of ground truth contact forces data for external datasets, we employ our trained model from *GroundLink* to predict GRF and CoP based solely on the available kinematics.

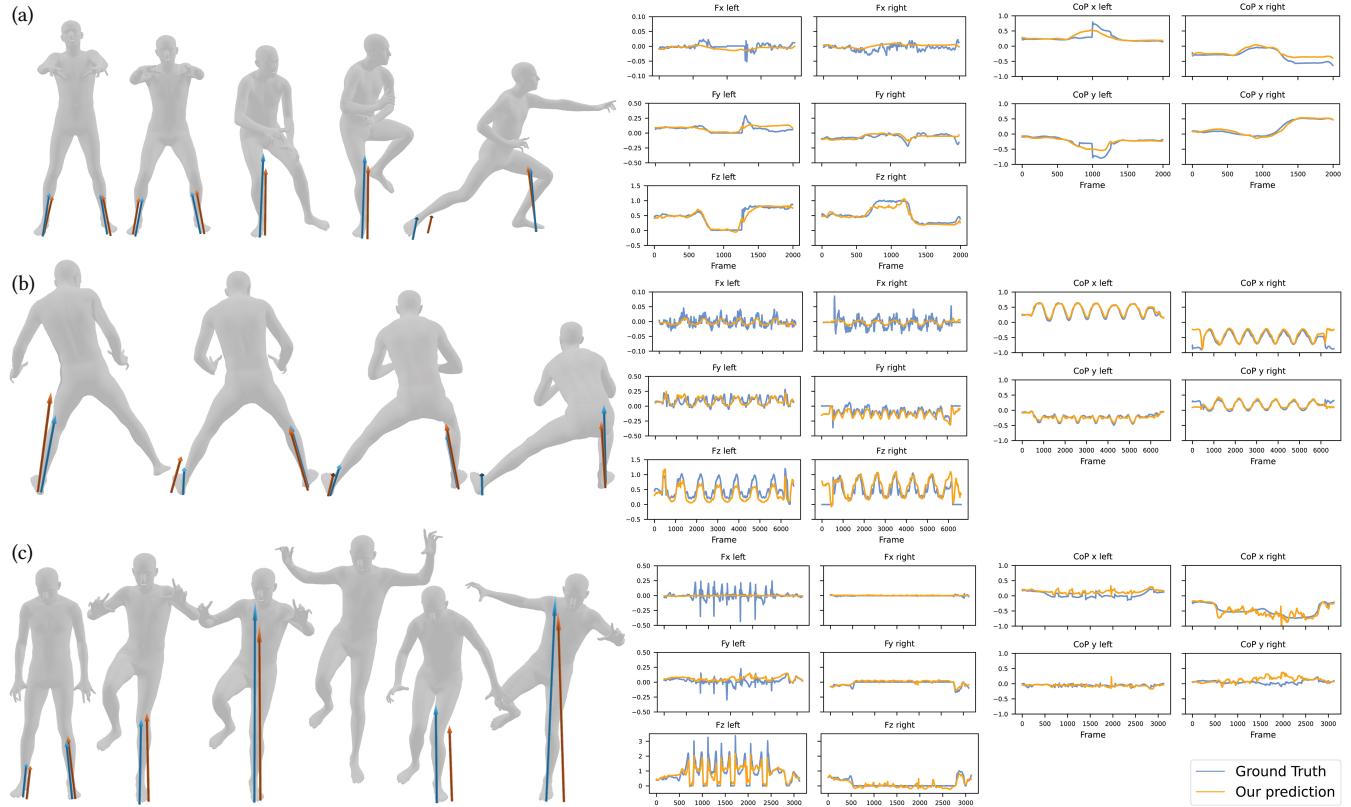


Figure 9: *GroundLinkNet* results for participant s004. The model is trained with the data from the other 6 participants. The plots show a comparison of the predicted value (orange) and the ground truth values (blue) for GRF (normalized by subject weight) and CoP (m). Horizontal axis: video frame #; vertical axis (middle): force normalized by weight (subject weighs 71.67kg); vertical axis (right): relative position with projected pelvis (m). Example motions: (a) taichi, (b) side stretching, and (c) stationary hopping. Overall, the models generate reasonable prediction patterns for GRF for various motions. For CoP prediction, it performs better with smooth motions such as side stretching with consistent weight shifting patterns but is relatively less convincing for motions when a single foot is lifted and shifts dramatically in the air such as taichi movement.